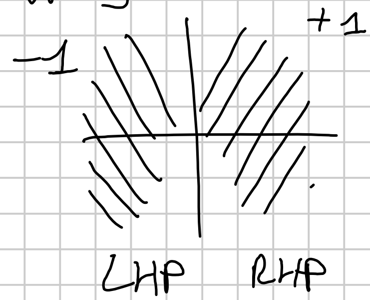
Note Title 2024-04-19

[Lezioni eliminate: 22 aprile, 3 maggio]

$$\operatorname{sign}(x) = \begin{cases} 1 & \operatorname{Re}(x) > 0 \\ -1 & \operatorname{Re}(x) < 0 \\ \text{non def.} & \operatorname{Re}(x) = 0 \end{cases}$$



L'iterazione
$$\begin{cases} X_0 = A \\ X_{k+1} = \frac{1}{2}\left(X_k + X_k^{-1}\right) \end{cases}$$
converge a $\operatorname{sign}(A)$ per ogni A che non ha autoval. immaginari puri. (quadraticamente)

Cominciamo a studiare la versione scalare di questa iterazione
$$x_{k+1} = f(x_k) \qquad f(x) = \frac{1}{2}\left(x + \frac{1}{x}\right) = \frac{x^2+1}{2x}.$$

Oss: questa iterazione è quello che si ottiene applicando il metodo di Newton all'equazione $x^2-1=0$, che ha soluzioni $x = \pm 1$:
$$\boxed{x_{k+1} = x_k - \frac{x_k^2-1}{2x_k}} = \frac{2x_k^2-(x_k^2-1)}{2x_k} = \frac{x_k^2+1}{2x_k}$$

Oss: i punti fissi sono ± 1: da $\boxed{}$, x_k punto fisso $\Leftrightarrow x_k^2 - 1 = 0$.

In particolare, la convergenza è localmente quadratica.

Studiamo la convergenza globale tramite il cambio di variabile
$$y = \frac{x-1}{x+1}$$

Definiamo $\quad y_k = \dfrac{x_k - 1}{x_k + 1}$

Allora

$$y_{k+1} = \frac{x_{k+1} - 1}{x_{k+1} + 1} = \frac{\frac{1}{2}\left(x_k + \frac{1}{x_k}\right) - 1}{\frac{1}{2}\left(x_k + \frac{1}{x_k}\right) + 1} = \frac{x_k^2 + 1 - 2x_k}{x_k^2 + 1 + 2x_k}$$

$$= \frac{(x_k - 1)^2}{(x_k + 1)^2} = y_k^2$$

$$y_k = \frac{x_k - 1}{x_k + 1} \qquad \begin{tikzcd} x_0 \arrow[r, "\frac{1}{2}(x + \frac{1}{x})"] \arrow[d] & x_1 \arrow[r] \arrow[d] & x_2 \arrow[r] \arrow[d] & x_3 \arrow[r] & \dots \\ y_0 \arrow[r, "y^2"'] & y_1 \arrow[r, "y^2"'] & y_2 \arrow[r, "y^2"'] & y_3 \arrow[r] & \dots \end{tikzcd}$$

Se $|y_0| > 1$, la successione $y_0,\ y_1 = y_0^2,\ y_2 = y_1^2, \dots \to \infty$

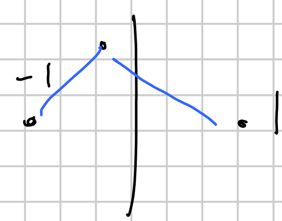
Se $|y_0| < 1$, la successione $y_0,\ y_0^2,\ y_0^4, \dots \to 0$

Se $|y_0| = 1$, $|y_k| = 1 \quad \forall k$

<u>Oss</u>: se $x_0 \in LHP$, $\quad |y_0| = \left|\dfrac{x_0 - 1}{x_0 + 1}\right| = \dfrac{\text{dist}(x_0, 1)}{\text{dist}(x_0, -1)}$

$x_0 \in LHP \iff \text{dist}(x_0, 1) > \text{dist}(x_0, -1)$

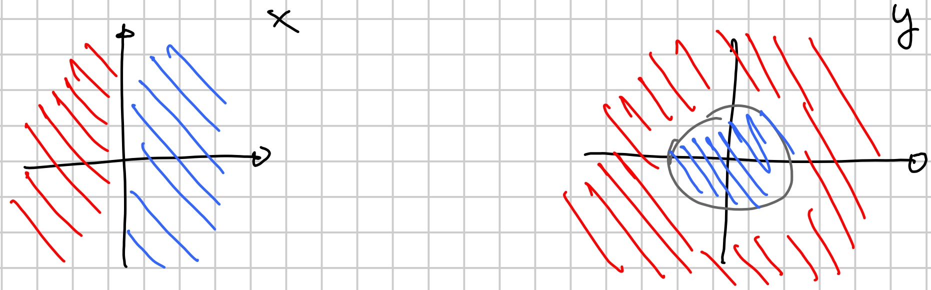
$\iff |y_0| > 1$



$x_0 \in RHP \iff |y_0| < 1$

Abbiamo quindi dimostrato che:

- se $x_0 \in LHP$, la successione $\{x_k\}$ converge a -1
- se $x_0 \in RHP$, la successione $\{x_k\}$ converge a 1
- se $x_0 \in$ asse immaginario, $\{x_k\}$ resta sempre sull'asse immaginario



Versione matriciale di Newton per il segno:

Caso facile: se X_0 è diagonalizzabile, $X_0 = V \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} V^{-1}$

$$f(X_0) = \frac{1}{2}\left(X_0 + X_0^{-1}\right) = V \begin{bmatrix} f(\lambda_1) & & \\ & \ddots & \\ & & f(\lambda_n) \end{bmatrix} V^{-1}$$

$$f^k(X_0) = V \begin{bmatrix} f^k(\lambda_1) & & \\ & \ddots & \\ & & f^k(\lambda_n) \end{bmatrix} V^{-1}$$

$$\lim_{k \to \infty} f^k(X_0) = V \begin{bmatrix} \operatorname{sign}(\lambda_1) & & \\ & \ddots & \\ & & \operatorname{sign}(\lambda_n) \end{bmatrix} V^{-1}.$$

Caso generale:

Teo: Se X_0 non ha autoval. immaginari, la seguente

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

converge a $\operatorname{sign}(X_0)$.

Dim: variante dell'argomento usato sugli scalari:

Chiamiamo $S = \operatorname{sign}(X_0)$

$$Y_k = (X_k - S)(X_k + S)^{-1} = (X_k + S)^{-1}(X_k - S)$$

(commutano perché X_k sono funzioni razionali di X_0, e S è una funzione di matrice).

L'inverso esiste? Gli autoval. di X_k sono

$$f^k(\lambda_0),\ f^k(\lambda_1),\ \dots\ f^k(\lambda_n)$$

Prendiamo una forma di Schur $X_0 = UTU^*$

$$X_k = f^k(X_0) = U \left[\ \nabla\ \right] U^*$$

e sulla diagonale ha $f^k(\lambda_1),\ f^k(\lambda_2),\ \dots\ f^k(\lambda_n)$

$$S = UTU^* = U \left[\ \nabla\ \right] U^*,$$ e sulla diagonale ha

$$\mathrm{sign}(\lambda_1),\ \mathrm{sign}(\lambda_2), \dots\ \mathrm{sign}(\lambda_n)$$

$$X_k + S = U \left[\ \nabla\ \right] U^*$$ e sulla diagonale ha

$$f^k(\lambda_1) + \mathrm{sign}(\lambda_1),\ \dots\ f^k(\lambda_n) + \mathrm{sign}(\lambda_n)$$

Osserviamo che $f^k(\lambda_i)$ e $\mathrm{sign}(\lambda_i)$ stanno tutti nello stesso semipiano: la funzione f manda LHP in LHP, e RHP in RHP, quindi sommando $f^k(\lambda_i)$ e $\mathrm{sign}(\lambda_i)$ resto nello stesso semipiano.

Quindi è ben definito il cambio di variabile

$$Y_k = (X_k - S)(X_k + S)^{-1}.$$

$$Y_k = U\ \nabla\ U^*,$$ e sulla diagonale ha

$$\frac{f^k(\lambda_i) - \mathrm{sign}(\lambda_i)}{f^k(\lambda_i) + \mathrm{sign}(\lambda_i)} \quad \begin{cases} \lambda \in LHP \quad f^k(\lambda_i) \in LHP \text{ e } \left| \dfrac{f^k(\lambda_i)+1}{f^k(\lambda_i)-1} \right| < 1 \\ \lambda \in RHP: \ f^k(\lambda_i) \in RHP \end{cases}$$

che ha modulo strettamente < 1:

$$\left|\frac{f^*(\lambda_i)-1}{f^*(\lambda_i)+1}\right| < 1$$

$$Y_{k+1} = (X_{k+1}-S)(X_{k+1}+S)^{-1} = \left(\tfrac{1}{2}(X_k+X_k^{-1})-S\right)\left(\tfrac{1}{2}(X_k+X_k^{-1})+S\right)^{-1}$$

$$= \left(\left(\tfrac{1}{2}(X_k^2+1)-SX_k\right)X_k^{-1}\right)\left(\left(\tfrac{1}{2}(X_k^2+1)+SX_k\right)X_k^{-1}\right)^{-1}$$

$$= \left(\tfrac{1}{2}(X_k^2+1)-SX_k\right)\cancel{X_k^{-1}}\,\cancel{X_k}\left(\tfrac{1}{2}(X_k^2+1)+SX_k\right)^{-1}$$

$$= (X_k^2+1-2SX_k)(X_k^2+1+2SX_k)^{-1}$$

Notiamo che $S^2=I$, perché $S = V\begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}V^{-1}$

e che $X_k S = S X_k$

Quindi posso scrivere quest'ultima formula ottenuta come

$$(X_k-S)^2\left((X_k+S)^2\right)^{-1} = Y_k^2$$

Questa catena di uguaglianze mostra $Y_{k+1} = Y_k^2$

Y_0 ha tutti autovalori <1 $\Rightarrow$ la successione

$Y_0,\ Y_1 = Y_0^2,\ Y_2 = Y_1^2, \dots$ converge a 0.

Visto che tutte le matrici commutano tra loro,

$$Y_k = (X_k-S)(X_k+S)^{-1} \iff$$

$$X_kY_k + SY_k = X_k - S \iff$$

$$X_k(Y_k-1) = S(-Y_k-1) \iff$$

$$X_k = S(-Y_k-1)(Y_k-1)^{-1} = S(1+Y_k)(1-Y_k)^{-1}$$

$Y_k \to 0$ implica $X_k \to S$.

<u>Oss:</u>
- ci serve davvero inv( ), non solo risolvere sistemi lineari
- su matrici "ben scalate", 10-15 iterazioni bastano, indipendentemente dalla dimensione
- su matrici con norma molto grande o molto piccola, la convergenza è più lenta.

Questo si vede bene dall'iterazione scalare:

se $x \gg 1$, $\frac{1}{2}\left(x+\frac{1}{x}\right) \sim \frac{1}{2}x$, e dimezziamo ad ogni passo. Analogamente

$$x \ll 1 \quad , \quad \frac{1}{2}\left(x+\frac{1}{x}\right) \sim \frac{1}{2x}$$

<u>Soluzione:</u> scalatura. Notiamo che $\operatorname{sign}(X) = \operatorname{sign}(\alpha X)$ per ogni $\alpha \in \mathbb{R}$, $\alpha > 0$

Quindi posso rimpiazzare X_0 con αX_0 per un $\alpha > 0$ scelto opportunamente in modo che gli autovalori abbiano moduli non troppo grandi / piccoli, il più vicino possibile. Ad esempio, se $|\lambda_i|$ sono compresi tra 10^5 e 10^7, vorrò moltiplicare per $\alpha = 10^{-6}$ in modo che diventino tra 10^{-1} e 10^1.

Posso stimare i moduli degli autovalori con norme, oppure posso usare qualche iterazione del metodo delle potenze per avere un'idea della loro dimensioni:

(Potenze su X_0 e su X_0^{-1}, che tanto devo calcolare lo stesso).

Altra alternativa: "determinantal scaling": scelgo α in modo che $\det(\alpha X_0) = 1$.

Stabilità: non è facile dimostrare qualcosa, ma i casi di instabilità dell'algoritmo (ad es. X_0 mal condizionata) corrispondono a casi in cui $\operatorname{sign}(X_0)$ è mal condizionata.

Nella pratica, l'algoritmo funziona bene.