Algoritmi iterativi per sqrt(A)

Newton su $X^2-A=0$:

$$X_{k+1}=X_k-E$$

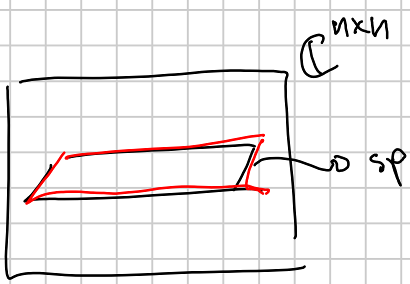
E risolve $X_kE+EX_k=X_k^2-A$

Se X_0 commuta con A, è equivalente a

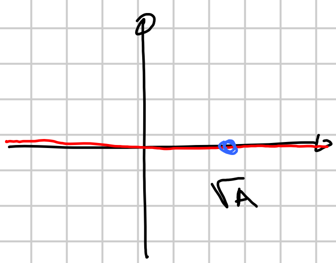
$$X_{k+1}=\frac{1}{2}\left(X_k+X_k^{-1}A\right)$$

In aritmetica esatta, producono la stessa successione

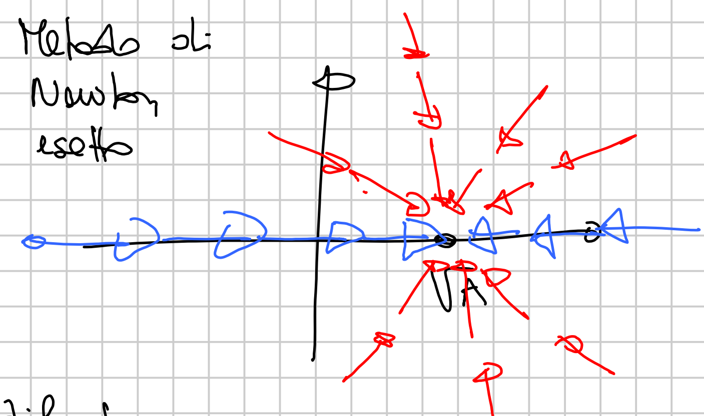
In realtà, una converge quadraticamente anche in aritmetica di macchina, l'altra ha forti problemi di stabilità.



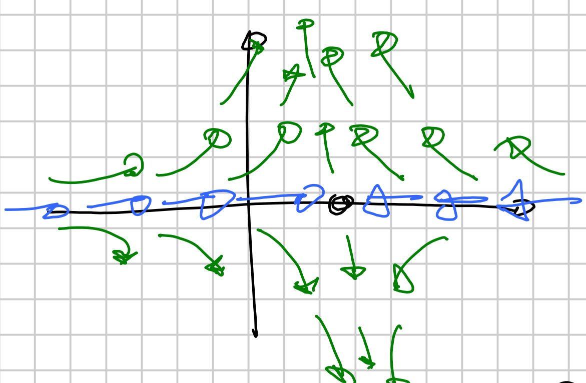
→ spazio delle matrici che commutano con A



Metodo di Newton esatto

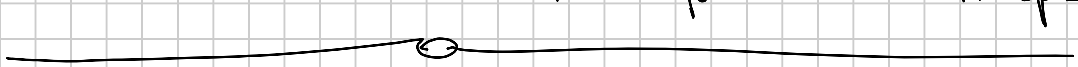


Metodo di Newton modificato



← Coincide con il Newton esatto sul sottospazio delle matrici che commutano con A

fa qualcosa di diverso (non converge) al di fuori del sottospazio



Sistemi dinamici (tempo discreto): studiamo

$x_{k+1} = f(x_k)$ attorno a un punto fisso α

Caso in dimensione 1: $f: \mathbb{R} \to \mathbb{R}$, $\alpha \in \mathbb{R}$

Se $x_k = \alpha + h$, $x_{k+1} = f(\alpha + h) = f(\alpha) + f'(\alpha)h + O(h^2)$

$$= \alpha + f'(\alpha)h + O(h^2)$$

Se $|f'(\alpha)| > 1$, mi allontano da α all'iterazione successiva,

Se $|f'(\alpha)| < 1$, mi avvicino ad α, e convergo localmente.

Nel primo caso, chiamo α punto fisso repulsivo,
nel secondo attrattivo. La convergenza è lineare
(o secondo alcuni "esponenziale") cioè $\lim_{k \to \infty} \frac{|x_{k+1} - \alpha|}{|x_k - \alpha|} = |f'(\alpha)|$

In dimensione superiore, la situazione è simile:

$f: \mathbb{C}^n \to \mathbb{C}^n$ $\quad x_{k+1} = f(x_k)$ $\quad \alpha \in \mathbb{C}^n$ punto fisso

Se $x_k = \alpha + h$ $\quad h \in \mathbb{C}^n$, allora

$$x_{k+1} = f(\alpha) + J_{f,\alpha} \cdot h + O(h^2) = \alpha + J_{f,\alpha} \cdot h + O(h^2)$$

Se $\rho(J_{f,\alpha}) < 1$, la successione si comporta come
moltiplicare ad ogni passo per $J_{f,\alpha}$ a meno di $O(h^2)$, e

$$\lim_{k \to \infty} \frac{\|x_{k+1} - \alpha\|}{\|x_k - \alpha\|} = \rho(J_{f,\alpha})$$

Se $\rho(J_{f,\alpha}) > 1$ la successione diverge.

Nel nostro caso abbiamo una successione di matrici

$X_0, X_1, X_2, \dots$ $\qquad X_{k+1} = F(X_k) \qquad X_k \in \mathbb{C}^{n\times n}$

Per studiare il comportamento della successione attorno a un punto fisso, guardo la derivata di Fréchet.

Per il metodo di Newton esatto, $J_{F,\alpha} = 0$, corrispondentemente al fatto che il metodo di Newton ha convergenza quadratica (o "doppiamente esponenziale").

Per il metodo di Newton modificato,

$$X_{k+1} = F(X_k) \qquad F(X) = \frac{1}{2}(X + X^{-1}A)$$

F ha derivata di Fréchet data da

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

$$= \frac{1}{2}\left(X+E + \left(X^{-1} - X^{-1}EX^{-1} + O(\|E\|^2)\right)A\right)$$

$$= \underbrace{\frac{1}{2}(X + X^{-1}A)}_{F(X)} + \underbrace{\frac{1}{2}(E - X^{-1}EX^{-1}A)}_{L_{F,X}[E]} + O(\|E\|^2)$$

La sua forma con prodotto di Kronecker è

$$K_{F,X} = \frac{1}{2}\left(I - (X^{-1}A)^T \otimes X^{-1}\right)$$

(cioè la matrice tale che
$K_{F,X} \cdot \mathrm{vec}(E) = \mathrm{vec}\, L_{F,X}[E]$)

Quindi $K_{F,A^{\frac{1}{2}}} = \frac{1}{2}\left(I - A^{\frac{1}{2}} \otimes A^{-\frac{1}{2}}\right) \in \mathbb{C}^{n^2\times n^2}$

Se chiamiamo $\lambda_1, \lambda_2, \dots \lambda_n$ gli autovalori di A, gli autoval. di $A^{\frac{1}{2}}$ sono $\lambda_1^{\frac{1}{2}}, \ldots \lambda_n^{\frac{1}{2}}$

e di $A^{\frac{1}{2}} \otimes A^{-\frac{1}{2}}$ sono $\lambda_i^{\frac{1}{2}} \cdot \lambda_j^{-\frac{1}{2}}$ per $i,j = 1, \ldots n$

di $K_{F, A^{\frac{1}{2}}}$ sono $\frac{1}{2}\left(1 + \lambda_i^{\frac{1}{2}} \lambda_j^{-\frac{1}{2}}\right)$ per $i,j = 1, \ldots n$.

Se la matrice è abbastanza mal condizionata, $\frac{\lambda_{max}}{\lambda_{min}}$ è grande e questa quantità è $>> 1$.

(In generale non è mai < 1, basta prendere $i = j$)

$\to$ la successione $X_{k+1} = \frac{1}{2}\left(X_k + X_k^{-1} A\right)$ non è mai localmente convergente

---

Metodo di Denman-Beavers

$$\begin{cases} X_{k+1} = \frac{1}{2}\left(X_k + Y_k^{-1}\right) & X_0 = A & \lim_{k \to \infty} X_k = A^{\frac{1}{2}} \\ Y_{k+1} = \frac{1}{2}\left(Y_k + X_k^{-1}\right) & Y_0 = I & \lim_{k \to \infty} Y_k = A^{-\frac{1}{2}} \end{cases}$$

Si ottiene dalla precedente definendo $Y_k = A^{-1} X_k$.

Oppure, usando Newton per il segno su

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

Se applichiamo l'iterazione di Newton per il segno, abbiamo

$$Z_0 = \begin{bmatrix} 0 & A \\ I & 0 \end{bmatrix}, \qquad Z_{k+1} = \frac{1}{2}\left(Z_k + Z_k^{-1}\right)$$

Si può dimostrare che le iterate sono nella forma

$Z_k = \begin{bmatrix} 0 & X_k \\ Y_k & 0 \end{bmatrix}$ con X_k, Y_k che soddisfano la ricorsione vista sopra.

Calcoliamo la derivata di Fréchet anche di questa iterazione:

$$\begin{cases} X_{k+1} = \frac{1}{2}(X_k + Y_k^{-1}) \\ Y_{k+1} = \frac{1}{2}(Y_k + X_k^{-1}) \end{cases} \qquad L_{F,\begin{bmatrix} X \\ Y \end{bmatrix}}\begin{bmatrix} E \\ F \end{bmatrix} = \dots$$

$$F\left(\begin{bmatrix} X \\ Y \end{bmatrix}\right) = \begin{bmatrix} \frac{1}{2}(X+Y^{-1}) \\ \frac{1}{2}(Y+X^{-1}) \end{bmatrix} \qquad F\left(\begin{bmatrix} X+E \\ Y+F \end{bmatrix}\right) = \dots$$

$$L_{F,\begin{bmatrix} X \\ Y \end{bmatrix}}\begin{bmatrix} E \\ F \end{bmatrix} = \begin{bmatrix} \frac{1}{2}(E - Y^{-1}FY^{-1}) \\ \frac{1}{2}(F - X^{-1}EX^{-1}) \end{bmatrix}$$

$$L_{F,\begin{bmatrix} A^{\frac{1}{2}} \\ A^{-\frac{1}{2}} \end{bmatrix}}\begin{bmatrix} E \\ F \end{bmatrix} = \begin{bmatrix} \frac{1}{2}(E - A^{\frac{1}{2}}FA^{\frac{1}{2}}) \\ \frac{1}{2}(F - A^{-\frac{1}{2}}EA^{-\frac{1}{2}}) \end{bmatrix}$$

← operatore lineare da uno spazio di dimensione $2n^2$ in sé

Si riesce a dimostrare che $\left(L_{F,\begin{bmatrix} A^{1/2} \\ A^{-1/2} \end{bmatrix}}\right)^2 = L_{F,\begin{bmatrix} A^{\frac{1}{2}} \\ A^{-\frac{1}{2}} \end{bmatrix}}$: infatti:

$$\left(L_{F,\begin{bmatrix} A^{\frac{1}{2}} \\ A^{-\frac{1}{2}} \end{bmatrix}}\right)^2\begin{bmatrix} E \\ F \end{bmatrix} = \begin{bmatrix} \frac{1}{2}\left(\frac{1}{2}(E - A^{\frac{1}{2}}FA^{\frac{1}{2}}) - A^{\frac{1}{2}}\left(\frac{1}{2}(F - A^{-\frac{1}{2}}EA^{-\frac{1}{2}})\right)A^{\frac{1}{2}}\right) \\ " \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{4}E - \frac{1}{4}A^{\frac{1}{2}}FA^{\frac{1}{2}} - \frac{1}{4}A^{\frac{1}{2}}FA^{\frac{1}{2}} + \frac{1}{4}E \\ " \end{bmatrix} = L_{F,\begin{bmatrix} A^{\frac{1}{2}} \\ A^{-\frac{1}{2}} \end{bmatrix}}\begin{bmatrix} E \\ F \end{bmatrix}$$

Visto che l'operatore è idempotente, $K^2 = K$, e gli autoval. sono $\varnothing$ e 1 (semplici, perché $K^2 - K = 0$).

Calcolare funzioni di matrici larghe e sparse.

$A \in \mathbb{C}^{n \times n}$ sparsa $\quad nnz(A)$

Poiché $f(A)$ tipicamente è densa, vorremmo metodi per calcolare $f(A)b$ dato un vettore b.

Idea 1: rimpiazzo f con un polinomio: se $p(x)$ e $f(x)$ sono vicini, $p(A)b$ e $f(A)b$ mi aspetto siano vicini, e $p(A)b$ si calcola con prodotti matrice sparsa/vettore.

Idea 2:
$$f(A)b = \frac{1}{2\pi i} \int f(z)(zI - A)^{-1} dz \cdot b$$
$$\approx \sum w_k \frac{1}{2\pi i} f(x_k)(x_k I - A)^{-1} b$$
dato una formula di quadratura con pesi/nodi w_k, x_k

Idea 3: Arnoldi

Def: $K_j(A,b) = \text{span}\{b, Ab, \dots A^{j-1}b\}$

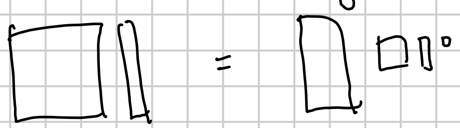
Arnoldi calcola una successione $u_1, u_2, \dots u_j, \dots$ di vettori ortonormali tali che $(u_1, u_2, \dots u_j)$ è una base di $K_j(A,b)$ $\forall j = 1, 2, \dots$

Dato $U_j = [u_1 | u_2 | \dots u_j]$, posso calcolare una versione proiettata di A come
$$A_j = U_j^* A U_j = \begin{bmatrix} & \\ 0 & \end{bmatrix}$$
matrice di Hessenberg i cui coefficienti vengono calcolati dall'algoritmo di Arnoldi insieme agli u_i.

In particolare, $u_1 = \frac{1}{\beta} \cdot b$, con $\beta = \|b\|$. $b = U_j e_1 \beta$

Lemma: per ogni polinomio di grado $\deg p < j$, si ha

$$p(A)b = U_j \cdot p(A_j) e_1 \beta.$$



Dim: per linearità, mi basta farlo per i monomi

$$p(x) = 1, x, x^2, \dots x^{j-1}.$$

$p(x)=1$: $b = U_j e_1 \beta$

$p(x)=x$: $$U_j A_j e_1 \beta = U_j U_j^T A \underbrace{U_j e_1 \beta} = U_j U_j^T \underbrace{Ab} = Ab$$

perché $U_j U_j^*$ è la matrice di proiez. ort. su $\mathrm{Im}\, U_j$,

e $Ab \in \mathrm{Im}\, U_j = K_j(A,b)$

Similmente, per $p(x) = x^{i+1}$ per induzione

$$U_j (A_j)^{i+1} e_1 \beta = U_j U_j^T A \underbrace{\boxed{U_j A_j^i e_1 \beta}}_{= A^i b} = U_j U_j^T A^{i+1} b$$

$= A^{i+1} b$ se $A^{i+1} b \in K_j(A,b)$, cioè $i+1 < j$.

Questo suggerisce di usare una formula analoga per le funzioni di matrici:

$$f(A)b \approx U_j f(A_j) \cdot e_1 \beta$$

Notiamo che

$f(A_j) = p_j(A_j)$, dove p_j è il polinomio che interpola f sullo spettro di A_j

Per il lemma appena dimostrato,

$$U_j f(A_j) e_1 \beta = U_j p_j(A_j) e_1 \beta = p_j(A) b.$$

Quindi non abbiamo $f(A) = p(A)$, dove p è il polinomio di interpolazione sullo spettro di A, ma un altro polinomio.

Gli autovalori di A_j sono i cosiddetti <u>valori di Ritz</u> di A, e approssimano gli autovalori più esterni dello spettro di A.