

Appunti di Programmazione Matematica

Canale 1 · Prof. Marco Sciandrone

Capitolo 2: Da $\mathbb{R}$ a $\mathbb{R}^n$

Autore: Tommaso C.

Appunti tratti dalle lezioni in aula e dal materiale del corso

Argomenti: Ordine totale ed estremi · Condizioni necessarie e sufficienti · Interni
Retta reale estesa · Spazio $\mathbb{R}^n$ · Funzioni di più variabili, dominio e grafico
Prodotto scalare · Norme · Distanza · Cauchy-Schwarz · Angolo e ortogonalità

Indice

1	Richiami su $\mathbb{R}$	1
1.1	$\mathbb{R}$ come insieme totalmente ordinato	1
1.2	Estremo superiore e inferiore	1
1.3	Condizioni necessarie e sufficienti	2
1.4	Intorni e punti di accumulazione in $\mathbb{R}$	3
2	La retta reale estesa	4
3	Lo spazio $\mathbb{R}^n$	5
3.1	Prodotto cartesiano e $\mathbb{R}^n$	5
3.2	L'ordine in $\mathbb{R}^n$	5
4	Funzioni di più variabili	7
5	Operazioni in $\mathbb{R}^n$	8
6	Il prodotto scalare	10
7	Norme	10
7.1	Definizione di norma	10
7.2	Esempi di norme	11
7.3	La “norma zero” e la sparsità	11
7.4	Distanza	12
7.5	Intorni in $\mathbb{R}^n$	13
8	Norma e prodotto scalare	13
8.1	Angolo fra due vettori e ortogonalità	14
9	Esercizi	15
10	Soluzioni	17

1 Richiami su $\mathbb{R}$

1.1 $\mathbb{R}$ come insieme totalmente ordinato

Definizione – Relazione d'ordine totale

Una relazione $\leq$ su un insieme X è una **relazione d'ordine** se è

- (i) **riflessiva**: $a \leq a$ per ogni a ;
- (ii) **antisimmetrica**: se $a \leq b$ e $b \leq a$, allora $a = b$;
- (iii) **transitiva**: se $a \leq b$ e $b \leq c$, allora $a \leq c$.

È una relazione d'**ordine totale** se è anche

- (iv) **totale**: per ogni $a \neq b$ vale $a \leq b$ oppure $b \leq a$, cioè due elementi qualsiasi sono confrontabili.

Con l'usuale $\leq$, $\mathbb{R}$ è **totalmente ordinato**: fissate sulla retta un'origine e un'unità di misura, $a \leq b$ se il punto di a sta a sinistra di quello di b o coincide. Su questa proprietà poggiano estremo superiore e inferiore, e poi minimo e massimo di una funzione di una variabile.

1.2 Estremo superiore e inferiore

Definizione – Maggioranti e minoranti

Sia $A \subseteq \mathbb{R}$. Un numero $\lambda \in \mathbb{R}$ è un **minorante** di A se

$$\lambda \leq x \quad \forall x \in A;$$

un numero $\Omega \in \mathbb{R}$ è un **maggiorante** di A se $x \leq \Omega$ per ogni $x \in A$. Se esiste almeno un minorante, A è **limitato inferiormente**; se esiste almeno un maggiorante, è **limitato superiormente**; se valgono entrambe, è **limitato**.

Definizione – Estremo inferiore e superiore

Sia $A \subseteq \mathbb{R}$. Il numero λ è l'**estremo inferiore** di A , $\lambda = \inf A$, se

- (1) λ è un minorante di A ;
- (2) $\lambda \geq L$ per ogni minorante L di A .

È cioè il *più grande dei minoranti*. Allo stesso modo $\Omega = \sup A$, l'**estremo superiore**, è il *più piccolo dei maggioranti*: Ω è un maggiorante e $\Omega \leq U$ per ogni maggiorante U di A .

Teorema – Esistenza degli estremi

Sia $A \subseteq \mathbb{R}$ non vuoto. Esiste $\sup A \in \mathbb{R}$ se e solo se A è limitato superiormente; esiste $\inf A \in \mathbb{R}$ se e solo se A è limitato inferiormente. In particolare esistono entrambi se e solo se A è limitato.

Se A non è limitato superiormente si pone $\sup A = +\infty$; se non è limitato inferiormente, $\inf A = -\infty$.

Esempio – Estremi che appartengono o no all'insieme

Per $A = \{-2, 3, 7, 1, 1000\}$ si ha $\inf A = -2$ e $\sup A = 1000$, entrambi in A . Per $A = \{x \in \mathbb{R} : x > 3\}$ si ha $\inf A = 3 \notin A$, mentre A non è limitato superiormente e $\sup A = +\infty$.

1.3 Condizioni necessarie e sufficienti**Definizione – Condizione necessaria, sufficiente, necessaria e sufficiente**

Siano P una proprietà e C una condizione.

- C è **necessaria** per P se $P \Rightarrow C$: se vale P , vale sicuramente C (“ P solo se C ”).
- C è **sufficiente** per P se $C \Rightarrow P$: basta C per avere P (“ P se C ”).
- C è **necessaria e sufficiente** per P se $P \Leftrightarrow C$ (“ P se e solo se C ”).

Dimostrare una condizione necessaria e sufficiente significa dimostrare due implicazioni, e bisogna saper dire quale parte della dimostrazione riguarda ciascuna. Nel teorema di esistenza degli estremi il verso “solo se” (se esiste $\sup A$ finito, allora A è limitato superiormente) è la condizione necessaria; il verso “se” è quella sufficiente.

Esempio – Tre condizioni a parole

- Avere almeno 18 anni è **necessario** per avere la patente, ma non sufficiente: si può avere più di 18 anni senza aver superato l'esame, o con la patente ritirata.
- Vincere tutte le partite è **sufficiente** per vincere il campionato, ma non necessario: si può vincerlo anche perdendo qualche partita.
- Il diploma di maturità è **necessario e sufficiente** per potersi iscrivere all'università.

1.4 Interni e punti di accumulazione in $\mathbb{R}$

Definizione – Interno sferico e intorno

Dati $x_0 \in \mathbb{R}$ e $\rho > 0$, l'**intorno sferico** (aperto) di centro x_0 e raggio ρ è

$$\mathcal{B}_\rho(x_0) = \{x \in \mathbb{R} : |x - x_0| < \rho\} = (x_0 - \rho, x_0 + \rho).$$

Un **intorno** $I(x_0)$ è un qualsiasi intervallo che contiene un intorno sferico di x_0 : per esempio, ogni intervallo aperto (a, b) che contiene x_0 .

Per esempio $(2, 8)$ è un intorno di $x_0 = 3$, perché contiene $\mathcal{B}_1(3) = (2, 4)$.

Definizione – Punto di accumulazione

Sia $A \subseteq \mathbb{R}$. Un punto $x_0 \in \mathbb{R}$ è un **punto di accumulazione** di A se ogni intorno sferico di x_0 contiene almeno un punto di A diverso da x_0 :

$$\forall \rho > 0 \quad \exists x \in A, x \neq x_0 : x \in \mathcal{B}_\rho(x_0).$$

L'insieme dei punti di accumulazione di A si chiama **derivato** di A e si indica con $\mathcal{D}(A)$.

Definizione – Punto isolato

Un punto x_0 è un **punto isolato** di A se $x_0 \in A$ e x_0 non è punto di accumulazione di A , cioè $x_0 \in A \setminus \mathcal{D}(A)$.

Un punto di accumulazione appartiene a $\mathbb{R}$ ma *non necessariamente* ad A ; un punto isolato appartiene sempre ad A .

Esempio – Accumulazione in un intervallo semiaperto

Sia $A = \{x \in \mathbb{R} : 0 < x \leq 1\} = (0, 1]$. Allora

$$\mathcal{D}(A) = \{x \in \mathbb{R} : 0 \leq x \leq 1\} = [0, 1].$$

Sono di accumulazione tutti i punti di A e anche 0, che non appartiene ad A : ogni intorno $(-\rho, \rho)$ contiene punti di $(0, 1]$. Un punto come $x_0 = 1,5$ non è di accumulazione: l'intorno $\mathcal{B}_{0,25}(1,5)$ non contiene punti di A . Se si aggiunge ad A il punto 3, cioè si considera $(0, 1] \cup \{3\}$, il punto 3 è isolato.



$0 \notin A$ è di accumulazione: ogni intorno contiene punti di A . Il punto 3 è isolato.

2 La retta reale estesa

Definizione – Retta reale estesa

La **retta reale estesa** è l'insieme

$$\mathbb{R}^* = \mathbb{R} \cup \{-\infty\} \cup \{+\infty\},$$

dove i due elementi aggiunti sono tali che

$$-\infty < x < +\infty \quad \forall x \in \mathbb{R}.$$

Queste relazioni hanno senso grazie all'ordine totale di $\mathbb{R}$, che si estende a $\mathbb{R}^*$.

Definizione – Intorni di $-\infty$ e di $+\infty$

Dato un qualsiasi numero $a \in \mathbb{R}$, un **intorno di $-\infty$** è l'insieme

$$I(-\infty) = \{x \in \mathbb{R} : x < a\},$$

e un **intorno di $+\infty$** è l'insieme

$$I(+\infty) = \{x \in \mathbb{R} : x > a\}.$$

La definizione di punto di accumulazione poggia solo sul concetto di intorno: ora $-\infty$ o $+\infty$ possono essere punti di accumulazione di $A \subseteq \mathbb{R}$, se ogni loro intorno contiene punti di A . È la condizione che serve per studiare il limite di una funzione per $x \rightarrow +\infty$ o $x \rightarrow -\infty$: $\pm\infty$ deve essere punto di accumulazione del dominio.

Esempio – Accumulazione all'infinito

Sia $A = \{x \in \mathbb{R} : x^2 \geq 1\} = \{x \leq -1\} \cup \{x \geq 1\}$. Ogni intorno $\{x < a\}$ di $-\infty$ contiene punti di A (tutti quelli minori di $\min\{a, -1\}$), e lo stesso vale per $+\infty$: $-\infty$ e $+\infty$ sono punti di accumulazione di A . Invece nessuno dei due è di accumulazione per $[0, 1]$.

3 Lo spazio $\mathbb{R}^n$

3.1 Prodotto cartesiano e $\mathbb{R}^n$

Definizione – Prodotto cartesiano

Dati due insiemi A e B , il loro **prodotto cartesiano** è l'insieme delle *coppie ordinate*

$$A \times B = \{(a, b) : a \in A, b \in B\}.$$

Più in generale, per $A_1, A_2, \dots, A_n$,

$$A_1 \times A_2 \times \dots \times A_n = \{(a_1, a_2, \dots, a_n) : a_1 \in A_1, a_2 \in A_2, \dots, a_n \in A_n\}.$$

“Ordinate” significa che conta quale elemento sta al primo posto e quale al secondo: in generale $A \times B \neq B \times A$, a meno che $A = B$. In quel caso si scrive $A \times A = A^2$, e analogamente A^n per il prodotto di n copie di A .

Definizione – Spazio $\mathbb{R}^n$

$\mathbb{R}^n$ è il prodotto cartesiano di n copie di $\mathbb{R}$, cioè l'insieme delle n -uple ordinate di numeri reali:

$$\mathbb{R}^n = \underbrace{\mathbb{R} \times \mathbb{R} \times \dots \times \mathbb{R}}_{n \text{ volte}} = \{(x_1, x_2, \dots, x_n) : x_1 \in \mathbb{R}, x_2 \in \mathbb{R}, \dots, x_n \in \mathbb{R}\}.$$

Per $n = 2$ è il piano $\{(x, y)\}$, per $n = 3$ lo spazio $\{(x, y, z)\}$.

Nota – Convenzione: vettori colonna

Un elemento $x \in \mathbb{R}^n$ è sempre un **vettore colonna**:

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Con due o tre variabili si usano anche i nomi x, y, z ; da quattro in poi $x_1, x_2, x_3, \dots$

3.2 L'ordine in $\mathbb{R}^n$

Definizione – Ordine componente per componente

Dati $x, y \in \mathbb{R}^n$,

$$x \leq y \iff x_i \leq y_i \quad i = 1, \dots, n.$$

La relazione è riflessiva, antisimmetrica e transitiva, perché lo è $\leq$ su ogni componente: $\mathbb{R}^n$ è **ordinato**. Per $n > 1$ non è però **totale**.

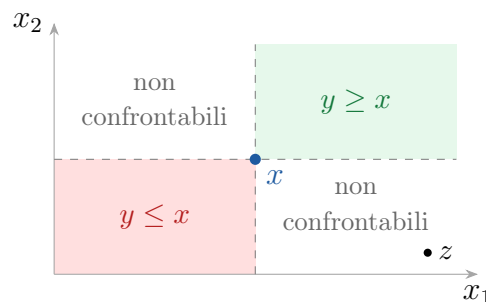
Esempio – Due confronti in $\mathbb{R}^5$

$$\begin{pmatrix} 2 \\ -1 \\ 3 \\ 4 \\ 1 \end{pmatrix} \leq \begin{pmatrix} 2 \\ 0 \\ 4 \\ 4 \\ 2 \end{pmatrix} \quad \text{perché } 2 \leq 2, -1 \leq 0, 3 \leq 4, 4 \leq 4, 1 \leq 2.$$

Se nel secondo vettore la prima componente diventa -1 , i vettori

$$\begin{pmatrix} 2 \\ -1 \\ 3 \\ 4 \\ 1 \end{pmatrix} \quad \text{e} \quad \begin{pmatrix} -1 \\ 0 \\ 4 \\ 4 \\ 2 \end{pmatrix}$$

non sono più confrontabili: la prima componente del primo è maggiore, le altre sono minori o uguali. Nessuno dei due è $\leq$ dell'altro.



In $\mathbb{R}^2$ i vettori confrontabili con x stanno in due soli quadranti; un punto come z non lo è.

Osservazione – La prima differenza fra $\mathbb{R}$ e $\mathbb{R}^n$

Per $n > 1$ in $\mathbb{R}^n$ non si riesce a definire una relazione d'ordine che sia anche totale: $\mathbb{R}^n$ è ordinato ma **non totalmente ordinato**. Estremo superiore e inferiore, massimo e minimo di un *insieme di vettori* non si definiscono come in $\mathbb{R}$: cercare l'estremo superiore di un insieme di $\mathbb{R}^2$ è un errore.

4 Funzioni di più variabili

Definizione – Funzioni a valori scalari e vettoriali

Una funzione di n variabili riceve un vettore $x \in \mathbb{R}^n$. Si distingue per lo spazio di arrivo:

- **a valori scalari:** $f : \mathbb{R}^n \rightarrow \mathbb{R}$, a ogni x associa un numero $f(x)$;
- **a valori vettoriali:** $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ con $m > 1$, a ogni x associa un vettore $f(x) \in \mathbb{R}^m$.

Esempio – Due funzioni

- $f : \mathbb{R}^3 \rightarrow \mathbb{R}$, $f(x, y, z) = x^2 + yz - 3z$: a valori scalari.
- $h : \mathbb{R} \rightarrow \mathbb{R}^2$, $h(x) = \begin{pmatrix} x^2 \\ e^x \end{pmatrix}$: anche con una sola variabile ($n = 1$) la funzione può essere a valori vettoriali. Conta lo spazio di arrivo.

Definizione – Dominio

Il **dominio** di una funzione f di n variabili è l'insieme $D \subseteq \mathbb{R}^n$ dei punti in cui f è definita: $f(x)$ esiste per ogni $x \in D$ e non esiste per $x \notin D$. Si scrive

$$f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m.$$

È la definizione di una variabile, con $\mathbb{R}^n$ al posto di $\mathbb{R}$; il dominio può essere tutto $\mathbb{R}^n$. Per esempio $f(x, y) = \log(x + y - 1)$ è definita dove l'argomento del logaritmo è positivo:

$$D = \{(x, y) \in \mathbb{R}^2 : x + y - 1 > 0\}.$$

È uno dei due semipiani in cui la retta $x + y - 1 = 0$ divide il piano. Per sapere quale, basta provare un punto fuori dalla retta: $(0, 0)$ dà $-1 > 0$, che è falso, quindi D è il semipiano che non contiene l'origine. La retta non appartiene a D .

Definizione – Grafico

Il **grafico** di una funzione a valori scalari $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ è l'insieme

$$G_f = \{(x, y) \in \mathbb{R}^{n+1} : x \in D, y = f(x)\}.$$

Il grafico sta in uno spazio con una dimensione in più di quello del dominio. Per $n = 1$ è una curva di $\mathbb{R}^2$, per $n = 2$ una superficie di $\mathbb{R}^3$; da $n = 3$ in poi non si disegna, ma si definisce allo stesso modo. Dominio e grafico sono insiemi, f è una funzione: aperto e chiuso si dicono del dominio, non della funzione.

Osservazione – Dove ha senso parlare di minimo e massimo

Una funzione a valori scalari *atterra* in $\mathbb{R}$, che è totalmente ordinato: anche con n variabili ha senso parlare di estremo superiore e inferiore dei suoi valori, e quindi di punti di minimo e di massimo. Per una funzione a valori vettoriali i valori stanno in $\mathbb{R}^m$, $m > 1$, che non è totalmente ordinato: non si parla di minimo o massimo, ma di **ottimo di Pareto**.

L'Esercizio 2 costruisce una funzione di tre variabili a valori scalari da minimizzare.

5 Operazioni in $\mathbb{R}^n$

Definizione – Somma e prodotto per uno scalare

Dati $x, y \in \mathbb{R}^n$ e $\alpha \in \mathbb{R}$,

$$x + y = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix} \in \mathbb{R}^n, \quad \alpha x = \begin{pmatrix} \alpha x_1 \\ \alpha x_2 \\ \vdots \\ \alpha x_n \end{pmatrix} \in \mathbb{R}^n.$$

Entrambe si eseguono componente per componente; fra vettori di dimensione diversa la somma non è definita. Il risultato resta in $\mathbb{R}^n$: le due operazioni sono **chiuse**, e $\mathbb{R}^n$ con somma e prodotto per uno scalare è uno **spazio vettoriale**.

Definizione – Tre prodotti fra vettori

Dati $x, y \in \mathbb{R}^n$:

- il **prodotto di Hadamard** (componente per componente) è il vettore

$$x \circ y = \begin{pmatrix} x_1 y_1 \\ x_2 y_2 \\ \vdots \\ x_n y_n \end{pmatrix} \in \mathbb{R}^n;$$

- il **prodotto scalare** è il numero

$$x^\top y = (x_1 \ x_2 \ \dots \ x_n) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = x_1 y_1 + x_2 y_2 + \dots + x_n y_n = \sum_{i=1}^n x_i y_i \in \mathbb{R};$$

- il **prodotto diadico** (o esterno) è la matrice $x y^\top \in \mathbb{R}^{n \times n}$, di elementi $(x y^\top)_{ij} = x_i y_j$.

I tre prodotti atterrano in spazi diversi. Nel seguito serve quasi sempre il prodotto scalare.

Esempio – I tre prodotti a confronto

Con

$$x = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad y = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} :$$

$$x \circ y = \begin{pmatrix} 1 \\ 0 \\ -3 \end{pmatrix}, \quad x^\top y = 1 + 0 - 3 = -2, \quad x y^\top = \begin{pmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 3 & 0 & -3 \end{pmatrix}.$$

6 Il prodotto scalare

Definizione – Prodotto scalare

Un **prodotto scalare** (o prodotto interno) su $\mathbb{R}^n$ è una funzione

$$\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$$

tale che, per ogni $x, y, z \in \mathbb{R}^n$ e $\alpha \in \mathbb{R}$:

- (1) $\langle x, x \rangle \geq 0$;
- (2) $\langle x, x \rangle = 0$ se e solo se $x = 0$;
- (3) $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$;
- (4) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$;
- (5) $\langle x, y \rangle = \langle y, x \rangle$.

La definizione nasce da $x^\top y$, che ha tutte e cinque le proprietà. Per esempio $x^\top x = x_1^2 + \dots + x_n^2 \geq 0$, e una somma di quadrati è nulla solo se ogni termine è nullo, cioè $x = 0$. Le altre tre seguono dalle regole del prodotto fra numeri.

$\langle x, y \rangle = x^\top y$ è il prodotto scalare a cui si farà sempre riferimento, ma non è l'unico: ogni funzione con le proprietà (1)–(5) è un prodotto scalare.

7 Norme

7.1 Definizione di norma

Definizione – Norma

Una **norma** su $\mathbb{R}^n$ è una funzione $\|\cdot\| : \mathbb{R}^n \rightarrow \mathbb{R}_+$, con $\mathbb{R}_+ = \{t \in \mathbb{R} : t \geq 0\}$, tale che, per ogni $x, y \in \mathbb{R}^n$ e $\alpha \in \mathbb{R}$:

- (1) $\|x\| \geq 0$;
- (2) $\|x\| = 0$ se e solo se $x = 0$;
- (3) $\|\alpha x\| = |\alpha| \|x\|$ (omogeneità);
- (4) $\|x + y\| \leq \|x\| + \|y\|$ (disuguaglianza triangolare).

La norma generalizza il valore assoluto, che misura quanto un numero dista dall'origine e ha le stesse quattro proprietà. A differenza del prodotto scalare, che parte da $\mathbb{R}^n \times \mathbb{R}^n$ e può essere negativo, la norma parte da $\mathbb{R}^n$ e non è mai negativa.

7.2 Esempi di norme

Nel piano, per il teorema di Pitagora, (x, y) dista dall'origine $\sqrt{x^2 + y^2}$. In $\mathbb{R}^n$ la stessa idea dà la **norma euclidea**, o norma L_2 ; per $n = 1$ si riduce a $\sqrt{x^2} = |x|$. Due altre norme molto usate nella pratica sono L_1 e L_∞ :

$$\|x\|_2 = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} = \sqrt{\sum_{i=1}^n x_i^2} \quad (\text{norma euclidea}),$$

$$\|x\|_1 = |x_1| + |x_2| + \cdots + |x_n| = \sum_{i=1}^n |x_i| \quad (\text{norma } L_1),$$

$$\|x\|_\infty = \max_{i=1, \dots, n} |x_i| \quad (\text{norma } L_\infty).$$

Per dire che una funzione è una norma bisogna verificare le quattro proprietà: per $\|\cdot\|_1$ e $\|\cdot\|_\infty$ lo fa l'Esercizio 1. Si userà quasi sempre la norma euclidea, ma ciò che si dimostra per una norma qualsiasi vale per tutte.

Esempio – Tre norme dello stesso vettore

Per

$$x = \begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix} :$$

$$\|x\|_2 = \sqrt{1 + 4 + 4} = 3, \quad \|x\|_1 = 1 + 2 + 2 = 5, \quad \|x\|_\infty = \max\{1, 2, 2\} = 2.$$

7.3 La “norma zero” e la sparsità

Definizione – Norma zero

$\|x\|_0$ è il numero di componenti $x_i \neq 0$ del vettore x . Va da $\mathbb{R}^n$ in $\mathbb{R}_+$, ma **non è una norma**: viola l'omogeneità.

Esempio – Un vettore di $\mathbb{R}^7$

Per

$$x = \begin{pmatrix} 0 \\ 3 \\ 0 \\ -1 \\ 2 \\ 0 \\ 5 \end{pmatrix} \in \mathbb{R}^7$$

le componenti non nulle sono quattro: $\|x\|_0 = 4$. Anche $2x$ ha quattro componenti non nulle, quindi $\|2x\|_0 = 4 \neq 2\|x\|_0 = 8$.

Moltiplicare per uno scalare non nullo non cambia il numero di componenti non nulle. Pur non essendo una norma, $\|\cdot\|_0$ misura la **sparsità**: per governare un sistema con il minor numero possibile di elementi si cerca una soluzione con $\|x\|_0$ piccola. Gli Esercizi 3 e 4 chiedono di riconoscere quale proprietà manca a funzioni che sembrano norme.

7.4 Distanza

Definizione – Distanza indotta da una norma

Data una norma $\|\cdot\|$ su $\mathbb{R}^n$, la **distanza** fra $x, y \in \mathbb{R}^n$ indotta dalla norma è

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

In $\mathbb{R}$ la distanza fra due numeri a e b è $|a - b|$; in $\mathbb{R}^n$ al valore assoluto si sostituisce una norma. La distanza dipende dalla norma scelta (L_1 , L_2 , L_∞); si userà sempre quella indotta dalla norma euclidea, $d(x, y) = \|x - y\|_2$.

Definizione – Distanza (metrica)

Una **distanza** (o metrica) su $\mathbb{R}^n$ è una funzione $d : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}_+$ tale che, per ogni $x, y, z \in \mathbb{R}^n$:

- (1) $d(x, y) \geq 0$;
- (2) $d(x, y) = 0$ se e solo se $x = y$;
- (3) $d(x, y) = d(y, x)$ (simmetria);
- (4) $d(x, y) \leq d(x, z) + d(z, y)$ (disuguaglianza triangolare).

La definizione nasce dalla distanza indotta da una norma, che ha le quattro proprietà: (1) e (2) seguono dalle proprietà (1) e (2) della norma; (3) dall'omogeneità con $\alpha = -1$,

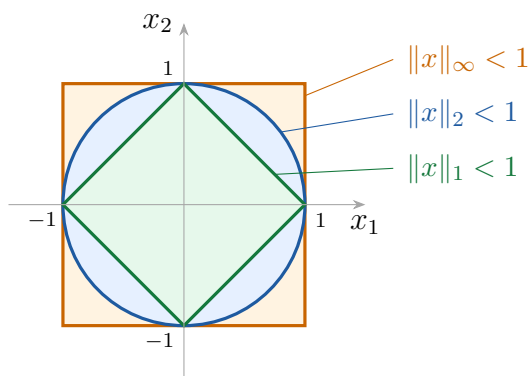
perché $\|x - y\| = \|-(y - x)\| = \|y - x\|$; (4) dalla disuguaglianza triangolare applicata a $x - y = (x - z) + (z - y)$.

7.5 Intorni in $\mathbb{R}^n$

Intorno sferico, intorno e punto di accumulazione poggiano sulla distanza $|x - x_0|$. In $\mathbb{R}^n$ si usa la distanza indotta da una norma: l'intorno sferico di centro x_0 e raggio $\rho > 0$ è l'insieme dei punti che distano da x_0 meno di ρ ,

$$\mathcal{B}_\rho(x_0) = \{x \in \mathbb{R}^n : \|x - x_0\| < \rho\}.$$

Fra $\mathbb{R}$ e $\mathbb{R}^n$ la differenza, nella scrittura, è una stanghetta in più: da $|\cdot|$ a $\|\cdot\|$. La forma dell'intorno dipende dalla norma; per $x_0 = 0$ e $\rho = 1$, in $\mathbb{R}^2$:



$\mathcal{B}_1(0)$ in $\mathbb{R}^2$: rombo con L_1 , cerchio con L_2 , quadrato con L_∞ .

8 Norma e prodotto scalare

Proposizione – Prodotto scalare e norma euclidea

Per ogni $x \in \mathbb{R}^n$

$$x^\top x = x_1^2 + x_2^2 + \cdots + x_n^2 = \|x\|_2^2.$$

Teorema – Disuguaglianza di Cauchy-Schwarz

Per ogni $x, y \in \mathbb{R}^n$

$$|x^\top y| \leq \|x\|_2 \|y\|_2,$$

e vale l'uguaglianza se e solo se i due vettori sono paralleli, cioè uno è multiplo dell'altro.

8.1 Angolo fra due vettori e ortogonalità

Definizione – Angolo fra due vettori

Dati $x, y \in \mathbb{R}^n$ non nulli, l'**angolo** fra x e y è il numero $\theta \in [0, \pi]$ tale che

$$\cos \theta = \frac{x^\top y}{\|x\|_2 \|y\|_2}.$$

La definizione ha senso per Cauchy-Schwarz: dividendo per $\|x\|_2 \|y\|_2$ e togliendo il valore assoluto,

$$-1 \leq \frac{x^\top y}{\|x\|_2 \|y\|_2} \leq 1,$$

quindi il rapporto è il coseno di un angolo.

Definizione – Vettori ortogonali

Due vettori non nulli $x, y \in \mathbb{R}^n$ sono **ortogonali** se il coseno del loro angolo è nullo ($\theta = \pi/2$), cioè se e solo se

$$x^\top y = 0.$$

Un vettore non nullo non è mai ortogonale a sé stesso, perché $x^\top x = \|x\|_2^2 > 0$; l'Esercizio 5 lo mette alla prova su un vettore di $\mathbb{R}^4$.

Esempio – Angoli in $\mathbb{R}^2$ e in $\mathbb{R}^3$

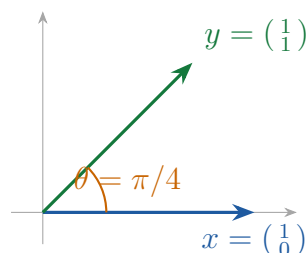
Per

$$x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad y = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

si ha $x^\top y = 1$, $\|x\|_2 = 1$, $\|y\|_2 = \sqrt{2}$, quindi $\cos \theta = 1/\sqrt{2}$ e $\theta = \pi/4$ (figura). Per

$$x = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \quad y = \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix}$$

si ha $x^\top y = 2 - 4 + 2 = 0$: i due vettori di $\mathbb{R}^3$ sono ortogonali.



Riepilogo

- $\mathbb{R}$ è totalmente ordinato; per $A \neq \emptyset$, $\sup A$ ($\inf A$) è finito se e solo se A è limitato superiormente (inferiormente).
- C necessaria per P : $P \Rightarrow C$ (“solo se”); sufficiente: $C \Rightarrow P$ (“se”).
- $\mathcal{B}_\rho(x_0) = \{x - x_0 \mid x - x_0 < \rho\}$; x_0 di accumulazione se ogni intorno contiene punti di A diversi da x_0 ; isolato se $x_0 \in A \setminus \mathcal{D}(A)$.
- $\mathbb{R}^* = \mathbb{R} \cup \{\pm\infty\}$, con intorni $\{x < a\}$ e $\{x > a\}$.
- $\mathbb{R}^n$ ($n > 1$) non è totalmente ordinato: minimo e massimo solo per $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$, con dominio D e grafico $G_f \subseteq \mathbb{R}^{n+1}$.
- Prodotto scalare e norma sono funzioni con le proprietà (1)–(5) e (1)–(4); $\|\cdot\|_1$, $\|\cdot\|_2$, $\|\cdot\|_\infty$ sono norme, $\|\cdot\|_0$ no; $d(x, y) = \|x - y\|$ è una distanza.
- $x^\top x = \|x\|_2^2$ e $|x^\top y| \leq \|x\|_2 \|y\|_2$ (Cauchy-Schwarz).
- $\cos \theta = x^\top y / (\|x\|_2 \|y\|_2)$; ortogonali se $x^\top y = 0$.

9 Esercizi

Esercizio 1 – Norma 1 e norma infinito

Verificare che $\|\cdot\|_1$ e $\|\cdot\|_\infty$ sono norme.

Soluzione a pagina 17 →

Esercizio 2 – La migliore parabola

Sono dati quattro punti del piano $(x_1, y_1), \dots, (x_4, y_4)$ per cui non passa nessuna parabola. Tradurre in un problema di ottimizzazione la richiesta di trovare *la parabola* $y = ax^2 + bx + c$ che meglio li approssima.

Soluzione a pagina 17 →

Esercizio 3 – Appelli del 23/02, 11/07 e 16/09/2024, Canale M-Z, Prof. Giampaolo Liuzzi, domanda 3

Sia $\sigma(x) = x_1 + x_2 + \dots + x_n$. Dire se ciascuna affermazione è vera o falsa.

- La funzione σ definisce una norma in $\mathbb{R}^n$.
- $\sigma(x)$ è sempre positiva.
- $\sigma(x) = 0$ se e solamente se $x = 0$.

(d) $\sigma(x)$ non soddisfa le proprietà delle norme in $\mathbb{R}^n$.

Soluzione a pagina 17 →

Esercizio 4 – Appello del 29/01/2024, Canale M-Z, Prof. Giampaolo Liuzzi, domanda 7

Sia $\sigma(x_1, x_2) = \max\{0, x_1 + x_2\}$. Dire se ciascuna affermazione è vera o falsa.

- (a) $\sigma(-1, 0) < 0$.
- (b) $\sigma(0, -1) < 0$.
- (c) σ definisce una valida norma in $\mathbb{R}^2$.
- (d) σ non definisce una valida norma in $\mathbb{R}^2$.

Soluzione a pagina 18 →

Esercizio 5 – Appelli del 23/02, 11/07 e 16/09/2024, Canale M-Z, Prof. Giampaolo Liuzzi, domanda 6

Sia

$$v = \begin{pmatrix} 0 \\ -2 \\ 1 \\ 3 \end{pmatrix}$$

un vettore in $\mathbb{R}^4$. Dire se ciascuna affermazione è vera o falsa.

- (a) v è ortogonale a sé stesso.
- (b) Ogni vettore $x \in \mathbb{R}^4$ è ortogonale a v .
- (c) Non esiste un vettore $w \in \mathbb{R}^4$ ortogonale a v .
- (d) L'angolo tra v e il vettore $e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ è di 90° .

Soluzione a pagina 18 →

10 Soluzioni

Soluzione dell'Esercizio 1 – Norma 1 e norma infinito

Le proprietà (1)–(3) seguono da quelle del valore assoluto: una somma o un massimo di valori assoluti è ≥ 0 ed è nullo solo se ogni componente è nulla; inoltre $|\alpha x_i| = |\alpha| |x_i|$ per ogni i . Per la disuguaglianza triangolare si parte da $|x_i + y_i| \leq |x_i| + |y_i|$. Sommando su i si ottiene $\|x + y\|_1 \leq \|x\|_1 + \|y\|_1$. Maggiorando ogni termine con il massimo, $|x_i + y_i| \leq \|x\|_\infty + \|y\|_\infty$ per ogni i , quindi anche il massimo dei primi membri: $\|x + y\|_\infty \leq \|x\|_\infty + \|y\|_\infty$.

[← Torna all'Esercizio 1, pagina 15](#)

Soluzione dell'Esercizio 2 – La migliore parabola

Le variabili di decisione non sono le coordinate dei punti, che sono dati, ma i tre coefficienti a, b, c che individuano la parabola. Per ogni punto si misura l'errore commesso, una funzione di a, b, c :

$$e_i(a, b, c) = (ax_i^2 + bx_i + c - y_i)^2, \quad i = 1, \dots, 4.$$

Si scelgono a, b, c minimizzando la somma degli errori:

$$\min_{(a,b,c) \in \mathbb{R}^3} f(a, b, c) = \sum_{i=1}^4 (ax_i^2 + bx_i + c - y_i)^2.$$

È il minimo, senza vincoli, di una funzione $\mathbb{R}^3 \rightarrow \mathbb{R}$ a valori scalari: il metodo dei **minimi quadrati**.

Il quadrato serve: sommando le differenze con segno, errori positivi e negativi si compenserebbero, e diminuendo c la somma diventerebbe arbitrariamente negativa, quindi il problema non avrebbe minimo. Anche il valore assoluto renderebbe ogni errore non negativo, ma darebbe un modello diverso.

[← Torna all'Esercizio 2, pagina 15](#)

Soluzione dell'Esercizio 3 – Appelli del 23/02, 11/07 e 16/09/2024, Canale M-Z, Prof. Giampaolo Liuzzi, domanda 3

(b) Falsa: per il vettore con prima componente -1 e le altre nulle si ha $\sigma(x) = -1 < 0$, quindi manca anche la proprietà (1). (c) Falsa: con $n \geq 2$, il vettore con componenti 1 e -1 nelle prime due posizioni e 0 altrove è non nullo, ma $\sigma(x) = 1 - 1 = 0$; manca la proprietà (2). Manca anche l'omogeneità, perché $\sigma(-x) = -\sigma(x)$

mentre dovrebbe valere $|-1| \sigma(x) = \sigma(x)$; la disuguaglianza triangolare invece vale, con l'uguaglianza $\sigma(x + y) = \sigma(x) + \sigma(y)$. Quindi (a) è falsa e (d) è vera.

[← Torna all'Esercizio 3, pagina 15](#)

Soluzione dell'Esercizio 4 – Appello del 29/01/2024, Canale M-Z, Prof. Giampaolo Liuzzi, domanda 7

(a) e (b) sono false: $\sigma(-1, 0) = \max\{0, -1\} = 0$ e $\sigma(0, -1) = \max\{0, -1\} = 0$; in generale il massimo con 0 rende σ sempre non negativa, quindi la proprietà (1) vale. Il calcolo mostra però due vettori non nulli con $\sigma = 0$, quindi manca la (2). Manca anche l'omogeneità: per $x = (1, 0)$ si ha $\sigma(x) = 1$ ma $\sigma(-x) = 0 \neq |-1| \sigma(x)$. Quindi (c) è falsa e (d) è vera.

[← Torna all'Esercizio 4, pagina 16](#)

Soluzione dell'Esercizio 5 – Appelli del 23/02, 11/07 e 16/09/2024, Canale M-Z, Prof. Giampaolo Liuzzi, domanda 6

(a) Falsa: $v^T v = 0 + 4 + 1 + 9 = 14 \neq 0$. (b) Falsa: basta $x = v$, per il punto (a). (c) Falsa: per esempio $w = e_1$ è non nullo e $e_1^T v = 0$; ce ne sono infiniti, tutti i vettori non nulli con $-2x_2 + x_3 + 3x_4 = 0$. (d) Vera: poiché $e_1^T v = 0$ si ha $\cos \theta = 0$, quindi $\theta = \pi/2$, cioè 90° .

[← Torna all'Esercizio 5, pagina 16](#)