Spazi vettoriali Euclidei

Ortogonalizzazione Lo spazio vettoriale $\mathbb{R}^n$ ha una base ortonormale ovvia, la base canonica $\mathbf{e}_1, \dots, \mathbf{e}_n$. Ogni sottospazio di $\mathbb{R}^n$ possiede delle basi ortonormali, anche se in generale non ne ha di ovvie. Di seguito mostraimo come a partire da una base di un sottospazio si possa sempre costruire una base ortogonale del sottospazio.

Proposizione 1 *Per ogni due vettori $\mathbf{f}, \mathbf{g}$ linearmente indipendenti di $\mathbb{R}^n$, si ha che il vettore*

$$\mathbf{h} = \mathbf{g} - \mathrm{pr}_{\mathbf{f}}(\mathbf{g}) = \mathbf{g} - \frac{\mathbf{f}^T\mathbf{g}}{\mathbf{f}^T\mathbf{f}}\,\mathbf{f}$$

e′ ortogonale ad $\mathbf{f}$; inoltre il sottospazio generato da $\mathbf{f}, \mathbf{h}$ coincide col sottospazio generato da $\mathbf{f}, \mathbf{g}$.

Questa proposizione e′ una riformulazione della proposizione sulla proiezione ortogonale di un vettore su un vettore non nullo. Si puo′ provare anche direttamente. Il fatto che $\mathbf{f}$ ed $\mathbf{h}$ siano ortogonali si verifica direttamente come segue

$$\mathbf{f}^T\mathbf{h} = \mathbf{f}^T\left(\mathbf{g} - \frac{\mathbf{f}^T\mathbf{g}}{\mathbf{f}^T\mathbf{f}}\,\mathbf{f}\right) = \mathbf{f}^T\mathbf{g} - \frac{\mathbf{f}^T\mathbf{g}}{\mathbf{f}^T\mathbf{f}}\,\mathbf{f}^T\mathbf{f} = \mathbf{f}^T\mathbf{g} - \mathbf{f}^T\mathbf{g} = 0.$$

Esempio I vettori $(1,0,2)$ e $(0,1,3)$ di $\mathbb{R}^3$ sono linearmente indipendenti. Allora il vettore

$$\begin{aligned}(0,1,3) - \mathrm{pr}_{(1,0,2)}((0,1,3)) &= (0,1,3) - \frac{1\cdot 0 + 0\cdot 1 + 3\cdot 2}{1\cdot 1 + 0\cdot 0 + 2\cdot 2}\,(1,0,2)\\ &= (0,1,3) - \frac{6}{5}\,(1,0,2) = (-\frac{6}{5}, 1, \frac{3}{5})\end{aligned}$$

e′ ortogonale al vettore $(1,0,2)$; inoltre, il sottospazio generato dai vettori $(1,0,2)$, $(-\frac{6}{5}, 1, \frac{3}{5})$ coincide col sottospazio generato dai vettori $(1,0,2)$, $(0,1,2)$.

Teorema 1 *Ogni sottospazio V di $\mathbb{R}^n$ possiede qualche base ortogonale, di piu′: per ogni sequenza $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m$ base di V, la sequenza $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m$ definita di seguito e′ una base ortogonale di V.* [1]

$$\begin{aligned}\mathbf{b}_1 &= \mathbf{a}_1;\\ \mathbf{b}_2 &= \mathbf{a}_2 - \mathrm{pr}_{\mathbf{b}_1}(\mathbf{a}_2);\\ \mathbf{b}_3 &= \mathbf{a}_3 - \mathrm{pr}_{\mathbf{b}_1}(\mathbf{a}_3) - \mathrm{pr}_{\mathbf{b}_2}(\mathbf{a}_3);\\ &\vdots\\ \mathbf{b}_m &= \mathbf{a}_m - \sum_{i=1}^{m} \mathrm{pr}_{\mathbf{b}_i}(\mathbf{a}_m);\end{aligned}$$

[1] In realta′ vale una proprieta′ piu′ forte, che caratterizza i vettori $\mathbf{b}_i$ a meno di uno scalare: per ogni $\ell = 1, \dots, m$ la sequenza $\mathbf{b}_1, \dots, \mathbf{b}_\ell$ e′ una base ortogonale del sottospazio di $\mathbb{R}^n$ generato dai vettori $\mathbf{a}_1, \dots, \mathbf{a}_\ell$.

Il processo di costruzione della sequenza dei vettori $\mathbf{b}_*$ a partire dalla sequenza dei vettori $\mathbf{a}_*$ si dice "processo di ortogonalizzazione di Gram-Schmidt."

Esempio A partire dai vettori

$$(1,0,1),(0,1,1),(0,0,1),$$

che sono una base di $\mathbb{R}^3$, il processo di Gram-Schmidt fornisce la seguente sequenza di vettori che e' una base ortogonale di $\mathbb{R}^3$.

$$\begin{aligned}
\mathbf{b}_1 &= (1,0,1) \\
\mathbf{b}_2 &= (0,1,1) - \mathrm{pr}_{(1,0,1)}((0,1,1)) \\
&= (0,1,1) - \frac{1}{2}(1,0,1) \\
&= (-\frac{1}{2},1,\frac{1}{2}) \\
\mathbf{b}_3 &= (0,0,1) - \mathrm{pr}_{(1,0,1)}((0,0,1)) - \mathrm{pr}_{(-\frac{1}{2},1,\frac{1}{2})}((0,0,1)) \\
&= (0,0,1) - \frac{1}{2}(1,0,1) - \frac{1}{3}(-\frac{1}{2},1,\frac{1}{2}) \\
&= (-\frac{1}{3},-\frac{1}{3},\frac{1}{3})
\end{aligned}$$

Matrici ortogonali Sia A una matrice quadrata di ordine n. La condizione "le colonne A_{*h} ($h = 1,\ldots,n$) di A sono una base ortonormale di $\mathbb{R}^n$" in simboli si esprimere nella forma

$$A_{*i}^T A_{*j} = \begin{cases} 1 & \text{per } i = j \\ 0 & \text{per } i \neq j \end{cases} \qquad (i,j = 1,\ldots,n)$$

e si puo' sitetizzare nella forma

$$A^T A = I_n,$$

che a sua volta significa "la matrice trasposta A^T e' un'inversa sinistra di A." Analogamente, la condizione "le righe A_{h*} ($h = 1,\ldots,n$) di A sono una base ortonormale di $\mathbb{R}^n$" in simboli si esprimere nella forma

$$A_{i*} A_{j*}^T = \begin{cases} 1 & \text{per } i = j \\ 0 & \text{per } i \neq j \end{cases} \qquad (i,j = 1,\ldots,n)$$

e si puo' sintetizzare nella forma

$$AA^T = I_n,$$

che a sua volta significa "la matrice trasposta A^T e' un'inversa destra di A." Approfondendo un poco, si ottiene il

Teorema 2 *Per una matrice A quadrata di ordine n sono equivalenti:*
(1) le colonne di A sono base ortonormale di R^n;
(1') $A^T A = I_n$;
(2) le righe di A sono base ortonormale di R^n;
(2') $AA^T = I_n$;
(3) A e' invertibile e la sua inversa e' uguale alla sua trasposta: $A^{-1} = A^T$.

Definizione 1 *Una matrice quadrata che soddisfa una, e quindi ciascuna, delle condizioni del precedente teorema si dice "matrice ortogonale".*

Segnaliamo il fatto che per una matrice A quadrata di ordine n, le condizioni "le colonne di A sono una base ortogonale di $\mathbb{R}^n$" e "le righe di A sono una base ortogonale di $\mathbb{R}^n$" non sono collegate. Ad esempio, la matrice

$$\begin{bmatrix} 1 & -4 \\ 2 & 2 \end{bmatrix}$$

ha le colonne ortogonali, ma non ha le righe ortogonali. Se si normalizzano le colonne, si ottiene la matrice

$$\begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{4}{\sqrt{20}} \\ \frac{2}{\sqrt{5}} & \frac{2}{\sqrt{20}} \end{bmatrix}$$

che ha sia le colonne che le righe ortonormali.

Algebra delle matrici

Le funzioni piu' semplici $f : \mathbb{R} \to \mathbb{R}$ sono quelle date dalla moltiplicazione per una costante $f : \ x \mapsto ax$ (con a costante in $\mathbb{R}$), che hanno per grafico le rette passanti per l'origine; le funzioni piu' semplici $f : \mathbb{R}^n \to \mathbb{R}$ sono quelle date da un polinomio omogeneo di I grado in n variabili

$$f : \ (x_1, \ldots, x_n) \mapsto a_1 x_1 + \cdots + a_n x_n \qquad (a_j \text{ costanti in } \mathbb{R});$$

le funzioni piu' semplici $f : \mathbb{R}^n \to \mathbb{R}^m$ sono quelle date da m polinomi omogenei di I grado in n variabili

$$f : \ (x_1, \ldots, x_n) \mapsto (a_{11}x_1 + \cdots + a_{1n}x_n, \ldots, a_{m1}x_1 + \cdots + a_{mn}x_n), \qquad (a_{ij} \text{ costanti in } \mathbb{R}).$$

Scrivendo i vettori come vettori colonna, questa descrizione diviene

$$f : \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \mapsto \begin{bmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{bmatrix} = \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

Dunque le funzioni piu' semplici $f : \mathbb{R}^n \to \mathbb{R}^m$ sono quelle date dalla moltiplicazione per una matrice di tipo $m \times n$ costante

$$f : \mathbf{x} \mapsto A\mathbf{x}, \qquad (A \text{ costante in } \mathbb{R}^{m \cdot n}).$$

Si osservi che per una tale funzione si ha

$$f(\mathbf{u}+\mathbf{v}) = A(\mathbf{u}+\mathbf{v}) = A\mathbf{u} + A\mathbf{v} = f(\mathbf{u}) + f(\mathbf{v}) \qquad \forall \mathbf{u}, \mathbf{v} \in \mathbb{R}^n$$

e

$$f(\alpha\mathbf{u}) = A(\alpha\mathbf{u}) = \alpha A\mathbf{u} = \alpha f(\mathbf{u}) \qquad \forall \mathbf{u} \in \mathbb{R}^n,\ \alpha \in \mathbb{R}.$$

Questa fatto suggerisce di dare la seguente

Definizione 2 *Una funzione $f : V \to W$ da uno spazio vettoriale V ad uno spazio vettoriale W si dice funzione lineare se e solo se*

$$\begin{aligned} f(\mathbf{u}+\mathbf{v}) &= f(\mathbf{u}) + f(\mathbf{v}) && \forall \mathbf{u}, \mathbf{v} \in V \\ f(\alpha\mathbf{u}) &= \alpha f(\mathbf{u}) && \forall \mathbf{u} \in V,\ \alpha \in \mathbb{R}. \end{aligned}$$

Non solo le funzioni di sopra sono i primi esempi di funzioni lineari $R^n \to \mathbb{R}^m$, ma sono anche gli unici, nel senso della seguente

Proposizione 2 *Una funzione $f : \mathbb{R}^n \to \mathbb{R}^m$ e' lineare se e solo se e' del tipo $f : \mathbf{x} \mapsto A\mathbf{x}$, con A matrice di tipo $m \cdot n$ costante.*

Dimostrazione. Dobbiamo in sostanza provare che se $f : \mathbb{R}^n \to \mathbb{R}^m$ e' lineare, allora e' del tipo $f : \mathbf{x} \mapsto A\mathbf{x}$, con A matrice di tipo $m \cdot n$ costante. Cio' si prova come segue

$$\begin{aligned} f(\mathbf{x}) &= f(\mathbf{e}_1 x_1 + \cdots + \mathbf{e}_n x_n) \\ &= f(\mathbf{e}_1)x_1 + \cdots + f(\mathbf{e}_n)x_n \\ &= \mathbf{a}_1 x_1 + \cdots + \mathbf{a}_n x_n \\ &= [\, \mathbf{a}_1 \ \ldots \ \mathbf{a}_n] \, \mathbf{x} \\ &= A\mathbf{x}. \end{aligned}$$

Abbiamo in realta' visto anche qualcosa di piu': la matrice A e' la matrice che ha per colonne le immagini $\mathbf{a}_1, \ldots, \mathbf{a}_n$ che f crea in $\mathbb{R}^m$ dei vettori della base canonica $\mathbf{e}_1, \ldots, \mathbf{e}_n$ di $\mathbb{R}^n$.