Capitolo 5

Matrici e sistemi lineari di equazioni

5.1 Spazi delle righe e delle colonne

Sia

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

una matrice $m \times n$ a coefficienti in K. Useremo la seguente notazione per le righe di A e le interpreteremo come elementi di K^n:

$a_1 = (a_{11}, \dots, a_{1n})$

$\dots$

$a_m = (a_{m1}, \dots, a_{mn})$;

e similmente per le colonne di A, che interpreteremo come elementi di K^m:

$$a^1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \ \cdots, a^n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}.$$

Diremo che una matrice A è quadrata se $m = n$.

Definizione 5.1.1 (Spazio delle righe e delle colonne). Spazio delle righe di A è il sottospazio vettoriale di K^n generato dalle righe di A: $< a_1, \dots, a_m >$. La sua dimensione è detta **rango per righe** di A. Spazio delle colonne di A è il sottospazio vettoriale di K^m generato dalle colonne di A: $< a^1, \dots, a^n >$. La sua dimensione è detta **rango per colonne** di A.

Chiaramente si ha che entrambi i ranghi sono minori o uguali sia di m sia di n (perchè?). Anticipiamo un teorema che sarà dimostrato successivamente.

Teorema 5.1.2. *Sia A una matrice a coefficienti in un campo K. Il suo rango per righe coincide con il suo rango per colonne.*

Esempio 5.1.3.

1. **Matrice identica o matrice identità**: per ogni $n \geq 1$ è la matrice quadrata $n \times n$

$$E_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} = (\delta_{ij})_{i,j=1,\dots,n},$$

dove i simboli δ_{ij} detti **simboli di Kronecker** valgono 1 se $i = j$ e 0 se $i \neq j$. Sia le righe sia le colonne di E_n sono i vettori $e_1, \dots, e_n$ della base canonica di K^n, dunque il rango per righe e il rango per colonne di E_n valgono n.

2.
$$A = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 2 \end{pmatrix}:$$

i ranghi per righe e per colonne sono uguali a 2.

$$B = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 4 & -2 \end{pmatrix}:$$

i ranghi per righe e per colonne di B sono uguali a 1.

3.
$$A = \begin{pmatrix} 0 & 1 & 3 & 4 \\ 2 & \frac{1}{2} & 1 & -1 \\ 4 & 0 & -1 & -6 \end{pmatrix}:$$

i ranghi sono uguali a 2; consideriamo il rango per righe: siccome vi sono righe non proporzionali è ≥ 2, inoltre si vede abbastanza facilmente che la terza riga è combinazione lineare delle prime due, e precisamente $a_3 = -a_1 + 2a_2$. Dal Teorema 5.1.2 segue che anche il rango per colonne è 2: verificarlo direttamente per esercizio.

5.2 Sistemi lineari di equazioni I

Sia $A = (a_{ij})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ una matrice $m \times n$ a entrate in K, e siano $b_1, \dots, b_m \in K$ scalari in numero uguale al numero delle righe di A. Siano $x_1, \dots, x_n$ incognite in numero pari al numero delle colonne di A.

Consideriamo il seguente sistema lineare di m equazioni nelle n incognite $x_1, \dots, x_n$:

$$\begin{cases} a_{11}x_1 + & \cdots \quad a_{1n}x_n & = b_1 \\ \vdots & \qquad\;\; \vdots & \;\vdots \\ a_{m1}x_1 + & \cdots \quad a_{mn}x_n & = b_m. \end{cases} \tag{5.1}$$

Si dice che A è la **matrice dei coefficienti** del sistema lineare (5.1). **La matrice completa** del sistema è invece la matrice $m \times (n+1)$ ottenuta aggiungendo ad A una colonna b formata dagli scalari $b_1, \dots, b_m$; la si denota $(A \mid b)$.

Una **soluzione del sistema** è un vettore $v = (v_1, \dots, v_n)$ di K^n tale che sostituendo al posto di ogni incognita x_i la coordinata v_i di v le relazioni in (5.1) siano tutte soddisfatte, ossia si abbia

$$\begin{cases} a_{11}v_1+ & \cdots & a_{1n}v_n & = b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1}v_1+ & \cdots & a_{mn}v_n & = b_m. \end{cases} \tag{5.2}$$

Un sistema lineare si dice **compatibile** se ha almeno una soluzione.

Se $b = (b_1, \dots, b_m)$ è il vettore nullo, ossia se $b_1 = \cdots = v_m = 0$ si dice che il sistema lineare è **omogeneo**. Si dice anche che

$$\begin{cases} a_{11}x_1+ & \cdots & a_{1n}x_n & = 0 \\ \vdots & & \vdots & \vdots \\ a_{m1}x_1+ & \cdots & a_{mn}x_n & = 0 \end{cases} \tag{5.3}$$

è il sistema lineare omogeneo associato al sistema (5.1).

Proposizione 5.2.1. *Un sistema omogeneo (5.3) è sempre compatibile, e l'insieme delle sue soluzioni è un sottospazio vettoriale di K^n.*

Dimostrazione. Osserviamo che il vettore nullo è una soluzione, in quanto

$$\begin{cases} a_{11}0+ & \cdots & a_{1n}0 & = 0 \\ \vdots & & \vdots & \vdots \\ a_{m1}0+ & \cdots & a_{mn}0 & = 0. \end{cases}$$

Sia W l'insieme delle soluzione del sistema (5.3); W risulta chiuso rispetto alle due operazioni di somma e di prodotto esterno, infatti se $v = (v_1, \dots, v_n)$, $w = (w_1, \dots, w_n)$ sono soluzioni, anche $(v_1, \dots, v_n) + (w_1, \dots, w_n)$ e $\lambda(v_1, \dots, v_n)$ lo sono. Dunque W è un sottospazio vettoriale di K^n. □

Un sistema lineare non omogeneo può essere compatibile o meno; se lo è può avere una o più soluzioni. Per esempio nel primo caso, in cui $m = n = 1$, cioè un'equazione in un'incognita, per l'equazione $ax = b$ si possono presentare i seguenti casi:

1. $a \neq 0$, c'è una e una sola soluzione $\frac{b}{a}$;
2. $a = 0$, $b \neq 0$, il sistema non è compatibile;
3. $a = b = 0$: ogni elemento $v \in K$ è soluzione.

Osservazione 5. Il sistema lineare (5.1) si può riscrivere nella forma

$$x_1a^1 + x_2a^2 + \cdots + x_na^n = b$$

ossia

$$x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + \cdots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

Se ne deduce immediatamente la seguente prima condizione perchè un sistema sia compatibile.

Proposizione 5.2.2. *Il sistema lineare (5.1) è compatibile se e solo se il vettore b appartiene allo spazio delle colonne della matrice A, cioè è combinazione lineare delle colonne di A. In tal caso una soluzione* $(v_1, \ldots, v_n)$ *è data dai coefficienti di una combinazione lineare di* $a^1, \ldots, a^n$ *uguale a b. Se in più* $a^1, \ldots, a^n$ *sono linearmente indipendenti, la soluzione è unica.*

Vediamo qualche esempio di conseguenze della Proposizione 5.2.2.

Esempio 5.2.3.
1. Consideriamo un sistema omogeneo (5.3): è sempre compatibile e ha quanto meno la soluzione nulla. Ha anche soluzioni non nulle se e solo se le colonne di A sono linearmente dipendenti. Ciò si esprime dicendo che il rango per colonne di A è strettamente minore del numero n delle colonne. Dal momento che il rango per colonne è minore o uguale al minimo fra m e n, certamente il sistema ha soluzioni non nulle se $m < n$.
2. Caso $m = n$. Se le colonne di A sono linearmente indipendenti costituiscono una base di K^n, allora **qualunque sia** b, il sistema è compatibile e ha una e una sola soluzione.

5.3 Trasformazioni elementari sulle righe di una matrice

Il nostro obiettivo è di dare un metodo per risolvere i sistemi di equazioni lineari, cioè per determinare innanzitutto se un dato sistema lineare (5.1) è compatibile, e in caso affermativo per trovarne tutte le soluzioni.

A tale scopo introdurremo delle trasformazioni che mutano un dato sistema lineare in un sistema equivalente.

Definizione 5.3.1. Due sistemi lineari di m equazioni in n incognite (stesso numero di equazioni e di incognite) si dicono **equivalenti** se hanno lo stesso insieme di soluzioni, in particolare uno è compatibile se e solo se lo è l'altro.

Vedremo una classe di trasformazioni, dette trasformazioni elementari sulle righe di una matrice, che, se applicate alle righe della matrice completa $(A \mid b)$ di un sistema lineare, lo mutano in un sistema equivalente.

Introduciamo ora quattro tipi di trasformazioni elementari. Sia

$$C = (c_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}} = \begin{pmatrix} c_1 \\ \vdots \\ c_m \end{pmatrix}$$

una matrice $m \times n$.

I tipo: moltiplicazione di una riga per uno scalare $\lambda \neq 0$:

$$\begin{pmatrix} c_1 \\ \vdots \\ c_i \\ \vdots \\ c_m \end{pmatrix} \xrightarrow{I} \begin{pmatrix} c_1 \\ \vdots \\ \lambda c_i \\ \vdots \\ c_m \end{pmatrix}$$

II tipo: addizione della j-esima riga alla i-esima riga:

$$\begin{pmatrix} c_1 \\ \vdots \\ c_i \\ \vdots \\ c_j \\ \vdots \\ c_m \end{pmatrix} \xrightarrow{II} \begin{pmatrix} c_1 \\ \vdots \\ c_i + c_j \\ \vdots \\ c_j \\ \vdots \\ c_m \end{pmatrix}$$

III tipo: addizione di un multiplo della j-esima riga alla i-esima riga:

$$\begin{pmatrix} c_1 \\ \vdots \\ c_i \\ \vdots \\ c_j \\ \vdots \\ c_m \end{pmatrix} \xrightarrow{III} \begin{pmatrix} c_1 \\ \vdots \\ c_i + \lambda c_j \\ \vdots \\ c_j \\ \vdots \\ c_m \end{pmatrix}$$

IV tipo: scambio della j-esima con la i-esima riga:

$$\begin{pmatrix} c_1 \\ \vdots \\ c_i \\ \vdots \\ c_j \\ \vdots \\ c_m \end{pmatrix} \xrightarrow{IV} \begin{pmatrix} c_1 \\ \vdots \\ c_j \\ \vdots \\ c_i \\ \vdots \\ c_m \end{pmatrix}$$

Osservazione 6. Una trasformazione del III tipo si può ottenere applicando successivamente trasformazioni del I e del II tipo; precisamente si moltiplica dapprima la riga j-esima per λ (I tipo), poi si somma la j-esima alla i-esima riga (II tipo), e infine si moltiplica la riga j-esima per λ^{-1} (I tipo). Anche una trasformazione del IV tipo si può ottenere applicando trasformazioni del I e del II tipo (come?).

Proposizione 5.3.2. *Sia C una matrice. Lo spazio delle righe di C non cambia se si applicano a C trasformazioni elementari sulle righe.*

Dimostrazione. La dimostrazione si basa sulla seguente semplice osservazione:

Osservazione 7. Siano $U =< u_1, \dots, u_s >$, $W =< w_1, \dots, w_r >$ sottospazi finitamente generati di uno spazio vettoriale V. Allora $U = W$ se e solo se $w_1, \dots, w_r \in U$ e $u_1, \dots, u_s \in W$. Infatti, ricordiamo che U è il minimo sottospazio che contiene $u_1, \dots, u_s$ e W è il minimo sottospazio che contiene $w_1, \dots, w_r$. Dunque se $w_1, \dots, w_r \in U$, allora anche $< w_1, \dots, w_r > \subset U$, ossia $W \subset U$. Similmente si ha l'inclusione opposta.

Nel caso di una trasformazione del I tipo, basta dunque verificare che

$$c_i \in < c_1, \dots, \lambda c_i, \dots, c_m >;$$

ma questo è vero perchè $c_i = 0c_1 + \cdots + \lambda^{-1}(\lambda c_i) + \cdots + 0c_m$. Per le trasformazioni del II tipo, osserviamo che $c_i = 0c_1 + \cdots + 1(c_i + c_j) + \cdots - c_j \cdots$. Per quelle del III tipo si ha $c_i = 0c_1 + \cdots + 1(c_i + \lambda c_j) + \cdots - \lambda c_j \cdots$. Il caso delle trasformazioni del IV tipo è chiaro, perchè si è solo cambiato l'ordine in cui si considerano i generatori. □

Abbiamo ora l'applicazione ai sistemi di equazioni lineari.

Proposizione 5.3.3. *Dato il sistema lineare (5.1), consideriamo la sua matrice completa $A' = (A \mid b)$. Se si applicano ad A' trasformazioni elementari sulle righe, si ottiene la matrice completa di un sistema lineare equivalente a quello di partenza.*

Dimostrazione. Infatti con una trasformazione del I tipo si moltiplica una delle equazioni per una costante non nulla; con una del II tipo si sostituisce ad un'equazione la sua somma con una delle altre equazioni. Il III tipo è una combinazione dei tipi precedenti, e con il IV tipo si scambia l'ordine di due delle equazioni. □

5.4 Algoritmo di eliminazione di Gauss

L'algoritmo di eliminazione di Gauss fa passare, mediante trasformazioni elementari sulle righe della matrice completa, da un sistema lineare (5.1) ad un sistema ad esso equivalente che è "facile da risolvere". Precisamente, la matrice completa $(A \mid b)$ viene trasformata in una matrice "a gradini o a scala".

Definizione 5.4.1. Una **matrice a gradini** o **a scala** è una matrice della forma

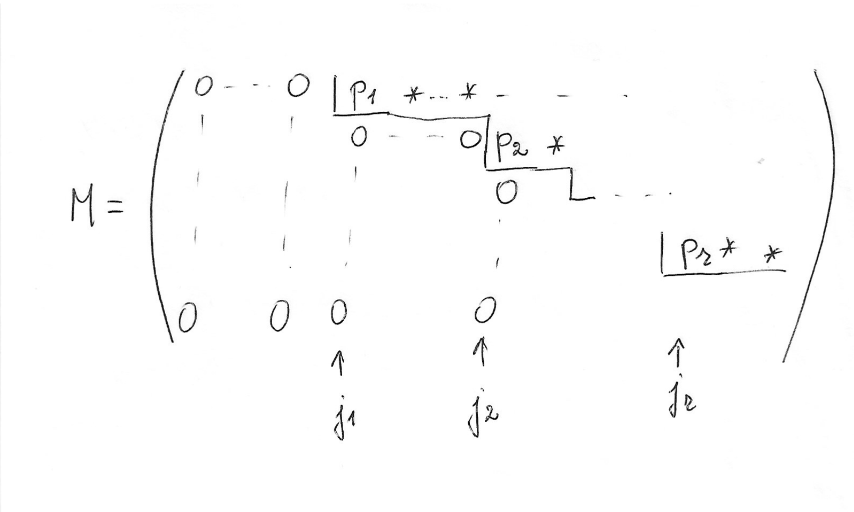

Figura 5.1: Matrice a gradini

dove $p_1, \dots, p_r$ sono scalari diversi da 0. Sotto la "scala" è tutto 0. Gli asterischi significano che in quelle posizioni vi può essere qualunque cosa. Gli r elementi $p_1, \dots, p_r$ si trovano nelle colonne di indici $j_1, \dots, j_r$, e si chiamano i "pivot" della matrice M.

Il pivot p_1 si trova nella prima riga ed è il "primo" elemento non nullo della prima riga, è detto primo pivot. Sotto al primo pivot nella colonna j_1-esima è tutto 0; p_2 è il primo elemento non nullo della seconda riga, e sotto è tutto 0, e così via.

Proposizione 5.4.2. *Sia M una matrice a gradini come nella Figura 5.1. Le sue prime r righe $m_1, \dots, m_r$ formano una base dello spazio delle righe di M, dunque il rango per righe di M è uguale a r.*

Dimostrazione. Le righe $m_1, \dots, m_r$ sono linearmente indipendenti; infatti una loro combinazione lineare nulla $\lambda_1 m_1 + \cdots + \lambda_r m_r = 0$ ha la forma

$$(0, \dots, 0, \lambda_1 p_1, \dots, \lambda_2 p_2 + \lambda_1(*), \dots, \lambda_r p_r + \lambda_{r-1}(*) + \cdots, 0, \dots, 0) = 0;$$

essendo i pivot tutti non nulli, segue $\lambda_1 = \lambda_2 = \cdots = \lambda_r = 0$. Le righe successive $m_{r+1}, \dots, m_n$ sono tutte nulle. □

Passiamo ora alla descrizione dell'algoritmo di eliminazione di Gauss.

Algoritmo. Si parte da una matrice $M \in M(s \times n, K)$ e la si trasforma in una matrice M' a scala con trasformazioni elementari sulle righe.

- Se M è la matrice nulla abbiamo finito.
- Se M non è la matrice nulla, sia m^{j_1} la prima colonna non nulla. Scambiando eventualmente la prima riga con una riga successiva (trasformazione del IV tipo) possiamo supporre che sia $m_{1j_1} \neq 0$, e poniamo $p_1 = m_{1j_1}$.
- Per ogni $h \geq 2$ sommiamo alla riga h-esima un opportuno multiplo della I riga, in modo da annullare tutti gli elementi sotto a p_1 (trasformazione del III tipo). La prima riga rimane fissa.
- Consideriamo le righe $m_2, \dots, m_s$: se sono tutte nulle abbiamo finito. Altrimenti sia j_2 l'indice della prima colonna che contiene un elemento non nullo nella righe $m_2, \dots, m_s$. Scambiando eventualmente la seconda riga con una riga successiva (trasformazione del IV tipo) possiamo supporre che sia $m_{2j_2} \neq 0$, e poniamo $p_2 = m_{2j_2}$.
- Con trasformazioni elementari del III tipo, per ogni $h \geq 3$ sommiamo alla riga h-esima un opportuno multiplo della seconda riga, in modo da annullare tutti gli elementi sotto a p_2.
- Proseguiamo in questo modo, finchè arriviamo o ad avere l'ultimo pivot nell'ultima riga, oppure le ultime righe tutte nulle.

Esempio 5.4.3. Nell'esempio della Figura 5.2 sono mostrate le trasformazioni elementari eseguite; la matrice ha rango (per righe) uguale a 3; i pivot sono nelle colonne di indici $j_1 = 2$, $j_2 = 4$, $j_3 = 5$. Dopo la trasformazione, le righe m_1, m_2, m_3 sono linearmente indipendenti, mentre la riga m_4 è nulla.

Un sistema lineare è detto a gradini se lo è la sua matrice completa. Un sistema lineare a gradini si risolve con il metodo di "sostituzione all'indietro". Lo illustriamo con un esempio.

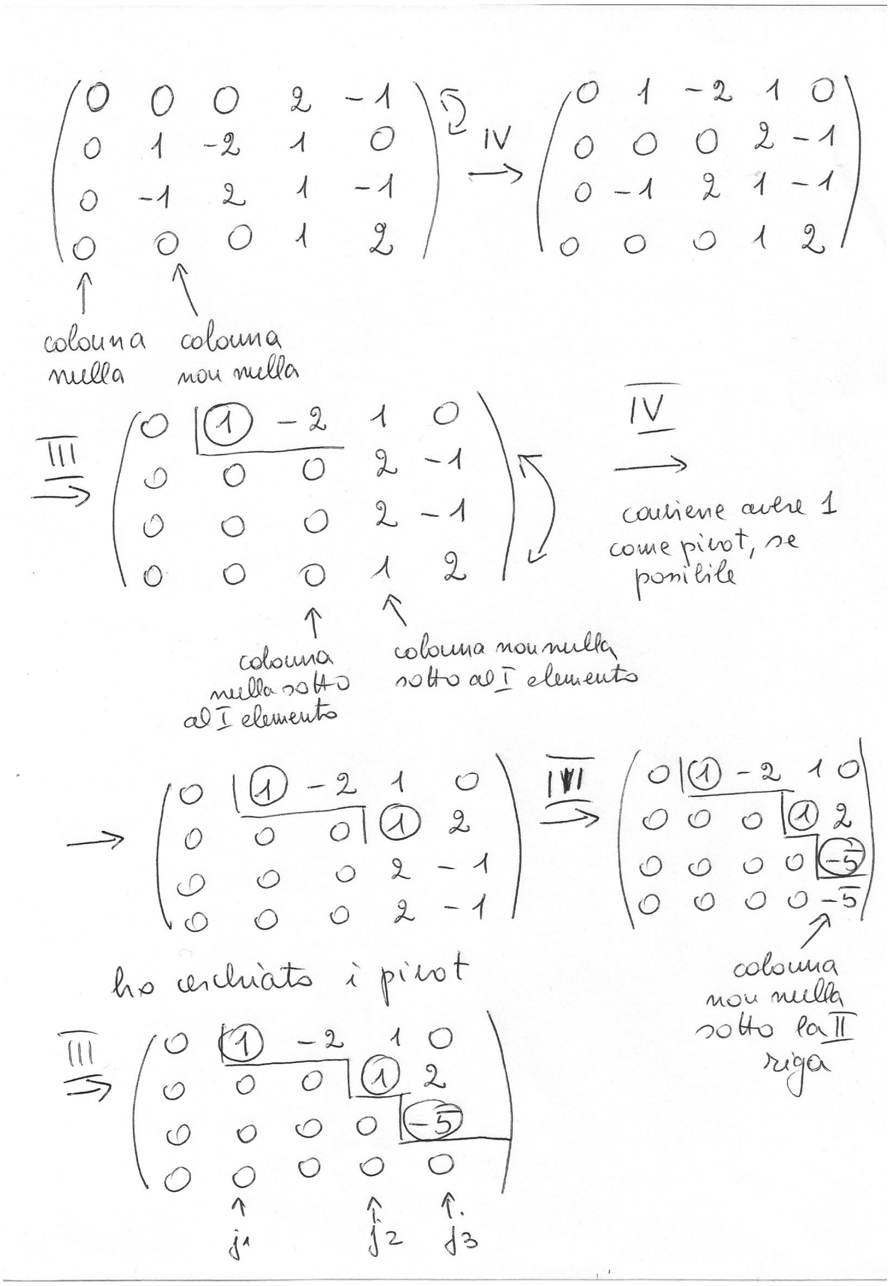

Figura 5.2: Esempio di applicazione dell'algoritmo di eliminazione di Gauss

Esempio 5.4.4. Consideriamo il sistema lineare

$$\begin{cases} x_1 + 2x_2 + x_3 = -1 \\ 2x_1 + x_2 - x_3 = 1 \\ -x_1 - x_2 + 2x_3 = 1 \end{cases} \tag{5.4}$$

Consideriamo $A' = (A \mid b)$ la sua matrice completa e le applichiamo l'algoritmo di eliminazione di Gauss:

$$\left(\begin{array}{rrr|r} 1 & 2 & 1 & -1 \\ 2 & 1 & -1 & 1 \\ -1 & -1 & 2 & 1 \end{array}\right) \xrightarrow{III} \left(\begin{array}{rrr|r} 1 & 2 & 1 & -1 \\ 0 & -3 & -3 & 3 \\ 0 & 1 & 3 & 0 \end{array}\right) \xrightarrow{I}$$

$$\rightarrow \left(\begin{array}{rrr|r} 1 & 2 & 1 & -1 \\ 0 & 1 & 1 & -1 \\ 0 & 1 & 3 & 0 \end{array}\right) \xrightarrow{III} \left(\begin{array}{rrr|r} \boxed{1} & 2 & 1 & -1 \\ 0 & \boxed{1} & 1 & -1 \\ 0 & 0 & \boxed{2} & 1 \end{array}\right):$$

abbiamo ottenuto una matrice a gradini con 3 pivot, dunque è di rango 3. Il sistema (5.4) è equivalente al seguente:

$$\begin{cases} x_1 + 2x_2 + x_3 = -1 \\ \quad\quad\ x_2 + x_3 = -1 \\ \quad\quad\quad\quad 2x_3 = 1 \end{cases} \tag{5.5}$$

Dall'ultima equazione ricaviamo $x_3 = \frac{1}{2}$; sostituiamo il valore di x_3 trovato nella seconda e otteniamo $x_2 = -1 - \frac{1}{2} = -\frac{3}{2}$, e infine sostituiamo nella prima i valori di x_2 e di x_3 e otteniamo $x_1 = 3 - \frac{1}{2} - 1 = \frac{3}{2}$. Quindi il sistema ha l'unica soluzione $(\frac{3}{2}, -\frac{3}{2}, \frac{1}{2})$. In effetti, le tre colonne della matrice dei coefficienti del sistema (5.5) sono chiaramente linearmente indipendenti, quindi il sistema ha una e una sola soluzione qualunque sia la colonna dei termini noti (Esempio 5.2.3, 2.). Osserviamo che trasformando A' in una matrice a gradini, si è trasformata in forma a gradini anche A.

5.5 Soluzioni di un sistema lineare compatibile

D'ora in poi per indicare il sistema lineare (5.1) useremo spesso la notazione compatta $AX = b$, dove con X intendiamo la n-upla delle incognite $x_1, \ldots, x_n$ scritte in colonna: $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$. Similmente, se v è una soluzione, scriveremo $Av = b$. Il sistema omogeneo associato è $AX = 0$: abbiamo osservato nella Proposizione 5.2.1 che è sicuramente compatibile e che l'insieme delle sue soluzioni è un sottospazio vettoriale di K^n, che denoteremo con W.

Proposizione 5.5.1. *Supponiamo che il sistema lineare $AX = b$ sia compatibile e denotiamo con $S \subset K^n$ l'insieme delle sue soluzioni.*

1. Se $b \neq 0$, cioè se il sistema lineare (5.1) non è omogeneo, allora S non è un sottospazio vettoriale di K^n.

2. Se $\bar{v} \in K^n$ è una soluzione particolare del sistema, si ha che

$$S = \bar{v} + W = \{\bar{v} + w \mid w \in W\},$$

dove W è lo spazio delle soluzioni del sistema omogeneo associato.

Dimostrazione. 1. Se il sistema non è omogeneo, il vettore nullo 0 non è una soluzione, quindi S non è sottospazio vettoriale.

2. Se w è una soluzione del sistema omogeneo associato, $Aw = 0$; per ipotesi $A\bar{v} = b$, dunque si ha $A(\bar{v} + w) = b$. Viceversa, se $A\bar{v} = b$ e $Av = b$, si ha $A(v - \bar{v}) = 0$ e quindi $v - \bar{v} \in W$, da cui segue che $v \in \bar{v} + W$. □

Osservazione 8. Essendo W un sottospazio vettoriale di K^n si può considerare lo **spazio vettoriale quoziente** K^n/W introdotto nel Foglio 2 di esercizi (Esercizio 2): è il quoziente rispetto alla relazione d'equivalenza secondo cui $v \sim v'$ se e solo se $v - v' \in W$, ed è un K-spazio vettoriale rispetto alle operazioni indotte dalle operazioni di K^n. Allora $\bar{v} + W$ risulta essere la classe d'equivalenza del vettore $\bar{v}$ in K^n/W.

Si ha anche la seguente definizione di sottospazio affine di uno spazio vettoriale.

Definizione 5.5.2. Sia V un K-spazio vettoriale. Un **sottospazio affine** di V è un sottinsieme di V del tipo $S = v + W$, dove $v \in V$ e W è un sottospazio vettoriale detto **giacitura di** S. Si definisce la dimensione di S, $\dim S$, come la dimensione della sua giacitura: $\dim S = \dim W$.

Poichè $0 \in W$, allora $v \in S$: si dice che S è un sottospazio affine passante per v. Osserviamo che se $u \in S = v + W$ è un qualunque vettore di S, allora si ha anche $S = u + W = v + W$. Infatti sia $u = v + w$, con $w \in W$; per ogni $w' \in W$, $u + w' = v + (w + w') \in v + W$; viceversa $v + w' = (u - w) + w' = u + (w' - w) \in u + W$.

Esempio 5.5.3. Se $W = (0)$ è il sottospazio nullo, i sottospazi affini di giacitura W sono i singoli vettori di V.

Se $\dim W = 1$, i sottospazi di dimensione 1 di giacitura W sono le rette "parallele" a W, o di direzione W.

E così via, se $\dim W = 2$ si ottengono i piani paralleli a W.

Con questo linguaggio possiamo riformulare come segue la Proposizione 5.5.1.

Corollario 5.5.4. *L'insieme delle soluzioni di un sistema lineare compatibile è un sottospazio affine avente come giacitura lo spazio delle soluzioni del sistema omogeneo associato, e passante per una soluzione particolare del sistema.*

5.6 Risoluzione dei sistemi omogenei

Vediamo ora come trovare la dimensione e una base dello spazio W delle soluzioni di un sistema lineare omogeneo.

Iniziamo con un esempio.

Esempio 5.6.1. Un sistema lineare omogeneo di 5 equazioni in 7 incognite $x_1, \ldots, x_7$ a coefficienti in $\mathbb{R}$.

$$\begin{cases} x_2 + 2x_4 - x_5 - 4x_6 = 0 \\ x_3 - x_4 - x_5 + 2x_6 + x_7 = 0 \\ x_2 + 2x_4 + x_5 - 2x_6 = 0 \\ x_3 - x_4 + 2x_6 - x_7 = 0 \\ x_2 + x_3 + x_4 + x_7 = 0 \end{cases} \tag{5.6}$$

Essendo il sistema omogeneo, la matrice completa ha soltanto una colonna nulla a destra della matrice dei coefficienti A; tale colonna rimane invariata per trasformazioni elementari; si lavora soltanto con la matrice dei coefficienti A.

$$\begin{pmatrix} 0 & \boxed{1} & 0 & 2 & -1 & -4 & 0 \\ 0 & 0 & 1 & -1 & -1 & 2 & 1 \\ 0 & 1 & 0 & 2 & 1 & -2 & 0 \\ 0 & 0 & 1 & -1 & 0 & 2 & -1 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{III} \begin{pmatrix} 0 & \boxed{1} & 0 & 2 & -1 & -4 & 0 \\ 0 & 0 & \boxed{1} & -1 & -1 & 2 & 1 \\ 0 & 0 & 0 & 0 & 2 & 2 & 0 \\ 0 & 0 & 1 & -1 & 0 & 2 & -1 \\ 0 & 0 & 1 & -1 & 1 & 4 & 1 \end{pmatrix} \xrightarrow{III}$$

$$\rightarrow \begin{pmatrix} 0 & \boxed{1} & 0 & 2 & -1 & -4 & 0 \\ 0 & 0 & \boxed{1} & -1 & -1 & 2 & 1 \\ 0 & 0 & 0 & 0 & 2 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 0 & 2 & 2 & 0 \end{pmatrix} \xrightarrow{I} \begin{pmatrix} 0 & \boxed{1} & 0 & 2 & -1 & -4 & 0 \\ 0 & 0 & \boxed{1} & -1 & -1 & 2 & 1 \\ 0 & 0 & 0 & 0 & \boxed{1} & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \end{pmatrix} \xrightarrow{III}$$

$$\rightarrow \begin{pmatrix} 0 & \boxed{1} & 0 & 2 & -1 & -4 & 0 \\ 0 & 0 & \boxed{1} & -1 & -1 & 2 & 1 \\ 0 & 0 & 0 & 0 & \boxed{1} & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & \boxed{1} & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Siamo arrivati a una matrice a gradini con $j_1 = 2$, $j_2 = 3$, $j_3 = 5$, $j_4 = 6$, $r = 4$.

Il sistema a gradini corrispondente, equivalente a (5.6), è il seguente.

$$\begin{cases} \boxed{x_2} + 2x_4 - x_5 - 4x_6 = 0 \\ \boxed{x_3} - x_4 - x_5 + 2x_6 + x_7 = 0 \\ \boxed{x_5} + x_6 = 0 \\ \boxed{x_6} + 2x_7 = 0 \\ 0 = 0 \end{cases} \tag{5.7}$$

Notiamo che l'ultima equazione non è significativa e non occorre scriverla. Dall'ultima equazione non banale ricaviamo x_6 in funzione di x_7, su cui non vi sono condizioni: $x_6 = -2x_7$, poi sostituiamo nella terza: $x_5 = -x_6 = 2x_7$; su x_4 non vi sono condizioni, sostituiamo nella seconda per ricavare x_3: $x_3 = x_4 + 2x_7 + 4x_7 - x_7 = x_4 + 5x_7$; infine dalla prima ricaviamo $x_2 = -2x_4 + x_5 + 4x_6 = -2x_4 + 2x_7 - 8x_7 = -2x_4 - 6x_7$. Abbiamo ricavato i valori delle incognite corrispondenti ai pivot, x_2, x_3, x_5, x_6 in funzione delle altre, che acquistano il significato di parametri e sono dette **parametri liberi.** In particolare su x_1 non vi sono condizioni, è un parametro libero. Le incognite che abbiamo calcolato in funzione delle altre sono quelle che stanno nelle colonne dei pivot, pertanto sono in numero di r; i parametri liberi sono quindi $n - r$, e corrispondono alle colonne in cui non ci sono pivot.

La soluzione generale del sistema si ottiene dando ai parametri liberi tutti i possibili valori in $\mathbb{R}$, ed è pertanto della forma $(\mathbf{t_1}, -2t_4 - 6t_7, t_4 + 5t_7, \mathbf{t_4}, 2t_7, -2t_7, \mathbf{t_7})$, che può essere riscritta come $t_1(1,0,0,0,0,0,0) + t_4(0,-2,1,1,0,0,0) + t_7(0,-6,5,0,2,-2,1)$, al variare di t_1, t_4, t_7 in $\mathbb{R}$. Introduciamo i tre vettori di $\mathbb{R}^7$ $w_1 = (1,0,0,0,0,0,0)$, $w_2 =$

$(0,-2,1,1,0,0,0)$, $w_3 = (0,-6,5,0,2,-2,1)$. Osserviamo che w_1 si ottiene per $t_1 = 1$, $t_4 = t_7 = 0$, w_2 si ottiene per $t_1 = t_7 = 0$ e $t_4 = 1$, e w_3 si ottiene per $t_1 = t_4 = 0$ e $t_7 = 1$. Un vettore di $\mathbb{R}^7$ è soluzione del sistema se e solo se è una combinazione lineare di w_1, w_2, w_3: $t_1w_1 + t_4w_2 + t_7w_3$. Quindi lo spazio delle soluzioni W è generato da w_1, w_2, w_3, inoltre w_1, w_2, w_3 sono linearmente indipendenti; infatti una loro combinazione lineare nulla $t_1w_1 + t_4w_2 + t_7w_3 = 0$ deve avere $t_1 = t_4 = t_7 = 0$, in quanto t_1, t_4, t_7 sono alcune delle sue coordinate.

Abbiamo dunque trovato una base (w_1, w_2, w_3) dello spazio W delle soluzioni, e la sua dimensione è $3 = 7 - 4$, ossia il numero delle incognite meno il numero dei pivot, o anche il numero delle incognite meno il rango per righe della matrice dei coefficienti.

Questo esempio si generalizza al caso di un qualunque sistema lineare omogeneo e ci permette di ottenere direttamente il seguente risultato generale.

Teorema 5.6.2. *Sia $AX = 0$ un sistema lineare omogeneo in n incognite a coefficienti in un campo K. Sia $W \subset K^n$ il suo spazio delle soluzioni. Si ha:*

1. $\dim W = n - r$*, dove r è il rango (per righe) della matrice A;*

2. *sia B una matrice a gradini ottenuta da A con trasformazioni elementari sulle righe; con il metodo di sostituzione all'indietro applicato al sistema $BX = 0$, equivalente a quello di partenza, si trova una soluzione generale del sistema, espressa in funzione di $n - r$ parametri liberi corrispondenti alle incognite non contenute nel colonne dei pivot;*

3. *assegnando agli $n-r$ parametri liberi i valori* $(1,0,\dots,0), (0,1,0,\dots,0),\dots,(0,0,\dots,1)$ *si trova una base di W.*

5.7 Teorema di Rouché - Capelli e risoluzione dei sistemi lineari

Teorema 5.7.1 (Teorema di Rouché-Capelli). *Sia $AX = b$ un sistema lineare a coefficienti in un campo K. Il sistema è compatibile se e solo se il rango per righe della matrice dei coefficienti A coincide con quello della matrice completa $(A \mid b)$.*

Dimostrazione. Con trasformazioni elementari sulle righe di $(A \mid b)$, trasformiamo il sistema $AX = b$ in un sistema equivalente $\bar{A}X = \bar{b}$, dove $\bar{A}$ è una matrice a gradini e $(\bar{A} \mid \bar{b})$ è della seguente forma (per l'equivalenza vedere la Proposizione 5.3.3):

$$\left(\begin{array}{cccccc|c}
\cdots & \boxed{\overline{a_{1j_1}}} & * & * & * & * & \overline{b_1} \\
 & & \boxed{\overline{a_{2j_2}}} & * & * & * & \overline{b_2} \\
 & & & \ddots & & & \cdots \\
 & & & & \boxed{\overline{a_{rj_r}}} & * & \overline{b_r} \\
\cdots & 0 & 0 & \cdots & 0 & 0 & \overline{b_{r+1}} \\
 & & & \ddots & & & \cdots \\
\cdots & 0 & 0 & \cdots & 0 & 0 & \overline{b_m}
\end{array}\right)$$

Chiaramente il sistema è compatibile se e solo se $\overline{b_{r+1}} = \cdots = \overline{b_m} = 0$. Sia r il rango di $\bar{A}$: r differisce dal rango di $(\bar{A} \mid \bar{b})$ se e solo se nell'ultima colonna ci sono più di r elementi non nulli, cioè se c'è un pivot di $(\bar{A} \mid \bar{b})$ nella riga $(r+1)$-esima, e ciò equivale a dire che l'equazione $(r+1)$-esima non è compatibile. Osserviamo che i due ranghi possono differire al più di uno. $\square$

Ricapitolando, sia A una matrice $m \times n$ a entrate in K, denotiamo con $\mathrm{rg}(A)$ il suo rango per righe.

- Il sistema lineare $AX = b$ è compatibile se e solo se $\mathrm{rg}(A) = \mathrm{rg}(A \mid b)$;
- in particolare, se il sistema è omogeneo, esso è compatibile e l'insieme delle soluzioni è un sottospazio vettoriale di K^n di dimensione $n - \mathrm{rg}(A)$;
- se il sistema non è omogeneo ed è compatibile, l'insieme delle soluzioni è un sottospazio affine di K^n del tipo $S = \bar{v} + W$, dove W è lo spazio delle soluzioni del sistema omogeneo associato $AX = 0$; inoltre $\dim S = \dim W = n - r$ dove $r = \mathrm{rg}(A) = \mathrm{rg}(A \mid\ b)$;
- per risolvere il sistema, si riduce $(A \mid b)$ a gradini; dopo aver verificato che i ranghi di A e di $(A \mid b)$ siano uguali, si procede con il metodo di sostituzione all'indietro; questo fornisce una soluzione particolare e una base di W.

Illustriamo l'ultimo punto con il seguente esempio.

Esempio 5.7.2.

$$\begin{cases} x_1 - 2x_2 + x_3 = 1 \\ x_1 - 2x_2 - x_4 = 2 \\ x_3 + x_4 = -1 \end{cases} \tag{5.8}$$

$$\left(\begin{array}{cccc|c} \boxed{1} & -2 & 1 & 0 & 1 \\ 1 & -2 & 0 & -1 & 2 \\ 0 & 0 & 1 & 1 & -1 \end{array}\right) \xrightarrow{III} \left(\begin{array}{cccc|c} \boxed{1} & -2 & 1 & 0 & 1 \\ 0 & 0 & \boxed{-1} & -1 & 1 \\ 0 & 0 & 1 & 1 & -1 \end{array}\right) \xrightarrow{I}$$

$$\to \left(\begin{array}{cccc|c} \boxed{1} & -2 & 1 & 0 & 1 \\ 0 & 0 & \boxed{-1} & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Il sistema è compatibile, $r = 2$, $j_1 = 1$, $j_2 = 3$. Il sistema a gradini è

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

(non abbiamo scritto l'equazione identica $0 = 0$). Ci sono due parametri liberi t_2, t_4 relativi alle colonne senza pivot. La soluzione generale è

$$(x_1, x_2, x_3, x_4) = (2t_2 + t_4 + 2, t_2, -t_4 - 1, t_4) = t_2(2, 1, 0, 0) + t_4(1, 0, -1, 1) + (2, 0, -1, 0).$$

Poniamo $w = (2, 1, 0, 0)$, $w' = (1, 0, -1, 1)$, $\bar{v} = (2, 0, -1, 0)$. $W =< w, w' >$ è lo spazio delle soluzioni del sistema omogeneo associato, $\bar{v}$ è una soluzione particolare del sistema (5.8), ottenuta per $t_2 = t_4 = 0$.

Capitolo 6

Applicazioni lineari

6.1 Prodotto di matrici

Premettiamo la definizione e le prime proprietà dell'operazione di prodotto righe per colonne di matrici, che sarà importante nella descrizione delle applicazioni lineari.

Definizione 6.1.1. Siano $A = (a_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}}$ e $B = (b_{jk})_{\substack{1 \le j \le n \\ 1 \le k \le p}}$, due matrici a coefficienti in un campo K, la prima di ordine $m \times n$ e la seconda di ordine $n \times p$. Il loro **prodotto righe per colonne** è la seguente matrice $m \times p$:

$$C = AB = (c_{ik})_{\substack{1 \le i \le m \\ 1 \le k \le p}}$$

dove

$$c_{ik} = a_{i1}b_{1k} + a_{i2}b_{2k} + \cdots + a_{in}b_{nk} = \sum_{j=1}^{n} a_{ij}b_{jk}.$$

Abbiamo così definito per ogni terna di interi positivi m, n, p un'applicazione

$$M(m \times n, K) \times M(n \times p, K) \to M(m \times p, K)$$

tale che

$$(A, B) \to C = AB.$$

In particolare, se $m = n$ abbiamo definito un prodotto interno in $M(n \times n, K)$.

Esempio 6.1.2. 1.

$$A = \begin{pmatrix} 2 & -1 & 3 \\ 0 & 5 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & -1 \\ 0 & 1 \\ 1 & 2 \end{pmatrix}$$

Il loro prodotto è una matrice 2×2:

$$AB = \begin{pmatrix} 3 & 7 \\ 1 & 7 \end{pmatrix}.$$

Ha senso anche il prodotto BA, che risulta essere la seguente matrice 3×3:

$$BA = \begin{pmatrix} 0 & -5 & -1 \\ 0 & 5 & 1 \\ 2 & 9 & 5 \end{pmatrix}.$$

Se consideriamo invece una matrice 2×2 C, esiste solo CA ma non AC.

2. A matrice riga $1 \times n$: $A = (a_{11}, \ldots, a_{1n})$, B matrice colonna $n \times 1$: $B = \begin{pmatrix} b_{11} \\ \cdots \\ b_{n1} \end{pmatrix}$, allora $AB = (a_{11}b_{11} + a_{12}b_{21} + \cdots + a_{1n}b_{n1})$, matrice 1×1.

3. A matrice colonna $n \times 1$: $A = \begin{pmatrix} a_{11} \\ \cdots \\ a_{n1} \end{pmatrix}$, B matrice riga $1 \times m$: $B = (b_{11}, \ldots, b_{1m})$, allora AB è una matrice $n \times m$:

$$\begin{pmatrix} a_{11}b_{11} & a_{11}b_{12} & \cdots & a_{11}b_{1m} \\ a_{21}b_{11} & a_{21}b_{12} & \cdots & a_{21}b_{1m} \\ \vdots & & & \vdots \\ a_{n1}b_{11} & a_{n1}b_{12} & \cdots & a_{n1}b_{1m} \end{pmatrix},$$

4. Se A è una matrice $m \times n$, ed E_n è la matrice identica $n \times n$, si ha $AE_n = A$; analogamente si ha $E_mA = A$ dove E_m è la matrice identica di ordine m. In particolare se A è una matrice $n \times n$ $AE_n = E_nA = A$. Dunque E_n è elemento neutro per il prodotto in $M(n \times n, K)$.

5. Il prodotto righe per colonne di matrici non gode della proprietà commutativa, come mostra il seguente esempio. Siano

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \ B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Allora

$$AB = \begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix}, \ BA = \begin{pmatrix} 3 & 4 \\ 1 & 2 \end{pmatrix}.$$

6. Se A è una matrice $m \times n$, b è una matrice $m \times 1$ e X denota una colonna di indeterminate $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, ha senso scrivere $AX = b$, ossia per esteso

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

e si ottiene così la (5.1). Questo spiega la notazione usata per indicare i sistemi lineari di equazioni.

7.

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}\begin{pmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{pmatrix} = \text{per la formula di addizione} =$$

$$\begin{pmatrix} \cos(\alpha+\beta) & -\sin(\alpha+\beta) \\ \sin(\alpha+\beta) & \cos(\alpha+\beta) \end{pmatrix}$$

Definizione 6.1.3. Data una matrice $m \times n$ $A = (a_{ij})$, la **trasposta di** A, denotata con ${}^tA$, è la matrice $n \times m$ ottenuta "scambiando" le righe e le colonne di A, ossia l'elemento di indici ij di ${}^tA$ è a_{ji}.

Proposizione 6.1.4. *Il prodotto righe per colonne di matrici verifica le seguenti proprietà, per ogni scelta di matrici A, B, A', B' con il numero appropriato di righe e colonne, e di uno scalare λ :*

a) ${}^t(AB) = {}^tB{}^tA$;

b) $A(B + B') = AB + AB'$,
$(A + A')B = AB + A'B$;

c) $(AB)C = A(BC)$;

d) $(\lambda A)B = \lambda(AB) = A(\lambda B)$.

Dimostrazione. Le verifiche seguono direttamente dalla definizione. Per esempio, per la a), osserviamo che l'elemento di indici ih di $C = AB$ è

$$c_{ih} = \sum_{j=1}^{n} a_{ij}b_{jh}$$

mentre l'elemento di indici hi di $C' = {}^tB{}^tA$ è

$$c'_{hi} = {}^t(b^h){}^t(a_i) = \sum_{j=1}^{n} b_{jh}a_{ij} \text{ che coincide con } c_{ih}.$$

□

6.2 Applicazioni lineari

Siano V, V' spazi vettoriali su uno stesso campo K.

Definizione 6.2.1. Un'applicazione $f : V \to V'$ è detta **applicazione lineare** o omomorfismo se

1. $f(v_1 + v_2) = f(v_1) + f(v_2)$, per ogni $v_1, v_2 \in V$;
2. $f(\lambda v) = \lambda f(v)$ per ogni $\lambda \in K$, $v \in V$.

La condizione 1. si esprime dicendo che f conserva la somma, o è compatibile con la somma, o anche è additiva; la condizione 2. che f conserva il prodotto esterno o è compatibile con il prodotto esterno, o ancora che è omogenea. Si dice anche che un'applicazione lineare è un'applicazione che conserva la struttura di spazio vettoriale.

Osservazione 9. Un'applicazione $f : V \to V'$ è lineare se e solo se per ogni scelta di $\lambda_1, \lambda_2 \in K$, $v_1, v_2 \in V$ si ha $f(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 f(v_1) + \lambda_2 f(v_2)$, ossia f conserva le combinazioni lineari. (Verifica per esercizio.)

Proposizione 6.2.2. *Sia $f : V \to V'$ un'applicazione lineare. Denotiamo con 0_V il vettore nullo di V e $0_{V'}$ il vettore nullo di V'. Si ha*

(i) $f(0_V) = 0_{V'}$;

(ii) $f(-v) = -f(v)$ *per ogni* $v \in V$.

Dimostrazione. (i) $f(0_V) = f(0 \cdot 0_V) =$ per la 2. $= 0f(0_V) = 0_{V'}$, dove 0 denota lo zero di K.

(ii) $f(-v) = f((-1)(v)) =$ per la 2. $= (-1)f(v) = -f(v)$.

□

Come immediata conseguenza abbiamo che se $f(0_V) \neq 0_{V'}$ allora f non può essere lineare. In particolare un'applicazione lineare è costante se e solo se $f(v) = 0_{V'}$ per ogni $v \in V$: tale applicazione è detta **applicazione lineare nulla** ed esiste per ogni coppia di K-spazi vettoriali.

Esempio 6.2.3.

1. Consideriamo $\mathbb{R}$ con la sua struttura naturale di $\mathbb{R}$-spazio vettoriale. Sia $a \in \mathbb{R}$ un numero reale fissato. Consideriamo l'applicazione di moltiplicazione per a:

$$f_a : \mathbb{R} \xrightarrow{\cdot a} \mathbb{R}, \text{ tale che } f_a(x) = ax \text{ per ogni } x \in \mathbb{R}.$$

Si ha $f(\lambda x + \mu y) = a(\lambda x + \mu y) = \lambda(ax) + \mu(ay) = \lambda f_a(x) + \mu f_a(y)$, per ogni $x, y \in \mathbb{R}$, pertanto f_a è un'applicazione lineare.

2. L'applicazione $g : \mathbb{R} \to \mathbb{R}$ definita da $g(x) = 2x + 1$ non è lineare, in quanto $g(0) \neq 0$.

3. L'applicazione $h : \mathbb{R} \to \mathbb{R}$ definita da $h(x) = x^2$ non è lineare; infatti non è additiva: $h(1 + 1) = h(2) = 4 \neq h(1) + h(1) = 1 + 1 = 2$.

4. Sia $f : \mathbb{R}^2 \to \mathbb{R}^3$ definita da

$$f\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3x_1 + 2x_2 \\ x_1 \\ 4x_1 + x_2 \end{pmatrix}.$$

E` facile verificare che f è lineare. Osserviamo che f si può rappresentare anche in forma matriciale mediante un prodotto righe per colonne di matrici come segue:

$$f\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ 1 & 0 \\ 4 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

5. L'applicazione $\varphi : \mathbb{R}^2 \to \mathbb{R}^3$ definita da

$$\varphi \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3x_1 + 1 \\ x_1 + x_2 \\ 4x_1 \end{pmatrix}$$

non è lineare, infatti $\varphi \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$.

Gli esempi 1. e 4. fanno parte di una classe fondamentale di applicazioni lineari **associate a matrici** che ora definiremo.

6.3 L'applicazione lineare $L(A)$.

Definizione 6.3.1. Sia $A \in M(m \times n, K)$ una matrice $m \times n$ a entrate in K. Definiamo l'applicazione lineare associata alla matrice A

$$L(A) : K^n \to K^m$$

ponendo

$$L(A) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum_j a_{1j} x_j \\ \vdots \\ \sum_j a_{mj} x_j \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{pmatrix}.$$

Denotando con X il vettore colonna $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, scriveremo anche $L(A)(X) = AX$.

La linearità dell'applicazione $L(A)$ segue dalla Proposizione 6.1.4:

$$L(A)\Big(\lambda \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \mu \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}\Big) \overset{\text{per definizione}}{=} A\Big(\lambda \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \mu \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}\Big) =$$

$$\overset{\text{per 6.1.4 b)}}{=} A\Big(\lambda \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\Big) + A\Big(\mu \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}\Big) \overset{\text{per 6.1.4 d)}}{=} \lambda A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \mu A \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

$$\overset{\text{per definizione}}{=} \lambda L(A) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \mu L(A) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Osservazione 10. Vediamo chi sono i corrispondenti dei vettori della base canonica di K^n nell'applicazione lineare $L(A)$:

$$L(A)(e_1) = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} = a^1 :$$

è la prima colonna di A. Similmente $L(A)(e_i) = a^i$, la colonna i-esima di A, per ogni indice i. Quindi i corrispondenti dei vettori della base canonica sono le colonne della matrice A.

6.4 Proprietà delle applicazioni lineari

Proposizione 6.4.1. *Sia $f : V \to V'$ un'applicazione lineare.*

(i) Se $W \subset V$ è un sottospazio vettoriale, allora $f(W) \subset V'$ è un sottospazio vettoriale;

(ii) se $W' \subset V'$ è un sottospazio vettoriale, allora $f^{-1}(W') \subset V$ è un sottospazio vettoriale.

Dimostrazione. (i) Per esercizio.

(ii) Siano $v_1, v_2 \in f^{-1}(W')$, dobbiamo verificare che $\lambda_1 v_1 + \lambda_2 v_2 \in f^{-1}(W')$, cioè che $f(\lambda_1 v_1 + \lambda_2 v_2) \in W'$. Ma $f(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 f(v_1) + \lambda_2 f(v_2)$ per la linearità di f. Per ipotesi $f(v_1) \in W'$ e $f(v_2) \in W'$; poichè W' è sottospazio di V' segue che anche $\lambda_1 f(v_1) + \lambda_2 f(v_2) \in W'$; ciò conclude la dimostrazione.

□

Dunque immagini e controimmagini di sottospazi vettoriali in un'applicazione lineare sono sottospazi vettoriali.

Corollario 6.4.2. *Se $f : V \to V'$ è lineare, l'immagine di f, $f(V)$, è un sottospazio vettoriale di V', detto* **sottospazio immagine** *di f. Lo si denota anche* $\operatorname{Im} f$.

La controimmagine del sottospazio nullo di V', $f^{-1}(0)$, è un sottospazio di V, detto **nucleo di** *f. Lo si denota* $\ker f$. *Tale notazione viene dal termine tedesco kern, che significa nucleo, nocciolo; analogamente in inglese kernel.*

Per definizione di applicazione suriettiva, l'applicazione lineare $f : V \to V'$ è suriettiva se e solo se $\operatorname{Im} f = V'$, cioè se l'immagine di f coincide con il codominio V'.

Il nucleo di f

$$\ker f = f^{-1}(0) = \{v \in V \mid f(v) = 0\}$$

dà invece informazioni sull'iniettività di f. Si ha:

Teorema 6.4.3. *Sia $f : V \to V'$ un'applicazione lineare di K-spazi vettoriali; f è iniettiva se e solo se* $\ker f = (0)$, *il sottospazio nullo di V.*

Dimostrazione. Ricordiamo che, essendo f lineare, $f(0_V) = 0_{V'}$. Se f è iniettiva, 0_V è l'unico vettore la cui immagine in f è $0_{V'}$, e ciò significa che $\ker f$, che è uguale a $f^{-1}(0_{V'})$, contiene solo il vettore nullo.

Viceversa supponiamo che $\ker f = (0)$. Siano $v_1, v_2 \in V$ due vettori tali che $f(v_1) = f(v_2)$; dobbiamo dimostrare che $v_1 = v_2$. Da $f(v_1) = f(v_2)$ segue $f(v_1) - f(v_2) = 0$; ma f è lineare, perciò $0 = f(v_1) - f(v_2) = f(v_1 - v_2)$. Ne segue che $v_1 - v_2 \in \ker f$, che è (0) per ipotesi, dunque $v_1 - v_2 = 0$, ossia $v_1 = v_2$. Concludiamo che f è iniettiva. □

Il Teorema 6.4.3 è un criterio molto utile per dimostrare l'iniettività di un'applicazione lineare.

6.5 Spazi Hom

Siano V, V' K-spazi vettoriali. Abbiamo visto che c'è sempre almeno un'applicazione lineare da V a V', l'applicazione nulla. Useremo la notazione $\mathrm{Hom}(V, V')$ per indicare l'insieme non vuoto delle applicazioni lineari di V in V':

$$\mathrm{Hom}(V, V') = \{f : V \to V' \mid f \text{ è lineare}\}.$$

La notazione deriva dal termine "omomorfismo" che è un sinonimo di applicazione lineare.

Introduciamo in $\mathrm{Hom}(V, V')$ due operazioni, di somma e di prodotto esterno con operatori in K, definite punto per punto. Con queste operazioni $\mathrm{Hom}(V, V')$ acquista struttura di K-spazio vettoriale. Date due applicazioni lineari $f, g \in \mathrm{Hom}(V, V')$, e dato $\lambda \in K$, definiamo:

- $f + g$ è l'applicazione di V in V' che tale che, se $v \in V$ è un vettore del dominio, $(f + g)(v) = f(v) + g(v)$;
- λf è l'applicazione di V in V' che tale che, se $v \in V$ è un vettore del dominio, $(\lambda f)(v) = \lambda f(v)$.

Proposizione 6.5.1. $\mathrm{Hom}(V, V')$ *è un* $K-$*spazio vettoriale.*

Dimostrazione. Dobbiamo innanzitutto dimostrare che la somma e il prodotto per uno scalare sono operazioni interne in $Hom(V, V')$, ossia che, se f, g sono lineari e $\lambda \in K$, allora $f + g$ e λf sono lineari. Si ha: $(f + g)(av + bw) \overset{\text{def}}{=} f(av + bw) + g(av + bw) \overset{\text{linearità}}{=} af(v) + bf(w) + ag(v) + bg(w) = a(f(v) + g(v)) + b(f(w) + g(w)) \overset{\text{def}}{=} a(f + g)(v) + b(f + g)(w)$. Similmente $(\lambda f)(av + bw) = a(\lambda f)(v) + b(\lambda f)(w)$ (esercizio).

Osserviamo poi che l'applicazione nulla è elemento neutro per la somma in $\mathrm{Hom}(V, V')$, e che l'opposto di f è l'applicazione lineare $-f = (-1)f$ tale che $(-f)(v) = -f(v)$, per ogni $v \in V$. Si dimostrano poi in maniera simile gli altri assiomi di spazio vettoriale. □

Un caso particolare molto importante è lo spazio vettoriale duale.

Definizione 6.5.2. Sia V un $K-$spazio vettoriale. Lo **spazio vettoriale duale** di V, denotato V^*, è $\mathrm{Hom}(V, K) = \{\varphi : V \to K \mid \varphi \text{ lineare}\}$. Gli elementi di V^* sono chiamati **forme lineari**, o funzionali, su V.

Introduciamo la terminologia usata per denotare vari tipi di applicazioni lineari. Sia $f : V \to V'$ un'applicazione lineare. Diremo che f è un

- epimorfismo se è suriettiva; si dice anche che f è un'applicazione lineare di V **su** V';
- monomorfismo se è iniettiva;
- isomorfismo se è sia iniettiva sia suriettiva;
- endomorfismo se $V = V'$;
- automorfismo se $V = V'$ e f è un isomorfismo.

Nel seguito avremo modo di studiare in maniera approfondita la nozione di endomorfismo.

Vediamo ora una proprietà significativa degli isomorfismi. Sia $f : V \to V'$ un isomorfismo. Ricordiamo che, essendo f biiettiva, esiste l'applicazione inversa $f^{-1} : V' \to V$ tale che, preso $v' \in V'$, si ha: $f^{-1}(v') = v$ se e solo se $f(v) = v'$.

Proposizione 6.5.3. *Se f è un isomorfismo, anche f^{-1} è lineare, e dunque anche f^{-1} è un isomorfismo.*

Dimostrazione. Per semplificare la scrittura poniamo $f^{-1} = g$. Dobbiamo verificare che, se $v', w' \in V'$, $\lambda, \mu \in K$, si ha $g(\lambda v' + \mu w') = \lambda g(v') + \mu g(w')$. Poichè f è biiettiva, esistono unici $v \in V$ e $w \in V$ tali che $f(v) = v'$ e $f(w) = w'$, e dunque $g(v') = v$, $g(w') = w$. Allora

$$g(\lambda v' + \mu w') = g(\lambda f(v) + \mu f(w)) \overset{f \text{ lineare}}{=} g(f(\lambda v + \mu w)) \overset{g=f^{-1}}{=} \lambda v + \mu w = \lambda g(v') + \mu g(w').$$

$\square$

In un isomorfismo due spazi vettoriali si corrispondono mediante una biiezione che conserva anche la loro struttura; se esiste un isomorfismo di V su V', si dice che V e V' sono **isomorfi** e si scrive $V \simeq V'$. Spesso spazi isomorfi vengono addirittura identificati tramite l'isomorfismo f, cioè vengono trattati come se fossero lo stesso spazio.