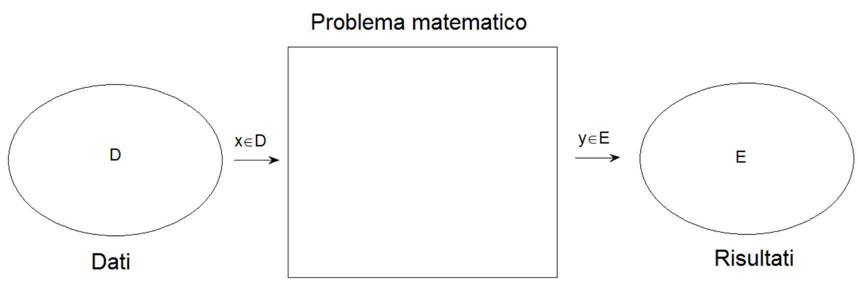
Condizionamento

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

June 15, 2018

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.

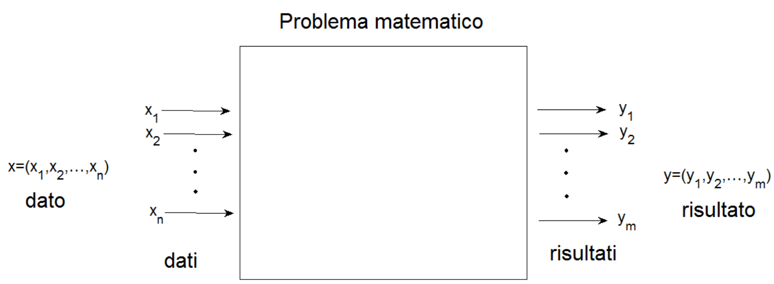


Pertanto, un problema matematico è caratterizzato dalla (e può essere identificato con la) *funzione dato-risultato* $f : D \to E$ che associa ad ogni dato $x \in D$ il corrispondente risultato $y \in E$.

Si assumerà ora 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 $\widetilde{x}$ 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 $\widetilde{y}$ a partire dall'ordine di grandezza dell'errore che ha il dato disponibile $\widetilde{x}$.

2 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 dipende dagli errori ε_i, $i \in \{1, \ldots, n\}$.

2.1 Indici di condizionamento

Essendo f di classe C^2, si ha

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

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

$$R(\widetilde{x}, x) = O\left(\|\widetilde{x} - x\|^2\right),\ \|\widetilde{x} - x\| \to 0, \tag{1}$$

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

In Analisi II si è visto che (1) vale quando si usa $\|\cdot\| = \|\cdot\|_2$. Dall'equivalenza delle norme su $\mathbb{R}^n$, si ha che (1) vale per una norma arbitraria $\|\cdot\|$ su $\mathbb{R}^n$. Usando la norma $\|\cdot\|_\infty$, (1) vuol dire: 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

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

Per cui, se

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

e quindi

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

si ha

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

Essendo interessati solo all'ordine di grandezza di δ ed 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 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

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

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 $\doteq$ significa uguale a meno di un termine il cui valore assoluto è minore o uguale di un multiplo di $\|\varepsilon\|_\infty^2$, per $\|\varepsilon\|_\infty = \max_{i\in\{1,\ldots,n\}} |\varepsilon_i|$ sufficientemente piccolo.

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.

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

Nel caso $n = 1$, si ha

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

dove

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

e

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

Come primo esempio di indice di condizionamento, 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 $x \neq 0$,

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

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

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. Quale relazione intercorre 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 $a \neq 0$.

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

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

2.2 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.

Il dire che un numero a ha ordine di grandezza non superiore all'unità è un modo per dire che non è un numero grande, cioè che non si ha $|a| >> 1$. 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 numeri in mezzo possono essere interpretati sia come non grandi che come grandi.

Se f è ben condizionata sul dato x, allora da

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

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

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.

Se f è mal condizionata sul dato x, allora δ ha "in generale" ordine di grandezza superiore al massimo ordine di grandezza degli ε_i, $i \in \{1, \dots, n\}$.

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

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

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

e quindi $|\delta| >> |\varepsilon_j|$. Quindi, in tal caso, f è mal condizionata sul dato x e δ ha ordine di grandezza superiore al massimo ordine di grandezza degli ε_i, $i \in \{1, \dots, n\}$.

D'altra parte, se esistono $j, k \in \{1, \dots, n\}$ con $j \neq k$ tali che $K_j(f, x) = K_k(f, x)$ e $|K_j(f, x)| = |K_k(f, x)| >> 1$, allora, per errori ε_i, $i \in \{1, \dots, n\}$, tali che $\varepsilon_i = 0$ per $i \neq j, k$ e $\varepsilon_j = -\varepsilon_k \neq 0$, 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 δ non ha ordine di grandezza superiore al massimo ordine di grandezza degli ε_i, $i \in \{1, \dots, n\}$, anzi ha ordine di grandezza inferiore. Esercizio. Qual è l'ordine di grandezza di δ?

Il motivo per cui δ può non avere ordine di grandezza superiore al massimo ordine di grandezza degli ε_i, $i \in \{1, \dots, n\}$, è che termini $K_i\,(f, x) \cdot \varepsilon_i$ con $|K_i\,(f, x)\,| >> 1$ 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 mal condizionata sul dato $x \Rightarrow$
δ ha ordine di grandezza superiore al massimo ordine di grandezza degli ε_i?

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

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

e

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

Esercizio. Sempre nel caso $n = 1$, cosa si può dire del buon o mal condizionamento di $g \circ f$ a partire dal buon o mal condizionamento di f e g? E del buon o mal condizionamento di fg e $f + g$ a partire dal buon o mal condizionamento di f e g?

2.3 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 $x_1 \neq -x_2$,

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

e quindi

$$\delta \doteq \frac{x_1}{x_1 + x_2} \cdot \varepsilon_1 + \frac{x_2}{x_1 + x_2} \cdot \varepsilon_2 = \frac{1}{1 + \frac{x_2}{x_1}} \cdot \varepsilon_1 + \frac{1}{1 + \frac{x_1}{x_2}} \cdot \varepsilon_2.$$

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 $x_1 \neq x_2$,

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

e quindi

$$\delta \doteq \frac{x_1}{x_1 - x_2} \cdot \varepsilon_1 - \frac{x_2}{x_1 - x_2} \cdot \varepsilon_2 = \frac{1}{1 - \frac{x_2}{x_1}} \cdot \varepsilon_1 + \frac{1}{1 - \frac{x_1}{x_2}} \cdot \varepsilon_2.$$

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

$$\delta \doteq \frac{x_1}{x_1 \pm x_2} \cdot \varepsilon_1 \pm \frac{x_2}{x_1 + x_2} \cdot \varepsilon_2 = \frac{1}{1 \pm \frac{x_2}{x_1}} \cdot \varepsilon_1 + \frac{1}{1 \pm \frac{x_1}{x_2}} \cdot \varepsilon_2.$$

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

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

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

$$\delta \doteq \varepsilon.$$

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

$$\delta \doteq -\varepsilon.$$

Per cui, con le operazione aritmetiche il mal condizionamento è presente solo nel caso di un'addizione $x_1 + x_2$ con $x_1 \simeq -x_2$, o nel caso di una sottrazione $x_1 - x_2$ con $x_1 \simeq x_2$. Ricordare che, comunque, stiamo assumendo $x_1 \neq -x_2$ per l'addizione e $x_1 \neq x_2$ per la sottrazione in modo da avere $y \neq 0$.

Per capire cosa accade, si consideri la sottrazione $y = x_1 - x_2$ con

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

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

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

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

Si ha

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

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 all'errore relativo su x_1 e x_2 è dovuta alla cancellazione delle cifre significative iniziali uguali $b_0, b_1 \dots, b_{k-1}$ 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

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

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 addizione, sottrazione e $f(x)=ax$, si può sostituire $\doteq$ con $=$ nella relazione tra l'errore δ e gli errori ε_i. Suggerimento: considerare come può essere espresso il resto $R\left(\widetilde{x},x\right)$ in termini di derivate di f oppure osservare che

$$R(\widetilde{x},x)=f(\widetilde{x})-f(x)-\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(x)(\widetilde{x}-x).$$

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

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

Suggerimento: si ha

$$\begin{aligned} |\delta| &= \left|\frac{x_1}{x_1+x_2}\varepsilon_1+\frac{x_2}{x_1+x_2}\varepsilon_2\right|\leq\left|\frac{x_1}{x_1+x_2}\varepsilon_1\right|+\left|\frac{x_2}{x_1+x_2}\varepsilon_2\right| \\ &= \frac{|x_1|}{|x_1+x_2|}\left|\varepsilon_1\right|+\frac{|x_2|}{|x_1+x_2|}\left|\varepsilon_2\right| \\ &\leq \frac{|x_1|}{|x_1+x_2|}\max\left\{|\varepsilon_1|,|\varepsilon_2|\right\}+\frac{|x_2|}{|x_1+x_2|}\max\left\{|\varepsilon_1|,|\varepsilon_2|\right\}. \end{aligned}$$

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

$$f\left(x\right)=x_1+x_2+\cdots+x_n,\;\; g\left(x\right)=x_1x_2\cdots x_n,\;\; x\in\mathbb{R}^n.$$

2.4 Indici di condizionamento delle funzioni matematiche elementari

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

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

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

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 la potenza ad esponente reale α è ben condizionata su ogni dato (a meno che $|\alpha| >> 1$). 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 condizionanta se e solo se $|x| >> 1$.

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 $x \simeq 1$.

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 cui, 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 $x \simeq 0$ la funzione seno è ben condizionata in quanto

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

<u>Coseno</u>:

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

Si ha, per $x \neq 0$ e x che non è $\frac{\pi}{2}$ 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 cui, 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) $f(x) = (x_1 - x_2)\mathrm{e}^{x_1 - x_2}, \; x \in \mathbb{R}^2.$

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

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 x piccolo e su x grande. Suggerimento. Considerare i limiti di $K(p,x)$ per $x \to 0$ e per $x \to \infty$.

3 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, $i \in \{1, \ldots, m\}$, e, per ciascuna di esse, applicare le considerazioni svolte nella precedente sezione. Così facendo, però, la funzione f avrebbe mn indici di condizionamento.

Sarebbe invece auspicabile 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.

3.1 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 $\|\cdot\|$ è la norma di operatore su $\mathbb{R}^{m\times n}$ relativa 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 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 $\dot{\leq}$ significa minore o uguale a meno di un termine non negativo minore o uguale di un multiplo di ε^2, per ε sufficientemente piccolo.

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. Quale relazione intercorre tra l'unico indice di condizionamento di una funzione $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$, dove $n = 1$, visto nella precedente sezione, e l'indice di condizionamento della stessa funzione vista come $f : D \subseteq \mathbb{R}^n \to \mathbb{R}^m$, dove $m = n = 1$, nella norma $|\cdot|$ su $\mathbb{R}^n = \mathbb{R}$ e $\mathbb{R}^m = \mathbb{R}$, visto in questa sezione?

3.2 Buon e mal condizionamento

Le nozioni di buon e mal condizionamento sono analoghe a quelle viste nel caso di 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.

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.

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.

Se f è mal condizionata sul dato x, allora δ ha "in generale" ordine di grandezza superiore all'ordine di grandezza di ε.

Sull'uso di "in generale" fatto sopra, osserviamo quanto segue.

Se $K(f,x) >> 1$, allora non si può concludere che δ ha ordine di grandezza superiore all'ordine di grandezza di ε, dal momento che si ha $K(f,x) \cdot \varepsilon$ con ordine di grandezza superiore all'ordine di grandezza di ε , ma $\delta \dot{\leq} K(f,x) \cdot \varepsilon$ e quindi δ potrebbe avere ordine di grandezza inferiore a quello di $K(f,x) \cdot \varepsilon$.

Ora però mostriamo che per un opportuno dato perturbato $\widetilde{x}$ si ha davvero che δ ha ordine di grandezza superiore all'ordine di grandezza di ε.

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

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

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 $\widetilde{x} - x \in \operatorname{span}(z)$.

Proveremo che

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

dove $\doteq$ significa uguale a meno di un termine il cui valore assoluto è minore o uguale di un multiplo di ε^2, per ε sufficientemente piccolo.

Si ha

$$\delta = K(f,x)\,\varepsilon + \delta - K(f,x)\,\varepsilon,$$

dove la differenza $\delta - K(f,x)\,\varepsilon$ è tale che

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

Sopra si è mostrato che

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

Quindi, se

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

vale a dire

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

si ha

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

da cui segue

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

Si conclude che se $K(f,x) >> 1$, allora δ ha ordine di grandezza superiore all'ordine di grandezza di ε per un dato perturbato $\widetilde{x} = x + tz \in D$ con $|t| \leq \frac{d}{\|z\|_{\mathbb{R}^n}}$.

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

Date norme $\|\cdot\|_{\mathbb{R}^n}^{(1)}$ e $\|\cdot\|_{\mathbb{R}^n}^{(2)}$ su $\mathbb{R}^n$ e $\|\cdot\|_{\mathbb{R}^m}^{(1)}$ e $\|\cdot\|_{\mathbb{R}^m}^{(2)}$ su $\mathbb{R}^m$, siano $\|\cdot\|^{(1)}$ e $\|\cdot\|^{(2)}$ le norme di operatore su $\mathbb{R}^{m\times n}$ relative alle norme $\|\cdot\|_{\mathbb{R}^n}^{(1)}$ su $\mathbb{R}^n$ e $\|\cdot\|_{\mathbb{R}^m}^{(1)}$ su $\mathbb{R}^m$ e alle norme $\|\cdot\|_{\mathbb{R}^n}^{(2)}$ su $\mathbb{R}^n$ e $\|\cdot\|_{\mathbb{R}^m}^{(2)}$ su $\mathbb{R}^m$, rispettivamente.

Siano poi

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

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

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

Per quanto visto in un precedente esercizio di “Analisi Numerica: norme” si ha allora anche

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

Quindi

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

e

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

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.

Pertanto si ha che

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

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

Esercizio. Trovare la costante $\frac{CD}{cd}$ nel caso in cui le norme $^{(1)}$ siano norme ∞ e le norme $^{(2)}$ siano norme 1.

3.3 Esempi

Riprendiamo il caso 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 $x\in\mathbb{R}^2\setminus\{0\}$ con $x_1\neq x_2$, si ha

$$f'(x)=\begin{bmatrix}\frac{\partial f}{\partial x_1}(x) & \frac{\partial f}{\partial x_2}(x)\end{bmatrix}=\begin{bmatrix}1 & 1\end{bmatrix}$$
$$\|f'(x)\|=\|f'(x)\|_\infty=2$$

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: si ha mal condizionamento se e solo se $x_i\simeq -x_j$.

Riprendiamo ora il caso della moltiplicazione

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

Come nell'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

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

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 $|x_i| >> |x_j|$, 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

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

Questa esplosione dell'errore sul prodotto non si può vedere nella precedente analisi, dal momento che in essa si considerano i contributi all'errore δ dati dagli errori relativi

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

delle componenti del dato.

Pur essendo in tale analisi la moltiplicazione ben condizionata, si ha un grande errore δ a causa del grande errore ε_2 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 grande 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 $m_S >> m_T$.

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 $x \in \mathbb{R}^2 \setminus \{0\}$,

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

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

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

Per cui, f è ben condizionata su ogni dato.

Esercizio. Determinare l'indice di condizionamento della funzione

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

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

$$f(x)=(x_1+x_2,x_1x_2),\ x\in\mathbb{R}^2,$$

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

4 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.

4.1 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

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

Quanto ora trovato corrisponde a considerare il caso particolare di una funzione lineare $f(x) = Ax$ nelle considerazioni svolte per una generale funzione dato-risultato $f : D \subseteq \mathbb{R}^n \to \mathbb{R}^m$.

Infatti, si ha $f'\left(x\right) = A$ e quindi

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

e, nella relazione $\delta \stackrel{\cdot}{\leq} K(f,x)\varepsilon$ che vale per una funzione f generale, si può rimpiazzare $\stackrel{\cdot}{\leq}$ con $\leq$ in questo caso particolare avendosi

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

4.2 Buon e mal condizionamento

Le nozioni di buon e mal condizionamento sono del tutto analoghe a quelle precedentemente viste. La matrice A (il problema matematico corrispondente) 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.

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.

La matrice A (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 A relativo al dato x ha ordine di grandezza superiore all'unità.

Se A è mal condizionata sul dato x, allora δ ha "in generale" ordine di grandezza superiore all'ordine di grandezza di ε.

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

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

risulta

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

Questo segue dall'aversi

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

(Esercizio. Provare questo) e quindi

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

Si conclude che se $K(A,x)>>1$, allora, per un dato perturbato $\widetilde{x}=x+tz$, $t\in\mathbb{R}$, si ha $\delta>>\varepsilon$.

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

Esercizio. Nell'esercizio precedente determinare dati perturbati $\widetilde{x}$ per i quali $\delta = K(A,x)\varepsilon$. Suggerimento: in "Analisi Numerica: norme", nel teorema che fornisce la formula per la norma $\|A\|_1$ di una matrice A veniva anche dato un vettore z tale che $\|Az\|_1 = \|A\|_1\|z\|_1$.

5 Indice di condizionamento (assoluto) di una matrice quadrata

Sia $A \in \mathbb{R}^{n\times n} \setminus \{0\}$ e sia $\|\cdot\|$ una norma su $\mathbb{R}^n$. Si considera $A \neq 0$, altrimenti $Ax = \underline{0}$ per ogni $x \in \mathbb{R}^n$.

Definiamo l'*indice di condizionamento (assoluto)* 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 nelle norme $\|\cdot\|$ su $\mathbb{R}^n$ e $\|\cdot\|$ su $\mathbb{R}^m = \mathbb{R}^n$.

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 su $\mathbb{R}^{n\times n}$ 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.

Cominciamo con il provare che

$$\sup_{\substack{x\in\mathbb{R}^n\setminus\{\underline{0}\}\\ Ax\neq\underline{0}}} \frac{\|x\|_2}{\|Ax\|_2} = +\infty.$$

Data una SVD $A = V\Sigma U^T$ di A, risulta

$$\begin{aligned} \sup_{\substack{x\in\mathbb{R}^n\setminus\{\underline{0}\}\\ Ax\neq\underline{0}}} \frac{\|x\|_2}{\|Ax\|_2} &= \sup_{\substack{x\in\mathbb{R}^n\setminus\{\underline{0}\}\\ Ax\neq\underline{0}}} \frac{\|x\|_2}{\|V\Sigma U^T x\|_2} = \sup_{\substack{x\in\mathbb{R}^n\setminus\{\underline{0}\}\\ Ax\neq\underline{0}}} \frac{\|x\|_2}{\|\Sigma U^T x\|_2} \quad \text{essendo } V \text{ ortogonale} \\ &= \sup_{\substack{z\in\mathbb{R}^n\setminus\{\underline{0}\}\\ AUz=V\Sigma z\neq\underline{0}}} \frac{\|Uz\|_2}{\|\Sigma z\|_2} \quad \text{posto } z = U^T x \text{ ed essendo } U^T \text{ non singolare} \\ &= \sup_{\substack{z\in\mathbb{R}^n\setminus\{\underline{0}\}\\ \Sigma z\neq\underline{0}}} \frac{\|Uz\|_2}{\|\Sigma z\|_2} \quad \text{essendo } V \text{ non singolare} \\ &= \sup_{\substack{z\in\mathbb{R}^n\setminus\{\underline{0}\}\\ \Sigma z\neq\underline{0}}} \frac{\|z\|_2}{\|\Sigma z\|_2} \quad \text{essendo } U \text{ ortogonale.} \end{aligned}$$

Sia $\varepsilon > 0$ e sia $z^* \in \mathbb{R}^n \setminus \{\underline{0}\}$ tale che

$$z^* = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \varepsilon \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \begin{array}{l} \\ \\ \\ \leftarrow r-\text{esima componente},\ r = \operatorname{rank}(A) \\ \\ \\ \\ \leftarrow n-\text{esima componente} \end{array}$$

Essendo A singolare, si ha $r < n$ e quindi ε non è sull'ultima componente, la quale è uguale a 1.

Si ha

$$\Sigma z^* = \operatorname{diag}(\sigma_1, \ldots, \sigma_r, 0, \ldots, 0) \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \varepsilon \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \sigma_r\varepsilon \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}$$

e quindi

$$\frac{\|z^*\|_2}{\|\Sigma z^*\|_2} = \frac{\sqrt{1+\varepsilon^2}}{\sigma_r\varepsilon}.$$

Avendosi

$$\sup_{\substack{x \in \mathbb{R}^n \setminus \{\underline{0}\} \\ Ax \neq \underline{0}}} \frac{\|x\|_2}{\|Ax\|_2} = \sup_{\substack{z \in \mathbb{R}^n \setminus \{\underline{0}\} \\ \Sigma z \neq \underline{0}}} \frac{\|z\|_2}{\|\Sigma z\|_2} \geq \frac{\|z^*\|_2}{\|\Sigma z^*\|_2} = \frac{\sqrt{1+\varepsilon^2}}{\sigma_r\varepsilon} \quad \text{per ogni } \varepsilon > 0$$

ed essendo $\lim_{\varepsilon \to 0} \frac{\sqrt{1+\varepsilon^2}}{\sigma_r\varepsilon} = +\infty$, si ottiene $\sup_{\substack{x \in \mathbb{R}^n \setminus \{\underline{0}\} \\ Ax \neq \underline{0}}} \frac{\|x\|_2}{\|Ax\|_2} = +\infty$.

Proviamo ora che anche per la norma generica $\|\cdot\|$ si ha

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

Essendo $\|\cdot\|_2$ e $\|\cdot\|$ norme su $\mathbb{R}^n$ equivalenti, esistono costanti $c, C > 0$ tali che

$$\forall u \in \mathbb{R}^n : c\,\|u\|_2 \leq \|u\| \leq C\,\|u\|_2$$

da cui

$$\frac{\|x\|}{\|Ax\|} \geq \frac{c\,\|x\|_2}{C\,\|Ax\|_2} \quad \text{per ogni } x \in \mathbb{R}^n \setminus \{\underline{0}\} \text{ tale che } Ax \neq \underline{0}$$

e quindi

$$\sup_{\substack{x \in \mathbb{R}^n \setminus \{\underline{0}\} \\ Ax \neq \underline{0}}} \frac{\|x\|}{\|Ax\|} \geq \sup_{\substack{x \in \mathbb{R}^n \setminus \{\underline{0}\} \\ Ax \neq \underline{0}}} \frac{c\,\|x\|_2}{C\,\|Ax\|_2} = \frac{c}{C} \sup_{\substack{x \in \mathbb{R}^n \setminus \{\underline{0}\} \\ Ax \neq \underline{0}}} \frac{\|x\|_2}{\|Ax\|_2} = +\infty.$$

■

Esercizio. Sia $A \in \mathbb{R}^{m \times n} \setminus \{0\}$ e sia $K(A, x)$ l'indice di condizionamento di A relativo al dato $x \in \mathbb{R}^n \setminus \{\underline{0}\}$ tale che $Ax \neq \underline{0}$ nelle norme $\|\cdot\|_{\mathbb{R}^n}$ su $\mathbb{R}^n$ e $\|\cdot\|_{\mathbb{R}^m}$ su $\mathbb{R}^m$ Guardando alla dimostrazione del precedente teorema, prima mostrare che

$$\sup_{\substack{x \in \mathbb{R}^n \setminus \{\underline{0}\} \\ Ax \neq \underline{0}}} \frac{\|x\|_2}{\|Ax\|_2} \begin{cases} < +\infty \text{ se } \operatorname{rank}(A) = n \\ = +\infty \text{ se } \operatorname{rank}(A) < n \end{cases},$$

poi concludere che

$$\sup_{\substack{x \in \mathbb{R}^n \setminus \{\underline{0}\} \\ Ax \neq \underline{0}}} \frac{\|x\|_{\mathbb{R}^n}}{\|Ax\|_{\mathbb{R}^m}} \begin{cases} < +\infty \text{ se } \operatorname{rank}(A) = n \\ = +\infty \text{ se } \operatorname{rank}(A) < n \end{cases}$$

e infine che

$$\sup_{\substack{x \in \mathbb{R}^n \setminus \{\underline{0}\} \\ Ax \neq \underline{0}}} K(A, x) \begin{cases} < +\infty \text{ se } \operatorname{rank}(A) = n \\ = +\infty \text{ se } \operatorname{rank}(A) < n \end{cases}$$

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

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

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

Come conseguenza della formula per l'indice di condizionamento di A data nel precedente teorema si ottiene

$$K_{\| \cdot \|_2}(A) = \frac{\sigma_1}{\sigma_n},$$

dove $\sigma_1, \sigma_2, \ldots, \sigma_n$ sono i valori singolari di A (poichè A è non singolare, A ha $r = \operatorname{rank}(A) = n$ valori singolari non nulli).

Infatti,

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

con

$$\|A\|_2 = \text{massimo valore singolare di } A = \sigma_1$$

e

$$\|A^{-1}\|_2 = \text{massimo valore singolare di } A^{-1} = \frac{1}{\sigma_n}.$$

Per cui, se A è una matrice ortogonale si ha

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

dal momento che A ha i valori singolari $\sigma_1 = \cdots = \sigma_n = 1$, essendo questi le radici quadrate degli autovalori di $A^T A = I_n$.

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 $x \in \mathbb{R}^n \setminus \{\underline{0}\}$ con $Ax \neq \underline{0}$,

$$\delta \leq \varepsilon$$

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

Osserviamo poi che se A è simmetrica si ha

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

dove

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

e

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

Infatti, per tali matrici, i valori singolari sono i moduli degli autovalori.

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

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

Suggerimento. Usare il fatto che AB è non singolare e quindi $K_{\|\,\cdot\,\|}(AB) = \|AB\| \cdot \|(AB)^{-1}\|$.

Esercizio. Siano date norme $\|\cdot\|^{(1)}$ e $\|\cdot\|^{(2)}$ su $\mathbb{R}^n$. Essendo tali norme equivalenti, esistono $c, C > 0$ tali che

$$\forall u \in \mathbb{R}^n : c\,\|u\|^{(1)} \leq \|u\|^{(2)} \leq C\,\|u\|^{(1)}.$$

Mostrare che, per ogni $A \in \mathbb{R}^{n\times n} \setminus \{0\}$, si ha

$$\left(\frac{c}{C}\right)^2 K_{\|\,\cdot\,\|^{(1)}}(A) \leq K_{\|\,\cdot\,\|^{(2)}}(A) \leq \left(\frac{C}{c}\right)^2 K_{\|\,\cdot\,\|^{(1)}}(A).$$

Suggerimento: maggiorare e minorare

$$K^{(2)}(A, x) = \frac{\|A\|^{(2)}\,\|x\|^{(2)}}{\|Ax\|^{(2)}}$$

in termini di

$$K^{(1)}(A, x) = \frac{\|A\|^{(1)}\,\|x\|^{(1)}}{\|Ax\|^{(1)}}.$$

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 $\alpha \in \mathbb{R}\backslash\{0\}$, 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 $K_{\|\,\cdot\,\|_\infty}(A)$.

5.1 Buon e mal condizionamento (assoluto)

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.

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

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.

Se A è mal condizionata, allora esiste un dato $x \in \mathbb{R}^n \setminus \{\underline{0}\}$ tale che δ ha "in generale" ordine di grandezza superiore all'ordine di grandezza di ε.

Infatti, se consideriamo il dato $x = A^{-1}y$, dove $y \in \mathbb{R}^n \setminus \{\underline{0}\}$ è tale che

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

si ha, come si è visto,

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

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

$$\|A\| = \frac{\|Az\|}{\|z\|},$$

risulta, come si è visto, $\delta = K(A, x)\,\varepsilon$ e quindi

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

Per cui, se $K_{\|\,\cdot\,\|}(A) >> 1$, allora per un tale dato x e per un tale dato perturbato $\widetilde{x}$ si ha che δ ha ordine di grandezza superiore all'ordine di grandezza di ε.

Esercizio. Esaminare la questione se le nozioni di buon e mal condizionamento di una matrice $A \in \mathbb{R}^{n\times n}$ possono essere considerate indipendenti dalla particolare norma $\|\cdot\|$ su $\mathbb{R}^n$ usata. Suggerimento: considerare quanto provato in un precedente esercizio.

5.2 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.

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.

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

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 $\widetilde{x} \in \mathbb{R}^2$ 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

$$C = \begin{bmatrix} a+b & a-b \\ a-b & a+b \end{bmatrix} \in \mathbb{R}^{2\times 2}$$

con $|a| \geq |b| > 0$, determinare $K_{\|\cdot\|_2}(C)$ e dire per quali a e b si ha C mal condizionata.