Note Title 2025-03-21

Fréchet derivative

The Fréchet derivative of a matrix function $f(A)$ is the linear operator $L_{f,A} : \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$

$$f(A+H) = f(A) + L_{f,A}[H] + o(\|H\|)$$

e.g. $f(x) = x^2$

$$(A+H)^2 = A^2 + \underbrace{HA + AH}_{L_{f,A}[H]} + \underbrace{H^2}_{o(\|H\|)}$$

$$K \in \mathbb{C}^{n^2\times n^2} \qquad K = A^T \otimes I + I \otimes A$$

Theorem: for all $A, H \in \mathbb{C}^{n\times n}$

$$f\left(\begin{bmatrix} A & H \\ 0 & A \end{bmatrix}\right) = \begin{bmatrix} f(A) & L_{f,A}[H] \\ 0 & f(A) \end{bmatrix}$$

assuming f is defined and continuous on $\begin{bmatrix} A & H \\ 0 & A \end{bmatrix}$ and $L_{f,A}[H]$ exists.

(cfr. $f\left(\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}\right) = \begin{bmatrix} f(\lambda) & f'(\lambda) \\ 0 & f(\lambda) \end{bmatrix}$, this is a "block version").

Proof: we evaluate

$$f\left(\begin{bmatrix} A+\varepsilon H & H \\ 0 & A \end{bmatrix}\right)$$

and then we will let $\varepsilon \to 0$.

Idea: block diagonalization. We have

$$\begin{bmatrix} A+\varepsilon H & 0 \\ 0 & A \end{bmatrix} = \begin{bmatrix} I & -X \\ 0 & I \end{bmatrix} \begin{bmatrix} A+\varepsilon H & H \\ 0 & A \end{bmatrix} \begin{bmatrix} I & X \\ 0 & I \end{bmatrix}$$

if X solves the Sylvester equation

$$(A+\varepsilon H)X - XA + H = 0$$

$X = -\frac{1}{\varepsilon} I$ is a solution!

$$-(\cancel{A}+\varepsilon H)\frac{1}{\varepsilon}I + \cancel{\frac{1}{\varepsilon}I\cdot A} + H = -\varepsilon\frac{1}{\varepsilon}H + H = 0.$$

$$f\left(\begin{bmatrix} A+\varepsilon H & H \\ 0 & A \end{bmatrix}\right) = f\left(\begin{bmatrix} 1 & X \\ 0 & I \end{bmatrix}\begin{bmatrix} A+\varepsilon H & 0 \\ 0 & A \end{bmatrix}\begin{bmatrix} 1 & -X \\ 0 & 1 \end{bmatrix}\right)$$

$$= \begin{bmatrix} I & X \\ 0 & I \end{bmatrix} f\left(\begin{bmatrix} A+\varepsilon H & 0 \\ 0 & A \end{bmatrix}\right)\begin{bmatrix} 1 & -X \\ 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & X \\ 0 & 1 \end{bmatrix}\begin{bmatrix} f(A+\varepsilon H) & 0 \\ 0 & A \end{bmatrix}\begin{bmatrix} 1 & -X \\ 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & -\frac{1}{\varepsilon}1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} f(A+\varepsilon H) & 0 \\ 0 & f(A) \end{bmatrix}\begin{bmatrix} 1 & \frac{1}{\varepsilon}1 \\ 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} f(A+\varepsilon H) & \frac{1}{\varepsilon}\left(f(A+\varepsilon H) - f(A)\right) \\ 0 & f(A) \end{bmatrix}$$

If we let $\varepsilon \to 0$, we have

$$\begin{bmatrix} f(A) & L_{f,A}[H] \\ 0 & f(A) \end{bmatrix}$$

$\square$

What conditions do we need for $f\left(\begin{bmatrix} A & H \\ 0 & A \end{bmatrix}\right)$ to exist and be continuous?

$\Lambda\left(\begin{bmatrix} A & H \\ 0 & A \end{bmatrix}\right)$ is two copies of $\Lambda(A)$

The algebraic multiplicities satisfy $m_a\left(\lambda, \begin{bmatrix} A & H \\ 0 & A \end{bmatrix}\right) = 2m_a(\lambda, A)$

If f is differentiable $2m_a(\lambda_i, A)$ times in each $\lambda_i \in \Lambda(A)$,
then $f\left(\begin{bmatrix} A & H \\ 0 & A \end{bmatrix}\right)$ is defined

If f is <u>continuously</u> differentiable $2m_a(\lambda_i, A)$ times in a <u>neighborhood</u> of λ_i for each $\lambda_i \in \Lambda(A)$,
then $f\left(\begin{bmatrix} A & H \\ 0 & A \end{bmatrix}\right)$ exists and is continuous in its argument.

<u>Theorem</u>: if f is differentiable continuously $2m_a(\lambda_i, A)$ times in a neighborhood of each $\lambda_i \in \Lambda(A)$, then $L_{f,A}[H]$ exists for each H (f is Fréchet-differentiable in A).

<u>Proof</u>:

$$f\left(\begin{bmatrix} A & H \\ 0 & A \end{bmatrix}\right) = \lim_{\varepsilon \to 0} f\left(\begin{bmatrix} A+\varepsilon H & H \\ 0 & A \end{bmatrix}\right)$$

$$= \lim_{\varepsilon \to 0} \begin{bmatrix} f(A+\varepsilon H) & \boxed{\frac{1}{\varepsilon}\left(f(A+\varepsilon H) - f(A)\right)} \\ 0 & f(A) \end{bmatrix}$$

The limit in the (1,2) block exists for each $H \in \mathbb{C}^n$
So all the <u>directional</u> derivatives $\frac{1}{\varepsilon}\left(f(A+\varepsilon H) - f(A)\right)$ exist for $f(A)$.

If all directional derivatives exist and are continuous in a point, then f is differentiable, by a Analysis 2 result.

$$f\left(\begin{bmatrix} \lambda & 1 & \\ & \lambda & 1 \\ & & \lambda \end{bmatrix}\right) = \begin{bmatrix} f(\lambda) & f'(\lambda) & \frac{1}{2}f''(\lambda) \\ & f(\lambda) & f'(\lambda) \\ & & f(\lambda) \end{bmatrix}$$

$$f(x) = x^{5/2}$$

$$K_{abs} = \limsup_{\|\tilde{x}-x\|\to 0} \frac{\|f(\tilde{x})-f(x)\|}{\|\tilde{x}-x\|} = \|Jac_x f\|.$$

$$K_{rel} = \limsup_{\|\tilde{x}-x\|\to 0} \frac{\frac{\|f(\tilde{x})-f(x)\|}{\|f(x)\|}}{\frac{\|\tilde{x}-x\|}{\|x\|}} = \frac{\|Jac_x f\|}{\|f(x)\|} \cdot \|x\|.$$

For a matrix function, $\frac{\|L_{f,A}\|}{\|f(A)\|_F} \cdot \|A\|_F$ is the condition number. We use the natural norm

$$\|A\|_F = \|vec(A)\|_2$$

$\|K\|_2$ $\overset{= \sigma_{max}(K)}{}$ is the operator norm induced by this norm, i.e.

$$\max_{H\neq 0} \frac{\|L_f(A)[H]\|_F}{\|H\|_F} = \|L_{f,A}\|$$

It is not trivial to compute this norm, without having access to the full matrix K.

[Chapter 3 in Higham's book "Functions of matrices" gives some more strategies]

normest( )

<u>Theorem</u>: Let $L_{f,A}$ be the Fréchet derivative of f in A. Let $\lambda_1, \lambda_2 \dots \lambda_n$ be the eigenvalues of $A \in \mathbb{C}^{n\times n}$ (counted with multiplicity).

Then, the n^2 eigenvalues of $L_{f,A}$ are given by

$$f[\lambda_i, \lambda_j] = \begin{cases} \dfrac{f(\lambda_i) - f(\lambda_j)}{\lambda_i - \lambda_j} & \text{if } \lambda_i \neq \lambda_j \\ f'(\lambda_i) & \text{if } \lambda_i = \lambda_j \end{cases} \qquad i,j = 1, 2, \dots n$$

Proof: we replace f with a polynomial that interpolates f and its derivatives up to twice the algebraic multiplicity of each eigenvalue

$$p(A) = f(A) \qquad p\left(\begin{bmatrix} A & H \\ 0 & A \end{bmatrix}\right) = f\left(\begin{bmatrix} A & H \\ 0 & A \end{bmatrix}\right) \quad \forall H$$

$$\text{i.e.,} \qquad L_{p,A} = L_{f,A}.$$

$$p(x) = c_0 + c_1 x + c_2 x^2 + \dots + c_d x^d$$

$$p(A) = c_0 I + c_1 A + c_2 A^2 + \dots + c_d A^d$$

$$p(A+H) = c_0 I + c_1 (A+H) + c_2 (A+H)^2 + \dots + c_d (A+H)^d$$

$$= \underbrace{c_0 I + c_1 A + c_2 A^2 + \dots + c_d A^d}_{p(A)} + \underbrace{c_1 H + c_2 (AH + HA) + c_3 (A^2 H + AHA + HA^2) + \dots + c_d (A^{d-1} H + \dots + H A^{d-1})}_{L_{p,A}[H]} + o(\|H\|)$$

$$K = c_1 (I \otimes I) + c_2 (I \otimes A + A^T \otimes I) + c_3 (I \otimes A^2 + A^T \otimes A + (A^T)^2 \otimes I) + \dots + c_d (I \otimes A^{d-1} + \dots + (A^{d-1})^T \otimes I).$$

We take Schur forms $A = Q U Q^*$ $\qquad U =$ 

$$A^T = \hat{Q}\hat{U}\hat{Q}^*$$

$$K = (\hat{Q}\otimes Q)\Big(c_1 I\otimes I + c_2(I\otimes U + \hat{U}\otimes I) + c_3(1\otimes U^2 + \hat{U}\otimes U + \hat{U}^2\otimes 1)$$
$$+ \dots + c_d(1\otimes U^{d-1} + \dots + \hat{U}^{d-1}\otimes 1\Big)(\hat{Q}\otimes Q)^*$$

$$= (\hat{Q}\otimes Q)\underbrace{\left(\sum_{k=1}^{d} c_k \sum_{h=0}^{k-1} \hat{U}^{h}\otimes U^{k-1-h}\right)}_{\text{upper triangular}}(\hat{Q}\otimes Q)^*$$

We can read off eigenvalues on the diagonal of the inner factor: the term in position $i+n(j-1)$ (i,j) is

$$\boxed{\sum_{k=1}^{d} c_k \sum_{h=0}^{k-1} \lambda_i^{h}\lambda_j^{k-1-h}}$$

because $\operatorname{diag}(U) = \operatorname{diag}(\hat{U}) = (\lambda_1, \lambda_2, \dots \lambda_n)$

$$\begin{cases} = \displaystyle\sum_{k=1}^{d} c_k \frac{\lambda_i^k - \lambda_j^k}{\lambda_i - \lambda_j} = \frac{p(\lambda_i) - p(\lambda_j)}{\lambda_i - \lambda_j} & \text{if } \lambda_i \neq \lambda_j \\ = \displaystyle\sum_{k=1}^{d} c_k\, k\,\lambda_i^{k-1} = p'(\lambda_i) & \text{if } \lambda_i = \lambda_j \end{cases}$$

$\square$

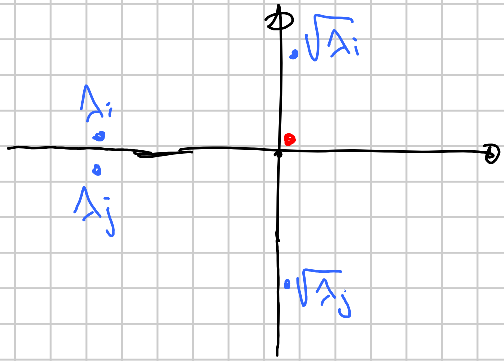
This gives us an idea of when K is large

E.g. let $f(x) = x^{1/2}$ principal square root

For which values of λ_i, λ_j is

$$f[\lambda_i, \lambda_j] = \begin{cases} \dfrac{\lambda_i^{1/2} - \lambda_j^{1/2}}{\lambda_i - \lambda_j} & \lambda_i \neq \lambda_j \\ \dfrac{1}{2\lambda_i^{1/2}} & \lambda_i = \lambda_j \end{cases}$$

very large?



- when $\lambda_i \approx 0$, $f[\lambda_i, \lambda_i]$ is large $\Rightarrow f(A)$ is ill-conditioned

- when λ_i and λ_j are close but on opposite sides of the negative real half-line, $f[\lambda_i, \lambda_j]$ is large.
$\Rightarrow f(A)$ is ill-conditioned.

Two main strategies to compute matrix functions:

- diagonalization
- Taylor series

<u>Diagonalization</u>: if $A = VDV^{-1}$ $\quad D = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$,

then
$$f(A) = V \begin{bmatrix} f(\lambda_1) & & & \\ & f(\lambda_2) & & \\ & & \ddots & \\ & & & f(\lambda_n) \end{bmatrix} V^{-1}$$

If V were orthogonal (for instance because A is Hermitian), then this method is perfectly stable.
One can prove that it is backward stable, if V is orthogonal, and our implementation of $f(\text{scalar})$ is backward stable.
Otherwise, the method can be very unstable.

<u>Example</u>: $A = \begin{bmatrix} 3 & -1 \\ 1 & 1+\varepsilon \end{bmatrix}$ $\qquad \varepsilon = 10^{-15}$

<u>Note</u>

$$\left.\begin{bmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ \varepsilon & & & 0 \end{bmatrix}\right\} K$$

has eigenvalues the K-th roots of ε. So a perturbation of magnitude ε to a Jordan block causes a perturbation of magnitude $|\varepsilon^{1/K}|$ on the eigenvalues.

$f = x^{1/2}$ $\quad f(A) \approx V \begin{bmatrix} \lambda_1^{1/2} & \\ & \lambda_2^{1/2} \end{bmatrix} V^{-1}$ has very poor accuracy!

$$f(x) = x^{1/2} = \sqrt{2} + \frac{1}{2\sqrt{2}}(x-2) + O\left((x-2)^2\right)$$