Monica Bianchi - Grazia Messineo

Appunti di matematica per l'analisi economica

ELEMENTI DI ALGEBRA LINEARE
OTTIMIZZAZIONE STATICA

Giappichelli

Prefazione

Questa opera è una rielaborazione di una dispensa del corso di Matematica per le Applicazioni Economiche e Finanziarie tenuto dagli autori presso l'Università Cattolica del Sacro Cuore di Milano. È rivolto a studenti di corsi di laurea triennale e magistrale in Economia interessati ad approfondire alcuni strumenti matematici utili nella trattazione di problemi economici e finanziari.
La presentazione degli argomenti è principalmente di carattere teorico, anche se si è preferito non soffermarsi sulle dimostrazioni degli enunciati, ma accompagnare i singoli argomenti con numerosi esempi ed esercizi che comprendono anche alcune tipiche applicazioni alla microeconomia e alla macroeconomia.
L'opera è strutturata in due volumi. Nel primo sono trattati alcuni elementi di algebra lineare utili per lo sviluppo delle parti successive e l'ottimizzazione statica, sia libera che vincolata; nel secondo volume viene esposta la teoria dei sistemi dinamici, sia nel caso continuo che nel caso discreto. I volumi sono entrambi accompagnati da una ampia raccolta di esercizi. Infine completano i volumi alcune appendici che richiamano argomenti propedeutici al materiale presentato.
Alcuni esercizi raccolti nel capitolo 6 sono ispirati o tratti da temi d'esame di vari insegnamenti dell'Università Cattolica del Sacro Cuore di Milano e dell'Università degli Studi di Milano - Bicocca o da dispense scritte per gli studenti delle stesse Università (Foresti e Vassallo (2000), Vassallo (2006), Vassallo (2007), Bianchi e Pini (2008), Bianchi e Torriero (2001)).
Ringraziamo tutti gli autori che ci hanno messo a disposizione il loro materiale: la prof. Paola Foresti, il prof. Nicolò Pecora, la prof. Rita Pini, la prof. Anna Torriero. Un ringraziamento particolare va al prof. Salvatore Vassallo e alla prof. Carla Peri, per il materiale messo a disposizione, per aver letto il manoscritto e per gli utili suggerimenti: il loro contributo ci ha permesso di migliorare sensibilmente la presentazione degli argomenti.
Non escludiamo, in ogni caso, la presenza di refusi o omissioni, di cui siamo interamente responsabili. Si ringrazia fin d'ora chi volesse evidenziarli.

Parte I

Elementi di Algebra Lineare

CAPITOLO 1

Elementi introduttivi di algebra lineare

1.1 Spazi vettoriali reali

Si consideri un insieme $\mathcal{V}$ e l'insieme $\mathbb{R}$ dei numeri reali, nel seguito denominati anche *scalari*. Si dice che $\mathcal{V}$ é uno *spazio vettoriale* (o *spazio lineare*) *reale* [(1)] se in esso sono definite due operazioni

i) $\mathcal{V} \times \mathcal{V} \to \mathcal{V}, \quad (\mathbf{v}, \mathbf{w}) \mapsto \mathbf{v} + \mathbf{w}, \quad \forall \mathbf{v}, \mathbf{w} \in \mathcal{V}$

ii) $\mathbb{R} \times \mathcal{V} \to \mathcal{V}, \quad (a, \mathbf{w}) \mapsto a \cdot \mathbf{w}, \quad \forall \mathbf{w} \in \mathcal{V}, \forall a \in \mathbb{R}$

denominate, rispettivamente, *somma* e *prodotto esterno (o moltiplicazione per uno scalare)*, con le seguenti proprietà:

i) la somma è commutativa, associativa, ammette un elemento neutro, indicato con $\mathbf{O}_{\mathcal{V}}$, e ogni elemento $\mathbf{v} \in \mathcal{V}$ ammette un inverso, indicato con $-\mathbf{v}$, tale che $\mathbf{v} + (-\mathbf{v}) = \mathbf{O}_{\mathcal{V}}$;

ii) il prodotto esterno è associativo, $(ab) \cdot \mathbf{v} = a \cdot \mathbf{v} + b \cdot \mathbf{v}$, e inoltre $1 \cdot \mathbf{v} = \mathbf{v}$ per ogni $\mathbf{v} \in \mathcal{V}$;

iii) valgono le proprietà distributive del prodotto esterno rispetto alla somma di vettori e della somma di scalari rispetto al prodotto esterno:

$$a \cdot (\mathbf{v} + \mathbf{w}) = a \cdot \mathbf{v} + a \cdot \mathbf{w};$$
$$(a + b) \cdot \mathbf{v} = a \cdot \mathbf{v} + b \cdot \mathbf{v}.$$

[(1)] Si usa il termine reale perché gli scalari in questione sono numeri reali. In modo analogo, si potrebbero considerare spazi vettoriali complessi supponendo che gli scalari siano numeri complessi.

L'esempio più noto di spazio vettoriale reale è costituito dall'insieme $\mathbb{R}^n$ delle n–ple ordinate di numeri reali

$$\mathbb{R}^n = \{\mathbf{x} = (x_1, x_2, \cdots, x_n), \quad x_i \in \mathbb{R}, \quad i = 1, 2, \ldots, n\},$$

provvisto delle seguenti operazioni: per ogni $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, la somma è definita da

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, x_2 + y_2, \cdots, x_n + y_n)$$

e, per ogni $\mathbf{x} \in \mathbb{R}^n$, $a \in \mathbb{R}$, il prodotto esterno è definito da

$$a\mathbf{x} = (ax_1, ax_2, \cdots, ax_n).$$

L'elemento neutro della somma è il vettore nullo $\mathbf{0} = (0, 0, \cdots, 0)$, e l'opposto di un elemento $\mathbf{x} = (x_1, x_2, \cdots, x_n)$ è l'elemento $-\mathbf{x} = (-x_1, -x_2, \cdots, -x_n)$. I vettori

$$\mathbf{e_1} = (1, 0, \cdots, 0),\ \mathbf{e_2} = (0, 1, 0, \cdots, 0), \quad \ldots \quad, \mathbf{e_n} = (0, 0, \cdots, 1)$$

sono detti *versori fondamentali.*

Per $n = 2, 3$, gli elementi degli spazi vettoriali $\mathbb{R}^2$ e $\mathbb{R}^3$ possono essere rappresentati graficamente come i punti di un piano cartesiano e di uno spazio cartesiano, rispettivamente, come si può vedere nella Figura 1

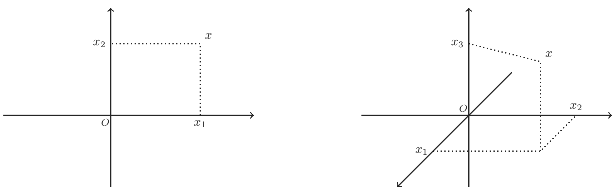

Figura 1 Gli elementi degli spazi $\mathbb{R}^2$ e $\mathbb{R}^3$

Tali punti possono essere posti in corrispondenza biunivoca con i vettori spiccati dall'origine degli assi nello spazio bidimensionale e tridimensionale; per questo motivo è stato introdotto il termine di spazio vettoriale e il termine *vettore* per indicare il generico elemento di tale spazio.

In $\mathbb{R}^2$ le operazioni di somma e moltiplicazione scalare hanno un'interpretazione geometrica, illustrata nella Figura 2:

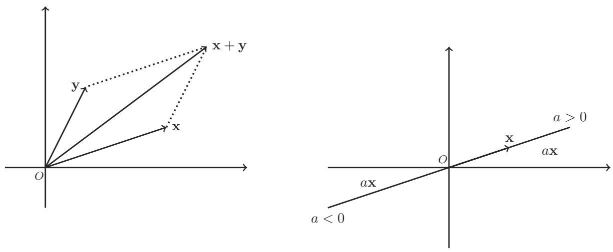

Figura 2 Interpretazione geometrica della somma e del prodotto esterno in $\mathbb{R}^2$

Nel seguito il generico vettore di $\mathbb{R}^n$, in accordo a notazioni standard di algebra lineare, verrà indicato come un vettore colonna o in modo equivalente come il trasposto di un vettore riga:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{bmatrix} = [x_1, x_2, \cdots, x_n]^T$$

Altri esempi importanti di spazi vettoriali reali sono l'insieme $\mathcal{C}^k(a, b)$ delle funzioni derivabili con continuità fino all'ordine k sull'intervallo (a, b), l'insieme $\mathcal{C}^\infty(a, b)$ delle funzioni derivabili infinite volte sull'intervallo (a, b) e l'insieme delle successioni reali $x_t, t \in \mathbb{N}$. In tutti questi casi, i vettori sono funzioni reali e le operazioni sono le usuali operazioni di somma di funzioni e di moltiplicazione di una funzione per uno scalare.

Utilizzando le operazioni di somma e prodotto esterno, si può introdurre la seguente definizione:

Definizione 1.1.1. Sia $\mathcal{V}$ uno spazio vettoriale reale. Si chiama *combinazione lineare* dei vettori $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_k} \in \mathcal{V}$, con coefficienti reali $a_1, a_2, \ldots, a_k$, il vettore $\mathbf{v} \in \mathcal{V}$ dato da

$$\mathbf{v} = a_1 \cdot \mathbf{v_1} + a_2 \cdot \mathbf{v_2} + \cdots a_k \cdot \mathbf{v_k}. \tag{1.1.1}$$

L'insieme di tutte le combinazioni lineari dei vettori $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_k}$ viene indicata con $\mathcal{S}(\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_k})$:

$$\mathcal{S}(\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_k}) = \left\{ \mathbf{v} \in \mathcal{V} : \mathbf{v} = \sum_{i=1}^{k} a_i \cdot \mathbf{v_i}, \quad a_i \in \mathbb{R} \right\}.$$

In particolare, se per ogni $i = 1, 2, \ldots, k$ si ha $a_i \geq 0$ e $\sum_{i=1}^{k} a_i = 1$, la combinazione lineare $\sum_{i=1}^{k} a_i \cdot \mathbf{v_i}$ si chiama *combinazione lineare convessa* dei vettori $\{\mathbf{v_i}\}$.
Nel seguito la notazione

$$\mathrm{co}(\mathbf{v_1}, \mathbf{v_2}, \cdots, \mathbf{v_k})$$

indicherà l'insieme delle combinazioni lineari convesse dei k vettori $\mathbf{v_1}, \mathbf{v_2}, \cdots, \mathbf{v_k}$.
Si osservi che il risultato di una combinazione lineare è sempre un elemento di $\mathcal{V}$, in virtù della chiusura dell'insieme $\mathcal{V}$ rispetto alle operazioni di somma e moltiplicazione scalare.

Esempio 1.1.2.

1) Se $\mathcal{V} = \mathbb{R}^3$, $\mathbf{x_1} = [1, 3, 2]^T$, $\mathbf{x_2} = [2, -1, 0]^T$, $a_1 = 3$, $a_2 = 2$ allora

$$a_1\mathbf{x_1} + a_2\mathbf{x_2} = 3[1, 3, 2]^T + 2[2, -1, 0] = [7, 7, 6]^T.$$

2) Se $\mathcal{V} = \mathcal{C}^\infty(0, +\infty)$, dati i vettori $\mathrm{e}^x, x^2, \ln x$, e gli scalari $2, -1, 4$, la loro combinazione lineare è la funzione $v(x)$ data da

$$v(x) = 2\,\mathrm{e}^x - x^2 + 4\ln x.$$

Si noti che se $a_i = 0$ per ogni $i = 1, 2, \cdots, k$, la combinazione lineare (1.1.1) è nulla, per ogni scelta dei vettori $\mathbf{v_i}$. In modo analogo, se $\mathbf{v_i} = \mathbf{0}_\mathcal{V}$ per ogni $i = 1, 2, \cdots, k$, la combinazione lineare (1.1.1) coincide con il vettore nullo $\mathbf{0}_\mathcal{V}$.
Consideriamo ora la seguente combinazione lineare di vettori di $\mathbb{R}^2$:

$$3[3, 2]^T - [-3, 4]^T - 2[6, 1]^T = [0, 0]^T.$$

Nonostante né i coefficienti né i vettori siano nulli, il risultato è il vettore nullo. Si osservi che uno dei tre vettori coincide con una combinazione lineare dei rimanenti: $[-3, 4]^T = 3[3, 2]^T - 2[6, 1]^T$. Quanto detto motiva la seguente definizione:

Definizione 1.1.3. I vettori $\mathbf{v_1}, \mathbf{v_2}, \cdots, \mathbf{v_k} \in \mathcal{V}$ si dicono *linearmente indipendenti* se nessuno di essi può essere espresso come combinazione lineare dei rimanenti.

È allora immediato verificare che i vettori $\mathbf{v_1}, \mathbf{v_2}, \cdots, \mathbf{v_k} \in \mathcal{V}$ sono linearmente indipendenti se e solo se è soddisfatta la seguente implicazione

$$a_1 \cdot \mathbf{v_1} + a_2 \cdot \mathbf{v_2} + \cdots a_k \cdot \mathbf{v_k} = \mathbf{0}_\mathcal{V} \Longrightarrow a_i = 0, \quad \forall i = 1, 2, \cdots k.$$

Se i vettori $\mathbf{v_1}, \mathbf{v_2}, \cdots, \mathbf{v_k}$ di $\mathcal{V}$ **non** sono linearmente indipendenti, si dice che sono *linearmente dipendenti.* In questo caso, esiste almeno un vettore $\mathbf{v_i}$ che può essere espresso come combinazione lineare dei rimanenti:

$$\mathbf{v_i} = \tilde{a}_1 \cdot \mathbf{v_1} + \tilde{a}_2 \cdot \mathbf{v_2} + \cdots \tilde{a}_{i-1} \cdot \mathbf{v_{i-1}} + \tilde{a}_{i+1} \cdot \mathbf{v_{i+1}} + \cdots \tilde{a}_n \cdot \mathbf{v_n}.$$

Osservazione 1.1.4.

i) Il concetto di dipendenza e indipendenza lineare è relativo a un insieme di vettori e non ad un singolo vettore.

ii) Un insieme di vettori che contiene il vettore nullo è sempre un insieme di vettori linearmente dipendenti.

iii) La dipendenza lineare di due vettori non nulli $\mathbf{v_1}, \mathbf{v_2}$ equivale al fatto che ciascuno di essi è multiplo dell'altro (cioè: $\mathbf{v_2} = a \cdot \mathbf{v_1}$, oppure $\mathbf{v_1} = a \cdot \mathbf{v_2}$, per un opportuno $a \in \mathbb{R}$).

iv) Se uno tra i vettori $\mathbf{v_1}, \mathbf{v_2}, \cdots, \mathbf{v_k}$ può essere espresso come combinazione lineare degli altri, non è detto che questo valga per tutti.

v) Se ad un insieme di vettori linearmente dipendenti si aggiunge un vettore, la dipendenza lineare si conserva.

vi) Se ad un insieme di vettori linearmente indipendenti se ne toglie uno, l'indipendenza lineare si conserva.

vii) Se $\mathbf{v_1}, \mathbf{v_2}, \cdots, \mathbf{v_k}$ sono linearmente indipendenti, ma $\mathbf{v_1}, \mathbf{v_2}, \cdots, \mathbf{v_k}, \mathbf{w}$ sono dipendenti, allora $\mathbf{w}$ può essere espresso come combinazione lineare dei vettori $\mathbf{v_1}, \mathbf{v_2}, \cdots, \mathbf{v_k}$.

viii) Gli n versori fondamentali di $\mathbb{R}^n$ sono linearmente indipendenti.

Vediamo ora come poter verificare l'indipendenza lineare di vettori appartenenti ad opportuni spazi vettoriali.

1) Sia $\mathcal{V} = \mathbb{R}^n$. Dati k vettori $\mathbf{x^1}, \mathbf{x^2}, \cdots, \mathbf{x^k}$, l'uguaglianza

$$c_1\mathbf{x^1} + c_2\mathbf{x^2} + \cdots + c_k\mathbf{x^k} = \mathbf{0} \tag{1.1.2}$$

si esprime in forma compatta come $X\mathbf{c} = \mathbf{0}$ dove X è la matrice di tipo (n, k) ottenuta accostando i k vettori $\mathbf{x^1}, \mathbf{x^2}, \cdots, \mathbf{x^k}$

$$X = [\mathbf{x^1}, \mathbf{x^2}, \cdots, \mathbf{x^k}]$$

e $\mathbf{c} = [c_1, c_2, \cdots, c_k]^T$.

Se i vettori $\mathbf{x^1}, \mathbf{x^2}, \cdots, \mathbf{x^k}$ sono linearmente dipendenti, esistono degli scalari $c_1, c_2, \cdots, c_k$ non tutti nulli che soddisfano la (1.1.2) e quindi il sistema omogeneo $X\mathbf{c} = \mathbf{0}$ ammette soluzioni non banali. Questo si verifica se il rango $r(X)$ della matrice X risulta inferiore a k. Se, invece, il rango della matrice X è pari a k, il sistema $X\mathbf{c} = \mathbf{0}$ ammette solo la soluzione banale $\mathbf{c} = \mathbf{0}$ ed i vettori $\mathbf{x^1}, \mathbf{x^2}, \cdots, \mathbf{x^k}$ sono linearmente indipendenti.

2) Sia $\mathcal{V} = C^\infty(a,b)$. Date k funzioni $\varphi_1, \varphi_2, \cdots \varphi_k \in C^\infty(a,b)$, per ogni $t \in (a,b)$, si consideri la *matrice Wronskiana*

$$W(t) = \begin{bmatrix} \varphi_1(t) & \varphi_2(t) & \cdots & \varphi_k(t) \\ \varphi_1^{(1)}(t) & \varphi_2^{(1)}(t) & \cdots & \varphi_k^{(1)}(t) \\ \cdots & \cdots & \cdots & \cdots \\ \varphi_1^{(k-1)}(t) & \varphi_2^{(k-1)}(t) & \cdots & \varphi_k^{(k-1)}(t) \end{bmatrix}$$

costruita derivando $k-1$ volte le funzioni $\varphi_1, \varphi_2, \cdots \varphi_k$ nell'intervallo (a,b).

Se le funzioni $\varphi_1, \varphi_2, \cdots \varphi_k$ sono linearmente dipendenti nell'intervallo (a,b), esistono degli scalari $c_1, c_2, \cdots, c_k$ non tutti nulli tali che

$$c_1\varphi_1(t) + c_2\varphi_2(t) + \cdots + c_k\varphi_k(t) = \mathbf{0}, \quad \forall t \in (a,b)$$

e derivando $(k-1)$ volte, si ottiene anche

$$c_1\varphi_1^{(j)}(t) + c_2\varphi_2^{(j)}(t) + \cdots + c_k\varphi_k^{(j)}(t) = \mathbf{0}, \quad \forall t \in (a,b), 1 \le j \le (k-1).$$

Posto $\mathbf{c} = [c_1, c_2, \cdots, c_k]^T$, il sistema omogeneo $W(t)\mathbf{c} = \mathbf{0}$ ammette quindi soluzioni non banali per ogni $t \in (a,b)$ e quindi il determinante della matrice Wronskiana $W(t)$ deve risultare identicamente nullo per ogni $t \in (a,b)$. Se invece il determinante non si annulla in almeno un punto di (a,b), le funzioni sono linearmente indipendenti.

Se ad esempio consideriamo le funzioni

$$\varphi_1(t) = \mathrm{e}^{\lambda_1 t}, \varphi_2(t) = \mathrm{e}^{\lambda_2 t}, \cdots, \varphi_k(t) = \mathrm{e}^{\lambda_k t}$$

il determinante della matrice Wronskiana risulta

$$\det(W(t)) = \mathrm{e}^{(\lambda_1+\lambda_2+\cdots+\lambda_k)t} \det \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_k \\ \cdots & \cdots & \cdots & \cdots \\ \lambda_1^{(k-1)} & \lambda_2^{(k-1)} & \cdots & \lambda_k^{(k-1)} \end{bmatrix}$$

Poiché tale determinante risulta non nullo in (a,b) se e solo se gli scalari $\lambda_i, i = 1, 2, \cdots, k$ sono tutti distinti, questa condizione garantisce l'indipendenza lineare delle funzioni considerate.

Attenzione: non si può concludere nulla sulla dipendenza/indipendenza lineare delle funzioni se il determinante della matrice Wronskiana si annulla identicamente in (a,b). Consideriamo ad esempio le funzioni $\varphi_1(t) = t^3$ e $\varphi_2(t) = |t|^3$. È facile verificare che le funzioni sono linearmente indipendenti in $\mathbb{R}$, mentre la matrice Wronskiana

$$W(t) = \begin{bmatrix} t^3 & |t|^3 \\ 3t^2 & 3t|t| \end{bmatrix}$$

ha $\det W(t) = 3t^4|t| - 3t^2|t|^3 = 0$ per ogni $t \in \mathbb{R}$.

Passiamo ora ad introdurre la nozione di sottospazio vettoriale.

Definizione 1.1.5. Un sottoinsieme $\mathcal{W}$ di uno spazio vettoriale $\mathcal{V}$ che sia chiuso rispetto alla somma e alla moltiplicazione scalare, cioè tale che

1) $\mathbf{v} + \mathbf{w} \in \mathcal{W}, \quad \forall \mathbf{v}, \mathbf{w} \in \mathcal{W};$

2) $a \cdot \mathbf{w} \in \mathcal{W}, \quad \forall \mathbf{w} \in \mathcal{W}, \quad \forall a \in \mathbb{R},$

si chiama *sottospazio vettoriale di* $\mathcal{V}$ (o *sottospazio lineare di* $\mathcal{V}$)

In pratica, un sottospazio $\mathcal{W}$ di $\mathcal{V}$ risulta a sua volta uno spazio vettoriale rispetto alle operazioni definite in $\mathcal{V}$.
È facile verificare che i criteri 1) e 2) possono essere unificati richiedendo che

$$a \cdot \mathbf{v} + b \cdot \mathbf{w} \in \mathcal{W}, \quad \forall \mathbf{v}, \mathbf{w} \in \mathcal{W}, \quad \forall a, b \in \mathbb{R}.$$

Più in generale se $\mathcal{W}$ è un sottospazio vettoriale e $\mathbf{v_1}, \mathbf{v_2}, \cdots, \mathbf{v_k}$ sono elementi di $\mathcal{W}$ si ha

$$\mathcal{S}(\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_k}) \subseteq \mathcal{W}.$$

Esempio 1.1.6.

1) $\{\mathbf{0}\}$, $\mathcal{V}$ sono sottospazi vettoriali di $\mathcal{V}$, detti sottospazi *banali*.

2) Dati k vettori $\mathbf{v_1}, \mathbf{v_2}, \cdots, \mathbf{v_k}$, appartenenti ad uno spazio vettoriale $\mathcal{V}$, il sottoinsieme

$$\mathcal{W} = \mathcal{S}(\mathbf{v_1}, \mathbf{v_2}, \cdots, \mathbf{v_k})$$

è un sottospazio vettoriale di $\mathcal{V}$. Inoltre, se $\mathcal{V} = \mathbb{R}^n$, si dimostra che tutti i sottospazi vettoriali sono di questo tipo per un'opportuna scelta dei vettori $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$, $k \leq n$. In particolare, se $k = 1$ e $\mathbf{v}_1 \neq \mathbf{0}$,

$$\mathcal{S}(\mathbf{v}_1) = \{\mathbf{v} \in \mathbb{R}^n : \ \mathbf{v} = t\mathbf{v}_1, \quad t \in \mathbb{R}\}$$

viene detta *retta orientata* passante per l'origine con direzione e verso assegnati dal vettore $\mathbf{v}_1$.

3) Sia A una matrice non nulla di tipo (m, n) ed $\mathbf{x} \in \mathbb{R}^n$. È facile verificare che l'insieme delle soluzioni del sistema lineare omogeneo $A\mathbf{x} = \mathbf{0}$ costituisce un sottospazio vettoriale di $\mathbb{R}^n$, detto *nucleo di A*, banale nel caso in cui $r(A) = n$.

Inoltre, anche l'*immagine* di A

$$\mathrm{Im}(A) = \{\mathbf{y} \in \mathbb{R}^m : \ \mathbf{y} = A\mathbf{x} \text{ per qualche } \mathbf{x} \in \mathbb{R}^n\}$$

è un sottospazio vettoriale di $\mathbb{R}^m$, banale nel caso in cui $r(A) = m$.

Se $\mathcal{W} = \mathcal{S}(\mathbf{v_1}, \mathbf{v_2}, \cdots, \mathbf{v_k})$ si dice che $\mathcal{W}$ *è generato* dai vettori $\{\mathbf{v_i}\}_1^k$, o che i vettori $\{\mathbf{v_i}\}_1^k$ *generano* lo spazio $\mathcal{W}$.
In caso di dipendenza lineare dei generatori, si può eliminare almeno uno di essi, senza modificare il sottospazio vettoriale generato e si può iterare il procedimento di eliminazione fino a quando l'insieme dei generatori rimasti risulta un insieme di vettori linearmente indipendenti. Quanto detto suggerisce la seguente definizione:

> **Definizione 1.1.7.** Sia $\mathcal{V}$ uno spazio vettoriale e $\mathcal{W} = \mathcal{S}(\mathbf{v_1}, \mathbf{v_2}, \ldots \mathbf{v_k})$.
> Se i vettori $\mathbf{v_1}, \mathbf{v_2}, \cdots, \mathbf{v_k}$ sono linearmente indipendenti, si dice che essi costituiscono una *base* per il sottospazio vettoriale $\mathcal{W}$.

Se i vettori $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_k}$ costituiscono una base per il sottospazio vettoriale $\mathcal{W} = \mathcal{S}(\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_k})$, ogni $\mathbf{w} \in \mathcal{W}$ può essere scritto come combinazione lineare dei vettori $\{\mathbf{v_i}\}_1^k$

$$\mathbf{w} = a_1 \cdot \mathbf{v_1} + \cdots + a_k \cdot \mathbf{v_k}$$

ed, inoltre, i coefficienti a_i sono univocamente determinati e vengono detti *coordinate*, o *componenti* di $\mathbf{w}$ rispetto alla base $\{\mathbf{v_i}\}_1^k$.
Si osservi che la base di un sottospazio vettoriale $\mathcal{W} = \mathcal{S}(\mathbf{v_1}, \mathbf{v_2}, \ldots \mathbf{v_k})$ non è univocamente determinata.
Ad esempio il sottospazio $\mathcal{W} = \{\mathbf{x} \in \mathbb{R}^3 : x_1 + x_2 + x_3 = 0\}$ ammette come base $\{[-1,1,0]^T, [-1,0,1]^T\}$, ma anche $\{[1,-1,0]^T, [0,-1,1]^T\}$.
A questo proposito si dimostra il seguente teorema

> **Teorema 1.1.8.** *Tutte le basi di un sottospazio* $\mathcal{W} = \mathcal{S}(\mathbf{v_1}, \mathbf{v_2}, \cdots, \mathbf{v_k})$ *hanno lo stesso numero di elementi.*

Dato che il numero di elementi di una base di un sottospazio $\mathcal{W} = \mathcal{S}(\mathbf{v_1}, \mathbf{v_2}, \cdots, \mathbf{v_k})$ di $\mathcal{V}$ è una caratteristica del sottospazio stesso, si introduce la seguente:

> **Definizione 1.1.9.** Se $\mathcal{W} = \mathcal{S}(\mathbf{v_1}, \mathbf{v_2}, \cdots, \mathbf{v_k})$ ha una base costituita da r vettori, $r \leq k$, il numero intero r è detto *dimensione* di $\mathcal{W}$ ed è indicato con il simbolo $\dim(\mathcal{W})$.

Se $\mathcal{W}$ è un sottospazio vettoriale di $\mathcal{V}$ di dimensione r, allora ogni insieme contente più di r vettori estratti da $\mathcal{W}$ è un insieme di vettori linearmente dipendenti; in altri termini, la dimensione di un sottospazio vettoriale indica il massimo numero di vettori linearmente indipendenti che si possono estrarre da $\mathcal{W}$.
Uno spazio vettoriale $\mathcal{V}$ si dice *finitamente generato* se esistono k vettori di $\mathcal{V}$, $\mathbf{v_1}, \mathbf{v_2}, \cdots, \mathbf{v_k}$, tali che

$$S(\mathbf{v_1}, \mathbf{v_2}, \cdots, \mathbf{v_k}) = \mathcal{V}. \tag{1.1.3}$$

Per quanto detto in precedenza, ogni spazio vettoriale finitamente generato ammette una base e quindi nella (1.1.3) si può supporre che i vettori risultino linearmente indipendenti. In questo caso $\dim(\mathcal{V}) = k$.
Poiché

$$\mathbb{R}^n = \mathcal{S}(\mathbf{e_1}, \mathbf{e_2}, \cdots, \mathbf{e_n})$$

ed i versori fondamentali sono linearmente indipendenti, si può concludere che $\mathbb{R}^n$ è uno spazio vettoriale di dimensione n. L'insieme dei vettori $\mathbf{e_1}, \mathbf{e_2}, \cdots, \mathbf{e_n}$ viene detto *base canonica.*
Assegnato un sottospazio vettoriale $\mathcal{W} = \mathcal{S}(\mathbf{v_1}, \mathbf{v_2}, \cdots, \mathbf{v_k})$ di $\mathbb{R}^n$, i suoi elementi ammettono la seguente rappresentazione

$$\mathcal{W} = \{\mathbf{x} \in \mathbb{R}^n : \ \mathbf{x} = X\mathbf{a}, \quad \mathbf{a} \in \mathbb{R}^k\}.$$

dove X è la matrice di tipo (n, k) ottenuta accostando i k vettori $\mathbf{v_1}, \mathbf{v_2}, \cdots, \mathbf{v_k}$.
Esiste una importante relazione tra la dimensione del sottospazio $\mathcal{W}$ ed il rango $r(X)$ della matrice X. Si dimostra infatti che

$$\dim(\mathcal{W}) = r(X).$$

Inoltre, per determinare una base di $\mathcal{W}$ (la scelta non è unica, in generale come abbiamo già sottolineato) basta considerare i vettori che costituiscono le colonne di una delle sottomatrici quadrate non singolari di ordine pari a $r(X)$ estraibili da X.
In particolare se $k = n$ e $r(X) = n$, ossia $\det(X) \neq 0$, i vettori $\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ sono linearmente indipendenti e costituiscono una base per $\mathbb{R}^n$.

Esempio 1.1.10.
Sia A una matrice di tipo (m, n) con $r(A) = k$. Si può dimostrare che il nucleo di A, $N(A)$, è un sottospazio vettoriale di dimensione $n - r(A)$ mentre l'immagine di A, $\mathrm{Im}(A)$, è un sottospazio vettoriale di dimensione $r(A)$. Quindi

$$\dim(N(A)) + \dim(\mathrm{Im}(A)) = n.$$

1.2 Lo spazio euclideo $\mathbb{R}^n$

In questa sezione e nelle seguenti considereremo esclusivamente lo spazio vettoriale $\mathbb{R}^n$ e tratteremo i suoi elementi, come già anticipato, come vettori colonna.
In $\mathbb{R}^n$, oltre alle operazioni di somma e prodotto esterno introdotte nel precedente paragrafo, si può definire anche l'operazione di *prodotto scalare (o prodotto interno)* che associa ad una coppia di vettori un numero reale:

$$\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \quad (\mathbf{x}, \mathbf{y}) \to \sum_{i=1}^{n} x_i y_i$$

Il prodotto scalare tra i due vettori $\mathbf{x}$ e $\mathbf{y}$ verrà indicato con il simbolo $\langle \mathbf{x}, \mathbf{y} \rangle$.

Il prodotto scalare gode delle seguenti proprietà:

- $\langle \mathbf{x}, \mathbf{x} \rangle = \sum_{i=1}^{n} x_i^2 \geq 0, \quad \forall \mathbf{x} \in \mathbb{R}^n;$
- $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ se e solo se $\mathbf{x} = 0;$

- $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle, \quad \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$;
- $\langle \mathbf{x}, (\mathbf{y} + \mathbf{z}) \rangle = \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{z} \rangle, \quad \forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$;
- $\langle a\mathbf{x}, \mathbf{y} \rangle = a\langle \mathbf{x}, \mathbf{y} \rangle, \quad \forall a \in \mathbb{R}, \quad \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

Dato un vettore $\mathbf{x} \in \mathbb{R}^n$, si chiama *norma euclidea* di $\mathbf{x}$, e la si indica con il simbolo $||\mathbf{x}||$, il numero reale

$$||\mathbf{x}|| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = \sqrt{\sum_{i=1}^{n} x_i^2}.$$

In $\mathbb{R}^2$ e in $\mathbb{R}^3$, la norma euclidea di un vettore $\mathbf{x}$ rappresenta la lunghezza del segmento i cui estremi sono l'origine e il punto del piano o dello spazio le cui coordinate cartesiane sono le componenti del vettore $\mathbf{x}$.
La norma euclidea gode delle seguenti proprietà:

- $||\mathbf{x}|| \geq 0$, per ogni $\mathbf{x} \in \mathbb{R}^n$, $\quad ||\mathbf{x}|| = 0$ se e solo se $\mathbf{x} = \mathbf{0}$;
- $||a\mathbf{x}|| = |a|\, ||\mathbf{x}||$, per ogni $a \in \mathbb{R}$, $\mathbf{x} \in \mathbb{R}^n$;
- $||\mathbf{x} + \mathbf{y}|| \leq ||\mathbf{x}|| + ||\mathbf{y}||$, per ogni $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ (disuguaglianza triangolare);
- $|\langle \mathbf{x}, \mathbf{y} \rangle| \leq ||\mathbf{x}||\, ||\mathbf{y}||$, per ogni $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ (disuguaglianza di Cauchy–Schwarz).

Ogni funzione definita in $\mathbb{R}^n$ a valori non negativi che soddisfa le prime tre proprietà del precedente elenco, viene detta *norma*. Le seguenti funzioni rappresentano altre possibili norme in $\mathbb{R}^n$:

1) $\mathbf{x} \to \max\{|x_1|, |x_2|, \cdots, |x_n|\}$,

2) $\mathbf{x} \to \sum_{i=1}^{n} |x_i|$,

3) $\mathbf{x} \to (\sum_{i=1}^{n} |x_i|^p)^{\frac{1}{p}}, p = 1, 2, ...$

Nel seguito utilizzeremo esclusivamente la norma euclidea. Grazie alla disuguaglianza di Cauchy-Schwartz che essa soddisfa, si può introdurre il concetto di angolo θ tra due vettori non nulli $\mathbf{x}$ e $\mathbf{y}$ di $\mathbb{R}^n$ ponendo:

$$\cos(\theta) = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{||\mathbf{x}||\, ||\mathbf{y}||}, \quad 0 \leq \theta \leq \pi.$$

In particolare due vettori $\mathbf{x}$ e $\mathbf{y}$ che formano un angolo di $\frac{\pi}{2}$ vengono detti *ortogonali*. Poiché $\cos(\frac{\pi}{2}) = 0$, due vettori sono ortogonali se e solo se il loro prodotto scalare è nullo. Si dimostra facilmente il seguente risultato

Proposizione 1.2.1. *Due vettori non nulli ortogonali sono linearmente indipendenti.*

Dimostrazione. Dati due vettori $\mathbf{x}$ ed $\mathbf{y}$ non nulli e ortogonali supponiamo che

$$\alpha\mathbf{x} + \beta\mathbf{y} = \mathbf{0}.$$

Si ha allora

$$\langle \mathbf{x}, \alpha\mathbf{x} + \beta\mathbf{y} \rangle = \alpha\langle \mathbf{x}, \mathbf{x} \rangle = 0$$

Essendo per ipotesi $\mathbf{x} \neq \mathbf{0}$, si ha necessariamente $\alpha = 0$. Analogamente

$$\langle \mathbf{y}, \alpha\mathbf{x} + \beta\mathbf{y} \rangle = \beta\langle \mathbf{y}, \mathbf{y} \rangle = 0$$

e quindi anche $\beta = 0$. Questo prova l'indipendenza lineare dei due vettori. ■

Si noti, invece, che vettori linearmente indipendenti non sono necessariamente ortogonali fra loro. Comunque, dati k vettori $\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_k}$ linearmente indipendenti in $\mathbb{R}^n$, è sempre possibile costruire un sistema di vettori a due a due ortogonali $\mathbf{y_1}, \mathbf{y_2}, \ldots, \mathbf{y_k}$ tali che

$$\mathcal{S}(\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_k}) = \mathcal{S}(\mathbf{y_1}, \mathbf{y_2}, \ldots, \mathbf{y_k})$$

Tale procedimento è noto come **processo di ortogonalizzazione di Gram-Schmidt**. Vediamo come si può procedere. Inizialmente si pone

$$\mathbf{y_1} = \mathbf{x_1},$$

e si cerca un elemento $\mathbf{y_2} \in \mathcal{S}(\mathbf{x_1}, \mathbf{x_2})$ che sia ortogonale a $\mathbf{y_1}$. Detto

$$\mathbf{y_2} = \mathbf{x_2} + a_1\mathbf{y_1},$$

si determina a_1 in modo tale che $\langle \mathbf{y_1}, \mathbf{y_2} \rangle = 0$. Con semplici passaggi algebrici si verifica che

$$a_1 = -\frac{\langle \mathbf{y_1}, \mathbf{x_2} \rangle}{\langle \mathbf{y_1}, \mathbf{y_1} \rangle}.$$

Si osservi che l'indipendenza lineare di $\mathbf{x_1}$ e $\mathbf{x_2}$ garantisce che $\mathbf{y_2} \neq \mathbf{0}$.
Il procedimento è concluso se $k = 2$. In generale, per $k \geq 2$, si prova che la scelta

$$\mathbf{y_p} = \mathbf{x_p} - \sum_{j=1}^{p-1} \frac{\langle \mathbf{y_j}, \mathbf{x_p} \rangle}{\langle \mathbf{y_j}, \mathbf{y_j} \rangle}\mathbf{y_j}, \quad p = 2, 3, \ldots, k$$

fornisce un sistema di vettori a due a due ortogonali con la proprietà richiesta.

Assegnato un vettore non nullo $\mathbf{a} \in \mathbb{R}^n$, l'insieme dei vettori ad esso ortogonali

$$\Pi_{\mathbf{a}} = \{\mathbf{x} \in \mathbb{R}^n : \langle \mathbf{a}, \mathbf{x} \rangle = 0\}$$

è un sottospazio vettoriale di dimensione $(n-1)$ detto *iperpiano passante per l'origine.* Si dice invece *iperpiano passante per* $\mathbf{x_0}$ l'insieme

$$\Pi_{\mathbf{a}}(x_0) = \{\mathbf{x} \in \mathbb{R}^n : \langle \mathbf{a}, \mathbf{x} - \mathbf{x_0} \rangle = 0\} = \{\mathbf{x} \in \mathbb{R}^n : \langle \mathbf{a}, \mathbf{x} \rangle = \alpha\}$$

dove $\alpha = \langle \mathbf{a}, \mathbf{x_0} \rangle$.
Ricordiamo infine che un vettore non nullo $\mathbf{x} \in \mathbb{R}^n$ è detto *normalizzato* se $||\mathbf{x}|| = 1$. Un vettore normalizzato viene anche detto *versore*. L'operazione di *normalizzazione* consente di costruire, partendo da un vettore non nullo $\mathbf{x} \in \mathbb{R}^n$, il vettore normalizzato di norma euclidea unitaria

$$\mathbf{x}^* = \frac{\mathbf{x}}{||\mathbf{x}||}$$

che mantiene la direzione ed il verso del vettore $\mathbf{x}$ originario.

1.3 Autovalori ed autovettori

Sia A una matrice quadrata di ordine n ad elementi reali.

Definizione 1.3.1. Un vettore non nullo $\mathbf{v} \in \mathbb{R}^n$ è detto *autovettore* di A corrispondente all'*autovalore* λ se

$$A\mathbf{v} = \lambda \mathbf{v},$$

ossia se risulta una soluzione non nulla del sistema lineare omogeneo

$$(A - \lambda I)\mathbf{v} = \mathbf{0}. \tag{1.3.1}$$

In base alla teoria dei sistemi lineari, il sistema (1.3.1) ammette soluzioni non nulle se e solo se il rango $r(A)$ della matrice A è inferiore ad n, ossia se e solo se

$$\det(A - \lambda I) = 0. \tag{1.3.2}$$

L'equazione (1.3.2) viene detta *equazione caratteristica* di A ed il suo primo membro è un polinomio di grado n in λ detto *polinomio caratteristico* di A:

$$p(\lambda) = p_0 + p_1\lambda + p_2\lambda^2 + \cdots + p_n\lambda^n.$$

In particolare, $p_0 = p(0) = \det(A)$, $p_1 = (-1)^{n-1}\mathrm{tr}(A)$ (2) e $p_n = (-1)^n$. Per $n = 2$, il polinomio caratteristico diviene

$$p(\lambda) = \lambda^2 - \mathrm{tr}(A)\lambda + \det(A).$$

Per determinare gli autovettori di una matrice, si devono quindi prima determinare le radici dell'equazione caratteristica, ossia gli *autovalori* di A. Poiché l'insieme dei numeri reali non è un campo algebricamente chiuso, non è detto che tutti gli autovalori di A siano reali (3).
Le eventuali radici complesse dell'equazione caratteristica sono comunque sempre in numero pari, poiché se un numero complesso è radice dell'equazione (1.3.2) lo è anche il suo complesso coniugato.

(2) Con $\mathrm{tr}(A)$ si indica la traccia di A, ossia la somma dei suoi elementi diagonali.
(3) Un caso significativo in cui questo invece si verifica si ha quando la matrice A è simmetrica