Generalized Matrix Functions

Sónia Carvalho,Pedro J. Freitas

UL-IPL

Seminário CEAFEL

Contents

- Derivatives
  - Generalized matrix function
  - Formulas for higher order derivatives
  - ξ- symmetric tensor power
  - Derivatives for operators
  - Derivatives for matrices
- Norms
  - Known results
  - Case $G = S_m$
  - General case

Derivatives

Fréchet derivative

Let $A, X \in \mathbb{C}^{n\times n}$ and $f : \mathbb{C}^{n\times n} \to \mathbb{C}$. We define

$$Df(A)(X) := \left.\frac{d}{dt}\right|_{t=0} f(A+tX)$$

If $f(A) = \det A$, then we have the *Jacobi formula*

$$D\det(A)(X) = \operatorname{tr}(\operatorname{adj}(A)X),$$

where $\operatorname{adj}(A)$ is the transpose of the matrix of cofactors.

For the permanent, we have (R. Bhatia, P. Grover)

$$D\operatorname{per}(A)(X) = \operatorname{tr}(\operatorname{padj}(A)X)$$

where $\operatorname{padj}(A)$ is the transpose of the matrix of permanental cofactors.

Generalized matrix function

Let G be a subgroup of S_n and P a representation of G.
A character of G is a map $\xi : G \mapsto \mathbb{C}$ afforded by the representation P defined as

$$\xi(\sigma) = \operatorname{tr} P(\sigma).$$

Definition

Let $A \in \mathbb{C}^{n\times n}$ and ξ a character of the subgroup G. The generalized matrix function determined by ξ and G is

$$d_\xi^G(A) = \sum_{\sigma \in G} \xi(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)}.$$

This is a multinear map in the columns of the matrix A and a polynomial map in the matrix entries.

Generalized matrix function

Suppose ξ is a irreducible character.

$G = S_n$

- $d_\xi^G(A) = d_\xi(A)$ is the immanant of A.
- If $\xi = 1$ then $d_\xi^G(A) = \text{per}(A)$ the permanent of A.
- If $\xi = sgn$ then $d_\xi^G(A) = \det(A)$ the determinant of A.

Definitions

Suppose $k \leq n$, define:

- $\Gamma_{k,n}$: the set of all maps $\{1, \ldots, k\} \longrightarrow \{1, \ldots, n\}$.
- $G_{k,n}$: the set of increasing maps $\{1, \ldots, k\} \longrightarrow \{1, \ldots, n\}$.
- $Q_{k,n}$: the set of strictly increasing maps $\{1, \ldots, k\} \longrightarrow \{1, \ldots, n\}$.

For $f : \mathbb{C}^{n \times n} \to \mathbb{C}$, we define

$$D^k f(A)(X^1, \ldots, X^k) := \frac{\partial^k}{\partial t_1 \ldots \partial t_k}\Big|_{t_1 = \ldots = t_k = 0} f(A + t_1 X^1 + \ldots + t_k X^k).$$

If f is multilinear this derivative is the coefficient of $t_1 \ldots t_k$ in the polynomial $f(A + t_1 X^1 + \ldots + t_k X^k)$.

Formula (1) for the k-th derivative

Let $\alpha \in Q_{k,n}$. Define $A(\alpha; X^1, \ldots, X^k)$ as the matrix of order n obtained from A by replacing the $\alpha(j)$ column of A by the $\alpha(j)$ column of X^j.

First expression

For every $1 \leq k \leq n$,

$$D^k d_\xi^G(A)(X^1, \ldots, X^k) = \sum_{\sigma \in S_k} \sum_{\alpha \in Q_{k,n}} d_\xi^G A(\alpha; X^{\sigma(1)}, \ldots, X^{\sigma(k)}).$$

Already known for $d_\xi = \det, \mathrm{per}$ (R. Bhatia, T. Jain, P. Grover).

Formula (1), rewritten

Define the mixed generalized matrix function of $X^1, \dots, X^n$ as

$$\Delta_\xi^G(X^1, \dots, X^n) := \frac{1}{n!} \sum_{\sigma \in S_n} d_\xi^G(X_{[1]}^{\sigma(1)}, \dots, X_{[n]}^{\sigma(n)}).$$

where $Y_{[i]}$ is the i-th column of Y.

For $k < n$, we abbreviate $\Delta_\xi^G(A, \dots, A, X^1, \dots, X^k)$ by $\Delta_\xi^G(A; X^1, \dots, X^k)$.

First expression, rewritten

$$D^k d_\xi^G(A)(X^1, \dots, X^k) = \frac{n!}{(n-k)!} \Delta_\xi^G(A; X^1, \dots, X^k).$$

Derivatives of the ξ-symmetric power

The ξ-symmetric tensors

Let V be a finite-dimensional vector space with inner product and consider its m-fold tensor power $\otimes^m V$. Define the ξ symmetriser:

$$T(G,\xi) = \frac{\xi(\mathrm{id})}{|G|} \sum_{\sigma \in G} \xi(\sigma) P(\sigma),$$

where $P(\sigma)(v_1 \otimes \ldots \otimes v_m) = v_{\sigma^{-1}(1)} \otimes \ldots \otimes v_{\sigma^{-1}(m)}$.

The range of $T(G,\xi)$, denoted $V_\xi(G) \leq \bigotimes^m V$ is called the ξ-symmetric class of tensors.

- For $\xi \equiv 1$, we get symmetric tensors.
- For $\xi = \mathrm{sgn}$, we get anti-symmetric tensors.

$$v_1 * v_2 * \ldots * v_m = T(G,\xi)(v_1 \otimes v_2 \otimes \ldots \otimes v_m).$$

These vectors are called decomposable symmetrised tensors.

The ξ-symmetric tensor power of T

Given $T_1, \ldots T_m \in L(V)$, the space $V_\xi(G)$ is invariant for the operator

$$T_1 \tilde{\otimes} \cdots \tilde{\otimes} T_m := \frac{1}{m!} \sum_{\sigma \in S_m} T_{\sigma(1)} \otimes \cdots \otimes T_{\sigma(m)}$$

We denote its restriction to $V_\xi(G)$ as $T_1 * \cdots * T_m$.

For $T \in L(V)$, define *the ξ-symmetric tensor power of T* as

$$K_\xi^G(T) = (\otimes^m T)|_{V_\xi(G)} = \underbrace{T * T * \cdots * T}_{m\,\text{times}}$$

Now we establish a formula for the directional derivative of the map $K_\xi^G : L(V_\xi(G)) \to L(V_\xi(G))$.

Formula for the derivative of K_ξ^G

Derivative for operators $\otimes^m T$

$$D^k \otimes^m T(X_1, \ldots, X_k) = \frac{m!}{(m-k)!} \underbrace{T \tilde{\otimes} \cdots \tilde{\otimes} T}_{m-k \text{ times}} \tilde{\otimes} X_1 \tilde{\otimes} \cdots * \tilde{\otimes} X_k$$

Known for $\vee, \wedge$ (Bhatia, Grover, Jain) We have proved that:

Derivative for operators

$$D^k K_\xi^G(T)(X_1, \ldots, X_k) = \frac{m!}{(m-k)!} \underbrace{T * \cdots * T}_{m-k \text{ times}} * X_1 * \cdots * X_k$$

What about for matrices?

Bases

In general, $V_\xi(G)$ does not have an orthonormal basis formed by decomposable symmetrised tensors.

Recall that $\Gamma_{m,n}$ is the set of all maps $\{1,\ldots,m\} \to \{1,\ldots,n\}$. Take $\{e_1,\ldots,e_n\}$ an o.n. basis of V, and, for $\alpha \in \Gamma$, define

$$e^*_\alpha := e_{\alpha(1)} * \ldots * e_{\alpha(m)}.$$

Take

- $\mathcal{E}'$: the basis of $V_\xi(G)$ formed by elements of the from e^*_α, with indexing set $\widehat{\Delta}$:
$$\mathcal{B} = \{e^*_\alpha : \alpha \in \widehat{\Delta} \subseteq \Gamma_{m,n}\}.$$
- $\mathcal{E}$: the orthonormal basis of $V_\xi(G)$ obtained by applying the Gram-Schmidt process to $\mathcal{E}'$,
- B: the change of basis matrix from $\mathcal{E}$ to $\mathcal{E}'$ — does not depend on the original basis of V.

Formula for the derivative of $K_\xi^G(A)$

Derivative for operators

$$D^k K_\xi^G(T)(X_1,\dots,X_k) = \frac{m!}{(m-k)!} \underbrace{T*\cdots*T}_{m-k \text{ times}} * X_1 * \cdots * X_k$$

Using the bases, it is possible to define $K_\xi^G(A)$ for a matrix A.

Denote by $\text{mixgmm}_\xi^G(A; X^1,\dots,X^k)$ the matrix indexed by $\widehat{\Delta}$ whose (γ,δ) entry is $\Delta_\xi^G(A\,[\gamma|\delta]\,; X^1\,[\gamma|\delta]\,,\dots, X^k\,[\gamma|\delta])$.

Derivative for matrices

$$D^k K_\xi^G(A)(X^1,\dots,X^k) = \frac{\xi(id)m!}{|G|(m-k)!} B^* \,\text{mixgmm}_\xi(A; X^1,\dots,X^k)B.$$

Norms

Two results

Definition

The norm of a multilinear operator Φ is given by

$$\|\Phi\| = \sup_{\|X^1\|=\ldots=\|X^k\|=1} \|\Phi(X^1, \ldots, X^k)\|.$$

We want to estimate $\|D^k f(T)\|$.

1. When $G = S_m$ we calculate the exact value for the norm considering $f(T) = K_\xi(T)$.
2. When G is a subgroup of S_m we obtain an upper bound for the norm when $f(T) = K_\xi^G(T)$.

Known Results

Let $k \leq m \leq n$.

- $\|D \wedge^m T\| = p_{m-1}(\nu_1, \ldots, \nu_m)$;
- $\|D \vee^m T\| = m\|T\|^{m-1} = m{\nu_1}^{m-1}$;
- $\|D^k \otimes^m T\| = \dfrac{m!}{(m-k)!}\|T\|^{m-k} = \dfrac{m!}{(m-k)!}{\nu_1}^{m-k}$;
- $\|D^k \vee^m T\| = \dfrac{m!}{(m-k)!}\|T\|^{m-k} = \dfrac{m!}{(m-k)!}{\nu_1}^{m-k}$;
- $\|D^k \wedge^m T\| = k!p_{m-k}(\nu_1, \ldots, \nu_m)$;
- $\|DK_\xi(T)\| = \sum_{j=1}^{m} \prod_{\substack{i=1 \\ i \neq j}}^{m} \nu_{\omega(\xi)(i)}$.

where $p_{m-k}(x_1, \ldots, x_m)$ is the elementary symmetric polynomial of degree $m-k$.

Partitions

The map $K_\xi(T)$ has an associated irreducible character, and hence a partition of m.

For $x \in \mathbb{C}^n$ and $(\pi_1, \ldots, \pi_m) \vdash m$, $\alpha \in \Gamma_{m,n}$, define

$$x_\alpha = (x_{\alpha(1)}, \ldots, x_{\alpha(m)})$$

$$\omega(\pi) := (\underbrace{1, \ldots, 1}_{\pi_1 \text{ times}}, \underbrace{2, \ldots, 2}_{\pi_2 \text{ times}}, \ldots, \underbrace{l(\pi), \ldots, l(\pi)}_{\pi_{l(\pi)} \text{ times}}) \in \Gamma_{m,n}.$$

Let $\operatorname{Im}\alpha = \{i_1, \ldots, i_l\}$, with $|\alpha^{-1}(i_1)| \geq \ldots \geq |\alpha^{-1}(i_l)|$.

$$\mu(\alpha) := (|\alpha^{-1}(i_1)|, \ldots, |\alpha^{-1}(i_l)|)$$

is called the multiplicity partition of α.

We have $\mu(\omega(\pi)) = \pi$ and $e^*_\alpha \neq 0$ iff ξ majorizes $\mu(\alpha)$.

Main Theorems: $G = S_m$

Norm of $D^k K_\xi(T)$

$$\|D^k K_\xi(T)\| = k!\, p_{m-k}(\nu_{\omega(\xi)})$$

- For $\xi \equiv 1$, $\xi = (m, 0, \ldots, 0)$, $\omega(\xi) = (1, 1, \ldots, 1)$.
- For $\xi = \mathrm{sgn}$, $\xi = (1, 1, \ldots, 1)$ and $\omega(\xi) = (1, 2, \ldots, m)$.

Norm of the derivative of the immanant

$$\|D^k d_\xi(A)\| \leq k!\, p_{n-k}(\nu_{\omega(\xi)})$$

Done for $k = 1$ by R. Bhatia and J. Dias da Silva.

The inequality becomes equality for $\xi = \mathrm{sgn}$, and $d_\xi = \det$.

Steps of the proof

1. Let $T = UP$ the polar decomposition of the operator T. First we prove that $\|D^k K_\chi(T)\| = \|D^k K_\chi(P)\|$. (unitarily invariance)
2. The map $D^k K_\chi(P)$ is multilinear and positive then its norm is attained at $(I, I, \ldots, I)$.
3. We calculate the largest singular value of $D^k K_\chi(P)(I, I, \ldots, I)$ that equals the norm.

For step (2) we use the following theorem.(R. Bhatia and T. Jain)

Russo-Dye Multilinear version

Suppose $\Phi : M_n^k(\mathbb{C}) \longrightarrow M_l(\mathbb{C})$ is a positive multilinear operator. Then

$$\|\Phi\| = \|\Phi(I, I, \ldots, I)\|.$$

General Case

When G is any subgroup of S_m this relation between partitions and irreducible characters of G does not exist.

Definition

Suppose ξ is an irreducible character of G. The multilinearity partition of the character ξ, $\mathrm{MP}(\xi)$, is the least upper bound of the partitions π of m for which $(\xi, \chi_\pi)_G \neq 0$.

When $G = S_m$, the multilinearity partition is the partition usually associated with ξ.

Proposition

Suppose ξ is an irreducible character of G and let $\alpha \in \Gamma_{m,n}$. If $e^*_\alpha \neq 0$, then $\mu(\alpha) \preceq \mathrm{MP}(\xi)$.

Lemma

Let α, β be elements of $\widehat{\Delta} \cap G_{m,n}$ and π be a partition of m.

1. $\lambda(\alpha) \geq \lambda(\beta)$ if and only if α precedes β in the lexicographic order.
2. If $\mu(\alpha) \preceq \pi$ then $\omega(\pi)$ precedes α in the lexicographic order.

Main Theorem

Let ξ be an irreducible character of G. Consider the map $T \to K_\xi^G(T)$. Then

$$\|D^k K_\xi(T)\| \leq k!\, p_{m-k}(\nu_{\omega(\mathrm{MP}(\xi))})$$

where p_{m-k} is the symmetric polynomial of degree $m-k$ in m variables, $\nu_1 \geq \ldots \geq \nu_n$ are the singular values of T and $\mathrm{MP}(\xi)$ the multilinearity partition of ξ.

Bibliography

- R. Bhatia, J. A. Dias da Silva, Variation of induced linear operators, *Linear Algebra and its Applications*, 341 (2002) 391–402.
- R. Bhatia, S. Friedland, Variation of Grassman powers and spectra, *Linear Algebra and its Applications*, 40 (1981) 1–18.
- R. Bhatia, P. Grover, T. Jain, Derivatives of tensor powers and their norms, *Electronic Linear Algebra*.
- S. Carvalho, P. J. Freitas, The *k*-th derivatives of the immanant and the χ-symmetric power of an operator, *The Electronic Journal of Linear Algebra.*, 19 (2014).
- S. Carvalho, P. J. Freitas, The norm of the *k*-th derivative of the χ-symmetric power of an operator, *The Electronic Journal of Linear Algebra*,19,(2014).
- T. Jain, Derivatives for antisymmetric tensor powers and perturbation bounds *Linear Algebra and its Applications*, 435 (2011) 1111–1121.