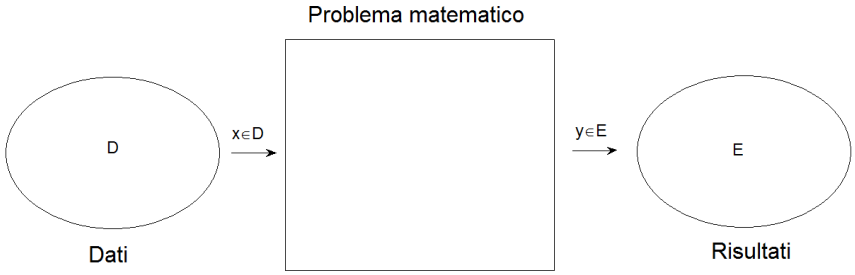
Condizionamento

S. Maset
Dipartimento di Matematica e Geoscienze
Università di Trieste
maset@units.it

April 30, 2020

1 Introduzione

Un *problema matematico* è caratterizzato da un *insieme di possibili dati* D e dal fatto che per ogni *dato* $x \in D$ del problema vi è un corrispondente *risultato* (una corrispondente soluzione) y appartenente ad un insieme E.

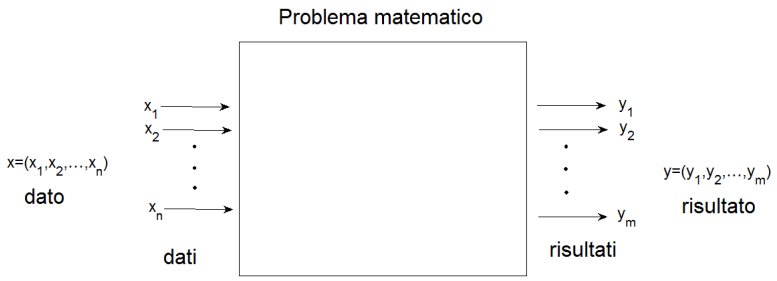


Pertanto, un problema matematico è caratterizzato dalla (e può essere identificato con la) *funzione dato-risultato* $f : D \to E$ che associa ad ogni dato $x \in D$ il corrispondente risultato $y \in E$. Si ha $y = f(x)$.

Si assumerà che i dati siano vettori di $\mathbb{R}^n$, vale a dire $D \subseteq \mathbb{R}^n$, e che i corrispondenti risultati siano vettori di $\mathbb{R}^m$, vale a dire $E \subseteq \mathbb{R}^m$.

In questo modo, se un problema ha molti dati $x_1, \ldots, x_n \in \mathbb{R}$ e molti risultati $y_1, \ldots, y_m \in \mathbb{R}$, questi vengono raggruppati in unico vettore di dati, il dato

$x = (x_1, \ldots, x_n)$, e in un unico vettore di risultati, il risultato $y = (y_1, \ldots, y_m)$.



Esempio. Il problema matematico di risolvere l'equazione di secondo grado

$$ay^2 + by + c = 0,$$

ha come insieme di possibili dati

$$D = \left\{(a, b, c) \in \mathbb{R}^3 : a \neq 0 \ \text{ e } \ b^2 - 4ac \geq 0\right\}$$

e il risultato corrispondente al dato (a, b, c) è la coppia (y_1, y_2) di radici dell'equazione.

Per cui, la funzione dato-risultato $f : D \subseteq \mathbb{R}^3 \to \mathbb{R}^2$ è data da

$$f(a, b, c) = \left(\frac{-b - \sqrt{b^2 - 4ac}}{2a}, \frac{-b + \sqrt{b^2 - 4ac}}{2a} \right), \ (a, b, c) \in D.$$

La *teoria del condizionamento* studia, in un problema matematico caratterizzato da una funzione dato-risultato f, come viene perturbato il risultato $y = f(x)$ a fronte di una perturbazione del dato $x \in D$.

Questa teoria è estremamente importante dal punto di vista applicativo, in quanto un dato non sarà mai noto in manierà esatta, ma si disporrà invece solo di una sua approssimazione.

Si pensi al caso in cui il dato è un vettore le cui componenti derivano da misurazioni di grandezze fisiche. Poichè il dato disponibile (misurato) $\widetilde{x}$ è solo un'approssimazione del dato vero x, anche il risultato disponibile $\widetilde{y} = f(\widetilde{x})$ è solo un'approssimazione del risultato vero $y = f(x)$.

La teoria del condizionamento permette di dedurre l'ordine di grandezza dell'errore che ha il risultato disponibile $\widetilde{y}$ rispetto al risultato vero y a partire dall'ordine di grandezza dell'errore che ha il dato disponibile $\widetilde{x}$ rispetto al dato vero x.

Occorre ora precisare cosa si intende per errore.

2 Errore assoluto e errore relativo

Sia x un numero reale e sia $\widetilde{x}$ una sua approssimazione o una sua perturbazione. L'*errore assoluto* dell'approssimazione $\widetilde{x}$ è

$$\varepsilon_a = \widetilde{x} - x,$$

mentre l'*errore relativo*, definito per $x \neq 0$, è

$$\varepsilon_r = \frac{\varepsilon_a}{x} = \frac{\widetilde{x} - x}{x}. \tag{1}$$

Per applicazioni scientifiche ed ingegneristiche l'errore relativo è la nozione più importante. Un errore assoluto di 5m sull'altezza di una montagna è sicuramente meno grave di un errore assoluto di 5m sull'altezza di un ponte: nel primo caso l'errore assoluto è molto minore che nel secondo caso.

La precisione dell'approssimazione viene colta solo dall'errore relativo. Un errore assoluto di 10^{-5} m $= 10^4$ nm non può essere considerato nè grande nè piccolo finchè non lo si raffronta a qualche lunghezza: è piccolo per lunghezze dell'ordine del metro ed è grande per lunghezze dell'ordine del nanometro. Solo l'errore relativo, che è una quantità adimensionale, può essere considerato grande o piccolo.

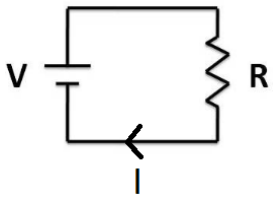
Invece, negli ambiti che hanno a che fare con il denaro è l'errore assoluto la nozione più importante. Ad esempio, un individuo A può concedere che un individuo B gli restituisca solo 10 Euro su un debito di 13 Euro, ma con più difficoltà concederebbe gli vengano restituiti solo 10 000 Euro su un debito di 13 000 Euro, pur essendo l'ammanco relativo il medesimo.

Si noti che la (1) può essere riscritta come

$$\widetilde{x} = x\,(1 + \varepsilon_r)$$

ed è in questa forma che verrà spesso usata.

Consideriamo il circuito elettrico a corrente continua in figura



Supponiamo che il valore misurato della tensione ai capi della batteria sia $V_m = 12$ V con un'incertezza del 1% e che il valore misurato della resistenza sia $R_m = 10$ kΩ con un'incertezza del 2%. Allora l'intensità di corrente elettrica è

$$I_m = \frac{V_m}{R_m} = 1.2 \text{ mA}.$$

Qual è l'incertezza di I_m?

Più specificatamente, denotati con V, R e I i valori veri di tensione, resistenza e intensità di corrente, rispettivamente, si ha

$$V = V_m\,(1 + \widetilde{\varepsilon}_V) = V_m + \widetilde{\varepsilon}_V V_m \text{ con } |\widetilde{\varepsilon}_V| \leq 1\%$$
$$R = R_m\,(1 + \widetilde{\varepsilon}_R) = R_m + \widetilde{\varepsilon}_R R_m \text{ con } |\widetilde{\varepsilon}_R| \leq 2\%$$

e

$$I = I_m\,(1 + \widetilde{\varepsilon}_I) = I_m + \widetilde{\varepsilon}_I I_m,$$

dove

$$I = \frac{V}{R},$$

e ci si domanda quanto grande può essere $|\widetilde{\varepsilon}_I|$.

Cercheremo di rispondere a questo nella prossima sezione.

In questa situazione, gli errori relativi vanno interpretati vedendo V, R e I come approssimazioni o perturbazioni di V_m, R_m e I_m, rispettivamente. Sarebbe più naturale vedere V_m, R_m e I_m come approssimazioni o perturbazioni di V, R e I, rispettivamente, ma la sostanza cambia poco. Infatti, se

$$x = \widetilde{x}\,(1 + \widetilde{\varepsilon}_r)$$

allora

$$\widetilde{x} = x\,(1 + \varepsilon_r)$$

con

$$\varepsilon_r = \frac{\widetilde{x}}{x} - 1 = \frac{1}{1 + \widetilde{\varepsilon}_r} - 1 = -\frac{\widetilde{\varepsilon}_r}{1 + \widetilde{\varepsilon}_r}$$

e, assumendo $|\widetilde{\varepsilon}_r| < 1$,

$$|\varepsilon_r| = \frac{|\widetilde{\varepsilon}_r|}{|1 + \widetilde{\varepsilon}_r|} \leq \frac{|\widetilde{\varepsilon}_r|}{1 - |\widetilde{\varepsilon}_r|} \leq |\widetilde{\varepsilon}_r|\,(1 + |\widetilde{\varepsilon}_r|) = |\widetilde{\varepsilon}_r| + |\widetilde{\varepsilon}_r|^2\,.$$

Per cui,

$$V_m = V\,(1 + \varepsilon_V) \;\text{ con }\; |\varepsilon_V| \leq |\widetilde{\varepsilon}_V| + |\widetilde{\varepsilon}_V|^2 \leq 1\% + 1\%1\% = 1.01\%$$
$$R_m = R\,(1 + \varepsilon_R) \;\text{ con }\; |\varepsilon_R| \leq |\widetilde{\varepsilon}_R| + |\widetilde{\varepsilon}_R|^2 \leq 2\% + 2\%2\% = 2.04\%.$$

3 Analisi con più indici di condizionamento

Consideriamo un problema matematico caratterizzato da una funzione dato-risultato $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$, dove D è un aperto.

Sia $x \in D$ e sia $y = f\,(x)$. Assumiamo ora di perturbare il dato x in $\widetilde{x} \in D$ e sia $\widetilde{y} = f\,(\widetilde{x})$ il risultato perturbato.

Il fatto che D sia aperto non impone vincoli sulla direzione della perturbazione $\widetilde{x} - x$, in quanto x è sempre circondato da una palla di centro x interamente inclusa in D.

Vogliamo studiare quanto la perturbazione sul dato influisca sul risultato.

Nel seguito si assumerà f di classe C^2.

Assumiamo $x_1, \ldots, x_n \neq 0$ e $y \neq 0$. Introduciamo gli errori relativi

$$\varepsilon_i = \frac{\widetilde{x}_i - x_i}{x_i},\; i \in \{1, \ldots, n\}\,,$$

sulle componenti del dato e l'errore relativo

$$\delta = \frac{\widetilde{y} - y}{y}.$$

sul risultato. Vogliamo vedere come l'errore δ dipenda dagli errori ε_i, $i \in \{1, \ldots, n\}$.

3.1 Indici di condizionamento

Essendo f di classe C^2, si ha

$$\begin{aligned} \widetilde{y} - y &= f(\widetilde{x}) - f(x) \\ &= \nabla f(x)^T (\widetilde{x} - x) + R(\widetilde{x}, x) \\ &= \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x) \cdot (\widetilde{x}_i - x_i) + R(\widetilde{x}, x), \end{aligned}$$

dove il resto $R(\widetilde{x}, x)$ dello sviluppo di Taylor è tale che

$$R(\widetilde{x}, x) = O\left(\|\widetilde{x} - x\|^2\right), \ \|\widetilde{x} - x\| \to 0, \tag{2}$$

dove $\| \cdot \|$ è una norma arbitraria su $\mathbb{R}^n$. Questo vuol dire

$$\exists d > 0 \ \exists C \geq 0 \ \forall \widetilde{x} \in D : \ \|\widetilde{x} - x\| \leq d \Rightarrow |R(\widetilde{x}, x)| \leq C \|\widetilde{x} - x\|^2.$$

In Analisi II si è visto che (2) vale quando si usa $\| \cdot \| = \| \cdot \|_2$. Dall'equivalenza delle norme su $\mathbb{R}^n$, si ha che (2) vale per una norma arbitraria $\| \cdot \|$ su $\mathbb{R}^n$. Usando la norma $\| \cdot \|_\infty$, la (2) dice che: esistono $d > 0$ e $C \geq 0$ tali che

$$|R(\widetilde{x}, x)| \leq C \|\widetilde{x} - x\|_\infty^2, \text{ per } \widetilde{x} \in D \text{ tale che } \|\widetilde{x} - x\|_\infty \leq d.$$

Quindi

$$\begin{aligned} \delta &= \frac{\widetilde{y} - y}{y} = \frac{\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x) \cdot \underbrace{(\widetilde{x}_i - x_i)}_{=\varepsilon_i x_i} + R(\widetilde{x}, x)}{y} \\ &= \sum_{i=1}^{n} \frac{\frac{\partial f}{\partial x_i}(x)\, x_i}{y} \cdot \varepsilon_i + \frac{R(\widetilde{x}, x)}{y}. \end{aligned}$$

Mostriamo ora che il termine $\frac{R(\widetilde{x}, x)}{y}$ può essere trascurato.

Introducendo il vettore $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)$, si ottiene

$$\begin{aligned} \|\widetilde{x} - x\|_\infty &= \max_{i \in \{1, \ldots, n\}} |\widetilde{x}_i - x_i| = \max_{i \in \{1, \ldots, n\}} |\varepsilon_i x_i| \\ &= \max_{i \in \{1, \ldots, n\}} \underbrace{|\varepsilon_i|\, |x_i|}_{\leq \|\varepsilon\|_\infty \|x\|_\infty} \leq \|\varepsilon\|_\infty \|x\|_\infty. \end{aligned}$$

Per cui, se

$$\|\varepsilon\|_\infty \leq \frac{d}{\|x\|_\infty}$$

e quindi

$$\|\widetilde{x} - x\|_\infty \leq \|\varepsilon\|_\infty \|x\|_\infty \leq \frac{d}{\|x\|_\infty} \|x\|_\infty = d,$$

si ha

$$\left|\frac{R(\widetilde{x}, x)}{y}\right| = \frac{|R(\widetilde{x}, x)|}{|y|} \leq \frac{C \|\widetilde{x} - x\|_\infty^2}{|y|} \leq \frac{C \|x\|_\infty^2}{|y|} \|\varepsilon\|_\infty^2 .$$

Essendo gli errori relativi ε_i, $i \in \{1, \ldots, n\}$, numeri piccoli, si può trascurare il termine $\frac{R(\widetilde{x},x)}{y}$ rispetto agli altri termini in

$$\delta = \sum_{i=1}^{n} \frac{\frac{\partial f}{\partial x_i}(x)\, x_i}{y} \cdot \varepsilon_i + \frac{R(\widetilde{x}, x)}{y}$$

in quanto esso è maggiorato in modulo da un multiplo di $\|\varepsilon\|_\infty^2 = \left(\max_{i \in \{1,\ldots,n\}} |\varepsilon_i|\right)^2$, mentre gli altri termini sono multipli degli errori ε_i.

Scriviamo allora

$$\delta \doteq \sum_{i=1}^{n} K_i(f, x) \cdot \varepsilon_i, \tag{3}$$

dove

$$K_i(f, x) := \frac{\frac{\partial f}{\partial x_i}(x)\, x_i}{y} = \frac{\frac{\partial f}{\partial x_i}(x)\, x_i}{f(x)},\ i \in \{1, \ldots, n\},$$

e il segno $\doteq$ significa uguale a meno di un termine il cui valore assoluto è minore o uguale di una costante volte $\|\varepsilon\|_\infty^2$, per $\|\varepsilon\|_\infty = \max_{i \in \{1,\ldots,n\}} |\varepsilon_i|$ sufficientemente piccolo: in (3), i due membri sono uguali a meno del termine

$$\frac{R(\widetilde{x}, x)}{y}$$

per il quale

$$\left|\frac{R(\widetilde{x}, x)}{y}\right| \leq \frac{C \|x\|_\infty^2}{|y|} \|\varepsilon\|_\infty^2 \text{ per } \|\varepsilon\|_\infty \leq \frac{d}{\|x\|_\infty}.$$

I numeri $K_i(f, x)$, $i \in \{1, \ldots, n\}$, sono detti *indici di condizionamento* di f (del problema matematico corrispondente) relativi al dato x e sono definiti per $x \in D$ tale che $x_i \neq 0$ per ogni $i \in \{1, \ldots, n\}$ e $y = f(x) \neq 0$, così da poter considerare gli errore relativi ε_i, $i \in \{1, \ldots, n\}$, e l'errore relativo δ.

Gli indici di condizionamento indicano come l'errore relativo sul risultato y sia condizionato dagli (sia sensibile agli) errori relativi sui dati $x_1, \ldots, x_n$.

Esercizio. Se $x_1, \ldots, x_n, y$ sono grandezze fisiche con delle dimensioni, qual è la dimensione degli indici di condizionamento?

3.1.1 Il caso $n = 1$

Nel caso $n = 1$, si ha

$$\delta \doteq K(f, x) \cdot \varepsilon,$$

dove

$$\varepsilon = \frac{\widetilde{x} - x}{x}$$

e

$$K(f, x) = \frac{f'(x)\, x}{f(x)}.$$

Esempio. Consideriamo la funzione $f : \mathbb{R} \to \mathbb{R}$ data da

$$f(x) = ax, \ x \in \mathbb{R},$$

dove $a \in \mathbb{R} \setminus \{0\}$. Si ha, per $x \in \mathbb{R}$ con $x \neq 0$,

$$K(f, x) = \frac{f'(x)\, x}{f(x)} = \frac{ax}{ax} = 1.$$

Quindi si ha

$$\delta \doteq \varepsilon.$$

In questo caso si ha

$$\delta = \varepsilon$$

come può essere verificato direttamente:

$$\delta = \frac{f(\widetilde{x}) - f(x)}{f(x)} = \frac{a\widetilde{x} - ax}{ax} = \frac{\widetilde{x} - x}{x} = \varepsilon.$$

Esercizio. Siano $f : D \subseteq \mathbb{R} \to \mathbb{R}$ e $g : E \subseteq \mathbb{R} \to \mathbb{R}$, con D e E aperti e $f(D) \subseteq E$. Consideriamo la funzione composta $g \circ f : D \to \mathbb{R}$. Provare che per $x \in D$ tale che $x \neq 0$, $f(x) \neq 0$ e $g(f(x)) \neq 0$ si ha

$$K(g \circ f, x) = K(g, f(x)) \cdot K(f, x).$$

Esercizio. Sia $f : D \subseteq \mathbb{R} \to \mathbb{R}$ strettamente monotona, dove D è un aperto. Consideriamo la funzione inversa $f^{-1} : E = f(D) \to D$. Provare che per $x \in D$ tale che $x \neq 0$ e $f(x) \neq 0$ si ha

$$K(f^{-1}, f(x)) = \frac{1}{K(f, x)}.$$

Suggerimento: usare l'esercizio precedente.

Esercizio. Provare che tra l'indice di condizionamento di $f : D \subseteq \mathbb{R} \to \mathbb{R}$ e gli indici di condizionamento di

$$g(x) = af(x), \ x \in D,$$

e

$$h(x) = f(ax),\ x \in \frac{1}{a}D,$$

dove $a \neq 0$, sussistono le relazioni

$$K(g,x) = K(f,x),\ x \in D \text{ con } x \neq 0 \text{ e } f(x) \neq 0$$

e

$$K(h,x) = K(f,ax),\ x \in \frac{1}{a}D \text{ con } x \neq 0 \text{ e } f(ax) \neq 0.$$

Esercizio. Siano $f, g : D \subseteq \mathbb{R} \to \mathbb{R}$, con D aperto. Consideriamo le funzioni prodotto e somma $fg, f+g : D \to \mathbb{R}$. Provare che per $x \in D$ tale che $x \neq 0$, $f(x) \neq 0$ e $g(x) \neq 0$ si ha

$$K(fg,x) = K(f,x) + K(g,x)$$

e per $x \in D$ tale che $x \neq 0$, $f(x) \neq 0$, $g(x) \neq 0$ e $f(x) + g(x) \neq 0$ si ha

$$K(f+g,x) = \frac{f(x)}{f(x)+g(x)}K(f,x) + \frac{g(x)}{f(x)+g(x)}K(g,x).$$

Mostrare inoltre che se $f(x)$ e $g(x)$ hanno lo stesso segno, allora

$$|K(f+g,x)| \leq \max\{|K(f,x)|, |K(g,x)|\}.$$

3.2 Buon e mal condizionamento

La funzione f (il problema matematico corrispondente) si dice *ben condizionata* sul dato x se tutti gli indici di condizionamento di f relativi al dato x hanno ordine di grandezza non superiore all'unità.

Formalmente, l'ordine di grandezza di un numero a con *rappresentazione normalizzat*a in base 10 (o *rappresentazione scientifica*)

$$a = \pm A \cdot 10^p,$$

dove $A \in [1, 10)$ e $p \in \mathbb{Z}$, è la potenza 10^p. Quindi dire che a ha ordine di grandezza non superiore all'unità significa $p \leq 0$, cioè $|a| < 10$.

Tuttavia non vogliamo essere così formali: il dire che a ha ordine di grandezza non superiore all'unità qui semplicemente significa: a non è un numero grande, cioè non si ha $|a| \gg 1$. Vi è un po' di imprecisione voluta in questo, nel senso che tutti accordano nel dire che $0.1, 1, 10$ sono numeri non grandi e $10^3, 10^4, 10^5$ sono numeri grandi, ma i numeri che stanno in mezzo possono essere interpretati sia come non grandi che come grandi.

Per cui, dire che f è ben condizionata sul dato x significa

$$\forall i \in \{1, \ldots, n\} :\ \text{non } |K_i(f,x)| \gg 1.$$

Se f è ben condizionata sul dato x, allora da

$$\delta \doteq \sum_{i=1}^{n} K_i\left(f, x\right) \cdot \varepsilon_i$$

segue che l'errore relativo δ sul risultato ha ordine di grandezza non superiore al massimo ordine di grandezza degli errori relativi ε_i, $i \in \{1, \ldots, n\}$, sui dati. Qui, con questo intendiamo dire che non si ha

$$|\delta| \gg \|\varepsilon\|_\infty = \max_{i \in \{1,\ldots,n\}} |\varepsilon_i|.$$

Infatti,

$$\begin{aligned} |\delta| &\doteq \left| \sum_{i=1}^{n} K_i\left(f, x\right) \cdot \varepsilon_i \right| \leq \sum_{i=1}^{n} \left| K_i\left(f, x\right) \cdot \varepsilon_i \right| = \sum_{i=1}^{n} \left| K_i\left(f, x\right) \right| \cdot |\varepsilon_i| \\ &\leq \sum_{i=1}^{n} \left| K_i\left(f, x\right) \right| \cdot \|\varepsilon\|_\infty = \left(\sum_{i=1}^{n} \left| K_i\left(f, x\right) \right| \right) \cdot \|\varepsilon\|_\infty . \end{aligned}$$

e

$$\sum_{i=1}^{n} \left| K_i\left(f, x\right) \right|$$

non è un numero grande. Questo naturalmente risulta vero supponendo che n non sia grande, come assumeremo nel seguito.

La funzione f (il problema matematico corrispondente) si dice *mal condizionata* sul dato x se non è ben condizionata sul dato x, vale a dire se esiste un indice di condizionamento di f relativo al dato x che ha ordine di grandezza superiore all'unità, vale a dire se

$$\exists i \in \{1, \ldots, n\} : \ |K_i(f, x)| \gg 1.$$

Se f è mal condizionata sul dato x, allora "in generale" si ha che δ ha ordine di grandezza superiore al massimo ordine di grandezza degli ε_i, $i \in \{1, \ldots, n\}$, vale a dire "in generale" si ha che

$$|\delta| \gg \|\varepsilon\|_\infty = \max_{i \in \{1,\ldots,n\}} |\varepsilon_i|.$$

Specifichiamo meglio l'uso di "in generale" fatto sopra.

Se esiste $j \in \{1, \ldots, n\}$ tale che $|K_j\left(f, x\right)| \gg 1$, allora, per errori ε_i, $i \in \{1, \ldots, n\}$, tali che $\varepsilon_j \neq 0$ e $\varepsilon_i = 0$ per $i \neq j$, si ha

$$\delta \doteq K_j\left(f, x\right) \cdot \varepsilon_j$$

e quindi $|\delta| \gg |\varepsilon_j| = \|\varepsilon\|_\infty$. Quindi, in tal caso, f è mal condizionata sul dato x e $|\delta| \gg \|\varepsilon\|_\infty$.

D'altra parte, se esistono $j, k \in \{1, \ldots, n\}$ con $j \neq k$ tali che $K_j(f,x) = K_k(f,x)$ e $|K_j(f,x)| = |K_k(f,x)| \gg 1$, allora, per errori ε_i, $i \in \{1, \ldots, n\}$, tali che $\varepsilon_j = -\varepsilon_k \neq 0$ e $\varepsilon_i = 0$ per $i \neq j, k$, si ha

$$\delta \doteq K_j(f,x) \cdot \varepsilon_j + K_k(f,x) \cdot \varepsilon_k = 0.$$

Quindi, in tal caso, f è mal condizionata sul dato x ma non si ha $|\delta| \gg \|\varepsilon\|_\infty$.

Esercizio. Qual è l'ordine di grandezza di δ?

Il motivo per cui, in caso di mal condizionamento, si può anche non avere $|\delta| \gg \|\varepsilon\|_\infty$ è che termini $K_i(f,x) \cdot \varepsilon_i$ con $|K_i(f,x)| \gg 1$ possono compensarsi nella somma $\sum_{i=1}^{n} K_i(f,x) \cdot \varepsilon_i$.

Esercizio. Nel caso $n = 1$, si può dire che

$$f \text{ mal condizionata sul dato } x \quad \Rightarrow \quad |\delta| \gg \|\varepsilon\|_\infty ?$$

Esercizio. Ancora nel caso $n = 1$, si può dire, per una f strettamente monotona, che

$$f \text{ mal condizionata sul dato } x \quad \Rightarrow \quad f^{-1} \text{ ben condizionata sul dato } f(x)$$

e

$$f \text{ ben condizionata sul dato } x \quad \Rightarrow \quad f^{-1} \text{ mal condizionata sul dato } f(x)?$$

3.3 Indici di condizionamento delle operazioni aritmetiche

Vediamo ora gli indici di condizionamento delle operazioni aritmetiche.

Addizione:

$$f(x) = x_1 + x_2, \ x \in \mathbb{R}^2.$$

Si ha, per $x \in \mathbb{R}^2$ con $x_1, x_2 \neq 0$ e $x_1 \neq -x_2$,

$$\begin{aligned} K_1(f,x) &= \frac{\frac{\partial f}{\partial x_1}(x)\, x_1}{f(x)} = \frac{1 \cdot x_1}{x_1 + x_2} = \frac{x_1}{x_1 + x_2} \\ K_2(f,x) &= \frac{\frac{\partial f}{\partial x_2}(x)\, x_2}{f(x)} = \frac{1 \cdot x_2}{x_1 + x_2} = \frac{x_2}{x_1 + x_2} \end{aligned}$$

e quindi

$$\delta \doteq \frac{x_1}{x_1 + x_2} \cdot \varepsilon_1 + \frac{x_2}{x_1 + x_2} \cdot \varepsilon_2.$$

Sottrazione:

$$f(x) = x_1 - x_2, \ x \in \mathbb{R}^2.$$

Si ha, per $x \in \mathbb{R}^2$ con $x_1, x_2 \neq 0$ e $x_1 \neq x_2$,

$$\begin{aligned} K_1(f,x) &= \frac{\frac{\partial f}{\partial x_1}(x)\, x_1}{f(x)} = \frac{1 \cdot x_1}{x_1 - x_2} = \frac{x_1}{x_1 - x_2} \\ K_2(f,x) &= \frac{\frac{\partial f}{\partial x_2}(x)\, x_2}{f(x)} = \frac{(-1) \cdot x_2}{x_1 - x_2} = -\frac{x_2}{x_1 - x_2} \end{aligned}$$

e quindi

$$\delta \doteq \frac{x_1}{x_1 - x_2} \cdot \varepsilon_1 - \frac{x_2}{x_1 - x_2} \cdot \varepsilon_2.$$

Per cui, per addizione e sottrazione $\pm$ si ha

$$\delta \doteq \frac{x_1}{x_1 \pm x_2} \cdot \varepsilon_1 \pm \frac{x_2}{x_1 \pm x_2} \cdot \varepsilon_2.$$

Moltiplicazione:

$$f(x) = x_1 x_2,\ x \in \mathbb{R}^2.$$

Si ha, per $x \in \mathbb{R}^2$ con $x_1, x_2 \neq 0$,

$$\begin{aligned} K_1(f,x) &= \frac{\frac{\partial f}{\partial x_1}(x)\, x_1}{f(x)} = \frac{x_2 \cdot x_1}{x_1 x_2} = 1 \\ K_2(f,x) &= \frac{\frac{\partial f}{\partial x_2}(x)\, x_2}{f(x)} = \frac{x_1 \cdot x_2}{x_1 x_2} = 1 \end{aligned}$$

e quindi

$$\delta \doteq \varepsilon_1 + \varepsilon_2.$$

Divisione:

$$f(x) = \frac{x_1}{x_2},\ x \in D = \left\{ x \in \mathbb{R}^2 : x_2 \neq 0 \right\}.$$

Si ha, per $x \in D$ con $x_1 \neq 0$,

$$\begin{aligned} K_1(f,x) &= \frac{\frac{\partial f}{\partial x_1}(x)\, x_1}{f(x)} = \frac{\frac{1}{x_2} \cdot x_1}{\frac{x_1}{x_2}} = 1 \\ K_2(f,x) &= \frac{\frac{\partial f}{\partial x_2}(x)\, x_2}{f(x)} = \frac{\left(-\frac{x_1}{x_2^2}\right) \cdot x_2}{\frac{x_1}{x_2}} = -1 \end{aligned}$$

e quindi

$$\delta \doteq \varepsilon_1 - \varepsilon_2.$$

Riconsideriamo la questione dell'incertezza delle misure V_m, R_m e I_m di V, R e I, rispettivamente, nel circuito elettrico precedentemente considerato. Si ha

$$\begin{aligned} V &= V_m\left(1 + \widetilde{\varepsilon}_V\right) \text{ con } |\widetilde{\varepsilon}_V| \leq 1\% \\ R &= R_m\left(1 + \widetilde{\varepsilon}_R\right) \text{ con } |\widetilde{\varepsilon}_R| \leq 2\% \end{aligned}$$

e

$$I = I_m \left(1 + \widetilde{\varepsilon}_I\right),$$

dove

$$I = \frac{V}{R} \text{ e } I_m = \frac{V_m}{R_m},$$

e ci si domanda quanto grande può essere $|\widetilde{\varepsilon}_I|$. Risulta

$$\widetilde{\varepsilon}_I \dot{=} \widetilde{\varepsilon}_V - \widetilde{\varepsilon}_R$$

e quindi

$$|\widetilde{\varepsilon}_I| \dot{=} |\widetilde{\varepsilon}_V - \widetilde{\varepsilon}_R| \leq |\widetilde{\varepsilon}_V| + |\widetilde{\varepsilon}_R| \leq 1\% + 2\% = 3\%.$$

Usando la notazione

$$a \dot{\leq} b \text{ per dire } \ a \dot{=} c \text{ e } c \leq b \text{ per qualche numero } c,$$

possiamo dire che $|\widetilde{\varepsilon}_I| \dot{\leq} 3\%$. L'incertezza di I_m è pertanto del 3%.

Opposto:

$$f(x) = -x, \ x \in \mathbb{R}.$$

Questo è il caso $a = -1$ della funzione $f(x) = ax$ vista in precedenza: si ha, per $x \in \mathbb{R}$ con $x \neq 0$,

$$K(f, x) = 1$$

e quindi

$$\delta \dot{=} \varepsilon.$$

Reciproco:

$$f(x) = \frac{1}{x}, \ x \in D = \{x \in \mathbb{R} : x \neq 0\}.$$

Si ha, per $x \in D$,

$$K(f, x) = \frac{f'(x)\, x}{f(x)} = \frac{-\frac{1}{x^2} \cdot x}{\frac{1}{x}} = -1$$

e quindi

$$\delta \dot{=} -\varepsilon.$$

Esaminiamo ora il buon e mal condizionamento delle operazioni aritmetiche. Riscrivendo gli indici di condizionamento dell'addizione come

$$\begin{aligned} K_1(f, x) &= \frac{x_1}{x_1 + x_2} = \frac{1}{1 + \frac{x_2}{x_1}} \\ K_2(f, x) &= \frac{x_2}{x_1 + x_2} = \frac{1}{1 + \frac{x_1}{x_2}} \end{aligned}$$

si vede che il mal condizionamento è presente solo nel caso di un'addizione $x_1 + x_2$ con $\frac{x_1}{x_2} \approx -1$. Analogamente, riscrivendo gli indici di condizionamento della sottrazione come

$$\begin{aligned} K_1(f,x) &= \frac{x_1}{x_1 - x_2} = \frac{1}{1 - \frac{x_2}{x_1}} \\ K_2(f,x) &= -\frac{x_2}{x_1 - x_2} = \frac{1}{1 - \frac{x_1}{x_2}} \end{aligned}$$

si vede che il mal condizionamento è presente solo nel caso di una sottrazione $x_1 - x_2$ con $\frac{x_1}{x_2} \approx 1$. Ricordare che, comunque, stiamo assumendo $\frac{x_1}{x_2} \neq -1$ per l'addizione e $\frac{x_1}{x_2} \neq 1$ per la sottrazione in modo da avere $y \neq 0$.

Le altre operazioni aritmetiche di moltiplicazione, divisione, opposto e reciproco sono ben condizionate su ogni dato.

Per capire il mal condizionamento nel caso di una sottrazione con $\frac{x_1}{x_2} \approx 1$, si consideri $y = x_1 - x_2$ con

$$\begin{aligned} x_1 &= b_0.b_1 \ldots b_{k-1} b_k \\ x_2 &= b_0.b_1 \ldots b_{k-1} c_k, \end{aligned}$$

dove $b_0, b_1, \ldots, b_{k-1}, b_k, c_k$ sono cifre decimali con $b_0 \neq 0$ e $b_k \neq c_k$. Si perturbi adesso x_1 e x_2 in

$$\begin{aligned} \widetilde{x}_1 &= b_0.b_1 \ldots b_{k-1} b_k b_{k+1} \\ \widetilde{x}_2 &= b_0.b_1 \ldots b_{k-1} c_k c_{k+1}, \end{aligned}$$

dove b_{k+1} e c_{k+1} sono cifre decimali con $b_{k+1} \neq c_{k+1}$.

Si ha

$$\begin{aligned} \varepsilon_1 &= \frac{\widetilde{x}_1 - x_1}{x_1} = \frac{b_{k+1} \cdot 10^{-k-1}}{b_0.b_1 \ldots b_{k-1} b_k} = \frac{b_{k+1}}{b_0.b_1 \ldots b_{k-1} b_k} \cdot 10^{-k-1} \\ \varepsilon_2 &= \frac{\widetilde{x}_2 - x_2}{x_2} = \frac{c_{k+1} \cdot 10^{-k-1}}{b_0.b_1 \ldots b_{k-1} c_k} = \frac{c_{k+1}}{b_0.b_1 \ldots b_{k-1} c_k} \cdot 10^{-k-1} \end{aligned}$$

ma

$$\begin{aligned} \delta &= \frac{\widetilde{y} - y}{y} \\ &= \frac{\underbrace{(b_k - c_k) \cdot 10^{-k} + (b_{k+1} - c_{k+1}) \cdot 10^{-k-1}}_{=\widetilde{y} = \widetilde{x}_1 - \widetilde{x}_2} - \underbrace{(b_k - c_k) \cdot 10^{-k}}_{=y = x_1 - x_2}}{(b_k - c_k) \cdot 10^{-k}} \\ &= \frac{(b_{k+1} - c_{k+1}) \cdot 10^{-k-1}}{(b_k - c_k) \cdot 10^{-k}} = \frac{b_{k+1} - c_{k+1}}{b_k - c_k} \cdot 10^{-1} \end{aligned}$$

e quindi δ è più grande di ε_1 e ε_2 di k ordini di grandezza.

In situazioni come queste si parla di *cancellazione numerica*, in quanto l'esplosione dell'errore relativo su $y = x_1 - x_2$ rispetto agli errori relativi su x_1 e x_2 è dovuta alla cancellazione delle cifre significative iniziali uguali $b_0, b_1 \ldots, b_{k-1}$ nel momento in cui si fa la differenza.

Tale cancellazione porta in posizione più significativa le perturbazioni b_{k+1} e c_{k+1}: si ha

$$\begin{array}{ll} x_1: & b_0.b_1 \ldots b_{k-1} b_k \quad - \\ x_2: & b_0.b_1 \ldots b_{k-1} c_k \\ y: & 0.\ 0\ \ldots \quad 0\,(b_k - c_k) \end{array}$$

e

$$\begin{array}{ll} \widetilde{x}_1: & b_0.b_1 \ldots b_{k-1} b_k b_{k+1} \quad - \\ \widetilde{x}_2: & b_0.b_1 \ldots b_{k-1} c_k c_{k+1} \\ \widetilde{y}: & 0.\ 0\ \ldots \quad 0\,(b_k - c_k)\,(b_{k+1} - c_{k+1})\,. \end{array}$$

Mentre in $\widetilde{x}_1$ e $\widetilde{x}_2$ le perturbazioni sono nella $k+2-$esima cifra significativa, in $\widetilde{y}$ finiscono nella seconda cifra significativa.

Esercizio. Spiegare perchè nel caso di una funzione lineare $f : \mathbb{R}^n \to \mathbb{R}$, vale a dire

$$f(x) = a^T x = \sum_{i=1}^{n} a_i x_i,\ x \in \mathbb{R}^n,$$

dove $a \in \mathbb{R}^n$, si può sostituire $\dot{=}$ con $=$ nella relazione tra l'errore δ e gli errori ε_i. Osservare che addizione, sottrazione e

$$f(x) = ax,\ x \in \mathbb{R},$$

dove $a \in \mathbb{R}$ hanno questa forma.

Esercizio. Provare che nel caso di una somma $x_1 + x_2$ con x_1 e x_2 dello stesso segno, risulta

$$|\delta| \leq \max\{|\varepsilon_1|, |\varepsilon_2|\}.$$

Esercizio. Determinare gli indici di condizionamento delle funzioni $f, g : \mathbb{R}^n \to \mathbb{R}$ date da

$$f(x) = x_1 + x_2 + \cdots + x_n,\ g(x) = x_1 x_2 \cdots x_n,\ x \in \mathbb{R}^n.$$

3.4 Indici di condizionamento delle funzioni matematiche elementari

Vediamo ora gli indici di condizionamento di alcune funzioni matematiche elementari.

Potenza ad esponente $\alpha \in \mathbb{R}$:

$$f(x) = x^{\alpha} = e^{\alpha \log x},\ x > 0.$$

Si ha, per $x > 0$,

$$K(f, x) = \frac{f'(x)\, x}{f(x)} = \frac{\alpha x^{\alpha - 1} \cdot x}{x^{\alpha}} = \alpha$$

e quindi

$$\delta \doteq \alpha \cdot \varepsilon.$$

Per cui, a meno che non si abbia α grande (vale a dire $|\alpha| \gg 1$), la potenza ad esponente reale α è ben condizionata su ogni dato.

Nel caso particolare $\alpha = \frac{1}{n}$, dove n è un intero positivo, si ha che la radice $n-$esima è ben condizionata su ogni dato.

Esponenziale:

$$f(x) = e^{x},\ x \in \mathbb{R}.$$

Si ha, per $x \in \mathbb{R}$ con $x \neq 0$,

$$K(f, x) = \frac{f'(x)\, x}{f(x)} = \frac{e^{x} x}{e^{x}} = x$$

e quindi

$$\delta \doteq x \cdot \varepsilon.$$

Per cui, la funzione esponenziale è mal condizionata se e solo se x è grande (vale a dire $|x| \gg 1$).

Logaritmo:

$$f(x) = \log x,\ x > 0.$$

Si ha per $x > 0$ con $x \neq 1$,

$$K(f, x) = \frac{f'(x)\, x}{f(x)} = \frac{\frac{1}{x} \cdot x}{\log x} = \frac{1}{\log x}$$

e quindi

$$\delta \doteq \frac{1}{\log x} \cdot \varepsilon.$$

Per cui, la funzione logaritmo è mal condizionata se e solo se $x \approx 1$.

Seno:

$$f(x) = \sin x,\ x \in \mathbb{R}.$$

Si ha, per x che non è un multiplo di π,

$$K(f, x) = \frac{f'(x)\, x}{f(x)} = \frac{\cos x \cdot x}{\sin x} = \frac{x}{\tan x}$$

e quindi

$$\delta \doteq \frac{x}{\tan x} \cdot \varepsilon.$$

Per x grande, a parte il caso in cui x è vicino a $\frac{\pi}{2}$ più un multiplo di π, si ha che la funzione seno è mal condizionata. Per x non grande, la funzione seno è mal condizionata se e solo se x è vicino a un multiplo non nullo di π. Per x piccolo (vale a dire $|x| \ll 1$), la funzione seno è ben condizionata in quanto

$$\lim_{x \to 0} \frac{x}{\tan x} = 1.$$

Coseno:

$$f(x) = \cos x, \ x \in \mathbb{R}.$$

Si ha, per $x \neq 0$ e x non uguale a $\frac{\pi}{2}$ più un multiplo di π,

$$K(f, x) = \frac{f'(x)\, x}{f(x)} = \frac{(-\sin x) \cdot x}{\cos x} = -x \tan x$$

e quindi

$$\delta \doteq -x \tan x \cdot \varepsilon.$$

Per x grande, a parte il caso in cui x è vicino a un multiplo di π, si ha che la funzione coseno è mal condizionata. Per x non grande, la funzione coseno è mal condizionata se e solo se x è vicino a $\frac{\pi}{2}$ più un multiplo di π.

Esercizio. Per ognuna delle seguenti funzioni di una variabile, determinare l'indice di condizionamento e i punti x in cui è mal condizionata:

1) tan;

2) arcsin, arccos, arctan;

3)
$$f(x) = \frac{ax + b}{x + c}, \ x \in \mathbb{R} \setminus \{-c\},$$
dove $a, b, c \in \mathbb{R}$;

4) sinh, cosh, tanh;

5)
$$f(x) = \sin x \cos x, \ x \in \mathbb{R}.$$

Esercizio. Per ognuna delle seguenti funzioni di due variabili, determinare gli indici di condizionamento e i punti x in cui è mal condizionata:

1) $f(x) = x_1^n - x_2^n,\ x \in \mathbb{R}^2$, dove n è un intero positivo.

2) $f(x) = g(x_1 x_2),\ x \in \mathbb{R}^2$, con $g : \mathbb{R} \to \mathbb{R}$.

3) $f(x) = x_1^{x_2},\ x \in D = \{x \in \mathbb{R}^2 : x_1 > 0\}$.

Esercizio. Determinare gli indici di condizionamento della funzione norma euclidea $\| \cdot \|_2 : \mathbb{R}^n \to \mathbb{R}$ e i punti di $\mathbb{R}^n$ in cui è mal condizionata.

Esercizio. Si consideri un polinomio

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0,\ x \in \mathbb{R},$$

di grado n. Dire se p è ben o mal condizionato su un dato x piccolo e su un dato x grande.

4 Analisi con un unico indice di condizionamento

Consideriamo ora il caso di un problema matematico caratterizzato da una funzione dato-risultato $f : D \subseteq \mathbb{R}^n \to \mathbb{R}^m$, dove D è un aperto.

Si potrebbero considerare le componenti $f_i : D \subseteq \mathbb{R}^n \to \mathbb{R}$ di f, $i \in \{1, \ldots, m\}$, date da

$$f_i(x) = f(x)_i = i\text{-esima componente di } f(x) \in \mathbb{R}^m,\ x \in D,$$

e, per ciascuna di esse, applicare le considerazioni svolte nella precedente sezione. Però, così facendo, la funzione f avrebbe mn indici di condizionamento.

Sarebbe invece auspicabile, per un problema, misurare la sensibilità del risultato del problema rispetto a perturbazioni sul dato utilizzando pochi, se non un solo indice. Questo anche nel caso di funzioni $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$, qualora n non sia un numero naturale piccolo.

Procediamo allora come segue dove, si assumerà che f sia di classe C^2.

4.1 Indice di condizionamento

Sia $x \in D$ il dato, $y = f(x)$ il risultato, $\widetilde{x} \in D$ il dato perturbato e $\widetilde{y} = f(\widetilde{x})$ il risultato perturbato. Osservare che x e $\widetilde{x}$ sono vettori di $\mathbb{R}^n$ e che y e $\widetilde{y}$ sono vettori di $\mathbb{R}^m$.

Assumiamo $x \neq \underline{0}$ e $y \neq \underline{0}$ e fissiamo una norma $\| \cdot \|_{\mathbb{R}^n}$ su $\mathbb{R}^n$ e una norma $\| \cdot \|_{\mathbb{R}^m}$ su $\mathbb{R}^m$.

Sia

$$\varepsilon = \frac{\|\widetilde{x} - x\|_{\mathbb{R}^n}}{\|x\|_{\mathbb{R}^n}}$$

l'errore relativo sul dato e

$$\delta = \frac{\|\widetilde{y} - y\|_{\mathbb{R}^m}}{\|y\|_{\mathbb{R}^m}}$$

l'errore relativo sul risultato.

Vogliamo vedere come l'errore δ dipende dall'errore ε.

Essendo f di classe C^2 si ha

$$\widetilde{y} - y = f'(x)(\widetilde{x} - x) + R(\widetilde{x}, x),$$

dove

$$f'(x) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(x) & \cdots & \frac{\partial f_1}{\partial x_n}(x) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(x) & \cdots & \frac{\partial f_m}{\partial x_n}(x) \end{bmatrix} \in \mathbb{R}^{m \times n}$$

è la matrice jacobiana di f in x e il resto $R(\widetilde{x}, x) \in \mathbb{R}^m$ dello sviluppo di Taylor è tale che

$$\|R(\widetilde{x}, x)\|_{\mathbb{R}^m} = O\left(\|\widetilde{x} - x\|^2_{\mathbb{R}^n}\right), \ \|\widetilde{x} - x\|_{\mathbb{R}^n} \to 0,$$

cioè esistono $d > 0$ e $C \geq 0$ per cui

$$\|R(\widetilde{x}, x)\|_{\mathbb{R}^m} \leq C \|\widetilde{x} - x\|^2_{\mathbb{R}^n}, \text{ per } \widetilde{x} \in D \text{ tale che } \|\widetilde{x} - x\|_{\mathbb{R}^n} \leq d.$$

Pertanto,

$$\begin{aligned} \|\widetilde{y} - y\|_{\mathbb{R}^m} &= \|f'(x)(\widetilde{x} - x) + R(\widetilde{x}, x)\|_{\mathbb{R}^m} \\ &\leq \|f'(x)(\widetilde{x} - x)\|_{\mathbb{R}^m} + \|R(\widetilde{x}, x)\|_{\mathbb{R}^m} \\ &\leq \|f'(x)\| \, \|\widetilde{x} - x\|_{\mathbb{R}^n} + \|R(\widetilde{x}, x)\|_{\mathbb{R}^m}, \end{aligned}$$

dove $\| f'(x) \|$ è la norma di operatore di $f'(x)$ relativa alle norme alle norme $\| \cdot \|_{\mathbb{R}^n}$ su $\mathbb{R}^n$ e $\| \cdot \|_{\mathbb{R}^m}$ su $\mathbb{R}^m$. Quindi

$$\begin{aligned} \delta &= \frac{\|\widetilde{y} - y\|_{\mathbb{R}^m}}{\|y\|_{\mathbb{R}^m}} \leq \frac{\|f'(x)\|}{\|y\|_{\mathbb{R}^m}} \cdot \underbrace{\|\widetilde{x} - x\|_{\mathbb{R}^n}}_{=\varepsilon \|x\|_{\mathbb{R}^n}} + \frac{\|R(\widetilde{x}, x)\|_{\mathbb{R}^m}}{\|y\|_{\mathbb{R}^m}} \\ &= \frac{\|f'(x)\| \, \|x\|_{\mathbb{R}^n}}{\|y\|_{\mathbb{R}^m}} \cdot \varepsilon + \frac{\|R(\widetilde{x}, x)\|_{\mathbb{R}^m}}{\|y\|_{\mathbb{R}^m}}. \end{aligned}$$

Mostriamo ora che il secondo termine $\frac{\|R(\widetilde{x}, x)\|_{\mathbb{R}^m}}{\|y\|_{\mathbb{R}^m}}$ è trascurabile rispetto al primo termine.

Se

$$\varepsilon \leq \frac{d}{\|x\|_{\mathbb{R}^n}}$$

e quindi

$$\|\widetilde{x} - x\|_{\mathbb{R}^n} = \varepsilon \|x\|_{\mathbb{R}^n} \leq \frac{d}{\|x\|_{\mathbb{R}^n}} \|x\|_{\mathbb{R}^n} = d,$$

si ha

$$\frac{\|R(\widetilde{x}, x)\|_{\mathbb{R}^m}}{\|y\|_{\mathbb{R}^m}} \leq \frac{C \|\widetilde{x} - x\|^2_{\mathbb{R}^n}}{\|y\|_{\mathbb{R}^m}} = \frac{C (\varepsilon \|x\|_{\mathbb{R}^n})^2}{\|y\|_{\mathbb{R}^m}} = \frac{C \|x\|^2_{\mathbb{R}^n}}{\|y\|_{\mathbb{R}^m}} \cdot \varepsilon^2.$$

Si può quindi trascurare il termine $\frac{\|R(\widetilde{x},x)\|_{\mathbb{R}^m}}{\|y\|_{\mathbb{R}^m}}$ in

$$\delta \leq \frac{\|f'(x)\| \|x\|_{\mathbb{R}^n}}{\|y\|_{\mathbb{R}^m}} \cdot \varepsilon + \frac{\|R(\widetilde{x},x)\|_{\mathbb{R}^m}}{\|y\|_{\mathbb{R}^m}}.$$

in quanto esso è maggiorato da un multiplo di ε^2, mentre l'altro termine è un multiplo dell'errore ε.

Scriviamo allora

$$\delta \dot{\leq} K(f,x) \cdot \varepsilon$$

dove

$$K(f,x) := \frac{\|f'(x)\| \|x\|_{\mathbb{R}^n}}{\|y\|_{\mathbb{R}^m}} = \frac{\|f'(x)\| \|x\|_{\mathbb{R}^n}}{\|f(x)\|_{\mathbb{R}^m}}$$

e la notazione $a \dot{\leq} b$ significa

$$a \dot{=} c \text{ e } c \leq b \text{ per qualche numero } c,$$

o equivalentemente

$$a \leq d \text{ e } d \dot{=} b \text{ per qualche numero } d.$$

Qui, $\dot{=}$ significa uguale a meno di un termine il cui valore assoluto è minore o uguale di una costante volte ε^2, per ε sufficientemente piccolo.

Infatti, sopra si ha

$$\delta \leq \frac{\|f'(x)\| \|x\|_{\mathbb{R}^n}}{\|y\|_{\mathbb{R}^m}} \cdot \varepsilon + \frac{\|R(\widetilde{x},x)\|_{\mathbb{R}^m}}{\|y\|_{\mathbb{R}^m}}. \tag{4}$$

e

$$\frac{\|f'(x)\| \|x\|_{\mathbb{R}^n}}{\|y\|_{\mathbb{R}^m}} \cdot \varepsilon + \frac{\|R(\widetilde{x},x)\|_{\mathbb{R}^m}}{\|y\|_{\mathbb{R}^m}} \dot{=} \frac{\|f'(x)\| \|x\|_{\mathbb{R}^n}}{\|y\|_{\mathbb{R}^m}} \cdot \varepsilon.$$

Il numero $K(f,x)$ è detto *indice di condizionamento* di f (del problema matematico corrispondente) relativo al dato x nelle norme $\|\cdot\|_{\mathbb{R}^n}$ su $\mathbb{R}^n$ e $\|\cdot\|_{\mathbb{R}^m}$ su $\mathbb{R}^m$.

Esercizio. Provare l'equivalenza di

$$a \dot{=} c \text{ e } c \leq b \text{ per qualche numero } c,$$

e

$$a \leq d \text{ e } d \dot{=} b \text{ per qualche numero } d.$$

Esercizio. Consideriamo una funzione $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ con $n = 1$. Nell'analisi con più indici di condizionamento, una tale funzione ha un unico indice di condizionamento $K(f,x)$. Quale relazione intercorre tra questo unico indice di condizionamento $K(f,x)$ e l'indice di condizionamento, introdotto nell'analisi con un unico indice di condizionamento, di questa stessa funzione $f : D \subseteq \mathbb{R}^n \to \mathbb{R}^m$, con $m = n = 1$, nella norma $|\cdot|$ su $\mathbb{R}^n = \mathbb{R}$ e $\mathbb{R}^m = \mathbb{R}$?

Proveremo che per un opportuno dato perturbato $\widetilde{x}$ si ha

$$\delta \doteq K(f,x)\,\varepsilon.$$

Ricordando che

$$\|f'(x)\| = \max_{u\in\mathbb{R}^n\setminus\{\underline{0}\}} \frac{\|f'(x)\,u\|_{\mathbb{R}^m}}{\|u\|_{\mathbb{R}^n}},$$

sia $z \in \mathbb{R}^n \setminus \{\underline{0}\}$ tale che

$$\|f'(x)\| = \frac{\|f'(x)\,z\|_{\mathbb{R}^m}}{\|z\|_{\mathbb{R}^n}}.$$

Per ogni $u \in \operatorname{span}(z) = \{tz : t \in \mathbb{R}\}$ si ha

$$\|f'(x)\,u\|_{\mathbb{R}^m} = \|f'(x)\|\,\|u\|_{\mathbb{R}^n}\,. \tag{5}$$

Infatti, per $u = tz$ con $t \in \mathbb{R}$, si ha

$$\begin{aligned}\|f'(x)\,u\|_{\mathbb{R}^m} &= \|f'(x)\,tz\|_{\mathbb{R}^m} = \|t\ f'(x)\,z\|_{\mathbb{R}^m}\\ &= |t|\,\|f'(x)\,z\|_{\mathbb{R}^m} = |t|\,\|f'(x)\|\,\|z\|_{\mathbb{R}^n}\\ &= \|f'(x)\|\,\|tz\|_{\mathbb{R}^n} = \|f'(x)\|\,\|u\|_{\mathbb{R}^n}\,.\end{aligned}$$

Si consideri ora un dato perturbato $\widetilde{x} \in D$ tale che $\widetilde{x} - x \in \operatorname{span}(z)$, vale a dire $\widetilde{x} - x = tz$ per qualche $t \in \mathbb{R}$.

Per provare che

$$\delta \doteq K(f,x)\,\varepsilon$$

basta mostrare che $|\delta - K(f,x)\,\varepsilon|$ è minore o uguale di una costante volte ε^2, per ε sufficientemente piccolo. Si ha

$$\begin{aligned}|\delta - K(f,x)\,\varepsilon| &= \left|\frac{\|\widetilde{y}-y\|_{\mathbb{R}^m}}{\|y\|_{\mathbb{R}^m}} - \frac{\|f'(x)\|\,\|x\|_{\mathbb{R}^n}}{\|y\|_{\mathbb{R}^m}}\cdot\frac{\|\widetilde{x}-x\|_{\mathbb{R}^n}}{\|x\|_{\mathbb{R}^n}}\right|\\ &= \left|\frac{\|\widetilde{y}-y\|_{\mathbb{R}^m} - \|f'(x)\|\,\|\widetilde{x}-x\|_{\mathbb{R}^n}}{\|y\|_{\mathbb{R}^m}}\right|\\ &= \frac{|\,\|\widetilde{y}-y\|_{\mathbb{R}^m} - \|f'(x)\|\,\|\widetilde{x}-x\|_{\mathbb{R}^n}\,|}{\|y\|_{\mathbb{R}^m}}\\ &= \frac{|\,\|\widetilde{y}-y\|_{\mathbb{R}^m} - \|f'(x)\,(\widetilde{x}-x)\|_{\mathbb{R}^m}\,|}{\|y\|_{\mathbb{R}^m}} \qquad \text{essendo } \widetilde{x}-x = \operatorname{span}(z)\\ &\le \frac{\|\widetilde{y}-y-f'(x)\,(\widetilde{x}-x)\|_{\mathbb{R}^m}}{\|y\|_{\mathbb{R}^m}} \qquad \text{segue da } |\,\|u\|-\|v\|\,| \le \|u-v\|\\ &= \frac{\|R(\widetilde{x},x)\|_{\mathbb{R}^m}}{\|y\|_{\mathbb{R}^m}}.\end{aligned}$$

Sopra si è mostrato che

$$\varepsilon \le \frac{d}{\|x\|_{\mathbb{R}^n}} \quad\Rightarrow\quad \frac{\|R(\widetilde{x},x)\|_{\mathbb{R}^m}}{\|y\|_{\mathbb{R}^m}} \le \frac{C\,\|x\|^2_{\mathbb{R}^n}}{\|y\|_{\mathbb{R}^m}}\cdot\varepsilon^2.$$

Quindi, se

$$\varepsilon \leq \frac{d}{\|x\|_{\mathbb{R}^n}}, \text{ cioè } \varepsilon \text{ è sufficientemente piccolo}$$

si ha

$$|\delta - K(f,x)\,\varepsilon| \leq \frac{C\,\|x\|_{\mathbb{R}^n}^2}{\|y\|_{\mathbb{R}^m}} \cdot \varepsilon^2, \text{ cioè } |\delta - K(f,x)\,\varepsilon| \leq \text{costante} \cdot \varepsilon^2.$$

Si conclude che

$$\delta \doteq K(f,x)\,\varepsilon.$$

Osservare che

$$\varepsilon = \frac{\|\widetilde{x} - x\|_{\mathbb{R}^n}}{\|x\|_{\mathbb{R}^n}} = \frac{\|tz\|_{\mathbb{R}^n}}{\|x\|_{\mathbb{R}^n}} = \frac{|t|\,\|z\|_{\mathbb{R}^n}}{\|x\|_{\mathbb{R}^n}} \leq \frac{d}{\|x\|_{\mathbb{R}^n}},$$

equivale a

$$|t| \leq \frac{d}{\|z\|_{\mathbb{R}^n}}.$$

4.2 Buon e mal condizionamento

Le nozioni di buon e mal condizionamento sono analoghe a quelle viste nell'analisi con più indici di condizionamento.

La funzione f (il problema matematico corrispondente) si dice *ben condizionata* sul dato x se l'indice di condizionamento di f relativo al dato x ha ordine di grandezza non superiore all'unità, vale a dire se non si ha $K(f,x) \gg 1$.

Per cui, se f è ben condizionata sul dato x, allora da

$$\delta \dot{\leq} K(f,x) \cdot \varepsilon$$

segue che l'errore relativo δ sul risultato ha ordine di grandezza non superiore all'ordine di grandezza dell'errore relativo ε sul dato, vale a dire non si ha $\delta \gg \varepsilon$.

La funzione f (il problema matematico corrispondente) si dice *mal condizionata* sul dato x se non è ben condizionata sul dato x, vale a dire se l'indice di condizionamento di f relativo al dato x ha ordine di grandezza superiore all'unità, vale a dire se si ha $K(f,x) \gg 1$.

Se f è mal condizionata sul dato x, allora $K(f,x)\varepsilon \gg \varepsilon$, ma non è detto che si abbia $\delta \gg \varepsilon$, vale a dire che δ abbia ordine di grandezza superiore all'ordine di grandezza di ε, dal momento che vale $\delta \dot{\leq} K(f,x) \cdot \varepsilon$. Quel che si ha è che δ può avere ordine di grandezza superiore all'ordine di grandezza di ε.

Infatti, come visto nella precedente sezione, si ha $\delta \doteq K(f,x)\,\varepsilon$ se il dato x è perturbato in $\widetilde{x} = x + tz \in D$, dove $z \in \mathbb{R}^n \setminus \{\underline{0}\}$ è tale che

$$\|f'(x)\| = \frac{\|f'(x)\,z\|_{\mathbb{R}^m}}{\|z\|_{\mathbb{R}^n}}$$

e $t \in \mathbb{R}$ è sufficientemente piccolo. Per cui, se f è mal condizionata sul dato x e il dato è perturbato in questo modo, allora $\delta \gg \varepsilon$, vale a dire δ ha ordine di grandezza superiore all'ordine di grandezza di ε.

Esaminiamo ora la questione se le nozioni di buon e mal condizionamento di f su un dato x possano essere ritenute indipendenti dalle particolari norme $\|\cdot\|_{\mathbb{R}^n}$ su $\mathbb{R}^n$ e $\|\cdot\|_{\mathbb{R}^m}$ su $\mathbb{R}^m$ che sono usate per definire l'indice di condizionamento.

Date norme $\|\cdot\|^{(1)}_{\mathbb{R}^n}$ su $\mathbb{R}^n$ e $\|\cdot\|^{(1)}_{\mathbb{R}^m}$ su $\mathbb{R}^m$ e altre norme $\|\cdot\|^{(2)}_{\mathbb{R}^n}$ su $\mathbb{R}^n$ e $\|\cdot\|^{(2)}_{\mathbb{R}^m}$ su $\mathbb{R}^m$, siano

- $\|\cdot\|^{(1)}$ la norme di operatore di matrici in $\mathbb{R}^{m\times n}$ relativa alle norme $\|\cdot\|^{(1)}_{\mathbb{R}^n}$ su $\mathbb{R}^n$ e $\|\cdot\|^{(1)}_{\mathbb{R}^m}$ su $\mathbb{R}^m$;
- $\|\cdot\|^{(2)}$ la norma di operatore di matrici in $\mathbb{R}^{m\times n}$ relativa alle norme $\|\cdot\|^{(2)}_{\mathbb{R}^n}$ su $\mathbb{R}^n$ e $\|\cdot\|^{(2)}_{\mathbb{R}^m}$ su $\mathbb{R}^m$.

Siano poi

$$K^{(1)}(f,x) = \frac{\|f'(x)\|^{(1)}\|x\|^{(1)}_{\mathbb{R}^n}}{\|f(x)\|^{(1)}_{\mathbb{R}^m}} \text{ e } K^{(2)}(f,x) = \frac{\|f'(x)\|^{(2)}\|x\|^{(2)}_{\mathbb{R}^n}}{\|f(x)\|^{(2)}_{\mathbb{R}^m}}$$

i corrispondenti indici di condizionamento.

Dall'equivalenza delle norme su $\mathbb{R}^n$ e $\mathbb{R}^m$ si ha che esistono $c, C, d, D > 0$ tali che

$$\begin{aligned} \forall u &\in \mathbb{R}^n : c\|u\|^{(1)}_{\mathbb{R}^n} \le \|u\|^{(2)}_{\mathbb{R}^n} \le C\|u\|^{(1)}_{\mathbb{R}^n} \\ \forall v &\in \mathbb{R}^m : d\|v\|^{(1)}_{\mathbb{R}^m} \le \|v\|^{(2)}_{\mathbb{R}^m} \le D\|v\|^{(1)}_{\mathbb{R}^m}. \end{aligned}$$

Si ha allora che

$$\forall A \in \mathbb{R}^{m\times n} : \frac{d}{C}\|A\|^{(1)} \le \|A\|^{(2)} \le \frac{D}{c}\|A\|^{(1)}.$$

Infatti, per $A \in \mathbb{R}^{m\times n}$ risulta

$$\|A\|^{(2)} = \max_{x\in\mathbb{R}^n\setminus\{\underline{0}\}} \underbrace{\frac{\|Ax\|^{(2)}_{\mathbb{R}^m}}{\|x\|^{(2)}_{\mathbb{R}^n}}}_{\le \frac{D\|Ax\|^{(1)}_{\mathbb{R}^m}}{c\|x\|^{(1)}_{\mathbb{R}^n}}} \le \max_{x\in\mathbb{R}^n\setminus\{\underline{0}\}} \frac{D\|Ax\|^{(1)}_{\mathbb{R}^m}}{c\|x\|^{(1)}_{\mathbb{R}^n}} = \frac{D}{c}\max_{x\in\mathbb{R}^n\setminus\{\underline{0}\}} \frac{\|Ax\|^{(1)}_{\mathbb{R}^m}}{\|x\|^{(1)}_{\mathbb{R}^n}} = \frac{D}{c}\|A\|^{(1)}$$

e

$$\|A\|^{(2)} = \max_{x\in\mathbb{R}^n\setminus\{\underline{0}\}} \underbrace{\frac{\|Ax\|^{(2)}_{\mathbb{R}^m}}{\|x\|^{(2)}_{\mathbb{R}^n}}}_{\ge \frac{d\|Ax\|^{(1)}_{\mathbb{R}^m}}{C\|x\|^{(1)}_{\mathbb{R}^n}}} \ge \max_{x\in\mathbb{R}^n\setminus\{\underline{0}\}} \frac{d\|Ax\|^{(1)}}{C\|x\|^{(1)}} = \frac{d}{C}\max_{x\in\mathbb{R}^n\setminus\{\underline{0}\}} \frac{\|Ax\|^{(1)}_{\mathbb{R}^m}}{\|x\|^{(1)}_{\mathbb{R}^n}} = \frac{d}{C}\|A\|^{(1)}.$$

Quindi

$$K^{(2)}(f,x) = \frac{\|f'(x)\|^{(2)} \|x\|^{(2)}_{\mathbb{R}^n}}{\|f(x)\|^{(2)}_{\mathbb{R}^m}} \leq \frac{\frac{D}{c} \|f'(x)\|^{(1)} C \|x\|^{(1)}_{\mathbb{R}^n}}{d \|f(x)\|^{(1)}_{\mathbb{R}^m}} = \frac{CD}{cd} K^{(1)}(f,x)$$

e

$$K^{(2)}(f,x) = \frac{\|f'(x)\|^{(2)} \|x\|^{(2)}_{\mathbb{R}^n}}{\|f(x)\|^{(2)}_{\mathbb{R}^m}} \geq \frac{\frac{d}{C} \|f'(x)\|^{(1)} c \|x\|^{(1)}_{\mathbb{R}^n}}{D \|f(x)\|^{(1)}_{\mathbb{R}^m}} = \frac{cd}{CD} K^{(1)}(f,x).$$

Per cui

$$\frac{1}{\frac{CD}{cd}} = \frac{cd}{CD} \leq \frac{K^{(2)}(f,x)}{K^{(1)}(f,x)} \leq \frac{CD}{cd}.$$

Si può concludere che $K^{(1)}(f,x)$ e $K^{(2)}(f,x)$ hanno lo stesso ordine di grandezza se la costante $\frac{CD}{cd}$, la quale è indipendente da f e da x, è dell'ordine di grandezza dell'unità, vale a dire non si ha $\frac{CD}{cd} \gg 1$ (osservare che $\frac{CD}{cd} \geq 1$).

Pertanto, se la costante $\frac{CD}{cd}$ è dell'ordine di grandezza dell'unità, allora

$$\begin{array}{c} f \text{ è ben condizionata sul dato } x \text{ nelle norme } {}^{(1)} \\ \Updownarrow \\ f \text{ è ben condizionata sul dato } x \text{ nelle norme } {}^{(2)}. \end{array}$$

Esercizio. Trovare la costante $\frac{CD}{cd}$ nel caso in cui:

a) le norme $^{(1)}$ siano norme ∞ e le norme $^{(2)}$ siano norme 1;

b) le norme $^{(1)}$ siano norme ∞ e le norme $^{(2)}$ siano norme 2;

c) le norme $^{(1)}$ siano norme 2 e le norme $^{(2)}$ siano norme 1.

4.3 Esempi

Calcoliamo l'indice di condizionamento dell'addizione

$$f(x) = x_1 + x_2, \ x \in \mathbb{R}^2.$$

Usiamo le norme $\|\cdot\|_\infty$ su $\mathbb{R}^2$ e $\|\cdot\|_\infty = |\cdot|$ su $\mathbb{R}$. Per $x \in \mathbb{R}^2 \setminus \{\underline{0}\}$ con $x_1 \neq -x_2$, si ha

$$\begin{aligned} f'(x) &= \left[\begin{array}{cc} \frac{\partial f}{\partial x_1}(x) & \frac{\partial f}{\partial x_2}(x) \end{array} \right] = \left[\begin{array}{cc} 1 & 1 \end{array} \right] \\ \|f'(x)\| &= \|f'(x)\|_\infty = 2 \end{aligned}$$

e quindi

$$K(f,x) = \frac{\|f'(x)\|_\infty \|x\|_\infty}{|f(x)|} = \frac{2 \max\{|x_1|, |x_2|\}}{|x_1 + x_2|}.$$

Siano ora $i, j \in \{1, 2\}$ tali che $|x_i| = \max\{|x_1|, |x_2|\}$ e $|x_j| = \min\{|x_1|, |x_2|\}$. Risulta

$$K(f, x) = \frac{2\max\{|x_1|, |x_2|\}}{|x_1 + x_2|} = \frac{2|x_i|}{|x_i + x_j|} = \frac{2}{\left|1 + \frac{x_j}{x_i}\right|}.$$

Viene confermato quanto trovato precedentemente nell'analisi a più indici di condizionamento: si ha mal condizionamento se e solo se $\frac{x_j}{x_i} \approx -1$.

Calcoliamo ora l'indice di condizionamento della moltiplicazione

$$f(x) = x_1 x_2, \ x \in \mathbb{R}^2.$$

Come per l'addizione usiamo le norme $\|\cdot\|_\infty$ su $\mathbb{R}^2$ e $\|\cdot\|_\infty = |\cdot|$ su $\mathbb{R}$.

Per $x \in \mathbb{R}^2$ con $x_1 \neq 0$ e $x_2 \neq 0$, si ha

$$f'(x) = \left[\begin{array}{cc} \frac{\partial f}{\partial x_1}(x) & \frac{\partial f}{\partial x_2}(x) \end{array} \right] = \left[\begin{array}{cc} x_2 & x_1 \end{array} \right]$$
$$\|f'(x)\| = \|f'(x)\|_\infty = |x_2| + |x_1|$$

e quindi

$$K(f, x) = \frac{\|f'(x)\|_\infty \|x\|_\infty}{|f(x)|} = \frac{(|x_2| + |x_1|)\max\{|x_1|, |x_2|\}}{|x_1 x_2|}.$$

Usando $|x_i|$ e $|x_j|$ come nell'addizione, si ha

$$K(f, x) = \frac{(|x_2| + |x_1|)\max\{|x_1|, |x_2|\}}{|x_1 x_2|} = \frac{(|x_i| + |x_j|)|x_i|}{|x_i||x_j|} = 1 + \frac{|x_i|}{|x_j|}.$$

Per cui, si ha mal condizionamento se e solo se $|x_i| \gg |x_j|$, in apparente contrasto con quanto stabilito nella precedente analisi a più indici di condizionamento, vale a dire che la moltiplicazione è ben condizionata su ogni dato.

Per spiegare questo, consideriamo

$$x = (1, 0.001) \ \text{ e } \widetilde{x} = (1.001, 0.002).$$

e quindi

$$y = x_1 x_2 = 0.001 \ \text{ e } \ \widetilde{y} = \widetilde{x}_1 \widetilde{x}_2 = 0.002002.$$

Risulta

$$\varepsilon = \frac{\|\widetilde{x} - x\|_\infty}{\|x\|_\infty} = \frac{\|(0.001, 0.001)\|_\infty}{\|(1, 0.001)\|_\infty} = \frac{0.001}{1} = 0.001$$

e

$$\delta = \frac{|\widetilde{y} - y|}{|y|} = \frac{0.001002}{0.001} = 1.002.$$

L'errore δ è di tre ordini di grandezza superiore all'errore ε, in accordo con il fatto che la moltiplicazione è malcondizionata sul dato x.

Questa esplosione dell'errore sul prodotto non si può vedere nella precedente analisi a più indici di condizionamento, dal momento che in essa si considerano i contributi all'errore δ dati dagli errori relativi sulle componenti del dato:

$$\begin{aligned}\varepsilon_1 &= \frac{\widetilde{x}_1 - x_1}{x_1} = \frac{0.001}{1} = 0.001\\ \varepsilon_2 &= \frac{\widetilde{x}_2 - x_2}{x_2} = \frac{0.001}{0.001} = 1.\end{aligned}$$

Pur essendo in tale analisi la moltiplicazione ben condizionata, si ha un errore δ non piccolo a causa dell'errore ε_2 non piccolo della seconda componente: ricordare che $\delta \doteq \varepsilon_1 + \varepsilon_2$.

Invece, qui in questa nuova analisi si considera il solo errore relativo sul dato:

$$\varepsilon = \frac{\|\widetilde{x} - x\|_\infty}{\|x\|_\infty} = \frac{\max\{|\widetilde{x}_1 - x_1|, |\widetilde{x}_2 - x_2|\}}{\max\{|x_1|, |x_2|\}},$$

dove gli errori assoluti di ciascuna componente non sono rapportati alla grandezza della componente, ma alla massima grandezza delle componenti.

L'errore ε è piccolo in quanto l'errore assoluto $\widetilde{x}_2 - x_2$, pur essendo non piccolo rispetto a x_2, è piccolo rispetto a $\max\{|x_1|, |x_2|\}$.

Una situazione di questo tipo si ha considerando la forza di attrazione gravitazionale che il Sole esercità sulla Terra. Tale forza è proporzionale al prodotto $m_T m_S$, dove m_T è la massa della Terra e m_S è la massa del Sole. Se la massa della Terra dovesse raddoppiare, anche la forza di attrazione gravitazionale raddoppierebbe, ma questo accadrebbe a fronte di una piccolissima perturbazione della massa complessiva del sistema Terra-Sole, essendo $m_S \gg m_T$.

Come ulteriore esempio, consideriamo $f : \mathbb{R}^2 \to \mathbb{R}^2$ data da

$$f(x) = \left(x_1^2 + x_2^2, x_1^2 - x_2^2\right), \ x \in \mathbb{R}^2.$$

Usiamo la norma $\|\cdot\|_{\mathbb{R}^2} = \|\cdot\|_1$ su $\mathbb{R}^2$.

Si ha, per $x \in \mathbb{R}^2 \setminus \{\underline{0}\}$,

$$\begin{aligned}f'(x) &= \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(x) & \frac{\partial f_1}{\partial x_2}(x) \\ \frac{\partial f_2}{\partial x_1}(x) & \frac{\partial f_2}{\partial x_2}(x) \end{bmatrix} = \begin{bmatrix} 2x_1 & 2x_2 \\ 2x_1 & -2x_2 \end{bmatrix}\\ \|f'(x)\| &= \|f'(x)\|_1 = \max\{ \underbrace{|2x_1| + |2x_1|}_{=2|x_1|+2|x_1|=4|x_1|}, \underbrace{|2x_2| + |-2x_2|}_{=2|x_2|+2|x_2|=4|x_2|} \}\\ &= 4\max\{|x_1|, |x_2|\}\end{aligned}$$

da cui

$$K(f,x) = \frac{\|f'(x)\|_1 \|x\|_1}{\|f(x)\|_1} = \frac{4\max\{|x_1|, |x_2|\}(|x_1| + |x_2|)}{\underbrace{|x_1^2 + x_2^2|}_{=|x_1|^2+|x_2|^2} + |x_1^2 - x_2^2|}.$$

Con $|x_i|$ e $|x_j|$ come sopra, si ha

$$\begin{aligned} K(f,x) &= \frac{4\max\{|x_1|,|x_2|\}(|x_1|+|x_2|)}{|x_1|^2+|x_2|^2+|x_1^2-x_2^2|} \\ &= \frac{4|x_i|(|x_i|+|x_j|)}{|x_i|^2+|x_j|^2+|x_i|^2-|x_j|^2} \\ &= \frac{4(|x_i|+|x_j|)}{2|x_i|} \\ &= 2\left(1+\frac{|x_j|}{|x_i|}\right) \leq 4. \end{aligned}$$

Per cui, f è ben condizionata su ogni dato.

Esercizio. Determinare l'indice di condizionamento della divisione

$$f(x) = \frac{x_1}{x_2},\ x \in D = \left\{x \in \mathbb{R}^2 : x_2 \neq 0\right\},$$

nelle norme $\|\cdot\|_\infty$ su $\mathbb{R}^2$ e $\|\cdot\|_\infty = |\cdot|$ su $\mathbb{R}$. Dire dove la funzione è mal condizionata.

Esercizio. Determinare l'indice di condizionamento della funzione $f : \mathbb{R}^2 \to \mathbb{R}^2$ data da

$$f(x) = (x_1 + x_2, x_1 x_2),\ x \in \mathbb{R}^2,$$

nella norma $\|\cdot\|_1$ su $\mathbb{R}^2$. Dire dove la funzione è mal condizionata.

5 Analisi del problema lineare

Consideriamo ora un problema matematico caratterizzato da una funzione dato-risultato $f : \mathbb{R}^n \to \mathbb{R}^m$ lineare, cioè si ha

$$f(x) = Ax \text{ per ogni } x \in \mathbb{R}^n$$

per un'unica matrice $A \in \mathbb{R}^{m\times n}$, detta la matrice associata a f.

Poichè la funzione lineare f è individuata dalla sua matrice associata A, d'ora in avanti si farà riferimento alla matrice A, piuttosto che alla funzione f.

Nonostante il caso del problema lineare possa essere inquadrato nel trattamento delle precedente sezione per problemi con funzioni dato-risultato $f : D \subseteq \mathbb{R}^n \to \mathbb{R}^m$, qui ripartiremo dall'inizio per meglio evidenziare le caratteristiche del problema lineare e solo dopo lo ricondurremo a caso speciale di quanto precedentemente visto.

5.1 Indice di condizionamento

Sia $x \in \mathbb{R}^n$ un dato, $y = Ax$ il corrispondente risultato, $\widetilde{x} \in \mathbb{R}^n$ il dato perturbato e $\widetilde{y} = A\widetilde{x}$ il risultato perturbato.

Assumiamo $x \neq \underline{0}$ e $y \neq \underline{0}$ e fissiamo una norma $\|\cdot\|_{\mathbb{R}^n}$ su $\mathbb{R}^n$ e una norma $\|\cdot\|_{\mathbb{R}^m}$ su $\mathbb{R}^m$. Sia poi $\|\cdot\|$ la norma di operatore su $\mathbb{R}^{m\times n}$ relativa alla norme $\|\cdot\|_{\mathbb{R}^n}$ su $\mathbb{R}^n$ e $\|\cdot\|_{\mathbb{R}^m}$ su $\mathbb{R}^m$.

Introduciamo l'errore relativo sul dato

$$\varepsilon = \frac{\|\widetilde{x} - x\|_{\mathbb{R}^n}}{\|x\|_{\mathbb{R}^n}}$$

e l'errore relativo sul risultato

$$\delta = \frac{\|\widetilde{y} - y\|_{\mathbb{R}^m}}{\|y\|_{\mathbb{R}^m}}.$$

Vogliamo vedere come l'errore δ dipende dall'errore ε.

Si ha

$$\widetilde{y} - y = A\widetilde{x} - Ax = A\left(\widetilde{x} - x\right)$$

e quindi

$$\delta = \frac{\|\widetilde{y} - y\|_{\mathbb{R}^m}}{\|y\|_{\mathbb{R}^m}} = \frac{\|A\left(\widetilde{x} - x\right)\|_{\mathbb{R}^m}}{\|y\|_{\mathbb{R}^m}} \leq \frac{\|A\| \overbrace{\|\widetilde{x} - x\|_{\mathbb{R}^n}}^{=\varepsilon\|x\|_{\mathbb{R}^n}}}{\|y\|_{\mathbb{R}^m}} = \frac{\|A\| \|x\|_{\mathbb{R}^n}}{\|y\|_{\mathbb{R}^m}} \cdot \varepsilon.$$

Per cui,

$$\delta \leq K\left(A, x\right)\varepsilon$$

dove

$$K\left(A, x\right) := \frac{\|A\| \|x\|_{\mathbb{R}^n}}{\|y\|_{\mathbb{R}^m}} = \frac{\|A\| \|x\|_{\mathbb{R}^n}}{\|Ax\|_{\mathbb{R}^m}}$$

è detto l'*indice di condizionamento di* A relativo al dato x nelle norme $\|\cdot\|_{\mathbb{R}^n}$ su $\mathbb{R}^n$ e $\|\cdot\|_{\mathbb{R}^m}$ su $\mathbb{R}^m$.

Notare che si ha

$$K\left(A, x\right) \geq 1,$$

essendo

$$\|Ax\|_{\mathbb{R}^m} \leq \|A\| \|x\|_{\mathbb{R}^n}.$$

Quanto ora trovato corrisponde a considerare il caso particolare della funzione lineare

$$f(x) = Ax,\ x \in \mathbb{R}^n,$$

nelle considerazioni svolte per una generale funzione dato-risultato $f : D \subseteq \mathbb{R}^n \to \mathbb{R}^m$.

Infatti, si ha $f'\left(x\right) = A$ in quanto, la componente i-esima di f , $i \in \{1, \ldots, m\}$, risulta essere

$$f_i\left(x\right) = \sum_{j=1}^{n} a_{ij}x_j$$

e quindi

$$\frac{\partial f_i}{\partial x_j}(x) = a_{ij},\ j \in \{1, \ldots, n\}.$$

Pertanto

$$K(f,x) = \frac{\|f'(x)\| \, \|x\|_{\mathbb{R}^m}}{\|Ax\|_{\mathbb{R}^m}} = \frac{\|A\| \, \|x\|_{\mathbb{R}^n}}{\|Ax\|_{\mathbb{R}^m}} = K(A,x).$$

Inoltre, nella relazione $\delta \dot{\leq} K(f,x)\varepsilon$, che vale per una funzione f generale, si può rimpiazzare $\dot{\leq}$ con $\leq$ in questo caso particolare della funzione lineare avendosi (ricordare (4))

$$R(\widetilde{x}, x) = f(\widetilde{x}) - f(x) - f'(x)(\widetilde{x} - x) = A\widetilde{x} - Ax - A(\widetilde{x} - x) = \underline{0}.$$

Per un opportuno dato perturbato $\widetilde{x}$ risulta

$$\delta = K(A,x)\,\varepsilon.$$

Infatti, per un dato perturbato $\widetilde{x} = x + tz \in \mathbb{R}^n$, $t \in \mathbb{R}$, con $z \in \mathbb{R}^n \setminus \{\underline{0}\}$ tale che

$$\|A\| = \frac{\|Az\|_{\mathbb{R}^m}}{\|z\|_{\mathbb{R}^n}},$$

si ha (ricordare (5))

$$\|A(\widetilde{x} - x)\|_{\mathbb{R}^m} = \|A\| \, \|\widetilde{x} - x\|_{\mathbb{R}^n}.$$

Quindi

$$\begin{aligned} \delta &= \frac{\|\widetilde{y} - y\|_{\mathbb{R}^m}}{\|y\|_{\mathbb{R}^m}} = \frac{\|A\widetilde{x} - Ax\|_{\mathbb{R}^m}}{\|Ax\|_{\mathbb{R}^m}} = \frac{\|A(\widetilde{x} - x)\|_{\mathbb{R}^m}}{\|Ax\|_{\mathbb{R}^m}} \\ &= \frac{\|A\| \overbrace{\|\widetilde{x} - x\|_{\mathbb{R}^n}}^{=\varepsilon\|x\|_{\mathbb{R}^n}}}{\|Ax\|_{\mathbb{R}^m}} = \frac{\|A\| \, \|x\|_{\mathbb{R}^n}}{\|Ax\|_{\mathbb{R}^m}} \cdot \varepsilon = K(A,x) \cdot \varepsilon. \end{aligned}$$

5.2 Buon e mal condizionamento

Le nozioni di buon e mal condizionamento sono del tutto analoghe a quelle precedentemente viste. La matrice A si dice *ben condizionata* sul dato x se l'indice di condizionamento di A relativo al dato x ha ordine di grandezza non superiore all'unità, vale a dire se non si ha $K(A,x) \gg 1$.

Per cui, se A è ben condizionata sul dato x, allora da

$$\delta \leq K(A,x)\varepsilon$$

segue che l'errore relativo δ sul risultato ha ordine di grandezza non superiore all'ordine di grandezza dell'errore relativo ε sul dato, vale a dire non si ha $\delta \gg \varepsilon$.

La matrice A si dice *mal condizionata* sul dato x se non è ben condizionata sul dato x, vale a dire se l'indice di condizionamento di A relativo al dato x ha ordine di grandezza superiore all'unità, vale a dire se si ha $K(f,x) \gg 1$.

Se A è mal condizionata sul dato x, allora δ può avere ordine di grandezza superiore all'ordine di grandezza di ε. Infatti, come visto sopra, si ha

$$\delta = K(A,x)\,\varepsilon$$

se il dato x è perturbato in $\widetilde{x} = x + tz$, dove $z \in \mathbb{R}^n \setminus \{\underline{0}\}$ è tale che

$$\|A\| = \frac{\|Az\|_{\mathbb{R}^m}}{\|z\|_{\mathbb{R}^n}}$$

e $t \in \mathbb{R}$. Per cui, se A è mal condizionata sul dato x e il dato è perturbato in questo modo, allora $\delta \gg \varepsilon$, vale a dire δ ha ordine di grandezza superiore all'ordine di grandezza di ε.

Esercizio. Provare che per un dato perturbato $\widetilde{x} = x + tx$, $t \in \mathbb{R}$, si ha $\delta = \varepsilon$. Quindi, anche se $K(A,x) \gg 1$, per questa particolare perturbazione di x risulta $\delta = \varepsilon$.

Esercizio. Determinare l'indice di condizionamento della matrice

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \in \mathbb{R}^{2\times 3}$$

nelle norma $\|\cdot\|_1$ su $\mathbb{R}^3$ e $\mathbb{R}^2$. Dire per quali dati la matrice è mal condizionata. Determinare anche dati perturbati $\widetilde{x}$ per i quali $\delta = K(A,x)\varepsilon$.

6 Indice di condizionamento (indipendente dal dato) di una matrice quadrata

Sia $A \in \mathbb{R}^{n\times n} \setminus \{0\}$ e sia $\|\cdot\|$ una norma su $\mathbb{R}^n$.

Definiamo l'*indice di condizionamento (indipendente dal dato)* di A nella norma $\|\cdot\|$ come

$$K_{\|\cdot\|}(A) := \sup_{\substack{x\in\mathbb{R}^n\setminus\{\underline{0}\} \\ Ax\neq\underline{0}}} K(A,x)\,,$$

dove $K(A,x)$ è l'indice di condizionamento di A relativo al dato x nella norma $\|\cdot\|$ su $\mathbb{R}^m = \mathbb{R}^n$. Si considera $A \neq 0$, altrimenti $Ax = \underline{0}$ per ogni $x \in \mathbb{R}^n$.

Notare che qualunque sia il dato $x \in \mathbb{R}^n \setminus \{\underline{0}\}$ con $Ax \neq \underline{0}$ risulta

$$\delta \leq K_{\|\cdot\|}(A) \cdot \varepsilon$$

essendo $\delta \leq K(A,x)\cdot\varepsilon$ e $K(A,x) \leq K_{\|\cdot\|}(A)$.

Osserviamo poi che si ha $K_{\|\cdot\|}(A) \geq 1$, essendo $K(A,x) \geq 1$ per qualunque dato $x \in \mathbb{R}^n \setminus \{\underline{0}\}$ con $Ax \neq \underline{0}$.

Teorema 1 *Si ha*

$$K_{\|\ \cdot\ \|}(A) = \begin{cases} \|A\| \cdot \|A^{-1}\| \text{ se } A \text{ è non singolare} \\ \\ +\infty \text{ se } A \text{ è singolare,} \end{cases}$$

dove $\|\ \cdot\ \|$ *in* $\|A\|$ *e* $\|A^{-1}\|$ *è la norma di operatore di matrici* $\mathbb{R}^{n\times n}$ *relativa alla norma* $\|\cdot\|$ *su* $\mathbb{R}^n$.

Dimostrazione. Risulta

$$K_{\|\ \cdot\ \|}(A) = \sup_{\substack{x\in\mathbb{R}^n\setminus\{\underline{0}\} \\ Ax\neq\underline{0}}} K(A,x) = \sup_{\substack{x\in\mathbb{R}^n\setminus\{\underline{0}\} \\ Ax\neq\underline{0}}} \frac{\|A\|\,\|x\|}{\|Ax\|} = \|A\| \sup_{\substack{x\in\mathbb{R}^n\setminus\{\underline{0}\} \\ Ax\neq\underline{0}}} \frac{\|x\|}{\|Ax\|}.$$

Proviamo ora che

$$\sup_{\substack{x\in\mathbb{R}^n\setminus\{\underline{0}\} \\ Ax\neq\underline{0}}} \frac{\|x\|}{\|Ax\|} = \begin{cases} \left\|A^{-1}\right\| \text{ se } A \text{ è non singolare} \\ \\ +\infty \text{ se } A \text{ è singolare.} \end{cases}$$

Sia A non singolare. Si ha

$$\begin{aligned} \sup_{\substack{x\in\mathbb{R}^n\setminus\{\underline{0}\} \\ Ax\neq\underline{0}}} \frac{\|x\|}{\|Ax\|} &= \sup_{y\in\mathbb{R}^n\setminus\{\underline{0}\}} \frac{\left\|A^{-1}y\right\|}{\|y\|} \quad \text{posto } y = Ax \\ &= \|A^{-1}\|. \end{aligned}$$

Sia A singolare. Si ha che esiste $x^* \in \mathbb{R}^n \setminus \{\underline{0}\}$ tale che $Ax^* = \underline{0}$ e quindi

$$\frac{\|x^*\|}{\|Ax^*\|} = +\infty.$$

Tuttavia, questo non permette di concludere che

$$\sup_{\substack{x\in\mathbb{R}^n\setminus\{\underline{0}\} \\ Ax\neq\underline{0}}} \frac{\|x\|}{\|Ax\|} = +\infty$$

in quanto l'estremo superiore è preso sugli $x \in \mathbb{R}^n \setminus \{\underline{0}\}$ tali che $Ax \neq \underline{0}$ e quindi bisogna argomentare in un altro modo.

Poichè $A \neq 0$, esiste $u \in \mathbb{R}^n$ tale che

$$Au \neq \underline{0}.$$

Consideriamo, per $t \in \mathbb{R}\setminus\{0\}$,

$$v = x^* + tu.$$

Si ha

$$Av = A\left(x^* + tu\right) = \underbrace{Ax^*}_{=\underline{0}} + Atu = Atu = tAu \neq \underline{0}$$

e quindi

$$\begin{aligned}
\sup_{\substack{x\in\mathbb{R}^n\setminus\{\underline{0}\}\\ Ax\neq\underline{0}}} \frac{\|x\|}{\|Ax\|} &\geq \frac{\|v\|}{\|Av\|} = \frac{\|x^* + tu\|}{\|tAu\|} \\
&\geq \frac{\left|\|x^*\| - \|tu\|\right|}{\|tAu\|} \\
&\quad \text{segue da } \left|\|u\| - \|v\|\right| = \left|\|u\| - \|-v\|\right| \leq \|u - (-v)\| = \|u + v\| \\
&= \frac{\left|\|x^*\| - |t|\,\|u\|\right|}{|t|\,\|Au\|}.
\end{aligned}$$

Poichè

$$\sup_{\substack{x\in\mathbb{R}^n\setminus\{\underline{0}\}\\ Ax\neq\underline{0}}} \frac{\|x\|}{\|Ax\|} \geq \lim_{t\to 0} \frac{\left|\|x^*\| - |t|\,\|u\|\right|}{|t|\,\|Au\|}$$

e

$$\lim_{t\to 0} \frac{\left|\|x^*\| - |t|\,\|u\|\right|}{|t|\,\|Au\|} = +\infty,$$

si ottiene

$$\sup_{\substack{x\in\mathbb{R}^n\setminus\{\underline{0}\}\\ Ax\neq\underline{0}}} \frac{\|x\|}{\|Ax\|} = +\infty.$$

■

Sia $A \in \mathbb{R}^{n\times n}$ non singolare. Si ha

$$K_{\|\,\cdot\,\|}(A) = \sup_{\substack{x\in\mathbb{R}^n\setminus\{\underline{0}\}\\ Ax\neq\underline{0}}} K(A,x) = \sup_{x\in\mathbb{R}^n\setminus\{\underline{0}\}} K(A,x) = \|A\|\,\|A^{-1}\|$$

(osservare che per $x \in \mathbb{R}^n \setminus \{\underline{0}\}$ si ha $Ax \neq \underline{0}$ essendo A non singolare).

Osserviamo che esiste $x^* \in \mathbb{R}^n \setminus \{\underline{0}\}$ tale che

$$K(A,x^*) = K_{\|\,\cdot\,\|}(A)$$

e quindi

$$K_{\|\,\cdot\,\|}(A) = \max_{x\in\mathbb{R}^n\setminus\{\underline{0}\}} K(A,x).$$

Infatti, basta considerare $y \in \mathbb{R}^n \setminus \{\underline{0}\}$ tale che

$$\|A^{-1}\| = \frac{\|A^{-1}y\|}{\|y\|}$$

e prendere $x^* = A^{-1}y$. Risulta

$$K(A, x^*) = \frac{\|A\| \, \|x^*\|}{\|Ax^*\|} = \frac{\|A\| \, \|A^{-1}y\|}{\|y\|} = \|A\| \, \|A^{-1}\| = K_{\|\cdot\|}(A).$$

Per cui, per ogni dato $x \in \mathbb{R}^n \setminus \{0\}$ perturbato in $\widetilde{x}$, si ha

$$\delta \leq K_{\|\cdot\|}(A)\varepsilon$$

e si ha

$$\delta = K_{\|\cdot\|}(A)\varepsilon$$

quando

$$x = x^* = A^{-1}y, \quad \text{dove } y \in \mathbb{R}^n \setminus \{\underline{0}\} \text{ è tale che } \left\|A^{-1}\right\| = \frac{\left\|A^{-1}y\right\|}{\|y\|},$$

e

$$\widetilde{x} = x^* + tz, \quad \text{dove } t \in \mathbb{R} \text{ e } z \in \mathbb{R}^n \setminus \{\underline{0}\} \text{ è tale che } \|A\| = \frac{\|Az\|}{\|z\|}.$$

Infatti, con un tale $\widetilde{x}$ si ha

$$\delta = K(A, x)\,\varepsilon$$

e con un tale x risulta

$$K_{\|\cdot\|}(A) = K(A, x).$$

Esercizio. Siano $A, B \in \mathbb{R}^{n\times n}$ non singolari e sia $\|\cdot\|$ una norma su $\mathbb{R}^n$. Mostrare che

$$K_{\|\cdot\|}(AB) \leq K_{\|\cdot\|}(A) \cdot K_{\|\cdot\|}(B).$$

Esercizio. Sia $A \in \mathbb{R}^{n\times n}$ non singolare e sia $\|\cdot\|$ una norma su $\mathbb{R}^n$. Quale relazione sussiste tra $K_{\|\cdot\|}(\alpha A)$, dove $\alpha \in \mathbb{R}\setminus\{0\}$, e $K_{\|\cdot\|}(A)$? Quale relazione sussiste tra $K_{\|\cdot\|}\left(A^{-1}\right)$ e $K_{\|\cdot\|}(A)$?

Esercizio. Sia $A \in \mathbb{R}^{n\times n}$ non singolare. Quale relazione sussiste tra $K_{\|\cdot\|_1}\left(A^T\right)$ e $K_{\|\cdot\|_\infty}(A)$? Quale relazione sussiste tra $K_{\|\cdot\|_2}\left(A^T\right)$ e $K_{\|\cdot\|_2}(A)$?

Esercizio. Sia $A = \operatorname{diag}(a_{11}, a_{22}, \ldots, a_{nn})$, dove $a_{11}, a_{22}, \ldots, a_{nn}$ sono tutti non nulli, una matrice diagonale non singolare. Determinare $K_{\|\cdot\|_1}(A)$, $K_{\|\cdot\|_2}(A)$ e $K_{\|\cdot\|_\infty}(A)$.

6.1 Buon e mal condizionamento (indipendente dal dato)

Sia $A \in \mathbb{R}^{n\times n}$ non singolare e sia $\|\cdot\|$ una norma su $\mathbb{R}^n$.

La matrice A si dice *ben condizionata* se l'indice di condizionamento di A ha ordine di grandezza non superiore all'unità, vale a dire se non si ha $K_{\|\cdot\|}(A) \gg 1$.

Per cui, se A è ben condizionata, allora, per ogni dato $x \in \mathbb{R}^n \setminus \{\underline{0}\}$, da

$$\delta \leq K_{\|\cdot\|}(A) \cdot \varepsilon$$

segue che l'errore relativo δ sul risultato ha ordine di grandezza non superiore all'ordine di grandezza dell'errore relativo ε sul dato, vale a dire non si ha $\delta \gg \varepsilon$, qualunque sia il dato.

La matrice A si dice *mal condizionata* se non è ben condizionata, vale a dire se l'indice di condizionamento di A ha ordine di grandezza superiore all'unità, vale a dire se si ha $K_{\|\cdot\|}(A) \gg 1$.

Se A è mal condizionata, allora esistono un dato $x \in \mathbb{R}^n \setminus \{\underline{0}\}$ e un dato perturbato $\widetilde{x}$ con cui risulta che δ ha ordine di grandezza superiore all'ordine di grandezza di ε, vale a dire si ha $\delta \gg \varepsilon$. Infatti, come precedentemente visto, si ha $\delta = K_{\|\cdot\|}(A)\varepsilon$ per

$$x = A^{-1}y, \quad \text{dove } y \in \mathbb{R}^n \setminus \{\underline{0}\} \text{ è tale che } \left\|A^{-1}\right\| = \frac{\left\|A^{-1}y\right\|}{\|y\|},$$

e

$$\widetilde{x} = x + tz, \quad \text{dove } t \in \mathbb{R} \text{ e } z \in \mathbb{R}^n \setminus \{\underline{0}\} \text{ è tale che } \|A\| = \frac{\|Az\|}{\|z\|}.$$

6.2 Indice di condizionamento in norma 2

Sia A non singolare e consideriamo $\|\cdot\| = \|\cdot\|_2$.

Teorema 2 *Si ha*

$$K_{\|\cdot\|_2}(A) = \frac{\sigma_{\max}}{\sigma_{\min}},$$

dove

$$\sigma_{\max} = \sqrt{\max\{\lambda : \lambda \text{ è autovalore di } A^TA\}} = \sqrt{\max\{\lambda : \lambda \text{ è autovalore di } AA^T\}}$$

e

$$\sigma_{\min} = \sqrt{\min\{\lambda : \lambda \text{ è autovalore di } A^TA\}} = \sqrt{\min\{\lambda : \lambda \text{ è autovalore di } AA^T\}}.$$

$\sigma_{\max}$ e $\sigma_{\min}$ sono il massimo e il minimo di quelli che sono chiamati i *valori singolari* di A, che sono gli autovalori non nulli di A^TA e che coincidono con gli autovalori non nulli di AA^T.

Dimostrazione. Si ha

$$K_{\|\cdot\|_2}(A) = \|A\|_2 \|A^{-1}\|_2$$

dove

$$\begin{aligned} \|A\|_2 &= \sqrt{\max\{\lambda : \lambda \text{ è autovalore di } A^TA\}} \\ &= \sqrt{\max\{\lambda : \lambda \text{ è autovalore di } AA^T\}} \\ &= \sigma_{\max} \end{aligned}$$

e

$$\begin{aligned}\|A^{-1}\|_2 &= \sqrt{\max\left\{\lambda : \lambda \text{ è autovalore di } \left(A^{-1}\right)^T A^{-1}\right\}} \\ &= \sqrt{\max\left\{\lambda : \lambda \text{ è autovalore di } A^{-1}\left(A^{-1}\right)^T\right\}}.\end{aligned}$$

Avendosi

$$\left(A^T\right)^{-1} = \left(A^{-1}\right)^T,$$

risulta

$$\begin{aligned}\left(A^{-1}\right)^T A^{-1} &= \left(A^T\right)^{-1} A^{-1} = \left(AA^T\right)^{-1} \\ A^{-1}\left(A^{-1}\right)^T &= A^{-1}\left(A^T\right)^{-1} = \left(A^T A\right)^{-1}\end{aligned}$$

e quindi

$$\begin{aligned}\|A^{-1}\|_2 &= \sqrt{\max\left\{\lambda : \lambda \text{ è autovalore di } \left(AA^T\right)^{-1}\right\}} \\ &= \sqrt{\max\left\{\lambda : \lambda \text{ è autovalore di } \left(A^T A\right)^{-1}\right\}}.\end{aligned}$$

Dal momento che gli autovalori dell'inversa B^{-1} di una matrice B sono i reciproci degli autovalori di B, risulta

$$\begin{aligned}\|A^{-1}\|_2 &= \sqrt{\max\left\{\lambda : \lambda \text{ è autovalore di } \left(AA^T\right)^{-1}\right\}} \\ &= \sqrt{\frac{1}{\min\left\{\lambda : \lambda \text{ è autovalore di } AA^T\right\}}} \\ &= \frac{1}{\sqrt{\min\left\{\lambda : \lambda \text{ è autovalore di } AA^T\right\}}} \\ &= \frac{1}{\sigma_{\min}}\end{aligned}$$

e

$$\begin{aligned}\|A^{-1}\|_2 &= \sqrt{\max\left\{\lambda : \lambda \text{ è autovalore di } \left(A^T A\right)^{-1}\right\}} \\ &= \sqrt{\frac{1}{\min\left\{\lambda : \lambda \text{ è autovalore di } A^T A\right\}}} \\ &= \frac{1}{\sqrt{\min\left\{\lambda : \lambda \text{ è autovalore di } A^T A\right\}}} = \frac{1}{\sigma_{\min}}.\end{aligned}$$

Per cui

$$K_{\|\cdot\|_2}(A) = \|A\|_2 \|A^{-1}\|_2 = \sigma_{\max} \cdot \frac{1}{\sigma_{\min}}.$$

■

6.2.1 Matrici ortogonali

Sia A una matrice ortogonale, cioè le colonne $v^{(1)}, \ldots, v^{(n)}$ di A costituiscono una base ortonormale di $\mathbb{R}^n$, vale a dire

$$\left\langle v^{(i)}, v^{(j)} \right\rangle = \left(v^{(i)}\right)^T v^{(j)} = \begin{cases} 1 \text{ se } i = j \\ 0 \text{ se } i \neq j \end{cases}, \ i, j \in \{1, \ldots, n\},$$

vale a dire

$$A^T A = I$$

in quanto

$$\left(A^T A\right)(i,j) = \left(v^{(i)}\right)^T v^{(j)} = \begin{cases} 1 \text{ se } i = j \\ 0 \text{ se } i \neq j \end{cases}, \ i, j \in \{1, \ldots, n\}.$$

Risulta

$$K_{\| \cdot \|_2}(A) = 1$$

in quanto

$$A^T A = I$$

ha tutti gli autovalori uguali a 1 e pertanto $\sigma_{\max} = \sigma_{\min} = 1$.

Questo è uno dei motivi per cui le matrici ortogonali sono importanti e implica che:

- le matrici ortogonali minimizzano tra tutte le matrici quadrate l'indice di condizionamento nella norma $\| \cdot \|_2$ avendosi

$$\forall B \in \mathbb{R}^{n \times n} : K_{\| \cdot \|_2}(B) \geq 1;$$

- per una matrice ortogonale A risulta, qualunque sia il dato $x \in \mathbb{R}^n \setminus \{\underline{0}\}$

$$\delta \leq \varepsilon$$

quando gli errori δ e ε sono misurati nella norma $\| \cdot \|_2$.

In realtà, per una matrice ortogonale A si ha

$$\delta = \varepsilon.$$

Infatti, se A è ortogonale si ha

$$\forall u \in \mathbb{R}^n : \|Au\|_2 = \|u\|_2$$

avendosi

$$\|Au\|_2^2 = (Au)^T Au = u^T \underbrace{A^T A}_{=I} u = u^T u = \|u\|_2^2.$$

Quindi

$$\delta = \frac{\|A\widetilde{x} - Ax\|_2}{\|Ax\|_2} = \frac{\|A(\widetilde{x} - x)\|_2}{\|Ax\|_2} = \frac{\|\widetilde{x} - x\|_2}{\|x\|_2} = \varepsilon.$$

6.2.2 Matrici simmetriche

Sia ora A una matrice simmetrica. Si ha

$$K_{\|\cdot\|_2}(A) = \frac{\rho_{\max}}{\rho_{\min}},$$

dove

$$\rho_{\max} = \max\{|\mu| : \mu \text{ è autovalore di } A\}$$

e

$$\rho_{\min} = \min\{|\mu| : \mu \text{ è autovalore di } A\}.$$

Infatti, dal momento che gli autovalori di $A^TA = AA^T = A^2$ sono i quadrati degli autovalori di A, risulta

$$\begin{aligned}
\sigma_{\max} &= \sqrt{\max\{\lambda : \lambda \text{ è autovalore di } A^TA\}} \\
&= \sqrt{\max\{\lambda : \lambda \text{ è autovalore di } A^2\}} \\
&= \max\{|\mu| : \mu \text{ è autovalore di } A\} \\
&= \rho_{\max}
\end{aligned}$$

e

$$\begin{aligned}
\sigma_{\min} &= \sqrt{\min\{\lambda : \lambda \text{ è autovalore di } A^TA\}} \\
&= \sqrt{\min\{\lambda : \lambda \text{ è autovalore di } A^2\}} \\
&= \min\{|\mu| : \mu \text{ è autovalore di } A\} \\
&= \rho_{\min}.
\end{aligned}$$

6.3 Un esempio

Consideriamo la matrice

$$A = \begin{bmatrix} a & 1 \\ 1 & b \end{bmatrix} \in \mathbb{R}^{2\times 2}$$

con $|a| \geq |b|$.

Assumiamo A non singolare, cioè $\det(A) = ab - 1 \neq 0$.

Determiniamo $K_{\|\cdot\|_\infty}(A)$. Si ha

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} b & -1 \\ -1 & a \end{bmatrix}.$$

Quindi, essendo $|a| \geq |b|$, risulta

$$\begin{aligned}
\|A\|_\infty &= \max\{|a| + 1, |b| + 1\} = |a| + 1 \\
\|A^{-1}\|_\infty &= \frac{1}{|ab-1|} \max\{|b| + 1, |a| + 1\} = \frac{|a| + 1}{|ab-1|}.
\end{aligned}$$

Si conclude che

$$K_{\|\cdot\|_\infty}(A) = \|A\|_\infty \|A^{-1}\|_\infty = \frac{(|a|+1)^2}{|ab-1|}.$$

Per cui:

1) Se $p = ab$ è vicino a 1, allora A è mal condizionata:

$$K_{\|\cdot\|_\infty}(A) = \frac{(|a|+1)^2}{|p-1|} \geq \frac{1}{|p-1|}.$$

2) Se p non è vicino a 1 e $|p|$ non è grande, allora A è mal condizionata se e solo se $|a|^2$ è grande:

$$\frac{4\max\{|a|^2,1\}}{|p-1|} \geq K_{\|\cdot\|_\infty}(A) = \frac{(|a|+1)^2}{|p-1|} \geq \frac{|a|^2}{|p|+1}.$$

3) Se $|p|$ è grande, allora A è mal condizionata se e solo se $\frac{|a|^2}{|p|} = \frac{|a|}{|b|}$ è grande:

$$\frac{4\max\left\{\frac{|a|^2}{|p|}, \frac{1}{|p|}\right\}}{|1-\frac{1}{p}|} \geq K_{\|\cdot\|_\infty}(A) = \frac{(|a|+1)^2}{|p-1|} \geq \frac{|a|^2}{|p|\cdot|1-\frac{1}{p}|}$$

Esercizio. Trovare, nelle tre situazioni 1), 2) e 3) sopra, dei valori a e b per cui $K_{\|\cdot\|_\infty}(A)$ ha ordine di grandezza 10^4.

Esercizio. Si consideri $a = 1.01$ e $b = 0.99$. Determinare un dato $x \in \mathbb{R}^2$ e una dato perturbato $\widetilde{x}$ tale che $\delta = K_{\|\cdot\|_\infty}(A) \cdot \varepsilon$.

Esercizio. Determinare $K_{\|\cdot\|_1}(A)$ e $K_{\|\cdot\|_2}(A)$.

Esercizio. Data la matrice

$$B = \begin{bmatrix} 1 & -1 \\ 1 & c \end{bmatrix} \in \mathbb{R}^{2\times 2}$$

con $c \neq -1$, determinare $K_{\|\cdot\|_1}(B)$ e dire per quali c si ha B mal condizionata.

Esercizio. Data la matrice

$$C = \begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix} \in \mathbb{R}^{2\times 2},$$

determinare $K_{\|\cdot\|_2}(C)$ e dire per quali a si ha C mal condizionata.