5. Richiami: $\mathbb{R}^n$ come spazio euclideo

5.a Prodotto scalare

Dati due vettori in $\mathbb{R}^n$, si misurano le rispettive lunghezze e l'angolo (senza il segno) fra essi mediante il *prodotto scalare*.

Il *prodotto scalare* di due vettori $x = (x^1, x^2, \dots, x^n)$ e $y = (y^1, y^2, \dots, y^n) \in \mathbb{R}^n$ è definito da

$$x \bullet y := \sum_{i=1}^{n} x^i y^i.$$

$(x, y) \to x \bullet y$ è dunque una mappa $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ con le seguenti proprietà

(i) (BILINEARITÀ) $x \bullet y$ è lineare su ciascun fattore, i.e.,

$$\begin{aligned}(\alpha x_1 + \beta x_2) \bullet y &= \alpha\, x_1 \bullet y + \beta\, x_2 \bullet y\,,\\ x \bullet (\alpha y_1 + \beta y_2) &= \alpha\, x \bullet y_1 + \beta\, x \bullet y_2\,,\end{aligned}$$

per ogni $x_1, x_2, y_1, y_2 \in \mathbb{R}^n$ $\forall \alpha, \beta \in \mathbb{R}$.

(ii) (SIMMETRIA) $x \bullet y = y \bullet x$ $\forall x, y \in \mathbb{R}^n$,

(iii) (POSITIVITÀ) $\forall x = (x^1, x^2, \dots, x^n) \in \mathbb{R}^n$ $x \bullet x = \sum_{i=1}^n (x^i)^2 \geq 0$ e $x \bullet x = 0$ se e solo se $x = 0$.

Il nunero reale positivo

$$|x| := \sqrt{x \bullet x}$$

si chiama *norma* o *lunghezza* del vettore $x \in \mathbb{R}^n$.

5.1 Definizione. *Si dice che due vettori $x, y \in \mathbb{R}^n$ sono perpendicolari o ortogonali e si scrive $x \perp y$, se $x \bullet y = 0$. Un insieme di vettori $\{a_1, a_2, \dots, a_k\} \subset \mathbb{R}^n$ si dice* ortonormale *se i vettori $a_1, a_2, \dots, a_k$ hanno lunghezza 1 e sono a due a due perpendicolari, in formula se*

$$a_i \bullet a_j = \delta_{ij} \qquad \forall i, j = 1, \dots, k.$$

5.2 Proposizione (Formula di Carnot). *Per ogni $x, y \in \mathbb{R}^n$ si ha*

$$|x + y|^2 = |x|^2 + |y|^2 + 2\, x \bullet y\,.$$

In particolare x e y sono perpendicolari se e solo se vale il teorema di Pitagora,

$$x \bullet y = 0 \qquad \text{se e solo se} \qquad |x + y|^2 = |x|^2 + |y|^2.$$

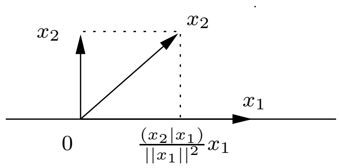

Figura 5.1.

5.b Prodotto scalare in coordinate

Un insieme ortonormale di vettori $\{e_1, e_2, \dots, e_k\} \subset \mathbb{R}^n$ è una base ortonormale dello spazio generato

$$W := \mathrm{Span}\Big\{e_1, e_2, \dots, e_k\Big\}.$$

Dunque ogni $x \in W$ si scrive in modo unico come $x = \sum_{i=1}^k x^i e_i$. D'altra parte, essendo $(e_1, e_2, \dots, e_k)$ ortonormale, per ogni $i = 1, \dots, k$

$$x \bullet e_i = \Big(\sum_{j=1}^n x^j e_j\Big) \bullet e_i = \sum_{j=1}^n x^j \, e_j \bullet e_i = \sum_{j=1}^n x^j \delta_{ji} = x^i,$$

i.e.,

$$x = \sum_{i=1}^k (\, x \bullet e_i \,) e_i.$$

I numeri $x \bullet e_i$, $i = 1, \dots, k$, detti i *coseni direttori* di x (rispetto al sistema ortonormale $(e_1, e_2, \dots, e_n)$) sono le coordinate di x nella base $(e_1, e_2, \dots, e_k)$. Si calcola quindi per $x = (x_1, x_2, \dots, x_n)$, $y = (y_1, y_2, \dots, y_n)$,

$$\begin{aligned} x \bullet y &= \Big(\sum_{i=1}^n x^i e_i\Big) \bullet \Big(\sum_{i=1}^n y^i e_i\Big) = \sum_{i,j=1}^n x^i y^j (\, e_i \bullet e_j \,) \\ &= \sum_{i=1}^n x^i y^i = \sum_{i=1}^n (\, x \bullet e_i \,)(\, y \bullet e_i \,), \end{aligned}$$

da cui segue il *teorema di Pitagora* nella forma

$$|x|^2 = x \bullet x = \sum_{i=1}^n |\, x \bullet e_i \,|^2.$$

5.c Basi ortonormali e algoritmo di Gram–Schmidt

A partire da un sistema di vettori linearmente indipendenti, si può costruire un sistema di vettori ortonormali che generano lo stesso sottospazio. Si dimostra

5.3 Teorema (Algoritmo di Gram–Schmidt). *Siano $(v_1, v_2, \dots, v_k)$ vettori linearmente indipendenti di $\mathbb{R}^n$. Esistono allora vettori ortonormali $(w_1, w_2, \dots, w_k)$ di $\mathbb{R}^n$ tali che per ogni $j = 1, 2, \dots, k$*

$$\mathrm{Span}\Big\{w_1, w_2, \dots, w_j\Big\} = \mathrm{Span}\Big\{v_1, v_2, \dots, v_j\Big\}.$$

5.d Proiezione ortogonale su un sottospazio

Sia V un sottospazio di $\mathbb{R}^n$ e sia $(v_1, v_2, \dots, v_k)$ una sua base ortonormale. Per ogni $x \in X$ sia $P(x) \in V \subset \mathbb{R}^n$ il vettore dato da

$$P(x) := \sum_{j=1}^{k} (\, x \bullet v_j \,)\, v_j \tag{5.1}$$

La mappa $P : \mathbb{R}^n \to \mathbb{R}^n$, $x \to P(x)$, si chiama la *proiezione ortogonale* su V. La ragione del nome è nel seguente

5.4 Teorema (della proiezione ortogonale). *Sono fatti equivalenti*

(i) *z è la proiezione perpendicolare di x su V, $z = P(x)$,*
(ii) *$x - z$ è perpendicolare a V, i.e. $(x - z) \bullet v = 0\ \forall v \in V$,*
(iii) *z è l'unico punto in V di minima distanza da x, i.e., $|z - x| < |v - x|$ per ogni $v \in V$, $v \neq z$.*

5.e Sottospazi ortogonali

Sia W un sottospazio di $\mathbb{R}^n$. Si chiama *ortogonale di W* il sottoinsieme

$$W^\perp := \Big\{ y \in \mathbb{R}^n \ \Big|\ y \bullet x = 0\ \forall x \in W \Big\}$$

È presto visto che

(i) $W^\perp$ è un sottospazio vettoriale di $\mathbb{R}^n$,
(ii) $W \cap W^\perp = \{0\}$,
(iii) $\dim W + \dim W^\perp = n$,
(iv) $(W^\perp)^\perp = W$,
(v) se P_W e $P_{W^\perp}$ sono rispettivamente le prioiezioni ortogonali su W e $W^\perp$, allora
$$P_{W^\perp}(x) = x - P_W(x) \qquad \forall x \in \mathbb{R}^n.$$

Dimostrazione. Dimostriamo la (iii). Sia $k := \dim W$ e sia $(v_1, v_2, \dots, v_k)$ una base ortonornale di W. Completiamo $v_1, v_2, \dots, v_k$ con altri $n-k$ vettori $v_{k+1}, \dots, v_n$ di $\mathbb{R}^n$ a formare una base $(v_1, v_2, \dots, v_n)$ di $\mathbb{R}^n$. Applicando a $(v_1, v_2, \dots, v_n)$ l'algoritmo di Gram–Schmidt, si trova una base ortonormale $(w_1, w_2, \dots, w_n)$ di $\mathbb{R}^n$. Per costruzione $w_i = v_i \in W\ \forall i = 1, \dots, k$, $w_{k+1}, \dots, w_n \in W^\perp$. Pertanto $w_{k+1}, \dots, w_n \in W^\perp$ è una base di $W^\perp$ e $\dim W^\perp = n - k = n - \dim W$. Le altre affermazioni sono ovvie. □

5.f Matrice trasposta

Sia $\mathbf{A} \in M_{m,n}(\mathbb{R})$ una matrice $m \times n$. La *trasposta* $\mathbf{A}^T$ *di* $\mathbf{A}$ è la matrice $n \times n$ ottenuta da $\mathbf{A}$ scambiando le righe con le colonne, in formula

$$(\mathbf{A}^T)^i_j = \mathbf{A}^j_i \qquad \forall i = 1, \dots, n,\ \forall j = 1, \dots, m.$$

5.5 Esercizio. Se $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 1 & 3 \\ 2 & 1 \end{pmatrix}$, allora $\mathbf{A}^T = \begin{pmatrix} 1 & 1 & 2 \\ 2 & 3 & 1 \end{pmatrix}$.

Il prodotto scalare permette di caratterizzare la matrice trasposta senza perdersi negli indici. Si ha

5.6 Proposizione. *Sia $\mathbf{A} \in M_{m,n}(\mathbb{R})$. Allora la trasposta è l'unica matrice $\mathbf{B} \in M_{n,m}(\mathbb{R})$ tale che*

$$\mathbf{B}y \bullet x = y \bullet \mathbf{A}x \qquad \forall x \in \mathbb{R}^n,\ \forall y \in \mathbb{R}^m. \tag{5.2}$$

5.g Vettore associato ad una applicazione lineare

Sia $L : \mathbb{R}^n \to \mathbb{R}$ lineare e $\mathbf{L} := [L(e_1) \,|\, \dots \,|\, L(e_n)]$ la matrice $1 \times n$ associata a L. Sia $F \in \mathbb{R}^n$ il vettore colonna

$$F := \begin{pmatrix} L(e_1) \\ L(e_2) \\ \vdots \\ L(e_n) \end{pmatrix} = \mathbf{L}^T$$

allora per ogni $x \in \mathbb{R}^n$, cfr. la (4.5),

$$L(x) = \mathbf{L}x = \sum_{i=1}^n x^i L(e_i) = F \bullet x\,.$$

La formula precedente è usata ad esempio in cinematica per definire le forze a partire dal lavoro elementare.

Più in generale, sia $\mathbf{A} = (a^i_j) \in M_{m,n}(\mathbb{R})$ una matrice $m \times n$ e siano $a^1, a^2, \dots, a^m$ gli m vettori-riga di $\mathbf{A}$ o, equivalentemente, le colonne di $\mathbf{A}^T$,

$$\mathbf{A} = \begin{pmatrix} a^1 \\ a^2 \\ \vdots \\ a^m \end{pmatrix}, \qquad \mathbf{A}^T = \Big[a^1 \,\Big|\, a^2 \,\Big|\, \dots \,\Big|\, a^m \Big].$$

Allora

$$\mathbf{A}x = \begin{pmatrix} a^1_1 x^1 + a^1_2 x^2 + \dots + a^1_n x^n \\ a^2_1 x^1 + a^2_2 x^2 + \dots + a^2_n x^n \\ \vdots \\ a^m_1 x^1 + a^m_2 x^2 + \dots + a^m_n x^n \end{pmatrix} = \begin{pmatrix} a^1 \bullet x \\ a^2 \bullet x \\ \vdots \\ a^m \bullet x \end{pmatrix}$$

Segue che

$$\mathbf{A}x = 0 \qquad \text{se e solo se} \qquad x \perp a^i\ \forall i = 1, \dots, m$$

o, equivalentemente,

$$\ker \mathbf{A} = \operatorname{Span}\Big\{ a^1, a^2, \dots, a^m \Big\}^\perp = \operatorname{Im}(\mathbf{A}^T)^\perp. \tag{5.3}$$

Pertanto, osservando anche che $(\mathbf{A}^T)^T = \mathbf{A}$ e $(W^\perp)^\perp = W$,

5.7 Proposizione. *Sia* $\mathbf{A} \in M_{m,n}(\mathbb{R})$. *Allora*

(i) $\ker \mathbf{A} = (\operatorname{Im} \mathbf{A}^T)^\perp$, $\operatorname{Im} \mathbf{A}^T = \ker \mathbf{A}^\perp$,
(ii) $\ker \mathbf{A}^T = \operatorname{Im} \mathbf{A}^\perp$, $\operatorname{Im} \mathbf{A} = (\ker \mathbf{A}^T)^\perp$.

Segue l'importante

5.8 Teorema. *Sia* $\mathbf{A} \in M_{m,n}(\mathbb{R})$. *Allora*

(i) $\operatorname{Rank} \mathbf{A}^T = \operatorname{Rank} \mathbf{A}$,
(ii) $\mathbf{A}$ *ha un minore* $p \times p$, $p := \operatorname{Rank} \mathbf{A}$, *non singolare.*

Dimostrazione. (i) Dalla formula del rango e dalla Proposizione 5.7 segue che

$$\begin{aligned}\operatorname{Rank}(\mathbf{A}) &= \dim \operatorname{Im} \mathbf{A} = \dim(\ker \mathbf{A}^T)^\perp \\ &= m - \dim \ker \mathbf{A}^T = \dim \operatorname{Im} \mathbf{A}^T = \operatorname{Rank}(\mathbf{A}^T).\end{aligned}$$

(ii) Sia $\mathbf{B}$ una sottomatrice $m \times p$ di colonne linearmente indipendenti di $\mathbf{A}$ formanti una base di $\operatorname{Im} \mathbf{A}$. Allora da (i) $\operatorname{Rank} \mathbf{B}^T = \operatorname{Rank} \mathbf{B} = \operatorname{Rank} \mathbf{A} = p$. $\mathbf{B}$ ha allora p righe linearmente indipendenti. Se $\mathbf{S}$ è la sottomatrice $p \times p$ di $\mathbf{B}$ formata da queste righe, allora da (i) $\operatorname{Rank} \mathbf{S} = \operatorname{Rank} \mathbf{S}^T = p$ e $\mathbf{S}$ è nonsingolare. □

Come conseguenza immediata della Proposizione 5.7, si ottiene

5.9 Teorema (dell'alternativa). *Sia* $\mathbf{A} \in M_{m,n}(\mathbb{R})$. *Allora* $x \to \mathbf{A}x$ *è una bigezione tra* $\ker \mathbf{A}^\perp = \operatorname{Im} \mathbf{A}^T$ *e* $\operatorname{Im} \mathbf{A} = (\ker \mathbf{A}^T)^\perp$. *In particolare*

(i) $\mathbf{A}x = y$ *ha almeno una soluzione se e solo se* y *è perpendicolare a* $\ker \mathbf{A}^T$.
(ii) y *è perpendicolare a* $\operatorname{Im} \mathbf{A}$ *se e solo se* $\mathbf{A}^T y = 0$.
(iii) $x \to \mathbf{A}x$ *è iniettiva se e solo se* $y \to \mathbf{A}^T y$ *è surgettiva.*
(iv) $x \to \mathbf{A}x$ *è surgettiva se e solo se* $y \to \mathbf{A}^T y$ *è iniettiva.*

5.10 Corollario. *Sia* $\mathbf{A} \in M_{m,n}(\mathbb{R})$. *Allora*

$$\operatorname{Rank} \mathbf{A}\mathbf{A}^T = \operatorname{Rank} \mathbf{A}^T = \operatorname{Rank} \mathbf{A} = \operatorname{Rank} \mathbf{A}^T \mathbf{A} \le \min(m, n).$$

5.h Sottospazi

Sia W un sottospazio di dimensione k in $\mathbb{R}^n$. Se $(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)$ è una base di W, allora $W = \operatorname{Im} \mathbf{L}$ dove

$$\mathbf{L} := \Big[\mathbf{v}_1 \Big| \mathbf{v}_2 \Big| \dots \Big| \mathbf{v}_k\Big].$$

La mappa $\mathbf{t} \to \mathbf{L}\mathbf{t}$ la cui immagine è W si chiama l'*equazione parametrica* di W generata da $(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)$. Osserviamo che scelte differenti della base di W individuano parametrizzazioni differenti e che le colomnne di $\mathbf{L}$ sono una base di W e $\operatorname{Rank} \mathbf{L} = k$.

Si puo' anche scrivere W come $W = \ker \mathbf{A}$ con $\mathbf{A} \in M_{n-k,n}(\mathbb{R})$. Diremo che l'equazione $\mathbf{A}x = 0$ è una *rappresentazione implicita di* W. Si noti che, essendo $\ker \mathbf{A} = W$, si ha $\operatorname{Rank} \mathbf{A}^T = \operatorname{Rank} \mathbf{A} = n - k$ percio' le righe di $\mathbf{A}$ sono linearmente indipendenti e generano l'ortogonale di W.

Si passa da una rappresentazione parametrica ad una implicita nel seguente modo. Supponiamo che $W = \operatorname{Im} \mathbf{L}$ dove $\mathbf{L} \in M_{n,k}(\mathbb{R})$ abbia rango massimo, $\operatorname{Rank} \mathbf{L} = k$. Esiste dunque un minore $\mathbf{M}$ di $\mathbf{L}$ di dimensione $k \times k$ non singolare. Supponendo che $\mathbf{M}$ sia costituito dalle prime k righe di $\mathbf{L}$,

$$\mathbf{L} = \begin{pmatrix} \boxed{\mathbf{M}} \\ \boxed{\mathbf{N}} \end{pmatrix}.$$

e decomponendo ξ come $x = (x', x'')$ con $x' \in \mathbb{R}^k$ and $x'' \in \mathbb{R}^{n-k}$, l'equazione

$$\mathbf{x} = \mathbf{Lt} \qquad \textit{si riscrive come} \qquad \begin{cases} \mathbf{Mt} = \mathbf{x}', \\ \mathbf{Nt} = \mathbf{x}''. \end{cases} \tag{5.4}$$

Essendo $\mathbf{M}$ invertibile,

$$\begin{cases} t = \mathbf{M}^{-1}\mathbf{x}', \\ \mathbf{NM}^{-1}\mathbf{x}' = \mathbf{x}'' \end{cases}$$

e dunque $\mathbf{x} \in \operatorname{Im} \mathbf{L} = W$ se e solo se $\mathbf{NM}^{-1}x' = x''$. Quelt'ultima equazione è una equazione implicita per W, i.e, della forma $\mathbf{A}x = 0$ se di definesce $\mathbf{A} \in M_{n-k,n}(\mathbb{R})$ come

$$\mathbf{A} = \begin{pmatrix} \boxed{\mathbf{NM}^{-1}} & \boxed{-\operatorname{Id}_n} \end{pmatrix}.$$

Supponiamo ora che $W = \ker \mathbf{A}$ dove $\mathbf{A} \in M_{n,k}(\mathbb{R})$ abbia rango massimo, $\operatorname{Rank} \mathbf{A} = n - k$. Selezioniamo $n - k$ colonne linearmente indipendenti. Supponendo che siano le prime $n-k$ a sinistra, chiamiamo $\mathbf{B} \in M_{n-k,n-k}(\mathbb{R})$ il minore di dimensione $(n-k) \times (n-k)$ formato da queste colonne e indichiamo x come $x = (x', x'')$ dove $x' \in \mathbb{R}^{n-k}$ e $x'' \in \mathbb{R}^k$. Allora l'equazione implicita $\mathbf{A}x = 0$ si riscrive come

$$\begin{pmatrix} \boxed{\mathbf{B}} & \boxed{\mathbf{C}} \end{pmatrix} \begin{pmatrix} x' \\ x'' \end{pmatrix} = 0, \qquad \text{i.e.,} \qquad \mathbf{B}\mathbf{x}' + \mathbf{C}\mathbf{x}'' = 0.$$

Essendo $\mathbf{B}$ invertibile, l'ultima equazione si riscrive come $x' = -\mathbf{B}^{-1}\mathbf{C}x''$, e dunque $x \in \ker \mathbf{A}$ se e solo se

$$x = \begin{pmatrix} \boxed{-\mathbf{B}^{-1}\mathbf{C}} \\ \boxed{\operatorname{Id}_k} \end{pmatrix} x'' := \mathbf{L}x'', \qquad x'' \in \mathbb{R}^k,$$

i.e., $W = \operatorname{Im} \mathbf{L}$.