Metodo di Newton per $X^{1/2}$

$f:\mathbb{R}^{n^2}\to\mathbb{R}^{n^2}$ $\qquad f(\mathrm{vec}\,X)=\mathrm{vec}(X^2-A)$

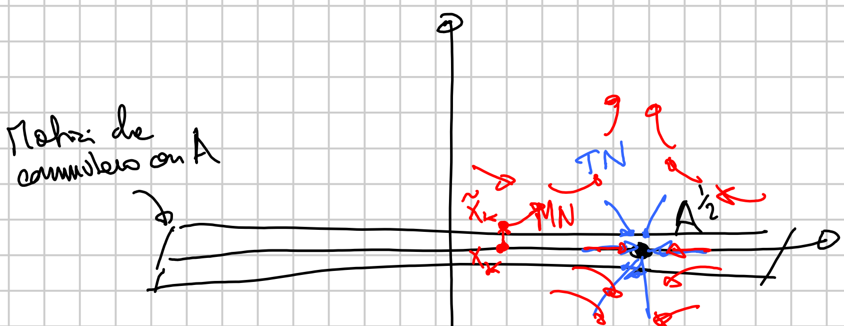
$$\boxed{X_{k+1}=X_k-H_k \qquad X_kH_k+H_kX_k=X_k^2-A} \qquad \text{(TN)}$$

Oppure: $H_k=(2X_k)^{-1}(X_k^2-A)$, e quindi

$$X_{k+1}=\frac{1}{2}\left(X_k+X_k^{-1}A\right) \qquad \text{(MN)}$$

In arit. esatta, se X_0 e A commutano (ad es. $X_0=\alpha A$, αI) allora coincidono

Geometricamente:



Comportamento intorno al punto fisso $X=A^{1/2}$ di $X_{k+1}=h(X_k)$

$$h(X)=\frac{1}{2}\left(X+X^{-1}A\right)$$

$$h(X+H)=h(X)+L_{h,X}[H]+o(\|H\|^2)$$

$\mathrm{Jac}(h)$ $\qquad L_{h,X}[H]=\frac{1}{2}\left(H-X^{-1}HX^{-1}A\right)$

$$\hat{L}_{h,X}=\frac{1}{2}\left(I\otimes I-(X^{-1}A)^T\otimes X^{-1}\right)$$

$$\hat{L}_{h,A^{1/2}}=\frac{1}{2}\left(I\otimes I-(A^{1/2})^T\otimes A^{-1/2}\right)$$

Per calcolare autovalori, fatt. Schur $(A^{\frac{1}{2}})^T = Q_1 T_1 Q_1^*$

$$A^{-\frac{1}{2}} = Q_2 T_2 Q_2^*$$

$$\hat{L}_{h, A^{\frac{1}{2}}} = (Q_1 \otimes Q_2)\Big(\underbrace{\tfrac{1}{2}\big(I \otimes I - T_1 \otimes T_2}_{\text{Triangolare}}\big)\Big)(Q_1 \otimes Q_2)^*$$

Se $\lambda_1, \dots \lambda_n$ autoval. di A, $\operatorname{diag}(T_1) = \lambda_1^{\frac{1}{2}}, \dots \lambda_n^{\frac{1}{2}}$

$$\operatorname{diag}(T_2) = \lambda_1^{-\frac{1}{2}}, \dots \lambda_n^{-\frac{1}{2}}$$

$$\operatorname{diag}\left(\tfrac{1}{2}(I \otimes I - T_1 \otimes T_2)\right) = \tfrac{1}{2}\left(1 - \lambda_i^{\frac{1}{2}} \lambda_j^{-\frac{1}{2}}\right) \quad \text{per } i,j = 1, \dots n$$

se tutti questi autoval. sono in modulo ≤ 1, $A^{\frac{1}{2}}$ è un punto fisso stabile

$$\lambda_1 = 25 \quad \lambda_2 = 1$$

$$\tfrac{1}{2}\left(1 - \lambda_1^{\frac{1}{2}} \lambda_2^{-\frac{1}{2}}\right) = \tfrac{1}{2}(1 - 5 \cdot 1) = -2$$

Altro metodo: Iterazione di Denman-Beavers (DB)

Ricordiamo che $\operatorname{sign}\left(\begin{bmatrix} 0 & A \\ I & 0 \end{bmatrix}\right) = \begin{bmatrix} 0 & A^{\frac{1}{2}} \\ A^{-\frac{1}{2}} & 0 \end{bmatrix}$

$X_0 = A$, $Y_0 = I$ Iterazione Newton per il segno:

$$\begin{bmatrix} 0 & X_k \\ Y_k & 0 \end{bmatrix} \longmapsto \frac{1}{2}\left(\begin{bmatrix} 0 & X_k \\ Y_k & 0 \end{bmatrix} + \begin{bmatrix} 0 & Y_k^{-1} \\ X_k^{-1} & 0 \end{bmatrix}\right)$$

cioè
$$\begin{cases} X_{k+1} = \frac{1}{2}(X_k + Y_k^{-1}) \\ Y_{k+1} = \frac{1}{2}(Y_k + X_k^{-1}) \end{cases} \qquad X_0 = A, \quad Y_0 = I \qquad \text{(DB)}$$

Per convergenza della fun. segno, $X_k \to A^{\frac{1}{2}}$ $Y_k \to A^{-\frac{1}{2}}$

(se $\begin{bmatrix} 0 & A \\ I & 0 \end{bmatrix}$ non ha autoval. sull'asse immaginario) $\leadsto$ autovalori: $\pm\sqrt{\lambda}$ per $\lambda \in \Lambda(A)$

ed è asintoticam. stabile

<u>Teo</u>: $(A^{\frac12}, A^{-\frac12})$ è un pto fisso stabile per (DB)

<u>dim</u>:

$$h: \begin{bmatrix} X \\ Y \end{bmatrix} \to \begin{bmatrix} \frac12 (X+Y^{-1}) \\ \frac12 (Y+X^{-1}) \end{bmatrix}$$

$$h\left(\begin{bmatrix} X+E \\ Y+F \end{bmatrix}\right) = \begin{bmatrix} \frac12 \left((X+E)+(Y+F)^{-1}\right) \\ \frac12 \left((Y+F)+(X+E)^{-1}\right) \end{bmatrix} = h\left(\begin{bmatrix} X \\ Y \end{bmatrix}\right) + \underbrace{\begin{bmatrix} \frac12 E - Y^{-1}FY^{-1} \\ \frac12 F - X^{-1}EX^{-1} \end{bmatrix}}_{L_{h,\begin{bmatrix} X \\ Y \end{bmatrix}}} + O\left(\left\|\begin{bmatrix} E \\ F \end{bmatrix}\right\|^2\right)$$

$$L_{h,\begin{bmatrix} A^{\frac12} \\ A^{-\frac12} \end{bmatrix}} = \begin{bmatrix} \frac12 \left(E - A^{\frac12} F A^{\frac12}\right) \\ \frac12 \left(F - A^{-\frac12} E A^{-\frac12}\right) \end{bmatrix}$$

Questa mappa è (quasi) idempotente: $L_{h,\begin{bmatrix} A^{\frac12} \\ A^{-\frac12} \end{bmatrix}}\left(L_{h\begin{bmatrix} A^{\frac12} \\ A^{-\frac12} \end{bmatrix}} \begin{bmatrix} E \\ F \end{bmatrix}\right) = \begin{bmatrix} E \\ F \end{bmatrix}$

$$\begin{cases} \frac12 \left(\frac12 \left(E - A^{\frac12} F A^{\frac12}\right)\right) - \frac12 A^{\frac12} \left(\frac12 \left(F - A^{-\frac12} E A^{-\frac12}\right)\right) A^{\frac12} = \frac14 E \cancel{- \frac14 A^{\frac12} F A^{\frac12}} \cancel{- \frac14 A^{\frac12} F A^{\frac12}} + \frac14 E \\ \frac12 \left(\frac12 \left(F - A^{-\frac12} E A^{-\frac12}\right)\right) - \frac12 A^{-\frac12} \left(\frac12 \left(E - A^{\frac12} F A^{\frac12}\right)\right) A^{-\frac12} = \frac14 F \cancel{- \frac14 A^{-\frac12} E A^{-\frac12}} \cancel{+ \frac14 A^{-\frac12} E A^{-\frac12}} + \frac14 F \end{cases}$$

$$= \begin{bmatrix} \frac12 E \\ \frac12 F \end{bmatrix}$$

Quindi, vettorizzandola, è una matrice $M \in \mathbb{C}^{2n^2 \times 2n^2}$ tale che $M^2 = \frac12 I$

$\Rightarrow$ autovalori sono tutti $\pm\sqrt{\frac12}$ $\Rightarrow$ stabile

---

$f(A)$ per matrici grandi e sparse

$f(A)$ densa, solitamente, quindi spesso si vuole calcolare $f(A)b$ per un certo vettore b

Tecniche principali

1) rimpiazzo f con un polinomio che la approssima bene in una regione che contiene $\Lambda(A)$

<u>es</u>: se so che A è simmetrica con autovalori in $[a,b]$ (che stimo, ad es. con $|\lambda| \leq \|A\|$), allora mi basta un polinomio tale che $\max_{x\in[a,b]} |f(x)-p(x)| < \varepsilon$, che esiste (Stone-W.)

e calcolo $p(A)b$ con $(\deg p)$ prodotti mat-vec

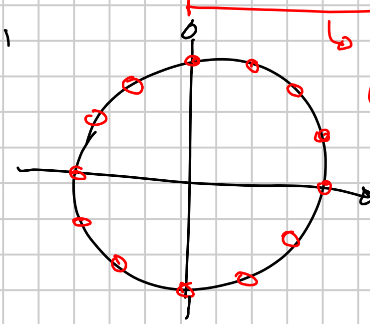
<u>Variante</u>: invece di p, uso una funzione razionale $\frac{p(x)}{q(x)}$,
e in questo caso mi servono anche soluzioni di sistemi lineari
(per esempio, se scrivo $r(x) = \sum_{i=1}^{d} \frac{w_i}{x-p_i}$ (fratti semplici),

allora per calcolare $r(A)b = \sum_{i=1}^{d} w_i (A-p_i I)^{-1} b$
mi servono d soluz. di sistemi lineari

2) Integrali di Cauchy:

$$f(A)b = \frac{1}{2\pi i}\int_{\Gamma} f(z)(zI-A)^{-1}b\,dz \approx \sum_{k=1}^{n} w_i f(z_i)\underbrace{(z_i I - A)^{-1} b}$$

↳ n sist. lineari con A shiftata



3) metodi ad-hoc per funzioni specifiche; ad es.,
$\exp(A)b$ è $v(1)$, dove v soddisfa $\begin{cases} \dot v = Av \\ v(0) = b \end{cases}$.

e posso applicare un metodo di soluzione numerico di ODE

odes.
$$\begin{cases} v_{k+1} = v_k + hAv_k \\ v_0 = b \end{cases} \Rightarrow v_n = (I+hA)^n b$$

4) Arnoldi:

Def: dati $A \in \mathbb{C}^{n\times n}$, $b \in \mathbb{C}^n$, spazio di Krylov di indice n è

$$K_n(A,b) = \text{span}\{b, Ab, A^2b, \dots A^{n-1}b\}$$

$$= \{v \in \mathbb{C}^n : v = \alpha_0 b + \alpha_1 Ab + \dots + \alpha_{n-1}A^{n-1}b = p(A)b\}$$

p polinomio di grado $\leq n-1$

Facendo n prodotti matrice-vettore, riesco a rappresentare l'azione dell'operatore lineare A su $K_n(A,b)$

$$A\underbrace{[b|Ab|\dots|A^{n-1}b]}_{V} = [Ab|A^2b|\dots|A^nb] = \underbrace{[b|Ab|\dots|A^nb]}_{W} \cdot \begin{bmatrix} 0 & 0 & \dots & 0 \\ 1 & & & \\ & \ddots & & \\ & & & 1 \end{bmatrix}$$

matrice che rappresenta A con basi V in partenza e W in arrivo sul sottosp. $\text{Im}\,V = K_n(A,b)$

Dato un vettore $v \in \text{Im}\,V = K_n(A,b)$, posso calcolarmi i coefficienti

$$\begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_{n-1} \end{bmatrix} = \boxed{V^+} v$$ e poi calcolarmi Av come

$$Av = A(\alpha_0 b + \alpha_1 Ab + \dots + \alpha_{n-1}A^{n-1}b) = \alpha_0 Ab + \dots + \alpha_{n-1}A^n b.$$

molto mal condizionato

Arnoldi: algoritmo per calcolare (iterativamente) basi ortonormali di $K_1(A,b) = \text{span}(b)$, $K_2(A,b) = \text{span}(b, Ab)$, $K_3(A,b) = \text{span}(b, Ab, A^2b)$

e così via.

Passo base: $v_1 = \frac{b}{\|b\|}$ è una base o.n. di span(b)

Passo $j \to j+1$: se ho calcolato $(v_1, v_2, \dots v_j)$ base o.n. di $K_j(A,b)$, calcolo v_{j+1} tale che $(v_1, v_2, \dots v_{j+1})$ è base o.n. di $K_{j+1}(A,b)$

Per farlo, parto da v_j, costruisco $w = Av_j$, e lo ortogonalizzo rispetto a tutti quelli precedenti:

$$w = Av_j \,, \quad w \leftarrow w - v_1(v_1^* w)$$
$$w \leftarrow w - v_2(v_2^* w)$$
$$\vdots$$
$$w \leftarrow w - v_j(v_j^* w)$$

$$v_j \in K_j(A,b) \Rightarrow Av_j = w \in K_{j+1}(A,b)$$

w_{finale}

$$v_j = \alpha_0 b + \dots + \alpha_{j-1} A^{j-1} b \Rightarrow Av_j = \alpha_0 Ab + \dots + \alpha_{j-1} A^j b$$

Alla fine del processo, w_{finale} è ortogonale a $v_1, v_2, \dots v_j$

$\Rightarrow v_{j+1} = \frac{w_{finale}}{\|w_{finale}\|}$ è tale che $(v_1, v_2, \dots v_j, v_{j+1})$ sono $j+1$ vettori ortonormali in $K_{j+1}(A,b)$ $\Rightarrow$ sono una base ortonormale

Perché parto da Av_j e non da Av_1, Av_2, o un altro elemento di $K_j(A,b)$? Perché in questo modo posso garantire che $\|w_{finale}\| \neq 0$

<u>Lemma</u>: supponendo che V_j abbia rango pieno (cioè che $b, Ab, \dots A^{j-1}b$ sono lin. indipendenti). Allora l'ultimo vettore costruito da GS v_j, appartiene a $K_j(A,b) \setminus K_{j-1}(A,b)$.

Dim: $j=1$: $v_1 = \frac{b}{\|b\|} \in$ span(b) ma non a span$(\emptyset) = \{0\}$

$j \to j+1$ $\quad v_j \in K_j(A,b) \setminus K_{j-1}(A,b)$ vuol dire che

$$v_j = \alpha_0 b + \alpha_1 Ab + \dots + \alpha_{j-1} A^{j-1} b \quad \text{ha } \alpha_{j-1} \neq 0$$

$v_j = p(A)b$ per un p di grado esattamente $j-1$.

Allora, $w = Av_j = \alpha_0 Ab + \dots + \alpha_{j-1} A^j b = q(A)b$ di grado esattamente j

$$\Rightarrow \quad w \in K_{j+1}(A,b) \setminus K_j(A,b)$$

Quando ortogonalizziamo, $\quad w \leftarrow w - v_i (v_i^* w)$,

sottraiamo $v_i \in K_j(A,b)$, quindi il risultato starà sempre

in $K_{j+1}(A,b) \setminus K_j(A,b)$, e in particolare

$w_{\text{finale}} \neq 0$.