

LEZIONE 20

Condizionamento di matrici e sistemi lineari

Grazie allo strumento delle norme vettoriali e matriciali indotte introdotto nella scorsa lezione, possiamo "misurare" gli errori su vettori e matrici.

In questa lezione, utilizzando le due disuguaglianze fondamentali per le norme matriciali indotte

$$(i) \|Ax\| \leq \|A\| \cdot \|x\| \text{ (1° disuguaglianza fondamentale)}$$

$$(ii) \|AB\| \leq \|A\| \cdot \|B\| \text{ (2° disuguaglianza fondamentale)}$$

vedremo come sia possibile stimare la "risposta" di un sistema lineare non singolare agli errori sui dati

$$Ax = b, \quad A \in \mathbb{R}^{n \times n}, \quad b, x \in \mathbb{R}^n, \quad \det(A) \neq 0$$

Nelle applicazioni infatti sia la matrice A che il vettore termine noto b possono essere (e di solito sono) affetti da errori (in ogni caso gli arrotondamenti nel sistema floating-point, errori di misura sperimentale che sono di solito molto più grandi degli errori di arrotondamento, ...).

Quindi in pratica ci si trova a risolvere un sistema "perturbato"

$$\tilde{A}\tilde{x} = \tilde{b}$$

dove $\tilde{b} = b + \delta b$, $\tilde{A} = A + \delta A$ cioè

$$\tilde{b} = \begin{pmatrix} \tilde{b}_1 \\ \vdots \\ \tilde{b}_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} + \begin{pmatrix} \delta b_1 \\ \vdots \\ \delta b_n \end{pmatrix}$$

$$\tilde{A} = \begin{pmatrix} \tilde{a}_{11} & \dots & \tilde{a}_{1n} \\ \vdots & & \vdots \\ \tilde{a}_{n1} & \dots & \tilde{a}_{nn} \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} + \begin{pmatrix} \delta a_{11} & \dots & \delta a_{1n} \\ \vdots & & \vdots \\ \delta a_{n1} & \dots & \delta a_{nn} \end{pmatrix}$$

sono il vettore termine noto "perturbato" e la matrice "perturbata" a cui corrisponde un vettore soluzione "perturbato" $\tilde{x} = x + \delta x$

$$\tilde{x} = \begin{pmatrix} \tilde{x}_1 \\ \vdots \\ \tilde{x}_n \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} \delta x_1 \\ \vdots \\ \delta x_n \end{pmatrix}$$

Il problema di stimare δx in funzione di δb e δA è detto problema di "condizionamento" del sistema lineare corrispondente ad un'analisi di stabilità "a monte" (stabilità del problema), prima di applicare algoritmi di calcolo della soluzione. Cominciamo da una stima di base

PROPOSIZIONE 1 (errore sulla soluzione di una sistema lineare generato da una perturbazione del termine noto)

Sia $A \in \mathbb{R}^{n \times n}$ una matrice non singolare, $x \in \mathbb{R}^n$ soluzione del sistema $Ax = b$, $b \neq 0$ e $\tilde{x} = x + \delta x$ soluzione del sistema perturbato $A\tilde{x} = \tilde{b} = b + \delta b$.

Fissata una norma vettoriale $\|\cdot\|$ in \mathbb{R}^n , vale la seguente stima dell'errore "relativo" su x

$$\frac{\|\delta x\|}{\|x\|} \leq K(A) \frac{\|\delta b\|}{\|b\|}$$

dove

$$k(A) = \|A\| \cdot \|A^{-1}\|$$

prodotto della norma indotta di A e A^{-1} , è detto "INDICE (o anche numero) DI CONDIZIONAMENTO" della matrice A (nella norma indotta).

Dimostrazione

Osserviamo che $x = A^{-1}b \neq 0$ quindi ha senso stimare l'errore relativo in norma (cioè l'errore assoluto $\|\delta x\|$ diviso per la "lunghezza" di x , cioè $\|x\|$).

Ora

$$\tilde{x} = x + \delta x = A^{-1}\tilde{b} = A^{-1}(b + \delta b) = A^{-1}b + A^{-1}\delta b$$

da cui otteniamo la stima dell'errore assoluto

$$\|\delta x\| = \|A^{-1}\delta b\| \underset{1^{\circ} \text{dis.fond.}}{\leq} \|A^{-1}\| \cdot \|\delta b\|$$

Per stimare l'errore relativo dobbiamo stimare da sopra $\frac{1}{\|x\|}$, cioè da sotto $\|x\|$.
Siccome x è la soluzione

$$\|b\| = \|Ax\| \underset{1^{\circ} \text{dis.fond.}}{\leq} \|A\| \cdot \|x\|$$

da cui

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

e

$$\frac{1}{\|x\|} \leq \frac{\|A\|}{\|b\|}$$

perciò

$$\frac{\|\delta x\|}{\|x\|} \leq \frac{\|A^{-1}\| \cdot \|\delta b\|}{\|x\|} \leq \|A^{-1}\| \cdot \|A\| \cdot \frac{\|\delta b\|}{\|b\|} = k(A) \cdot \frac{\|\delta b\|}{\|b\|}$$

Fine dimostrazione

Facciamo subito alcune osservazioni:

1) $k(A) \geq 1$: infatti

$$k(A) = \|A^{-1}\| \cdot \|A\| \underset{1^{\circ} \text{dis.fond.}}{\geq} \|A^{-1}A\| = \|I\| = 1$$

Quindi $k(A)$ nella stima ha il ruolo di "amplificatore" dell'errore sui dati.

In molte applicazioni accade che $k(A) \gg 1$: in questi casi si dice che il sistema è MAL CONDIZIONATO, il che significa che piccole perturbazioni sui dati possono portare ad errori molto grandi sulla soluzione.

Non sono infrequenti sistemi di interesse in modelli applicativi per cui $k(A) \approx 10^{20}, 10^{30}$.

In questi casi il sistema è "estremamente" mal condizionato ed amplifica in modo tale perfino gli errori di arrotondamento delle componenti di b , che la soluzione può perdere completamente di significato perchè

$$\frac{\|\delta x\|}{\|x\|} > 1$$

(cioè l'errore diventa maggiore del 100%).

Per queste situazioni estreme (e comunque quando $k(A) \cdot \frac{\|\delta b\|}{\|b\|} > 1$) non ha molto senso calcolare la soluzione del sistema con algoritmi classici (che porterebbero alla soluzione perturbata con un errore inaccettabile) ma vanno costruiti metodi particolari che "limitano" la perdita di precisione (come analogia, si pensi alla minimizzazione dell'errore in un altro classico problema molto instabile, quello del calcolo della derivata di una funzione perturbata; faremo un cenno facoltativo a questo tipo di approccio alla fine della lezione).

Quando invece la situazione non è così estrema e $k(A) \cdot \frac{\|\delta b\|}{\|b\|}$ è ancora abbastanza piccolo, è comunque importante stimare $k(A)$ in qualche norma per avere un'idea di quanta precisione ci si aspetta di poter perdere rispetto alla precisione con cui sono noti i dati.

In effetti $k(A)$ dipende dalla norma indotta che si usa: ad esempio chiameremo $k_{\infty}(A)$ l'indice di condizionamento rispetto alla norma ∞ , $k_2(A)$ quello rispetto alla norma 2 (indotte dalla corrispondenti norme vettoriali).

In Matlab (come in tutti gli ambienti evoluti di calcolo) esistono funzioni predefinite per stimare l'ordine di grandezza dell'indice di condizionamento, ad es. in Matlab $\text{cond}(A) \approx K_2(A) = \|A\|_2 \|A^{-1}\|_2$

e $\text{cond}(A, \text{Inf}) \approx K_\infty(A) = \|A\|_\infty \|A^{-1}\|_\infty$.

Vale la pena di fare subito un semplice esempio, per fissare le idee.

ESEMPIO

Consideriamo il sistema 2×2

$$\begin{cases} 7x_1 + 10x_2 = b_1 \\ 5x_1 + 7x_2 = b_2 \end{cases}$$

con matrice (e inversa)

$$A = \begin{pmatrix} 7 & 10 \\ 5 & 7 \end{pmatrix}, \quad A^{-1} = \begin{pmatrix} -7 & 10 \\ 5 & -7 \end{pmatrix}$$

Prendiamo $b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0.7 \end{pmatrix}$ a cui corrisponde la soluzione $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = A^{-1}b = \begin{pmatrix} 0 \\ 0.1 \end{pmatrix}$ e consideriamo il sistema perturbato

$A\tilde{x} = \tilde{b} = \begin{pmatrix} \tilde{b}_1 \\ \tilde{b}_2 \end{pmatrix} = \begin{pmatrix} 1.01 \\ 0.69 \end{pmatrix}$ con soluzione $\tilde{x} = \begin{pmatrix} \tilde{x}_1 \\ \tilde{x}_2 \end{pmatrix} = A^{-1}\tilde{b} = \begin{pmatrix} -0.17 \\ 0.22 \end{pmatrix}$ (soluzioni calcolabili "a mano" per sostituzione)

Si vede che $\delta b = \tilde{b} - b = \begin{pmatrix} 0.01 \\ -0.01 \end{pmatrix}$ e

$$\frac{\|\delta b\|_\infty}{\|b\|_\infty} = \frac{\max\{|\delta b_1|, |\delta b_2|\}}{\max\{|b_1|, |b_2|\}} = \frac{0.01}{1} = 10^{-2} = 1\%$$

mentre $\delta x = \tilde{x} - x = \begin{pmatrix} -0.17 \\ 0.12 \end{pmatrix}$ e

$$\frac{\|\delta x\|_\infty}{\|x\|_\infty} = \frac{\max\{|\delta x_1|, |\delta x_2|\}}{\max\{|x_1|, |x_2|\}} = \frac{0.17}{0.1} = 1.7 = 170\%$$

Quindi un errore relativo ($\|\cdot\|_\infty$) dell'1% su b ha come effetto un errore relativo del 170% su x , rendendo la soluzione \tilde{x} inaccettabile.

Il motivo è ben spiegato calcolando $K_\infty(A) = \|A\|_\infty \|A^{-1}\|_\infty$. Infatti

$$\|A\|_\infty = \max \left\{ \begin{array}{l} \text{somme moduli 1° riga} \\ 10 + 7 \end{array}, \begin{array}{l} \text{somme moduli 2° riga} \\ 5 + 7 \end{array} \right\} = 17$$

$$\|A^{-1}\|_\infty = \max\{7 + 10, 5 + 7\} = 17$$

(il fatto che siano uguali è casuale), quindi

$$K_\infty(A) = 17 \cdot 17 = 289$$

Questo valore dell'indice di condizionamento spiega molto bene come mai si siano persi più di 2 ordini nella precisione sulla soluzione rispetto alla precisione sui dati.

E' il caso di osservare che un indice di condizionamento come questo (dell'ordine di 10^2) è in realtà piccolo (non si può parlare qui di sistema mal condizionato), ma quello che conta è il prodotto

$$k(A) \cdot \frac{\|\delta b\|}{\|b\|}$$

che qui è > 1 perchè l'errore sui dati è solo dell'1%.

Con

$$\frac{\|\delta b\|_\infty}{\|b\|_\infty}$$

dell'ordine di 10^{-k} ci aspetteremmo un errore

$$\frac{\|\delta x\|_\infty}{\|x\|_\infty}$$

dell'ordine di 10^{2-k} (quindi ad esempio per $k=6$ avremmo una soluzione \tilde{x} che fa un errore relativo dell'ordine di $10^{-4} = 0.01\%$, che potrebbe essere più che accettabile in molte applicazioni pratiche). Quello che conta è che se abbiamo una stima degli errori sui dati tipo

$$\frac{\|\delta b\|}{\|b\|} \approx \varepsilon$$

allora crescendo l'ordine di grandezza di $k(A)$ sappiamo che possiamo aspettarci un errore

$$\frac{\|\delta x\|}{\|x\|} \lesssim k(A) \cdot \varepsilon$$

e quindi possiamo decidere se la soluzione sarà o meno accettabile

Scriviamo una seconda proprietà (è una stima da sotto) dell'indice di condizionamento
2)

$$k(A) \geq \frac{\max_i |\lambda_i(A)|}{\min_i |\lambda_i(A)|}$$

Con λ_i autovalori. Si osservi che A è invertibile quindi non ha autovalori nulli e $\min_i |\lambda_i(A)|$ non può essere 0.

La dimostrazione è semplice ricordando che per qualsiasi norma indotta vale la stima (localizzazione)

$$|\lambda_i(B)| \leq \|B\|$$

Quindi applicando la stima con $B = A$ otteniamo

$$\|A\| \geq \max_i |\lambda_i(A)|$$

D'altra parte, chi sono gli autovalori di A^{-1} ? Se λ è autovalore di A , allora $\exists x \neq 0$ (autovettore) tale che $Ax = \lambda x$. Moltiplicando a sx per A^{-1}

$$A^{-1}Ax = Ix = x = A^{-1}\lambda x = \lambda A^{-1}x$$

da cui

$$A^{-1}x = \frac{1}{\lambda}x$$

cioè gli autovalori di A^{-1} sono i reciproci degli autovalori di A (ricordiamo che A non ha autovalori nulli) mentre gli autovettori restano gli stessi.

Applicando la stima di localizzazione a $B = A^{-1}$

$$\|A^{-1}\| \geq \max_i |\lambda_i(A^{-1})| = \max_i \left| \frac{1}{\lambda_i(A)} \right| = \frac{1}{\min_i |\lambda_i(A)|}$$

quindi

$$K(A) = \|A\| \cdot \|A^{-1}\| \geq \frac{\max_i |\lambda_i(A)|}{\min_i |\lambda_i(A)|}$$

Osserviamo anche da questa stima da sotto otteniamo subito $K(A) \geq 1$ perché

$$\max_i |\lambda_i(A)| \geq \min_i |\lambda_i(A)|$$

Osserviamo anche che la disuguaglianza può diventare uguaglianza in certi casi: ad esempio, se A è simmetrica si ha $\|A\|_2 = \max_i |\lambda_i(A)|$ e lo stesso vale per A^{-1} che resta simmetrica

$$\|A^{-1}\|_2 = \max_i |\lambda_i(A^{-1})| = \frac{1}{\min_i |\lambda_i(A)|}$$

quindi

$$K_2(A) = \frac{\max_i |\lambda_i(A)|}{\min_i |\lambda_i(A)|}$$

[parte facoltativa, vedi pdf prof.]

Consideriamo ora situazioni in cui ci sono errori sulla matrice.

Tratteremo prima il caso in cui $\delta b = 0$ e poi il caso generale cioè $\delta b \neq 0$ e $\delta A = \tilde{A} - A \neq 0$

PROPOSIZIONE 2 (errore sulla soluzione di un sistema lineare generato da una perturbazione della matrice)

Sia $A \in \mathbb{R}^{n \times n}$ una matrice non singolare, $x \in \mathbb{R}^n$ soluzione del sistema $Ax = b$, $b \neq 0$ e $\tilde{x} = x + \delta x$ soluzione del sistema perturbato $\tilde{A}\tilde{x} = b$, $\tilde{A} = A + \delta A$.

Fissata una norma vettoriale $\|\cdot\|$ in \mathbb{R}^n , vale la seguente stima dell'errore "relativo" su x

$$\frac{\|\delta x\|}{\|\tilde{x}\|} \leq k(A) \cdot \frac{\|\delta A\|}{\|A\|}$$

Dimostrazione:

Da $\tilde{A}\tilde{x} = (A + \delta A)(x + \delta x) = b$ otteniamo

$$Ax + A\delta x + \delta A\tilde{x} = b$$

cioè

$$\delta x = A^{-1}(-\delta A\tilde{x}) = -A^{-1}\delta A\tilde{x}$$

Quindi

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

e perciò

$$\frac{\|\delta x\|}{\|\tilde{x}\|} \leq \|A^{-1}\| \cdot \|\delta A\| = \|A\| \cdot \|A^{-1}\| \cdot \frac{\|\delta A\|}{\|A\|} = k(A) \cdot \frac{\|\delta A\|}{\|A\|}$$

c.v.d.

Si noti che in sostanza la stima ottenuta ci dice che la risposta del sistema ad errori sulla matrice è anch'essa determinata dall'indice di condizionamento, che ci permette di quantificare la perdita di precisione.

Passiamo ora al caso generale.

TEOREMA

Sia $A \in \mathbb{R}^{n \times n}$ una matrice non singolare, $x \in \mathbb{R}^n$ soluzione del sistema $Ax = b$, $b \neq 0$ e $\tilde{x} = x + \delta x$ soluzione del sistema perturbato $\tilde{A}\tilde{x} = \tilde{b}$, $\tilde{A} = A + \delta A$, $\tilde{b} = b + \delta b$.

Fissata una norma vettoriale $\|\cdot\|$ in \mathbb{R}^n , vale la seguente stima dell'errore "relativo" su x per $k(A) \cdot \frac{\|\delta A\|}{\|A\|} < 1$

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

[dimostrazione facoltativa, vedi pdf prof] Concludiamo la lezione con qualche cenno (facoltativo) ... [il resto della lezione è facoltativo, vedi il pdf del prof.]