1. Sistemi lineari.

Definizione. Un sistema lineare di m equazioni in n incognite è un sistema di equazioni della forma

$$\begin{cases} a_{11}x_1 & +a_{12}x_2 & +\ldots & +a_{1n}x_n & = b_1 \\ a_{21}x_1 & +a_{22}x_2 & +\ldots & +a_{2n}x_n & = b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1}x_1 & +a_{m2}x_2 & +\ldots & +a_{mn}x_n & = b_m, \end{cases}$$

dove gli a_{ij} sono numeri assegnati detti *coefficienti*, i b_j sono numeri assegnati detti *termini noti* e le incognite $x_1, \ldots, x_n$ appaiono al primo grado.

Un sistema omogeneo è un sistema i cui termini noti sono nulli: $b_1 = \ldots = b_m = 0$.

Definizione. Una soluzione del sistema

$$\begin{cases} a_{11}x_1 & +a_{12}x_2 & +\ldots & +a_{1n}x_n & = b_1 \\ a_{21}x_1 & +a_{22}x_2 & +\ldots & +a_{2n}x_n & = b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1}x_1 & +a_{m2}x_2 & +\ldots & +a_{mn}x_n & = b_m, \end{cases}$$

è un' ennupla

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}$$

che soddisfa tutte le equazioni del sistema. La *soluzione generale del sistema* è l'insieme di tutte le soluzioni del sistema.

Esempio. Le triple

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 5 \\ 2 \\ 0 \end{pmatrix}$$

sono soluzioni del sistema di due equazioni in tre incognite

$$\begin{cases} x_1 - 2x_2 + x_3 = 1 \\ 2x_1 - 4x_2 = 2. \end{cases}$$

• Un sistema che ammette soluzioni si dice *compatibile*. Vedremo in seguito che un sistema compatibile ammette un'unica soluzione oppure ammette infinite soluzioni. Un sistema che non ammette soluzioni si dice *incompatibile*.

• Un sistema lineare omogeneo

$$\begin{cases} a_{11}x_1 & +a_{12}x_2 & +\ldots & +a_{1n}x_n & = 0 \\ a_{21}x_1 & +a_{22}x_2 & +\ldots & +a_{2n}x_n & = 0 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1}x_1 & +a_{m2}x_2 & +\ldots & +a_{mn}x_n & = 0 \end{cases}$$

è sempre compatibile: infatti ammette almeno la soluzione nulla

$$\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

• Un sistema lineare non omogeneo può essere incompatibile, come ad esempio:

$$\begin{cases} x_1 + 2x_2 + x_3 - x_4 = 1 \\ x_1 + 2x_2 + x_3 - x_4 = 2. \end{cases}$$

Definizione. Due sistemi lineari si dicono equivalenti se hanno esattamente le stesse soluzioni.

Operazioni elementari sulle equazioni di un sistema:
- scambiare tra loro due equazioni,
- moltiplicare un'equazione per uno scalare non nullo,
- sostituire un'equazione con quella ottenuta sommando ad essa un multiplo di un'altra equazione.

Proposizione. *Effettuando una qualsiasi operazione elementare sulle equazioni di un sistema lineare dato si ottiene un sistema ad esso equivalente.*

Notazioni:
Sia dato un sistema lineare di m equazioni in n incognite

$$\begin{cases} a_{11}x_1 & +a_{12}x_2 & +\dots & +a_{1n}x_n & = b_1 \\ a_{21}x_1 & +a_{22}x_2 & +\dots & +a_{2n}x_n & = b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1}x_1 & +a_{m2}x_2 & +\dots & +a_{mn}x_n & = b_m, \end{cases}$$

Matrice dei coefficienti:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

Vettore dei termini noti:

$$\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

Matrice completa:

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right)$$

Definizione. Un sistema lineare si dice *a scala* se la matrice dei coefficienti è della forma

$$\begin{pmatrix} 0 & \cdots & 0 & a_{1k_1} & * & * & * & * & \cdots & \cdots & * \\ \vdots & & \vdots & 0 & \cdots & 0 & a_{2k_2} & * & \cdots & \cdots & * \\ \vdots & & \vdots & \vdots & & \vdots & 0 & \cdots & \cdots & \cdots & * \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & \ddots & & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & a_{rk_r} & \cdots & * \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 \end{pmatrix}$$

dove i coefficienti a_{jk_j}, $j = 1, \ldots, r$ sono diversi da zero, ed i coefficienti $*$ sono arbitrari.

• Un sistema a scala gode delle seguenti proprietà:

- tutte le sue equazioni sono indipendenti (un sistema, ottenuto da un sistema a scala togliendo un'equazione, non è mai equivalente al sistema di partenza);
- è incompatibile se e solo se contiene almeno un'equazione del tipo $0 = b$, con $b \neq 0$.
- le soluzioni di un sistema lineare *a scala compatibile* si ottengono per sostituzione da sotto in su.

• Dato un sistema lineare, è possibile ottenere un sistema a scala ad esso equivalente, mediante una successione di operazioni elementari. Questo procedimento si chiama "metodo di eliminazione di Gauss". ("elimina" le equazioni superflue) (Apostol, sezione 4.18).

Lemma. *Sia*

$$\begin{cases} a_{11}x_1 & +a_{12}x_2 & + \ldots & +a_{1n}x_n & = 0 \\ a_{21}x_1 & +a_{22}x_2 & + \ldots & +a_{2n}x_n & = 0 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1}x_1 & +a_{m2}x_2 & + \ldots & +a_{mn}x_n & = 0 \end{cases}$$

un sistema lineare omogeneo di m equazioni in n incognite. Se $n > m$, allora ammette soluzioni non nulle.

Dimostrazione. Per il metodo di eliminazione di Gauss, il sistema è equivalente a un sistema omogeneo *a scala* con matrice dei coefficienti uguale a

$$\begin{pmatrix} 0 & \cdots & 0 & a_{1k_1} & * & * & * & * & \cdots & \cdots & * \\ \vdots & & \vdots & 0 & \cdots & 0 & a_{2k_2} & * & \cdots & \cdots & * \\ \vdots & & \vdots & \vdots & & \vdots & 0 & \cdots & \cdots & \cdots & * \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & \ddots & & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & a_{rk_r} & \cdots & * \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 \end{pmatrix}$$

dove $a_{jk_j} \neq 0$, $j = 1, \ldots, r$ mentre gli elementi $*$ sono arbitrari. Ci sono r equazioni non nulle, per un certo $r \leq m$, e se $n > m$, allora $n - r \geq n - m > 0$. Si esprimono le incognite $x_{k_1}, x_{k_2}, \ldots, x_{k_r}$, corrispondenti agli elementi $a_{jk_j} \neq 0$, in funzione delle rimanenti $n - r$. Le soluzioni del sistema dipendono così da $n - r$ parametri liberi. In particolare, esistono soluzioni non banali. Questo completa la dimostrazione.

• Dato un sistema lineare di n equazioni in n incognite a scala e compatibile, esso ammette un'unica soluzione se e solo se la matrice dei coefficienti ha tutti gli elementi sulla diagonale principale diversi da zero.

• Le soluzioni di un sistema lineare *a scala compatibile* dipendono da un numero di parametri liberi uguale al numero delle incognite meno il numero delle equazioni non banali del sistema.

2. Spazi vettoriali e sottospazi.
Indichiamo con $\mathbf{R}^n$ l'insieme delle ennuple reali

$$\mathbf{R}^n = \left\{ X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \mid x_1, \ldots, x_n \in \mathbf{R} \right\}.$$

A volte identifichiamo $\mathbf{R}^n$ con i vettori uscenti dall'origine, ossia i segmenti orientati aventi il primo estremo nell'origine $O = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ e il secondo nel punto $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$.

• In $\mathbf{R}^n$ si possono definire la *somma fra vettori* ed il *prodotto di un vettore per un numero reale* nel modo seguente:

$$\text{dati} \quad X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \; Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \in \mathbf{R}^n, \quad \lambda \in \mathbf{R},$$

si definiscono

$$X + Y := \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix} \qquad \lambda X := \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \\ \vdots \\ \lambda x_n \end{pmatrix}. \tag{1}$$

In $\mathbf{R}^2$, il vettore $X+Y$ coincide con il vettore ottenuto sommando X e Y con la regola del parallelogramma. Questa interpretazione geometrica in realtà vale in generale. Dati due vettori $X,\ Y \in \mathbf{R}^n$, esiste infatti un piano π passante per $X,\ Y$ e per l'origine O, e il vettore $X+Y$ coincide col vettore ottenuto applicando la regola del parallelogramma ad X e Y nel piano π.

Si verifica facilmente che le operazioni di somma e prodotto per uno scalare su $\mathbf{R}^n$ hanno una serie di proprietà, che derivano essenzialmente dalle analoghe proprietà della somma e del prodotto fra numeri reali:

1. (commutatività della somma) $\forall X, Y \in \mathbf{R}^n \quad X + Y = Y + X$;
2. (associatività della somma) $\forall X, Y, Z \in \mathbf{R}^n \quad (X + Y) + Z = X + (Y + Z)$;
3. (elemento neutro per la somma) $\exists O \in \mathbf{R}^n : \quad O + X = X + O = X, \quad \forall X \in \mathbf{R}^n$, precisamente

$$O = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix};$$

4. (opposto per la somma) $\forall X \in \mathbf{R}^n, \quad \exists (-X) : \quad X + (-X) = (-X) + X = O$, precisamente

$$(-X) = \begin{pmatrix} -x_1 \\ \vdots \\ -x_n \end{pmatrix};$$

5. (associatività del prodotto) $\forall \lambda, \mu \in \mathbf{R},\ \forall X \in \mathbf{R}^n \quad (\lambda\mu)X = \lambda(\mu X)$;
6. (distributività del prodotto) $\forall \lambda, \mu \in \mathbf{R},\ \forall X \in \mathbf{R}^n \quad (\lambda + \mu)X = \lambda X + \mu X$;
7. (distributività del prodotto) $\forall \lambda \in \mathbf{R},\ \forall X, Y \in \mathbf{R}^n \quad \lambda(X + Y) = \lambda X + \lambda Y$;
8. (elemento neutro per il prodotto) $1X = X, \quad \forall X \in \mathbf{R}^n$

Esercizio. Far vedere che vale il principio dell'annullamento

$$\lambda X = O \quad \Rightarrow \quad \lambda = 0 \quad \text{oppure} \quad X = O.$$

Le proprietà $1-8$ sono esattamente gli assiomi che definiscono uno spazio vettoriale reale.

Definizione. Uno spazio vettoriale reale V è un insieme su cui sono definite un'operazione di somma

$$V \times V \longrightarrow V, \quad (v,w) \mapsto v+w$$

e di moltiplicazione per uno scalare reale

$$\mathbf{R} \times V \longrightarrow V, \quad (\lambda, v) \mapsto \lambda v$$

che soddisfano le proprietà

1. (commutatività della somma) $\forall u, v \in V \quad u+v = v+w$;
2. (associatività della somma) $\forall u, v, w \in V \quad (u+v)+w = u+(v+w)$;
3. (elemento neutro per la somma) $\exists O \in V: \quad O+v = v+O = v, \quad \forall v \in V$;
4. (opposto per la somma) $\forall v \in V, \quad \exists (-v): \quad (-v)+v = v+(-v) = O$;
5. (associatività del prodotto) $\forall \lambda, \mu \in \mathbf{R}, \ \forall v \in V \quad (\lambda\mu)v = \lambda(\mu v)$;
6. (distributività del prodotto) $\forall \lambda, \mu \in \mathbf{R}, \ \forall v \in V \quad (\lambda+\mu)v = \lambda v + \mu v$;
7. (distributività del prodotto) $\forall \lambda \in \mathbf{R}, \ \forall v \in V \quad \lambda(v+w) = \lambda v + \lambda w$;
8. (elemento neutro per il prodotto) $1v = v, \quad \forall v \in V$

Dunque $\mathbf{R}^n$, con la somma fra vettori e il prodotto di un vettore per uno scalare definiti in (1), è un esempio di spazio vettoriale reale.

Esempio. Un esempio analogo è dato dalle matrici $m \times n$ a coefficienti reali

$$\mathcal{M}(m,n,\mathbf{R}) = \{M = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \mid a_{ij} \in \mathbf{R}, \quad i = 1, \dots, m, \ j = 1, \dots, n\},$$

dove la somma fra matrici e il prodotto di una matrice per uno scalare sono definiti nel modo seguente

$$M + N := \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \dots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \dots & a_{2n}+b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \dots & a_{mn}+b_{mn} \end{pmatrix}$$

dove $M = \{a_{ij}\}$, $N = \{b_{ij}\} \in \mathcal{M}(m,n,\mathbf{R})$, e

$$\lambda M := \begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \dots & \lambda a_{1n} \\ \lambda a_{21} & \lambda a_{22} & \dots & \lambda a_{2n} \\ \vdots & \vdots & & \vdots \\ \lambda a_{m1} & \lambda a_{m2} & \dots & \lambda a_{mn} \end{pmatrix},$$

dove $M = \{b_{ij}\} \in \mathcal{M}(m,n,\mathbf{R})$, $\lambda \in \mathbf{R}$.

Molti esempi di spazi vettoriali vengono dall'analisi e sono dati da spazi di funzioni.

Esempio. La famiglia delle funzioni di una variabile reale a valori reali

$$F(\mathbf{R},\mathbf{R}) = \{f \colon \mathbf{R} \longrightarrow \mathbf{R}\},$$

con la somma e il prodotto per uno scalare reale definiti da

$$(f+g)(x) := f(x) + g(x) \qquad (\lambda f)(x) := \lambda f(x), \qquad f, g \in F(\mathbf{R}, \mathbf{R}),\ \lambda \in \mathbf{R},$$

è uno spazio vettoriale reale.

• Gli spazi vettoriali si presentano per lo piú come *sottospazi* di spazi vettoriali dati, ossia come sottoinsiemi non vuoti $S \subset V$ di uno spazio vettoriale, *chiusi rispetto alla somma e al prodotto per uno scalare*:
(i) $s + t \in S, \quad \forall s, t \in S$
(ii) $\lambda s \in S, \quad \forall \lambda \in \mathbf{R},\ \forall s \in S.$

Osservazione.
(i) Sia V uno spazio vettoriale. L'elemento neutro per la somma $\{O\}$ é un sottospazio vettoriale di V (detto sottospazio banale): infatti, $O + O = O$, e per ogni $\lambda \in \mathbf{R}$ si ha $\lambda O = O$.
(ii) Ogni sottospazio vettoriale $S \subset V$ contiene l'elemento neutro O (per (ii), con $\lambda = 0$).

Esempio. Le seguenti famiglie di funzioni sono tutti sottospazi vettoriali di $F(\mathbf{R}, \mathbf{R})$.

$$C^0(\mathbf{R}, \mathbf{R}) = \{f\colon \mathbf{R} \longrightarrow \mathbf{R},\ \text{continue}\}.$$

Infatti, se f, g sono funzioni continue, $f + g$ è una funzione continua e per ogni $\lambda \in \mathbf{R}$, λf è una funzione continua. Analogamente,

$$C^1(\mathbf{R}, \mathbf{R}) = \{f\colon \mathbf{R} \longrightarrow \mathbf{R},\ \text{con derivata continua}\}$$

$$C^+(\mathbf{R}, \mathbf{R}) = \{f\colon \mathbf{R} \longrightarrow \mathbf{R},\ \text{crescenti}\}$$

$$\mathbf{R}[x] = \{p(x) = a_0 + a_1 x + \ldots + a_n x^n,\ n \geq 0\}, \quad \text{polinomi}$$

$$\mathbf{R}[x]_m = \{p(x) = a_0 + a_1 x + \ldots + a_m x^m,\ m \geq 0\}, \quad \text{polinomi di grado} \leq m$$

Inoltre, valgono le seguenti inclusioni fra sottospazi

$$\mathbf{R}[x]_m \subset \mathbf{R}[x] \subset C^1(\mathbf{R}, \mathbf{R}) \subset C^0(\mathbf{R}, \mathbf{R}) \subset F(\mathbf{R}, \mathbf{R})$$

Definizione. Siano $v_1, \ldots, v_k$ elementi in uno spazio vettoriale V. L'insieme delle combinazioni lineari di $v_1, \ldots, v_k$ (indicato $\text{span}\{v_1, \ldots, v_k\}$) è per definizione

$$\text{span}\{v_1, \ldots, v_k\} = \{\lambda_1 v_1 + \ldots + \lambda_k v_k \mid \lambda_1, \ldots, \lambda_k \in \mathbf{R}\}.$$

Proposizione. *L'insieme delle combinazioni lineari di un insieme di elementi $v_1, \ldots, v_k$ in uno spazio vettoriale V è un sottospazio vettoriale di V.*

Dimostrazione. Siano $X = \lambda_1 v_1 + \ldots + \lambda_k v_k$, $Y = \mu_1 v_1 + \ldots + \mu_k v_k$ elementi arbitrari di $\text{span}\{v_1, \ldots, v_k\}$. Facciamo vedere che $X + Y \in \text{span}\{v_1, \ldots, v_k\}$. Inoltre, per ogni $\alpha \in R$, anche $\alpha X \in \text{span}\{v_1, \ldots, v_k\}$. Infatti

$$\begin{aligned} X + Y &= (\lambda_1 v_1 + \ldots + \lambda_k v_k) + (\mu_1 v_1 + \ldots + \mu_k v_k), \\ &= (\lambda_1 + \mu_1) v_1 + \ldots + (\lambda_k + \mu_k) v_k \in \text{span}\{v_1, \ldots, v_k\}; \end{aligned}$$

$$\alpha X = \alpha(\lambda_1 v_1 + \ldots + \lambda_k v_k) = \alpha\lambda_1 v_1 + \ldots + \alpha\lambda_k v_k \in \text{span}\{v_1, \ldots, v_k\}.$$

Esercizio.

(i) In $\mathbf{R}^2$, determinare e disegnare

$$\text{span}\{\begin{pmatrix}0\\0\end{pmatrix}\},\ \text{span}\{\begin{pmatrix}2\\2\end{pmatrix}\},\ \text{span}\{\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}2\\2\end{pmatrix}\},\ \text{span}\{\begin{pmatrix}1\\1\end{pmatrix},\begin{pmatrix}2\\2\end{pmatrix}\},\ \text{span}\{\begin{pmatrix}1\\-1\end{pmatrix},\begin{pmatrix}2\\2\end{pmatrix}\}.$$

(ii) Determinare se $\begin{pmatrix}3\\2\end{pmatrix} \in \text{span}\{\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}2\\2\end{pmatrix}\}$.

(iii) Determinare se $\begin{pmatrix}3\\2\end{pmatrix} \in \text{span}\{\begin{pmatrix}1\\-1\end{pmatrix},\begin{pmatrix}2\\2\end{pmatrix}\}$.

(iv) Determinare se $\begin{pmatrix}0\\0\end{pmatrix} \in \text{span}\{\begin{pmatrix}1\\-1\end{pmatrix},\begin{pmatrix}2\\2\end{pmatrix}\}$.

(v) Far vedere che $\begin{pmatrix}3\\2\end{pmatrix} \notin \text{span}\{\begin{pmatrix}1\\1\end{pmatrix},\begin{pmatrix}2\\2\end{pmatrix}\}$.

(vi) Far vedere che un arbitrario vettore in $X \in \mathbf{R}^2$ appartiene a $\text{span}\{\begin{pmatrix}1\\-1\end{pmatrix},\begin{pmatrix}1\\2\end{pmatrix}\}$.

Esercizio.

(i) In $\mathbf{R}^3$, determinare i sottospazi $\quad \text{span}\{\begin{pmatrix}0\\0\\0\end{pmatrix}\},\ \text{span}\{\begin{pmatrix}2\\2\\0\end{pmatrix}\},$

$$\text{span}\{\begin{pmatrix}0\\0\\0\end{pmatrix},\begin{pmatrix}2\\2\\1\end{pmatrix}\},\ \text{span}\{\begin{pmatrix}1\\1\\1\end{pmatrix},\begin{pmatrix}2\\2\\2\end{pmatrix}\},\ \text{span}\{\begin{pmatrix}1\\-1\\0\end{pmatrix},\begin{pmatrix}2\\2\\0\end{pmatrix}\}.$$

(ii) Determinare se $\begin{pmatrix}3\\2\\1\end{pmatrix} \in \text{span}\{\begin{pmatrix}0\\0\\1\end{pmatrix},\begin{pmatrix}2\\2\\0\end{pmatrix}\}$.

(iii) Determinare se $\begin{pmatrix}3\\2\\1\end{pmatrix} \in \text{span}\{\begin{pmatrix}1\\-1\\0\end{pmatrix},\begin{pmatrix}2\\2\\0\end{pmatrix},\begin{pmatrix}0\\0\\1\end{pmatrix}\}$.

(iv) Determinare se $\begin{pmatrix}0\\0\\0\end{pmatrix} \in \text{span}\{\begin{pmatrix}1\\-1\\0\end{pmatrix},\begin{pmatrix}2\\2\\2\end{pmatrix}\}$.

(v) Far vedere che $\begin{pmatrix}3\\2\\1\end{pmatrix} \notin \text{span}\{\begin{pmatrix}1\\1\\1\end{pmatrix},\begin{pmatrix}2\\2\\1\end{pmatrix},\begin{pmatrix}0\\0\\1\end{pmatrix}\}$.

(vi) Far vedere che un arbitrario vettore in $\mathbf{R}^3$ appartiene a $\text{span}\{\begin{pmatrix}1\\-1\\1\end{pmatrix},\begin{pmatrix}1\\2\\1\end{pmatrix},\begin{pmatrix}3\\0\\1\end{pmatrix}\}$.

Proposizione. *Le soluzioni di un sistema lineare omogeneo di m equazioni in n incognite sono un sottospazio vettoriale di $\mathbf{R}^n$.*

Dimostrazione. Siano $S = \begin{pmatrix}s_1\\ \vdots\\ s_n\end{pmatrix}$ e $T = \begin{pmatrix}t_1\\ \vdots\\ t_n\end{pmatrix}$ due soluzioni di un sistema lineare omogeneo di m equazioni in n incognite. Facciamo vedere che

$$S+T = \begin{pmatrix}s_1+t_1\\ \vdots\\ s_n+t_n\end{pmatrix} \qquad \text{e} \qquad \lambda S = \begin{pmatrix}\lambda s_1\\ \vdots\\ \lambda s_n\end{pmatrix},\quad \forall\lambda \in \mathbf{R},$$

sono ancora soluzioni del sistema. Sia

$$a_{j1}x_1 + \ldots + a_{jn}x_n = 0, \qquad j \in \{1, \ldots, m\},$$

una equazione qualunque del sistema. Per ipotesi,

$$a_{j1}s_1 + \ldots + a_{jn}s_n = a_{j1}t_1 + \ldots + a_{jn}t_n = 0.$$

Sostituendo $S + T$ nell'equazione, troviamo

$$a_{j1}(s_1 + t_1) + \ldots + a_{jn}(s_n + t_n) = (a_{j1}s_1 + \ldots + a_{jn}s_n) + (a_{j1}t_1 + \ldots + a_{jn}t_n) = 0 + 0 = 0.$$

Similmente, sostituendo nell'equazione λS troviamo

$$a_{j1}\lambda s_1 + \ldots + a_{jn}\lambda s_n = \lambda(a_{j1}s_1 + \ldots + a_{jn}s_n) = \lambda 0 = 0,$$

come richiesto.

Osservazione. Osserviamo che le soluzioni di un sistema lineare di m equazioni in n incognite formano un sottospazio vettoriale di $\mathbf{R}^n$ se e solo se il sistema è omogeneo: infatti $O = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$ non è mai soluzione di un sistema non omogeneo, mentre un sottospazio vettoriale deve sempre contenere l'elemento O.

3. Dipendenza e indipendenza lineare.

Consideriamo il sottospazio delle combinazioni lineari di un insieme di elementi $\{v_1, \ldots, v_k\}$ in uno spazio vettoriale V

$$\mathrm{span}\{v_1, \ldots, v_k\} = \{\lambda_1 v_1 + \ldots + \lambda_k v_k \mid \lambda_i \in \mathbf{R}\}.$$

Osserviamo subito che:

(i) $\mathrm{span}\{v_1, \ldots, v_k\} = \mathrm{span}\{v_1, \ldots, v_k, O\}$;

(ii) Se l'elemento v_k è combinazione lineare degli elementi $v_1, \ldots, v_{k-1}$, allora

$$\mathrm{span}\{v_1, \ldots, v_k\} = \mathrm{span}\{v_1, \ldots, v_{k-1}\}.$$

• *Quali condizioni devono soddisfare gli elementi* $\{v_1, \ldots, v_k\}$ *di* V*, affinché il sottospazio* $\mathrm{span}\{v_1, \ldots, v_k\}$ *risulti espresso nel modo piú "efficiente" possibile?*

• *Qual è il numero minimo di elementi necessario a generare un dato sottospazio ?*

Definizione. Dati $v_1, \ldots, v_k$ in uno spazio vettoriale V, si dicono linearmente indipendenti se

$$\lambda_1 v_1 + \ldots + \lambda_k v_k = O \quad \Leftrightarrow \lambda_1 = \ldots = \lambda_k = 0.$$

In altre parole, nessuna combinazione lineare di $v_1, \ldots, v_k$, a coefficienti non tutti nulli, può dare il vettore $O \in V$.

Esempio.

(i) Un elemento $\{v_1\}$ è linearmente indipendente se e solo se è diverso dal vettore nullo O. Infatti $\lambda v_1 = O$ implica $\lambda = 0$ se e solo se $v_1 \neq O$.

(ii) Se $v_1, \ldots, v_k$ sono linearmente indipendenti, allora sono tutti non nulli. Altrimenti, se fosse ad esempio $v_k = O$,

$$0v_1 + \ldots + 0v_{k-1} + \lambda_k v_k = 0v_1 + \ldots + 0v_{k-1} + \lambda_k O = O, \quad \lambda_k \neq 0,$$

sarebbe una combinazione lineare di $v_1, \ldots, v_k$, a coefficienti non tutti nulli, con risultato il vettore nullo.

(iii) Due elementi (non nulli) $\{v_1, v_2\}$ sono linearmente indipendenti se e solo se non sono uno multiplo dell'altro, ossia non esiste $\lambda \neq 0$ tale che $v_1 = \lambda v_2$. Geometricamente, se e solo se $\{v_1, v_2\}$ non stanno sulla stessa retta.

(iii) Tre elementi (non nulli) $\{v_1, v_2, v_3\}$ sono linearmente indipendenti se e solo se nessuno di essi è come combinazione lineare dei due rimanenti, ossia non esistono λ, μ (non entrambi nulli) tali che $v_3 = \lambda v_1 + \mu v_2$. Geometricamente, se e solo se $\{v_1, v_2, v_3\}$ non stanno sullo stesso piano.

Esempio.

(i) Il vettore $\begin{pmatrix} 0 \\ 0 \end{pmatrix} \in \mathbf{R}^2$ non è linearmente indipendente.

(ii) I vettori $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} \in \mathbf{R}^3$ non sono linearmente indipendenti.

(iii) I vettori $\begin{pmatrix} 2 \\ 0 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} \in \mathbf{R}^3$ non sono linearmente indipendenti, mentre i vettori $\begin{pmatrix} 2 \\ 0 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}$ lo sono.

(iv) I vettori $\begin{pmatrix} 2 \\ 0 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 3 \\ 6 \end{pmatrix} \in \mathbf{R}^3$ non sono linearmente indipendenti, mentre i vettori $\begin{pmatrix} 2 \\ 0 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ lo sono.

In generale, si ha che:

Proposizione. *Gli elementi (non nulli) $v_1, \ldots, v_k \in V$ sono linearmente indipendenti se e solo nessuno di essi può essere espresso come combinazione lineare dei rimanenti. Equivale a dire che gli elementi (non nulli) $v_1, \ldots, v_k \in V$ sono linearmente dipendenti se e solo almeno uno di essi può essere espresso come combinazione lineare dei rimanenti.*

Dimostrazione. Supponiamo che $v_k = \alpha_1 v_1 + \ldots + \alpha_{k-1} v_{k-1}$, con $\alpha_j \in \mathbf{R}$. Allora

$$\alpha_1 v_1 + \ldots + \alpha_{k-1} v_{k-1} - v_k = O$$

è una combinazione lineare di $v_1, \ldots, v_k$, a coefficienti non tutti nulli, che dà il vettore $O \in V$. Di conseguenza $v_1, \ldots, v_k$ non sono linearmente indipendenti.

Viceversa supponiamo che $v_1, \ldots, v_k$ siano linearmente dipendenti e sia

$$\alpha_1 v_1 + \ldots + \alpha_{k-1} v_{k-1} + \alpha_k v_k = O$$

una combinazione lineare di $v_1, \ldots, v_k$, a coefficienti non tutti nulli, che dà il vettore $O \in V$. Supponiamo per esempio che $\alpha_k \neq 0$. Allora v_k può essere espresso come combinazione lineare dei rimanenti vettori

$$v_k = -\frac{1}{\alpha_k}(\alpha_1 v_1 + \ldots + \alpha_{k-1} v_{k-1}).$$

• *Il sottospazio* $\mathrm{span}\{v_1, \ldots, v_k\} \subset V$ *è espresso nel modo più "efficiente" possibile se e solo se gli elementi $v_1, \ldots, v_k$ sono linearmente indipendenti.* In questo caso, e solo in questo caso, il sottospazio generato da un qualunque sottoinsieme proprio di $\{v_1, \ldots, v_k\}$ è un sottospazio proprio di $\mathrm{span}\{v_1, \ldots, v_k\}$.

Esercizio. Se $v_1, \ldots, v_k$ sono linearmente indipendenti, allora ogni loro sottoinsieme è formato da elementi linearmente indipendenti.

Esercizio. Le righe non nulle di una matrice S di ordine $m \times n$ *a scala* sono vettori linearmente indipendenti in $\mathbf{R}^n$.

Osservazione. Ricordiamo il metodo di eliminazione di Gauss sulle righe $R_j,\ j = 1, \ldots, m$ di una matrice S di ordine $m \times n$. Esso consiste in una successione di operazioni elementari quali
- scambiare tra loro due righe,
- moltiplicare una riga per uno scalare non nullo,
- sostituire una riga con quella ottenuta sommando ad essa un multiplo di un'altra riga.

Ad ogni passo, il sottospazio

$$\mathrm{span}\{R_1, \ldots, R_m\} \subset \mathbf{R}^n$$

generato dalle righe della matrice rimane inalterato. Di conseguenza, le righe non nulle della matrice a scala S ottenuta da M mediante l'eliminazione di Gauss (per righe) costituiscono un insieme di generatori linearmente indipendenti di $\mathrm{span}\{R_1, \ldots, R_m\}$ in $\mathbf{R}^n$.

Possiamo usare l'osservazione precedente per risolvere il seguente problema:
Dati vettori $X_1, \ldots, X_k$ in $\mathbf{R}^n$, determinare vettori $Y_1, \ldots, Y_h$, con $h \leq k$, *linearmente indipendenti* tali che

$$\mathrm{span}\{X_1, \ldots, X_k\} = \mathrm{span}\{Y_1, \ldots, Y_h\}.$$

Un modo per farlo è questo:
– Considerare la matrice M, di ordine $k \times n$ che ha per righe i vettori $\{X_1, \ldots, X_k\}$;
– fare l'eliminazione di Gauss per righe su M;
– Le righe non nulle $Y_1, \ldots, Y_h$ della matrice a scala cosi' ottenuta sono linearmente indipendenti e

$$\mathrm{span}\{X_1, \ldots, X_k\} = \mathrm{span}\{Y_1, \ldots, Y_h\}.$$

4. Basi, dimensione e coordinate.

Definizione. Uno spazio vettoriale V si dice finitamente generato se esistono un numero finito di elementi $v_1, \ldots, v_k \in V$ (generatori) , tali che

$$V = \mathrm{span}\{v_1, \ldots, v_k\}.$$

Esempio.
(i) Per ogni n, lo spazio $\mathbf{R}^n$ è finitamente generato:

$$\mathbf{R}^n = \mathrm{span}\{ \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \}.$$

(ii) Lo spazio vettoriale $\mathbf{R}[x]_m$ dei polinomi di grado $\leq m$ è finitamente generato:

$$\mathbf{R}[x]_m = \mathrm{span}\{1, x, x^2, \ldots, x^m\}.$$

(iii) Lo spazio di tutti i polinomi $\mathbf{R}[x]$ non è finitamente generato:

$$\mathbf{R}[x] = \mathrm{span}\{1, x, x^2, \ldots, x^n, \ldots\},$$

e tanto meno lo sono gli spazi $\mathbf{R}[x] \subset C^1(\mathbf{R}, \mathbf{R}) \subset C^0(\mathbf{R}, \mathbf{R}) \subset F(\mathbf{R}, \mathbf{R})$.

Consideriamo da ora in poi spazi vettoriali $V \neq O$.

Definizione. Una base di uno spazio vettoriale è un insieme *ordinato* di generatori linearmente indipendenti.

Osservazione. Pur contenendo gli stessi elementi, due basi $\mathcal{B}_1 = \{X_1, X_2, X_3\}$ e $\mathcal{B}_2 = \{X_2, X_1, X_3\}$ di $\mathbf{R}^3$ sono da considerarsi distinte, in quanto l'ordine degli elementi non è lo stesso.

Proposizione. *Ogni spazio vettoriale V finitamente generato ammette una base.*

Dimostrazione. Sia $\{v_1, \ldots, v_k\}$ un insieme di generatori di V. Se sono linearmente indipendenti, sono una base di V e la proposizione è dimostrata. Se non lo sono, almeno uno di essi è combinazione lineare dei rimanenti. Lo eliminiamo. Gli elementi che restano generano ancora V. Se sono linearmente indipendenti, sono una base di V e la proposizione è dimostrata. Se non lo sono, ne possiamo eliminare un altro e i rimanenti generano ancora V. Dopo un numero finito di passi otteniamo una base di V.

Osservazione. Ogni spazio vettoriale V, finitamente generato o no, ammette una base. Nel caso di uno spazio vettoriale non finitamente generato, la dimostrazione non è elementare.

Proposizione. *Tutte le basi di uno spazio vettoriale V finitamente generato hanno la stessa cardinalità.*

Dimostrazione. Siano $\mathcal{B}_1 = \{v_1, \ldots, v_n\}$ e $\mathcal{B}_2 = \{w_1, \ldots, w_m\}$ due basi di V e supponiamo per assurdo che $n > m$. Scriviamo gli elementi di $\mathcal{B}_1$ come combinazioni degli elementi di $\mathcal{B}_2$:

$$\begin{aligned} v_1 =& a_{11}w_1 + \ldots a_{1m}w_m \\ v_2 =& a_{21}w_1 + \ldots a_{2m}w_m \\ \vdots& \\ v_n =& a_{n1}w_1 + \ldots a_{nm}w_m \end{aligned} \quad . \tag{1}$$

Poiché $v_1, \ldots, v_n$ sono linearmente indipendenti,

$$\alpha_1 v_1 + \ldots + \alpha_n v_n = O \tag{2}$$

se e solo se $\alpha_1 = \ldots = \alpha_n = 0$. D'altra parte, sostituendo le relazioni (1) nell'equazione (2) e raccogliendo i coefficienti di $w_1, \ldots, w_m$, si ha la (2) vale se e solo se

$$\begin{cases} a_{11}\alpha_1 + \ldots a_{n1}\alpha_n = 0 \\ a_{12}\alpha_1 + \ldots a_{n2}\alpha_n = 0 \\ \vdots \quad \vdots \\ a_{1m}\alpha_1 + \ldots a_{nm}\alpha_n = 0 \\ . \end{cases} \tag{3}$$

Questo è assurdo, perchè il sistema (3) è un sistema lineare *omogeneo* di m equazioni in n incognite, con $n > m$ (piu' incognite che equazioni). Tale sistema ammette soluzioni $(\alpha_1, \ldots, \alpha_n) \neq (0, \ldots, 0)$ (vedi appunti 1), e questo contraddice la lineare indipendenza di $\{v_1, \ldots, v_n\}$.

Definizione. Sia V uno spazio vettoriale finitamente generato. La dimensione di V è per definizione la cardinalità di una sua qualunque base.

Esempio. La dimensione di $\mathbf{R}^n$ è uguale a n. Un qualunque insieme di n vettori linearmente indipendenti in uno spazio vettoriale di dimensione n è una base di V.

Osservazione. Sia V uno spazio vettoriale di dimensione n. Con lo stesso ragionamento usato nella proposizione precedente si dimostra che k elementi di V, con $k > n$, sono necessariamente linearmente dipendenti.

Proposizione. *(completamento di una base). Sia V uno spazio vettoriale di dimensione n, sia $U \subset V$ un sottospazio e sia $\{u_1, \ldots, u_k\}$ una base di U. Allora esistono $v_{k+1}, \ldots, v_n$ elementi di V tali che $\{u_1, \ldots, u_k, v_{k+1}, \ldots, v_n\}$ sia una base di V.*

Dimostrazione. Se $U = V$, non c'è più niente da dimostrare. Sia allora $U \neq V$. In questo caso, esiste $v_{k+1} \in V \setminus \mathrm{span}\{u_1, \ldots, u_k\}$. Poiché i vettori $\{u_1, \ldots, u_k, v_{k+1}\}$ sono linearmente indipendenti, se $\text{span}\{u_1, \ldots, u_k, v_{k+1}\} = V$, formano una base di V e abbiamo finito. Altrimenti possiamo trovare $v_{k+2} \in V \setminus \text{span}\{u_1, \ldots, u_k, v_{k+1}\}$ e considerare i vettori linearmente indipendenti $\{u_1, \ldots, u_k, v_{k+1}, v_{k+2}\}$. Se $\text{span}\{u_1, \ldots, u_k, v_{k+1}, v_{k+2}\} = V$, abbiamo finito. Altrimenti, ripetendo lo stesso ragionamento, dopo numero finito di passi avremo trovato una base di V, i cui primi k elementi sono la base data di U.

Osservazione. Gli elementi $v_{k+1}, \ldots, v_n$ possono essere scelti in infiniti modi.

Proposizione. *Sia V uno spazio vettoriale di dimensione n e sia $\mathcal{B}_1 = \{v_1, \ldots, v_n\}$ una sua base. Allora ogni elemento $v \in V$ si scrive in modo unico come*

$$v = x_1 v_1 + \ldots + x_n v_n, \quad x_i \in \mathbf{R}.$$

Dimostrazione. Che si possa scrivere $v = x_1 v_1 + \ldots + x_n v_n$, per oppurtuni $x_i \in \mathbf{R}$, segue dal fatto che $\{v_1, \ldots, v_n\}$ sono generatori di V. La loro lineare indipendenza garantisce l'unicità degli scalari x_i: supponiamo infatti che si possa scrivere

$$v = x_1 v_1 + \ldots + x_n v_n = y_1 v_1 + \ldots + y_n v_n, \quad x_i, y_i \in \mathbf{R}.$$

Questo equivale a $(x_1 - y_1)v_1 + \ldots + (x_n - y_n)v_n = O$, che a sua volta implica $x_1 = y_1, \ldots, x_n = y_n$.

Conseguenza. Per la proposizione precedente, fissare una base $\mathcal{B} = \{v_1, \ldots, v_n\}$ in uno spazio vettoriale V di dimensione n induce una identificazione di V con $\mathbf{R}^n$ mediante

$$v = x_1 v_1 + \ldots + x_n v_n \quad \leftrightarrow \quad \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Questa identificazione rispetta la operazioni su V, nel senso che alla somma $v + w$ degli elementi $v = x_1 v_1 + \ldots + x_n v_n$ e $w = y_1 v_1 + \ldots + y_n v_n$ corrisponde la somma delle rispettive ennuple $\begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}$.
Analogamente al prodotto di λv di uno scalare λ per un elemento $v = x_1 v_1 + \ldots + x_n v_n$ corrisponde il prodotto $\lambda \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix}$.

Definizione. I numeri $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ sono per definizione le coordinate dell'elemento $v = x_1 v_1 + \ldots + x_n v_n$ nella base $\mathcal{B} = \{v_1, \ldots, v_n\}$ di V.

5. Somma e intersezione di sottospazi, formule di Grassmann.

Sia V uno spazio vettoriale di dimensione finita n e siano U e W sottospazi di V. A partire da essi costruiamo dei nuovi sottospazi di V.

Definizione. Il sottospazio *intersezione di* U *e* W

$$U \cap W = \{v \in V \mid v \in U \text{ e } v \in W\}$$

è formato dagli elementi di V che appartengono sia a U che a W.

È immediato verificare che $U \cap W$ è effettivamente un sottospazio vettoriale di V.
Se $V = \mathbf{R}^n$ e i sottospazi U, W sono dati rispettivamente dalle soluzioni di un sistema lineare omogeneo di m equazioni in n incognite e dalle soluzioni di un sistema lineare omogeneo di p equazioni in n incognite

$$U: \quad \begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n = 0 \\ \ldots \quad \ldots \\ a_{m1}x_1 + \ldots + a_{mn}x_n = 0 \end{cases} \qquad W: \quad \begin{cases} b_{11}x_1 + \ldots + b_{1n}x_n = 0 \\ \ldots \quad \ldots \\ b_{p1}x_1 + \ldots + b_{pn}x_n = 0 \end{cases},$$

allora il sottospazio $U \cap W$ è dato dalle soluzioni del sistema lineare omogeneo di $m + p$ equazioni

$$U \cap W: \quad \begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n = 0 \\ \ldots \quad \ldots \\ a_{m1}x_1 + \ldots + a_{mn}x_n = 0 \\ b_{11}x_1 + \ldots + b_{1n}x_n = 0 \\ \ldots \quad \ldots \\ b_{p1}x_1 + \ldots + b_{pn}x_n = 0. \end{cases}$$

Definizione. Il sottospazio *somma di* U *e* W

$$U + W = \{v \in V \mid v = u + w, \ u \in U, \ w \in W\}$$

è formato dagli elementi di V che si possono scrivere come somma di un elemento di U e di un elemento di W. Se vale $U \cap W = \{O\}$, allora si dice che $U + W$ è somma diretta di U e W e si indica con $U \oplus W$.

Segue immediatamente dalla definizione che $U + W$ è effettivamente un sottospazio vettoriale di V.
Osserviamo che se $U = \text{span}\{u_1, \ldots, u_k\}$ e $W = \text{span}\{w_1, \ldots, w_h\}$, allora

$$U + W = \text{span}\{u_1, \ldots, u_k, w_1, \ldots, w_h\}.$$

In altre parole, se $\{u_1, \ldots, u_k\}$ sono un insieme di generatori di U e $\{w_1, \ldots, w_h\}$ sono un insieme di generatori di W, allora l'unione $\{u_1, \ldots, u_k\} \cup \{w_1, \ldots, w_h\}$ costituisce un insieme di generatori di $U + W$.

• Valgono le seguenti inclusioni fra sottospazi

$$U \cap W \subset U \subset (U + W) \subset V, \qquad U \cap W \subset W \subset (U + W) \subset V,$$

e le seguenti disuguaglianze fra le rispettive dimensioni

$$\dim U \cap W \le \dim U \le \dim(U + W) \le \dim V, \quad \dim U \cap W \le \dim W \le \dim(U + W) \le \dim V.$$

Esercizio. Chi sono i sottospazi $U + W$ e $U \cap W$ quando $W \subset U$?

Proposizione. *(formule di Grassmann). Sia V uno spazio vettoriale di dimensione finita. Siano U, W sottospazi di V. Allora vale la seguente relazione fra le dimensioni di U, V, $U+V$, $U \cap V$*

$$\dim(U+W) = \dim(U) + \dim(W) - \dim(U \cap W).$$

Dimostrazione. Siano $\dim(U \cap W) = q$, $\dim U = t$ e $\dim W = s$. Sia $\{z_1, \ldots, z_q\}$ una base di $U \cap W$ (se $q = 0$, allora $\{z_1, \ldots, z_q\} = \emptyset$) . Per il principio del completamento di una base, esistono $t - q$ elementi $\{u_{q+1}, \ldots, u_t\}$ di U tali che $\{z_1, \ldots, z_q, u_{q+1}, \ldots, u_t\}$ formano una base di U. Allo stesso modo, esistono $s-q$ elementi $\{w_{q+1}, \ldots, w_s\}$ di W tali che $\{z_1, \ldots, z_q, w_{q+1}, \ldots, w_s\}$ formano una base di W. Dalla definizione di $U+W$ segue che

$$U + W = \mathrm{span}\{z_1, \ldots, z_q, u_{q+1}, \ldots, u_t, w_{q+1}, \ldots, w_s\},$$

ossia che $\{z_1, \ldots, z_q, u_{q+1}, \ldots, u_t, w_{q+1}, \ldots, w_s\}$ generano $U+W$. Se dimostriamo che sono anche linearmente indipendenti (cioè formano una base di $U+W$), otteniamo $\dim(U+W) = t+s-q$ e la proposizione è dimostrata. Sia dunque

$$\alpha_1 z_1 + \ldots + \alpha_q z_q + \alpha_{q+1} u_{q+1} + \ldots + \alpha_t u_t + \beta_{q+1} w_{q+1} + \ldots + \beta_s w_s = O, \qquad (*)$$

una loro combinazione lineare che dà il vettore nullo, con $\alpha_1, \ldots, \alpha_t, \beta_{q+1} \ldots \beta_s \in \mathbf{R}$. La relazione (*) è quivalente a

$$\beta_{q+1} w_{q+1} + \ldots + \beta_s w_s = -(\alpha_1 z_1 + \ldots + \alpha_q z_q + \alpha_{q+1} u_{q+1} + \ldots + \alpha_t u_t).$$

Osserviamo che il vettore $X = \beta_{q+1} w_{q+1} + \ldots + \beta_s w_s$ è un elemento di $W \cap U$, in quanto può essere espresso sia come combinazione lineare di elementi di W che come combinazione lineare di elementi di U. Poiché $\{z_1, \ldots, z_q\}$ sono una base di $U \cap W$, esistono $\gamma_1, \ldots, \gamma_q \in \mathbf{R}$ tali che

$$X = \beta_{q+1} w_{q+1} + \ldots + \beta_s w_s = -(\alpha_1 z_1 + \ldots + \alpha_q z_q + \alpha_{q+1} u_{q+1} + \ldots + \alpha_t u_t) = \gamma_1 z_1 + \ldots + \gamma_q z_q. \qquad (**)$$

Poiché $\{z_1, \ldots, z_q, u_{q+1}, \ldots, u_t\}$ sono una base di U, accoppiando gli ultimi due termini della (**), si ricava

$$\alpha_{q+1} = \ldots = \alpha_t = 0.$$

Similmente, poiché $\{z_1, \ldots, z_q, w_{q+1}, \ldots, w_s\}$ sono una base di W, accoppiando i primi due termini della (**), si ricava che anche i rimanenti α_j e β_k devono essere nulli. Dunque i vettori $\{z_1, \ldots, z_q, u_{q+1}, \ldots, u_t, w_{q+1}, \ldots, w_s\}$ sono una base di $U+W$ e $\dim(U+W) = t+s-q = \dim U + \dim W - \dim U \cap W$ come richiesto.

Definizione. Si definisce *complemento di U in V* e si indica con U^c un qualunque sottospazio di V tale che

$$V = U \oplus U^c.$$

Sia $\{u_1, \ldots, u_k\}$ una base di U e sia $\{v_{k+1}, \ldots, v_n\}$ un suo qualunque completamento ad una base di V. Allora il sottospazio $\mathrm{span}\{v_{k+1}, \ldots, v_n\}$ è un complemento di U in V. Per costruzione valgono infatti le relazioni

$$V = U + U^c \quad \text{e} \quad U \cap U^c = \{O\}.$$

In altre parole, i complementi di U in V sono in corrispondenza $1-1$ con i completamenti di una base di U ad una base di V.

• Per ogni complemento di U in V si ha

$$\dim V = \dim U + \dim U^c.$$

Esempio. Sia $V = \mathbf{R}^2$ e sia $U = \mathrm{span}\{v\}$, con $v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Per ogni $w = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, che non sia multiplo di v, il sottospazio $W = \mathrm{span}\{w\}$ costituisce un complemento di U in $\mathbf{R}^2$.

Geometricamente: fissata una retta per l'origine in $\mathbf{R}^2$, una qualunque altra retta per l'origine, distinta da essa, costituisce un suo complemento.

Esempio. Qual è la generalizzazione di questi fatti in $\mathbf{R}^3$?
Sia $V = \mathbf{R}^3$ e sia $U = \mathrm{span}\{u, v\}$, con u, v linearmente indipendenti. Per ogni vettore non nullo w, linearmente indipendente da u, v, ossia $w \notin \mathrm{span}\{u, v\}$, il sottospazio $W = \mathrm{span}\{w\}$ costituisce un complemento di U in $\mathbf{R}^3$.
Geometricamente: fissato un piano per l'origine in $\mathbf{R}^3$, una qualunque retta per l'origine non contenuta nel piano, costituisce un suo complemento; fissata una retta per l'origine in $\mathbf{R}^3$, un qualunque piano per l'origine, che non la contiene, costituisce un suo complemento.

Esempio. Siano dati i sottospazi

$$U = \mathrm{span}\{\begin{pmatrix}1\\1\\0\\0\end{pmatrix}, \begin{pmatrix}1\\-1\\0\\0\end{pmatrix}\} \quad \text{e} \quad W = \mathrm{span}\{\begin{pmatrix}1\\1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\\1\end{pmatrix}, \begin{pmatrix}0\\0\\1\\-1\end{pmatrix}\}$$

in $\mathbf{R}^4$. In questo caso è evidente che

$$\dim U = 2 \quad \dim W = 3 \quad U + W = \mathbf{R}^4 \quad \dim U \cap W = 1 \quad U \cap W = \mathrm{span}\{\begin{pmatrix}1\\1\\0\\0\end{pmatrix}\}.$$

Osserviamo che i vettori

$$\{\begin{pmatrix}1\\1\\0\\0\end{pmatrix}, \begin{pmatrix}1\\-1\\0\\0\end{pmatrix}\} \cup \{\begin{pmatrix}1\\1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\\1\end{pmatrix}, \begin{pmatrix}0\\0\\1\\-1\end{pmatrix}\}$$

sono generatori di $U + W = \mathbf{R}^4$, ma non sono una base di $\mathbf{R}^4$ (anche se i primi due sono una base di U e gli ultimi tre sono una base di W).

Sia $\mathcal{M}(n, n, \mathbf{R})$ lo spazio vettoriale delle matrici quadrate $n \times n$. Indichiamo con

$$\mathcal{T}^+(n) = \{M = \begin{pmatrix}m_{11} & \dots & m_{1n}\\ \vdots & \ddots & \vdots\\ m_{n1} & \dots & m_{nn}\end{pmatrix} \mid m_{ij} = 0,\ i > j\} = \{\begin{pmatrix}m_{11} & m_{12} & \dots & m_{1n}\\ 0 & m_{22} & \dots & m_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & m_{nn}\end{pmatrix}\}$$

le matrici *triangolari superiori*, ossia le matrici che hanno tutti gli elementi sotto la diagonale principale uguali a zero, con

$$\mathcal{T}^-(n) = \{M = \begin{pmatrix}m_{11} & \dots & m_{1n}\\ \vdots & \ddots & \vdots\\ m_{n1} & \dots & m_{nn}\end{pmatrix} \mid m_{ij} = 0,\ i < j\} = \{\begin{pmatrix}m_{11} & 0 & \dots & 0\\ m_{21} & m_{22} & \dots & \vdots\\ \vdots & \vdots & \ddots & 0\\ m_{n1} & \dots & \dots & m_{nn}\end{pmatrix}\}$$

le matrici *triangolari inferiori*, ossia le matrici che hanno tutti gli elementi sopra la diagonale principale uguali a zero, ed infine con

$$\mathcal{D}(n) = \{M = \{m_{ij}\} \mid m_{ij} = 0,\ i \neq j\} = \{\begin{pmatrix}m_{11} & 0 & \dots & 0\\ 0 & m_{22} & \dots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & m_{nn}\end{pmatrix}\}.$$

le matrici *diagonali*, ossia le matrici che hanno tutti gli elementi fuori dalla diagonale principale uguali a zero.

Esercizio.
(a) Dimostrare che $\mathcal{T}^+(n)$, $\mathcal{T}^-(n)$, $\mathcal{D}(n)$ sono sottospazi vettoriali di $\mathcal{M}(n,n,\mathbf{R})$ di dimensione rispettivamente uguale a

$$\dim \mathcal{T}^+(n) = \frac{n(n+1)}{2}, \qquad \dim \mathcal{T}^-(n) = \frac{n(n+1)}{2}, \qquad \dim \mathcal{D}(n) = n.$$

(b) Dimostrare che $\mathcal{T}^+(n) + \mathcal{T}^-(n) = \mathcal{M}(n,n,\mathbf{R})$.
(c) Dimostrare che $\mathcal{T}^+(n) \cap \mathcal{T}^-(n) = \mathcal{D}(n)$.

6. Matrici.

Prodotto righe per colonne fra matrici.

Abbiamo visto in precedenza che matrici dello stesso ordine possono essere sommate tra loro e moltiplicate per uno scalare. In altre parole, le matrici $m \times n$ a coefficienti reali

$$\mathcal{M}(m,n,\mathbf{R}) = \{A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \mid a_{ij} \in \mathbf{R}, \quad i = 1,\dots,m, \ j = 1,\dots,n\}$$

formano uno spazio vettoriale, quando la somma fra matrici e il prodotto di una matrice per uno scalare sono definiti nel modo seguente

$$A + B := \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \dots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \dots & a_{2n}+b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \dots & a_{mn}+b_{mn} \end{pmatrix}$$

dove $A = \{a_{ij}\}$, $B = \{b_{ij}\} \in \mathcal{M}(m,n,\mathbf{R})$ e

$$\lambda A := \begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \dots & \lambda a_{1n} \\ \lambda a_{21} & \lambda a_{22} & \dots & \lambda a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda a_{m1} & \lambda a_{m2} & \dots & \lambda a_{mn} \end{pmatrix},$$

dove $A = \{a_{ij}\} \in \mathcal{M}(m,n,\mathbf{R})$, $\lambda \in \mathbf{R}$.

Siano date adesso due matrici

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}, \qquad B = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1k} \\ b_{21} & b_{22} & \dots & b_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nk} \end{pmatrix},$$

rispettivamente di ordine $m \times n$ ed $n \times k$.

DefinizioneIl *prodotto righe per colonne di A e B* è la matrice $m \times k$ così definita

$$A \cdot B = \begin{pmatrix} R_1 \cdot C_1 & R_1 \cdot C_2 & \dots & R_1 \cdot C_k \\ R_2 \cdot C_1 & R_2 \cdot C_2 & \dots & R_2 \cdot C_k \\ \vdots & \vdots & \ddots & \vdots \\ R_m \cdot C_1 & R_m \cdot C_2 & \dots & R_m \cdot C_k \end{pmatrix},$$

dove $R_i = (\, a_{i1} \quad a_{i2} \quad \dots \quad a_{in}\,)$ indica la i-sima riga di A, $C_j = \begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{pmatrix}$ indica la j-sima colonna di B e

$$R_i \cdot C_j = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj}.$$

È chiaro dalla definizione che il prodotto fra matrici è un'applicazione

$$\mathcal{M}(m, n, \mathbf{R}) \times \mathcal{M}(n, k, \mathbf{R}) \longrightarrow \mathcal{M}(m, k, \mathbf{R}), \quad (A, B) \mapsto A \cdot B,$$

ossia il prodotto $A \cdot B$ è definito solo se la lunghezza delle righe di A è uguale alla lunghezza delle colonne di B.

Indichiamo con I_n la matrice identità di ordine n

$$I_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

Proposizione. *Sia A una matrice $m \times n$ e siano B, C matrici di dimensioni tali che le somme e i prodotti elencati siano ben definiti. Allora valgono le seguenti identità:*
(i) $A \cdot (B \cdot C) = (A \cdot B) \cdot C$ *Associatività del prodotto*
(ii) $A \cdot (B + C) = (A \cdot B) + A \cdot C$ *Distributività a destra*
(iii) $(A + B) \cdot C = A \cdot C + B \cdot C$ *Distributività a sinistra*
(iv) $(\lambda A) \cdot B = A \cdot (\lambda B) = \lambda(A \cdot B)$
(v) $I_m \cdot A = A \cdot I_n = A$ *Elemento identità per il prodotto*

Indichiamo con $O_{m,n}$ la matrice identicamente nulla di ordine $m \times n$

$$O_{m,n} = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}.$$

Si verifica facilmente che per ogni matrice A di ordine $m \times n$

$$O_{h,m} \cdot A = O_{h,n}, \quad A \cdot O_{n,k} = O_{m,k}.$$

Osservazione. Per il prodotto fra matrici *non valgono* invece la *proprietà commutativa* e il *principio dell'annullamento.* Anche quando entrambi i prodotti $A \cdot B$ e $B \cdot A$ sono definiti, generalmente

$$A \cdot B \neq B \cdot A.$$

Inoltre, esistono matrici non nulle A e B il cui prodotto è la matrice nulla:

$$A \cdot B = O \quad \not\Rightarrow \quad A = O \text{ oppure } B = O.$$

Esempio.

(i) Siano $A = \begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix}$ e $B = \begin{pmatrix} 1 & 0 \\ 5 & 3 \end{pmatrix}$. È facile verificare che le matrici A e B non commutano. Infatti

$$A \cdot B = \begin{pmatrix} 17 & 9 \\ 21 & 12 \end{pmatrix} \neq B \cdot A = \begin{pmatrix} 2 & 3 \\ 13 & 27 \end{pmatrix}.$$

(ii) Le matrici $A = \begin{pmatrix} 0 & 3 \\ 0 & 0 \end{pmatrix}$ e $B = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix}$ sono matrici non nulle, il cui prodotto è la matrice nulla:

$$A \cdot B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Le operazioni di somma e prodotto fra matrici hanno un significato preciso che illustriamo brevemente. Sia A una matrice $m \times n$. Ad essa è corrisponde in modo naturale l'applicazione che ad un vettore $X \in \mathbf{R}^n$ (visto come una matrice $n \times 1$) associa il vettore di $\mathbf{R}^m$ dato dal prodotto $A \cdot X$

$$L_A \colon \mathbf{R}^n \longrightarrow \mathbf{R}^m$$

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto A \cdot X = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \ldots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n \end{pmatrix}.$$

Dalla Proposizione segue che l'applicazione L_A gode delle seguenti proprietà

$$L_A(X+Y) = L_A(X) + L_A(Y), \qquad L_A(\lambda X) = \lambda L_A(X), \qquad \forall X,\ Y \in \mathbf{R}^n,\ \forall \lambda \in \mathbf{R}.$$

Se B è un'altra matrice $m \times n$, le applicazioni $L_A \colon \mathbf{R}^n \longrightarrow \mathbf{R}^m$ ed $L_B \colon \mathbf{R}^n \longrightarrow \mathbf{R}^m$ si possono sommare fra loro e moltiplicare per uno scalare (come si fa per le funzioni a valori reali)

$$(L_A + L_B)(X) := L_A(X) + L_B(X), \quad (\lambda L_A)(X) = \lambda L_A(X).$$

L'applicazione L_{B+A} associata alla somma delle matrici $A+B$ coincide con la somma delle applicazioni $L_A + L_B$; l'applicazione $L_{\lambda A}$ associata al prodotto λA coincide con l'applicazione λL_A:

$$(L_A + L_B)(X) = L_{A+B}(X), \qquad (\lambda L_A)(X) = L_{\lambda A}(X), \qquad \forall X \in \mathbf{R}^n.$$

Se B è una matrice $k \times m$, e

$$L_B \colon \mathbf{R}^m \longrightarrow \mathbf{R}^k, \quad X \mapsto B \cdot X$$

è l'applicazione corrispondente, possiamo considerare la composizione delle applicazioni

$$L_B \circ L_A \colon \mathbf{R}^n \longrightarrow \mathbf{R}^m \longrightarrow \mathbf{R}^k, \quad L_B \circ L_A(X) := L_B(L_A(X)) = B \cdot (A \cdot X)).$$

Il significato del prodotto righe per colonne fra matrici è il seguente:
l'applicazione $L_{B \cdot A}$ associata al prodotto righe per colonne $B \cdot A$ coincide con l'applicazione composta $L_B \circ L_A$:

$$L_B \circ L_A = L_{B \cdot A},$$

Trasposta di una matrice.

Sia $A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$ una matrice $m \times n$.

Definizione. La trasposta di A è la matrice $n \times m$, indicata con ${}^tA$, che ha per colonne le righe di A

$$ {}^tA = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{pmatrix}. $$

• L'operazione di *trasposizione* gode delle seguenti proprietà

(i) ${}^t({}^tA) = A$;

(ii) ${}^t(A + B) = {}^tA + {}^tB$;

(iii) ${}^t(\lambda A) = \lambda\, {}^tA$;

(iv) ${}^t(A \cdot B) = {}^tB \cdot {}^tA$.

Esempio. Sia $A = \begin{pmatrix} 1 & 3 & 1 & 4 \\ 1 & 2 & 0 & 0 \end{pmatrix}$. Allora ${}^tA = \begin{pmatrix} 1 & 1 \\ 3 & 2 \\ 1 & 0 \\ 4 & 0 \end{pmatrix}$.

Se A è una matrice quadrata $n \times n$, la sua trasposta è ancora una matrice quadrata $n \times n$. In questo caso ha senso confrontare A con la sua trasposta ${}^tA$.

Definizione. Una matrice A quadrata $n \times n$ si dice *simmetrica* se $A = {}^tA$, ossia se coincide con la sua trasposta.

Definizione. Una matrice A quadrata $n \times n$ si dice *antisimmetrica* se $A = -{}^tA$, ossia se coincide con l'opposto della sua trasposta.

Esempio. Le matrici simmetriche 2×2 sono della forma $\begin{pmatrix} a & b \\ b & d \end{pmatrix}$, $a, b, d \in \mathbf{R}$; quelle antisimmetriche sono della forma $\begin{pmatrix} 0 & c \\ -c & 0 \end{pmatrix}$, $c \in \mathbf{R}$. Le matrici simmetriche 3×3 sono della forma $\begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix}$, $a, b, c, d, e, f \in \mathbf{R}$; quelle antisimmetriche della forma $\begin{pmatrix} 0 & b & c \\ -b & 0 & e \\ -c & -e & 0 \end{pmatrix}$, $b, c, e \in \mathbf{R}$.

Esercizio. Sia $\mathcal{M}(n, n, \mathbf{R})$ lo spazio vettoriale delle matrici quadrate $n \times n$.

(a) Dimostrare che le matrici simmetriche e le matrici antisimmetriche

$$ \mathcal{S}(n) = \{M \in \mathcal{M}(n, n, \mathbf{R}) \mid {}^tM = M\}, \quad \mathcal{A}(n) = \{M \in \mathcal{M}(n, n, \mathbf{R}) \mid {}^tM = -M\} $$

sono sottospazi vettoriali di $\mathcal{M}(n, n, \mathbf{R})$ di dimensione rispettivamente uguale a

$$ \dim \mathcal{S}(n) = \frac{n(n+1)}{2}, \qquad \dim \mathcal{A}(n) = \frac{n(n-1)}{2}. $$

(b) Dimostrare che $\mathcal{S}(n) \cap \mathcal{A}(n) = \{O\}$ e che

$$ \mathcal{M}(n, n, \mathbf{R}) = \mathcal{S}(n) \oplus \mathcal{A}(n). $$

Esercizio. Verificare che la trasposta di una matrice triangolare superiore è una matrice triangolare inferiore e che la trasposta di una matrice triangolare inferiore è una matrice triangolare superiore.

7. Applicazioni lineari.
Generalità.

Definizione. Una applicazione $L: V \longrightarrow W$ fra spazi vettoriali è un'applicazione lineare se

$$L(u+v) = L(u) + L(v), \quad L(\lambda v) = \lambda L(v), \qquad \forall u, v \in V, \ \forall \lambda \in \mathbf{R}.$$

Esempio. Siano $V = \mathbf{R}^n$ e $W = \mathbf{R}^m$ e sia $A = \{a_{ij}\}$ una matrice $m \times n$. L'applicazione $L_A: R^n \longrightarrow \mathbf{R}^m$ data dalla moltiplicazione matrice-vettore

$$L_A(X) = A \cdot X = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \qquad X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

è un'applicazione lineare. Questo segue dalle proprietà del prodotto fra matrici:

$$A \cdot (X+Y) = A \cdot X + A \cdot Y, \qquad A \cdot (\lambda X) = \lambda (A \cdot X), \qquad \forall X, Y \in \mathbf{R}^n, \ \forall \lambda \in \mathbf{R}.$$

Vedremo in seguito che vale anche il viceversa: *ogni applicazione lineare $L: R^n \longrightarrow \mathbf{R}^m$ è della forma $L(X) = AX$, per una opportuna matrice A di ordine $m \times n$.*

Esempio. Siano $V = W = C^\infty(\mathbf{R}, \mathbf{R})$, dove $C^\infty(\mathbf{R}, \mathbf{R})$ indica lo spazio vettoriale delle funzioni reali derivabili infinite volte. L'applicazione

$$D: C^\infty(\mathbf{R}, \mathbf{R}) \longrightarrow C^\infty(\mathbf{R}, \mathbf{R}), \quad f \mapsto f',$$

che associa ad una funzione f la sua derivata, è un'applicazione lineare. Questo segue dal fatto che

$$(f+g)' = f' + g', \quad (\lambda f)' = \lambda f', \qquad \forall f, g \in C^\infty(\mathbf{R}, \mathbf{R}), \ \forall \lambda \in \mathbf{R}.$$

Esempio. Siano $V = W = C^\infty(\mathbf{R}, \mathbf{R})$. L'applicazione

$$I: C^\infty(\mathbf{R}, \mathbf{R}) \longrightarrow C^\infty(\mathbf{R}, \mathbf{R}), \quad f \mapsto \int f,$$

che associa ad una funzione f il suo integrale indefinito, è un'applicazione lineare. Questo segue dal fatto che

$$\int (f+g) = \int f + \int g, \quad \int (\lambda f) = \lambda \int f, \qquad \forall f, g \in C^\infty(\mathbf{R}, \mathbf{R}), \ \forall \lambda \in \mathbf{R}.$$

Esempio. Siano $V = C^\infty(\mathbf{R}, \mathbf{R})$, $W = \mathbf{R}$, $x_0 \in \mathbf{R}$. L'applicazione

$$V_{x_0}: C^\infty(\mathbf{R}, \mathbf{R}) \longrightarrow \mathbf{R}, \quad f \mapsto f(x_0),$$

che associa ad una funzione f il suo valore in x_0, è un'applicazione lineare. Questo segue dal fatto che

$$(f+g)(x_0) = f(x_0) + g(x_0), \qquad (\lambda f)(x_0) = \lambda f(x_0), \quad \forall f, g \in C^\infty(\mathbf{R}, \mathbf{R}), \ \forall \lambda \in \mathbf{R}.$$

Sia $L: V \longrightarrow W$ un'applicazione lineare. I seguenti fatti seguono direttamente dalla definizione:

(i) *$L(O_V) = O_W$, dove O_V e O_W indicano rispettivamente lo zero in V e lo zero in W.*

Infatti, per ogni $v \in V$, si ha $L(O_V) = L(v - v) = L(v) - L(v) = O_W$.

(ii) *L manda sottospazi di V in sottospazi di W:*

$$U = \mathrm{span}\{u_1, \dots, u_k\} \subset V \quad \Rightarrow \quad L(U) = \mathrm{span}\{L(u_1), \dots, L(u_k)\}.$$

Se $u = \lambda_1 u_1 + \ldots + \lambda_k u_k \in U = \mathrm{span}\{u_1, \dots, u_k\}$, dalla linearità di L segue che $L(u) = \lambda_1 L(u_1) + \ldots + \lambda_k L(u_k)$, ossia $L(u) \in \mathrm{span}\{L(u_1), \dots, L(u_k)\}$. In altre parole, se $\{u_1, \dots, u_k\}$ sono generatori del sottospazio $U \subset V$, le immagini $\{L(u_1), \dots, L(u_k)\}$ sono generatori dell'immagine $L(U)$ in W.

(iii) *Se $\dim V = n < \infty$, allora L è completamente determinata dai valori assunti sugli elementi di una qualunque base di V.*

Infatti, se $\mathcal{B} = \{v_1, \dots, v_n\}$ è una base di V, ogni $v \in V$ si scrive in modo unico come come $v = x_1 v_1 = \ldots + x_n v_n$, per oppurtuni $x_1, \dots, x_n \in \mathbf{R}$. Per la linearità di L, si ha

$$L(v) = x_1 L(v_1) + \ldots + x_n L(v_n),$$

ossia l'immagine di v tramite L dipende solo dai valori assegnati agli elementi della base $L(v_1), \dots, L(v_n)$ e dalle coordinate di v.

Da ciò segue che:

Un'applicazione lineare $L \colon \mathbf{R}^n \longrightarrow \mathbf{R}^m$ è sempre data dalla moltiplicazione matrice-vettore, ossia esiste una matrice A di ordine $m \times n$ tale che $L(X) = AX$.

Dim. Scriviamo $X \in \mathbf{R}^n$ come

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \ldots + x_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Poiché L è lineare, abbiamo

$$L(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}) = x_1 L(\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}) + x_2 L(\begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}) + \ldots + x_n L(\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}). \qquad (*)$$

Sostituendo

$$L(\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}) = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \ L(\begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}) = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, \ldots\ldots, L(\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}) = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix},$$

nella (*), troviamo

$$L(X) = x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + \ldots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \ldots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n \end{pmatrix} = AX,$$

con A matrice $m \times n$ data da

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{11} & a_{12} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}.$$

Esercizio 1. Un'applicazione lineare $L: V \longrightarrow W$ manda insiemi di elementi linearmente dipendenti in V in insiemi di elementi linearmente dipendenti in W.

Sol. Siano $\{v_1, \ldots, v_k\}$ un insieme di elementi linearmente dipendenti in V. Per definizione, esistono $\alpha_1, \ldots, \alpha_k$ reali, non tutti nulli, tali che

$$\alpha_1 v_1 + \ldots + \alpha_k v_k = O_V.$$

Per la linearità di L, segue che

$$\alpha_1 L(v_1) + \ldots + \alpha_k L(v_k) = L(O_V) = O_W.$$

In altre parole, $L(v_1), \ldots, L(v_k)$ sono linearmente dipendenti.

Esercizio 2. Sia $L: V \longrightarrow W$ un'applicazione lineare. Dato U un sottospazio di V, la sua immagine $L(U)$ è un sottospazio di W con

$$\dim L(U) \leq \dim U.$$

Sol. Sia U un sottospazio di dimensione k. Supponiamo per assurdo che $\dim L(U) > \dim U$. Allora $L(U)$ ha almeno $k+1$ generatori linearmente indipendenti $w_1 = L(u_1), \ldots, w_{k+1} = L(u_{k+1})$, per qualche $u_1, \ldots, u_{k+1} \in U$. Poiché $\dim U = k$, i vettori $u_1, \ldots, u_{k+1}$ sono linearmente dipendenti. Per l'esercizio precedente, anche i vettori $w_1, \ldots, w_{k+1}$ sono sono linearmente dipendenti, contro l'ipotesi.

Definizione. Sia $L: V \to W$ un'applicazione lineare. L'*immagine di* L è per definizione l'immagine di V tramite L

$$L(V) = \{w \in W \mid w = L(v),\ v \in V\}.$$

Da quanto osservato nell'Esercizio 1 e nell'Esercizio 2, l'immagine $L(V)$ è un sottospazio di W di dimensione

$$\dim L(V) \leq \dim V.$$

Definizione. Un'applicazione $L: V \to W$ si dice *suriettiva* se l'immagine coincide col codominio, ossia $L(V) = W$.

Esercizio 3. Sia $L: V \longrightarrow W$ un' applicazione lineare fra spazi vettoriali di dimensione finita. Se L è *suriettiva*, allora

$$\dim V \geq \dim W.$$

Sol. Poiché $L(V)$ è un sottospazio vettoriale di W, si ha che L è suriettiva, cioè vale $L(V) = W$, se e solo se $\dim L(V) = \dim W$. D'altra parte, poiché L manda insiemi di elementi linearmente dipendenti in insiemi di elementi linearmente dipendenti, $\dim V \geq \dim L(V)$. Mettendo insieme le due osservazioni, si ha che se L è suriettiva, deve valere $\dim V \geq \dim W$.

Conseguenza. *Non esistono* applicazioni lineari suriettive $L: \mathbf{R}^n \longrightarrow \mathbf{R}^m$, se $n < m$.

Abbiamo visto che data un'applicazione lineare $L: V \longrightarrow W$, l'immagine tramite L di un sottospazio U di V è un sottospazio di W con $\dim L(U) \leq \dim U$. È naturale cercare di identificare e caratterizzare quelle applicazioni lineari che mandano sottospazi del dominio in sottospazi del codominio della stessa dimensione. Tali applicazioni sono precisamente le applicazioni lineari iniettive.

Definizione. Un'applicazione $L: V \to W$ si dice *iniettiva* se manda elementi distinti in elementi distinti: $\forall v_1 \neq v_2 \in V \Rightarrow L(v_1) \neq L(v_2)$.)

Definizione. Sia $L\colon V \longrightarrow W$ un'applicazione lineare. Si definisce *nucleo* di L, e si indica con $\ker L$, l'insieme dei vettori di V che hanno immagine O_W tramite L:

$$\ker L = \{v \in V \mid L(v) = O_W\} \subset V.$$

Esercizio 4. Il nucleo $\ker L$ è un sottospazio vettoriale di V.
Sol. Siano $u,\ v \in \ker L$ e $\alpha,\ \beta \in \mathbf{R}$. Poiché L è lineare,

$$L(\alpha u + \beta v) = \alpha L(u) + \beta L(v) = \alpha O_W + \beta O_W = O_W.$$

In altre parole, $\alpha u + \beta v \in \ker L$ e $\ker L$ è un sottospazio vettoriale di V.

Esempio. Se $L_A\colon R^n \longrightarrow \mathbf{R}^m,\ \ L_A(X) = AX$ è l'applicazione data dalla moltiplicazione matrice-vettore, il nucleo $\ker L_A$ è dato dalle soluzioni del sistema lineare omogeneo $AX = O$, con matrice dei coefficienti A.

Proposizione. *Un'applicazione lineare $L\colon V \longrightarrow W$ è iniettiva se e solo se $\ker L = \{O_V\}$.*

Dimostrazione. Se L è iniettiva, per definizione $\ker L = \{O_V\}$. Viceversa, supponiamo che sia $\ker L = \{O_V\}$ e facciamo vedere che L è necessariamente iniettiva. Infatti

$$L(u) = L(v) \Leftrightarrow L(u) - L(v) = L(u - v) = O_W \Leftrightarrow u - v \in \ker L = \{O_V\} \Leftrightarrow u = v.$$

Esercizio 5. Un'applicazione lineare $L\colon V \longrightarrow W$ *iniettiva* manda insiemi di elementi linearmente indipendenti in V in insiemi di elementi linearmente indipendenti in W.
Sol. Siano $\{v_1, \ldots, v_k\}$ elementi linearmente indipendenti in V e siano $L(v_1), \ldots, L(v_k)$ le loro immagini tramite L. Supponiamo per assurdo che $L(v_1), \ldots, L(v_k)$ siano linearmente dipendenti. Per definizione, esistono $\alpha_1, \ldots, \alpha_k$ reali, non tutti nulli, tali che

$$\alpha_1 L(v_1) + \ldots + \alpha_k L(v_k) = O_W.$$

Per la linearità e l'iniettività di L ciò equivale a

$$L(\alpha_1 v_1 + \ldots + \alpha_k v_k) = O_W \quad \Leftrightarrow \quad \alpha_1 v_1 + \ldots + \alpha_k v_k \in \ker L = \{O_V\}.$$

Poiché $\{v_1, \ldots, v_k\}$ sono linearmente indipendenti in V, tutti gli α_i devono essere nulli e questo implica anche la lineare indipendenza degli elementi $L(v_1), \ldots, L(v_k)$ in W.

Esercizio 6. Siano V, W spazi vettoriali di dimensione finita. Se $L\colon V \longrightarrow W$ è un'applicazione lineare *iniettiva*, allora

$$\dim V = \dim L(V) \leq \dim W.$$

Sol. Sia $\{v_1, \ldots, v_n\}$ una base di V. Poiché L è iniettiva, gli elementi $L(v_1), \ldots, L(v_n)$ sono linearmente indipendenti in W, da cui la tesi (in uno spazio vettoriale di dimensione n, ci sono al più n elementi linearmente indipendenti).

Conseguenza. *Non esistono* applicazioni lineari iniettive $L\colon \mathbf{R}^n \longrightarrow \mathbf{R}^m$, se $n > m$.

Proposizione. *Sia* $L: V \longrightarrow W$ *è un'applicazione lineare. Sia* $\dim V = n$. *Allora vale*

$$\dim V = \dim \ker L + \dim L(V).$$

Dimostrazione. Se $\ker L = \{O_V\}$, l'applicazione L è iniettiva e $\dim L(V) = \dim V$, come richiesto. Supponiamo adesso $\ker L \neq \{O_V\}$. Sia $\{v_1, \ldots, v_k\}$ una base del sottospazio $\ker L \subset V$ e siano $v_{k+1}, \ldots, v_n$ elementi di V tali che $\{v_1, \ldots, v_k, v_{k+1}, \ldots, v_n\}$ sia una base di V. Osserviamo innanzitutto che

$$L(V) = \mathrm{span}\{L(v_1), \ldots, L(v_k), L(v_{k+1}), \ldots, L(v_n)\} = \mathrm{span}\{L(v_{k+1}), \ldots, L(v_n)\}.$$

Dunque, i vettori $L(v_{k+1}), \ldots, L(v_n)$ sono generatori di $L(V)$.
Facciamo vedere che i vettori $L(v_{k+1}), \ldots, L(v_n)$ sono linearmente indipendenti in W: infatti

$$\alpha_{k+1}L(v_{k+1}) + \ldots + \alpha_n L(v_n) = O_W \Leftrightarrow L(\alpha_{k+1}v_{k+1} + \ldots + \alpha_n v_n) = O_W$$

$$\Leftrightarrow \alpha_{k+1}v_{k+1} + \ldots + \alpha_n v_n \in \ker L \cap \mathrm{span}\{v_{k+1}, \ldots, v_n\}. \qquad (*)$$

Per costruzione,

$$\ker L \cap \mathrm{span}\{v_{k+1}, \ldots, v_n\} = O_V$$

e, poiché gli elementi $\{v_{k+1}, \ldots, v_n\}$ sono linearmente indipendenti, l'equazione (*) è equivalente a $\alpha_{k+1} = \ldots = \alpha_n = 0$. Dunque $\{L(v_{k+1}), \ldots, L(v_n)\}$ sono una base di $L(V)$. Contando le dimensioni, troviamo

$$\dim V = \dim \ker L + \dim L(V).$$

Corollario. Sia $L: V \longrightarrow W$ un'applicazione lineare fra spazi vettoriali di dimensione finita:

• Se L è iniettiva e suriettiva (cioè *biiettiva*) allora necessariamente $\dim V = \dim W$.

• Se $\dim V = \dim W$, allora L *è iniettiva se e solo se è suriettiva se e solo se è biiettiva.* In particolare, un'applicazione lineare $L: V \longrightarrow V$ di uno spazio vettoriale in sè è iniettiva se e solo se è suriettiva se e solo se è biiettiva.

Se L è un'aplicazione biiettiva, esiste la sua inversa

$$L^{-1}: W \longrightarrow V,$$

cioè l'applicazione che a $w \in W$ associa l'unico elemento $v \in V$ tale che $L(v) = w$. L'applicazione inversa soddisfa le relazioni

$$L^{-1} \circ L = Id_V, \qquad L \circ L^{-1} = Id_W;$$

inoltre, se L è lineare anche L^{-1} è lineare. In questo caso si dice anche che $L: V \longrightarrow W$ è un *isomorfismo*.

Esempio. Sia V uno spazio vettoriale di dimensione n e sia $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base di V. L'applicazione che a ogni elemento $v \in V$ associa le coordinate di v in $\mathcal{B}$

$$v = x_1 v_1 + \ldots + x_n v_n \mapsto \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

è un isomorfismo di V con $\mathbf{R}^n$.

Applicazioni lineari e matrici.

Sia M una matrice $m \times n$

$$M = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

Definizione. Il *rango per colonne di M* è per definizione il massimo numero di colonne linearmente indipendenti di M. Il *rango per righe di M* è il massimo numero di righe linearmente indipendenti di M.

In questa sezione dimostreremo che il rango per righe e il rango per colonne di M coincidono, un risultato non cosi' evidente a priori visto che le righe di M sono vettori di $\mathbf{R}^n$, mentre le colonne sono vettori di $\mathbf{R}^m$. Dimostreremo il risultato sfruttando il fatto che il rango per righe e il rango per colonne di M sono legati rispettivamente alla dimensione del nucleo e alla dimensione dell'immagine dell'applicazione lineare

$$L_M \colon \mathbf{R}^n \longrightarrow \mathbf{R}^m, \quad X \mapsto MX,$$

determinata da M. Ricordiamo che le colonne della matrice M contengono le coordinate (nella base canonica di $\mathbf{R}^m$) dei trasformati dei vettori della base canonica di $\mathbf{R}^n$:

$$L_M(\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}) = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \ L_M(\begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}) = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, \dots\dots, L_M(\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}) = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}. \qquad (*)$$

Ne segue che la dimensione dell'immagine $L_M(\mathbf{R}^n)$ coincide col massimo numero di colonne linearmente indipendenti di M. In altre parole, *il rango per colonne della matrice M è uguale alla dimensione dell'immagine dell'applicazione lineare $L_M \colon \mathbf{R}^n \longrightarrow \mathbf{R}^m$, da essa determinata.*

Come abbiamo già osservato nella sezione precedente, il nucleo dell'applicazione lineare L_M è dato dalle soluzioni del sistema lineare omogeneo $MX = O$, con matrice dei coefficienti M. Il rango per righe di M è uguale al numero di righe non nulle di una qualunque ridotta a scala S di M mediante il metodo di eliminazione di Gauss per righe. Risolvendo per sostituzione il sistema a scala equivalente con matrice S, le soluzioni dipendono da un numero di parametri liberi uguale al numero delle incognite meno il rango per righe di M. Ne segue che *la dimensione del nuleo di L_M risulta minore o uguale ad n meno il rango per righe di M.* (Attenzione: mettendo in evidenza i parametri liberi nella soluzione generale trovata per sostituzione, si trovano dei generatori della soluzione generale. A priori non è chiaro che si tratti di generatori linearmente indipendenti).

Teorema. *Sia M una matrice $m \times n$. Il rango per righe di M è uguale al rango per colonne di M.*

Dimostrazione. Sia $L_M \colon \mathbf{R}^n \longrightarrow \mathbf{R}^m, \quad X \mapsto MX$ l'applicazione lineare associata ad M. Vale la relazione

$$\dim \mathbf{R}^n = \dim \ker(L_M) + \dim L_M(\mathbf{R}^n),$$

dove $\dim L_M(\mathbf{R}^n) = \text{rango}_{col}(M)$. Dunque

$$\dim \ker(L_M) = n - \text{rango}_{col}(M).$$

D'altra parte, poiché $\dim \ker(L_M)$ è uguale alla dimensione dello spazio delle soluzioni del sistema lineare omogeneo $MX = O$, si ha anche che

$$\dim \ker(L_M) \leq n - \text{rango}_{righe}(M),$$

da cui segue che

$$\text{rango}_{righe}(M) \leq \text{rango}_{col}(M).$$

Ripetendo gli stessi ragionamenti per la matrice trasposta ${}^tM$ e per l'applicazione da essa determinata $L_{{}^tM} \colon \mathbf{R}^m \longrightarrow \mathbf{R}^n$, troviamo

$$\text{rango}_{righe}({}^tM) \leq \text{rango}_{col}({}^tM) \quad \Leftrightarrow \quad \text{rango}_{righe}(M) \geq \text{rango}_{col}(M).$$

Dunque vale anche la disuguaglianza opposta e il teorema segue.

• Il rango per righe, che è uguale al rango per colonne di M, si chiama semplicemente il *rango di* M e si indica con $r(M)$. È chiaro che $r(M) \leq \min\{m, n\}$.

Applicazioni lineari, matrici e sistemi lineari.

Citiamo alcune conseguenze del teorema precedente nella teoria dei sistemi lineari.

Proposizione. *Sia $MX = O$ un sistema lineare omogeneo di m equazioni in n incognite. Allora il sottospazio delle soluzioni* $\text{Sol}(MX = O)$ *ha dimensione uguale a $n - r(M)$, dove $r(M)$ è il rango di M. In particolare, lo spazio delle soluzioni di un sistema lineare omogeneo a scala, di r equazioni (non banali) in n incognite ha dimensione $n - r$.*

Dimostrazione. Sia $L_M \colon \mathbf{R}^n \longrightarrow \mathbf{R}^m \quad X \mapsto MX$ l'applicazione lineare associata ad M. Abbiamo che $\text{Sol}(MX = O) = \ker(L_M)$. Inoltre vale la relazione

$$\dim \mathbf{R}^n = \dim \ker(L_M) + \dim L_M(\mathbf{R}^n).$$

Ciò equivale a dire che

$$\dim \text{Sol}(MX = O) = n - \text{rango}_{col}(M) = n - \text{rango}_{righe}(M) = n - r(M).$$

Esempio. Lo spazio delle soluzioni di un sistema lineare omogeneo *a scala*, di 2 equazioni (non banali) in 6 incognite ha dimensione 4. Lo spazio delle soluzioni di un sistema lineare omogeneo *a scala*, di 3 equazioni (non banali) in 5 incognite ha dimensione 2.

Corollario. Per esprimere un sottospazio vettoriale di dimensione k in $\mathbf{R}^n$, ci vogliono precisamente $n - k$ equazioni indipendenti.

Esempio. Per esprimere un sottospazio vettoriale di dimensione 1 in $\mathbf{R}^n$ (una retta per l'origine) ci vogliono precisamente $n - 1$ equazioni indipendenti, Per esprimere un sottospazio vettoriale di dimensione 2 in $\mathbf{R}^n$ (un piano per l'origine) ci vogliono precisamente $n - 2$ equazioni indipendenti.

Un piano per l'origine in $\mathbf{R}^3$ è dato dalle soluzioni di un'equazione lineare omogenea in 3 variabili; una retta per l'origine in $\mathbf{R}^3$ è data dalle soluzioni un sistema di due equazioni omogenee indipendenti in tre variabili: geometricamente significa che una retta per l'origine in $\mathbf{R}^3$ si esprime come intersezione di due piani distinti (per l'origine).

Un piano per l'origine in $\mathbf{R}^4$ è dato dalle soluzioni un sistema di due equazioni omogenee indipendenti in quattro variabili.

Teorema. *(Rouché-Capelli). Sia $MX = b$ un sistema lineare di m equazioni in n incognite. Il sistema è compatibile se e solo se*

$$r(M) = r(M|b).$$

Dimostrazione. Sia $L_M \colon \mathbf{R}^n \longrightarrow \mathbf{R}^m, \ \ X \mapsto MX$ l'applicazione lineare associata alla matrice M. È chiaro che il sistema è compatibile se e solo se il vettore dei termini noti b appartiene all'immagine di $L_M(\mathbf{R}^n) \subset \mathbf{R}^m$. Se chiamiamo $C_1, \ldots, C_n$ le colonne di M e riscriviamo

$$MX = x_1C_1 + \ldots + x_nC_n = b, \qquad X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

è evidente che il sistema $MX = b$ è compatibile se e solo se

$$b \in \operatorname{span}\{C_1, \ldots, C_n\} \quad \Leftrightarrow \quad \operatorname{rango}(M) = \operatorname{rango}(M|b). \tag{$*$}$$

Concludiamo con una osservazione sulla struttura delle soluzioni di un sistema lineare non omogeneo compatibile:

Sia ξ una soluzione del sistema $MX = b$. Allora ξ è della forma

$$\xi \in \xi_0 + \operatorname{Sol}(MX = O),$$

ove ξ_0 è una soluzione fissata del sistema, i.e. $M\xi_0 = b$.

In altre parole, se ξ_1, ξ_2 sono soluzioni del sistema $MX = b$, allora la loro differenza $\xi_1 - \xi_2$ è soluzione del sistema lineare omogeneo associato $MX = O$.

Esempio. Le soluzioni del sistema lineare non omogeneo in tre incognite

$$\begin{cases} x + 2y - z = 1 \\ y + z = -1 \end{cases}$$

sono date dall'insieme

$$\{ \begin{pmatrix} 3+3z \\ -1-z \\ z \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \\ 0 \end{pmatrix} + z \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix}, \quad z \in \mathbf{R}\}$$

e sono costituite dai punti di una retta passante per $\begin{pmatrix} 3 \\ -1 \\ 0 \end{pmatrix}$, parallela alla retta per l'origine $\operatorname{span}\{ \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix} \}$.

Matrici rappresentative di un'applicazione lineare.
Sia $L\colon V \longrightarrow W$ un'applicazione lineare fra spazi vettoriali di dimensione finita:

$$\dim V = n, \qquad \dim W = m.$$

Ricordiamo che fissate una base $\mathcal{B} = \{v_1, \ldots, v_n\}$ in V e una base $\mathcal{B}' = \{w_1, \ldots, w_m\}$ in W si ottengono un'identificazione di V con $\mathbf{R}^n$ e un'identificazione di W con $\mathbf{R}^m$, date rispettivamente da:

$$v = x_1 v_1 + \ldots + x_n v_n \leftrightarrow X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \qquad w = y_1 w_1 + \ldots + y_m w_m \leftrightarrow Y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}.$$

Il fatto importante è che L, vista come applicazione lineare fra $\mathbf{R}^n$ ed $\mathbf{R}^m$, *coincide con la moltiplicazione per una opportuna matrice:* $X \mapsto MX$.

Proposizione. *Esiste una matrice $M = \{a_{ij}\}_{i,j}$ reale $m \times n$ che moltiplicata per il vettore delle coordinate di $v = \sum x_i v_i$ nella base $\mathcal{B}$ restituisce il vettore delle coordinate dell'immagine $L(v) = \sum y_j w_j$ nella base $\mathcal{B}'$*

$$\begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

La matrice M dipende dalla scelta delle basi $\mathcal{B}$ e $\mathcal{B}'$ e si chiama la matrice rappresentativa di L rispetto a $\mathcal{B}$ e $\mathcal{B}'$.

Dimostrazione. Siano infatti

$$L(v_1) = \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix}, \ldots, L(v_n) = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

i vettori delle coordinate di $L(v_1), \ldots, L(v_n)$ nella base $\mathcal{B}'$ di W. Sia $v = x_1 v_1 + \ldots + x_n v_n \in V$. Poiché $L(v) = x_1 L(v_1) + \ldots x_n L(v_n)$ (per la linearità di L), le coordinate di $L(v)$ in $\mathcal{B}'$ sono date da

$$x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + \ldots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \ldots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n \end{pmatrix} = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

La matrice rappresentativa di L in $\mathcal{B}$ e $\mathcal{B}'$ è dunque la matrice

$$M = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix}.$$

Osserviamo che M *che ha nelle sue colonne le coordinate dei trasformati degli elementi della base $\mathcal{B}$*

$$\{L(v_1), \ldots, L(v_n)\},$$

nella base $\mathcal{B}'$.

Esempio. Sia $L\colon \mathbf{R}^2 \longrightarrow \mathbf{R}^3$ l'applicazione lineare data da $L(\begin{pmatrix} x \\ y \end{pmatrix}) = \begin{pmatrix} x+y \\ y \\ x+2y \end{pmatrix}$.

• La matrice rappresentativa di L nelle basi canoniche $\mathcal{C}$ di $\mathbf{R}^2$ e $\mathcal{C}$ di $\mathbf{R}^3$ è data da

$$M = M_{\mathcal{C},\mathcal{C}} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 2 \end{pmatrix}.$$

Le colonne di $M_{\mathcal{C},\mathcal{C}}$ contengono le coordinate dei vettori

$$L(\begin{pmatrix} 1 \\ 0 \end{pmatrix}) = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \qquad L(\begin{pmatrix} 0 \\ 1 \end{pmatrix}) = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$$

nella base canonica di $\mathbf{R}^3$.

• La matrice rappresentativa di L nelle basi canoniche $\mathcal{C}$ di $\mathbf{R}^2$ e $\mathcal{B}'$ di $\mathbf{R}^3$ è data da

$$M_{\mathcal{C},\mathcal{B}'} = \begin{pmatrix} 1/2 & 1 \\ 1/3 & 2/3 \\ 1/6 & -2/3 \end{pmatrix}, \qquad \mathcal{B}' = \{\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}\}.$$

Le colonne di $M_{\mathcal{C},\mathcal{B}'}$ contengono le coordinate dei vettori $L(\begin{pmatrix} 1 \\ 0 \end{pmatrix})$ e $L(\begin{pmatrix} 0 \\ 1 \end{pmatrix})$ nella base $\mathcal{B}'$.

• La matrice rappresentativa di L nelle basi canoniche $\mathcal{B}$ di $\mathbf{R}^2$ e $\mathcal{C}$ di $\mathbf{R}^3$ è data da

$$M_{\mathcal{B},\mathcal{C}} = \begin{pmatrix} 0 & 3 \\ -1 & 2 \\ -1 & 5 \end{pmatrix}, \qquad \mathcal{B} = \{\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}\}.$$

Le colonne di $M_{\mathcal{B},\mathcal{C}}$ contengono le coordinate dei vettori

$$L(\begin{pmatrix} 1 \\ -1 \end{pmatrix}) = \begin{pmatrix} 0 \\ -1 \\ -1 \end{pmatrix}, \qquad L(\begin{pmatrix} 1 \\ 2 \end{pmatrix}) = \begin{pmatrix} 3 \\ 2 \\ 5 \end{pmatrix}$$

nella base canonica di $\mathbf{R}^3$.

• La matrice rappresentativa di L nelle basi canoniche $\mathcal{B}$ di $\mathbf{R}^2$ e $\mathcal{B}'$ di $\mathbf{R}^3$ è data da

$$M_{\mathcal{B},\mathcal{B}'} = \begin{pmatrix} -3 & -3 \\ -1 & -1 \\ 3 & 4 \end{pmatrix}, \qquad \mathcal{B} = \{\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}\}, \quad \mathcal{B}' = \{\begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}\};$$

Le colonne di $M_{\mathcal{B},\mathcal{B}'}$ contengono le coordinate dei vettori

$$L(\begin{pmatrix} 1 \\ 1 \end{pmatrix}) = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}, \qquad L(\begin{pmatrix} 2 \\ 1 \end{pmatrix}) = \begin{pmatrix} 3 \\ 1 \\ 4 \end{pmatrix}$$

nella base $\mathcal{B}'$ di $\mathbf{R}^3$.

Esercizio. Per la stessa applicazione lineare L, calcolare

$$M_{\mathcal{C},\mathcal{B}'}, \qquad \mathcal{B}' = \{\begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}\};$$

$$M_{\mathcal{B},\mathcal{B}'}, \qquad \mathcal{B} = \{\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -3 \\ 1 \end{pmatrix}\}, \quad \mathcal{B}' = \{\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}\};$$

$$M_{\mathcal{B},\mathcal{C}}, \qquad \mathcal{B} = \{\begin{pmatrix} 3 \\ 1 \end{pmatrix}, \begin{pmatrix} -3 \\ 0 \end{pmatrix}\};$$

Esempio. Sia $L = Id_V : V \longrightarrow V$ l'applicazione *identità*

$$Id_V(v) = v, \quad \forall v \in V.$$

Siano $\mathcal{B}$ e $\mathcal{B}'$ due basi di V. La matrice rappresentativa dell'applicazione identità rispetto a $\mathcal{B}$ e $\mathcal{B}'$ si chiama la *matrice del cambiamento di base da* $\mathcal{B}$ *a* $\mathcal{B}'$. Questa matrice moltiplicata per il vettore delle coordinate di un elemento v in $\mathcal{B}$ restituisce le coordinate dello stesso vettore in $\mathcal{B}'$.

Esempio. Sia $\mathcal{B}$ una base di $\mathbf{R}^n$, formata da n vettori le cui coordinate nella base canonica $\mathcal{C}$ sono date da

$$\mathcal{B} = \{\begin{pmatrix} a_{11} \\ a_{21} \vdots \\ a_{n1} \end{pmatrix}, \begin{pmatrix} a_{12} \\ a_{22} \vdots \\ a_{n2} \end{pmatrix}, \ldots, \begin{pmatrix} a_{1n} \\ a_{2n} \vdots \\ a_{nn} \end{pmatrix}\}.$$

La matrice $C_{\mathcal{B},\mathcal{C}}$ del cambiamento di base, dalla base $\mathcal{B}$ alla base canonica $\mathcal{C}$, ha per colonne le coordinate dei vettori della base $\mathcal{B}$ nella base canonica $\mathcal{C}$. Risulta pertanto

$$C_{\mathcal{B},\mathcal{C}} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}.$$

La matrice $C_{\mathcal{C},\mathcal{B}}$ del cambiamento di base, dalla base canonica $\mathcal{C}$ alla base $\mathcal{B}$, ha per colonne le coordinate dei vettori della base canonica $\mathcal{C}$ nella base $\mathcal{B}$. Per determinare questa matrice conviene osservare che è l'inversa della matrice $C_{\mathcal{B},\mathcal{C}}$ del cambiamento di base, dalla base $\mathcal{B}$ alla base canonica $\mathcal{C}$ (quella più facile da determinare). Risulta pertanto

$$C_{\mathcal{C},\mathcal{B}} = C^{-1}_{\mathcal{B},\mathcal{C}} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}^{-1}.$$

Esercizio. Calcolare le matrici dei seguenti cambiamenti di base:

$$V = \mathbf{R}^2, \quad \mathcal{B} = \mathcal{C}, \quad \mathcal{B}' = \{\begin{pmatrix} 3 \\ 1 \end{pmatrix}, \begin{pmatrix} -3 \\ 0 \end{pmatrix}\};$$

$$V = \mathbf{R}^2, \quad \mathcal{B} = \{\begin{pmatrix} 3 \\ 1 \end{pmatrix}, \begin{pmatrix} -3 \\ 0 \end{pmatrix}\}, \quad \mathcal{B}' = \mathcal{C};$$

$$V = \mathbf{R}^3, \quad \mathcal{B} = \mathcal{C}, \quad \mathcal{B}' = \{\begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}\};$$

$$V = \mathbf{R}^3, \quad \mathcal{B} = \{\begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}\}, \quad \mathcal{B}' = \mathcal{C};$$

Problema. Sia $L\colon \mathbf{R}^n \longrightarrow \mathbf{R}^m$ l'applicazione data dalla matrice M, rispetto alle basi canoniche di $\mathbf{R}^n$ ed $\mathbf{R}^m$. Siano $\mathcal{B}$ e $\mathcal{B}'$ altre due basi di $\mathbf{R}^n$ e di $\mathbf{R}^m$ rispettivamente. Come determinare la matrice $\widetilde{M}$ di L rispetto alle basi $\mathcal{B}$ e $\mathcal{B}'$? Date le coordinate di un vettore $v \in \mathbf{R}^n$ nella base $\mathcal{B}$, vogliamo la matrice che ci restituisce le coordinalte di $L(v) \in \mathbf{R}^m$ nella base $\mathcal{B}'$, conoscendo la matrice M che partendo dalle coordinate di $v \in \mathbf{R}^n$ nella base canonica ci restituisce le coordinalte di $L(v) \in \mathbf{R}^m$ nella base canonica. Un metodo è quello di considerare il diagramma commutativo

$$\begin{tikzcd} \mathbf{R}^n, \mathcal{C} \arrow[r, "M"] & \mathbf{R}^m, \mathcal{C} \arrow[d, "C_{\mathcal{C},\mathcal{B}'}"] \\ \mathbf{R}^n, \mathcal{B} \arrow[u, "C_{\mathcal{B},\mathcal{C}}"'] \arrow[r, "\widetilde{M}=?"] & \mathbf{R}^m, \mathcal{B}' \end{tikzcd}$$

da cui segue che

$$\widetilde{M} = C_{\mathcal{C},\mathcal{B}'} M C_{\mathcal{B},\mathcal{C}}.$$

Il significato del diagramma è questo: dalle coordinate di v in $\mathcal{B}$ troviamo le coordinate di v nella base canonica di $\mathbf{R}^n$ mediante la matrice del cambiamento di base $C_{\mathcal{B},\mathcal{C}}$; poi applicando M troviamo le coordinate di $L(v)$ nella base canonica di $\mathbf{R}^m$; infine applicando la matrice del cambiamento di base $C_{\mathcal{C},\mathcal{B}'}$, troviamo le coordinate di $L(v)$ nella base $\mathcal{B}'$ di $\mathbf{R}^m$.
Attenzione: osservare l'ordine con cui agiscono le varie matrici e l'ordine con cui sono scritte nella formula che da $\widetilde{M}$.

Esercizio. Determinare la matrice M di L, rispetto alle basi canoniche di $\mathbf{R}^n$ ed $\mathbf{R}^m$, conoscendo la matrice $\widetilde{M}$ di L rispetto alle basi $\mathcal{B}$ e $\mathcal{B}'$.
Suggerimento: considerare questa volta il diagramma

$$\begin{tikzcd} \mathbf{R}^n, \mathcal{C} \arrow[r, "M=?"] \arrow[d, "C_{\mathcal{C},\mathcal{B}}"] & \mathbf{R}^m, \mathcal{C} \\ \mathbf{R}^n, \mathcal{B} \arrow[r, "\widetilde{M}"] & \mathbf{R}^m, \mathcal{B}' \arrow[u, "C_{\mathcal{B}',\mathcal{C}}"'] \end{tikzcd}$$

Osservazione. Sia $L\colon V \longrightarrow W$ un'applicazione lineare e sia M una sua qualunque matrice rappresentativa. Allora

$$\dim \mathrm{Im} L = r(M), \qquad \dim \ker L = \dim V - r(M).$$

Applicazioni lineari invertibili.

Sia $L_M\colon \mathbf{R}^n \longrightarrow \mathbf{R}^n, \quad X \mapsto MX$ un'applicazione lineare biiettiva, dove M è una matrice quadrata $n \times n$. L'inversa di L_M è lineare ed è data da $(L_M)^{-1} = L_A$, ossia dalla moltiplicazione matrice-vettore per una matrice A quadrata $n \times n$ che soddisfa le relazioni

$$AM = MA = I_n.$$

La matrice A è per definizione l'inversa di M, e si indica con M^{-1}.

Calcolo dell'inversa di una matrice con l'eliminazione di Gauss.

Sia M una matrice quadrata $n \times n$ a coefficienti reali o complessi:

$$M = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

La matrice M si dice invertibile se esiste una matrice X quadrata $n \times n$ che soddisfa

$$M \cdot X = X \cdot M = I_n, \tag{1}$$

dove I_n è la matrice identità di ordina n. In tal caso la matrice X si dice l'inversa di M e viene indicata con M^{-1}.

Osserviamo che se X esiste, allora è unica: infatti se X e Y soddifano la (1), vale

$$X = X \cdot I_n = X \cdot (M \cdot Y) = (X \cdot M) \cdot Y = I_n \cdot Y = Y.$$

Si può anche dimostrare che se una matrice X soddisfa $M \cdot X = I_n$, allora soddisfa necessariamente anche $X \cdot M = I_n$, cioè è proprio l'inversa cercata.

Se M è invertibile, come determinare l'inversa di M?

Siano $X_1,\ X_2, \dots, X_n$ le n colonne della matrice X e siano $E_1\ E_2 \dots E_n$ le n colonne della matrice I_n, date rispettivamente dai vettori

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Allora l'equazione matriciale $M \cdot X = I_n$ si può riscrivere come

$$\begin{pmatrix} | & | & & | \\ MX_1 & MX_2 & \dots & MX_n \\ | & | & & | \end{pmatrix} = \begin{pmatrix} | & | & & | \\ E_1 & E_2 & \dots & E_n \\ | & | & & | \end{pmatrix}.$$

Quando esiste, l'inversa di M è la matrice le cui colonne $X_1,\ X_2, \dots, X_n$ soddisfano rispettivamente le equazioni

$$M \cdot X_i = E_i \quad \Longleftrightarrow \quad \begin{cases} a_{11}x_{i1} & + & \dots & + & a_{1n}x_{in} & = & 0 \\ \vdots & & & & \vdots & & \vdots \\ a_{i1}x_{i1} & + & \dots & + & a_{in}x_{in} & = & 1 \\ \vdots & & & & \vdots & & \vdots \\ a_{n1}x_{i1} & + & \dots & + & a_{nn}x_{in} & = & 0 \end{cases} \tag{2}$$

Al variare di $i = 1, \ldots, n$, si hanno n sistem lineari di n equazioni le cui n incognite sono gli elementi della i-sima colonna della matrice inversa X

$$X_i = \begin{pmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{in} \end{pmatrix}.$$

Poiché questi sistemi hanno in comune la matrice dei coefficienti M e differiscono solo per i termini noti, possiamo applicare l'eliminazione di Gauss a tutti simultaneamente. Completando la matrice M con le varie colonne dei termini noti, otteniamo

$$M|I_n = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} & | & 1 & 0 & \ldots & 0 \\ a_{21} & a_{22} & \ldots & a_{2n} & | & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots & | & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} & | & 0 & 0 & \ldots & 1 \end{pmatrix}.$$

Procedendo nell'eliminazione di Gauss dall'alto in basso come al solito, ad un certo punto avremo ridotto M a scala, cioè avremo davanti una matrice della forma

$$\begin{pmatrix} s_{11} & s_{12} & \ldots & s_{1n} & | & * & * & \ldots & * \\ 0 & s_{22} & \ldots & s_{2n} & | & * & * & \ldots & * \\ \vdots & \vdots & \ddots & \vdots & | & \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & s_{nn} & | & * & * & \ldots & * \end{pmatrix}. \tag{3}$$

Affinché M sia invertibile, gli n sistemi (2) devono avere soluzione *unica*: questo avviene se e solo se gli elementi

$$s_{11},\ s_{22}, \ldots,\ s_{nn} \neq 0, \quad i = 1, \ldots, n.$$

Le soluzioni degli n sistemi formano le colonne della matrice inversa.

Osservazione. *la matrice M è invertibile se e solo se ogni sua ridotta a scala ha tutti gli elementi sulla diagonale principale diversi da zero se e solo se ha rango massimo* $\text{rango}(M) = n$.

Osserviamo che un sistema lineare quadrato non omogeneo compatibile, con soluzione unica, ridotto a scala, può essere risolto procedendo direttamente per sostituzione, oppure proseguendo con l'eliminazione di Gauss "dal basso in alto" fino ad ottenere una matrice del tipo

$$\begin{pmatrix} 1 & 0 & \ldots & 0 & | & \alpha_1 \\ 0 & 1 & \ldots & 0 & | & \alpha_2 \\ \vdots & \vdots & \ddots & \vdots & | & \vdots \\ 0 & 0 & \ldots & 1 & | & \alpha_n \end{pmatrix}.$$

Il sistema di equazioni corrispondente è $x_1 = \alpha_1,\ x_2 = \alpha_2, \ldots,\ x_n = \alpha_n$. In altre parole, se la matrice dei coefficienti è stata ridotta all'identità mediante operazioni ammissibili sulle righe, la colonna dei termini noti contiene le soluzioni del sistema.

Applicando l'eliminazione di Gauss "da sotto in su" alla *famiglia* di sistemi (3), arriveremo ad una matrice della forma $(I_n|X_0)$ ove X_0 è esattamente l'inversa cercata.

Esempio 1. Sia $M = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$. Vediamo se innanzitutto se è una matrice invertibile. Procedendo con l'eliminazione di Gauss dall'alto in basso sulla matrice $\begin{pmatrix} 1 & 2 & | & 1 & 0 \\ 2 & 4 & | & 0 & 1 \end{pmatrix}$, troviamo $\begin{pmatrix} 1 & 2 & | & 1 & 0 \\ 0 & 0 & | & -2 & 1 \end{pmatrix}$. Dunque M non è invertibile.

Esempio 2. Sia $M = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}$. Procedendo con l'eliminazione di Gauss dall'alto in basso sulla matrice $\begin{pmatrix} 1 & 2 & | & 1 & 0 \\ -1 & 1 & | & 0 & 1 \end{pmatrix}$ troviamo $\begin{pmatrix} 1 & 2 & | & 1 & 0 \\ 0 & 3 & | & 1 & 1 \end{pmatrix}$. Questa volta M è invertibile e possiamo quindi procedere al calcolo dell'inversa. Sostituendo la prima riga R_1 con $R_1 - \frac{2}{3}R_2$ troviamo $\begin{pmatrix} 1 & 0 & | & 1/3 & -2/3 \\ 0 & 3 & | & 1 & 1 \end{pmatrix}$ e finalmente, dividento la seconda riga per 3, otteniamo

$$\begin{pmatrix} 1 & 0 & | & 1/3 & -2/3 \\ 0 & 1 & | & 1/3 & 1/3 \end{pmatrix}.$$

La matrice $X_0 = \begin{pmatrix} 1/3 & -2/3 \\ 1/3 & 1/3 \end{pmatrix}$ è l'inversa di M. *Prova:* $X_0 \cdot M = I_2$.

8. Determinanti (cenni).

Sia M una matrice quadrata $n \times n$ e sia $L_M\colon \mathbf{R}^n \longrightarrow \mathbf{R}^n$, $X \mapsto MX$ l'applicazione lineare associata. L'applicazione L_M è biiettiva (se e solo se è iniettiva se e solo è suriettiva) se e solo se la matrice M ha rango massimo $r(M) = n$. In questo caso, L_M ammettte inversa e l'inversa è data dalla moltiplicazione per la matrice inversa $(L_M)^{-1} = L_{M^{-1}}$. La funzione *determinante* permette di caratterizzare le matrici invertibili direttamente in termini dei coefficienti. Inoltre, grazie alle proprietà della funzione determinante, è possibile esprimere l'inversa di una matrice in funzione dei suoi coefficienti.

Sia $M(n, n, \mathbf{R})$ lo spazio delle matrici quadrate $n \times n$ a coefficienti reali (la stessa teoria vale per le matrici a coefficienti complessi). Il *determinante* è la funzione $\det\colon M(n, n, \mathbf{R}) \longrightarrow \mathbf{R}$

$$M = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \quad \mapsto \quad \det M = \sum_\sigma \epsilon(\sigma) a_{1\sigma(1)} \dots a_{n\sigma(n)},$$

ossia il polinomio omogeneo di grado n nei coefficienti della matrice dato dalla somma di tutti i possibili i prodotti di coefficienti che stanno su righe e su colonne distinte (sono $n!$ termini) con un segno $+$ o $-$ davanti che dipende dalla posizione dei coefficienti $a_{1\sigma(1)}, \dots, a_{n\sigma(n)}$ nella matrice ($\epsilon(\sigma) = \pm 1$ è il segno della permutazione σ di $\{1, \dots, n\}$). Ad esempio,

$n = 1, \quad \det(a_{11}) = a_{11};$

$n = 2, \quad \det(\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}) = a_{11}a_{22} - a_{12}a_{21};$

$n = 3, \quad \det(\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}) =$

$$= a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{12}a_{23}a_{31} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{23}a_{32}a_{11}.$$

• È chiaro dalla definizione che il determinante di una matrice è una funzione molto complicata da calcolare, a meno che non si tratti di una matrice speciale: ad esempio, il determinante di una matrice triangolare superiore, di una matrice triangolare inferiore o di una matrice diagonale è dato dal prodotto degli elementi sulla diagonale principale:

$$\det \begin{pmatrix} m_{11} & m_{12} & \dots & m_{1n} \\ 0 & m_{22} & \dots & m_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & m_{nn} \end{pmatrix} = \det \begin{pmatrix} m_{11} & 0 & \dots & 0 \\ m_{21} & m_{22} & \dots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ m_{n1} & \dots & \dots & m_{nn} \end{pmatrix} =$$

$$= \det \begin{pmatrix} m_{11} & 0 & \dots & 0 \\ 0 & m_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & m_{nn} \end{pmatrix} = m_{11}m_{22}\dots m_{nn}.$$

Inoltre, il determinante di una matrice con una riga o una colonna nulla è uguale a 0:

$$\det \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} = 0, \quad \det \begin{pmatrix} a_{11} & \dots & 0 & \dots & a_{1n} \\ a_{21} & \dots & 0 & \dots & a_{2n} \\ \dots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \dots & 0 & \dots & a_{nn} \end{pmatrix} = 0$$

• A partire dalla definizione, si può dimostrare che il determinante gode delle seguenti proprietà:

(1) Il determinante della matrice identità è uguale a 1: $\det(I_n) = 1$;

(2) *(Alternanza)* Scambiando fra loro due righe (o due colonne) il determinante cambia di segno ;

$$\det(\begin{pmatrix} \dots & \dots & \dots \\ \dots & R_i & \dots \\ \dots & \dots & \dots \\ \dots & R_j & \dots \\ \dots & \dots & \dots \end{pmatrix}) = -\det(\begin{pmatrix} \dots & \dots & \dots \\ \dots & R_j & \dots \\ \dots & \dots & \dots \\ \dots & R_i & \dots \\ \dots & \dots & \dots \end{pmatrix});$$

$$\det(\begin{pmatrix} \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & C_i & \vdots & C_j & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}) = -\det(\begin{pmatrix} \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & C_j & \vdots & C_i & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}).$$

In particolare, una matrice M con due righe o due colonne uguali ha determinante uguale a 0: infatti scambiando fra loro le due rige (o le due colonne) uguali la matrice resta invariata, ma al tempo stesso si ha

$$\det(M) = -\det(M) \quad \Leftrightarrow \quad \det(M) = 0.$$

(3) *(Multilinearità)* Se la i-sima riga (o la i-sima colonna) di una matrice è data dalla combinazione lineare di due vettori $\alpha T_i + \beta S_i$, allora il determinante si spezza nel modo seguente:

$$\det(\begin{pmatrix} \dots & \dots & \dots \\ \dots & \alpha T_i + \beta S_i & \dots \\ \dots & \dots & \dots \end{pmatrix}) = \alpha \det(\begin{pmatrix} \dots & \dots & \dots \\ \dots & T_i & \dots \\ \dots & \dots & \dots \end{pmatrix}) + \beta \det(\begin{pmatrix} \dots & \dots & \dots \\ \dots & S_i & \dots \\ \dots & \dots & \dots \end{pmatrix});$$

$$\det(\begin{pmatrix} \vdots & \vdots & \vdots \\ \vdots & \alpha T_i + \beta S_i & \vdots \\ \vdots & \vdots & \vdots \end{pmatrix}) = \alpha \det(\begin{pmatrix} \vdots & \vdots & \vdots \\ \vdots & T_i & \vdots \\ \vdots & \vdots & \vdots \end{pmatrix}) + \beta \det(\begin{pmatrix} \vdots & \vdots & \vdots \\ \vdots & S_i & \vdots \\ \vdots & \vdots & \vdots \end{pmatrix}).$$

• Conseguenze di (1)(2)(3):

(i) sia S una matrice a scala ottenuta da M mediante l'eliminazione di Gauss per righe, *limitata a scambio di righe e sostituzione di una riga R_i con $R_i + \alpha R_j$, per $\alpha \in \mathbf{R}$ e $i \neq j$.* Allora

$$\det(M) = \begin{cases} \det(S) & \text{se } S \text{ è ottenuta con un numero pari di scambi di righe} \\ -\det(S) & \text{se } S \text{ è ottenuta con un numero dispari di scambi di righe.} \end{cases}$$

Questo fatto offre un metodo effettivo per calcolare il determinante di una matrice M di ordine arbitrario.

(ii) Se una riga (o una colonna) di M è combinazione lineare delle rimanenti, allora

$$\det(M) = 0.$$

Vale anche il viceversa: se $\det(M) = 0$ allora $r(M) < n$ e le sue righe (e le sue colonne) sono linearmente dipendenti.

(iii) Una matrice M di ordine n è invertibile se e solo se

$$r(M) = n \quad \Leftrightarrow \quad \det(M) \neq 0.$$

Tramite il determinante si ottiene dunque una una condizione necessaria e sufficiente all'invertibilità di M, direttamente in termini dei suoi coefficienti.

• Sia $M = \{a_{ij}\}_{ij}$ una matrice $n \times n$. Fissato a_{ij}, il suo complemento algebrico (o cofattore) A_{ij} è per definizione il prodotto di $(-1)^{i+j}$ per il determinante della matrice ottenuta da M eliminando la i-sima riga e la j-sima colonna:

$$A_{ij} = (-1)^{i+j} \det \begin{pmatrix} a_{11} & \dots & * & \dots & a_{1n} \\ a_{21} & \dots & * & \dots & a_{2n} \\ \dots & \dots & * & \dots & \dots \\ ** & ** & ** & ** & ** \\ \dots & \dots & * & \dots & \dots \\ a_{n1} & \dots & * & \dots & a_{nn} \end{pmatrix}$$

Teorema di Laplace. *Valgono le seguenti identità:*

(1) Sviluppo del determinante lungo la i-sima riga:

$$\det(M) = a_{i1}A_{i1} + a_{i2}A_{i2} + \ldots + a_{in}A_{in}, \qquad 1 \leq i \leq n;$$

(2) Sviluppo del determinante lungo la k-sima colonna:

$$\det(M) = a_{1k}A_{1k} + a_{2k}A_{2k} + \ldots + a_{nk}A_{nk}, \qquad 1 \leq k \leq n.$$

(3)

$$a_{i1}A_{j1} + a_{i2}A_{j2} + \ldots + a_{in}A_{jn} = 0, \qquad 1 \leq i \neq j \leq n;$$

(4)

$$a_{1k}A_{1h} + a_{2k}A_{2h} + \ldots + a_{nk}A_{nh} = 0, \qquad 1 \leq h \neq k \leq n.$$

Il Teorema di Laplace riduce il calcolo di un determinante di ordine n al calcolo di un certo numero di determinanti di ordine $n-1$. In questo modo, applicando ripetutamente il teorema di Laplace, si ottiene un altro metodo effettivo per calcolare il determinante di una matrice. Questo metodo risulta particolarmente conveniente quando la matrice ha molti coefficienti uguali a 0.

Esempio. Calcoliamo il determinante della matrice $M = \begin{pmatrix} 1 & 2 & 0 & 0 \\ 2 & 1 & 4 & 2 \\ 3 & 0 & 0 & 1 \\ 4 & -1 & 0 & 2 \end{pmatrix}$ col Teorema di Laplace, sviluppandolo prima lungo la terza colonna e poi lungo la prima riga della matrice 3×3 rimanente.

$$\det M = 4 \cdot (-1)^{2+3} \det \begin{pmatrix} 1 & 2 & 0 \\ 3 & 0 & 1 \\ 4 & -1 & 2 \end{pmatrix} = -4(1(-1)^{1+1} \det \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix} + 2(-1)^{1+2} \det \begin{pmatrix} 3 & 1 \\ 4 & 2 \end{pmatrix}) =$$

$$= -4(1-4) = -12.$$

• Un'altra applicazione delle identità del teorema di Laplace è la *formula dell'inversa di una matrice* M, direttamente in termini dei suoi coefficienti:

$$M^{-1} = \frac{1}{\det(M)} \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \vdots & \ddots & \vdots & \\ A_{n1} & A_{n2} & \dots & A_{nn} \end{pmatrix}^t = \frac{1}{\det(M)} \begin{pmatrix} A_{11} & A_{21} \dots & A_{n1} \\ A_{12} & A_{22} \dots & A_{n2} \\ \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} \dots & A_{nn} \end{pmatrix}.$$

Esempio. Sia $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ una matrice invertibile 2×2, con $\det M = ade - bc \neq 0$. La sua inversa è data da

$$M^{-1} = \frac{1}{\det M} \begin{pmatrix} A_{11} & A_{21} \\ A_{12} & A_{22} \end{pmatrix} = \begin{pmatrix} \frac{d}{ad-bc} & \frac{-b}{ad-bc} \\ \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{pmatrix}.$$

• *Identità del determinante.*

(i) $\det(M^t) = \det(M)$: segue direttamente dalla definizione;
(ii) $\det(MN) = \det(M)\det(N)$ (qui M ed N sono matrici quadrate);
(iii) $\det(M^{-1}) = \frac{1}{\det(M)}$ (se M è invertibile):
infatti, poiché vale $MM^{-1} = I_n$, applicando la proprietà (ii), troviamo

$$\det(MM^{-1}) = \det M \det M^{-1} = \det I_n = 1 \quad \Leftrightarrow \quad \det(M^{-1}) = \frac{1}{\det(M)}.$$

• *Identità dell'inversa.*

(i) $(M^{-1})^{-1} = M$.
(ii) $(M^t)^{-1} = (M^{-1})^t$;
(iii) $(MN)^{-1} = N^{-1}M^{-1}$ (qui M ed N sono matrici quadrate).

Dimostrazione.
(i) L'equazione $MM^{-1} = I_n$ ci dice al tempo stesso che M^{-1} è l'inversa di M e che M è l'inversa di M^{-1}.
(ii) Poiché $\det M^t = \det M$, si ha che la trasposta di M è invertibile se e solo se M è invertibile. Poiché l'inversa, quando esiste è unica, per dimostrare il punto (ii) è sufficiente verificare che $(M^{-1})^t$ funziona: ed infatti

$$(M^{-1})^t M^t = (MM^{-1})^t - = I_n^t = I_n.$$

(iii) Poiché $\det(MN) = \det(M)\det(N)$, si ha che il prodotto MN è invertibile se e solo se sono invertibili sia M che N. Poiché l'inversa, quando esiste è unica, per dimostrare il punto (iii) è sufficiente verificare che $N^{-1}M^{-1}$ funziona: ed infatti

$$N^{-1}M^{-1}MN = N^{-1}I_nN = N^{-1}N = I_n.$$

• *Regola di Sarrus.* La regola di Sarrus è un trucco per ricostruire la formula del determinante di una matrice 3×3 (*solo* 3×3). Sia $M = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$. Scriviamo la tabella

$$\begin{pmatrix} a & b & c & a & b \\ d & e & f & d & e \\ g & h & i & g & h \end{pmatrix}.$$

Allora il determinante di M è dato dai prodotti degli elementi sulle diagonali che vanno da sinistra in alto a destra in basso $aei + bfg + cdf$ meno i prodotti degli elementi sulle diagonali che vanno da destra in alto a sinistra in basso $bdi + afh + ceg$. In totale:

$$\det M = aei + bfg + cdf - bdi - afh - ceg.$$

• *Formule di Cramer.* Sia $MX = b$ un sistema di n equazioni in n incognite. Se M ha rango massimo, ossia se M è invertibile, allora il sistema è compatibile ed ha soluzione unica data da

$$X = M^{-1}b.$$

Esprimendo l'inversa di M come sopra, si riottengono le formule di Cramer.

9. Numeri complessi.

I numeri complessi $\mathbf{C}$ sono l'insieme delle combinazioni formali $z = x + iy$, dove x, y sono numeri reali e $i = \sqrt{-1}$ indica l'unità immaginaria che soddisfa $i^2 = -1$. I numeri x e y sono detti rispettivamente la parte reale e la parte immaginaria di $z = x + iy$ e sono indicati con

$$x = \mathrm{Re}z, \qquad y = \mathrm{Im}z.$$

Vale l'inclusione $\mathbf{R} \subset \mathbf{C}$, identificando i numeri reali $\mathbf{R}$ con i numeri complessi con parte immaginaria nulla. I numeri complessi si posono sommare e moltiplicare fra loro, mediante le formule

$$(a + ib) + (c + id) = (a + c) + i(b + d), \qquad (a + ib)(c + id) = (ac - bd) + i(ad + bc).$$

Si verifica facilmente che somma e prodotto così definiti hanno le seguenti proprietà:

s1) $\forall z_1, z_2 \in \mathbf{C} \quad z_1 + z_2 = z_2 + z_1$ (proprietà commutativa della somma)
s2) $\forall z_1, z_2, z_3 \in \mathbf{C} \quad z_1 + (z_2 + z_3) = (z_1 + z_2) + z_3$ (proprietà associativa della somma)
s3) $\exists \mathbf{0} = 0 + i0 \in \mathbf{C} : \quad z + \mathbf{0} = \mathbf{0} + z = z \;\; \forall z \in \mathbf{C}$ (elemento neutro per la somma)
s4) $\forall z = x + iy \in \mathbf{C} \quad z + (-z) = (-z) + z = \mathbf{0}$, dove $-z = -x - iy$ (opposto per la somma)
p1) $\forall z_1, z_2 \in \mathbf{C} \quad z_1 z_2 = z_2 z_1$ (proprietà commutativa del prodotto)
p2) $\forall z_1, z_2, z_3 \in \mathbf{C} \quad z_1(z_2 z_3) = (z_1 z_2)z_3$ (proprietà associativa del prodotto)
p3) $\exists \mathbf{1} = 1 + i0 \in \mathbf{C} : \quad z\mathbf{1} = \mathbf{1}z = z \;\; \forall z \in \mathbf{C}$ (elemento neutro per il prodotto)
p4) $\forall z = x + iy \neq \mathbf{0} \; \exists z^{-1} = \frac{x}{x^2+y^2} - i\frac{y}{x^2+y^2} : \quad z^{-1}z = zz^{-1} = \mathbf{1}$ (inverso per il prodotto)
d) $\forall z_1, z_2, z_3 \in \mathbf{C} \quad z_1(z_2 + z_3) = z_1 z_2 + z_1 z_3$ (proprietà distributiva di somma e prodotto)

Osservazione. Ricostruiamo la formula dell'inverso di un numero complesso $z = x+iy \neq \mathbf{0}$: per definizione l'inverso di z è un numero complesso $w = a + ib$ tale che

$$zw = (xa - yb) + i(xb + ya) = 1, \quad \text{che equivale a} \quad \begin{cases} xa - yb = 1 \\ xb + ya = 0. \end{cases} \tag{$*$}$$

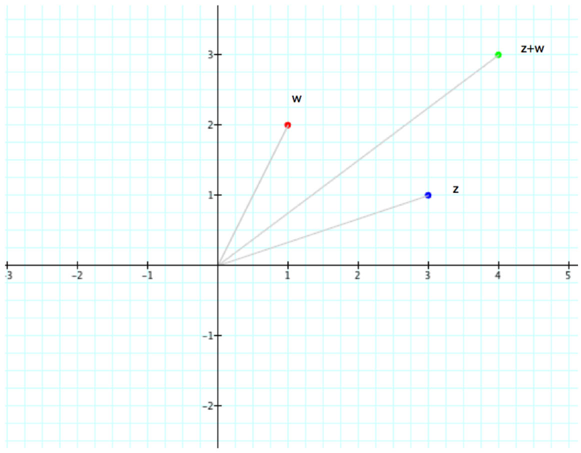
Si verifica facilmente che il sistema lineare $(*)$, nelle incognite $a = \mathrm{Re}w$ e $b = \mathrm{Im}w$, ammette soluzioni se e solo se x e y non sono entrambi nulli, quando cioè $z \neq \mathbf{0}$. In tal caso la soluzione è anche unica ed è data da $a = \frac{x}{x^2+y^2}$ e $b = \frac{-y}{x^2+y^2}$. Da ciò segue appunto la formula per l'inverso z^{-1}.

Si dice coniugato di $z = x + iy$ il numero complesso $\overline{z} = x - iy$.
L'operazione di *coniugio* ha le seguenti proprietà:

$\overline{\overline{z}} = z$
$\mathrm{Re}z = (z + \overline{z})/2, \quad \mathrm{Im}z = (z - \overline{z})/2i, \quad z^{-1} = \overline{z}/z\overline{z}$

$\overline{z+w} = \overline{z} + \overline{w}$
$\overline{zw} = \overline{z} \cdot \overline{w}$
$\overline{z^{-1}} = \overline{z}^{-1}$
$z\overline{z} = x^2 + y^2 \geq 0$

L'insieme $\mathbf{C}$ si può rappresentare nel piano cartesiano $\mathbf{R}^2$, identificando $z = x+iy$ con il vettore di coordinate $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \mathrm{Re}\, z \\ \mathrm{Im}\, \mathrm{z} \end{pmatrix}$. In questa rappresentazione, la somma fra due numeri complessi coincide con la somma fra i rispettivi vettori con la regola del parallelogramma, il coniugato $\overline{z}$ è il simmetrico di z rispetto all'asse della ascisse e il modulo di z, dato da $|z| = \sqrt{z\overline{z}} = \sqrt{(\mathrm{Re}\, z)^2 + (\mathrm{Im}\, \mathrm{z})^2}$, è la lunghezza del vettore z.



La somma $z + w = 4 + 3i$ fra i numeri complessi $z = 3 + i$ e $w = 1 + 2i$.

Per il modulo di un numero complesso valgono le seguenti proprietà:

$\mathrm{Re} z \leq |z|, \quad \mathrm{Im} z \leq |z|$
$|z| = |\overline{z}|$
$|z_1 z_2| = |z_1||z_2|$
$|z_1 + z_2| \leq |z_1| + |z_2|$ (disuguaglianza triangolare).

Se si indica con θ l'angolo tra l'asse x e il segmento congiungente z con l'origine, si ha $x = |z| \cos\theta$ e $y = |z| \sin\theta$. Si ottiene così la forma trigonometrica del numero complesso z

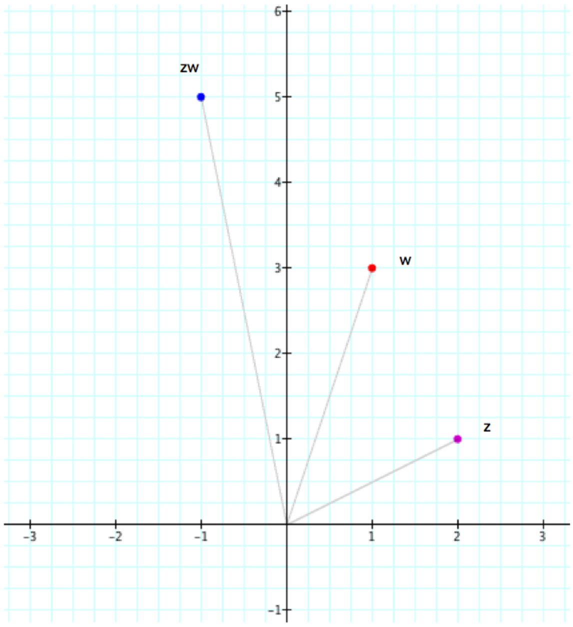
$$z = |z|(\cos\theta + i \sin\theta).$$

L'angolo θ si chiama l'argomento di z e si indica con $\theta = \arg(z)$. Se $z \neq \mathbf{0}$ esso è determinato solo a meno di multipli interi di 2π (per convenzione si può scegliere $\theta \in [0, 2\pi[$), mentre se $z = \mathbf{0}$ l'argomento è indeterminato. La forma trigonometrica è particolarmente conveniente per esprimere prodotti di numeri complessi. Se $z_1 = |z_1|(\cos\theta_1 + i \sin\theta_1)$ e $z_2 = |z_2|(\cos\theta_2 + i \sin\theta_2)$, allora

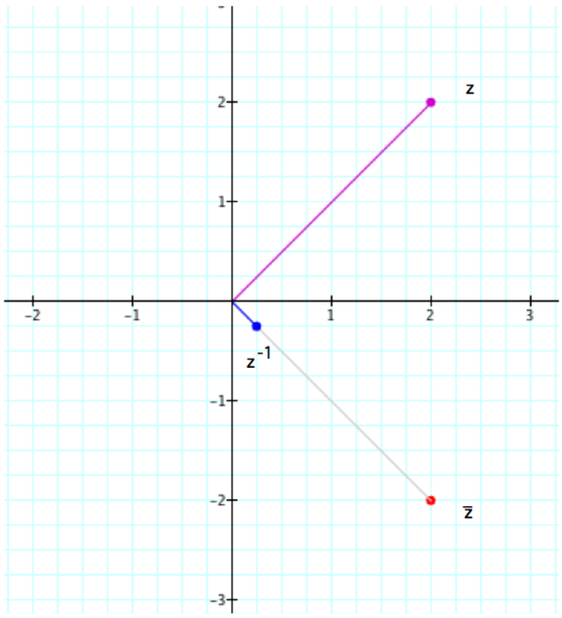
$$z_1 z_2 = |z_1 z_2|(\cos(\theta_1 + \theta_2) + i \sin(\theta_1 + \theta_2)).$$

In particolare,

$$z^n = |z|^n(\cos(n\theta) + i \sin(n\theta)).$$



Il prodotto $zw = -1 + 5i$ fra i numeri complessi $z = 2 + i$ e $w = 1 + 3i$.



Il coniugato $\overline{z}$ del numero complesso $z =$ e il suo inverso $z^{-1} =$.

L'importanza dei numeri complessi risiede nel fatto che sono un campo *algebricamente chiuso*: tutte le radici di un qualunque polinomio a coefficienti complessi appartengono a $\mathbf{C}$.

Teorema Fondamentale dell'Algebra. *Un'equazione polinomiale di grado n a coefficienti complessi*

$$a_0 + a_1 z + \ldots + a_n z^n = 0, \qquad a_i \in \mathbf{C},\ i = 1, \ldots n,$$

ha esattamente n radici complesse.

Esempio. *Le radici n-sime di un numero complesso.*

Per il Teorema Fondamentale dell'Algebra, per ogni $a_0 \in \mathbf{C}$ non nullo, l'equazione

$$z^n = a_0 \tag{**}$$

ha esattamente n radici in $\mathbf{C}$. Se scriviamo $z = |z|(\cos\theta + i\sin\theta)$ e $a_0 = |a_0|(\cos\phi_0 + i\sin\phi_0)$, l'equazione (**) diventa

$$|z|^n(\cos n\theta + i\sin n\theta) = |a_0|(\cos\phi_0 + i\sin\phi_0)$$

da cui si ricava $|z| = \sqrt[n]{|a_0|}$ ed $n\theta = \phi_0 + 2\pi k$, per $k \in \mathbf{Z}$. Al variare di $k \in \mathbf{Z}$ ci sono in realtà solo n angoli che danno luogo a numeri complessi distinti. Essi sono $\theta_i = \frac{\phi}{n} + \frac{2\pi k}{n}$, per $k = 0, 1, \ldots, n-1$. Le n radici dell'equazione (**) sono dunque

$$z_k = \sqrt[n]{|a_0|}(\cos(\frac{\phi}{n} + \frac{2\pi k}{n}) + i\sin(\frac{\phi}{n} + \frac{2\pi k}{n})), \quad k = 0, 1, \ldots, n-1.$$

Esempio. Le radici dell'equazione $z^4 = 4$ sono:
$z_1 = \sqrt{2}, \quad z_2 = \sqrt{2}(\cos\frac{\pi}{2} + i\sin\frac{\pi}{2}) = i\sqrt{2},$
$z_3 = \sqrt{2}(\cos\pi + i\sin\pi) = -\sqrt{2}, \quad z_4 = \sqrt{2}(\cos\frac{3\pi}{2} + i\sin\frac{3\pi}{2}) = -i\sqrt{2}.$

Esempio. Le radici dell'equazione $z^2 + 3iz + 4 = 0$ sono (usando la formula risolutiva delle equazioni di secondo grado)

$$z_{1,2} = \frac{-3i + \sqrt{-25}}{2} = \frac{-3i \pm 5i}{2} = -4i, \; i.$$

Esempio. Le radici dell'equazione $z^2 + 2z + i = 0$ sono (usando la formula risolutiva delle equazioni di secondo grado) $z_1 = -1 + w_1$ e $z_2 = -1 + w_2$, dove

$$w_1 = \sqrt[4]{2}(\cos\frac{-\pi}{8} + i\sin\frac{-\pi}{8}), \quad w_2 = \sqrt[4]{2}(\cos\frac{7\pi}{8} + i\sin\frac{7\pi}{8})$$

sono le due radici dell'equazione $w^2 = 1 - i$.

Dal Teorema Fondamentale dell'Algebra segue che un polinomio a coefficienti complessi è completamente riducibile su $\mathbf{C}$:

Corollario 1. *Un polinomio di grado n a coefficienti complessi $p(z) = z^n + a_{n-1}z^{n-1} + \ldots + a_1 z + a_0$ si decompone in modo unico (a meno dell'ordine dei fattori) nel prodotto di n polinomi di grado 1 a coefficienti complessi*

$$p(z) = (z - \alpha_1)(z - \alpha_2)\ldots(z - \alpha_n),$$

dove $\alpha_1, \ldots, \alpha_n$ sono le radici di p.

Consideriamo adesso un polinomio di grado n a coefficienti reali

$$p(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0, \qquad a_i \in \mathbf{R}.$$

In questo caso particolare, vale il seguente fatto:

• *Se α è una radice di p, allora anche $\overline{\alpha}$ lo è.*

Infatti, se $\alpha^n + a_{n-1}\alpha^{n-1} + \ldots + a_1\alpha + a_0 = 0$, coniugando si trova

$$\overline{\alpha}^n + \overline{a}_{n-1}\overline{\alpha}^{n-1} + \ldots + \overline{a}_1\overline{\alpha} + \overline{a}_0 = \overline{\alpha}^n + a_{n-1}\overline{\alpha}^{n-1} + \ldots + a_1\overline{\alpha} + a_0 = 0.$$

Dunque anche $\overline{\alpha}$ è radice di p.

Da questo fatto segue per esempio che un polinomio a coefficienti reali di grado dispari ha almeno una radice reale.

Corollario 2. *Un polinomio a coefficienti reali di grado n si decompone in modo unico (a meno dell'ordine dei fattori) nel prodotto di polinomi di grado 1 e di grado 2 a coefficienti reali.*

Dimostrazione. Per il teorema fondamentale dell'algebra, p ha n radici in $\mathbf{C}$: un certo numero di esse saranno a due a due complesse coniugate $\beta_1, \overline{\beta}_1, \ldots \beta_k, \overline{\beta}_k \in$, altre saranno possibilmente reali $\lambda_1, \ldots, \lambda_h$. A partire dalla decomposizione del polinomio su $\mathbf{C}$, troviamo

$$(z - \beta_1)(z - \overline{\beta}_1) \ldots (z - \beta_k)(z - \overline{\beta}_k)(z - \lambda_1) \ldots (z - \lambda_h) =$$

$$= (z^2 - 2\mathrm{Re}\beta_1 z + |\beta_1|^2) \ldots (z^2 - 2\mathrm{Re}\beta_k z + |\beta_k|^2)(z - \lambda_1) \ldots (z - \lambda_h),$$

che è proprio la decomposizione cercata.

10. Autovalori e autospazi.

Sia V uno spazio vettoriale *reale* di dimensione n e sia $L: V \longrightarrow V$ un'applicazione lineare.
Consideriamo il seguente problema:

esiste una base di V, *la stessa in dominio e codominio*, rispetto alla quale la matrice rappresentativa di L è più semplice possibile? Ad esempio:
Esiste una base $\mathcal{B}$ di V, la stessa in dominio e codominio, rispetto alla quale la matrice rappresentativa di L è diagonale?

In caso affermativo, l'applicazione si dice *diagonalizzabile*. Una caratterizzazione delle applicazioni lineari diagonalizzabili è data dalla seguente proposizione:

Proposizione. *Sia $L: V \to V$ un'applicazione lineare di uno spazio vettoriale reale V di dimensione n in sè. La matrice rappresentativa di L rispetto ad una base $\mathcal{B} = \{v_1, \ldots, v_n\}$ di V è diagonale se e solo se esistono numeri reali $\lambda_1, \ldots, \lambda_n$ tali che*

$$L(v_1) = \lambda_1 v_1, \ldots, L(v_n) = \lambda_n v_n. \tag{2}$$

Dimostrazione. La matrice rappresentativa M di L rispetto ad una base $\mathcal{B} = \{v_1, \ldots, v_n\}$ di V (la stessa in dominio e codominio) è quella matrice che moltiplicata per le coordinate di un vettore v in $\mathcal{B}$ restituisce le coordinate di $L(v)$ in $\mathcal{B}$. La matrice M ha per colonne le coordinate di $\{L(v_1), \ldots, L(v_n)\}$ in $\mathcal{B}$. Le coordinate di $\{v_1, \ldots, v_n\}$ in $\mathcal{B}$ sono date rispettivamente da

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Adesso è evidente che la matrice rappresentativa di L è una matrice diagonale reale

$$M = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}$$

se e solo se le coordinate di $\{L(v_1), \ldots, L(v_n)\}$ in $\mathcal{B}$ sono date da

$$\begin{pmatrix} \lambda_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \lambda_2 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \lambda_n \end{pmatrix},$$

ossia se e solo se

$$L(v_1) = \lambda_1 v_1, \ldots, L(v_n) = \lambda_n v_n, \tag{$*$}$$

con $\lambda_1, \ldots, \lambda_n$ numeri reali.

Definizione. Sia λ un numero reale. Si dice che λ è un *autovalore* (reale) di L se esiste un vettore non nullo $v \in V$ tale che

$$L(v) = \lambda v.$$

In questo caso v è per definizione un *autovettore di L di autovalore* λ. L'insieme di tutti gli autovettori di L di autovalore λ

$$V_\lambda = \{v \in V \mid L(v) = \lambda v\}$$

si chiama l'*autospazio* di L di autovalore λ.

Esercizio. L'autospazio V_λ è un sottospazio vettoriale di V.

Sol. Siano $v_1, v_2 \in V_\lambda$. Facciamo vedere che per ogni $a, b \in \mathbf{R}$, il vettore $av_1 + bv_2$ appartiene ancora a V_λ. Infatti

$$L(av_1 + bv_2) = aL(v_1) + bL(v_2) = a\lambda v_1 + b\lambda v_2 = \lambda(av_1 + bv_2).$$

Osservazione. La proposizione precedente può essere così riformulata: *L'applicazione L è diagonalizzabile se e solo se esiste una base di V formata da autovettori di L, relativi ad autovalori reali. In una qualunque base $\mathcal{B}$ che soddisfa (*), la matrice rappresentativa di L è una matrice diagonale*

$$M = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

che ha sulla diagonale principale gli autovalori (reali) di L.

Esempio. Sia $L_A \colon \mathbf{R}^2 \longrightarrow \mathbf{R}^2$ l'applicazione data da $X \mapsto A \cdot X$, con $A = \begin{pmatrix} 2 & 3 \\ 0 & -1 \end{pmatrix}$. Si ha che

$$L_A(\begin{pmatrix} 1 \\ 0 \end{pmatrix}) = \begin{pmatrix} 2 \\ 0 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Dunque il vettore $X = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ è autovettore di L_A, di autovalore $\lambda = 2$: $\begin{pmatrix} 1 \\ 0 \end{pmatrix} \in V_2$. Poiché V_2 è un sottospazio vettoriale di V, tutti i vettori della forma $tX = \begin{pmatrix} t \\ 0 \end{pmatrix}$, $t \in \mathbf{R}$ sono autovettori di L_A, di autovalore $\lambda = 2$.

Esempio. Sia $L \colon V \longrightarrow V$ un'applicazione lineare. L'autospazio V_0 dell'autovalore $\lambda = 0$ è esattamente il nucleo di L

$$V_0 = \{v \in V \mid L(v) = O = 0 \cdot v\} = \ker L.$$

In particolare, L è invertibile se e solo se $\lambda = 0$ non è un autovalore di L.

Esempio. Sia $L \colon V \longrightarrow V$ l'applicazione *identità*

$$L(v) = v, \qquad \forall v \in V.$$

In questo caso, tutti i vettori di V sono autovettori di autovalore $\lambda = 1$, ossia l'autospazio V_1 coincide con l'intero spazio V. In qualunque base $\mathcal{B}$ di V, *la stessa in dominio e codominio*, la matrice rappresentativa dell'applicazione *identità* è data da

$$I_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

Esempio. Sia $L \colon V \longrightarrow V$ l'applicazione *identicamente nulla*

$$L(v) = O, \qquad \forall v \in V.$$

In questo caso, tutti i vettori di V sono autovettori di autovalore $\lambda = 0$, e l'autospazio V_0 coincide con l'intero spazio V. In qualunque base $\mathcal{B}$ di V, la matrice rappresentativa dell'applicazione *nulla* è data da

$$O_n = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}.$$

Esempio. Sia $V = \mathbf{R}^2$ e sia r una retta per l'origine in $\mathbf{R}^2$. Sia $S_r \colon \mathbf{R}^2 \longrightarrow \mathbf{R}^2$ l'applicazione data dalla riflessione ortogonale rispetto ad r. Tutti i punti della retta r sono mandati in se stessi, mentre tutti i punti sulla retta $r^\perp$, ortogonale ad r, sono mandati nel loro opposto:

$$S_r(v) = v, \ \forall v \in r, \qquad S_r(v) = -v, \ \forall v \in r^\perp.$$

In altre parole la retta r coincide con l'autospazio V_1, mentre la retta $r^\perp$ coincide con l'autospazio V_{-1}. Se $\mathcal{B} = \{v_1, v_2\}$ è una base di $\mathbf{R}^2$, con $v_1 \in r$ e $v_2 \in r^\perp$, allora la matrice rappresentativa di S_r rispetto a $\mathcal{B}$ (in dominio e codominio) è data da

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Non tutte le applicazioni lineari di uno spazio vettoriale reale in sè sono diagonalizzabili. Negli esempi che seguono, lo possiamo vedere da semplici considerazioni geometriche.

Esempio. Sia $R_\theta \colon \mathbf{R}^2 \longrightarrow \mathbf{R}^2$, $\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$. Si può verificare facilmente che R_θ è la rotazione di un angolo θ attorno all'origine (basta controllare l'effetto di R_θ sui vettori della base canonica $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$). Per $\theta \neq 0, \pi$, nessuna retta del piano è mandata in se stessa: in particolare non esistono $\lambda \in \mathbf{R}$ e $\begin{pmatrix} x \\ y \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix} \in \mathbf{R}^2$ tali che $R_\theta \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix}$ e l'applicazione R_θ non è diagonalizzabile.

Esempio. Sia $R_\theta \colon \mathbf{R}^3 \longrightarrow \mathbf{R}^3$, $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$. Questa volta R_θ è la rotazione di un angolo θ attorno all'asse z. Tutti i vettori del piano $z = 0$ vengono ruotati di un angolo θ. Tutti i vettori della forma $v = \begin{pmatrix} 0 \\ 0 \\ z \end{pmatrix}$ vengono lasciati fissi

$$R_\theta(v) = v,$$

ossia sono autovettori di R_θ di autovalore $\lambda = 1$. Per $\theta \neq 0, \pi$, l'asse z è l'unica retta dello spazio mandata in se stessa. Ne segue che non esiste una base di $\mathbf{R}^3$ formata da autovettori di R_θ: non c'è un numero sufficiente di autovettori linearmente indipendenti e l'applicazione non è diagonalizzabile.

• *Volendo determinare esplicitamente autovalori e autovettori di una applicazione lineare $L \colon V \to V$, procediamo come segue:*

(a) Assumiamo $\dim V = n$. Sia $\mathcal{C}$ una base arbitraria ma fissata di V, e sia $A = \{a_{ij}\}_{i,j=1,\ldots,n}$ la corrispondente matrice rappresentativa di L. Determiniamo tutti i numeri reali λ, per cui esistono soluzioni non nulle $X \in \mathbf{R}^n$, $X \neq 0$, dell'equazione vettoriale

$$AX = \lambda X.$$

Riscriviamo tale equazione come

$$AX - \lambda X = O \quad \Leftrightarrow \quad (A - \lambda I_n)X = O.$$

Condizione necessaria e sufficiente affinché il sistema lineare omogeneo $(A - \lambda I_n)X = O$ abbia soluzioni non nulle è che la matrice dei coefficienti $(A - \lambda I_n)$ sia non invertibile, ossia che

$$\det(A - \lambda I_n) = 0. \tag{3}$$

Il determinante (3) è un polinomio a coefficienti reali di grado n in λ, detto il *polinomio caratteristico* di A:

$$P_\lambda(A) = \det(A - \lambda I_n) = (-1)^n \lambda^n \pm \mathrm{traccia}(A)\lambda^{n-1} + \ldots + \det(A),$$

dove per definizione $\mathrm{traccia}(A) := (a_{11} + \ldots + a_{nn})$. Le radici del polinomio caratteristico $P_\lambda(A)$ sono per definizione gli autovalori di L. Trattandosi di un'applicazione lineare di uno spazio vettoriale *reale* in sè, siamo interessati alle *radici reali* di $P_\lambda(A)$, ossia agli autovalori reali di L.

(b) Per ogni autovalore reale λ_0, l'autospazio corrispondente è dato da

$$V_{\lambda_0} = \{X \in \mathbf{R}^n \mid (A - \lambda_0 I_n)X = O\},$$

ossia dalle soluzioni del sistema lineare omogeneo

$$(A - \lambda_0 I_n)X = O,$$

con matrice dei coefficienti $(A - \lambda_0 I_n)$. Per costruzione, questa matrice ha rango *minore* di n (infatti $\det(A - \lambda_0 I_n) = 0$) e il sistema corrispondente ammette soluzioni non nulle. Precisamente,

$$1 \le \dim(V_{\lambda_0}) = n - \mathrm{rango}(A - \lambda I_n).$$

La dimensione dell'autospazio V_{λ_0} si chiama anche la *molteplicità geometrica* dell'autovalore λ_0.

• Date due basi $\mathcal{B}_1$ e $\mathcal{B}_2$ di V, la relazione fra le rispettive matrici corrispondenti è illustrata dal seguente diagramma commutativo:

$$\begin{array}{ccc} V, \mathcal{B}_1 & \xrightarrow{M_1} & V, \mathcal{B}_1 \\ \uparrow C_{\mathcal{B}_2,\mathcal{B}_1} & & \downarrow C^{-1}_{\mathcal{B}_2,\mathcal{B}_1} \\ V, \mathcal{B}_2 & \xrightarrow{M_2} & V, \mathcal{B}_2. \end{array}$$

(moltiplicando le coordinate di v in $\mathcal{B}_2$ per la matrice rappresentativa M_2 si ottengono le coordinate di $L(v)$ in $\mathcal{B}_2$. Possiamo arrivare allo stesso risultato anche "passando da sopra": moltiplicando le coordinate di v in $\mathcal{B}_2$ per la matrice del cambiamento di base $C_{\mathcal{B}_2,\mathcal{B}_1}$ otteniamo le coordinate di v in $\mathcal{B}_1$; moltiplicando le coordinate di v in $\mathcal{B}_1$ per la matrice rappresentativa M_1 otteniamo le coordinate di $L(v)$ in $\mathcal{B}_1$; infine, moltiplicando le coordinate di $L(v)$ in $\mathcal{B}_1$ per la matrice del cambiamento di base $C_{\mathcal{B}_1,\mathcal{B}_2} = C^{-1}_{\mathcal{B}_2,\mathcal{B}_1}$, otteniamo come prima le coordinate di $L(v)$ in $\mathcal{B}_2$.)

Di conseguenza M_1 ed M_2 stanno nella seguente relazione:

$$M_2 = C^{-1}_{\mathcal{B}_2,\mathcal{B}_1} \cdot M_1 \cdot C_{\mathcal{B}_2,\mathcal{B}_1}, \qquad (1)$$

dove $C_{\mathcal{B}_2,\mathcal{B}_1}$ è la matrice del cambiamento di base dalla base $\mathcal{B}_2$ alla base $\mathcal{B}_1$.

Definizione. Due matrici $M_1,\ M_2$ si dicono *coniugate* se esiste un matrice invertibile C tale che

$$M_2 = C^{-1} M_1 C.$$

Esempio. Le matrici $M_1,\ M_2$ in (1) sono matrici coniugate.

Fatto importante. *Se $M_1,\ M_2$ sono matrici coniugate, allora hanno lo stesso polinomio caratteristico:*

$$M_2 = C^{-1} M_1 C, \quad C \text{ invertibile} \quad \Rightarrow \quad P_\lambda(M_2) = P_\lambda(M_1).$$

In particolare $P_\lambda(M_2)$ e $P_\lambda(M_1)$ hanno gli stessi coefficienti e le stesse radici. Di conseguenza, *il polinomio caratteristico, così come i suoi coefficienti e le sue radici, sono effettivamente associati all'applicazione L, e non dipendono dalla matrice rappresentativa scelta per calcolarli.*

Dim. Per definizione, $P_\lambda(M_2) = \det(M_2 - \lambda I_n) = \det(C^{-1}M_1C - \lambda I_n)$. Poiché $\lambda I_n = C^{-1}\lambda I_n C$, grazie alla distributività del prodotto fra matrici, possiamo scrivere

$$\det(C^{-1}M_1C - \lambda I_n) = \det((C^{-1}(M_1 - \lambda I_n)C) = \det C^{-1}\det(M_1 - \lambda I_n)\det C = \det(M_1 - \lambda I_n).$$

In conclusione, $P_\lambda(M_2) = P_\lambda(M_1)$, come richiesto.

Conseguenza. Due matrici coniugate hanno la stessa traccia e lo stesso determinante.
In generale vale

$$\text{traccia}(M) = (a_{11} + \ldots + a_{nn}) = \lambda_1 + \ldots + \lambda_n, \qquad \det(M) = \lambda_1 \cdot \ldots \cdot \lambda_n,$$

dove $\lambda_1, \ldots, \lambda_n$ sono le n radici del polinomio caratteristico, reali o complesse.
In particolare, se M è coniugata ad una matrice diagonale reale

$$\begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix},$$

allora

$$\text{traccia}(M) = (a_{11} + \ldots + a_{nn}) = \lambda_1 + \ldots + \lambda_n, \qquad \det(M) = \lambda_1 \cdot \ldots \cdot \lambda_n.$$

Osservazione. Per il Teorema Fondamentale dell'Algebra il polinomio caratteristico $P_\lambda(A)$ ha esattamente n radici in $\mathbf{C}$. Poiché i coefficienti di $P_\lambda(A)$ sono reali, le radici non reali sono sempre in numero pari, perché vengono a coppie $\alpha, \bar{\alpha}$. In particolare, se n è *dispari*, il polinomio $P_\lambda(A)$ ha almeno una radice reale.

Osservazione. Sia $m(\lambda_0)$ la *molteplicità algebrica* di λ_0, cioè la sua molteplicità come radice del polinomio $P_\lambda(A)$. Allora vale

$$1 \leq \dim(V_{\lambda_0}) \leq m(\lambda_0).$$

Dimostrazione. Supponiamo $\dim V_{\lambda_0} = k$. Siano $\{v_1, \ldots, v_k\}$ vettori linearmente indipendenti in V_{λ_0} e sia

$$\mathcal{B} = \{v_1, \ldots, v_k, w_{k+1}, \ldots, w_n\}$$

un loro completamento ad una base di V. La matrice rappresentativa di L rispetto alla base $\mathcal{B}$ (in dominio e codominio) è una matrice della forma

$$M = \begin{pmatrix} \lambda_0 I_k & * \\ O & * \end{pmatrix},$$

dove I_k è la matrice identità $k \times k$, O è la matrice nulla $(n-k) \times k$ e gli altri blocchi sono arbitrari. Il polinomio caratteristico di M è della forma

$$P_\lambda(M) = (\lambda_0 - \lambda)^k q(\lambda),$$

con $q(\lambda)$ polinomio in λ. Dunque, se la dimensione dell'autospazio V_{λ_0} è k, allora la molteplicità algebrica dell'autovalore λ_0 è almeno k:

$$\dim V_{\lambda_0} \leq m(\lambda_0).$$

Proposizione. *Siano $v_1, \dots, v_k$ autovettori relativi ad autovalori distinti*
$$v_1 \in V_{\lambda_1}, \dots, v_k \in V_{\lambda_k},$$
con $\lambda_1, \dots, \lambda_k \in \mathbf{R}$ autovalori distinti di L. Allora $v_1, \dots, v_k$ sono linearmente indipendenti.

Osservazione. Sia V uno spazio vettoriale reale di dimensione n e sia $L : V \longrightarrow V$ un'applicazione lineare. Se L ha k autovalori reali e distinti, per la proposizione precedente, L ha almeno k autovettori *linearmente indipendenti.* In particolare, un'applicazione lineare L con n autovalori reali e distinti è diagonalizzabile.
Attenzione: La condizione di avere n autovalori reali e distinti è *sufficiente, ma non necessaria* alla diagonalizzabilità di L. Consideriamo per esempio l'applicazione
$$L_M : \mathbf{R}^3 \to \mathbf{R}^3, \quad X \mapsto MX, \qquad M = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{pmatrix}.$$
Gli autovalori dell'applicazione sono $\lambda_1 = \lambda_2 = \lambda_3 = 1$, ma l'applicazione non è diagonalizzabile (provare per credere).

La caratterizzazione degli endomorfismi diagonalizzabili di uno spazio vettoriale reale può essera formulata anche così :

Proposizione. *Sia V uno spazio vettoriale reale. Un'applicazione lineare $L : V \longrightarrow V$ è diagonalizzabile se e solo se*
(i) Tutti gli autovalori di L sono reali;
(ii) Per ogni autovalore λ di L, la dimensione dell'autospazio $\dim V_\lambda$ *è massima, ossia uguale alla molteplicità algebrica di λ.*

Esercizio. Sia $S_\theta : \mathbf{R}^2 \longrightarrow \mathbf{R}^2$, $\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$. Si può verificare facilmente che S_θ è la riflessione rispetto ad una retta inclinata di un angolo θ rispetto all'asse delle x positive (basta controllare l'effetto di S_θ sui vettori della base canonica $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$). Calcolare autovalori e autospazi di S_θ e darne un'interpretazione geometrica.

Esercizio. Sia $L_A : \mathbf{R}^3 \longrightarrow \mathbf{R}^3$, $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$. Determinare 3 basi distinte di $\mathbf{R}^3$, formate da autovettori di L_A.

Esercizio. Sia $L : \mathbf{R}^3 \longrightarrow \mathbf{R}^3$, $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \wedge \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ (dove $\wedge$ il prodotto vettoriale). Calcolare autovalori e autospazi di S_θ e darne un'interpretazione geometrica.

Esercizio. Sia $L_M : \mathbf{R}^2 \longrightarrow \mathbf{R}^2$, $\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$. Calcolare autovalori e autospazi di L_M. Dire se è diagonalizzabile.

• Una matrice simmetrica reale è sempre diagonalizzabile.

Teorema. *Sia M una matrice reale simmetrica $n \times n$.*
(i) Tutte le radici del polinomio caratteristico $P_\lambda(M)$ sono reali.
(ii) Per ogni autovalore λ, si ha che la dimensione dell'autospazio V_λ è uguale alla molteplicità algebrica di λ.
(iii) Siano $\lambda_1, \dots, \lambda_k \in \mathbf{R}$ autovalori distinti di M, e siano $v_1, \dots, v_k$ autovettori, con $v_i \in V_{\lambda_i}$. Allora $v_1, \dots, v_k$ sono a due a due ortogonali.
(iv) Esiste una base ortonormale di $\mathbf{R}^n$, fatta di autovettori di M.