Norme

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

July 5, 2018

Consideriamo le quadruple $x, y \in \mathbb{R}^4$ date da

$$x = (1, 2, -1, 0) \text{ e } \ y = (0, -1, 0, 1)$$

Quanto “distano” e quanto sono “grandi” queste quadruple?

Consideriamo le funzioni $f, g : [0, 2\pi] \to \mathbb{R}$ date da

$$f(t) = \sin t \quad \text{e} \quad g(t) = \cos t, \ t \in [0, 2\pi].$$

Quanto “distano” e quanto sono “grandi” queste funzioni?

Per rispondere a questo, introduciamo una nozione astratta di “distanza” tra due oggetti matematici, come pure una nozione astratta di “grandezza” di un oggetto matematico. La prima nozione è catturata nella definizione di *metrica* su un insieme, la seconda nella definizione di *norma* su uno spazio vettoriale.

1 Metriche e spazi metrici

La nozione di metrica su un insieme permette di quantificare la distanza tra due elementi dell'insieme.

Definizione 1 *Sia M un insieme e sia $\rho : M \times M \to \mathbb{R}$. Si dice che ρ è una* metrica *su M se*

1) $\forall u, v \in M$: $\rho(u, v) \geq 0 \quad e \quad \rho(u, v) = 0 \Leftrightarrow u = v$ *(*positività*);*

2) $\forall u, v \in M$: $\rho(u, v) = \rho(v, u)$ *(*simmetria*);*

3) $\forall u, v, w \in M$: $\rho(u, v) \leq \rho(u, w) + \rho(w, v)$ *(*disuguaglianza triangolare*).*

Se ρ è una metrica su M e $u, v \in M$, il numero $\rho(u, v)$ è detto la *distanza* tra u e v nella metrica ρ.

L'insieme M con la metrica ρ su M, cioè la coppia (M, ρ), è detta uno *spazio metrico*.

Gli spazi metrici sono un'astrazione del seguente esempio elementare di spazio metrico.

Consideriamo la retta ($M = \mathbb{R}$), il piano ($M = \mathbb{R}^2$) o lo spazio ($M = \mathbb{R}^3$) e sia $d : M \times M \to \mathbb{R}$ l'ordinaria maniera di misurare la distanza tra punti: per $x, y \in M$

$$d(x, y) = \text{lunghezza del segmento di estremi } x \text{ e } y.$$

Le proprietà nella definizione di metrica sono ovviamente soddisfatte per d: in particolare la disuguaglianza triangolare esprime il noto fatto che in un triangolo un lato non è maggiore della somma degli altri due. Quindi, d è una metrica su $M = \mathbb{R}, \mathbb{R}^2, \mathbb{R}^3$ detta la *metrica euclidea*.

Si osservi che, per $x, y \in \mathbb{R}^n$ con $n \in \{1, 2, 3\}$, si ha

$$d(x, y) = \sqrt{\sum_{i=1}^{n} (y_i - x_i)^2}.$$

Per $n = 1$ si ha

$$d(x, y) = |y - x|.$$

Vedremo nel seguito parecchie metriche diverse dalla metrica euclidea. Qui, vediamo un esempio di metrica abbastanza particolare.

Sia M un insieme arbitrario e sia $\rho : M \times M \to \mathbb{R}$ data da

$$\rho(u, v) = \begin{cases} 1 \text{ se } u \neq v \\ 0 \text{ se } u = v \end{cases}, \quad u, v \in M.$$

Per tale ρ, le proprietà 1) e 2) nella definizione di metrica sono evidenti.

Esercizio. Provare la proprietà 3), vale a dire la disuguaglianza triangolare.

Chiameremo tale metrica la *metrica 0 - 1* su M.

1.1 Palle aperte e palle chiuse

Sia (M, ρ) uno spazio metrico. Per $a \in M$ e $r \geq 0$, i sottoinsiemi di M

$$B(a, r) := \{u \in M : \rho(a, u) < r\}$$

e

$$\overline{B}(a, r) := \{u \in M : \rho(a, u) \leq r\}$$

sono detti *palla aperta* e *palla chiusa*, rispettivamente, di centro a e raggio r.

Nello spazio metrico $(\mathbb{R}, d)$, si ha

$$\begin{aligned} B(a, r) &= (a - r, a + r) \\ \overline{B}(a, r) &= [a - r, a + r]. \end{aligned}$$

Nello spazio metrico $(\mathbb{R}^2, d)$, $B(a, r)$ è il cerchio di centro a e raggio r senza la circonferenza e $\overline{B}(a, r)$ è il cerchio di centro a e raggio r con la circonferenza.

Nello spazio metrico $(\mathbb{R}^3, d)$, $B(a,r)$ è la sfera di centro a e raggio r senza la superfice sferica e $\overline{B}(a,r)$ è la sfera di centro a e raggio r con la superfice sferica.

Esercizio. Descrivere le palle aperte e le palle chiuse per la metrica 0 - 1 su un insieme M.

2 Norme e spazi normati

La nozione di norma su uno spazio vettoriale permette di quantificare la grandezza di un vettore dello spazio vettoriale.

Nel seguito il vettore nullo di uno spazio vettoriale, cioè l'elemento neutro dell'addizione vettoriale, sarà denotato con $\underline{0}$.

Definizione 2 *Sia V uno spazio vettoriale e sia $\|\cdot\| : V \to \mathbb{R}$. Si dice che $\|\cdot\|$ è una* norma *su V se*

*1) $\forall u \in V$: $\|u\| \geq 0$ e $\|u\| = 0 \Leftrightarrow u = \underline{0}$ (*positività*);*

*2) $\forall \alpha \in \mathbb{R}\ \forall u \in V : \|\alpha u\| = |\alpha|\,\|u\|$ (*omogeneità*);*

*3) $\forall u, v \in V$: $\|u + v\| \leq \|u\| + \|v\|$ (*disuguaglianza triangolare*).*

Se $\|\cdot\|$ è una norma su V e $u \in V$, il numero $\|u\|$ è detto la *grandezza* di u nella norma $\|\cdot\|$.

Lo spazio vettoriale V con la norma $\|\cdot\|$ su V, cioè la coppia $(V, \|\cdot\|)$, è detta uno *spazio normato*.

Per una norma $\|\cdot\|$ su uno spazio vettoriale V, come conseguenza delle proprietà 1), 2) e 3) nella definizione, si ottengono le seguenti ulteriori proprietà 4) e 5).

4) $\forall u \in V$: $\|-u\| = \|u\|$.

5) $\forall u, v \in V$: $|\,\|u\| - \|v\|\,| \leq \|u - v\|$.

Dimostrazione di 4): per $u \in V$, si ha

$$\|-u\| = \|(-1)\,u\| = |-1|\,\|u\| = \|u\|.$$

Dimostrazione di 5): per $u, v \in V$, si tratta di provare che

$$-\|u - v\| \leq \|u\| - \|v\| \leq \|u - v\|;$$

Dalla disuguaglianza triangolare si ha

$$\|u\| = \|u - v + v\| \leq \|u - v\| + \|v\|,$$

da cui

$$\|u\| - \|v\| \leq \|u - v\|.$$

Poichè tale disuguaglianza vale anche con u e v scambiati, si ha

$$\|v\| - \|u\| \leq \|v - u\| = \|-(u - v)\| = \|u - v\|$$

e pertanto

$$\|u\| - \|v\| \geq - \|u - v\| .$$

Esercizio. Sia $(V, \|\cdot\|)$ uno spazio normato. Provare che per ogni $n \in \{1, 2, 3, \ldots\}$ e per ogni $u^{(1)}, u^{(2)}, \ldots, u^{(n)} \in V$ si ha

$$\|u^{(1)} + u^{(2)} + \cdots + u^{(n)}\| \leq \|u^{(1)}\| + \|u^{(2)}\| + \cdots + \|u^{(n)}\|.$$

Dimostrare per induzione su n. Osservare che questa proprietà è una generalizzazione della disuguaglianza triangolare, che è il caso $n = 2$.

Gli spazi normati sono astrazioni del seguente esempio elementare di spazio normato.

Consideriamo l'insieme dei vettori della retta ($V = V_1$), l'insieme dei vettori del piano ($V = V_2$) o l'insieme dei vettori dello spazio ($V = V_3$) e sia $|\cdot| : V \to \mathbb{R}$ la funzione modulo: per $\overrightarrow{v} \in V$

$$|\overrightarrow{v}| = \text{lunghezza del segmento orientato rappresentante } \overrightarrow{v}.$$

Le proprietà nella definizione di norma sono soddisfatte per $|\cdot|$. Quindi, $|\cdot|$ è una norma su V.

Introdotti dei versori ortogonali e considerando le componenti dei vettori rispetto a tali versori, lo spazio vettoriale V_n, $n \in \{1, 2, 3\}$, diventa lo spazio vettoriale $\mathbb{R}^n$ e la funzione modulo diventa la funzione $\|\cdot\|_2 : \mathbb{R}^n \to \mathbb{R}$ detta *norma euclidea* su $\mathbb{R}^n$ e data da

$$\|x\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2}, \ x \in \mathbb{R}^n.$$

Per $n = 1$ si ha

$$\|x\|_2 = |x| .$$

Esercizio. La norma euclidea su $\mathbb{R}$ (caso $n = 1$) è il valore assoluto (o modulo) $|\cdot|$. Sia $\|\cdot\| : \mathbb{R} \to \mathbb{R}$. Provare che $\|\cdot\|$ è una norma su $\mathbb{R}$ se e solo se ha la forma

$$\|x\| = C|x|, \ x \in \mathbb{R},$$

per qualche $C > 0$. Per la parte "solo se" esprimere $x \in \mathbb{R}$ come $x = x \cdot 1$.

Sia $(V, \|\cdot\|)$ uno spazio normato e sia U un sottospazio di V.

Se consideriamo la restrizione

$$\|\cdot\| |_U : U \to \mathbb{R}$$

a U di $\|\cdot\| : V \to \mathbb{R}$, questa risulta essere una norma su U, detta la *norma eredidata* dal sopraspazio V.

La restrizione a U della norma su V è una norma su U in quanto le tre proprietà nella definizione di norma sono valide per ogni $u, v, w \in V$ e, quindi, in particolare anche per ogni $u, v, w \in U \subseteq V$.

3 Metrica dedotta da una norma

Sia $(V, \|\cdot\|)$ uno spazio normato. Si può introdurre su V una metrica nel seguente modo.

Sia $\rho : V \times V \to \mathbb{R}$ data da

$$\rho(u, v) = \|v - u\|, \ u, v \in V.$$

Proviamo che ρ è una metrica su V. Positività: per $u, v \in V$, si ha, dalla positività della norma,

$$\rho(u, v) = \|v - u\| \geq 0$$

e

$$\rho(u, v) = \|v - u\| = 0 \quad \Leftrightarrow \quad v - u = \underline{0} \quad \Leftrightarrow \quad u = v.$$

Simmetria: per $u, v \in V$, si ha, dalla proprietà 4) della norma,

$$\rho(u, v) = \|v - u\| = \|-(u - v)\| = \|u - v\| = \rho(v, u).$$

Disuguaglianza triangolare: per $u, v, w \in V$, si ha, dalla proprietà di disuguaglianza triangolare della norma,

$$\begin{aligned} \rho(u, v) &= \|v - u\| = \|v - w + w - u\| \\ &\leq \|v - w\| + \|w - u\| \\ &= \rho(w, v) + \rho(u, w) = \rho(u, w) + \rho(w, v). \end{aligned}$$

La metrica ρ è detta la *metrica* su V *dedotta* dalla norma $\|\cdot\|$ su V.

Osservare che la metrica euclidea in $\mathbb{R}^n$, $n \in \{1, 2, 3\}$, data da

$$d(x, y) = \sqrt{\sum_{i=1}^{n} (y_i - x_i)^2} = \|y - x\|_2, \ x, y \in \mathbb{R}^n,$$

è la metrica dedotta dalla norma euclidea.

Esercizio. Sia V uno spazio vettoriale con almeno un elemento diverso da $\underline{0}$. Provare che la metrica $0 - 1$ su V non è una metrica dedotta da una qualche norma.

Sia $(V, \|\cdot\|)$ uno spazio normato e sia ρ la metrica su V dedotta dalla norma $\|\cdot\|$.

La palla aperta di centro $a \in V$ e raggio $r \geq 0$ è data da

$$B\,(a,r) = \{u \in V : \rho(a,u) = \|u-a\| < r\}$$

e la palla chiusa da

$$\overline{B}\,(a,r) = \{u \in V : \rho(a,u) = \|u-a\| \leq r\}\,.$$

Le palle

$$B\,(\underline{0},1) = \{u \in V : \|u\| < 1\}$$

e

$$\overline{B}\,(\underline{0},1) = \{u \in V : \|u\| \leq 1\}$$

sono dette la *palla aperta unitaria* e la *palla chiusa unitaria*, rispettivamente, dello spazio normato $(V, \|\cdot\|)$.

Ogni palla aperta (chiusa) nello spazio normato $(V, \|\cdot\|)$ si ottiene traslando la palla aperta (chiusa) unitaria opportunamente scalata.

Infatti, per ogni $a \in V$ e $r \geq 0$, si ha

$$B\,(a,r) = a + rB\,(\underline{0},1) \ \text{ se } r > 0$$

e

$$\overline{B}\,(a,r) = a + r\overline{B}\,(\underline{0},1)\,.$$

Qui si usano la notazioni:

- αF, con $\alpha \in \mathbb{R}$ e $F \subseteq V$, per denotare l'insieme $\{\alpha v : v \in F\}$ dei vettori ottenuti moltiplicando per α i vettori di F (cioè scalando con fattore α i vettori di F); questo spiega che cosa siano gli insiemi $rB\,(\underline{0},1)$ e $r\overline{B}\,(\underline{0},1)$;
- $u + E$, con $u \in V$ e $E \subseteq V$, per denotare l'insieme $\{u + v : v \in E\}$ dei vettori ottenuti sommando u ai vettori di E (cioè traslando di u i punti di E); questo spiega che cosa siano gli insiemi $a + rB\,(\underline{0},1)$ e $a + r\overline{B}\,(\underline{0},1)$.

Nel caso $r > 0$, questo fatto è provato esprimendo un elemento $u \in V$ come

$$u = a + rv, \ \text{con } v = \frac{1}{r}(u-a),$$

e notando che

$$\begin{aligned} u \in B\,(a,r) \quad &\Leftrightarrow \quad \|u-a\| < r \\ &\Leftrightarrow \quad \frac{1}{r}\,\|u-a\| = \left\|\frac{1}{r}\,(u-a)\right\| = \|v\| < 1 \\ &\Leftrightarrow \quad v \in B\,(\underline{0},1)\,. \end{aligned}$$

Analogamente si ha $u \in \overline{B}\,(a,r)$ se e solo se $v \in \overline{B}\,(\underline{0},1)$ (i segni $<$ diventano $\leq$).

Esercizio. Completare la dimostrazione nel caso $r = 0$ per le palle chiuse. Dire inoltre perché occorre assumere $r > 0$ per le palle aperte.

Concludendo si può dire che per descrivere in uno spazio normato le palle aperte (chiuse) basta descrivere la palla aperta (chiusa) unitaria.

4 Norme su $\mathbb{R}^n$

Consideriamo lo spazio vettoriale $\mathbb{R}^n$.

L'addizione vettoriale su $\mathbb{R}^n$ è definita come segue:

$$\text{per } x, y \in \mathbb{R}^n,\ \ x + y = (x_1 + y_1, \ldots, x_n + y_n)\,.$$

Il vettore nullo è la $n-$upla nulla:

$$\underline{0} = (0, 0, \ldots, 0).$$

La moltiplicazione per scalari è definita come segue:

$$\text{per } \alpha \in \mathbb{R} \text{ e } x \in \mathbb{R}^n,\ \ \alpha x = (\alpha x_1, \ldots, \alpha x_n)\,.$$

4.1 Norma infinito

La funzione $\|\cdot\|_\infty : \mathbb{R}^n \to \mathbb{R}$ data da

$$\|x\|_\infty = \max_{i \in \{1,\ldots,n\}} |x_i|\,,\ x \in \mathbb{R}^n,$$

è una norma su $\mathbb{R}^n$.

Verifichiamo le tre proprietà richieste dalla definizione. Positività: per $x \in \mathbb{R}^n$, si ha

$$\|x\|_\infty = \max_{i \in \{1,\ldots,n\}} \underbrace{|x_i|}_{\geq 0} \geq 0$$

e

$$\begin{aligned} \|x\|_\infty = \max_{i \in \{1,\ldots,n\}} \underbrace{|x_i|}_{\geq 0} = 0 \quad &\Leftrightarrow \forall i \in \{1, \ldots, n\} : |x_i| = 0 \\ \Leftrightarrow \forall i \in \{1, \ldots, n\} : x_i = 0 \ \ &\Leftrightarrow x = \underline{0}. \end{aligned}$$

Omogeneità: per $\alpha \in \mathbb{R}$ e $x \in \mathbb{R}^n$, si ha

$$\begin{aligned} \|\alpha x\|_\infty &= \max_{i \in \{1,\ldots,n\}} |\alpha x_i| = \max_{i \in \{1,\ldots,n\}} |\alpha|\,|x_i| \\ &= |\alpha| \max_{i \in \{1,\ldots,n\}} |x_i| = |\alpha|\, \|x\|_\infty\,. \end{aligned}$$

Disuguaglianza triangolare: per $x, y \in \mathbb{R}^n$, si ha

$$\|x + y\|_\infty = \max_{i \in \{1,\ldots,n\}} |x_i + y_i|$$

e avendosi, per $i \in \{1, \ldots, n\}$,

$$|x_i + y_i| \leq |x_i| + |y_i| \leq \max_{j \in \{1,\ldots,n\}} |x_j| + \max_{j \in \{1,\ldots,n\}} |y_j| = \|x\|_\infty + \|y\|_\infty$$

si ottiene

$$\|x + y\|_\infty \leq \|x\|_\infty + \|y\|_\infty\,.$$

La norma $\|\cdot\|_\infty$ è detta la *norma* ∞ o *norma del massimo* su $\mathbb{R}^n$.

Esempio. Consideriamo le quadruple di $\mathbb{R}^4$

$$x = (1, 2, -1, 0) \text{ e } y = (0, -1, 0, 1)$$

di cui si voleva determinare la "grandezza". Le grandezze di tali quadruple nella norma ∞ sono

$$\|x\|_\infty = \max\{1, 2, 1, 0\} = 2 \text{ e } \|y\|_\infty = \max\{0, 1, 0, 1\} = 1.$$

4.2 Norma 1

La funzione $\|\cdot\|_1 : \mathbb{R}^n \to \mathbb{R}$ data da

$$\|x\|_1 = \sum_{i=1}^n |x_i|\,, \; x \in \mathbb{R}^n,$$

è una norma su $\mathbb{R}^n$.

Verifichiamo le tre proprietà richieste dalla definizione. Positività: per $x \in \mathbb{R}^n$, si ha

$$\|x\|_1 = \sum_{i=1}^n \underbrace{|x_i|}_{\geq 0} \geq 0$$

e

$$\begin{aligned} \|x\|_1 = \sum_{i=1}^n \underbrace{|x_i|}_{\geq 0} = 0 \quad &\Leftrightarrow \forall i \in \{1, \dots, n\} : |x_i| = 0 \\ \Leftrightarrow \forall i \in \{1, \dots, n\} : x_i = 0 \Leftrightarrow x = \underline{0}.& \end{aligned}$$

Omogeneità: per $\alpha \in \mathbb{R}$ e $x \in \mathbb{R}^n$, si ha

$$\|\alpha x\|_1 = \sum_{i=1}^n |\alpha x_i| = \sum_{i=1}^n |\alpha|\,|x_i| = |\alpha| \sum_{i=1}^n |x_i| = |\alpha|\,\|x\|_1\,.$$

Disuguaglianza triangolare: per $x, y \in \mathbb{R}^n$, si ha

$$\begin{aligned} \|x+y\|_1 &= \sum_{i=1}^n \underbrace{|x_i + y_i|}_{\leq |x_i| + |y_i|} \leq \sum_{i=1}^n (|x_i| + |y_i|) = \sum_{i=1}^n |x_i| + \sum_{i=1}^n |y_i| \\ &= \|x\|_1 + \|y\|_1\,. \end{aligned}$$

La norma $\|\cdot\|_1$ è detta la *norma* 1 su $\mathbb{R}^n$.

Esempio. Consideriamo di nuovo le quadruple di $\mathbb{R}^4$

$$x = (1, 2, -1, 0) \quad \text{e} \quad y = (0, -1, 0, 1).$$

Le grandezze di tali quadruple nella norma 1 sono

$$\|x\|_1 = 1 + 2 + 1 + 0 = 4 \;\text{ e }\; \|y\|_1 = 0 + 1 + 0 + 1 = 2.$$

4.3 Norma p e norma 2

Sia $p \in \mathbb{R}$ con $p \geq 1$. La funzione $\|\cdot\|_p : \mathbb{R}^n \to \mathbb{R}$ data da

$$\|x\|_p := \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}}, \ x \in \mathbb{R}^n,$$

è una norma su $\mathbb{R}^n$.

Verifichiamo le prime due proprietà richieste dalla definizione. Positività: per $x \in \mathbb{R}^n$, si ha

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}} \geq 0$$

e

$$\begin{aligned} &\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}} = 0 \quad \Leftrightarrow \sum_{i=1}^{n} |x_i|^p = 0 \Leftrightarrow \forall i \in \{1, \ldots, n\} : |x_i|^p = 0 \\ &\Leftrightarrow \forall i \in \{1, \ldots, n\} : |x_i| = 0 \Leftrightarrow \forall i \in \{1, \ldots, n\} : x_i = 0 \ \Leftrightarrow x = \underline{0}. \end{aligned}$$

Omogeneità: per $\alpha \in \mathbb{R}$ e $x \in \mathbb{R}^n$, si ha

$$\begin{aligned} \|\alpha x\|_p &= \left(\sum_{i=1}^{n} |\alpha x_i|^p \right)^{\frac{1}{p}} = \left(\sum_{i=1}^{n} |\alpha|^p |x_i|^p \right)^{\frac{1}{p}} \\ &= \left(|\alpha|^p \sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}} = |\alpha| \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}} \\ &= |\alpha| \, \|x\|_p . \end{aligned}$$

La dimostrazione della disuguaglianza triangolare

$$\forall x, y \in \mathbb{R}^n : \|x + y\|_p \leq \|x\|_p + \|y\|_p ,$$

nota in questo caso di $\|\cdot\|_p$ come *disuguaglianza di Minkowski*, è più complicata e non sarà data. Più avanti si vedrà la dimostrazione per il caso $p = 2$.

La norma $\|\cdot\|_p$ è detta la *norma* p su $\mathbb{R}^n$.

Casi particolari sono la *norma* 1

$$\|x\|_1 = \left(\sum_{i=1}^{n} |x_i|^1 \right)^{\frac{1}{1}} = \sum_{i=1}^{n} |x_i| , \ x \in \mathbb{R}^n,$$

già introdotta sopra e la *norma* 2, o *norma euclidea*,

$$\|x\|_2 = \left(\sum_{i=1}^{n} |x_i|^2 \right)^{\frac{1}{2}} = \sqrt{\sum_{i=1}^{n} |x_i|^2} = \sqrt{\sum_{i=1}^{n} x_i^2}, \ x \in \mathbb{R}^n.$$

già introdotta sopra nel caso $n \in \{1,2,3\}$.

Ancora una volta consideriamo le quadruple di $\mathbb{R}^4$

$$x = (1,2,-1,0) \quad \text{e} \quad y = (0,-1,0,1).$$

Le grandezze di tali quadruple nella norma 2 sono

$$\|x\|_2 = \sqrt{1^2+2^2+1^2+0^2} = \sqrt{6} = 2.45 \text{ e } \|y\|_2 = \sqrt{0^2+1^2+0^2+1^2} = \sqrt{2} = 1.41.$$

Ricordando quanto precedentemente computato si ha:

$$\begin{array}{c} \|x\|_\infty = 2, \ \|x\|_1 = 4, \ \|x\|_2 = 2.45 \\ \|y\|_\infty = 1, \ \|y\|_1 = 2, \ \|y\|_2 = 1.41. \end{array}$$

Ovviamente, le grandezze di x e y dipendono dalla norma che si usa per misurarla.

La norma ∞ può essere vista come la norma p per $p = +\infty$. Si ha infatti

$$\forall x \in \mathbb{R}^n : \|x\|_\infty = \lim_{p \to +\infty} \|x\|_p .$$

Proviamo questo fatto.

Sia $x \in \mathbb{R}^n$. Se $x = \underline{0}$, l'uguaglianza è banalmente vera avendosi $\|x\|_\infty = 0$ e $\|x\|_p = 0$ per ogni $p \geq 1$.

Supponiamo adesso $x \neq \underline{0}$ e quindi $\|x\|_\infty > 0$. Si ha, per $p \geq 1$,

$$\begin{aligned} \|x\|_p &= \left(\sum_{i=1}^n |x_i|^p\right)^{\frac{1}{p}} = \left(\|x\|_\infty^p \sum_{i=1}^n \left(\frac{|x_i|}{\|x\|_\infty}\right)^p\right)^{\frac{1}{p}} \\ &= \|x\|_\infty \left(\sum_{i=1}^n \left(\frac{|x_i|}{\|x\|_\infty}\right)^p\right)^{\frac{1}{p}} . \end{aligned}$$

Sia ora m il numero di indici $i \in \{1,\dots,n\}$ tali che

$$|x_i| = \|x\|_\infty = \max_{j \in \{1,\dots,n\}} |x_j| .$$

Risulta, per $p \to +\infty$,

$$\left(\sum_{i=1}^n \left(\frac{|x_i|}{\|x\|_\infty}\right)^p\right)^{\frac{1}{p}} = \left(\underbrace{\sum_{\substack{i=1 \\ |x_i| < \|x\|_\infty}}^n \left(\frac{|x_i|}{\|x\|_\infty}\right)^p}_{\to 0} + \underbrace{\sum_{\substack{i=1 \\ |x_i| = \|x\|_\infty}}^n 1}_{=m}\right)^{\frac{1}{p}} \to m^0 = 1$$

e quindi

$$\|x\|_p = \|x\|_\infty \left(\sum_{i=1}^{n} \left(\frac{|x_i|}{\|x\|_\infty} \right)^p \right)^{\frac{1}{p}} \to \|x\|_\infty .$$

Per $p \in (0,1)$, la funzione $\| \cdot \|_p$ data da

$$\|x\|_p := \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}}, \ x \in \mathbb{R}^n.$$

risulta ancora essere definita e soddisfa le proprietà di positività e omogeneità. Tuttavia, essa non è più una norma, in quanto viene a mancare la disuguaglianza triangolare.

Infatti, per $x = e^{(1)} = (1,0,0,\ldots,0)$ e $y = e^{(2)} = (0,1,0,\ldots,0)$, si ha

$$\begin{aligned} \|x\|_p &= \|(1,0,0\ldots,0)\|_p = 1 \\ \|y\|_p &= \|(0,1,0,\ldots,0)\|_p = 1 \\ \|x+y\|_p &= \|(1,1,0,\ldots,0)\|_p = 2^{\frac{1}{p}} \end{aligned}$$

e quindi

$$\|x\|_p + \|y\|_p = 2 < 2^{\frac{1}{p}} = \|x+y\|_p .$$

4.4 Confronto tra norme p

Sopra, per le quadruple di $\mathbb{R}^4$

$$x = (1,2,-1,0) \quad \text{e} \quad y = (0,-1,0,1),$$

si è visto che

$$\begin{aligned} &\|x\|_1 = 4, \ \|x\|_2 = 2.45, \ \|x\|_\infty = 2, \\ &\|y\|_1 = 2, \ \|y\|_2 = 1.41, \ \|y\|_\infty = 1, \end{aligned}$$

e quindi $\|x\|_1 \geq \|x\|_2 \geq \|x\|_\infty$ e $\|y\|_1 \geq \|y\|_2 \geq \|y\|_\infty$. Questo fatto vale in generale. Si ha

$$\forall x \in \mathbb{R}^n : \|x\|_1 \geq \|x\|_2 \geq \|x\|_\infty .$$

Infatti, sia $x \in \mathbb{R}^n$ e sia $j \in \{1,\ldots,n\}$ tale che

$$|x_j| = \max_{i \in \{1,\ldots,n\}} |x_i| = \|x\|_\infty .$$

Si ha

$$\|x\|_\infty = |x_j| = \sqrt{|x_j|^2} \leq \sqrt{\sum_{i=1}^{n} |x_i|^2} = \|x\|_2 .$$

Inoltre

$$\|x\|_1^2 = \left(\sum_{i=1}^{n} |x_i| \right)^2 = \left(\sum_{i=1}^{n} |x_i| \right) \left(\sum_{j=1}^{n} |x_j| \right) = \sum_{i=1}^{n} \sum_{j=1}^{n} |x_i| \, |x_j|$$

$$= \sum_{i=1}^{n} |x_i|^2 + \underbrace{\sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} |x_i| \, |x_j|}_{\geq 0} \geq \sum_{i=1}^{n} |x_i|^2 = \|x\|_2^2$$

e quindi $\|x\|_1 \geq \|x\|_2$.

Più in generale ancora, si ha che, per ogni $x \in \mathbb{R}^n$, la funzione

$$p \mapsto \|x\|_p \, , \; p \in [1, +\infty) \, , \tag{1}$$

è decrescente. Come conseguenza si ottiene che, per ogni $x \in \mathbb{R}^n$,

$$\|x\|_\infty = \lim_{p \to +\infty} \|x\|_p = \inf_{p \in [1, +\infty)} \|x\|_p \, .$$

Per provare che (1) è decrescente, si usa il fatto che, per $a, b \geq 0$ e $\alpha \geq 1$, risulta

$$(a+b)^\alpha \geq a^\alpha + b^\alpha.$$

Osserviamo che la dimostrazione di questo fatto è banale per α intero positivo, in quanto al membro destro della disequazione sono presenti solo alcuni dei termini non negativi dello sviluppo di $(a+b)^\alpha$.

Per provare questo fatto in generale consideriamo, per $\alpha \geq 1$ e $b \geq 0$ fissati, la funzione $f : [0, +\infty) \to \mathbb{R}$ data da

$$f(a) = (a+b)^\alpha - a^\alpha - b^\alpha, \;\; a \geq 0.$$

Si ha

$$f(0) = b^\alpha - b^\alpha = 0$$

e, per $a \geq 0$,

$$f'(a) = \alpha \, (a+b)^{\alpha-1} - \alpha a^{\alpha-1} = \alpha \left((a+b)^{\alpha-1} - a^{\alpha-1} \right) \geq 0$$

essendo $a + b \geq a$ e quindi $(a+b)^{\alpha-1} \geq a^{\alpha-1}$ dal momento che $\alpha - 1 \geq 0$. Da ciò segue, per $a \geq 0$,

$$f(a) = \underbrace{f(0)}_{=0} + \underbrace{\int_0^a \underbrace{f'(c)}_{\geq 0} dc}_{\geq 0} \geq 0.$$

Come conseguenza del precedente fatto si ha (provarlo come esercizio per induzione su n): per $a_1, a_2, \ldots, a_n \geq 0$ e $\alpha \geq 1$, risulta

$$\left(\sum_{i=1}^{n} a_i\right)^{\alpha} \geq \sum_{i=1}^{n} a_i^{\alpha}.$$

Ora, per provare (1) occorre mostrare che, per $x \in \mathbb{R}^n$ e $p, q \in [1, +\infty)$ con $p \leq q$, si ha

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} \geq \|x\|_q = \left(\sum_{i=1}^{n} |x_i|^q\right)^{\frac{1}{q}},$$

esprimibile equivalentemente come

$$\left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{q}{p}} \geq \sum_{i=1}^{n} |x_i|^q.$$

Si vede subito come questo sussista avendosi

$$\left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{q}{p}} \geq \sum_{i=1}^{n} \left(|x_i|^p\right)^{\frac{q}{p}} = \sum_{i=1}^{n} |x_i|^q.$$

Come esempio della funzione (1) si consideri $x \in \mathbb{R}^n$ con m componenti uguali a 1 e $n-m$ componenti nulle. Si ha

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} = m^{\frac{1}{p}} \text{ per ogni } p \in [1, +\infty).$$

Esercizio. Trovare la funzione (1) per $x \in \mathbb{R}^n$ con m componenti uguali a α e $n-m$ componenti nulle.

Esercizio. Provare che, per $a, b > 0$ e $\alpha > 1$, si ha

$$(a+b)^{\alpha} > a^{\alpha} + b^{\alpha}.$$

Esercizio. Usando l'esercizio precedente mostrare che la funzione (1) è strettamente decrescente se x ha (almeno) due componenti non nulle.

4.5 Metrica p

Sia $p \in [1, +\infty]$. La metrica ρ_p su $\mathbb{R}^n$ dedotta dalla norma p su $\mathbb{R}^n$ è detta *metrica* p su $\mathbb{R}^n$ ed è data da

$$\rho_p(x, y) = \|y - x\|_p = \begin{cases} \left(\sum\limits_{i=1}^{n} |y_i - x_i|^p\right)^{\frac{1}{p}} & \text{se } p < +\infty \\ \max_{i \in \{1, \ldots, n\}} |y_i - x_i| & \text{se } p = +\infty \end{cases}, \quad x, y \in \mathbb{R}^n.$$

La metrica 2

$$\rho_2(x,y) = \sqrt{\sum_{i=1}^{n} |y_i - x_i|^2} = \sqrt{\sum_{i=1}^{n} (y_i - x_i)^2},\ x, y \in \mathbb{R}^n,$$

è detta anche *metrica euclidea* ed è già stata introdotta sopra nel caso $n \in \{1, 2, 3\}$.

Notare che la metrica euclidea su $\mathbb{R}^n$ è quella usata nell'analisi delle funzioni reali di n variabili reali.

Le quadruple di $\mathbb{R}^4$ viste all'inizio

$$x = (1, 2, -1, 0) \quad \text{e} \quad y = (0, -1, 0, 1)$$

distano:

-
$$\begin{aligned} \rho_1(x,y) &= \|y - x\|_1 = \|(1, 2, -1, 0) - (0, -1, 0, 1)\|_1 \\ &= \|(1, 3, -1, -1)\|_1 = 1 + 3 + 1 + 1 = 6 \end{aligned}$$

  nella metrica 1;

-
$$\rho_2(x,y) = \|y - x\|_2 = \|(1, 3, -1, -1)\|_2 = \sqrt{1^2 + 3^2 + 1^2 + 1^2} = \sqrt{12} = 3.46$$

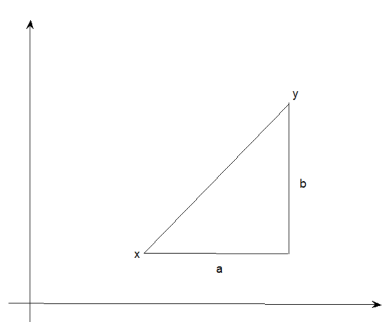
  nella metrica euclidea;

-
$$\rho_\infty(x,y) = \|y - x\|_\infty = \|(1, 3, -1, -1)\|_\infty = \max\{1, 3, 1, 1\} = 3$$

  nella metrica ∞.

Ovviamente, la distanza tra x e y dipende dalla metrica che si usa per misurarla.

Vediamo ora come una distanza in $\mathbb{R}^2$ o $\mathbb{R}^3$ varia a seconda della metrica 1, 2 o ∞ considerata.

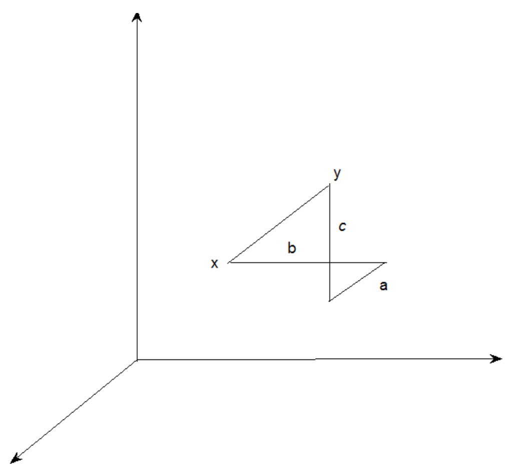


Per i punti $x, y \in \mathbb{R}^2$ in figura, si ha

$$\rho_1(x,y) \quad = \quad a+b, \ \ \rho_2(x,y) = \sqrt{a^2+b^2}, \ \ \rho_\infty(x,y) = \max\{a,b\}.$$

La distanza $\rho_2(x,y)$ è la distanza "in linea retta". La distanza $\rho_1(x,y)$ la distanza sotto il vincolo che ci si possa muovere solo orizzontalmente o verticalmente: come accade nelle grandi città dove si hanno strade parallele tagliate a 90° da altre strade parallele; per tale motivo la norma (o la metrica) 1 è chiamata anche *norma (metrica) Manhattan.*

Esercizio. Dare un significato alla distanza $\rho_\infty(x,y)$ in una grande città analogo a quello fatto per $\rho_1(x,y)$.



Per i punti $x, y \in \mathbb{R}^3$ in figura, si ha

$$\rho_1(x,y) = a+b+c, \ \ \rho_2(x,y) = \sqrt{a^2+b^2+c^2}, \ \ \rho_\infty(x,y) = \max\{a,b,c\}.$$

4.6 Palle nella norma p

Nel seguito denotiamo con $B_p(a,r)$ e $\overline{B}_p(a,r)$ la palla aperta e la palla chiusa, rispettivamente, di centro $a \in \mathbb{R}^n$ e raggio $r \geq 0$ nello spazio normato $\mathbb{R}^n$ con la norma $p \in [1,+\infty]$.

Come visto in precedenza si ha

$$B_p(a,r) = a + rB_p(\underline{0},1) \quad \text{e } \overline{B}_p(a,r) = a + r\overline{B}_p(\underline{0},1)$$

dove $B_p(\underline{0},1)$ e $\overline{B}_p(\underline{0},1)$ sono le palle unitarie. Quindi basta descrivere le palle unitarie per descrivere tutte le palle.

Per $p, q \in [1,+\infty]$ con $p \leq q$, si ha

$$B_p(\underline{0},1) \subseteq B_q(\underline{0},1) \quad \text{e} \quad \overline{B}_p(\underline{0},1) \subseteq \overline{B}_q(\underline{0},1)$$

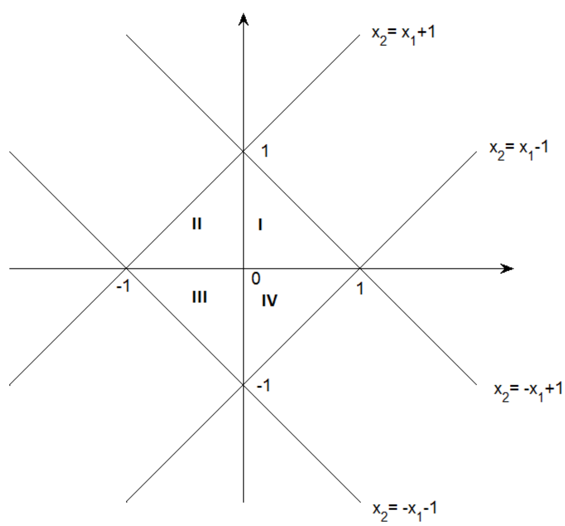
Infatti, se $x \in B_p\left(\underline{0},1\right)$, si ha, essendo $\|x\|_p$ funzione decrescente di p,

$$\|x\|_q \leq \|x\|_p < 1$$

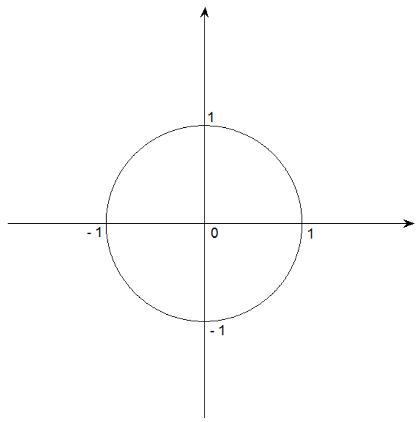
e quindi $x \in B_q\left(\underline{0},1\right)$. Il caso delle palle chiuse è analogo.

Descriviamo ora le palle unitarie in $\mathbb{R}^2$ nelle norme 1, 2 e ∞.
Si ha

$$\begin{aligned}
B_1\left(\underline{0},1\right) &= \left\{x \in \mathbb{R}^2 : \|x\|_1 = |x_1| + |x_2| < 1\right\} \\
&= \left\{x \in \mathbb{R}^2 : x_1 \geq 0 \text{ e } x_2 \geq 0 \text{ e } x_1 + x_2 < 1\right\} \\
&\quad \cup \left\{x \in \mathbb{R}^2 : x_1 < 0 \text{ e } x_2 \geq 0 \text{ e } -x_1 + x_2 < 1\right\} \\
&\quad \cup \left\{x \in \mathbb{R}^2 : x_1 < 0 \text{ e } x_2 < 0 \text{ e } -x_1 - x_2 < 1\right\} \\
&\quad \cup \left\{x \in \mathbb{R}^2 : x_1 \geq 0 \text{ e } x_2 < 0 \text{ e } x_1 - x_2 < 1\right\} \\
&= \left\{x \in \mathbb{R}^2 : x_1 \geq 0 \text{ e } x_2 \geq 0 \text{ e } x_2 < -x_1 + 1\right\} \ \mathbf{I} \\
&\quad \cup \left\{x \in \mathbb{R}^2 : x_1 < 0 \text{ e } x_2 \geq 0 \text{ e } x_2 < x_1 + 1\right\} \ \mathbf{II} \\
&\quad \cup \left\{x \in \mathbb{R}^2 : x_1 < 0 \text{ e } x_2 < 0 \text{ e } x_2 > -x_1 - 1\right\} \ \mathbf{III} \\
&\quad \cup \left\{x \in \mathbb{R}^2 : x_1 \geq 0 \text{ e } x_2 < 0 \text{ e } x_2 > x_1 - 1\right\} \ \mathbf{IV}.
\end{aligned}$$



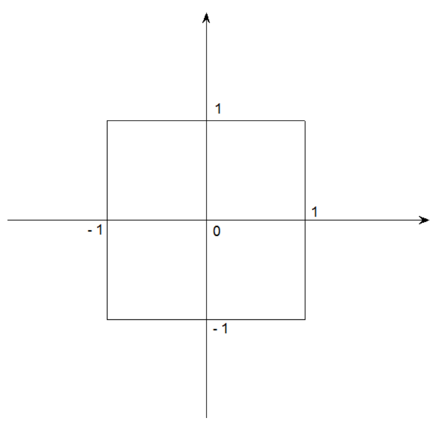
La palla aperta $B_1(\underline{0},1)$ è l'interno del quadrato costituito da **I**, **II**, **III** e **IV**. La palla chiusa $\overline{B}_1(\underline{0},1)$ è il quadrato comprensivo del bordo: ora si considera $\leq$ invece di $<$.



Si ha poi

$$B_2\left(\underline{0},1\right)=\left\{x\in\mathbb{R}^2:\|x\|_2=\sqrt{\left|x_1\right|^2+\left|x_2\right|^2}<1\right\}=\left\{x\in\mathbb{R}^2:x_1^2+x_2^2<1\right\}.$$

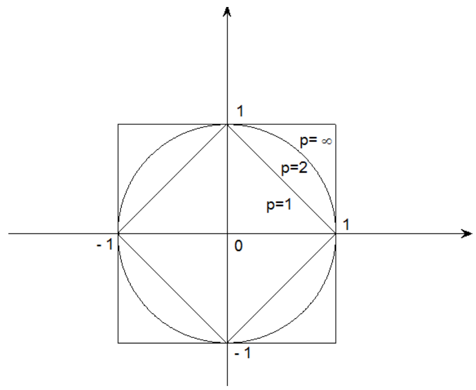
La palla aperta $B_2(\underline{0},1)$ è l'interno del cerchio. La palla chiusa $\overline{B}_2(\underline{0},1)$ è il cerchio comprensivo della circonferenza.



Si ha infine

$$B_\infty\left(\underline{0},1\right)=\left\{x\in\mathbb{R}^2:\|x\|_\infty=\max\left\{\left|x_1\right|,\left|x_2\right|\right\}<1\right\}$$
$$=\left\{x\in\mathbb{R}^2:\left|x_1\right|<1 \text{ e } \left|x_2\right|<1\right\}=\left\{x\in\mathbb{R}^2:-1<x_1<1 \text{ e } -1<x_2<1\right\}.$$

La palla aperta $B_\infty(\underline{0},1)$ è l'interno del quadrato. La palla chiusa $\overline{B}_\infty(\underline{0},1)$ è il quadrato comprensivo del bordo.



Riassumendo:

- La palla unitaria (aperta o chiusa) nella norma 2 è il cerchio di centro 0 e raggio 1.
- La palla unitaria nella norma 1 è il quadrato inscritto nella circonferenza con diagonali sugli assi.
- La palla unitaria nella norma ∞ è il quadrato circoscritto alla circonferenza con lati paralleli agli assi.

Esercizio. Descrivere in $\mathbb{R}^3$ le palle unitarie nelle norme 1, 2 e ∞.

4.7 Altre norme su $\mathbb{R}^n$

Sia $\|\cdot\|$ una norma su $\mathbb{R}^n$. A partire da $\|\cdot\|$, si può ottenere una nuova norma utilizzando una matrice $A \in \mathbb{R}^{n\times n}$ non singolare.

Precisamente, si consideri la funzione $\|\cdot\|_A : \mathbb{R}^n \to \mathbb{R}$ data da

$$\|x\|_A = \|Ax\|\,,\ x \in \mathbb{R}^n.$$

Proviamo che $\|\cdot\|_A$ è una norma su $\mathbb{R}^n$.

Positività: per $x \in \mathbb{R}^n$, si ha

$$\|x\|_A = \|Ax\| \geq 0$$

e

$$\|x\|_A = \|Ax\| = 0 \quad \Leftrightarrow Ax = \underline{0} \Leftrightarrow x = \underline{0} \quad \text{(essendo } A \text{ non singolare).}$$

Omogeneità: per $\alpha \in \mathbb{R}$ e $x \in \mathbb{R}^n$, si ha

$$\|\alpha x\|_A = \|A(\alpha x)\| = \|\alpha Ax\| = |\alpha|\,\|Ax\| = |\alpha|\,\|x\|_A\,.$$

Disuguaglianza triangolare: per $x, y \in \mathbb{R}^n$, si ha

$$\|x+y\|_A = \|A(x+y)\| = \|Ax+Ay\| \leq \|Ax\| + \|Ay\| = \|x\|_A + \|y\|_A\,.$$

Esercizio. Se A è singolare, allora la funzione $\|\cdot\|_A$ non è una norma su $\mathbb{R}^n$. Quale proprietà nella definizione di norma viene a mancare?

Esercizio. Dire se la seguente funzione $\mathbb{R}^2 \to \mathbb{R}$:

$$x \mapsto |2x_1 + 3x_2| + |x_1 - x_2|\,,\ x \in \mathbb{R}^2,$$

è una norma su $\mathbb{R}^2$. Poi, dire anche se la seguente funzione $\mathbb{R}^3 \to \mathbb{R}$:

$$x \mapsto \max\left\{|x_1 + 2x_2 + 3x_3|\,, |2x_1 + x_2 + 2x_3|\,, |-x_1 + x_2 + x_3|\right\},\ x \in \mathbb{R}^3,$$

è una norma su $\mathbb{R}^3$.

Esercizio. Sia $\|\cdot\|$ una norma su $\mathbb{R}^n$ e sia $C > 0$. Provare che $\|\cdot\|' : \mathbb{R}^n \to \mathbb{R}$ data da

$$\|x\|' = C\|x\|,\ x \in \mathbb{R}^n$$

è una norma su $\mathbb{R}^n$, interpretandola come norma $\|\cdot\|_A$ per una idonea matrice $A \in \mathbb{R}^{n\times n}$.

Come esempi di norme $\|\cdot\|_A$ si hanno le norme p pesate.

Cnsiderando la norma $\|\cdot\|_p$, con $p \in [1, +\infty)$, e la matrice non singolare

$$A = \operatorname{diag}(w_1^{\frac{1}{p}}, \ldots, w_n^{\frac{1}{p}}),$$

dove $w_i > 0$ per $i \in \{1, \ldots, n\}$, si ottiene la *norma p w-pesata* (con $w = (w_1, \ldots, w_n)$):

$$\begin{aligned} \|x\|_{p,w} &:= \|Ax\|_p = \left\|(w_1^{\frac{1}{p}} x_1, \ldots, w_n^{\frac{1}{p}} x_n)\right\|_p \\ &= \left(\sum_{i=1}^n |w_i^{\frac{1}{p}} x_i|^p\right)^{\frac{1}{p}} = \left(\sum_{i=1}^n w_i|x_i|^p\right)^{\frac{1}{p}},\ x \in \mathbb{R}^n. \end{aligned}$$

Considerando poi la norma $\|\cdot\|_\infty$ e la matrice non singolare

$$A = \operatorname{diag}(w_1, \ldots, w_n)\,,$$

dove $w_i > 0$ per $i \in \{1, \ldots, n\}$, si ottiene la *norma ∞ w-pesata*:

$$\begin{aligned} \|x\|_{\infty,w} &:= \|Ax\|_\infty = \|(w_1x_1, \ldots, w_nx_n)\|_\infty \\ &= \max_{i\in\{1,\ldots,n\}} |w_ix_i| = \max_{i\in\{1,\ldots,n\}} w_i|x_i|,\ x \in \mathbb{R}^n. \end{aligned}$$

5 Norme su $C([a, b])$

Consideriamo lo spazio vettoriale $C([a, b])$ delle funzioni $[a, b] \to \mathbb{R}$ continue.

L'addizione vettoriale su $C([a, b])$ è definita come segue:

$$\text{per } f, g \in C([a, b])\,,\ \ (f + g)(t) = f(t) + g(t)\,,\ t \in [a, b]\,.$$

Il vettore nullo $\underline{0}$ è la funzione nulla

$$\underline{0}(t) = 0, \ t \in [a, b].$$

La moltiplicazione per scalari è definita come segue:

$$\text{per } \alpha \in \mathbb{R} \text{ e } f \in C([a, b]), \ (\alpha f)(t) = \alpha f(t), \ t \in [a, b].$$

5.1 Norma infinito

La funzione $\|\cdot\|_\infty : C([a, b]) \to \mathbb{R}$ data da

$$\|f\|_\infty = \max_{t \in [a,b]} |f(t)|, \ f \in C([a, b]),$$

è una norma su $C([a, b])$.

Esercizio. Guardando alla dimostrazione del fatto che la funzione $\|\cdot\|_\infty : \mathbb{R}^n \to \mathbb{R}$ è una norma, provare che anche $\|\cdot\|_\infty : C([a, b]) \to \mathbb{R}$ è una norma.

La norma $\|\cdot\|_\infty$ è detta la *norma* ∞ o *norma del massimo* su $C([a, b])$.

La grandezza della funzione

$$f(t) = t^\alpha, \ t \in [0, 1],$$

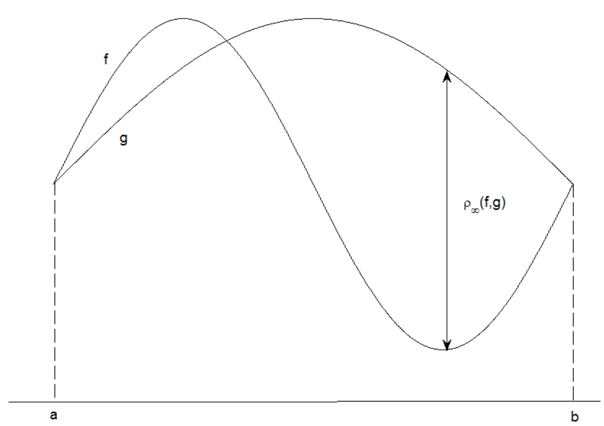
dove $\alpha > 0$, nella norma ∞ è

$$\|f\|_\infty = \max_{t \in [0,1]} |t^\alpha| = \max_{t \in [0,1]} t^\alpha = 1.$$

Nella norma ∞ la grandezza di una funzione è data dal suo massimo valore assoluto e quindi, nella metrica dedotta

$$\rho_\infty(f, g) = \|g - f\|_\infty = \max_{t \in [a,b]} |g(t) - f(t)|, \ f, g \in C([a, b]),$$

la distanza tra due funzioni è data dal massimo valore assoluto della loro differenza.



Esercizio. Si considerino le funzioni $f, g : [0, 2\pi] \to \mathbb{R}$ viste all'inizio:

$$f(t) = \sin t \;\; \text{e} \;\; g(t) = \cos t,\;\; t \in [0, 2\pi].$$

Determinare la loro grandezza nella norma ∞ e la loro distanza nella metrica ∞.

5.2 Norma 1

Vi sono altro modi, oltre alla norma ∞, di quantificare la grandezza di una funzione in $C([a, b])$ e quindi la distanza tra due funzioni in $C([a, b])$.

La funzione $\|\cdot\|_1 : C([a, b]) \to \mathbb{R}$ data da

$$\|f\|_1 = \int_a^b |f(t)|\, dt,\;\; f \in C([a, b]),$$

è una norma su $C([a, b])$.

Esercizio. Guardando alla dimostrazione del fatto che la funzione $\|\cdot\|_1 : \mathbb{R}^n \to \mathbb{R}$ è una norma, provare che anche $\|\cdot\|_1 : C([a, b], \mathbb{R}) \to \mathbb{R}$ è una norma.

La norma $\|\cdot\|_1$ è detta *norma* 1 o *norma integrale* su $C([a, b])$.

La grandezza della funzione

$$f(t) = t^\alpha,\;\; t \in [0, 1],$$

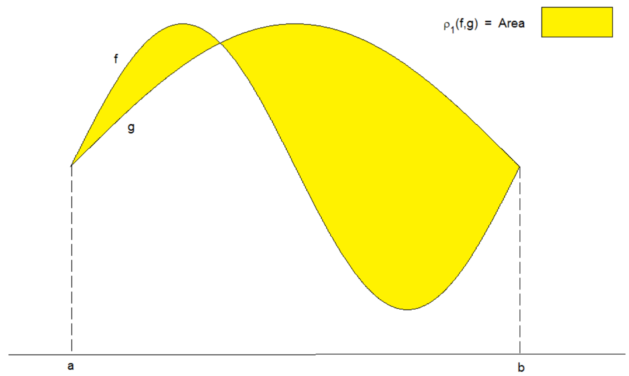
dove $\alpha > 0$, nella norma 1 è

$$\|f\|_1 = \int_0^1 |t^\alpha|\, dt = \int_0^1 t^\alpha dt = \left[\frac{1}{\alpha + 1} t^{\alpha+1}\right]_0^1 = \frac{1}{\alpha + 1}.$$

Nella norma 1 la grandezza di una funzione è data dall'integrale del suo valore assoluto e quindi, nella metrica dedotta

$$\rho_1(f, g) = \int_a^b |g(t) - f(t)|\, dt,\;\; f, g \in C([a, b]),$$

la distanza tra due funzioni è data dall'integrale del valore assoluto della loro differenza.



Esercizio. Si considerino di nuovo le funzioni $f, g : [0, 2\pi] \to \mathbb{R}$:

$$f(t) = \sin t \quad \text{e} \quad g(t) = \cos t, \ t \in [0, 2\pi].$$

Determinare la loro grandezza nella norma 1 e la loro distanza nella metrica 1.

Le grandezze e le distanze in norma ∞ e 1 possono essere molto diverse.

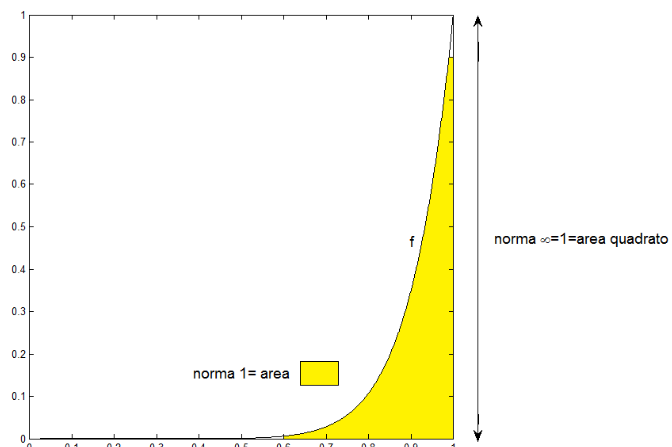
Si è visto che per

$$f(t) = t^{\alpha}, \ t \in [0, 1],$$

risulta

$$\|f\|_{\infty} = 1 \text{ e } \|f\|_1 = \frac{1}{\alpha + 1}.$$

Per α grande, le grandezze di f nelle due norme sono molte diverse.



5.3 Norma p e norma 2

Sia $p \in [1, +\infty)$. La funzione $\|\cdot\|_p : C([a, b]) \to \mathbb{R}$ data da

$$\|f\|_p = \left(\int_a^b |f(t)|^p \, dt \right)^{\frac{1}{p}}, \ f \in C([a, b]),$$

è una norma su $C([a,b])$, detta la *norma* p su $C([a,b])$.

Esercizio. Guardando alla dimostrazione del fatto che la funzione $\|\cdot\|_p : \mathbb{R}^n \to \mathbb{R}$ soddisfa le proprietà di positività e omogeneità, provare che anche $\|\cdot\|_p : C([a,b],\mathbb{R}) \to \mathbb{R}$ soddisfa tali proprietà.

Come nel caso di $\mathbb{R}^n$, la disuguaglianza triangolare per $\|\cdot\|_p : C([a,b],\mathbb{R}) \to \mathbb{R}$ prende il nome di *disuguaglianza di Minkowski*. La sua dimostrazione non sarà data.

Casi particolari sono la norma 1 già introdotta in precedenza e la norma 2

$$\|f\|_2 = \sqrt{\int_a^b f(t)^2\,dt},\ f \in C([a,b]).$$

La grandezza della funzione

$$f(t) = t^\alpha,\ t \in [0,1],$$

dove $\alpha > 0$, nella norma p è

$$\begin{aligned}\|f\|_p &= \left(\int_0^1 |t^\alpha|^p\,dt\right)^{\frac{1}{p}} = \left(\int_0^1 (t^\alpha)^p\,dt\right)^{\frac{1}{p}} = \left(\int_0^1 t^{\alpha p}dt\right)^{\frac{1}{p}} \\ &= \left(\left[\frac{1}{\alpha p+1}t^{\alpha p+1}\right]_0^1\right)^{\frac{1}{p}} = \left(\frac{1}{\alpha p+1}\right)^{\frac{1}{p}}.\end{aligned}$$

Esercizio. Si considerino ancora una volta le funzioni $f,g : [0,2\pi] \to \mathbb{R}$ viste all'inizio:

$$f(t) = \sin t \ \text{ e } \ g(t) = \cos t,\ t \in [0,2\pi].$$

Determinare la loro grandezza nella norma 2 e la loro distanza nella metrica 2.

Come nel caso dello spazio $\mathbb{R}^n$, si può dimostrare che si ha

$$\forall f \in C([a,b]) : \lim_{p\to+\infty} \|f\|_p = \|f\|_\infty. \tag{2}$$

Inoltre, si ha che, per $p \in (0,1)$, la funzione $\|\cdot\|_p : C([a,b],\mathbb{R}) \to \mathbb{R}$, pur soddisfacendo le proprietà di positività e omogeneità, non è una norma in quanto non soddisfa la disuguaglianza triangolare. Esercizio. Provare questo utilizzando le funzioni

$$f(t) = \frac{t}{b-a} \ \text{ e } \ g(t) = \frac{b-t}{b-a},\ t \in [a,b].$$

Infine, si può dimostrare che si ha

$$\forall f \in C([a,b])\ \forall p,q \in [1,+\infty] \text{ con } p \le q : \|f\|_p \le (b-a)^{\frac{1}{p}-\frac{1}{q}} \|f\|_q. \tag{3}$$

Esercizio. Considerare la disuguglianza (3) per $[a,b]=[0,1]$. Cosa c'è di diverso rispetto allo spazio $\mathbb{R}^n$?

Esercizio. Provare la disuguaglianza nel caso $q=+\infty$.

Esercizio. Per la funzione

$$f(t)=t^{\alpha},\ t\in[0,1],$$

dove $\alpha>0$, verificare che le due proprietà sopra (2) e (3) sussistono.

Come per lo spazio $\mathbb{R}^n$, anche nello spazio $C([a,b])$ si introducono le norme p pesate.

Sia $p\in[1,+\infty]$. La *norma p w-pesata* è data da

$$\|f\|_{p,w}=\begin{cases}\left(\int\limits_a^b w\left(t\right)\left|f\left(t\right)\right|^p dt\right)^{\frac{1}{p}} & \text{se } p<+\infty \\ \max_{t\in[a,b]} w\left(t\right)\left|f\left(t\right)\right| & \text{se } p=+\infty\end{cases}\quad,\ f\in C([a,b]),$$

dove il peso $w:[a,b]\to\mathbb{R}$ è una funzione non negativa integrabile per $p<+\infty$, non negativa limitata per $p=+\infty$, che si annulla in un insieme di misura zero.

Si dimostra che tale funzione $\|\ \cdot\|_{p,w}:C([a,b])\to\mathbb{R}$ è una norma.

6 Equivalenza di norme

L'equivalenza di norme stabilisce quando due norme forniscono grandezze comparabili.

Definizione 3 *Sia V uno spazio vettoriale e siano $\|\cdot\|$ e $\|\cdot\|'$ norme su V. Si dice che $\|\cdot\|$ è equivalente a $\|\cdot\|'$ se esistono $c,C>0$ tali che*

$$\forall u\in V: c\left\|u\right\|\leq\left\|u\right\|'\leq C\left\|u\right\|.$$

Esercizio. La nozione sopra introdotta è una relazione sull'insieme delle norme su V. Provare che si tratta di una relazione di equivalenza, cioè che gode delle proprietà riflessiva, simmetrica e transitiva. Questo spiega il nome "equivalenza" usato per questa relazione.

Notare che nella definizione di equivalenza di norme la condizione

$$\forall u\in V: c\left\|u\right\|\leq\left\|u\right\|'\leq C\|u\|$$

può essere sostituita da

$$\forall u\in V\setminus\{\underline{0}\}: c\left\|u\right\|\leq\left\|u\right\|'\leq C\left\|u\right\|,$$

vale a dire

$$\forall u\in V\setminus\{\underline{0}\}: c\leq\frac{\|u\|'}{\|u\|}\leq C,$$

in quanto $c\,\|u\| \leq \|u\|' \leq C\,\|u\|$ sussiste per $u = \underline{0}$.

Per cui, si ha che $\|\cdot\|$ è equivalente a $\|\cdot\|'$ se e solo se

$$\inf \mathcal{A} > 0 \quad \text{e} \quad \sup \mathcal{A} < +\infty,$$

dove

$$\mathcal{A} = \left\{ \frac{\|u\|'}{\|u\|} : u \in V \setminus \{\underline{0}\} \right\}.$$

6.1 Equivalenza in $\mathbb{R}^n$

Sia $p \in [1, +\infty)$. In $\mathbb{R}^n$, si ha che la norma ∞ è equivalente alla norma p.

Infatti, per $x \in \mathbb{R}^n$, si ha

$$c\|x\|_\infty \leq \|x\|_p \leq C\|x\|_\infty$$

con $c = 1$ e $C = n^{\frac{1}{p}}$:

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}} \leq \left(\sum_{i=1}^{n} \|x\|_\infty^p \right)^{\frac{1}{p}} = \left(n\,\|x\|_\infty^p\right)^{\frac{1}{p}} = n^{\frac{1}{p}}\,\|x\|_\infty$$

$$\text{si ha} \leq \text{in quanto } \forall i \in \{1, \ldots, n\} : |x_i|^p \leq \|x\|_\infty^p \text{ e quindi } \sum_{i=1}^{n} |x_i|^p \leq \sum_{i=1}^{n} \|x\|_\infty^p$$

e

$$\|x\|_\infty = \inf_{q \in [1, +\infty)} \|x\|_q \leq \|x\|_p\,.$$

Come conseguenza del fatto che la norma ∞ è equivalente a ogni norma p, si ha che in $\mathbb{R}^n$ tutte le norme p sono equivalenti.

Infatti, per $p, q \in [1, +\infty)$, la norma p è equivalente alla norma ∞ (proprietà simmetrica) e la norma ∞ è equivalente alla norma q, per cui la norma p è equivalente alla norma q (proprietà transitiva).

In $\mathbb{R}^n$, la norma 2 non è solo equivalente a tutte le norme p, ma è equivalente a tutte le norme.

Teorema 4 *Sia $\|\cdot\|$ una norma su $\mathbb{R}^n$. Si ha che $\|\cdot\|_2$ è equivalente a $\|\cdot\|$.*

Dimostrazione. Sia

$$\mathcal{A} = \left\{ \frac{\|x\|}{\|x\|_2} : x \in \mathbb{R}^n \setminus \{\underline{0}\} \right\}.$$

e proviamo che

$$\inf \mathcal{A} > 0 \quad \text{e} \quad \sup \mathcal{A} < +\infty.$$

Si ha

$$\mathcal{A} = \left\{ \frac{\|x\|}{\|x\|_2} : x \in \mathbb{R}^n \setminus \{\underline{0}\} \right\} = \{\|y\| : y \in S\} \quad \text{con} \quad S = \{y \in \mathbb{R}^n : \|y\|_2 = 1\}\,.$$

avendosi, per $x \in \mathbb{R}^n \setminus \{\underline{0}\}$,

$$\frac{\|x\|}{\|x\|_2} = \|\frac{1}{\|x\|_2}x\| \quad \text{con} \quad \|\frac{1}{\|x\|_2}x\|_2 = \frac{\|x\|_2}{\|x\|_2} = 1.$$

Consideriamo ora la funzione $f = \| \cdot \| : \mathbb{R}^n \to \mathbb{R}$ data da

$$f(y) = \|y\|, \; y \in \mathbb{R}^n,$$

per la quale

$$\mathcal{A} = \{f(y) : y \in S\} = f(S).$$

Se proviamo che f è continua e che S è chiuso e limitato, allora dal Teorema di Weierstrass seguirà che $\mathcal{A} = f(S)$ ha massimo e minimo. Per cui,

$$\inf \mathcal{A} = \min \mathcal{A} \geq 0 \quad \text{e} \quad \sup \mathcal{A} = \max \mathcal{A} < +\infty.$$

Esercizio. Provare che $\min \mathcal{A} > 0$.

Si sarà quindi ottenuto

$$\inf \mathcal{A} > 0 \quad \text{e} \quad \sup \mathcal{A} < +\infty$$

e pertanto l'equivalenza di $\| \cdot \|_2$ e $\| \cdot \|$ risulterà provata.

Proviamo che f è continua.

Ricordiamo ora che in Analisi II si è usata la metrica euclidea per definire la nozione di continuità di una funzione $g : D \subseteq \mathbb{R}^n \to \mathbb{R}$: si dice che g è continua in $y \in D$ se

$$\forall \varepsilon > 0 \; \exists \delta > 0 \; \forall x \in D : \; \|x - y\|_2 \leq \delta \; \Rightarrow \; |g(x) - g(y)| \leq \varepsilon;$$

poi si dice che g è continua se è continua in ogni $y \in D$.

Particolari funzioni continue sono le funzioni Lipschitz continue (o lipschitziane): si dice che g è Lipschitz continua se

$$\exists L \geq 0 \; \forall x, y \in D : |g(x) - g(y)| \leq L\|x - y\|_2.$$

Proviamo che f è Lipschitz continua.

Osserviamo che si ha

$$\forall z \in \mathbb{R}^n : \|z\| \leq \sum_{i=1}^{n} \|e^{(i)}\| \cdot \|z\|_2 \,, \tag{4}$$

Infatti, per $z \in \mathbb{R}^n$, si ha

$$\begin{aligned} \|z\| &= \|\sum_{i=1}^{n} z_i e^{(i)}\| \leq \sum_{i=1}^{n} \|z_i e^{(i)}\| = \sum_{i=1}^{n} \underbrace{|z_i|}_{\leq \|z\|_\infty \leq \|z\|_2} \|e^{(i)}\| \\ &\leq \sum_{i=1}^{n} \|z\|_2 \, \|e^{(i)}\| = \sum_{i=1}^{n} \|e^{(i)}\| \cdot \|z\|_2 \,. \end{aligned}$$

Dalla (4) segue immediatamente che f è Lipschitz continua: per $x, y \in \mathbb{R}^n$ si ha

$$|f(x) - f(y)| = |\|x\| - \|y\|| \leq \|x - y\| \leq \sum_{i=1}^{n} \|e^{(i)}\| \cdot \|x - y\|_2 .$$

dove il primo $\leq$ segue dalla proprietà 5) delle norme e il secondo dalla (4) con $z = x - y$.

Esercizio. Provare che l'insieme S è chiuso. Suggerimento: usare la caratterizzazione di un insieme chiuso tramite successioni e usare la continuità della norma euclidea (la Lipschitz continuità segue dalla 5) nel caso $\|\cdot\| = \|\cdot\|_2$: per $x, y \in \mathbb{R}^n$ si ha $|\|x\|_2 - \|y\|_2| \leq \|x - y\|_2$).

Esercizio. Provare che l'insieme S è limitato. ■

Dal precedente teorema segue immediatamente che tutte le norme su $\mathbb{R}^n$ sono equivalenti.

Infatti, se $\|\cdot\|$ e $\|\cdot\|'$ sono norme su $\mathbb{R}^n$, si ha che $\|\cdot\|$ è equivalente a $\|\cdot\|_2$ e $\|\cdot\|_2$ è equivalente a $\|\cdot\|'$ e, quindi, $\|\cdot\|$ è equivalente a $\|\cdot\|'$.

Osservare che nella precedente dimostrazione si è anche mostrato che una qualsiasi norma $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$, la funzione f nella dimostrazione, è una funzione Lipschitz continua.

6.2 Non equivalenza in $C([a, b])$

Siano $p, q \in [1, +\infty]$ con $p < q$. Nonostante si abbia

$$\forall f \in C([a, b]) : \|f\|_p \leq (b - a)^{\frac{1}{p} - \frac{1}{q}} \|f\|_q ,$$

la norma q non è equivalente alla norma p, mancando l'altra disuguaglianza richiesta dalla definizione.

Infatti, consideriamo il caso $[a, b] = [0, 1]$. Posto

$$\mathcal{A} = \left\{ \frac{\|f\|_p}{\|f\|_q} : f \in C([0, 1]) \setminus \{\underline{0}\} \right\},$$

proviamo che

$$\inf \mathcal{A} = 0.$$

Esercizio. Si ha $\sup \mathcal{A} = +\infty$?

Si è visto che, dato $\alpha > 0$, per la funzione

$$\widehat{f}_\alpha(t) = t^\alpha,\ t \in [0, 1],$$

risulta

$$\|\widehat{f}_\alpha\|_p = \left(\frac{1}{\alpha p + 1}\right)^{\frac{1}{p}} \quad \text{e} \quad \|\widehat{f}_\alpha\|_q = \begin{cases} \left(\frac{1}{\alpha q + 1}\right)^{\frac{1}{q}} & \text{se } q < +\infty \\ 1 \text{ se } q = +\infty. \end{cases}$$

Nel caso $q < +\infty$, risulta

$$\begin{aligned} \frac{\|\widehat{f}_\alpha\|_p}{\|\widehat{f}_\alpha\|_q} &= \frac{\left(\frac{1}{\alpha p+1}\right)^{\frac{1}{p}}}{\left(\frac{1}{\alpha q+1}\right)^{\frac{1}{q}}} = \frac{(\alpha q+1)^{\frac{1}{q}}}{(\alpha p+1)^{\frac{1}{p}}} \\ &= (\alpha q+1)^{\frac{1}{q}-\frac{1}{p}} \left(\frac{\alpha q+1}{\alpha p+1}\right)^{\frac{1}{p}} \\ &\to 0 \cdot \left(\frac{q}{p}\right)^{\frac{1}{p}} = 0,\ \alpha \to +\infty, \end{aligned}$$

e quindi $\inf \mathcal{A} = 0$.

Esercizio. Provare $\inf \mathcal{A} = 0$ nel caso $q = +\infty$.

Nel caso generale di un intervallo $[a, b]$ qualsiasi, si considera, per $\alpha > 0$, la funzione

$$f_\alpha(s) = \widehat{f}_\alpha\left(\frac{s-a}{b-a}\right),\ s \in [a, b],$$

per cui

$$\begin{aligned} \|f_\alpha\|_p &= \left(\int_a^b |f_\alpha(s)|^p\, ds\right)^{\frac{1}{p}} = \left(\int_a^b \left|\widehat{f}_\alpha\left(\frac{s-a}{b-a}\right)\right|^p ds\right)^{\frac{1}{p}} \\ &= \left(\int_0^1 \left|\widehat{f}_\alpha(t)\right|^p (b-a)\, dt\right)^{\frac{1}{p}},\ t = \frac{s-a}{b-a}, \\ &= (b-a)^{\frac{1}{p}} \left(\int_0^1 \left|\widehat{f}_\alpha(t)\right|^p dt\right)^{\frac{1}{p}} = (b-a)^{\frac{1}{p}}\, \|\widehat{f}_\alpha\|_p \end{aligned}$$

e, analogamente, per $q < +\infty$,

$$\|f_\alpha\|_q = (b-a)^{\frac{1}{q}}\, \|\widehat{f}_\alpha\|_q.$$

Quindi, per $q < +\infty$,

$$\frac{\|f_\alpha\|_p}{\|f_\alpha\|_q} = \frac{(b-a)^{\frac{1}{p}}\|\widehat{f}_\alpha\|_p}{(b-a)^{\frac{1}{q}}\|\widehat{f}_\alpha\|_q} = (b-a)^{\frac{1}{p}-\frac{1}{q}}\, \frac{\|\widehat{f}_\alpha\|_p}{\|\widehat{f}_\alpha\|_q} \to 0,\ \alpha \to +\infty,$$

da cui $\inf \mathcal{A} = 0$.

Esercizio. Provare $\inf \mathcal{A} = 0$ nel caso $q = +\infty$.

In conclusione, si ha che in $C([a, b])$ due qualsiasi norme p non sono equivalenti.

6.3 Equivalenza di norme in spazi di dimensione finita

L'equivalenza di tutte le norme non si ha solo nello spazio $\mathbb{R}^n$, ma in ogni spazio vettoriale di dimensione finita come ora proveremo.

Sia V uno spazio vettoriale di dimensione finita n.

Fissata una base $v^{(1)}, \dots, v^{(n)}$ per V, lo spazio vettoriale V è isomorfo allo spazio vettoriale $\mathbb{R}^n$ tramite l'isomorfismo $\Phi : V \to \mathbb{R}^n$ dato da

$$\Phi(u) = (x_1, \dots, x_n), \ u \in V,$$

dove $x_1, \dots, x_n \in \mathbb{R}$ sono le componenti di u nella base $v^{(1)}, \dots, v^{(n)}$, cioè sono gli unici scalari tali che

$$u = x_1 v^{(1)} + \cdots + x_n v^{(n)}.$$

Osservare che $\Phi^{-1} : \mathbb{R}^n \to V$ è dato da

$$\Phi^{-1}(x) = x_1 v^{(1)} + \cdots + x_n v^{(n)}, \ x \in \mathbb{R}^n.$$

Sia ora $\| \cdot \|$ una norma su V. Consideriamo la funzione $\| \cdot \|^* : \mathbb{R}^n \to \mathbb{R}$ data da

$$\|x\|^* = \|\Phi^{-1}(x)\| = \|x_1 v^{(1)} + \cdots + x_n v^{(n)}\|, \ x \in \mathbb{R}^n.$$

Osservare che, per $u \in V$, si ha $\|u\| = \|\Phi(u)\|^*$.

Proviamo che questa è una norma su $\mathbb{R}^n$.

Positività. Per $x \in \mathbb{R}^n$, si ha

$$\|x\|^* = \left\| x_1 v^{(1)} + \cdots + x_n v^{(n)} \right\| \geq 0$$

e

$$\begin{aligned}
&\left\| x_1 v^{(1)} + \cdots + x_n v^{(n)} \right\| = 0 \ \Leftrightarrow \ x_1 v^{(1)} + \cdots + x_n v^{(n)} = \underline{0} \\
&\Leftrightarrow \ x_1 = \cdots = x_n = 0 \ \ \text{(essendo } v^{(1)}, \dots, v^{(n)} \text{ linearmente indipendenti)} \\
&\Leftrightarrow x = \underline{0}.
\end{aligned}$$

Omogeneità. Per $\alpha \in \mathbb{R}$ e $x \in \mathbb{R}^n$, si ha

$$\begin{aligned}
\|\alpha x\|^* &= \left\| \alpha x_1 v^{(1)} + \cdots + \alpha x_n v^{(n)} \right\| = \left\| \alpha \left(x_1 v^{(1)} + \cdots + x_n v^{(n)} \right) \right\| \\
&= |\alpha| \left\| x_1 v^{(1)} + \cdots + x_n v^{(n)} \right\| = |\alpha| \, \|x\|^*.
\end{aligned}$$

Disuguaglianza triangolare. Per $x, y \in \mathbb{R}^n$, si ha

$$\begin{aligned}
\|x+y\|^* &= \left\| (x_1 + y_1) v^{(1)} + \cdots + (x_n + y_n) v^{(n)} \right\| \\
&= \left\| x_1 v^{(1)} + \cdots + x_n v^{(n)} + y_1 v^{(1)} + \cdots + y_n v^{(n)} \right\| \\
&\leq \left\| x_1 v^{(1)} + \cdots + x_n v^{(n)} \right\| + \left\| y_1 v^{(1)} + \cdots + y_n v^{(n)} \right\| \\
&= \|x\|^* + \|y\|^*.
\end{aligned}$$

Come esempio di spazio vettoriale a dimensione finita e di norma $\|\ \cdot\ \|^*$, consideriamo lo spazio vettoriale $\Pi_n([a,b])$ delle funzioni $[a,b] \to \mathbb{R}$ polinomiali di grado $\leq n$, che è un sottospazio di $C([a,b])$ di dimensione finita $n+1$.

Una sua base è data dalle funzioni polinomiali

$$t \mapsto t^i,\ t \in [a,b],\ \text{per } i \in \{0,1,\ldots,n\}.$$

Le componenti di un elemento $f \in \Pi_n([a,b])$ in tale base sono i coefficienti $a_0, a_1, \ldots, a_n$ nella forma canonica del polinomio f:

$$f(t) = a_0 + a_1 t + \cdots + a_n t^n,\ t \in [a,b].$$

Sia $\Pi_n([a,b])$ è dotato della norma $\|\cdot\|_p$, dove $p \in [1,+\infty]$, ereditata dal sopraspazio $C([a,b])$. La norma $\|\cdot\|_p^*$ su $\mathbb{R}^{n+1}$ è data da: per $(a_0, a_1, \ldots, a_n) \in \mathbb{R}^{n+1}$,

$$\begin{aligned}
\|(a_0,a_1,\ldots,a_n)\|_p^* &= \|\Phi^{-1}(a_0,a_1,\ldots,a_n)\|_p \\
&= \|f\|_p \\
&\quad (\text{con } f(t) = a_0 + a_1 t + \cdots + a_n t^n,\ t \in [a,b]) \\
&= \begin{cases} \left(\int\limits_a^b |a_0 + a_1 t + \cdots + a_n t^n|^p\, dt\right)^{\frac{1}{p}} & \text{se } p < +\infty \\ \\ \max_{t \in [a,b]} |a_0 + a_1 t + \cdots + a_n t^n| & \text{se } p = +\infty. \end{cases}
\end{aligned}$$

Esercizio. Descrivere nel caso $n = 1$, cioè nel caso di polinomi di grado ≤ 1 e $[a,b] = [0,1]$ le norme $\|\ \cdot\ \|_1^*$, $\|\ \cdot\ \|_2^*$ e $\|\ \cdot\ \|_\infty^*$. Questo vuol dire, dato $(a_0,a_1) \in \mathbb{R}^2$, determinare $\|(a_0,a_1)\|_1^*$, $\|(a_0,a_1)\|_2^*$ e $\|(a_0,a_1)\|_\infty^*$ come funzioni di a_0 e a_1.

Tornando a parlare dello spazio vettoriale generale V di dimensione finita n, proviamo ora che in V tutte le norme sono equivalenti.

Siano $\|\cdot\|$ e $\|\cdot\|'$ norme su V. Poichè in $\mathbb{R}^n$ tutte le norme sono equivalenti, le norme $\|\cdot\|^*$ e $\|\cdot\|'^*$ su $\mathbb{R}^n$ sono equivalenti, cioè esistono $c, C > 0$ tali che

$$\forall x \in \mathbb{R}^n : c\,\|x\|^* \leq \|x\|'^* \leq C\,\|x\|^*,$$

da cui

$$\forall u \in V : c\,\|\Phi(u)\|^* \leq \|\Phi(u)\|'^* \leq C\,\|\Phi(u)\|^*$$

vale a dire

$$\forall u \in V : c\,\|u\| \leq \|u\|' \leq C\,\|u\|.$$

essendo

$$\|\Phi(u)\|^* = \|u\| \ \text{ e } \ \|\Phi(u)\|'^* = \|u\|'.$$

Quindi sono equivalenti anche le norme $\|\cdot\|$ e $\|\cdot\|'$ su V.

Come conseguenza si ha che, a differenza del sopraspazio $C([a,b])$, nello spazio $\Pi_n([a,b])$ tutte le norme p ereditate da tale sopraspazio sono equivalenti.

6.4 Topologie dedotte da norme equivalenti

Il modo usuale di definire gli intorni di punti di $\mathbb{R}^n$ (visto in Analisi II) è basato sulla metrica euclidea.

Gli intorni di un punto $x \in \mathbb{R}^n$ sono i sottoinsiemi U di $\mathbb{R}^n$ tali che esiste una palla aperta $B(x, r)$ nella metrica euclidea, con $r > 0$, inclusa in U.

L'applicazione che associa ad ogni punto di $\mathbb{R}^n$ la famiglia dei suoi intorni è detta la topologia di $\mathbb{R}^n$ dedotta dalla metrica euclidea.

In generale, una *topologia* $\mathcal{T}$ su un insieme M è un'applicazione che associa ad ogni $u \in M$ una famiglia $\mathcal{T}_u$ di sottoinsiemi di M e che soddisfa certe proprietà che qui non riportiamo (gli interessati vedano il paragrafo "spazio di Hausdorff" nella la voce "topologia" su wikipedia).

I sottoinsiemi di M appartenti alla famiglia $\mathcal{T}_u$ sono detti *intorni* di u nella topologia $\mathcal{T}$. Per cui, $U \in \mathcal{T}_u$ vuol dire che U è un intorno di u nella topologia $\mathcal{T}$.

L'insieme M con la topologia $\mathcal{T}$ su M, cioè la coppia $(M, \mathcal{T})$, è detta uno *spazio topologico*.

L'introduzione di una topologia e quindi degli intorni permette di definire nozioni come: insieme aperto, insieme chiuso, successione convergente e suo limite, funzione continua, ecc.

Infatti, dato uno spazio topologico $(M, \mathcal{T})$:

- un sottoinsieme di M si dice aperto se è intorno di ogni suo punto e si dice chiuso se il suo complementare è aperto;
- una successione $\{u_n\}$ in M si dice convergente al limite $u^* \in M$ se

$$\forall U \in \mathcal{T}_{u^*} \ \exists \overline{n} \ \ \forall n \geq \overline{n} : \ u_n \in U.$$

Inoltre, dati spazi topologici $(M, \mathcal{T})$ e $(N, \mathcal{S})$, una funzione $g : D \subseteq M \to N$ si dice continua in $u \in D$ se

$$\forall U \in \mathcal{S}_{g(u)} \ \exists V \in \mathcal{T}_u : \ g\,(V \cap D) \subseteq U.$$

In Analisi II sono state date questi definizioni per $M = \mathbb{R}^n$ e $N = \mathbb{R}^m$ dotati della topologia dedotta dalla metrica euclidea.

Vediamo ora qual è il modo usuale di introdurre una topologia su uno spazio metrico. Si tratta semplicemente di generalizzare a una metrica qualsiasi la nozione di topologia dedotta dalla metrica euclidea.

Sia (M, ρ) uno spazio metrico. Consideriamo l'applicazione $\mathcal{T}$ che associa ad ogni $u \in M$ la famiglia di sottoinsiemi di M

$$\begin{aligned} \mathcal{T}_u \quad = \quad & \{U \subseteq M : \text{esiste una palla } B(u, r) \text{ nella metrica } \rho, \text{ con } r > 0, \\ & \qquad \text{inclusa in } U\}\,. \end{aligned}$$

Si prova che questa è una topologia su M, in quanto soddisfa le proprietà (qui non riportate) richieste dalla definizione di topologia. Essa è detta la *topologia su M dedotta dalla metrica ρ.*

Esercizio. Descrivere la topologia su M dedotta dalla metrica 0-1, vale a dire: quali sono gli intorni di un punto $u \in M$ in tale topologia?

Si può ora introdurre una topologia anche in uno spazio normato e quindi poi parlare di aperti, chiusi, successioni convergenti a un limite e funzioni continue: basta considerare la topologia dedotta dalla metrica dedotta dalla norma, detta topologia dedotta dalla norma.

Sia ora V uno spazio vettoriale e siano $\|\cdot\|$ e $\|\cdot\|'$ norme su V. Siano poi ρ e ρ' le metriche su V dedotte dalle norme $\|\cdot\|$ e $\|\cdot\|'$, rispettivamente. Siano infine $\mathcal{T}$ e $\mathcal{T}'$ le topologie su V dedotte dalle metriche ρ e ρ', rispettivamente.

Teorema 5 *Se $\|\cdot\|$ è equivalente a $\|\cdot\|'$, allora*

$$\mathcal{T} = \mathcal{T}', \quad \textit{cioè} \ \ \forall u \in V : \mathcal{T}_u = \mathcal{T}'_u.$$

In altri termini, se due norme su V sono equivalenti, allora gli intorni di ogni punto di V sono gli stessi nelle due topologie dedotte.

Dimostrazione. Assumiamo $\|\cdot\|$ equivalente a $\|\cdot\|'$, cioè esistono $c, C > 0$ tali che

$$\forall u \in V : c\,\|u\| \leq \|u\|' \leq C\,\|u\|\,.$$

Si ha quanto segue. Qui palla vuol dire palla aperta con raggio > 0.

(*) Per ogni $u \in V$: ogni palla di centro u nella metrica ρ include una palla di centro u nella metrica ρ'; e ogni palla di centro u nella metrica ρ' include una palla di centro u nella metrica ρ.

Infatti, dato $u \in V$ e $r > 0$, la palla $B_\rho(u, r)$ nella metrica ρ include la palla $B_{\rho'}(u, cr)$ nella metrica ρ'; e la palla $B_{\rho'}(u, r)$ nella metrica ρ' include la palla $B_\rho\left(u, \frac{r}{C}\right)$ nella metrica ρ.

Esercizio. Provare queste inclusioni.

Proviamo ora che (*) implica $\forall u \in V : \mathcal{T}_u = \mathcal{T}'_u$. Osserviamo che $\mathcal{T}_u = \mathcal{T}'_u$ (uguaglianza di insiemi) vuol dire: per ogni $U \in \mathcal{T}_u$ si ha $U \in \mathcal{T}'_u$ e per ogni $U \in \mathcal{T}'_u$ si ha $U \in \mathcal{T}_u$.

Sia $u \in V$. Proviamo che per ogni $U \in \mathcal{T}_u$ si ha $U \in \mathcal{T}'_u$. Sia $U \in \mathcal{T}_u$. U include una palla di centro u nella metrica ρ e quindi, valendo (*), anche una palla di centro u nella metrica ρ'. Quindi $U \in \mathcal{T}'_u$. Analogamente si prova che per ogni $U \in \mathcal{T}'_u$ si ha $U \in \mathcal{T}_u$. ■

Come conseguenza del precedente teorema si ha che se V ha dimensione finita, allora le topologie dedotte da norme su V sono tutte uguali essendo le norme tutte equivalenti.

Così, in uno spazio vettoriale di dimensione finita V (ad esempio $V = \mathbb{R}^n$), si hanno gli stessi insiemi aperti, gli stessi insiemi chiusi e le stesse successioni convergenti con lo stesso limite, indipendentemente da quale particolare norma si sia scelta su V.

Inoltre, dati spazi vettoriali di dimensione finita V e W (ad esempio $V = \mathbb{R}^n$ e $W = \mathbb{R}^m$), si hanno le stesse funzioni continue $D \subseteq V \to W$, indipendentemente da quali particolari norme si siano scelte su V e W.

Questo perchè tutte tali nozioni sono formulate in termini di intorni e gli intorni di ogni punto sono gli stessi, indipendentemente dalla particolare norma scelta.

Si osservi che l'equivalenza di tutte le norme implica che se una successione $\{x_n\}$ in uno spazio vettoriale di dimensione finita V converge a un limite $x \in V$ in una norma $\|\cdot\|$ su V, cioè

$$\lim_{n\to\infty} \|x_n - x\| = 0,$$

allora converge al limite x in una qualsiasi altra norma $\|\cdot\|'$ su V, cioè

$$\lim_{n\to\infty} \|x_n - x\|' = 0.$$

Si ha anche che gli ordini di convergenza a zero (gli ordini di infinitesimo) sono gli stessi: si ha

$$\|x_n - x\|' = O\left(\|x_n - x\|\right) \quad \text{e} \quad \|x_n - x\| = O\left(\|x_n - x\|'\right), \ n \to \infty.$$

Esercizio. Provare quest'ultimo fatto. Ricordare che, date successioni $\{f_n\}$ e $\{g_n\}$, la notazione $f_n = O(g_n),\ n \to \infty$, significa

$$\exists C \geq 0 \ \forall n : \ |f_n| \leq C|g_n|.$$

Se le successioni $\{f_n\}$ e $\{g_n\}$ sono infinitesime, questo vuol dire che l'ordine di infinitesimo di $\{f_n\}$ è $\geq$ dell'ordine di infinitesimo di $\{g_n\}$.

7 Prodotti scalari e spazi euclidei

Introduciamo ora la nozione di prodotto scalare su uno spazio vettoriale, la quale permette di definire la nozione di ortogonalità tra due vettori.

Definizione 6 *Sia V uno spazio vettoriale e sia $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{R}$. Si dice che $\langle\,\cdot\,\rangle$ è un* prodotto scalare *su V se*

*1) $\forall u \in V : \langle u, u\rangle \geq 0$ e $\langle u, u\rangle = 0 \Leftrightarrow u = \underline{0}$ (*positività*);*

*2) $\forall u, v \in V : \langle u, v\rangle = \langle v, u\rangle$ (*simmetria*);*

*3) $\forall \alpha \in \mathbb{R}\ \forall u, v \in V : \langle u, \alpha v\rangle = \alpha\langle u, v\rangle$ (*linearità*);*

4) $\forall u, v, w \in V : \langle u, v + w\rangle = \langle u, v\rangle + \langle u, w\rangle$ *(*linearità*).*

Lo spazio vettoriale V con il prodotto scalare $\langle\cdot,\cdot\rangle$ su V, cioè la coppia $(V, \langle\cdot,\cdot\rangle)$, è detta uno *spazio euclideo.*

Osserviamo che se $\langle\cdot,\cdot\rangle$ è un prodotto scalare su V, allora, accanto alle proprietà 1), 2), 3) e 4) della definizione, valgono anche le seguenti ulteriori proprietà 3bis), 4bis), 5) e 6), che ne sono una conseguenza:

3bis) $\forall\alpha \in \mathbb{R}\ \forall u, v \in V : \langle\alpha u, v\rangle = \alpha\langle u, v\rangle$;

Infatti, per $\alpha \in \mathbb{R}$ e $u, v \in V$, si ha

$$\begin{aligned} \langle\alpha u, v\rangle &= \langle v, \alpha u\rangle \quad \text{(simmetria)} \\ &= \alpha\langle v, u\rangle \quad \text{(linearità 3))} \\ &= \alpha\langle u, v\rangle \quad \text{(simmetria).} \end{aligned}$$

4bis) $\forall u, v, w \in V : \langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$.

Infatti, per $u, v, w \in V$ si ha

$$\begin{aligned} \langle u + v, w\rangle &= \langle w, u + v\rangle \quad \text{(simmetria)} \\ &= \langle w, u\rangle + \langle w, v\rangle \quad \text{(linearità 4))} \\ &= \langle u, w\rangle + \langle v, w\rangle \quad \text{(simmetria).} \end{aligned}$$

5) $\forall u \in V : \langle u, \underline{0}\rangle = \langle\underline{0}, u\rangle = 0.$

Infatti, per $u \in V$ si ha $\langle u, \underline{0},\rangle = \langle\underline{0}, u\rangle$ (simmetria) e

$$\begin{aligned} \langle u, \underline{0}\rangle &= \langle u, 0u\rangle = 0\langle u, u\rangle \quad \text{(linearità 3))} \\ &= 0. \end{aligned}$$

6) $\forall u, v \in V$: $\langle u + v, u + v\rangle = \langle u, u\rangle + 2\langle u, v\rangle + \langle v, v\rangle$.

Infatti, per $u, v \in V$ si ha

$$\begin{aligned} &\langle u + v, u + v\rangle = \langle u + v, u\rangle + \langle u + v, v\rangle \quad \text{(linearità 4))} \\ &= \langle u, u\rangle + \langle v, u\rangle + \langle u, v\rangle + \langle v, v\rangle \quad \text{(linearità 4bis))} \\ &= \langle u, u\rangle + \langle u, v\rangle + \langle u, v\rangle + \langle v, v\rangle \quad \text{(simmetria)} \\ &= \langle u, u\rangle + 2\langle u, v\rangle + \langle v, v\rangle. \end{aligned}$$

Esercizio. Sia $(V, \langle\,\cdot\,\rangle)$ uno spazio euclideo. Provare che per ogni $n \in \{1, 2, 3, \ldots\}$ e per ogni $u, v^{(1)}, v^{(2)}, \ldots, v^{(n)} \in V$ si ha

$$\langle u, v^{(1)} + v^{(2)} + \cdots + v^{(n)}\rangle = \langle u, v^{(1)}\rangle + \langle u, v^{(2)}\rangle + \cdots + \langle u, v^{(n)}\rangle.$$

Dimostrare per induzione su n. Osservare che questa proprietà è una generalizzazione della linearità 4), che è il caso $n = 2$. Si formuli poi e si dimostri la corrispondente generalizzazione della linearità 4bis).

Sia $(V, \langle \cdot \rangle)$ uno spazio euclideo e sia U un sottospazio di V.

Se consideriamo la restrizione

$$\langle \cdot \rangle|_U : U \times U \to \mathbb{R}$$

a U di $\langle \cdot \rangle : V \times V \to \mathbb{R}$, questa risulta essere un prodotto scalare su U, detto il *prodotto scalare ereditato* dal sopraspazio V.

La restrizione a U del prodotto scalare su V è un prodotto scalare su U in quanto le quattro proprietà nella definizione di prodotto scalare sono valide per ogni $u, v, w \in V$ e, quindi, in particolare anche per ogni $u, v, w \in U \subseteq V$.

8 Trasposizione di matrice

Ricordiamo che la *trasposta* di una matrice $A \in \mathbb{R}^{m \times n}$ è la matrice $A^T \in \mathbb{R}^{n \times m}$ con colonne sono le righe di A (equivalentemente, con righe le colonne di A):

$$A^T(i,j) = A(j,i), \ (i,j) \in \{1, \dots, n\} \times \{1, \dots, m\}.$$

Proprietà della trasposizione di matrice:

1) $\forall A \in \mathbb{R}^{m \times n} : \left(A^T\right)^T = A$;
2) $\forall A, B \in \mathbb{R}^{m \times n} : (A+B)^T = A^T + B^T$;
3) $\forall \alpha \in \mathbb{R} \ \forall A \in \mathbb{R}^{m \times n} : (\alpha A)^T = \alpha A^T$;
4) $\forall A \in \mathbb{R}^{m \times n} \ \forall B \in \mathbb{R}^{n \times p} : (AB)^T = B^T A^T$;
5) $\forall A \in \mathbb{R}^{n \times n} : A$ non singolare $\Rightarrow A^T$ non singolare e $(A^T)^{-1} = (A^{-1})^T$.

Queste proprietà sono note dal corso di Geometria. Esercizio. Provare la 5) a partire dalla 4).

9 Prodotti scalari su $\mathbb{R}^n$

Consideriamo lo spazio vettoriale $\mathbb{R}^n$ e la funzione $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ data da

$$\langle x, y \rangle = x^T y = \sum_{i=1}^{n} x_i y_i, \ x, y \in \mathbb{R}^n,$$

Proviamo che $\langle \cdot, \cdot \rangle$ è un prodotto scalare su $\mathbb{R}^n$.

Positività. Per $x \in \mathbb{R}^n$ si ha

$$\langle x, x \rangle = \sum_{i=1}^{n} \underbrace{x_i^2}_{\geq 0} \geq 0$$

e

$$\langle x,x\rangle = \sum_{i=1}^{n} \underbrace{x_i^2}_{\geq 0} = 0 \quad \Leftrightarrow \quad x_i^2 = 0 \text{ per } i \in \{1,\ldots,n\}$$
$$\Leftrightarrow \quad x_i = 0 \text{ per } i \in \{1,\ldots,n\} \Leftrightarrow \quad x = \underline{0}.$$

Simmetria. Per $x, y \in \mathbb{R}^n$ si ha

$$\langle x,y\rangle = \sum_{i=1}^{n} x_i y_i = \sum_{i=1}^{n} y_i x_i = \langle y,x\rangle$$

oppure anche

$$\begin{aligned}
\langle x,y\rangle &= x^T y = \left(x^T y\right)^T \\
&\qquad x^T y \text{ è uno scalare, cioè una matrice } 1\times 1\text{, e quindi coincide con la sua trasposta} \\
&= y^T \left(x^T\right)^T = y^T x \\
&= \langle y,x\rangle\,.
\end{aligned}$$

Linearità 3). Per $\alpha \in \mathbb{R}$ e $x, y \in \mathbb{R}^n$ si ha

$$\langle x,\alpha y\rangle = \sum_{i=1}^{n} x_i \alpha y_i = \alpha \sum_{i=1}^{n} x_i y_i = \alpha \langle x,y\rangle$$

oppure, usando una proprietà del prodotto di matrici (la $A(\alpha B) = \alpha AB$),

$$\langle x,\alpha y\rangle = x^T(\alpha y) = \alpha(x^T y) = \alpha \langle x,y\rangle\,.$$

Linearità 4). Per $x, y, z \in \mathbb{R}^n$ si ha

$$\langle x,y+z\rangle = \sum_{i=1}^{n} x_i\left(y_i+z_i\right) = \sum_{i=1}^{n}\left(x_i y_i + x_i z_i\right) = \sum_{i=1}^{n} x_i y_i + \sum_{i=1}^{n} x_i z_i = \langle x,y\rangle + \langle x,z\rangle$$

oppure, usando un'altra proprietà del prodotto di matrici (la distributività della moltiplicazione matriciale rispetto all'addizione matriciale: $A(B+C) = AB + AC$),

$$\langle x,y+z\rangle = x^T(y+z) = x^T y + x^T z = \langle x,y\rangle + \langle x,z\rangle\,.$$

Il prodotto scalare $\langle\cdot,\cdot\rangle$ è detto il *prodotto scalare canonico* su $\mathbb{R}^n$.

Un'importante proprietà del prodotto scalare canonico è la seguente, che spiega anche il ruolo che ha la matrice trasposta nell'algebra lineare.

Siano $\langle\,\cdot\,,\,\cdot\,\rangle_{\mathbb{R}^m}$ e $\langle\,\cdot\,,\,\cdot\,\rangle_{\mathbb{R}^n}$ i prodotti scalari canonici su $\mathbb{R}^n$ e $\mathbb{R}^m$. Si ha

$$\forall A \in \mathbb{R}^{m\times n}\ \forall x \in \mathbb{R}^m\ \forall y \in \mathbb{R}^n : \langle x, Ay\rangle_{\mathbb{R}^m} = \langle A^T x, y\rangle_{\mathbb{R}^n}.$$

Infatti, per $A \in \mathbb{R}^{m\times n}$, $x \in \mathbb{R}^m$ e $y \in \mathbb{R}^n$, si ha

$$\langle x, Ay\rangle_{\mathbb{R}^m} = x^T Ay = \underbrace{(A^T x)^T}_{=x^T(A^T)^T=x^TA} y = \langle A^T x, y\rangle_{\mathbb{R}^n}.$$

Naturalmente, per la simmetria del prodotto scalare si ha anche

$$\langle Ay, x\rangle_{\mathbb{R}^m} = \langle x, Ay\rangle_{\mathbb{R}^m} = \langle A^T x, y\rangle_{\mathbb{R}^n} = \langle y, A^T x\rangle_{\mathbb{R}^n}.$$

Pertanto, cambiandola con la sua trasposta, una matrice può essere spostata da un argomento all'altro di un prodotto scalare canonico mantenendo lo stesso valore del prodotto scalare.

9.1 Altri prodotti scalari su $\mathbb{R}^n$

Sia $\langle\cdot,\cdot\rangle$ un prodotto scalare su $\mathbb{R}^n$. A partire da $\langle\cdot,\cdot\rangle$, si può ottenere un nuovo prodotto scalare su $\mathbb{R}^n$ utilizzando una matrice $A \in \mathbb{R}^{n\times n}$ non singolare.

Precisamente, si consideri la funzione $\langle\cdot,\cdot\rangle_A : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ data da

$$\langle x, y\rangle_A = \langle Ax, Ay\rangle\,,\ x, y \in \mathbb{R}^n.$$

Proviamo che $\langle\cdot,\cdot\rangle_A$ è un prodotto scalare su $\mathbb{R}^n$.

Positività: per $x \in \mathbb{R}^n$, si ha

$$\langle x, x\rangle_A = \langle Ax, Ax\rangle \geq 0$$

e

$$\langle x, x\rangle_A = \langle Ax, Ax\rangle = 0 \quad \Leftrightarrow Ax = \underline{0} \Leftrightarrow x = \underline{0} \ \ (A \text{ è non singolare}).$$

Simmetria: per $x, y \in \mathbb{R}^n$ si ha

$$\langle x, y\rangle_A = \langle Ax, Ay\rangle = \langle Ay, Ax\rangle = \langle y, x\rangle_A\,.$$

Linearità 3). Per $\alpha \in \mathbb{R}$ e $x, y \in \mathbb{R}^n$ si ha

$$\langle x, \alpha y\rangle_A = \langle Ax, A(\alpha y)\rangle = \langle Ax, \alpha Ay\rangle = \alpha\,\langle Ax, Ay\rangle = \alpha\,\langle x, y\rangle_A\,.$$

Linearità 4). Per $x, y, z \in \mathbb{R}^n$ si ha

$$\begin{aligned}\langle x, y+z\rangle_A &= \langle Ax, A(y+z)\rangle = \langle Ax, Ay + Az\rangle \\ &= \langle Ax, Ay\rangle + \langle Ax, Az\rangle = \langle x, y\rangle_A + \langle x, z\rangle_A\,.\end{aligned}$$

Esercizio. Se A è singolare, allora la funzione $\langle\cdot,\cdot\rangle_A$ non è un prodotto scalare su $\mathbb{R}^n$. Quale proprietà nella definizione di prodotto scalare viene a mancare?

Esercizio. Sia $\langle\cdot,\cdot\rangle$ il prodotto scalare canonico su $\mathbb{R}^n$. Provare che

$$\forall A \in \mathbb{R}^{n\times n} \text{ non singolare } \forall x, y \in \mathbb{R}^n \ : \langle x, y\rangle_A = \langle x, A^T Ay\rangle = \langle A^T Ax, y\rangle.$$

Esercizio. Si consideri il prodotto scalare canonico $\langle\cdot,\cdot\rangle$ su $\mathbb{R}^2$ e la matrice non singolare

$$A = \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix}.$$

Trovare un'espressione per $\langle x,y\rangle_A$, dove $x, y \in \mathbb{R}^2$, cioè esprimere $\langle x,y\rangle_A$ in termini delle componenti x_1, x_2, y_1, y_2.

Considerando il prodotto scalare canonico $\langle\cdot,\cdot\rangle$ e la matrice non singolare

$$A = \operatorname{diag}\left(\sqrt{w_1}, \ldots, \sqrt{w_n}\right),$$

dove $w_i > 0$ per $i \in \{1, \ldots, n\}$, si ottiene il prodotto scalare *w-pesato* (con $w = (w_1, \ldots, w_n)$) su $\mathbb{R}^n$:

$$\begin{aligned} \langle x,y\rangle_w &:= \langle Ax, Ay\rangle \\ &= \langle(\sqrt{w_1}x_1, \ldots, \sqrt{w_n}x_n), (\sqrt{w_1}y_1, \ldots, \sqrt{w_n}y_n)\rangle \\ &= \sum_{i=1}^{n} \sqrt{w_i}x_i \cdot \sqrt{w_i}y_i \\ &= \sum_{i=1}^{n} w_i x_i y_i, \;\; x, y \in \mathbb{R}^n. \end{aligned}$$

Esercizio. Sia $\langle\cdot,\cdot\rangle$ il prodotto scalare canonico su $\mathbb{R}^n$ e sia $A \in \mathbb{R}^{n\times n}$ non singolare. Mostrare che $A^T A$ è non singolare. Poi, data una matrice $B \in \mathbb{R}^{n\times n}$, trovare una matrice $C \in \mathbb{R}^{n\times n}$ tale che

$$\forall x, y \in \mathbb{R}^n : \langle x, By\rangle_A = \langle Cx, y\rangle_A.$$

Cosa risulta essere questa matrice C nel caso del prodotto scalare w-pesato?

10 Prodotti scalari su $C([a,b])$

Consideriamo ora lo spazio vettoriale $C([a,b])$ e consideriamo la funzione $\langle\cdot,\cdot\rangle : C([a,b]) \times C([a,b]) \to \mathbb{R}$ data da

$$\langle f,g\rangle = \int_a^b f(t)\, g(t)\, dt, \;\; f, g \in C([a,b]).$$

Esercizio. Provare che $\langle\cdot,\cdot\rangle$ è un prodotto scalare.
Tale prodotto scalare è detto il *prodotto scalare canonico* su $C([a,b])$.

Il prodotto scalare *w-pesato* su $C([a,b])$ è dato da

$$\langle f,g\rangle_w := \int_a^b w(t) f(t)\, g(t)\, dt, \;\; f, g \in C([a,b]).$$

dove il peso $w : [a, b] \to \mathbb{R}$ è una funzione integrabile e non negativa che si annulla solo in un insieme di misura zero.

Esercizio. Generalizzando quanto fatto nel precedente esercizio, provare che la funzione $\langle \cdot, \cdot \rangle_w : C([a, b]) \times C([a, b]) \to \mathbb{R}$ è un prodotto scalare.

11 La disuguaglianza di Cauchy-Schwarz

Negli spazi euclidei vale la disuguaglianza di Cauchy-Schwarz.

Teorema 7 *(*Disuguaglianza di Cauchy-Schwarz*). Sia $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo. Si ha*

$$\forall u, v \in V : |\langle u, v \rangle| \leq \sqrt{\langle u, u \rangle}\ \sqrt{\langle v, v \rangle}.$$

Dimostrazione. Siano $u, v \in V$. Distinguiamo i due casi $u = \underline{0}$ e $u \neq \underline{0}$.

Se $u = \underline{0}$, allora $\langle u, u \rangle = 0$ per la positività del prodotto scalare e $\langle u, v \rangle = 0$ per la proprietà 5) del prodotto scalare. Pertanto la disuguaglianza vale con l'uguale.

Sia $u \neq \underline{0}$ e quindi $\langle u, u \rangle > 0$ per la positività. Per $t \in \mathbb{R}$, si ha

$$\langle tu + v, tu + v \rangle \geq 0 \quad \text{(positività)}$$

e

$$\begin{aligned} \langle tu + v, tu + v \rangle &= \langle tu, tu \rangle + 2\langle tu, v \rangle + \langle v, v \rangle \quad \text{(proprietà 6))} \\ &= t^2 \langle u, u \rangle + 2t \langle u, v \rangle + \langle v, v \rangle \quad \text{(linearità 3) e 3bis)).} \end{aligned}$$

Per cui, il trinomio di secondo grado

$$p(t) = \underbrace{\langle u, u \rangle}_{=a>0} t^2 + \underbrace{2\langle u, v \rangle}_{=b} t + \underbrace{\langle v, v \rangle}_{=c} = at^2 + bt + c = \langle tu + v, tu + v \rangle, \ t \in \mathbb{R},$$

ha per grafico una parabola rivolta verso l'alto che non va sotto l'asse t e quindi tale trinomio non ha due radici reali distinte. Si ha così

$$0 \geq \frac{\Delta}{4} = \left(\frac{b}{2}\right)^2 - ac = \langle u, v \rangle^2 - \langle u, u \rangle \langle v, v \rangle$$

da cui $\langle u, v \rangle^2 \leq \langle u, u \rangle \langle v, v \rangle$ e infine $|\langle u, v \rangle| \leq \sqrt{\langle u, u \rangle}\sqrt{\langle v, v \rangle}$. ■

Esercizio. Siano $u, v \in V$ con $u \neq \underline{0}$. Osservare che si ha

$$|\langle u, v \rangle| = \sqrt{\langle u, u \rangle}\sqrt{\langle v, v \rangle}$$

se e solo se

$$\frac{\Delta}{4} = \langle u, v \rangle^2 - \langle u, u \rangle \langle v, v \rangle = 0,$$

e questo si ha se e solo se il trinomio di secondo grado

$$p\left(t\right)=\left\langle u,u\right\rangle t^{2}+2\left\langle u,v\right\rangle t+\left\langle v,v\right\rangle =\left\langle tu+v,tu+v\right\rangle ,\ t\in\mathbb{R},$$

ha un'unica radice reale. Mostrare che il trinomio ha un'unica radice reale se e solo se v è un multiplo di u. Per cui si è mostrato che, per $u\neq\underline{0}$, la disuguaglianza di Cauchy-Schwarz

$$\left|\left\langle u,v\right\rangle \right|\leq\sqrt{\left\langle u,u\right\rangle }\sqrt{\left\langle v,v\right\rangle }$$

vale con l'uguale se e solo se v è un multiplo di u. Poi, come visto nella dimostrazione, essa vale con l'uguale anche nel caso $u=\underline{0}$.

Mostrare poi che se v è un multiplo non negativo di u, allora

$$\left\langle u,v\right\rangle =\sqrt{\left\langle u,u\right\rangle }\sqrt{\left\langle v,v\right\rangle }$$

e se v è un multiplo non positivo di u, allora

$$\left\langle u,v\right\rangle =-\sqrt{\left\langle u,u\right\rangle }\sqrt{\left\langle v,v\right\rangle }.$$

Nel caso dello spazio euclideo $\left(\mathbb{R}^{n},\left\langle \cdot,\cdot\right\rangle _{w}\right)$ (il prodotto scalare canonico corrisponde a $w=(1,\ldots,1)$), la disuguaglianza di Cauchy-Schwarz si legge:

$$\forall x,y\in\mathbb{R}^{n}:\left|\sum_{i=1}^{n}w_{i}x_{i}y_{i}\right|\leq\sqrt{\sum_{i=1}^{n}w_{i}x_{i}^{2}}\sqrt{\sum_{i=1}^{n}w_{i}y_{i}^{2}}.$$

Nel caso dello spazio euclideo $\left(C\left(\left[a,b\right]\right),\left\langle \cdot,\cdot\right\rangle _{w}\right)$ (il prodotto scalare canonico corrisponde a $w(t)=1$, $t\in[a,b]$), la disuguaglianza di Cauchy-Schwarz si legge:

$$\forall f,g\in C\left(\left[a,b\right]\right):\left|\int_{a}^{b}w(t)f(t)g(t)dt\right|\leq\sqrt{\int_{a}^{b}w(t)f(t)^{2}dt}\sqrt{\int_{a}^{b}w(t)g(t)^{2}dt}.$$

Esercizio. Usando la disuguaglianza di Cauchy-Schwarz provare che

$$\forall x\in\mathbb{R}^{n}:\|x\|_{1}\leq\sqrt{n}\|x\|_{2}\ \text{ e }\forall f\in C\left(\left[a,b\right]\right):\|f\|_{1}\leq\sqrt{b-a}\|f\|_{2}.$$

La seconda è la disuguaglianza già vista

$$\left\Vert f\right\Vert _{p}\leq\left(b-a\right)^{\frac{1}{p}-\frac{1}{q}}\left\Vert f\right\Vert _{q}$$

per $p=1$ e $q=2$.

Esercizio. Generalizzare le disuguaglianze del precedente esercizio alle norme 1 e 2 w-pesate.

12 Norma dedotta da un prodotto scalare

Sia $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo.

Si può introdurre su V una norma nel seguente modo. Sia $\| \cdot \| : V \to \mathbb{R}$ data da

$$\|u\| = \sqrt{\langle u, u \rangle}, \ u \in V.$$

Notare che per la positività del prodotto scalare si ha $\langle u, u \rangle \geq 0$, $u \in V$, e quindi la radice quadrata è ben definita.

Proviamo che $\| \cdot \|$ è una norma su V.

Positività: per $u \in V$ si ha

$$\|u\| = \sqrt{\langle u, u \rangle} \geq 0$$

e, per la positività del prodotto scalare,

$$\|u\| = \sqrt{\langle u, u \rangle} = 0 \quad \Leftrightarrow \ \langle u, u \rangle = 0 \quad \Leftrightarrow \quad u = \underline{0}.$$

Omogeneità: per $\alpha \in \mathbb{R}$ e $u \in V$ si ha

$$\begin{aligned} \|\alpha u\| &= \sqrt{\langle \alpha u, \alpha u \rangle} \\ &= \sqrt{\alpha^2 \langle u, u \rangle} \ \text{ (linearità 3) e 3bis) del prodotto scalare)} \\ &= |\alpha| \sqrt{\langle u, u \rangle} = |\alpha| \, \|u\| \, . \end{aligned}$$

Disuguaglianza triangolare: per $u, v \in V$ si ha,

$$\begin{aligned} \|u + v\|^2 &= \langle u + v, u + v \rangle \\ &= \langle u, u \rangle + 2 \langle u, v \rangle + \langle v, v \rangle \text{ (proprietà 6) del prodotto scalare)} \\ &= \|u\|^2 + 2 \langle u, v \rangle + \|v\|^2 \end{aligned}$$

con

$$\begin{aligned} \langle u, v \rangle &\leq |\langle u, v \rangle| \\ &\leq \sqrt{\langle u, u \rangle} \ \sqrt{\langle v, v \rangle} \ \text{ disuguaglianza di Cauchy-Schwarz} \\ &= \|u\| \ \|v\| \, . \end{aligned}$$

Si ha così

$$\begin{aligned} \|u + v\|^2 &= \|u\|^2 + 2 \langle u, v \rangle + \|v\|^2 \\ &\leq \|u\|^2 + 2 \|u\| \, \|v\| + \|v\|^2 = (\|u\| + \|v\|)^2 \end{aligned}$$

da cui si ottiene

$$\|u + v\| \leq \|u\| + \|v\| \, .$$

La norma $\| \cdot \|$ così ottenuta è detta la *norma* su V *dedotta* dal prodotto scalare $\langle \cdot, \cdot \rangle$.

Esercizio. Guardando alla dimostrazione della disuguaglianza triangolare per la norma $\| \cdot \|$ dedotta dal prodotto scalare, provare che, per $u, v \in V$ si ha

$$\|u + v\| = \|u\| + \|v\|$$

se e solo se $u = \overline{0}$ oppure v è un multiplo non negativo di u. Suggerimento: guardare all'esercizio dopo la dimostrazione della disuguaglianza di Cauchy-Schwarz. Dire inoltre quando si ha

$$\|u + v\| = |\ \|u\| - \|v\|\ |$$

e poi quando si ha

$$\|u - v\| = |\ \|u\| - \|v\|\ |.$$

Osservare che, per la proprietà 4) e 5) delle norme, risulta comunque

$$\|u - v\| \geq |\ \|u\| - \|v\|\ |$$

e

$$\|u + v\| = \|u - (-v)\| \geq |\ \|u\| - \| -v\|\ | = |\ \|u\| - \|v\|\ |.$$

Riassumendo si può dire che:

- da un prodotto scalare si deduce una norma (da uno spazio euclideo si deduce uno spazio normato);
- da una norma si deduce una metrica (da uno spazio normato si deduce uno spazio metrico);
- da una metrica si deduce una topologia (da uno spazio metrico si deduce uno spazio topologico).

La norma dedotta dal prodotto scalare $w-$pesato su $\mathbb{R}^n$ è la norma 2 $w-$pesata su $\mathbb{R}^n$ precedentemente introdotta: per $x \in \mathbb{R}^n$, si ha

$$\|x\| = \sqrt{\langle x, x \rangle_w} = \sqrt{\sum_{i=1}^{n} w_i x_i^2} = \|x\|_{2,w}\,.$$

In particolare, per $w = (1, \dots, 1)$ si ha che la norma dedotta dal prodotto scalare canonico su $\mathbb{R}^n$ è la norma 2 su $\mathbb{R}^n$.

La norma dedotta dal prodotto scalare $\langle \cdot, \cdot \rangle_A$ su $\mathbb{R}^n$, dove $\langle \cdot, \cdot \rangle$ è il prodotto scalare canonico su $\mathbb{R}^n$ e $A \in \mathbb{R}^{n \times n}$ è non singolare, è la norma $\| \cdot \|_{2,A}$ su $\mathbb{R}^n$ precedentemente introdotta: per $x \in \mathbb{R}^n$, si ha

$$\|x\| = \sqrt{\langle x, x \rangle_A} = \sqrt{\langle Ax, Ax \rangle} = \|Ax\|_2 = \|x\|_{2,A}\,.$$

Esercizio. Si consideri di nuovo la matrice non singolare

$$A = \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix}.$$

di un esercizio precedente. Trovare un'espressione per $\|x\|_{2,A}$, dove $x \in \mathbb{R}^2$, cioè esprimere $\|x\|_{2,A}$ in termini delle componenti x_1, x_2.

La norma dedotta dal prodotto scalare $w-$pesato su $C([a,b])$ è la norma 2 $w-$pesata su $C([a,b])$ precedentemente introdotta: per $f \in C([a,b])$ si ha

$$\|f\| = \sqrt{\langle f, f\rangle_w} = \sqrt{\int_a^b w(t)\, f(t)^2\, dt} = \|f\|_{2,w}\,.$$

In particolare, per $w(t) = 1$, $t \in [a,b]$, si ha che la norma dedotta dal prodotto scalare canonico su $C([a,b])$ è la norma 2 su $C([a,b])$.

Notare che usando la norma $\|\cdot\|$ dedotta dal prodotto scalare, la proprietà 6) del prodotto scalare può scriversi:

$$\forall u, v \in V : \|u+v\|^2 = \|u\|^2 + 2\langle u, v\rangle + \|v\|^2\,;$$

e la disuguaglianza di Cauchy-Schwarz può scriversi:

$$\forall u, v \in V : |\langle u, v\rangle| \le \|u\|\,\|v\|\,.$$

Per il prodotto scalare canonico su $\mathbb{R}^n$ la disuguaglianza di Cauchy-Schwarz si legge

$$\forall x, y \in \mathbb{R}^n : |x^T y| \le \|x\|_2\,\|y\|_2\,.$$

Questa disuguaglianza è un caso particolare della *disuguaglianza di Holder* (che non dimostreremo): per ogni $p, q \in [1, +\infty]$ con

$$\frac{1}{p} + \frac{1}{q} = 1$$

si ha

$$\forall x, y \in \mathbb{R}^n : |x^T y| \le \|x\|_p\,\|y\|_q\,.$$

La disuguaglianza di Cauchy-Schwarz è il caso $p = q = 2$.

Esercizio. Provare il caso $p = 1$ e $q = +\infty$ della disuguaglianza di Holder:

$$\forall x, y \in \mathbb{R}^n : |x^T y| \le \|x\|_1\,\|y\|_\infty\,.$$

La disuguaglianza di Holder vale anche in $C([a,b])$: per ogni $p, q \in [1, +\infty]$ con

$$\frac{1}{p} + \frac{1}{q} = 1$$

si ha

$$\forall f, g \in C([a,b], \mathbb{R}) : \left| \int_a^b f(t)g(t)dt \right| \leq \|f\|_p \, \|g\|_q \, .$$

Esercizio. Provare il caso $p = 1$ e $q = +\infty$ della disuguaglianza di Holder:

$$\forall f, g \in C([a,b]) : \left| \int_a^b f(t)g(t)dt \right| \leq \|f\|_1 \, \|g\|_\infty \, .$$

12.1 Identità del parallelogramma

In uno spazio euclideo $(V, \langle \cdot, \cdot \rangle)$, per la norma dedotta $\| \cdot \|$ dal prodotto scalare $\langle \cdot, \cdot \rangle$ vale la cosidetta *identità del parallelogramma*:

$$\forall u, v \in V : \|u + v\|^2 + \|u - v\|^2 = 2 \left(\|u\|^2 + \|v\|^2 \right).$$

Infatti, per $u, v \in V$ si ha, dalla proprietà 6) del prodotto scalare,

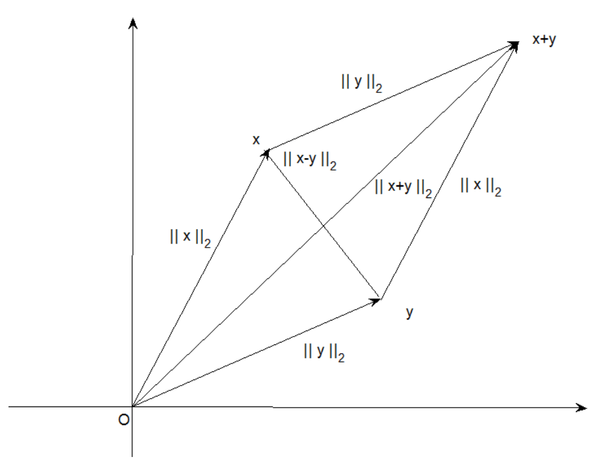
$$\|u + v\|^2 = \|u\|^2 + 2 \langle u, v \rangle + \|v\|^2$$

e

$$\begin{aligned} \|u - v\|^2 &= \|u\|^2 + 2 \langle u, -v \rangle + \|-v\|^2 \\ &= \|u\|^2 - 2 \langle u, v \rangle + \|v\|^2 \end{aligned}$$

Per cui, sommando membro a membro, si ottiene l'identità del parallelogramma.

L'identità del parallelogramma ha questo nome in quanto, nello spazio euclideo $V = \mathbb{R}^2$ o $V = \mathbb{R}^3$ con il prodotto scalare canonico, essa esprime il fatto che in un parallelogramma la somma dei quadrati delle lunghezze dei lati è uguale alla somma dei quadrati delle lunghezze delle diagonali.



La norma 2 su $\mathbb{R}^n$ è la norma dedotta dal prodotto scalare canonico su $\mathbb{R}^n$. Siamo ora in grado di provare che essa è l'unica norma p su $\mathbb{R}^n$, dove $n \geq 2$, che si deduce da un prodotto scalare su $\mathbb{R}^n$.

Infatti, per la norma p, con $p \in [1, +\infty] \setminus \{2\}$, non vale l'identità del parallelogramma: per $x = e^{(1)}$ e $y = e^{(2)}$, si ha

$$\begin{aligned}
&\|x\|_p = \|(1,0,0,\ldots,0)\|_p = 1 \\
&\|y\|_p = \|(0,1,0,\ldots,0)\|_p = 1 \\
&\|x+y\|_p = \|(1,1,0,\ldots,0)\|_p = 2^{\frac{1}{p}} \\
&\|x-y\|_p = \|(1,-1,0,\ldots,0)\|_p = 2^{\frac{1}{p}}
\end{aligned}$$

e quindi

$$\begin{aligned}
&\|x+y\|_p^2 + \|x-y\|_p^2 = 2 \cdot 2^{\frac{2}{p}} = 2^{1+\frac{2}{p}} \\
&2\left(\|x\|_p^2 + \|y\|_p^2\right) = 2 \cdot 2 = 2^2.
\end{aligned}$$

Pertanto, avendosi $1 + 2/p = 2$ se e solo se $p = 2$, si ha

$$\|x+y\|_p^2 + \|x-y\|_p^2 \neq 2\left(\|x\|_p^2 + \|y\|_p^2\right).$$

13 Ortogonalità

Sia $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo e sia $\|\cdot\|$ la norma dedotta dal prodotto scalare.

Siano $u, v \in V$. Si dice che u e v sono *ortogonali* se $\langle u, v \rangle = 0$.

La nozione di ortogonalità può essere estesa a più di due vettori come segue. I vettori $v^{(1)}, \ldots, v^{(l)} \in V$ si dicono

- *ortogonali* se $\forall i, j \in \{1, \ldots, l\}$ con $i \neq j : \left\langle v^{(i)}, v^{(j)} \right\rangle = 0$;
- *ortonormali* se sono ortogonali e $\forall i \in \{1, \ldots, l\} : \left\langle v^{(i)}, v^{(i)} \right\rangle = 1$, vale a dire $\left\| v^{(i)} \right\| = 1$.

Teorema 8 *Vettori ortogonali non nulli sono linearmente indipendenti.*

Dimostrazione. Siano $v^{(1)}, \ldots, v^{(l)} \in V$ ortogonali e non nulli. Sia

$$\sum_{j=1}^{l} \alpha_j v^{(j)} = \underline{0},$$

dove $\alpha_1, \ldots, \alpha_l \in \mathbb{R}$, una combinazione lineare nulla di $v^{(1)}, \ldots, v^{(l)}$. Proviamo che $\alpha_1 = \cdots = \alpha_l = 0$.

Per $i \in \{1, \ldots, l\}$, si ha

$$\left\langle v^{(i)}, \sum_{j=1}^{l} \alpha_j v^{(j)} \right\rangle = \left\langle v^{(i)}, \underline{0} \right\rangle = 0$$

e

$$\begin{aligned}
\left\langle v^{(i)}, \sum_{j=1}^{l} \alpha_j v^{(j)} \right\rangle &= \sum_{j=1}^{l} \left\langle v^{(i)}, \alpha_j v^{(j)} \right\rangle = \sum_{j=1}^{l} \alpha_j \left\langle v^{(i)}, v^{(j)} \right\rangle \\
&= \alpha_i \left\langle v^{(i)}, v^{(i)} \right\rangle + \sum_{\substack{j=1 \\ j \neq i}}^{n} \alpha_j \underbrace{\left\langle v^{(i)}, v^{(j)} \right\rangle}_{=0} \text{ (ortogonalità di } v^{(1)}, \ldots, v^{(l)}) \\
&= \alpha_i \left\langle v^{(i)}, v^{(i)} \right\rangle .
\end{aligned}$$

Per cui,

$$\alpha_i \left\langle v^{(i)}, v^{(i)} \right\rangle = 0.$$

Essendo $v^{(i)} \neq \underline{0}$, si ha $\left\langle v^{(i)}, v^{(i)} \right\rangle \neq 0$ per la positività del prodotto scalare. Quindi $\alpha_i = 0$. ■

Sia V di dimensione finita n. Una base $v^{(1)}, \ldots, v^{(n)}$ di V si dice una *base ortogonale* se $v^{(1)}, \ldots, v^{(n)}$ sono ortogonali e si dice una *base ortonormale* se $v^{(1)}, \ldots, v^{(n)}$ sono ortonormali.

Dal teorema precedente si ha che:

- se $v^{(1)}, \ldots, v^{(n)} \in V$ sono ortogonali non nulli, allora $v^{(1)}, \ldots, v^{(n)}$ è una base ortogonale di V;
- se $v^{(1)}, \ldots, v^{(n)} \in V$ sono ortonormali, allora $v^{(1)}, \ldots, v^{(n)}$ è una base ortonormale di V.

Nello spazio euclideo $(\mathbb{R}^n, \langle \cdot, \cdot \rangle_w)$, un esempio di base ortonormale è

$$\frac{1}{\sqrt{w_1}} e^{(1)}, \ldots, \frac{1}{\sqrt{w_n}} e^{(n)}.$$

Infatti, per $i, j \in \{1, \ldots, n\}$, si ha

$$\begin{aligned}
\left\langle \frac{1}{\sqrt{w_i}} e^{(i)}, \frac{1}{\sqrt{w_j}} e^{(j)} \right\rangle_w &= \sum_{k=1}^{n} w_k \cdot \frac{1}{\sqrt{w_i}} e_k^{(i)} \cdot \frac{1}{\sqrt{w_j}} e_k^{(j)} \\
&= \begin{cases} 0 \text{ se } i \neq j \\ w_i \cdot \frac{1}{\sqrt{w_i}} \cdot \frac{1}{\sqrt{w_i}} = 1 \text{ se } i = j. \end{cases}
\end{aligned}$$

Ricordiamo che nel corso di Geometria è stato presentato il *processo di ortogonalizzazione di Gram-Schmidt*, il quale permette di costruire una base ortogonale $w^{(1)}, \ldots, w^{(n)}$ e una base ortonormale $v^{(1)}, \ldots, v^{(n)}$ di V a partire da una qualsiasi base $u^{(1)}, \ldots, u^{(n)}$ di V.

Il processo è il seguente. I vettori $w^{(1)}, \ldots, w^{(n)} \in V$ definiti ricorsivamente a partire da $u^{(1)}, \ldots, u^{(n)}$ come segue

$$\begin{aligned} w^{(1)} &= u^{(1)} \\ w^{(k+1)} &= u^{(k+1)} - \sum_{j=1}^{k} \alpha_{k+1,j} w^{(j)} \quad \text{per } k = 1, 2, \ldots, n-1, \end{aligned}$$

dove

$$\alpha_{k+1,j} = \frac{\left\langle u^{(k+1)}, w^{(j)} \right\rangle}{\left\langle w^{(j)}, w^{(j)} \right\rangle},\ j \in \{1, \ldots, k\},$$

costituiscono una base ortogonale di V. Poi, i vettori $v^{(1)}, \ldots, v^{(n)}$ definiti da

$$v^{(i)} = \frac{1}{\left\| w^{(i)} \right\|} \cdot w^{(i)},\ i \in \{1, \ldots, n\},$$

costituiscono una base ortonormale di V.

Esercizio. Si consideri il sottospazio V di $\mathbb{R}^4$, avente per base $u^{(1)} = (1, 1, 1, 1)$, $u^{(2)} = (1, 1, 1, -1)$ e $u^{(3)} = (1, 1, -1, -1)$, con il prodotto scalare canonico ereditato da $\mathbb{R}^4$. Determinare una base ortonormale di questo sottospazio.

Esercizio. Si consideri lo spazio $\Pi_1([0, 1])$ dotato del prodotto scalare canonico ereditato dal sopraspazio $C([0, 1])$. Determinare una base ortonormale per $\Pi_1([0, 1])$.

Esercizio. Si ripeta per lo spazio $\Pi_2([0, 1])$ quanto fatto nel precedente esercizio.

14 Norme di matrice

Ricordiamo che $\mathbb{R}^{m \times n}$, l'insieme delle matrici $m \times n$ ad elementi in $\mathbb{R}$, è uno spazio vettoriale.

L'addizione vettoriale su $\mathbb{R}^{m \times n}$ è l'addizione di matrici con vettore nullo la matrice nulla, indicata con 0, e la moltiplicazione di scalari per vettori è la moltiplicazione di scalari per matrici.

Introducendo una norma sullo spazio vettoriale $\mathbb{R}^{m \times n}$ si ottiene una norma che viene applicata a matrici $m \times n$ e quindi siamo in grado di misurare la "grandezza" di tali matrici.

Tuttavia, per norma di matrice si intende qualcosa di più di una norma sullo spazio vettoriale $\mathbb{R}^{m \times n}$.

Definizione 9 *Una* norma di matrice *è una funzione*

$$\| \cdot \| : \bigcup_{\substack{m,n \text{ interi} \\ \text{positivi}}} \mathbb{R}^{m \times n} = \{\text{matrici di dimensioni qualunque}\} \to \mathbb{R}$$

tale che:

1) $\forall m, n$ *interi positivi: la restrizione di* $\|\cdot\|$ *a* $\mathbb{R}^{m\times n}$ *è una norma sullo spazio vettoriale* $\mathbb{R}^{m\times n}$*, cioè*

- $\forall A \in \mathbb{R}^{m\times n} : \|A\| \geq 0$ *e* $\|A\| = 0 \Leftrightarrow A = 0$ (positività)*;*
- $\forall \alpha \in \mathbb{R}\ \forall A \in \mathbb{R}^{m\times n} : \|\alpha A\| = |\alpha|\, \|A\|$ (omogeneità) *;*
- $\forall A, B \in \mathbb{R}^{m\times n} : \|A + B\| \leq \|A\| + \|B\|$ (disuguaglianza triangolare)*;*

2) $\forall m, n, p$ *interi positivi* $\forall A \in \mathbb{R}^{m\times n}\ \forall B \in \mathbb{R}^{n\times p} : \|AB\| \leq \|A\|\, \|B\|$ (proprietà submoltiplicativa)*.*

Si osservi che la proprietà submoltiplicativa coinvolge $\|AB\|$, $\|A\|$ e $\|B\|$ che sono valori della funzione $\|\ \cdot\ \|$ su matrici di dimensioni diverse. Questo è il motivo per cui la funzione $\|\ \cdot\ \|$ deve essere definita sull'insieme delle matrici di dimensioni qualunque.

14.1 Norma di Frobenius

Un importante esempio di norma di matrice è il seguente.

Consideriamo la seguente funzione

$$\|\cdot\|_F : \bigcup_{\substack{m,n \text{ interi} \\ \text{positivi}}} \mathbb{R}^{m\times n} \to \mathbb{R}$$

data da

$$\|A\|_F = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2},\ A \in \mathbb{R}^{m\times n}.$$

La proprietà 1) della norma di matrice è soddisfatta per $\|\cdot\|_F$.

Infatti, per m e n interi positivi, vedendo lo spazio $\mathbb{R}^{m\times n}$ come lo spazio $\mathbb{R}^{mn}$ (isomorfo a $\mathbb{R}^{m\times n}$), cioè vedendo una matrice $m \times n$ come un vettore di dimensione mn (mettendo le righe, o le colonne, una dopo l'altra), si ha che la restrizione di $\|\cdot\|_F$ a $\mathbb{R}^{m\times n}$ è la norma 2 su $\mathbb{R}^{mn}$.

Proviamo che $\|\cdot\|_F$ soddisfa anche la proprietà submoltiplicativa.

Dimostrazione. Siano $A \in \mathbb{R}^{m\times n}$ e $B \in \mathbb{R}^{n\times p}$. Scriviamo

$$A = \left[a^{(1)} \dots\ a^{(n)}\right],\ B = \begin{bmatrix} \left(b^{(1)}\right)^T \\ \vdots \\ \left(b^{(n)}\right)^T \end{bmatrix},$$

dove $a^{(1)}, \dots,\ a^{(n)} \in \mathbb{R}^m$ sono le colonne di A e $\left(b^{(1)}\right)^T, \dots, \left(b^{(n)}\right)^T \in \mathbb{R}^{1\times p}$ sono le righe di B, con $b^{(1)}, \dots,\ b^{(n)} \in \mathbb{R}^p$.

Si ha, con una moltiplicazione a blocchi,

$$AB = \sum_{k=1}^{n} a^{(k)} \left(b^{(k)}\right)^T,$$

dove, per $k \in \{1,\dots,n\}$, $a^{(k)}\left(b^{(k)}\right)^T \in \mathbb{R}^{m\times p}$ ha elementi

$$a^{(k)}\left(b^{(k)}\right)^T(i,j) = a_{ik}b_{kj},\ (i,j) \in \{1,\dots,m\}\times\{1,\dots,p\}.$$

Dalla disuguaglianza triangolare (ricordare che vale la la proprietà 1) per $\|\cdot\|_F$) si ottiene

$$\|AB\|_F = \left\|\sum_{k=1}^n a^{(k)}\left(b^{(k)}\right)^T\right\|_F \leq \sum_{k=1}^n \left\|a^{(k)}\left(b^{(k)}\right)^T\right\|_F.$$

Ora, per $k \in \{1,\dots,n\}$, si ha

$$\begin{aligned}
\left\|a^{(k)}\left(b^{(k)}\right)^T\right\|_F &= \sqrt{\sum_{i=1}^m\sum_{j=1}^p (a_{ik}b_{kj})^2} = \sqrt{\sum_{i=1}^m\sum_{j=1}^p a_{ik}^2 b_{kj}^2} \\
&= \sqrt{\sum_{i=1}^m a_{ik}^2 \cdot \sum_{j=1}^p b_{kj}^2} = \sqrt{\sum_{i=1}^m a_{ik}^2}\cdot\sqrt{\sum_{j=1}^p b_{kj}^2} \\
&= \left\|a^{(k)}\right\|_2 \left\|b^{(k)}\right\|_2.
\end{aligned}$$

Quindi

$$\begin{aligned}
\|AB\|_F &\leq \sum_{k=1}^n \left\|a^{(k)}\left(b^{(k)}\right)^T\right\|_F = \sum_{k=1}^n \left\|a^{(k)}\right\|_2 \left\|b^{(k)}\right\|_2 \\
&\leq \sqrt{\sum_{k=1}^n \left\|a^{(k)}\right\|_2^2}\cdot\sqrt{\sum_{k=1}^n \left\|b^{(k)}\right\|_2^2} \\
&\quad \text{questo segue in quanto dalla disuguaglianza di Cauchy-Schwarz risulta:} \\
&\quad \sum_{k=1}^n u_k v_k \leq \sqrt{\sum_{k=1}^n u_k^2}\cdot\sqrt{\sum_{k=1}^n v_k^2} \text{ con } u_k, v_k \geq 0,\ k \in \{1,\dots,n\}, \\
&= \sqrt{\sum_{k=1}^n\sum_{i=1}^m a_{ik}^2}\cdot\sqrt{\sum_{k=1}^n\sum_{j=1}^p b_{kj}^2} \\
&= \|A\|_F\,\|B\|_F.
\end{aligned}$$

■

La norma di matrice $\|\cdot\|_F$ è detta *norma di Frobenius*.

Esercizio. Sia $p \in [1,+\infty]$ e sia

$$\|\cdot\|_{(p)} : \bigcup_{\substack{m,n \text{ interi} \\ \text{positivi}}} \mathbb{R}^{m\times n} \to \mathbb{R}$$

data da

$$\|A\|_{(p)} = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^p \right)^{\frac{1}{p}}, \ A \in \mathbb{R}^{m \times n}.$$

1) Ripetendo e addattando la dimostrazione del fatto che $\|\cdot\|_F$ è una norma di matrice, provare che $\|\cdot\|_{(p)}$ è una norma di matrice per $p = 1$. Ad un certo punto si deve usare

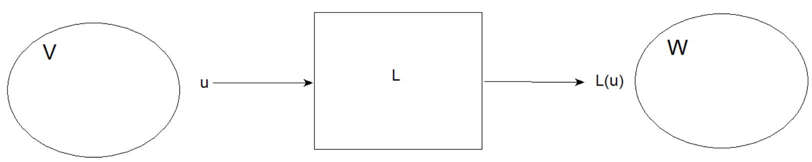
$$\sum_{k=1}^{n} u_k v_k \leq \sum_{k=1}^{n} u_k \cdot \sum_{k=1}^{n} v_k \text{ con } u_k, v_k \geq 0, \ k \in \{1, \ldots, n\}.$$

2) Considerando una matrice $A \in \mathbb{R}^{n \times n}$ con elementi tutti uguali a 1 calcolare $\|A\|_{(p)}$ e $\|A^2\|_{(p)}$ e, utilizzando questo calcolo, mostrare che $\|\cdot\|$ è una norma di matrice solo se $p \leq 2$.

15 Norme di operatore

Siano $(V, \|\cdot\|_V)$ e $(W, \|\cdot\|_W)$ spazi normati.

Sia L un *operatore lineare* da V in W, cioè una funzione lineare $V \to W$.



Al fine di definire la "grandezza" di L consideriamo i rapporti

$$\frac{\text{grandezza dell'output } L(u)}{\text{grandezza dell'input } u} = \frac{\|L(u)\|_W}{\|u\|_V}$$

al variare di $u \in V$.

Esercizio. Com'è tale rapporto quando $u = \underline{0}$?

Definiamo

$$\|L\| := \sup_{u \in V \setminus \{\underline{0}\}} \frac{\|L(u)\|_W}{\|u\|_V}$$

la "grandezza" di L relativa alle norme $\|\cdot\|_V$ e $\|\cdot\|_W$. L'operatore lineare L si dice *limitato* se $\|L\| < +\infty$.

Fissati gli interi positivi m e n, vediamo ora di introdurre una norma su $\mathbb{R}^{m \times n}$ che rispecchi il fatto che gli elementi di $\mathbb{R}^{m \times n}$ sono operatori lineari da $\mathbb{R}^n$ in $\mathbb{R}^m$.

Infatti, ogni matrice $A \in \mathbb{R}^{m\times n}$ rappresenta l'operatore lineare $L_A : \mathbb{R}^n \to \mathbb{R}^m$ dato da

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

Per cui, data una norma $\|\cdot\|_{\mathbb{R}^n}$ su $\mathbb{R}^n$, una norma $\|\cdot\|_{\mathbb{R}^m}$ su $\mathbb{R}^m$ e una matrice $A \in \mathbb{R}^{m\times n}$, definiamo

$$\|A\| := \|L_A\| = \sup_{x\in\mathbb{R}^n\setminus\{\underline{0}\}} \frac{\|Ax\|_{\mathbb{R}^m}}{\|x\|_{\mathbb{R}^n}}.$$

Teorema 10 *Siano $\|\cdot\|_{\mathbb{R}^n}$ e $\|\cdot\|_{\mathbb{R}^m}$ norme su $\mathbb{R}^n$ e $\mathbb{R}^m$, rispettivamente, e sia $A \in \mathbb{R}^{m\times n}$. L'operatore L_A è limitato, cioè $\|A\| < +\infty$, e si ha*

$$\|A\| = \max_{x\in\mathbb{R}^n\setminus\{\underline{0}\}} \frac{\|Ax\|_{\mathbb{R}^m}}{\|x\|_{\mathbb{R}^n}} = \max_{\substack{y\in\mathbb{R}^n \\ \|y\|_{\mathbb{R}^n}=1}} \|Ay\|_{\mathbb{R}^m}.$$

Dimostrazione. Osserviamo che $\|A\| = \sup \mathcal{A}$, dove

$$\mathcal{A} = \left\{ \frac{\|Ax\|_{\mathbb{R}^m}}{\|x\|_{\mathbb{R}^n}} : x \in \mathbb{R}^n \setminus \{\underline{0}\} \right\}.$$

Si ha

$$\mathcal{A} = \left\{ \frac{\|Ax\|_{\mathbb{R}^m}}{\|x\|_{\mathbb{R}^n}} : x \in \mathbb{R}^n \setminus \{\underline{0}\} \right\} = \{\|Ay\|_{\mathbb{R}^m} : y \in S\}$$

con

$$S = \{y \in \mathbb{R}^n : \|y\|_{\mathbb{R}^n} = 1\},$$

avendosi, per $x \in \mathbb{R}^n \setminus \{\underline{0}\}$,

$$\frac{\|Ax\|_{\mathbb{R}^m}}{\|x\|_{\mathbb{R}^n}} = \left\| \frac{1}{\|x\|_{\mathbb{R}^n}} Ax \right\|_{\mathbb{R}^m} = \left\| A\left(\frac{1}{\|x\|_{\mathbb{R}^n}} x \right) \right\|_{\mathbb{R}^m}$$

con

$$\left\| \frac{1}{\|x\|_{\mathbb{R}^n}} x \right\|_{\mathbb{R}^n} = \frac{\|x\|_{\mathbb{R}^n}}{\|x\|_{\mathbb{R}^n}} = 1.$$

Il teorema è provato se mostriamo che $\mathcal{A}$ ha massimo.

A tal proposito, consideriamo la funzione $f : \mathbb{R}^n \to \mathbb{R}$ data da

$$f(y) = \|Ay\|_{\mathbb{R}^m}, \ y \in \mathbb{R}^n,$$

per la quale

$$\mathcal{A} = \{f(y) : y \in S\} = f(S).$$

Se proviamo che f è continua e che S è chiuso e limitato, allora dal Teorema di Weierstrass seguirà che $\mathcal{A} = f(S)$ ha massimo.

Proviamo che f è Lipschitz continua.

Osserviamo che si ha

$$\forall z \in \mathbb{R}^n : \|Az\|_{\mathbb{R}^m} \leq \sum_{i=1}^{n} \left\|Ae^{(i)}\right\|_{\mathbb{R}^m} \cdot \|z\|_2\,, \tag{5}$$

Infatti, per $z \in \mathbb{R}^n$, si ha

$$\begin{aligned}
\|Az\|_{\mathbb{R}^m} &= \left\|A\sum_{i=1}^{n} z_i e^{(i)}\right\|_{\mathbb{R}^m} = \left\|\sum_{i=1}^{n} A\left(z_i e^{(i)}\right)\right\|_{\mathbb{R}^m} = \left\|\sum_{i=1}^{n} z_i Ae^{(i)}\right\|_{\mathbb{R}^m} \\
&\leq \sum_{i=1}^{n} \left\|z_i Ae^{(i)}\right\|_{\mathbb{R}^m} = \sum_{i=1}^{n} \underbrace{|z_i|}_{\leq \|z\|_\infty \leq \|z\|_2} \left\|Ae^{(i)}\right\|_{\mathbb{R}^m} \\
&\leq \sum_{i=1}^{n} \|z\|_2 \left\|Ae^{(i)}\right\|_{\mathbb{R}^m} = \sum_{i=1}^{n} \left\|Ae^{(i)}\right\|_{\mathbb{R}^m} \cdot \|z\|_2\,.
\end{aligned}$$

Esercizio. Da (5) e dall'equivalenza delle norme 2 e $\| \cdot \|_{\mathbb{R}^n}$ su $\mathbb{R}^n$ dedurre $\|A\| < +\infty$.

Dalla (5) segue immediatamente che f è Lipschitz continua: per $x, y \in \mathbb{R}^n$ si ha

$$\begin{aligned}
|f(x) - f(y)| &= |\|Ax\|_{\mathbb{R}^m} - \|Ay\|_{\mathbb{R}^m}| \\
&\leq \|Ax - Ay\|_{\mathbb{R}^m} \\
&= \|A(x-y)\|_{\mathbb{R}^m} \\
&\leq \sum_{i=1}^{n} \left\|Ae^{(i)}\right\|_{\mathbb{R}^m} \cdot \|x-y\|_\infty
\end{aligned}$$

dove il primo $\leq$ segue dalla proprietà 5) delle norme e il secondo dalla (5) con $z = x - y$.

Esercizio. Provare che l'insieme S è chiuso. Suggerimento: usare la caratterizzazione di un insieme chiuso tramite successioni e usare la continuità della norma $\| \cdot \|_{\mathbb{R}^n}$ (ogni norma su $\mathbb{R}^n$ è una funzione Lipschitz continua, come osservato alla fine della Sezione "Equivalenza di norme").

Esercizio. Provare che l'insieme S è limitato. Suggerimento: usare l'equivalenza delle norme 2 e $\| \cdot \|_{\mathbb{R}^n}$ su $\mathbb{R}^n$. ■

Siano $\| \cdot \|_{\mathbb{R}^n}$ e $\| \cdot \|_{\mathbb{R}^m}$ norme su $\mathbb{R}^n$ e $\mathbb{R}^m$, rispettivamente. Visto il precedente teorema si può introdurre la funzione $\| \cdot \| : \mathbb{R}^{m\times n} \to \mathbb{R}$ data da

$$\|A\| = \sup_{x \in \mathbb{R}^n \setminus \{\underline{0}\}} \frac{\|Ax\|_{\mathbb{R}^m}}{\|x\|_{\mathbb{R}^n}} = \max_{x \in \mathbb{R}^n \setminus \{\underline{0}\}} \frac{\|Ax\|_{\mathbb{R}^m}}{\|x\|_{\mathbb{R}^n}} = \max_{\substack{y \in \mathbb{R}^n \\ \|y\|_{\mathbb{R}^n} = 1}} \|Ay\|_{\mathbb{R}^m}\,,\ A \in \mathbb{R}^{m\times n}.$$

Proviamo che $\| \cdot \|$ è una norma su $\mathbb{R}^{m\times n}$.

Positività. Per $A \in \mathbb{R}^{m\times n}$, si ha

$$\|A\| = \max_{x\in\mathbb{R}^n\setminus\{\underline{0}\}} \underbrace{\frac{\|Ax\|_{\mathbb{R}^m}}{\|x\|_{\mathbb{R}^n}}}_{\geq 0} \geq 0$$

e

$$\begin{aligned}
\|A\| = \max_{x\in\mathbb{R}^n\setminus\{\underline{0}\}} \underbrace{\frac{\|Ax\|_{\mathbb{R}^m}}{\|x\|_{\mathbb{R}^n}}}_{\geq 0} = 0 \quad & \Leftrightarrow \quad \forall x \in \mathbb{R}^n \setminus \{\underline{0}\} : \|Ax\|_{\mathbb{R}^m} = 0 \\
& \Leftrightarrow \quad \forall x \in \mathbb{R}^n \setminus \{\underline{0}\} : Ax = \underline{0} \\
& \Leftrightarrow \quad A = 0 \\
(\quad & \Rightarrow \quad \text{segue da } \forall j \in \{1,\ldots,n\} : \\
& \qquad A(:,j) = Ae^{(j)} = \underline{0}).
\end{aligned}$$

Omogeneità. Per $\alpha \in \mathbb{R}$ e $A \in \mathbb{R}^{m\times n}$, si ha

$$\begin{aligned}
\|\alpha A\| &= \max_{x\in\mathbb{R}^n\setminus\{\underline{0}\}} \frac{\|(\alpha A)\,x\|_{\mathbb{R}^m}}{\|x\|_{\mathbb{R}^n}} \\
&= \max_{x\in\mathbb{R}^n\setminus\{\underline{0}\}} \frac{\|\alpha\,(Ax)\|_{\mathbb{R}^m}}{\|x\|_{\mathbb{R}^n}} \\
&= \max_{x\in\mathbb{R}^n\setminus\{\underline{0}\}} \frac{|\alpha|\,\|Ax\|_{\mathbb{R}^m}}{\|x\|_{\mathbb{R}^n}} \\
&= |\alpha| \max_{x\in\mathbb{R}^n\setminus\{\underline{0}\}} \frac{\|Ax\|_{\mathbb{R}^m}}{\|x\|_{\mathbb{R}^n}} \\
&= |\alpha|\,\|A\|\,.
\end{aligned}$$

Disuguaglianza triangolare. Per $A, B \in \mathbb{R}^{m\times n}$, si ha

$$\begin{aligned}
\|A+B\| &= \max_{x\in\mathbb{R}^n\setminus\{\underline{0}\}} \frac{\|(A+B)\,x\|_{\mathbb{R}^m}}{\|x\|_{\mathbb{R}^n}} \\
&= \max_{x\in\mathbb{R}^n\setminus\{\underline{0}\}} \frac{\|Ax+Bx\|_{\mathbb{R}^m}}{\|x\|_{\mathbb{R}^n}} \\
&\leq \max_{x\in\mathbb{R}^n\setminus\{\underline{0}\}} \frac{\|Ax\|_{\mathbb{R}^m}+\|Bx\|_{\mathbb{R}^m}}{\|x\|_{\mathbb{R}^n}} \\
&= \max_{x\in\mathbb{R}^n\setminus\{\underline{0}\}} \underbrace{\left(\frac{\|Ax\|_{\mathbb{R}^m}}{\|x\|_{\mathbb{R}^n}} + \frac{\|Bx\|_{\mathbb{R}^m}}{\|x\|_{\mathbb{R}^n}}\right)}_{\leq \max_{z\in\mathbb{R}^n\setminus\{\underline{0}\}} \frac{\|Az\|_{\mathbb{R}^m}}{\|z\|_{\mathbb{R}^n}} + \max_{z\in\mathbb{R}^n\setminus\{\underline{0}\}} \frac{\|Bz\|_{\mathbb{R}^m}}{\|z\|_{\mathbb{R}^n}} = \|A\|+\|B\|} \\
&\leq \|A\|+\|B\|\,.
\end{aligned}$$

Tale norma $\|\cdot\|$ su $\mathbb{R}^{m\times n}$ è detta la *norma di operatore* o *norma naturale* o *norma indotta* su $\mathbb{R}^{m\times n}$ relativa alle norme $\|\cdot\|_{\mathbb{R}^n}$ su $\mathbb{R}^n$ e $\|\cdot\|_{\mathbb{R}^m}$ su $\mathbb{R}^m$.

Quando $m = n$ e $\|\cdot\|_{\mathbb{R}^n} = \|\cdot\|_{\mathbb{R}^m}$, si parla di norma di operatore su $\mathbb{R}^{n\times n}$ relativa alla norma $\|\cdot\|_{\mathbb{R}^n}$ su $\mathbb{R}^n$.

Esercizio. Siano $\|\cdot\|_{\mathbb{R}^n}$ e $\|\cdot\|'_{\mathbb{R}^n}$ norme su $\mathbb{R}^n$, siano $\|\cdot\|_{\mathbb{R}^m}$ e $\|\cdot\|'_{\mathbb{R}^m}$ norme su $\mathbb{R}^m$, e siano $c, C, d, D > 0$ tali che

$$\forall x \in \mathbb{R}^n : c\|x\|_{\mathbb{R}^n} \leq \|x\|'_{\mathbb{R}^n} \leq C\|x\|_{\mathbb{R}^n}$$

e

$$\forall y \in \mathbb{R}^m : d\|y\|_{\mathbb{R}^m} \leq \|y\|'_{\mathbb{R}^m} \leq D\|y\|_{\mathbb{R}^m}.$$

Provare che

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

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

Esercizio. Calcolare $\|I_n\|$, dove $I_n \in \mathbb{R}^{n\times n}$ è la matrice identità e $\|\cdot\|$ è la norma di operatore su $\mathbb{R}^{n\times n}$ relativa alla norma $\|\cdot\|_{\mathbb{R}^n}$ su $\mathbb{R}^n$.

Esercizio. Calcolare $\|I_n\|_F$, dove $\|\cdot\|_F$ è la norma di Frobenius. Concludere che, per $n > 1$, la norma di Frobenius, quando considerata come una norma su $\mathbb{R}^{n\times n}$, non è la norma di operatore relativa a una qualche norma su $\mathbb{R}^n$.

Esercizio. Sia $\|\cdot\|$ la norma di operatore su $\mathbb{R}^{n\times n}$ relativa alla norma $\|\cdot\|_{\mathbb{R}^n}$ su $\mathbb{R}^n$ e sia $A \in \mathbb{R}^{n\times n}$. Provare che per ogni autovalore $\lambda \in \mathbb{R}$ di A si ha

$$\|A\| \geq |\lambda|.$$

Suggerimento: determinare il rapporto

$$\frac{\|Ax\|_{\mathbb{R}^n}}{\|x\|_{\mathbb{R}^n}}$$

per un autovettore $x \in \mathbb{R}^n$ relativo a λ.

Esercizio. Siano $\|\cdot\|_{\mathbb{R}^n}$ e $\|\cdot\|_{\mathbb{R}^m}$ norme su $\mathbb{R}^n$ e $\mathbb{R}^m$, rispettivamente, e sia $\|\cdot\|$ la norma di operatore su $\mathbb{R}^{m\times n}$ relativa a tali norme. Data $A \in \mathbb{R}^{m\times n}$, sia $x^* \in \mathbb{R}^n \setminus \{\underline{0}\}$ tale che

$$\|A\| = \max_{x\in\mathbb{R}^n\setminus\{\underline{0}\}} \frac{\|Ax\|_{\mathbb{R}^m}}{\|x\|_{\mathbb{R}^n}} = \frac{\|Ax^*\|_{\mathbb{R}^m}}{\|x^*\|_{\mathbb{R}^n}}.$$

Anche αx^*, dove $\alpha \in \mathbb{R} \setminus \{0\}$, soddisfa questa proprietà?

Esercizio. Si consideri l'operatore lineare integrale $L : C([a, b]) \to C([a, b])$ che associa ad ogni funzione $f \in C([a, b])$ la sua primitiva

$$L(f)(t) = \int_a^t f(s)\, ds,\ t \in [a, b].$$

Provare che L è limitato relativamente alla norma del massimo su $C([a, b])$.

Esercizio. Si consideri l'operatore lineare derivata $L : C^1([a,b]) \to C([a,b])$ che associa ad ogni funzione $f \in C^1([a,b])$ la sua derivata

$$L(f) = f'.$$

Provare che L non è limitato relativamente alla norma del massimo su $C^1([a,b])$ (ereditata dal sopraspazio $C([a,b])$) e alla norma del massimo su $C([a,b])$. Suggerimento: considerare le funzioni

$$f(t) = \left(\frac{t-a}{b-a}\right)^{\alpha}, \ t \in [a,b],$$

con $\alpha \to +\infty$.

Si osservi che la norma di operatore $\|\cdot\|$ su $\mathbb{R}^{n\times n}$ relativa alle norme $\|\cdot\|_{\mathbb{R}^n}$ su $\mathbb{R}^n$ e $\|\cdot\|_{\mathbb{R}^m}$ su $\mathbb{R}^m$ gode della seguente *proprietà di compatibilità*:

$$\forall A \in \mathbb{R}^{m\times n} \ \forall x \in \mathbb{R}^n : \|Ax\|_{\mathbb{R}^m} \leq \|A\| \, \|x\|_{\mathbb{R}^n}.$$

Tale proprietà segue immediatamente dalla definizione di norma di operatore. Infatti, sia $A \in \mathbb{R}^{m\times n}$ e $x \in \mathbb{R}^n$. Se $x = \underline{0}$, allora

$$\|Ax\|_{\mathbb{R}^m} = \|\underline{0}\|_{\mathbb{R}^m} = 0 = \|A\| \underbrace{\|x\|_{\mathbb{R}^n}}_{=\|\underline{0}\|_{\mathbb{R}^n}=0}$$

e se $x \neq \underline{0}$, allora

$$\frac{\|Ax\|_{\mathbb{R}^m}}{\|x\|_{\mathbb{R}^n}} \leq \max_{z \in \mathbb{R}^n \setminus \{\underline{0}\}} \frac{\|Az\|_{\mathbb{R}^m}}{\|z\|_{\mathbb{R}^n}} = \|A\|.$$

Le norme di operatore soddisfano la seguente *proprietà submoltiplicativa*:

Sia $\|\cdot\|_{\mathbb{R}^m}$ una norma su $\mathbb{R}^m$, sia $\|\cdot\|_{\mathbb{R}^n}$ una norma su $\mathbb{R}^n$ e sia $\|\cdot\|_{\mathbb{R}^l}$ una norma su $\mathbb{R}^l$. Si ha

$$\forall A \in \mathbb{R}^{m\times n} \ \forall B \in \mathbb{R}^{n\times l} : \|AB\| \leq \|A\| \cdot \|B\|,$$

dove:

$\|AB\|$ è la norma di operatore su $\mathbb{R}^{m\times l}$ relativa alle norme $\|\cdot\|_{\mathbb{R}^l}$ e $\|\cdot\|_{\mathbb{R}^m}$,

$\|A\|$ è la norma di operatore su $\mathbb{R}^{m\times n}$ relativa alle norme $\|\cdot\|_{\mathbb{R}^n}$ e $\|\cdot\|_{\mathbb{R}^m}$,

$\|B\|$ è la norma di operatore su $\mathbb{R}^{n\times l}$ relativa alle norme $\|\cdot\|_{\mathbb{R}^l}$ e $\|\cdot\|_{\mathbb{R}^n}$.

Infatti, per $A \in \mathbb{R}^{m\times n}$ e $B \in \mathbb{R}^{n\times l}$, si ha

$$\begin{aligned}
\|AB\| &= \max_{x\in\mathbb{R}^l\setminus\{\underline{0}\}} \frac{\|(AB)\,x\|_{\mathbb{R}^m}}{\|x\|_{\mathbb{R}^l}} \\
&= \max_{x\in\mathbb{R}^l\setminus\{\underline{0}\}} \frac{\|A\,(Bx)\|_{\mathbb{R}^m}}{\|x\|_{\mathbb{R}^l}} \\
&\leq \max_{x\in\mathbb{R}^l\setminus\{\underline{0}\}} \frac{\|A\|\,\|Bx\|_{\mathbb{R}^n}}{\|x\|_{\mathbb{R}^l}} \text{ (dalla proprietà di compatibilità)} \\
&= \|A\| \max_{x\in\mathbb{R}^l\setminus\{\underline{0}\}} \frac{\|Bx\|_{\mathbb{R}^n}}{\|x\|_{\mathbb{R}^l}} \\
&= \|A\|\,\|B\|\,.
\end{aligned}$$

Osservare che la proprietà submoltiplicativa sopra presentata è diversa dalla proprietà submoltiplicativa della definizione di norma di matrice.

Nella proprietà submoltiplicativa sopra $\|AB\|$, $\|A\|$ e $\|B\|$ sono tre valori di tre funzioni diverse definite su $\mathbb{R}^{m\times l}$, $\mathbb{R}^{m\times n}$ e $\mathbb{R}^{n\times l}$, rispettivamente. Invece, nella proprietà submoltiplicativa della definizione di norma di matrice $\|AB\|$, $\|A\|$ e $\|B\|$ sono tre valori di una stessa funzione definita sull'insieme delle matrici di dimensione qualunque.

Esercizio. Sia $\|\cdot\|$ la norma di operatore su $\mathbb{R}^{n\times n}$ relativa alla norma $\|\cdot\|_{\mathbb{R}^n}$ su $\mathbb{R}^n$ e sia $A \in \mathbb{R}^{n\times n}$ non singolare. Provare che

$$\|A^{-1}\| \geq \frac{1}{\|A\|}.$$

Suggerimento: usare la proprietà submoltiplicativa e il fatto che $A^{-1}A = \ldots$

16 Norme p di matrice

Sia $p \in [1, +\infty]$. Consideriamo la funzione

$$\|\cdot\|_p : \bigcup_{\substack{m,n \text{ interi} \\ \text{positivi}}} \mathbb{R}^{m\times n} \to \mathbb{R}$$

data da

$$\begin{aligned}
\|A\,\|_p &= \text{norma di operatore di } A \text{ relativa alle norme } p \text{ su } \mathbb{R}^n \text{ e } \mathbb{R}^m \\
&= \max_{x\in\mathbb{R}^n\setminus\{\underline{0}\}} \frac{\|Ax\|_p}{\|x\|_p},\ A \in \mathbb{R}^{m\times n}.
\end{aligned}$$

Proviamo che $\|\cdot\|_p$ è una norma di matrice.

La proprietà 1) delle norme di matrice è verificata: per ogni m, n interi positivi, la restrizione di $\|\cdot\|_p$ a $\mathbb{R}^{m\times n}$ è una norma di operatore e quindi una norma su $\mathbb{R}^{m\times n}$.

Anche la proprietà submoltiplicativa delle norme di matrice è verificata: segue dalla precedente proprietà submoltiplicativa delle norme di operatore prendendo la norma p su $\mathbb{R}^m$, $\mathbb{R}^n$ e $\mathbb{R}^l$.

La norma di matrice $\|\cdot\|_p$ è detta *norma p di matrice.*

Esercizio. Fissati gli interi positivi m e n, trovare costanti $c, C > 0$ tali che

$$\forall A \in \mathbb{R}^{m\times n} : c\|A\|_p \leq \|A\|_q \leq C\|A\|_p$$

per $(p, q) = (1, 2), (2, 1), (1, +\infty), (+\infty, 1), (2, +\infty), (+\infty, 2)$. Suggerimento: usare un esercizio precedente.

In generale, non è facile calcolare la norma p di una matrice $A \in \mathbb{R}^{m\times n}$ utilizzando la sua definizione

$$\|A\|_p = \max_{x\in\mathbb{R}^n\setminus\{\underline{0}\}} \frac{\|Ax\|_p}{\|x\|_p}.$$

Per $p = 1$ o $p = +\infty$ esistono però delle formule che permettono di calcolarla in modo immediato.

Teorema 11 *Per ogni $A \in \mathbb{R}^{m\times n}$ si ha*

$$\|A\|_1 = \max_{j\in\{1,\ldots,n\}} \sum_{i=1}^{m} |a_{ij}|$$

Inoltre si ha

$$\|A\|_1 = \frac{\|Ax^*\|_1}{\|x^*\|_1}$$

per $x^* = e^{(j^*)}$, dove $j^* \in \{1, \ldots, n\}$ è tale che

$$\sum_{i=1}^{m} |a_{ij^*}| = \max_{j\in\{1,\ldots,n\}} \sum_{i=1}^{m} |a_{ij}| .$$

Così $\|A\|_1$ si ottiene sommando in ogni colonna i moduli degli elementi e prendendo il massimo di tutte queste somme.

Esempio: per

$$A = \begin{bmatrix} 1 & -2 & 5 \\ 3 & 0 & 2 \\ -1 & 1 & -4 \end{bmatrix}$$

si ha

$$\|A\|_1 = \max\{1 + 3 + 1, 2 + 0 + 1, 5 + 2 + 4\} = \max\{5, 3, 11\} = 11.$$

Inoltre $j^* = 3$ e quindi per $x^* = e^{(3)}$ si ha

$$\|A\|_1 = \frac{\|Ax^*\|_1}{\|x^*\|_1}.$$

Dimostrazione. Sia $A \in \mathbb{R}^{m\times n}$. Posto

$$L := \max_{j\in\{1,\dots,n\}} \sum_{i=1}^{m} |a_{ij}|$$

mostriamo che valgono i seguenti due fatti:

A) $\forall x \in \mathbb{R}^n : \|Ax\|_1 \le L\,\|x\|_1$.

B) per x^* come nell'enunciato si ha: $\|Ax^*\|_1 = L\,\|x^*\|_1$.

Da A) e B) seguirà

$$\|A\|_1 = \max_{x\in\mathbb{R}^n\setminus\{\underline{0}\}} \frac{\|Ax\|_1}{\|x\|_1} = \frac{\|Ax^*\|_1}{\|x^*\|_1} = L.$$

Dimostrazione di A). Sia $x \in \mathbb{R}^n$. Si ha

$$\begin{aligned}
\|Ax\|_1 &= \sum_{i=1}^{m} |(Ax)_i| = \sum_{i=1}^{m} \underbrace{\left|\sum_{j=1}^{n} a_{ij}x_j\right|}_{\le \sum_{j=1}^{n}|a_{ij}x_j| = \sum_{j=1}^{n}|a_{ij}|\cdot|x_j|} \\
&\le \sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|\cdot|x_j| = \sum_{j=1}^{n}\sum_{i=1}^{m} |a_{ij}|\cdot|x_j| \\
&= \sum_{j=1}^{n} \underbrace{\sum_{i=1}^{m} |a_{ij}|}_{\le \max_{k\in\{1,\dots,n\}} \sum_{i=1}^{m}|a_{ik}| = L} \cdot |x_j| \\
&\le \sum_{j=1}^{n} L\,|x_j| = L\sum_{j=1}^{n} |x_j| = L\,\|x\|_1\,.
\end{aligned}$$

Dimostrazione di B). Per x^* come nell'enunciato si ha

$$\|x^*\|_1 = \|e^{(j^*)}\|_1 = 1$$

e

$$\begin{aligned}
\|Ax^*\|_1 &= \|Ae^{(j^*)}\|_1 = \|A(:,j^*)\|_1 \\
&= \sum_{i=1}^{m} |a_{ij^*}| = L = L\|x^*\|_1.
\end{aligned}$$

■

Teorema 12 *Per ogni $A \in \mathbb{R}^{m \times n}$ si ha*

$$\|A\|_\infty = \max_{i \in \{1,\dots,m\}} \sum_{j=1}^{n} |a_{ij}| .$$

Inoltre, si ha

$$\|A\|_\infty = \frac{\|Ax^*\|_\infty}{\|x^*\|_\infty}$$

per $x^* \in \mathbb{R}^n$ tale che

$$x_j^* = \begin{cases} \frac{a_{i^*j}}{|a_{i^*j}|} = \mathrm{sign}\,(a_{i^*j}) \ \textit{se}\ a_{i^*j} \neq 0 \\ 1 \ \textit{se}\ a_{i^*j} = 0 \end{cases}, \ j \in \{1, \dots, n\},$$

dove $i^* \in \{1, \dots, m\}$ è tale che

$$\sum_{j=1}^{n} |a_{i^*j}| = \max_{i \in \{1,\dots,m\}} \sum_{j=1}^{n} |a_{ij}| .$$

Così $\|A\|_\infty$ si ottiene sommando in ogni riga i moduli degli elementi e prendendo il massimo di tutte queste somme.

Esempio: di nuovo per

$$A = \begin{bmatrix} 1 & -2 & 5 \\ 3 & 0 & 2 \\ -1 & 1 & -4 \end{bmatrix}$$

si ha

$$\|A\|_\infty = \max\{1+2+5, 3+0+2, 1+1+4\} = \max\{8,5,6\} = 8.$$

Inoltre $i^* = 1$ e quindi per $x^* \in \mathbb{R}^3$ tale che

$$x_j^* = \begin{cases} \frac{a_{i^*j}}{|a_{i^*j}|} = \mathrm{sign}\,(a_{i^*j}) \ \text{se}\ a_{i^*j} \neq 0 \\ 1 \ \text{se}\ a_{i^*j} = 0 \end{cases} = \begin{cases} 1 \ \text{se}\ j = 1 \\ -1 \ \text{se}\ j = 2 \\ 1 \ \text{se}\ j = 3 \end{cases}, \ j \in \{1,2,3\},$$

si ha

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

Dimostrazione. Sia $A \in \mathbb{R}^{m \times n}$. Posto

$$M := \max_{i \in \{1,\dots,m\}} \sum_{j=1}^{n} |a_{ij}|$$

mostriamo che valgono i seguenti due fatti:

A) $\forall x \in \mathbb{R}^n : \|Ax\|_\infty \leq M \|x\|_\infty$;

B) per x^* come nell'enunciato si ha: $\|Ax^*\|_\infty = M\,\|x^*\|_\infty$.

Da A) e B) seguirà

$$\|A\|_\infty = \max_{x\in\mathbb{R}^n\setminus\{\underline{0}\}} \frac{\|Ax\|_\infty}{\|x\|_\infty} = \frac{\|Ax^*\|_\infty}{\|x^*\|_\infty} = M.$$

Dimostrazione di A). Sia $x \in \mathbb{R}^n$. Si ha

$$\begin{aligned}
\|Ax\|_\infty &= \max_{i\in\{1,\ldots,m\}} |(Ax)_i| = \max_{i\in\{1,\ldots,m\}} \underbrace{\left|\sum_{j=1}^n a_{ij}x_j\right|}_{\leq \sum_{j=1}^n |a_{ij}x_j| = \sum_{j=1}^n |a_{ij}||x_j|} \\
&\leq \max_{i\in\{1,\ldots,m\}} \sum_{j=1}^n |a_{ij}| \underbrace{|x_j|}_{\leq\|x\|_\infty} \leq \max_{i\in\{1,\ldots,m\}} \sum_{j=1}^n |a_{ij}|\,\|x\|_\infty \\
&= \max_{i\in\{1,\ldots,m\}} \left(\left(\sum_{j=1}^n |a_{ij}|\right)\|x\|_\infty\right) = \underbrace{\left(\max_{i\in\{1,\ldots,m\}} \sum_{j=1}^n |a_{ij}|\right)}_{=M} \|x\|_\infty\,.
\end{aligned}$$

Dimostrazione di B). Per x^* come nell'enunciato si ha

$$\|x^*\|_\infty \;=\; \max_{j\in\{1,\ldots,n\}} |x_j^*| = 1,$$

in quanto tutte le componenti di x^* sono 1 o -1, e

$$\begin{aligned}
\|Ax^*\|_\infty &= \max_{i\in\{1,\ldots,m\}} |(Ax^*)_i| \geq |(Ax^*)_{i^*}| = \left|\sum_{j=1}^n a_{i^*j}x_j^*\right| \\
&= \left|\sum_{\substack{j=1\\a_{i^*j}\neq 0}}^n a_{i^*j}x_j^*\right| = \left|\sum_{\substack{j=1\\a_{i^*j}\neq 0}}^n a_{i^*j}\mathrm{sign}\,(a_{i^*j})\right| = \left|\sum_{\substack{j=1\\a_{i^*j}\neq 0}}^n |a_{i^*j}|\right| \\
&= \left|\sum_{j=1}^n |a_{i^*j}|\right| = \sum_{j=1}^n |a_{i^*j}| = M = M\|x^*\|_\infty.
\end{aligned}$$

Quindi $\|Ax^*\|_\infty \geq M\|x^*\|_\infty$. Avendosi anche

$$\|Ax^*\|_\infty \leq M\|x^*\|_\infty$$

dal fatto A), si conclude che $\|Ax^*\|_\infty = M\|x^*\|_\infty$. ■

Esercizio. Nel precedente teorema, è fondamentale che il vettore $x^* \in \mathbb{R}^n \setminus \{\underline{0}\}$ tale che $\|A\|_\infty = \frac{\|Ax^*\|_\infty}{\|x^*\|_\infty}$ abbia componenti x_j^* uguali a 1 per $a_{i^*j} = 0$? Tali componenti possono anche essere diverse da 1?

Più avanti daremo una formula anche per $\|A\|_2$, però più complicata rispetto alle due precedenti, nel senso che $\|A\|_2$ non è immediatamente calcolabile come sopra dagli elementi di A.

Una delle ragioni per considerare norme su $\mathbb{R}^n$ diverse dalla norma euclidea, come la norma 1 e la norma ∞, è che, per quest'ultime, le corrispondenti norme di matrice si trovano immediatamente con le formule date sopra. Si osservi che anche la norma di Frobenius è immediatamente calcolabile dagli elementi della matrice.

Esercizio. Sia $x \in \mathbb{R}^n$. Vedendo x come vettore colonna, cioè come matrice $n \times 1$, determinare le norme di matrice 1 e ∞ di x con le precedenti formule. Determinare poi le norme di matrice 1 e ∞ del vettore riga x^T, che è una matrice $1 \times n$.

Esercizio. Determinare la norma p di matrice, $p \in [1, +\infty]$, del vettore colonna $x \in \mathbb{R}^{n\times 1}$ usando la definizione di norma p come norma di operatore.

Esercizio. Usando la definizione di norma p come norma di operatore, provare che la norma p di matrice del vettore riga $x^T \in \mathbb{R}^{1\times n}$ è minore o uguale a $\|x\|_q$, dove $q \in [1, +\infty]$ è tale che $\frac{1}{p} + \frac{1}{q} = 1$. Suggerimento: usare la disuguaglianza di Holder.

Esercizio. Data $A \in \mathbb{R}^{m\times n}$, quale relazione intercorre tra $\|A\|_1$ e $\|A^T\|_\infty$?

Esercizio. Data

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

disegnare nel piano $\mathbb{R}^2$ l'insieme

$$\left\{(a, b) \in \mathbb{R}^2 : \|A\|_1 \geq \|A\|_\infty\right\}.$$

Esercizio. Sia $A \in \mathbb{R}^{n\times n}$. Si consideri una successione di vettori $x^{(0)}, x^{(1)}, x^{(2)}, \ldots$ data ricorsivamente da

$$x^{(k+1)} = Ax^{(k)}, k = 0, 1, 2, \ldots$$

Sia $\|\cdot\|$ la norma di operatore su $\mathbb{R}^{n\times n}$ relativa alla norma $\|\cdot\|_{\mathbb{R}^n}$ su $\mathbb{R}^n$. Sotto quale condizione su $\|A\|$ si ha

$$\|x^{(0)}\|_{\mathbb{R}^n} \geq \|x^{(1)}\|_{\mathbb{R}^n} \geq \|x^{(2)}\|_{\mathbb{R}^n} \geq \cdots ?$$

Sotto quale altra condizione su $\|A\|$ si ha

$$\lim_{k\to\infty} \|x^{(k)}\|_{\mathbb{R}^n} = 0?$$

Suggerimento: usare la proprietà di compatibilità.

Esercizio. Consideriamo un paese dove non si ha nè immigrazione nè emigrazione. Si vuol studiare come evolve nel tempo la popolazione di questo paese.

Dividiamo la popolazione in quattro fasce d'età: la prima da 0 a 20 anni, la seconda da 21 a 40, la terza da 41 a 60 e la quarta da 61 a 80. Non consideriamo gli individui con più di 80 anni.

Per $k \in \{0, 1, 2, \ldots\}$ e $i \in \{1, 2, 3, 4\}$, sia $x_i^{(k)}$ il numero di individui nella fascia $i-$esima al tempo $20k$ e sia

$$x^{(k)} = \left(x_1^{(k)}, x_2^{(k)}, x_3^{(k)}, x_4^{(k)} \right) \in \mathbb{R}^4.$$

Per $k \in \{0, 1, 2, \ldots\}$, risulta

$$\begin{cases} x_1^{(k+1)} = a_1 x_1^{(k)} + a_2 x_2^{(k)} + a_3 x_3^{(k)} + a_4 x_4^{(k)} \\ x_2^{(k+1)} = (1 - b_1)\, x_1^{(k)} \\ x_3^{(k+1)} = (1 - b_2)\, x_2^{(k)} \\ x_4^{(k+1)} = (1 - b_3)\, x_3^{(k)} \end{cases}$$

dove

- $a_i \geq 0$, $i \in \{1, 2, 3, 4\}$, è il tasso di natalità della fascia $i-$esima in 20 anni (a_i è il numero di figli generati in 20 anni da individui nella fascia $i-$esima rapportato al totale degli individui nella fascia)
- $b_i \in [0, 1]$, $i \in \{1, 2, 3\}$,è il tasso di mortalità della fascia $i-$esima in 20 anni (b_i è il numero degli individui nella fascia $i-$esima che muoiono in 20 anni rapportato al totale degli individui nella fascia).

Si ha così

$$x^{(k+1)} = Ax^{(k)}, k = 0, 1, 2, \ldots$$

con

$$A = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ 1 - b_1 & 0 & 0 & 0 \\ 0 & 1 - b_2 & 0 & 0 \\ 0 & 0 & 1 - b_3 & 0 \end{bmatrix} \in \mathbb{R}^{4\times 4}.$$

Mostrare che se

$$a_1 < b_1, \ a_2 < b_2, \ a_3 < b_3 \ \text{ e } \ a_4 < 1$$

allora la popolazione si estinguerà. Suggerimento: considerare la norma 1 nell'esercizio precedente.

Mostrare che la popolazione si estinguerà anche quando

$$a_1 + a_2 + a_3 + a_4 < 1 \ \text{ e } \ \ b_1, b_2, b_3, b_4 > 0.$$

Suggerimento: considerare la norma ∞ nell'esercizio precedente.