Instituto Superior Técnico
Departamento de Matemática

ÁLGEBRA LINEAR

Sumário

0. Introdução

Este texto regista de forma razoavelmente fiel o conteúdo das aulas de Álgebra Linear dadas nos anos letivos 2021-2024 aos cursos de LMAC e LEFT no Instituto Superior Técnico. Não substitui a consulta dos livros de texto indicados na bibliografia na página da cadeira. Cada secção conclui com uma lista de exercícios propostos cujas soluções são apresentadas no final do texto. Os exercícios estão ordenados de forma a que possam indo ser realizados à medida que uma secção é lida, não sendo necessário ler toda a secção antes de começar a fazer exercícios.

0.1. **O que é a Álgebra Linear? - I.** A Álgebra é a parte da Matemática que estuda a resolução de equações (e estruturas relacionadas). O nome vem do árabe *Al-Jabr* - reunião de pedaços - que foi a expressão usada pelo matemático persa Al-Khwarizmi (no início do século IX) para designar a passagem de um termo numa equação para o outro lado do sinal de igual. Foi o nome deste matemático (e portanto o nome de uma região do atual Uzbequistão!) que deu origem às palavras algarismo e algoritmo.

O termo Al-Jabr aparecia no título de um famoso livro escrito por Al-Khwarizmi - O Compêndio de cálculo por Al-Jabr e Al-Muqabalah - que estabeleceu a Álgebra como disciplina da Matemática e foi um dos principais livros de texto de Matemática nas universidades europeias até ao século XVI!

A Álgebra Linear é (numa primeira aproximação) a parte da Matemática que estuda a resolução de equações lineares. Os sistemas de equações lineares devem já ser-vos familiares. Eis um exemplo:

$$\begin{cases} 2x + 3y - z + w = 4 \\ -x + 2z - w = 1 \end{cases}$$

Estamos interessados em saber se o sistema tem soluções e, em caso afirmativo, em descrever as soluções de uma forma conveniente. Mesmo uma questão tão simples pode ser encarada de vários pontos de vista, todos eles importantes:

(1) Tradicional: Pretende-se determinar as quantidades codificadas pelas variáveis que satisfazem as relações dadas,
(2) Geométrica: Cada uma das equações define um plano num espaço de dimensão 4 e o sistema descreve a sua intersecção,
(3) Funcional: O sistema define a pré-imagem do ponto $(4, 1) \in \mathbb{R}^2$ pela função linear $(x, y, z, w) \mapsto (2x + 3y - z + w, -x + 2z - w)$.

A relevância do primeiro ponto de vista é familiar. Iremos ver muitos exemplos de aplicação ao longo do curso incluindo o fundamento do algoritmo de busca do Google.

O segundo ponto de vista sugere que a Álgebra Linear nos abre uma janela sobre a geometria dos espaços de dimensão arbitrária!

O terceiro ponto de vista alude a uma das razões pelas quais a Álgebra Linear é fundamental dentro da própria Matemática e à razão prática pela qual a vão estudar agora: a ideia fundamental do Cálculo é o estudo das funções através das suas aproximações lineares (ou derivadas). A Álgebra Linear descreve o comportamento destas funções lineares.

Para poder dar uma ideia mais precisa daquilo em que vai consistir o nosso estudo da Álgebra Linear é necessário tratar primeiro uma questão mais geral.

0.2. **O que é a Matemática?**

A Matemática é a poesia das ideias lógicas.

Einstein

Não há nenhuma definição comumente aceite de Matemática, mas como irão ver é algo muito diferente da Matemática com a qual tiveram contacto no ensino secundário. Apesar de não ser fácil dizer o que é exatamente Matemática é possível enumerar algumas das suas características. As mais relevantes para o que queremos tentar comunicar são as seguintes:

0.2.1. *É um mundo alternativo povoado por objetos abstratos, cheio de beleza e mistério.* Há certos conceitos abstratos, como o de números naturais, que são utilizados por vários animais. À medida que a civilização se foi desenvolvendo tornou-se necessário resolver certos problemas práticos como o cálculo de áreas e a Matemática surgiu destas necessidades práticas. Com o decorrer do tempo (notavelmente na Grécia antiga) houve pessoas que se dedicaram a estudar os métodos de resolução em si mesmos (desligados das suas aplicações) e começou assim a Matemática "pura".

Foi também (tanto quanto se sabe) na Grécia, no âmbito da Geometria, que surgiu o conceito de *demonstração*, isto é, de dedução lógica de proposições a partir de proposições anteriormente estabelecidas, tomando como partida um conjunto de termos e suas propriedades tidas como evidentes - os axiomas.

A ideia de demonstração é central na Matemática. Uma demonstração consiste numa *explicação* da validade de uma dada afirmação que não admite qualquer objeção. É a existência das demonstrações que torna a Matemática o conhecimento humano mais seguro a que temos acesso.

A Matemática consiste no estudo de certos objetos abstratos (números, operações entre estes, triângulos, funções,...), na descoberta das suas propriedades, e na justificação destas através de demonstrações. Este próprio processo de estudo ou necessidades das aplicações (notavelmente à Física mas também a outros domínios) conduzem frequentemente à descoberta (ou invenção) de novos objetos de estudo. Ao longo da História foi sendo construído um edifício gigante que hoje em dia cresce exponencialmente.

O crescimento exponencial da Matemática requer uma constante reavaliação, simplificação e síntese dos conhecimentos anteriores. É isso que permite que voçês tenham entendido com 12 anos factos que só estavam ao alcance dos mais destacados intelectuais de há algumas centenas de anos e que vários de vós vão, daqui a 4 ou 5 anos, perceber muito melhor a Relatividade Geral do que Einstein no final da sua vida.

A Matemática tem muitos paralelos com a Biologia. Em vez de estudar esquilos, estudamos triângulos. Tal como na Biologia todos os objetos de estudo estão relacionados, por muito diferentes que pareçam. Tal como a Biologia, a Matemática tem aplicações mas é independente delas.

0.2.2. *É uma atividade humana e uma forma de pensar.* O objetivo da Matemática é o *entendimento* dos objetos matemáticos, que é algo humano, psicológico e progressivo. As definições e as demonstrações são os *mecanismos* através do qual esse entendimento é transmitido. São mecanismos imperfeitos mas são os melhores de que dispomos (não é fácil transmitir a intuição sobre objetos abstratos que se ganha depois de muitos anos a pensar neles).

Exemplo 0.3. *Vários tipos de entendimento da fórmula resolvente das equações do segundo grau:*

(1) *Saber que existe uma fórmula.*
(2) *Saber a fórmula de cor e saber aplicá-la.*
(3) *Saber justificar a validade da fórmula (isto é, saber demonstrá-la).*
(4) *Perceber que a essência da fórmula se reduz à definição de raíz quadrada (resolução das equações* $x^2 = A$*) e à ideia de "completar o quadrado"(que geometricamente é uma translação) para reduzir o caso geral a este caso particular.*

A história não para aqui. À medida que forem aprendendo mais Matemática o vosso entendimento desta fórmula tornar-se-á mais sofisticado.

A Matemática é uma atividade criativa em que temos uma enorme liberdade. A criatividade reside na identificação de novos objetos, na descoberta das suas propriedades e na construção dos argumentos que as justificam [1].

Da mesma forma que a criatividade nas artes não pode ser exercida se não estiver apoiada nalguma técnica, a atividade matemática também tem o seu conjunto de técnicas - certos hábitos de raciocínio e modos de pensamento particulares.

São estes modos de pensamento, que se ganham ao estudar matemática e que podem ser aplicados a outros domínios, que constituem um dos principais valores sociais e económicos da Matemática. É possível argumentar que este valor é até superior ao das aplicações da Matemática, normalmente usado para justificar o investimento de tempo e dinheiro nos matemáticos.

Escusado será dizer que estes modos de pensamento não se adquirem com a aplicação cega de fórmulas e métodos que não se entendem na resolução de "problemas-tipo".

Finalmente, o estudo da Matemática tem um valor cultural análogo ao do estudo da música por pessoas que não irão necessariamente ser músicos profissionais. Tal como a música, a matemática é um dos pilares da civilização humana.

0.4. **A necessidade do rigor.** Durante quase toda a história da Matemática os objetos matemáticos não tinham uma definição rigorosa e eram manipulados intuitivamente. O ideal grego da demonstração existia (pelo menos para alguns matemáticos) mas era impossível implementá-lo completamente.

Exemplo 0.5. *É fácil cometer erros quando apelamos à intuição mesmo em situações muito simples: suponham que a Terra é uma esfera perfeita e que se tem uma corda que dá uma volta ao planeta ajustada à superfície. Se aumentarmos o comprimento da corda*

[1]Também há criatividade na aplicação da Matemática a outros domínios.

em um metro e a afastarmos uniformemente da superfície, conseguiria um gato passar por debaixo da corda? Intuitivamente, a maioria das pessoas (eu incluído) diria que não, mas é fácil de verificar com uma conta muito elementar que o gato passa facilmente.

No século XIX tornou-se claro que este estado das coisas era uma obstrução ao progresso - certas perguntas naturais não tinham resposta sem uma clarificação dos objetos que estavam a ser estudados. Vocês já beneficiaram desta clarificação quando estudaram o conceito abstrato de função que data do século XIX.

Isso levou à criação da Matemática tal como ela existe hoje: Partindo de um conjunto de axiomas que descrevem propriedades básicas dos conjuntos[2], todos os objetos matemáticos são definidos rigorosamente (começando pelos números naturais, as operações entre eles, etc.) e as suas propriedades são deduzidas dos axiomas da teoria dos conjuntos usando as regras da lógica.

O parágrafo anterior descreve um ideal pouco prático[3] - as demonstrações seriam gigantes e humanamente incompreensíveis! Na prática, para abreviar, a Matemática usa linguagem corrente (embora com notação especializada) mas de tal forma que, com paciência suficiente, os argumentos numa demonstração possam ser reduzidos a uma aplicação mecânica das regras da lógica.

0.6. **A importância das definições.** As definições desempenham um papel absolutamente essencial no entendimento e progresso da Matemática. São, de alguns pontos de vista, principalmente quando a complexidade dos objetos em questão é muito grande, o mais importante.

Uma definição isola as propriedades centrais de um dado objeto (e nesse processo cria-o enquanto objeto matemático!). Geralmente estas propriedades são derivadas por *abstração* a partir de vários exemplos concretos. Este processo de abstração permite identificar relações inesperadas entre objetos e situações aparentemente muito diferentes, desempenhando assim um papel fundamental na clarificação, organização e síntese da informação em Matemática.

Este processo de isolar propriedades que descrevem a essência de um dado tipo de objeto, explorando depois as consequência lógicas destas propriedades chama-se o *método axiomático* (as propriedades incluídas na definição são tomadas como ponto de partida ou axiomas). Surpreendentemente, é com frequência mais fácil efetuar esta exploração em abstrato do que em casos concretos que tenham motivado a definição.

Exemplo 0.7. *É bastante mais fácil verificar a associatividade da composição de funções* $((f\circ g)\circ h)(x) = (f\circ(g\circ h))(x)$ *fazendo as contas em abstrato do que para funções específicas dadas por fórmulas complicadas!*

A partir do século XIX o método axiomático revelou-se incrivelmente poderoso no estabelecimento de novas verdades matemáticas e de pontes entre as várias áreas da Matemática. Tornou-se assim o "método habitual" em Matemática.

[2]O sistema axiomático mais habitual é o ZFC - Zermelo-Fraenkel-Choice.

[3]Mais prático hoje em dia - ver o projecto Xena `https://www.ma.imperial.ac.uk/~buzzard/xena/`

Uma razão psicológica para o sucesso deste método é que as definições permitem encapsular conceitos arbitrariamente complexos de forma a que possam ser manuseados mentalmente como blocos. A criação destas "caixas" na nossa mente, assim como de uma teia de ligações com outras caixas e procedimentos é um processo físico no cérebro, análogo ao treino desportivo, e como este, requer tempo, prática, e envolve a superação de frustração.

É a este método axiomático a que muitos de vós irão agora começar a ser expostos e é de esperar que isso suscite algum choque. É necessário ultrapassar este choque para fazer progresso no vosso conhecimento e entendimento da Matemática.

Algumas definições resultam de um esforço de síntese de muitas pessoas ao longo de séculos e têm muito pouco de intuitivo quando as vemos pela primeira vez. A sua justificação é a utilidade que têm na promoção do entendimento da Matemática. É necessário ter confiança que as definições que estudam se virão a revelar úteis por mais incompreensíveis que pareçam inicialmente. Com o hábito e a experiência acabarão por tornar-se naturais.

Quando confrontados com uma nova definição temos que a tornar familiar, pensando em exemplos e não exemplos, no que aconteceria se alguma condição da definição fosse alterada e também na maneira como a definição é utilizada na demonstração de propriedades do objeto que está a ser definido. É este processo que nos permite ganhar familiaridade com uma definição e adquirir assim um novo conceito.

Note-se que este entendimento não é independente da memorização, apesar do que vos possa ter sido dito no passado. A natureza hierárquica e cumulativa da Matemática faz com que não seja possível ir progredindo sem que certas noções e procedimentos se tornem automáticos, o que requer, em particular, alguma memorização.

0.8. **O que é a Álgebra Linear? - II.** Cada área da matemática tem os seus objetos de estudo específicos. Na Álgebra Linear os objetos que são estudados (os esquilos por assim dizer) são os *espaços vetoriais* e as relações entre eles que se manifestam através de *transformações lineares*. Seguindo a lógica descrita acima, a cadeira de Álgebra Linear devia iniciar-se com a apresentação da definição de um espaço vetorial (ou equivalentemente dos axiomas para um espaço vetorial) e a dedução lógica das suas propriedades. É essa a abordagem em [Ax] e [FIS].

Não é isso que vou fazer para vos dar algum tempo de adaptação ao ritmo de trabalho mais elevado no ensino superior e à novidade do ambiente. Vamos começar nas primeiras aulas por aprender a fazer algumas contas que estão mais próximas da vossa experiência matemática anterior e que de qualquer forma teriam que ser estudadas um pouco mais à frente. É também nesse contexto que vamos introduzir os argumentos formais - as demonstrações. Iremos dar como adquirido (como axiomas se quiserem) várias propriedades dos sistemas de números (reais, complexos) e construções com conjuntos que aprenderam no vosso estudo prévio da Matemática. Esta é a abordagem seguida em [HK] que é uma excelente fonte para quem estiver interessado em aprofundar o seu estudo de Álgebra Linear.

0.9. **Conselhos práticos para o estudo.** O ritmo de avanço na matéria será muito mais rápido do que aquele a que provavelmente estão habituados. O número de horas *médio* de estudo que se calcula necessário para seguirem a matéria é de 8h por semana *além* das 4h de aulas (ou 12h se não vierem às aulas). Isto é o valor para um aluno ou aluna média. Alguns de vós precisarão de menos, outros de mais. Mesmo que deixem dois dias inteiros de estudo para a época de exames continua a ser de 7h por semana de estudo *concentrado*. Comecem imediatamente a trabalhar. Se se atrasarem será muito mais difícil recuperar.

Eu e os restantes professores estamos aqui para vos ajudar a aprender. A aprendizagem é um trabalho de construção interior que cada um de vós terá de realizar mas nós podemos ajudar a superar bloqueios e sugerir estratégias úteis. A discussão com os vossos colegas também ajuda (e a partilha da frustração é terapêutica).

(1) Aproveitem as várias formas de contacto comigo e com os vossos colegas: o diálogo nas e após ou antes das aulas, as aulas de dúvidas.
(2) Façam perguntas sem medo de que possam ser tontas! O medo a errar ou a fazer má figura é um dos principais obstáculos à aprendizagem. Posso garantir-vos por experiência própria que até os melhores matemáticos do mundo fazem de vez em quando perguntas tontas.
(3) Explorem a bibliografia recomendada (aconselho principalmente as referências [HK, Ax, FIS]) que contem mais exemplos, exercícios e pontos de vista. A quantidade de recursos online, desde videos a listas de exercícios e exercícios resolvidos é praticamente ilimitada mas é fundamental passar uma boa parte do tempo de estudo a ler e refletir sobre a matéria e a tentar resolver exercícios por vós próprios.
(4) Tentem ser pró-ativos na aprendizagem. Não se cinjam a adquirir técnicas de resolução de exercícios. Interroguem-se uns aos outros e a vós próprios sobre o significado do que estão a aprender. Tentem inventar problemas. Façam conjeturas. Encarem a aprendizagem como a exploração de um mundo novo!
(5) O guia de apoio ao estudante do Técnico tem bons conselhos sobre a organização do estudo. Em particular é muito importante tomar conta da vossa saúde física e mental.

Uma nota final. Tentem não se preocupar demasiado com as classificações que vão obter nas cadeiras. Para entrarem no Técnico tiveram provavelmente que se preocupar bastante com as notas até ao momento, mas a partir de agora as notas não são verdadeiramente importantes. Elas medem o vosso desempenho durante a avaliação que é em grande parte função de características como a rapidez e a capacidade de lidar com a pressão. O verdadeiro objetivo do estudo universitário é o vosso desenvolvimento e crescimento intelectual que está apenas fracamente correlacionado com o resultado da avaliação. Por exemplo, a meu ver, é mais proveitoso dedicar tempo a lutar com exercícios com os quais sintam dificuldade do que a fazer muitos exercícios de um certo tipo para aumentar a rapidez de resolução. Isto independentemente de a segunda estratégia poder resultar numa melhor classificação final.

Posso garantir-vos com base na experiência de muitos anos que:

(1) Ninguém estará interessado nas classificações que obtiveram uma vez que saiam do IST.
(2) A correlação entre a média de curso e o desempenho profissional posterior (mesmo na academia) não é muito elevada. Isto deve-se ao facto de as qualidades que são requeridas para ter notas muito elevadas não serem *nem necessárias, nem suficientes* para ter sucesso em tarefas de grande exigência (mesmo na investigação científica).

Não quero com isto encorajar-vos a ter más notas! Só vos quero pedir que não as levem demasiado a sério nem as usem como único instrumento de medida daquilo que é verdadeiramente importante: o vosso desenvolvimento intelectual (e ainda menos como um julgamento pessoal).

Outro aspeto que vale a pena realçar é o carácter cultural da atribuição das classificações. Há países (EUA, Itália por exemplo) nos quais a grande maioria de vós obteria provavelmente a classificação máxima possível e outros (França por exemplo) onde a nota mais alta dificilmente passaria de 65% ou 70%. Eu irei seguir a tradição corrente em Portugal que é intermédia entre estas: devem esperar relativamente poucas notas acima de 17 e uma média à volta de 14.

Para a maioria de vós serão notas muito abaixo daquelas a que estão habituados. Não desanimem nem levem isso muito a peito! As notas não são importantes!

Bom trabalho!

1. O método de Gauss

O método de Gauss (ou método de eliminação de Gauss) é um método para resolver sistemas lineares cuja ideia é a simplificação do sistema através da eliminação sucessiva de variáveis.

Definição 1.1. *Um* sistema linear *de m equações a n incógnitas é uma expressão da forma*

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m \end{cases} \tag{1}$$

onde a_{ij}, x_j, b_i para $1 \leq i \leq m$, $1 \leq j \leq n$ denotam números reais (ou complexos). Os números a_{ij} chamam-se os coeficientes do sistema*, os x_i são as* incógnitas *e os b_i os* termos independentes*. Se os termos independentes são nulos (isto é $b_i = 0$ para todo o i) o sistema diz-se* homogéneo*.*

Estamos interessados em saber se um sistema admite soluções (isto é, se existem números $x_1, \ldots, x_n$ tais que as relações (1) são satisfeitas). Quando isto acontece diz-se que o sistema é *possível*, senão é *impossível*. Quando existem soluções, queremos descrevê-las. Em particular queremos saber se a solução é única (nesse caso diz-se que o sistema é *determinado*) ou não, caso em que o sistema se diz *indeterminado*.

Observe-se que um sistema homogéneo é sempre possível. Tem pelo menos a solução $x_j = 0$ para todo o j, que se chama a *solução trivial*.

Observação 1.2. *Toda a teoria que vamos desenvolver durante o próximo par de meses aplica-se mais geralmente. Os números reais ou complexos podem ser substituídos pelos elementos de qualquer* corpo *(um conjunto com duas operações binárias - soma e multiplicação - que são comutativas, associativas, têm elemento neutro, a multiplicação é distributiva relativamente à soma, todos os elementos têm inverso relativamente à soma e todos os elementos excepto o elemento neutro da soma têm inverso multiplicativo).*

Um exemplo familiar de corpo além dos conjuntos $\mathbb{R}$ e $\mathbb{C}$ dos números reais e complexos com as suas operações habituais é o conjunto $\mathbb{Q}$ dos números racionais, também com a soma e produto habituais.

Exemplos provavelmente menos familiares são o conjunto $\mathbb{F}_2 = \{0, 1\}$ com a soma e produto definidas tomando o resto da divisão por 2 da soma e produto usuais e $\mathbb{Q}[\sqrt{2}] = \{a + b\sqrt{2} \colon a, b \in \mathbb{Q}\} \subset \mathbb{R}$ com as operações habituais.

Mais geralmente, se p é um número primo, o conjunto $\mathbb{F}_p = \{0, 1, \ldots, p-1\}$ com a soma e produto definidos tomando o resto da divisão por p após o cálculo da soma e produto usuais nos inteiros forma um corpo.

O *método de eliminação de Gauss* é o seguinte algoritmo para simplificar um sistema de equações lineares:

(1) Identificar a primeira variável que ocorre de facto no sistema (isto é, que tem coeficiente não nulo nalguma das equações do sistema).
(2) Se o coeficiente dessa variável na primeira equação for nulo, trocar a primeira equação com outra na qual o coeficiente não é nulo
(3) Subtrair um múltiplo conveniente da primeira equação às restantes de forma a eliminar nelas a variável em questão (isto é tornar o coeficiente dessa variável nulo)
(4) Regressar ao passo (1) considerando apenas o sistema que se obtém esquecendo a primeira equação, a não ser que o sistema fique reduzido a uma única equação ou que deixe de haver variáveis, caso em que o algoritmo termina.

Exemplo 1.3. *Considere-se o sistema*

$$\begin{cases} 0x_1 + 0x_2 + 2x_3 - x_4 = 5 \\ 0x_1 + x_2 + 0x_3 + 3x_4 = 1 \\ 0x_1 + 2x_2 + x_3 + x_4 = 2 \end{cases}$$

A primeira variável que ocorre no sistema é x_2. Uma vez que o coeficiente de x_2 na primeira equação é 0, trocamos a primeira equação com a segunda (também poderíamos trocar com a terceira). Obtemos então o sistema

$$\begin{cases} x_2 \phantom{{}+2x_3} + 3x_4 = 1 \\ \phantom{x_2 + {}} 2x_3 - x_4 = 5 \\ 2x_2 + x_3 + x_4 = 2 \end{cases}$$

Subtraímos agora à terceira equação o dobro da primeira para eliminar a variável x_2 obtendo

$$\begin{cases} x_2 \quad\quad + 3x_4 = 1 \\ \quad 2x_3 - x_4 = 5 \\ \quad x_3 - 5x_4 = 0 \end{cases}$$

Voltamos agora ao início mas consideramos apenas as duas últimas equações. A primeira variável é agora x_3 e o seu coeficiente na primeira linha (que é a segunda linha do sistema inicial) é não nulo, pelo que não é necessário trocar a ordem das equações. Subtraindo metade da segunda equação à terceira obtemos o sistema

$$\begin{cases} x_2 \quad\quad + 3x_4 = 1 \\ \quad 2x_3 - x_4 = 5 \\ \quad\quad -\frac{9}{2}x_4 = -\frac{5}{2} \end{cases} \tag{2}$$

O sistema (2) é fácil de resolver começando pela equação de baixo e substituindo repetidamente os resultados obtidos nas equações de cima: da última equação obtemos $x_4 = \frac{5}{9}$ e substituindo na segunda equação obtemos

$$2x_3 = 5 + \frac{5}{9} \Leftrightarrow x_3 = \frac{25}{9}$$

Finalmente substituindo na primeira equação (em geral precisaríamos também do valor de x_3 mas neste sistema isso não acontece) obtemos

$$x_2 = 1 - 3 \cdot \frac{5}{9} = -\frac{2}{3}$$

O conjunto das soluções do sistema é portanto

$$\{(x_1, -\tfrac{2}{3}, \tfrac{25}{9}, \tfrac{5}{9}) \colon x_1 \in \mathbb{R}\} \tag{3}$$

Em particular o sistema é possível e indeterminado.

É um desperdício de tempo escrever as variáveis durante a aplicação dos passos do algoritmo acima. Podemos apenas escrever os coeficientes e termos independentes dos vários sistemas. O procedimento aplicado no exemplo anterior pode então ser abreviado da seguinte forma:

$$\left[\begin{array}{cccc|c} 0 & 0 & 2 & -1 & 5 \\ 0 & 1 & 0 & 3 & 1 \\ 0 & 2 & 1 & 1 & 2 \end{array}\right] \xrightarrow{L_1 \leftrightarrow L_2} \left[\begin{array}{cccc|c} 0 & 1 & 0 & 3 & 1 \\ 0 & 0 & 2 & -1 & 5 \\ 0 & 2 & 1 & 1 & 2 \end{array}\right] \xrightarrow{L_3 - 2L_1} \left[\begin{array}{cccc|c} 0 & 1 & 0 & 3 & 1 \\ 0 & 0 & 2 & -1 & 5 \\ 0 & 0 & 1 & -5 & 0 \end{array}\right]$$

$$\xrightarrow{L_3 - \frac{1}{2}L_2} \left[\begin{array}{cccc|c} 0 & 1 & 0 & 3 & 1 \\ 0 & 0 & 2 & -1 & 5 \\ 0 & 0 & 0 & -\frac{9}{2} & -\frac{5}{2} \end{array}\right] \tag{4}$$

As tabelas de números que aparecem acima chamam-se *matrizes* e são objetos fundamentais na álgebra linear. A linha a tracejado antes da última coluna destina-se a lembrar que estamos a resolver um sistema não homogéneo e que a última coluna é formada pelos termos independentes. Quando é claro do contexto a linha a tracejado é por vezes omitida. Quando o sistema é homogéneo a última coluna (formada só por 0s) é omitida.

Exemplo 1.4. *Vamos resolver o sistema*

$$\begin{cases} x+3y+2z=0 \\ 4y+z=2 \\ -2x-2y-3z=1 \end{cases}$$

Aplicando o método de Gauss obtemos

$$\left[\begin{array}{ccc|c} 1 & 3 & 2 & 0 \\ 0 & 4 & 1 & 2 \\ -2 & -2 & -3 & 1 \end{array}\right] \xrightarrow{L_3+2L_1} \left[\begin{array}{ccc|c} 1 & 3 & 2 & 0 \\ 0 & 4 & 1 & 2 \\ 0 & 4 & 1 & 1 \end{array}\right] \xrightarrow{L_3-L_2} \left[\begin{array}{ccc|c} 1 & 3 & 2 & 0 \\ 0 & 4 & 1 & 2 \\ 0 & 0 & 0 & -1 \end{array}\right]$$

A última equação do sistema descrito pela matriz em que termina o método de Gauss é $0x+0y+0z=-1$*, que é impossível. Conclui-se que o sistema inicial é impossível.*

Definição 1.5. *Sejam* m,n *números naturais. Uma* matriz $m \times n$ *de números reais ou complexos é uma função*[4] $\{1,\ldots,m\}\times\{1,\ldots,n\} \to \mathbb{R}$ *(ou* $\mathbb{C}$*). É habitual representar uma tal função por uma tabela de números*

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \qquad \text{onde } a_{ij} \text{ é o valor da função em } (i,j).$$

m *é o* número de linhas *da matriz, enquanto que* n *é o* número de colunas. *Diz-se que uma matriz está* em escada de linhas *se todas as linhas nulas estão em baixo e se a primeira entrada não nula de cada linha, que se denomina por* pivot, *está para a esquerda do pivot da linha abaixo. Isto é,* $[a_{ij}]_{1\leq i\leq m, 1\leq j\leq n}$ *está em escada de linhas quando, para todos os* $1 \leq i \leq m-1$ *e* $0 \leq j,k \leq m$,

$$a_{ij}=0 \text{ para } j \leq k \quad \Rightarrow \quad a_{i+1\ j}=0 \text{ para } j \leq k+1$$

onde estamos a deixar j,k *tomar o valor* 0 *convencionando que entradas que têm um dos índices iguais a* 0 *são nulas.*

Note-se que, em termos das matrizes associadas aos sistemas, o que o método de Gauss faz é colocar a matriz do sistema em escada de linhas.

Após a aplicação do método de Gauss temos ainda que resolver iterativamente as equações do sistema, começando pela que está mais abaixo. Este processo pode ser feito de forma muito mais eficiente, efetuando operações semelhantes às do método de Gauss. Este novo algoritmo, uma continuação do método de Gauss, chama-se *Método de Gauss-Jordan* e consiste em, dada uma matriz *em escada de linhas*,

[4] Sendo X,Y conjuntos arbitrários, $X\times Y$ designa o *produto cartesiano* de X e Y, que é por definição o conjunto dos pares ordenados (x,y) em que x é um elemento de X e y um elemento de Y. Por exemplo $\mathbb{R}\times\mathbb{R}$ é o conjunto dos pares ordenados de números reais que se abrevia normalmente por $\mathbb{R}^2$.

(1) Multiplicar cada linha não nula pelo inverso do pivot de forma a tornar o pivot igual a 1.
(2) Subtrair múltiplos apropriados das linhas acima de cada linha com pivot até que todas as entradas acima dos pivots fiquem nulas.

Vamos aplicar este algoritmo à matriz em escada de linhas (4) que resultou do Exemplo 1.3.

Exemplo 1.6.

$$\left[\begin{array}{cccc|c} 0 & 1 & 0 & 3 & 1 \\ 0 & 0 & 2 & -1 & 5 \\ 0 & 0 & 0 & -\frac{9}{2} & -\frac{5}{2} \end{array}\right] \underset{-\frac{2}{9}L_3}{\xrightarrow{\frac{1}{2}L_2}} \left[\begin{array}{cccc|c} 0 & 1 & 0 & 3 & 1 \\ 0 & 0 & 1 & -\frac{1}{2} & \frac{5}{2} \\ 0 & 0 & 0 & 1 & \frac{5}{9} \end{array}\right] \underset{L_2+\frac{1}{2}L_3}{\xrightarrow{L_1-3L_3}} \left[\begin{array}{cccc|c} 0 & 1 & 0 & 0 & -\frac{2}{3} \\ 0 & 0 & 1 & 0 & \frac{25}{9} \\ 0 & 0 & 0 & 1 & \frac{5}{9} \end{array}\right]$$

Recuperamos assim o conjunto de soluções (3) *obtido acima.*

Quando há muitas equações, o algoritmo de Gauss-Jordan é muito mais eficiente do que o processo de substituições sucessivas que usámos antes. Observamos ainda que, na aplicação do passo (2) do método de Gauss-Jordan, os cálculos serão geralmente mais rápidos se eliminarmos as entradas acima do último pivot (isto é do pivot da linha não nula mais abaixo) e repetirmos sucessivamente o processo para as matrizes que se obtêm esquecendo a última linha; isto é, se procedermos à eliminação das entradas acima dos pivots percorrendo os pivots da direita para a esquerda ou, equivalentemente, de baixo para cima.

Definição 1.7. *Diz-se que uma matriz está* em escada de linhas reduzida *se está em escada de linhas, os pivots são todos iguais a* 1 *e as entradas acima dos pivots são todas* 0.

O algoritmo de Gauss-Jordan coloca portanto uma matriz em escada de linhas numa matriz em escada de linhas reduzida.

Exemplo 1.8. *Vamos resolver o sistema homogéneo*

$$\begin{cases} y + 4w = 0 \\ x - 2y + 3z = 0 \\ 2x - 6y + 16w = 0 \end{cases}$$

Recorde-se que neste caso não incluímos a coluna de 0*s correspondente aos termos dependentes. Obtemos assim*

$$\begin{bmatrix} 0 & 1 & 0 & 4 \\ 1 & -2 & 3 & 0 \\ 2 & -6 & 0 & 16 \end{bmatrix} \xrightarrow{L_1 \leftrightarrow L_2} \begin{bmatrix} 1 & -2 & 3 & 0 \\ 0 & 1 & 0 & 4 \\ 2 & -6 & 0 & 16 \end{bmatrix} \xrightarrow{L_3-2L_1} \begin{bmatrix} 1 & -2 & 3 & 0 \\ 0 & 1 & 0 & 4 \\ 0 & -2 & -6 & 16 \end{bmatrix}$$

$$\xrightarrow{L_3+2L_2} \begin{bmatrix} 1 & -2 & 3 & 0 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & -6 & 24 \end{bmatrix} \xrightarrow{-\frac{1}{6}L_3} \begin{bmatrix} 1 & -2 & 3 & 0 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & -4 \end{bmatrix} \xrightarrow{L_1-3L_3} \begin{bmatrix} 1 & -2 & 0 & 12 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & -4 \end{bmatrix}$$

$$\xrightarrow{L_1+2L_2} \begin{bmatrix} 1 & 0 & 0 & 20 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & -4 \end{bmatrix}$$

Obtemos assim a seguinte solução para o sistema:

$$\begin{cases} x = -20w \\ y = -4w \\ z = 4w \end{cases} \qquad \text{com } w \in \mathbb{R} \text{ qualquer.}$$

Exemplo 1.9. *Vamos resolver o sistema linear homogéneo*

$$\begin{cases} x - y + 2z + w - v = 0 \\ 2x - 2y + z - w + 2v = 0 \\ x - y + 5z + 4w - 5v = 0 \end{cases}$$

Aplicando os métodos de Gauss e Gauss-Jordan temos

$$\begin{bmatrix} 1 & -1 & 2 & 1 & -1 \\ 2 & -2 & 1 & -1 & 2 \\ 1 & -1 & 5 & 4 & -5 \end{bmatrix} \underset{L_3-L_1}{\overset{L_2-2L_1}{\longrightarrow}} \begin{bmatrix} 1 & -1 & 2 & 1 & -1 \\ 0 & 0 & -3 & -3 & 4 \\ 0 & 0 & 3 & 3 & -4 \end{bmatrix} \overset{L_3+L_2}{\longrightarrow} \begin{bmatrix} 1 & -1 & 2 & 1 & -1 \\ 0 & 0 & -3 & -3 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\overset{-\frac{1}{3}L_2}{\longrightarrow} \begin{bmatrix} 1 & -1 & 2 & 1 & -1 \\ 0 & 0 & 1 & 1 & -\frac{4}{3} \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \overset{L_1-2L_2}{\longrightarrow} \begin{bmatrix} 1 & -1 & 0 & -1 & \frac{5}{3} \\ 0 & 0 & 1 & 1 & -\frac{4}{3} \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Ou seja, o conjunto solução deste sistema é

$$\{(y + w - \tfrac{5}{3}v, y, -w + \tfrac{4}{3}v, w, v) \colon y, w, v \in \mathbb{R}\}$$

Os dois exemplos acima ilustram a seguinte observação relativa à solução de sistemas homogéneos por este método:

- As colunas com pivots correspondem às *variáveis dependentes* do sistema que são expressas em função das restantes.
- As colunas sem pivots correspondem às *variáveis livres* cujo valor pode ser atribuído arbitrariamente numa solução.

Num sistema não homogéneo, o sistema é impossível se houver um pivot na última coluna (como acontece no Exemplo 1.4). Quando o sistema é possível, as colunas com pivot correspondem às variáveis dependentes e as restantes, com excepção da última, às variáveis livres.

Definição 1.10. *A* característica *de uma matriz*[5] *A é o número de pivots que se obtém no final da aplicação do método de Gauss (ou Gauss-Jordan).*

Alternativamente, a característica é o número de linhas não nulas na matriz que resulta da aplicação do método de Gauss (ou Gauss-Jordan). Veremos que a característica nos indica *o número mínimo de equações necessárias para descrever a solução do sistema.*

Note-se que não é imediatamente claro que a definição de característica faça sentido pois há alguma indeterminação no método de Gauss relativa à escolha das trocas de linha. Podia acontecer que escolhas diferentes durante a aplicação do algoritmo conduzissem a

[5] Em inglês "rank of a matrix".

matrizes com números diferentes de pivots no final. Vamos ver que isso não pode acontecer, mas primeiro comecemos por analisar exatamente a razão pela qual os métodos de Gauss e Gauss-Jordan produzem sistemas *equivalentes* ao inicial, isto é, sistemas que têm exatamente as mesmas soluções.

Suponhamos que temos um sistema linear

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m \end{cases} \tag{5}$$

Se $(x_1, \ldots, x_n)$ é uma solução do sistema, então para qualquer escolha de $c_1, \ldots, c_m \in \mathbb{R}$ (ou $\mathbb{C}$ consoante os escalares que estejamos a considerar) a seguinte relação será verificada

$$c_1(a_{11}x_1 + \ldots + a_{1n}x_n) + \ldots + c_m(a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n) = c_1b_1 + \ldots + c_mb_m \tag{6}$$

A expressão (6) diz-se uma *combinação linear* das equações do sistema (5). Obtém-se multiplicando a i-ésima equação pela constante c_i e somando as equações resultantes. Os números c_i dizem-se os *coeficientes* da combinação linear. Concretizando, a combinação linear com coeficientes 2 e -3 das equações

$$x + y = 3 \qquad 2x - 5y = 2 \tag{7}$$

é a equação

$$2(x + y) - 3(2x - 5y) = 2 \cdot 3 - 3 \cdot 2 \Leftrightarrow -4x + 17y = 0 \tag{8}$$

Qualquer solução do sistema formado pelas duas equações (7) será também uma solução da equação (8).

Observação 1.11. *O conceito de combinação linear é talvez o conceito central da Álgebra Linear. Informalmente, uma combinação linear de coisas é uma expressão que se obtém multiplicando cada coisa por um escalar e somando tudo. Por exemplo, admitindo que se pode multiplicar mamíferos por escalar e somá-los,* 2*morcego-*3*castor é uma combinação linear de mamíferos.*

Quando executamos um passo do algoritmo de Gauss ou Gauss-Jordan, as equações do novo sistema são (por definição do algoritmo) combinações lineares das do sistema anterior. Portanto uma solução do sistema antes da aplicação do passo é ainda uma solução do sistema seguinte.

As combinações lineares envolvidas nos algoritmos são muito simples. Chamando S ao sistema inicial e S' ao sistema obtido após aplicação de um passo do algoritmo e usando a notação L_i (respetivamente L'_i) para a i-ésima equaçao do sistema S (respetivamente S'), temos após um passo do método

$$\begin{cases} L'_i = L_j \\ L'_j = L_i \end{cases} \text{com } j \neq i, \quad L'_i = \alpha L_i \text{ com } \alpha \neq 0, \quad \text{ou } L'_i = L_i - \alpha L_j \text{ com } j \neq i \tag{9}$$

permanecendo as restantes equações inalteradas. Note-se em particular, para utilização posterior, que na terceira operação só a i-ésima linha é alterada. Em particular, $L'_j = L_j$.

A observação fulcral é que as expressões acima permitem também escrever as linhas do sistema S como combinações lineares das linhas de S':

$$\begin{cases} L_i = L'_j \\ L_j = L'_i \end{cases} \quad \text{com } j \neq i, \quad L_i = \frac{1}{\alpha} L'_i \text{ com } \alpha \neq 0, \quad \text{ou } L_i = L'_i + \alpha L'_j \text{ com } j \neq i$$

(onde no último caso usámos o facto de L_j e L'_j serem iguais). Conclui-se que as soluções do sistema S' são também soluções do sistema S e portanto que os sistemas S e S' têm exatamente as mesmas soluções. Uma vez que isto acontece durante todas os passos do método conclui-se que *todos os sistemas que ocorrem ao longo da aplicação dos métodos de Gauss e Gauss-Jordan são equivalentes*, isto é, todos têm exatamente o mesmo conjunto de soluções.

Para terminar esta nossa discussão inicial dos sistemas lineares vamos agora provar que a matriz em escada de linhas reduzida no final do método de Gauss-Jordan é independente de quaisquer escolhas, o que mostra que a Definição 1.10 faz sentido (diz-se que a característica de uma matriz está *bem definida*).

Na realidade iremos demonstrar algo mais geral. As operações (9) sobre as linhas de uma matriz usadas durante o método de Gauss e Gauss-Jordan designam-se por *operações elementares*. Duas matrizes A, B com as mesmas dimensões dizem-se *equivalentes mediante operações elementares* se existe uma sequência finita de operações elementares que transforma A em B. Qualquer aplicação dos métodos de Gauss e Gauss-Jordan a uma matriz A produz uma sequência de matrizes equivalentes[6] mediante operações elementares que termina numa matriz em escada de linhas reduzida. Iremos provar que para cada matriz A existe exatamente uma matriz em escada de linhas reduzida que é equivalente a A mediante operações elementares.

A demonstração utilizará um género de argumento que se diz *por redução ao absurdo* e que se baseia no seguinte facto simples da lógica: Se uma afirmação P implica outra afirmação Q e Q é falsa, então P é necessariamente falsa. Em símbolos:

$$((P \Rightarrow Q) \wedge \neg Q) \Rightarrow \neg P$$

Este facto permite-nos provar a validade de uma afirnação A se conseguirmos deduzir uma falsidade a partir da sua negação $\neg A$ (isto é se $\neg A$ se reduzir ao absurdo). Conclui-se então que a afirmação $\neg A$ é falsa, ou seja que A é verdadeira.

[6] Esta relação entre duas matrizes é um exemplo de uma *relação de equivalência* num conjunto. Dado um conjunto X, uma relação $\sim$ entre pares de elementos de X, diz-se uma relação de equivalência se para todos os $x, y, z \in X$ se verifica:

(1) Reflexividade: $x \sim x$
(2) Simetria: $x \sim y \Rightarrow y \sim x$
(3) Transitividade: $x \sim y \wedge y \sim z \Rightarrow x \sim z$

Teorema 1.12. *Sejam m, n números naturais e A uma matriz $m \times n$ de números reais ou complexos. Existe uma única matriz em escada de linhas reduzida equivalente a A mediante operações elementares.*

Dem. A existência é garantida pelos algoritmos de Gauss e Gauss-Jordan. Resta-nos demonstrar a unicidade. A demonstração é por indução no número n das colunas de A.

Para a base da indução precisamos de mostrar que se A é uma matriz com uma única coluna, o resultado é verdadeiro. As únicas matrizes em escada de linhas reduzidas com uma coluna são a matriz nula e a matriz

$$\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{10}$$

pelo que é suficiente ver que uma matriz coluna A não pode ser simultaneamente equivalente a estas duas. Isso é verdade porque a única matriz que é equivalente à matriz nula mediante operações elementares é a própria matriz nula. Isto conclui a prova da base da indução.

Para o passo da indução vamos admitir que a unicidade é válida para matrizes com n colunas. Queremos concluir que é também válida para matrizes com $n+1$ colunas. Designemos por $X_{\leq n}$ a matriz $m \times n$ que se obtém da matriz $m \times (n+1)$ X suprimindo a última coluna[7]. Observemos que:

(i) Se X está em escada de linhas reduzida, o mesmo acontece com $X_{\leq n}$.
(ii) Se X' resulta da aplicação de uma operação elementar a X então $X'_{\leq n}$ resulta da aplicação da mesma operação elementar a $X_{\leq n}$.

Seja A uma matriz $m\times(n+1)$ e suponhamos que B e C são matrizes em escada de linhas reduzida que se obtêm de A por operações elementares. Por (i), as matrizes $B_{\leq n}$ e $C_{\leq n}$ estão também em escada de linhas reduzida. Por (ii), $B_{\leq n}$ e $C_{\leq n}$ resultam da aplicação de operações elementares a $A_{\leq n}$. Por hipótese de indução conclui-se que $B_{\leq n} = C_{\leq n}$.

Vamos admitir com vista a chegar a um absurdo que $B \neq C$. Então B e C diferem numa entrada da última coluna. Seja i tal que $b_{i\ n+1} \neq c_{i\ n+1}$. Recorde-se que os sistemas homogéneos determinados por A, B, e C são equivalentes. Subtraindo as i-ésimas equações dos sistemas correspondentes a B e C obtemos a equação

$$(b_{i\ n+1} - c_{i\ n+1})x_{n+1} = 0$$

(pois $b_{ij} = c_{ij}$ para $j \leq n$). Como o coeficiente de x_{n+1} é não nulo, conclui-se que todas as soluções do sistema determinado por A (ou B ou C) satisfazem $x_{n+1} = 0$.

Tal só sucede num sistema de m equações a $(n+1)$ incógnitas cuja matriz dos coeficientes está em escada de linhas quando existe um pivot na última coluna (só assim x_{n+1} não é uma variável livre). Conclui-se que tanto B como C têm um pivot na coluna $n+1$.

Numa matriz em escada de linhas reduzida, um pivot na última coluna ocorre exatamente à direita da primeira linha de 0s na matriz obtida ao suprimir a última coluna. Ou seja, sabendo que B e C têm um pivot na última coluna, a posição do pivot é determinada por

[7]Esta notação *ad hoc* não voltará a ser usada depois desta demonstração.

$B_{\leq n} = C_{\leq n}$ e portanto é igual para B e C. Como B e C estão em escada de linhas reduzida, todas as entradas da última coluna de B e C são 0 excepto a entrada correspondente ao pivot, que é 1. Conclui-se que as últimas colunas de B e de C são iguais e portanto que $B = C$.

A afirmação anterior contradiz a nossa hipótese que $B \neq C$ e conclui portanto a demonstração do passo de indução. □

Corolário 1.13. *A característica de uma matriz está bem definida, isto é, é independente das escolhas realizadas durante a aplicação do método de Gauss.*

1.14. **Exercícios.**

1. Resolva os seguintes sistemas (se possível)
   (a) $\begin{cases} x - 2y + 3z - w = 1 \\ 3x - y + 2z + 5w = 2 \\ -3x + 6y - 9z + 3w = -6 \end{cases}$ (b) $\begin{cases} -2v + 3w = 3 \\ 3u + 6v - 3w = -2 \\ 6u + 6v + 3w = 5 \end{cases}$
   Resolução.
2. Determine o conjunto dos polinómios reais p de grau menor ou igual a 3 tais que $p(1) = 0$, $p(2) = 3$ e $p'(0) = 0$. Resolução.
3. (Hefferon Ex. I.2.32) Num planeta distante vivem extraterrestres de três cores: amarelos, azuis e verdes. Quando dois extraterrestres de cores diferentes se encontram mudam imediatamente para a cor restante. Admitindo que inicialmente há 15 criaturas amarelas, 13 azuis e 17 verdes, é ou não possível que venham a ficar todas da mesma cor? Resolução.
4. (Hefferon Ex. 1.2.33) Determine em função de $a \in \mathbb{R}$ quando é que o sistema
   $$\begin{cases} ax + y = a^2 \\ x + ay = 1 \end{cases}$$
   tem solução e quando é que esta é única. Resolução.
5. Determine a natureza do sistema
   $$\begin{cases} \alpha x + \beta z = 2 \\ \alpha x + \alpha y + 4z = 4 \\ \alpha y + 2z = \beta \end{cases}$$
   em função dos parâmetros $\alpha, \beta \in \mathbb{R}$. Resolução.
6. Determine um sistema que tenha como solução o conjunto
   $$\{(2t + 3s, t + s - 1, 2s + 1, t - 1) \colon t, s, \in \mathbb{R}\}$$
   Resolução.
7. Aplique o método de Gauss-Jordan para colocar as seguintes matrizes em escada de linhas reduzida

(a) $\begin{bmatrix} 2 & 3 \\ 1 & -2 \\ 1 & 5 \end{bmatrix}$ (b) $\begin{bmatrix} i & -(1+i) & 0 \\ 1 & -2 & 1 \\ 1 & 2i & -1 \end{bmatrix}$ (c) $\begin{bmatrix} 1 & 0 & 3 & 1 & 2 \\ 1 & 4 & 2 & 1 & 5 \\ 3 & 4 & 8 & 1 & 2 \end{bmatrix}$

Resolução.

8. Determine os valores de a, b, c para os quais o seguinte sistema tem solução e, nesse caso, determine as soluções.

$$\begin{cases} x + y + z - w = a \\ x - 2z + w = b \\ x - y - 5z + 3w = c \end{cases}$$

Resolução.

9. Considere a matriz

$$A = \begin{bmatrix} 0 & 1 & 1 \\ 2 & 0 & 1 \\ -1 & 0 & 3 \end{bmatrix}$$

(a) Obtenha a partir de A duas matrizes em escada de linhas distintas por aplicação do método de Gauss.
(b) Aplique o método de Gauss-Jordan a cada uma das matrizes da alínea anterior e verifique que obtém a mesma matriz em escada de linhas reduzida.

Resolução.

10. Determine a característica das seguintes matrizes.

(a) $\begin{bmatrix} 0 & 1 & 1 \\ 2 & 3 & -3 \\ -1 & 0 & 3 \\ 0 & 4 & 4 \\ 1 & 0 & 1 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix}$

Resolução.

11. Determine em função do(s) parâmetro(s) a característica das seguintes matrizes:

(a) $\begin{bmatrix} \alpha & 1 & 1 \\ 1 & \alpha & 1 \\ 1 & 1 & \alpha \end{bmatrix}$ (b) $\begin{bmatrix} 1 & i & \alpha & 0 \\ i & 0 & 1 & -i \\ 1+i & i & \beta & \beta \end{bmatrix}$

Resolução.

12. Mostre que um sistema linear homogéneo cujo número de incógnitas é maior do que número de equações tem sempre infinitas soluções distintas. O que pode acontecer se o sistema não for homogéneo? Resolução.

13. Considere o sistema linear S

$$\begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n = b_1 \\ \quad \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n = b_m \end{cases}$$

e o sistema homogéneo associado S_h (que tem os mesmos coeficientes a_{ij} e $b_i = 0$ para todo o i).

(a) Mostre que se $(x_1, \ldots, x_n)$ e $(x'_1, \ldots, x'_n)$ são soluções de S então $(x_1 - x'_1, \ldots, x_n - x'_n)$ é uma solução de S_h.

(b) Mostre que se $(x_1, \ldots, x_n)$ é uma solução de S e $(x'_1, \ldots, x'_n)$ é uma solução de S_h então $(x_1 + x'_1, \ldots, x_n + x'_n)$ é uma solução de S.

(c) Suponha que $(1, 2, -1)$ e $(1, 3.4)$ são soluções de um sistema linear. Determine outra solução do sistema linear.

Resolução.

(Os dois exercícios seguintes são para alunos mais interessados em aprofundar os conhecimentos matemáticos - não fazem parte do programa.)

14. Seja $\mathbb{F}_3 = \{\underline{0}, \underline{1}, \underline{2}\}$ o corpo[8] com 3 elementos que se obtém definindo o produto e a soma tomando o resto da divisão por 3.

(a) Escreva a tabela da soma e multiplicação e verifique que $\mathbb{F}_3$ é um corpo.

(b) Ache as soluções $(x, y, z) \in \mathbb{F}_3^3$ do sistema

$$\begin{cases} x + \underline{2}z = \underline{2} \\ x + y + z = \underline{0} \\ \underline{2}x + y + z = \underline{1} \end{cases}$$

(c) Determine a característica da seguinte matriz com entradas em $\mathbb{F}_3$: $\begin{bmatrix} \underline{2} & \underline{0} & \underline{1} & \underline{1} \\ \underline{1} & \underline{2} & \underline{1} & \underline{0} \\ \underline{1} & \underline{1} & \underline{0} & \underline{1} \end{bmatrix}$

Resolução.

15. Quantas matrizes $m \times n$ com entradas em $\mathbb{F}_2 = \{\underline{0}, \underline{1}\}$ há com característica 1? Resolução.

2. O produto de matrizes

Vimos acima que qualquer combinação linear (6) das equações de um sistema linear (5) é satisfeita por uma solução do sistema. Mais geralmente, começando com um sistema linear (5), podemos considerar um novo sistema cujas equações são combinações lineares das equações do sistema inicial. No caso homogéneo (ou seja com $b_i = 0$) um tal sistema com k equações tem o aspecto seguinte

(11)
$$\begin{cases} c_{11}(a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n) + \ldots + c_{1m}(a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n) = 0 \\ c_{21}(a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n) + \ldots + c_{2m}(a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n) = 0 \\ \vdots \\ c_{k1}(a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n) + \ldots + c_{km}(a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n) = 0 \end{cases}$$

[8]Ver a definição de corpo na Observação 1.2. Os sublinhados servem para lembrar que as operações de soma e produto com $0, 1, 2$ não são as habituais.

onde $c_{i1},\ldots,c_{im}$ são os coeficientes da combinação linear que produz a i-ésima equação do novo sistema. Estes escalares podem ser dispostos numa matriz $k \times m$.

$$\begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1m} \\ c_{21} & c_{22} & \cdots & c_{2m} \\ \vdots & & & \vdots \\ c_{k1} & c_{k2} & \cdots & c_{km} \end{bmatrix}$$

Identificando o sistema inicial com a matriz $[a_{ij}]_{1\leq i\leq m, 1\leq j\leq n}$ dos seus coeficientes, podemos pensar neste processo de combinação linear de equações como uma operação que partindo de duas matrizes, $C=[c_{pq}]$ de tipo $k\times m$ e $A=[a_{ij}]$ de tipo $m\times n$ produz uma nova matriz que tem por entradas os coeficientes das equações do sistema (11). Esta nova matriz é de tipo $k\times n$ e tem como entrada ij (correspondente ao coeficiente de x_j na i-ésima equação de (11))

$$c_{i1}a_{1j}+c_{i2}a_{2j}+\ldots+c_{im}a_{mj}=\sum_{l=1}^{m} c_{il}a_{lj} \tag{12}$$

Definição 2.1. *Sejam k,m,n números naturais, C uma matriz $k\times m$ e A uma matriz $m\times n$ de números reais (ou complexos). O* produto da matriz C pela matriz A *é a matriz $k\times n$, denotada por CA, cuja entrada ij é dada pela expressão* (12).

Note-se que a expressão (12) não é mais do que o *produto escalar* da linha i da matriz C com a coluna j da matriz A.

$$\begin{bmatrix} \vdots & & & \vdots \\ c_{i1} & c_{i2} & \cdots & c_{im} \\ \vdots & & & \vdots \end{bmatrix} \begin{bmatrix} \ldots & a_{1j} & \ldots \\ \ldots & a_{2j} & \ldots \\ & \vdots & \\ \ldots & a_{mj} & \ldots \end{bmatrix}$$

Exemplo 2.2.

$$\begin{bmatrix} 2 & 0 & 3 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 & 0 \\ -1 & 1 & -1 & 3 \\ 0 & 3 & 0 & 1 \end{bmatrix} =$$

$$= \begin{bmatrix} 2\cdot 1+0\cdot(-1)+3\cdot 0 & 2\cdot 2+0\cdot 1+3\cdot 3 & 2\cdot 0+0\cdot(-1)+3\cdot 0 & 2\cdot 0+0\cdot 3+3\cdot 1 \\ 1\cdot 1+(-1)\cdot(-1)+0\cdot 0 & 1\cdot 2+(-1)\cdot 1+0\cdot 3 & 1\cdot 0+(-1)\cdot(-1)+0\cdot 0 & 1\cdot 0-1\cdot 3+0\cdot 1 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 13 & 0 & 3 \\ 2 & 1 & 1 & -3 \end{bmatrix}$$

A fórmula (12) para o produto de matrizes admite várias interpretações que facilitam muitas vezes o cálculo e que são já patentes no exemplo anterior:

- A i-ésima linha do produto CA é a combinação linear das linhas de A cujos coeficientes são as entradas da i-ésima linha de C (foi esta aliás a maneira como chegámos à fórmula para o produto de matrizes). Concretamente, no exemplo acima, a primeira linha do produto é igual a

$$2\cdot\begin{bmatrix} 1 & 2 & 0 & 0 \end{bmatrix}+0\cdot\begin{bmatrix} -1 & 1 & -1 & 3 \end{bmatrix}+3\cdot\begin{bmatrix} 0 & 3 & 0 & 1 \end{bmatrix}$$

- A j-ésima coluna do produto CA é a combinação linear das colunas de C cujos coeficientes são as entradas da j-ésima coluna de A. No exemplo acima, a primeira coluna do produto é igual a

$$1 \cdot \begin{bmatrix} 2 \\ 1 \end{bmatrix} - 1 \cdot \begin{bmatrix} 0 \\ -1 \end{bmatrix} + 0 \cdot \begin{bmatrix} 3 \\ 0 \end{bmatrix}$$

Em muitos exemplos (como no Exemplo 2.2 acima) o produto calcula-se muito mais rapidamente fazendo as contas por linhas ou colunas do que aplicando a fórmula (12) entrada a entrada.

Usando o produto de matrizes, podemos escrever um sistema (5) usando matrizes para os coeficientes, incógnitas e termos independentes. A expressão (5) é equivalente à igualdade de matrizes

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} \tag{13}$$

que se pode abreviar

$$AX = B$$

Uma vez que entendamos as propriedades do produto de matrizes, poderemos manipular sistemas e resolvê-los de forma análoga à que é já familiar do estudo anterior da resolução de equações numéricas.

Os métodos de Gauss e Gauss-Jordan podem também ser descritos em termos do produto de matrizes. Por exemplo, tendo em conta a descrição do produto de matrizes em termos de combinação linear de linhas, a aplicação da operação $L_2 + 3L_1$ ao sistema (13) consiste na multiplicação à esquerda (em ambos os lados da igualdade) pela matriz de tipo $m \times m$

$$\begin{bmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 3 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$$

De forma semelhante, a operação $-2L_2$ corresponde à multiplicação de (13) pela matriz $m \times m$

$$\begin{bmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 0 & -2 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$$

enquanto que a operação $L_1 \leftrightarrow L_2$ corresponderia à multiplicação de (13) pela matriz $m \times m$

$$\begin{bmatrix} 0 & 1 & \cdots & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$$

2.3. **Multiplicação por blocos.** Uma outra observação sobre o produto de matrizes que é por vezes muito útil é que este pode ser realizado "por blocos". Se decompusermos duas matrizes A, B em "matrizes de matrizes", por exemplo,

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \qquad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$$

onde A_{11} é $m \times k$, A_{12} é $m \times l$, A_{21} é $n \times k$, A_{22} é $n \times l$, B_1 é $k \times p$ e B_2 é $l \times p$ então

$$AB = \begin{bmatrix} A_{11}B_1 + A_{12}B_2 \\ A_{21}B_1 + A_{22}B_2 \end{bmatrix}$$

Em geral, o produto de matrizes pode ser realizado desta forma desde que a decomposição em blocos dos fatores seja escolhida apropriadamente (de forma a que as multiplicações de matrizes façam sentido). O número e posição dos blocos em cada fator é arbitrário, sujeito apenas à condição anterior que as multiplicações dos blocos façam sentido[9]. Vejamos um exemplo concreto. Sejam

$$A = \begin{bmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \\ 6 & 7 & 8 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 1 \\ 2 & 3 \\ 4 & 5 \end{bmatrix}$$

Decompondo A e B como

$$A = \begin{bmatrix} \begin{bmatrix} 0 & 1 \\ 3 & 4 \end{bmatrix} & \begin{bmatrix} 2 \\ 5 \end{bmatrix} \\ \\ \begin{bmatrix} 6 & 7 \end{bmatrix} & \begin{bmatrix} 8 \end{bmatrix} \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} \begin{bmatrix} 0 \\ 2 \end{bmatrix} & \begin{bmatrix} 1 \\ 3 \end{bmatrix} \\ \\ \begin{bmatrix} 4 \end{bmatrix} & \begin{bmatrix} 5 \end{bmatrix} \end{bmatrix}$$

temos

$$AB = \begin{bmatrix} \begin{bmatrix} 10 \\ 28 \end{bmatrix} & \begin{bmatrix} 13 \\ 40 \end{bmatrix} \\ \begin{bmatrix} 46 \end{bmatrix} & \begin{bmatrix} 67 \end{bmatrix} \end{bmatrix}$$

[9] No limite, a multiplicação usual de matrizes pode ser vista como a multiplicação de matrizes que foram decompostas em blocos 1×1. O cálculo do produto de matrizes "por linhas"e "por colunas"que descrevemos antes é também um caso particular da multiplicação de matrizes por blocos: num dos fatores os blocos são as linhas (ou as colunas) e no outro a matriz está decomposta em blocos 1×1.

Por exemplo, a matriz 2×1 correspondente às entradas 11 e 21 do produto AB é dada por

$$\begin{bmatrix} 10 \\ 28 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 0 \\ 2 \end{bmatrix} + \begin{bmatrix} 2 \\ 5 \end{bmatrix} \begin{bmatrix} 4 \end{bmatrix}$$

2.4. **Propriedades do produto de matrizes.**

Definição 2.5. *Seja n um número natural.* A matriz identidade *de tipo $n \times n$ é a matriz I_n que tem como entrada ij*

$$\delta_{ij} = \begin{cases} 1 & se\ i = j \\ 0 & se\ i \neq j \end{cases}$$

ou seja

$$I_n = \begin{bmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$$

Teorema 2.6 (Associatividade e existência de elementos neutros)**.** *Sejam k, m, n, p números naturais e A, B, C matrizes de tipo $k \times m$, $m \times n$ e $n \times p$ respetivamente.*

(i) Propriedade associativa do produto: $A(BC) = (AB)C$.
(ii) Elemento neutro para o produto: $I_k A = A$ e $AI_m = A$.

Dem. (i) Temos a verificar que para cada i, j com $1 \leq i \leq k$ e $1 \leq j \leq p$, a entrada ij das matrizes $A(BC)$ e $(AB)C$ são iguais. Escrevendo $(AB)_{ij}$ para a entrada ij do produto das matrizes A e B e aplicando (duas vezes) a fórmula (12) que define o produto de matrizes obtemos

$$\begin{aligned} (A(BC))_{ij} &= \sum_{x=1}^{m} a_{ix}(BC)_{xj} \\ &= \sum_{x=1}^{m} a_{ix} \left(\sum_{y=1}^{n} b_{xy} c_{yj} \right) \\ &= \sum_{x=1}^{m} \sum_{y=1}^{n} a_{ix} b_{xy} c_{yj} \end{aligned}$$

onde na última igualdade aplicámos as propriedades distributiva da soma em relação ao produto (de números) e também as propriedade associativas da soma e multiplicação (de números). De forma inteiramente análoga temos

$$\begin{aligned} ((AB)C))_{ij} &= \sum_{z=1}^{n} (AB)_{iz} c_{zj} \\ &= \sum_{z=1}^{n} \left(\sum_{w=1}^{m} a_{iw} b_{wz} \right) c_{zj} \\ &= \sum_{z=1}^{n} \sum_{w=1}^{m} a_{iw} b_{wz} c_{zj} \end{aligned}$$

As expressões obtidas para $(A(BC))_{ij}$ e $((AB)C)_{ij}$ são idênticas[10] (pelas propriedades associativa e comutativa da soma de números) o que conclui a demonstração da igualdade $A(BC) = (AB)C$.

(ii) A demonstração é análoga (mas mais fácil). Alternativamente, as afirmações são evidentes se calcularmos o produto $I_k A$ por linhas e o produto AI_m por colunas. Exercício.

□

Na proposição anterior vimos propriedades importantes que a multiplicação de matrizes partilha com a multiplicação de números, (embora seja importante notar que a complexidade da multiplicação de matrizes é superior: há matrizes de vários tipos e só quando o número de colunas do fator da esquerda é igual ao número de linhas do fator da direita se pode efetuar a multiplicação). Há também diferenças importantes:

Exemplo 2.7 (A multiplicação de matrizes não é comutativa)**.** *Note-se que os produtos AB e BA só poderão ser matrizes do mesmo tipo se A e B forem matrizes quadradas com igual número de linhas. Se escolhermos duas destas matrizes ao acaso (com mais de uma linha!), a probabilidade de os produtos serem diferentes é 100%. Por exemplo,*

$$\begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -2 & 1 \end{bmatrix} = \begin{bmatrix} -3 & 3 \\ 5 & 2 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix} = \begin{bmatrix} 4 & 1 \\ 1 & -5 \end{bmatrix}$$

Uma das propriedades da multiplicação de números que é muito útil é a chamada *lei do corte*:

$$\text{Se } a \neq 0 \text{ e } ab = ac \text{ então } b = c.$$

Definição 2.8. *A matriz $m \times n$* nula *é a matriz que tem todas as entradas iguais a* 0*. É denotada por* 0 *(deixando implícitas as dimensões).*

[10] Os índices dos somatórios são *variáveis mudas*. Obtém-se uma expressão da outra substituindo o índice x por w e y por z.

É imediato da definição do produto que (sempre que os produtos façam sentido) temos

$$A \cdot 0 = 0 \qquad 0 \cdot A = 0$$

Exemplo 2.9 (A lei do corte não é válida para o produto de matrizes). *Seja A a matriz* $\begin{bmatrix} 2 & -1 \\ 4 & -2 \end{bmatrix}$. *Então*

$$A^2 \stackrel{def}{=} AA = \begin{bmatrix} 2 & -1 \\ 4 & -2 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 4 & -2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

portanto, apesar de $A \neq 0$ temos

$$AA = A \cdot 0.$$

Definição 2.10. *Uma matriz $n \times n$, A diz-se* invertível *se existe uma matriz B (necessariamente também $n \times n$) tal que*

$$AB = BA = I_n$$

Uma tal matriz B diz-se uma inversa *de A.*

Proposição 2.11. *Seja A uma matriz $n \times n$* invertível, *C, D matrizes $n \times m$ e E, F matrizes $m \times n$. Então*

$$AC = AD \Rightarrow C = D \qquad e \qquad EA = FA \Rightarrow E = F$$

Dem. Provamos apenas a primeira implicação deixando a segunda como exercício. Seja B uma inversa de A. Então

$$AC = AD \Rightarrow B(AC) = B(AD) \Leftrightarrow (BA)C = (BA)D \Leftrightarrow I_nC = I_nD \Leftrightarrow C = D$$

□

Vejamos algumas propriedades das matrizes invertíveis.

Proposição 2.12 (Unicidade da inversa). *Seja A uma matriz $n \times n$. Se B e C são inversas de A então $B = C$.*

Dem. Temos

$$B = BI_n = B(AC) = (BA)C = I_nC = C$$

□

A partir de agora escrevemos

$$A^{-1} \quad \text{para } a \text{ inversa da matriz } A.$$

Notemos as seguintes consequências da unicidade da inversa.

Proposição 2.13. *Sejam A, B matrizes $n \times n$ invertíveis. Então*

(i) AB é invertível e $(AB)^{-1} = B^{-1}A^{-1}$
(ii) A^{-1} é invertível e $(A^{-1})^{-1} = A$.

Dem. Mostramos apenas a primeira afirmação deixando a segunda como exercício. Uma vez que a inversa é única, tudo o que é necessário fazer é verificar que as relações na Definição 2.10 são satisfeitas:

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}I_nB = B^{-1}B = I_n$$

e, analogamente,

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AI_nA^{-1} = AA^{-1} = I_n$$

□

2.14. **Soma e produto por escalar.** Vamos também necessitar de outras operações com matrizes que têm uma natureza muito mais elementar do que o produto.

Definição 2.15. *Sejam A, B matrizes $m \times n$. A* soma *das matrizes A e B é a matriz do mesmo tipo $A + B$ que tem como entrada ij*

$$(A + B)_{ij} = a_{ij} + b_{ij}$$

O produto de uma matriz A $m \times n$ pelo escalar $\lambda \in \mathbb{R}$ (ou $\mathbb{C}$) *é a matriz λA também de tipo $m \times n$ cuja entrada ij é*

$$(\lambda A)_{ij} = \lambda a_{ij}$$

Por exemplo

$$\begin{bmatrix} 2 & -1 & 2 \\ 0 & -3 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 4 & 2 \\ 2 & 3 & -1 \end{bmatrix} = \begin{bmatrix} 2+1 & -1+4 & 2+2 \\ 0+2 & -3+3 & 0-1 \end{bmatrix} = \begin{bmatrix} 3 & 3 & 4 \\ 2 & 0 & -1 \end{bmatrix}$$

e

$$\sqrt{2}\begin{bmatrix} 1 & 1 \\ -1 & 2 \\ 4 & 0 \end{bmatrix} = \begin{bmatrix} \sqrt{2} & \sqrt{2} \\ -\sqrt{2} & 2\sqrt{2} \\ 4\sqrt{2} & 0 \end{bmatrix}$$

Vejamos algumas propriedades fundamentais destas operações cujas demonstrações são imediatas e ficam como exercício.

Proposição 2.16 (Propriedades da soma de matrizes)**.** *Sejam A, B, C matrizes $m \times n$. Então*

(i) (Associatividade) $A + (B + C) = (A + B) + C$
(ii) (Comutatividade) $A + B = B + A$
(iii) (Existência de elemento neutro para a soma) $A + 0 = A$
(iv) (Existência de simétricos[11]) Existe D tal que $A + D = 0$

É fácil verificar (exercício) que o simétrico de uma matriz é único. Usa-se a notação $-A$ para o simétrico de uma matriz e claramente a componente ij da matriz $-A$ é dada por $-a_{ij}$.

Proposição 2.17 (Propriedades do produto por escalar)**.** *Sejam A, B matrizes $m \times n$ e λ, μ escalares reais (ou complexos). Então*

[11]Isto é, de inversos para a operação soma.

(i) $1 \cdot A = A$
(ii) $\lambda(\mu A) = (\lambda\mu)A$
(iii) $\lambda(A + B) = \lambda A + \lambda B$
(iv) $(\lambda + \mu)A = \lambda A + \mu A$

Outras propriedades do produto por escalar que são muitas vezes utilizadas são as seguintes

$$0 \cdot A = 0, \qquad (-1) \cdot A = -A$$

Estas propriedades são de verificação imediata a partir da definição do produto por escalar mas podem também ser deduzidas das propriedades indicadas nas Proposições acima (sem usar a definição). Fica como exercício a realização dessas deduções.

Vejamos agora algumas relações entre a soma e o produto por escalar com o produto de matrizes.

Proposição 2.18 (Distributividade)**.** *Sejam A uma matrix $m \times n$, B e C matrizes $n \times p$ e D uma matriz $p \times q$. Então*

$$A(B + C) = AB + AC \qquad (B + C)D = BD + CD$$

Dem. Verificamos apenas a primeira igualdade dado que a demonstração da segunda é inteiramente análoga. Temos a ver que para cada i, j com $1 \leq i \leq m$ e $1 \leq j \leq p$, as entradas ij das matrizes $A(B + C)$ e $AB + AC$ são iguais. De acordo com (12) a entrada ij de $A(B + C)$ é dada pela expressão

$$\begin{aligned} \sum_{k=1}^{n} a_{ik}(B + C)_{kj} &= \sum_{k=1}^{n} a_{ik}(b_{kj} + c_{kj}) \\ &= \sum_{k=1}^{n} a_{ik}b_{kj} + a_{ik}c_{kj} \\ &= (AB)_{ij} + (AC)_{ij} \end{aligned}$$

o que mostra a igualdade pretendida. □

Podemos usar as propriedades acima para desenvolver e simplificar expressões como estamos habituados a fazer com os números mas devido às diferenças indicadas acima, isto requer algum cuidado. Por exemplo, se A e B são matrizes $n \times n$ temos

$$(A + B)^2 = (A + B)(A + B) = A(A + B) + B(A + B) = A^2 + AB + BA + B^2$$

Esta expressão é (pela lei do corte para a soma de matrizes) igual à expressão habitual

$$A^2 + 2AB + B^2$$

se e só se for satisfeita a seguinte igualdade pelas matrizes A, B

$$AB = BA$$

o que, como já indicámos acima, quase nunca se verifica.

Definição 2.19. *Sejam A, B matrizes $n \times n$. Diz-se que A e B* comutam *se $AB = BA$.*

É imediato verificar que a matriz λI_n comuta com qualquer outra matriz $n \times n$, uma vez que, pela interpretação do produto de matrizes em termos de combinações lineares de linhas e colunas, multiplicar A à esquerda por λI_n consiste em multiplicar cada linha de A por λ, enquanto que multiplicar por λI_n à direita consiste em multiplicar por λ cada coluna de A. Portanto

$$(\lambda I_n)A = \lambda A = A(\lambda I_n)$$

O Exercício 13 deste capítulo pede-vos para verificar que estas matrizes - os múltiplos escalares da matriz identidade - são na realidade as únicas matrizes que têm esta propriedade de comutar com todas as outras. A igualdade acima é um caso particular da seguinte propriedade que relaciona o produto de matrizes com o produto por escalar. A demonstração (muito fácil) é deixada como exercício.

Proposição 2.20. *Sejam A uma matriz $m \times n$, B uma matriz $n \times p$ e λ um escalar real (ou complexo). Então*

$$\lambda(AB) = A(\lambda B) = (\lambda A)B$$

Exemplo 2.21. *Seja A uma matriz $n \times n$. Então (uma vez que $3I_n$ comuta com A)*

$$(A + 3I_n)^2 = A^2 + 2(3I_n)A + (3I_n)^2 = A^2 + 6A + 9I_n$$

2.22. **Cálculo da inversa.** Já vimos que a invertibilidade de uma matriz é uma propriedade útil, permitindo-nos por exemplo a aplicação da lei do corte. Põe-se agora a questão de como saber se uma matriz é invertível e, nesse caso, calcular a matriz inversa. Na realidade já aprendemos a calcular a inversa! Se B é a inversa de A então

$$AB = I_n$$

Tendo em conta a interpretação do produto AB como um cálculo de combinações lineares de colunas de A, isto diz-nos que as entradas da i-ésima coluna de B são coeficientes de uma combinação linear das colunas de A que produz a i-ésima coluna da matriz identidade. Se denotarmos a i-ésima coluna de B por X_i, isto diz-nos que a seguinte relação é satisfeita

$$AX_i = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{14}$$

(onde a entrada não nula da matriz à direita está na i-ésima linha). Assim *podemos calcular a i-ésima coluna da inversa resolvendo o sistema linear* (14) para o que podemos usar os métodos de Gauss e Gauss-Jordan. Para calcular a inversa temos que resolver n sistemas lineares mas não há qualquer razão para o fazer separadamente. Como os coeficientes do sistema são os mesmos para todos os sistemas, podemos resolver todos ao mesmo tempo:

Exemplo 2.23. *Vamos calcular A^{-1} para a matriz* $A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 3 & 0 \\ 4 & 0 & 5 \end{bmatrix}$

Aplicamos o método de Gauss-Jordan aos sistemas com termos independentes $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ *e* $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ *simultaneamente:*

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 3 & 0 & 0 & 1 & 0 \\ 4 & 0 & 5 & 0 & 0 & 1 \end{array}\right] \xrightarrow{L_3-4L_1} \left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 3 & 0 & 0 & 1 & 0 \\ 0 & 0 & -3 & -4 & 0 & 1 \end{array}\right] \underset{-\frac{1}{3}L_3}{\xrightarrow{\frac{1}{3}L_2}} \left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 1 & \frac{4}{3} & 0 & -\frac{1}{3} \end{array}\right]$$

$$\xrightarrow{L_1-2L_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{5}{3} & 0 & \frac{2}{3} \\ 0 & 1 & 0 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 1 & \frac{4}{3} & 0 & -\frac{1}{3} \end{array}\right]$$

As colunas da matriz à direita são as soluções de cada um dos sistemas e portanto as colunas da matriz inversa. Assim, se a matriz A for invertível então teremos necessariamente

$$A^{-1} = \begin{bmatrix} -\frac{5}{3} & 0 & \frac{2}{3} \\ 0 & \frac{1}{3} & 0 \\ \frac{4}{3} & 0 & -\frac{1}{3} \end{bmatrix}$$

Exemplo 2.24. *Vamos calcular A^{-1} para a matriz* $A = \begin{bmatrix} 1 & 3 & 1 \\ 0 & -1 & 0 \\ 2 & 0 & 1 \end{bmatrix}$

Temos

$$\left[\begin{array}{ccc|ccc} 1 & 3 & 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 \\ 2 & 0 & 1 & 0 & 0 & 1 \end{array}\right] \xrightarrow{L_3-2L_1} \left[\begin{array}{ccc|ccc} 1 & 3 & 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 \\ 0 & -6 & -1 & -2 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{L_3-6L_2} \left[\begin{array}{ccc|ccc} 1 & 3 & 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & -2 & -6 & 1 \end{array}\right] \underset{-L_3}{\xrightarrow{-L_2}} \left[\begin{array}{ccc|ccc} 1 & 3 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 2 & 6 & -1 \end{array}\right]$$

$$\xrightarrow{L_1-L_3} \left[\begin{array}{ccc|ccc} 1 & 3 & 0 & -1 & -6 & 1 \\ 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 2 & 6 & -1 \end{array}\right] \xrightarrow{L_1-3L_2} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & -3 & 1 \\ 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 2 & 6 & -1 \end{array}\right]$$

Assim, se a matriz A for invertível então teremos necessariamente

$$A^{-1} = \begin{bmatrix} -1 & -3 & 1 \\ 0 & -1 & 0 \\ 2 & 6 & -1 \end{bmatrix}$$

Resta perceber porque é que a matriz B calculada nos exemplos anteriores é de facto uma inversa de A. A maneira como foi determinada torna claro que $AB = I_n$, mas para que B seja a inversa é ainda necessário que $BA = I_n$. Isto está longe de ser óbvio (embora seja fácil de verificar nos exemplos acima ou em qualquer exemplo concreto).

Antes de explicar a razão pela qual o método anterior pode ser sempre usado para achar a inversa (ou ver que uma matriz não é invertível) vamos primeiro responder à seguinte pergunta natural: Porque não achar a inversa por linhas resolvendo o sistema determinado pela equação $BA = I_n$ linha a linha? De facto podemos fazê-lo, mas a matriz dos coeficientes do sistema não será A, e dado que o método de Gauss-Jordan (tal como nós o apresentámos) se aplica imediatamente apenas à solução de sistemas $Ax = b$ com x e b matrizes coluna, é mais prático fazer as contas como fizemos acima.

Esta questão aponta no entanto para um aspeto básico do cálculo matricial que diz respeito à simetria entre linhas e colunas. A atribuição do primeiro índice às linhas e do segundo às colunas é claramente apenas uma convenção pelo que é natural considerar a seguinte *simetria* do conjunto das matrizes que troca linhas com colunas.

Definição 2.25. *Seja A uma matriz $m \times n$. A* matriz transposta *de A é a matriz A^T, de tipo $n \times m$ cuja entrada ij é*

$$(A^T)_{ij} = a_{ji}$$

Por exemplo

$$\begin{bmatrix} 1 & -1 & 2 \\ 0 & 3 & 2 \end{bmatrix}^T = \begin{bmatrix} 1 & 0 \\ -1 & 3 \\ 2 & 2 \end{bmatrix} \qquad \text{e} \qquad \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}^T = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$$

Proposição 2.26 (Propriedades da transposição). *(i)* $(A^T)^T = A$
(ii) $(\alpha A)^T = \alpha A^T$
(iii) $(A + B)^T = A^T + B^T$
(iv) $(AB)^T = B^T A^T$.

Dem. As primeiras três propriedades são muito fáceis de demonstrar e ficam como exercício. Quanto à última, suponhamos que A é uma matriz $m \times n$ e B é uma matriz $n \times p$, de forma a que $(AB)^T$ é uma matriz $p \times m$. Dados i, j com $1 \leq i \leq p$ e $1 \leq j \leq m$ temos então que a entrada ij da matriz $(AB)^T$ é

$$\left((AB)^T\right)_{ij} = (AB)_{ji} = \sum_{k=1}^{n} a_{jk} b_{ki} = \sum_{k=1}^{n} (A^T)_{kj} (B^T)_{ik} = \sum_{k=1}^{n} (B^T)_{ik} (A^T)_{kj} = (B^T A^T)_{ij}$$

conforme queríamos demonstrar. □

Usando esta simetria e a propriedade (iv) acima, é imediato verificar que a solução do sistema para uma linha da matriz inversa mencionado anteriormente não é mais do que a solução do sistema

$$A^T x = b$$

com b a coluna correspondente da matriz identidade. Isto sugere uma relação entre a transposição e a inversão... Qual?

Justifiquemos então finalmente o nosso método de cálculo de inversas:

Teorema 2.27. *Seja A uma matriz $n \times n$ de números reais ou complexos.*[12] *As seguintes afirmações são equivalentes:*

(i) A é invertível,
(ii) Para cada matriz $n \times 1$, B, o sistema $AX = B$ tem solução e esta é única.
(iii) Para cada matriz $n \times 1$, B, o sistema $AX = B$ tem no máximo uma solução.
(iv) O sistema $AX = 0$ tem no máximo uma solução.
(v) A tem característica n

Dem. Vamos ver que (i)⇒(ii)⇒(iii)⇒(iv)⇒(v)⇒(i).

(i)⇒(ii): Multiplicando o sistema dos dois lados por A^{-1} temos

$$A^{-1}AX = A^{-1}B \Rightarrow I_n X = A^{-1}B \Rightarrow X = A^{-1}B$$

Logo se a solução existe, ela é única e é dada por $X = A^{-1}B$. Mas é fácil verificar que $A^{-1}B$ é de facto uma solução:

$$A(A^{-1}B) = I_n B = B$$

o que conclui a prova desta implicação.

(ii)⇒(iii) Evidente.

(iii)⇒(iv) Evidente.

(iv)⇒(v): Esta implicação é equivalente à implicação $\neg$(v) $\Rightarrow$ $\neg$(iv) que passamos a demonstrar. Se a característica de A não é igual a n, então no final do método de Gauss-Jordan, alguma das colunas não tem pivot. A variável correspondente é então livre na solução do sistema homogéneo $AX = 0$, que tem portanto infinitas soluções.

(v)⇒(i): Se A tem característica n, a aplicação dos métodos de Gauss e Gauss-Jordan transformam a matriz A na matriz I_n (uma vez que esta é a única matriz $n \times n$ em escada de linhas reduzida com característica n). Mas, como já observámos, cada passo do método de Gauss-Jordan consiste na multiplicação à esquerda por uma matriz. Nomeadamente:

- A operação $L_i \leftrightarrow L_j$, com $i \neq j$ corresponde à multiplicação à esquerda pela matriz

$$S_{ij} = \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & & 1 & & \\ & & & \ddots & & & \\ & & 1 & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix}$$

em que os $\ddots$ indicam 1s, todas as entradas não indicadas são 0 e os $0s$ na diagonal ocorrem nas linhas i e j.

[12] Ou mais geralmente com entradas num corpo.

- A operação αL_i com $\alpha \neq 0$ corresponde à multiplicação pela matriz
$$D_{i,\alpha} = \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & \alpha & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix}$$
com todas as entradas fora da diagonal 0 e todas as entradas na diagonal 1 exceto a i-ésima que é α.
- A operação $L_i + \alpha L_j$ com $i \neq j$ e $\alpha \neq 0$ corresponde à multiplicação pela matriz
$$I_n + \alpha E_{ij} = \begin{bmatrix} 1 & & & & \\ & \ddots & \alpha & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix}$$
em que todas as entradas da diagonal são 1 e todas as entradas fora da diagonal são 0 exceto a entrada ij, que é igual a α. O esquema acima corresponde ao caso em que $i < j$ e portanto à fase final do método de Gauss-Jordan. A fase inicial do método de Gauss consiste na multiplicação por estas matrizes com $i > j$, caso em que a entrada não nula fora da diagonal está abaixo da diagonal.

Em termos do produto de matrizes, a observação que o método de Gauss-Jordan termina na matriz I_n expressa a igualdade

$$E_k \cdots E_2 E_1 A = I_n \tag{15}$$

em que k é o número de passos do método de Gauss-Jordan e a multiplicação pela matriz E_i executa o passo i do método (cada E_i é uma das matrizes referidas acima). Ora cada matriz E_i é invertível! De facto, é imediato verificar que

- $S_{ij}^{-1} = S_{ij}$
- $D_{i,\alpha}^{-1} = D_{i,\frac{1}{\alpha}}$
- $(I_n + \alpha E_{ij})^{-1} = I_n - \alpha E_{ij}$

Multiplicando à esquerda a igualdade (15) sucessivamente pelas inversas das matrizes $E_k, E_{k-1}, \ldots$ obtemos

$$A = E_1^{-1} E_2^{-1} \cdots E_k^{-1}$$

Uma vez que A é um produto de matrizes invertíveis, pela Proposição 2.13, A é invertível.

□

Vemos assim que, quando aplicamos o método de Gauss-Jordan para resolver simultaneamente os n sistemas lineares correspondentes à equação $AB = I_n$, só há duas possibilidades: ou a aplicação do método mostra que a característica de A é menor do que n e então A

não é invertível ou, a característica de A é n e então a matriz A é invertível. Neste último caso, uma vez que a matriz B calculada pelo método de Gauss-Jordan satisfaz $AB = I_n$, temos

$$A^{-1}(AB) = A^{-1}I_n \Leftrightarrow B = A^{-1}.$$

Observação 2.28. *As matrizes que aparecem na demonstração da implicação ((v)⇒ (i)) no Teorema 2.27 (que correspondem às operações elementares sobre as linhas) designam-se por* matrizes elementares. *Vimos durante a demonstração que qualquer matriz com característica n, ou seja, qualquer matriz invertível, se pode escrever como um produto de matrizes elementares*[13].

Para terminar esta discussão sobre as matrizes observemos ainda a seguinte condição equivalente à invertibilidade.

Corolário 2.29. *Seja A uma matriz quadrada. As seguintes condições são equivalentes.*

(i) A é invertível.

(ii) Para cada matriz $n \times 1$, B, o sistema $AX = B$ tem solução.

Demonstração. Claramente (i) ⇒ (ii) (pela equivalência entre (i) e (ii) no Teorema 2.27), Para demonstrar a implicação recíproca vamos ver que ¬(i)⇒ ¬(ii). Suponhamos então que A não é invertível. Pelo Teorema 2.27 a característica de A é menor do que n. Seja S uma matriz invertível que se obtém multiplicando sucessivamente as matrizes elementares correspondentes aos passos do método de Gauss, de forma que SA está em escada de linhas. Como a característica de A é menor do que n, a última linha de SA é nula. Escrevendo

$$C = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

vemos que o sistema

$$(SA)X = C \tag{16}$$

não tem solução. Ora (16) é equivalente ao sistema $AX = S^{-1}C$, logo o sistema $AX = B$ não tem solução quando $B = S^{-1}C$. Isto conclui a demonstração. ☐

Sendo A uma matriz quadrada, podemos considerar a função $X \mapsto AX$ que leva matrizes coluna em matrizes coluna. O Teorema 2.27 e o Corolário 2.29 mostram que, para uma tal função, as condições de bijetividade, injetividade e sobrejetividade são equivalentes!

2.30. **Exercícios.**

1. Calcule os seguintes produtos matriciais

(a) $\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 2 & 3 & -1 \\ 2 & 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ -1 & 0 & 0 \\ 0 & -2 & 1 \\ 3 & 2 & -1 \end{bmatrix}$

[13] É um bom exercício estimar o número máximo de fatores necessário para uma tal fatorização.

(c) $\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix}$ e $\begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$ (d) $\begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix}^{50}$ (e) $\begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix}^{n}$

Resolução.

2. Considere as matrizes $A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 0 & 1 \\ -1 & -2 & 3 \end{bmatrix}$ e $B = \begin{bmatrix} 0 & 1 & 1 \\ -1 & 1 & 1 \\ -1 & 0 & 3 \end{bmatrix}$. Calcule
   (a) A entrada 12 de A^2.
   (b) A segunda linha de AB.
   (c) A terceira coluna de BA.
   Resolução.
3. Considere as matrizes $A = \begin{bmatrix} 0 & 1 & 1 \\ 2 & 0 & 1 \\ -1 & 0 & 3 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & 1 \end{bmatrix}$, $C = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}$ e $D = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$. Efetue se possível as seguintes operações com matrizes
   (a) $A^2 - 2A + I_3$ (b) $AD - D$ (c) $C^3 + D + 2A$
   (d) $3A + CD$ (e) $BA - 3B$ (f) $DC - \sqrt{2}I_1$
   Resolução.
4. Determine (se existirem) as inversas das seguintes matrizes usando o método de Gauss-Jordan
   (a) $\begin{bmatrix} 2 & 1 \\ -1 & 4 \end{bmatrix}$ (b) $\begin{bmatrix} 6 & -4 \\ -3 & 2 \end{bmatrix}$ (c) $\begin{bmatrix} 2 & 6 & 6 \\ 2 & 7 & 6 \\ 2 & 7 & 7 \end{bmatrix}$
   (d) $\begin{bmatrix} 1 & 1 & 2 & -1 \\ 0 & -1 & 3 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ (e) $\begin{bmatrix} -2 & 7 & 2 & \frac{1}{3} \\ 1 & 0 & \frac{2}{5} & -1 \\ 0 & 0 & 0 & 0 \\ -1 & 13 & 4 & 2 \end{bmatrix}$
   Resolução.
5. Determine todas as soluções das seguintes equações matriciais
   (a) $(A - 2I_2)^{-1} = \begin{bmatrix} -3 & 6 \\ 4 & 5 \end{bmatrix}$ (b) $(A - 3I_3)^{-1}\begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 3 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & -3 & 1 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
   (c) $A^2 = I_2$
   Resolução.
6. Sejam a, b, c, d números reais. Verifique que se $ad - bc \neq 0$, então
$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$
   Resolução.
7. Seja A a matriz $n \times n$ cuja entrada ij é $a_{ij} = 2^{i+j}$. Dê uma expressão para a entrada ij da matriz AB quando B é a matriz $n \times n$ com entradas
   (a) $b_{ij} = 2^{-i-j}$ (b) $b_{ij} = 1$
   Resolução.

8. Determine o conjunto das matrizes que comutam com $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Resolução.
9. Mostre que se A e B comutam então A^{-1} e B^{-1} também comutam. Resolução.
10. Uma matriz quadrada A diz-se *simétrica* se $A = A^T$ e *anti-simétrica* se $A + A^T = 0$. Mostre que toda a matriz quadrada se pode escrever de forma única como a soma de uma matriz simétrica e de uma anti-simétrica. Resolução.
11. Uma matriz quadrada A diz-se *triangular superior* se $a_{ij} = 0$ para $i > j$ (isto é, se só tem zeros abaixo da diagonal principal). Diz-se *triangular inferior* se A^T é triangular superior.
    (a) Mostre que o produto de matrizes triangulares superiores (resp. inferiores) é uma matriz triangular superior (resp. inferior).
    (b) Quando é que uma matriz triangular superior (ou inferior) é invertível? A inversa é triangular superior (inferior)?
    (c) Suponha que A é invertível e que durante a aplicação do método de Gauss a A para achar a inversa nunca é necessário trocar linhas. Mostre que então A se pode escrever de forma única como o produto de uma matriz triangular inferior L com todas as entradas na diagonal iguais a 1 e de uma matriz triangular superior U (na forma $A = LU$).

    Resolução.
12. Seja $A = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix}$ uma matriz descrita por blocos com B e D matrizes quadradas. Mostre que A é invertível sse B e D são invertíveis e, nesse caso, determine uma fórmula para a inversa de A em termos das matrizes B, C, D. Resolução.
13. Seja A uma matriz $n \times n$. Mostre que se A comuta com todas as matrizes $n \times n$ então existe um escalar λ tal que $A = \lambda I_n$. *Sugestão: Considere as matrizes E_{ij} que têm* 1 *como entrada ij e* 0 *em todas as outras entradas.* Resolução.
14. Se A é uma matriz $n \times n$ invertível, define-se para k um inteiro não positivo

$$A^k = \begin{cases} I_n & \text{se } k = 0 \\ (A^{-1})^{-k} & \text{se } k < 0 \end{cases}$$

    Mostre que é válida a regra dos expoentes: Dados k, l inteiros $A^k A^l = A^{k+l}$ Resolução.
15. Sejam m, n números naturais e A uma matriz $m \times n$. Uma matriz $n \times m$ B diz-se uma *inversa à direita* para A se $AB = I_m$. B diz-se uma *inversa à esquerda* para A se $BA = I_n$.
    (a) Mostre que se A tem inversa à direita e à esquerda entâo as inversas à direita e esquerda coincidem.
    (b) Dê um exemplo de uma matriz invertível à direita mas não à esquerda e vice-versa.
    (c) Mostre que as seguintes afirmações são equivalentes
        - A é invertível à direita
        - Para toda a matriz $m \times 1$, C, a equação $AX = C$ tem solução
        - A característica de A é m.
    (d) Uma matriz $m \times n$ com $m > n$ pode ser invertível à direita? Justifique.

(e) Mostre que se A tem inversa à direita e à esquerda então $m = n$ (e A é invertível).
(f) Mostre que se A é uma matriz $n \times n$ e tem inversa à direita ou inversa à esquerda, então A é invertível.

Resolução.

3. Espaços vetoriais

Um espaço vetorial é um "sítio onde se podem fazer combinações lineares". Para que tal seja possível, precisamos de saber como somar e como multiplicar por escalar os objetos do espaço vetorial. Para que estas combinações lineares se comportem da forma que estamos habituados nos exemplos que vimos até agora é necessário que satisfaçam certas propriedades que são especificadas na definição de espaço vetorial.

O arquétipo de um espaço vetorial é $\mathbb{R}^n = \{(x_1, \ldots, x_n) \colon x_i \in \mathbb{R}\}$ com a multiplicação por escalar definida pela expressão

$$\alpha \cdot (x_1, \ldots, x_n) = (\alpha x_1, \ldots, \alpha x_n)$$

e a soma por

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$$

Nos casos em que $n = 1, 2$ ou 3, estamos habituados a identificar $\mathbb{R}^n$ geometricamente com o conjunto dos vetores com origem em $(0, \ldots, 0)$, e sabemos interpretar geometricamente o produto por escalar e a soma.

Por exemplo, o conjunto de todas as combinações lineares de dois vetores não colineares em $\mathbb{R}^3$ formam um plano que passa pela origem e contém os dois vetores.

A definição de espaço vetorial vai-nos permitir transferir a nossa intuição geométrica sobre o comportamento de vetores no espaço para um sem-fim de novas situações!

Definição 3.1. *Um* espaço vetorial real *é um terno $(V, +, \cdot)$ constituído por um conjunto V, cujos elementos se designam por* vetores*, juntamente com duas funções:*

- *Multiplicação por escalar: $\mathbb{R} \times V \xrightarrow{\cdot} V$ que a um par (α, v) associa um vetor $\alpha \cdot v$.*
- *Soma de vetores: $V \times V \xrightarrow{+} V$ que a um par de vetores (v, w) associa um vetor $v + w$*

satisfazendo as seguintes relações:

(i) Para todos os $u, v, w \in V$, $u + (v + w) = (u + v) + w$.
(ii) Para todos os $u, v \in V$, $u + v = v + u$.
(iii) Existe um elemento $0 \in V$ tal que, para todo o $v \in V$ se tem $v + 0 = v$.
(iv) Para todo o $v \in V$ existe um elemento $w \in V$ tal que $v + w = 0$.
(v) Para todo o $v \in V$, tem-se $1 \cdot v = v$.
(vi) Para todos os $\alpha, \beta \in \mathbb{R}$, e $v \in V$ tem-se $\alpha \cdot (\beta \cdot v) = (\alpha\beta) \cdot v$.
(vii) Para todos os $\alpha \in \mathbb{R}$ e $v, w \in V$ tem-se $\alpha \cdot (v + w) = \alpha \cdot v + \alpha \cdot w$.
(viii) Para todos os $\alpha, \beta \in \mathbb{R}$ e $v \in V$ tem-se $(\alpha + \beta) \cdot v = \alpha \cdot v + \beta \cdot v$.

Há uma possível ambiguidade no axioma (iv) uma vez que o axioma (iii) não garante a unicidade do elemento neutro 0 para a soma. Mas na realidade é imediato verificar que este elemento (que se diz o *vetor zero*) é único: se $0'$ for um outro elemento neutro temos

$0 + 0' = 0$ (aplicando o axioma (iii) ao elemento neutro $0'$ e ao elemento $v = 0$); por outro lado, $0+0' = 0'+0 = 0'$ (aplicando primeiro o axioma (ii) e depois (iii) ao elemento neutro 0 e ao elemento $v = 0'$).

Também não é difícil mostrar que o elemento w tal que $v + w = 0$ é único: se $v + w = v + w' = 0$ então

$$w' = w' + 0 = w' + (v + w) = (w' + v) + w = (v + w') + w = 0 + w = w + 0 = w$$

O único w tal que $w + v = 0$ chama-se o *simétrico* de v e denota-se por $-v$.

Observação 3.2. *(i) Substituindo na definição acima $\mathbb{R}$ por $\mathbb{C}$ obtemos a definição de um* espaço vetorial complexo. *Mais geralmente se $\mathbb{K}$ é um corpo (ver Observação 1.2) e substituirmos $\mathbb{R}$ por $\mathbb{K}$ obtemos a noção de espaço vetorial sobre o corpo $\mathbb{K}$.*

(ii) É também comum usar a terminologia espaço linear *em vez de* espaço vetorial.

(iii) O produto por escalar $\alpha \cdot v$ denota-se normalmente simplesmente por αv e passaremos a denotá-lo desta forma quando não haja possibilidade de confusão.

Definição 3.3. *Seja V um espaço vetorial e $v_1, \ldots, v_k$ elementos de V. Diz-se que $v \in V$ é uma* combinação linear *dos vetores $v_1, \ldots, v_k$ se existem $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ tais que*

$$v = \alpha_1 v_1 + \ldots + \alpha_k v_k$$

Os escalares $\alpha_1, \ldots, \alpha_k$ chamam-se os coeficientes *da combinação linear.*

Exemplo 3.4. *(1) $\mathbb{R}^n$ com a soma e produto por escalar definidos coordenada a coordenada é um espaço vetorial real. A validade dos axiomas na Definição 3.1 é uma consequência imediata das propriedades das operações de soma e produto de números reais. Por exemplo a propriedade associativa da soma de vetores segue imediatamente da propriedade associativa da soma de números reais. Analogamente $\mathbb{C}^n = \{(z_1, \ldots, z_n) \colon z_i \in \mathbb{C}\}$ é um espaço vetorial complexo, com as operações de soma e produto por escalar definidas componente a componente. Mais geralmente $\mathbb{K}^n$ é um espaço vetorial sobre o corpo $\mathbb{K}$.*

(2) O conjunto $M_{m\times n}(\mathbb{R})$ das matrizes $m \times n$ reais é um espaço vetorial real. É esse o conteúdo das Proposições 2.16 e 2.17. Analogamente, o conjunto das matrizes $M_{m\times n}(\mathbb{K})$ é um espaço vetorial sobre o corpo $\mathbb{K}$.

(3) Seja S um conjunto não vazio. O conjunto $F(S; \mathbb{R}) = \{f \colon S \to \mathbb{R}\}$ das funções de S para $\mathbb{R}$ munido das operações

$$(f + g)(x) \stackrel{def}{=} f(x) + g(x) \qquad (\alpha f)(x) \stackrel{def}{=} \alpha f(x)$$

é um espaço vetorial real. Analogamente o conjunto das funções com valores complexos é um espaço vetorial complexo. Note-se que este exemplo contém os dois exemplos anteriores. De facto $\mathbb{R}^n$ é basicamente o caso em que o conjunto S é $\{1, \ldots, n\}$ e $M_{m\times n}(\mathbb{R})$ é, por definição, o caso em que $S = \{1, \ldots, m\} \times \{1, \ldots, n\}$.

Observação 3.5. *É habitual referirmo-nos a um espaço vetorial apenas pelo conjunto subjacente deixando implícitas a estrutura de soma de vetores e multiplicação por escalares quando estas são claras do contexto. Por exemplo, quando falamos do espaço vetorial*

$M_{m\times n}(\mathbb{R})$ referimo-nos a este conjunto com as operações habituais de soma e multiplicação por escalar.

Exemplo 3.6. *Sejam $v, w \in \mathbb{R}^3$ dois vetores não colineares. Pelo significado geométrico da soma de vetores e produto por escalar, o conjunto das combinações lineares de v e w é o plano que passa pela origem e contém v e w. Dado um ponto u desse plano, o significado dos coeficientes α, β na combinação linear $u = \alpha v + \beta w$ é o seguinte (familiar da noção de coordenadas cartesianas)*

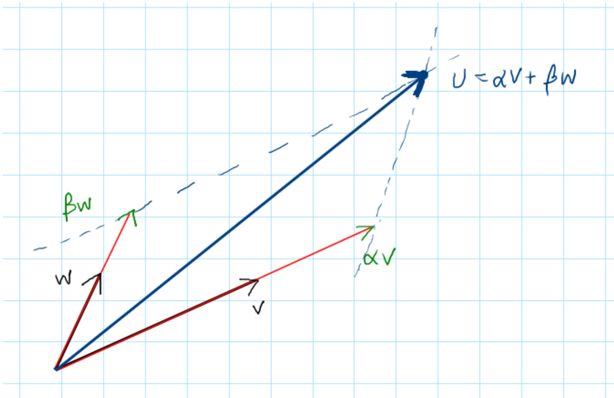


FIGURA 1. Uma combinação linear de dois vetores em $\mathbb{R}^3$

- *αv é o ponto de interseção da reta paralela a w que passa por u, com a reta determinada por v e pela origem (que é o conjunto $\{\lambda v \colon \lambda \in \mathbb{R}\}$).*
- *βw é o ponto de interseção da reta paralela a v que passas por u, com a reta $\{\lambda w \colon \lambda \in \mathbb{R}\}$*

Vejamos mais alguns exemplos e não-exemplos de espaços vetoriais.

Exemplo 3.7.

(i) O conjunto V de todos os polinómios reais com as operações de soma e produto por escalar habituais é um espaço vetorial. Note-se que V está contido no conjunto das funções reais $F(\mathbb{R}, \mathbb{R})$ e que as operações de soma e produto por escalar são a restrição aos polinómios das operações definidas para as funções. Isso torna a verificação da maioria dos axiomas na Definição 3.1 automáticas. De facto, uma vez que se observe que a soma de polinómios e a multiplicação de um escalar por um polinómio são polinómios, a validade das propriedades (i)-(ii) e (v)-(viii) é imediata e resta apenas observar que a função nula é um polinómio logo (iii) é satisfeito e que a função simétrica de um polinómio é um polinómio logo (iv) é também satisfeito.

(ii) Seja $V = \{(x, y) \in \mathbb{R}^2 \colon x \geq 0, y \geq 0\}$ com a soma habitual de vetores em $\mathbb{R}^2$ e com o produto por escalar definido por

$$\alpha \cdot (x, y) \overset{def}{=} (|\alpha| x, |\alpha| y)$$

Com estas operações V não é um espaço vetorial porque os axiomas (iv) e (vii) não são verificados. Por exemplo o vetor $(1,0)$ não tem simétrico e $(0,0) = 0 \cdot (1,0) = (1+(-1)) \cdot (1,0) \neq 1 \cdot (1,0) + (-1) \cdot (1,0) = (2,0)$. Em geral, se α e β têm sinais contrários e $v \neq 0$, a igualdade $(\alpha+\beta) \cdot v = \alpha \cdot v + \beta \cdot v$ não se verifica.

3.8. **Subespaços vetoriais.**

Definição 3.9. *Seja V um espaço vetorial sobre $\mathbb{K}$. Um subconjunto $W \subset V$ diz-se um* subespaço vetorial *de V se W é fechado*[14] *para as operações de V e, munido destas operações, é um espaço vetorial.*

Exemplo 3.10. *O Exemplo 3.7 (i) verifica que o conjunto dos polinómios é um subespaço vetorial de $F(\mathbb{R};\mathbb{R})$.*

Como observámos no Exemplo 3.7(i), quando $W \subset V$ é um subconjunto de um espaço vetorial fechado para a soma e multiplicação por escalar, a verificação de que W é um espaço vetorial pode reduzir-se à verificação que o elemento neutro da soma e os simétricos (em V) de elementos de W pertencem a W. A próxima proposição mostra que mesmo estas verificações não são na realidade necessárias.

Proposição 3.11. *Seja V um espaço vetorial. Se W é um subconjunto não vazio de V fechado para a soma e multiplicação por escalar, então W é um subespaço vetorial de V.*

Demonstração. Como já observámos, a verificação dos axiomas (i)-(ii) e (v)-(viii) é imediata. É um exercício verificar que, para qualquer $v \in V$, o produto por escalar $0v$ é o elemento neutro para a soma (ver Exercício 8(ii)). Como W é não vazio e fechado para o produto por escalar conclui-se que $0 \in W$ e portanto o axioma (iii) é verificado. É também um exercício (ver Exercício 8(v)) verificar que o simétrico de $v \in V$ é o produto por escalar $(-1)v$. Uma vez que W é fechado para o produto por escalar conclui-se que o axioma (iv) é verificado em W. □

Exemplo 3.12. *(i) Seja V o espaço vetorial de todos os polinómios reais. O subconjunto $W \subset V$ formado pelos polinómios de grau menor ou igual a 3 é um subespaço vetorial. De facto, de acordo com a proposição anterior basta observar que $W \neq \emptyset$ (por exemplo o polinómio 0 está em W), que a soma de polinómios de grau ≤ 3 tem grau ≤ 3 e que o produto de um polinómio de grau ≤ 3 por um escalar tem ainda grau ≤ 3.*

(ii) O plano $W = \{(x,y,z) \in \mathbb{R}^3 \colon x+y+z=0\}$ é um subespaço vetorial de $\mathbb{R}^3$. De acordo com a Proposição acima basta notar que se $(x,y,z),(x',y',z') \in W$ e $\alpha \in \mathbb{R}$ então $(x+x')+(y+y')+(z+z') = (x+y+z)+(x'+y'+z') = 0$ e $(\alpha x)+(\alpha y)+(\alpha z) = \alpha(x+y+z) = 0$ logo $(x+x', y+y', z+z') \in W$ e $(\alpha x, \alpha y, \alpha z) \in W$.

(iii) Seja A uma matriz $m \times n$. O núcleo *de A é o conjunto*

$$N(A) = \{x \in \mathbb{R}^n \colon A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = 0\}$$

[14] Isto é, se dados $w_1, w_2 \in W$ e $\alpha \in \mathbb{K}$, temos $w_1 + w_2 \in W$ e $\alpha w_1 \in W$. Por palavras: a soma em V de vetores em W está em W, e a multiplicação em V de um vetor de W por um escalar permanece em W.

Este conjunto é um subespaço vetorial de $\mathbb{R}^n$ (o argumento é exatamente o mesmo que no exemplo anterior, que corresponde ao núcleo da matriz $\begin{bmatrix}1 & 1 & 1\end{bmatrix}$). Note-se que $N(A)$ é exactamente o conjunto das soluções do sistema linear homogéneo que tem A como matriz de coeficientes.

Intuitivamente, devemos pensar nos espaços vetoriais como sendo objetos que se comportam de forma semelhante ao espaço euclidiano usual - $\mathbb{R}^3$ - e nos subespaços vetoriais como sendo subconjuntos com comportamento semelhante ao das retas e planos em $\mathbb{R}^3$ que passam pela origem.

3.13. **Expansão linear.**

Definição 3.14. *Seja V um espaço vetorial e $S \subset V$ um subconjunto. A* expansão linear *de S em V é o conjunto $L(S)$ das combinações lineares de elementos de S, isto é*

$$L(S) = \{\alpha_1 v_1 + \ldots + \alpha_n v_n \colon \alpha_1, \ldots, \alpha_n \in \mathbb{R}, v_1, \ldots, v_n \in S, n \in \mathbb{N}\}$$

Por convenção $L(\emptyset) = \{0\}$.

Exemplo 3.15. *(i) Seja V o espaço vetorial dos polinómios reais. Vamos determinar se $x + 2x^3 \in L(S)$ onde $S = \{1 - x, x + x^2 + x^3, x^2\}$. Por definição, a pergunta é se existem escalares $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$ tais que*

$$x + 2x^3 = \alpha_1(1 - x) + \alpha_2(x + x^2 + x^3) + \alpha_3 x^2$$

Como dois polinómios são iguais precisamente quando têm os mesmos coeficientes, a igualdade anterior é equivalente ao sistema

$$\begin{cases} \alpha_1 = 0 \\ -\alpha_1 + \alpha_2 = 1 \\ \alpha_2 + \alpha_3 = 0 \\ \alpha_2 = 2 \end{cases} \Leftrightarrow \begin{cases} \alpha_1 = 0 \\ \alpha_2 = 1 \\ \alpha_3 = -1 \\ \alpha_2 = 2 \end{cases}$$

Uma vez que o sistema é impossível, conclui-se que $x+2x^3 \notin L(S)$. Neste caso não se justificava a utilização do método de Gauss para a resolução do sistema. Mas note-se que se tivéssemos escrito o sistema acima da forma habitual, a matriz à qual iríamos aplicar o método de Gauss seria

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 2 \end{array}\right]$$

Os coeficientes dos polinómios em S aparecem nas primeiras três colunas. A última coluna contém os coeficientes do polinómio $x + 2x^3$.

(ii) Sendo $S = \{(1, 3, 2), (0, 1, 4), (1, 4, 6)\} \subset \mathbb{R}^3$, vamos determinar equações cartesianas que definam $L(S)$. Os elementos de $L(S)$ são os vetores $(a, b, c) \in \mathbb{R}^3$ para os quais é possível achar $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$ tais que

$$\alpha_1(1, 3, 2) + \alpha_2(0, 1, 4) + \alpha_3(1, 4, 6) = (a, b, c)$$

Ou seja, são os vetores (a,b,c) tais que o seguinte sistema é possível

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & a \\ 3 & 1 & 4 & b \\ 2 & 4 & 6 & c \end{array}\right] \underset{L_3-2L_1}{\overset{L_2-3L_1}{\longrightarrow}} \left[\begin{array}{ccc|c} 1 & 0 & 1 & a \\ 0 & 1 & 1 & b-3a \\ 0 & 4 & 4 & c-2a \end{array}\right] \overset{L_3-4L_2}{\longrightarrow} \left[\begin{array}{ccc|c} 1 & 0 & 1 & a \\ 0 & 1 & 1 & b-3a \\ 0 & 0 & 0 & c-4b+10a \end{array}\right]$$

Conclui-se que $(a,b,c) \in L(S) \Leftrightarrow c-4b+10a = 0$. Geometricamente, $L(S)$ é um plano que passa pela origem. Normalmente, esperaríamos que três vetores em $\mathbb{R}^3$ formassem um referencial e que qualquer outro vetor se pudesse escrever como combinação linear deles mas neste caso $(1,3,2)+(0,1,4) = (1,4,6)$ e portanto podemos escrever qualquer combinação linear dos três vetores de S usando apenas os dois primeiros. A expansão linear destes dois vetores é um plano que tem equação paramétrica

$$(x,y,z) = \alpha_1(1,3,2) + \alpha_2(0,1,4), \qquad \text{com } \alpha_1, \alpha_2 \in \mathbb{R}$$

e, como vimos acima, equação cartesiana

$$10x - 4y + z = 0.$$

Proposição 3.16. *Seja V um espaço vetorial e $S \subset V$ um subconjunto. Então $L(S)$ é o mais pequeno subespaço vetorial de V que contém S. Mais precisamente:*

- *$L(S)$ é um subespaço vetorial de V e $S \subset L(S)$.*
- *Se $W \subset V$ é um subespaço vetorial de V que contém S, então $L(S) \subset W$.*

Dem. Se S é vazio então as condições são claramente verificadas (recorde-se que $L(\emptyset) = \{0\}$). Suponhamos que S é não vazio. $L(S)$ contém S porque dado $v \in S$ temos que $1 \cdot v = v$ é uma combinação linear de elementos de S e portanto pertence a $L(S)$. Como S é não vazio, conclui-se que $L(S) \neq \emptyset$. Para ver que $L(S)$ é um subespaço vetorial precisamos agora de ver que $L(S)$ é fechado para a soma e para o produto por escalar. Seja $\lambda \in \mathbb{R}$ um escalar e $\alpha_1 v_1 + \ldots + \alpha_n v_n$ um elemento de $L(S)$. Então

$$\lambda(\alpha_1 v_1 + \ldots + \alpha_n v_n) = (\lambda\alpha_1)v_1 + \ldots + (\lambda\alpha_n)v_n$$

é também uma combinação linear de elementos de S e portanto pertence a $L(S)$. Conclui-se que $L(S)$ é fechado para o produto por escalar. Por outro lado, dados dois elementos $\alpha_1 v_1 + \ldots + \alpha_n v_n$ e $\beta_1 w_1 + \ldots + \beta_m w_m$ em $L(S)$ a sua soma é

$$\alpha_1 v_1 + \ldots + \alpha_n v_n + \beta_1 w_1 + \ldots + \beta_m w_m$$

que é ainda uma combinação linear de elementos de S. Conclui-se que $L(S)$ também é fechado para a soma de vetores e portanto é um subespaço vetorial de V.

Finalmente, seja W um qualquer subespaço vetorial de V que contém S. Então dados $v_1, \ldots, v_n \in S$ e $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ temos que $\alpha_i v_i \in W$ (pois W é fechado para o produto por escalar) e portanto

$$\alpha_1 v_1 + \ldots + \alpha_n v_n \in W$$

(porque W é fechado para a soma). Conclui-se que W contém qualquer combinação linear de elementos de S, ou seja, que W contém $L(S)$. $\square$

Devido ao resultado enunciado na Proposição anterior, chamamos a $L(S)$ o *subespaço gerado por* S e se $W = L(S)$ dizemos que W *é gerado por* S e que S é um *conjunto de geradores* para W.

Note-se que dado um subespaço vetorial W de V, podemos sempre encontrar um conjunto $S \subset W$ tal que $W = L(S)$: de facto, podemos sempre tomar $S = W$. Esta solução não é na prática muito útil pois normalmente estaremos interessados em encontrar um conjunto de geradores tão pequeno quanto possível.

Exemplo 3.17. *(i) Vamos achar um conjunto de geradores para o subespaço*

$$W = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} : a + b - 2c = 0, d - c + a = 0 \right\} \subset M_{2\times 2}(\mathbb{R})$$

(é imediato verificar que W é de facto um subespaço vetorial de $M_{2\times 2}(\mathbb{R})$).

Podemos resolver o sistema dado pelas condições que definem W (aqui não se justifica a aplicação do método de Gauss)

$$\begin{cases} a + b - 2c = 0 \\ d - c + a = 0 \end{cases} \Leftrightarrow \begin{cases} c = \frac{1}{2}a + \frac{1}{2}b \\ d = -\frac{1}{2}a + \frac{1}{2}b \end{cases}$$

O elemento típico de W pode portanto escrever-se na forma

$$\begin{bmatrix} a & b \\ \frac{1}{2}a + \frac{1}{2}b & -\frac{1}{2}a + \frac{1}{2}b \end{bmatrix} = a \begin{bmatrix} 1 & 0 \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix} + b \begin{bmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \qquad \text{com } a, b \in \mathbb{R}$$

logo

$$S = \left\{ \begin{bmatrix} 1 & 0 \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \right\}$$

é um conjunto de geradores para W.

3.18. **Subespaços de $\mathbb{R}^n$ associados a uma matriz.** Seja A uma matriz $m \times n$. Chama-se *espaço das linhas de* A, e denota-se por $EL(A)$ ao subespaço de $\mathbb{R}^n$ gerado pelas linhas de A. Por exemplo, para

$$A = \begin{bmatrix} 2 & 0 & 1 & 4 \\ 0 & 3 & 1 & 2 \end{bmatrix} \tag{17}$$

temos

$$EL(A) = L(\{(2, 0, 1, 4), (0, 3, 1, 2)\}) \subset \mathbb{R}^4$$

Quando aplicamos o método de Gauss(-Jordan) a uma matriz, o espaço das linhas não muda. De facto suponhamos que

$$A = A_1 \to A_2 \to \cdots \to A_k$$

é uma sucessão de matrizes obtida por aplicação do método de Gauss(-Jordan) à matriz A. Uma vez que as linhas de A_{i+1} são combinações lineares das linhas da matriz A_i temos que

$$\{\text{linhas de } A_{i+1}\} \subset EL(A_i)$$

e portanto, pela Proposição 3.16 temos $EL(A_{i+1}) \subset EL(A_i)$. Mas, as linhas de A_i também são combinações lineares das linhas de A_{i+1}, logo $EL(A_i) \subset EL(A_{i+1})$ e conclui-se que $EL(A_i) = EL(A_{i+1})$. O método de Gauss(-Jordan) dá-nos portanto um método para determinar um conjunto de geradores particularmente simples para o espaço das linhas de uma matriz: as linhas não nulas da matriz em escada de linhas reduzida obtida como output do algoritmo.

Analogamente definimos o *espaço das colunas* de uma matriz A de tipo $m \times n$ como o subespaço de $\mathbb{R}^m$ gerado pelas colunas de A. Por exemplo, para a matriz (17) temos

$$EC(A) = L(\{(2,0),(0,3),(1,1),(4,2)\}) = \mathbb{R}^2.$$

Note-se que *não é verdade* que o espaço das colunas permaneça inalterado ao longo da aplicação do método de Gauss.

Definição 3.19. *Um espaço vetorial V diz-se* finitamente gerado *se existe um conjunto finito $S \subset V$ tal que $V = L(S)$.*

Exemplo 3.20. *O espaço vetorial V formado por todos os polinómios reais não é finitamente gerado. De facto, sendo $S = \{p_1, \ldots, p_k\} \subset V$ um conjunto finito de polinómios, e n_i o grau do polinómio p_i podemos tomar*

$$N = \max\{n_1, \ldots, n_k\}$$

e claramente x^{N+1} não pode ser escrito como combinação linear de elementos de S. Isto mostra que não existe um conjunto finito de geradores para V.

3.21. **Dependência linear.** Chegamos agora a um conceito fundamental da Álgebra Linear que generaliza os conceitos de colinearidade e complanaridade para vetores de $\mathbb{R}^3$.

Definição 3.22. *Seja V um espaço vetorial. Um conjunto $S \subset V$ diz-se* linearmente dependente *se existem $v_1, \ldots, v_n \in S$* todos distintos *e escalares $\alpha_1, \ldots, \alpha_n$ não todos nulos tais que*

$$\alpha_1 v_1 + \ldots + \alpha_n v_n = 0$$

Caso contrário, S diz-se linearmente independente. *Um conjunto $B \subset V$ diz-se uma* base *de V se é linearmente independente e gera V.*

Note-se que a negação da condição de dependência linear é logicamente equivalente à seguinte condição, que utilizamos normalmente para testar independência linear:

> S é linearmente independente se e só se dados $v_1, \ldots, v_n$ elementos distintos de S e escalares $\alpha_1, \ldots, \alpha_n$ tais que $\alpha_1 v_1 + \ldots + \alpha_n v_n = 0$ temos necessariamente $\alpha_1 = \cdots = \alpha_n = 0$

Exemplo 3.23.

(i) Seja $S = \{v\}$ um conjunto com um único elemento. Se $v = 0$ então S é linearmente dependente uma vez que $1 \cdot 0$ é uma combinação linear com coeficientes não nulos de elementos de S que produz o vetor 0. Se $v \neq 0$, então S é linearmente independente. De facto, uma combinação linear de elementos de S com coeficientes não nulos é da

forma αv *com* $\alpha \neq 0$ *e é uma consequência dos axiomas de espaço vetorial que sendo* $\alpha \neq 0$ *e* $v \neq 0$ *então* $\alpha v \neq 0$ *(ver o Exercício 8(vi)).*

(ii) *Se* S *contém o vetor nulo então* S *é linearmente dependente (pois* $1 \cdot 0 = 0$*).*

(iii) *Mais geralmente, se* $S \subset S'$ *e* S *é linearmente dependente, o mesmo é verdade para* S' *(pois a combinação linear com coeficientes não todos nulos que certifica a dependência linear de* S*, certifica também a dependência linear de* S'*). Equivalentemente, se* S' *é um conjunto linearmente independente e* $S \subset S'$ *então* S *é também linearmente independente.*

(iv) *Seja* $S = \{v, w\}$ *um conjunto com dois elementos (distintos). Então* S *é linearmente dependente se e só se* v *e* w *são colineares, isto é se um deles é um múltiplo escalar do outro. De facto, se existem* α_1, α_2 *não ambos nulos tais que*

$$\alpha_1 v + \alpha_2 w = 0$$

ou $\alpha_1 \neq 0$ *e então* $v = -\frac{\alpha_2}{\alpha_1} w$*, ou* $\alpha_2 \neq 0$ *e* $w = -\frac{\alpha_1}{\alpha_2} v$*. Reciprocamente se, por exemplo* $v = \alpha w$*,* $v - \alpha w$ *é uma combinação linear nula de* v, w *cujos coeficientes não são todos nulos, pelo que* $\{v, w\}$ *é linearmente dependente.*

(v) *Generalizando o exemplo anterior vemos que um conjunto* $S \subset V$ *é linearmente dependente se e só se um dos elementos de* S *pode ser expresso como uma combinação linear dos restantes elementos de* S*. De facto uma das implicações é imediata e para ver a outra, se* S *é linearmente dependente podemos escolher* $v_1, \ldots, v_n \in S$ *e escalares* $\alpha_1, \ldots, \alpha_n$ *não todos nulos de tal forma que*

$$\alpha_1 v_1 + \ldots + \alpha_n v_n = 0$$

Assumindo, por exemplo, que $\alpha_i \neq 0$ *temos que*

$$v_i = -\frac{\alpha_1}{\alpha_i} v_1 - \ldots - \frac{\alpha_{i-1}}{\alpha_i} v_{i-1} - \frac{\alpha_{i+1}}{\alpha_i} v_{i+1} - \ldots - \frac{\alpha_n}{\alpha_i} v_n$$

é uma combinação linear de $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n$*.*

(vi) *O subconjunto* $\{(1,2),(0,3),(1,0)\} \subset \mathbb{R}^2$ *é linearmente dependente uma vez que*

$$(1,2) - (1,0) - \tfrac{2}{3}(0,3) = (0,0)$$

Como nenhum par de vetores do conjunto é colinear, se retirarmos qualquer dos vetores ao conjunto obtemos um conjunto linearmente independente, que claramente gera $\mathbb{R}^2$ *e constitui portanto uma base para* $\mathbb{R}^2$*.*

(vii) *Sejam* $e_1 = (1, 0, \ldots, 0), e_2 = (0, 1, 0, \ldots, 0), \ldots, e_n = (0, \ldots, 0, 1) \in \mathbb{R}^n$*. O conjunto* $B_{can} = \{e_1, e_2, \ldots, e_n\}$ *é uma base de* $\mathbb{R}^n$ *chamada a* base canónica*. De facto, dado* $(x_1, \ldots, x_n) \in \mathbb{R}^n$ *temos*

$$(x_1, \ldots, x_n) = x_1 e_1 + \ldots + x_n e_n$$

logo $L(B_{can}) = \mathbb{R}^n$ *e se* $\alpha_1, \ldots, \alpha_n$ *são números reais e* $\alpha_1 e_1 + \ldots + \alpha_n e_n = 0$ *então dado que*

$$\alpha_1 e_1 + \ldots + \alpha_n e_n = (\alpha_1, \ldots, \alpha_n)$$

temos $\alpha_1 = \cdots = \alpha_n = 0$ *o que mostra que* B_{can} *é linearmente independente.*

(viii) *Se A é uma matriz $m \times n$ em escada de linhas, então as linhas não nulas constituem uma base para $EL(A)$. De facto já vimos acima que as linhas não nulas geram $EL(A)$ e se uma combinação linear das linhas se anular, o sistema para os coeficientes da combinação linear que se obtém considerando apenas as componentes correspondentes às colunas que contêm pivots implica imediatamente que os coeficientes da combinação linear são todos nulos. Por exemplo, para*

$$A = \begin{bmatrix} 2 & 1 & 1 & 4 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

olhando apenas para a primeira e terceira componente dos vetores na equação

$$\alpha_1(2,1,1,4) + \alpha_2(0,0,1,2) = (0,0,0,0)$$

vemos que

$$2\alpha_1 = 0 \qquad e \qquad \alpha_1 + \alpha_2 = 0$$

pelo que $\alpha_1 = \alpha_2 = 0$.

O método de Gauss dá-nos portanto uma maneira prática de determinar uma base para o espaço das linhas de uma matriz (e, na prática, para qualquer subespaço de um espaço vetorial finitamente gerado).

(ix) *É um exercício simples verificar que $\{1, x, x^2, \ldots, x^n, \ldots\}$ é uma base para o espaço vetorial dos polinómios reais.*

Intuitivamente, uma base para um espaço vetorial é um "referencial". De facto, se B é uma base de V, os coeficientes da combinação linear que exprime um vetor $v \in V$ em termos dos elementos de B são *únicos*: Admitindo que $B = \{v_1, \ldots, v_n\}$, qualquer vetor v pode ser escrito na forma

$$v = \alpha_1 v_1 + \ldots + \alpha_n v_n$$

(porque B gera V) mas se tivermos também

$$v = \beta_1 v_1 + \ldots + \beta_n v_n$$

então subtraindo as duas igualdades temos

$$0 = (\alpha_1 - \beta_1)v_1 + \ldots + (\alpha_n - \beta_n)v_n$$

e, uma vez que, B é um conjunto linearmente independente, isto implica que $\alpha_1 - \beta_1 = 0, \ldots, \alpha_n - \beta_n = 0$. Os coeficientes dos elementos da base chamam-se as *coordenadas de v na base B*. Uma base permite assim identificar os vetores de V com listas de n escalares (ou seja com $\mathbb{K}^n$).

3.24. **Bases e dimensão.** O primeiro Teorema da Álgebra Linear é que todo o espaço vetorial tem uma base e que todas as bases têm o mesmo número de elementos. Esse número chama-se a *dimensão de V*. Vamos apenas mostrar este Teorema no caso de espaços finitamente gerados, para os quais a dimensão é um número finito. O caso geral é tratado em [FIS, Secção 1.7].

Este teorema será uma consequência de certas propriedades da relação de dependência linear que passamos a explicar. Sugerimos que ao ler os enunciados que se seguem se tenha em mente o exemplo de $\mathbb{R}^3$ e a interpretação geométrica usual da combinação linear de vetores no espaço assim como dos subespaços lineares de $\mathbb{R}^3$ - retas, planos, etc.

Proposição 3.25. *Seja V um espaço vetorial e $S \subset V$ um conjunto linearmente independente. Se $v \notin L(S)$ então $S \cup \{v\}$ é linearmente independente.*

Dem. Temos a mostrar que dados $v_1, \ldots, v_n \in S \cup \{v\}$ vetores distintos e $\alpha_1, \ldots, \alpha_n$ escalares tais que $\alpha_1 v_1 + \ldots + \alpha_n v_n = 0$ então todos os α_i se anulam. Uma vez que S é linearmente independente, esta afirmação é verdadeira quando os vetores v_i pertencem todos a S. Resta portanto considerar o caso em que um deles é v e podemos sem perda de generalidade assumir que $v_n = v$.

Se $n = 1$, então dado que $0 \in L(S)$ e portanto $v \neq 0$ conclui-se que $\alpha_1 v_1 = \alpha_1 v = 0 \Rightarrow \alpha_1 = 0$ (Exemplo 3.23(i)). Se $n > 1$ começamos por notar que α_n é necessariamente 0 pois senão

$$v = v_n = -\frac{\alpha_1}{\alpha_n} v_1 - \cdots - \frac{\alpha_{n-1}}{\alpha_n} v_{n-1}$$

é uma combinação linear de elementos de S, contrariando a hipótese da Proposição. Mas então

$$\alpha_1 v_1 + \ldots + \alpha_{n-1} v_{n-1} = 0$$

Como S é linearmente independente e $v_1, \ldots, v_{n-1}$ são vetores distintos de S temos então $\alpha_1 = \cdots = \alpha_{n-1} = 0$. $\square$

Podemos usar o resultado anterior para construir indutivamente um conjunto linearmente independente com a mesma expansão linear que um dado conjunto finito S de vetores. Mais geralmente temos o seguinte resultado.

Proposição 3.26. *Seja V um espaço vetorial e $S \subset V$ um subconjunto finito. Se $T \subset S$ é linearmente independente, podemos escolher $T' \subset S$ tal que*

- *$T \cup T'$ é linearmente independente*
- *$L(T \cup T') = L(S)$*

Dem. A demonstração é por indução no número de elementos de $S \setminus T$. Suponhamos que $S \setminus T = \{v\}$ tem apenas um elemento. Se $v \in L(T)$ então $L(S) = L(T)$ e podemos tomar $T' = \emptyset$. Se $v \notin L(T)$ podemos aplicar a Proposição 3.25 para concluir $T \cup \{v\}$ é linearmente independente e então $T' = \{v\}$ satisfaz as condições requeridas.

Suponhamos agora que a afirmação do enunciado é válida sempre que o conjunto de geradores tem mais n elementos do que o seu subconjunto linearmente independente. Suponhamos que $S \setminus T$ tem $n+1$ elementos. Se S for linearmente independente, então podemos tomar $T' = S \setminus T$. Se S é linearmente dependente podemos escolher $v_1, \ldots, v_k \in S$ distintos e escalares não todos nulos $\alpha_1, \ldots, \alpha_k$ tais que

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_k v_k = 0$$

Sem perda de generalidade podemos assumir que todos os escalares α_i são não nulos (senão podemos simplesmente suprimir os termos para os quais $\alpha_i = 0$). Uma vez que T é linearmente independente, nem todos os elementos v_i podem pertencer a T. Seja j tal que $v_j \in S \setminus T$ e defina-se $S' = S \setminus \{v_j\}$. Então

$$v_j = -\frac{1}{\alpha_j} \sum_{i \neq j} \alpha_i v_i \in L(S'),$$

e portanto $L(S) = L(S')$. Note-se agora que $T \subset S'$ e $S' \setminus T$ tem n elementos. Aplicando a hipótese de indução ao conjunto de geradores S' e ao subconjunto linearmente independente $T \subset S'$ obtemos um conjunto linearmente independente $T' \subset S'$ tal que $T \cup T'$ é linearmente independente e $L(T \cup T') = L(S') = L(S)$. □

Por palavras, a Proposição anterior diz que dado um conjunto finito de vetores S, todo o subconjunto linearmente independente $T \subset S$ pode ser completado numa base de $L(S)$ contida em S. Uma consequência imediata da Proposição anterior é a existência de bases para espaços finitamente gerados:

Corolário 3.27. *Se V é um espaço vetorial e $S \subset V$ é um conjunto finito, S contém uma base de $L(S)$. Em particular todo o espaço vetorial finitamente gerado tem uma base com um número finito de elementos.*

Dem. Aplicando a Proposição 3.26 com $T = \emptyset$ obtemos uma base T' para $L(S)$ contida em S. Se V for um espaço vetorial finitamente gerado, podemos escolher um subconjunto finito $S \subset V$ tal que $L(S) = V$, e portanto V tem uma base finita. □

Vejamos agora que todas as bases de um espaço vetorial finitamente gerado têm o mesmo número de elementos. A chave da demonstração é o seguinte resultado. Recomendamos que leia o enunciado seguinte com um exemplo simples em mente: $V = \mathbb{R}^2$ ou $\mathbb{R}^3$ e $n = 1$ ou $n = 2$.

Lema 3.28 (Lema da substituição)**.** *Seja V um espaço vetorial e S um subconjunto de V com m elementos. Seja T um subconjunto linearmente independente de $L(S)$ com n elementos. Então $n \leq m$ e existe um subconjunto $T' \subset S$ com $m - n$ elementos tal que $L(T \cup T') = L(S)$.*

Dem. A demonstração é por indução no número de elementos de T. Quando $n = 0$ não há nada a provar, pois $0 \leq m$ e podemos tomar $T' = S$.

Suponhamos que o resultado é válido quando o subconjunto linearmente independente tem n elementos. Seja S um conjunto com m elementos e seja $T = \{v_1, \ldots, v_{n+1}\} \subset L(S)$ um conjunto linearmente independente com $n + 1$ elementos. Uma vez que $\{v_1, \ldots, v_n\}$ é linearmente independente, por hipótese de indução temos que $n \leq m$. Se n fosse igual a m, a hipótese de indução garantiria que $L(\{v_1, \ldots, v_n\}) = L(S)$ (pois o conjunto T' teria 0 elementos) mas então v_{n+1} pertenceria a $L(\{v_1, \ldots, v_n\})$ contradizendo a hipótese de T ser linearmente independente. Conclui-se que $n + 1 \leq m$.

Por hipótese de indução, existem $u_1, \ldots, u_{m-n}$ vetores de S tais que

$$L(\{v_1, \ldots, v_n, u_1, \ldots, u_{m-n}\}) = L(S)$$

Como $v_{n+1} \in L(S)$ existem escalares $\alpha_1, \ldots, \alpha_n$ e $\beta_1, \ldots, \beta_{m-n}$ tais que

$$v_{n+1} = \alpha_1 v_1 + \ldots + \alpha_n v_n + \beta_1 u_1 + \ldots + \beta_{m-n} u_{m-n}$$

Mas T é linearmente independente, logo algum dos coeficientes β_i tem de ser não nulo. Então

$$u_i = \frac{1}{\beta_i} v_{n+1} - \frac{1}{\beta_i}\left(\sum_{j=1}^{n} \alpha_j v_j + \sum_{j\neq i} \beta_j u_j\right) \in L(T \cup \{u_j \colon j \neq i\})$$

Segue-se que $L(S) = L(T \cup \{u_j \colon j \neq i\})$. Portanto podemos tomar $T' = \{u_j \colon j \neq i\}$, o que conclui a demonstração. □

Corolário 3.29. *Seja V um espaço vetorial finitamente gerado. Então*

(i) Todo o subconjunto linearmente independente $T \subset V$ é finito.
(ii) Todo o subconjunto linearmente independente $T \subset V$ está contido numa base de V.

Dem. Seja S um conjunto finito de geradores para V.

(i) Uma vez que um subconjunto de um conjunto linearmente independente é linearmente independente, se existisse um conjunto linearmente independente infinito haveria conjuntos linearmente independentes finitos com um número arbitrário de elementos. Mas sendo n o número de elementos de S, pelo Lema 3.28 um conjunto linearmente independente contido em $L(S) = V$ tem no máximo n elementos.
(ii) Por (i) sabemos que T é finito, logo $S' = S \cup T$ é um conjunto finito de geradores para V. Aplicando a Proposição 3.26 ao conjunto finito de geradores S' e ao seu subconjunto linearmente independente T, obtemos um subconjunto $T' \subset S'$ tal que $T \cup T'$ é uma base de V.

□

Teorema 3.30. *Seja V um espaço vetorial finitamente gerado. Então todas as bases de V são conjuntos finitos com o mesmo número de elementos.*

Dem. Seja $S \subset V$ um conjunto finito com $L(S) = V$. Pelo Corolário 3.27, S contém uma base B para V. Seja n o número de elementos de B. Uma vez que $L(B) = V$, o Lema 3.28 garante que qualquer subconjunto linearmente independente de V tem no máximo n elementos. Seja B' outra base para V. Uma vez que B' é linearmente independente, B' tem no máximo n elementos. Mas $L(B') = V$, e $B \subset V$ é linearmente independente pelo que uma nova aplicação do Lema 3.28 mostra que $n = \sharp B \leq \sharp B'$. Conclui-se que B e B' têm o mesmo número de elementos. □

Definição 3.31. *O número de elementos de qualquer base de um espaço finitamente gerado chama-se a* dimensão de V *e denota-se por* $\dim V$ *ou* $\dim_{\mathbb{K}}(V)$ *para enfatizar o corpo dos escalares. Se um espaço vetorial V não tem uma base finita, diz-se que tem* dimensão infinita.

Note-se que um espaço vetorial tem dimensão infinita se e só se não é finitamente gerado, ou, equivalentemente, é de dimensão finita se e só se é finitamente gerado. Esta consequência imediata do Corolário 3.27 fica como exercício.

Exemplo 3.32. *Tendo em conta o Exemplo 3.23(vii),(viii) e (ix) temos*

(i) $\dim \mathbb{R}^n = n$.
(ii) Se A é uma matriz, então $\dim EL(A)$ *é igual à característica da matriz A.*
(iii) O espaço dos polinómios tem dimensão infinita.

Intuitivamente, a dimensão de um conjunto é o número de parâmetros reais (ou coordenadas) que necessitamos para descrever os pontos do conjunto. Por exemplo a superfície da Terra tem dimensão 2 pois um ponto à superfície da terra é descrito por dois números reais - a latitude e a longitude. Estas questões serão discutidas mais tarde na disciplina de Cálculo 2. O Teorema 3.30 encoraja esta nossa intuição ao afirmar que numa gama restrita de exemplos - aqueles em que o conjunto em questão tem a estrutura de um espaço vetorial finitamente gerado - não há qualquer ambiguidade quanto ao número de parâmetros necessários para descrever o conjunto.

Exemplo 3.33. *A dimensão do espaço $M_{2\times 4}(\mathbb{R})$ é 8. De facto é imediato verificar que as oito matrizes*

$$E_{11} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, E_{12} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \ldots, E_{24} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

constituem uma base. Mais geralmente $\dim M_{m\times n}(\mathbb{R}) = mn$. *Uma base é dada pelas matrizes $\{E_{ij}\}_{1\leq i\leq m, 1\leq j\leq n}$ onde E_{ij} designa a matriz que tem 1 como entrada ij e todas as restantes entradas iguais a 0.*

Conhecer a dimensão de um espaço vetorial dá-nos muita informação sobre esse espaço que, como veremos em seguida, pode ser muito útil para a realização de cálculos no espaço em questão.

Corolário 3.34. *Seja V um espaço vetorial de dimensão n.*

(i) Qualquer conjunto linearmente independente com n vetores é uma base de V.
(ii) Qualquer conjunto de geradores de V tem pelo menos n elementos.
(iii) Qualquer conjunto linearmente independente tem no máximo n elementos. Equivalentemente, todo o conjunto com mais de n elementos é linearmente dependente.

Dem. (i) Seja T um conjunto linearmente independente com n elementos. Pelo Corolário 3.29(ii) existe uma base B contendo T. Como todas as bases têm n elementos, $B = T$ portanto T é uma base.
(ii) Seja S um conjunto de geradores. Se S é infinito não há nada a provar. Se S é finito, pelo Corolário 3.27, S contém uma base e tem portanto pelo menos n elementos.
(iii) Pelo Corolário 3.29(ii) todo o conjunto linearmente independente está contido numa base e portanto tem no máximo n elementos.

□

Observação 3.35. *Com excepção do Corolário 3.34(i), todos os resultados demonstrados acima que assumem que o espaço vetorial é finitamente gerado admitem versões para espaços vetoriais arbitrários. Por exemplo em qualquer espaço vetorial é verdade que duas* *bases têm o mesmo número de elementos, no sentido em que é possível definir uma correspondência bijetiva entre os elementos de uma base e da outra. A demonstração destas versões mais gerais requer alguns conhecimentos de Teoria dos Conjuntos pelo que não discutiremos estes resultados.*

Vejamos como as propriedades dos conjuntos linearmente independentes e bases demonstrados acima podem auxiliar no cálculo de bases e na determinação se um conjunto é ou não linearmente dependente.

Exemplo 3.36. *Vamos verificar que o conjunto* $B = \{(1,0,1),(1,1,0),(0,0,3)\}$ *é uma base para* $\mathbb{R}^3$ *e determinar as componentes de* $(1,2,1)$ *nesta base.*

Uma vez que $\dim \mathbb{R}^3 = 3$*, de acordo com o Corolário 3.34(i) para ver que* B *é uma base basta-nos verificar que* B *é um subconjunto linearmente independente de* $\mathbb{R}^3$*. Podemos fazer isto (pelo menos) de duas formas:*

- *Usando a definição:* B *é linearmente independente se e só se*

$$\alpha(1,0,1)+\beta(1,1,0)+\gamma(0,0,3)=(0,0,0) \Rightarrow \alpha=\beta=\gamma=0$$

  A equação à esquerda da implicação é um sistema linear homogéneo cujas incógnitas são os coeficientes α, β, γ*. Resolvendo o sistema vemos se o conjunto é ou não linearmente independente:*

$$\begin{cases} \alpha+\beta=0 \\ \beta=0 \\ \alpha+3\gamma=0 \end{cases} \Leftrightarrow \begin{cases} \alpha=0 \\ \beta=0 \\ \gamma=0 \end{cases}$$

  o que mostra que B *é linearmente independente. Neste caso não se justificava utilizar o método de Gauss para resolver o sistema, mas vale a pena notar (para quando as contas sejam mais complicadas) que o sistema em questão tem como coeficientes a matriz cujas colunas são os elementos do conjunto* B*. No exemplo acima:*

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 3 \end{bmatrix}$$

- *Alternativamente podemos usar a observação feita no Exemplo 3.23(viii) acima. Se escrevermos os elementos de* B *nas linhas de uma matriz e aplicarmos o método de Gauss à matriz obteremos, no final, uma base para* $L(B)$ *e, em particular, calcularemos a dimensão da expansão linear de* B*.* B *será linearmente independente se e só se* $\dim L(B)$ *for igual ao número de elementos de* B*. De facto, se* $\dim L(B) < \#B$ *então pela Corolário 3.34 (iii)* B *será linearmente dependente. Por outro lado, se* $\dim L(B) = \#B$*,* B *não pode ser linearmente dependente porque, se assim fosse, o Corolário3.27 (i) garantiria a existência de uma base para* $L(B)$ *com menos elementos que* B *o que contradiria o Teorema 3.30.*

Finalmente, a determinação das componentes de um vetor numa dada base consiste na solução de um sistema linear:

$$\alpha(1,0,1)+\beta(1,1,0)+\gamma(0,0,3)=(1,2,1)$$

que podemos escrever na forma de uma matriz aumentada

$$\left[\begin{array}{ccc|c} 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 1 & 0 & 3 & 1 \end{array}\right] \xrightarrow{L_3-L_1} \left[\begin{array}{ccc|c} 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & -1 & 3 & 0 \end{array}\right] \xrightarrow{L_3+L_2} \left[\begin{array}{ccc|c} 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 3 & 2 \end{array}\right]$$

donde obtemos os coeficientes $\alpha=-1, \beta=2, \gamma=\frac{2}{3}$.

Exemplo 3.37. *Consideremos o conjunto* $S=\left\{\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} 0 & 2 \\ 1 & 3 \end{bmatrix}\right\} \subset M_{2\times 2}(\mathbb{R})$. *Vamos determinar uma base para o subespaço* $L(S) \subset M_{2\times 2}(\mathbb{R})$ *e completá-la de forma a obter uma base para* $M_{2\times 2}(\mathbb{R})$.

A observação básica para realizar estes cálculos é que estas matrizes se identificam naturalmente com vetores de $\mathbb{R}^4$ *através da correspondência*

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \Leftrightarrow (a,b,c,d)$$

De facto tanto a soma como o produto por escalar são, em ambos os casos, efetuados coordenada a coordenada. Para determinar uma base para $L(S)$ *podemos portanto (conforme o Exemplo 3.23(viii)) aplicar o método de Gauss a uma matriz cujas linhas são os vetores de* $\mathbb{R}^4$ *correspondentes aos elementos de* S*:*

$$\begin{bmatrix} 1 & 2 & 0 & 1 \\ -1 & 0 & 1 & 2 \\ 0 & 2 & 1 & 3 \end{bmatrix} \xrightarrow{L_2+L_1} \begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 2 & 1 & 3 \\ 0 & 2 & 1 & 3 \end{bmatrix} \xrightarrow{L_3-L_2} \begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 2 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Conclui-se que uma base para $L(S)$ *é*

$$\left\{\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 2 \\ 1 & 3 \end{bmatrix}\right\}$$

(e portanto $L(S)$ *tem dimensão 2). Para completar este conjunto de forma a obter uma base de* $M_{2\times 2}(\mathbb{R})$ *precisamos de juntar dois vetores ao conjunto acima de forma a que o conjunto resultante seja ainda linearmente independente. Isto porque* $\dim M_{2\times 2}(\mathbb{R})=4$ *e portanto, pelo Corolário 3.34, qualquer subconjunto linearmente independente de* $M_{2\times 2}(\mathbb{R})$ *com quatro elementos constitui uma base para* $M_{2\times 2}(\mathbb{R})$.

Podemos novamente apoiar-nos na correspondência entre $M_{2\times 2}(\mathbb{R})$ *e* $\mathbb{R}^4$ *e no facto de as linhas de uma matriz em escada de linhas serem linearmente independentes. Uma vez que*

$$\begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 2 & 1 & 3 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

está em escada de linhas, o conjunto

$$\left\{ \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 2 \\ 1 & 3 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$

é uma base de $M_{2\times 2}(\mathbb{R})$ *contendo a base achada para* $L(S)$*.*

3.38. **Decomposições em soma direta.** Como iremos ver mais tarde, é frequentemente útil ”decompor”espaços vetoriais em subespaços mais pequenos. Trata-se de um procedimento que abstrai a familiar possibilidade de decompor um vetor no plano ou no espaço como a soma das suas componentes ao longo de eixos de um referencial dado (ver Figura 1).

Definição 3.39. *Seja* V *um espaço vetorial e* U, W *subespaços de* V*. Diz-se que* V *é a* soma direta *de* U *com* W*, e escreve-se* $V = U \oplus W$ *se*

(i) Todo o elemento $v \in V$ *se pode escrever como uma soma* $v = u + w$ *com* $u \in U$ *e* $w \in W$*.*

(ii) A decomposição do ponto anterior é única, isto é se $v = u + w$ *e* $v = u' + w'$ *com* $u, u' \in U$ *e* $w, w' \in W$ *então* $u = u'$ *e* $w = w'$*.*

Se U, W *verificam apenas a propriedade* (i) *diz-se que* V *é a* soma *de* U *com* W *e escreve-se* $V = U + W$*.*

Na Figura 1 o plano $V = L(\{v, w\})$ é a soma direta dos seus subespaços de dimensão 1 dados por $U = L(\{v\})$ e $W = L(\{w\})$. A seguinte proposição dá critérios úteis para verificar que temos uma decomposição em soma direta.

Proposição 3.40. *Seja* V *um espaço vetorial e* U, W *subespaços de* V*. As seguintes afirmações são equivalentes:*

(i) $V = U \oplus W$*.*

(ii) $V = U + W$ *e* $U \cap W = \{0\}$*.*

Se V *é finitamente gerado*[15] *então as afirmações anteriores são ainda equivalentes a*

(iii) Existe uma base $\{v_1, \ldots, v_n\}$ *para* V *e* $k \leq n$ *tais que* $\{v_1, \ldots, v_k\}$ *é uma base para* U *e* $\{v_{k+1}, \ldots, v_n\}$ *é uma base para* W*.*

Dem: $(i) \Rightarrow (ii)$: Suponhamos que $V = U \oplus W$. Então, em particular, $V = U + W$. Dado $v \in U \cap W$ temos duas decomposições de v como soma de um vetor de U e de um vetor de W, nomeadamente $v = v + 0$ e $v = 0 + v$. Uma vez que a decomposição é única conclui-se que $v = 0$ e portanto $U \cap W = \{0\}$.

$(ii) \Rightarrow (i)$: Temos a mostrar que se $V = U + W$ e $U \cap W = \{0\}$ então a decomposição de um vetor $v \in V$ como a soma de um vetor em U e outro em W é única. Suponhamos que $v \in V$ se pode decompor como

$$v = u + w \quad \text{e} \quad v = u' + w' \qquad \text{com } u, u' \in U,\ w, w' \in W$$

[15]Para V arbitrário pode mostrar-se que V é a soma direta de U e W se e só se existe uma base B para V e uma decomposição de B como a união de dois subconjuntos B_1 e B_2 tais que B_1 é uma base para U e B_2 é uma base para V.

Então temos

$$u + w = u' + w' \Rightarrow u - u' = w' - w$$

O vetor do lado esquerdo do sinal de igual está em U (porque a diferença de vetores em U está em U) e o do lado direito em W. Como são iguais, trata-se de um vetor em $U \cap W = \{0\}$. Portanto $u - u' = 0 = w' - w$, ou seja, $u = u'$ e $w = w'$ pelo que a decomposição de v é única.

Suponhamos agora que V é finitamente gerado e vejamos primeiro que $(i) \Rightarrow (iii)$: Começamos por notar que se um espaço vetorial Z não é finitamente gerado então aplicações sucessivas da Proposição 3.25 começando com $S = \emptyset$ mostram que Z contém conjuntos linearmente independentes de tamanho arbitrário. Pelo Corolário 3.34(iii) isto implica que um subespaço de um espaço de dimensão finita é necessariamente finitamente gerado. Logo U, W têm dimensão finita.

Sejam $\{u_1, \ldots, u_k\}$ e $\{w_1, \ldots, w_l\}$ bases de U e W. É suficiente verificar que $B = \{u_1, \ldots, u_k, w_1, \ldots, w_l\}$ é uma base para V.

Como B contém uma base de U temos $U \subset L(B)$ e analogamente $W \subset L(B)$. Uma vez que $L(B)$ é fechado para a soma conclui-se que $V = U + W \subset L(B)$ logo B gera V. Resta ver que B é linearmente independente. Sejam $\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_l$ escalares tais que

$$\alpha_1 u_1 + \ldots + \alpha_k u_k + \beta_1 w_1 + \ldots + \beta_l w_l = 0$$

Então

$$\alpha_1 u_1 + \ldots + \alpha_k u_k = -\beta_1 w_1 - \ldots - \beta_l w_l$$

O vetor do lado esquerdo do sinal de igual pertence a U e o do lado direito a W. Sendo iguais, pertencem a $U \cap W = \{0\}$. Uma vez que $\{u_1, \ldots, u_k\}$ e $\{w_1, \ldots, w_l\}$ são linearmente independentes segue-se que os α_i's e os β_j's são 0 e portanto B é linearmente independente.

Resta ver que $(iii) \Rightarrow (ii)$. Dado $v \in V$ podemos escrevê-lo de forma única como

$$v = \alpha_1 v_1 + \ldots + \alpha_n v_n = (\alpha_1 v_1 + \ldots + \alpha_k v_k) + (\alpha_{k+1} v_{k+1} + \ldots + \alpha_n v_n)$$

logo v é a soma de um vetor em U com um vetor em W, e portanto $V = U + W$. Dado $v \in U \cap W$, têm que existir escalares $\alpha_1, \ldots, \alpha_k$ tais que $v = \alpha_1 v_1 + \ldots + \alpha_k v_k$ e escalares $\beta_1, \ldots, \beta_{n-k}$ tais que $v = \beta_1 v_{k+1} + \ldots + \beta_{n-k} v_n$. Mas então

$$\alpha_1 v_1 + \ldots + \alpha_k v_k - \beta_1 v_{k+1} - \ldots - \beta_{n-k} v_n = 0$$

e sendo B linearmente independente isto só pode acontecer se todos os α_i's e β_j's forem 0. Conclui-se que $v = 0$ e portanto $U \cap W = \{0\}$ conforme pretendido. □

Exemplo 3.41. *Seja $V = F(\mathbb{R}, \mathbb{R})$ o espaço vetorial de todas as funções reais de variável real. Se $U = \{f \in V : f(x) = f(-x)\}$ é o subespaço vetorial das funções pares e $W = \{f \in V : f(x) = -f(-x)\}$ o subespaço vetorial formado pelas funções ímpares temos*

$$V = U \oplus W$$

De facto a expressão

$$f(x) = \frac{f(x) + f(-x)}{2} + \frac{f(x) - f(-x)}{2}$$

mostra que qualquer função $f \in V$ se pode escrever como a soma de uma função par e de uma função ímpar e, dado que a única função que é simultaneamente par e ímpar é a função identicamente nula, temos $U \cap W = \{0\}$.

Podemos mais geralmente dizer que V é a soma direta dos seus subespaços $U_1, \ldots, U_k$ se todo o vetor $v \in V$ se puder escrever de forma única como uma soma de vetores $u_i \in U_i$. Fica como exercício enunciar e demonstrar a generalização natural da Proposição 3.40 a estas decomposições em mais fatores.

3.42. **Mudanças de coordenadas.**

Definição 3.43. *Uma* base ordenada *B de um espaço vetorial de dimensão finita V é uma sequência finita $B = (v_1, \ldots, v_n)$ de vetores distintos $v_i \in V$ tais que o conjunto $\{v_1, \ldots, v_n\}$ é linearmente independente e gera V*

Como o nome indica, a diferença entre base e base ordenada é que numa base ordenada escolhemos explicitamente uma ordem para os vetores da base. Há um primeiro vetor da base, um segundo, etc... Na realidade, até agora, ao fazer cálculos já escolhemos implicitamente uma ordem para os vetores das bases envolvidas de forma a poder identificar o espaço vetorial em questão com $\mathbb{R}^n$.

Uma base ordenada $B = (v_1, \ldots, v_n)$ determina uma bijeção natural

$$V \longleftrightarrow \mathbb{R}^n$$

que faz corresponder a um vetor $v \in V$ os seus coeficientes na base B, na ordem indicada, (cf. discussão após o Exemplo 3.23)

$$v = \alpha_1 v_1 + \ldots + \alpha_n v_n \longleftrightarrow (\alpha_1, \ldots, \alpha_n)$$

O escalar α_i diz-se *a i-ésima coordenada* de v na base ordenada B. Convém ter uma notação para as coordenadas de um vetor $v \in V$ numa dada base ordenada.

Definição 3.44. *Seja $B = (v_1, \ldots, v_n)$ uma base ordenada para o espaço vetorial V e $v = \alpha_1 v_1 + \ldots + \alpha_n v_n$ um vetor de V. Escrevemos*

$$[v]_B = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}$$

para o vetor das coordenadas *de v na base B.*

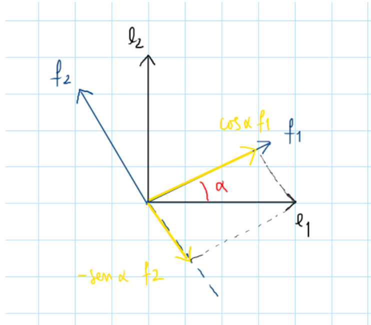
Exemplo 3.45.

(i) A base ordenada canónica de $\mathbb{R}^n$ é $B_{can} = (e_1, \ldots, e_n)$, onde $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$ (com o 1 na posição i). Uma vez que

$$(x_1, \ldots, x_n) = x_1 e_1 + \ldots + x_n e_n$$

as coordenadas de $(x_1, \ldots, x_n)$ na base canónica são $(x_1, \ldots, x_n)$.

(ii) Para $0 < \alpha < \frac{\pi}{2}$, seja $B = ((\cos\alpha, \operatorname{sen}\alpha), (-\operatorname{sen}\alpha, \cos\alpha))$ a base ordenada de $\mathbb{R}^2$ que se obtém rodando os vetores da base canónica um ângulo α no sentido anti-horário. Vamos achar as coordenadas do vetor $(1, 0)$ na base B.



Podemos fazê-lo usando a interpretação geométrica das coordenadas (conforme o Exemplo 3.6) e trigonometria elementar obtendo $(\cos\alpha, -\operatorname{sen}\alpha)$ *ou, alternativamente, resolvendo o sistema*

$$(1,0) = c_1(\cos\alpha, \operatorname{sen}\alpha) + c_2(-\operatorname{sen}\alpha, \cos\alpha) \Leftrightarrow \begin{cases} c_1\cos\alpha - c_2\operatorname{sen}\alpha = 1 \\ c_1\operatorname{sen}\alpha + c_2\cos\alpha = 0 \end{cases}$$

A combinação linear $\cos\alpha L_1 + \operatorname{sen}\alpha L_2$ *das duas equações do sistema produz* $c_1 = \cos\alpha$, *e substituindo na segunda equação temos*

$$\cos\alpha\operatorname{sen}\alpha + c_2\cos\alpha = 0 \Leftrightarrow c_2 = -\operatorname{sen}\alpha$$

(uma vez que $\cos\alpha > 0$*). Em geral, podemos ver geometricamente qual é a relação entre as coordenadas* (a,b) *de um vetor na base canónica e as suas coordenadas na base* B*. As coordenadas na base* B *obtêm-se de* (a,b) *rodando este vetor um ângulo* α no sentido horário.

Vimos no exemplo anterior que as coordenadas na nova base B podiam ser obtidas a partir das coordenadas noutra base (a base canónica) através de uma certa transformação. É natural perguntar em geral qual é a relação entre as coordenadas de um vetor $v \in V$ em duas bases ordenadas $B_1 = (v_1, \ldots, v_n)$ e $B_2 = (w_1, \ldots, w_n)$ de V dadas.

Seja

$$v = \alpha_1 v_1 + \ldots + \alpha_n v_n$$

Para achar as coordenadas de v na base B_2 podemos escrever os vetores v_i na base B_2:

$$\begin{aligned} v_1 &= a_{11}w_1 + a_{21}w_2 + \ldots + a_{n1}w_n \\ v_2 &= a_{12}w_1 + a_{22}w_2 + \ldots + a_{n2}w_n \\ &\vdots \\ v_n &= a_{1n}w_1 + a_{2n}w_2 + \ldots + a_{nn}w_n \end{aligned}$$

Substituindo na fórmula para v obtemos

$$\begin{aligned} v &= \alpha_1(a_{11}w_1 + a_{21}w_2 + \ldots + a_{n1}w_n) + \alpha_2(a_{12}w_1 + a_{22}w_2 + \ldots + a_{n2}w_n) + \\ &\quad \ldots + \alpha_n(a_{1n}w_1 + a_{2n}w_2 + \ldots + a_{nn}w_n) \\ &= (a_{11}\alpha_1 + a_{12}\alpha_2 + \ldots + a_{1n}\alpha_n)w_1 + (a_{21}\alpha_1 + a_{22}\alpha_2 + \ldots + a_{2n}\alpha_n)w_2 + \\ &\quad \ldots + (a_{n1}\alpha_1 + a_{n2}\alpha_2 + \ldots + a_{nn}\alpha_n)w_n \end{aligned}$$

Escrevendo $(\beta_1, \ldots, \beta_n)$ para as coordenadas do vetor v na base B_2 temos portanto

$$\begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}$$

onde na coluna j da matriz $[a_{ij}]$ aparecem as coordenadas do vetor v_j na base B_2. Usando a notação da Definição 3.44 a igualdade anterior escreve-se

$$[v]_{B_2} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix} [v]_{B_1}$$

Proposição 3.46. *Seja V um espaço vetorial de dimensão n e B_1 e B_2 bases ordenadas para V. Existe uma única matriz $n \times n$, denotada por $S_{B_1 \to B_2}$, tal que para todo o vetor $v \in V$, as coordenadas de v na base B_2 e as coordenadas de v na base B_1 estão relacionadas por*

$$[v]_{B_2} = S_{B_1 \to B_2}[v]_{B_1}$$

A esta matriz chama-se a matriz de mudança de coordenadas da base B_1 para a base B_2.

Dem. Já observámos acima que é possível relacionar as coordenadas através de uma matriz. Para ver que a matriz é única note-se que se existir uma tal matriz S então a j-ésima coluna da matriz terá necessariamente de consistir nas coordenadas do j-ésimo vetor da base B_1 na base B_2. De facto, as coordenadas desse vetor (chamemos-lhe v_j) na base B_1 são $(0, \ldots, 0, 1, 0, \ldots, 0)$ com o 1 na j-ésima posição, e ao multiplicarmos a matriz S por este vetor de coordenadas obtemos a j-ésima coluna de S que tem então que conter as coordenadas de v_j na base B_2. $\square$

Exemplo 3.47. *A matriz de mudança de base da base canónica B_{can} de $\mathbb{R}^2$ para a base B do Exemplo 3.45 é dada por*

$$S_{B_{can} \to B} = \begin{bmatrix} \cos\alpha & \operatorname{sen}\alpha \\ -\operatorname{sen}\alpha & \cos\alpha \end{bmatrix}$$

De facto, a primeira coluna contém as componentes do primeiro vetor da base canónica na base B como vimos no Exemplo 3.45 e da mesma forma podemos verificar que a segunda coluna contém as coordenadas do vetor $(0,1)$ na base B. Note-se que o efeito que tem a multiplicação desta matriz por um vetor coluna é a rotação do vetor um ângulo α no sentido horário conforme tínhamos previsto geometricamente.

Proposição 3.48. *Seja V um espaço vetorial de dimensão finita e B_1, B_2, B_3 bases ordenadas para V. Temos as seguintes relações entre as matrizes de mudança de coordenadas:*

(i) $S_{B_1 \to B_3} = S_{B_2 \to B_3} S_{B_1 \to B_2}$
(ii) $S_{B_2 \to B_1} = (S_{B_1 \to B_2})^{-1}$

Dem. (i) Por definição das matrizes de mudança de coordenadas temos para todo o $v \in V$

$$[v]_{B_2} = S_{B_1 \to B_2}[v]_{B_1}, \qquad [v]_{B_3} = S_{B_2 \to B_3}[v]_{B_2}$$

Substituindo a primeira equação na segunda obtemos

$$[v]_{B_3} = S_{B_2 \to B_3}\left(S_{B_1 \to B_2}[v]_{B_1}\right) = \left(S_{B_2 \to B_3} S_{B_1 \to B_2}\right)[v]_{B_1}$$

Uma vez que a equação anterior é válida para qualquer vetor $v \in V$ e a matriz de mudança de coordenadas é única conclui-se que

$$S_{B_1 \to B_3} = S_{B_2 \to B_3} S_{B_1 \to B_2}$$

(ii) Claramente, para qualquer base ordenada B com n elementos, temos que a matriz de mudança de coordenadas da base B para ela própria é a matriz identidade I_n. Aplicando o ponto (i) com $B_3 = B_1$ obtemos

$$I_n = S_{B_2 \to B_1} S_{B_1 \to B_2}$$

e da mesma forma, trocando B_1 com B_2

$$I_n = S_{B_1 \to B_2} S_{B_2 \to B_1}$$

o que mostra que $S_{B_1 \to B_2}$ e $S_{B_2 \to B_1}$ são matrizes inversas.

□

Exemplo 3.49. *Vamos determinar a matriz de mudança de coordenadas $S_{B_{can} \to B}$ da base canónica de $\mathbb{R}^3$ para a base $B = ((1,0,1),(1,0,-1),(1,1,1))$. É fácil escrever a matriz $S_{B \to B_{can}}$ (só temos que escrever os vetores da base B por ordem nas colunas da matriz):*

$$S_{B \to B_{can}} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & -1 & 1 \end{bmatrix}$$

Pela Proposição 3.48 temos

$$S_{B_{can} \to B} = S^{-1}_{B \to B_{can}} = \begin{bmatrix} \frac{1}{2} & -1 & \frac{1}{2} \\ \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 1 & 0 \end{bmatrix}$$

Por exemplo, as coordenadas do vetor $(1,2,3)$ na base B são $(0,-1,2)$.

Observação 3.50. *Note-se que o ponto (ii) da Proposição anterior diz, em particular, que uma matriz de mudança de base é sempre invertível. Reciprocamente, no Exercício 34 irão verificar que qualquer matriz invertível é uma matriz de mudança de base (a partir de qualquer base dada).*

3.51. **Exercícios.**

1. Para cada um dos conjuntos V abaixo com as operações indicadas, mostre que se trata de um espaço vetorial ou indique justificadamente quais os axiomas de espaço vetorial (real ou complexo conforme o caso) que não são satisfeitos. Neste exercício vamos excepcionalmente usar o símbolo $\oplus$ para a soma de vetores e $\odot$ para o produto por escalar numa tentativa de evitar confusão com as operações habituais.
   (i) $V = \mathbb{R}^2$ com a operação de soma definida por $(x, y) \oplus (x', y') \stackrel{\text{def}}{=} (x - 2x', y - 3y')$ e a multiplicação por escalar usual.
   (ii) $V = \mathbb{C}^n$ com a operação de soma usual e a operação de multiplicação por escalar (complexo) definida por $\alpha \odot (z_1, \ldots, z_n) \stackrel{\text{def}}{=} (\bar{\alpha} z_1, \ldots, \bar{\alpha} z_n)$.
   (iii) $V = \mathbb{R}^n$ com a soma usual e a multiplicação por escalar definida por $\alpha \odot (x_1, \ldots, x_n) \stackrel{\text{def}}{=} (0, \ldots, 0)$.
   (iv) $V = \mathbb{R}^+ = \{x \in \mathbb{R} \colon x > 0\}$ com a soma definida por $x \oplus y \stackrel{\text{def}}{=} xy$ e a multiplicação por escalar real definida por $\alpha \odot x \stackrel{\text{def}}{=} x^\alpha$.

   Resolução.
2. Seja $F(S, \mathbb{R})$ o espaço vetorial das funções do conjunto S para $\mathbb{R}$ com a soma e produto por escalar usuais e $M_{m \times n}(\mathbb{R})$ o espaço vetorial das matrizes $m \times n$ com a soma e produto por escalar usuais. Para cada subconjunto $W \subset V$ indique justificadamente se W é um subespaço vetorial de V
   (i) $V = F(\mathbb{R}, \mathbb{R})$, e W o subconjunto das funções contínuas.
   (ii) $V = F(\mathbb{R}, \mathbb{R})$ e W o subconjunto das funções duas vezes diferenciáveis tais que $f''(0) = 1$ e $f'(0) = 0$.
   (iii) V o espaço vetorial dos polinómios reais e W o conjunto dos polinómios de grau exatamente 3 juntamente com o polinómio nulo.
   (iv) V o espaço vetorial dos polinómios reais e W o conjunto dos polinómios reais tais que $p''(1) = p(3)$.
   (v) $V = M_{n \times n}(\mathbb{R}))$ e $W = \{A \in M_{n \times n}(\mathbb{R}) \colon A^T = A\}$ o subconjunto das matrizes *simétricas*.
   (vi) $W = \{A \in M_{m \times n}(\mathbb{R}) \colon a_{ij} \geq 0 \text{ para todo o } i, j\}$

   Resolução.
3. Para cada espaço vetorial V e conjunto S indicados, determine se o elemento $v \in V$ pertence a $L(S)$ e nesse caso ache os coeficientes de uma combinação linear que exprima v em termos de S.
   (a) $V = \mathbb{R}^3$, $v = (1, 2, 1)$, $S = \{(1, 0, 1), (2, 1, 3), (-1, -1, -2)\}$
   (b) $V = M_{2 \times 2}(\mathbb{R})$, $v = \begin{bmatrix} 2 & 5 \\ 4 & 1 \end{bmatrix}$ e $S = \left\{ \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}, \begin{bmatrix} -1 & 2 \\ 0 & 3 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \right\}$
   (c) $V = F(\mathbb{R}, \mathbb{R})$, $v = e^x$, $S = \{x, \operatorname{sen} x, x^3, \cos x\}$.
   (d) V o espaço vectorial dos polinómios reais, $v = 1 + x - 2x^2$ e $S = \{1 - x + x^2, x^2 - 1, x + 2, x^2 + x\}$.

   Resolução.

4. Determine um conjunto finito S de vetores tal que $L(S)$ seja igual ao núcleo da matriz A dada.
(a) $A = \begin{bmatrix} 1 & 1 & 1 \\ -1 & -2 & 3 \end{bmatrix}$ (b) $A = \begin{bmatrix} 1 & 1 & 0 & 2 \\ -1 & 0 & 1 & 1 \\ -1 & 1 & -2 & 0 \end{bmatrix}$
Resolução.
5. Determine uma matriz A cujo núcleo seja
(i) $L(\{(1,0,2),(-1,1,0)\})$
(ii) $L(\{(1,1,1,1),(-1,1,0,1)\})$
Resolução.
6. Determine (sistemas de) equações homogéneas que descrevam os conjuntos $L(S) \subset V$ para o conjunto S e o espaço vetorial V indicados.
(i) $S = \{(1,0,2,1),(-1,3,2,0)\}$ com $V = \mathbb{R}^4$.
(ii) $S = \{1 - x + x^3, x + x^2\}$ com V o espaço vetorial dos polinómios de grau ≤ 4.
(iii) $S = \left\{ \begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & 3 \end{bmatrix}, \begin{bmatrix} 2 & 0 & 1 \\ -1 & -2 & 3 \end{bmatrix} \right\}$ com $V = M_{2\times 3}(\mathbb{R})$.
Resolução.
7. Seja $V = F(\mathbb{R},\mathbb{R})$ o espaço vetorial das funções reais. Mostre que se $\alpha, \beta, \gamma \in \mathbb{R}$ são tais que
$$\alpha e^x + \beta e^{2x} + \gamma e^{3x}$$
é a função nula então $\alpha = \beta = \gamma = 0$. Resolução.
8. Seja V um espaço vetorial. Mostre a partir dos axiomas de espaço vetorial:
(i) Se $u + v = u + w$ então $v = w$.
(ii) Para todo o $v \in V$ tem-se $0v = 0$.
(iii) Para todo o $\alpha \in \mathbb{R}$ tem-se $\alpha 0 = 0$ (onde 0 designa o vetor nulo).
(iv) O simétrico de um vetor é único.
(v) Para todo o $v \in V$ tem-se $(-1)v = -v$.
(vi) Se $\alpha v = 0$ então $\alpha = 0$ ou $v = 0$.
(vii) Se $v \neq 0$ e $\alpha v = \beta v$ então $\alpha = \beta$.
Resolução.
(O exercício seguinte é para alunos mais interessados em aprofundar os conhecimentos matemáticos - não faz parte do programa.)
9. Recorde que para um primo p, $\mathbb{F}_p = \{\underline{0}, \underline{1}, \ldots, \underline{p-1}\}$ denota o corpo[16] com p elementos cujas operações são dadas pela soma e multiplicação nos inteiros seguidas do cálculo do resto da divisão por p. Quantos subespaços vetoriais distintos tem $\mathbb{F}_p^2$? Resolução.
10. Seja V um espaço vetorial. Mostre que se $W_1, W_2 \subset V$ são subespaços vetoriais então $W_1 \cap W_2$ também é um subespaço vetorial de V. Mais geralmente mostre que se $\{W_\alpha\}$ é uma família de subespaços vetoriais de V então a sua interseção $\cap_\alpha W_\alpha$ é um subespaço vetorial de V. Resolução.
11. Sejam V_1 e V_2 espaços vetoriais sobre $\mathbb{R}$ (ou $\mathbb{C}$). Mostre que o produto cartesiano $V_1 \times V_2$ com a soma e produto por escalar definidos componente a componente é um

[16] Ver a definição de corpo na Observação 1.2.

espaço vetorial. Este espaço chama-se a *soma direta* de V_1 e V_2 e denota-se por $V_1 \oplus V_2$.
Resolução.

12. Seja V um espaço vetorial sobre $\mathbb{R}$ e W um subespaço vetorial. Este exercício define o *espaço vetorial quociente* V/W. Dados $v \in V$ escrevemos
$$v + W = \{v + w \colon w \in W\} \subset V$$
Intuitivamente, $v + W$ é o plano paralelo a W que passa por v. Seja
$$V/W = \{v + W \colon v \in V\}$$
Note que V/W é um conjunto cujos elementos são subconjuntos de V.
(a) Mostre que $v + W = v' + W$ se e só se $v - v' \in W$.
(b) Mostre que as operações
$$\mathbb{R} \times V/W \to V/W \qquad V/W \times V/W \to V/W$$
dadas por
$$(\alpha, v + W) \mapsto (\alpha v) + W, \qquad (v_1 + W, v_2 + W) \mapsto (v_1 + v_2) + W$$
estão bem definidas. Por exemplo, para a primeira operação, isto significa verificar que se $v + W = v' + W$ então para todo o escalar α temos $(\alpha v) + W = (\alpha v') + W$.
(c) Mostre que o conjunto V/W com as operações definidas na alínea anterior é um espaço vetorial sobre $\mathbb{R}$ (ou $\mathbb{C}$).
(d) Tomando $V = \mathbb{R}^3$ e $W = \{(x, y, z) \in V \colon x = y = z\}$ determine se $x = (1, -1, 1) + W$ pertence à expansão linear de $S = \{(0, 1, 2) + W, (1, 2, -1) + W\}$ e em caso afirmativo determine os coeficientes de uma expressão de x como combinação linear dos elementos de S.
Resolução.

13. Verifique se os seguintes conjuntos de vetores são linearmente independentes. Se forem linearmente dependentes, indique uma combinação linear não nula de vetores distintos do conjunto que se anula.
(a) $\{(1, 0, -2, 1), (2, 2, 1, 0), (0, 1, 2, 1), (0, -1, -3, 3)\}$ em $\mathbb{R}^4$
(b) $\{1 + t + t^2, -t + t^3, 2 + t\}$ no espaço vetorial dos polinómios reais.
(c) $\{(1, i, 0), (i, -1, 1 + i), (0, 0, 2i)\}$ em $\mathbb{C}^3$.
(d) $\left\{ \begin{bmatrix} 1 & 1 & 0 & 2 \\ -1 & 1 & -2 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 0 & 3 \\ -1 & 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 0 & 2 \\ -1 & -1 & 1 & 0 \end{bmatrix} \right\}$ em $M_{2\times 4}(\mathbb{R})$.
Resolução.

14. Determine as coordenadas dos vetores na base indicada
(a) $1 - t^2$ na base $\{1, 1 + t, 1 + t + t^2\}$ para o espaço dos polinómios reais de grau ≤ 2.
(b) $(1, 4, 3, 2)$ na base $\{(1, 0, 0, 1), (1, 0, -1, 0), (1, 1, 0, 1), (1, 0, 0, -1)\}$ para $\mathbb{R}^4$
(c) $\begin{bmatrix} 1 & 2 \\ -1 & -1 \end{bmatrix}$ na base $\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$ para $M_{2\times 2}(\mathbb{R})$.
Resolução.

15. Determine bases para os núcleos, espaços das linhas e espaço das colunas das seguintes matrizes:

(a) $\begin{bmatrix} 1 & 1 & 0 \\ -1 & 2 & 1 \\ 1 & 0 & -1 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 1 & 0 & 2 \\ -1 & -1 & 1 & 0 \end{bmatrix}$ (c) $\begin{bmatrix} 1 & 1 & 0 & 2 \\ -1 & -1 & 1 & 0 \\ 0 & 0 & 2 & 4 \end{bmatrix}$

Resolução.

16. Determine uma base para os subespaços vetoriais U e complete-as de forma a obter uma base do espaço vetorial indicado.
    (a) $U = \{p(t)\colon p'(1) = 0\}$ no espaço vetorial dos polinómios de grau ≤ 2.
    (b) $U = L\left(\left\{\begin{bmatrix} 1 & 2 \\ -1 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}\right\}\right)$ em $M_{2\times 2}(\mathbb{R})$.

    Resolução.
17. Sendo V e W os subespaços vetoriais indicados. Determine uma base para $U \cap V$.
    (a) $U = \{(x, y, z) \in \mathbb{R}^3\colon x + y + z = 0\}$ e $V = L(\{(1, 3, 4), (0, 1, 1)\})$ em $\mathbb{R}^3$.
    (b) $U = \{(x, y, z, w) \in \mathbb{R}^4\colon x + z - w = 0\}$ e $V = \{(x, y, z, w) \in \mathbb{R}^4\colon x - y + 2z - w = 0\}$ em $\mathbb{R}^4$.
    (c) $U = L(\{(-1, 3, 1, 0), (0, 3, 3, 1)\}$ e $V = L(\{(0, 1, 1, 0), (1, 1, 3, 1)\}$.
    (d) $U = \{p(t)\colon p(0) = 0, p'(1) = 0\}$ e $V = L(\{1 + t, t^2, t + t^4\})$ no espaço vetorial dos polinómios de grau ≤ 4.

    Resolução.
18. Seja V um espaço vetorial finitamente gerado. Mostre que se W é um subespaço vetorial de V e $W \neq V$ então $\dim W < \dim V$. Resolução.
19. (a) Seja A uma mat8riz $m \times n$. Mostre que $\dim EL(A) + \dim N(A) = n$.
    (b) Sendo $A = \begin{bmatrix} 1 & a & b & c \\ d & e & f & 3 \end{bmatrix}$ determine a, b, c, d, e, f tais que $(2, 0, 3, -1) \in EL(A)$ e $\dim N(A) = 3$.

    Resolução.
20. Seja V um espaço vetorial complexo. Mostre que V é naturalmente um espaço vetorial real e que $\dim_{\mathbb{R}} V = 2 \dim_{\mathbb{C}} V$. Resolução.
21. Seja W um espaço vetorial e $U, V \subset W$ subespaços vetoriais.
    (a) A *soma* dos subespaços U, V é $U + V = \{u + v\colon u \in U, v \in V\}$. Mostre que $U + V$ é um subespaço vetorial de W.
    (b) Considere a operação da alínea anterior no conjunto dos subespaços vetoriais de W. Esta operação é comutativa? associativa? tem elemento neutro? tem inversos?
    (c) Mostre que se W é finitamente gerado então $\dim(U + V) = \dim(U) + \dim(V) - \dim(U \cap V)$.
    (d) Recorde que se $W = U + V$ e $U \cap V = \{0\}$ se diz que W é a soma direta de U e V e escrevemos $W = U \oplus V$. Sendo

    $$W = M_{n\times n}(\mathbb{R}) \qquad U = \{A \in W\colon A = A^T\} \qquad V = \{A \in W\colon A + A^T = 0\}$$

    mostre que $W = U \oplus V$ e dê uma fórmula para a componente de $A \in W$ segundo U e V.
    (e) Mostre que se W é finitamente gerado, para todo o subespaço $U \subset W$ existe um subespaço $V \subset W$ tal que $U \oplus V = W$.

    Resolução.

22. (Interpolação de Lagrange) Sejam $t_0, t_1, \ldots t_n$ números reais distintos. Considere os polinómios de grau n definidos por
$$p_i(t) = \frac{\prod_{j \neq i}(t - t_j)}{\prod_{j \neq i}(t_i - t_j)}$$
para $i = 0, \ldots, n$. Mostre que $\{p_0, p_1, \ldots, p_n\}$ é uma base para o espaço dos polinómios de grau menor ou igual a n. *Sugestão: Avalie em $t = t_k$ para cada k.* Quais são as coordenadas de um polinómio $p(t)$ nesta base? Resolução.
23. Mostre que $\{1, 1+t, 1+t+t^2, 1+t+t^2+t^3, \ldots, 1+t+t^2+\ldots+t^n, \ldots\}$ é uma base para o espaço vetorial de todos os polinómios reais. Resolução.
24. O conjunto dos números reais $\mathbb{R}$ é um espaço vetorial sobre $\mathbb{Q}$ com a soma e multiplicação habituais. Mostre que $\dim_{\mathbb{Q}} \mathbb{R} = \infty$. Resolução.
25. Considere o espaço vetorial $F(\mathbb{R}, \mathbb{R})$ das funções reais de variável real. Mostre que o subconjunto $C(\mathbb{R})$ formado pelas funções contínuas é um subespaço vetorial e que o subconjunto $\{e^{\alpha x} \colon \alpha \in \mathbb{R}\}$ é linearmente independente em $C(\mathbb{R})$. Resolução.
26. Seja V um espaço vetorial finitamente gerado e $W \subset V$. Recordando o Exercício 12, calcule $\dim(V/W)$ em função de $\dim(V)$ e $\dim(W)$. Resolução.
27. Sejam $B_1 = ((1,1),(1,-1))$ e B_2 bases ordenadas para $\mathbb{R}^2$. Supondo que a matriz de mudança de coordenadas
$$S_{B_1 \to B_2} = \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}$$
Determine as coordenadas na base B_2 do vetor $(3,4) \in \mathbb{R}^2$. Resolução.
28. Sabendo que os vetores $u, v \in \mathbb{R}^2$ têm respetivamente coordenadas $(1,2)$ e $(-1,3)$ na base B_1 e $(2,2)$ e $(-1,4)$ na base B_2, determine a matriz mudança de base $S_{B_1 \to B_2}$. Resolução.
29. Considere as seguintes bases ordenadas de $M_{2\times 2}(\mathbb{R})$:
$$B_1 = \left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right)$$
e
$$B_2 = \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix}, \begin{bmatrix} 0 & \frac{1}{2} \\ -\frac{1}{2} & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \right)$$
Determine a matriz mudança de base $S_{B_1 \to B_2}$. Resolução.
30. Considere a base ordenada $B_1 = ((1,0,0),(1,1,0),(1,1,1))$ de $\mathbb{R}^3$. Supondo que a matriz de mudança de base é
$$S_{B_1 \to B_2} = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 2 \\ 0 & 1 & 0 \end{bmatrix}$$
determine B_2. Resolução.
31. Seja V o espaço vetorial dos polinómios de grau ≤ 2 e $B_1 = (1, x, x^2)$ a base ordenada habitual. Seja $t \in \mathbb{R}$ um número real qualquer
    (a) Mostre que o conjunto $B_2 = (1, (x-t), (x-t)^2)$ é uma base de V.
    (b) Determine a matriz de mudança de base $S_{B_1 \to B_2}$.

Resolução.

32. Sejam B_1, B_2, B_3 bases ordenadas de $\mathbb{R}^3$. Sendo

$$S_{B_1 \to B_3} = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 2 \\ 0 & -1 & 0 \end{bmatrix} \qquad S_{B_2 \to B_3} = \begin{bmatrix} 0 & 1 & 0 \\ 2 & -1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

Determine $S_{B_3 \to B_2}$ e $S_{B_1 \to B_2}$. Resolução.

33. Seja B_1 a base canónica de $\mathbb{R}^2$ e B_2 a base tal que

$$S_{B_1 \to B_2} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$

Escrevendo (x, y) para as coordenadas de um vetor na base B_1 e (u, v) para as coordenadas do mesmo vetor na base B_2 expresse os seguintes conjuntos, função e equação em termos das coordenadas (u, v).

(a) $\{(x, y) \in \mathbb{R}^2 \colon x^2 + y^2 = 1\}$
(b) $\{(x, y) \in \mathbb{R}^2 \colon x^2 + 2xy + y^2 + x + 2y = 0\}$
(c) $f(x, y) = x^2 + \cos(xy)$
(d) $\frac{dx}{dt} + x\frac{dy}{dt} = x + y$

Resolução.

34. Seja V um espaço vetorial de dimensão n. Mostre que se B_1 é uma base ordenada para V e S é uma matriz $n \times n$ invertível, existe uma base ordenada B_2 tal que $S_{B_1 \to B_2} = S$. Resolução.

4. Transformações lineares

Em cada área da Matemática estudamos um certos tipo de objetos matemáticos cuja natureza é variável. Por exemplo, em Álgebra Linear estudamos espaços vetoriais, enquanto que em Geometria se pode estudar, por exemplo, curvas e superfícies. Normalmente estes objetos consistem em conjuntos munidos de certa estrutura adicional. No caso dos espaços vetoriais esta estrutura adicional toma a forma das operações de soma de vetores e o produto de vetores por escalares. Para estudar os objetos em questão é sempre necessário pensar em como se relacionam entre eles. As relações entre os objetos manifestam-se através de funções entre os conjuntos subjacentes que preservam a estrutura adicional. No caso que nos interessa agora isso leva-nos à seguinte definição.

Definição 4.1. *Sejam V e W espaços vetoriais. Uma função $f \colon V \to W$ diz-se uma* transformação linear *de V para W se*

(i) $f(v_1 + v_2) = f(v_1) + f(v_2)$ para todos os $v_1, v_2 \in V$.
(ii) $f(\alpha v) = \alpha f(v)$ para todo o $v \in V$ e escalar α.

As transformações lineares são portanto as funções entre os conjuntos subjacentes aos espaços vetoriais que preservam a soma e o produto por escalar. Note-se que na definição acima aparecem duas somas (em geral) distintas no axioma (i): do lado esquerdo do sinal de igual, a soma é a soma de vetores em V, enquanto que do lado direito se trata da soma em W. Analogamente para os dois produtos por escalar que aparecem no axioma (ii).

Chamamos a atenção para as seguintes consequências imediatas dos axiomas acima: uma transformação linear leva necessariamente o vetor $0 \in V$ no vetor $0 \in W$. De facto, sendo $v \in V$ um vetor qualquer sabemos que $0 \cdot v = 0$. Como f preserva o produto por escalar temos então

$$f(0) = f(0 \cdot v) = 0 \cdot f(v) = 0 \in W$$

A outra observação importante é que uma transformação linear leva combinações lineares em V para combinações lineares em W: dados escalares $\alpha_1, \ldots, \alpha_n$ e vetores $v_1, \ldots, v_n$ temos

$$\begin{aligned} f(\alpha_1 v_1 + \ldots + \alpha_n v_n) &= f(\alpha_1 v_1) + f(\alpha_2 v_2) + \ldots + f(\alpha_n v_n) \\ &= \alpha_1 f(v_1) + \ldots + \alpha_n f(v_n) \end{aligned}$$

Vejamos alguns exemplos de transformações lineares $f \colon V \to W$.

Exemplo 4.2.

(1) Sejam $V = W = \mathbb{R} = \mathbb{R}^1$. A função $f \colon \mathbb{R} \to \mathbb{R}$ definida pela expressão $f(x) = 2x$ é uma transformação linear. De facto temos

$$f(x_1 + x_2) = 2(x_1 + x_2) = 2x_1 + 2x_2 = f(x_1) + f(x_2)$$

$$f(\alpha x) = 2(\alpha x) = \alpha(2x) = \alpha f(x)$$

O gráfico de f é uma linha reta que passa pela origem. Mais geralmente, é fácil ver (exercício) que uma função $f \colon \mathbb{R} \to \mathbb{R}$ é uma transformação linear se e só se f é uma função linear, isto é, da forma $f(x) = ax$ para algum número real $a \in \mathbb{R}$. Assim, as transformações lineares de $\mathbb{R}$ para $\mathbb{R}$ são as funções reais de variável real cujos gráficos são retas que passam pela origem.

Por exemplo, a expressão $f(x) = 3x + 1$ não define uma transformação linear de $\mathbb{R}$ para $\mathbb{R}$. De facto $f(0 + 0) = 1$ é diferente de $f(0) + f(0) = 1 + 1 = 2$. Alternativamente, $f(0) = 1 \neq 0$ e vimos acima que uma transformação linear leva sempre o vetor nulo do conjunto de partida no vetor nulo do conjunto de chegada.

(2) Sejam $V = W = \mathbb{R}^2$ e identifiquemos como habitualmente $\mathbb{R}^2$ com o plano. Considere-se a função $f \colon \mathbb{R}^2 \to \mathbb{R}^2$ definida geometricamente como "rotação de 90 graus em torno da origem no sentido anti-horário". Apelando ao significado geométrico da soma de vetores e produto por escalar é imediato verificar que esta transformação preserva a soma de vetores e o produto por escalar pelo que é uma transformação linear.

Podemos verificar a afirmação anterior obtendo uma expressão analítica para a função f. Sendo (a, b) um vetor no primeiro quadrante é imediato verificar que após a rotação o vetor fica com coordenadas $(-b, a)$. É fácil verificar que o mesmo sucede para qualquer vetor pelo que a expressão analítica para a rotação é

$$f(a, b) = (-b, a)$$

Podemos agora ver que f é uma transformação linear:

$$\begin{aligned} f((a_1,b_1)+(a_2,b_2)) &= f(a_1+a_2,b_1+b_2) \\ &= (-b_1-b_2,a_1+a_2)=(-b_1,a_1)+(-b_2,a_2) \\ &= f(a_1,b_1)+f(a_2,b_2) \end{aligned}$$

e

$$f(\alpha(a,b))=f(\alpha a,\alpha b)=(-\alpha b,\alpha a)=\alpha(-b,a)=\alpha f(a,b)$$

Note-se que identificando os vetores de $\mathbb{R}^2$ com matrizes coluna 2×1, podemos escrever f da seguinte forma

$$f\left(\begin{bmatrix} a \\ b \end{bmatrix}\right)=\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix}$$

(3) *Seja $V=\mathbb{R}^n, W=\mathbb{R}^m$ e A uma matriz $m\times n$. Identificando como habitualmente vetores de $\mathbb{R}^n$ com matrizes coluna podemos definir uma transformação linear $f\colon \mathbb{R}^n\to\mathbb{R}^m$ através da fórmula*

$$f(x)=Ax$$

O exemplo anterior é um caso particular deste. De facto, o primeiro exemplo também é. Nesse caso, $A=[a]$ é uma matriz 1×1.

(4) *Seja $W=F(\mathbb{R},\mathbb{R})$ o espaço vetorial das funções reais de variável real e*

$$V=\{f\in W\colon f \text{ é diferenciável}\}$$

o subespaço vetorial formado pelas funções diferenciáveis. Então a aplicação $T\colon V\to W$ definida por

$$T(f)=f'$$

ou seja a operação de derivação, é uma transformação linear. De facto temos

$$T(f+g)=(f+g)'=f'+g'=T(f)+T(g)$$

e

$$T(\alpha f)=(\alpha f)'=\alpha f'=\alpha T(f)$$

pelas regras de derivação para a soma e para o produto por escalar. Estas regras dizem precisamente que a operação de derivação é uma transformação linear. Este exemplo é, pelo menos aparentemente, muito diferente dos anteriores. O conceito de transformação linear estabelece assim uma analogia entre operações tão diferentes como uma rotação do plano e a operação de derivação de uma função.

(5) *Seja $V=M_{m\times n}(\mathbb{R})$ e $W=M_{p\times q}(\mathbb{R})$ e sejam B uma matriz $p\times m$ e C uma matriz $n\times q$. Então a aplicação $T\colon V\to W$ definida pela fórmula*

$$T(A)=BAC$$

é uma transformação linear:

$$\begin{aligned} T(A_1+A_2) &= B(A_1+A_2)C=(BA_1+BA_2)C \\ &= BA_1C+BA_2C=T(A_1)+T(A_2) \end{aligned}$$

(pela distributividade do produto de matrizes em relação à soma, e associatividade da multiplicação de matrizes) e

$$T(\alpha A) = B(\alpha A)C = (\alpha BA)C = \alpha BAC = \alpha T(A)$$

pela relação entre o produto de matrizes e o produto por escalar. Um exemplo concreto é por exemplo a transformação $T\colon M_{2\times 2}(\mathbb{R}) \to M_{4\times 3}(\mathbb{R})$ *determinada pelas matrizes*

$$B = \begin{bmatrix} 1 & 3 \\ -2 & 0 \\ -1 & 1 \\ 2 & 0 \end{bmatrix}, \qquad C = \begin{bmatrix} 0 & 1 & 2 \\ -1 & 1 & 0 \end{bmatrix}$$

que é dada pela fórmula

$$T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} 1 & 3 \\ -2 & 0 \\ -1 & 1 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 \\ -1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} -b-3d & a+b+3c+3d & 2a+6c \\ 2b & -2a-2b & -4a \\ b-d & -a-b+c+d & -2a+2c \\ -2b & 2a+2b & 4a \end{bmatrix}$$

(6) *Seja* V *o espaço vetorial dos polinómios e* $W = \mathbb{R}^2$*. Então a função* $f\colon V \to \mathbb{R}^2$ *definida por*

$$f(p) = (p(1), p''(2))$$

é uma transformação linear:

$$\begin{aligned} f(p+q) &= ((p+q)(1), (p+q)''(2)) = (p(1)+q(1), p''(2)+q''(2)) \\ &= (p(1), p''(2)) + (q(1), q''(2)) = f(p) + f(q) \end{aligned}$$

$$f(\alpha p) = ((\alpha p)(1), (\alpha p)''(2)) = (\alpha p(1), \alpha p''(2)) = \alpha(p(1), p''(2)) = \alpha f(p)$$

porque a soma de funções e a multiplicação de uma função por escalar são calculadas ponto a ponto e pelas regras de derivação. Note-se que este exemplo é, pelo menos aparentemente, de uma natureza bastante diferente dos exemplos (1)-(5) acima.

Proposição 4.3. *Sejam* V, W *espaços vetoriais,* $B = \{v_1, \ldots, v_n\}$ *uma base para* V *e* $w_1, \ldots, w_n$ *vetores quaisquer de* W*. Então existe uma única transformação linear* $f\colon V \to W$ *tal que* $f(v_i) = w_i$*.*

Dem. Começamos por mostrar a unicidade. Suponhamos que $f\colon V \to W$ é uma transformação linear tal que $f(v_i) = w_i$. Dado um vetor $v \in V$ qualquer, existem escalares $\alpha_1, \ldots, \alpha_n$ únicos tais que

$$v = \alpha_1 v_1 + \ldots + \alpha_n v_n$$

Uma vez que uma transformação linear preserva combinações lineares, teremos necessariamente

$$f(v) = f(\alpha_1 v_1 + \ldots + \alpha_n v_n) = \alpha_1 f(v_1) + \ldots + \alpha_n f(v_n) = \alpha_1 w_1 + \ldots + \alpha_n w_n \tag{18}$$

Obtemos assim uma fórmula para f que mostra a unicidade da transformação linear (caso exista). Para verificar que existe e completar a demonstração resta ver que a expressão

(18) define efetivamente uma transformação linear com as propriedades requeridas. Seja então $f\colon V \to W$ a função definida pela expressão (18).

- f envia o vetor $v_i \in B$ em w_i: Temos $v_i = 0\cdot v_1+\ldots+0\cdot v_{i-1}+1\cdot v_i+0\cdot v_{i+1}+\ldots 0\cdot v_n$ logo $f(v_i) = 0\cdot w_1+\ldots+0\cdot w_{i-1}+1\cdot w_i+0\cdot w_{i+1}+\ldots 0\cdot w_n = w_i$.
- f é uma transformação linear: Sejam $v = \alpha_1 v_1 + \ldots + \alpha_n v_n$ e $w = \beta_1 v_1 + \ldots + \beta_n v_n$ dois vetores quaisquer de V. Então $v + w = (\alpha_1 + \beta_1)v_1 + \ldots + (\alpha_n + \beta_n)v_n$ pelo que

$$\begin{aligned} f(v+w) &= (\alpha_1+\beta_1)w_1+\ldots+(\alpha_n+\beta_n)w_n \\ &= (\alpha_1 w_1+\ldots+\alpha_n w_n)+(\beta_1 w_1+\ldots+\beta_n w_n) = f(v)+f(w) \end{aligned}$$

e, dado um escalar α temos $\alpha v = \alpha\alpha_1 v_1 + \ldots + \alpha\alpha_n v_n$ e portanto

$$f(\alpha v) = \alpha\alpha_1 w_1+\ldots+\alpha\alpha_n w_n = \alpha(\alpha_1 w_1+\ldots+\alpha_n w_n) = \alpha f(v)$$

o que conclui a demonstração.

□

O resultado anterior pode ser visto (pelo menos) de duas maneiras diferentes. Por um lado, dá-nos um método para construir transformações lineares: basta escolher uma base para o espaço de partida e decidir qual o valor que irá tomar em cada vetor da base. Além disso a demonstração acima dá-nos uma fórmula ((18)) para a transformação linear assim obtida. Por outro lado, a Proposição diz-nos que as transformações lineares são funções excepcionalmente simples. Para definir uma função de V para W é normalmente necessário decidir o seu valor individualmente para cada vetor de V. A Proposição anterior diz que quando f é linear, todo o comportamento da função é completamente determinado pelos valores que toma num número finito de elementos do domínio (os vetores constituintes de uma base).

Observação 4.4. *A Proposição 4.3 é ainda válida quando a base de V é um conjunto infinito, sendo a demonstração essencialmente a mesma. Deixamos esta verificação como exercício.*

Exemplo 4.5. *A transformação linear $T\colon \mathbb{R}^2 \to \mathbb{R}^3$ tal que $T(1,0) = (2,1,-3)$ e $T(0,1) = (4,1,5)$ é a função definida pela expressão*

$$T(a,b) = a(2,1,-3)+b(4,1,5) = (2a+4b, a+b, -3a+5b)$$

que pode ser representada matricialmente por

$$T\left(\begin{bmatrix} a \\ b \end{bmatrix}\right) = \begin{bmatrix} 2 & 4 \\ 1 & 1 \\ -3 & 5 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}$$

Claramente o exemplo anterior pode ser generalizado a qualquer transformação linear de $\mathbb{R}^m$ para $\mathbb{R}^n$ e vemos assim que o Exemplo 4.2 (3) é na realidade exaustivo.

4.6. **Representação matricial de uma transformação linear.** Vamos agora ver que, em completa generalidade, desde que os espaços vetoriais envolvidos tenham dimensão finita, uma transformação linear é determinada por uma matriz. Recordamos a notação introduzida na Definição 3.44 para as coordenadas de um vetor $v \in V$ numa base ordenada $B = (v_1, \ldots, v_n)$ para V: se $v = \alpha_1 v_1 + \ldots + \alpha_n v_n$ escrevemos

$$[v]_B = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}$$

Note-se que uma base B com n elementos determina uma função $f \colon V \to M_{n\times 1}(\mathbb{R})$ definida por

$$f(v) = [v]_B$$

Pela unicidade das coordenadas, a função f é uma bijecção, e aliás é esta a identificação que temos usado, informalmente, para efetuar cálculos em espaços vetoriais de polinómios e matrizes tratando-os como se fossem algum dos espaços $\mathbb{R}^n$.

Exercício 4.7. *Dado um espaço vetorial V e uma base $B = (v_1, \ldots, v_n)$ para V, verifique que a função $f \colon V \to M_{n\times 1}(\mathbb{R})$ definida por $f(v) = [v]_B$ é uma transformação linear (cf. Exercício 9).*

Proposição 4.8. *Sejam V, W espaços vetoriais e $B_1 = (v_1, \ldots, v_n)$ e $B_2 = (w_1, \ldots, w_m)$ bases ordenadas para V e W respetivamente. Seja $f \colon V \to W$ uma transformação linear. Então existe uma única matriz $A_{f,B_1,B_2} \in M_{m\times n}(\mathbb{R})$ tal que, para todo o vetor $v \in V$ se tem*

$$[f(v)]_{B_2} = A_{f,B_1,B_2}[v]_{B_1}$$

A matriz A_{f,B_1,B_2} diz-se a matriz que *representa a transformação linear f com respeito às bases B_1 e B_2.*

Exemplo 4.9. *(i) Seja V um espaço vetorial com bases $B_1 = (v_1, \ldots, v_n)$ e $B_2 = (w_1, \ldots, w_n)$ e* $\mathrm{Id} \colon V \to V$ *a função identidade (definida por* $\mathrm{Id}(v) = v$*). É imediato verificar que* Id *é uma transformação linear. Temos então, por definição de matriz mudança de base*

$$A_{\mathrm{Id},B_1,B_2} = S_{B_1 \to B_2}$$

De facto, a identidade

$$[\mathrm{Id}(v)]_{B_2} = A_{\mathrm{Id},B_1,B_2}[v]_{B_1} \Leftrightarrow [v]_{B_2} = A_{\mathrm{Id},B_1,B_2}[v]_{B_1}$$

mostra que A_{Id,B_1,B_2} satisfaz a relação que caracteriza a matriz de mudança de coordenadas.

(ii) Seja V o espaço vetorial dos polinómios de grau ≤ 3 e considere-se a transformação linear $T \colon V \to V$ definida por $T(p) = p'$. Uma vez que

$$T(a + bx + cx^2 + dx^3) = b + 2cx + 3dx^2,$$

sendo $B = (1, x, x^2, x^3)$ *a base canónica, a equação* $[T(p)]_B = A_{T,B,B}[p]_B$ *para a matriz* $A_{T,B,B}$ *fica*

$$\begin{bmatrix} b \\ 2c \\ 3d \\ 0 \end{bmatrix} = A_{T,B,B} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$$

e conclui-se então que

$$A_{T,B,B} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Vale a pena refletir durante um momento no facto de a matriz acima representar a operação de derivação (embora no contexto restrito dos polinómios de grau menor ou igual a 3*).*

Dem. da Proposição 4.8. Vejamos primeiro que se a matriz A_{f,B_1,B_2} existir, ela é única. Para o i-ésimo vetor da base B_1, $v = v_i$, a equação que caracteriza a matriz A_{f,B_1,B_2} é

$$[f(v_i)]_{B_2} = A_{f,B_1,B_2}[v_i]_{B_1}$$

mas, uma vez que $[v_i]_{B_1}$ tem todas as entradas iguais a 0 exceto a i-ésima que é igual a 1, o produto no termo direito da equação acima é a i-ésima coluna da matriz A_{f,B_1,B_2}. Isto mostra que a matriz A_{f,B_1,B_2} fica univocamente determinada: se existir, a sua i-ésima coluna é necessariamente igual a $[f(v_i)]_{B_2}$.

Para completar a demonstração basta agora verificar que a matriz $m \times n$ A_{f,B_1,B_2} cuja i-ésima coluna é $[f(v_i)]_{B_2}$ satisfaz a equação do enunciado. Seja $v = \alpha_1 v_1 + \ldots + \alpha_n v_n$ um vetor de V. Então

$$\begin{aligned} [f(v)]_{B_2} &= [f(\alpha_1 v_1 + \ldots + \alpha_n v_n)]_{B_2} \\ &= [\alpha_1 f(v_1) + \ldots + \alpha_n f(v_n)]_{B_2} \\ &= \alpha_1 [f(v_1)]_{B_2} + \ldots + \alpha_n [f(v_n)]_{B_2} \end{aligned}$$

onde na segunda igualdade usámos o facto de f ser uma transformação linear e na terceira o Exercício 4.7. Pela definição do produto de matrizes a expressão

$$\alpha_1 [f(v_1)]_{B_2} + \ldots + \alpha_n [f(v_n)]_{B_2}$$

é exatamente o produto da matriz que tem por i-ésima coluna $[f(v_i)]_{B_2}$ pelo vetor coluna com componentes $(\alpha_1, \ldots, \alpha_n)$, ou seja, é exatamente $Af, B_1, B_2[v]_{B_1}$. Isto conclui a demonstração. □

A Proposição 4.8 permite identificar uma transformação linear entre espaços vetoriais de dimensão finita com uma matriz mediante a escolha de bases para o espaço vetorial de partida e de chegada. Além disso explica como obter a matriz em questão: é a matriz cuja i-ésima coluna contém as coordenadas da imagem do i-ésimo vetor da base do espaço de partida na base do espaço de chegada.

Isto é extremamente útil para fazer contas com transformações lineares como iremos ver em seguida. Convém no entanto notar que a Proposição não se aplica a todos os exemplos de transformação linear que queremos considerar - por exemplo, à operação de derivação de funções diferenciáveis arbitrárias. Por outro lado, o objeto em que normalmente estamos interessados é a transformação linear ela própria e não uma (das muitas possíveis) representações matriciais que usamos para calcular. Uma analogia que pode ser útil é que uma transformação linear é como uma ideia, que se pode exprimir em várias línguas, as bases nos espaços de partida e de chegada são como uma escolha de língua, e a matriz que representa a transformação linear é a palavra que representa a ideia na língua escolhida.

4.10. **Operações com transformações lineares e a sua tradução em matrizes.** As transformações lineares podem ser combinadas através de várias operações que agora passamos a descrever.

Definição 4.11. *Sejam V e W espaços vetoriais. Escrevemos $L(V,W)$ para o conjunto das transformações lineares de V para W. Dadas $f,g \in L(V,W)$ e um escalar α definimos a* soma *de f e g como sendo a função $f+g\colon V \to W$ definida pela expressão*

$$(f+g)(v) = f(v) + g(v)$$

e definimos o produto de uma transformação linear f pelo escalar α *como sendo a função $\alpha f\colon V \to W$ definida pela expressão*

$$(\alpha f)(v) = \alpha \cdot f(v).$$

Proposição 4.12. *Sejam V e W espaços vetoriais. Com as operações de soma e produto por escalar definidas acima, o conjunto $L(V,W)$ é um espaço vetorial.*

Dem. Temos a verificar que as operações de soma e produto por escalar estão bem definidas, isto é, que dadas $f,g \in L(V,W)$ e um escalar f, as funções $f+g$ e αf estão ainda em $L(V,W)$ e depois os oito axiomas que estas operações devem satisfazer num espaço vetorial.

Vemos primeiro que $f+g$ é uma transformação linear: dados $v_1, v_2 \in V$ temos

$$\begin{aligned}(f+g)(v_1+v_2) &= f(v_1+v_2) + g(v_1+v_2) = f(v_1) + f(v_2) + g(v_1) + g(v_2) \\ &= f(v_1) + g(v_1) + f(v_2) + g(v_2) = (f+g)(v_1) + (f+g)(v_2)\end{aligned}$$

e dado um escalar α e $v \in V$ temos

$$(f+g)(\alpha v) = f(\alpha v) + g(\alpha v) = \alpha f(v) + \alpha g(v) = \alpha(f(v)+g(v)) = \alpha((f+g)(v))$$

A verificação que $(\alpha f) \in L(V,W)$ é análoga e fica como exercício. A verificação dos axiomas de espaço vetorial é também deixada como exercício. Notamos apenas que o vetor $0 \in L(V,W)$ é a transformação linear identicamente nula que envia todos os vetores $v \in V$ para $0 \in W$. □

Proposição 4.13. *Sejam V, W, U espaços vetoriais e $f\colon V \to W$, e $g\colon W \to U$ transformações lineares. Então a função composta*

$$g \circ f\colon V \to U$$

é uma transformação linear.

Dem. Temos a verificar que $g \circ f$ preserva a soma e o produto por escalar.

- Dados $v_1, v_2 \in V$ temos

$$(g \circ f)(v_1+v_2) = g(f(v_1+v_2)) = g(f(v_1)+f(v_2)) = g(f(v_1))+g(f(v_2)) = (g \circ f)(v_1)+(g \circ f)(v_2)$$

  onde na segunda igualdade usámos o facto de f ser uma transformação linear, e na terceira, o facto de g ser uma transformação linear.
- Dados um escalar α e um vetor $v \in V$ temos

$$(g \circ f)(\alpha v) = g(f(\alpha v)) = g(\alpha f(v)) = \alpha g(f(v)) = \alpha (g \circ f)(v)$$

  onde, tal como acima, na segunda igualdade usámos o facto de f ser uma transformação linear, e na terceira, o facto de g ser uma transformação linear.

□

Proposição 4.14. *Sejam V, W espaços vetoriais e $f \colon V \to W$ uma transformação linear. Se a função f é invertível (isto é, se é bijetiva) então a função inversa $f^{-1} \colon W \to V$ é uma transformação linear.*

Demonstração. Temos a verificar que a função inversa f^{-1} preserva a soma e a multiplicação por escalar. Sejam w_1, w_2 vetores de W. Como f é sobrejetiva existem vetores v_1 e v_2 de V tais que $f(v_1) = w_1$ e $f(v_2) = w_2$. Então

$$f^{-1}(w_1 + w_2) = f^{-1}(f(v_1) + f(v_2)) = f^{-1}(f(v_1 + v_2)) = (f^{-1} \circ f)(v_1 + v_2) = v_1 + v_2$$

onde na segunda igualdade usámos o facto de f ser uma transformação linear. Por definição de função inversa temos que $v_1 = f^{-1}(w_1)$ e $v_2 = f^{-1}(w_2)$. Substituindo na igualdade acima concluímos que $f^{-1} \colon W \to V$ preserva a soma de vetores. A verificação que f^{-1} preserva o produto por escalar é análoga e fica como exercício. □

Observação 4.15. *Alternativamente, na demonstração anterior poderíamos ter aplicado a função injetiva (por hipótese) f às expressões $f^{-1}(w_1 + w_2)$ e $f^{-1}(w_1) + f^{-1}(w_2)$ e verificado que essas contas produziam o mesmo resultado. A injetividade de f garante então que $f^{-1}(w_1 + w_2) = f^{-1}(w_1) + f^{-1}(w_2)$.*

Definição 4.16. *Sejam V, W espaços vetoriais. Uma transformação linear invertível $f \colon V \to W$ diz-se um* isomorfismo *de espaços vetoriais.*

A palavra isomorfismo vem de "iso- igual - e "morphos- forma. Um isomorfismo entre dois espaços vetoriais é uma equivalência entre eles. O isomorfismo estabelece uma correspondência bijetiva entre os conjuntos subjacentes (um "dicionário" entre os vetores de um dos espaços e os vetores do outro). Uma vez que a função e a sua inversa preservam as operações dos espaços vetoriais ou, equivalentemente, as combinações lineares, qualquer propriedade ou afirmação acerca de um dos espaços (que se possa expressar usando combinações lineares) será verdadeira se e só se for verdadeira no outro. Por exemplo um conjunto será linearmente (in)dependente num espaço se e só se a sua imagem através do isomorfismo for linearmente (in)dependente no outro - ver o Exercício 14.

Exemplo 4.17.

(i) As funções $M_{n\times 1}(\mathbb{R}) \to \mathbb{R}^n$ *e* $M_{1\times n}(\mathbb{R}) \to \mathbb{R}^n$ *definidas por*

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \mapsto (x_1, \ldots, x_n) \qquad e \qquad \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} \mapsto (x_1, \ldots, x_n)$$

são isomorfismos de espaços vetoriais. De facto as funções descritas acima são claramente bijetivas e também transformações lineares (pela definição de soma e produto por escalar nos vários espaços envolvidos).

(ii) Seja V *um espaço vetorial com base ordenada* $B = (v_1, \ldots, v_n)$*. A função* $f\colon V \to M_{n\times 1}(\mathbb{R})$ *definida por*

$$f(v) = [v]_B$$

que calcula a matriz coluna das coordenadas na base ordenada B *é um isomorfismo. Que* f *é uma transformação linear, é o conteúdo do Exercício 4.7. A função* f *é também bijetiva: a sobrejetividade de* f *traduz o facto que qualquer* n*-tuplo* $(\alpha_1, \ldots, \alpha_n)$ *de escalares formar as coordenadas de um vetor de* V *(nomeadamente de* $v = \alpha_1 v_1 + \ldots \alpha_n v_n$*), enquanto que a injetividade de* f *é uma consequència da unicidade das coordenadas de um vetor (que por sua vez é uma consequência de* B *ser um conjunto linearmente independente).*

(iii) Sejam V, W *espaços vetoriais e* $B_1 = (v_1, \ldots, v_n), B_2 = (w_1, \ldots, w_m)$ *bases ordenadas para* V *e* W *respetivamente. A função*

$$\Phi\colon L(V, W) \to M_{m\times n}(\mathbb{R})$$

definida por (ver Proposição 4.8 para o significado da notação)

$$\Phi(f) = A_{f,B_1,B_2}$$

é um isomorfismo de espaços vetoriais. Portanto uma transformação linear entre espaços vetoriais finitamente gerados pode ser identificada com uma matriz, uma vez escolhidas bases ordenadas para o domínio e conjunto de chegada da transformação linear.

Temos que verificar que Φ *é uma transformação linear e que é invertível (ou bijetiva) enquanto função.*

- *Sejam* $f, g\colon V \to W$ *transformações lineares. Dado* $v \in V$ *temos*

$$[(f+g)(v)]_{B_2} = [f(v) + g(v)]_{B_2} = [f(v)]_{B_2} + [g(v)]_{B_2} \tag{19}$$

onde na primeira igualdade usámos a definição de soma de transformações lineares e na segunda o facto que a operação de calcular as coordenadas é linear. Por definição das matrizes que representam f, g*, e pela distributividade em relação à soma do produto de matrizes obtemos*

$$[f(v)]_{B_2} + [g(v)]_{B_2} = A_{f,B_1,B_2}[v]_{B_1} + A_{g,B_1,B_2}[v]_{B_1} = (A_{f,B_1,B_2} + A_{g,B_1,B_2})[v]_{B_1} \tag{20}$$

Das igualdades (19) *e* (20) *obtemos, novamente por definição da matriz que representa* $(f+g)$*,*

$$A_{f+g,B_1,B_2} = A_{f,B_1,B_2} + A_{g,B_1,B_2}$$

ou seja

$$\Phi(f+g) = \Phi(f) + \Phi(g)$$

A demonstração que $\Phi(\alpha f) = \alpha\Phi(f)$ é análoga e fica como exercício. Concluímos que Φ é uma transformação linear.

- *Recorde-se da demonstração da Proposição 4.8 que a matriz $\Phi(f)$ tem como i-ésima coluna $[f(v_i)]_{B_2}$. Dada uma matriz A, pela Proposição 4.3 e o exemplo (ii) acima existe uma transformação linear f tal que $[f(v_i)]_{B_2}$ é a i-ésima coluna de A. Temos então $\Phi(f) = A$, o que mostra que Φ é sobrejetiva. Por outro lado, suponhamos que f e g são transformações lineares tais que $\Phi(f) = \Phi(g)$ então, para cada $i = 1, \ldots, n$, as coordenadas de $f(v_i)$ e $g(v_i)$ são iguais. Mas isto significa que $f(v_i) = g(v_i)$ para cada i, e então pela Proposição 4.3 temos que $f = g$. Isto mostra que Φ é uma função injetiva e portanto, dado que também é sobrejetiva, invertível.*

Conclui-se que Φ é um isomorfismo de espaços vetoriais. Em linguagem corrente, esta afirmação significa que os dicionários entre transformações lineares e matrizes (dados pela escolha de uma base no espaço de chegada e outra no espaço de partida) traduzem a soma e a multiplicação de um dos lados na soma e multiplicação do outro.

Os exemplos anteriores dizem-nos que qualquer espaço vetorial real finitamente gerado é equivalente a $\mathbb{R}^n$ e que uma transformação linear entre tais espaços pode ser identificada com uma matriz. Estes factos são muito úteis para fazer contas. Já foram usados muitas vezes e continuarão a ser usados até ao final do semestre para esse efeito. No entanto não seria uma boa ideia concluir daqui que nos podemos concentrar exclusivamente em $\mathbb{R}^n$ e nas matrizes. Apesar de ser possível identificar um espaço finitamente gerado com algum $\mathbb{R}^n$ *não há em geral nenhuma maneira canónica de o fazer.* A identificação é feita através de uma *escolha* de base e há muitas escolhas possíveis. Um espaço vetorial geral não possui coordenadas especiais (ao contrário do que acontece em $\mathbb{R}^n$ e em vários outros exemplos que temos vindo a considerar como os espaços de matrizes) e esta é uma diferença muito importante. Veremos em breve que as soluções de certas equações diferenciais formam espaços vetoriais nos quais não há habitualmente qualquer "base canónica". O mesmo se pode dizer para um subespaço vetorial típico de $\mathbb{R}^n$. Pensar num vetor em termos das suas coordenadas numa base é análogo a pensar numa ideia através da palavra que é usada para designar a ideia numa dada língua. É muitas vezes útil mas há que ter consciência de que por vezes tem também desvantagens.

Proposição 4.18. *Sejam V, W, U espaços vetoriais, B_1, B_2, B_3 bases ordenadas para V, W, U respetivamente, e $f\colon V \to W$, $g\colon W \to U$ transformações lineares. Então a matriz que representa a transformação linear $g \circ f$ nas bases dadas é o produto da matriz que representa g pela matriz que representa f. Isto é,*

$$A_{g\circ f, B_1, B_3} = A_{g, B_2, B_3} A_{f, B_1, B_2}$$

Dem. Dado $v \in V$ temos pela definição das matrizes que representam f e g

$$\begin{aligned}[(g \circ f)(v)]_{B_3} &= [g(f(v))]_{B_3} = A_{g,B_2,B_3}[f(v)]_{B_2} \\ &= A_{g,B_2,B_3}(A_{f,B_1,B_2}[v]_{B_1}) = (A_{g,B_2,B_3}A_{f,B_1,B_2})[v]_{B_1}\end{aligned}$$

donde, pela unicidade da matriz que representa $g \circ f$ se conclui que

$$A_{g \circ f,B_1,B_3} = A_{g,B_2,B_3}A_{f,B_1,B_2}$$

conforme pretendido. □

A proposição anterior "explica"a associatividade do produto de matrizes: o produto de matrizes é a tradução através dos isomorfismos do Exemplo 4.17(iii) da composição de funções, que é uma operação associativa.

Observação 4.19. *É possível pensar visualmente na correspondência entre transformações lineares e matrizes, e em particular na Proposição anterior da seguinte forma. Considere-se o diagrama*

(21)
$$\begin{tikzcd} V \arrow[r, "f"] \arrow[d, "{[\cdot]_{B_1}}"', "\cong"] & W \arrow[d, "{[\cdot]_{B_2}}"', "\cong"] \\ M_{n\times 1}(\mathbb{R}) \arrow[r, "A_{f,B_1,B_2}"] & M_{m\times 1}(\mathbb{R}) \end{tikzcd}$$

onde as setas representam transformações lineares com domínio a origem da seta e conjunto de chegada o término da seta. As setas pretendem representar visualmente que os vetores do espaço da origem são "transportados"pela transformação linear do seu domínio até ao espaço vetorial de chegada. O símbolo $\cong$ *designa isomorfismo e os isomorfismos no diagrama acima são os do Exemplo 4.17(ii) que calculam a matriz coluna das coordenadas, ou seja,* $v \mapsto [v]_{B_1}$ *para a seta da esquerda e* $w \mapsto [w]_{B_2}$ *para a seta da direita. A equação*

(22)
$$[f(v)]_{B_2} = A_{f,B_1,B_2}[v]_{B_1}$$

diz que se obtém o mesmo resultado quando se faz um vetor $v \in V$ *seguir os dois possíveis trajetos do canto superior esquerdo até ao canto inferior direito em* (21)*: do lado esquerdo de* (22) *temos o efeito de seguir primeiro a seta de cima e depois a seta da direita; do lado direito de* (22) *segue-se primeiro a seta da esquerda e depois a de baixo.*

Quando independentemente do caminho seguido entre dois nós do diagrama se obtém sempre o mesmo resultado diz-se que o diagrama é comutativo. *Portanto a equação* (22) *traduz a comutatividade de* (21).

Nestes termos, a Proposição 4.18 traduz a comutatividade do retângulo exterior no seguinte diagrama

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[d, "{[\cdot]_{B_1}}"', "\cong"] & W \arrow[r, "g"] \arrow[d, "{[\cdot]_{B_2}}"', "\cong"] & U \arrow[d, "{[\cdot]_{B_3}}"', "\cong"] \\ M_{n\times 1}(\mathbb{R}) \arrow[r, "A_{f,B_1,B_2}"] & M_{m\times 1}(\mathbb{R}) \arrow[r, "A_{g,B_2,B_3}"] & M_{p\times 1}(\mathbb{R}) \end{tikzcd}$$

que é claramente uma consequência da comutatividade dos dois quadrados.

Corolário 4.20. *Sejam V, W espaços vetoriais, $f\colon V \to W$ uma transformação linear invertível e B_1, B_2 bases para V e W respetivamente. Então $A_{f^{-1},B_2,B_1} = (A_{f,B_1,B_2})^{-1}$.*

Dem. Uma vez que $f \circ f^{-1} = \mathrm{Id}_W$ e $f^{-1} \circ f = \mathrm{Id}_V$, e que a matriz que representa a transformação linear identidade com respeito a uma mesma base num espaço vetorial é a matriz identidade, pela Proposição anterior temos

$$A_{f,B_1,B_2} A_{f^{-1},B_2,B_1} = I \qquad A_{f^{-1},B_2,B_1} A_{f,B_1,B_2} = I$$

(onde I designa a matriz identidade). □

Para terminar, registamos a interação das operações de composição de transformações lineares com as de soma e produto por escalar.

Proposição 4.21. *Sejam V, W, U espaços vetoriais, $f, g \in L(V, W), h, k \in L(W, U)$ e $\alpha \in \mathbb{K}$. Então*

(i) $h \circ (f + g) = h \circ f + h \circ g$ *e* $(h + k) \circ f = h \circ f + k \circ f$
(ii) $h \circ (\alpha f) = (\alpha h) \circ f = \alpha(h \circ f)$

Dem. Exercício. □

Estas propriedades traduzem-se na distributividade do produto de matrizes em relação à soma e na comutatividade do produto por escalar com o produto de matrizes. Notamos ainda que estas propriedades implicam que uma transformação linear $h\colon W \to U$ determina uma transformação linear

$$h_*\colon L(V, W) \to L(V, U)$$

definida pela expressão $h_*(\phi) = h \circ \phi$ e analogamente, $f\colon V \to W$ determina uma transformação linear

$$f^*\colon L(W, U) \to L(V, U)$$

definida por $f^*(\psi) = \psi \circ f$.

4.22. **Subespaços vetoriais associados a uma transformação linear.**

Definição 4.23. *Seja $f\colon V \to W$ uma transformação linear. O* núcleo de f *é o conjunto*

$$N(f) = \{v \in V\colon f(v) = 0\}$$

e a imagem de f *é o conjunto*

$$f(V) = \{f(v)\colon v \in V\} \subset W$$

Proposição 4.24. *Seja $f\colon V \to W$ uma transformação linear. Então $N(f)$ é um subespaço vetorial de V e $f(V)$ é um subespaço vetorial de W.*

Dem. Uma vez que $f(0) = 0$ temos que $0 \in N(f)$ e $0 \in f(V)$ pelo que estes conjuntos são não vazios. Vejamos que $N(f)$ é um subespaço vetorial:

- Sendo $v_1, v_2 \in N(f)$ temos $f(v_1+v_2) = f(v_1)+f(v_2) = 0+0 = 0$ logo $v_1+v_2 \in N(f)$.
- Sendo α um escalar e $v \in N(f)$ temos $f(\alpha v) = \alpha f(v) = \alpha 0 = 0$ logo $\alpha v \in N(f)$.

Quanto a $f(V)$:

- Dados $w_1, w_2 \in f(V)$, existem $v_1, v_2 \in V$ tais que $f(v_1) = w_1$ e $f(v_2) = w_2$. Então $f(v_1 + v_2) = w_1 + w_2$ logo $w_1 + w_2 \in f(V)$.
- Dado um escalar α e $w = f(v) \in f(V)$ temos $\alpha w = f(\alpha v) \in f(V)$.

$\square$

Por definição de sobrejetividade, uma transformação linear é sobrejetiva se e só se $f(V) = W$. A injetividade de f pode ser determinada em termos do núcleo como explica o seguinte resultado.

Proposição 4.25. *Uma transformação linear $f\colon V \to W$ é injetiva se e só se $N(f) = \{0\}$.*

Dem. Suponhamos que f é injetiva. Se $v \in N(f)$ então $f(v) = 0 = f(0)$. Uma vez que f é injetiva conclui-se que $v = 0$, logo $N(f) = \{0\}$.

Suponhamos agora que $N(f) = \{0\}$. Então se $f(v_1) = f(v_2)$ temos $f(v_1 - v_2) = 0$ e portanto $v_1 - v_2 \in N(f) = \{0\}$, ou seja, $v_1 = v_2$. $\square$

A Proposição anterior pode ser vista como mais uma manifestação do “bom comportamento”das transformações lineares. A condição $N(f) = \{0\}$ é equivalente (uma vez que $f(0) = 0$) à proposição

$$f(x) = f(0) \Rightarrow x = 0$$

que é um caso particular da condição geral de injetividade

$$f(x) = f(y) \Rightarrow x = y.$$

A Proposição 4.25 diz que, quando uma função é linear, para verificar a condição de injetividade podemos assumir que um dos elementos do domínio é 0. Se for verdade nesse caso particular então é verdade em geral.

Exemplo 4.26. *Seja $V = \{f \in F(\mathbb{R}, \mathbb{R})\colon f$ é diferenciável$\}$ o espaço vetorial das funções diferenciáveis e $T\colon V \to F(\mathbb{R}, \mathbb{R})$ a transformação linear definida por $T(f) = f'$.*

O núcleo de T é o subespaço das funções constantes de V, e tem portanto dimensão 1. A transformação T não é portanto injetiva. T também não é sobrejetiva. De facto a derivada de uma função tem necessariamente a seguinte propriedade do valor médio (cuja demonstração é um bom exercício de aplicação do Teorema de Rolle): se $f'(a) < 0$ e $f'(b) > 0$ então existe um real c entre a e b tal que $f'(c) = 0$.

Segue-se que, por exemplo, a função definida por

$$g(x) = \begin{cases} -1 & \text{se } x < 0 \\ 1 & \text{se } x \geq 0 \end{cases}$$

não é a derivada de qualquer função e não está portanto na imagem de T. É interessante notar que não foi encontrada até à data qualquer caracterização conveniente da imagem da transformação T.

É natural perguntar a que correspondem o núcleo e a imagem de uma transformação linear em termos de coordenadas, ou seja através do “dicionário”descrito no diagrama (21).

Quanto ao núcleo, temos
$$v \in N(f) \Leftrightarrow f(v) = 0 \Leftrightarrow [f(v)]_{B_2} = 0$$
uma vez que um vetor é nulo se e só se as suas coordenadas numa base são todas nulas. Por (22) isto acontece se e só se
$$A_{f,B_1,B_2}[v]_{B_1} = 0$$
ou seja, se o vetor de $\mathbb{R}^n$ formado pelas coordenadas de v pertence ao núcleo da matriz A_{f,B_1,B_2} que representa a transformação linear f. Assim, não muito surpreendentemente, em coordenadas, o núcleo de uma transformação linear corresponde ao núcleo da matriz que representa a transformação linear.

Quanto à imagem de f, a sua tradução em coordenadas é o conjunto
$$\{[f(v)]_{B_2} : v \in V\} \subset M_{m\times 1}(\mathbb{R})$$
Novamente por (22) temos que este conjunto é igual a
$$\{A_{f,B_1,B_2}[v]_{B_1} : v \in V\}$$
Mas sendo v um vector arbitrário de V, a sua matriz coluna de coordenadas é uma matriz arbitrária em $M_{n\times 1}(\mathbb{R})$ e portanto este conjunto não é mais do que o espaço das colunas da matriz A_{f,B_1,B_2}. Ou seja, em coordenadas, *a imagem de uma transformação linear f é o espaço das colunas da matriz que representa f.*

4.27. **O Teorema da característica-nulidade.** Chegamos agora a um dos resultados básicos da Álgebra Linear, cuja importância se irá tornando clara com o desenrolar do semestre.

Teorema 4.28. *Seja V um espaço vetorial finitamente gerado, W um espaço vetorial e $f : V \to W$ uma transformação linear. Então*
$$\dim N(f) + \dim f(V) = \dim V$$

Dem. Seja $\{v_1, \ldots, v_k\}$ uma base para o subespaço $N(f) \subset V$ (que é finitamente gerado porque V é). Pelo Corolário 3.29(ii) podemos completar este conjunto com um número finito de vetores distintos $\{v_{k+1} \ldots, v_n\}$ de tal forma que $\{v_1, \ldots, v_n\}$ seja uma base para V. Vamos verificar que $\{f(v_{k+1}), \ldots, f(v_n)\}$ é uma base de $f(V)$. Teremos então
$$\dim N(f) = k, \quad \dim f(V) = n - k, \quad \dim V = n$$
o que verifica a afirmação do enunciado. O caso em que $n = k$ verifica-se imediatamente pelo que vamos assumir a partir de agora que $n > k$.

- $\{f(v_{k+1}), \ldots, f(v_n)\}$ gera V: Seja w um vetor em $f(V)$. Então existe $v \in V$ tal que $f(v) = w$. Uma vez que $\{v_1, \ldots, v_n\}$ é uma base, existem escalares $\alpha_1, \ldots, \alpha_n$ tais que $v = \alpha_1 v_1 + \ldots + \alpha_n v_n$. Então

$$\begin{aligned} f(\alpha_1 v_1 + \ldots + \alpha_k v_k + \alpha_{k+1} v_{k+1} + \ldots + \alpha_n v_n) &= f(\alpha_1 v_1 + \ldots + \alpha_k v_k) + \alpha_{k+1} f(v_{k+1}) + \ldots + \alpha_n f(v_n) \\ &= 0 + \alpha_{k+1} f(v_{k+1}) + \ldots + \alpha_n f(v_n) \end{aligned}$$

onde na segunda igualdade usámos o facto de o vetor $\alpha_1 v_1 + \ldots + \alpha_k v_k$ pertencer ao núcleo de f. A expressão acima mostra que w é uma combinação linear de $f(v_{k+1}), \ldots, f(v_n)$ pelo que estes vetores geram $f(V)$.

- $\{f(v_{k+1}), \ldots, f(v_n)\}$ é linearmente independente: Suponhamos que $\beta_1, \ldots, \beta_{n-k}$ são escalares tais que
$$\beta_1 f(v_{k+1}) + \ldots + \beta_{n-k} f(v_n) = 0$$
Então $f(\beta_1 v_{k+1} + \ldots + \beta_{n-k} v_n) = 0$, logo $\beta_1 v_{k+1} + \ldots + \beta_{n-k} v_n \in N(f)$. Portanto existem escalares $\alpha_1, \ldots, \alpha_k$ tais que $\alpha_1 v_1 + \ldots + \alpha_k v_k = \beta_1 v_{k+1} + \ldots + \beta_{n-k} v_n$ ou seja tais que
$$\alpha_1 v_1 + \ldots + \alpha_k v_k - \beta_1 v_{k+1} - \ldots - \beta_{n-k} v_n = 0$$
Uma vez que $\{v_1, \ldots, v_n\}$ é uma base de V tal só pode acontecer se $\alpha_1 = \ldots = \alpha_k = -\beta_1 = \ldots = -\beta_{n-k} = 0$. Conclui-se que $\beta_1 = \cdots = \beta_{n-k} = 0$ e portanto que $\{f(v_{k+1}), \ldots, f(v_n)\}$ é linearmente independente.

$\square$

Exemplo 4.29. *Vamos determinar uma base para o núcleo e imagem da transformação linear $f \colon M_{2\times 2}(\mathbb{R}) \to M_{3\times 2}(\mathbb{R})$ definida por*
$$f\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} 1 & -2 \\ 3 & -6 \\ 2 & -4 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$
Temos
$$f\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = 0 \Leftrightarrow \begin{bmatrix} a-2c & b-2d \\ 3a-6c & 3b-6d \\ 2a-4c & 2b-4d \end{bmatrix} = 0 \Leftrightarrow \begin{cases} a = 2c \\ b = 2d \end{cases}$$
logo
$$N(f) = \left\{ \begin{bmatrix} 2c & 2d \\ c & d \end{bmatrix} : b, d \in \mathbb{R} \right\} = L\left(\left\{ \begin{bmatrix} 2 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 2 \\ 0 & 1 \end{bmatrix} \right\}\right)$$
Conclui-se que $\left\{ \begin{bmatrix} 2 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 2 \\ 0 & 1 \end{bmatrix} \right\}$ *é uma base para $N(f)$. Pelo Teorema 4.28*
$$\dim(\operatorname{Im}(f)) = \dim M_{2\times 2}(\mathbb{R}) - \dim N(f) = 4 - 2 = 2$$
Podemos obter uma base para $\operatorname{Im}(f)$ completando uma base de $N(f)$ com dois elementos de forma a obter uma base de $M_{2\times 2}(\mathbb{R})$ e calculando a imagem por f desses dois novos elementos. Claramente
$$\left\{ \begin{bmatrix} 2 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 2 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \right\}$$
é uma base de $M_{2\times 2}(\mathbb{R})$, pelo que
$$\left\{ f\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right), f\left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right) \right\} = \left\{ \begin{bmatrix} 1 & 0 \\ 3 & 0 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 3 \\ 0 & 2 \end{bmatrix} \right\}$$

é uma base da imagem de f.

Definição 4.30. *Sendo V um espaço finitamente gerado, W um espaço vetorial e $f\colon V \to W$ uma transformação linear, o número* $\dim f(V)$ *chama-se a* característica *da transformação linear f (rank em inglês) e o número* $\dim N(f)$ *chama-se a* nulidade *de f (nullity em inglês).*

O Teorema 4.28 é conhecido em inglês por "the rank-nullity Theorem". Por palavras, afirma que a soma da característica de uma transformação linear com a sua nulidade é igual à dimensão do espaço de partida. Tem o seguinte corolário extremamente útil:

Corolário 4.31. *Sejam V e W espaços vetoriais finitamente gerados* com a mesma dimensão *e seja $f\colon V \to W$ uma transformação linear. Então as seguintes afirmações são equivalentes:*

(i) f é invertível (isto é, f é bijetiva).
(ii) f é injetiva (equivalentemente, $N(f) = \{0\}$).
(iii) f é sobrejetiva (isto é, $f(V) = W$).

Dem. É claro que a afirmação (i) implica as afirmações (ii) e (iii), e, por definição (ii) juntamente com (iii) implicam (i). Para demonstrar a equivalência das afirmações basta assim ver que quando (ii) se verifica, (iii) também se verifica e vice-versa.

Suponhamos que f é injetiva. Então $\dim N(f) = 0$ e portanto pelo Teorema 4.28 e a hipótese sobre a dimensão dos espaços V e W temos

$$\dim f(V) = \dim V = \dim W$$

Ou seja $f(V)$ é um subespaço de W com a mesma dimensão que W. Então temos necessariamente $f(V) = W$ (por exemplo, pelo Corolário 3.34(i)) e portanto f é também sobrejetiva.

Suponhamos agora que f é sobrejetiva, ou seja que $\dim f(V) = \dim W$. Aplicando o Teorema 4.28 e a hipótese $\dim V = \dim W$ temos

$$\dim f(V) + \dim N(f) = \dim V \Leftrightarrow \dim V + \dim N(f) = \dim V \Leftrightarrow \dim N(f) = 0$$

logo $N(f) = \{0\}$ e portanto, pela Proposição 4.25, f é injetiva. $\square$

4.32. **Aplicações ao estudo das matrizes.** Tendo em conta a interpretação da imagem de uma transformação linear f como o espaço das colunas da matriz que a representa, o Teorema 4.28 tem a seguinte consequência importante (que está longe de ser óbvia!).[17]

Proposição 4.33. *Seja A uma matriz $m \times n$. Então o espaço das linhas e o espaço das colunas de A têm a mesma dimensão (que é a característica de A). Isto é,*

$$\dim EC(A) = \dim EL(A) = \textit{característica de } A$$

[17]Para uma explicação conceptual desta igualdade que é independente da nossa discussão inicial dos sistemas lineares e do método de Gauss ver o Exercício 34(c).

Demonstração. A dimensão do espaço das linhas é o número de pivots da matriz A após aplicação do método de Gauss, enquanto que a dimensão do núcleo de A é o número de variáveis livres no sistema homogéneo associado a A, ou seja, o número de colunas de A sem pivot. Isto significa que

$$\dim EL(A) = n - \dim N(A)$$

Por outro lado, no caso da transformação linear $f\colon M_{n\times 1}(\mathbb{R}) \to M_{m\times 1}(\mathbb{R})$ definida por $f(x) = Ax$, o Teorema 4.28 diz que

$$\dim EC(A) + \dim N(A) = n \Leftrightarrow \dim EC(A) = n - \dim N(A)$$

Conclui-se portanto que $\dim EC(A) = \dim EL(A)$ e este número é a característica de A. □

A Proposição anterior justifica também a atribuição do nome "característica"de f à dimensão de $f(V)$.

Observação 4.34. *Também podemos deduzir a Proposição 4.33 da seguinte forma. Ela é claramente verdadeira se A é uma matriz em escada de linhas visto que as dimensões de $EC(A)$ e $EL(A)$ são nesse caso ambas iguais ao número de pivots. Ora, embora o espaço das colunas se vá alterando quando aplicamos o método de Gauss, a sua* dimensão *não se altera!*

De facto, duas matrizes A_i e A_{i+1} que estejam relacionadas por uma operação elementar sobre as linhas satisfazem $A_{i+1} = SA_i$ para alguma matriz invertível S. Sendo m o número de linhas das matrizes, a multiplicação à esquerda por S dá-nos um isomorfismo $\phi_S\colon \mathbb{R}^m \to \mathbb{R}^m$. Pela definição do produto matricial, ϕ_S leva a j-ésima coluna de A_i na j-ésima coluna de A_{i+1} e, em particular, transforma $EC(A_i)$ em $EC(A_{i+1})$. Uma vez que ϕ_S é um isomorfismo, a sua restrição a $EC(A_i)$ é ainda um isomorfismo e portanto temos

$$\dim EC(A_i) = \dim \phi_S(EC(A_i)) = \dim EC(A_{i+1})$$

Conclui-se que a dimensão do espaço das colunas não se altera ao longo do método de Gauss. Mais, uma vez que as colunas respetivas da matriz inicial e da matriz em escada de linhas no final do método de Gauss estão relacionadas por um isomorfismo, as colunas da matriz inicial correspondentes às colunas com pivot no final do método de Gauss formam uma base para $EC(A)$.

Exemplo 4.35. *Consideremos a matriz*

$$A = \begin{bmatrix} 1 & 2 & -1 & 2 \\ 2 & 4 & 1 & 3 \\ 1 & 2 & 2 & 1 \end{bmatrix}$$

Aplicando o método de Gauss obtemos

$$\begin{bmatrix} 1 & 2 & -1 & 2 \\ 2 & 4 & 1 & 3 \\ 1 & 2 & 2 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & -1 & 2 \\ 0 & 0 & 3 & -1 \\ 0 & 0 & 3 & -1 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & -1 & 2 \\ 0 & 0 & 3 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Uma vez que os pivots ocorrem nas colunas 1 e 3, temos que uma base para o $EC(A)$ *é dado por* $\{(1,2,1),(-1,1,2)\}$ *enquanto que uma base para* $EL(A)$ *é dado por* $\{(1,2,-1,2),(0,0,3,-1)\}$.

Podemos agora aproveitar para atualizar os nossos critérios para a invertibilidade de uma matriz (comparem com o Teorema 2.27)

Proposição 4.36. *Seja* A *uma matriz* $n \times n$*. Então as seguintes afirmações são equivalentes*

(i) A *é invertível.*
(ii) A *característica de* A *é* n *(equivalentemente* $\dim EL(A) = n$*).*
(iii) Para cada matriz $b \in M_{n\times 1}(\mathbb{R})$ *a equação* $Ax = b$ *tem solução única (equivalentemente, a função* $x \mapsto Ax$ *é bijetiva).*
(iv) $N(A) = 0$
(v) $EC(A) = \mathbb{R}^n$
(vi) Existe $B \in M_{n\times n}(\mathbb{R})$ *tal que* $AB = I_n$
(vii) Existe $B \in M_{n\times n}(\mathbb{R})$ *tal que* $BA = I_n$

Dem. A equivalência das primeiras três afirmações foi já vista no Teorema 2.27 embora a equivalência de (i) com (iii) possa agora ser interpretada conceptualmente como uma consequência da Proposição 4.14 e Corolário 4.20. A equivalência de (iii), (iv) e (v) é uma consequência do Corolário 4.31 e da interpretação do núcleo e espaço das colunas da matriz como núcleo e imagem da transformação linear associada.

É claro da definição de invertibilidade que $(i) \Rightarrow (vi)$ e (vii). Reciprocamente se existe B tal que $AB = I_n$ então o espaço das colunas de A contém as colunas da matriz identidade, e portanto $EC(A) = \mathbb{R}^n$, que é a condição (v). Por outro lado se existe B tal que $BA = I_n$ então dado $x \in N(A)$ temos $x = I_n x = BAx = B0 = 0$ pelo que $N(A) = \{0\}$ que é a condição (iv). Vemos assim que (vi) e (vii) são também equivalentes às restantes condições. □

4.37. **Equações lineares.**

Definição 4.38. *Uma* equação linear *é uma equação da forma*

$$f(x) = w$$

onde $f\colon V \to W$ *é uma transformação linear,* w *é um vetor de* W *e a incógnita* x *é um vetor de* V *a determinar. A equação diz-se* homogénea *quando* $w = 0$.

É claro que uma equação linear tem solução se e só e $w \in f(V)$. O conjunto das soluções é controlado pelo núcleo de f no seguinte sentido.

Proposição 4.39 (Princípio da sobreposição)**.** *Seja* $f\colon V \to W$ *uma transformação linear. Se* v *é uma solução da equação linear* $f(x) = w$*, o conjunto de todas as soluções é*

$$v + N(f) = \{v + z \colon z \in N(f)\} \subset V$$

Dem. Se v é uma solução e $z \in N(f)$ temos que $f(v+z) = f(v) + f(z) = w + 0 = w$ logo $v + z$ é uma solução. Assim

$$v + N(f) \subset \{u \in V \colon f(u) = w\}$$

Reciprocamente, seja u uma solução qualquer da equação. Então $u = v + (u - v)$ e $f(u - v) = f(u) - f(v) = w - w = 0$ pelo que $u - v \in N(f)$ e portanto $u \in v + N(f)$. Conclui-se que

$$\{u \in V : f(u) = w\} \subset v + N(f)$$

o que termina a demonstração. □

Geometricamente, o resultado anterior diz que o conjunto das soluções é o "plano"paralelo a $N(f)$ (que é um "plano"em V contendo a origem) que passa por uma solução particular qualquer da equação.

É costume enunciar o resultado da Proposição 4.39 da seguinte forma;

> A solução geral de uma equação linear é obtida somando a uma solução particular da equação a solução geral da equação homogénea.

Por uma solução particular entende-se uma qualquer solução v fixada para a equação. Por solução geral entende-se o conjunto das soluções. Assim a afirmação acima diz apenas que o conjunto das soluções de uma equação linear é obtido somando todas as soluções da equação homogénea a uma qualquer solução da equação que consigamos determinar.

Exemplo 4.40 (O oscilador harmónico)**.** *Seja $x : \mathbb{R} \to \mathbb{R}$ uma função que descreve a posição de uma partícula presa a uma mola em função do tempo. A partícula é atuada unicamente pela força determinada pela extensão ou contração da mola, que é proporcional ao deslocamento da mola em relação à sua posição de repouso. Assumindo que* 0 *é a coordenada da posição de repouso, a equação de Newton diz-nos que*

$$x''(t) + kx(t) = 0 \tag{23}$$

onde k é uma constante positiva determinada pelas características físicas da mola e a massa da partícula (recorde que x'' é a aceleração e note que a força exercida pela mola, mx'' tem o sentido contrário ao deslocamento x). Para simplificar as contas vamos assumir a partir de agora que $k = 1$.

Sendo $V \subset F(\mathbb{R}, \mathbb{R})$ o subespaço vetorial formado pelas funções duas vezes diferenciáveis e T a transformação linear

$$T : V \to F(\mathbb{R}, \mathbb{R})$$

definida pela expressão

$$T(x) = x'' + x$$

vemos que o núcleo de T é exatamente o conjunto das soluções da equação diferencial (23) *(com $k = 1$) que formam portanto um subespaço vetorial de V.*

É fácil adivinhar duas soluções para a equação

$$x'' + x = 0 \tag{24}$$

pois claramente $x(t) = \cos t$ e $x(t) = \operatorname{sen} t$ são soluções. Como o conjunto das soluções é um espaço vetorial temos mais geralmente que

$$x(t) = \alpha_1 \cos t + \alpha_2 \operatorname{sen} t, \qquad \text{com } \alpha_1, \alpha_2 \in \mathbb{R} \tag{25}$$

são soluções.

Para o ano que vem irão aprender que uma solução de uma equação diferencial como (23) *é completamente determinada por* $x(0)$ *e* $x'(0)$ *(fisicamente isto diz que a evolução da posição da partícula é completamente determinada pela sua posição e velocidade iniciais). Assim o conjunto das soluções é um espaço vetorial de dimensão 2 (um vetor é determinado por dois números reais) e portanto a fórmula* (25) *descreve a solução geral da equação* (24).

No caso da equação (24) *podemos verificar a afirmação anterior diretamente recorrendo à* conservação da energia. *Definindo a quantidade*

$$E(t) = (x')^2 + x^2$$

(correspondendo à soma das energia cinética e potencial) temos

$$\frac{dE}{dt} = 2x'x'' + 2xx' = 2x'(-x) + 2xx' = 0$$

logo a quantidade $(x')^2 + x^2$ *é conservada ao longo do tempo para qualquer solução da equação diferencial* (24). *Em particular se* $x(t)$ *for uma solução com* $x(0) = x'(0) = 0$ *teremos* $(x'(t))^2 + x(t)^2 = 0$ *para todo o* t *e portanto* $x(t) = 0$.

Isto permite-nos concluir que os valores de $x(0)$ *e* $x'(0)$ *determinam completamente a solução* $x(t)$ *para todo o* t*: se* $x(t)$ *e* $y(t)$ *forem soluções de* (24) *com* $x(0) = y(0)$ *e* $x'(0) = y'(0)$ *então* $u(t) = x(t) - y(t)$ *é também uma solução de* (24) *(porque se trata de uma equação linear!) que satisfaz* $u(0) = u'(0) = 0$. *Mas então* $u(t) = 0$ *e portanto* $x(t) = y(t)$.

É agora imediato verificar que as soluções (25) *permitem atribuir valores arbitrários a* $x(0)$ *e* $x'(0)$ *mediante variação dos coeficientes* α_1 *e* α_2 *(na realidade* $\alpha_1 = x(0)$ *e* $\alpha_2 = x'(0)$*) e portanto descrevem todas as soluções de* (24).

Suponhamos agora que queremos resolver a equação[18]

$$x'' + x = t^3 \tag{26}$$

Trata-se agora de uma equação linear não homogénea. Não é no entanto difícil descobrir uma solução particular desta equação tentando encontrar um polinómio que a satisfaça. Se o fizer irá ver que o único polinómio que satisfaz esta equação é

$$x(t) = t^3 - 6t$$

A Proposição 4.39 diz-nos então que a solução geral da equação (26) *é*

$$x(t) = t^3 - 6t + \alpha_1 \cos t + \alpha_2 \operatorname{sen} t, \qquad \text{com } \alpha_1, \alpha_2 \in \mathbb{R}.$$

4.41. **Exercícios.**

1. Indique justificando se as seguintes aplicações são transformações lineares
   (i) $f\colon \mathbb{R}^3 \to \mathbb{R}^4$ definida por $f(x, y, z) = (x - 2y, 0, x + y - 3z, 0)$.
   (ii) $f\colon M_{m\times n}(\mathbb{R}) \to M_{n\times m}(\mathbb{R})$ definida por $f(A) = A^T$.
   (iii) $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ definida por $f(x, y) = (x^2 y, x - y + 1)$.
   (iv) $f\colon M_{2\times 2}(\mathbb{R}) \to \mathbb{R}$ definida por $f\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = a + b$.

[18] Fisicamente esta equação corresponde a adicionar ao sistema mecânico considerado anteriormente uma força exterior dependente do tempo que actua com intensidade t^3/m (onde m é a massa da partícula).

(v) $T\colon \mathbb{R}^3 \to \mathbb{R}^2$ definida por $T(x, y, z) = (1, 3)$.
(vi) $f\colon M_{3\times 3}(\mathbb{R}) \to M_{3\times 3}(\mathbb{R})$ definida por $T(A) = A^2 + A$.
(vii) $T\colon F(\mathbb{R}, \mathbb{R}) \to \mathbb{R}^2$ definida por $T(f) = (f(0), f(2))$.
(viii) Sendo V um espaço vetorial e α um escalar, a aplicação $T\colon V \to V$ definida por $T(v) = \alpha v$.

Resolução.

2. Seja $T\colon \mathbb{R}^2 \to \mathbb{R}^3$ uma transformação linear tal que $T(1, 1) = (2, 0, 1)$ e $T(1, 2) = (1, 4, 6)$.
   (a) Calcule $T(2, 5)$.
   (b) Descreva a imagem por T do triângulo com vértices $(0, 1)$, $(1, 2)$ e $(0, 0)$.
   (c) Determine a matriz que representa T com respeito às bases canónicas de $\mathbb{R}^2$ e $\mathbb{R}^3$ e escreva a expressão para $T(x, y) \in \mathbb{R}^3$.

   Resolução.
3. Determine uma transformação linear $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ tal que
   (a) $f([0, 2] \times [0, 1])$ (isto é, a imagem por f do retângulo
   $$[0, 2] \times [0, 1] = \{(x, y) \in \mathbb{R}^2\colon 0 \leq x \leq 2 \text{ e } 0 \leq y \leq 1\})$$
   seja o quadrado com vértices $(0, 0), (1, 2), (-2, 1)$ e $(-1, 3)$.
   (b) A imagem por f do triângulo com vértices $(0, 0), (1, 0), (0, 1)$ seja o segmento de reta $\{(x, x)\colon x \in [0, 2]$.

   Resolução.
4. Seja V um espaço vetorial. Existe alguma transformação linear $f\colon \mathbb{R}^2 \to V$ satisfazendo as seguintes condições?
   - O conjunto $\{f(1, 0), f(0, 1)\} \subset V$ é linearmente independente.
   - $f(1, 3) = 2f(1, 1) + f(0, 3)$

   E se a primeira condição for omitida? Justifique.Resolução.
5. Seja V o espaço dos polinómios de grau ≤ 3 e considere a função $f\colon V \to V$ definida por
   $$f(p(t)) = p''(t) - 2p(t)$$
   (a) Verifique que f é uma transformação linear.
   (b) Determine a matriz que representa f com respeito à base canónica no espaço dos polinómios.

   Resolução.
6. Determine a matriz que representa a transformação linear $T : V \to W$ indicada com respeito às bases ordenadas B_1 para V e B_2 para W indicadas.
   (a) $V = \mathbb{R}^3, W = \mathbb{R}$ com as bases canónicas, $T(x, y, z) = x - 2y + 6z$.
   (b) $V = \mathbb{R}^2, W = \mathbb{R}^4$, $B_1 = ((1, 1), (0, 1))$ e B_2 a base canónica, $T(x, y) = (x-2y, 0, x+y, -x)$.
   (c) V o espaço dos polinómios de grau ≤ 2, W o espaço dos polinómios de grau ≤ 4, B_1 e B_2 as bases canónicas e T a transformação linear definida por $T(p(t)) = p(t)(1-t^2)$

(d) $V = M_{2\times 2}(\mathbb{R})$, $W = M_{3\times 1}(\mathbb{R})$ com as bases canónicas e

$$T(A) = \begin{bmatrix} 1 & 1 \\ -1 & 0 \\ 3 & 2 \end{bmatrix} A \begin{bmatrix} 4 \\ -2 \end{bmatrix}$$

Resolução.

7. Seja P o espaço vetorial dos polinómios de grau menor ou igual a 2 e $T\colon M_{2\times 2}(\mathbb{R}) \to P$ uma transformação linear tal que

$$T\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right) = 1+t, T\left(\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}\right) = t-2t^2, T\left(\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}\right) = -3t^2, T\left(\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right) = 4.$$

(a) Calcule $T\left(\begin{bmatrix} 3 & 1 \\ 1 & 0 \end{bmatrix}\right)$.

(b) Determine a matriz que representa T nas bases canónicas para as matrizes e polinómios.

Resolução.

8. Seja $V = L(\{e^{-2x}, e^{-x}, 1, e^{x}, e^{2x}) \subset F(\mathbb{R},\mathbb{R})$ e considere a transformação linear $T\colon V \to V$ definida por $T(f) = f'' + f' - 2f$

(a) Verifique que T está bem definida, isto é que $T(f) \in V$ quando $f \in V$.

(b) Determine a matriz que representa T com respeito a uma base para V por si escolhida.

(c) Determine $T^{-1}(0) = \{f \in V : T(f) = 0\}$.

(d) Determine $T^{-1}(e^{-x}) = \{f \in V : T(f) = e^{-x}\}$.

Note que nas duas últimas alíneas resolveu equações diferenciais! Resolução.

9. Seja V um espaço vetorial e $B = (v_1, \ldots, v_n)$ uma base ordenada para V. Mostre que a função $f\colon V \to M_{n\times 1}(\mathbb{R})$ definida por $f(v) = [v]_B$ (que associa a um vetor v a matriz coluna que tem por entradas as suas coordenadas na base B) é uma transformação linear. Resolução.

10. Sejam V e W espaços vetoriais com bases ordenadas B_1 e B_2 respetivamente, e $T\colon V \to W$ uma transformação linear. Se a matriz que representa T nas bases B_1 e B_2 é

$$\begin{bmatrix} 1 & 1 & 0 \\ -1 & 3 & 2 \end{bmatrix}$$

e temos matrizes de mudanças de coordenadas

$$S_{B_1\to B_1'} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}, \qquad S_{B_2\to B_2'} = \begin{bmatrix} 1 & -2 \\ 3 & 1 \end{bmatrix}$$

(a) Quais são as dimensões dos espaços vetoriais V e W?

(b) Determine a matriz que representa a transformação linear T nas bases B_1' e B_2'.

Resolução.

11. Seja $T\colon V \to W$ uma transformação linear e $S \subset V$ um subconjunto. Se S for linearmente dependente, $T(S)$ é um conjunto linearmente dependente em W? Se S for

linearmente independente, $T(S)$ é necessariamente linearmente independente? Justifique. Resolução.

12. Determine uma base para o espaço vetorial $L(\mathbb{R}^2, \mathbb{R}^3)$ das transformações lineares de $\mathbb{R}^2$ para $\mathbb{R}^3$. Resolução.
13. Seja V um espaço vetorial real. O espaço vetorial das transformações lineares de V para $\mathbb{R}$ chama-se o *dual* de V e denota-se por $V^* = L(V, \mathbb{R})$.
    (a) Sendo V o espaço dos polinómios de grau ≤ 2, mostre que os vetores $\phi_0, \phi_1, \phi_2 \in V^*$ definidos por
    $$\phi_0(p) = p(0), \quad \phi_1(p) = p(1), \quad \phi_2(p) = p(2)$$
    formam uma base para V^*.
    (b) Determine as coordenadas do vetor ψ definido por $\psi(p) = p'(1)$ na base ordenada (ϕ_0, ϕ_1, ϕ_2).

    Resolução.
14. Seja $f\colon V \to W$ um isomorfismo de espaços vetoriais.
    (a) Mostre que um conjunto $S \subset V$ é linearmente (in)dependente se e só se $f(S)$ é linearmente (in)dependente em W.
    (b) Mostre que um conjunto $S \subset V$ gera V se e só se $f(S)$ gera W.
    (c) Mostre que V é finitamente gerado se e só se W é e nesse caso V e W têm a mesma dimensão.
    (d) Mostre que $U \subset V$ é um subespaço vetorial se e só se $f(U)$ é um subespaço vetorial e nesse caso a restrição de f a U é um isomorfismo entre U e $f(U)$.

    Resolução.
15. O espaço vetorial livremente gerado por um conjunto.
    (a) Seja S um conjunto não vazio. Recorde que $F(S, \mathbb{R})$ denota o espaço vetorial das funções de S para $\mathbb{R}$ com a soma e produto por escalar definidos da forma usual. O *suporte* de $f \in F(S, \mathbb{R})$ é o conjunto
    $$\operatorname{supp}(f) = \{x \in S\colon f(x) \neq 0\} \subset S$$
    O espaço vetorial sobre $\mathbb{R}$ livremente gerado por S é o subespaço
    $$\mathbb{R} \cdot S = \{f \in F(S, \mathbb{R})\colon \operatorname{supp}(f) \text{ é finito}\} \subset F(S, \mathbb{R})$$
    Mostre que se trata realmente de um subespaço de $F(S, \mathbb{R})$.
    (b) Para cada $x \in S$, seja $\phi_x\colon S \to \mathbb{R}$ a função definida por
    $$\phi_x(y) = \begin{cases} 1 & \text{se } y = x \\ 0 & \text{caso contrário.} \end{cases}$$
    Mostre que $\{\phi_x\colon x \in S\}$ é uma base de $\mathbb{R} \cdot S$. Um elemento $f \in \mathbb{R} \cdot S$ é visto como uma combinação linear formal dos elementos do suporte de f. A combinação linear $\sum_{x \in S} \alpha_x \phi_x$ (em que apenas finitos coeficientes são não nulos) é normalmente escrita $\sum_{x \in S} \alpha_x \cdot x$.
    (c) Seja V um espaço vetorial sobre $\mathbb{R}$ e B uma base de V. Defina um isomorfismo natural $\mathbb{R} \cdot B \to V$.

(d) Sendo V um espaço vetorial sobre $\mathbb{R}$ estabeleça um isomorfismo natural entre $L(\mathbb{R} \cdot S, V)$ e $F(S, V)$.

Resolução.

16. Para as seguintes transformações lineares, determine se são injetivas ou sobrejetivas, ache uma base para o núcleo e para a imagem da transformação linear, e caso sejam invertíveis determine uma expressão para a transformação inversa.
 (a) $f\colon \mathbb{R}^2 \to \mathbb{R}^3$ definida por $f(x, y) = (x + 2y, 0, -2x - 4y)$.
 (b) $f\colon \mathbb{R}^4 \to \mathbb{R}^2$ definida por $f(x, y, z, w) = (x + y, z + w)$.
 (c) $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ definida por $f(x, y, z) = (x + z, y - z, 2x + 2y)$.
 (d) $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ definida por $f(x, y, z) = (y, x + z, x - y)$.

Resolução.

17. Indique justificadamente quais das seguintes transformações lineares são isomorfismos.
 (a) $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ definida por $f(x, y, z) = (x + y, -y + z, x + 2y + 3z)$.
 (b) $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ definida por $f(x, y, z) = (x + y, -y + z, x + z)$.
 (c) $f\colon \mathbb{R}^5 \to \mathbb{R}^5$ definida por $f(x, y, z, u, v) = (x - y + z + u + v, -x + 2y + u, z + u - 2v, x + 4y + u, -2y + 3z - 2u - v)$
 (d) $f\colon \{p(t)\colon p \text{ polinómio real de grau } \leq 3\} \to M_{2\times 2}(\mathbb{R})$ definida por

$$f(p) = \begin{bmatrix} p(0) & p'(0) \\ p''(0) + p(1) & p'''(1) \end{bmatrix}$$

 (e) Uma transformação linear representada por uma matriz 4×4 cuja terceira coluna é a soma das duas primeiras colunas.
 (f) A transformação linear do espaço de todos os polinómios reais nele próprio definida por $T(p) = p'$.
 (g) A transformação linear do espaço de todos os polinómios reais nele próprio definida por $T(p) = p' - p$.

Resolução.

18. Sejam V, W espaços vetoriais finitamente gerados e B_1, B_2 bases para V, W respetivamente. Sejam $f, g\colon V \to V$ e $h\colon V \to W$ transformações lineares com representações matriciais

$$A_{f,B_1,B_1} = \begin{bmatrix} 1 & 1 \\ -2 & 1 \end{bmatrix} \quad A_{g,B_1,B_1} = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \quad A_{h,B_1,B_2} = \begin{bmatrix} 2 & 1 \\ 3 & 0 \\ -1 & 1 \end{bmatrix}$$

 (a) Quais são as dimensões dos espaços V, W?
 (b) Quais das transformações lineares são isomorfismos?
 (c) Determine as representações matriciais de $h \circ (f + 3(g \circ g))$ e de $(f + f^{-1}) \circ g$.

Resolução.

19. Seja V o espaço vetorial dos polinómios de grau ≤ 2 e considere a base $B_1 = (1 - t^2, t, t + t^2)$ para este espaço. Seja $f\colon V \to \mathbb{R}^2$ a transformação linear com representação matricial

$$A_{f,B_1,B_{can}} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & -1 \end{bmatrix}$$

e $g\colon \mathbb{R}^2 \to M_{2\times 2}(\mathbb{R})$ a transformação linear determinada por

$$g(1,1) = \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix} \qquad g(2,1) = \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix}$$

(a) Calcule $f(1+t)$.
(b) Sendo $B_{can} = (1, t, t^2)$ a base canónica dos polinómios, determine $S_{B_{can}\to B_1}$ e aproveite para calcular $A_{f,B_{can},B_{can}}$.
(c) Calcule $A_{g\circ f,B_1,B_{can}}$ onde tomamos para base canónica das matrizes $\left(\left[\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right], \left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right], \left[\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right], \left[\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right]\right)$.

Resolução.

20. Seja V o espaço dos polinómios de grau ≤ 2. Seja $f\colon \mathbb{R}^2 \to V$ a transformação linear tal que

$$f(1,2) = 1 - t^2, \quad f(1,1) = 2 + t$$

e $g\colon V \to M_{2\times 2}(\mathbb{R})$ a transformação linear definida por

$$g(p) = \begin{bmatrix} 0 & p(1) \\ p(0) & p'(0) \end{bmatrix}$$

(a) Determine $(g \circ f)(1,0)$.
(b) Determine a expressão geral para $(g \circ f)(x,y)$.
(c) Determine a matriz que representa $g \circ f$ com respeito às base $B_1 = ((1,1),(1,-1))$ de $\mathbb{R}^2$ e a base canónica de $M_{2\times 2}(\mathbb{R})$.
(d) Determine uma base para o espaço $W = (g \circ f)(\mathbb{R}^2)$.
(e) Determine a matriz que representa $(g\circ f)^{-1}\colon W \to \mathbb{R}^2$ com respeito à base da alínea anterior e a base canónica em $\mathbb{R}^2$.

Resolução.

21. Sendo $f\colon V \to W$ uma transformação linear, quais são os possíveis valores que $(\dim N(f), \dim f(V))$ pode tomar quando f
(a) aplica $\mathbb{R}^7$ em $\mathbb{R}^4$.
(b) aplica $\mathbb{R}^3$ em $\mathbb{R}^5$.

Resolução.

22. Dê um exemplo de uma transformação linear $f\colon V \to V$ tal que $N(f) = f(V)$. *Sugestão: Pode tomar* $V = \mathbb{R}^2$. Resolução.
23. Seja $f\colon M_{2\times 2}(\mathbb{R}) \to M_{2\times 2}(\mathbb{R})$ a transformação linear definida pela expressão

$$f(A) = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} A - A \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

Determine o núcleo e a imagem de f. Resolução.

24. Seja V um espaço vetorial finitamente gerado e $f\colon V \to V$ uma transformação linear. Sejam B_1 e B_2 bases para V.
(a) Mostre que se A é a matriz que representa f na base B_1 e $S = S_{B_1\to B_2}$ é a matriz de mudança de coordenadas da base B_1 para a base B_2 então a matriz que representa f na base B_2 é SAS^{-1}.
(b) Determine a matriz que representa a transformação linear $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ definida por

$$f(x,y,z) = (x - 2y, x + z, y + z)$$

com respeito à base ordenada $B = ((1,0,1),(0,1,1),(-1,0,1))$ de $\mathbb{R}^3$.

(c) As matrizes $\begin{bmatrix} 2 & -1 \\ -1 & 0 \end{bmatrix}$ e $\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}$ podem representar a mesma transformação linear em bases distintas de $\mathbb{R}^2$? Justifique.

Resolução.

25. Uma transformação linear $P\colon V \to V$ diz-se uma *projeção* se $P^2 = P$.
   (a) Mostre que $(\mathrm{Id} - P)$ também é uma projeção.
   (b) Mostre que $P(V) = N(\mathrm{Id} - P)$ (e portanto $(\mathrm{Id} - P)(V) = N(P)$).
   (c) Mostre que $V = P(V) \oplus (\mathrm{Id} - P)(V)$, portanto uma projeção dá azo a uma decomposição de V em soma direta.
   (d) Reciprocamente, sendo $U, W \subset V$ subespaços vetoriais tais que $V = U \oplus W$ mostre que existe uma projeção única $P\colon V \to V$ com $P(V) = U$ e $N(P) = W$, pelo que há uma correspondência bijetiva entre projeções e decomposições de V como a soma direta de dois espaços.
   (e) Note que, geometricamente, o efeito de P é projetar no plano dado pela imagem de P ao longo das direções no núcleo de P. Numa base adequada para V, a expressão da projeção fica muito simples. Escreva uma expressão para as seguintes projeções (usando por exemplo o Exercício 24 (a))
      (i) A projeção de $\mathbb{R}^2$ na reta $\{(x,y) \in \mathbb{R}^2 \colon x + 2y = 0\}$ ao longo da direção do vetor $(1,1)$,
      (ii) A projeção de $\mathbb{R}^3$ no plano $\{(x,y,z) \in \mathbb{R}^3 \colon x - y + z = 0\}$ ao longo da direção do vetor $(1,3,-2)$.

Resolução.

26. Sejam A uma matriz $m \times n$ e B uma matriz $n \times k$. Mostre que

$$\mathrm{car}(AB) \leq \min(\mathrm{car}(A), \mathrm{car}(B))$$

Consegue dar uma fórmula para a característica do produto? *Sugestão: Considere as transformações lineares determinadas pelas matrizes e recorde que a característica de uma matriz é a dimensão da imagem da transformação associada.* Resolução.

27. Sejam V e W espaços vetoriais de dimensão finita e $f\colon V \to W$ uma transformação linear com característica k. Mostre que existem bases B_1 para V e B_2 para W tais que a matriz $A = A_{f,B_1,B_2}$ tem entradas

$$a_{ij} = \begin{cases} 1 & \text{se } i = j \text{ e } 1 \leq i \leq k \\ 0 & \text{caso contrário.} \end{cases}$$

Resolução.

28. Recorde do exercício 3.20 que um espaço vetorial complexo tem uma estrutura natural de espaço vetorial real. Sejam V, W espaços vetoriais complexos. Uma transformação linear $T\colon V \to W$ de espaços vetoriais complexos é claramente também uma transformação linear entre V e W quando estes são considerados como espaços vetoriais reais.

(a) Mostre que se $T\colon \mathbb{C} \to \mathbb{C}$ é a transformação linear definida por $T(z) = (a+bi)z$ então a matriz que representa T com respeito à base $(1, i)$ para $\mathbb{C}$ como espaço vetorial real é $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$.

(b) Mais geralmente mostre que se $B_1 = (v_1, \ldots, v_n)$ e $B_2 = (w_1, \ldots, w_m)$ são bases para os espaços vetoriais complexos V e W respetivamente, $B_1' = (v_1, iv_1, \ldots, v_n, iv_n)$, $B_2' = (w_1, iw_1, \ldots, w_m, iw_m)$ são as bases associadas para V e W enquanto espaços vetoriais reais e $A = A_{T,B_1,B_2} \in M_{n\times m}(\mathbb{C})$ tem entrada jk dada por $a_{jk} = x_{jk} + iy_{jk}$ (com x_{jk} e y_{jk} as partes real e imaginária de a_{jk}) então a matriz $A_{T,B_1',B_2'} \in M_{2m\times 2n}(\mathbb{R})$ toma a forma

$$A_{T,B_1',B_2'} = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & \ddots & & A_{2n} \\ \vdots & & & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{bmatrix}$$

onde cada A_{jk} é o bloco 2×2 dado por $A_{jk} = \begin{bmatrix} x_{jk} & -y_{jk} \\ y_{jk} & x_{jk} \end{bmatrix}$.

Resolução.

29. Sejam V, W espaços vetoriais e $f\colon V \to W$ uma transformação linear. O *gráfico* de uma transformação linear é o subconjunto do espaço vetorial $V \times W$ (cujas operações são definidas componente a componente) definido por

$$G(f) = \{(v, f(v))\colon v \in V\} \subset V \times W$$

(a) Verifique que $G(f)$ é um subespaço vetorial de $V \times W$.

(b) Seja U um espaço vetorial de dimensão finita e $V \subset U$ um subespaço vetorial com $\dim V < \dim U$. Mostre que se W é um subespaço vetorial de U com $\dim W + \dim V = \dim U$ e $W \cap V = 0$, então a função $\phi\colon V \times W \to U$ definida por $(v, w) \mapsto v + w$ é um isomorfismo de espaços vetoriais.

(c) Nas condições da alínea anterior, seja $Z \subset U$ um espaço vetorial de U com $\dim Z = \dim V$. Mostre que $\phi^{-1}(Z)$ é o gráfico de uma transformação linear $f\colon V \to W$ se e só se $Z \cap W = \{0\}$.

Resolução.

30. Escreva cada subespaço de $V \subset \mathbb{R}^n$ indicado como o gráfico de uma transformação linear $f\colon \mathbb{R}^p \to \mathbb{R}^q$ para os subespaços $\mathbb{R}^p, \mathbb{R}^q \subset \mathbb{R}^n$ indicados. Ou seja, determine uma transformação linear $f\colon \mathbb{R}^p \to \mathbb{R}^q$ tal que $V = \{(x, f(x)) \in \mathbb{R}^n\colon x \in \mathbb{R}^p\}$ onde estamos a escrever um vetor de $\mathbb{R}^n$ na forma (x, y) com $x \in \mathbb{R}^p$ e $y \in \mathbb{R}^q$.

(a) $V = \{(x, y, z) \in \mathbb{R}^3\colon x+2y+3z = 0\}$ como gráfico de uma função de $\mathbb{R}^2 = \{(x, y, 0)\}$ para $\mathbb{R} = \{(0, 0, z)\}$ e de $\mathbb{R}^2 = \{(x, 0, z)\}$ para $\mathbb{R} = \{(0, y, 0)\}$.

(b) $V = \{(x, y, z, w) \in \mathbb{R}^4\colon x - y + z + w = 0, 2x + z - w = 0\}$ como gráfico de uma função de $\mathbb{R}^2 = \{(x, y, 0, 0)\}$ para $\mathbb{R}^2 = \{(0, 0, z, w)\}$ e de $\mathbb{R}^2 = \{(0, 0, z, w)\}$ para $\mathbb{R}^2 = \{(x, y, 0, 0)\}$.

(c) $V = \{(x,y,z,u,v) \in \mathbb{R}^5 \colon x+y+v=0, y-3u+v=0, z-u+v=0\}$ como gráfico de uma função de $\mathbb{R}^2 = \{(0,0,0,u,v)\}$ para $\mathbb{R}^3 = \{(x,y,z,0,0)\}$.

Resolução.

31. Fatorização canónica de uma transformação linear.
    (a) Mostre que toda a transformação linear $T \colon V \to W$ se fatoriza como uma composição $T = i \circ p$ com $p \colon V \to U$ sobrejetiva $i \colon U \to W$ injetiva. *Sugestão: Pode tomar para U um espaço quociente apropriado de V.*
    (b) Seja $T \colon V \to W$ uma transformação linear e $Z \subset V$ um subespaço vetorial tal que $Z \subset N(T)$. Mostre que existe uma única transformação linear $\overline{T} \colon V/Z \to W$ tal que $T = \overline{T} \circ \pi$, onde $\pi \colon V \to V/Z$ denota a aplicação canónica definida por $\pi(v) = v + Z$.
    (c) Mostre que a fatorização da alínea (a) é única a menos de isomorfismo canónico no seguinte sentido: se $p' \colon V \to U'$ é sobrejetiva, $i' \colon U' \to W$ é injetiva e $T = i' \circ p'$ então existe um único isomorfismo $\phi \colon U \to U'$ tal que $i' \circ \phi = i$ e $\phi \circ p = p'$.

Resolução.

32. Significado geométrico dos pivots.
    (a) Seja $A \in M_{m\times n}(\mathbb{R})$ e considere as projeções $\pi_j \colon \mathbb{R}^n \to \mathbb{R}^j$ nas primeiras j coordenadas. Mostre que a função $L \colon \{1,\ldots,n\} \to \mathbb{N}_0$ definida por
    $$L(j) = \dim(\pi_j(EL(A))$$
    é crescente (isto é $j \leq k \Rightarrow L(j) \leq L(k)$) sendo $L(n) = \mathrm{car}(A)$.
    (b) Mostre que A tem um pivot na coluna i se e só se $L(i) = L(i-1)+1$ (convencionando que $L(0) = 0$).
    (c) Mostre que A tem pivots nas colunas $i_1 < \ldots < i_k$ se e só se a projeção nas coordenadas $i_1,\ldots,i_k$
    $$\pi_{i_1,\ldots,i_k} \colon \mathbb{R}^n \to \mathbb{R}^k$$
    se restringe a um isomorfismo entre $EL(A)$ e $\mathbb{R}^k$ (ou seja, se e só se $EL(A)$ é o gráfico de uma transformação linear de $\mathbb{R}^k$ para $\mathbb{R}^{n-k}$, com $\mathbb{R}^k$ correspondendo às variáveis $i_1,\ldots,i_k$ de $\mathbb{R}^n$, e $\mathbb{R}^{n-k}$ às restantes) e cada um dos índices i_j toma o mais pequeno valor possível que torna esta afirmação verdadeira.
    (d) Descreva concretamente as várias possíbilidades para os planos contidos em $\mathbb{R}^2, \mathbb{R}^3$ e $\mathbb{R}^4$.

Resolução.

33. Sejam V, W, U espaços vetoriais e $f \colon V \to W$ uma transformação linear.
    (a) Mostre que a aplicação $f^* \colon L(W,U) \to L(V,U)$ definida por $f^*(T) = T \circ f$ é uma transformação linear.
    (b) Recorde que o dual de um espaço vetorial V é $V^* = L(V,\mathbb{R})$. Verifique que se $B = (v_1,\ldots,v_n)$ é uma base para V e definirmos $\phi_1,\ldots,\phi_n \in V^*$ por
    $$\phi_j(v_i) = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases}$$

então $B^* = (\phi_1, \ldots, \phi_n)$ é uma base para V^* que se chama a *base dual* de B. Note que $\phi_i(v)$ é a i-ésima coordenada de v na base B.

(c) O dual pode ser visto como o espaço das funções coordenadas em V juntamente com a função nula: mostre que se V é finitamente gerado e $\phi \in V^* \setminus \{0\}$ então existe uma base ordenada $B = (v_1, \ldots, v_n)$ de V tal que $\phi(v)$ é a primeira coordenada de v na base B.

(d) Sendo B_1 e B_2 bases ordenadas para V e W e A a matriz que representa f com respeito a estas bases, mostre que a matriz que representa f^* com respeito às bases duais é A^T.

(e) Mostre que a aplicação $\Phi\colon V \to (V^*)^*$ definida por

$$(\Phi(v))(\psi) = \psi(v)$$

é uma transformação linear injetiva, que é um isomorfismo se V é finitamente gerado.

(f) Seja V o espaço vetorial de todos os polinómios reais. Mostre que a transformação linear da alínea anterior não é sobrejetiva neste caso.

Resolução.

34. Seja V um espaço vetorial de dimensão finita e $W \subset V$ um subespaço vetorial. O *aniquilador* de W é o subconjunto do dual $V^* = L(V, \mathbb{R})$ definido por

$$W^o = \{\phi \in V^* \colon \phi(w) = 0 \text{ para todo o } w \in W\}$$

Note que se encararmos a expressão $\phi(w) = 0$ como uma equação linear satisfeita por w, o aniquilador é o espaço vetorial das equações lineares que são satisfeitas por todos os elementos de W.

(a) Mostre que $\dim W + \dim W^o = \dim V$.

(b) Sejam $f\colon U \to V$ uma transformação linear e $f^*\colon V^* \to U^*$ a transformação linear dual definida por $f^*\phi = \phi \circ f$. Mostre que o núcleo de f^* é $N(f^*) = f(U)^o$.

(c) Use o resultado das alíneas anteriores e a relação entre as dimensões do núcleo e da imagem de f^* para concluir que

$$\dim f(U) = \dim f^*(V^*)$$

Se f é representada por uma matriz A, no Exercício 33(d) viu que f^* é representada por A^T. Assim a equação anterior mostra que as dimensões do espaço das linhas e das colunas de uma matriz A coincidem. Note que este argumento é completamente independente da discussão dos sistemas lineares no início do semestre.

(d) Recorde do Exercício 33(e) que há um isomorfismo canónico $\psi\colon V \to (V^*)^*$ definido por $(\psi(v))(\phi) = \phi(v)$. Mostre que sendo $W \subset V$ um subespaço se tem $\psi(W) = (W^o)^o$.

(e) Supondo que U é também de dimensão finita mostre que mediante a identificação de U e V com $(U^*)^*$ e $(V^*)^*$ se tem, para uma transformação linear $f\colon U \to V$ que $(f^*)^* = f$. Formalmente isto significa que sendo ψ_U e ψ_V os isomorfismos canónicos da alínea (d), temos $\psi_V \circ f = (f^*)^* \circ \psi_U$

Resolução.

35. Usando apenas uma vez o método de Gauss, calcule de forma eficiente bases para os espaços das linhas e colunas da matriz

$$\begin{bmatrix} 1 & -3 & 1 & 4 & 0 \\ -2 & 6 & 0 & 3 & -1 \\ 1 & -3 & -1 & 0 & 4 \end{bmatrix}$$

Resolução.

36. Seja $V = \{f \in F(\mathbb{R}, \mathbb{R})\colon f$ é diferenciável$\}$ e $T\colon V \to F(\mathbb{R}, \mathbb{R})$ a transformação linear definida por $Tf = f'$.
    (a) Determine o conjunto dos $f \in V$ tais que $Tf = \lambda f$ para algum $\lambda \in \mathbb{R}$. *Sugestão: Para mostrar que encontrou todas as soluções calcule* $\frac{d}{dt}\left(e^{-\lambda t} f(t)\right)$ *para* f *uma solução.*
    (b) Determine a solução geral da equação diferencial $f' - f = t + t^2$.
    Resolução.

37. Para cada uma das transformações lineares do exercício 16(a) e (c) determine as soluções das duas equações lineares $f(v) = (1, 0, -2)$ e $f(v) = (1, 1, 4)$. Resolução.

38. Seja V o espaço vetorial dos polinómios de grau ≤ 3 e considere a transformação linear $T\colon V \to M_{3\times 2}(\mathbb{R})$ definida por

$$T(1+t) = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ -1 & 0 \end{bmatrix}, T(2+t^2) = \begin{bmatrix} 0 & 2 \\ 2 & 1 \\ -1 & 0 \end{bmatrix}, T(t-t^3) = \begin{bmatrix} 1 & 1 \\ 2 & 0 \\ 0 & 1 \end{bmatrix}, T(2+t) = \begin{bmatrix} 2 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Determine o conjunto das soluções das seguintes equações lineares.

(a) $T(x) = 0$ (b) $T(x) = \begin{bmatrix} 2 & -1 \\ 5 & 3 \\ -1 & 0 \end{bmatrix}$ (c) $T(x) = \begin{bmatrix} 1 & -1 \\ -2 & 0 \\ 0 & 0 \end{bmatrix}$

Resolução.

39. Seja $V \subset F(\mathbb{R}, \mathbb{R})$ o subespaço vetorial formado por todas as funções duas vezes diferenciáveis e considere a transformação linear $T\colon V \to F(\mathbb{R}, \mathbb{R})$ definida por $T(f) = f'' - f$.
    (a) Verifique que e^x e e^{-x} pertencem ao núcleo de T. Pode mostrar-se estas duas funções geram o núcleo de T.
    (b) Determine uma solução particular da equação $f'' - f = x^2 + x$. *Sugestão: Determine um polinómio que satisfaça esta equação.*
    (c) Determine todas as soluções da equação $f'' - f = x^2 + x$.
    (d) Determine todas as soluções do sistema de equações diferenciais

$$\begin{cases} f'' - f = 0 \\ f' + 2e^{2x} f = 4e^{3x} \end{cases}$$

Resolução.

5. O Determinante

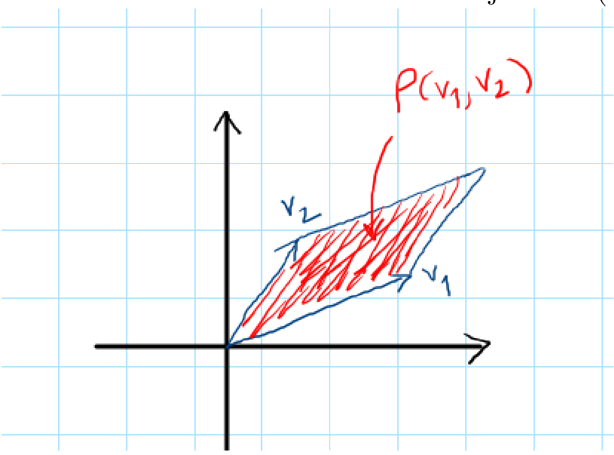
O nosso objetivo seguinte é compreender em completo detalhe as transformações lineares $T\colon V \to V$ onde V é um espaço vetorial complexo de dimensão finita. Para atingir este objetivo vai ser útil ter um critério para a invertibilidade de uma matriz quadrada em termos das suas entradas. O critério irá dizer que uma matriz $A \in M_{n\times n}(\mathbb{K})$ é invertível se e só se uma certa expressão complicada das entradas da matriz (chamada o determinante da matriz) não se anula. No que se segue $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$.

5.1. **Motivação.** Para motivar esta expressão vamos começar por discutir o caso em que o corpo $\mathbb{K}$ é o dos números reais, caso em que o determinante tem uma interpretação geométrica. Consideremos primeiro os casos $n = 2$ e $n = 3$.

A uma matriz $A \in M_{2\times 2}(\mathbb{R})$ podemos associar um paralelogramo

$$P(v_1, v_2) = \{\alpha v_1 + \beta v_2 \colon 0 \le \alpha, \beta \le 1\} \subset \mathbb{R}^2$$

gerado pelas linhas v_1 e v_2 da matriz. A matriz será invertível se o seu espaço das linhas for $\mathbb{R}^2$, ou, equivalentemente, se o paralelogramo $P(v_1, v_2)$ não degenerar num segmento de reta (ou até na origem). Apelando ao conceito intuitivo de área podemos dizer que a matriz será invertível se a área do conjunto $P(v_1, v_2)$ for não nula.



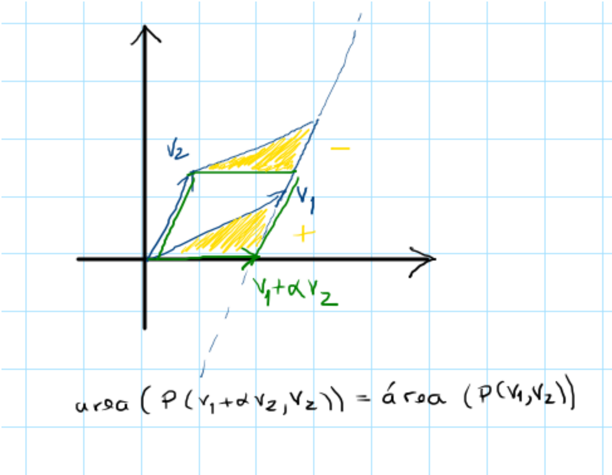
Analogamente, uma matriz 3×3 terá característica menor que 3 se e só se o paralelipípedo

$$P(v_1, v_2, v_3) = \{\alpha v_1 + \beta v_2 + \gamma v_3 \colon 0 \le \alpha, \beta, \gamma \le 1\}$$

gerado pelas linhas v_1, v_2, v_3 da matriz tiver volume nulo. Mais geralmente pode definir-se uma noção de volume n-dimensional para um subconjunto de $\mathbb{R}^n$ (como irão ver em Cálculo 2) e então a condição para a invertibilidade de uma matriz em $M_{n\times n}(\mathbb{R})$ é que o volume n-dimensional do paralelipípedo n-dimensional $P(v_1, \dots, v_n)$ gerado pelas linhas da matriz seja não nulo.

Se conseguirmos obter uma fórmula para o volume n-dimensional do paralelipípedo gerado por n vetores em $\mathbb{R}^n$ isso dar-nos-á um critério para a invertibilidade da matriz: que o volume do paralelipípedo gerado pelas linhas seja não nulo. A observação básica que nos permite obter esta fórmula é a seguinte:

Ao deslizar em conjunto duas arestas de um paralelogramo ao longo da reta gerada por outra das outras arestas, a área do paralelogramo não se altera



ou seja

$$\text{área}(P(v_1, v_2)) = \text{área}(P(v_1 + \alpha v_2, v_2)) \tag{27}$$

(e claro que o mesmo se verifica se deslizarmos o ponto final de v_2 ao longo da direção v_1). Esta fórmula diz-nos por exemplo que as áreas dos paralelogramos determinados pelas linhas das matrizes

$$\begin{bmatrix} a & 0 \\ c & d \end{bmatrix} \qquad \text{e} \qquad \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}$$

são iguais, pois $(0, d)$ pode obter-se de (c, d) deslizando ao longo de $(a, 0)$ (a não ser que $a = 0$, mas nesse caso as áreas são nulas e a afirmação permanece verdadeira). Assim, a área do paralelogramo com arestas $(a, 0)$ e (c, d) é a área do retângulo com arestas $(a, 0)$ e $(0, d)$, ou seja $|ad|$ (mesmo que a ou d sejam 0). Mas a fórmula (27) diz-nos mais geralmente que quando aplicamos o método de Gauss a uma matriz 2×2, a área do paralelogramo associado não muda! Supondo que $a \neq 0$ temos

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \xrightarrow{L_2 - \frac{c}{a} L_1} \begin{bmatrix} a & b \\ 0 & d - \frac{bc}{a} \end{bmatrix}$$

logo concluímos que a área de um paralelogramo com arestas (a, b) e (c, d) é

$$\text{área}\,(P((a, b), (c, d))) = |a| \cdot \left| d - \frac{bc}{a} \right| = |ad - bc|$$

É um exercício simples verificar que esta fórmula permanece válida mesmo quando $a = 0$. Obtemos assim a condição desejada nas entradas da matriz:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \text{ é invertível sse } ad - bc \neq 0$$

Podemos fazer um raciocínio análogo para matrizes 3×3 mas a fórmula obtida será agora mais complicada. Novamente o volume de um paralelipípedo $P(v_1, v_2, v_3)$ em $\mathbb{R}^3$

não se alterará se deslizarmos o ponto final de uma das arestas paralelamente ao plano determinado pelas outras duas, ou seja, por exemplo

$$\text{volume } P(v_1 + \alpha v_2, v_2, v_3) = \text{volume } P(v_1, v_2, v_3)$$

Portanto o volume de um paralelipípedo com arestas as linhas da matriz

$$\begin{bmatrix} a & b & c \\ 0 & e & f \\ 0 & 0 & i \end{bmatrix}$$

será o volume do paralelipípedo reto com arestas de comprimento $|a|, |e|$ e $|i|$, e podemos reduzir a este caso usando eliminação de Gauss:

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \underset{L_3 - \frac{g}{a}L_1}{\overset{L_2 - \frac{d}{a}L_1}{\longrightarrow}} \begin{bmatrix} a & b & c \\ 0 & e - \frac{db}{a} & f - \frac{dc}{a} \\ 0 & h - \frac{gb}{a} & i - \frac{gc}{a} \end{bmatrix} \xrightarrow{L_3 - \frac{h - \frac{gb}{a}}{e - \frac{db}{a}} L_2} \begin{bmatrix} a & b & c \\ 0 & e - \frac{db}{a} & f - \frac{dc}{a} \\ 0 & 0 & i - \frac{gc}{a} - \frac{h - \frac{gb}{a}}{e - \frac{db}{a}}(f - \frac{dc}{a}) \end{bmatrix}$$

Obtemos assim a fórmula

$$\text{volume } (P((a,b,c),(d,e,f),(g,h,i))) = |a| \left| e - \frac{db}{a} \right| \left| i - \frac{gc}{a} - \frac{h - \frac{gb}{a}}{e - \frac{db}{a}}(f - \frac{dc}{a}) \right|$$

que, simplificando, se transforma em:

$$\text{volume } (P((a,b,c),(d,e,f),(g,h,i))) = |aei + bfg + cdh - ceg - bdi - afh|$$

Fica como exercício verificar que esta fórmula é válida mesmo nos casos em que $a = 0$, ou $a \neq 0$ mas $e - \frac{db}{a} = 0$, nos quais a eliminação de Gauss feita acima tem de ser modificada.

O cálculo anterior sugere que não será prático obter e manipular diretamente uma expressão para o volume de um paralelipípedo n-dimensional. Com efeito, para $n = 4$ veremos que a fórmula análoga tem 24 termos, para $n = 5$, 120 termos, e em geral o número de termos é $n!$. Uma expressão de tal complexidade só pode ser manipulada conceptualmente.

5.2. **A função determinante.** Abstraindo as propriedades, não do volume, mas da expressão mais fundamental que obtivemos acima para $n = 2, 3$ cujo módulo é o volume, obtemos a seguinte definição, que faz sentido para um corpo arbitrário.

Definição 5.3. *Seja $\mathbb{K}$ um corpo. Uma função* determinante *para as matrizes $n \times n$ com entradas em $\mathbb{K}$ é uma função*

$$\det \colon M_{n \times n}(\mathbb{K}) \to \mathbb{K}$$

que se denota por

$$\det \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \qquad \textit{ou} \qquad \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$$

que satisfaz as seguintes propriedades.

(i) **Multilinearidade:** *Para cada* $1 \leq i \leq n$ *temos*

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1}+b_{i1} & & a_{in}+b_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & & a_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ b_{i1} & & b_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$$

e, para α *um escalar qualquer,*

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ \alpha a_{i1} & & \alpha a_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = \alpha \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & & a_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$$

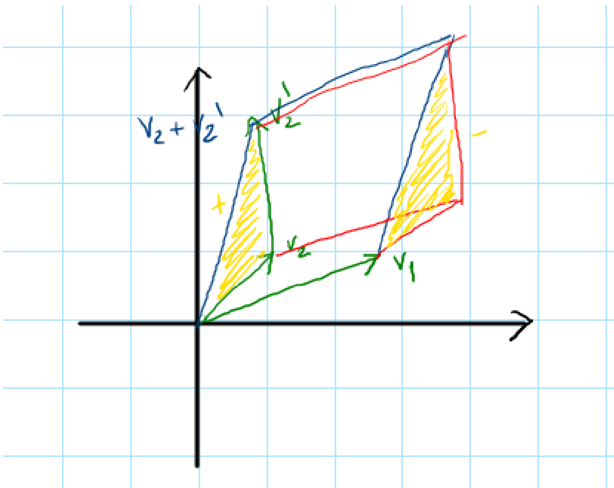
(ii) **Alternância:** $\det A = 0$ *se duas linhas da matriz* A *forem iguais.*

(iii) **Normalização:** $\det I_n = 1$.

Em concreto, no caso das matrizes 2×2, a primeira propriedade diz por exemplo que

$$\begin{vmatrix} 2 & 1 \\ 1+3 & 2+4 \end{vmatrix} = \begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} + \begin{vmatrix} 2 & 1 \\ 3 & 4 \end{vmatrix} \quad \text{e} \quad \begin{vmatrix} 2 & 1 \\ -2 \cdot 3 & -2 \cdot 4 \end{vmatrix} = -2 \begin{vmatrix} 2 & 1 \\ 3 & 4 \end{vmatrix}$$

e corresponde à *aditividade* da área (ou mais geralmente, volume n-dimensional).



Identificando as linhas de uma matriz $n \times n$ com vetores de $\mathbb{K}^n$, podemos pensar na função determinante como uma função $D\colon \mathbb{K}^n \times \cdots \times \mathbb{K}^n \to \mathbb{K}$ que associa um escalar a um n-tuplo $(v_1, \ldots, v_n)$ de vetores de $\mathbb{K}^n$ (v_i é a i-ésima linha da matriz). Deste ponto de vista, a propriedade de multilinearidade escreve-se

$$(28) \qquad D(v_1, \ldots, \alpha v_i + \beta v_i', \ldots, v_n) = \alpha D(v_1, \ldots, v_i, \ldots, v_n) + \beta D(v_1, \ldots, v_i', \ldots, v_n)$$

onde $v_1, \ldots, v_n \in \mathbb{K}^n$ são vetores arbitrários e α, β escalares arbitrários. A equação (28) diz que, para cada i entre 1 e n, a função $D_i \colon \mathbb{K}^n \to \mathbb{K}$ que se obtém quando fixamos todos os vectores excepto o i-ésimo,

$$D_i(v) = D(v_1, \ldots, v_{i-1}, v, v_{i+1}, \ldots, v_n)$$

é linear (ou seja, um elemento do dual de $\mathbb{K}^n$).

Em geral, uma função $D \colon V \times \cdots \times V \to \mathbb{K}$ satisfazendo (28) diz-se uma *função multilinear*[19] (é linear em cada variável independentemente).

A razão para o nome da segunda propriedade na definição de determinante é a seguinte.

Proposição 5.4. *Seja $D \colon V \times \cdots \times V \to \mathbb{K}$ uma função multilinear e suponhamos*[20] *que $1 + 1 \neq 0$ em $\mathbb{K}$. Então as seguintes condições são equivalentes:*

(i) $D(v_1, \ldots, v_n) = 0$ se $v_i = v_j$ para algum $i \neq j$.
(ii) Se $i < j$, então $D(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) = -D(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n)$ para todos os $v_1, \ldots, v_n$ (isto é, a troca de dois argumentos tem como efeito a troca de sinal do valor da função).

Dem.

$(i) \Rightarrow (ii)$ Supondo que $i < j$, e aplicando a linearidade primeiro na i-ésima variável e depois na j-ésima obtemos

$$\begin{aligned} D(v_1, \ldots, v_i + v_j, \ldots, v_i + v_j, \ldots, v_n) &= D(v_1, \ldots, v_i, \ldots, v_i + v_j, \ldots, v_n) + \\ &\quad D(v_1, \ldots, v_j, \ldots, v_i + v_j, \ldots, v_n) = \\ &= D(v_1, \ldots, v_i, \ldots, v_i, \ldots, v_n) + D(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) \\ &\quad + D(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n) + D(v_1, \ldots, v_j, \ldots, v_j, \ldots, v_n) \end{aligned}$$

Substituindo os termos com argumentos repetidos por 0 obtém-se

$$0 = 0 + D(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) + D(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n) + 0$$

que é equivalente à condição (ii).

$(ii) \Rightarrow (i)$ Se $v_i = v_j$, então a troca do i-ésimo argumento com o j-ésimo não tem nenhum efeito. Portanto

$$D(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) = -D(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n) = -D(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n)$$

e portanto $(1 + 1)D(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) = 0$.

□

[19] Também se chama um tensor-n covariante em V.

[20] Diz-se que a *característica* do corpo $\mathbb{K}$ é diferente de 2. Notamos no entanto que a demonstração da implicação $(i) \Rightarrow (ii)$ permanece válida mesmo que esta hipótese não seja verificada.

5.5. **Existência e unicidade do determinante.** Vamos ver que as propriedades (*i*) a (*iii*) na definição de determinante especificam completamente uma função.

Teorema 5.6. *Existe uma única função determinante* $\det M_{n\times n}(\mathbb{K}) \to \mathbb{K}$

A demonstração deste teorema segue o padrão usual: iremos ver que só há uma possibilidade para uma tal função (obtendo no processo uma fórmula para o determinante) e depois verificar que essa única possibilidade satisfaz de facto os axiomas da definição. Começamos por ilustrar este processo usando os axiomas para ver que a única função determinante nas matrizes 2×2 é

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$$

Sendo $a, b, c, d \in \mathbb{K}$ quaisquer e aplicando a linearidade do determinante na primeira linha da matriz temos

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = a \begin{vmatrix} 1 & 0 \\ c & d \end{vmatrix} + b \begin{vmatrix} 0 & 1 \\ c & d \end{vmatrix}$$

e aplicando agora a linearidade na segunda linha obtemos

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = a \left(c \begin{vmatrix} 1 & 0 \\ 1 & 0 \end{vmatrix} + d \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} \right) + b \left(c \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} + d \begin{vmatrix} 0 & 1 \\ 0 & 1 \end{vmatrix} \right)$$

Os primeiro e último termos do lado direito do sinal de igual na expressão acima são nulos porque as linhas das matrizes em questão estão repetidas. Pelas propriedades (*iii*) e (*ii*) respetivamente temos

$$\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1 \qquad \text{e} \qquad \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = -1$$

portanto

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

é a única função real das matrizes 2×2 que satisfaz as condições da Definição 5.3.

Façamos agora o caso mais realista de uma matriz 3×3. Assumindo que existe a função determinante e usando linearidade na primeira linha obtemos

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = a \begin{vmatrix} 1 & 0 & 0 \\ d & e & f \\ g & h & i \end{vmatrix} + b \begin{vmatrix} 0 & 1 & 0 \\ d & e & f \\ g & h & i \end{vmatrix} + c \begin{vmatrix} 0 & 0 & 1 \\ d & e & f \\ g & h & i \end{vmatrix} \tag{29}$$

Desenvolvendo o primeiro termo do lado direito do sinal de igual usando linearidade na segunda linha obtemos

$$a \begin{vmatrix} 1 & 0 & 0 \\ d & e & f \\ g & h & i \end{vmatrix} = a \left(d \begin{vmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ g & h & i \end{vmatrix} + e \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ g & h & i \end{vmatrix} + f \begin{vmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ g & h & i \end{vmatrix} \right)$$

O primeiro termo na soma do lado direito é nulo porque a primeira linha está repetida. Da mesma forma, cada parcela do lado direito em (29) vai dar origem a dois termos não nulos quando aplicarmos linearidade ao longo da segunda linha da matriz. Podemos agora aplicar linearidade ao longo da terceira linha a cada um destes 6 termos. Por exemplo, para o primeiro dos seis resultaria

$$ae\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ g & h & i \end{vmatrix} = ae\left(g\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{vmatrix} + h\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{vmatrix} + i\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix}\right) = aei$$

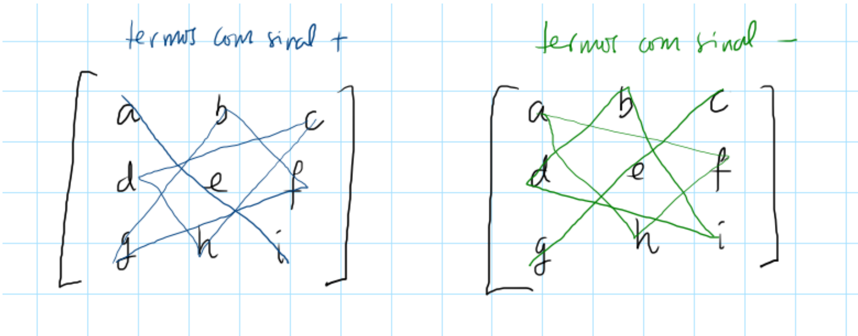
uma vez que os dois primeiros termos da soma anterior têm linhas repetidas e o determinante da matriz identidade é 1. Aplicando o mesmo raciocínio para os restantes termos não nulos na expansão até à segunda linha obtemos a seguinte expressão para o determinante:

$$aei + afh\begin{vmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{vmatrix} + bdi\begin{vmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{vmatrix} + bfg\begin{vmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{vmatrix} + cdh\begin{vmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{vmatrix} + ceg\begin{vmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{vmatrix}$$

Os determinantes das matrizes com 0s e $1s$ são ± 1 consoante o número de vezes que temos que trocar um par de linhas para transformar a matriz na identidade é par ou ímpar. Recuperamos assim a expressão para o determinante de uma matriz 3×3:

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = aei - afh - bdi + bfg + cdh - ceg$$

Os sinais da fórmula anterior podem ser memorizados usando a seguinte mnemónica (que se chama a *regra de Sarrus*):



Procedendo desta forma para uma matriz $n \times n$ é agora claro que vamos obter uma expressão para o determinante. Haverá um termo não nulo na expressão para cada matriz de $1s$ e 0s que tenha exatamente um 1 em cada linha e em cada coluna. Para descrever estes termos por meio de uma expressão necessitamos de alguma terminologia.

Definição 5.7. *Uma* permutação *do conjunto* $\{1, \dots, n\}$ *é uma função bijetiva*

$$\sigma \colon \{1, \dots, n\} \to \{1, \dots, n\}$$

Designamos por Σ_n *o conjunto de todas estas permutações.*

Uma permutação descreve uma troca de ordem. Deve ser familiar do ensino secundário que o número de elementos de Σ_n é $n!$. Os termos na expansão do determinante vão corresponder precisamente às permutações: se chamarmos $\sigma(i)$ à coluna em que aparece o 1 na linha i, a condição que não apareçam dois 1s na mesma coluna é $\sigma(i) \neq \sigma(j)$ para $i \neq j$, ou seja é a injetividade da função σ. Como uma função injetiva de um conjunto com n elementos para ele próprio é necessariamente uma bijeção, conclui-se que a função determinada por uma matriz de 0s e 1s satisfazendo as condições indicadas é uma bijeção.

O termo do determinante de A correspondente a uma permutação σ será dado pelo produto das entradas de A que ocorriam nas posições onde estão os 1s, ou seja o produto dos $a_{i\sigma(i)}$ com $i = 1, \ldots, n$. O termo terá um sinal que será $\pm$ consoante o número de vezes que temos que trocar pares de linhas para transformar a matriz de 0s e 1s na identidade é par ou impar. Chamando a este sinal $\operatorname{sgn}(\sigma)$ - *o sinal da permutação* σ - obtemos a seguinte expressão para o determinante:

$$\det(A) = \sum_{\sigma \in \Sigma_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)} \tag{30}$$

O argumento anterior torna claro que se existir uma função determinante, ela é única (tem que ser dada pela fórmula (30)!). Mas neste momento não é ainda claro que uma tal função exista. Há muitas maneiras de trocar pares de linhas de forma a obter a matriz identidade a partir de uma matriz de 0s e 1s. Se para uma das maneiras o número de trocas fosse par e para outra maneira fosse ímpar concluir-se-ia que a função determinante não podia existir.

Não é fácil verificar diretamente que o sinal de uma permutação está bem definido. Em vez disso vamos dar uma construção indutiva do determinante. Uma vez que isto esteja feito teremos implicitamente provado que o sinal de uma permutação está bem definido! Será necessariamente

$$\operatorname{sgn}(\sigma) = \det A(\sigma) \qquad \text{com } A(\sigma) \text{ a matriz com entradas } a_{ij} = \begin{cases} 1 & \text{se } j = \sigma(i) \\ 0 & \text{caso contrário.} \end{cases} \tag{31}$$

A matriz $A(\sigma)$ diz-se uma *matriz de permutação.* O efeito que tem nas coordenadas de um vetor linha ou coluna é uma permutação das coordenadas. Por exemplo,

$$A(\sigma) \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_{\sigma(1)} \\ x_{\sigma(2)} \\ \vdots \\ x_{\sigma(n)} \end{bmatrix}$$

É um bom exercício ver o que acontece quando se multiplica $A(\sigma)$ à esquerda por um vetor linha.

Dem. do Teorema 5.6. Já vimos que se existir uma função determinante ela é única (e dada pela fórmula (30)). Vamos ver por indução em n que existe uma função determinante para matrizes $n \times n$. Quando $n = 1$, é imediato que

$$\det([a_{11}]) = a_{11}$$

Suponhamos que já definimos uma função determinante nas matrizes $n \times n$. Dada uma matriz A de tipo $(n+1) \times (n+1)$, seja A_{1i} a matriz $n \times n$ que se obtém de A suprimindo a primeira linha e a i-ésima coluna. Vamos definir

$$\det(A) = a_{11}\det(A_{11}) - a_{12}\det(A_{12}) + \ldots + (-1)^n a_{1(n+1)} \det A_{1(n+1)} \tag{32}$$

fórmula esta que é motivada pela relação entre os determinantes para matrizes 3×3 e 2×2 que obtivemos anteriormente.

Temos a verificar que $\det A$ verifica as condições $(i) - (iii)$ da Definição 5.3. A condição (i) é verificada porque a expressão (32) é claramente linear na primeira linha da matriz A e, por hipótese de indução, nas restantes, uma vez que as funções $\det(A_{1i})$ são multilineares.

A condição (iii) também é verificada porque as entradas na primeira linha da matriz identidade I_{n+1} com excepção da primeira são todas nulas. Uma vez que $(I_{(n+1)})_{11} = I_n$ obtemos

$$\det(I_{n+1}) = 1 \cdot \det(I_n) = 1.$$

Resta-nos verificar que se uma das linhas de A está repetida então $\det A = 0$. Se a repetição ocorrer nas linhas i e j com $i, j \geq 2$ então todos os termos $\det(A_{1i})$ em (32) se anulam (por hipótese de indução) e portanto $\det A = 0$. Se $i = 1$, podemos assumir que $j = 2$ uma vez que, por hipótese de indução, o termo direito da equação (32) troca de sinal quando trocamos a linha j de A com a segunda linha.

Suponhamos assim que A tem a primeira e segunda linha iguais. Se A é uma matriz 2×2 a expressão (32) é

$$\det(A) = a_{11}a_{22} - a_{12}a_{21} = a_{11}a_{12} - a_{12}a_{11} = 0$$

Se $n > 1$, podemos, por hipótese de indução aplicar a expressão (32) às matrizes $n \times n$ A_{1i}. A entrada $1j$ na primeira linha de A_{1i} é igual a

$$\begin{cases} a_{2j} & \text{se } j < i \\ a_{2(j+1)} & \text{se } j > i \end{cases}$$

portanto

$$\det(A_{1j}) = \sum_{j=1}^{i-1} (-1)^{j-1} a_{2j} \det(A_{12|ij}) + \sum_{j=i+1}^{n+1} (-1)^j a_{2j} \det(A_{12|ij})$$

onde $A_{12|ij}$ denota a matriz $(n-1) \times (n-1)$ que se obtém de A suprimindo as primeiras duas linhas e as colunas i e j. Substituindo esta expressão em (32) vemos que há dois termos nos quais aparece $\det(A_{12|ij})$ para i, j dados com $1 \leq i < j \leq n$:

$$(-1)^{i-1} a_{1i} \cdot (-1)^{j-2} a_{2j} \det(A_{12|ij})$$

que é o $(j-1)$-ésimo termo da expansão do termo $(-1)^{i-1} a_{1i} \det(A_{1i})$ à direita do sinal de igual em (32) e

$$(-1)^{j-1} a_{1j} \cdot (-1)^{i-1} a_{2i} \det(A_{12|ij})$$

que vem da expansão do termo $(-1)^{j-1} a_{1j} \det(A_{1j})$. Uma vez que as primeiras duas linhas da matriz são iguais, temos

$$(-1)^{i-1} a_{1i} \cdot (-1)^{j-2} a_{2j} \det(A_{12|ij}) + (-1)^{j-1} a_{1j} \cdot (-1)^{i-1} a_{2i} \det(A_{12|ij}) = 0$$

o que conclui a demonstração. $\square$

Observação 5.8. *Uma função $f\colon M_{n\times n}(\mathbb{K}) \to \mathbb{K}$ satisfazendo as propriedades (i) e (ii) na Definição 5.3 chama-se uma função multilinear alternante das linhas da matriz. O argumento usado na demonstração de unicidade do determinante aplicado a uma tal função (sem qualquer alteração) mostra que*

$$f(A) = \sum_{\sigma\in\Sigma_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)}\cdots a_{n\sigma(n)} f(I_n)$$

pelo que o valor de uma tal função em qualquer matriz é completamente determinado pelo valor que assume na matriz identidade. Mas sendo $\lambda \in \mathbb{K}$ qualquer, a função $A \mapsto \lambda\det(A)$ é uma função multilinear alternante que assume o valor λ em I_n, pelo que se conclui que toda a função multilinear alternante é da forma

$$f(A) = \lambda\det(A)$$

sendo $\lambda = f(I_n)$.

5.9. **Propriedades do determinante.** Vamos agora ver algumas propriedades importantes do determinante que nos ajudam a calculá-lo.

Definição 5.10. *Seja A uma matriz $n\times n$. Para $1 \leq i, j \leq n$ designamos por A_{ij} a matriz $(n-1)\times(n-1)$ que se obtém de A omitindo a i-ésima linha e a j-ésima coluna. O* menor-ij *de A é o número* $\det A_{ij}$ *e o* cofator-ij *de A é* $(-1)^{i+j}\det A_{ij}$*. A matriz $n\times n$ cuja entrada ij é o cofator-ij diz-se* a matriz dos cofatores de A *e denota-se por* $\operatorname{cof} A$.

Proposição 5.11 (Propriedades do determinante)**.** *Sejam A e B matrizes $n\times n$.*

(i) **Expansão de Laplace:** *Sendo $1 \leq i \leq n$, temos*

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij}\det(A_{ij})$$

onde A_{ij} é a matriz que se obtém de A omitindo a linha i e a coluna j. A fórmula acima chama-se a expansão de Laplace para o determinante ao longo da linha i.

(ii) $\det(AB) = \det(A)\det(B)$

(iii) $\det(A^T) = \det(A)$

(iv) $A(\operatorname{cof}(A))^T = \det(A)I_n$.

Antes de vermos a demonstração destas propriedades notemos as seguintes consequências.

Corolário 5.12 (Expansão de Laplace ao longo de colunas)**.** *Sendo $1 \leq j \leq n$, temos*

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij}\det(A_{ij})$$

Dem. A expansão ao longo da coluna j no enunciado é exatamente a expansão ao longo da linha j de A^T. Logo calcula $\det A^T = \det A$. □

Corolário 5.13. *Uma matriz quadrada A é invertível sse $\det A \neq 0$ e nesse caso*

$$A^{-1} = \frac{1}{\det A}(\operatorname{cof} A)^T$$

Dem. Se A é invertível então $\det(A)\det(A^{-1}) = \det(AA^{-1}) = \det(I_n) = 1$ logo $\det(A) \neq 0$ e

$$\det(A^{-1}) = \frac{1}{\det A}$$

Reciprocamente se $\det A \neq 0$, a Proposição 5.11 (iv) diz-nos que

$$A\left(\frac{1}{\det A}(\operatorname{cof} A)^T\right) = I_n$$

pelo que A é invertível (cf. Proposição 4.36 (vi)) sendo a inversa descrita pela fórmula no enunciado. □

Esta fórmula para a inversa de uma matriz tem mais utilidade teórica do que prática porque não é fácil calcular determinantes de matrizes grandes. É no entanto muito útil para matrizes 2×2, caso em que afirma que

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \qquad \text{quando } ad - bc \neq 0$$

É ainda razoável aplicar a fórmula para matrizes 3×3 mas se o tamanho das matrizes é maior ou igual a 4 a fórmula começa a tornar-se impraticável e é muito mais rápido usar o método de Gauss-Jordan.

Dem. da Proposição 5.11. (i) Para $i = 1$ a expansão de Laplace é simplesmente a expressão indutiva (32) usada para demonstrar a existência do determinante. Se $i > 1$, seja $\tilde{A}$ a matriz que se obtém de A trocando a linha 1 com a linha i. Aplicando (32) obtemos

$$\det(A) = -\det(\tilde{A}) = -\sum_{j=1}^{n}(-1)^{1+j}\tilde{a}_{1j}\det(\tilde{A}_{1j}) = -\sum_{j=1}^{n}(-1)^{1+j}a_{ij}\det(\tilde{A}_{1j}) \tag{33}$$

Notamos agora que as matrizes $\tilde{A}_{1j}$ e A_{ij} diferem pela troca da $(i-1)$-ésima linha com o bloco formado pelas linhas que a precedem - o que corresponde a $(i-2)$-trocas de pares de linhas à medida que a linha $(i-1)$ “flutua até chegar à superfície”. Portanto

$$\det(\tilde{A}_{1j}) = (-1)^{i-2}\det A_{ij}$$

Substitituindo em (33) obtemos a fórmula pretendida.

(ii) Fixada uma matriz B, considere-se a função $f\colon M_{n\times n}(\mathbb{R}) \to \mathbb{R}$ definida por

$$f(A) = \det(AB)$$

Trata-se de uma função multilinear e alternante das linhas de A pela definição do produto de matrizes e pelas propriedades (i) e (ii) na definição de função determinante. Uma vez que $f(I_n) = \det(B)$, a Observação 5.8 diz-nos que $f(A) = \det(A)\det(B)$.

(iii) A expressão (30) diz-nos que

$$\det(A^T) = \sum_{\sigma\in\Sigma_n}\operatorname{sgn}(\sigma)a^T_{1\sigma(1)}\cdots a^T_{n\sigma(n)} = \sum_{\sigma\in\Sigma_n}\operatorname{sgn}(\sigma)a_{\sigma(1)1}\cdots a_{\sigma(n)n}$$

Seja $\sigma^{-1}\colon \{1,\dots,n\} \to \{1,\dots,n\}$ a permutação inversa de σ (isto é, a permutação que verifica $\sigma^{-1}(\sigma(i)) = i$ para $i = 1,\dots,n$). Então

$$\sigma(i) = j \Leftrightarrow i = \sigma^{-1}(j)$$

e portanto

$$a_{\sigma(1)1}\cdots a_{\sigma(n)n} = a_{1\sigma^{-1}(1)}\cdots a_{n\sigma^{-1}(n)}$$

(do lado direito do sinal de igual aparecem as mesmas entradas da matriz que do lado esquerdo mas por outra ordem; estão agora ordenados pelo primeiro índice, enquanto que à esquerda estão ordenados pelo segundo). Temos assim

$$\det(A^T) = \sum_{\sigma\in\Sigma_n} \operatorname{sgn}(\sigma) a_{1\sigma^{-1}(1)}\cdots a_{n\sigma^{-1}(n)} \tag{34}$$

As matrizes $A(\sigma)$ associadas às permutações (ver (31)) colocam na coordenada i de um vetor coluna a coordenada que estava na posição $\sigma(i)$. Logo o efeito de $A(\sigma)A(\sigma^{-1})$ num vetor coluna é colocar na coordenada i a componente $x_{\sigma^{-1}(\sigma(i))} = x_i$. Portanto

$$A(\sigma)A(\sigma^{-1}) = I_n \Rightarrow \det(A(\sigma))\det A(\sigma^{-1}) = 1 \Rightarrow \det(A(\sigma)) = \det(A(\sigma^{-1}))$$

onde a última implicação usa que o determinante de uma matriz de permutação é necessariamente ± 1. Notando que $\operatorname{sgn}(\sigma) = \det A(\sigma)$ e substituindo em (34) temos

$$\det(A^T) = \sum_{\sigma\in\Sigma_n} \operatorname{sgn}(\sigma^{-1}) a_{1\sigma^{-1}(1)}\cdots a_{n\sigma^{-1}(n)}$$

Quando σ percorre todos os elementos de Σ_n, o mesmo sucede com a sua inversa σ^{-1} pelo que a expressão à direita na igualdade acima é exatamente a fórmula (30) para o determinante de A. Isto conclui a demonstração.

(iv) A fórmula no enunciado diz-nos que o produto da linha i da matriz A pela coluna j da matriz $(\operatorname{cof} A)^T$ é $\det(A)$ se $i = j$ e 0 caso contrário. A expressão para este produto é

$$\sum_{k=1}^n a_{ik}((\operatorname{cof} A)^T)_{kj} = \sum_{k=1}^n a_{ik}(\operatorname{cof} A)_{jk} = \sum_{k=1}^n a_{ik}(-1)^{j+k}\det(A_{jk})$$

Quando $i = j$, a expressão anterior é a expansão de Laplace para o determinante de A ao longo da linha i e é portanto igual a $\det A$. Para $i \neq j$, a expressão é a expansão de Laplace ao longo da linha j da matriz que se obtém de A repetindo a linha j na linha i, e é portanto igual a 0.

□

Observação 5.14. *É instrutivo pensar em escrever explicitamente a igualdade indicada na Proposição 5.11(ii) em termos das entradas das matrizes envolvidas. Mesmo para matrizes* 3×3 *a complexidade é enorme! É fácil no entanto convencer-se que, pelo menos a menos de sinal, a igualdade se deve verificar:*

Atendendo à Proposição 5.11(iii), $|\det A|$ *é o volume do paralelípipedo que tem por arestas as colunas da matriz* A*, paralelipípedo este que é a imagem do cubo com arestas unitárias em* $\mathbb{R}^n$ *pela transformação linear* $x \mapsto Ax$ *. Segue-se que a imagem de um cubo qualquer em* $\mathbb{R}^n$ *por esta transformação tem volume igual a* $|\det(A)|$ *vezes o volume do* *cubo original. Verão em Cálculo 2 que o volume de um subconjunto (razoável) de* $\mathbb{R}^n$ *se define aproximando esse conjunto por cubos muito pequenos e passando ao limite. Segue-se então que* $|\det A|$ *é o fator pelo qual a transformação linear* $x \mapsto Ax$ *multiplica volumes.*

Uma vez que AB *é a matriz que representa a composta das transformações lineares representadas por* A *e* B*, segue-se que o fator pela qual* AB *multiplica volumes é* $|\det(A)||\det(B)|$.

Exemplo 5.15. *Vamos calcular o determinante*

$$\begin{vmatrix} 2 & 0 & 3 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 4 & 5 & 7 \\ 1 & 8 & 9 & 3 \end{vmatrix}$$

usando a expansão de Laplace. Uma vez que a segunda linha tem 3 zeros, é mais eficiente fazer a expansão ao longo dessa linha. Obtemos

$$0\cdot(-1)^{2+1}\begin{vmatrix} 0 & 3 & 0 \\ 4 & 5 & 7 \\ 8 & 9 & 3 \end{vmatrix} + 0\cdot(-1)^{2+2}\begin{vmatrix} 2 & 3 & 0 \\ 1 & 5 & 7 \\ 1 & 9 & 3 \end{vmatrix} + 1\cdot(-1)^{2+3}\begin{vmatrix} 2 & 0 & 0 \\ 1 & 4 & 7 \\ 1 & 8 & 3 \end{vmatrix} + 0\cdot(-1)^{2+4}\begin{vmatrix} 2 & 0 & 3 \\ 1 & 4 & 5 \\ 1 & 8 & 9 \end{vmatrix}$$

e fazendo agora a expansão de Laplace do único termo não nulo ao longo da primeira linha obtém-se

$$-\begin{vmatrix} 2 & 0 & 0 \\ 1 & 4 & 7 \\ 1 & 8 & 3 \end{vmatrix} = -2\cdot(-1)^{1+1}\begin{vmatrix} 4 & 7 \\ 8 & 3 \end{vmatrix} = -2(4\cdot 3 - 7\cdot 8) = 88.$$

5.16. **Aplicação à solução de sistemas lineares.** A fórmula para a inversa de uma matriz em termos do determinante conduz à seguinte fórmula explícita para a solução de um sistema linear quando a matriz dos coeficiente do sistema é invertível.

Proposição 5.17 (Regra de Cramer). *Seja* A *uma matriz* $n \times n$ *invertível e* b *uma matriz* $n \times 1$*. Então a componente* x_i *da solução do sistema linear*

$$Ax = b$$

é dada pela fórmula

$$x_i = \frac{\det A_i}{\det A}$$

onde A_i *é a matriz que se obtém de* A *substituindo a coluna* i *de* A *por* b*.*

Dem. A componente x_i da solução do sistema é a i-ésima entrada de $A^{-1}b$ e é portanto dada por

$$x_i = \sum_{j=1}^{n} c_{ij} b_j$$

onde c_{ij} é a entrada ij da matriz A^{-1}. Pelo Corolário 5.13 esta entrada é

$$c_{ij} = (-1)^{i+j}\frac{\det(A_{ji})}{\det A}$$

pelo que

$$x_i = \frac{1}{\det A}\sum_{j=1}^{n}(-1)^{i+j} b_j \det(A_{ji})$$

O somatório na expressão anterior é exatamente o desenvolvimento de Laplace ao longo da coluna i da matriz A_i do enunciado. Isto conclui a demonstração. □

Exemplo 5.18. *Vamos achar a coordenada y da solução do sistema*

$$\begin{cases} 2x+3y+z=3 \\ x-y+z=4 \\ x+2y-z=5 \end{cases}$$

Pela regra de Cramer temos

$$y = \frac{\begin{vmatrix} 2 & 3 & 1 \\ 1 & 4 & 1 \\ 1 & 5 & -1 \end{vmatrix}}{\begin{vmatrix} 2 & 3 & 1 \\ 1 & -1 & 1 \\ 1 & 2 & -1 \end{vmatrix}} = -\frac{11}{7}$$

5.19. **O determinante de uma matriz triangular por blocos.** Recorde que uma matriz quadrada A diz-se *triangular superior* se $a_{ij}=0$ para $i>j$ (isto é se todas as entradas abaixo da diagonal principal são nulas) e *triangular inferior* se $a_{ij}=0$ para $i<j$ (isto é se todas as entradas acima da diagonal principal são nulas).

Usando a expansão de Laplace e indução, é imediato verificar que o determinante de uma matriz triangular (superior ou inferior) é igual ao produto das entradas na diagonal

$$\begin{vmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_n \end{vmatrix} = \lambda_1 \cdots \lambda_n$$

Uma generalização da última propriedade que é muito útil diz respeito ao cálculo de determinantes de matrizes escritas *por blocos.*

Proposição 5.20. *O determinante de uma matriz triangular por blocos com blocos quadrados na diagonal é o produto dos determinantes dos blocos diagonais*

$$\begin{vmatrix} A_1 & * & \cdots & * \\ 0 & A_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & A_n \end{vmatrix} = |A_1|\cdots|A_n|$$

Dem. Exercício. □

Exemplo 5.21.

$$\begin{vmatrix} 1 & 0 & 5 & 11 & 6 \\ 3 & 2 & 3 & 27 & 5 \\ 0 & 0 & 4 & 2 & 2 \\ 0 & 0 & 0 & 3 & 2 \\ 0 & 0 & 0 & 1 & 4 \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ 3 & 2 \end{vmatrix} \cdot 4 \cdot \begin{vmatrix} 3 & 2 \\ 1 & 4 \end{vmatrix} = 2 \cdot 4 \cdot 10 = 80$$

5.22. **O produto externo de vetores.**

Definição 5.23. *Sejam $v, w \in \mathbb{R}^3$. O* produto externo *de v e w é o vetor $v \times w \in \mathbb{R}^3$ definido por*

$$\begin{aligned} v \times w &= \begin{vmatrix} \mathbf{e_1} & \mathbf{e_2} & \mathbf{e_3} \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} = (v_2w_3 - v_3w_2)\mathbf{e_1} + (v_3w_1 - v_1w_3)\mathbf{e_2} + (v_1w_2 - v_2w_1)\mathbf{e_3} \\ &= (v_2w_3 - v_3w_2, v_3w_1 - v_1w_3, v_1w_2 - v_2w_1) \end{aligned}$$

onde $\mathbf{e_i}$ designa o i-ésimo vetor da base canónica de $\mathbb{R}^3$ e a expressão à direita se obtém expandindo o determinante ao longo da primeira linha.

O determinante que aparece na definição anterior é apenas uma (boa) mnemónica, uma vez que formalmente não consideramos matrizes que tenham vetores como algumas das suas entradas. A verdadeira definição do produto externo são as expressões equivalentes que aparecem a seguir ao determinante mas para efetuar cálculos a mnemónica é extremamente útil.

Exemplo 5.24.

$$(1, -3, 2) \times (5, 0, 2) = \left| \begin{matrix} \mathbf{e_1} & \mathbf{e_2} & \mathbf{e_3} \\ 1 & -3 & 2 \\ 5 & 0 & 2 \end{matrix} \right] = (-6, 8, 15)$$

O produto externo tem inúmeras aplicações em Matemática e Física. Por exemplo, será usado em Cálculo 2 para calcular fluxos de campos vetoriais através de superfícies; em Mecânica aparece por exemplo na expressão para o momento angular de uma partícula em torno de um ponto, que é dado pela expressão $\vec{L} = \vec{r} \times \vec{p}$ com $\vec{r}$ o vetor de posição e $\vec{p}$ o momento linear; finalmente, a Força de Lorentz a que uma carga elétrica em movimento é sujeita ao interagir com um campo magnético $\vec{B}$ é $\vec{F} = q\vec{v} \times \vec{B}$, com $\vec{v}$ a velocidade e q a carga da partícula em questão.

As propriedades do determinante implicam imediatamente certas propriedades do produto externo. Vamos usar a notação $\langle v, w \rangle$ para o produto interno de dois vetores de $v, w \in \mathbb{R}^3$ familiar do ensino secundário e definido por

$$\langle (v_1, v_2, v_3), (w_1, w_2, w_3) \rangle = v_1w_1 + v_2w_2 + v_3w_3$$

No Capítulo 7 iremos estudar em detalhe este produto interno assim como as suas generalizações a espaços vetoriais reais e complexos.

Proposição 5.25 (Propriedades do produto externo). *(i) O produto externo é linear em cada um dos seus argumentos.*

(ii) $v \times w = -w \times v$

(iii) $v \times v = 0$

(iv) $\langle u, (v \times w) \rangle = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}$

Demonstração. A primeira afirmação é verdadeira porque o determinante é multilinear, a segunda porque o determinante troca de sinal quando se trocam linhas, e a terceira porque o determinante é zero se houver uma linha repetida. A quarta é uma consequência das definições de produto interno e externo juntamente com a expansão de Laplace ao longo da primeira linha. ☐

A Proposição anterior dá-nos também o significado geométrico do produto externo. De facto, uma vez que o determinante de uma matriz com linhas repetidas é 0, pela propriedade (iv) temos

$$\langle v, v \times w \rangle = \langle w, v \times w \rangle = 0$$

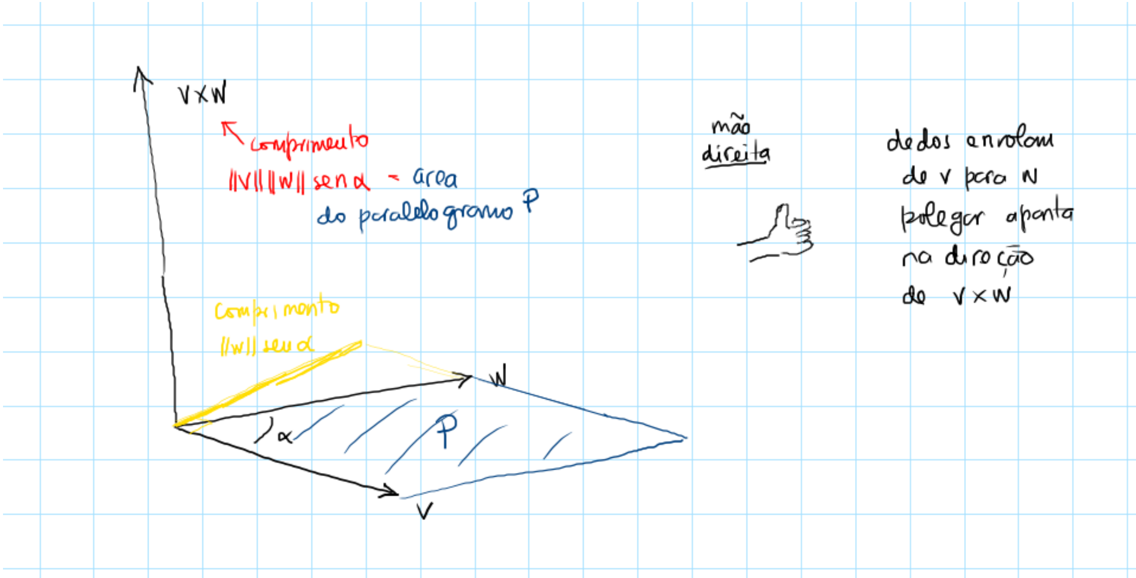
pelo que $v \times w$ *é ortogonal ao plano gerado por* v *e* w (se v e w são colineares, então as propriedade (i) e (iii) dizem-nos que o produto externo é o vetor nulo). Além disso, dada a interpretação do determinante como o volume do paralelipípedo temos que

$$\|v \times w\|^2 = \langle v \times w, v \times w \rangle = \begin{vmatrix} - & v \times w & - \\ - & v & - \\ - & w & - \end{vmatrix}$$

é o volume do paralelipípedo com base o paralelogramo formado por v e w sendo a outra aresta perpendicular ao paralelogramo com comprimento $\|v \times w\|$. Este volume é a área da base vezes o comprimento da aresta perpendicular à base pelo que $\|v \times w\|$ *é a área do paralelogramo com arestas* v *e* w. Note-se que no caso degenerado em que v e w são colineares a afirmação anterior continua a ser válida.

Em suma, quando v, w não são colineares, o produto externo $v \times w$ é um vetor perpendicular ao plano determinado por v e w, cujo comprimento é a área do paralelogramo com arestas v e w. Se α for a o ângulo entre v e w, a área do paralelogramo é a mesma que a área do retângulo com arestas de comprimento $\|v\|$ e $\|w\| \operatorname{sen} \alpha$ (isto vê-se deslizando a aresta w ao longo de uma reta paralela a v até que fique perpendicular a v - movimento que não afeta a área do paralelogramo). Portanto

$$\|v \times w\| = \|v\| \|w\| \operatorname{sen} \alpha \qquad \text{com } \alpha \text{ o ângulo entre } v \text{ e } w$$



Há dois vetores com a propriedade que acabámos de descrever, que diferem apenas no seu sentido. O sentido do produto externo é dado pela *regra da mão direita:* se colocarmos a mão *direita* aberta, com os dedos que não o polegar juntos apontando na direção de v e a rodarmos de modo a que esses dedos apontem para w, o polegar aponta na direção de $v \times w$.

A razão pela qual isto é assim prende-se com o *significado geométrico do sinal do determinante de uma matriz* 3×3 *invertível*, que é precisamente

$$\begin{vmatrix} - & v_1 & - \\ - & v_2 & - \\ - & v_3 & - \end{vmatrix} > 0 \quad \Leftrightarrow \quad v_1, v_2 \text{ e } v_3 \text{ satisfazem a regra da mão direita.}$$

Nesse caso diz-se que a *orientação* do referencial (v_1, v_2, v_3) é positiva. Note-se que o referencial canónico formado pela base canónica de $\mathbb{R}^3$ tem esta propriedade. Assim podemos pensar nos referenciais positivamente orientados como sendo "semelhantes"ao referencial habitual.

Para perceber a afirmação anterior recorde-se que podemos transformar a matriz com linhas v_1, v_2 e v_3 na matriz identidade aplicando o método do Gauss-Jordan. Cada passo do método consiste numa operação

$$L_i - \alpha L_j, \quad \alpha L_i, \quad L_i \leftrightarrow L_j \tag{35}$$

que, em termos da matriz dos coeficientes do sistema, corresponde à multiplicação à esquerda por uma matriz elementar. No primeiro caso trata-se de uma matriz triangular com uma única entrada não nula fora da diagonal, no segundo caso por uma matriz diagonal com α na posição i e 1 nas restantes, e no último por uma matriz de permutação que troca as linhas i e j. O sinal do determinante da matriz dos coeficientes não é alterado pelas operações do primeiro tipo, permanece igual ou é alterado pelas do segundo tipo consoante α é positivo ou negativo, e é sempre alterado por operações do terceiro tipo (com $i \neq j$).

Resta agora observar que o efeito que as operações (35) têm relativamente à verificação da regra da mão direita por um referencial é exatamente o mesmo: operações do primeiro tipo não têm efeito no que diz respeito à verificação da regra da mão direita pelas linhas da matriz; operações do segundo tipo não têm efeito se $\alpha > 0$ e têm efeito se $\alpha < 0$; as operações do terceiro tipo têm sempre efeito. Conclui-se que o determinante é positivo sse as linhas satisfazem a regra da mão direita.

Observação 5.26. *A fórmula da Definição 5.23 pode ser usada para definir o produto externo de $(n-1)$ vetores em $\mathbb{R}^n$, para $n \geq 1$. Sendo $\mathbf{e_1}, \ldots, \mathbf{e_n}$ a base canónica de $\mathbb{R}^n$ e $v_1, \ldots v_{n-1}$ vetores de $\mathbb{R}^n$, define-se*

$$v_1 \times \cdots \times v_{n-1} = \begin{vmatrix} \mathbf{e_1} & \cdots & \mathbf{e_n} \\ - & v_1 & - \\ - & \cdots & - \\ - & v_{n-1} & - \end{vmatrix}$$

Por exemplo, se $n = 2$, o produto externo de um único vetor $v_1 \in \mathbb{R}^2$ dá o vetor que se obtém de v_1 rodando 90 graus no sentido horário. Em geral, os argumentos acima mostram que o produto externo é nulo sse os vetores $v_1, \ldots, v_{n-1}$ forem linearmente dependentes e senão é perpendicular ao plano $(n-1)$-dimensional gerado por $v_1, \ldots, v_{n-1}$. Além disso, o comprimento do produto externo é o volume $(n-1)$-dimensional do paralelipípedo com arestas $v_1, \ldots, v_{n-1}$ e o seu sentido é tal que a orientação do referencial $(v_1 \times \cdots \times v_{n-1}, v_1, \ldots, v_{n-1})$ coincide com a da base canónica de $\mathbb{R}^n$.

5.27. **Exercícios.**

1. Qual é a área da imagem pela função $f(x, y) = (x + 2y, 3x - y)$ do paralelogramo com vértice na origem e arestas $(1, 3)$ e $(2, 5)$? Resolução.
2. Calcule os seguintes determinantes indicando se as matrizes em questão são invertíveis

(a) $\begin{vmatrix} -1 & 2 \\ 3 & 1 \end{vmatrix}$ (b) $\begin{vmatrix} -1 & 2 & 1 \\ -3 & 2 & -1 \\ 1 & 4 & 3 \end{vmatrix}$ (c) $\begin{vmatrix} -1 & 2 & 1 & 7 \\ -3 & 2 & -1 & 5 \\ 1 & 4 & 3 & 2 \\ -1 & 2 & 1 & 7 \end{vmatrix}$

(d) $\begin{vmatrix} 0 & 2 & 1 & 2 \\ -3 & 0 & -1 & 0 \\ 0 & 1 & 3 & 0 \\ -1 & 2 & 1 & 1 \end{vmatrix}$ (e) $\begin{vmatrix} 2 & 0 & 1 & 1 & 6 \\ 0 & 1 & -1 & 5 & 1 \\ 0 & 0 & 3 & 2 & -3 \\ 0 & 0 & 0 & 7 & 1 \\ 0 & 0 & 0 & 0 & -1 \end{vmatrix}$

(f) Uma matriz $n \times n$ tal que a soma das entradas de cada linha seja 0.

Resolução.

3. Calcule o seguinte determinante usando o método de Gauss para reduzir o cálculo ao do determinante de uma matriz triangular

$$\begin{vmatrix} 0 & -1 & 1 & 1 & 0 \\ 1 & 1 & -1 & 2 & 1 \\ -1 & 0 & 3 & 2 & -3 \\ -2 & 1 & 0 & 1 & 1 \\ 1 & 0 & 2 & 0 & -1 \end{vmatrix}$$

Resolução.

4. Sendo $A = \begin{bmatrix} -1 & 2 \\ 1 & 1 \end{bmatrix}$ e $B = \begin{bmatrix} 0 & 2 \\ -1 & 3 \end{bmatrix}$, calcule $\det(A^T B^{-2} A B^T A^{-2} (A^T)^3)$. Resolução.

5. Considere a matriz $A = \begin{bmatrix} -1 & 2 & 1 \\ 1 & 2 & -1 \\ 1 & 4 & 3 \end{bmatrix}$ Mostre que não existe uma matriz (real) B tal que $A = BB^T$ nem tal que $A = B^2$. Resolução.

6. Sabendo que $\begin{vmatrix} a & 0 & b & c \\ d & 2 & e & f \\ g & 0 & h & i \\ j & 0 & k & l \end{vmatrix} = 6$, determine $\begin{vmatrix} c & i & l \\ b-3c & h-3i & k-3l \\ 2a & 2g & 2j \end{vmatrix}$ Resolução.

7. Sem calcular o determinante, ache o coeficiente de x^3 nas seguintes expressões

(a) $\begin{vmatrix} -1 & x & 1 \\ x & 2x & -1 \\ 1 & 4 & 3x \end{vmatrix}$ (b) $\begin{vmatrix} x & 1 & 1 & -2 \\ -1 & 0 & x & 0 \\ 0 & x & -x & 1 \\ -x & 3 & 1 & 1 \end{vmatrix}$

Resolução.

8. (O determinante de Vandermonde). Sejam $x_1, \ldots, x_n$ escalares. Mostre que

$$\begin{vmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & & \vdots & \\ 1 & x_{n-1} & x_{n-1}^2 & \cdots & x_{n-1}^{n-1} \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{vmatrix} = (x_2 - x_1) \cdots (x_n - x_1)(x_3 - x_2) \cdots (x_n - x_2) \cdots (x_n - x_{n-1})$$

pelo que o determinante se anula se e só se $x_i = x_j$ para algum $i \neq j$.

Sugestão: Sendo C_i a coluna i da matriz, subtraia $x_1 C_{n-1}$ a C_n, depois $x_1 C_{n-2}$ a C_{n-1}, e assim por diante de forma a ficar com uma expressão para o determinante pretendido dada pelo determinante de uma matriz em que a primeira linha é $(1, 0, \ldots, 0)$. Use as propriedades do determinante e indução. Resolução.

9. Calcule a entrada 23 da inversa da matriz $\begin{bmatrix} 0 & 1 & 1 & -2 \\ -1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & -1 & 1 & 1 \end{bmatrix}$ Resolução.

10. Seja $A \in M_{n\times n}(\mathbb{K})$. O *traço* de A é a soma das entradas diagonais da matriz.

$$\operatorname{tr} A = a_{11} + a_{22} + \ldots + a_{nn}$$

(a) Mostre que dadas $A, B \in M_{n\times n}(\mathbb{K})$ temos $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.
(b) Usando a alínea anterior conclua que se S é invertível, então $\operatorname{tr}(SAS^{-1}) = \operatorname{tr}(A)$.
Resolução.

11. Considere uma função $A\colon \mathbb{R} \to M_{k\times l}(\mathbb{R})$ definida por

$$A(t) = [a_{ij}(t)]_{\substack{1\leq i\leq k \\ 1\leq j\leq l}}$$

portanto A consiste em kl funções reais de variável real $t \mapsto a_{ij}(t)$. Define-se a derivada de $A(t)$ entrada a entrada. Isto é, A é diferenciável se cada $a_{ij}(t)$ é diferenciável e então $A'(t)$ é, por definição, $[a'_{ij}(t)]$.

(a) Mostre que a regra de derivação do produto se mantém, isto é que dadas duas funções matriciais $A(t)$ e $B(t)$ então

$$\frac{d}{dt}\left(A(t)B(t)\right) = A'(t)B(t) + A(t)B'(t)$$

(b) Mostre que se $A\colon \mathbb{R} \to M_{n\times n}(\mathbb{R})$ é diferenciável em $t = 0$ e $A(0) = I_n$ então

$$\frac{d}{dt}\left(\det A(t)\right)_{|t=0} = \operatorname{tr} A'(0)$$

Sugestão: Use a expressão para o determinante em termos de uma soma indexada por permutações e considere primeiro o caso em que as funções $a_{ij}(t)$ são afins (isto é da forma $\alpha + \beta t$ com $\alpha, \beta \in \mathbb{R}$).
Resolução.

12. Mostre que o determinante da matriz triangular por blocos $\begin{bmatrix} A & B \\ 0 & C \end{bmatrix}$ (com A e C matrizes quadradas) é igual a $\det(A)\det(C)$.
Sugestão: Considere a função multilinear das linhas de C definida por

$$f(C) = \begin{vmatrix} A & B \\ 0 & C \end{vmatrix}$$

Aproveite para calcular os seguintes determinantes

$$\text{(i)} \begin{vmatrix} 2 & 1 & -5 & 4 & 10 \\ 0 & 3 & 6 & -1 & -3 \\ 0 & 7 & 8 & -2 & 3 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & -3 & 1 \end{vmatrix} \qquad \text{(ii)} \begin{vmatrix} 2 & 1 & -5 & 4 & 10 & 7 \\ 1 & 2 & -1 & -1 & -3 & 8 \\ 1 & 0 & 0 & -2 & 3 & -3 \\ 0 & 0 & 0 & -2 & 1 & 3 \\ 0 & 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{vmatrix}$$

Resolução.

13. Considere uma matriz por blocos $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$ com A e D quadradas e A invertível. Mostre que
$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = \det(A)\det(-CA^{-1}B + D)$$
Sugestão: Multiplique a matriz $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$ *por uma matriz da forma* $\begin{bmatrix} I & X \\ 0 & I \end{bmatrix}$ *apropriada de forma*[21] *a que o produto seja uma matriz triangular inferior por blocos.* Resolução.
14. Seja V um espaço vetorial sobre o corpo $\mathbb{K}$. Uma função $h\colon V\times V \to \mathbb{K}$ diz-se *bilinear* se para cada $v \in V$, as funções $w \mapsto h(v,w)$ e $w \mapsto h(w,v)$ são lineares (isto é, pertencem ao dual de V). Diz-se também que um tal h é uma *forma bilinear.*
   (a) Sendo $B = (v_1, \dots, v_n)$ uma base ordenada para V com n elementos, mostre que existe uma única matriz $A \in M_{n\times n}(\mathbb{K})$ tal que
   $$h(v,w) = [v]_B^T A [w]_B$$
   Diz-se que a matriz A representa a função bilinear h com respeito à base B.
   (b) Considere a função de $\overline{h}\colon V \to V^*$ de V para o seu dual, definida por
   $$v \mapsto h(v,\cdot)$$
   Mostre que se trata de uma transformação linear que é representada pela matriz transposta à da alínea (a) quando se toma para base de V^* a base B^* dual [22] à base B de V.
   (c) Diz-se que a forma bilinear $h\colon V\times V \to \mathbb{K}$ é *não degenerada* se
   $$h(v,w) = 0 \text{ para todo o } w \Rightarrow v = 0$$
   Mostre que h é não degenerada se e só se a transformação linear $\overline{h}$ da alínea anterior é um isomorfismo. Portanto *uma forma bilinear não degenerada dá uma identificação de um espaço vetorial com o seu dual.*
   (d) Determine a matriz que representa a função bilinear $h\colon \mathbb{R}^3\times\mathbb{R}^3 \to \mathbb{R}$ definida por
   $$h((x,y,z),(u,v,w)) = xu + 3xw - yv + 2yw + 5yu - 3zu + zw$$
   com respeito à base canónica de $\mathbb{R}^3$. Esta forma é não-degenerada?

   Resolução.
15. Usando a regra de Cramer determine a componente y da solução do sistema linear
$$\begin{cases} x + 2y - z = 2 \\ x - y + 3z = 1 \\ x - 2z = 1 \end{cases}$$
Resolução.

[21] Essencialmente está a aproveitar-se o facto de A ser invertível para eliminar B com operações elementares sobre as colunas.

[22] Ver o exercício 4.33.

16. Calcule os seguintes produtos externos
   (i) $(1, 0, 2) \times (2, 0, 1)$
   (ii) $(1, -1, 1) \times (1, 2, 4)$
   (iii) $(2v_1 + 3v_2) \times (-v_1 + v_2)$

   Resolução.
17. Use o produto externo para achar a equação cartesiana do plano que passa pelo ponto $(0, 1, 3)$ e que é paralelo a $L(\{(1, 1, 2), (-1, 0, 1)\})$. Resolução.
18. Seja $u = (a, b, c)$ e considere a transformação linear $T \colon \mathbb{R}^3 \to \mathbb{R}^3$ definida por $T(v) = u \times v$. Determine a representação matricial de T com respeito à base canónica de $\mathbb{R}^3$. Resolução.
19. Dados $u, v, w \in \mathbb{R}^3$, mostre que $u \times (v \times w) = \langle u, w\rangle v - \langle u, v\rangle w$. Resolução.
20. **Os quaterniões.** A multiplicação dos números complexos pode ser vista como uma multiplicação de vetores de $\mathbb{R}^2$. Em 1843, depois de tentar durante 20 anos encontrar uma multiplicação de vetores de $\mathbb{R}^3$ com propriedades análogas às dos números complexos, Hamilton descobriu a seguinte multiplicação em $\mathbb{R}^4$ que, com esta multiplicação se designa por conjunto dos *quaterniões* ou números de Hamilton.

   Escrevemos um elemento $(t, x, y, z) \in \mathbb{R}^4$ como $t + v$ com $v = (x, y, z) \in \mathbb{R}^3$. t chama-se a *parte real* do quaternião $q = t + v$ e v chama-se a *parte vetorial* de q. Define-se o produto de dois quaterniões pela fórmula[23]

$$(t + v) \cdot (s + w) = (ts - \langle v, w\rangle) + (tw + sv + v \times w) \tag{36}$$

   (a) Considere os quaterniões com parte real nula $\mathbf{i} = 0 + (1, 0, 0)$, $\mathbf{j} = 0 + (0, 1, 0)$ e $\mathbf{k} = 0 + (0, 0, 1)$. Verifique que a fórmula (36) implica as relações

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -1, \quad \mathbf{ij} = \mathbf{k} = -\mathbf{ji}, \quad \mathbf{jk} = \mathbf{i} = -\mathbf{kj}, \quad \mathbf{ki} = \mathbf{j} = -\mathbf{ik}.$$

   Reciprocamente, as relações acima juntamente com o requerimento que a multiplicação dos quaterniões com parte vetorial nula coincida com a multiplicação dos números reais e que o produto seja linear sobre os reais em cada uma das variáveis implicam a fórmula (36).

   (b) Verifique que os quaterniões da forma $a + b\mathbf{i}$ são uma cópia dos números complexos.
   (c) Verifique que a multiplicação dada por (36) é associativa e tem $1 + 0$ como elemento neutro mas não é comutativa.
   (d) Define-se o *conjugado* de $q = t + v$ por $\overline{q} = t - v$ e o comprimento (ou norma) de q por $\|q\| = \sqrt{t^2 + \|v\|^2}$. Mostre que se $q \neq 0$ então q tem um inverso multiplicativo dado pela fórmula

$$q^{-1} = \frac{\overline{q}}{\|q\|^2}$$

   Resolução.

[23]Foi esta multiplicação dos quaterniões que deu origem ao produto interno e externo de vetores de $\mathbb{R}^3$ e a formulação inicial das leis do Eletromagnetismo por Maxwell era inteiramente em termos de operações com quaterniões!

6. Endomorfismos

Vamos agora iniciar um estudo detalhado das transformações lineares $T\colon V \to V$ em que o espaço de chegada é o mesmo que o espaço de partida. Estas transformações designam-se por endomorfismos (do grego endon - dentro) porque aplicam o espaço V para dentro de si próprio.

Estas transformações desempenham um papel especialmente importante na Matemática e nas suas aplicações, em parte porque codificam simetrias (por exemplo rotações do espaço $\mathbb{R}^3$), e talvez principalmente, porque podem ser usadas para descrever evolução temporal: se o vetor $v \in V$ codificar o estado de um sistema, podemos por vezes codificar o estado desse sistema após a passagem de uma unidade de tempo por $T(v)$, após outra unidade de tempo por $T^2(v) = T(T(v))$, etc...

Exemplo 6.1. *Suponhamos que um vetor $(x, y) \in \mathbb{R}^2$ codifica o estado de um sistema e que a evolução após uma unidade de tempo é descrita pela transformação linear*

$$T(x, y) = (2x, \tfrac{y}{2})$$

Para todo o $n \in \mathbb{Z}$ temos $T^n(x, y) = (2^n x, \frac{y}{2^n})$ pelo que os estados atingidos por um sistema com estado inicial (x_0, y_0) fora dos eixos são todos pontos da hipérbole $xy = x_0 y_0$. Quando $n \to +\infty$ o estado tende para ∞ ao longo do eixo dos xx (quando x_0 é 0, converge para $(0, 0)$).

Exemplo 6.2. *(Cadeias de Markov) Suponhamos que um sistema tem n estados possíveis e que temos uma população de tais sistemas. Seja x_i a percentagem da população que se encontra no estado i e suponhamos que podemos determinar a probabilidade p_{ij} de, ao longo de uma unidade de tempo, um sistema evoluir do estado j para o estado i.*

Por exemplo os sistemas podem ser pessoas, os estados podem ser 1-viver em Lisboa; 2-viver fora de Lisboa; x_1 é então a percentagem da população que vive em Lisboa e $x_2 = 1 - x_1$. Se cada ano, 3% da população se mudar para fora de Lisboa e 1% da população exterior se mudar para Lisboa, a evolução anual da distribuição da população é codificada pela operação

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \mapsto \begin{bmatrix} 0.97 & 0.01 \\ 0.03 & 0.99 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

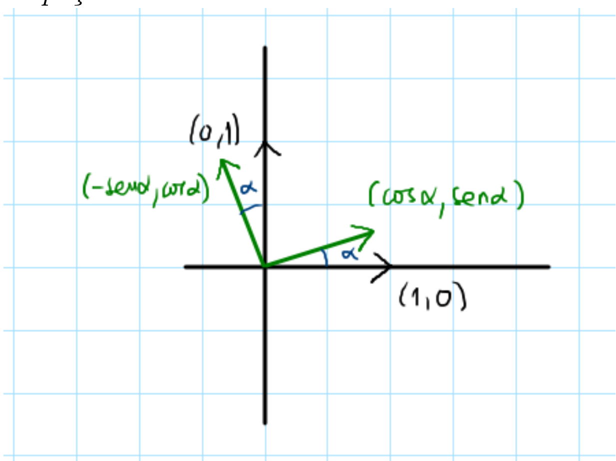
Note-se que uma matriz $n \times n$ com entradas p_{ij} construída pelo procedimento descrito acima tem entradas todas não negativas e que a soma dos elementos em cada coluna é 1. Uma tal matriz chama-se uma matriz de Markov.

6.3. **Subespaços invariantes.** Para analisar um endomorfismo $T\colon V \to V$ vamos usar uma tática habitual em matemática: decompor T em objetos da mesma natureza, mas tão simples quanto possível.

Definição 6.4. *Seja $T\colon V \to V$ um endomorfismo. Um subespaço vetorial $W \subset V$ diz-se um* subespaço invariante de T *se $T(W) \subset W$.*

Exemplo 6.5.

(i) *Se* $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ *é uma rotação de um ângulo* α *em torno de um eixo* L *que passe pela origem, tanto o eixo* L *como o plano perpendicular que passa pela origem são subespaços invariantes.*



Concretamente, no caso em que o eixo é o dos zz*, temos*

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} \mapsto \begin{bmatrix} \cos\alpha & -\operatorname{sen}\alpha & 0 \\ \operatorname{sen}\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

e claramente $W_1 = \{(x,y,0)\colon x,y \in \mathbb{R}\}$ *e* $W_2 = \{(0,0,z)\colon z \in \mathbb{R}\}$ *são subespaços invariantes.*

(ii) *Para qualquer endomorfismo* $T\colon V \to V$ *temos que* $\{0\}$ *e* V *são espaços invariantes, ditos* triviais. *Exemplos mais interessantes são o núcleo* $N(T)$ *e a imagem* $\operatorname{Im}(T)$*. De facto,* $T(N(T)) = \{0\} \subset N(T)$ *e claramente* $T(\operatorname{Im}(T)) \subset \operatorname{Im}(T)$*.*

(iii) *Dado* $v \in V$*, consideremos o conjunto* $S = \{v, T(v), T^2(v), \ldots\} \subset V$*. Então* $L(S) \subset V$ *é um subespaço invariante: de facto, dado*

$$\alpha_0 v + \ldots + \alpha_n T^n(v) \in L(S)$$

temos

$$T(\alpha_0 v + \ldots + \alpha_n T^n(v)) = \alpha_0 T(v) + \ldots + \alpha_n T^{n+1}(v) \in L(S)$$

Este espaço chama-se o subespaço cíclico *determinado pelo vetor* $v \in V$*.*

Recorde da Proposição 3.40 que um espaço V se diz a soma direta de dois subespaços $W_1, W_2 \subset V$ se $V = W_1 + W_2$ e $W_1 \cap W_2 = \{0\}$. Nesse caso escrevemos $V = W_1 \oplus W_2$. Cada vetor de V pode então decompor-se de forma única como a soma de um vetor de W_1 e de um vetor de W_2. No Exemplo 6.5 (i) acima, o espaço $\mathbb{R}^3$ é a soma direta do eixo de rotação e do plano ortogonal ao eixo.

Se conseguirmos decompor V como uma soma direta $W_1 \oplus W_2$ de espaços invariantes, então sendo $T_1 = T_{|W_1}\colon W_1 \to W_1$, $T_2 = T_{|W_2}\colon W_2 \to W_2$ as transformações induzidas por T nos subespaços invariantes teremos, em termos da decomposição única de um vetor

$v \in V$ como a soma $v = x + y$ com $x \in W_1$ e $y \in W_2$

$$T(x+y) = T_1(x) + T_2(y)$$

e a análise do comportamento de T reduz-se à análise do comportamento de T_1 e T_2. Podemos pensar neste processo como uma "separação de variáveis": o comportamento de T como função das duas variáveis x e y é completamente determinado pelo comportamento de duas funções de apenas uma variável.

Indutivamente podemos tentar decompor analogamente T_1 e T_2 até que isso deixe de ser possível, ou seja até expressar T como uma soma direta de endomorfismos "atómicos". O nosso objetivo vai ser descrever estes últimos. Os detalhes de como fazer isso dependem muito do corpo de base $\mathbb{K}$ pois, como iremos ver, este problema está fortemente relacionado com o problema de fatorizar polinómios com coeficientes no corpo. Por questões de tempo teremos de nos concentrar nos casos em que o corpo é $\mathbb{R}$ ou $\mathbb{C}$ mas tentaremos dar alguma indicação de como o problema se pode resolver em geral.

6.6. **Valores próprios e vetores próprios.** Os subespaços invariantes (não triviais) mais simples são os que têm dimensão 1. Se v for um elemento não nulo de um tal espaço, teremos necessariamente $T(v) = \lambda v$ para algum escalar λ.

Definição 6.7. *Seja V um espaço vetorial sobre o corpo $\mathbb{K}$ e $T\colon V \to V$ um endomorfismo. Um vetor não nulo $v \in V \setminus \{0\}$ diz-se um* vetor próprio de T *se existe $\lambda \in \mathbb{K}$ tal que $T(v) = \lambda v$. Nesse caso λ diz-se um* valor próprio de T *associado ao vetor próprio v.*

Também podemos falar de valores e vetores próprios de uma matriz $A \in M_{n\times n}(\mathbb{K})$: são os vetores e valores próprios do endomorfismo de $\mathbb{K}^n$ que é representado na base canónica por A.

Exemplo 6.8. *(i) Um vetor não nulo segundo um eixo de uma rotação de $\mathbb{R}^3$ é um vetor próprio com valor próprio 1.*

(ii) Se $P\colon V \to V$ é uma projeção (isto é se $P^2 = P$ cf. Exercício 4.25) então todos os vetores não nulos do plano de projeção $\operatorname{Im}(P)$ são vetores próprios com valor próprio 1.

(iii) Os vetores não nulos do núcleo $N(T)$ (se existirem) são vetores próprios de T com valor próprio 0.

(iv) Seja $V = C^\infty(\mathbb{R})$ o espaço vetorial das funções reais de variável real indefinidamente diferenciáveis e $D\colon V \to V$ a operação de derivação definida por $D(f) = f'$. Todos os números reais λ são valores próprios de D. Os vetores próprios correspondentes a λ são as funções exponenciais $ce^{\lambda t}$ com $c \neq 0$ (cf. Exercício 4.36).

Uma vez que

$$T(v) = \lambda v \Leftrightarrow T(v) = \lambda \operatorname{Id}(v) \Leftrightarrow (T - \lambda \operatorname{Id})(v) = 0$$

vemos que λ é um valor próprio se e só se $N(T - \lambda \operatorname{Id}) \neq \{0\}$. Desde que V tenha dimensão finita, é possível determinar os valores próprios e vetores próprios recorrendo ao determinante: escolhendo uma base B para V e considerando a matriz $A = A_{T,B,B}$ a condição acima traduz-se em

$$N(A-\lambda I) \neq \{0\} \Leftrightarrow (A-\lambda I) \text{ não é invertível} \Leftrightarrow \det(A-\lambda I)=0$$

Quando $\det(A-\lambda I)=0$, os elementos não nulos de $N(A-\lambda I)$ corresponderão aos vetores próprios associados a λ.

Exemplo 6.9. *Consideremos a transformação linear* $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ *definida por* $T(x,y) = (x+2y, 2x+y)$*. Considerando a base canónica temos*

$$A-\lambda I = \begin{bmatrix} 1-\lambda & 2 \\ 2 & 1-\lambda \end{bmatrix}$$

Esta matriz não é invertível exatamente quando

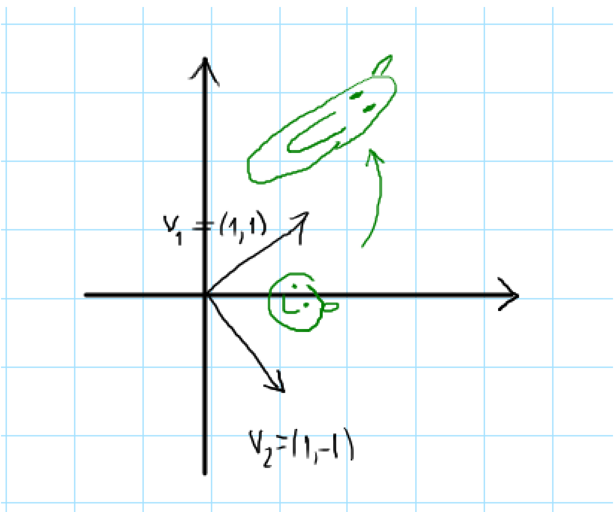
$$\begin{vmatrix} 1-\lambda & 2 \\ 2 & 1-\lambda \end{vmatrix} = 0 \Leftrightarrow (1-\lambda)^2-4=0 \Leftrightarrow \lambda=-1 \text{ ou } \lambda=3.$$

São estes os valores próprios de T*. Os vetores próprios de* $\lambda=-1$ *são os elementos não nulos de* $N(A+I)$*:*

$$(A+I)v=0 \Leftrightarrow \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = 0 \Leftrightarrow b=-a$$

ou seja, os vetores da forma $a(1,-1)$ *com* $a \neq 0$*. Analogamente, os vetores próprios de* $\lambda = 3$ *(os elementos não nulos do núcleo de* $A-3I$*) são os vetores da forma* $a(1,1)$ *com* $a \neq 0$*.*

Os vetores $v_1=(1,1)$ *e* $v_2=(1,-1)$ *formam uma base de* $\mathbb{R}^2$ *em termos da qual é extremamente simples compreender o efeito que a transformação linear* T *tem sobre os vetores de* $\mathbb{R}^2$*: Ao longo da direção de* v_1 *(a diagonal do primeiro quadrante)* T *expande por um fator de 3, enquanto que na direção ortogonal, (a diagonal do quarto quadrante),* T *reflete. Com base nisto é fácil descrever o efeito que* T *teria num desenho qualquer no plano (ver figura).*



Note-se ainda que, uma vez que $T(v_1) = 3v_1$ e $T(v_2) = -v_2$, a representação matricial de T com respeito à base $B = (v_1, v_2)$ é diagonal:

$$A_{T,B,B} = \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix}$$

Definição 6.10. *Um endomorfismo $T\colon V \to V$ diz-se* diagonalizável *se existe uma base para V constituída por vetores próprios de T. Uma matriz $A \in M_{n\times n}(\mathbb{K})$ diz-se diagonalizável, se a transformação linear de $\mathbb{K}^n$ representada por A (com respeito à base canónica) é diagonalizável.*

A razão da palavra diagonalizável é, claro, que a representação de uma transformação linear diagonalizável numa base $B = (v_1, \ldots, v_n)$ formada por vetores próprios é uma matriz diagonal

$$\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}$$

onde λ_i é o valor próprio associado a v_i. Note-se que os valores próprios não são necessariamente distintos dois a dois.

Uma matriz quadrada A é diagonalizável se existe uma matriz invertível S (uma matriz de mudança de coordenadas de uma base formada por vetores próprios de A para a base canónica cf. Exercício 4.24(a)) tal que $A = SDS^{-1}$ com D uma matriz diagonal (que tem como entradas não nulas valores próprios de A). Diz-se que A é *semelhante* a uma matriz diagonal.

Recorde-se a nossa estratégia de decompor um endomorfismo T como uma soma direta de endomorfismos "atómicos". Um endomorfismo $T\colon V \to V$ é diagonalizável se V se decompõe numa soma direta de espaços invariantes de dimensão 1. Neste caso a nossa estratégia é bem sucedida.

Pode, no entanto, haver algo de arbitrário na decomposição assim obtida. Se na base formada por vetores próprios houver mais do que um vetor próprio associado a um valor próprio λ podemos substituir esses vetores por qualquer outra base para o subespaço que eles geram.

Definição 6.11. *Seja $T\colon V \to V$ um endomorfismo. O* espaço próprio *associado ao valor próprio λ de T é*

$$E(\lambda) = N(T - \lambda\,\mathrm{Id})$$

A dimensão de $E(\lambda)$ chama-se a multiplicidade geométrica *de λ.*

O espaço próprio de λ é o conjunto de todos os vetores próprios de λ juntamente com o vetor 0. A multiplicidade geométrica de λ é o número máximo de vetores próprios de λ linearmente independentes que conseguimos obter. Note-se que $E(0)$ é o núcleo de T.

Pela Proposição 3.40, um endomorfismo T de um espaço vetorial finitamente gerado é diagonalizável se e só se existem valores próprios $\lambda_1, \ldots, \lambda_k$ *distintos* de T tais que

$$V = E(\lambda_1) \oplus \ldots \oplus E(\lambda_k)$$

A restrição de T a cada um dos espaços $E(\lambda)$ é simplesmente a operação de multiplicação por λ. Quando alguma das multiplicidades geométricas é maior do que um, esta decomposição é mais natural que a decomposição em espaços de dimensão 1 determinada por uma escolha de bases para cada um dos espaços próprios.

Exemplo 6.12. *Seja* $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ *a transformação linear definida por*

$$T(x, y, z) = (3x - y + z, 2y, x - y + 3z)$$

Sendo A *a matriz que representa* T *na base canónica temos*

$$\begin{aligned}\det(A - \lambda I) &= \begin{vmatrix} 3-\lambda & -1 & 1 \\ 0 & 2-\lambda & 0 \\ 1 & -1 & 3-\lambda \end{vmatrix} \\ &= (3-\lambda)^2(2-\lambda) - (2-\lambda) = (2-\lambda)((\lambda-3)^2 - 1)\end{aligned}$$

Portanto a matriz $A - \lambda I$ *tem característica* < 3 *se e só se* $\lambda = 2$ *ou* $(\lambda - 3)^2 = 1 \Leftrightarrow \lambda = 2$ *ou* $\lambda = 4$. *Os valores próprios são portanto* 2 *e* 4. *Os vetores próprios de* 2 *são as soluções de* $(A - 2I)v = 0$*:*

$$\begin{bmatrix} 1 & -1 & 1 \\ 0 & 0 & 0 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 0 \Leftrightarrow b = a + c$$

Os vetores próprios de 2 *são portanto os vetores não nulos da forma* $(a, a + c, c)$ *pelo que* $E(2) = L(\{(1, 1, 0), (0, 1, 1)\})$. *Os vetores próprios de* 4 *são as soluções de*

$$\begin{bmatrix} -1 & -1 & 1 \\ 0 & -2 & 0 \\ 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 0 \Leftrightarrow \begin{cases} b = 0 \\ c = a \end{cases}$$

Temos portanto $E(4) = L(\{(1, 0, 1)\})$. *Escrevendo* $v_1 = (1, 1, 0), v_2 = (0, 1, 1)$ *e* $v_3 = (1, 0, 1)$ *e tomando* $B = (v_1, v_2, v_3)$ *temos que*

$$A_{T,B.B} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

Temos $\mathbb{R}^3 = E(2) \oplus E(4)$. *No subespaço invariante* $E(2)$ *o efeito de* T *é multiplicar por* 2, *enquanto que em* $E(4)$ *é multiplicar por* 4. *A multiplicidade geométrica do valor próprio* 2 *é* 2, *enquanto que a multiplicidade geométrica do valor próprio* 4 *é* 1.

Exemplo 6.13. *Uma projeção, isto é um endomorfismo* $P\colon V \to V$ *tal que* $P^2 = P$, *é diagonalizável. De facto, vimos no Exercício 4.25 que* V *se decompõe como a soma direta* $N(P) \oplus \mathrm{Im}(P)$. *Uma vez que* $N(P)$ *é o espaço próprio de* 0 *e que* $\mathrm{Im}(P)$ *é o espaço próprio de* 1 *temos* $V = E(0) \oplus E(1)$. *A multiplicidade geométrica de* 0 *é a dimensão do núcleo (ou nulidade) de* P. *A multiplicidade geométrica de* 1 *é a dimensão de* $\mathrm{Im}(P)$ *(ou característica de* P*).*

Infelizmente, nem sempre é possível diagonalizar uma transformação linear. Na realidade, nem é garantido que exista algum vetor próprio!

Exemplo 6.14. *É geometricamente óbvio (e fácil de confirmar algebricamente) que, desde que α não seja um múltiplo de π, a rotação $R_\alpha \colon \mathbb{R}^2 \to \mathbb{R}^2$ determinada (na base canónica) pela matriz*

$$\begin{bmatrix} \cos\alpha & -\operatorname{sen}\alpha \\ \operatorname{sen}\alpha & \cos\alpha \end{bmatrix}$$

não tem qualquer vetor próprio (não há nenhuma direção do plano que permaneça invariante pela rotação).

6.15. **Existência de valores próprios.** É no entanto um resultado básico da Álgebra Linear que um endomorfismo de um espaço vetorial complexo (de dimensão finita) tem sempre um valor próprio. Esta afirmação é uma consequência do seguinte Teorema cuja demonstração, por involver Análise (ou então Álgebra mais avançada), será omitida. As pessoas de Matemática verão uma demonstração extremamente elegante para o ano na cadeira de Introdução à Análise Complexa.

Teorema 6.16 (Teorema Fundamental da Álgebra)**.** *Todo o polinómio não constante*

$$p(x) = a_0 + a_1 x + \ldots + a_k x^k \tag{37}$$

com coeficientes $a_i \in \mathbb{C}$ tem uma raíz complexa. Isto é, existe $z_0 \in \mathbb{C}$ tal que $p(z_0) = 0$.

Observação 6.17. *(i) O Teorema Fundamental da Álgebra 6.16 implica facilmente por indução que, assumindo $a_k \neq 0$, o polinómio* (37) *pode ser escrito de forma única a menos de troca de ordem dos fatores na forma*

$$a_k(x-\lambda_1)^{n_1}(x-\lambda_2)^{n_2}\cdots(x-\lambda_k)^{n_k} \tag{38}$$

com $\lambda_i \in \mathbb{C}$ distintos, e n_i números naturais. Os números λ_i na expressão (38) *são as raízes do polinómio $p(x)$. O expoente n_i diz-se a* multiplicidade *da raíz λ_i.*

Vemos assim que o Teorema Fundamental da Álgebra é análogo ao Teorema Fundamental da Aritmética que diz que qualquer número natural maior do que 1 *se pode escrever de forma única como um produto de potências de números primos a menos de troca da ordem dos fatores.*

(ii) Um corpo $\mathbb{K}$ diz-se algebricamente fechado *se todo o polinómio de grau > 0 com coeficientes em $\mathbb{K}$ tem uma raíz em $\mathbb{K}$. O Teorema 6.16 afirma que $\mathbb{C}$ é algebricamente fechado. Todos os resultados desta secção enunciados para espaços vetoriais complexos são verdade mais geralmente (com a mesma demonstração) para espaços vetoriais sobre corpos algebricamente fechados.*

Teorema 6.18. *Seja $V \neq \{0\}$ um espaço vetorial complexo de dimensão finita e $T \colon V \to V$ um endomorfismo. Então T tem um valor próprio.*

Dem. Vamos dar duas demonstrações, uma usando o determinante e outra que tem interesse conceptual e não usa o determinante.
Primeira demonstração: Seja $B = (v_1, \ldots, v_n)$ uma base para V e $A = A_{T,B,B} \in M_{n\times n}(\mathbb{C})$ a representação matricial de T com respeito a esta base. Os valores próprios de

T são os valores de λ para os quais $\det(A - \lambda I) = 0$. Temos

$$\det(A - \lambda I) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix}$$

Se aplicarmos a fórmula (30) para o determinante, cada parcela nesta fórmula é um polinómio em λ (de grau $\leq n$). A única parcela de grau exatamente n é o produto das entradas ao longo da diagonal principal $(a_{11} - \lambda) \cdots (a_{nn} - \lambda)$. O coeficiente de λ^n é $(-1)^n$ pelo que o termo de grau n no polinómio $p(\lambda) = \det(A - \lambda I)$ é $(-1)^n \lambda^n$. Em particular $p(\lambda)$ não é um polinómio constante. O Teorema Fundamental da Álgebra 6.16 garante que $p(\lambda)$ tem pelo menos uma raíz, pelo que T tem pelo menos um valor próprio.
Segunda demonstração: Seja $v \in V \setminus \{0\}$ um vetor não nulo. Seja k o maior inteiro não negativo tal que

$$S_k = \{v, Tv, \ldots, T^k(v)\}$$

é um conjunto linearmente independente de vetores distintos. Este número existe porque $S_0 = \{v\}$ é linearmente independente (logo o conjunto dos inteiros k para os quais S_k é linearmente independente é não vazio) e porque todo o subconjunto de V com $1 + \dim V$ elementos é linearmente dependente (logo $k \leq \dim(V) - 1$). Temos então

$$T^{k+1}(v) = a_0 v + a_1 T(v) + \ldots + a_k T^k(v)$$

para alguns $a_i \in \mathbb{C}$. Podemos escrever a expressão anterior na forma

$$(T^{k+1} - a_k T^k - \ldots - a_1 T - a_0 \operatorname{Id})(v) = 0 \tag{39}$$

Seja λ uma raíz do polinómio $p(x) = x^{k+1} - a_k x^k - \ldots - a_1 x - a_0$ (que existe pelo Teorema Fundamental da Álgebra). Então

$$p(x) = x^{k+1} - a_k x^k - \ldots - a_1 x - a_0 = (x - \lambda)(x^k + b_{k-1} x^{k-1} + \ldots + b_1 x + b_0)$$

para alguns $b_0, \ldots, b_{k-1} \in \mathbb{C}$. Aplicando esta fatorização a (39) (o que podemos fazer pela distributividade da composição de transformações lineares com respeito à soma cf. Proposição 4.21) obtemos

$$(T - \lambda \operatorname{Id})(T^k + b_{k-1} T^{k-1} + \ldots + b_1 T + b_0 \operatorname{Id})(v) = 0$$

Seja $w = (T^k + b_{k-1} T^{k-1} + \ldots + b_1 T + b_0 \operatorname{Id})(v) = T^k(v) + b_{k-1} T^{k-1}(v) + \ldots + b_0 v$. Como S_k é linearmente independente temos que $w \neq 0$. Uma vez que $(T - \lambda \operatorname{Id}) w = 0$ conclui-se que λ é um valor próprio de T. □

Exemplo 6.19. *Consideremos o endomorfismo de $\mathbb{C}^2$ determinado pela matriz R_α do Exemplo 6.14. Recorde-se que estamos a assuir que* $\operatorname{sen} \alpha \neq 0$. *Temos*

$$\det(R_\alpha - \lambda I) = \begin{vmatrix} \cos(\alpha) - \lambda & -\operatorname{sen} \alpha \\ \operatorname{sen} \alpha & \cos(\alpha) - \lambda \end{vmatrix} = (\cos \alpha - \lambda)^2 + \operatorname{sen}^2 \alpha$$

Os valores próprios de R_α são as raízes deste polinómio do segundo grau:

$$\lambda = \cos \alpha \pm i \operatorname{sen} \alpha$$

Um vetor próprio de $\cos\alpha + i\operatorname{sen}\alpha$ *é um elemento não nulo do núcleo de* $R_\alpha - (\cos\alpha + i\operatorname{sen}\alpha)\operatorname{Id}$*:*

$$(R_\alpha - (\cos\alpha + i\operatorname{sen}\alpha)\operatorname{Id})(v) = 0 \Leftrightarrow \begin{bmatrix} -i\operatorname{sen}\alpha & -\operatorname{sen}\alpha \\ \operatorname{sen}\alpha & -i\operatorname{sen}\alpha \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = 0 \Leftrightarrow b = -ia$$

O espaço próprio $E(\cos\alpha + i\operatorname{sen}\alpha)$ *é portanto gerado pelo vetor* $(1, -i)$*. Da mesma forma vemos que* $(1, i)$ *gera o espaço próprio* $E(\cos\alpha - i\operatorname{sen}\alpha)$*.*

Observação 6.20. *Não é uma coincidência que os valores próprios e vetores próprios do exemplo anterior sejam conjugados. Dados dois números complexos* z_1, z_2*, temos* $\overline{z_1 z_2} = \overline{z_1}\,\overline{z_2}$ *e* $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$*, logo, se* A *é uma matriz* $n \times n$ *real (portanto* $\overline{A} = A$*) e* $v \in \mathbb{C}^n$ *é tal que* $Av = \lambda v$ *então* $A\overline{v} = \overline{\lambda}\overline{v}$*. Assim, os valores (e vetores) próprios complexos de um endomorfismo de* $\mathbb{C}^n$ *representado por uma matriz com entradas reais ocorrem em pares conjugados. Esta observação facilita o cálculo dos valores e vetores próprios nesta situação.*

Vejamos um exemplo concreto do procedimento utilizado na segunda demonstração do Teorema 6.18.

Exemplo 6.21. *Consideremos a transformação linear* $T(x, y) = (x + 2y, 2x + y)$*. Considerando o vetor* $v = (1, 0)$ *temos* $T(v) = (1, 2)$*. Assim* v *e* $T(v)$ *formam uma base para* $\mathbb{R}^2$*. Uma vez que* $T^2(v) = T(1, 2) = (5, 4) = 3v + 2T(v)$ *temos*

$$(T^2 - 2T - 3\operatorname{Id})(v) = 0$$

As raízes do polinómio $x^2 - 2x - 3$ *são* $x = 3$ *e* $x = -1$*. A equação acima pode portanto escrever-se na forma*

$$(T - 3\operatorname{Id})(T + \operatorname{Id})(v) = 0 \qquad \text{ou equivalentemente } (T + \operatorname{Id})(T - 3\operatorname{Id})(v) = 0$$

Segue-se que $(T + \operatorname{Id})(1, 0) = (1, 2) + (1, 0) = (2, 2)$ *é um vetor próprio com valor próprio* 3 *e que* $(T - 3\operatorname{Id})(1, 0) = (1, 2) - 3(1, 0) = (-2, 2)$ *é um vetor próprio com valor próprio* -1*.*

Experimente fazer a conta acima com outro vetor inicial v*. O que conclui?*

Observação 6.22. *Vale a pena sublinhar outra ideia utilizada na segunda demonstração do Teorema 6.18. Seja* V *um espaço vetorial sobre o corpo* $\mathbb{K}$ *e* $T\colon V \to V$ *um endomorfismo. Dado um polinómio* $p(x) = a_0 + a_1 x + \ldots + a_k x^k$ *com coeficientes* $a_i \in \mathbb{K}$*,* $p(x)$ *determina um endomorfismo*

$$p(T) = a_0 \operatorname{Id} + a_1 T + \ldots + a_k T^k \colon V \to V$$

Uma vez que a composição de transformações lineares é linear em cada argumento e as potências de T *comutam entre si, uma fatorização* $p(x) = q(x)r(x)$ *determina uma fatorização* $p(T) = q(T) \circ r(T)$ *da transformação linear* $p(T)$*.*

6.23. **O determinante de um endomorfismo e o polinómio característico.** A primeira demonstração do Teorema 6.18 fez uso de um polinómio associado a um endomorfismo de um espaço vetorial de dimensão finita V. Começamos por notar que este polinómio é independente da escolha de base B usada na sua definição.

Definição 6.24. *Seja V um espaço vetorial de dimensão finita e $T\colon V \to V$ um endomorfismo. Sendo B uma base ordenada qualquer de V definimos o* determinante de T *por*

$$\det(T) = \det(A_{T,B,B})$$

Temos que verificar que o escalar $\det(T)$ é independente da escolha de B: se B' é outra base para V e $S = S_{B\to B'}$ a matriz de mudança de coordenadas então

$$A_{T,B',B'} = SA_{T,B,B}S^{-1}$$

e portanto

$$\det(A_{T,B',B'}) = \det(S)\det(A_{T,B,B})\det(S^{-1}) = \det(S)\det(A_{T,B,B})\frac{1}{\det(S)} = \det(A_{T,B,B})$$

Infelizmente, não seria neste momento fácil explicar-vos como definir intrinsecamente o determinante de um endomorfismo sem apelar às representações matriciais pelo que a Definição 6.24 terá de servir.

Uma vez que o determinante deteta a invertibilidade de um endomorfismo (T é invertível se e só se $A_{T,B,B}$ é invertível para qualquer base B se e só se $\det A_{T,B,B} \neq 0$) podemos usar o determinante para calcular os valores próprios de um endomorfismo T, como foi feito na primeira demonstração do Teorema 6.18.

Definição 6.25. *Seja V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{K}$ e $T\colon V \to V$ um endomorfismo. O* polinómio característico *de T é o polinómio (com coeficientes em $\mathbb{K}$) definido por*

$$p(\lambda) = \det(T - \lambda\,\mathrm{Id})$$

O polinómio característico de uma matriz quadrada $A \in M_{n\times n}(\mathbb{K})$ é o polinómio característico da transformação linear representada por A na base canónica de $\mathbb{K}^n$, ou seja,

$$p(\lambda) = \det(A - \lambda I)$$

Note-se que o termo constante do polinómio característico (que é igual a $p(0)$) é o determinante de T.

Exemplo 6.26. *O polinómio característico da matriz*

$$A = \begin{bmatrix} \frac{5}{2} & 0 & -\frac{1}{2} & 0 \\ \frac{1}{2} & 1 & -\frac{1}{2} & 2 \\ \frac{1}{2} & 0 & \frac{3}{2} & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

é

$$\begin{aligned} p(\lambda) &= \begin{vmatrix} \frac{5}{2}-\lambda & 0 & -\frac{1}{2} & 0 \\ \frac{1}{2} & 1-\lambda & -\frac{1}{2} & 2 \\ \frac{1}{2} & 0 & \frac{3}{2}-\lambda & 0 \\ 0 & 0 & 0 & -1-\lambda \end{vmatrix} = -(1+\lambda)(1-\lambda) \begin{vmatrix} \frac{5}{2}-\lambda & -\frac{1}{2} \\ \frac{1}{2} & \frac{3}{2}-\lambda \end{vmatrix} \\ &= (\lambda^2-1)(\lambda^2-4\lambda+4) = (\lambda^2-1)(\lambda-2)^2 \end{aligned}$$

Este polinómio anula-se quando $\lambda = 2$ ou $\lambda = \pm 1$. São portanto estes os valores próprios de A.

Registamos um par de propriedades imediatas do polinómio característico de um endomorfismo de um espaço vetorial de dimensão finita (sobre um corpo qualquer, não necessariamente o dos números complexos).

Proposição 6.27. *Seja V um espaço vetorial de dimensão finita, $T\colon V \to V$ um endomorfismo, e $p(\lambda)$ o seu polinómio característico.*

(a) Sendo $n = \dim V$, o polinómio característico é da forma

$$p(\lambda) = (-1)^n\lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_0$$

e portanto tem grau n.

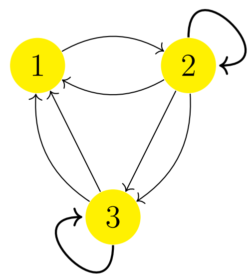
(b) As raízes do polinómio característico de T são os valores próprios de T.

Dem. (a) O argumento utilizado na primeira demonstração do Teorema 6.18 aplica-se literalmente.

(b) Conforme já explicámos, λ é um valor próprio se e só se $(T - \lambda\,\mathrm{Id})$ não é invertível, o que acontece sse $p(\lambda) = 0$.

□

6.28. **O algoritmo PageRank.** Vamos agora fazer um pequeno interlúdio para discutir uma aplicação famosa do conceito de vetor próprio. Consideremos uma internet com apenas três páginas ligadas de acordo com o diagrama



Como no Exemplo 6.2 podemos considerar a população de todos os internautas e atribuir a cada um o estado i num dado instante se estiverem a visitar a página i. Sejam n_1, n_2 e n_3 o número de pessoas em cada página num dado instante. Se cada pessoa clicar num link ao acaso em cada página, a distribuição das pessoas pelas páginas no instante seguinte seria

$$\begin{bmatrix} 0 & \frac{1}{4} & \frac{2}{3} \\ 1 & \frac{1}{4} & 0 \\ 0 & \frac{1}{2} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix}$$

A entrada ij da matriz é a probabilidade de uma internauta que está na página j carregar numa ligação que a leva à página i, e é portanto igual a $\frac{\ell(j,i)}{\ell(j)}$ onde $\ell(j,i)$ é o número de ligações que une a página j à página i e $\ell(j)$ é o número de total de ligações de j para outras páginas.[24]

Esta matriz que descreve a evolução temporal do nosso sistema é uma matriz de Markov (cf. Exemplo 6.2): as entradas são não negativas porque codificam probabilidades e a soma das entradas de cada coluna é 1 (porque é a soma das probabilidades de ir parar a cada destino possível partindo da página correspondente à coluna).

Quando é que o número de internautas em cada página permanece constante ao longo do tempo? Quando o vetor (n_1, n_2, n_3) é um vetor próprio da matriz

$$\begin{bmatrix} 0 & \frac{1}{4} & \frac{2}{3} \\ 1 & \frac{1}{4} & 0 \\ 0 & \frac{1}{2} & \frac{1}{3} \end{bmatrix} \tag{40}$$

com valor próprio 1. Um tal vetor próprio existe necessariamente porque a soma das entradas de cada coluna é um! De facto, isto significa exatamente que $(1, 1, 1)$ é um vetor próprio da matriz transposta, com valor próprio 1. Uma vez que as dimensões dos espaços das linhas e colunas de uma matriz quadrada coincidem, a característica de $A - \lambda\,\mathrm{Id}$ e de $(A - \lambda\,\mathrm{Id})^T = A^T - \lambda\,\mathrm{Id}$ coincidem. Isto implica que os conjuntos de valores próprios de A e A^T coincidem.

Pode mostrar-se que existe necessariamente um vetor próprio de 1 com componentes todas não negativas, e (com bastante generalidade) que se normalizarmos os vetores que indicam o estado das páginas de modo a que a soma das entradas[25] seja 1, o limite quando o tempo tende para $+\infty$ do estado do sistema é o vetor próprio de 1 (normalizado), que é único.

Mais precisamente, se A é a matriz (40) que controla a transição entre estados e (p_1, p_2, p_3) é um estado inicial qualquer (com $p_i \geq 0$ e $p_1 + p_2 + p_3 = 1$), temos

$$\lim_{k\to\infty} A^k \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} = v$$

com v o único vetor próprio de 1 com entradas não negativas cuja soma é 1. Pode mostrar-se que (com grande generalidade) o significado das componentes de v é a seguinte: v_i é a percentagem do tempo que uma internauta surfando ao acaso naquelas páginas passaria

[24] Se uma página não tem ligações para outras assume-se que tem uma ligação para cada página.

[25] Isto corresponde a considerar a percentagem dos internautas em cada página em vez do número absoluto.

na página i. É este número que é usado como medida da relevância da página i - o seu PageRank.

No exemplo acima teríamos que os vetores próprios de 1 da matriz (40) são as soluções de

$$(A - I_3)v = 0 \Leftrightarrow \begin{bmatrix} -1 & \frac{1}{4} & \frac{2}{3} \\ 1 & -\frac{3}{4} & 0 \\ 0 & \frac{1}{2} & -\frac{2}{3} \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Leftrightarrow \begin{cases} a = \frac{3}{4}b \\ c = \frac{3}{4}b \end{cases}$$

Logo um vetor próprio de 1 é um múltiplo não nulo de $(\frac{3}{4}, 1, \frac{3}{4})$. Normalizando obtemos

$$(0.3, 0.4, 0.3)$$

pelo que a página mais relevante é a página 2, sendo as outras duas igualmente relevantes. Uma internauta surfando aleatoriamente entre estas três páginas passaria 40% do seu tempo na página 2 e 30% em cada uma das outras duas páginas.

O algoritmo utilizado pelo Google para ordenar as páginas por relevância é seguramente muito mais complicado mas o princípio básico é o que foi explicado acima. Ao pesquisarmos um termo, o algoritmo começa por selecionar as páginas relacionadas com esse termo (utilizando as etiquetas previamente atribuídas a cada página) e analisa depois as ligações entre essas páginas conforme descrito acima, listando-as por ordem de relevância.

Na realidade, no algoritmo original de Larry Page e Sergey Brin é também levada em conta a possibilidade de uma internauta não seguir nenhum link na página em que se encontra (e em vez disso usar um bookmark ou escrever diretamente um URL). Esta possibilidade é considerada atribuindo uma probabilidade d de ir para qualquer outra página da internet a partir de uma dada página, sendo $(1 - d)$ a probabilidade de carregar numa das ligações da página. O parâmetro d é medido experimentalmente (e é cerca de 15%). Tente descrever analiticamente este algoritmo modificado. A solução encontra-se na página da Wikipedia do algoritmo PageRank.

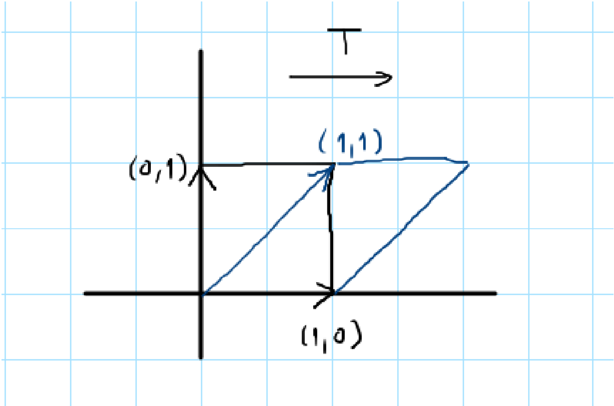
6.29. **Vetores próprios generalizados.** Apesar de todo o endomorfismo $T\colon V \to V$ de um espaço vetorial complexo ter pelo menos um valor próprio, não é em geral possível encontrar uma base para T formada por vetores próprios.

Exemplo 6.30 (A aplicação de "shear"ou deslizamento)**.** *Seja $T\colon \mathbb{C}^2 \to \mathbb{C}^2$ a transformação linear definida por $T(x, y) = (x + y, y)$. Este endomorfismo é representado na base canónica pela matriz*

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

É imediato verificar que o único valor próprio de A é 1 pelo que A não é diagonalizável (uma transformação diagonalizável com um único valor próprio é necessariamente um múltiplo escalar da identidade). O espaço próprio $E(1)$ é claramente o primeiro eixo coordenado $\{(x, 0)\colon x \in \mathbb{C}\}$ pelo que o valor próprio 1 tem multiplicidade geométrica 1.

Podemos no entanto observar que $(T - \mathrm{Id})(0, 1) = (1, 0)$ e portanto $(T - \mathrm{Id})^2(1, 0) = 0$.



Geometricamente, o efeito que T tem num vetor de $\mathbb{R}^2$ é deslizá-lo ao longo do eixo dos xx (o eixo gerado pelo vetor próprio) uma distância igual à sua ordenada.

O principal Teorema sobre endomorfismos complexos afirma que qualquer endomorfismo não diagonalizável se pode decompor nestas aplicações de "shear"ou deslizamento.

Notação 6.31. *A partir de agora, para simplificar a notação usaremos muitas vezes o escalar λ para denotar a transformação linear $\lambda\,\mathrm{Id}$. Assim, por exemplo, $(T-\lambda\,\mathrm{Id})$ aparecerá como $(T-\lambda)$. Também omitiremos alguns parêntesis quando a sua omissão não cause confusão: por exemplo um vetor $T(v)$ poderá ser denotado simplesmente por Tv.*

Definição 6.32. *Seja $T\colon V\to V$ um endomorfismo e λ um valor próprio de T. O* espaço próprio generalizado *de λ é*

$$E^g(\lambda)=\{v\in V\colon (T-\lambda)^k v=0 \text{ para algum } k\}$$

Os elementos não nulos de $E^g(\lambda)$ designam-se por vetores próprios generalizados *associados ao valor próprio λ. A dimensão de $E^g(\lambda)$ chama-se a* multiplicidade algébrica *de λ.*

Note-se que $E(\lambda)\subset E^g(\lambda)$. É fácil verificar que $E^g(\lambda)$ é de facto um subespaço vetorial de V: claramente $0\in E^g(\lambda)$ e $E^g(\lambda)$ é fechado para o produto por escalar. Vejamos que também é fechado para a soma: se $v_1,v_2\in E^g(\lambda)$ então existem k_1,k_2 tais que $(T-\lambda)^{k_1}(v_1)=0$ e $(T-\lambda)^{k_2}(v_2)=0$. Seja k o máximo de $\{k_1,k_2\}$. Então

$$(T-\lambda)^k(v_1+v_2)=(T-\lambda)^{k-k_1}(T-\lambda)^{k_1}(v_1)+(T-\lambda)^{k-k_2}(T-\lambda)^{k_2}(v_2)=0+0=0$$

Finalmente, observe-se que $E^g(\lambda)$ é um subespaço invariante: dado $v\in E^g(\lambda)$, existe k tal que $(T-\lambda)^k(v)=0$. Logo

$$(T-\lambda)^k(Tv)=T(T-\lambda)^k(v)=T(0)=0 \qquad\Rightarrow\qquad Tv\in E^g(\lambda) \tag{41}$$

Exemplo 6.33. *Aplicando a nossa nova terminologia ao Exemplo 6.30 vemos que nesse caso $E^g(1)=\mathbb{C}^2$ é estritamente maior do que $E(1)=L(\{(1,0)\})$. Logo a multiplicidade algébrica de $\lambda=1$ é 2, enquanto que a multiplicidade geométrica é apenas 1.*

Exemplo 6.34. *Consideremos a transformação linear* $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ *representada com respeito à base canónica pela matriz*

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & -1 \\ 1 & 1 & 1 \end{bmatrix}$$

Calculando os valores próprios vemos que estes são $\lambda = 1$ *e* $\lambda = 2$. *Os vectores próprios de* 1 *são as soluções da equação* $(A - I)v = 0$, *ou seja*

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & -1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 0 \Leftrightarrow \begin{cases} a+b=0 \\ a+b-c=0 \\ a+b=0 \end{cases} \Leftrightarrow \begin{cases} b=-a \\ c=0 \end{cases}$$

Assim $(1,-1,0)$ *forma uma base para o espaço próprio de* 1. *Uma vez que*

$$(A-I)^2 = \begin{bmatrix} 2 & 2 & -1 \\ 1 & 1 & -1 \\ 2 & 2 & -1 \end{bmatrix}$$

tem característica 2, *o seu núcleo tem dimensão* 1 *e portanto consiste no espaço próprio de* 1. *Podemos facilmente verificar que o mesmo acontece com as matrizes* $(A-I)^k$ *para todos os valores de* k *para* $k \geq 2$ *pelo que os únicos vetores próprios generalizados de* 1 *são os vetores próprios de* 1. *Na realidade, iremos ver em breve no Lema 6.48, o facto de* $N((A-I)^2) = N(A-I)$ *implica que* $N((A-I)^k) = N(A-I)$ *para todo o* k, *e não há portanto necessidade de realizar mais cálculos.*

Temos assim que a multiplicidade algébrica de 1 *é igual à multiplicidade geométrica, que é* 1.

Os vectores próprios de 2 *são as soluções da equação*

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & -1 \\ 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 0 \Leftrightarrow \begin{cases} b=0 \\ a-c=0 \\ a+b-c=0 \end{cases} \Leftrightarrow \begin{cases} b=0 \\ a=c \end{cases}$$

Uma base para o espaço próprio de 2 *é* $(1,0,1)$ *e portanto a multiplicidade geométrica de* 2 *é apenas* 1. *Isto significa que a matriz* A *não é diagonalizável (não conseguimos achar mais do que dois vetores próprios independentes).*

No entanto

$$(A-2I)^2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & -1 \\ 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & -1 \\ 1 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 \\ -1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

pelo que $(A-2I)^2 v = 0 \Leftrightarrow v \in L(\{(1,0,1),(0,1,0)\})$. *Claramente* $(A-2I)^k = (A-2I)^2$ *para* $k \geq 2$ *pelo que não há mais vetores próprios generalizados de* 2. *Veremos mais à frente que este último cálculo é na realidade desnecessário uma vez que a soma das dimensões dos espaços próprios generalizados é sempre menor ou igual à dimensão do espaço vetorial. Concluímos que* $E^g(2) = L(\{(1,0,1),(0,1,0)\})$. *Temos assim que* *a multiplicidade algébrica de 2 é igual a 2. Note-se que juntando as bases dos espaços próprios generalizados de 1 e 2 obtemos uma base de $\mathbb{R}^3$.*

Para utilização posterior notamos que, quando V tem dimensão finita, podemos escolher um expoente para $(T-\lambda)$ que anula todos os vetores de $E^g(\lambda)$ simultaneamente (como vimos no exemplo que acabamos de considerar).

Lema 6.35. *Seja V um espaço de dimensão finita, $T\colon V \to V$ um endomorfismo e λ um valor próprio de T. Então existe $n > 0$ tal que*

$$E^g(\lambda) = \{v \in V \colon (T-\lambda)^n v = 0\} = N((T-\lambda)^n)$$

Dem. Seja $\{v_1, \ldots, v_p\}$ uma base de $E^g(\lambda)$ e $k_1, \ldots, k_p$ tais que $(T-\lambda)^{k_i}(v_i) = 0$. Tomando para n o máximo de todos os k_i e escrevendo $v \in E^g(\lambda)$ como uma combinação linear dos v_i's temos, para $v = \alpha_1 v_1 + \ldots + \alpha_p v_p$,

$$(T-\lambda)^n(\alpha_1 v_1 + \ldots + \alpha_p v_p) = \alpha_1 (T-\lambda)^n(v_1) + \ldots + \alpha_p (T-\lambda)^n(v_p) = 0$$

pelo que $E^g(\lambda) \subset N((T-\lambda)^n)$. A inclusão recíproca é imediata da definição de $E^g(\lambda)$. $\square$

6.36. **Decomposição primária de um endomorfismo.** Podemos agora dar um grande passo na direção da decomposição de um endomorfismo em "endomorfismos atómicos".

Teorema 6.37. *(Decomposição primária) Seja V um espaço vetorial complexo de dimensão finita e $T\colon V \to V$ um endomorfismo. Sejam $\lambda_1, \ldots, \lambda_k$ os valores próprios de T. Então*

$$V = E^g(\lambda_1) \oplus E^g(\lambda_2) \oplus \ldots \oplus E^g(\lambda_k)$$

A afirmação do Teorema anterior pode ser dividida em duas partes: a primeira é que a soma dos espaços próprios é direta, ou seja que cada um dos espaços próprios generalizados não interseta os vetores que se podem escrever como somas de vetores dos restantes espaços.

$$E^g(\lambda_i) \cap \left(\sum_{j \neq i} E^g(\lambda_j) \right) = 0$$

A segunda parte é a afirmação que os espaços próprios generalizados geram todo o V. Para provar a primeira afirmação vamos usar o seguinte resultado.

Lema 6.38. *Seja V um espaço vetorial de dimensão finita, $T\colon V \to V$ um endomorfismo, μ um valor próprio de T e λ um escalar diferente de μ. Então a restrição de $(T-\lambda)$ a $E^g(\mu)$ é um isomorfismo.*

Dem. Uma vez que $E^g(\mu)$ é invariante para T, também é invariante para $T-\lambda$. Como V e portanto $E^g(\mu)$ têm dimensão finita só temos que verificar que $N(T-\lambda) \cap E^g(\mu) = \{0\}$.

Suponhamos que $v \in N(T-\lambda)$. Então $Tv = \lambda v$, e portanto

$$(T-\mu)(v) = Tv - \mu v = (\lambda - \mu)v$$

Se v pertence também a $E^g(\mu)$ temos $(T-\mu)^k(v) = (\lambda-\mu)^k v = 0$ para algum $k > 0$. Uma vez que $\mu \neq \lambda$, conclui-se que $v = 0$ conforme queríamos demonstrar. $\square$

Para demonstrar que os vetores própios generalizados geram V vamos usar a noção de espaço vetorial quociente. Recorde-se do Exercício 3.12 que dado um subespaço $W \subset V$, o conjunto V/W é o conjunto dos planos paralelos a W em V, que se escrevem na forma $v + W$ com $v \in V$. Com as operações de soma e produto por escalar definidas pelas fórmulas

$$(v_1 + W) + (v_2 + W) = (v_1 + v_2) + W, \qquad \alpha(v + W) = (\alpha v) + W$$

o conjunto V/W adquire a estrutura de um espaço vetorial - o espaço vetorial quociente de V por W. Este espaço "descarta"o subespaço W no sentido em que dois vetores v_1 e v_2 de V que difiram por um elemento de W correspondem ao mesmo elemento $v_1 + W = v_2 + W$ no espaço quociente.

Se $T : V \to V$ é uma transformação linear e $W \subset V$ é um subespaço invariante, então T determina um endomorfismo $\overline{T} : V/W \to V/W$, definido pela expressão

$$\overline{T}(v + W) = T(v) + W$$

De facto, se $v_1 + W = v_2 + W \Leftrightarrow v_1 - v_2 \in W$, então $T(v_1) - T(v_2) = T(v_1 - v_2) \in W$ pelo que $T(v_1) + W = T(v_2) + W$. Isto mostra que a fórmula acima para $\overline{T}$ define de facto uma função de V/W para V/W e é então imediato verificar que esta função preserva a soma e o produto por escalar em V/W.

O nosso plano será ver que, sendo W o subespaço gerado por todos os subespaços próprios generalizados de um dado endomorfismo $T : V \to V$, temos $V/W = \{0\}$ e portanto $W = V$.

Dem. do Teorema 6.37. Começamos por ver que a soma dos espaços próprios generalizados é direta. Suponhamos que $v \in E^g(\lambda_1)$ e que $v = w_2 + \ldots + w_p$ com $w_j \in E^g(\lambda_j)$ vetores próprios generalizados de valores próprios distintos de λ_1. Sejam $n_2, \ldots, n_p$ tais que $(T - \lambda_j)^{n_j} w_j = 0$, e

$$q(T) = (T - \lambda_2)^{n_2} \cdots (T - \lambda_p)^{n_p}$$

Uma vez que a ordem dos fatores na fatorização de $q(T)$ é arbitrária, temos que $q(T)w_j = 0$ para todo o j. Portanto

$$q(T)v = q(T)w_2 + \ldots + q(T)w_p = 0$$

Pelo Lema 6.38 a restrição de cada um dos fatores de $q(T)$ a $E^g(\lambda_1)$ é um isomorfismo. Segue-se que a restrição de $q(T)$ a $E^g(\lambda_1)$ é também um isomorfismo e portanto

$$q(T)v = 0 \Rightarrow v = 0.$$

Conclui-se assim que $E^g(\lambda_1) \cap \sum_{j=2}^{n} E^g(\lambda_j) = \{0\}$ e portanto, indutivamente, que

$$V = \oplus_{i=1}^{k} E^g(\lambda_i)$$

Vejamos agora que os espaços próprios generalizados geram V. Seja $W = \oplus_{i=1}^{k} E^g(\lambda_i)$ o subespaço de V gerado por todos os espaços próprios generalizados. Claramente W é um subespaço invariante. Consideremos a transformação linear $\overline{T} : V/W \to V/W$ determinada por T no espaço quociente.

Se $V/W \neq 0$, pelo Teorema 6.18, o endomorfismo $\overline{T}$ tem um valor próprio α. Seja $v+W$ um vetor próprio de $\overline{T}$ associado a α. Então

$$\overline{T}(v+W) = \alpha(v+W) \Leftrightarrow T(v) = \alpha v + w \text{ com } w \in W$$

Suponhamos primeiro que α não é um dos valores próprios de T. Então, pelo lema 6.38, a restrição de $(T-\alpha)$ a W é um isomorfismo. Seja $w' \in W$ tal que $(T-\alpha)w' = w$. Então

$$(T-\alpha)(v-w') = (T-\alpha)v - (T-\alpha)w' = w - w = 0$$

Uma vez que $v - w' \neq 0$ (senão $v = w' \in W$ e $v + W = 0 + W$ contrariamente à nossa hipótese que $v+W$ é um vetor próprio de $\overline{T}$) conclui-se que α é um valor próprio de T, o que contradiz a hipótese inicial sobre α.

Suponhamos então que α é um dos valores próprios de T. Sem perda de generalidade podemos assumir que $\alpha = \lambda_1$. Podemos escrever

$$w = w_1 + w_2 \quad \text{ com } w_1 \in E^g(\lambda_1), \quad w_2 \in \bigoplus_{j=2}^{k} E^g(\lambda_j)$$

Pelo Lema 6.38 podemos escolher $w_2' \in \bigoplus_{j=2}^{k} E^g(\lambda_j)$ tal que $(T-\alpha)w_2' = w_2$. Então

$$(T-\alpha)(v-w_2') = w - w_2 = w_1$$

Sendo m tal que $(T-\lambda_1)^m(w_1) = (T-\alpha)^m(w_1) = 0$ temos então

$$(T-\alpha)^{m+1}(v-w_2') = (T-\alpha)^m(T-\alpha)(v-w_2') = (T-\alpha)^m(w_1) = 0$$

Portanto $v - w_2' \in E^g(\alpha) = E^g(\lambda_1) \subset W$. Uma vez que $w_2' \in W$ conclui-se que $v \in W$, o que novamente contradiz a hipótese de $v+W$ ser um vetor próprio. Esta contradição mostra que $V/W = \{0\}$ e conclui a demonstração. ☐

6.39. **Endomorfismos nilpotentes.** Tendo em conta o Teorema 6.37 resta-nos entender a restrição de uma transformação linear T a um espaço próprio generalizado $E^g(\lambda)$. Escrevendo $N = T - \lambda$, basta-nos entender a restrição de N a $E^g(\lambda)$ pois

$$T = \lambda + (T-\lambda) = \lambda + N$$

Definição 6.40. *Seja W um espaço vetorial. Um endomorfismo $N\colon W \to W$ diz-se* nilpotente *se existe $k > 0$ tal que $N^k = 0$. O menor k tal que $N^k = 0$ diz-se o* índice de nilpotência *de N.*

Note-se que, se $N\colon W \to W$ é nilpotente, então $W = E^g(0)$. Por outro lado, se W é finitamente gerado, o Lema 6.35 implica que quando $W = E^g(0)$ o endomorfismo N é nilpotente. Do Teorema da Decomposição Primária conclui-se então que, para W de dimensão finita, $N\colon W \to W$ é nilpotente se e só se 0 é o único valor próprio de N.

Os endomorfismos nilpotentes podem ser encarados para certos efeitos como sendo "negligíveis". Afinal, são "raízes índice k"de 0! Num espaço próprio generalizado (que não é um espaço próprio) T expressa-se não como um múltiplo da identidade mas como um múltiplo da identidade mais um endomorfismo nilpotente.

Exemplo 6.41. *O índice de nilpotência de N é 1 se e só se N é a transformação linear identicamente nula. O endomorfismo N representado na base canónica de $\mathbb{R}^3$ pela matriz*

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

tem índice de nilpotência 3, uma vez que

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}^2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}^3 = 0$$

Mais geralmente, um endomorfismo $\mathbb{R}^k$ representado por uma matriz $k \times k$

$$N = \begin{bmatrix} 0 & 1 & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \ddots & 1 \\ 0 & 0 & \cdots & 0 \end{bmatrix} \tag{42}$$

tem índice de nilpotência k. De facto, o efeito de multiplicar N por uma matriz com k linhas é

$$\begin{bmatrix} 0 & 1 & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \ddots & 1 \\ 0 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} \cdots & L_1 & \cdots \\ \cdots & L_2 & \cdots \\ & \vdots & \\ \cdots & L_k & \cdots \end{bmatrix} = \begin{bmatrix} \cdots & L_2 & \cdots \\ \cdots & L_3 & \cdots \\ & \vdots & \\ \cdots & L_k & \cdots \\ 0 & 0 & 0 \end{bmatrix}$$

logo, à medida que o expoente i em N^i aumenta, a linha de 1s vai subindo até que finalmente desaparece quando i chega a k.

O nosso objetivo nesta secção é mostrar o seguinte resultado, que, em conjunto com o Teorema da Decomposição Primária, levará imediatamente a uma forma normal para todos os endomorfismos de espaços vetoriais complexos de dimensão finita - a forma canónica de Jordan.

Teorema 6.42. *Seja W um espaço vetorial de dimensão finita e $N: W \to W$ um endomorfismo nilpotente. Então existe uma base B para W tal que a matriz $A_{N,B,B}$ que representa N com respeito à base B é diagonal por blocos, sendo cada bloco da forma (42).*

Note-se que se $v_1, \ldots, v_k$ são os vetores da base correspondentes às colunas de um dos blocos temos

$$Nv_1 = 0, \quad Nv_2 = v_1, \quad \ldots \quad Nv_k = v_{k-1} \tag{43}$$

Definição 6.43. *Seja $N: W \to W$ um endomorfismo nilpotente. Uma sucessão de vetores não nulos $(v_1, \ldots, v_k)$ satisfazendo (43) chama-se uma* cadeia de Jordan *para N.*

Proposição 6.44. *Seja $N: W \to W$ um endomorfismo nilpotente. Uma cadeia de Jordan $(v_1, \ldots, v_k)$ para N é um conjunto linearmente independente.*

Dem. Note-se que $N^{k-i}v_k = v_i$, e que $N^j v_i = 0$ se $j \geq i$. Logo se $\alpha_1 v_1 + \ldots + \alpha_k v_k = 0$, temos

$$0 = N^{k-1}(\alpha_1 v_1 + \ldots + \alpha_k v_k) = 0 + \ldots 0 + \alpha_k v_1$$

portanto $\alpha_k = 0$. Aplicando sucessivamente N^{k-i} com $i = 2, \ldots k$ obtemos $\alpha_j = 0$ para todo o j, o que conclui a demonstração. $\square$

Observação 6.45. *Seja $(v_1, \ldots, v_k)$ uma cadeia de Jordan para N.*

(i) $W = L(\{v_1, \ldots, v_k\})$ é o subespaço cíclico gerado pelo vetor v_k.
(ii) Na base $(v_1, \ldots, v_k)$ a expressão matricial da restrição de N a W é (42).

Corolário 6.46. *Se V é um espaço vetorial de dimensão n, e $N\colon V \to V$ é um endomorfismo nilpotente, o índice de nilpotência de N é no máximo n. Em particular, no Lema 6.35 podemos tomar para expoente n a dimensão de V.*

Dem. Seja k o índice de nilpotência de N e $v \in V$ tal que $N^{k-1}v \neq 0$. Então $(v, Nv, \ldots, N^{k-1}v)$ forma uma cadeia de Jordan e portanto um conjunto linearmente independente. Conclui-se que $\dim V \geq k$. $\square$

A chave para entender que tamanho têm os blocos no Teorema 6.42 é a seguinte. Um endomorfismo nilpotente $N\colon W \to W$ determina uma *filtração* do espaço W, isto é uma sucessão encadeada de subespaços

$$\{0\} = W(0) \subset W(1) \subset W(2) \subset \cdots \subset W(k) = W$$

onde $W(i)$ é o núcleo de N^i, e k é o índice de nilpotência de N. De facto,

$$v \in W(i) \Leftrightarrow N^i v = 0 \Rightarrow N^{i+1}v = 0 \Leftrightarrow v \in W(i+1)$$

logo $W(i) \subset W(i+1)$. Definindo $W(-1) = \emptyset$, para cada $w \in W$, existe exatamente um $i \in \{0, \ldots, k\}$ tal que $w \in W(i)$ mas $w \notin W(i-1)$. Este número chama-se a *filtração* de w (diz qual é o primeiro $W(i)$ a que w pertence). Um elemento de $W(i)$ tem portanto filtração $\leq i$.

Exemplo 6.47. *A filtração de um vetor $v \in \mathbb{R}^5$ determinada pelo endomorfismo nilpotente representado pela matriz*

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

é

- 0 *se* $v = (0, 0, 0, 0, 0)$,
- 1 *se* $v = (x, 0, 0, w, 0)$ *com* $x \neq 0$ *ou* $w \neq 0$
- 2 *se* $v = (x, y, 0, w, v)$ *com* $y \neq 0$ *ou* $v \neq 0$
- 3 *caso contrário, isto é se* $v = (x, y, z, w, v)$ *com* $z \neq 0$.

A seguinte observação é importante. Confirme-a no exemplo anterior.

Lema 6.48. *Seja $N\colon W \to W$ um endomorfismo nilpotente com índice de nilpotência k. Um elemento $w \in W$ tem filtração i, com $1 \leq i \leq k$ se e só se Nw tem filtração $i-1$.*

Dem. Para $1 \leq i \leq k$ temos $w \in W(i) \Leftrightarrow N^i w = 0 \Leftrightarrow N^{i-1}(Nw) = 0 \Leftrightarrow Nw \in W(i-1)$. Se $i = 1$ isto é a afirmação que pretendemos demonstrar. Se $i > 1$, um elemento $w \in W$ tem filtração i se e só se $w \in W(i) \wedge w \notin W(i-1) \Leftrightarrow Nw \in W(i-1) \wedge Nw \notin W(i-2) \Leftrightarrow Nw$ tem filtração $i-1$. $\square$

Lema 6.49. *Seja $W \neq \{0\}$ e $N\colon W \to W$ um endomorfismo nilpotente com índice de nilpotência k e $d_i = \dim W(i) - \dim W(i-1)$, $n = \dim W$. Então*

$$d_1 \geq d_2 \geq \ldots \geq d_k > 0$$

Dem. Se o índice de nilpotência é 1, não temos nada a mostrar. Suponhamos que $k \geq 2$ e sejam i tal que $2 \leq i \leq k$, B uma base de $W(i-1)$ e $\{v_1, \ldots, v_{d_i}\}$ um subconjunto linearmente independente de $W(i)$ de tal forma que $B' = B \cup \{v_1, \ldots, v_{d_i}\}$ é uma base de $W(i)$ (este subconjunto existe porque podemos completar uma base B de $W(i-1)$ numa base B' de $W(i)$). Seja

$$v = \alpha_1 v_1 + \ldots + \alpha_{d_i} v_{d_i}$$

uma combinação linear com os coeficientes não todos nulos. Trata-se de um elemento não nulo de $W(i)$ que não pode pertencer a $W(i-1) = L(B)$ devido à independência linear de B'. Então v tem filtração i e portanto pelo Lema 6.48

$$Nv = \alpha_1 N v_1 + \ldots + \alpha_{d_i} N v_{d_i}$$

tem filtração $i-1$. Em particular Nv não pode ser zero. Isto mostra que

$$\{Nv_1, \ldots, Nv_{d_i}\} \text{ é lin indep em } W(i-1) \text{ e } L(\{Nv_1, \ldots, Nv_{d_i}\}) \cap W(i-2) = \{0\} \tag{44}$$

e portanto que

$$\dim W(i-2) + d_i \leq \dim W(i-1) \Leftrightarrow d_i \leq d_{i-1}$$

$\square$

Note-se que $d_1 + d_2 + \ldots + d_k = \dim W$. Os números $d_1, \ldots, d_k$ formam aquilo a que se chama uma *partição* do inteiro $n = \dim W$. Esta pode representar-se visualmente como no exemplo seguinte, no qual $n = 17, k = 6, d_1 = 5, d_2 = d_3 = 4, d_4 = 2$ e $d_5 = d_6 = 1$.

$$\begin{array}{|c|c|c|c|c|c|}
\hline
1 & 1 & 1 & 1 & 1 & 1 \\
\hline
2 & 2 & 2 & 2 \\
\cline{1-4}
3 & 3 & 3 \\
\cline{1-3}
4 & 4 & 4 \\
\cline{1-3}
5 \\
\cline{1-1}
\end{array} \tag{45}$$

O diagrama acima explica como a dimensão dos espaços $W(i)$ vai aumentando à medida que i aumenta: a dimensão de $W(i)$ é igual a $d_1 + \ldots + d_i$ que é o número total de quadrados nas primeiras i colunas do diagrama.

Dem. do Teorema 6.42. A demonstração vai consistir na construção de uma base B para W que obedeça às condições do enunciado. Para cada i entre 1 e $k-1$ seja $U(i) = W(i-1) + N(W(i+1)) \subset W(i)$ e $U(k) = W(k-1)$. Temos então

$$W(i-1) \subset U(i) \subset W(i)$$

$U(i)$ consiste nos vetores de filtração $\leq i-1$ juntamente com os que têm filtração i e estão na imagem por N de elementos de filtração $i+1$. Uma base para $U(i)$ obtém-se juntando a uma base de $W(i-i)$ a imagem por N de vetores $\{v_1, \ldots, v_{d_{i+1}}\}$ que completem uma base de $W(i)$ a uma base de $W(i+1)$ (cf. (44)). Em particular, $\dim U(i) = \dim W(i-1) + d_{i+1}$

Para cada $i = 1, \ldots, k-1$ seja $b_i = d_i - d_{i+1}$ (≥ 0 pelo Lema 6.49) e seja $b_k = d_k$. Sejam $v(i)_1, \ldots v(i)_{b_i}$ vetores tais que

$$U(i) + L(\{v(i)_1, \ldots, v(i)_{b_i}) = W(i)$$

(estes vetores obtêm-se completando uma base de $U(i)$ a uma de $W(i)$). Note-se que cada vetor $v(i)_j$ tem filtração i e que o conjunto $\{v(i)_j\}$ é linearmente independente. Intuitivamente, os vetores $v(i)_j$ contabilizam o aumento de dimensão na passagem de $W(i-1)$ para $W(i)$ que é justificada pela existência de "novos vetores de filtração exatamente i", isto é de vetores de filtração i que não provêm de vetores com filtração superior.

Vamos ver que a base B pretendida é a união das cadeias de Jordan com início em $v(i)_j$ para cada i, j (em particular b_i é o número de blocos de tamanho i).

Definimos $B = \{N^m v(i)_j \colon 1 \leq i \leq k, 0 \leq m \leq i-1, 1 \leq j \leq b_i\}$. Uma vez que o número de elementos de B é

$$b_1 + 2b_2 + \ldots + kb_k = d_1 - d_2 + 2(d_2 - d_3) + \ldots + (k-1)(d_{k_1} - d_k) + kd_k = d_1 + \ldots + d_k$$

que é a dimensão de W, basta-nos demonstrar que B é linearmente independente. Note-se que, pelo Lema 6.48, a filtração do vetor $N^m v(i)_j$ é $i-m$. Suponhamos que

$$\sum_{i=1}^{k} \sum_{j=1}^{b_i} \sum_{m=1}^{i-1} \alpha_{i,j,m} N^m v(i)_j = 0 \tag{46}$$

Aplicando N^{k-1} à equação (46) todos os termos excepto os de filtração k são aniquilados e ficamos com

$$N^{k-1}\left(\sum_{j=1}^{b_k} \alpha_{k,j,0} v(k)_j\right) = 0$$

Como (pelo Lema 6.48) o conjunto $\{N^{k-1}v(k)_j\}$ é linearmente independente conclui-se que todos os coeficientes de vetores de B com filtração k se anulam. Aplicando N^{k-2} a (46) obtemos agora uma igualdade que envolve apenas os termos em (46) com filtração $k-1$:

$$N^{k-2}\left(\sum_{j=1}^{b_k} \alpha_{k,j,1} N v(k)_j + \sum_{j=1}^{b_{k-1}} \alpha_{k-1,j,0} v(k-1)_j\right) = 0$$

Uma vez que $U(k-1) = W(k-2) + L(\{Nv(k)_1, \ldots, Nv(k)_{b_k}\})$ e $U(k-1) + L(\{v(k-1)_1, \ldots, v(k-1)_{b_{k-1}}\}) = W(k-1)$ vemos que o conjunto $\{Nv(k)_1, \ldots, Nv(k)_{b_k}, v(k-$

$1)_1, \ldots, v(k-1)_{b_{k-1}}\}$ é linearmente independente e portanto todos os coeficientes de vetores de B com filtração $k-1$ se anulam.

Aplicando sucessivamente N^{k-i} com $i = 3, \ldots k$ vemos como acima que todos os coeficientes se anulam, o que conclui a demonstração. □

Observação 6.50. *A demonstração anterior dá-nos um algoritmo (não muito prático) para determinar a base com respeito à qual N se expressa como uma matriz diagonal por blocos da forma* (42)*:*

(i) Determinar o índice de nilpotência k de N.
(ii) Determinar os subespaços $W(i) = N(N^i)$ e $U(i) = W(i-1) + N(W(i+1))$ e bases para estes compatíveis com as inclusões

$$0 \subset U(1) \subset W(1) \subset U(2) \subset W(2) \subset W(k) = W$$

no sentido em que a base de cada subespaço contém as bases dos subespaços mais pequenos.
(iii) Tomar para base B a união das cadeias de Jordan (de comprimento i) dos elementos da base de $W(i)$ que não pertencem a $U(i)$. Em particular, o número de blocos de Jordan de tamanho i é $\dim W(i) - \dim U(i)$.

Note-se que em termos do diagrama (45) o número de blocos de tamanho i é a diferença entre os comprimentos das colunas i e $i+1$. Ou seja, os blocos de Jordan correspondem precisamente às *linhas* do diagrama, sendo o seu tamanho o comprimento da linha correspondente.

6.51. **A forma canónica de Jordan.** Podemos agora terminar a nossa classificação dos endomorfismos de um espaço vetorial complexo de dimensão finita.

Definição 6.52. *Uma matriz $n \times n$ diz-se um* bloco de Jordan*, se é da forma*

$$J = \begin{bmatrix} \lambda & 1 & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \ddots & 1 \\ 0 & 0 & \cdots & \lambda \end{bmatrix} \tag{47}$$

Sendo V um espaço vetorial, $T\colon V \to V$ um endomorfismo e λ um escalar, uma cadeia de Jordan *é uma sucessão de vetores não nulos $(v_1, \ldots, v_k)$ em V tal que*

$$Tv_1 = \lambda v_1, \quad Tv_2 = \lambda v_2 + v_1, \quad \ldots \quad Tv_k = \lambda v_k + v_{k-1} \tag{48}$$

As condições (48) podem ser re-escritas como

$$(T-\lambda)v_1 = 0, \quad (T-\lambda)v_2 = v_1, \quad \ldots \quad (T-\lambda)v_k = v_{k-1} \tag{49}$$

Em particular $(T-\lambda)^i v_i = 0$ pelo que os vetores de uma cadeia de Jordan são vetores próprios generalizados de λ, e formam uma cadeia de Jordan no sentido da Definição 43 para o operador nilpotente dado pela restrição de $(T-\lambda)$ a $E^g(\lambda)$.

Teorema 6.53 (Forma canónica de Jordan)**.** *Seja V um espaço vetorial complexo de dimensão finita e $T\colon V \to V$ um endomorfismo. Então existe uma base B para V tal que a matriz $A_{T,B,B}$ que representa T com respeito à base B é diagonal por blocos*

$$A_{T,B,B} = \begin{bmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & & J_m \end{bmatrix}$$

com cada J_i um bloco de Jordan.

Dem. Pelo Teorema 6.37 é suficiente ver que cada espaço próprio generalizado $E^g(\lambda)$ tem uma base com respeito à qual a restrição de T a $E^g(\lambda)$ é representada por uma matriz diagonal por blocos com cada bloco um bloco de Jordan. Pelo Teorema 6.42, a restrição de $(T - \lambda)$ a $E^g(\lambda)$ admite uma tal representação, tendo os blocos 0 na diagonal. Mas isso é exatamente equivalente à afirmação que a restrição de T a $E^g(\lambda)$ é representada por uma matriz diagonal por bloco, sendo os blocos todos eles blocos de Jordan com λ na diagonal. □

Observação 6.54. *Se aplicarmos o Teorema 6.53 à transformação linear $T\colon \mathbb{C}^n \to \mathbb{C}^n$ representada na base canónica pela matriz $A \in M_{n\times n}(\mathbb{C})$ ele diz-nos que existe uma matriz invertível S (a matriz de mudança de base da base B para a base canónica) e uma matriz diagonal por blocos J cujos blocos são de Jordan, tais que*

$$A = SJS^{-1}$$

O Teorema é muitas vezes enunciado nesta forma.

Observação 6.55. *É fácil ver que a matriz diagonal por blocos no enunciado do Teorema 6.53 é única a menos de troca da ordem dos blocos. Isso está longe de ser verdade para a base B, como aliás já era verdade no caso em que a transformação é diagonalizável.*

O Teorema 6.53 classifica *todos os endomorfismos de espaços vetoriais complexos de dimensão finita a menos de isomorfismo. Se $T_1, T_2\colon V \to V$ têm a mesma forma canónica de Jordan, então sendo S o isomorfismo que envia os vetores da base B_1 dada pelo Teorema para T_1, por ordem, para os vetores da base B_2 correspondente a T_2 teremos $T_2 = S \circ T_1 \circ S^{-1}$, pelo que o isomorfismo S "traduz"T_1 em T_2.*

Reciprocamente, se existe um isomorfismo S tal que $T_2 = ST_1S^{-1}$, os endomorfismos T_1 e T_2 têm a mesma forma canónica de Jordan (exercício).

Observação 6.56. *Seja V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{K}$ e $T\colon V \to V$ um endomorfismo. Se existir uma base B para V satisfazendo as condições do Teorema 6.53 diz-se que T admite uma forma canónica de Jordan. Quando $\mathbb{K}$ não é algebricamente fechado tal pode não acontecer mas, quando acontece, é um facto que pode ser aproveitado.*

Para determinar a forma canónica de Jordan de um endomorfismo T (o que significa calcular a matriz diagonal por blocos e a base) podemos começar por calcular os valores próprios de T. Estes dizem-nos que escalares aparecem na diagonal da matriz $A_{T,B,B}$.

Embora seja possível determinar os espaços próprios generalizados de cada valor próprio λ e depois aplicar o algoritmo da Observação 6.50 à restrição de $T-\lambda$ a cada $E^g(\lambda)$, é mais prático proceder da seguinte forma.

Uma vez que os elementos da base correspondentes às primeiras colunas de cada bloco J_i são vetores próprios, a multiplicidade geométrica de cada valor próprio λ dá-nos o número de blocos com λ na diagonal.

Para determinar o resto da base B e o comprimento dos blocos tentamos resolver recursivamente as equações (49) começando com um vetor próprio v_1. Isto pode requerer algum cuidado na escolha do vetor v_1 como iremos ver nos exemplos que se seguem.

Finalmente, o polinómio característico de T dá informação sobre a forma canónica de Jordan que pode facilitar a determinação da base:

Proposição 6.57. *Seja V um espaço vetorial de dimensão finita, $T\colon V \to V$ um endomorfismo, e $p(\lambda)$ o seu polinómio característico.*

(a) A multiplicidade do valor próprio λ_i enquanto raíz de p (isto é, o maior expoente m tal que $(\lambda-\lambda_i)^m$ divide $p(\lambda)$) é igual à multiplicidade algébrica de λ_i, $\dim E^g(\lambda_i)$.

(b) Se $\mathbb{K}=\mathbb{C}$ então

$$p(\lambda) = (\lambda_1-\lambda)^{n_1}(\lambda_2-\lambda)^{n_2}\cdots(\lambda_k-\lambda)^{n_k}$$

onde $\lambda_1,\dots,\lambda_k$ são os valores próprios (distintos) de T e n_i as suas multiplicidades algébricas.

Dem. (a) Quando o corpo de base é $\mathbb{C}$ esta afirmação é uma consequência imediata da forma canónica de Jordan. De facto a multiplicidade de λ_i como raíz é igual aos número de vezes que λ_i aparece na diagonal na forma canónica de Jordan, que é precisamente o número de elementos de uma base para o espaço próprio generalizado de λ_i.

Para demonstrar a afirmação em geral (para um espaço vetorial sobre um corpo qualquer) note-se que pelo Teorema 6.42 existe uma base B_1 para $E^g(\lambda_i)$ formada por cadeias de Jordan para o valor próprio λ_i. Completando esta base com um conjunto B_2 de vetores de V tal que $B=B_1\cup B_2$ é uma base de V obtemos uma representação matricial diagonal por blocos

$$A_{T,B,B} = \begin{bmatrix} J & X \\ 0 & T' \end{bmatrix}$$

onde as colunas de J correspondem aos vetores do conjunto B_1 (J é diagonal por blocos, com blocos de Jordan correspondentes a λ_i na diagonal) e as restantes colunas correspondem aos vetores de B_2. Pela Proposição 5.20 o polinómio característico de T é igual a $(\lambda_i-\lambda)^m q(\lambda)$, onde $m=\#B_1=\dim E^g(\lambda_i)$, e $q(\lambda)$ é o polinómio característico de T'.

Deixamos como exercício mostrar, argumentando como na demonstração do Teorema 6.37, que λ_i não pode ser um valor próprio de T', pelo que $(\lambda_i-\lambda)$ não divide $q(\lambda)$. Isto conclui a demonstração.

(b) É uma consequência imediata do facto de podermos calcular o polinómio característico usando a forma canónica de Jordan de T.

□

Observação 6.58. *A Proposição 6.57(a) justifica a terminologia* multiplicidade algébrica*.*

A Proposição anterior ajuda a calcular a forma canónica de Jordan de uma matriz complexa. De facto, uma fatorização do polinómio característico em fatores do primeiro grau dá-nos a multiplicidade algébrica de cada valor próprio λ_i e portanto o número de vezes que λ_i aparece na diagonal da forma canónica de Jordan (que é a soma dos comprimentos dos blocos correspondentes a λ_i).

Exemplo 6.59. *No Exemplo 6.26 a fatorização $p(\lambda) = (\lambda-1)(\lambda+1)(\lambda-2)^2$ do polinómio característico implica que a multiplicidade algébrica de 2 é igual a 2. Assim, ao verificar que a multiplicidade geométrica de $\lambda = 2$ é apenas 1, podemos concluir, sem quaisquer outros cálculos adicionais, que a forma canónica de Jordan desta matriz é*

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

Exemplo 6.60. *Retomando o Exemplo 6.34 vemos que, uma vez que a multiplicidade geométrica de cada valor próprio é 1, teremos um bloco para cada valor próprio. À partida haveria duas possibilidades para a matriz diagonal por blocos (a menos de troca de ordem dos blocos):*

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \qquad ou \qquad \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

No entanto vimos já que o espaço próprio generalizado de $\lambda = 2$ tem dimensão 2 (e isto é também uma consequência da Proposição 6.57(a)) pelo que teremos necessariamente a segunda opção. A base $B = (v_1, v_2, v_3)$ terá que ser formada por um vetor próprio v_1 de $\lambda = 2$, um vetor próprio v_3 de $\lambda = 1$ e um vetor próprio generalizado v_2 de 2 satisfazendo

$$(A - 2I)v_2 = v_1$$

Neste exemplo as escolhas possíveis para v_1 e v_3 são únicas a menos de um escalar não nulo. Se tomarmos $v_1 = (1, 0, 1)$ e $v_3 = (1, -1, 0)$ temos que resolver a equação

$$(A - 2I)v_2 = v_1 \Leftrightarrow \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & -1 \\ 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

Uma solução é por exemplo $v_2 = (0, 1, 0)$ (mas poderíamos somar a este vetor qualquer elemento do núcleo de $(A-2I)$, isto é, qualquer vetor próprio de 2). Conclui-se assim que neste exemplo podemos tomar para a base B no Teorema 6.53

$$B = ((1, 0, 1), (0, 1, 0), (1, -1, 0))$$

Exemplo 6.61. *Seja A uma matriz com forma canónica de Jordan*

$$J=\begin{bmatrix}1&1&0\\0&1&0\\0&0&1\end{bmatrix} \tag{50}$$

O espaço próprio de 1 *tem dimensão* 2*. Seja* $\{v_1, v_1'\}$ *uma base para o espaço próprio de* 1*. Tem que se ter cuidado na escolha do vector próprio* v *de* 1 *que se põe na primeira coluna da matriz* S*. De facto, só será possível resolver a equação*

$$(A-I)v_2 = v$$

para achar a segunda coluna se v *estiver no espaço das colunas da matriz* $(A-I)$*, que tem dimensão* 1 *(tem a mesma dimensão que o espaço das colunas de* $(J-I)$*). É portanto necessário achar uma combinação linear* $v=\alpha v_1+\beta v_1'$ *que pertença ao espaço das colunas de* $A-I$*. Feito isto, a terceira coluna poderá ser qualquer vector próprio de* 1 *que juntamente com* v *forme uma base para o espaço próprio.*

Vejamos um exemplo concreto. Considere-se a matriz

$$A=\begin{bmatrix}0&-1&2\\-1&0&2\\-1&-1&3\end{bmatrix}$$

Verifica-se facilmente que 1 *é o único valor próprio. Os vectores próprios de* 1 *são as soluções de*

$$\begin{bmatrix}-1&-1&2\\-1&-1&2\\-1&-1&2\end{bmatrix}\begin{bmatrix}a\\b\\c\end{bmatrix}=0\Leftrightarrow 2c=a+b$$

O espaço próprio de 1 *é portanto o conjunto dos vectores*

$$\begin{bmatrix}a\\b\\\frac{1}{2}(a+b)\end{bmatrix}=a\begin{bmatrix}1\\0\\\frac{1}{2}\end{bmatrix}+b\begin{bmatrix}0\\1\\\frac{1}{2}\end{bmatrix}$$

e $\lambda=1$ *tem multiplicidade geométrica* 2*. Há portanto dois blocos de Jordan e a forma canónica de Jordan de* A *é necessariamente* (50) *(a menos de troca de ordem dos blocos).*

Não é no entanto possível resolver a equação

$$(A-I)v_2 = v_1$$

quando v_1 *é um dos vectores* $\left(1,0,\frac{1}{2}\right)$ *ou* $\left(0,1,\frac{1}{2}\right)$ *da base "natural"do espaço próprio de* 1*. Como observámos acima, para que a equação tenha solução é necessário que* v_1 *pertença ao espaço das colunas de* $A-I$*, que é o espaço gerado por* $(1,1,1)$*. A soma dos dois vectores da "base natural"é exactamente* $(1,1,1)$*. Resolvendo a equação*

$$\begin{bmatrix}-1&-1&2\\-1&-1&2\\-1&-1&2\end{bmatrix}\begin{bmatrix}a\\b\\c\end{bmatrix}=\begin{bmatrix}1\\1\\1\end{bmatrix}=0\Leftrightarrow 2c=a+b+1$$

obtemos as soluções

$$\begin{bmatrix} a \\ b \\ \frac{1}{2}+\frac{a}{2}+\frac{b}{2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \frac{1}{2} \end{bmatrix} + a \begin{bmatrix} 1 \\ 0 \\ \frac{1}{2} \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \\ \frac{1}{2} \end{bmatrix}$$

Podemos tomar por exemplo $v_2 = \left(0,0,\frac{1}{2}\right)$. *Para* v_3 *podemos tomar qualquer vector próprio de* 1 *que juntamente com* $(1,1,1)$ *forme uma base do espaço próprio, por exemplo,* $\left(1,0,\frac{1}{2}\right)$. *Obtemos assim a base*

$$B = ((1,1,1),(0,0,\tfrac{1}{2}),(1,0,\tfrac{1}{2}))$$

e temos

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}^{-1}$$

Exemplo 6.62. *Suponhamos que* V *é um espaço vetorial de dimensão* 5 *e* $T\colon V \to V$ *tem um único valor próprio* λ *com multiplicidade geométrica* 2. *As possíveis formas canónicas de Jordan (a menos de troca de blocos) são*

$$\begin{bmatrix} \lambda & 1 & 0 & 0 & 0 \\ 0 & \lambda & 1 & 0 & 0 \\ 0 & 0 & \lambda & 1 & 0 \\ 0 & 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & 0 & \lambda \end{bmatrix} \qquad ou \qquad \begin{bmatrix} \lambda & 1 & 0 & 0 & 0 \\ 0 & \lambda & 1 & 0 & 0 \\ 0 & 0 & \lambda & 0 & 0 \\ 0 & 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & 0 & \lambda \end{bmatrix}$$

Sendo $B = (v_1,v_2,v_3,v_4,v_5)$ *uma base na qual* T *fica em forma de Jordan, podemos distinguir os dois casos da seguinte forma. Na matriz da direita, uma vez que os blocos têm dimensão* ≤ 3 *o endomorfismo* $(T-\lambda)^3$ *é identicamente nulo. Isso não acontece na matriz da esquerda, onde* $(T-\lambda)^3 v_4 = v_1$.

Para acharmos a base B *resolvendo as equações* (49) *indutivamente teremos de ter o cuidado de começar com um vetor próprio* v_1 *de* λ *que esteja na imagem de* $(T-\lambda)^3$ *no caso da matriz da esquerda, e na imagem de* $(T-\lambda)^2$ *no caso da matriz da direita. O segundo vetor próprio (*v_5 *no caso da esquerda e* v_4 *no caso da direita) tem unicamente que ser escolhido de forma a gerar o espaço próprio de* λ *juntamente com* v_1.

6.63. **O Teorema de Cayley-Hamilton.** Vamos agora ver um resultado fundamental sobre endomorfismos de espaços vetoriais de dimensão finita que é muitas vezes útil para fazer cálculos com matrizes quadradas. Seja V um espaço vetorial de dimensão finita n e $T\colon V \to V$ um endomorfismo. Dado um vetor $v \in V$, sabemos que existem necessariamente escalares não todos nulos $c_0, c_1, \ldots, c_n$ tais que

$$c_0 v + c_1 T v + \ldots + c_n T^n v = 0$$

(ou porque há alguma repetição na sequência $v, Tv, \ldots, T^n v$, ou porque um conjunto com mais de n vetores de V é necessariamente linearmente dependente). O próximo Teorema garante, em particular, que existe uma escolha para os coeficientes c_i que funciona para todos os vetores $v \in V$ simultaneamente.

Teorema 6.64 (Teorema de Cayley-Hamilton)**.** *Seja V um espaço vetorial de dimensão finita, $T\colon V \to V$ um endomorfismo e $p(\lambda)$ o seu polinómio característico. Então $p(T) = 0$ (isto é, $p(T)\colon V \to V$ é o endomorfismo identicamente nulo).*

Dem. Seja $A \in M_{n\times n}(\mathbb{K})$ a matriz que representa T com respeito a uma qualquer base B de V. É suficiente ver que $p(A) = 0$. Pela Proposição 5.11(iv) temos[26]

$$(A - \lambda I)\operatorname{cof}(A - \lambda I)^T = p(\lambda)I \tag{51}$$

Seja

$$p(\lambda) = (-1)^n\lambda^n + c_{n-1}\lambda^{n-1} + \ldots + c_1\lambda + c_0$$

Cada entrada da matriz $\operatorname{cof}(A-\lambda I)^T$ é o determinante de uma submatriz $(n-1)\times(n-1)$ de $(A - \lambda I)^T$ e portanto é uma expressão polinomial de grau $\leq (n - 1)$ em λ. Pondo as potências de λ em evidência vemos que existem matrizes $B_i \in M_{n\times n}(\mathbb{K})$ tais que

$$\operatorname{cof}(A - \lambda I)^T = B_0 + B_1\lambda + \ldots + B_{n-1}\lambda^{n-1}$$

Substituindo em (51) obtemos

$$(A - \lambda I)(B_0 + B_1\lambda + \ldots + \lambda^{n-1}B_{n-1}) = ((-1)^n\lambda^n + \ldots + c_1\lambda + c_0)I$$

Distribuindo o produto do lado esquerdo do sinal de igual e igualando os coeficientes de λ^k obtemos as seguintes igualdades

$$\begin{cases} AB_0 = c_0 I \\ AB_1 - B_0 = c_1 I \\ AB_2 - B_1 = c_2 I \\ \vdots \\ AB_{n-1} - B_{n-2} = c_{n-1}I \\ -B_{n-1} = (-1)^n I \end{cases}$$

Multiplicando a i-ésima equação por A^{i-1} e somando obtemos

$$\begin{aligned} AB_0 + A(AB_1 - B_0) + A^2(AB_2 - B_1) + \ldots + A^{n-1}(AB_{n-1} - B_{n-2}) - A^nB_{n-1} = \\ = c_0I + c_1A + \ldots + c_{n-1}A^{n-1} + (-1)^nA^n \end{aligned}$$

A soma telescópica do lado esquerdo do sinal de igual anula-se, o que conclui a demonstração. □

O seguinte exemplo ilustra os cálculos efetuados na demonstração anterior.

[26] O significado de (51) é algo subtil. No caso em que o corpo $\mathbb{K}$ tem infinitos elementos, os polinómios são determinados pelas funções $\mathbb{K} \to \mathbb{K}$ que definem e portanto a validade de (51) segue do facto de a igualdade se verificar para todo o $\lambda \in \mathbb{K}$. Quando $\mathbb{K}$ é um corpo finito, é necessário observar que, na dedução da igualdade descrita na Proposição 5.11(iv), nunca usámos a invertibilidade dos elementos não nulos de $\mathbb{K}$ mas apenas os restantes axiomas de corpo. O resultado é portanto válido para matrizes cujas entradas são polinómios com coeficientes em $\mathbb{K}$.

Exemplo 6.65. *Seja A a matriz* $\begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$ *que tem polinómio característico* $p(\lambda) = \lambda^2 - 3\lambda - 4$. *A identidade* (51) *afirma neste caso que*

$$\begin{bmatrix} 1-\lambda & 2 \\ 3 & 2-\lambda \end{bmatrix} \begin{bmatrix} 2-\lambda & -2 \\ -3 & 1-\lambda \end{bmatrix} = \begin{bmatrix} \lambda^2 - 3\lambda - 4 & 0 \\ 0 & \lambda^2 - 3\lambda - 4 \end{bmatrix}$$

$$\left(\begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \left(\begin{bmatrix} 2 & -2 \\ -3 & 1 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) = \lambda^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \lambda \begin{bmatrix} -3 & 0 \\ 0 & -3 \end{bmatrix} + \begin{bmatrix} -4 & 0 \\ 0 & -4 \end{bmatrix}$$

Distribuindo a soma à esquerda e igualando os coeficientes de λ^k *para* $k = 0, 1, 2$ *obtemos três equações matriciais que implicam a relação* $A^2 - 3A - 4I = 0$.

Observação 6.66. *Para endomorfismos de espaços complexos de dimensão finita, o Teorema de Cayley-Hamilton é uma consequência imediata da forma canónica de Jordan.*

Exemplo 6.67. *Consideremos a matriz*

$$A = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 2 & 1 \\ -1 & 0 & 2 \end{bmatrix}$$

O seu polinómio característico é

$$p(\lambda) = \begin{vmatrix} 1-\lambda & 0 & -1 \\ 0 & 2-\lambda & 1 \\ -1 & 0 & 2-\lambda \end{vmatrix} = -\lambda^3 + 5\lambda^2 - 7\lambda + 2$$

Portanto

$$-A^3 + 5A^2 - 7A + 2I = 0 \Leftrightarrow A(-A^2 + 5A - 7I) = -2I \Rightarrow A^{-1} = \tfrac{1}{2}(A^2 - 5A + 7I)$$

É fácil calcular o valor de $t(A)$ *para qualquer polinómio* $t(x)$. *Dividindo* $t(x)$ *pelo polinómio característico obtemos*

$$t(x) = q(x)p(x) + r(x)$$

sendo o grau de r *menor ou igual a* 2. *Uma vez que* $p(A) = 0$, *temos* $t(A) = q(A)p(A) + r(A) = r(A)$.

6.68. **Endomorfismos de espaços vetoriais reais.** Vamos agora usar o nosso entendimento dos endomorfismos de espaços complexos para descrever os endomorfismos de espaços vetoriais reais de dimensão finita.

Uma vez que os números reais são um subconjunto dos números complexos, temos uma inclusão natural $M_{n\times n}(\mathbb{R}) \subset M_{n\times n}(\mathbb{C})$ e portanto uma matriz real pode ser encarada como uma matriz complexa cujas entradas têm parte imaginária nula. Uma matriz $A \in M_{n\times n}(\mathbb{R})$ determina portanto um endomorfismo $x \mapsto Ax$ de $\mathbb{C}^n$ ao qual podemos aplicar o Teorema 6.53.

O processo descrito no parágrafo anterior tem uma versão para transformações lineares: todo o espaço vetorial real pode ser extendido a um espaço vetorial complexo - a sua complexificação - de uma maneira que generaliza a inclusão $\mathbb{R}^n \subset \mathbb{C}^n$ e um endomorfismo de um espaço vetorial real determina um endomorfismo da sua complexificação.

Definição 6.69. *Seja V um espaço vetorial real. A* complexificação *de V é o espaço vetorial complexo $V^{\mathbb{C}}$ que tem por conjunto subjacente $V \times V$ e operações de soma e multiplicação por escalar definidas por*

$$(v_1, v_2) + (w_1, w_2) = (v_1 + w_1, v_2 + w_2) \qquad (a + bi) \cdot (v_1, v_2) = (av_1 - bv_2, av_2 + bv_1)$$

Definimos as partes real e imaginária de $(v_1, v_2) \in V^{\mathbb{C}}$ por $\mathrm{Re}((v_1, v_2)) = v_1$ *e* $\mathrm{Im}((v_1, v_2)) = v_2$ *e a operação de* conjugação *em $V^{\mathbb{C}}$ por*

$$\overline{(v_1, v_2)} = (v_1, -v_2)$$

Note-se que $(0+1i)\cdot(v, 0) = (0, v)$. Podemos identificar V com o subconjunto $\{(v, 0) : v \in V\} \subset V^{\mathbb{C}}$ e mediante esta identificação é habitual escrever o elemento $(v_1, v_2) \in V^{\mathbb{C}}$ como $v_1 + iv_2$. Nestes termos a multiplicação por escalar é definida pela fórmula familiar

$$(a + bi)(v_1 + iv_2) = (av_1 - bv_2) + i(av_2 + bv_1)$$

e a conjugação escreve-se

$$\overline{v_1 + iv_2} = v_1 - iv_2$$

Deixamos como exercício a verificação que com as operações de soma e multiplicação definidas acima $V^{\mathbb{C}}$ é de facto um espaço vetorial complexo.

Note-se que a conjugação é um isomorfismo entre o espaço vetorial real subjacente a $V^{\mathbb{C}}$ e ele próprio.

Definição 6.70. *Seja V um espaço vetorial real e $T : V \to V$ um endomorfismo. A* complexificação *de T é a transformação linear $T^{\mathbb{C}} : V^{\mathbb{C}} \to V^{\mathbb{C}}$ de espaços vetoriais complexos definida por*

$$T^{\mathbb{C}}(v_1 + iv_2) = T(v_1) + iT(v_2)$$

Deixamos como exercício a verificação que $T^{\mathbb{C}}$ é de facto uma transformação linear de espaços vetoriais complexos.

Exemplo 6.71.

(i) Se $V = \mathbb{R}^n$ podemos identificar $V^{\mathbb{C}}$ naturalmente com $\mathbb{C}^n$ usando o isomorfismo de espaços vetoriais complexos dado por

$$((x_1, x_2, \ldots, x_n), (y_1, y_2 \ldots, y_n)) \mapsto (x_1 + iy_1, x_2 + iy_2, \ldots, x_n + iy_n)$$

Se $T : \mathbb{R}^2 \to \mathbb{R}^2$ for a transformação linear definida pela expressão

$$T(x, y) = (2x + y, x - 3y)$$

mediante a identificação acima de $(\mathbb{R}^2)^{\mathbb{C}}$ com $\mathbb{C}^2$, a transformação linear $T^{\mathbb{C}} : \mathbb{C}^2 \to \mathbb{C}^2$ é dada pela mesma expressão. De facto, temos

$$T^{\mathbb{C}}(1 + 0i, 0 + 0i) = T(1, 0) + iT(0, 0) = (2 + 0i, 1 + 0i) = (2, 1)$$

$$T^{\mathbb{C}}(0 + 0i, 1 + 0i) = T(0, 1) + iT(0, 0) = (1 + 0i, -3 + 0i) = (1, -3)$$

logo

$$T^{\mathbb{C}}(x, y) = (2x + y, x - 3y)$$

(ii) Se V for o espaço dos polinómios reais, $V^{\mathbb{C}}$ identifica-se naturalmente com o espaço dos polinómios com coeficientes complexos mediante o isomorfismo

$$(a_0 + a_1x + \ldots + a_nx^n, b_0 + b_1x + \ldots + b_nx^n) \mapsto (a_0 + ib_0) + (a_1 + ib_1)x + \ldots + (a_n + ib_n)x^n$$

Se $B = (v_1, \ldots, v_n)$ é uma base para V é imediato verificar que B (ou mais precisamente o conjunto correspondente $((v_1, 0), \ldots, (v_n, 0)) \subset V^{\mathbb{C}}$) é também uma base para $V^{\mathbb{C}}$ (agora como espaço vetorial complexo) e que a representação matricial de $T^{\mathbb{C}}$ na base B coincide com a representação matricial de T na base B:

$$A_{T^{\mathbb{C}},B,B} = A_{T,B,B}$$

sendo a matriz real $A_{T,B,B}$ encarada como uma matriz complexa cujas entradas têm parte imaginária nula. No Exemplo (6.71)(i) acima temos

$$A_{T^{\mathbb{C}},B_{can},B_{can}} = A_{T,B_{can},B_{can}} = \begin{bmatrix} 2 & 1 \\ 1 & -3 \end{bmatrix}$$

Proposição 6.72. *Seja V um espaço vetorial real e $T\colon V \to V$ um endomorfismo.*

(i) $\lambda \in \mathbb{C}$ é um valor próprio de $T^{\mathbb{C}}$ se e só se $\overline{\lambda}$ é um valor próprio de $T^{\mathbb{C}}$

(ii) $v \in V^{\mathbb{C}}$ é um vetor próprio (generalizado) de $T^{\mathbb{C}}$ se e só se $\overline{v}$ é um vetor próprio (generalizado) de $T^{\mathbb{C}}$.

(iii) $(v_1, v_2, \ldots, v_k)$ formam uma cadeia de Jordan em $V^{\mathbb{C}}$ associada ao valor próprio λ se e só se $(\overline{v_1}, \overline{v_2}, \ldots, \overline{v_k})$ formam uma cadeia de Jordan em $V^{\mathbb{C}}$ associada ao valor próprio $\overline{\lambda}$.

Dem. Escrevendo $\lambda = a + bi$ e $v = x + iy \in V^{\mathbb{C}}$ com $x, y \in V$ temos

$$(52)\quad T(x+iy) = (a+ib)(x+iy) \Leftrightarrow Tx+iTy = ax-by+i(ay+bx) \Leftrightarrow \begin{cases} Tx = ax - by \\ Ty = ay + bx \end{cases}$$

o que claramente é equivalente a

$$T(x - iy) = (a - ib)(x - iy) \Leftrightarrow Tx - iTy = ax - by - i(ay + bx)$$

logo λ é um valor próprio se e só se $\overline{\lambda}$ é um valor próprio, e $v = x + iy$ é um vetor próprio associado a λ se e só se $\overline{v} = x - iy$ é um vetor próprio associado a $\overline{\lambda}$. Isto mostra (i) e uma parte de (ii). A demonstração de (iii) é inteiramente análoga e é deixada como exercício. Uma vez que um vetor $v \neq 0$ é um vetor próprio generalizado de λ se e só se $(v, (T-\lambda)v, \ldots, (T-\lambda)^k v)$ é uma cadeia de Jordan para algum k, (iii) implica a afirmação que falta em (ii). $\square$

O resultado anterior diz que os espaços próprios generalizados de $T^{\mathbb{C}}$ correspondentes a valores próprios complexos (isto é, com parte imaginária não nula) ocorrem em pares conjugados (relacionados pelo isomorfismo de espaços vetoriais reais dado pela conjugação $V^{\mathbb{C}} \to V^{\mathbb{C}}$). Podemos portanto tomar para base de $E^g(\overline{\lambda})$ os vetores conjugados de uma base de $E^g(\lambda)$, pelo que na forma canónica de Jordan para $T^{\mathbb{C}}$ há uma correspondência bijetiva entre os blocos com λ na diagonal e os blocos com $\overline{\lambda}$ na diagonal.

A observação anterior juntamente com o resultado seguinte permite-nos obter uma forma canónica para os endomorfismos reais.

Proposição 6.73. *Seja V um espaço vetorial real e $\{v_1, \ldots, v_n\}$ um subconjunto de vetores (distintos) de $V^{\mathbb{C}}$. Se $\{v_1, \ldots, v_n, \overline{v_1}, \ldots, \overline{v_n}\}$ é um subconjunto linearmente independente do espaço vetorial complexo $V^{\mathbb{C}}$ formado por vetores distintos então*

$$\{\mathrm{Re}(v_1), \mathrm{Im}(v_1), \ldots, \mathrm{Re}(v_n), \mathrm{Im}(v_n)\}$$

é um subconjunto linearmente independente do espaço vetorial real V formado por vetores distintos.

Dem. Usando a identificação de V com o conjunto $\{(x, 0) \colon x \in V\}$ de $V^{\mathbb{C}}$, temos

$$\mathrm{Re}(v) = \frac{v + \overline{v}}{2} \qquad \mathrm{Im}(v) = \frac{v - \overline{v}}{2i}$$

Se $\{v_1, \ldots, v_n, \overline{v_1}, \ldots, \overline{v_n}\}$ é formado por vetores distintos o mesmo é verdade a propósito do conjunto $\{\mathrm{Re}(v_1), \mathrm{Im}(v_1), \ldots, \mathrm{Re}(v_n), \mathrm{Im}(v_n)\}$. Dados escalares reais $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n$, a expressão

$$\alpha_1 \mathrm{Re}(v_1) + \ldots + \alpha_n \mathrm{Re}(v_n) + \beta_1 \mathrm{Im}(v_1) + \ldots + \beta_n \mathrm{Im}(v_n) = 0$$

é equivalente a

$$\begin{aligned} & \alpha_1 \frac{v_1 + \overline{v_1}}{2} + \ldots + \alpha_n \frac{v_n + \overline{v_n}}{2} + \beta_1 \frac{v_1 - \overline{v_1}}{2i} + \ldots + \beta_n \frac{v_n - \overline{v_n}}{2i} = 0 \\ \Leftrightarrow \quad & \frac{\alpha_1 - i\beta_1}{2} v_1 + \ldots + \frac{\alpha_n - i\beta_n}{2} v_n + \frac{\alpha_1 + i\beta_1}{2} \overline{v_1} + \ldots + \frac{\alpha_n + i\beta_n}{2} \overline{v_n} = 0 \end{aligned}$$

Se $\{v_1, \ldots, v_n, \overline{v_1}, \ldots, \overline{v_n}\}$ é linearmente dependente em $V^{\mathbb{C}}$, a equação anterior implica que os escalares $\alpha_i + i\beta_i$ são todos nulos, ou seja, todos os α_i e β_i são nulos. Conclui-se que o conjunto $\{\mathrm{Re}(v_1), \mathrm{Im}(v_1), \ldots, \mathrm{Re}(v_n), \mathrm{Im}(v_n)\}$ é linearmente dependente em V. □

Definição 6.74. *Um* bloco de Jordan real *é uma matriz quadrada da forma* (47) *com λ real ou da forma*

$$(53) \qquad \begin{bmatrix} a & -b & 1 & 0 & \cdots & & 0 & 0 \\ b & a & 0 & 1 & 0 & \cdots & & \vdots \\ 0 & 0 & a & -b & 1 & 0 & \ddots & \\ 0 & 0 & b & a & 0 & 1 & \ddots & \\ \vdots & & \ddots & & \ddots & & 1 & 0 \\ & & & & & & 0 & 1 \\ 0 & 0 & \cdots & & 0 & 0 & a & -b \\ 0 & 0 & \cdots & & 0 & 0 & b & a \end{bmatrix}$$

com $a, b \in \mathbb{R}$ e $b \neq 0$.

Teorema 6.75. *(Forma canónica de Jordan real) Seja V um espaço vetorial real de dimensão finita e $T\colon V \to V$ um endomorfismo. Então existe uma base B para V tal que a matriz $A_{T,B,B}$ que representa T com respeito à base B é diagonal por blocos*

$$A_{T,B,B} = \begin{bmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & & J_m \end{bmatrix}$$

com cada J_i um bloco de Jordan real.

Dem. Seja $T^{\mathbb{C}}\colon V^{\mathbb{C}} \to V^{\mathbb{C}}$ a complexificação de T.

Se $\lambda \in \mathbb{R}$ então $(T^{\mathbb{C}} - \lambda)^k v = 0 \Leftrightarrow (T-\lambda)^k \operatorname{Re}(v) = (T-\lambda)^k \operatorname{Im}(v) = 0$ pelo que o espaço próprio generalizado de λ para $T^{\mathbb{C}}$ é a complexificação do espaço próprio generalizado de λ para T. Em particular, podemos assumir que as cadeias de Jordan para $T^{\mathbb{C}}$ que correspondem a blocos com $\lambda \in \mathbb{R}$ na diagonal são formadas por vetores de $V \subset V^{\mathbb{C}}$.

Pela Prop 6.72, os blocos de Jordan de $T^{\mathbb{C}}$ correspondentes a valores próprios complexos $\lambda = a + bi$ com $b \neq 0$ ocorrem em pares conjugados e para cada par $J, \overline{J}$ podemos escolher cadeias de Jordan conjugadas $(v_1, \ldots, v_k)$ e $(\overline{v_1}, \ldots, \overline{v_k})$ com

$$T^{\mathbb{C}} v_1 = \lambda v_1, \ldots, T^{\mathbb{C}} v_k = \lambda v_k + v_{k-1}$$

Seja B' a base de $V^{\mathbb{C}}$ formada por todas as cadeias de Jordan para $V^{\mathbb{C}}$ escolhidas da forma acima (isto é, escolhendo vetores de V para as cadeias correspondentes a valores próprios reais e escolhendo cadeias conjugadas para blocos de valores próprios conjugados correspondentes). Tomamos para B o subconjunto de V formado por

(i) cadeias de Jordan em B' correspondentes a valores próprios reais,
(ii) $\{\operatorname{Re}(v_1), \operatorname{Im}(v_1), \ldots, \operatorname{Re}(v_k), \operatorname{Im}(v_k)\}$ para cada par conjugado de cadeias $(v_1, \ldots, v_k)$ e $(\overline{v_1}, \ldots, \overline{v_k})$

Pelo Lema 6.73, os conjuntos $\{\operatorname{Re}(v_1), \operatorname{Im}(v_1), \ldots, \operatorname{Re}(v_k), \operatorname{Im}(v_k)\}$ são linearmente independentes. Contas inteiramente análogas a (52) mostram que a restrição de T a

$$L(\{\operatorname{Re}(v_1), \operatorname{Im}(v_1), \ldots, \operatorname{Re}(v_k), \operatorname{Im}(v_k)\})$$

na base $(\operatorname{Re}(v_1), \operatorname{Im}(v_1), \ldots, \operatorname{Re}(v_k), \operatorname{Im}(v_k))$ é dado pelo bloco de Jordan real $(2k) \times (2k)$ (53) com a e b tais que $v_1 \in V^{\mathbb{C}}$ é um vetor próprio de $a - bi$.

B é um conjunto linearmente independente porque a soma dos subespaços gerados pelos vários conjuntos de vetores de tipo (i) e (ii) é direta (isso é claramente verdade para as suas complexificações pelo Teorema da Decomposição Primária). Uma vez que o número de elementos de B é igual à dimensão de V conclui-se que B é uma base para V com as propriedades requeridas. $\square$

Observação 6.76. *Ao contrário da forma canónica de Jordan complexa, a forma real não é única a menos da ordem dos blocos. Nos blocos correspondentes a valores próprios complexos, o sinal da parte imaginária b pode ser escolhido arbitrariamente (mas a escolha tem de ser constante em cada bloco). A troca do sinal de b corresponde à simetria dada* *pela conjugação que troca a ordem do par $(\lambda, \overline{\lambda})$. Esta indeterminação pode ser resolvida convencionando que nos blocos com entrada complexa temos sempre $b > 0$.*

Exemplo 6.77. *Vamos determinar a forma canónica de Jordan real da transformação linear $T\colon \mathbb{R}^4 \to \mathbb{R}^4$ representada na base canónica pela matriz*

$$A = \begin{bmatrix} 2 & 0 & 0 & -2 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & -1 \\ 2 & 1 & -1 & -1 \end{bmatrix}$$

Calcula-se que os valores próprios são 0, e $1 \pm i$. A multiplicidade geométrica de 0 é 2, sendo uma base para o espaço próprio de 0 dada pelos vetores $(0, 1, 1, 0)$ e $(1, -1, 0, 1)$. A matriz A é portanto diagonalizável enquanto matriz complexa. A forma canónica de Jordan real de A é

$$J = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

Um vetor próprio de $1 - i$ é $(1 - i, 0, -i, 1)$ logo uma matriz S tal que $A = SJS^{-1}$ é, por exemplo,

$$S = \begin{bmatrix} 0 & 1 & 1 & -1 \\ 1 & -1 & 0 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

6.78. **A classificação geral dos endomorfismos de espaços vetoriais de dimensão finita.** Nesta secção, por uma questão de cultura geral, vamos indicar brevemente como se generaliza a classificação dos endomorfismos que vimos quando o corpo de base é $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ a corpos arbitrários. Seja V um espaço vetorial de dimensão finita sobre o corpo $\mathbb{K}$ e $T\colon V \to V$ um endomorfismo qualquer.

Todo o polinómio com coeficientes em $\mathbb{K}$ pode ser fatorizado de forma única como o produto de polinómios ditos *irredutíveis* que desempenham um papel análogo ao dos números primos nos inteiros. Quando $\mathbb{K} = \mathbb{C}$ o Teorema Fundamental da Álgebra garante que os únicos polinómios (mónicos) irredutíveis são da forma $(x - \lambda)$ com $\lambda \in \mathbb{C}$ e daqui segue fácilmente que os únicos polinómios (mónicos) reais irredutíveis são $(x - \lambda)$ com $\lambda \in \mathbb{R}$ e $(x - a)^2 + b^2$ com $a, b \in \mathbb{R}$ e $b \neq 0$.

A versão geral do Teorema da Decomposição Primária 6.37 afirma que todo o endomorfismo T de um espaço vetorial de dimensão finita pode ser decomposto numa soma direta de endomorfismos cujos polinómios característicos são potências de polinómios irredutíveis, havendo uma parcela para cada polinómio irredutível que divide o polinómio característico de T. Quando $\mathbb{K} = \mathbb{C}$ este resultado é precisamente o Teorema da Decomposição Primária 6.37. Quando $\mathbb{K} = \mathbb{R}$, este resultado corresponde à decomposição do endomorfismo em blocos de Jordan reais e complexos (cada valor próprio real e cada par de valores próprios complexos conjugados corresponde a uma parcela).

Vejamos que forma tomam os blocos de uma decomposição primária. Dado $v \in V \setminus \{0\}$, associamos a v o subespaço cíclico $W = L(\{T^i v \colon i \in \mathbb{N}_0\})$. Sendo $p(x) = c_0 + c_1 x + \ldots + x^k$ um polinómio com coeficientes em $\mathbb{K}$ de grau mínimo (dito o polinómio mínimo de v) tal que

$$p(T)v = 0$$

temos que $B = (v, Tv, T^2 v, \ldots, T^{k-1}v)$ é uma base para W e, nessa base para W, a matriz que representa $T_{|W}$ é

$$A_{T_{|W},B,B} = \begin{bmatrix} 0 & 0 & \ldots & -c_0 \\ 1 & 0 & & -c_1 \\ 0 & \ddots & \ddots & \vdots \\ 0 & \ddots & 1 & -c_{k-1} \end{bmatrix}$$

que se chama a *matriz companheira* do polinómio p. Estas matrizes têm uma forma bastante simples: a grande maioria das entradas é igual a zero. O seu polinómio característico é (possivelmente a menos de sinal) exatamente o polinómio $p(\lambda)$ como verificarão no Exercício 28. Usando o Teorema de Cayley-Hamilton, não é díficil verificar que o polinómio mínimo de um dado $v \in V \setminus \{0\}$ é sempre um fator do polinómio característico de T.

Um Teorema fundamental da Álgebra Linear afirma que para todo o endomorfismo T de um espaço vetorial de dimensão finita, existe uma decomposição de V como a soma direta de subespaços cíclicos, ou alternativamente, que existe uma base B de V tal que $A_{T,B,B}$ é diagonal por blocos sendo cada bloco diagonal uma matriz companheira de um fator do polinómio característico[27].

Verão no Exercício 28 que uma fatorização de um polinómio leva a uma decomposição em blocos da sua matriz companheira. Aplicando o resultado mencionado no parágrafo anterior aos blocos de uma decomposição primária (cujos polinómios característicos são potências de polinómios irredutíveis) obtemos uma decomposição diagonal por blocos na qual, ao longo da diagonal, temos matrizes companheiras de polinómios irredutíveis. Quando $\mathbb{K} = \mathbb{C}$ isto é exatamente a forma canónica de Jordan, enquanto que no caso em que $\mathbb{K} = \mathbb{R}$ obtemos uma matriz semelhante à forma canónica de Jordan real.

Para um excelente tratamento elementar dos temas descritos nos parágrafos anteriores recomendamos [HK]. Quem continuar a estudar Álgebra irá provavelmente ver todos estes resultados como um caso particular do *Teorema de classificação dos módulos finitamente gerados sobre domínios de ideais principais.* Um endomorfismo $T \colon V \to V$ dá a V a estrutura de um *módulo*[28] sobre o domínio de ideais principais $\mathbb{K}[t]$ dos polinómios com coeficientes em $\mathbb{K}$: define-se a multiplicação de $p(t) \in \mathbb{K}[t]$ por um elemento $v \in V$ como

[27]Além disso é possível escolher estes fatores de uma forma natural, o que leva à chamada *Forma canónica racional* de um endomorfismo (assim conhecida porque é válida, em particular, sobre o corpo dos racionais).

[28]Se suprimirmos na definição de corpo o requisito da invertibilidade dos elementos não nulos obtemos a noção de anel. Os números inteiros e os polinómios com coeficientes num corpo são exemplos de anéis. Um módulo sobre um anel é o análogo neste contexto de um espaço vetorial sobre um corpo.

$p(t) \cdot v = p(T)v$. As formas normais para os endomorfismos brevemente descritas acima são uma consequência imediata da classificação dos módulos finitamente gerados sobre $\mathbb{K}[t]$.

6.79. **Exercícios.**

1. Determine os valores próprios e os vetores próprios das seguintes matrizes/transformações lineares indicando se as mesmas são diagonalizáveis. Se os valores próprios das matrizes forem complexos, encare a matriz como sendo complexa.
   (a) $\begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix}$ (b) $\begin{bmatrix} 2 & -1 \\ 0 & 1 \end{bmatrix}$ (c) $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ (d) $\begin{bmatrix} 2 & -3 \\ 1 & 2 \end{bmatrix}$
   (e) $\begin{bmatrix} -1 & 2 & 1 \\ 0 & 2 & -1 \\ 0 & 0 & 3 \end{bmatrix}$ (f) $\begin{bmatrix} 2 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ (g) $\begin{bmatrix} 1+i & 0 \\ 1 & -i \end{bmatrix}$
   (h) $f\colon M_{2\times 2}(\mathbb{R}) \to M_{2\times 2}(\mathbb{R})$ definida por $f(A) = \begin{bmatrix} -1 & 2 \\ 0 & 1 \end{bmatrix} A$
   (i) Sendo V o espaço vetorial dos polinómios (reais) de grau ≤ 3, $T\colon V \to V$ é definida por
   $$T(p) = p' + \left(\int_0^1 p(x)dx\right) \frac{3x^2}{10}$$
   Resolução.
2. Considere a transformação linear $T\colon M_{2\times 2}(\mathbb{R}) \to M_{2\times 2}\mathbb{R}$ definida por
   $$T(A) = \begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix} A - A \begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix}$$
   Determine os valores próprios de T, os espaços próprios e indique se T é diagonalizável
   Resolução.
3. Considere a base ordenada $B = ((1, 2, -1), (0, 2, 1), (1, -2, -1))$ de $\mathbb{R}^3$, assim como a base canónica B_c. Sendo $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ a transformação linear tal que
   $$A_{T,B_c,B} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 3 & -1 \\ 1 & 1 & 2 \end{bmatrix}$$
   Determine os valores próprios de T. Resolução.
4. Seja $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ uma transformação linear que tem $(1, 0, 1), (0, 1, 1)$ e $(-1, 0, 1)$ como vetores próprios com valores próprios $1, 2, -1$ respetivamente. Determine a expressão geral de T. Resolução.
5. Determine a expressão geral para a transformação linear $R\colon \mathbb{R}^3 \to R^3$ que geometricamente é a reflexão no plano $P = \{(x, y, z) \in \mathbb{R}^3 \colon x + y + z = 0\}$. Resolução.
6. Suponha que $T\colon \mathbb{R}^4 \to \mathbb{R}^4$ é uma transformação linear que tem como espaços próprios
   $$L(\{(1, 2, -1, 0), (1, 3, 2, 1)\}) \quad \text{e} \quad L(\{(0, 1, 0, 1), (-1, 2, 0, 0)\})$$
   correspondentes a valores próprios distintos.
   (a) T é diagonalizável? Justifique.
   (b) Indique justificadamente se os seguintes vetores são vetores próprios de T.

(i) $(1,1,-4,-1)$ (ii) $(3,2,1,-2)$

Resolução.

7. Seja $V = C^\infty(\mathbb{R}) \subset F(\mathbb{R},\mathbb{R})$ o espaço das funções de classe C^∞ e seja $D^2\colon V \to V$ a transformação linear definida por $D^2(f) = f''$. Calcule os valores próprios e vetores próprios de D^2. *Sugestão: Veja o Exercício 4.39 da ficha sobre transformações lineares e o Exemplo 4.40.* Resolução.
8. Sendo A uma matriz quadrada e $p(x) = a_0 + a_1x + \ldots + a_nx^n$ um polinómio, define-se

$$p(A) = a_0 + a_1A + \ldots + a_nA^n$$

   (a) Sendo S uma matriz invertível, mostre que $p(SAS^{-1}) = S(p(A))S^{-1}$.

   (b) Calcule $p\left(\begin{bmatrix} 1 & -2 \\ 2 & -4 \end{bmatrix}\right)$ para $p(x) = 2 - 3x^4 + x^{11}$. *Sugestão: É muito fácil aplicar um polinómio a uma matriz diagonal.*

   Resolução.
9. Determine A^{721} para a matriz A do Exercício 1(a). Resolução.
10. Resolva os seguintes sistemas de equações diferenciais lineares usando uma mudança de coordenadas apropriada que reescreva o sistema como duas equações independentes (diz-se que estamos a diagonalizar o sistema):

   (a) $\begin{cases} x'(t) = x(t) + 2y(t) \\ y'(t) = 2x(t) + y(t) \end{cases} \Leftrightarrow \begin{bmatrix} x'(t) \\ y'(t) \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x(t) \\ y(t) \end{bmatrix}$ (b) $\begin{cases} x'(t) = x(t) - y(t) \\ y'(t) = -x(t) + y(t) \end{cases}$

   Resolução.
11. Mostre que um endomorfismo $P\colon V \to V$ é uma projeção se e só se P é diagonalizável e tem apenas $0, 1$ como valores próprios. Resolução.
12. Seja $A \in M_{n\times n}(\mathbb{C})$ uma matriz tal que todas as entradas são reais.

   (a) Mostre que se λ é um valor próprio de A com parte imaginária diferente de 0 então qualquer vetor próprio v associado a λ tem necessariamente alguma componente com parte imaginária diferente de 0.

   (b) Mostre que $\overline{\lambda}$ é também um valor próprio de A e que $\overline{v}$ (o vetor que se obtém de v conjugando cada componente) é um vetor próprio de A (linearmente independente de v pelo exercício seguinte).

   (c) Use o resultado da alínea anterior para diagonalizar eficientemente a matriz $\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$

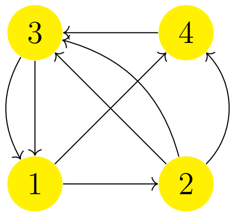
   Resolução.
13. Suponha que $v_1, \ldots, v_n$ são vetores próprios de uma transformação linear f com valores próprios $\lambda_1, \ldots, \lambda_n$ distintos (isto é, $\lambda_i \neq \lambda_j$ para $i \neq j$).

   (a) Mostre que $\{v_1, \ldots, v_n\}$ é um conjunto linearmente independente. *Sugestão: Dada uma combinação linear dos v_i's é possível tornar todos os coeficientes dos v_i excepto um nulos por aplicação de uma transformação linear apropriada! Considere primeiro o caso em que $n = 2$.*

   (b) Seja V um espaço vetorial de dimensão finita. Mostre que $T\colon V \to V$ é diagonalizável se e só se a soma das multiplicidades geométricas dos seus valores próprios for igual a $\dim(V)$.

Resolução.

14. Numa internet com 4 páginas ligadas de acordo com o diagrama seguinte, qual seria a página mais relevante?



Resolução.

15. Determine todos os subespaços invariantes da transformação linear do Exercício 2. Resolução.
16. Seja $T\colon V \to V$ um endomorfismo e W_1, W_2 subespaços invariantes de T. Mostre que $W_1 + W_2$ e $W_1 \cap W_2$ são invariantes. Mais geralmente, mostre que uma soma qualquer e uma intersecção qualquer de subespaços invariantes é invariante. Resolução.
17. Determine os espaços próprios generalizados e a forma canónica de Jordan para as seguintes matrizes:

(a) $\begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & -1 \\ 1 & 1 & 1 \end{bmatrix}$ (b) $\begin{bmatrix} 3 & 2 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -1 & 3 \end{bmatrix}$ (c) $\begin{bmatrix} \frac{5}{2} & 0 & -\frac{1}{2} \\ -1 & 2 & 1 \\ \frac{1}{2} & 0 & \frac{3}{2} \end{bmatrix}$

(d) $\begin{bmatrix} 2 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & 1 & 0 & 2 \end{bmatrix}$

Resolução.

18. Determine o polinómio característico, os valores próprios, indique as respetivas multiplicidades algébricas e geométricas [29] das seguintes matrizes:

(a) $\begin{bmatrix} -1 & 2 \\ 3 & 1 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 0 & 0 \\ 7 & -2 & 0 \\ 1 & 9 & -2 \end{bmatrix}$ (c) $\begin{bmatrix} -1 & 0 & 1 \\ 1 & 0 & -1 \\ -1 & -1 & 2 \end{bmatrix}$

(d) $\begin{bmatrix} -1 & -2 & 2 \\ 0 & 2 & 1 \\ 0 & -1 & 2 \end{bmatrix}$ (e) $\begin{bmatrix} -1 & -2 & -2 \\ 1 & 2 & 1 \\ -1 & -1 & 0 \end{bmatrix}$ (f) $\begin{bmatrix} 1 & 1 & 0 & 1 \\ 2 & 2 & -1 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 3 & 0 \end{bmatrix}$

Resolução.

19. Seja $A \in M_{n\times n}(\mathbb{K})$. O *traço* de A é a soma das entradas diagonais da matriz A:

$$\operatorname{tr} A = a_{11} + a_{22} + \ldots + a_{nn}$$

(a) Mostre que dadas $A, B \in M_{n\times n}(\mathbb{K})$ temos $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.
(b) Usando a alínea anterior mostre que para S invertível temos $\operatorname{tr}(SAS^{-1}) = \operatorname{tr}(A)$.

[29] Se os valores próprios forem números complexos, considere as matrizes como sendo complexas

(c) Para $\mathbb{K} = \mathbb{C}$, use a alínea anterior para concluir que $\operatorname{tr} A$ é a soma dos valores próprios de A repetidos de acordo com as suas multiplicidades algébricas. Observe ainda que $\det A$ é o produto dos valores próprios de A repetidos de acordo com a sua multiplicidade algébrica.

(d) Aproveite os resultados da alínea anterior para achar os valores próprios das seguintes matrizes 2×2

$$(i) \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \qquad (ii) \begin{bmatrix} 3 & -2 \\ 2 & -1 \end{bmatrix}$$

(Para matrizes de maior dimensão, o traço e o determinante dão informação útil sobre os valores próprios mas esta não é suficiente para achar todos os valores próprios como no caso 2×2)

Resolução.

20. Quando é que uma matriz triangular superior (ou inferior) é nilpotente? Resolução.
21. Seja V o espaço dos polinómios (reais) de grau menor ou igual a 2 e seja $T\colon V \to V$ definida por
$$T(p) = p + p(0)t + p(1)t^2$$
Determine se T admite uma forma canónica de Jordan e em caso afirmativo calcule-a. Resolução.
22. Seja V um espaço vetorial real de dimensão finita e seja $T\colon V \to V$ um endomorfismo. Dê uma condição necessária e suficiente para que T admita uma forma canónica de Jordan. Resolução.
23. Seja $T\colon V \to V$ um endomorfismo de um espaço vetorial complexo de dimensão 7. Suponha que 1 é o único valor próprio de T e que a multiplicidade geométrica de 1 é 3. Quais são os possíveis valores para a dimensão do núcleo de $(T - \mathrm{Id})^2$? Resolução.
24. Sendo
$$A = \begin{bmatrix} 1 & 2 & -1 \\ -1 & -1 & 0 \\ 0 & 1 & 2 \end{bmatrix}$$
e $p(x) = 1 - 2x + x^2 - x^5 + x^7$, calcule $p(A)$.
Sugestão: use o Teorema de Cayley-Hamilton Resolução.
25. Usando apenas a existência de valores próprios (isto é, sem usar a forma canónica de Jordan) demonstre o **Teorema de Schur:** Seja V um espaço vetorial complexo de dimensão finita. Então existe uma base B para V tal que $A_{T,B,B}$ é uma matriz triangular superior. Resolução.
26. Seja $V = C^\infty(\mathbb{R})$ e considere o endomorfismo $D\colon V \to V$ dado pelo operador de derivação $D(f) = f'$.
(a) Mostre que o espaço próprio generalizado de $\lambda \in \mathbb{R}$ é
$$E(\lambda) = L(\{e^{\lambda t}, te^{\lambda t}, \ldots, t^n e^{\lambda t}, \ldots\})$$
Sugestão: $y' - \lambda y = f(t) \Leftrightarrow \frac{d}{dt}(e^{-\lambda t} y) = e^{-\lambda t} f(t)$. *Use indução.*

(b) Mostre que a soma de todos os espaços próprios generalizados de D não é todo o V. [30]

Resolução.

27. Seja V um espaço vetorial (não necessariamente finitamente gerado) sobre um corpo $\mathbb{K}$.
   (a) Seja $N\colon V \to V$ um endomorfismo nilpotente e λ um escalar não nulo. Mostre que $\lambda \operatorname{Id} + N\colon V \to V$ é um isomorfismo. *Sugestão:* $(1-x)(1+x+x^2+\ldots+x^{k-1}) = 1-x^k$.
   (b) Dê um exemplo de um espaço vetorial V e de um endomorfismo $T\colon V \to V$ tais que
      - $V = E^g(1)$
      - $(T - \operatorname{Id})\colon V \to V$ não é nilpotente.
   (c) Apesar do exemplo da alínea anterior, mostre que sendo V e $T\colon V \to V$ arbitrários, λ um valor próprio de T e μ um escalar diferente de λ, a restrição de $(T-\mu)$ a $E^g(\lambda)$ é sempre um isomorfismo. Conclua que a soma dos espaços próprios generalizados de $T\colon V \to V$ é sempre direta.

Resolução.

28. Seja V um espaço vetorial de dimensão finita e $v \in V \setminus \{0\}$. Seja k o maior natural tal que
$$\{v, Tv, \ldots, T^{k-1}v\}$$
é um conjunto linearmente independente formado por vetores distintos e $W = L(\{v, Tv, \ldots, T^{k-1}v\})$ o subespaço cíclico gerado por v. Sejam $a_0, \ldots, a_{k-1} \in \mathbb{K}$ os escalares tais que
$$T^k v + a_{k-1}T^{k-1}v + \ldots + a_1 Tv + a_0 v = 0$$
e $p(x)$ o polinómio $x^k + a_{k-1}x^{k-1} + \ldots + a_1 x + a_0$.
   (a) Determine a matriz que representa a restrição da transformação T ao subespaço invariante W com respeito à base $(v, Tv, \ldots T^{k-1}v)$. Esta matriz chama-se a *matriz companheira* do polinómio p.
   (b) Mostre que o polinómio característico da matriz companheira é $(-1)^n p(\lambda)$.
   (c) Determine a matriz companheira que obtemos se v é um vetor próprio de T.
   (d) Sejam $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ a transformação linear definida por
   $$T(x, y, z) = (2x + y + z, x - y + z, x + z)$$
   e $v = (1, 0, 0)$. Determine o espaço cíclico determinado por v e a matriz companheira correspondente.
   (e) Suponha que $p(x) = q(x)r(x)$ com q e r mónicos. Escrevendo A_t para a matriz companheira do polinómio t, determine uma base B para o espaço cíclico determinado por v de tal forma que
   $$A_{T,B,B} = \begin{bmatrix} A_q & 0 \\ E & A_r \end{bmatrix}$$
   sendo E uma matriz que tem todas as entradas nulas, com excepção da entrada no canto superior direito que é igual a 1.

[30] Isto seria também verdade se substituíssemos V pelo espaço vetorial complexo determinado pelas funções C^∞ de $\mathbb{R}$ para $\mathbb{C}$, pela mesma razão.

(f) Recorde-se que duas matrizes A e B se dizem semelhantes, se existe uma matriz invertível S tal que $B = SAS^{\;1}$. Determine uma matriz companheira semelhante a um bloco de Jordan.

Resolução.

29. Seja $T\colon V \to V$ um endomorfismo.

(a) Sendo $v \in V$, considere o conjunto dos polinómios $p(x)$ com coeficientes no corpo $\mathbb{K}$ tais que $p(T)v = 0$

$$I_v = \{a_0 + a_1 x + \ldots + a_n x^n \colon a_i \in \mathbb{K}, a_0 v + a_1 T v + \ldots + a_n T^n v = 0\}$$

Sendo $m(x) \in I_v$ um polinómio (não nulo) de grau *mínimo*, mostre que I_v é o conjunto de todos os múltiplos de $m(x)$, isto é

$$I_v = \{m(x)q(x) \colon q(x) \text{ qualquer }\}$$

O polinómio $m(x)$ fica determinado a menos de multiplicação por escalar. Podemos escolhê-lo de forma única tomando o único polinómio *mónico*, isto é, com coeficiente do termo de maior grau igual a 1. Esse polinómio chama-se o *polinómio mínimo de v* e será denotado por $m_v(x)$ no resto desta pergunta.

Sugestão: Sendo $p(x)$ um elemento de I_v use o algoritmo da divisão para escrever $p(x) = m(x)q(x) + r(x)$ com o grau de $r(x)$ inferior ao de $m(x)$.

(b) Mostre que $m_v(x)$ divide o polinómio caraterístico de T.

(c) O menor múltiplo comum de todos os $m_v(x)$ chama-se o *polinómio mínimo de T*. É o polinómio de menor grau que anula todos os elementos de V. Pela alínea anterior o polinómio mínimo é um fator do polinómio característico. Supondo que $\mathbb{K} = \mathbb{C}$ dê uma condição necessária e suficiente sobre T para que o polinómio mínimo coincida com o polinómio característico.

(d) Determine a menos de semelhança todas as matrizes $A \in M_{4\times 4}(\mathbb{C})$ cujo polinómio mínimo é $(x-1)^2$.

(e) Mostre que, sendo $\mathbb{K} = \mathbb{C}$, existe[31] um vetor $v \in V$ tal que $m_v(x)$ é o polinómio mínimo de T.

Resolução.

30. Seja $A \in M_{8\times 8}(\mathbb{R})$ uma matriz tal que
   - 2 é um valor próprio de A
   - $1 + i$ é um valor próprio de A (vista como matriz complexa) com multiplicidade geométrica 2
   - A não tem outros valores próprios além de 2 e $1 \pm i$

Quais são as possíveis formas canónicas de Jordan reais de A?

Resolução.

31. Determine a forma canónica Jordan real das seguintes matrizes assim como as bases que as colocam em forma canónica.

[31]Isto é na realidade verdade sem qualquer restrição sobre o corpo $\mathbb{K}$.

(a) $\begin{bmatrix} 1 & -1 & 1 \\ 2 & 1 & -1 \\ 1 & 1 & 1 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 2 & 1 & 1 \\ -1 & 3 & 0 & 0 \\ 0 & 0 & 2 & -1 \\ 0 & 0 & 1 & 2 \end{bmatrix}$

Resolução.

7. Espaços vetoriais com produto interno

Neste capítulo final vamos introduzir mais estrutura geométrica nos objetos que temos vindo a estudar. Esta estrutura incluirá por exemplo noções de distância e ângulo entre vetores. Isto é feito com recurso à noção de *produto interno*, um conceito que generaliza a noção de produto escalar em $\mathbb{R}^2$ e $\mathbb{R}^3$ familiar do ensino secundário. Neste capítulo o corpo de base será sempre $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

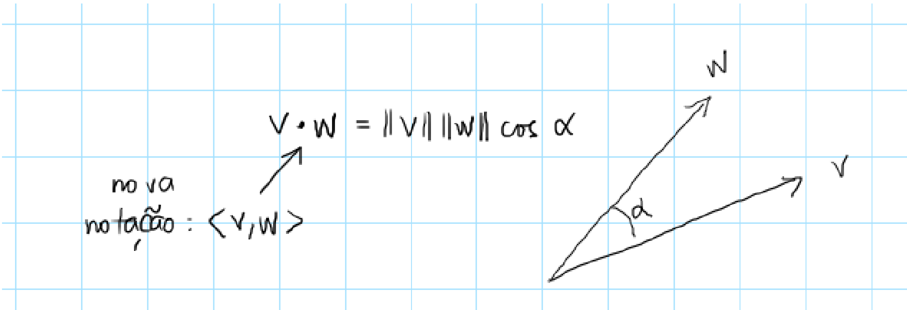
7.1. **Definição de produto interno num espaço vetorial.** Recordemos o *produto interno* (ou escalar) de vetores de $\mathbb{R}^2$ e $\mathbb{R}^3$. Trata-se de uma operação que produz um número real[32] $\langle v, w\rangle$ a partir de dois vetores v e w. É dado pelas fórmulas

$$\langle (x_1, x_2), (y_1, y_2)\rangle = x_1y_1 + x_2y_2 \qquad \text{para } (x_1, x_2), (y_1, y_2) \in \mathbb{R}^2$$

e

$$\langle (x_1, x_2, x_3), (y_1, y_2, y_3)\rangle = x_1y_1 + x_2y_2 + x_3y_3 \qquad \text{para } (x_1, x_2, x_3), (y_1, y_2, y_3) \in \mathbb{R}^3$$

respetivamente. Em ambos os casos, o significado geométrico, do produto interno $\langle v, w\rangle$ é $\|v\|\|w\| \cos\alpha$ em que $\|x\|$ designa o comprimento do vetor x e α é o ângulo entre v e w.



As propriedades destes produtos podem ser abstraídas nos seguintes axiomas simples.

Definição 7.2. *Seja V um espaço vetorial* real. *Um* produto interno *em V é uma função*

$$\langle\cdot,\cdot\rangle \colon V \times V \to \mathbb{R}$$

satisfazendo

(1) **Bilinearidade:** *Para todos os $\alpha_1, \alpha_2 \in \mathbb{R}$ e $v_1, v_2, w \in V$.*
- $\langle \alpha_1 v_1 + \alpha_2 v_2, w\rangle = \alpha_1\langle v_1, w\rangle + \alpha_2\langle v_2, w\rangle$
- $\langle w, \alpha_1 v_1 + \alpha_2 v_2\rangle = \alpha_1\langle w, v_1\rangle + \alpha_2\langle w, v_2\rangle$

(2) **Simetria:** $\langle v, w\rangle = \langle w, v\rangle$ *para todos os $v, w \in V$.*

[32]No ensino secundário, a notação normalmente utilizada para o produto interno dos vetores v e w é $v \cdot w$ mas aqui usaremos a notação $\langle v, w\rangle$, que é a standard.

(3) **Positividade:** *$\langle v, v\rangle > 0$ para todo o $v \neq 0$.*

Observação 7.3. *Tendo em conta a simetria do produto interno, para verificar a bilinearidade basta verificar a primeira (ou a segunda) das igualdades que a caracterizam.*

Exemplo 7.4. *O produto interno usual (ou standard) em $\mathbb{R}^n$ é definido por*

$$\langle (x_1, \ldots, x_n), (y_1, \ldots, y_n)\rangle = x_1y_1 + x_2y_2 + \ldots + x_ny_n$$

É imediato verificar que as propriedades (1)-(3) na Definição 7.2 são verificadas. Este produto interno generaliza o produto interno já conhecido nos casos em que $n = 2$ e 3.

Exemplo 7.5. *Seja $[a, b]$ um intervalo limitado e fechado de $\mathbb{R}$ e $V = C([a, b], \mathbb{R})$ o espaço vetorial das funções contínuas de $[a, b]$ para $\mathbb{R}$ (que é um subespaço vetorial do espaço vetorial de todas as funções de $\mathbb{R}$ para $\mathbb{R}$). Define-se $\langle \cdot, \cdot\rangle \colon V \times V \to \mathbb{R}$ pela expressão*

$$\langle f, g\rangle = \int_a^b f(x)g(x)dx$$

A expressão anterior faz sentido porque o produto de funções contínuas é contínua e uma função contínua é integrável num intervalo compacto. Verifiquemos as propriedades (1)-(3) da Definição 7.2:

(1) $\langle \alpha_1 f_1 + \alpha_2 f_2, g\rangle = \int_a^b (\alpha_1 f_1(x) + \alpha_2 f_2(x))g(x)dx = \alpha_1 \int_a^b f_1(x)g(x)dx + \alpha_2 \int_a^b f_2(x)g(x)dx = \alpha_1\langle f_1, g\rangle + \alpha_2\langle f_2, g\rangle$

(2) É imediato uma vez que $f(x)g(x) = g(x)f(x)$.

(3) $\langle f, f\rangle = \int_a^b f^2(x)dx \geq 0$ por monotonia do integral. Se $f(x) \neq 0$ então existe $x_0 \in [a, b]$ tal que $f(x_0) \neq 0$. Como f é contínua isso significa que existe $\epsilon > 0$ e um intervalo J contendo x_0 com interior não vazio tal que $f(x)^2 \geq \epsilon$ quando $x \in J$. Mas então $\int_a^b f(x)^2dx \geq \int_J f(x)^2dx \geq \int_J \epsilon\, dx > 0$.

Observação 7.6. *Se pensarmos numa função f como um "vetor com componentes indexadas pelos números reais"cuja componente x é o número $f(x)$, e no integral como uma "soma em x"o Exemplo 7.5 é uma generalização natural do Exemplo 7.4.*

Existe também uma versão do conceito de produto interno para um espaço vetorial complexo, que se chama um *produto interno Hermitiano*, ou simplesmente um produto interno. O modelo será $\mathbb{C}^n$, mas agora não podemos usar exatamente a mesma fórmula que nos dá o produto interno standard em $\mathbb{R}^n$ porque perderíamos a propriedade da positividade (que é a chave para definir o comprimento de vetores). A solução é conjugar um dos argumentos coordenada a coordenada, uma vez que $\overline{z}z = |z|^2 \geq 0$. Esta solução requer a modificação dos restantes dois axiomas da forma seguinte.

Definição 7.7. *Seja V um espaço vetorial complexo. Um* produto interno *em V é uma função*

$$\langle \cdot, \cdot\rangle \colon V \times V \to \mathbb{C}$$

satisfazendo

(1) **Sesquilinearidade ou linearidade conjugada:** *Para todos os* $\alpha_1, \alpha_2 \in \mathbb{C}$ *e* $v_1, v_2, w \in V$.
- $\langle \alpha_1 v_1 + \alpha_2 v_2, w\rangle = \overline{\alpha_1}\langle v_1, w\rangle + \overline{\alpha_2}\langle v_2, w\rangle$
- $\langle w, \alpha_1 v_1 + \alpha_2 v_2\rangle = \alpha_1\langle w, v_1\rangle + \alpha_2\langle w, v_2\rangle$

(2) **Simetria conjugada:** $\langle v, w\rangle = \overline{\langle w, v\rangle}$ *para todos os* $v, w \in V$.

(3) **Positividade:** $\langle v, v\rangle$ *é real e positivo para todo o* $v \neq 0$.

Observação 7.8. *Tendo em conta a simetria conjugada de um produto interno complexo, para verificar a sesquilinearidade basta verificar a primeira (ou a segunda) das igualdades que a caracterizam.*

Exemplo 7.9. *O produto interno standard em* $\mathbb{C}^n$ *é a função* $\langle\cdot,\cdot\rangle\colon \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ *definida pela expressão*

$$\langle (z_1, \ldots, z_n), (w_1, \ldots, w_n)\rangle = \overline{z_1}w_1 + \overline{z_2}w_2 + \ldots + \overline{z_n}w_n$$

É imediato verificar as condições (1)-(3) da Definição 7.7. Por exemplo,

$$\langle (z_1, \ldots, z_n), (z_1, \ldots, z_n)\rangle = |z_1|^2 + \ldots + |z_n|^2 \geq 0$$

e só se anula se $z_1 = \cdots = z_n = 0$.

Um produto interno num espaço vetorial real ou complexo permite-nos introduzir noções de comprimento e distância no espaço em questão.

Definição 7.10. *Seja* V *um espaço vetorial e* $\langle\cdot,\cdot\rangle$ *um produto interno em* V. *A* norma *ou* comprimento *de um vetor* $v \in V$ *é o número real não negativo* $\|v\| = \sqrt{\langle v, v\rangle}$. *Sendo* $v, w \in V$, *a* distância de v a w *é o número real não negativo* $\|v - w\|$.

Note-se que as noções de norma e comprimento para o produto interno usual em $\mathbb{R}^2$ ou $\mathbb{R}^3$ são as habituais. Por exemplo:

$$\|(x, y, z)\| = \sqrt{x^2 + y^2 + z^2}$$

Exemplo 7.11. *Em* $\mathbb{C}^2$ *com o produto interno usual,*

$$\|(1 + i, -1)\| = \sqrt{|1 + i|^2 + 1} = \sqrt{2 + 1} = \sqrt{3}$$

Em $C([0, 1], \mathbb{R})$ *a distância entre as funções* x *e* 1 *é*

$$\|x - 1\| = \sqrt{\int_0^1 (x - 1)^2 dx} = \sqrt{\left.\frac{(x - 1)^3}{3}\right|_0^1} = \frac{1}{\sqrt{3}}$$

Além da distância, um outro conceito fundamental de que passamos a dispôr num espaço com produto interno é a noção de ortogonalidade ou perpendicularidade.

Definição 7.12. *Seja* V *um espaço vetorial e* $\langle\cdot,\cdot\rangle$ *um produto interno em* V. *Um subconjunto* $S \subset V$ *diz-se* ortogonal *se* $\langle v, w\rangle = 0$ *para todos os* $v, w \in S$ *distintos. Um subconjunto* $S \subset V$ *diz-se* ortonormado *se* S *é ortogonal e* $\|v\| = 1$ *para todo o* $v \in S$.

Exemplo 7.13. *O conjunto* $\{(1,1),(1,-1)\}$ *é ortogonal em* $\mathbb{R}^2$ *para o produto interno usual, uma vez que* $\langle(1,1),(1-1)\rangle = 1-1 = 0$. *Não é ortonormado uma vez que* $\|(1,1)\| = \sqrt{2} \neq 1$, *mas dividindo cada um dos vetores pelo seu comprimento obtemos o conjunto ortonormado* $\{\left(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}\right),\left(\frac{1}{\sqrt{2}},-\frac{1}{\sqrt{2}}\right)\}$.

As funções $\operatorname{sen} x$ *e* 1 *são ortogonais em* $C([0,2\pi],\mathbb{R})$ *uma vez que*

$$\langle \operatorname{sen} x, 1\rangle = \int_0^{2\pi} \operatorname{sen} x dx = -\cos x|_0^{2\pi} = 0$$

As bases canónicas de $\mathbb{R}^n$ *e* $\mathbb{C}^n$ *são conjuntos ortonormados para os produtos internos usuais.*

7.14. **Representação matricial de um produto interno.** Seja V um espaço vetorial com produto interno $\langle\cdot,\cdot\rangle$ e suponhamos que $B = (v_1,\ldots,v_n)$ é uma base para V.

Podemos escrever dois vetores $v, w \in V$ em função da base B

$$v = \alpha_1 v_1 + \ldots + \alpha_n v_n, \qquad w = \beta_1 v_1 + \ldots + \beta_n v_n,$$

Vamos agora usar a bilinearidade/sesquilinearidade para obter uma fórmula para o produto interno em termos do produto de matrizes. Consideraremos o caso complexo notando que este contém o caso real: os números reais são os números complexos tais que $\overline{\alpha} = \alpha$; para obter as fórmulas no caso real basta suprimir todas as conjugações das fórmulas.

Usando linearidade conjugada na primeira variável temos

$$\langle v, w\rangle = \langle \alpha_1 v_1 + \ldots \alpha_n v_n, w\rangle = \overline{\alpha_1}\langle v_1, w\rangle + \ldots + \overline{\alpha_n}\langle v_n, w\rangle$$

Usando a linearidade na segunda coordenada temos para cada i

$$\langle v_i, w\rangle = \langle v_i, \beta_1 v_1 + \ldots \beta_n v_n\rangle = \beta_1\langle v_i, v_1\rangle + \ldots + \beta_n\langle v_i, v_n\rangle$$

e substituindo na primeira expressão obtemos a seguinte expressão para o produto interno

$$\langle v, w\rangle = \overline{\alpha_1}\beta_1\langle v_1,v_1\rangle + \ldots + \overline{\alpha_1}\beta_n\langle v_1,v_n\rangle + \overline{\alpha_2}\beta_1\langle v_2,v_1\rangle + \ldots + \overline{\alpha_n}\beta_1\langle v_n,v_1\rangle + \ldots + \overline{\alpha_n}\beta_n\langle v_n,v_n\rangle$$

Vemos assim que o produto interno é completamente determinado pelo conjunto de n^2 escalares $\langle v_i, v_j\rangle$ com $i,j = 1,\ldots,n$. Identificando escalares com matrizes 1×1 a expressão anterior pode ser escrita matricialmente na forma

$$\begin{bmatrix} \overline{\alpha_1} & \overline{\alpha_2} & \cdots & \overline{\alpha_n} \end{bmatrix} \begin{bmatrix} \langle v_1,v_1\rangle & \langle v_1,v_2\rangle & \cdots & \langle v_1,v_n\rangle \\ \langle v_2,v_1\rangle & \langle v_2,v_2\rangle & \cdots & \langle v_2,v_n\rangle \\ \vdots & \vdots & & \vdots \\ \langle v_n,v_1\rangle & \langle v_n,v_2\rangle & \cdots & \langle v_n,v_n\rangle \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} \tag{54}$$

Proposição 7.15. *Seja* V *um espaço vetorial de dimensão finita sobre* $\mathbb{K} = \mathbb{R}$ *ou* $\mathbb{C}$ *e seja* $\langle\cdot,\cdot\rangle$ *um produto interno em* V. *Dada uma base ordenada* $B = (v_1,\ldots,v_n)$ *para* V, *existe uma única matriz* $G_B \in M_{n\times n}(\mathbb{K})$, *chamada a* matriz da métrica *ou* matriz de Gram *tal que*

$$\langle v, w\rangle = \begin{cases} [v]_B^T G_B [w]_B & \text{se } \mathbb{K} = \mathbb{R} \\ \overline{[v]_B}^T G_B [w]_B & \text{se } \mathbb{K} = \mathbb{C} \end{cases}$$

Esta matriz é dada pela fórmula

$$G_B = [\langle v_i, v_j\rangle] \tag{55}$$

Dem. Resta-nos apenas provar a unicidade. Consideremos apenas o caso real uma vez que o caso complexo é inteiramente análogo: se $A \in M_{n\times n}(\mathbb{R})$ é tal que $\langle v, w\rangle = [v]_B^T A[w]_B$ então tomando $v = v_i$, $w = w_j$ obtemos

$$\langle v_i, w_j\rangle = \begin{bmatrix} 0 & \cdots & 1 & \cdots & 0 \end{bmatrix} A \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = a_{ij}$$

pelo que a matriz A é necessariamente dada pela fórmula (55). $\square$

Observação 7.16. *Para chegar à expressão* (54) *usámos apenas a propriedade (1) das Definições 7.2 e 7.7 pelo que a expressão matricial* (54) *se aplica a funções de $V \times V \to \mathbb{K}$ que satisfaçam o axioma (1) (ditas funções bilineares no caso real, e sesquilineares no caso complexo).*

As propriedades (2) e (3) na definição de produto interno impõem algumas condições sobre a matriz G_B. Quanto à condição (2), escrevendo g_{ij} para a entrada ij da matriz G_B, temos no caso real

$$g_{ij} = \langle v_i, v_j\rangle = \langle v_j, v_i\rangle = g_{ji} \quad \Leftrightarrow \quad G_B = G_B^T$$

ou seja, a matriz da métrica é *simétrica*. No caso complexo temos

$$g_{ij} = \langle v_i, v_j\rangle = \overline{\langle v_j, v_i\rangle} = \overline{g_{ji}} \quad \Leftrightarrow \quad G_B = \overline{G_B}^T$$

Diz-se que a matriz G_B é *hermitiana*. Reciprocamente, se G é uma matriz que satisfaz estas condições é imediato verificar que a função

$$\langle v, w\rangle = \overline{[v]_B}^T G[w]_B$$

satisfaz as condições (1) e (2) nas definições 7.2 e 7.7 (uma tal função diz-se uma forma bilinear (resp. sesquilinear) simétrica).

Quanto à condição (3), ela claramente implica que os valores próprios de uma matriz da métrica têm que ser *reais* positivos (mesmo no caso em que $\mathbb{K} = \mathbb{C}$): se $G_B[v]_B = \lambda[v]_B$ então

$$\|v\|^2 = \langle v, v\rangle = \overline{[v]_B}^T G_B[v]_B = \lambda\overline{[v]_B}^T [v]_B = \lambda\|[v]_B\|^2 > 0 \quad \Rightarrow \lambda > 0$$

Observação 7.17. *Veremos em breve que uma matriz simétrica ou hermitiana é sempre diagonalizável. Admitindo este facto, é imediato verificar que dada uma base B para V e uma matriz G simétrica (resp. hermitiana) com valores próprios todos positivos, a expressão $\langle v, w\rangle = \overline{[v]_B}^T G[w]_B$ define um produto interno em V.*

Assim, uma base B para o espaço vetorial V estabelece uma correspondência biunívoca entre os produtos internos em V e as matrizes simétricas (resp. hermitianas) com valores próprios todos positivos.

Exemplo 7.18. *Consideremos a restrição do produto interno usual em* $\mathbb{R}^3$ *ao subespaço* $V = \{(x,y,z) \in \mathbb{R}^3 \colon x+y+z=0\}$. *Uma base para* V *é dada, por exemplo, pelos vetores* $v_1 = (1,-1,0)$ *e* $v_2 = (0,1,-1)$. *A matriz da métrica para o produto interno em* V *com respeito à base* $B = (v_1, v_2)$ *é portanto*

$$G_B = \begin{bmatrix} \langle v_1, v_1\rangle & \langle v_1, v_2\rangle \\ \langle v_2, v_1\rangle & \langle v_2, v_2\rangle \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

Dados vetores $v, w \in V$ *com* $[v]_B = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ *e* $[w]_B = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ *temos*

$$\langle v, w\rangle = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} -3 \\ 3 \end{bmatrix} = 3$$

Podemos confirmar este resultado fazendo as contas em $\mathbb{R}^3$*: Temos*

$$v = 1\cdot(1,-1,0)+2(0,1,-1) = (1,1,-2), \quad w = -1\cdot(1,-1,0)+1\cdot(0,1,-1) = (-1,2,-1)$$

logo

$$\langle v, w\rangle = 1\cdot(-1) + 1\cdot 2 + (-2)\cdot(-1) = -1+2+2 = 3.$$

O ponto do exemplo anterior é o seguinte: mesmo que estejamos interessados apenas no produto interno usual em $\mathbb{R}^n$ (isto é na noção usual de comprimento e ângulo) em certas situações estaremos interessados em considerar apenas vetores que estão em certos subespaços (imaginemos por exemplo que um avião voa num dado plano) e para fazer contas nesse plano é mais prático escolher coordenadas no plano (da mesma forma que à superfície da Terra utilizamos duas coordenadas para descrever um ponto). No plano não há em geral coordenadas canónicas como em $\mathbb{R}^n$ e numas coordenadas arbitrárias que escolhamos, a expressão do produto interno não será aquela a que estamos acostumados, mesmo que o produto interno em questão provenha do produto interno usual em $\mathbb{R}^n$.

Observação 7.19. *Note-se que uma base* B *para um espaço vetorial* V *é ortogonal com respeito a um produto interno sse a matriz da métrica* G_B *é diagonal (e então as entradas diagonais são positivas e iguais às normas dos vetores da base ao quadrado) e que* B *é ortonormada (isto é um conjunto ortonormado) sse* G_B *é a matriz identidade (e portanto, nessas coordenadas, o produto interno calcula-se exatamente da mesma forma que o produto interno usual em* $\mathbb{R}^n$ *ou* $\mathbb{C}^n$*).*

Suponhamos agora que B, B' são duas bases para o espaço vetorial V com produto interno. Como se relacionam as matrizes da métrica com respeito às duas bases?

Sendo $S = S_{B\to B'}$ a matriz de mudança de coordenadas da base B para a base B' temos para qualquer $x \in V$

$$[x]_{B'} = S[x]_B$$

substituindo na expressão para a matriz da métrica na base B' temos (novamente o caso real obtém-se omitindo os conjugados)

$$\langle v, w\rangle = \overline{[v]_{B'}}^T G_{B'}[w]_{B'} = \overline{S[v]_B}^T G_{B'}(S[w]_B) = \overline{[v]_B}^T \overline{S}^T G_{B'} S[w]_B$$

onde usámos que $\overline{A}\,\overline{B} = \overline{AB}$ e $(AB)^T = B^T A^T$. Tendo em conta a expressão

$$\langle v, w\rangle = \overline{[v]_B}^T G_B [w]_B$$

que caracteriza a matriz da métrica com respeito à base B conclui-se que

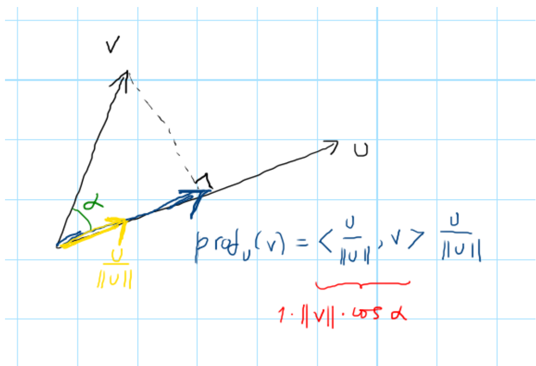
$$G_B = \overline{S}^T G_{B'} S \quad \text{ou, no caso real,} \quad G_B = S^T G_{B'} S \tag{56}$$

Estas fórmulas explicam como a expressão para o produto interno é alterada por uma mudança de coordenadas e são inteiramente análogas à fórmula que descreve a maneira como as expressões matriciais de um endomorfismo em duas bases distintas estão relacionadas.

7.20. **As desigualdades triangular e de Cauchy-Schwarz.**

Definição 7.21. *Seja V um espaço vetorial com produto interno, $v \in V$ e $u \in V \setminus \{0\}$ um vetor não nulo. Define-se a* projeção ortogonal de v sobre u *(com respeito ao produto interno dado) por*

$$\operatorname{proj}_u(v) = \langle u, v\rangle \frac{u}{\|u\|^2} = \frac{\langle u, v\rangle}{\langle u, u\rangle} u = \langle \frac{u}{\|u\|}, v\rangle \frac{u}{\|u\|} \tag{57}$$



As expressões do lado direito do primeiro sinal de igual em (57) são todas iguais pela definição de norma e pela linearidade na primeira variável (no caso complexo note-se que o escalar $\frac{1}{\|u\|}$ é real e portanto igual ao seu conjugado).

Quando $V = \mathbb{R}^2$ ou $\mathbb{R}^3$ com o produto interno usual, a definição anterior coincide com a noção de projeção ortogonal já estudada no ensino secundário. De facto o vetor $\frac{u}{\|u\|}$ é um versor da direção determinada por u (isto é, tem a mesma direção e sentido e comprimento 1). O escalar que multiplica este versor é

$$\langle \frac{u}{\|u\|}, v\rangle = \left\| \frac{u}{\|u\|} \right\| \|v\| \cos\alpha = 1 \cdot \|v\| \cos\alpha = \|v\| \cos\alpha$$

com α o ângulo entre u e v, pelo que a expressão 57 é, neste caso, a expressão familiar do ensino secundário.

Observação 7.22. *No caso complexo é importante que o produto interno na expressão para* $\operatorname{proj}_u(v)$ *seja realizado na ordem indicada, caso contrário a expressão não será linear em v.*

Exemplo 7.23. *A projeção ortogonal de* $(1,-1,2)$ *sobre o vetor* $(0,1,1)$ *com respeito ao produto interno usual em* $\mathbb{R}^3$ *é*

$$\frac{\langle(1,-1,2),(0,1,1)\rangle}{\langle(0,1,1),(0,1,1)\rangle}(0,1,1)=\tfrac{1}{2}(0,1,1)=\left(0,\tfrac{1}{2},\tfrac{1}{2}\right)$$

Note-se que $\mathrm{proj}_u(v)$ é colinear com u (por definição, $\mathrm{proj}_u(v)$ é um múltiplo escalar de u). Claramente proj_u é uma transformação linear (pois o produto interno é linear na segunda coordenada) e, como

$$\mathrm{proj}_u(\lambda u)=\langle u,\lambda u\rangle\frac{u}{\|u\|^2}=\lambda\|u\|^2\frac{u}{\|u\|^2}=\lambda u$$

temos que

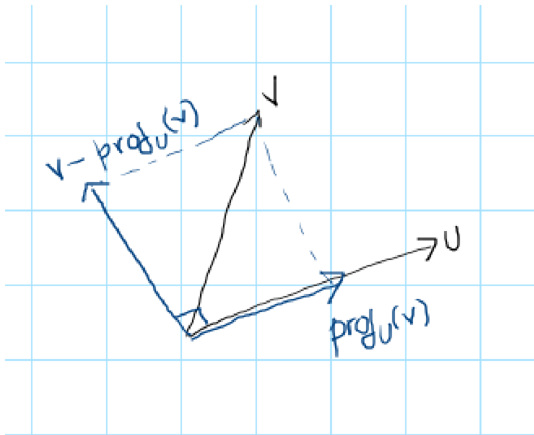
- A imagem de proj_u é exatamente $L(\{u\})$
- proj_u é a identidade na sua imagem, isto é $\mathrm{proj}_u^2=\mathrm{proj}_u$ (ou seja, proj_u é uma projeção).

A transformação linear proj_u é portanto uma projeção na reta gerada por u. As direções de projeção são as que estão contidas no núcleo de proj_u que é precisamente o plano perpendicular a u:

$$\mathrm{proj}_u(v)=0\Leftrightarrow\langle u,v\rangle\frac{u}{\|u\|^2}=0\Leftrightarrow\langle u,v\rangle=0$$

A transformação proj_u permite assim escrever qualquer vetor v como a soma de um vetor colinear com u e outro ortogonal a u:

$$v=(v-\mathrm{proj}_u(v))+\mathrm{proj}_u(v)$$



Da consideração da componente ortogonal a um vetor u, vêm duas desigualdades fundamentais.

Proposição 7.24. *Seja* V *um espaço vetorial (real ou complexo) com produto interno* $\langle\cdot,\cdot\rangle$, *e* $u,v\in V$. *Então*

(i) **Desigualdade de Cauchy-Schwarz:** $|\langle u,v\rangle|\leq\|u\|\|v\|$

(ii) **Desigualdade triangular:** $\|u+v\|\leq\|u\|+\|v\|$

A igualdade verifica-se na primeira desigualdade se e só se u *e* v *são colineares.*

Dem. (i) Podemos assumir sem perda de generalidade que $u \neq 0$ (pois nesse caso $0 = |\langle u, v\rangle| = \|u\|\|v\|$ e u, v são colineares). Nesse caso temos, pela positividade do produto interno

$$\begin{aligned} 0 \leq \|v - \mathrm{proj}_u(v)\|^2 &= \langle v - \frac{\langle u,v\rangle}{\langle u,u\rangle}u, v - \frac{\langle u,v\rangle}{\langle u,u\rangle}u\rangle \\ &= \langle v,v\rangle - \frac{\overline{\langle u,v\rangle}}{\langle u,u\rangle}\langle u,v\rangle - \frac{\langle u,v\rangle}{\langle u,u\rangle}\langle v,u\rangle + \frac{\overline{\langle u,v\rangle}}{\langle u,u\rangle}\frac{\langle u,v\rangle}{\langle u,u\rangle}\langle u,u\rangle \\ &= \|v\|^2 - \frac{|\langle u,v\rangle|^2}{\|u\|^2} \end{aligned}$$

e esta desigualdade é equivalente a

$$|\langle u,v\rangle|^2 \leq \|u\|^2\|v\|^2$$

Tomando raízes quadradas obtemos a desigualdade de Cauchy-Schwarz. A igualdade verifica-se apenas quando $v - \mathrm{proj}_u(v) = 0$, o que acontece se e só se v é um múltiplo escalar de u.

(ii) Temos

$$\|u+v\|^2 = \langle u+v, u+v\rangle = \langle u,u\rangle + \langle u,v\rangle + \langle v,u\rangle + \langle v,v\rangle \tag{58}$$

Uma vez que $z + \overline{z} = 2\,\mathrm{Re}(z) \leq 2|z|$ temos

$$\langle u,v\rangle + \langle v,u\rangle = 2\,\mathrm{Re}(\langle u,v\rangle) \leq 2|\langle u,v\rangle| \leq 2\|u\|\|v\|$$

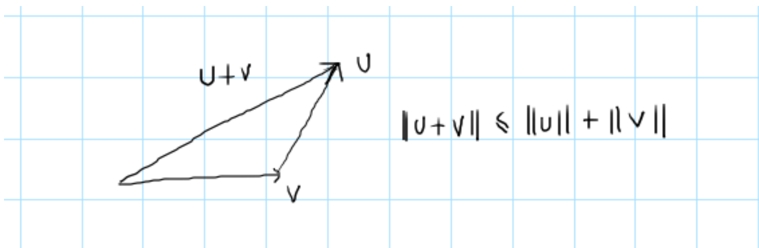
onde na segunda desigualdade aplicámos a desigualdade de Cauchy-Schwarz. Substituindo em (58) obtemos

$$\|u+v\|^2 \leq \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 = (\|u\| + \|v\|)^2$$

que é equivalente à desigualdade triangular.

□

Observação 7.25. *(i) A desigualdade triangular chama-se assim porque $u, v, u+v$ formam as arestas de um triângulo em V e a desigualdade diz precisamente que o comprimento de um dos lados de um triângulo é sempre menor ou igual à soma do comprimento dos dois outros lados.*



(ii) Quando u, v são ortogonais, a expressão (58) *é o* **Teorema de Pitágoras:** $\|u+v\|^2 = \|u\|^2 + \|v\|^2$**.**

7.26. **A definição de ângulo. Complementos ortogonais.**

Definição 7.27. *Seja V um espaço vetorial* real *e $v, w \in V$ vetores não nulos. Define-se o* ângulo *entre v e w como o único $\alpha \in [0, \pi]$ tal que*

$$\cos \alpha = \frac{\langle v, w \rangle}{\|v\|\|w\|}$$

(Isto faz sentido porque, pela desigualdade de Cauchy-Schwarz a expressão do lado direito do sinal de igual pertence ao intervalo $[-1, 1]$.)

Note-se que $v, w \in V$ são ortogonais se e só se o ângulo entre v e w é $\frac{\pi}{2}$.

Exemplo 7.28. *O ângulo entre as funções x e x^2 em $C([0,1], \mathbb{R})$ é*

$$\arccos \frac{\langle x, x^2 \rangle}{\|x\|\|x^2\|} = \frac{\int_0^1 x^3 dx}{\sqrt{\int_0^1 x^2 dx \int_0^1 x^4 dx}} = \arccos \frac{\frac{1}{4}}{\sqrt{\frac{1}{3}\frac{1}{5}}} = \arccos \frac{\sqrt{15}}{4}$$

Definição 7.29. *Seja V um espaço vetorial com um produto interno e $S \subset V$ um subconjunto qualquer. Define-se o subespaço ortogonal a S por*

$$S^{\perp} = \{v \in V : \langle v, x \rangle = 0 \text{ para todo o } x \in S\}$$

Se $W \subset V$ é um subespaço vetorial, chama-se a $W^{\perp}$ o complemento ortogonal *de W em V.*

É imediato verificar que $S^{\perp}$ é um subespaço vetorial de V: claramente $0 \in S^{\perp}$ e se $v_1, v_2 \in S^{\perp}$ e $\alpha_1, \alpha_2 \in \mathbb{R}$ temos $\langle \alpha_1 v_1 + \alpha_2 v_2, x \rangle = \alpha_1 \langle v_1, x \rangle + \alpha_2 \langle v_2, x \rangle = 0$ para todo o $x \in S$, pelo que $\alpha_1 v_1 + \alpha_2 v_2 \in S^{\perp}$.

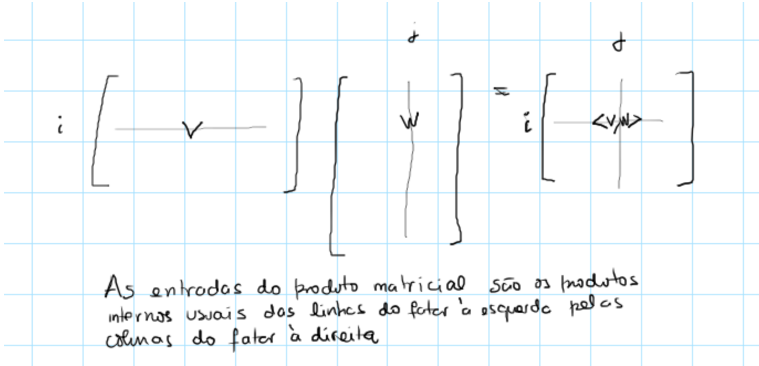
Proposição 7.30. $S^{\perp} = L(S)^{\perp}$

Dem. Uma vez que $S \subset L(S)$, é evidente que $L(S)^{\perp} \subset S^{\perp}$ (se um vetor é ortogonal a todos os elementos de $L(S)$, certamente é também ortogonal a todos os vetores de S). Reciprocamente, seja $w \in S^{\perp}$. Dado $v \in L(S)$, existem vetores $v_1, \ldots, v_k$ em S e escalares $\alpha_1, \ldots, \alpha_k$ tais que $v = \alpha_1 v_1 + \ldots + \alpha_k v_k$. Então

$$\langle w, v \rangle = \langle w, \alpha_1 v_1 + \ldots + \alpha_k v_k \rangle = \alpha_1 \langle w, v_1 \rangle + \ldots + \alpha_k \langle w, v_k \rangle = 0$$

logo $w \in L(S)^{\perp}$. Isso mostra que $S^{\perp} \subset L(S)^{\perp}$ e conclui a demonstração. $\square$

Exemplo 7.31. *(i) Se $A \in M_{m \times n}(\mathbb{R})$ então $N(A) = EL(A)^{\perp} \subset \mathbb{R}^n$ (onde o produto interno considerado é o usual). De facto, pela definição do produto de matrizes,*



$x \in \mathbb{R}^n$ está no núcleo de A sse é ortogonal às linhas de A para o produto interno usual em $\mathbb{R}^n$, e pela Proposição anterior isto é o mesmo que ser ortogonal ao espaço das linhas.

(ii) Se B é uma base de V (ou mais geralmente um conjunto de geradores) então $B^\perp = \{0\}$. De facto, $B^\perp = L(B)^\perp = V^\perp$. Mas a positividade do produto interno diz-nos que o único vetor que é perpendicular a si próprio é o vetor 0. Logo $V^\perp = \{0\}$.

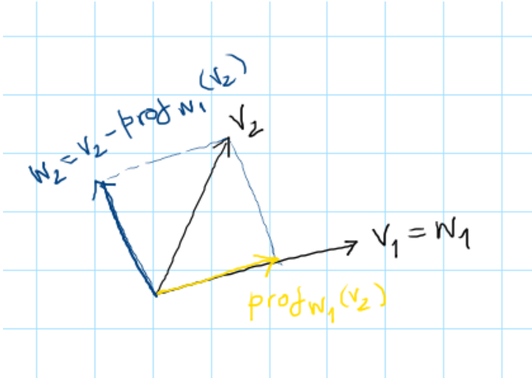
7.32. **O método de ortogonalização de Gram-Schmidt.** A operação de projeção ortogonal segundo um vetor permite-nos facilmente achar uma base ortogonal a partir de uma base qualquer.

Proposição 7.33 (**Método de ortogonalização de Gram-Schmidt).** *Seja V um espaço vetorial com produto interno e $\{v_1, \ldots, v_k\} \subset V$ um conjunto linearmente independente. Então os vetores definidos indutivamente pelas fórmulas*

$$\begin{aligned}
w_1 &= v_1 \\
w_2 &= v_2 - \mathrm{proj}_{w_1}(v_2) \\
w_3 &= v_3 - \mathrm{proj}_{w_1}(v_3) - \mathrm{proj}_{w_2}(v_3) \\
\vdots \quad & \quad \vdots \\
w_k &= v_k - \mathrm{proj}_{w_1}(v_k) - \ldots - \mathrm{proj}_{w_{k-1}}(v_k)
\end{aligned}$$

formam um conjunto ortogonal $\{w_1, \ldots, w_k\}$ tal que, para cada $i = 1, \ldots, k$, temos

$$L(\{v_1, \ldots, v_i\}) = L(\{w_1, \ldots, w_i\})$$



Dem. Vamos usar indução em i para ver que $\{w_1, \ldots, w_i\}$ é um conjunto ortogonal e $L(\{v_1, \ldots, v_i\}) = L(\{w_1, \ldots, w_i\})$. A base da indução é o caso $i = 1$, que é óbvio porque um conjunto com um único vetor é ortogonal e, por definição, $w_1 = v_1$.

Seja $i > 1$ e assumamos por indução que o resultado é válido para $i - 1$. Vejamos primeiro que $L(\{v_1, \ldots, v_i\}) = L(\{w_1, \ldots, w_i\})$. Temos que verificar duas inclusões

- Por hipótese de indução $v_1, \ldots, v_{i-1} \in L(\{w_1, \ldots, w_{i-1}\}) \subset L(\{w_1, \ldots, w_i\})$. Uma vez que $\text{proj}_u(v)$ é um múltiplo de u, a seguinte reformulação da definição de w_i
$$v_i = w_i + \text{proj}_{w_1}(v_i) + \ldots + \text{proj}_{w_{i-1}}(v_i)$$
mostra que $v_i \in L(\{w_1, \ldots, w_i\})$. Conclui-se que $L(\{v_1, \ldots, v_i\}) \subset L(\{w_1, \ldots, w_i\})$
- Novamente, por hipótese de indução, temos $L(\{w_1, \ldots, w_{i-1}\}) \subset L(\{v_1, \ldots, v_i\})$. Na expressão para w_i
$$w_i = v_i - \text{proj}_{w_1}(v_i) - \ldots - \text{proj}_{w_{i-1}}(v_i)$$
os termos precedidos por um sinal menos formam uma combinação linear de $w_1, \ldots, w_{i-1}$ e portanto, por hipótese de indução, pertencem a $L(\{v_1, \ldots, v_{i-1}\})$. Conclui-se que $w_i \in L(\{v_1, \ldots, v_i\})$ e portanto que $L(\{w_1, \ldots, w_i\}) \subset L(\{v_1, \ldots, v_i\})$.

Para ver que $\{w_1, \ldots, w_i\}$ é um conjunto ortogonal basta-nos ver que $\langle w_j, w_i\rangle = 0$ para $j < i$ pois a hipótese de indução diz-nos que $\langle w_j, w_l\rangle = 0$ para $j \neq l$ quando $j, l < i$. Ora

$$\begin{aligned}
\langle w_j, w_i\rangle &= \langle w_j, v_i - \text{proj}_{w_1}(v_i) - \ldots - \text{proj}_{w_{i-1}}(v_i)\rangle \\
&= \langle w_j, v_i\rangle - \langle w_j, \frac{\langle w_1, v_i\rangle}{\langle w_1, w_1\rangle} w_1\rangle - \ldots - \langle w_j, \frac{\langle w_{i-1}, v_i\rangle}{\langle w_{i-1}, w_{i-1}\rangle} w_{i-1}\rangle \\
&= \langle w_j, v_i\rangle - \frac{\langle w_1, v_i\rangle}{\langle w_1, w_1\rangle}\langle w_j, w_1\rangle - \ldots - \frac{\langle w_{i-1}, v_i\rangle}{\langle w_{i-1}, w_{i-1}\rangle}\langle w_j, w_{i-1}\rangle
\end{aligned}$$

Do lado direito do sinal de igual, novamente pela hipótese de indução que $\{w_1, \ldots, w_{i-1}\}$ é ortogonal, o único termo $\langle w_j, w_k\rangle$ que é não nulo é o termo correspondente a $k = j$ portanto

$$\langle w_j, w_i\rangle = \langle w_j, v_i\rangle - 0 - \ldots - \frac{\langle w_j, v_i\rangle}{\langle w_j, w_j\rangle}\langle w_j, w_j\rangle - \ldots - 0 = \langle w_j, v_i\rangle - \langle w_j, v_i\rangle = 0$$

o que conclui a demonstração. $\square$

Exemplo 7.34. *Vamos achar uma base ortonormada para o subespaço*

$$V = \{(x, y, z, w) \in \mathbb{R}^4 \colon x + y + w = 0\} \subset \mathbb{R}^4$$

Uma base para este subespaço é por exemplo

$$\{(1, 0, 0, -1), (0, 1, 0, -1), (0, 0, 1, 0)\}$$

Vamos aplicar o processo de ortogonalização de Gram-Schmidt, dividindo depois os vetores resultantes pelas suas normas para obter uma base ortonormada.

O primeiro vetor da base ortonormada será

$$w_1 = \frac{(1, 0, 0, -1)}{\|(1, 0, 0, -1)\|} = (\tfrac{1}{\sqrt{2}}, 0, 0, -\tfrac{1}{\sqrt{2}})$$

Obtemos um vetor ortogonal através da expressão

$$w_2 = (0, 1, 0, -1) - \langle w_1, (0, 1, 0, -1)\rangle w_1 = (0, 1, 0, -1) - \tfrac{1}{\sqrt{2}}(\tfrac{1}{\sqrt{2}}, 0, 0, -\tfrac{1}{\sqrt{2}}) = (-\tfrac{1}{2}, 1, 0, -\tfrac{1}{2})$$

Na expressão anterior não foi necessário dividir por $\langle w_1, w_1\rangle$ *porque* $\|w_1\| = 1$*. Dividindo pela norma obtemos o segundo vetor da base ortonormada*

$$\tilde{w}_2 = \frac{1}{\|w_2\|}(-\tfrac{1}{2}, 1, 0, -\tfrac{1}{2}) = \sqrt{\tfrac{2}{3}}(-\tfrac{1}{2}, 1, 0, -\tfrac{1}{2})$$

O vetor $v_3 = (0, 0, 1, 0)$ *já é ortogonal a* w_1 *e* $\tilde{w}_2$ *e tem norma* 1*, pelo que podemos tomar para base ortonormada de* V *o conjunto*

$$\{(\tfrac{1}{\sqrt{2}}, 0, 0, -\tfrac{1}{\sqrt{2}}), (-\tfrac{1}{\sqrt{6}}, \sqrt{\tfrac{2}{3}}, 0, -\tfrac{1}{\sqrt{6}}), (0, 0, 1, 0)\}$$

7.35. **Cálculos em bases ortogonais e projeção ortogonal.** As bases ortogonais são extremamente úteis porque tornam os cálculos muito mais fáceis. Começamos por observar que um conjunto ortogonal sem vetores nulos é necessariamente linearmente independente

Proposição 7.36. *Seja* V *um espaço vetorial com produto interno e* $S \subset V \setminus \{0\}$ *um conjunto ortogonal de vetores não nulos. Então* S *é linearmente independente.*

Dem. Sejam $v_1, \ldots, v_k$ elementos de S e suponhamos que

$$\alpha_1 v_1 + \ldots + \alpha_k v_k = 0 \tag{59}$$

Queremos ver que os coeficientes α_i são todos nulos. Como S é ortogonal temos $\langle v_i, v_j\rangle = 0$ para $i \neq j$. Fazendo o produto interno da equação com v_i obtemos

$$\langle v_i, \alpha_1 v_1 + \ldots + \alpha_k v_k\rangle = \langle v_i, 0\rangle = 0$$

Do lado esquerdo temos

$$\alpha_1\langle v_i, v_1\rangle + \ldots + \alpha_i\langle v_i, v_i\rangle + \ldots + \alpha_k\langle v_i, v_k\rangle = \alpha_1 \cdot 0 + \ldots + \alpha_i\|v_i\|^2 + \ldots + \alpha_k \cdot 0$$

Portanto $\alpha_i\|v_i\|^2 = 0$. Como $v_i \neq 0$, conclui-se que $\alpha_i = 0$. $\square$

O resultado seguinte, embora muito simples, é uma das principais razões para a utilização de bases ortogonais ou ortonormadas. Juntamente com as noções de valor e vetor próprio será um dos resultados de Álgebra Linear que mais vezes será utilizado nas aplicações. Diz essencialmente que é muito fácil calcular as coordenadas de um vetor numa base ortogonal. Não é necessário resolver um sistema linear, basta fazer uma conta muito simples.

Proposição 7.37. *Seja $B = (v_1, \ldots, v_n)$ uma base ortogonal para o espaço com produto interno V. Então dado $v \in V$ as coordenadas de v na base B são dadas pela expressão*

$$[v]_B = \begin{bmatrix} \frac{\langle v_1, v\rangle}{\langle v_1, v_1\rangle} \\ \vdots \\ \frac{\langle v_n, v\rangle}{\langle v_n, v_n\rangle} \end{bmatrix}$$

Dem. Sendo $v \in V$, temos a mostrar que

$$v = \frac{\langle v_1, v\rangle}{\langle v_1, v_1\rangle} v_1 + \ldots + \frac{\langle v_n, v\rangle}{\langle v_n, v_n\rangle} v_n \Leftrightarrow v - \frac{\langle v_1, v\rangle}{\langle v_1, v_1\rangle} v_1 - \ldots - \frac{\langle v_n, v\rangle}{\langle v_n, v_n\rangle} v_n = 0$$

De acordo com o Exemplo 7.31(ii) basta ver que o vetor do lado esquerdo da segunda igualdade é ortogonal aos elementos da base B. Ora

$$\begin{aligned} \langle v_i, v - \frac{\langle v_1, v\rangle}{\langle v_1, v_1\rangle} v_1 - \ldots - \frac{\langle v_n, v\rangle}{\langle v_n, v_n\rangle} v_n\rangle &= \langle v_i, v\rangle - \frac{\langle v_1, v\rangle}{\langle v_1, v_1\rangle}\langle v_i, v_1\rangle - \ldots - \frac{\langle v_n, v\rangle}{\langle v_n, v_n\rangle}\langle v_i, v_n\rangle \\ &= \langle v_i, v\rangle - 0 - \ldots - \frac{\langle v_i, v\rangle}{\langle v_i, v_i\rangle}\langle v_i, v_i\rangle - \ldots - 0 \\ &= \langle v_i, v\rangle - \langle v_i, v\rangle = 0 \end{aligned}$$

□

Exemplo 7.38. *Numa base ortonormada as contas da Proposição anterior são ainda mais simples porque os denominadores das expressões para as coordenadas são* 1. *Considerando a base ortonormada*

$$B = \left((\tfrac{1}{\sqrt{2}}, 0, 0, -\tfrac{1}{\sqrt{2}}), (-\tfrac{1}{\sqrt{6}}, \sqrt{\tfrac{2}{3}}, 0, -\tfrac{1}{\sqrt{6}}), (0, 0, 1, 0) \right)$$

do Exemplo 7.34 e o vetor $(1, 1, 3, -2) \in V$ temos

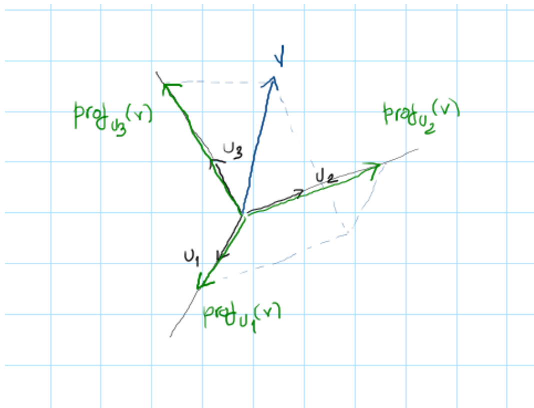
$$[(1,1,3,-2)]_B = \begin{bmatrix} \langle (\frac{1}{\sqrt{2}}, 0, 0, -\frac{1}{\sqrt{2}}), (1,1,3,-2)\rangle \\ \langle (-\frac{1}{\sqrt{6}}, \sqrt{\frac{2}{3}}, 0, -\frac{1}{\sqrt{6}}), (1,1,3,-2)\rangle \\ \langle (0,0,1,0), (1,1,3,-2)\rangle \end{bmatrix} = \begin{bmatrix} \frac{3}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} + \sqrt{\frac{2}{3}} \\ 3 \end{bmatrix}$$

Uma base ortogonal para um subespaço pode ser usada para definir a projeção ortogonal nesse subespaço.

Definição 7.39. *Seja V um espaço vetorial com produto interno e $U \subset V$ um subespaço finitamente gerado. A* projeção ortogonal *de V em U é a transformação linear $P_U \colon V \to V$ definida pela fórmula*

(60) $$P_U(v) = \operatorname{proj}_{u_1}(v) + \ldots + \operatorname{proj}_{u_k}(v)$$

onde $\{u_1, \ldots, u_k\}$ *é uma base ortogonal de* U.



P_U é uma transformação linear pois é uma soma de transformações lineares. Não é no entanto imediatamente óbvio que a fórmula (60) seja independente da escolha da base ortogonal para o subespaço U. Isso é uma consequência do seguinte resultado.

Proposição 7.40. *Seja* V *um espaço com produto interno e* U *um subespaço vetorial finitamente gerado. A transformação linear* $P_U\colon V \to V$ *definida pela expressão* (60) *verifica*

(1) $P_U^2 = P_U$ *(ou seja,* P_U *é uma projeção).*
(2) $P_U(V) = U$ *e* $N(P_U) = U^\perp$.

Portanto $V = U \oplus U^\perp$ *(isto é* $V = U + U^\perp$ *e* $U \cap U^\perp = \{0\}$*) sendo a decomposição única de um vetor de* V *em vetores de* U *e* $U^\perp$ *dada pela expressão*

$$v = \overbrace{P_U(v)}^{\in U} + \overbrace{(v - P_U(v))}^{\in U^\perp}$$

Dem. Exercício. □

A Proposição anterior mostra que P_U é independente da escolha da base ortogonal para U que aparece na fórmula 60 pois uma projeção é completamente determinada pela sua imagem e o seu núcleo (como vimos no Exercício 4.25).

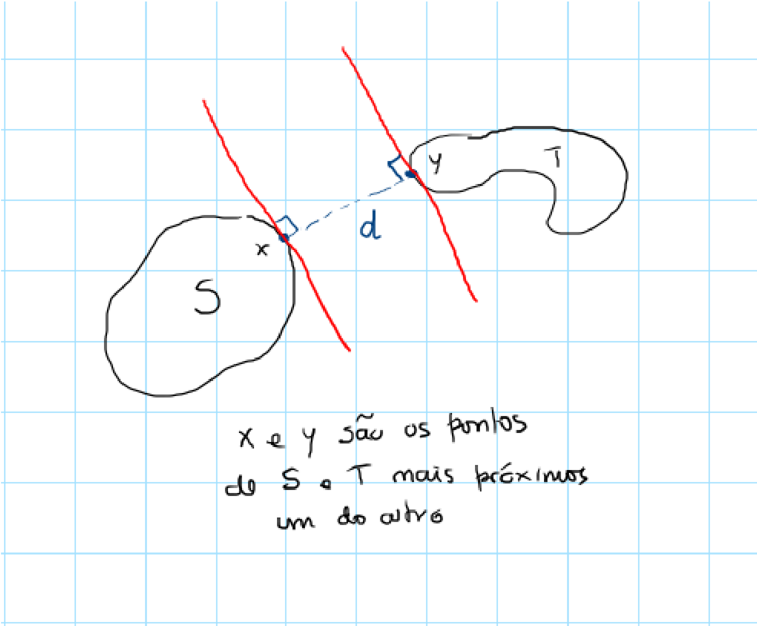
Uma aplicação interessante das projeções ortogonais é o cálculo da distância de um ponto x de um espaço vetorial com produto interno V a um subespaço U de V.

Definição 7.41. *Seja* V *um espaço vetorial com produto interno* $x \in V$ *e* $S \subset V$ *um subconjunto não vazio. A* distância de x a S *é*

$$d(x, S) = \inf\{\|x - u\| \colon u \in S\}$$

Mais geralmente, a distância entre dois subconjuntos não vazios $S, T \subset V$ *define-se por*

$$d(S, T) = \inf\{\|x - y\| \colon x \in S, y \in T\}$$



Note-se que o ínfimo existe porque o conjunto $\{\|x-y\| : x \in S, y \in T\}$ é não vazio (porque S e T são não vazios) e limitado inferiormente (por 0).

Uma característica fundamental da distância entre conjuntos pode resumir-se no seguinte

Slogan: As distâncias medem-se na perpendicular

Veremos agora algumas instâncias do aforismo anterior deixando outras para os exercícios. Verão outras instâncias quando estudarem curvas e superfícies em Cálculo 2.

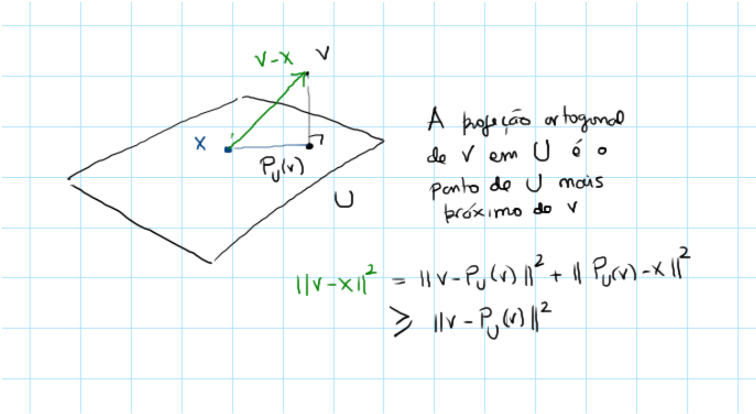
No caso em que $S = U$ é um subespaço vetorial de V e $v \notin U$, dado $x \in U$ podemos escrever o vetor $v - x$ como

$$v - x = (v - P_U(v)) + (P_U(v) - x)$$

uma vez que $v - P_U(v) \in U^\perp$ e $P_U(v) - x \in U$, pelo Teorema de Pitágoras, temos

$$\|v - x\|^2 = \|v - P_U(v)\|^2 + \|P_U(v) - x\|^2 \geq \|v - P_U(v)\|^2 \Leftrightarrow \|v - x\| \geq \|v - P_U(v)\|$$

Uma vez que $P_U(v) \in U$, isso mostra que $d(v, U) = \|v - P_U(v)\|$ e, portanto, que $P_U(v)$ **é o ponto de U mais próximo de v.**



Este argumento pode facilmente ser adaptado para calcular distâncias de pontos a planos $v+U$ que não passam pela origem ou a distância entre planos que não se intersetem.

Exemplo 7.42. *Vamos achar a distância (para o produto interno usual) do ponto* $(1,2,-1)$ *ao plano* $H=\{(x,y,z)\in\mathbb{R}^3\colon x+y+2z=2\}$.

A direção ortogonal ao plano é $(1,1,2)$. *A reta ortogonal ao plano que passa por* $(1,2,-1)$ *tem equação paramétrica*

$$(1,2,-1)+t(1,1,2)=(1+t,2+t,-1+2t)$$

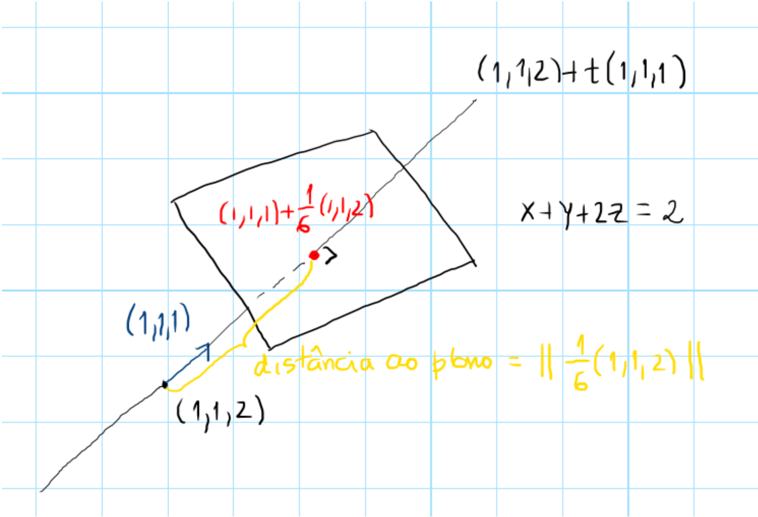
e interseta H *quando*

$$(1+t)+(2+t)+2(-1+2t)=2 \Leftrightarrow 6t=1 \Leftrightarrow t=\tfrac{1}{6}$$

O ponto $v=(\frac{7}{6},\frac{13}{6},-\frac{2}{3})$ *de interseção desta reta com* H *é o ponto de* H *mais próximo de* $(1,2,-1)$. *De facto se* $w\in H$ *for outro ponto, temos como antes, pelo Teorema de Pitágoras, que*

$$\|w-(1,2,-1)\|^2=\|w-v\|^2+\|v-(1,2,-1)\|^2\geq\|v-(1,2,-1)\|^2$$

pois $v-(1,2,-1)$ *(que tem a direção de* $(1,1,2)$*) e* $w-v$ *(que pertence ao plano paralelo a* H *que passa pela origem) são perpendiculares.*



Conclui-se que a distância de $(1,2,-1)$ *a* H *é* $\|\frac{1}{6}(1,1,2)\| = \frac{1}{\sqrt{6}}$.

7.43. **O método dos quadrados mínimos.** [33] Seja A uma matriz $m \times n$. Mesmo que o sistema linear $Ax = b$ seja impossível, podemos tentar encontrar o valor de x que está mais próximo de constituir uma solução no sentido em que a distância de Ax a b é minimizada.

O conjunto $\{Ax\colon x \in \mathbb{R}^n\}$ é um subespaço de $\mathbb{R}^m$, nomeadamente o espaço das colunas de A, $EC(A)$. Como vimos acima, Ax estará o mais próximo possível de um ponto $b \in \mathbb{R}^m$ quando

$$Ax - b \in EC(A)^\perp$$

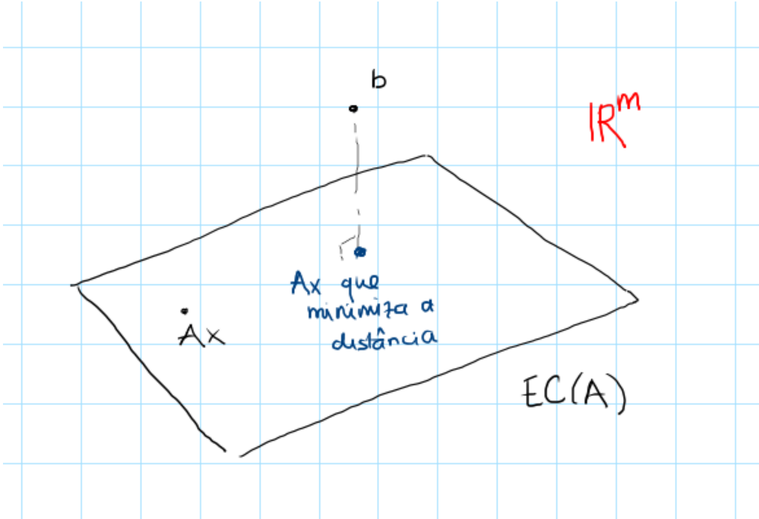
mas, uma vez que $EC(A) = EL(A^T)$, pelo Exemplo 7.31(i) temos

$$EC(A)^\perp = EL(A^T)^\perp = N(A^T)$$

Assim, Ax será o ponto mais próximo de b quando se verifica a **equação dos quadrados mínimos** para x

(61) $$A^T(Ax - b) = 0 \Leftrightarrow A^TAx = A^Tb$$

[33]Esta discussão é adaptada do tratamento deste método em [D].



Note-se que a solução pode não ser única (se $N(A) \neq 0$) mas o sistema (61) tem sempre solução pois traduz *exatamente* a condição de Ax ser o ponto de $EC(A)$ mais próximo de b, e este ponto existe sempre.

Este método é extremamente útil na prática. Frequentemente temos dados experimentais que queremos ajustar a uma lei conhecida, que depende de parâmetros. Os inevitáveis erros experimentais terão como consequência que nenhuma escolha dos parâmetros se adequará exatamente às medições. Este método permite achar quais os valores dos parâmetros que melhor se adequam às medições efetuadas.

Exemplo 7.44. *Vamos determinar a reta $y = ax + b$ que melhor aproxima os três pontos (não colineares) $(0,-2), (1,3), (4,5) \in \mathbb{R}^2$. Se existisse uma reta que passasse pelos três pontos, os coeficientes a, b seriam soluções do sistema*

$$\begin{cases} a \cdot 0 + b = -2 \\ a \cdot 1 + b = 3 \\ a \cdot 4 + b = 5 \end{cases} \Leftrightarrow \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} -2 \\ 3 \\ 5 \end{bmatrix}$$

Este sistema não tem solução mas o método dos quadrados mínimos dá-nos os coeficientes a, b tais que a soma

$$(a \cdot 0 + b - (-2))^2 + (a \cdot 1 + b - 3)^2 + (a \cdot 4 + b - 5)^2$$

é mínima (é isto que dá o nome ao método). Temos que achar a solução do sistema

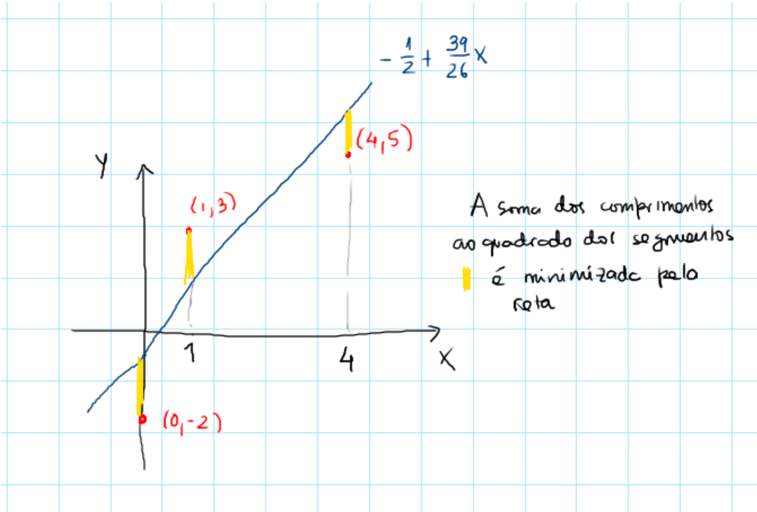
$$\begin{bmatrix} 0 & 1 & 4 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0 & 1 & 4 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} -2 \\ 3 \\ 5 \end{bmatrix} \Leftrightarrow \begin{bmatrix} 17 & 5 \\ 5 & 3 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 23 \\ 6 \end{bmatrix}$$

que é

$$\begin{bmatrix} a \\ b \end{bmatrix} = \frac{1}{26} \begin{bmatrix} 3 & -5 \\ -5 & 17 \end{bmatrix} \begin{bmatrix} 23 \\ 6 \end{bmatrix} = \begin{bmatrix} \frac{39}{26} \\ -\frac{1}{2} \end{bmatrix}$$

pelo que a reta que melhor aproxima os pontos dados (no sentido dos quadrados mínimos) é

$$y = \tfrac{39}{26}x - \tfrac{1}{2}$$



Observação 7.45. *Pouco após a sua descoberta, em 1801, o planeta anão Ceres (na cintura dos asteróides) ficou tapado pelo Sol. Foi para prever (com sucesso) o sítio onde Ceres iria aparecer depois de passar por detrás do Sol, com base nas poucas observações que se tinham conseguido anteriormente, que Gauss inventou o método dos quadrados mínimos.*

7.46. **Uma fórmula para o volume de um paralelipípedo k-dimensional.** Vamos descrever uma fórmula para o volume k-dimensional de um paralelipípedo de dimensão k em $\mathbb{R}^n$ que será útil em Cálculo 2 quando se estudar a integração em superfícies (k-dimensionais) curvas.

Proposição 7.47. *Sejam $v_1, \dots, v_k \in \mathbb{R}^n$ vetores linearmente independentes. Então o volume k-dimensional do paralelipípedo P com arestas $v_1, \dots, v_k$ é*

$$\mathrm{Vol}_k(P) = \sqrt{\det A^T A}$$

onde $A \in M_{n\times k}(\mathbb{R})$ é a matriz que tem $v_1, \dots, v_k$ por colunas.

Dem. Sejam $w_{k+1}, \dots, w_n$ uma base ortonormada para o complemento ortogonal do plano gerado por $v_1, \dots, v_k$. Então (para qualquer noção razoável de volume k-dimensional) o volume do paralelipípedo n-dimensional com arestas $v_1, \dots, v_k, w_{k+1}, \dots, w_n$ é igual ao volume k-dimensional que queremos calcular. Sendo $B \in M_{n\times n}(\mathbb{R})$ a matriz que tem por colunas os vetores $v_1, \dots, v_k, w_{k+1}, \dots, w_n$ (por ordem) e escrevendo B por blocos na forma $[A \mid C]$ com A a matriz formada pelas primeiras k colunas, temos

$$B^T B = \begin{bmatrix} A^T A & 0 \\ 0 & I_{n-k} \end{bmatrix}$$

(onde $C^TC = I_{n-k}$ porque os vetores w_i constituem uma base ortonormada para o plano que geram). Portanto

$$(\det B)^2 = \det(B^TB) = \det(A^TA) \quad \Leftrightarrow \quad \sqrt{\det A^TA} = |\det B|$$

Como $|\det B|$ é o volume do paralelipípedo n-dimensional com arestas $v_1, \ldots, v_k, w_{k+1}, \ldots, w_n$, isto conclui a demonstração. $\square$

Notamos que a matriz A^TA na Proposição 7.47 é exatamente a *matriz da métrica* com respeito à base $(v_1, \ldots, v_k)$ para a restrição do produto interno usual ao plano gerado por $\{v_1, \ldots, v_k\}$.

Exemplo 7.48. *A área do paralelogramo em* $\mathbb{R}^3$ *com arestas* $(1, -2, 1)$ *e* $(2, 3, 0)$ *é*

$$\sqrt{\det \begin{bmatrix} 1 & -2 & 1 \\ 2 & 3 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ -2 & 3 \\ 1 & 0 \end{bmatrix}} = \sqrt{\begin{vmatrix} 6 & -4 \\ -4 & 13 \end{vmatrix}} = \sqrt{62}$$

7.49. **Projeção ortogonal e compressão de dados.** A ideia fundamental utilizada na compressão de dados (por exemplo som, ou imagem) é a projeção ortogonal e baseia-se na descoberta por Joseph Fourier, um engenheiro, matemático e físico do século XIX, durante as suas investigações sobre a propagação do calor, que é possível descrever funções arbitrários através de somas de funções trigonométricas.

Na sua expressão mais simples, suponhamos que pretendemos descrever uma função real contínua $f\colon [0, \pi] \to \mathbb{R}$ (que pode representar por exemplo, uma linha numa imagem, ou a intensidade do som). É fácil verificar, que com respeito ao produto interno $\langle \cdot, \cdot \rangle$ no espaço vetorial $C([0, \pi], \mathbb{R})$ das funções contínuas em $[0, \pi]$ definido por

$$\langle f, g \rangle = \int_0^\pi f(x)g(x)dx$$

o conjunto

$$\{\operatorname{sen} x, \operatorname{sen}(2x), \ldots, \operatorname{sen}(nx), \ldots\}$$

é ortogonal. Fourier descobriu que é possível expressar qualquer função contínua como "combinação linear"destas funções[34] - aquilo a que se chama hoje uma *série de Fourier*. Intuitivamente isto significa que o conjunto acima forma uma "base ortogonal"para o espaço das funções contínuas em $[0, \pi]$.

A ortogonalidade permite determinar facimente os coeficientes da combinação linear correspondente a uma função f: o coeficiente segundo $\operatorname{sen}(nx)$ da função f é dado pela expressão

$$P_{\operatorname{sen}(nx)}(f) = \frac{\langle \operatorname{sen}(nx), f(x) \rangle}{\| \operatorname{sen}(nx)\|^2}$$

[34] Trata-se de uma combinação linear infinita, ou desenvolvimento em série. A análise da convergência destas séries é delicada e constitui ainda hoje uma área vibrante da Matemática que se designa por Análise Harmónica.

A ideia básica da compressão de dados é que, para armazenar a informação contida no gráfico de f basta armazenar um número suficientemente grande destes coeficientes. Quanto maior o número de coeficientes, maior a fidelidade com que conseguimos reproduzir a função f. Dados os coeficientes, reproduzir a função f consiste em somar a expressão obtida a partir dos coeficientes armazenados. Desde que o número de coeficientes utilizado seja suficientemente grande será impossível ao ouvido ou olho humano distinguir entre a função original e a soma finita de funções trigonométricas usada para a aproximar.

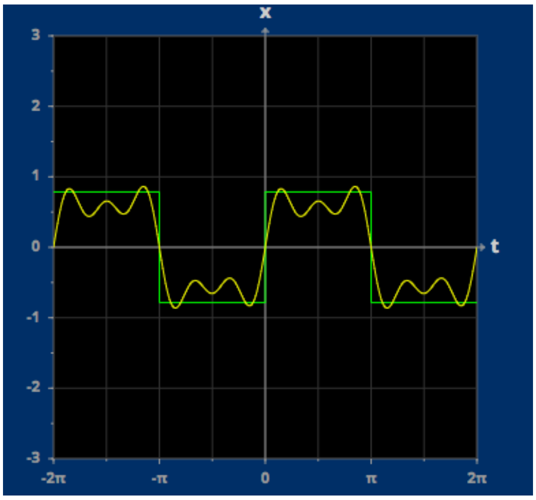


FIGURA 2. Aproximação de um sinal retangular por uma soma de Fourier com 5 termos.

Recomendamos a utilização do applet disponível em `http://mathlets.org/mathlets/fourier-coefficients/` (parte dos MIT Mathlets) para explorar esta ideia, que será descrita com mais detalhe e utilizada no próximo ano, no curso de Cálculo 3.

7.50. **Transformações unitárias e (anti)-hermitianas.** Para terminar, vamos falar um pouco das transformações lineares entre espaços vetoriais munidos de produto interno. Começamos por estudar as transformações que preservam o produto interno e portanto ângulos e distâncias.

Definição 7.51. *Seja V um espaço vetorial com produto interno $\langle\cdot,\cdot\rangle$. Uma transformação linear $T\colon V \to V$ tal que*

$$\langle T(v), T(w)\rangle = \langle v, w\rangle \qquad \text{para todos os } v, w \in V$$

diz-se **ortogonal** *quando V é um espaço vetorial real e* **unitária** *quando V é um espaço vetorial complexo.*

As transformações unitárias são exatamente as que preservam a distância determinada pelo produto interno - isto é as *isometrias* (ver o Exercício 28). O caso mais importante da definição anterior é o dos produtos internos usuais em $\mathbb{R}^n$ e $\mathbb{C}^n$.

Exemplo 7.52. *Consideremos $\mathbb{R}^n$ com o seu produto interno usual e $T\colon \mathbb{R}^n \to \mathbb{R}^n$ a transformação linear definida por $T(x) = Ax$ com A uma matriz $n \times n$ (onde, como habitualmente, estamos a identificar $\mathbb{R}^n$ com as matrizes coluna $n \times 1$). O produto interno de dois vetores x e y de $\mathbb{R}^n$ pode escrever-se matricialmente na forma $x^T y$. Portanto T é ortogonal se e só se*

$$(Ax)^T(Ay) = x^T y \Leftrightarrow x^T A^T A y = x^T y \qquad \text{para todos os } x, y \in \mathbb{R}^n \tag{62}$$

Isto acontece se e só se

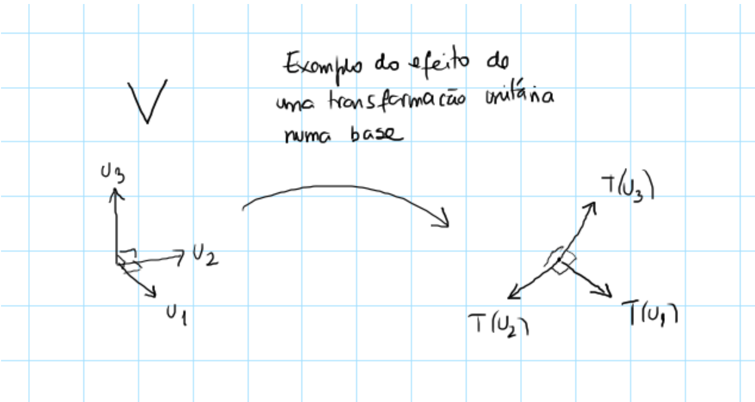
$$A^T A = I_n \tag{63}$$

De facto, é claro que se A satisfaz a condição (63) então satisfaz (62). Reciprocamente se (62) é satisfeita então tomando para x e y o i-ésimo e j-ésimo vetores da base canónica de $\mathbb{R}^n$ respetivamente, a expressão $x^T A^T A y$ calcula a entrada ij da matriz $A^T A$ que é portanto 1 quando $i = j$ e 0 caso contrário. $A^T A$ é portanto a matriz identidade.

As matrizes de $M_{n\times n}(\mathbb{R})$ que satisfazem (63) chamam-se **matrizes ortogonais**. *Note-se que esta equação é também equivalente a dizer que A* **é invertível com inversa** A^T.

Uma vez que as linhas da matriz A^T são as colunas de A, a condição (63) diz que **uma matriz é ortogonal sse as suas colunas formam uma base ortonormada para $\mathbb{R}^n$.**

Assim, quando multiplicamos a matriz A por um vetor $x \in \mathbb{R}^n$, obtemos um vetor que tem as mesmas coordenadas que x na base ortonormada dada pelas colunas de A. Podemos pensar que o referencial ortonormado inicial foi "rodado"para um novo referencial ortornormado. Isto generaliza o conceito de rotação e reflexão no espaço tridimensional como veremos no Exercício 40.



Consideremos agora o caso inteiramente análogo em que $V = \mathbb{C}^n$ com o produto interno usual, e $Tx = Ax$ com $x \in \mathbb{C}^n$. Temos agora que o produto interno é definido matricialmente pela expressão $\langle x, y\rangle = \overline{x}^T y$ e então T é unitária se

$$\overline{x}^T \overline{A}^T A y = \overline{x}^T y \Leftrightarrow \overline{A}^T A = I_n$$

As matrizes que satisfazem esta condição dizem-se **unitárias**. *Novamente uma matriz é unitária sse é invertível e a sua inversa é $\overline{A}^T$, sse as suas colunas formam uma base ortonormada para $\mathbb{C}^n$.*

É conveniente simplificar a notação para a matriz transposta conjugada.

Definição 7.53. *Seja $A \in M_{m\times n}(\mathbb{C})$. A matriz transposta conjugada $\overline{A}^T$ é denotada por A^* e designada por matriz* adjunta *de A* [35]*. Assim, $A^* \in M_{n\times m}(\mathbb{C})$ é a matriz com entrada ij dada por $\overline{a_{ji}}$.*

Proposição 7.54. *Seja V um espaço vetorial complexo com produto interno $\langle\cdot,\cdot\rangle$ e $T\colon V \to V$ uma transformação unitária. Então:*

(i) Os valores próprios de T são complexos com módulo 1.
(ii) Vetores próprios de T correspondentes a valores próprios distintos são ortogonais.

Dem. Seja v um vetor próprio de T. Sendo $T(v) = \lambda v$ temos

$$\|T(v)\|^2 = \langle T(v), T(v)\rangle = \langle \lambda v, \lambda v\rangle = \overline{\lambda}\lambda\langle v, v\rangle = |\lambda|^2\|v\|^2$$

Por outro lado, como T é unitária temos $\langle T(v), T(v)\rangle = \langle v, v\rangle = \|v\|^2$. Portanto $\|v\|^2 = |\lambda|^2\|v\|^2$, e como $v \neq 0$, isto significa que $|\lambda| = 1$.

Suponhamos agora que $T(v) = \lambda v$ e $T(w) = \mu w$ com λ, μ distintos. Então

$$\langle v, w\rangle = \langle Tv, Tw\rangle = \langle \lambda v, \mu w\rangle = \overline{\lambda}\mu\langle v, w\rangle$$

ou seja

$$(1 - \overline{\lambda}\mu)\langle v, w\rangle = 0 \Leftrightarrow \overline{\lambda}\mu = 1 \quad \text{ou} \quad \langle v, w\rangle = 0$$

Como λ é um complexo com módulo 1, $\overline{\lambda} = \frac{1}{\lambda}$ logo a primeira condição na disjunção acima é equivalente a $\mu = \lambda$ pelo que não se verifica. Conclui-se que $\langle v, w\rangle = 0$, isto é, que v e w são ortogonais. □

Observação 7.55. *Se encararmos uma matriz $n \times n$ real A como uma matriz complexa, dizer que A é ortogonal ou unitária é equivalente (uma vez que $\overline{A} = A$). Vemos portanto que os valores próprios de uma matriz ortogonal (encarada como uma matriz complexa) são complexos unitários e que os seus vetores próprios são ortogonais em $\mathbb{C}^n$.*

Exemplo 7.56. *A matriz*

$$\begin{bmatrix} \cos\alpha & -\operatorname{sen}\alpha \\ \operatorname{sen}\alpha & \cos\alpha \end{bmatrix}$$

é ortogonal, como se verifica imediatamente. Geometricamente corresponde à rotação de um ângulo α no sentido anti-horário.

Note-se que os valores próprios (complexos) desta matriz são as soluções de

$$(\cos\alpha - \lambda)^2 + \operatorname{sen}^2\alpha = 0 \Leftrightarrow \lambda = \cos\alpha \pm i\operatorname{sen}\alpha$$

Os vetores próprios (também necessariamente complexos) são as soluções de

$$\begin{bmatrix} \cos\alpha - (\cos\alpha \pm i\operatorname{sen}\alpha) & -\operatorname{sen}\alpha \\ \operatorname{sen}\alpha & \cos\alpha - (\cos\alpha \pm i\operatorname{sen}\alpha) \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = 0 \Leftrightarrow \begin{bmatrix} a \\ b \end{bmatrix} = \gamma \begin{bmatrix} \mp i \\ 1 \end{bmatrix}$$

(com $\gamma \in \mathbb{C}$) e são ortogonais para o produto interno usual em $\mathbb{C}^2$.

[35] Por vezes A^* diz-se também a matriz transconjugada de A

Vamos agora falar de outras duas classes importantes de endomorfismos de espaços vetoriais com produto interno que são "bem comportados" com respeito ao produto interno. A verdadeira justificação para considerarmos estas classes é que endomorfismos deste tipo são ubíquos em Matemática e nas suas aplicações, pelo que é importante conhecer as suas propriedades.

Definição 7.57. *Seja V um espaço vetorial de dimensão finita com produto interno. Uma transformação linear $T\colon V \to V$ diz-se* auto adjunta[36] *se para todos os $v, w \in V$ temos*

$$\langle T(v), w\rangle = \langle v, T(w)\rangle \tag{64}$$

e anti-adjunta[37] *se para todos os $v, w \in V$ temos*

$$\langle T(v), w\rangle = -\langle v, T(w)\rangle \tag{65}$$

No caso em que $V = \mathbb{C}^n$ com o produto interno usual, e $Tv = Av$ é determinada por uma matriz $A \in M_{n\times n}(\mathbb{C})$, as equações (64) e (65) traduzem-se em

$$(\overline{Av})^T w = \pm\overline{v}^T(Aw) \Leftrightarrow \overline{v}^T A^* w = \pm\overline{v}^T Aw \qquad \text{para todos os } v, w \in \mathbb{C}^n$$

Vemos assim que uma matriz $A \in M_{n\times n}(\mathbb{C})$ determina uma transformação auto-adjunta sse $A = A^*$, isto é, se A é uma matriz hermitiana. Uma transformação é anti-adjunta sse $A = -A^*$ e diz-se então que A é uma matriz *anti-hermitiana*. Analogamente uma matriz $A \in M_{n\times n}(\mathbb{R})$ determina uma transformação auto-adjunta de $\mathbb{R}^n$ sse é uma matriz simétrica e anti-adjunta sse é *anti-simétrica*, isto é se $A^T + A = 0$.

Proposição 7.58. *Os valores próprios de uma transformação linear auto-adjunta são reais, e os de uma transformação linear anti-adjunta são imaginários puros. Em qualquer dos casos, vetores próprios de valores próprios distintos são ortogonais.*

Dem. Suponhamos que T é auto-adjunta e v é um vetor próprio de T então

$$\overline{\lambda}\langle v, v\rangle = \langle \lambda v, v\rangle = \langle Tv, v\rangle = \langle v, Tv\rangle = \langle v, \lambda v\rangle = \lambda\langle v, v\rangle$$

Como $v \neq 0$ temos que $\overline{\lambda} = \lambda$ pelo que λ é real. No caso anti-adjunto obteríamos a igualdade $\overline{\lambda} + \lambda = 0$ que diz que λ é imaginário puro.

Sejam λ e μ valores próprios distintos de T auto-adjunta com vetores próprios v, w. Então

$$\lambda\langle v, w\rangle = \langle \lambda v, w\rangle = \langle Tv, w\rangle = \langle v, Tw\rangle = \mu\langle v, w\rangle$$

onde na primeira igualdade usámos o facto de λ ser real e portanto igual ao seu conjugado. A igualdade anterior traduz-se em $(\lambda - \mu)\langle v, w\rangle = 0$. Uma vez que $\lambda \neq \mu$, conclui-se que v e w são ortogonais.

No caso anti-adjunto obtemos análogamente $(\overline{\lambda} + \mu)\langle v, w\rangle = 0$. Como λ e μ são imaginários puros $\overline{\lambda} + \mu = -\lambda + \mu$ pelo que novamente vemos que v e w são ortogonais. $\square$

[36] Em inglês *self-adjoint*. Ver o Exercício 34 para a justificação desta terminologia.

[37] Em inglês *skew-adjoint*.

Teorema 7.59 (Teorema espetral). [38]

(i) Seja V um espaço vetorial complexo de dimensão finita com produto interno e $T\colon V \to V$ uma transformação linear unitária, auto-adjunta ou anti-adjunta. Então T é diagonalizável por uma base ortogonal de V.

(ii) Seja V um espaço vetorial real de dimensão finita com produto interno e $T\colon V \to V$ uma transformação linear auto-adjunta. Então T é diagonalizável por uma base ortogonal de V.

Dem. As demonstrações são todas análogas pelo que vamos apenas ilustrar o caso de uma transformação auto-adjunta deixando os restantes como exercício.

A demonstração é por indução na dimensão do espaço vetorial complexo V, sendo que o caso de dimensão 1 é trivial. Supondo que a dimensão de V é maior do que 1, seja v um vetor próprio de T e consideremos o subespaço $W = v^\perp \subset V$. Então $T_{|W}$ é também auto-adjunta como endomorfismo de W para a restrição do produto interno a W. De facto, a igualdade

$$\overline{\lambda}\langle v, w\rangle = \langle \lambda v, w\rangle = \langle Tv, w\rangle = \langle v, Tw\rangle$$

mostra que, se $w \in v^\perp$ então $Tw \in v^\perp$. Dado que W é invariante para T, é evidente que $T_{|W}$ é auto-adjunta. Por hipótese de indução, existe uma base ortogonal de W formada por valores próprios de $T_{|W}$ que juntamente com v forma a base ortogonal desejada para V.

Seja agora V um espaço vetorial real e $T\colon V \to V$ uma transformação linear auto-adjunta. Para mostrarmos que T é diagonalizável por uma base ortogonal seguindo o argumento do parágrafo anterior basta-nos mostrar que T tem um vetor próprio $v \in V$ (pois então $v^\perp$ será invariante e a restrição de T a $v^\perp$ auto-adjunta).

Seja B uma base ortonormada para V, e $A = A_{T,B,B}$ a matriz que representa a transformação linear T com respeito à base B. A matriz A é simétrica e portanto a transformação linear $T'\colon \mathbb{C}^n \to \mathbb{C}^n$ definida por $T'(x) = Ax$ é auto-adjunta com respeito ao produto interno usual em $\mathbb{C}^n$. Seja $\lambda \in \mathbb{R}$ um valor próprio de T'. Como A é uma matriz real podemos escolher um vetor próprio y para λ com coordenadas reais. Seja $v \in V$ tal que $[v]_B = y$. Então

$$[T(v)]_B = A_{T,B,B}[v]_B = Ay = \lambda y = \lambda[v]_B = [\lambda v]_B \Rightarrow T(v) = \lambda v$$

o que conclui a demonstração. □

Sumarizamos agora a informaçao sobre matrizes quadradas que resulta do Teorema anterior, aplicando-o à transformação linear definida por $Tx = Ax$ com A uma matriz $n \times n$ real ou complexa. Em cada caso a matriz A pode ser escrita na forma

$$A = SDS^{-1}$$

com S uma matriz, unitária quando A é diagonalizável sobre $\mathbb{C}$, e ortogonal quando A é diagonalizável sobre $\mathbb{R}$, (cujas colunas formam uma base ortonormada para $\mathbb{C}^n$ ou $\mathbb{R}^n$

[38] O *espetro* de um endomorfismo é o conjunto dos seus valores próprios. O nome vem do espetro das cores - as cores são frequências de ondas eletromagnéticas que por sua vez são os valores próprios de um certo endomorfismo - o operador das ondas.

consoante o caso, constituída por vetores próprios de A), e D uma matriz diagonal cujas entradas são os valores próprios de A.

Tipo de matriz	Definição	Diagonalizável	Valores próprios
ortogonal	$AA^T = I_n$	sobre $\mathbb{C}$	$\lambda \in \mathbb{C}, \lvert\lambda\rvert = 1$
simétrica	$A = A^T$	sobre $\mathbb{R}$	reais
anti-simétrica	$A + A^T = 0$	sobre $\mathbb{C}$	imaginários puros

Matrizes $n \times n$ reais especiais.

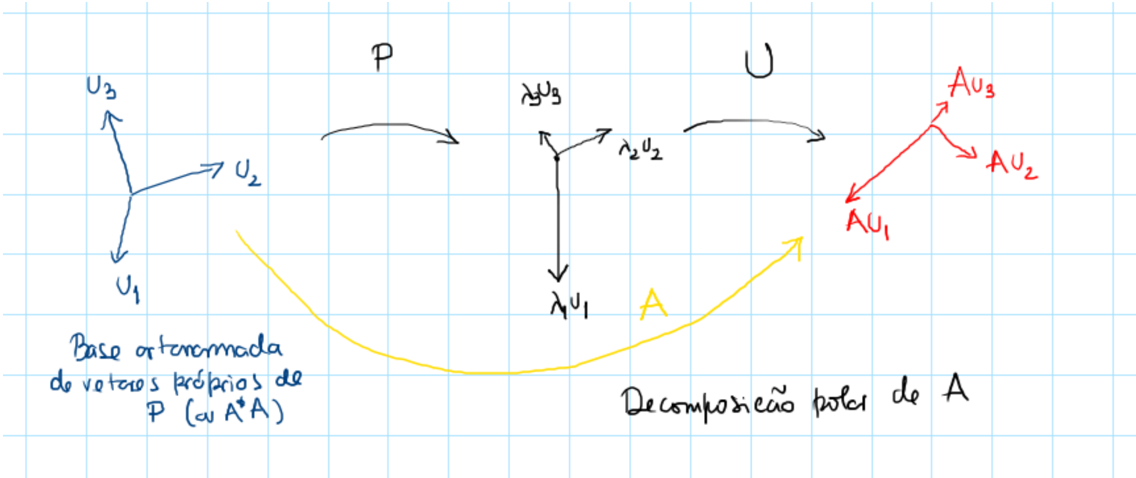
Tipo de matriz	Definição	Valores próprios
unitária	$AA^* = I_n$	$\lambda \in \mathbb{C}, \lvert\lambda\rvert = 1$
hermitiana	$A = A^*$	reais
anti-hermitiana	$A + A^* = 0$	imaginários puros

Matrizes $n \times n$ complexas especiais.

Observação 7.60. *Embora não haja nenhum critério útil para ver se uma matriz é diagonalizável, há um critério muito simples para ver se uma matriz complexa A é diagonalizável por uma base* ortonormada. *Isto acontece sse $AA^* = A^*A$, caso em que se diz que a matriz A é* normal. *Ver o Exercício 42 para uma demonstração.*

7.61. **A decomposição em valores singulares.** Vamos agora usar o Teorema espetral para obter uma decomposição muito útil para uma transformação linear de $\mathbb{K}^n$ em $\mathbb{K}^m$ para $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Deixamos como exercício a adaptação desta discussão ao caso em que os espaços $\mathbb{K}^n$ são substituído por espaços de dimensão finita com produto interno.

Começamos por considerar o caso de uma matriz quadrada. O seguinte resultado generaliza a decomposição polar de um número complexo não nulo. Geometricamente, exprime um endomorfismo invertível de $\mathbb{K}^n$ como a composição de um conjunto de expansões e contrações (reais) nas direções de uma base ortonormada (pelo Teorema Espetral é esse o efeito de uma matriz hermitiana com todos os valores próprios positivos) seguido de uma "rotação"(uma transformação unitária).



Proposição 7.62 (Decomposição polar)**.** *Seja $A \in M_{n\times n}(\mathbb{C})$ uma matriz invertível. Então existem uma matriz hermitiana P com valores próprios todos positivos e uma matriz unitária U únicas tais que $A = UP$.*

Dem. A matriz A^*A é hermitiana e todos os seus valores próprios são positivos uma vez que $A^*Av = \lambda v \Rightarrow \|Av\|^2 = v^*A^*Av = \lambda\|v\|^2$ (e o núcleo de A é trivial logo $\lambda \neq 0$). Pelo Teorema espetral existe uma matriz unitária S e uma matriz diagonal D com entradas diagonais positivas tais que $A^*A = SDS^*$. Seja $P = S\sqrt{D}S^*$, onde $\sqrt{D}$ denota a matriz que se obtém de D tomando a raíz quadrada positiva de cada uma das entradas diagonais de D.

É imediato verificar que P é hermitiana com todos os valores próprios positivos (os seus valores próprios são as raízes quadradas dos valores próprios de D). Além disso

$$P^*P = (S\sqrt{D}S^*)^*(S\sqrt{D}S^*) = S\sqrt{D}^*S^*S\sqrt{D}S^* = S\sqrt{D}\sqrt{D}S^* = SDS^* = A^*A$$

Para concluir a existência da decomposição polar, verifiquemos que $U = AP^{-1}$ é unitária:

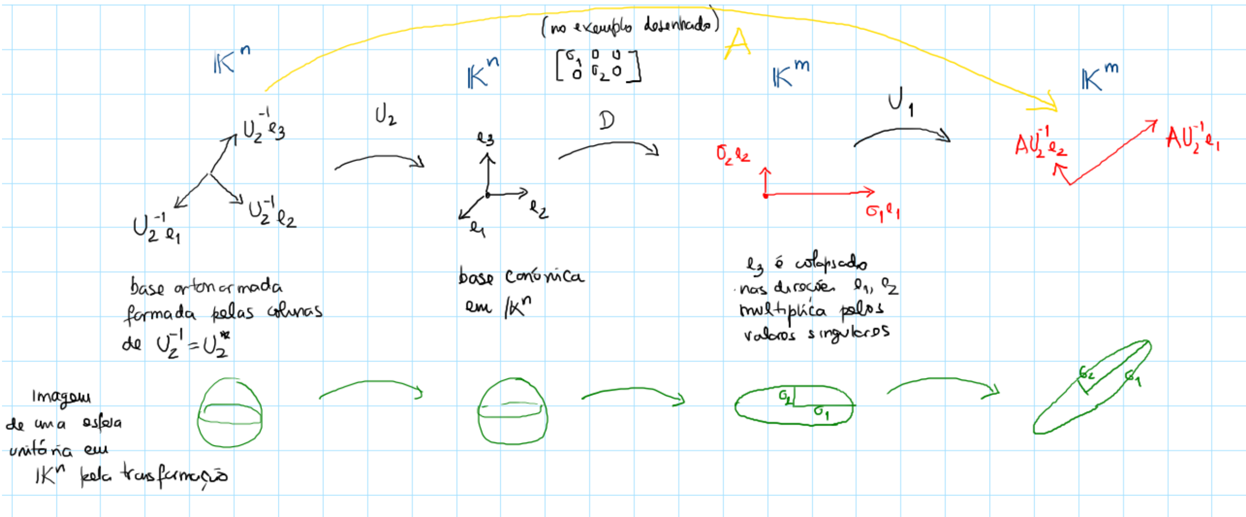
$$U^*U = (P^{-1})^*A^*AP^{-1} = (P^{-1})^*P^*PP^{-1} = I$$

Resta-nos demonstrar a unicidade das matrizes U e P. Para tal basta notar que se $A = UP$ é uma decomposição polar então $A^*A = P^*P = P^2$. Se B for uma base ortonormada que diagonaliza P então a mesma base ortonormada diagonaliza A^*A. Conclui-se que o espaço próprio de λ para P coincide com o espaço próprio de λ^2 para A^*A. Logo P é completamente determinada por A, e portanto o mesmo sucede com U. □

Observação 7.63. *No Teorema anterior, se $A \in M_{n\times n}(\mathbb{R})$ podemos escolher P na demonstração acima real (pelo Teorema espetral para matrizes simétricas) e portanto U será também real. Conclui-se assim que toda a matriz real invertível A se fatoriza de forma única como $A = OP$ com O uma matriz ortogonal e P uma matriz simétrica com todos os valores próprios positivos.*

Exemplo 7.64. *Se $A = [a + bi] \in M_{1\times 1}(\mathbb{C})$ então $A^*A = [a^2 + b^2]$, $P = [\sqrt{a^2 + b^2}]$ e escrevendo $a + bi = r(\cos\theta + i \operatorname{sen}\theta)$ temos que $U = P^{-1}A = [\cos\theta + i \operatorname{sen}\theta]$. Vemos assim que a decomposição polar generaliza a forma trigonométrica dos números complexos.*

O seguinte Teorema pode ser visto como uma generalização da decomposição polar a transformações lineares arbitrárias. Afirma que qualquer transformação linear de $\mathbb{K}^n$ para $\mathbb{K}^m$ pode ser decomposta numa rotação inicial em $\mathbb{K}^n$, seguida de uma transformação que expande ou contrai ao longo de alguns eixos coordenados, colapsando os restantes eixos coordenados, seguida de outra rotação em $\mathbb{K}^m$.



Teorema 7.65 (Decomposição em valores singulares)**.** *Seja $A \in M_{m\times n}(\mathbb{C})$. Existe uma fatorização de A na forma*

$$A = U_1 D U_2$$

com $U_1 \in M_{m\times m}(\mathbb{C})$ e $U_2 \in M_{n\times n}(\mathbb{C})$ unitárias e $D \in M_{m\times n}(\mathbb{R})$ uma matriz cujas únicas entradas não nulas são reais positivos $\sigma_1 \geq \ldots \geq \sigma_k > 0$ ao longo da diagonal, isto é

$$d_{ij} = \begin{cases} \sigma_i & \text{se } j = i \text{ e } i \leq k, \\ 0 & \text{caso contrário} \end{cases}$$

Os números σ_i chamam-se os valores singulares *de A.*

Demonstração. A matriz $A^*A \in M_{n\times n}(\mathbb{C})$ é hermitiana e tem todos os valores próprios ≥ 0 (uma vez que $\langle A^*Av, v\rangle = \|Av\|^2 \geq 0$). Sejam U_2 uma matriz unitária e Λ uma matriz diagonal tais que

$$A^*A = U_2^{-1}\Lambda U_2 \tag{66}$$

com as entradas diagonais não nulas de Λ aparecendo na ordem $\sigma_1^2 \geq \ldots \geq \sigma_k^2 > 0$ (note-se que podemos alterar a ordem das entradas na diagonal com uma matriz de permutação, que é unitária). Seja $D \in M_{m\times n}(\mathbb{R})$ uma matriz cujas únicas entradas não nulas são $\sigma_1 \geq \ldots \geq \sigma_k > 0$ ao longo da diagonal (isto é $d_{ii} = \sigma_i$ para $1 \leq i \leq k$ e todas as outras entradas são nulas). Temos então $D^TD = \Lambda$.

Recorde que $U_i^{-1} = U_i^*$ uma vez que as matrizes U_i são unitárias. A equação (66) diz-nos que $(AU_2^{-1})^*(AU_2^{-1}) = \Lambda$ pelo que

- As últimas $n-k$ colunas de AU_2^{-1} são nulas.
- As primeiras k colunas $v_1,\ldots,v_k \in \mathbb{C}^m$ são um conjunto ortogonal em $\mathbb{C}^m$ sendo o comprimento da coluna i igual a σ_i

As colunas não nulas de AU_2^{-1} são uma base ortogonal para o espaço das colunas de A (uma vez que geram $EC(A)$ e são ortogonais).

Seja U_1 uma matriz unitária cujas primeiras k colunas são $\frac{v_i}{\|v_i\|}$ (tal matriz obtém-se completando a base ortonormada $\{\frac{v_1}{\|v_1\|},\ldots,\frac{v_k}{\|v_k\|}\}$ para $EC(A)$ a uma base ortonormada de $\mathbb{C}^n$.

Então $A = U_1DU_2$ conforme desejado: esta equação é equivalente a $AU_2^{-1} = U_1D$ que se verifica por definição da matriz U_1! □

Observação 7.66. *(i) Se A for uma matriz real, podemos tomar para U_1 e U_2 matrizes ortogonais obtendo assim a decomposição em valores singulares real.*

(ii) Embora a decomposição em valores singulares não seja única, os valores singulares *de A são completamente determinados por A: são as raízes quadradas positivas dos valores próprios não nulos de A^*A.*

(iii) Em termos muito concretos:
- *As colunas de U_2^{-1} são uma base ortonormada $B_1 = \{w_1,\ldots,w_n\}$ para $\mathbb{C}^n$ (cujos elementos se chamam* vetores singulares à direita*)*
- *As colunas de U_1 são uma base ortonormada $B_2 = \{x_1,\ldots,x_m\}$ para $\mathbb{C}^m$ (cujos elementos se chamam* vetores singulares à esquerda*)*
- *O vetor w_i é enviado por A em $\sigma_i x_i$ para $1 \le i \le k$ e para 0 se $i > k$.*

Exemplo 7.67. *Consideremos a matriz $A = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & -1 \end{bmatrix}$. Temos*

$$\begin{aligned}
A^TA &= \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & -1 \\ 0 & -1 & 2 \end{bmatrix} = U_2^{-1}\Lambda U_2 \\
&= \begin{bmatrix} -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \end{bmatrix}\begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \end{bmatrix}^T
\end{aligned}$$

Portanto

$$U_2 = \begin{bmatrix} -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \end{bmatrix}^T, \qquad D = \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & \sqrt{2} & 0 \end{bmatrix}$$

Os valores singulares são $\sigma_1 = \sqrt{3}$, $\sigma_2 = \sqrt{2}$. Neste caso, uma vez que o espaço das colunas é todo o $\mathbb{R}^2$, não temos escolha para a matriz U_1:

$$AU_2^{-1} = \begin{bmatrix} 0 & \sqrt{2} & 0 \\ -\sqrt{3} & 0 & 0 \end{bmatrix} \quad \Rightarrow \quad U_1 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

A decomposição singular de A é portanto

$$A = U_1 D U_2 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & \sqrt{2} & 0 \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{bmatrix}$$

O efeito de A nos vetores de $\mathbb{R}^3$ pode ser descrito da seguinte forma: colapsa $\mathbb{R}^3$ ortogonalmente no plano gerado pelas duas primeiras colunas de U_2^{-1}, enviando depois um círculo unitário neste plano numa elipse com eixos $\sqrt{3}$ e $\sqrt{2}$ segundo as direções dadas pelas colunas de U_1 (neste caso o eixo dos yy, que é a imagem da reta gerada pela primeira coluna de U_2^{-1} em $\mathbb{R}^3$ e e o eixo dos xx que é a imagem da segunda coluna de U_2^{-1})

A decomposição em valores singulares desempenha um papel importante na análise de dados. Os dados podem normalmente ser organizados numa matriz (gigante) e a decomposição em valores singulares ajuda a organizá-los. A decomposição em valores singulares também pode ser usada para comprimir dados, descartando os valores singulares de tamanho inferior a um dado limite fixado.

7.68. **Formas quadráticas.** Como outra aplicação do Teorema espetral vamos aproveitar para classificar a menos de mudança de variável linear os polinómios homogéneos de grau 2 de várias variáveis. Podemos pensar nestes como as funções de várias variáveis mais simples a seguir às funções lineares.

Definição 7.69. *Uma* forma quadrática *em $\mathbb{R}^n$ é uma função $f\colon \mathbb{R}^n \to \mathbb{R}$ da forma*

$$f(x) = x^T A x \tag{67}$$

com $A \in M_{n\times n}(\mathbb{R})$ (estamos como habitualmente a identificar uma matriz 1×1 com um escalar).

Por exemplo

$$f(x,y) = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 2 & 5 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 2x^2 + 6xy + 4y^2 \tag{68}$$

é uma forma quadrática em $\mathbb{R}^2$. Note-se que a forma quadrática depende apenas da *parte simétrica* $\frac{A+A^T}{2}$ da matriz A. De facto uma vez que a transposição de matrizes 1×1 não tem qualquer efeito temos $x^T A x = (x^T A x)^T = x^T A^T x$. Substituindo a matriz A em (67) por $\frac{A+A^T}{2}$ obtemos portanto a mesma expressão. Por outro lado, uma vez que a soma das entradas ij e ji da matriz A é o coeficiente de $x_i x_j$ na expressão (67) matrizes simétricas distintas dão azo a formas quadráticas distintas. Há assim uma *correspondência biunívoca* entre formas quadráticas e matrizes quadradas reais simétricas.

Tendo em conta o Teorema espetral, dada uma matriz simétrica A, existe uma matriz ortogonal S e uma matriz diagonal (real) D tal que

$$A = SDS^{-1}$$

E dado que S é ortogonal, $S^{-1} = S^T$. Usando coordenadas y na base ortonormada formada pelas colunas de S a expressão para a forma quadrática simplifica-se muito. Temos $x = Sy$ e então

$$f(x) = x^T Ax = (y^T S^T)A(Sy) = (y^T S^T)SDS^T(Sy) = y^T Dy = \lambda_1 y_1^2 + \ldots + \lambda_n y_n^2 \tag{69}$$

onde $\lambda_1, \ldots, \lambda_n$ são as entradas diagonais de D, ou seja, os valores próprios de A. Nas aplicações (por exemplo para a determinação de extremos de funções de várias variáveis como verão em Cálculo 2) é importante determinar o "sinal"de uma forma quadrática no seguinte sentido.

Definição 7.70. *Uma forma quadrática $f\colon \mathbb{R}^n \to \mathbb{R}$ diz-se*

(i) definida positiva *se $f(x) > 0$ para $x \neq 0$.*
(ii) semi-definida positiva *se $f(x) \geq 0$ para todo o $x \in \mathbb{R}^n$.*
(iii) definida negativa *se $f(x) < 0$ para $x \neq 0$.*
(iv) semi-definida negativa *se $f(x) \leq 0$ para todo o $x \in \mathbb{R}^n$.*
(v) indefinida *se $f(x)$ assume valores positivos e negativos.*

Da expressão (69) obtemos imediatamente o seguinte resultado.

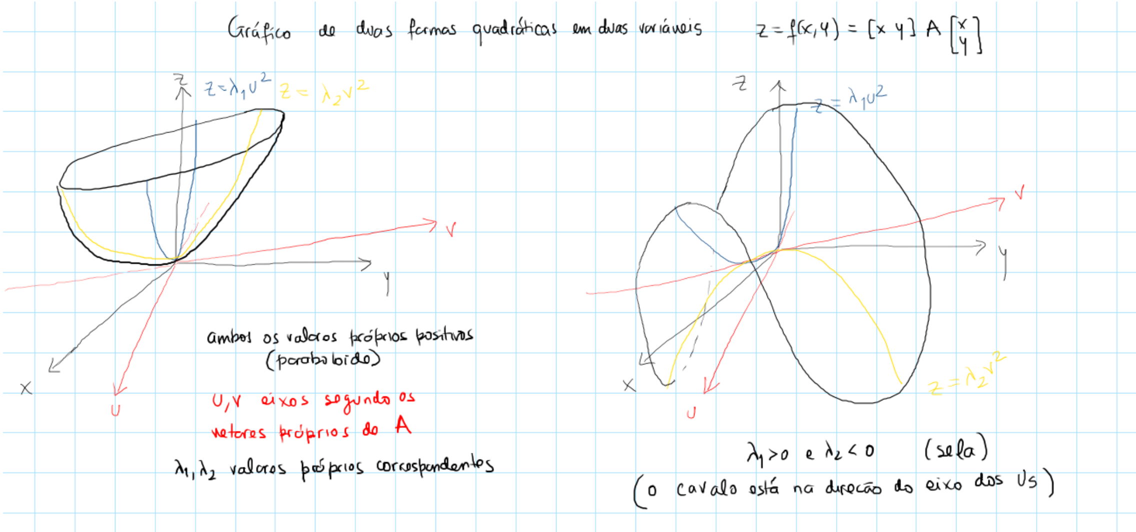
Proposição 7.71. *Uma forma quadrática $f(x) = x^T Ax$ com $A \in M_{n\times n}(\mathbb{R})$ simétrica é*

(i) definida positiva sse todos os valores próprios de A são positivos.
(ii) semidefinida positiva sse todos os valores próprios de A são maiores ou iguais a zero.
(iii) definida negativa sse todos os valores próprios de A são negativos.
(iv) semidefinida negativa sse todos os valores próprios de A são menores ou iguais a zero.
(v) indefinida sse A tem valores próprios de sinal contrário.

Exemplo 7.72. *A forma quadrática* (68) *é indefinida uma vez que a matriz simétrica que a representa*

$$\begin{bmatrix} 2 & 3 \\ 3 & 4 \end{bmatrix}$$

tem determinante negativo e portanto valores próprios de sinais contrários.



Observação 7.73. *A expressão* (69) *mostra também que toda a matriz simétrica com valores próprios positivos é a matriz da métrica de um produto interno, pois a positividade do produto interno corresponde precisamente ao facto da forma quadrática determinada pela matriz da métrica ser definida positiva.*

Seja A uma matriz simétrica $n \times n$. Dado $1 \leq i \leq n$ escrevemos A_i para a matriz que se obtém de A tomando apenas as primeiras i linhas e colunas de A. Os determinantes destas submatrizes de A chamam-se os *menores principais* de A.

Proposição 7.74 (Critério de Sylvester)**.** *Seja $f\colon \mathbb{R}^n \to \mathbb{R}$ a forma quadrática determinada pela matriz simétrica $A \in M_{n\times n}(\mathbb{R})$. Então*

- *f é definida positiva sse $\det A_i > 0$ para $i = 1, \ldots, n$.*
- *f é definida negativa sse $\det A_i$ é positivo para i par e negativo para i ímpar.*

Dem. Note-se que $f(x) = x^T A x$ é definida positiva sse $-f(x) = x^T(-A)x$ é definida negativa. Uma vez que $\det(-A_i) = (-1)^i \det A_i$ (em geral, a multilinearidade do determinante implica que $\det(\lambda A)_i = \lambda^i \det A_i$), vemos que as duas afirmações do enunciado são equivalentes. Basta portanto demonstrar a primeira.

Se f é definida positiva, a sua restrição a $\mathbb{R}^i = \{(x_1, \ldots, x_i, 0, \ldots, 0)\colon x_1, \ldots, x_i \in \mathbb{R}\}$ será também definida positiva. Mas claramente esta restrição é dada pela fórmula (com $x \in \mathbb{R}^i$)

$$f_{|\mathbb{R}^i}(x) = x^T A_i x$$

logo, para que f seja definida positiva, é necessário que $\det A_i > 0$.

Reciprocamente, suponhamos que $\det A_i > 0$ para cada $i = 1, \ldots n$. Seja $i > 1$ e suponhamos indutivamente que já verificámos que $f_{|\mathbb{R}^k}$ é definida positiva para todo $k < i$ (para $k = 1$ é claro que se $\det A_1 = a_{11} > 0$ então $f_{|\mathbb{R}^1}(x_1) = a_{11}x_1^2$ é definida positiva).

Suponhamos por absurdo que $f_{|\mathbb{R}^i}$ não era definida positiva. Uma vez que, por hipótese, $\det A_i > 0$, a matriz A_i teria que ter pelo menos dois valores próprios negativos (contados com multiplicidade). Sendo $W \subset \mathbb{R}^i$ um plano gerado por dois vetores próprios independentes de A_i com valores próprios negativos, teríamos $f_{|W}(y) < 0$ para $y \in W \setminus \{0\}$.

Mas a intersecção de W com $\mathbb{R}^{i-1} \subset \mathbb{R}^i$ tem dimensão pelo menos 1 pelo que existiria um vetor $y \in \mathbb{R}^{i-1} \setminus \{0\}$ com $f(y) < 0$, contradizendo a hipótese de indução que $f_{|\mathbb{R}^{i-1}}$ é definida positiva. $\square$

Exemplo 7.75. *Consideremos a forma quadrática $f\colon \mathbb{R}^3 \to \mathbb{R}$ definida por*

$$f(x,y,z) = 10x^2 + 10y^2 + 10z^2 + 2xy + 2yz$$

A matriz simétrica que lhe está associada é

$$A = \begin{bmatrix} 10 & 1 & 0 \\ 1 & 10 & 1 \\ 0 & 1 & 10 \end{bmatrix}$$

Os menores principais

$$|10| = 10, \qquad \begin{vmatrix} 10 & 1 \\ 1 & 10 \end{vmatrix} = 99, \qquad \begin{vmatrix} 10 & 1 & 0 \\ 1 & 10 & 1 \\ 0 & 1 & 10 \end{vmatrix} = 1000 - 20 = 980$$

são todos positivos, pelo que a forma quadrática é definida positiva.

7.76. **A classificação das quádricas.** Uma *quádrica* é uma curva em $\mathbb{R}^2$ ou uma superfície em $\mathbb{R}^3$ definida por uma equação quadrática. Podemos usar a diagonalização de matrizes simétricas para entender geometricamente estas curvas e superfícies (que irão ser exemplos básicos em Cálculo 2).

Quádricas em $\mathbb{R}^2$. A expressão geral de uma quádrica é

$$ax^2 + bxy + cy^2 + dx + ey + f = 0 \tag{70}$$

em que $a, b, c, d, e, f \in \mathbb{R}$. Devemos excluir alguns casos degenerados: se $a = b = c = 0$ então o conjunto descrito pela expressão anterior é uma reta se $(d,e) \neq (0,0)$, vazio se $d = e = 0$ e $f \neq 0$, e todo o plano se $d = e = f = 0$. Consideremos portanto o caso em que os termos de grau 2 não se anulam todos. Temos

$$ax^2 + bxy + cy^2 = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Sejam $\lambda_1, \lambda_2 \in \mathbb{R}$ os valores próprios da matriz associada à forma quadrática anterior e $(u_1, u_2), (v_1, v_2)$ os vetores próprios correspondentes, que podemos assumir formarem uma base ortonormada para $\mathbb{R}^2$. Sendo (u, v) as coordenadas no referencial determinado pelos vetores próprios temos

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} u_1 & v_1 \\ u_2 & v_2 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} \tag{71}$$

Nestas coordenadas, temos

$$\begin{aligned} ax^2+bxy+cy^2 &= \begin{bmatrix} u & v \end{bmatrix} \begin{bmatrix} u_1 & u_2 \\ v_1 & v_2 \end{bmatrix} \begin{bmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{bmatrix} \begin{bmatrix} u_1 & v_1 \\ u_2 & v_2 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} \\ &= \begin{bmatrix} u & v \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \lambda_1 u^2 + \lambda_2 v^2 \end{aligned}$$

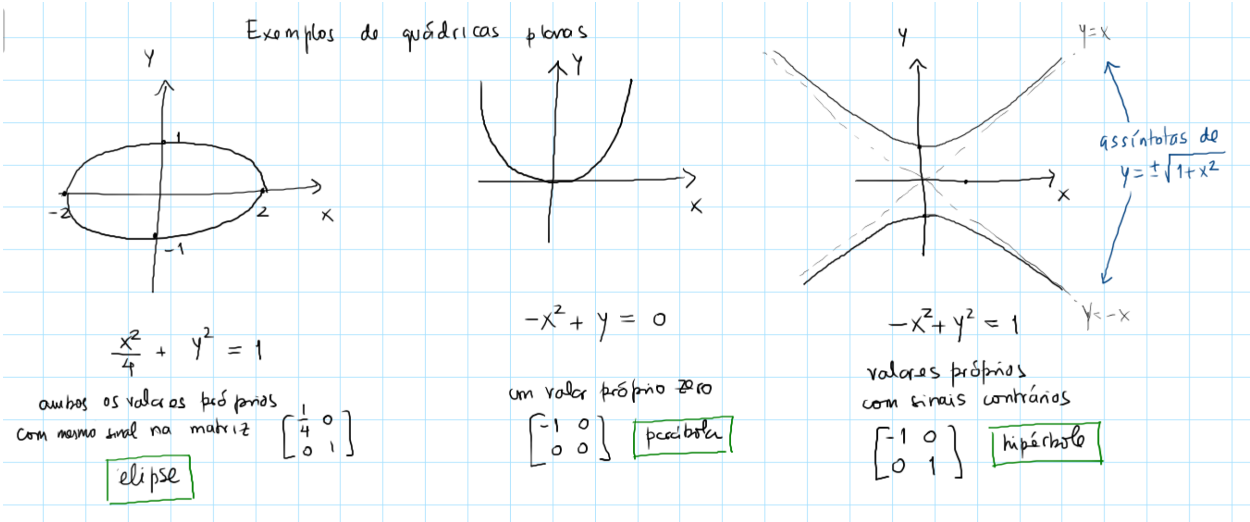
O termo linear em (70) transforma-se mediante a mudança de coordenadas (71) num termo linear em u e v, pelo que esta mudança de coordenadas transforma (70) na seguinte equação:

$$\lambda_1 u^2 + \lambda_2 v^2 + d'u + e'v + f = 0 \tag{72}$$

Temos agora a considerar três casos:

- λ_1, λ_2 ambos diferentes de 0, com o mesmo sinal: Multiplicando (72) por -1 se necessário podemos assumir que λ_1 e λ_2 são positivos. Completando os quadrados podemos escrever a expressão na forma
$$\lambda_1(u+\tfrac{d'}{2\lambda_1})^2 + \lambda_2(v+\tfrac{e'}{2\lambda_2})^2 = f'$$
onde $f' = -f + \frac{d'^2}{4\lambda_1^2} + \frac{e'^2}{4\lambda_2^2}$. Se $f' < 0$ este conjunto é vazio, se $f' = 0$ este conjunto consiste no ponto $(-\frac{d'}{2\lambda_1}, -\frac{e'}{2\lambda_2})$, e se $f' > 0$, o conjunto é uma elipse com centro em $(-\frac{d'}{2\lambda_1}, -\frac{e'}{2\lambda_2})$ (ou uma circunferência quando $\lambda_1 = \lambda_2$).
- λ_1, λ_2 ambos diferentes de 0, com sinais opostos: Multiplicando (72) por -1 se necessário podemos assumir que λ_1 é positivo. Com uma manipulação semelhante à do caso anterior obtemos uma expressão da forma
$$\lambda_1(u-u_0)^2 + \lambda_2(v-v_0)^2 = f'$$
que, para $f' \neq 0$ é a equação de uma hipérbole[39] com "centro" em (u_0, v_0) e assíntotas dadas pelas retas
$$v - v_0 = \pm\frac{\lambda_1}{\lambda_2}(u-u_0)$$
Quando $f' = 0$ a equação define a união das duas retas dadas pela expressão anterior.
- λ_1 ou λ_2 são 0: Sem perda de generalidade podemos assumir que $\lambda_2 = 0$ e que $\lambda_1 > 0$. Manipulando a expressão (72) como antes obtemos
$$\lambda_1(u-u_0)^2 + e'v + f' = 0$$
Se $e' \neq 0$, trata-se da equação de uma parábola, cujo sentido é determinado pelo sinal de e'. Se $e' = 0$ obtemos o conjunto vazio, a reta $u = u_0$, ou duas retas paralelas a esta última, consoante $f' > 0$, $f' = 0$ ou $f' < 0$ respetivamente.

[39] Note-se que a equação $x^2 - y^2 = 1$ se pode escrever na forma $(x-y)(x+y) = 1$ e portanto, mediante a mudança de variável linear $u = x-y, v = x+y$ é equivalente à equação mais familiar para uma hipérbole $uv = 1$.



Exemplo 7.77. *Consideremos o exemplo concreto da equação*

$$x^2 + 2xy + y^2 + x + 2y + 3 = 0$$

A matriz simétrica associada à forma quadrática determinada pelos termos quadráticas é

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

que tem valores próprios 0 *e* 2 *com vetores próprios* $(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})$ *e* $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$*. Fazendo a mudança de coordenadas*

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} \Leftrightarrow \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

obtemos a equação

$$0 \cdot u^2 + 2v^2 + \tfrac{1}{\sqrt{2}}u + \tfrac{1}{\sqrt{2}}v + 2(-\tfrac{1}{\sqrt{2}}u + \tfrac{1}{\sqrt{2}}v) + 3 = 0 \Leftrightarrow 2v^2 - \tfrac{1}{\sqrt{2}}u + \tfrac{3}{\sqrt{2}}v + 3 = 0$$

que se pode escrever na forma

$$u = 2\sqrt{2}(v + \tfrac{3}{4\sqrt{2}})^2 + 3\sqrt{2} - \tfrac{9}{8\sqrt{2}}$$

Quádricas em $\mathbb{R}^3$. A expressão geral de uma quádrica é

(73) $$ax^2 + by^2 + cz^2 + dxy + exz + fyz + gx + hy + iz + j = 0$$

Novamente a análise desta superfície baseia-se na análise dos termos de grau 2 (se estes se anulam identicamente a equação define um plano, o vazio ou todo o $\mathbb{R}^3$) que constituem a forma quadrática

$$\begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} a & \frac{d}{2} & \frac{e}{2} \\ \frac{d}{2} & b & \frac{f}{2} \\ \frac{e}{2} & \frac{f}{2} & c \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Num referencial ortonormado formado por vetores próprios da matriz simétrica que ocorre na expressão acima, a expressão (73) transforma-se em

$$\lambda_1 u^2 + \lambda_2 v^2 + \lambda_3 w^2 + g'u + h'v + i'z + j = 0$$

Módulo translações nos eixos dos u, v, w podemos assumir que as constantes $g' = h' = i'$ se anulam, desde que o λ_i correspondente não se anule. Temos então os seguintes casos:

- $\lambda_1, \lambda_2, \lambda_3$ todos diferentes de 0 e com sinais iguais (que podemos assumir positivos): A equação define o conjunto vazio se $j < 0$, um ponto se $j = 0$ e um *elipsóide* se $j > 0$ (trata-se da superfície que se obtém de uma superfície esférica reescalando os eixos).
- $\lambda_1, \lambda_2, \lambda_3$ todos diferentes de 0 e com sinais não todos iguais (podemos assumir que $\lambda_1, \lambda_2 > 0$ e $\lambda_3 < 0$: Os protótipos destas superfícies são as definidas pelas equações

  $$x^2 + y^2 - z^2 = 1, \qquad x^2 + y^2 - z^2 = 0, \qquad x^2 + y^2 - z^2 = -1$$

  Para entender a sua forma convém observar que o significado geométrico de $\sqrt{x^2 + y^2}$ é (pelo Teorema de Pitágoras) a distância do ponto (x, y, z) ao eixo dos zz. Num qualquer semiplano limitado pelo eixo dos zz podemos usar $r = \sqrt{x^2 + y^2} \geq 0$ como coordenada ao longo do semi-eixo perpendicular a Oz e a equação da interseção da nossa superfície com esse semiplano é determinada pela equação

  $$r^2 - z^2 = 1, \qquad r^2 - z^2 = 0, \qquad r^2 - z^2 = -1$$

  ou seja, trata-se de uma hipérbole nos casos em que o termo direito é ± 1 e de um par de semi-retas no caso restante. As superfícies que pretendemos descrever obtêm-se *rodando estas curvas em torno do eixo Oz*. Denominam-se respetivamente um *hiperbolóide*, um *cone* e um *hiperbolóide de duas folhas*.
- $\lambda_1 = 0$ e $\lambda_2, \lambda_3 \neq 0$ com o mesmo sinal que podemos assumir positivo: Os protótipos são agora da forma

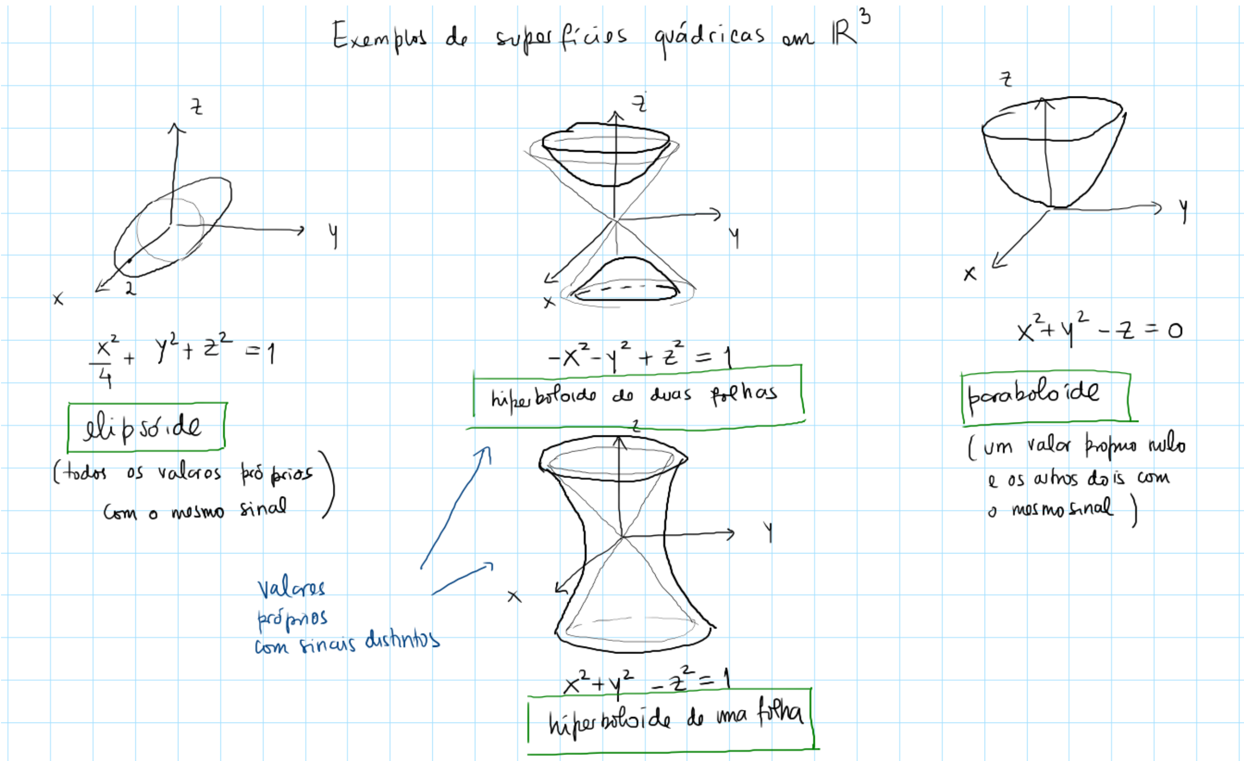
  $$x^2 + y^2 = j', \quad x^2 + y^2 - z = j'$$

  que são respetivamente o vazio, o eixo dos zz ou um *cilindro* em torno do eixo dos zz no primeiro caso, ou um *parabolóide* (uma parábola $z = r^2 - j'$ rodada em torno do eixo dos zz).
- $\lambda_1 = 0$ e $\lambda_2, \lambda_3 \neq 0$ com sinais diferentes (podemos assumir $\lambda_2 > 0, \lambda_3 < 0$): Os protótipos são

  $$x^2 - y^2 = j', \quad x^2 - y^2 - z = j'$$

  No primeiro caso trata-se de um *cilindro hiperbólico*, isto é, de uma hipérbole transladada ao longo do eixo dos zz (ou no caso degenerado em que $j' = 0$, da união de dois planos concorrentes no eixo dos zz), enquanto que no segundo a superfície designa-se por uma *sela* uma vez que tem o aspeto de uma sela de um cavalo (há uma parábola virada para cima ao longo do eixo dos xx e uma decrescente ao longo do eixo dos yy).

- $\lambda_1 = \lambda_2 = 0$ e $\lambda_3 > 0$. Os protótipos são agoras as equações da forma
$$z^2 + g'x + h'y = j'$$
Se $g' = h' = 0$ esta equação define o vazio, um plano ou dois planos paralelos consoante o sinal de j'. No caso em que $(g', h') \neq 0$ define *um cilindro parabólico*, isto é a translação de uma parábola, ao longo de um eixo no plano xy perpendicular ao vetor (g', h').



7.78. **Exercícios.**

1. Indique justificadamente se as seguintes aplicações são produtos internos nos espaços vetoriais indicados
   (a) $\langle (x, y), (u, v) \rangle = xu + 2xv + yv$ em $\mathbb{R}^2$.
   (b) $\langle (x, y), (u, v) \rangle = xu + 2xv + 2yu + yv$ em $\mathbb{R}^2$.
   (c) $\langle (x, y), (u, v) \rangle = 2xu - xv - yu + 2yv$ em $\mathbb{R}^2$.
   (d) $\langle (x, y), (u, v) \rangle = x^2u + xv^2$ em $\mathbb{R}^2$.
   (e) $\langle A, B \rangle = \mathrm{tr}(AB)$ em $M_{n\times n}(\mathbb{R})$.
   Resolução.
2. Considere o produto interno em $\mathbb{R}^3$ definido por
$$\langle (x, y, z), (u, v, w) \rangle = 2xu - yu + 2yv - xv + zw$$
   (a) Calcule $\langle (1, 0, 2), (1, -2, 1) \rangle$.
   (b) Calcule $\|(1, 1, 1)\|$.

(c) Determine a distância de $(0,0,1)$ a $(1,0,0)$.
(d) Determine o conjunto dos vetores ortogonais a $(1,2,3)$.
(e) Determine se os seguintes conjuntos são ortogonais ou ortonormados:

$$S = \{(1,1,1),(2,2,2)\} \qquad X = \{(1,2,1),(0,-1,3),(11,1,-3)\}$$

Resolução.

3. Considere o espaço vetorial $\mathbb{C}^3$ com o produto interno usual
   (a) Calcule $\langle (1, 2+i, 1-i), (i, 1-3i, i) \rangle$.
   (b) Determine $\|(i+2, 3i, -1)\|$.
   (c) Calcule a distância de $(1, 2+i, i)$ a $(1, 3+2i, 2)$.
   (d) Determine a equação cartesiana do plano perpendicular a $(i+2, 2, 3-i)$.
   Resolução.
4. Seja $V = \{a + bt + ct^2 \colon a, b, c \in \mathbb{R}\}$ o espaço vetorial dos polinómios de grau menor ou igual a 2.
   (a) Mostre que a função $\langle \cdot, \cdot \rangle \colon V \times V \to \mathbb{R}$ definida por
   $$\langle p, q \rangle = p(-1)q(-1) + p(0)q(0) + p(1)q(1)$$
   é um produto interno em V.
   (b) Determine uma base ortogonal para V (isto é, uma base de V que seja um conjunto ortogonal).
   (c) Determine a matriz da métrica deste produto interno com respeito à base ordenada $B = (1, t, t^2)$
   Resolução.
5. Considere o espaço vetorial $C([0,\pi],\mathbb{R})$ das funções contínuas de $[0,\pi]$ para $\mathbb{R}$ com o produto interno usual dado por
   $$\langle f, g \rangle = \int_0^\pi f(x)g(x)dx$$
   (a) Calcule $\langle x, 2 + x^2 \rangle$.
   (b) Determine uma base ortonormada para o subespaço formado pelos polinómios de grau ≤ 2.
   (c) Mostre que os conjuntos $S_1 = \{\cos(nx) \colon n \geq 0\}$ e $S_2 = \{\text{sen}(nx) \colon n \geq 1\}$ são ambos conjuntos ortogonais e calcule as normas dos seus elementos.
   Resolução.
6. Determine um produto interno em $\mathbb{R}^2$ para o qual $\{(1,2),(-1,1)\}$ seja uma base ortonormada (isto é uma base que forma um conjunto ortonormado). Resolução.
7. Mostre que um conjunto ortogonal que não contém 0 é linearmente independente.Resolução.
8. Considere o subespaço vetorial
   $$V = \{(z_1, z_2, z_3, z_4) \in \mathbb{C}^4 \colon z_1 + z_2 + iz_3 = 0\}$$
   de $\mathbb{C}^4$.
   (a) Determine uma base B para V.
   (b) Determine a matriz da métrica para a restrição do produto interno usual em $\mathbb{C}^4$ a V.

(c) Determine a equação nas coordenadas dadas pela base B para o plano $W \subset V$ perpendicular ao vetor $(0, 1, i, 2)$.

Resolução.

9. Seja $V = M_{n\times n}(\mathbb{R})$.
 (a) Mostre que $\langle A, B\rangle = \mathrm{tr}(A^T B)$ é um produto interno em V.
 (b) Determine a matriz da métrica para este produto interno com respeito à base
$$\left(\begin{bmatrix} 1 & 2 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 2 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 2 & -3 \end{bmatrix}\right)$$
 para $M_{2\times 2}(\mathbb{R})$.
 (c) Qual é a matriz simétrica mais próxima de $\begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}$?

Resolução.

10. Seja V um espaço vetorial (real ou complexo) com produto interno $\langle\cdot,\cdot\rangle$, e $S \subset V$.
 (a) Mostre que $S^\perp = \{v \in V : \langle x, v\rangle = 0 \text{ para todo o } x \in S\}$ é um subespaço vetorial de V.
 (b) Mostre que $L(S) \subset (S^\perp)^\perp$.

Resolução.

11. Seja V um espaço vetorial e $\langle\cdot,\cdot\rangle$ um produto interno em V. Recorde que o dual de V é o espaço vetorial V^* das transformações lineares de V para o corpo dos escalares ($\mathbb{R}$ ou $\mathbb{C}$). Para cada $v \in V$ seja $\phi_v \in V^*$ o elemento definido pela fórmula
$$\phi_v(w) = \langle v, w\rangle$$
 (a) Mostre que se V é real e tem dimensão finita, a aplicação $v \mapsto \phi_v$ é um isomorfismo entre V e o seu dual.
 (b) Se V é um espaço vetorial complexo define-se o *espaço vetorial conjugado* $\overline{V}$ como sendo o mesmo conjunto V, com a mesma operação de soma de vetores mas com uma nova operação $\cdot_{\overline{V}}$ de multiplicação por escalar definida em termos da operação de produto por escalar em V por
$$\alpha \cdot_{\overline{V}} v = \overline{\alpha} \cdot v$$
 Mostre que se V é um espaço vetorial complexo de dimensão finita, a aplicação $v \mapsto \phi_v$ determina um isomorfismo entre $\overline{V}$ e V^*.

Resolução.

12. Seja V um espaço vetorial real ou complexo. Uma função $\|\cdot\| : V \to \mathbb{R}_0^+$ diz-se uma *norma* se
$$\|v\| = 0 \Rightarrow v = 0, \qquad \|\alpha v\| = |\alpha|\|v\|, \qquad \|v + w\| \leq \|v\| + \|w\| \quad \text{para todos os } \alpha, v, w$$
 (a) Verifique que se uma norma provém de um produto interno então é válida a *identidade do paralelogramo* (que relaciona o comprimento das diagonais de um paralelogramo com os comprimentos das arestas):
$$2\|v\|^2 + 2\|w\|^2 = \|v + w\|^2 + \|v - w\|^2$$

Pode mostrar-se que, reciprocamente, se a identidade acima é satisfeita, então a norma provém de um produto interno dado pela expressão:

$$\langle v, w\rangle = \frac{1}{4}(\|v+w\|^2 - \|v-w\|^2)$$

(b) Mostre que a função de $\mathbb{R}^n$ para $\mathbb{R}_0^+$ definida pela expressão

$$\|(x_1, \ldots, x_n)\| = |x_1| + \ldots + |x_n|$$

é uma norma que não provém de um produto interno. Para $n = 2, 3$ esboce o conjunto $\{x \in \mathbb{R}^n : \|x\| \leq 1\}$ (que se chama a *bola* de raio 1 para a norma dada).

Resolução.

13. **Relação entre os produtos internos real e complexo.** Seja $(V, \langle \cdot, \cdot \rangle)$ um espaço vetorial complexo de dimensão finita. Recorde que V tem uma estrutura natural de espaço vetorial real, restringindo o produto por escalar aos números reais. Neste exercício designamos este espaço vetorial real por $V_{\mathbb{R}}$.

(a) Considere as funções $V_{\mathbb{R}} \times V_{\mathbb{R}} \to \mathbb{R}$ definidas por

$$g(v, w) = \operatorname{Re}\langle v, w\rangle, \qquad \omega(v, w) = \operatorname{Im}\langle v, w\rangle$$

Mostre que g é um produto interno em $V_{\mathbb{R}}$ e ω é uma forma bilinear não degenerada em $V_{\mathbb{R}}$ (ver Exercício 5.14(c)) tal que $\omega(v, w) = -\omega(w, v)$ para todos os $v, w \in V_{\mathbb{R}}$ (diz-se que ω é *anti-simétrica*).

(b) Verifique que g e ω satisfazem

$$\omega(v, w) = -g(v, iw), \qquad \omega(v, iw) = g(v, w)$$

onde iw denota a multiplicação (em V) de w pelo escalar complexo i.

(c) Sendo g e ω respetivamente um produto interno e uma forma bilinear anti-simétrica num espaço vetorial real W de dimensão finita mostre que existe um único endomorfismo $J : W \to W$ tal que

$$g(v, Jw) = \omega(v, w) \quad \text{para todos os } v, w \in W$$

e que um tal endomorfismo satisfaz necessariamente a relação

$$g(v, Jw) = -g(Jv, w) \quad \text{para todos os } v, w \in W$$

Que condição tem J que verificar para que g e ω possam ser obtidos de um produto interno complexo pelo procedimento da alínea (a)?

Resolução.

14. Seja V um espaço vetorial com produto interno. Determine em que condições é que a igualdade $\|x+y\| = \|x\| + \|y\|$ se verifica, e mostre que $\|x+y\| \geq |\|x\| - \|y\||$ para todos os $x, y \in V$. Resolução.

15. Mostre que a reflexão no plano perpendicular a um vetor $u \in \mathbb{R}^n \setminus \{0\}$ é dada pela fórmula

$$R(v) = v - 2\operatorname{proj}_u(v) = v - 2\frac{\langle u, v\rangle}{\langle u, u\rangle}u$$

Resolução.

16. Determine uma base ortonormada para $\mathbb{R}^4$ (com o produto interno usual) que contenha os vetores $(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}, 0)$ e $(-\frac{1}{\sqrt{3}}, 0, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$, e calcule as coordenadas do vetor $(1,2,3,4)$ nessa base. Resolução.
17. Considere o produto interno usual em $\mathbb{R}^n$. Sendo $(u_1, \ldots, u_k)$ uma base ortonormada para o subespaço $U \subset \mathbb{R}^n$ e A a matriz $n \times k$ que tem os vetores $u_1, \ldots, u_k$ como colunas, mostre que a matriz que representa a projeção ortogonal P_U com respeito à base canónica é AA^T. *Sugestão: Pense qual é o significado da matriz coluna que se obtém multiplicando um vetor por* A^T*.* Resolução.
18. Determine a expressão para a projeção ortogonal de $\mathbb{R}^4$ (com o produto interno usual) no plano
$$\{(x,y,z,w) \in \mathbb{R}^4 \colon x+y+w=0, y+z+2w=0\}$$
Resolução.
19. A transformação linear definida por $P(x,y) = \left(\frac{x}{2}+\frac{y}{2}, \frac{x}{2}+\frac{y}{2}\right)$ é uma projeção ortogonal? Justifique. Resolução.
20. Considere o produto interno em $\mathbb{R}^3$ que tem $\{(1,0,1),(0,1,0),(1,0,-1)\}$ como base ortonormada.
   (a) Determine uma base ortonormada (relativamente à restrição deste produto interno) para o subespaço $U = \{(x,y,z) \colon x+y+z=0\}$.
   (b) Decomponha o vetor $(1,2,-1)$ como a soma de um vetor em U e de outro em $U^\perp$.
   Resolução.
21. Considere o produto interno em $M_{2\times 2}(\mathbb{R})$ dado por $\langle A, B\rangle = \operatorname{tr}(A^T B)$.
   (a) Calcule o ângulo entre as matrizes $\begin{bmatrix} 1 & 2 \\ -1 & 0 \end{bmatrix}$ e $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$.
   (b) Determine a projeção ortogonal de $\begin{bmatrix} 1 & 2 \\ -1 & 0 \end{bmatrix}$ no subespaço $U = \{A \in M_{2\times 2}(\mathbb{R}) \colon \operatorname{tr} A = 0\}$.
   Resolução.
22. Determine a expressão para a projeção ortogonal do espaço dos polinómios de grau menor ou igual que 2 no subespaço dos polinómios de grau menor ou igual a 1 com respeito ao produto interno definido por $\langle p, q\rangle = \int_0^1 p(x)q(x)dx$. Resolução.
23. Considere o espaço $C([0,\pi],\mathbb{R})$ das funções contínuas com o produto interno usual. Determine a projeção ortogonal de $f(x) = x$ no subespaço gerado pelas funções $\operatorname{sen} x$, $\operatorname{sen}(2x)$ e $\operatorname{sen}(3x)$. Veja o gráfico da função obtida num computador. Resolução.
24. Determine
   (a) A equação cartesiana de uma reta que é perpendicular ao plano $-x+2y+3z=1$ e que passa pelo ponto $(1,0,0)$.
   (b) A equação cartesiana de um plano $U \subset \mathbb{R}^4$ que é perpendicular aos vetores $(1,0,2,1)$ e $(0,1,2,-1)$ e passa pelo ponto $(1,1,1,1)$.
   Resolução.
25. Calcule as seguintes distâncias (sempre para o produto interno usual no $\mathbb{R}^n$ indicado).
   (a) A distância em $\mathbb{R}^3$ do ponto $(0,1,2)$ ao plano definido pela equação $x+y+z=1$.
   (b) A distância em $\mathbb{R}^3$ de $(1,2,1)$ à reta $\{\alpha(0,3,-1) \colon \alpha \in \mathbb{R}\}$.

(c) A distância em $\mathbb{R}^4$ entre os planos definidos pelas equações $x+y-z+3w=5$ e $x+y-z+3w=2$.
(d) A distância em $\mathbb{R}^3$ do ponto $(-1,0,1)$ à reta definida pelas equações $x+z=2, x+y+z=1$.

Resolução.

26. **Mais sobre ângulos.** Considere $\mathbb{R}^n$ com o seu produto interno usual. O *ângulo* entre dois subespaços não nulos $U,V\subset\mathbb{R}^n$ tais que $U\cap V=\{0\}$ define-se como o mais pequeno ângulo entre vetores (não nulos) $u\in U$ e $v\in V$.
    (a) Mostre que o ângulo entre dois planos U,V pertence a $[0,\frac{\pi}{2}]$.
    (b) Em $\mathbb{C}^n$ o ângulo entre dois vetores v,w não nulos define-se como o ângulo entre eles para o produto interno real em $\mathbb{C}^n$ definido por $(v,w)\mapsto \mathrm{Re}(\langle v,w\rangle)$ (ver o Exercício 13 (a)). Calcule o ângulo entre $(1+i,i)$ e $(2-i,1+i)$ em $\mathbb{C}^2$.
    (c) Mostre que o ângulo entre as linhas complexas geradas por vetores não colineares v e w em $\mathbb{C}^n$ (encaradas como planos em $\mathbb{R}^{2n}$) é arccos $\frac{|\langle v,w\rangle|}{\|v\|\|w\|}$
    (d) O *ângulo dihedral* entre dois planos U,V que se intersetam num subespaço não nulo $W=U\cap V$ (e tais que nenhum deles é igual à interseção) é definido como o ângulo entre os planos $U\cap W^\perp$ e $V\cap W^\perp$. Calcule o ângulo dihedral entre os planos $U=\{(x,y,z,w)\colon x-y+z-w=0\}$ e $V=\{(x,y,z,w)\colon 2x+y-z+3w=0\}$.

    Resolução.
27. Determine pelo método dos quadrados mínimos.
    (a) A reta $y=ax+b$ que melhor aproxima os pontos $(-1,2),(0,3),(2,6)\in\mathbb{R}^2$.
    (b) A função da forma $\alpha\cos x+\beta\,\mathrm{sen}\,x$ cujo que melhor aproxima os pontos $(-\frac{\pi}{2},1),(0,0),(\frac{\pi}{2},2)$.
    (c) O polinómio do segundo grau que melhor aproxima os pontos $(0,0),(1,1),(2,2),(-1,1)$.
    (d) O plano $z=a+bx+cy$ que melhor aproxima os pontos $(0,0,0),(1,0,1),(-1,2,0),(1,2,3)$.

    Resolução.
28. Seja V um espaço vetorial com produto interno. Mostre que as seguintes afirmações acerca de uma transformação $T\colon V\to V$ são equivalentes:
    (i) T é unitária/ortogonal.
    (ii) T é uma isometria, isto é $\|T(x)-T(y)\|=\|x-y\|$ para todos os $x,y\in V$.
    (iii) T envia a esfera unitária de V nela própria, isto é $\|T(x)\|=1$ para todo o $x\in V$ com $\|x\|=1$.

    Resolução.
29. Diga se existe uma matriz simétrica real com os vetores próprios indicados e em caso afirmativo ache uma
    (a) Vetores próprios $(1,-1,1),(1,1,-1)$ com valores próprios $1,2$.
    (b) Vetores próprios $(1,-1,1),(1,1,-1)$ ambos com valor próprio 1.
    (c) Vetores próprios $(1,0,-1),(1,1,1),(1,-2,1)$ com valores próprios $-1,0,1$ respetivamente.

    Resolução.
30. Classifique as seguintes formas quadráticas.
    (a) $f(x,y)=3x^2+xy-y^2$
    (b) $f(x,y)=x^2+2xy+y^2$

(c) $f(x,y) = -x^2 + 3xy - 3y^2$
(d) $f(x,y) = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$
(e) $f(x,y,z) = x^2 + 2xy + xz - y^2 + yz + z^2$

Resolução.

31. Determine em cada caso se a curva plana é uma elipse, uma hipérbole ou uma parábola e ache os eixos.
(a) $2x^2 + xy + 3y^2 = 5$
(b) $x^2 + 2xy + y^2 + x - y = 3$
(c) $x^2 + 4xy + y^2 = 1$

Resolução.

32. Escreva uma expressão para as seguintes transformações ortogonais de $\mathbb{R}^3$:
(a) Rotação de 30 graus em torno do eixo $(1,1,0)$ num sentido que parece o horário quando visto do ponto $(10,10,0)$.
(b) Rotação de 45 graus em torno do eixo dos zz num sentido que parece o anti-horário quando visto de muito acima do plano xy, seguida de uma reflexão no plano ortogonal ao vetor $(0,1,-1)$.

Resolução.

33. Determine o eixo e o ângulo da rotação descrita pela matriz $\begin{bmatrix} \frac{1}{2}+\frac{1}{2\sqrt{2}} & -\frac{1}{2}+\frac{1}{2\sqrt{2}} & -\frac{1}{2} \\ -\frac{1}{2}+\frac{1}{2\sqrt{2}} & \frac{1}{2}+\frac{1}{2\sqrt{2}} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{\sqrt{2}} \end{bmatrix}$

Resolução.

34. Sejam V e W espaços vetoriais com produto interno e $T\colon V \to W$ uma transformação linear. Uma transformação linear $T^*\colon W \to V$ diz-se *adjunta* de T se para todos os $v \in V$ e $w \in W$ temos

$$\langle T(v), w\rangle = \langle v, T^*(w)\rangle$$

(a) Mostre que se V é finitamente gerado então a transformação adjunta existe e é única, sendo dada pela fórmula

$$T^*w = \langle Tv_1, w\rangle v_1 + \ldots + \langle Tv_n, w\rangle v_n$$

onde $\{v_1, \ldots, v_n\}$ denota uma qualquer base ortonormada para V.
(b) Mostre que se $V = \mathbb{C}^n, W = \mathbb{C}^m$ e $T(v) = Av$ para $A \in M_{m\times n}(\mathbb{C})$ então $T^*(w) = A^*w$ (onde A^* denota a matriz adjunta de A).
(c) Mostre que em termos da identificação de um espaço vetorial finitamente gerado com produto interno com o seu dual explicada no Exercício 11, T^* corresponde à transformação dual induzida por T entre os espaços duais W^* e V^*.
(d) Seja V o espaço vetorial dos polinómios de grau ≤ 2 com o produto interno definido por

$$\langle p, q\rangle = p(-1)q(-1) + p(0)q(0) + p(1)q(1)$$

e $W = M_{2\times 2}(\mathbb{R})$ com o produto interno definido por $\langle A, B\rangle = \mathrm{tr}(A^T B)$. Sendo $T\colon V \to W$ a transformação linear definida por

$$T(p) = \begin{bmatrix} p(1) & p(0)+p(1) \\ p(-1) & p(2) \end{bmatrix}$$

calcule a transformação adjunta $T^*\colon W \to V$.

Resolução.

35. (a) Sendo $A \in M_{n\times n}(\mathbb{R})$ uma matriz simétrica semi-definida positiva, mostre que existe uma única matriz simétrica semi-definida positiva B tal que $B^2 = A$ (que se chama a raíz quadrada simétrica de A). *Sugestão: Diagonalize.*
 (b) Conclua que toda a matriz simétrica definida positiva é a matriz da métrica do produto interno *usual* em $\mathbb{R}^n$ com respeito a alguma base de $\mathbb{R}^n$.
 (c) Calcule

$$\sqrt{\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}}$$

 Resolução.

36. Determine as matrizes hermitianas $A \in M_{2\times 2}(\mathbb{C})$ tais que

$$A^2 + A = \begin{bmatrix} 1 & -i \\ i & 1 \end{bmatrix}$$

Sugestão: Diagonalize. Resolução.

37. Determine a decomposição polar das seguintes matrizes

(a) $\begin{bmatrix} 2+i & 1+2i \\ 1+2i & 2+i \end{bmatrix}$; (b) $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ com $(a, b) \neq (0, 0)$

Resolução.

38. Determine uma decomposição em valores singulares das seguintes matrizes

(a) $\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & -1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$ (c) $\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$

Resolução.

39. Seja $A \in M_{m\times n}(\mathbb{R})$ e $A = U_1 D U_2$ uma decomposição singular de A e suponha que $\sigma_1, \ldots, \sigma_k$ são os valores singulares de A. Mostre que:
 - A característica de A é k.
 - As primeiras k colunas de U_1 são uma base ortonormada para $EC(A)$ e as restantes $m - k$ uma base ortonormada para $N(A^T)$.
 - As primeiras k linhas de U_2 são uma base ortonormada para $EL(A)$ e as últimas $n - k$ são uma base ortonormada para $N(A)$.

Resolução.

40. **Mais sobre rotações.** Seja $O(n) = \{A \in M_{n\times n}(\mathbb{R})\colon A^t A = I\}$ o conjunto das transformações ortogonais de $\mathbb{R}^n$ (que se costuma chamar *o grupo ortogonal*).
 (a) Mostre que se $A \in O(n)$ então $\det(A) = \pm 1$. Escrevemos

$$SO(n) = \{A \in O(n)\colon \det(A) = 1\}, \qquad O(n)_- = \{A \in O(n)\colon \det(A) = -1\}$$

(b) Mostre que

$$SO(2) = \left\{ \begin{bmatrix} \cos\alpha & -\operatorname{sen}\alpha \\ \operatorname{sen}\alpha & \cos\alpha \end{bmatrix} : \alpha \in [0, 2\pi] \right\}$$

$$O(2)_- = \left\{ \begin{bmatrix} \cos\alpha & \operatorname{sen}\alpha \\ \operatorname{sen}\alpha & -\cos\alpha \end{bmatrix} : \alpha \in [0, 2\pi] \right\}$$

(c) Mostre que os elementos de $O(2)_-$ são reflexões (na reta gerada por $(\cos\frac{\alpha}{2}, \operatorname{sen}\frac{\alpha}{2})$) e que os elementos de $SO(2)$ se podem escrever como o produto de duas reflexões.

(d) Dada A in $O(3)$ mostre que existe uma base ortonormada de $\mathbb{R}^3$ na qual a expressão de A é da forma

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{bmatrix} \quad \text{ou} \quad \begin{bmatrix} -1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{bmatrix}$$

consoante $A \in SO(3)$ ou $A \in O(3)_-$ respetivamente.

(e) Dado $A \in O(n)$, mostre que existe uma base ortonormada de $\mathbb{R}^n$ na qual a expressão de A é da forma

$$\begin{bmatrix} A_1 & 0 & 0 & 0 \\ 0 & A_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & A_{\frac{n}{2}} \end{bmatrix} \quad \text{ou} \quad \begin{bmatrix} \pm 1 & 0 & 0 & 0 \\ 0 & A_1 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & A_{\frac{n-1}{2}} \end{bmatrix}$$

consoante n é par ou ímpar, onde $A_i \in SO(2)$ para $i \geq 2$; $A_1 \in O(2)$ se n é par e $A_1 \in SO(2)$ quando n é ímpar.

(f) Mostre que todo o elemento de $O(n)$ pode ser escrito como um produto de k reflexões com $k \leq n$.

Resolução.

41. **A norma de uma matriz quadrada.** Seja $V \neq \{0\}$ um espaço vetorial de dimensão finita com produto interno e $T \colon V \to V$ um endomorfismo. Define-se a norma de T por

$$\|T\| = \sup_{v \in V \setminus \{0\}} \frac{\|T(v)\|}{\|v\|}$$

(a) Mostre que $\|T\|$ está bem definida, isto é, que o conjunto $\{\frac{\|T(v)\|}{\|v\|} : v \in V \setminus \{0\}\}$ é limitado superiormente.

(b) Mostre que para todo o $v \in V$ se tem $\|T(v)\| \leq \|T\|\|v\|$.

(c) Mostre que $T \mapsto \|T\|$ é de facto uma norma no espaço vetorial $L(V, V)$ (veja o Exercício 12).

(d) Mostre que se T é diagonalizável por uma base ortogonal então $\|T\| = \max\{|\lambda_1|, \ldots, |\lambda_k|\}$ onde λ_i são os valores próprios de T.

(e) Mostre que, em geral, $\|T\| = \sqrt{\|T^* T\|}$, isto é, a norma é o maior dos valores singulares de T.

(f) Para o produto interno usual em $\mathbb{R}^2$, calcule a norma de $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$.

(g) **A exponencial de matrizes.** O limite de uma sucessão de vetores num espaço de dimensão finita é definido coordenada a coordenada (em qualquer base). Mostre que para qualquer $T \in L(V, V)$ a expressão

$$v \mapsto \sum_{n=0}^{\infty} \frac{1}{n!} T^n(v)$$

define um endomorfismo de V (que se denota por e^T). Calcule $e^{\begin{bmatrix} 3t & t \\ 0 & 3t \end{bmatrix}}$.

Resolução.

42. Este exercício caracteriza as transformações lineares que são diagonalizáveis com respeito a uma base *ortogonal*. Uma matriz $A \in M_{n\times n}(\mathbb{C})$ diz-se **normal** se $A^*A = AA^*$.
   (a) Verifique que se A é unitária, hermitiana ou anti-hermitiana então A é normal.
   (b) Verifique que se $A = UDU^{-1}$ com U unitária e D diagonal então A é normal.
   (c) Nas restantes alíneas, assuma que A é normal. Mostre que AA^* é hermitiana e semi-definida positiva.
   (d) Verifique que A comuta com AA^* e portanto os espaços próprios de AA^* são invariantes por A.
   (e) Mostre que se $E(\lambda)$ é o espaço próprio de AA^* correspondente ao valor próprio $\lambda > 0$ então a restrição de $\frac{1}{\sqrt{\lambda}}A$ a $E(\lambda)$ é uma transformação unitária.
   (f) Conclua que se A é normal existe uma matriz unitária U e uma matriz diagonal D tal que $A = UDU^{-1}$.
   (g) Determine se é possível achar uma matriz unitária que diagonalize as seguintes matrizes:

$$(i) \begin{bmatrix} 2 & 3 \\ -1 & 2 \end{bmatrix} \qquad (ii) \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

Resolução.

8. Soluções dos exercícios

8.1. Sistemas lineares.

(1) Enunciado Aplicamos o método de Gauss:
   (a)

$$\begin{cases} x - 2y + 3z - w = 1 \\ 3x - y + 2z + 5w = 2 \\ -3x + 6y - 9z + 3w = -6 \end{cases} \Leftrightarrow \begin{cases} x - 2y + 3z - w = 1 \\ 5y - 7z + 8w = -1 \\ 0 = -3 \end{cases}$$

logo o sistema é impossível.

(b)

$$\begin{cases} -2v+3w=3 \\ 3u+6v-3w=-2 \\ 6u+6v+3w=5 \end{cases} \Leftrightarrow \begin{cases} 3u+6v-3w=-2 \\ -2v+3w=3 \\ 6u+6v+3w=5 \end{cases} \Leftrightarrow \begin{cases} 3u+6v-3w=-2 \\ -2v+3w=3 \\ -6v+9w=9 \end{cases}$$

$$\Leftrightarrow \begin{cases} 3u+6v-3w=-2 \\ -2v+3w=3 \\ 0=0 \end{cases} \Leftrightarrow \begin{cases} u=\frac{1}{3}\left(-2-6(-\frac{3}{2}+\frac{3}{2}w)+3w\right)=\frac{7}{3}-2w \\ v=-\frac{3}{2}+\frac{3}{2}w \end{cases}$$

com $w \in \mathbb{R}$ arbitrário.

(2) Enunciado Podemos escrever um polinómio na forma $p(x)=ax^3+bx^2+cx+d$ de forma que $p'(x)=3ax^2+2bx+c$ e substituindo as condições do enunciado obtemos o sistema

$$\begin{cases} a+b+c+d=0 \\ 8a+4b+2c+d=3 \\ c=0 \end{cases} \Leftrightarrow \begin{cases} a+b+d=0 \\ -4b-7d=3 \\ c=0 \end{cases} \Leftrightarrow \begin{cases} a=\frac{3}{4}+\frac{7}{4}d-d=\frac{3}{4}+\frac{3}{4}d \\ b=-\frac{3}{4}-\frac{7}{4}d \\ c=0 \end{cases}$$

Logo o conjunto pedido é

$$\left\{(\tfrac{3}{4}+\tfrac{3}{4}d)x^3-(\tfrac{3}{4}+\tfrac{7}{4}d)x^2+d\colon d\in\mathbb{R}\right\}$$

(3) Enunciado Após cada encontro o número de extraterrestres mantém-se constante pelo que se ficarem todos da mesma cor haverá 45 extraterrestres da mesma cor.

Sejam x, y, z o número de encontros entre extraterrestres amarelos e azuis, amarelos e verdes, e azuis e verdes respetivamente. Após um tal número de encontros o número total de extraterrestres amarelos é dado por

$$15-x-y+2z$$

Analogamente os números de extraterrestres azuis e verdes são, respetivamente,

$$13-x+2y-z \quad \text{e} \quad 17+2x-y-z$$

Para que, por exemplo, todos os extraterrestres fiquem amarelos é necessário que o seguinte sistema tenha solução com x, y, z números naturais não negativos:

$$\begin{cases} 15-x-y+2z=45 \\ 13-x+2y-z=0 \\ 17+2x-y-z=0 \end{cases} \Leftrightarrow \begin{cases} -x-y+2z=30 \\ -x+2y-z=-13 \\ 2x-y-z=-17 \end{cases} \Leftrightarrow \begin{cases} -x-y+2z=30 \\ 3y-3z=-43 \\ -3y+3z=43 \end{cases}$$

$$\Leftrightarrow \begin{cases} x=-30-y+\frac{86}{3}+2y=-\frac{4}{3}+y \\ z=\frac{43}{3}+y \end{cases}$$

Uma vez que o sistema não tem soluções naturais conclui-se que os extraterrestres não podem ficar todos amarelos. De forma análoga vemos que não podem ficar todos verdes nem todos azuis. Alternativamente podemos notar que, por exemplo, com dois encontros podemos passar do estado inicial para 13 criaturas amarelas, 17 azuis e 15 verdes e usar simetria para ver que não podem ficar todos verdes e, de forma análoga, ver que não podem ficar todos azuis.

(4) Enunciado Vamos aplicar o método de Gauss para resolver o sistema. Se $a = 0$ o sistema produz imediatamente a solução única $y = 0, x = 1$. Caso contrário obtemos

$$\begin{cases} ax + y = a^2 \\ x + ay = 1 \end{cases} \Leftrightarrow \begin{cases} ax + y = a^2 \\ (a - \frac{1}{a})y = 1 - a \end{cases}$$

Como $a - \frac{1}{a} = 0 \Leftrightarrow a = \pm 1$, nestes casos, a segunda equação transforma-se em $0 = 0$ quando $a = 1$ e em $0 = 2$ quando $a = -1$. No primeiro caso, o sistema é equivalente à equação $x + y = 1$ que tem solução não única, no segundo caso, o sistema é impossível. Finalmente se $a \neq \pm 1, a \neq 0$ o sistema é determinado (tem solução única). Conclui-se que o sistema tem solução se e só se $a \neq -1$ e que esta é única exceto quando $a = 1$.

(5) Enunciado Se $\alpha = 0$ o sistema fica

$$\begin{cases} \beta z = 2 \\ 4z = 4 \\ 2z = \beta \end{cases} \Leftrightarrow \begin{cases} \beta = 2 \\ z = 1 \\ \beta = 2 \end{cases}$$

pelo que o sistema é possível e indeterminado se $\beta = 2$ e impossível se $\beta \neq 2$. Se $\alpha \neq 0$, aplicando o método de Gauss obtemos

$$\begin{cases} \alpha x + \beta z = 2 \\ \alpha y + (4 - \beta)z = 2 \\ \alpha y + 2z = \beta \end{cases} \Leftrightarrow \begin{cases} \alpha x + \beta z = 2 \\ \alpha y + (4 - \beta)z = 2 \\ (\beta - 2)z = \beta - 2 \end{cases}$$

Se $\beta = 2$ o sistema é possível indeterminado. Caso contrário o sistema é possível e determinado. Resumindo, o sistema é

- Impossível se $\alpha = 0$ e $\beta \neq 2$.
- Possível determinado se $\alpha \neq 0$ e $\beta \neq 2$
- Possível indeterminado se $\beta = 2$

(6) Enunciado Temos

$$\begin{cases} x = 2t + 3s \\ y = t + s - 1 \\ z = 2s + 1 \\ w = t - 1 \end{cases} \Leftrightarrow \begin{cases} x = 2(w + 1) + \frac{3}{2}(z - 1) \\ y = (w + 1) + \frac{z-1}{2} - 1 \\ s = \frac{z-1}{2} \\ t = w + 1 \end{cases} \Leftrightarrow \begin{cases} x - \frac{3}{2}z - 2w = \frac{1}{2} \\ y - \frac{1}{2}z - w = -\frac{1}{2} \\ s = \frac{z-1}{2} \\ t = w + 1 \end{cases}$$

logo o conjunto dado é (por exemplo) o conjunto das soluções do sistema

$$\begin{cases} x - \frac{3}{2}z - 2w = \frac{1}{2} \\ y - \frac{1}{2}z - w = -\frac{1}{2} \end{cases}$$

(7) Enunciado

(a)

$$\begin{bmatrix}2 & 3\\ 1 & -2\\ 1 & 5\end{bmatrix} \to \begin{bmatrix}2 & 3\\ 0 & -\frac{7}{2}\\ 0 & \frac{7}{2}\end{bmatrix} \to \begin{bmatrix}2 & 3\\ 0 & -\frac{7}{2}\\ 0 & 0\end{bmatrix} \to \begin{bmatrix}1 & \frac{3}{2}\\ 0 & 1\\ 0 & 0\end{bmatrix} \to \begin{bmatrix}1 & 0\\ 0 & 1\\ 0 & 0\end{bmatrix}$$

(b)

$$\begin{bmatrix}i & -(1+i) & 0\\ 1 & -2 & 1\\ 1 & 2i & -1\end{bmatrix} \to \begin{bmatrix}i & -(1+i) & 0\\ 0 & -1-i & 1\\ 0 & 1+i & -1\end{bmatrix} \to \begin{bmatrix}i & -(1+i) & 0\\ 0 & -1-i & 1\\ 0 & 0 & 0\end{bmatrix} \to$$

$$\begin{bmatrix}1 & -1+i & 0\\ 0 & 1 & \frac{-1+i}{2}\\ 0 & 0 & 0\end{bmatrix} \to \begin{bmatrix}1 & 0 & i\\ 0 & 1 & \frac{-1+i}{2}\\ 0 & 0 & 0\end{bmatrix}$$

(c)

$$\begin{bmatrix}1 & 0 & 3 & 1 & 2\\ 1 & 4 & 2 & 1 & 5\\ 3 & 4 & 8 & 1 & 2\end{bmatrix} \to \begin{bmatrix}1 & 0 & 3 & 1 & 2\\ 0 & 4 & -1 & 0 & 3\\ 0 & 4 & -1 & -2 & -4\end{bmatrix} \to \begin{bmatrix}1 & 0 & 3 & 1 & 2\\ 0 & 4 & -1 & 0 & 3\\ 0 & 0 & 0 & -2 & -7\end{bmatrix} \to$$

$$\begin{bmatrix}1 & 0 & 3 & 1 & 2\\ 0 & 1 & -\frac{1}{4} & 0 & \frac{3}{4}\\ 0 & 0 & 0 & 1 & \frac{7}{2}\end{bmatrix} \to \begin{bmatrix}1 & 0 & 3 & 0 & -\frac{3}{2}\\ 0 & 1 & -\frac{1}{4} & 0 & \frac{3}{4}\\ 0 & 0 & 0 & 1 & \frac{7}{2}\end{bmatrix}$$

(8) Enunciado Aplicando o método de Gauss obtemos

$$\begin{bmatrix}1 & 1 & 1 & -1 & | & a\\ 1 & 0 & -2 & 1 & | & b\\ 1 & -1 & -5 & 3 & | & c\end{bmatrix} \to \begin{bmatrix}1 & 1 & 1 & -1 & | & a\\ 0 & -1 & -3 & 2 & | & b-a\\ 0 & -2 & -6 & 4 & | & c-a\end{bmatrix} \to \begin{bmatrix}1 & 1 & 1 & -1 & | & a\\ 0 & -1 & -3 & 2 & | & b-a\\ 0 & 0 & 0 & 0 & | & c+a-2b\end{bmatrix}$$

Conclui-se que o sistema tem solução se e só se $c + a - 2b = 0$. Para achar as soluções nesse caso continuamos a aplicação do método de Gauss-Jordan:

$$\begin{bmatrix}1 & 1 & 1 & -1 & | & a\\ 0 & -1 & -3 & 2 & | & b-a\\ 0 & 0 & 0 & 0 & | & c+a-2b\end{bmatrix} \to \begin{bmatrix}1 & 1 & 1 & -1 & | & a\\ 0 & 1 & 3 & -2 & | & -b+a\\ 0 & 0 & 0 & 0 & | & 0\end{bmatrix} \to \begin{bmatrix}1 & 0 & -2 & 1 & | & b\\ 0 & 1 & 3 & -2 & | & -b+a\\ 0 & 0 & 0 & 0 & | & 0\end{bmatrix}$$

pelo que o conjunto das soluções é o conjunto

$$\{(2z - w + b, -3z + 2w - b + a, z, w) \colon z, w \in \mathbb{R}\}$$

(9) Enunciado

(a) No primeiro passo do método podemos trocar a primeira linha com a segunda ou com a terceira, obtendo respetivamente

$$\begin{bmatrix}2 & 0 & 1\\ 0 & 1 & 1\\ -1 & 0 & 3\end{bmatrix} \to \begin{bmatrix}2 & 0 & 1\\ 0 & 1 & 1\\ 0 & 0 & \frac{7}{2}\end{bmatrix}$$

ou

$$\begin{bmatrix} -1 & 0 & 3 \\ 2 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \to \begin{bmatrix} -1 & 0 & 3 \\ 0 & 0 & 7 \\ 0 & 1 & 1 \end{bmatrix} \to \begin{bmatrix} -1 & 0 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & 7 \end{bmatrix}$$

(b) Aplicando Gauss-Jordan à primeira matriz obtemos

$$\begin{bmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

e à segunda

$$\begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

(10) EnunciadoAplicamos o método de Gauss

(a)

$$\begin{bmatrix} 0 & 1 & 1 \\ 2 & 3 & -3 \\ -1 & 0 & 3 \\ 0 & 4 & 4 \\ 1 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 1 \\ 2 & 3 & -3 \\ -1 & 0 & 3 \\ 0 & 4 & 4 \\ 0 & 1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 1 \\ 0 & 3 & -5 \\ 0 & 0 & 4 \\ 0 & 4 & 4 \\ 0 & 1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 1 \\ 0 & 3 & -5 \\ 0 & 0 & 4 \\ 0 & 0 & \frac{32}{3} \\ 0 & 0 & \frac{8}{3} \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 1 \\ 0 & 3 & -5 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Conclui-se que a matriz tem característica 3.

(b)

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & -4 & -8 & -12 \\ 0 & -8 & -16 & -24 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & -4 & -8 & -12 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Conclui-se que a matriz tem característica 2.

(11) Enunciado

(a) É conveniente trocar a primeira linha com a terceira antes de começar a aplicar o método de Gauss (o que não afeta a característica).

$$\begin{bmatrix} 1 & 1 & \alpha \\ 1 & \alpha & 1 \\ \alpha & 1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & \alpha \\ 0 & \alpha-1 & 1-\alpha \\ 0 & 1-\alpha & 1-\alpha^2 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & \alpha \\ 0 & \alpha-1 & 1-\alpha \\ 0 & 0 & 2-\alpha-\alpha^2 \end{bmatrix}$$

Se $\alpha = 1$ a segunda e terceira linha anulam-se pelo que a característica da matriz é 1. Quando $\alpha \neq 1$, a matriz tem característica pelo menos 2 e tem característica 3 exceto quando

$$2 - \alpha - \alpha^2 = 0 \Leftrightarrow \alpha = 1 \text{ ou } \alpha = -2$$

Conclui-se assim que a característica é 2 se $\alpha = -2$ e 3 se $\alpha \neq 1, -2$.

(b)

$$\begin{bmatrix} 1 & i & \alpha & 0 \\ i & 0 & 1 & -i \\ 1+i & i & \beta & \beta \end{bmatrix} \to \begin{bmatrix} 1 & i & \alpha & 0 \\ 0 & 1 & 1-i\alpha & -i \\ 0 & 1 & \beta-(1+i)\alpha & \beta \end{bmatrix} \to \begin{bmatrix} 1 & i & \alpha & 0 \\ 0 & 1 & 1-i\alpha & -i \\ 0 & 0 & \beta-\alpha-1 & \beta+i \end{bmatrix}$$

Conclui-se que a característica da matriz é 3 exceto quando

$$\begin{cases} \beta-\alpha-1=0 \\ \beta+i=0 \end{cases} \Leftrightarrow \begin{cases} \alpha=-1-i \\ \beta=-i \end{cases}$$

caso em que é igual a 2.

(12) Enunciado Um sistema homogéneo é sempre possível uma vez que tem sempre pelo menos a solução nula. Se há mais incógnitas do que equações, o número de pivots no final do método de Gauss será inferior ao número de colunas da matriz pelo que haverá pelo menos uma variável livre e portanto infinitas soluções.

Se o sistema não for homogéneo o sistema pode ser impossível. Por exemplo

$$\begin{cases} x+y+z=1 \\ x+y+z=2 \end{cases}$$

é impossível. Se for possível será necessariamente indeterminado porque, como acima, terá de haver pelo menos uma variável livre.

(13) Enunciado

(a) Por definição de solução temos

$$\begin{cases} a_{11}x_1+\ldots+a_{1n}x_n=b_1 \\ \quad\vdots \\ a_{m1}x_1+\ldots+a_{mn}x_n=b_m \end{cases} \quad \text{e} \quad \begin{cases} a_{11}x'_1+\ldots+a_{1n}x'_n=b_1 \\ \quad\vdots \\ a_{m1}x'_1+\ldots+a_{mn}x'_n=b_m \end{cases}$$

Subtraindo as equações correspondentes nos dois sistemas obtemos

$$\begin{cases} a_{11}(x_1-x'_1)+\ldots+a_{1n}(x_n-x'_n)=b_1-b_1=0 \\ \quad\vdots \\ a_{m1}(x_1-x'_1)+\ldots+a_{mn}(x_n-x'_n)=b_m-b_m=0 \end{cases}$$

pelo que $(x_1-x'_1,\ldots,x_n-x'_n)$ é uma solução do sistema homogéneo associado.

(b) Temos

$$\begin{cases} a_{11}x_1+\ldots+a_{1n}x_n=b_1 \\ \quad\vdots \\ a_{m1}x_1+\ldots+a_{mn}x_n=b_m \end{cases} \quad \text{e} \quad \begin{cases} a_{11}x'_1+\ldots+a_{1n}x'_n=0 \\ \quad\vdots \\ a_{m1}x'_1+\ldots+a_{mn}x'_n=0 \end{cases}$$

Somando as equações correspondentes dos dois sistemas obtemos

$$\begin{cases} a_{11}(x_1 + x_1') + \ldots + a_{1n}(x_n + x_n') = b_1 + 0 \\ \vdots \\ a_{m1}(x_1 + x_1') + \ldots + a_{mn}(x_n + x_n') = b_m + 0 \end{cases}$$

pelo que $(x_1 + x_1', \ldots, x_n + x_n')$ é uma solução do sistema S.

(c) Pela alínea (a), $(1, 2, -1) - (1, 3, 4) = (0, -1, -5)$ é uma solução do sistema homogéneo associado e, pela alínea (b),

$$(1, 2, -1) + (0, -1, -5) = (1, 1, -6)$$

é outra solução do sistema linear. Na realidade $(1, 2, -1) + \alpha(0, -1, -5)$ é uma solução para qualquer $\alpha \in \mathbb{R}$.

(14) Enunciado

(a)

$+$	$\underline{0}$	$\underline{1}$	$\underline{2}$
$\underline{0}$	$\underline{0}$	$\underline{1}$	$\underline{2}$
$\underline{1}$	$\underline{1}$	$\underline{2}$	$\underline{0}$
$\underline{2}$	$\underline{2}$	$\underline{0}$	$\underline{1}$

$\cdot$	$\underline{0}$	$\underline{1}$	$\underline{2}$
$\underline{0}$	$\underline{0}$	$\underline{0}$	$\underline{0}$
$\underline{1}$	$\underline{0}$	$\underline{1}$	$\underline{2}$
$\underline{2}$	$\underline{0}$	$\underline{2}$	$\underline{1}$

As propriedades comutativas de $+$ e $\cdot$ são manifestas (simetria das tabelas) assim como a existência dos elementos neutros para a soma ($\underline{0}$) e multiplicação ($\underline{1}$). A existência de simétricos para a soma e inversos de elementos não nulos é também manifesta (por exemplo o inverso de $\underline{2}$ é $\underline{2}$).

Resta verificar as propriedades associativas das duas operações e a distributividade de $+$ em relação a $\cdot$. Isto pode ser feito exaustivamente (para cada propriedade há $3^3 = 27$ casos a verificar) mas podemos também deduzi-las das propriedades correspondentes das operações com inteiros. A título de exemplo vemos a associatividade da multiplicação, sendo as outras propriedades inteiramente análogas.

Sejam m, n, p inteiros. Seja $R(x) \in \{0, 1, 2\}$ o resto da divisão por 3 do inteiro x. As operações em $\mathbb{F}_3$ realizam-se aplicando R à operação habitual com os inteiros. Para verificar a associatividade de $\cdot$ basta portanto mostrar que

$$R(mnp) = R(R(mn)p)$$

para todos os inteiros m, n, p (pois analogamente teremos também $R(mnp) = R(mR(np))$). Por definição de resto da divisão por 3, existem inteiros a, b tais que

$$mn = 3a + R(mn), \qquad R(mn)p = 3b + R(R(mn)p)$$

Então

$$mnp = (3a + R(mn))p = 3ap + 3b + R(R(mn)p) = 3(ap + b) + R(R(mn)p)$$

pelo que $R(mnp) = R(R(mn)p)$ como desejado.

(b) Podemos aplicar o método de Gauss tendo o cuidado de usar as operações descritas na alínea anterior

$$\left[\begin{array}{ccc|c} \underline{1} & \underline{0} & \underline{2} & \underline{2} \\ \underline{1} & \underline{1} & \underline{1} & \underline{0} \\ \underline{2} & \underline{1} & \underline{1} & \underline{1} \end{array}\right] \to \left[\begin{array}{ccc|c} \underline{1} & \underline{0} & \underline{2} & \underline{2} \\ \underline{0} & \underline{1} & \underline{2} & \underline{1} \\ \underline{0} & \underline{1} & \underline{0} & \underline{0} \end{array}\right] \to \left[\begin{array}{ccc|c} \underline{1} & \underline{0} & \underline{2} & \underline{2} \\ \underline{0} & \underline{1} & \underline{2} & \underline{1} \\ \underline{0} & \underline{0} & \underline{1} & \underline{2} \end{array}\right] \to \left[\begin{array}{ccc|c} \underline{1} & \underline{0} & \underline{0} & \underline{1} \\ \underline{0} & \underline{1} & \underline{0} & \underline{0} \\ \underline{0} & \underline{0} & \underline{1} & \underline{2} \end{array}\right]$$

Conclui-se que $x = \underline{1}$, $y = \underline{0}$ e $z = \underline{2}$ é a única solução do sistema.

(c) Aplicando o método de Gauss obtemos

$$\begin{bmatrix} \underline{2} & \underline{0} & \underline{1} & \underline{1} \\ \underline{1} & \underline{2} & \underline{1} & \underline{0} \\ \underline{1} & \underline{1} & \underline{0} & \underline{1} \end{bmatrix} \to \begin{bmatrix} \underline{2} & \underline{0} & \underline{1} & \underline{1} \\ \underline{0} & \underline{2} & \underline{2} & \underline{1} \\ \underline{0} & \underline{1} & \underline{1} & \underline{2} \end{bmatrix} \to \begin{bmatrix} \underline{2} & \underline{0} & \underline{1} & \underline{1} \\ \underline{0} & \underline{2} & \underline{2} & \underline{1} \\ \underline{0} & \underline{0} & \underline{0} & \underline{0} \end{bmatrix}$$

pelo que a característica desta matriz é 2. Note-se que a matriz de inteiros correspondente tem característica 3.

(15) Enunciado Se a característica da matriz é 1 todas as linhas são proporcionais, e pelo menos uma delas é não nula. Podemos pensar numa linha não nula como um número binário entre 1 e $2^n - 1$ pelo que há $2^n - 1$ possibilidades para uma linha não nula.

Como a constante de proporcionalidade entre as linhas só pode ser $\underline{0}$ ou $\underline{1}$, a linha não nula associada à matriz aparecerá ou não em cada uma das m linhas da matriz, tendo que ocorrer pelo menos numa linha. As possibilidades podem novamente ser vistas como um número binário entre 1 e $2^m - 1$. Conclui-se que há exatamente $(2^m - 1)(2^n - 1)$ matrizes com entradas em $\mathbb{F}_2$ e característica 1.

8.2. **Matrizes.**

1. Enunciado

(a) $\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

(b) $\begin{bmatrix} 1 & 2 & 3 & -1 \\ 2 & 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ -1 & 0 & 0 \\ 0 & -2 & 1 \\ 3 & 2 & -1 \end{bmatrix} = \begin{bmatrix} -4 & -6 & 4 \\ 5 & 2 & 1 \end{bmatrix}$

(c) $\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ 5 & 6 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 5 & 8 \end{bmatrix}$

(d) Calculando as primeiras potências de $\begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix}$ obtemos

$$\begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix}^2 = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} \qquad \begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix}^3 = \begin{bmatrix} 0 & 8 \\ 8 & 0 \end{bmatrix} \qquad \begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix}^4 = \begin{bmatrix} 16 & 0 \\ 0 & 16 \end{bmatrix}$$

O padrão é claro

$$\begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix}^n = \begin{cases} \begin{bmatrix} 2^n & 0 \\ 0 & 2^n \end{bmatrix} & \text{se } n \text{ é par} \\ \begin{bmatrix} 0 & 2^n \\ 2^n & 0 \end{bmatrix} & \text{se } n \text{ é ímpar} \end{cases}$$

e pode facilmente ser demonstrado por indução: supondo o resultado válido para um dado natural n temos

$$\begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix}^{n+1} = \begin{cases} \begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 2^n & 0 \\ 0 & 2^n \end{bmatrix} = \begin{bmatrix} 0 & 2^{n+1} \\ 2^{n+1} & 0 \end{bmatrix} & \text{se } n \text{ é par} \\ \begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 0 & 2^n \\ 2^n & 0 \end{bmatrix} = \begin{bmatrix} 2^{n+1} & 0 \\ 0 & 2^{n+1} \end{bmatrix} & \text{se } n \text{ é ímpar} \end{cases}$$

Em particular temos

$$\begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix}^{50} = \begin{bmatrix} 2^{50} & 0 \\ 0 & 2^{50} \end{bmatrix}$$

(e) Novamente, calculando as potências para os primeiros valores de n obtemos

$$\begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix}^2 = \begin{bmatrix} 1 & 2x \\ 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix}^3 = \begin{bmatrix} 1 & 3x \\ 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix}^4 = \begin{bmatrix} 1 & 4x \\ 0 & 1 \end{bmatrix}$$

Assumindo a hipótese de indução

$$\begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix}^n = \begin{bmatrix} 1 & nx \\ 0 & 1 \end{bmatrix}$$

temos

$$\begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix}^{n+1} = \begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix}^n \begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & nx \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & (n+1)x \\ 0 & 1 \end{bmatrix}$$

pelo que a fórmula

$$\begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix}^n = \begin{bmatrix} 1 & nx \\ 0 & 1 \end{bmatrix}$$

fica provada por indução.

2. Enunciado
   (a) A entrada obtém-se multiplicando a primeira linha de A pela segunda coluna de A logo é igual a $1 \cdot 1 + 1 \cdot 0 + 1 \cdot (-2) = -1$.
   (b) A segunda linha de AB obtém-se combinando as linhas de B de acordo com os coeficientes dados pela segunda linha de A logo é dada por

$$2 \begin{bmatrix} 0 & 1 & 1 \end{bmatrix} + 0 \cdot \begin{bmatrix} -1 & 1 & 1 \end{bmatrix} + 1 \begin{bmatrix} -1 & 0 & 3 \end{bmatrix} = \begin{bmatrix} -1 & 2 & 5 \end{bmatrix}$$

(c) A terceira coluna de BA obtém-se combinando as colunas de B de acordo com os coeficientes dados pela terceira coluna de A logo é dada por

$$1 \cdot \begin{bmatrix} 0 \\ -1 \\ -1 \end{bmatrix} + 1 \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + 3 \cdot \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \\ 8 \end{bmatrix}$$

3. Enunciado

(a)

$$\begin{bmatrix} 0 & 1 & 1 \\ 2 & 0 & 1 \\ -1 & 0 & 3 \end{bmatrix}^2 - 2\begin{bmatrix} 0 & 1 & 1 \\ 2 & 0 & 1 \\ -1 & 0 & 3 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 4 \\ -1 & 2 & 5 \\ -3 & -1 & 8 \end{bmatrix} - \begin{bmatrix} 0 & 2 & 2 \\ 4 & 0 & 2 \\ -2 & 0 & 6 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} 2 & -2 & 2 \\ -5 & 3 & 3 \\ -1 & -1 & 3 \end{bmatrix}$$

(b) O produto AD não faz sentido porque A tem 3 colunas e D só tem uma linha.

(c) O produto C^3 não faz sentido porque C não tem o mesmo número de linhas e colunas.

(d)

$$3\begin{bmatrix} 0 & 1 & 1 \\ 2 & 0 & 1 \\ -1 & 0 & 3 \end{bmatrix} + \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}\begin{bmatrix} 1 & 2 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 3 & 3 \\ 6 & 0 & 3 \\ -3 & 0 & 9 \end{bmatrix} + \begin{bmatrix} -1 & -2 & -3 \\ 1 & 2 & 3 \\ 2 & 4 & 6 \end{bmatrix} = \begin{bmatrix} -1 & 1 & 0 \\ 7 & 2 & 6 \\ -1 & 4 & 15 \end{bmatrix}$$

(e)

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & 1 \end{bmatrix}\begin{bmatrix} 0 & 1 & 1 \\ 2 & 0 & 1 \\ -1 & 0 & 3 \end{bmatrix} - 3\begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 5 \\ -3 & 0 & 2 \end{bmatrix} - \begin{bmatrix} 3 & 3 & 3 \\ 0 & -3 & 3 \end{bmatrix} = \begin{bmatrix} -2 & -2 & 2 \\ -3 & 3 & -1 \end{bmatrix}$$

(f)

$$\begin{bmatrix} 1 & 2 & 3 \end{bmatrix}\begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} - \sqrt{2}\begin{bmatrix} 1 \end{bmatrix} = \begin{bmatrix} 7 - \sqrt{2} \end{bmatrix}$$

4. Enunciado

(a)

$$\left[\begin{array}{cc|cc} 2 & 1 & 1 & 0 \\ -1 & 4 & 0 & 1 \end{array}\right] \to \left[\begin{array}{cc|cc} 2 & 1 & 1 & 0 \\ 0 & \frac{9}{2} & \frac{1}{2} & 1 \end{array}\right] \to \left[\begin{array}{cc|cc} 1 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 1 & \frac{1}{9} & \frac{2}{9} \end{array}\right] \to \left[\begin{array}{cc|cc} 1 & 0 & \frac{4}{9} & -\frac{1}{9} \\ 0 & 1 & \frac{1}{9} & \frac{2}{9} \end{array}\right]$$

A inversa é $\begin{bmatrix} \frac{4}{9} & -\frac{1}{9} \\ \frac{1}{9} & \frac{2}{9} \end{bmatrix}$.

(b)

$$\left[\begin{array}{cc|cc} 6 & -4 & 1 & 0 \\ 3 & -2 & 0 & 1 \end{array}\right] \to \left[\begin{array}{cc|cc} 6 & -4 & 1 & 0 \\ 0 & 0 & -\frac{1}{2} & 1 \end{array}\right]$$

Uma vez que a matriz tem característica 1, não é invertível.

(c)

$$\left[\begin{array}{ccc|ccc} 2 & 6 & 6 & 1 & 0 & 0 \\ 2 & 7 & 6 & 0 & 1 & 0 \\ 2 & 7 & 7 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 2 & 6 & 6 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 1 & 0 \\ 0 & 1 & 1 & -1 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 2 & 6 & 6 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & 0 & -1 & 1 \end{array}\right] \to$$

$$\left[\begin{array}{ccc|ccc} 1 & 3 & 3 & \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & 0 & -1 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 3 & 0 & \frac{1}{2} & 3 & -3 \\ 0 & 1 & 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & 0 & -1 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{7}{2} & 0 & -3 \\ 0 & 1 & 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & 0 & -1 & 1 \end{array}\right]$$

A inversa é $\begin{bmatrix} \frac{7}{2} & 0 & -3 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$.

(d)

$$\left[\begin{array}{cccc|cccc} 1 & 1 & 2 & -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 3 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{cccc|cccc} 1 & 1 & 2 & -1 & 1 & 0 & 0 & 0 \\ 0 & 1 & -3 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right] \to$$

$$\left[\begin{array}{cccc|cccc} 1 & 1 & 2 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & -3 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{cccc|cccc} 1 & 1 & 0 & 0 & 1 & 0 & -1 & 2 \\ 0 & 1 & 0 & 0 & 0 & -1 & \frac{3}{2} & -\frac{3}{2} \\ 0 & 0 & 1 & 0 & 0 & 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right] \to$$

$$\left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 1 & -\frac{5}{2} & \frac{7}{2} \\ 0 & 1 & 0 & 0 & 0 & -1 & \frac{3}{2} & -\frac{3}{2} \\ 0 & 0 & 1 & 0 & 0 & 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right]$$

A inversa é $\begin{bmatrix} 1 & 1 & -\frac{5}{2} & \frac{7}{2} \\ 0 & -1 & \frac{3}{2} & -\frac{3}{2} \\ 0 & 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 0 & 1 \end{bmatrix}$

(e) A matriz tem uma linha nula logo tem característica < 4 e portanto não é invertível.

5. Enunciado

(a) A matriz do lado direito do sinal de igual é invertível uma vez que tem característica 2 (as linhas não são proporcionais). Usando a fórmula do exercício seguinte temos

$$\begin{bmatrix} -3 & 6 \\ 4 & 5 \end{bmatrix}^{-1} = \frac{1}{-15-24} \begin{bmatrix} 5 & -6 \\ -4 & -3 \end{bmatrix} = \begin{bmatrix} -\frac{5}{39} & \frac{2}{13} \\ \frac{4}{39} & \frac{1}{13} \end{bmatrix}$$

logo a equação é equivalente a

$$A - 2I_2 = \begin{bmatrix} -\frac{5}{39} & \frac{2}{13} \\ \frac{4}{39} & \frac{1}{13} \end{bmatrix} \Leftrightarrow A = \begin{bmatrix} -\frac{5}{39} & \frac{2}{13} \\ \frac{4}{39} & \frac{1}{13} \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} \frac{73}{39} & \frac{2}{13} \\ \frac{4}{39} & \frac{27}{13} \end{bmatrix}$$

(b) A matriz $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 3 \\ 1 & 0 & 1 \end{bmatrix}$ tem característica 3 logo é invertível pelo que a equação tem solução (a matriz do lado direito do sinal de igual tem claramente característica 3). Temos

$$(A-3I_3)^{-1}\begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 3 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & -3 & 1 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \Leftrightarrow (A-3I_3)^{-1} = \begin{bmatrix} 2 & -3 & 1 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 3 \\ 1 & 0 & 1 \end{bmatrix}^{-1}$$

$$\Leftrightarrow A-3I_3 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 3 \\ 1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 2 & -3 & 1 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1}$$

Vamos portanto calcular $\begin{bmatrix} 2 & -3 & 1 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1}$:

$$\left[\begin{array}{ccc|ccc} 2 & -3 & 1 & 1 & 0 & 0 \\ 0 & 3 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & -\frac{3}{2} & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & -\frac{3}{2} & 0 & \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 1 & 0 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right] \to$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ 0 & 1 & 0 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right]$$

Pelo que

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 3 \\ 1 & 0 & 1 \end{bmatrix}\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{5}{6} & -\frac{1}{2} \\ 0 & \frac{2}{3} & 3 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix} + \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix} = \begin{bmatrix} \frac{7}{2} & \frac{5}{6} & -\frac{1}{2} \\ 0 & \frac{11}{3} & 3 \\ \frac{1}{2} & \frac{1}{2} & \frac{7}{2} \end{bmatrix}$$

(c) A matriz A tem de ser uma matriz 2×2. Escrevendo $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ obtemos

$$A^2 = I_2 \Leftrightarrow \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \Leftrightarrow \begin{bmatrix} a^2+bc & ab+bd \\ ca+cd & cb+d^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

donde resulta o sistema

$$\begin{cases} a^2+bc = 1 \\ b(a+d) = 0 \\ c(a+d) = 0 \\ cb+d^2 = 1 \end{cases}$$

Se $a+d=0 \Leftrightarrow d=-a$ a primeira e última equações são iguais pelo que ficamos apenas com a equação $a^2+bc=1$. Caso contrário, $b=c=0$ e substituindo nas outras equações obtemos o sistema

$$\begin{cases} a^2 = 1 \\ d^2 = 1 \end{cases}$$

Conclui-se que o conjunto de soluções da equação é

$$\{I_2, -I_2\} \cup \left\{ \begin{bmatrix} a & b \\ c & -a \end{bmatrix} : a^2 + bc = 1 \right\}$$

(as matrizes diagonais com entradas 1 e -1 já estão incluídas no conjunto do lado direito).

6. Enunciado

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \frac{1}{ad - bc} \begin{bmatrix} ad - bc & -ab + ba \\ cd - dc & -cb + da \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Veremos mais adiante que o cálculo anterior é suficiente para verificar que a matriz indicada é a inversa. No entanto, neste momento, usando a definição de inversa temos ainda que verificar que a multiplicação na ordem inversa produz também a identidade:

$$\frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \frac{1}{ad - bc} \begin{bmatrix} da - bc & db - bd \\ -ca + ac & -cb + ad \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

7. Enunciado
   (a) A entrada ij da matriz AB é dada pela expressão

$$\sum_{k=1}^{n} a_{ik} b_{kj} = \sum_{k=1}^{n} 2^{i+k} 2^{-k-j} = \sum_{k=1}^{n} 2^{i-j} = n 2^{i-j}$$

   (b) A entrada ij da matriz AB é dada pela expressão

$$\sum_{k=1}^{n} a_{ik} b_{kj} = \sum_{k=1}^{n} 2^{i+k} = 2^{i+1} \frac{1 - 2^n}{1 - 2} = 2^{i+1}(2^n - 1)$$

   onde usámos a fórmula para a soma de n termos de uma progressão geométrica.

8. Enunciado Uma matriz $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ comuta com $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ se e só se

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \Leftrightarrow \begin{bmatrix} b & a \\ d & c \end{bmatrix} = \begin{bmatrix} c & d \\ a & b \end{bmatrix}$$

   isto é, se e só se, $b = c$ e $a = d$.

9. Enunciado Supondo que $AB = BA$ e que A, B são invertíveis então, uma vez que o produto de matrizes invertíveis é invertível e que $(XY)^{-1} = Y^{-1}X^{-1}$ temos

$$AB = BA \Rightarrow (AB)^{-1} = (BA)^{-1} \Rightarrow B^{-1}A^{-1} = A^{-1}B^{-1}$$

   pelo que A^{-1} e B^{-1} comutam.

10. Enunciado Suponhamos que $A = B + C$ com B simétrica e C antisimétrica de forma que $b_{ij} = b_{ji}$ e $c_{ij} = -c_{ji}$ (em particular $c_{ii} = 0$ para todo o i). Então para cada $i < j$ temos
$$\begin{cases} a_{ij} = b_{ij} + c_{ij} \\ a_{ji} = b_{ji} + c_{ji} \end{cases} \Leftrightarrow \begin{cases} a_{ij} = b_{ij} + c_{ij} \\ a_{ji} = b_{ij} - c_{ij} \end{cases} \Leftrightarrow \begin{cases} b_{ij} = \frac{a_{ij}+a_{ji}}{2} \\ c_{ij} = \frac{a_{ij}-a_{ji}}{2} \end{cases}$$
Como $c_{ii} = 0$ temos necessariamente $b_{ii} = a_{ii}$ para todo o i. Isto mostra que se existir solução, ela é única, sendo B a matriz que tem entrada ij dada por $\frac{a_{ij}+a_{ji}}{2}$ para todo o i, j (mesmo que $i \geq j$) e C a única matriz anti-simétrica que tem entrada $c_{ij} = \frac{a_{ij}-a_{ji}}{2}$ para $i < j$. Por outro lado é imediato verificar que a soma destas matrizes é de facto A pelo que a solução existe e é única.

    Note-se que temos a fórmula
$$B = \frac{1}{2}(A + A^T) \qquad C = \frac{1}{2}(A - A^T)$$
e estas fórmulas são inteiramente análogas à fórmula que exprime uma função real de variável real como a soma de uma função par com uma função ímpar.

11. Enunciado
    (a) Sejam A e B matrizes triangulares superiores (portanto $a_{ij} = b_{ij} = 0$ para $i > j$). Temos a verificar que a entrada ij de AB é zero para $i > j$. Esta entrada é dada pela expressão
$$\sum_{k=1}^{n} a_{ik}b_{kj} = \sum_{k=1}^{i-1} a_{ik}b_{kj} + \sum_{k=i}^{n} a_{ik}b_{kj}$$
    Os termos do primeiro dos dois somatórios são nulos porque $i > k$, o que implica $a_{ik} = 0$ e os do segundo são zero porque $k \geq i > j$, o que implica que $b_{kj} = 0$. Conclui-se que a soma é zero conforme pretendido.
    Para as matrizes triangulares inferiores podemos argumentar analogamente ou observar que A é triangular inferior se e só se A^T é triangular superior. Dadas matrizes triangulares inferiores A e B temos que $B^T A^T = (AB)^T$ é triangular superior, logo AB é triangular inferior.
    (b) Um matriz triangular é invertível se e só se as entradas na diagonal são todas não nulas. Basta considerar o caso triangular superior uma vez que uma matriz é invertível se e só se a sua transposta é invertível. Se uma matriz triangular superior tem todas as entradas na diagonal não nulas, então está em escada de linhas, tem característica máxima e é portanto invertível. Se alguma das entradas na diagonal é zero então a coluna correspondente ao primeiro zero na diagonal não conterá um pivot, pelo que a característica da matriz não é máxima e a matriz não é portanto invertível.
    A inversa de uma matriz triangular é triangular do mesmo tipo. Novamente basta considerar o caso triangular superior. Ao aplicar o método de Gauss-Jordan para calcular a inversa, as operações sobre as linhas que aplicamos à matriz identidade só afetam as entradas diagonais ou acima da diagonal pelo que a matriz resultante é triangular superior. O mesmo dito matricialmente: existem matrizes triangulares

superiores $E_1, \ldots, E_k$ (correspondentes às operações do método de Gauss-Jordan) tais que $E_k \cdots E_1 A = I \Rightarrow A = E_1^{-1} \cdots E_k^{-1}$ é um produto de matrizes triangulares superiores, logo triangular superior.

(c) Se não há troca de linhas, durante a aplicação do método de Gauss multiplicamos A sucessivamente por matrizes elementares $E_1, E_2, \ldots, E_k$ correspondentes a operações sobre as linhas da forma $L_i - \alpha L_j$ com $i > j$. Estas matrizes são triangulares inferiores e têm todas as entradas na diagonal iguais a 1; isto é, são da forma

$$E_i = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & -\alpha & 1 & 0 \\ 0 & \cdots & & 1 \end{bmatrix}$$

com α na linha i e coluna j. No final do método de Gauss obtemos uma matriz triangular superior U. É portanto satisfeita uma equação matricial da forma

$$E_k \cdots E_2 E_1 A = U$$

As matrizes E_i são invertíveis (a inversa obtém-se trocando o sinal da entrada não nula fora da diagonal) logo a equação anterior é equivalente a

$$A = E_1^{-1} E_2^{-1} \cdots E_k^{-1} U$$

Pela alínea (a) a matriz $L = E_1^{-1} E_2^{-1} \cdots E_k^{-1}$ é triangular inferior. Tem 1s na diagonal porque quando multiplicamos duas matrizes triangulares (do mesmo tipo) as entradas diagonais do produto obtêm-se multiplicando as entradas diagonais correspondentes dos fatores. Vemos assim que A se pode escrever na forma LU como pretendido.

Suponhamos que $A = L'U'$ é outra fatorização da mesma forma. Notamos que pela alínea (b) L e L' são invertíveis e portanto U e U' também são invertíveis. Então

$$LU = L'U' \Leftrightarrow L'^{-1} L = U' U^{-1}$$

Pela alínea (b) L'^{-1} é triangular inferior e U^{-1} é triangular superior. Temos assim uma matriz triangular inferior com 1s na diagonal do lado esquerdo e uma matriz triangular superior do lado direito. Isto só pode acontecer se todas as entradas fora da diagonal forem 0 (e as da diagonal têm que ser 1 claro) isto é, se

$$L'^{-1} L = I = U' U^{-1} \Leftrightarrow L = L' \text{ e } U = U'$$

Conclui-se que a decomposição $A = LU$ é única.

12. Enunciado Se A é invertível então tem de haver um pivot em cada coluna no final da aplicação do método de Gauss. Em particular, o mesmo tem que acontecer a B (porque ao aplicarmos o método de Gauss a A estamos também a fazê-lo a B) logo B é invertível. Sendo B invertível, após colocar as linhas correspondentes a B em escada de linhas resta-nos aplicar o método de Gauss a D pelo que se A é invertível D terá de ter um pivot em cada coluna e ser portanto invertível.

Queremos achar uma matriz T tal que $AT = TA = I$. Podemos decompor T em blocos do mesmo tipo que A e efetuar o produto das matrizes por blocos: escrevendo

$$T = \begin{bmatrix} X & Y \\ Z & W \end{bmatrix}$$

temos

$$AT = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix} \begin{bmatrix} X & Y \\ Z & W \end{bmatrix} = \begin{bmatrix} BX + CZ & BY + CW \\ DZ & DW \end{bmatrix}$$

e

$$TA = \begin{bmatrix} X & Y \\ Z & W \end{bmatrix} \begin{bmatrix} B & C \\ 0 & D \end{bmatrix} = \begin{bmatrix} XB & XC + YD \\ ZB & ZC + WD \end{bmatrix}$$

As equações $AT = I = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$ e $TA = I$ traduzem-se nos sistemas matriciais:

$$\begin{cases} BX + CZ = I \\ BY + CW = 0 \\ DZ = 0 \\ DW = I \end{cases} \qquad \begin{cases} XB = I \\ XC + YD = 0 \\ ZB = 0 \\ ZC + WD = I \end{cases}$$

Uma vez que D é invertível, o primeiro sistema diz-nos que $Z = 0$ e $W = D^{-1}$. A primeira equação fica então $BX = I \Leftrightarrow X = B^{-1}$ e a segunda $BY + CW = 0 \Leftrightarrow Y = -B^{-1}CD^{-1}$. É imediato verificar que estes valores para X, Y, Z, W satisfazem também o segundo sistema logo A é invertível com inversa dada pela expressão:

$$\begin{bmatrix} B^{-1} & -B^{-1}CD^{-1} \\ 0 & D^{-1} \end{bmatrix} \tag{74}$$

Reciprocamente, se B e D são invertíveis verifica-se imediatamente que a matriz dada pela expressão (74) é uma inversa para A logo A é invertível.

13. Enunciado Se $A = \lambda I_n$ então sendo X uma matriz $n \times n$ qualquer temos $AX = \lambda I_n X = \lambda X = X(\lambda I_n) = XA$ e portanto A comuta com todas as matrizes $n \times n$.

Se A comuta com todas as matrizes então em particular para todos os $1 \leq i, j \leq n$ temos

$$AE_{ij} = E_{ij}A \Leftrightarrow \begin{bmatrix} 0 & \cdots & 0 & a_{1i} & 0 & \cdots & 0 \\ \vdots & & \vdots & a_{2i} & \vdots & & \vdots \\ \vdots & & & \vdots & & & \vdots \\ 0 & \cdots & 0 & a_{ni} & 0 & \cdots & 0 \end{bmatrix} = \begin{bmatrix} 0 & \cdots & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & & 0 \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ 0 & \cdots & & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & & 0 \end{bmatrix}$$

onde do lado esquerdo do sinal de igual as entradas não nulas ocorrem na coluna j e do lado direito na linha i. A igualdade significa que $a_{ki} = 0$ para $k \neq i$ (logo A é diagonal) e que a_{ii} (a entrada ij da matriz da esquerda) é igual a a_{jj} (a entrada ij da matriz da direita). Isto significa que todas as entradas diagonais de A são iguais, ou seja que A se pod escrever na forma λI_n.

14. Enunciado Quando k, l são positivos é fácil mostrar por indução em l que $A^{k+l} = A^k A^l$. Se um dos expoentes for nulo então, como $A^0 = I_n$ é evidente que a fórmula se verifica (por exemplo $A^{k+0} = A^k = A^k I_n = A^k A^0$). O caso em que ambos os expoentes são negativos é uma consequência da definição e do caso em que ambos são positivos: dados $k, l < 0$ temos, aplicando a definição e depois o caso em que ambos os expoentes são positivos,
$$A^{k+l} = (A^{-1})^{-k-l} = (A^{-1})^{-k}(A^{-1})^{-l} = A^k A^l$$
Resta considerar o caso em que os expoentes têm sinais contrários. Fazemos o caso em que $k > 0$ e $l < 0$ sendo o caso oposto inteiramente análogo. Consideramos três sub-casos:
   - $k = -l$: Então $A^{k+l} = A^0 = A^k(A^{-1})^k = A^k A^l$
   - $k > -l$ (portanto $k + l > 0$): Então $A^{k+l} = A^{k+l}A^{-l}(A^{-1})^{-l} = A^k(A^{-1})^{-l} = A^k A^l$.
   - $k < -l$ (portanto $k + l < 0$): Então $A^{k+l} = (A^{-1})^{-k-l} = A^k(A^{-1})^k(A^{-1})^{-k-l} = A^k(A^{-1})^{-l} = A^k A^l$
15. Enunciado
   (a) Suponhamos que $AB = I_m$ e $CA = I_n$. Note-se que tanto B como C são necessariamente matrizes $n \times m$. Então pela associatividade do produto de matrizes temos
   $$C = CI_m = C(AB) = (CA)B = I_nB = B$$
   (b) É claramente impossível achar um exemplo com matrizes 1×1. O caso mais simples que se segue é o de matrizes 2×1 (ou 1×2). Considerando as equações
   $$\begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \end{bmatrix} \Leftrightarrow ax + by = 1$$
   vemos que uma matriz 1×2 não nula tem um inverso à direita (por exemplo, os inversos à direita de $\begin{bmatrix} 1 & 0 \end{bmatrix}$ são as matrizes da forma $\begin{bmatrix} 1 \\ x \end{bmatrix}$) e que as matrizes 2×1 não nulas têm um inverso à esquerda. No entanto uma matriz 1×2 nunca tem um inverso à esquerda:
   $$\begin{bmatrix} a \\ b \end{bmatrix} \begin{bmatrix} x & y \end{bmatrix} = \begin{bmatrix} ax & ay \\ bx & by \end{bmatrix}$$
   Para que este produto seja a identidade temos $ay = 0$ e $bx = 0$ e em conjunto estas condições implicam que a matriz produto não pode ser a identidade. Por exemplo, se $a = 0$ e $x = 0$ todas as entradas exceto possivelmente a 22 anulam-se. Analogamente, uma matriz 2×1 não pode ter um inverso à direita.
   (c) Mostramos que cada afirmação implica as restantes:
   (i) $\Rightarrow$ (ii): Seja B um inverso à direita para A. Então tomando $X = BC$ temos $AX = A(BC) = (AB)C = I_mC = C$.
   (ii) $\Rightarrow$ (iii): Provamos a afirmação contra-recíproca. Suponhamos que a característica de A é menor que m. Então aplicando o método de Gauss a A obtemos uma

matriz cuja última linha é nula. Isso significa que existe uma matriz invertível L (o produto das matrizes elementares correspondentes às operações sobre as linhas efetuadas durante a aplicação do método) tal que LA tem última linha nula. Portanto a equação

$$LAX = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

é impossível. Esta equação é equivalente a

$$AX = L^{-1} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

logo existe uma matriz $m \times 1$ C tal que a equação $AX = C$ não tem solução.

(iii) $\Rightarrow$ (i): Uma vez que a característica de A é m, pelo método de Gauss, todo o sistema $AX = C$ tem solução. Se E_i for a i-ésima coluna de I_m e X_i uma solução da equação $AX = E_i$ então a matriz $n \times m$ B que tem por i-ésima coluna X_i satisfaz a equação $AB = I_m$ e é portanto uma inversa à direita para A.

(d) Não, porque a característica não pode ser maior do que o número de colunas, que é n, e é portanto inferior a m.

(e) Se A tem inversa à direita, pela alínea anterior $m \leq n$. Se A tem inversa à esquerda B, então B, que é uma matriz $n \times m$ tem uma inversa à direita (nomeadamente A) e portanto, pela alínea anterior temos $n \leq m$. Conclui-se que $n = m$. Pela alínea (a) as inversas à direita e esquerda são iguais (quando existem) logo quando uma matriz tem inversa à esquerda e à direita é invertível.

(f) Suponhamos que A é uma matriz $n \times n$ com uma inversa à direita. Então a característica de A é n pela alínea (c) e como tal a matriz é invertível. Se A tem uma inversa à esquerda B, então B tem uma inversa à direita e portanto B é invertível. Como $BA = I$ temos $A = B^{-1}$ logo A é invertível (com inversa B).

8.3. **Espaços Vetoriais.**

1. Enunciado Nesta resolução vamos usar a numeração dos axiomas usada nos apontamentos

   (i) Não é um espaço vetorial. Os axiomas (i) e (ii) não se verificam: Por exemplo

   $$(1,0) \oplus (0,1) = (1,-3) \neq (-2,1) = (0,1) \oplus (1,0)$$

   e

   $$(1,0) \oplus ((1,0) \oplus (0,1)) = (1,0) \oplus (1,-3) = (-1,9)$$

   que é diferente de

   $$((1,0) \oplus (1,0)) \oplus (0,1) = (-1,0) \oplus (0,1) = (-1,-3)$$

Os restantes axiomas verificam-se (embora os axiomas (iii) e (iv) - a existência de elemento neutro e simétrico - teriam uma formulação diferente para uma operação não comutativa como esta e nessa formulação natural também não seriam verificados). De facto as igualdades

$$(x,y) \oplus (0,0) = (x,y) \qquad (x,y) \oplus (\tfrac{x}{2}, \tfrac{y}{3}) = (0,0)$$

mostram que (iii) e (iv) se verificam. Os axiomas (v),(vi) verificam-se porque envolvem apenas a multiplicação por escalar, que é a usual. O axioma (vii) também se verifica:

$$\begin{aligned} \alpha((x,y) \oplus (x',y')) &= \alpha(x-2x', y-3y') \\ &= (\alpha x - 2\alpha x', \alpha y - 3\alpha y') \\ &= (\alpha x, \alpha y) \oplus (\alpha x', \alpha y') \\ &= \alpha(x,y) \oplus \alpha(x',y') \end{aligned}$$

mas o axioma (viii) não:

$$2 \cdot (1,0) = (1+1) \cdot (1,0) \neq 1 \cdot (1,0) \oplus 1 \cdot (1,0) = (-1,0)$$

(ii) Trata-se de um espaço vetorial. Os axiomas (i)-(iv) verificam-se porque a operação de soma é a usual em $\mathbb{C}^n$. O elemento neutro para este produto por escalar continua a ser $1 \in \mathbb{C}$ uma vez que $\overline{1} = 1$ logo o axioma (v) é verificado. Quanto ao (vi):

$$\begin{aligned} \alpha \odot (\beta \odot (z_1, \ldots, z_n)) &= \alpha \odot (\overline{\beta} z_1, \ldots, \overline{\beta} z_n) \\ &= (\overline{\alpha}\overline{\beta} z_1, \ldots, \overline{\alpha}\overline{\beta} z_n) \\ &= (\overline{\alpha\beta} z_1, \ldots, \overline{\alpha\beta} z_n) \\ &= (\alpha\beta) \cdot (z_1, \ldots, z_n) \end{aligned}$$

A verificação dos axiomas (vii) e (viii) é inteiramente análoga; (vii) usa apenas a distributividade da multiplicação usual dos números complexos em relação à soma e (viii) o facto que $\overline{\alpha + \beta} = \overline{\alpha} + \overline{\beta}$.

(iii) Não é um espaço vetorial. Os axiomas (i)-(iv) verificam-se porque a soma é a usual. O axioma (v) não se verifica pois para $v \in \mathbb{R}^n$ não nulo temos $1 \odot v = 0 \neq v$. Os axioma (vi)-(viii) verificam-se trivialmente já que todos os produtos por escalar têm como resultado o vetor nulo.

(iv) É um espaço vetorial. A operação $\oplus$ é comutativa e associativa porque é o produto habitual de números reais positivos, que tem estas propriedades. O elemento neutro da operação $\oplus$ é 1 (o elemento neutro para o produto usual, que é um vetor de V) e o simétrico de x para a operação $\oplus$ existe, sendo dado por $\frac{1}{x}$ (o inverso para o produto de números reais, que é um real positivo porque x também é). O axioma (v) é também verificado já que $1 \odot x = x^1 = x$ e os axiomas (vi)-(viii) são verificados porque são as regras habituais da exponenciação. Por exemplo, no caso de (vii) temos

$$\alpha \odot (x \oplus y) = (x \oplus y)^\alpha = (xy)^\alpha = x^\alpha y^\alpha = (\alpha \odot x)(\alpha \odot y) = (\alpha \odot x) \oplus (\alpha \odot y)$$

Nota: O logaritmo natural dá uma bijeção $\log\colon \mathbb{R}^+ \to \mathbb{R}$. Estas operações em $\mathbb{R}^+$ são simplesmente as operações usuais de soma e produto em $\mathbb{R}$ traduzidas por meio deste "dicionário"dado pelo logaritmo. Este espaço vetorial é portanto essencialmente equivalente a $\mathbb{R}^1$ (mais tarde introduziremos um termo para esta noção de equivalência - isomorfismo).

2. Enunciado
   (i) Sim, a função 0 é contínua logo $W \neq \emptyset$ e uma vez que a soma de funções contínuas é contínua e o produto de uma função constante por uma função contínua é uma função contínua, W é fechado para a soma e para o produto por escalar.
   (ii) W não contém o vetor nulo (que é a função constante igual a zero) uma vez que esta não satisfaz a condição $f''(0) = 1$. Logo W não é fechado para o produto por escalar e portanto não é um subespaço vetorial de V.
   (iii) W não é um subespaço vetorial porque não é fechado para a soma. Por exemplo $t^3 + t^2$ e $-t^3$ pertencem a W mas a sua soma t^2 não pertence.
   (iv) Sim, W é um subespaço vetorial de V. O polinómio nulo satisfaz $p''(1) = p(3) = 0$ logo $W \neq \emptyset$. Se $p, q \in W$ e $\alpha \in \mathbb{R}$ temos

$$(p+q)''(1) = p''(1) + q''(1) = p(3) + q(3) = (p+q)(3)$$

$$(\alpha p)''(1) = \alpha p''(1) = \alpha p(3) = (\alpha p)(3)$$

   logo W é fechado para a soma e para o produto por escalar.
   (v) A matriz nula é simétrica logo $W \neq \emptyset$. Dadas $A, B \in W$ e $\alpha \in \mathbb{R}$ temos

$$(A+B)^T = A^T + B^T = A + B \qquad (\alpha A)^T = \alpha A^T = \alpha A$$

   logo W é fechado para a soma e para o produto por escalar e é portanto um subespaço vetorial de V.
   (vi) W não é um subespaço vetorial porque não é fechado para o produto por escalar: se $A \in W$ é uma matriz não nula qualquer então $(-1) \cdot A$ tem alguma entrada negativa e portanto não pertence a W.

3. Enunciado
   (a) O vetor v pertence a $L(S)$ se existem $\alpha, \beta, \gamma \in \mathbb{R}$ tais que

$$(1,2,1) = \alpha(1,0,1) + \beta(2,1,3) + \gamma(-1,-1,-2)$$

   A equação anterior é equivalente ao sistema

$$\left[\begin{array}{ccc|c} 1 & 2 & -1 & 1 \\ 0 & 1 & -1 & 2 \\ 1 & 3 & -2 & 1 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 2 & -1 & 1 \\ 0 & 1 & -1 & 2 \\ 0 & 1 & -1 & 0 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 2 & -1 & 1 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & -2 \end{array}\right]$$

   Uma vez que o sistema é impossível conclui-se que $v \notin L(S)$.
   (b) O vetor v pertence a $L(S)$ se existirem $\alpha, \beta, \gamma \in \mathbb{R}$ tais que

$$\begin{bmatrix} 2 & 5 \\ 4 & 1 \end{bmatrix} = \alpha \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix} + \beta \begin{bmatrix} -1 & 2 \\ 0 & 3 \end{bmatrix} + \gamma \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

e a equação anterior traduz-se no sistema

$$\left[\begin{array}{ccc|c}1 & -1 & 1 & 2\\ 1 & 2 & 1 & 5\\ 2 & 0 & 1 & 4\\ 0 & 3 & -1 & 1\end{array}\right] \to \left[\begin{array}{ccc|c}1 & -1 & 1 & 2\\ 0 & 3 & 0 & 3\\ 0 & 2 & -1 & 0\\ 0 & 3 & -1 & 1\end{array}\right] \to \left[\begin{array}{ccc|c}1 & -1 & 1 & 2\\ 0 & 3 & 0 & 3\\ 0 & 0 & -1 & -2\\ 0 & 0 & -1 & -2\end{array}\right] \to$$

$$\left[\begin{array}{ccc|c}1 & -1 & 1 & 2\\ 0 & 3 & 0 & 3\\ 0 & 0 & -1 & -2\\ 0 & 0 & 0 & 0\end{array}\right]$$

Como o sistema é possível, vemos que $v \in L(S)$, sendo os coeficientes da combinação linear $\gamma = 2, \beta = 1, \alpha = 2 + 1 - 2 = 1$.

(c) Uma vez que

$$\lim_{x \to +\infty} \frac{\alpha x + \beta \operatorname{sen} x + \gamma x^3 + \delta \cos x}{e^x} = 0$$

para quaisquer $\alpha, \beta, \gamma, \delta$, a função e^x não pode ser expressa como combinação linear de elementos de S; se tal fosse o caso, existiria uma escolha dos coeficientes para a qual o limite daria 1. Conclui-se que $v \notin L(S)$.

(d) Igualando os coeficientes na equação

$$1 + x - 2x^2 = \alpha(1 - x + x^2) + \beta(x^2 - 1) + \gamma(x + 2) + \delta(x^2 + x)$$

obtemos o sistema

$$\left[\begin{array}{cccc|c}1 & -1 & 2 & 0 & 1\\ -1 & 0 & 1 & 1 & 1\\ 1 & 1 & 0 & 1 & -2\end{array}\right] \to \left[\begin{array}{cccc|c}1 & -1 & 2 & 0 & 1\\ 0 & -1 & 3 & 1 & 2\\ 0 & 2 & -2 & 1 & -3\end{array}\right] \to \left[\begin{array}{cccc|c}1 & -1 & 2 & 0 & 1\\ 0 & -1 & 3 & 1 & 2\\ 0 & 0 & 4 & 3 & 1\end{array}\right]$$

Este sistema é possível pelo que $v \in L(S)$. Há infinitas possibilidades para os coeficientes uma vez que o sistema é indeterminado. Tomando por exemplo $\delta = 0$ obtemos $\gamma = \frac{1}{4}$, $\beta = -2 + \frac{3}{4} = -\frac{5}{4}$, $\alpha = 1 - \frac{5}{4} - \frac{1}{2} = -\frac{3}{4}$.

4. Enunciado

(a) O núcleo é o conjunto das soluções do sistema homogéneo associado a A. Temos

$$\begin{bmatrix}1 & 1 & 1\\ -1 & -2 & 3\end{bmatrix} \to \begin{bmatrix}1 & 1 & 1\\ 0 & -1 & 4\end{bmatrix} \to \begin{bmatrix}1 & 0 & 5\\ 0 & -1 & 4\end{bmatrix}$$

logo

$$N(A) = \{(-5z, 4z, z) \colon z \in \mathbb{R}\} = \{z(-5, 4, 1) \colon z \in \mathbb{R}\}$$

Podemos portanto tomar, por exemplo, $S = \{(-5, 4, 1)\}$.

(b) Analogamente à alínea anterior temos

$$\begin{bmatrix}1 & 1 & 0 & 2\\ -1 & 0 & 1 & 1\\ -1 & 1 & -2 & 0\end{bmatrix} \to \begin{bmatrix}1 & 1 & 0 & 2\\ 0 & 1 & 1 & 3\\ 0 & 2 & -2 & 2\end{bmatrix} \to \begin{bmatrix}1 & 1 & 0 & 2\\ 0 & 1 & 1 & 3\\ 0 & 0 & -4 & -4\end{bmatrix} \to \begin{bmatrix}1 & 1 & 0 & 2\\ 0 & 1 & 1 & 3\\ 0 & 0 & 1 & 1\end{bmatrix} \to$$

$$\begin{bmatrix}1 & 1 & 0 & 2\\ 0 & 1 & 0 & 2\\ 0 & 0 & 1 & 1\end{bmatrix} \to \begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 2\\ 0 & 0 & 1 & 1\end{bmatrix}$$

logo

$$N(A) = \{(0, -2w, -w, w) \colon w \in \mathbb{R}\} = L(S) \text{ com } S = \{(0, -2, -1, 1)\}$$

5. EnunciadoDeterminar uma matriz cujo núcleo é um dado subespaço é o mesmo do que encontrar um sistema homogéneo cujas soluções sejam os vetores do subespaço, ou seja, é encontrar equações cartesianas que definam esse subespaço.
   (i) Procuramos condições sobre (x, y, z) que são satisfeitas pelos vetores $\alpha(1, 0, 2) + \beta(-1, 1, 0)$. Podemos obtê-las resolvendo o sistema

$$\begin{cases} x = \alpha - \beta \\ y = \beta \\ z = 2\alpha \end{cases} \Leftrightarrow \begin{cases} x + y - \frac{z}{2} = 0 \\ \beta = y \\ \alpha = \frac{z}{2} \end{cases}$$

   A equivalência mostra-nos que os vetores de $L(\{(1, 0, 2), (-1, 1, 0)\})$ são precisamente as soluções da equação homogénea $x + y - \frac{z}{2} = 0$, que formam o núcleo de (por exemplo)

$$A = \begin{bmatrix}1 & 1 & -\frac{1}{2}\end{bmatrix}$$

   (ii) Vamos usar o método de Gauss para resolver o sistema equivalente à equação

$$(x, y, z, w) = \alpha(1, 1, 1, 1) + \beta(-1, 1, 0, 1)$$

   Temos

$$\left[\begin{array}{cc|c}1 & -1 & x\\ 1 & 1 & y\\ 1 & 0 & z\\ 1 & 1 & w\end{array}\right] \to \left[\begin{array}{cc|c}1 & -1 & x\\ 0 & 2 & y - x\\ 0 & 1 & z - x\\ 0 & 2 & w - x\end{array}\right] \to \left[\begin{array}{cc|c}1 & -1 & x\\ 0 & 2 & y - x\\ 0 & 0 & z - \frac{1}{2}y - \frac{1}{2}x\\ 0 & 0 & w - y\end{array}\right]$$

   logo os vetores (x, y, z, w) que pertencem ao plano dado (isto é, para os quais o sistema anterior é possível) são exatamente os que satisfazem as condições

$$\begin{cases} z - \frac{1}{2}y - \frac{1}{2}x = 0 \\ w - y = 0 \end{cases}$$

   Podemos portanto tomar, por exemplo,

$$A = \begin{bmatrix}-\frac{1}{2} & -\frac{1}{2} & 1 & 0\\ 0 & -1 & 0 & 1\end{bmatrix}$$

6. Enunciado
   (i) Os vetores de $L(S)$ são os vetores (x, y, z, w) para os quais existem $\alpha, \beta \in \mathbb{R}$ tais que

$$(x, y, z, w) = \alpha(1, 0, 2, 1) + \beta(-1, 3, 2, 0)$$

Pensando em α, β como incógnitas vemos que estes vetores são aqueles para os quais o seguinte sistema tem solução

$$\begin{bmatrix} 1 & -1 & | & x \\ 0 & 3 & | & y \\ 2 & 2 & | & z \\ 1 & 0 & | & w \end{bmatrix} \to \begin{bmatrix} 1 & -1 & | & x \\ 0 & 3 & | & y \\ 0 & 4 & | & z-2x \\ 0 & 1 & | & w-x \end{bmatrix} \to \begin{bmatrix} 1 & -1 & | & x \\ 0 & 1 & | & w-x \\ 0 & 3 & | & y \\ 0 & 4 & | & z-2x \end{bmatrix} \to$$

$$\begin{bmatrix} 1 & -1 & | & x \\ 0 & 1 & | & w-x \\ 0 & 0 & | & y-3w+3x \\ 0 & 0 & | & z+2x-4w \end{bmatrix}$$

Logo $L(S)$ é descrito pelo sistema de equações

$$\begin{cases} 3x+y-3w=0 \\ 2x+z-4w=0 \end{cases}$$

(ii) Um polinómio $a+bx+cx^2+dx^3+ex^4$ pertence a $L(S)$ se e só se o seguinte sistema homogéneo tem solução:

$$\begin{bmatrix} 1 & 0 & | & a \\ -1 & 1 & | & b \\ 0 & 1 & | & c \\ 1 & 0 & | & d \\ 0 & 0 & | & e \end{bmatrix} \to \begin{bmatrix} 1 & 0 & | & a \\ 0 & 1 & | & b+a \\ 0 & 1 & | & c \\ 0 & 0 & | & d-a \\ 0 & 0 & | & e \end{bmatrix} \to \begin{bmatrix} 1 & 0 & | & a \\ 0 & 1 & | & b+a \\ 0 & 0 & | & c-b-a \\ 0 & 0 & | & d-a \\ 0 & 0 & | & e \end{bmatrix}$$

logo

$$L(S) = \{a+bx+cx^2+dx^3+ex^4 \colon \begin{cases} c=a+b \\ d=a \\ e=0 \end{cases} \}$$

(iii) Analogamente às alíneas anteriores, é suficiente ver quando o sistema

$$\begin{bmatrix} 1 & 2 & | & x \\ 1 & 0 & | & y \\ 1 & 1 & | & z \\ -1 & -1 & | & w \\ -1 & -2 & | & u \\ 3 & 3 & | & v \end{bmatrix} \to \begin{bmatrix} 1 & 2 & | & x \\ 0 & -2 & | & y-x \\ 0 & -1 & | & z-x \\ 0 & 1 & | & w+x \\ 0 & 0 & | & u+x \\ 0 & -3 & | & v-3x \end{bmatrix} \to \begin{bmatrix} 1 & 2 & | & x \\ 0 & 1 & | & w+x \\ 0 & 0 & | & y+x+2w \\ 0 & 0 & | & z+w \\ 0 & 0 & | & u+x \\ 0 & 0 & | & v+3w \end{bmatrix}$$

donde

$$L(S) = \{ \begin{bmatrix} x & y & z \\ w & u & v \end{bmatrix} : \begin{cases} x+y+2w=0 \\ z+w=0 \\ x+u=0 \\ v+3w=0 \end{cases} \}$$

7. Enunciado Suponhamos que $\alpha e^x + \beta e^{2x} + \gamma e^{3x} = 0$. Então

$$\lim_{x\to+\infty} \frac{\alpha e^x + \beta e^{2x} + \gamma e^{3x}}{e^{3x}} = 0 \Leftrightarrow \lim_{x\to+\infty} \alpha e^{-2x} + \beta e^{-x} + \gamma = 0 \Leftrightarrow \gamma = 0$$

Portanto $\alpha e^x + \beta e^{2x} = 0$. Da mesma forma $\lim_{x\to+\infty} \frac{\alpha e^x+\beta e^{2x}}{e^{2x}} = 0 \Rightarrow \beta = 0$ e então α é também necessariamente 0.

Alternativamente podíamos substituir três valores de x distintos na igualdade $\alpha e^x + \beta e^{2x} + \gamma e^{3x} = 0$ e resolver o sistema resultante para α, β, γ. A única solução será $\alpha = \beta = \gamma = 0$.

8. Enunciado
   (i) Pelo axioma (iv) existe um elemento $t \in V$ tal que $t + u = 0$. Somando t a ambos os membros da equação $u + v = u + w$ obtemos

$$t + (u + v) = t + (u + w) \Leftrightarrow (t + u) + v = (t + u) + w \Leftrightarrow 0 + v = 0 + w \Leftrightarrow v = w$$

   onde a primeira equivalência é válida pela associatividade da soma (axioma (i)), a segunda pela definição de t e a terceira pela comutatividade e definição de 0 (axiomas (ii) e (iii)).

   (ii) Pelo axioma (viii) temos $0v = (0 + 0)v = 0v + 0v$. Pelo axioma (iii) podemos escrever esta igualdade na forma

$$0v + 0 = 0v + 0v$$

   e então pela lei do corte (a propriedade (i) demonstrada acima) temos $0 = 0v$ como requerido.

   (iii) Pelo axioma (vii) temos $\alpha \cdot 0 = \alpha \cdot (0 + 0) = \alpha \cdot 0 + \alpha \cdot 0$. Aplicando a lei do corte à equação

$$\alpha \cdot 0 + 0 = \alpha \cdot 0 + \alpha \cdot 0$$

   vemos que $\alpha \cdot 0 = 0$.

   (iv) Sejam w, w' simétricos de v. Então

$$w = w + 0 = w + (v + w') = (w + v) + w' = (v + w) + w' = 0 + w' = w'$$

   (v) Queremos ver que $(-1) \cdot v$ é o simétrico de v. Pelo axioma (v) temos $1 \cdot v = v$. Logo

$$v + (-1) \cdot v = 1 \cdot v + (-1) \cdot v = (1 + (-1)) \cdot v = 0 \cdot v = 0$$

   onde na segunda igualdade usámos o axioma (viii) e na quarta a alínea (ii).

   (vi) A afirmação é logicamente equivalente a: Se $\alpha \neq 0$ e $\alpha v = 0$ então $v = 0$. Se $\alpha \neq 0$ temos

$$\tfrac{1}{\alpha}(\alpha v) = (\tfrac{1}{\alpha}\alpha)v = 1v = v$$

   onde usámos o axioma (vi) e depois o axioma (v). Mas se $(\alpha v) = 0$ temos que o termo à esquerda nas igualdades acima é

$$\tfrac{1}{\alpha}0 = 0$$

   (pela propriedade (iv) acima) logo $v = 0$.

(vii) Temos

$$\alpha v = \beta v \Rightarrow \alpha v + (-(\beta v)) = \beta v + (-(\beta v)) = 0$$

Como

$$-(\beta v) = (-1)(\beta v) = (-\beta)v$$

onde na primeira igualdade usámos a alínea (v) e na segunda o axioma (vii), obtemos

$$\alpha v + (-\beta)v = 0 \Rightarrow (\alpha - \beta)v = 0$$

(pelo axioma (viii)). Finalmente, dado que $v \neq 0$, a alínea (vi) diz-nos que $\alpha - \beta = 0$, isto é, $\alpha = \beta$.

9. EnunciadoVamos resolver este exercício sem usar a noção de dimensão (que tornaria a resolução mais rápida) uma vez que este exercício é normalmente atribuído antes de este conceito ter sido coberto nas aulas.

Uma possibilidade para subespaço é o subespaço nulo $\{(0,0)\}$. Se $W \subset \mathbb{F}_p^2$ é um subespaço e contém um vetor $v \neq 0$ então contém todos os seus múltiplos escalares e de facto

$$L(v) = \{\lambda v \colon \lambda \in \mathbb{F}_p\}$$

é claramente um subespaço de $\mathbb{F}_p^2$ (a "reta"gerada pelo vetor v no "plano"$\mathbb{F}_p^2$). Uma possibilidade é a "reta vertical"$\{(0,y) \colon y \in \mathbb{F}_p\}$ que é gerada por qualquer vetor $(0,b)$ com $b \neq 0$. Se $L(v)$ contém um vetor (a,b) com primeira coordenada não nula então contém um vetor $\frac{1}{a}(a,b)$ com primeira coordenada igual a $\underline{1}$ e dois vetores desta forma distintos $(\underline{1},\lambda) \neq (\underline{1},\mu)$ não são colineares e portanto geram retas distintas.

Conclui-se que há $p+1$ subespaços da forma $L(v)$: a reta vertical com "declive infinito"e uma reta $L((\underline{1},\lambda))$ para cada possível "declive"$\lambda \in \mathbb{F}_p$.

Finalmente se W é um subespaço que contém um vetor não nulo v mas não é da forma $L(v)$ então $W = \mathbb{F}_p^2$: dado $v \in W \setminus \{0\}$, o subespaço W tem que conter um vetor v' que não é múltiplo de v. Se $v = (\underline{0},\lambda)$ então $v' = (\underline{1},\mu)$ para algum μ. Se $v = (\underline{1},\lambda)$ para algum λ então ou $v' = (\underline{0},\mu)$ com $\mu \neq 0$ ou, multiplicando por um escalar podemos assumir que $v' = (\underline{1},\mu)$ com $\mu \neq \lambda$ de forma que $v - v' = (\underline{0},\lambda-\mu) \in W$ com $\lambda - \mu \neq 0$. Claramente dois vetores da forma $(\underline{1},\alpha)$ e $(\underline{0},\beta)$ com $\beta \neq 0$ geram $\mathbb{F}_p^2$ donde se conclui que $W = \mathbb{F}_p^2$.

Vemos assim que $\mathbb{F}_p^2$ contém exatamente $1 + (p+1) + 1 = p+3$ subespaços vetoriais distintos.

10. Enunciado Seja $\{W_\alpha\}_{\alpha \in A}$ uma família de subespaços vetoriais de V indexados por um conjunto A (o primeiro caso do exercício é o caso em que $A = \{1,2\}$). Consideremos a interseção de todos os subespaços da família

$$\cap_{\alpha \in A} W_\alpha \overset{def}{=} \{v \in V \colon v \in W_\alpha \text{ para todo o } \alpha \in A\}$$

Temos a verificar que $\cap_{\alpha \in A} W_\alpha$ é não vazio, fechado para a soma e para o produto por escalar.

- Sendo $\alpha \in A$, o conjunto W_α é não-vazio e fechado para o produto por escalar, logo contém o vetor nulo de V. Conclui-se que $0 \in \cap_\alpha W_\alpha$ e portanto $\cap_\alpha W_\alpha$ é não vazio.

- Sejam $v, w \in \cap_{\alpha \in A} W_\alpha$. Então para cada $\alpha \in A$ temos que $v, w \in W_\alpha$. Como W_α é um subespaço vetorial de V temos então $v + w \in W_\alpha$. Como isto acontece para todo o $\alpha \in A$, conclui-se que $v + w \in \cap_{\alpha \in A} W_\alpha$.
- Sejam $v \in \cap_{\alpha \in A} W_\alpha$ e $\lambda \in \mathbb{R}$. Para cada $\alpha \in A$ temos que $v \in W_\alpha$. Como W_α é um subespaço vetorial vemos que $\lambda v \in W_\alpha$. Isto acontece para todo o α logo $\lambda v \in \cap_{\alpha \in A} W_\alpha$.

11. EnunciadoOs axiomas são todos consequências imediatas dos mesmos axiomas para V_1 e V_2 e a definição das operações. A título de exemplo verificamos alguns:
   (i) Sejam $(v_1, v_2), (w_1, w_2), (u_1, u_2) \in V_1 \times V_2$. Então

$$\begin{aligned} (v_1, v_2) + ((w_1, w_2) + (u_1, u_2)) &= (v_1, v_2) + (w_1 + u_1, w_2 + u_2) \\ &= (v_1 + (w_1 + u_1), v_2 + (w_2 + u_2)) \\ &= ((v_1 + w_1) + u_1, (v_2 + w_2) + u_2) \\ &= (v_1 + w_1, v_2 + w_2) + (u_1, u_2) \\ &= ((v_1, v_2) + (w_1, w_2)) + (u_1, u_2) \end{aligned}$$

   (iii) Sendo $0_1 \in V_1$ e $0_2 \in V_2$ os elementos neutros para a soma temos para todos os $v_1 \in V_1, v_2 \in V_2$:

$$(v_1, v_2) + (0_1, 0_2) = (v_1 + 0_1, v_2 + 0_2) = (v_1, v_2)$$

   logo $(0_1, 0_2)$ é um elemento neutro para a soma em $V_1 \times V_2$.
   (vii) Dados $\alpha \in \mathbb{R}(ou\mathbb{C})$, $(v_1, v_2), (w_1, w_2) \in V_1 \times V_2$ temos

$$\begin{aligned} \alpha \cdot ((v_1, v_2) + (w_1, w_2)) &= \alpha \cdot (v_1 + w_1, v_2 + w_2) \\ &= (\alpha \cdot (v_1 + w_1), \alpha \cdot (v_2 + w_2)) \\ &= (\alpha \cdot v_1 + \alpha \cdot w_1, \alpha \cdot v_2 + \alpha \cdot w_2) \\ &= (\alpha \cdot v_1, \alpha \cdot v_2) + (\alpha \cdot w_1, \alpha \cdot w_2) \\ &= \alpha \cdot (v_1, v_2) + \alpha \cdot (w_1, w_2) \end{aligned}$$

12. Enunciado
   (a) Se $v + W = v' + W$ então $v' \in v + W$ e portanto existe $w \in W$ tal que $v' = v + w$, pelo que $v - v' = -w \in W$. Reciprocamente se $v - v' \in W$ então dado qualquer $w \in W$ temos $v + w = v' + ((v - v') + w) \in v' + W$ logo $v + W \subset v' + W$ e, analogamente $v' + w = v + (w - (v - v')) \in v + W$ logo $v' + W \subset v + W$. Concluímos que $v + W = v' + W$.
   (b) Suponhamos que $v + W = v' + W$. Então, pela alínea anterior, $v - v' \in W$. Dado $\alpha \in \mathbb{R}$ temos $\alpha(v - v') = \alpha v - \alpha v' \in W$ portanto, novamente pela alínea anterior, $(\alpha v) + W = (\alpha v') + W$. Isto mostra que o produto por escalar está bem definido. Para ver que a soma está bem definida, suponhamos que $v + W = v' + W$ e $u + W = u' + W$, de forma que $v - v', u - u' \in W$. Então $(v + u) - (v' + u') = (v - v') + (u - u') \in W$ e portanto $(v + u) + W = (v' + u') + W$, o que mostra que a soma está bem definida em V/W.

(c) Os axiomas são uma consequência imediata dos mesmos axiomas para V desde que as operações estejam bem definidas como verificámos na alínea anterior. Verificamos apenas alguns dos axiomas a título de exemplo:

(ii) Sejam $u+W, v+W \in V/W$. Então

$$(u+W)+(v+W)=(u+v)+W=(v+u)+W=(v+W)+(u+W)$$

(iii) Vejamos que $0+W=W$ é o elemento neutro de V/W: dado $v+W \in V/W$ temos

$$v+W+(0+W)=(v+0)+W=v+W$$

(vii) Sejam $\alpha,\beta \in \mathbb{R}$ e $v+W \in V/W$. Então

$$\alpha\cdot(\beta\cdot(v+W))=\alpha\cdot((\beta v)+W)=(\alpha(\beta v))+W=((\alpha\beta)v)+W=(\alpha\beta)\cdot(v+W)$$

(d) O vetor $(1,-1,1)+W$ pertence à expansão linear de S quando existem $\alpha,\beta \in \mathbb{R}$ tais que

$$(1,-1,1)+W=\alpha((0,1,2)+W)+\beta((1,2,-1)+W)$$

De acordo com a alínea (a) isto acontece se $(1,-1,1)-\alpha(0,1,2)-\beta(1,2,-1) \in W$. Uma vez que $W=\{\lambda(1,1,1)\colon \lambda \in \mathbb{R}\}$ a afirmação anterior é equivalente à existência de $\alpha,\beta,\lambda \in \mathbb{R}$ tais que

$$(1,-1,1)-\alpha(0,1,2)-\beta(1,2,-1)=\lambda(1,1,1)$$

$$\Leftrightarrow (1,-1,1)=\alpha(0,1,2)+\beta(1,2,-1)+\lambda(1,1,1)$$

A igualdade anterior corresponde ao sistema

$$\begin{bmatrix}0&1&1&|&1\\1&2&1&|&-1\\2&-1&1&|&1\end{bmatrix}\to\begin{bmatrix}1&2&1&|&-1\\0&1&1&|&1\\2&-1&1&|&1\end{bmatrix}\to\begin{bmatrix}1&2&1&|&-1\\0&1&1&|&1\\0&-5&-1&|&3\end{bmatrix}\to$$

$$\begin{bmatrix}1&2&1&|&-1\\0&1&1&|&1\\0&0&4&|&8\end{bmatrix}\to\begin{bmatrix}1&2&1&|&-1\\0&1&1&|&1\\0&0&1&|&2\end{bmatrix}\to\begin{bmatrix}1&2&0&|&-3\\0&1&0&|&-1\\0&0&1&|&2\end{bmatrix}\to\begin{bmatrix}1&0&0&|&-1\\0&1&0&|&-1\\0&0&1&|&2\end{bmatrix}$$

Concluímos que

$$(1,-1,1)+W=(-1)\cdot((0,1,2)+W)+(-1)\cdot((1,2,-1)+W)$$

Nota: A ideia do espaço quociente por W é que estamos a considerar o espaço vetorial V mas "desprezando"os vetores (ou direções) de W no sentido em que não distinguimos dois vetores de V se a diferença entre eles pertencer a W. É uma maneira de pensar nos vetores de V com "menos detalhe", que é frequentemente útil.

13. Enunciado

(a) Temos que verificar se a equação

$$\alpha(1,0,-2,1)+\beta(2,2,1,0)+\gamma(0,1,2,1)+\delta(0,-1,-3,3)=(0,0,0,0)$$

para as incógnitas $\alpha, \beta, \gamma, \delta$ tem alguma solução além da trivial. A equação anterior corresponde ao sistema homogéneo

$$\begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 2 & 1 & -1 \\ -2 & 1 & 2 & -3 \\ 1 & 0 & 1 & 3 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 2 & 1 & -1 \\ 0 & 5 & 2 & -3 \\ 0 & -2 & 1 & 3 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 2 & 1 & -1 \\ 0 & 0 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 2 & 2 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 2 & 1 & -1 \\ 0 & 0 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Conclui-se que o conjunto não é linearmente independente. As soluções do sistema verificam $\gamma = -\delta$, $\beta = \frac{1}{2}\delta - \frac{1}{2}\gamma = \delta$ e $\alpha = -2\beta = -2\delta$ pelo que uma combinação linear não trivial que se anula é por exemplo

$$-2(1,0,-2,1) + (2,2,1,0) - (0,1,2,1) + (0,-1,-3,3)$$

(b) Queremos resolver a equação

$$\alpha(1+t+t^2) + \beta(-t+t^3) + \gamma(2+t) = 0$$

Igualando os coeficientes dos polinómios dos dois lados do sinal de igual obtemos o seguinte sistema homogéneo para α, β, γ:

$$\begin{bmatrix} 1 & 0 & 2 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 2 \\ 0 & -1 & -1 \\ 0 & 0 & -2 \\ 0 & 1 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 2 \\ 0 & -1 & -1 \\ 0 & 0 & -2 \\ 0 & 0 & -1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 2 \\ 0 & -1 & -1 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{bmatrix}$$

Este sistema tem apenas a solução trivial pelo que o conjunto de polinómios dados é linearmente independente.

Alternativamente podíamos notar que uma combinação linear dos polinómios dados corresponde a uma combinação linear das linhas da matriz

$$\begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 2 & 1 & 0 & 0 \end{bmatrix}$$

e que estas são obviamente linearmente independentes (se trocarmos a primeira com a segunda linha a matriz fica "em escada de linhas da direita para a esquerda").

(c) Resolvemos o sistema homogéneo

$$\begin{bmatrix} 1 & i & 0 \\ i & -1 & 0 \\ 0 & 1+i & 2i \end{bmatrix} \to \begin{bmatrix} 1 & i & 0 \\ 0 & 0 & 0 \\ 0 & 1+i & 2i \end{bmatrix}$$

A linha de zeros diz-nos que irá haver combinações lineares nulas não triviais. Os coeficientes α, β, γ satisfazem $\alpha = -i\beta, \gamma = -\frac{1+i}{2i}\beta = \frac{-1+i}{2}\beta$ logo uma combinação linear que se anula é, por exemplo,

$$-i(1,i,0) + (i,-1,1+i) + \frac{i-1}{2}(0,0,2i)$$

(d) A equação

$$\alpha \begin{bmatrix} 1 & 1 & 0 & 2 \\ -1 & 1 & -2 & 0 \end{bmatrix} + \beta \begin{bmatrix} 1 & 2 & 0 & 3 \\ -1 & 0 & 1 & 0 \end{bmatrix} + \gamma \begin{bmatrix} 1 & 1 & 0 & 2 \\ -1 & -1 & 1 & 0 \end{bmatrix} = 0$$

corresponde ao sistema homogéneo

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & 0 & 0 \\ 2 & 3 & 2 \\ -1 & -1 & -1 \\ 1 & 0 & -1 \\ -2 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & -1 & -2 \\ 0 & 3 & 3 \\ 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

que tem apenas a solução nula. Conclui-se que o conjunto dado é linearmente independente.

14. Enunciado

(a) Temos que resolver a equação

$$1 - t^2 = \alpha + \beta(1+t) + \gamma(1+t+t^2)$$

É imediato que $\gamma = -1$ e que $\beta + \gamma = 0 \Rightarrow \beta = 1$. Finalmente $\alpha + \beta + \gamma = 1 \Rightarrow \alpha = 1$. Conclui-se que as coordenadas de $1 - t^2$ nesta base são $(1, 1, -1)$.

(b) Para achar as coordenadas $(\alpha, \beta, \gamma, \delta)$ temos que resolver o sistema

$$\left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 4 \\ 0 & -1 & 0 & 0 & 3 \\ 1 & 0 & 1 & -1 & 2 \end{array}\right]$$

Claramente temos $\beta = -3, \gamma = 4$. A primeira e última equações ficam portanto

$$\begin{cases} \alpha + \delta = 0 \\ \alpha - \delta = -2 \end{cases} \Leftrightarrow \begin{cases} \alpha = -1 \\ \delta = 1 \end{cases}$$

logo as coordenadas são $(-1, -3, 4, 1)$.

(c) Sendo $(\alpha, \beta, \gamma, \delta)$ as coordenadas, é evidente que $\alpha = 1$ e $\delta = -1$. As restantes duas coordenadas satisfazem as relações (que vêm de igualar as entradas 12 e 21 na equação matricial)

$$\begin{cases} \beta + \gamma = 2 \\ \beta - \gamma = -1 \end{cases} \Leftrightarrow \begin{cases} \beta = \frac{1}{2} \\ \gamma = \frac{3}{2} \end{cases}$$

Conclui-se que as coordenadas são $(1, \frac{1}{2}, \frac{3}{2}, -1)$.

15. Enunciado

(a) Aplicando o método de Gauss temos

$$\begin{bmatrix} 1 & 1 & 0 \\ -1 & 2 & 1 \\ 1 & 0 & -1 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 0 \\ 0 & 3 & 1 \\ 0 & -1 & -1 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & -\frac{2}{3} \end{bmatrix}$$

logo o núcleo da matriz é $\{0\}$ sendo uma base o conjunto vazio. Uma base para o espaço das linhas é dada pelas próprias linhas da matriz (ou equivalentemente pelas linhas da matriz no final, ou em qualquer passo, do método de Gauss).
Mais tarde iremos ver que os espaços das linhas e das colunas têm sempre a mesma dimensão. Portanto o espaço das colunas desta matriz tem dimensão 3. Sendo assim as colunas têm necessariamente que formar uma base. Sem usar este resultado podemos simplesmente aplicar o método de Gauss à matriz transposta para verificar que as colunas da matriz são linearmente independentes e portanto formam uma base.

(b) As linhas da matriz não são colineares logo são independentes e portanto formam uma base para o espaço das linhas. O espaço das colunas é claramente $\mathbb{R}^2$ e uma base é, por exemplo, formada pelas duas últimas colunas. Para achar uma base para o núcleo podemos usar o método de Gauss:

$$\begin{bmatrix} 1 & 1 & 0 & 2 \\ -1 & -1 & 1 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 0 & 2 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

Vemos que o núcleo é definido pelas condições

$$\begin{cases} x = -y - 2w \\ z = -2w \end{cases}$$

Um elemento típico do núcleo escreve-se portanto na forma

$$(-y - 2w, y, -2w, w) = y(-1, 1, 0, 0) + w(-2, 0, -2, 1)$$

e uma base é dada, por exemplo, por $\{(-1, 1, 0, 0), (-2, 0, -2, 1)\}$.

(c) Aplicando o método de Gauss temos

$$\begin{bmatrix} 1 & 1 & 0 & 2 \\ -1 & -1 & 1 & 0 \\ 0 & 0 & 2 & 4 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 0 & 2 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 2 & 4 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 0 & 2 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Assim, uma base para o espaço das linhas é $\{(1, 1, 0, 2), (0, 0, 1, 2)\}$ e o núcleo, que é igual ao calculado na alínea anterior, tem base $\{(-1, 1, 0, 0), (-2, 0, -2, 1)\}$. Se nos permitirmos usar que o espaço das colunas tem também dimensão 2 podemos escolher para base quaisquer duas colunas não colineares. Caso contrário podemos aplicar o método de Gauss:

$$\begin{bmatrix} 1 & -1 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & 2 \\ 2 & 0 & 4 \end{bmatrix} \to \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 4 \end{bmatrix} \to \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

e concluir que $\{(1,-1,0),(0,1,2)\}$ forma uma base para o espaço das colunas da matriz.

16. Enunciado
   (a) Dado $p(t) = a + bt + ct^2$ temos $p'(t) = b + 2ct$,
   $$p'(1) = 0 \Leftrightarrow b + 2c = 0$$
   logo $U = \{a - 2ct + ct^2 \colon a, c \in \mathbb{R}\} = \{a \cdot 1 + c \cdot (-2t + t^2) \colon a, c \in \mathbb{R}\}$ tem $\{1, -2t+t^2\}$ como base. Esta base de U pode ser completada a uma base do espaço dos polinómios de grau ≤ 2 adicionando por exemplo o polinómio t: o conjunto $\{1, t, -2t + t^2\}$ é claramente linearmente independente (por exemplo, quando os coeficientes dos polinómios são escritos nas linhas de uma matriz ela está em escada de linhas) e portanto constitui uma base para este espaço de dimensão 3.
   (b) As duas matrizes são linearmente independentes (porque não são colineares) e por definição geram U logo constituem uma base para U. Para completar este conjunto numa base de $M_{2\times 2}(\mathbb{R})$ podemos notar que a matriz
   $$\begin{bmatrix} 1 & 2 & -1 & -1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$
   está em escada de linhas, logo as suas linhas constituem uma base de $\mathbb{R}^4$. Uma vez que as operações em $M_{2\times 2}(\mathbb{R}^4)$ se identificam com as operações em $\mathbb{R}^4$ isto diz-nos que
   $$\left\{ \begin{bmatrix} 1 & 2 \\ -1 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$
   é uma base para $M_{2\times 2}(\mathbb{R})$.

17. Enunciado
   (a) Impondo a equação que define U a um elemento $\alpha(1,3,4)+\beta(0,1,1)$ de V obtemos
   $$\alpha + (3\alpha + \beta) + (4\alpha + \beta) = 0 \Leftrightarrow \beta = -4\alpha$$
   Portanto
   $$U \cap V = \{\alpha((1,3,4) - 4(0,1,1)) \colon \alpha \in \mathbb{R}\} = L(\{(1,-1,0)\}$$
   pelo que uma base para $U \cap V$ é $\{(1,-1,0)\}$.
   (b) O espaço $U \cap V$ é o conjunto das soluções do sistema
   $$\begin{bmatrix} 1 & 0 & 1 & -1 \\ 1 & -1 & 2 & -1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 1 & -1 \\ 0 & -1 & 1 & 0 \end{bmatrix}$$
   logo $U \cap V = \{(-z+w, z, z, w) \colon z, w \in \mathbb{R}\} = \{z(-1,1,1,0)+w(1,0,0,1) \colon z, w \in \mathbb{R}\}$ tem $\{(-1,1,1,0),(1,0,0,1)\}$ como base.
   Alternativamente poderíamos ter achado uma base para um dos espaços e usado o método da alínea anterior.

(c) Para intersetar os dois espaços é conveniente expressar um deles em termos de equações cartesianas. Um vetor (x, y, z, w) está em U se e só se é uma solução do sistema

$$\begin{bmatrix} -1 & 0 & | & x \\ 3 & 3 & | & y \\ 1 & 3 & | & z \\ 0 & 1 & | & w \end{bmatrix} \to \begin{bmatrix} -1 & 0 & | & x \\ 0 & 3 & | & y+3x \\ 0 & 3 & | & z+x \\ 0 & 1 & | & w \end{bmatrix} \to \begin{bmatrix} -1 & 0 & | & x \\ 0 & 3 & | & y+3x \\ 0 & 0 & | & z-y-2x \\ 0 & 0 & | & w-\frac{y}{3}-x \end{bmatrix}$$

logo $U = \{(x, y, z, w) \in \mathbb{R}^4 \colon z - y - 2x = 0, 3w - y - 3x = 0\}$, e

$$U \cap V = U \cap \{\alpha(0,1,1,0) + \beta(1,1,3,1) \colon \alpha, \beta \in \mathbb{R}\} = U \cap \{(\beta, \alpha+\beta, \alpha+3\beta, \beta) \colon \alpha, \beta \in \mathbb{R}\}$$

é determinado pelas equações

$$\begin{cases} \alpha + 3\beta - (\alpha+\beta) - 2\beta = 0 \\ 3\beta - (\alpha+\beta) - 3\beta = 0 \end{cases} \Leftrightarrow \beta = -\alpha$$

Logo

$$U \cap V = \{\alpha((0,1,1,0) - (1,1,3,1)) \colon \alpha \in \mathbb{R}\} = L(\{(-1,0,-2,-1)\})$$

e uma base para $U \cap V$ é $\{(-1, 0, -2, -1)\}$.
Alternativamente, podemos encarar como um sistema linear a equação que expressa um ponto ser comum aos dois planos:

$$\alpha(-1,3,1,0) + \beta(0,3,3,1) = \gamma(0,1,1,0) + \delta(1,1,3,1)$$

Esta traduz-se no sistema homogéneo

$$\begin{bmatrix} -1 & 0 & 0 & -1 \\ 3 & 3 & -1 & -1 \\ 1 & 3 & -1 & -3 \\ 0 & 1 & 0 & -1 \end{bmatrix} \to \begin{bmatrix} -1 & 0 & 0 & -1 \\ 0 & 3 & -1 & -4 \\ 0 & 3 & -1 & -4 \\ 0 & 1 & 0 & -1 \end{bmatrix} \to \begin{bmatrix} -1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Conclui-se que um vetor pertence à interseção quando $\delta = -\gamma$ o que reproduz o resultado obtido anteriormente.

(d) Se $p(t) = \alpha(1+t) + \beta t^2 + \gamma(t+t^4)$ é um elemento de V então $p'(t) = \alpha + 2\beta t + \gamma(1+4t^3)$. A condição para pertencer a U expressa-se então através do sistema

$$\begin{cases} \alpha = 0 \\ \alpha + 2\beta + 5\gamma = 0 \end{cases} \Leftrightarrow \begin{cases} \alpha = 0 \\ \beta = -\frac{5}{2}\gamma \end{cases}$$

Logo $U \cap V = \{\gamma(-\frac{5}{2}t^2 + t + t^4) \colon \gamma \in \mathbb{R}\}$ e uma base para $U \cap V$ é $\{t - \frac{5}{2}t^2 + t^4\}$.

18. Enunciado Recorde-se da demonstração da Proposição 3.40 que o subespaço W é necessariamente finitamente gerado. Uma base de W é um conjunto linearmente independente em V e portanto tem no máximo $\dim V$ elementos. Isto mostra que $\dim W \leq \dim V$ para qualquer subespaço $W \subset V$. Se $\dim W = \dim V$ então uma base para W é um conjunto linearmente independente em V com $\dim V$ elementos e portanto necessariamente uma base para V pelo que $W = V$. Da nossa hipótese que $V \neq W$ conclui-se portanto que $\dim W < \dim V$.

19. Enunciado
   (a) A dimensão do espaço das linhas de A é igual à característica da matriz e portanto igual ao número de colunas com pivot numa matriz em escada de linhas que se obtenha de A por aplicação do método de Gauss.
   O núcleo de A é o conjunto das soluções do sistema homogéneo com matriz dos coeficientes igual a A. Este sistema pode ser resolvido exprimindo as variáveis correspondentes às colunas com pivot em termos das restantes variáveis que ficam livres. As variáveis livres parametrizam univocamente os vetores do núcleo pelo que a dimensão do núcleo é igual ao número de colunas sem pivot.
   Portanto $\dim EL(A) + \dim N(A) =$ número total de colunas de A=n.
   Versão mais formal (e portanto mais difícil de seguir) do segundo parágrafo: $N(A) = \{X \colon AX = 0\}$ é o conjunto das soluções do sistema homogéneo com matriz de coeficientes A. Sejam $1 \leq i_1 < i_2 < \ldots < i_k \leq n$ os indíces das colunas com pivot e $1 \leq j_1 < \ldots < j_{n-k} \leq n$ os índices das colunas sem pivot na matriz B obtida ao final do método de Gauss-Jordan. Então o sistema $AX = 0$ é equivalente a $BX = 0$ que, resolvido em ordem às variáveis correspondentes às colunas com pivot, toma a forma
$$\begin{cases} x_{i_1} = -\sum_{l=1}^{n-k} b_{1j_l} x_{j_l} \\ \vdots \\ x_{i_k} = -\sum_{l=1}^{n-k} b_{kj_l} x_{j_l} \end{cases}$$
   Conclui-se que
$$N(A) = \left\{ \sum_{l=1}^{n-k} x_{j_l} v_l \colon x_{j_l} \in \mathbb{R} \right\}$$
   em que as componentes dos vetores $v_l \in \mathbb{R}^m$ são
$$(v_l)_\alpha = \begin{cases} 1 & \text{se } \alpha = j_l \\ 0 & \text{se } \alpha = j_p \text{ com } p \neq l \\ -b_{pj_l} & \text{se } \alpha = i_p \end{cases}$$
   Claramente os vetores $\{v_1, \ldots, v_{n-k}\}$ geram $N(A)$. Mas eles são também linearmente independentes pois sendo
$$\lambda_1 v_1 + \ldots + \lambda_{n-k} v_{n-k} = 0$$
   a componente j_l desta equação em $\mathbb{R}^n$ é
$$\lambda_1 0 + \ldots + \lambda_{l-1} 0 + \lambda_l 1 + \lambda_{l+1} 0 + \ldots + \lambda_{n-k} 0 = 0 \Leftrightarrow \lambda_l = 0$$
   (b) Pela alínea anterior, se $\dim N(A) = 3$ então $\dim EL(A) = 1$ e portanto as linhas da matriz são colineares com $(2, 0, 3, -1)$. Para determinar a constante de proporcionalidade basta-nos uma das coordenadas, logo
$$(1, a, b, c) = \tfrac{1}{2}(2, 0, 3, -1) \Leftrightarrow a = 0, b = \tfrac{3}{2}, c = -\tfrac{1}{2}$$
   e
$$(d, e, f, 3) = -3(2, 0, 3, -1) \Leftrightarrow d = -6, e = 0, f = -9$$

20. Enunciado V é naturalmente um espaço vetorial real porque a multiplicação por escalar complexo se restringe a uma multiplicação por escalar real e os axiomas que têm de se verificar para que V com estas operações seja um espaço vetorial real são casos particulares dos axiomas que se verificam por V ser um espaço vetorial complexo.

    Vou responder apenas no caso em que V é finitamente gerado (que era o que eu tinha em mente na pergunta) embora o resultado seja verdade em geral, com a mesma demonstração (entendendo-se que $2\infty = \infty$).

    Seja $B = \{v_1, \ldots, v_n\}$ uma base para o espaço vetorial complexo V. Vamos verificar que $BR = \{v_1, iv_1, \ldots, v_n, iv_n\}$ é uma base para o espaço vetorial V encarado como um espaço vetorial real:

    - BR gera V enquanto espaço vetorial real: Seja $v \in V$. Então existem escalares $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$ tais que
    $$v = \alpha_1 v_1 + \ldots + \alpha_n v_n$$
    Escrevendo $\alpha_j = a_j + ib_j$ com $a_j, b_j \in \mathbb{R}$ temos então que
    $$v = (a_1 + ib_1)v_1 + \ldots + (a_n + ib_n)v_n = a_1 v_1 + b_1(iv_1) + \ldots + a_n v_n + b_n(iv_n)$$
    logo BR é um conjunto de geradores do espaço vetorial real V.
    - BR é linearmente independente: Sejam $c_1, \ldots, c_{2n} \in \mathbb{R}$ tais que
    $$c_1 v_1 + c_2(iv_1) + \ldots + c_{2n-1} v_n + c_{2n}(iv_n) = 0$$
    Então
    $$(c_1 + ic_2)v_1 + \ldots + (c_{2n-1} + ic_{2n})v_n = 0$$
    Uma vez que $\{v_1, \ldots, v_n\}$ são linearmente independentes sobre $\mathbb{C}$, os escalares complexos $c_1 + ic_2, \ldots, c_{2n-1} + ic_{2n}$ são todos nulos. Mas isto significa que os escalares reais $c_1, \ldots, c_{2n}$ são todos nulos. Logo BR é linearmente independente.

21. Enunciado
    (a) $U + V$ é não vazio porque $0 = 0 + 0 \in U + V$. Dados $u_1 + v_1, u_2 + v_2 \in U + V$ com $u_i \in U$ e $v_i \in V$ temos $u_1 + u_2 \in U$ e $v_1 + v_2 \in V$ logo $(u_1 + v_1) + (u_2 + v_2) = (u_1 + u_2) + (v_1 + v_2) \in U + V$ o que mostra que $U + V$ é fechado para a soma. Finalmente dado um escalar α e vetores $u \in U, v \in V$ temos $\alpha(u + v) = \alpha u + \alpha v$. Como U e V são fechados para o produto por escalar, $\alpha u \in U$ e $\alpha v \in V$ logo $\alpha(u+v) \in U+V$ e portanto $U+V$ é fechado para o produto por escalar. Conclui-se que $U + V$ é um subespaço vetorial de W.
    (b) A operação é comutativa uma vez que $\{u+v \colon u \in U, v \in V\} = \{v+u \colon v \in V, u \in U\}$ pela comutatividade da soma em W. Exatamente o mesmo argumento mostra que a operação é associativa. O espaço vetorial trivial é um elemento neutro para esta operação: $U + \{0\} = \{u + 0 \colon u \in U\} = U$. O único subespaço para o qual existe um inverso para esta operação é o espaço nulo (que é o seu próprio inverso). De facto, se $U \neq \{0\}$ então como $U + V \supset U$, $U + V$ não pode conter apenas o vetor nulo.
    (c) Uma vez que W é finitamente gerado, o mesmo se aplica aos seus subespaços $U, V, U \cap V$ e $U + V$ (cf. demonstração da Proposição 3.40(iii)). Seja $\{x_1, \ldots, x_k\}$

uma base para $U \cap V$ (de forma que $k = \dim U \cap V$). Uma vez que U, V são finitamente gerados podemos completar esta base a bases

$$\{x_1, \ldots, x_k, u_1, \ldots, u_m\} \text{ para } U \quad \text{e} \quad \{x_1, \ldots, x_k, v_1, \ldots, v_n\} \text{ para } V$$

de forma que $\dim(U) = k + m$ e $\dim(V) = k + n$. Queremos ver que $\dim(U + V) = k + m + n$ pelo que é suficiente verificar que o conjunto

$$B = \{x_1, \ldots, x_k, u_1, \ldots, u_m, v_1, \ldots, v_n\}$$

é uma base para $U + V$.

- B gera $U + V$: Dados $u \in U, v \in V$ existem escalares $\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_m$ e $\alpha'_1, \ldots, \alpha'_k, \gamma_1, \ldots, \gamma_n$ tais que

$$u = \alpha_1 x_1 + \ldots + \alpha_k x_k + \beta_1 u_1 + \ldots + \beta_m u_m$$

e

$$v = \alpha'_1 x_1 + \ldots + \alpha'_k x_k + \gamma_1 v_1 + \ldots + \gamma_n v_n$$

Então

$$u + v = (\alpha_1 + \alpha'_1) x_1 + \ldots + (\alpha_k + \alpha'_k) x_k + \beta_1 u_1 + \ldots + \beta_m u_m + \gamma_1 v_1 + \ldots + \gamma_n v_n$$

- B é linearmente independente: Suponhamos que $\alpha_i, \beta_j, \gamma_k$ são escalares tais que

$$\alpha_1 x_1 + \ldots + \alpha_k x_k + \beta_1 u_1 + \ldots + \beta_m u_m + \gamma_1 v_1 + \ldots + \gamma_n v_n = 0$$

Então

$$\gamma_1 v_1 + \ldots + \gamma_n v_n = -(\alpha_1 x_1 + \ldots + \alpha_k x_k + \beta_1 u_1 + \ldots + \beta_m u_m)$$

pertence a U e também a V, logo existem escalares $\alpha'_1, \ldots, \alpha'_k$ tais que

$$\gamma_1 v_1 + \ldots + \gamma_n v_n = \alpha'_1 x_1 + \ldots + \alpha'_k x_k \Leftrightarrow -\alpha'_1 x_1 - \ldots - \alpha'_k x_k + \gamma_1 v_1 + \ldots + \gamma_n v_n = 0$$

Como $\{x_1, \ldots, x_k, v_1, \ldots, v_n\}$ é um conjunto linearmente independente conclui-se (em particular) que $\gamma_1 = \cdots = \gamma_n = 0$. Portanto

$$0 = \alpha_1 x_1 + \ldots + \alpha_k x_k + \beta_1 u_1 + \ldots + \beta_m u_m$$

Mas $\{x_1, \ldots, x_k, u_1, \ldots, u_m\}$ é também um conjunto linearmente independente, logo $\alpha_1 = \cdots = \alpha_k = \beta_1 = \cdots = \beta_m = 0$, o que mostra que B é linearmente independente.

(d) Temos a verificar que $U \cap V = \{0\}$. Dada $A \in U \cap V$ temos

$$\begin{cases} A = A^T \\ A = -A^T \end{cases} \Rightarrow A = -A \Rightarrow A = 0$$

U é o subespaço das matrizes simétricas e V o das matrizes anti-simétricas. Vimos já na ficha sobre matrizes que dada $A \in M_{n \times n}(\mathbb{R})$ temos a decomposição única

$$A = \frac{A + A^T}{2} + \frac{A - A^T}{2}$$

de uma matriz como soma de uma matriz simétrica e uma anti-simétrica.

(e) Seja $\{u_1, \ldots, u_k\}$ uma base para U e sejam $w_1, \ldots, w_m$ vetores de W (distintos) tais que $\{u_1, \ldots, u_k, w_1, \ldots, w_m\}$ é uma base para W. Seja $V = L(\{w_1, \ldots, w_m\})$. Então claramente $W = U + V$. Por outro lado se $w \in U \cap V$ então existem escalares α_i e β_j tais que

$$w = \alpha_1 u_1 + \ldots + \alpha_k u_k \quad \text{e} \quad w = \beta_1 w_1 + \ldots + \beta_m w_m$$

(respetivamente porque $w \in U$ e porque $w \in V$). Então

$$\alpha_1 u_1 + \ldots + \alpha_k u_k - (\beta_1 w_1 + \ldots + \beta_m w_m) = 0$$

e uma vez que $\{u_1, \ldots, u_k, w_1, \ldots, w_m\}$ é um conjunto linearmente independente, isto significa que $\alpha_1 = \ldots = \alpha_k = \beta_1 = \ldots = \beta_m = 0$, o que implica que $w = 0$. Portanto $U \cap V = 0$ e logo $U \oplus V = W$.

22. EnunciadoUma vez que o espaço dos polinómios de grau $\leq n$ tem dimensão $n+1$, para ver que estes $n+1$ polinómios formam uma base basta verificar que são linearmente independentes. Note-se que o polinómio p_i se anula em $t = t_j$ sempre que $j \neq i$ (porque o numerador da fração se anula) e que $p_i(t_i) = 1$ (o numerador e denominador ficam iguais).

Suponhamos que $\alpha_0 p_0 + \ldots + \alpha_n p_n = 0$. Então avaliando esta igualdade em $t = t_0, \ldots, t_n$ sucessivamente obtemos

$$\begin{cases} \alpha_0 \cdot 1 + \alpha_1 \cdot 0 + \ldots + \alpha_n \cdot 0 = 0 \\ \alpha_0 \cdot 0 + \alpha_1 \cdot 1 + \ldots + \alpha_n \cdot 0 = 0 \\ \vdots \\ \alpha_0 \cdot 0 + \alpha_1 \cdot 0 + \ldots + \alpha_n \cdot 1 = 0 \end{cases} \Leftrightarrow \begin{cases} \alpha_0 = 0 \\ \alpha_1 = 0 \\ \vdots \\ \alpha_n = 0 \end{cases}$$

donde se conclui que $\{p_0, \ldots, p_n\}$ é um conjunto linearmente independente. Podemos usar o mesmo argumento para achar as coordenadas nesta base de um polinómio qualquer: para achar α_i tais que $\alpha_0 p_0 + \ldots + \alpha_n p_n = p$ podemos novamente avaliar esta igualdade em $t = t_i$ obtendo agora o sistema

$$\begin{cases} \alpha_0 \cdot 1 + \alpha_1 \cdot 0 + \ldots + \alpha_n \cdot 0 = p(t_0) \\ \alpha_0 \cdot 0 + \alpha_1 \cdot 1 + \ldots + \alpha_n \cdot 0 = p(t_1) \\ \vdots \\ \alpha_0 \cdot 0 + \alpha_1 \cdot 0 + \ldots + \alpha_n \cdot 1 = p(t_n) \end{cases} \Leftrightarrow \begin{cases} \alpha_0 = p(t_0) \\ \alpha_1 = p(t_1) \\ \vdots \\ \alpha_n = p(t_n) \end{cases}$$

portanto as coordenadas de um polinómio $p(t)$ nesta base são precisamente os valores que p assume nos pontos $t_0, \ldots, t_n$. Note-se que isto significa que um polinómio de grau $\leq n$ é completamente determinado pelos valores que assume em $n+1$ pontos distintos - generalizando o caso bem conhecido quando o grau é 1 que diz que uma reta (não vertical) é determinada por dois pontos (com abcissas distintas).

Interpolar significa "passar por pontos". Esta base permite rapidamente escrever uma fórmula para um polinómio de grau n que passa por $n+1$ pontos dados do plano (desde que não haja dois na mesma reta vertical claro).

23. EnunciadoSeja V o espaço vetorial de todos os polinómios reais e B o conjunto do enunciado. B é um conjunto de geradores porque $L(B)$ contém $t^k = 1 + t + \ldots + t^k - (1 + t + \ldots + t^{k-1})$ para todo o k e portanto $L(B) \supset L(\{t^k \colon k \in \mathbb{N}_0\}) = V$. Resta ver que B é linearmente independente. Sejam $v_1, \ldots, v_N \in B$ vetores distintos e $\alpha_1, \ldots, \alpha_N$ escalares tais que

$$\alpha_1 v_1 + \ldots + \alpha_N v_N = 0 \tag{75}$$

Sem perda de generalidade podemos assumir que $v_1 = 1, \ldots, v_N = 1+t+\ldots+t^{N-1}$ (caso contrário podemos simplesmente acrescentar vetores a $\{v_1, \ldots, v_N\}$ e reordenar de forma a que isso suceda e se um conjunto é linearmente independente todo o seu subconjunto é também linearmente independente). Igualando os termos de grau $N-1, N-2, \ldots, 0$ sucessivamente em (75) obtemos o sistema

$$\begin{cases} \alpha_N = 0 \\ \alpha_{N-1} + \alpha_N = 0 \\ \vdots \\ \alpha_1 + \ldots + \alpha_N = 0 \end{cases}$$

que tem apenas a solução nula. Isto mostra que o conjunto B é linearmente independente e portanto é uma base para V.

24. Enunciado Se a dimensão fosse finita (digamos n) então existiria uma bijeção entre $\mathbb{R}$ e $\mathbb{Q}^n$ que associaria a um número real as suas coordenadas numa base qualquer fixada. Mas $\mathbb{Q}^n$ é um produto finito de conjuntos numeráveis, logo é numerável, enquanto que $\mathbb{R}$ não é numerável.

25. Enunciado A função constante igual a 0 é contínua logo $C(\mathbb{R})$ é não vazio. Como a soma de funções contínuas é contínua e o produto de um escalar por uma função contínua é uma função contínua, $C(\mathbb{R})$ é um subespaço vetorial de $F(\mathbb{R}, \mathbb{R})$.

   Sejam $\alpha_1, \ldots, \alpha_k$ reais distintos, que supomos ordenados de forma crescente. Suponhamos que c_i são escalares tais que

$$c_1 e^{\alpha_1 x} + \ldots + c_k e^{\alpha_k x} = 0$$

Então

$$\lim_{x \to +\infty} e^{-\alpha_k x}(c_1 e^{\alpha_1 x} + \ldots + c_k e^{\alpha_k x}) = 0 \Leftrightarrow \lim_{x \to +\infty} c_1 e^{(\alpha_1 - \alpha_k)x} + \ldots + c_{k-1} e^{(\alpha_{k-1} - \alpha_k)x} + c_k = 0$$

Para $i < k$ temos $\alpha_i - \alpha_k < 0$ logo todos os termos do limite anterior excepto o último anulam-se. Obtemos assim

$$0 + \ldots + 0 + c_k = 0$$

donde

$$c_1 e^{\alpha_1 x} + \ldots + c_{k-1} e^{\alpha_{k-1} x} = 0.$$

Prosseguindo da mesma forma, ou argumentando por indução concluímos que o conjunto $\{e^{\alpha_1 x}, \ldots, e^{\alpha_k x}\}$ é linearmente independente. Uma vez que $\alpha_1, \ldots, \alpha_k$ são arbitrários, isto significa que $\{e^{\alpha x} : \alpha \in \mathbb{R}\}$ é linearmente independente.

Nota: Por palavras, as funções $e^{\alpha x}$ são linearmente independentes porque "crescem"todas de forma distinta quando $x \to +\infty$.

26. EnunciadoSeja $\{w_1, \ldots, w_k\}$ uma base para W e $v_1, \ldots, v_m$ vetores de V tais que $\{w_1, \ldots, w_k, v_1, \ldots, v_m\}$ é uma base para V. Vamos verificar que $B = \{v_1 + W, \ldots, v_m + W\}$ é uma base para V/W pelo que $\dim(V/W) = m = (m+k) - k = \dim(V) - \dim(W)$.
    - B gera V/W: Seja $v + W \in V/W$. Então existem escalares $\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_m$ tais que $v = \alpha_1 w_1 + \ldots + \alpha_k w_k + \beta_1 v_1 + \ldots + \beta_m v_m$ e portanto
    $$v - \beta_1 v_1 - \ldots - \beta_m v_m \in W$$
    Portanto
    $$v + W = (\beta_1 v_1 + \ldots + \beta_m v_m) + W = \beta_1(v_1 + W) + \ldots + \beta_m(v_m + W)$$
    pelo que $L(B) = V/W$.
    - B é linearmente independente: Suponhamos que $\beta_1(v_1+W)+\ldots+\beta_m(v_m+W) = 0$. Então $(\beta_1 v_1 + \ldots + \beta_m v_m) + W = 0 \Leftrightarrow \beta_1 v_1 + \ldots + \beta_m v_m \in W$. Isto significa que existem $\alpha_1, \ldots, \alpha_k$ tais que
    $$\beta_1 v_1 + \ldots + \beta_m v_m = \alpha_1 w_1 + \ldots + \alpha_k w_k$$
    Uma vez que $\{w_1, \ldots, w_k, v_1, \ldots, v_m\}$ é linearmente independente, isto só pode acontecer se todos os escalares forem nulos. Em particular $\beta_1 = \cdots = \beta_m = 0$ pelo que B é linearmente independente.

27. Enunciado Começamos por achar as coordenadas do vetor $(3, 4)$ na base B_1:
$$\alpha(1,1) + \beta(1,-1) = (3,4) \Leftrightarrow \begin{cases} \alpha + \beta = 3 \\ \alpha - \beta = 4 \end{cases} \Leftrightarrow \begin{cases} \alpha = \frac{7}{2} \\ \beta = -\frac{1}{2} \end{cases}$$
logo
$$[(3,4)]_{B_2} = S_{B_1 \to B_2} \begin{bmatrix} \frac{7}{2} \\ -\frac{1}{2} \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} \frac{7}{2} \\ -\frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{5}{2} \\ -4 \end{bmatrix}$$

28. EnunciadoTemos
$$\begin{bmatrix} 2 & -1 \\ 2 & 4 \end{bmatrix} = S_{B_1 \to B_2} \begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix}$$
logo
$$S_{B_1 \to B_2} = \begin{bmatrix} 2 & -1 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix}^{-1} = \begin{bmatrix} 2 & -1 \\ 2 & 4 \end{bmatrix} \frac{1}{5} \begin{bmatrix} 3 & 1 \\ -2 & 1 \end{bmatrix} = \frac{1}{5} \begin{bmatrix} 8 & 1 \\ -2 & 6 \end{bmatrix}$$

29. EnunciadoUma vez que B_1 é a base canónica é imediato obter a matriz $S_{B_2 \to B_1}$ que podemos depois inverter para achar $S_{B_1 \to B_2}$. Não se justifica no entanto efetuar estes cálculos porque é imediato escrever diretamente os elementos da base B_1 na base B_2:
$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \qquad \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix} + \begin{bmatrix} 0 & \frac{1}{2} \\ -\frac{1}{2} & 0 \end{bmatrix}$$
$$\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix} - \begin{bmatrix} 0 & \frac{1}{2} \\ -\frac{1}{2} & 0 \end{bmatrix} \qquad \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Logo

$$S_{B_1\to B_2} = \begin{bmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & 1 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ \frac{1}{2} & 0 & 0 & -\frac{1}{2} \end{bmatrix}$$

30. EnunciadoA matriz $S_{B_1\to B_2}$ tem por i-ésima coluna as coordenadas do i-ésimo vetor da base B_1 na base B_2. Isto dá-nos um sistema para achar os vetores v_1, v_2, v_3 da base B_2:

$$\begin{cases} (1,0,0) = v_1 - v_2 \\ (1,1,0) = v_1 + v_3 \\ (1,1,1) = 2v_2 \end{cases} \Leftrightarrow \begin{cases} v_1 = (1,0,0) + v_2 = (\frac{3}{2}, \frac{1}{2}, \frac{1}{2}) \\ v_3 = (1,1,0) - v_1 = (-\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}) \\ v_2 = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2}) \end{cases}$$

portanto $B_2 = ((\frac{3}{2}, \frac{1}{2}, \frac{1}{2}), (\frac{1}{2}, \frac{1}{2}, \frac{1}{2}), (-\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}))$.

Alternativamente, podemos achar a base B_2 pensando que os vetores desta base formam as colunas da matriz $S_{B_2\to B_{can}}$. Uma vez que

$$S_{B_2\to B_{can}} S_{B_1\to B_2} = S_{B_1\to B_{can}}$$

temos

$$S_{B_2\to B_{can}} = S_{B_1\to B_{can}} S^{-1}_{B_1\to B_2} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 2 \\ 0 & 1 & 0 \end{bmatrix}^{-1} =$$

$$= \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{3}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{bmatrix}$$

o que reproduz o resultado obtido anteriormente.

31. Enunciado
   (a) Se $p(x)$ é um polinómio de grau 2 então, para qualquer $t \in \mathbb{R}$ temos $p(x) = p(t) + p'(t)(x-t) + \frac{p''(t)}{2}(x-t)^2$. Isto mostra que $L(B_2) = V$. Uma vez que a dimensão de V é 3 e B_2 contém uma base, B_2 é necessariamente uma base.
   (b) As coordenadas de $p(x)$ na base B_2 calculam-se avaliando o polinómio e as suas primeiras derivadas em t como indicado na alínea anterior. Aplicando este procedimento aos três elementos da base B_1 obtemos

$$S_{B_1\to B_2} = \begin{bmatrix} 1 & t & t^2 \\ 0 & 1 & 2t \\ 0 & 0 & 1 \end{bmatrix}$$

32. EnunciadoA matriz $S_{B_3\to B_2}$ é a inversa de $S_{B_2\to B_3}$:

$$\begin{bmatrix} 0 & 1 & 0 & | & 1 & 0 & 0 \\ 2 & -1 & 1 & | & 0 & 1 & 0 \\ 0 & 1 & 1 & | & 0 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 2 & -1 & 1 & | & 0 & 1 & 0 \\ 0 & 1 & 0 & | & 1 & 0 & 0 \\ 0 & 1 & 1 & | & 0 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 2 & -1 & 1 & | & 0 & 1 & 0 \\ 0 & 1 & 0 & | & 1 & 0 & 0 \\ 0 & 0 & 1 & | & -1 & 0 & 1 \end{bmatrix} \to$$

$$\begin{bmatrix} 1 & -\frac{1}{2} & \frac{1}{2} & | & 0 & \frac{1}{2} & 0 \\ 0 & 1 & 0 & | & 1 & 0 & 0 \\ 0 & 0 & 1 & | & -1 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & -\frac{1}{2} & 0 & | & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ 0 & 1 & 0 & | & 1 & 0 & 0 \\ 0 & 0 & 1 & | & -1 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 & | & 1 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 1 & 0 & | & 1 & 0 & 0 \\ 0 & 0 & 1 & | & -1 & 0 & 1 \end{bmatrix}$$

logo

$$S_{B_3 \to B_2} = \begin{bmatrix} 1 & \frac{1}{2} & -\frac{1}{2} \\ 1 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}$$

Por outro lado

$$S_{B_1 \to B_2} = S_{B_3 \to B_2} S_{B_1 \to B_3} = \begin{bmatrix} 1 & \frac{1}{2} & -\frac{1}{2} \\ 1 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 2 \\ 0 & -1 & 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{3}{2} & 1 \\ 1 & 1 & 0 \\ -1 & -2 & 0 \end{bmatrix}$$

33. EnunciadoOs dois sistemas de coordenadas estão relacionados por

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \Leftrightarrow \begin{cases} u = x + y \\ v = x + 2y \end{cases}$$

e

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}^{-1} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} \Leftrightarrow \begin{cases} x = 2u - v \\ y = -u + v \end{cases}$$

Uma expressão transforma-se de umas coordenadas nas outras substituindo as relações dadas.

(a) Uma vez que $x = 2u - v$ e $y = -u + v$, um vetor de $\mathbb{R}^2$ é tal que as suas coordenadas (x, y) na base B_1 satisfazem $x^2 + y^2 = 1$ se e só se as suas coordenadas (u, v) na base B_2 satisfazem

$$(2u - v)^2 + (-u + v)^2 = 1 \Leftrightarrow 5u^2 - 6uv + 2v^2 = 1$$

(que é portanto a equação da circunferência nas coordenadas (u, v)).

(b) Podemos argumentar como na alínea anterior ou alternativamente notar que

$$x^2 + 2xy + y^2 + x + 2y = 0 \Leftrightarrow (x + y)^2 + (x + 2y) = 0 \Leftrightarrow u^2 + v = 0$$

(portanto a curva indicada é uma parábola no novo sistema de coordenadas).

(c) O valor da função f no vetor com coordenadas (u, v) na base B_2 é

$$(2u - v)^2 + \cos((2u - v)(-u + v))$$

o que normalmente (apesar de ser um pouco abusivo) se escreve

$$f(u, v) = (2u - v)^2 + \cos((2u - v)(-u + v)) = 4u^2 - 4uv + v^2 + \cos(-2u^2 + 3uv - v^2)$$

(d) Uma vez que $x = 2u - v, y = -u + v$ temos

$$\frac{dx}{dt} = 2\frac{du}{dt} - \frac{dv}{dt} \qquad \frac{dy}{dt} = -\frac{du}{dt} + \frac{dv}{dt}$$

logo a equação diferencial nas coordenadas (x, y) é transformada na nova equação

$$2\frac{du}{dt} - \frac{dv}{dt} + (2u - v)(-\frac{du}{dt} + \frac{dv}{dt}) = 2u - v + (-u + v) \Leftrightarrow (2 - 2u + v)\frac{du}{dt} + (2u - v - 1)\frac{dv}{dt} = u$$

34. Enunciado Seja $S = [a_{ij}]_{i,j=1}^n$ uma matriz invertível com inversa $S^{-1} = [b_{ij}]$ e $B_1 = (v_1, \ldots, v_n)$ uma base ordenada de V. Procuramos uma base $B_2 = (w_1, \ldots, w_n)$ tal que $S_{B_1 \to B_2} = S$. Isto significa que procuramos vetores $w_i \in V$ que constituam uma base de V e satisfaçam as seguintes relações:

$$\begin{array}{rcl} v_1 & = & a_{11}w_1 + a_{21}w_2 + \ldots + a_{n1}w_n \\ \vdots & & \vdots \\ v_n & = & a_{1n}w_1 + a_{2n}w_2 + \ldots + a_{nn}w_n \end{array}$$

Estas relações podem ser vistas como um sistema de equações para as incógnitas w_j. Vamos ver que a invertibilidade da matriz S garante que este sistema tem uma solução única e as entradas da matriz S^{-1} permitem escrever a solução (os vetores w_j) à custa dos v_i.

Para tal convém escrever as relações acima na seguinte forma:

$$v_i = \sum_{k=1}^n a_{ki} w_k \qquad i = 1, \ldots, n \tag{76}$$

Note-se que a afirmação que S e S^{-1} são inversas equivale às condições

$$\sum_{k=1}^n a_{ik} b_{kj} = \delta_{ij} \qquad \sum_{l=1}^n b_{il} a_{lj} = \delta_{ij} \qquad \text{para todos os } i, j = 1, \ldots n$$

onde

$$\delta_{ij} = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{caso contrário} \end{cases}$$

(chamado o símbolo de Kronecker) nos dá as entradas da matriz identidade.

Combinando as equações (76) usando como coeficientes as entradas da matriz inversa obtemos

$$\sum_{i=1}^n b_{ij} v_i = \sum_{i=1}^n b_{ij} \left(\sum_{k=1}^n a_{ki} w_k \right) = \sum_{k=1}^n \left(\sum_{i=1}^n a_{ki} b_{ij} \right) w_k = \sum_{k=1}^n \delta_{kj} w_k = w_j$$

(onde na segunda igualdade trocámos a ordem da soma e dos fatores). Deduzimos que os vetores w_j que procuramos têm necessariamente que satisfazer as relações:

$$\begin{array}{rcl} w_1 & = & b_{11}v_1 + b_{21}v_2 + \ldots + b_{n1}v_n \\ \vdots & = & \vdots \\ w_n & = & b_{1n}v_1 + b_{2n}v_2 + \ldots + b_{nn}v_n \end{array}$$

Na realidade este sistema é equivalente ao inicial, uma vez que podemos obter o sistema inicial a partir deste se combinarmos as equações usando os coeficientes da matriz S (as contas são inteiramente análogas às anteriores).

Vemos assim que existe exatamente uma solução possível para o sistema (76), dada pela fórmula

$$w_j = \sum_{i=1}^{n} b_{ij} v_i$$

Resta-nos apenas justificar que $\{w_1, \ldots, w_n\}$ constitui uma base de V. Mas (76) diz-nos que $\{v_1, \ldots, v_n\} \subset L(\{w_1, \ldots, w_n\})$ logo $\{w_1, \ldots, w_n\}$ é um conjunto de geradores para V. Uma vez que um conjunto de geradores contém sempre uma base e que as bases de V têm todas n elementos, vemos que $\{w_1, \ldots, w_n\}$ é necessariamente uma base.
Resolução alternativa: Seja $S = [a_{ij}]_{i,j=1}^n$ uma matriz invertível com inversa $S^{-1} = [b_{ij}]$ e $B_1 = (v_1, \ldots, v_n)$ uma base ordenada de V. Procuramos uma base $B_2 = (w_1, \ldots, w_n)$ tal que $S_{B_1 \to B_2} = S$. Isto significa que procuramos vetores $w_i \in V$ que constituam uma base de V e que sejam tais que as seguintes relações sejam satisfeitas:

$$\begin{array}{rcl} v_1 & = & a_{11}w_1 + a_{21}w_2 + \ldots + a_{n1}w_n \\ \vdots & & \vdots \\ v_n & = & a_{1n}w_1 + a_{2n}w_2 + \ldots + a_{nn}w_n \end{array}$$

Estas relações podem ser vistas como um sistema de equações para as incógnitas w_j. Vamos resolver este sistema matricialmente, traduzindo-o em coordenadas numa base auxiliar B de V (arbitrária). Escrevendo $[v]_B$ para a matriz coluna das coordenadas de $v \in V$ na base B, o sistema acima é equivalente a

$$\begin{array}{rcl} [v_1]_B & = & a_{11}[w_1]_B + a_{21}[w_2]_B + \ldots + a_{n1}[w_n]_B \\ \vdots & & \vdots \\ [v_n]_B & = & a_{1n}[w_1]_B + a_{2n}[w_2]_B + \ldots + a_{nn}[w_n]_B \end{array}$$

que se pode escrever matricialmente como

$$\begin{bmatrix} | & | & \cdots & | \\ [v_1]_B & [v_2]_B & \cdots & [v_n]_B \\ | & | & \cdots & | \end{bmatrix} = \begin{bmatrix} | & | & \cdots & | \\ [w_1]_B & [w_2]_B & \cdots & [w_n]_B \\ | & | & \cdots & | \end{bmatrix} S \Leftrightarrow$$

$$\Leftrightarrow \begin{bmatrix} | & | & \cdots & | \\ [w_1]_B & [w_2]_B & \cdots & [w_n]_B \\ | & | & \cdots & | \end{bmatrix} = \begin{bmatrix} | & | & \cdots & | \\ [v_1]_B & [v_2]_B & \cdots & [v_n]_B \\ | & | & \cdots & | \end{bmatrix} S^{-1}$$

Novamente, a segunda equação matricial é equivalente a

$$\begin{array}{rcl} w_1 & = & b_{11}v_1 + b_{21}v_2 + \ldots + b_{n1}v_n \\ \vdots & & \vdots \\ w_n & = & b_{1n}v_1 + b_{2n}v_2 + \ldots + b_{nn}v_n \end{array}$$

e vemos assim que existe exatamente uma solução $(w_1, \ldots, w_n)$ do sistema que queríamos resolver. Podemos facilmente justificar que os vetores w_i formam uma base, como no final da primeira resolução.

8.4. **Transformações lineares.**

1. Enunciado

(i) Temos

$$\begin{aligned} f((x_1,y_1,z_1)+(x_2,y_2,z_2)) &= f(x_1+x_2,y_1+y_2,z_1+z_2) \\ &= (x_1+x_2-2(y_1+y_2),0,x_1+x_2+y_1+y_2-3(z_1+z_2),0) \\ &= (x_1-2y_1,0,x_1+y_1-3z_1,0)+(x_2-2y_2,0,x_2+y_2-3z_2,0) \\ &= f(x_1,y_1,z_1)+f(x_2,y_2,z_2) \end{aligned}$$

e

$$\begin{aligned} f(\alpha(x,y,z)) &= f(\alpha x,\alpha y,\alpha z) \\ &= (\alpha x-2\alpha y,0,\alpha x+\alpha y-3\alpha z,0) \\ &= \alpha(x-2y,0,x+y-3z,0)=\alpha f(x,y,z) \end{aligned}$$

logo f é uma transformação linear.

(ii) Dadas matrizes A, B e um escalar α temos
- $f(A+B)=(A+B)^T=A^T+B^T=f(A)+f(B)$
- $f(\alpha A)=(\alpha A)^T=\alpha A^T=\alpha f(A)$

logo f é uma transformação linear.

(iii) Uma vez que $f(0,0)=(0,1)$ não é o vetor zero do espaço vetorial de chegada, conclui-se que f não é uma transformação linear.

(iv) Temos

$$\begin{aligned} f\left(\begin{bmatrix} a & b \\ c & d\end{bmatrix}+\begin{bmatrix} a' & b' \\ c' & d'\end{bmatrix}\right) &= f\left(\begin{bmatrix} a+a' & b+b' \\ c+c' & d+d'\end{bmatrix}\right) \\ &= a+a'+b+b'=(a+b)+(a'+b') \\ &= f\left(\begin{bmatrix} a & b \\ c & d\end{bmatrix}\right)+f\left(\begin{bmatrix} a' & b' \\ c' & d'\end{bmatrix}\right) \end{aligned}$$

e

$$\begin{aligned} f\left(\alpha\begin{bmatrix} a & b \\ c & d\end{bmatrix}\right) &= f\left(\begin{bmatrix} \alpha a & \alpha b \\ \alpha c & \alpha d\end{bmatrix}\right) \\ &= \alpha a+\alpha b=\alpha(a+b) \\ &= \alpha f\left(\begin{bmatrix} a & b \\ c & d\end{bmatrix}\right) \end{aligned}$$

logo f é uma transformação linear.

(v) T não é uma transformação linear porque $T(0,0,0)\neq(0,0)$ (também não preserva a soma nem o produto por escalar).

(vi) T não é uma transformação linear porque não preserva o produto por escalar. Por exemplo $T(2I)=4I+2I=6I$ é diferente de $2T(I)=2(I+I)=4I$. T também não preserva a soma.

(vii) Temos

$$\begin{aligned} T(f+g) &= ((f+g)(0),(f+g)(2)) = (f(0)+g(0), f(2)+g(2)) \\ &= (f(0),f(2)) + (g(0),g(2)) = T(f)+T(g), \\ T(\alpha f) &= ((\alpha f)(0),(\alpha f)(2)) = (\alpha f(0), \alpha f(2)) = \\ &= \alpha(f(0),f(2)) = \alpha T(f) \end{aligned}$$

logo T é uma transformação linear.

(viii) Temos

$$T(v_1+v_2) = \alpha(v_1+v_2) = \alpha v_1 + \alpha v_2 = T(v_1)+T(v_2)$$

e

$$T(\beta v) = \alpha(\beta v) = \beta(\alpha v) = \beta T(v)$$

logo T é uma transformação linear.

2. Enunciado

(a) Podemos escrever $(2,5)$ como combinação linear de $(1,1)$ e $(1,2)$:

$$(2,5) = \alpha(1,1)+\beta(1,2) \Leftrightarrow \begin{cases} \alpha+\beta = 2 \\ \alpha+2\beta = 5 \end{cases} \Leftrightarrow \begin{cases} \beta = 3 \\ \alpha = -1 \end{cases}$$

Uma vez que T é uma transformação linear temos portanto que

$$T(2,5) = -T(1,1)+3T(1,2) = -(2,0,1)+3(1,4,6) = (1,12,17)$$

(b) Começamos por observar que o segmento de reta que une dois vetores v_1 e v_2 num espaço vetorial é parametrizado por

$$v_1 + t(v_2 - v_1) = (1-t)v_1 + tv_2 \qquad \text{com } t \in [0,1]$$

Em particular, o segmento de reta que une $(0,1)$ a $(1,2)$ é descrito pelas combinações lineares $(1-t)(0,1)+t(1,2), t \in [0,1]$. O triângulo do enunciado é a união dos segmentos de reta cujas extremidades são a origem $(0,0)$ e um ponto do segmento $(1-t)(0,1)+t(1,2), t \in [0,1]$ ou seja, é o conjunto dos pontos

$$S = \{r((1-t)(0,1)+t(1,2))\colon 0 \leq r \leq 1, 0 \leq t \leq 1\}$$

Uma vez que T preserva combinações lineares, a imagem $T(S)$ é o conjunto

$$T(S) = \{r((1-t)T(0,1)+tT(1,2))\colon 0 \leq r \leq 1, 0 \leq t \leq 1\}$$

que é o triângulo dado pela união de todos os segmentos de reta que vão de $T(0,0) = (0,0,0)$ ao segmento de reta de $T(0,1) = T(1,2)-T(1,1) = (-1,4,5)$ a $T(1,2) = (1,4,6)$, ou seja, o triângulo com vértices $(0,0,0), (-1,4,5)$ e $(1,4,6)$.

(c) Já vimos que $T(0,1) = (-1,4,5)$ e temos que $T(1,0) = T(1,1)-T(0,1) = (2,0,1)-(-1,4,5) = (3,-4,-4)$ logo

$$A_{T,B_{can},B_{can}} = \begin{bmatrix} 3 & -1 \\ -4 & 4 \\ -4 & 5 \end{bmatrix}$$

A expressão para $T(x,y)$ é portanto

$$T(x,y) = (3x - y, -4x + 4y, -4x + 5y)$$

3. Enunciado
   (a) O retângulo $[0,2] \times [0,1]$ é o conjunto

   $$\{\alpha(2,0) + \beta(0,1) \colon 0 \leq \alpha, \beta \leq 1\}$$

   Uma vez que f é uma transformação linear e portanto preserva combinações lineares, a imagem do retângulo é o conjunto

   $$\{\alpha f(2,0) + \beta f(0,1) \colon 0 \leq \alpha, \beta, \leq 1\}$$

   que é um paralelogramo com vértice na origem e arestas $f(2,0)$ e $f(0,1)$ (pela interpretação geométrica das operações com vetores de $\mathbb{R}^2$). Para que este paralelogramo seja o quadrado com os vértices indicados temos que escolher para imagem dos vértices $(2,0)$ e $(0,1)$ os vectores $(1,2)$ e $(-2,1)$ (numa ordem qualquer). Há duas maneiras de o fazer, tomando por exemplo $f(2,0) = (1,2) \Leftrightarrow f(1,0) = (\frac{1}{2}, 1)$ e $f(0,1) = (-2,1)$ obtemos a transformação linear

   $$f(x,y) = xf(1,0) + yf(0,1) = x(\tfrac{1}{2}, 1) + y(-2,1) = (\tfrac{x}{2} - 2y, x + y)$$

   (b) Como na pergunta 2(b), a imagem do triângulo pela transformação linear é dada pela união de todos os segmentos de reta que unem a origem a um ponto da forma

   $$(1-t)f(1,0) + tf(0,1)$$

   Para que a imagem esteja contida no segmento de reta dado precisamos que $f(1,0)$ e $f(0,1)$ pertençam ao segmento; para que a imagem seja todo o segmento que vai de $(0,0)$ a $(2,2)$ precisamos que pelo menos um dos pontos $f(1,0)$ ou $f(0,1)$ seja $(2,2)$. Tomando, por exemplo, $f(1,0) = (2,2)$ e $f(0,1) = (0,0)$ obtemos

   $$f(x,y) = xf(1,0) + yf(0,1) = (2x, 2x)$$

4. Enunciado Escrevendo $(1,3), (1,1)$ e $(0,3)$ como combinações lineares de $(1,0)$ e $(0,1)$ e usando que f é uma transformação linear, a segunda condição escreve-se

   $$f(1,0) + 3f(0,1) = 2(f(1,0) + f(0,1)) + 3f(0,1) \Leftrightarrow -f(1,0) - 2f(0,1) = 0$$

   o que contradiz a independência linear de $f(1,0)$ e $f(0,1)$. Conclui-se que não existe nenhuma transformação linear satisfazendo as duas condições.

   Se a primeira condição for omitida existe uma tal transformação linear. Por exemplo a transformação linear identicamente nula. Mais geralmente, vimos que a segunda condição é verificada se e só se $f(1,0) = -2f(0,1)$ pelo que as transformações lineares que satisfazem a segunda condição são determinadas por uma escolha arbitrária de $v = f(1,0)$, sendo que $f(0,1)$ é necessariamente igual a $-\frac{1}{2}v$.

5. Enunciado
   (a) Dados polinómios $p, q \in V$ e um escalar α temos

   $$\begin{aligned} f((p+q)(t)) &= (p+q)''(t) - 2(p+q)(t) = p''(t) + q''(t) - 2p(t) - 2q(t) \\ &= p''(t) - 2p(t) + q''(t) - 2q(t) = f(p(t)) + f(q(t)) \end{aligned}$$

e

$$f((\alpha p)(t)) = (\alpha p)''(t) - 2(\alpha p)(t) = \alpha p''(t) - 2\alpha p(t) = \alpha(p''(t) - 2p(t)) = \alpha f(p(t))$$

logo f é uma transformação linear.

(b) A base canónica no espaço dos polinómios é $B = (1, t, t^2, t^3)$. Uma vez que

$$f(1) = 0 - 2 = -2, \quad f(t) = 0 - 2t = -2t, \quad f(t^2) = 2 - 2t^2, \quad f(t^3) = 6t - 2t^3$$

temos que a representação matricial de f é

$$A_{f,B,B} = \begin{bmatrix} -2 & 0 & 2 & 0 \\ 0 & -2 & 0 & 6 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & -2 \end{bmatrix}$$

6. Enunciado

(a) As bases canónicas são $B_1 = ((1,0,0),(0,1,0),(0,0,1))$ e $B_2 = (1)$. Uma vez que

$$T(1,0,0) = 1, \quad T(0,1,0) = -2, \quad T(0,0,1) = 6$$

temos

$$A_{T,B_1,B_2} = \begin{bmatrix} 1 & -2 & 6 \end{bmatrix}$$

(b) Temos $T(1,1) = (-1,0,2,-1)$ e $T(0,1) = (-2,0,1,0)$ logo

$$A_{T,B_1,B_2} = \begin{bmatrix} -1 & -2 \\ 0 & 0 \\ 2 & 1 \\ -1 & 0 \end{bmatrix}$$

(c) Sendo $B_1 = (1, t, t^2)$ e $B_2 = (1, t, t^2, t^3, t^4)$ temos

$$T(1) = 1 - t^2, \quad T(t) = t - t^3, \quad T(t^2) = t^2 - t^4$$

logo

$$A_{T,B_1,B_2} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

(d) Tomando $B_1 = \left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right)$ e $B_2 = \left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right)$ temos

$$T\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right) = \begin{bmatrix} 1 & 1 \\ -1 & 0 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 4 \\ -2 \end{bmatrix} = \begin{bmatrix} 4 \\ -4 \\ 12 \end{bmatrix}, \qquad T\left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \right) = \begin{bmatrix} -2 \\ 2 \\ -6 \end{bmatrix}$$

$$T\left(\begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix}\right)=\begin{bmatrix}4\\0\\8\end{bmatrix},\qquad T\left(\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}\right)=\begin{bmatrix}-2\\0\\-4\end{bmatrix}$$

logo

$$A_{T,B_1,B_2}=\begin{bmatrix}4 & -2 & 4 & -2\\ -4 & 2 & 0 & 0\\ 12 & -6 & 8 & -4\end{bmatrix}$$

7. Enunciado
   (a) Temos

$$T\left(\begin{bmatrix}3 & 1\\ 1 & 0\end{bmatrix}\right)=T\left(\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}\right)+T\left(\begin{bmatrix}1 & 1\\ 0 & 0\end{bmatrix}\right)+T\left(\begin{bmatrix}1 & 0\\ 1 & 0\end{bmatrix}\right)=1+t+(t-2t^2)-3t^2=1+2t-5t^2$$

   (b) Uma vez que

$$T\left(\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}\right)=(t-2t^2)-(1+t)=-1-2t^2,\qquad T\left(\begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix}\right)=-3t^2-(1+t)=-1-t-3t^2$$

   obtemos para $B_1=\left(\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix},\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix},\begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix},\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}\right)$ e $B_2=(1,t,t^2)$

$$A_{T,B_1,B_2}=\begin{bmatrix}1 & -1 & -1 & 4\\ 1 & 0 & -1 & 0\\ 0 & -2 & -3 & 0\end{bmatrix}$$

8. Enunciado
   (a) As funções $e^{-2x},\dots,e^{2x}$ são linearmente independentes (ver Exercício 3.25) logo formam uma base para V. Temos $T(e^{-2x})=4e^{-2x}-2e^{-2x}-2e^{-2x}=0$, $T(e^{-x})=-2e^{-x}$, $T(1)=-2$, $T(e^x)=0$ e $T(e^{2x})=4e^{2x}$. Uma vez que T envia uma base de V para V, envia todo o V para o subespaço V das funções reais de variável real, logo está bem definida.
   (b) Pelos cálculos da alínea anterior temos

$$A_{T,B,B}=\begin{bmatrix}0 & 0 & 0 & 0 & 0\\ 0 & -2 & 0 & 0 & 0\\ 0 & 0 & -2 & 0 & 0\\ 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 4\end{bmatrix}$$

   (c) Em coordenadas, a equação $T(f)=0$ traduz-se na equação matricial

$$\begin{bmatrix}0 & 0 & 0 & 0 & 0\\ 0 & -2 & 0 & 0 & 0\\ 0 & 0 & -2 & 0 & 0\\ 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 4\end{bmatrix}\begin{bmatrix}a\\b\\c\\d\\e\end{bmatrix}=\begin{bmatrix}0\\0\\0\\0\\0\end{bmatrix}$$

que tem solução definida por $b = c = e = 0$. Conclui-se que

$$T^{-1}(0) = \{ae^{-2x} + de^x \colon a, d \in \mathbb{R}\}$$

(d) Em coordenadas, a equação $T(f) = e^{-x}$ fica

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \\ e \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

que tem solução definida por $b = -\frac{1}{2}, c = 0, e = 0$ logo

$$T^{-1}(e^{-x}) = \left\{ae^{-2x} - \tfrac{1}{2}e^{-x} + de^x \colon a, d \in \mathbb{R}\right\}$$

9. EnunciadoSejam $v, w \in V$ e α um escalar. Podemos escrever de forma única $v = \alpha_1 v_1 + \ldots + \alpha_n v_n$ e $w = \beta_1 v_1 + \ldots + \beta_n v_n$. Então $v + w = (\alpha_1 + \beta_1)v_1 + \ldots + (\alpha_n + \beta_n)v_n$ e $\alpha v = (\alpha\alpha_1)v_1 + \ldots + (\alpha\alpha_n)v_n$ logo

$$f(v + w) = [v + w]_B = \begin{bmatrix} \alpha_1 + \beta_1 \\ \vdots \\ \alpha_n + \beta_n \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} + \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix} = f(v) + f(w)$$

e

$$f(\alpha v) = [\alpha v]_B = \begin{bmatrix} \alpha\alpha_1 \\ \vdots \\ \alpha\alpha_n \end{bmatrix} = \alpha \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = \alpha[v]_B$$

logo f é uma transformação linear.

10. Enunciado
   (a) V tem dimensão 3 porque $S_{B_1 \to B_1'}$ é uma matrix 3×3 e analogamente W tem dimensão 2.
   (b) A matriz $A_{T,B_1',B_2'}$ é a única matriz que satisfaz

$$[T(v)]_{B_2'} = A_{T,B_1',B_2'}[v]_{B_1'}$$

para todo o $v \in V$. Uma vez que

$$S_{B_2 \to B_2'} A_{T,B_1,B_2} S_{B_1' \to B_1}[v]_{B_1'} = S_{B_2 \to B_2'} A_{T,B_1,B_2}[v]_{B_1} = S_{B_2 \to B_2'}[T(v)]_{B_2} = [T(v)]_{B_2'}$$

vemos que $A_{T,B_1',B_2'} = S_{B_2 \to B_2'} A_{T,B_1,B_2} S_{B_1' \to B_1}$ e portanto

$$\begin{aligned} A_{T,B_1',B_2'} &= \begin{bmatrix} 1 & -2 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ -1 & 3 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} 3 & -5 & -4 \\ 2 & 6 & 2 \end{bmatrix} \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 3 & -8 & 9 \\ 2 & 4 & -8 \end{bmatrix} \end{aligned}$$

11. Enunciado Se S é linearmente dependente então $T(S)$ não é necessariamente linearmente dependente. Esta afirmação é bastante surpreendente pois seria natural pensar que dados $v_1, \dots, v_n$ vetores distintos de V e $\alpha_1, \dots \alpha_n$ escalares não todos nulos tais que $\alpha_1 v_1 + \dots + \alpha_n v_n = 0$ então, como $\alpha_1 T(v_1) + \dots + \alpha_n T(v_n) = 0$ e os vetores $T(v_1), \dots, T(v_n)$ pertencem a $T(S)$, o conjunto $T(S)$ é linearmente dependente. Este argumento não está correto porque os vetores $T(v_1), \dots, T(v_n)$ podem não ser distintos e essa condição é fundamental na definição de dependência linear.

    De facto a transformação linear $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ definida por $T(x,y) = (x,0)$ (que projeta no eixo dos xx) leva o conjunto linearmente dependente $S = \{(1,1),(1,2),(1,3)\}$ no conjunto $T(S) = \{(1,0)\}$ que é linearmente independente.

    Se S é linearmente independente não há qualquer razão para que $T(S)$ seja. Por exemplo a transformação linear identicamente nula leva qualquer conjunto linearmente independente em $\{0\}$ que é linearmente dependente.

12. Enunciado $L(\mathbb{R}^2, \mathbb{R}^3)$ é isomorfo ao espaço $M_{3\times 2}(\mathbb{R})$ por exemplo pelo isomorfismo que leva $f \in L(\mathbb{R}^2, \mathbb{R}^3)$ em $A_{f,B_{can},B_{can}}$. Como tal, uma base de $L(\mathbb{R}^2, \mathbb{R}^3)$ é dada (por exemplo) pelas transformações lineares que correspondem pelo isomorfismo à base

$$B = \left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \right)$$

    de $M_{2\times 2}(\mathbb{R})$. Estas são, respetivamente,

$$f_1(x,y) = (x,0,0), \quad f_2(x,y) = (y,0,0), \quad f_3(x,y) = (0,x,0)$$

$$f_4(x,y) = (0,y,0), \quad f_5(x,y) = (0,0,x), \quad f_6(x,y) = (0,0,y)$$

    Uma base para $L(\mathbb{R}^2, \mathbb{R}^3)$ é assim, por exemplo, $(f_1, \dots, f_6)$.

13. Enunciado
    (a) Usando as bases $B_1 = (1, t, t^2)$ e $B_2 = (1)$ para V e $\mathbb{R}$ respetivamente, e notando que

$$\phi_0(1) = 1, \quad \phi_0(t) = 0, \quad \phi_0(t^2) = 0$$

    temos

$$A_{\phi_0,B_1,B_2} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$$

e analogamente, uma vez que $\phi_1(1) = 1, \phi_1(t) = 1, \phi_1(t^2) = 1$ e $\phi_2(1) = 1, \phi_2(t) = 2, \phi_2(t^2) = 4$, obtemos

$$A_{\phi_1,B_1,B_2} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \qquad A_{\phi_2,B_1,B_2} = \begin{bmatrix} 1 & 2 & 4 \end{bmatrix}$$

As três matrizes linhas acima são claramente uma base de $M_{1\times 3}(\mathbb{R})$ (pois são um conjunto de três vetores linearmente independentes num espaço de dimensão 3) e correspondem a ϕ_0, ϕ_1, ϕ_2 pelo isomorfismo $L(V, \mathbb{R}) \to M_{1\times 3}(\mathbb{R})$ dado pelas bases B_1 e B_2. Conclui-se que $\{\phi_0, \phi_1, \phi_2\}$ é uma base.

(b) Uma vez que $\psi(1) = 0, \psi(t) = 1$ e $\psi(t^2) = 2$ temos

$$A_{\psi,B_1,B_2} = \begin{bmatrix} 0 & 1 & 2 \end{bmatrix}$$

Portanto $A_{\psi,B_1,B_2} = \frac{1}{2}\left(A_{\phi_2,B_1,B_2} - A_{\phi_0,B_1,B_2}\right)$ e consequentemente $\psi = -\frac{1}{2}\phi_0 + \frac{1}{2}\phi_2$. Conclui-se que

$$[\psi]_{(\phi_0,\phi_1,\phi_2)} = \begin{bmatrix} -\frac{1}{2} \\ 0 \\ \frac{1}{2} \end{bmatrix}$$

14. Enunciado

(a) Uma vez que ser linearmente independente é a negação de ser linearmente dependente as afirmações relativas à preservação por f da dependência e independência linear são logicamente equivalentes.

Seja $S \subset V$ um conjunto linearmente dependente. Vamos começar por ver que $f(S)$ é linearmente dependente. Sejam $v_1, \ldots, v_n$ vetores distintos de S e $\alpha_1, \ldots, \alpha_n$ escalares não todos nulos tais que $\alpha_1 v_1 + \ldots + \alpha_n v_n = 0$. Então $f(\alpha_1 v_1 + \ldots + \alpha_n v_n) = \alpha_1 f(v_1) + \ldots + \alpha_n f(v_n) = 0$. Como f é uma função injetiva $f(v_1), \ldots, f(v_n)$ são vetores distintos de $f(S)$ e portanto a igualdade anterior mostra que $f(S)$ é um conjunto linearmente dependente.

Reciprocamente se $f(S)$ é um conjunto linearmente dependente em W, então, dado que f^{-1} é também um isomorfismo, o parágrafo anterior permite-nos concluir que $f^{-1}(f(S)) = S$ é um conjunto linearmente dependente em V.

(b) Suponhamos que S gera V e vejamos que $f(S)$ gera W. Temos a mostrar que todo o $w \in W$ pode ser expresso como uma combinação linear de elementos de $f(S)$. Seja $w \in W$. Uma vez que f é uma função sobrejetiva, existe $v \in V$ tal que $f(v) = w$. Como S gera V, existem $v_1, \ldots, v_n \in S$ e escalares $\alpha_1, \ldots, \alpha_n$ tais que $v = \alpha_1 v_1 + \ldots + \alpha_n v_n$. Então

$$w = f(v) = \alpha_1 f(v_1) + \ldots + \alpha_n f(v_n) \in L(f(S))$$

Isto mostra que $W = L(f(S))$ conforme pretendido.

Reciprocamente, suponhamos que $f(S)$ gera W. Então aplicando o parágrafo anterior ao isomorfismo $f^{-1}\colon W \to V$ vemos que $f^{-1}(f(S)) = S$ gera V.

(c) Suponhamos que V é finitamente gerado. Então existe um conjunto finito $S \subset V$ tal que $V = L(S)$. Pela alínea (b) temos que $W = L(f(S))$. Uma vez que S é finito, a sua imagem $f(S)$ é também um conjunto finito logo W é finitamente gerado. A afirmação recíproca obtém-se como acima aplicando a implicação já demonstrada ao isomorfismo f^{-1}.

Se B é uma base para V então B é um conjunto linearmente independente que gera V. Logo pela alínea (a), $f(B)$ é um conjunto linearmente independente de W e, pela alínea (b) $f(B)$ gera W, logo $f(B)$ é uma base para W. Como V é finitamente gerado, B é necessariamente um conjunto finito e, uma vez que f é uma função bijetiva, B e $f(B)$ têm o mesmo número de elementos. Conclui-se que V e W têm a mesma dimensão.

(d) $f(U)$ é a imagem da restrição de f a U, $f_{|U}\colon U \to W$ logo é um subespaço vetorial. Por definição $f_{|U}\colon U \to f(U)$ é sobrejectiva. Como f é um isomorfismo, f é injetiva e portanto $f_{|U}$ é também uma função injetiva. Conclui-se que $f_{|U}\colon U \to f(U)$ é um isomorfismo.

Finalmente, se $f(U)$ é um subespaço de W então $f^{-1}(f(U)) = U$ é um subespaço de V.

15. Enunciado

(a) O conjunto $\mathbb{R} \cdot S$ é não vazio uma vez que a função constante igual a 0 tem suporte vazio e portanto pertence a $\mathbb{R} \cdot S$.

Sejam f, g elementos de $\mathbb{R} \cdot S$ e $\alpha \in \mathbb{R}$. Temos a verificar que as funções $f + g$ e αf estão em $\mathbb{R} \cdot S$, isto é, que têm suporte finito.

Uma vez que se $f(x) = 0$ então $(\alpha f)(x) = \alpha f(x) = 0$ vemos que o suporte de αf está contido no suporte de f (na realidade, $\operatorname{supp}(\alpha f) = \operatorname{supp}(f)$ se $\alpha \neq 0$ e $\operatorname{supp}(\alpha f) = \emptyset$ se $\alpha = 0$). Portanto $\operatorname{supp}(\alpha f)$ é também finito, ou seja $\alpha f \in \mathbb{R} \cdot S$.

Analogamente, se $f(x) = 0$ e $g(x) = 0$ temos que $(f + g)(x) = 0$ e portanto, se $(f + g)(x) \neq 0$ então, ou $f(x) \neq 0$ ou $g(x) \neq 0$. Isto mostra que $\operatorname{supp}(f + g) \subset \operatorname{supp}(f) \cup \operatorname{supp}(g)$. Como $\operatorname{supp}(f + g)$ está contido numa união de dois conjuntos finitos é um conjunto finito e portanto $f + g \in \mathbb{R} \cdot S$.

(b) Temos a ver que o conjunto $\{\phi_x \colon x \in S\}$ gera $\mathbb{R} \cdot S$ e é linearmente independente. Seja $f \in \mathbb{R} \cdot S$ e suponhamos que $\operatorname{supp}(f) = \{x_1, \ldots, x_k\}$ (com $x_i \in S$ distintos). Então existem $\alpha_i \in \mathbb{R}$ não nulos tais que

$$f(x) = \begin{cases} \alpha_i & \text{se } x = x_i \\ 0 & \text{se } x \notin \{x_1, \ldots, x_k\} \end{cases}$$

Vejamos que $f = \alpha_1\phi_{x_1} + \ldots + \alpha_k\phi_{x_k}$. Para isto basta calcular o valor de cada uma das duas funções em $x \in S$: Se $x = x_i$ temos

$$\alpha_1\phi_{x_1}(x_i) + \ldots + \alpha_i\phi_{x_i}(x_i) + \ldots \alpha_n\phi_{x_n}(x_i) = \alpha_1 \cdot 0 + \ldots + \alpha_i \cdot 1 + \ldots \alpha_n \cdot 0 = \alpha_i = f(x_i)$$

Se $x \notin \{x_1, \ldots, x_k\}$ então $\phi_{x_i}(x) = 0$ para todo o $i = 1, \ldots, k$ logo

$$\alpha_1\phi_{x_1}(x) + \ldots + \alpha_k\phi_{x_k}(x) = \alpha_1 0 + \ldots + \alpha_k 0 = 0 = f(x)$$

Conclui-se que f e $\alpha_1\phi_{x_1} + \ldots + \alpha_k\phi_{x_k}$ tomam o mesmo valor em todos os elementos $x \in S$, logo são iguais. Portanto $\{\phi_x \colon x \in S\}$ gera $\mathbb{R} \cdot S$.

Para ver que o conjunto $B = \{\phi_x \colon x \in S\}$ é linearmente independente, suponhamos que $\phi_{x_1}, \ldots, \phi_{x_n}$ são elementos distintos de B, ou seja, que os elementos $x_1, \ldots, x_n \in$

S são distintos. Sejam $\alpha_1, \ldots, \alpha_n$ escalares tais que $\alpha_1\phi_{x_1} + \ldots + \alpha_n\phi_{x_n} = 0$. Então avaliando a combinação linear $\alpha_1\phi_{x_1} + \ldots + \alpha_n\phi_{x_n}$ no ponto x_i obtemos

$$\alpha_1\phi_{x_1}(x_i) + \ldots + \alpha_i\phi_{x_i}(x_i) + \ldots + \alpha_n\phi_{x_n}(x_i) = \alpha_1 \cdot 0 + \ldots + \alpha_i \cdot 1 + \ldots + \alpha_n \cdot 0 = \alpha_i$$

Logo $\alpha_1\phi_{x_1} + \ldots + \alpha_k\phi_{x_n} = 0 \Rightarrow \alpha_i = 0$ para cada $i = 1, \ldots, n$ e portanto B é linearmente independente.

(c) O isomorfismo natural é a transformação linear $\Psi \colon \mathbb{R} \cdot B \to V$ que envia o elemento $\phi_b \in \mathbb{R} \cdot B$ no elemento $b \in B$ em V. Mais precisamente definimos

$$\Psi\left(\alpha_1\phi_{b_1} + \ldots + \alpha_k\phi_{b_k}\right) = \alpha_1 b_1 + \ldots + \alpha_k b_k$$

onde $\alpha_i \in \mathbb{R}$ e b_i são elementos distintos de B. Esta função associa a uma função $f \in \mathbb{R} \cdot B$ com suporte $\{b_1, \ldots, b_k\}$ o elemento $f(b_1)b_1 + \ldots f(b_k)b_k \in V$ e pode por isso escrever-se

$$\Psi(f) = \sum_{b \in B} f(b)b$$

(onde na soma os termos com coeficientes nulos são ignorados). Em termos desta expressão é claro que Ψ é uma transformação linear: dados α, β escalares e $f, g \in \mathbb{R} \cdot B$ temos

$$\begin{aligned} \Psi(\alpha f + \beta g) &= \sum_{b \in B} (\alpha f + \beta g)(b) = \sum_{b \in B} (\alpha f(b) + \beta g(b)) \\ &= \sum_{b \in B} \alpha f(b) + \sum_{b \in B} \beta g(b) = \alpha\Psi(f) + \beta\Psi(g) \end{aligned}$$

Uma vez que B gera V e está contido na imagem de Ψ (note-se que $b = \Psi(\phi_b)$) a função Ψ é sobrejetiva. Uma vez que B é linearmente independente, o núcleo de Ψ é constituído apenas pela função nula. Conclui-se que Ψ é também injetiva e portanto um isomorfismo.

(d) Uma transformação linear é determinada pelo efeito que tem numa base e vimos na alínea (b) uma base de $\mathbb{R} \cdot S$ que se identifica naturalmente com S. Seja $\Phi \colon L(\mathbb{R} \cdot S, V) \to F(S, V)$ a função definida por

$$(\Phi(T))(x) = T(\phi_x) \qquad (\text{onde } T \in L(\mathbb{R} \cdot S, V), x \in S)$$

Então é imediato verificar que Φ é linear por definição das operações em $L(\mathbb{R} \cdot S, V)$ e $F(S, V)$. Se $\Phi(T) = 0$ então T anula-se na base $\{\phi_x \colon x \in S\}$ e é por isso a transformação linear nula. Conclui-se que Φ é injetiva. Dada uma função $f \colon S \to V$ seja $T \colon \mathbb{R} \cdot S \to V$ a transformação linear definida por

$$T(g) = \sum_{x \in S} g(s)f(s)$$

onde $g \in \mathbb{R} \cdot S$ (como antes, a soma está bem definida porque o suporte de g é finito). Então, como na alínea (c), $T \colon \mathbb{R} \cdot S \to V$ é uma transformação linear, e para todo o $x \in S$ temos

$$\Phi(T)(x) = T(\phi_x) = \sum_{y \in S} \phi_x(y) f(y) = \phi_x(x) f(x) = f(x)$$

Assim $\Phi(T) = f$, o que mostra que Φ é sobrejetiva e portanto um isomorfismo.

16. Enunciado
   (a) Temos

$$A_{f,B_{can},B_{can}} = \begin{bmatrix} 1 & 2 \\ 0 & 0 \\ -2 & -4 \end{bmatrix}$$

O núcleo desta matriz é determinado pela equação $x + 2y = 0$ pelo que $\{(-2,1)\}$ é uma base para o núcleo. Como o núcleo é não trivial f não é injetiva. Claramente também não é sobrejetiva (por exemplo $(0,1,0)$ não está na imagem). Na realidade nunca poderia ser sobrejetiva, a imagem de qualquer transformação linear $f\colon \mathbb{R}^2 \to \mathbb{R}^3$ tem no máximo dimensão 2 como se vê por exemplo aplicando o Teorema da característica-nulidade. Neste caso específico, o Teorema da característica-nulidade diz que a imagem tem dimensão $\dim(\mathbb{R}^2) - \dim N(f) = 2 - 1 = 1$. Uma base para a imagem é por exemplo $\{f(1,0)\} = \{(1,0,-2)\}$

   (b) A representação matricial de f com respeito às bases canónicas é

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

Claramente f é sobrejetiva (o espaço das colunas da matriz é $\mathbb{R}^2$). Uma base para a imagem é por exemplo $\{(1,0),(0,1)\}$. Pelo Teorema da característica nulidade a dimensão do núcleo é $4 - 2 = 2$ pelo que f não é injetiva. Uma base para o núcleo é por exemplo $\{(1,-1,0,0),(0,0,1,-1)\}$ uma vez que estes vetores estão no núcleo e não são colineares.

   (c) Temos

$$A_{f,B_{can},B_{can}} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 2 & 2 & 0 \end{bmatrix}$$

Aplicando o método de Gauss temos

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 2 & 2 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 2 & -2 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

logo o núcleo é definido pelas equações $x + z = 0$, $y - z = 0$ e uma base é dada por $\{(1,-1,-1)\}$. Como o núcleo não é trivial a função não é injetiva. Pelo Teorema da característica-nulidade a imagem de f tem dimensão 2 (e portanto f não é sobrejetiva). Uma vez que $f(1,0,0) = (1,0,2)$ e $f(0,1,0) = (0,1,2)$ não são colineares, constituem uma base para $\operatorname{Im} f$.

(d) A representação matricial de f com respeito às bases canónicas é

$$A_{f,B_{can},B_{can}} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix}$$

Aplicando o método de Gauss obtemos

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & -1 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & -1 & -1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

logo a matriz é invertível e portanto f é injetiva e sobrejetiva. Sendo assim a base para o núcleo é o conjunto vazio e uma base para imagem é por exemplo a base canónica de $\mathbb{R}^3$. Seria fácil aplicar o método de Gauss-Jordan para calcular a matriz inversa e portanto a transformação inversa mas neste caso é ainda mais rápido aplicar a definição de função inversa:

$$\begin{cases} u = y \\ v = x + z \\ w = x - y \end{cases} \Leftrightarrow \begin{cases} y = u \\ v = x + z \\ x = w + u \end{cases} \Leftrightarrow \begin{cases} y = u \\ z = v - (w + u) = v - w - u \\ x = w + u \end{cases}$$

A transformação inversa é portanto dada pela expressão $f^{-1}(u, v, w) = (w+u, u, v-w-u)$.

17. Enunciado

(a) A matriz que representa f é

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 2 & 3 \end{bmatrix}$$

Usando o método de Gauss obtemos

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & 3 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 4 \end{bmatrix}$$

Uma vez que a matriz $A_{f,B,B}$ tem característica 3 ela é invertível. Logo f é invertível e portanto um isomorfismo.

(b) Uma vez que $f(1, -1, -1) = (0, 0, 0)$ o núcleo de f não é trivial e portanto f não é um isomorfismo.

(c) Aplicando o método de Gauss à matriz que representa f com respeito às bases canónicas obtemos

$$\begin{bmatrix} 1 & -1 & 1 & 1 & 1 \\ -1 & 2 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & -2 \\ 1 & 4 & 0 & 1 & 0 \\ 0 & -2 & 3 & -2 & -1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 2 & 1 \\ 0 & 0 & 1 & 1 & -2 \\ 0 & 5 & -1 & 0 & -1 \\ 0 & -2 & 3 & -2 & -1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 2 & 1 \\ 0 & 0 & 1 & 1 & -2 \\ 0 & 0 & -6 & -10 & -6 \\ 0 & 0 & 5 & 2 & 1 \end{bmatrix}$$

$$\to \begin{bmatrix} 1 & -1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 2 & 1 \\ 0 & 0 & 1 & 1 & -2 \\ 0 & 0 & 0 & -4 & -18 \\ 0 & 0 & 0 & -3 & 11 \end{bmatrix}$$

Uma vez que a matriz tem característica 5, f é um isomorfismo

(d) Temos

$$f(1) = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \quad f(t) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad f(t^2) = \begin{bmatrix} 0 & 0 \\ 3 & 0 \end{bmatrix}, \quad f(t^3) = \begin{bmatrix} 0 & 0 \\ 1 & 6 \end{bmatrix}$$

logo com respeito às bases $B_1 = (1, t, t^2, t^3)$ e $B_2 = \left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right)$ temos

$$A_{f,B_1,B_2} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 3 & 1 \\ 0 & 0 & 1 & 6 \end{bmatrix}$$

É imediato verificar que esta matriz tem característica 4 pelo que f é um isomorfismo.

(e) Não, a dimensão do espaço das colunas de uma tal matriz tem dimensão no máximo 3 logo a matriz não é invertível.

(f) Não porque os polinómios constantes pertencem ao núcleo da transformação.

(g) A transformação linear é injetiva por que o único polinómio que é igual à sua derivada é o polinómio constante igual a zero. Uma vez que $T(-t^k - kt^{k-1} - k(k-1)t^{k-2} - \ldots - k!) = t^k$, a imagem de T contém uma base do espaço dos polinómios. Logo T é sobrejetiva e portanto um isomorfismo.

18. Enunciado

(a) 2 e 3 respetivamente, dadas as dimensões da matriz que representa h.

(b) Apenas f pois é representada por uma matriz de característica 2 que é portanto invertível. A matriz que representa g tem característica 1 e h não pode ser um isomorfismo uma vez que os espaços de partida e chegada têm dimensões diferentes.

(c) A representação matricial de $h \circ (f + 3(g \circ g))$ é dada por

$$\begin{aligned} A_{h,B_1,B_2}\left(A_{f,B_1,B_1} + 3A^2_{g,B_1,B_1}\right) &= \begin{bmatrix} 2 & 1 \\ 3 & 0 \\ -1 & 1 \end{bmatrix} \left(\begin{bmatrix} 1 & 1 \\ -2 & 1 \end{bmatrix} + 3 \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}^2 \right) \\ &= \begin{bmatrix} 2 & 1 \\ 3 & 0 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 16 & 31 \\ 28 & 61 \end{bmatrix} = \begin{bmatrix} 60 & 123 \\ 48 & 93 \\ 12 & 30 \end{bmatrix} \end{aligned}$$

A transformação f^{-1} é representada pela matrix

$$A_{f^{-1},B_1,B_1} = \tfrac{1}{3} \begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix}$$

logo a representação matricial de $(f+f^{-1})\circ g$ é dada por

$$\begin{aligned}\left(A_{f,B_1,B_1}+A^{-1}_{f,B_1,B_1}\right)A_{g,B_1,B_1} &= \left(\begin{bmatrix}1 & 1\\ -2 & 1\end{bmatrix}+\begin{bmatrix}\frac{1}{3} & -\frac{1}{3}\\ \frac{2}{3} & \frac{1}{3}\end{bmatrix}\right)\begin{bmatrix}1 & 2\\ 2 & 4\end{bmatrix}\\ &= \begin{bmatrix}\frac{4}{3} & \frac{2}{3}\\ \frac{4}{3} & \frac{4}{3}\end{bmatrix}\begin{bmatrix}1 & 2\\ 2 & 4\end{bmatrix}=\frac{1}{3}\begin{bmatrix}8 & 16\\ 12 & 24\end{bmatrix}\end{aligned}$$

19. Enunciado

(a) Uma vez que $1+t=(1-t^2)+(t+t^2)$ temos $f(1+t)=f(1-t^2)+f(t+t^2)=(1,0)+(0,-1)=(1,-1)$.

(b) Achamos as coordenadas dos vetores da base canónica na base B_1. Temos

$$[t]_{B_1}=\begin{bmatrix}0\\1\\0\end{bmatrix}$$

Como $t^2=-t+(t+t^2)$ temos

$$[t^2]_{B_1}=\begin{bmatrix}0\\-1\\1\end{bmatrix}$$

e finalmente

$$[1]_{B_1}=[1-t^2]_{B_1}+[t^2]_{B_1}=\begin{bmatrix}1\\0\\0\end{bmatrix}+\begin{bmatrix}0\\-1\\1\end{bmatrix}=\begin{bmatrix}1\\-1\\1\end{bmatrix}$$

Conclui-se que

$$S_{B_{can}\to B_1}=\begin{bmatrix}1 & 0 & 0\\ -1 & 1 & -1\\ 1 & 0 & 1\end{bmatrix}$$

e

$$A_{f,B_{can},B_{can}}=A_{f,B_1,B_{can}}S_{B_{can}\to B_1}=\begin{bmatrix}1 & 1 & 0\\ 0 & 2 & -1\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\ -1 & 1 & -1\\ 1 & 0 & 1\end{bmatrix}=\begin{bmatrix}0 & 1 & -1\\ -3 & 2 & -3\end{bmatrix}$$

(c) Temos $A_{g\circ f,B_1,B_{can}}=A_{g,B_{can},B_{can}}A_{f,B_1,B_{can}}$. Uma vez que

$$g(1,0)=g(2,1)-g(1,1)=\begin{bmatrix}0 & 1\\ 1 & -1\end{bmatrix}-\begin{bmatrix}1 & 0\\ 2 & 3\end{bmatrix}=\begin{bmatrix}-1 & 1\\ -1 & -4\end{bmatrix}$$

e

$$g(0,1)=g(1,1)-g(1,0)=\begin{bmatrix}1 & 0\\ 2 & 3\end{bmatrix}-\begin{bmatrix}-1 & 1\\ -1 & -4\end{bmatrix}=\begin{bmatrix}2 & -1\\ 3 & 7\end{bmatrix}$$

temos

$$A_{g,B_{can},B_{can}}=\begin{bmatrix}-1 & 2\\ 1 & -1\\ -1 & 3\\ -4 & 7\end{bmatrix}$$

e

$$A_{g\circ f,B_1,B_{can}} = \begin{bmatrix} -1 & 2 \\ 1 & -1 \\ -1 & 3 \\ -4 & 7 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & -1 \end{bmatrix} = \begin{bmatrix} -1 & 3 & -2 \\ 1 & -1 & 1 \\ -1 & 5 & -3 \\ -4 & 10 & -7 \end{bmatrix}$$

20. Enunciado

(a) $f(1,0) = 2f(1,1) - f(1,2) = 4 + 2t - (1-t^2) = 3 + 2t + t^2$ logo

$$(g\circ f)(1,0) = g(3+2t+t^2) = \begin{bmatrix} 0 & 6 \\ 3 & 2 \end{bmatrix}$$

(b) Temos $f(0,1) = f(1,2) - f(1,1) = -1 - t - t^2$ logo

$$(g\circ f)(0,1) = \begin{bmatrix} 0 & -3 \\ -1 & -1 \end{bmatrix}$$

e portanto

$$(g\circ f)(x,y) = x(g\circ f)(1,0) + y(g\circ f)(0,1) = x\begin{bmatrix} 0 & 6 \\ 3 & 2 \end{bmatrix} + y\begin{bmatrix} 0 & -3 \\ -1 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 6x-3y \\ 3x-y & 2x-y \end{bmatrix}$$

(c) Pela alínea anterior $(g\circ f)(1,1) = \begin{bmatrix} 0 & 3 \\ 2 & 1 \end{bmatrix}$ e $(g\circ f)(1,-1) = \begin{bmatrix} 0 & 9 \\ 4 & 3 \end{bmatrix}$ logo

$$A_{g\circ f,B_1,B_{can}} = \begin{bmatrix} 0 & 0 \\ 3 & 9 \\ 2 & 4 \\ 1 & 3 \end{bmatrix}$$

(d) Uma vez que as duas colunas da matriz da alínea (b) não são colineares, formam uma base para o espaço das colunas. Uma vez que o espaço das colunas é a tradução em coordenadas da imagem de uma transformação linear vemos que

$$B_2 = \left(\begin{bmatrix} 0 & 3 \\ 2 & 1 \end{bmatrix}, \begin{bmatrix} 0 & -9 \\ -4 & -3 \end{bmatrix} \right)$$

é uma base para $W = (g\circ f)(\mathbb{R}^2)$.

(e) A aplicação $g\circ f\colon \mathbb{R}^2 \to W$ é um isomorfismo já que é uma transformação linear sobrejetiva entre espaços de dimensão 2. Uma vez que $B_2 = ((g\circ f)(1,1),(g\circ f)(1,-1))$ temos que $(g\circ f)^{-1}$ envia esta base em $(1,1)$ e $(1,-1)$ logo

$$A_{(g\circ f)^{-1},B_2,B_{can}} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

21. Enunciado

(a) Pelo Teorema da característica-nulidade temos

$$\dim N(f) + \dim f(\mathbb{R}^7) = 7$$

Como $f(\mathbb{R}^7) \subset \mathbb{R}^4$ este subespaço tem dimensão ≤ 4, logo

$$(\dim N(f), \dim f(\mathbb{R}^7)) \in \{(3,4),(4,3),(5,2),(6,1),(7,0)\}$$

Todos estes valores são de facto possíveis: podemos por exemplo tomar as transformações

$$f_1(x_1,\ldots,x_7)=(x_1,x_2,x_3,x_4),\quad f_2(x_1,\ldots,x_7)=(x_1,x_2,x_3,0),$$
$$f_3(x_1,\ldots,x_7)=(x_1,x_2,0,0),\quad f_4(x_1,\ldots,x_7)=(x_1,0,0,0)$$
$$f_4(x_1,\ldots,x_7)=(0,0,0,0)$$

(b) Pelo Teorema da característica-nulidade temos

$$\dim N(f)+\dim f(\mathbb{R}^3)=3$$

logo

$$(\dim N(f),\dim f(\mathbb{R}^3))\in\{(0,3),(1,2),(2,1),(3,0)\}$$

Todos estes valores são possíveis. Podemos por exemplo considerar as transformações

$$f_1(x_1,x_2,x_3)=(x_1,x_2,x_3,0,0),\quad f_2(x_1,x_2,x_3)=(x_1,x_2,0,0,0),$$
$$f_3(x_1,x_2,x_3)=(x_1,0,0,0,0),\quad f_4(x_1,x_2,x_3)=(0,0,0,0,0)$$

22. Enunciado Para $V=\mathbb{R}^2$, pelo Teorema da característica-nulidade temos necessariamente $\dim N(f)=\dim f(\mathbb{R}^2)=1$. Se $(a,b)\neq(0,0)$ for um gerador da imagem, f será representada por uma matriz da forma

$$\begin{bmatrix}\alpha a & \beta a\\ \alpha b & \beta b\end{bmatrix}$$

para alguns $\alpha,\beta\in\mathbb{R}$ (pois a imagem de f corresponde ao espaço das colunas da matriz). Para que (a,b) esteja no núcleo de f é necessário que

$$\begin{bmatrix}\alpha a & \beta a\\ \alpha b & \beta b\end{bmatrix}\begin{bmatrix}a\\ b\end{bmatrix}=0\Leftrightarrow\begin{cases}a(\alpha a+\beta b)=0\\ b(\alpha a+\beta b)=0\end{cases}\Leftrightarrow\alpha a+\beta b=0$$

Podemos assim tomar por exemplo $(a,b)=(1,1)$ e $\alpha=1,\beta=-1$ obtendo $f\colon\mathbb{R}^2\to\mathbb{R}^2$ dada pela expressão

$$f(x,y)=(x-y,x-y)$$

23. Enunciado Consideremos a base $B=\left(\begin{bmatrix}1&0\\0&0\end{bmatrix},\begin{bmatrix}0&1\\0&0\end{bmatrix},\begin{bmatrix}0&0\\1&0\end{bmatrix},\begin{bmatrix}0&0\\0&1\end{bmatrix}\right)$ para $M_{2\times2}(\mathbb{R})$. Temos

$$f\left(\begin{bmatrix}1&0\\0&0\end{bmatrix}\right)=\begin{bmatrix}1&0\\0&0\end{bmatrix}-\begin{bmatrix}1&1\\0&0\end{bmatrix}=\begin{bmatrix}0&-1\\0&0\end{bmatrix},\quad f\left(\begin{bmatrix}0&1\\0&0\end{bmatrix}\right)=\begin{bmatrix}0&1\\0&0\end{bmatrix}-\begin{bmatrix}0&1\\0&0\end{bmatrix}=\begin{bmatrix}0&0\\0&0\end{bmatrix}$$
$$f\left(\begin{bmatrix}0&0\\1&0\end{bmatrix}\right)=\begin{bmatrix}1&0\\1&0\end{bmatrix}-\begin{bmatrix}0&0\\1&1\end{bmatrix}=\begin{bmatrix}1&0\\0&-1\end{bmatrix},\quad f\left(\begin{bmatrix}0&0\\0&1\end{bmatrix}\right)=\begin{bmatrix}0&1\\0&1\end{bmatrix}-\begin{bmatrix}0&0\\0&1\end{bmatrix}=\begin{bmatrix}0&1\\0&0\end{bmatrix}$$

Claramente a imagem de f é gerada pelos dois vetores não colineares $\begin{bmatrix}1&0\\0&-1\end{bmatrix}$ e $\begin{bmatrix}0&1\\0&0\end{bmatrix}$ logo tem dimensão 2. Pelo Teorema da característica-nulidade, o núcleo de f tem também dimensão 2.

Uma vez que $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ e $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ estão no núcleo e são linearmente independentes conclui-se que

$$f(M_{2\times 2}(\mathbb{R})) = L\left(\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right), \qquad N(f) = L\left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right).$$

24. Enunciado
   (a) Uma vez que $S^{-1} = S_{B_2\to B_1}$, dado $v \in V$ temos
   $$SAS^{-1}[v]_{B_2} = SA[v]_{B_1} = S[f(v)]_{B_1} = [f(v)]_{B_2}$$
   logo SAS^{-1} representa f com respeito à base B_2.
   (b) Usando a alínea anterior com $B_1 = B_{can}$ e $B_2 = B$ temos
   $$S^{-1} = S_{B\to B_1} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$
   logo
   $$S = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}^{-1} = \frac{1}{2}\begin{bmatrix} 1 & -1 & 1 \\ 0 & 2 & 0 \\ -1 & -1 & 1 \end{bmatrix}$$
   e
   $$\begin{aligned} A_{f,B,B} &= \frac{1}{2}\begin{bmatrix} 1 & -1 & 1 \\ 0 & 2 & 0 \\ -1 & -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & -2 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \\ &= \frac{1}{2}\begin{bmatrix} 0 & -1 & 0 \\ 2 & 0 & 2 \\ -2 & 3 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 0 & -1 & 0 \\ 4 & 2 & 0 \\ -2 & 3 & 2 \end{bmatrix} \end{aligned}$$
   (c) Não. Sendo A a primeira matriz e B a segunda, se representassem a mesma transformação linear em bases distintas teríamos $SAS^{-1} = B$ para alguma matriz invertível S. Mas A, e portanto SAS^{-1} é uma matriz invertível, enquanto que B tem característica 1.
25. Enunciado
   (a) Dado $v \in V$ temos
   $$\begin{aligned} (\mathrm{Id} - P)^2(v) &= (\mathrm{Id} - P)\circ(\mathrm{Id} - P)(v) = (\mathrm{Id} - P)(v - P(v)) \\ &= v - P(v) - P(v) + P^2(v) = v - P(v) - P(v) + P(v) \\ &= v - P(v) = (\mathrm{Id} - P)(v) \end{aligned}$$
   logo $(\mathrm{Id} - P)^2 = (\mathrm{Id} - P)$ e portanto $\mathrm{Id} - P$ é uma projeção.
   (b) Para estabelecer a igualdade $P(V) = N(\mathrm{Id} - P)$ vamos verificar as duas inclusões. Dado $w = P(v)$ em $P(V)$ temos
   $$(\mathrm{Id} - P)(w) = (\mathrm{Id} - P)(P(v)) = P(v) - P^2(v) = P(v) - P(v) = 0$$

logo $w \in N(\mathrm{Id} - P)$, o que mostra que $P(V) \subset N(\mathrm{Id} - P)$. Reciprocamente se $v \in N(\mathrm{Id} - P)$ então $(\mathrm{Id} - P)(v) = 0 \Leftrightarrow v - P(v) = 0 \Leftrightarrow P(v) = v$ logo v é a imagem de si próprio por P. Conclui-se que $N(\mathrm{Id} - P) \subset P(V)$, e portanto que $P(V) = N(\mathrm{Id} - P)$.

Uma vez que $\mathrm{Id} - P$ também é uma projeção (pela alínea (a)), o parágrafo anterior mostra que $(\mathrm{Id} - P)(V) = N(\mathrm{Id} - (\mathrm{Id} - P)) = N(P)$.

(c) Recorde da Proposição 3.40 que temos a mostrar que $V = P(V) + (\mathrm{Id} - P)(V)$ e que $P(V) \cap (\mathrm{Id} - P)(V) = \{0\}$. A primeira afirmação é imediata uma vez que, dado $v \in V$ temos

$$v = P(v) + (v - P(v)) = P(v) + (\mathrm{Id} - P)(v) \in P(V) + (\mathrm{Id} - P)(V)$$

Pela alínea (b) resta mostrar que $P(V) \cap N(P) = 0$. Seja $w = P(v)$ um elemento de $P(V)$. Se w pertence a $N(P)$ então temos $P(w) = 0 \Leftrightarrow P(P(v)) = 0 \Leftrightarrow P(v) = 0 \Leftrightarrow w = 0$. Logo $P(V) \cap N(P) = 0$.

(d) Se $V = U \oplus W$ cada elemento $v \in V$ pode escrever-se de forma única como $v = u + w$ com $u \in U$ e $w \in W$. Se $P \colon V \to V$ for uma projeção satisfazendo as condições do enunciado então $P(v) = P(u) + P(w) = P(u) + 0 = P(u)$ (uma vez que w pertence ao núcleo de P. Por outro lado, como $u \in \mathrm{Im}(P)$ temos $u = P(x) \Rightarrow P(u) = P^2(x) = P(x) = u$ logo $P(v) = u$ (isto é, P associa necessariamente a v o único vetor $u \in U$ tal que $v = u + w$ com $u \in U$ e $w \in W$). Vejamos que esta fórmula define a projeção requerida:

- P é uma transformação linear: Sejam α, β escalares e $v_1, v_2 \in V$. Escrevendo $v_1 = u_1 + w_1, v_2 = u_2 + w_2$ com $u_i \in U$ e $w_i \in W$ temos

$$\alpha v_1 + \beta v_2 = \alpha(u_1 + w_1) + \beta(u_2 + w_2) = (\alpha u_1 + \beta u_2) + (\alpha w_1 + \beta w_2)$$

  que é uma decomposição de $\alpha v_1 + \beta v_2$ como uma soma de um vetor em U com um vetor em W. Conclui-se que $P(\alpha v_1 + \beta v_2) = \alpha u_1 + \beta u_2 = \alpha P(v_1) + \beta P(v_2)$ pelo que P é uma transformação linear.
- P é uma projeção com imagem U e núcleo W: Por definição, para todo o $v \in V$ temos $P(v) \in U$. Portanto a decomposição de $P(v)$ como uma soma de um vetor em U com um vetor em W é $P(v) = P(v) + 0$, donde $P(P(v)) = P(v)$. A imagem de P está contida em U por definição e se $u \in U$ temos $P(u) = u$ logo $P(V) = U$. Finalmente escrevendo um vetor $v \in V$ como $v = u + w$ vemos que $P(v) = 0 \Leftrightarrow u = 0 \Leftrightarrow v = w$ logo $N(P) = W$.

(e) Vimos na alínea anterior que uma projeção P permite decompor um vetor $v \in V$ de forma única na forma $v = u + w$ com $u = P(v) \in P(V)$ e $w = v - P(v) \in N(P)$, e, descrevendo v nestes termos, a projeção é dada por $P(v) = u$.

Geometricamente, a projeção de v é o único elemento do plano $P(V)$ que pertence ao plano $v + N(P)$, pelo que P projeta um vetor de V no plano $P(V)$ ao longo das direções contidas no plano $N(P)$.

Para calcular uma projeção podemos usar uma base adaptada à situação, em que os vetores da base pertençam todos a $N(P)$ ou a $P(V)$ (cf. Proposição 3.40). Claro que P tem de enviar os vetores de $N(P)$ para 0. Quanto aos vetores de $P(V)$, têm de ser necessariamente enviados para si mesmos.

(i) Uma base para a reta (que será a imagem da projeção) é $(-2,1)$ e a direção de projeção é uma base para o núcleo de P. Conclui-se que a projeção é determinada por

$$P(-2,1) = (-2,1), \qquad P(1,1) = (0,0)$$

Uma vez que $P(1,0) = \frac{1}{3}(P(1,1) - P(-2,1)) = \frac{1}{3}(2,-1)$ e $P(0,1) = P(1,1) - P(1,0) = \frac{1}{3}(-2,1)$ vemos que $P(x,y) = \frac{1}{3}(2x-2y, -x+y)$.

(ii) Uma base para o plano é dada por $(1,1,0)$ e $(0,1,1)$ já que estes são dois vetores não colineares que pertencem ao plano. A projeção P é portanto determinada por

$$P(1,1,0) = (1,1,0), \quad P(0,1,1) = (0,1,1), \quad P(1,3,-2) = (0,0,0)$$

Na base $B = ((1,1,0),(0,1,1),(1,3,-2))$ a representação matricial de P é dada por

$$A_{P,B,B} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Na base canónica, de acordo com o Exercício 24(a), a projeção é dada por

$$\begin{aligned}
A_{P,B_{can},B_{can}} &= \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 3 \\ 0 & 1 & -2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 3 \\ 0 & 1 & -2 \end{bmatrix}^{-1} \\
&= \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 3 \\ 0 & 1 & -2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \frac{1}{4} \begin{bmatrix} 5 & -1 & 1 \\ -2 & 2 & 2 \\ -1 & 1 & -1 \end{bmatrix} \\
&= \frac{1}{4} \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 3 \\ 0 & 1 & -2 \end{bmatrix} \begin{bmatrix} 5 & -1 & 1 \\ -2 & 2 & 2 \\ 0 & 0 & 0 \end{bmatrix} \\
&= \frac{1}{4} \begin{bmatrix} 5 & -1 & 1 \\ 3 & 1 & 3 \\ -2 & 2 & 2 \end{bmatrix}
\end{aligned}$$

logo $P(x,y,z) = \frac{1}{4}(5x - y + z, 3x + y + 3z, -2x + 2y + 2z)$.

26. EnunciadoConsideremos as transformações lineares $T\colon \mathbb{R}^n \to \mathbb{R}^m$ e $S\colon \mathbb{R}^k \to \mathbb{R}^n$ representadas nas bases canónicas relevantes respetivamente pelas matrizes A e B. Então AB representa a transformação $T \circ S$ e portanto $\operatorname{car}(AB) = \dim(T \circ S)(\mathbb{R}^k)$

Uma vez que $(T \circ S)(\mathbb{R}^k) \subset T(\mathbb{R}^n)$ temos $\dim(T \circ S)(\mathbb{R}^k) \leq \dim T(\mathbb{R}^n) = \operatorname{car}(A)$. Pelo Teorema da característica nulidade aplicada à restrição de T a $S(\mathbb{R}^k)$ temos $\dim(T \circ S)(\mathbb{R}^k) + \dim N(T_{|S(\mathbb{R}^k)}) = \dim S(\mathbb{R}^k)$ pelo que $\dim(T \circ S)(\mathbb{R}^k) \leq \dim S(\mathbb{R}^k) = \operatorname{car}(B)$.

Conclui-se que $\operatorname{car}(AB) \leq \min(\operatorname{car}(A), \operatorname{car}(B))$

Podemos obter uma fórmula para a característica do produto aplicando o Teorema da característica-nulidade à restrição de T ao subespaço $S(\mathbb{R}^k)$ como no parágrafo acima:

$$\begin{aligned} \operatorname{car}(AB) &= \dim(T\circ S)(\mathbb{R}^k) = \dim S(\mathbb{R}^k) - \dim N(T_{|S(\mathbb{R}^k)}) \\ &= \operatorname{car}(B) - \dim(N(T)\cap S(\mathbb{R}^k)) \\ &= \operatorname{car}(B) - \dim(N(A)\cap EC(B)) \end{aligned}$$

27. Enunciado A expressão matricial pretendida é essencialmente a expressão com respeito às bases construídas na demonstração do Teorema da Característica-Nulidade. Sejam $n = \dim V$ e $m = \dim W$, de forma que $k \leq n$ e $k \leq m$. Pelo Teorema da característica-nulidade $\dim N(f) = n - k$. Seja $\{v_{k+1}, \ldots, v_n\}$ uma base para $N(f)$ e $\{v_1, \ldots, v_k\}$ um conjunto de vetores que completa esta base de $N(f)$ numa base para V.

   De acordo com a demonstração do Teorema da Característica-Nulidade, os vetores $f(v_1), \ldots, f(v_k)$ formam uma base para a imagem de f. Tomamos $w_j = f(v_j)$ para $j \leq k$ e, para $j = k+1, \ldots, m$ escolhemos vetores w_j tais que $\{w_1, \ldots, w_m\}$ formem uma base para W.

   Uma vez que
$$f(v_i) = \begin{cases} w_i & \text{se } i \leq k \\ 0 & \text{caso contrário,} \end{cases}$$
   com respeito às bases $B_1 = (v_1, \ldots, v_n)$ e $B_2 = (w_1, \ldots, w_m)$ obtemos a representação matricial do enunciado.
28. Enunciado
   (a) Temos $T(1) = a + bi$ e $T(i) = (a+bi)i = -b + ai$ logo
$$A_{T,(1,i),(1,i)} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$
   (b) Temos
$$T(v_k) = \sum_{j=1}^m a_{jk}w_k = \sum_{j=1}^m (x_{jk} + iy_{jk})w_j = \sum_{j=1}^m x_{jk}w_j + y_{jk}(iw_j)$$
$$T(iv_k) = \sum_{j=1}^m ia_{jk}w_k = \sum_{j=1}^m i(x_{jk} + iy_{jk})w_j = \sum_{j=1}^m (-y_{jk} + ix_{jk})w_j = \sum_{j=1}^m -y_{jk}w_j + x_{jk}(iw_j)$$
   Isto signfica que as colunas $2k-1$ e $2k$ da matrix A (que são dadas pelas coordenadas de $T(v_k)$ e $T(iv_k)$ na base B_2') são dadas por
$$\begin{bmatrix} x_{1k} & -y_{1k} \\ y_{1k} & x_{1k} \\ \vdots & \vdots \\ x_{mk} & -y_{mk} \\ y_{mk} & x_{mk} \end{bmatrix}$$
   conforme pretendido.
29. Enunciado

(a) $G(f) \neq \emptyset$ uma vez que $(0,0) = (0, f(0)) \in G(f)$. Dados $(v_1, f(v_1)), (v_2, f(v_2)) \in G(f)$ temos

$$(v_1, f(v_1)) + (v_2, f(v_2)) = (v_1 + v_2, f(v_1) + f(v_2)) = (v_1 + v_2, f(v_1 + v_2)) \in G(f)$$

e, para um escalar α,

$$\alpha(v, f(v)) = (\alpha v, \alpha f(v)) = (\alpha v, f(\alpha v)) \in G(f)$$

Logo $G(f)$ é não vazio, fechado para a soma e produto por escalar, e portanto um subespaço vetorial de $V \times W$.

(b) ϕ é uma transformação linear uma vez que

$$\begin{aligned} \phi((v_1, w_1) + (v_2, w_2)) &= \phi(v_1 + v_2, w_1 + w_2) = (v_1 + v_2) + (w_1 + w_2) \\ &= (v_1 + w_1) + (v_2 + w_2) = \phi(v_1, w_1) + \phi(v_2, w_2) \end{aligned}$$

e

$$\phi(\alpha(v_1, w_1)) = \phi(\alpha v_1, \alpha w_1) = \alpha v_1 + \alpha w_1 = \alpha(v_1 + w_1) = \alpha\phi(v_1, w_1)$$

Uma vez que $\dim(W + V) = \dim(W) + \dim(V) - \dim(W \cap V)$ temos que $\dim(W + V) = \dim(U)$ e portanto o subespaço $W + V \subset U$ é na realidade igual a U. Isto mostra que a transformação linear ϕ é sobrejetiva. Se $\phi(v, w) = 0$ temos $v + w = 0 \Leftrightarrow v = -w$. Portanto $v \in V \cap W = \{0\}$, o mesmo sucedendo com w. Conclui-se que $N(\phi) = \{(0,0)\}$ logo ϕ é injetiva, e portanto um isomorfismo. Alternativamente, podemos concluir que ϕ é um isomorfismo pelo facto de ϕ ser sobrejetiva e os conjuntos de partida e chegada terem a mesma dimensão.

(c) Suponhamos primeiro que $\phi^{-1}(Z)$ é o gráfico de $f\colon V \to W$. Então

$$Z \cap W = \phi(\phi^{-1}(Z \cap W)) = \phi(\phi^{-1}(Z) \cap \phi^{-1}(W))$$

(uma vez que ϕ^{-1} é uma bijeção e portanto preserva intersecções). Uma vez que $\phi^{-1}(W) = \{0\} \times W$ e $G(f) \cap (\{0\} \times W) = \{0\} \times \{0\}$ vemos que

$$Z \cap W = \phi(\{0\} \times \{0\}) = \{0\}.$$

Reciprocamente, suponhamos que $Z \cap W = \{0\}$. Então $\phi^{-1}(Z) \cap (\{0\} \times W) = \phi^{-1}(Z) \cap \phi^{-1}(W) = \phi^{-1}(Z \cap W) = \{(0,0)\}$. Seja $p\colon \phi^{-1}(Z) \to V$ a restrição a $\phi^{-1}(Z)$ da projeção $\pi_1\colon V \times W \to V$ definida por $\pi(v, w) = v$. Então

$$N(p) = \{(v, w) \in \phi^{-1}(Z) \colon v = 0\} = \phi^{-1}(Z) \cap (\{0\} \times W) = \{(0,0)\}$$

Isto mostra que p é injetiva. Uma vez que $\dim \phi^{-1}(Z) = \dim Z = \dim V$ concluimos que p é um isomorfismo. Como p é sobrejetiva, para cada $v \in V$ existe $w \in W$ com $(v, w) \in \phi^{-1}(Z)$; como p é também injetiva este elemento é único. Isto significa que

$$f(v) = \text{ único } w \in W \text{ tal que } (v, w) \in \phi^{-1}(Z)$$

é uma função de V para W e que $\phi^{-1}(Z)$ é o gráfico de f. Resta ver que f é linear. Sendo $q\colon \phi^{-1}(Z) \to W$ a restrição a $\phi^{-1}(Z)$ da projeção $\pi_2\colon V \times W \to W$ (definida por $\pi_2(v, w) = w$), temos $f = q \circ p^{-1}$ logo f, sendo a composição de duas transformações lineares, é uma transformação linear.

30. Enunciado

(a) Temos $x+2y+3z=0 \Leftrightarrow z=-\frac{1}{3}(x+2y)$ logo

$$V=\{(x,y,-\tfrac{1}{3}(x+2y))\colon (x,y)\in\mathbb{R}^2\}=G(f)$$

quando $f\colon \mathbb{R}^2\to\mathbb{R}$ é definida por $f(x,y)=-\frac{1}{3}(x+2y)$.
Analogamente, $x+2y+3z=0\Leftrightarrow y=-\frac{1}{2}(x+3z)$ logo

$$V=\{(x,-\tfrac{1}{2}(x+3z),z)\colon (x,z)\in\mathbb{R}^2\}=G(g)$$

para $g\colon\mathbb{R}^2\to\mathbb{R}$ definida por $g(x,z)=-\frac{1}{2}(x+3z)$.

(b) Temos

$$\begin{cases} x-y+z+w=0 \\ 2x+z-w=0 \end{cases} \Leftrightarrow \begin{cases} z+w=y-x \\ z-w=-2x \end{cases} \Leftrightarrow \begin{cases} 2z=-3x+y \\ w=z+2x=\frac{1}{2}x+\frac{1}{2}y \end{cases}$$

logo V é o gráfico da função $(z,w)=f(x,y)=\frac{1}{2}(-3x+y,x+y)$.
Analogamente,

$$\begin{cases} x-y+z+w=0 \\ 2x+z-w=0 \end{cases} \Leftrightarrow \begin{cases} x-y=-z-w \\ x=\frac{1}{2}(-z+w) \end{cases} \Leftrightarrow \begin{cases} y=x+z+w=\frac{1}{2}(z+3w) \\ x=\frac{1}{2}(-z+w) \end{cases}$$

logo V é o gráfico da função $(x,y)=g(z,w)=\frac{1}{2}(-z+w,z+3w)$.

(c) Temos

$$\begin{cases} x+y+v=0 \\ y-3u+v=0 \\ z-u+v=0 \end{cases} \Leftrightarrow \begin{cases} x=-y-v=-3u+v-v=-3u \\ y=3u-v \\ z=u-v \end{cases}$$

Portanto

$$V=\{(-3u,3u-v,u-v,u,v)\colon u,v\in\mathbb{R}\}=G(f)$$

para $f\colon\mathbb{R}^2\to\mathbb{R}^3$ definida por $(x,y,z)=f(u,v)=(-3u,3u-v,u-v)$.

31. Enunciado

(a) Podemos simplesmente tomar $U=T(V)$, $p\colon V\to T(V)$ a aplicação definida por $p(v)=T(v)$ e $i\colon T(V)\to W$ a inclusão do subespaço $T(V)$ em W. Claramente $i(p(v))=p(v)=T(v)$ logo $T=i\circ p$ e, certamente p é sobrejetiva e i injetiva.
No entanto, para mostrar a unicidade da fatorização é útil pensar nela da seguinte forma. Uma vez que i é injetiva, temos necessariamente, em qualquer fatorização, $N(p)=N(T)$. Há uma aplicação canónica a partir de V que tem este núcleo, nomeadamente a projecção

$$\pi\colon V\to V/N(T)$$

definida por $\pi(v)=v+N(T)$ (ver o Exercício 3.12 da ficha sobre espaços vetoriais).
Verifiquemos que π é uma transformação linear: dados $v_1,v_2\in V$ e $\alpha\in\mathbb{R}$ temos

$$\pi(v_1+v_2)=v_1+v_2+N(T)=(v_1+N(T))+(v_2+N(T))=\pi(v_1)+\pi(v_2)$$

$$\pi(\alpha v)=(\alpha v)+N(T)=\alpha(v+N(T))=\alpha\pi(v)$$

Claramente π é sobrejetiva e $v \in N(\pi) \Leftrightarrow \pi(v) = 0 + N(T) \Leftrightarrow v + N(T) = 0 + N(T) \Leftrightarrow v \in N(T)$. Seja $i\colon V/N(T) \to W$ a aplicação definida por

$$i(v + N(T)) = T(v)$$

Vejamos que i está bem definida: se $v + N(T) = v' + N(T)$ então $v - v' \in N(T)$ e portanto $T(v) = T(v - v' + v') = T(v - v') + T(v') = 0 + T(v') = T(v')$. É imediato verificar que i é uma transformação linear e claramente

$$i(\pi(v)) = i(v + N(T)) = T(v)$$

pelo que $i \circ \pi = T$. Resta ver que i é injetiva: $i(v + N(T)) = T(v) = 0 \Leftrightarrow v \in N(T) \Leftrightarrow v + N(T) = 0 + N(T)$, logo $N(i) = \{0 + N(T)\}$ e portanto i é injetiva.

(b) Uma vez que a aplicação π é sobrejetiva, se $\overline{T}$ existir, ela é única e dada pela fórmula

$$\overline{T}(v + Z) = \overline{T}(\pi(v)) = T(v)$$

A fórmula acima define de facto uma função $\overline{T}\colon V/Z \to W$: se $v + Z = v' + Z$ então $v - v' \in Z$, portanto $\overline{T}(v' + Z) = T(v') = T(v') + 0 = T(v') + T(v - v') = T(v' + (v - v')) = T(v) = \overline{T}(v + Z)$. É imediato verificar que $\overline{T}$ é uma transformação linear e, por definição, $T = \overline{T} \circ \pi$.

(c) É suficiente demonstrar o resultado para $p = \pi\colon V \to V/N(T)$ e $i\colon V/N(T) \to W$ as aplicações canónicas definidas na segunda parte da resposta à alínea (a).
Uma vez que i e i' são injetivas e $i \circ p = i' \circ p'$, temos $N(p) = N(p') = N(T)$. Pela alínea (b) existe uma única transformação linear $\phi\colon U = V/N(T) \to U'$ tal que $\phi \circ p = p'$, que é definida por $\phi(v + N(T)) = p'(v)$. A aplicação ϕ é necessariamente sobrejetiva porque p' é sobrejetiva. Uma vez que $N(p') = N(T)$

$$\phi(v + N(T)) = 0 \Leftrightarrow p'(v) = 0 \Leftrightarrow v \in N(p') \Leftrightarrow v \in N(T) \Leftrightarrow v + N(T) = 0 + N(T)$$

logo ϕ é injetiva e portanto um isomorfismo.
Finalmente uma vez que p é sobrejetiva, para verificar que $i' \circ \phi = i$ basta ver que $i' \circ \phi \circ p = i \circ p$, o que é verdade porque $i' \circ \phi \circ p = i' \circ p' = i \circ p$.

32. Enunciado

(a) Seja $j \leq k$ e $q\colon \mathbb{R}^k \to \mathbb{R}^j$ a projeção nas primeiras j coordenadas. Então $\pi_j = q \circ \pi_k$ logo $\pi_j(EL(A)) = q(\pi_k(EL(A)))$ e portanto, pelo Teorema da característica nulidade aplicado a $q_{|\pi_k(EL(A))}$ temos

$$\begin{aligned} L(j) &= \dim \pi_j(EL(A)) = \dim q(\pi_k(EL(A))) \\ &= \dim \pi_k(EL(A)) - \dim N(q_{|\pi_k(EL(A))}) = L(k) - \dim N(q_{|\pi_k(EL(A))}) \end{aligned}$$

Conclui-se que $L(j) \leq L(k)$.

(b) O espaço das linhas é invariante mediante operações elementares pelo que basta verificar a afirmação para matrizes A em escada de linhas, onde ela é imediata: Seja A_i a matriz que se obtém da matriz em escada de linhas A considerando apenas as primeiras i colunas. Então $\dim \pi_i(EL(A)) = \operatorname{car}(A_i)$. A matriz em escada de linhas A tem um pivot na coluna i se e só se $\operatorname{car}(A_i) = \operatorname{car}(A_{i-1}) + 1 \Leftrightarrow L(i) = L(i-1) + 1$.

Ou seja, os pivots estão nas colunas em que a dimensão da projeção nos eixos aumenta e o j-ésimo pivot ocorre na primeira coluna i tal que $L(i) = j$.

(c) Se os pivots estão nas colunas $i_1, \dots, i_k$ então usando para base de $EL(A)$ as linhas não nulas da matriz em escada de linhas obtida no final do método de Gauss e para base de $\mathbb{R}^k$ a base canónica, a matriz que representa $\pi_{i_1,\dots,i_k} \colon EL(A) \to \mathbb{R}^k$ é triangular inferior com entradas não nulas (os pivots!) na diagonal (é a matriz quadrada que se obtém de A descartando todas as linhas nulas e todas as colunas que não contenham pivot e depois transpondo). Logo a restrição de $\pi_{i_1,\dots,i_k}$ é um isomorfismo.
Isto significa que $EL(A)$ é o gráfico de uma função das coordenadas $i_1, \dots, i_k$ para as restantes pelo Exercício 29(c), pois a intersecção de $EL(A)$ com o núcleo de $\pi_{i_1,\dots,i_k}$ (que é o plano formado pelas restantes coordenadas) é $\{0\}$.
Se $j_1, \dots, j_k$ é um qualquer conjunto de coordenadas tal que $\pi_{j_1,\dots,j_k} \colon EL(A) \to \mathbb{R}^k$ é um isomorfismo (note-se que têm que ser k coordenadas uma vez que $\dim EL(A) = k$), vamos ver que $j_1 \geq i_1, \dots, j_k \geq i_k$. Seja $m \in \{1, \dots, k\}$ e considere-se o diagrama comutativo:

$$\begin{tikzcd}
EL(A) \arrow[d, "\pi_{j_1,\dots,j_k}"'] \arrow[dr, "\pi_{j_m}"] \arrow[dd, bend right=60, "\pi_{j_1,\dots,j_m}"'] & \\
\mathbb{R}^k \arrow[d, "\pi_m"'] & \mathbb{R}^{j_m} \arrow[dl, "\pi_{j_1,\dots,j_m}"] \\
\mathbb{R}^m &
\end{tikzcd}$$

Uma vez que $\pi_{j_1,\dots,j_k}$ é um isomorfismo e π_m é sobrejetiva, a composta $\pi_{j_1,\dots,j_m}$ é sobrejetiva, logo a imagem de $\pi_{j_m} \colon EL(A) \to \mathbb{R}^{j_m}$ tem dimensão $\geq m$. Isto significa, na notação da alínea (a), que $L(j_m) \geq m$. Pela alínea (b), i_m é o mais pequeno índice no qual L atinge o valor m, portanto $j_m \geq i_m$.
Isto conclui a demonstração.

(d) Vamos descartar os casos triviais de planos $EL(A)$ com dimensão 0 ou iguais a todo o espaço.
   (i) Seja A uma matriz com 2 colunas e com característica 1. A tem o pivot na primeira coluna se e só se a projeção no eixo dos xx de $EL(A)$ é não trivial. Nesse caso $EL(A)$ é o gráfico de uma função $y = ax$ para algum $a \in \mathbb{R}$. Se A tem o pivot na segunda coluna então $EL(A)$ é a reta vertical.
   (ii) Seja A uma matriz com 3 colunas e com característica 1. A tem o pivot na primeira coluna se a reta $EL(A)$ é um gráfico sobre o eixo dos xx, isto é pode ser descrita por equações da forma $y = ax, z = bx$ com $a, b \in \mathbb{R}$. Tem o pivot na segunda coluna se está contida no plano $x = 0$ (isto é no plano yz) e é dada por $z = ay$ para algum $a \in \mathbb{R}$. Finalmente A tem o pivot na última coluna se $EL(A)$ coincide com o eixo dos zz.
   Suponhamos agora que A tem característica 2, então se os pivots estão nas primeira e segunda coluna, o plano $EL(A)$ é um gráfico sobre o plano xy, da

forma $z = ax + by$. Se os pivots estão na primeira e terceira coluna então a projecção sobre o plano xy é uma reta, ou seja, o plano contém o eixo dos zz mas não coincide com o plano yz (pois a projeção no eixo dos xx é sobrejetiva). Finalmente, se os pivots estão nas segunda e terceira colunas $EL(A)$ é o plano yz.

(iii) Seja A uma matrix com 4 colunas. Se $\text{car}(A) = 1$ e A tem um pivot na primeira coluna, é um gráfico sobre o eixo dos xx. Se o pivot está na segunda coluna $EL(A)$ está contido no plano yzw e é um gráfico sobre o eixo dos yy, etc.
Os restantes casos são análogos (lamento mas acabou-se me a paciência...). Indiquemos apenas um par de exemplos. Se os pivots estão nas colunas 2 e 4 isso significa que $EL(A)$ está contido no plano yzw e que se projeta numa reta no plano yz, ou seja, contém o eixo dos w.
Se os pivots estão nas colunas $1, 2$ e 4, o plano $EL(A)$ (de dimensão 3) projeta-se num plano de dimensão 2 da forma $z = ax + by$ no plano xyz e contém o eixo dos ww.

33. Enunciado

(a) A função f^* está bem definida porque a composição de duas transformações lineares é uma transformação linear. Dadas $T_1, T_2 \in L(W, U)$ e escalares α, β temos para todo o $v \in V$

$$\begin{aligned} f^*(\alpha T_1 + \beta T_2)(v) &= (\alpha T_1 + \beta T_2) \circ f(v) \\ &= (\alpha T_1 + \beta T_2)(f(v)) = \alpha T_1(f(v)) + \beta T_2(f(v)) \\ &= \alpha f^*(T_1)(v) + \beta f^*(T_2)(v) \end{aligned}$$

logo $f^*(\alpha T_1 + \beta T_2) = \alpha f^*(T_1) + \beta f^*(T_2)$. Conclui-se que f^* é uma transformação linear.

(b) Seja f um elemento de V^* então as transformações lineares f e $\sum_{i=1}^n f(v_i)\phi_i$ tomam os mesmos valores na base B e portanto coincidem. Isto mostra que B^* gera V^*. Vale a pena registar esta fórmula:

$$f = \sum_{i=1}^{n} f(v_i)\phi_i \tag{77}$$

Se $\sum_{i=1}^n \alpha_i\phi_i = 0$ então avaliando esta expressão em v_j obtemos $\alpha_j = 0$, logo B^* é linearmente independente e portanto uma base para V^*.
Uma vez que $\phi_i(\sum_{j=1}^n \alpha_j v_j) = \alpha_i$, ϕ_i é a i-ésima função coordenada com respeito à base B.

(c) Suponhamos que $\dim(V) = n$. Seja $W = N(\phi)$. Uma vez que ϕ não é o elemento identicamente nulo, $\phi\colon V \to \mathbb{R}$ é sobrejetivo e portanto, pelo Teorema da característica-nulidade $\dim(W) = n - 1$. Seja $\{v_2, \ldots, v_n\}$ uma base para W e $v_1 \in V$ um vetor tal que $\phi(v_1) = 1$ (que existe porque ϕ é não nulo e linear). Então $v_1 \notin W = L(\{v_2, \ldots, v_n\})$ logo $B = \{v_1, v_2, \ldots, v_n\}$ é linearmente independente e portanto uma base para V. Uma vez que $\phi(v_1) = 1$ e $\phi(v_i) = 0$ para $i \geq 2$, ϕ é o primeiro elemento da base dual de B e portanto é a primeira função coordenada com respeito à base B.

(d) Seja $B_1 = (v_1, \ldots, v_n)$, $B_2 = (w_1, \ldots, w_m)$, $B_1^* = (\phi_1, \ldots, \phi_n)$ e $B_2^* = (\psi_1, \ldots, \psi_m)$. Seja $A = A_{f,B_1,B_2}$ de forma que

$$f(v_j) = \sum_{i=1}^{m} a_{ij} w_i$$

(a coluna j de A contém as coordenadas de $f(v_j)$ na base B_2). Para calcular $f^*(\psi_i)$ podemos usar a expressão (77). Por definição de f^* temos

$$f^*(\psi_i)(v_j) = \psi_i(f(v_j)) = \psi_i(\sum_{k=1}^{m} a_{kj} w_k) = a_{ij}$$

logo, por (77),

$$f^*(\psi_i) = \sum_{j=1}^{n} (f^*(\psi_i)(v_j))\phi_j = \sum_{j=1}^{n} a_{ij}\phi_j$$

Isto mostra que $A_{f^*,B_2^*,B_1^*} = A^T$ conforme pretendido.

(e) Para ver que Φ é uma transformação linear, sejam $v_1, v_2 \in V$ e $\alpha, \beta \in \mathbb{R}$. Então para todo o $\psi \in V^*$ temos

$$\begin{aligned}(\Phi(\alpha v_1 + \beta v_2))(\psi) &= \psi(\alpha v_1 + \beta v_2) \\ &= \alpha\psi(v_1) + \beta\psi(v_2) \\ &= (\alpha\Phi(v_1))(\psi) + (\beta\Phi(v_2))(\psi) \\ &= (\alpha\Phi(v_1) + \beta\Phi(v_2))(\psi)\end{aligned}$$

e portanto $\Phi(\alpha v_1 + \beta v_2) = \alpha\Phi(v_1) + \beta\Phi(v_2)$ e Φ é linear.
Se $\Phi(v) = 0$ então $\psi(v) = 0$ para todo o $\psi \in V^*$, o que só acontece se $v = 0$ (de outra forma, v é o primeiro vetor de uma base[40] de B e tomando para ψ o primeiro elemento do conjunto B^* temos $\psi(v) = 1$). Conclui-se que $N(\Phi) = 0$ logo Φ é injetiva.
Se V é finitamente gerado, digamos com dimensão n então V^* tem também dimensão n (uma base para V determina um isomorfismo de V^* com $M_{1\times n}(\mathbb{R})$). Logo $(V^*)^*$ tem também dimensão n e sendo Φ uma transformação linear injetiva entre espaços da mesma dimensão, é um isomorfismo.

(f) Se V é o espaço dos polinómios reais, um elemento de $f \in V^*$ é dado por uma sucessão arbitrária de números reais $f(t^n)$ (na realidade V^* é isomorfo ao espaço das sucessões reais com as operações habituais de soma de sucessões e produto por escalar). O espaço V^* não tem uma base contável. De facto o subconjunto de sucessões $\{(e^{n\alpha}) \colon \alpha \in \mathbb{R}\}$ é um subconjunto linearmente independente não contável

[40] Agora que escrevo a solução noto que só provámos isto quando V é finitamente gerado pelo que só este caso devia ser considerado nesta alínea...

e está contido numa base[41] para V^* pelo que V^* tem uma base B não contável. Podemos agora definir um conjunto linearmente independente não contável B^* em $(V^*)^*$ pelo que $(V^*)^*$ também não tem uma base contável, e portanto não pode ser isomorfo a V.

34. Enunciado
    (a) Seja $B = (v_1, \ldots, v_k, v_{k+1}, \ldots, v_n)$ uma base para V tal que $(v_1, \ldots, v_k)$ é uma base para W. Seja $B^* = (\phi_1, \ldots, \phi_n)$ a base dual. Um elemento $\psi \in V^*$ pertence ao aniquilador de W se e só se $\psi(v_1) = \cdots = \psi(v_k) = 0$. Isto é equivalente a dizer que as primeiras k coordenadas de ψ na base dual são nulas (cf. (77)). Logo $W^o = L(\{\phi_{k+1}, \ldots, \phi_n\})$ tem dimensão $n-k$ e portanto $\dim W + \dim W^o = k + (n-k) = n$.
    (b) Temos

$$\begin{aligned} \phi \in N(f^*) &\Leftrightarrow f^*(\phi) = 0 \\ &\Leftrightarrow f^*(\phi)(u) = 0 \text{ para todo o } u \in U \\ &\Leftrightarrow \phi(f(u)) = 0 \text{ para todo o } u \in U \\ &\Leftrightarrow \phi \in f(U)^o \end{aligned}$$

    (c) Temos

$$\begin{aligned} \dim f^*(V^*) &= \dim V^* - \dim N(f^*) \quad \text{pelo Teorema da característica-nulidade aplicado a } f^* \\ &= \dim V - \dim f(U)^o \quad \text{pela alínea (b)} \\ &= \dim V - (\dim V - \dim f(U)) \quad \text{pela alínea (a)} \\ &= \dim f(U) \end{aligned}$$

    (d) Vejamos primeiro que $\psi(W) \subset (W^o)^o$: dado $w \in W$ temos que ver que $\psi(w)$ aniquila o subespaço W^o de V^*. Seja $\phi \in W^o$. Então

$$(\psi(w))(\phi) = \phi(w) = 0$$

    (pois $w \in W$ e $\phi \in W^o$) conforme pretendido.
    Note-se agora que, pela alínea (a), os espaços W e $(W^o)^o$ têm a mesma dimensão: $\dim W^o = \dim V - \dim W$ e portanto $\dim (W^o)^o = \dim V^* - \dim W^o = \dim V^* - \dim V + \dim W = \dim W$. Uma vez que a restrição

$$\psi_{|W} \colon W \to (W^o)^o$$

    é uma transformação injetiva (é a restrição de um isomorfismo) entre espaços da mesma dimensão, conclui-se que $\psi_{|W}$ é um isomorfismo, logo $\psi_{|W}$ é sobrejetivo, o que significa que $\psi(W) = (W^o)^o$.
    (e) Para verificar que $\psi_V \circ f = (f^*)^* \circ \psi_U$ temos a verificar que para cada $u \in U$ os elementos $\psi_V(f(u))$ e $(f^*)^*(\psi_U(u))$ de $(V^*)^*$ coincidem: seja ϕ um elemento de V^*, então

$$(\psi_V(f(u)))(\phi) = \phi(f(u))$$
$$((f^*)^*(\psi_U(u)))(\phi) = (\psi_U(u))(f^*(\phi)) = f^*(\phi)(u) = \phi(f(u))$$

[41]Novamente, não demonstrámos isto no caso de dimensão infinita.

conforme pretendido.

35. EnunciadoAplicando o método de Gauss obtemos

$$\begin{bmatrix} 1 & -3 & 1 & 4 & 0 \\ -2 & 6 & 0 & 3 & -1 \\ 1 & -3 & -1 & 0 & 4 \end{bmatrix} \to \begin{bmatrix} 1 & -3 & 1 & 4 & 0 \\ 0 & 0 & 2 & 11 & -1 \\ 0 & 0 & -2 & -4 & 4 \end{bmatrix} \to \begin{bmatrix} 1 & -3 & 1 & 4 & 0 \\ 0 & 0 & 2 & 11 & -1 \\ 0 & 0 & 0 & 7 & 3 \end{bmatrix}$$

Uma base para o espaço das linhas é $\{(1, -3, 1, 4, 0), (0, 0, 2, 11, -1), (0, 0, 0, 7, 3)\}$ (as três linhas da matriz inicial ou de qualquer das matrizes intermédias também são bases uma vez que são conjuntos linearmente independentes com 3 elementos distintos num espaço de dimensão 3).

Uma base para o espaço das colunas é dada pelas colunas da matriz inicial correspondentes às colunas com pivot na matriz final: $\{(1, -2, 1), (1, 0, -1), (4, 3, 0)\}$.

36. Enunciado
   (a) Se f é uma solução de $f' = \lambda f$ então

$$\frac{d}{dt}\left(e^{-\lambda t} f(t)\right) = -\lambda e^{-\lambda t} f(t) + e^{-\lambda t} f'(t) = -\lambda e^{-\lambda t} f(t) + e^{-\lambda t} \lambda f(t) = 0$$

   Logo existe uma constante $c \in \mathbb{R}$ tal que $e^{-\lambda t} f(t) = c \Leftrightarrow f(t) = ce^{-\lambda t}$. Por outro lado é imediato verificar que as funções desta forma satisfazem $f' = \lambda f$ logo

$$Tf = \lambda f \Leftrightarrow f \in L(\{e^{\lambda t}\})$$

   (b) Esta equação diferencial é um exemplo de uma equação linear uma vez que é da forma $S(x) = b$ com S uma transformação linear. Neste caso, a transformação linear S tem como domínio o subespaço de $F(\mathbb{R}, \mathbb{R})$ formado pelas funções diferenciáveis e toma valores em $F(\mathbb{R}, \mathbb{R})$, sendo definida pela expressão $S(f) = f' - f$ e o termo independente é a função $b(t) = t + t^2$.
   Tentamos então encontrar uma solução particular e é natural procurar uma que seja um polinómio. Claramente termos de grau > 2 não irão ajudar (pois contribuirão do lado esquerdo do sinal de igual com termos de grau > 2). Substituindo $f(t) = a + bt + ct^2$ na equação obtemos

$$b + 2ct - a - bt - ct^2 = t + t^2 \Leftrightarrow b - a + (2c - b)t - ct^2 = t + t^2$$

   e portanto $c = -1, b = -3, a = -3$. Logo $-3 - 3t - t^2$ é uma solução da equação. Pela alínea (a), a solução geral da equação homogénea $f' - f = 0$ é $f(t) = ce^t$ com $c \in \mathbb{R}$. Logo, pelo princípio da sobreposição, a solução geral é

$$f(t) = -3 - 3t - t^2 + ce^t, c \in \mathbb{R}$$

37. Enunciado 16(a): Uma solução particular da equação $f(v) = (1, 0, -2)$ é por exemplo $(1, 0)$. Vimos nesse exercício que o núcleo de f é $L(\{(-2, 1)\})$ logo pelo princípio da sobreposição o conjunto das soluções desta equação é $\{(1, 0) + \alpha(-2, 1) \colon \alpha \in \mathbb{R}\}$.

A equação $f(v) = (1, 1, 4)$ não tem solução uma vez que $(1, 1, 4)$ não pertence à imagem de f.

16(c): A equação $f(v) = (1, 0, -2)$ não tem solução uma vez que a imagem de f é gerada por $(1, 0, 2)$ e $(0, 1, 2)$ e o sistema

$$\alpha(1, 0, 2) + \beta(0, 1, 2) = (1, 0, -2)$$

não tem solução (das duas primeiras coordenadas obtemos $\alpha = 1, \beta = 0$ e a equação para a terceira coordenada fica impossível).

Uma solução particular de $f(v) = (1, 1, 4)$ é por exemplo $(1, 1, 0)$. Uma vez que o núcleo de f é $L(\{(1, -1, -1)\})$ conclui-se que o conjunto das soluções é $\{(1, 1, 0) + \alpha(1, -1, -1)\}$.

38. Enunciado Considerando a base $B_1 = (1+t, 2+t^2, t-t^3, 2+t)$ para V e a base canónica B_2 para as matrizes, a matriz que representa T com respeito a estas bases é

$$A_{T,B_1,B_2} = \begin{bmatrix} 1 & 0 & 1 & 2 \\ 1 & 2 & 1 & 0 \\ 0 & 2 & 2 & 0 \\ 1 & 1 & 0 & 0 \\ -1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

Podemos calcular o núcleo usando o método de Gauss:

$$\begin{bmatrix} 1 & 0 & 1 & 2 \\ 1 & 2 & 1 & 0 \\ 0 & 2 & 2 & 0 \\ 1 & 1 & 0 & 0 \\ -1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 1 & 2 \\ 0 & 2 & 0 & -2 \\ 0 & 2 & 2 & 0 \\ 0 & 1 & -1 & -2 \\ 0 & -1 & 1 & 2 \\ 0 & 0 & 1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 1 & 2 \\ 0 & 2 & 0 & -2 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & -1 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 1 & 2 \\ 0 & 2 & 0 & -2 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Portanto $(a, b, c, d) \in N(A_{T,B_1,B_2})$ se $c = -d$, $b = d$ e $a + c = -2d \Leftrightarrow a = -d$. Ou seja o núcleo da matriz é o conjunto dos vetores $\{(-d, d, -d, d) \colon d \in \mathbb{R}\}$. Uma vez que o núcleo desta matriz é a tradução nas coordenadas da base B_1 do núcleo de T vemos que

$$N(T) = \{-d(1+t) + d(2+t^2) - d(t-t^3) + d(2+t) \colon d \in \mathbb{R}\} = L(\{3 - t + t^2 + t^3\})$$

(a) O conjunto das soluções é o núcleo de T indicado acima.

(b) Esta equação é impossível pois as matrizes na imagem de T têm necessariamente entradas 22 e 31 simétricas, o que não acontece com o termo independente desta equação. Alternativamente podemos tentar resolver a equação em coordenadas:

$$\left[\begin{array}{cccc|c} 1 & 0 & 1 & 2 & 2 \\ 1 & 2 & 1 & 0 & -1 \\ 0 & 2 & 2 & 0 & 5 \\ 1 & 1 & 0 & 0 & 3 \\ -1 & -1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 1 & 0 \end{array}\right] \to \left[\begin{array}{cccc|c} 1 & 0 & 1 & 2 & 2 \\ 0 & 2 & 0 & -2 & -3 \\ 0 & 2 & 2 & 0 & 5 \\ 0 & 1 & -1 & -2 & 1 \\ 0 & -1 & 1 & 2 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{array}\right] \to \left[\begin{array}{cccc|c} 1 & 0 & 1 & 2 & 2 \\ 0 & 2 & 0 & -2 & -3 \\ 0 & 0 & 2 & 2 & 8 \\ 0 & 0 & -1 & -1 & \frac{5}{2} \\ 0 & 0 & 1 & 1 & -\frac{1}{2} \\ 0 & 0 & 1 & 1 & 0 \end{array}\right]$$

e dado que as últimas três equações do sistema são claramente incompatíveis, vemos novamente que a equação do enunciado não tem solução.

(c) Observamos que o termo independente da equação do enunciado é a diferença $T(1+t)-T(2+t^2)$ logo uma solução particular desta equação é dada por $1+t-(2+t^2) = -1+t-t^2$. Pelo princípio da sobreposição conclui-se que o conjunto das soluções é

$$\{-1+t-t^2+\alpha(3-t+t^2+t^3) \colon \alpha \in \mathbb{R}\}$$

39. Enunciado

(a) $T(e^x) = (e^x)'' - e^x = e^x - e^x = 0$ e $T(e^{-x}) = (e^{-x})'' - e^{-x} = e^{-x} - e^{-x} = 0$.

(b) Um polinómio de grau > 2 não pode satisfazer a equação (pois substituindo na equação teríamos do lado esquerdo do sinal de igual um polinómio de grau > 2). Tentamos encontrar uma solução da forma $f(x) = a + bx + cx^2$. Substituindo na equação obtemos:

$$2c - (a + bx + cx^2) = x^2 + x \Leftrightarrow -cx^2 - bx + (2c - a) = x^2 + x$$

pelo que $c = -1$, $b = -1$ e $a = 2c = -2$. Vemos assim que $f(x) = -2 - x - x^2$ é uma solução particular da equação.

(c) Tendo em conta as alíneas (a) e (b), pelo princípio da sobreposição o conjunto das soluções é

$$\{2 - x - x^2 + \alpha e^x + \beta e^{-x} \colon \alpha, \beta \in \mathbb{R}\}$$

(d) As soluções da primeira equação são $f(x) = \alpha e^x + \beta e^{-x}$ com $\alpha, \beta \in \mathbb{R}$. Substituindo na segunda equação obtemos

$$\begin{aligned}\left(\alpha e^x + \beta e^{-x}\right)' + 2e^{2x}\left(\alpha e^x + \beta e^{-x}\right) &= 4e^{3x}\\ \alpha e^x - \beta e^{-x} + 2\alpha e^{3x} + 2\beta e^x &= 4e^{3x}\\ (\alpha + 2\beta)e^x - \beta e^{-x} + (2\alpha - 4)e^{3x} &= 0\end{aligned}$$

Uma vez que as exponenciais são funções linearmente independentes (cf. Exercício 3.25) a equação anterior é equivalente ao sistema

$$\begin{cases}\alpha + 2\beta = 0\\ -\beta = 0\\ 2\alpha - 4 = 0\end{cases} \Leftrightarrow \begin{cases}\alpha = -2\beta\\ \beta = 0\\ \alpha = 2\end{cases}$$

Como este sistema é impossível conclui-se que o sistema de equações diferenciais dado é também impossível.

8.5. **O Determinante.**

1. Enunciado Seja P o paralelogramo com arestas $(1,3)$ e $(2,5)$. A função f é representada pela matriz $\begin{bmatrix}1 & 2\\ 3 & -1\end{bmatrix}$. O módulo do determinante desta matriz é $|-1-6| = 7$. Pela aditividade da área, a função f transforma um conjunto $A \subset \mathbb{R}^2$ com área a num conjunto $f(A)$ com área $7a$. Uma vez que a área de P é

$$|\begin{vmatrix}1 & 2\\ 3 & 5\end{vmatrix}| = 1$$

vemos que a área pretendida é 7.

Alternativamente, P é a imagem do quadrado unitário pela transformação linear representada na base canónica pela matriz $\begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}$. Logo a imagem de P por f é a imagem do quadrado unitário pela transformação linear representada na base canónica pela matriz

$$\begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix} = \begin{bmatrix} 7 & 12 \\ 0 & 1 \end{bmatrix}$$

ou seja, o paralelogramo com arestas $(7, 0)$ e $(12, 1)$ que tem área 7.

2. Enunciado As matrizes em questão são invertíveis se e só se o determinante for diferente de zero.
   (a) $(-1) \cdot 1 - 2 \cdot 3 = -7 \neq 0$.
   (b) Usando a regra de Sarrus

$$\begin{aligned} \begin{vmatrix} -1 & 2 & 1 \\ -3 & 2 & -1 \\ 1 & 4 & 3 \end{vmatrix} &= (-1) \cdot 2 \cdot 3 + 1 \cdot 2 \cdot (-1) + 1 \cdot (-3) \cdot 4 \\ &\quad -(1 \cdot 2 \cdot 1 + 2 \cdot (-3) \cdot 3 + (-1) \cdot (-1) \cdot 4) \\ &= -20 - (-12) = -8 \end{aligned}$$

   (c) Usando a regra de Laplace ao longo da primeira linha obtemos

$$(-1)\begin{vmatrix} 2 & -1 & 5 \\ 4 & 3 & 2 \\ 2 & 1 & 7 \end{vmatrix} - 2\begin{vmatrix} -3 & -1 & 5 \\ 1 & 3 & 2 \\ -1 & 1 & 7 \end{vmatrix} + 1\begin{vmatrix} -3 & 2 & 5 \\ 1 & 4 & 2 \\ -1 & 2 & 7 \end{vmatrix} - 7\begin{vmatrix} -3 & 2 & -1 \\ 1 & 4 & 3 \\ -1 & 2 & 1 \end{vmatrix} = -52 + 56 - 60 + 56 = 0$$

   (d) Usando a expansão de Laplace ao longo da última coluna obtemos

$$-2 \cdot \begin{vmatrix} -3 & 0 & -1 \\ 0 & 1 & 3 \\ -1 & 2 & 1 \end{vmatrix} + \begin{vmatrix} 0 & 2 & 1 \\ -3 & 0 & -1 \\ 0 & 1 & 3 \end{vmatrix} = -2 \cdot 14 + 15 = -13$$

   (e) Usando repetidos desenvolvimentos de Laplace ao longo da primeira coluna, ou notando que na fórmula para o determinante em termos de permutações o único termo não nulo é o da diagonal principal, vemos que o determinante de uma matriz triangular (superior ou inferior) é o produto das entradas na diagonal. Neste caso, esse produto é -42.
   (f) Cada coluna é uma combinação linear das restantes (sendo os coeficientes da combinação todos -1) logo a matriz tem característica no máximo $n-1$ e não é portanto invertível. Conclui-se que o seu determinante é igual a 0.
3. Enunciado Aplicando o método de Gauss obtemos

$$\begin{bmatrix} 0 & -1 & 1 & 1 & 0 \\ 1 & 1 & -1 & 2 & 1 \\ -1 & 0 & 3 & 2 & -3 \\ -2 & 1 & 0 & 1 & 1 \\ 1 & 0 & 2 & 0 & -1 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & -1 & 2 & 1 \\ 0 & -1 & 1 & 1 & 0 \\ 0 & 1 & 2 & 4 & -2 \\ 0 & 3 & -2 & 5 & 3 \\ 0 & -1 & 3 & -2 & -2 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & -1 & 2 & 1 \\ 0 & -1 & 1 & 1 & 0 \\ 0 & 0 & 3 & 5 & -2 \\ 0 & 0 & 1 & 8 & 3 \\ 0 & 0 & 2 & -3 & -2 \end{bmatrix} \to$$

$$\begin{bmatrix} 1 & 1 & -1 & 2 & 1 \\ 0 & -1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 8 & 3 \\ 0 & 0 & 0 & -19 & -11 \\ 0 & 0 & 0 & -19 & -8 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & -1 & 2 & 1 \\ 0 & -1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 8 & 3 \\ 0 & 0 & 0 & -19 & -11 \\ 0 & 0 & 0 & 0 & 3 \end{bmatrix}$$

Nos passos 1 e 4 trocámos um par de linhas pelo que o determinante da matriz inicial é $(-1)^2 \cdot (1 \cdot (-1) \cdot 1 \cdot (-19) \cdot 3) = 57$

4. EnunciadoTemos

$$\begin{aligned} \det(A^T B^{-2} A B^T A^{-2} (A^T)^3) &= \det(A^T) \det(B^{-1})^2 \det(A) \det(B^T) \det(A^{-1})^2 \det(A^T)^3 \\ &= \det(A) \frac{1}{(\det B)^2} \det(A) \det(B) \frac{1}{(\det A)^2} \det(A)^3 \\ &= \frac{\det(A)^3}{\det B} = \frac{(-3)^3}{2} = -\frac{27}{2} \end{aligned}$$

5. EnunciadoSe existisse uma tal matriz teríamos $(\det B)^2 = \det(A)$. Mas $\det(A) = -6 - 2 + 4 - 2 - 6 - 4 = -16 < 0$ não é o quadrado de um número real.
6. Enunciado Aplicando a regra de Laplace ao longo da segunda coluna da matriz 4×4 vemos que

$$2 \begin{vmatrix} a & b & c \\ g & h & i \\ j & k & l \end{vmatrix} = 6 \Leftrightarrow \begin{vmatrix} a & b & c \\ g & h & i \\ j & k & l \end{vmatrix} = 3 \tag{78}$$

Por outro lado, usando as propriedades do determinante temos

$$\begin{aligned} \begin{vmatrix} c & i & l \\ b-3c & h-3i & k-3l \\ 2a & 2g & 2j \end{vmatrix} &= \begin{vmatrix} c & i & l \\ b & h & k \\ 2a & 2g & 2j \end{vmatrix} - 3 \begin{vmatrix} c & i & l \\ c & i & l \\ 2a & 2g & 2j \end{vmatrix} \\ &= 2 \begin{vmatrix} c & i & l \\ b & h & k \\ a & g & j \end{vmatrix} - 0 \end{aligned}$$

A última matriz obtém-se da matriz em (78) transpondo e depois trocando a primeira linha com a terceira logo

$$\begin{vmatrix} c & i & l \\ b-3c & h-3i & k-3l \\ 2a & 2g & 2j \end{vmatrix} = 2 \cdot (-1) \cdot \begin{vmatrix} a & b & c \\ g & h & i \\ j & k & l \end{vmatrix} = -6$$

7. Enunciado O determinante é uma soma de termos que se obtêm multiplicando todas as entradas ao longo de um caminho que percorre todas as linhas e colunas sem repetir nenhuma com o sinal dado pelo determinante da matriz de permutação correspondente (é este o conteúdo da fórmula (30)).

(a) Para obter um termo em x^3 temos que escolher a segunda entrada da primeira linha e a terceira da terceira linha. Para que não haja repetições na segunda linha temos que escolher a primeira entrada. O termo correspondente no determinante é

$$x \cdot x \cdot (3x) \begin{vmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 3x^3 \cdot (-1) = -3x^3$$

O coeficiente pedido é portanto -3.

(b) Para obter um termo em x^3 temos que escolher uma entrada com x em todas as linhas excepto uma. Escolhendo um caminho que passe pela primeira entrada da primeira linha, o único caminho que produz mais dois termos com x é o que passa pelas entradas $(2,3),(3,2),(4,4)$. Se o caminho passa por qualquer outra entrada na primeira linha, terá que passar por três x's nas restantes linhas (sem repetições) e a única hipótese é então a anti-diagonal. Conclui-se que os termo envolvendo x^3 no determinante são

$$x \cdot x \cdot x \cdot 1 \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} + (-2) \cdot x \cdot x \cdot (-x) \begin{vmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{vmatrix} = x^3 \cdot (-1) + 2x^3 \cdot (-1)^2 = x^3$$

onde os determinantes foram calculados vendo quantas linhas temos de trocar para obter a matriz identidade a partir da matriz de 0s e 1s. O sinal $-$ do primeiro termo corresponde à troca da segunda com a terceira linha enquanto que o sinal $(-1)^2$ no segundo corresponde à troca da primeira com a quarta e segunda com terceira. Conclui-se que o coeficiente de x^3 no determinante é 1.

8. Enunciado Subtraindo $x_1 C_{n-1}$ a C_n, $x_2 C_{n-2}$ a C_{n-1}, etc. sucessivamente não alteramos o determinante e obtemos

$$\begin{vmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & x_2 - x_1 & x_2(x_2 - x_1) & \cdots & x_2^{n-2}(x_2 - x_1) \\ \vdots & \vdots & & \vdots & \\ 1 & x_{n-1} - x_1 & x_{n-1}(x_{n-1} - x_1) & \cdots & x_{n-1}^{n-2}(x_{n-1} - x_1) \\ 1 & x_n - x_1 & x_n(x_n - x_1) & \cdots & x_n^{n-2}(x_n - x_1) \end{vmatrix}$$

Fazendo o desenvolvimento de Laplace ao longo da primeira linha obtemos

$$\begin{vmatrix} x_2 - x_1 & x_2(x_2 - x_1) & \cdots & x_2^{n-2}(x_2 - x_1) \\ \vdots & & \vdots & \\ x_{n-1} - x_1 & x_{n-1}(x_{n-1} - x_1) & \cdots & x_{n-1}^{n-2}(x_{n-1} - x_1) \\ x_n - x_1 & x_n(x_n - x_1) & \cdots & x_n^{n-2}(x_n - x_1) \end{vmatrix}$$

e pondo em evidência sucessivamente os fatores comuns $(x_j - x_1)$ em cada linha

$$(x_2-x_1)\cdots(x_n-x_1)\begin{vmatrix} 1 & x_2 & \cdots & x_2^{n-2} \\ \vdots & & \vdots & \\ 1 & x_{n-1} & \cdots & x_{n-1}^{n-2} \\ 1 & x_n & \cdots & x_n^{n-2} \end{vmatrix}$$

Por hipótese de indução o determinante à direita é o produto de todos os termos $(x_j - x_i)$ com $2 \leq i < j \leq n$, o que conclui a demonstração.

9. Enunciado A entrada 23 é dada pelo quociente do cofator 32 da matriz pelo determinante, ou seja, por

$$\frac{(-1)^{3+2}\begin{vmatrix} 0 & 1 & -2 \\ -1 & 2 & 0 \\ 1 & 1 & 1 \end{vmatrix}}{\begin{vmatrix} 0 & 1 & 1 & -2 \\ -1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & -1 & 1 & 1 \end{vmatrix}} = \frac{-(2+4+1)}{-(-1)\cdot\begin{vmatrix} 1 & 1 & -2 \\ 1 & 0 & 1 \\ -1 & 1 & 1 \end{vmatrix} - \begin{vmatrix} 1 & 1 & -2 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{vmatrix}} = \frac{-7}{-5-6} = \frac{7}{11}$$

10. Enunciado
    (a) A entrada ij da matriz AB é dada pela expressão $\sum_{k=1}^{n} a_{ik}b_{kj}$ logo

$$\operatorname{tr}(AB) = \sum_{i=1}^{n}\sum_{k=1}^{n} a_{ik}b_{ki} = \sum_{k=1}^{n}\sum_{i=1}^{n} b_{ki}a_{ik} = \operatorname{tr}(BA)$$

    onde, na segunda igualdade trocámos a ordem da soma, o que é válido pelas propriedades comutativa e associativa da soma.
    (b) Temos $\operatorname{tr}(SAS^{-1}) = \operatorname{tr}((SA)S^{-1}) = \operatorname{tr}(S^{-1}(SA)) = \operatorname{tr}(S^{-1}SA) = \operatorname{tr}(A)$ (onde na segunda igualdade usámos o resultado da alínea (a)).

11. Enunciado
    (a) A derivada da entrada ij da matriz produto é dada por

$$\frac{d}{dt}\left(\sum_{k=1}^{n} a_{ik}(t)b_{kj}(t)\right) = \sum_{k=1}^{n} a'_{ik}(t)b_{kj}(t) + a_{ik}(t)b'_{kj}(t) = \sum_{k=1}^{n} a'_{ik}(t)b_{kj}(t) + \sum_{k=1}^{n} a_{ik}(t)b'_{kj}(t)$$

    que é precisamente a soma das entradas ij das matrizes $A'(t)B(t)$ e $A(t)B'(t)$. Logo $\frac{d}{dt}(A(t)B(t)) = A'(t)B(t) + A(t)B'(t)$.
    (b) Pela definição de derivada de uma função real de variável real, cada entrada ij da matriz $A(t)$ pode escrever-se na forma

$$a_{ij}(t) = a_{ij}(0) + a'_{ij}(0)t + A_{ij}(t)$$

    com $\lim_{t\to 0}\frac{A_{ij}(t)}{t} = 0$. Uma vez que $A(0) = I_n$ temos

$$a_{ij}(0) = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases}$$

Assim, todas as entradas fora da diagonal se anulam em 0 pelo menos tão rápido como t. Quando escrevemos o determinante como uma soma de termos indexados pelas permutações de n, todos os termos excepto o diagonal têm como fatores pelo menos duas entradas fora da diagonal e portanto anulam-se pelo menos tão rápido como t^2. Como tal esses termos não contribuem para a derivada em 0 e portanto

$$\begin{aligned}
\frac{d}{dt}\left(\det A(t)\right)_{|t=0} &= \frac{d}{dt}\left(a_{11}(t)a_{22}(t)\cdots a_{nn}(t)\right)_{|t=0} \\
&= \frac{d}{dt}\left((1+a'_{11}(0)t+A_{11}(t))\cdots(1+a'_{nn}(0)t+A_{nn}(t))\right)_{|t=0} \\
&= \lim_{t\to 0}\frac{(1+a'_{11}(0)t+A_{11}(t))\cdots(1+a'_{nn}(0)t+A_{nn}(t))-1}{t} \\
&= \lim_{t\to 0}\frac{1+a'_{11}(0)t+\ldots+a'_{nn}(0)t+B(t)-1}{t}
\end{aligned}$$

onde $B(t) = A_{11}(t) + \ldots + A_{nn}(t) + a'_{11}(0)ta'_{22}(0)t + \ldots$ é uma função tal que $\lim_{t\to 0}\frac{B(t)}{t} = 0$. Conclui-se que

$$\frac{d}{dt}\left(\det A(t)\right)_{|t=0} = a'_{11}(0) + \ldots + a'_{nn}(0) = \operatorname{tr}(A'(0))$$

12. Enunciado Uma vez que o determinante é linear em cada linha e troca de sinal com uma troca de linhas, aplicando estas propriedades apenas às linhas que envolvem a matriz C vemos que a função

$$f(C) = \begin{vmatrix} A & B \\ 0 & C \end{vmatrix}$$

é multilinear e alternante. Vimos na demonstração da unicidade da função determinante que uma tal função é necessariamente da forma

$$f(C) = \det(C)f(I)$$

pelo que resta mostrar que $f(I) = \det(A)$. Mas

$$f(I) = \begin{vmatrix} A & B \\ 0 & I \end{vmatrix}$$

e este determinante é de facto $\det(A)$ como se pode ver aplicando sucessivamente a fórmula de Laplace à última linha, ou usando a fórmula para o determinante como uma soma de termos indexados por permutações (as únicas permutações que contribuem termos não nulas são a identidade no conjunto de índices correspondente às linhas de C).

Note-se que, indutivamente vemos que o determinante de uma matriz triangular por blocos (com qualquer número de blocos na diagonal) é o produto dos determinantes dos blocos na diagonal.

(i)

$$\begin{vmatrix} 2 & 1 & -5 & 4 & 10 \\ 0 & 3 & 6 & -1 & -3 \\ 0 & 7 & 8 & -2 & 3 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & -3 & 1 \end{vmatrix} = |2| \begin{vmatrix} 3 & 6 \\ 7 & 8 \end{vmatrix} \begin{vmatrix} 1 & 3 \\ -3 & 1 \end{vmatrix} = 2 \cdot (-18) \cdot 10 = -360$$

(ii)

$$\begin{vmatrix} 2 & 1 & -5 & 4 & 10 & 7 \\ 1 & 2 & -1 & -1 & -3 & 8 \\ 1 & 0 & 0 & -2 & 3 & -3 \\ 0 & 0 & 0 & -2 & 1 & 3 \\ 0 & 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{vmatrix} = \begin{vmatrix} 2 & 1 & -5 \\ 1 & 2 & -1 \\ 1 & 0 & 0 \end{vmatrix} | - 2| \begin{vmatrix} 2 & 1 \\ 1 & 4 \end{vmatrix} = 9 \cdot (-2) \cdot 7 = -126$$

13. Enunciado Temos

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} I & X \\ 0 & I \end{bmatrix} = \begin{bmatrix} A & AX + B \\ C & CX + D \end{bmatrix}$$

Uma vez que A é invertível podemos resolver a equação $AX + B = 0$ obtendo $X = -A^{-1}B$. Para esta escolha de X temos

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = \begin{vmatrix} A & B \\ C & D \end{vmatrix} \begin{vmatrix} I & -A^{-1}B \\ 0 & I \end{vmatrix} = \begin{vmatrix} A & 0 \\ C & -CA^{-1}B + D \end{vmatrix} = \det(A)\det(-CA^{-1}B + D)$$

onde na primeira e na última igualdade usámos o resultado do exercício anterior, e na segunda igualdade usámos a regra para o determinante de um produto.

14. Enunciado
   (a) Escrevendo $v = \alpha_1 v_1 + \ldots + \ldots + \alpha_n v_n$ e $w = \beta_1 v_1 + \ldots + \beta_n v_n$ com $\alpha_i, \beta_i \in \mathbb{K}$ e usando sucessivamente a linearidade na primeira e depois na segunda variável obtemos

$$\begin{aligned} h(v,w) &= h\left(\sum_{i=1}^n \alpha_i v_i, \sum_{j=1}^n \beta_j v_j\right) \\ &= \sum_{i=1}^n \alpha_i h\left(v_i, \sum_{j=1}^n \beta_j v_j\right) \\ &= \sum_{i=1}^n \sum_{j=1}^n \alpha_i \beta_j h(v_i, v_j) \\ &= \begin{bmatrix} \alpha_1 & \cdots & \alpha_n \end{bmatrix} \begin{bmatrix} h(v_1,v_1) & \cdots & h(v_1,v_n) \\ \vdots & \ddots & \vdots \\ h(v_n,v_1) & \cdots & h(v_n,v_n) \end{bmatrix} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix} \end{aligned}$$

pelo que podemos tomar para A a matrix $[h(v_i, v_j)]$. Por outro lado, se $A = [a_{ij}]$ for uma matriz satisfazendo a condição do enunciado temos

$$h(v_i, v_j) = \begin{bmatrix} 0 & \cdots & 1 & \cdots & 0 \end{bmatrix} A \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = a_{ij}$$

(onde o 1 está na posição i do vetor da esquerda e na posição j do vetor da direita) pelo que a matriz A é única.

(b) A linearidade na segunda variável garante que para todo o $v \in V$ a função $v \mapsto h(v, \cdot)$ é um elemento de V^*. Para verificar que $\overline{h}$ é uma transformação linear temos a mostrar que dados $\alpha, \beta \in \mathbb{K}$ e $v_1, v_2 \in V$ se verifica

$$\overline{h}(\alpha v_1 + \beta v_2) = \alpha \overline{h}(v_1) + \beta \overline{h}(v_2)$$

o que podemos confirmar avaliando as duas funções num elemento $w \in V$:

$$\begin{aligned} \overline{h}(\alpha v_1 + \beta v_2)(w) &= h(\alpha v_1 + \beta v_2, w) \\ &= \alpha h(v_1, w) + \beta h(v_2, w) \\ &= \alpha \overline{h}(v_1)(w) + \beta \overline{h}(v_2)(w) \\ &= (\alpha \overline{h}(v_1) + \beta \overline{h}(v_2))(w) \end{aligned}$$

A entrada ij da representação matricial de $\overline{h}$ com respeito às bases B e B^* é a coordenada i de $\overline{h}(v_j)$ na base B^*, que se calcula avaliando o funcional $\overline{h}(v_j)$ em v_i:

$$\overline{h}(v_j)(v_i) = h(v_j, v_i)$$

Conclui-se que a representação matricial de $\overline{h}$ é a transposta da matriz da alínea anterior.

(c) h é não degenerada se e só se $\overline{h}(v) = 0 \Rightarrow v = 0$. Uma vez que V e V^* têm a mesma dimensão esta condição (de o núcleo de $\overline{h}$ ser 0) é equivalente à condição de $\overline{h}$ ser um isomorfismo.

(d) Quando expandimos a expressão matricial $[v]_B^T A [w]_B$ a i-ésima coordenada de v e a j-ésima coordenada de w ficam a multiplicar por a_{ij} logo

$$xu + 3xw - yv + 2yw + 5yu - 3zu + zw = \begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} 1 & 0 & 3 \\ 5 & -1 & 2 \\ -3 & 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix}$$

Uma vez que a matriz é invertível (tem característica 3), a transformação $\overline{h}$ que é representada pela matriz transposta é um isomorfismo e portanto, pela alínea anterior, esta forma bilinear é não degenerada.

15. Enunciado

$$y = \frac{\begin{vmatrix} 1 & 2 & -1 \\ 1 & 1 & 3 \\ 1 & 1 & -2 \end{vmatrix}}{\begin{vmatrix} 1 & 2 & -1 \\ 1 & -1 & 3 \\ 1 & 0 & -2 \end{vmatrix}} = \frac{5}{11}$$

16. Enunciado
(i) $\begin{vmatrix} e_1 & e_2 & e_3 \\ 1 & 0 & 2 \\ 2 & 0 & 1 \end{vmatrix} = (0, 3, 0)$
(ii) $\begin{vmatrix} e_1 & e_2 & e_3 \\ 1 & -1 & 1 \\ 1 & 2 & 4 \end{vmatrix} = (-6, -3, 3)$
(iii) $(2v_1 + 3v_2) \times (-v_1 + v_2) = 2v_1 \times (-v_1 + v_2) + 3v_2 \times (-v_1 + v_2) = -2v_1 \times v_1 + 2v_1 \times v_2 - 3v_2 \times v_1 + 3v_2 \times v_2 = 0 + 2v_1 \times v_2 + 3v_1 \times v_2 + 0 = 5v_1 \times v_2$

17. Enunciado Uma vez que o plano é paralelo aos vetores $(1, 1, 2)$ e $(-1, 0, 1)$, um vetor perpendicular ao plano é dado por

$$(1, 1, 2) \times (-1, 0, 1) = \begin{vmatrix} e_1 & e_2 & e_3 \\ 1 & 1 & 2 \\ -1 & 0 & 1 \end{vmatrix} = (1, -3, 1)$$

Logo o plano tem equação $x - 3y + z = C$. Para que o ponto $(0, 1, 3)$ pertença ao plano temos $0 - 3 \cdot 1 + 3 = C$ logo a equação do plano é $x - 3y + z = 0$.

18. Enunciado Uma vez que $u \times e_1 = (0, c, -b)$, $u \times e_2 = (-c, 0, a)$ e $u \times e_3 = (b, -a, 0)$ a matriz que representa T é

$$\begin{bmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{bmatrix}$$

19. Enunciado Ambas as expressões são tri-lineares e anti-simétricas em v, w pelo que basta verificar que produzem o mesmo resultado quando u percorre a base canónica (e_1, e_2, e_3) e (v, w) tomam os valores $(e_1, e_2), (e_1, e_3)$ e (e_2, e_3):

u	v	w	$u \times (v \times w)$	$\langle u, w\rangle v - \langle u, v\rangle w$
e_1	e_1	e_2	$-e_2$	$-e_2$
e_1	e_1	e_3	$-e_3$	$-e_3$
e_1	e_2	e_3	0	0
e_2	e_1	e_2	e_1	e_1
e_2	e_1	e_3	0	0
e_2	e_2	e_3	$-e_3$	$-e_3$
e_3	e_1	e_2	0	0
e_3	e_1	e_3	e_1	e_1
e_3	e_2	e_3	e_2	e_2

20. Enunciado
   (a) Por exemplo $\mathbf{i}^2 = -\langle \mathbf{i}, \mathbf{i}\rangle + \mathbf{i}\times\mathbf{i} = -1+0$; $\mathbf{ij} = -\langle \mathbf{i}, \mathbf{j}\rangle + \mathbf{i}\times\mathbf{j} = -0+\mathbf{k}$, e analogamente para os restantes produtos.
   Reciprocamnente aplicando as relações postuladas obtemos

$$\begin{aligned}(a + x\mathbf{i} + y\mathbf{j} + z\mathbf{k})(A + X\mathbf{i} + Y\mathbf{j} + Z\mathbf{k}) &= aA - xX - yY - zZ + (aX + xA + yZ \\ &\quad -zY)\mathbf{i} + (aY + yA - xZ + zX)\mathbf{j} + \\ &\quad (aZ + zA + xY - yX)\mathbf{k}\end{aligned}$$

   pelo que a parte real é de facto $aA - \langle (x, y, z), (X, Y, Z)\rangle$ enquanto que a diferença entre a parte vetorial e $a(X, Y, Z) + A(x, y, z)$ é dada por

$$\begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x & y & z \\ X & Y & Z \end{vmatrix} = (yZ - zY)\mathbf{i} + (zX - xZ)\mathbf{j} + (xY - yX)\mathbf{i}$$

   (b) Temos $(a + b\mathbf{i})(c + d\mathbf{i}) = ac - bd + (ad + bc)\mathbf{i}$ pelo que a multiplicação coincide com a multiplicação nos complexos.
   (c) $(t+v)\cdot(1+0) = t-0+t0-1v+0\times v = t+v$ e analogamente $(1+0)(t+v) = (t+v)$. A multiplicação não é comutativa porque, por exemplo $\mathbf{ij} = \mathbf{i}\times\mathbf{j} \neq \mathbf{j}\times\mathbf{i} = \mathbf{ji}$. Para verificar a associatividade, pela multilinearidade da multiplicação, é suficiente substituir na igualdade $x(yz) = (xy)z$ os elementos da base $\mathbf{i}, \mathbf{j}$ e $\mathbf{k}$ pois a verificação é imediata se x, y ou z forem iguais à identidade 1. Dois casos representativos (entre os 27) são:

$$\mathbf{i}(\mathbf{jk}) = \mathbf{ii} = -1 = \mathbf{kk} = (\mathbf{ij})\mathbf{k}, \qquad \mathbf{j}(\mathbf{jk}) = \mathbf{ji} = -\mathbf{k} = (-1)\mathbf{k} = (\mathbf{jj})\mathbf{k}$$

   (d) Temos

$$(t + v)\frac{t - v}{t^2 + \|v\|^2} = \frac{(t^2 - \langle v, -v\rangle) + t(-v) + tv - v \times v}{t^2 + \|v\|^2} = \frac{t^2 + \|v\|^2}{t^2 + \|v^2\|} = 1$$

   e analogamente para a multiplicação na ordem inversa.

8.6. **Endomorfismos.**

1. Enunciado
   (a) O polinómio característico é $(1 - \lambda)^2 - 9$ e tem raízes $1 - \lambda = \pm 3 \Leftrightarrow \lambda = -2$ ou $\lambda = 4$. Os vetores próprios de $\lambda = -2$ são as soluções não nulas de

$$\begin{bmatrix} 3 & 3 \\ 3 & 3 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = 0 \Leftrightarrow b = -a$$

   ou seja $a(1, -1)$ com $a \neq 0$. Os vetores próprios de $\lambda = 4$ são as soluções não nulas de

$$\begin{bmatrix} -3 & 3 \\ 3 & -3 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = 0 \Leftrightarrow a = b$$

   ou seja $a(1, 1)$ com $a \neq 0$. Há uma base para $\mathbb{R}^2$ constituída por vetores próprios logo A é diagonalizável.

(b) O polinómio característico é $(2-\lambda)(1-\lambda)$ logo os valores próprios são $\lambda = 2$ e $\lambda = 1$. Claramente os vetores próprios de $\lambda = 2$ são os vetores da forma $a(1,0)$ com $a \neq 0$. Os vetores próprios de $\lambda = 1$ são as soluções não nulas de

$$\begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = 0 \Leftrightarrow a = b$$

ou seja são os vetores $a(1,1)$ com $a \neq 0$. Uma vez que há uma base para $\mathbb{R}^2$ formada por vetores próprios, A é diagonalizável.

(c) O único valor próprio é 1 (pois ambas as entradas na diagonal são iguais a 1) e claramente os vetores próprios de 1 são os vetores $a(1,0)$ com $a \neq 0$. Uma vez que não há uma base para $\mathbb{R}^2$ formada por vetores próprios, a matriz não é diagonalizável. Note-se que uma matriz diagonalizável com um único valor próprio é necessariamente um múltiplo da identidade (é multiplicação pelo mesmo escalar em todas as direções).

(d) O polinómio característico é $(2-\lambda)^2+3$ que tem apenas as raízes complexas $2\pm\sqrt{3}i$. Os vetores próprios de $2+\sqrt{3}i$ (quando a matriz é vista como uma matriz complexa) são as soluções não nulas de

$$\begin{bmatrix} -\sqrt{3}i & -3 \\ 1 & -\sqrt{3}i \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = 0 \Leftrightarrow a = \sqrt{3}ib$$

que são $b(\sqrt{3}i, 1)$ com $b \in \mathbb{C} \setminus \{0\}$.
Os vetores próprios de $2-\sqrt{3}i$ são os vetores conjugados dos vetores próprios de $2+\sqrt{3}i$, ou seja, $a(-\sqrt{3}i, 1)$ com $a \in \mathbb{C} \setminus \{0\}$. Uma vez que há uma base para $\mathbb{C}^2$ formada por vetores próprios, esta matriz complexa é diagonalizável.

(e) Os valores próprios são as entradas diagonais $-1, 2$ e 3. Os vetores próprios de -1 são os vetores da forma $a(1,0,0,)$ com $a \neq 0$. Os vetores próprios de 2 são as soluções não nulas de

$$\begin{bmatrix} -3 & 2 & 1 \\ 0 & 0 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 0 \Leftrightarrow \begin{cases} b = \frac{3}{2}a \\ c = 0 \end{cases}$$

ou seja os vetores $a(1, \frac{3}{2}, 0)$ com $a \neq 0$. Os vetores próprios de 3 são as soluções não nulas de

$$\begin{bmatrix} -4 & 2 & 1 \\ 0 & -1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 0 \Leftrightarrow \begin{cases} c = -2b + 4a \\ b = -c \end{cases} \quad \Leftrightarrow \begin{cases} c = -4a \\ b = 4a \end{cases}$$

ou seja os vetores $a(1,4,-4)$ com $a \neq 0$. Uma vez que há uma base para $\mathbb{R}^3$ formada por vetores próprios a matriz é diagonalizável.

(f) O polinómio característico é $(2-\lambda)((1-\lambda)^2-1) = (2-\lambda)(\lambda^2-2\lambda) = -\lambda(\lambda-2)^2$. Os valores próprios são portanto $0, 2$. Os vetores próprios de 2 são os elementos não nulos do núcleo de

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{bmatrix}$$

que são os vetores da forma $(a, 0, 0)$ com $\neq 0$. Os vetores próprios de 0 são os elementos não nulos do núcleo da matriz que são as soluções de

$$\begin{cases} 2a + c = 0 \\ b + c = 0 \end{cases} \Leftrightarrow \begin{cases} c = -2a \\ b = -c = 2a \end{cases}$$

Logo os vetores próprios são os vetores da forma $a(1, 2, -2)$ com $a \neq 0$. Uma vez que não há uma base para $\mathbb{R}^3$ formada por vetores próprios conclui-se que a matriz não é diagonalizável.

(g) A matriz é triangular logo os valores próprios são $(1 + i)$ e $-i$. Os vetores próprios de $-i$ são os vetores da forma $(0, a)$ com $a \in \mathbb{C} \setminus \{0\}$. Os vetores próprios de $1 + i$ são os elementos não nulos do núcleo de

$$\begin{bmatrix} 0 & 0 \\ 1 & -1 - 2i \end{bmatrix}$$

que são os vetores da forma $b(1 + 2i, 1)$ com $b \in \mathbb{C} \setminus \{0\}$. Uma vez que existe uma base para $\mathbb{C}^2$ formada por vetores próprios a matriz é diagonalizável.

(h) Consideramos a base $B = (\left[\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right], \left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right], \left[\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right], \left[\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right])$ para $M_{2\times 2}(\mathbb{R})$. Temos

$T\left(\left[\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right]\right) = \left[\begin{smallmatrix} -1 & 0 \\ 0 & 0 \end{smallmatrix}\right]$, $\quad T\left(\left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right]\right) = \left[\begin{smallmatrix} 0 & -1 \\ 0 & 0 \end{smallmatrix}\right]$, $\quad T\left(\left[\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right]\right) = \left[\begin{smallmatrix} 2 & 0 \\ 1 & 0 \end{smallmatrix}\right]$, $\quad T\left(\left[\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right]\right) = \left[\begin{smallmatrix} 0 & 2 \\ 0 & 1 \end{smallmatrix}\right]$

logo

$$A_{T,B,B} = \begin{bmatrix} -1 & 0 & 2 & 0 \\ 0 & -1 & 0 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Os valores próprios são portanto 1 e -1. Os vetores próprios de 1 são os vetores não nulos cujas coordenadas estão no núcleo de

$$\begin{bmatrix} -2 & 0 & 2 & 0 \\ 0 & -2 & 0 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

ou seja, com coordenadas (c, d, c, d) com c, d não ambos nulos. Os vetores próprios são portanto as matrizes $\left[\begin{smallmatrix} c & d \\ c & d \end{smallmatrix}\right]$ com c, d não ambos nulos.

Os vetores próprios de -1 são os vetores não nulos cujas coordenadas estão no núcleo de

$$\begin{bmatrix} 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

ou seja, com coordenadas $(a, b, 0, 0)$ com a, b não ambos nulos. Os vetores próprios são portanto as matrizes $\left[\begin{smallmatrix} a & b \\ 0 & 0 \end{smallmatrix}\right]$ com a, b não ambos nulos. Uma vez que há uma base para $M_{2\times 2}(\mathbb{R})$ formada por vetores próprios a transformação T é diagonalizável.

(i) Consideramos a base $B = (1, x, x^2, x^3)$ para V. Temos

$$T(1) = \frac{3x^2}{10}, \quad T(x) = 1 + \frac{3x^2}{20}, \quad T(x^2) = 2x + \frac{x^2}{10}, \quad T(x^3) = 3x^2 + \frac{3x^2}{40}$$

logo

$$A_{T,B,B} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ \frac{3}{10} & \frac{3}{20} & \frac{1}{10} & \frac{123}{40} \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

O polinómio característico desta matriz é

$$\begin{vmatrix} -\lambda & 1 & 0 & 0 \\ 0 & -\lambda & 2 & 0 \\ \frac{3}{10} & \frac{3}{20} & \frac{1}{10} - \lambda & \frac{123}{40} \\ 0 & 0 & 0 & -\lambda \end{vmatrix} = -\lambda \left(-\lambda^3 + \frac{1}{10}\lambda^2 + \frac{3}{10}\lambda + \frac{6}{10} \right)$$

(onde aplicámos a regra de Laplace ao longo da última linha e depois a regra de Sarrus). O fator de grau 3 tem $\lambda = 1$ como raíz e portanto aplicando o algoritmo da divisão ou a regra de Ruffini vemos que o polinómio caracteristico é

$$\lambda^4 - \frac{4}{10}\lambda^2 + \frac{6}{10}\lambda = \lambda(\lambda - 1)\left(\lambda^2 + \frac{9}{10}\lambda + \frac{6}{10}\right)$$

O fator quadrático não tem raízes reais pelo que os únicos valores próprios de T são 0 e 1. Os vetores de coordenadas dos vetores próprios de 0 são os elementos não nulos (a, b, c, d) do núcleo da matriz, ou seja as soluções do sistema

$$\begin{cases} b = 0 \\ 2c = 0 \\ 12a + 6b + 4c + 123d = 0 \end{cases}$$

que são os vetores da forma $(-\frac{41}{4}d, 0, 0, d)$ com $d \neq 0$. Os vetores próprios de 0 são portanto os polinómios da forma $d(-\frac{41}{4} + x^3)$ com $d \neq 0$.

Os vetores das coordenadas dos vetores próprios de 1 são os elementos não nulos do núcleo da matriz

$$\begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 2 & 0 \\ \frac{3}{10} & \frac{3}{20} & -\frac{9}{10} & \frac{123}{40} \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

que são as soluções não nulas (a,b,c,d) do sistema

$$\begin{cases} a=b \\ b=2c \\ \frac{3}{10}a+\frac{3}{20}b-\frac{9}{10}c+\frac{123}{40}d=0 \\ d=0 \end{cases} \Leftrightarrow \begin{cases} a=2c \\ b=2c \\ d=0 \end{cases}$$

que são os vetores $c(2,2,1,0)$ com $c\neq 0$. Os vetores próprios de 1 são portanto os polinómios $c(2+2x+x^2)$ com $c\neq 0$. Uma vez que não existe uma base para B formada por vetores próprios, T não é diagonalizável.

2. Enunciado Consideramos a base $B=(\left[\begin{smallmatrix}1&0\\0&0\end{smallmatrix}\right],\left[\begin{smallmatrix}0&1\\0&0\end{smallmatrix}\right],\left[\begin{smallmatrix}0&0\\1&0\end{smallmatrix}\right],\left[\begin{smallmatrix}0&0\\0&1\end{smallmatrix}\right])$ para $M_{2\times 2}(\mathbb{R})$. Temos

$$T\left(\left[\begin{smallmatrix}1&0\\0&0\end{smallmatrix}\right]\right)=\left[\begin{smallmatrix}-1&0\\1&0\end{smallmatrix}\right]-\left[\begin{smallmatrix}-1&1\\0&0\end{smallmatrix}\right]=\left[\begin{smallmatrix}0&-1\\1&0\end{smallmatrix}\right],\qquad T\left(\left[\begin{smallmatrix}0&1\\0&0\end{smallmatrix}\right]\right)=\left[\begin{smallmatrix}0&-1\\0&1\end{smallmatrix}\right]-\left[\begin{smallmatrix}1&0\\0&0\end{smallmatrix}\right]=\left[\begin{smallmatrix}-1&-1\\0&1\end{smallmatrix}\right],$$

$$T\left(\left[\begin{smallmatrix}0&0\\1&0\end{smallmatrix}\right]\right)=\left[\begin{smallmatrix}1&0\\0&0\end{smallmatrix}\right]-\left[\begin{smallmatrix}0&0\\-1&1\end{smallmatrix}\right]=\left[\begin{smallmatrix}1&0\\1&-1\end{smallmatrix}\right],\qquad T\left(\left[\begin{smallmatrix}0&0\\0&1\end{smallmatrix}\right]\right)=\left[\begin{smallmatrix}0&1\\0&0\end{smallmatrix}\right]-\left[\begin{smallmatrix}0&0\\1&0\end{smallmatrix}\right]=\left[\begin{smallmatrix}0&1\\-1&0\end{smallmatrix}\right],$$

logo

$$A_{T,B,B}=\begin{bmatrix}0&-1&1&0\\-1&-1&0&1\\1&0&1&-1\\0&1&-1&0\end{bmatrix}$$

O polinómio característico de T é

$$\begin{vmatrix}-\lambda&-1&1&0\\-1&-1-\lambda&0&1\\1&0&1-\lambda&-1\\0&1&-1&-\lambda\end{vmatrix}=\lambda^4-5\lambda^2=\lambda^2(\lambda^2-5)$$

logo os valores próprios são 0 e $\pm\sqrt{5}$. O espaço próprio de 0 é o núcleo de T que, em coordenadas, é o conjunto das soluções do sistema

$$\begin{bmatrix}0&-1&1&0\\-1&-1&0&1\\1&0&1&-1\\0&1&-1&0\end{bmatrix}\to\begin{bmatrix}1&0&1&-1\\0&-1&1&0\\-1&-1&0&1\\0&1&-1&0\end{bmatrix}\to\begin{bmatrix}1&0&1&-1\\0&-1&1&0\\0&-1&1&0\\0&1&-1&0\end{bmatrix}\to\begin{bmatrix}1&0&1&-1\\0&-1&1&0\\0&0&0&0\\0&0&0&0\end{bmatrix}$$

O núcleo consiste portanto nos vetores com coordenadas (a,b,c,d) tais que $c=b$ e $d=a+c=a+b$, isto é, $(a,b,b,a+b)$. O espaço próprio de 0 é portanto

$$\left\{\left[\begin{smallmatrix}a&b\\b&a+b\end{smallmatrix}\right]:a,b\in\mathbb{R}\right\}=L\left(\left\{\left[\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right],\left[\begin{smallmatrix}0&1\\1&1\end{smallmatrix}\right]\right\}\right)$$

Alternativamente podíamos ter observado que claramente as matrizes $\left[\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right]$ e $\left[\begin{smallmatrix}-1&1\\1&0\end{smallmatrix}\right]$ estão no núcleo de T e, sendo linearmente independentes, têm de gerar o espaço próprio de 0.

Em coordenadas os valores próprios de $\sqrt{5}$ são os elementos do núcleo da matriz

$$\begin{bmatrix}-\sqrt{5}&-1&1&0\\-1&-1-\sqrt{5}&0&1\\1&0&1-\sqrt{5}&-1\\0&1&-1&-\sqrt{5}\end{bmatrix}\to\begin{bmatrix}1&0&1-\sqrt{5}&-1\\-\sqrt{5}&-1&1&0\\-1&-1-\sqrt{5}&0&1\\0&1&-1&-\sqrt{5}\end{bmatrix}$$

$$\to \begin{bmatrix} 1 & 0 & 1-\sqrt{5} & -1 \\ 0 & -1 & \sqrt{5}-4 & -\sqrt{5} \\ 0 & -1-\sqrt{5} & 1-\sqrt{5} & 0 \\ 0 & 1 & -1 & -\sqrt{5} \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 1-\sqrt{5} & -1 \\ 0 & 1 & -1 & -\sqrt{5} \\ 0 & -1 & \sqrt{5}-4 & -\sqrt{5} \\ 0 & -1-\sqrt{5} & 1-\sqrt{5} & 0 \end{bmatrix}$$

$$\to \begin{bmatrix} 1 & 0 & 1-\sqrt{5} & -1 \\ 0 & 1 & -1 & -\sqrt{5} \\ 0 & 0 & \sqrt{5}-5 & -2\sqrt{5} \\ 0 & 0 & -2\sqrt{5} & -5-\sqrt{5} \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 1-\sqrt{5} & -1 \\ 0 & 1 & -1 & -\sqrt{5} \\ 0 & 0 & \sqrt{5}-5 & -2\sqrt{5} \\ 0 & 0 & 0 & 0 \end{bmatrix} \overset{-\frac{5+\sqrt{5}}{20}L_3}{\longrightarrow} \begin{bmatrix} 1 & 0 & 1-\sqrt{5} & -1 \\ 0 & 1 & -1 & -\sqrt{5} \\ 0 & 0 & 1 & \frac{1+\sqrt{5}}{2} \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

As soluções são da forma $(-d, \frac{-1+\sqrt{5}}{2}d, -\frac{1+\sqrt{5}}{2}d, d)$ com $d \in \mathbb{R}$. Conclui-se que o espaço próprio de $\sqrt{5}$ é

$$L\left(\left\{\begin{bmatrix} -1 & \frac{-1+\sqrt{5}}{2} \\ -\frac{1+\sqrt{5}}{2} & 1 \end{bmatrix}\right\}\right)$$

Contas inteiramente análogas (basta trocar o sinal em todas as ocorrências de $\sqrt{5}$!) mostram que o espaço próprio de $-\sqrt{5}$ é

$$L\left(\left\{\begin{bmatrix} -1 & \frac{-1-\sqrt{5}}{2} \\ -\frac{1-\sqrt{5}}{2} & 1 \end{bmatrix}\right\}\right)$$

3. Enunciado Para calcular os valores próprios de T temos de usar uma representação matricial com respeito à mesma base no espaço de partida e chegada! Temos

$$A_{T,B_c,B_c} = S_{B\to B_c}A_{T,B_c,B} = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 2 & -2 \\ -1 & 1 & -1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 \\ 1 & 3 & -1 \\ 1 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 6 & -4 \\ 0 & 1 & -4 \end{bmatrix}$$

que tem polinómio característico $(1-\lambda)((6-\lambda)(-4-\lambda)+4) = (1-\lambda)(\lambda^2-2\lambda-20)$. Os valores próprios de T são portanto $\lambda = 1$ e $\lambda = \frac{2\pm\sqrt{4+80}}{2} = 1 \pm \sqrt{21}$.

4. Enunciado Na base $B = ((1,0,1),(0,1,1),(-1,0,1))$ a transformação T é representada por

$$A_{T,B,B} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

logo a representação com respeito à base canónica é dada por $A_{T,B_{can},B_{can}} = S_{B\to B_{can}}A_{T,B,B}S_{B_{can}\to B}$ Temos

$$S_{B\to B_{can}} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

e obtemos $S_{B_{can}\to B}$ calculando a inversa:

$$\begin{bmatrix} 1 & 0 & -1 & | & 1 & 0 & 0 \\ 0 & 1 & 0 & | & 0 & 1 & 0 \\ 1 & 1 & 1 & | & 0 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -1 & | & 1 & 0 & 0 \\ 0 & 1 & 0 & | & 0 & 1 & 0 \\ 0 & 1 & 2 & | & -1 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -1 & | & 1 & 0 & 0 \\ 0 & 1 & 0 & | & 0 & 1 & 0 \\ 0 & 0 & 2 & | & -1 & -1 & 1 \end{bmatrix}$$

$$\to \left[\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{array}\right]$$

Logo

$$\begin{aligned} A_{T,B_{can},B_{can}} &= \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 \\ -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 2 & -1 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 \\ -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} 0 & -1 & 1 \\ 0 & 2 & 0 \\ 1 & 2 & 0 \end{bmatrix} \end{aligned}$$

Conclui-se que T é dada pela expressão $T(x,y,z) = (-y+z, 2y, x+2y)$.

5. Enunciado A reflexão é completamente determinada por enviar o vetor perpendicular ao plano no seu simétrico e os vetores do plano neles próprios. Isto significa que $(1,1,1)$ (um vetor perpendicular ao plano) é um vetor próprio de -1 e que uma base do plano, por exemplo $\{(1,-1,0),(0,1,-1)\}$ consiste em vetores próprios de 1. Na base $B = ((1,-1,0),(0,1,-1),(1,1,1))$ para $\mathbb{R}^3$ temos portanto

$$A_{R,B,B} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

A representação com respeito à base canónica de $\mathbb{R}^3$ é portanto dada por

$$A_{R,B_{can},B_{can}} = \begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & 1 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & 1 \\ 0 & -1 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 & -1 \\ -1 & 1 & -1 \\ 0 & -1 & -1 \end{bmatrix} \frac{1}{3} \begin{bmatrix} 2 & -1 & -1 \\ 1 & 1 & -2 \\ 1 & 1 & 1 \end{bmatrix}$$

ou seja

$$A_{R,B_{can},B_{can}} = \frac{1}{3} \begin{bmatrix} 1 & -2 & -2 \\ -2 & 1 & -2 \\ -2 & -2 & 1 \end{bmatrix}$$

Logo a expressão geral para R é

$$R(x,y,z) = \tfrac{1}{3}(x-2y-2z, -2x+y-2z, -2x-2y+z)$$

6. Enunciado
   (a) Sim, uma vez que os espaços próprios correspondem a valores próprios distintos estes planos são distintos e têm intersecção nula (como também se pode facilmente verificar diretamente). Cada um dos subespaços tem dimensão 2 e uma vez que não se intersetam, a sua soma tem dimensão 4 e portanto consiste em todo o $\mathbb{R}^4$. Tomando a união de bases de cada um dos planos obtemos uma base para $\mathbb{R}^4$ formada por vetores próprios, logo T é diagonalizável.
   (b) Um vetor será um vetor próprio de T se e só se pertencer a um dos planos dados. De facto, qualquer vetor de $\mathbb{R}^4$ se pode escrever de forma única como $v_1 + v_2$ com v_i pertencente ao i-ésimo plano. Sendo λ_1, λ_2 os valores próprios correspondentes aos

dois planos obtemos $T(v_1+v_2) = \lambda_1 v_1 + \lambda_2 v_2$ que só é igual a $\mu(v_1+v_2)$ para algum μ se $v_1 = 0$ e $\mu = \lambda_2$ ou $v_2 = 0$ e $\mu = \lambda_1$ (se v_1 e v_2 forem ambos não nulos são necessariamente linearmente independentes e portanto duas combinações lineares de v_1 e v_2 só podem ser iguais se tiverem os mesmos coeficientes).

Resta portanto ver se cada um dos vetores pertence a algum dos planos: começando com $(1,1,-4,-1)$ temos

$$\begin{bmatrix} 1 & 1 & | & 1 \\ 2 & 3 & | & 1 \\ -1 & 2 & | & -4 \\ 0 & 1 & | & -1 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & | & 1 \\ 0 & 1 & | & -1 \\ 0 & 3 & | & -3 \\ 0 & 1 & | & -1 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & | & 1 \\ 0 & 1 & | & -1 \\ 0 & 0 & | & 0 \\ 0 & 0 & | & 0 \end{bmatrix}$$

Conclui-se que $(1,1,-4,-1) \in L(\{(1,2,-1,0),(1,3,2,1)\})$ e portanto é um vetor próprio.

Analogamente para $(3,2,1,-2)$ temos

$$\begin{bmatrix} 1 & 1 & | & 3 \\ 2 & 3 & | & 2 \\ -1 & 2 & | & 1 \\ 0 & 1 & | & -2 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & | & 3 \\ 0 & 1 & | & -4 \\ 0 & 3 & | & 4 \\ 0 & 1 & | & -2 \end{bmatrix}$$

O sistema anterior é impossível pelo que $(3,2,1,-2) \notin L(\{(1,2,-1,0),(1,3,2,1)\})$.

Finalmente

$$\begin{bmatrix} 0 & -1 & | & 3 \\ 1 & 2 & | & 2 \\ 0 & 0 & | & 1 \\ 1 & 0 & | & -2 \end{bmatrix}$$

é claramente um sistema impossível pelo que $(3,2,1,-2) \notin L(\{(0,1,0,1),(-1,2,0,0)\})$.

Conclui-se que $(3,2,1,-2)$ não é um vetor próprio de T.

7. Enunciado No Exercício 4.39 vimos que as soluções de $y'' - y = 0$ são $c_1 e^t + c_2 e^{-t}$. Isto diz-nos que 1 é um valor próprio de D^2 e o espaço próprio de 1 é $L(\{e^t, e^{-t}\})$. No exemplo do oscilador harmónico vimos que as soluções da equação $y'' + y = 0$ são $y(t) = c_1 \cos t + c_2 \operatorname{sen} t$ logo -1 é um valor próprio de D^2 com espaço próprio $L(\{\cos t, \operatorname{sen} t\})$.

Note-se que $\frac{d}{dt}^2 (y(\alpha t)) = \alpha^2 y(\alpha t)$ logo, por exemplo, $y(t)$ é uma solução de $y''+y=0$ se e só se $z(t) = y(\alpha t)$ é uma solução de $z'' + \alpha^2 z = 0$.

Desta forma vemos que, para $\lambda > 0$ as soluções de $(D^2+\lambda)f = 0$ são

$$f(t) = c_1 \cos(\sqrt{\lambda}t) + c_2 \operatorname{sen}(\sqrt{\lambda}t)$$

e portanto todos os números negativos são valores próprios de D^2 e o espaço próprio do número real negativo $-\lambda$ para D^2 é $L(\{\cos(\sqrt{\lambda}t), \operatorname{sen}(\sqrt{\lambda}t)\})$.

Analogamente todos os números reais positivos são valores próprios, sendo o espaço próprio de $\lambda > 0$, $L(\{e^{\sqrt{\lambda}t}, e^{-\sqrt{\lambda}t}\})$. Finalmente $\lambda = 0$ também é um valor próprio: o seu espaço próprio é o núcleo de D^2 que é o conjunto $L(\{1,t\})$ dos polinómios de grau ≤ 1.

8. Enunciado
   (a) Temos $(SAS^{-1})^k = (SAS^{-1})(SAS^{-1})\cdots(SAS^{-1}) = SA(S^{-1}S)A(S^{-1}\cdots S)AS^{-1} = SA^kS^{-1}$ logo

$$\begin{aligned} p(SAS^{-1}) &= a_0I + a_1SAS^{-1} + \ldots + a_n(SAS^{-1})^n \\ &= a_0SIS^{-1} + a_1SAS^{-1} + \ldots + a_nSA^nS^{-1} \\ &= S(a_0I + a_1A + \ldots + a_nA^n)S^{-1} \\ &= Sp(A)S^{-1} \end{aligned}$$

   (b) Seguindo a sugestão, diagonalizamos a matriz em questão. Os valores próprios são as soluções de $(1-\lambda)(-4-\lambda)+4 = \lambda^2+3\lambda = 0$ que são 0 e -3. O vetor $(2,1)$ é um vetor próprio de 0 e o vetor $(1,2)$ é um vetor próprio de -3. Logo

$$\begin{bmatrix}1 & -2\\ 2 & -4\end{bmatrix} = \begin{bmatrix}2 & 1\\ 1 & 2\end{bmatrix}\begin{bmatrix}0 & 0\\ 0 & -3\end{bmatrix}\begin{bmatrix}2 & 1\\ 1 & 2\end{bmatrix}^{-1} = \begin{bmatrix}2 & 1\\ 1 & 2\end{bmatrix}\begin{bmatrix}0 & 0\\ 0 & -3\end{bmatrix}\frac{1}{3}\begin{bmatrix}2 & -1\\ -1 & 2\end{bmatrix}$$

   e portanto

$$p\left(\begin{bmatrix}1 & -2\\ 2 & -4\end{bmatrix}\right) = \begin{bmatrix}2 & 1\\ 1 & 2\end{bmatrix}\begin{bmatrix}2 & 0\\ 0 & 2-3(-3)^4+(-3)^{11}\end{bmatrix}\frac{1}{3}\begin{bmatrix}2 & -1\\ -1 & 2\end{bmatrix}$$

9. Enunciado Pelos cálculos efetuados nesse exercício temos

$$A = \begin{bmatrix}1 & 3\\ 3 & 1\end{bmatrix} = \begin{bmatrix}1 & -1\\ 1 & 1\end{bmatrix}\begin{bmatrix}4 & 0\\ 0 & -2\end{bmatrix}\begin{bmatrix}1 & -1\\ 1 & 1\end{bmatrix}^{-1} = \begin{bmatrix}1 & -1\\ 1 & 1\end{bmatrix}\begin{bmatrix}4 & 0\\ 0 & -2\end{bmatrix}\frac{1}{2}\begin{bmatrix}1 & 1\\ -1 & 1\end{bmatrix}$$

   Logo, pela alínea (a) do exercício anterior temos

$$A^{721} = \begin{bmatrix}1 & -1\\ 1 & 1\end{bmatrix}\begin{bmatrix}4^{721} & 0\\ 0 & -2^{721}\end{bmatrix}\frac{1}{2}\begin{bmatrix}1 & 1\\ -1 & 1\end{bmatrix}$$

10. Enunciado
   (a) Podemos escrever o sistema na forma

$$\frac{d}{dt}\left(\begin{bmatrix}x(t)\\ y(t)\end{bmatrix}\right) = \begin{bmatrix}1 & 2\\ 2 & 1\end{bmatrix}\begin{bmatrix}x(t)\\ y(t)\end{bmatrix} \tag{79}$$

   Se mudarmos de coordenadas para uma base formada por vetores próprios da matriz dos coeficientes, o sistema vai transformar-se em duas equações independentes, uma em cada variável:
   Os valores próprios da matriz são as soluções de $(1-\lambda)^2 - 4 = 0 \Leftrightarrow \lambda = 1 \pm 2$, ou seja 3 e -1. O vetor $(1,1)$ é um vetor próprio de 3, enquanto $(1,-1)$ é um vetor próprio de -1. Logo

$$\begin{bmatrix}1 & 2\\ 2 & 1\end{bmatrix} = \begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}\begin{bmatrix}3 & 0\\ 0 & -1\end{bmatrix}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}^{-1}$$

Nas coordenadas (u, v) da base dada pelos vetores próprios, que estão relacionadas com (x, y) pelas fórmulas

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} \Leftrightarrow \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}^{-1} \begin{bmatrix} x \\ y \end{bmatrix}$$

o sistema transforma-se em

$$\frac{d}{dt}\left(\begin{bmatrix} u(t) \\ v(t) \end{bmatrix}\right) = \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} u(t) \\ v(t) \end{bmatrix}$$

De facto, o sistema acima obtém-se do sistema (79) fazendo a mudança de variável indicada (isto é substituindo $x(t)$ e $y(t)$ pelas combinações lineares relevantes de $u(t)$ e $v(t)$):

$$\begin{aligned}
\frac{d}{dt}\left(\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} u(t) \\ v(t) \end{bmatrix}\right) &= \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} u(t) \\ v(t) \end{bmatrix} \\
\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\frac{d}{dt}\left(\begin{bmatrix} u(t) \\ v(t) \end{bmatrix}\right) &= \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}^{-1}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} u(t) \\ v(t) \end{bmatrix} \\
\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\frac{d}{dt}\left(\begin{bmatrix} u(t) \\ v(t) \end{bmatrix}\right) &= \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} u(t) \\ v(t) \end{bmatrix} \\
\frac{d}{dt}\left(\begin{bmatrix} u(t) \\ v(t) \end{bmatrix}\right) &= \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} u(t) \\ v(t) \end{bmatrix}
\end{aligned}$$

que tem solução

$$\begin{cases} u(t) = c_1 e^{3t} \\ v(t) = c_2 e^{-t} \end{cases} \qquad c_1, c_2 \in \mathbb{R}$$

Dado que $x(t) = u(t) + v(t), y(t) = u(t) - v(t)$ obtemos a seguinte solução para o sistema inicial

$$\begin{cases} x(t) = c_1 e^{3t} + c_2 e^{-t} \\ y(t) = c_1 e^{3t} - c_2 e^{-t} \end{cases} \qquad c_1, c_2 \in \mathbb{R}$$

(b) Tal como na alínea anterior diagonalizamos a matriz dos coeficientes do sistema

$$\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

Os valores próprios são as soluções de $(1-\lambda)^2 - 1 = 0 \Leftrightarrow \lambda = 0, 2$. $(1, 1)$ é um vetor próprio de 0 e $(1, -1)$ é um vetor próprio de 2. Nas coordenadas (u, v) dadas pela base dos vetores próprios o sistema escreve-se

$$\begin{cases} \frac{du}{dt} = 0 \\ \frac{dv}{dt} = 2v \end{cases} \Leftrightarrow \begin{cases} u(t) = c_1 \\ v(t) = c_2 e^{2t} \end{cases} \qquad c_1, c_2 \in \mathbb{R}$$

A relação entre as coordenadas é

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}$$

logo a solução do sistema é

$$\begin{cases} x(t) = c_1 + c_2 e^{2t} \\ y(t) = c_1 - c_2 e^{2t} \end{cases} \qquad c_1, c_2 \in \mathbb{R}$$

11. Enunciado Suponhamos que P é uma projeção. Então vimos no Exercício 4.25 que o espaço vetorial V se decompõe na soma direta $V = N(P) \oplus P(V)$ sendo os elementos de $P(V)$ vetores próprios de 1 (uma vez que $P(P(v)) = P(v) = 1 \cdot P(v)$) e os vetores do núcleo vetores próprios de 0. Tomando a união de bases para o núcleo e imagem de P obtemos uma base de V formada por vetores próprios de 0 e de 1 pelo que P é diagonalizável e tem como valores próprios 0 e 1.

    Reciprocamente se $P\colon V \to V$ é uma transformação linear diagonalizável que tem apenas 0 e 1 como valores próprios, seja $B = (v_1, \dots, v_n)$ uma base ordenada de V formada por vetores próprios de P. Podemos escolher a ordenação de tal forma que $v_1, \dots, v_k$ sejam vetores próprios de 0 e os restantes elementos da base sejam vetores próprios de 1. Para ver que P é uma projeção temos que verificar a equação $P^2 = P$, para o qual é suficiente ver que P^2 e P têm o mesmo efeito na base B:

$$P^2(v_i) = P(P(v_i)) = \begin{cases} P(0) = 0 = P(v_i) & \text{se } 1 \leq i \leq k \\ P(1 \cdot v_i) = P(v_i) & \text{se } i > k \end{cases}$$

    conforme pretendido.
12. Enunciado
    (a) Seja $\lambda = a + bi$ com $b \neq 0$. Imaginemos por absurdo que $v \in M_{n\times 1}(\mathbb{C})$ era um vetor próprio de λ com todas as componentes reais. Uma vez que A é uma matriz real temos que Av seria também um vetor real mas $\lambda v = (a + bi)v$ tem pelo menos uma entrada com parte imaginária não nula (pois v é real e não nulo e $b \neq 0$). É portanto impossível que $Av = \lambda v$ contráriamente à nossa hipótese inicial.
    (b) Uma vez que a conjugação preserva a soma e o produto por escalares o mesmo é verdade para a soma e produto de matrizes. Como A é real temos $A = \overline{A}$ e portanto

$$\overline{Av} = \overline{A}\,\overline{v} = A\overline{v}$$

    Por outro lado

$$\overline{Av} = \overline{\lambda v} = \overline{\lambda}\,\overline{v}$$

    Conclui-se que se v é um vetor próprio de λ então $\overline{v}$ é um vetor próprio de $\overline{\lambda}$ (note-se que $\overline{v} \neq 0$ porque $v \neq 0$).
    (c) Os valores próprios são as soluções de $(1-\lambda)^2 + 1 = 0 \Leftrightarrow \lambda = 1 \pm i$. Os vetores próprios de $1 + i$ são as soluções de

$$\begin{bmatrix} -i & -1 \\ 1 & -i \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = 0 \Leftrightarrow b = -ia.$$

Conclui-se que $(1,-i)$ é um vetor próprio de $1+i$ e portanto $(1,i)=\overline{(1,-i)}$ é um vetor próprio de $1-i$. Temos portanto

$$\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix} \begin{bmatrix} 1+i & 0 \\ 0 & 1-i \end{bmatrix} \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix}^{-1}$$

13. Enunciado

(a) Por indução em n, sendo que o caso $n=1$ é imediato. Assumimos que o resultado é válido para um conjunto de $(n-1)$ vetores como no enunciado.
Sejam $v_1,\ldots,v_n$ como no enunciado (e portanto distintos). Sejam $\alpha_1,\ldots,\alpha_n \in \mathbb{K}$ tais que

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n = 0$$

Aplicando a transformação linear $(T-\lambda_1 \operatorname{Id})$ à equação anterior eliminamos o vetor v_1:

$$\begin{aligned} & (T-\lambda_1)(\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n) = 0 \\ \Leftrightarrow \quad & \alpha_1 (T-\lambda_1)(v_1) + \ldots + \alpha_n (T-\lambda_1)(v_n) = 0 \\ \Leftrightarrow \quad & \alpha_1 \cdot 0 + \alpha_2 (T-\lambda_1)(v_2) + \ldots + \alpha_n (T-\lambda_1)(v_n) = 0 \end{aligned}$$

Notando que

$$(T-\lambda_1)(v_i) = T(v_i) - \lambda_1 v_i = \lambda_i v_i - \lambda_1 v_i = (\lambda_i - \lambda_1) v_i$$

obtemos a equação

$$\alpha_2(\lambda_2-\lambda_1)v_2 + \ldots + \alpha_n(\lambda_n-\lambda_1)v_n = 0$$

Uma vez que, por hipótese de indução $v_2,\ldots,v_n$ são linearmente independentes vemos que

$$\alpha_2(\lambda_2-\lambda_1) = \ldots = \alpha_n(\lambda_n-\lambda_1) = 0$$

Mas por hipótese $\lambda_2-\lambda_1 \neq 0, \ldots, \lambda_n-\lambda_1 \neq 0$ logo $\alpha_2 = \ldots = \alpha_n = 0$. Consequentemente temos também $\alpha_1 = 0$ o que mostra que $v_1,\ldots,v_n$ são linearmente independentes (conforme pretendido).

(b) Sejam $\lambda_1,\ldots,\lambda_k$ os vetores próprios de T (assumidos distintos). A alínea anterior mostra que a soma dos espaços próprios $E(\lambda_i)$ é direta. De facto, se $x = v_1+\ldots+v_k$ com $v_i \in E(\lambda_i)$ então os vetores v_i são únicos: dada outra decomposição $x = v_1'+\ldots+v_k'$ com $v_i' \in E(\lambda_i)$ teríamos

$$0 = (v_1-v_1') + (v_2-v_2') + \ldots + (v_k-v_k)'$$

e se algum dos vetores $v_i - v_i'$ (que pertence a $E(\lambda_i)$) fosse não nulo teríamos uma combinação linear de vetores próprios de T com valores próprios distintos que se anulava apesar de os coeficientes da combinação não serem todos nulos.
Se T é diagonalizável, existe uma base para V formada por vetores próprios. Logo $E(\lambda_1)+\ldots+E(\lambda_k) = V$. Uma vez que a soma é direta, uma base para V pode obter-se reunindo bases de cada um dos subespaços $E(\lambda_i)$ e portanto $\dim(V) = \dim E(\lambda_1) + \ldots + \dim E(\lambda_k)$, isto é, $\dim(V)$ é a soma das multiplicidades geométricas.

Reciprocamente se a soma das multiplicidades geométricas é $\dim(V)$ então $\dim E(\lambda_1)\oplus \ldots \oplus E(\lambda_k) = \dim E(\lambda_1) + \ldots + \dim E(\lambda_k) = \dim(V)$ pelo que o subespaço $E(\lambda_1)\oplus\ldots\oplus E(\lambda_k)$ tem que coincidir com V. Reunindo bases de cada um dos $E(\lambda_i)$ obtemos uma base para V formada por vetores próprios pelo que T é diagonalizável.

14. Enunciado A matriz de Markov que codifica a probabilidade de transição entre as páginas é

$$M = \begin{bmatrix} 0 & 0 & 1 & 0 \\ \frac{1}{2} & 0 & 0 & 0 \\ 0 & \frac{2}{3} & 0 & 1 \\ \frac{1}{2} & \frac{1}{3} & 0 & 0 \end{bmatrix}$$

O vetor próprio de 1 que codifica o estado estacionário é a solução do seguinte sistema que tem soma das entradas igual a 1:

$$\begin{bmatrix} -1 & 0 & 1 & 0 \\ \frac{1}{2} & -1 & 0 & 0 \\ 0 & \frac{2}{3} & -1 & 1 \\ \frac{1}{2} & \frac{1}{3} & 0 & -1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = 0 \Leftrightarrow \begin{cases} c = a \\ b = \frac{a}{2} \\ d = c - \frac{2}{3}b = \frac{2}{3}a \end{cases}$$

Os vetores próprios de 1 são os múltiplos de $(1, \frac{1}{2}, 1, \frac{2}{3})$ pelo que o estado estacionário normalizado é

$$\frac{1}{1 + \frac{1}{2} + 1 + \frac{2}{3}}(1, \tfrac{1}{2}, 1, \tfrac{2}{3}) = (\tfrac{6}{19}, \tfrac{3}{19}, \tfrac{6}{19}, \tfrac{4}{19})$$

Conclui-se que as páginas mais relevantes são a 1 e a 3 igualmente (um(a) internauta passeando ao acaso passa aproximadamente $\frac{6}{19} \simeq 32\%$ do tempo em cada uma dessas páginas).

15. Enunciado Em geral, numa transformação linear diagonalizável, em que temos $V = E(\lambda_1) \oplus \cdots \oplus E(\lambda_k)$ um subespaço $W \subset V$ é invariante se e só se W é a soma das suas intersecções com os espaços próprios, isto é, se

$$W = (W \cap E(\lambda_1)) \oplus \cdots \oplus (W \cap E(\lambda_k))$$

De facto, uma soma de espaços invariantes é invariante e qualquer subespaço de $E(\lambda_i)$ é invariante por T (uma vez que T atua em $E(\lambda_i)$ multiplicando pelo escalar λ_i). Por outro lado, se W é invariante e escrevermos

$$w = w_1 + w_2 + \ldots + w_k$$

com $w_i \in E(\lambda_i)$ então, uma vez que W é invariante temos que $(T-\lambda_2)\cdots(T-\lambda_k)(w) \in W$. Mas

$$\begin{aligned} (T-\lambda_2)\cdots(T-\lambda_k)(w) &= (T-\lambda_2)\cdots(T-\lambda_k)(w_1) + \ldots + (T-\lambda_2)\cdots(T-\lambda_k)(w_k) \\ &= (\lambda_1-\lambda_2)\cdots(\lambda_1-\lambda_k)w_1 + 0 + \ldots + 0 \end{aligned}$$

logo $w_1 \in W$. Da mesma forma vemos que todos os termos $w_i \in W$. Portanto W é a soma das suas intersecções com os $E(\lambda_i)$.

No exemplo em questão temos portanto que além dos subespaços invariantes triviais $\{0\}$ e $M_{2\times 2}(\mathbb{R})$, os subespaços invariantes de dimensão 1 são $E(\sqrt{5})$, $E(-\sqrt{5})$ e qualquer reta que passe pela origem em $E(0) = N(T)$. Os subespaços invariantes de dimensão 2 são $E(0)$, $E(\sqrt{5}) \oplus E(-\sqrt{5})$ e a soma de qualquer reta que passe pela origem em $E(0)$ com $E(\sqrt{5})$ ou $E(-\sqrt{5})$. Finalmente os subespaços invariantes de dimensão 3 são $E(0) \oplus E(\sqrt{5})$, $E(0) \oplus E(-\sqrt{5})$ e a soma de qualquer reta que passe pela origem em $E(0)$ com $E(\sqrt{5}) \oplus E(-\sqrt{5})$.

16. Enunciado Suponhamos que W_α é uma família de espaços invariantes. Dado $w \in \cap_\alpha W_\alpha$ temos que $w \in W_\alpha$ para todo o α (por definição de intersecção de conjuntos). Uma vez que W_α é invariante vemos que $T(w) \in W_\alpha$ para todo o α e portanto que $T(w) \in \cap_\alpha W_\alpha$. Isto mostra que $\cap W_\alpha$ (que pelo Exercício 3.10 é um subespaço de V) é um subespaço invariante.

    Analogamente seja $W = \sum W_\alpha$ o conjunto de todos os vetores de V que se podem escrever como somas (finitas) de elementos dos subespaços W_α. Dado $w \in W$ existe um conjunto finito $\alpha_1, \ldots, \alpha_k$ de índices tais que $w = w_1 + \ldots + w_k$ com $w_i \in W_{\alpha_i}$. Então

$$T(w) = T(w_1 + \ldots + w_k) = T(w_1) + \ldots + T(w_k)$$

    Uma vez que W_{α_i} são invariantes, $T(w_i) \in W_{\alpha_i}$. Conclui-se que $T(w)$ é uma soma finita de vetores que pertencem aos espaços W_α e portanto pertence a W. Logo W é um subespaço invariante.

17. Enunciado
    (a) O polinómio característico é

$$\begin{vmatrix} 2-\lambda & 1 & 0 \\ 1 & 2-\lambda & -1 \\ 1 & 1 & 1-\lambda \end{vmatrix} = (2-\lambda)^2(1-\lambda) - 1 + 0 - 0 - (1-\lambda) + 2 - \lambda = (2-\lambda)^2(1-\lambda)$$

logo os valores próprios são 2 e 1 com multiplicidades algébricas 2 e 1 respetivamente. Os vetores próprios de 2 são dados por

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & -1 \\ 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 0 \Leftrightarrow \begin{cases} b = 0 \\ a = c \end{cases}$$

Logo a multiplicidade geométrica de 2 é 1. Conclui-se que a forma canónica de Jordan tem um único bloco com $\lambda = 2$ na diagonal, de tamanho 2 e portanto

$$J = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Os vetores próprios de 1 são as soluções de

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & -1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 0 \Leftrightarrow \begin{cases} b = -a \\ c = 0 \end{cases}$$

Sendo $B = (v_1, v_2, v_3)$ uma base com respeito à qual a matriz está em forma canónica de Jordan podemos tomar $v_1 = (1, 0, 1), v_3 = (1, -1, 0)$ e para v_2 uma solução da equação $(A-2I)v_2 = v_1$:

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & -1 \\ 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

Por exemplo $v_3 = (0,1,0)$. Conclui-se que os espaços próprios generalizados são

$$E^g(1) = E(1) = L(\{(1,-1,0)\}), \qquad E^g(2) = L(\{(1,0,1),(0,1,0)\})$$

(b) O polinómio característico é

$$\begin{vmatrix} 3 & 2 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -1 & 3 \end{vmatrix} = (3-\lambda)^2((1-\lambda)(3-\lambda)+1) = (3-\lambda)^2(\lambda^2-4\lambda+4) = (\lambda-2)^2(\lambda-3)^2$$

Os vetores próprios de 2 são os vetores da forma $(0,0,c,c)$ com $c \neq 0$ e os vetores próprios de 3 são os vetores da forma $(a,0,0,0)$ com $a \neq 0$. Uma vez que ambos os vetores próprios têm multiplicidade algébrica 2 e multiplicidade geométrica 1, a forma canónica de Jordan é

$$J = \begin{bmatrix} 3 & 1 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

Sendo $B = (v_1, v_2, v_3, v_4)$ uma base que coloca a matriz nesta forma podemos tomar $v_1 = (1,0,0,0)$ e $v_3 = (0,0,1,1)$, sendo v_2 e v_4 soluções de $(A-3I)v_2 = v_1$ e $(A-2I)v_4 = v_3$ ou seja

$$\begin{bmatrix} 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad \begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$$

Podemos portanto tomar por exemplo $v_2 = (0, \frac{1}{2}, 0, 0)$ e $v_4 = (0,0,0,1)$. Conclui-se que os espaços próprios generalizados são

$$E^g(3) = L(\{(1,0,0,0),(0,1,0,0)\}) \qquad E^g(2) = L(\{(0,0,1,0),(0,0,0,1)\})$$

(c) O polinómio característico é

$$\begin{vmatrix} \frac{5}{2}-\lambda & 0 & -\frac{1}{2} \\ -1 & 2-\lambda & 1 \\ \frac{1}{2} & 0 & \frac{3}{2} \end{vmatrix} = (2-\lambda)((\tfrac{5}{2}-\lambda)(\tfrac{3}{2}-\lambda)+\tfrac{1}{4}) = (2-\lambda)(\lambda^2-4\lambda+4) = (2-\lambda)^3$$

Os vetores próprios de 2 são as soluções de

$$\begin{bmatrix} \frac{1}{2} & 0 & -\frac{1}{2} \\ -1 & 0 & 1 \\ \frac{1}{2} & 0 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 0 \Leftrightarrow c = a$$

Portanto $E(2) = \{(a,b,a)\colon a,b \in \mathbb{R}\}$ tem dimensão 2. Uma vez que a multiplicidade geométrica de $\lambda = 2$ é 2 a forma canónica de Jordan da matriz é (a menos de uma possível troca da ordem dos blocos)

$$J = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

O espaço próprio generalizado de 2 é todo o $\mathbb{R}^3$. Apesar de não ser pedido vamos achar a base $B = (v_1, v_2, v_3)$ com respeito à qual a matriz fica nesta forma canónica. Os vetores v_1 e v_3 serão uma base para o espaço próprio de 3 e o vetor v_2 tem que satisfazer a equação $(A - 2I)v_2 = v_1$. Para que esta equação tenha solução v_1 terá de pertencer ao espaço das colunas de $(A - 2I)$ que claramente é $L(\{(\frac{1}{2}, -1, \frac{1}{2})\})$. Podemos assim tomar $v_1 = (\frac{1}{2}, -1, \frac{1}{2})$ e achar v_2 resolvendo a equação

$$\begin{bmatrix} \frac{1}{2} & 0 & -\frac{1}{2} \\ -1 & 0 & 1 \\ \frac{1}{2} & 0 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ -1 \\ \frac{1}{2} \end{bmatrix}$$

Uma possível solução é $v_2 = (1, 0, 0)$. Para v_3 podemos tomar qualquer vetor próprio de 2 que seja linearmente independente de v_1, por exemplo $v_3 = (0, 1, 0)$.

(d) O polinómio característico é

$$\begin{vmatrix} 2-\lambda & 1 & 0 & 0 \\ 1 & 2-\lambda & 1 & 0 \\ 0 & -1 & 2-\lambda & 0 \\ 0 & 1 & 0 & 2-\lambda \end{vmatrix} = (2-\lambda) \begin{vmatrix} 2-\lambda & 1 & 0 \\ 1 & 2-\lambda & 1 \\ 0 & -1 & 2-\lambda \end{vmatrix} = (\lambda - 2)^4$$

O espaço próprio generalizado de 2 é todo o $\mathbb{R}^4$ uma vez que só há um valor próprio. Os vetores próprios de 2 são as soluções de

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = 0 \Leftrightarrow \begin{cases} b = 0 \\ c = -a \end{cases}$$

Logo $E(2) = \{(a, 0, -a, d)\colon a, d \in \mathbb{R}\}$. Uma vez que a multiplicidade geométrica de $\lambda = 2$ é 2, há dois blocos de Jordan logo temos as seguintes duas possibilidades para a forma canónica de Jordan (a menos de troca de ordem dos blocos):

$$J = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix} \qquad \text{ou} \qquad J = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

No primeiro caso temos $(J-2I)^2=0$ e portanto $(A-2I)^2=0$, o que não acontece no segundo caso. Uma vez que

$$(A-2I)^2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & -1 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix}$$

é não nula concluímos que a forma canónica de Jordan é a segunda das opções indicadas acima. Embora não seja pedido vamos achar a base $B=(v_1,v_2,v_3,v_4)$ que coloca a matriz nesta forma. Os vetores v_1 e v_4 formam uma base para o espaço próprio enquanto que v_2 e v_3 têm de satisfazer as equações $(A-2I)v_2=v_1$ e $(A-2I)v_3=v_2$. Estas equações implicam que $(A-2I)^2v_3=v_1$ logo v_1 terá de estar no espaço das colunas de $(A-2I)^2$. Desde que esta condição seja verificada conseguiremos resolver estas equações.
Podemos portanto tomar $v_1=(1,0,-1,1)$ e então $v_3=(1,0,0,0)$, por exemplo, será uma solução de $(A-2I)^2v_3=v_1$. Para v_2 podemos então tomar $(A-2I)v_3=(0,1,0,0)$. Para v_4 podemos tomar qualquer vetor próprio de 2 não colinear com v_1, por exemplo $v_4=(1,0,-1,0)$.

18. Enunciado
    (a) O polinómio característico é

$$\begin{vmatrix} -1-\lambda & 2 \\ 3 & 1-\lambda \end{vmatrix} = (\lambda^2-1)-6=\lambda^2-7$$

logo os valores próprios são $\pm\sqrt{7}$ ambos com multiplicidade algébrica, e portanto geométrica, igual a 1.

    (b) O polinómio característico é

$$\begin{vmatrix} 1-\lambda & 0 & 0 \\ 7 & -2-\lambda & 0 \\ 1 & 9 & -2-\lambda \end{vmatrix} = (1-\lambda)(2+\lambda)^2$$

logo os valores próprios são $\lambda=1$ e $\lambda=-2$. A multiplicidade algébrica e geométrica de 1 é igual a 1. A multiplicidade algébrica de -2 é 2. Os vetores próprios de -2 são as soluções não nulas de $(A+2I)v=0$. Uma vez que $A+2I$ tem característica 2, o espaço próprio de -2 tem dimensão 1 e portanto a multiplicidade geométrica de -2 é 1.

    (c) O polinómio característico é

$$\begin{vmatrix} -1-\lambda & 0 & 1 \\ 1 & -\lambda & -1 \\ -1 & -1 & 2-\lambda \end{vmatrix} = (1+\lambda)\lambda(2-\lambda)-1-\lambda+1+\lambda=\lambda(1+\lambda)(2-\lambda)$$

Portanto os valores próprios são $0,-1$ e 2 todos com multiplicidade algébrica e geométrica 1.

(d) O polinómio característico é

$$\begin{vmatrix} -1-\lambda & -2 & -2 \\ 0 & 2-\lambda & 1 \\ 0 & -1 & 2-\lambda \end{vmatrix} = -(1+\lambda)((2-\lambda)^2+1)$$

logo os valores próprios (enquanto matriz complexa) são -1 e $2 \pm i$ todos com multiplicidade algébrica e portanto geométrica iguais a 1.

(e) O polinómio característico é

$$\begin{aligned} \begin{vmatrix} -1-\lambda & -2 & -2 \\ 1 & 2-\lambda & 1 \\ -1 & -1 & -\lambda \end{vmatrix} &= (\lambda^2+\lambda)(2-\lambda)+2+2-2(2-\lambda)-2\lambda-(1+\lambda) \\ &= (1+\lambda)\left(\lambda(2-\lambda)-1\right) = -(1+\lambda)(\lambda-1)^2 \end{aligned}$$

logo os valores próprios são -1 com multiplicidade algébrica e geométrica 1 e 1 com multiplicidade algébrica 2. Os vetores próprios de 1 são as soluções não nulas de

$$\begin{bmatrix} -2 & -2 & -2 \\ 1 & 1 & 1 \\ -1 & -1 & -1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 0 \Leftrightarrow c = a+b$$

Logo o espaço próprio de 1 é $E(1) = \{(a, b, -a-b) \colon a, b \in \mathbb{R}\}$. Como $E(2)$ tem dimensão 2, a multiplicidade geométrica de 1 é também 2.

(f) O polinómio característico é

$$\begin{aligned} \begin{vmatrix} 1-\lambda & 1 & 0 & 1 \\ 2 & 2-\lambda & -1 & 0 \\ 0 & 0 & 1-\lambda & 2 \\ 0 & 0 & 3 & -\lambda \end{vmatrix} &= ((1-\lambda)(2-\lambda)-2)(\lambda(\lambda-1)-6) \\ &= (\lambda^2-3\lambda)(\lambda^2-\lambda-6) = \lambda(\lambda-3)^2(\lambda+2) \end{aligned}$$

logo os valores próprios são $0, 3$ e -2. As multiplicidades algébricas e geométricas de 0 e -2 são ambas 1 e a multiplicidade algébrica de 3 é 2. Os vetores próprios de 3 são as soluções não nulas de

$$\begin{bmatrix} -2 & 1 & 0 & 1 \\ 2 & -1 & -1 & 0 \\ 0 & 0 & -2 & 2 \\ 0 & 0 & 3 & -3 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = 0 \Leftrightarrow \begin{cases} b = 2a-c \\ c = d \end{cases}$$

Logo $E(3) = \{(a, 2a-c, c, c) \colon a, c \in \mathbb{R}\}$ tem dimensão 2 pelo que a multiplicidade geométrica de 3 é também 2.

19. Enunciado

(a) A entrada ii de AB é dada por $\sum_{j=1}^{n} a_{ij}b_{ji}$ logo

$$\operatorname{tr}(AB) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}b_{ji} = \sum_{j=1}^{n}\sum_{i=1}^{n} b_{ji}a_{ij} = \operatorname{tr}(BA)$$

(onde usámos a comutatividade e associatividade da soma de escalares, e a comutatividade do produto de escalares).

(b) Pela alínea anterior temos $\operatorname{tr}(SAS^{-1}) = \operatorname{tr}((AS^{-1})S) = \operatorname{tr}(A)$.

(c) Dada uma matriz $A \in M_{n\times n}(\mathbb{C})$, a forma canónica de Jordan garante a existência de uma matriz invertível S tal que SAS^{-1} está em forma canónica de Jordan. Uma vez que os valores próprios de A aparecem na diagonal de SAS^{-1} repetidos de acordo com a sua multiplicidade algébrica, temos o resultado pretendido. A afirmação relativa ao determinante é uma consequência da igualdade

$$\det(SAS^{-1}) = \det(S)\det(A)\det(S^{-1}) = \det(S)\det(A)\frac{1}{\det(S)} = \det(A)$$

Uma vez que a forma canónica de Jordan SAS^{-1} é triangular superior, o seu determinante é o produto das entradas na diagonal, ou seja, é o produto dos valores próprios de A repetidos de acordo com a sua multiplicidade algébrica.

(d) Sendo λ_1, λ_2 as entradas na diagonal da forma canónica de Jordan das matrizes dadas temos

(i) $\lambda_1 + \lambda_2 = 1 + 4 = 5$ e $\lambda_1\lambda_2 = 4 - 4 = 0$. Logo um dos valores próprios é 0 e o outro é então necessariamente igual a 5.

(ii) $\lambda_1 + \lambda_2 = 3 - 1 = 2$ e $\lambda_1\lambda_2 = -3 + 4 = 1$. As únicas soluções deste sistema são $\lambda_1 = \lambda_2 = 1$ (pois $\lambda_1 + \frac{1}{\lambda_1} = 2 \Leftrightarrow (\lambda_1 - 1)^2 = 0$) logo o único valor próprio de A é 1. Uma vez que A não é a matriz identidade, vemos que A não é diagonalizável pelo que a multiplicidade geométrica de 1 é apenas 1 e a multiplicidade algébrica é 2.

20. Enunciado Se A e B são matrizes triangulares superiores (respetivamente inferiores) a entrada ii do produto é dada por $a_{ii}b_{ii}$ - os restantes termos na fórmula para esta entrada no produto de matrizes anulam-se devido aos zeros abaixo (respetivamente acima) da diagonal. Logo se alguma das entradas diagonais de $A \in M_{n\times n}(\mathbb{K})$ for não nula, a mesma entrada diagonal de A^k será não nula para todo o $k \geq 1$. Portanto A não será nilpotente. Isto mostra que se A é nilpotente então as entradas diagonais são nulas.

Reciprocamente se $A \in M_{n\times n}(\mathbb{K})$ é triangular (superior digamos - podemos deduzir o resultado correspondente para matrizes triangulares inferiores transpondo) e as entradas diagonais são todas nulas vamos ver que $A^n = 0$ pelo que A é nilpotente.

Isto é uma consequência do seguinte facto que iremos demonstrar: quando multiplicamos uma matriz A que consiste em 0s até m linhas acima da diagonal (isto é, $a_{ij} = 0$ para $i \geq j - m$) por uma matriz que consiste em 0s até p linhas acima da diagonal (isto é $b_{ij} = 0$ para $i \geq j - p$) então AB consiste em 0s até $p+m+1$ linhas acima da diagonal.

Aplicando isto a uma matriz triangular superior A com 0s na diagonal vemos que A^2 consiste em 0s até à linha acima da diagonal (inclusivamente), A^3 terá duas linhas de 0s acima da diagonal, etc... até que finalmente $A^n = 0$

Sejam então A e B matrizes $n \times n$ com $a_{ij} = 0$ para $i \geq j - m$ e $b_{ij} = 0$ para $i \geq j - p$. Vamos calcular a entrada ij de AB assumindo que $i \geq j - m - p - 1$:

$$(AB)_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} = \sum_{k=1}^{i+m} a_{ik} b_{kj} + \sum_{k=i+m+1}^{n} a_{ik} b_{kj}$$

O primeiro dos somatórios à direita é zero porque por hipótese $a_{ik} = 0$ para $i \geq k - m \Leftrightarrow k \leq i + m$. O segundo dos somatórios é zero porque $k \geq i + m + 1$ e $i \geq j - m - p - 1$ (que assumimos) implica $k \geq j - p$ e portanto temos $b_{kj} = 0$. Conclui-se que $(AB)_{ij} = 0$ para $i \geq j - p - m - 1$ conforme pretendido.

21. Enunciado Seja $B = (1, t, t^2)$ a base canónica dos polinómios de grau ≤ 2. Uma vez que $T(1) = 1 + t + t^2, T(t) = t + t^2$ e $T(t^2) = 2t^2$ vemos que

$$A_{T,B,B} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 2 \end{bmatrix}$$

Os valores próprios de T são as soluções de

$$\begin{vmatrix} 1-\lambda & 0 & 0 \\ 1 & 1-\lambda & 0 \\ 1 & 1 & 2-\lambda \end{vmatrix} = 0 \Leftrightarrow (1-\lambda)^2(2-\lambda) = 0$$

Uma vez que os valores próprios da matriz são reais, a forma canónica de Jordan da matriz enquanto matriz complexa será real e, uma vez que podemos escolher vetores reais para a base que coloca a matriz em forma canónica de Jordan, T vai também admitir uma forma canónica de Jordan (a base que coloca T em forma canónica de Jordan é formada pelos vetores cujas coordenadas colocam a representação matricial em forma canónica de Jordan).

Os vetores próprios de 1 da matriz que representa T na base B são as soluções não nulas de

$$\begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 0 \Leftrightarrow \begin{cases} a = 0 \\ c = -b \end{cases}$$

Os vetores próprios de 2 são claramente os vetores não nulos da forma $(0, 0, c)$.

A multiplicidade geométrica do valor próprio 1 é apenas 1, o que significa que a forma canónica de Jordan de T tem apenas um bloco com 1 na diagonal. Logo, a menos de uma possível troca da ordem dos blocos,

$$J = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Sendo $B' = (v_1, v_2, v_3)$ a base com respeito à qual $A_{T,B',B'} = J$ podemos tomar $[v_1]_B = (0, 1, -1) \Leftrightarrow v_1 = t - t^2$. As coordenadas de v_2 têm de satisfazer a equação

$(A_{T,B,B} - I)[v_2]_B = [v_1]_B$ ou seja

$$\begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \Leftrightarrow \begin{cases} a = 1 \\ b + c = -2 \end{cases}$$

Uma possível solução é $[v_2]_B = (1, -2, 0) \Leftrightarrow v_2 = 1 - 2t$. Finalmente podemos tomar $[v_3]_B = (0, 0, 1) \Leftrightarrow v_3 = t^2$. Conclui-se que uma base que coloca T em forma canónica de Jordan é

$$B' = (t - t^2, 1 - 2t, t^2)$$

22. Enunciado A condição necessária e suficiente é que todos os valores próprios da matriz real que representa T com respeito a uma (qualquer) base B de V sejam reais (quando a matriz é encarada como uma matriz complexa).

    A necessidade da condição é evidente: dizer que T admite uma forma canónica de Jordan significa que existe uma base B para V com respeito à qual $A_{T,B,B}$ é representada por uma matriz diagonal por blocos com blocos de Jordan (reais) na diagonal. Se isso acontece, então claro que todos os valores próprios de $A_{T,B,B}$ (enquanto matriz complexa) são reais e se B' for qualquer outra base para V, os valores próprios de $A_{T,B',B'}$ (sendo os mesmos que os de $A_{T,B,B}$) serão também reais.

    Para ver a suficiência seja $A = A_{T,B,B}$ a representação matricial de T e $n = \dim(V)$. Enquanto matriz $n \times n$ complexa, A tem uma forma canónica de Jordan, isto é, existem uma matriz invertível $S \in M_{n \times n}(\mathbb{C})$ e J em forma de Jordan tais que $A = SJS^{-1}$. Por hipótese J é real porque estamos a assumir que todos os valores próprios de A são reais. Vejamos que podemos escolher a matriz S de forma a que seja também real.

    As colunas de S são vetores próprios generalizados, que se obtêm resolvendo as equações $(A - \lambda I)v_1 = 0$, $(A - \lambda I)v_2 = v_1$, etc..., com λ um valor próprio de A. Uma vez que A e λ são reais podemos encontrar soluções reais: de facto, o número de vetores próprios generalizados de λ que pertencem à base B' de $\mathbb{C}^n$ formada pelas colunas de S é $\dim N((A - \lambda I)^n)$ e este número é o mesmo quer calculemos o núcleo da matriz $(A - \lambda I)^n$ como matriz complexa quer o façamos como matriz real.

    Seja então $S \in M_{n \times n}(\mathbb{R})$ tal que $A = SJS^{-1}$ e $w_1, \ldots, w_n \in V$ os vetores cujas coordenadas na base B são as colunas de S (isto é $[w_i]_B$ é a i-ésima coluna de S). Então $B'' = (w_1, \ldots, w_n)$ é uma base ordenada para V com respeito à qual $A_{T,B'',B''} = J$ o que mostra que T admite uma forma canónica de Jordan.

23. Enunciado Uma vez que 1 é o único vetor próprio de T, a forma canónica de Jordan para T terá todas as entradas diagonais iguais a 1. Como a multiplicidade geométrica de 1 é 3 a forma canónica de Jordan terá 3 blocos. As possibilidades para as dimensões dos blocos (a menos de troca de ordem dos blocos) são: $5 + 1 + 1, 4 + 2 + 1$ e $3 + 2 + 2$ correspondendo às formas canónicas de Jordan

$$\begin{bmatrix} 1&1&0&0&0&0&0 \\ 0&1&1&0&0&0&0 \\ 0&0&1&1&0&0&0 \\ 0&0&0&1&1&0&0 \\ 0&0&0&0&1&0&0 \\ 0&0&0&0&0&1&0 \\ 0&0&0&0&0&0&1 \end{bmatrix}, \quad \begin{bmatrix} 1&1&0&0&0&0&0 \\ 0&1&1&0&0&0&0 \\ 0&0&1&1&0&0&0 \\ 0&0&0&1&0&0&0 \\ 0&0&0&0&1&1&0 \\ 0&0&0&0&0&1&0 \\ 0&0&0&0&0&0&1 \end{bmatrix}, \quad \begin{bmatrix} 1&1&0&0&0&0&0 \\ 0&1&1&0&0&0&0 \\ 0&0&1&0&0&0&0 \\ 0&0&0&1&1&0&0 \\ 0&0&0&0&1&0&0 \\ 0&0&0&0&0&1&1 \\ 0&0&0&0&0&0&1 \end{bmatrix}$$

Nas bases em que a representação matricial de T é dada pelas matrizes acima é fácil de calcular o núcleo de $(T-\mathrm{Id})^2$. Obtemos as seguintes dimensões para o núcleo: $4, 5$ e 6 respetivamente. São estas as possibilidades.

24. Enunciado O polinómio característico de A é

$$\begin{vmatrix} 1-\lambda & 2 & -1 \\ -1 & -1-\lambda & 0 \\ 0 & 1 & 2-\lambda \end{vmatrix} = (\lambda^2-1)(2-\lambda)+1+2(2-\lambda) = -\lambda^3+2\lambda^2-\lambda+3$$

Aplicando o algoritmo da divisão vemos que

$$p(x) = x^7-x^5+x^2-2x+1 = (-x^4-2x^3-2x^2-5x-14)(-x^3+2x^2-x+3)+30x^2-x+43$$

Pelo Teorema de Cayley Hamilton sabemos que $-A^3+2A^2-A+3I=0$ logo

$$\begin{aligned} p(A) &= (-A^4-2A^3-2A^2-5A-14I)\cdot 0+30A^2-A+43I \\ &= 30\begin{bmatrix} -1 & -1 & -3 \\ 0 & -1 & 1 \\ -1 & 1 & 4 \end{bmatrix} - \begin{bmatrix} 1 & 2 & -1 \\ -1 & -1 & 0 \\ 0 & 1 & 2 \end{bmatrix} + \begin{bmatrix} 43 & 0 & 0 \\ 0 & 43 & 0 \\ 0 & 0 & 43 \end{bmatrix} \\ &= \begin{bmatrix} 12 & -32 & -89 \\ 1 & 14 & 30 \\ -30 & 29 & 161 \end{bmatrix} \end{aligned}$$

25. Enunciado Vamos demonstrar o resultado por indução na dimensão de V. O caso em que $\dim V = 1$ é imediato. Assumimos como hipótese que o resultado é verdadeiro sempre que o espaço vetorial tem dimensão $n-1$ e tomamos para V um espaço vetorial de dimensão n.

Uma vez que V é um espaço vetorial complexo de dimensão finita, T tem um vetor próprio v. Seja W o espaço vetorial quociente $W = V/L(\{v\})$. Uma vez que $L(\{v\})$ é um subespaço invariante, T induz um endomorfismo

$$\overline{T}\colon W \to W$$

definido por $\overline{T}(v+L(\{v\})) = T(v)+L(\{v\})$. Dado que W tem dimensão $n-1$, por hipótese de indução existe uma base $B' = (v_2+L(\{v\}),\ldots,v_n+L(\{v\}))$ para W tal que $A_{\overline{T},B',B'}$ é triangular superior. Isto significa que existem escalares α_{ij} para $2 \leq j \leq i \leq n$

$$\overline{T}(v_i+L(\{v\})) = \alpha_{ii}v_i+\alpha_{i(i-1)}v_{i-1}+\ldots+\alpha_{i2}v_2+L(\{v\})$$

A equação anterior significa que existem escalares α_{i1} tais que

$$T(v_i) = \alpha_{ii}v_i+\alpha_{i(i-1)}v_{i-1}+\ldots+\alpha_{i2}v_2+\alpha_{i1}v$$

Mas então $B = (v, v_2, \ldots, v_n)$ é uma base de V tal que $A_{T,B,B}$ é triangular superior.

26. Enunciado
   (a) Temos que calcular as soluções da equação diferencial $(D-\lambda)^n y = 0$ uma vez que $y(t)$ é um vetor próprio generalizado de D se e só se é uma solução desta equação para algum natural n.

Sabemos que $(D-\lambda)y = 0 \Leftrightarrow y(t) = c_1 e^{\lambda t}$. Portanto para $n = 2$ obtemos

$$(D-\lambda)^2 y = 0 \Leftrightarrow (D-\lambda)((D-\lambda)y) = 0 \Leftrightarrow (D-\lambda)y = c_1 e^{\lambda t}$$

Esta equação pode ser resolvida usando a sugestão que começamos por verificar:

$$y' - \lambda y = f(t) \Leftrightarrow e^{-\lambda t}y' - \lambda e^{-\lambda t}y = e^{-\lambda t}f(t) \Leftrightarrow \frac{d}{dt}\left(e^{-\lambda t}y\right) = e^{-\lambda t}f(t)$$

Tomando agora $f(t) = c_1 e^{\lambda t}$ obtemos então

$$y' - \lambda y = c_1 e^{\lambda t} \Leftrightarrow \frac{d}{dt}\left(e^{-\lambda t}y(t)\right) = c_1 \Leftrightarrow e^{-\lambda t}y(t) = c_1 t + c_2 \Leftrightarrow y(t) = (c_1 t + c_2)e^{\lambda t}$$

Conclui-se assim que

$$(D-\lambda)^2 y = 0 \Leftrightarrow y(t) = (c_1 t + c_2)e^{\lambda t}$$

Vamos então mostrar por indução que

$$(D-\lambda)^n y = 0 \Leftrightarrow y(t) = c_0 e^{\lambda t} + c_1 t e^{\lambda t} + \ldots + c_{n-1}t^{n-1}e^{\lambda t}$$

o que implica o resultado pretendido. Já observámos que a base de indução é verificada. Assumindo que o resultado é válido para a equação $(D-\lambda)^{n-1}y = 0$ temos

$$(D-\lambda)^n y = 0 \Leftrightarrow (D-\lambda)^{n-1}\left((D-\lambda)y\right) = 0 \Leftrightarrow (D-\lambda)y = (c_1 + c_2 t + \ldots c_{n-1}t^{n-2})e^{\lambda t}$$

para algumas constantes $c_1, \ldots, c_{n-1}$. Usando a sugestão, a equação anterior é equivalente a

$$\frac{d}{dt}\left(e^{-\lambda t}y(t)\right) = c_1 + c_2 t + \ldots c_{n-1}t^{n-2} \Leftrightarrow e^{-\lambda t}y(t) = c_0 + c_1 t + c_2\frac{t^2}{2} + \ldots + c_{n-1}\frac{t^{n-1}}{n-1}$$

Uma vez que $c_0, c_1, \ldots c_{n-1}$ são constantes arbitrárias, isto conclui a demonstração do passo de indução.

(b) Vamos ver que e^{e^t} não se pode escrever como uma combinação linear de funções de algum $E(\lambda)$. De facto se $f(t) \in E(\lambda)$ então claramente

$$\lim_{t\to+\infty}\frac{f(t)}{e^{e^t}} = 0$$

Suponhamos por absurdo que $e^{e^t} = f_1(t) + \ldots + f_k(t)$ com $f_i(t) \in E(\lambda_i)$. Dividindo esta igualdade por e^{e^t} e tomando o limite quando $t \to +\infty$ obtemos a contradição $1 = 0$. Conclui-se que não é possível escrever e^{e^t} como uma soma (finita) de vetores próprios generalizados de D.

27. Enunciado

(a) Suponhamos que $N^k = 0$. Dado que $\lambda \operatorname{Id} + N = \lambda\left(\operatorname{Id} + \frac{1}{\lambda}N\right)$, para ver que este endomorfismo é invertível é suficiente ver que $\operatorname{Id} + \frac{1}{\lambda}N$ é invertível, sendo que $(\frac{1}{\lambda}N)^k = 0$. Mas pela sugestão temos

$$(\operatorname{Id} + \frac{1}{\lambda}N)\left(\operatorname{Id} - \frac{1}{\lambda}N + \frac{1}{\lambda^2}N^2 - \ldots + (-1)^{k-1}\frac{1}{\lambda^{k-1}}N^{k-1}\right) = \operatorname{Id} + \frac{(-1)^k}{\lambda^k}N^k = \operatorname{Id} + 0 = \operatorname{Id}$$

Conclui-se que $\operatorname{Id} - \frac{1}{\lambda}N + \frac{1}{\lambda^2}N^2 - \ldots + (-1)^{k-1}\frac{1}{\lambda^{k-1}}N^{k-1}$ é a inversa de $\operatorname{Id} + \frac{1}{\lambda}N$ e portanto que

$$\frac{1}{\lambda}\left(\operatorname{Id} - \frac{1}{\lambda}N + \frac{1}{\lambda^2}N^2 - \ldots + (-1)^{k-1}\frac{1}{\lambda^{k-1}}N^{k-1}\right) \tag{80}$$

é a inversa de $\lambda \operatorname{Id} + N$.

(b) Tendo em conta o exercício anterior, podemos tomar $V = L(\{e^t, te^t, \ldots, t^k e^t, \ldots\})$ e para T a restrição da operação de derivação D ao subespaço V. Todos os elementos de V são vetores próprios generalizados de 1 para T logo $V = E^g(1)$.
No entanto $T - \operatorname{Id} = D - \operatorname{Id}$ não é nilpotente em V: dado um natural n arbitrário temos $(D - \operatorname{Id})^n(t^n e^t) = n!e^t \neq 0$ pelo que nenhuma potência de $(D - \operatorname{Id})$ se anula em todo o espaço V.

(c) Podemos escrever $(T - \mu) = (\lambda - \mu) + (T - \lambda)$. Apesar de $(T - \lambda)$ não ser nilpotente em $E^g(\lambda)$ continua a ser verdade que para cada $v \in E^g(\lambda)$ individualmente $(T - \lambda)^k v = 0$ para algum k, portanto a expressão

$$\frac{1}{\lambda - \mu}\left(\operatorname{Id} - \frac{1}{\lambda - \mu}(T - \lambda) + \frac{1}{(\lambda - \mu)^2}(T - \lambda)^2 + \ldots\right) v$$

faz sentido para todo o $v \in E^g(\lambda)$: apesar de aparentemente haver infinitos termos na soma, eles são todos 0 a partir de certa ordem *que depende de v*.
É agora imediato verificar que a função definida em $E^g(\lambda)$ pela expressão $S(v) = \frac{1}{\lambda-\mu}\left(\operatorname{Id} - \frac{1}{\lambda-\mu}(T - \lambda) + \frac{1}{(\lambda-\mu)^2}(T - \lambda)^2 + \ldots\right) v$ (que toma valores em $E^g(\lambda)$ porque $E^g(\lambda)$ é um subespaço invariante por T) é linear e satisfaz $S(T - \mu) = (T - \mu)S = \operatorname{Id}$ pelo que a restrição de $T - \mu$ a $E^g(\lambda)$ é um isomorfismo.
A demonstração que a soma dos espaços próprios generalizados é direta (parte da demonstração do Teorema 6.37) só utilizou o facto de a restrição dos endomorfismos $(T - \mu)$ a $E^g(\lambda)$ serem isomorfismos quando $\lambda \neq \nu$ e portanto permanece válida.

28. Enunciado

(a) Para $i < k-1$ o efeito de T no $(i-1)$-ésimo vetor da base escolhida é $T(T^i v) = T^{i+1} v$ que é ainda um vetor da base de W (é o vetor seguinte). O efeito no k-ésimo vetor é

$$T(T^{k-1}v) = T^k v = -a_0 v - \ldots - a_{k-1}T^{k-1}v$$

logo a matriz que representa T com respeito à base $(v, Tv, \ldots T^{k-1}v)$ é

$$\begin{bmatrix} 0 & 0 & \cdots & & 0 & -a_0 \\ 1 & 0 & \cdots & & 0 & -a_1 \\ 0 & 1 & 0 & & \vdots & \vdots \\ 0 & 0 & \ddots & \ddots & & \\ \vdots & & \ddots & & & \\ 0 & \cdots & & 0 & 1 & -a_{k-1} \end{bmatrix}$$

(b) Temos de calcular o determinante

$$\begin{vmatrix} -\lambda & 0 & \cdots & & 0 & -a_0 \\ 1 & -\lambda & \cdots & & 0 & -a_1 \\ 0 & 1 & -\lambda & & \vdots & \vdots \\ 0 & 0 & \ddots & \ddots & & \\ \vdots & & \ddots & & & \\ 0 & \cdots & & 0 & 1 & -a_{k-1}-\lambda \end{vmatrix}$$

o que fazemos por indução na dimensão da matriz. Se $k=1$ a matriz é simplesmente $\begin{bmatrix}-a_0\end{bmatrix}$ pelo que o resultado está correto. Assumindo que o resultado está correto para matrizes companheiras de dimensão $k-1$ e aplicando a regra de Laplace ao longo da primeira linha obtemos a seguinte fórmula para o determinante

$$-\lambda \begin{vmatrix} -\lambda & \cdots & & 0 & -a_1 \\ 1 & -\lambda & & \vdots & \vdots \\ 0 & \ddots & \ddots & & \\ \vdots & \ddots & & & \\ 0 & \cdots & 0 & 1 & -a_{k-1}-\lambda \end{vmatrix} + (-1)^{k+1}(-a_0) \begin{vmatrix} 1 & -\lambda & \cdots & & 0 \\ 0 & 1 & -\lambda & & \vdots \\ 0 & 0 & \ddots & \ddots & \\ \vdots & & \ddots & & \\ 0 & \cdots & & 0 & 1 \end{vmatrix}$$

que, por hipótese de indução, é

$$(-\lambda)((-1)^{k-1}(a_1 + a_2\lambda + \ldots + a_{k-1}\lambda^{k-2} + \lambda^{k-1}) + (-1)^k a_0 = (-1)^k p(\lambda)$$

conforme pretendido.

(c) Se v é um vetor próprio com valor próprio λ, o subespaço cíclico gerado por v é $L(\{v\})$ e $Tv = \lambda v$ logo a matriz companheira é $\begin{bmatrix}\lambda\end{bmatrix}$.

(d) Temos $T(v) = (2,1,1)$ e $T^2(v) = T(2,1,1) = (6,2,3)$. Os vetores $\{v, T(v), T^2(v)\}$ são uma base de $\mathbb{R}^3$ pelo que o espaço cíclico gerado por v é todo o $\mathbb{R}^3$. Temos $T^3(v) = T(6,2,3) = (17,7,9)$ e resolvendo o sistema $T^3(v) = av + bT(v) + cT^2(v)$ obtemos

$$\begin{bmatrix} 1 & 2 & 6 & | & 17 \\ 0 & 1 & 2 & | & 7 \\ 0 & 1 & 3 & | & 9 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 6 & | & 17 \\ 0 & 1 & 2 & | & 7 \\ 0 & 0 & 1 & | & 2 \end{bmatrix}$$

pelo que $c = 2$, $b = 7 - 2c = 3$ e $a = 17 - 2b - 6c = -1$. Ou seja $T^3(v) = -v + 3T(v) + 2T^2(v)$ e a matriz companheira (que representa T na base (v, Tv, T^2v) é

$$\begin{bmatrix} 0 & 0 & -1 \\ 1 & 0 & 3 \\ 0 & 1 & 2 \end{bmatrix}$$

(e) Notamos primeiro que se $\{v, T(v), \ldots, T^{k-1}(v)\}$ é linearmente independente, o mesmo será verdade para o conjunto $\{v, p_1(T)(v), \ldots, p_{k-1}(T)(v)\}$ se o polinómio p_j tiver grau j (a matriz de mudança de coordenadas de uma base para a outra é triangular com entradas não nulas na diagonal).

Consequentemente se $p(x) = q(x)r(x)$, o grau de $q(x)$ for m e o de $r(x)$ for ℓ (com $m + \ell = k$) então

$$B = (v, T(v), T^2(v), \ldots, T^{m-1}(v), q(T)(v), Tq(T)(v), \ldots, T^{\ell-1}q(T)(v))$$

é uma base para o espaço cíclico determinado por v e é imediato verificar que $A_{T,B,B}$ é a matriz indicada: note-se que se $r(x) = b_0 + b_1 x + \ldots + x^\ell$ então dado que $x^\ell q(x) = r(x)q(x) - b_{\ell-1}x^{\ell-1}q(x) - \ldots - b_1 x q(x) - b_0 q(x)$ temos

$$\begin{aligned} T(T^{\ell-1}q(T)(v)) &= T^\ell q(T)(v) \\ &= p(T)(v) - b_{\ell-1}T^{\ell-1}q(T)(v) - \ldots - b_1 T q(T)(v) - b_0 q(T)(v) \\ &= -b_{\ell-1}T^{\ell-1}q(T)(v) - \ldots - b_1 T q(T)(v) - b_0 q(T)(v) \end{aligned}$$

(f) Seja J um bloco de Jordan de tamanho k com λ na diagonal e $T\colon \mathbb{R}^k \to \mathbb{R}^k$ a transformação definida por $T(x) = Jx$. Seja $(v_1, \ldots, v_k)$ a base canónica de $\mathbb{R}^k$. Então dado que $(T - \lambda)^i(v_k) = v_{k-i}$ para $i = 1, \ldots, k-1$, vemos que $\{v_k, (T - \lambda)(v_k), \ldots, (T-\lambda)^{k-1}(v_k)\}$ é linearmente independente, e portanto $\{v_k, T(v_k), \ldots, T^{k-1}(v_k)\}$ é linearmente independente (conforme a alínea anterior). O espaço cíclico gerado por v_k é portanto todo o $\mathbb{R}^k$. Uma vez que $(T - \lambda)^k(v_k) = 0$ a matriz companheira determinada pelo vetor v_k é a do polinómio $p(x) = (x - \lambda)^k = \sum_{i=0}^k (-1)^i \binom{k}{i} \lambda^i x^{k-i}$:

$$\begin{bmatrix} 0 & 0 & \cdots & 0 & (-1)^{k+1}\lambda^k \\ 1 & 0 & & 0 & (-1)^k \lambda^{k-1} \\ 0 & 1 & 0 & 0 & (-1)^{k-1}\binom{k}{2}\lambda^{k-2} \\ \vdots & & \ddots & 0 & \vdots \\ 0 & \cdots & & 1 & k\lambda \end{bmatrix}$$

esta matriz representa a transformação T com respeito à base $(v_k, \ldots, T^{k-1}(v_k))$ de $\mathbb{R}^k$ e é portanto semelhante à matriz J que representa T na base canónica.

29. Enunciado

(a) Dado $p(x) \in I_v$ podemos usar o algoritmo de divisão para escrever $p(x) = q(x)m(x) + r(x)$ tendo $r(x)$ grau menor que $m(x)$. Uma vez que

$$0 = p(T)(v) = q(T)m(T)(v) + r(T)(v) = q(T)(0) + r(T)(v) = 0 + r(T)(v) = r(T)(v)$$

vemos que $r(x) \in I_v$. Como escolhemos $m(x)$ com o menor grau possível, isto só pode acontecer se $r(x) = 0$, isto é se $p(x)$ é um múltiplo de $m(x)$. Reciprocamente é evidente que todos os múltiplos de $m(x)$ estão em I_v.

(b) Pelo Teorema de Cayley-Hamilton o polinómio característico de T pertence a I_v logo pela alínea anterior, $m_v(x)$ divide o polinómio característico de T.

(c) O polinómio mínimo de uma transformação linear pode ser calculado usando a forma canónica de Jordan da seguinte forma:
O polinómio mínimo de um bloco de Jordan de tamanho k com λ na diagonal coincide com o polinómio característico e é dado por $p(x) = (\lambda - x)^k$ (ver a alínea (f) do exercício anterior - é o polinómio mínimo do elemento da base correspondente à última coluna do bloco).

O polinómio característico de uma matriz é o produto dos polinómios característicos de cada bloco na forma canónica de Jordan. No entanto se k for o maior tamanho de um bloco com λ na diagonal, o polinómio $(T-\lambda)^k$ aniquilará todos os vetores próprios generalizados de λ. Portanto se $\lambda_1, \ldots, \lambda_k$ forem os valores próprios distintos de T e $m_1, \ldots, m_k$ os tamanhos dos maiores blocos correspondentes a estes vetores próprios, o polinómio $(T-\lambda_1)^{m_1} \cdots (T-\lambda_k)^{m_k}$ anular-se-á.
Por outro lado nenhum fator próprio deste polinómio aniquilará a transformação T (se faltar algum fator $(T-\lambda_i)$ então o polinómio não aniquilará o vetor da base correspondente à última coluna de um bloco com λ_i na diagonal e comprimento máximo). Conclui-se que $(T-\lambda_1)^{m_1} \cdots (T-\lambda_k)^{m_k}$ é o polinómio mínimo de T e este coincide com o polinómio característico se e só se existe exatamente um bloco de Jordan para cada valor próprio.

(d) Cada matriz é semelhante à sua forma canónica de Jordan e duas formas de Jordan com uma estrutura de blocos diferentes não são semelhantes pelo que temos apenas que encontrar as formas canónicas de Jordan que têm polinómio mínimo $(x-1)^2$. Pela alínea anterior estas são as formas canónicas de Jordan que têm 1 em todas as entradas diagonais e cujos blocos de Jordan têm comprimento máximo 2. As três possibilidades (note-se que matrizes que se obtenham por troca de ordem dos blocos são semelhantes) são

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

(e) Vimos na alínea (c) que o polinómio mínimo é $(x-\lambda_1)^{m_1} \cdots (x-\lambda_k)^{m_k}$ onde λ_i são os valores próprios distintos de T e m_i é o tamanho do maior bloco correspondente a λ_i. Seja v_i o vetor da base correspondente à última coluna de um dos maiores blocos com λ_i na diagonal e $v = v_1 + \ldots + v_k$. Dado um polinómio $p(x)$, $P(T)v_i$ pertence ao espaço cíclico gerado por v_i. A soma dos espaços cíclicos gerados pelos v_i é direta (pois os espaços cíclicos estão contidos em espaços próprios generalizados correspondentes a valores próprios distintos) logo $p(T)v = 0$ se e só se $p(T)v_i = 0$ para todo o i, o que acontece se e só se $(x-\lambda_i)^{m_i}$ divide $p(x)$. Conclui-se que I_v consiste nos múltiplos do polinómio mínimo de T, isto é que o polinómio mínimo de T é $m_v(x)$.

30. Enunciado Nesta resolução vamos escrever $m_a(\lambda)$ e $m_g(\lambda)$ respetivamente para as multiplicidades algébrica e geométrica do valor próprio λ e convencionamos que as entradas abaixo da diagonal na forma canónica de Jordan real são escolhidas positivas (cf. Observação 6.76). Notamos que $m_g(1-i) = m_g(1+i) = 2$ e que $m_a(1+i) = m_a(1-i) \geq m_g(1+i) = 2$. Além disso $m_a(2) \geq m_g(2) \geq 1$ e

$$m_a(2) + 2m_a(1+i) = 8.$$

Conclui-se que ou $m_a(1+i) = m_a(1-i) = 2$ e então $m_a(2) = 4$, ou $m_a(1+i) = m_a(1-i) = 3$ e então $m_a(2) = 2$. No primeiro caso os espaços próprios generalizados de $1 \pm i$ coincidem com os espaços próprios e portanto a forma canónica de Jordan real toma (a menos de troca da ordem dos blocos) uma das seguintes formas dependendo da ação de A no espaço próprio generalizado de $\lambda = 2$:

$$\begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix}$$

No segundo caso, na forma canónica de Jordan complexa cada um dos valores próprios $1 \pm i$ tem dois blocos, um de comprimento 2 e outro de comprimento 1. Há portanto apenas duas possibilidades consoante o espaço próprio de 2 coincide ou não com o respetivo espaço próprio generalizado:

$$\begin{bmatrix} 1 & -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix} \quad \text{ou} \quad \begin{bmatrix} 1 & -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix}$$

31. Enunciado

(a) O polinómio característico da matriz é

$$\begin{vmatrix} 1-\lambda & -1 & 1 \\ 2 & 1-\lambda & -1 \\ 1 & 1 & 1-\lambda \end{vmatrix} = -\lambda^3 + 3\lambda^2 - 5\lambda + 6$$

Vemos que $\lambda = 2$ é uma raíz, donde obtemos a fatorização

$$-\lambda^3 + 3\lambda^2 - 5\lambda + 6 = -(\lambda - 2)(\lambda^2 - \lambda + 3)$$

Os valores próprios são portanto 2 e $\frac{1 \pm i\sqrt{11}}{2}$ todos com multiplicidade algébrica 1. Conclui-se que a forma canónica de Jordan real é (a menos de troca de ordem de blocos e do sinal das entradas fora da diagonal):

$$J = \begin{bmatrix} 2 & 0 & 0 \\ 0 & \frac{1}{2} & -\frac{\sqrt{11}}{2} \\ 0 & \frac{\sqrt{11}}{2} & \frac{1}{2} \end{bmatrix} \tag{81}$$

Os vetores próprios de 2 são as soluções não nulas de

$$\begin{bmatrix} -1 & -1 & 1 \\ 2 & -1 & -1 \\ 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 0 \Leftrightarrow \begin{cases} c = a + b \\ c = 2a - b \end{cases} \Leftrightarrow \begin{cases} a = 2b \\ c = 3b \end{cases}$$

ou seja, os múltiplos de $(2, 1, 3)$. Os vetores próprios de $\frac{1+i\sqrt{11}}{2}$ são as soluções não nulas de

$$\begin{bmatrix} \frac{1-i\sqrt{11}}{2} & -1 & 1 \\ 2 & \frac{1-i\sqrt{11}}{2} & -1 \\ 1 & 1 & \frac{1-i\sqrt{11}}{2} \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 0 \Leftrightarrow \begin{cases} c = \frac{-1+i\sqrt{11}}{2} a + b \\ c = 2a + \frac{1-i\sqrt{11}}{2} b \end{cases} \Leftrightarrow \begin{cases} b = -\frac{1+i\sqrt{11}}{2} a \\ c = -a \end{cases}$$

ou seja, os múltiplos de $(1, -\frac{1+i\sqrt{11}}{2}, -1) = (1, -\frac{1}{2}, -1) + i(0, -\frac{\sqrt{11}}{2}, 0)$. De acordo com a demonstração do Teorema 6.75 uma base que coloca a matriz dada na forma (81) é

$$((2,1,3),(1,-\tfrac{1}{2},-1),(0,\tfrac{\sqrt{11}}{2},0))$$

(b) O polinómio característico da matriz dada é

$$\begin{vmatrix} 1 & 2 & 1 & 1 \\ -1 & 3 & 0 & 0 \\ 0 & 0 & 2 & -1 \\ 0 & 0 & 1 & 2 \end{vmatrix} = ((1-\lambda)(3-\lambda)+2)((2-\lambda)^2+1) = ((\lambda-2)^2+1)^2$$

logo os valores próprios complexos são $2 \pm i$ ambos com multiplicidade algébrica 2. Os vetores próprios de $2+i$ são as soluções não nulas de

$$\begin{bmatrix} -1-i & 2 & 1 & 1 \\ -1 & 1-i & 0 & 0 \\ 0 & 0 & -i & -1 \\ 0 & 0 & 1 & -i \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = 0 \Leftrightarrow \begin{cases} -(1+i)(1-i)b + 2b + c - ic = 0 \\ a = (1-i)b \\ d = -ic \end{cases} \Leftrightarrow \begin{cases} c = 0 \\ a = (1-i)b \\ d = 0 \end{cases}$$

ou seja, os múltiplos não nulos de $(1-i,1,0,0) = (1,1,0,0) + i(-1,0,0,0)$. Uma vez que a multiplicidade geométrica de $2+i$ é apenas um vemos que a forma canónica de Jordan real toma a forma (novamente escolhendo as entradas abaixo da diagonal positivas cf. Observação 6.76)

$$J = \begin{bmatrix} 2 & -1 & 1 & 0 \\ 1 & 2 & 0 & 1 \\ 0 & 0 & 2 & -1 \\ 0 & 0 & 1 & 2 \end{bmatrix} \tag{82}$$

Um vetor próprio generalizado de $2+i$ é uma solução de

$$\begin{bmatrix} -1-i & 2 & 1 & 1 \\ -1 & 1-i & 0 & 0 \\ 0 & 0 & -i & -1 \\ 0 & 0 & 1 & -i \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} 1-i \\ 1 \\ 0 \\ 0 \end{bmatrix} \Leftrightarrow \begin{cases} -2b + (1+i) + 2b + (1+i)d = 1-i \\ a = (1-i)b - 1 \\ c = id \end{cases}$$

$$\Leftrightarrow \begin{cases} d = -1-i \\ a = (1-i)b - 1 \\ c = 1-i \end{cases}$$

Escolhendo por exemplo $(-1,0,1-i,-1-i) = (-1,0,1,-1) + i(0,0,-1,-1)$ vemos (de acordo com a demonstração do Teorema 6.75) que podemos escolher para base que coloca a matriz na forma (82)

$$((1,1,0,0),(1,0,0,0),(-1,0,1,-1),(0,0,1,1))$$

8.7. **Espaços vetoriais com produto interno.**

1. Enunciado
   (a) Esta função é bilinear mas não simétrica: por exemplo $\langle(1,0),(0,1)\rangle = 2 \neq 0 = \langle(0,1),(1,0)\rangle$, logo não é um produto interno.
   (b) Esta função é bilinear (é claramente uma função linear quando fixamos (x,y) ou (u,v)) e é simétrica visto que $\langle(u,v),(x,y)\rangle = ux+2uy+2vx+vy = \langle(x,y),(u,v)\rangle$. Temos
$$\langle(x,y),(x,y)\rangle = x^2+2xy+2yx+y^2 = (x+y)^2+2xy$$
   Esta função não é sempre positiva. Por exemplo tomando $(x,y)=(1,-1)$ obtemos -2 logo o axioma de positividade não se verifica. Conclui-se que esta função não é um produto interno.
   (c) A função é bilinear (a expressão é linear nas restantes variáveis quando fixamos x e y ou u e v) e é simétrica:
$$\langle(u,v),(x,y)\rangle = 2ux-uy-vx+2vy = \langle(x,y),(u,v)\rangle$$
   Por outro lado
$$\begin{aligned}\langle(x,y),(x,y)\rangle &= 2x^2-xy-yx+2y^2 = x^2+(x^2-2xy+y^2)+y^2\\ &= x^2+(x-y)^2+y^2 \geq 0 \text{ para todos os } (x,y)\end{aligned}$$
   e claramente a expressão acima só se anula se $(x,y)=(0,0)$. A função satisfaz portanto a condição de positividade e portanto é um produto interno.
   (d) Esta função não é linear nas variáveis (x,y) (nem nas (u,v) na realidade) e portanto não é um produto interno.
   (e) Esta função é bilinear (porque o produto de matrizes é bilinear e o traço é linear):
$$\begin{aligned}\langle A,\alpha B_1+\beta B_2\rangle &= \operatorname{tr}(A(\alpha B_1+\beta B_2)) = \operatorname{tr}(\alpha(AB_1)+\beta(AB_2))\\ &= \alpha\operatorname{tr}(AB_1)+\beta\operatorname{tr}(AB_2) = \alpha\langle A,B_1\rangle+\beta\langle A,B_2\rangle\end{aligned}$$
   A linearidade na primeira variável é análoga (e segue também da simetria).
   Esta função também é simétrica uma vez que $\operatorname{tr}(AB)=\operatorname{tr}(BA)$ (ver Exercício 5.10). No entanto, a não ser no caso trivial em que $n=1$, esta função não satisfaz positividade. Por exemplo,
$$\operatorname{tr}\left(\begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}\begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}\right) = \operatorname{tr}\left(\begin{bmatrix}-1 & 0\\ 0 & -1\end{bmatrix}\right) = -2$$
   Para qualquer $n \geq 2$, a matriz A que é diagonal por blocos com primeiro bloco dado pela matriz $\left[\begin{smallmatrix}0 & 1\\ -1 & 0\end{smallmatrix}\right]$ e que tem todas as restantes entradas nulas dá um contra-exemplo para a positividade. Conclui-se que esta função não é um produto interno a não ser quando $n=1$.
2. Enunciado
   (a) $\langle(1,0,2),(1,-2,1)\rangle = 2.1-0.1+2.0.(-2)-1.(-2)+2.1 = 6$.
   (b) $\|(1,1,1)\| = \sqrt{\langle(1,1,1),(1,1,1)\rangle} = \sqrt{2-1+2-1+1} = \sqrt{3}$.

(c) A distância de $(0,0,1)$ a $(1,0,0)$ é dada por

$$\|(0,0,1)-(1,0,0)\| = \|(-1,0,1)\| = \sqrt{3}.$$

(d) Os vetores ortogonais a $(1,2,3)$ verificam

$$\langle (1,2,3),(u,v,w)\rangle = 0 \Longleftrightarrow 2u-2u+4v-v+3w=0 \Longleftrightarrow v+w=0,$$

logo o conjunto dos vetores ortogonais a $(1,2,3)$ é

$$\{(u,v,w)\in\mathbb{R}^3 : \ v=-w\}.$$

(e) Como $\langle (1,1,1),(2,2,2)\rangle = 6$ o conjunto S não é ortogonal. Uma vez que os produtos internos entre vetores distintos de X são todos iguais a 0, o conjunto é ortogonal. Mas não é ortonormado pois, por exemplo, $\|(1,2,1)\| = \sqrt{7}\neq 1$.

3. Enunciado
   (a) $\langle (1,2+i,1-i),(i,1-3i,i)\rangle = 1.i+(2-i)(1-3i)+(1+i)i = -2-5i.$
   (b) $\|(i+2,3i,-1)\| = \sqrt{|i+2|^2+|3i|^2+1} = \sqrt{15}.$
   (c) A distância de $(1,2+i,i)$ a $(1,3+2i,2)$ é dada por

$$\|(1,2+i,i)-(1,3+2i,2)\| = \|(0,-1-i,i-2)\| = \sqrt{0+2+5} = \sqrt{7}.$$

   (d) O plano perpendicular a $(i+2,2,3-i)$ tem equação

$$\langle (i+2,2,3-i),(z_1,z_2,z_3)\rangle = 0 \Longleftrightarrow (2-i)z_1+2z_2+(3+i)z_3=0.$$

4. Enunciado
   (a) Para provar que a função é um produto interno temos de verificar a bilinearidade, simetria e positividade. A simetria é óbvia e tendo em conta esta propriedade é suficiente verificar a linearidade num dos argumentos: se $\alpha_1,\alpha_2\in\mathbb{R}$ e $p_1,p_2,q\in V$ então

$$\begin{aligned}\langle \alpha_1p_1+\alpha_2p_2,q\rangle = \ & (\alpha_1p_1+\alpha_2p_2)(-1)q(-1)+(\alpha_1p_1+\alpha_2p_2)(0)q(0)\\ &+(\alpha_1p_1+\alpha_2p_2)(1)q(1) = \alpha_1p_1(-1)q(-1)+\alpha_1p_1(0)q(0)+\alpha_1p_1(1)q(1)\\ &+\alpha_2p_2(-1)q(-1)+\alpha_2p_2(0)q(0)+\alpha_2p_2(1)q(1) = \alpha_1\langle p_1,q\rangle+\alpha_2\langle p_2,q\rangle.\end{aligned}$$

Para verificar a positividade notamos que

$$\langle p,p\rangle = (p(-1))^2+(p(0))^2+(p(1))^2 \geq 0$$

e

$$\langle p,p\rangle = 0 \Leftrightarrow p(-1)=p(0)=p(1)=0.$$

Se $p(t)=a+bt+ct^2$ com $a,b,c\in\mathbb{R}$ então estas condições são equivalentes a

$$\begin{cases} a-b+c=0\\ a=0\\ a+b+c=0\end{cases}$$

e portanto $a=b=c=0$, ou seja, $p=0$.

(b) Consideramos os polinómios $p_1 = t(t+1), p_2 = 1-t^2$ e $p_3 = t(t-1)$. O polinómio p_1 tem zeros em $t = 0$ e $t = -1$, p_2 tem zeros em 1 e -1 e p_3 em 0 e 1, logo o produto interno de dois destes polinómios é igual a 0 (desde que não sejam o mesmo, claro). Por outro lado é claro que estes 3 polinómios formam uma base de V, portanto o conjunto $\{p_1, p_2, p_3\}$ é uma base ortogonal de V.

(c) A entrada ij da matriz da métrica G_B com respeito à base B é dada por $\langle v_i, v_j\rangle$, onde $1 \leq i, j \leq 3$ e os vetores v_i são os vetores da base B. Por exemplo, $\langle v_1, v_3\rangle = \langle 1, t^2\rangle = 1 + 0 + 1 = 2$. Sabendo que G_B é simétrica e calculando os restantes produtos internos obtemos

$$G_B = \begin{bmatrix} 3 & 0 & 2 \\ 0 & 2 & 0 \\ 2 & 0 & 2 \end{bmatrix}.$$

5. Enunciado

(a) Usando a definição de produto interno neste espaço vetorial obtemos

$$\langle x, 2+x^2\rangle = \int_0^\pi x(2+x^2)\,dx = \left(x^2 + \frac{x^4}{4}\right)\Bigg|_0^\pi = \pi^2 + \frac{\pi^4}{4}.$$

(b) Consideramos a base $B = \{1, x, x^2\} = \{v_1, v_2, v_3\}$ para o subespaço e aplicamos o processo de ortogonalização de Gram-Schmidt a B. Definimos $w_1 = v_1 = 1$. Temos

$$\|w_1\|^2 = \int_0^\pi 1\ dx = \pi.$$

A seguir definimos

$$w_2 = v_2 - \operatorname{proj}_{w_1}(v_2) = v_2 - \langle w_1, v_2\rangle \frac{w_1}{\|w_1\|^2}$$

Uma vez que

$$\langle w_1, v_2\rangle = \langle 1, x\rangle = \int_0^\pi x\,dx = \frac{x^2}{2}\Bigg|_0^\pi = \frac{\pi^2}{2}$$

obtemos

$$w_2 = x - \frac{\pi^2}{2}\frac{1}{\pi} = x - \frac{\pi}{2}$$

e

$$\|w_2\|^2 = \int_0^\pi \left(x - \frac{\pi}{2}\right)^2 dx = \frac{1}{3}\left(x - \frac{\pi}{2}\right)^3\Bigg|_0^\pi = \frac{2}{3}\left(\frac{\pi}{2}\right)^3 = \frac{\pi^3}{12}.$$

O terceiro vetor ortogonal é

$$w_3 = v_3 - \operatorname{proj}_{w_1}(v_3) - \operatorname{proj}_{w_2}(v_3) = v_3 - \langle w_1, v_3\rangle \frac{w_1}{\|w_1\|^2} - \langle w_2, v_3\rangle \frac{w_2}{\|w_2\|^2}.$$

Calculando os produtos internos

$$\langle w_1, v_3\rangle = \langle 1, x^2\rangle = \int_0^\pi x^2\,dx = \frac{x^3}{3}\Bigg|_0^\pi = \frac{\pi^3}{3} \quad \text{e}$$

$$\langle w_2, v_3\rangle = \langle x - \frac{\pi}{2}, x^2\rangle = \int_0^\pi \left(x - \frac{\pi}{2}\right) x^2\, dx = \left(\frac{x^4}{4} - \frac{\pi}{2}\frac{x^3}{3}\right)\Big|_0^\pi = \frac{\pi^4}{4} - \frac{\pi^4}{6} = \frac{\pi^4}{12},$$

obtemos

$$w_3 = x^2 - \frac{\pi^3}{3}\frac{1}{\pi} - \frac{\pi^4}{12}\left(x - \frac{\pi}{2}\right)\frac{12}{\pi^3} = x^2 - \pi x + \frac{\pi^2}{6} \quad \text{e}$$

$$\|w_3\|^2 = \int_0^\pi \left(x^2 - \pi x + \frac{\pi^2}{6}\right)^2 dx = \left(\frac{x^5}{5} - 2\pi\frac{x^4}{4} + \frac{4\pi^2}{3}\frac{x^3}{3} - \frac{\pi^3}{3}\frac{x^2}{2} + \frac{\pi^4}{36}x\right)\Big|_0^\pi = \frac{\pi^5}{180}.$$

Finalmente obtemos uma base ortonormada

$$\left\{\frac{w_1}{\|w_1\|}, \frac{w_2}{\|w_2\|}, \frac{w_3}{\|w_3\|}\right\} = \left\{\frac{1}{\sqrt{\pi}}, \frac{2\sqrt{3}}{\pi\sqrt{\pi}}\left(x - \frac{\pi}{2}\right), \frac{6\sqrt{5}}{\pi^2\sqrt{\pi}}\left(x^2 - \pi x + \frac{\pi^2}{6}\right)\right\}.$$

(c) Começamos por mostrar que elementos distintos de S_1 são ortogonais

$$\begin{aligned}\langle \cos(n_1 x), \cos(n_2 x)\rangle &= \int_0^\pi \cos(n_1 x)\cos(n_2 x)\, dx = \\ &= \int_0^\pi \frac{1}{2}\cos((n_1 - n_2)x) + \frac{1}{2}\cos((n_1 + n_2)x)\, dx = \\ &= \frac{\operatorname{sen}((n_1 - n_2)x)}{2(n_1 - n_2)} + \frac{\operatorname{sen}((n_1 + n_2)x)}{2(n_1 + n_2)}\Big|_0^\pi = 0.\end{aligned}$$

Para os elementos de S_2 os cálculos são análogos

$$\begin{aligned}\langle \operatorname{sen}(n_1 x), \operatorname{sen}(n_2 x)\rangle &= \int_0^\pi \operatorname{sen}(n_1 x)\operatorname{sen}(n_2 x)\, dx = \\ &= \int_0^\pi \frac{1}{2}\cos((n_1 - n_2)x) - \frac{1}{2}\cos((n_1 + n_2)x)\, dx = \\ &= \frac{\operatorname{sen}((n_1 - n_2)x)}{2(n_1 - n_2)} - \frac{\operatorname{sen}((n_1 + n_2)x)}{2(n_1 + n_2)}\Big|_0^\pi = 0.\end{aligned}$$

Para as normas obtemos, se $n > 0$,

$$\|\cos(nx)\|^2 = \int_0^\pi \cos^2(nx)\, dx = \int_0^\pi \frac{1 + \cos(2nx)}{2}\, dx = \left(\frac{x}{2} + \frac{\operatorname{sen}(2nx)}{4n}\right)\Big|_0^\pi = \frac{\pi}{2}$$

e se $n = 0$ temos $\|\cos(nx)\| = \|1\| = \sqrt{\pi}$. Para os elementos do conjunto S_2 obtemos

$$\|\operatorname{sen}(nx)\|^2 = \int_0^\pi \operatorname{sen}^2(nx)\, dx = \int_0^\pi \frac{1 - \cos(2nx)}{2}\, dx = \left(\frac{x}{2} - \frac{\operatorname{sen}(2nx)}{4n}\right)\Big|_0^\pi = \frac{\pi}{2}.$$

6. Enunciado Se $B = \{(1,2), (-1,1)\}$ é uma base ortonormada então a matriz da métrica nesta base é $G_B = I$. Para obter a expressão do produto interno podemos calcular a matriz da métrica com respeito à base canónica, que é dada pela fórmula

$$G_{B_{can}} = S^T_{B_{can}\to B} G_B S_{B_{can}\to B}$$

Temos

$$S_{B_{can}\to B} = S^{-1}_{B\to B_{can}} = \begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix}^{-1} = \tfrac{1}{3}\begin{bmatrix} 1 & 1 \\ -2 & 1 \end{bmatrix}$$

logo

$$G_{B_{can}} = \tfrac{1}{3}\begin{bmatrix} 1 & 1 \\ -2 & 1 \end{bmatrix}^T \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tfrac{1}{3}\begin{bmatrix} 1 & 1 \\ -2 & 1 \end{bmatrix} = \tfrac{1}{9}\begin{bmatrix} 1 & -2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ -2 & 1 \end{bmatrix} = \tfrac{1}{9}\begin{bmatrix} 5 & -1 \\ -1 & 2 \end{bmatrix}$$

Portanto o produto interno pretendido é dado por

$$\langle (x,y),(u,v)\rangle = \tfrac{1}{9}\begin{bmatrix} x & y \end{bmatrix}\begin{bmatrix} 5 & -1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} u \\ v \end{bmatrix} = \tfrac{1}{9}\left(5xu - xv - yu + 2yv\right).$$

7. Enunciado Sejam $v_1,\dots,v_n$ vetores distintos de um conjunto ortogonal e $\alpha_1,\dots,\alpha_n$ escalares. Supondo $\alpha_1 v_1 + \ldots + \alpha_n v_n = 0$ queremos mostrar que $\alpha_k = 0$ para todo k, com $1 \leq k \leq n$. Temos

$$\begin{aligned} 0 = \langle 0, v_k\rangle &= \langle \alpha_1 v_1 + \ldots + \alpha_n v_n, v_k\rangle = \alpha_1\langle v_1, v_k\rangle + \ldots + \alpha_k\langle v_k, v_k\rangle + \ldots + \alpha_n\langle v_n, v_k\rangle \\ &= \alpha_k\langle v_k, v_k\rangle = \alpha_k\|v_k\|^2, \end{aligned}$$

onde usámos a linearidade do produto interno na terceira igualdade, e na igualdade seguinte usámos o facto do conjunto ser ortogonal, ou seja $\langle v_i, v_k\rangle = 0$ para $i \neq k$. Uma vez que o conjunto não contém 0, a norma $\|v_k\| \neq 0$ e portanto $\alpha_k = 0$ para todo o $1 \leq k \leq n$.

8. Enunciado
   (a) Uma base para V pode ser $B = ((1,0,i,0),(0,1,i,0),(0,0,0,1))$. É claro que este conjunto é linearmente independente e gera V.
   (b) Calculando o produto interno entre os vetores da base obtemos

$$\begin{aligned} \langle (1,0,i,0),(1,0,i,0)\rangle &= \langle (0,1,i,0),(0,1,i,0)\rangle = 2 \\ \langle (1,0,i,0),(0,1,i,0)\rangle &= \langle (0,0,0,1),(0,0,0,1)\rangle = 1 \\ \langle (1,0,i,0),(0,0,0,1)\rangle &= \langle (0,1,i,0),(0,0,0,1)\rangle = 0. \end{aligned}$$

   Logo a matriz da métrica com respeito a esta base é

$$G_B = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

   (c) O vetor $z = (0,1,i,2)$ tem cordenadas $(0,1,2)$ na base B. Nestas coordenadas o plano W tem equação $\langle (0,1,2),(u,v,w)\rangle = 0$, ou seja,

$$[z]_B^T\, G_B \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 0 & 1 & 2 \end{bmatrix}\begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} u \\ v \\ w \end{bmatrix} = 0.$$

   Logo a equação do plano W nas coordenadas dadas pela base B é $u + 2v + 2w = 0$.

9. Enunciado

(a) Esta função é bilinear (porque o produto de matrizes é bilinear e o traço e a transposição são lineares):

$$\begin{aligned}\langle A, \alpha B_1 + \beta B_2\rangle &= \operatorname{tr}(A^T(\alpha B_1 + \beta B_2)) = \operatorname{tr}(\alpha(A^T B_1) + \beta(A^T B_2)) \\ &= \alpha \operatorname{tr}(A^T B_1) + \beta \operatorname{tr}(A^T B_2) = \alpha\langle A, B_1\rangle + \beta\langle A, B_2\rangle\end{aligned}$$

A linearidade na primeira variável é análoga (e segue também da simetria). A simetria é dada pelas seguintes igualdades

$$\langle B, A\rangle = \operatorname{tr}(B^T A) = \operatorname{tr}((A^T B)^T) = \operatorname{tr}(A^T B) = \langle A, B\rangle$$

onde a penúltima igualdade segue de o traço de uma matriz ser igual ao da sua transposta. Para provar o axioma da positividade notamos que se as colunas de A são dadas por n vetores $v_1, \ldots, v_n$ então as entradas na diagonal da matriz $A^T A$ são $(A^T A)_{jj} = \|v_j\|^2$, com $1 \leq j \leq n$. Logo

$$\langle A, A\rangle = \operatorname{tr}(A^T A) = \|v_1\|^2 + \ldots + \|v_n\|^2 \geq 0$$

e se $\langle A, A\rangle = 0$ então $\|v_j\| = 0$ para todo o j e portanto $A = 0$.
Note-se que este produto interno (que é o produto interno usual nas matrizes, não necessariamente quadradas) se identifica com o produto interno usual em $\mathbb{R}^{n^2}$ quando pensamos nas matrizes $n \times n$ como listas de n^2 números. Isto é, para calcular o produto interno temos que multiplicar as entradas correspondentes das matrizes e depois somá-las. Por exemplo, para $n = 2$ temos

$$\operatorname{tr}\left(\begin{bmatrix} a & b \\ c & d\end{bmatrix}^T \begin{bmatrix} x & y \\ z & w\end{bmatrix}\right) = \operatorname{tr}\left(\begin{bmatrix} ax+cz & ay+cw \\ bx+cz & by+dw\end{bmatrix}\right) = ax + by + cz + dw$$

É portanto fácil calcular de cabeça este produto interno de matrizes.

(b) Sabendo que as entradas da matriz da métrica são dadas por $(G_B)_{ij} = \langle v_i, v_j\rangle$ onde $B = (v_1, v_2, v_3, v_4)$, usando o produto interno definido na alínea (a) e recordando que G_B é simétrica obtemos

$$G_B = \begin{bmatrix} 6 & 1 & -4 & -2 \\ 1 & 2 & 0 & -3 \\ -4 & 0 & 5 & 4 \\ -2 & -3 & 4 & 13\end{bmatrix}$$

Por exemplo,

$$(G_B)_{13} = \langle v_1, v_3\rangle = \operatorname{tr}\left(\begin{bmatrix} 1 & -1 \\ 2 & 0\end{bmatrix}\begin{bmatrix} 0 & -1 \\ 2 & 0\end{bmatrix}\right) = \operatorname{tr}\left(\begin{bmatrix} -2 & -1 \\ 0 & -2\end{bmatrix}\right) = -4,$$

$$(G_B)_{44} = \langle v_4, v_4\rangle = \operatorname{tr}\left(\begin{bmatrix} 0 & 2 \\ 0 & -3\end{bmatrix}\begin{bmatrix} 0 & 0 \\ 2 & -3\end{bmatrix}\right) = \operatorname{tr}\left(\begin{bmatrix} 4 & -6 \\ -6 & 9\end{bmatrix}\right) = 13.$$

(c) A matriz simétrica mais próxima de $A = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}$ é a projeção ortogonal de A sobre o subespaço W das matrizes simétricas, $P_W(A)$. Uma base para W pode ser

$$B = \left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right).$$

Sendo v_1, v_2, v_3 as matrizes de B é fácil de verificar que esta base é ortogonal, $\|v_1\| = \|v_3\| = 1$ e $\|v_2\|^2 = 2$. Usando esta base para calcular $P_W(A)$ obtemos

$$\begin{aligned} P_W(A) &= \operatorname{proj}_{v_1}(A) + \operatorname{proj}_{v_2}(A) + \operatorname{proj}_{v_3}(A) \\ &= \langle v_1, A\rangle \frac{v_1}{\|v_1\|^2} + \langle v_2, A\rangle \frac{v_2}{\|v_2\|^2} + \langle v_3, A\rangle \frac{v_3}{\|v_3\|^2} \end{aligned}$$

Uma vez que $\langle v_1, A\rangle = 1$, $\langle v_2, A\rangle = 3$ e $\langle v_3, A\rangle = 0$ a matriz simétrica mais próxima de A é a matriz

$$P_W(A) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \frac{3}{2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 3/2 \\ 3/2 & 0 \end{bmatrix}$$

Alternativamente, podemos achar a interseção com W da reta perpendicular a W que passa por $\left[\begin{smallmatrix} 1 & 2 \\ 1 & 0 \end{smallmatrix}\right]$: A direção desta reta é uma matriz perpendicular às três matrizes da base B, por exemplo $\left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right]$. A reta com esta direção que passa por$\left[\begin{smallmatrix} 1 & 2 \\ 1 & 0 \end{smallmatrix}\right]$ tem equação paramétrica

$$\begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} + t \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

Interseta o plano W quando $2 + t = 1 - t \Leftrightarrow t = -\frac{1}{2}$. A projeção em W é portanto dada por

$$\begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} - \tfrac{1}{2} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & \frac{3}{2} \\ \frac{3}{2} & 0 \end{bmatrix}$$

10. Enunciado
    (a) Primeiro notamos que $0 \in S^\perp$, portanto $S^\perp$ é não vazio. Se $u, v \in S^\perp$ então é claro que $\langle x, u+v\rangle = \langle x, u\rangle + \langle x, v\rangle = 0$ para todo o $x \in S$, ou seja, $S^\perp$ é fechado para a soma. Por outro lado, se $v \in S^\perp$ e α é um escalar então $\langle x, \alpha v\rangle = \alpha\langle x, v\rangle = 0$ para todo o $x \in S$, portanto $S^\perp$ é fechado para o produto por escalar. Concluimos que $S^\perp$ é um subespaço vetorial de V.
    (b) Dado $v \in L(S)$ existem vetores $v_1, \ldots, v_n \in S$ e escalares $\alpha_1, \ldots, \alpha_n$ tais que $v = \alpha_1 v_1 + \ldots + \alpha_n v_n$. Para todo o $w \in S^\perp$ temos

    $$\langle w, v\rangle = \langle w, \alpha_1 v_1 + \ldots + \alpha_n v_n\rangle = \alpha_1\langle w, v_1\rangle + \ldots + \alpha_n\langle w, v_n\rangle = 0,$$

    usando a linearidade do produto interno e $\langle w, v_k\rangle = 0$ para todo o k. Logo concluimos que $v \in (S^\perp)^\perp$.
11. Enunciado Vamos resolver as duas alíneas simultaneamente. Seja $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Seja $T\colon V \to V^*$ a função definida por $T(v) = \phi_v$. Dados $\alpha, \beta \in \mathbb{K}$ e $v_1, v_2 \in V$ então para todo o $w \in V$ temos

$$\phi_{\alpha v_1 + \beta v_2}(w) = \langle \alpha v_1 + \beta v_2, w\rangle = \overline{\alpha}\langle v_1, w\rangle + \overline{\beta}\langle v_2, w\rangle = \overline{\alpha}\phi_{v_1}(w) + \overline{\beta}\phi_{v_2}(w)$$

Logo

$$T(\alpha v_1 + \beta v_2) = \overline{\alpha} T(v_1) + \overline{\beta} T(v_2) \tag{83}$$

Quando $\mathbb{K} = \mathbb{R}$ a equação anterior mostra que T é uma transformação linear. Quando $\mathbb{K} = \mathbb{C}$ o argumento de T pode re-escrever-se

$$\alpha v_1 + \beta v_2 = \overline{\alpha} \cdot_{\overline{V}} v_1 + \overline{\beta} \cdot_{\overline{V}} v_2$$

e dado que $\overline{\alpha}, \overline{\beta}$ são escalares complexos arbitrários, a equação (83) significa que T é linear como aplicação de $\overline{V}$ para V^*.

Note-se que $\phi_v(v) = \|v\|^2 \neq 0$ para $v \neq 0$ logo $T(v) \neq 0$ para $v \neq 0$, portanto a transformação T é sempre injetiva. Se V tem dimensão finita, dado que $\dim V^* = \dim V$ vemos que T é um isomorfismo.

12. Enunciado
   (a) Se a norma provém de um produto interno temos $\|v\| = \sqrt{\langle v, v\rangle}$, logo obtemos

$$\begin{aligned} \|v+w\|^2 &+ \|v-w\|^2 = \langle v+w, v+w\rangle + \langle v-w, v-w\rangle \\ &= \langle v,v\rangle + \langle v,w\rangle + \langle w,v\rangle + \langle w,w\rangle + \langle v,v\rangle - \langle v,w\rangle - \langle w,v\rangle + \langle w,w\rangle \\ &= 2\langle v,v\rangle + 2\langle w,w\rangle = 2\|v\|^2 + 2\|w\|^2 \end{aligned}$$

   (b) Começamos por provar que esta função é uma norma. Se $\|x\| = \|(x_1, \ldots, x_n)\| = 0$ então $|x_1| + \ldots + |x_n| = 0$ e portanto $x_i = 0$ para todo o i, com $1 \leq i \leq n$. Por outro lado,

$$\|\alpha x\| = \|(\alpha x_1, \ldots, \alpha x_n)\| = |\alpha x_1| + \ldots + |\alpha x_n| = |\alpha| \|(x_1, \ldots, x_n)\| = |\alpha| \|x\|.$$

   Finalmente

$$\|x+y\| = |x_1 + y_1| + \ldots + |x_n + y_n| \leq |x_1| + |y_1| + \ldots + |x_n| + |y_n| = \|x\| + \|y\|.$$

   Logo a função é uma norma. No entanto, sejam $x = (1, 0, \ldots, 0)$ e $y = (0, \ldots, 0, 1)$. Temos $\|x\| = 1$, $\|y\| = 1$, $\|x+y\| = \|(1, 0, \ldots, 0, 1)\| = 2$ e $\|x-y\| = \|(1, 0, \ldots, 0, -1)\| = 2$. Logo $\|x+y\|^2 + \|x-y\|^2 = 4 + 4 = 8$ e $2\|x\|^2 + 2\|y\|^2 = 4$. Vemos que a igualdade da alínea (a) não se verifica, por isso esta norma não provém de um produto interno. Para $n = 2$ a bola de raio 1 é um quadrado com vértices nos pontos $(1,0), (0,1), (-1,0), (0,-1)$. Para $n = 3$ a bola de raio 1 é um octaedro com vértices nos pontos $(1,0,0), (0,1,0), (0,0,1), (-1,0,0), (0,-1,0), (0,0,-1)$.

13. Enunciado
   (a) A função g é bilinear, porque $\langle \cdot, \cdot \rangle$ é bilinear em $V_{\mathbb{R}}$ e a função que calcula a parte real é linear (sobre $\mathbb{R}$). A simetria segue da seguinte igualdade

$$g(v,w) = \operatorname{Re}\langle v, w\rangle = \operatorname{Re} \overline{\langle w, v\rangle} = \operatorname{Re} \langle w, v\rangle = g(w,v),$$

   enquanto a positividade segue de

$$g(v,v) = \operatorname{Re}\langle v, v\rangle = \operatorname{Re} \|v\|^2 = \|v\|^2 > 0 \quad \text{se} \quad v \neq 0.$$

A função ω também é bilinear porque $\langle \cdot, \cdot \rangle$ é bilinear em $V_{\mathbb{R}}$ e a função parte imaginária é linear (sobre $\mathbb{R}$). Uma vez que

$$\omega(v,w) = \operatorname{Im}\langle v, w\rangle = \operatorname{Im}\overline{\langle w, v\rangle} = -\operatorname{Im}\langle w, v\rangle = -\omega(w,v)$$

ω é anti-simétrica. Finalmente, se para algum $w \in V_{\mathbb{R}}$

$$\omega(v,w) = 0 \quad \forall v \in V_{\mathbb{R}} \quad \Leftrightarrow \quad \operatorname{Im}\langle v, w\rangle = 0 \quad \forall v \in V_{\mathbb{R}}$$

então fazendo $v = iw$ temos

$$0 = \operatorname{Im}\langle v, w\rangle = \operatorname{Im}(-i\langle w, w\rangle) = \operatorname{Im}(-i\|w\|^2) = -\|w\|^2,$$

portanto $w = 0$. Conclui-se que ω é não degenerada.

(b) Por definição das funções g e ω temos

$$\begin{aligned} g(v,iw) &= \operatorname{Re}\langle v, iw\rangle = \operatorname{Re} i\langle v, w\rangle = -\operatorname{Im}\langle v, w\rangle = -\omega(v,w) \quad \text{e} \\ \omega(v,iw) &= \operatorname{Im}\langle v, iw\rangle = \operatorname{Im} i\langle v, w\rangle = \operatorname{Re}\langle v, w\rangle = g(v,w) \end{aligned}$$

como queriamos demonstrar.

(c) Dado $w \in W$, a aplicação $v \mapsto \omega(v,w)$ é um elemento de W^* (o dual de W) e portanto, pelo Exercício 11(a) existe um único elemento $J(w) \in W$ tal que $g(v, J(w)) = \omega(v,w)$ para todo o $v \in V$.

É imediato verificar que a função $J \colon W \to W$ assim definida é linear: dados escalares $\alpha, \beta \in \mathbb{R}$ e $w_1, w_2 \in W$ temos

$$\begin{aligned} g(v, J(\alpha w_1 + \beta w_2)) &= \omega(v, \alpha w_1 + \beta w_2) = \alpha\omega(v,w_1) + \beta\omega(v,w_2) \\ &= \alpha g(v, J(w_1)) + \beta g(v, J(w_2)) = g(v, \alpha J(w_1) + \beta J(w_2)) \end{aligned}$$

e portanto a unicidade do vetor $J(\alpha w_1 + \beta w_2)$ mostra que

$$J(\alpha w_1 + \beta w_2) = \alpha J(w_1) + \beta J(w_2)$$

Alternativamente, podemos descrever J matricialmente: seja $B = (v_1, \ldots, v_n)$ uma base para V, G_B a matriz da métrica do produto interno g e $\Omega \in M_{n\times n}(\mathbb{R})$ a matriz $[\omega(v_i, v_j)]$ (o análogo para ω da matriz da métrica) que satisfaz

$$\omega(v,w) = [v]_B^T \Omega [w]_B$$

Então a identidade $g(v, Jw) = \omega(v,w)$ é equivalente a

$$[v]_B^T G_B A_{J,B,B}[w]_B = [v]_B^T \Omega [w]_B \qquad \text{para todos os } v, w \in W$$

Por sua vez esta identidade é equivalente à equação matricial

$$G_B A_{J,B,B} = \Omega \Leftrightarrow A_{J,B,B} = G_B^{-1}\Omega$$

(note-se que a matriz da métrica é sempre invertível - se $[v]_B$ pertence ao núcleo de G_B então $\langle v, v\rangle = 0$ logo $v = 0$).

Temos

$$g(v, Jw) = \omega(v,w) = -\omega(w,v) = -g(w, Jv) = -g(Jv, w) \quad \text{para todos os } v, w \in W$$

(isto é J é anti-adjunta com respeito ao produto interno g).

Para que g e ω possam ser obtidos de um produto interno complexo em W (para alguma estrutura de espaço vetorial complexo em W) a alínea (b) juntamente com a unicidade do vetor $J(w)$ mostram que J tem de corresponder à multiplicação pelo escalar $(-i)$ no espaço vetorial complexo W e portanto tem de verificar $J^2 = -I$ Reciprocamente, se J verificar $J^2 = -I$ então podemos definir em W uma estrutura de espaço vetorial complexo definindo

$$(a + bi) \cdot w = aw - bJ(w)$$

e um produto interno complexo por $\langle v, w\rangle = g(v, w) + i\omega(v, w)$ (deixamos como exercício a verificação que estas fórmulas definem de facto uma estrutura de espaço vetorial complexo e um produto interno nesse espaço vetorial complexo).

14. Enunciado Temos

$$\begin{aligned} \|x + y\| = \|x\| + \|y\| &\Leftrightarrow \|x + y\|^2 = (\|x\| + \|y\|)^2 \\ &\Leftrightarrow \langle x + y, x + y\rangle = \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \\ &\Leftrightarrow \|x\|^2 + 2\langle x, y\rangle + \|y\|^2 = \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \\ &\Leftrightarrow \langle x, y\rangle = \|x\|\|y\| \end{aligned}$$

Pela desigualdade de Cauchy-Schwarz para que a última igualdade se verifique x e y têm de ser colineares. Suponhamos sem perda de generalidade que $y = \alpha x$ com $\alpha \in \mathbb{R}$, vemos que além disso $\alpha \geq 0$. Reciprocamente se $y = \alpha x$ ou $x = \beta y$ com α, respetivamente β, ≥ 0 a igualdade verifica-se.

Concluímos que $\|x + y\| = \|x\| + \|y\|$ se e só se x, y são dois vetores com a mesma direção e sentido (ou um deles é nulo).

A desigualdade

$$\|x + y\| \geq |\|x\| - \|y\||$$

é equivalente às duas desigualdades

$$(i) \;\; \|x + y\| \geq \|x\| - \|y\| \quad \text{e} \quad (ii) \;\; \|x + y\| \geq \|y\| - \|x\|$$

(de facto, dados $a, b \in \mathbb{R}$ temos $a \geq |b| \Leftrightarrow (a \geq b \wedge a \geq -b)$) que passamos a demonstrar: Pela desigualdade triangular temos

$$\|x\| = \|(x + y) + (-y)\| \leq \|x + y\| + \| - y\| = \|x + y\| + \|y\|$$

e portanto $\|x\| - \|y\| \leq \|x + y\|$, que é a desigualdade (i). A desigualdade (ii) obtém-se de (i) trocando x e y.

15. Enunciado Sabemos que qualquer vetor se decompõe numa soma $v = \operatorname{proj}_u(v) + (v - \operatorname{proj}_u(v))$ onde $\operatorname{proj}_u(v)$ pertence à reta gerada por u e $v - \operatorname{proj}_u(v)$ pertence ao plano perpendicular a u. A reflexão no plano perpendicular é a identidade nos vetores que pertencem ao plano e passa ao simétrico um vetor que pertença à reta gerada por u, logo, porque a reflexão é uma transformação linear, obtemos

$$\begin{aligned} R(v) &= R(\operatorname{proj}_u(v) + (v - \operatorname{proj}_u(v))) = R(\operatorname{proj}_u(v)) + R(v - \operatorname{proj}_u(v))) \\ &= -\operatorname{proj}_u(v) + v - \operatorname{proj}_u(v) = v - 2\operatorname{proj}_u(v) = v - 2\frac{\langle u, v\rangle}{\langle u, u\rangle}u, \end{aligned}$$

como queriamos provar.

16. Enunciado Sejam $v_1 = (\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}, 0)$ e $v_2 = (-\frac{1}{\sqrt{3}}, 0, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$. Se escolhermos $v_3 = (0,1,0,0)$ é fácil de verificar que $\langle v_1, v_3\rangle = \langle v_2, v_3\rangle = 0$ e $\|v_3\| = 1$, portanto falta-nos apenas um vetor para obter uma base ortonormada. Seja $w = (0,0,0,1)$. Este vetor é obviamente ortogonal a v_1 e v_3, mas não é ortogonal a v_2. No entanto o vetor

$$\begin{aligned} u &= w - \operatorname{proj}_{v_2}(w) = w - \langle v_2, w\rangle v_2 = (0,0,0,1) - \frac{1}{\sqrt{3}}\left(-\frac{1}{\sqrt{3}}, 0, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right) \\ &= \left(\frac{1}{3}, 0, -\frac{1}{3}, \frac{2}{3}\right) \end{aligned}$$

é ortogonal a v_2. A norma de u é $\sqrt{\frac{2}{3}}$ e portanto definimos

$$v_4 = \frac{u}{\|u\|} = \sqrt{\frac{3}{2}}\left(\frac{1}{3}, 0, -\frac{1}{3}, \frac{2}{3}\right) = \left(\frac{1}{\sqrt{6}}, 0, -\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}\right)$$

e $B = (v_1, v_2, v_3, v_4)$ é uma base ortonormada de $\mathbb{R}^4$ como pedido.

As coordenadas do vetor $v = (1,2,3,4)$ na base B são dadas por

$$[(1,2,3,4)]_B = \begin{bmatrix} \langle v, v_1\rangle \\ \langle v, v_2\rangle \\ \langle v, v_3\rangle \\ \langle v, v_4\rangle \end{bmatrix} = \begin{bmatrix} 2\sqrt{2} \\ 2\sqrt{3} \\ 2 \\ \sqrt{6} \end{bmatrix}.$$

17. Enunciado Multiplicando A^T por um vetor v obtemos um vetor coluna em que as entradas são os produtos internos com os vetores u_i, ou seja, $(A^T v)_i = \langle u_i, v\rangle$. Logo

$$AA^T v = \langle u_1, v\rangle u_1 + \ldots + \langle u_k, v\rangle u_k = \operatorname{proj}_{u_1}(v) + \ldots + \operatorname{proj}_{u_k}(v) = P_U(v)$$

e portanto a matriz AA^T representa a projeção ortogonal P_U com respeito à base canónica.

18. Enunciado Começamos por considerar a base $B = ((1,-2,0,1),(1,-1,1,0)) = (v_1, v_2)$ do plano que denotamos por W. Esta base não é ortogonal, mas podemos obter uma base ortogonal usando o processo de ortogonalização de Gram-Schmidt. O primeiro vetor da base ortogonal é $w_1 = v_1 = (1,-2,0,1)$ cuja norma é $\sqrt{6}$. O segundo vetor é

$$w_2 = v_2 - \operatorname{proj}_{w_1}(v_2) = v_2 - \langle w_1, v_2\rangle \frac{w_1}{\|w_1\|^2} = (1,-1,1,0) - \frac{3}{6}(1,-2,0,1) = \left(\frac{1}{2}, 0, 1, -\frac{1}{2}\right),$$

cuja norma é $\sqrt{\frac{3}{2}}$. A projeção ortogonal de $v = (x, y, z, w) \in \mathbb{R}^4$ sobre W é então dada por

$$\begin{aligned} P_W(v) &= \operatorname{proj}_{w_1}(v) + \operatorname{proj}_{w_2}(v) = \langle w_1, v\rangle \frac{w_1}{\|w_1\|^2} + \langle w_2, v\rangle \frac{w_2}{\|w_2\|^2} = \\ &= \frac{1}{6}(x - 2y + w)(1, -2, 0, 1) + \frac{2}{3}\left(\frac{x}{2} + z - \frac{w}{2}\right)\left(\frac{1}{2}, 0, 1, -\frac{1}{2}\right) = \\ &= \frac{1}{3}(x - y + z, -x + 2y - w, x + 2z - w, -y - z + w). \end{aligned}$$

Alternativamente, podemos usar o Exercício anterior. A matriz que representa a projeção ortogonal na base canónica é

$$\begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ -\frac{2}{\sqrt{6}} & 0 \\ 0 & \sqrt{\frac{2}{3}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} & 0 & \sqrt{\frac{2}{3}} & -\frac{1}{\sqrt{6}} \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & -\frac{1}{3} & \frac{1}{3} & 0 \\ -\frac{1}{3} & \frac{2}{3} & 0 & -\frac{1}{3} \\ \frac{1}{3} & 0 & \frac{2}{3} & -\frac{1}{3} \\ 0 & -\frac{1}{3} & -\frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

o que reproduz a expressão achada acima.

19. Enunciado Começamos por calcular

$$P^2(x, y) = P\left(\frac{x+y}{2}, \frac{x+y}{2}\right) = \left(\frac{x+y}{4} + \frac{x+y}{4}, \frac{x+y}{4} + \frac{x+y}{4}\right) = \left(\frac{x+y}{2}, \frac{x+y}{2}\right),$$

portanto $P^2 = P$. Por outro lado o núcleo de P é $N(P) = \{\alpha(1, -1) : \alpha \in \mathbb{R}\}$ e a imagem de P é $P(\mathbb{R}^2) = \{\alpha(1, 1) : \alpha \in \mathbb{R}\}$, logo o núcleo e a imagem são ortogonais. Concluímos que P é uma projeção ortogonal.

20. Enunciado
    (a) Primeiro obtemos uma expressão para este produto interno. Para isso calculamos a matriz da métrica com respeito à base canónica. Sendo $B = ((1, 0, 1), (0, 1, 0), (1, 0, -1))$ temos

$$G_B = S^T G_{B_c} S \iff G_{B_c} = (S^T)^{-1} G_B S^{-1}$$

onde S é a matriz mudança de base $S_{B \to B_c}$, ou seja,

$$S = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{bmatrix} \quad \text{e} \quad S^{-1} = \frac{1}{2}\begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & -1 \end{bmatrix}$$

Uma vez que $G_B = I$ e S é simétrica obtemos

$$G_{B_c} = (S^{-1})^2 = \frac{1}{2}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Observe-se que, com um pouco de reflexão, podíamos ter escrito imediatamente G_{B_c}: a base dada é ortogonal para o produto interno usual. A única diferença para

o produto interno usual é que, no plano xz, os comprimentos são divididos por $\sqrt{2}$ (de forma que os vetores ortogonais $(1, 0, \pm 1)$ tenham comprimento 1).
Logo

$$\langle (x, y, z), (u, v, w) \rangle = \begin{bmatrix} x & y & z \end{bmatrix} G_{B_c} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \frac{1}{2} xu + yv + \frac{1}{2} zw.$$

Agora consideramos a base de U formada pelos vetores $v_1 = (1, 0, -1)$ e $v_2 = (0, 1, -1)$. Com este produto interno $\langle v_1, v_2 \rangle = \frac{1}{2}$. Para obter uma base ortogonal usamos o processo de ortogonalização de Gram-Schmidt: definimos $w_1 = v_1$ e

$$w_2 = v_2 - \langle w_1, v_2 \rangle \frac{w_1}{\|w_1\|^2} = (0, 1, -1) - \frac{1}{2}(1, 0, -1) = \left(-\frac{1}{2}, 1, -\frac{1}{2} \right),$$

pois $\|w_1\| = 1$. A norma de w_2 é

$$\|w_2\| = \sqrt{\frac{1}{2}.\frac{1}{4} + 1 + \frac{1}{2}.\frac{1}{4}} = \sqrt{\frac{5}{4}} = \frac{\sqrt{5}}{2}.$$

Assim obtemos uma base ortonormada para o subespaço U

$$B = ((1, 0, -1), \tfrac{1}{\sqrt{5}}(-1, 2, -1)).$$

Acima fizemos as contas nas coordenadas da base canónica. Alternativamente, podemos fazer as contas nas coordenadas da base B onde o produto interno se calcula como habitualmente: a relação entre as coordenadas (u, v, w) na base B e as coordenadas (x, y, z) na base canónica é dada por

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} \Leftrightarrow \begin{cases} x = u + w \\ y = v \\ z = u - w \end{cases}$$

A equação do plano U nas novas coordenadas obtém-se substitutindo x, y, z pelas expressões acima na equação $x + y + z = 0$. Obtemos

$$(u + w) + v + (u - w) = 0 \Leftrightarrow 2u + v = 0$$

Note-se o significado preciso da equação anterior: é a tradução de U nas novas coordenadas, isto é,

$$\{(u, v, w) \in \mathbb{R}^3 \colon 2u + v = 0\} = \{[x]_B \in \mathbb{R}^3 \colon x \in U\}$$

O plano U nas novas coordenadas é então $\{(u, -2u, w) \colon u, w \in \mathbb{R}\}$. Uma base é $\{(1, -2, 0), (0, 0, 1)\}$ e esta base é já ortogonal (recorde-se que nestas coordenadas o produto interno se calcula da forma usual). Uma base ortonormada obtém-se dividindo pela norma: $\{\frac{1}{\sqrt{5}}(1, -2, 0), (0, 0, 1)\}$. A base ortonormada pretendida é agora a tradução desta base nas coordenadas iniciais (as da base canónica). Temos

$$(-1, 2, 0) \leftrightarrow -(1, 0, 1) + 2(0, 1, 0) = (-1, 2, -1) \qquad (0, 0, 1) \Leftrightarrow (1, 0, -1)$$

Conclui-se que uma base ortogonal para U com respeito ao produto interno do enunciado é dada por

$$\{\tfrac{1}{\sqrt{5}}(-1,2,-1),(1,0,-1)\}$$

(b) Usando a base obtida na alínea anterior podemos calcular a projeção ortogonal de $v=(1,2,-1)$ sobre o subespaço U

$$\begin{aligned}P_U(v) &= \mathrm{proj}_{w_1}(v)+\mathrm{proj}_{w_2}(v)=\langle w_1,v\rangle w_1+\langle w_2,v\rangle w_2=\\ &= \langle(1,0,-1),(1,2,-1)\rangle w_1+\langle\tfrac{1}{\sqrt{5}}(-1,2,-1)),(1,2,-1)\rangle w_2=\\ &= (1,0,-1)+\tfrac{4}{5}(-1,2,-1)=\left(\tfrac{1}{5},\tfrac{8}{5},-\tfrac{9}{5}\right)\end{aligned}$$

O vetor v decompõe-se então como $v=P_U(v)+(v-P_U(v))$ onde $P_U(v)\in U$ e $v-P_U(v)=\left(\frac{4}{5},\frac{2}{5},\frac{4}{5}\right)\in U^\perp$.

21. Enunciado

(a) O ângulo é dado pela expressão

$$\alpha=\arccos\frac{\left\langle\begin{bmatrix}1&2\\-1&0\end{bmatrix},\begin{bmatrix}1&1\\1&1\end{bmatrix}\right\rangle}{\left\|\begin{bmatrix}1&2\\-1&0\end{bmatrix}\right\|\left\|\begin{bmatrix}1&1\\1&1\end{bmatrix}\right\|}$$

Recordamos da resoluçãõ do Exercício 9 que este produto interno se calcula da forma usual (isto é multiplicando as entradas correspondentes das duas matrizes e somando). Logo

$$\left\langle\begin{bmatrix}1&2\\-1&0\end{bmatrix},\begin{bmatrix}1&1\\1&1\end{bmatrix}\right\rangle=2$$

$$\left\|\begin{bmatrix}1&2\\-1&0\end{bmatrix}\right\|=\sqrt{6},\qquad\left\|\begin{bmatrix}1&1\\1&1\end{bmatrix}\right\|=\sqrt{4}=2$$

Conclui-se que o ângulo pretendido é $\arccos\frac{1}{\sqrt{6}}$.

(b) Temos uma fórmula para a projeção ortogonal em U que envolve determinar uma base ortogonal para U. Como $U^\perp$ tem dimensão 1 é mais simples calcular a fórmula para a projeção em $U^\perp$ e depois fazer uso da relação $P_U(x)+P_{U^\perp}(x)=x$.
Temos

$$U=\left\{\begin{bmatrix}a&b\\c&d\end{bmatrix}: a+d=0\right\}$$

Dado que o produto interno considerado nas matrizes se calcula como o produto interno em $\mathbb{R}^4$ (ver resolução do Exercício 9) temos

$$a+d=0\Leftrightarrow\left\langle\begin{bmatrix}a&b\\c&d\end{bmatrix},\begin{bmatrix}1&0\\0&1\end{bmatrix}\right\rangle=0$$

logo $U^\perp=L(\{\left[\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right]\})$. Portanto

$$P_{U^\perp}\left(\left[\begin{smallmatrix}1&2\\-1&0\end{smallmatrix}\right]\right)=\mathrm{proj}_{\left[\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right]}\left[\begin{smallmatrix}1&2\\-1&0\end{smallmatrix}\right]=\frac{\langle\left[\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right],\left[\begin{smallmatrix}1&2\\-1&0\end{smallmatrix}\right]\rangle}{\langle\left[\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right],\left[\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right]\rangle}\left[\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right]=\tfrac{1}{2}\left[\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right]$$

Logo

$$P_U\left(\left[\begin{smallmatrix}1&2\\-1&0\end{smallmatrix}\right]\right)=\left[\begin{smallmatrix}1&2\\-1&0\end{smallmatrix}\right]-P_{U^\perp}\left(\left[\begin{smallmatrix}1&2\\-1&0\end{smallmatrix}\right]\right)=\left[\begin{smallmatrix}1&2\\-1&0\end{smallmatrix}\right]-\begin{bmatrix}\frac12&0\\0&\frac12\end{bmatrix}=\begin{bmatrix}\frac12&2\\-1&-\frac12\end{bmatrix}$$

22. Enunciado Aplicando o método de ortogonalização de Gram-Schmidt à base $(1,x)$ do subespaço $U=\{a+bx\colon a,b\in\mathbb{R}\}$ obtemos a base ortogonal formada pelos polinómios 1 e

$$x-\frac{\langle 1,x\rangle}{\langle 1,1\rangle}1=x-\frac{\int_0^1 x\ dx}{\int_0^1 1\ dx}1=x-\frac{\frac12}{1}1=x-\tfrac12.$$

Dado que $a+bx\in U$ temos $P_U(a+bx)=a+bx$ logo resta-nos calcular $P_U(x^2)$:

$$\begin{aligned}P_U(x^2)&=\operatorname{proj}_1(x^2)+\operatorname{proj}_{x-\frac12}(x^2)\\&=\frac{\int_0^1 x^2\ dx}{\int_0^1 1\ dx}+\frac{\int_0^1(x-\frac12)x^2\ dx}{\int_0^1(x-\frac12)^2\ dx}(x-\tfrac12)\\&=\tfrac13+\frac{\int_0^1-\frac12x^2+x^3\ dx}{\frac13\ (x-\frac12)^3\big|_0^1}(x-\tfrac12)\\&=\tfrac13+12\left(-\tfrac16+\tfrac14\right)(x-\tfrac12)=\tfrac13+(x-\tfrac12)\end{aligned}$$

logo

$$P_U(a+bx+cx^2)=a+bx+\tfrac{c}{3}+c(x-\tfrac12)=a-\tfrac{c}{6}+(b+c)x$$

23. Enunciado As funções $\operatorname{sen}x,\operatorname{sen}(2x)$ e $\operatorname{sen}(3x)$ são ortogonais (cf. Exercício 5 (c)) pelo que basta calcular

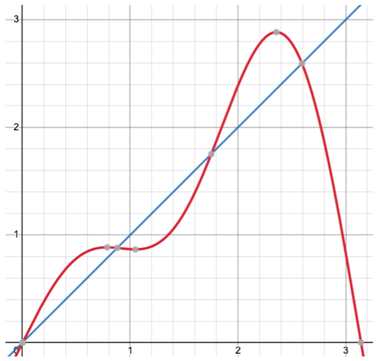
$$\begin{aligned}\langle\operatorname{sen}(nx),x\rangle&=\int_0^\pi x\operatorname{sen}(nx)\ dx\\&=-x\frac{\cos(nx)}{n}\Big|_0^\pi-\int_0^\pi(-\frac{\cos nx}{n})\ dx\\&=\frac{(-1)^{n+1}\pi}{n}+\frac{\operatorname{sen}nx}{n^2}\Big|_0^\pi\\&=\frac{(-1)^{n+1}\pi}{n}\end{aligned}$$

e novamente do exercício 5(c)

$$\|\operatorname{sen}nx\|^2=\frac{\pi}{2}.$$

Assim, sendo $U=L(\{\operatorname{sen}x,\operatorname{sen}2x,\operatorname{sen}3x\})$ temos

$$\begin{aligned}P_U(x)&=\operatorname{proj}_{\operatorname{sen}x}(x)+\operatorname{proj}_{\operatorname{sen}2x}(x)+\operatorname{proj}_{\operatorname{sen}3x}(x)\\&=\frac{\pi}{\frac{\pi}{2}}\operatorname{sen}x-\frac{\frac{\pi}{2}}{\frac{\pi}{2}}\operatorname{sen}2x+\frac{\frac{\pi}{3}}{\frac{\pi}{2}}\operatorname{sen}3x\\&=2\operatorname{sen}x-\operatorname{sen}2x+\tfrac23\operatorname{sen}3x\end{aligned}$$



24. Enunciado
   (a) A equação cartesiana de uma reta que passa pelo ponto x_0 e é perpendicular ao plano gerado pelos vetores v_1 e v_2 é
$$\begin{cases} \langle v_1, x - x_0 \rangle = 0 \\ \langle v_2, x - x_0 \rangle = 0 \end{cases}$$
   Precisamos de dois vetores não colineares perpendiculares à reta do enunciado. Uma vez que a direção da reta é a do vetor perpendicular ao plano $(-1, 2, 3)$ podemos por exemplo considerar os vetores $v_1 = (2, 1, 0)$ e $v_2 = (3, 0, 1)$. Concluímos que uma equação cartesiana da reta é por exemplo
$$\begin{cases} 2(x-1) + y = 0 \\ 3(x-1) + z = 0 \end{cases} \Leftrightarrow \begin{cases} 2x + y = 2 \\ 3x + z = 3 \end{cases}$$
   (b) A equação cartesiana acha-se exatamente como na alínea anterior excepto que agora foram-nos dados os vetores perpendiculares. Uma equação é por exemplo
$$\begin{cases} x + 2z + w = 1 + 2 + 1 \\ y + 2z - w = 1 + 2 - 1 \end{cases} \Leftrightarrow \begin{cases} x + 2z + w = 4 \\ y + 2z - w = 2 \end{cases}$$
25. Enunciado
   (a) A reta perpendicular ao plano que passa pelo ponto dado tem equação paramétrica $(0, 1, 2) + t(1, 1, 1)$ com $t \in \mathbb{R}$. O ponto de intersecção desta reta com o plano é o que satisfaz
$$t + (1 + t) + (2 + t) = 1 \Leftrightarrow t = -\tfrac{2}{3}$$
   Pelo Teorema de Pitágoras, a distância do ponto ao plano é $\|(0, 1, 2) - P\|$ onde P é o ponto de intersecção. Uma vez que
$$(0, 1, 2) - P = \tfrac{2}{3}(1, 1, 1)$$
   conclui-se que a distância é $\frac{2}{\sqrt{3}}$.
   (b) Novamente pelo Teorema de Pitágoras, a distância pretendida é igual à distância entre $(1, 2, 1)$ e o ponto de interseção do plano perpendicular à reta que passa por

$(1,2,1)$ com a reta. Este plano tem equação $3y - z = 3\cdot 2 - 1 = 5$. A sua interseção com a reta é o ponto com parâmetro α satisfazendo

$$3(3\alpha) - (-\alpha) = 5 \Leftrightarrow \alpha = \tfrac{1}{2}$$

logo a distância pretendida é

$$\|(1,2,1) - \tfrac{1}{2}(0,3,-1)\| = \|(1,\tfrac{1}{2},\tfrac{3}{2})\| = \sqrt{\tfrac{7}{2}}$$

(c) Pelo Teorema de Pitágoras a distância entre os planos será a distância de qualquer ponto do primeiro plano ao ponto do segundo que se obtém intersetando-o com uma reta perpendicular aos dois planos que passa pelo ponto inicial. Consideremos por exemplo o ponto $(5,0,0,0)$ do primeiro plano. A reta perpendicular aos dois planos que passa por ele tem equação paramétrica

$$(5,0,0,0) + t(1,1,-1,3)$$

O ponto de interseção desta reta com o segundo plano é o que tem parâmetro

$$(5+t) + t - (-t) + 3(3t) = 2 \Leftrightarrow t = -\tfrac{1}{4}$$

Logo a distância entre $(5,0,0,0)$ e o ponto de interseção com o segundo plano é

$$\|\tfrac{1}{4}(1,1,-1,3)\| = \tfrac{\sqrt{3}}{2}$$

(d) Pelo Teorema de Pitágoras, a distância é a distância de $(-1,0,1)$ ao ponto de intersecção da reta com o plano perpendicular à reta que passa por $(-1,0,1)$. As direções perpendiculares à reta podem ler-se da equação cartesiana: formam o plano $L(\{(1,0,1),(1,1,1)\})$; assim o plano perpendicular à reta que passa por $(-1,0,1)$ tem equação paramétrica

$$(-1,0,1) + t(1,0,1) + s(1,1,1) = (-1+t+s, s, 1+t+s) \qquad t,s,\in\mathbb{R}$$

O ponto de intersecção deste plano com a reta dada corresponde aos parâmetros que satisfazem

$$\begin{cases}(-1+t+s) + (1+t+s) = 2\\ (-1+t+s) + s + (1+t+s) = 1\end{cases} \Leftrightarrow \begin{cases}t = 2\\ s = -1\end{cases}$$

A distância é portanto

$$\|2(1,0,1) - (1,1,1)\| = \|(1,-1,1)\| = \sqrt{3}$$

26. Enunciado

(a) O ângulo entre dois vetores $u \in U$ e $v \in V$ está por definição entre 0 e π. Uma vez que

$$\angle(-u,v) = \arccos\frac{\langle -u,v\rangle}{\|-u\|\|v\|} = \arccos\left(-\frac{\langle u,v\rangle}{\|u\|\|v\|}\right) = \pi - \arccos\frac{\langle u,v\rangle}{\|u\|\|v\|} = \pi - \angle(u,v)$$

e $-u \in U$, é sempre possível encontrar vetores em U e V que fazem um ângulo entre 0 e $\frac{\pi}{2}$ e portanto o ângulo entre os planos tem de estar neste intervalo.

(b) O ângulo é dado por

$$\begin{aligned}\arccos\frac{\mathrm{Re}(\langle(1+i,i),(2-i,1+i)\rangle)}{\|(1+i,i)\|\|(2-i,1+i)\|} &= \arccos\frac{\mathrm{Re}((1-i)(2-i)-i(1+i))}{\sqrt{2+1}\sqrt{5+2}}\\ &= \arccos\frac{2}{\sqrt{21}}\end{aligned}$$

(c) Seja $V = L(\{v\})$ e $W = L(\{w\})$ as linhas complexas geradas pelos vetores não nulos $v, w \in \mathbb{C}^n$. Os vetores não nulos em V e W podem escrever-se na forma αv e βw com $\alpha, \beta \in \mathbb{C}\setminus\{0\}$. Temos portanto que minimizar

$$\arccos\mathrm{Re}\left(\frac{\langle\alpha v,\beta w\rangle}{|\alpha|\|v\||\beta|\|w\|}\right) = \arccos\mathrm{Re}\left(\frac{\overline{\alpha}\beta}{|\alpha||\beta|}\frac{\langle v,w\rangle}{\|v\|\|w\|}\right)$$

o que corresponde a maximizar

$$\mathrm{Re}\left(\frac{\overline{\alpha}\beta}{|\alpha||\beta|}\frac{\langle v,w\rangle}{\|v\|\|w\|}\right)$$

como função de α, β. Escrevendo

$$\frac{\overline{\alpha}\beta}{|\alpha||\beta|} = \cos\theta + i\,\mathrm{sen}\,\theta$$

e $\langle v,w\rangle = |\langle v,w\rangle|(\cos\phi + i\,\mathrm{sen}\,\phi)$ temos

$$\begin{aligned}\mathrm{Re}\left(\frac{\overline{\alpha}\beta}{|\alpha||\beta|}\frac{\langle v,w\rangle}{\|v\|\|w\|}\right) &= \mathrm{Re}\left((\cos\theta+i\,\mathrm{sen}\,\theta)\frac{(\cos\phi+i\,\mathrm{sen}\,\phi)|\langle v,w\rangle|}{\|v\|\|w\|}\right)\\ &= \mathrm{Re}\left((\cos\theta+i\,\mathrm{sen}\,\theta)(\cos\phi+i\,\mathrm{sen}\,\phi)\right)\frac{|\langle v,w\rangle|}{\|v\|\|w\|}\\ &= \cos(\theta+\phi)\frac{|\langle v,w\rangle|}{\|v\|\|w\|}\end{aligned}$$

Claramente o máximo ocorre quando $\theta = -\phi$ e é dado por $\frac{|\langle v,w\rangle|}{\|v\|\|w\|}$ conforme pretendido.

(d) Podíamos começar por achar a intersecção W (que terá dimensão 2 dado que $U + V = \mathbb{R}^4$), depois achar $W^\perp$ (que terá também dimensão 2) e depois calcular o ângulo entre as retas $U\cap W^\perp, V\cap W^\perp \subset W^\perp$.
No entanto os cálculos são facilitados se notarmos que $W^\perp = U^\perp + V^\perp$: de facto, as linhas $U^\perp = L(\{(1,-1,1,-1)\})$ e $V^\perp = L(\{(2,1,-1,3)\})$ são perpendiculares a $W = U\cap V$ (porque este subespaço está contido em U e V respetivamente) logo $U^\perp + V^\perp \subset W^\perp$ e, tendo ambos os espaços dimensão 2, eles são iguais.
Conclui-se assim que $W^\perp = L(\{(1,-1,1,-1),(2,1,-1,3)\})$. Um vetor

$$\alpha(1,-1,1,-1)+\beta(2,1,-1,3) = (\alpha+2\beta,-\alpha+\beta,\alpha-\beta,-\alpha+3\beta)$$

pertence a U se

$$(\alpha+2\beta)-(-\alpha+\beta)+(\alpha-\beta)-(-\alpha+3\beta) = 0 \Leftrightarrow 4\alpha-3\beta=0 \Leftrightarrow \beta = \tfrac{4}{3}\alpha$$

Logo $U \cap W^{\perp} = L(\{(1,-1,1,-1) + \frac{4}{3}(2,1,-1,3)\}) = L(\{(\frac{11}{3}, \frac{1}{3}, -\frac{1}{3}, 3)\})$.
O vetor

$$\alpha(1,-1,1,-1) + \beta(2,1,-1,3) = (\alpha + 2\beta, -\alpha + \beta, \alpha - \beta, -\alpha + 3\beta)$$

pertence a V se

$$2(\alpha + 2\beta) + (-\alpha + \beta) - (\alpha - \beta) + 3(-\alpha + 3\beta) = 0 \Leftrightarrow -3\alpha + 15\beta = 0 \Leftrightarrow \beta = \frac{1}{5}\alpha$$

Logo $V \cap W^{\perp} = L(\{(1,-1,1,-1) + \frac{1}{5}(2,1,-1,3)\}) = L(\{(\frac{7}{5}, -\frac{4}{5}, \frac{4}{5}, -\frac{2}{5})\})$.
Temos

$$\frac{\langle(\frac{11}{3}, \frac{1}{3}, -\frac{1}{3}, 3), (\frac{7}{5}, -\frac{4}{5}, \frac{4}{5}, -\frac{2}{5})\rangle}{\|(\frac{11}{3}, \frac{1}{3}, -\frac{1}{3}, 3)\|\|(\frac{7}{5}, -\frac{4}{5}, \frac{4}{5}, -\frac{2}{5})\|} = \frac{\frac{51}{15}}{\sqrt{\frac{204}{9}}\sqrt{\frac{17}{5}}} = \frac{\sqrt{3}}{2\sqrt{5}}$$

Logo o ângulo dihedral é $\arccos \frac{\sqrt{3}}{2\sqrt{5}}$.

27. Enunciado

(a) O sistema que queremos resolver aproximadamente é

$$\begin{bmatrix} -1 & 1 \\ 0 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 6 \end{bmatrix}$$

A equação dos quadrados mínimos é

$$\begin{bmatrix} -1 & 0 & 2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 0 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} -1 & 0 & 2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 6 \end{bmatrix} \Leftrightarrow \begin{bmatrix} 5 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 10 \\ 11 \end{bmatrix}$$

Logo

$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 5 & 1 \\ 1 & 3 \end{bmatrix}^{-1} \begin{bmatrix} 10 \\ 11 \end{bmatrix} = \frac{1}{14} \begin{bmatrix} 3 & -1 \\ -1 & 5 \end{bmatrix} \begin{bmatrix} 10 \\ 11 \end{bmatrix} = \begin{bmatrix} \frac{19}{14} \\ \frac{45}{14} \end{bmatrix}$$

e portanto a reta pedida é $y = \frac{19}{14}x + \frac{45}{14}$

(b) O sistema que queremos resolver aproximadamente é

$$\begin{bmatrix} \cos -\frac{\pi}{2} & \operatorname{sen} -\frac{\pi}{2} \\ \cos 0 & \operatorname{sen} 0 \\ \cos \frac{\pi}{2} & \operatorname{sen} \frac{\pi}{2} \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \Leftrightarrow \begin{bmatrix} 0 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$$

A equação dos quadrados mínimos é

$$\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \Leftrightarrow \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Logo a função pedida é $y = \frac{1}{2} \operatorname{sen} x$

(c) Escrevendo $y = a + bx + cx^2$ para o polinómio em questão, o sistema que queremos resolver aproximadamente é

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 1 \end{bmatrix}$$

A equação dos quadrados mínimos é

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & -1 \\ 0 & 1 & 4 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & -1 \\ 0 & 1 & 4 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 2 \\ 1 \end{bmatrix} \Leftrightarrow \begin{bmatrix} 4 & 2 & 6 \\ 2 & 6 & 8 \\ 6 & 8 & 18 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \\ 10 \end{bmatrix}$$

que produz $a = \frac{3}{10}, b = -\frac{1}{10}, c = \frac{1}{2}$. Logo o polinómio pretendido é $\frac{3}{10} - \frac{1}{10}x + \frac{1}{2}x^2$.

(d) Queremos resolver aproximadamente o sistema

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & -1 & 2 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 3 \end{bmatrix}$$

A equação dos quadrados mínimos é

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 2 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & -1 & 2 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 2 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 3 \end{bmatrix} \Leftrightarrow \begin{bmatrix} 4 & 1 & 4 \\ 1 & 3 & 0 \\ 4 & 0 & 8 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \\ 6 \end{bmatrix}$$

que produz $a = -\frac{1}{5}, b = \frac{7}{5}, c = \frac{17}{20}$ logo o plano da forma dada que melhor aproxima estes quatro pontos é

$$z = -\tfrac{1}{5} + \tfrac{7}{5}x + \tfrac{17}{20}y$$

28. Enunciado $(i) \Rightarrow (ii)$: Dados $x, y \in V$ e uma transformação unitária/ortogonal $T\colon V \to V$ temos

$$\|T(x) - T(y)\|^2 = \langle T(x) - T(y), T(x) - T(y)\rangle = \langle T(x-y), T(x-y)\rangle = \langle x-y, x-y\rangle = \|x-y\|^2$$

onde na penúltima igualdade usámos o facto que T preserva o produto interno.
$(ii) \Rightarrow (iii)$: Tomando $y = 0$ na afirmação (ii) vemos que $\|T(x)\| = \|x\|$ para todo o $x \in V$. Em particular, se $\|x\| = 1$ temos $\|T(x)\| = 1$.

$(iii) \Rightarrow (i)$: Começamos por ver que a afirmação (iii) implica que $\|T(x)\| = \|x\|$ para todo o $x \in V$. Isto é evidente se $x = 0$. Para $x \neq 0$, temos

$$\begin{aligned} \left\|\frac{x}{\|x\|}\right\| = 1 \quad &\Rightarrow \quad \left\|T\left(\frac{x}{\|x\|}\right)\right\| = 1 \\ &\Leftrightarrow \quad \left\|\frac{1}{\|x\|}T(x)\right\| = 1 \\ &\Leftrightarrow \quad \frac{1}{\|x\|}\|T(x)\| = 1 \Leftrightarrow \|T(x)\| = \|x\| \end{aligned}$$

O resultado segue agora da *igualdade de polarização* que relaciona o produto interno com a norma: dados $x, y \in V$ temos

$$\langle x+y, x+y\rangle = \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle$$

Portanto, considerando primeiro o caso real

$$\|x+y\|^2 = \|x\|^2 + 2\langle x, y\rangle + \|y\|^2 \Leftrightarrow \langle x, y\rangle = \frac{\|x+y\|^2 - \|x\|^2 - \|y\|^2}{2}$$

Esta fórmula, que descreve o produto interno em função da norma, mostra que qualquer transformação que preserve a norma vai também preservar o produto interno: assumindo que $\|T(v)\| = \|v\|$ para todo o $v \in V$ temos

$$\begin{aligned} \langle T(x), T(y)\rangle &= \frac{\|T(x)+T(y)\|^2 - \|T(x)\|^2 - \|T(y)\|^2}{2} \\ &= \frac{\|T(x+y)\|^2 - \|T(x)\|^2 - \|T(y)\|^2}{2} \\ &= \frac{\|x+y\|^2 - \|x\|^2 - \|y\|^2}{2} = \langle x, y\rangle \end{aligned}$$

No caso complexo, uma vez que $\langle x, y\rangle + \langle y, x\rangle = \langle x, y\rangle + \overline{\langle x, y\rangle} = 2\,\mathrm{Re}(\langle x, y\rangle)$, a hipótese que $\|T(v)\| = \|v\|$ garante que $\mathrm{Re}(\langle T(x), T(y)\rangle) = \mathrm{Re}(\langle x, y\rangle)$. Daqui se conclui facilmente que T preserva o produto interno conforme requerido, pois

$$\begin{aligned} \mathrm{Im}(\langle T(x), T(y)\rangle) &= \mathrm{Re}(-i\langle T(x), T(y)\rangle) = \mathrm{Re}(\langle iT(x), T(y)\rangle) \\ &= \mathrm{Re}(\langle T(ix), T(y)\rangle) = \mathrm{Re}(\langle ix, y\rangle) = \mathrm{Re}(-i\langle x, y\rangle) = \mathrm{Im}(\langle x, y\rangle) \end{aligned}$$

29. Enunciado
    (a) Não existe porque para uma matriz simétrica, vetores próprios de valores próprios distintos têm de ser ortogonais e $\langle(1, -1, 1), (1, 1, -1)\rangle = -1 \neq 0$.
    (b) Sim, se um espaço próprio tiver dimensão > 1 haverá sempre vetores nesse espaço que não são ortogonais. Uma matriz simétrica que tem estes dois vetores próprios é por exemplo a matriz identidade.
    É um exercício instrutivo achar *todas* as matrizes que satisfazem estas condições. Pelo Teorema espetral, uma tal matriz A é diagonalizável. Se $A \neq I$ terá que ter um valor próprio real $\lambda \neq 1$. O espaço próprio $E(\lambda)$ tem de ser o complemento ortogonal

de $E(1) = L(\{(1,-1,1),(1,1,-1)\})$ e é portanto (recorde que $EL(A)^{\perp} = N(A)$) o núcleo da matriz

$$\begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix} \to \begin{bmatrix} 1 & -1 & 1 \\ 0 & 2 & -2 \end{bmatrix}$$

Concluímos que $E(\lambda) = L(\{(0,1,1)\})$.

As condições acima já determinam completamente uma matriz A! Terá de ser

$$A = S \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \lambda \end{bmatrix} S^{-1}$$

com S uma matriz invertível que tem como primeiras duas colunas uma base de $E(1)$ e terceira coluna uma base de $E(\lambda)$. Não é no entanto claro porque é que uma tal matriz terá de ser simétrica. A razão é a seguinte: a fórmula acima é válida para qualquer matriz S satisfazendo as condições indicadas e podemos escolher para S uma matriz ortogonal, cujas colunas formam uma base ortonormada de vetores próprios de A. Então $S^{-1} = S^T$ e a fórmula para A produz manifestamente uma matriz simétrica: sendo D uma matriz diagonal,

$$\left(SDS^T\right)^T = (S^T)^T D^T S^T = SDS^T$$

Vamos então calcular a matriz A. Usando o método de Gram-Schmidt obtemos umas base ortonormada para $E(1)$:

$$(1,1,-1) - \operatorname{proj}_{(1,-1,1)}(1,1,-1) = (1,1,-1) - \tfrac{-1}{3}(1,-1,1) = (\tfrac{4}{3}, \tfrac{2}{3}, -\tfrac{2}{3})$$

logo uma base ortonormada para $E(1)$ é dada por $\{(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}), (\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}})\}$. Uma base ortonormada para $E(\lambda)$ é dada por $\{0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$. Logo as matrizes distintas da identidade que satisfazem as condições do enunciado são as matrizes

$$A = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \lambda \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 1 & \frac{2}{3} & -\frac{2}{3} \\ \frac{2}{3} & \frac{\lambda+1}{2} & \frac{\lambda-1}{2} \\ -\frac{2}{3} & \frac{\lambda-1}{2} & \frac{\lambda+1}{2} \end{bmatrix}$$

com $\lambda \in \mathbb{R} \setminus \{1\}$.

(c) Os vetores dados são ortogonais dois a dois logo conforme explicado na alínea anterior são os vetores próprios de uma matriz simétrica. Normalizando os vetores obtemos

$$A = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{bmatrix} = \frac{1}{3} \begin{bmatrix} -1 & -1 & 2 \\ -1 & 2 & -1 \\ 2 & -1 & -1 \end{bmatrix}$$

30. Enunciado

(a) A matriz simétrica associada a esta forma quadrática é $A = \begin{bmatrix} 3 & \frac{1}{2} \\ \frac{1}{2} & -1 \end{bmatrix}$. Uma vez que o determinante de uma matriz é o produto dos valores próprios (repetidos de acordo com a sua multiplicidade) e como neste caso temos $\det A = -\frac{13}{4} < 0$, podemos concluir que os dois valores próprios têm sinais contrários e portanto a forma quadrática é indefinida.

(b) A matriz simétrica associada a esta forma quadrática é $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, que tem determinante igual a zero e traço igual a 2. Como o traço de uma matriz é a soma dos valores próprios (repetidos de acordo com a a sua multiplicidade), concluimos que um dos valores próprios é 0 e o outro é 2, portanto a forma quadrática é semi-definida positiva.

(c) A matriz simétrica associada a esta forma quadrática é $A = \begin{bmatrix} -1 & \frac{3}{2} \\ \frac{3}{2} & -3 \end{bmatrix}$ e como $\det A = \frac{3}{4} > 0$ e $\operatorname{tr} A = -4 < 0$, podemos concluir que os valores próprios são ambos negativos, ou seja, a forma quadrática é definida negativa.

(d) Esta forma quadrática é dada pela expressão $f(x, y) = x^2+4xy+y^2$. Para a analisar temos que usar a matriz simétrica associada à forma quadrática (e não a matriz que aparece no enunciado, que nem sequer é diagonalizável). A matriz simétrica associada é $\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$. Como o determinante desta matriz é negativo, concluimos que a forma quadrática é indefinida, pois os valores próprios têm sinais contrários.

(e) A matriz simétrica associada a esta forma quadrática é $A = \begin{bmatrix} 1 & 1 & \frac{1}{2} \\ 1 & -1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 1 \end{bmatrix}$. Temos $\det A = -\frac{3}{2}$, logo podem ser 3 valores próprios negativos ou um negativo e dois positivos. No entanto como o traço da matriz é igual a 1, não podem ser os 3 negativos. Concluimos que há valores próprios com sinais contrários e portanto a forma quadrática é indefinida.

31. Enunciado

(a) A forma quadrática $2x^2 + xy + 3y^2$ corresponde à matriz simétrica $\begin{bmatrix} 2 & \frac{1}{2} \\ \frac{1}{2} & 3 \end{bmatrix}$. O polinómio característico é $(2 - \lambda)(3 - \lambda) - \frac{1}{4} = \lambda^2 - 5\lambda + \frac{23}{4}$. Os valores próprios são portanto

$$\lambda = \frac{5 \pm \sqrt{2}}{2}$$

Os vetores próprios correspondentes são por exemplo $(1, 1 \pm \sqrt{2})$. Se (u, v) forem as coordenadas numa base ortonormada formada por vetores próprios ou seja

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{4+2\sqrt{2}}} & \frac{1}{\sqrt{4-2\sqrt{2}}} \\ \frac{1+\sqrt{2}}{\sqrt{4+2\sqrt{2}}} & \frac{1-\sqrt{2}}{\sqrt{4-2\sqrt{2}}} \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}$$

então

$$2x^2 + xy + 3y^2 = 5 \Leftrightarrow \tfrac{5+\sqrt{2}}{2}u^2 + \tfrac{5-\sqrt{2}}{2}v^2 = 5 \Leftrightarrow \left(\frac{u}{\sqrt{\frac{10}{5+\sqrt{2}}}}\right)^2 + \left(\frac{v}{\sqrt{\frac{10}{5-\sqrt{2}}}}\right)^2 = 1$$

logo a curva é uma elipse com eixos nas direções $(1, 1 \pm \sqrt{2})$ com eixos de comprimento $\sqrt{\frac{10}{5\pm\sqrt{2}}}$ respetivamente.

(b) Neste caso, sendo a expressão tão simples, não é sequer necessário achar os valores e vetores próprios da forma quadrática $x^2 + 2xy + y^2$ (esse procedimento levaria exatamente ao mesmo resultado). O termo quadrátiico pode escrever-se $(x + y)^2$ logo fazendo a mudança de coordenadas ortogonal

$$\begin{cases} u = \frac{x+y}{\sqrt{2}} \\ v = \frac{x-y}{\sqrt{2}} \end{cases} \Leftrightarrow \begin{cases} x = \frac{u-v}{\sqrt{2}} \\ y = \frac{u+v}{\sqrt{2}} \end{cases}$$

obtemos a seguinte expressão para a curva:

$$2u^2 + \sqrt{2}v = 1 \Leftrightarrow v = \frac{1}{\sqrt{2}} - \sqrt{2}u^2$$

Com respeito aos eixos u, v nas direções $(1, 1)$ e $(-1, 1)$ respetivamente, a curva é portanto uma parábola voltada para baixo.

(c) A matriz simétrica associada à forma quadrática $x^2 + 4xy + y^2$ é

$$\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$$

Os valores próprios são as soluções de $(1 - \lambda)^2 - 4 = 0 \Leftrightarrow \lambda = 1 \pm 2 \Leftrightarrow \lambda = -1, 3$. Um vetores próprio de -1 é por exemplo $(1, -1)$ e um vetor próprio de 3 é $(1, 1)$. Nas coordenadas (u, v) da base ortonormada de vetores próprios, definidas pelo sistema

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}$$

temos

$$x^2 + 4xy + y^2 = 1 \Leftrightarrow -u^2 + 3v^2 = 1 \Leftrightarrow v = \pm\sqrt{\frac{1 + u^2}{3}}$$

que é uma hipérbole orientada segundo os eixos gerados por $(1, 1)$ e $(1, -1)$ e que tem por assíntotas as retas $v = \pm\frac{u}{\sqrt{3}}$.

32. Enunciado

(a) Uma base ortonormada adequada a esta rotação é

$$B = ((\tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}}, 0), (\tfrac{1}{\sqrt{2}}, -\tfrac{1}{\sqrt{2}}, 0), (0, 0, 1))$$

O primeiro vetor gera o eixo de rotação e será portanto um vetor próprio de 1. Se o sentido da rotação é o horário quando observado de $(10, 10, 0)$ o terceiro vetor da base terá que girar um ângulo de $\frac{\pi}{6}$ afastando-se do segundo vetor da base. Na base B obtemos assim a seguinte expressão para a transformação desejada:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\frac{\pi}{6} & -\operatorname{sen}\frac{\pi}{6} \\ 0 & \operatorname{sen}\frac{\pi}{6} & \cos\frac{\pi}{6} \end{bmatrix}$$

A matriz que representa a rotação na base canónica é portanto

$$\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ 0 & \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

(b) A reflexão no plano ortogonal ao vetor $(0,1,-1)$ tem esse plano como espaço próprio de 1 e o vetor $(0,1,-1)$ como vetor próprio de -1 logo é dada pela expressão

$$\begin{bmatrix} 0 & 0 & 1 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Alternativamente podíamos ter notado que o plano de reflexão é o plano $y = z$ e portanto a reflexão troca as coordenadas y e z e mantém a coordenada x.
A rotação do enunciado é descrita na base canónica pela matriz

$$\begin{bmatrix} \cos\frac{\pi}{4} & -\operatorname{sen}\frac{\pi}{4} & 0 \\ \operatorname{sen}\frac{\pi}{4} & \cos\frac{\pi}{4} & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

A matriz que representa a transformação pedida é portanto

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix}$$

33. Enunciado Se R é uma rotação de um ângulo α em torno de um eixo, então numa base ortonormada em que o primeiro vetor aponta na direção do eixo, a expressão de R é

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\operatorname{sen}\theta \\ 0 & \operatorname{sen}\theta & \cos\theta \end{bmatrix}$$

com $\theta = \pm\alpha$ (o sinal depende da orientação da base escolhida) e portanto os valores próprios de R são 1 e $\cos\alpha \pm i \operatorname{sen}\alpha$. O vetor próprio de 1 é a direção do eixo de rotação.

O vetor próprio de 1 para a matriz do enunciado é um elemento não nulo do núcleo de

$$\begin{bmatrix} -\frac{1}{2} + \frac{1}{2\sqrt{2}} & -\frac{1}{2} + \frac{1}{2\sqrt{2}} & -\frac{1}{2} \\ -\frac{1}{2} + \frac{1}{2\sqrt{2}} & -\frac{1}{2} + \frac{1}{2\sqrt{2}} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{\sqrt{2}} - 1 \end{bmatrix} \to \begin{bmatrix} -\frac{1}{2} + \frac{1}{2\sqrt{2}} & -\frac{1}{2} + \frac{1}{2\sqrt{2}} & -\frac{1}{2} \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

O núcleo é portanto $\{(a,-a,0)\colon a \in \mathbb{R}\}$ e constitui o eixo desta rotação.

Como parece complicado calcular o polinómio característico podemos em vez disso ver o efeito que a matriz tem no plano perpendicular ao eixo, que tem $(v_2, v_3) = ((\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0), (0,0,1))$ como base ortonormada. O efeito em $(0,0,1)$ é

$$(-\tfrac{1}{2}, -\tfrac{1}{2}, \tfrac{1}{\sqrt{2}}) = -\tfrac{1}{\sqrt{2}} v_2 + \tfrac{1}{\sqrt{2}} v_3 = \cos(\tfrac{\pi}{4}) v_3 - \operatorname{sen}(\tfrac{\pi}{4}) v_2$$

logo a matriz representa uma rotação de um ângulo de $\frac{\pi}{4}$ em torno do eixo gerado por $(1,-1,0)$. Note-se que o sentido da rotação é o anti-horário quando olhamos para o plano de rotação que passa pela origem de um ponto no eixo dos xx.

34. Enunciado
   (a) Seja $B=(v_1,\ldots,v_n)$ uma base ortonormada para V. Supondo que a adjunta T^* existe temos, para todo o $w\in W$,
   $$\langle v_i, T^*w\rangle = \langle Tv_i, w\rangle$$
   Uma vez que $\langle v_i, x\rangle$ é a i-ésima coordenada de x na base ortonormada B conclui-se que
   $$T^*(w) = \langle Tv_1, w\rangle v_1 + \ldots + \langle Tv_n, w\rangle v_n$$
   que é a fórmula do enunciado. Isto mostra que, se existir, a adjunta é única. Resta verificar que a fórmula anterior (que pela linearidade do produto interno na segunda variável define de facto uma transformação linear) é a adjunta de T: dado $v\in V$ podemos escrever $v=\alpha_1v_1+\ldots+\alpha_nv_n$, com $\alpha_i=\langle v_i, v\rangle$. Então
   $$\begin{aligned}\langle v, T^*w\rangle &= \left\langle v, \sum_{i=1}^n \langle Tv_i, w\rangle v_i\right\rangle = \sum_{i=1}^n \langle Tv_i, w\rangle\langle v, v_i\rangle\\ &= \sum_{i=1}^n \langle Tv_i, w\rangle\overline{\alpha_i} = \sum_{i=1}^n \langle \alpha_i Tv_i, w\rangle\\ &= \left\langle T\left(\sum_{i=1}^n \alpha_i v_i\right), w\right\rangle = \langle Tv, w\rangle\end{aligned}$$
   conforme pretendido.
   (b) De facto temos
   $$\langle Tv, w\rangle = \overline{(Av)}^T w = \overline{v}^T A^* w = \langle v, A^*w\rangle$$
   pelo que a transformação linear $T^*(w)=A^*w$ é a adjunta de T.
   (c) Recorde-se que o isomorfismo entre V (ou $\overline{V}$ no caso complexo) envia $v\in V$ no funcional $\phi_v(w)=\langle v,w\rangle$. Escrevendo $T^\vee\colon W^*\to V^*$ para a aplicação dual a T (que é definida por $T^\vee(\psi)=\psi\circ T$), a afirmação que queremos demonstrar é
   $$\phi_{T^*(w)} = T^\vee(\phi_w) \quad \text{para todo o } w\in W$$
   para o que basta avaliar estes dois elementos de V^* num elemento $v\in V$:
   $$\phi_{T^*(w)}(v) = \langle T^*(w), v\rangle = \overline{\langle v, T^*(w)\rangle} = \overline{\langle T(v), w\rangle} = \langle w, T(v)\rangle = \phi_w(T(v)) = (T^\vee(\phi_w))(v)$$
   (d) Para calcular a transformação adjunta, iremos determinar a matriz que representa a transformação adjunta com respeito à base canónica de W,
   $$B_c = \left(\begin{bmatrix}1&0\\0&0\end{bmatrix}, \begin{bmatrix}0&1\\0&0\end{bmatrix}, \begin{bmatrix}0&0\\1&0\end{bmatrix}, \begin{bmatrix}0&0\\0&1\end{bmatrix}\right)$$

e à base $B = (t(t+1), 1-t^2, t(1-t))$ de V (vimos no Exercício 4(b) que B é uma base ortogonal de V). Recordemos que a transformação adjunta é a única transformação linear que verifica a igualdade $\langle Tv, w\rangle = \langle v, T^*w\rangle$. Em coordenadas esta igualdade escreve-se como

$$[Tv]_{B_c}^T\, G_{B_c}\, [w]_{B_c} = [v]_B^T\, G_B\, [T^*w]_B,$$

onde G_{B_c} é a matriz da métrica em W, com respeito à base B_c e G_B é a matriz da métrica em V com respeito à base B. Sendo A_{T,B,B_c} e $A_{T^*,B_c,B}$ as matrizes que representam as transformações T e T^*, respectivamente, nestas bases, obtemos

$$(A_{T,B,B_c}[v]_B)^T\, G_{B_c}\, [w]_{B_c} = [v]_B^T\, G_B\, A_{T^*,B_c,B}[w]_{B_c} \Leftrightarrow$$

$$[v]_B^T\, (A_{T,B,B_c})^T\, G_{B_c}\, [w]_{B_c} = [v]_B^T\, G_B\, A_{T^*,B_c,B}[w]_{B_c},$$

donde podemos concluir que

$$(A_{T,B,B_c})^T\, G_{B_c} = G_B\, A_{T^*,B_c,B} \Leftrightarrow A_{T^*,B_c,B} = G_B^{-1}(A_{T,B,B_c})^T\, G_{B_c}.$$

Para obter a matriz da métrica G_B falta apenas calcular as normas ao quadrado dos vetores da base, que nos vão dar as entradas na diagonal (as outras entradas são zero porque a base é ortogonal). Calculando essas normas obtemos $G_B = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{bmatrix}$.

Para obter a matriz da métrica G_{B_c} é necessário calcular o produto interno entre as matrizes da base B_c. É fácil de verificar que a base é ortonormada, por isso $G_{B_c} = I$. Para obter a matriz que representa T calculamos a imagem dos vetores da base B:

$$T(t(1+t)) = \begin{bmatrix} 2 & 2 \\ 0 & 6 \end{bmatrix}, \quad T(1-t^2) = \begin{bmatrix} 0 & 1 \\ 0 & -3 \end{bmatrix} \quad \text{e} \quad T(t(1-t)) = \begin{bmatrix} 0 & 0 \\ -2 & -2 \end{bmatrix}$$

portanto $A_{T,B,B_c} = \begin{bmatrix} 2 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & -2 \\ 6 & -3 & -2 \end{bmatrix}$. Finalmente obtemos

$$A_{T^*,B_c,B} = G_B^{-1}(A_{T,B,B_c})^T = \begin{bmatrix} \frac{1}{4} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{4} \end{bmatrix} \begin{bmatrix} 2 & 2 & 0 & 6 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & -2 & -2 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & \frac{3}{2} \\ 0 & 1 & 0 & -3 \\ 0 & 0 & -\frac{1}{2} & -\frac{1}{2} \end{bmatrix}$$

As coordenadas de $T^*\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right)$ na base B obtêm-se multiplicando a última matriz pelo vetor com as coordenadas (a, b, c, d). Portanto

$$T^*\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = (a+b+3d)\frac{t(1+t)}{2} + (b-3d)(1-t^2) + (c+d)\frac{t(t-1)}{2}.$$

Alternativamente, podemos usar a fórmula que usámos para justificar a existência e unicidade da transfomação adjunta: Como $(\frac{t}{2}(t+1), 1-t^2, \frac{t}{2}(t-1))$ é uma base ortonormada de V, a transformação adjunta é dada por

$$\begin{aligned} T^*(\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]) &= \langle T\left(\tfrac{t}{2}(t+1)\right), \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]\rangle \tfrac{t}{2}(t+1) + \langle T\left(1-t^2\right), \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]\rangle (1-t^2) + \\ &\quad \langle T\left(\tfrac{t}{2}(t-1)\right), \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]\rangle \tfrac{t}{2}(t-1) \\ &= \langle \left[\begin{smallmatrix} 1 & 1 \\ 0 & 3 \end{smallmatrix}\right], \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]\rangle \tfrac{t}{2}(t+1) + \langle \left[\begin{smallmatrix} 0 & 1 \\ 0 & -3 \end{smallmatrix}\right], \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]\rangle (1-t^2) + \langle \left[\begin{smallmatrix} 0 & 0 \\ 1 & 1 \end{smallmatrix}\right], \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]\rangle \tfrac{t}{2}(t-1) \\ &= (a+b+3d)\tfrac{t}{2}(t+1) + (b-3d)(1-t^2) + (c+d)\tfrac{t}{2}(t-1) \end{aligned}$$

35. Enunciado
   (a) Começamos por ver a unicidade. Suponhamos que B é uma matriz simétrica tal que $B^2 = A$. Pelo Teorema Espetral existem uma matriz ortogonal S e uma matriz diagonal Λ tal que
   $$B = S\Lambda S^T$$
   Recorde-se que as colunas de S formam uma base ortonormada para $\mathbb{R}^n$ (pois S é ortogonal) formada por vetores próprios de B e que Λ tem como entradas diagonais os valores próprios associados às colunas correspondentes de S (é esse o significado da equação equivalente $BS = S\Lambda$).
   Uma vez que $A = B^2 = S\Lambda S^T S\Lambda S^T = S\Lambda^2 S^T$ a matriz S também diagonaliza A e os valores próprios de A são os quadrados dos valores próprios de B (logo não negativos). Seja $E^X(\lambda)$ o espaço próprio de λ para a matriz X. Se admitirmos que os valores próprios de B são não negativos então $E^A(\mu) = E^B(\sqrt{\mu})$ - ambos os espaços são gerados pelas colunas de S correspondentes às de Λ^2 que têm μ na diagonal ou, equivalentemente, às colunas de Λ que têm $\sqrt{\mu}$ na diagonal.
   Dado que A é diagonalizável, e portanto $\oplus_\mu E^A(\mu) = \mathbb{R}^n$, a igualdade $E^A(\mu) = E^B(\sqrt{\mu})$ determina completamente B - ela identifica os valores que B toma numa base de $\mathbb{R}^n$ (uma base formada por vetores próprios de A). Conclui-se que, se existir, a matriz B é única.
   Para ver a existência basta notar que a matriz identificada na demonstração da unicidade é de facto uma solução do problema. Pelo Teorema espetral temos $A = SDS^T$ com S ortogonal e D diagonal contendo os valores próprios (não negativos) de A. Seja Λ a matriz diagonal com entradas não negativas tais que $\Lambda^2 = D$ e $B = S\Lambda S^T$. Então
   - B é simétrica: $B^T = (SDS^T)^T = (S^T)^T D^T S^T = SDS^T = B$,
   - os valores próprios de B são não negativos pois são as entradas diagonais de Λ,
   - $B^2 = S\Lambda S^T S\Lambda S = S\Lambda^2 S^T = SDS^T = A$.

   (b) Seja A é uma matriz simétrica com todos os valores próprios positivos. Pela alínea anterior existe uma matriz simétrica B com todos os valores próprios positivos (as raízes quadradas dos valores próprios de A), e portanto invertível, tal que $A = B^2$. Como B é simétrica isto significa que $A = B^T B$. Dado que o produto de matrizes calcula o produto interno usual das linhas da matriz à esquerda pelas colunas da matriz à direita, A é a matriz da métrica do produto interno usual na base para $\mathbb{R}^n$ formada pelas colunas da matriz B (note-se que as colunas de B formam uma base para $\mathbb{R}^n$ porque a matriz B é invertível).

(c) O polinómio característico de $\left[\begin{smallmatrix}2 & 1\\ 1 & 2\end{smallmatrix}\right]$ é $(2-\lambda)^2-1$ logo os valores próprios são $2\pm 1 = 1, 3$. Os vetores próprios de 1 são os múltiplos não nulos de $(1,-1)$ e os de 3 os múltiplos não nulos de $(1,1)$. Portanto

$$\begin{bmatrix}2 & 1\\ 1 & 2\end{bmatrix} = \begin{bmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & 3\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\end{bmatrix}$$

e

$$\sqrt{\begin{bmatrix}2 & 1\\ 1 & 2\end{bmatrix}} = \begin{bmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & \sqrt{3}\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\end{bmatrix} = \begin{bmatrix}\frac{1+\sqrt{3}}{2} & \frac{-1+\sqrt{3}}{2}\\ \frac{-1+\sqrt{3}}{2} & \frac{1+\sqrt{3}}{2}\end{bmatrix}$$

36. Enunciado Uma vez que A é hermitiana, pelo Teorema Espetral é diagonalizável, ou seja, $A = SDS^{-1}$ onde S é uma matriz unitária e D é uma matriz diagonal com os valores próprios de A na diagonal. Segue que $A^2 + A = SD^2S^{-1} + SDS^{-1} = S(D^2+D)S^{-1}$. Se $D = \begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix}$ então $D^2 + D = \begin{bmatrix}\lambda_1^2+\lambda_1 & 0\\ 0 & \lambda_2^2+\lambda_2\end{bmatrix}$.

A equação do enunciado diz-nos então que

$$S\begin{bmatrix}\lambda_1^2+\lambda_1 & 0\\ 0 & \lambda_2^2+\lambda_2\end{bmatrix}S^{-1} = \begin{bmatrix}1 & -i\\ i & 1\end{bmatrix}$$

logo os valores próprios da matriz $B = \begin{bmatrix}1 & -i\\ i & 1\end{bmatrix}$ têm de ser $\{\lambda_1^2+\lambda_1, \lambda_2^2+\lambda_2\}$ e as colunas correspondentes de S têm de ser vetores próprios de B associados a estes valores próprios.

Os valores próprios de B são 0 e 2 (as raízes do polinómio de $(1-\lambda)^2 - 1 = 0$). Os vetores próprios de 0 são $\{\alpha(i,1) : \alpha \neq 0\}$ e os vetores próprios de 2 são $\{\alpha(1,i) : \alpha \neq 0\}$, logo, assumindo sem perda de generalidade que a ordem dos valores próprios na diagonal é $0, 2$ (a ordem não irá afetar as matrizes resultantes) temos $S = \frac{1}{\sqrt{2}}\begin{bmatrix}i & 1\\ 1 & i\end{bmatrix}$. A igualdade $S(D^2+D)S^{-1} = S\Lambda S^{-1}$ implica que $\begin{bmatrix}\lambda_1^2+\lambda_1 & 0\\ 0 & \lambda_2^2+\lambda_2\end{bmatrix} = \begin{bmatrix}0 & 0\\ 0 & 2\end{bmatrix}$, portanto resolvendo o sistema seguinte obtemos os valores possíveis para as entradas na diagonal de D.

$$\begin{cases}\lambda_1^2+\lambda_1 = 0\\ \lambda_2^2+\lambda_2 = 2\end{cases} \iff \begin{cases}\lambda_1 = 0 \ \vee\ \lambda_1 = -1\\ \lambda_2 = 1 \ \vee\ \lambda_2 = -2\end{cases}$$

Temos assim 4 possibilidades para a matriz D e portanto 4 soluções para a $A = SDS^{-1}$ (como S é unitária $S^{-1} = S^*$):

$$\begin{aligned}
A &= \frac{1}{\sqrt{2}}\begin{bmatrix} i & 1 \\ 1 & i \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} -i & 1 \\ 1 & -i \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & -i \\ i & 1 \end{bmatrix} \\
A &= \frac{1}{\sqrt{2}}\begin{bmatrix} i & 1 \\ 1 & i \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 0 & -2 \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} -i & 1 \\ 1 & -i \end{bmatrix} = \begin{bmatrix} -1 & i \\ -i & -1 \end{bmatrix} \\
A &= \frac{1}{\sqrt{2}}\begin{bmatrix} i & 1 \\ 1 & i \end{bmatrix}\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} -i & 1 \\ 1 & -i \end{bmatrix} = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \\
A &= \frac{1}{\sqrt{2}}\begin{bmatrix} i & 1 \\ 1 & i \end{bmatrix}\begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} -i & 1 \\ 1 & -i \end{bmatrix} = \frac{1}{2}\begin{bmatrix} -3 & i \\ -i & -3 \end{bmatrix}
\end{aligned}$$

37. Enunciado
   (a) Sendo A a matriz do enunciado temos

$$A^*A = \begin{bmatrix} 2-i & 1-2i \\ 1-2i & 2-i \end{bmatrix}\begin{bmatrix} 2+i & 1+2i \\ 1+2i & 2+i \end{bmatrix} = \begin{bmatrix} 10 & 8 \\ 8 & 10 \end{bmatrix}$$

Os valores próprios são as soluções de $(10-\lambda)^2 - 64 = 0 \Leftrightarrow \lambda = 2, 18$. Os valores próprios são por exemplo $(1,-1)$ e $(1,1)$ respetivamente. Portanto

$$\begin{bmatrix} 10 & 8 \\ 8 & 10 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 0 & 18 \end{bmatrix}\begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

e portanto a parte hermitiana da decomposição polar é

$$P = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}\begin{bmatrix} \sqrt{2} & 0 \\ 0 & 3\sqrt{2} \end{bmatrix}\begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 2\sqrt{2} & \sqrt{2} \\ \sqrt{2} & 2\sqrt{2} \end{bmatrix}$$

A parte unitária é então

$$U = AP^{-1} = \begin{bmatrix} 2+i & 1+2i \\ 1+2i & 2+i \end{bmatrix}\frac{1}{6}\begin{bmatrix} 2\sqrt{2} & -\sqrt{2} \\ -\sqrt{2} & 2\sqrt{2} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

   (b) Temos

$$A^TA = \begin{bmatrix} a^2+b^2 & 0 \\ 0 & a^2+b^2 \end{bmatrix}$$

que já é diagonal, logo a parte simétrica da decomposição polar é

$$P = \begin{bmatrix} \sqrt{a^2+b^2} & 0 \\ 0 & \sqrt{a^2+b^2} \end{bmatrix}$$

e a parte unitária é

$$U = AP^{-1} = \begin{bmatrix} \frac{a}{\sqrt{a^2+b^2}} & -\frac{b}{\sqrt{a^2+b^2}} \\ \frac{b}{\sqrt{a^2+b^2}} & \frac{a}{\sqrt{a^2+b^2}} \end{bmatrix}$$

Escrevendo $(\frac{a}{\sqrt{a^2+b^2}}, \frac{b}{\sqrt{a^2+b^2}}) = (\cos\theta, \operatorname{sen}\theta)$ e $r = \sqrt{a^2+b^2}$ vemos que esta decomposição corresponde precisamente à decomposição polar $(a+bi) = r(\cos\theta + i\operatorname{sen}\theta)$ do número complexo $a+bi$.

38. Enunciado

(a) Temos

$$A^TA = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}$$

cujos valores próprios são as soluções de $(5-\lambda)^2 - 16 = 0 \Leftrightarrow \lambda = 5 \pm 4 = 9, 1$. Os vetores próprios são por exemplo $(1,1)$ e $(1,-1)$ respetivamente. Conclui-se que

$$A^TA = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 9 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

Os valores singulares de A são portanto $\sigma_1 = 3$ e $\sigma_2 = 1$ e a matriz U_2 é dada por

$$U_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

Uma vez que

$$AU_2^{-1} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \frac{3}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{3}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

a matriz U_1 que é obtida normalizando as colunas de AU_2^{-1} é

$$\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

A decomposição SVD da matriz dada é portanto

$$\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

(b) Temos

$$A^TA = \begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

cujos valores próprios são as soluções de $(2-\lambda)^2 - 1 = 0 \Leftrightarrow \lambda = 2 \pm 1 = 3, 1$. Os vetores próprios são por exemplo $(1,-1)$ e $(1,1)$ respetivamente. Conclui-se que

$$A^TA = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

Os valores singulares de A são portanto $\sigma_1 = \sqrt{3}$ e $\sigma_2 = 1$ e a matriz U_2 é dada por

$$U_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

Uma vez que

$$AU_2^{-1} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 2 & 0 \\ -1 & 1 \\ 1 & 1 \end{bmatrix}$$

Um vetor perpendicular ao espaço das colunas de A é por exemplo $(1, 1, -1)$ logo a matriz U_1 que se obtém normalizando as colunas não nulas de AU_2^{-1} e completando a uma base ortonormada de $\mathbb{R}^3$ é, por exemplo

$$\begin{bmatrix} \frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \end{bmatrix}$$

A decomposição SVD da matriz dada é portanto

$$\begin{bmatrix} 1 & -1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

(c) Temos

$$A^T A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

O polinómio característico desta matriz é

$$\begin{aligned} \begin{vmatrix} 1-\lambda & 0 & 1 & 0 \\ 0 & 1-\lambda & 0 & 1 \\ 1 & 0 & 1-\lambda & 0 \\ 0 & 1 & 0 & 1-\lambda \end{vmatrix} &= (1-\lambda) \begin{vmatrix} 1-\lambda & 0 & 1 \\ 0 & 1-\lambda & 0 \\ 1 & 0 & 1-\lambda \end{vmatrix} + \begin{vmatrix} 0 & 1-\lambda & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1-\lambda \end{vmatrix} \\ &= (1-\lambda)((1-\lambda)^3 - (1-\lambda)) - ((1-\lambda)^2 - 1) \\ &= ((1-\lambda)^2 - 1)^2 = \lambda^2(\lambda - 2)^2 \end{aligned}$$

Logo os valores próprios são 0 e 2. O núcleo da matriz $A^T A$ é $\{(a, b, -a, -b) \colon a, b \in \mathbb{R}\}$ logo uma base ortonormada para este espaço próprio de 0 é

$$\{(\tfrac{1}{\sqrt{2}}, 0, -\tfrac{1}{\sqrt{2}}, 0), (0, \tfrac{1}{\sqrt{2}}, 0, -\tfrac{1}{\sqrt{2}})\}$$

O espaço próprio de 2 é o núcleo de

$$\begin{bmatrix} -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix}$$

que é igual a $\{(a,b,a,b)\colon a,b\in\mathbb{R}\}$ logo uma base ortonormada para este espaço próprio de 0 é

$$\{(\tfrac{1}{\sqrt{2}},0,\tfrac{1}{\sqrt{2}},0),(0,\tfrac{1}{\sqrt{2}},0,\tfrac{1}{\sqrt{2}})\}$$

Conclui-se que

$$A^TA=\begin{bmatrix}\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 0\\ 0 & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} & 0\\ 0 & \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}}\end{bmatrix}\begin{bmatrix}2&0&0&0\\0&2&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 0\\ 0 & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} & 0\\ 0 & \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}}\end{bmatrix}$$

Os valores singulares de A são portanto $\sigma_1=\sigma_2=\sqrt{2}$ e a matriz U_2 é dada por

$$U_2=\begin{bmatrix}\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 0\\ 0 & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} & 0\\ 0 & \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}}\end{bmatrix}$$

Uma vez que

$$AU_2^{-1}=\begin{bmatrix}1&0&1&0\\0&1&0&1\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 0\\ 0 & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} & 0\\ 0 & \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}}\end{bmatrix}=\begin{bmatrix}\sqrt{2}&0&0&0\\0&\sqrt{2}&0&0\end{bmatrix}$$

a matriz U_1 que é obtida normalizando as colunas de AU_2^{-1} é portanto a matriz identidade 2×2 e a decomposição SVD da matriz do enunciado é então

$$\begin{bmatrix}1&0&1&0\\0&1&0&1\end{bmatrix}=\begin{bmatrix}1&0\\0&1\end{bmatrix}\begin{bmatrix}\sqrt{2}&0&0&0\\0&\sqrt{2}&0&0\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 0\\ 0 & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} & 0\\ 0 & \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}}\end{bmatrix}$$

39. Enunciado Uma vez que U_1 e U_2 são invertíveis, a característica de A é igual à característica de D que é k (note-se que a matriz "diagonal" D está em escada de linhas).

    A equação $A=U_1DU_2$ é equivalente a $AU_2^{-1}=U_1D$ e esta última descreve o efeito de A numa base ortonormada para $\mathbb{R}^n$ dada pelas colunas de $U_2^{-1}=U_2^T$: os primeiros k vetores da base são enviados respetivamente para os vetores $\sigma_1u_1,\ldots,\sigma_ku_k$ onde $u_1,\ldots,u_k$ são as primeiras k colunas de U_1, sendo os restantes enviados para zero. Vemos assim que $u_1,\ldots,u_k$ são uma base ortonormada para o espaço das colunas de A. As restantes colunas de U_1 são portanto uma base ortonormada para $EC(A)^\perp=N(A^T)$.

    A última afirmação é uma consequência da segunda pois $U_2^TD^TU_1^T$ é uma decomposição em valores singulares para A^T. Logo as primeiras k colunas de U_2^T, que são as primeiras k linhas de U_2, são uma base ortonormada para $EC(A^T)=EL(A)$ enquanto que as restantes linhas de U_2 são uma base ortonormada para $EL(A)^\perp=N(A)$.
40. Enunciado

(a) Temos $\det(A^TA) = 1 \Leftrightarrow \det(A^T)\det(A) = 1 \Leftrightarrow \det(A)^2 = 1$. Logo $\det(A) = \pm 1$.

(b) Seja
$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in O(2)$$
A equação $A^TA = I$ diz que os vetores (a,c) e (b,d) de $\mathbb{R}^2$ têm comprimento 1 e são perpendiculares. Escrevendo $(a,c) = (\cos\alpha, \operatorname{sen}\alpha)$ para algum $\alpha \in [0, 2\pi[$ temos portanto
$$(b,d) = \pm(-\operatorname{sen}\alpha, \cos\alpha)$$
Se $(b,d) = (-\operatorname{sen}\alpha, \cos\alpha)$ então $\det A = 1$ logo $A \in SO(2)$, caso contrário o determinante é -1 e $A \in O(2)_-$.

(c) Para verificar que um elemento $A \in O(2)_-$ é a reflexão na reta indicada basta calcular o efeito que tem num vetor da reta e num vetor não colinear.
$A\left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right] = \left[\begin{smallmatrix} \cos\alpha \\ \operatorname{sen}\alpha \end{smallmatrix}\right]$ que é de facto a reflexão de $(1,0)$ com respeito à reta que bissecta o ângulo α com o eixo dos xx.
Por outro lado
$$\begin{bmatrix} \cos\alpha & \operatorname{sen}\alpha \\ \operatorname{sen}\alpha & -\cos\alpha \end{bmatrix} \begin{bmatrix} \cos\frac{\alpha}{2} \\ \operatorname{sen}\frac{\alpha}{2} \end{bmatrix} = \begin{bmatrix} \cos\alpha\cos\frac{\alpha}{2} + \operatorname{sen}\alpha\operatorname{sen}\frac{\alpha}{2} \\ \operatorname{sen}\alpha\cos\frac{\alpha}{2} - \cos\alpha\operatorname{sen}\frac{\alpha}{2} \end{bmatrix} = \begin{bmatrix} \cos(\alpha - \frac{\alpha}{2}) \\ \operatorname{sen}(\alpha - \frac{\alpha}{2}) \end{bmatrix}$$
onde na última igualdade usámos as fórmulas trigonométricas para a diferença de ângulos. Conclui-se que a matriz A fixa a reta gerada por $(\cos\frac{\alpha}{2}, \operatorname{sen}\frac{\alpha}{2})$ e é portanto a reflexão com respeito a esta reta.
Seja E_α a reflexão na reta gerada por $(\cos\alpha, \operatorname{sen}\alpha)$ e consideremos o elemento $E_0 = \left[\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right]$. Dada $A = \left[\begin{smallmatrix} \cos\alpha & \operatorname{sen}\alpha \\ \operatorname{sen}\alpha & -\cos\alpha \end{smallmatrix}\right] \in SO(2)$ e tendo em conta o parágrafo anterior temos
$$A = E_{\frac{\alpha}{2}}E_0$$
logo A é o produto de duas reflexões.

(d) Vamos escrever $T \colon \mathbb{R}^3 \to \mathbb{R}^3$ para o endomorfismo representado por A na base canónica.
O polinómio característico de $A \in O(3)$ tem grau 3. As suas raízes são os valores próprios de A que são complexos de módulo 1. Uma vez que as raízes complexas deste polinómio real aparecem aos pares conjugados, o polinómio tem pelo menos uma raíz real (o que também se pode ver observando o comportamento do polinómio em $\pm\infty$), raíz esta que será necessariamente ± 1. Temos duas possibilidades para as restantes raízes: ou são complexas e conjugadas da forma $\cos\alpha \pm \operatorname{sen}\alpha$ ou são reais e iguais a ± 1.
Consideremos primeiro o caso em que todas as raízes λ_1, λ_2 e λ_3 são reais. Se $\det A = \lambda_1\lambda_2\lambda_3 = -1$ então pelo menos uma das raízes é igual a -1 (e as restantes duas são iguais). Escolhendo uma base ortonormada $B = (v_1, v_2, v_3)$ de vetores próprios de A com v_1 um vetor próprio de -1 vemos que
$$A_{T,B,B} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & s & 0 \\ 0 & 0 & s \end{bmatrix} \quad \text{com } s = \pm 1$$

tem a forma indicada no enunciado (com $\alpha = 0$ ou π). Se $\det A = 1$ então uma das raízes é necessariamente igual a 1 e procedendo da mesma forma vemos que

$$A_{T,B,B} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & s & 0 \\ 0 & 0 & s \end{bmatrix} \qquad \text{com } s = \pm 1$$

como pretendido.
Suponhamos agora que A tem um único valor próprio real $\lambda = \pm 1$ que é então necessariamente igual a $\det A$. Seja v_1 um vetor próprio de λ com $\|v_1\| = 1$ e $P = \{v\}^\perp$. Seja $\{v_2, v_3\}$ uma base ortonormada para P. Então, uma vez que T é ortogonal, o plano P é invariante por T e (com a restrição do produto interno usual em $\mathbb{R}^3$) a restrição de T a P é uma transformação ortogonal. Uma vez que a expressão para o produto interno na base ortonormada (v_2, v_3) para P é a usual, a matriz que representa a restrição de T a P nesta base é uma matriz de $O(2)$. Conclui-se que na base ortonormada $B = (v_1, v_2, v_3)$ a expressão de T é diagonal por blocos da forma

$$A_{T,B,B} = \begin{bmatrix} \lambda & 0 \\ 0 & R \end{bmatrix} \qquad \text{com } R \in O(2)$$

Temos $\det(A) = \det(T) = \lambda \det R = \det A \det R$, logo $\det R = 1$, isto é $R \in SO(2)$. $A_{T,B,B}$ tem portanto a forma indicada no enunciado.

(e) Argumentamos por indução em n notando que a afirmação foi demonstrada para $n = 2, 3$ nas alíneas (c) e (d) (e é evidente para $n = 1$).
Suponhamos primeiro que o polinómio característico de A tem apenas raízes reais (necessariamente iguais a ± 1). Este caso pode ser tratado diretamente sem recurso à indução: existe uma base ortonormada de $\mathbb{R}^n$ com respeito à qual a expressão de A é diagonal com entradas diagonais ± 1. Seja k a multiplicidade geométrica de -1 de forma que $\det(A) = (-1)^k$. Para obter uma expressão para A da forma desejada (observe-se que $\left[\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right] \in SO(2)$ - é uma rotação de um ângulo π). tomamos

- Vetores próprios de -1 para os primeiros k vetores da base quando k e n são ambos pares ou k e n são ambos ímpares,
- Se n é par e k ímpar, tomamos para primeiro vetor da base um vetor próprio de -1, para segundo um vetor próprio de 1 pondo depois $k-1$ vetores próprios de -1 e $n-k-1$ vetores próprios de 1.
- Se n é ímpar e k par, tomamos para primeiro vetor da base um vetor próprio de 1, os restantes k de -1 e os restantes $n-k-1$ de 1.

Seja então $n \geq 4$ e suponhamos que o polinómio característico de A tem pelo menos uma raíz complexa $\cos\alpha + i \operatorname{sen}\alpha$. Seja $v \in \mathbb{C}^n$ um vetor próprio de $\cos\alpha + i \operatorname{sen}\alpha$. Escrevendo $v = x + iy$ com $x, y \in \mathbb{R}^n$ temos a seguinte identidade em $\mathbb{C}^n$:

$$A(x+iy) = (\cos\alpha + i \operatorname{sen}\alpha)(x+iy) \Leftrightarrow Ax + iAy = \cos(\alpha)x - \operatorname{sen}(\alpha)y + i(\operatorname{sen}(\alpha)x + \cos(\alpha)y)$$

ou seja temos as seguintes duas condições em $\mathbb{R}^n$:

$$\begin{cases} Ax = \cos(\alpha)x - \operatorname{sen}(\alpha)y \\ Ay = \operatorname{sen}(\alpha)x + \cos(\alpha)y \end{cases} \tag{84}$$

Em particular o espaço $L(\{x,y\}) \subset \mathbb{R}^n$ é invariante por A.
Note-se agora que $x - iy$ é um vetor próprio de A correspondendo ao vetor próprio $\cos\alpha - i\operatorname{sen}\alpha$ e vetores próprios de valores próprios distintos são ortogonais. Logo para o produto interno usual em $\mathbb{C}^n$ temos $\langle x+iy, x-iy\rangle = 0$. Podemos escrever esta condição em termos do produto interno usual em $\mathbb{R}^n$:

$$\langle x,x\rangle - \langle y,y\rangle + i(-2\langle x,y\rangle) = 0 \Leftrightarrow \begin{cases} \|x\|^2 = \|y\|^2 \\ \langle x,y\rangle = 0 \end{cases}$$

Conclui-se que $\{x,y\}$ é uma base ortogonal para $L(\{x,y\})$ formada por vetores com o mesmo comprimento ℓ. Na base ortonormada $(\frac{x}{\ell}, \frac{y}{\ell})$ a expressão da restrição de A a este plano é de acordo com (84) dada pela expressão

$$\begin{bmatrix} \cos\alpha & \operatorname{sen}\alpha \\ -\operatorname{sen}\alpha & \cos\alpha \end{bmatrix} \in SO(2)$$

Podemos então tomar $\frac{x}{\ell}, \frac{y}{\ell}$ como os dois últimos vetores da base para $\mathbb{R}^n$ que buscamos. O complemento ortogonal $W = L(\{\frac{x}{\ell}, \frac{y}{\ell}\})^\perp$ é um subespaço invariante de $\mathbb{R}^n$ de dimensão $n-2$. Por hipótese de indução existe uma base ortonormada para W na qual A se expressa como indicado no enunciado e juntando a esta base os dois vetores $\frac{x}{\ell}, \frac{y}{\ell}$ (como os dois últimos vetores da base) obtemos a expressão pretendida para A.

(f) Esta afirmação é uma consequência da alínea anterior e da alínea (c): pela alínea (c) cada bloco diagonal 2×2 corresponde a um produto de no máximo duas reflexões e, no caso em que n é ímpar, a entrada ± 1 corresponde ao produto com uma reflexão se o sinal for negativo e com a identidade se for positivo.

41. Enunciado

(a) Seja $B = (v_1, \ldots, v_n)$ uma base ortonormada para V e $M = \max\{\|T(v_i)\| : i = 1, \ldots, n\}$. Dado $v = \alpha_1 v_1 + \ldots + \alpha_n v_n \in V$ temos (pelo Teorema de Pitágoras)

$$\|v\|^2 = |\alpha_1|^2 + \ldots + |\alpha_n|^2$$

e em particular $|\alpha_i| \leq \|v\|$. Por outro lado aplicando a desigualdade triangular vemos que

$$\begin{aligned} \|T(v)\| &= \|\alpha_1 T(v_1) + \ldots + \alpha_n T(v_n)\| \\ &\leq |\alpha_1|\|T(v_1)\| + \ldots + |\alpha_n|\|T(v_n)\| \\ &\leq (|\alpha_1| + \ldots + |\alpha_n|)M \\ &\leq n\|v\|M \end{aligned}$$

pelo que nM é um majorante para o conjunto $\{\frac{\|T(v)\|}{\|v\|} : v \in V \setminus \{0\}\}$.

(b) Se $v=0$ a desigualdade é obviamente verificada e, para $v \neq 0$ é uma consequência de $\|T\|$ ser um majorante do conjunto $\frac{\|T(v)\|}{\|v\|}$.

(c) **Positividade:** Claramente $\|T\| \geq 0$ sendo o supremo de um conjunto de números não negativos. Se $\|T\|=0$ então, para $v \neq 0$, os números não negativos $\frac{\|T(v)\|}{\|v\|}$ tem que ser majorados por 0 e portanto iguais a 0. Isto significa que $T=0$ conforme pretendido.

Homogeneidade: Dado $\alpha \in \mathbb{K}$ temos $(\alpha T)(v) = \alpha T(v)$ logo

$$\|\alpha T\| = \sup_{v \in V \setminus \{0\}} \frac{\|(\alpha T)(v)\|}{\|v\|} = \sup_{v \in V \setminus \{0\}} \frac{|\alpha|\|T(v)\|}{\|v\|} = |\alpha| \sup_{v \in V \setminus \{0\}} \frac{\|T(v)\|}{\|v\|} = |\alpha|\|T\|$$

Desigualdade triangular: Dados endomorfismos $S, T\colon V \to V$ temos

$$\|S+T\| = \sup_{v \in V \setminus \{0\}} \frac{\|T(v)+S(v)\|}{\|v\|}$$

Pela desigualdade triangular (para vetores) $\|T(v)+S(v)\| \leq \|T(v)\| + \|S(v)\|$, e pela alínea anterior $\|T(v)\| + \|S(v)\| \leq \|T\|\|v\| + \|S\|\|v\|$. Logo

$$\frac{\|T(v)+S(v)\|}{\|v\|} \leq \|T\| + \|S\| \quad \text{para todo o } v \neq 0$$

Conclui-se que $\|T+S\| \leq \|T\| + \|S\|$ conforme pretendíamos demonstrar.

(d) Seja $B=(v_1, \ldots, v_n)$ uma base ortonormada formada por vetores próprios de T com $T(v_i) = \lambda_i v_i$. Seja $M = \max\{|\lambda_1|, \ldots, |\lambda_n|\}$. Dado $v = \alpha_1 v_1 + \ldots + \alpha_n v_n$ temos

$$\begin{aligned}
\|T(v)\|^2 &= \|\alpha_1 T(v_1) + \ldots + \alpha_n T(v_n)\|^2 = \|\alpha_1 \lambda_1 v_1 + \ldots + \alpha_n \lambda_n v_n\|^2 \\
&= |\alpha_1 \lambda_1|^2 \|v_1\|^2 + \ldots + |\alpha_n \lambda_n|^2 \|v_n\|^2 \qquad \text{(pelo Teorema de Pitágoras)} \\
&= |\alpha_1|^2 |\lambda_1|^2 + \ldots + |\alpha_n|^2 |\lambda_n|^2 \\
&\leq (|\alpha_1|^2 + \ldots + |\alpha_n|^2) M^2 \\
&= M^2 \|v\|^2 \qquad \text{(porque } B \text{ é uma base ortonormada).}
\end{aligned}$$

Daqui concluímos que $\|T\| \leq M$. Mas por outro lado, sendo j tal que $|\lambda_j| = \max\{|\lambda_1|, \ldots, |\lambda_n|\}$ temos $M = \frac{\|T(v_j)\|}{\|v_j\|}$ logo $M \leq \|T\|$ e portanto $M = \|T\|$.

(e) Note-se que, por definição de transformação adjunta (ver o Exercício 34)

$$\|T(v)\|^2 = \langle T(v), T(v) \rangle = \langle T^*T(v), v \rangle$$

Seja $B = (v_1, \ldots, v_n)$ uma base ortonormada para V formada por vetores próprios da transformação auto-adjunta T^*T e $\lambda_1, \ldots, \lambda_n$ os valores próprios correspondentes (que recorde-se são todos não negativos). Suponhamos que j é tal que $\lambda_j = \max\{\lambda_1, \ldots, \lambda_n\}$. Pela alínea anterior temos $\|T^*T\| = \lambda_j$ e este número é, por definição de valores singulares, o quadrado do maior valor singular de T.

Escrevendo $v = \alpha_1 v_1 + \ldots + \alpha_n v_n$ temos

$$\begin{aligned}\|T(v)\|^2 &= \langle T^*T(v), v\rangle \\ &= \langle \alpha_1\lambda_1 v_1 + \ldots + \alpha_n\lambda_n v_n, \alpha_1 v_1 + \ldots + \alpha_n v_n\rangle \\ &= |\alpha_1|^2\lambda_1 + \ldots + |\alpha_n|^2\lambda_n \\ &\leq (|\alpha_1|^2 + \ldots + |\alpha_n|^2)\lambda_j = \lambda_j\|v\|^2\end{aligned}$$

donde concluímos que $\|T\| \leq \sqrt{\lambda_j}$. Por outro lado, tomando $v = v_j$ temos $\frac{\|T(v_j)\|^2}{\|v_j\|^2} = \lambda_j$ logo $\|T\| \geq \sqrt{\lambda_j}$, pelo que

$$\|T\| = \sqrt{\lambda_j} = \sqrt{\|T^*T\|}$$

conforme pretendido.

(f) Temos

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}^T \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$

O polinómio característico desta matriz é $(1-\lambda)(2-\lambda) - 1 = \lambda^2 - 3\lambda + 1$. Os valores próprios são $\lambda = \frac{3\pm\sqrt{5}}{2}$ logo, de acordo com a alínea anterior

$$\left\| \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \right\| = \sqrt{\frac{3+\sqrt{5}}{2}}$$

(g) Seja $B = (v_1, \ldots, v_n)$ uma base ortonormada. Recorde-se que a i-ésima coordenada do vetor v é dada pela expressão $\langle v_i, v\rangle$. Temos a mostrar que para cada i a série numérica

$$\sum_{n=0}^{\infty} \frac{1}{n!}\langle v_i, T^n(v)\rangle \tag{85}$$

converge. Dado que

$$|\langle v_i, T^n(v)\rangle\| \leq \|v_i\|\|T^n(v)\| = \|T^n(v)\| \leq \|T\|^n\|v\|$$

(onde na última desigualdade usámos a alínea (b)), a série numérica (85) converge absolutamente (podemos por exemplo aplicar o critério de d'Alembert).

Para verificar que a expressão é linear em v, basta verificar que é linear em cada coordenada e isso é uma consequência da linearidade do limite (de sucessões numéricas):

$$\begin{aligned}
\sum_{n=0}^{\infty} \frac{1}{n!}\langle v_i, T^n(\alpha v + \beta w)\rangle &= \lim_{N\to\infty} \sum_{n=0}^{\mathbb{N}} \frac{1}{n!}\langle v_i, T^n(\alpha v + \beta w)\rangle \\
&= \lim_{N\to\infty} \sum_{n=0}^{\mathbb{N}} \frac{1}{n!}\left(\alpha\langle v_i, T^n(v)\rangle + \beta\langle v_i, T^n(w)\rangle\right) \\
&= \alpha \lim_{N\to\infty} \sum_{n=0}^{\mathbb{N}} \frac{1}{n!}\langle v_i, T^n(v)\rangle + \beta \lim_{N\to\infty} \sum_{n=0}^{\mathbb{N}} \frac{1}{n!}\langle v_i, T^n(w)\rangle \\
&= \alpha \sum_{n=0}^{\infty} \frac{1}{n!}\langle v_i, T^n(v)\rangle + \beta \sum_{n=0}^{\infty} \frac{1}{n!}\langle v_i, T^n(w)\rangle
\end{aligned}$$

Finamente para $T = \left[\begin{smallmatrix} 3t & t \\ 0 & 3t \end{smallmatrix}\right]$ temos

$$\begin{aligned}
T^n &= \left(\begin{bmatrix} 3t & 0 \\ 0 & 3t \end{bmatrix} + \begin{bmatrix} 0 & t \\ 0 & 0 \end{bmatrix}\right)^n \\
&= \begin{bmatrix} 3t & 0 \\ 0 & 3t \end{bmatrix}^n + n \begin{bmatrix} 3t & 0 \\ 0 & 3t \end{bmatrix}^{n-1} \begin{bmatrix} 0 & t \\ 0 & 0 \end{bmatrix} + 0 \\
&= \begin{bmatrix} 3^n t^n & 3^{n-1} n t^n \\ 0 & 3^n t^n \end{bmatrix}
\end{aligned}$$

onde aplicámos o binómio de Newton (o que podemos fazer porque o primeiro termo da soma é um múltiplo da identidade e portanto comuta com o segundo termo) e usámos a identidade $\left[\begin{smallmatrix} 0 & t \\ 0 & 0 \end{smallmatrix}\right]^2 = 0$ que faz com que todos os termos do binómio a partir do segundo se anulem.
Dado que $\sum_{n=0}^{\infty} \frac{3^n t^n}{n!} = e^{3t}$ e $\sum_{n=0}^{\infty} \frac{n 3^{n-1} t^n}{n!} = t \sum_{n=1}^{\infty} \frac{3^{n-1} t^{n-1}}{(n-1)!} = t e^{3t}$ concluímos que

$$e^{\left[\begin{smallmatrix} 3t & t \\ 0 & 3t \end{smallmatrix}\right]} = \begin{bmatrix} e^{3t} & t e^{3t} \\ 0 & e^{3t} \end{bmatrix}$$

42. Enunciado
   (a) Se A é unitária então $A^*A = I = AA^*$ logo A é normal. Se A é (anti-)hermitiana temos $A^*A = \pm A^2 = AA^*$ logo A é normal.
   (b) Uma vez que $U^{-1} = U^*$ temos

$$A^*A = (UDU^*)^*(UDU^*) = UD^*U^*UDU^* = UD^*DU^*$$

$$AA^* = UDU^*(UDU^*)^* = UDU^*UD^*U^* = UDD^*U^*$$

Sendo $D = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{bmatrix}$ temos

$$D^*D = DD^* = \begin{bmatrix} |\lambda_1|^2 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & |\lambda_n|^2 \end{bmatrix}$$

e portanto A é normal.

(c) AA^* é hermitiana porque $(AA^*)^* = (A^*)^*A^* = AA^*$. Se v é um vetor próprio de AA^* com valor próprio λ temos

$$\lambda\|v\|^2 = \langle v, \lambda v\rangle = \langle v, AA^*v\rangle = \langle A^*v, A^*v\rangle = \|A^*v\|^2$$

logo os valores próprios de AA^* são não negativos.

(d) Temos

$$A(AA^*) = A(A^*A) = (AA^*)A$$

logo A comuta com AA^*. Se v é um vetor próprio de AA^* então $A(AA^*)v = A\lambda v = \lambda Av$ portanto $(AA^*)Av = \lambda Av$ o que significa que Av é também um vetor próprio de AA^* com o mesmo valor próprio.

(e) Sejam $v, w \in E(\lambda)$. Temos de verificar que

$$\langle \tfrac{1}{\sqrt{\lambda}}Av, \tfrac{1}{\sqrt{\lambda}}Aw\rangle = \langle v, w\rangle$$

Ora como λ é real e portanto igual a $\overline{\lambda}$

$$\langle \tfrac{1}{\sqrt{\lambda}}Av, \tfrac{1}{\sqrt{\lambda}}Aw\rangle = \tfrac{1}{\lambda}\langle Av, Aw\rangle = \tfrac{1}{\lambda}\langle v, A^*Aw\rangle = \tfrac{1}{\lambda}\langle v, \lambda w\rangle = \langle v, w\rangle$$

(f) A equação $A = UDU^*$ significa que as colunas de U são uma base ortonormada para $\mathbb{C}^n$ formada por vetores próprios de A. Pelo Teorema espetral e a alínea anterior podemos escolher bases ortonormadas para cada $E(\lambda)$ (o espaço próprio de λ para $AA^* = A^*A$) formada por vetores próprios de A. Uma vez que espaços próprios de valores próprios distintos da matriz hermitiana AA^* são ortogonais, juntando estas bases para todos os $\lambda > 0$ obtermos uma base ortonormada para $N(AA^*)^\perp \subset \mathbb{C}^n$ formada por vetores próprios de A.
Mas $N(AA^*) = N(A)$ (como $AA^* = A^*A$, temos $N(AA^*) \supset N(A)$; por outro lado se $AA^*v = 0$ então $\|Av\|^2 = \langle A^*Av, v\rangle = 0$ logo $v \in N(A)$) e qualquer base ortonormada para $N(A)$ será formada por vetores próprios de A.
Juntando as bases ortonormadas de vetores próprios de A para $N(AA^*)^\perp$ e $N(A)$ obtemos a base para $\mathbb{C}^n$ que pretendíamos.

(g) Temos que verificar se $A^*A = AA^*$:

$$\begin{bmatrix} 2 & 3 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 2 & -1 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 13 & 4 \\ 4 & 5 \end{bmatrix} \neq \begin{bmatrix} 5 & 4 \\ 4 & 13 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} 2 & 3 \\ -1 & 2 \end{bmatrix}$$

logo a matriz (i) não é normal e portanto não é diagonalizável por uma matriz unitária.

Para a matriz (ii) temos

$$\begin{bmatrix}1&1&0\\0&1&1\\1&0&1\end{bmatrix}\begin{bmatrix}1&0&1\\1&1&0\\0&1&1\end{bmatrix}=\begin{bmatrix}2&1&1\\1&2&1\\1&1&2\end{bmatrix}=\begin{bmatrix}1&0&1\\1&1&0\\0&1&1\end{bmatrix}\begin{bmatrix}1&1&0\\0&1&1\\1&0&1\end{bmatrix}$$

logo a matriz (ii) é normal e portanto é diagonalizável por uma matriz unitária.

Referências

[Ax] S. Axler, *Linear Algebra Done Right*, Springer UTM (1997).

[D] E. Dias, *Álgebra Linear*, `https://www.math.tecnico.ulisboa.pt/~edias/TextosNet/ALbookfin_Net.pdf`

[FIS] S. Friedberg, A. Insel and L. Spence, *Linear Algebra* (4th edition), Pearson Education (2003).

[HK] K. Hoffman and R. Kunze, *Linear Algebra*, Prentice-Hall (1961)