

Spazi Vettoriali e Applicazioni lineari.

Appunti dalle lezioni di Geometria I - Corso di laurea in Matematica

Giorgio PATRIZIO - A.A. 2010/11

Queste note sono una traccia delle lezioni svolte e costituiscono solo un aiuto allo studio senza pretesa di completezza. Quando ci si riferisce a ABATE - de FABRITIIS si intende il testo: Abate - de Fabritiis, Geometria analitica con elementi di algebra lineare, Seconda Edizione, McGraw-Hill, 2010

1. Spazi vettoriali: definizioni e esempi.

Un campo \mathbb{K} è un insieme su cui sono definite due operazioni $+$ e \cdot che godono delle seguenti proprietà:

- (1) $a + (b + c) = (a + b) + c$ per ogni $a, b, c \in \mathbb{K}$ “associatività di $+$ ”
- (2) esiste $0 \in \mathbb{K}$ tale che $a + 0 = 0 + a$ per ogni $a \in \mathbb{K}$ “esistenza dello zero 0”
- (3) per ogni $a \in \mathbb{K}$ esiste $-a \in \mathbb{K}$ tale che $a + (-a) = -a + a = 0$ “esistenza dell’opposto”
- (4) $a + b = b + a$ per ogni $a, b \in \mathbb{K}$ “commutatività di $+$ ”
- (5) $a \cdot (b + c) = a \cdot b + a \cdot c$ e $a \cdot (b + c) \cdot a = b \cdot a + c \cdot a$ per ogni $a, b, c \in \mathbb{K}$ “distributività”
- (6) $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ per ogni $a, b, c \in \mathbb{K}$ “associatività di \cdot ”
- (7) esiste $1 \in \mathbb{K}$ tale che $a \cdot 1 = 1 \cdot a$ per ogni $a \in \mathbb{K}$ “esistenza dell’uno 1”
- (8) $\forall a \in \mathbb{K}$ con $a \neq 0$ esiste $a^{-1} \in \mathbb{K}$ tale che $a \cdot a^{-1} = a^{-1} \cdot a = 1$ “esistenza inverso”
- (9) $a \cdot b = b \cdot a$ per ogni $a, b \in \mathbb{K}$ “commutatività di \cdot ”

Nomenclatura: Un insieme G dotato di una operazione che soddisfa le proprietà (1),(2) e (3) si dice *gruppo*; se vale anche la (4) si dice *gruppo commutativo o abeliano*. Dunque \mathbb{K} con l’operazione $+$ è un gruppo abeliano e si vede facilmente che $\mathbb{K}^* = \mathbb{K} \setminus \{0\}$ considerato con l’operazione \cdot è un gruppo abeliano.

Esempi: L’insieme dei numeri interi \mathbb{Z} con la usuale somma è un gruppo abeliano. L’insieme \mathbb{R} dei numeri reali, l’insieme \mathbb{C} dei numeri complessi e l’insieme \mathbb{Q} dei numeri razionali, con le usuali operazioni di somma e prodotto, sono esempi di campi.

Anche se molte delle considerazioni che faremo qui di seguito valgono in maggiore generalità, in quello che segue \mathbb{K} sarà sempre il campo dei numeri reali \mathbb{R} o il campo dei numeri complessi \mathbb{C} .

Definizione. Uno spazio vettoriale su \mathbb{K} è un insieme V , i cui elementi chiameremo *vettori*, su cui sono definite due operazioni

$$\begin{array}{l} \text{somma} \quad +: V \times V \rightarrow V \\ \quad \quad \quad (v, w) \mapsto v + w \end{array} \quad e \quad \begin{array}{l} \text{prodotto per uno scalare} \quad p: \mathbb{K} \times V \rightarrow V \\ \quad \quad \quad (\lambda, w) \mapsto \lambda w \end{array}$$

in modo che V con la somma $+$ sia un *gruppo abeliano* ossia valgano le seguenti

- (1) $u + (v + w) = (u + v) + w$ per ogni $u, v, w \in V$ “*proprietà associativa della somma*”
- (2) esiste $O \in V$ tale che $u + O = O + u$ per ogni $u \in V$ “*esistenza del vettore nullo O* ”
- (3) per ogni $u \in V$ esiste $-u \in V$ tale che $u + (-u) = -u + u = O$ “*esistenza dell’opposto*”
- (4) $u + v = v + u$ per ogni $u, v \in V$ “*proprietà commutativa della somma*”

e che inoltre valgano le seguenti proprietà del prodotto per uno scalare:

- (5) $\lambda(u + v) = \lambda u + \lambda v$ per ogni $\lambda \in \mathbb{K}$ e $u, v \in V$
- (6) $(\lambda + \mu)u = \lambda u + \mu u$ per ogni $\lambda, \mu \in \mathbb{K}$ e $u \in V$
- (7) $(\lambda\mu)u = \lambda(\mu u)$ per ogni $\lambda, \mu \in \mathbb{K}$ e $u \in V$
- (8) $1u = u$ per ogni $u \in V$

Immediatamente dalla definizione abbiamo le seguenti conseguenze formali:

Proposizione 1.1: *Sia V uno spazio vettoriale su \mathbb{K} . Allora:*

- (i) *il vettore nullo O è unico;*
- (ii) *l’opposto di ogni $u \in V$ è unico;*
- (iii) *per ogni $u \in V$ si ha $0u = O$;*
- (iv) *per ogni $\lambda \in \mathbb{K}$ si ha $\lambda O = O$;*
- (v) *per ogni $u \in V$ si ha $-u = (-1)u$;*
- (vi) *per $\lambda \in \mathbb{K}$ e $u \in V$ si ha $\lambda u = O \iff \lambda = 0$ oppure $u = O$.*

Dimostrazione: (i) Il vettore nullo è unico dato che se O e O' verificano (2) allora

$$O = O + O' = O'.$$

(ii) Sia $u \in V$ e siano $-u, -u'$ due vettori che soddisfano (3). Allora

$$-u = -u + O = -u + (u + (-u')) = (-u + u) + (-u') = O + (-u') = -u'.$$

(iii) Per $u \in V$ si ha $0u + u = 0u + 1u = (0 + 1)u = 1u = u$; dunque

$$0u = 0u + u + (-u) = u + (-u) = O.$$

(iv) Sia $\lambda \in \mathbb{K}$. Per $u \in V$ arbitrario abbiamo $\lambda O = \lambda(0u) = (\lambda 0)u = 0u = O$.

(v) Per ogni $u \in V$ si ha $(-1)u + u = (-1)u + 1u = (-1 + 1)u = O$ e quindi la tesi segue per l’unicità dell’opposto.

(vi) Basta dimostrare che vale l’implicazione \implies . Supponiamo che per qualche $\lambda \in \mathbb{K}$ e $u \in V$ si abbia $\lambda u = O$. Se $\lambda \neq 0$ esiste λ^{-1} e quindi $O = \lambda^{-1}O = \lambda^{-1}\lambda u = 1u = u$. \square

Esempio 1. Lo spazi $\mathcal{V}_O^1, \mathcal{V}_O^2, \mathcal{V}_O^3$ dei vettori applicati in un punto della retta, del piano e dello spazio, con le operazioni di somma e di prodotto per uno scalare usuali, sono spazi vettoriali sul campo dei reali \mathbb{R} .

Esempio 2. Lo spazio delle ennuple di scalari \mathbb{K}^n con la somma e il prodotto per uno scalare definito da:

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix} \quad \forall x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{K}^n$$

e

$$\lambda \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix} \quad \forall \lambda \in \mathbb{K}, x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n.$$

è uno spazio vettoriale sul campo \mathbb{K} . Più in generale lo spazio $M_{m,n}(\mathbb{K})$ delle matrici m righe e n colonne con entrate nel campo \mathbb{K} è uno spazio vettoriale sul campo \mathbb{K} se equipaggiato con somma e prodotto per uno scalare definito da:

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \dots & b_{mn} \end{pmatrix} = \begin{pmatrix} a_{11} + b_{11} & \dots & a_{1n} + b_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & \dots & a_{mn} + b_{mn} \end{pmatrix}$$

e

$$\lambda \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} = \begin{pmatrix} \lambda a_{11} & \dots & \lambda a_{1n} \\ \vdots & \ddots & \vdots \\ \lambda a_{m1} & \dots & \lambda a_{mn} \end{pmatrix}$$

per ogni $A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$, $B = \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \dots & b_{mn} \end{pmatrix} \in M_{m,n}(\mathbb{K})$ e $\lambda \in \mathbb{K}$.

Esempio 3. Sia A un insieme arbitrario e sia W l'insieme di tutte le funzioni $f: A \rightarrow \mathbb{K}$ definite su A a valori nel campo \mathbb{K} . Se $f, g \in W$ e $\lambda \in \mathbb{K}$, allora le funzioni $f + g: A \rightarrow \mathbb{K}$ e $\lambda f: A \rightarrow \mathbb{K}$ sono le funzioni definite rispettivamente da:

$$(f + g)(a) = f(a) + g(a) \quad \text{e} \quad (\lambda f)(a) = \lambda f(a) \quad \forall a \in A.$$

Con le operazioni di somma di funzioni e di prodotto di una funzione per uno scalare definite in tal modo, W è uno spazio vettoriale su \mathbb{K} .

Più in generale sia V uno spazio vettoriale su \mathbb{K} . Sia Z l'insieme di tutte le applicazioni $f: A \rightarrow V$ definite su un insieme A a valori in V . Usando le operazioni di V si possono definire una somma e un prodotto per uno scalare su Z . Se $F, G \in Z$ e $\lambda \in \mathbb{K}$, allora le applicazioni $F + G: A \rightarrow V$ e $\lambda F: A \rightarrow V$ sono le applicazioni definite rispettivamente da:

$$(F + G)(a) = F(a) + G(a) \quad \text{e} \quad (\lambda F)(a) = \lambda F(a) \quad \forall a \in A.$$

Con le operazioni definite in tal modo, Z è uno spazio vettoriale su \mathbb{K} .

Esempio 4. Sia $\mathbb{K}[t]$ l'insieme dei polinomi nell'indeterminata t con coefficienti in \mathbb{K} . I polinomi sono funzioni da \mathbb{K} in \mathbb{K} e quindi, come nell'esempio 3, sono definite la somma di polinomi e il prodotto di un polinomio per uno scalare, che come è facile convincersi sono ancora polinomi. Siano $\lambda \in \mathbb{K}$ uno scalare e $p(t) = \sum_{k=0}^m a_k t^k, q(t) = \sum_{k=0}^n b_k t^k \in \mathbb{K}[t]$ due

polinomi dove si intende che $a_m \neq 0$ e $b_n \neq 0$. Supponiamo inoltre che $m \leq n$ (altrimenti basta scambiare i ruoli di $p(t)$ e $q(t)$) e si ponga $a_{m+1} = \dots = a_n = 0$; allora

$$(\lambda p)(t) = \lambda p(t) = \sum_{k=0}^m \lambda a_k t^k \quad \text{e} \quad (p+q)(t) = p(t) + q(t) = \sum_{k=0}^n (a_k + b_k) t^k.$$

Con queste operazioni $\mathbb{K}[t]$ è uno spazio vettoriale su \mathbb{K} .

Esempio 5. Sia \mathbb{R}^∞ lo spazio delle successioni reali. Dato che una successione $a_n = \{a_0, a_1, \dots, a_n, \dots\} \in \mathbb{R}^\infty$ non è altro che la funzione $a: \mathbb{N} \rightarrow \mathbb{R}$ definita da $a(n) = a_n$, allora \mathbb{R}^∞ è uno spazio vettoriale su \mathbb{R} con la struttura definita nell'esempio 3.

Esempio 6. Siano V, W due spazi vettoriali sullo stesso campo \mathbb{K} . Sul prodotto cartesiano $V \times W$ si può definire una struttura di spazio vettoriale definendo una somma e un prodotto per uno scalare utilizzando le operazioni definite su V e W . Infatti se $(v_1, w_1), (v_2, w_2) \in V \times W$ e $\lambda \in \mathbb{K}$, basta porre:

$$(v_1, w_1) + (v_2, w_2) = (v_1 + v_2, w_1 + w_2) \quad \text{e} \quad \lambda(v_1, w_1) = (\lambda v_1, \lambda w_1)$$

dove le operazioni su ciascuna componenti sono quelle definite su V e W . Con queste operazioni, $V \times W$ si dice spazio vettoriale prodotto di V e W . Si osservi che se $n = l + m$, definendo le struttura di spazio vettoriale di $\mathbb{K}^n, \mathbb{K}^l$ e \mathbb{K}^m come nell'esempio 2, si ha che \mathbb{K}^n è esattamente lo spazio vettoriale prodotto $\mathbb{K}^l \times \mathbb{K}^m$.

Definizione. Sia V uno spazio vettoriale. Un sottoinsieme $U \subset V$ si dice *sottospazio vettoriale* (o semplicemente *sottospazio*) di V se è *chiuso rispetto alla somma e al prodotto per uno scalare* ossia se

$$u + v \in U \quad \forall u, v \in U \quad \text{e} \quad \lambda u \in U \quad \forall \lambda \in \mathbb{K}, \forall u \in U.$$

Se $v_0 \in V$ e U è un sottospazio di V , l'insieme $S = v_0 + U = \{v = v_0 + u \in V \mid u \in U\}$ si dice *sottospazio affine* (o *sottovarietà affine*) di V . Il sottospazio U si dice *sottospazio di giacitura* di S .

Osservazione. Se U è un sottospazio vettoriale di uno spazio vettoriale V su \mathbb{K} , allora U , con le operazioni di somma e prodotto per uno scalare indotte da V , è uno spazio vettoriale su \mathbb{K} . Un sottospazio affine $S = v_0 + U$ è un sottospazio vettoriale e quindi è uno spazio vettoriale con le operazioni indotte da V se e solo se $O \in S$ e quindi se e solo se $v_0 \in U$.

Esempio 6. Si consideri \mathcal{V}_O^3 lo spazio dei vettori dello spazio applicati in O . Siano r, π rispettivamente una retta e un piano passanti per O . Allora i sottoinsiemi V_r e V_π di \mathcal{V}_O^3 dei vettori con secondo estremo rispettivamente in r e π sono sottospazi vettoriali di \mathcal{V}_O^3 . Se invece r, π sono una retta non passante per O , i sottoinsiemi S_r e S_π di \mathcal{V}_O^3 dei vettori con secondo estremo rispettivamente in r e π sono sottospazi affini di \mathcal{V}_O^3 .

Esempio 7. L'insieme $\mathbb{K}_n[t]$ dei polinomi nell'indeterminata t con coefficienti in \mathbb{K} di grado al più n è un sottospazio vettoriale di $\mathbb{K}[t]$.

Una classe di esempi particolarmente importante di sottospazi vettoriali e affini è quella degli insiemi di soluzioni dei sistemi lineari. A tal proposito abbiamo il seguente risultato:

Proposizione 1.2: (Teorema di struttura per i sistemi lineari) Siano $A \in M_{m,n}(\mathbb{K})$ una matrice con m righe e n colonne e $b \in \mathbb{K}^m$. L'insieme U delle soluzioni del sistema omogeneo di m equazioni e n incognite $Ax = 0$ è un sottospazio vettoriale di \mathbb{K}^n . Se v è una soluzione qualunque del sistema $Ax = b$, ogni altra soluzione ha la forma $w = v + u$ per qualche $u \in U$. In altre parole l'insieme delle soluzioni del sistema di m equazioni e n incognite $Ax = b$ è il sottospazio affine $L = v + U$.

Dimostrazione: Sia $A = (A^1, \dots, A^n)$ l'espressione in colonne della matrice A . Se $u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$

e $z = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} \in U$ sono soluzioni del sistema omogeneo $Ax = 0$, si ha $u + z \in U$ dato che

$$\begin{aligned} A(u + z) &= (u^1 + z^1)A^1 + \dots + (u^n + z^n)A^n = (u^1A^1 + \dots + u^nA^n) + (z^1A^1 + \dots + z^nA^n) \\ &= Au + Az = 0. \end{aligned}$$

Allo stesso modo si verifica banalmente che per ogni $\lambda \in \mathbb{K}$ si ha $\lambda u \in U$. Per concludere, si osservi che $w \in L$ se e solo se $Aw = b$ e quindi se e solo se

$$\begin{aligned} 0 &= Aw - Av = (w^1A^1 + \dots + w^nA^n) - (v^1A^1 + \dots + v^nA^n) \\ &= ((w^1 - v^1)A^1 + \dots + (w^n - v^n)A^n) = A(w - v) \end{aligned}$$

ossia se e solo se $u = w - v \in U$. □

2. Combinazioni lineari e lineare indipendenza.

Definizione. Sia V uno spazio vettoriale su \mathbb{K} , Dati $v_1, \dots, v_n \in V$ e $\alpha_1, \dots, \alpha_n \in \mathbb{K}$, il vettore

$$v = \alpha_1 v_1 + \dots + \alpha_n v_n$$

si dice *combinazione lineare* di v_1, \dots, v_n con *coefficienti* (o *pesi*) $\alpha_1, \dots, \alpha_n$.

Definizione. Sia V uno spazio vettoriale su \mathbb{K} . Se $v_1, \dots, v_n \in V$, il *sottospazio generato* da $v_1, \dots, v_n \in V$ è l'insieme di tutte le combinazioni lineari dei vettori v_1, \dots, v_n e si denota con

$$\text{Span}(v_1, \dots, v_n) = \{\alpha_1 v_1 + \dots + \alpha_n v_n \mid \alpha_1, \dots, \alpha_n \in \mathbb{K}\}.$$

Più in generale se $S \subset V$ (non necessariamente finito) il *sottospazio generato* da S è l'insieme di tutte le combinazioni lineari fdi un numero finito di elementi di S :

$$\text{Span}(S) = \{\alpha_1 v_1 + \dots + \alpha_k v_k \mid k \in \mathbb{N}, v_1, \dots, v_k \in S \text{ e } \alpha_1, \dots, \alpha_k \in \mathbb{K}\}.$$

Osservazione. $\text{Span}(S)$ è un sottospazio vettoriale di V .

Definizione. Se $V = \text{Span}(S)$, S si dice *insieme* (o *sistema*) di *generatori* di V . Uno spazio vettoriale si dice *finitamente generato* se $V = \text{Span}(S)$ per qualche sottoinsieme finito $S \subset V$.

Esempio 8. Gli spazi \mathbb{K}^n e $M_{m,n}(\mathbb{K})$ sono finitamente generati.

Esempio 9. Lo spazio dei polinomi $\mathbb{K}[t]$ con coefficienti in \mathbb{K} non è finitamente generato, mentre, per ogni n , lo spazio $\mathbb{K}[t]$ dei polinomi con coefficienti in \mathbb{K} di grado al più n è finitamente generato.

Esempio 10. Gli spazi $\mathcal{V}_O^1, \mathcal{V}_O^2, \mathcal{V}_O^3$ dei vettori applicati in O rispettivamente della retta del piano e dello spazio sono finitamente generati. Infatti per quanto riguarda \mathcal{V}_O^1 se $A \neq O$ è un qualunque punto della retta distinto dall'origine O e $\vec{i} = \overrightarrow{OA}$ allora per ogni vettore $\vec{v} \in \mathcal{V}_O^1$ esiste un unico scalare $t \in \mathbb{R}$ tale che $\vec{v} = t\vec{i}$. Per quanto riguarda \mathcal{V}_O^2 siano A, B sono punti del piano tali che O, A, B non sono allineati e si ponga $\vec{i} = \overrightarrow{OA}$ e $\vec{j} = \overrightarrow{OB}$. In questo caso si dirà anche che \vec{i}, \vec{j} sono *vettori non allineati*. Si ha che $\mathcal{V}_O^2 = \text{Span}(\vec{i}, \vec{j})$. In effetti si dimostra che dato un qualunque $\vec{v} \in \mathcal{V}_O^2$ esistono unici scalari $t_1, t_2 \in \mathbb{R}$ tali che $\vec{v} = t_1\vec{i} + t_2\vec{j}$. Infine per quanto riguarda \mathcal{V}_O^3 se A, B, C sono punti del piano tali che O, A, B, C non sono complanari allora si ponga $\vec{i} = \overrightarrow{OA}, \vec{j} = \overrightarrow{OB}$ e $\vec{k} = \overrightarrow{OC}$. In questo caso si dirà che $\vec{i}, \vec{j}, \vec{k}$ sono *vettori non complanari*. Si ha che $\mathcal{V}_O^3 = \text{Span}(\vec{i}, \vec{j}, \vec{k})$. In effetti si dimostra che dato un qualunque $\vec{v} \in \mathcal{V}_O^3$ esistono unici scalari $t_1, t_2, t_3 \in \mathbb{R}$ tali che $\vec{v} = t_1\vec{i} + t_2\vec{j} + t_3\vec{k}$. (Per le dimostrazioni si guardi ad esempio ABATE - de FABRITIIS pg 24 e seguenti)

Per i sistemi lineari le nozioni che abbiamo definito possono essere usate per interpretare la risolubilità di un sistema:

Proposizione 2.1: Siano $A \in M_{m,n}(\mathbb{K})$ una matrice con m righe e n colonne e $b \in \mathbb{K}^m$. Se $A = (A^1, \dots, A^n)$ è l'espressione di A secondo le sue colonne, il sistema lineare $Ax = b$ è compatibile (ossia ammette soluzione) se e solo se $b \in \text{Span}(A^1, \dots, A^n)$.

Dimostrazione: Il sistema ha una soluzione se e solo se esiste $x \in \mathbb{K}$ tale che $Ax = b$ ossia

se e solo se per qualche $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}$ si ha $b = x_1A^1 + \dots + x_nA^n$ e quindi se e solo se

$b \in \text{Span}(A^1, \dots, A^n)$. □

Definizione. Sia V uno spazio vettoriale su \mathbb{K} . I vettori $v_1, \dots, v_n \in V$ si dicono *linearmente dipendenti* se esistono $\alpha_1, \dots, \alpha_n \in \mathbb{K}$ **non tutti nulli** tali che

$$\alpha_1v_1 + \dots + \alpha_nv_n = O.$$

Invece $v_1, \dots, v_n \in V$ si dicono *linearmente indipendenti* se non sono linearmente dipendenti ossia se

$$\alpha_1v_1 + \dots + \alpha_nv_n = O \implies \alpha_1 = \dots = \alpha_n = 0.$$

Se v_1, \dots, v_n sono linearmente indipendenti, si dice che $\{v_1, \dots, v_n\}$ è un *sottoinsieme (o sistema) libero* di V .

Proposizione 2.2: I vettori $v_1, \dots, v_n \in V$ sono *linearmente dipendenti* se e solo se almeno uno fra di essi è combinazione lineare degli altri.

Dimostrazione: Se v_1, \dots, v_n sono linearmente dipendenti, allora esistono $\alpha_1, \dots, \alpha_n \in \mathbb{K}$ non tutti nulli tali che $\alpha_1v_1 + \dots + \alpha_nv_n = O$. A meno di riordinare i vettori, si può supporre che $\alpha_1 \neq 0$ e quindi $v_1 = -\frac{\alpha_2}{\alpha_1}v_2 - \dots - \frac{\alpha_n}{\alpha_1}v_n$. D'altra parte se un vettore, ad esempio v_1 , è combinazione lineare degli altri, allora $v_1 = \beta_2v_2 + \dots + \beta_nv_n$ e quindi $v_1 - \beta_2v_2 - \dots - \beta_nv_n = 0$ da cui segue che v_1, \dots, v_n sono linearmente dipendenti. □

Esempio 11. Un sistema libero in \mathcal{V}_O^1 è costituito da un solo vettore diverso dal vettore nullo. Ogni altro sottoinsieme di vettori di \mathcal{V}_O^1 è costituito da vettori linearmente dipendenti. In \mathcal{V}_O^2 un sottoinsieme costituito da due vettori non allineati è libero. Ogni sottoinsieme di vettori di \mathcal{V}_O^2 costituito da almeno tre vettori non è libero, infatti almeno uno dei vettori che lo compongono si può scrivere come combinazione lineare degli altri. In \mathcal{V}_O^3 un sottoinsieme costituito da due vettori non complanari è libero. Ogni sottoinsieme di vettori di \mathcal{V}_O^3 costituito da almeno quattro vettori non è libero, infatti almeno uno dei vettori che lo compongono si può scrivere come combinazione lineare degli altri.

Un'applicazione della nozione di lineare indipendenza ai sistemi lineari:

Proposizione 2.3: Sia $A \in M_{m,n}(\mathbb{K})$ una matrice con m righe e n colonne. e $b \in \mathbb{K}^m$. Se il sistema lineare $Ax = b$ ha soluzione, questa è unica se e solo se le colonne A^1, \dots, A^n di A sono linearmente indipendenti.

Dimostrazione: Dalla Proposizione 1.2 segue che basta dimostrare il risultato per $b = 0$. In questo caso il vettore nullo è sicuramente una soluzione. È l'unica soluzione se e solo se si ha $Ax = 0$, ossia $x_1 A^1 + \dots + x_n A^n = 0$, solo nel caso che $x_1 = \dots = x_n = 0$. Per definizione di lineare indipendenza abbiamo dunque la tesi. \square

3. Basi di uno spazio vettoriale.

Definizione. Sia V uno spazio vettoriale su \mathbb{K} finitamente generato. Un sottoinsieme $\mathcal{B} = \{v_1, \dots, v_n\}$ di V è una base di V se v_1, \dots, v_n sono linearmente indipendenti e \mathcal{B} è un insieme di generatori di V ossia $V = \text{Span}(v_1, \dots, v_n)$.

Proposizione 3.1: Siano V uno spazio vettoriale su \mathbb{K} e $v_1, \dots, v_n \in V$ vettori linearmente indipendenti. Se per $\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_n \in \mathbb{K}$ si ha

$$\alpha_1 v_1 + \dots + \alpha_n v_n = \beta_1 v_1 + \dots + \beta_n v_n$$

allora $\alpha_j = \beta_j$ per ogni $j = 1, \dots, n$. In particolare, allora, se $\mathcal{B} = \{v_1, \dots, v_n\}$ è una base di V , allora ogni vettore $v \in V$ si esprime in modo unico come combinazione lineare dei vettori di \mathcal{B}

Dimostrazione: Se $\alpha_1 v_1 + \dots + \alpha_n v_n = \beta_1 v_1 + \dots + \beta_n v_n$ allora $(\alpha_1 - \beta_1)v_1 + \dots + (\alpha_n - \beta_n)v_n = 0$, e quindi, dato che v_1, \dots, v_n sono vettori linearmente indipendenti, la tesi segue. \square

Definizione. Sia $\mathcal{B} = \{v_1, \dots, v_n\}$ una base dello spazio vettoriale V e $v \in V$. Gli scalari $\alpha_1, \dots, \alpha_n \in \mathbb{K}$ tali che $v = \alpha_1 v_1 + \dots + \alpha_n v_n$ si dicono coordinate di v rispetto alla base \mathcal{B} .

Esempio 12. In \mathbb{K}^n siano e_1, \dots, e_n i vettori definiti da

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \dots, e_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

Allora $\{e_1, \dots, e_n\}$ è una base di \mathbb{K}^n spesso chiamata la *base canonica* di \mathbb{K}^n . Naturalmente la base canonica è solo una delle infinite basi di \mathbb{K}^n . La sua particolarità consiste nel fatto che le coordinate di un vettore di \mathbb{K}^n rispetto alla base canonica sono particolarmente facili da individuare dato che

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ 0 \end{pmatrix} = x_1 e_1 + x_2 e_2 + \dots + x_n e_n = \sum_{i=1}^n x_i e_i.$$

Esempio 13. Nello spazio $M_{m,n}(\mathbb{K})$ delle matrici m righe, n colonne, per $i = 1, \dots, m$ e $j = 1, \dots, n$, si considerino le matrici $E^{i,j} = ((e^{i,j})_{kl})$ definite da

$$((e^{i,j})_{kl}) = \begin{cases} 1 & \text{se } k = i \text{ e } j = l \\ 0 & \text{altrimenti} \end{cases}.$$

Allora l'insieme delle mn matrici $E^{i,j}$ costituisce una base per $M_{m,n}(\mathbb{K})$

Esempio 14. L'insieme $\{1 = t^0, t, t^2, \dots, t^n\}$ costituisce una base per lo spazio $\mathbb{K}_n[t]$ dei polinomi di grado al più n .

Esempio 15. Se $\vec{i} \neq \vec{0}$, allora $\{\vec{i}\}$ è una base di \mathcal{V}_O^1 . Se \vec{i}, \vec{j} sono vettori non allineati, allora $\{\vec{i}, \vec{j}\}$ è una base di \mathcal{V}_O^2 . Se $\vec{i}, \vec{j}, \vec{k}$ sono vettori non complanari, allora $\{\vec{i}, \vec{j}, \vec{k}\}$ è una base di \mathcal{V}_O^3 .

La seguente osservazione sarà spesso utile:

Lemma 3.2: Sia $\{v_1, \dots, v_n\}$ un sistema di generatori di uno spazio vettoriale V . Se per $\{w_1, \dots, w_k\}$ si ha $\{v_1, \dots, v_n\} \subset \text{Span}(w_1, \dots, w_k)$, allora $\{w_1, \dots, w_k\}$ è un sistema di generatori di uno spazio vettoriale V .

Dimostrazione: Se per ogni j si ha $v_j \in \text{Span}(w_1, \dots, w_k)$, allora per $i = 1, \dots, k$ e $j = 1, \dots, n$, esistono scalari t_i^j tali che, per ogni $j = 1, \dots, n$, si abbia

$$v_j = t_1^j w_1 + \dots + t_k^j w_k = \sum_{i=1}^k t_i^j w_i.$$

Se $v \in V = \text{Span}(v_1, \dots, v_n)$, si ha

$$v = \sum_{j=1}^n a_j v_j = \sum_{j=1}^n a_j \left(\sum_{i=1}^k t_i^j w_i \right) = \sum_{i=1}^k \left(\sum_{j=1}^n a_j t_i^j \right) w_i$$

e quindi $v \in V = \text{Span}(w_1, \dots, w_k)$. □

Lemma 3.3: *Se uno spazio vettoriale $V \neq \{O\}$ ha una base composta da n elementi allora ogni sottoinsieme di V con più di n elementi non è libero.*

Dimostrazione: Sia $\mathcal{B} = \{v_1, \dots, v_n\}$ una base di V e, per assurdo sia $\{w_1, \dots, w_{n+1}\} \subset V$ un sistema libero. Dimostriamo, per induzione, che per ogni $k = 1, \dots, n$ risulta

$$V = \text{Span}(w_1, \dots, w_k, v_{k+1}, \dots, v_n). \quad (3.1)$$

Cominciamo esaminando il caso $k = 1$. Dato che \mathcal{B} è una base e $w_1 \neq O$ (perché?), per opportuni scalari t_1, \dots, t_n non tutti nulli, si deve avere $w_1 = t_1 v_1 + \dots + t_n v_n$. A meno di riordinare \mathcal{B} , possiamo supporre che $t_1 \neq 0$ e quindi si ha

$$v_1 = \frac{1}{t_1} w_1 - \frac{t_2}{t_1} v_2 - \dots - \frac{t_n}{t_1} v_n.$$

Ma allora $\{v_1, v_2, \dots, v_n\} \subset \text{Span}(w_1, v_2, \dots, v_n)$ e quindi, per il Lemma 3.2, segue che $V = \text{Span}(w_1, v_2, \dots, v_n)$.

Argomentando per induzione, supponiamo ora che l'affermazione (3.1) sia vera fino al passo $(k-1)$ -esimo – ossia $V = \text{Span}(w_1, \dots, w_{k-1}, v_k, \dots, v_n)$ – e dimostriamo $V = \text{Span}(w_1, \dots, w_k, v_{k+1}, \dots, v_n)$.

Dunque $w_k = s_1 w_1 + \dots + s_{k-1} w_{k-1} + s_k v_k + \dots + s_n v_n$, per opportuni scalari s_1, \dots, s_n . Si osservi che s_k, \dots, s_n non sono tutti nulli perché w_1, \dots, w_k sono linearmente indipendenti. Di nuovo a meno di riordinare v_k, \dots, v_n , possiamo supporre $s_k \neq 0$. Ma allora

$$v_k = -\frac{s_1}{s_k} w_1 - \dots - \frac{s_{k-1}}{s_k} w_{k-1} + \frac{1}{s_k} w_k - \frac{s_{k+1}}{s_k} v_{k+1} - \dots - \frac{s_n}{s_k} v_n$$

e quindi, argomentando come prima, $V = \text{Span}(w_1, \dots, w_k, v_{k+1}, \dots, v_n)$.

Dunque, per induzione, la (3.1) è vera e possiamo concludere che $V = \text{Span}(w_1, \dots, w_n)$. Ma allora $w_{n+1} \in \text{Span}(w_1, \dots, w_n)$, contro l'ipotesi che $\{w_1, \dots, w_{n+1}\}$ fosse libero. \square

Teorema 3.4: *Ogni spazio vettoriale finitamente generato $V \neq \{O\}$ ha una base. Due basi dello stesso spazio vettoriale hanno la stessa cardinalità (ossia lo stesso numero di elementi).*

Dimostrazione: Supponiamo che $V = \text{Span}(v_1, \dots, v_n)$. Sia k il massimo numero di vettori linearmente indipendenti in $\{v_1, \dots, v_n\}$. A meno di riordinare l'insieme, si può supporre dunque che v_1, \dots, v_k siano linearmente indipendenti e che per ogni $j = k+1, \dots, n$ i vettori v_1, \dots, v_k, v_j siano linearmente dipendenti e quindi $v_j \in \text{Span}(v_1, \dots, v_k)$. Dunque, per il Lemma 3.2, \mathcal{B} è un sistema di generatori per V ed è quindi una base perché è anche libero.

Sia ora \mathcal{C} un'altra base di V . Osserviamo, per cominciare, che \mathcal{C} ha al più k elementi. Infatti la base \mathcal{C} non può contenere più di k elementi perché se esistessero $w_1, \dots, w_{k+1} \in \mathcal{C}$, allora $\{w_1, \dots, w_{k+1}\}$ sarebbe un sistema libero: questo non è possibile per il Lemma 3.3. Scambiando i ruoli di \mathcal{B} e \mathcal{C} , la stessa dimostrazione prova che \mathcal{B} non può avere più elementi di \mathcal{C} e quindi la tesi segue. \square

Definizione. Sia V uno spazio vettoriale finitamente generato. Se $V \neq \{O\}$, la cardinalità, ossia il numero degli elementi, di una base di V si dice *dimensione* di V e si denota $\dim(V)$. Per lo spazio nullo $V = \{O\}$ si pone $\dim(\{O\}) = 0$.

Osservazione. La dimostrazione del Teorema 3.4 suggerisce una tecnica per trovare una base di uno spazio vettoriale qualora se ne conosca un sistema di generatori. L'idea è quella

di selezionare nel sistema di generatori un insieme massimale di vettori linearmente indipendenti con un procedimento induttivo. Se $\{O\} \neq V = \text{Span}(v_1, \dots, v_n)$ allora almeno uno fra i v_1, \dots, v_n è diverso dal vettore nullo. A meno di riordinare i vettori v_1, \dots, v_n , si può supporre $v_1 \neq O$. Si presentano ora due possibilità mutualmente esclusive. La prima è che, per ogni indice $j = 2, \dots, n$, l'insieme $\{v_1, v_j\}$ non è libero. In questo caso $V = \text{Span}(v_1)$ e quindi $\{v_1\}$ è una base di V . La seconda è che, per qualche $j = 2, \dots, n$, l'insieme $\{v_1, v_j\}$ sia un sistema libero. Sia j_0 il primo indice j per cui questo avviene. A meno di riordinare v_2, \dots, v_n , possiamo supporre che sia $j_0 = 2$. Dunque $\{v_1, v_2\}$ è un sistema libero. Di nuovo ci sono due possibilità mutualmente esclusive. La prima è che, per ogni indice $j = 3, \dots, n$, l'insieme $\{v_1, v_2, v_j\}$ non è libero. In questo caso $V = \text{Span}(v_1, v_2)$ e quindi $\{v_1, v_2\}$ è una base di V . La seconda è che, per qualche $j = 3, \dots, n$, l'insieme $\{v_1, v_2, v_j\}$ sia un sistema libero. Sia j_0 il primo indice j per cui questo avviene. A meno di riordinare v_3, \dots, v_n , possiamo supporre che sia $j_0 = 3$. Dunque $\{v_1, v_2, v_3\}$ è un sistema libero.....

Procedendo in questo modo, dopo un numero finito di passi (al più n) si determina un sistema libero di generatori di V .

Teorema 3.5: (del Completamento) *Sia $\mathcal{B} = \{v_1, \dots, v_n\}$ una base di uno spazio vettoriale V e, per $1 \leq p \leq n$, siano w_1, \dots, w_p siano vettori linearmente indipendenti di V . Esistono allora $n - p$ vettori di \mathcal{B} che insieme ai w_1, \dots, w_p formano una base di V .*

Dimostrazione: Si procede per induzione sull'intero p . Supponiamo che sia $p = 1$. Allora deve essere $w_1 \neq 0$ e quindi $w_1 = t_1 v_1 + \dots + t_n v_n$ con i coefficienti t_j non tutti nulli. A meno di riordinare \mathcal{B} si può supporre $t_1 \neq 0$. Allora

$$v_1 = \frac{1}{t_1} w_1 - \frac{t_2}{t_1} v_2 - \dots - \frac{t_n}{t_1} v_n.$$

Dunque $V = \text{span}(\mathcal{B}) = \text{Span}(w_1, v_2, \dots, v_n)$ per il Lemma 3.2. D'altra parte, se, per opportuni scalari s_1, \dots, s_n , risulta $s_1 w_1 + s_2 v_2 + \dots + s_n v_n = O$, allora

$$\begin{aligned} O &= s_1 w_1 + s_2 v_2 + \dots + s_n v_n = s_1 \left(\sum_{i=1}^n t_i v_i \right) + s_2 v_2 + \dots + s_n v_n \\ &= s_1 t_1 v_1 + (s_1 t_2 + s_2) v_2 \dots + (s_1 t_n + s_n) v_n \end{aligned}$$

Dato che \mathcal{B} è una base, deve essere

$$s_1 t_1 = s_1 t_2 + s_2 = \dots = s_1 t_n + s_n = 0.$$

Visto che $t_1 \neq 0$ allora segue $s_1 = 0$ e, di conseguenza, $s_2 = \dots = s_n = 0$. Segue che $\{w_1, \dots, v_n\}$ è una base e quindi se $p = 1$ il risultato è provato. Assumiamo ora che il risultato sia vero per $p = k - 1$ e proviamolo per $p = k$. Dunque supporremo che, a meno di riordinamenti di \mathcal{B} , $\{w_1, \dots, w_{k-1}, v_k, \dots, v_n\}$ sia una base di V . Allora $w_k = t_1 w_1 + \dots + t_{k-1} w_{k-1} + t_k v_k + \dots + t_n v_n$ con i coefficienti t_k, \dots, t_n non sono tutti nulli perché w_1, \dots, w_k sono linearmente indipendenti. A meno di riordinare t_k, \dots, v_n , si può supporre $t_k \neq 0$. Allora

$$v_k = -\frac{t_1}{t_k} w_1 - \dots - \frac{t_{k-1}}{t_k} w_{k-1} + \frac{1}{t_k} w_k - \frac{t_{k+1}}{t_k} v_{k+1} - \dots - \frac{t_n}{t_k} v_n$$

e quindi, ancora per il Lemma 3.2 si ha

$$V = \text{Span}(w_1, \dots, w_{k-1}, v_k, \dots, v_n) = \text{Span}(w_1, \dots, w_k, v_{k+1}, \dots, v_n).$$

Rimane da dimostrare che $w_1, \dots, w_k, v_{k+1}, \dots, v_n$ sono linearmente indipendenti. Se, per opportuni scalari s_1, \dots, s_n , risulta $s_1 w_1 + \dots + s_k w_k + s_{k+1} v_{k+1} + \dots + s_n v_n = O$, allora

$$\begin{aligned} O &= s_1 w_1 + \dots + s_k w_k + s_{k+1} v_{k+1} + \dots + s_n v_n \\ &= s_1 w_1 + \dots + s_k (t_1 w_1 + \dots + t_{k-1} v_{k-1} + t_k v_k + \dots + t_n v_n) s_{k+1} v_{k+1} + \dots + s_n v_n \\ &= (s_k t_1 + s_1) w_1 + \dots + (s_k t_{k-1} + s_{k-1}) w_{k-1} + s_k t_k v_k \\ &\quad + (s_k t_{k+1} + s_{k+1}) v_{k+1} \dots + (s_k t_n + s_n) v_n. \end{aligned}$$

Dato che $\{w_1, \dots, w_{k-1}, v_k, \dots, v_n\}$ è una base di V , allora

$$s_k t_1 + s_1 = \dots = s_k t_{k-1} + s_{k-1} = s_k t_k = s_k t_{k+1} + s_{k+1} = \dots = s_k t_n + s_n = 0.$$

Visto che $t_k \neq 0$ allora segue $s_k = 0$ e, di conseguenza, anche

$$s_1 = \dots = s_{k-1} = s_k = \dots = s_n = 0.$$

□

Corollario 3.6: Sia V uno spazio vettoriale di dimensione n e sia $\mathcal{B} = \{v_1, \dots, v_n\} \subset V$.

- (i) Se \mathcal{B} è un sistema libero in V , è una base di V ;
- (ii) se \mathcal{B} è un sistema di generatori di V , è una base di V .

Dimostrazione: (i) Se \mathcal{B} è un sistema libero in V allora, dato che in V ogni insieme di vettori con più di $n = \dim(V)$ vettori non è libero, per ogni $v \in V$ si deve avere $v \in \text{Span}(v_1, \dots, v_n)$, e quindi \mathcal{B} è una base di V .

(ii) Supponiamo che \mathcal{B} sia un sistema di generatori di V ma non sia una base di V . Allora \mathcal{B} conterrebbe un sottoinsieme proprio libero \mathcal{B}' e $\mathcal{B} \subset \text{Span}(\mathcal{B}')$. Ma allora per il Lemma 3.2, \mathcal{B}' sarebbe anche un sistema di generatori di V e quindi una base con un cardinalità strettamente minore della dimensione di V : assurdo! □

Corollario 3.7: Sia V uno spazio vettoriale di dimensione n e sia $W \subset V$ un sottospazio.

- (i) W è finitamente generato e $\dim(W) \leq \dim(V)$;
- (ii) $\dim(W) = \dim(V) \iff W = V$.

Dimostrazione: (i) Se W non fosse finitamente generato, in W esisterebbero $n + 1$ vettori ciascuno dei quali non è combinazione lineare degli altri. Dunque ci sarebbero $n + 1$ linearmente indipendenti in W e quindi in V . Questo è impossibile per il Lemma 3.3 dato che $\dim(V) = n$. In W non si possono dunque trovare più di n vettori linearmente indipendenti e quindi, necessariamente $\dim(W) \leq \dim(V)$.

(ii) Se $W = V$ ovviamente $\dim(W) = \dim(V)$. Viceversa, se $\dim(W) = \dim(V)$, allora una base $\mathcal{B} = \{w_1, \dots, w_n\}$ di W , è un sistema libero in V formato da $n = \dim(V)$ vettori ed è quindi una base di V per il Corollario 3.6. Ma allora $V = \text{Span}(\mathcal{B}) = W$. □

Un'osservazione finale per evitare confusione. In alcuni testi – fra questi quello di M. Abate e C. de Fabritiis, *Geometria analitica con elementi di algebra lineare* – dato che ci si occupa esclusivamente di spazi finitamente generati, si richiede che una base di uno spazio vettoriale debba essere necessariamente un insieme finito. Sotto questa condizione ovviamente gli spazi non finitamente generati non possono avere una base. Per spazi vettoriali non finitamente generati le cose sono più complicate. Una generalizzazione a spazi vettoriali arbitrari della nozione di base è data dalla seguente:

Definizione. Sia V uno spazio vettoriale su \mathbb{K} . Un sottoinsieme \mathcal{B} di V è una *base* di V se è un insieme di generatori di V tale che ogni suo sottoinsieme finito è libero.

Si osservi che, nel caso di spazi finitamente generati, questa definizione di base è completamente equivalente a quella data sopra. Come abbiamo osservato, ad esempio lo spazio dei polinomi $\mathbb{K}[t]$ con coefficienti in \mathbb{K} non è finitamente generato e ha base nel senso ora definito data da $\mathcal{B} = \{p_k(t) = t^k \mid k \in \mathbb{N}\}$. Si può dimostrare che ogni spazio vettoriale ammette una base in questo senso. Chi fosse curioso di vedere una dimostrazione dell'esistenza di basi per spazi non finitamente generati e qualcosa sulla nozione di dimensione in questo caso, può consultare ad esempio il testo M. Abate, *Geometria* alle pagine 93 – 103. Una parola di avvertenza. Nel caso non finitamente generato la prova dell'esistenza di una base non è costruttiva e non si hanno indicazioni su come trovarle. Inoltre in genere non si tratta di una nozione utile come nel caso finitamente generato. Infatti, persino per esempi semplici, ci si deve aspettare che una base sia un insieme troppo grande per essere maneggevole. Ad esempio si può dimostrare che una base per lo spazio \mathbb{R}^∞ delle successioni reali deve essere necessariamente un insieme non numerabile. Per evitare questi problemi, per spazi non finitamente generati si introducono nozioni più appropriate e maneggevoli.

Appendice: Riferimenti affini per piano e spazio, coordinate, equazioni per rette e piani.

Diamo per scontato che il lettore abbia familiarità con le nozioni di base sullo spazio dei vettori applicati del piano e dello spazio. Per queste si può consultare il capitolo 2 del testo di M. Abate e C. de Fabritiis, *Geometria analitica con elementi di algebra lineare*. Prima di tutto occorre introdurre la nozione di riferimento affine. Un *riferimento affine* del piano è una terna $\mathcal{R} = RA(O, \vec{i}, \vec{j})$ dove O è un punto del piano, e $\{\vec{i} = \overrightarrow{OA}, \vec{j} = \overrightarrow{OB}\}$ è una base dello spazio \mathcal{V}_O dei vettori applicati in O . Assegnare un riferimento affine equivale dunque ad assegnare tre punti O, A, B non allineati del piano dato che in questo caso i vettori $\vec{i} = \overrightarrow{OA}$ e $\vec{j} = \overrightarrow{OB}$ sono non allineati e quindi linearmente indipendenti.

Fissato un riferimento affine $\mathcal{R} = RA(O, \vec{i}, \vec{j})$ del piano, le coordinate $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ di un punto P del piano, per definizione, sono le coordinate del vettore \overrightarrow{OP} relative alla base $\{\vec{i}, \vec{j}\}$ ossia sono unicamente determinate dalla relazione:

$$\overrightarrow{OP} = x_1 \vec{i} + x_2 \vec{j}. \quad (3.2)$$

Per quanto riguarda lo spazio si procede nel modo seguente. Un *riferimento affine* dello spazio è una quaterna $\mathcal{R} = RA(O, \vec{i}, \vec{j}, \vec{k})$ dove O è un punto dello spazio, e $\{\vec{i} = \overrightarrow{OA}, \vec{j} = \overrightarrow{OB}, \vec{k} = \overrightarrow{OC}\}$ è una base dello spazio \mathcal{V}_O dei vettori applicati in O . Assegnare un riferimento affine nello spazio equivale dunque ad assegnare quattro punti O, A, B, C non complanari del piano dato che in questo caso i vettori $\vec{i} = \overrightarrow{OA}, \vec{j} = \overrightarrow{OB}$ e $\vec{k} = \overrightarrow{OC}$ sono non complanari e quindi linearmente indipendenti.

Fissato un riferimento affine $\mathcal{R} = RA(O, \vec{i}, \vec{j}, \vec{k})$ dello spazio, le coordinate $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ di un punto P dello spazio, per definizione, sono le coordinate del vettore \overrightarrow{OP} relative alla base $\{\vec{i}, \vec{j}, \vec{k}\}$ ossia sono unicamente determinate dalla relazione:

$$\overrightarrow{OP} = x_1 \vec{i} + x_2 \vec{j} + x_3 \vec{k}. \quad (3.3)$$

Diamo ora una rapidissima applicazione dell'idea di assegnare coordinate ai punti di piano e spazio, descrivendo come si possano descrivere rette e piani mediante equazioni parametriche. Rimandiamo per maggiori dettagli al capitolo 2 del testo di M. Abate e C. de Fabritiis citato sopra.

Cominciamo con le rette. Nel piano o nello spazio si fissi un punto O . Sia \mathbf{r}_0 una retta (del piano o dello spazio) passante per O e sia $Q_1 \neq O$ un altro punto di \mathbf{r}_0 . Allora il vettore $\overrightarrow{OQ_1}$ è una base per lo spazio dei vettori applicati in O della retta \mathbf{r}_0 . Dunque un qualunque altro punto Q giace sulla retta \mathbf{r}_0 se e solo se esiste, ed è unicamente determinato, $t \in \mathbb{R}$ si ha

$$\overrightarrow{OQ} = t \overrightarrow{OQ_1}. \quad (3.4)$$

Se \mathbf{r} è una retta arbitraria, siano P_0, P_1 due punti distinti di \mathbf{r} . Se \mathbf{r}_0 è la retta parallela a \mathbf{r} passante per O , esiste un unico punto $Q_1 \neq O$ su \mathbf{r}_0 tale che $\overrightarrow{OP_0} + \overrightarrow{OQ_1} = \overrightarrow{OP_1}$ ossia tale che $\overrightarrow{OQ_1} = \overrightarrow{OP_1} - \overrightarrow{OP_0}$. Dunque se P è un qualunque punto e Q è l'unico punto tale che $\overrightarrow{OP} = \overrightarrow{OP_0} + \overrightarrow{OQ}$, usando la (3.4), si può concludere che un punto P è giace sulla retta \mathbf{r} se e solo se esiste $t \in \mathbb{R}$ tale che

$$\overrightarrow{OP} = \overrightarrow{OP_0} + t(\overrightarrow{OP_1} - \overrightarrow{OP_0}). \quad (3.5)$$

La (3.5) si dice *equazione vettoriale* della retta \mathbf{r} . Si osservi che per ottenere la “descrizione” (3.5) della retta \mathbf{r} ci è bastato fissare il punto O e non è stato necessario precisare se eravamo nel piano o nello spazio. Supponiamo ora di essere nel piano e sia $\mathcal{R} = RA(O, \vec{\mathbf{i}}, \vec{\mathbf{j}})$ un riferimento del piano dove O è il punto che abbiamo fissato in precedenza. Se

$$x^0 = \begin{pmatrix} x_1^0 \\ x_2^0 \end{pmatrix}, \quad x^1 = \begin{pmatrix} x_1^1 \\ x_2^1 \end{pmatrix} \quad \text{e} \quad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

sono rispettivamente le coordinate di P_0, P_1 e di un punto arbitrario P di \mathbf{r} , allora da (3.5) si ha immediatamente che

$$\begin{cases} x_1 = x_1^0 + t(x_1^1 - x_1^0) \\ x_2 = x_2^0 + t(x_2^1 - x_2^0) \end{cases} \quad (3.6)$$

per $t \in \mathbb{R}$. Le (3.6) si dicono *equazioni parametriche* della retta \mathbf{r} .

Nello spazio, fissato un riferimento $\mathcal{R} = RA(O, \vec{\mathbf{i}}, \vec{\mathbf{j}}, \vec{\mathbf{k}})$, si procede allo stesso modo. Se

$$x^0 = \begin{pmatrix} x_1^0 \\ x_2^0 \\ x_3^0 \end{pmatrix}, \quad x^1 = \begin{pmatrix} x_1^1 \\ x_2^1 \\ x_3^1 \end{pmatrix} \quad \text{e} \quad x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

sono rispettivamente le coordinate di P_0, P_1 e di un punto arbitrario P di \mathbf{r} , allora da (3.5) si ha immediatamente che

$$\begin{cases} x_1 = x_1^0 + t(x_1^1 - x_1^0) \\ x_2 = x_2^0 + t(x_2^1 - x_2^0) \\ x_3 = x_3^0 + t(x_3^1 - x_3^0) \end{cases} \quad (3.7)$$

per $t \in \mathbb{R}$. Le (3.7) si dicono *equazioni parametriche* della retta \mathbf{r} .

Per quanto riguarda piani nello spazio lavoreremo in modo del tutto analogo. Se π_O è un piano per l'origine O , siano Q_1, Q_2 due punti arbitrari di π_O presi in modo che O, Q_1, Q_2 non siano allineati. Allora un qualunque altro punto Q giace in π_O se e solo se esistono, e sono unicamente determinati, $s, t \in \mathbb{R}$ tali che

$$\overrightarrow{OQ} = s\overrightarrow{OQ_1} + t\overrightarrow{OQ_2}. \quad (3.8)$$

Se π è un piano arbitrario, siano $P_0, P_1, P_2 \in \pi$ tre punti non allineati. Se π_O è il piano passante per O parallelo a π , siano $Q_1, Q_2 \in \pi_O$ i punti tali che

$$\overrightarrow{OP_0} + \overrightarrow{OQ_1} = \overrightarrow{OP_1} \quad \text{e} \quad \overrightarrow{OP_0} + \overrightarrow{OQ_2} = \overrightarrow{OP_2}$$

ossia tali che

$$\overrightarrow{OQ_1} = \overrightarrow{OP_1} - \overrightarrow{OP_0} \quad \text{e} \quad \overrightarrow{OQ_2} = \overrightarrow{OP_2} - \overrightarrow{OP_0}.$$

Dunque $\overrightarrow{OQ_1}, \overrightarrow{OQ_2}$ sono non allineati e quindi se P è un qualunque punto Q è l'unico punto tale che $\overrightarrow{OP} = \overrightarrow{OP_0} + \overrightarrow{OQ}$, usando la (3.8), si può concludere che un punto P è giace sulla retta \mathbf{r} se e solo se esistono – e sono unicamente determinati – $s, t \in \mathbb{R}$ tali che

$$\overrightarrow{OP} = \overrightarrow{OP_0} + s(\overrightarrow{OP_1} - \overrightarrow{OP_0}) + t(\overrightarrow{OP_2} - \overrightarrow{OP_0}). \quad (3.9)$$

La (3.9) si dice *equazione vettoriale del piano* π .

Fissato un riferimento $\mathcal{R} = RA(O, \vec{\mathbf{i}}, \vec{\mathbf{j}}, \vec{\mathbf{k}})$, come per le rette, è ora semplice ricavare equazioni parametriche per il piano π . Se

$$x^0 = \begin{pmatrix} x_1^0 \\ x_2^0 \\ x_3^0 \end{pmatrix}, \quad x^1 = \begin{pmatrix} x_1^1 \\ x_2^1 \\ x_3^1 \end{pmatrix}, \quad x^2 = \begin{pmatrix} x_1^2 \\ x_2^2 \\ x_3^2 \end{pmatrix} \quad \text{e} \quad x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

sono rispettivamente le coordinate di P_0, P_1, P_2 e di un punto arbitrario P di π , allora da (3.9) si ha immediatamente che

$$\begin{cases} x_1 = x_1^0 + s(x_1^1 - x_1^0) + t(x_1^2 - x_1^0) \\ x_2 = x_2^0 + s(x_2^1 - x_2^0) + t(x_2^2 - x_2^0) \\ x_3 = x_3^0 + s(x_3^1 - x_3^0) + t(x_3^2 - x_3^0) \end{cases} \quad (3.10)$$

per $s, t \in \mathbb{R}$. Le (3.10) si dicono *equazioni parametriche* del piano π .

4. Operazioni fra spazi vettoriali.

Sia V uno spazio vettoriale e siano $U, W \subset V$ suoi sottospazi. Allora $U \cap W$ è un sottospazio di V (esercizio!) mentre in generale $U \cup W$ non è un sottospazio di V .

Esempio. Se $U = \text{Span} \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix} \right) \subset \mathbb{R}^2$ e $W = \text{Span} \left(\begin{pmatrix} 0 \\ 1 \end{pmatrix} \right) \subset \mathbb{R}^2$, allora $U \cup W$ non è un sottospazio di \mathbb{R}^2 .

In effetti l'unione di due sottospazi è un sottospazio se e solo uno dei due è contenuto nell'altro:

Proposizione 4.1: *Siano U, W sottospazi vettoriali di uno spazio vettoriale V . Allora $U \cup W$ è un sottospazio di $V \iff o U \subset W \text{ o } W \subset U$.*

Dimostrazione: Se uno dei due sottospazi è contenuto nell'altro, è evidente che l'unione è un sottospazio! Supponiamo che $U \cup W$ sia un sottospazio di V ma che $U \not\subset W$ e $W \not\subset U$. Siano allora $u \in U \setminus W$ e $w \in W \setminus U$. Ma allora $u, w \in U \cup W$ e quindi anche $u + w \in U \cup W$. Ma questo è assurdo perché se fosse $u + w \in U$ allora $w = -u + u + w \in U$ mentre se fosse $u + w \in W$ allora $u = -w + u + w \in W$. \square

La nozione "giusta" di sottospazio che contiene due sottospazi dati è quella di somma. Se $U, W \subset V$ sono sottospazi di uno spazio vettoriale, definiamo la loro *somma* mediante:

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

Si vede subito che $U + W$ è un sottospazio di V (esercizio!). La seguente osservazione è immediata:

Lemma 4.2: *Se per sottospazi vettoriali U, W di uno spazio vettoriale V si ha $U = \text{Span}(\mathcal{B})$ e $W = \text{Span}(\mathcal{C})$ allora $U + W = \text{Span}(\mathcal{B} \cup \mathcal{C})$. In particolare se U e W sono finitamente generati, $U + W$ è finitamente generato.*

Il seguente risultato chiarisce la relazione fra le dimensioni di due sottospazi, della loro intersezione e della loro somma.

Teorema 4.3: (Formula di Grassmann) *Siano U, W sottospazi vettoriali finitamente generati di uno spazio vettoriale V . Allora*

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

Dimostrazione: Per la dimostrazione, si veda ABATE - de FABRITIIS pg. 75. \square

Se $U, W \subset V$ sono sottospazi di uno spazio vettoriale tali che $U \cap W = \{O\}$ la loro somma si dice *somma diretta* di U e W e si denota mediante $U \oplus W$. Dunque, in questo caso, la formula di Grassmann diventa:

$$\dim(U \oplus W) = \dim(U) + \dim(W).$$

Nel caso in cui $U \oplus W = V$ si dice che U e W sono supplementari.

Osservazione. La somma di V due sottospazi U e W è una somma diretta se e solo se ogni vettore di V si scrive in modo unico come somma di un vettore di U e uno di W .

Proposizione 4.4: Siano V uno spazio vettoriale finitamente generato, e $U, W \subset V$ sottospazi vettoriali. Allora $U \oplus W = V$ se e solo se per basi $\{u_1, \dots, u_r\}$ di U e $\{w_1, \dots, w_s\}$ di W si ha che $\{u_1, \dots, u_r, w_1, \dots, w_s\}$ è una base di V .

Dimostrazione: Siano $\{u_1, \dots, u_r\}$ e $\{w_1, \dots, w_s\}$ basi rispettivamente di U e di W tali che $\{u_1, \dots, u_r, w_1, \dots, w_s\}$ sia una base di V . Allora si ha immediatamente che $U + W = V$. D'altro canto $U \cap W = \{O\}$ perché $\dim(U \cap W) = \dim(U) + \dim(W) - \dim(U + W) = 0$. Viceversa, se $U \oplus W = V$ e $\{u_1, \dots, u_r\}$ e $\{w_1, \dots, w_s\}$ sono basi di U e di W rispettivamente, allora $\{u_1, \dots, u_r, w_1, \dots, w_s\}$ è una base di V perché è un sistema di generatori con $r + s$ vettori e $\dim(V) = r + s$. \square

Osservazione. Siano V uno spazio vettoriale finitamente generato, e $U \subset V$ un sottospazio vettoriale. Dalla Proposizione 4.4 segue che per costruire un supplementare di U si può procedere come segue. Si consideri una base $\{u_1, \dots, u_r\}$ di U e la si completi a una base $\{u_1, \dots, u_r, w_1, \dots, w_s\}$ una base di V . Allora $W = \text{Span}(w_1, \dots, w_s)$ è un supplementare di U .

Infine osserviamo che la nozione di somma di sottospazi e di somma diretta si estende facilmente al caso di un numero finito di sottospazi. Siano U_1, \dots, U_k sottospazi di uno spazio vettoriale V . Allora si definisce

$$U_1 + \dots + U_k = \{v = u_1 + \dots + u_k \mid u_j \in U_j \forall j = 1, \dots, k\}.$$

Si dice che V è *somma diretta* di U_1, \dots, U_k e si scrive

$$V = U_1 \oplus \dots \oplus U_k$$

se $V = U_1 + \dots + U_k$ e se per ogni $j = 1, \dots, k$ si ha

$$U_j \cap (U_1 + \dots + U_{j-1} + U_{j+1} + \dots + U_k) = \{O\}.$$

Questo succede esattamente se e solo se ogni vettore di $v \in V$ si scrive in modo unico come $v = u_1 + \dots + u_k$ con $u_j \in U_j$ per ogni $j = 1, \dots, k$. In questo caso, la formula di Grassmann diventa:

$$\dim(V) = \dim(U_1 \oplus \dots \oplus U_k) = \sum_{j=1}^k \dim(U_j).$$

Infine si generalizza facilmente la Proposizione 4.4:

Proposizione 4.5: Siano V uno spazio vettoriale finitamente generato, e $U_1, \dots, U_k \subset V$ sottospazi vettoriali. Allora $V = U_1 \oplus \dots \oplus U_k$ se e solo se per ogni $j = 1, \dots, k$, fissate basi $\{u_1^j, \dots, u_{r_j}^j\}$ di U_j si ha che $\{u_1^1, \dots, u_{r_1}^1, \dots, u_1^k, \dots, u_{r_k}^k\}$ è una base di V .

5. Applicazioni lineari.

Definiamo ora le trasformazioni compatibili con la struttura di spazio vettoriale:

Definizione. Siano V, W due spazi vettoriali sullo stesso campo \mathbb{K} . Un'applicazione $L: V \rightarrow W$ si dice *applicazione lineare* se

$$L(u + v) = L(u) + L(v) \quad \forall u, v \in V \quad \text{e} \quad L(\lambda v) = \lambda L(v) \quad \forall \lambda \in \mathbb{K}, \forall v \in V.$$

Osservazione 1. Segue immediatamente dalla definizione che se $L: V \rightarrow W$ è un'applicazione lineare e O_V, O_W sono rispettivamente i vettori nulli di V e W , allora $L(O_V) = O_W$

Osservazione 2. Se V, W, Z sono spazi vettoriali su \mathbb{K} e $T: V \rightarrow W, S: W \rightarrow Z$ sono applicazioni lineari, la composizione $S \circ T: V \rightarrow Z$ definita da $S \circ T(v) = S(T(v))$ per ogni v in V , è un'applicazione lineare.

Osservazione 3. Siano V, W due spazi vettoriali sullo stesso campo \mathbb{K} e $\mathcal{L}(V, W)$ l'insieme delle applicazioni lineari da V a W . E' immediato verificare che $\mathcal{L}(V, W)$ è un sottospazio vettoriale dello spazio delle applicazioni da V a W e quindi che, con le operazioni definite come nell'esempio 3 sopra, è uno spazio vettoriale. Una applicazione lineare $f: V \rightarrow \mathbb{K}$ a valori nel campo degli scalari (considerato come spazio vettoriale) a volte si dice *funzionale lineare* e lo spazio $\mathcal{L}(V, \mathbb{K})$ si dice *duale* di V e si denota $V^* = \mathcal{L}(V, \mathbb{K})$.

Esempio 1. Siano V e W due spazi vettoriali. L'applicazione nulla $N: V \rightarrow W$ definita per ogni $w \in W$ da $N(w) = O_W$ è lineare.

Esempio 2. Per ogni spazio vettoriale V l'applicazione identica $Id_V: V \rightarrow V$ è lineare.

Esempio 3. Siano O e O' due punti dello spazio e siano \mathcal{V}_O e $\mathcal{V}_{O'}$ gli spazi dei vettori dello spazio applicati in O e O' rispettivamente. La traslazione $\tau_{OO'}: \mathcal{V}_O \rightarrow \mathcal{V}_{O'}$ è un'applicazione lineare. Si osservi che $\tau_{OO'}$ è biettiva e che l'inversa è la traslazione $\tau_{O'O}: \mathcal{V}_{O'} \rightarrow \mathcal{V}_O$, anch'essa un'applicazione lineare.

Esempio 4. Sia $A \in M_{m,n}(\mathbb{K})$ una matrice con m righe e n colonne. Allora per ogni $x \in \mathbb{K}^n$ si ha $Ax \in \mathbb{K}^m$. Dunque è definita un'applicazione $L_A: \mathbb{K}^n \rightarrow \mathbb{K}^m$ mediante $L_A(x) = Ax$. Siano $x, y \in \mathbb{K}^n$ e $\lambda \in \mathbb{K}$. Se $A = (A^1, \dots, A^n)$ è l'espressione di A secondo le sue colonne, allora $L_A(x) = x_{11}A^1 + \dots + x_{n1}A^n$ e quindi

$$\begin{aligned} L_A(x + y) &= A(x + y) = \sum_{j=1}^n (x_j + y_j)A^j = \sum_{j=1}^n x_j A^j + \sum_{j=1}^n y_j A^j \\ &= Ax + Ay = L_A(x) + L_A(y) \end{aligned}$$

e

$$L_A(\lambda x) = A(\lambda x) = \sum_{j=1}^n (\lambda x^j)A^j = \lambda \sum_{j=1}^n x^j A^j = \lambda Ax = \lambda L_A(x)$$

e quindi L_A è lineare. Si osservi che, se $A = (A^1, \dots, A^n)$ è l'espressione di A secondo le sue colonne e $\{e_1, \dots, e_n\}$ è la base canonica di \mathbb{K}^n , allora $L_A(e_j) = A^j$ per ogni $j = 1, \dots, n$.

Esempio 5. Siano V uno spazio vettoriale su \mathbb{K} e $\mathcal{B} = \{v_1, \dots, v_n\}$ una sua base. L'applicazione $F_{\mathcal{B}}: V \rightarrow \mathbb{K}^n$ che ogni vettore di V associa le sue coordinate relative a \mathcal{B} è lineare.

Esempio 6. Sia $x_0 \in \mathbb{K}$. La funzione $v_{x_0}: \mathbb{K}[t] \rightarrow \mathbb{K}$, detta *valutazione in x_0* , definita da $v_{x_0}(p(t)) = p(x_0)$ è lineare. La derivata $D: \mathbb{K}[t] \rightarrow \mathbb{K}[t]$ definita da $D\left(\sum_{k=0}^n a_k t^k\right) = \sum_{k=1}^n k a_k t^{k-1}$ è un'applicazione lineare.

Una proprietà fondamentale delle applicazioni lineari consiste nel fatto che sono completamente determinate dal loro comportamento su una base:

Teorema 5.1: Siano V uno spazio vettoriale finitamente generato su \mathbb{K} e $\mathcal{B} = \{v_1, \dots, v_n\}$ una sua base. Se W è uno spazio vettoriale su \mathbb{K} e $w_1, \dots, w_n \in W$ allora esiste una unica applicazione lineare $T: V \rightarrow W$ tale che $T(v_j) = w_j$ per ogni $j = 1, \dots, n$. L'applicazione T è definita da

$$T\left(\sum_{j=1}^n x_j v_j\right) = \sum_{j=1}^n x_j w_j.$$

Dimostrazione: Vedi ABATE - de FABRITIIS pg. 96. □

Corollario 5.2: Siano V uno spazio vettoriale finitamente generato su \mathbb{K} , $\mathcal{B} = \{v_1, \dots, v_n\}$ una sua base e W un altro spazio vettoriale su \mathbb{K} . Allora se $S, T: V \rightarrow W$ sono applicazioni lineari, si ha che $S = T$ (ossia che $S(v) = T(v)$ per ogni $v \in V$) se e solo se $T(v_j) = S(v_j)$ per ogni $v_j \in \mathcal{B}$.

Dimostrazione: Immediata dal Teorema 5.1. □

Corollario 5.3: Siano $A, B \in M_{m,n}(\mathbb{K})$, allora $L_A = L_B$ se e solo se $A = B$.

Dimostrazione: Immediata dal Teorema 5.1 e dal fatto che $L_A(e_j) = A^j$ e $L_B(e_j) = B^j$ per ogni $j = 1, \dots, n$. □

Infine possiamo dimostrare che le applicazioni lineari fra spazi vettoriali numerici sono tutte e sole quelle definite da matrici:

Corollario 5.4: Sia $T: \mathbb{K}^n \rightarrow \mathbb{K}^m$ un'applicazione lineare. Se si pone $T(e_j) = A^j$ e $A = (A^1, \dots, A^n) \in M_{m,n}(\mathbb{K})$, allora $T = L_A$.

Dimostrazione: Immediata dal Teorema 5.1. □

Osservazione 4. Dal Corollario 5.4 segue immediatamente che l'applicazione

$$\Phi: M_{m,n}(\mathbb{K}) \rightarrow \mathcal{L}(\mathbb{K}^n, \mathbb{K}^m),$$

definita per ogni $A \in M_{m,n}(\mathbb{K})$ da $\Phi(A) = L_A$, è biettiva. In particolare l'applicazione Φ nel caso particolare in cui $m = 1$ permette di identificare il duale $(\mathbb{K}^n)^*$ con lo spazio delle righe $M_{1,n}(\mathbb{K})$ composte da n scalari.

Prima di andare avanti ricordiamo alcune nozioni generali sulle applicazioni. Un'applicazione $F: A \rightarrow B$ si dice *iniettiva* se per ogni $a_1, a_2 \in A$ si ha $F(a_1) = F(a_2)$ se e solo

se $a_1 = a_2$ o, equivalentemente, se $a_1 \neq a_2$ implica $F(a_1) \neq F(a_2)$. L'applicazione F si dice *suriettiva* se per ogni $b \in B$ esiste $a \in A$ tale che $F(a) = b$ ossia se, posto $Im(F) = \{b = F(a) \in B \mid a \in A\}$, risulta $Im(F) = B$. Infine l'applicazione F si dice *biettiva* se è sia iniettiva sia suriettiva. Le applicazioni biettive si dicono anche *invertibili*. Questa nomenclatura è giustificata dal fatto che se F è biettiva per ogni $b \in B$ esiste un unico elemento $a_b \in A$ tale che $F(a_b) = b$. Se si pone allora $F^{-1}(b) = a_b$ per ogni $b \in B$ rimane definita un'applicazione $F^{-1}: B \rightarrow A$ detta *inversa* di F . Per costruzione allora si ha $F^{-1} \circ F = Id_A$ e $F \circ F^{-1} = Id_B$.

A volte è utile la seguente osservazione:

Proposizione 5.5: *Sia $F: A \rightarrow B$ fra due insiemi. Allora*

- (i) F è iniettiva se e solo se esiste una inversa sinistra ossia un'applicazione $G: B \rightarrow A$ tale che $G \circ F = Id_A$;
- (ii) F è suriettiva se e solo se esiste una inversa destra ossia un'applicazione $H: B \rightarrow A$ tale che $F \circ H = Id_B$;
- (iii) F è biettiva se e solo se esiste un'applicazione $L: B \rightarrow A$ tale che $L \circ F = Id_A$ e $F \circ L = Id_B$. Se esiste, $L = F^{-1}$.

Dimostrazione: (i) Se esiste $G: B \rightarrow A$ tale che $G \circ F = Id_A$, allora per ogni $a_1, a_2 \in A$ si ha $F(a_1) = F(a_2)$ se e solo se $a_1 = G(F(a_1)) = G(F(a_2)) = a_2$ e quindi F è iniettiva. Se F è iniettiva sia $G: B \rightarrow A$ definita nel modo seguente. Si fissi arbitrariamente $a_0 \in A$, Sia $b \in B$. Allora o $b \in Im(F)$ o $b \notin Im(F)$. Se $b \in Im(F)$ esiste un unico elemento $a_b \in A$ tale che $F(a_b) = b$ e poniamo $G(b) = a_b$. Se $b \notin Im(F)$, poniamo $G(b) = a_0$. L'applicazione $G: B \rightarrow A$ è tale che $G \circ F = Id_A$.

(ii) Se esiste $H: B \rightarrow A$ tale che $F \circ H = Id_B$, per ogni $b \in B$ si ha $b = F(H(b))$ e quindi F è suriettiva. Se F è suriettiva, per ogni $b \in B$ si scelga un elemento $a_b \in \{a \in A \mid F(a) = b\} \neq \emptyset$ e si definisca $H(b) = a_b$. L'applicazione $H: B \rightarrow A$ è tale che $F \circ H = Id_B$.

(iii) Se esiste $L: B \rightarrow A$ tale che $L \circ F = Id_A$ e $F \circ L = Id_B$, allora F è biettiva per (i) e (ii). Inoltre per ogni $b \in B$ si ha $L(b) = F^{-1} \circ F(L(b)) = F^{-1}(F \circ L(b)) = F^{-1}(b)$ e quindi $L = F^{-1}$. Se F è biettiva, basta porre $L = F^{-1}$. \square

Sia $T: V \rightarrow W$ un'applicazione lineare. L'immagine $Im(T)$ e il nucleo $Ker(T)$ di T sono definiti rispettivamente da:

$$Im(T) = \{w = T(v) \in W \mid v \in V\} \quad \text{e} \quad Ker(T) = \{v \in V \mid T(v) = O_W\}.$$

Esempio 7. Per una matrice $A \in M_{m,n}(\mathbb{K})$ si consideri l'applicazione $L_A: \mathbb{K}^n \rightarrow \mathbb{K}^m$. Allora, se A^1, \dots, A^n sono le colonne di A allora $Im(L_A) = Span(A^1, \dots, A^n)$ e $Ker(L_A)$ è lo spazio delle soluzioni del sistema omogeneo $Ax = 0$. In seguito si scriverà spesso $ImA = Im(L_A)$ e $Ker(A) = Ker(L_A)$.

Abbiamo la seguente

Proposizione 5.6: *Sia $T: V \rightarrow W$ un'applicazione lineare. Allora*

- (i) $Im(T)$ e $Ker(T)$ sono sottospazi vettoriali rispettivamente di W e di V ;
- (ii) T è suriettiva se e solo se $Im(T) = W$;
- (iii) T è iniettiva se e solo se $Ker(T) = \{O_V\}$.

Dimostrazione: (i) è una verifica ovvia lasciata al lettore.

(ii) È la definizione di suriettività.

(iii) Se T è iniettiva allora è immediato dalla definizione che $\text{Ker}(T) = \{O_V\}$. Supponiamo, viceversa, che si abbia $\text{Ker}(T) = \{O_V\}$. Allora se, $v_1, v_2 \in V$ si ha $T(v_1) = T(v_2)$, dato che T è lineare, segue che $T(v_1 - v_2) = T(v_1) - T(v_2) = O_W$ e quindi che $v_1 - v_2 \in \text{Ker}(T) = \{O_V\}$. \square

È utile segnalare come applicazioni iniettive, suriettive e biettive trasformano sistemi di generatori e sistemi liberi:

Proposizione 5.7: Sia $T: V \rightarrow W$ un'applicazione lineare. Allora:

- (i) Sia $\{v_1, \dots, v_n\}$ un sistema di generatori di V . Allora $\{T(v_1), \dots, T(v_n)\}$ è un sistema di generatori di $\text{Im}(T)$. Dunque T è suriettiva se e solo se $\{T(v_1), \dots, T(v_n)\}$ è un sistema di generatori di W ;
- (ii) Se T è iniettiva e $\{v_1, \dots, v_n\}$ è un sistema libero di vettori di V allora $\{T(v_1), \dots, T(v_n)\}$ è un sistema libero di vettori di W ;
- (iii) Sia $\{v_1, \dots, v_n\}$ una base di V . Allora T è biettiva se e solo se $\{T(v_1), \dots, T(v_n)\}$ è una base di W .

Dimostrazione: (i) Sia $w \in \text{Im}(T)$. Allora per qualche $v \in V$ si deve avere $w = T(v)$. Dato che $\{v_1, \dots, v_n\}$ è un sistema di generatori di V , allora $v = x_1v_1 + \dots + x_nv_n$ per opportuni scalari x_1, \dots, x_n e quindi, per la linearità di T si ha $w = T(v) = T(x_1v_1 + \dots + x_nv_n) = x_1T(v_1) + \dots + x_nT(v_n)$ e quindi $\{T(v_1), \dots, T(v_n)\}$ è un sistema di generatori di $\text{Im}(T)$. Dato che T è suriettiva se e solo se $W = \text{Im}(T)$ il resto della tesi è immediato.

(ii) Siano s_1, \dots, s_n scalari tali che $x_1T(v_1) + \dots + x_nT(v_n) = O_W$. Per la linearità di T , allora $T(x_1v_1 + \dots + x_nv_n) = O_W$ ossia $x_1v_1 + \dots + x_nv_n \in \text{Ker}(T) = \{O_V\}$ e la tesi segue.

(iii) Da (i) e (ii) segue immediatamente che se T è biettiva allora $\{T(v_1), \dots, T(v_n)\}$ è una base di W . Se $\{T(v_1), \dots, T(v_n)\}$ è una base di W allora da (i) segue che T è suriettiva. D'altra parte è anche iniettiva dato che se $v = x_1v_1 + \dots + x_nv_n \in \text{Ker}(T)$ allora

$$O_W = T(v) = T(x_1v_1 + \dots + x_nv_n) = x_1T(v_1) + \dots + x_nT(v_n)$$

e quindi, dato che $\{T(v_1), \dots, T(v_n)\}$ è una base di W , necessariamente $x_1 = \dots = x_n = 0$ e di conseguenza $v = x_1v_1 + \dots + x_nv_n = \{O_V\}$. \square

Le applicazioni lineari biettive hanno un ruolo speciale nella teoria degli spazi vettoriali. Una importante proprietà consiste nel fatto che l'inversa di una applicazione lineare biettiva è automaticamente lineare:

Proposizione 5.8: Sia $T: V \rightarrow W$ un'applicazione lineare biettiva. Allora $T^{-1}: W \rightarrow V$ è un'applicazione lineare.

Dimostrazione: Siano $w_1, w_2 \in W$. Allora $T^{-1}(w_1) = v_1$ e $T^{-1}(w_2) = v_2$ sono gli unici vettori di V tali che $T(v_i) = w_i$ per $i = 1, 2$. Dunque, utilizzando la linearità di T :

$$T^{-1}(w_1 + w_2) = T^{-1}(T(v_1) + T(v_2)) = T^{-1}(T(v_1 + v_2)) = v_1 + v_2 = T^{-1}(w_1) + T^{-1}(w_2).$$

Il fatto che $T^{-1}(\lambda w) = \lambda T^{-1}(w)$ per ogni scalare λ e $w \in W$ si verifica in modo simile. \square

Un'applicazione lineare biettiva $T: V \rightarrow W$ si dice *isomorfismo*. Se esiste un isomorfismo fra due spazi vettoriali V e W , questi si dicono *isomorfi*. Per riconoscere se due spazi vettoriali finitamente generati sono isomorfi, basta confrontare le loro dimensioni:

Proposizione 5.9: *Due spazi vettoriali finitamente generati sono isomorfi se e solo se hanno la stessa dimensione.*

Dimostrazione: Se due spazi vettoriali V e W finitamente generati sono isomorfi, allora hanno la stessa dimensione perché un'applicazione lineare biettiva trasforma basi in basi. Supponiamo viceversa che due spazi vettoriali V e W abbiano la stessa dimensione: $\dim(V) = n = \dim(W)$ e siano $\{v_1, \dots, v_n\}$ e $\{w_1, \dots, w_n\}$ basi rispettivamente per V e W . Sia $T: V \rightarrow W$ l'unica applicazione lineare, esistente per il Teorema 5.1, tale che $T(v_j) = w_j$ per ogni $j = 1, \dots, n$. L'applicazione T è definita da

$$T \left(\sum_{j=1}^n x_j v_j \right) = \sum_{j=1}^n x_j w_j.$$

Rimane da dimostrare che T è biettiva. Dato che per costruzione T trasforma una base in una base, questo segue immediatamente dalla Proposizione 5.7. \square

Un importante invariante delle applicazioni lineari è il rango:

Definizione. Sia V uno spazio vettoriale finitamente generato e sia $T: V \rightarrow W$ un'applicazione lineare. Si dice *rango* di T il numero $rg(T) = \dim(\text{Im}(T))$. Per una matrice $A \in M_{m,n}(\mathbb{K})$ si dice *rango* di A il numero $rg(A) = \dim(\text{Im}(L_A))$.

La relazione fra le dimensioni del dominio di un'applicazione lineare, della sua immagine e del suo nucleo è dato dal seguente importante risultato:

Teorema 5.10: (Teorema della dimensione) *Sia V uno spazio vettoriale finitamente generato e sia $T: V \rightarrow W$ un'applicazione lineare. Allora si ha*

$$\dim(V) = rg(T) + \dim(\text{Ker}(T)).$$

Dimostrazione: Vedi ABATE - de FABRITIIS pg. 99. \square

Come conseguenza immediata del Teorema della dimensione si ha una corrispondente formula per matrici. Se $A \in M_{m,n}(\mathbb{K})$, posto $\text{Ker}(A) = \text{Ker}(L_A)$, dato che il dominio di L_A è \mathbb{K}^n , si ha

$$n = rg(A) + \dim(\text{Ker}(A)).$$

Corollario 5.11: *Sia $T: V \rightarrow W$ un'applicazione lineare fra spazi vettoriali finitamente generati. Allora:*

- (i) T è iniettiva se e solo se $rg(T) = \dim(V)$;
- (ii) T è suriettiva se e solo se $rg(T) = \dim(W)$;
- (iii) Se $\dim(V) = \dim(W)$, allora T è biettiva se e solo se è iniettiva se e solo se è suriettiva.

Dimostrazione: Vedi ABATE - de FABRITIIS pg. 100. \square

Teorema 5.12: (di Rouché-Capelli) *Un sistema lineare $Ax = b$ di m equazioni in n incognite ha soluzione se e solo se il rango della matrice $A = (A^1, \dots, A^n)$ del sistema è uguale al rango della matrice completa $A' = (A^1, \dots, A^n, b)$ del sistema. Qualora esista soluzione, questa è unica se e solo se $rg(A) = n$.*

Dimostrazione: Vedi ABATE - de FABRITIIS pg. 100. \square

6. Dualità e rango.

Sia V uno spazio vettoriale finitamente generato. Si dimostra che il suo duale $V^* = \mathcal{L}(V, \mathbb{K})$ è isomorfo a V . Precisamente:

Proposizione 6.1: *Sia V uno spazio vettoriale di dimensione n . Allora il suo duale V^* ha dimensione n ed è quindi isomorfo a V . Infatti per ogni base $\mathcal{B} = \{v_1, \dots, v_n\}$ di V esiste una base $\mathcal{B}^* = \{v_1^*, \dots, v_n^*\}$ di V^* , chiamata **base duale** di \mathcal{B} , dove i funzionali lineari v_1^*, \dots, v_n^* sono unicamente determinati da*

$$v_i^*(v_j) = \delta_{ij} = \begin{cases} 0 & \text{se } i \neq j \\ 1 & \text{se } i = j \end{cases}.$$

L'espressione di $f \in V^*$ rispetto alla base \mathcal{B}^* è data da

$$f = \sum_{i=1}^n f(v_i)v_i^*.$$

Dimostrazione: Dimostriamo per cominciare che \mathcal{B}^* è un sistema libero. Siano t_1, \dots, t_n scalari tali che

$$t_1v_1^* + \dots + t_nv_n^* = O_{V^*}$$

dove O_{V^*} è il funzionale lineare nullo, vettore nullo di V^* . Ma allora, per ogni $j = 1, \dots, n$ si ha

$$0 = O_{V^*}(v_j) = t_1v_1^*(v_j) + \dots + t_nv_n^*(v_j) = t_j$$

e quindi \mathcal{B}^* è libero. Sia ora $f \in V^*$. Allora $f = \sum_{i=1}^n f(v_i)v_i^*$ perché i due funzionali lineari coincidono sulla base \mathcal{B} . Infatti per ogni $j = 1, \dots, n$ si ha

$$\left(\sum_{i=1}^n f(v_i)v_i^* \right) (v_j) = \sum_{i=1}^n f(v_i)v_i^*(v_j) = \sum_{i=1}^n f(v_i)\delta_{ij} = f(v_j).$$

Dunque \mathcal{B}^* è anche un sistema di generatori e quindi è una base. □

Osservazione. Se $\mathcal{B} = \{v_1, \dots, v_n\}$ è base di uno spazio vettoriale V e $\mathcal{B}^* = \{v_1^*, \dots, v_n^*\}$ è la base di V^* duale di \mathcal{B} , allora, per costruzione, per ogni vettore $v = x_1v_1 + \dots + x_nv_n \in V$ si ha $v_j^*(v) = x_j$ per ogni $j = 1, \dots, n$ ossia v_j^* è il funzionale lineare che assegna a ogni vettore di V la sua j -esima coordinata relativa alla base \mathcal{B} .

Osservazione. Abbiamo già osservato che l'applicazione

$$\Phi: M_{m,n}(\mathbb{K}) \rightarrow \mathcal{L}(\mathbb{K}^n, \mathbb{K}^m),$$

definita per ogni $A \in M_{m,n}(\mathbb{K})$ da $\Phi(A) = L_A$, è biettiva e quindi nel caso particolare in cui $m = 1$, l'isomorfismo Φ permette di identificare il duale $(\mathbb{K}^n)^*$ con lo spazio delle righe $M_{1,n}(\mathbb{K})$ composte da n scalari. In questo caso, utilizzando l'identificazione di funzionali lineari (\mathbb{K}^n) con vettori riga, la base duale della base canonica $\mathcal{C} = \{e_1, \dots, e_n\}$ è la base $\mathcal{C}^* = \{e_1^*, \dots, e_n^*\}$ dove per ogni $j = 1, \dots, n$

$$e_j^* = (0, \dots, 0, 1, 0 \dots 0).$$

Per uno sottospazio vettoriale di uno spazio vettoriale vogliamo considerare l'insieme di tutti i funzionali lineari che si annullano su di esso:

Definizione. Sia $W \subset V$ un sottospazio vettoriale di uno spazio vettoriale V , L'annullatore di W è definito da

$$W^0 = \{f \in V^* \mid F|_W = 0\} = \{f \in V^* \mid F(w) = 0 \quad \forall w \in W\}.$$

Osservazione. L'annullatore di un sottospazio è un sottospazio vettoriale del duale.

Per spazi finitamente generati la dimensione dell'annullatore è uguale al numero atteso di condizioni lineari che un sottospazio "deve" soddisfare. Precisamente abbiamo:

Teorema 6.2: Sia V uno spazio vettoriale finitamente generato e sia $W \subset V$ un sottospazio vettoriale. Allora

$$\dim(V) = \dim(W) + \dim(W^0)$$

Dimostrazione: Sia $\mathcal{B}' = \{w_1, \dots, w_k\}$ una base di W e sia $\mathcal{B} = \{w_1, \dots, w_k, w_{k+1}, \dots, w_n\}$ un suo completamento a una base di V . Sia $\mathcal{B}^* = \{w_1^*, \dots, w_k^*, w_{k+1}^*, \dots, w_n^*\}$ la base duale di \mathcal{B} di V^* . Dimostriamo che $\mathcal{C}^* = \{w_{k+1}^*, \dots, w_n^*\}$ è una base di $\dim(W^0)$, la tesi segue allora immediatamente. Per cominciare $\mathcal{C}^* \subset W^0$. Infatti per ogni $i = k+1, \dots, n$ se $j = 1, \dots, k$ si ha $w_i^*(w_j) = 0$ e quindi necessariamente $w_i^* \in W^0$. Visto che \mathcal{C}^* è sicuramente un sistema libero dato che è un sottoinsieme di una base, rimane solo da dimostrare che è un sistema di generatori per W^0 . Sia $f \in W^0$. Allora si deve avere $f(w_i) = 0$ per ogni $i = 1, \dots, k$ e quindi

$$f = \sum_{i=1}^n f(w_i)w_i^* = \sum_{k+i=1}^n f(w_i)w_i^* \in \text{Span}(\mathcal{C}^*).$$

□

Vogliamo ora applicare queste considerazioni allo studio del rango delle matrici. In particolare vogliamo dimostrare che lo spazio generato dalle colonne di una matrice ha la stessa dimensione dello spazio dalle righe.

Cominciamo ricordando che se $A = (a_{ij}) \in M_{m,n}(\mathbb{K})$ è una matrice con m righe e n colonne, la *trasposta* di A è la matrice $A^T = (b_{ji}) \in M_{n,m}(\mathbb{K})$ con n righe e m colonne tale che $b_{ji} = a_{ij}$. In parole povere A^T è la matrice ottenuta "scrivendo ordinatamente in colonne" le righe di A .

La proprietà cui accennavamo sopra più precisamente si esprime dicendo che i ranghi di una matrice e della sua trasposta coincidono. Un fatto cruciale per la dimostrazione di questo importante risultato, è il seguente:

Lemma 6.3: $(\text{Im}(A))^0$ è isomorfo a $\text{Ker}(A^T)$

Dimostrazione: Sia $\{e_1, \dots, e_m\}$ la base canonica di \mathbb{K}^m e sia $\{\epsilon_1, \dots, \epsilon_m\}$ la base di $(\mathbb{K}^m)^*$ duale della base canonica. Ricordiamo che dalla definizione segue immediatamente

che, per ogni $i = 1, \dots, m$, se $x = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \in \mathbb{K}^m$, si ha $\epsilon_i(x) = x_i$.

Allora l'applicazione lineare $T: \mathbb{K}^m \rightarrow (\mathbb{K}^m)^*$ unicamente definita ponendo $T(e_i) = \epsilon_i$ per ogni $i = 1, \dots, m$ è un isomorfismo. Per dimostrare la tesi ci basterà dimostrare che $T(Ker(A^T)) = (Im(A))^0$. In questo caso infatti la restrizione a $Ker(A^T)$ di T è un isomorfismo fra $Ker(A^T)$ e $(Im(A))^0$.

Sia $A = (a_{ij}) = (A^1, \dots, A^n) \in M_{m,n}(\mathbb{K})$. Dobbiamo provare che $z = \begin{pmatrix} z_1 \\ \vdots \\ z_m \end{pmatrix} \in Ker(A^T)$

se e solo se $f = T(z) = \sum_{i=1}^m z_i \epsilon_i \in (Im(A))^0$. Ricordando che $\epsilon_i(A^j) = a_{ij}$ per ogni $i = 1, \dots, m$ e $j = 1, \dots, n$, abbiamo

$$z = \begin{pmatrix} z_1 \\ \vdots \\ z_m \end{pmatrix} \in Ker(A^T) \iff 0 = \sum_{i=1}^m a_{ij} z_i = \sum_{i=1}^m z_i \epsilon_i(A^j) = f(A^j) \text{ per ogni } j = 1, \dots, n$$

e quindi se e solo se $f = T(z) \in (Im(A))^0$.

□

Abbiamo allora il seguente importante risultato:

Teorema 6.4: Sia $A \in M_{m,n}(\mathbb{K})$, allora $rg(A) = rg(A^T)$.

Dimostrazione: Utilizzando il Teorema della dimensione, il Teorema 6.2 e il Lemma 6.3, si ha

$$rg(A^T) = m - dim(Ker(A^T)) = m - dim((Im(A))^0) = m - (m - dim(Im(A))) = rg(A).$$

□

7. Matrici.

7.1 Riduzione a scala e rango

Una matrice $S \in M_{m,n}(\mathbb{K})$ si dice una *matrice a scala* se è la matrice nulla o se è della forma

$$S = \begin{pmatrix} 0 & \dots & 0 & s_{1j_1} & \dots & \dots & \dots & \dots & \dots & \dots & s_{1n} \\ 0 & \dots & 0 & 0 & \dots & 0 & s_{2j_2} & \dots & \dots & \dots & s_{2n} \\ 0 & \dots & 0 & 0 & \dots & \dots & 0 & \dots & \dots & \dots & s_{2n} \\ \vdots & \dots & \vdots & \vdots & \dots & \vdots & \dots & \dots & \dots & \dots & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 & \dots & 0 & s_{rj_r} & \dots & s_{rn} \\ 0 & \dots & 0 & 0 & \dots & \dots & \dots & \dots & 0 & \dots & 0 \\ \vdots & \vdots \\ 0 & \dots & 0 & 0 & \dots & \dots & \dots & \dots & \dots & \dots & 0 \end{pmatrix} \quad (7.1)$$

dove $1 \leq r \leq m$ e j_1, \dots, j_r sono interi tali che $1 \leq j_1 < j_2 < \dots < j_r \leq n$ e, per ogni $k = 1, \dots, r$ si ha $s_{kj} = 0$ per ogni $1 \leq j < j_k$ e $s_{kj_k} \neq 0$ mentre per $k > r$ si ha $s_{kj} = 0$ per ogni j . I numeri $s_{kj_k} \neq 0$ si dicono *pivot* della matrice S .

Per le matrici a scala è molto semplice calcolare il rango, trovare una base per lo spazio generato dalle colonne e determinare la risolubilità dei sistemi lineari associati. Come al solito denotiamo con e_1, \dots, e_n gli elementi della base canonica di \mathbb{K}^n :

Proposizione 7.1: Sia $S \in M_{m,n}(\mathbb{K})$ una matrice a scala con r pivot $s_{1j_1}, \dots, s_{rj_r}$. Allora

- (i) $rg(S) = r$ e $Im(S) = Span(e_1, \dots, e_r)$; inoltre $\{S^{j_1}, \dots, S^{j_r}\}$ è una base di $Im(S)$.
- (ii) $Ker(S)$, il sottospazio delle soluzioni del sistema omogeneo $Sx = 0$, ha dimensione $n - r$.
- (iii) Se $c \in \mathbb{K}^m$ il sistema $Sx = c$ ha soluzione se e solo se le ultime $m - r$ coordinate di c sono zero.

Dimostrazione: Per costruzione tutte le colonne di S sono in $Span(e_1, \dots, e_r)$ e quindi $Im(S) \subset Span(e_1, \dots, e_r)$. Dato che $dim(Span(e_1, \dots, e_r)) = r$, per provare (i), basterà verificare che le colonne $\{S^{j_1}, \dots, S^{j_r}\}$ sono linearmente indipendenti. Sia

$$S' = (S^{j_1}, \dots, S^{j_r}) = \begin{pmatrix} s_{1j_1} & \dots & \dots & \dots & \dots & \dots \\ 0 & s_{2j_2} & \dots & \dots & \dots & \dots \\ 0 & 0 & s_{3j_3} & \dots & \dots & \dots \\ \vdots & \vdots & \ddots & \dots & \dots & \dots \\ \vdots & \vdots & \ddots & \ddots & \dots & \dots \\ 0 & \dots & 0 & \dots & 0 & s_{rj_r} \\ 0 & \dots & \dots & \dots & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & \dots & \dots & \dots & 0 \end{pmatrix}.$$

Allora il sistema $S'x = 0$ ha soluzione unica e questo equivale al fatto che le colonne S^{j_1}, \dots, S^{j_r} sono linearmente indipendenti. L'enunciato (ii) è immediato da (i) e dal teorema della dimensione mentre (iii) segue da (i) e dal fatto che $Sx = c$ ha soluzione se e solo $c \in Im(S)$. \square

Mediante l'eliminazione di Gauss è sempre possibile ridurre a scala una matrice:

Proposizione 7.2: Sia $A \in M_{m,n}(\mathbb{K})$. Se A non è a scala, con una successione finita delle due seguenti operazioni sulle righe:

- (I) sostituire una riga con la somma di questa più un multiplo di un'altra,
- (II) scambiare due righe,

si può sempre passare da A a una matrice a scala $S \in M_{m,n}(\mathbb{K})$.

Dimostrazione: Sia $0 \leq k \leq n$ l'intero più grande tale che la matrice B costituita dalle prime k colonne di A è a scala. Se B non ha righe nulle, allora A è a scala e non c'è nulla da dimostrare. Se $k = n$, di nuovo A è a scala e non c'è nulla da dimostrare. Possiamo dunque assumere $0 \leq k < n$. Sia r l'indice della prima riga identicamente nulla di B se $k > 0$ e $r = 1$ se $k = 0$. Allora i numeri $a_{rk+1}, \dots, a_{mk+1}$ non possono essere tutti nulli, altrimenti esisterebbe una matrice a scala costituita da $k + 1$ colonne di A . A meno di scambiare le righe di A (operazione effettuabile mediante una successione finita di scambi di righe!) si può supporre che $a_{rk+1} \neq 0$. Se per ogni indice $r < i \leq m$ si somma alla riga i -esima di A la r -esima riga moltiplicata per $-\frac{a_{ik+1}}{a_{rk+1}}$, si ottiene una matrice A' tale che esiste una matrice a scala B' costituita da $k + 1$ colonne di A' . Iterando la procedura, con un numero finito di operazioni si ottiene una matrice a scala S come richiesto. \square

Una matrice a scala S ottenuta da A utilizzando l'eliminazione di Gauss, ossia la procedura illustrata nella Proposizione 7.2, si dice *riduzione a scala* di A . La riduzione a scala di una matrice ci permette di calcolare facilmente il rango di una matrice qualunque e di risolvere sistemi lineari.

Teorema 7.3: Sia $A \in M_{m,n}(\mathbb{K})$ e sia $S \in M_{m,n}(\mathbb{K})$ una sua riduzione a scala. Se $b \in \mathbb{K}^m$, sia $c \in \mathbb{K}^m$ l'unico vettore tale che effettuando i passi dell'eliminazione di Gauss che riduce A a S , si ottiene la matrice $(S|c)$ partendo dalla matrice $(A|b)$. Allora:

- (i) l'insieme delle soluzioni del sistema $Ax = b$ coincide con l'insieme delle soluzioni del sistema $Sx = c$ (ossia i due sistemi sono equivalenti);
- (ii) $\text{Ker}(A) = \text{Ker}(S)$;
- (iii) $\text{rg}(A) = \text{rg}(S)$;
- (iv) se S^{j_1}, \dots, S^{j_r} sono le colonne della matrice a scala S corrispondenti ai pivot $s_{1j_1}, \dots, s_{rj_r}$ di S , allora $\{A^{j_1}, \dots, A^{j_r}\}$ è una base di $\text{Im}(A)$.

Dimostrazione: (i) è immediata dato che in un sistema lineare sostituendo a un'equazione la somma di questa con un multiplo di un'altra oppure scambiando equazioni si ottiene un sistema equivalente. (ii) è (i) nel caso in cui $b = 0$. Usando il teorema della dimensione, si ha (iii) dato che

$$\text{rg}(A) = n - \dim \text{Ker}(A) = n - \dim \text{Ker}(S) = \text{rg}(S).$$

Infine per dimostrare (iv) basta provare che le colonne A^{j_1}, \dots, A^{j_r} sono linearmente indipendenti. Se $A' = (A^{j_1}, \dots, A^{j_r})$, questo equivale a dimostrare che il sistema $A'x = 0$ ha soluzione unica. Questo è vero dato che la matrice $S' = (S^{j_1}, \dots, S^{j_r})$ è una riduzione a scala di A' e che il sistema $S'x = 0$ ha soluzione unica. \square

Osservazione Si osservi che se $A \in M_{m,n}(\mathbb{K})$ e $S \in M_{m,n}(\mathbb{K})$ è una sua riduzione a scala, anche se $\text{rg}(A) = \text{rg}(S)$, in generale $\text{Im}(A) \neq \text{Im}(S)$!

7.2 Prodotto di matrici, invertibilità

Ricordiamo per cominciare la definizione di prodotto di matrici. Se $A = (a_1 \ \dots \ a_n) \in M_{1,n}(\mathbb{K})$ è una riga (con n colonne) e $B = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in M_{n,1}(\mathbb{K})$ è una colonna (con n righe), allora il prodotto (riga per colonna) di A per B è lo scalare

$$AB = \sum_{l=1}^n a_l b_l. \quad (7.2)$$

Più in generale siano $A = (a_{il}) \in M_{m,n}(\mathbb{K})$ e $B = (b_{lj}) \in M_{n,p}(\mathbb{K})$. Se A_1, \dots, A_m sono le righe di A e B^1, \dots, B^n sono le colonne di B , dato che A ha n colonne e B ha n righe, utilizzando (7.2) ha senso considerare il prodotto $A_i B^j$ per ogni $i = 1, \dots, m$ e $j = 1, \dots, p$. Per definizione il prodotto di A per B è la matrice $C = AB = (c_{ij}) \in M_{m,p}(\mathbb{K})$ tale che per ogni $i = 1, \dots, m$ e $j = 1, \dots, p$ si ha

$$c_{ij} = A_i B^j = \sum_{l=1}^n a_{il} b_{lj}. \quad (7.3)$$

Il prodotto definito da (7.3) si può caratterizzare in termini di applicazioni lineari nel modo seguente:

Proposizione 7.4: Il prodotto $C = AB \in M_{m,p}(\mathbb{K})$ di $A = (a_{il}) \in M_{m,n}(\mathbb{K})$ e $B = (b_{lj}) \in M_{n,p}(\mathbb{K})$ è l'unica matrice tale che $L_C: \mathbb{K}^p \rightarrow \mathbb{K}^m$ è la composizione $L_A \circ L_B$ delle applicazioni lineari $L_B: \mathbb{K}^p \rightarrow \mathbb{K}^n$ e $L_A: \mathbb{K}^n \rightarrow \mathbb{K}^m$.

Dimostrazione: Basta verificare che per ogni vettore e_j della base canonica di \mathbb{K}^p si abbia $L_C(e_j) = L_A \circ L_B(e_j)$. Usando (7.3), si ha

$$L_C(e_j) = C^j = \begin{pmatrix} c_{1j} \\ \vdots \\ c_{mj} \end{pmatrix} = \begin{pmatrix} A_1 B^j \\ \vdots \\ A_m B^j \end{pmatrix}.$$

D'altra parte

$$L_A \circ L_B(e_j) = L_A(L_B(e_j)) = L_A(Be_j) = L_A(B^j) = AB^j = \begin{pmatrix} A_1 B^j \\ \vdots \\ A_m B^j \end{pmatrix}.$$

□

Osservazione. Il prodotto AB di due matrici A, B non è sempre definito dato che occorre che il numero delle colonne di A sia uguale al numero delle righe di B . Anche quando hanno senso ambedue, come nel caso di matrici quadrate dello stesso ordine, in generale $AB \neq BA$.

Chiamiamo

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

la matrice identica di ordine n . A volte invece di I_n si scrive semplicemente I . Indicheremo inoltre con O la matrice nulla di qualunque dimensione. Valgono le seguenti proprietà per il prodotto:

Proposizione 7.5: Siano A, B, C matrici delle dimensioni giuste affinché i prodotti che scriveremo abbiano senso e sia $\lambda \in \mathbb{K}$

- (i) $A(B + C) = AB + AC$ e $(B + C)A = BA + CA$;
- (ii) $(\lambda A)B = \lambda(AB)$;
- (iii) $A(BC) = (AB)C$;
- (iv) $AI = A, IA = A, AO = O$ e $OA = O$
- (v) $(AB)^T = B^T A^T$.

Dimostrazione: per le dimostrazioni vedi ABATE - de FABRITIIS pg 134. □

Applichiamo ora queste nozioni a matrici quadrate. Denotiamo con $M_n(\mathbb{K})$ lo spazio delle matrici quadrate di ordine n . Se $A, B \in M_n(\mathbb{K})$, allora il prodotto AB è sempre definito ed è una matrice quadrata di ordine n . Dunque il prodotto di matrici ristretta a matrici quadrate dello stesso ordine, definisce una vera e propria operazione. In particolare la matrice svolge il ruolo di elemento neutro per questa operazione dato che per ogni matrice $A \in M_n(\mathbb{K})$ si ha: $AI_n = I_n A = A$. Dunque la matrice I_n si comporta come il numero 1 per il prodotto di numeri. Per ogni numero $a \neq 0$ esiste un unico inverso a^{-1} tale che

$aa^{-1} = a^{-1}a = 1$. Per il prodotto di matrici la nozione di invertibilità si esprime nel modo seguente:

Definizione. Una matrice quadrata $A \in M_n(\mathbb{K})$ si dice *invertibile* se esiste una matrice $B \in M_n(\mathbb{K})$ tale che

$$AB = BA = I_n. \quad (7.4)$$

Se $A \in M_n(\mathbb{K})$ è invertibile, la matrice $B \in M_n$ tale che valga (7.4) è unica, si dice *inversa* di A e si denota A^{-1} . L'unicità si dimostra facilmente. Se ci fossero due matrici $B, B' \in M_n(\mathbb{K})$ tali che $AB = BA = I_n = AB' = B'A$, allora si avrebbe:

$$B' = B'I_n = B'(AB) = (B'A)B = I_nB = B.$$

L'insieme delle matrici di ordine n invertibili, con l'operazione di prodotto di matrici, forma dunque un *gruppo* ossia è un insieme con una operazione associativa, con elemento neutro e tale che ogni elemento ha un inverso. Il gruppo delle matrici invertibili di ordine n si chiama *gruppo lineare* e si denota $GL_n(K)$. Come abbiamo già osservato il gruppo lineare, se $n > 1$, non è commutativo.

Osservazione. Sia $A \in GL_n(K)$. Allora $A^{-1} \in GL_n(K)$ con inversa $(A^{-1})^{-1} = A$ e $A^T \in GL_n(K)$ con inversa $(A^T)^{-1} = (A^{-1})^T$. Inoltre se anche $B \in GL_n(K)$ allora $AB \in GL_n(K)$ con inversa $(AB)^{-1} = B^{-1}A^{-1}$. La verifica è un semplice esercizio.

Proposizione 7.6: Sia $A \in M_n(\mathbb{K})$ una matrice quadrata. Le seguenti affermazioni sono equivalenti:

- (i) A è invertibile;
- (ii) per ogni colonna $b \in \mathbb{K}^n$, il sistema $Ax = b$ ha soluzione unica;
- (iii) ogni riduzione a scala di A ha gli elementi sulla diagonale diversi da zero (in questo caso si dice che A è **non singolare**);
- (iv) esiste una riduzione a scala di A che ha gli elementi sulla diagonale diversi da zero;
- (v) il sistema omogeneo $Ax = 0$ ha soluzione unica $x = 0$;
- (vi) le colonne di A sono linearmente indipendenti;
- (vii) le righe di A sono linearmente indipendenti;
- (viii) $rg(A) = n$;
- (ix) $L_A: \mathbb{K}^n \rightarrow \mathbb{K}^n$ è suriettiva;
- (x) $L_A: \mathbb{K}^n \rightarrow \mathbb{K}^n$ è iniettiva;
- (xi) $L_A: \mathbb{K}^n \rightarrow \mathbb{K}^n$ è biettiva.

Dimostrazione: (i) \Rightarrow (ii): Se A è invertibile, per ogni colonna $b \in \mathbb{K}^n$, il sistema $Ax = b$ ha soluzione unica $x = A^{-1}b$.

(ii) \Rightarrow (iii): Ogni riduzione a scala di A produce un sistema che ha soluzione unica e quindi gli elementi sulla diagonale devono essere non nulli.

(iii) \Rightarrow (iv): Ovvio!

(iv) \Rightarrow (v): Ovvio!

(v) \Rightarrow (vi): Le colonne di A sono linearmente indipendenti se e solo se il sistema omogeneo $Ax = 0$ ha soluzione unica $x = 0$.

(vi) \Rightarrow (vii): Il massimo numero di colonne linearmente dipendenti è uguale al massimo numero di righe linearmente dipendenti.

(vii) \Rightarrow (viii): Il rango di una matrice è la dimensione dello spazio delle colonne.

(viii) \Rightarrow (ix): Il rango di una matrice è la dimensione dell'immagine di L_A .

(ix) \Rightarrow (x): $L_A: \mathbb{K}^n \rightarrow \mathbb{K}^n$ è suriettiva se e solo se è iniettiva.

(x) \Rightarrow (xi): $L_A: \mathbb{K}^n \rightarrow \mathbb{K}^n$ è iniettiva se e solo se è biettiva.

(xi) \Rightarrow (i): Se L_A è biettiva sia L_A^{-1} la sua inversa. Allora $L_A^{-1} = L_B$ per qualche matrice $B \in M_n(\mathbb{K})$: Allora A è invertibile con inversa B . Infatti per ogni $x \in \mathbb{K}^n$:

$$x = L_B(L_A(x)) = B(Ax) \quad \text{e} \quad x = L_A(L_B(x)) = A(Bx)$$

e quindi $L_B \circ L_A = Id_{\mathbb{K}^n} = L_{I_n}$ e $L_A \circ L_B = Id_{\mathbb{K}^n} = L_{I_n}$. Dunque $BA = I_n$ e $AB = I_n$. \square

Corollario 7.7: Una matrice $A \in M_n(\mathbb{K})$ è invertibile se e solo se ha un'inversa destra $B \in M_n(\mathbb{K})$ ossia se $AB = I_n$ oppure se e solo se ha un'inversa sinistra $C \in M_n(\mathbb{K})$ ossia se $CA = I_n$.

Dimostrazione: Esiste $B \in M_n(\mathbb{K})$ ossia se $AB = I_n$ se e solo se $L_A \circ L_B = Id_{\mathbb{K}^n}$. Questo succede se e solo se L_A è suriettiva (vedi Proposizione 5.5) e quindi se e solo se A è invertibile (Proposizione 7.6). Esiste $C \in M_n(\mathbb{K})$ ossia se $CA = I_n$ se e solo se $L_A \circ L_B = Id_{\mathbb{K}^n}$. Questo succede se e solo se L_A è iniettiva (vedi Proposizione 5.5) e quindi se e solo se A è invertibile (Proposizione 7.6). \square

Il Corollario 7.7 indica anche come procedere per stabilire se una matrice A è invertibile e nel caso trovare l'inversa:

I passo: Mediante un'eliminazione di Gauss si riduce a scala A e si determina se A è non singolare.

Il passo: Se A è non singolare i sistemi $Ax = e_1, \dots, Ax = e_n$ hanno soluzioni uniche B^1, \dots, B^n . L'inversa di A è la matrice $A^{-1} = (B^1 \dots B^n)$ dato che $A(B^1 \dots B^n) = (e_1 \dots e_n) = I_n$.

Questa idea si può trasformare in un algoritmo che utilizzando l'eliminazione di Gauss calcola l'inversa delle matrici invertibili (vedi ad esempio ABATE - de FABRITIIS pgg. 137-138)

8. Matrici di cambiamento di base.

In questo paragrafo ci proponiamo di scoprire con che legge cambiano le coordinate un vettore quando si considerano basi diverse. Precisamente sia V uno spazio vettoriale e siano

$$\mathcal{B} = \{v_1, \dots, v_n\} \qquad \mathcal{B}' = \{v'_1, \dots, v'_n\}$$

due basi per V . Vogliamo risolvere il seguente

Problema: Al variare di $v \in V$, se $x \in \mathbb{K}^n$ è il vettore delle coordinate di v alla base \mathcal{B} e $x' \in \mathbb{K}^n$ è il vettore delle coordinate di v alla base \mathcal{B}' , esprimere x in funzione di x' .

Procediamo nel modo seguente. Per ogni $j = 1, \dots, n$, esistono, e sono unici, scalari b_{1j}, \dots, b_{nj} tali che

$$v'_j = b_{1j}v_1 + \dots + b_{nj}v_n = \sum_{i=1}^n b_{ij}v_i.$$

Dunque per $v \in V$, si deve avere:

$$\sum_{i=1}^n x_i v_i = v = \sum_{j=1}^n x'_j v'_j = \sum_{j=1}^n x'_j \left(\sum_{i=1}^n b_{ij} v_i \right) = \sum_{i=1}^n \left(\sum_{j=1}^n b_{ij} x'_j \right) v_i.$$

Per l'unicità delle coordinate, allora si deve avere:

$$x_i = \sum_{j=1}^n b_{ij} x'_j.$$

Dunque abbiamo dimostrato la seguente

Teorema 8.1: Sia

$$B = \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \dots & b_{nn} \end{pmatrix} \qquad (8.1)$$

la matrice che ha per colonne le coordinate dei vettori della base \mathcal{B}' relative alla base \mathcal{B} . Allora per ogni vettore $v \in V$, se $x \in \mathbb{R}^n$ è il vettore delle coordinate di v alla base \mathcal{B} e $x' \in \mathbb{R}^n$ è il vettore delle coordinate di v alla base \mathcal{B}' , allora

$$x = Bx'.$$

La matrice B in (8.1) si dice *matrice del cambio di base* dalla base \mathcal{B} alla base \mathcal{B}' (in questo ordine!).

Osservazione. La matrice B è invertibile dato che il sistema $Bx' = x$ ha soluzione unica per ogni colonna x . La sua inversa B^{-1} è esattamente la matrice del cambio di base dalla base \mathcal{B}' alla base \mathcal{B} , ossia quella che ha per colonne le coordinate dei vettori della base \mathcal{B} relative alla base \mathcal{B}' .

Osservazione. Le basi di uno spazio vettoriale V di dimensione n su \mathbb{K} sono in corrispondenza biunivoca con il gruppo delle matrici invertibili $GL_n(\mathbb{K})$. Infatti fissata una base $\mathcal{B} = \{v_1, \dots, v_n\}$ di V , ogni altra base $\mathcal{B}' = \{v'_1, \dots, v'_n\}$ si ottiene scegliendo una

matrice $B = (b_{ij}) \in GL_n(\mathbb{K})$ e definendo $v'_j = b_{1j}v_1 + \dots + b_{nj}v_n$ per ogni $j = 1, \dots, n$. Moltiplicando formalmente righe di vettori per colonne di scalari questo equivale a dire che

$$(v'_1, \dots, v'_n) = (v_1, \dots, v_n)B. \quad (8.2)$$

È facile convincersi che $\mathcal{B}' = \{v'_1, \dots, v'_n\}$ è una base. Basta dimostrare che v'_1, \dots, v'_n sono linearmente indipendenti. A tal fine sia $\alpha = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$ tale che

$$O_V = \alpha_1 v'_1 + \dots + \alpha_n v'_n = (v'_1, \dots, v'_n) \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = (v_1, \dots, v_n)B\alpha.$$

Dato che v_1, \dots, v_n sono linearmente indipendenti, da $(v_1, \dots, v_n)B\alpha = O_V$ segue che necessariamente deve essere $B\alpha = 0_n$. Visto che $B = (b_{ij}) \in GL_n(\mathbb{K})$, questo succede se e solo se $\alpha = 0$ ossia se e solo se $\alpha_j = 0$ per ogni j . Viceversa quanto fatto sopra per discutere la matrice del cambio di base prove che se $\mathcal{B}' = \{v'_1, \dots, v'_n\}$ è una base di V allora vale la (8.2) per un'opportuna $B = (b_{ij}) \in GL_n(\mathbb{K})$.

In particolare se $V = \mathbb{K}^n$, questo argomento, scegliendo come base \mathcal{B} la base canonica, implica che ogni altra base di \mathbb{K}^n è formata dalle colonne di una matrice invertibile ossia le colonne A^1, \dots, A^n formano una base di \mathbb{K}^n se e solo se la matrice $A = (A^1, \dots, A^n)$ è invertibile.

Osservazione. Un modo alternativo e elegante per dimostrare la Teorema 8.1 è il seguente. Fissate le basi $\mathcal{B} = \{v_1, \dots, v_n\}$ e $\mathcal{B}' = \{v'_1, \dots, v'_n\}$ dello spazio vettoriale V , siano

$$F_{\mathcal{B}}: V \rightarrow \mathbb{K}^n \quad F_{\mathcal{B}'}: V \rightarrow \mathbb{K}^n$$

le applicazioni lineari definite nel modo seguente. Per ogni $v \in V$, se $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ sono le

coordinate di v rispetto a \mathcal{B} e $x' = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$ sono le coordinate di v rispetto a \mathcal{B}' , ossia se

$$\sum_{i=1}^n x_i v_i = v = \sum_{j=1}^n x'_j v'_j$$

allora

$$F_{\mathcal{B}}(v) = x \quad F_{\mathcal{B}'}(v) = x'.$$

Allora le coordinate x e x' di v relative a \mathcal{B} e \mathcal{B}' rispettivamente, si ha

$$x = F_{\mathcal{B}}(v) = F_{\mathcal{B}}(F_{\mathcal{B}'}^{-1}(x')) = F_{\mathcal{B}} \circ F_{\mathcal{B}'}^{-1}(x').$$

D'altra parte $F_{\mathcal{B}} \circ F_{\mathcal{B}'}^{-1}: \mathbb{K}^n \rightarrow \mathbb{K}^n$ è un isomorfismo di \mathbb{K}^n in sé e quindi per qualche matrice invertibile B si deve avere $F_{\mathcal{B}} \circ F_{\mathcal{B}'}^{-1} = L_B$: La matrice B ha colonne B^1, \dots, B^n che si

calcolano facilmente. Per ogni $j = 1, \dots, n$, se e_j è il j -esimo vettore della base canonica di \mathbb{K}^n , si ha:

$$B^j = F_{\mathcal{B}} \circ F_{\mathcal{B}'}^{-1}(e_j) = F_{\mathcal{B}}(F_{\mathcal{B}'}^{-1}(e_j)) = F_{\mathcal{B}}(v'_j)$$

e $F_{\mathcal{B}}(v'_j)$ è la colonna delle coordinate rispetto a \mathcal{B} del il j -esimo vettore della base \mathcal{B}' .

Le idee introdotte in questo paragrafo si usano anche per stabilire le relazioni fra le coordinate di punti dello spazio o del piano rispetto a diversi sistemi di riferimento. Per semplicità di notazioni descriviamo la situazione nel piano. Un *riferimento affine* del piano è una terna $\mathcal{R} = RA(O, \vec{\mathbf{i}}, \vec{\mathbf{j}})$ dove O è un punto del piano, e $\{\vec{\mathbf{i}} = \overrightarrow{OA}, \vec{\mathbf{j}} = \overrightarrow{OB}\}$ è una base dello spazio \mathcal{V}_O dei vettori applicati in O . Assegnare un riferimento affine equivale dunque ad assegnare tre punti O, A, B non allineati del piano. Fissato un riferimento affine \mathcal{R} , le coordinate $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ di un punto P del piano sono determinate dalla relazione

$$\overrightarrow{OP} = x_1 \vec{\mathbf{i}} + x_2 \vec{\mathbf{j}}. \quad (8.3)$$

Sia $\mathcal{R}' = RA(O', \vec{\mathbf{i}}', \vec{\mathbf{j}}')$ un altro riferimento affine. Se rispetto a \mathcal{R}' il punto P del piano ha coordinate $x' = \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix}$, ci proponiamo di trovare la relazione che lega x a x' .

Se $\vec{\mathbf{i}}' = \overrightarrow{O'A'}$ e $\vec{\mathbf{j}}' = \overrightarrow{O'B'}$, siano

$$\vec{\mathbf{i}}'' = \overrightarrow{OA''} = \overrightarrow{OA'} - \overrightarrow{OO'} \quad \text{e} \quad \vec{\mathbf{j}}'' = \overrightarrow{OB''} = \overrightarrow{OB'} - \overrightarrow{OO'}. \quad (8.4)$$

Dato che $\{\vec{\mathbf{i}}', \vec{\mathbf{j}}'\}$ è una base dello spazio $\mathcal{V}_{O'}$ dei vettori applicati in O' , allora i vettori $\vec{\mathbf{i}}'', \vec{\mathbf{j}}''$, ottenuti per traslazione mediante il vettore $-\overrightarrow{OO'}$, sono paralleli ai vettori $\vec{\mathbf{i}}', \vec{\mathbf{j}}'$ e quindi costituiscono una base \mathcal{V}_O dei vettori applicati in O . Sia B la matrice del cambio di base da $\{\vec{\mathbf{i}}, \vec{\mathbf{j}}\}$ a $\{\vec{\mathbf{i}}'', \vec{\mathbf{j}}''\}$. Dunque $B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$ è la matrice invertibile tale che

$$\vec{\mathbf{i}}'' = b_{11} \vec{\mathbf{i}} + b_{21} \vec{\mathbf{j}} \quad \text{e} \quad \vec{\mathbf{j}}'' = b_{12} \vec{\mathbf{i}} + b_{22} \vec{\mathbf{j}}. \quad (8.5)$$

Infine sia $c = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$ il vettore delle coordinate del punto O' rispetto al riferimento \mathcal{R} .

Teorema 8.2: Siano $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ e $x' = \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix}$ i vettori delle coordinate di un punto P del piano relative rispettivamente ai riferimenti \mathcal{R} e \mathcal{R}' . Allora

$$x = Bx' + c. \quad (8.6)$$

Viceversa se $B \in GL_2(\mathbb{R})$ è una matrice invertibile e $c \in \mathbb{R}^2$ allora $\mathcal{R}' = RA(O', \vec{\mathbf{i}}', \vec{\mathbf{j}}')$ dove O' è il punto di coordinate c relative a \mathcal{R} , e $\vec{\mathbf{i}}', \vec{\mathbf{j}}'$ sono date da (8.4) e (8.5), è un riferimento affine con coordinate x' date dalla (8.6).

Dimostrazione: Ricordiamo che, per definizione, le coordinate x di P relative al riferimento \mathcal{R} sono le coordinate del vettore \overrightarrow{OP} relative alla base $\{\vec{\mathbf{i}}, \vec{\mathbf{j}}\}$ di \mathcal{V}_O , mentre le coordinate x' relative al riferimento \mathcal{R}' sono le coordinate del vettore $\overrightarrow{O'P'}$ relative alla base $\{\vec{\mathbf{i}}', \vec{\mathbf{j}}'\}$ di $\mathcal{V}_{O'}$.

Per costruzione, le coordinate del vettore $\overrightarrow{O'P}$ relative alla base $\{\vec{i}', \vec{j}'\}$ di $\mathcal{V}_{O'}$ sono uguali alle coordinate del vettore $\overrightarrow{OP''} = \overrightarrow{OP} - \overrightarrow{OO'}$ $\in \mathcal{V}_O$ relative alla base $\{\vec{i}'', \vec{j}''\}$. Dunque

$$\begin{aligned} x_1 \vec{i}' + x_2 \vec{j}' &= \overrightarrow{OP} = \overrightarrow{OP''} + \overrightarrow{OO'} = x'_1 \vec{i}'' + x'_2 \vec{j}'' + c_1 \vec{i} + c_2 \vec{j} \\ &= x'_1 (b_{11} \vec{i} + b_{21} \vec{j}) + x'_2 (b_{12} \vec{i} + b_{22} \vec{j}) + c_1 \vec{i} + c_2 \vec{j} \\ &= (b_{11} x'_1 + b_{12} x'_2 + c_1) \vec{i} + (b_{21} x'_1 + b_{22} x'_2 + c_2) \vec{j} \end{aligned}$$

e quindi, per l'unicità delle coordinate,

$$\begin{cases} x_1 = b_{11} x'_1 + b_{12} x'_2 + c_1 \\ x_2 = b_{21} x'_1 + b_{22} x'_2 + c_2 \end{cases} \quad (8.7)$$

ossia vale la (8.6). Il resto dell'enunciato è immediato. \square

Con dimostrazione simile, il Teorema 8.2 si estende al caso dello spazio affine e si ha:

Teorema 8.3: Siano $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ e $x' = \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix}$ i vettori delle coordinate di un punto P dello spazio relative rispettivamente ai riferimenti $\mathcal{R} = RA(O, \vec{i}, \vec{j}, \vec{k})$ e $\mathcal{R}' = RA(O', \vec{i}', \vec{j}', \vec{k}')$. Allora

$$x = Bx' + c \quad (8.8)$$

dove $B \in GL_3(\mathbb{R})$ è la matrice del cambio di base dalla base $\{\vec{i}, \vec{j}, \vec{k}\}$ alla base $\{\vec{i}'', \vec{j}'', \vec{k}''\}$ di \mathcal{V}_O dove se $\vec{i}'' = \overrightarrow{OA'}$, $\vec{j}'' = \overrightarrow{OB'}$ e $\vec{k}'' = \overrightarrow{OC'}$,

$$\vec{i}'' = \overrightarrow{OA'} - \overrightarrow{OO'} \quad e \quad \vec{j}'' = \overrightarrow{OB'} - \overrightarrow{OO'} \quad \vec{k}'' = \overrightarrow{OC'} - \overrightarrow{OO'} \quad (8.9)$$

e c è la colonna delle coordinate del vettore $\overrightarrow{OO'}$ rispetto alla base $\{\vec{i}, \vec{j}, \vec{k}\}$.

9. Applicazioni lineari e matrici.

In questa sezione illustreremo come studiare applicazioni lineari fra spazi vettoriali utilizzando matrici.

Siano V e W spazi vettoriali finitamente generati e sia $T: V \rightarrow W$ un'applicazione lineare. Siano fissate una base $\mathcal{B} = \{v_1, \dots, v_n\}$ per V e una base $\mathcal{C} = \{w_1, \dots, w_m\}$ per W . Vogliamo trovare il modo di calcolare le coordinate rispetto alla base \mathcal{C} dell'immagine $T(v)$ mediante T di un vettore $v \in V$ in funzione delle coordinate di v relative alla base \mathcal{B} .

Per ogni $j = 1, \dots, n$ si ha per $v_j \in \mathcal{B}$

$$T(v_j) = a_{1j} w_1 + \dots + a_{mj} w_m \quad (9.1)$$

dove a_{1j}, \dots, a_{mj} sono le coordinate di $T(v_j)$ relative alla base \mathcal{C} . Chiamiamo la matrice

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \quad (9.2)$$

che ha per colonne le coordinate relative alla base \mathcal{C} di W delle immagini dei vettori della base \mathcal{B} di V la *matrice associata a T relativa alle basi \mathcal{B} di V e \mathcal{C} di W* . Il risultato che cerchiamo è il seguente:

Teorema 9.1: Per ogni $v \in V$ se $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ è la colonna delle coordinate di v relative

alla base \mathcal{B} di V e $y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$ è la colonna delle coordinate di $T(v)$ relative alla base \mathcal{C} di W , allora si ha:

$$y = Ax \quad (9.3)$$

dove A è la matrice associata a T definita in (9.2).

Dimostrazione: Utilizzando la linearità di T e (9.1), abbiamo

$$\begin{aligned} y_1 w_1 + \dots + y_m w_m &= T(v) = T(x_1 v_1 + \dots + x_n v_n) \\ &= x_1 T(v_1) + \dots + x_n T(v_n) \\ &= x_1 (a_{11} w_1 + \dots + a_{1m} w_m) + \dots + x_n (a_{n1} w_1 + \dots + a_{nm} w_m) \\ &= (a_{11} x_1 + \dots + a_{1n} x_n) w_1 + \dots + (a_{m1} x_1 + \dots + a_{mn} x_n) w_m \end{aligned}$$

e quindi per ogni $i = 1, \dots, m$, per l'unicità delle coordinate, si ha

$$y_i = a_{i1} x_1 + \dots + a_{in} x_n$$

che equivale a (9.3). □

Osservazione. Come nel caso di Teorema 8.1, vi è una dimostrazione alternativa del Teorema 9.1. Fissate le basi $\mathcal{B} = \{v_1, \dots, v_n\}$ di V e $\mathcal{C} = \{w_1, \dots, w_m\}$ dello spazio vettoriale V , siano

$$F_{\mathcal{B}}: V \rightarrow \mathbb{K}^n \quad F_{\mathcal{C}}: W \rightarrow \mathbb{K}^m$$

gli isomorfismi definiti da

$$F_{\mathcal{B}}\left(\sum_{j=1}^n x_j v_j\right) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad F_{\mathcal{C}}\left(\sum_{i=1}^m y_i w_i\right) = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}. \quad (9.4)$$

Se $v \in V$ ha coordinate $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ rispetto a \mathcal{B} e $y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$ sono le coordinate di

$T(v)$ rispetto a \mathcal{C} , allora

$$F_{\mathcal{B}}(v) = x \quad F_{\mathcal{C}}(T(v)) = y$$

e quindi si ha

$$y = F_{\mathcal{C}}(T(v)) = F_{\mathcal{C}}(T(F_{\mathcal{B}}^{-1}(x))) = F_{\mathcal{C}} \circ T \circ F_{\mathcal{B}}^{-1}(x).$$

D'altra parte $F_{\mathcal{C}} \circ T \circ F_{\mathcal{B}}^{-1}: \mathbb{K}^n \rightarrow \mathbb{K}^m$ è un'applicazione lineare di \mathbb{K}^n in \mathbb{K}^m e quindi per qualche matrice $A \in M_{n,m}(\mathbb{K})$ si deve avere $F_{\mathcal{C}} \circ T \circ F_{\mathcal{B}}^{-1} = L_A$. Le colonne A^1, \dots, A^n della matrice A sono le immagini dei vettori della base canonica e quindi, per $j = 1, \dots, n$, si calcolano nel modo seguente:

$$A^j = F_{\mathcal{C}} \circ T \circ F_{\mathcal{B}}^{-1}(e_j) = F_{\mathcal{C}}(T(F_{\mathcal{B}}^{-1}(e_j))) = F_{\mathcal{C}}(T(v_j))$$

e $F_{\mathcal{C}}(T(v_j))$ è la colonna delle coordinate rispetto a \mathcal{C} del j -esimo vettore della base \mathcal{B} e quindi $y = Ax$ dove A è la matrice definita in (9.2) come si voleva dimostrare.

I problemi incontrati nello studio di applicazioni lineari si possono tradurre in questioni analoghe per la matrice associata come è illustrato dalla seguente

Proposizione 9.2: Siano V e W spazi vettoriali finitamente generati e $T:V \rightarrow W$ un'applicazione lineare. Siano fissate una base $\mathcal{B} = \{v_1, \dots, v_n\}$ di V e una base $\mathcal{C} = \{w_1, \dots, w_m\}$ di W e siano $F_{\mathcal{B}}$ e $F_{\mathcal{C}}$ gli isomorfismi definiti in (9.4). Sia A la matrice associata a T relativa alle basi \mathcal{B} e \mathcal{C} .

- (i) Se $\{\eta_1, \dots, \eta_r\}$ è una base di $Im(A)$, allora $\{F_{\mathcal{C}}^{-1}(\eta_1), \dots, F_{\mathcal{C}}^{-1}(\eta_r)\}$ è una base di $Im(T)$;
- (ii) Se $\{\epsilon_1, \dots, \epsilon_k\}$ è una base di $Ker(A)$, allora $\{F_{\mathcal{B}}^{-1}(\epsilon_1), \dots, F_{\mathcal{B}}^{-1}(\epsilon_k)\}$ è una base di $Ker(T)$.

Dimostrazione: (i): Per $w \in W$ si ha $w \in Im(T) \Leftrightarrow w = T(v)$ per qualche $v \in V$. Se $y = F_{\mathcal{C}}(w)$ allora

$$w \in Im(T) \Leftrightarrow y = F_{\mathcal{C}}(w) = F_{\mathcal{C}}(T(v)) = F_{\mathcal{C}}(T(F_{\mathcal{B}}^{-1}(x))) = L_A(x) = Ax \in Im(A)$$

per qualche $x \in \mathbb{K}^n$. Dunque $F_{\mathcal{C}}(Im(T)) = Im(A)$ e quindi la restrizione dell'isomorfismo $F_{\mathcal{C}}$ a $Im(T)$ è un isomorfismo di $Im(T)$ in $Im(A)$. Il suo inverso $F_{\mathcal{C}}^{-1}$ è anche un isomorfismo e quindi trasforma basi di $Im(A)$ in basi di $Im(T)$.

(ii): Per $v \in V$ si ha $v \in Ker(T) \Leftrightarrow T(v) = 0_W$. Se $x = F_{\mathcal{B}}(v)$ allora

$$v \in Ker(T) \Leftrightarrow O_m = F_{\mathcal{C}}(O_W) = F_{\mathcal{C}}(T(v)) = F_{\mathcal{C}}(T(F_{\mathcal{B}}^{-1}(x))) = L_A(x) = Ax.$$

Segue che $F_{\mathcal{B}}(Ker(T)) = Ker(A)$ e quindi che la restrizione dell'isomorfismo $F_{\mathcal{B}}$ a $Ker(T)$ è un isomorfismo di $Ker(T)$ in $Ker(A)$. Il suo inverso $F_{\mathcal{B}}^{-1}$ è anche un isomorfismo e quindi trasforma basi di $Ker(A)$ in basi di $Ker(T)$. \square

Il Teorema 9.1 permette di dimostrare in generale che lo spazio vettoriale delle applicazioni lineari fra spazi vettoriali finitamente generati è isomorfo a uno spazio di matrici:

Teorema 9.3: Siano V e W spazi vettoriali sul campo \mathbb{K} con $dim(V) = n$ e $dim(W) = m$. Allora lo spazio $\mathcal{L}(V, W)$ delle applicazioni lineari da V a W è isomorfo a $M_{m,n}(\mathbb{K})$ e quindi, in particolare ha dimensione mn .

Dimostrazione: Si fissino basi $\mathcal{B} = \{v_1, \dots, v_n\}$ di V e $\mathcal{C} = \{w_1, \dots, w_m\}$ di W . Per ogni $T \in \mathcal{L}(V, W)$ sia A_T la matrice associata a T relativa alle basi \mathcal{B} e \mathcal{C} . Si consideri l'applicazione $\Phi: \mathcal{L}(V, W) \rightarrow M_{m,n}(\mathbb{K})$ definita da $\Phi(T) = A_T$. Dato che si vede immediatamente che per $S, T \in \mathcal{L}(V, W)$ si ha $A_{S+T} = A_S + A_T$ e che per $\lambda \in \mathbb{K}$ si ha $A_{\lambda T} = \lambda A_T$, segue che Φ è lineare. Inoltre Φ è iniettiva. Infatti due applicazioni lineari $S, T \in \mathcal{L}(V, W)$ hanno la stessa matrice A relativa alle basi \mathcal{B} e \mathcal{C} , allora coincidono su \mathcal{B} e quindi sono uguali. Infine Φ è anche suriettiva. Infatti se $A = (a_{ij}) \in M_{m,n}(\mathbb{K})$, allora sia $T \in \mathcal{L}(V, W)$ l'unica applicazione lineare tale che per ogni vettore $v_j \in \mathcal{B}$ si abbia $T(v_j) = a_{1j}w_1 + \dots + a_{mj}w_m$. Allora $\Phi(T) = A$. Dunque Φ è l'isomorfismo cercato. \square

Osservazione. Si osservi che l'isomorfismo costruito in Teorema 9.3 dipende dalla scelta fatta delle basi di V e W .

Concludiamo il paragrafo stabilendo come varia la matrice associata a un'applicazione lineare cambiando basi:

Proposizione 9.4: Siano V e W spazi vettoriali finitamente generati e sia $T:V \rightarrow W$ un'applicazione lineare. Siano $\mathcal{B}, \mathcal{B}'$ basi di V e $\mathcal{C}, \mathcal{C}'$ basi di W . Se A e A' sono le matrici

associate a T rispettivamente a \mathcal{B}, \mathcal{C} e $\mathcal{B}', \mathcal{C}'$, B è la matrice del cambio di base da \mathcal{B} a \mathcal{B}' e C è la matrice del cambio di base da \mathcal{C} a \mathcal{C}' , allora

$$A' = C^{-1}AB. \quad (9.5)$$

Dimostrazione: Sia $v \in V$ arbitrario. Siano x e x' le coordinate di v rispettivamente relative a \mathcal{B} e \mathcal{B}' e siano y e y' le coordinate di $F(v)$ rispettivamente relative a \mathcal{C} e \mathcal{C}' . Allora le matrici A e A' sono le uniche matrici tali che $y = Ax$ e $y' = A'x'$. Dato che $x = Bx'$ e $y = Cy'$, allora abbiamo

$$A'x' = y' = C^{-1}y = C^{-1}Ax = C^{-1}ABx'. \quad (9.6)$$

Per l'arbitrarietà di v , la (9.6) vale per ogni $x' \in \mathbb{K}^n$ e quindi $A' = C^{-1}AB$. \square

La formula (9.5) prende una forma particolarmente semplice nel caso in cui $V = W$, $\mathcal{B} = \mathcal{C}$ e $\mathcal{B}' = \mathcal{C}'$, ossia nel caso si considerino endomorfismi di uno spazio vettoriale e stesse basi nel dominio e nel codominio. Se A è la matrice di $T: V \rightarrow V$ relativa a \mathcal{B} (nel dominio e nel codominio) e A' è la matrice di T relativa a \mathcal{B}' (nel dominio e nel codominio), allora

$$A' = B^{-1}AB. \quad (9.7)$$

Definizione. Due matrici quadrate A, A' di ordine n si dicono *simili* se esiste una matrice $B \in GL_n(\mathbb{K})$ tale che valga (9.7).

Si ha immediatamente la seguente

Proposizione 9.5: *Due matrici quadrate A, A' di ordine n sono le matrici associate a un endomorfismo di uno spazio vettoriale di dimensione n se e solo se sono simili.*