Sistemi lineari – Malcondizionamento

Calcolo Numerico - Ing. Inf. - Lezione 7

Outline

Outline

Esempio 1

Sono dati i due sistemi lineari

$$\begin{pmatrix} 1 & -1 \\ 1 & -1.000001 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & -1 \\ 1 & -1.000001 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0.999999 \\ 1 \end{pmatrix}$$

Le due soluzioni **esatte** sono

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} -10^{-6} \\ -1 \end{pmatrix}$$

Esempio 2

Sono dati i due sistemi lineari

$$\begin{pmatrix} 3 & 4 \\ 3 & 4.00001 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 7 \\ 7.00001 \end{pmatrix}$$

$$\begin{pmatrix} 3 & 4 \\ 3 & 3.99999 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 7 \\ 7.00004 \end{pmatrix}$$

Le due soluzioni **esatte** sono

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \frac{23}{3} \\ -4 \end{pmatrix}$$

In entrambi i precedenti esempi il secondo sistema differisce dal primo per delle **perturbazioni** su elementi della matrice dei coefficienti e/o elementi del vettore dei termini noti

Le perturbazioni sono dell'ordine di 10^{-6} nel primo esempio e dell'ordine di 10^{-5} nel secondo

Queste perturbazioni hanno portato ad avere variazioni delle soluzioni dell'ordine dell'unità

Quando gli errori introdotti nei dati sono amplificati nella soluzione si parla di **malcondizionamento**

Si osserva che anche nello svolgimento di un algoritmo di risoluzione del problema dato, in generale, si introducono perturbazioni sui valori ottenuti

Se la matrice dei coefficienti di un sistema lineare è perturbata significa che andiamo a risolvere un sistema con matrice $A + \delta A$ dove δA rappresenta la matrice di perturbazione

Analogamente, se il vettore dei termini noti di un sistema lineare é perturbato significa che andiamo a risolvere un sistema vettore dei termini noti $b + \delta b$ dove δb rappresenta il vettore di perturbazione

La conseguenza è che anche il vettore incognita x risulterà perturbato per, in effetti, il sistema che si risolve è del tipo

$$(A + \delta A)(x + \delta x) = b + \delta b$$

Supponiamo di non avere perturbazioni sulla matrice A ($\delta A = \mathbf{O}$) per cui il sistema lineare si presenta nella forma

$$A(x + \delta x) = b + \delta b$$

Eseguendo il prodotto a primo membro si ha

$$A\,x + A\,\delta x = b + \delta b$$

Ricordando che $Ax = b$ si ha

$$A\,\delta x = \delta b$$

Nella ipotesi che sia $\det(A) \neq 0$, premoltiplicando per la matrice A^{-1} otteniamo

$$\delta x = A^{-1}\,\delta b$$

Passando alle norme, utilizzando una norma matriciale coerente o compatibile con la norma vettoriale scelta, risulta

$$\|\delta x\| = \|A^{-1}\,\delta b\| \leq \|A^{-1}\|\,\|\delta b\|$$

Momentaneamente, mettiamo da parte questo risultato

Riprendiamo il sistema di partenza $Ax = b$ da cui si ha

$$\|b\| = \|A\,x\| \leq \|A\|\,\|x\|$$

Quindi otteniamo

$$\|x\| \geq \frac{\|b\|}{\|A\|}$$

Adesso, facciamo il rapporto membro a membro tra le relazioni trovate

$$\|\delta x\| \leq \|A^{-1}\|\,\|\delta b\| \quad \text{e} \quad \|x\| \geq \frac{\|b\|}{\|A\|}$$

Il risultato è

$$\frac{\|\delta x\|}{\|x\|} \leq \|A\|\,\|A^{-1}\|\,\frac{\|\delta b\|}{\|b\|}$$

L'ultima maggiorazione trovata ci dice che l'errore relativo su x risulta maggiorato dall'errore relativo su b amplificato dal fattore

$$\|A\|\|A^{-1}\|$$

Definizione

Il fattore di amplificazione si chiama **numero di condizionamento** del sistema lineare o della matrice A e si pone

$$\mu(A) = \|A\|\,\|A^{-1}\|$$

Un sistema lineare (o una matrice A) si dice **malcondizionato** se

$$\mu(A) \gg 1$$

Dire che un sistema lineare è malcondizionato significa dire che il sistema ha la predisposizione ad amplificare notevolmente gli errori relativi introdotti sui dati (o nello sviluppo dell'algoritmo) conducendo ad un errore relativo sulla soluzione con un ordine di grandezza molto superiore

Il numero di condizionamento $\mu(A)$ non può risultare minore di 1

Infatti, ricordando una proprietà delle norme matriciali e il Toerema di Hirsh, si ha

$$\mu(A) = \|A\|\|A^{-1}\| \geq \|A\,A^{-1}\| = \|I\| \geq \rho(I) = 1$$

da cui ricaviamo

$$\mu(A) \geq 1$$

Tornando al caso generale del sistema lineare perturbato

$$(A + \delta A)\,(x + \delta x) = b + \delta b$$

si può dimostrare il teorema che segue

Teorema

Nell'ipotesi che la matrice $A + \delta A$ sia non singolare e che, rispetto ad una data norma, sia $\|\delta A\| < 1/\|A^{-1}\|$, vale la relazione:

$$\frac{\|\delta x\|}{\|x\|} \leq \frac{\mu(A)}{1 - \mu(A)\frac{\|\delta A\|}{\|A\|}} \left(\frac{\|\delta A\|}{\|A\|} + \frac{\|\delta b\|}{\|b\|} \right)$$

(se $\delta A = \mathbf{O}$ si ritrova la maggiorazione determinata in precedenza)

Quella che abbiamo visto è una maggiorazione a priori dell'errore relativo sulla soluzione del sistema lineare se risulta nota una stima della perturbazione sui dati o la perturbazione introdotta utilizzando un metodo diretto

Di validità più generale è invece una maggiorazione a posteriori

Sia $\tilde{x}$ la soluzione ottenuta per il sistema $Ax = b$ risolto con un qualunque metodo e si abbia

$$b - A\,\tilde{x} = r$$

dove r è il vettore residuo

In corrispondenza alla soluzione esatta, r risulta nullo, cioè si ha $b - Ax = 0$
Eseguendo la differenza tra i due sistemi otteniamo

$$A\,(\tilde{x} - x) = -r \quad \Longrightarrow \quad \tilde{x} - x = -A^{-1}\,r$$

Per una qualunque coppia di norme matriciale e vettoriale coerenti vale la disugusglianza

$$\|\tilde{x} - x\| \leq \|A^{-1}\|\|r\|$$

Ricordando che, da $Ax = b$, si ha

$$\|x\| \geq \frac{\|b\|}{\|A\|}$$

Dalle ultime due relazioni segue la maggiorazione

$$\frac{\|\tilde{x} - x\|}{\|x\|} \leq \mu(A)\,\frac{\|r\|}{\|b\|}$$

L'ultima maggiorazione dell'errore relativo sulla soluzione mostra che la dipendenza dell'errore finale da $\mu(A)$ è un fatto generale e mette in guardia dal ritenere buona una approssimazione $\tilde{x}$ quando il corrispondente residuo sia piccolo

In generale non si conosce la matrice inversa di A e quindi $\mu(A)$
Tuttavia la maggiorazione può essere effettivamente utilizzata ricorrendo ad appositi procedimenti per il calcolo approssimato di $\mu(A)$

Esempio 3

Sia $A \in \mathbb{C}^{n \times n}$ una matrice hermitiana
Si vuole determinare il suo numero di condizionamento calcolato in norma 2 e cioé $\mu_2(A)$

Abbiamo visto che, nel caso di matrici hermitiane, risulta

$$\|A\|_2 = \rho(A)$$

Segue che

$$\mu_2(A) = \|A\|_2 \, \|A^{-1}\|_2 = \rho(A) \, \rho(A^{-1})$$

Esempio 3

Ricordando che gli autovalori della matrice inversa sono i reciproci degli autovalori della matrice data, otteniamo

$$\mu_2(A) = \max_{1 \le i \le n} |\lambda_i| \, \frac{1}{\min_{1 \le i \le n} |\lambda_i|} = \frac{\max_{1 \le i \le n} |\lambda_i|}{\min_{1 \le i \le n} |\lambda_i|}$$