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 $\mathrm{H_2SO_4}$:

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

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 $a_{11}, \dots, a_{1n}, \dots, a_{m1}, \dots, a_{mn}$ si chiamano coefficienti del sistema, $b_1, \dots, b_m$ termini noti. Se $b_i = 0$ per ogni $i = 1, \dots, m$ il sistema si dice *omogeneo*. Una *soluzione* del sistema lineare (1.2) è una n-upla $(s_1, s_2, \dots, s_n)$ 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:

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

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:

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

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. □

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:

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

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:

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

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

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

- 4^a riga $\rightarrow$ 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:

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

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

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 $\forall \alpha \in \mathbb{R}\backslash\{0,1\}$, 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:

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

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

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

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 $k = 2$:

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

Riducendo questa matrice in forma a scala otteniamo la matrice:

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

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, disegnamo su fogli essenzialmente bidimensionali, e così via.

2.1 Definizione di spazio vettoriale reale

Cos'è uno spazio vettoriale (reale)? Uno spazio vettoriale è un insieme non vuoto V dotato di una somma, su cui 'agisce dall'esterno' l'insieme $\mathbb{R}$ dei numeri reali. Ma cosa vuol dire agire dall'esterno? Vuol dire che, preso un elemento v di questo insieme V e preso un qualsiasi elemento c di $\mathbb{R}$, viene associato a questa coppia un elemento di V che denoteremo con cv. Passiamo alla definizione vera e propria.

Definizione 2.1.1 *Diremo che un insieme non vuoto V è uno spazio vettoriale su $\mathbb{R}$ o, equivalentemente, un $\mathbb{R}$-spazio vettoriale, se:*

- *su V è definita una operazione detta somma, $+_V$ (operazione interna), che è: commutativa, associativa, ammette elemento neutro, $\mathbf{0}_V$, detto vettore nullo, e tale che ogni elemento v di V ammetta opposto denotato con $-v$;*
- *è definita un'azione esterna di $\mathbb{R}$ su V, cioè un'applicazione $\mathbb{R} \times V \to V$, che ad ogni coppia (r, v) con $r \in \mathbb{R}$ e $v \in V$ associa un unico elemento (che denoteremo con) $rv \in V$. Questa operazione è detta operazione esterna.*

Le precedenti operazioni godono delle seguenti proprietà di compatibilità:
(i) $1v = v$, *per ogni* $v \in V$.
(ii) $\forall v \in V$ *e* $\forall \alpha, \beta \in \mathbb{R}$, $(\alpha + \beta)v = \alpha v +_V \beta v$.
(iii) $\forall v, w \in V$ *e* $\forall \alpha \in \mathbb{R}$, $\alpha(v +_V w) = \alpha v +_V \alpha w$.
(iv) $\forall v \in V$ *e* $\forall \alpha, \beta \in \mathbb{R}$, $(\alpha\beta)v = \alpha(\beta v)$.

Gli elementi di V sono detti vettori, gli elementi di $\mathbb{R}$ scalari; l'operazione esterna si dice anche prodotto per scalari. Talvolta, qualora non vi sia pericolo di confusione, indicheremo la somma $+_V$ in V semplicemente con $+$. Anche se lavoreremo sempre con i numeri reali è importante sapere che nella definizione di spazio vettoriale che abbiamo appena dato, l'insieme dei numeri reali può essere sostituito con un altro insieme avente analoghe proprietà, ad esempio $\mathbb{Q}$ (campo dei numeri razionali).

2.2 Proprietà degli spazi vettoriali

Elenchiamo in questo paragrafo alcune utili proprietà degli spazi vettoriali.

1. Calcoliamo il prodotto del numero reale 0 per il vettore $v \in V$: $0v$. Chi è questo elemento? Per la proprietà *(ii)* della definizione 2.1.1 si ha $0v = (0 + 0)v = 0v +_V 0v$. Sommando a destra e a sinistra l'opposto di $0v$, che denoteremo con $-0v$, si ha: $\mathbf{0}_V = 0v +_V (-0v) = 0v +_V 0v +_V (-0v) = 0v$; la prima uguaglianza vale poiché $(-0v)$ è l'opposto di $0v$, la seconda perché si aggiunge a due elementi uguali lo stesso elemento e infine l'ultima perché $0v +_V (-0v) = \mathbf{0}_V$ e $0v +_V \mathbf{0}_V = 0v$. Per la proprietà transitiva dell'uguaglianza si ha $0v = \mathbf{0}_V$ vettore nullo di V.

2. Per la definizione 2.1.1 (i) $1v = v$ per ogni $v \in V$. Ora considerando $(1+(-1))v$, per la proprietà distributiva si ottiene $\mathbf{0}_V = 0v = 1v +_V (-1)v = v +_V (-1)v$ quindi $(-1)v = -v$ è l'opposto di v. Tale opposto verrà indicato semplicemente con $-v$. Notiamo che se nel precedente ragionamento avessimo sostituito 1 con un qualsiasi numero reale α avremmo ottenuto che l'opposto di αv è $(-\alpha)v = -\alpha v$.

3. Per ogni $\alpha \in \mathbb{R}$ vale $\alpha\mathbf{0}_V = \mathbf{0}_V$. Infatti dalle proprietà precedenti degli spazi vettoriali si ha che:

$$\alpha\mathbf{0}_V = \alpha(\mathbf{0}_V +_V \mathbf{0}_V) = \alpha\mathbf{0}_V +_V \alpha\mathbf{0}_V.$$

Sommando ad ambedue i membri dell'uguaglianza l'opposto di $\alpha\mathbf{0}_V$, che indichiamo con $-\alpha\mathbf{0}_V$, otteniamo

$$\alpha\mathbf{0}_V +_V (-\alpha\mathbf{0}_V) = (\alpha\mathbf{0}_V +_V \alpha\mathbf{0}_V) +_V (-\alpha\mathbf{0}_V)$$

che, per la proprietà associativa e per la definizione di elemento neutro, equivale a

$$\mathbf{0}_V = \alpha\mathbf{0}_V + (\alpha\mathbf{0}_V +_V -\alpha\mathbf{0}_V) = \alpha\mathbf{0}_V +_V \mathbf{0}_V = \alpha\mathbf{0}_V.$$

4. Se $\alpha \neq \beta$ con $\alpha, \beta \in \mathbb{R}$, allora $\alpha v \neq \beta v$ per ogni $v \in V$, $v \neq \mathbf{0}_V$. Infatti se $\alpha v = \beta v$, allora sommando a destra e a sinistra per l'opposto di βv, si avrebbe $(\alpha - \beta)v = \mathbf{0}_V$. Da cui moltiplicando ambedue i membri per $1/(\alpha - \beta)$ (essendo $\alpha - \beta \neq 0$) si otterrebbe $v = 1/(\alpha - \beta)\mathbf{0}_V = \mathbf{0}_V$ che contraddirebbe l'ipotesi $v \neq \mathbf{0}_V$. Dunque è stato assurdo pensare che αv e βv fossero eguali.

5. Può esistere uno spazio vettoriale su $\mathbb{R}$ con un numero finito di vettori? Certamente un esempio è lo spazio vettoriale banale, $V = \{\mathbf{0}_V\}$, cioè lo spazio vettoriale costituito dal solo elemento nullo in cui le operazioni interna ed esterna sono banali: $\mathbf{0}_V + \mathbf{0}_V = \mathbf{0}_V$, $\alpha\mathbf{0}_V = \mathbf{0}_V$ per ogni α in $\mathbb{R}$. È facile verificare la validità degli assiomi della definizione 2.1.1. La domanda che ci poniamo ora è la seguente: esistono altri esempi di spazi vettoriali reali con un numero finito di elementi? Supponiamo che V sia uno spazio vettoriale su $\mathbb{R}$ con un numero finito di elementi, i.e. un numero finito di vettori, diciamo $n \in \mathbb{N}$. Sia $v \in V$, $v \neq \mathbf{0}_V$. Per la proprietà precedente $v, 2v, 3v, 4v, \ldots, nv$ sono tutti diversi fra loro e poiché sono n devono essere tutti e soli i vettori di V. Ne segue che, essendo $(n + 1)v$ un vettore di V, $(n + 1)v$ dovrà essere uguale ad uno dei precedenti, ma ciò è assurdo perché si avrebbe $(n+1)v = sv$ con $s \in \mathbb{N}$, $s \leq n$, cioè $s \neq n + 1$. È stato assurdo pensare che V contenesse solo un numero finito di elementi. Quindi l'unico spazio vettoriale su $\mathbb{R}$ con un numero finito di vettori è lo spazio vettoriale banale. Il metodo appena illustrato è detto delle "gabbie dei piccioni": se si ha un numero di gabbie inferiore al numero di piccioni e se si vogliono mettere tutti i piccioni in gabbia, allora necessariamente almeno una gabbia dovrà alloggiare più di un piccione.

2.3 Esempi

2.3.1 Spazio vettoriale banale

Ogni spazio vettoriale contiene il vettore nullo cioè l'elemento neutro rispetto alla somma. Prendiamo l'insieme formato dal solo vettore nullo $\{\mathbf{0}_V\}$, cioè lo spazio vettoriale banale introdotto in 2.2.5. In esso la somma e il prodotto per scalari sono definiti in modo ovvio:

$$\mathbf{0}_V + \mathbf{0}_V = \mathbf{0}_V; \qquad \lambda\mathbf{0}_V = \mathbf{0}_V, \quad \forall\ \lambda \in \mathbb{R}.$$

A titolo di esempio si noti che, con le definizioni date, per ogni $\alpha, \beta \in \mathbb{R}$ $\mathbf{0}_V = (\alpha+\beta)\mathbf{0}_V = \alpha\mathbf{0}_V + \beta\mathbf{0}_V$. Torneremo ancora su questo esempio.

2.3.2 Spazi vettoriali $\mathbb{R}^n$

Consideriamo l'insieme delle coppie ordinate di numeri reali. (Ricordiamo che ordinare le coppie significa che in generale l'elemento (a, b) è diverso dall'elemento (b, a), ad esempio $(2, 3) \neq (3, 2)$.) Tale insieme viene indicato con $\mathbb{R}^2 = \{(\alpha_1, \alpha_2) \mid \alpha_1, \alpha_2 \in \mathbb{R}\}$. Due elementi (α_1, α_2) e (β_1, β_2) di $\mathbb{R}^2$ sono uguali se e solo se $\alpha_1 = \beta_1$ e $\alpha_2 = \beta_2$. Vogliamo definire su $\mathbb{R}^2$ una struttura di spazio vettoriale. Occorre innanzitutto introdurre un'operazione interna che definiamo "componente per componente" nel modo seguente:

$$(\alpha_1, \alpha_2) +_{\mathbb{R}^2} (\beta_1, \beta_2) = (\alpha_1 + \beta_1, \alpha_2 + \beta_2)$$

dove la somma in ogni componente è la somma di numeri reali. L'operazione $+_{\mathbb{R}^2}$ appena definita è commutativa e associativa poiché essa viene definita componente per componente mediante una operazione (la somma di numeri reali) che gode di tali proprietà. L'elemento neutro è $\mathbf{0}_{\mathbb{R}^2} = (0, 0)$. Definiamo ora il prodotto per scalari: se $\lambda \in \mathbb{R}$ e $(\alpha_1, \alpha_2) \in \mathbb{R}^2$ definiamo $\lambda(\alpha_1, \alpha_2) = (\lambda\alpha_1, \lambda\alpha_2)$. Anche in questo caso è chiaro che tutti gli assiomi della definizione di spazio vettoriale sono rispettati.

Analogamente indicheremo con $\mathbb{R}^3$ l'insieme delle terne ordinate di numeri reali e, più in generale, con $\mathbb{R}^n$, con $n \in \mathbb{N}$, l'insieme delle $n-$*uple* ordinate di numeri reali. Generalizzando quanto descritto sopra, definiamo su $\mathbb{R}^n$ una struttura di spazio vettoriale reale definendo la somma ed il prodotto per scalari come segue:

$$(\alpha_1, \alpha_2, \ldots, \alpha_n) +_{\mathbb{R}^n} (\beta_1, \beta_2, \ldots, \beta_n) = (\alpha_1 + \beta_1, \alpha_2 + \beta_2, \ldots, \alpha_n + \beta_n)$$

e

$$\lambda(\alpha_1, \alpha_2, \ldots, \alpha_n) = (\lambda\alpha_1, \lambda\alpha_2, \ldots, \lambda\alpha_n).$$

Notiamo che tra gli spazi vettoriali così definiti vi è anche $\mathbb{R}^1 = \mathbb{R}$: l'insieme dei numeri reali inteso come spazio vettoriale su se stesso. In tal caso denotiamo i vettori di $\mathbb{R}^1$ con (a), $a \in \mathbb{R}$.

2.3.3 Matrici

Fissati $m, n \in \mathbb{R}$ l'insieme $\mathcal{M}_{m,n}(\mathbb{R})$, può essere munito di una struttura di spazio vettoriale. Cominciamo con l'esempio di $\mathcal{M}_{2,3}(\mathbb{R})$, gli altri casi sono del tutto analoghi e si differenziano solo per la forma delle matrici. Definiamo l'operazione interna componente per componente:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} +_{\mathcal{M}_{2,3}(\mathbb{R})} \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \end{pmatrix}$$

dove $c_{11} = a_{11} + b_{11}$ come somma di numeri reali, $c_{12} = a_{12} + b_{12}$ e così via $c_{ij} = a_{ij} + b_{ij}$. L'elemento neutro rispetto alla somma così definita è la matrice nulla

$$\mathbf{0}_{\mathcal{M}_{2,3}(\mathbb{R})} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Definiamo l'operazione esterna sempre componente per componente: dati $\lambda \in \mathbb{R}$ ed $A = (a_{ij}) \in \mathcal{M}_{2,3}(\mathbb{R})$ poniamo

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

Usando ancora una volta le proprietà dei numeri reali si verifica facilmente che le operazioni introdotte definiscono su $\mathcal{M}_{2,3}(\mathbb{R})$ una struttura di spazio vettoriale reale. Si noti che, sia a livello di insiemi che a livello di spazi vettoriali, è possibile identificare $\mathcal{M}_{1,3}(\mathbb{R})$ con $\mathbb{R}^3$ e più in generale $\mathcal{M}_{1,n}(\mathbb{R})$ con $\mathbb{R}^n$, per ogni $n \in \mathbb{N}$. Con una piccola rotazione... della testa... si potrebbe pure identificare $\mathcal{M}_{n,1}(\mathbb{R})$ con $\mathbb{R}^n$, per ogni $n \in \mathbb{N}$, e in effetti questo è vero anche a livello di spazi vettoriali. Preciseremo meglio in seguito che cosa intendiamo.

In generale, date $A = (a_{ij})$ e $B = (b_{ij})$ matrici $m \times n$ definiamo $A+B = C$ dove $C = (c_{ij})$ è la matrice $m \times n$ di termine generico $c_{ij} = a_{ij} + b_{ij}$; inoltre, per ogni numero reale λ, definiamo λA come la matrice $m \times n$ ottenuta moltiplicando ogni entrata di A per λ:

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

In questo modo abbiamo dotato $\mathcal{M}_{m,n}(\mathbb{R})$ di una struttura di $\mathbb{R}$-spazio vettoriale.

2.3.4 Insiemi di polinomi

Sia $\mathbb{R}[X]$ l'insieme dei polinomi nella indeterminata X a coefficienti in $\mathbb{R}$ e indichiamo con $\mathbb{R}^{\leq n}[X]$ i polinomi di $\mathbb{R}[X]$ di grado minore o uguale a n.

Anche su $\mathbb{R}[X]$ si può definire una struttura di $\mathbb{R}$-spazio vettoriale: presi due polinomi $P(X)$ e $Q(X)$, di grado rispettivamente n e m, con $n, m \in \mathbb{N}$ e $n \leq m$: $P(X) = a_nX^n + a_{n-1}X^{n-1} + \ldots + a_1X + a_0$, $Q(X) = b_mX^m + b_{m-1}X^{m-1} + \ldots + b_1X + b_0$, definiamo la loro somma in $\mathbb{R}[X]$ come

$$P(X) +_{\mathbb{R}[X]} Q(X) = \sum_{i=0}^{m} c_iX^i$$

con $c_i = b_i$ se $m \geq i \geq n+1$ e $c_i = b_i + a_i$ per $i \leq n$. Ad esempio, se $Q(X) = -X^2 + 4X - \sqrt{3}$ e $P(X) = 3X^4 + 2X + 6$, $P(X) +_{\mathbb{R}[X]} Q(X) = 3X^4 - X^2 + 6X + (6 - \sqrt{3})$.

La somma di polinomi così definita è commutativa e associativa; l'elemento neutro $\mathbf{0}_{\mathbb{R}[X]}$ per la somma è il polinomio di grado 0 il cui termine noto è zero, cioè il polinomio identicamente nullo. Ogni polinomio ammette opposto per tale somma: se $P(X) = a_nX^n + a_{n-1}X^{n-1} + \ldots + a_1X + a_0$, allora il suo opposto è dato da $-P(X) = -a_nX^n - a_{n-1}X^{n-1} - \ldots - a_1X - a_0$. Sia λ un numero reale e sia $P(X) = a_nX^n + a_{n-1}X^{n-1} + \ldots + a_1X + a_0$ un polinomio, definiamo $\lambda P(X) = (\lambda a_n)X^n + (\lambda a_{n-1})X^{n-1} + \ldots + (\lambda a_1)X + (\lambda a_0)$, cioè moltiplichiamo ogni monomio di $P(X)$ per λ. Si vede facilmente che il prodotto per scalari così definito gode della proprietà distributiva rispetto alla somma, e che $1P(X) = P(X)$. L'insieme $\mathbb{R}[X]$ dotato delle operazioni sopra descritte è quindi uno spazio vettoriale reale. In particolare $(-1)P(X)$ è l'opposto di $P(X)$.

Osservazione. Consideriamo ora il sottoinsieme $\mathbb{R}^{\leq n}[X]$ di $\mathbb{R}[X]$. Le operazioni che abbiamo definito su $\mathbb{R}[X]$ sono definite anche in $\mathbb{R}^{\leq n}[X]$ e osserviamo che la somma di due polinomi di grado minore o uguale a n è ancora un polinomio di grado minore o uguale a n e se $\lambda \in \mathbb{R}$ e $P(X) \in \mathbb{R}^{\leq n}[X]$, allora il polinomio $\lambda P(X)$ appartiene a $\mathbb{R}^{\leq n}[X]$. In altre parole $\mathbb{R}^{\leq n}[X]$ è un sottoinsieme di $\mathbb{R}[X]$ chiuso rispetto alle operazioni di $\mathbb{R}[X]$. Dunque $\mathbb{R}^{\leq n}[X]$ ha, con queste operazioni, una struttura di spazio vettoriale.

Questo è un esempio di sottoinsieme di vettori di uno spazio vettoriale su cui le operazioni dello spazio ambiente inducono una struttura di spazio vettoriale. Un tale sottoinsieme si dice sottospazio vettoriale, ma vedremo in dettaglio questa definizione nella Lezione 3.

Notiamo infine che le nostre conoscenze matematiche ci permettono di definire un'altra operazione all'interno dell'insieme dei polinomi $\mathbb{R}[X]$: la moltiplicazione di due polinomi (i.e. $(3X^2 + 2)(5X - 1) = 15X^3 - 3X^2 + 10X - 2$). Tale operazione non interviene tuttavia nella definizione di spazio vettoriale.

2.3.5 Spazi di funzioni

Costruiamo infine un ultimo esempio di spazio vettoriale. Consideriamo l'insieme delle applicazioni continue dall'intervallo $[0, 1] \subset \mathbb{R}$ verso $\mathbb{R}$ e indichiamo tale insieme con $\mathcal{C}_{\mathbb{R}}([0, 1])$. Muniamo tale insieme di una struttura di spazio vettoriale. Dati f, g due elementi di $\mathcal{C}_{\mathbb{R}}([0, 1])$ definiamo la funzione $h \in \mathcal{C}_{\mathbb{R}}([0, 1])$ somma di f e g

$$f +_{\mathcal{C}_{\mathbb{R}}([0,1])} g = h$$

come l'applicazione da $[0, 1]$ a valori in $\mathbb{R}$ tale che il valore di h in $s \in [0, 1]$ sia dato da $h(s) = f(s) +_{\mathbb{R}} g(s)$, cioè dalla somma in $\mathbb{R}$ dei valori delle due funzioni f e g in s. Essendo la somma una applicazione continua allora anche h è una applicazione continua e quindi appartiene a $\mathcal{C}_{\mathbb{R}}([0, 1])$. Tale somma è commutativa e associativa (prese tre funzioni f_1, f_2, f_3 di $\mathcal{C}_{\mathbb{R}}([0, 1])$, allora

$$(f_1 +_{\mathcal{C}_{\mathbb{R}}([0,1])} f_2) +_{\mathcal{C}_{\mathbb{R}}([0,1])} f_3 = f_1 +_{\mathcal{C}_{\mathbb{R}}([0,1])} (f_2 +_{\mathcal{C}_{\mathbb{R}}([0,1])} f_3)$$

poiché il valore in $s \in [0, 1]$ di ciascuna somma è: $(f_1(s) + f_2(s)) + f_3(s) = f_1(s) + (f_2(s) + f_3(s))$ per l'associatività della somma di numeri reali. L'elemento neutro rispetto alla somma è la funzione $\mathbf{0}_{\mathcal{C}_{\mathbb{R}}([0,1])}$, cioè la funzione identicamente uguale a 0; ogni $f \in \mathcal{C}_{\mathbb{R}}([0,1])$ ammette opposto, cioè esiste $g \in \mathcal{C}_{\mathbb{R}}([0,1])$ tale che

$$f +_{\mathcal{C}_{\mathbb{R}}([0,1])} g = \mathbf{0}_{\mathcal{C}_{\mathbb{R}}([0,1])}.$$

Infatti, essendo $\mathbf{0}_{\mathcal{C}_{\mathbb{R}}([0,1])}$ la funzione identicamente nulla, si ha che l'opposto di una funzione f è la funzione g tale che $g(s) = -f(s)$ per $s \in [0,1]$. Dati $\lambda \in \mathbb{R}$ e $f \in \mathcal{C}_{\mathbb{R}}([0,1])$ definiamo $(\lambda f)(s) = \lambda f(s)$. Tale applicazione è continua, quindi λf è ancora un elemento di $\mathcal{C}_{\mathbb{R}}([0,1])$. Valutando di volta in volta le funzioni nei punti $s \in [0,1]$, si vede che le proprietà di spazio vettoriale sono verificate. Si noti ad esempio che la proprietà distributiva del prodotto per scalari rispetto alla somma è data dalla uguaglianza delle due funzioni continue

$$\lambda(f_1 +_{\mathcal{C}_{\mathbb{R}}([0,1])} f_2) = \lambda f_1 +_{\mathcal{C}_{\mathbb{R}}([0,1])} \lambda f_2,$$

il che si verifica valutando le due funzioni in $s \in [0,1]$:

$$\lambda(f_1(s) + f_2(s)) = \lambda f_1(s) + \lambda f_2(s).$$

In conclusione la definizione astratta di spazio vettoriale è un formalismo che si adatta a molte situazioni in apparenza ben diverse: numeri reali, n-uple, matrici, polinomi ... Quindi non ci interessano necessariamente quali siano le operazioni introdotte per definire una struttura di spazio vettoriale perché il comportamento di ogni spazio vettoriale risulta sempre lo stesso. Ad esempio in 2.3.5 il vettore opposto a f, che sappiamo essere $(-1)f$, è $-f$. Così come in $\mathbb{R}^3$ il vettore opposto a $v = (1,-1,5)$, e formalmente dato da $(-1)v$, si determina attraverso la legge di spazio vettoriale data su $\mathbb{R}^3$: $(-1)(1,-1,5) = (-1,1,-5)$.

Lezione 3

Combinazioni lineari e sottospazi

3.1 Sottospazi vettoriali

In ogni costruzione matematica, dopo aver definito una classe di oggetti aventi delle proprietà omogenee, risulta naturale introdurre la nozione di sottooggetti che rispettino tali proprietà. Da un punto di vista insiemistico, cioè quando gli oggetti sono insiemi, un sottooggetto di un insieme W è un sottoinsieme U, cioè un insieme i cui elementi sono elementi di W: $U \subseteq W$. Vogliamo allora che nella definizione di sottooggetto di uno spazio vettoriale si tenga conto della sua struttura, più ricca di quella di un semplice insieme: dovremo dunque fare in modo che ogni sottooggetto erediti questa struttura.

Definizione 3.1.1 *Dato un $\mathbb{R}$-spazio vettoriale V, un sottospazio vettoriale U di V è un sottoinsieme non vuoto $U \subseteq V$ su cui le operazioni di V inducono una struttura di spazio vettoriale reale.*

Che cosa vuol dire esattamente questa definizione? Sia V uno spazio vettoriale reale e sia U un suo sottoinsieme non vuoto. Dati due vettori u e v di U possiamo considerare $u +_V v$ dal momento che $U \subseteq V$ e quindi gli elementi di U sono, in particolare, vettori di V. In generale il vettore somma $u +_V v$ NON sarà un elemento di U. Dire che le operazioni di V inducono una struttura di spazio vettoriale su U significa che la somma in V ed il prodotto per scalari:

$$+_V : V \times V \longrightarrow V \qquad \cdot : \mathbb{R} \times V \longrightarrow V$$

si restringono ad operazioni su U:

$$+_V : U \times U \longrightarrow U \qquad \cdot : \mathbb{R} \times U \longrightarrow U.$$

In particolare il vettore $u +_V v$ appartiene ancora ad U. In questo caso si usa dire che U è *chiuso* rispetto alla somma di V. Analogamente, dati $\lambda \in \mathbb{R}$ e $v \in U$, il vettore λv è un elemento di V ma, in generale, se U è solo un sottoinsieme di V, λu NON è un elemento di U. Se, però, U è un sottospazio vettoriale di V allora λu appartiene ancora ad U. Diremo in questo caso che U è *chiuso* rispetto al prodotto per scalari di V. La Definizione 3.1.1 è dunque equivalente alla seguente:

Definizione 3.1.2 *Dato un $\mathbb{R}$-spazio vettoriale V, un sottospazio vettoriale U di V è un sottoinsieme non vuoto $U \subseteq V$ chiuso rispetto alla somma e al prodotto per scalari definiti in V, cioè un sottoinsieme tale che:*

1. *per ogni $u, v \in U$ la somma $u +_V v \in U$;*

2. *per ogni elemento $u \in U$ e per ogni $\lambda \in \mathbb{R}$, $\lambda u \in U$.*

Dato un $\mathbb{R}$-spazio vettoriale V per indicare che un sottoinsieme U di V è un suo sottospazio scriveremo $U \leq V$.

Osservazione 3.1.3 Ogni spazio vettoriale V ha sempre almeno due sottospazi: V stesso e il sottospazio vettoriale banale $\{\mathbf{0}_V\}$. Chiaramente ogni sottospazio vettoriale contiene il vettore nullo $\mathbf{0}_V$.

Vediamo ora in alcuni esempi come applicare la Definizione 3.1.2.

Esempi 3.1.4 Supponiamo che il nostro spazio ambiente sia $V = \mathbb{R}^2$:

1. Consideriamo l'insieme costituito dal solo vettore nullo: $\{(0, 0)\}$. Esso è un sottospazio vettoriale di $\mathbb{R}^2$: il sottospazio banale.

2. Sia $U = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$. L'insieme U non può essere un sottospazio vettoriale poiché non contiene il vettore nullo di $\mathbb{R}^2$ che è $(0, 0)$.

3. Cerchiamo di sopperire a questo fatto considerando ora $U = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$, che è il cerchio di centro l'origine e raggio 1. Questa volta $(0, 0) \in U$, ma U non è chiuso rispetto a $+_{\mathbb{R}^2}$: infatti, presi, ad esempio, $(1, 0), (0, 1) \in U$, la loro somma $(1, 0) +_{\mathbb{R}^2} (0, 1) = (1, 1)$ non è un elemento del cerchio, quindi U non è un sottospazio di $\mathbb{R}^2$.

4. Prendiamo adesso $U = \{(x, y) \in \mathbb{R}^2 \mid x \geq 0,\ y \geq 0\}$, cioè l'insieme dei punti del primo quadrante. In questo caso la somma di due vettori di U è ancora nel primo quadrante, l'elemento neutro $(0, 0)$ ci sta, ma se consideriamo, ad esempio, $\lambda = -3$ e $u = (2, 9)$, allora $\lambda u = (-6, -27)$ non appartiene ad U che quindi non è un sottospazio.

5. Si consideri $U = \{(x, y) \in \mathbb{R}^2 \mid x - 2y = 0\}$. Verifichiamo che U è un sottospazio vettoriale di $\mathbb{R}^2$. Siano $v_1 = (x_1, y_1)$ e $v_2 = (x_2, y_2)$ due elementi di U, tali che, cioè, $x_1 - 2y_1 = 0$ e $x_2 - 2y_2 = 0$. La somma dei due vettori: $v_1 + v_2 = (x_1 + x_2, y_1 + y_2)$ appartiene ad U, infatti $x_1 + x_2 - 2(y_1 + y_2) = (x_1 - 2y_1) + (x_2 - 2y_2) = 0 + 0 = 0$. Analogamente, per ogni numero reale λ, $\lambda v_1 = (\lambda x_1, \lambda y_1)$ è un elemento di U, dal momento che $\lambda x_1 - 2\lambda y_1 = \lambda(x_1 - 2y_1) = \lambda 0 = 0$. Usando la Definizione 3.1.2 concludiamo che U è un sottospazio vettoriale di $\mathbb{R}^2$.

6. Sia $U = \{(x, y) \in \mathbb{R}^2 \mid yx = 0\}$ cioè l'insieme delle coppie (x, y) di numeri reali con almeno una componente uguale a 0. Osserviamo che l'insieme U non è chiuso rispetto alla somma, infatti i vettori $(1, 0)$ e $(0, 1)$ appartengono ad U ma non la loro somma: $(1, 1)$. Dunque U non è un sottospazio vettoriale di $\mathbb{R}^2$.

7. Si prenda $U = \{(x, y) \in \mathbb{R}^2 \mid y + 3x - 1 = 0\}$. Allora U non è un sottospazio vettoriale di $\mathbb{R}^2$ dal momento che il vettore nullo di $\mathbb{R}^2$, $(0, 0)$, non appartiene ad U.

Osservazione sugli esempi fatti
Per stabilire se un sottoinsieme S di uno spazio vettoriale V è un sottospazio vettoriale di V possono essere utilizzati due approcci:

1°. Se pensiamo che l'insieme considerato non sia un sottospazio vettoriale basterà portare un esempio di vettori che non soddisfino le condizioni della Definizione 3.1.2 come fatto negli esempi 2, 3, 4, 6, 7.

2°. Se invece pensiamo che l'insieme considerato sia un sottospazio vettoriale allora si procederà come nell'esempio 5 verificando la validità delle condizioni della Definizione 3.1.2.

Nell'incertezza, ovvero quando non abbiamo a priori un'idea sul fatto che l'insieme in oggetto sia o meno un sottospazio vettoriale, si procede sempre come in 2°. Se l'insieme non è un sottospazio si arriva ad un ostacolo il che consente di trovare un controesempio. Ad esempio, consideriamo l'insieme dell'esempio 3. Procedendo come in 5, una delle condizioni da verificare è che per ogni $v = (x, y)$ elemento di $U = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$ e $\forall \lambda \in \mathbb{R}$ si ha $\lambda v \in U$. Ora da $v \in U$ sappiamo che $x^2+y^2 \leq 1$ quindi $\lambda v = (\lambda x, \lambda y)$ e $(\lambda x)^2+(\lambda y)^2 = \lambda^2(x^2+y^2) \leq \lambda^2$ ma $\lambda v \in U$ se e solo se $\lambda^2(x^2+y^2) \leq 1$. Ora è evidente che se v è tale che $x^2+y^2 = 1$, ad esempio $v = (1, 0)$, per λ "grandi" $\lambda v \notin U$. Ad esempio $v = (1, 0) \in U$, ma preso $\lambda = 100$, $100v = (100, 0) \notin U$ perché $100^2 > 1$.

3.2 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 per ogni indice $i = 1, 2, \ldots, n$ il vettore $\lambda_i v_i \in V$ e poi fare la somma in V dei vettori ottenuti:

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

(si noti che, per la proprietà commutativa della somma, l'ordine in cui vengono sommati gli elementi non altera il risultato finale).

Definizione 3.2.1 *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 3.2.2 Si consideri lo spazio vettoriale $\mathbb{R}^2$.

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

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

Che cosa succederebbe se tutti i vettori di V si potessero scrivere come combinazione lineare di un numero finito di vettori come nell'esempio 2 di 3.2.2?

Definizione 3.2.3 *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 3.2.4 Prendiamo la nostra vecchia conoscenza $\mathbb{R}^3$ e tre suoi vettori $u_1 = (1, 1, -1)$, $u_2 = (0, 1, 2)$, $u_3 = (0, -2, 3)$. Vogliamo dimostrare che $\mathbb{R}^3$ è finitamente generato mostrando 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

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

Esplicitiamo 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 $\mathbb{R}^3$. Si noti che anche l'insieme $\{(2, 1, 0), (0, 1, 0), (0, 0, 1)\}$ è un insieme di generatori di $\mathbb{R}^3$ come pure $\{(1, 1, 0), (0, 1, 1), (1, 0, 1), (2, 1, 1)\}$. Quindi gli insiemi di generatori non sono unici.

Esempio 3.2.5 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

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

Osservazione 3.2.6 Sia $\mathbf{u} = \{u_1, u_2, \ldots, u_s\}$ un insieme di generatori di un $\mathbb{R}$-spazio vettoriale V. Allora ogni vettore v di V si scrive come loro combinazione lineare:

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

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 3.2.7 Osserviamo che, dato un $\mathbb{R}$-spazio vettoriale V generato dall'insieme di vettori $\{u_1, u_2, \ldots, u_s\}$ e preso qualsiasi vettore $u \in V$, l'insieme di vettori $\{u_1, u_2, \ldots, u_s, u\}$ è pure un insieme di generatori di V. Ovviamente si possono continuare ad aggiungere vettori ottenendo sempre un insieme di generatori.

Definizione 3.2.8 *Siano V un $\mathbb{R}$-spazio vettoriale e $\{u_1, u_2, \ldots, u_k\}$ un insieme di vettori di V. Il più piccolo sottospazio di V contenente $u_1, u_2, \ldots, u_k$ si dirà il sottospazio da essi generato e si indicherà con $\langle u_1, u_2, \ldots, u_k\rangle$. Tale sottospazio è costituito da tutte le combinazioni lineari di $u_1, u_2, \ldots, u_k$:*

$$\langle u_1, u_2, \ldots, u_k\rangle = \left\{ \sum_{i=1}^{k} \alpha_i u_i = \alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_k u_k \mid \alpha_i \in \mathbb{R}, i = 1, \ldots k \right\}$$

Il contenuto di questa definizione va giustificato. Prima di tutto verificheremo che l'insieme formato da tutte le combinazioni lineari di $u_1, u_2, \ldots, u_k$ è un sottospazio vettoriale di V; poi vedremo che esso è contenuto in ogni sottospazio contenente $u_1, u_2, \ldots, u_k$, e che quindi è il più piccolo sottospazio soddisfacente questa proprietà. 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 che è pertanto un sottospazio vettoriale di V. D'altro canto se consideriamo un sottospazio $T \leq V$ contenente i vettori $u_1, u_2, \ldots, u_k$, per definizione di sottospazio, T contiene ogni loro combinazione lineare e quindi tutto $\langle u_1, u_2, \ldots, u_k \rangle$.

Osserviamo che, vista la definizione 3.2.8, è chiaro che uno spazio vettoriale V è finitamente generato se e solo se esistono $v_1, \ldots, v_n$ vettori di V tali che $V = \langle v_1, \ldots, v_n \rangle$.

Se invece di partire da un numero finito di vettori $\{u_1, u_2, \ldots, u_k\}$ di uno spazio vettoriale V partissimo da un generico sottoinsieme $S \subset V$, potremmo chiederci qual è il più piccolo sottospazio vettoriale di V che contiene S?

Definizione 3.2.9 *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 $\mathbb{R}$.*

Esempio 3.2.10 Sia V un $\mathbb{R}$-spazio vettoriale, allora $\emptyset \subset V$ e $\langle \emptyset \rangle = \{\mathbf{0}_V\}$. Ma anche $\{\mathbf{0}_V\} \subset V$ e $\langle \mathbf{0}_V \rangle = \{\mathbf{0}_V\}$.

Esempio 3.2.11 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.

3.3 Esercizi svolti

Esercizio 3.3.1 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 3.3.2 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 3.3.3 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 3.3.4 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 3.3.5 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 3.3.6 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 3.3.7 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 3.1.2. 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.

3.4 Esercizi proposti

Esercizio 3.4.1 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 3.4.2 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 3.4.3 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 3.4.4 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 4

Basi e dimensione

4.1 Dipendenza e indipendenza lineare

Nell'Osservazione 3.2.7 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 per il vettore nullo $\mathbf{0}_V$.

Definizione 4.1.1 *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 4.1.2 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.

4. 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 4.1.3 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 4.1.4 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 4.1.5 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 4.1.1 risponde al quesito introdotto all'inizio della lezione, infatti vale la seguente proposizione.

Proposizione 4.1.6 *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:

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

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

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

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

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

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): $\mathbf{0}_V = 0u_1 + 0u_2 + \cdots + 0u_k$ 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.**

4.2 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 4.2.1 *In un $\mathbb{R}$-spazio vettoriale V un insieme (finito) ordinato di vettori $\{b_1, b_2, \ldots, b_s\}$ 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 4.2.2 *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. Sia

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

una relazione di dipendenza tra i vettori $u_1, u_2, \ldots, u_k$, cioè una scrittura del vettore nullo in cui non tutti i coefficienti sono nulli e supponiamo, per fissare le idee, $\lambda_i \neq 0$. Allora sommando il vettore $-\lambda_i u_i$ ad ambedue i membri della precedente relazione si ha:

$$\lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_{i-1} u_{i-1} + \lambda_{i+1} u_{i+1} \cdots + \lambda_k u_k = -\lambda_i u_i$$

e, poiché $\lambda_i \neq 0$, possiamo moltiplicare entrambi i membri per $-\frac{1}{\lambda_i}$ ottenendo

$$u_i = -\frac{\lambda_1}{\lambda_i} u_1 - \frac{\lambda_2}{\lambda_i} u_2 - \cdots - \frac{\lambda_{i-1}}{\lambda_i} u_{i-1} - \frac{\lambda_{i+1}}{\lambda_i} u_{i+1} - \cdots - \frac{\lambda_k}{\lambda_i} u_k.$$

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:

$$v = \beta_1 u_1 + \beta_2 u_2 + \cdots + \beta_{i-1} u_{i-1} + \beta_i u_i + \beta_{i+1} u_{i+1} \cdots + \beta_k u_k$$

e sostituendo ad u_i l'espressione precedente otteniamo:

$$\begin{aligned} v &= \beta_1 u_1 + \beta_2 u_2 + \cdots + \beta_{i-1} u_{i-1} + \beta_{i+1} u_{i+1} \cdots + \beta_k u_k + \\ &+ \beta_i \Big(-\frac{\lambda_1}{\lambda_i} u_1 - \frac{\lambda_2}{\lambda_i} u_2 - \cdots - \frac{\lambda_{i-1}}{\lambda_i} u_{i-1} - \frac{\lambda_{i+1}}{\lambda_i} u_{i+1} \cdots - \frac{\lambda_k}{\lambda_i} u_k \Big) \end{aligned}$$

che è una combinazione lineare dei vettori $u_1, \dots, u_{i-1}, u_{i+1}, \dots u_k$, infatti:

$$\begin{aligned} v &= \left(\beta_1 - \beta_i \frac{\lambda_1}{\lambda_i}\right)u_1 + \left(\beta_2 - \beta_i \frac{\lambda_2}{\lambda_i}\right)u_2 + \cdots + \left(\beta_{i-1} - \beta_i \frac{\lambda_{i-1}}{\lambda_i}\right)u_{i-1} + \\ &+ \left(\beta_{i+1} - \beta_i \frac{\lambda_{i+1}}{\lambda_i}\right)u_{i+1} + \cdots + \left(\beta_k - \beta_i \frac{\lambda_k}{\lambda_i}\right)u_k \end{aligned}$$

e quindi $u_1, \dots, u_{i-1}, u_{i+1}, \dots u_k$ sono un insieme di generatori di V. **C.V.D.**

Abbiamo dunque dimostrato che, dato un insieme di generatori, se questi sono linearmente dipendenti possiamo eliminarne alcuni ed ottenere ancora un insieme di generatori. Ma se l'insieme di generatori I fosse un insieme di vettori linearmente indipendenti? Quanto detto prima non sarebbe più vero. Infatti se in un $\mathbb{R}$-spazio vettoriale V i vettori $v_1, v_2, \dots, v_k$ sono linearmente indipendenti esistono dei vettori di v che sono combinazione lineare di $v_1, v_2, \dots, v_{k-1}, v_k$ ma non di $v_1, v_2, \dots, v_{k-1}$. Ad esempio v_k non è combinazione lineare di $v_1, v_2, \dots, v_{k-1}$ (ma bensì di $v_1, v_2, \dots, v_{k-1}, v_k$, infatti $v_k = 0v_1 + 0v_2 + \cdots + 0v_{k-1} + (1)v_k$). Concludendo, un insieme di generatori può essere raffinato se tali vettori sono linearmente dipendenti. Ne deduciamo che la costruzione di una base di uno spazio vettoriale V è legata all'individuazione del minimo numero di generatori di V.

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

Dimostrazione. "$\Rightarrow$" Supponiamo che $u_1, u_2, \dots, u_k, u$ siano linearmente indipendenti e che, per assurdo, u appartenga a $\langle u_1, u_2, \dots, u_k \rangle$, allora $u = \alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_k u_k$ con $\alpha_i \in \mathbb{R}$. Sommando $-u$ ai due membri dell'uguaglianza otteniamo $\mathbf{0}_V = \alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_k u_k + (-1)u$ e, poiché il coefficiente di u è diverso da zero, i vettori $u_1, u_2, \dots, u_k, u$ sono linearmente dipendenti, contro l'ipotesi.

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

$$\alpha_1 u_1 + \alpha_2 u_2 + \dots + \alpha_k u_k + \alpha u = \mathbf{0}_V \tag{4.1}$$

con qualcuno dei numeri reali $\alpha_1, \alpha_2, \dots, \alpha_k, \alpha$ diverso da zero. Ora α non può essere uguale a 0, poiché allora avremmo una relazione di dipendenza tra $u_1, u_2, \dots, u_k$ che sono linearmente indipendenti, quindi $\alpha \neq 0$ e $\alpha u = -\alpha_1 u_1 - \alpha_2 u_2 - \dots - \alpha_k u_k$. Moltiplicando tale uguaglianza per $1/\alpha$, si ottiene u come combinazione lineare di $u_1, u_2, \dots, u_k$, 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. Riassumiamo quanto detto nel seguente corollario:

Corollario 4.2.4 *Dati uno spazio vettoriale V ed un insieme B di vettori di V, i fatti seguenti sono equivalenti:*

1. *B è una base di V;*

2. *B è un insieme massimale di vettori linearmente indipendenti;*

3. *B è un insieme minimale di generatori di V ;*

(qui massimale e minimale si intendono rispetto all'ordine dato dalle inclusioni come sottoinsiemi di V).

Siamo al risultato culmine della teoria: ogni spazio vettoriale finitamente generato V possiede sempre una base ed ogni base di V ha lo stesso numero di elementi, che chiameremo *dimensione* dello spazio vettoriale V.

Teorema 4.2.5 **(Senza dimostrazione)** *In un $\mathbb{R}$-spazio vettoriale V sia $I = \{u_1, u_2, \dots, u_k\}$ un insieme di vettori linearmente indipendenti e sia $G = \{w_1, w_2, \dots, w_p\}$ 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 4.2.6 *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 4.2.7 *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 4.2.8 *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 loro stessi una base; se invece sono linearmente dipendenti per la Proposizione 4.2.2 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 4.2.3 $u_1, \ldots, u_s, u$ 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.

Esempi 4.2.9

1. Consideriamo in $\mathbb{R}^2$ i vettori $e_1 = (1, 0)$ ed $e_2 = (0, 1)$. Essi generano ovviamente $\mathbb{R}^2$, infatti 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 $\mathbb{R}^n$: $e_1 = (1, 0, \ldots, 0)$, $e_2 = (0, 1, 0, \ldots, 0)$, $\ldots, e_n = (0, 0, \ldots, 0, 1)$. Allora $\{e_1, e_2, \ldots, e_n\}$ è una base di $\mathbb{R}^n$, detta base canonica. La dimensione di $\mathbb{R}^n$ è dunque n. Dato il vettore $(x_1, x_2, \ldots, x_n)$ di $\mathbb{R}^n$ vale: $(x_1, x_2, \ldots, x_n) = x_1 e_1 + \cdots + x_n e_n$.

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 4.2.10 Sia V uno spazio vettoriale finitamente generato. V ammette quindi una base e supponiamo che V abbia dimensione $n \geq 1$. Sia $T \leq V$ 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, ma anche in V, poiché le operazioni in T sono indotte dalle operazioni di 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.

Definizione 4.2.11 *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 $(\sigma_1, \dots, \sigma_n) \in \mathbb{R}^n$ e posto $w = \sigma_1 u_1 + \sigma_2 u_2 + \cdots + \sigma_n u_n$, w è un elemento di V e $f(w) = (\sigma_1, \dots, \sigma_n)$. L'applicazione f è dunque biunivoca.

Esempio 4.2.12 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 4.2.13 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 $\{e_{ij}, i = 1, \ldots, n, j = 1, \ldots, m\}$ dove e_{ij} è la matrice avente tutte le entrate nulle tranne quella di posto i, j che è uguale a 1.

4.3 Strumenti di calcolo

Cocentriamoci ora sulla seguente domanda: dati k vettori $v_1, \ldots, v_k$ 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 $v_1, \ldots, v_k$ con qualche coefficiente diverso da 0. Abbiamo utilizzato questo metodo nell'Esempio 4.1.2(2.). Vediamo un altro esempio:

Esempio 4.3.1 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 4.1.5 che se i vettori $v_1,\dots,v_k$ sono linearmente indipendenti allora qualsiasi sottoinsieme di $\{v_1,\dots,v_k\}$ continua ad essere formato da vettori linearmente indipendenti. Per stabilire dunque se certi vettori assegnati $v_1,\dots,v_k$ 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 $\{v_1, \dots, v_i\}$ per $i = 1, \dots, k$.

Esempio 4.3.2 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 $v_1, \dots, v_4$. 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 4.3.3 *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 4.3.4 *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 4.3.5 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 4.3.6 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 4.3.7 *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 $a v_j + b(v_i + \alpha v_j) = (a + b\alpha) v_j + b v_i$, 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 $h v_i + k v_j \in \langle v_j, v_i \rangle$ con $h, k \in \mathbb{R}$, possiamo scrivere $h v_i + k v_j = h(v_i + \alpha v_j) + (k - h\alpha) v_j$ 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 $\langle v_j, v_i \rangle = \langle v_j, v_i + \alpha v_j \rangle$. □

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 4.3.7, 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 4.3.7, 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 4.3.8 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 $v_1, \dots, v_4$ sono linearmente indipendenti. Perciò i vettori $v_1, \dots, v_4$ sono linearmente dipendenti. Sappiamo, inoltre, che $\langle v_1, \dots, v_4 \rangle = \langle (1, 2, 0, 1), (0, 3, 1, 2) \rangle$.

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

Esempio 4.3.9 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.

4.4 Esercizi svolti

Esercizio 4.4.1 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 $\mathbb{R}^3$.

Per mostrare la lineare indipendenza dobbiamo verificare che, presa una combinazione lineare dei vettori $(1,2,3)$, $(0,1,0)$, $(0,0,1)$, questa è uguale al vettore nullo se e solo se tutti i coefficienti sono nulli. Sia, dunque, $\alpha(1,2,3)+\beta(0,1,0)+\gamma(0,0,1)=(\alpha, 2\alpha+\beta, 3\alpha+\gamma)=(0,0,0)$, cioè

$$\begin{cases} \alpha = 0 \\ 2\alpha+\beta=0 \\ 3\alpha+\gamma=0 \end{cases}.$$

Risolvendo il sistema otteniamo: $\alpha=\beta=\gamma=0$. Dunque i vettori $\{(1,2,3),\ (0,1,0),\ (0,0,1)\}$ sono linearmente indipendenti. Mostriamo ora che essi generano $\mathbb{R}^3$. Sia (a,b,c) un qualunque vettore di $\mathbb{R}^3$ e cerchiamo $\alpha,\beta,\gamma \in \mathbb{R}$ tali che $(a,b,c) = \alpha(1,2,3)+\beta(0,1,0)+\gamma(0,0,1) = (\alpha, 2\alpha+\beta, 3\alpha+\gamma)$. Otteniamo: $\alpha=a$, $\beta=b-2a$ $\gamma=c-3a$.

Esercizio 4.4.2 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 4.2.3 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 4.4.3 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 3.3.1. Infine si può mostrare che l'insieme $\{(0,1,1), (1,0,-1), (2,1,0)\}$ è una base di $\mathbb{R}^3$.

Esercizio 4.4.4 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 4.4.5 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 4.4.6 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 4.4.7 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 $v = (0, -2, 1)_{\mathcal{B}}$.

4.5 Esercizi proposti

Esercizio 4.5.1 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 4.5.2 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 4.5.3 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 4.5.4 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 4.5.5 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 5

Applicazioni lineari e matrici

5.1 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 preservano le operazioni definite sul dominio.

Definizione 5.1.1 *Una applicazione $L : V \to W$ tra due $\mathbb{R}$-spazi vettoriali V, W si dice lineare se*

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

Questa è la definizione di applicazione lineare. Si noti che dalla definizione segue subito che, per qualunque vettore v di V, $L(\mathbf{0}_V) = L(0v) = 0L(v) = \mathbf{0}_W$. Si noti anche che l'opposto di un vettore v viene mandato nell'opposto del vettore $L(v)$: $L((-1)v) = (-1)L(v)$ che è l'opposto in W del vettore $L(v)$.

Ricordiamo che, data una applicazione qualsiasi f tra due insiemi X, Y, $f : X \to Y$ (quindi X è il dominio dell'applicazione f e Y il suo codominio), per ogni sottoinsieme X_1 di X possiamo definire un sottoinsieme di Y ponendo:

$$f(X_1) = \{y \in Y \mid \exists\, x \in X_1 \text{ tale che } f(x) = y\}.$$

Tale insieme si dirà *insieme immagine di X_1 tramite f* . Analogamente se Y_1 è un sottoinsieme di Y possiamo definire il seguente sottoinsieme di X:

$$f^{-1}(Y_1) = \{x \in X \mid f(x) \in Y_1\}$$

(attenzione: si tratta solo di una notazione, NON stiamo richiedendo che la funzione f sia invertibile); $f^{-1}(Y_1)$ si dirà *immagine inversa (o controimmagine o antiimmagine) di* Y_1 *tramite* f.

Proposizione 5.1.2 *Data una applicazione lineare* $L : V \to W$ *tra due* $\mathbb{R}$-*spazi vettoriali* V, W *se* $V_1 \leq V$ *è un sottospazio vettoriale di* V *allora la sua immagine tramite* L, $L(V_1)$, *è un sottospazio vettoriale di* W. *Se* $W_1 \leq W$ *è un sottospazio vettoriale di* W *allora la sua controimmagine tramite* L, $L^{-1}(W_1)$, *è un sottospazio di* V.

Dimostrazione. Dimostriamo la prima affermazione: $L(V_1) = \{w \in W \mid \exists\, v \in V_1 \text{ tale che } L(v) = w\}$ è un sottospazio di W. Per fare questo dobbiamo usare la Definizione 3.1.2. Cioè, presi w_1 e w_2 in $L(V_1)$ (che è in ogni caso un insieme di vettori di W), dobbiamo vedere innanzitutto che $w_1 +_W w_2 \in L(V_1)$. Per definizione $w = w_1 +_W w_2 \in L(V_1)$ se esiste un vettore $v \in V_1$ che ha w come immagine mediante L, cioè: $L(v) = w$. Ora: $w_1, w_2 \in L(V_1)$ quindi esistono $v_1, v_2 \in V_1$ tali che $L(v_1) = w_1$ e $L(v_2) = w_2$, pertanto $w_1 +_W w_2 = L(v_1) +_W L(v_2)$; per la linearità di L si ha che $L(v_1) +_W L(v_2) = L(v_1 +_V v_2)$, ma V_1 è un sottospazio di V e quindi $v_1 +_V v_2 \in V_1$ e allora $w_1 +_W w_2 = L(v_1 +_V v_2)$ sta in $L(V_1)$ poiché è l'immagine di $v_1 +_V v_2 \in V_1$. Analogamente, se $\alpha \in \mathbb{R}$ e $w \in L(V_1)$, allora esiste $v \in V_1$ tale che $L(v) = w$, quindi $\alpha w = \alpha L(v)$ ed essendo L lineare $\alpha w = L(\alpha v)$, ma V_1 è un sottospazio di V, pertanto $\alpha v \in V_1$. Dunque αw è immagine del vettore $\alpha v \in V_1$. Quindi $L(V_1)$ è un sottospazio vettoriale di W.

Dimostriamo la seconda affermazione: siano allora $v_1, v_2 \in L^{-1}(W_1)$, vale a dire $L(v_1) \in W_1$ e $L(v_2) \in W_1$; $v_1 +_V v_2$ appartiene ancora a $L^{-1}(W_1)$? Cioè $L(v_1 +_V v_2) \in W_1$? In effetti L è lineare, quindi $L(v_1 +_V v_2) = L(v_1) +_W L(v_2)$, ma sia $L(v_1)$ che $L(v_2)$ stanno in W_1 ed essendo W_1 un sottospazio, anche la loro somma appartiene a W_1. Per il prodotto per scalari: se $\alpha \in \mathbb{R}$ e $v \in L^{-1}(W_1)$ allora $\alpha v \in L^{-1}(W_1)$ se $L(\alpha v) \in W_1$, ma, ancora, L è lineare e dunque: $L(\alpha v) = \alpha L(v)$. Dal momento che $L(v) \in W_1$ per ipotesi e W_1 è un sottospazio, $\alpha L(v) \in W_1$. **C.V.D.**

Alcuni dei sottospazi costruiti nella proposizione precedente ci interesseranno più degli altri. In particolare l'insieme immagine mediante L del dominio V di L si chiama **immagine di L** e si indica con $\mathrm{Im}L$: $L(V) = \mathrm{Im}L$. Osserviamo che l'applicazione L è suriettiva se e solo se $\mathrm{Im}L = W$.

Consideriamo la controimmagine del vettore nullo di W: $L^{-1}(\{\mathbf{0}_W\})$. Questo è il sottospazio degli elementi di V che sono mandati mediante L nel vettore nullo di W. Tale sottospazio di V si chiama **nucleo** dell'applicazione lineare L e si indica con Ker$L$:

$$\mathrm{Ker}L = \{v \in V \mid L(v) = \mathbf{0}_W\}.$$

Osserviamo che Ker$L$ contiene sempre $\mathbf{0}_V$.

La Proposizione 5.1.2 afferma, in particolare, che il nucleo e l'immagine di una applicazione lineare L sono sottospazi vettoriali rispettivamente del dominio e del codominio di L.

Abbiamo visto il legame tra l'immagine di una applicazione lineare e la sua suriettività. Che cosa possiamo dire sulla iniettività? Sappiamo 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 5.1.3 *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 che sono mandati mediante L nel 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.**

Riflessione. Una definizione equivalente di applicazione iniettiva $f : X \to Y$ consiste nel dire che l'antiimmagine di ogni elemento di Imf è costituita da uno ed un solo elemento del dominio. Data un'applicazione lineare $L : V \to W$, il vettore $\mathbf{0}_W$ appartiene sempre all'immagine di L e la proposizione precedente mostra che L è iniettiva se e solo se l'antiimmagine del vettore nullo $\mathbf{0}_W$ è costituita solamente da $\mathbf{0}_V$. Per verificare l'iniettività di un'applicazione lineare L dunque non occorre considerare l'antiimagine di ogni vettore dell'immagine ma solo di $\mathbf{0}_W$.

Quest'ultima proposizione ci aiuta anche a risolvere il seguente esercizio:

Esercizio 5.1.4 *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:

$$L^{-1}(\{w\}) = v + \mathrm{Ker}L. \tag{5.1}$$

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 (5.1) 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 5.1.5 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 $\{v_1, v_2, \ldots, v_n\}$ una sua base. Consideriamo un altro spazio vettoriale W e n vettori qualsiasi $w_1, w_2, \ldots, w_n$ di W (i vettori $w_1, w_2, \ldots, w_n$ 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

$$f(v) = \lambda_1 w_1 + \lambda_2 w_2 + \cdots + \lambda_n w_n. \tag{5.2}$$

La condizione (5.2) 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.

5.2 Struttura dimensionale

Per il momento abbiamo solo informazioni qualitative riguardo ad una applicazione lineare. Adesso vogliamo dare qualche informazione 'quantitativa'. Supponiamo quindi di avere una applicazione lineare $L : V \to W$ e supponiamo inoltre 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 ImL. Infatti, sia $w \in$ ImL; 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 ImL. 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 5.2.1 (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 4.2.8): $\{t_1, \ldots, t_k, t_{k+1}, \ldots, t_n\}$. I vettori immagine: $L(t_1), \ldots, L(t_k), L(t_{k+1}), \ldots, L(t_n)$ 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 $\beta_1 t_{k+1} + \cdots + \beta_{n-k} t_n \in$ 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 5.2.2 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 ($\mathrm{Im}L = V$) allora $\dim(\mathrm{Im}L) = \dim(V)$ il che implica, per il Teorema 5.2.1, $\dim(\mathrm{Ker}L) = 0$ cioè il nucleo è il sottospazio banale, quindi la funzione è iniettiva. Viceversa, se l'endomorfismo è iniettivo, $\mathrm{Ker}L = \{\mathbf{0}_V\}$ e dunque la sua dimensione è zero. Allora $\dim V = \dim \mathrm{Im}L$ e quindi $\mathrm{Im}L$ 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 5.2.1, $n = \dim \mathrm{Ker}(L) + \dim \mathrm{Im}(L)$, 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 $\mathrm{Im}L = W$, poiché la funzione è iniettiva $\dim(\mathrm{Ker}L) = 0$ 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.

5.3 Applicazioni lineari, basi e matrici

Abbiamo visto che, dato uno spazio vettoriale V e fissata una sua base $\mathbf{v} = \{v_1, v_2, \ldots, v_n\}$, 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 $(\alpha_1, \alpha_2, \ldots, \alpha_n)_{\mathbf{v}}$. 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 $\mathbf{v} = \{v_1 = (1,2), v_2 = (1,1)\}$ e $\mathbf{w} = \{w_1 = (1,0), w_2 = (1,5)\}$, il vettore di coordinate $(1,1)_{\mathbf{v}}$ nella base $\mathbf{v}$ è il vettore $1v_1 + 1v_2 = 1(1,2) + 1(1,1) = (2,3)$, mentre $(1,1)_{\mathbf{w}}$ 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 e siano $\dim V = n$ e $\dim W = m$; scegliamo poi una base per ciascuno spazio: $\mathbf{v} = \{v_1, v_2, \ldots, v_n\}$ e $\mathbf{w} = \{w_1, w_2, \ldots, w_m\}$. Abbiamo visto in 5.1.5.iv) (e lo ricordiamo) che conoscere i valori di L sui vettori di una base significa conoscere l'applicazione lineare interamente. Infatti, supponiamo di sapere chi siano $L(v_1), L(v_2), \ldots, L(v_n) \in W$ e di voler calcolare l'immagine di ogni vettore $v \in V$. Poiché i vettori $v_1, v_2, \ldots, v_n$ sono una base di V, v si scrive in modo unico come loro combinazione lineare: $v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n$ e, applicando L, si ha: $L(v) = L(\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n) = \lambda_1 L(v_1) + \lambda_2 L(v_2) + \cdots + \lambda_n L(v_n)$. Dunque conoscendo $L(v_1), L(v_2), \ldots, L(v_n)$ e sfruttando l'unicità della scrittura di v nella base $v_1, v_2, \ldots, v_n$, si conosce la funzione lineare L completamente. 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:

$$L(v_1) \quad = \quad a_{11} w_1 + a_{21} w_2 + \cdots + a_{m1} w_m$$

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

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 $\mathbf{v}$ e $\mathbf{w}$**. Osserviamo che, una volta fissate le basi $\mathbf{v}$ e $\mathbf{w}$, le colonne di A rappresentano le coordinate rispetto alla base $\mathbf{w}$ dei vettori $L(v_1), \ldots, L(v_n)$. A questo punto 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 $\{w_1, w_2, \ldots, w_m\}$, come segue:

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

(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 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 5.3.1 *i*) Consideriamo la base $\mathbf{v} = \{v_1 = (1,1,0),\ v_2 = (0,1,0),\ v_3 = (2,0,1)\}$ di $\mathbb{R}^3$ e la base $\mathbf{w} = \{w_1 = (1,0), w_2 = (1,1)\}$ 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 v_1, v_2, v_3. 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 w_1, w_2 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 5.3.2 Sia V uno spazio vettoriale di dimensione n su $\mathbb{R}$. Fissare una base di V significa costruire un isomorfismo tra V e $\mathbb{R}^n$. 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 $v_1, \dots, v_n$: $(\lambda_1, \lambda_2, \dots, \lambda_n)$. 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 $\varphi(v + v') = (\lambda_1 + \lambda'_1, \lambda_2 + \lambda'_2, \dots, \lambda_n + \lambda'_n) = (\lambda_1, \lambda_2, \dots, \lambda_n) + (\lambda'_1, \lambda'_2, \dots, \lambda'_n) = \varphi(v) + \varphi(v')$. 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 $\varphi(\lambda v) = (\lambda\lambda_1, \lambda\lambda_2, \dots, \lambda\lambda_n) = \lambda\varphi(v)$. 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 $\varphi(v) = \mathbf{0}_{\mathbb{R}^n} = (0,0,\dots,0)$ 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, $\mathbf{v} = \{v_1, \dots, v_n\}$ una base di V e $\mathbf{w} = \{w_1, w_2, \dots, w_m\}$ una base di W. Le coordinate rispetto a $\mathbf{w}$ dei vettori $L(v_1)$, $L(v_2), \dots, L(v_n)$ sono le colonne della matrice $A \in \mathcal{M}_{m,n}(\mathbb{R})$ associata ad L. Ora dobbiamo calcolare $\dim(\mathrm{Im}L) = \dim\langle L(v_1), L(v_2), \dots, L(v_n)\rangle$ (vedi 5.2), 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 $\{w_1, w_2, \dots, w_m\}$, e quindi come un vettore di $\mathbb{R}^m$. Grazie alla Definizione 4.3.3 e al Teorema 4.3.4 , 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.

5.4 Sistemi lineari e funzioni lineari

Che legame c'è tra sistemi lineari e funzioni lineari?

Abbiamo visto che, dati due spazi vettoriali V e W di dimensione rispettivamente n e m, e scelte due basi $\mathcal{V} = \{v_1, \dots, v_n\}$ e $\mathcal{W} = \{w_1, \dots, w_m\}$ rispettivamente di V e W, un'applicazione lineare $f : V \to W$ è rappresentata da una matrice $A = (a_{ij})_{i=1,\dots m}^{j=1,\dots,n} \in \mathcal{M}_{m,n}(\mathbb{R})$ in cui la colonna $j-esima$ è data dalle coordinate, rispetto alla base $\mathcal{W}$, dell'immagine del $j-esimo$ vettore della base $\mathcal{V}$, i.e., $f(v_j) = a_{1j}w_1 + a_{2j}w_2 + \cdots + a_{mj}w_m = (a_{1j}, a_{2j}, \dots, a_{mj})_{\mathcal{W}}$. Risulta altresì che per calcolare l'immagine mediante f del vettore $v \in V$ si devono innanzitutto calcolare le sue coordinate nella base $\mathcal{V}$: $v = \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_n v_n = (\beta_1, \beta_2, \dots, \beta_n)_{\mathcal{V}}$, dopodiché il prodotto $A \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix}$ fornirà le coordinate nella base $\mathcal{W}$ del vettore $f(v)$ (attenzione:

espresse in colonna, non in riga). La scelta delle basi ci è servita solamente per interpretare le applicazioni lineari in termini di matrici.

Prendiamo ora una matrice $A = (a_{ij}) \in \mathcal{M}_{m,n}(\mathbb{R})$ ed una $m - upla$ $\underline{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{R}^m$ (colonna dei termini noti). Allora cercare le $n - uple$ $\underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ (colonna delle incognite) tali che $A\underline{x} = \underline{b}$, cioè

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{m-1\,1} & a_{m-1\,2} & a_{m-1\,3} & \dots & \dots & a_{m-1\,n} \\ a_{m1} & a_{m2} & a_{m3} & \dots & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ \dots \\ \dots \\ \dots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ \dots \\ \dots \\ b_m \end{pmatrix},$$

equivale a cercare i vettori di coordinate $(x_1, x_2, \dots, x_n)$ che sono inviati in $(b_1, b_2, \dots, b_m)$ tramite l'applicazione lineare descritta dalla matrice A. Rispetto a quali basi? Dal momento che stiamo esprimendo tutto in coordinate stiamo fissando implicitamente ed una volta per tutte la base canonica di $\mathbb{R}^n$ nel dominio e la base canonica di $\mathbb{R}^m$ nel codominio. Denotiamo con $\varphi_A : \mathbb{R}^n \to \mathbb{R}^m$ l'applicazione lineare associata ad A rispetto alle basi canoniche del dominio e del codominio.

A questo punto l'interpretazione del problema iniziale è chiara: risolvere il sistema lineare $A\underline{x} = \underline{b}$ significa determinare tutti i vettori di $\mathbb{R}^n$ la cui immagine tramite l'applicazione lineare φ_A è il vettore $(b_1, \dots, b_m) \in \mathbb{R}^m$. Dall'Esercizio 5.1.4 conosciamo già tutti i risultati possibili:

-) se il vettore $(b_1, \dots, b_m) \in \mathbb{R}^m$ non appartiene all'immagine di φ_A il sistema non è risolubile,

-) se $(b_1, \dots, b_m)$ appartiene all'immagine di φ_A, l'insieme delle soluzioni del sistema è $(a_1, a_2, \dots, a_n) + \mathrm{Ker}\varphi_A$ (vedi (5.1)) dove $(a_1, a_2, \dots, a_n) \in \mathbb{R}^n$

è una soluzione particolare del sistema, i.e.,

$$A \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Non ci sono altre possibilità.

Teorema 5.4.1 (Teorema di Rouché-Capelli) *Dato un sistema lineare*

$$A\underline{x} = \underline{b}, \quad \textit{con } A \in \mathcal{M}_{m,n}(\mathbb{R}), \quad \underline{x} \in \mathbb{R}^n, \quad \underline{b} \in \mathbb{R}^m \ \textit{(vettori colonna)},$$

esso ammette soluzione se e solo se:

$$\mathrm{rg}A = \mathrm{rg}(A \mid \underline{b})$$

cioè se e solo se il rango della matrice incompleta del sistema coincide con il rango della matrice completa. Nel caso di uguaglianza l'insieme delle soluzioni è il sottoinsieme $v + \mathrm{Ker}A$ *di* $\mathbb{R}^n$*, ove* v *è una soluzione particolare del sistema e* $\mathrm{Ker}A$ *è il nucleo dell'applicazione lineare* φ_A*. Il sottospazio vettoriale* $\mathrm{Ker}A$ *ha dimensione* $n - \mathrm{rg}A$*.*

Dimostrazione. L'idea è la seguente: nell'interpretazione data del sistema in termini dell'applicazione lineare $\varphi_A : \mathbb{R}^n \to \mathbb{R}^m$, dire che il sistema ha soluzione significa che il vettore w che ha coordinate $\underline{b}$ rispetto alla base canonica di $\mathbb{R}^m$ appartiene all'immagine dell'applicazione φ_A e cioè che la colonna $\underline{b}$ è combinazione lineare delle colonne di A (che generano l'immagine di φ_A). Pertanto dire che il sistema ammette soluzione equivale a dire che il numero di colonne linearmente indipendenti della matrice A non cambia se si aggiunge ad A la colonna dei termini noti, cioè che il rango della matrice completa $(A \mid \underline{b}) \in \mathcal{M}_{m,n+1}(\mathbb{R})$ è uguale al rango della matrice incompleta A.

La struttura della controimmagine di un vettore tramite una applicazione lineare è già stata studiata nell'Esercizio 5.1.4, quindi il resto del contenuto del teorema è già stato dimostrato. **C.V.D.**

Osservazione 5.4.2 Abbiamo già osservato che in uno spazio vettoriale V un sottoinsieme della forma $v + \ker A$ è un sottospazio vettoriale di V se e solo se $v \in \ker A$. In tal caso $v + \ker A = \ker A$. Sia ora $v + \ker A$ l'insieme delle soluzioni di un sistema lineare della forma $A\underline{x} = \underline{b}$. Allora tale insieme è un sottospazio vettoriale se e solo se il sistema è omogeneo, vale a dire se e solo se $\underline{b} = \underline{0}$.

Osserviamo inoltre che l'insieme $v+\ker A$ è costituito da un solo elemento se $\ker A$ è banale altrimenti è costituito da infiniti elementi, perché se $\ker A$ non è banale contiene infiniti elementi. Abbiamo così ottenuto un'ulteriore dimostrazione del fatto che se un sistema lineare a coefficienti reali ammette soluzioni, tali soluzioni sono infinite o una sola. Non ci sono altre possibilità.

5.5 Esercizi svolti

Esercizio 5.5.1 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 5.5.2 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 $\mathrm{Ker} f$ e $\mathrm{Im} f$.

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 $\mathrm{Im} f$ è generata dai vettori $f(1,0,0)$, $f(0,1,0)$ e $f(0,0,1)$. Dunque $\mathrm{Im} f = \langle (1,1,0), (0,0,1) \rangle$. Usando il Teorema delle dimensioni 5.2.1 otteniamo che $\mathrm{Ker}(f)$ 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 $\mathrm{Ker}(f) = \langle (1,-1,0) \rangle$.

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 5.5.3 Sia $D : \mathbb{R}^{\leq 3}[x] \longrightarrow \mathbb{R}^{\leq 3}[x]$ l'applicazione derivata (rispetto alla variabile x). Determinare Ker$D$, ImD 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,

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

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 5.2.1 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 5.5.4 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 5.5.5 Tra le applicazioni lineari dell'esercizio 5.5.1 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 5.2.2.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 5.5.6 Sia $L : V \longrightarrow W$ un'applicazione lineare e siano $\{v_1, \ldots, v_n\}$ vettori linearmente indipendenti di V. Mostrare che se L è iniettiva allora i vettori $\{L(v_1), \ldots, L(v_n)\}$ di W sono linearmente indipendenti.

Svolgimento. Consideriamo una combinazione lineare dei vettori $L(v_1)$, $\ldots, L(v_n)$ 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 $v_1, \ldots, v_n$ sono linearmente indipendenti. L'esercizio è concluso.

Esercizio 5.5.7 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 5.2.2iii), 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 5.2.2iii), 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 5.5.8 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 5.5.9 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 $\text{Ker}L = \mathbf{0}_{\mathbb{R}^3}$ 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 $\text{rg}A = 3$.

5.6 Esercizi proposti

Esercizio 5.6.1 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 5.6.2 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 5.6.3 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 5.6.4 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 5.6.5 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 6

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.

6.1 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 6.1.1 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 6.1.2 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 6.1.3 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 6.1.4 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.

6.2 Matrici invertibili

Proposizione 6.2.1 *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 $\mathcal{V} = \{v_1, \ldots, v_n\}$ del dominio di f e una base $\mathcal{W} = \{w_1, \ldots, w_n\}$ 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 $\{v_1, \ldots, v_n\}$. Le sue coordinate sono $(1, 0, 0, \ldots, 0)$, così come l'immagine di v_2 è v_2 e le sue coordinate nella base $\mathcal{V}$ sono $(0, 1, 0, \ldots, 0)$. 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 6.2.2 Si noti che se $C \in \mathcal{M}_{n,p}(\mathbb{R})$ e $I_n \in \mathcal{M}_n(\mathbb{R})$ allora $I_n C = C$. (Attenzione: il prodotto CI_n non è definito! Se, però, I_p è la matrice identica di ordine p allora $CI_p = C$).

Definizione 6.2.3 *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}.*

Osservazione 6.2.4 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.

6.3 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:

$$\begin{array}{rccc} \det: & \mathcal{M}_1(\mathbb{R}) & \longrightarrow & \mathbb{R} \\ & A = (a) & \longmapsto & a. \end{array}$$

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 6.3.1 *Data la matrice* $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{R})$*, poniamo*

$$\det(A) = a_{11} \quad \textit{se} \ \ n = 1$$

e

$$\begin{aligned} \det A = & \ \det(a_{ij}) = (-1)^{k+1} a_{k1} \det(\mathcal{A}_{k1}) + (-1)^{k+2} a_{k2} \det(\mathcal{A}_{k2}) + \\ + & \ (-1)^{k+3} a_{k3} \det(\mathcal{A}_{k3}) + \cdots + (-1)^{k+n} a_{kn} \det(\mathcal{A}_{kn}) \end{aligned} \tag{6.1}$$

se $n > 1$*, con* $1 \leq k \leq n$*.*

Attenzione: l'espressione (6.1) si chiama sviluppo del determinante secondo la riga k-esima. Ogni riga può essere scelta per il calcolo del determinante:

il risultato della formula (6.1) 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 6.3.2 *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 6.3.3 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.

6.4 Esercizi svolti

Esercizio 6.4.1 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 6.4.2 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 6.4.3 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 6.4.4 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.

Svolgimento. Si verifica facilmente che il rango della matrice H è uguale a 3, pertanto la funzione h è invertibile.

6.5 Esercizi proposti

Esercizio 6.5.1 Calcolare $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^n$.

Esercizio 6.5.2 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 6.5.3 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 6.5.4 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 6.5.5 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 6.5.6 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 7

Matrici diagonalizzabili

Ci concentreremo ora sugli endomorfismi di uno spazio vettoriale V di dimensione n. Sappiamo che, scelta una base di V, ad ogni endomorfismo resta associata una matrice $A \in M_n(\mathbb{R})$. Ci porremo due problemi diversi:

1) trovare una base rispetto alla quale la matrice del nostro endomorfismo abbia la forma più semplice possibile;

2) trovare un modo per decidere se due matrici rappresentino o meno lo stesso endomorfismo.

Le risposte che daremo sono parziali; per una risposta completa bisognerebbe ricorrere alla teoria di Jordan.

Sottolineiamo che useremo in maniera equivalente matrici quadrate ed endomorfismi.

7.1 Autovalori e autovettori

Consideriamo un endomorfismo $\varphi : V \to V$.

Definizione 7.1.1 *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 $\varphi : V \to V$. Consideriamo l'insieme V_λ^* degli autovettori di autovalore λ (che, per definizione di autovalore, non è mai 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 7.1.2 *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 $\alpha v \in V_\lambda$. Dunque V_λ è un sottospazio vettoriale di V. **C. V. D.**

Definizione 7.1.3 *Dati un endomorfismo φ di uno spazio vettoriale V ed un suo autovalore λ, il sottospazio vettoriale V_λ è detto l'autospazio di φ relativo all'autovalore λ.*

Come si calcolano autovalori e autovettori di un endomorfismo? Innanzitutto osserviamo che, scelta una base $\mathcal{B} = \{v_1, \ldots, v_n\}$ di V ($\dim V = n$) e indicata con $A \in M_n(\mathbb{R})$ la matrice associata a φ rispetto a questa base sia nel dominio che nel codominio di φ, dire che un vettore v è un autovettore di autovalore λ, significa che per la $n-upla$ delle sue coordinate $(a_1, \ldots, a_n)$ nella base $\mathcal{B}$ si ha

$$A \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \lambda \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}.$$

In altre parole λ è autovalore di φ se e solo se esiste un vettore non nullo $v \in V$ tale che:

$$Av = \lambda v$$

dove il vettore v è pensato come il vettore colonna delle sue coordinate rispetto alla base $\mathcal{B}$. Abbiamo:

$$Av = \lambda v \Leftrightarrow Av - \lambda v = \underline{0} \Leftrightarrow (A - \lambda I_n)v = \underline{0},$$

dunque λ è autovalore di φ se e solo se esiste un vettore non nullo v nel nucleo della matrice $A - \lambda I_n$, cioè se e solo se tale matrice ha nucleo non banale o, equivalentemente, NON è invertibile, cioè se e solo se $\det(A - \lambda I_n) = 0$. Analogamente, il ragionamento appena illustrato mostra che $V_\lambda = \ker(A - \lambda I_n)$.

Cerchiamo ora di capire meglio che oggetto è $\det(A - \lambda I_n)$.

Esempio 7.1.4 i) Sia $A = \begin{pmatrix} 1 & 2 \\ 3 & -1 \end{pmatrix}$ e calcoliamo $\det(A - \lambda I_2)$:

$$\det(A-\lambda I_2) = \det \begin{pmatrix} 1 & 2 \\ 3 & -1 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \det \begin{pmatrix} 1-\lambda & 2 \\ 3 & -1-\lambda \end{pmatrix} =$$

$$= (1-\lambda)(-1-\lambda) - 6 = \lambda^2 - 7.$$

$\det(A - \lambda I_2)$ è dunque un polinomio nella variabile λ di grado 2.

ii) Sia $B = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 2 & -1 & 1 \end{pmatrix}$ e calcoliamo $\det(B - \lambda I_3)$:

$$\det(B - \lambda I_3) = \det \begin{pmatrix} -\lambda & -1 & 0 \\ 1 & -\lambda & 0 \\ 2 & -1 & 1-\lambda \end{pmatrix} = -\lambda^3 + \lambda^2 - \lambda + 1.$$

$\det(B - \lambda I_3)$ è dunque un polinomio nella variabile λ di grado 3.

Osservazione 7.1.5 Sulla base degli esempi appena fatti, lo studente non farà fatica a convincersi del fatto che, data $A \in M_n(\mathbb{R})$, $det(A - \lambda I_n)$ è un polinomio di grado n nella variabile λ. Tale polinomio si chiama *polinomio caratteristico* della matrice A o, equivalentemente, dell'endomorfismo φ. È molto importante sottolineare che il polinomio caratteristico di φ non dipende dalla base scelta per costruire la matrice A. In altre parole, se B è la matrice associata a φ rispetto ad un'altra base, allora il polinomio caratteristico di B coincide con quello di A.

Abbiamo dunque dimostrato la seguente proposizione:

Proposizione 7.1.6 *Un numero reale λ è un autovalore dell'endomorfismo φ se e solo se è radice del suo polinomio caratteristico.*

Osservazione 7.1.7 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) 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 che descrivono lo stesso endomorfismo rispetto a basi diverse hanno lo stesso polinomio caratteristico (Osservazione 7.1.5), quanto appena osservato implica che matrici che descrivono lo stesso endomorfismo rispetto a basi diverse hanno lo stesso determinante.

iii) Poiché il polinomio caratteristico di un endomorfismo φ di uno spazio vettoriale n-dimensionale ha grado n, 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 7.1.8 *Sia $f : V \to V$ un endomorfismo di V. Autovettori di f relativi ad autovalori diversi sono linearmente indipendenti.*

7.2 Matrici/endomorfismi diagonalizzabili

Mettiamoci ora in una situazione "idilliaca": sia $\varphi : V \to V$ un endomorfismo e supponiamo che esista una base di V costituita da autovettori: $\mathcal{B} = \{v_1, \ldots, v_n\}$ con $\varphi(v_i) = \lambda_i v_i$, $i = 1, \ldots, n$ (dove gli autovalori λ_i non sono necessariamente tutti distinti). Adesso scriviamo la matrice dell'applicazione lineare φ rispetto alla base $\mathcal{B}$. Cominciamo da v_1: $\varphi(v_1) = \lambda_1 v_1$ che in coordinate nella base $\mathcal{B}$ si scrive come $\lambda_1 v_1 + 0v_2 + \cdots + 0v_n = (\lambda_1, 0, \ldots, 0)_{\mathcal{B}}$. Per v_2 si ha $\varphi(v_2) = \lambda_2 v_2 = 0v_1 + \lambda_2 v_2 + 0v_3 + \cdots + 0v_n = (0, \lambda_2, 0, \ldots, 0)$. In definitiva la matrice associata a φ rispetto ad una base di autovettori è diagonale e gli elementi sulla diagonale sono gli autovalori:

$$D = \begin{pmatrix} \lambda_1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda_2 & 0 & \ldots & 0 \\ 0 & 0 & \ddots & \ddots & 0 \\ 0 & 0 & \ddots & \lambda_{n-1} & 0 \\ 0 & 0 & \ldots & 0 & \lambda_n \end{pmatrix}.$$

Definizione 7.2.1 *Sia λ un autovalore di un endomorfismo φ di uno spazio vettoriale V. Si dice molteplicità algebrica di λ la sua molteplicità come radice del polinomio caratteristico. Si dice molteplicità geometrica di λ la dimensione dell'autospazio relativo all'autovalore λ.*

Definizione 7.2.2 *Un endomorfismo φ di uno spazio vettoriale V si dice diagonalizzabile se esiste una base di V costituita da autovettori di φ. Equivalentemente, una matrice $A \in M_n(\mathbb{R})$ si dice diagonalizzabile se è la matrice di un endomorfismo diagonalizzabile (rispetto ad una certa base).*

Molteplicità algebrica e geometrica di un autovalore sono legate dal seguente risultato:

Teorema 7.2.3 *La molteplicità algebrica di un autovalore è sempre maggiore della sua molteplicità geometrica o uguale ad essa.*

Teorema 7.2.4 *Una matrice $A \in M_n(\mathbb{R})$ (o, equivalentemente, un endomorfismo φ di uno spazio vettoriale n-dimensionale V) è diagonalizzabile (su $\mathbb{R}$) se e solo se valgono le seguenti due condizioni:*

1. *il polinomio caratteristico di A (o di φ) ha n autovalori reali (non necessariamente distinti);*

2. *la molteplicità algebrica di ogni autovalore λ coincide con la sua molteplicità geometrica.*

Osservazione 7.2.5 *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 nel prodotto di polinomi di grado 1, 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 $z = 3$. Per il Teorema 7.2.4 dunque, la matrice A non è diagonalizzabile (su $\mathbb{R}$).

ii) Anche se il polinomio caratteristico di una matrice $n \times n$ ha n autovalori reali, può accadere che la matrice non sia diagonalizzabile. Ad esempio 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 algebrica 2. La molteplicità geometrica dell'autovalore 0, tuttavia, cioè la dimensione di $V_0 = \ker B$, è uguale ad 1, quindi, per il Teorema 7.2.4, la matrice 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 algebrica. Ovviamente tale forma diagonale non è unica, bensì unica a meno di permutazioni degli elementi sulla diagonale.

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 $n \times n$ ha n autovalori distinti, cioè se la molteplicità algebrica 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 il Teorema 7.2.3. 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 ora al metodo per verificare se un endomorfismo (o, equivalentemente, una matrice) sia o meno diagonalizzabile e trovare, in caso affermativo, una base di autovettori che lo diagonalizzi. Siano dati l'endomorfismo

$\varphi : V \to V$ e la matrice $A \in M_n(\mathbb{R})$ associata a φ rispetto ad una base fissata di V.

1) Si calcola il polinomio caratteristico $\det(A - zI_n)$. Se le sue radici non sono tutte reali, allora φ non è diagonalizzabile. Altrimenti: $\det(A - zI_n) = (\alpha_1 - z)^{n_1}(\alpha_2 - z)^{n_2} \cdots (\alpha_k - z)^{n_k}$, $\alpha_i \in \mathbb{R}$.

2) Si considerano gli autospazi V_{α_i}. Si ha $V_{\alpha_i} = Ker(A - \alpha_i I_n)$ quindi V_{α_i} è un sottospazio di V di dimensione $n - \mathrm{rg}(A - \alpha_i I_n)$, dato dalle soluzioni del sistema

$$(A - \alpha_i I_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Nel caso in cui si voglia soltanto sapere se la matrice $A \in M_n(\mathbb{R})$ è diagonalizzabile o meno, basterà verificare se per ogni autovalore α_i di molteplicità algebrica n_i si ha $\dim(V_{\alpha_i}) = n - \mathrm{rg}(A - \alpha_i I_n) = n_i$. (Abbiamo già osservato che per gli autovalori con molteplicità algebrica 1 tale uguaglianza è sempre verificata). In caso affermativo la matrice è diagonalizzabile, altrimenti non lo è.

Per trovare poi una base di autovettori di A si dovrà scegliere una base di ogni autospazio e prendere l'unione delle basi trovate.

Esempio 7.2.6 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 base di autovettori di T.

Svolgimento. Seguiamo esattamente le linee guida del procedimento che abbiamo illustrato. Calcoliamo dunque il polinomio caratteristico di T nella incognita λ:

$$\det(\begin{pmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{pmatrix} - \lambda I_3) = \det \begin{pmatrix} 3-\lambda & 2 & 4 \\ 2 & -\lambda & 2 \\ 4 & 2 & 3-\lambda \end{pmatrix} = P_T(\lambda).$$

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 $\mathbb{R}[\lambda]$ in fattori di primo grado. Le sue radici sono: 8, di molteplicità algebrica 1, e -1 di molteplicità algebrica 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 algebrica dell'autovalore -1. Per l'autovalore 8 il risultato è vero automaticamente dal momento che esso ha molteplicità algebrica 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 la matrice T è 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 $V_{-1} = \langle (1, -2, 0), (0, -2, 1) \rangle$.

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 $V_8 = \langle(2, 1, 2)\rangle$. Una base di autovettori di T è quindi data da $\{(1, -2, 0), (0, -2, 1), (2, 1, 2)\}$.

Esempio 7.2.7 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 $\langle(1, 0, 0)\rangle$.

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 algebrica due e quindi, affinché essa sia diagonalizzabile, occorre che l'autospazio V_2 relativo all'autovalore 2 abbia dimensione 2. Si ha:

$$\dim(V_2) = 3 - \mathrm{rg}(B_3 - 2I_3) = 3 - \mathrm{rg}(\begin{pmatrix} 3-1-2 & 3-3 & 3 \\ 0 & \frac{5}{2}-2 & -\frac{1}{2} \\ 0 & -\frac{1}{2} & \frac{5}{2}-2 \end{pmatrix}) = 3 - 2 = 1.$$

Quindi la dimensione di V_2 non coincide con la molteplicità algebrica dell'autovalore 2: la matrice B_3 non è diagonalizzabile.

Consideriamo la matrice B_4: essa ha autovalore 3 di molteplicità algebrica 2. Calcoliamo la dimensione di V_3:

$$\dim(V_3) = 3 - \mathrm{rg}(B_4 - 3I_3) = 3 - \mathrm{rg}(\begin{pmatrix} 4-1-3 & 4-3 & 4 \\ 0 & \frac{5}{2}-3 & -\frac{1}{2} \\ 0 & -\frac{1}{2} & \frac{5}{2}-3 \end{pmatrix}) = 3 - 2 = 1.$$

Anche in questo caso la dimensione dell'autospazio V_3 non coincide con la molteplicità algebrica dell'autovalore 3. La matrice B_4 non è diagonalizzabile.

Esempio 7.2.8 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 algebrica 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 $(0, 0, 1, 1)$: $V_1 = \langle(1, 0, -2, 0), (0, 0, 1, 1)\rangle$.

Si noti che se la matrice A fosse stata diagonalizzabile la sua forma diagonale sarebbe stata la matrice identica I_4. La matrice A avrebbe dunque descritto la funzione identica $id : \mathbb{R}^4 \to \mathbb{R}^4$, $id(v) = v$. Ma la matrice della funzione identica rispetto ad una qualsiasi base è la matrice identica, dunque avremmo potuto concludere che A non è diagonalizzabile usando questo ragionamento.

7.3 Esercizi svolti

Esercizio 7.3.1 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 7.3.2 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 5.5.3 abbiamo determinato la matrice associata all'applicazione lineare D rispetto alla base $\{1, x, x^2, x^3\}$ di $\mathbb{R}^{\leq 3}[x]$ (detta base canonica):

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

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 la matrice nulla descrive, rispetto a qualsiasi base, solo l'applicazione identicamente nulla e D non è certamente la funzione identicamente 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 7.3.3 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 base di autovettori.

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. La sua forma diagonale è:

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

Determiniamo una base di $\mathbb{R}^3$ di autovettori di A. Si tratta dunque di determinare gli autospazi della matrice A. Cominciamo con $V_0 = \mathrm{Ker}A = \{(x, y, z) \mid A(x, y, z)^t = (0, 0, 0)^t\}$. Facendo i calcoli si ottiene $V_0 = \langle(1, -2, -1)\rangle$.

Analogamente $V_2 = \{(x, y, z) \mid (A-2I_3)(x, y, z)^t = (0, 0, 0)^t\} = \langle(1, 0, 1)\rangle$.

Infine, $V_3 = \{(x, y, z) \mid (A - 3I_3)(x, y, z)^t = (0, 0, 0)^t\} = \langle(1, 1, -1)\rangle$.

Una base di $\mathbb{R}^3$ costituita da autovettori di A è dunque: $\mathcal{B} = \{(1, -2, -1), (1, 0, 1), (1, 1, -1)\}$.

Osservazione Si noti l'ordine in cui abbiamo scritto gli autovettori nella base $\mathcal{B}$: il primo vettore è un elemento v_0 di V_0, il secondo un elemento v_2 di V_2, infine, il terzo è un elemento v_3 di V_3. Tale ordine è coerente con l'ordine in cui sono disposti gli autovalori nella forma diagonale D che abbiamo scritto. In altre parole se interpretiamo A come la matrice di un endomorfismo f_A di $\mathbb{R}^3$ rispetto alla base canonica, allora D è la matrice di f_A rispetto alla base di autovettori $\mathcal{B}$.

Esercizio 7.3.4 Sia $V = M_2(\mathbb{R})$ 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 $M_2(\mathbb{R})$: $\mathcal{C} = \{e_{11}, e_{12}, e_{21}, e_{22}\}$ 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 $V = M_2(\mathbb{R})$ 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 7.3.5

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):

$$\dim(V_1) = 3 - \mathrm{rg}(A_0 - I_3) = 3 - \mathrm{rg} \begin{pmatrix} 0 & 0 & 0 \\ 1 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix} = 1.$$

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:

$$\dim(V_{\frac{1}{2}}) = 3 - \mathrm{rg}(A_{-\frac{1}{4}} - \frac{1}{2}I_3) = 3 - \mathrm{rg} \begin{pmatrix} \frac{1}{2} & -\frac{1}{4} & 0 \\ 1 & -\frac{1}{2} & 0 \\ \frac{5}{4} & 0 & \frac{1}{2} \end{pmatrix} = 3 - 2 = 1.$$

Anche in questo caso, dunque, la matrice A_k non è diagonalizzabile.

In conclusione A_k è diagonalizzabile se e solo se $k \neq 0, -\frac{1}{4}$.

7.4 Esercizi proposti

Esercizio 7.4.1 Stabilire per quali valori del parametro reale k la matrice $A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & k \\ 0 & -k & 0 \end{pmatrix}$ è diagonalizzabile.

Esercizio 7.4.2 Data la matrice $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$:

1. determinare gli autovalori di A e gli autospazi ad essi relativi;

2. stabilire se la matrice A è diagonalizzabile e, in caso affermativo, determinare una forma diagonale della matrice ed una base di autovettori.

Esercizio 7.4.3 Stabilire per quali valori del parametro reale h la matrice $A = \begin{pmatrix} 1 & 0 & h \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$ è diagonalizzabile. Per ognuno dei valori trovati determinare una base di $\mathbb{R}^3$ costituita da autovettori di A.

Esercizio 7.4.4 Stabilire se esistono valori del parametro reale k tali che la matrice

$$K = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ k & 0 & 0 & 0 \end{pmatrix}$$

sia diagonalizzabile su $\mathbb{R}$.

Appendice A

Aritmetica elementare

In questo capitolo introdurremo lo studente all'aritmetica degli interi. Inizieremo con il principio di induzione, un risultato di fondamentale rilievo che ha numerose applicazioni nei vari ambiti della matematica. Procederemo poi con l'algoritmo della divisione, l'algoritmo di Euclide, fino ad arrivare alle congruenze, l'argomento più importante della matematica discreta elementare. Lo studente avrà l'impressione (corretta) che il contenuto di questa appendice non abbia molto a che fare con il contenuto delle lezioni precedenti. In effetti questi argomenti hanno caratteristiche completamente diverse dal resto del corso.

A.1 Il principio di Induzione

Il principio di induzione è la tecnica principe per dimostrare enunciati riguardanti i numeri naturali cioè l'insieme $\mathbb{N} = \{1, 2, \ldots\}$. Facciamo innanzitutto un esempio.

Supponiamo di voler dimostrare che la somma dei numeri naturali da 1 a n è $n(n+1)/2$:

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}. \tag{A.1}$$

Indicheremo tale enunciato con $P(n)$. Verifichiamo innanzitutto la sua validità per $n = 1$: si tratta di mostrare che la somma dei numeri naturali da 1 a 1 (cioè 1), è $1(1+1)/2 = 1$. Dunque l'affermazione $P(1)$ è vera. Analogamente vale $P(2)$: $1 + 2 = 2(2+1)/2 = 3$. È chiaro che, con un po' di pazienza, potremmo andare avanti così verificando la formula (A.1) per n anche molto grande, ma noi vogliamo dimostrare che essa vale *per tutti i numeri naturali* n. Il principio di induzione ci viene in aiuto permettendoci di dimostrare la validità di un enunciato per tutti i numeri naturali.

Principio di Induzione. Supponiamo che una affermazione $P(n)$ venga formulata per ogni numero naturale. Se si verifica che:

1) $P(1)$ è vera;

2) se $P(k)$ è vera allora $P(k+1)$ è vera;

allora $P(n)$ è vera per ogni $n \in \mathbb{N}$.

Esempio A.1.1 Vediamo come utilizzare il principio di induzione per dimostrare la formula (A.1), ossia l'enunciato $P(n)$:

la somma dei primi n *numeri naturali* è $\frac{n(n+1)}{2}$.

Abbiamo già visto che tale enunciato è vero per $n = 1$, cioè vale l'ipotesi 1) del principio di induzione. Ora supponiamo che $P(k)$ sia vero e cioè che:

$$1 + 2 + \ldots + k = \frac{k(k+1)}{2}.$$

Dobbiamo mostrare che $P(k) \implies P(k+1)$ (ipotesi 2) del principio di induzione), cioè che:

$$1 + 2 + 3 + \ldots + k + (k+1) = \frac{(k+1)(k+2)}{2}.$$

Dal momento che vale $P(k)$, abbiamo:

$$1 + \ldots + k + (k+1) = \frac{k(k+1)}{2} + k + 1 = \frac{k^2 + k + 2k + 2}{2} = \frac{(k+1)(k+2)}{2},$$

quindi $P(k+1)$ è vero. A questo punto il principio di induzione garantisce che l'enunciato $P(n)$ valga per ogni n.

Esempio A.1.2 Dimostriamo, utilizzando il Principio di induzione, la seguente affermazione $P(n)$:

$$\sum_{j=1}^{n} j^2 = \frac{n(n+1)(2n+1)}{6}.$$

1) Verifichiamo la validità di $P(1)$: $1 = \frac{1 \cdot 2 \cdot 3}{6}$ su cui non ci sono dubbi.

2) Supponiamo (ipotesi induttiva) che valga $P(k)$, ossia $\sum_{j=1}^{k} j^2 = \frac{k(k+1)(2k+1)}{6}$, e calcoliamo

$$\sum_{j=1}^{k+1} j^2 = \sum_{j=1}^{k} j^2 + (k+1)^2 = \frac{k(k+1)(2k+1)}{6} + (k+1)^2 =$$

$$\frac{(k+1)(2k^2+7k+6)}{6} = \frac{(k+1)(k+2)(2k+3)}{6}$$

che è esattamente il contenuto dell'affermazione $P(k+1)$. Dunque, per il Principio di induzione, l'affermazione $P(n)$ è vera per ogni $n \in \mathbb{N}$.

Esistono varie formulazioni equivalenti del principio di induzione.

Principio di Induzione Completa. Supponiamo che un'affermazione $P(n)$ venga formulata per ogni numero naturale. Se si verifica che:

1. $P(1)$ è vera;

2. se $P(j)$ è vera per ogni $j < k$ allora $P(k)$ è vera;

allora $P(n)$ è vera per ogni $n \in \mathbb{N}$.

Assioma del buon ordinamento. Ogni sottoinsieme non vuoto dell'insieme dei numeri naturali contiene un elemento minimo.

Teorema A.1.3 *Sono equivalenti:*

1. *Il principio di induzione.*

2. *Il principio di induzione completa.*

3. *L' assioma del buon ordinamento.*

Dimostrazione. Dimostreremo che (1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\rightarrow$ (1).

(1) $\Rightarrow$ (2) Assumiamo che valga il Principio di Induzione e supponiamo che venga formulata un'affermazione $P(n)$, per $n \in \mathbb{N}$, tale che:

a) valga $P(1)$;

b) se vale $P(j)$ per ogni $0 \leq j < k$ allora vale $P(k)$.

Dimostriamo che vale $P(n)$ per ogni $n \in \mathbb{N}$. Formuliamo la seguente affermazione $B(n)$:

$$P(j) \text{ è vera per ogni } 0 \leq j \leq n.$$

L'affermazione $B(n)$ soddisfa le ipotesi del Principio di Induzione, infatti:

1) $B(1)$ è vera, dal momento che $P(1)$ è vera;

2) se $B(k)$ è vera, cioè $P(j)$ è vera per ogni $0 \leq j \leq k$, per *b*) anche $P(k+1)$ è vera, quindi $B(k+1)$ è vera.

Dunque, per il Principio di Induzione, $B(n)$ è vera per ogni $n \in \mathbb{N}$, in particolare $P(n)$ è vera per ogni $n \in \mathbb{N}$.
(2) $\Rightarrow$ (3) Sia $S \subset \mathbb{N}$ un sottoinsieme di $\mathbb{N}$ privo di elemento minimo. Vogliamo dimostrare che $S = \emptyset$ o, equivalentemente, che vale l'affermazione $P(n)$: $n \notin S$ per ogni $n \in \mathbb{N}$. Mostriamo che $P(n)$ è vera utilizzando il principio di induzione completa. $P(1)$ è vera, cioè $1 \notin S$, altrimenti 1 sarebbe l'elemento minimo di S (essendo 1 l'elemento minimo di $\mathbb{N}$). Supponiamo che $P(j)$ sia vera per ogni $j \leq k$, cioè supponiamo che $1, 2, \ldots, k \notin S$. Allora $k+1 \notin S$, altrimenti $k+1$ sarebbe il minimo di S. Dunque, per il principio di induzione completa, $P(n)$ vale per ogni $n \in \mathbb{N}$, cioè vale l'assioma del buon ordinamento.
(3) $\Rightarrow$ (1). In questo caso supponiamo valido l'assioma del buon ordinamento e supponiamo che un'affermazione $P(n)$, formulata per ogni $n \in \mathbb{N}$, soddisfi le ipotesi:

a) $P(1)$ è vera;

b) $P(k)$ vera $\Rightarrow$ $P(k+1)$ vera.

Vogliamo dimostrare che $P(n)$ è vera per ogni $n \in \mathbb{N}$. Sia

$$S = \{j \in \mathbb{N} \mid P(j) \text{ è falsa}\}.$$

Vogliamo mostrare che $S = \emptyset$. Supponiamo per assurdo che S sia non vuoto. Allora, per l'assioma del buon ordinamento, S contiene un elemento minimo m. Questo significa che $m - 1 \notin S$, cioè $P(m-1)$ è vera. Per l'ipotesi *b*), quindi, $P(m-1+1) = P(m)$ è vera, cioè $m \notin S$. Assurdo. □

L'utilizzo dell'enunciato $P(k)$ per dimostrare $P(k+1)$ prende il nome di *ipotesi induttiva*.

Remark A.1.4 Il Principio di Induzione (così come una delle sue formulazioni equivalenti) può essere utilizzato anche per dimostrare la validità di affermazioni $P(n)$ che vengono formulate per $n \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$ (resp. $\{k \in \mathbb{N} \mid k \geq n_0\}$). In tal caso basterà sostituire 0 (resp. n_0) ad 1 nel primo passo del procedimento induttivo.

Come esempio di applicazione del Principio di Induzione Completa dimostriamo il seguente ben noto algoritmo di divisione:

Teorema A.1.5 *Siano $n \in \mathbb{N}_0$, $b \in \mathbb{N}$. Allora esistono e sono unici due numeri naturali q ed r, detti rispettivamente quoziente e resto, tali che:*

$$n = qb + r, \qquad \text{con} \quad 0 \leq r < b.$$

Dimostrazione. Dimostriamo il teorema con il Principio di induzione completa su $\mathbb{N}_0$. $P(0)$ è vero in quanto $n = 0 = 0b+0$, con $q = r = 0$. Supponiamo che $P(j)$ sia vero per ogni j tale che $0 \leq j < k$ e vogliamo mostrare $P(k)$ (vogliamo, cioè verificare l'ipotesi 2) del principio di induzione completa). Se $k < b$ allora:

$$k = 0 \cdot b + k, \qquad \text{con resto } r = k, \quad 0 \leq k < b,$$

dunque $P(k)$ è vero per $k < b$. Sia ora $k \geq b$. Poiché $0 \leq k - b < k$, applicando l'ipotesi induttiva a $k - b$ abbiamo:

$$k - b = q_1 \cdot b + r_1, \qquad \text{con} \quad 0 \leq r_1 < b,$$

da cui ricaviamo:

$$k = (q_1 + 1)b + r_1, \qquad \text{con} \quad 0 \leq r_1 < b,$$

che dimostra quanto vogliamo. □

A.2 L'algoritmo di divisione e l'algoritmo di Euclide

Vogliamo ora introdurre il concetto di *divisibilità* e di *massimo comun divisore*.

Definizione A.2.1 *Siano a e b due numeri interi. Diciamo che b divide a se esiste un numero intero c tale che $a = bc$ e scriviamo $b|a$. Diciamo che d è il massimo comun divisore tra due numeri interi a e b se li divide entrambi ed è diviso da ogni altro divisore comune di a e b. Indicheremo il massimo comun divisore tra a e b con M.C.D.(a, b).*

Osservazione A.2.2 Osserviamo che se a, b e c sono numeri interi e a divide sia b che c, allora a divide anche $b+c$ e $b-c$. Lasciamo allo studente la facile verifica di questa proprietà che in seguito useremo più volte.

Come si calcola il massimo comun divisore tra due numeri naturali?

Teorema A.2.3 *(**Algoritmo di Euclide**) Siano a e b due numeri naturali tali che $b \leq a$ e b non divide a. Abbiamo allora:*

$$\begin{aligned}
a &= q_0 b + r_0 & 0 &\leq r_0 < b \\
b &= q_1 r_0 + r_1 & 0 &\leq r_1 < r_0 \\
r_0 &= q_2 r_1 + r_2 & 0 &\leq r_2 < r_1 \\
&\vdots \\
r_{t-2} &= q_t r_{t-1} + r_t & 0 &< r_t < r_{t-1} \\
r_{t-1} &= q_{t+1} r_t
\end{aligned}$$

e l'ultimo resto non nullo r_t è il massimo comun divisore tra a e b.

Dimostrazione. Per il Teorema A.1.5 possiamo scrivere: $a = qb + r$. Ora vogliamo mostrare che M.C.D. (a, b) = M.C.D. (b, r). Infatti se $c|a$ e $c|b$ allora $c|r$, poiché $r = a - qb$. Analogamente, se $c|b$ e $c|r$ allora $c|a = qb + r$. Dunque l'insieme degli interi che dividono sia a che b coincide con l'insieme degli interi che dividono sia b che r e dunque il massimo comune divisore delle due coppie (a, b) e (b, r) è lo stesso. Detto questo, il risultato segue immediatamente dalla catena di uguaglianze:

$$\text{M.C.D. } (a, b) = \text{M.C.D. } (b, r_0) = \text{M.C.D. } (r_0, r_1) = \ldots = \text{M.C.D. } (r_{t-1}, r_t) = r_t.$$

$\square$

Vediamo concretamente come utilizzare questo algoritmo per determinare il massimo comun divisore di due numeri dati.

Esempio A.2.4 calcoliamo M.C.D. $(603, 270)$. Utilizziamo l'algoritmo di Euclide (Teorema A.2.3).

$$\begin{aligned} 603 &= 2 \cdot 270 + 63 \\ 270 &= 4 \cdot 63 + 18 \\ 63 &= 3 \cdot 18 + 9 \\ 18 &= 2 \cdot 9 \end{aligned}$$

Dunque M.C.D. $(603, 270) = 9$.

Il teorema che segue è una conseguenza dell'algoritmo di Euclide e sarà lo strumento fondamentale per la risoluzione delle congruenze che studieremo nel paragrafo A.4.

Teorema A.2.5 Identità di Bézout. *Sia $d = M.C.D.\ (a, b)$. Allora esistono due numeri interi u, v (non unici) tali che:*

$$d = ua + vb.$$

Dimostrazione. La dimostrazione di questo risultato ripercorre l'algoritmo di Euclide A.2.3. Dimostriamo che ad ogni passo esistono $u_i, v_i \in \mathbb{Z}$ tali che $r_i = u_i a + v_i b$.

Per r_0 il risultato è vero, basta prendere $u_0 = 1$ e $v_0 = q_0$, infatti

$$r_0 = a - q_0 b.$$

Si ha poi che:

$$r_1 = b - r_0 q_1 = b - (u_0 a + v_0 b) q_1 = -u_0 a q_1 + (1 - v_o q_1) b$$

e il risultato è vero anche per r_1, basta prendere $u_1 = -u_0 q_1$ e $v_1 = 1 - v_0 q_1$. In generale dopo il passo $i - 1$ conosciamo $u_{i-2}, u_{i-1}, v_{i-2}, v_{i-1} \in \mathbb{Z}$ tali che:

$$r_{i-2} = u_{i-2} a + v_{i-2} b, \quad r_{i-1} = u_{i-1} a + v_{i-1} b.$$

Allora abbiamo che:

$$r_i = r_{i-2} - r_{i-1} q_i = u_{i-2} a + v_{i-2} b - (u_{i-1} a + v_{i-1} b) q_i$$

$$= (u_{i-2} - u_{i-1}q_i)a + (v_{i-2} - v_{i-1}q_i)b,$$

e il risultato è vero per r_i, basta prendere $u_i = u_{i-2} - u_{i-1}q_i$ e $v_i = v_{i-2} - v_{i-1}q_i$. Poiché M.C.D. (a, b) è l'ultimo resto non nullo r_t, dopo il passo t conosciamo u_t e v_t tali che $r_t =$ M.C.D. $(a, b) = u_t a + v_t b$ e $u = u_t$, $v = v_t$ sono gli interi cercati. □

Attenzione: l'esistenza di due numeri u, v tali che $d = ua + vb$ non garantisce che $d =$ M.C.D. (a, b).

Esempio A.2.6 Dall'Esempio A.2.4 sappiamo che M.C.D. $(603, 270) = 9$. Vogliamo ora ricavare due numeri u e v tali che $u\,603 + v\,270 = 9$. Procediamo a ritroso sostituendo attentamene i resti nella sequenza di equazioni ricavate nell'Esempio A.2.4.

$$\begin{aligned}
9 &= 63 - 3 \cdot 18 = \\
&= 63 - 3 \cdot [270 - 4 \cdot 63] = 63 - 3 \cdot 270 + 12 \cdot 63 = \\
&= (-3) \cdot 270 + 13 \cdot 63 == (-3) \cdot 270 + 13 \cdot [603 - 2 \cdot 270] = \\
&= (-3) \cdot 270 + 13 \cdot 603 + (-26) \cdot 270 \\
&= 13 \cdot 603 + (-29) \cdot 270.
\end{aligned}$$

Dunque $9 = 13 \cdot 603 + (-29) \cdot 270$ cioè vale la relazione $u\,603 + v\,270 = 9$ con $u = 13$ e $v = -29$.

Concludiamo questa sezione con un risultato che non dimostriamo e che non utilizzeremo in seguito, ma che riveste un ruolo fondamentale per la teoria degli interi e le cui generalizzazioni sono estremamente importanti in teoria dei numeri.

Diciamo che un intero p è *primo* se ha come unici divisori $\pm p$ e ± 1.

Teorema A.2.7 *Ogni intero è prodotto di primi in modo unico a meno dell'ordine e del segno:*

$$n = p_1 p_2 \dots p_r,$$

ove p_1, …, p_r sono numeri primi (non necessariamente distinti).

A.3 Classi di congruenza

Le classi di congruenza rappresentano un modo di contare, di sommare e di moltiplicare i numeri diverso da quello che conosciamo dall'aritmetica elementare e che tuttavia utilizziamo in ambiti particolari ogni giorno. Ad esempio se chiediamo a qualcuno di incontrarci precisamente tra 10 ore e sono le 23, la persona con cui stiamo parlando sa benissimo che l'incontro avverrà alle 9 e non alle 33, come la somma di numeri interi suggerirebbe. Analogamente se le lancette del nostro orologio segnano le 8 e abbiamo un appuntamento tra 6 ore, sappiamo che all'ora dell'appuntamento l'orologio segnerà le 2.

Vogliamo formalizzare questo modo di effettuare operazioni, che nei casi esaminati consiste semplicemente nel fare la somma dei numeri, dividere il risultato per un certo intero (che nel primo esempio è 24, nel secondo 12) e poi prendere il resto della divisione.

Definizione A.3.1 *Siano a, b ed n tre interi, con $n > 0$. Diciamo che a è congruo a b modulo n e scriviamo $a \equiv b \pmod{n}$ se $a - b = kn$ con $k \in \mathbb{Z}$.*

La relazione di congruenza tra numeri interi verifica le seguenti proprietà:

(i) Proprietà riflessiva: $a \equiv a \pmod{n}$, infatti $a - a = 0n$.

(ii) Proprietà simmetrica: se $a \equiv b \pmod{n}$ allora $b \equiv a \pmod{n}$. Infatti se $a - b = kn$ allora $b - a = -kn$.

(iii) Proprietà transitiva: se $a \equiv b \pmod{n}$ e $b \equiv c \pmod{n}$ allora $a \equiv c \pmod{n}$. Infatti se $a - b = kn$ e $b - c = hn$ per qualche $k, h \in \mathbb{Z}$ allora $a - c = a - b + b - c = (k + h)n$.

Una relazione soddisfacente la proprietà riflessiva, la proprietà simmetrica e la proprietà transitiva si chiama *relazione di equivalenza*. Non ci soffermeremo sulle relazioni di equivalenza in generale, tuttavia utilizzeremo le tre proprietà sopracitate delle congruenze.

Enunciamo ora un risultato che ci sarà utile in seguito e la cui dimostrazione lasciamo per esercizio.

Proposizione A.3.2 *Siano $a, b, c, d, n \in \mathbb{Z}$, con $n > 0$. Se $a \equiv b \pmod{n}$ e $c \equiv d \pmod{n}$ allora:*

1. $a + c \equiv b + d \pmod{n}$,
2. $ac \equiv bd \pmod{n}$.

Definiamo ora le classi di congruenza cioè gli insiemi che contengono numeri interi congrui tra loro.

Definizione A.3.3 *Siano $a, n \in \mathbb{Z}$. Si chiama classe di congruenza di a modulo n e si indica con $[a]_n$ l'insieme di numeri interi congrui ad a modulo n:*

$$[a]_n = \{b \in \mathbb{Z} \mid b \equiv a \pmod{n}\} = \{a + kn \mid k \in \mathbb{Z}\}.$$

Esempio A.3.4 Ad esempio se prendiamo $n = 4$ abbiamo che:

$$\begin{aligned}
[0]_4 &= \{0, 4, -4, 8, -8 \ldots\} = \text{multipli di } 4 \\
[1]_4 &= \{1, 5, -3, 9, -7 \ldots\} = 1 + \text{multipli di } 4 \\
[2]_4 &= \{2, 6, -2, 10, -6 \ldots\} = 2 + \text{multipli di } 4 \\
[3]_4 &= \{3, 7, -1, 11, -5 \ldots\} = 3 + \text{multipli di } 4 \\
[4]_4 &= [0]_4 \\
[5]_4 &= [1]_4 \\
[6]_4 &= [2]_4 \\
[7]_4 &= [3]_4 \\
&\ldots
\end{aligned}$$

Notiamo che in questo esempio:

- non ci sono elementi comuni a due classi di congruenza diverse;
- l'unione delle classi di congruenza è l'insieme di tutti i numeri interi;
- c'è un numero finito di classi di congruenza.

Questi fatti, come vedremo, valgono in generale.

Proposizione A.3.5 *Siano $[a]_n$ e $[b]_n$ due classi di congruenza modulo n. Allora $[a]_n = [b]_n$ oppure $[a]_n$ e $[b]_n$ sono disgiunte, cioè non hanno elementi in comune.*

Dimostrazione. Supponiamo che ci sia un elemento in comune tra $[a]_n$ e $[b]_n$, cioè $x \equiv a \pmod n$ e $x \equiv b \pmod n$. Allora, per la proprietà transitiva $a \equiv b \pmod n$ cioè $[a]_n = [b]_n$. $\square$

Dalla definizione di relazione di congruenza seguono facilmente i seguenti fatti:

Proposizione A.3.6
1. *Sia r il resto della divisione di a per n. Allora $[a]_n = [r]_n$.*
2. *$[0]_n$, $[1]_n$, $\dots$, $[n-1]_n$ sono tutte e sole le classi di congruenza distinte modulo n.*
3. *$[0]_n \cup [1]_n \cup \dots \cup [n-1]_n = \mathbb{Z}$.*

Siamo giunti alla definizione più importante di questo capitolo: l'insieme $\mathbb{Z}_n$.

Definizione A.3.7 *Si chiama insieme degli interi modulo n, e si indica con $\mathbb{Z}_n$, l'insieme delle classi di congruenza modulo n:*

$$\mathbb{Z}_n := \{[0]_n, [1]_n, \dots, [n-1]_n\}.$$

Definizione A.3.8 *Definiamo sull'insieme $\mathbb{Z}_n$ le seguenti operazioni di somma e prodotto:*

$$[a]_n + [b]_n := [a+b]_n, \qquad [a]_n\,[b]_n := [ab]_n.$$

Osservazione A.3.9 Le operazioni appena definite non dipendono dai numeri a e b che scegliamo per rappresentare le classi di congruenza che sommiamo o moltiplichiamo, ma soltanto dalla loro classe di congruenza. Si dice in tal caso che le operazioni sono *ben definite*. Ad esempio, in $\mathbb{Z}_4$ si ha: $[1]_4 = [5]_4$ e $[2]_4 = [6]_4$. Per definizione $[1]_4 + [2]_4 = [3]_4 = [11]_4 = [5]_4 + [6]_4$.

Esempio A.3.10 Calcoliamo le tavole dell'addizione e della moltiplicazione per $\mathbb{Z}_3$ e $\mathbb{Z}_4$ invitando lo studente ad esercitarsi a costruire le tavole analoghe per $\mathbb{Z}_5$ e $\mathbb{Z}_6$:

$\mathbb{Z}_3$

$+$	$[0]_3$	$[1]_3$	$[2]_3$
$[0]_3$	$[0]_3$	$[1]_3$	$[2]_3$
$[1]_3$	$[1]_3$	$[2]_3$	$[0]_3$
$[2]_3$	$[2]_3$	$[0]_3$	$[1]_3$

$\cdot$	$[0]_3$	$[1]_3$	$[2]_3$
$[0]_3$	$[0]_3$	$[0]_3$	$[0]_3$
$[1]_3$	$[0]_3$	$[1]_3$	$[2]_3$
$[2]_3$	$[0]_3$	$[2]_3$	$[1]_3$

$\mathbb{Z}_4$

$+$	$[0]_4$	$[1]_4$	$[2]_4$	$[3]_4$
$[0]_4$	$[0]_4$	$[1]_4$	$[2]_4$	$[3]_4$
$[1]_4$	$[1]_4$	$[2]_4$	$[3]_4$	$[0]_4$
$[2]_4$	$[2]_4$	$[3]_4$	$[0]_4$	$[1]_4$
$[3]_4$	$[3]_4$	$[0]_4$	$[1]_4$	$[2]_4$

$\cdot$	$[0]_4$	$[1]_4$	$[2]_4$	$[3]_4$
$[0]_4$	$[0]_4$	$[0]_4$	$[0]_4$	$[0]_4$
$[1]_4$	$[0]_4$	$[1]_4$	$[2]_4$	$[3]_4$
$[2]_4$	$[0]_4$	$[2]_4$	$[0]_4$	$[2]_4$
$[3]_4$	$[0]_4$	$[3]_4$	$[2]_4$	$[1]_4$

Notiamo alcuni fatti molto importanti: in $\mathbb{Z}_3$ ogni elemento diverso da $[0]_3$ ammette un *inverso*, cioè per ogni $[a]_3 \neq [0]_3$ esiste un elemento $[b]_3$ tale che $[a]_3[b]_3 = [1]_3$. Tale inverso si indica con $[a]_3^{-1}$. Abbiamo dunque: $[1]_3^{-1} = [1]_3$, $[2]_3^{-1} = [2]_3$. Tale proprietà invece non vale nel caso di $\mathbb{Z}_4$. Infatti la tavola moltiplicativa mostra che non esiste alcun inverso della classe $[2]_4$.

Proposizione A.3.11 *In $\mathbb{Z}_n$ l'elemento $[a]_n$ è invertibile se e solo se vale $M.C.D.(a, n) = 1$.*

Dimostrazione. Se $M.C.D.(a, n) = 1$, per l'Identità di Bézout esistono $x, y \in \mathbb{Z}$ tali che $1 = ax + ny$ cioè $[1]_n = [a]_n[x]_n$, quindi $[a]_n$ è invertibile con inverso $[x]_n$.

Se, viceversa, $[a]_n$ è invertibile, esiste $[x]_n \in \mathbb{Z}_n$ tale che $[a]_n[x]_n = [1]_n$, cioè $ax \equiv 1 \pmod{n}$ o, equivalentemente, $ax - kn = 1$ per qualche $k \in \mathbb{Z}$. Se fosse $M.C.D.(a, n) = d > 1$, d dividerebbe $ax - kn$ ma non dividerebbe 1 e questo, chiaramente, non è possibile. □

A.4 Congruenze

Definizione Si chiama congruenza lineare una relazione della forma

$$ax \equiv b \pmod{n}$$

con $a, b, n \in \mathbb{N}$.

Risolvere la congruenza lineare significa determinare $x \in \mathbb{Z}$ in modo tale che la relazione $ax \equiv b \pmod{n}$ sia soddisfatta. Notiamo che la relazione $ax \equiv b \pmod{n}$ è equivalente a $[a]_n[x]_n = [b]_n$. Quello che si fa dunque, in generale, è cercare le soluzioni distinte modulo n della congruenza assegnata.

Proposizione 1. Condizione necessaria affinché la congruenza

$$ax \equiv b \quad (mod\ n)$$

sia soddisfatta è che il $M.C.D.(a, n)$ divida b.
Dimostrazione. Se $ax \equiv b \quad (mod\ n)$, $ax - b = kn$ per un certo $k \in \mathbb{Z}$, cioè $b = ax - kn$. Quindi se $d|a, n$, $d|b$. □

Proposizione 2. Si consideri la congruenza $ax \equiv b \quad (mod\ n)$.

1) Se $M.C.D.(a, n) = 1$ allora la congruenza data ha una sola soluzione modulo n.

2) Se $M.C.D.(a, n) = d > 1$ e $d|b$ allora la congruenza data ha esattamente d soluzioni distinte modulo n.

Dimostrazione. Abbiamo già notato che la congruenza assegnata è equivalente alla relazione $[a]_n[x]_n = [b]_n$ in $\mathbb{Z}_n$. Nel caso 1), cioè se $M.C.D.(a, n) = 1$, sappiamo che $[a]_n$ è invertibile in $\mathbb{Z}_n$ perciò possiamo moltiplicare entrambi i membri dell'uguaglianza per $[a]_n^{-1}$ trovando una uguaglianza equivalente: $[x]_n = [a]_n^{-1}[b]_n$ il che mostra l'esistenza e l'unicità $(mod\ n)$ della soluzione cercata.

Ora supponiamo che sia $M.C.D.(a, n) = d > 1$ e che $d|b$ Allora abbiamo: $a = a'd$, $b = b'd$, $n = n'd$ con $a', b', n' \in \mathbb{N}$ e $M.C.D.(a', n') = 1$. La congruenza assegnata è equivalente a quella che otteniamo dividendo tutto per d, ossia:

$$a'x \equiv b' \quad (mod\ n').$$

Ora però $M.C.D.(a', n') = 1$ quindi, per il punto 1), tale congruenza ha un'unica soluzione modulo n': indichiamola semplicemente con $[c]_{n'}$. A noi interessano tuttavia le soluzioni modulo n perciò dobbiamo stabilire quante sono le classi di congruenza modulo n distinte degli elementi $c + kn'$. È facile osservare che si tratta esattamente delle classi degli elementi

$$c, c + n', \ldots, c + (d - 1)n'$$

dal momento che $dn' = d$. $\square$

Esempio. Risolvere la congruenza lineare $33x \equiv 5 \pmod{74}$.

Applichiamo l'algoritmo di Euclide:

$$\begin{aligned} 74 &= 2 \cdot 33 + 8 \\ 33 &= 4 \cdot 8 + 1 \\ 8 &= 8 \cdot 1 \end{aligned}$$

Dunque $M.C.D(33, 74) = 1$ quindi la soluzione della congruenza assegnata esiste ed è unica, e i conti appena fatti ci permettono di calcolare l'inverso di $[33]_{74}$. Abbiamo infatti che:

$$1 = 33 - 4 \cdot 8 = 33 - 4 \cdot (74 - 2 \cdot 33) = (-4) \cdot 74 + 9 \cdot 33.$$

Abbiamo perciò trovato: $1 = (-4) \cdot 74 + 9 \cdot 33$. Scriviamo questa equazione in $\mathbb{Z}_{74}$:

$$[1]_{74} = [9]_{74}[33]_{74}$$

Quindi la soluzione cercata è $x = [9]_{74} \cdot [5]_{74} = [45]_{74}$.

Esempio. Vogliamo determinare tutte le soluzioni della congruenza $63x \equiv 24 \pmod{375}$.

Calcoliamo $M.C.D.(63, 375)$:

$$\begin{aligned} 375 &= 5 \cdot 63 + 60 \\ 63 &= 1 \cdot 60 + 3 \\ 60 &= 20 \cdot 3 \end{aligned}$$

Dunqe $M.C.D.(63, 375) = 3$. Poiché 3 divide 24, la congruenza ammette soluzioni. Abbiamo che $63 = 3 \cdot 21$, $375 = 3 \cdot 125$, $24 = 3 \cdot 8$, quindi la congruenza data è equivalente a $[21]_{125}[x]_{125} = [8]_{125}$. Vogliamo trovare l'inverso di $[21]_{125}$, e par far ciò come prima cosa utilizziamo l'algoritmo di Euclide per calcolare $M.C.D.(125, 21)$.

$$\begin{aligned} 125 &= 5 \cdot 21 + 20 \\ 21 &= 1 \cdot 20 + 1 \\ 20 &= 20 \cdot 1 \end{aligned}$$

Procedendo a ritroso:

$$1 = 21 - 1 \cdot 20 = 21 - (125 - 5 \cdot 21) = -125 + 6 \cdot 21.$$

Otteniamo quindi che:

$$[1]_{125} = [6]_{125}[21]_{125}$$

Quindi la soluzione della congruenza $[21]_{125}[x]_{125} = [8]_{125}$ è $x = [6]_{125} \cdot [8]_{125} = [48]_{125}$. In altre parole le soluzioni della congruenza data sono della forma $x = 48 + k125$, con $k \in \mathbb{Z}$. Se vogliamo le soluzioni *mod* 375 abbiamo: $[48]_{375}$, $[173]_{375}$, $[298]_{375}$.

A.5 Esercizi proposti

1. Si dimostri che un insieme di n elementi ha 2^n sottoinsiemi.

2. Si dimostri (per induzione) che se n è un intero non negativo vale $2^n > n$.

3. Si considerino le due classi di congruenza $[0]_6$ e $[3]_{12}$. Si dica se se sono uguali o se sono diverse tra loro o se una è contenuta nell'altra.

4. Stabilire se gli elementi $[3]_{12}$, $[5]_{12}$, $[7]_{12}$ sono invertibili in $\mathbb{Z}_{12}$. In caso affermativo se ne calcolino gli inversi.

5. Si risolvano (se possibile) le seguenti congruenze lineari:

   i) $23x \equiv 7 \ (mod\ 40)$
   ii) $3x \equiv 13 \ (mod\ 17)$
   iii)$15x \equiv 9 \ (mod\ 53)$
   iv) $4x \equiv 1 \ (mod\ 400)$
   v) $18x \equiv 30 \ (mod\ 42)$
   vi) $16x \equiv 36 \ (mod\ 56)$
   vii) $20x \equiv 56 \ (mod\ 178)$