Note di Algebra Lineare

Nicoletta Cantarini[1]

[1]Liberamente (es)tratto da: Un Corso di Matematica, N. Cantarini, B. Chiarellotto, L. Fiorot, Ed. Progetto, Padova 2006

Indice

Introduzione

Queste note non hanno la pretesa di sostituirsi ad uno dei numerosi testi di Algebra Lineare disponibili in letteratura ma semplicemente di offrire agli studenti del corso di Laurea in Informatica dell'Università di Bologna un supporto nella preparazione dell'esame di Algebra e Geometria.

Descriviamo un paio di problemi che gli studenti saranno in grado di risolvere alla fine del corso.

Problema A. (Problema enigmistico di basso livello)

i) Calcolare le età di due sorelle, età che indicheremo con E_1 ed E_2, sapendo che l'età della prima sommata a 2 volte l'età della seconda è pari a 22 e che 3 volte l'età della prima meno 2 volte l'età della seconda è pari a 2. Risolvere tale problema significa trovare E_1 ed E_2 tali che le due equazioni: $E_1+2E_2 = 22$ e $3E_1-2E_2 = 2$ siano soddisfatte contemporaneamente. Dalla prima equazione si ottiene $E_1 = 22 - 2E_2$ e, sostituendo questa espressione nella seconda equazione, si ottiene $E_2 = 8$ da cui $E_1 = 6$.

Potremmo rendere le cose più complicate facendo entrare in gioco anche l'età di una zia che indichiamo con Z. Allora il quesito potrebbe essere il seguente: calcolare le tre età E_1, E_2, Z, sapendo che l'età della prima sorella meno l'età della seconda meno quella della zia è pari a 2, e che 2 volte l'età della zia meno l'età della prima sorella più l'età della seconda è pari a 4. Allora $Z = 6$, $E_2 = 2$ ed $E_1 = 10$ è una soluzione, ma anche $Z = 6$, $E_2 = 4$ ed $E_1 = 12$ lo è. Quindi tali problemi possono avere diverse soluzioni, ma quante esattamente? Quando possiamo affermare con certezza che il problema ha una sola soluzione, come nel primo caso?

ii) Un secondo esempio di applicazione dei sistemi lineari viene dalla chimica. Supponiamo di voler bilanciare un'equazione chimica. Ad esempio, consideriamo la reazione tra sale comune NaCl e acido sulfureo $\mathrm{H_2SO_4}$:

$$\mathrm{NaCl} + \mathrm{H_2SO_4} \to \mathrm{Na_2SO_4} + \mathrm{HCl}.$$

È immediato vedere che per bilanciare tale equazione si trova

$$2\mathrm{NaCl} + \mathrm{H_2SO_4} \to \mathrm{Na_2SO_4} + 2\mathrm{HCl}$$

Bilanciare un'equazione chimica equivale a richiedere che il numero di atomi di ogni elemento prima di una reazione sia pari al numero di atomi presente dopo la reazione. Quindi per bilanciare l'equazione

$$x\mathrm{C_7H_8} + y\mathrm{HNO_3} \to z\mathrm{C_7H_5O_6N_3} + w\mathrm{H_2O}$$

dovremo risolvere il sistema lineare

$$\begin{cases} 7x = 7z \\ 8x + 1y = 5z + 2w \\ 1y = 3z \\ 3y = 6z + 1w. \end{cases}$$

Problema B. (Evoluzione del sistema)

Supponiamo che in un'isola vi siano volpi in numero pari a V e galline in numero pari a G. Supponiamo sia dato un modello per cui in un anno le galline si riproducono determinando un aumento della popolazione del 60 per cento mentre le volpi si mangiano le galline per un fattore 1 rispetto al loro numero. Come si sarà evoluto il sistema dopo un anno? Il numero di galline, che indichiamo con G_1, sarà pari a $G_1 = 1,6G_0 - V_0$ ovvero al numero iniziale di galline G_0 a cui si è aggiunto il numero di pulcini $0,6G_0$ meno il numero di galline mangiate dalle volpi, pari al numero iniziale di volpi, cioè V_0. D'altro canto supponiamo che il tasso di natalità delle volpi sia del 10 per cento e che le galline abbiano una malattia che si trasmette alle volpi che se le mangiano in modo tale che la mortalità delle volpi a causa di questa malattia sia proporzionale a metà del numero di galline. Questo significa che dopo un anno il numero di volpi V_1 sarà pari a $V_1 = 1,1V_0 - 0,5G_0$ (dove $0,5G_0$ è la quantità di volpi uccise dalla malattia). Cosa potrebbe succedere a questo punto? Se le volpi fossero troppe alla fine si mangerebbero tutte le galline e non resterebbe più nulla da mangiare per le volpi, così nell'isola non vi sarebbe più nessuno. Quante galline ci vogliono e quante volpi occorrono per avere un sistema che non si esaurisca? Oppure, in tale situazione, per ogni scelta iniziale di galline e volpi alla fine l'isola rimarrà deserta? Ovviamente bisognerebbe conoscere a priori l'evoluzione del nostro sistema, cioè sapere a priori quello che avverrà.

Lezione 1

Introduzione ai sistemi lineari

In questa lezione ci proponiamo di risolvere un qualsiasi sistema lineare a coefficienti reali attraverso un metodo noto come algoritmo di Gauss. In seguito useremo questo metodo anche per risolvere problemi diversi dai sistemi lineari e, nello stesso tempo, interpreteremo i sistemi lineari come casi particolari di una teoria molto più ampia.

1.0.1 Sistemi lineari: primi esempi

Una *equazione lineare* è un'equazione in cui le incognite compaiono con grado 1, cioè una equazione della forma:

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = b, \tag{1.1}$$

dove $a_1, a_2, \ldots, a_n$ e b sono numeri assegnati e $x_1, x_2, \ldots, x_n$ sono le incognite. I numeri $a_1, \ldots, a_n$ si dicono *coefficienti* della equazione lineare, b si chiama *termine noto*. Se $b = 0$ l'equazione si dice *omogenea*. Una *soluzione* della equazione (1.1) è una n-upla di numeri $(s_1, s_2, \ldots, s_n)$ che sostituiti ordinatamente alle incognite verificano l'uguaglianza, cioè tali che

$$a_1s_1 + a_2s_2 + \ldots + a_ns_n = b.$$

Ad esempio $(3, -1, 4)$ è una soluzione dell'equazione $2x_1 + 7x_2 - x_3 = -5$ perché $2 \cdot 3 + 7 \cdot (-1) - 4 = -5$.

Un *sistema lineare di m equazioni nelle n incognite* $x_1, x_2, \ldots, x_n$ è un insieme di m equazioni lineari nelle n incognite $x_1, x_2, \ldots, x_n$ che devono essere soddisfatte contemporaneamente:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \qquad\qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \tag{1.2}$$

I numeri $a_{11}, \ldots, a_{1n}, \ldots, a_{m1}, \ldots, a_{mn}$ si chiamano coefficienti del sistema, $b_1, \ldots, b_m$ termini noti. Se $b_i = 0$ per ogni $i = 1, \ldots, m$ il sistema si dice *omogeneo*. Una *soluzione* del sistema lineare (1.2) è una n-upla $(s_1, s_2, \ldots, s_n)$ di numeri che sia soluzione di tutte le equazioni del sistema. Ad esempio $(1, 2)$ è soluzione del sistema lineare

$$\begin{cases} x_1 + x_2 = 3 \\ x_1 - x_2 = -1 \end{cases}$$

In questo corso ci occuperemo esclusivamente di sistemi lineari a **coefficienti reali** cioè di sistemi della forma (1.2) in cui tutti i coefficienti a_{ij} delle incognite e tutti i termini noti b_i sono numeri reali. Le soluzioni che cercheremo, dunque, saranno sempre n-uple (ordinate) di numeri reali.

Dato un sistema lineare, ci prefiggiamo di rispondere alle seguenti domande:

1. Il sistema ammette soluzioni?

2. In caso affermativo, quante e quali sono?

In certi casi rispondere a queste domande è particolarmente facile. Vediamo qualche esempio:

Esempio 1.0.1 Consideriamo il seguente sistema lineare nelle incognite x_1, x_2:

$$\begin{cases} x_1 + x_2 = 3 \\ x_1 + x_2 = 1 \end{cases}$$

È immediato osservare che la somma di due numeri reali non può essere contemporaneamente uguale a 3 e ad 1. Dunque il sistema non ammette soluzioni. In altre parole, le condizioni assegnate dalle due equazioni del sistema sono incompatibili perciò il sistema non può avere soluzioni.

L'esempio appena fatto giustifica la seguente definizione:

Definizione 1.0.2 *Un sistema si dice compatibile se ammette soluzioni.*

Esempio 1.0.3 Consideriamo il seguente sistema lineare nelle incognite x_1, x_2:

$$\begin{cases} x_1 + x_2 = 3 \\ x_2 = -1 \end{cases}$$

Sostituendo nella prima equazione il valore di x_2 fissato dalla seconda equazione, otteniamo: $x_1 = 3 - x_2 = 3 + 1 = 4$. Il sistema è dunque compatibile e ammette un'unica soluzione: $(4, -1)$. In questo esempio vengono assegnate due variabili (le incognite x_1 e x_2) e due condizioni (le due equazioni del sistema). Tali condizioni sono compatibili, cioè non si contraddicono, e sono 'indipendenti' nel senso che non possono essere ottenute una dall'altra. In sintesi: due variabili reali + due condizioni compatibili $\Rightarrow$ 1 sola soluzione.

Esempio 1.0.4 Consideriamo ora il sistema lineare

$$\begin{cases} x_1 + x_2 = 3 \\ 2x_1 + 2x_2 = 6 \end{cases}$$

nelle incognite x_1, x_2. Diversamente da quanto succedeva nell'esempio precedente, qui le condizioni assegnate dalle due equazioni non sono 'indipendenti' nel senso che la seconda equazione si ottiene moltiplicando la prima per 2. Le due equazioni stabiliscono dunque la stessa relazione tra le variabili x_1 e x_2, quindi risolvere il sistema lineare assegnato equivale a risolvere semplicemente l'equazione $x_1 + x_2 = 3$. Tale equazione ha certamente soluzioni: ad esempio abbiamo visto nell'esempio precedente che $(4, -1)$ è soluzione dell'equazione, ma certamente anche $(1, 2)$ o $(0, 3)$ sono soluzioni dell'equazione. Quante sono allora esattamente le soluzioni di questa equazione? E come sono fatte? In questo caso abbiamo due variabili ed una sola condizione su di esse. Questo significa che una variabile è libera e siccome varia nell'insieme dei numeri reali, che sono infiniti, essa può assumere infiniti valori diversi. L'equazione assegnata ci permette di esprimere una variabile in funzione della variabile libera. Le soluzioni dell'equazione saranno tutte e sole della forma: $(x_1, 3 - x_1)$. Con questa scrittura si intende che la variabile x_1 può assumere tutti gli infiniti valori reali e che, affinché l'equazione $x_1 + x_2 = 3$ sia soddisfatta, deve essere $x_2 = 3 - x_1$. Naturalmente potevamo decidere di far variare liberamente la variabile x_2 e di esprimere x_1 in funzione di x_2. In tal caso avremmo descritto ogni soluzione del sistema nella forma $(3-x_2, x_2)$. Il sistema assegnato ha dunque infinite soluzioni. In sintesi: due variabili reali + 1 sola condizione $\Rightarrow$ infinite soluzioni.

Definizione 1.0.5 *Due sistemi lineari si dicono equivalenti se hanno le stesse soluzioni.*

Nell'Esempio 1.0.4 abbiamo osservato che il sistema lineare

$$\begin{cases} x_1 + x_2 = 3 \\ 2x_1 + 2x_2 = 6 \end{cases}$$

è equivalente all'equazione $x_1 + x_2 = 3$. Naturalmente riuscire a capire se due sistemi sono equivalenti può essere molto utile, per esempio potremmo tentare di risolvere un sistema lineare riducendolo ad uno ad esso equivalente ma più semplice da risolvere.

Nel prossimo paragrafo introdurremo alcune nozioni utili per semplificare la scrittura di un sistema lineare.

1.1 Matrici

Dati due numeri naturali m, n, si chiama *matrice* $m \times n$ a coefficienti reali una tabella di mn numeri reali collocati su m righe e n colonne. Ad esempio:

$$\begin{pmatrix} 5 & -6 & 0 \\ 4 & 3 & -1 \end{pmatrix}$$

è una matrice 2×3.

Se $m = n$ la matrice si dice *quadrata* di ordine n. Ad esempio:

$$\begin{pmatrix} 1 & 0 \\ \frac{2}{3} & 3 \end{pmatrix}$$

è una matrice quadrata di ordine 2.

Indicheremo con $\mathrm{M}_{m,n}(\mathbb{R})$ l'insieme delle matrici $m \times n$ a coefficienti reali e semplicemente con $\mathrm{M}_n(\mathbb{R})$ l'insieme delle matrici quadrate di ordine n a coefficienti reali.

Data una matrice A, il numero reale che compare nella i-esima riga e nella j-esima colonna di A viene detto *elemento di posto* (i, j) di A. Ad esempio nella matrice

$$A = \begin{pmatrix} 5 & -6 & 0 \\ 4 & 3 & -1 \end{pmatrix}$$

l'elemento di posto $(1, 3)$ è 0 e l'elemento di posto $(2, 2)$ è 3. Naturalmente due matrici $m \times n$ A e B sono uguali se coincidono entrata per entrata, cioè se l'elemento di posto (i, j) di A coincide con l'elemento di posto (i, j) di B, per ogni $i = 1, \dots, m$ e per ogni $j = 1, \dots, n$.

Data una generica matrice $m \times n$:

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

essa può essere sinteticamente indicata nel modo seguente: $A = (a_{ij})$ dove a_{ij} è l'elemento di posto (i, j), $i = 1, \dots, m$, $j = 1, \dots, n$.

Se $A = (a_{ij})$ è una matrice $m \times n$ e $B = \begin{pmatrix} b_{11} \\ \vdots \\ b_{n1} \end{pmatrix}$ una matrice $n \times 1$, cioè una matrice costituita da una sola colonna, le matrici A e B possono essere moltiplicate. Il risultato è una matrice C con m righe ed una sola colonna: $C = (c_{i1})$, in cui l'elemento di posto $(i, 1)$ si ottiene nel modo seguente: si fissa la i-esima riga di A

$$(a_{i1} \dots a_{in})$$

e si moltiplicano, nell'ordine, le sue entrate per le entrate dell'unica colonna di B, dopodiché si sommano i numeri ottenuti. Si ha, cioè:

$$c_{i1} = a_{i1}b_{11} + a_{i2}b_{21} + \dots + a_{in}b_{n1} = \sum_{h=1}^{n} a_{ih}b_{h1}.$$

Ad esempio, se

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

si ha

$$c_{11} = 1 \cdot 0 + 0 \cdot (-3) + 3 \cdot 1 + (-1) \cdot 2 = 1$$

$$c_{21} = 0 \cdot 0 + (-2) \cdot (-3) + 2 \cdot 1 + 1 \cdot 2 = 10$$

$$c_{31} = 1 \cdot 0 + 0 \cdot (-3) + (-1) \cdot 1 + 0 \cdot 2 = -1$$

cioè:

$$C = \begin{pmatrix} 1 & 0 & 3 & -1 \\ 0 & -2 & 2 & 1 \\ 1 & 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ -3 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 10 \\ -1 \end{pmatrix}$$

Vedremo in seguito che il prodotto appena definito è solo un caso particolare del cosiddetto prodotto *righe per colonne* di due matrici. Ci limitiamo per ora a questa definizione 'ridotta' perché è quella di cui ci serviremo nella trattazione dei sistemi lineari. Infatti un sistema lineare della forma:

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

può essere pensato in forma matriciale:

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

e quindi può essere riscritto utilizzando il prodotto righe per colonne nel modo seguente:

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

o, più sinteticamente, come

$$A\underline{x} = \underline{b},$$

dove $A = (a_{ij})$ è la matrice $m \times n$ dei coefficienti delle incognite, $\underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ è la colonna delle n incognite, e $\underline{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$ è la colonna degli m termini noti. La matrice $A = (a_{ij})$ si chiama matrice *incompleta* associata al sistema e la matrice

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

si chiama matrice *completa* associata al sistema.

Esempio 1.1.1 Consideriamo il sistema lineare

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

nelle incognite x_1, x_2, x_3. Allora la matrice incompleta e la matrice completa associate al sistema sono, rispettivamente:

$$A = \begin{pmatrix} 2 & \sqrt{2} & -1 \\ 1 & 0 & -1 \end{pmatrix} \text{ e } (A|\underline{b}) = \left(\begin{array}{ccc|c} 2 & \sqrt{2} & -1 & 2 \\ 1 & 0 & -1 & 1 \end{array}\right).$$

Usare le matrici è semplicemente un modo più comodo di scrivere e trattare i sistemi lineari. Ogni riga della matrice completa associata ad un sistema lineare equivale ad una equazione del sistema in cui vengono sottintese le incognite.

Definizione 1.1.2 *Una matrice si dice in forma a scala (per righe) o, semplicemente, a scala se sono soddisfatte le seguenti condizioni:*

(a) eventuali righe nulle si trovano in fondo alla matrice;

(b) il primo elemento non nullo di ogni riga (non nulla) si trova più a destra del primo elemento non nullo della riga precedente.

Esempio 1.1.3 La matrice

$$A = \begin{pmatrix} 1 & -1 & -1 & 2 & -4 \\ 0 & 0 & -1 & 3 & 5 \\ 0 & 0 & 0 & \frac{1}{3} & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

è una matrice in forma a scala (per righe) perché soddisfa le condizioni (a) e (b) della Definizione 1.1.2.

Al contrario, la matrice

$$B = \begin{pmatrix} 2 & -1 & -1 & 2 & -4 \\ 0 & 1 & -1 & 3 & 5 \\ 0 & 2 & 0 & 1 & \frac{1}{5} \end{pmatrix}$$

non è in forma a scala perché il primo elemento non nullo della terza riga non si trova più a destra del primo elemento non nullo della seconda riga (ma sotto di esso).

Definizione 1.1.4 *Sia A una matrice in forma a scala (per righe). Si chiama pivot di A il primo elemento non nullo di ogni riga (non nulla) di A. Si chiama rango di A e si indica con $rg(A)$ il numero di righe non nulle di A o, equivalentemente, il numero dei suoi pivot.*

Esempio 1.1.5 Data

$$A = \begin{pmatrix} 1 & -1 & -1 & 2 & -4 \\ 0 & 0 & -1 & 3 & 5 \\ 0 & 0 & 0 & \frac{1}{3} & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

i pivot di A sono 1, -1, $\frac{1}{3}$, perciò $rg(A) = 3$.

Osservazione 1.1.6 Sia $A \in \mathrm{M}_{m,n}(\mathbb{R})$ una matrice a scala. Per definizione di rango si ha

$$rg(A) \leq m. \tag{1.3}$$

Vale però anche la disuguaglianza

$$rg(A) \leq n. \tag{1.4}$$

Se $m \leq n$ (1.4) segue ovviamente da (1.3). Se $m > n$ è facile rendersi conto, disegnando una matrice a scala con un numero di righe m maggiore del numero n di colonne, che la proprietà (b) della Definizione 1.1.2 implica che il massimo numero di pivot di A è n.

Definizione 1.1.7 *Il sistema lineare $A\underline{x} = \underline{b}$ si dice a scala se la matrice $(A|\underline{b})$ è in forma a scala.*

Mostreremo ora come risolvere velocemente un sistema lineare a scala.

Esempio 1.1.8 Consideriamo il seguente sistema lineare nelle incognite x_1, x_2, x_3, x_4:

$$\begin{cases} 4x_1 + 2x_2 + 3x_3 + 4x_4 = 1 \\ x_2 - 2x_3 = 2 \\ x_3 - x_4 = 0 \\ x_4 = 1 \end{cases}$$

La matrice completa associata al sistema è:

$$(A|\underline{b}) = \left(\begin{array}{rrrr|r} 4 & 2 & 3 & 4 & 1 \\ 0 & 1 & -2 & 0 & 2 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right),$$

che è in forma a scala e ha rango 4. Ovviamente anche la matrice incompleta A è in forma a scala e notiamo che anch'essa ha rango 4. Il fatto che la matrice A sia in forma a scala indica che in ogni equazione del sistema compare una incognita che non compare nelle equazioni successive. Il sistema lineare può dunque essere facilmente risolto per sostituzioni successive dal basso, cioè a partire dall'ultima equazione e risalendo verso la prima: dalla quarta equazione abbiamo $x_4 = 1$; sostituendo $x_4 = 1$ nella terza equazione otteniamo $x_3 = x_4 = 1$. Sostituendo $x_3 = 1$ nella seconda equazione otteniamo $x_2 = 2+2x_3 = 2+2 = 4$. Infine, sostituendo $x_2 = 4$ e $x_3 = x_4 = 1$ nella prima equazione, otteniamo $x_1 = \frac{1}{4}(1 - 2x_2 - 3x_3 - 4x_4) = \frac{1}{4}(1 - 8 - 3 - 4) = -\frac{7}{2}$. Il sistema assegnato ha dunque una sola soluzione: $(-\frac{7}{2}, 4, 1, 1)$.

Esempio 1.1.9 Consideriamo il sistema lineare nelle incognite x_1, x_2, x_3, x_4 ottenuto da quello dell'esempio precedente cancellando l'ultima equazione:

$$\begin{cases} 4x_1 + 2x_2 + 3x_3 + 4x_4 = 1 \\ x_2 - 2x_3 = 2 \\ x_3 - x_4 = 0 \end{cases}$$

La matrice completa associata al sistema è:

$$(A|\underline{b}) = \left(\begin{array}{rrrr|r} 4 & 2 & 3 & 4 & 1 \\ 0 & 1 & -2 & 0 & 2 \\ 0 & 0 & 1 & -1 & 0 \end{array}\right),$$

che è in forma a scala e ha rango 3. Anche la matrice incompleta A è in forma a scala e anch'essa ha rango 3. Naturalmente la soluzione $(-\frac{7}{2}, 4, 1, 1)$, trovata nell'esempio precedente, continua ad essere una soluzione del sistema che quindi è senz'altro compatibile. Quante sono, tuttavia, in questo caso le soluzioni del sistema? Anche in questo caso possiamo procedere per sostituzioni successive dal basso perché, come prima, in ogni equazione compare una incognita che non compare nelle equazioni successive. Dall'ultima equazione abbiamo $x_3 = x_4$. Sostituendo $x_3 = x_4$ nella seconda equazione, otteniamo $x_2 = 2+2x_3 = 2+2x_4$. Sostituendo x_2 e x_3 nella prima equazione otteniamo $x_1 = \frac{1}{4}(1-2x_2-3x_3-4x_4) = \frac{1}{4}(1-4-4x_4-3x_4-4x_4) = \frac{1}{4}(-3-11x_4)$. Il sistema ha dunque infinite soluzioni della forma $(\frac{1}{4}(-3-11x_4), 2+2x_4, x_4, x_4)$ al variare di x_4 nell'insieme dei numeri reali.

Quanto illustrato negli esempi 1.1.8, 1.1.9 è un fatto del tutto generale. Vale infatti la seguente proposizione:

Proposizione 1.1.10 *Sia $A\underline{x} = \underline{b}$ un sistema lineare a scala nelle n incognite $x_1, \ldots, x_n$. Allora:*

(a) il sistema ammette soluzioni se e solo se $rg(A) = rg(A|\underline{b})$;

(b) se $rg(A) = rg(A|\underline{b}) = n$ il sistema ammette una sola soluzione;

(c) se $rg(A) = rg(A|\underline{b}) < n$ il sistema ammette infinite soluzioni.

Dimostrazione. Osserviamo innanzitutto che cancellando la colonna $\underline{b}$ dalla matrice $(A|\underline{b})$ si ottiene ancora una matrice in forma a scala, quindi anche la matrice incompleta A associata al sistema è una matrice a scala. Inoltre, cancellando la colonna $(\underline{b})$ dalla matrice $(A|\underline{b})$, il numero di pivot può diminuire al più di 1. Più precisamente questo succede se e soltanto se la matrice A ha almeno una riga nulla, diciamo la i-esima, e l'elemento b_i è diverso da 0. In termini di equazioni questo equivale alla condizione $0 = b_i \neq 0$ che, evidentemente, non può essere soddisfatta. Pertanto se $rg(A) \neq rg(A|\underline{b})$, il sistema non ammette soluzioni. Supponiamo ora che sia $rg(A) = rg(A|\underline{b}) = n$. Questo significa che il numero dei pivot, ossia il numero degli 'scalini', coincide con il numero delle incognite, quindi il sistema è costituito esattamente da n equazioni, l'incognita x_1 compare solo nella prima equazione, x_2 solo nelle prime due equazioni, x_3 solo nelle prime tre e così via. In particolare l'ultima equazione del sistema contiene solo l' incognita x_n e quindi ne fissa il valore. Sostituendo tale valore nella penultima equazione si ottiene l'unico valore della variabile x_{n-1} e così via, procedendo per sostituzioni successive dal basso come nell'Esempio 1.1.8, si ottiene l'unica soluzione del sistema. Se, invece, $rg(A) = rg(A|\underline{b}) = k < n$ è possibile, procedendo per sostituzioni successive dal basso, esprimere k variabili in funzione delle altre $n - k$ che restano libere di variare nell'insieme dei numeri reali. Si ottengono così infinite soluzioni. $\square$

Esempio 1.1.11 Risolviamo il seguente sistema lineare a scala nelle incognite x_1, x_2, x_3, x_4:

$$\begin{cases} x_1 - x_2 + x_3 - x_4 = 1 \\ x_3 + \frac{1}{2}x_4 = 0 \end{cases}$$

La matrice completa associata al sistema è:

$$(A|\underline{b}) = \left(\begin{array}{cccc|c} 1 & -1 & 1 & -1 & 1 \\ 0 & 0 & 1 & \frac{1}{2} & 0 \end{array}\right).$$

Notiamo che $rg(A) = rg(A|\underline{b}) = 2$ quindi, per la Proposizione 1.1.10(a), il sistema ammette soluzioni. Dal momento che il numero delle variabili è $4 > 2$, per la Proposizione 1.1.10(c), il sistema ammette infinite soluzioni. In sostanza abbiamo 4 variabili e due condizioni su di esse, perciò due variabili restano libere di variare nell'insieme dei numeri reali e potremo esprimere due variabili in funzione delle altre due. Procedendo per sostituzioni successive dal basso abbiamo:

$$x_3 = -\frac{1}{2}x_4$$

$$x_1 = x_2 - x_3 + x_4 + 1 = x_2 + \frac{1}{2}x_4 + x_4 + 1 = x_2 + \frac{3}{2}x_4 + 1.$$

Le infinite soluzioni del sistema sono pertanto della forma $(x_2 + \frac{3}{2}x_4 + 1,\, x_2,\, -\frac{1}{2}x_4,\, x_4)$, con $x_2, x_4 \in \mathbb{R}$.

Osserviamo che avremmo potuto decidere di esprimere la variabile x_4 in funzione della variabile x_3 ($x_4 = -2x_3$) e, ad esempio, la variabile x_2 in funzione delle variabili x_1 e x_3 ($x_2 = x_1 + 3x_3 - 1$). In altri termini, la scelta delle variabili 'libere' non è obbligata. Tuttavia è sempre possibile scegliere come variabili libere quelle corrispondenti alle colonne della matrice A non contenenti pivot ed esprimere in funzione di queste le incognite corrispondenti alle colonne contenenti i pivot. Ad esempio, in questo caso i pivot, entrambi uguali ad 1, si trovano sulla prima e sulla terza colonna di A e nella nostra prima scelta noi abbiamo lasciato libere le variabili x_2 e x_4 ed espresso x_1 e x_3 in funzione di x_2 e x_4.

1.2 Algoritmo di Gauss

Abbiamo stabilito come risolvere un sistema lineare a scala. Cosa succede nel caso di un qualsiasi sistema lineare $A\underline{x} = \underline{b}$? Sarebbe comodo poter ottenere un nuovo sistema lineare $A'\underline{x} = \underline{b}'$, questa volta a scala, equivalente al sistema di partenza, in modo tale da poter calcolare le soluzioni di $A\underline{x} = \underline{b}$ risolvendo il sistema a scala $A'\underline{x} = \underline{b}'$. Questo è esattamente quello che faremo.

Esempio 1.2.1 I seguenti sistemi lineari nelle incognite x_1, x_2, sono equivalenti:

$$\begin{cases} x_1 - x_2 = 1 \\ x_1 + x_2 = 2 \end{cases} \qquad \begin{cases} x_1 - x_2 = 1 \\ 2x_2 = 1 \end{cases}$$

Si verifica, infatti, facilmente che l'unica soluzione di ciascuno dei due sistemi è: $(\frac{3}{2}, \frac{1}{2})$. Notiamo che la prima equazione è la stessa nei due sistemi e che il secondo sistema può essere ottenuto sostituendo la seconda equazione del primo con la differenza tra questa e la prima equazione:

$$2^a\text{equazione} \to 2^a\text{equazione} - 1^a\text{equazione}.$$

Come si può passare da un sistema ad uno ad esso equivalente? Per esempio eseguendo le seguenti operazioni:

(a) scambio di due equazioni;

(b) moltiplicazione di un'equazione per un numero reale diverso da 0;

(c) sostituzione della equazione i-esima con la somma dell'equazione i-esima e della j-esima moltiplicata per un numero reale α qualsiasi. In sintesi:

$$i\text{-esima equazione} \longrightarrow i\text{-esima equazione}+\alpha(j\text{-esima equazione}).$$

È immediato verificare che le operazioni (a) e (b) non alterano le soluzioni del sistema. In quanto alla operazione (c) basta osservare che essa coinvolge soltanto la i-esima e la j-esima equazione del sistema, quindi basta osservare che i sistemi

$$\begin{cases} j\text{-esima equazione} \\ i\text{-esima equazione} \end{cases} \qquad \begin{cases} j\text{-esima equazione} \\ i\text{-esima equazione} + \alpha(j\text{-esima equazione}) \end{cases}$$

sono equivalenti.

Traduciamo ora le operazioni (a), (b) e (c) in termini matriciali: scambiare due equazioni del sistema equivale a scambiare due righe della matrice completa associata al sistema; moltiplicare una equazione per un numero reale diverso da 0 equivale a moltiplicare una riga della matrice completa associata al sistema per un numero reale diverso da 0, cioè moltiplicare per tale numero ogni elemento della riga; infine l'operazione (c) equivale a sostituire la riga i-esima della matrice completa associata al sistema con la somma della riga i-esima e della j-esima moltiplicata per un numero reale α. Spieghiamo un po' meglio che cosa si intende con tale somma: siano $(a_{i1} \, \dots \, a_{in} \, b_i)$ e $(a_{j1} \, \dots \, a_{jn} \, b_j)$ rispettivamente la i-esima e la j-esima riga della matrice $(A|\underline{b})$. Sommare la i-esima riga con la j-esima moltiplicata per un numero α, significa effettuare la somma entrata per entrata:

$$(a_{i1} \, \dots \, a_{in} \; b_i) + \alpha(a_{j1} \, \dots \, a_{jn} \; b_j) = (a_{i1} + \alpha a_{j1} \, \dots \, a_{in} + \alpha a_{jn} \; b_i + \alpha b_j).$$

In virtù dell'importanza che tali operazioni avranno in seguito, diamo ad esse un nome:

Definizione 1.2.2 *Data una matrice A si chiamano operazioni elementari sulle righe di A le seguenti operazioni:*

(a) scambio di due righe;

(b) moltiplicazione di una riga per un numero reale diverso da 0;

(c) sostituzione della riga i-esima con la somma della riga i-esima e della j-esima moltiplicata per un numero reale α qualsiasi.

Osservazione 1.2.3 Osserviamo che nella operazione elementare (c) non richiediamo che il numero α sia non nullo. In effetti se $\alpha = 0$ l'operazione (c) equivale a lasciare la riga i-esima inalterata.

Data una qualsiasi matrice $A = (a_{ij})$ è possibile trasformare A in una matrice a scala attraverso operazioni elementari sulle righe di A. Tale procedimento è noto come *riduzione di Gauss* e l'algoritmo che si utilizza si chiama **Algoritmo di Gauss** e funziona nel modo seguente:

1. Se $a_{11} = 0$ si scambia la prima riga di A con una riga in cui il primo elemento è non nullo. Indichiamo con a tale elemento non nullo. Se il primo elemento di ogni riga di A è nullo, si va direttamente al punto 3.

2. Si controllano una dopo l'altra tutte le righe tranne la prima. Se il primo elemento di una riga è nullo si lascia quella riga inalterata. Se il primo elemento di una riga, diciamo la i-esima ($i > 1$), è uguale a $b \neq 0$, si sostituisce la riga i-esima con la somma della riga i-esima e della prima riga moltiplicata per $-\frac{b}{a}$.

3. A questo punto tutti gli elementi della prima colonna, tranne eventualmente il primo, sono nulli. Si considera dunque la matrice che si ottiene cancellando la prima riga e la prima colonna della matrice ottenuta e si ricomincia dal punto 1.

Esempio 1.2.4 Sia

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

Utilizziamo l'algoritmo di Gauss per ridurre A a scala.

Dal momento che l'elemento di posto $(1,1)$ è nullo, scambiamo la prima con la seconda riga, ottenendo la matrice:

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

Il primo elemento della seconda riga della matrice ottenuta è 0, perciò lasciamo questa riga inalterata. Il primo elemento della terza riga, invece, è 2, quindi sostituiamo la terza riga con la somma della terza riga e della prima moltiplicata per -2. Otteniamo così la matrice:

$$\begin{array}{c} -2 \\ \\ \\ \end{array} \begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & 1 & -1 & 0 \\ 2 & -1 & 1 & 2 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & -5 & 1 & 0 \end{pmatrix}.$$

Ora ogni elemento della prima colonna tranne il primo è uguale a 0. Passiamo a considerare la matrice che otteniamo cancellando la prima riga e la prima colonna della matrice ottenuta:

$$\begin{pmatrix} 1 & -1 & 0 \\ -5 & 1 & 0 \end{pmatrix}.$$

Applichiamo un'altra volta l'algoritmo di Gauss: il primo elemento della prima riga, questa volta, è diverso da 0, perciò lasciamo inalterata la prima riga. Ora sostituiamo la seconda riga con la somma della seconda e della prima moltiplicata per 5, ottenendo:

$$\begin{array}{c} 5 \\ \\ \end{array} \begin{pmatrix} 1 & -1 & 0 \\ -5 & 1 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & -1 & 0 \\ 0 & -4 & 0 \end{pmatrix}.$$

Abbiamo così ottenuto la matrice a scala:

$$B = \begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & -4 & 0 \end{pmatrix}.$$

A questo punto siamo in grado di risolvere qualsiasi sistema lineare $A\underline{x} = \underline{b}$. La matrice completa associata al sistema è $(A|\underline{b})$. Utilizzando l'algoritmo di Gauss possiamo 'ridurre' $(A|\underline{b})$ a scala ottenendo una matrice $(A'|\underline{b}')$. Il sistema lineare $A'\underline{x} = \underline{b}'$ è equivalente al sistema $A\underline{x} = \underline{b}$ dal momento che ogni operazione elementare sulle righe di $(A|\underline{b})$ equivale ad una operazione sulle equazioni del sistema che ne preserva le soluzioni. Quindi per trovare le soluzioni del sistema di partenza risolveremo il sistema a scala $A'\underline{x} = \underline{b}'$, tenendo conto della Proposizione 1.1.10. Notiamo in particolare che, in conseguenza del ragionamento appena illustrato e del contenuto della Proposizione 1.1.10, dato un qualsiasi sistema lineare $A\underline{x} = \underline{b}$ a coefficienti reali **SOLO UNA** delle seguenti situazioni si può presentare:

1. il sistema NON ha soluzioni;
2. il sistema ha UNA SOLA soluzione;
3. il sistema ha INFINITE soluzioni.

Questo significa che non esiste alcun sistema lineare a coefficienti reali con un numero finito di soluzioni strettamente più grande di 1. Nel momento in cui un sistema lineare a coefficienti reali ha 2 soluzioni allora ne ha infinite.

Osservazione 1.2.5 Le mosse dell'algoritmo di Gauss non sono necessariamente obbligate. Nell'Esempio 1.2.4, ad esempio, anziché scambiare la prima con la seconda riga, avremmo potuto scambiare la prima con la terza riga. In questo modo, portando a termine l'algoritmo, avremmo ottenuto una matrice a scala diversa dalla matrice B. Dal punto di vista dei sistemi lineari questo significa semplicemente ottenere sistemi a scala diversi ma tutti equivalenti al sistema di partenza (e quindi equivalenti tra loro).

Esempio 1.2.6 Risolvere il seguente sistema lineare di quattro equazioni nelle cinque incognite u, v, w, x, y:

$$\begin{cases} u + 2v + 3w + x + y = 4 \\ u + 2v + 3w + 2x + 3y = -2 \\ u + v + w + x + y = -2 \\ -3u - 5v - 7w - 4x - 5y = 0 \end{cases}$$

La matrice completa associata al sistema è:

$$(A|\underline{b}) = \left(\begin{array}{rrrrr|r} 1 & 2 & 3 & 1 & 1 & 4 \\ 1 & 2 & 3 & 2 & 3 & -2 \\ 1 & 1 & 1 & 1 & 1 & -2 \\ -3 & -5 & -7 & -4 & -5 & 0 \end{array}\right).$$

Si tratta ora di ridurre la matrice $(A|\underline{b})$ in forma a scala utilizzando l'algoritmo di Gauss e, successivamente, di risolvere il sistema lineare associato alla matrice ridotta.

In questo primo esempio riportiamo i passi dell'algoritmo di Gauss descrivendo contemporaneamente le operazioni sulle equazioni del sistema che equivalgono ad ogni passo. Naturalmente il vantaggio dell'algoritmo di Gauss consiste proprio nel poter dimenticare equazioni ed incognite concentrandosi solo sulle matrici, quindi questa descrizione è puramente esplicativa.

L'elemento di posto $(1, 1)$ è non nullo perciò lasciamo la prima riga inalterata. Dopodiché effettuiamo le seguenti operazioni elementari sulle righe di $(A|\underline{b})$:

- 2^a riga $\to$ 2^a riga - 1^a riga;

- 3^a riga $\to$ 3^a riga - 1^a riga;

- 4^a riga $\to$ 4^a riga + $3(1^a$ riga).

Otteniamo così la seguente matrice (e l'equivalente sistema lineare):

$$\left(\begin{array}{rrrrr|r} 1 & 2 & 3 & 1 & 1 & 4 \\ 0 & 0 & 0 & 1 & 2 & -6 \\ 0 & -1 & -2 & 0 & 0 & -6 \\ 0 & 1 & 2 & -1 & -2 & 12 \end{array}\right) \Leftrightarrow \begin{cases} u + 2v + 3w + x + y = 4 \\ x + 2y = -6 \\ -v - 2w = -6 \\ v + 2w - x - 2y = 12 \end{cases}$$

Ora scambiamo la seconda con la quarta riga:

$$\left(\begin{array}{rrrrr|r} 1 & 2 & 3 & 1 & 1 & 4 \\ 0 & 1 & 2 & -1 & -2 & 12 \\ 0 & -1 & -2 & 0 & 0 & -6 \\ 0 & 0 & 0 & 1 & 2 & -6 \end{array}\right) \Leftrightarrow \begin{cases} u + 2v + 3w + x + y = 4 \\ v + 2w - x - 2y = 12 \\ -v - 2w = -6 \\ x + 2y = -6 \end{cases}$$

Ora sostituiamo alla terza riga la somma della terza riga e della seconda:

$$\left(\begin{array}{rrrrr|r} 1 & 2 & 3 & 1 & 1 & 4 \\ 0 & 1 & 2 & -1 & -2 & 12 \\ 0 & 0 & 0 & -1 & -2 & 6 \\ 0 & 0 & 0 & 1 & 2 & -6 \end{array}\right) \Leftrightarrow \begin{cases} u + 2v + 3w + x + y = 4 \\ v + 2w - x - 2y = 12 \\ -x - 2y = 6 \\ x + 2y = -6 \end{cases}$$

Infine sostituiamo alla quarta riga la somma della quarta riga e della terza:

$$\left(\begin{array}{rrrrr|r} 1 & 2 & 3 & 1 & 1 & 4 \\ 0 & 1 & 2 & -1 & -2 & 12 \\ 0 & 0 & 0 & -1 & -2 & 6 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \Leftrightarrow \begin{cases} u + 2v + 3w + x + y = 4 \\ v + 2w - x - 2y = 12 \\ -x - 2y = 6 \\ 0 = 0 \end{cases}$$

Il sistema di partenza è equivalente al sistema a scala che abbiamo ottenuto, in cui l'ultima equazione è diventata un'identità. Il rango della matrice incompleta e il rango della matrice completa della matrice a scala ottenuta coincidono e sono uguali a 3. Il numero delle incognite del sistema è 5, quindi il sistema ammette infinite soluzioni che dipenderanno da $5 - 3 = 2$ variabili libere. Risolviamo il sistema per sostituzioni successive dal basso: usando la terza equazione possiamo esprimere la variabile x in funzione di y:

$$x = -2y - 6.$$

Nella seconda equazione sostituiamo x con la sua espressione in funzione di y e ricaviamo v in funzione di w e di y:

$$v = -2w + 6.$$

Infine nella prima equazione sostituiamo x con la sua espressione in funzione di y, v con la sua espressione in funzione di w e ricaviamo u in funzione di w e di y:

$$u = -2v - 3w - x - y + 4 = -2(-2w + 6) - 3w - (-2y - 6) - y + 4 = w + y - 2.$$

Dunque il sistema ha infinite soluzioni del tipo $(w + y - 2, -2w + 6, w, -2y - 6, y)$ che dipendono da due variabili libere $w, y \in \mathbb{R}$.

1.3 Esercizi svolti

Esercizio 1.3.1 Risolvere il seguente sistema lineare nelle quattro incognite x, y, z, t:

$$\begin{cases} x - 2y = 5 \\ -x + 2y - 3z = -2 \\ -2y + 3z - 4t = -11 \\ -3z + 4t = 15 \end{cases}$$

Svolgimento. La matrice completa associata al sistema è:

$$(A|\underline{b}) = \left(\begin{array}{rrrr|r} 1 & -2 & 0 & 0 & 5 \\ -1 & 2 & -3 & 0 & -2 \\ 0 & -2 & 3 & -4 & -11 \\ 0 & 0 & -3 & 4 & 15 \end{array}\right).$$

Riduciamo la matrice $(A|\underline{b})$ in forma a scala utilizzando l'algoritmo di Gauss:

$$\left(\begin{array}{rrrr|r} 1 & -2 & 0 & 0 & 5 \\ 0 & 0 & -3 & 0 & 3 \\ 0 & -2 & 3 & -4 & -11 \\ 0 & 0 & -3 & 4 & 15 \end{array}\right) \to \left(\begin{array}{rrrr|r} 1 & -2 & 0 & 0 & 5 \\ 0 & -2 & 3 & -4 & -11 \\ 0 & 0 & -3 & 0 & 3 \\ 0 & 0 & -3 & 4 & 15 \end{array}\right)$$

$$\to \left(\begin{array}{rrrr|r} 1 & -2 & 0 & 0 & 5 \\ 0 & -2 & 3 & -4 & -11 \\ 0 & 0 & -3 & 0 & 3 \\ 0 & 0 & 0 & 4 & 12 \end{array}\right) = (A'|\underline{b}').$$

La matrice a scala ottenuta è la matrice completa associata al sistema lineare:

$$\begin{cases} x - 2y = 5 \\ -2y + 3z - 4t = -11 \\ -3z = 3 \\ 4t = 12 \end{cases}$$

Notiamo che $rg(A') = rg(A'|\underline{b}') = 4 =$ numero incognite. Il sistema di partenza ammette dunque un'unica soluzione che possiamo calcolare procedendo per sostituzioni successive dal basso: dalla quarta equazione abbiamo

$$t = 3;$$

e dalla terza equazione abbiamo

$$z = -1;$$

sostituendo questi valori di t e di z nella seconda equazione otteniamo

$$y = -2;$$

infine, sostituendo i valori di t, z, y nella prima equazione otteniamo

$$x = 1.$$

Dunque il sistema ha come unica soluzione la quaterna $(1, -2, -1, 3)$.

Esercizio 1.3.2 Risolvere il seguente sistema lineare nelle incognite x, y, z, t, al variare del parametro reale α:

$$\begin{cases} x + y + z + t = 0 \\ x - z - t = -1 \\ x + 2y + (2\alpha + 1)z + 3t = 2\alpha - 1 \\ 3x + 4y + (3\alpha + 2)z + (\alpha + 5)t = 3\alpha - 1. \end{cases}$$

Svolgimento. In questo esercizio si ha a che fare con un sistema lineare in cui compare un parametro reale α. Questo significa che al variare di $\alpha \in \mathbb{R}$ si ottengono infiniti sistemi lineari diversi che noi risolveremo trattandoli il più possibile come un solo sistema lineare. Il modo di procedere resta sempre lo stesso, come se il parametro fosse un numero reale fissato. Per prima cosa, dunque, scriviamo la matrice completa associata al sistema:

$$(A|\underline{b}) = \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 0 \\ 1 & 0 & -1 & -1 & -1 \\ 1 & 2 & 2\alpha + 1 & 3 & 2\alpha - 1 \\ 3 & 4 & 3\alpha + 2 & \alpha + 5 & 3\alpha - 1 \end{array}\right)$$

e riduciamola in forma a scala mediante l'algoritmo di Gauss:

$$\left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 0 \\ 0 & -1 & -2 & -2 & -1 \\ 0 & 1 & 2\alpha & 2 & 2\alpha - 1 \\ 0 & 1 & 3\alpha - 1 & \alpha + 2 & 3\alpha - 1 \end{array}\right) \to \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 0 \\ 0 & -1 & -2 & -2 & -1 \\ 0 & 0 & 2\alpha - 2 & 0 & 2\alpha - 2 \\ 0 & 0 & 3\alpha - 3 & \alpha & 3\alpha - 2 \end{array}\right) \to$$

$$\left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 0 \\ 0 & -1 & -2 & -2 & -1 \\ 0 & 0 & \alpha-1 & 0 & \alpha-1 \\ 0 & 0 & 3\alpha-3 & \alpha & 3\alpha-2 \end{array}\right) \rightarrow \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 0 \\ 0 & -1 & -2 & -2 & -1 \\ 0 & 0 & \alpha-1 & 0 & \alpha-1 \\ 0 & 0 & 0 & \alpha & 1 \end{array}\right) = (A'|\underline{b}')$$

Dobbiamo ora stabilire cosa succede al variare del parametro α nell'insieme dei numeri reali. Dobbiamo cioè rispondere alle seguenti domande:

1. Per quali valori di α il sistema è compatibile?
2. Per i valori di α per cui il sistema è compatibile, quante soluzioni ammette e quali sono?

Come sappiamo la risposta viene fornita dalla Proposizione 1.1.10(a): dobbiamo confrontare il rango di A' con il rango di $(A'|\underline{b}')$. Notiamo che questi ranghi dipendono dal valore di α. Più precisamente: $rg(A') = rg(A'|\underline{b}') = 4$ per $\alpha \neq 0, 1$. In questo caso il sistema ha un'unica soluzione che possiamo calcolare procedendo per sostituzioni successive dal basso: la soluzione è $(\frac{1}{\alpha}, -\frac{\alpha+2}{\alpha}, 1, \frac{1}{\alpha})$.

Per $\alpha = 0$ si ha

$$(A'|\underline{b}') = \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 0 \\ 0 & -1 & -2 & -2 & -1 \\ 0 & 0 & -1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right),$$

perciò $rg(A') = 3$ e $rg(A'|\underline{b}') = 4$, dunque il sistema non è risolubile.

Per $\alpha = 1$ si ha

$$(A'|\underline{b}') = \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 0 \\ 0 & -1 & -2 & -2 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right)$$

perciò $rg(A') = 3 = rg(A'|\underline{b}')$, quindi il sistema ha infinite soluzioni dipendenti da una variabile libera. Come al solito possiamo determinare tali soluzioni procedendo per sostituzioni successive dal basso: $(x_3, -1-2x_3, x_3, 1)$, $x_3 \in \mathbb{R}$.

Esercizio 1.3.3 Si consideri il sistema lineare Σ_α nelle incognite x_1, x_2, x_3 dipendente dal parametro reale α:

$$\Sigma_\alpha : \begin{cases} \alpha x_1 + (\alpha+3)x_2 + 2\alpha x_3 = \alpha + 2 \\ \alpha x_1 + (2\alpha+2)x_2 + 3\alpha x_3 = 2\alpha + 2 \\ 2\alpha x_1 + (\alpha+7)x_2 + 4\alpha x_3 = 2\alpha + 4 \end{cases}$$

1. Determinare le soluzioni del sistema lineare Σ_α al variare di $\alpha \in \mathbb{R}$.
2. Determinare le soluzioni del sistema lineare Σ_α interpretato ora come sistema lineare nelle 4 incognite x_1, x_2, x_3, x_4.

Svolgimento.

1. Consideriamo la matrice completa $(A|\underline{b})$ associata al sistema lineare Σ_α:
$$\left(\begin{array}{ccc|c} \alpha & \alpha+3 & 2\alpha & \alpha+2 \\ \alpha & 2\alpha+2 & 3\alpha & 2\alpha+2 \\ 2\alpha & \alpha+7 & 4\alpha & 2\alpha+4 \end{array}\right)$$
e riduciamola a scala con l'algoritmo di Gauss:
$$\left(\begin{array}{ccc|c} \alpha & \alpha+3 & 2\alpha & \alpha+2 \\ \alpha & 2\alpha+2 & 3\alpha & 2\alpha+2 \\ 2\alpha & \alpha+7 & 4\alpha & 2\alpha+4 \end{array}\right) \to \left(\begin{array}{ccc|c} \alpha & \alpha+3 & 2\alpha & \alpha+2 \\ 0 & \alpha-1 & \alpha & \alpha \\ 0 & -\alpha+1 & 0 & 0 \end{array}\right) \to$$
$$\left(\begin{array}{ccc|c} \alpha & \alpha+3 & 2\alpha & \alpha+2 \\ 0 & \alpha-1 & \alpha & \alpha \\ 0 & 0 & \alpha & \alpha \end{array}\right) = (A'|\underline{b}').$$
Se $\alpha \neq 0$ e $\alpha - 1 \neq 0$, cioè $\forall \alpha \in \mathbb{R}\setminus\{0,1\}$, si ha $rg(A') = rg(A'|b') = 3$ quindi, per la Proposizione 1.1.10, il sistema Σ_α ammette un'unica soluzione: $(\frac{2-\alpha}{\alpha}, 0, 1)$;
se $\alpha = 0$ otteniamo la matrice:
$$(A'|\underline{b}') = \left(\begin{array}{ccc|c} 0 & 3 & 0 & 2 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$
che non è in forma a scala ma può essere ridotta a scala sostituendo la seconda riga con la somma della seconda riga e della prima moltiplicata per $\frac{1}{3}$:
$$\left(\begin{array}{ccc|c} 0 & 3 & 0 & 2 \\ 0 & 0 & 0 & 2/3 \\ 0 & 0 & 0 & 0 \end{array}\right).$$
Il sistema di partenza risulta dunque equivalente al sistema lineare
$$\begin{cases} 3x_2 = 2 \\ 0 = 2/3 \end{cases}$$

che, ovviamente, non ha soluzioni;

infine, se $\alpha = 1$ si ha:

$$(A'|\underline{b}') = \left(\begin{array}{ccc|c} 1 & 4 & 2 & 3 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right)$$

che possiamo ridurre in forma a scala ottenendo la matrice

$$(A''|\underline{b}'') = \left(\begin{array}{ccc|c} 1 & 4 & 2 & 3 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

quindi $rg(A'') = rg(A''|\underline{b}'') = 2$ perciò il sistema Σ_1 è equivalente al sistema lineare di 2 equazioni in 3 incognite

$$\begin{cases} x_1 + 4x_2 + 2x_3 = 3 \\ x_3 = 1 \end{cases}$$

Tale sistema ha infinite soluzioni dipendenti da una variabile e l'insieme delle soluzioni risulta: $\{(1 - 4x_2, x_2, 1) \mid x_2 \in \mathbb{R}\}$.

2. Aggiungere l'incognita x_4 significa aggiungere alla matrice completa $(A|\underline{b})$ associata al sistema una colonna di zeri corrispondenti ai coefficienti di x_4. Pertanto, riducendo $(A|\underline{b})$ in forma a scala, si otterrà la matrice:
$$(A'|\underline{b}') = \left(\begin{array}{cccc|c} \alpha & \alpha+3 & 2\alpha & 0 & \alpha+2 \\ 0 & \alpha-1 & \alpha & 0 & \alpha \\ 0 & 0 & \alpha & 0 & \alpha \end{array}\right).$$

   Quindi, ragionando come sopra ma tenendo conto che in questo caso il numero di variabili è 4, si ottiene che:

   per $\alpha \in \mathbb{R} \backslash \{0, 1\}$ il sistema ha infinite soluzioni della forma $(\frac{2-\alpha}{\alpha}, 0, 1, x_4)$ con $x_4 \in \mathbb{R}$;

   per $\alpha = 0$ il sistema non ha soluzioni;

   per $\alpha = 1$ il sistema ha infinite soluzioni della forma $(1 - 4x_2, x_2, 1, x_4)$, con $x_2, x_4 \in \mathbb{R}$.

Esercizio 1.3.4 Stabilire se esistono valori del parametro reale k tali che il sistema lineare

$$\Sigma : \begin{cases} 2x_1 + x_2 - x_3 = 0 \\ 4x_1 - x_2 = 0 \\ x_1 + \frac{1}{2}x_2 - x_3 = -\frac{3}{2} \end{cases}$$

sia equivalente al sistema lineare

$$\Pi_k : \begin{cases} x_1 + x_2 - \frac{1}{2}x_3 = 1 \\ 2x_1 - x_2 + x_3 = 2 \\ kx_1 - 4x_2 + 3x_3 = k \end{cases}$$

(tutti i sistemi si intendono nelle variabili x_1, x_2, x_3).

Svolgimento. Due sistemi sono equivalenti se hanno le stesse soluzioni. Risolviamo innanzitutto il sistema lineare Σ. La matrice completa associata al sistema è:

$$(A|\underline{b}) = \left(\begin{array}{ccc|c} 2 & 1 & -1 & 0 \\ 4 & -1 & 0 & 0 \\ 1 & \frac{1}{2} & -1 & -\frac{3}{2} \end{array} \right).$$

Utilizzando l'algoritmo di Gauss possiamo ridurre $(A|\underline{b})$ in forma a scala, ottenendo la matrice

$$(A'|\underline{b}') = \left(\begin{array}{ccc|c} 2 & 1 & -1 & 0 \\ 0 & -3 & 2 & 0 \\ 0 & 0 & -\frac{1}{2} & -\frac{3}{2} \end{array} \right).$$

Abbiamo: $rg(A') = rg(A'|\underline{b}') = 3$, dunque il sistema Σ ammette una sola soluzione che possiamo determinare procedendo per sostituzioni successive dal basso:

$$x_3 = 3; \quad x_2 = 2; \quad x_1 = \frac{1}{2}.$$

L'unica soluzione del sistema Σ è pertanto $(\frac{1}{2}, 2, 3)$.

Affinché Σ sia equivalente a Π_k è dunque necessario che $(\frac{1}{2}, 2, 3)$ soddisfi contemporaneamente tutte le equazioni di Π_k. Sostituiamo dunque $x_1 = \frac{1}{2}$, $x_2 = 2$ e $x_3 = 3$ nelle equazioni di Π_k:

$$\begin{cases} \frac{1}{2} + 2 - \frac{3}{2} = 1 \\ 1 - 2 + 3 = 2 \\ \frac{k}{2} - 8 + 9 = k \end{cases}$$

Abbiamo così ottenuto due identità e la condizione necessaria $k = 2$. Pertanto $(\frac{1}{2}, 2, 3)$ è soluzione del sistema Π_k solo se $k = 2$. Possiamo quindi affermare che per $k \neq 2$ i sistemi Π_k e Σ non sono equivalenti ma non sappiamo ancora se i sistemi Σ e Π_2 sono equivalenti. Infatti ciò accade se e solo se $(\frac{1}{2}, 2, 3)$ è l'unica soluzione anche del sistema Π_2. Consideriamo allora la matrice completa associata al sistema Π_k per $k = 2$:

$$\left(\begin{array}{ccc|c} 1 & 1 & -\frac{1}{2} & 1 \\ 2 & -1 & 1 & 2 \\ 2 & -4 & 3 & 2 \end{array}\right).$$

Riducendo questa matrice in forma a scala otteniamo la matrice:

$$(A''|\underline{b}'') = \left(\begin{array}{ccc|c} 1 & 1 & -\frac{1}{2} & 1 \\ 0 & -3 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

Si ha dunque $rg(A'') = rg(A''|\underline{b}'') = 2 < 3$, dunque il sistema Π_2 ha infinite soluzioni. Possiamo allora concludere che non esistono valori di k tali che i sistemi Σ e Π_k siano tra loro equivalenti.

1.4 Esercizi proposti

1. Risolvere i seguenti sistemi lineari nelle incognite x, y, z:

   (a) $\begin{cases} x + y + z = 1 \\ 2x + 2y + z = 1 \\ 3y + z = 1 \end{cases}$

   (b) $\begin{cases} x - y + 4z = 10 \\ 3x + y + 5z = 15 \\ x + 3y - 3z = 6 \end{cases}$

2. Risolvere i seguenti sistemi lineari nelle incognite x, y, z, w:

   (a) $\begin{cases} x - y + 2z - 3w = 0 \\ 2x + y - w = 3 \\ 2y + z + w = -3 \\ 2x + z = 0 \end{cases}$

(b) $\begin{cases} x+y-z+w=0 \\ 2x-z-w=0 \\ x-y-2w=0 \\ 3x+y-2z=0 \end{cases}$

(c) $\begin{cases} x+z=7 \\ x+y=2 \\ 4x+12y+z=1 \end{cases}$

3. Discutere, al variare del parametro reale k, il seguente sistema lineare nelle incognite x, y, z:
$$\begin{cases} x+2y+kz=0 \\ x+y=-1 \\ x+ky=-2. \end{cases}$$
Stabilire per quali valori di k il sistema ammette soluzioni e, quando possibile, determinarle.

4. Si dica per quali valori di $k \in \mathbb{R}$ il seguente sistema lineare ammette soluzioni:
$$\begin{cases} x_1+2x_2+5x_4=0 \\ 2x_1+5x_2+11x_4=4 \\ x_1+x_2+(k-k^2)x_3+(5-k^2)x_4=-k^2-3 \\ -x_2+2(k^2-k)x_3+(3k^2-4)x_4=2(k^2+k-4) \end{cases}$$
Risolvere il sistema per i valori di k per i quali la soluzione non è unica.

5. Si dica per quali valori di $a \in \mathbb{R}$ il seguente sistema lineare, nelle incognite x, y, z, t, ammette soluzioni e, quando possibile, determinare tali soluzioni:
$$\begin{cases} 2x+y-z=1 \\ -2x+3z+t=1 \\ 2x+3y+(a^2+2a+3)z+(a^2-2)t=a+6 \\ y+2(a^2+2a+1)z+(3a^2-2a-7)t=3a+4 \end{cases}$$

6. Dato il sistema lineare nelle incognite x, y, z:
$$\Sigma_{a,b} : \begin{cases} x+(2+a)y=b \\ (2+2a)x+3y-(b+1)z=1+b \\ bx+by-(b+4)z=b^2+3b. \end{cases}$$

(a) Stabilire per quali valori di $a, b \in \mathbb{R}$ il sistema omogeneo associato ammette la soluzione $(a, -a, 0)$. (Si chiama sistema omogeneo associato al sistema lineare $A\underline{x} = \underline{b}$ il sistema $A\underline{x} = \underline{0}$.)

(b) Dire per quali tra i valori a, b trovati al punto precedente il sistema $\Sigma_{a,b}$ è risolubile e determinarne le soluzioni.

7. Stabilire per quali valori del parametro reale k il seguente sistema lineare nelle incognite x_1, x_2, x_3, x_4 è compatibile. Determinare, quando possibile, le soluzioni del sistema.

$$\begin{cases} x_1 + 3x_2 + kx_3 + 2x_4 = k \\ x_1 + 6x_2 + kx_3 + 3x_4 = 2k + 1 \\ -x_1 - 3x_2 + (k-2)x_4 = 1 - k \\ kx_3 + (2-k)x_4 = 1 \end{cases}$$

Lezione 2

Spazi vettoriali

Questa lezione è dedicata allo studio degli spazi vettoriali. Attraverso il concetto di spazio vettoriale vogliamo innanzitutto costruire un modello di uno spazio di dimensione qualsiasi. Non dobbiamo dimenticare che, aldilà di qualsiasi astrazione, ognuno di noi possiede una idea intuitiva di dimensione legata alla vita quotidiana: viviamo e ci muoviamo in uno spazio (fisico) tridimensionale, disegniamo su fogli essenzialmente bidimensionali, e così via.

2.1 Premessa: l'insieme dei numeri reali

Ricordiamo brevemente le principali proprietà delle operazioni che siamo abituati ad effettuare con i numeri, in particolare con i numeri reali.

La somma di due numeri reali è un'operazione che associa ad ogni coppia di numeri reali a e b un altro numero reale, indicato con $a + b$. Quindi la somma è una funzione che ha come dominio $\mathbb{R} \times \mathbb{R}$ e come codominio $\mathbb{R}$:

$$+ : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$$

$$(a, b) \mapsto a + b.$$

La somma di numeri reali è:

- commutativa: $a + b = b + a \quad \forall a, b \in \mathbb{R}$;
- associativa: $(a + b) + c = a + (b + c) \quad \forall a, b, c \in \mathbb{R}$;
- ammette *elemento neutro*, cioè esiste un numero, lo 0, tale che $0 + a = a + 0 = a \quad \forall a \in \mathbb{R}$;

- ogni numero reale a ammette *opposto*, cioè esiste un altro numero, che indichiamo con $-a$, tale che $a + (-a) = 0$.

Il prodotto di due numeri reali è un'operazione che associa ad ogni coppia di numeri reali a e b un altro numero reale, indicato con ab. Quindi anche il prodotto è una funzione che ha come dominio $\mathbb{R} \times \mathbb{R}$ e come codominio $\mathbb{R}$:

$$\cdot : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$$

$$(a, b) \mapsto ab.$$

Il prodotto di numeri reali è:

- commutativo: $ab = ba \quad \forall a, b \in \mathbb{R}$;

- associativo: $(ab)c = a(bc) \quad \forall a, b, c \in \mathbb{R}$;

- ammette elemento neutro, cioè esiste un numero, 1, tale che $1a = a1 = a \quad \forall a \in \mathbb{R}$;

- distributivo rispetto alla somma: $a(b + c) = ab + ac$.

Una delle proprietà più importanti dei numeri reali, che li distingue da altri insiemi di numeri, è la loro continuità. Geometricamente questo significa che pensiamo i numeri reali distribuiti lungo una retta. Più precisamente, data una retta e fissati su di essa un punto (origine) ed una unità di misura, esiste una corrispondenza biunivoca tra i punti sulla retta e l'insieme dei numeri reali. In altre parole ogni numero reale individua univocamente uno ed un solo punto sulla retta.

2.2 Un esempio: lo spazio vettoriale $\mathbb{R}^n$

Indichiamo con il simbolo $\mathbb{R}^2$ l'insieme delle coppie ordinate di numeri reali:

$$\mathbb{R}^2 = \{(x, y) \mid x, y \in \mathbb{R}\}.$$

Il fatto che le coppie siano ordinate significa, ad esempio, che l'elemento $(1, 2)$ è diverso dall'elemento $(2, 1)$.

Fissato nel piano un sistema di riferimento cartesiano, esiste una corrispondenza biunivoca tra $\mathbb{R}^2$ e l'insieme dei punti del piano. Fissare un riferimento cartesiano nel piano significa fissare due rette ortogonali orientate r ed s ed una unità di misura. Il punto di intersezione tra le due rette si chiama origine del sistema di riferimento. Ogni punto del piano è allora univocamente individuato da una coppia di numeri reali, detti coordinate del punto, che indicano, rispettivamente, la distanza del punto dalla retta s e la sua distanza dalla retta r. Lo studente che non avesse familiarità con il piano cartesiano, può pensare alla battaglia navale.

È naturale cercare di estendere alle coppie di numeri reali le operazioni che sappiamo eseguire con i numeri. Definiamo allora le seguenti operazioni:

$$+ : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2 \quad (\textit{somma})$$

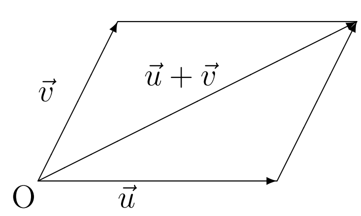
$$(x, y) + (x', y') := (x + x', y + y').$$

$$\cdot : \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}^2 \quad (\textit{prodotto per un numero reale})$$

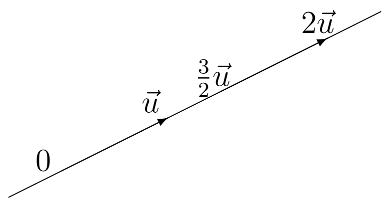
$$\lambda(x, y) := (\lambda x, \lambda y).$$

Si noti che pur avendo indicato l'applicazione di prodotto per un numero reale con il simbolo $\cdot$, nel prodotto $\lambda(x, y)$ abbiamo omesso il simbolo, esattamente come si fa di solito quando si moltiplicano due numeri reali.

Cerchiamo di interpretare geometricamente le operazioni definite nel caso di $\mathbb{R}^2$. A tal fine pensiamo ogni elemento (a, b) di $\mathbb{R}^2$ come il secondo estremo di un *vettore* applicato nell'origine, cioè di un segmento orientato uscente dall'origine con la freccia diretta verso il punto (a, b). In questo caso il modo di sommare due elementi di $\mathbb{R}^2$ coincide con la ben nota *regola del parallelogramma* con cui si sommano le forze in fisica. Secondo questa regola la somma di due vettori $\vec{u}$ e $\vec{v}$ applicati in un punto è un vettore applicato nello stesso punto avente direzione, verso e lunghezza della diagonale del parallelogramma avente per lati $\vec{u}$ e $\vec{v}$, uscente dal punto di applicazione.



Moltiplicando invece un vettore $\vec{v}$ per un numero reale α si ottiene un vettore avente la stessa direzione di $\vec{v}$, lunghezza moltiplicata per il valore assoluto di α e verso concorde o discorde da quello di $\vec{v}$ a seconda che α sia positivo o negativo.



Vale la pena di sottolineare fin da questo momento la diversa natura delle operazioni di somma di due elementi di $\mathbb{R}^2$ e di prodotto di un elemento di $\mathbb{R}^2$ per un numero reale: la somma associa a due elementi di $\mathbb{R}^2$ un elemento di $\mathbb{R}^2$. Per questo viene spesso detta *operazione interna*: associa ad una coppia di elementi *dentro* $\mathbb{R}^2$ un elemento di $\mathbb{R}^2$. Il prodotto per i numeri reali associa ad un numero reale e ad un elemento di $\mathbb{R}^2$ un altro elemento di $\mathbb{R}^2$. Questa operazione viene spesso detta *esterna* perché prende un elemento che sta dentro $\mathbb{R}^2$ ed uno che sta *fuori*, ed associa ad essi un elemento dentro $\mathbb{R}^2$.

È pressocché immediato verificare che la somma di elementi di $\mathbb{R}^2$ soddisfa le seguenti proprietà:

1. *commutativa*: $(x, y) + (x', y') = (x', y') + (x, y) \quad \forall (x, y), (x', y') \in \mathbb{R}^2$;

2. *associativa*: $((x, y) + (x', y')) + (x'', y'') = (x, y) + ((x', y') + (x'', y''))$ $\forall (x, y), (x', y'), (x'', y'') \in \mathbb{R}^2$;

3. *esistenza dell'elemento neutro* $(0, 0)$: $(x, y) + (0, 0) = (0, 0) + (x, y) = (x, y) \quad \forall (x, y) \in \mathbb{R}^2$.

4. *esistenza dell'opposto*: per ogni (x, y) esiste un elemento (a, b), detto *opposto* di (x, y), tale che $(a, b) + (x, y) = (x, y) + (a, b) = (0, 0)$. Ovviamente si ha: $(a, b) = (-x, -y)$.

Altrettanto immediato è verificare le seguenti proprietà del prodotto per i numeri reali:

5. *Proprietà distributive*:
   $a((x,y)+(x',y')) = a(x,y)+a(x',y')$, $\forall (x,y)$, $(x',y') \in \mathbb{R}^2$ e $\forall a \in \mathbb{R}$.
   $(a+b)(x,y) = a(x,y)+b(x,y)$, $\forall (x,y) \in \mathbb{R}^2$ e $\forall a,b \in \mathbb{R}$.
6. $(ab)(x,y) = a(b(x,y))$, $\forall (x,y) \in \mathbb{R}^2$ e $\forall a,b \in \mathbb{R}$.
7. $1(x,y) = (x,y)$, $\forall (x,y) \in \mathbb{R}^2$.

Naturalmente possiamo generalizzare quanto fatto su $\mathbb{R}^2$ all'insieme delle n-uple ordinate di numeri reali, per ogni $n \in \mathbb{N}$:

$$\mathbb{R}^n := \{(x_1,\dots,x_n) \,|\, x_1,\dots,x_n \in \mathbb{R}\}.$$

Definiamo le seguenti operazioni di somma $+$ e di prodotto per i numeri reali $\cdot$:

$$+ : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$$
$$(x_1,\dots,x_n)+(x'_1,\dots,x'_n) := (x_1+x'_1,\dots,x_n+x'_n);$$

$$\cdot : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$$
$$\lambda(x_1,\dots,x_n) := (\lambda x_1,\dots,\lambda x_n).$$

Con un po' di pazienza si possono verificare le proprietà $1. - 7.$ elencate sopra per la somma ed il prodotto per i numeri reali in $\mathbb{R}^2$.

Chiameremo $\mathbb{R}^n$, con le operazioni di somma e di prodotto per i numeri reali appena definite, *spazio vettoriale*. Chiameremo *vettori* gli elementi di $\mathbb{R}^n$ e *scalari* i numeri reali. In tal modo il prodotto per i numeri reali si dirà, equivalentemente, prodotto per scalari. Chiameremo, infine, *vettore nullo* di $\mathbb{R}^n$, e lo indicheremo con $\mathbf{0}_{\mathbb{R}^n}$, l'elemento neutro rispetto alla somma, cioè: $\mathbf{0}_{\mathbb{R}^n} = (0,\dots,0)$.

Lezione 3

Sottospazi vettoriali

Come possiamo riconoscere e descrivere uno spazio vettoriale dentro un altro? Come distinguiamo un sottoinsieme qualsiasi di uno spazio vettoriale da uno che ne possiede le stesse caratteristiche? Per rispondere a queste domande è necessario introdurre la definizione di sottospazio vettoriale.

3.1 Definizione e prime proprietà

Definizione 3.1.1 *Sia W un sottoinsieme non vuoto di uno spazio vettoriale V. Diremo che W è un sottospazio vettoriale di V se soddisfa le seguenti proprietà:*

1) W è chiuso rispetto alla somma, cioè per ogni $u, v \in W$ si ha che $u + v \in W$;

2) W è chiuso rispetto al prodotto per scalari, cioè per ogni $u \in W$ e ogni $\alpha \in \mathbb{R}$ si ha che $\alpha u \in W$.

È importante notare che W, con le operazioni di V opportunamente ristrette, è esso stesso uno spazio vettoriale. Infatti, la proprietà 1) della Definizione 3.1.1 assicura che la restrizione a W della somma definita in V dia come risultato un vettore di W:

$$+_V|_{W\times W} : W \times W \to W.$$

Analogamente la proprietà 2) della Definizione 3.1.1 assicura che la restrizione a $\mathbb{R} \times W$ del prodotto per scalari definito su $\mathbb{R} \times V$ dia come risultato un vettore di W. Dopodiché le proprietà della Definizione 2.3.1 continuano a valere perché valgono in V.

In particolare, dunque, ogni spazio vettoriale V possiede sempre almeno due sottospazi: V stesso e il sottospazio banale, costituito solo dal vettore nullo $\mathbf{0}_V$. Per quanto abbiamo osservato in 2.4, se V stesso non è banale, ogni sottospazio vettoriale di V non banale conterrà infiniti elementi.

Cerchiamo di chiarire il concetto di sottospazio vettoriale attraverso degli esempi e dei controesempi.

Esempio 3.1.2 L'insieme $X = \{(x, y) \in \mathbb{R}^2 \mid y = 0\}$ è un sottospazio vettoriale di $\mathbb{R}^2$. X è infatti:

1) non vuoto: contiene le infinite coppie di numeri reali $(x, 0)$;

2) chiuso rispetto alla somma: presi due elementi $(x_1, 0)$, $(x_2, 0)$ in X, la loro somma $(x_1, 0) + (x_2, 0) = (x_1 + x_2, 0)$ appartiene ancora all'insieme X;

3) chiuso rispetto al prodotto per scalari: presi un qualsiasi numero reale α ed un qualsiasi elemento $(x, 0) \in X$, il prodotto $\alpha(x, 0) = (\alpha x, 0)$ appartiene a X.

Geometricamente, fissato un sistema di riferimento cartesiano in $\mathbb{R}^2$, l'insieme X può essere pensato come l'asse delle ascisse. Allora sommando due vettori che giacciono sull'asse delle x o moltiplicando per uno scalare uno di essi, si ottiene ancora un vettore che giace sull'asse delle x.

Osservazione 3.1.3 Un sottospazio vettoriale W di uno spazio vettoriale V contiene sempre il vettore nullo di V. Infatti se W è un sottospazio dello spazio vettoriale V e $w \in W$, allora anche $\mathbf{0}_V = 0w \in W$, per la proprietà 2) della Definizione 3.1.1. In altre parole $\mathbf{0}_V \in S$ è una condizione necessaria affinché un sottoinsieme S sia un sottospazio di V.

Esempio 3.1.4 Consideriamo l'insieme $S = \{(x, y) \in \mathbb{R}^2 \mid y = x + 1\}$. Possiamo immediatamente affermare che S NON è un sottospazio vettoriale di $\mathbb{R}^2$ poiché non contiene $\mathbf{0}_{\mathbb{R}^2} = (0, 0)$. Dunque non tutte le rette del piano individuano dei sottospazi vettoriali di $\mathbb{R}^2$ (ma solo quelle passanti per $(0, 0)$).

ATTENZIONE: se S è un sottoinsieme di uno spazio vettoriale V la condizione $\mathbf{0}_V \in S$ è necessaria ma non sufficiente affinché S sia un sottospazio di V. Facciamo subito un controesempio:

Esempio 3.1.5 Sia $S = \{(x, y, z) \in \mathbb{R}^3 \mid xy = z\}$. Nonostante l'insieme S contenga il vettore nullo $(0, 0, 0)$ di $\mathbb{R}^3$, S non è un sottospazio vettoriale di $\mathbb{R}^3$ perché non è chiuso rispetto alla somma. Infatti i vettori $v = (1, 1, 1)$ e $w = (-1, -1, 1)$ appartengono ad S, dal momento che soddisfano l'equazione $xy = z$, tuttavia la loro somma $v + w = (1, 1, 1) + (-1, -1, 1) = (0, 0, 2)$ non appartiene ad S dal momento che $0 \cdot 0 \neq 2$.

Osservazione 3.1.6 Dire che un sottoinsieme S di uno spazio vettoriale V è chiuso rispetto al prodotto per scalari, significa dire che se S contiene un vettore non nullo allora *deve contenere anche tutti i suoi multipli.* Se pensiamo alla interpretazione geometrica del prodotto per scalari che abbiamo dato nel caso di $\mathbb{R}^2$, questo significa che se un sottospazio di $\mathbb{R}^2$ contiene un vettore non nullo v, allora esso contiene tutta la retta del piano individuata da v. Questo ragionamento geometrico ci permette di dire immediatamente che alcuni sottoinsiemi del piano NON sono spazi vettoriali, ad esempio l'insieme

$$P = \{(x, y) \in \mathbb{R}^2 \mid y = x^2\}.$$

Osservazione 3.1.7 Gli esempi fatti finora mettono in luce due diversi tipi di ragionamento. Per dimostrare che un sottoinsieme S di uno spazio vettoriale V è un sottospazio vettoriale di V, bisogna dimostrare che valgono le proprietà 1) e 2) della Definizione 3.1.1. Tali proprietà devono valere *sempre* cioè per ogni coppia di vettori di S (proprietà 1) e per tutti i numeri reali (proprietà 2).

Per dimostrare, al contrario, che un sottoinsieme di uno spazio vettoriale V NON è un sottospazio vettoriale di V, basta mostrare che una delle proprietà 1) e 2) della Definizione 3.1.1 fallisce, ossia, se $S \neq \emptyset$, che esiste (anche solo) una coppia di vettori di S la cui somma non sta in S o che esistono (anche solo) un vettore di S ed uno scalare il cui prodotto non sta in S.

3.2 Esercizi svolti

Esercizio 3.2.1 Si stabilisca se l'insieme $X := \{(r, s, r - s) \in \mathbb{R}^3\}$ è un sottospazio vettoriale di $\mathbb{R}^3$.

Svolgimento. Innanzitutto osserviamo che X è diverso dall'insieme vuoto perché $(0, 0, 0) \in X$ (basta prendere $r = s = 0$).

Consideriamo ora due elementi generici di X: $(r_1, s_1, r_1 - s_1)$ e $(r_2, s_2, r_2 - s_2)$. La loro somma è: $(r_1, s_1, r_1 - s_1) + (r_2, s_2, r_2 - s_2) = (r_1 + r_2, s_1 + s_2, r_1 -$

$s_1 + r_2 - s_2) = (r_1 + r_2, s_1 + s_2, r_1 + r_2 - (s_1 + s_2))$ e appartiene ancora ad X in quanto è del tipo $(r, s, r - s)$, con $r = r_1 + r_2$ e $s = s_1 + s_2$.

Siano poi $(r_1, s_1, r_1 - s_1) \in X$ e $\alpha \in \mathbb{R}$. Allora $\alpha(r_1, s_1, r_1 - s_1) = (\alpha r_1, \alpha s_1, \alpha(r_1 - s_1)) = (\alpha r_1, \alpha s_1, \alpha r_1 - \alpha s_1)$ appartiene ancora ad X in quanto è del tipo $(r, s, r - s)$, con $r = \alpha r_1$ e $s = \alpha s_1$. Quindi X è un sottospazio vettoriale di $\mathbb{R}^3$.

Esercizio 3.2.2 Si stabilisca se l'insieme $W := \{(x, z) \in \mathbb{R}^2 \mid 2x + z^2 = 0\} \subseteq \mathbb{R}^2$ è un sottospazio vettoriale di $\mathbb{R}^2$.
Svolgimento. L'insieme W non è chiuso rispetto alla somma. Infatti, i vettori $(-2, 2)$ e $(-2, -2)$ appartengono a W ma la loro somma $(-2, 2) + (-2, -2) = (-4, 0)$ non appartiene a W. Quindi W non è un sottospazio vettoriale di $\mathbb{R}^2$.

Esercizio 3.2.3 Si determini un sottoinsieme non vuoto di $\mathbb{R}^3$ chiuso rispetto alla somma ma non al prodotto per scalari.
Svolgimento. L'insieme $X := \{(x, y, z) | x, y, z \in \mathbb{R}, x \geq 0\}$ ha questa proprietà. Infatti X è non vuoto poiché, ad esempio, $(0, 0, 0) \in X$. Verifichiamo ora che sia chiuso rispetto alla somma. Siano $(x_1, y_1, z_1), (x_2, y_2, z_2) \in X$, con $x_1, x_2 \geq 0$. Allora $(x_1, y_1, z_1) + (x_2, y_2, z_2) = (x_1 + x_2, y_1 + y_2, z_1 + z_2) \in X$ perché $x_1 + x_2 \geq 0$ (la somma di due numeri reali non negativi è un numero reale non negativo).

Tuttavia X non è chiuso rispetto al prodotto per scalari. Infatti $(1, 0, 0) \in X$ ma $-1(1, 0, 0) = (-1, 0, 0)$ non appartiene a X.

3.3 Esercizi proposti.

1. Si dica quali dei seguenti sottoinsiemi di spazi vettoriali sono sottospazi:

   i) $S = \{(x, y, z) \in \mathbb{R}^3 \mid x = 0\}$.
   ii) $T = \{(x, y, z) \in \mathbb{R}^3 \mid x = y\}$.
   iii) $X = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y = 0\}$.
   iv) $X = \{(x, y, z) \in \mathbb{R}^3 \mid x + y + z = -1\}$.

2. Si dimostri che l'insieme delle soluzioni del sistema lineare omogeneo

$$\begin{cases} x_1 + 3x_2 - x_3 = 0 \\ 2x_1 + 4x_2 - 4x_3 - x_4 = 0 \\ x_2 + x_3 + 2x_4 = 0 \end{cases}$$

nelle incognite x_1, x_2, x_3, x_4 è un sottospazio vettoriale di $\mathbb{R}^4$.

3. Si scriva, se possibile, l'insieme $S = \{(x, y) \in \mathbb{R}^2 \mid x^2 + xy - 2y^2 = 0\}$ come unione di due sottospazi vettoriali di $\mathbb{R}^2$ e si dica se S è un sottospazio vettoriale di $\mathbb{R}^2$.

Lezione 4

Generatori

Sia V un $\mathbb{R}$-spazio vettoriale e siano $v_1, v_2, \ldots, v_n$ vettori di V e $\lambda_1, \lambda_2, \ldots, \lambda_n$ numeri reali. Allora possiamo considerare il vettore $\lambda_i v_i \in V$ per ogni indice $i = 1, 2, \ldots, n$, e poi fare la somma in V dei vettori ottenuti:

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n.$$

Definizione 4.0.1 *Dato un $\mathbb{R}$-spazio vettoriale V e dato un insieme di vettori $v_1, v_2, \ldots, v_n$ di V, si dirà che un vettore $v \in V$ è loro* **combinazione lineare** *se esistono dei numeri reali $\lambda_1, \lambda_2, \ldots, \lambda_n$ tali che*

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n.$$

Gli scalari $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ si dicono coefficienti della combinazione lineare.

Esempio 4.0.2 Si consideri lo spazio vettoriale $\mathbb{R}^2$.

1. Il vettore $(\sqrt{5}, \sqrt{5})$ è combinazione lineare dei vettori $(1, -1), (0, \sqrt{5})$ poiché $(\sqrt{5}, \sqrt{5}) = \sqrt{5}(1, -1) + 2(0, \sqrt{5})$;

2. una generica combinazione lineare dei vettori $(1, 0)$, $(0, 1)$ è $a(1, 0) + b(0, 1) = (a, b)$ con $a, b \in \mathbb{R}$. Quindi ogni vettore di $\mathbb{R}^2$ è combinazione lineare dei vettori $(1, 0), (0, 1)$.

In ogni spazio vettoriale V il vettore nullo è combinazione lineare di qualunque insieme di vettori. Infatti se $v_1, \ldots, v_n$ sono vettori di V, allora $\mathbf{0}_V = 0v_1 + 0v_2 + \cdots + 0v_n$.

Definizione 4.0.3 *Un $\mathbb{R}$-spazio vettoriale V si dice* **finitamente generato** *se esiste un insieme finito di vettori $v_1, v_2, \ldots, v_n$ tali che ogni vettore di V sia loro combinazione lineare. I vettori $v_1, v_2, \ldots, v_n$ si dicono allora* **generatori** *di V.*

Esempio 4.0.4 La parte 2. dell'Esempio 4.0.2 mostra che lo spazio vettoriale $\mathbb{R}^2$ è finitamente generato. Analogamente lo spazio vettoriale $\mathbb{R}^n$ è finitamente generato per ogni $n \in \mathbb{N}$. Infatti gli n vettori $\{(1, 0, \ldots, 0), \ldots, (0, \ldots, 0, 1)\}$ generano $\mathbb{R}^n$.

Esempio 4.0.5 Dato uno spazio vettoriale finitamente generato V, è unico l'insieme di generatori di V? Certamente no! Facciamo subito un esempio. Abbiamo già notato che i vettori $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ generano $\mathbb{R}^3$. Mostriamo ora che possiamo individuare un'altra terna di generatori di $\mathbb{R}^3$: consideriamo i vettori $u_1 = (1, 1, -1)$, $u_2 = (0, 1, 2)$, $u_3 = (0, -2, 3)$. Vogliamo mostrare che $\{u_1, u_2, u_3\}$ è un insieme di generatori di $\mathbb{R}^3$. Dobbiamo verificare che **ogni** vettore di $\mathbb{R}^3$ è combinazione lineare di u_1, u_2, u_3 . Prendiamo un vettore qualsiasi di $\mathbb{R}^3$: (α, β, γ) con $\alpha, \beta, \gamma \in \mathbb{R}$. Dobbiamo dimostrare che esistono ben definiti $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$ tali che

$$(\alpha, \beta, \gamma) = \lambda_1 u_1 +_{\mathbb{R}^3} \lambda_2 u_2 +_{\mathbb{R}^3} \lambda_3 u_3.$$

Calcoliamo tale somma usando la somma ed il prodotto per scalari che abbiamo definito in $\mathbb{R}^3$:

$$\begin{aligned}(\alpha, \beta, \gamma) &= \lambda_1(1, 1, -1) +_{\mathbb{R}^3} \lambda_2(0, 1, 2) +_{\mathbb{R}^3} \lambda_3(0, -2, 3) = \\ &= (\lambda_1, \lambda_1, -\lambda_1) +_{\mathbb{R}^3} (0, \lambda_2, 2\lambda_2) +_{\mathbb{R}^3} (0, -2\lambda_3, 3\lambda_3) = \\ &= (\lambda_1, \lambda_1 + \lambda_2 - 2\lambda_3, -\lambda_1 + 2\lambda_2 + 3\lambda_3).\end{aligned}$$

Per calcolare i λ_i richiesti dobbiamo risolvere il sistema

$$\begin{cases} \alpha = \lambda_1 \\ \beta = \lambda_1 + \lambda_2 - 2\lambda_3 \\ \gamma = -\lambda_1 + 2\lambda_2 + 3\lambda_3. \end{cases}$$

Il sistema ottenuto ha soluzione: $\lambda_1 = \alpha$, $\lambda_2 = \frac{2}{7}\gamma - \frac{1}{7}\alpha + \frac{3}{7}\beta$ e $\lambda_3 = \frac{1}{7}(\gamma + 3\alpha - 2\beta)$, quindi i vettori u_1, u_2, u_3 sono dei generatori di $\mathbb{R}^3$.

Esempio 4.0.6 Non tutti gli spazi vettoriali sono finitamente generati. Prendiamo ad esempio lo spazio vettoriale dei polinomi $\mathbb{R}[X]$ nella indeterminata

X. Se fosse finitamente generato esisterebbe un numero finito di polinomi $P_1(X), \dots, P_n(X)$ che generano $\mathbb{R}[X]$, quindi ogni polinomio a coefficienti reali si scriverebbe come loro combinazione lineare. Sia g_i con $i = 1, \dots, n$, il grado del polinomio $P_i(x)$ e sia g il massimo tra questi gradi. Allora ogni combinazione lineare di $P_1(X), \dots, P_n(X)$ avrà al più grado g. Di conseguenza nessun polinomio di grado $g + 1$ si potrà scrivere come combinazione lineare di $P_1(X), \dots, P_n(X)$. Dunque i polinomi $P_1(X), \dots, P_n(X)$ non sono dei generatori di $\mathbb{R}[X]$ per nessun n.

Osservazione 4.0.7 Sia $S = \{u_1, u_2, \dots, u_s\}$ un insieme di generatori di un $\mathbb{R}$-spazio vettoriale V. Allora ogni vettore v di V si scrive come loro combinazione lineare:

$$v = \lambda_1 u_1 +_V \lambda_2 u_2 +_V \dots +_V \lambda_s u_s.$$

Tale scrittura è unica? In generale la risposta è no. Ad esempio l'insieme $\{(0), (1)\}$ è chiaramente un insieme di generatori dello spazio vettoriale reale $\mathbb{R}$ (perché ogni $(\alpha) \in \mathbb{R}$ è $(\alpha) = \alpha(1) + (0)$, ma la scrittura non è unica infatti $(0) = 0(1) + 0(0) = 0(1) + 1(0)$. Come ulteriore esempio si consideri l'insieme ordinato $\mathcal{B} = \{(1, 1, 0), (0, 1, 1), (1, 0, 1), (2, 1, 1)\}$ di generatori di $\mathbb{R}^3$. Il vettore $(0, 1, 1)$ si può scrivere come $0(1, 1, 0) + 1(0, 1, 1) + 0(1, 0, 1) + 0(2, 1, 1)$, ma pure come $1(1, 1, 0) + 1(0, 1, 1) + 1(1, 0, 1) - 1(2, 1, 1)$.

Osservazione 4.0.8 Osserviamo che, dato un $\mathbb{R}$-spazio vettoriale V generato dall'insieme di vettori $\{u_1, u_2, \dots, u_s\}$ e preso qualsiasi vettore $u \in V$, l'insieme di vettori $\{u_1, u_2, \dots, u_s, u\}$ è pure un insieme di generatori di V. Infatti, preso qualsiasi vettore $v \in V$, esso è combinazione lineare di $u_1, \dots, u_s$: $v = a_1 u_1 + a_2 u_2 + \dots + a_s u_s$, dunque $v = a_1 u_1 + a_2 u_2 + \dots + a_s u_s + 0u$.

Definizione 4.0.9 *Siano $\{u_1, u_2, \dots, u_k\}$ vettori di un $\mathbb{R}$-spazio vettoriale V. Si chiama sottospazio di V generato da $u_1, \dots, u_k$ e si indica con*

$$\langle u_1, u_2, \dots, u_k \rangle$$

l'insieme di tutte le combinazioni lineari di $u_1, u_2, \dots, u_k$:

$$\langle u_1, u_2, \dots, u_k \rangle = \Big\{ \sum_{i=1}^{k} \alpha_i u_i = \alpha_1 u_1 + \alpha_2 u_2 + \dots + \alpha_k u_k \mid \alpha_i \in \mathbb{R}, i = 1, \dots k \Big\}.$$

Osservazione 4.0.10 Osserviamo prima di tutto che l'insieme delle combinazioni lineari di $u_1, u_2, \ldots, u_k$ è un sottospazio vettoriale di V.

Siano $\lambda \in \mathbb{R}$ e $v = \alpha_1 u_1 + \alpha_2 u_2 + \ldots + \alpha_k u_k$ una combinazione lineare di $u_1, u_2, \ldots, u_k$. Allora $\lambda v = \lambda(\alpha_1 u_1 + \alpha_2 u_2 + \ldots + \alpha_k u_k) = \lambda\alpha_1 u_1 + \lambda\alpha_2 u_2 + \ldots + \lambda\alpha_k u_k$ è pure una combinazione lineare di $u_1, u_2, \ldots, u_k$ e quindi un elemento del nostro insieme.

Se poi abbiamo due combinazioni lineari $v = \alpha_1 u_1 + \alpha_2 u_2 + \ldots + \alpha_k u_k$ e $w = \beta_1 u_1 + \beta_2 u_2 + \ldots + \beta_k u_k$, allora la loro somma $v+w = (\alpha_1 u_1 + \alpha_2 u_2 + \ldots + \alpha_k u_k) + (\beta_1 u_1 + \beta_2 u_2 + \ldots + \beta_k u_k) = (\alpha_1 + \beta_1)u_1 + (\alpha_2 + \beta_2)u_2 + \ldots + (\alpha_k + \beta_k)u_k$ è ancora una combinazione lineare dei vettori $u_1, u_2, \ldots, u_k$, quindi appartiene al nostro insieme.

Possiamo concludere che $\langle u_1, \ldots, u_k \rangle$ è un sottospazio vettoriale di V, essendo chiuso rispetto alla somma ed al prodotto per scalari.

Osserviamo inoltre che se T è un altro sottospazio di V contenente i vettori $u_1, u_2, \ldots, u_k$, per definizione di sottospazio T contiene ogni loro combinazione lineare e quindi tutto $\langle u_1, u_2, \ldots, u_k \rangle$. Quindi $\langle u_1, u_2, \ldots, u_k \rangle$ è il più piccolo sottospazio di V contenente i vettori $u_1, \ldots, u_k$.

Osserviamo infine che, per la Definizione 4.0.9, uno spazio vettoriale V è finitamente generato se e solo se esistono $v_1, \ldots, v_n$ vettori di V tali che $V = \langle v_1, \ldots, v_n \rangle$.

Esempio 4.0.11 Dato un $\mathbb{R}$-spazio vettoriale V, lo spazio vettoriale generato da un vettore non nullo v è l'insieme $\langle v \rangle = \{\alpha v \mid \alpha \in \mathbb{R}\}$ di tutti i multipli scalari di v.

Vale, analogamente, la seguente definizione:

Definizione 4.0.12 *Dato un qualsiasi sottoinsieme S di uno spazio vettoriale reale V, indichiamo con $\langle S \rangle$ il più piccolo sottospazio vettoriale di V contenente S. Allora $\langle S \rangle$ è l'insieme di tutte le combinazioni lineari di elementi di S a coefficienti in $\mathbb{R}$.*

4.1 Esercizi svolti

Esercizio 4.1.1 Stabilire se i seguenti insiemi di vettori generano $\mathbb{R}^3$:

(i) $(0, 0, 1)$, $(2, 1, 0)$, $(1, 1, 1)$;

(ii) $(2, 3, 4)$, $(3, 2, 1)$;

(iii) $(1,0,0)$, $(1,2,0)$, $(0,0,2)$, $(1,3,4)$.

Svolgimento. L'esercizio consiste nel stabilire se ogni vettore di $\mathbb{R}^3$ si possa scrivere come combinazione lineare dei vettori di volta in volta indicati.

(i) Sia (α,β,γ) un generico elemento di $\mathbb{R}^3$. Ci chiediamo se esistono λ_1, λ_2, λ_3 in $\mathbb{R}$ tali che sia $(\alpha,\beta,\gamma) = \lambda_1(0,0,1)+\lambda_2(2,1,0)+\lambda_3(1,1,1)$, cioè $(\alpha,\beta,\gamma) = (2\lambda_2+\lambda_3, \lambda_2+\lambda_3, \lambda_1+\lambda_3)$. Si tratta di risolvere il seguente sistema:

$$\begin{cases} \alpha = 2\lambda_2+\lambda_3 \\ \beta = \lambda_2+\lambda_3 \\ \gamma = \lambda_1+\lambda_3. \end{cases}$$

Sottraendo la seconda equazione dalla prima otteniamo: $\lambda_2 = \alpha-\beta$; sottraendo la prima equazione da due volte la seconda otteniamo: $\lambda_3 = 2\beta-\alpha$. Infine, sostituendo nella terza equazione l'espressione ottenuta di λ_3, otteniamo: $\lambda_1 = \gamma-2\beta+\alpha$. Dunque per ogni $(\alpha,\beta,\gamma) \in \mathbb{R}^3$ abbiamo determinato i numeri reali λ_i che cercavamo. Pertanto l'insieme di vettori (i) genera tutto $\mathbb{R}^3$.

(ii) In questo esercizio l'insieme che ci viene proposto contiene solo due elementi. Anche in questo caso potremmo procedere come prima cercando di risolvere un sistema, ma questa volta il sistema avrà 3 equazioni e 2 incognite. Quindi è lecito aspettarsi che un tale sistema non abbia sempre soluzioni, cioè che vi siano dei vettori di $\mathbb{R}^3$ che non sono generati dall'insieme (ii). In questo caso per risolvere il quesito sarà sufficiente esibire un vettore di $\mathbb{R}^3$ che non è combinazione lineare di $(2,3,4)$ e $(3,2,1)$.

Consideriamo il vettore $(1,0,0)$ di $\mathbb{R}^3$. Ci chiediamo: è possibile scrivere $(1,0,0)$ come combinazione lineare di $(2,3,4)$, $(3,2,1)$? Se ciò fosse possibile esisterebbero λ_1 e λ_2 in $\mathbb{R}$ tali che $(1,0,0) = \lambda_1(2,3,4)+\lambda_2(3,2,1) = (2\lambda_1+3\lambda_2, 3\lambda_1+2\lambda_2, 4\lambda_1+\lambda_2)$. Quindi il sistema

$$\begin{cases} 1 = 2\lambda_1+3\lambda_2 \\ 0 = 3\lambda_1+2\lambda_2 \\ 0 = 4\lambda_1+\lambda_2. \end{cases}$$

dovrebbe ammettere soluzione, ma dalla seconda e dalla terza equazione ricaviamo $\lambda_1 = 0 = \lambda_2$ che non soddisfano la prima equazione. Il sistema non ha soluzioni cioè il vettore $(1,0,0)$ non è combinazione lineare dei vettori $(2,3,4)$ e $(3,2,1)$. Questo dimostra che $(2,3,4)$ e $(3,2,1)$ non generano $\mathbb{R}^3$.

(iii) Anche in questo caso potremmo precedere come in (i) e, preso $(\alpha,\beta,\gamma)\in\mathbb{R}^3$, determinare λ_1, λ_2, λ_3, λ_4 in $\mathbb{R}$ tali che $(\alpha,\beta,\gamma)=\lambda_1(1,0,0)+\lambda_2(1,2,0)+\lambda_3(0,0,2)+\lambda_4(1,3,4)$.

Mostriamo invece un modo alternativo di procedere. Prima di tutto osserviamo un fatto di carattere generale: dato un insieme I di vettori di uno spazio vettoriale V e fissato un insieme S di generatori di V, se ogni elemento di S è combinazione lineare degli elementi di I, allora anche I è un insieme di generatori di V. Infatti ogni vettore di V si scrive come combinazione lineare dei vettori di S che a loro volta sono combinazioni lineari dei vettori di I e quindi ogni vettore di V è combinazione lineare dei vettori di I. Osserviamo inoltre che i vettori $(1,0,0)$, $(0,1,0)$, $(0,0,1)$ generano $\mathbb{R}^3$, infatti ogni elemento (α,β,γ) è loro combinazione lineare:

$$(\alpha,\beta,\gamma)=\alpha(1,0,0)+\beta(0,1,0)+\gamma(0,0,1).$$

Notiamo, dunque, che il vettore $(1,0,0)$ appartiene all'insieme dato e che si ha: $(0,1,0)=\frac{1}{2}((1,2,0)-(1,0,0))$; $(0,0,1)=\frac{1}{2}(0,0,2)$. Concludiamo, per quanto appena osservato, che l'insieme di vettori (iii) genera $\mathbb{R}^3$. Esplicitamente per un vettore qualsiasi (α,β,γ) di $\mathbb{R}^3$ si ha $(\alpha,\beta,\gamma)=\alpha(1,0,0)+\beta(0,1,0)+\gamma(0,0,1)=\alpha(1,0,0)+\beta(\frac{1}{2}((1,2,0)-(1,0,0)))+\gamma(\frac{1}{2}(0,0,2))=(\alpha-\frac{\beta}{2})(1,0,0)+\frac{\beta}{2}(1,2,0)+\frac{\gamma}{2}(0,0,2)$.

Esercizio 4.1.2 Determinare due insiemi distinti di generatori dello spazio vettoriale $\mathbb{R}^{\leq 2}[x]$ dei polinomi a coefficienti reali nella variabile x di grado ≤ 2.

Svolgimento. L'insieme $\{1,x,x^2\}$ è certamente un insieme di generatori di $\mathbb{R}^{\leq 2}[x]$. Infatti ogni elemento di $\mathbb{R}^{\leq 2}[x]$ è un polinomio della forma $a+bx+cx^2$ con $a,b,c\in\mathbb{R}$ quindi $a+bx+cx^2=a\cdot 1+b\cdot x+c\cdot x^2$.

È molto facile, a questo punto, costruire un altro insieme di generatori: basterà aggiungere all'insieme già individuato un qualsiasi altro polinomio di $\mathbb{R}^{\leq 2}[x]$. Così, ad esempio, l'insieme $\{1,x,x^2,4x+69x^2\}$ genera $\mathbb{R}^{\leq 2}[x]$. Ma pure $\{1,x,4x+69x^2\}$ è un insieme di generatori, mentre $\{x,x^2,4x+69x^2\}$ non lo è (verificarlo per esercizio).

Esercizio 4.1.3 Mostrare che l'insieme dei polinomi $3 + x$, x^2, $1 + x^2 + x^3$ non è un insieme di generatori di $\mathbb{R}^{\leq 3}[x]$.

Svolgimento. Prendiamo un monomio di grado 0, ad esempio 1, e vediamo se è possibile scriverlo come combinazione lineare dei polinomi $3 + x$, x^2, $1 + x^2 + x^3$, cerchiamo cioè α, β, γ tali che sia $1 = \alpha(3 + x) + \beta x^2 + \gamma(1 + x^2 + x^3) = 3\alpha + \gamma + \alpha x + (\beta + \gamma)x^2 + \gamma x^3$. Ricordiamo che due polinomi sono uguali se i coefficienti dei termini dello stesso grado dei due polinomi sono ordinatamente uguali. Dobbiamo pertanto risolvere il sistema

$$\begin{cases} 1 = 3\alpha + \gamma \\ 0 = \alpha \\ 0 = \beta + \gamma \\ 0 = \gamma. \end{cases}$$

Si vede immediatamente che il sistema trovato non ha soluzioni, pertanto i polinomi $3 + x$, x^2, $1 + x^2 + x^3$ non individuano un insieme di generatori di $\mathbb{R}^{\leq 3}[x]$.

Esercizio 4.1.4 Verificare che l'insieme delle matrici $e_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $e_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $e_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, $e_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ genera lo spazio vettoriale $\mathcal{M}_2(\mathbb{R})$.

Svolgimento. Sia $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ una qualunque matrice di $\mathcal{M}_2(\mathbb{R})$. Allora $A = ae_{11} + be_{12} + ce_{21} + de_{22}$ (notiamo inoltre che tale scrittura è unica). Questo prova che $\{e_{11}, e_{12}, e_{21}, e_{22}\}$ è un insieme di generatori di $\mathcal{M}_2(\mathbb{R})$. Osserviamo che i coefficienti della combinazione lineare trovata coincidono con le entrate di A. Un altro insieme di generatori di $\mathcal{M}_2(\mathbb{R})$ è l'insieme costituito dalle matrici $e_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $e_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $e_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, $T = \begin{pmatrix} 3 & 0 \\ 2 & 1 \end{pmatrix}$.

Esercizio 4.1.5 L'insieme dei polinomi di grado 2 a coefficienti reali nella variabile x è un sottospazio vettoriale di $\mathbb{R}[x]$?

Svolgimento. Condizione necessaria affinché un sottoinsieme S di uno spazio vettoriale V sia un sottospazio di V è che S contenga il vettore nullo $\mathbf{0}_V$ di V. Il vettore nullo nello spazio vettoriale $\mathbb{R}[x]$ è il polinomio identicamente nullo, cioè il polinomio di grado 0 con termine noto uguale a 0, pertanto esso non è contenuto nell'insieme dei polinomi di grado 2. Concludiamo che l'insieme dei polinomi di grado 2 non è un sottospazio vettoriale di $\mathbb{R}[x]$.

Esercizio 4.1.6 Stabilire quali tra i seguenti sottoinsiemi sono sottospazi vettoriali di $\mathbb{R}^3$:

(i) $A = \{(x, y, z) \in \mathbb{R}^3 \mid x + y + z = 1\}$;

(ii) $B = \{(x, y, z) \in \mathbb{R}^3 \mid x + y + z = 0\}$;

(iii) $C = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y = 0\}$.

Svolgimento.

(i) Come nell'esercizio precedente, osserviamo immediatamente che l'insieme A non contiene il vettore $\mathbf{0}_{\mathbb{R}^3} = (0, 0, 0)$ pertanto A non è un sottospazio vettoriale di $\mathbb{R}^3$.

Gli insiemi B e C contengono $(0, 0, 0)$ ma questo non è sufficiente a dimostrare che essi siano sottospazi vettoriali di $\mathbb{R}^3$. Dobbiamo stabilire se B e C sono chiusi rispetto alle operazioni di $\mathbb{R}^3$, cioè se presi due vettori v, w in B (risp. C) la loro somma appartiene ancora a B (risp. C) e se preso un qualunque vettore v in B (risp. C) e un qualunque numero reale λ il prodotto λv appartiene a B (risp. C).

(ii) Siano $v = (x, y, z)$ e $w = (\tilde{x}, \tilde{y}, \tilde{z})$ elementi di B, cioè: $x + y + z = 0$ e $\tilde{x} + \tilde{y} + \tilde{z} = 0$. Allora il vettore $v + w = (x + \tilde{x}, y + \tilde{y}, z + \tilde{z})$ appartiene a B dal momento che le sue componenti soddisfano l'equazione di B: $(x + \tilde{x}) + (y + \tilde{y}) + (z + \tilde{z}) = (x + y + z) + (\tilde{x} + \tilde{y} + \tilde{z}) = 0 + 0 = 0$. Analogamente, per ogni $\lambda \in \mathbb{R}$, $\lambda v = (\lambda x, \lambda y, \lambda z)$ appartiene a B dal momento che $\lambda x + \lambda y + \lambda z = \lambda(x + y + z) = \lambda 0 = 0$. Quindi B è sottospazio vettoriale di $\mathbb{R}^3$.

(iii) Consideriamo ora l'insieme C: i vettori $v = (1, -1, 0)$ e $w = (-1, -1, 0)$ appartengono a C ma la loro somma $v + w = (0, -2, 0)$ non appartiene a C (perché $0^2 - 2 \neq 0$). Pertanto C non è chiuso rispetto alla somma e quindi non è un sottospazio vettoriale di $\mathbb{R}^3$.

Quest'ultimo esercizio ci fornisce un'indicazione che potrà essere utile: se un sottoinsieme di uno spazio vettoriale è descritto da equazioni, tale sottoinsieme difficilmente sarà un sottospazio vettoriale se le equazioni coinvolte non sono lineari nelle incognite oppure se appare un termine noto. Per ora prenderemo questa osservazione solo come un'indicazione di massima.

Esercizio 4.1.7 Verificare che i seguenti insiemi di matrici sono sottospazi vettoriali di $\mathcal{M}_3(\mathbb{R})$ e per ciascuno di essi esibire un insieme di generatori.

(i) $A = \left\{ \begin{pmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{pmatrix} \mid a, b, c \in \mathbb{R} \right\}$;

(ii) $B = \left\{ \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \mid \lambda_1, \lambda_2, \lambda_3 \in \mathbb{R} \right\}$;

(iii) $C = \left\{ \begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix} \mid a, b, c, d, e, f \in \mathbb{R} \right\}$.

Svolgimento. Come nell'esercizio precedente, per verificare che A, B e C sono sottospazi vettoriali di $\mathcal{M}_3(\mathbb{R})$ basta utilizzare la Definizione **??**. Effettuiamo questa verifica soltanto per l'insieme A e lasciamo allo studente la verifica per gli insiemi B e C.

(i) L'insieme A è l'insieme delle matrici quadrate $M = (m_{ij})$ di ordine 3 ($i, j \in \{1, 2, 3\}$) ad entrate reali, triangolari strettamente superiori, cioè delle matrici quadrate di ordine 3 per cui $m_{ij} = 0$ se $i \geq j$ (i.e. le matrici i cui elementi diagonali sono nulli e i cui elementi al di sotto della diagonale principale sono anch'essi nulli). Siano dunque $M_1 = \begin{pmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{pmatrix}$ e $M_2 = \begin{pmatrix} 0 & \alpha & \beta \\ 0 & 0 & \gamma \\ 0 & 0 & 0 \end{pmatrix}$ due elementi di A, con $a, b, c, \alpha, \beta, \gamma \in \mathbb{R}$. Allora la matrice $M_1 + M_2 = \begin{pmatrix} 0 & a+\alpha & b+\beta \\ 0 & 0 & c+\gamma \\ 0 & 0 & 0 \end{pmatrix}$ è evidentemente una matrice triangolare strettamente superiore e quindi appartiene all'insieme A. Analogamente, per ogni $\lambda \in \mathbb{R}$, $\lambda M_1 = \begin{pmatrix} 0 & \lambda a & \lambda b \\ 0 & 0 & \lambda c \\ 0 & 0 & 0 \end{pmatrix}$ appartiene all'insieme A. Pertanto A è chiuso rispetto alla somma e al prodotto per scalari ed quindi è un sottospazio vettoriale di $\mathcal{M}_3(\mathbb{R})$. Per determinare un insieme di generatori di A osserviamo che un suo generico elemento $M = \begin{pmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{pmatrix}$, con $a, b, c \in \mathbb{R}$ si può sempre scrivere come:

$$M = \begin{pmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{pmatrix} = a \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + b \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + c \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

e quindi l'insieme delle matrici $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ genera il sottospazio A.

(ii) Procedendo nello stesso modo otteniamo che $\left\{ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right\}$ è un insieme di generatori del sottospazio B.

(iii) Infine, le matrici $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$ generano C.

C è detto l'insieme delle *matrici simmetriche* di ordine 3.

4.2 Esercizi proposti

Esercizio 4.2.1 Costruire un esempio di un sottoinsieme di $\mathcal{M}_2(\mathbb{R})$ costituito da infiniti elementi, che non sia un sottospazio vettoriale di $\mathcal{M}_2(\mathbb{R})$.

Esercizio 4.2.2 Determinare i valori di $k \in \mathbb{R}$ tali che l'insieme

$$S_k = \{(x, y, z, t) \in \mathbb{R}^4 \mid x + 2y - kz + 8t = k\}$$

sia un sottospazio vettoriale di $\mathbb{R}^4$. Per i valori di k trovati determinare un insieme di generatori di S_k e, se possibile, esibire un vettore di $\mathbb{R}^4$ che non sia una combinazione lineare di questi.

Esercizio 4.2.3 Sia S l'insieme dei polinomi di grado 3 a coefficienti reali nella variabile x.

1. S è un sottospazio vettoriale di $\mathbb{R}[x]$?
2. Determinare il più piccolo sottospazio vettoriale T di $\mathbb{R}[x]$ contenente S.

Esercizio 4.2.4 Sia $S = \{(x, y) \in \mathbb{R}^2 \mid 2x^2 - 7xy + 3y^2 = 0\}$.

1. Mostrare che S non è un sottospazio vettoriale di $\mathbb{R}^2$.
2. Scrivere, se possibile, S come unione di sottospazi vettoriali di $\mathbb{R}^2$.
3. Determinare il più piccolo sottospazio vettoriale di $\mathbb{R}^2$ contenente S.

Lezione 5

Basi e dimensione

5.1 Dipendenza e indipendenza lineare

Nell'Osservazione 4.0.8 abbiamo visto che, dato un $\mathbb{R}$-spazio vettoriale V generato dai vettori u_1, u_2, …, u_s (quindi $V = \langle u_1, u_2, \ldots, u_s \rangle$) si possono aggiungere all'insieme $\{u_1, u_2, \ldots, u_s\}$ altri vettori e ottenere ancora un insieme di generatori. Inoltre abbiamo visto che ogni vettore $v \in V$ si può scrivere come combinazione lineare di u_1, u_2, …, u_s tramite $\lambda_1, \lambda_2, \ldots, \lambda_s \in \mathbb{R}$: $v = \lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_s u_s$, ma che tali λ_i non sono in generale unici. In questo paragrafo cercheremo di caratterizzare gli insiemi di vettori I per i quali valga il fatto che se un vettore si scrive come combinazione lineare dei vettori di I allora tale scrittura è unica. Se questi vettori sono anche dei generatori dello spazio vettoriale V potremo scrivere ogni vettore di V in maniera unica come loro combinazione lineare. Vedremo ora come per testare l'unicità della scrittura basti testarla sul vettore nullo $\mathbf{0}_V$.

Definizione 5.1.1 *In un $\mathbb{R}$-spazio vettoriale V i vettori $u_1, u_2, \ldots, u_k$, $k \in \mathbb{N}$, si dicono* **linearmente indipendenti** *se il solo modo di scrivere il vettore nullo $\mathbf{0}_V$ come loro combinazione lineare è con tutti i coefficienti nulli, i.e.*

$$\mathbf{0}_V = \lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_k u_k \Leftrightarrow \lambda_1 = \lambda_2 = \ldots = \lambda_k = 0.$$

Se, al contrario, il vettore nullo si può scrivere in modi diversi come combinazione lineare dei vettori $u_1, u_2, \ldots, u_k$, allora diremo che i vettori $u_1, u_2, \ldots,$ u_k sono **linearmente dipendenti**.

Esempi 5.1.2 1. In $\mathbb{R}^3$ i due vettori $(1,0,1),(2,1,1)$ sono linearmente indipendenti: infatti se dovessimo scrivere il vettore nullo di $\mathbb{R}^3$ come loro combinazione lineare avremmo $(0,0,0)=\lambda_1(1,0,1)+\lambda_2(2,1,1)$ da cui otterremmo che $(0,0,0)=(\lambda_1+2\lambda_2,\lambda_2,\lambda_1+\lambda_2)$ cioè $\lambda_1=\lambda_2=0$.

2. Consideriamo ora in $\mathbb{R}^2$ l'insieme $I=\{(1,1),(2,1),(1,-1)\}$. Il vettore nullo si scrive in modi diversi come combinazione lineare dei vettori di I:

$$3(1,1)-2(2,1)+(1,-1)=0(1,1)+0(2,1)+0(1,-1)=(0,0),$$

quindi i vettori di I sono linearmente dipendenti.

3. In $\mathbb{R}^n$ consideriamo un insieme di vettori con la seguente proprietà: ciascun vettore ha una componente diversa da zero, diciamo la i-esima, e i rimanenti vettori hanno invece entrata i-esima nulla. Tali vettori sono linearmente indipendenti. Ad esempio consideriamo in $\mathbb{R}^4$ i vettori $(2,1,0,0),(0,4,0,1),(0,5,1,0)$ (la prima componente del primo vettore è non nulla mentre la prima componente degli altri due vettori è nulla; il secondo vettore ha la quarta entrata diversa da zero mentre gli altri hanno quarta entrata nulla; infine il terzo vettore ha la terza entrata non nulla e i primi due hanno la terza entrata uguale a 0). Allora una loro combinazione lineare: $\alpha(2,1,0,0)+\beta(0,4,0,1)+\gamma(0,5,1,0)=(2\alpha,\alpha+4\beta+5\gamma,\gamma,\beta)$ è uguale a zero se e solo se $\alpha=\beta=\gamma=0$. Pertanto i vettori $(2,1,0,0),(0,4,0,1),(0,5,1,0)$ sono linearmente indipendenti.

Osservazione 5.1.3 Tra tutti gli insiemi di vettori possiamo scegliere anche l'insieme formato da un solo vettore. La domanda è la seguente: quando un vettore v di uno spazio vettoriale V è linearmente indipendente? Per definizione dobbiamo vedere in che modo possiamo scrivere il vettore nullo come sua combinazione lineare: $\lambda v=\mathbf{0}_V$. Ora, abbiamo già visto che se $v\neq\mathbf{0}_V$ allora $\lambda v=\mathbf{0}_V$ se e solo se $\lambda=0$. Se, invece, $v=\mathbf{0}_V$, allora per ogni $\lambda\in\mathbb{R}$, si ha $\lambda\mathbf{0}_V=\mathbf{0}_V$. Quindi un vettore è linearmente indipendente se e solo se $v\neq\mathbf{0}_V$. Osserviamo inoltre che se un insieme di vettori contiene il vettore nullo, allora è un insieme linearmente dipendente: considerato infatti l'insieme $\{\mathbf{0}_V,v_2,v_3,\dots v_n\}$, la combinazione lineare $(50)\mathbf{0}_V+0v_2+0v_3+\cdots+0v_n=\mathbf{0}_V$ ma i suoi coefficienti non sono tutti nulli.

Osservazione 5.1.4 Due vettori sono linearmente dipendenti se e solo se uno dei due è multiplo dell'altro. Infatti se $v_1, v_2 \in V$ spazio vettoriale, allora v_1 e v_2 sono linearmente dipendenti se e solo se esistono $\lambda_1, \lambda_2 \in \mathbb{R}$ non entrambi nulli tali che $\lambda_1 v_1 + \lambda_2 v_2 = \mathbf{0}_V$. Supponiamo che sia $\lambda_1 \neq 0$, allora $v_1 = -(\lambda_2/\lambda_1)v_2$ quindi v_1 è multiplo di v_2.

Osservazione 5.1.5 Supponiamo che i vettori $v_1, \ldots, v_k$ di V siano linearmente dipendenti. Per definizione questo significa che è possibile scrivere

$$\mathbf{0}_V = a_1 v_1 + \ldots + a_k v_k$$

con qualcuno dei coefficienti reali a_i diverso da 0. Tanto per fissare le idee, supponiamo che sia $a_1 \neq 0$. Allora $v_1 = -\frac{a_2}{a_1}v_2 - \ldots - \frac{a_k}{a_1}v_k$, cioè v_1 è combinazione lineare di $v_2, \ldots, v_k$. Abbiamo dunque mostrato che dire che k vettori sono linearmente dipendenti significa dire che uno di essi è combinazione lineare degli altri.

Osservazione 5.1.6 È importante notare che se i vettori $v_1, \ldots, v_k$ sono linearmente indipendenti allora qualsiasi sottoinsieme di $\{v_1, \ldots, v_k\}$ continua ad essere costituito da vettori linearmente indipendenti. Infatti dire che $v_1, \ldots, v_k$ sono linearmente indipendenti significa che nessuno di questi k vettori è combinazione lineare degli altri pertanto questa proprietà resta vera se si eliminano vettori dall'insieme $\{v_1, \ldots, v_k\}$.

La Definizione 5.1.1 risponde al quesito introdotto all'inizio della lezione, infatti vale la seguente proposizione:

Proposizione 5.1.7 *In un $\mathbb{R}$-spazio vettoriale V i vettori $u_1, u_2, \ldots, u_k$, $k \in \mathbb{N}$, sono linearmente indipendenti se e solo se ogni loro combinazione lineare si scrive in modo unico:*

$$\lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_k u_k = \lambda_1' u_1 + \lambda_2' u_2 + \cdots + \lambda_k' u_k, \qquad \lambda_i, \lambda_i' \in \mathbb{R}$$

$$\Updownarrow$$

$$\lambda_1 = \lambda_1';\ \lambda_2 = \lambda_2'; \ldots; \lambda_k = \lambda_k'.$$

Dimostrazione. "$\Rightarrow$" Per ipotesi sappiamo che i vettori $u_1, \ldots, u_k$ sono linearmente indipendenti e vogliamo mostrare che ogni vettore in $\langle u_1, \ldots, u_k \rangle$

si scrive in modo unico come combinazione lineare dei vettori $u_1, \dots, u_k$. Supponiamo di poter scrivere:

$$v = \lambda_1 u_1 + \lambda_2 u_2 + \dots + \lambda_k u_k = \lambda_1' u_1 + \lambda_2' u_2 + \dots + \lambda_k' u_k$$

per qualche $\lambda_i, \lambda_i' \in \mathbb{R}$, $i = 1, 2, \dots, k$. Allora possiamo scrivere

$$\mathbf{0}_V = v - v = \lambda_1 u_1 + \lambda_2 u_2 + \dots + \lambda_k u_k - (\lambda_1' u_1 + \lambda_2' u_2 + \dots + \lambda_k' u_k)$$

e per la proprietà distributiva del prodotto per scalari rispetto alla somma:

$$\begin{aligned} \mathbf{0}_V &= \lambda_1 u_1 + \lambda_2 u_2 + \dots + \lambda_k u_k - \lambda_1' u_1 - \lambda_2' u_2 + \dots - \lambda_k' u_k \\ &= (\lambda_1 - \lambda_1') u_1 + (\lambda_2 - \lambda_2') u_2 + \dots + (\lambda_k - \lambda_k') u_k. \end{aligned}$$

Ora, dal momento che $u_1, \dots, u_k$ sono linearmente indipendenti, l'unico modo di scrivere il vettore nullo come loro combinazione lineare è a coefficienti tutti nulli. Pertanto $\lambda_i = \lambda_i'$ per ogni $i = 1, \dots, k$, cioè esiste un solo modo di scrivere v come combinazione lineare di $u_1, \dots, u_k$.

"$\Leftarrow$" In questo caso la nostra ipotesi è che ogni vettore che è combinazione lineare dei vettori $u_1, u_2, \dots, u_k$ lo sia in modo unico. Automaticamente il vettore nullo è loro combinazione lineare (lo è di ogni insieme di vettori): $\mathbf{0}_V = 0u_1 + 0u_2 + \dots + 0u_k$ e questa scrittura è unica per ipotesi. Quindi abbiamo dimostrato che i vettori $u_1, u_2, \dots, u_k$ sono linearmente indipendenti. **C. V. D.**

5.2 Basi e dimensione

Siano V uno spazio vettoriale e $I = \{u_1, u_2, \dots, u_n\}$ un insieme di vettori. Nelle sezioni precedenti abbiamo risolto i seguenti quesiti:

Quesito 1: quando è possibile scrivere ogni vettore di V come combinazione lineare dei vettori di I? Risposta: quando I è insieme di generatori di V.

Quesito 2: quando è possibile scrivere ogni combinazione lineare di vettori di I in modo unico? Risposta: quando I è insieme di vettori linearmente indipendenti.

In questo paragrafo uniremo queste due nozioni: vogliamo che ogni vettore di V si scriva in modo unico come combinazione lineare degli elementi di I. In questa frase il termine "ogni" indica che stiamo cercando un insieme di generatori e il termine "unico" che stiamo cercando un insieme di vettori lineramente indipendenti.

Definizione 5.2.1 *In un $\mathbb{R}$-spazio vettoriale V un insieme (finito) ordinato di vettori $\{w_1, w_2, \ldots, w_s\}$ si dice base (finita) di V se è un insieme di generatori linearmente indipendenti di V.*

Si noti che se V ha una base finita allora è finitamente generato, ma non tutti gli insiemi di generatori sono basi. Mostreremo come, dato un insieme di generatori di uno spazio vettoriale, sia possibile trovarne un sottoinsieme che risulti ancora un insieme di generatori. Cerchiamo di essere più precisi:

Proposizione 5.2.2 *Sia $I = \{u_1, u_2, \ldots, u_k\}$ un insieme di generatori di un $\mathbb{R}$-spazio vettoriale V. Se i vettori $u_1, u_2, \ldots, u_k$ sono linearmente dipendenti esiste un sottoinsieme proprio di I costituito da generatori di V.*

Dimostrazione. Dal momento che i vettori $u_1, \ldots, u_k$ sono linearmente dipendenti, uno di essi è combinazione lineare degli altri. Tanto per fissare le idee, supponiamo che sia

$$u_k = \lambda_1 u_1 + \ldots + \lambda_{k-1} u_{k-1}.$$

Sia dunque v un qualunque vettore di V. Allora possiamo scrivere v come combinazione lineare dei vettori $u_1, u_2, \ldots, u_k$, dal momento che questi generano V:

$$v = \beta_1 u_1 + \beta_2 u_2 + \ldots + \beta_k u_k$$

e sostituendo a u_k l'espressione precedente otteniamo:

$v = \beta_1 u_1 + \ldots + \beta_k(\lambda_1 u_1 + \ldots + \lambda_{k-1} u_{k-1}) =$

$(\beta_1 + \beta_k \lambda_1) u_1 + \ldots + (\beta_{k-1} + \beta_k \lambda_{k-1}) u_{k-1},$

che è una combinazione lineare dei vettori $u_1, \ldots, u_{k-1}$. Quindi $u_1, \ldots, \ldots u_{k-1}$ sono generatori di V. **C.V.D.**

Abbiamo visto quando è possibile eliminare un vettore da un insieme di generatori ottenendo ancora un insieme di generatori, ora vediamo quando è possibile aggiungere un vettore ad un insieme di vettori linearmente indipendenti ottenendo ancora un insieme di vettori linearmente indipendenti.

Osservazione 5.2.3 *In un $\mathbb{R}$-spazio vettoriale V sia dato un insieme di vettori $\{u_1, u_2, \ldots, u_k\}$ linearmente indipendenti. Allora i vettori $u_1, u_2, \ldots, u_k, u$ sono linearmente indipendenti se e solo se u non appartiene al sottospazio $\langle u_1, u_2, \ldots, u_k \rangle$.*

Dimostrazione. "$\Rightarrow$" Se $u_1, u_2, \ldots, u_k, u$ sono linearmente indipendenti, nessuno di essi è combinazione lineare degli altri, perciò, in particolare, $u \notin \langle u_1, \ldots, u_k \rangle$.

"$\Leftarrow$" Supponiamo ora che u non appartenga al sottospazio $\langle u_1, u_2, \ldots, u_k \rangle$ e supponiamo, per assurdo, che $u_1, u_2, \ldots, u_k, u$ siano linearmente dipendenti, pur essendo $u_1, u_2, \ldots, u_k$ linearmente indipendenti. Una relazione di dipendenza per $u_1, u_2, \ldots, u_k, u$ sarà data da

$$\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_k u_k + \alpha u = \mathbf{0}_V \tag{5.1}$$

con qualcuno dei numeri reali $\alpha_1, \alpha_2, \ldots, \alpha_k, \alpha$ diverso da zero. Ora α non può essere uguale a 0, poiché allora avremmo una relazione di dipendenza tra $u_1, u_2, \ldots, u_k$ che sono linearmente indipendenti, quindi $\alpha \neq 0$ e $\alpha u = -\alpha_1 u_1 - \alpha_2 u_2 - \cdots - \alpha_k u_k$. Moltiplicando tale uguaglianza per $1/\alpha$, si ottiene u come combinazione lineare di $u_1, u_2, \ldots, u_k$, contro l'ipotesi. **C.V.D.**

Finora quello che abbiamo fatto è stato da una parte, nel caso in cui avessimo un insieme di generatori, mostrare come togliere alcuni vettori e mantenere ancora un insieme di generatori, e dall'altra, nel caso in cui avessimo un insieme di vettori linearmente indipendenti, mostrare come aggiungere altri vettori e mantenere il fatto che fossero linearmente indipendenti. Ma questi due procedimenti hanno un termine? Sì: essi terminano nella individuazione di una base. Siamo al risultato culmine della teoria: ogni spazio vettoriale finitamente generato V possiede sempre una base.

Teorema 5.2.4 (**Senza dimostrazione**) *In un $\mathbb{R}$-spazio vettoriale V sia $I = \{u_1, u_2, \ldots, u_k\}$ un insieme di vettori linearmente indipendenti e sia $G = \{w_1, w_2, \ldots, w_p\}$ un insieme di generatori, con $p, k \in \mathbb{N}$. Allora $k \leq p$, cioè la cardinalità di un insieme di generatori è sempre maggiore della cardinalità di un insieme di vettori linearmente indipendenti o uguale ad essa.*

La principale conseguenza del precedente teorema è il seguente corollario:

Corollario 5.2.5 *Ogni base di uno spazio vettoriale ha la medesima cardinalità, cioè lo stesso numero di elementi.*

Dimostrazione. I vettori di una base sono nello stesso tempo linearmente indipendenti e generatori, pertanto se $B_1 = \{c_1, c_2, \ldots, c_l\}$ e $B_2 = \{b_1, b_2, \ldots, b_m\}$ sono due basi di uno spazio vettoriale V, con $l, m \in \mathbb{N}$,

allora $l \leq m$ se si pensano $c_1, c_2 \ldots, c_l$ come vettori linearmente indipendenti e $b_1, b_2, \ldots, b_m$ come generatori e $m \leq l$ se si pensano $b_1, b_2, \ldots, b_m$ come linearmente indipendenti e $c_1, c_2 \ldots, c_l$ come generatori. Quindi $l = m$.
C.V.D.

Questo corollario ci permette di introdurre la seguente definizione:

Definizione 5.2.6 *Si dice* **dimensione di uno spazio vettoriale** *la cardinalità (i.e. il numero di elementi) di una sua qualunque base. Convenzionalmente poniamo la dimensione dello spazio vettoriale banale uguale a 0.*

Teorema 5.2.7 *Ogni spazio vettoriale finitamente generato ammette (almeno) una base. Più precisamente:*

1. *ogni insieme di generatori contiene almeno una base dello spazio;*

2. *ogni insieme di vettori linearmente indipendenti si può completare ad una base.*

Dimostrazione. Per quanto visto, dato un insieme di generatori di uno spazio vettoriale, si può estrarre da questo insieme una base: se i vettori di partenza sono linearmente indipendenti allora sono essi stessi una base; se invece sono linearmente dipendenti per la Proposizione 5.2.2 posso toglierne qualcuno e mantenere il fatto che siano dei generatori. Continuando finché non è più possibile eliminare vettori si ottiene un insieme di generatori linearmente indipendenti, quindi una base.

Analogamente in uno spazio vettoriale V ogni insieme di vettori linearmente indipendenti $u_1, \ldots, u_s$ si può considerare parte di una base: infatti si hanno due possibilità: o $\langle u_1, \ldots, u_s \rangle = V$ e allora $u_1, \ldots, u_s$ sono pure generatori e quindi una base di V, oppure $\langle u_1, \ldots, u_s \rangle$ è contenuto propriamente in V, e dunque esiste $u \in V$, $u \notin \langle u_1, \ldots, u_s \rangle$ e per l'Osservazione 5.2.3 $u_1, \ldots, u_s, u$ sono linearmente indipendenti. Così $\langle u_1, \ldots, u_s \rangle \subset \langle u_1, \ldots, u_s, u \rangle$. Continuando in questo modo si ottiene una base di V nel momento in cui si ha un numero di vettori linearmente indipendenti uguale alla dimensione di V. Diremo allora che ogni insieme di vettori linearmente indipendenti di V si può completare in una base di V.

Osservazione 5.2.8 Sia V uno spazio vettoriale di dimensione n. Allora:

i) n è il massimo numero di vettori linearmente indipendenti in V. Infatti, se V contenesse k vettori linearmente indipendenti, con $k > n$, si potrebbe costruire una base di V contenente i k vettori dati ottenendo una base contenente più di n vettori e questo non è possibile.

ii) n è il minimo numero di vettori che servono a generare V. Infatti, se $k < n$ vettori generassero V, eventualmente eliminando vettori linearmente dipendenti si potrebbe costruire una base contenente meno di n vettori e questo non è possibile.

Esempi 5.2.9 1. Consideriamo in $\mathbb{R}^2$ i vettori $e_1 = (1, 0)$ ed $e_2 = (0, 1)$. Abbiamo già visto che essi generano $\mathbb{R}^2$ dal momento che per ogni vettore $v = (a, b)$ di $\mathbb{R}^2$ si ha: $v = (a, b) = a(1, 0) + b(0, 1)$. Inoltre e_1 ed e_2 sono linearmente indipendenti, infatti

$$\alpha e_1 + \beta e_2 = (\alpha, \beta) = (0, 0) \quad \Leftrightarrow \quad \alpha = 0 = \beta.$$

Dunque $B = \{e_1, e_2\}$ è una base di $\mathbb{R}^2$ e $\mathbb{R}^2$ ha pertanto dimensione due. La base B si dice base canonica di $\mathbb{R}^2$. Si noti che le coordinate di un vettore nella base canonica coincidono con le componenti del vettore.

Analogamente, per ogni $n \in \mathbb{N}$ definiamo i seguenti vettori di $\mathbb{R}^n$: $e_1 = (1, 0, \dots, 0)$, $e_2 = (0, 1, 0, \dots, 0)$, $\dots, e_n = (0, 0, \dots, 0, 1)$. Allora $\{e_1, e_2, \dots, e_n\}$ è una base di $\mathbb{R}^n$, detta base canonica. La dimensione di $\mathbb{R}^n$ è dunque n. Dato il vettore $(x_1, x_2, \dots, x_n)$ di $\mathbb{R}^n$ vale: $(x_1, x_2, \dots, x_n) = x_1e_1 + \cdots + x_ne_n$.

2. Se un sottospazio L di uno spazio vettoriale V di dimensione n ha dimensione n allora coincide con lo spazio ambiente V. Infatti se L ha dimensione n allora esso contiene n vettori linearmente indipendenti: tali vettori, dal momento che le operazioni in L sono quelle di V, risultano linearmente indipendenti anche in V e sono quindi una base di V stesso. Dunque lo spazio da essi generato è $L = V$.

Osservazione 5.2.10 Sia V uno spazio vettoriale finitamente generato. V ammette quindi una base e supponiamo che V abbia dimensione $n \geq 1$. Sia T un sottospazio non banale di V. Allora anche T è finitamente generato ed ha dimensione $n_1 \leq n$. Mostriamo come costruire una base di T. Sia $v_1 \in T$ un vettore non nullo. Vi sono due possibilità: o $\langle v_1 \rangle = T$ e allora v_1 è un insieme di generatori di T linearmente indipendente, cioè una base, oppure

$\langle v_1 \rangle$ è contenuto propriamente in T. Allora esiste in T un vettore v_2 che non è combinazione lineare di v_1 e quindi v_1, v_2 sono linearmente indipendenti in T e quindi in V. Allora abbiamo ancora due possibilità: $\langle v_1, v_2 \rangle = T$, e allora $\{v_1, v_2\}$ è un insieme di generatori linearmente indipendenti di T e quindi una sua base, oppure $\langle v_1, v_2 \rangle \subsetneq T$ e possiamo continuare il ragionamento. Il procedimento deve avere una fine poiché i vettori linearmente indipendenti in T sono linearmente indipendenti anche in V, e quindi sono al più n.

Osservazione 5.2.11 Sia $n = \dim(V)$ e siano $\{v_1, \dots, v_n\}$ vettori di V.

i) Se $\{v_1, \dots, v_n\}$ generano V, allora $\{v_1, \dots, v_n\}$ sono automaticamente linearmente indipendenti. Se non lo fossero, infatti, uno di essi sarebbe combinazione lineare degli altri. Scartandolo otterremmo un insieme di generatori di V contenente meno di n elementi e questo non è possibile.

ii) Se $\{v_1, \dots, v_n\}$ sono linearmente indipendenti, allora automaticamente generano V. Se così non fosse, infatti, esisterebbe un vettore v di V non appartenente al sottospazio $\langle v_1, \dots, v_n \rangle$. I vettori $\{v_1, \dots, v_n, v\}$ sarebbero dunque $n+1$ vettori linearmente indipendenti in uno spazio vettoriale di dimensione n e questo non è possibile.

In conclusione, se la dimensione di uno spazio vettoriale V è nota ed è uguale ad n, per stabilire se n vettori dati di V individuano una base non c'è bisogno di verificare sia il fatto che generino V sia la loro lineare indipendenza perchè ognuna di queste due proprietà implica l'altra.

Definizione 5.2.12 *Sia V uno spazio vettoriale reale di dimensione n, e sia $\mathcal{B} = \{u_1, \dots, u_n\}$ una base di V. Allora per ogni vettore $v \in V$ sono univocamente determinati gli scalari $\lambda_i \in \mathbb{R}$ tali che $v = \sum_{i=1}^{n} \lambda_i u_i$; la n-upla $(\lambda_1, \dots, \lambda_n)$ si dice n-upla delle coordinate di v rispetto alla base $\mathcal{B}$.*

Per indicare che $(\lambda_1, \dots, \lambda_n)$ è la n-upla di coordinate del vettore v rispetto alla base $\mathcal{B}$, scriveremo $v = (\lambda_1, \dots, \lambda_n)_{\mathcal{B}}$. Possiamo allora definire l'applicazione

$$f : V \longrightarrow \mathbb{R}^n$$

che associa ad ogni vettore $v \in V$ la n-upla delle sue coordinate nella base $\mathcal{B}$:

$$f(v) = (\lambda_1, \lambda_2, \dots, \lambda_n).$$

L'applicazione f è ben definita perché $\mathcal{B}$ è una base; è iniettiva, infatti se due vettori hanno le stesse coordinate in una base fissata vuol dire che essi coincidono. Inoltre f è suriettiva, infatti, preso un qualsiasi elemento $(\sigma_1, \ldots, \sigma_n) \in \mathbb{R}^n$ e posto $w = \sigma_1 u_1 + \sigma_2 u_2 + \cdots + \sigma_n u_n$, w è un elemento di V e $f(w) = (\sigma_1, \ldots, \sigma_n)$. L'applicazione f è dunque biunivoca.

Esempio 5.2.13 Consideriamo l'insieme $L = \{(x, y, z) \in \mathbb{R}^3 \mid 2x - y + z = 0\}$. Si verifichi per esercizio che L è un sottospazio vettoriale di $\mathbb{R}^3$. Vogliamo determinare una base di L. Essendo un sottospazio di uno spazio di dimensione 3, L potrà avere dimensione: 0,1,2,3. Esso non ha dimensione 3 perché allora coinciderebbe con $\mathbb{R}^3$ ma, ad esempio, il vettore $(1, 1, 1) \in \mathbb{R}^3$ non appartiene a L dal momento che le sue coordinate non soddisfano l'equazione di L. Del resto L non è banale poiché contiene almeno il vettore $(1, 2, 0) \neq (0, 0, 0)$. Dunque L avrà dimensione 1 o 2. Abbiamo già notato che $(1, 2, 0)$ è un vettore di L. Osserviamo che $(0, 1, 1)$ è un altro vettore che appartiene a L. Dal momento che $(1, 2, 0)$ e $(0, 1, 1)$ sono linearmente indipendenti, $\dim L = 2$ e $\{(1, 2, 0), (0, 1, 1)\}$ è una base di L.

Osserviamo che, sebbene abbiamo costruito una funzione biiettiva tra ogni spazio vettoriale di dimensione n e $\mathbb{R}^n$, abbiamo definito una base "canonica" solo per $\mathbb{R}^n$.

Esempio 5.2.14 Determiniamo una base dello spazio vettoriale $\mathcal{M}_{2,3}(\mathbb{R})$ delle matrici 2×3 ad entrate reali. Consideriamo allora le seguenti matrici:

$$e_{11} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \; e_{12} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \; e_{13} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

$$e_{21} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \; e_{22} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \; e_{23} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Si verifica facilmente che e_{11}, e_{12}, e_{13}, e_{21}, e_{22}, e_{23} sono linearmente indipendenti: ognuna ha una entrata diversa da zero in una posizione in cui tutte le altre matrici hanno una entrata nulla. Inoltre ogni matrice di $\mathcal{M}_{2,3}(\mathbb{R})$ è della forma

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} = a_{11}e_{11} + a_{12}e_{12} + a_{13}e_{13} + a_{21}e_{21} + a_{22}e_{22} + a_{23}e_{23}$$

quindi e_{11}, e_{12}, e_{13}, e_{21}, e_{22}, e_{23} generano $\mathcal{M}_{2,3}(\mathbb{R})$. La dimensione di $\mathcal{M}_{2,3}(\mathbb{R})$ è pertanto 6. Analogamente si può dimostrare, in generale, che la dimensione di $\mathcal{M}_{n,m}(\mathbb{R})$ è nm e che una base di $\mathcal{M}_{n,m}(\mathbb{R})$ è data dall'insieme delle matrici $\{e_{ij}, i = 1, \dots, n, j = 1, \dots, m\}$ dove e_{ij} è la matrice avente tutte le entrate nulle tranne quella di posto i, j che è uguale a 1.

5.3 Strumenti di calcolo

Concentriamoci ora sulla seguente domanda: dati k vettori $v_1, \dots, v_k$ di uno spazio vettoriale V, come facciamo a stabilire se questi sono o meno linearmente indipendenti? Si può rispondere a questa domanda utilizzando diversi metodi:

Metodo 1 Possiamo certamente utilizzare la definizione e cercare quindi di stabilire se è possibile scrivere $\mathbf{0}_V$ come combinazione lineare di $v_1, \dots, v_k$ con qualche coefficiente diverso da 0. Abbiamo utilizzato questo metodo nell'Esempio 5.1.2(2.). Vediamo un altro esempio:

Esempio 5.3.1 Stabilire se i vettori $v_1 = (1, 2, 0, 1)$, $v_2 = (-1, 1, 1, 1)$, $v_3 = (3, 3, -1, 1)$, $v_4 = (-3, 0, 2, 1)$ di $\mathbb{R}^4$ sono linearmente indipendenti.

Usando il Metodo 1, dobbiamo scrivere il vettore $(0, 0, 0, 0)$ come combinazione lineare dei vettori v_1, v_2, v_3, v_4:

$$a(1, 2, 0, 1) + b(-1, 1, 1, 1) + c(3, 3, -1, 1) + d(-3, 0, 2, 1) = (0, 0, 0, 0) \Leftrightarrow$$

$$(a - b + 3c - 3d, 2a + b + 3c, b - c + 2d, a + b + c + d) = (0, 0, 0, 0) \Leftrightarrow$$

$$\begin{cases} a - b + 3c - 3d = 0 \\ 2a + b + 3c = 0 \\ b - c + 2d = 0 \\ a + b + c + d = 0 \end{cases}$$

Abbiamo dunque trovato un sistema lineare omogeneo nelle quattro incognite a, b, c, d. Per risolvere il sistema lineare trovato costruiamo la matrice completa associata al sistema e riduciamola a scala:

$$\left(\begin{array}{cccc|c} 1 & -1 & 3 & -3 & 0 \\ 2 & 1 & 3 & 0 & 0 \\ 0 & 1 & -1 & 2 & 0 \\ 1 & 1 & 1 & 1 & 0 \end{array}\right) \to \left(\begin{array}{cccc|c} 1 & -1 & 3 & -3 & 0 \\ 0 & 3 & -3 & 6 & 0 \\ 0 & 1 & -1 & 2 & 0 \\ 0 & 2 & -2 & 4 & 0 \end{array}\right) \to \left(\begin{array}{cccc|c} 1 & -1 & 3 & -3 & 0 \\ 0 & 1 & -1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

La matrice incompleta ridotta a scala ha rango 2 (uguale al rango della matrice completa), pertanto il sistema ha infinite soluzioni. Questo significa che esistono infinite quaterne (a, b, c, d) di coefficienti tali che $\mathbf{0}_{\mathbb{R}^4} = av_1 + bv_2 + cv_3 + dv_4$. Perciò i vettori v_1, v_2, v_3 e v_4 sono linearmente dipendenti. (Al contrario, se il sistema trovato avesse avuto l'unica soluzione $(0, 0, 0, 0)$, allora i vettori v_1, v_2, v_3 e v_4 sarebbero stati linearmente indipendenti).

Metodo 2 Abbiamo notato nella Osservazione 5.1.6 che se i vettori $v_1, \ldots, v_k$ sono linearmente indipendenti allora qualsiasi sottoinsieme di $\{v_1, \ldots, v_k\}$ continua ad essere formato da vettori linearmente indipendenti. Per stabilire dunque se certi vettori assegnati $v_1, \ldots, v_k$ sono linearmente indipendenti si può procedere nel modo seguente:

(1) consideriamo il vettore v_1: questo è linearmente indipendente se e solo se è non nullo. Dunque se $v_1 = \mathbf{0}_V$ allora i vettori assegnati sono linearmente dipendenti. Se, invece, $v_1 \neq \mathbf{0}_V$ allora passiamo al punto (2);

(2) consideriamo i vettori $\{v_1, v_2\}$ e ci chiediamo se v_2 è linearmente dipendente da v_1 cioè se v_2 è un multiplo di v_1. In caso affermativo possiamo concludere che i vettori assegnati sono linearmente dipendenti. Se, invece, $v_2 \neq \alpha v_1$ per qualsiasi $\alpha \in \mathbb{R}$, allora passiamo al punto (3);

(3) consideriamo i vettori $\{v_1, v_2, v_3\}$ e ci chiediamo se v_3 è linearmente dipendente da v_1 e v_2 cioè se v_3 è combinazione lineare di v_1 e v_2. In caso affermativo possiamo concludere che i vettori assegnati sono linearmente dipendenti. Se, invece, $v_3 \notin \langle v_1, v_2 \rangle$, allora andiamo avanti e procediamo in modo analogo fino ad aver analizzato uno dopo l'altro tutti i sottoinsiemi $\{v_1, \ldots, v_i\}$ per $i = 1, \ldots, k$.

Esempio 5.3.2 Stabilire se i vettori $v_1 = (1, 2, 0, 1)$, $v_2 = (-1, 1, 1, 1)$, $v_3 = (3, 3, -1, 1)$, $v_4 = (-3, 0, 2, 1)$ sono linearmente indipendenti.

Utilizziamo questa volta il Metodo 2 per stabilire la lineare dipendenza dei vettori $v_1, \ldots, v_4$. Il vettore v_1 è non nullo, quindi linearmente indipendente. Il vettore v_2 non è un multiplo del vettore v_1 (ad esempio perché la sua terza coordinata è non nulla mentre la terza coordinata di v_1 è nulla), quindi v_1 e v_2 sono linearmente indipendenti. Ora devo stabilire se v_3 è combinazione lineare dei vettori v_1 e v_2, cioè se esistono $\alpha, \beta \in \mathbb{R}$ tali che $v_3 = \alpha v_1 + \beta v_2$. Scriviamo per esteso la relazione appena scritta:

$$(3, 3, -1, 1) = \alpha(1, 2, 0, 1) + \beta(-1, 1, 1, 1) \Longrightarrow$$

$$(3,3,-1,1) = (\alpha-\beta, 2\alpha+\beta, \beta, \alpha+\beta) \Longrightarrow$$
$$\begin{cases} \alpha-\beta=3 \\ 2\alpha+\beta=3 \\ \beta=-1 \\ \alpha+\beta=1 \end{cases}$$

Si vede facilmente che il sistema trovato ha soluzione $\beta=-1, \alpha=2$. Questo significa che possiamo scrivere il vettore v_3 come combinazione lineare dei vettori v_1 e v_2, pertanto i vettori v_1, v_2, v_3 sono linearmente dipendenti. Di conseguenza anche i vettori v_1, v_2, v_3, v_4 sono linearmente dipendenti.

I Metodi 1 e 2 appena illustrati sono più o meno equivalenti da un punto di vista dei calcoli ed entrambi piuttosto laboriosi quando i vettori assegnati sono molti e/o dipendenti da parametri. Esiste in effetti un metodo decisamente più efficace dei precedenti per stabilire se certi vettori sono o meno linearmente indipendenti. Per illustrare questo terzo metodo abbiamo però bisogno di introdurre una nuova definizione e di riflettere su di essa.

Data una matrice $A \in M_{m,n}(\mathbb{R})$ possiamo leggere le righe di A come vettori di $\mathbb{R}^n$ e le sue colonne come vettori di $\mathbb{R}^m$. Questo è il punto di vista adottato nella seguente definizione:

Definizione 5.3.3 *Data una matrice $A \in M_{m,n}(\mathbb{R})$ si chiama rango righe di A il massimo numero di righe linearmente indipendenti di A (come vettori di $\mathbb{R}^n$); si chiama rango colonne di A il massimo numero di colonne linearmente indipendenti di A (come vettori di $\mathbb{R}^m$).*

Il rango righe ed il rango colonne di una matrice non sono indipendenti, al contrario, il legame tra questi due numeri è forte ed espresso dal seguente teorema che NON dimostreremo:

Teorema 5.3.4 *Data una matrice $A \in M_{m,n}(\mathbb{R})$, il rango righe ed il rango colonne di A coincidono.*

Ha senso dunque assegnare al rango righe e al rango colonne un unico nome. Chiameremo semplicemente *rango* di A sia il rango righe che il rango colonne di A e lo indicheremo con $rg(A)$.

Osservazione 5.3.5 Nella Lezione 1 abbiamo definito il rango di una matrice in forma a scala per righe. Naturalmente la definizione di rango (Definizione 1.1.4) data per le matrici a scala coincide con quella più generale appena fornita. È facile infatti mostrare che in una matrice a scala le righe non nulle sono linearmente indipendenti: basta usare il Metodo 2 partendo dall'ultima riga non nulla e risalendo verso l'alto nella matrice. Ad esempio, consideriamo la matrice

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

La matrice A è in forma a scala con 3 pivot (1, -5, -2) e ha pertanto rango 3. Mostriamo che i vettori riga $(0,0,0,-2), (0,0,-5,3), (1,2,3,-1)$ sono linearmente indipendenti: il vettore $(0,0,0,-2)$ è non nullo e il vettore $(0,0,-5,3)$ non è un multiplo di $(0,0,0,-2)$ perché la sua terza coordinata è -5 mentre la terza coordinata di $(0,0,0,-2)$ è nulla. Infine il vettore $(1,2,3,-1)$ non è combinazione lineare di $(0,0,0,-2)$ e $(0,0,-5,3)$ perché la sua prima coordinata è 1 mentre la prima coordinata di $(0,0,0,-2)$ e $(0,0,-5,3)$ è nulla.

Osservazione 5.3.6 Il rango di una matrice $A \in M_{m,n}(\mathbb{R})$ fornisce, per definizione, la dimensione del sottospazio di $\mathbb{R}^n$ generato dalle righe di A e la dimensione del sottospazio di $\mathbb{R}^m$ generato dalle colonne di A.

Proposizione 5.3.7 *Sia $A \in M_{m,n}(\mathbb{R})$. Le operazioni elementari sulle righe di A preservano il sottospazio di $\mathbb{R}^n$ generato dalle righe di A quindi preservano il rango.*

Dimostrazione Sia $A \in M_{m,n}(\mathbb{R})$ ed indichiamo con $v_1, \dots, v_m$ le sue righe, pensate come vettori di $\mathbb{R}^n$. Vogliamo mostrare che le operazioni elementari sulle righe della matrice non modificano il sottospazio $\langle v_1, \dots, v_m \rangle$. Ora, scambiare due righe di A certamente non cambia questo sottospazio. Analogamente sostituire ad una riga v_i un suo multiplo non nullo preserva il sottospazio, vale a dire:

$$\langle v_1, \dots, v_i, \dots, v_m \rangle = \langle v_1, \dots, \alpha v_i, \dots, v_m \rangle \quad \forall \alpha \in \mathbb{R} \setminus \{0\}.$$

Infatti un vettore genera tutto e solo ciò che è generato da qualsiasi suo multiplo non nullo. Si tratta infine di dimostrare che sostituendo alla riga i-esima la somma della riga i-esima con un multiplo della j-esima, ancora lo spazio generato dalle righe non cambia. Poiché questa operazione coinvolge solo la i-esima e la j-esima riga, si tratta di dimostrare che, per $i \neq j$,

$$\langle v_j, v_i \rangle = \langle v_j, v_i + \alpha v_j \rangle.$$

Naturalmente

$$\langle v_j, v_i \rangle \supseteq \langle v_j, v_i + \alpha v_j \rangle,$$

infatti ogni elemento di $\langle v_j, v_i + \alpha v_j \rangle$ è della forma $av_j + b(v_i + \alpha v_j) = (a+b\alpha)v_j + bv_i$, per qualche $a, b \in \mathbb{R}$, e quindi è una combinazione lineare di v_i e v_j, cioè appartiene a $\langle v_j, v_i \rangle$. Viceversa, dato un elemento $hv_i + kv_j \in \langle v_j, v_i \rangle$ con $h, k \in \mathbb{R}$, possiamo scrivere $hv_i + kv_j = h(v_i + \alpha v_j) + (k - h\alpha)v_j$ che è una combinazione lineare di $v_i + \alpha v_j$ e v_j e quindi appartiene a $\langle v_j, v_i + \alpha v_j \rangle$. Abbiamo dunque mostrato che $\langle v_j, v_i \rangle \subseteq \langle v_j, v_i + \alpha v_j \rangle$, quindi $\langle v_j, v_i \rangle = \langle v_j, v_i + \alpha v_j \rangle$. $\square$

Siamo ora pronti per illustrare il terzo metodo per stabilire se certi vettori assegnati $\{v_1, \dots, v_k\}$ sono o meno linearmente indipendenti:

Metodo 3. Si costruisce la matrice A che ha i vettori $v_1, \dots, v_k$ come righe. Il rango di A fornisce, per definizione, il numero di vettori linearmente indipendenti tra $v_1, \dots, v_k$. Per calcolare tale rango, grazie alla Proposizione 5.3.7, si riduce la matrice A in forma a scala attraverso l'algoritmo di Gauss e si calcola il rango della matrice ridotta. Notiamo anche che, sempre per la Proposizione 5.3.7, lo spazio generato dalle righe di A coincide con lo spazio generato dalle righe della matrice ridotta, proprietà, questa, che in generale può rivelarsi molto utile.

Esempio 5.3.8 Stabilire se i vettori $v_1 = (1, 2, 0, 1)$, $v_2 = (-1, 1, 1, 1)$, $v_3 = (3, 3, -1, 1)$, $v_4 = (-3, 0, 2, 1)$ sono linearmente indipendenti.

Utilizziamo infine il Metodo 3 per stabilire la lineare dipendenza dei vettori $v_1, \dots, v_4$. Si tratta di ridurre in forma a scala la matrice che ha sulle righe i vettori $v_1, \dots, v_4$:

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

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

Si ha dunque: $rg(A) = rg(A') = 2$ il che significa che solo due tra i vettori $v_1, \ldots, v_4$ sono linearmente indipendenti. Perciò i vettori $v_1, \ldots, v_4$ sono linearmente dipendenti. Sappiamo, inoltre, che $\langle v_1, \ldots, v_4 \rangle = \langle (1, 2, 0, 1), (0, 3, 1, 2) \rangle$.

Per meglio illustrare l'efficacia del Metodo 3, facciamo ancora un altro esempio:

Esempio 5.3.9 Stabilire per quali valori del parametro reale k i vettori $(1, k, 1)$, $(k, 1, -k)$, $(k, k, 1)$ sono linearmente indipendenti.

Per rispondere a questo quesito basta determinare i valori di $k \in \mathbb{R}$ tali che

$$rg \begin{pmatrix} 1 & k & 1 \\ k & 1 & -k \\ k & k & 1 \end{pmatrix} = 3.$$

Si ha:

$$rg \begin{pmatrix} 1 & k & 1 \\ k & 1 & -k \\ k & k & 1 \end{pmatrix} = rg \begin{pmatrix} 1 & k & 1 \\ 0 & 1-k^2 & -2k \\ 0 & k-1 & 1+k \end{pmatrix} \to rg \begin{pmatrix} 1 & k & 1 \\ 0 & k-1 & 1+k \\ 0 & 0 & k^2+1 \end{pmatrix}.$$

Notiamo che per ogni valore di $k \in \mathbb{R}$ il numero $k^2 + 1$ è positivo, pertanto per ogni $k \neq 1$ il rango della matrice è uguale a 3, cioè, per ogni $k \neq 1$ i vettori assegnati sono linearmente indipendenti. Per $k = 1$, invece, il rango della matrice trovata è uguale a 2, quindi i vettori assegnati sono linearmente dipendenti.

5.4 Esercizi svolti

Esercizio 5.4.1 Si verifichi che $\{(1, 2, 3), (0, 1, 0), (0, 0, 1)\}$ è una base di $\mathbb{R}^3$.

Svolgimento. Dobbiamo verificare che

1) i vettori dati sono linearmente indipendenti;

2) i vettori dati generano $\mathbb{R}^3$.

In effetti, poiché sappiamo che $\dim(\mathbb{R}^3) = 3$, basterà verificare una o l'altra proprietà (vedi Osservazione 5.2.11). Verifichiamo, ad esempio, che i tre vettori dati sono linearmente indipendenti. Cominciamo dal vettore $(0, 0, 1)$: esso è non nullo perciò linearmente indipendente; $(0, 1, 0)$ non è un multiplo di $(0, 0, 1)$, perciò $(0, 0, 1)$ e $(0, 1, 0)$ sono linearmente indipendenti. Infine $(1, 2, 3)$ non è una combinazione lineare di $(0, 1, 0)$ e $(0, 0, 1)$ dal momento che combinando linearmente questi due vettori si ottiene un vettore la cui prima componente è nulla. Possiamo quindi concludere che i tre vettori assegnati sono linearmente indipendenti e, dunque, individuano una base di $\mathbb{R}^3$.

Esercizio 5.4.2 In $\mathcal{M}_2(\mathbb{R})$ si considerino le matrici $A = \begin{pmatrix} 3 & 0 \\ 2 & -1 \end{pmatrix}$ e $B = \begin{pmatrix} -5 & 1 \\ 4 & 3 \end{pmatrix}$. Stabilire se A e B sono linearmente indipendenti. Eventualmente completare l'insieme $\{A, B\}$ in una base di $\mathcal{M}_2(\mathbb{R})$.

Svolgimento. Le matrici A e B sono linearmente indipendenti perché non sono una multipla dell'altra. Completare l'insieme $\{A, B\}$ in una base di $\mathcal{M}_2(\mathbb{R})$ significa individuare due elementi C, $D \in \mathcal{M}_2(\mathbb{R})$ tali che $\{A, B, C, D\}$ sia una base di $\mathcal{M}_2(\mathbb{R})$. Procedendo come indicato in 5.2.3 si verifichi che possiamo scegliere $C = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ e $D = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$.

Esercizio 5.4.3 Dato l'insieme $\{(0, 1, 1), (1, 0, -1), (2, 3, 4), (2, 1, 0)\}$ di vettori di $\mathbb{R}^3$, stabilire quale delle seguenti affermazioni è vera:

(i) Ogni insieme che contiene quello dato genera $\mathbb{R}^3$.

(ii) Esiste un insieme che contiene quello dato ed è costituito da vettori linearmente indipendenti.

(iii) L'insieme dato è una base di $\mathbb{R}^3$.

Estrarre, se possibile, una base di $\mathbb{R}^3$ dall'insieme dato.

Svolgimento. Ricordiamo innanzitutto che la dimensione di $\mathbb{R}^3$ è 3 quindi 3 è la cardinalità di ogni base di $\mathbb{R}^3$ e quindi il massimo numero di vettori linearmenti indipendenti di $\mathbb{R}^3$. Per questo possiamo immediatamente asserire che le affermazioni (ii) e (iii) sono false. Per convincerci ora della veridicità della affermazione (i) basta verificare che i vettori $(0,1,1)$, $(1,0,-1),(2,3,4),(2,1,0)$ generano $\mathbb{R}^3$. Per questo basta procedere come nell'esercizio 4.1.1. Infine si può mostrare che l'insieme $\{(0,1,1),(1,0,-1),(2,1,0)\}$ è una base di $\mathbb{R}^3$.

Esercizio 5.4.4 Sia $W = \{(2s+t, s-t, s+t, s+2t) \mid s,t \in \mathbb{R}\}$.

(i) Verificare che W è un sottospazio vettoriale di $\mathbb{R}^4$;

(ii) determinare una base $\mathcal{B}$ di W;

(iii) completare $\mathcal{B}$ ad una base $\tilde{\mathcal{B}}$ di $\mathbb{R}^4$.

Svolgimento.

(i) Siano $v = (2s+t, s-t, s+t, s+2t)$ e $w = (2r+p, r-p, r+p, r+2p)$, con $s,t,r,p \in \mathbb{R}$, due elementi di W. Allora $v+w$ è ancora un elemento di W, infatti $v+w = (2s+t+2r+p, s-t+r-p, s+t+r+p, s+2t+r+2p) = (2(s+r)+t+p, (s+r)-(t+p), (s+r)+(t+p), (s+r)+2(t+p))$. Analogamente per ogni $\lambda \in \mathbb{R}$ e per ogni $v \in W$, $\lambda v \in W$. Pertanto W è un sottospazio vettoriale di $\mathbb{R}^4$.

(ii) Individuiamo innanzitutto un insieme di generatori di W. Per questo osserviamo che ogni vettore di W è della forma $(2s+t, s-t, s+t, s+2t) = s(2,1,1,1) + t(1,-1,1,2)$ il che ci consente di affermare che i vettori $(2,1,1,1)$ e $(1,-1,1,2)$ generano W. D'altra parte i vettori $(2,1,1,1)$, $(1,-1,1,2)$ sono linearmente indipendenti (non sono uno multiplo dell'altro!) e quindi individuano una base di W.

(iii) Indicata con $\mathcal{B} = \{(2,1,1,1),(1,-1,1,2)\}$ la base di W individuata in (ii), per ottenere una base $\tilde{\mathcal{B}}$ di $\mathbb{R}^4$ possiamo aggiungere alla base $\mathcal{B}$ i vettori $(0,0,1,0)$ e $(0,0,0,1)$. Si verifica infatti facilmente che i vettori $(2,1,1,1),(1,-1,1,2)$, $(0,0,1,0)$ e $(0,0,0,1)$ sono linearmente indipendenti. Dal momento che $\dim(\mathbb{R}^4) = 4$ questo basta per concludere che essi individuano una base di $\mathbb{R}^4$.

Esercizio 5.4.5 Sia $\mathbb{R}^{\leq 3}[x]$ lo spazio vettoriale reale dei polinomi a coefficienti reali nella variabile x di grado minore o uguale a 3.

(i) Mostrare che l'insieme dei monomi $\{1, x, x^2, x^3\}$ è una base di $\mathbb{R}^{\leq 3}[x]$. Dedurre che la dimensione di $\mathbb{R}^{\leq 3}[x]$ è 4.

(ii) I vettori $2x^2+1$, $2x+1$, x^3 sono linearmente indipendenti? Completare, se possibile, l'insieme $\{2x^2+1, 2x+1, x^3\}$ in una base di $\mathbb{R}^{\leq 3}[x]$.

(iii) Esistono basi di $\mathbb{R}^{\leq 3}[x]$ costituite da polinomi di grado 3? In caso affermativo esibire un esempio.

(iv) Esistono basi di $\mathbb{R}^{\leq 3}[x]$ costituite da polinomi di grado minore o uguale a 2? In caso affermativo esibire un esempio.

Svolgimento.

(i) Un generico elemento di $\mathbb{R}^{\leq 3}[x]$ è un polinomio $p(x)$ della forma $p(x) = a_0 + a_1x + a_2x^2 + a_3x^3 = a_0 \cdot 1 + a_1 \cdot x + a_2 \cdot x^2 + a_3 \cdot x^3$. Dunque $\mathcal{B} = \{1, x, x^2, x^3\}$ è un insieme di generatori di $\mathbb{R}^{\leq 3}[x]$. Del resto una combinazione lineare $a_0 + a_1x + a_2x^2 + a_3x^3$ di $1, x, x^2, x^3$ è uguale al polinomio nullo se e solo se tutti i coefficienti a_0, a_1, a_2, a_3 sono nulli. Questo dimostra che i vettori $1, x, x^2, x^3$ sono linearmente indipendenti e quindi individuano una base di $\mathbb{R}^{\leq 3}[x]$. La dimensione di $\mathbb{R}^{\leq 3}[x]$ è pertanto uguale a 4.

(ii) Sia ora $\alpha(2x^2+1)+\beta(2x+1)+\gamma x^3 = 0$, cioè $\alpha+\beta+2\beta x+2\alpha x^2+\gamma x^3 = 0$. Allora, necessariamente, $\alpha = \beta = \gamma = 0$, pertanto i vettori $2x^2 + 1$, $2x + 1$, x^3 sono linearmente indipendenti. Per completare l'insieme $S = \{2x^2+1, 2x+1, x^3\}$ ad una base di $\mathbb{R}^{\leq 3}[x]$ basta allora aggiungere all'insieme S un polinomio di $\mathbb{R}^{\leq 3}[x]$ che non sia una combinazione lineare di $2x^2 + 1$, $2x + 1$, x^3, ad esempio il polinomio 1 (verificare!). Così l'insieme $\{1, 2x^2 + 1, 2x + 1, x^3\}$ è una base di $\mathbb{R}^{\leq 3}[x]$.

(iii) L'insieme $X = \{x^3, x^3 + x^2, x^3 + x, x^3 + 1\}$ è un esempio di una base di $\mathbb{R}^{\leq 3}[x]$ costituita da polinomi di grado 3. Per rendersi conto che si tratta di una base basta osservare che gli elementi della base $\mathcal{B}$ di (i) si ottengono facilmente come combinazione lineare degli elementi di X. Dunque l'insieme X genera $\mathbb{R}^{\leq 3}[x]$. Dal momento che X ha 4 elementi e che $\dim(\mathbb{R}^{\leq 3}[x]) = 4$, X è una base di $\mathbb{R}^{\leq 3}[x]$.

(iv) Non esiste una base di $\mathbb{R}^{\leq 3}[x]$ costituita da polinomi di grado minore o uguale a 2. Infatti combinando linearmente polinomi di grado minore o uguale a 2 non è possibile ottenere polinomi di grado 3.

Esercizio 5.4.6 Calcolare la dimensione del sottospazio di $\mathbb{R}^4$ generato dai vettori $v_1 = (2, 1, 0, 3)$, $v_2 = (-5, 3, 4, 0)$, $v_3 = (-1, -1, 0, 2)$, $v_4 = (3, 2, 0, 1)$.

Svolgimento. I vettori v_1 e v_2 sono certo linearmente indipendenti dal momento che non sono uno multiplo dell'altro. Il vettore v_3 è combinazione lineare dei vettori v_1 e v_2? Esistono, cioè, α e $\beta \in \mathbb{R}$ tali che $(-1,-1,0,2) = \alpha(2,1,0,3) + \beta(-5,3,4,0)$? Si tratta di stabilire se esistono α e $\beta \in \mathbb{R}$ tali che

$$\begin{cases} -1 = 2\alpha - 5\beta \\ -1 = \alpha + 3\beta \\ 0 = 4\beta \\ 2 = 3\alpha \end{cases}.$$

Il sistema individuato non ha soluzioni ($\alpha = \frac{2}{3},\ \beta = 0,\ \alpha = -1!$), quindi i vettori v_1, v_2, v_3 sono linearmente indipendenti. Resta, infine, da stabilire se v_4 è combinazione lineare dei vettori v_1, v_2, v_3. Si vede facilmente che $v_4 = v_1 - v_3$. Dunque $\langle v_1, v_2, v_3, v_4\rangle = \langle v_1, v_2, v_3\rangle$, pertanto $\dim\langle v_1, v_2, v_3, v_4\rangle = 3$.

Esercizio 5.4.7 Sia W il sottospazio vettoriale di $\mathbb{R}^5$ definito da $W = \{(x_1, x_2, x_3, x_4, x_5) \mid x_1 - x_2 - 2x_5 = 0,\ x_3 + x_4 + x_5 = 0\}$. Si determini una base $\mathcal{B}$ di W e si calcolino le coordinate del vettore $v = (-4, 0, 1, 1, -2)$ rispetto alla base $\mathcal{B}$.

Svolgimento. Osserviamo innanzitutto che il nostro spazio W non è tutto lo spazio vettoriale ambiente poiché, ad esempio, il vettore $(0,0,0,0,1)$ non soddisfa le due equazioni. Non è neanche lo spazio nullo, poiché il vettore $(1,1,0,0,0)$ appartiene a W. Studiamo le due equazioni che lo caratterizzano: dalla prima otteniamo che $x_1 = x_2 + 2x_5$, dalla seconda $x_3 = -x_4 - x_5$, si vede quindi che scegliendo in modo indipendente x_2, x_4, x_5, possiamo calcolare di conseguenza x_1 e x_3 per ottenere una 5-upla in W. Se scegliamo $x_2 = 1, x_4 = 0, x_5 = 0$ otteniamo il vettore $v_1 = (1,1,0,0,0)$; se scegliamo $x_2 = 0, x_4 = 1, x_5 = 0$ otteniamo il vettore $v_2 = (0,0,-1,1,0)$; se scegliamo $x_2 = 0, x_4 = 0, x_5 = 1$ otteniamo il vettore $v_3 = (2,0,-1,0,1)$. I vettori v_1, v_2, v_3 sono tre vettori di W e sono linearmente indipendenti:

$$\alpha v_1 + \beta v_2 + \gamma v_3 = 0_{\mathbb{R}^5} \quad \Leftrightarrow \quad (\alpha + 2\gamma, \alpha, -\beta - \gamma, \beta, \gamma) = 0_{\mathbb{R}^5}$$

$$\Leftrightarrow \quad \alpha = \beta = \gamma = 0.$$

Quindi W può avere dimensione 3 o 4. Per determinare una base di W osserviamo che gli elementi di W sono tutti e soli i vettori di $\mathbb{R}^5$ della forma $(x_2 + 2x_5, x_2, -x_4 - x_5, x_4, x_5)$. Ciascuno di questi vettori è esprimibile nel modo seguente:

$$(x_2 + 2x_5, x_2, -x_4 - x_5, x_4, x_5) = x_2(1,1,0,0,0)+$$

$$+x_5(2,0,-1,0,1)+x_4(0,0,-1,1,0).$$

Pertanto i vettori $v_1=(1,1,0,0,0)$, $v_2=(0,0,-1,1,0)$, $v_3=(2,0,-1,0,1)$ generano W. Abbiamo così individuato una base di W: $\mathcal{B}=\{v_1,v_2,v_3\}$. In particolare W ha dimensione 3.

Per calcolare le coordinate del vettore $v=(-4,0,1,1,-2)$ rispetto alla base $\mathcal{B}$ dobbiamo determinare α, β, $\gamma\in\mathbb{R}$ tali che $v=\alpha v_1+\beta v_2+\gamma v_3$, cioè $(-4,0,1,1,-2)=(\alpha+2\gamma,\alpha,-\beta-\gamma,\beta,\gamma)$. Otteniamo:

$$\beta=1,\quad \gamma=-2,\quad \alpha=0.$$

Pertanto $v=(0,-2,1)_{\mathcal{B}}$.

5.5 Esercizi proposti

Esercizio 5.5.1 Costruire una base $\mathcal{B}$ di $\mathbb{R}^2$ diversa dalla base canonica e scrivere le coordinate dei vettori della base canonica rispetto alla base $\mathcal{B}$.

Esercizio 5.5.2 Determinare una base $\mathcal{B}=\{v_1,v_2,v_3\}$ di $\mathbb{R}^3$ soddisfacente le seguenti condizioni:

1. le coordinate del vettore $(1,1,1)$ rispetto alla base $\mathcal{B}$ sono $(1,0,0)$;
2. i vettori v_1, v_2 generano il sottospazio S di $\mathbb{R}^3$: $S=\{(x,y,z)\in\mathbb{R}^3 \mid x-y=0\}$;
3. le coordinate del vettore $(1,0,1)$ rispetto alla base $\mathcal{B}$ sono $(1,0,1)$.

La base $\mathcal{B}$ richiesta è unica?

Esercizio 5.5.3 Si considerino i vettori di $\mathbb{R}^3$: $v_1=(1,2,0)$, $v_2=(1,1,1)$, $v_3=(0,-1,1)$, $v_4=(2,3,1)$.

1. Stabilire se i vettori v_1, v_2, v_3, v_4 sono linearmente indipendenti.
2. Stabilire se i vettori v_1, v_2, v_3, v_4 generano $\mathbb{R}^3$.
3. Determinare una base del sottospazio di $\mathbb{R}^3$ generato dai vettori v_1, v_2, v_3, v_4.
4. Completare la base trovata in 3. in una base di $\mathbb{R}^3$.

Esercizio 5.5.4 Sia

$$S = \left\{ A = \begin{pmatrix} a & b & c \\ d & f & g \end{pmatrix} \in \mathcal{M}_{2,3}(\mathbb{R}) \mid a + b + d = 0,\ d + f + c = 0 \right\}.$$

1. Mostrare che S è un sottospazio vettoriale di $\mathcal{M}_{2,3}(\mathbb{R})$.
2. Determinare una base di S.
3. Determinare un sottospazio T di $\mathcal{M}_{2,3}(\mathbb{R})$ tale che $S + T = \mathcal{M}_{2,3}(\mathbb{R})$.

Esercizio 5.5.5 Nell'insieme $V = \mathbb{R}[x, y]$ dei polinomi a coefficienti reali nelle variabili x e y, con le usuali operazioni di somma di polinomi e di prodotto di un polinomio per un numero reale, si consideri il sottoinsieme S dei polinomi di grado minore o uguale a 2.

1. Dopo aver verificato che V è un $\mathbb{R}$-spazio vettoriale e che S è un suo sottospazio, calcolare la dimensione di S ed esibire una sua base $\mathcal{B}$.
2. Calcolare le coordinate del polinomio $x + y - x^2$ nella base $\mathcal{B}$.
3. Mostrare che i polinomi $x - y$, $1 + x - y$, $1 - xy$ sono linearmente indipendenti.
4. Completare l'insieme $\{x - y, 1 + x - y, 1 - xy\}$ in una base di V.

Lezione 6

Intersezione e somma di sottospazi

In questo capitolo studieremo che cosa succede effettuando le operazioni insiemistiche sui sottospazi di uno spazio vettoriale, vale a dire considerando l'intersezione e l'unione di due (o più) sottospazi.

6.1 Intersezione di sottospazi

Proposizione 6.1 *Siano S e T sottospazi di uno spazio vettoriale V. Allora $S \cap T$ è un sottospazio vettoriale di V.*

Dimostrazione. Osserviamo preliminarmente che 0_V appartiene sia ad S che a T, perciò 0_V appartiene a $S \cap T$. Mostriamo ora che $S \cap T$ è chiuso rispetto alla somma: siano v, w vettori di $S \cap T$, cioè $v, w \in S$ e $v, w \in T$. Poiché S è un sottospazio di V, S è chiuso rispetto alla somma, pertanto $v + w \in S$; analogamente $v + w \in T$, perciò $v + w \in S \cap T$. Siano ora $v \in S \cap T$ ed $\alpha \in \mathbb{R}$. Dal momento che S è un sottospazio di V, esso è chiuso rispetto al prodotto per scalari, perciò $\alpha v \in S$; analogamente $\alpha v \in T$, pertanto $\alpha v \in S \cap T$. Essendo chiuso rispetto alla somma e al prodotto per scalari, $S \cap T$ è un sottospazio di V.

Esempio 6.1.1 Siano $S = \langle (1, 1, 1), (2, 1, 0) \rangle$ e $T = \{(x, y, z) \in \mathbb{R}^3 \mid x - y + 2z = 0\}$. Vogliamo determinare $S \cap T$. Si tratta di determinare tutti gli elementi di $\mathbb{R}^3$ che appartengono sia ad S che a T, cioè tutti gli elementi di S che appartengono anche a T. Osserviamo che i vettori $(1, 1, 1)$ e $(2, 1, 0)$

di S sono linearmente indipendenti perché non sono uno multiplo dell'altro, pertanto individuano una base di S. Ogni elemento v di S è dunque della forma: $v = a(1,1,1) + b(2,1,0) = (a+2b, a+b, a)$, con $a, b \in \mathbb{R}$. Tra tutti i vettori di S cerchiamo quelli che appartengono a T, vale a dire quelli che soddisfano l'equazione di T: $S \cap T = \{(a+2b, a+b, a) \in \mathbb{R}^3 \mid a+2b-a-b+2a = 0\} = \{(a+2b, a+b, a) \in \mathbb{R}^3 \mid b = -2a\} = \{(-3a, -a, a) \in \mathbb{R}^3\} = \langle(-3,-1,1)\rangle$.

Esempio 6.1.2 Consideriamo i sottospazi $S = \{(x,0) \in \mathbb{R}^2\}$ e $T = \{(0,y) \in \mathbb{R}^2\}$ di $\mathbb{R}^2$. Notiamo che $S \cup T$ non è chiuso rispetto alla somma, infatti sommando, ad esempio, il vettore $(1,0)$ di S con il vettore $(0,1)$ di T otteniamo il vettore $(1,1)$ che non appartiene né ad S né a T e che quindi non appartiene ad $S \cup T$.

Proposizione 6.2 *Siano S e T sottospazi di uno spazio vettoriale V. Allora $S \cup T$ è un sottospazio di V se e solo se $S \subseteq T$ oppure $T \subseteq S$.*

Dimostrazione. Certamente se $S \subseteq T$ (resp. $T \subseteq S$), $S \cup T = T$ (resp. $S \cup T = S$), pertanto in questi casi l'unione di due sottospazi è un sottospazio. Supponiamo ora che S non sia contenuto in T e che T non sia contenuto in S. Esisteranno dunque almeno un elemento $v \in S$, $v \notin T$, ed un elemento $w \in T$, $w \notin S$. Allora $v+w$ non appartiene ad S, altrimenti S, essendo chiuso rispetto alla somma ed al prodotto per scalari, conterrebbe anche il vettore $(v+w) - v = w$, contro le ipotesi. Analogamente, $v+w$ non appartiene a T, altrimenti T conterrebbe il vettore $(v+w) - w = v$, contro le ipotesi. Dunque $v+w$ non appartiene a $S \cup T$, dal momento che non appartiene né ad S né a T, pertanto $S \cup T$ non è chiuso rispetto alla somma e quindi non è un sottospazio vettoriale di V.

Dal momento che in generale l'unione di due sottospazi S, T di uno spazio vettoriale V non è un sottospazio di V, è naturale chiedersi quale sia il più piccolo sottospazio di V contenente $S \cup T$.

Definizione 6.1.3 *Dati due sottospazi S e T di uno spazio vettoriale V, si chiama somma di S e T il seguente sottoinsieme di V:*

$$S + T = \{v = s + t \mid s \in S, t \in T\}.$$

Proposizione 6.1.4 *La somma $S + T$ di due sottospazi vettoriali S e T di uno spazio vettoriale V è un sottospazio vettoriale di V.*

Dimostrazione. Osserviamo preliminarmente che il vettore nullo di V appartiene ad $S+T$, dal momento che $0_V \in S$, $0_V \in T$ e $0_V = 0_V + 0_V$.

Mostriamo ora che $S+T$ è chiuso rispetto alla somma: siano v, w elementi di $S+T$, cioè $v = s_1 + t_1$, $w = s_2 + t_2$, con $s_1, s_2 \in S$, $t_1, t_2 \in T$. Allora: $v+w = (s_1+t_1)+(s_2+t_2) = (s_1+s_2)+(t_1+t_2)$ dove s_1+s_2 appartiene ad S essendo somma di elementi di S e t_1+t_2 appartiene a T essendo somma di elementi di T. Dunque $v+w$ appartiene ad $S+T$.

Sia ora, come prima, $v = s_1 + t_1 \in S+T$ e sia $\alpha \in \mathbb{R}$. Allora $\alpha v = \alpha s_1 + \alpha t_1$, dove αs_1 appartiene ad S, essendo S chiuso rispetto al prodotto per scalari e, similmente, αt_1 appartiene a T. Si ha dunque $\alpha v \in S+T$.

Abbiamo così mostrato che $S+T$ è chiuso rispetto alla somma e al prodotto per scalari e pertanto è un sottospazio vettoriale di V.

Osservazione 6.1.5 Osserviamo che se W è un sottospazio di V contenente i sottospazi S e T, certamente W dovrà contenere anche la somma di tutti gli elementi di S e di tutti gli elementi di T, cioè tutti gli elementi di $S+T$. $S+T$ è pertanto il più piccolo sottospazio di V contenente sia S che T.

Osservazione 6.1.6 Dati due sottospazi S e T di uno spazio vettoriale V, come si determina concretamente il sottospazio $S+T$? Sia $\mathcal{B} = \{v_1, \dots, v_k\}$ una base di S e sia $\mathcal{C} = \{u_1, \dots, u_h\}$ una base di T. Allora $\{v_1, \dots, v_k, u_1, \dots, u_h\}$ è un insieme di generatori di $S+T$. Infatti ogni vettore v di $S+T$ è della forma: $v = s + t$ con $s \in S$ e $t \in T$. Allora $s = \alpha_1 v_1 + \ldots + \alpha_k v_k$, essendo $\mathcal{B}$ una base di S, e $t = \beta_1 u_1 + \ldots + \beta_h u_h$, essendo $\mathcal{C}$ una base di T. Dunque $v = \alpha_1 v_1 + \ldots + \alpha_k v_k + \beta_1 u_1 + \ldots + \beta_h u_h$.

Naturalmente NON stiamo affermando che i vettori $\{v_1, \dots, v_k, u_1, \dots, u_h\}$ sono una base di $S+T$ ma solo che essi generano $S+T$. Per individuare una base di $S+T$ bisognerà scartare gli eventuali vettori linearmente dipendenti.

Esempio 6.1.7 Consideriamo i sottospazi vettoriali di $\mathbb{R}^4$:

$$S = \langle (1,0,1,1), (2,1,3,-1) \rangle, \quad T = \langle (0,-1,-1,3), (-1,1,2,1) \rangle$$

e determiniamo una base di $S+T$. Grazie all'Osservazione 6.1.6, sappiamo che $S+T = \langle (1,0,1,1), (2,1,3,-1), (0,-1,-1,3), (-1,1,2,1) \rangle$, pertanto

$$\dim(S+T) = rg \begin{pmatrix} 1 & 0 & 1 & 1 \\ 2 & 1 & 3 & -1 \\ 0 & -1 & -1 & 3 \\ -1 & 1 & 2 & 1 \end{pmatrix} = rg \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & -3 \\ 0 & -1 & -1 & 3 \\ 0 & 1 & 3 & 2 \end{pmatrix} =$$

$$= rg \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & -3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 5 \end{pmatrix} = 3.$$

Il sottospazio $S+T$ ha pertanto dimensione 3. Dal momento che le operazioni elementari sulle righe di una matrice preservano lo spazio generato dalle righe, una base di $S+T$ è data, ad esempio dalle righe non nulle della matrice ridotta in forma a scala: $\{(1,0,1,1),(0,1,1,-3),(0,0,2,5)\}$.

Se lo spazio vettoriale W possiede due sottospazi T_1 e T_2 tali che $T_1+T_2 = W$ ma $T_1 \cap T_2 \neq \{\mathbf{0}_V\}$, allora uno stesso vettore di W si può scrivere in due modi diversi come somma di un elemento di T_1 ed uno di T_2 (si veda, a questo proposito, l'Esempio 6.1.8). Naturalmente il sottospazio W può anche coincidere con lo spazio ambiente.

Esempio 6.1.8 In $\mathbb{R}^3$ siano dati i due sottospazi $T_1 = \langle(0,1,-1),(1,1,0)\rangle$ e $T_2 = \langle(1,2,-1),(0,0,1)\rangle$. Usando la definizione di spazio vettoriale generato da un insieme di vettori possiamo descrivere esplicitamente gli elementi di ciascun sottospazio: ogni elemento di T_1 è del tipo $\lambda_1(0,1,-1) + \lambda_2(1,1,0)$ al variare di $\lambda_1, \lambda_2 \in \mathbb{R}$, mentre ogni elemento di T_2 è della forma $\eta_1(1,2,-1) + \eta_2(0,0,1)$ al variare di $\eta_1, \eta_2 \in \mathbb{R}$. Per determinare un elemento nella intersezione di T_1 e T_2 si dovranno determinare $\lambda_1, \lambda_2, \eta_1, \eta_2 \in \mathbb{R}$ tali che $\eta_1(1,2,-1) + \eta_2(0,0,1) = \lambda_1(0,1,-1) + \lambda_2(1,1,0)$, vale a dire $(\eta_1, 2\eta_1, -\eta_1 + \eta_2) = (\lambda_2, \lambda_1 + \lambda_2, -\lambda_1)$. Due vettori di $\mathbb{R}^3$ sono uguali se e solo se le loro componenti sono ordinatamente uguali, in questo caso:

$$\begin{cases} \eta_1 = \lambda_2 \\ 2\eta_1 = \lambda_1 + \lambda_2 \\ -\eta_1 + \eta_2 = -\lambda_1. \end{cases}$$

Risolvendo otteniamo: $\eta_1 = \lambda_1 = \lambda_2$ e $\eta_2 = 0$. (Abbiamo trovato tutte le possibili soluzioni del sistema? Perché?). Abbiamo trovato che $T_1 \cap T_2 = \{\eta_1(1,2,-1) \mid \eta_1 \in \mathbb{R}\} = \langle\langle(1,2,-1)\rangle$. In particolare si ha allora che $1(0,1,-1)+1(1,1,0) = (1,2,-1) \in T_1$ e $1(1,2,-1)+0(0,0,1) = (1,2,-1) \in T_2$.

Osserviamo che il vettore $(1,3,-1)$ si può scrivere in due modi diversi come somma di un elemento di T_1 e di un elemento di T_2: $(1,3,-1) = (1,3,-2) + (0,0,1) = (0,1,-1) + (1,2,0)$.

Definizione 6.1.9 *La somma di due sottospazi S e T di uno spazio vettoriale V si dice diretta se $S \cap T = \{0_V\}$. In tal caso indicheremo la somma di S e T con $S \oplus T$.*

Osservazione 6.1.10 Supponiamo che V sia uno spazio vettoriale finitamente generato e che T_1 e T_2 siano sottospazi di V tali che $V = T_1 \oplus T_2$. Allora anche T_1 e T_2 sono finitamente generati. Siano dunque $\{u_1, \dots, u_k\}$ e $\{v_1, \dots, v_p\}$ rispettivamente una base di T_1 ed una base di T_2. L'unione delle due basi $\{u_1, \dots, u_k, v_1, \dots, v_p\}$ è una base per lo spazio somma V.

Abbiamo già visto nell'Osservazione 6.1.6 che lo spazio somma V è generato dai vettori $\{u_1, \dots, u_k, v_1, \dots, v_p\}$. Si tratta ora di vedere che essi sono linearmente indipendenti. Supponiamo che esista una relazione di dipendenza tra essi: $\beta_1 u_1 + \cdots + \beta_k u_k + \gamma_1 v_1 + \cdots + \gamma_p v_p = \mathbf{0}_V$ per determinati numeri reali γ_j e β_i non tutti nulli. Se tutti i γ_j fossero eguali a 0 avremmo qualche β_i diverso da zero e quindi una relazione di dipendenza che coinvolge solo $u_1, \dots, u_k$ che sono linearmente indipendenti il che è assurdo. Analogamente otterremmo una contraddizione se tutti i β_i fossero uguali a 0. Dobbiamo quindi concludere che qualche β_i e contemporaneamente qualche γ_j sia diverso da zero, allora, posto

$$w = \beta_1 u_1 + \cdots + \beta_k u_k = -\gamma_1 v_1 - \cdots - \gamma_p v_p,$$

si avrebbe $w \in T_1 \cap T_2$, ma l'intersezione tra T_1 e T_2 contiene solo il vettore nullo 0_V, quindi $\beta_1 u_1 + \cdots + \beta_k u_k = \mathbf{0}_V = -\gamma_1 v_1 - \cdots - \gamma_p v_p$. Essendo $u_1, \dots, u_k$ linearmente indipendenti, così come $v_1, \dots, v_p$, si ottiene $\beta_1 = \ldots = \beta_k = \gamma_1 = \ldots = \gamma_p = 0$. È stato dunque assurdo pensare che qualche coefficiente potesse essere diverso da 0.

6.2 La formula di Grassmann

Presi due sottoinsiemi finiti di un insieme se vogliamo contare il numero di elementi della loro unione possiamo innanzitutto sommare il numero di elementi di uno al numero di elementi dell'altro: in questo modo però gli elementi che stanno nella intersezione vengono contati due volte. Per calcolare allora il numero esatto di elementi dell'unione dobbiamo sottrarre alla somma precedente il numero degli elementi dell'intersezione. Questo è essenzialmente ciò che succede anche per i sottospazi, sostituendo al numero di elementi del sottoinsieme la dimensione del sottospazio e alla unione la somma di sottospazi.

Teorema 6.2.1 (Formula di Grassmann) *Siano V un $\mathbb{R}$-spazio vettoriale, T_1 e T_2 due suoi sottospazi di dimensione finita, rispettivamente n_1 e n_2. Allora anche l'intersezione $T_1 \cap T_2$ ha dimensione finita e si ha:*

$$\dim(T_1 + T_2) = \dim T_1 + \dim T_2 - \dim(T_1 \cap T_2).$$

Dimostrazione. Certamente $T_1 \cap T_2$ ha dimensione finita in quanto sottospazio vettoriale di uno spazio di dimensione finita (T_1, ad esempio, o T_2). Sia ora $B = \{v_1, \dots, v_k\}$ una base di $T_1 \cap T_2$. B è un insieme di vettori linearmente indipendenti di T_1 (e di T_2) e quindi si può completare sia in una base di T_1: $\mathcal{B}_1 = \{v_1, \dots, v_k, v'_{k+1}, \dots, v'_{n_1}\}$, che in una base di T_2: $\mathcal{B}_2 = \{v_1, \dots, v_k, v''_{k+1}, \dots, v''_{n_2}\}$. Con gli indici sopra introdotti abbiamo $\dim T_1 = n_1$, $\dim T_2 = n_2$ e $\dim T_1 \cap T_2 = k$. Il teorema sarà dimostrato se mostreremo che $\mathcal{B}_1 \cup \mathcal{B}_2 = \{v_1, \dots, v_k, v'_{k+1}, \dots, v'_{n_1}, v''_{k+1}, \dots, v''_{n_2}\}$, che ha proprio n_1+n_2-k elementi, è una base di T_1+T_2. Dall'osservazione 6.1.6 sappiamo che $\mathcal{B}_1 \cup \mathcal{B}_2$ è un insieme di generatori di T_1+T_2. Vediamo ora che esso è un insieme di vettori linearmenti indipendenti. Se non lo fosse potremmo scrivere $\mathbf{0}_V = \delta_1 v_1 + \dots + \delta_k v_k + \delta_{k+1} v'_{k+1} + \dots + \delta_{n_1} v'_{n_1} + \delta_{n_1+1} v''_{k+1} + \dots + \delta_{n_1+n_2-k} v''_{n_2}$ per opportuni $\delta_j \in \mathbb{R}$ non tutti nulli. Avremmo cioè $\delta_1 v_1 + \dots + \delta_k v_k + \delta_{k+1} v'_{k+1} + \dots + \delta_{n_1} v'_{n_1} = -\delta_{n_1+1} v''_{k+1} - \dots - \delta_{n_1+n_2-k} v''_{n_2}$. Al primo membro dell'uguaglianza ottenuta abbiamo un vettore di T_1 e al secondo membro un elemento di T_2, quindi un elemento di $T_1 \cap T_2$. Ma se questo vettore appartiene all'intersezione di T_1 e T_2 esso è combinazione lineare dei soli vettori $v_1, \dots, v_k$ dal momento che $\{v_1, \dots, v_k\}$ è una base di $T_1 \cap T_2$. Quindi esistono $\alpha_1, \dots, \alpha_k \in \mathbb{R}$ tali che $\alpha_1 v_1 + \dots + \alpha_k v_k = -\delta_{n_1+1} v''_{k+1} - \dots - \delta_{n_1+n_2-k} v''_{n_2}$ da cui otteniamo $\alpha_1 v_1 + \dots + \alpha_k v_k + \delta_{n_1+1} v''_{k+1} + \dots + \delta_{n_1+n_2-k} v''_{n_2} = \mathbf{0}_V$, ma allora $\alpha_1 = \dots = \alpha_k = \delta_{n_1+1} = \delta_{n_1+n_2-k} = 0$ poiché $\mathcal{B}_2$ è base di T_2 e quindi insieme di vettori linearmente indipendenti.

Resta così : $\delta_1 v_1 + \dots + \delta_k v_k + \delta_{k+1} v'_{k+1} + \dots + \delta_{n_1} v'_{n_1} = \mathbf{0}_V$, ma adesso $v_1, \dots, v_k, v'_{k+1}, \dots v'_{n_1}$ sono linearmente indipendenti e quindi il solo modo di scrivere il vettore nullo come loro combinazione lineare è attraverso tutti i coefficienti uguali a zero. **C.V.D.**

Esempio 6.2.2 In $\mathbb{R}^3$ consideriamo i due sottospazi $Z_1 = \langle (0,1,0), (1,1,1) \rangle$ e $Z_2 = \langle (1,2,0), (1,1,-1) \rangle$. Z_1 e Z_2 hanno entrambi dimensione 2 e $\dim(Z_1 + Z_2) = \dim Z_1 + \dim Z_2 - \dim(Z_1 \cap Z_2)$. Ora, dal momento che $Z_1 + Z_2$ è un sottospazio di $\mathbb{R}^3$, avrà al più dimensione 3, quindi da $\dim(Z_1 + Z_2) = 4 - \dim(Z_1 \cap Z_2)$ si deduce che $Z_1 \cap Z_2$ non può essere lo spazio vettoriale banale. D'altra parte $Z_1 \cap Z_2$ è un sottospazio di Z_1 (e di Z_2) e dunque potrà avere dimensione 1 o 2. Se avesse dimensione 2 allora sarebbe un sottospazio di dimensione due in uno spazio di dimensione 2 cioè coinciderebbe con Z_1 (e con Z_2), si avrebbe quindi: $Z_1 = Z_2 = Z_1 \cap Z_2$, ma questo non è vero, perché $(0,1,0) \notin Z_2$, quindi Z_1 ha dimensione 1 e $\dim(Z_1 + Z_2) = 3$, i.e., $Z_1 + Z_2 = \mathbb{R}^3$. Chi è $Z_1 \cap Z_2$? Sappiamo che è uno spazio vettoriale di dimensione 1, quindi per trovarne una base basterà esibire un suo elemento diverso dal vettore nullo. Si tratta quindi di determinare $\eta_1, \eta_2, \epsilon_1, \epsilon_2 \in \mathbb{R}$, per cui $\eta_1(0,1,0) + \eta_2(1,1,1) = \epsilon_1(1,2,0) + \epsilon_2(1,1,-1)$, cioè: $(\eta_2, \eta_2 + \eta_1, \eta_2) = (\epsilon_2 + \epsilon_1, \epsilon_2 + 2\epsilon_1, -\epsilon_2)$, i.e. $\eta_2 = -\epsilon_2, \epsilon_1 = -2\epsilon_2, \eta_1 = -2\epsilon_2$, ad esempio $\epsilon_2 = 1, \eta_2 = -1, \epsilon_1 = -2, \eta_1 = -2 : (-1,-3,-1) \in Z_1 \cap Z_2$.

Osservazione 6.2.3 Per definizione di somma diretta, dal Teorema 6.2.1 si ha subito:

$$\dim(S \oplus T) = \dim(S) + \dim(T).$$

6.3 Esercizi svolti

Esercizio 6.3.1 In $\mathcal{M}_2(\mathbb{R})$ si consideri l'insieme $U = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathbb{R}) \mid a^2 = d^2 \right\}$. L'insieme U è un sottospazio vettoriale di $\mathcal{M}_2(\mathbb{R})$? In caso di risposta negativa, si determini il sottospazio vettoriale generato da U.

Svolgimento. La prima cosa da verificare è che U contenga il vettore nullo. Il vettore nullo in $\mathcal{M}_2(\mathbb{R})$ è la matrice nulla che appartiene a U. Osserviamo che le matrici $A = \begin{pmatrix} -1 & 3 \\ 4 & 1 \end{pmatrix}$ e $B = \begin{pmatrix} 2 & -3 \\ 0 & 2 \end{pmatrix}$ appartengono ad U. Se U fosse un sottospazio vettoriale di $\mathcal{M}_2(\mathbb{R})$ dovrebbe essere chiuso rispetto alla somma. Osserviamo però che se sommiamo le matrici A e B:

$$A + B = \begin{pmatrix} -1 & 3 \\ 4 & 1 \end{pmatrix} +_{\mathcal{M}_2(\mathbb{R})} \begin{pmatrix} 2 & -3 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 4 & 3 \end{pmatrix}$$

la matrice $A + B$ non appartiene ad U dal momento che $1^2 \neq 3^2$! Quindi U non è un sottospazio vettoriale di $\mathcal{M}_2(\mathbb{R})$. Si noti, però, che se $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in U$ (i.e. $a^2 = d^2$), preso $\alpha \in \mathbb{R}$ allora $\alpha \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \alpha a & \alpha b \\ \alpha c & \alpha d \end{pmatrix} \in U$: infatti se $a^2 = d^2$ allora $(\alpha a)^2 = (\alpha d)^2$.

Il sottospazio generato da U è, per definizione, il più piccolo sottospazio vettoriale che contenga U. Tale sottospazio deve pertanto contenere tutti gli elementi di U ed ogni loro combinazione lineare. Osserviamo che U contiene i vettori $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ e i vettori $C = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ e $D = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Lo spazio che stiamo cercando contiene allora le matrici A, B, C, D e tutte le loro combinazioni lineari. Ad esempio conterrà la somma di C e D: $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$ e la loro differenza $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + (-1)\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$. Lo spazio generato da U contiene allora le matrici $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ e $\begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$, $\begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$: ma questi sono quattro vettori linearmente indipendenti di $\mathcal{M}_2(\mathbb{R})$ e quindi una base di $\mathcal{M}_2(\mathbb{R})$ (dal momento che $\mathcal{M}_2(\mathbb{R})$ ha dimensione 4). Pertanto lo spazio generato da U è tutto $\mathcal{M}_2(\mathbb{R})$.

Esercizio 6.3.2 Siano $U = \{(x, y, z) \in \mathbb{R}^3 \mid x+y+z = 0\}$ e $V = \{(x, y, z) \in \mathbb{R}^3 \mid x = y\}$.

(i) Verificare che U e V sono sottospazi vettoriali di $\mathbb{R}^3$.

(ii) Determinare una base di U ed una base di V.

(iii) Determinare $U \cap V$ e $U + V$.

(iv) Completare la base di $U \cap V$ ad una base di $\mathbb{R}^3$.

Svolgimento. A questo punto dello studio, lo studente dovrebbe essere in grado di risolvere la parte (i) senza colpo ferire, se non è questo il caso ....si riparta dalla Lezione 3.

(ii) Gli elementi di U sono i vettori di $\mathbb{R}^3$ della forma $(x, y, -x - y) = x(1, 0, -1)+y(0, 1, -1)$ con $x, y \in \mathbb{R}$. Pertanto l'insieme $\{(1, 0, -1), (0, 1, -1)\}$ genera U. D'altra parte i vettori $(1, 0, -1)$, $(0, 1, -1)$ sono linearmente indipendenti e quindi individuano una base di U.

Analogamente gli elementi di V sono tutti e soli della forma (a, a, z) con a e z in $\mathbb{R}$, cioè $a(1, 1, 0) + z(0, 0, 1)$. Così $\{(1, 1, 0), (0, 0, 1)\}$ genera V ed è una sua base dal momento che i vettori $(1, 1, 0)$ e $(0, 0, 1)$ sono linearmente indipendenti.

(iii) Il sottospazio $U+V$ contiene tutti gli elementi di U e tutti gli elementi di V. In particolare contiene i vettori $(1,0,-1),(0,1,-1)$, $(0,0,1)$. Questi tre vettori sono linearmente indipendenti e generano quindi $\mathbb{R}^3$. Di conseguenza $U+V$ coincide con $\mathbb{R}^3$ e l'insieme $\{(1,0,-1),(0,1,-1),(0,0,1)\}$ è una sua base (come pure $\{(1,0,0),(0,1,0),(0,0,1)\}$ o qualsiasi altra base di $\mathbb{R}^3$). Il Teorema 6.2.1 ci consente ora di calcolare la dimensione di $U \cap V$: $\dim(U \cap V) = \dim U + \dim V - \dim(U+V) = 2+2-3 = 1$. Pertanto per determinare una base di $U \cap V$ sarà sufficiente determinare un vettore non nullo appartenente sia ad U che a V. Ad esempio, il vettore $(1,1,-2)$ soddisfa sia l'equazione di U che quella di V e appartiene pertanto alla loro intersezione. Dunque $U \cap V = \langle (1,1,-2) \rangle$.

(iv) Per completare $\{(1,1,-2)\}$ in una base di $\mathbb{R}^3$ possiamo prendere, ad esempio, i vettori $(0,1,0)$ e $(0,0,1)$. Perché? Basta verificare che l'insieme $\{(1,1,-2),(0,1,0),(0,0,1)\}$ è un insieme di vettori linearmente indipendenti.

Esercizio 6.3.3 In $\mathbb{R}^{\leq 3}[x]$ si considerino i polinomi $p_1 = x^2-1$, $p_2 = 3$, $p_3 = 2x^2+1$, $q_1 = x^3$, $q_2 = 5$, $q_3 = x+1$. Siano S e T i sottospazi vettoriali di $\mathbb{R}^{\leq 3}[x]$ generati rispettivamente da p_1, p_2, p_3 e q_1, q_2, q_3. Determinare:

(i) la dimensione di S e la dimensione di T;

(ii) la dimensione di $S \cap T$ e $S+T$ ed una base di ciascuno di essi.

Svolgimento.

(i) S è il sottospazio vettoriale di $\mathbb{R}^{\leq 3}[x]$ generato dai polinomi $p_1 = x^2 - 1$, $p_2 = 3$, $p_3 = 2x^2+1$. Cerchiamo di stabilire se questi sono o meno linearmente indipendenti. I polinomi p_1 e p_2 sono certamente linearmente indipendenti dal momento che p_1 ha grado 2 e p_2 ha grado 0. Il polinomio p_3 è combinazione lineare di p_1 e p_2? Esistono, in altre parole, α e β in $\mathbb{R}$ tali che sia $p_3 = 2x^2+1 = \alpha(x^2-1)+\beta(3) = 3\beta - \alpha + \alpha x^2$? In effetti basta prendere $\alpha = 2$ e $\beta = 1$. Dunque p_3 è combinazione lineare di p_1 e p_2 e $S = \langle p_1, p_2 \rangle$ ha dimensione 2.

Procedendo in modo analogo si verifica che i polinomi q_1, q_2 e q_3 sono linearmente indipendenti e che, quindi, T ha dimensione 3.

(ii) A questo punto sappiamo già che la somma di S e T non può essere diretta, cioè che la loro intersezione non può essere banale. Infatti

$$\dim S \cap T = \dim S + \dim T - \dim(S+T) = 2+3-\dim(S+T) \geq$$

$\geq 2+3-4=1$

dal momento che $S+T$ è un sottospazio di $\mathbb{R}^{\leq 3}[x]$ e ha pertanto dimensione minore o uguale a quella di $\mathbb{R}^{\leq 3}[x]$ cioè 4. Determiniamo $S\cap T$: un elemento di $S\cap T$ è un polinomio che si scrive contemporaneamente come combinazione lineare di p_1 e p_2 e di q_1, q_2, q_3:

$$a(x^2-1)+b(3)=\alpha x^3+\beta(5)+\gamma(x+1).$$

Otteniamo dunque: $\alpha=a=\gamma=0$ e $5\beta=3b$. Di conseguenza, gli unici polinomi che appartengono sia a T che a S sono i polinomi di grado 0: $S\cap T=\langle 1\rangle$. Otteniamo allora immediatamente che $\dim(S+T)=4$ cioè $S+T=\mathbb{R}^{\leq 3}[x]$.

Esercizio 6.3.4 In $V=\mathbb{R}^2$ consideriamo i sottospazi vettoriali $T_1=\langle(1,1)\rangle$ e $T_2=\langle(1,2)\rangle$. Si vede immediatamente che T_1 e T_2 non hanno vettori in comune diversi dal vettore nullo. Inoltre l'insieme $\{(1,1),(1,2)\}$ genera $\mathbb{R}^2$. Dunque: $\mathbb{R}^2=T_1\oplus T_2$. Sia ora $S=\langle(1,3)\rangle$. Allora, come prima, $\mathbb{R}^2=T_1\oplus S$, ma $S\neq T_2$. Inoltre vale anche: $\mathbb{R}^2=S\oplus T_2$ e questo mostra che non solo S e T_2 non hanno vettori in comune, ma addirittura sono talmente diversi da generare tutto lo spazio in somma diretta. D'altro canto non avendo S e T_2 vettori in comune ed essendo lo spazio ambiente $\mathbb{R}^2$, così deve essere. Da questo esercizio risulta chiaro che dato $T\leq V$ (con V di dimensione finita) esiste, ma non è unico, un sottospazio $S\leq V$ tale che $T\oplus S=V$.

Esercizio 6.3.5 Consideriamo i seguenti sottospazi vettoriali di $\mathbb{R}^4$: $S=\{(x,y,z,t)\in\mathbb{R}^4 \mid 3x+y+2z=0\}$ e $T=\langle(2,1,0,-5),(3,4,0,0)\rangle$. Stabilire se la somma di S e T è diretta e calcolare la dimensione di $S+T$.

Svolgimento. La somma di due sottospazi S e T di uno spazio vettoriale V è diretta se $S\cap T=\{\mathbf{0}_V\}$. Determiniamo dunque l'intersezione di S e T. Un elemento generico di T è un elemento della forma $v=(2\alpha+3\beta,\alpha+4\beta,0,-5\alpha)$ con $\alpha,\beta\in\mathbb{R}$. Il vettore v appartiene a S se e solo se le sue coordinate soddisfano l'equazione di S: $6\alpha+9\beta+\alpha+4\beta=0$, cioè $7\alpha+13\beta=0$. Dunque il vettore $v=13(2,1,0,-5)-7(3,4,0,0)=(5,-15,0,-65)$ appartiene a $S\cap T$, così come ogni suo multiplo. Pertanto la somma di S e T non è diretta e $S\cap T=\langle v\rangle$. Si ha dunque $\dim(S+T)=\dim S+\dim T-\dim S\cap T=3+2-1=4$, cioè $S+T=\mathbb{R}^4$.

6.4 Esercizi proposti

Esercizio 6.4.1 Costruire due sottospazi S e T di $\mathbb{R}^3$ la cui somma non sia diretta e sia uguale ad $\mathbb{R}^3$.

Esercizio 6.4.2 Sia $S = \langle (1,1,1,1), (1,0,-1,2), (-1,0,-1,1), (1,2,1,3) \rangle$.

1. Determinare la dimensione di S.
2. Determinare, se possibile, un sottospazio T di $\mathbb{R}^4$ tale che $S \oplus T = \mathbb{R}^4$.
3. Determinare, se possibile, un sottospazio non banale W di $\mathbb{R}^4$ tale che la somma di S e W sia diretta e propriamente contenuta in $\mathbb{R}^4$.
4. Determinare, se possibile, un sottospazio V di $\mathbb{R}^4$ tale che la somma di S e V non sia diretta e sia uguale a $\mathbb{R}^4$.

Esercizio 6.4.3 Si considerino i seguenti sottospazi vettoriali di $\mathcal{M}_2(\mathbb{R})$:

$$U = \left\{ \begin{pmatrix} a & b \\ 0 & -a \end{pmatrix} \mid a, b \in \mathbb{R} \right\}, \quad V = \left\{ \begin{pmatrix} d & 0 \\ d & d \end{pmatrix} \mid d \in \mathbb{R} \right\}.$$

1. Determinare una base di U ed una base di V.
2. Determinare una base di $U + V$ ed una base di $U \cap V$.
3. Completare la base di $U + V$ trovata in 2. in una base $\mathcal{B}$ di $\mathcal{M}_2(\mathbb{R})$.
4. Determinare le coordinate del vettore $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ nella base $\mathcal{B}$.

Esercizio 6.4.4 Sia V un $\mathbb{R}$ spazio vettoriale di dimensione 3 e sia $\{u, v, w\}$ una base di V. Siano $S = \langle u - v, v - w \rangle$ e $T = \langle u + w, 2u - w \rangle$.

1. Calcolare la dimensione di S e la dimensione di T.
2. Determinare $S + T$ e $S \cap T$ ed esibire una base per ciascuno di essi.
3. Completare la base trovata di $S \cap T$ in una base di V.

Esercizio 6.4.5 In $\mathbb{R}^{\leq 2}[x]$ si considerino i sottospazi $S = \langle 1 + x, 1 - x^2 \rangle$ e $T = \langle 2 + x + x^2, x + x^2 \rangle$.

1. Determinare $S + T$ e $S \cap T$.
2. Stabilire se ogni vettore di $\mathbb{R}^{\leq 2}[x]$ si scrive come somma di un vettore di S e di un vettore di T e, in caso affermativo, se tale scrittura è unica.

Lezione 7

Applicazioni lineari e matrici

7.1 Applicazioni lineari

Un modo naturale di confrontare insiemi diversi è attraverso le funzioni da un insieme all'altro. Se si vogliono confrontare spazi vettoriali diversi occorre studiare quelle funzioni che in qualche modo preservino le operazioni definite sul dominio.

Ricordiamo che, dati due insiemi X e Y, una funzione $f : X \to Y$ è una corrispondenza che ad ogni elemento x di X associa un elemento $f(x)$ di Y. L'elemento $f(x)$ si dice *immagine* di x mediante f. Gli insiemi X e Y si chiamano, rispettivamente, dominio e codominio della funzione f.

Definizione 7.1.1 *Siano V e W due $\mathbb{R}$-spazi vettoriali. Una funzione $L : V \to W$ si dice lineare se*

- $\forall\ v_1, v_2 \in V$, $L(v_1 +_V v_2) = L(v_1) +_W L(v_2)$.
- $\forall\ v \in V$, $\forall \alpha \in \mathbb{R}$, $L(\alpha v) = \alpha L(v)$.

Osservazione 7.1.2 Dalla definizione segue subito che, per qualunque vettore v di V, $L(\mathbf{0}_V) = L(0v) = 0L(v) = \mathbf{0}_W$.

Notiamo anche che l'opposto di un vettore v ha come immagine l'opposto del vettore $L(v)$: $L((-1)v) = (-1)L(v)$ che è l'opposto in W del vettore $L(v)$.

Definizione 7.1.3 *Sia $f : V \to W$ una funzione lineare. Si chiama nucleo di f e si indica con* $\ker(f)$ *il sottoinsieme di V così definito:*

$$\ker(f) = \{v \in V \mid f(v) = 0_W\}.$$

Proposizione 7.1.4 *Il nucleo di una funzione lineare $f : V \to W$ è un sottospazio vettoriale di V.*

Dimostrazione. Nell'Osservazione 7.1.2 abbiamo visto che $f(0_V) = 0_W$, perciò $0_V \in \ker(f)$.

Ora dobbiamo mostrare che $\ker f$ è chiuso rispetto alla somma e al prodotto per scalari. Siano $v_1, v_2 \in \ker(f)$, cioè $f(v_1) = 0_W = f(v_2)$. Allora $f(v_1+v_2) = f(v_1)+f(v_2) = 0_W+0_W = 0_W$, quindi $v_1+v_2 \in \ker(f)$, pertanto $\ker(f)$ è chiuso rispetto alla somma.

Siano ora $\alpha \in \mathbb{R}$ e $v_1 \in \ker(f)$. Allora $f(\alpha v_1) = \alpha f(v_1) = \alpha 0_W = 0_W$, pertanto $\alpha v_1 \in \ker(f)$, cioè $\ker(f)$ è chiuso rispetto al prodotto per scalari.

Il nucleo di una funzione lineare è indicativo della iniettività della funzione. Ricordiamo che una funzione $f : X \to Y$ è iniettiva se per ogni coppia di elementi distinti di X $x_1 \neq x_2$ si ha $f(x_1) \neq f(x_2)$. La medesima definizione vale ovviamente per applicazioni lineari iniettive. Come possiamo caratterizzare una applicazione lineare iniettiva? Vale il seguente risultato:

Proposizione 7.1.5 *Data una applicazione lineare $L : V \to W$ tra due $\mathbb{R}$-spazi vettoriali V, W, L è iniettiva se e solo se* Ker$L$ *è il sottospazio banale di V.*

Dimostrazione. "$\Rightarrow$" Sia L iniettiva. Sappiamo che il sottospazio Ker$L$ è costituito dai vettori di V la cui immagine mediante L è il vettore nullo di W. Per definizione di applicazione lineare tale spazio contiene sempre il vettore nullo di V. Se contenesse un altro vettore $v \in V$, $v \neq \mathbf{0}_V$, vorrebbe dire $L(v) = L(\mathbf{0}_V) = \mathbf{0}_W$ e questo non è possibile perché la funzione è iniettiva.

"$\Leftarrow$" Sia Ker$L = \{\mathbf{0}_V\}$ e supponiamo che la funzione non sia iniettiva cioè che esistano due vettori $v_1, v_2 \in V$, $v_1 \neq v_2$, tali che $L(v_1) = L(v_2)$ cioè $L(v_1)-L(v_2) = \mathbf{0}_W$, ma L è lineare dunque $L(v_1-v_2) = L(v_1)-L(v_2) = \mathbf{0}_W$, pertanto $v_1 - v_2 \in$ Ker$L$. Essendo $v_1 \neq v_2$, $v_1 - v_2 \neq \mathbf{0}_V$ e questo non è possibile perché Ker$L = \{\mathbf{0}_V\}$. La funzione L è quindi iniettiva. **C.V.D.**

Definizione 7.1.6 *Sia $f : V \to W$ una funzione lineare tra due $\mathbb{R}$-spazi vettoriali. Si chiama immagine di f e si indica con $Im(f)$ il seguente sottoinsieme di W:*

$$Im(f) = \{w \in W \mid w = f(v) \;\; \textit{per qualche } v \in V\}.$$

Proposizione 7.1.7 *L'immagine di una funzione lineare $f : v \to W$ è un sottospazio vettoriale di W.*

Dimostrazione. Abbiamo più volte osservato che $0_W = f(0_V)$, pertanto $0_W \in Im(f)$.

Ora mostriamo che $Im(f)$ è chiusa rispetto alla somma e al prodotto per scalari. Siano w_1, w_2 in $Im(f)$, cioè, $w_1 = f(v_1)$, $w_2 = f(v_2)$ con $v_1, v_2 \in V$. Allora $w_1 + w_2 = f(v_1) + f(v_2) = f(v_1 + v_2)$, perciò $w_1 + w_2$ appartiene a $Im(f)$ e $Im(f)$ è chiuso rispetto alla somma.

Siano poi $\alpha \in \mathbb{R}$ e $w \in Im(f)$, cioè $w = f(v)$ con $v \in V$. Allora $\alpha w = \alpha f(v) = f(\alpha v)$, quindi $alphaw \in Im(f)$ e $Im(f)$ è chiuso anche rispetto al prodotto per scalari.

Osservazione 7.1.8 Ricordiamo che una funzione si dice suriettiva se ogni elemento del codominio è immagine di almeno un elemento del dominio. Segue subito dalla definizione, dunque, che una funzione lineare $f : V \to W$ è suriettiva se e solo se $Im(f) = W$, cioè, essendo $Im(f)$ un sottospazio vettoriale di W, se e solo se $\dim(Im(f)) = \dim(W)$.

Definizione 7.1.9 *Sia $f : V \to W$ una funzione lineare e sia w un elemento di W. Si chiama antiimmagine (o preimmagine o controimmagine) di w mediante f, e si indica con $f^{-1}(w)$, il seguente sottoinsieme di V:*

$$f^{-1}(w) = \{v \in V \mid f(v) = w\}.$$

Esempio 7.1.10 Il nucleo di una funzione lineare $f : V \to W$ è la controimmagine del vettore nullo del codominio: $f^{-1}(0_W) = \ker(f)$.

Esercizio 7.1.11 *Data un'applicazione lineare $L : V \to W$ e dato un vettore $w \in W$, in che modo possiamo descrivere $L^{-1}(\{w\})$, cioè l'insieme di tutti i vettori di V la cui immagine mediante L è w?*

Soluzione Se $w = \mathbf{0}_W$ allora $L^{-1}(\{w\}) = \mathrm{Ker}L$.

Se $w \neq \mathbf{0}_W$? Se $w \notin \mathrm{Im}L$ la risposta è facile: $L^{-1}(\{w\}) = \emptyset$.

Se, invece, $w \in \mathrm{Im}L$ esiste sicuramente un elemento $v \in V$ tale che $L(v) = w$. Si ha allora:

$$L^{-1}(\{w\}) = v + \mathrm{Ker}L. \tag{7.1}$$

La scrittura $v + \mathrm{Ker}L$ indica l'insieme dei vettori della forma $v + k$ con $k \in \mathrm{Ker}L$. Per dimostrare l'uguaglianza tra i due insiemi in (7.1) dobbiamo mostrare che vale la doppia inclusione $L^{-1}(\{w\}) \subseteq v + \mathrm{Ker}L$ e $v + \mathrm{Ker}L \subseteq L^{-1}(\{w\})$. Dato $v+k \in v+\mathrm{Ker}L$ la sua immagine è $L(v+k)$ che, per linearità, coincide con $L(v) + L(k) = w + \mathbf{0}_W = w$, dunque $v + k \in L^{-1}(\{w\})$; così abbiamo visto che $v+\mathrm{Ker}L \subseteq L^{-1}(\{w\})$. Per l'altra inclusione prendiamo un vettore $s \in L^{-1}(\{w\})$ e consideriamo la somma di s con l'opposto di v: $s-v$. Applichiamo L: $L(s-v) = L(s) - L(v) = w - w = \mathbf{0}_W$, cioè $s - v \in \mathrm{Ker}L$. Così, posto $k = s - v$, si ha $s = v + k \in v + \mathrm{Ker}L$; vale dunque l'inclusione $L^{-1}(\{w\}) \subseteq v + \mathrm{Ker}L$.

Osservazione 7.1.12 i) Nella costruzione di $L^{-1}(\{w\}) = v + \mathrm{Ker}L$ abbiamo scelto un vettore v tale che $L(v) = w$. Se ne avessimo scelto un altro? Il risultato sarebbe stato lo stesso. Infatti scelto $\overline{v} \neq v$ tale che $L(\overline{v}) = w$, avremmo trovato $L^{-1}(\{w\}) = \overline{v} + \mathrm{Ker}L$, quindi $\overline{v} + \mathrm{Ker}L = v + \mathrm{Ker}L$.

ii) Che struttura ha $v + \mathrm{Ker}L$? Non possiamo aspettarci molto. Se $v \notin \mathrm{Ker}L$ allora nessuna delle somme $v + k$, $k \in \mathrm{Ker}L$, può essere il vettore nullo, quindi $v + \mathrm{Ker}L$ non è uno spazio vettoriale (se fosse $v + k = \mathbf{0}_V$ si avrebbe $v = -k \in \mathrm{Ker}L$, assurdo). L'altra possibilità è che $v \in \mathrm{Ker}L$, ma allora $v + \mathrm{Ker}L = \mathrm{Ker}L$ è un sottospazio. (Per mostrare l'uguaglianza basta osservare che ogni elemento z di $\mathrm{Ker}L$ si può scrivere come $z = v + (z - v)$ ove $-v \in \mathrm{Ker}L$ per ipotesi e dunque $z - v \in \mathrm{Ker}L$).

iii) Data una applicazione lineare $L : V \to W$ l'antiimmagine tramite L di un vettore $w \in \mathrm{Im}L$ è data da un vettore particolare $v \in V$ che soddisfi la condizione $L(v) = w$ e da tutti i vettori che si ottengono sommando v agli elementi di $\mathrm{Ker}L$. Cosicché se L è iniettiva allora $L^{-1}(\{w\})$ è costituito da un solo vettore.

iv) Sia V uno spazio vettoriale di dimensione finita n e sia $\{v_1, v_2, \dots, v_n\}$ una sua base. Consideriamo un altro spazio vettoriale W e n vettori qualsiasi $w_1, w_2, \dots, w_n$ di W (i vettori $w_1, w_2, \dots, w_n$ possono essere scelti del tutto arbitrariamente: possono essere tutti uguali, tutti uguali al vettore nullo, tutti diversi, etc.). Possiamo allora costruire una applicazione lineare $f : V \longrightarrow W$ semplicemente ponendo

$f(v_1) = w_1, f(v_2) = w_2, \ldots, f(v_n) = w_n$. Non solo: esiste un'unica applicazione lineare $f : V \longrightarrow W$ tale che $f(v_i) = w_i$ con $i = 1, \ldots, n$.

Come si costruisce f? Ogni vettore v di V si scrive in modo unico come combinazione lineare dei vettori $v_1, v_2, \ldots, v_n$: $v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n$. Se vogliamo che f sia lineare dovrà dunque essere

$$f(v) = \lambda_1 w_1 + \lambda_2 w_2 + \cdots + \lambda_n w_n. \qquad (7.2)$$

La condizione (7.2) definisce una funzione lineare. Siano infatti $v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n$ e $t = \gamma_1 v_1 + \gamma_2 v_2 + \cdots + \gamma_n v_n$, con $\lambda_i, \gamma_i \in \mathbb{R}$, due elementi di V. Allora la loro somma si scrive in modo unico come $t + v = (\gamma_1 + \lambda_1)v_1 + (\gamma_2 + \lambda_2)v_2 + \cdots + (\gamma_n + \lambda_n)v_n$. E dunque $f(t + v) = (\gamma_1 + \lambda_1)w_1 + (\gamma_2 + \lambda_2)w_2 + \cdots + (\gamma_n + \lambda_n)w_n = (\gamma_1 w_1 + \gamma_2 w_2 + \cdots + \gamma_n w_n) + (\lambda_1 w_1 + \lambda_2 w_2 + \cdots + \lambda_n w_n) = f(t) + f(v)$. Inoltre, per ogni $\alpha \in \mathbb{R}$, $\alpha v = \alpha\lambda_1 v_1 + \alpha\lambda_2 v_2 + \cdots + \alpha\lambda_n v_n$ e $f(\alpha v) = \alpha\lambda_1 w_1 + \alpha\lambda_2 w_2 + \cdots + \alpha\lambda_n w_n = \alpha(\lambda_1 w_1 + \lambda_2 w_2 + \cdots + \lambda_n w_n) = \alpha f(v)$. Quindi la applicazione f è lineare. L'unicità di f dipende dal fatto che la condizione $f(v_i) = w_i$ fissa le immagini degli elementi di una base e, per linearità, determina univocamente l'immagine di qualsiasi elemento di V. Osserviamo che il dato delle immagini di un insieme di vettori linearmente indipendenti di V che non sia una base non caratterizza univocamente un'applicazione lineare.

7.2 Struttura dimensionale

Per il momento abbiamo solo informazioni qualitative riguardo ad una applicazione lineare. Adesso vogliamo dare qualche informazione 'quantitativa'.

Osservazione 7.2.1 Sia $L : V \to W$ una funzione lineare e supponiamo che V abbia dimensione finita e che $\{v_1, \ldots, v_n\}$ sia una sua base. Allora le immagini dei vettori della base, $L(v_1), \ldots, L(v_n)$, sono un insieme di generatori di $\mathrm{Im}L$. Infatti, sia $w \in \mathrm{Im}L$; questo vuol dire che esiste un vettore $v \in V$ tale che $L(v) = w$, allora $v = \lambda_1 v_1 + \cdots + \lambda_n v_n$, essendo $v_1, \ldots, v_n$ una base di V. Applicando L, per linearità, si ottiene: $w = L(v) = \lambda_1 L(v_1) + \cdots + \lambda_n L(v_n)$ e quindi, $L(v_1), \ldots, L(v_n)$ sono un insieme di generatori di $\mathrm{Im}L$. Attenzione: NON stiamo dicendo che $L(v_1), \ldots, L(v_n)$ sono una base dell'immagine cioè che sono linearmente indipendenti! Ma dal loro insieme (come da ogni insieme di generatori) si potrà estrarre una base dell'immagine, il che consentirà di calcolare la dimensione di ImL.

Il risultato seguente risponderà a tutti i nostri quesiti in merito.

Teorema 7.2.2 (Teorema delle dimensioni) *Data una applicazione lineare $L : V \to W$ tra due $\mathbb{R}$-spazi vettoriali V, W, con $\dim V = n$, allora*

$$\dim V = \dim(\mathrm{Im}L) + \dim(\mathrm{Ker}L).$$

Dimostrazione. (N.B. Non stiamo facendo alcuna ipotesi su W, in ogni caso tutto dipende dal dominio!). Essendo Ker$L \leq V$ e $\dim V = n$, anche Ker$L$ ha dimensione finita, sia essa $k \leq n$ e sia $\{t_1, \ldots, t_k\}$ una base di Ker$L$. I vettori $t_1, \ldots, t_k$ sono linearmente indipendenti anche in V e si possono allora completare in una base di V (Teorema 5.2.7): $\{t_1, \ldots, t_k, t_{k+1}, \ldots, t_n\}$. Per l'Osservazione 7.2.1, i vettori: $L(t_1), \ldots, L(t_k), L(t_{k+1}), \ldots, L(t_n)$ sono un insieme di generatori di ImL; per costruzione $L(t_1) = \ldots = L(t_k) = \mathbf{0}_W$ e, quindi, non partecipano alla "generazione". Allora i vettori $L(t_{k+1}), \ldots, L(t_n)$ sono generatori di ImL. La proposizione sarà dimostrata se dimostreremo che $L(t_{k+1}), \ldots, L(t_n)$ sono vettori linearmente indipendenti nel qual caso $\dim(\mathrm{Im}L) = n-k$, quindi $n = \dim V = k+(n-k) = \dim(\mathrm{Ker}L)+\dim(\mathrm{Im}L)$. Sia dunque $\beta_1 L(t_{k+1}) + \cdots + \beta_{n-k} L(t_n) = \mathbf{0}_W$ una relazione di dipendenza tra i vettori $L(t_{k+1}), \ldots, L(t_n)$, con $\beta_i \in \mathbb{R}$, cioè, per la linearità di L, $L(\beta_1 t_{k+1} + \cdots + \beta_{n-k} t_n) = \mathbf{0}_W$. Questo significa che $\beta_1 t_{k+1} + \cdots + \beta_{n-k} t_n \in \mathrm{Ker}L$. Essendo $\{t_1, \ldots, t_k\}$ una base di Ker$L$, si ha: $\beta_1 t_{k+1} + \cdots + \beta_{n-k} t_n = \alpha_1 t_1 + \ldots + \alpha_k t_k$, cioè: $\alpha_1 t_1 + \ldots + \alpha_k t_k - \beta_1 t_{k+1} - \cdots - \beta_{n-k} t_n = \mathbf{0}_V$. Dal momento che i vettori $t_1, \ldots, t_n$ sono linearmente indipendenti, tutti i coefficienti della combinazione lineare trovata sono nulli, in particolare i β_i sono nulli. Quindi i vettori $L(t_{k+1}), \ldots, L(t_n)$ sono linearmente indipendenti. **C.V.D.**

Osservazione 7.2.3

i) La dimensione dell'immagine di una applicazione lineare è sempre più piccola della dimensione del dominio o uguale ad essa.

ii) Si dice che una applicazione lineare è un *endomorfismo di* V se è un'applicazione lineare in cui dominio e codominio coincidono con V cioè: $L : V \to V$. Se un endomorfismo di V è suriettivo allora è pure iniettivo: infatti se la funzione è suriettiva (Im$L = V$) allora $\dim(\mathrm{Im}L) =$

dim(V) il che implica, per il Teorema 7.2.2, dim(Ker$L$) $= 0$ cioè il nucleo è il sottospazio banale, quindi la funzione è iniettiva. Viceversa, se l'endomorfismo è iniettivo, Ker$L = \{\mathbf{0}_V\}$ e dunque la sua dimensione è zero. Allora $\dim V = \dim \mathrm{Im} L$ e quindi ImL ha dimensione uguale alla dimensione dello spazio di cui è sottospazio, dunque coincide con esso e la funzione è suriettiva. In conclusione, un endomorfismo di uno spazio vettoriale V è iniettivo se e solo se è suriettivo e quindi biiettivo. Si osservi che questa proprietà differenzia in modo sostanziale il comportamento delle funzioni lineari da quello delle funzioni non lineari. Si faccia, per esercizio, un esempio di una funzione (non lineare) $g : \mathbb{R} \to \mathbb{R}$ iniettiva ma non suriettiva o suriettiva ma non iniettiva.

iii) Sia $L : V \to W$ un'applicazione lineare tra spazi vettoriali di dimensione finita.

Se $\dim V = n > \dim W = m$, per il Teorema 7.2.2, $n = \dim \mathrm{Ker}(L) + \dim \mathrm{Im}(L)$, ma, essendo $\mathrm{Im}(L) \leq W$, si ha $\dim \mathrm{Im}(L) \leq m$ quindi $\dim \mathrm{Ker}(L) = n - \dim \mathrm{Im}(L) \geq n - m > 0$. Essendo $\dim \mathrm{Ker}(L) > 0$ l'applicazione lineare L non può mai essere iniettiva.

Se $\dim \mathrm{Ker}(L) + \dim \mathrm{Im}(L) = n < m$ si ha $\dim \mathrm{Im}(L) = n - \dim \mathrm{Ker}(L) \leq n < m$, quindi l'applicazione lineare L non può mai essere suriettiva.

iv) Una applicazione lineare $L : V \to W$ tra due $\mathbb{R}$-spazi vettoriali V, W di dimensione finita, si dice un **isomorfismo** se è biiettiva. Poiché la funzione è suriettiva Im$L = W$, poiché la funzione è iniettiva dim(Ker$L$) $=$ 0 e dal Teorema delle dimensioni deduciamo che $\dim V = \dim(\mathrm{Im} L) = \dim W$. Si noti che non è detto a priori che la funzione inversa di un'applicazione lineare sia lineare, ma dimostreremo più avanti che, di fatto, lo è.

7.3 Applicazioni lineari, basi e matrici

Abbiamo visto che, dato uno spazio vettoriale V e fissata una sua base $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$, ogni vettore $v \in V$ è univocamente individuato dalla n-upla delle sue coordinate rispetto alla base scelta: se $v = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n$, allora v può essere indicato con $(\alpha_1, \alpha_2, \ldots, \alpha_n)_{\mathcal{B}}$. Questa n-upla è una specie di numero di maglia che diamo ad ogni vettore/giocatore: il colore della maglia ci dice di quale squadra si tratta e il numero sulla maglia individua univocamente il giocatore. Lo stesso numero su maglie di colore diverso individua giocatori diversi così come la stessa n-upla di numeri reali rispetto a basi diverse individua vettori diversi. Ad esempio, se in $\mathbb{R}^2$ scegliamo le due basi $\mathcal{B} = \{v_1 = (1,2), v_2 = (1,1)\}$ e $\mathcal{C} = \{w_1 = (1,0), w_2 = (1,5)\}$, il vettore di coordinate $(1,1)_\mathcal{B}$ è il vettore $1v_1 + 1v_2 = 1(1,2) + 1(1,1) = (2,3)$, mentre $(1,1)_\mathcal{C}$ è il vettore $1w_1 + 1w_2 = 1(1,0) + 1(1,5) = (2,5)$, stesso numero ma su maglie diverse: giocatori diversi. In ogni caso, se lo spazio ha dimensione n, la scelta di una base ci assicura di poter scrivere ogni vettore come una $n - upla$ di numeri reali.

Come possiamo costruire una applicazione lineare nel modo più semplice possibile? Sia $L : V \to W$ una applicazione lineare tra due spazi di dimensione finita, con $\dim V = n$ e $\dim W = m$; scegliamo poi una base per ciascuno spazio: $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ e $\mathcal{C} = \{w_1, w_2, \ldots, w_m\}$. Abbiamo visto in 7.1.12.iv) (e lo ricordiamo) che conoscere i valori di L sui vettori di una base significa conoscere l'applicazione lineare interamente. Ora, poiché abbiamo fissato una base di W, ogni vettore $L(v_1), \ldots, L(v_n)$ si può scrivere in modo unico come combinazione lineare dei vettori di questa base:

$$\begin{aligned}
L(v_1) &= a_{11}w_1 + a_{21}w_2 + \cdots + a_{m1}w_m \\
L(v_2) &= a_{12}w_1 + a_{22}w_2 + \cdots + a_{m2}w_m \\
&\vdots \\
L(v_n) &= a_{1n}w_1 + a_{2n}w_2 + \ldots + a_{mn}w_m.
\end{aligned}$$

I termini a_{ij} sono univocamente determinati dalle scelte da noi fatte delle basi di V e W (si noti che fissare una base significa fissare anche l'ordine dei suoi elementi). Possiamo allora costruire la seguente matrice:

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

Tale matrice in $\mathcal{M}_{m,n}(\mathbb{R})$ dipende dalla scelta delle due basi $\{v_1, v_2, \ldots, v_n\}$ e $\{w_1, w_2, \ldots, w_m\}$, ma caratterizza completamente la nostra applicazione lineare L. La matrice A si dice **la matrice associata alla applicazione lineare L rispetto alle basi $\mathcal{B}$ e $\mathcal{C}$**. Osserviamo che, una volta fissate le basi $\mathcal{B}$ e $\mathcal{C}$, le colonne di A sono le coordinate rispetto alla base $\mathcal{C}$ dei vettori

$L(v_1), \dots, L(v_n)$. Quindi l'immagine di un vettore $v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n \in V$ si può esprimere in coordinate rispetto alla base $\{w_1, w_2, \dots, w_m\}$, come segue:

$$\begin{aligned} L(v) &= \lambda_1 L(v_1) + \lambda_2 L(v_2) + \cdots + \lambda_n L(v_n) = \\ &= \lambda_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \dots \\ \dots \\ a_{m1} \end{pmatrix} + \lambda_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \dots \\ \dots \\ a_{m2} \end{pmatrix} + \cdots + \lambda_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \dots \\ \dots \\ a_{mn} \end{pmatrix} = \\ &= \begin{pmatrix} \lambda_1 a_{11} + \lambda_2 a_{12} + \ldots + \lambda_n a_{1n} \\ \lambda_1 a_{21} + \lambda_2 a_{22} + \ldots + \lambda_n a_{2n} \\ \dots \quad \dots \quad \dots \\ \dots \quad \dots \quad \dots \\ \lambda_1 a_{m1} + \lambda_2 a_{m2} + \ldots + \lambda_n a_{mn} \end{pmatrix}. \end{aligned}$$

(Si noti che le coordinate dei vettori $L(v_i)$ sono indicate come vettori colonna).

Questo ci permette di affermare che l'immagine del vettore $v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n$ è data, in coordinate rispetto alla base $\mathcal{C}$, dal prodotto righe per colonne

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \dots \\ \dots \\ \lambda_n \end{pmatrix}.$$

Esempio 7.3.1 *i*) Consideriamo la base $\mathcal{B} = \{v_1 = (1,1,0),\ v_2 = (0,1,0),\ v_3 = (2,0,1)\}$ di $\mathbb{R}^3$ e la base $\mathcal{C} = \{w_1 = (1,0), w_2 = (1,1)\}$ di $\mathbb{R}^2$ e sia $L : \mathbb{R}^3 \to \mathbb{R}^2$ l'applicazione lineare definita da $L(v_1) = (1,-1), L(v_2) = (0,-1), L(v_3) = (2,1)$. Chi è l'immagine mediante L di un vettore di $\mathbb{R}^3$, ad esempio $v = (1,1,1)$? Scriviamo prima di tutto v rispetto alla base $\mathcal{B}$. Si ha $v = (1,1,1) = \lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 = \lambda_1(1,1,0) + \lambda_2(0,1,0) + \lambda_3(2,0,1) = (\lambda_1 + 2\lambda_3, \lambda_1 + \lambda_2, \lambda_3)$, cioè: $\lambda_3 = 1$, $\lambda_2 = 2$, $\lambda_1 = -1$. Pertanto, per la linearità di L, abbiamo: $L(1,1,1) = -1L(v_1) + 2L(v_2) + 1L(v_3) = -(1,-1) + 2(0,-1) + (2,1) = (1,0)$. Più in generale, per ogni vettore di coordinate $(\lambda_1, \lambda_2, \lambda_3)$ nella base v_1, v_2, v_3, si ha $L(\lambda_1, \lambda_2, \lambda_3) = \lambda_1(1,-1) + \lambda_2(0,-1) + \lambda_3(2,1) = (\lambda_1 + 2\lambda_3, -\lambda_1 + \lambda_3 - \lambda_2)$. Cerchiamo la matrice associata a L rispetto alle basi fissate. Quale forma avrà tale matrice? Il numero di colonne è uguale alla dimensione del dominio e il numero di righe uguale alla dimensione del codominio. Quindi stiamo cercando una matrice in $\mathcal{M}_{2,3}(\mathbb{R})$. Le sue colonne sono date dalle coordinate nella base $\mathcal{C}$ delle immagini dei vettori della base fissata nel dominio: $L(v_1) = (1, -1) = 2(1, 0) - (1, 1)$, $L(v_2) = (0, -1) = (1, 0) - (1, 1)$, $L(v_3) = (2, 1) = (1, 0) + (1, 1)$. Otteniamo così la matrice

$$\begin{pmatrix} 2 & 1 & 1 \\ -1 & -1 & 1 \end{pmatrix}.$$

L'immagine del vettore di $\mathbb{R}^3$ che nella base $\{v_1, v_2, v_3\}$ ha coordinate $(1, 2, 3)$ è il vettore

$$\begin{pmatrix} 2 & 1 & 1 \\ -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 7 \\ 0 \end{pmatrix}$$

espresso in coordinate rispetto alla base w_1 e w_2, si tratta cioè del vettore $7(1, 0) + 0(1, 1) = (7, 0)$.

ii) Cosa succede se nell'esempio precedente cambiamo le basi degli spazi V e W? Prendiamo ad esempio la base $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, $e_3 = (0, 0, 1)$ per il dominio e la base $e' = \{e'_1 = (1, 0),\ e'_2 = (0, 1)\}$ per il codominio. Allora per descrivere la matrice associata a L rispetto alle nuove basi dobbiamo calcolare le coordinate dei vettori $L(e_1), L(e_2), L(e_3)$ nella base e'_1, e'_2. Ora $e_1 = v_1 - v_2$ e quindi $L(e_1) = L(v_1 - v_2) = L(v_1) - L(v_2) = (1, -1) - (0, -1) = (1, 0)$; $e_2 = v_2$ e dunque $L(v_2) = L(e_2) = (0, -1)$, infine $e_3 = -2v_1 + 2v_2 + v_3$ quindi $L(e_3) = L(-2v_1 + 2v_2 + v_3) = -2L(v_1) + 2L(v_2) + L(v_3) = -2(1, -1) + 2(0, -1) + (2, 1) = (0, 1)$. Tali immagini sono espresse nella base $\{e'_1, e'_2\}$ come segue: $L(e_1) = 1e'_1 + 0e'_2 = (1, 0)_{e'}$, $L(e_2) = 0e'_1 + (-1)e'_2 = (0, -1)_{e'}$, $L(e_3) = 0e'_1 + 1e'_2 = (0, 1)_{e'}$. Otteniamo pertanto la matrice

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

Osservazione 7.3.2 Sia V uno spazio vettoriale di dimensione n su $\mathbb{R}$. Fissare una base di V significa costruire un'applicazione lineare biunivoca da V in $\mathbb{R}^n$. Infatti sia $v_1, \dots, v_n$ una base di V. Allora ogni vettore v di V si scrive in modo unico come combinazione lineare di $v_1, \dots, v_n$: $v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n$. Definiamo la funzione

$$\varphi : V \longrightarrow \mathbb{R}^n$$

che associa ad ogni vettore $v \in V$ le sue coordinate nella base $v_1, \ldots, v_n$: $(\lambda_1, \lambda_2, \ldots, \lambda_n)$. La funzione φ è lineare: se $v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n$ e $v' = \lambda'_1 v_1 + \lambda'_2 v_2 + \cdots + \lambda'_n v_n$ sono due vettori di V allora $v + v' = (\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n) + (\lambda'_1 v_1 + \lambda'_2 v_2 + \cdots + \lambda'_n v_n) = (\lambda_1 + \lambda'_1) v_1 + (\lambda_2 + \lambda'_2) v_2 + \cdots + (\lambda_n + \lambda'_n) v_n$, cosicché $\varphi(v + v') = (\lambda_1 + \lambda'_1, \lambda_2 + \lambda'_2, \ldots, \lambda_n + \lambda'_n) = (\lambda_1, \lambda_2, \ldots, \lambda_n) + (\lambda'_1, \lambda'_2, \ldots, \lambda'_n) = \varphi(v) + \varphi(v')$. Ancora, se $\lambda \in \mathbb{R}$ e $v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n$, allora $\lambda v = \lambda(\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n) = \lambda\lambda_1 v_1 + \lambda\lambda_2 v_2 + \cdots + \lambda\lambda_n v_n$, dunque $\varphi(\lambda v) = (\lambda\lambda_1, \lambda\lambda_2, \ldots, \lambda\lambda_n) = \lambda\varphi(v)$. Quindi φ è lineare. Ed è in particolare una funzione lineare tra due spazi della stessa dimensione. Per il Teorema delle dimensioni se φ è iniettiva allora è pure suriettiva. Del resto φ è iniettiva perché se $\varphi(v) = \mathbf{0}_{\mathbb{R}^n} = (0, 0, \ldots, 0)$ significa che le coordinate del vettore v sono nulle cioè: $v = 0v_1 + 0v_2 + \ldots + 0v_n = \mathbf{0}_V$. La funzione è iniettiva e quindi un isomorfismo. In particolare questo significa che ragionare sui vettori di V equivale a ragionare sui vettori pensati attraverso le loro coordinate rispetto ad una base fissata o, equivalentemente, che ogni spazio vettoriale reale di dimensione n è “sostanzialmente” $\mathbb{R}^n$.

Da quanto detto sopra dovrebbe risultare chiaro come poter calcolare la dimensione dell'immagine di una applicazione lineare quando è data la matrice ad essa associata rispetto a basi fissate. Spieghiamolo meglio: siano allora $L : V \to W$ un'applicazione lineare, $\mathbf{v} = \{v_1, \ldots, v_n\}$ una base di V e $\mathbf{w} = \{w_1, w_2, \ldots, w_m\}$ una base di W. Le coordinate rispetto a $\mathbf{w}$ dei vettori $L(v_1)$, $L(v_2), \ldots, L(v_n)$ sono le colonne della matrice $A \in \mathcal{M}_{m,n}(\mathbb{R})$ associata ad L. Ora dobbiamo calcolare $\dim(\mathrm{Im}L) = \dim\langle L(v_1), L(v_2), \ldots, L(v_n)\rangle$ (vedi 7.2), cioè la dimensione del sottospazio vettoriale generato da $L(v_1)$, $L(v_2)$, …, $L(v_n)$. Questo è equivalente a determinare il massimo numero di vettori linearmente indipendenti tra $L(v_1)$, $L(v_2)$, …, $L(v_n)$ cioè il massimo numero di colonne di A linearmente indipendenti, pensando ogni colonna di A come un vettore di coordinate rispetto alla base $\{w_1, w_2, \ldots, w_m\}$, e quindi come un vettore di $\mathbb{R}^m$. Grazie alla Definizione 5.3.3 e al Teorema 5.3.4 , si ha dunque: $\dim(\mathrm{Im}L) = \mathrm{rg}A$.

Riassumendo, se $L : V \to W$ è una applicazione lineare tra due $\mathbb{R}$-spazi vettoriali, con $\dim V = n$ e $\dim W = m$, e se la matrice associata a tale applicazione lineare rispetto ad una base fissata del dominio ed una del codominio è $A \in \mathcal{M}_{m,n}(\mathbb{R})$, allora $\mathrm{rg}A = \dim(\mathrm{Im}L)$. Inoltre per ogni possibile scelta di basi di V, W tutte le matrici associate a L appartengono all'insieme $\mathcal{M}_{m,n}(\mathbb{R})$ e tutte hanno rango uguale a $\dim(\mathrm{Im}L)$. Quindi il rango dipende solo dall'applicazione lineare L e non dalle basi scelte nel dominio e nel codominio.

7.4 Esercizi svolti

Esercizio 7.4.1 Stabilire quali delle seguenti applicazioni sono lineari:

i) $f_1 : \mathbb{R} \longrightarrow \mathbb{R}$, $f_1(x) = x + 3$.

ii) $f_2 : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$, $f_2(x, y) = (x^2, y)$.

iii) $f_3 : \mathbb{R}^2 \longrightarrow \mathbb{R}$, $f_3(x, y) = 2x + 3y$.

iv) $f_4 : \mathcal{M}_2(\mathbb{R}) \longrightarrow \mathbb{R}^2$, $f_4 \begin{pmatrix} a & b \\ c & d \end{pmatrix} = (a + c, b + d)$.

v) $f_5 : \mathcal{M}_n(\mathbb{R}) \longrightarrow \mathcal{M}_n(\mathbb{R})$, $f_5(A) = 2A$.

Svolgimento.

i) L'immagine mediante un'applicazione lineare del vettore nullo del dominio è sempre il vettore nullo del codominio, ma $f_1(0) = 3$ quindi f_1 non è un'applicazione lineare.

ii) L'applicazione f_2 non è lineare dal momento che $f_2((1, 0) + (-1, 0)) = f_2(0, 0) = (0, 0) \neq f_2(1, 0) + f_2(-1, 0) = (1, 0) + (1, 0) = (2, 0)$.

iii) Per verificare che un'applicazione $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ è un'applicazione lineare occorre verificare le due seguenti condizioni:

   1) per ogni coppia di vettori v e w in $\mathbb{R}^2$: $f(v + w) = f(v) + f(w)$;
   2) per ogni $v \in \mathbb{R}^2$ e per ogni $\alpha \in \mathbb{R}$ $f(\alpha v) = \alpha f(v)$.

   Consideriamo dunque l'applicazione f_3 e siano $v = (a, b)$ e $w = (c, d)$ due elementi di $\mathbb{R}^2$. Abbiamo:

   $f_3(v+w) = f_3(a+c, b+d) = 2(a+c)+3(b+d) = (2a+3b)+(2c+3d) = f_3(v) + f_3(w)$, quindi la proprietà 1) è verificata.

   Analogamente, preso α in $\mathbb{R}$,

$$f_3(\alpha v) = f_3(\alpha a, \alpha b) = 2\alpha a + 3\alpha b = \alpha(2a + 3b) = \alpha f_3(v).$$

   Possiamo concludere che l'applicazione f_3 è lineare.

Analogamente si procede per dimostrare che le applicazioni f_4 e f_5 sono lineari.

Esercizio 7.4.2 Sia $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ l'applicazione definita da:

$$f(x, y, z) = (x + y, x + y, z).$$

i) Scrivere la matrice associata a f rispetto alla base canonica nel dominio e nel codominio.

ii) Determinare Ker$f$ e Imf.

iii) Scrivere la matrice associata a f rispetto alla base canonica nel dominio e alla base $\{v_1 = (1, 1, 1), v_2 = (1, 1, 0), v_3 = (1, -1, 0)\}$ nel codominio.

iv) Scrivere la matrice associata a f rispetto alla base $\{v_1, v_2, v_3\}$ nel dominio e alla base canonica nel codominio.

Svolgimento.

i) La matrice associata a f rispetto alla base canonica nel dominio e nel codominio è la matrice che ha sulle colonne le coordinate rispetto alla base canonica di $\mathbb{R}^3$ delle immagini mediante f dei vettori della stessa base. Calcoliamo dunque:

$$f(1, 0, 0) = (1, 1, 0) = 1(1, 0, 0) + 1(0, 1, 0) + 0(0, 0, 1)$$

$$f(0, 1, 0) = (1, 1, 0) = 1(1, 0, 0) + 1(0, 1, 0) + 0(0, 0, 1)$$

$$f(0, 0, 1) = (0, 0, 1) = 0(1, 0, 0) + 0(0, 1, 0) + 1(0, 0, 1).$$

Dunque la matrice associata a f rispetto alla base canonica nel dominio e nel codominio è:

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

ii) Sappiamo che Imf è generata dai vettori $f(1, 0, 0)$, $f(0, 1, 0)$ e $f(0, 0, 1)$. Dunque Im$f = \langle(1, 1, 0), (0, 0, 1)\rangle$. Usando il Teorema delle dimensioni 7.2.2 otteniamo che Ker(f) ha dimensione 1. Del resto, dal momento che $f(1, 0, 0) = (1, 1, 0) = f(0, 1, 0)$, per la linearità di f il vettore $(1, 0, 0) - (0, 1, 0) = (1, -1, 0)$ ha come immagine mediante f il vettore nullo: $f((1, 0, 0) - (0, 1, 0)) = f(1, 0, 0) - f(0, 1, 0) = 0_{\mathbb{R}^3}$. Dunque Ker$(f) = \langle(1, -1, 0)\rangle$.

iii) La matrice richiesta ha sulle colonne le coordinate nella base $\{v_1, v_2, v_3\}$ delle immagini dei vettori della base canonica. Abbiamo già determinato le immagini, tramite f, dei vettori della base canonica. Si tratta ora di esprimere queste immagini in coordinate rispetto alla base $\{v_1, v_2, v_3\}$. Abbiamo:

$$\begin{array}{rcccl} f(1,0,0) & = & (1,1,0) & = & 0v_1 + 1v_2 + 0v_3 \\ f(0,1,0) & = & (1,1,0) & = & 0v_1 + 1v_2 + 0v_3 \\ f(0,0,1) & = & (0,0,1) & = & 1v_1 - 1v_2 + 0v_3. \end{array}$$

La matrice richiesta è dunque:

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

iv) In questo caso la base fissata nel dominio è la base $\{v_1, v_2, v_3\}$. Calcoliamo allora le immagini dei vettori v_1, v_2, v_3 e determiniamo le coordinate dei vettori trovati rispetto alla base canonica:

$$\begin{array}{rcccl} f(v_1) & = & (2,2,1) & = & 2(1,0,0) + 2(0,1,0) + 1(0,0,1) \\ f(v_2) & = & (2,2,0) & = & 2(1,0,0) + 2(0,1,0) + 0(0,0,1) \\ f(v_3) & = & (0,0,0). & & \end{array}$$

La matrice richiesta è dunque:

$$\begin{pmatrix} 2 & 2 & 0 \\ 2 & 2 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Esercizio 7.4.3 Sia $D : \mathbb{R}^{\leq 3}[x] \longrightarrow \mathbb{R}^{\leq 3}[x]$ l'applicazione derivata (rispetto alla variabile x). Determinare $\mathrm{Ker}D$, $\mathrm{Im}D$ e la matrice associata a D rispetto alla base $\mathcal{B} = \{1, x, x^2, x^3\}$ di $\mathbb{R}^{\leq 3}[x]$.

Svolgimento. Ricordiamo la definizione di derivata di un polinomio in una variabile (di grado ≤ 3):

$$D(a_0 + a_1x + a_2x^2 + a_3x^3) = a_1 + 2a_2x + 3a_3x^2.$$

Si verifica immediatamente, usando questa definizione, che la derivata è un'applicazione lineare. Per definizione di nucleo di un'applicazione lineare,

$$\begin{array}{rcl} \mathrm{Ker}(D) & = & \{p(x) = a_0 + a_1x + a_2x^2 + a_3x^3 \mid D(p(x)) = 0\} = \\ & = & \{p(x) = a_0 + a_1x + a_2x^2 + a_3x^3 \mid a_1 + 2a_2x + 3a_3x^2 = 0\}. \end{array}$$

Dunque

$$\begin{aligned}\mathrm{Ker}(D) &= \{p(x) = a_0 + a_1x + a_2x^2 + a_3x^3 \mid a_1 = a_2 = a_3 = 0\} = \\ &= \mathbb{R}.\end{aligned}$$

Usando 7.2.2 deduciamo immediatamente che l'immagine di D ha dimensione 3. Del resto

$$\begin{aligned}\mathrm{Im}D &= \langle D(1), D(x), D(x^2), D(x^3)\rangle = \\ &= \langle 1, 2x, 3x^2\rangle.\end{aligned}$$

La matrice associata a D rispetto alla base $\mathcal{B}$ è:

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

Esercizio 7.4.4 Possiamo, nelle ipotesi precedenti, considerare l'applicazione composta $D \circ D = D^2$, cioè la derivata seconda nella variabile x, come applicazione di $\mathbb{R}^{\leq 3}[x]$ in se stesso. Si verifichi che D^2 è una applicazione lineare. Definiamo allora l'applicazione che ad ogni polinomio di $\mathbb{R}^{\leq 3}[x]$ associa la sua derivata seconda meno la sua derivata prima: $P \longmapsto D^2(P) - D(P)$. Si verifichi che anche questo è un endomorfismo di $\mathbb{R}^{\leq 3}[x]$. In particolare si studi il suo nucleo. Indicando, secondo la consuetudine, la derivata prima e la derivata seconda di P rispettivamente con P' e P'', il nucleo dell'applicazione costruita è allora dato dall'insieme di polinomi P che soddisfano la relazione

$$P'' - P' = 0.$$

L'equazione trovata è detta equazione differenziale.

Esercizio 7.4.5 Tra le applicazioni lineari dell'esercizio 7.4.1 si determinino quelle iniettive e quelle suriettive.

Svolgimento. Le funzioni f_1 e f_2 non sono lineari.
La funzione f_3 non è iniettiva:$f_3(-3,2) = 0$, dunque $(-3,2)$ è un vettore non nullo del nucleo di f_3. D'altra parte dal teorema delle dimensioni si può concludere che $\dim \mathbb{R}^2 = \dim \mathrm{Ker}(f_3) + \dim \mathrm{Im}(f_3)$ da cui $\dim \mathrm{Ker}(f_3) = 2 - \dim \mathrm{Im}(f_3) \geq 1$, quindi non esistono applicazioni lineari da $\mathbb{R}^2$ a $\mathbb{R}$ iniettive (vedi anche 7.2.3.iii). Nel nostro caso f_3 è suriettiva, infatti $\mathrm{Im}f_3$ contiene il vettore $f_3(1,0) = 2$ dunque ha dimensione almeno 1. Del resto $\dim(\mathbb{R}) = 1$, quindi $\mathrm{Im}f_3 = \mathbb{R}$.

Nello stesso modo f_4 è un'applicazione lineare suriettiva ma non iniettiva.

La funzione f_5 è un endomorfismo, dunque essa è iniettiva se e solo se è suriettiva. Del resto f_5 è ovviamente suriettiva e quindi biiettiva.

Esercizio 7.4.6 Sia $L : V \longrightarrow W$ un'applicazione lineare e siano $\{v_1, \dots, v_n\}$ vettori linearmente indipendenti di V. Mostrare che se L è iniettiva allora i vettori $\{L(v_1), \dots, L(v_n)\}$ di W sono linearmente indipendenti.

Svolgimento. Consideriamo una combinazione lineare dei vettori $L(v_1)$, $\dots, L(v_n)$ e supponiamo che essa sia uguale al vettore nullo di W:

$$\alpha_1 L(v_1) +_W \cdots +_W \alpha_n L(v_n) = \mathbf{0}_W.$$

Per la linearità di L si ha:

$$0_W = \alpha_1 L(v_1) +_W \cdots +_W \alpha_n L(v_n) = L(\alpha_1 v_1 +_V \cdots +_V \alpha_n v_n).$$

Essendo L iniettiva per ipotesi, l'unico vettore di V che ha come immagine 0_W è il vettore nullo di V, quindi:

$$\alpha_1 v_1 +_V \cdots +_V \alpha_n v_n = \mathbf{0}_V$$

e questo implica $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$ dal momento che i vettori $v_1, \dots, v_n$ sono linearmente indipendenti. L'esercizio è concluso.

Esercizio 7.4.7

i) È possibile costruire una applicazione lineare iniettiva da $\mathbb{R}^3$ in $\mathbb{R}^2$? E un'applicazione suriettiva? In caso affermativo si facciano degli esempi.

ii) È possibile costruire una applicazione lineare iniettiva da $\mathbb{R}^2$ in $\mathbb{R}^3$? E un'applicazione suriettiva? In caso affermativo si facciano degli esempi.

Svolgimento.

i) Dall'osservazione 7.2.3iii), sappiamo già che non vi sono applicazioni lineari iniettive fra $\mathbb{R}^3$ e $\mathbb{R}^2$. Come ulteriore verifica, notiamo che nell'esercizio precedente abbiamo mostrato che un'applicazione lineare iniettiva manda vettori linearmente indipendenti in vettori linearmente indipendenti dunque non esiste un'applicazione lineare iniettiva da uno spazio di dimensione n in uno spazio di dimensione m se $m < n$.

   Al contrario è certamente possibile costruire un'applicazione lineare suriettiva da $\mathbb{R}^3$ in $\mathbb{R}^2$, ad esempio l'applicazione lineare f definita da: $f(1,0,0) = (1,0)$, $f(0,1,0) = (0,1)$, $f(0,0,1) = (1,1)$ è suriettiva dal momento che $\mathrm{Im} f = \langle (1,0), (0,1) \rangle = \mathbb{R}^2$.

ii) Sia $g : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ l'applicazione lineare definita da: $g(1,0) = (1,0,0)$, $g(0,1) = (0,1,0)$. L'applicazione lineare g è iniettiva, infatti $\mathrm{Im} g = \langle g(1,0), g(0,1)\rangle = \langle (1,0,0),(0,1,0)\rangle$ ha dimensione 2, di conseguenza il nucleo di g ha dimensione 0, cioè è banale.

D'altra parte, sempre per 7.2.3iii), l'immagine di un'applicazione lineare ha dimensione minore della dimensione del dominio o uguale ad essa. Dunque non è possibile costruire un'applicazione lineare suriettiva da $\mathbb{R}^2$ in $\mathbb{R}^3$.

Esercizio 7.4.8 i) Esiste un'applicazione lineare $g : \mathbb{R}^4 \longrightarrow \mathbb{R}^2$ tale che $g(1,0,0,0) = (3,4)$, $g(2,3,4,0) = (1,2)$, e $g(0,3,4,0) = (6,5)$? In caso affermativo se ne faccia un esempio.

ii) Esiste un'applicazione lineare $h : \mathbb{R}^4 \longrightarrow \mathbb{R}^2$ tale che $h(1,0,0,0) = (3,4)$, $h(2,3,4,0) = (1,2)$, $h(0,0,3,2) = (2,8)$? In caso affermativo le si descrivano tutte.

Svolgimento.

i) Osserviamo innanzitutto che $(0,3,4,0) = (2,3,4,0) - 2(1,0,0,0)$. Pertanto, se esistesse una funzione lineare g come richiesta, si avrebbe: $g(0,3,4,0) = g(2,3,4,0) - 2g(1,0,0,0)$, cioè $(6,5) = (1,2) - 2(3,4)$, ma questo non è possibile.

ii) Osserviamo che i vettori $(1,0,0,0)$, $(2,3,4,0)$, $(0,0,3,2)$ sono linearmente indipendenti (si verifichi!) dunque è possibile definire un'applicazione lineare fissando liberamente le loro immagini. In particolare esiste sicuramente un'applicazione lineare h come richiesta.

Per definire un'applicazione lineare è sufficiente definire le immagini dei vettori di una base del dominio. Nel nostro caso possiamo aggiungere all'insieme $(1,0,0,0)$, $(2,3,4,0)$, $(0,0,3,2)$ il vettore $(0,0,0,1)$ per ottenere una base di $\mathbb{R}^4$. Allora una qualsiasi applicazione h soddisfacente le ipotesi dell'esercizio sarà definita da $h(1,0,0,0) = (3,4)$, $h(2,3,4,0) = (1,2)$, $h(0,0,3,2) = (2,8)$ $h(0,0,0,1) = (a,b)$ al variare di a e b in $\mathbb{R}$. Notiamo, in particolare, che esistono infinite applicazioni lineari soddisfacenti le ipotesi.

Esercizio 7.4.9 Sia L l'endomorfismo di $\mathbb{R}^3$ che rispetto alla base $\mathcal{B} = \{(2,1,2),\ (3,1,1),\ (1,0,1)\}$ ha come matrice

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

Determinare l'immagine mediante L del vettore $w = (1,1,1)$. Determinare, inoltre, nucleo e immagine di L.

Svolgimento. La matrice A ha sulle colonne le coordinate nella base $\mathcal{B}$ delle immagini dei vettori della base $\mathcal{B}$ stessa. Pertanto, se (a,b,c) sono le coordinate di un vettore v nella base $\mathcal{B}$, $L(v) = A \begin{pmatrix} a \\ b \\ c \end{pmatrix}$ saranno le coordinate nella base $\mathcal{B}$ del vettore $L(v)$.

Dobbiamo dunque, innanzitutto, calcolare le coordinate del vettore $(1,1,1)$ nella base $\mathcal{B}$. Si verifica facilmente che $(1,1,1) = 1(2,1,2) + 0(3,1,1) - 1(1,0,1)$, cioè $(1,1,1) = (1,0,-1)_{\mathcal{B}}$. Pertanto

$$L(w) = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 3 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}$$

cioè $L(w) = (2,-1,0)_{\mathcal{B}} = 2(2,1,2) - (3,1,1) = (1,1,3)$.

Chi è il nucleo dell'applicazione L? È, per definizione, l'insieme dei vettori $(x_1,x_2,x_3)_{\mathcal{B}}$ tali che sia

$$\begin{pmatrix} 2 & 1 & 0 \\ 0 & 3 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

cioè i vettori $(x_1,x_2,x_3)_{\mathcal{B}}$ tali che $(2x_1+x_2, 3x_2+x_3, x_1+x_2+x_3) = (0,0,0)$, i.e., $x_1 = x_2 = x_3 = 0$. Dunque $\mathrm{Ker}L = \mathbf{0}_{\mathbb{R}^3}$ e la funzione L è iniettiva. Di conseguenza L è suriettiva (L è un endomorfismo) pertanto $ImL = \mathbb{R}^3$ e quindi il rango colonne di A è $\mathrm{rg}A = 3$.

7.5 Esercizi proposti

Esercizio 7.5.1 Stabilire quali delle seguenti applicazioni sono lineari:

1. $f : \mathbb{R}^2 \to \mathbb{R}^2,\ \ f(x, y) = (x - y + 2, x + y)$;
2. $g : \mathbb{R}^3 \to \mathbb{R}^2,\ \ g(x, y, z) = (x^2 - y^2, x + y)$;
3. $h : \mathbb{R}^2 \to \mathcal{M}_2(\mathbb{R}),\ \ h(x, y) = \begin{pmatrix} x - y & 0 \\ 0 & 2x - 3y \end{pmatrix}$.

Esercizio 7.5.2 Si consideri, al variare di $k \in \mathbb{R}$, l'applicazione $f_k : \mathbb{R}^2 \to \mathbb{R}^3$, $f_k(x, y) = (x - y, ky, ky)$. Stabilire per quali valori di k l'applicazione f_k è lineare. Per i valori di k trovati:

1. Scrivere la matrice associata ad f_k rispetto alla base $\mathcal{B} = \{(1, 1), (1, 2)\}$ di $\mathbb{R}^2$ ed alla base canonica di $\mathbb{R}^3$.
2. Determinare nucleo e immagine di f_k.
3. Determinare i valori di k tali che il vettore $(1, 0, 0)$ appartenga all'immagine di f_k.

Esercizio 7.5.3 Costruire una applicazione lineare $f : \mathbb{R}^3 \to \mathbb{R}^2$ non nulla tale che:

1. f non sia suriettiva;
2. il nucleo di f contenga i vettori $(2, 2, 2)$, $(1, 1, -1)$.

Scrivere la matrice associata all'applicazione f costruita, rispetto alle basi canoniche di $\mathbb{R}^3$ ed $\mathbb{R}^2$. L'applicazione f richiesta è unica?

Esercizio 7.5.4 Si consideri l'applicazione lineare $L : \mathbb{R}^{\leq 2}[x] \to \mathbb{R}^3,\ \ L(a + bx + cx^2) = (a + b, a + c, b - c)$.

1. Dopo aver fissato una base di $\mathbb{R}^{\leq 2}[x]$ ed una base di $\mathbb{R}^3$, determinare la matrice associata ad L rispetto a tali basi.
2. Determinare nucleo ed immagine di L. L'applicazione L è iniettiva? È suriettiva?

Esercizio 7.5.5 Sia $f : \mathcal{M}_2(\mathbb{R}) \to \mathbb{R}^2$ l'applicazione così definita:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \longmapsto (b - c, b + c).$$

1. Mostrare che l'applicazione f è lineare.
2. Scrivere la matrice associata ad f rispetto alle basi canoniche di $\mathcal{M}_2(\mathbb{R})$ e $\mathbb{R}^2$.
3. Determinare nucleo e immagine di f.
4. Determinare la controimmagine mediante f del sottospazio vettoriale $S = \langle (0, 1) \rangle$ di $\mathbb{R}^2$.

Lezione 8

Matrici

Abbiamo visto come le matrici $\mathcal{M}_{m,n}(\mathbb{R})$ rappresentino l'insieme delle applicazioni lineari tra due spazi V e W di dimensione rispettivamente n e m, una volta scelta una base per ciascuno dei due spazi (ricordiamo che la scelta delle basi corrisponde ad identificare ogni vettore di V con un elemento di $\mathbb{R}^n$ e ogni elemento di W con uno di $\mathbb{R}^m$).

Se componessimo due applicazioni lineari, la funzione composta sarebbe ancora lineare? E che cosa potremmo dire della matrice associata all'applicazione composta?

In questo paragrafo mostreremo che la composizione di due funzioni lineari è ancora un'applicazione lineare ed introdurremo il prodotto (righe per colonne) di matrici.

8.1 Prodotto righe per colonne

Problema. Dati tre spazi vettoriali V,W e Z, di dimensione, rispettivamente, n, m e p, per ciascuno di essi scegliamo una base: $\mathcal{V} = \{v_1, \ldots, v_n\}$, $\mathcal{W} = \{w_1, \ldots, w_m\}$ e $\mathcal{Z} = \{z_1, \ldots, z_p\}$. Siano poi $f : V \to W$ e $g : W \to Z$ funzioni lineari. Grazie alla scelta delle basi possiamo associare ad f una matrice $A_f \in \mathcal{M}_{m,n}(\mathbb{R})$ e a g una matrice $A_g \in \mathcal{M}_{p,m}(\mathbb{R})$.

1. La composizione $g \circ f : V \to Z$ è lineare?

2. Se la risposta alla precedente domanda è affermativa, possiamo associare a $g \circ f$ la matrice $A_{g\circ f} \in \mathcal{M}_{p,n}(\mathbb{R})$. È possibile determinare la matrice $A_{g\circ f}$ conoscendo A_f e A_g?

Soluzione del problema. Per prima cosa dimostriamo che la composizione di due applicazioni lineari è ancora un'applicazione lineare. Siano quindi f e g due applicazioni lineari e sia $g \circ f : V \to Z$ la loro composizione. Per ogni $v \in V$ si ha $(g \circ f)(v) = g(f(v))$ con $f(v) \in W$. Ora, se prendiamo $v_1, v_2 \in V$,

$$\begin{aligned}(g \circ f)(v_1 +_V v_2) &= g(f(v_1 +_V v_2)) = g(f(v_1) +_W f(v_2)) = \\ &= g(f(v_1)) +_Z g(f(v_2)) = (g \circ f)(v_1) +_Z (g \circ f)(v_2)\end{aligned}$$

dove abbiamo usato la linearità di f e g. Sempre per linearità, se $\lambda \in \mathbb{R}$ e $v \in V$ allora $(g \circ f)(\lambda v) = g(f(\lambda v)) = g(\lambda f(v)) = \lambda g(f(v)) = \lambda(g \circ f)(v)$. Quindi la funzione composta $g \circ f$ è lineare.

Come calcolare la matrice ad essa associata rispetto alle basi $\mathcal{V}$ del dominio e $\mathcal{Z}$ del codominio? Sfruttiamo quello che sappiamo: le colonne di tale matrice sono le coordinate nella base $\mathcal{Z}$ delle immagini tramite $g \circ f$ dei vettori della base $\mathcal{V}$. Ad esempio, nella prima colonna dobbiamo mettere le coordinate del vettore $(g \circ f)(v_1)$, cioè $g(f(v_1))$, ma $f(v_1)$ è un vettore di W le cui coordinate nella base $\mathcal{W}$ sono le entrate della prima colonna di A_f. Quindi $g(f(v_1))$ è dato dal prodotto della matrice A_g per la prima colonna di A_f. Cioè, detta c_1 la prima colonna di A_f, la prima colonna di $A_{g\circ f}$ è $A_g c_1$, quindi una colonna formata da p righe. Si può (e si deve) fare lo stesso ragionamento per ogni vettore della base $\mathcal{V}$ e, quindi, per ogni colonna $c_1, c_2, \ldots, c_n$ di A_f. Mettendo una dopo l'altra tali colonne (nell'ordine) formate da p-entrate si ottiene una matrice a p-righe e n-colonne, in cui ogni colonna rappresenta le coordinate dei vettori $(g \circ f)(v_i)$, $i = 1, \ldots, n$, nella base $\mathcal{Z}$: è la matrice $A_{g\circ f}$ associata alla applicazione lineare $g \circ f$ rispetto alle basi $\mathcal{V}$ e $\mathcal{Z}$. Tale matrice è pertanto legata alle matrici A_f e A_g e si ha formalmente che, se

$$A_f = (a_{ij})_{i=1,\ldots,m}^{j=1,\ldots,n} \qquad A_g = (b_{ki})_{k=1,\ldots,p}^{i=1,\ldots,m},$$

allora

$$A_{g\circ f} = (c_{kj})_{k=1,\ldots,p}^{j=1,\ldots,n}$$

dove

$$c_{kj} = \sum_{i=1}^{m} b_{ki} a_{ij}.$$

In altre parole l'elemento che si trova nella riga k-esima e nella colonna j-esima di $A_{g\circ f}$ è il prodotto della riga k-esima di A_g per la colonna j-esima di A_f. In questo senso si chiama prodotto righe per colonne:

$$A_{g\circ f} = A_g A_f$$

con $A_f \in \mathcal{M}_{m,n}(\mathbb{R})$, $A_g \in \mathcal{M}_{p,m}(\mathbb{R})$ e $A_{g\circ f} \in \mathcal{M}_{p,n}(\mathbb{R})$.

Esempio 8.1.1 Consideriamo le due matrici

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

Calcoliamo il prodotto AB. La matrice AB rappresenta, secondo l'interpretazione precedente e con una opportuna scelta delle basi, la composizione di una applicazione tra due spazi di dimensione 2 (nelle notazioni precedenti $B = A_f$) e di una applicazione da uno spazio di dimensione 3 ad uno di dimensione 2 ($A = A_g$). Moltiplicare A per B non è possibile! Lo si può vedere direttamente scrivendo le matrici

$$AB = \begin{pmatrix} 2 & 1 & 3 \\ 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}:$$

il numero di entrate di ogni riga di A (i.e. il numero di colonne di A) non è uguale al numero di entrate di ogni colonna di B (i.e. il numero di righe di B). Viceversa è possibile calcolare il prodotto BA. Questo corrisponde alla composizione di una funzione lineare da uno spazio di dimensione 3 ad uno spazio di dimensione 2 seguita da un endomorfismo di uno spazio di dimensione 2 (nelle notazioni precedenti questa volta $A_g = B$ e $A_f = A$). Tutto questo è formalmente compatibile. Si ha:

$$BA = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 2 & 1 & 3 \\ 0 & -1 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 5 \\ 2 & -1 & 7 \end{pmatrix}$$

dove, ad esempio, l'entrata di posto 2, 2 di BA è data dal "prodotto" della seconda riga di B per la seconda colonna di A: $1(1) + 2(-1) = -1$.

Esempio 8.1.2 Il nostro esempio mostra che, affinché il prodotto AB tra due matrici A e B sia definito, occorre che il numero di colonne di A sia eguale al numero di righe di B, cioè $A \in \mathcal{M}_{n_1,n_2}(\mathbb{R})$ e $B \in \mathcal{M}_{n_2,n_3}(\mathbb{R})$. Allora la matrice prodotto avrà tante righe quante quelle di A e tante colonne quante quelle di B cioè : $AB \in \mathcal{M}_{n_1,n_3}(\mathbb{R})$.

Notiamo che il prodotto righe per colonne di due matrici A e B era già stato introdotto nel paragrafo 1.1 nel caso particolare in cui $A \in \mathcal{M}_{m,n}(\mathbb{R})$ e $B \in \mathcal{M}_{n,1}(\mathbb{R})$. Per quanto appena osservato questo prodotto è ben definito e dà luogo ad una matrice $m \times 1$ cioè ad un vettore colonna di $\mathbb{R}^m$.

Osservazione 8.1.3 Dall'esempio precedente emerge con evidenza il fatto che il prodotto di matrici non è commutativo. Addirittura nell'esempio precedente il prodotto BA risultava definito, ma non risultava definito il prodotto AB. Facciamo un altro esempio: consideriamo le matrici

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

Allora $AB = \begin{pmatrix} 3 & 3 \\ 3 & 3 \end{pmatrix}$ e $BA = \begin{pmatrix} 1 & 5 \\ 1 & 5 \end{pmatrix}$.

Osserviamo che, in modo equivalente, la composizione di applicazioni (lineari) non è un'operazione commutativa. Ad esempio, se consideriamo gli endomorfismi f e g di $\mathbb{R}^2$ di matrici A e B rispetto alla base canonica di $\mathbb{R}^2$, naturalmente $f \circ g \neq g \circ f$.

Osservazione 8.1.4 Il prodotto righe per colonne di matrici gode della proprietà associativa e della proprietà distributiva rispetto alla somma. Questo significa che, prese $A \in \mathcal{M}_{m,n}(\mathbb{R})$, $B \in \mathcal{M}_{p,m}(\mathbb{R})$ e infine $C \in \mathcal{M}_{q,p}(\mathbb{R})$: i prodotti $BA \in \mathcal{M}_{p,n}(\mathbb{R})$ e $CB \in \mathcal{M}_{q,m}(\mathbb{R})$ sono ben definiti, come pure i prodotti $C(BA) \in \mathcal{M}_{q,n}(\mathbb{R})$ e $(CB)A \in \mathcal{M}_{q,n}(\mathbb{R})$; si ha allora

$$C(BA) = (CB)A.$$

Per questo indicheremo questo prodotto semplicemente con CBA.

Inoltre, date $A_1 \in \mathcal{M}_{m,n}(\mathbb{R})$, $A_2 \in \mathcal{M}_{m,n}(\mathbb{R})$ e $B \in \mathcal{M}_{p,m}(\mathbb{R})$, allora $B(A_1 + A_2) = BA_1 + BA_2 \in \mathcal{M}_{p,n}(\mathbb{R})$ (distributività del prodotto rispetto alla somma).

Infine, per ogni $\alpha \in \mathbb{R}$, $B(\alpha A) = \alpha(BA)$.

Tutte le proprietà qui elencate possono essere dimostrate utilizzando la definizione di prodotto righe per colonne.

8.2 Matrici invertibili

Proposizione 8.2.1 *Se $f : V \to V$ è un endomorfismo invertibile, l'applicazione inversa di f è anch'essa un'applicazione lineare che indicheremo con f^{-1}.*

Dimostrazione. Dobbiamo verificare che f^{-1} è lineare. Dati quindi $v_1, v_2 \in V$ dobbiamo vedere che $f^{-1}(v_1+v_2) = f^{-1}(v_1)+f^{-1}(v_2)$. Essendo f biiettiva esiste un unico $w_1 \in V$ tale che $f(w_1) = v_1$, così come esiste un unico $w_2 \in V$ tale che $f(w_2) = v_2$, cioè $f^{-1}(v_1) = w_1$ e $f^{-1}(v_2) = w_2$. In particolare allora $f(w_1 + w_2) = v_1 + v_2$ e quindi $f^{-1}(v_1 + v_2) = w_1 + w_2$, cioè $f^{-1}(v_1 + v_2) = f^{-1}(v_1) + f^{-1}(v_2)$. Ancora: dobbiamo mostrare che $f^{-1}(\lambda v) = \lambda f^{-1}(v)$ per ogni reale λ e ogni vettore $v \in V$. Sia dunque $w \in V$ il solo elemento di V tale che $f(w) = v$, cioè $f^{-1}(v) = w$. Per la linearità di f si ha che $f(\lambda w) = \lambda v$ e quindi, poiché la funzione è biiettiva, $f^{-1}(\lambda v) = \lambda w = \lambda f^{-1}(v)$. Concludiamo che f^{-1} è lineare. **C. V. D.**

Abbiamo dimostrato che se $f : V \to V$ è lineare, invertibile allora $f^{-1} : V \to V$ è anch'essa una funzione lineare. Scegliamo allora una base $\mathcal{V} = \{v_1, \dots, v_n\}$ del dominio di f e una base $\mathcal{W} = \{w_1, \dots, w_n\}$ del suo codominio, scegliamo cioè due basi di V (che magari possono coincidere). Allora se indichiamo con $A \in \mathcal{M}_n(\mathbb{R})$ la matrice associata a f rispetto alle basi $\mathcal{V}$ e $\mathcal{W}$, e se denotiamo con $B \in \mathcal{M}_n(\mathbb{R})$ la matrice di f^{-1} rispetto alla base $\mathcal{W}$ nel dominio e $\mathcal{V}$ nel codominio, il prodotto $BA \in \mathcal{M}_n(\mathbb{R})$ è la matrice della applicazione identica $f^{-1} \circ f = id_V$ rispetto alla base $\mathcal{V}$ sia nel dominio che nel codominio. Tale matrice avrà nella prima colonna le coordinate dell'immagine di v_1 rispetto alla applicazione identica, cioè v_1 stesso, nella base $\{v_1, \dots, v_n\}$. Le sue coordinate sono $(1, 0, 0, \dots, 0)$, così come l'immagine di v_2 è v_2 e le sue coordinate nella base $\mathcal{V}$ sono $(0, 1, 0, \dots, 0)$. La matrice ottenuta si chiama matrice *identica di ordine* n e si indica con I_n: è una matrice quadrata di ordine n in cui tutte le entrate sono nulle salvo quelle sulla diagonale che sono uguali a 1.

Per lo stesso motivo si ha che $AB = I_n$ rappresenta l'applicazione identica di V in se stesso espressa rispetto alla base $\mathcal{W}$ sia nel dominio che nel codominio.

Osservazione 8.2.2 Si noti che se $C \in \mathcal{M}_{n,p}(\mathbb{R})$ e $I_n \in \mathcal{M}_n(\mathbb{R})$ allora $I_nC = C$. (Attenzione: il prodotto CI_n non è definito! Se, però, I_p è la matrice identica di ordine p allora $CI_p = C$).

Definizione 8.2.3 *Una matrice $A \in \mathcal{M}_n(\mathbb{R})$ si dice invertibile se esiste una matrice $B \in \mathcal{M}_n(\mathbb{R})$ tale che $AB = BA = I_n$. La matrice B si dice inversa di A e si denota con A^{-1}.*

Quali sono le proprietà delle matrici invertibili? Possiamo fare alcune brevi osservazioni sfruttando l'interpretazione di una matrice quadrata come applicazione lineare o, più precisamente, come endomorfismo. Sia $A \in \mathcal{M}_n(\mathbb{R})$

una matrice invertibile. Allora essa descrive un endomorfismo invertibile di $\mathbb{R}^n$. In particolare il rango di A è n e il numero di pivots in una qualunque forma a scalini per righe di A è n. In questa situazione non vi sono molte possibilità dal momento che la matrice ha ordine n: la forma a scalini deve avere tutti i pivots sulla diagonale e ogni elemento della diagonale deve essere un pivot. Questo fa sì che si possa ulteriormente raffinare la forma a scalini. Infatti, dopo opportune operazioni elementari sulle righe, la matrice sarà del tipo

$$\begin{pmatrix} \alpha_1 & * & * & * & \dots & \beta_1 \\ 0 & \alpha_2 & * & * & \dots & \beta_2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \alpha_{n-1} & \beta_{n-1} \\ 0 & 0 & 0 & 0 & 0 & \alpha_n \end{pmatrix}$$

dove $\alpha_i \neq 0$, e dove $*$ e β_j sono numeri reali. A questo punto possiamo continuare a semplificare la matrice con altri cambiamenti di base nel codominio. Ad esempio se sostituiamo la prima riga con la somma della prima riga con l'ultima moltiplicata per $-\frac{\beta_1}{\alpha_n}$ (cosa che può essere fatta poiché $\alpha_n \neq 0$) otterremo una matrice in cui l'elemento di posto $1, n$ (1 è la riga e n la colonna) è uguale a zero. Possiamo ripetere l'operazione per la seconda riga sommandole l'ultima moltiplicata per $-\frac{\beta_2}{\alpha_n}$. In questo modo si trova una matrice in cui gli elementi di posto k, n (k è la riga e n la colonna) con $k \neq n$ sono nulli e quello di posto n, n è α_n. Procedendo in questo modo con le altre colonne si arriva ad una matrice diagonale cioè ad una matrice del tipo:

$$\begin{pmatrix} \alpha_1 & 0 & 0 & 0 & \dots & 0 \\ 0 & \alpha_2 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & \alpha_n \end{pmatrix}.$$

Possiamo fare ancora di meglio: se moltiplichiamo la prima riga per $\frac{1}{\alpha_1}$, la seconda per $\frac{1}{\alpha_2}$ e così via, alla fine otteniamo la matrice identica I_n.

Proposizione 8.2.4 *Una matrice $A \in \mathcal{M}_n(\mathbb{R})$ è invertibile se e solo se mediante operazioni elementari sulle righe può essere ridotta alla matrice identica I_n.*

Dimostrazione. "$\Rightarrow$" Questa implicazione è stata appena vista: se una matrice è invertibile allora esiste una forma a scalini per righe uguale alla matrice identica. "$\Leftarrow$" Supponiamo che una matrice A abbia una forma a scalini per righe uguale alla matrice identica I_n. La riduzione in forma a scalini per righe è prodotta tramite operazioni elementari che non alterano il rango della matrice che quindi è uguale a n. Pertanto l'applicazione associata è invertibile e quindi la matrice è invertibile. **C.V.D.**

Osservazione 8.2.5 Se una matrice $A \in \mathcal{M}_n(\mathbb{R})$ ammette inversa, i.e., se esiste una matrice $B \in \mathcal{M}_n(\mathbb{R})$ tale che $AB = BA = I_n$, allora B è unica. Supponiamo infatti che esista un'altra matrice $C \in \mathcal{M}_n(\mathbb{R})$ tale che $CA = AC = I_n$. Consideriamo allora la matrice CAB. Essa è una matrice quadrata di ordine n e, per la proprietà associativa del prodotto righe per colonne, abbiamo $CAB = C(AB) = (CA)B$. Del resto $AB = CA = I_n$. Dunque abbiamo $CAB = CI_n = I_nB$, cioè $C = B$. Dunque l'inversa di una matrice è unica.

Per riconoscere se una matrice sia invertibile o meno non occorre ridurla alla matrice identica (bastava già una forma a scalini in cui tutti i pivots fossero gli elementi della diagonale). Tuttavia il lavoro fatto non è superfluo: ci fornisce senza ulteriore fatica un metodo per calcolare l'inversa di una matrice (invertibile)! Perché? Quello che abbiamo fatto è stato trovare un modo per ridurre la matrice di partenza A alla matrice identica. In termini molto grossolani è come se avessimo moltiplicato A^{-1} per A. Ma allora se applicassimo le stesse trasformazioni alla matrice identica I_n otterremmo il prodotto di A^{-1} per I_n cioè: $A^{-1}I_n = A^{-1}$. Quindi otterremmo la matrice inversa A^{-1}! Illustriamo quanto appena detto attraverso un esempio.

Esempio 8.2.6 Prendiamo la matrice

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

cerchiamo di vedere se è invertibile ed eventualmente calcoliamone l'inversa. Procediamo come abbiamo detto prima e consideriamo la matrice che si ottiene formalmente scrivendo le colonne della matrice I_3 a destra di A in un'unica nuova matrice:

$$\left(\begin{array}{ccc|ccc} 2 & 3 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ -1 & 0 & 1 & 0 & 0 & 1 \end{array}\right).$$

Ora applichiamo a tutta la matrice esattamente le stesse trasformazioni per riga che useremmo per ottenere una forma a scalini per righe, diagonale, della matrice A. Cominciamo: scambiamo l'ultima riga di A con la prima:

$$\left(\begin{array}{ccc|ccc} -1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 2 & 3 & 1 & 1 & 0 & 0 \end{array}\right); \tag{8.1}$$

sostituiamo alla terza riga la somma della terza riga con la prima moltiplicata per 2:

$$\left(\begin{array}{ccc|ccc} -1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 3 & 3 & 1 & 0 & 2 \end{array}\right);$$

sostituiamo alla terza riga la somma della terza riga con la seconda moltiplicata per -3:

$$\left(\begin{array}{ccc|ccc} -1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 3 & 1 & -3 & 2 \end{array}\right);$$

sostituiamo alla prima riga la somma della prima con l'ultima moltiplicata per $-\frac{1}{3}$:

$$\left(\begin{array}{ccc|ccc} -1 & 0 & 0 & -\frac{1}{3} & 1 & \frac{1}{3} \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 3 & 1 & -3 & 2 \end{array}\right);$$

moltiplichiamo la prima riga per -1 e l'ultima per $\frac{1}{3}$:

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{3} & -1 & -\frac{1}{3} \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & \frac{1}{3} & -1 & \frac{2}{3} \end{array}\right).$$

La matrice

$$\begin{pmatrix} \frac{1}{3} & -1 & -\frac{1}{3} \\ 0 & 1 & 0 \\ \frac{1}{3} & -1 & \frac{2}{3} \end{pmatrix}$$

è dunque la matrice inversa della matrice di partenza. (N.B. Consideriamo 8.1: se moltiplichiamo la matrice A per la matrice che è formata dalle ultime 3 colonne di 8.1, abbiamo

$$\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 & 3 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 1 \\ 0 & 1 & 0 \\ 2 & 3 & 1 \end{pmatrix},$$

cioè la matrice data dalle prime tre colonne di 8.1. Questo significa che la matrice data dalle ultime tre colonne di 8.1 è la matrice che moltiplicata per A dà la matrice data dalle prime tre colonne di 8.1. La stessa considerazione può essere fatta per le successive trasformazioni per riga: di fatto tutte le trasformazioni di righe e colonne possono essere realizzate tramite moltiplicazione per opportune matrici invertibili). Possiamo verificare direttamente, usando il prodotto righe per colonne, che la matrice trovata è effettivamente la matrice inversa della matrice di partenza:

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

Esempio 8.2.7 Consideriamo il seguente sistema nelle incognite x_1, x_2, x_3:

$$A\mathbf{x} = \begin{pmatrix} 2 & 3 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix}.$$

La matrice $A \in \mathcal{M}_3(\mathbb{R})$ è quella dell'esempio precedente e quindi è invertibile. Il sistema avrà dunque un'unica soluzione data dall'unico vettore di $\mathbb{R}^3$ che è spedito, dall'applicazione rappresentata da A, nel vettore $(1, -2, 0)$, o, equivalentemente, dall'unico vettore immagine di $(1, -2, 0)$ mediante l'applicazione inversa. Poiché conosciamo la matrice dell'applicazione inversa siamo in grado di risolvere il sistema:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} & -1 & -\frac{1}{3} \\ 0 & 1 & 0 \\ \frac{1}{3} & -1 & \frac{2}{3} \end{pmatrix} \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{7}{3} \\ -2 \\ \frac{7}{3} \end{pmatrix}.$$

Esercizio 8.2.8 *(***Problema A**) A questo punto siamo in grado di affrontare il problema indicato nell'introduzione come problema A. Esso si può configurare come un sistema in due incognite E_1 e E_2 dato da

$$\begin{cases} E_1 + 2E_2 = 22 \\ 3E_1 - 2E_2 = 2 \end{cases}.$$

Il rango della matrice incompleta è due e quindi uguale al rango della matrice completa. Il sistema è risolubile. Le soluzioni sono date da una soluzione particolare sommata al nucleo dell'applicazione lineare associata alla matrice del sistema incompleto. Ma tale matrice è la matrice di una applicazione lineare tra due spazi vettoriali di dimensione 2. Poiché la matrice ha rango 2 l'applicazione è suriettiva e quindi è un isomorfismo. Il nucleo è pertanto banale. Il sistema ha quindi una e una sola soluzione: quella che abbiamo trovato nella introduzione ($E_1 = 6, E_2 = 8$)!

Per risolvere la seconda parte del problema usiamo tre incognite E_1, E_2 e Z. Il sistema associato al problema sarà un sistema di due equazioni in tre incognite:

$$\begin{cases} E_1 - E_2 - Z = 2 \\ -E_1 + E_2 + 2Z = 4. \end{cases}$$

La matrice completa del sistema è data da

$$\left(\begin{array}{ccc|c} 1 & -1 & -1 & 2 \\ -1 & 1 & 2 & 4 \end{array}\right)$$

e si vede subito che il rango della matrice completa e il rango della matrice incompleta sono uguali a 2. Il sistema ammette soluzioni.

Occorre determinare il nucleo della matrice incompleta

$$\begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & 2 \end{pmatrix}$$

che è la matrice di una applicazione lineare da $\mathbb{R}^3$ in $\mathbb{R}^2$, di rango due. Per la formula delle dimensioni il nucleo di questa applicazione ha dimensione 1 ed è dato da $\langle(1, 1, 0)\rangle$.

Una soluzione particolare del sistema è fornita già nella introduzione: $E_1 = 12$, $E_2 = 4$, $Z = 6$. Le soluzioni sono allora tutte e sole della forma: $E_1 = 12 + t$, $E_2 = 4 + t$, $Z = 6$ al variare di $t \in \mathbb{R}$.

8.3 Il determinante

Intendiamo definire una funzione f avente come dominio le matrici quadrate, diciamo di ordine n, e come codominio i numeri reali, in modo che l'immagine $f(A)$ di una matrice quadrata A sia diversa da 0 se e solo se A è invertibile. Tale funzione esiste e si chiama *determinante*. Non vogliamo qui esporre la teoria completa di tale funzione, ma limitarci a dare alcune giustificazioni per la sua costruzione.

Caso n=1. L'insieme delle matrici reali 1×1 è $\mathcal{M}_1(\mathbb{R}) = \{(a) \mid a \in \mathbb{R}\}$, quindi il rango di una matrice quadrata di ordine 1, $A = (a)$, è massimo (e uguale ad 1) se e solo se $a \neq 0$. È facile costruire in questo caso l'applicazione determinante:

$$\begin{array}{rrcl} \det : & \mathcal{M}_1(\mathbb{R}) & \longrightarrow & \mathbb{R} \\ & A = (a) & \longmapsto & a. \end{array}$$

In questo modo una matrice $A = (a)$ è invertibile se e solo se $\det(A) \neq 0$.

Caso n=2. Cerchiamo ora di capire quando una matrice 2×2:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

è invertibile o, equivalentemente, quando non lo è, cioè quando non ha rango massimo. Se non ha rango massimo, allora vuol dire che le sue righe sono linearmente dipendenti. Cioè

$$\alpha(a, b) + \beta(c, d) = (0, 0)$$

per qualche coppia di numeri reali $(\alpha, \beta) \neq (0, 0)$. Sia $\alpha \neq 0$ (il caso $\beta \neq 0$ porterebbe alle stesse conclusioni). Abbiamo $(a, b) = -\frac{\beta}{\alpha}(c, d)$, cioè $a = -\frac{\beta}{\alpha}c$ e $b = -\frac{\beta}{\alpha}d$, dunque $ad = cb$ perciò:

$$ad - cb = 0.$$

Questa è la condizione affinché le righe di A siano linearmente dipendenti! D'altra parte se $ad - cb \neq 0$ allora le righe di A sono linearmente indipendenti e il rango di A è massimo. Siamo dunque in grado di costruire l'applicazione determinante per le matrici quadrate di ordine 2:

$$\begin{array}{rcl} \det : \mathcal{M}_2(\mathbb{R}) & \longrightarrow & \mathbb{R} \\ A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} & \longmapsto & \det(A) = ad - bc. \end{array}$$

Così $A \in \mathcal{M}_2(\mathbb{R})$ è invertibile se e solo se $\det A \neq 0$.

Torniamo al caso generale. Quali proprietà richiediamo alla funzione determinante?

i) Innanzitutto dovrà essere una funzione definita su $\mathcal{M}_n(\mathbb{R})$ a valori nell'insieme dei numeri reali:

$$\det \ : \ \mathcal{M}_n(\mathbb{R}) \to \mathbb{R}$$

tale che alla matrice A venga associato il numero reale $\det(A)$.

ii) La matrice identica $I_n \in \mathcal{M}_n(\mathbb{R})$ è invertibile e quindi richiederemo che abbia determinante non nullo. In particolare chiediamo che $\det I_n = 1$.

iii) Se una matrice ha due righe uguali, oppure due colonne uguali, oppure, più in generale, se il rango di una matrice $A \in \mathcal{M}_n(\mathbb{R})$ non è massimo, (cioè è minore di n), allora $\det A = 0$ (in tutti questi casi infatti la matrice A non è invertibile!). In particolare $\det(\mathbf{0}_n) = 0$ ($\mathbf{0}_n$ è la matrice quadrata di ordine n nulla).

iv) NON possiamo aspettarci che l'applicazione determinante sia lineare. In effetti, già nel caso 2×2 in generale $\det(A + B) \neq \det A + \det B$. Ad esempio, prendiamo $A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ e $B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Allora

$$A + B = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$$

e si ha: $\det(A + B) = 0$ mentre $\det(A) + \det(B) = 1 + 0 = 1$ (si noti che B non ha rango massimo).

v) La composizione di due isomorfismi è ancora un isomorfismo, cioè un'applicazione invertibile! Vorremo allora che $\det(AB) \neq 0$ se $\det A \neq 0$ e $\det B \neq 0$. Posssiamo dire anche qualcosa in più: se A rappresenta un isomorfismo di $\mathbb{R}^n$ e A^{-1} è la matrice inversa della matrice A, cioè la matrice che rappresenta l'endomorfismo inverso rispetto alle stesse basi, allora la matrice dell'endomorfismo composto è la matrice identica I_n. Vorremo pertanto che: $1 = \det I_n = \det(AA^{-1}) = \det A \ \det A^{-1}$, cioè $\det A^{-1} = \frac{1}{\det A}$. Più in generale ci aspettiamo che valga $\det(AB) = \det A \det B$.

A colpo d'occhio sembrerebbe assai complicato trovare una funzione che goda di tutte queste proprietà ma in effetti una funzione con tali caratteristiche esiste e può essere definita, a partire dagli esempi che abbiamo illustrato, procedendo per induzione sull'ordine delle matrici (nel senso che usando il determinante delle matrici 2×2 possiamo costruire il determinante delle matrici 3×3 e utilizzando il determinante delle matrici 3×3 possiamo costruire il determinante delle matrici 4×4 e così via...). Si verifichi che la funzione determinante definita sopra su $\mathcal{M}_1(\mathbb{R})$ e $\mathcal{M}_2(\mathbb{R})$ verifica le proprietà $i) - v)$.

Costruiamo ora la funzione determinante per qualunque matrice quadrata di ordine n. Abbiamo già affrontato i casi $n = 1$, $n = 2$. Ora supponiamo di essere in grado di calcolare il determinante di una matrice quadrata di ordine

$(n-1)$ e calcoliamo il determinante di una matrice quadrata A di ordine n. Osserviamo preliminarmente che cancellando la riga $i-esima$ e la colonna $j-esima$ della matrice A si ottiene una matrice quadrata di ordine $n-1$ che indichiamo con $\mathcal{A}_{ij}$. Per ipotesi (induttiva) conosciamo allora $\det(\mathcal{A}_{ij})$, per ogni $i, j = 1, \ldots, n$.

Per ogni elemento $a_{k1}, a_{k2}, \ldots, a_{kn}$ della riga $k-esima$ della matrice $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{R})$ possiamo considerare la matrice quadrata di ordine $n-1$ ottenuta da A cancellando la riga e la colonna in cui quell'elemento si trova, nell'ordine: $\mathcal{A}_{k1}, \mathcal{A}_{k2}, \ldots, \mathcal{A}_{kn}$. Diamo allora la seguente definizione:

Definizione 8.3.1 *Data la matrice $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{R})$, poniamo*

$$\det(A) = a_{11} \qquad se \quad n = 1$$

e

$$\begin{aligned} \det A = \quad & \det(a_{ij}) = (-1)^{k+1} a_{k1} \det(\mathcal{A}_{k1}) + (-1)^{k+2} a_{k2} \det(\mathcal{A}_{k2}) + \\ + \quad & (-1)^{k+3} a_{k3} \det(\mathcal{A}_{k3}) + \cdots + (-1)^{k+n} a_{kn} \det(\mathcal{A}_{kn}) \end{aligned} \tag{8.2}$$

se $n > 1$, con $1 \leq k \leq n$.

Attenzione: l'espressione (8.2) si chiama sviluppo del determinante secondo la riga k-esima. Ogni riga può essere scelta per il calcolo del determinante: il risultato della formula (8.2) sarà sempre lo stesso! Vale ancora di più: se avessimo scelto una qualsiasi colonna avremmo potuto costruire una formula analoga e con lo stesso risultato! Vale cioè anche la seguente formula:

$$\begin{aligned} \det A = \det(a_{ij}) = \quad & (-1)^{k+1} a_{1k} \det(\mathcal{A}_{1k}) + (-1)^{k+2} a_{2k} \det(\mathcal{A}_{2k}) + \\ + \quad & (-1)^{k+3} a_{3k} \det(\mathcal{A}_{3k}) + \cdots + (-1)^{k+n} a_{nk} \det(\mathcal{A}_{nk}) \end{aligned}$$

(sviluppo del determinante di A rispetto alla colonna $k-esima$: gli elementi della colonna k-esima sono infatti $a_{1k}, a_{2k}, a_{3k}, \ldots, a_{nk}$).

Si osservi che la regola precedente può essere utilizzata per calcolare il determinante delle matrici 2×2. Il risultato sarà esattamente quello già descritto!

È sorprendente che una definizione così complicata possa implicare tutte le proprietà da noi richieste, ma in effetti questo è ciò che succede. La dimostrazione di questo fatto deriva da una teoria più avanzata di cui non ci occuperemo.

Esempi 8.3.2 *a*) Sappiamo già calcolare il determinante delle matrici di ordine 2. Utilizziamo la nostra formula nel caso di una matrice $B = (b_{ij}) \in \mathcal{M}_3(\mathbb{R})$. Sia ad esempio

$$B = \begin{pmatrix} 2 & 0 & -1 \\ 0 & 3 & 2 \\ 1 & 1 & 1 \end{pmatrix}.$$

Sviluppiamo il determinante rispetto alla seconda riga ($b_{21} = 0, b_{22} = 3, b_{23} = 2$):

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

$$+(-1)^{2+3}(2)\det\begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix} = 3(2-(-1)) - 2(2) = 5.$$

Potremmo adesso sviluppare il determinante di B secondo la terza colonna ($b_{13} = -1, b_{23} = 2, b_{33} = 1$), ottenendo:

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

$$+(-1)^{3+3}(1)\det\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} = (-1)(-3) + (-2)(2) + (6) = 3 - 4 + 6 = 5.$$

Come ci aspettavamo abbiamo trovato in entrambi i casi lo stesso risultato: i due modi di procedere sono equivalenti. Nello stesso modo avremmo potuto scegliere qualsiasi altra riga o colonna di B e avremmo trovato $\det(B) = 5$.

b) Come si è visto, se $a_{ij} = 0$, nello sviluppo del determinante rispetto alla i-esima riga o alla j-esima colonna l'addendo $(-1)^{i+j}a_{ij}\mathcal{A}_{ij}$ non dà alcun contributo. Pertanto, essendo liberi di scegliere la riga o la colonna rispetto a cui sviluppare, sceglieremo, se possibile, una riga o colonna con molte entrate nulle!

c) Se $A \in \mathcal{M}_n(\mathbb{R})$ è del tipo:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & \dots & a_{1n} \\ 0 & & & & \\ 0 & & & & \\ \vdots & & \mathcal{A}_{11} & & \\ 0 & & & & \\ 0 & & & & \end{pmatrix}$$

cioè $a_{k1} = 0$ per ogni $k \neq 1$, allora $\det A = a_{11} \det \mathcal{A}_{11}$ (basta sviluppare secondo gli elementi della prima colonna). Analogamente, se

$$A = \begin{pmatrix} a_{11} & 0 & \dots & \dots & 0 \\ a_{21} & & & & \\ a_{31} & & & & \\ \vdots & & \mathcal{A}_{11} & & \\ a_{n-11} & & & & \\ a_{n1} & & & & \end{pmatrix},$$

$\det A = a_{11} \det \mathcal{A}_{11}$ (sviluppo secondo la prima riga).

Osservazione 8.3.3 Sappiamo che una matrice quadrata triangolare superiore con tutti gli elementi sulla diagonale diversi da zero ha rango massimo quindi è invertibile e ha determinante diverso da zero. Più precisamente, il determinante è il prodotto degli elementi sulla diagonale. Infatti se la matrice $A \in \mathcal{M}_n(\mathbb{R})$ è del tipo:

$$A = \begin{pmatrix} a_1 & a_{12} & \dots & & & a_{1n} \\ 0 & a_2 & a_{23} & \dots & & \vdots \\ 0 & 0 & a_3 & \dots & & \\ \vdots & \vdots & 0 & \ddots & & \\ 0 & \vdots & & & \ddots & \vdots \\ 0 & 0 & \dots & & & a_n \end{pmatrix}$$

allora possiamo sviluppare il determinante rispetto alla prima colonna:

$$\det A = a_1 \det \begin{pmatrix} a_2 & a_{12} & \dots & \\ 0 & a_3 & \dots & \\ 0 & 0 & \ddots & \\ 0 & & & \\ 0 & & & a_n \end{pmatrix}$$

e poi procedere nello stesso modo. Alla fine si ottiene che il determinante è il prodotto degli elementi sulla diagonale di A.

8.4 Nota sulle trasformazioni elementari e il teorema di Binet

(di Fabrizio Caselli)

8.4.1 Le trasformazioni e le matrici elementari

Limitiamo in questa sezione la nostra attenzione alle matrici quadrate $n \times n$. Le trasformazioni elementari permettono di ridurre una qualunque matrice ad una matrice a scala. Esistono tre tipi di trasformazioni elementari che ora andiamo a richiamare.

- Scambiare due righe;
- Moltiplicare una riga per uno scalare non nullo;
- Aggiungere ad una riga un multiplo scalare di un'altra riga.

La matrice associata ad una trasformazione elementare è la matrice ottenuta applicando la trasformazione stessa alla matrice identità.

Proposizione 8.4.1 *Sia A una matrice quadrata $n \times n$, B la matrice ottenuta da A applicando una trasformazione elementare ed R la matrice associata alla trasformazione elementare stessa. Allora*

$$B = RA.$$

Dimostrazione. Questa dimostrazione è una semplice verifica. Per ragioni tipografiche la illustriamo nel caso in cui le trasformazioni elementari coivolgano le prime righe della matrice A. Se la trasformazione elementare consiste nello scambio delle prime due righe si ha

$$R = \begin{pmatrix} 0 & 1 & & & \\ 1 & 0 & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix}.$$

e

$$RA = \begin{pmatrix} 0 & 1 & & & \\ 1 & 0 & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix} \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ a_{31} & \cdots & a_{3n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = \begin{pmatrix} a_{21} & \cdots & a_{2n} \\ a_{11} & \cdots & a_{1n} \\ a_{31} & \cdots & a_{3n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}.$$

Se la trasformazione consiste nel moltiplicare la prima riga di A per c, con $c \neq 0$ si ha

$$R = \begin{pmatrix} c & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}$$

e quindi

$$RA = \begin{pmatrix} c & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix} \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = \begin{pmatrix} ca_{11} & \cdots & ca_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}.$$

Infine, se la trasformazione consiste nell'aggiungere alla prima riga la seconda moltiplicata per c abbiamo

$$R = \begin{pmatrix} 1 & c & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}$$

e quindi

$$RA = \begin{pmatrix} 1 & c & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix} \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = \begin{pmatrix} a_{11} + ca_{21} & \cdots & a_{1n} + ca_{2n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}.$$

Ricordiamo ora la definizione di determinante.

Definizione 8.4.2 (Laplace) *Sia $i = 1, \ldots, n$ e A una matrice quadrata $n \times n$. Il determinante di A è definito ricorsivamente da (sviluppo rispetto alla i-esima riga)*

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{i,j} \det A_{i,j},$$

dove $A_{i,j}$ è la matrice ottenuta da A concellando la riga i e la colonna j. Analogamente si ha (sviluppo rispetto alla i-esima colonna)

$$\det A = \sum_{j=1}^{n} (-1)^{j+i} a_{j,i} \det A_{j,i}.$$

Abbiamo già visto che il determinante di una matrice triangolare (superiore o inferiore) è dato dal prodotto dei coefficienti sulla diagonale. Vogliamo ora calcolare il determinante di una matrice elementare e capire come cambia il determinante di una matrice A se applichiamo ad essa una trasformazione elementare.

Lemma 8.4.3 *Sia T una trasformazione elementare e R la matrice elementare associata. Allora*

$$\det R = \begin{cases} -1 & \textit{se } T \textit{ è lo scambio di due righe} \\ c & \textit{se } T \textit{ è la moltiplicazione di una riga per uno scalare non nullo } c \\ 1 & \textit{se } T \textit{ è la somma di un multiplo scalare di una riga ad un'altra} \end{cases}$$

Dimostrazione. La dimostrazione è un semplice esercizio. Nel caso in cui T non è lo scambio di due righe allora R è triangolare e il calcolo del determinante è immediato. Mostriamo il caso dello scambio di due righe nel caso speciale in cui le righe coinvolte siano le prime due. Si ha

$$R = \begin{pmatrix} 0 & 1 & & & \\ 1 & 0 & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix}.$$

da cui, sviluppando il determinante rispetto alla prima riga otteniamo

$$\det R = (-1)^3 \det I_{n-1} = -1.$$

Proposizione 8.4.4 *Siano A una matrice quadrata e R una matrice elementare. Allora*

$$\det RA = \det R \det A.$$

Dimostrazione. Mostriamo ancora questa dimostrazione nel caso in cui siano coinvolte le prime righe nella trasformazione elementare in oggetto. Chiamiamo B la matrice ottenuta applicando la trasformazione elementare ad A.

Se la trasformazione elementare è lo scambio delle prime due righe abbiamo

$$B = \begin{pmatrix} a_{21} & \cdots & a_{2n} \\ a_{11} & \cdots & a_{1n} \\ a_{31} & \cdots & a_{3n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}.$$

Sviluppando il determinante di B rispetto alla seconda riga abbiamo

$$\det B = \sum_{j=1}^{n} (-1)^{2+j} b_{2j} \det B_{2j}.$$

Osservando ora che $B_{2j} = A_{1j}$ (cancellado la seconda riga a B si ottiene la stessa matrice che si otterrebbe cancellando la prima riga di A) e che $b_{2j} = a_{1j}$ abbiamo

$$\det B = \sum_{j=1}^{n} (-1)^{2+j} a_{1j} \det A_{1j} = (-1) \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \det A_{1j} = -\det A$$

avendo riconosciuto nell'ultima uguaglianza lo sviluppo del determinante di A rispetto alla prima riga. Come conseguenza di questo fatto abbiamo che se una matrice A ha due righe uguali, diciamo la riga i e la riga j, allora $\det A = 0$. Infatti, se R è la matrice elementare associata allo scambio delle righe i e j abbiamo $RA = A$. Ma per quanto dimostrato abbiamo anche

$$\det RA = -\det A$$

da cui $\det A = -\det A$ e quindi $\det A = 0$.

Se la trasformazione elementare è la moltiplicazione della prima riga per uno scalare non nullo c abbiamo

$$B = \begin{pmatrix} ca_{11} & \cdots & ca_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

e calcoliamo $\det B$ sviluppando rispetto alla prima riga

$$\det B = \sum_{j=1}^{n} (-1)^{1+j} b_{1j} \det B_{1j} = \sum_{j=1}^{n} (-1)^{1+j} c a_{1j} \det A_{1j}$$

dove abbiamo osservato che $B_{1j} = A_{1j}$ (cancellando la prima riga ad A o a B si ottiene la stessa matrice) e che $b_{1j} = ca_{1j}$. Concludiamo che

$$\det B = c \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \det A_{1j} = c \det A = \det R \det A.$$

Infine, se B è ottenuta da A aggiungendo alla prima riga la seconda moltiplicata per c abbiamo

$$B = \begin{pmatrix} a_{11} + ca_{21} & \cdots & a_{1n} + ca_{2n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

e calcolando $\det B$ sviluppando rispetto alla prima riga otteniamo

$$\det B = \sum_{j=1}^{n} (-1)^{1+j} (a_{1j} + ca_{2j}) \det A_{1j} =$$

$$= \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \det A_{1j} + c \sum_{j=1}^{n} (-1)^{1+j} a_{2j} \det A_{1j} = \det A + 0$$

dove nell'ultima uguaglianza abbiamo osservato che $\sum_{j=1}^{n} (-1)^{1+j} a_{2j} \det A_{1j}$ altri non è che il determinante della matrice

$$\begin{pmatrix} a_{21} & \cdots & a_{2n} \\ a_{21} & \cdots & a_{2n} \\ a_{31} & \cdots & a_{3n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

sviluppato secondo la prima riga. Avendo tale matrice due righe uguali il suo determinante è nullo. Una conseguenza immediata di questa proposizione è il seguente

Corollario 8.4.5 *Una matrice quadrata A è invertibile se e solo se il suo* $\det A \neq 0$.

Dimostrazione. Dalla proposizione abbiamo che una matrice ha determinante non nullo se e solo una sua riduzione a scala ha determinante non nullo. Ma una matrice quadrata a scala ha determinante non nullo se e solo se tutti i coefficienti sulla diagonale sono non nulli e questo è equivalente a richiedere che abbia rango massimo e che quindi sia invertibile.

Teorema 8.4.6 (Binet) *Siano A e B due matrici quadrate $n \times n$. Allora*

$$\det AB = \det A \det B.$$

Dimostrazione. Se A è invertibile sappiamo che $A = R_1 \cdots R_k$ dove $R_1, \ldots, R_k$ sono matrici elementari. Abbiamo quindi, utilizzando la proposizione precedente che

$$\det(AB) = \det(R_1 \cdots R_k B) = \det R_1 \det(R_2 \cdots R_k B) =$$

$$\det R_1 \det R_2 \det(R_3 \cdots R_k B) = \ldots = \det R_1 \det R_2 \cdots \det R_k \det B.$$

Applicando ancora la proposizione precedente abbiamo ora

$$\det R_1 \det R_2 \cdots \det R_k = \det R_1 \cdots \det R_{k-2} \det(R_{k-1} R_k) =$$

$$\det R_1 \cdots \det R_{k-3} \det(R_{k-2} R_{k-1} R_k) = \ldots = \det(R_1 \cdots R_k) = \det A$$

e il teorema è quindi dimostrato in questo caso.

Se invece A non è invertibile sappiamo che $\det A = 0$ e dobbiamo quindi mostrare che anche $\det AB = 0$, cioè che AB non è invertibile. Basterà quindi mostrare che esiste un vettore "colonna" non nullo $v \in \mathbb{R}^n$ tale che $ABv = 0_{\mathbb{R}_n}$. Se anche B non è invertibile basterà scegliere $v \neq 0_{\mathbb{R}_n}$ tale che $Bv = 0_{\mathbb{R}_n}$. Se invece B è invertibile, sia $w \neq 0_{\mathbb{R}_n}$ tale che $Aw = 0_{\mathbb{R}_n}$. Siccome B è invertibile sappiamo che esiste $v \in \mathbb{R}^n$ tale che $Bv = w$. Ma allora $ABv = Aw = 0_{\mathbb{R}_n}$.

8.5 Esercizi svolti

Esercizio 8.5.1 Calcolare i prodotti AB e BA delle matrici

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & -3 & 2 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & 1 \\ -2 & 1 \\ -1 & 0 \\ 3 & 5 \end{pmatrix}.$$

Indicato con $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ l'endomorfismo di $\mathbb{R}^2$ la cui matrice rispetto alla base canonica di $\mathbb{R}^2$ sia la matrice AB, determinare $f(2, -3)$. La matrice BA è la matrice di un endomorfismo?

Svolgimento. Per calcolare AB e BA dobbiamo semplicemente usare la definizione di prodotto righe per colonne:

$$AB = \begin{pmatrix} 5 & 23 \\ 7 & 11 \end{pmatrix}; \qquad BA = \begin{pmatrix} 0 & 1 & -3 & 2 \\ -2 & -3 & -9 & -6 \\ -1 & -2 & -3 & -4 \\ 3 & 11 & -6 & 22 \end{pmatrix}.$$

Dunque, AB è una matrice quadrata di ordine 2 e quindi la matrice associata ad un endomorfismo di $\mathbb{R}^2$. BA è una matrice quadrata di ordine 4 ed è pertanto la matrice associata ad un endomorfismo di $\mathbb{R}^4$. Dire che AB è la matrice associata ad un endomorfismo f rispetto alla base canonica significa che le colonne di f sono le coordinate rispetto alla base canonica e_1, e_2 di $\mathbb{R}^2$ dei vettori $f(e_1)$ e $f(e_2)$. Così $f(e_1) = 5e_1 + 7e_2 = (5, 7)$ e $f(e_2) = 23e_1 + 11e_2 = (23, 11)$. Usiamo la linearità di f per calcolare l'immagine del vettore $(2, -3)$:

$f(2, -3) = f(2e_1 - 3e_2) = 2f(e_1) - 3f(e_2) = 2(5, 7) - 3(23, 11) = (-59, -19)$.

Esercizio 8.5.2 Stabilire se i vettori $v_1 = (1, 3, 4)$, $v_2 = (4, 3, -2)$, $v_3 = (2, 3, 0)$ di $\mathbb{R}^3$ sono linearmente indipendenti.

Svolgimento. Abbiamo già risolto questo tipo di problema utilizzando la definizione di vettori linearmente indipendenti. Vogliamo proporre una soluzione diversa e decisamente più rapida che fa uso della nozione di determinante. Costruiamo la matrice che abbia come vettori riga i vettori v_1, v_2, v_3:

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

A questo punto dire che i vettori v_1, v_2, v_3 sono linearmente indipendenti equivale a dire che le righe della matrice A sono linearmente indipendenti cioè che la matrice A è invertibile. Del resto la matrice A è invertibile se e solo se il suo determinante è non nullo. Dunque i vettori v_1, v_2, v_3 sono linearmente indipendenti se e solo se $\det(A) \neq 0$. Non ci resta che calcolare $\det(A)$: sviluppando il determinante di A secondo gli elementi della terza colonna otteniamo

$$\det(A) = 4(12 - 6) + 2(3 - 6) = 24 - 6 = 18.$$

Possiamo concludere che i vettori v_1, v_2 e v_3 sono linearmente indipendenti.

Esercizio 8.5.3 In $\mathbb{R}^3$ siano dati i vettori $v_1 = (2, t, 1)$, $v_2 = (-1, 1, 0)$, $v_3 = (1, 1, t)$ dove t è un parametro reale. Si consideri l'endomorfismo $f_t : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ definito da: $f_t(e_1) = v_1$, $f_t(e_2) = v_2$, $f_t(e_3) = v_3$, dove $\{e_1, e_2, e_3\}$ è la base canonica di $\mathbb{R}^3$. Esistono valori del parametro t per i quali l'applicazione f_t è invertibile?

Svolgimento. Costruiamo la matrice associata all'applicazione lineare f_t rispetto alla base canonica di $\mathbb{R}^3$:

$$F_t = \begin{pmatrix} 2 & -1 & 1 \\ t & 1 & 1 \\ 1 & 0 & t \end{pmatrix}.$$

Sviluppiamo il determinante di F_t rispetto alla seconda colonna:

$\det(F_t) = 1(t^2 - 1) + (2t - 1) = t^2 + 2t - 2$. Allora per ogni $t \in \mathbb{R}$ tale che $t^2 + 2t - 2 \neq 0$, vale a dire per ogni $t \neq -1 \pm \sqrt{3}$, $\det(F_t) \neq 0$ e l'applicazione f_t è invertibile. Viceversa, se $t = -1 + \sqrt{3}$ oppure $t = -1 - \sqrt{3}$ l'applicazione f_t non è invertibile.

Esercizio 8.5.4 Determinare la matrice inversa di ciascuna delle seguenti matrici:

$$M_1 = \begin{pmatrix} 1 & 1 \\ -2 & 1 \end{pmatrix}, \qquad M_2 = \begin{pmatrix} -1 & 0 & 1 \\ 0 & -1 & 4 \\ 0 & 0 & 2 \end{pmatrix}, \qquad M_3 = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -3. \end{pmatrix}$$

Svolgimento. Procediamo come illustrato nell'esempio 8.2.6. Cominciamo dalla matrice M_1 e consideriamo la matrice $\left(\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ -2 & 1 & 0 & 1 \end{array}\right)$.

Sostituiamo alla seconda riga la somma della prima riga moltiplicata per due e della seconda: $\left(\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 0 & 3 & 2 & 1 \end{array}\right)$.

Ora sostituiamo alla prima riga la somma della prima riga moltiplicata per -3 e della seconda: $\left(\begin{array}{cc|cc} -3 & 0 & -1 & 1 \\ 0 & 3 & 2 & 1 \end{array}\right)$.

Infine, moltiplichiamo la prima riga per $-\frac{1}{3}$ e la seconda per $\frac{1}{3}$: $\left(\begin{array}{cc|cc} 1 & 0 & \frac{1}{3} & -\frac{1}{3} \\ 0 & 1 & \frac{2}{3} & \frac{1}{3} \end{array}\right)$.

Abbiamo così ottenuto la matrice inversa della matrice M_1:

$$M_1 = \begin{pmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{pmatrix}.$$

Analogamente si procede per la matrice M_2:
$\left(\begin{array}{ccc|ccc} -1 & 0 & 1 & 1 & 0 & 0 \\ 0 & -1 & 4 & 0 & 1 & 0 \\ 0 & 0 & 2 & 0 & 0 & 1 \end{array}\right) \Rightarrow \left(\begin{array}{ccc|ccc} 2 & 0 & 0 & -2 & 0 & 1 \\ 0 & -1 & 0 & 0 & 1 & -2 \\ 0 & 0 & 2 & 0 & 0 & 1 \end{array}\right) \Rightarrow$
$\Rightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 0 & \frac{1}{2} \\ 0 & 1 & 0 & 0 & -1 & 2 \\ 0 & 0 & 1 & 0 & 0 & \frac{1}{2} \end{array}\right)$.

Dunque $M_2^{-1} = \begin{pmatrix} -1 & 0 & \frac{1}{2} \\ 0 & -1 & 2 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}$. Osserviamo che M_2 è una matrice triangolare superiore e che la sua inversa è anch'essa una matrice triangolare superiore. Questo fatto è vero in generale.

Anche nel caso della matrice M_3 si procede come per qualsiasi altra matrice, ma, essendo essa una matrice diagonale, il numero di passaggi sarà inferiore: la forma diagonale è infatti già di per sé una forma a scalini per righe. Otteniamo dunque, immediatamente,

$$M_3^{-1} = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{3} \end{pmatrix}.$$

In generale, data una matrice diagonale di ordine n con elementi diagonali non nulli α_1, α_2, $\ldots, \alpha_n$ (se uno di questi fosse nullo la matrice non sarebbe invertibile!), la sua inversa è la matrice diagonale con elementi diagonali α_1^{-1}, α_2^{-1}, $\ldots, \alpha_n^{-1}$.

Esercizio 8.5.5 Sia $h : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ l'applicazione lineare associata, rispetto alla base canonica di $\mathbb{R}^3$, alla matrice:

$$H = \begin{pmatrix} 1 & 2 & 0 \\ -1 & 1 & -2 \\ 0 & -1 & -3 \end{pmatrix}.$$

Stabilire se h è ivertibile. e determinare $h^{-1}(2,1,2)$.

Svolgimento. Si verifica facilmente che il rango della matrice H è uguale a 3, pertanto la funzione h è invertibile. A questo punto un modo di risolvere l'esercizio è calcolare la matrice inversa della matrice H. Infatti se h è un endomorfismo invertibile di uno spazio vettoriale V e H è la matrice ad esso associata rispetto ad una base $\mathcal{V}$ del dominio e ad una base $\mathcal{W}$ del codominio, allora la matrice associata all'endomorfismo inverso h^{-1}, rispetto alla base $\mathcal{W}$ del dominio e alla base $\mathcal{V}$ del codominio, è la matrice H^{-1}.

Procediamo col calcolo della matrice H^{-1}:

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ -1 & 1 & -2 & 0 & 1 & 0 \\ 0 & -1 & -3 & 0 & 0 & 1 \end{array}\right) \Rightarrow \left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 3 & -2 & 1 & 1 & 0 \\ 0 & -1 & -3 & 0 & 0 & 1 \end{array}\right) \Rightarrow$$
$$\left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 3 & -2 & 1 & 1 & 0 \\ 0 & 0 & -11 & 1 & 1 & 3 \end{array}\right) \Rightarrow \left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & -33 & 0 & -9 & -9 & 6 \\ 0 & 0 & -11 & 1 & 1 & 3 \end{array}\right) \Rightarrow$$
$$\left(\begin{array}{ccc|ccc} \frac{33}{2} & 0 & 0 & \frac{15}{2} & -9 & 6 \\ 0 & -33 & 0 & -9 & -9 & 6 \\ 0 & 0 & -11 & 1 & 1 & 3 \end{array}\right) \Rightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{5}{11} & -\frac{6}{11} & \frac{4}{11} \\ 0 & 1 & 0 & \frac{3}{11} & \frac{3}{11} & -\frac{2}{11} \\ 0 & 0 & 1 & -\frac{1}{11} & -\frac{1}{11} & -\frac{3}{11} \end{array}\right).$$

Abbiamo quindi

$$H^{-1} = \begin{pmatrix} \frac{5}{11} & -\frac{6}{11} & \frac{4}{11} \\ \frac{3}{11} & \frac{3}{11} & -\frac{2}{11} \\ -\frac{1}{11} & -\frac{1}{11} & -\frac{3}{11} \end{pmatrix}.$$

Pertanto $h^{-1}(2,1,2) = \begin{pmatrix} \frac{5}{11} & -\frac{6}{11} & \frac{4}{11} \\ \frac{3}{11} & \frac{3}{11} & -\frac{2}{11} \\ -\frac{1}{11} & -\frac{1}{11} & -\frac{3}{11} \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} \frac{12}{11} \\ \frac{5}{11} \\ -\frac{9}{11} \end{pmatrix}.$

8.6 Esercizi proposti

Esercizio 8.6.1 Calcolare l'inversa della matrice

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

Esercizio 8.6.2 Calcolare $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^n$.

Esercizio 8.6.3 Sia $A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$. Costruire, se possibile, una matrice $B \neq I_2$ tale che $(A+B)(A-B) = A^2 - B^2$.

Esercizio 8.6.4 Una matrice quadrata A si dice *nilpotente* se $A^k = 0$ per qualche intero $k > 0$. Mostrare che se A è nilpotente allora $I + A$ è invertibile.

Esercizio 8.6.5 Data la matrice $A = \begin{pmatrix} 2 & 3 \\ 1 & 2 \\ 2 & 5 \end{pmatrix}$,

1. trovare, se possibile, una matrice B tale che $BA = I_2$. Una siffatta matrice B è unica?
2. Trovare, se possibile, una matrice C tale che $AC = I_3$.

Esercizio 8.6.6 Calcolare il determinante della seguente matrice:

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

Esercizio 8.6.7 Si consideri l'applicazione lineare $D : \mathbb{R}^3[x] \longrightarrow \mathbb{R}^3[x]$ che ad ogni polinomio a coefficienti reali di grado minore o uguale a 3 nella variabile x associa la sua derivata rispetto ad x. Scrivere la matrice associata a D rispetto alla base $\mathcal{C} = \{1, x, x^2, x^3\}$ nel dominio e alla base $\mathcal{B} = \{x^3, x^3 + x^2, x^3 + x^2 + x, x^3 + x^2 + x + 1\}$ nel codominio. Calcolare il determinante della matrice ottenuta.

Lezione 9

Cambiamenti di base

Proponiamo due semplici problemi per giustificare l'argomento di questo paragrafo:

Problema 1. In $\mathbb{R}^2$ consideriamo due basi: $\{v_1 = (1,2),\ v_2 = (0,-1)\}$ e $\{w_1 = (1,1),\ w_2 = (3,1)\}$. Poiché sono delle basi, ogni vettore di $\mathbb{R}^2$ si scrive in modo unico come combinazione lineare dei vettori di ciascuno dei due insiemi. Così, se prendiamo, ad esempio, $(-1,2) \in \mathbb{R}^2$, abbiamo l'unica scrittura $(-1,2) = (-1)(1,2) + (-4)(0,-1)$, quindi possiamo univocamente individuare $(-1,2)$ nella base v_1, v_2 mediante le coordinate $(-1,-4)_{\mathbf{v}}$. Analogamente, sempre lo stesso vettore $(-1,2)$ è espresso nella base w_1, w_2 dalle coordinate $(\frac{7}{2}, -\frac{3}{2})_{\mathbf{w}}$. Come possiamo passare da una scrittura all'altra? È possibile costruire una macchina che ci consenta di ottenere le coordinate di un vettore rispetto ad una data base conoscendo le sue coordinate rispetto ad una base diversa?

Ricordiamo che, fissata la base canonica $\mathcal{C}$: $\{e_1 = (1,0,\dots,0),\ e_2 = (0,1,0,\dots,0),\ \dots, e_n = (0,0,\dots,0,1)\}$ di $\mathbb{R}^n$, per ogni vettore v di $\mathbb{R}^n$ le coordinate di v rispetto alla base canonica coincidono con le entrate di v.

Problema 2. Dobbiamo descrivere una applicazione lineare f tra due spazi vettoriali V e W. Come abbiamo fatto finora? Abbiamo fissato una base del dominio ed una del codominio e abbiamo costruito una matrice tale che, scritto un vettore v di V nelle coordinate della base scelta del dominio, allora il prodotto della matrice per la matrice colonna data dalle coordinate di v, fornisce le coordinate del vettore immagine, $f(v)$, nella base scelta del codominio. Se ora cambiassimo base sia nel dominio che nel codominio, cosa succederebbe alla matrice associata all'applicazione f?

In questo paragrafo vogliamo dare una risposta ai quesiti or ora illustrati.

Sia V uno spazio vettoriale di dimensione n e consideriamo l'endomorfismo di V dato dalla applicazione lineare identica, cioè dall'applicazione lineare che ad ogni vettore $v \in V$ associa il vettore v stesso:

$$id_V : V \longrightarrow V$$

$$id_V(v) = v \qquad v \in V.$$

Fissiamo ora una base $\mathbf{v} = \{v_1, v_2, \dots, v_n\}$ nel dominio e la stessa nel codominio. Possiamo allora associare all'endomorfismo id_V una matrice: la prima colonna di tale matrice è data dalle coordinate dell'immagine del primo vettore della base scelta nel dominio rispetto alla base fissata nel codominio: $id_V(v_1) = v_1 = 1v_1 + 0v_2 + 0v_3 + \dots + 0v_n$. Dunque $id_V(v_1)$ ha coordinate $(1, 0, 0, \dots, 0)_{\mathbf{v}}$. Analogamente, $id_V(v_2) = v_2 = 0v_1 + 1v_2 + 0v_3 + \dots + 0v_n$, e dunque le coordinate di $id_V(v_2)$ sono $(0, 1, 0, \dots, 0)_{\mathbf{v}}$. In conclusione, se fissiamo la stessa base nel dominio e nel codominio, la matrice dell'endomorfismo identico è la matrice identica $I_n \in \mathcal{M}_n(\mathbb{R})$.

Supponiamo ora di voler cambiare la base nel codominio, cioè la base del dominio sarà sempre $\mathbf{v}$, mentre nel codominio prenderemo un'altra base: $\mathbf{w} = \{w_1, w_2, \dots, w_n\}$. Come sarà la matrice associata all'applicazione id_V rispetto alle nuove basi?

Per fare un esempio prendiamo $n = 3$ e supponiamo che i vettori della base $\mathbf{w}$ si scrivano, rispetto ai vettori della base $\mathbf{v}$, come segue: $w_1 = 2v_1 - v_3, w_2 = v_3, w_3 = v_1 - v_2$. Una cosa balza subito all'occhio: saremmo già in grado di scrivere la matrice associata all'applicazione identica rispetto alla base w_1, w_2, w_3 nel dominio e alla base v_1, v_2, v_3 nel codominio. In effetti la prima colonna di tale matrice è data dalle coordinate di $id_V(w_1) = w_1$ rispetto a v_1, v_2, v_3: $w_1 = 2v_1 - v_3$, cioè $id_V(w_1) = (2, 0, -1)_{\mathbf{v}}$. Lo stesso ragionamento vale per w_2 e w_3: la matrice della applicazione identica rispetto alla base $\mathbf{w}$ nel dominio e $\mathbf{v}$ nel codominio risulta allora essere

$$M_{\mathbf{v}}^{\mathbf{w}}(id_V) = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}.$$

Attenzione, però: il nostro problema di partenza non era questo! Tuttavia la matrice appena trovata (con poca fatica) ci sarà di aiuto. Vediamo perché. Osserviamo, intanto che $id_V \circ id_V = id_V$ e consideriamo il seguente diagramma:

$$(V, \mathbf{v}) \xrightarrow{id} (V, \mathbf{w}) \xrightarrow{id} (V, \mathbf{v})$$

che vogliamo tradurre in termini matriciali. Dovremmo quindi determinare $C = BA$, con $A, B, C \in \mathcal{M}_3(\mathbb{R})$, dove A è la matrice della applicazione identica rispetto alla base $\mathbf{v}$ nel dominio e $\mathbf{w}$ nel codominio, B è la matrice associata all'applicazione identica rispetto alla base $\mathbf{w}$ nel dominio e $\mathbf{v}$ nel codominio e quindi C è la matrice dell'applicazione identica rispetto alla base $\mathbf{v}$ nel dominio e nel codominio. Ora, per quanto osservato sopra, $C = I_3$, e $B = M_{\mathbf{v}}^{\mathbf{w}}(id_V)$. Pertanto la matrice A, cioè la matrice che stavamo cercando, è l'inversa di $B = M_{\mathbf{v}}^{\mathbf{w}}(id_V)$:

$$I_3 = BA,$$

quindi siamo in grado di calcolarla! Ma che cosa rappresenta la matrice A? È la matrice che trasforma un vettore scritto in coordinate rispetto alla base $\mathbf{v}$, nello stesso vettore scritto in coordinate rispetto alla base $\mathbf{w}$:

$$A = M_{\mathbf{w}}^{\mathbf{v}}(id_V) = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 1 \\ 0 & -1 & 0 \end{pmatrix}.$$

Per questo la matrice A si chiama **matrice del cambiamento di base dalla base $\mathbf{v}$ alla base $\mathbf{w}$**.

Se a questo punto vogliamo conoscere le coordinate rispetto alla base $\mathbf{w}$ del vettore $v_1 + 2v_2 - 3v_3 = (1, 2, -3)_{\mathbf{v}}$ basta applicare la matrice del cambiamento di base (dalla base $\mathbf{v}$ alla base $\mathbf{w}$) al vettore $(1, 2, -3)$:

$$\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 1 \\ 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix} = \begin{pmatrix} \frac{3}{2} \\ -\frac{3}{2} \\ -2 \end{pmatrix},$$

cioè $(1, 2, -3)_{\mathbf{v}} = (\frac{3}{2}, -\frac{3}{2}, -2)_{\mathbf{w}}$.

Conclusione del Problema 1. Possiamo dedurre un metodo per risolvere il problema 1. Abbiamo uno spazio vettoriale e due sue basi: $\mathbf{v}$ e $\mathbf{w}$. La matrice di passaggio dalla base $\mathbf{w}$ alla base $\mathbf{v}$, cioè la matrice $A \in \mathcal{M}_n(\mathbb{R})$ che ci permette di determinare le coordinate nella base $\mathbf{v}$ di un vettore espresso in coordinate rispetto alla base $\mathbf{w}$, è data dalla matrice nelle cui colonne si trovano, nell'ordine, le coordinate dei vettori $w_1, \ldots, w_n$ rispetto alla base $\mathbf{v}$. D'altro canto, se vogliamo la matrice del cambiamento di base dalla base $\mathbf{v}$ alla base $\mathbf{w}$, allora prenderemo A^{-1}.

Se abbiamo, ad esempio, in $\mathbb{R}^2$ la base $\mathbf{v}$ data da $v_1 = (1,1), v_2 = (1,-1)$ e la base $\mathbf{w}$ espressa nella base $\mathbf{v}$ come $w_1 = v_1 - 2v_2, w_2 = -v_1 + 3v_2$, allora la matrice di passaggio dalla base $\mathbf{w}$ alla base $\mathbf{v}$ è

$$A = \begin{pmatrix} 1 & -1 \\ -2 & 3 \end{pmatrix},$$

mentre la sua inversa è la matrice di passaggio dalla base $\mathbf{v}$ alla base $\mathbf{w}$:

$$A^{-1} = \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix}.$$

Nel Problema 1 la situazione è leggermente diversa poiché i vettori delle due basi date non sono espressi gli uni in coordinate rispetto agli altri. Si tratta di fare un passaggio in più e di determinare le coordinate dei vettori w_i nella base $\mathbf{v}$: $w_1 = (1,1) = \alpha_1 v_1 + \alpha_2 v_2 = \alpha_1(1,2) + \alpha_2(0,-1) = (\alpha_1, 2\alpha_1 - \alpha_2)$ e quindi $\alpha_1 = 1$ e $\alpha_2 = 1$; $w_2 = (3,1) = \beta_1 v_1 + \beta_2 v_2 = \beta_1(1,2) + \beta_2(0,-1) = (\beta_1, 2\beta_1 - \beta_2)$ cioè $\beta_1 = 3$ e $\beta_2 = 5$. Quindi la matrice di passaggio dalla base $\mathbf{w}$ alla base $\mathbf{v}$ è

$$\begin{pmatrix} 1 & 3 \\ 1 & 5 \end{pmatrix}.$$

Non resta che prendere la sua inversa per calcolare la matrice di passaggio dalla base $\mathbf{v}$ alla base $\mathbf{w}$:

$$\begin{pmatrix} \frac{5}{2} & -\frac{3}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Possiamo procedere anche in un modo diverso. In effetti sia i vettori della base $\mathbf{v}$ che i vettori della base $\mathbf{w}$ nel Problema 1 sono espressi in coordinate rispetto alla base canonica. Quindi i dati dell'esempio ci forniscono immediatamente la matrice di passaggio dalla base $\mathbf{v}$ alla base canonica $\mathbf{e}$: $T = \begin{pmatrix} 1 & 0 \\ 2 & -1 \end{pmatrix}$ e la matrice di passaggio dalla base $\mathbf{w}$ alla base $\mathbf{e}$: $S = \begin{pmatrix} 1 & 3 \\ 1 & 1 \end{pmatrix}$. In termini di composizione possiamo allora scrivere la matrice di passaggio dalla base $\mathbf{v}$ alla base $\mathbf{w}$ come il prodotto $S^{-1}T$. Questo significa che stiamo componendo l' applicazione identica con se stessa, una volta descrivendola mediante T rispetto alla base $\mathbf{v}$ nel dominio ed $\mathbf{e}$ nel codominio, una volta mediante S^{-1} rispetto alla base $\mathbf{e}$ nel dominio e $\mathbf{w}$ nel codominio: il prodotto $S^{-1}T$ è dunque la matrice della applicazione identica rispetto alla base $\mathbf{v}$ nel dominio e $\mathbf{w}$ nel codominio, cioè quello che cercavamo:

$$S^{-1} = \begin{pmatrix} -\frac{1}{2} & \frac{3}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}; \qquad S^{-1}T = \begin{pmatrix} \frac{5}{2} & -\frac{3}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Passiamo ora al Problema 2. Sia $f : V \to W$ una funzione lineare e siano $\dim V = m$ e $\dim W = n$. Sia $F \in \mathcal{M}_{n,m}(\mathbb{R})$ la matrice associata a f rispetto ad una base fissata $\mathbf{v}$ di V e ad una base fissata $\mathbf{w}$ di W. Sia $H \in \mathcal{M}_m(\mathbb{R})$ la matrice del cambiamento di base dalla base $\mathbf{v}'$ alla base $\mathbf{v}$ di V e sia $K \in \mathcal{M}_n(\mathbb{R})$ la matrice del cambiamento di base dalla base $\mathbf{w}$ alla base $\mathbf{w}'$. Cerchiamo la matrice $F' \in \mathcal{M}_{n,m}(\mathbb{R})$ della applicazione f rispetto alla base $\mathbf{v}'$ nel dominio e alla base $\mathbf{w}'$ nel codominio. Moltiplicando F' per il vettore colonna delle coordinate di un vettore v di V nella base $\mathbf{v}'$ otterremo le coordinate di $f(v)$ nella base $\mathbf{w}'$ di W.

Partiamo dunque da un vettore di V espresso nella base $\mathbf{v}'$, tramite la matrice H lo trasformiamo nello stesso vettore espresso però in coordinate nella base $\mathbf{v}$, adesso possiamo applicare F che spedisce il vettore trovato nel vettore immagine tramite f, espresso in coordinate rispetto alla base $\mathbf{w}$ del codominio, a questo vettore applichiamo la matrice K che ci fornisce le sue coordinate nella base $\mathbf{w}'$ e a questo punto abbiamo finito. In termini di prodotto di matrici si ha:

$$F' = KFH.$$

Conclusione del Problema 2. Se la matrice $F \in \mathcal{M}_{n,m}(\mathbb{R})$ rappresenta una applicazione lineare $f \in \mathcal{L}in(V, W)$ tra gli spazi vettoriali V e W, di dimensione rispettivamente m e n, rispetto alle basi $\mathbf{v}$ e $\mathbf{w}$ e se F' è la matrice della stessa applicazione rispetto alle basi $\mathbf{v}'$ nel dominio e $\mathbf{w}'$ nel codominio, allora si ha:

$$F' = KFH$$

dove $H \in \mathcal{M}_m(\mathbb{R})$ è la matrice invertibile che rappresenta il cambiamento di base dalla base $\mathbf{v}'$ alla base $\mathbf{v}$, mentre $K \in \mathcal{M}_n(\mathbb{R})$ è la matrice invertibile che rappresenta il cambiamento di base dalla base $\mathbf{w}$ alla base $\mathbf{w}'$.

Esempio. Sia $\varphi : \mathbb{R}^3 \to \mathbb{R}^2$ l'applicazione lineare che rispetto alle basi canoniche in $\mathbb{R}^3$ e $\mathbb{R}^2$ è data dalla matrice

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

(Denoteremo con $\mathbf{e}$ sia la base canonica di $\mathbb{R}^3$ che la base canonica di $\mathbb{R}^2$). In particolare l'immagine mediante φ del vettore $(\alpha, \beta, \gamma) \in \mathbb{R}^3$, che nella base canonica ha coordinate (α, β, γ), è il vettore

$$\begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 2 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} \alpha - \beta \\ -\alpha + \beta + 2\gamma \end{pmatrix}.$$

Vogliamo ora scrivere la matrice associata a φ rispetto alla base di $\mathbb{R}^3$ data da $v_1 = (2, 1, 0), v_2 = (1, 1, 1), v_3 = (1, 0, 1)$ e alla base di $\mathbb{R}^2$ data da $w_1 = (1, 2), w_2 = (2, 1)$. La matrice del cambiamento di base dalla base $\mathbf{v}$ alla base canonica di $\mathbb{R}^3$ è la matrice

$$H = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

e la matrice del cambiamento di base da $\mathbf{w}$ ad $\mathbf{e}$ è

$$K^{-1} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

La matrice inversa è la matrice di passaggio dalla base canonica di $\mathbb{R}^2$ alla base $\mathbf{w}$:

$$K = \begin{pmatrix} -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} \end{pmatrix}.$$

Così la matrice che cerchiamo è

$$F' = KFH = \begin{pmatrix} -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 2 \end{pmatrix} \begin{pmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} -1 & \frac{4}{3} & \frac{1}{3} \\ 1 & -\frac{2}{3} & \frac{1}{3} \end{pmatrix}.$$

Esempio 9.0.8 Sia V uno spazio vettoriale di dimensione n e consideriamo lo spazio vettoriale $\mathcal{E}nd(V)$ degli endomorfismi di V. Supponiamo che un endomorfismo φ ci venga dato nella base $v_1, v_2, \ldots, v_n$ dalla matrice $A \in \mathcal{M}_n(\mathbb{R})$ cioè scegliamo sia nel dominio che nel codominio la stessa base $\mathbf{v}$. Supponiamo di voler determinare la matrice della stessa applicazione rispetto ad una nuova base $w_1, w_2, \ldots, w_n$ di V (sia nel dominio che nel codominio). Se $H \in \mathcal{M}_n(\mathbb{R})$ è la matrice (invertibile) del cambiamento di base dalla base $\mathbf{w}$ alla base $\mathbf{v}$, allora la matrice che cerchiamo è $H^{-1}AH$.

Esempio 9.0.9 Data una base $\mathbf{v} = \{v_1, v_2, \ldots, v_n\}$ di uno spazio vettoriale, qualsiasi matrice invertibile $H \in \mathcal{M}_n(\mathbb{R})$ rappresenta la matrice di un cambiamento di base. Ma tra quali basi? E in che ordine? In effetti se consideriamo la matrice come la matrice di passaggio da una base $\mathbf{w}$ (che non conosciamo) alla base $\mathbf{v}$ allora le colonne di H rappresentano nell'ordine le coordinate nella base $\mathbf{v}$ dei vettori della base $\mathbf{w}$: $w_1, w_2, \ldots, w_n$.

D'altro canto possiamo pure interpretare H come la matrice del cambiamento di base dalla base $\mathbf{v}$ ad una base $\mathbf{w}$ che non conosciamo. Allora per determinare $\mathbf{w}$ basterà prendere la matrice inversa H^{-1} che rappresenterà la matrice del cambiamento di base dalla base $\mathbf{w}$ alla base conosciuta $\mathbf{v}$ e allora le sue colonne saranno le coordinate rispetto a $\mathbf{v}$ dei vettori (nell'ordine) della base $\mathbf{w}$.

In conclusione, fissata una base $\mathbf{v} = \{v_1, v_2, \ldots, v_n\}$ di uno spazio vettoriale V, possiamo identificare le matrici invertibili di $\mathcal{M}_n(\mathbb{R})$ e l'insieme delle basi di V: ogni matrice invertibile H può essere vista come la matrice di un cambiamento di base dalla base $\mathbf{w}$ alla base $\mathbf{v}$, nel senso che le colonne di H sono nell'ordine le coordinate dei vettori $w_1, w_2, \ldots, w_n$.

Osservazione 9.0.10 Abbiamo già osservato che le trasformazioni per riga (o per colonna) di una matrice sono equivalenti a dei cambiamenti di base. Un cambiamento di base è un isomorfismo e il suo determinante è diverso da zero. Dunque le trasformazioni per riga o per colonna sono trasformazioni invertibili. Presa una matrice quadrata A ed indicata con A' la matrice che si ottiene dalla matrice A mediante una trasformazione per riga (o colonna), si ha: $det(A') = \alpha det(A)$ dove $\alpha \neq 0$ è il determinante della matrice del cambiamento di base corrispondente alla trasformazione per riga (o per colonna) effettuata. Questo altera il determinante, ma non il fatto che sia o meno uguale a zero!

Definizione 9.0.11 *Due matrici $A, B \in \mathcal{M}_n(\mathbf{R})$ si dicono* **simili** *se rappresentano lo stesso endomorfismo rispetto a basi diverse, cioè, per quanto visto, se esiste una matrice invertibile $H \in \mathcal{M}_n(\mathbb{R})$ tale che $A = H^{-1}BH$.*

Nella definizione precedente le matrici A e B rappresentano lo stesso endomorfismo rispetto a basi diverse (la matrice $H^{-1}BH$ è legata alla matrice B tramite un cambiamento di base dato dalla matrice H). Indichiamo l'insieme delle matrici invertibili di ordine n con $GL_n(\mathbb{R})$.

Osservazione 9.0.12 Segue immediatamente dalla definizione che matrici simili hanno lo stesso determinante. Infatti se A e B sono matrici quadrate di ordine n e sono simili, esiste una matrice invertibile $H \in \mathcal{M}_n(\mathbb{R})$ tale che $A = H^{-1}BH$. Di conseguenza $det(A) = det(H^{-1}BH) = det(H^{-1})det(B)det(H) = \frac{1}{det(H)}det(B)det(H) = det(B)$.

Osservazione 9.0.13 Il calcolo del determinante e la sua applicazione al calcolo del rango dei minori quadrati di una matrice forniscono un ulteriore metodo per il calcolo del rango di una matrice e quindi per la soluzione di un sistema lineare. Consideriamo il seguente esempio: in $\mathbf{R}^4$ si calcoli, al variare di $\alpha \in \mathbf{R}$, lo spazio delle soluzioni Σ_α del sistema lineare S_α nelle incognite x, y, z, w:

$$S_\alpha = \begin{cases} 2x + 3y - 4z - w = 0 \\ x + 2y - 3z = 0 \\ \alpha x + z - 2w = 0 \end{cases}.$$

S_α è un sistema omogeneo di 3 equazioni in 4 incognite. Scriviamo la matrice (completa=incompleta in questo caso) ad esso associata:

$$A_\alpha = \begin{pmatrix} 2 & 3 & -4 & -1 \\ 1 & 2 & -3 & 0 \\ \alpha & 0 & 1 & -2 \end{pmatrix}$$

e consideriamo il minore individuato dalle prime tre colonne

$$\begin{pmatrix} 2 & 3 & -4 \\ 1 & 2 & -3 \\ \alpha & 0 & 1 \end{pmatrix}.$$

Il suo determinante è uguale a $-\alpha + 1$ quindi per ogni $\alpha \neq 1$ esso è diverso da zero! In questo caso la matrice A_α ha rango massimo (uguale a 3), quindi rappresenta una applicazione lineare di rango 3 di uno spazio di dimensione 4 in uno spazio di dimensione 3. Il sistema ammette soluzioni e lo spazio delle soluzioni è il nucleo dell'applicazione lineare individuata dalla matrice A_α e ha pertanto dimensione 1. Per determinare le soluzioni del sistema riduciamo la matrice A_α in forma a scalini per righe. Attraverso un certo numero di passaggi otteniamo la matrice:

$$\begin{pmatrix} 1 & 2 & -3 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & 1-\alpha & 2\alpha - 2 \end{pmatrix}$$

e, dunque, il sistema

$$\begin{cases} x + 2y - 3z = 0 \\ -y + 2z - w = 0 \\ (1-\alpha)z + (2\alpha - 2)w = 0. \end{cases}$$

Otteniamo $\Sigma_\alpha =< (0, 3, 2, 1) >$.

Dobbiamo ora considerare il caso $\alpha = 1$. Attenzione: non stiamo dicendo che la matrice ha rango minore di 3 in questo caso! Vi sono altri minori quadrati di ordine 3 di cui non abbiamo calcolato il determinante. Può ancora succedere che la matrice abbia rango 3: per avere rango minore di 3, tutti i minori di ordine 3 debbono avere determinante uguale a 0. Il calcolo delle soluzioni del sistema S_1 si può fare direttamente per sostituzione. Abbiamo il sistema:

$$S_1 = \begin{cases} 2x + 3y - 4z - w = 0 \\ x + 2y - 3z = 0 \\ x + z - 2w = 0. \end{cases}$$

Calcoliamo allora il rango del sistema col metodo della riduzione per righe:

$$A_1 = \begin{pmatrix} 2 & 3 & -4 & -1 \\ 1 & 2 & -3 & 0 \\ 1 & 0 & 1 & -2 \end{pmatrix} \Rightarrow \begin{pmatrix} 2 & 3 & -4 & -1 \\ 1 & 2 & -3 & 0 \\ 0 & -2 & 4 & -2 \end{pmatrix} \Rightarrow \begin{pmatrix} 0 & -1 & 2 & -1 \\ 1 & 2 & -3 & 0 \\ 0 & -2 & 4 & -2 \end{pmatrix}$$
$$\Rightarrow \begin{pmatrix} 0 & -1 & 2 & -1 \\ 1 & 2 & -3 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

La matrice A_1 ha dunque rango 2. Lo spazio delle sue soluzioni è uno spazio vettoriale di dimensione 2: le variabili libere sono z e w, dunque scegliendo $z = 1, w = 0$ si trova la soluzione $x = -1, y = 2, z = 1, w = 0$, mentre scegliendo $z = 0, w = 1$, si ha $x = 2, y = -1, z = 0, w = 1$. L'insieme delle soluzioni è: $\Sigma_1 =< (-1, 2, 1, 0), (2, -1, 0, 1) >$.

9.1 Esercizi svolti

Esercizio 9.1.1 Sia $id : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ l'applicazione identica. Si consideri la base $\mathcal{B}$ di $\mathbb{R}^3$: $\mathcal{B} = \{w_1 = (1, 1, 0), w_2 = (1, -1, 0), w_3 = (0, 1, 1)\}$. Si scriva la matrice associata all'applicazione id fissando:

1) la base $\mathcal{B}$ nel dominio e la base canonica nel codominio;

2) la base canonica nel dominio e la base $\mathcal{B}$ nel codominio;

3) la base $\mathcal{B}$ nel dominio e nel codominio.

Svolgimento. L'applicazione id manda ogni vettore di $\mathbb{R}^3$ in se stesso: $id(v) = v$ per ogni $v \in \mathbb{R}^3$. Nel caso 1) abbiamo dunque $id(w_i) = w_i$ per ogni $i = 1, 2, 3$ e, poiché le coordinate di un vettore rispetto alla base canonica non sono che le sue componenti, la matrice associata all'applicazione id rispetto alla base $\mathcal{B}$ nel dominio e alla base canonica nel codominio è:

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

Nel caso 2) abbiamo $id(e_i) = e_i$ per ogni vettore e_i della base canonica e dobbiamo esprimere i vettori e_i in coordinate rispetto alla base $\mathcal{B}$. Potremmo procedere direttamente determinando, per ogni $i = 1, 2, 3$, i numeri reali α, β, γ tali che $e_1 = \alpha w_1 + \beta w_2 + \gamma w_3$ e così per e_2 ed e_3. Tuttavia (astuti!) osserviamo che la matrice che ha sulle colonne le componenti dei vettori e_i rispetto ai vettori w_j è la matrice del cambiamento di base dalla base canonica alla base $\mathcal{B}$ pertanto essa è la matrice inversa della matrice determinata in 1):

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 \end{pmatrix}.$$

Nel caso 3) abbiamo, banalmente, $id(w_1) = 1w_1 + 0w_2 + 0w_3$, $id(w_2) = 0w_1 + 1w_2 + 0w_3$, $id(w_3) = 0w_1 + 0w_2 + 1w_3$. Pertanto la matrice richiesta è la matrice I_3.

Esercizio 9.1.2 Sia $\mathcal{B} = \{v_1 = (1, 1), v_2 = (1, 0)\}$ una base di $\mathbb{R}^2$ e sia $\mathcal{C}$ la base canonica di $\mathbb{R}^3$. Sia $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ l'applicazione lineare associata, rispetto alla base $\mathcal{B}$ del dominio e alla base $\mathcal{C}$ del codominio, alla matrice

$$F = \begin{pmatrix} 1 & 4 \\ 3 & 1 \\ 2 & 0 \end{pmatrix}.$$

Determinare la matrice associata all'applicazione lineare f rispetto alla base canonica $\mathcal{C}'$ di $\mathbb{R}^2$ e alla base $\mathcal{B}' = \{w_1 = (1, 1, 1), w_2 = (1, 1, 0), w_3 = (1, 0, 0)\}$ di $\mathbb{R}^3$.

Svolgimento. L'esercizio consiste nel tradurre in termini matriciali il seguente diagramma:

$$(\mathbb{R}^2, \mathcal{C}') \xrightarrow{id} (\mathbb{R}^2, \mathcal{B}) \xrightarrow{f} (\mathbb{R}^3, \mathcal{C}) \xrightarrow{id} (\mathbb{R}^3, \mathcal{B}').$$

La composizione di funzioni illustrata nel diagramma produce infatti il seguente effetto: si parte da un vettore di $\mathbb{R}^2$ espresso in coordinate rispetto alla base canonica di $\mathbb{R}^2$, lo si esprime in coordinate rispetto alla base $\mathcal{B}$, si applica ad esso f e si ottengono le coordinate del vettore immagine nella base canonica di $\mathbb{R}^3$; infine si determinano le coordinate del vettore immagine nella base $\mathcal{B}'$. La matrice della applicazione composta $id \circ f \circ id$ è la matrice richiesta dall'esercizio.

Indichiamo con $M_{\mathcal{C}'}^{\mathcal{B}}$ la matrice del cambiamento di base dalla base $\mathcal{C}'$ alla base $\mathcal{B}$, cioè la matrice associata all'applicazione identica di $\mathbb{R}^2$ rispetto alla base canonica del dominio e alla base $\mathcal{B}$ del codominio, e con $M_{\mathcal{C}}^{\mathcal{B}'}$ la matrice del cambiamento di base dalla base canonica alla base $\mathcal{B}'$ di $\mathbb{R}^3$. Allora $M_{\mathcal{B}}^{\mathcal{C}'} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix}$. Analogamente $M_{\mathcal{B}'}^{\mathcal{C}} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{pmatrix}$.

Così la matrice richiesta è:

$$M_{\mathcal{B}'}^{\mathcal{C}} F M_{\mathcal{B}}^{\mathcal{C}'} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 4 \\ 3 & 1 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ 1 & 0 \\ 3 & -5 \end{pmatrix}.$$

Esercizio 9.1.3 Stabilire se i vettori $v_1 = (1, 3, 4)$, $v_2 = (4, 3, -2)$, $v_3 = (2, 3, 0)$ di $\mathbb{R}^3$ sono linearmente indipendenti.

Svolgimento. Abbiamo già risolto questo tipo di problema utilizzando la definizione di insieme di vettori linearmente indipendente. Vogliamo proporre una soluzione diversa e decisamente più rapida che fa uso della nozione di determinante. Costruiamo la matrice che abbia come vettori riga i vettori v_1, v_2, v_3:

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

A questo punto dire che i vettori v_1, v_2, v_3 sono linearmente indipendenti equivale a dire che le righe della matrice A sono linearmente indipendenti cioè che la matrice A è invertibile. Del resto la matrice A è invertibile se e solo se il suo determinante è non nullo. Dunque i vettori v_1, v_2, v_3 sono linearmente indipendenti se e solo se $det(A) \neq 0$. Non ci resta che calcolare $det(A)$: sviluppando il determinante di A secondo gli elementi della terza colonna otteniamo

$$det(A) = 4(12 - 6) + 2(3 - 6) = 24 - 6 = 18.$$

Possiamo concludere che i vettori v_1, v_2 e v_3 sono linearmente indipendenti.

Esercizio 9.1.4 Calcolare il rango della matrice $M = \begin{pmatrix} 1 & 0 & -1 & 4 \\ -1 & 1 & 0 & 2 \\ -1 & 0 & 1 & 3 \end{pmatrix}$.

Svolgimento. Non vogliamo utilizzare il metodo della riduzione di una matrice in forma a scalini per righe ma la caratterizzazione del rango di una matrice come massimo ordine di minori invertibili della matrice. Il rango della matrice M è minore o uguale a 3. Sarà uguale a 3 se riusciremo a trovare un minore di ordine 3 di M con determinante non nullo.

Cominciamo col minore che si ottiene da M eliminando la quarta colonna: $\begin{pmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$. Esso ha determinante nullo e quindi non fa al caso nostro.

Consideriamo allora il minore di M che si ottiene eliminando la prima colonna: esso ha determinante uguale a 7. Dunque abbiamo individuato un minore invertibile di M di ordine 3. M ha pertanto rango 3.

Esercizio 9.1.5 In $\mathbb{R}^3$ siano dati i vettori $v_1 = (2, t, 1)$, $v_2 = (-1, 1, 0)$, $v_3 = (1, 1, t)$ dove t è un parametro reale. Si consideri l'endomorfismo $f_t : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ definito da: $f_t(e_1) = v_1$, $f_t(e_2) = v_2$, $f_t(e_3) = v_3$, dove $\{e_1, e_2, e_3\}$ è la base canonica di $\mathbb{R}^3$. Esistono valori del parametro t per i quali l'applicazione f_t è invertibile? In caso affermativo, per uno di questi valori si determini la matrice che rappresenta l'applicazione inversa di f_t rispetto alla base canonica.

Svolgimento. Costruiamo la matrice associata all'applicazione lineare f_t rispetto alla base canonica di $\mathbb{R}^3$:

$$F_t = \begin{pmatrix} 2 & -1 & 1 \\ t & 1 & 1 \\ 1 & 0 & t \end{pmatrix}.$$

Sviluppiamo il determinante di F_t rispetto alla seconda colonna:

$det(F_t) = 1(t^2-1)+(2t-1) = t^2+2t-2$. Allora per ogni $t \in \mathbb{R}$ tale che $t^2+2t-2 \neq 0$, vale a dire per ogni $t \neq -1 \pm \sqrt{3}$, $det(F_t) \neq 0$ e l'applicazione f_t è invertibile. Viceversa, se $t = -1+\sqrt{3}$ oppure $t = -1-\sqrt{3}$ l'applicazione f_t non è invertibile.

Poniamo ora $t = 1$ e consideriamo la matrice F_1 associata a f_1 rispetto alla base canonica. Si ha: $det(F_1) = 1$ e

$$F_1^{-1} = \begin{pmatrix} 2 & -1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 1 & -2 \\ 0 & 1 & -1 \\ -1 & -1 & 3 \end{pmatrix}.$$

Lezione 10

Matrici diagonalizzabili

Tratteremo ora di endomorfismi di uno spazio vettoriale V di dimensione n. Sappiamo che, scelta una base di V, ad ogni endomorfismo resta associata una matrice $A \in \mathcal{M}_n(\mathbb{R})$. D'altro canto sappiamo anche che un cambiamento di base trasforma la matrice dell'endomorfismo in una matrice simile $B \in \mathcal{M}_n(\mathbb{R})$ per cui $B = H^{-1}AH$ con $H \in GL_n(\mathbb{R})$ (9.3.4). Il problema su cui ci soffermeremo sarà in qualche senso a double-face: trovare una base rispetto alla quale la matrice del nostro endomorfismo abbia la forma più semplice possibile e trovare un modo per decidere se due matrici sono o non sono simili (se, cioè, rappresentano o meno lo stesso endomorfismo). La linea di risoluzione del problema è la seguente: date due matrici determinare per ciascuna di esse una forma per così dire canonica, a meno di similitudine, e dal confronto tra tali forme dedurre se le due matrici siano simili o no. Le risposte che daremo sono parziali ma sufficienti per attaccare uno dei problemi descritti nella introduzione. Per una risposta completa bisognerebbe ricorrere alla teoria di Jordan.

Sottolineiamo il fatto che qualche volta useremo in maniera intercambiabile matrici ed endomorfismi.

10.1 Autovalori e autovettori

Consideriamo φ endomorfismo di uno spazio vettoriale V.

Definizione 10.1.1 *Un numero reale λ si dice un autovalore dell'endomorfismo φ se esiste un vettore $v \neq \mathbf{0}_V$ tale che $\varphi(v) = \lambda v$. Il vettore v si dice allora autovettore di φ di autovalore λ.*

Si noti che abbiamo richiesto che un autovettore sia diverso dal vettore nullo: in effetti se accettassimo anche $\mathbf{0}_V$ come possibile autovettore allora si avrebbe $\mathbf{0}_V = \varphi(\mathbf{0}_V) = \alpha\mathbf{0}_V$ per ogni reale α. Dunque ogni numero reale sarebbe un autovalore e la definizione non avrebbe tanto senso. Si osservi inoltre che un vettore non nullo di V è un autovettore di φ di autovalore zero se e solo se appartiene al nucleo di φ.

Sia ora $\lambda \in \mathbb{R}$ un autovalore dell'endomorfismo φ di V. Consideriamo l'insieme V_λ^* degli autovettori di autovalore λ (non è vuoto...):

$$V_\lambda^* = \{v \in V \mid v \neq \mathbf{0}_V \, , \, \varphi(v) = \lambda v\}.$$

Aggiungiamo a tale insieme (che non può essere uno spazio vettoriale perché non contiene il vettore nullo) il vettore nullo: $V_\lambda = V_\lambda^* \cup \{\mathbf{0}_V\} = \{v \in V \mid \varphi(v) = \lambda v\}$. Allora

Proposizione 10.1.2 *L'insieme V_λ è un sottospazio vettoriale di V.*

Dimostrazione. Siano v_1 e v_2 due elementi di V_λ (i.e. $\varphi(v_1) = \lambda v_1$, $\varphi(v_2) = \lambda v_2$), dobbiamo vedere che pure $v_1 + v_2$ appartiene a V_λ. Calcoliamo quindi $\varphi(v_1 + v_2)$ che, per la linearità di φ, è uguale a $\varphi(v_1) + \varphi(v_2) = \lambda v_1 + \lambda v_2 = \lambda(v_1 + v_2)$. Dunque $v_1 + v_2$ è un autovettore di autovalore λ oppure il vettore nullo, in ogni caso $v_1 + v_2 \in V_\lambda$. Siano ora $\alpha \in \mathbb{R}$ e $v \in V_\lambda$, allora $\varphi(\alpha v) = \alpha\varphi(v) = \alpha(\lambda v)$, essendo φ lineare e $v \in V_\lambda$; del resto $\alpha(\lambda v) = \lambda(\alpha v)$ per le proprietà del prodotto per scalari, così $\alpha v \in V_\lambda$. V_λ è dunque un sottospazio vettoriale di V. **C. V. D.**

Definizione 10.1.3 *Dati φ endomorfismo di uno spazio vettoriale V ed un suo autovalore λ, il sottospazio vettoriale V_λ è detto l'autospazio di φ associato all'autovalore λ.*

Per ora tutto è formalmente chiaro, ma finché non sappiamo calcolare gli autovalori di un endomorfismo non possiamo fare nulla di concreto. Innanzitutto osserviamo che, scelta una base $\mathbf{v} = \{v_1, \dots, v_n\}$ di V ($\dim V = n$) e indicata con $A \in \mathcal{M}_n(\mathbb{R})$ la matrice associata a φ rispetto a questa base sia nel dominio che nel codominio di φ, dire che un vettore t è un autovettore di autovalore λ, significa che per la $n-upla$ delle sue coordinate $(t_1, \dots, t_n)$ nella base $\mathbf{v}$ si ha

$$\lambda \begin{pmatrix} t_1 \\ t_2 \\ \dots \\ t_n \end{pmatrix} = \begin{pmatrix} \lambda t_1 \\ \lambda t_2 \\ \dots \\ \lambda t_n \end{pmatrix} = A \begin{pmatrix} t_1 \\ t_2 \\ \dots \\ t_n \end{pmatrix}.$$

Questa condizione è indipendente dalla base scelta nel senso che, se $H \in GL_n(\mathbb{R})$ è la matrice del cambiamento di base dalla base $\mathbf{v}$ ad una base $\mathbf{w}$, allora le coordinate $H\begin{pmatrix} t_1 \\ t_2 \\ \dots \\ t_n \end{pmatrix}$ sono le coordinate di t nella base $\mathbf{w}$ e soddisfano l'analoga relazione dove al posto di A troveremo la matrice B associata all'endomorfismo φ rispetto alla base $\mathbf{w}$. Cerchiamo di essere più precisi: per costruzione B è simile ad A attraverso la relazione esplicita: $B = HAH^{-1}$. Ne consegue che

$$B(H\begin{pmatrix} t_1 \\ t_2 \\ \dots \\ t_n \end{pmatrix}) = BH\begin{pmatrix} t_1 \\ t_2 \\ \dots \\ t_n \end{pmatrix} = (HAH^{-1}H)\begin{pmatrix} t_1 \\ t_2 \\ \dots \\ t_n \end{pmatrix} = HA\begin{pmatrix} t_1 \\ t_2 \\ \dots \\ t_n \end{pmatrix} = H\begin{pmatrix} \lambda t_1 \\ \lambda t_2 \\ \dots \\ \lambda t_n \end{pmatrix} =$$

$= \lambda(H\begin{pmatrix} t_1 \\ t_2 \\ \dots \\ t_n \end{pmatrix})$ cioè $H\begin{pmatrix} t_1 \\ t_2 \\ \dots \\ t_n \end{pmatrix}$ sono le coordinate nella base $\mathbf{w}$ di un autovettore di φ relativo all'autovalore λ.

Da queste semplici considerazioni deduciamo che la matrice associata ad un endomorfismo rispetto ad una base fissata è strettamente legata all'esistenza di autovettori e autovalori. Il prossimo risultato illustra la natura di questo legame.

Sia φ un endomorfismo di uno spazio vettoriale V e sia $\mathbf{v}$ una base di V, $\mathbf{v} = \{v_1, v_2, \dots, v_n\}$. Allora, se $A \in \mathcal{M}_n(\mathbb{R})$ è la matrice associata a φ rispetto alla base $\mathbf{v}$ (sia nel dominio che nel codominio di φ), possiamo considerare la matrice $A - tI_n$ (cioè la matrice che si ottiene sottraendo agli elementi diagonali di A l'indeterminata t). Il determinante di questa matrice è un polinomio nella variabile t:

$$P(t) = det(A - tI_n)$$

detto **polinomio caratteristico di** φ. In effetti tale polinomio sembrerebbe dipendere dalla matrice associata all'endomorfismo (e quindi dalla base scelta).

Osservazione 10.1.4 *Il polinomio caratteristico di φ non dipende dalla base scelta per costruire la matrice A.*

Dimostrazione. Per mostrare che il polinomio caratteristico di φ è indipendente dalla base scelta basta dimostrare che due matrici simili hanno lo stesso polinomio caratteristico. Fissata, infatti, una base di V diversa dalla precedente, la matrice associata a φ rispetto alla nuova base è una matrice simile alla matrice A. Siano dunque A, B due matrici simili: $B = H^{-1}AH$, con $H \in GL_n(\mathbb{R})$ matrice invertibile. Il polinomio caratteristico della matrice B è: $det(B - tI_n) = det(H^{-1}AH - tI_n)$. Osserviamo che tI_n è una matrice i cui soli elementi non nulli si trovano sulla diagonale e sono tutti uguali a t. Si vede facilmente che per ogni matrice quadrata T di ordine n, $(tI_n)T = t(I_nT) = tT = T(tI_n)$. Si ha pertanto: $det(B - tI_n) = det(H^{-1}AH - H^{-1}(tI_n)H) = det(H^{-1}(A - tI_n)H) = detH^{-1}det(A - tI_n)detH = det(A - tI_n)$. **C.V.D.** Per analogia parleremo pure di polinomio caratteristico associato ad una matrice $A \in \mathcal{M}_n(\mathbb{R})$,

$$det(A - tI_n) = P_A(t) :$$

tale polinomio è lo stesso per ogni matrice simile ad A, i.e., per ogni matrice $B \in \mathcal{M}_n(\mathbb{R})$ tale che esiste $H \in Gl_n(\mathbb{R})$, per cui $H^{-1}AH = B$.

Proposizione 10.1.5 *Un numero reale λ è un autovalore dell'endomorfismo φ se e solo se è radice del suo polinomio caratteristico.*

Dimostrazione. Dire che $\alpha \in \mathbb{R}$ è un autovalore di φ equivale a dire, in coordinate rispetto ad una base $\mathbf{v}$, che esiste un vettore $(t_1, \ldots, t_n) \neq \mathbf{0}_{\mathbb{R}^n}$ tale che

$$A \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{pmatrix} = \alpha \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{pmatrix}$$

cioè

$$A \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{pmatrix} = \alpha I_n \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{pmatrix} \iff (A - \alpha I_n) \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

i.e., $\mathbf{0}_{\mathbb{R}^n} \neq (t_1, \ldots, t_n) \in Ker(A - \alpha I_n)$. Osserviamo che la matrice quadrata $(A - \alpha I_n)$ ha nucleo non banale se e solo se il suo rango non è massimo (se tale rango fosse massimo $Ker(A - \alpha I_n)$ conterrebbe solo il vettore nullo). In altre parole α è autovalore di φ se e solo se $A - \alpha I_n$ non è invertibile, se e solo se, quindi, il suo determinante è uguale a zero: $det(A - \alpha I_n) = 0$. Ma tale determinante è proprio il polinomio caratteristico di A, $P_A(t)$, calcolato in α: $P_A(\alpha)$. Cosicché α è un autovalore di φ se e solo se $P_A(\alpha) = 0$. **C.V.D.**

Definizione 10.1.6 *Sia λ un autovalore di un endomorfismo φ di uno spazio vettoriale V. Si chiama molteplicità algebrica di λ la sua molteplicità come radice del polinomio caratteristico. Si chiama molteplicità geometrica di λ la dimensione dell'autospazio relativo a λ.*

Osservazione 10.1.7 i) Esistono naturalmente matrici che non hanno autovalori reali! Consideriamo, ad esempio, la matrice $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Il suo polinomio caratteristico è il polinomio

$$det(A - \lambda I_2) = det \begin{pmatrix} -\lambda & 1 \\ -1 & -\lambda \end{pmatrix} = \lambda^2 + 1$$

che non ha radici reali.

ii) Il polinomio caratteristico di un endomorfismo φ di uno spazio vettoriale di dimensione n ha grado n, vale a dire, indicata con A la matrice associata a φ rispetto ad una base fissata, il polinomio $det(A - \lambda I_n)$ è un polinomio di grado n nella variabile λ. Questo significa che il grado massimo con cui compare la variabile λ nello sviluppo del determinante della matrice $A - \lambda I_n$ è n. Per renderci conto di questo calcoliamo $det(A - \lambda I_n)$ sviluppando il determinante rispetto alla prima riga: il termine di grado massimo (in λ) si ottiene moltiplicando il maggior numero possibile di termini contenenti λ. Dal momento che λ compare, con grado 1 e coefficiente -1, solo sulla diagonale della matrice e che gli elementi diagonali sono n, si deduce facilmente che il grado massimo con cui λ compare nella espressione del determinante è n e che il coefficiente di λ^n è esattamente $(-1)^n$.

Possiamo dire qualcosa sul termine noto del polinomio caratteristico? In questo caso dobbiamo considerare soltanto i termini che non contengono λ: basterebbe calcolare il polinomio caratteristico e porre $\lambda = 0$ o, equivalentemente, porre $\lambda = 0$ e calcolare il polinomio caratteristico. Ma quando $\lambda = 0$ si ha $det(A - 0I_n) = det(A)$. Dunque il termine noto del polinomio caratteristico della matrice A è $det(A)$. Poiché matrici simili hanno lo stesso polinomio caratteristico, quanto appena osservato è coerente col fatto, già ottenuto in precedenza (come conseguenza diretta della definizione di similitudine di matrici), che matrici simili hanno lo stesso determinante.

Vediamo il significato di altri coefficienti del polinomio caratteristico. Consideriamo il coefficiente di λ^{n-1}: sviluppiamo, come prima, il determinante della matrice $A - \lambda I_n$ rispetto alla prima riga. Per chiarezza di notazioni poniamo $B = A - \lambda I_n$, allora:

$$(*)\quad \begin{aligned} det(A - \lambda I_n) = \ & (a_{11} - \lambda)det(\mathcal{A}_{11} - \lambda I_{n-1}) - a_{12}det(\mathcal{B}_{12}) + \\ & + a_{13}det(\mathcal{B}_{13}) + \ldots + (-1)^{1+n}a_{1n}det(\mathcal{B}_{1n}). \end{aligned}$$

Ricordiamo che le sottomatrici $\mathcal{B}_{1k}$ della matrice B con $k \neq 1$ si ottengono eliminando la prima riga e la colonna k-esima della matrice $B = A - \lambda I_n$, quindi eliminando due entrate della matrice $A - \lambda I_n$ contenenti la variabile λ: l'elemento di posto $1, 1$ e l'elemento di posto k, k. Inoltre le entrate a_{1k} della matrice A per $k \neq 1$ non contengono la variabile λ. Di conseguenza nella espressione $(*)$ gli eventuali termini di grado $n - 1$ nella variabile λ sono contenuti nell'addendo $(a_{11} - \lambda)det(\mathcal{A}_{11} - \lambda I_{n-1})$. Proviamo allora che il coefficiente del termine di grado $n - 1$ è $(-1)^{n-1}(a_{11} + a_{22} + \ldots + a_{nn})$. Per calcolarlo senza troppa fatica si può procedere per induzione su n: se $n = 1$, $A = (a_{11})$ e si vede immediatamente che il polinomio caratteristico di A è $a_{11} - \lambda$ e il termine di grado 0 è proprio $(-1)^0 a_{11} = a_{11}$. Dopodiché, assumiamo di conoscere il risultato per qualunque matrice di ordine minore di n e calcoliamo il coefficiente del termine di grado $n - 1$ della nostra matrice B. Abbiamo detto che si tratta di calcolare i termini di grado $n - 1$ in λ che compaiono nel prodotto $(a_{11} - \lambda)det(\mathcal{A}_{11} - \lambda I_{n-1})$. Ora in $(a_{11} - \lambda)det(\mathcal{A}_{11} - \lambda I_{n-1})$ i termini di grado $n - 1$ sono quelli che si ottengono moltiplicando a_{11} per gli elementi di grado $n-1$ in $det(\mathcal{A}_{11}-\lambda I_{n-1})$ e quelli che si ottengono moltiplicando $-\lambda$ per i termini di grado $n-2$ in $det(\mathcal{A}_{11}-\lambda I_{n-1})$. Osserviamo che $det(\mathcal{A}_{11} - \lambda I_{n-1})$ è precisamente il polinomio caratteristico della matrice $\mathcal{A}_{11}$ che è una matrice quadrata di ordine $n - 1$. Quindi, per ipotesi induttiva e per quanto già osservato, i termini di grado $n - 1$ nel polinomio caratteristico di A sono: $a_{11}((-1)^{n-1}\lambda^{n-1}) - \lambda((-1)^{n-2}(a_{22} + \ldots + a_{nn})\lambda^{n-2}) = (-1)^{n-1}\lambda^{n-1}(a_{11} + a_{22} + \ldots + a_{nn})$.

La somma $a_{11} + a_{22} + \ldots + a_{nn}$ è la somma degli elementi sulla diagonale di A. Tale somma è detta la **traccia** di A e si indica di solito con $\mathrm{tr}(A)$. Abbiamo quindi mostrato che il coefficiente del monomio di grado $n - 1$ del polinomio caratteristico di una matrice quadrata A di ordine n è, a meno del segno (che dipende solo da n), la traccia della matrice A. In particolare matrici simili hanno la stessa traccia! (Esse hanno infatti lo stesso polinomio caratteristico.)

iii) Il polinomio caratteristico di un endomorfismo φ di uno spazio vettoriale n-dimensionale ha grado n ma le sue radici possono anche non essere tutte distinte. Ne consegue che gli autovalori distinti di φ sono in numero minore o uguale a n.

Che relazione c'è tra autospazi relativi ad autovalori diversi? Si intersecano? La prossima proposizione risponde a questa domanda:

Proposizione 10.1.8 *Sia φ un endomorfismo di V e sia $\mathbf{v}$ una base di V, $\dim V = n$. Siano $\alpha_1, \alpha_2, \ldots, \alpha_k$ autovalori distinti di φ. Allora i relativi autospazi sono in somma diretta: $V_{\alpha_1} \oplus V_{\alpha_2} \oplus \ldots \oplus V_{\alpha_k}$.*

Dimostrazione. Si tratta di verificare che l'intersezione fra una somma qualsiasi di alcuni autospazi e qualunque autospazio che non compaia fra gli addendi di questa somma è banale. Procediamo per induzione su k. Innanzitutto mostriamo che la somma di due autospazi distinti è diretta: sia $v \in V_{\alpha_1} \cap V_{\alpha_2}$. Allora $\varphi(v) = \alpha_1 v$ e $\varphi(v) = \alpha_2 v$, i.e., $(\alpha_1 - \alpha_2)v = \mathbf{0}_V$ il che implica, essendo $\alpha_1 \neq \alpha_2$, $v = \mathbf{0}_V$.

Ora supponiamo che la somma dei j autospazi $V_{\alpha_1} \oplus V_{\alpha_2} \oplus \ldots \oplus V_{\alpha_j}$ sia diretta e consideriamo V_{α_l} con $l \notin \{1, 2, \ldots, j\}$. Sia poi $v \in (V_{\alpha_1} \oplus V_{\alpha_2} \oplus V_{\alpha_j}) \cap V_{\alpha_l}$. Si ha dunque: $v = v_1 + v_2 + \ldots + v_j = v_l$, con $v_s \in V_{\alpha_s}$, pertanto $\varphi(v_1 + v_2 + \ldots + v_j) = \varphi(v_l)$, ma φ è lineare e i vettori v_i sono autovettori, dunque

$$\varphi(v_1 + \ldots + v_j) = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_j v_j = \alpha_l v_l = \alpha_l v = \alpha_l v_1 + \ldots + \alpha_l v_j$$

cioè

$$\alpha_l v_1 + \alpha_l v_2 + \ldots + \alpha_l v_j = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_j v_j.$$

Dunque $(\alpha_l - \alpha_1)v_1 + (\alpha_l - \alpha_2)v_2 + \ldots + (\alpha_l - \alpha_j)v_j = \mathbf{0}_V$. Dal momento che, per ipotesi, la somma di $V_1, \ldots, V_j$ è diretta, i vettori $v_1, \ldots, v_j$ sono linearmente indipendenti, ma essendo $\alpha_l - \alpha_s \neq 0$ se $s \neq l$, si ha che tutti i vettori sono nulli: $v_1 = v_2 = \ldots = v_j = v_l = \mathbf{0}_V$. Quindi $(V_1 \oplus \ldots \oplus V_j) \cap V_l = 0_V$. **C. V. D.**

La proposizione 10.1.8 ci assicura che autovettori relativi ad autovalori diversi sono linearmente indipendenti: essi appartengono infatti a spazi vettoriali in somma diretta.

10.2 Matrici/endomorfismi diagonalizzabili

Mettiamoci ora in una situazione "idilliaca": sia $\varphi \in \mathcal{E}nd(V)$ e supponiamo che la somma (diretta) degli autospazi $V_{\alpha_1} \oplus V_{\alpha_2} \oplus \ldots \oplus V_{\alpha_k}$ sia uguale a V, con $\alpha_1, \alpha_2, \ldots, \alpha_k$ autovalori diversi. Allora, come già sottolineato in 5.1.4, se scegliamo una base di ciascun autospazio: $v_1^1, v_2^1, \ldots, v_{n_1}^1$ per V_{α_1}, $v_1^2, v_2^2, \ldots, v_{n_2}^2$ per V_{α_2} e così via, fino a $v_1^k, v_2^k, \ldots, v_{n_k}^k$ per V_{α_k}, otteniamo una base di V come unione delle basi dei V_{α_i}. Adesso cerchiamo di scrivere la matrice dell' applicazione lineare φ rispetto a questa base. Cominciamo da v_1^1: v_1^1 appartiene al sottospazio V_{α_1}, ne segue che $\varphi(v_1^1) = \alpha_1 v_1^1$ che in coordinate, nella base fissata, si scrive come $\alpha_1 v_1^1 + 0v_2^1 + \ldots + 0v_{n_1}^1 + 0v_1^2 + \ldots + 0v_1^k + \ldots + 0v_{n_k}^k$. Per v_2^1 si ha $\varphi(v_2^1) = \alpha_1 v_2^1 = 0v_1^1 + \alpha_1 v_2^1 + 0v_3^1 + \ldots + 0v_{n_1}^1 + 0v_1^2 + \ldots + 0v_1^k + \ldots + 0v_{n_k}^k$. In definitiva la matrice associata a φ rispetto ad una base di autovettori è diagonale e gli elementi sulla diagonale sono gli autovalori:

$$D = \begin{pmatrix} \alpha_1 & 0 & 0 & \ldots & 0 \\ 0 & \alpha_1 & 0 & \ldots & 0 \\ 0 & 0 & \ddots & \ldots & 0 \\ 0 & 0 & \ddots & \alpha_k & 0 \\ 0 & 0 & \ldots & 0 & \alpha_k \end{pmatrix}.$$

Poiché la matrice trovata è la matrice dell'endomorfismo rispetto ad una base di autovettori di φ, il suo polinomio caratteristico è quello dell'endomorfismo φ pertanto gli zeri del polinomio caratteristico di D debbono avere la stessa molteplicità degli zeri del polinomio caratteristico di φ; in altre parole α_i compare sulla diagonale di D un numero di volte pari alla sua molteplicità come radice del polinomio caratteristico di φ.

Definizione 10.2.1 *Un endomorfismo φ di uno spazio vettoriale V si dice diagonalizzabile se V ammette una base di autovettori di φ. Una matrice $A \in \mathcal{M}_n(\mathbb{R})$ si dice diagonalizzabile se è simile ad una matrice diagonale $D \in \mathcal{M}_n(\mathbb{R})$.*

Abbiamo appena visto che se la somma (diretta) degli autospazi coincide con lo spazio vettoriale V allora l'endomorfismo è diagonalizzabile e la matrice ad esso associata rispetto ad una base di autovettori è diagonale.

Proposizione 10.2.2 *Sia α un autovalore di un endomorfismo φ di uno spazio vettoriale V. Allora la molteplicità algebrica di α è sempre maggiore della sua molteplicità geometrica o uguale ad essa.*

Dimostrazione. Indichiamo con k la molteplicità algebrica di α e sia A la matrice dell'endomorfismo φ rispetto ad una base fissata di V. Allora il polinomio caratteristico si può scrivere come: $det(A - zI_n) = (\alpha - z)^k Q(z)$, dove $Q(z)$ ha grado $n - k$ e $Q(\alpha) \neq 0$. Supponiamo ora, per assurdo, che l'autospazio relativo all'autovalore α abbia dimensione h strettamente maggiore di k. Allora possiamo scegliere una base $v_1, v_2, \dots, v_k, v_{k+1}, \dots, v_h$ di V_α e completare questa base in una base $\mathbf{v}$ di V. La matrice associata a φ rispetto alla base $\mathbf{v}$ sarà allora del tipo

$$A' = \left(\begin{array}{ccccc} \alpha & 0 & \dots & 0 & \\ 0 & \alpha & \dots & & \\ \vdots & 0 & \ddots & & * \\ & & & \alpha & \\ \vdots & \vdots & 0 & 0 & \\ \vdots & \vdots & \dots & \vdots & \mathrm{B} \\ 0 & 0 & \dots & 0 & \end{array}\right),$$

in cui gli α sulla diagonale sono h e B è una matrice quadrata di ordine $n - h$. Il polinomio caratteristico della matrice A' è $(\alpha - z)^h det(B - zI_{n-h})$. Ma A e A' sono matrici associate alla stessa applicazione lineare rispetto a basi diverse, pertanto sono simili e hanno lo stesso polinomio caratteristico. Questo non è possibile dal momento che α è radice del polinomio $(\alpha - z)^k Q(z)$ con molteplicità k ed è radice del polinomio $(\alpha - z)^h det(B - zI_{n-h})$ con molteplicità almeno h, ma per ipotesi $h > k$. Assurdo.

Per quanto detto finora possiamo affermare che se il polinomio caratteristico $p(z)$ di un endomorfismo φ di uno spazio vettoriale V di dimensione n si fattorizza in $\mathbb{R}[z]$ nel prodotto di fattori lineari (i.e. se le radici di $p(z)$ sono tutte reali) e la molteplicità di ogni autovalore come radice del polinomio caratteristico coincide con la dimensione dell'autospazio associato, allora l'endomorfismo è diagonalizzabile: in questo caso, infatti, sarà possibile individuare una base dello spazio vettoriale ambiente costituita da autovettori di φ. Più precisamente, il fatto che le radici $\alpha_1, \dots, \alpha_k$ del polinomio caratteristico siano reali assicura che, indicata con n_i la molteplicità dell'autovalore α_i, $n_1 + n_2 + \dots + n_k = n$. Del resto la somma $V_{\alpha_1} \oplus \dots \oplus V_{\alpha_k}$ degli autospazi è diretta e, dal momento che per ogni $i = 1, \dots, k$ la dimensione dell'autospazio V_{α_i} coincide con la molteplicità n_i dell'autovalore α_i, si ha $\dim(V_{\alpha_1} \oplus \dots \oplus V_{\alpha_k}) = n_1 + \dots + n_k = n$ quindi $V_{\alpha_1} \oplus \dots \oplus V_{\alpha_k} = V$.

Vogliamo ora mostrare che, viceversa, se un endomorfismo φ di uno spazio vettoriale V n-dimensionale è diagonalizzabile allora il suo polinomio caratteristico $p(z)$ si fattorizza nel prodotto di fattori lineari su $\mathbb{R}[z]$ e la molteplicità di ogni autovalore come radice del polinomio caratteristico coincide con la dimensione dell'autospazio associato. In effetti, sia A la matrice associata a φ rispetto ad una base fissata di V. Se A è diagonalizzabile, per definizione essa è simile ad una matrice diagonale, cioè esiste una matrice invertibile H tale che $H^{-1}AH = D$ è diagonale. In particolare le matrici A e D hanno lo stesso polinomio caratteristico e, essendo D diagonale, tale polinomio caratteristico è il prodotto degli elementi diagonali della matrice $D - zI_n$, quindi esso è il prodotto di fattori lineari in z:

$$(\alpha_1 - z)^{n_1}(\alpha_2 - z)^{n_2} \cdots (\alpha_k - z)^{n_k}$$

dove gli α_i sono gli autovalori di D (e quindi di A) e $n_1 + n_2 + \ldots + n_k = n$, essendo n_i il numero di volte in cui α_i appare nella matrice diagonale D. Abbiamo così mostrato che il polinomio caratteristico $p(z)$ di una matrice diagonalizzabile si fattorizza in $\mathbb{R}[z]$ in polinomi di grado 1 (non necessariamente distinti).

D'altra parte la matrice D è la matrice associata a φ rispetto ad una base $\mathcal{B}$ di autovettori di φ e la molteplicità con cui ogni autovalore α_i appare come radice del polinomio caratteristico, cioè il numero di volte in cui α_i compare sulla diagonale della matrice D, è uguale al numero di autovettori relativi all'autovalore α_i che appaiono nella base $\mathcal{B}$. Quindi la dimensione dell'autospazio relativo ad ogni autovalore è uguale alla molteplicità dell'autovalore come zero del polinomio caratteristico.

Abbiamo così dimostrato il seguente

Teorema 10.2.3 *Una matrice $A \in \mathcal{M}_n(\mathbb{R})$ (o, equivalentemente, un endomorfismo φ di uno spazio vettoriale n-dimensionale V) è diagonalizzabile se e solo se valgono le seguenti due condizioni:*

1. *Il polinomio caratteristico $p(z)$ di A (o di φ) si fattorizza in $\mathbb{R}[z]$ nel prodotto di polinomi di primo grado (non necessariamente distinti);*

2. *la molteplicità di ogni autovalore λ come radice del polinomio caratteristico coincide con la dimensione dell'autospazio V_λ.*

Osservazione 10.2.4 *i*) Abbiamo già osservato che esistono matrici ad entrate reali prive di autovalori reali. Nello stesso modo esistono matrici il cui polinomio caratteristico, pur avendo alcune radici reali, non si fattorizza completamente in polinomi di grado 1 in $\mathbb{R}[x]$, cioè le sue radici non sono tutte reali. Consideriamo, ad esempio, la matrice

$$A = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Il suo polinomio caratteristico è $(3-z)(z^2+1)$ che ha una sola radice reale – $z = 3$ – (e due radici immaginarie).

ii) Anche se il polinomio caratteristico di una matrice è completamente fattorizzabile in $\mathbb{R}[x]$, può accadere che la matrice non sia diagonalizzabile. In effetti prendiamo la matrice $B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Il suo polinomio caratteristico è $z^2 = (-z+0)^2$. Quindi B ha un solo autovalore: $\alpha = 0$ di molteplicità 2. Se B fosse diagonalizzabile sarebbe simile alla matrice diagonale con l'autovalore 0 sulla diagonale, cioè alla matrice nulla. Ma qualunque matrice simile alla matrice nulla è nulla: $H^{-1}\mathbf{O}_nH = \mathbf{O}_n$! Quindi la matrice B, che è non nulla, non può essere simile ad una diagonale, cioè B non è diagonalizzabile.

iii) Se A è una matrice diagonalizzabile la sua forma diagonale D è completamente determinata dai suoi autovalori: D ha come elementi (diagonali) gli autovalori di A in numero pari alla loro molteplicità come radici del polinomio caratteristico.

iv) Se α è radice del polinomio caratteristico di un endomorfismo φ, allora φ ammette almeno un autovettore di autovalore α, quindi la dimensione dell'autospazio V_α è maggiore o uguale ad 1. In particolare se il polinomio caratteristico di una matrice è completamente fattorizzabile in $\mathbb{R}[x]$ e tutte le sue radici sono distinte, cioè se la molteplicità di ogni autovalore è uguale ad uno, allora la dimensione di ogni autospazio è esattamente uguale ad 1, dovendo essere maggiore o uguale ad 1 per quanto appena detto e minore o uguale ad 1 per quanto osservato sopra. Quindi in questo caso la molteplicità di ogni autovalore coincide con la dimensione dell'autospazio corrispondente. In questo caso, dunque, l'endomorfismo è diagonalizzabile.

Veniamo dunque al metodo per verificare se un endomorfismo (e quindi una matrice) sia o meno diagonalizzabile e trovare, in caso affermativo, una base di autovettori che lo diagonalizzi. Siano dati l'endomorfismo φ di V e la matrice $A \in \mathcal{M}_n(\mathbb{R})$ associata a φ rispetto ad una base fissata di V.

1) Si calcola il polinomio caratteristico $det(A-zI_n)$. Se non è prodotto di fattori lineari in $\mathbb{R}[x]$, i.e., se le sue radici non sono tutte reali allora φ non è diagonalizzabile. Altrimenti: $det(A-zI_n) = (\alpha_1-z)^{n_1}(\alpha_2-z)^{n_2}\dots(\alpha_k-z)^{n_k}$, $\alpha_i \in \mathbb{R}$.

2) Si determinano gli autospazi: V_{α_i} è dato dalle soluzioni del sistema

$$(A-\alpha_i I_n)\begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} = \mathbf{0}_{\mathbb{R}^n}$$

e quindi è un sottospazio di V di dimensione $n - rg(A - \alpha_i I_n)$. Se tale dimensione concide con n_i per ogni autovalore α_i allora la matrice è diagonalizzabile. Per trovare una base che la diagonalizzi si dovrà scegliere una base per ogni autospazio e prendere l'unione delle basi trovate.

Nel caso in cui si voglia soltanto sapere se una matrice $A \in \mathcal{M}_n(\mathbb{R})$ è diagonalizzabile o meno, basterà verificare se per ogni autovalore α_i di molteplicità n_i come radice del polinomio caratteristico si ha $n_i = n-rg(A-\alpha_i I_n)$ (cioè se n_i coincide con la dimensione dell'autospazio associato). In caso affermativo la matrice è diagonalizzabile, altrimenti non lo è. Questa verifica è superflua per gli autovalori con molteplicità 1 dal momento che per questi l'uguaglianza è automatica!

Esempio 10.2.5 Si consideri la matrice

$$T = \begin{pmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{pmatrix}.$$

Dire se T è diagonalizzabile e, in caso affermativo, trovare una matrice che la diagonalizzi, i.e. una matrice $H \in Gl_3(\mathbb{R})$ per cui $H^{-1}TH$ sia una matrice diagonale.

Svolgimento. Seguiamo esattamente le linee guida del procedimento che abbiamo illustrato. Calcoliamo dunque il polinomio caratteristico di T nella incognita λ:

$$det(\begin{pmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{pmatrix} - \lambda I_3) = det\begin{pmatrix} 3-\lambda & 2 & 4 \\ 2 & -\lambda & 2 \\ 4 & 2 & 3-\lambda \end{pmatrix} = P_T(\lambda).$$

(En passant, voilá-voilu, sottolineiamo ancora una volta il fatto che tale polinomio è lo stesso per tutte le matrici simili a T – e quindi anche a $H^{-1}TH$, con $H \in Gl_3(\mathbb{R})$–). Sviluppando il determinante di $T - \lambda I_3$ rispetto alla prima riga e ricordando che il nostro obbiettivo è calcolare le radici del polinomio caratteristico di T, abbiamo: $P_T(\lambda) = (3-\lambda)(-\lambda(3-\lambda)-4) - 2(-2\lambda-2) + 4(4+4\lambda) = (\lambda+1)((3-\lambda)(\lambda-4)+20) = -(\lambda+1)^2(\lambda-8)$. Il polinomio caratteristico di T è dunque fattorizzabile in $\mathbb{R}[x]$ in fattori di primo grado. Le sue radici sono: 8, di molteplicità 1, e -1 di molteplicità 2. Per verificare che la matrice T sia diagonalizzabile occorre allora controllare che l'autospazio relativo all'autovalore -1 abbia dimensione due, cioè che la dimensione di V_{-1} coincida con la molteplicità dell'autovalore -1 come radice del polinomio caratteristico. Per l'autovalore 8 il risultato è vero automaticamente dal momento che esso ha molteplicità uguale ad 1.

Determiniamo l'autospazio relativo all' autovalore -1. Tale autospazio è dato dalle soluzioni del sistema:

$$\begin{pmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = (-1) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

cioè dall'insieme delle terne (x_1, x_2, x_3) che soddisfano il sistema

$$(T + I_3) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3+1 & 2 & 4 \\ 2 & 0+1 & 2 \\ 4 & 2 & 3+1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

A priori sappiamo che la dimensione di V_{-1}, cioè dello spazio delle soluzioni che stiamo cercando, è uguale ad 1 o a 2. Per calcolare tale dimensione osserviamo che le colonne della matrice $T + I_3$ sono una multipla dell'altra, pertanto il rango della matrice $T + I_3$ è uguale a 1. Di conseguenza $\dim(V_{-1}) = 3 - 1 = 2$. Quindi $\dim(V_{-1})$ coincide con la molteplicità dell'autovalore -1 come radice del polinomio caratteristico. La matrice T è pertanto diagonalizzabile

Determiniamo ora una base di autovettori di T. Cominciamo con l'autospazio relativo all'autovalore -1: dal momento che la matrice $T + I_3$ ha rango 1, V_{-1} è dato dalle soluzioni di una sola equazione: $2x_1 + x_2 + 2x_3 = 0$. Due soluzioni linearmente indipendenti di questa equazione sono, ad esempio, $(1, -2, 0)$ e $(0, -2, 1)$. Dunque $V_{-1} =< (1, -2, 0), (0, -2, 1) >$.

Calcoliamo ora l'autospazio relativo all'autovalore 8. Il sistema che caratterizza tale autospazio è allora:

$$(T-8I_3)\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}3-8 & 2 & 4\\ 2 & -8 & 2\\ 4 & 2 & 3-8\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}.$$

Sappiamo già che l'insieme delle soluzioni di questo sistema è uno spazio di dimensione 1 (la molteplicità dell'autovalore 8 come radice del polinomio caratteristico è infatti uguale ad 1 e la dimensione di V_8, che è certo non banale, non può essere più grande di 1). Quindi V_8 è l'insieme delle soluzioni di due equazioni linearmente indipendenti, ad esempio quelle individuate dalla prima e dalla seconda riga della matrice $T-8I_3$. Per sostituzione otteniamo: $x_1=4x_2-x_3$ e $-20x_2+5x_3+2x_2+4x_3=0$, cioè $x_3=2x_2$ e $x_1=2x_2$. Si ha pertanto $V_8 =< (2,1,2) >$. Una base di autovettori di T è quindi data da $v_1=(1,-2,0), v_2=(0,-2,1), v_3=(2,1,2)$ e la matrice di passaggio da questa base a quella di partenza è la matrice

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

In conclusione la forma diagonale della matrice T è

$$D=\begin{pmatrix}-1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & 8\end{pmatrix}=H^{-1}TH.$$

Osserviamo che la matrice D è la matrice dell'applicazione lineare di partenza (associata alla matrice T rispetto alla base canonica di $\mathbb{R}^3$) rispetto alla base v_1, v_2, v_3. Le sue colonne sono dunque date dalle coordinate dei vettori $T(v_1)$, $T(v_2)$, $T(v_3)$ rispetto alla base v_1, v_2, v_3. Poiché questi vettori sono autovettori si ha, effettivamente, $T(v_1)=-v_1=(-1)v_1+0v_2+0v_3$, $T(v_2)=-v_2=0v_1+(-1)v_2+0v_3$, $T(v_3)=8v_3=0v_1+0v_2+8v_3$. Nella lezione XII vedremo che una matrice simmetrica reale (cioè una matrice che coincide con la propria trasposta) è sempre diagonalizzabile.

Esempio 10.2.6 Stabilire, al variare di $t\in\mathbb{R}$, se la matrice

$$B_t=\begin{pmatrix}t-1 & t-3 & t\\ 0 & \frac{5}{2} & -\frac{1}{2}\\ 0 & -\frac{1}{2} & \frac{5}{2}\end{pmatrix}$$

è diagonalizzabile o meno.

Svolgimento. Calcoliamo innanzitutto il polinomio caratteristico della matrice B_t:

$$det(\begin{pmatrix} t-1-\lambda & t-3 & t \\ 0 & \frac{5}{2}-\lambda & -\frac{1}{2} \\ 0 & -\frac{1}{2} & \frac{5}{2}-\lambda \end{pmatrix}) = (t-1-\lambda)(\lambda^2-5\lambda+6).$$

Le sue radici sono : $t-1, 2, 3$. Ora se $t \neq 3, 4$ le tre radici sono diverse e quindi la matrice è diagonalizzabile.

Si noti che per ogni $t \neq 3, 4$ l'autospazio relativo all'autovalore $t-1$ è in ogni caso $< (1, 0, 0) >$.

Consideriamo ora i casi $t = 3, 4$: in questi casi dobbiamo studiare direttamente le matrici B_3 e B_4. La matrice B_3 ha autovalore 2 di molteplicità due e quindi, affinché essa sia diagonalizzabile, occorre che l'autospazio V_2 relativo all'autovalore 2 abbia dimensione 2. Si ha:

$$\dim(V_2) = 3 - rg(B_3 - 2I_3) = 3 - rg(\begin{pmatrix} 3-1-2 & 3-3 & 3 \\ 0 & \frac{5}{2}-2 & -\frac{1}{2} \\ 0 & -\frac{1}{2} & \frac{5}{2}-2 \end{pmatrix}) = 3-2 = 1.$$

Quindi la dimensione di V_2 non coincide con la molteplicità dell'autovalore 2: la matrice B_3 non è diagonalizzabile.

Consideriamo la matrice B_4: essa ha autovalore 3 di molteplicità 2. Calcoliamo la dimensione di V_3:

$$\dim(V_3) = 3 - rg(B_4 - 3I_3) = 3 - rg(\begin{pmatrix} 4-1-3 & 4-3 & 4 \\ 0 & \frac{5}{2}-3 & -\frac{1}{2} \\ 0 & -\frac{1}{2} & \frac{5}{2}-3 \end{pmatrix}) = 3-2 = 1.$$

Anche in questo caso la dimensione dell'autospazio V_3 non coincide con la molteplicità dell'autovalore 3. La matrice B_4 non è diagonalizzabile.

Esempio 10.2.7 Si determinino gli autovalori e gli autospazi della matrice

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

Svolgimento. Come al solito occorre calcolare il polinomio caratteristico della matrice A:

$$det(\begin{pmatrix} 1-\lambda & 0 & 0 & 0 \\ 0 & 1-\lambda & 0 & 0 \\ -2 & -2 & -\lambda & 1 \\ -2 & 0 & -1 & 2-\lambda \end{pmatrix}) = (1-\lambda)^2(\lambda^2 - 2\lambda + 1) = (1-\lambda)^4.$$

Dunque la matrice A ha un unico autovalore $\lambda = 1$ di molteplicità 4. Studiamo il relativo autospazio. Dobbiamo cercare le soluzioni del sistema

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -2 & -2 & 0 & 1 \\ -2 & 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix},$$

cioè le soluzioni del sistema

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -2 & -2 & -1 & 1 \\ -2 & 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

La matrice del sistema omogeneo trovato ha rango due e quindi l'autospazio relativo all'autovalore 1 ha dimensione 2, in particolare la matrice A non è diagonalizzabile. L'autospazio V_1 è l'insieme delle soluzioni del sistema

$$\begin{cases} -2x_1 - 2x_2 - x_3 + x_4 = 0 \\ -2x_1 - x_3 + x_4 = 0 \end{cases}$$

da cui otteniamo: $x_3 = -2x_1 + x_4$ e $x_2 = 0$. Una base di V_1 è allora data dai vettori $(1, 0, -2, 0)$ e $(0, 0, 1, 1)$: $V_1 =< (1, 0, -2, 0), (0, 0, 1, 1) >$

Si noti che se la matrice A fosse stata diagonalizzabile la sua forma diagonale sarebbe stata la matrice identica I_4. In questo caso avremmo avuto $H^{-1}AH = I_4$ per qualche matrice invertibile H di ordine 4. Allora, moltiplicando entrambi i membri dell'uguaglianza a destra per H^{-1} e a sinistra per H, avremmo avuto: $H(H^{-1}AH)H^{-1} = HI_4H^{-1}$ cioè $A = I_4$. Ma questo è ovviamente falso! Si noti ancora che se una matrice $B \in \mathcal{M}_n(\mathbb{R})$ è diagonalizzabile e il suo polinomio caratteristico è $(-1)^n(\lambda - \alpha)^n$, con $\alpha \in \mathbf{R}$, allora $B = \alpha I_n$, dove I_n è la matrice identica in $\mathcal{M}_n(\mathbb{R})$.

10.3 Esercizi svolti

Esercizio 10.3.1 Dimostrare che 0 è autovalore per la matrice $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$.

Svolgimento. Osserviamo che 0 è autovalore per la matrice A se esiste un vettore $v \neq 0_{\mathbb{R}^2}$ tale che $Av = 0v = 0_{\mathbb{R}^2}$, cioè se esiste un vettore $v \neq 0_{\mathbb{R}^2}$ nel nucleo della matrice A. In altre parole la matrice A ammette l'autovalore 0 se e solo se essa è non invertibile. In effetti le righe di A sono linearmente dipendenti (anzi, uguali!) perciò $det(A) = 0$. Dunque A non è invertibile.

Esercizio 10.3.2 Si consideri lo spazio vettoriale $\mathbb{R}^{\leq 3}[x]$ dei polinomi di grado minore o uguale a 3 nella variabile x, a coefficienti reali, e l'operatore di derivazione

$$D : \mathbb{R}^{\leq 3}[x] \longrightarrow \mathbb{R}^{\leq 3}[x]$$

che associa ad ogni polinomio la sua derivata prima. Determinare gli autovalori di D e calcolarne i relativi autospazi. Decidere se D è diagonalizzabile e, in caso affermativo, determinarne la forma diagonale.

Svolgimento. Nell'esercizio 3 del capitolo 6 abbiamo determinato la matrice associata all'applicazione lineare D rispetto alla base canonica $\{1, x, x^2, x^3\}$ di $\mathbb{R}^{\leq 3}[x]$:

$$D = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Il polinomio caratteristico della matrice D è dunque:

$$det(D - \lambda I_4) = det \begin{pmatrix} -\lambda & 1 & 0 & 0 \\ 0 & -\lambda & 2 & 0 \\ 0 & 0 & -\lambda & 3 \\ 0 & 0 & 0 & -\lambda \end{pmatrix} = \lambda^4.$$

La matrice D ha pertanto un solo autovalore $\lambda = 0$ di molteplicità algebrica 4. Allora D non è certamente diagonalizzabile: se lo fosse la sua forma diagonale sarebbe la matrice identicamente nulla, ma l'unica matrice simile alla matrice nulla è la matrice nulla, e D non è certamente la matrice nulla!

Determiniamo l'autospazio V_0 relativo all'autovalore $\lambda = 0$: $V_0 = KerD = \{p(x) \in \mathbb{R}^{\leq 3}[x] \mid D(p(x)) = 0\} = \mathbb{R}$. In particolare $\dim(V_0) = 1$.

Esercizio 10.3.3 Calcolare gli autovalori della matrice $A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & -1 \\ 0 & -1 & 2 \end{pmatrix}$. Stabilire se essa è diagonalizzabile su $\mathbb{R}$ ed in caso affermativo scrivere la forma diagonale ed una matrice diagonalizzante.

Svolgimento. Calcoliamo il polinomio caratteristico della matrice A:
$$det(A - tI_3) = det \begin{pmatrix} 2-t & 1 & 0 \\ 1 & 1-t & -1 \\ 0 & -1 & 2-t \end{pmatrix} = (2-t)((1-t)(2-t)-1) -$$
$(2-t) = t(2-t)(t-3)$.

La matrice A ha 3 autovalori distinti: $t_1 = 0$, $t_2 = 2$, $t_3 = 3$ ed è pertanto diagonalizzabile. Esiste cioè una matrice invertibile H tale che $H^{-1}AH = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$. Determiniamo la matrice H: le sue colonne costituiscono una base di $\mathbb{R}^3$ di autovettori di A. Si tratta dunque di determinare gli autospazi della matrice A. Cominciamo con $V_0 = KerA = \{(x,y,z) \mid A(x,y,z)^t = (0,0,0)^t\}$. Facendo i calcoli si ottiene $V_0 =< (1,-2,-1) >$.

Analogamente $V_2 = \{(x,y,z) \mid (A - 2I_3)(x,y,z)^t = (0,0,0)^t\} =< (1,0,1) >$.

Infine, $V_3 = \{(x,y,z) \mid (A - 3I_3)(x,y,z)^t = (0,0,0)^t\} =< (1,1,-1) >$.

Siamo ora in grado di costruire la matrice H:
$$H = \begin{pmatrix} 1 & 1 & 1 \\ -2 & 0 & 1 \\ -1 & 1 & -1 \end{pmatrix}.$$

Osservazione Si noti l'ordine in cui abbiamo scritto le componenti degli autovettori nelle colonne della matrice H: nella prima colonna appaiono le componenti di un elemento v_0 di V_0, nella seconda le componenti di un elemento v_2 di V_2, infine, nella terza colonna, le componenti di un elemento v_3 di V_3. Che cosa sarebbe successo se avessimo disposto i vettori sulle colonne della matrice in un ordine diverso? Naturalmente avremmo trovato un'altra matrice diagonalizzante la matrice di partenza A. Ad esempio, posto $K = (v_2, v_0, v_3)$ (dove con v_i intendiamo i vettori colonna), si ha:
$$K^{-1}AK = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Esercizio 10.3.4 Sia $V = \mathcal{M}_2(\mathbb{R})$ e sia f l'applicazione lineare di V in sé definita da:

$$f(X) = XA \qquad\qquad (X \in V)$$

dove $A = \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix}$. Determinare gli autovalori di f.

Svolgimento. Fissiamo innanzitutto la base canonica $\mathcal{C}$ di $\mathcal{M}_2(\mathbb{R})$: $\mathcal{C} = \{e_{11}, e_{12}, e_{21}, e_{22}\}$ dove $e_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $e_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $e_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, $e_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$.

Determiniamo ora la matrice F associata all'applicazione lineare f rispetto alla base $\mathcal{C}$ sia nel dominio che nel codominio. Dal momento che lo spazio vettoriale $V = \mathcal{M}_2(\mathbb{R})$ ha dimensione 4, la matrice F sarà una matrice quadrata di ordine 4. Abbiamo:

$f(e_{11}) = e_{11}A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = e_{12}$;

$f(e_{12}) = e_{12}A = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix} = e_{11} + 2e_{12}$;

$f(e_{21}) = e_{21}A = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = e_{22}$;

$f(e_{22}) = e_{22}A = \begin{pmatrix} 0 & 0 \\ 1 & 2 \end{pmatrix} = e_{21} + 2e_{22}$.

Otteniamo così la matrice

$$F = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 2 \end{pmatrix}$$

il cui polinomio caratteristico è:

$$det(F - tI_4) = det \begin{pmatrix} -t & 1 & 0 & 0 \\ 1 & 2-t & 0 & 0 \\ 0 & 0 & -t & 1 \\ 0 & 0 & 1 & 2-t \end{pmatrix} = (t^2 - 2t - 1)^2.$$

Gli autovalori di F sono le radici del polinomio $t^2 - 2t - 1$ cioè $t_1 = 1 + \sqrt{2}$, $t_2 = 1 - \sqrt{2}$.

Esercizio 10.3.5

Studiare, al variare del parametro reale k, la diagonalizzabilità della matrice $A_k = \begin{pmatrix} 1 & k & 0 \\ 1 & 0 & 0 \\ 1-k & 0 & 1 \end{pmatrix}$.

Svolgimento. Calcoliamo per prima cosa il polinomio caratteristico della matrice A_k:

$$p_k(t) = det(A_k - tI_3) = det \begin{pmatrix} 1-t & k & 0 \\ 1 & -t & 0 \\ 1-k & 0 & 1-t \end{pmatrix} = (1-t)(t^2 - t - k).$$

Il polinomio $p_k(t)$ ha radici $t_1 = 1$, $t_2 = \frac{1+\sqrt{1+4k}}{2}$, $t_3 = \frac{1-\sqrt{1+4k}}{2}$. Così se $k \neq 0, -\frac{1}{4}$ le radici di $p_k(t)$ sono distinte e la matrice A_k è pertanto diagonalizzabile.

Sia ora $k = 0$. Allora l'autovalore $t = 1$ ha molteplicità algebrica 2. Calcoliamo la sua molteplicità geometrica (la molteplicità algebrica di un autovalore è la sua molteplicità come radice del polinomio caratteristico; la molteplicità geometrica di un autovalore è la dimensione del corrispondente autospazio):

$$\dim(V_1) = 3 - rg(A_0 - I_3) = 3 - rg \begin{pmatrix} 0 & 0 & 0 \\ 1 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix} = 1.$$

Dal momento che la molteplicità geometrica di $t = 1$ non coincide con la sua molteplicità algebrica, per $k = 0$ la matrice A_k non è diagonalizzabile.

Sia, infine, $k = -\frac{1}{4}$. In questo caso l'autovalore $t = \frac{1}{2}$ ha molteplicità algebrica 2. Si ha:

$$\dim(V_{\frac{1}{2}}) = 3 - rg(A_{-\frac{1}{4}} - \frac{1}{2}I_3) = 3 - rg \begin{pmatrix} \frac{1}{2} & -\frac{1}{4} & 0 \\ 1 & -\frac{1}{2} & 0 \\ \frac{5}{4} & 0 & \frac{1}{2} \end{pmatrix} = 3 - 2 = 1.$$

Anche in questo caso, dunque, la matrice A_k non è diagonalizzabile.

In conclusione A_k è diagonalizzabile se e solo se $k \neq 0, -\frac{1}{4}$.