<u>Thm</u>: se $\|A\| \leq 5.4$, allora l'approssimante di Padé $(13, 13)$ soddisfa

$$D^{-1}(A)\, N(A) = \exp(A+H) \qquad \frac{\|H\|}{\|A\|} \leq u \approx 2.2 \cdot 10^{-16}$$

ma quanto distano $\exp(A+H)$ e $\exp(A)$?

$$\|\exp(A+H) - \exp(A)\| = \underbrace{K_{exp,abs}} \cdot H + O(H^2)$$

Se $\|A\| > 5.4$? <u>Idea</u>: per ogni $s \in \mathbb{N}$, $\exp(A) = \underbrace{\exp\left(\frac{1}{s}A\right)}_{B}{}^{s}$

comodo con $s = 2^k$, perché $B^{2^k} = \underbrace{(\cdots((B^2)^2)^2 \cdots)^2}_{k \text{ volte}}$

1) Trovo $s = 2^k$ abbastanza grande da $\|\frac{1}{s}A\| \leq 5.4$
2) Calcolo $C = D(\frac{1}{s}A)^{-1} N(\frac{1}{s}A)$ (approx. $(13, 13)$)
3) Calcolo C^{2^k} per quadrature successive

Se A in forma di Schur,

approssimaz. di $\exp(A_{ii})$

$$\exp(A) \approx \begin{bmatrix} \square & & * \\ & \ddots & \\ 0 & & \end{bmatrix}$$

Se $\|A\|$ già più piccola di 5.4, expm può usare approssimanti di grado inferiore a 13

Abbiamo visto che $D(A) = N(-A)$

$$N(A) = a_0 + a_1 A + a_2 A^2 + \cdots + a_{13} A^{13} = \overbrace{(a_0 + a_2 A^2 + a_4 A^4 + \cdots)}^{U} + A \overbrace{(a_1 + a_3 A^2 + \cdots + a_{13} A^{12})}^{V}$$

$$D(A) = a_0 - a_1 A + a_2 A^2 - \dots - a_{13} A^{13} = \underbrace{(a_0 + a_2 A^2 + a_4 A^4 + \dots)}_{U} - A\underbrace{(a_1 + a_3 A^2 + \dots + a_{12} A^{12})}_{V}$$

Mi basta valutare U e V, poi

$$N(A) = U + AV$$
$$D(A) = U - AV$$

Il calcolo di C^{2^k} a partire da C è stabile?

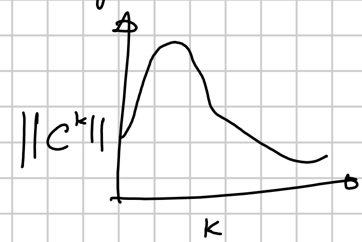
<u>Teo</u>: date M, N, l'approssimazione P di $M \cdot N$ calcolata in macchina soddisfa

$$\|P - MN\| \le \underbrace{\|M\| \cdot \|N\|}_{\text{qui c'è } \|M\|\cdot\|N\| \text{ e non } \|P\|} \cdot u \cdot O(n)$$

Usandolo ricorsivamente, l'approssimazione calcolata D di C^{2^k} soddisfa

$$\|D - C^{2^k}\| \le \|C\|^2 \|C^2\| \cdot \|C^4\| \cdots \|C^{2^{k-1}}\| \cdot u \cdot O(n) O(k)$$

potenze intermedie, e sappiamo che $\|C^{2^i}\|$ può essere molto maggiore di $\|C^{2^k}\|$ ("hump phenomenon")



$$A = \underbrace{V}_{} J \underbrace{V^{-1}}_{} \to K(V)$$

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1+\varepsilon \end{bmatrix}$$ per diagonalizzarla serve una V con $K(V) \approx \frac{1}{\varepsilon}$

autovett. di A: $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ con $\lambda = 1$

$$\ker(A - (1+\varepsilon)I) = \ker \begin{bmatrix} -\varepsilon & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 \\ \varepsilon \end{bmatrix}$$

$$V = \begin{bmatrix} 1 & 1 \\ 0 & \varepsilon \end{bmatrix} \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix} = \begin{bmatrix} d_1 & d_2 \\ 0 & \varepsilon d_2 \end{bmatrix}$$

$$V^{-1} = \begin{pmatrix} d_1^{-1} & 0 \\ 0 & d_2^{-1} \end{pmatrix} \begin{pmatrix} 1 & -\varepsilon^{-1} \\ 0 & \varepsilon^{-1} \end{pmatrix} = \begin{pmatrix} d_1^{-1} & -d_1^{-1}\varepsilon^{-1} \\ 0 & d_2^{-1}\varepsilon^{-1} \end{pmatrix}$$

Per esempio, in norma-inf, $\|V\| = \max(d_1 + d_2, \varepsilon d_2)$

$$\|V^{-1}\| = \max\left(d_1^{-1}(1+\varepsilon^{-1}),\ d_2^{-1}\varepsilon^{-1}\right)$$

<u>Teo</u>: se A, B commutano, allora $\exp(A+B) = \exp(A)\exp(B)$.

<u>Dim</u>:

$$\exp(A)\exp(B) = \left(\sum_n \frac{A^n}{n!}\right)\left(\sum_m \frac{B^m}{m!}\right) = \qquad = \sum_k \frac{(A+B)^k}{k!} = \exp(A+B)$$

uguali perché il coeff. di $A^n B^m$ è

$k = n+m$

coeff. di $A^n B^m$: $\quad \dfrac{A^n}{n!} \cdot \dfrac{B^m}{m!} \xrightarrow{\text{uguali}} \dbinom{n+m}{n} \cdot \dfrac{A^n B^m}{(m+n)!}$

Fissato ε, esiste N tale che

$$\sum_{m+n \geq N} \frac{\|A\|^n \cdot \|B\|^m}{(m+n)!} < \varepsilon$$

La somma parziale del LHS

$$\left(\sum_{n \leq N} \frac{A^n}{n!}\right)\left(\sum_{m \leq N} \frac{B^m}{m!}\right)$$ differisce da $\exp(A)\exp(B)$

per $$\Bigg\| \sum_{\substack{n > N \\ \text{oppure } m > N}} \frac{A^n}{n!} \frac{B^m}{m!} \Bigg\| \leq \Bigg\| \sum_{m+n \geq N} \frac{\|A\|^n \|B\|^m}{(m+n)!} \Bigg\| < \varepsilon$$

$$\exp(A - \tau I) = \exp(-\tau I) \cdot \exp(A) = \exp(-\tau) \cdot \exp(A)$$

Permette di ricondurre $\exp(A)$ a $\exp(A-\tau I)$ che può avere norma minore se scelgo τ opportunamente

---

A si dice matrice di Metzler (o "essenzialmente non-negativa") se $A_{ij} \geq 0$ per ogni $i \neq j$ (es. $\begin{bmatrix} 18 & 2 & 3 \\ 7 & -5 & 0 \\ 8 & 0 & -37 \end{bmatrix}$)

Teo: se A Metzler, allora $\exp(A) \geq 0$ (entrywise)

Dim: per τ opportuno, $B = A + \tau I \geq 0$

$$\exp(A) = \exp(B - \tau I) = e^{-\tau} \cdot \exp(B) = e^{-\tau} \sum_{k=0}^{\infty} \frac{B^k}{k!} \geq 0 \quad \square$$

Tra l'altro, se calcolo $\exp(A)$ in questo modo, $e^{-\tau} \sum_{k=0}^{M} \frac{B^k}{k!}$, non ho problemi di cancellazione perché sono tutti termini positivi

---

$$f(a,b) = 0 \quad \forall a, b \in \mathbb{R}$$

$$A, I, \ .. \ (\sqrt{A})^2 - A = 0$$

Numeri di condizionamento: $f: \mathbb{R}^n \to \mathbb{R}^n$

se f differenziabile,
$f(\tilde{x}) = f(x) + \nabla f \cdot (\tilde{x} - x) + o(\tilde{x} - x)$

$\tilde{x} - x = h$

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

$$= \lim_{\varepsilon \to 0} \sup_{\|h\| \leq \varepsilon} \frac{\|\nabla f \cdot h + \cancel{o(h)}\|}{\|h\|} = \|\nabla f\|$$

$$K_{rel}(f,x) = \lim_{\varepsilon \to 0} \sup_{\frac{\|\tilde{x}-x\|}{\|x\|} \leq \varepsilon} \frac{\|f(\tilde{x}) - f(x)\|}{\|f(x)\|} \Big/ \frac{\|\tilde{x} - x\|}{\|x\|} = \|\nabla f\| \cdot \frac{\|x\|}{\|f(x)\|}$$

Le funzioni di matrici sono $f: \mathbb{R}^{n^2} \to \mathbb{R}^{n^2}$

Def:

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

$$f(X+H) = f(X) + L_{f,X}(H) + o(\|H\|)$$

ES:

Derivata di Fréchet di $f(x) = x^2$ nella matrice X

$$f(X+H) = (X+H)^2 = \underbrace{X^2}_{f(X)} + \underbrace{XH + HX}_{L_{f,X}(H)} + \underbrace{H^2}_{o(\|H\|)} \qquad \|H^2\| \leq \|H\|^2 = o(\|H\|)$$

$L_{f,X}$ funzione da $\mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$

lineare: $L_{f,X}(H_1 + H_2) = L_{f,X}(H_1) + L_{f(x)}(H_2)$

In alternativa, uno può vedere la versione vettorizzata di questo L: cioè, la funzione

$$\hat{L}: \operatorname{vec}(H) \mapsto \operatorname{vec} L_{f,X}[H] \qquad \hat{L}: \mathbb{R}^{n^2} \to \mathbb{R}^{n^2}$$

Identità: $\operatorname{vec}(AXB) = (B^T \otimes A) \cdot \operatorname{vec}(X)$

$$\operatorname{vec} L_{f,X}[H] = \operatorname{vec}(XHI + IHX) =$$

$$= \underbrace{(I \otimes X + X^T \otimes I)}_{\text{matrice } n^2 \times n^2 \text{ che rappresenta } L_{f,X}} \operatorname{vec}(H)$$

$\hat{L}_{f,X}$

$f: x \mapsto x^3$, derivata della f. di matrice associata:

$$(X+H)^3 = \underbrace{X^3}_{f(X)} + \underbrace{X^2H + XHX + HX^2}_{L_{f,X}[H]} + \underbrace{H^2X + HXH + XH^2 + H^3}_{o(\|H\|)}$$

Vettorizzando,

$$\mathrm{vec}\, L_{f,X}[H] = \mathrm{vec}\left(X^2H + XHX + HX^2\right) = \left(I\otimes X^2 + X^T\otimes X + (X^2)^T\otimes I\right)\mathrm{vec}\, H$$

Proprietà:

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

2) $L_{f\circ g,X} = L_{f,g(X)} \circ L_{g,X}$

3) $L_{f^{-1},f(X)} = L_{f,X}^{-1}$ $\leftarrow$

ES: calcolo la derivata di Fréchet della funzione inversa di $f(x)=x^2$,

cioè $g(y)=\sqrt{y}$ per $x \in RHP = \{z\in\mathbb{C} : Re(z)>0\}$

$\uparrow$ radice quadrata principale, cioè quella con $Re\sqrt{y}>0$

$L_{f,X}[H] = XH + HX$ $\quad Y = X^2$

$L_{g,Y}$ è l'inverso di $L_{f,X}$, cioè l'operatore che prende E e restituisce H tale che $XH + HX = E$

$\uparrow$ soluzione di un'equazione di Sylvester

Per calcolare $L_{g,Y}[E]$, prima calcolo $X=$ radice quadrata princ. di Y, e poi risolvo l'eq. di Sylvester $XH+HX=E$, e la soluzione H è uguale a $H = L_{g,Y}[E]$

L'equazione è risolubile se $\Lambda(X)\cap\Lambda(-X)=\emptyset$, che è sempre vera perché ho preso la radice quadrata principale, cioè $\Lambda(X)\subseteq RHP$, $\Lambda(-X)\subseteq LHP$ (semipiani aperti)

In termini di matrici $n^2 \times n^2$,

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

$$\hat{L}_{g,Y} = \left(I \otimes X + X^T \otimes I\right)^{-1} = \left(I \otimes \sqrt{Y} + \sqrt{Y}^T \otimes I\right)^{-1}$$

radici principali

($\hat{L}$ è detta "forma di Kronecker" della derivata di Fréchet)

Teo: supponiamo che f ammetta derivata di Fréchet in X,

allora

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

(riduce il calcolo di $L_{f,X}[E]$ al calcolo di f su una matrice $2n \times 2n$).