O Teorema da Função Inversa

Recordemos que se $f : I \subset \mathbb{R} \to \mathbb{R}$ é uma função com derivada contínua, definida num intervalo I, e $f'(x) \neq 0$ para todo o $x \in I$, então f é injectiva, como consequência dos teoremas do Valor Intermédio e do Valor Médio: o primeiro garante que ou $f'(x) > 0,\ \forall x \in I$ ou $f'(x) < 0,\ \forall x \in I$ (uma vez que f' é contínua), enquanto que pelo segundo sabemos que, dados $x < y$ em I, se tem

$$f(y) - f(x) = f'(c)(y - x)$$

para algum $c \in]x, y[$; portanto f é estritamente crescente caso $f'(x) > 0,\ \forall x \in I$ e estritamente decrescente caso $f'(x) < 0,\ \forall x \in I$.
Além disso, nas condições indicadas, a imagem $f(I)$ é também um intervalo J e a inversa $f^{-1} : J \to I$ tem derivada contínua:

$$(f^{-1})'(y) = \frac{1}{f'(f^{-1}(y))}.$$

Observação 0.1 *De facto, a hipótese da derivada de f ser contínua não é necessária, pois a derivada de uma função definida num intervalo tem sempre a propriedade do Valor Intermédio, mesmo que tenha descontinuidades (este resultado é o **Teorema de Darboux**).*

Em particular, se $f : I \subset \mathbb{R} \to \mathbb{R}$ é uma função com derivada contínua, $x_0 \in I$ e $f'(x_0) \neq 0$, então existe um intervalo $I_0 \subset I$ contendo x_0 onde f' tem sempre o mesmo sinal e portanto a restrição $f_{|I_0}$ é injectiva, e a sua inversa tem derivada contínua.
O facto de $f'(x_0) \neq 0$ significa que a aproximação linear de f na vizinhança de x_0

$$x_0 + h \to f(x_0) + f'(x_0)h$$

é injectiva (ou seja, invertível) e portanto o resultado enunciado pode exprimir-se do seguinte modo: se f tem derivada contínua e a sua aproximação linear num ponto x_0 é injectiva, então f é igualmente injectiva numa vizinhança desse ponto e a inversa da sua restrição a essa vizinhança tem derivada contínua.

No caso de uma função $F : U \subset \mathbb{R}^n \to \mathbb{R}^n$ diferenciável, a aproximação linear na vizinhança de um ponto $x_0 \in U$

$$x_0 + h \to F(x_0) + DF(x_0)h$$

é injectiva se o determinante $\det(DF(x_0))$ da matriz jacobiana não se anular.
Se F for injectiva numa vizinhança de x_0 e a sua inversa F^{-1} for diferenciável, então tem-se necessariamente

$$DF^{-1}(F(x))DF(x) = I$$

onde I designa a matriz identidade, e portanto tem que se verificar $\det(DF(x)) \neq 0$ e

$$DF^{-1}(F(x)) = [DF(x)]^{-1}.$$

No entanto, como se verifica, considerando por exemplo

$$F(x,y) = (e^x \cos(y), e^x \sin(y)),$$

uma função pode ter aproximação linear injectiva em todos os pontos do seu domínio ($\mathbb{R}^2$ neste caso) e no entanto não ser injectiva. Ou seja, o primeiro resultado citado para funções de uma variável não se generaliza. Mas o resultado local vale em geral:

Teorema 0.2 ***Teorema da Função Inversa:*** *Se $F : U \subset \mathbb{R}^n \to \mathbb{R}^n$ é uma função de classe C^1 (isto é, tem derivadas parciais contínuas), x_0 é um ponto interior de U, e $\det(DF(x_0)) \neq 0$, então existe uma vizinhança aberta V de x_0 onde F é injectiva, $F(V)$ é aberto, e a inversa F^{-1} da restrição $F_{|V}$ é também de classe C^1 com derivada*

$$DF^{-1}(y) = \left(DF(F^{-1}(y))\right)^{-1}$$

para todo o $y \in F(V)$.

Observação 0.3 *Note-se que, para este resultado local, não podemos prescindir da hipótese da derivada ser contínua: por exemplo a função*

$$g(x) = \begin{cases} x + x^2 \sin(1/x) & \text{se } x \neq 0 \\ \\ 0 & \text{se } x = 0 \end{cases}$$

tem derivada para todo o $x \in \mathbb{R}$ e $g'(0) = 1$; no entanto, g não é injectiva em nenhuma vizinhança de 0.

Apresenta-se em seguida uma demonstração do Teorema da Função Inversa. Começamos por apresentar dois resultados que têm um papel importante nesta demonstração mas que são igualmente de grande interesse teórico e prático no estudo de funções de várias variáveis.

A igualdade do Teorema do Valor Médio não se generaliza para funções de mais do que uma variável, mas temos um resultado semelhante:

Teorema 0.4 ***Teorema do Valor Médio:*** *Seja $F : U \subset \mathbb{R}^n \to \mathbb{R}^n$ uma função diferenciável e $a, b \in U$ dois pontos tais que U contém o segmento*

$$a + t(b-a),\ t \in [0,1].$$

Então existe um ponto $c = a + t_0(b-a)$ tal que

$$\|F(b) - F(a)\| \leq \|DF(c)(b-a)\|.$$

Demonstração 0.5 *Definimos a função real de variável real*

$$\phi : [0,1] \to \mathbb{R}, \qquad \phi(t) =< F(a + t(b-a)), F(b) - F(a) >$$

onde $<,>$ designa o produto interno usual em $\mathbb{R}^n$.
ϕ é a composta de funções diferenciáveis:

$$t \to a + t(b-a) \to F(a+t(b-a)) \to < F(a+t(b-a)), F(b) - F(a) > .$$

Temos então

$$\phi'(t) = < DF(a+t(b-a))(b-a), F(b) - F(a) >;$$

por outro lado

$$\phi(1) - \phi(0) = < F(b) - F(a), F(b) - F(a) > = \|F(b) - F(a)\|^2$$

e o Teorema do Valor Médio para funções reais de variável real garante portanto que

$$\|F(b)-F(a)\|^2 = \phi'(t_0)(1-0) = < DF(a+t(b-a))(b-a), F(b)-F(a) > = < DF(c)(b-a), F(b)-F(a) >;$$

Mas pela desigualdade de Cauchy-Schwartz

$$| < DF(c)(b-a), F(b) - F(a) > | \leq \|DF(c)(b-a)\| \|F(b) - F(a)\|$$

donde se deduz a desigualdade pretendida.

Observação 0.6 *Suponhamos que para todo o $x \in U$ se verifica*

$$|\frac{\partial F_i}{\partial x_j}(x)| \leq K;$$

Podemos então obter a seguinte majoração em função de $\|b-a\|$:

$$\|DF(c)(b-a)\| = \sqrt{\sum_{i=1}^{n} \left(\sum_{j=1}^{n} \frac{\partial F_i}{\partial x_j}(c)(b_j - a_j) \right)^2} \leq \sqrt{\sum_{i=1}^{n} K^2 \left(\sum_{j=1}^{n} (b_j - a_j) \right)^2} = \sqrt{n} K |\sum_{j=1}^{n} (b_j - a_j)| \leq$$

$$\leq \sqrt{n} K \sum_{j=1}^{n} |b_j - a_j| \leq \sqrt{n} K n \max_j |b_j - a_j| \leq n^{3/2} K \|b-a\|.$$

O segundo resultado preliminar é o seguinte:

Proposição 0.7 ***Lema da Contracção:*** *Seja $B \subset \mathbb{R}^n$ um conjunto fechado e $h : B \to B$ uma função que satisfaz a condição seguinte: existe uma constante $0 < c < 1$ tal que*

$$\|h(y) - h(x)\| \leq c\|y - x\| \; \forall x, y \in B;$$

Então existe um e só um ponto $z \in B$ tal que $h(z) = z$ (ou seja z é ponto fixo de h).

Demonstração 0.8 *A demonstração é inteiramente construtiva: dado um ponto inicial qualquer $x \in B$, mostramos que a sucessão definida recursivamente pelas condições*

$$x_0 = x,\ x_{k+1} = h(x_k),$$

converge, e que o seu limite tem a propriedade desejada.
Para isso vamos mostrar que a sucessão x_k é uma sucessão de Cauchy, ou seja, que dado um $\epsilon > 0$ existe uma ordem a partir da qual a distância entre quaisquer dois termos da sucessão é menor que ϵ: começamos por verificar que, como para qualquer $k > 0$

$$\|x_{k+1} - x_k\| = \|h(x_k) - h(x_{k-1})\| \leq c\|x_k - x_{k-1}\|,$$

prova-se facilmente (por exemplo, por indução) que

$$\|x_{k+1} - x_k\| \leq c^k\|x_1 - x_0\|,\ \forall k \geq 0.$$

Temos então

$$\|x_{m+k} - x_m\| = \|\sum_{j=0}^{k-1}(x_{m+j+1} - x_{m+j})\| \leq \sum_{j=0}^{k-1}\|x_{m+j+1} - x_{m+j}\| \leq \sum_{j=0}^{k-1} c^{m+j}\|x_1 - x_0\|;$$

e usando a convergência da série geométrica de razão c,

$$\sum_{j=0}^{k-1} c^{m+j}\|x_1 - x_0\| = c^m(1 + c + \cdots + c^{k-1})\|x_1 - x_0\| \leq \frac{c^m}{1-c}\|x_1 - x_0\|.$$

Concluímos que a sucessão x_k é de facto uma sucessão de Cauchy: dado $\epsilon > 0$, escolhendo M tal que

$$\frac{c^M}{1-c}\|x_1 - x_0\| < \epsilon$$

temos que para quaisquer dois termos x_m, x_{m+k} com $m \geq M$, a distância entre eles é menor que ϵ:

$$\|x_{m+k} - x_m\| \leq \frac{c^m}{1-c}\|x_1 - x_0\| \leq \frac{c^M}{1-c}\|x_1 - x_0\| < \epsilon.$$

Portanto a sucessão converge para um ponto $z \in B$ uma vez que este conjunto é fechado. Como a condição do enunciado implica em especial que h é contínua, temos

$$h(z) = h(\lim_j x_k) = \lim_k h(x_k) = \lim_k x_{k+1} = z.$$

Além disso, se $w \in B$ fosse um outro ponto fixo de h chegaríamos a uma contradição:

$$\|w - z\| = \|h(w) - h(z)\| \leq c\|w - z\| < \|w - z\|.$$

Passamos então à demonstração do Teorema da Função Inversa:

Demonstração 0.9 *Começamos por notar que podemos supor que $DF(x_0) = I$ onde I designa a matriz identidade: de facto se F é uma função nas condições do Teorema e $DF(X_0) = A$, podemos considerar a função $F_1 = A^{-1} \circ F$ que está igualmente nessas condições e satisfaz a nossa hipótese adicional; evidentemente se F_1 é invertível numa vizinhança V de x_0, também $F = A \circ F_1$ o é e as restantes conclusões do Teorema aplicam-se igualmente.*
Seja V uma vizinhança de x_0 em que se possa aplica o Teorema do Valor Médio (uma bola aberta, por exemplo), satisfazendo as seguintes condições:

$$\det(DF(x)) \neq 0 \, \forall x \in V, \qquad \|(DF(x) - I)v\| \leq \frac{1}{2}\|v\|, \ \forall v \in \mathbb{R}^n.$$

Note-se que para a segunda condição basta, de acordo com a majoração estabelecida a seguir ao Teorema do Valor Médio, que

$$|\frac{\partial F_i}{\partial x_j}(x)| \leq \frac{1}{2n^{3/2}} \ \textit{ se } i \neq j, \ \textit{ e } |\frac{\partial F_i}{\partial x_i}(x) - 1| \leq \frac{1}{2n^{3/2}} \, \forall i,$$

o que se verifica numa bola de raio suficientemente pequeno com centro em x_0, uma vez que as derivadas parciais de F são contínuas. E, do mesmo modo, a primeira condição verifica-se numa bola de raio suficientemente pequeno com centro em x_0, uma vez que a função

$$x \to \det(DF(x))$$

é contínua.
A desigualdade do Teorema do Valor Médio, aplicada à função $F(x) - x$ dá-nos

$$\|(F(x) - x) - (F(y) - y)\| \leq \frac{1}{2}\|x - y\|;$$

mas por outro lado, como, em consequência da desigualdade triangular, $\|u - v\| \geq \|u\| - \|v\|$ para quaisquer vectores u e v,

$$\|(F(x) - x) - (F(y) - y)\| = \|y - x - (F(y) - F(x))\| \geq \|x - y\| - \|F(x) - F(y)\|;$$

juntando as duas desigualdades obtemos

$$\|F(x) - F(y)\| \geq \frac{1}{2}\|x - y\|$$

ou seja, F é injectiva em V. Designamos a partir daqui por F^{-1} a inversa da restrição $F_{|V}$.
A desigualdade anterior também se pode então escrever como

$$\|F^{-1}(z) - F^{-1}(w)\| \leq 2\|z - w\| \ \forall z, w \in F(V),$$

o que implica em particular que F^{-1} é contínua nesse conjunto.

Seja $x \in V$ e $z = F(x)$; vamos verificar que F^{-1} é diferenciável em z com derivada $[DF(x)]^{-1}$, ou seja, vamos mostrar que, para w numa vizinhança de z se tem

$$\lim_{w \to z} \frac{F^{-1}(w) - F^{-1}(z) - [DF(x)]^{-1}(w - z)}{\|w - z\|} = 0.$$

Temos, para qualquer y suficientemente próximo de x,

$$F(y) = F(x) + DF(x)(y-x) + R(y)), \quad com \ \lim_{y \to x} \frac{R(y)}{\|y-x\|} = 0;$$

dado w próximo de z temos $w = F(y)$ para algum $y \in V$ (ou seja, $F(V)$ é aberto, o que comprovaremos mais adiante) e então a igualdade anterior pode escrever-se

$$w - x = DF(x)(y-x) + R(y)$$

o que implica que

$$[DF(x)]^{-1}(w-z) = y - x + [DF(x)]^{-1}R(y)$$

ou seja

$$F^{-1}(w) = F^{-1}(z) + [DF(x)]^{-1}(w-z) - [DF(x)]^{-1}R(y);$$

se mostrarmos que a última parcela satisfaz a condição do resto, ou seja, que

$$\lim_{w \to z} \frac{[DF(x)]^{-1}R(y)}{\|w-z\|} = 0,$$

aquela igualdade mostra que F^{-1} é de facto diferenciável em z com $DF^{-1}(z) = [DF(x)]^{-1}$. Ora

$$\frac{[DF(x)]^{-1}R(y)}{\|w-z\|} = \frac{[DF(x)]^{-1}R(y)}{\|y-x\|} \frac{\|y-x\|}{\|w-z\|};$$

quando $w \to z$, temos $y \to x$ (F^{-1} é contínua, como já vimos) e, pela hipótese de diferenciabilidade de F, o primeiro factor converge para zero; e o segundo factor é limitado:

$$\frac{\|F^{-1}(w) - F^{-1}(z)\|}{\|w-z\|} \leq 2.$$

Podemos mesmo concluir que F^{-1} é de classe C^1: as derivadas parciais de F^{-1}, ou seja as entradas de $DF^{-1}(z) = [DF(x)]^{-1}(x)$, são funções contínuas das entradas de $DF(x)$, estas são funções contínuas de x e x é funcão contínua de z, ou seja, cada derivada parcial de F^{-1} no ponto z se obtém como resultado de uma composição de funções contínuas.

Resta verificar que $F(V)$ é aberto. Seja então $z = F(x)$ com $x \in V$. Como V é aberto, sabemos que existe $r > 0$ tal que a bola fechada de raio r centrada em x, que designamos $\overline{B_r(x)}$, está contida em V. O que vamos mostrar é que então a bola de raio $r/2$ centrada em z está contida na imagem pela função F daquela bola fechada, isto é

$$B_{r/2}(z) \subset F(\overline{B_r(x)}).$$

Para isso, dado um $w \in B_{r/2}(z)$ qualquer, consideramos a função auxiliar h definida em $\overline{B_r(x)}$ por

$$h(y) = w - F(y) + y;$$

e verificamos, por um lado, que

$$\|h(y)-x\| = \|w-F(y)+y-x\| = \|w-z+z-F(y)+y-x\| = \|w-z-(F(y)-y)+(F(x)-x)\| \leq$$

$$\leq \|w-z\| + \|(F(y)-y)-(F(x)-x)\|;$$

a primeira parcela é menor que $r/2$ por hipótese, enquanto que a segunda, pela aplicação do teorema do Valor Médio à função $F(x)-x$ feita atrás, é menor que $\frac{1}{2}\|y-x\| \leq r/2$; ou seja, $\|h(y)-x\| \leq r$ e portanto

$$h(\overline{B_r(x)}) \subset \overline{B_r(x)}.$$

Por outro lado,

$$\|h(y)-h(y')\| = \|(w-F(y)+y)-(w-F(y')+y')\| = \|(F(y')-y')-(F(y)-y)\| \leq \frac{1}{2}\|y'-y\|,$$

ou seja, h é uma contracção. Existe portanto um único ponto $y \in \overline{B_r(x)}$ tal que $h(y) = y$, isto é, temos $w - F(y) + y = y \Leftrightarrow w = F(y)$, o que confirma que $w \in F(\overline{B_r(x)})$.

Observação 0.10 *Se num certo ponto x_0 do domínio de F se tem $\det(DF(x_0)) = 0$, então, tal como no caso de uma função real de variável real, verifica-se uma de duas situações: ou F não é injectiva em nenhuma vizinhança de x_0, ou é injectiva mas a respectiva inversa não é diferenciável no ponto $F(x_0)$.*