Spazi Vettoriali e Applicazioni lineari.

Appunti dalle lezioni di Geometria - Corso di laurea in Fisica

Giorgio PATRIZIO - A.A. 2006/07

1. Spazi vettoriali: definizioni e esempi.

Denoteremo con $\mathbb{K}$ 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

<u>*somma*</u> $\begin{array}{rccc} +\colon & V\times V & \to & V \\ & (v,w) & \mapsto & v+w \end{array}$ *e* <u>*prodotto per uno scalare*</u> $\begin{array}{rccc} p\colon & \mathbb{K}\times V & \to & V \\ & (\lambda,w) & \mapsto & \lambda w \end{array}$

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 $\Longrightarrow$. 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:

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

e

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

è uno spazio vettroriale 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 vettroriale 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 $f\colon A \to \mathbb{K}$ definite su A a valori nel campo $\mathbb{K}$. Se $f, g \in W$ e $\lambda \in \mathbb{K}$, allora le funzioni $f + g\colon A \to \mathbb{K}$ e $\lambda f\colon A \to \mathbb{K}$ sono le funzioni definite rispettivamente da:

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

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 $f\colon A \to V$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 $F + G\colon A \to V$ e $\lambda F\colon A \to V$ sono le applicazioni definite rispettivamente da:

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

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 $a_{m+1} = \ldots = a_n = 0$; 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. 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

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

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

$$A(u+z) = ((u^1+z^1)A^1, \dots, (u^n+z^n)A^n) = (u^1A^1, \dots, u^nA^n)+(z^1A^1, \dots, z^nA^n) = Au+Az = 0.$$

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

$$0 = Aw-Av = (w^1A^1, \dots, w^nA^n)-(v^1A^1, \dots, v^nA^n) = ((w^1-v^1)A^1, \dots, (w^n-v^n)A^n) = A(w-v)$$

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

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

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 lineare dei vettori $v_1, \dots, v_n$ e si denota con

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

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:

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

Osservazione. $Span(S)$ è un sottospazio vettoriale di V.

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

Esempio 5. Gli spazi $\mathbb{K}^n$ e $M_{m,n}(\mathbb{K})$ sono finitamente generati.

Esempio 6. 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.

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 $A = (A^1, \ldots, A^n)$ è l'espressione di A secondo le sue colonne, il sistema lineare $Ax = b$ è compatibile (ossia ammette soluzione) se e solo se $b \in Span(A^1, \ldots, A^n)$.*

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 $b = x_1 A^1 + \ldots + x_n A^n$ e quindi se e solo se $b \in Span(A^1, \ldots, A^n)$. $\square$

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

$$\alpha_1 v_1 + \ldots + \alpha_n v_n = O.$$

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

$$\alpha_1 v_1 + \ldots + \alpha_n v_n = O \Longrightarrow \alpha_1 = \ldots = \alpha_n = 0.$$

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

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

Dimostrazione: Se $v_1, \ldots, v_n$ sono linearmente dipendenti, allora esistono $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ non tutti nulli tali che $\alpha_1 v_1 + \ldots + \alpha_n v_n = O$. A meno di riordinare i vettori, si può supporre che $\alpha_1 \neq 0$ e quindi $v_1 = -\frac{\alpha_2}{\alpha_1} v_2 \ldots - \frac{\alpha_n}{\alpha_1} v_n$. D'altra parte se un vettore, ad esempio v_1, è combinazione lineare degli altri, allora $v_1 = \beta_2 v_2 + \ldots + \beta_n v_n$ e quindi $v_1 - \beta_2 v_2 - \ldots - \beta_n v_n = 0$ da cui segue che $v_1, \ldots, v_n$ sono linearmente dipendenti. $\square$

Una 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 $A^1, \ldots, A^n$ 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 $x_1 A^1 + \ldots + x_n A^n = 0$, solo nel caso che $x_1 = \ldots = x_n = 0$. Per definizione di lineare indipendenza abbiamo dunque la tesi. $\square$

3. Basi di uno spazio vettoriale.

Definizione. Sia V uno spazio vettoriale su $\mathbb{K}$. 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. In particolare, nel caso in cui $\mathcal{B} = \{v_1, \ldots, v_n\}$ sia finito, allora $\mathcal{B}$ è una base se e solo se $v_1, \ldots, v_n$ sono linearmente indipendenti e $V = Span(v_1, \ldots, v_n)$.

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

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

allora $\alpha_j = \alpha_j$ per ogni $j = 1, \ldots, n$. In particolare, allora, se $\mathcal{B} = \{v_1, \ldots, v_n\}$ 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 $\alpha_1 v_1 + \ldots + \alpha_n v_n = \beta_1 v_1 + \ldots + \beta_n v_n$ allora $(\alpha_1 - \beta_1)v_1 + \ldots + (\alpha_n - \beta_n)v_n = O$, e quindi, dato che $v_1, \ldots, v_n$ sono vettori linearmente indipendenti, la tesi segue. □

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

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

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

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

Se $v \in V = Span(v_1, \ldots, v_n)$, 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 $v \in V = Span(w_1, \ldots, w_k)$. □

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 $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base di V e, per assurdo sia $\{w_1, \ldots, w_{n+1}\} \subset V$ un sistema libero. Dato che $\mathcal{B}$ è una base, si deve avere $w_1 = t_1 v_1 + \ldots + t_n v_n$, per opportuni scalari $t_1, \ldots, t_n$ non tutti nulli. A meno di riordinare $\mathcal{B}$, possiamo supporre che $t_1 \neq 0$ e quindi si ha

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

Segue allora che $V = Span(w_1, v_2, \ldots, v_n)$. Ma allora $w_2 = s_1 w_1 + s_2 v_2 + \ldots + s_n v_n$, per opportuni scalari $s_1, \ldots, s_n$. Si osservi che $s_2, \ldots, s_n$ non sono tutti nulli perché w_1, w_2 sono linearmente indipendenti. Di nuovo a meno di riordinare $v_2, \ldots, v_n$, possiamo supporre $s_2 \neq 0$. Ma allora

$$v_2 = -\frac{s_1}{s_2} w_1 + \frac{1}{s_2} w_2 - \ldots - \frac{s_n}{s_2} v_n$$

e quindi $V = Span(w_1, w_2, v_3, \ldots, v_n)$. Ripetendo l'argomento, per induzione, possiamo concludere che $V = Span(w_1, \ldots, w_n)$. Ma allora $w_{n+1} \in Span(w_1, \ldots, w_n)$, contro l'ipotesi che $\{w_1, \ldots, w_{n+1}\}$ fosse libero. □

Teorema 3.4: *Ogni spazio vettoriale $V \neq \{O\}$ ha una base e se V è finitamente generato allora ha base finita. Due basi dello stesso spazio vettoriale hanno la stessa cardinalità.*

Dimostrazione: Diamo la dimostrazione solo nel caso in cui V sia finitamente generato. Supponiamo che $V = Span(v_1, \ldots, v_n)$. Sia k il massimo numero di vettori linearmente indipendenti in $\{v_1, \ldots, v_n\}$. A meno di riordinare l'insieme, si può supporre dunque che $v_1, \ldots, v_k$ siano linearmente indipendenti e che per ogni $j = k+1, \ldots, n$ i vettori $v_1, \ldots, v_k, v_j$ siano linearmente dipendenti e quindi $v_j \in Span(v_1, \ldots, v_k)$. 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 $w_1, \ldots, w_{k+1} \in \mathcal{C}$, allora $\{w_1, \ldots, w_{k+1}\}$ 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 3.3 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 $\{O\} \neq V = Span(v_1, \ldots, v_n)$ allora almeno uno fra i $v_1, \ldots, v_n$ è diverso dal vettore nullo. A meno di riordinare i vettori $v_1, \ldots, v_n$, si può supporre $v_1 \neq O$. Si presentano ora due possibilità mutualmente esclusive. La prima è che, per ogni indice $j = 2, \ldots, n$, l'insieme $\{v_1, v_j\}$ non è libero. In questo caso $V = Span(v_1)$ e quindi $\{v_1\}$ è una base di V. La seconda è che, per qualche $j = 2, \ldots, n$, 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 $v_2, \ldots, v_n$, 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 $j = 3, \ldots, n$, l'insieme $\{v_1, v_2, v_j\}$ non è libero. In questo caso $V = Span(v_1, v_2)$ e quindi $\{v_1, v_2\}$ è una base di V. La seconda è che, per qualche $j = 3, \ldots, n$, 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 $v_3, \ldots, v_n$, 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 $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base di uno spazio vettoriale V e, per $1 \leq p \leq n$, siano $w_1, \ldots, w_p$ siano vettori linearmente indipendenti di V. Esistono allora $n - p$ vettori di $\mathcal{B}$ che insieme ai $w_1, \ldots, w_p$ formano una 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 $w_1 = t_1 v_1 + \ldots + t_n v_n$ con i coefficienti t_j non tutti nulli. A meno di riordinare $\mathcal{B}$ si può supporre $t_1 \neq 0$. Allora

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

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

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

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

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

Visto che $t_1 \neq 0$ allora segue $s_1 = 0$ e, di conseguenza, $s_2 = \ldots = s_n = 0$. Segue che $\{w_1, \ldots, v_n\}$ 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 $\mathcal{B}$, si abbia che $\{w_1, \ldots, w_{k-1}, v_k, \ldots v_n\}$ è una base di V. Allora $w_k = t_1 w_1 + \ldots + t_{k-1} w_{k-1} + t_k v_k + \ldots + t_n v_n$ con i coefficienti $t_k, \ldots, t_n$ non sono tutti nulli perché $w_1, \ldots, w_k$ sono linearmente indipendenti. A meno di riordinare $t_k, \ldots, v_n$, si può supporre $t_k \neq 0$. Allora

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

e quindi, ancora per il Lemma 3.2 si ha

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

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

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

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

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

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

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

□

Corollario 3.6: *Sia V uno spazio vettoriale di dimensione n e sia $\mathcal{B} = \{v_1, \ldots, v_n\} \subset V$.*
(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 $v \in Span(v_1, \ldots, v_n)$, 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 $\mathcal{B} \subset Span(\mathcal{B}')$. 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 $\mathcal{B} = \{w_1, \ldots, w_n\}$ 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 $V = Span(\mathcal{B}) = W$. $\square$

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 $U = Span\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) \subset \mathbb{R}^2$ e $W = Span\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) \subset \mathbb{R}^2$, 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 $U \subset W$ 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 $u = -w + u + w \in W$. □

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 $U = Span(\mathcal{B}$ e $W = Span(\mathcal{C}$ allora $U + W = Span(\mathcal{B} \cup \mathcal{C}$. 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. 69. □

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 esistono basi $\{u_1, \ldots, u_r\}$ di U e $\{w_1, \ldots, w_s\}$ di W tali che $\{u_1, \ldots, u_r, w_1, \ldots, w_s\}$ sia una base di V.*

Dimostrazione: Siano $\{u_1, \ldots, u_r\}$ e $\{w_1, \ldots, w_s\}$ basi rispettivamente di U di W tali che $\{u_1, \ldots, u_r, w_1, \ldots, w_s\}$ 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 $\{u_1, \ldots, u_r\}$ e $\{w_1, \ldots, w_s\}$ sono basi di U e di W rispettivamente, allora $\{u_1, \ldots, u_r, w_1, \ldots, w_s\}$ è una base di V perchè è un sistema di generatori con $r + s$ vettori e $dim(V) = r + s$. □

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 $\{u_1, \ldots, u_r\}$ di U e la si completi a una base $\{u_1, \ldots, u_r, w_1, \ldots, w_s\}$ una base di V. Allora $W = Span(w_1, \ldots, w_s)$ è un supplementare di U.

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 $L: V \to W$ si dice *applicazione lineare* se

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

Osservazione 1. Segue immediatamente dalla definizione che se $L: V \to W$ 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 $T: V \to W$, $S: W \to Z$ sono applicazioni lineari , la composizione $S \circ T: V \to Z$ 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 $f: V \to \mathbb{K}$ a valori nel campo degli scalari (considerato come spazio vattoriale) 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* $N: V \to W$ definita per ogni $w \in W$ da $N(w) = O_W$ è lineare.

Esempio 2. Per ogni spazio vettoriale V l'*applicazione identica* $Id_V: V \to V$ 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 $\tau_{OO'}: \mathcal{V}_O \to \mathcal{V}_{O'}$ è un'applicazione lineare. Si osservi che $\tau_{OO'}$ è biettiva e che l'inversa è la traslazione $\tau_{O'O}: \mathcal{V}_{O'} \to \mathcal{V}_O$, 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 $L_A\colon \mathbb{K}^n \to \mathbb{K}^m$ mediante $L_A(x) = Ax$. Siano $x, y \in \mathbb{K}^n$ e $\lambda \in \mathbb{K}$. Se $A = (A^1, \ldots, A^n)$ è l'espressione di A secondo le sue colonne, allora $L_A(x) = x^1A^1, \ldots, x^nA^n$ e quindi

$$\begin{aligned} L_A(x+y) &= A(x+y) = ((x^1+y^1)A^1, \ldots, (x^n+y^n)A^n) = (x^1A^1, \ldots, x^nA^n) + (y^1A^1, \ldots, y^nA^n) \\ &= Ax + Ay = L_A(x) + L_A(y) \end{aligned}$$

e

$$L_A(\lambda x)A(\lambda x) = ((\lambda x^1)A^1, \ldots, (\lambda x^n)A^n) = \lambda(x^1A^1, \ldots, x^nA^n) + (y^1A^1, \ldots, y^nA^n) = \lambda Ax = \lambda L_A(x)$$

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

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

Esempio 6. La funzione $v_0\colon \mathbb{K}[t] \to \mathbb{K}$, detta *valutazione in* 0, definita da $v_0(p(t) = p(0)$ è lineare. La derivata $D\left(\sum_{k=0}^{n} a_k t^k\right) = \sum_{k=1}^{n} k a_k t^{k-1}$ definisce un'applicazione lineare $D\colon \mathbb{K}[t] \to \mathbb{K}[t]$.

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 $\mathcal{B} = \{v_1, \ldots, v_n\}$ una sua base. Se W è uno spazio vettoriale su $\mathbb{K}$ e $w_1, \ldots, w_n \in W$ allora esiste una unica applicazione lineare $T\colon V \to W$ tale che $T(v_j) = w_j$ per ogni $j = 1, \ldots, n$. 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. 88. □

Corollario 5.2: *Siano V uno spazio vettoriale finitamente generato su $\mathbb{K}$, $\mathcal{B} = \{v_1, \ldots, v_n\}$ una sua base e W un altro spazio vettoriale su $\mathbb{K}$. Allora se $S, T\colon V \to W$ sono applicazioni lineari, si ha che $S = T$ (ossia che $S(v) = T(v)$ per ogni $v \in V$) se e solo se $T(v_j) = T(v_j)$ 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 dal Teorema 5.1 e dal fatto che $L_A(e_j) = A^j$ e $L_B(e_j) = B^j$ per ogni $j = 1, \ldots, n$. □

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

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

Dimostrazione: Immediata dal Teorema 5.1. □

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

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

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 $(\mathbb{K}^n)^*$ com 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 $F: A \to B$ 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 dicona 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 $F^{-1}: B \to A$ 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 $F: A \to B$ fra due insiemi. Allora*
(i) F è iniettiva se e solo se esiste un'applicazione $G: B \to A$ tale che $G \circ F = Id_A$;
(ii) F è suriettiva se e solo se esiste un'applicazione $H: B \to A$ tale che $F \circ H = Id_A$;
(iii) F è biettiva se e solo se esiste un'applicazione $L: B \to A$ tale che $L \circ F = Id_A$ e $F \circ L = Id_B$. Se esiste, $L = F^{-1}$.

Dimostrazione: (i) Se esiste $G: B \to A$ 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 $G: B \to A$ 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 $G: B \to A$ è tale che $G \circ F = Id_A$.

(ii) Se esiste $H: B \to A$ tale che $F \circ H = Id_A$, 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 $H: B \to A$ tale che $F \circ H = Id_A$.

(iii) Se esiste $L: B \to A$ 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 $L = F^{-1}$. $\square$

Torniamo ora alle applicazioni lineari. Sia $T: V \to W$ un'applicazione lineare. L'*immagine* $Im(T)$ e il *nucleo* $Im(T)$ di T sono definiti rispettivamente da:

$$Im(T) = \{w = T(v) \in B \mid v \in V\} \qquad \text{e} \qquad Ker(T) = \{v \in V \mid F(v) = O_W\}.$$

Esempio 7. Per una matrice $A \in M_{m,n}(\mathbb{K})$ si consideri l'applicazione $L_A: \mathbb{K}^n \to \mathbb{K}^m$. Allora, se $A^1, \ldots, A^n$ sono le colonne di A allora $Im(L_A) = Span(A^1, \ldots, A^n)$ 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 $T: V \to W$ 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 $Ker(T) = \{O_V\}$. Supponiamo, viceversa, che si abbia $Ker(T) = \{O_V\}$. 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 $v_1 - v_2 \in Ker(T) = \{O_V\}$. □

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

Proposizione 5.7: *Sia $T: V \to W$ un'applicazione lineare. Allora:*
(i) Sia $\{v_1, \ldots, v_n\}$ un sistema di generatori di V. Allora $\{T(v_1), \ldots, T(v_n)\}$ è un sistema di generatori di $Im(T)$. Dunque T è suriettiva se e solo se $\{T(v_1), \ldots, T(v_n)\}$ è un sistema di generatori di W;
(ii) Se T è iniettiva e $\{v_1, \ldots, v_n\}$ è un sistema libero di vettori di V allora $\{T(v_1), \ldots, T(v_n)\}$ è un sistema libero di vettori di W;
(iii) Se $\{v_1, \ldots, v_n\}$ è una base di V, T è biettiva se e solo se $\{T(v_1), \ldots, T(v_n)\}$ è una base di W.

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

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

(iii) Immediata da (i) e (ii). □

Un importante invariante delle applicazioni lineari è il rango:

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

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

$$dim(V) = rg(T) + dim(Ker(T)).$$

Dimostrazione: Vedi ABATE - de FABRITIIS pg. 91. □

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

$$n = rg(A) + dim(Ker(A)).$$

Corollario 5.9: *Sia $T: V \to W$ 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 soloè iniettiva se e solo è suriettiva.

Dimostrazione: Vedi ABATE- de FABRITIIS pg. 91. □

Teorema 5.10: (di Rouché-Capelli) *Un sistema lineare $Ax = b$ di m equazioni in n incognite ha soluzione se e solo se il rango della matrice $A = (A^1, \ldots, A^n)$ del sistema è uguale al rango della matrice completa $A = (A^1, \ldots, A^n, b)$ del sistema. Qualora esista soluzione, questa è unica se e solo se $rg(A) = n$.*

Dimostrazione: Vedi ABATE pg. 91. □

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

Proposizione 5.11: *Sia $T: V \to W$ un'applicazione lineare biettiva . Allora $T^{-1}: W \to V$ è un'applicazione lineare.*

Dimostrazione: Siano $w_1, w_2 \in W$. Allora esistono, e sono unici, $T^{-1}(w_1) = v_1$ 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:

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

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

Un'applicazione lineare biettiva $T: V \to W$ 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 genrati sono isomorfi, basta confrontare le loro dimensioni:

Proposizione 5.12: *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 $\{v_1, \ldots, v_n\}$ e $\{w_1, \ldots, w_n\}$ basi rispettivamente per V e W. Sia $T: V \to W$ l'unica applicazione lineare, esistente per il Teorema 5.1, tale che $T(v_j) = w_j$ per ogni $j = 1, \ldots, n$. L'applicazione T è definita da

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

Bisogna dimostrare che T è biettiva. Prima di tutto si osservi che T è suriettiva perché, per costruzione trasforma un sistema di generatori in un sistema di generatori. D'atra parte è anche iniettiva dato che se $v = x_1v_1 + \ldots + x_nv_n \in Ker(T)$, allora

$$O_W = T(v) = T\left(\sum_{j=1}^{n} x_jv_j\right) = \sum_{j=1}^{n} x_jw_j$$

e quindi $x_1 = \ldots = x_n = 0$ perché $\{w_1, \ldots, w_n\}$ è una base di W. □

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 $\mathcal{B} = \{v_1, \ldots, v_n\}$ di V esiste una base $\mathcal{B}^* = \{v_1^*, \ldots, v_n^*\}$ di V^*, chiamata* **base duale** *di $\mathcal{B}$, dove i funzionali lineari $v_1^*, \ldots, v_n^*$ 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 $t_1, \ldots, t_n$ scalari tali che

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

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

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

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 $j = 1, \ldots, n$ 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. □

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

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

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 $\mathcal{B}' = \{w_1, \ldots, w_k\}$ una base di W e sia $\mathcal{B} = \{w_1, \ldots, w_k, w_{k+1}, \ldots, n_n\}$ un suo completamento a una base di di V. Sia $\mathcal{B}^* = \{w_1^*, \ldots, w_k^*, w_{k+1}^* \ldots, w_n^*\}$ la base duale di $\mathcal{B}$ di V^*. Dimostriamo che $\mathcal{C}^* = \{w_{k+1}^* \ldots, w_n^*\}$ è una base di $dim(W^0)$, la tesi segue allora immediatamente. Per cominciare $\mathcal{C}^* \subset W^0$. Infatti per ogni $i = k+1, \ldots, n$ se $j = 1, \ldots, k$ 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 $i = 1, \ldots, k$ e quindi

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

□

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: $(Im(A))^0$ *è isomorfo a* $Ker(A^T)$

Dimostrazione: Sia $\{e_1, \ldots, e_m\}$ la base canonica di $\mathbb{K}^m$ e sia $= \{\epsilon_1, \ldots, \epsilon_n\}$ la base di $(\mathbb{K}^m)^*$ duale della base canonica. Allora l'applicazione lineare $T\colon \mathbb{K}^m \to (\mathbb{K}^m)^*$ unicamente definita ponendo $T(e_i) = \epsilon_i$ per ogni $i = 1, \ldots, n$ 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 $A = (a_{ij}) \in M_{m,n}(\mathbb{K})$. Cominciamo dimostrando che $T(Ker(A^T)) \subset (Im(A))^0$. Sia $z \in Ker(A^T)$, allora deve essere $A^T z = 0_n$ e quindi

$$0 = \sum_{i=1}^{m} a_{ij} z_i$$

per ogni $j = 1, \ldots, n$. Sia allora $f = T(z) = \sum_{i=1}^{m} z_i \epsilon_i$. Verifichiamo che $f \in (Im(A))^0$. A tal fine basta controllare che f sia nulla quando valutata sulle colonne di A che generano $Im(A)$. Sia dunque A^j una colonna di A, allora

$$f(A^j) = \sum_{i=1}^{m} z_i \epsilon_i(A^j) = \sum_{i=1}^{m} z_i a_{ij} = 0.$$

Per dimostrare che $(Im(A))^0 \subset T(Ker(A^T))$, basta ripercorrere a ritroso l'argomento. Sia $f = \sum_{i=1}^m f_i \epsilon_i \in (Im(A))^0$. Allora per ogni colonna A^j di A si deve avere

$$0 = f(A^j) = \sum_{i=1}^m f_i \epsilon_i(A^j) = \sum_{i=1}^m f_i a_{ij}$$

e quindi il vettore $z = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} \in Ker(A^T)$. Dato che evidentemente $T(z) = f$, segue $f \in T(Ker(A^T))$ $\square$

Possiamo ora provare :

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

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

$\square$

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

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

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 associati. Come al solito denotiamo con $e_1, \dots, e_n$ gli elementi della base canonica di $\mathbb{K}^n$:

Preposizione 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 $Span(e_1, \ldots, e_r)$ e quindi $Im(S) \subset Span(e_1, \ldots, e_r)$. Dato che $dim(Span(e_1, \ldots, e_r)) = r$, per provare (i), basterà verificare che le colonne $\{S^{j_1}, \ldots, S^{j_r}\}$ sono linearmente indipendenti. Sia

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

Allora il sistema $S'x = 0$ ha soluzione unica e questo equivale al fatto che le colonne $S^{j_1}, \ldots, S^{j_r}$ 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 $a_{r\,k+1}, \ldots, a_{m\,k+1}$ 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 che $a_{r\,k+1} \neq 0$. Se per ogni indice $r < i \leq m$ si somma alla riga i-esima di A la r-esima riga moltiplicata per $-\frac{a_{i\,k+1}}{a_{r\,k+1}}$, si ottiene una 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, di 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) $Ker(A) = Ker(S)$;
(iii) $rg(A) = rg(S)$;
(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 $Im(A)$.

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

$$rg(A) = n - dimKer(A) = n - dimKer(S) = rg(S).$$

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. □

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 $rg(A) = rg(S)$, in generale $Im(A) \neq Im(S)$!

7.2 Prodotto di matrici, invertibilità

Ricordiamo per cominciare la definizione di prodotto di matrici. Se $A = \begin{pmatrix} a_1 & \dots & a_n \end{pmatrix} \in M_{1,n}(\mathbb{K})$ 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

$$AB = \sum_{l=1}^{n} a_l b_l. \tag{7.2}$$

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

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

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 $L_C\colon \mathbb{K}^p \to \mathbb{K}^m$ è la composizione $L_A \circ L_B$ delle applicazioni lineari $L_B\colon \mathbb{K}^p \to \mathbb{K}^n$ e $L_A\colon \mathbb{K}^n \to \mathbb{K}^m$.*

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

$$L_A \circ L_B(e_j) = L_A(L_B(e_j)) = L_A(Be_j) = L_A(B^j) = AB^j = \begin{pmatrix} A_1B^j \\ \vdots \\ A_mB^j \end{pmatrix}.$$

$\square$

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

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

la matrice identica di ordine n. A volte si 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 Vedi ABATE - de FABRITIIS pg 123. $\square$

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: $AI_n = I_nA = A$. 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

$$AB = BA = I_n. \tag{7.4}$$

.

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) $L_A \colon \mathbb{K}^n \to \mathbb{K}^n$ è suriettiva;

(x) $L_A \colon \mathbb{K}^n \to \mathbb{K}^n$ è iniettiva;

(xi) $L_A \colon \mathbb{K}^n \to \mathbb{K}^n$ è biettiva.

Dimostrazione: $(i) \Rightarrow (ii)$: Se A è invertibile, er 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.

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

$(x) \Rightarrow (xi)$: $L_A \colon \mathbb{K}^n \to \mathbb{K}^n$ 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 $x \in \mathbb{K}^n$:

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

e quindi $L_B \circ L_A = Id_{K^n} = L_{I_n}$ e $L_A \circ L_B = Id_{K^n} = L_{I_n}$. 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 $L_A \circ L_B = Id_{K^n}$. 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 $L_A \circ L_B = Id_{K^n}$. 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.

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

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. 125-126)

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

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

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 $j = 1, \ldots, n$, esistono, e sono unici, scalari $b_{1j}, \ldots, b_{nj}$ tali che

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

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*

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

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 biuniuvoca 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 $j = 1, \dots, n$. In particolare se $V = \mathbb{K}^n$, questo argomento, scegliendo come base $\mathcal{B}$ la base canonica, dimostra 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

$$F_{\mathcal{B}}: V \to \mathbb{K}^n \qquad F_{\mathcal{B}'}: V \to \mathbb{K}^n$$

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

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

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 $F_{\mathcal{B}} \circ F_{\mathcal{B}'}^{-1} : \mathbb{K}^n \to \mathbb{K}^n$ 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

$$\overrightarrow{OP} = x_1 \vec{\mathbf{i}} + x_2 \vec{\mathbf{j}}. \tag{8.2}$$

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

$$\vec{\mathbf{i}}'' = \overrightarrow{OA''} = \overrightarrow{OA'} - \overrightarrow{OO'} \quad \text{e} \quad \vec{\mathbf{j}}'' = \overrightarrow{OB''} = \overrightarrow{OB'} - \overrightarrow{OO'}. \tag{8.3}$$

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

$$\vec{\mathbf{i}}'' = b_{11} \overrightarrow{\mathbf{i}} + b_{21} \overrightarrow{\mathbf{j}} \quad \text{e} \quad \vec{\mathbf{j}}'' = b_{12} \overrightarrow{\mathbf{i}} + b_{22} \overrightarrow{\mathbf{j}}. \tag{8.4}$$

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*

$$x = Bx' + c. \tag{8.5}$$

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 (8.3) e (8.4), è un riferimento affine con coordinate x' date dalla (8.5).

Dimostrazione: Ricordiamo che, per definizione, le coordinate x di P relative la 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 la riferimento $\mathcal{R}'$ sono le coordinate del vettore $\overrightarrow{OP'}$ 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 $\{\vec{\mathbf{i}}', \vec{\mathbf{j}}'\}$ di $\mathcal{V}_{O'}$ sono uguali alle coordinate del vettore $\overrightarrow{OP''} = \overrightarrow{OP} - \overrightarrow{OO'} \in \mathcal{V}_O$ relative alla base $\{\vec{\mathbf{i}}'', \vec{\mathbf{j}}''\}\}$. Dunque

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

e quindi, per l'unicità delle coordinate,

$$\begin{cases} x_1 = b_{11}x_1' + b_{12}x_2' + c_1 \\ x_2 = b_{21}x_1' + b_{22}x_2' + c_2 \end{cases} \tag{8.6}$$

ossia vale la (8.5). Il resto dell'enunciato è immediato. □

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* $\mathcal{R} = RA(O, \vec{\mathbf{i}}, \vec{\mathbf{j}}, \vec{\mathbf{k}})$ *e* $\mathcal{R}' = RA(O', \vec{\mathbf{i}}', \vec{\mathbf{j}}', \vec{\mathbf{k}}')$. *Allora*

$$x = Bx' + c \tag{8.7}$$

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

$$\vec{\mathbf{i}}'' = \overrightarrow{OA'} - \overrightarrow{OO'} \quad \text{e} \quad \vec{\mathbf{j}}'' = \overrightarrow{OB'} - \overrightarrow{OO'} \quad \vec{\mathbf{k}}'' = \overrightarrow{OC'} - \overrightarrow{OO'} \tag{8.8}$$

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

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 $T\colon V \to W$ un'applicazione lineare. Siano fissate una base $\mathcal{B} = \{v_1, \ldots, v_n\}$ per V e una base $\mathcal{C} = \{w_1, \ldots, w_m\}$ 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 $i = 1, \ldots, n$ si ha per $v_i \in \mathcal{B}$

$$T(v_i) = a_{1i}w_1 + \ldots + a_{mi}w_m \tag{9.1}$$

dove $a_{1i}w_1, \ldots, a_{mi}$ sono le coordinate di $T(v_i)$ relative alla base C. Chiamiamo la matrice

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \tag{9.2}$$

che ha per colonne le coordinate relative alla base 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* 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:*

$$y = Ax \tag{9.3}$$

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

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

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

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

$$y_j = a_{j1} x_1 + \ldots + a_{jn} x_n$$

che equivale a (9.3). $\square$

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

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

gli isomorfismi definiti da

$$F_{\mathcal{B}}(\sum_{i=1}^{n} x_i v_i) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \qquad F_{\mathcal{C}}(\sum_{j=1}^{n} y_j w_j) = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}. \tag{9.4}$$

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

$$F_{\mathcal{B}}(v) = x \qquad F_{\mathcal{C}}(F(v)) = y$$

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 $F_{\mathcal{C}} \circ T \circ F_{\mathcal{B}}^{-1}: \mathbb{K}^n \to \mathbb{K}^m$ 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 $A^1, \ldots, A^n$ della matrice A sono le immagini dei vettori della base canonica e quindi, per $j = 1, \ldots, n$, 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 il 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 $T: V \to W$ un'applicazione lineare. Siano fissate una base $\mathcal{B} = \{v_1, \ldots, v_n\}$ di V e una base $\mathcal{C} = \{w_1, \ldots, w_m\}$ 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 $\{\eta_1, \ldots, \eta_r\}$ è una base di $Im(A)$, allora $\{F_\mathcal{C}^{-1}(\eta_1), \ldots, F_\mathcal{C}^{-1}(\eta_r)\}$ è una base di $Im(T)$;

(ii) Se $\{\epsilon_1, \ldots, \epsilon_k\}$ è una base di $Ker(A)$, allora $\{F_\mathcal{B}^{-1}(\epsilon_1), \ldots, F_\mathcal{B}^{-1}(\epsilon_k)\}$ 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 isomorfimo 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 isomorfimo e quindi trasforma basi di $Ker(A)$ in basi di $Ker(T)$. □

Il Teorema 9.1 permette di dimostrare in generale che lo spazio vettoriale delle applicazioni lineari fra spazi vettorili 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 $\mathcal{B} = \{v_1, \ldots, v_n\}$ di V e $\mathcal{C} = \{w_1, \ldots, w_m\}$ 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 $\Phi: \mathcal{L}(V, W) \to M_{m,n}(\mathbb{K})$ 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 $T(v_j) = a_{1j}w_1 + \ldots + a_{mj}w_j$. Allora $\Phi(T) = A$. Dunque Φ è l'isomorfismo cercato. □

Osservazione. Si osservi che l'isomorfismo costruito in Teorema 9.3dipende 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 $T: V \to W$ un'applicazione lineare. Siano $\mathcal{B}$, $\mathcal{B}'$ basi di V $\mathcal{C}$, $\mathcal{C}'$ 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*

$$A' = C^{-1}AB. \tag{9.5}$$

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

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

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 ne ldominio e nel codominio. Se A è la matrice di $T \colon V \to V$ 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

$$A' = B^{-1}AB. \tag{9.7}$$

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.*