Funzioni di matrici e integrali di Cauchy

<u>Recap</u>: se f olomorfa all'interno di un contorno Γ, allora

$$f(a) = \frac{1}{2\pi i}\int_\Gamma \frac{f(z)}{z-a}\,dz \qquad \text{per ogni } a \in \operatorname{int}\Gamma$$

$$\frac{1}{k!} f^{(k)}(a) = \frac{1}{2\pi i}\int_\Gamma \frac{f(z)}{(z-a)^{k+1}}\,dz \qquad k \geq 1$$

<u>Teo</u>: f olomorfa in $\operatorname{int}\Gamma$, $A \in \mathbb{C}^{n\times n}$ con $\Lambda(A) \subseteq \operatorname{int}\Gamma$, allora

$$\underset{\substack{\uparrow \\ f \text{ di matrice}}}{f(A)} = \frac{1}{2\pi i}\int_\Gamma \underset{\substack{\uparrow \\ f \text{ scalare}}}{f(z)} (zI-A)^{-1}\,dz.$$

<u>Dim</u>: basta ridursi ai blocchi di Jordan: difatti $A = VJV^{-1}$

$J = \operatorname{diag}(J_1, J_2, \dots J_s)$

$$\frac{1}{2\pi i}\int_\Gamma f(z)(zI-A)^{-1}dz = \frac{1}{2\pi i}\int_\Gamma f(z)\, V\left(zI-J\right)^{-1}V^{-1}\,dz$$

$$= \frac{1}{2\pi i}\int_\Gamma f(z) V \begin{bmatrix} (zI-J_1)^{-1} & & 0 \\ & \ddots & \\ 0 & & (zI-J_s)^{-1} \end{bmatrix} V^{-1} dz$$

$$= V \begin{bmatrix} \frac{1}{2\pi i}\int f(z)(zI-J_1)^{-1}dz & & 0 \\ & \ddots & \\ 0 & & \frac{1}{2\pi i}\int f(z)(zI-J_s)^{-1}dz \end{bmatrix} V^{-1}$$

$$= V \begin{bmatrix} f(J_1) & & \\ & \ddots & \\ & & f(J_s) \end{bmatrix} V^{-1} = f(A).$$

se ho dimostrato il risultato sui blocchi

Su un blocco:

$$\frac{1}{2\pi i}\int f(z)(zI-J)^{-1}dz = \frac{1}{2\pi i}\int f(z) \begin{bmatrix} z-\lambda & -1 & & \\ & z-\lambda & \ddots & \\ & & \ddots & -1 \\ & & & z-\lambda \end{bmatrix}^{-1} dz$$

$$= \frac{1}{2\pi i}\int f(z) \begin{bmatrix} (z-\lambda)^{-1} & (z-\lambda)^{-2} & (z-\lambda)^{-3} & \cdots & (z-\lambda)^{-k} \\ & \ddots & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & (z-\lambda)^{-2} \\ 0 & & & & (z-\lambda)^{-1} \end{bmatrix} dz$$

$$= \begin{bmatrix} \frac{1}{2\pi i}\int \frac{f(z)}{z-\lambda}dz & \frac{1}{2\pi i}\int \frac{f(z)}{(z-\lambda)^2}dz & \cdots & \frac{1}{2\pi i}\int \frac{f(z)}{(z-\lambda)^k}dz \\ & \ddots & \ddots & \vdots \\ & & \ddots & \\ 0 & & & \frac{1}{2\pi i}\int \frac{f(z)}{z-\lambda}dz \end{bmatrix}$$

$$= \begin{bmatrix} f(z) & f'(z) & \frac{1}{2}f''(z) & \cdots & \frac{1}{(k-1)!}f^{(k-1)}(z) \\ & \ddots & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & \vdots \\ 0 & & & \ddots & \\ & & & & f(z) \end{bmatrix} = f(J)$$

$\square$

Corollario: se $\{A_n\} \to A$, allora $f(A_n) \to f(A)$: basta scriverlo con la formula integrale, e usare il fatto che l'integrando $f(z)(zI-A)^{-1}$ è continuo e quindi uniformemente continuo su un compatto $\Lambda(A) \subset K \subset \operatorname{int}\Gamma$

⚠ per f olomorfa.

Oss: questo è vero anche in casi più generali, per esempio $\Lambda(A) \subseteq \mathbb{R}$ e f derivabile un numero opportuno di volte.
La dimostrazione passa dal dimostrare che i polinomi di interpolazione $p_n(x)$ di A_n convergono al polinomio di interpolazione $p(x)$ di A.

Condizionamento di funzioni di matrici

Data $f: \mathbb{C}^m \to \mathbb{C}^n$

$$\lim_{\varepsilon \to 0} \sup_{\|\tilde{x}-x\|<\varepsilon} \frac{\|f(\tilde{x})-f(x)\|}{\|\tilde{x}-x\|} =: K_{abs}(f,x)$$

Se f è differenziabile,

$$f(\tilde{x})-f(x) = J_f(x)\cdot(\tilde{x}-x) + o(\|\tilde{x}-x\|)$$

Quindi $K_{abs}(f,x) = \|J_f(x)\|$

$$K_{rel}(f,x) = \lim_{\varepsilon \to 0} \sup_{\|\tilde{x}-x\|<\varepsilon} \frac{\frac{\|f(\tilde{x})-f(x)\|}{\|f(x)\|}}{\frac{\|\tilde{x}-x\|}{\|x\|}} = \frac{\|x\|}{\|f(x)\|}\cdot K_{abs}(f,x)$$

Per calcolare il condizionamento in più variabili quindi ci serve calcolare norme di Jacobiani.

Def: La <u>derivata di Fréchet</u> di una funzione di matrice f in $X \in \mathbb{C}^{n\times n}$ è l'operatore lineare $L_{f,X}: \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$ tale che

$$f(X+E) = f(X) + L_{f,X}[E] + o(\|E\|)$$

ES: $f(x) = x^2 \qquad f(X) = X^2$

$$f(X+E) = (X+E)^2 = \underbrace{X^2}_{f(X)} + \underbrace{XE + EX}_{L_{f,X}[E]} + \underbrace{E^2}_{o(\|E\|)}$$

$$L_{f,X} : E \longmapsto XE + EX$$

Posso anche definire la mappa vettorizzata

$$\hat{f} : \underset{\in \mathbb{C}^{n^2}}{x = \operatorname{vec} X} \longmapsto \underset{\in \mathbb{C}^{n^2}}{\operatorname{vec}(XE + EX)}$$

Lo Jacobiano $\hat{f}$ è la mappa $\operatorname{vec} E \mapsto \operatorname{vec} L_{f,X}[E]$

È una mappa lineare $\mathbb{C}^{n^2} \to \mathbb{C}^{n^2}$ che posso rappresentare con una matrice.

$$\operatorname{vec}(XEI) = (I^T \otimes X) \cdot \operatorname{vec} E \qquad \operatorname{vec}(I E X) = (X^T \otimes I) \operatorname{vec} E$$

"forma di Kronecker" della derivata di Fréchet:

$$J_{\hat{f},X} = I \otimes X + X^T \otimes I = \hat{L}_{f,X}$$

ES: $f(x) = x^{-1} \qquad f(X) = X^{-1}$

$$f(X+E) = (X+E)^{-1} = \left((I + EX^{-1}) X \right)^{-1} = X^{-1} (I + EX^{-1})^{-1}$$

$$= X^{-1} \left(I - EX^{-1} + EX^{-1}EX^{-1} - EX^{-1}EX^{-1}EX^{-1} + \dots \right)$$

$$= \underbrace{X^{-1}}_{f(X)} \underbrace{- X^{-1}EX^{-1}}_{L_{f,X}[E]} + \underbrace{X^{-1}EX^{-1}EX^{-1} - X^{-1}EX^{-1}EX^{-1}EX^{-1} + \dots}_{o(\|E\|)}$$

$$L_{f,X}[E] = -X^{-1}EX^{-1}$$

$$\hat{L}_{f,X} = -(X^{-1})^T \otimes X^{-1}.$$

Coerente con il caso scalare, $\frac{d}{dx} f(x) = \frac{1}{x^2}$

Proprietà delle derivate di Fréchet:

prodotto di matrici $n^2 \times n^2$

1) $L_{f+g,X} = L_{f,X} + L_{g,X}$

composiz. di funzioni

2) $L_{f\circ g,X} = L_{f,g(X)} \circ L_{g,X}$ $\qquad \hat{L}_{f\circ g,X} = \hat{L}_{f,g(X)} \cdot \hat{L}_{g,X}$

3) $L_{f^{-1},f(X)} = \left(L_{f,X}\right)^{-1}$ $\qquad \hat{L}_{f^{-1},f(X)} = \left(\hat{L}_{f,X}\right)^{-1}$

funzione inversa

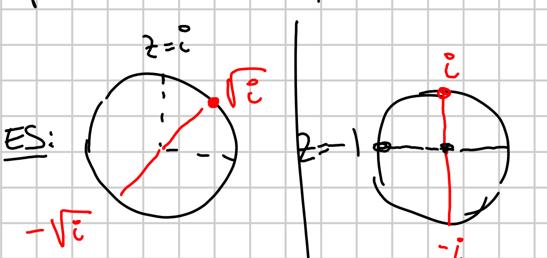
matrice inversa $n^2 \times n^2$

ES: derivata di Fréchet di $g(y) = y^{1/2}$ ← radice quadrata principale

inversa di $f(x) = x^2$

Se $z \in \mathbb{C}$ non è un reale negativo, allora esiste una e una sola radice quadrata di z nel semipiano destro, la chiamiamo radice quadrata principale

ES:



non ha una radice nel semipiano dx aperto

$Y = X^2$ $\qquad f: X \to Y = X^2$

$\qquad g: Y \to \sqrt{Y} = X$

per X con $\Lambda(X) \in RHP$ "right half-plane"

$$L_{f,X} : E \mapsto XE + EX$$

$$(L_{f,X})^{-1} : E \mapsto \text{soluzione } F \text{ di } XF + FX = E$$

Esiste una unica soluzione F? Sì, perché è un'equazione di Sylvester con $\underbrace{\Lambda(X)}_{\substack{\cap \\ \text{RHP}}} \cap \underbrace{\Lambda(-X)}_{\substack{\cap \\ \text{LHP}}} = \phi$

Quindi la derivata di Fréchet della funzione di matrice $g(Y) = Y^{1/2}$ è data da

$$L_{g,Y} : E \mapsto \text{soluzione } F \text{ di } Y^{1/2}F + FY^{1/2} = E$$

Teo: siano dati X, E, f sufficientemente regolare. Allora,

$$\underbrace{f\left(\begin{bmatrix} X & E \\ 0 & X \end{bmatrix}\right)}_{2n \times 2n} = \begin{bmatrix} f(X) & L_{f,X}[E] \\ 0 & f(X) \end{bmatrix}$$

(oss: questa formula è una sorta di "generalizzazione matriciale" di $f\left(\begin{bmatrix} x & 1 \\ 0 & x \end{bmatrix}\right) = \begin{bmatrix} f(x) & f'(x) \\ 0 & f(x) \end{bmatrix}$)

Dim: Idea: partiamo da $\begin{bmatrix} X+\varepsilon E & E \\ 0 & X \end{bmatrix}$ con $\varepsilon > 0$

e usiamo un'equazione di Sylvester per trasformarla in una diagonale a blocchi:

$$\begin{bmatrix} I & -Z \\ 0 & I \end{bmatrix} \cdot \begin{bmatrix} X+\varepsilon E & E \\ 0 & X \end{bmatrix} \cdot \begin{bmatrix} I & Z \\ 0 & I \end{bmatrix} =$$

$$\begin{bmatrix} X+\varepsilon E & E-ZX \\ 0 & X \end{bmatrix}\begin{bmatrix} I & Z \\ 0 & I \end{bmatrix} = \begin{bmatrix} X+\varepsilon E & \boxed{(X+\varepsilon E)Z - ZX + E} \\ 0 & X \end{bmatrix}$$

Se Z soddisfa $(X+\varepsilon E)Z - ZX + E = 0$, ho una matrice diagonale a blocchi. Notiamo che $Z = -\frac{1}{\varepsilon} I$ risolve l'equazione:

$$-(\cancel{X} + \cancel{\varepsilon} E)\frac{1}{\cancel{\varepsilon}} + \frac{1}{\cancel{\varepsilon}}\cancel{X} + E = -E + E = 0$$

Allora posso scrivere la catena di uguaglianze

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

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

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

Facendo tendere $\varepsilon \to 0$, ho

$$\lim f\left(\begin{bmatrix} X & E \\ 0 & X \end{bmatrix}\right) = \begin{bmatrix} f(X) & L_{f,X}[E] \\ 0 & f(X) \end{bmatrix}.$$

<u>Teo</u>: Sia X con autovalori λ_i, ognuno con molt. algebrica m_i, e sia f di classe C^{2m_i-1} in un intorno di ogni λ_i. Allora, $f(X)$ esiste ed è differenziabile nel senso di Fréchet in X.

<u>Dim</u>: data $\tilde{X}$ sufficientemente vicina a X, $\tilde{X}$ ha blocchi di Jordan di dimensione al più m_i.

e $f(\tilde{X})$ è ben definita.

La matrice $\begin{bmatrix} \tilde{X} & E \\ 0 & \tilde{X} \end{bmatrix}$ ha blocchi di Jordan di dimensione al più $2m_i$. Allora, $f\left(\begin{bmatrix} \tilde{X} & E \\ 0 & \tilde{X} \end{bmatrix}\right)$

è ben definita, perché compaiono derivate di f fino alla $(2m_i-1)$-esima. Inoltre, è continua.

Quindi, anche $\lim\limits_{\varepsilon \to 0} \dfrac{f(X+\varepsilon E) - f(X)}{\varepsilon}$ esiste ed è una funzione

continua della X. Questa è l'ipotesi che mi serve per applicare il teorema del differenziale totale e concludere che $\hat{f}$ è differenziabile come funzione di n^2 variabili. $\square$

Quindi in particolare $K_{abs}(f,X) = \|L_{f,X}\| = \|\hat{L}_{f,X}\|$.

La norma comoda da usare qui è $\|X\|_F = \|\mathrm{vec}\, X\|_2$

$$\|L_{f,X}\| = \max_{E \neq 0} \frac{\|L_{f,X}[E]\|_F}{\|E\|_F} \qquad \|\hat{L}_{f,X}\|_2 \quad \text{(norma 2 della matrice } n^2 \times n^2\text{)}$$

$$\sigma_{max}(\hat{L}_{f,X})$$

Ad es. se $f(x) = x^2$ $\quad \|X^T \otimes I + I \otimes X\|_2 \leq \|X^T \otimes I\| + \|I \otimes X\|$

$$= \|X^T\| + \|X\| = 2\|X\|.$$

Studiare $\sigma_{max}(\hat{L}_{f,X})$ spesso è difficile, però possiamo dare un teorema che caratterizza $\Lambda(\hat{L}_{f,X})$