

Note di Algebra Lineare

Nicoletta Cantarini¹

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

Indice

1	Introduzione ai sistemi lineari	1
1.1	Sistemi lineari: primi esempi	1
1.2	Matrici	4
1.3	Algoritmo di Gauss	12
1.4	Esercizi svolti	18
1.5	Esercizi proposti	24
2	Spazi e sottospazi vettoriali	27
2.1	Premessa: l'insieme dei numeri reali	27
2.2	Lo spazio vettoriale $\mathbb{R}^n$	28
2.3	Sottospazi vettoriali	31
2.4	Generatori	34
2.5	Esercizi svolti	37
2.6	Esercizi proposti.	41
3	Lineare indipendenza e basi	43
3.1	Vettori linearmente indipendenti	43
3.2	Basi e dimensione	46
3.3	Strumenti di calcolo	49
3.4	Esercizi svolti	54
3.5	Esercizi proposti	57
4	Sottovarietà lineari	59
4.1	Lo spazio n -dimensionale	59
4.2	Lo spazio tridimensionale	66
4.3	Esercizi proposti	76

5	Concetti metrici	79
5.1	Prodotto scalare	79
5.2	Il procedimento di Gram-Schmidt	81
5.3	Proiezioni ortogonali	83
5.4	Esercizi svolti	88
5.5	Esercizi proposti	90
6	Spazi metrici	93
6.1	Spazio Euclideo	93
6.2	Perpendicolarità	94
6.3	Distanze tra sottovarietà lineari	97
6.4	Il prodotto vettoriale	103
6.5	Esercizi proposti	107
7	Applicazioni lineari e matrici	109
7.1	Applicazioni lineari	109
7.2	Matrici	112
7.3	Prodotto righe per colonne	112
7.4	Matrici invertibili	115
7.5	Esercizi svolti	120
7.6	Esercizi proposti	124

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 Architettura dell'Università di Bologna un supporto nella preparazione dell'esame di Istituzioni di Matematiche I.

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.

1.1 Sistemi lineari: primi esempi

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

$$a_1x_1 + a_2x_2 + \dots + a_nx_n = b, \quad (1.1)$$

dove $a_1, a_2, \dots, a_n$ e b sono numeri assegnati e $x_1, x_2, \dots, x_n$ sono le incognite. I numeri $a_1, \dots, a_n$ 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 $(s_1, s_2, \dots, s_n)$ che sostituiti ordinatamente alle incognite verificano l'uguaglianza, cioè tali che

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

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* $x_1, x_2, \dots, x_n$ è un insieme di m equazioni lineari nelle n incognite $x_1, x_2, \dots, x_n$ che devono

essere soddisfatte contemporaneamente:

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

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 1.1.1 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 1.1.2 *Un sistema si dice compatibile se ammette soluzioni.*

Esempio 1.1.3 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 1.1.4 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 1.1.5 *Due sistemi lineari si dicono equivalenti se hanno le stesse soluzioni.*

Nell'Esempio 1.1.4 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.

1.2 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 $M_{m,n}(\mathbb{R})$ l'insieme delle matrici $m \times n$ a coefficienti reali e semplicemente con $M_n(\mathbb{R})$ 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

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

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:

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

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 1.2.1 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:

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

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 1.2.2 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 1.2.3 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 1.2.2.

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 1.2.4 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 1.2.5 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 1.2.6 Sia $A \in M_{m,n}(\mathbb{R})$ una matrice a scala. Per definizione di rango si ha

$$rg(A) \leq m. \quad (1.3)$$

Vale però anche la disuguaglianza

$$rg(A) \leq n. \quad (1.4)$$

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 1.2.2 implica che il massimo numero di pivot di A è n .

Definizione 1.2.7 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 1.2.8 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 1.2.9 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 1.2.8, 1.2.9 è un fatto del tutto generale. Vale infatti la seguente proposizione:

Proposizione 1.2.10 *Sia $A\underline{x} = \underline{b}$ un sistema lineare a scala nelle n incognite $x_1, \dots, x_n$. 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 1.2.8, si ottiene l'unica soluzione del sistema. Se, invece, $rg(A) = rg(A|\underline{b}) = k < n$ è possibile, procedendo per sostituzioni successive dal basso, esprimere k variabili in funzione delle altre $n - k$ che restano libere di variare nell'insieme dei numeri reali. Si ottengono così infinite soluzioni. $\square$

Esempio 1.2.11 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 1.2.10(a), il sistema ammette soluzioni. Dal momento che il numero delle variabili è $4 > 2$, per la Proposizione 1.2.10(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:

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

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 .

1.3 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 1.3.1 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:

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

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

- (a) scambio di due equazioni;
 - (b) moltiplicazione di un'equazione per un numero reale diverso da 0;
 - (c) sostituzione della equazione i -esima con la somma dell'equazione i -esima e della j -esima moltiplicata per un numero reale α qualsiasi.
- In sintesi:

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

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

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

sono equivalenti.

Traduciamo ora le operazioni (a), (b) e (c) in termini matriciali: scambiare due equazioni del sistema equivale a scambiare due righe della matrice completa associata al sistema; moltiplicare una equazione per un numero reale diverso da 0 equivale a moltiplicare una riga della matrice completa associata al sistema per un numero reale diverso da 0, cioè moltiplicare per tale numero ogni elemento della riga; infine l'operazione (c) equivale a sostituire la riga i -esima della matrice completa associata al sistema con la somma della riga i -esima e della j -esima moltiplicata per un numero reale α . Spieghiamo un po' meglio che cosa si intende con tale somma: siano $(a_{i1} \dots a_{in} \ b_i)$ e $(a_{j1} \dots a_{jn} \ b_j)$ rispettivamente la i -esima e la j -esima riga della matrice $(A|b)$. 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 1.3.2 *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 1.3.3 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 1.3.4 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:

$$-2 \begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & 1 & -1 & 0 \\ 2 & -1 & 1 & 2 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & -5 & 1 & 0 \end{pmatrix}.$$

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:

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

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 1.2.10. Notiamo in particolare che, in conseguenza del ragionamento appena illustrato e del contenuto della Proposizione 1.2.10, 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 1.3.5 Le mosse dell'algoritmo di Gauss non sono necessariamente obbligate. Nell'Esempio 1.3.4, 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 1.3.6 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:

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

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):

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

Ora scambiamo la seconda con la quarta riga:

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

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

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

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

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

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 $w, y \in \mathbb{R}$.

1.4 Esercizi svolti

Esercizio 1.4.1 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:

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

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

$$\begin{aligned} \left(\begin{array}{cccc|c} 1 & -2 & 0 & 0 & 5 \\ 0 & 0 & -3 & 0 & 3 \\ 0 & -2 & 3 & -4 & -11 \\ 0 & 0 & -3 & 4 & 15 \end{array} \right) &\rightarrow \left(\begin{array}{cccc|c} 1 & -2 & 0 & 0 & 5 \\ 0 & -2 & 3 & -4 & -11 \\ 0 & 0 & -3 & 0 & 3 \\ 0 & 0 & -3 & 4 & 15 \end{array} \right) \\ &\rightarrow \left(\begin{array}{cccc|c} 1 & -2 & 0 & 0 & 5 \\ 0 & -2 & 3 & -4 & -11 \\ 0 & 0 & -3 & 0 & 3 \\ 0 & 0 & 0 & 4 & 12 \end{array} \right) = (A'|\underline{b}'). \end{aligned}$$

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 $rg(A') = rg(A'|\underline{b}') = 4 = \text{numero incognite}$. 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 1.4.2 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:

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

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

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

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

Come sappiamo la risposta viene fornita dalla Proposizione 1.2.10(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 1.4.3 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:

$$\left(\begin{array}{ccc|c} \alpha & \alpha+3 & 2\alpha & \alpha+2 \\ \alpha & 2\alpha+2 & 3\alpha & 2\alpha+2 \\ 2\alpha & \alpha+7 & 4\alpha & 2\alpha+4 \end{array} \right) \rightarrow \left(\begin{array}{ccc|c} \alpha & \alpha+3 & 2\alpha & \alpha+2 \\ 0 & \alpha-1 & \alpha & \alpha \\ 0 & -\alpha+1 & 0 & 0 \end{array} \right) \rightarrow$$

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

Se $\alpha \neq 0$ e $\alpha - 1 \neq 0$, cioè $\forall \alpha \in \mathbb{R} \setminus \{0, 1\}$, si ha $rg(A') = rg(A'|\underline{b}') = 3$ quindi, per la Proposizione 1.2.10, 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 $\alpha \in \mathbb{R} \setminus \{0, 1\}$ 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 1.4.4 Stabilire se esistono valori del parametro reale k tali che il sistema lineare

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

sia equivalente al sistema lineare

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

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

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

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

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

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

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

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

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

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

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

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

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

Riducendo questa matrice in forma a scala otteniamo la matrice:

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

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

1.5 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 e sottospazi vettoriali

2.1 Premessa: l'insieme dei numeri reali

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

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

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

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

La somma di numeri reali è:

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

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

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

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

Il prodotto di numeri reali è:

- commutativo: $ab = ba \quad \forall a, b \in \mathbb{R}$;
- associativo: $(ab)c = a(bc) \quad \forall a, b, c \in \mathbb{R}$;
- ammette elemento neutro, cioè esiste un numero, 1, tale che $1a = a1 = a \quad \forall a \in \mathbb{R}$;
- distributivo rispetto alla somma: $a(b + c) = ab + ac$.

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

2.2 Lo spazio vettoriale $\mathbb{R}^n$

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

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

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

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

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

$$+ : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}^2 \quad (\text{somma})$$

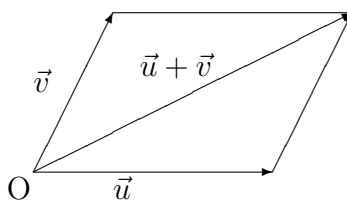
$$(x, y) + (x', y') := (x + x', y + y').$$

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

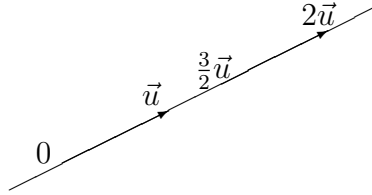
$$\lambda(x, y) := (\lambda x, \lambda y).$$

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

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



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



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

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

1. *commutativa*: $(x, y) + (x', y') = (x', y') + (x, y) \quad \forall (x, y), (x', y') \in \mathbb{R}^2$;
2. *associativa*: $((x, y) + (x', y')) + (x'', y'') = (x, y) + ((x', y') + (x'', y''))$
 $\forall (x, y), (x', y'), (x'', y'') \in \mathbb{R}^2$;
3. *esistenza dell'elemento neutro* $(0, 0)$: $(x, y) + (0, 0) = (0, 0) + (x, y) = (x, y) \quad \forall (x, y) \in \mathbb{R}^2$.
4. *esistenza dell'opposto*: per ogni (x, y) esiste un elemento (a, b) , detto *opposto* di (x, y) , tale che $(a, b) + (x, y) = (x, y) + (a, b) = (0, 0)$. Ovviamente si ha: $(a, b) = (-x, -y)$.

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

5. *Proprietà distributive:*

$$a((x, y) + (x', y')) = a(x, y) + a(x', y'), \forall (x, y), (x', y') \in \mathbb{R}^2 \text{ e } \forall a \in \mathbb{R}.$$

$$(a + b)(x, y) = a(x, y) + b(x, y), \forall (x, y) \in \mathbb{R}^2 \text{ e } \forall a, b \in \mathbb{R}.$$

6. $(ab)(x, y) = a(b(x, y)), \forall (x, y) \in \mathbb{R}^2 \text{ e } \forall a, b \in \mathbb{R}.$

7. $1(x, y) = (x, y), \forall (x, y) \in \mathbb{R}^2.$

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

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

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

$$+ : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$$

$$(x_1, \dots, x_n) + (x'_1, \dots, x'_n) := (x_1 + x'_1, \dots, x_n + x'_n);$$

$$\cdot : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$$

$$\lambda(x_1, \dots, x_n) := (\lambda x_1, \dots, \lambda x_n).$$

Con un po' di pazienza si possono verificare le proprietà 1. – 7. elencate sopra per la somma ed il prodotto per i numeri reali in $\mathbb{R}^2$. Chiameremo $\mathbb{R}^n$, con le operazioni di somma e di prodotto per i numeri reali appena definite, *spazio vettoriale*. Chiameremo *vettori* gli elementi di $\mathbb{R}^n$ e *scalari* i numeri reali. In tal modo il prodotto per i numeri reali si dirà, equivalentemente, prodotto per scalari. Chiameremo, infine, *vettore nullo* di $\mathbb{R}^n$, e lo indicheremo con $\mathbf{0}_{\mathbb{R}^n}$, l'elemento neutro rispetto alla somma, cioè: $\mathbf{0}_{\mathbb{R}^n} = (0, \dots, 0)$.

2.3 Sottospazi vettoriali

Definizione 2.3.1 *Sia W un sottoinsieme non vuoto di $\mathbb{R}^n$. Diremo che W è un sottospazio vettoriale di $\mathbb{R}^n$ se soddisfa le seguenti proprietà:*

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

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

Osservazione 2.3.2 Quanti elementi può avere un sottospazio vettoriale S di $\mathbb{R}^n$? Innanzitutto, per definizione, un sottospazio vettoriale di $\mathbb{R}^n$ non può essere vuoto. Notiamo inoltre che, per la proprietà 2) della Definizione 2.3.1, S contiene sempre almeno il vettore nullo di $\mathbb{R}^n$. Infatti, se $(x_1, \dots, x_n)$ è un elemento di S , per la proprietà 2), anche $0(x_1, \dots, x_n) = (0, \dots, 0)$ appartiene ad S .

È possibile che un sottospazio di $\mathbb{R}^n$ contenga SOLO $\mathbf{0}_{\mathbb{R}^n}$? Sì: il sottoinsieme $\{\mathbf{0}_{\mathbb{R}^n}\}$ di $\mathbb{R}^n$ è certamente chiuso rispetto alla somma e al prodotto per scalari, dal momento che si ha:

$$\mathbf{0}_{\mathbb{R}^n} + \mathbf{0}_{\mathbb{R}^n} = \mathbf{0}_{\mathbb{R}^n}$$

$$a\mathbf{0}_{\mathbb{R}^n} = \mathbf{0}_{\mathbb{R}^n}, \quad \forall a \in \mathbb{R}.$$

Chiameremo $\{\mathbf{0}_{\mathbb{R}^n}\}$ sottospazio banale di $\mathbb{R}^n$.

Supponiamo ora che V sia un sottospazio vettoriale di $\mathbb{R}^n$ NON banale, ossia che esso contenga almeno un vettore $v \neq \mathbf{0}_{\mathbb{R}^n}$. Quanti elementi contiene V ? Poiché disponiamo del prodotto per i numeri reali, che sono infiniti, V conterrà tutti gli infiniti multipli di v , cioè tutti i vettori della forma λv al variare di $\lambda \in \mathbb{R}$. Al variare di λ nell'insieme dei numeri reali si ottengono così infiniti elementi diversi di V .

Geometricamente quello che stiamo cercando di fare è molto semplice: stiamo cercando di costruire un modello algebrico per uno spazio di dimensione qualsiasi. Il concetto di dimensione, di cui daremo presto una definizione matematica rigorosa, è un concetto intuitivo che siamo abituati ad usare nella vita quotidiana. Sappiamo bene, ad esempio, cosa sia un film in 3D o che se la terra fosse piatta vivremmo su un piano infinito. Intuitivamente dunque è abbastanza chiaro che uno spazio di dimensione 0 è costituito solo da un punto (come lo spazio banale) ma nel momento in cui studiamo, ad esempio, il moto rettilineo di un punto, la traiettoria di questo punto sarà una retta (dimensione 1) ed una retta è costituita da infiniti punti, così come non appena lo spazio vettoriale è più grande del banale automaticamente contiene infiniti elementi.

In particolare, dunque, $\mathbb{R}^n$ possiede sempre almeno due sottospazi: $\mathbb{R}^n$ stesso e il sottospazio banale, costituito solo dal vettore nullo $\mathbf{0}_{\mathbb{R}^n}$.

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

Esempio 2.3.3 L'insieme $X = \{(x, y) \in \mathbb{R}^2 \mid y = 0\}$ è un sottospazio vettoriale di $\mathbb{R}^2$. X è infatti:

- 1) non vuoto: contiene le infinite coppie di numeri reali $(x, 0)$;
- 2) chiuso rispetto alla somma: presi due elementi $(x_1, 0)$, $(x_2, 0)$ in X , la loro somma $(x_1, 0) + (x_2, 0) = (x_1 + x_2, 0)$ appartiene ancora all'insieme X ;
- 3) chiuso rispetto al prodotto per scalari: presi un qualsiasi numero reale α ed un qualsiasi elemento $(x, 0) \in X$, il prodotto $\alpha(x, 0) = (\alpha x, 0)$ appartiene a X .

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

Esempio 2.3.4 Consideriamo l'insieme $S = \{(x, y) \in \mathbb{R}^2 \mid y = x + 1\}$. Possiamo immediatamente affermare che S NON è un sottospazio vettoriale di $\mathbb{R}^2$ poiché non contiene $\mathbf{0}_{\mathbb{R}^2} = (0, 0)$. Dunque non tutte le rette del piano individuano dei sottospazi vettoriali di $\mathbb{R}^2$ (ma solo quelle passanti per $(0, 0)$).

ATTENZIONE: se S è un sottospazio vettoriale di $\mathbb{R}^n$ la condizione $\mathbf{0}_{\mathbb{R}^n} \in S$ è necessaria ma non sufficiente affinché S sia un sottospazio di V . Facciamo subito un controesempio:

Esempio 2.3.5 Sia $S = \{(x, y, z) \in \mathbb{R}^3 \mid xy = z\}$. Nonostante l'insieme S contenga il vettore nullo $(0, 0, 0)$ di $\mathbb{R}^3$, S non è un sottospazio vettoriale di $\mathbb{R}^3$ perché non è chiuso rispetto alla somma. Infatti i vettori $v = (1, 1, 1)$ e $w = (-1, -1, 1)$ appartengono ad S , dal momento che soddisfano l'equazione $xy = z$, tuttavia la loro somma $v + w = (1, 1, 1) + (-1, -1, 1) = (0, 0, 2)$ non appartiene ad S dal momento che $0 \cdot 0 \neq 2$.

Osservazione 2.3.6 Dire che un sottoinsieme S di $\mathbb{R}^n$ è chiuso rispetto al prodotto per scalari, significa dire che se S contiene un vettore non nullo allora *deve contenere anche tutti i suoi multipli*. Se pensiamo alla interpretazione

geometrica del prodotto per scalari che abbiamo dato nel caso di $\mathbb{R}^2$, questo significa che se un sottospazio di $\mathbb{R}^2$ contiene un vettore non nullo v , allora esso contiene tutta la retta del piano individuata da v . Questo ragionamento geometrico ci permette di dire immediatamente che alcuni sottoinsiemi del piano che ben conosciamo NON sono spazi vettoriali, ad esempio

$$C = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$$

(una circonferenza),

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

(una parabola) e così via.

Osservazione 2.3.7 Gli esempi fatti finora mettono in luce due diversi tipi di ragionamento. Per dimostrare che un sottoinsieme S di $\mathbb{R}^n$ è un sottospazio vettoriale, bisogna dimostrare che valgono le proprietà 1) e 2) della Definizione 2.3.1. Tali proprietà devono valere *sempre* cioè per ogni coppia di vettori di S (proprietà 1)) e per tutti i numeri reali (proprietà 2)).

Per dimostrare, al contrario, che un sottoinsieme di $\mathbb{R}^n$ NON è un sottospazio vettoriale di $\mathbb{R}^n$, basta mostrare che una delle proprietà 1), 2) della Definizione 2.3.1 fallisce, ossia, se $S \neq \emptyset$, che esiste (anche solo) una coppia di vettori di S la cui somma non sta in S o che esistono (anche solo) un vettore di S ed uno scalare il cui prodotto non sta in S .

2.4 Generatori

Siano $v_1, v_2, \dots, v_k$ vettori di $\mathbb{R}^n$ e $\lambda_1, \lambda_2, \dots, \lambda_k$ numeri reali. Allora possiamo considerare, per ogni indice $i = 1, 2, \dots, k$, il vettore $\lambda_i v_i \in \mathbb{R}^n$ e poi fare la somma dei vettori ottenuti:

$$\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_k v_k.$$

Definizione 2.4.1 Dati k vettori $v_1, v_2, \dots, v_k$ di $\mathbb{R}^n$, si dirà che un vettore $v \in \mathbb{R}^n$ è loro combinazione lineare se esistono dei numeri reali $\lambda_1, \lambda_2, \dots, \lambda_k$ tali che

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_k v_k.$$

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

Esempio 2.4.2 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 $(1, 0), (0, 1)$;
3. il vettore nullo è combinazione lineare di qualunque insieme di vettori. Infatti se $v_1, \dots, v_k$ è un insieme di vettori di $\mathbb{R}^n$, allora $\mathbf{0}_{\mathbb{R}^n} = 0v_1 + 0v_2 + \dots + 0v_k$.

Che cosa succederebbe se tutti i vettori di $\mathbb{R}^n$ si potessero scrivere come combinazione lineare di un numero finito di vettori come nell'esempio 2 di 2.4.2?

Definizione 2.4.3 Diremo che $\mathbb{R}^n$ è generato dai vettori $v_1, \dots, v_k$ se ogni vettore di $\mathbb{R}^n$ è combinazione lineare di $v_1, \dots, v_k$.

Esempio 2.4.4 In $\mathbb{R}^3$ consideriamo i vettori $u_1 = (1, 1, -1)$, $u_2 = (0, 1, 2)$, $u_3 = (0, -2, 3)$. Vogliamo dimostrare che $\mathbb{R}^3$ è generato da $\{u_1, u_2, u_3\}$. 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

$$(\alpha, \beta, \gamma) = \lambda_1 u_1 + \lambda_2 u_2 + \lambda_3 u_3.$$

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

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

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 $\mathbb{R}^3$.

Osservazione 2.4.5 Siano $\{u_1, u_2, \dots, u_s\}$ generatori di $\mathbb{R}^n$. Allora ogni vettore v di $\mathbb{R}^n$ si scrive come loro combinazione lineare:

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

Tale scrittura è unica? In generale la risposta è no. Ad esempio i vettori $\{(1, 0, 0), (0, 1, 0), (0, 0, 1), (2, 1, 1)\}$ generano $\mathbb{R}^3$. Tuttavia notiamo che si ha, ad esempio:

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

come pure

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

Osservazione 2.4.6 Osserviamo che, dato un insieme di generatori $\{u_1, u_2, \dots, u_s\}$ di $\mathbb{R}^n$ e preso qualsiasi vettore $u \in \mathbb{R}^n$, i vettori $\{u_1, u_2, \dots, u_s, u\}$ pure generano $\mathbb{R}^n$. Ovviamente si possono continuare ad aggiungere vettori ottenendo sempre un insieme di generatori.

Definizione 2.4.7 Siano $\{u_1, u_2, \dots, u_k\}$ vettori di $\mathbb{R}^n$. Si chiama *sottospazio vettoriale generato da $u_1, \dots, u_k$* , e si indica con $\langle u_1, u_2, \dots, u_k \rangle$, l'insieme di tutte le combinazioni lineari di $u_1, u_2, \dots, u_k$:

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

Equivalentemente, diremo che i vettori $u_1, \dots, u_k$ generano $\langle u_1, u_2, \dots, u_k \rangle$.

Osservazione 2.4.8 Il contenuto di questa definizione va giustificato: dobbiamo infatti mostrare che l'oggetto appena definito è effettivamente un sottospazio vettoriale di $\mathbb{R}^n$. Poniamo $S = \langle u_1, u_2, \dots, u_k \rangle$. Dobbiamo mostrare che S è chiuso rispetto alla somma e al prodotto per scalari.

Consideriamo due combinazioni lineari $v = \alpha_1 u_1 + \alpha_2 u_2 + \dots + \alpha_k u_k$ e $w = \beta_1 u_1 + \beta_2 u_2 + \dots + \beta_k u_k$, allora la loro somma $v + w = (\alpha_1 u_1 + \alpha_2 u_2 + \dots + \alpha_k u_k) + (\beta_1 u_1 + \beta_2 u_2 + \dots + \beta_k u_k) = (\alpha_1 + \beta_1) u_1 + (\alpha_2 + \beta_2) u_2 + \dots + (\alpha_k + \beta_k) u_k$ è ancora una combinazione lineare dei vettori $u_1, u_2, \dots, u_k$, quindi appartiene ad S . Pertanto S è chiuso rispetto alla somma.

Siano ora $\lambda \in \mathbb{R}$ e $v = \alpha_1 u_1 + \alpha_2 u_2 + \dots + \alpha_k u_k$ una combinazione lineare di $u_1, u_2, \dots, u_k$. Allora $\lambda v = \lambda(\alpha_1 u_1 + \alpha_2 u_2 + \dots + \alpha_k u_k) = \lambda \alpha_1 u_1 + \lambda \alpha_2 u_2 + \dots +$

$\lambda \alpha_k u_k$ è pure una combinazione lineare di $u_1, u_2, \dots, u_k$ e quindi appartiene ad S . Dunque S è chiuso anche rispetto al prodotto per scalari ed è pertanto un sottospazio vettoriale di $\mathbb{R}^n$.

Osserviamo infine che se consideriamo un sottospazio T di $\mathbb{R}^n$ contenente i vettori $u_1, u_2, \dots, u_k$, per definizione di sottospazio, T contiene ogni loro combinazione lineare e quindi contiene anche tutto $S = \langle u_1, u_2, \dots, u_k \rangle$. Dunque S è il più piccolo sottospazio vettoriale di $\mathbb{R}^n$ contenente i vettori $u_1, \dots, u_k$.

Esempio 2.4.9 Il sottospazio vettoriale di $\mathbb{R}^n$ 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 .

2.5 Esercizi svolti

Esercizio 2.5.1 Stabilire se l'insieme $X := \{(r, s, r - s) \in \mathbb{R}^3\}$ è un sottospazio vettoriale di $\mathbb{R}^3$.

Svolgimento. Innanzitutto osserviamo che $(0, 0, 0) \in X$ (basta prendere $r = s = 0$).

Consideriamo ora due elementi generici di X : $(r_1, s_1, r_1 - s_1)$ e $(r_2, s_2, r_2 - s_2)$. La loro somma è: $(r_1, s_1, r_1 - s_1) + (r_2, s_2, r_2 - s_2) = (r_1 + r_2, s_1 + s_2, r_1 - s_1 + r_2 - s_2) = (r_1 + r_2, s_1 + s_2, r_1 + r_2 - (s_1 + s_2))$ e appartiene ancora ad X in quanto è del tipo $(r, s, r - s)$, con $r = r_1 + r_2$ e $s = s_1 + s_2$.

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

Esercizio 2.5.2 Si stabilisca se l'insieme $W := \{(x, y, z) \in \mathbb{R}^3 \mid 2x + z^2 = 0\} \subseteq \mathbb{R}^3$ è un sottospazio vettoriale di $\mathbb{R}^3$.

Svolgimento. L'insieme W non è un sottospazio vettoriale di $\mathbb{R}^3$ poiché non è chiuso rispetto alla somma. Infatti notiamo che, ad esempio, i vettori $(-2, 2)$ e $(-2, -2)$ appartengono a W dal momento che soddisfano entrambi l'equazione $2x + z^2 = 0$. Ma la loro somma: $(-2, 2) + (-2, -2) = (-4, 0)$ non appartiene a W poiché non soddisfa l'equazione $2x + z^2 = 0$.

Esercizio 2.5.3 Si determini un sottoinsieme non vuoto di $\mathbb{R}^2$ chiuso rispetto alla somma ma non al prodotto per scalari.

Svolgimento. L'insieme $X := \{(x, y) \in \mathbb{R}^2 | x \geq 0\}$ ha questa proprietà. Infatti X è non vuoto poiché, ad esempio, $(0, 0) \in X$. Inoltre X è chiuso rispetto alla somma: siano $(x_1, y_1), (x_2, y_2) \in X$, cioè tali che $x_1, x_2 \geq 0$. Allora $(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2) \in X$ perché $x_1 + x_2 \geq 0$ (la somma di due numeri reali non negativi è un numero reale non negativo).

Tuttavia X non è certamente chiuso rispetto al prodotto per scalari. Infatti, ad esempio, il vettore $(1, 0)$ appartiene ad X ma il vettore $-(1, 0) = (-1, 0) \notin X$.

Esercizio 2.5.4 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. Si tratta di stabilire se ogni vettore di $\mathbb{R}^3$ si può 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 $\lambda_1, \lambda_2, \lambda_3$ 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 caso l'insieme che ci viene proposto contiene solo due elementi. Potremmo procedere come prima cercando di risolvere un sistema, ma in qualche senso ci aspettiamo che due vettori non bastino a generare $\mathbb{R}^3$, 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 procedere come in (i) e, preso $(\alpha, \beta, \gamma) \in \mathbb{R}^3$, determinare $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ 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 2.5.5 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 $\lambda x + \lambda y + \lambda z = \lambda(x + y + z) = \lambda \cdot 0 = 0$. 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.

2.6 Esercizi proposti.

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

i) $S = \{(x, y, z) \in \mathbb{R}^3 \mid x = 0\}$.

ii) $T = \{(x, y, z) \in \mathbb{R}^3 \mid x = y\}$.

iii) $D = \{(a, 0, 0, d) \in \mathbb{R}^4\}$.

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

v) $X = \{(x, y, z) \in \mathbb{R}^3 \mid x + 2y + 2z = -3\}$.

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

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

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

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

4. 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.

Lezione 3

Lineare indipendenza e basi

3.1 Vettori linearmente indipendenti

Il problema che ci poniamo ora è: come possiamo stabilire quali sono gli eventuali vettori “superflui” nella descrizione di uno spazio vettoriale generato da un insieme di vettori? La risposta a questa domanda viene dal concetto di *indipendenza lineare*. Facciamo innanzitutto un esempio:

Esempio 3.1.1 Consideriamo il sottospazio vettoriale S di $\mathbb{R}^3$, $S = \langle (1, 0, 0), (0, 1, 0) \rangle$. Per definizione abbiamo:

$$S = \{(a, b, 0) \in \mathbb{R}^3\}.$$

Se eliminassimo uno tra i vettori $(1, 0, 0)$ e $(0, 1, 0)$ e considerassimo il sottospazio di $\mathbb{R}^3$ generato dal vettore rimasto, non otterremmo più S ma un sottospazio più piccolo:

$$\langle (1, 0, 0) \rangle = \{(a, 0, 0) \in \mathbb{R}^3\}$$

$$\langle (0, 1, 0) \rangle = \{(0, b, 0) \in \mathbb{R}^3\}.$$

Osserviamo ora che il vettore $(2, 1, 0)$ appartiene ad S , perciò avremmo potuto definire $S = \langle (1, 0, 0), (0, 1, 0), (2, 1, 0) \rangle$. Poiché $(2, 1, 0)$ è generato dai vettori $(1, 0, 0)$ e $(0, 1, 0)$, esso non aggiunge nulla all’insieme che già avevamo, vale a dire:

$$S = \langle (1, 0, 0), (0, 1, 0) \rangle = \langle (1, 0, 0), (0, 1, 0), (2, 1, 0) \rangle.$$

In altre parole se tra i vettori $u_1, \dots, u_k$ ce n'è uno che è combinazione lineare degli altri, lo si può scartare e lo spazio generato dai vettori rimasti coincide con $\langle u_1, \dots, u_k \rangle$. Al contrario se nessuno tra i vettori dati è combinazione lineare degli altri, nessun vettore può essere scartato senza determinare una perdita di informazioni.

Definizione 3.1.2 Sia V un sottospazio vettoriale di $\mathbb{R}^n$ (eventualmente $V = \mathbb{R}^n$):

- a) un vettore $v \in V$ si dice linearmente indipendente se $v \neq \mathbf{0}_{\mathbb{R}^n}$;
- b) k vettori $v_1, \dots, v_k$, con $k > 1$, si dicono linearmente indipendenti se nessuno di essi è combinazione lineare degli altri.

In caso contrario $v_1, \dots, v_k$ si dicono linearmente dipendenti.

Esempio 3.1.3 I vettori $\{(1, 0), (0, 1)\}$ di $\mathbb{R}^2$ sono linearmente indipendenti, poiché uno non è multiplo dell'altro.

Al contrario i vettori $\{(1, 0), (0, 1), (1, 1)\}$ sono linearmente dipendenti poiché

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

Osservazione 3.1.4 Siano $\{v_1, \dots, v_k\}$ vettori linearmente indipendenti di uno spazio vettoriale V . Che cosa succede eliminando vettori da questo insieme? Certamente i vettori che restano continueranno ad essere linearmente indipendenti: se nessuno dei vettori $v_1, \dots, v_k$ può essere espresso come combinazione lineare degli altri, questa proprietà continua a valere in un insieme di vettori più piccolo.

Che cosa succede, invece, aggiungendo un vettore v all'insieme $\{v_1, \dots, v_k\}$? I vettori $\{v_1, \dots, v_k, v\}$ continuano ad essere linearmente indipendenti se e solo se v non è combinazione lineare degli altri, vale a dire se e solo se $v \notin \langle v_1, \dots, v_k \rangle$.

Esempio 3.1.5 Per mostrare che i vettori di un insieme sono linearmente dipendenti è sufficiente mostrare che uno di essi è combinazione lineare degli altri. Ad esempio se un vettore è multiplo di un altro allora sappiamo subito che i vettori dell'insieme sono linearmente dipendenti. I vettori dei seguenti insiemi sono linearmente dipendenti e possiamo verificarlo senza fare troppi calcoli. Come?

- In $\mathbb{R}^3$: $\{(1, 0, 0), (0, 1, 0), (0, 0, 1), (0, 2, 0)\}$.
- In $\mathbb{R}^3$: $\{(1, 0, 0), (0, 1, 0), (-1, 2, 0), (1, 2, 1)\}$.

Attenzione. Un insieme di vettori contenente due vettori che siano uno multiplo dell'altro è un insieme di vettori linearmente dipendenti. Tuttavia *non è vero il contrario*: come mostra il secondo tra gli esempi appena fatti, possiamo avere un insieme di vettori linearmente dipendenti in cui *nessuno* è multiplo di un altro.

Proposizione 3.1.6 *Sia S un sottospazio vettoriale di $\mathbb{R}^n$ (eventualmente $S = \mathbb{R}^n$). Se k vettori $u_1, u_2, \dots, u_k$ di S , $k \in \mathbb{N}$, sono linearmente indipendenti allora ogni loro combinazione lineare si scrive in modo unico:*

$$\lambda_1 u_1 + \lambda_2 u_2 + \dots + \lambda_k u_k = \lambda'_1 u_1 + \lambda'_2 u_2 + \dots + \lambda'_k u_k, \quad \lambda_i, \lambda'_i \in \mathbb{R}$$

$$\Updownarrow$$

$$\lambda_1 = \lambda'_1; \lambda_2 = \lambda'_2; \dots; \lambda_k = \lambda'_k.$$

Dimostrazione. Per ipotesi sappiamo che i vettori $u_1, \dots, u_k$ sono linearmente indipendenti e vogliamo mostrare che ogni vettore in $\langle u_1, \dots, u_k \rangle$ si scrive in modo unico come combinazione lineare dei vettori $u_1, \dots, u_k$. Supponiamo di poter scrivere:

$$v = \lambda_1 u_1 + \lambda_2 u_2 + \dots + \lambda_k u_k = \lambda'_1 u_1 + \lambda'_2 u_2 + \dots + \lambda'_k u_k$$

per qualche $\lambda_i, \lambda'_i \in \mathbb{R}$, $i = 1, 2, \dots, k$. Se le due scritte fossero diverse, si avrebbe $\lambda_i \neq \lambda'_i$ per qualche $i = 1, \dots, n$. Tanto per fissare le idee supponiamo $\lambda_1 \neq \lambda'_1$. Allora possiamo scrivere

$$(\lambda_1 - \lambda'_1)u_1 = (\lambda'_2 - \lambda_2)u_2 + \dots + (\lambda'_k - \lambda_k)u_k,$$

cioè:

$$u_1 = \frac{(\lambda'_2 - \lambda_2)}{(\lambda_1 - \lambda'_1)}u_2 + \dots + \frac{(\lambda'_k - \lambda_k)}{(\lambda_1 - \lambda'_1)}u_k,$$

ma questo non è possibile perché i vettori $u_1, \dots, u_k$ sono linearmente indipendenti per ipotesi, perciò u_1 non può essere una combinazione lineare di $u_2, \dots, u_k$. **C. V. D.**

3.2 Basi e dimensione

Siano V un sottospazio vettoriale di $\mathbb{R}^n$ e $I = \{u_1, u_2, \dots, u_k\}$ un insieme di vettori di V . 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 un 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 un 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 linearmente indipendenti.

Definizione 3.2.1 *In un sottospazio vettoriale V di $\mathbb{R}^n$ un insieme (finito) ordinato di vettori $\{v_1, v_2, \dots, v_s\}$ si dice base di V se è un insieme di generatori linearmente indipendenti di V .*

Teorema 3.2.2 (senza dimostrazione). *Ogni base di uno spazio vettoriale V ha la medesima cardinalità, cioè lo stesso numero di elementi.*

Questo teorema ci permette di dare la seguente definizione:

Definizione 3.2.3 *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 3.2.4 *Ogni sottospazio vettoriale S di $\mathbb{R}^n$ ammette (almeno) una base. Più precisamente:*

1. ogni insieme di generatori di S contiene almeno una base di S ;
2. ogni insieme di vettori linearmente indipendenti di S si può completare ad una base.

Dimostrazione. Siano $\{u_1, \dots, u_k\}$ generatori di S . Se $\{u_1, \dots, u_k\}$ sono anche linearmente indipendenti, allora individuano una base di S . Altrimenti significa che uno di essi, supponiamo u_k , è combinazione lineare degli altri. Allora anche i vettori $\{u_1, \dots, u_{k-1}\}$ generano S . Se essi sono linearmente indipendenti, allora individuano una base di S , altrimenti uno di essi è combinazione lineare degli altri e lo si può scartare ottenendo un nuovo insieme di generatori di S . Questo ragionamento si può ripetere fino ad ottenere un insieme di generatori di S linearmente indipendenti, cioè una base di S . In questo modo abbiamo *estratto* una base dall'insieme di generatori di S da cui eravamo partiti.

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

Osservazione 3.2.5 La dimensione n di uno spazio vettoriale V è:

- a) il minimo numero di vettori che servono per generare V . Infatti, se così non fosse potrei costruire una base di V contenente meno di n elementi;
- b) il massimo numero di vettori linearmente indipendenti in V . Infatti, se così non fosse potrei costruire una base di V contenente più di n elementi.

Esempi 3.2.6 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, perché uno non è multiplo dell'altro. Dunque $B = \{e_1, e_2\}$ è una base di $\mathbb{R}^2$ e $\mathbb{R}^2$ ha pertanto dimensione due. La base B si dice base canonica di $\mathbb{R}^2$. Si noti che le coordinate di un vettore nella base canonica coincidono con le componenti del vettore.

Analogamente, per ogni $n \in \mathbb{N}$ definiamo i seguenti vettori di $\mathbb{R}^n$: $e_1 = (1, 0, \dots, 0)$, $e_2 = (0, 1, 0, \dots, 0)$, $\dots$, $e_n = (0, 0, \dots, 0, 1)$. Allora $\{e_1, e_2, \dots, e_n\}$ è una base di $\mathbb{R}^n$, detta base canonica. La dimen-

sione di $\mathbb{R}^n$ è dunque n . Dato il vettore $(x_1, x_2, \dots, x_n)$ di $\mathbb{R}^n$ vale:
 $(x_1, x_2, \dots, x_n) = x_1 e_1 + \dots + x_n e_n$.

2. Se un sottospazio L di $\mathbb{R}^n$ ha dimensione n allora $L = \mathbb{R}^n$. Infatti se L ha dimensione n allora esso contiene n vettori linearmente indipendenti: tali vettori risultano linearmente indipendenti anche in $\mathbb{R}^n$ e sono quindi una base di $\mathbb{R}^n$ stesso. Dunque lo spazio da essi generato è $L = \mathbb{R}^n$.

Osservazione 3.2.7 Sia $n = \dim(V)$ e siano $\{v_1, \dots, v_n\}$ vettori di V .

- i) Se $\{v_1, \dots, v_n\}$ generano V , allora $\{v_1, \dots, v_n\}$ sono automaticamente linearmente indipendenti. Se non lo fossero, infatti, uno di essi sarebbe combinazione lineare degli altri. Scartandolo otterremmo un insieme di generatori di V contenente meno di n elementi e questo non è possibile.
- ii) Se $\{v_1, \dots, v_n\}$ sono linearmente indipendenti, allora automaticamente generano V . Se così non fosse, infatti, esisterebbe un vettore v di V non appartenente al sottospazio $\langle v_1, \dots, v_n \rangle$. I vettori $\{v_1, \dots, v_n, v\}$ sarebbero dunque $n + 1$ vettori linearmente indipendenti in uno spazio vettoriale di dimensione n e questo non è possibile.

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

Definizione 3.2.8 Sia V un sottospazio vettoriale di $\mathbb{R}^n$ di dimensione k , e sia $\mathcal{B} = \{u_1, \dots, u_k\}$ una base di V . Allora per ogni vettore $v \in V$ sono univocamente determinati gli scalari $\lambda_i \in \mathbb{R}$ tali che $v = \sum_{i=1}^k \lambda_i u_i$; la k -upla $(\lambda_1, \dots, \lambda_k)$ si dice k -upla delle coordinate di v rispetto alla base $\mathcal{B}$.

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

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

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

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

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 $(\sigma_1, \dots, \sigma_k) \in \mathbb{R}^k$ e posto $w = \sigma_1 u_1 + \sigma_2 u_2 + \dots + \sigma_k u_k$, w è un elemento di V e $f(w) = (\sigma_1, \dots, \sigma_k)$. L'applicazione f è dunque biunivoca.

Esempio 3.2.9 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 $\mathbb{R}^n$.

3.3 Strumenti di calcolo

Cocentriamoci ora sulla seguente domanda: dati k vettori $v_1, \dots, v_k$ di $\mathbb{R}^n$, come facciamo a stabilire se questi sono o meno linearmente indipendenti? Si può rispondere a questa domanda utilizzando diversi metodi:

Metodo 1 Abbiamo notato nella Osservazione 3.1.4 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 3.3.1 Stabilire se i vettori $v_1 = (1, 2, 0, 1)$, $v_2 = (-1, 1, 1, 1)$, $v_3 = (3, 3, -1, 1)$, $v_4 = (-3, 0, 2, 1)$ sono linearmente indipendenti.

Utilizziamo il Metodo 1 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:

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

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

$$\begin{cases} \alpha - \beta = 3 \\ 2\alpha + \beta = 3 \\ \beta = -1 \\ \alpha + \beta = 1 \end{cases}$$

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.

Il metodo appena illustrato è piuttosto laborioso quando i vettori assegnati sono molti e/o dipendenti da parametri. Esiste infatti un metodo decisamente più efficace per stabilire se certi vettori sono o meno linearmente

indipendenti. Per illustrare questo secondo 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 3.3.2 *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 3.3.3 *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 3.3.4 Nella Lezione 1 abbiamo definito il rango di una matrice in forma a scala per righe. Naturalmente la definizione di rango (Definizione 1.2.4) 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 1 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 righe non nulle 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 3.3.5 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 3.3.6 Sia $A \in M_{m,n}(\mathbb{R})$. Le operazioni elementari sulle righe di A preservano il sottospazio di $\mathbb{R}^n$ generato dalle righe di A quindi preservano il rango.

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

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

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

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

Naturalmente

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

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

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

Metodo 2. 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 3.3.6, 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 3.3.6, 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 3.3.7 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 2 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 $v_1, \dots, v_4$:

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

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

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 2, facciamo ancora un altro esempio:

Esempio 3.3.8 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

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

Si ha:

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

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.

3.4 Esercizi svolti

Esercizio 3.4.1 Si verifichi che $\{(1, 2, 3), (0, 1, 0), (0, 0, 1)\}$ è una base di $\mathbb{R}^3$.

Svolgimento. Dobbiamo verificare che

- 1) i vettori dati sono linearmente indipendenti;
- 2) i vettori dati generano $\mathbb{R}^3$.

In effetti, poiché sappiamo che $\dim(\mathbb{R}^3) = 3$, basterà verificare una o l'altra proprietà (vedi Osservazione 3.2.7). Verifichiamo, ad esempio, che i tre vettori dati sono linearmente indipendenti. Cominciamo dal vettore $(0, 0, 1)$: esso è non nullo perciò linearmente indipendente; $(0, 1, 0)$ non è un multiplo di $(0, 0, 1)$, perciò $(0, 0, 1)$ e $(0, 1, 0)$ sono linearmente indipendenti. Infine $(1, 2, 3)$ non è una combinazione lineare di $(0, 1, 0)$ e $(0, 0, 1)$ dal momento che combinando linearmente questi due vettori si ottiene un vettore la cui prima componente è nulla. Possiamo quindi concludere che i tre vettori assegnati sono linearmente indipendenti e, dunque, individuano una base di $\mathbb{R}^3$.

Esercizio 3.4.2 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 linearmente 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 2.5.4. Infine si può mostrare che l'insieme $\{(0, 1, 1), (1, 0, -1), (2, 1, 0)\}$ è una base di $\mathbb{R}^3$.

Esercizio 3.4.3 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 $\mathcal{B}$ 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 3.4.4 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 3.4.5 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:

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

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

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:

$$\begin{aligned} (x_2 + 2x_5, x_2, -x_4 - x_5, x_4, x_5) &= x_2(1, 1, 0, 0, 0) + \\ &+ x_5(2, 0, -1, 0, 1) + x_4(0, 0, -1, 1, 0). \end{aligned}$$

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 $\alpha, \beta, \gamma \in \mathbb{R}$ 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 $v = (0, -2, 1)_{\mathcal{B}}$.

3.5 Esercizi proposti

Esercizio 3.5.1 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 3.5.2 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 3.5.3 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 $\mathbb{R}^3$.

Lezione 4

Sottovarietà lineari

In questa lezione vogliamo dare una interpretazione geometrica ai risultati illustrati nelle precedenti lezioni e completare la costruzione di un modello algebrico per uno spazio di dimensione qualsiasi.

4.1 Lo spazio n -dimensionale

Cominciamo con qualche considerazione su un paio di esempi:

Esempio 4.1.1 Risolviamo il sistema lineare

$$\begin{cases} x + y + z = 1 \\ x - y + 2z = 2 \end{cases}$$

come abbiamo imparato a fare nella Lezione 1. La matrice completa associata al sistema è

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

Per le Proposizioni 1.2.10 e 3.3.6, il sistema lineare assegnato ha soluzioni se e solo se $rg(A) = rg(A|\underline{b})$. Riduciamo la matrice in forma a scala per determinare le soluzioni del sistema:

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & -1 & 2 & 2 \end{array} \right) \rightarrow \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & -2 & 1 & 1 \end{array} \right).$$

Si ha dunque $rg A = rg(A|\underline{b}) = 2$, pertanto il sistema ammette infinite soluzioni dipendenti da una variabile. Il sistema di partenza è equivalente al

sistema lineare associato alla matrice ridotta:

$$\begin{cases} x + y + z = 1 \\ -2y + z = 1 \end{cases}$$

che può essere risolto per sostituzioni successive dal basso: $z = 1 + 2y$, $x = 1 - y - z = 1 - y - 1 - 2y = -3y$. L'insieme delle soluzioni è pertanto:

$$S = \{(-3y, y, 1 + 2y) \in \mathbb{R}^3\} = (0, 0, 1) + \langle(-3, 1, 2)\rangle.$$

Con la somma $(0, 0, 1) + \langle(-3, 1, 2)\rangle$ intendiamo il sottoinsieme di $\mathbb{R}^3$ costituito dai vettori che si scrivono come somma del vettore $(0, 0, 1)$ e di un qualsiasi multiplo del vettore $(-3, 1, 2)$, quindi esattamente ogni elemento di S .

Notiamo che il sottospazio $\langle(-3, 1, 2)\rangle$ è l'insieme delle soluzioni del sistema lineare omogeneo associato al sistema da cui siamo partiti:

$$\begin{cases} x + y + z = 0 \\ x - y + 2z = 0 \end{cases}$$

Esempio 4.1.2 Cosa succede se risolviamo semplicemente l'equazione

$$x + y + z = 1?$$

In tal caso possiamo usare l'equazione assegnata per esprimere una incognita in funzione delle altre due: $z = 1 - x - y$. L'insieme delle soluzioni dell'equazione assegnata dipende dunque da due variabili libere ed è:

$$T = \{(x, y, 1 - x - y) \in \mathbb{R}^3\} = (0, 0, 1) + \langle(1, 0, -1), (0, 1, -1)\rangle.$$

Anche in questo caso abbiamo indicato con $(0, 0, 1) + \langle(1, 0, -1), (0, 1, -1)\rangle$ la somma del vettore $(0, 0, 1)$ e di qualunque vettore appartenente al sottospazio $\langle(1, 0, -1), (0, 1, -1)\rangle$. Notiamo che il sottospazio $\langle(1, 0, -1), (0, 1, -1)\rangle$ è l'insieme delle soluzioni dell'equazione omogenea associata all'equazione di partenza:

$$x + y + z = 0.$$

In ciascuno degli esempi appena fatti abbiamo descritto l'insieme delle soluzioni del sistema lineare assegnato come la 'somma' di un vettore fissato e di un sottospazio vettoriale di $\mathbb{R}^3$. Qual è l'ambiente giusto per leggere ed interpretare geometricamente i risultati ottenuti?

Definizione 4.1.3 Sia $A\underline{x} = \underline{b}$ un sistema lineare in cui $\text{rg}(A) = \text{rg}(A|\underline{b}) = k$. Diremo in tal caso che il sistema $A\underline{x} = \underline{b}$ ha rango k .

Definizione 4.1.4 Chiameremo spazio n -dimensionale il dato di:

- un insieme di punti X che identificheremo con $\mathbb{R}^n$;
- lo spazio vettoriale $\mathbb{R}^n$;
- un'azione di $\mathbb{R}^n$ su X cioè una funzione:

$$\begin{aligned} X \times \mathbb{R}^n &\longrightarrow X \\ (P, v) &\longmapsto Q = P + v, \end{aligned}$$

essendo $Q = P + v$ il punto che si raggiunge applicando il vettore v nel punto P .

In altre parole, dati un punto $P = (x_1, \dots, x_n)$ ed un vettore $v = (\alpha_1, \dots, \alpha_n)$, associamo ad essi un punto Q , indicato con $P + v$, secondo la seguente legge:

$$Q = P + v = (x_1, \dots, x_n) + (\alpha_1, \dots, \alpha_n) = (x_1 + \alpha_1, \dots, x_n + \alpha_n).$$

Attenzione! La somma che abbiamo scritto è in effetti una strana somma: è la somma di un punto e di un vettore e produce un punto. Schematizzando con un disegno:

$$\begin{array}{ccc} & v & Q \\ & \nearrow & \\ P & & \end{array}$$

Vediamo alcune proprietà di questa legge.

i) Presi due punti P e Q in X esiste un unico vettore $v \in \mathbb{R}^n$ tale che sia $P + v = Q$. Indicheremo il vettore v con $Q - P$. (Attenzione: questa è solo una scrittura e indica un vettore di $\mathbb{R}^n$). Possiamo determinare il vettore $Q - P$ esplicitamente: se $P = (x_1, \dots, x_n)$ e $Q = (y_1, \dots, y_n)$, allora $Q - P = (y_1 - x_1, \dots, y_n - x_n)$ (ed è evidente che $Q - P$ è il vettore che sommato a P dà Q). Nello stesso modo $P - Q = (x_1 - y_1, \dots, x_n - y_n)$ è il vettore che sommato a Q dà P .

Esempio: nel piano (cioè nello spazio bidimensionale) siano $P = (2, -1)$, $Q = (3, 2)$, allora $P - Q = (-1, -3)$. Infatti $Q + (P - Q) = (3, 2) + (-1, -3) = (2, -1) = P$.

ii) Se si ha $P + v = P + w$ allora $v = w$. In particolare il solo vettore che sommato a P lo lascia invariato è il vettore nullo di $\mathbb{R}^n$.

Le proprietà *i)* e *ii)* evidenziano il fatto che, fissato un punto $L \in X$, possiamo definire una applicazione che ad ogni punto $Q \in X$ associa il vettore $Q - L$:

$$\begin{aligned}\alpha_L : X &\longrightarrow \mathbb{R}^n \\ Q &\longmapsto \alpha_L(Q) = Q - L.\end{aligned}$$

Questa funzione è una biiezione.

iii) Se prendiamo due vettori $v, w \in \mathbb{R}^n$ e un punto $P \in X$ allora $P + v$ è un punto di X a cui possiamo sommare un altro vettore w : otteniamo il punto $(P + v) + w$. Tale punto, come risulta evidente da una scrittura in coordinate, coincide con il punto che si ottiene sommando a P il vettore $v + \mathbb{R}^n w$.

Definizione 4.1.5 *Chiameremo lo spazio bidimensionale **piano**.*

Così come studiando gli spazi vettoriali abbiamo introdotto dei sottooggetti (i sottospazi vettoriali), ora definiremo i sottooggetti dello spazio n -dimensionale: le cosiddette **sottovarietà lineari**.

Definizione 4.1.6 *Una sottovarietà lineare L dello spazio n -dimensionale è l'insieme*

$$\{P + v \mid v \in W\}$$

dove P è un punto fissato e W è un sottospazio vettoriale di $\mathbb{R}^n$. Il sottospazio W si dice **giacitura o spazio direttore** della sottovarietà lineare L . La dimensione di W viene detta **dimensione della sottovarietà lineare**. Un punto è dunque una sottovarietà lineare di dimensione zero. Chiameremo **retta** una sottovarietà lineare di dimensione uno, **piano** una sottovarietà lineare di dimensione due.

4.1.1 Sottovarietà lineari del piano

Sia $L = P + W$ una sottovarietà lineare del piano, dove P è un punto fissato della sottovarietà e W è un sottospazio vettoriale di $\mathbb{R}^2$. Quale può essere la dimensione di W ? Essendo $\dim \mathbb{R}^2 = 2$, W può avere dimensione 0, 1 oppure 2. Se $\dim W = 0$, allora $L = P$ è un punto. Se $\dim W = 2$ allora L è tutto il piano. Resta da studiare il caso in cui L sia una sottovarietà lineare di dimensione uno, cioè una retta. Consideriamo dunque la retta r passante per il punto $P = (x_0, y_0)$, avente giacitura generata dal vettore $v = (a, b) \in \mathbb{R}^2$, $v \neq (0, 0)$:

$$r : P + \langle v \rangle = (x_0, y_0) + \langle (a, b) \rangle.$$

I punti di r sono tutti e soli della forma:

$$\begin{cases} x = x_0 + \lambda a \\ y = y_0 + \lambda b \end{cases} \quad (4.1)$$

al variare di $\lambda \in \mathbb{R}$. Le equazioni (4.1) si dicono *equazioni parametriche* della retta r . Al variare del parametro reale λ esse forniscono le coordinate di tutti e soli i punti di r .

Supponiamo $a \neq 0 \neq b$ ed eliminiamo il parametro dalle equazioni (4.1):

$$\begin{cases} x - x_0 = \lambda a \\ y - y_0 = \lambda b \end{cases} \Leftrightarrow \begin{cases} b(x - x_0) = \lambda ab \\ a(y - y_0) = \lambda ab \end{cases} \Leftrightarrow b(x - x_0) = a(y - y_0).$$

La retta r può dunque essere descritta mediante l'equazione lineare

$$bx - ay = bx_0 - ay_0. \quad (4.2)$$

Notiamo che se $a = 0$ allora $b \neq 0$ perché, per ipotesi, $(a, b) \neq 0$. In tal caso l'equazione (4.2) è equivalente all'equazione $x = x_0$ che descrive la retta r nel caso in cui sia $a = 0$. Analogamente se $b = 0$, allora $a \neq 0$ e l'equazione di r diventa semplicemente $y = y_0$.

L'equazione (4.2) si chiama *equazione cartesiana* della retta r . Osserviamo che essa è ovviamente soddisfatta dalle coordinate del punto $P = (x_0, y_0)$, che è un punto della retta, e che l'equazione omogenea ad essa associata:

$$bx - ay = 0$$

è l'equazione del sottospazio direttore $\langle (a, b) \rangle$ di r .

Esempio 4.1.7 In $\mathbb{A}^2(\mathbb{R})$ consideriamo la retta t passante per il punto $P = (1, 4)$, avente spazio direttore $V = \langle (1, -1) \rangle$. I punti di t sono tutti e soli della forma $(1, 4) + v$, dove $v \in V$. Dal momento che i vettori di V sono i multipli di $(1, -1)$, i punti $(1, 4) + (1, -1) = (2, 3) = P_1$ e $(1, 4) + (4, -4) = (5, 0) = P_2$ appartengono alla retta.

P
 P_1

P_2

Determiniamo ora equazioni parametriche e cartesiane della retta t : le sue equazioni cartesiane sono:

$$\begin{cases} x = 1 + \lambda \\ y = 4 - \lambda \end{cases} \quad (\lambda \in \mathbb{R}).$$

Ricaviamo il parametro λ dalla prima equazione: $\lambda = x - 1$ e sostituiamolo nella seconda equazione. Otteniamo così l'equazione lineare $y = 5 - x$ che è l'equazione cartesiana della retta t .

Naturalmente ogni equazione lineare equivalente a quella trovata descrive la stessa retta, cioè l'equazione di una retta è individuata a meno di un multiplo dei suoi coefficienti.

Notiamo che lo spazio direttore della retta t , cioè il sottospazio di $\mathbb{R}^2$ $\langle(1, -1)\rangle$, è l'insieme delle soluzioni dell'equazione lineare omogenea associata all'equazione trovata: $x + y = 0$. Scopriremo più avanti come questo non sia un caso fortuito.

Esercizio 4.1.8 Determinare l'equazione cartesiana della retta r che passa per i due punti $P_1 = (1, 3)$ e $P_2 = (3, -1)$.

La retta r passa per il punto P_1 e ha giacitura generata dal vettore $P_2 - P_1 = (2, -4)$: $r : (1, 3) + \langle(2, -4)\rangle = (1, 3) + \langle(1, -2)\rangle$. Le equazioni cartesiane di r sono dunque:

$$\begin{cases} x = 1 + \lambda \\ y = 3 - 2\lambda \end{cases} \quad (\lambda \in \mathbb{R}).$$

Dalla prima equazione otteniamo $\lambda = x - 1$ che, sostituito nella seconda equazione, dà: $y = 5 - 2x$. L'equazione cartesiana della retta r è quindi $2x + y - 5 = 0$.

Definizione 4.1.9 Due sottovarietà lineari $L_1 = P + V_1$ e $L_2 = Q + V_2$ dello spazio n -dimensionale si dicono **parallele** se uno dei due spazi direttori è contenuto nell'altro (i.e. se $V_1 \leq V_2$ oppure $V_2 \leq V_1$).

Si osservi che questa definizione di parallelismo risulta un po' diversa dalle solite definizioni intuitive. Ad esempio, secondo la definizione 4.1.9, un punto è parallelo a qualsiasi sottovarietà lineare dal momento che il suo spazio direttore è banale. Due sottovarietà lineari della stessa dimensione, invece, sono parallele se e solo se hanno lo stesso spazio direttore. In particolare due sottovarietà coincidenti sono parallele.

Esercizio 4.1.10 *La mutua posizione di due rette nel piano.* Utilizziamo ora la teoria dei sistemi lineari per studiare la posizione reciproca di due rette nel piano. Consideriamo dunque le rette $r_1 : a_1x_1 + b_1x_2 = c_1$ e $r_2 : a_2x_1 + b_2x_2 = c_2$ con $(a_1, b_1) \neq (0, 0)$ e $(a_2, b_2) \neq (0, 0)$. Gli eventuali punti di intersezione fra le due rette sono punti le cui coordinate soddisfano contemporaneamente le equazioni di r_1 e r_2 cioè il sistema:

$$\begin{cases} a_1x_1 + b_1x_2 = c_1 \\ a_2x_1 + b_2x_2 = c_2, \end{cases}$$

i.e.,

$$\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}.$$

Osserviamo che $\text{rg} \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} \geq 1$ dal momento che ciascuna delle due righe della matrice è diversa dalla riga nulla! (N.B. Affinché una matrice abbia rango maggiore di uno basta che abbia una entrata diversa da zero). Indichiamo il nostro sistema con $A\underline{x} = \underline{c}$. Analizziamo i casi possibili:

- $\text{rg}A = 1$: in questo caso la seconda riga di A è combinazione lineare della prima, quindi l'equazione dello spazio direttore della seconda retta è un multiplo scalare dell'equazione dello spazio direttore della prima. Questo significa che le due rette hanno il medesimo spazio direttore e sono quindi parallele. Adesso guardiamo il sistema totale. Se la matrice completa ha rango 2, il sistema non ha soluzioni ($\text{rg}(A) \neq \text{rg}(A|\underline{b})$). Le due rette sono parallele e non hanno punti in comune. Se la matrice completa ha rango 1 il sistema ha infinite soluzioni e queste soluzioni dipendono da una sola variabile. L'insieme delle soluzioni è della forma: $P + \langle v \rangle$. Questo significa che l'insieme dei punti comuni alle due rette è una retta e questa coincide necessariamente con ciascuna delle rette date: le due rette coincidono. Il fatto che la matrice completa abbia rango 1 ci assicura che le equazioni delle due rette siano proporzionali.

- $\text{rg}A = 2$: in questo caso la matrice completa, che è una matrice 2×3 , ha necessariamente rango 2: il suo rango è infatti sempre maggiore del rango della matrice incompleta o uguale ad esso ma è minore o uguale a due. Quindi il sistema ha una ed una sola soluzione. Questo significa che le due rette si intersecano in un punto. Diremo in questo caso che le rette sono *incidenti*.

4.2 Lo spazio tridimensionale

Studiamo ora le sottovarietà lineari $L = P + T$ dello spazio tridimensionale, con P punto dello spazio e $T \leq \mathbb{R}^3$. Essendo T un sottospazio vettoriale di $\mathbb{R}^3$, esso può avere dimensione 0, 1, 2 oppure 3. Se $\dim T = 0$ la varietà sarà costituita da un solo punto e se $\dim T = 3$ essa coinciderà con tutto lo spazio. Come possiamo descrivere rette e piani nello spazio tridimensionale?

4.2.1 *I piani nello spazio tridimensionale.* In questo caso T ha dimensione due e quindi è generato da due vettori linearmente indipendenti. I punti del piano $\pi : (x_0, y_0, z_0) + \langle (m_1, n_1, t_1), (m_2, n_2, t_2) \rangle$, dove i vettori (m_1, n_1, t_1) e (m_2, n_2, t_2) sono linearmente indipendenti, possono essere descritti mediante le seguenti equazioni dipendenti dai parametri reali λ e μ :

$$\begin{cases} x = x_0 + \lambda m_1 + \mu m_2 \\ y = y_0 + \lambda n_1 + \mu n_2 \\ z = z_0 + \lambda t_1 + \mu t_2. \end{cases} \quad (4.3)$$

Le equazioni (4.3) si dicono *equazioni parametriche* del piano π . Possiamo utilizzare due delle equazioni (4.3) per esprimere ciascuno dei parametri λ e μ in funzione delle variabili. Sostituendo tali espressioni di λ e μ nella terza equazione si ottiene la cosiddetta *equazione cartesiana* del piano π , vale a dire una equazione lineare nelle incognite x, y e z avente come soluzioni tutti e soli i punti del piano π . Facciamo un esempio: determiniamo equazioni parametriche e cartesiane del piano $\pi : (1, 2, 1) + \langle (1, 0, 1), (0, 1, 2) \rangle$. Le equazioni parametriche di π sono:

$$\begin{cases} x = 1 + \lambda \\ y = 2 + \mu \\ z = 1 + \lambda + 2\mu \end{cases} \quad (\lambda, \mu \in \mathbb{R}).$$

Dalla prima equazione otteniamo $\lambda = x - 1$ e dalla seconda equazione otteniamo $\mu = y - 2$. Sostituendo queste espressioni di λ e μ nella terza equazione otteniamo $z = 1 + x - 1 + 2y - 4$, cioè $x + 2y - z = 4$ che è l'equazione cartesiana del piano π .

Notiamo che, ovviamente, le coordinate del punto $(1, 2, 1)$ soddisfano l'equazione del piano π e che lo spazio direttore del piano, cioè il sottospazio $\langle (1, 0, 1), (0, 1, 2) \rangle$ di $\mathbb{R}^3$, è l'insieme delle soluzioni dell'equazione omogenea $x + 2y - z = 0$.

Proponiamo ora un altro metodo per determinare l'equazione cartesiana di un piano.

Esempio 4.2.2 Si determini l'equazione cartesiana del piano $\pi : (1, 0, 1) + \langle (1, 1, 1), (0, 1, 1) \rangle$. Dobbiamo descrivere i punti $(x, y, z) \in (1, 0, 1) + \langle (1, 1, 1), (0, 1, 1) \rangle$, cioè i punti (x, y, z) tali che $(x, y, z) - (1, 0, 1) \in \langle (1, 1, 1), (0, 1, 1) \rangle$. Questo equivale a determinare i punti (x, y, z) tali che la matrice

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ x-1 & y & z-1 \end{pmatrix}$$

abbia rango 2. Riduciamo la matrice in forma a scala:

$$rg \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & y+1-x & z-x \end{pmatrix} = rg \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & z-y-1 \end{pmatrix} = 2 \Leftrightarrow z-y=1.$$

L'equazione cartesiana del piano π è dunque $z-y=1$.

Notiamo che le coordinate del punto $(1, 0, 1)$ soddisfano ovviamente l'equazione trovata, essendo $(1, 0, 1)$ un punto di π . L'equazione omogenea $z-y=0$, d'altra parte, è proprio l'equazione della giacitura del piano, cioè l'equazione del sottospazio $\langle (1, 1, 1), (0, 1, 1) \rangle$.

Esempio 4.2.3 Nello spazio tridimensionale vogliamo descrivere il piano di equazione $3x-y=5$. Occorre quindi descrivere l'insieme delle soluzioni dell'equazione data, nella forma $(x_0, y_0, z_0) + W$, essendo W un sottospazio di $\mathbb{R}^3$ di dimensione 2. L'insieme delle soluzioni dell'equazione $3x-y=5$ è $\{(x, 3x-5, z) \in \mathbb{R}^3\} = (0, -5, 0) + \langle (1, 3, 0), (0, 0, 1) \rangle$. Il piano cercato passa dunque per il punto $(0, -5, 0)$ ed ha giacitura $\langle (1, 3, 0), (0, 0, 1) \rangle$.

4.2.4 Posizione reciproca di due piani nello spazio tridimensionale. Si tratta di studiare il sistema lineare

$$\begin{cases} ax + by + cz = d \\ a'x + b'y + c'z = d' \end{cases}$$

con $(a, b, c) \neq (0, 0, 0)$ e $(a', b', c') \neq (0, 0, 0)$. Possiamo scrivere il sistema in forma matriciale:

$$\begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} d \\ d' \end{pmatrix},$$

i.e., $A\underline{x} = \underline{d}$. Abbiamo due possibilità:

- $\text{rg}A = 1$. In questo caso le equazioni delle giaciture dei due piani sono proporzionali. I due piani sono paralleli (Definizione 4.1.9): ciò vuol dire che hanno la stessa giacitura. Per la matrice completa ci sono due possibilità: se essa ha rango due allora il sistema non ha soluzioni, e quindi i due piani sono paralleli senza punti in comune. Se la matrice completa ha rango 1 il sistema ammette infinite soluzioni dipendenti da due variabili libere: l'insieme delle soluzioni del sistema è un piano e i due piani di partenza coincidono.

- $\text{rg}A = 2$. I due piani non sono paralleli e il rango della matrice completa è necessariamente uguale a due. Quindi il sistema ammette soluzioni. Il sistema omogeneo associato è un sistema a 3 incognite, di rango due: l'insieme delle sue soluzioni dipende da una variabile libera. L'intersezione dei due piani è pertanto una varietà lineare di dimensione uno: una retta.

4.2.5 Le rette in $\mathbb{A}^3(\mathbb{R})$. In 4.2.4 abbiamo visto che due piani non paralleli si intersecano lungo una retta. È dunque possibile descrivere una retta in $\mathbb{A}^3(\mathbb{R})$ attraverso un sistema lineare in 3 incognite di rango 2:

$$\begin{cases} ax + by + cz = 0 \\ a'x + b'y + c'z = 0 \end{cases}$$

con $\text{rg} \begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix} = 2$. Tali equazioni si chiamano *equazioni cartesiane* della retta nello spazio affine tridimensionale.

Determinare esplicitamente le equazioni cartesiane di una retta $r : (x_0, y_0, z_0) + \langle (m, n, p) \rangle$ in $\mathbb{A}^3(\mathbb{R})$, significa determinare i punti (x, y, z) tali che $(x, y, z) - (x_0, y_0, z_0) \in \langle (m, n, p) \rangle$, cioè tali che la matrice

$$\begin{pmatrix} m & n & p \\ x - x_0 & y - y_0 & z - z_0 \end{pmatrix}$$

abbia rango 1. Riducendo la matrice in forma a scala ed imponendo che la matrice ridotta abbia rango uno si ottiene un sistema di equazioni cartesiane per r .

Ad esempio, determiniamo le equazioni cartesiane della retta $r : (1, 1, 1) + \langle (1, 2, 0) \rangle$. Usando il procedimento appena descritto, si tratta di determinare i punti (x, y, z) tali che la matrice

$$\begin{pmatrix} 1 & 2 & 0 \\ x - 1 & y - 1 & z - 1 \end{pmatrix}$$

abbia rango 1. Riducendo la matrice in forma a scala otteniamo:

$$\begin{aligned} \operatorname{rg} \begin{pmatrix} 1 & 2 & 0 \\ x-1 & y-1 & z-1 \end{pmatrix} &= \operatorname{rg} \begin{pmatrix} 1 & 2 & 0 \\ 0 & -2x+y+1 & z-1 \end{pmatrix} = 1 \Leftrightarrow \\ &\begin{cases} -2x+y+1=0 \\ z-1=0 \end{cases} \end{aligned}$$

La retta r ha dunque equazioni cartesiane

$$\begin{cases} -2x+y+1=0 \\ z-1=0 \end{cases}$$

Naturalmente ogni sistema lineare equivalente a quello trovato descrive la medesima retta r , interpretandola come intersezione di due piani diversi.

Notiamo che il sistema lineare omogeneo

$$\begin{cases} -2x+y=0 \\ z=0 \end{cases}$$

descrive lo spazio direttore $\langle(1, 2, 0)\rangle$ della retta r .

4.2.6 Naturalmente, così come accade in $\mathbb{A}^2(\mathbb{R})$, anche in $\mathbb{A}^3(\mathbb{R})$ possiamo descrivere una retta attraverso delle equazioni parametriche. Consideriamo la retta r passante per il punto $P = (x_0, y_0, z_0)$ e avente spazio direttore $W = \langle(a, b, c)\rangle$ con $(a, b, c) \neq (0, 0, 0)$. Allora le equazioni parametriche di r sono:

$$\begin{cases} x = x_0 + \lambda a \\ y = y_0 + \lambda b \\ z = z_0 + \lambda c \end{cases} \quad (\lambda \in \mathbb{R}). \quad (4.4)$$

Certamente le equazioni parametriche di r non sono uniche nel senso che al posto del punto P potremmo scegliere un qualsiasi altro punto di r e al posto del vettore (a, b, c) un suo qualsiasi multiplo non nullo. Il vettore (a, b, c) si chiama *vettore direttore* della retta r .

Il metodo più comodo per passare dalle equazioni parametriche alle equazioni cartesiane di r consiste nella eliminazione del parametro dalle equazioni (4.4): si utilizza una equazione per esprimere il parametro in funzione di una variabile e si sostituisce l'espressione trovata nelle restanti due equazioni.

Esempio 4.2.7 Descrivere mediante equazioni cartesiane la retta di equazioni parametriche:

$$\begin{cases} x = -1 + \lambda \\ y = -2 + \lambda \\ z = 2\lambda \end{cases} \quad (\lambda \in \mathbb{R}).$$

Usiamo la prima equazione per esprimere il parametro λ in funzione di x : $\lambda = x + 1$ e sostituiamo l'espressione di λ così ottenuta nelle altre due equazioni. Otteniamo il sistema:

$$\begin{cases} y = -1 + x \\ z = 2x + 2 \end{cases}$$

che è il sistema delle equazioni cartesiane di r . Naturalmente ogni sistema lineare equivalente a quello trovato (cioè avente lo stesso insieme di soluzioni) descrive la retta r .

Esercizio 4.2.8 Intersezione tra una retta e un piano nello spazio. Consideriamo la retta

$$r : \begin{cases} ax + by + cz = d \\ a'x + b'y + c'z = d' \end{cases}$$

dove $\text{rg} \begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix} = 2$, e il piano $\pi : a''x + b''y + c''z = d''$ con $(a'', b'', c'') \neq (0, 0, 0)$. Il sistema che descrive la loro intersezione è :

$$\begin{cases} ax + by + cz = d \\ a'x + b'y + c'z = d' \\ a''x + b''y + c''z = d'' \end{cases}$$

ovvero

$$\begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} d \\ d' \\ d'' \end{pmatrix}$$

con $\text{rg} \begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix} \geq 2$ (le prime due righe sono necessariamente linearmente indipendenti). Indichiamo il sistema con $A\underline{x} = \underline{d}$. Abbiamo le seguenti possibilità:

- $\text{rg} A = 2$. In questo caso l'equazione (omogenea) della giacitura del piano è combinazione lineare delle equazioni che individuano la giacitura della retta. Questo significa che la retta è parallela al piano. Per la matrice completa vi

sono due possibilità: se essa ha rango 3 il sistema non ha soluzioni, quindi la retta è parallela al piano e non ha punti in comune con esso. Se il rango della matrice completa è 2, il sistema ha infinite soluzioni che individuano una varietà lineare di dimensione uno, cioè una retta. La retta di partenza è pertanto contenuta nel piano.

- $\text{rg}A = 3$. In questo caso l'equazione del piano è indipendente dall'equazione della retta. Per la matrice completa abbiamo una sola possibilità: che abbia anch'essa rango 3. Il sistema ha quindi una ed una sola soluzione: il punto di intersezione tra la retta e il piano. Diremo in questo caso che la retta e il piano sono *incidenti*.

Esercizio 4.2.9 *Posizione reciproca di due rette in $\mathbb{A}^3(\mathbb{R})$.* Sempre più arditamente, studiamo l'intersezione tra due rette nello spazio affine tridimensionale. Consideriamo le rette r e s di equazioni:

$$r : \begin{cases} ax + by + cz = d \\ a'x + b'y + c'z = d' \end{cases} \quad s : \begin{cases} a''x + b''y + c''z = d'' \\ a'''x + b'''y + c'''z = d''' \end{cases}.$$

Ambedue i sistemi, descrivendo delle rette, hanno rango 2. Il sistema delle loro equazioni è allora dato, in termini matriciali, da

$$\begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \\ a''' & b''' & c''' \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} d \\ d' \\ d'' \\ d''' \end{pmatrix}$$

i.e., $A\underline{x} = \underline{d}$. Si ha subito $\text{rg}A \geq 2$, dal momento che le prime due righe (e le ultime due) sono linearmente indipendenti. Distinguiamo i seguenti casi:

- $\text{rg}A = 2$. Significa che le ultime due righe di A sono linearmente dipendenti dalle prime due. Quindi gli spazi direttori delle due rette sono uguali: le due rette sono parallele. Se la matrice completa ha anch'essa rango due il sistema ammette soluzioni e queste formano una varietà lineare di dimensione uno, pertanto le due rette coincidono. Se il rango della matrice completa è tre (il rango della matrice incompleta è diverso dal rango della matrice completa) allora il sistema non ha soluzioni. Le due rette sono parallele e non si incontrano. Si noti che la matrice completa è di ordine 4: può avere rango 4? Certamente no: la matrice completa si ottiene infatti aggiungendo alla matrice completa una sola colonna. Dunque il rango di A ed il rango di $(A|\underline{d})$ possono differire al più di uno.

- $\text{rg}A = 3$. In questo caso la matrice incompleta ha rango massimo. Se il rango della matrice completa è 3 allora il sistema ha una ed una sola soluzione: le due rette si intersecano in un punto. Se la matrice completa ha rango 4 allora le due rette non hanno punti in comune ma non sono parallele. In questo caso le due rette si dicono *sghembe*.

Esercizio 4.2.10 *Caratterizzazione geometrica delle rette sghembe.*

Abbiamo definito le rette sghembe utilizzando la struttura del sistema delle loro equazioni. Vogliamo ora dare di esse una caratterizzazione più geometrica. In effetti quello che vogliamo vedere è che due rette sono sghembe se e solo se non esiste alcun piano contenente entrambe. Nell'Esercizio 4.2.9 abbiamo esaminato tutte le possibili mutue posizioni di due rette in $\mathbb{A}^3(\mathbb{R})$. Date due rette, supponiamo che esista un piano che le contenga: allora le due rette o sono parallele senza alcun punto in comune, oppure hanno un punto in comune, oppure coincidono. In ogni caso non possono essere sghembe. Viceversa, se non esiste alcun piano che le contiene non possono essere parallele e non possono intersecarsi (perché sarebbero in entrambi i casi complanari - si vedano, a questo proposito, i quesiti 2 e 3 dell'Esercizio 4.2.12). L'unica possibilità è quindi che le rette siano sghembe.

Osservazione 4.2.11 Ci poniamo ora il seguente quesito: date due sotto-varietà lineari $L_1 = P + V_1$ e $L_2 = Q + V_2$ di $\mathbb{A}^n(\mathbb{R})$, come facciamo a stabilire se esse coincidono? Mostriamo che L_1 e L_2 coincidono se e solo se $P \in L_2$ e $V_1 = V_2$.

Se L_1 e L_2 coincidono allora $L_1 = P + V_1 = Q + V_2 = L_2$ e dunque $P = P + \mathbf{0}_{V_1} \in L_2$. D'altro canto se $L_1 = L_2$ allora $P + v_1 \in L_2$, per ogni $v_1 \in V_1$, cioè si può scrivere $P + v_1 = Q + v_2$ con $v_2 \in V_2$. Ma $P \in L_2$ e dunque $P = Q + s_2$ con $s_2 \in V_2$, ne deriva che $P + v_1 = (Q + s_2) + v_1 = Q + v_2$ e per le proprietà degli spazi affini si ha $Q + (s_2 + v_1) = Q + v_2$, dunque $s_2 + v_1 = v_2$, i.e., $v_1 = v_2 - s_2 \in V_2$ dunque ogni vettore di V_1 è un vettore di V_2 : $V_1 \leq V_2$. Notiamo che potevamo fare l'analogo ragionamento con $Q \in L_2$ e ottenere che $V_2 \leq V_1$, dunque $V_1 = V_2$.

Viceversa, prendiamo due varietà lineari $L_1 = P + V_1$ e $L_2 = Q + V_2$ tali che: $P \in L_2$ e $V_1 = V_2$; prendiamo un punto $P + v_1$ di L_1 , con $v_1 \in V_1 = V_2$. Dal momento che $P \in L_2$ possiamo scrivere $P = Q + v_2$ con $v_2 \in V_2$ e dunque: $P + v_1 = (Q + v_2) + v_1$, ma $v_1 + v_2 \in V_2$ e quindi $P + v_1 \in L_2$. Abbiamo ottenuto da $P \in L_2$ e $V_1 = V_2$ che $L_1 \subseteq L_2$. Ma se $P \in L_2$ e $V_1 = V_2$ allora vale anche $Q \in L_1$. Infatti se si ha $P = Q + v_2$ con $v_2 \in V_2 = V_1$ allora sommando a

P il vettore $-v_2 \in V_1 = V_2$ si ha che $P + (-v_2) = (Q + v_2) + (-v_2) = Q$ ottenendo che $Q = P + (-v_2) \in L_1$. Possiamo a questo punto ripetere per Q il ragionamento fatto sopra ottenendo $L_2 \subseteq L_1$. Così abbiamo dimostrato che le varietà $L_1 = P + V_1$ e $L_2 = Q + V_2$ coincidono se e solo se $P \in L_2$ e $V_1 = V_2$. Equivalentemente possiamo affermare che L_1 e L_2 coincidono se e solo se esiste un punto di L_1 che appartiene ad L_2 e le giaciture di L_1 e L_2 sono uguali.

Esercizio 4.2.12 Vogliamo ora risolvere alcuni quesiti utilizzando le costruzioni realizzate finora.

Quesito 1: determinare un piano che contenga la retta $r : (1, 1, 1) + \langle (2, 1, 0) \rangle$ e che sia parallelo alla direzione data dal vettore $v = (1, 0, -1)$.

Risoluzione: quello che cerchiamo di trovare è una scrittura del tipo: $P + \langle v_1, v_2 \rangle$. Ora, il piano cercato passa per $(1, 1, 1)$ dovendo contenere la retta $(1, 1, 1) + \langle (2, 1, 0) \rangle$, e la sua giacitura deve contenere la giacitura della retta r (cioè $(2, 1, 0)$) e la direzione $(1, 0, -1)$. Abbiamo fatto: il piano è dato da: $(1, 1, 1) + \langle (2, 1, 0), (1, 0, -1) \rangle$. Per determinare una sua equazione cartesiana cerchiamo gli elementi (x, y, z) tali che la matrice

$$\begin{pmatrix} 1 & 0 & -1 \\ 2 & 1 & 0 \\ x-1 & y-1 & z-1 \end{pmatrix}$$

abbia rango 2. Riducendo la matrice in forma a scala otteniamo la matrice:

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & x-2y+z \end{pmatrix}$$

che ha rango 2 se e solo se $x - 2y + z = 0$. Questa è quindi l'equazione cartesiana del piano cercato. Il piano passa quindi per l'origine $(0, 0, 0)$ e può essere descritto semplicemente mediante la scrittura $\langle (2, 1, 0), (1, 0, -1) \rangle$, intendendo con essa la varietà $(0, 0, 0) + \langle (2, 1, 0), (1, 0, -1) \rangle$.

Quesito 2: in $\mathbb{A}^3(\mathbb{R})$ siano date le rette $r_1: (1, 1, 0) + \langle (1, 0, -1) \rangle$ e $r_2: (3, 2, 1) + \langle (2, 1, 1) \rangle$; determinare, se possibile, un piano che contenga r_1 e r_2 .

Risoluzione: in questo caso abbiamo una descrizione esplicita dei punti della retta r_1 : essi sono della forma $(1 + t, 1, -t)$ al variare di t nell'insieme dei numeri reali, mentre i punti di r_2 sono della forma $(3 + 2q, 2 + q, 1 + q)$ con $q \in \mathbb{R}$. Studiamo la posizione reciproca delle due rette: r_1 e r_2 non sono certo

parallele dal momento che hanno giaciture diverse. I punti in comune alle due rette sono i punti tali che si possa scrivere $(1+t, 1, -t) = (3+2q, 2+q, 1+q)$, per qualche t e q in $\mathbb{R}$. Con qualche calcolo si ottiene $q = -1$ e $t = 0$. Le due rette hanno pertanto in comune il punto $(1, 1, 0)$. La retta r_2 si può allora scrivere come $(1, 1, 0) + \langle (2, 1, 1) \rangle$. Il piano che contiene r_1 e r_2 dovrà essere parallelo ad entrambe le rette e quindi la sua giacitura deve contenere la giacitura di entrambe: $(1, 0, -1)$, $(2, 1, 1)$, e passare per il loro punto di intersezione. Il piano è dunque: $(1, 1, 0) + \langle (1, 0, -1), (2, 1, 1) \rangle$.

Quesito 3: Date due rette parallele r_1 e r_2 , determinare il piano che le contenga. Ad esempio siano r_1 e r_2 le rette parallele $(1, 1, 0) + \langle (1, 1, -1) \rangle$ e $(3, 1, -1) + \langle (-2, -2, +2) \rangle$ rispettivamente.

Risoluzione: ovviamente il piano che contiene r_1 e r_2 dovrà contenere i punti $(1, 1, 0)$ e $(3, 1, -1)$ e la sua giacitura dovrà contenere lo spazio direttore delle due rette: $\langle (1, 1, -1) \rangle$. In altre parole il piano è del tipo $(1, 1, 0) + \langle (1, 1, -1), v \rangle$ (con v vettore per ora sconosciuto ma linearmente indipendente da $(1, 1, -1)$) e tale che il punto $(3, 1, -1)$ appartenga al piano. Ma ecco che allora il vettore differenza tra i due punti $(1, 1, 0) - (3, 1, -1)$ deve appartenere alla giacitura del piano, cioè $(-2, 0, 1) \in \langle (1, 1, -1), v \rangle$. Essendo tale vettore linearmente indipendente da $(1, 1, -1)$ ne segue che il piano cercato è: $(1, 1, 0) + \langle (1, 1, -1), (-2, 0, 1) \rangle$.

Quesito 4 Siano dati la retta r : $(1, 2, 0) + \langle (1, 1, 1) \rangle$ e il punto $M = (0, -1, -1)$; determinare il piano che li contenga.

Risoluzione: se il punto appartenesse alla retta allora non vi sarebbe un solo piano contenente la retta e il punto. Ma affinché il punto M appartenga alla retta deve esistere $t \in \mathbb{R}$ tale che sia $(1+t, 2+t, t) = (0, -1, -1)$, e si vede subito che questo non può avvenire per nessun t reale. Allora il piano cercato deve contenere i punti $(1, 2, 0)$, $(0, -1, -1)$ e la sua giacitura deve contenere $(1, 1, 1)$. Ora la giacitura dovrà anche contenere un vettore che sommato ad $(1, 2, 0)$ dia $(0, -1, -1)$ cioè dovrà contenere il vettore $(1, 2, 0) - (0, -1, -1) = (1, 3, 1)$. Tale vettore è linearmente indipendente da $(1, 1, 1)$ e deve stare nella giacitura del piano. Il piano è dato allora da $(1, 2, 0) + \langle (1, 3, 1), (1, 1, 1) \rangle$.

Esercizio 4.2.13 Qual è il piano in $\mathbb{A}^3(\mathbb{R})$ individuato da tre punti distinti e non allineati (significato: che non appartengono ad una stessa retta)? Consideriamo tre punti distinti, non allineati $P_1 = (x_1, y_1, z_1)$, $P_2 = (x_2, y_2, z_2)$ e $P_3 = (x_3, y_3, z_3)$. Il piano contenente P_1 , P_2 e P_3 potrà essere descritto come $P_1 + \langle v_1, v_2 \rangle$ (al posto di P_1 avremmo anche potuto usare P_2 o P_3). Come deve essere fatto lo spazio vettoriale giacitura? Sicuramente dovrà contenere i

vettori che sommati a P_1 danno P_2 e P_3 , cioè $P_3 - P_1$ e $P_2 - P_1$. Quindi il piano sarà $P_1 + \langle P_3 - P_1, P_2 - P_1 \rangle$. Siamo sicuri di aver determinato un piano perché i punti P_1, P_2, P_3 non sono allineati, quindi $\dim(\langle P_3 - P_1, P_2 - P_1 \rangle) = 2$. Se i punti fossero allineati $P_3 - P_1$ sarebbe un multiplo di $P_2 - P_1$ e quindi avremmo $\dim(\langle P_3 - P_1, P_2 - P_1 \rangle) = 1$. Questa non è che la traduzione algebrica del fatto geometrico che per tre punti allineati non passa un solo piano ma molti. Quanti?

Esercizio 4.2.14 Quando quattro punti di $\mathbb{A}^3(\mathbb{R})$ sono complanari? Prendiamo quattro punti distinti P_1, P_2, P_3 e P_4 di $\mathbb{A}^3(\mathbb{R})$. Se i quattro punti sono allineati, certo sono complanari perché esistono infiniti piani contenenti una retta. Se P_1, P_2 e P_3 sono allineati e P_4 non appartiene alla retta r da essi individuata, la retta r e il punto P_4 individuano un unico piano che contiene i 4 punti. Supponiamo ora che P_1, P_2, P_3 e P_4 siano a 3 a 3 non allineati. Consideriamo i punti P_1, P_2 e P_3 . Abbiamo visto prima che un piano che li contiene è dato da $P_1 + \langle P_3 - P_1, P_2 - P_1 \rangle$. Ora, affinché P_4 appartenga a questo piano, deve succedere che il vettore $P_4 - P_1$ stia nella giacitura del piano: $P_4 - P_1 \in \langle P_3 - P_1, P_2 - P_1 \rangle$, cioè che $P_4 - P_1$ sia combinazione lineare dei vettori $P_3 - P_1$ e $P_2 - P_1$. Questo significa che il rango della matrice

$$\begin{pmatrix} P_4 - P_1 \\ P_3 - P_1 \\ P_2 - P_1 \end{pmatrix}$$

è uguale a due (con questa scrittura intendiamo la matrice che ha come righe i vettori coordinate dei vettori $P_i - P_1$).

Osservazione 4.2.15 Vogliamo porci la seguente domanda: una varietà lineare può essere un sottospazio vettoriale? La domanda non è veramente ben posta. Da una parte uno spazio vettoriale è costituito da vettori e dall'altra una sottovarietà lineare è fatta semplicemente da punti dello spazio affine. È un po' come chiedersi se una cassa di aringhe sia uguale o meno ad una cassa di pesci San Pietro...la domanda non ha senso. Le nostre definizioni su spazi vettoriali e spazi affini non lo permettono. Ma potremmo in ogni caso osservare che la sottovarietà più vicina ad uno spazio vettoriale è quella data da $P + V$ ove $V \leq \mathbb{R}^3$ e P , pensato come un vettore di $\mathbb{R}^3$, appartiene a V . In tal caso, pensando ogni punto di $P + V$ come un vettore, si ha l'identificazione $P + V = V$. Quindi in modo poco ortodosso possiamo dire che una sottovarietà lineare $P + V$ è un sottospazio vettoriale se e solo se

$P \in V$. In tal caso $P + V$ coincide con V , nel senso che è una sottovarietà contenente il punto di coordinate $(0, 0, 0)$, a rigore: $V = (0, 0, 0) + V = P + V$.

4.3 Esercizi proposti

Esercizio 4.3.1 In $\mathbb{A}^3(\mathbb{R})$ si considerino i punti $P = (1, 2, 0)$ e $Q = (0, 0, -1)$.

1. Determinare la retta r passante per P e Q ;
2. determinare equazioni cartesiane della retta s parallela a r , passante per il punto $T = (1, 1, 1)$;
3. determinare un'equazione cartesiana del piano π contenente le rette r ed s ;
4. determinare il piano σ parallelo a π passante per il punto $R = (1, 1, 2)$.

Esercizio 4.3.2 Date le rette $r : \begin{cases} x - y + 2z = 0 \\ y + z = 0 \end{cases}$ ed $s : \begin{cases} x + y - 4z = 1 \\ x - y - 2z = 2, \end{cases}$ determinare:

1. la posizione reciproca delle rette r ed s ;
2. se possibile, una retta incidente la retta r in $(0, 0, 0)$ e non incidente s .

Esercizio 4.3.3 In $\mathbb{A}^3(\mathbb{R})$ si considerino i punti $P = (1, 1, 1)$ e $Q = (1, 0, 2)$.

1. Determinare, se possibile, un piano passante per i punti P e Q , parallelo al vettore $v = (1, 2, 1)$. Quanti piani esistono come quello richiesto?
2. Determinare, se possibile, un piano passante per i punti P e Q , parallelo al vettore $v = (0, 2, -2)$. Quanti piani esistono come quello richiesto?
3. Determinare, se possibile, un piano passante per i punti P e Q , contenente il punto $R = (1, 1, -1)$. Quanti piani esistono come quello richiesto?
4. Determinare, se possibile, un piano passante per i punti P , Q , R e $S = (2, -1, 0)$. Quanti piani esistono come quello richiesto?

Esercizio 4.3.4 Stabilire se esistono valori del parametro reale k tali che i punti $P = (1, 0, 0)$, $Q = (k, k, 1)$ ed $R = (2, 1, 2)$ siano allineati. Determinare, al variare di k , il piano π_k contenente i punti P , Q ed R . Determinare, se possibile, il valore di k tale che π_k sia parallelo alla retta r di equazioni $r : \begin{cases} 2x + 2y - z = 0 \\ x + y + 3z = 1. \end{cases}$

Esercizio 4.3.5 Date le rette $r : \begin{cases} 2x - y = 1 \\ y + z = 6 \end{cases}$ ed $s : \begin{cases} x = 2 \\ x + y - 2z = -1, \end{cases}$ determinare:

1. la posizione reciproca delle rette r ed s ;
2. un piano contenente r ed s ;
3. se possibile, una retta passante per $T = (1, 1, 1)$ incidente le rette r ed s .

Lezione 5

Concetti metrici

In questa lezione vogliamo introdurre la definizione di “prodotto scalare” tra vettori (il risultato di questo prodotto è un numero reale). Tale nozione è alla base sia della nozione di norma (o lunghezza) di vettori che di quella di (misura del coseno dell’) angolo tra due vettori. In particolare il prodotto scalare permette di definire l’ortogonalità tra vettori.

5.1 Prodotto scalare

Consideriamo dapprima la nozione di lunghezza di un vettore, nozione che ci è più familiare. Come definire la lunghezza di un vettore?

Incominciamo col definire la lunghezza di un vettore di $\mathbb{R}^2$ o di $\mathbb{R}^3$. Esprimeremo tale lunghezza attraverso le coordinate del vettore in una base fissata e assumeremo che i vettori $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$, e $(1, 0)$, $(0, 1)$ abbiano lunghezza 1. Sottolineiamo che questo è completamente arbitrario, nel senso che dipende dal modello che decidiamo di usare. Iniziamo con la definizione di un prodotto scalare su $\mathbb{R}^2$ e $\mathbb{R}^3$ che poi generalizzeremo al caso di dimensione superiore.

Definizione 5.1.1 *Il prodotto scalare usuale o euclideo in $\mathbb{R}^2$ (resp. $\mathbb{R}^3$) è la seguente applicazione*

$$\mathbb{R}^2 \times \mathbb{R}^2 \longrightarrow \mathbb{R}$$

$$(resp. \mathbb{R}^3 \times \mathbb{R}^3 \longrightarrow \mathbb{R})$$

$$((x_1, x_2), (y_1, y_2)) \longmapsto x_1y_1 + x_2y_2$$

$$(resp. ((x_1, x_2, x_3), (y_1, y_2, y_3)) \longmapsto x_1y_1 + x_2y_2 + x_3y_3.)$$

Definizione 5.1.2 Definiamo il prodotto scalare usuale su $\mathbb{R}^n$ come l'applicazione che associa ad ogni coppia di vettori $v = (x_1, \dots, x_n)$ e $w = (y_1, \dots, y_n)$ di $\mathbb{R}^n$ il numero reale

$$v \bullet w = x_1 y_1 + x_2 y_2 + \dots + x_n y_n.$$

Si vede subito che $v \bullet w = w \bullet v$. Usando la definizione si verificano facilmente le seguenti proprietà:

- 1) $\mathbf{x} \bullet \mathbf{y} = \mathbf{y} \bullet \mathbf{x}$ (simmetria);
- 2) $(\mathbf{x} + \mathbf{y}) \bullet \mathbf{z} = \mathbf{x} \bullet \mathbf{z} + \mathbf{y} \bullet \mathbf{z} = \mathbf{z} \bullet \mathbf{x} + \mathbf{z} \bullet \mathbf{y} = \mathbf{z} \bullet (\mathbf{x} + \mathbf{y})$ per ogni $\mathbf{x}, \mathbf{y}, \mathbf{z}$ in $\mathbb{R}^n$;
- 3) $(\lambda \mathbf{x}) \bullet \mathbf{y} = \lambda(\mathbf{x} \bullet \mathbf{y}) = \mathbf{x} \bullet (\lambda \mathbf{y})$ per ogni $\mathbf{x}, \mathbf{y}$ in $\mathbb{R}^n$, $\lambda \in \mathbb{R}$.

Dal momento che il prodotto scalare è simmetrico, le proprietà che valgono per un argomento valgono anche per l'altro.

Si noti che:

- 4) per ogni $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ si ha $\mathbf{x} \bullet \mathbf{x} = x_1^2 + x_2^2 + \dots + x_n^2 \geq 0$ e vale l'uguaglianza se e solo se $\mathbf{x} = \mathbf{0}$.

Le quattro proprietà appena elencate sono molto importanti in quanto sono in qualche modo la vera "essenza" di un prodotto scalare.

Definizione 5.1.3 Sia $\mathbb{R}^n$ dotato del prodotto scalare usuale. Dato $\mathbf{x} \in \mathbb{R}^n$ definiamo $\|\mathbf{x}\| = \sqrt{\mathbf{x} \bullet \mathbf{x}}$. Il numero reale positivo $\|\mathbf{x}\|$ si dice norma del vettore $\mathbf{x}$. I vettori di norma 1 si dicono versori.

La norma del vettore è la lunghezza che stavamo cercando.

Consideriamo l'applicazione

$$\|\cdot\| : \mathbb{R}^n \longrightarrow \mathbb{R}^{\geq 0}$$

che associa ad ogni vettore la sua norma (con $\mathbb{R}^{\geq 0}$ abbiamo indicato l'insieme dei numeri reali positivi o nulli). Essa gode di alcune interessanti proprietà:

- $\|\mathbf{x}\| = 0$ se e solo se $\mathbf{x} = \mathbf{0}_{\mathbb{R}^n}$
- se $\alpha \in \mathbb{R}$ allora $\|\alpha \mathbf{x}\| = |\alpha| \|\mathbf{x}\|$.

Infatti

$$\begin{aligned} \|\alpha \mathbf{x}\|^2 &= (\alpha \mathbf{x}) \bullet (\alpha \mathbf{x}) = (\alpha x_1, \alpha x_2, \dots, \alpha x_n) \bullet (\alpha x_1, \alpha x_2, \dots, \alpha x_n) = \alpha^2 x_1^2 + \\ &\alpha^2 x_2^2 + \dots + \alpha^2 x_n^2 = \alpha^2 (x_1^2 + x_2^2 + \dots + x_n^2) = \alpha^2 \|\mathbf{x}\|^2. \end{aligned}$$

Osservazione 5.1.4 Se α è un numero reale conosciamo il valore assoluto $|\alpha|$ fin dalle scuole medie. Ricordiamo che $|\alpha| = \alpha$ se $\alpha \geq 0$ e $|\alpha| = -\alpha$ se $\alpha < 0$. Del resto possiamo considerare $\mathbb{R}$ come uno spazio vettoriale su se stesso e scegliere 1 come sua base. Allora possiamo definire il prodotto scalare usuale su $\mathbb{R}$ ponendo $(\alpha) \bullet (\beta) = \alpha\beta$. Secondo questa definizione la norma di un vettore α di $\mathbb{R}$ (cioè di un numero reale α) è proprio il valore assoluto di α : $\sqrt{\alpha^2} = |\alpha|$.

Definizione 5.1.5 Si definisce misura del coseno dell'angolo $\hat{\mathbf{v}\mathbf{w}}$ tra due vettori non nulli $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$:

$$\cos \hat{\mathbf{v}\mathbf{w}} := \frac{\mathbf{v} \bullet \mathbf{w}}{\|\mathbf{v}\| \|\mathbf{w}\|}. \quad (5.1)$$

In base alla Definizione 5.1.5 definiremo vettori ortogonali quei vettori per cui $\cos \hat{\mathbf{v}\mathbf{w}}$ vale zero, quindi

Definizione 5.1.6 Dato $\mathbb{R}^n$ con il prodotto scalare usuale “ $\bullet$ ”, un vettore $\mathbf{v}$ si dice **ortogonale** ad un vettore $\mathbf{w}$ se vale $\mathbf{v} \bullet \mathbf{w} = 0$.

Definizione 5.1.7 Una base $\{v_1, v_2, \dots, v_n\}$ di $\mathbb{R}^n$ si dice **ortonormale** se $\|v_i\| = 1$ per ogni $i = 1, 2, \dots, n$ e $v_i \bullet v_j = 0$ per ogni $i \neq j$. Si dice **ortogonale** se $v_i \bullet v_j = 0$ per ogni $i \neq j$.

Esempio 5.1.8 La base canonica di $\mathbb{R}^n$ è una base ortonormale rispetto al prodotto scalare usuale.

5.2 Il procedimento di Gram-Schmidt

Il procedimento di Gram-Schmidt serve a costruire una base ortonormale di ogni sottospazio di $\mathbb{R}^n$ a partire da una sua base qualunque. Da questo dedurremo i seguenti risultati:

- i) ogni sottospazio di $\mathbb{R}^n$ ammette una base ortonormale;
- ii) definiremo la proiezione ortogonale di un vettore su un sottospazio e la calcoleremo esplicitamente.

Sia T un sottospazio vettoriale di $\mathbb{R}^n$. Consideriamo il vettore non nullo $\mathbf{t} \in T$ e prendiamo la sua norma: $\|\mathbf{t}\| = \sqrt{\mathbf{t} \bullet \mathbf{t}}$. Questo è un numero positivo quindi possiamo prendere $\mathbf{t}' = \frac{1}{\|\mathbf{t}\|} \mathbf{t}$ che è ancora un vettore di T per il quale si ha: $\mathbf{t}' \bullet \mathbf{t}' = \frac{1}{\|\mathbf{t}\|} \mathbf{t} \bullet \frac{1}{\|\mathbf{t}\|} \mathbf{t} = \frac{1}{\|\mathbf{t}\|^2} \mathbf{t} \bullet \mathbf{t} = 1$. Tale procedimento si chiama

normalizzazione del vettore $\mathbf{t}$: si parte da un qualunque vettore $\mathbf{t}$ non nullo e si ottiene un vettore ad esso proporzionale di norma unitaria.

In generale, dati un sottospazio vettoriale T di $\mathbb{R}^n$ ed una sua base $\{v_1, \dots, v_k\}$, è sempre possibile costruire una base ortonormale $\{u_1, \dots, u_k\}$ di T tale che $\langle v_1, \dots, v_i \rangle = \langle u_1, \dots, u_i \rangle$ (per ogni $i = 1, \dots, k$). Il procedimento è il seguente:

$$\begin{aligned} u_1 &= \frac{v_1}{\|v_1\|} \\ u'_2 &= v_2 - (v_2 \bullet u_1)u_1 \\ u_2 &= \frac{u'_2}{\|u'_2\|} \\ u'_3 &= v_3 - (v_3 \bullet u_1)u_1 - (v_3 \bullet u_2)u_2 \\ &\vdots \\ &\vdots \\ u'_k &= v_k - \sum_{i=1}^{k-1} (v_k \bullet u_i)u_i \\ u_k &= \frac{u'_k}{\|u'_k\|}. \end{aligned}$$

Esempio 5.2.1 Sia $T = \langle (1, 1, 1), (2, 1, 0) \rangle$. Vogliamo costruire una base ortonormale di T utilizzando il procedimento di ortonormalizzazione di Gram-Schmidt appena illustrato. Poniamo $v_1 = (1, 1, 1)$ e $v_2 = (2, 1, 0)$ e osserviamo che tali vettori sono linearmente indipendenti poiché uno non è multiplo dell'altro. Dunque $\dim(T) = 2$. Abbiamo:

$$\begin{aligned} u_1 &= \frac{v_1}{\|v_1\|} = \frac{1}{\sqrt{3}}(1, 1, 1); \\ u'_2 &= v_2 - (v_2 \bullet u_1)u_1 = (2, 1, 0) - ((2, 1, 0) \bullet \frac{1}{\sqrt{3}}(1, 1, 1))\frac{1}{\sqrt{3}}(1, 1, 1) = \\ &= (2, 1, 0) - (1, 1, 1) = (1, 0, -1); \\ u_2 &= \frac{u'_2}{\|u'_2\|} = \frac{1}{\sqrt{2}}(1, 0, -1). \end{aligned}$$

5.3 Proiezioni ortogonali

Definizione 5.3.1 Sia T un sottoinsieme di $\mathbb{R}^n$. Si dice **ortogonale** di T e si indica con $T^\perp$ l'insieme dei vettori di $\mathbb{R}^n$ ortogonali ad ogni vettore in T :

$$T^\perp = \{v \in \mathbb{R}^n \mid v \bullet t = 0 \ \forall t \in T\}.$$

Esempio 5.3.2 Sia $T = \langle \mathbf{0}_{\mathbb{R}^n} \rangle$. Allora $T^\perp = \mathbb{R}^n$, infatti $\langle \mathbf{0}_{\mathbb{R}^n} \rangle^\perp = \{v \in \mathbb{R}^n \mid v \bullet \mathbf{0}_{\mathbb{R}^n} = 0\} = \mathbb{R}^n$. Si ha, equivalentemente, $\mathbb{R}^{n\perp} = \mathbf{0}_{\mathbb{R}^n}$.

Osservazione 5.3.3 Studiamo alcune proprietà dell'ortogonale.

- 1) L'insieme $T^\perp$ è un sottospazio vettoriale di V . Infatti siano v_1 e v_2 elementi di $T^\perp$, allora la loro somma $v_1 + v_2$ è ancora ortogonale a ogni vettore di T : per ogni $t \in T$ si ha $(v_1 + v_2) \bullet t = v_1 \bullet t + v_2 \bullet t = 0$ (per la proprietà distributiva del prodotto scalare rispetto alla somma). Lo stesso succede se si moltiplica per uno scalare $\alpha \in \mathbb{R}$ un elemento $v \in T^\perp$: $(\alpha v) \bullet t = \alpha(v \bullet t) = 0$ per ogni $t \in T$. Quindi $T^\perp$ è un sottospazio di $\mathbb{R}^3$.

2)

$$T \cap T^\perp = \{\mathbf{0}_V\}.$$

Infatti se v è un vettore nell'intersezione di T e $T^\perp$ esso appartiene sia a T che a $T^\perp$, cioè v è ortogonale a tutti gli elementi di T e quindi, in particolare, a se stesso: $v \bullet v = 0$, ma questo avviene se e solo se $v = \mathbf{0}_V$.

- 3) Se $S \subset T$ allora $T^\perp \subset S^\perp$.
- 4) $T^\perp = \langle T \rangle^\perp$. Infatti essendo $T \subset \langle T \rangle$ è chiaro che $\langle T \rangle^\perp \subset T^\perp$. Viceversa se $v \in T^\perp$ allora v è ortogonale anche ad ogni vettore di $\langle T \rangle$ (verificarlo per esercizio).
- 5) Sia T un sottospazio vettoriale di $\mathbb{R}^n$. Notiamo che, affinché un vettore v appartenga a $T^\perp$, basterà che sia ortogonale ai vettori di una base di T . Supponiamo infatti che T abbia dimensione uno e sia generato da un vettore (non nullo) w : $T = \langle w \rangle = \{\lambda w \mid \lambda \in \mathbb{R}\}$. Allora se $v \bullet w = 0$ si ha $v \bullet (\lambda w) = \lambda(v \bullet w) = 0$ cioè v è ortogonale a T . Analogamente, se $\{w_1, w_2, \dots, w_m\}$ è una base di T allora ogni vettore di T è della forma: $\beta_1 w_1 + \beta_2 w_2 + \dots + \beta_m w_m$ e dunque: $v \bullet (\beta_1 w_1 + \beta_2 w_2 + \dots + \beta_m w_m) = \beta_1(v \bullet w_1) + \beta_2(v \bullet w_2) + \dots + \beta_m(v \bullet w_m) = 0$. Pertanto se v è ortogonale

a $w_1, w_2, \dots, w_m$ allora è ortogonale ad ogni altro vettore di T quindi appartiene a $T^\perp$.

Proposizione 5.3.4 *Sia T un sottospazio vettoriale di $\mathbb{R}^n$ con $\dim T = m$. Allora $\dim T^\perp = n - m$ ed ogni vettore di $\mathbb{R}^n$ si scrive come somma di un vettore di T e di un vettore di $T^\perp$.*

Dimostrazione. Fissiamo una base ortonormale $\{w_1, \dots, w_m\}$ di T e completiamola in una base di $\mathbb{R}^n$: $\mathcal{B} = \{w_1, \dots, w_m, v_{m+1}, \dots, v_n\}$. Utilizzando il procedimento di ortonormalizzazione di Gram-Schmidt possiamo costruire, a partire da $\mathcal{B}$, una base di $\mathbb{R}^n$ ortonormale: $\{w_1, \dots, w_m, w_{m+1}, \dots, w_n\}$. Notiamo che i primi m vettori non sono cambiati perché erano già ortonormali. Ora, per l'Osservazione 5.3.3(5), abbiamo: $T^\perp = \{v \in \mathbb{R}^n \mid v \bullet w_i = 0 \ \forall i = 1, \dots, m\}$. Ma ogni vettore di $\mathbb{R}^n$ si scrive in modo unico come $v = \alpha_1 w_1 + \dots + \alpha_n w_n$, quindi

$$\begin{aligned} T^\perp &= \left\{ \sum_{i=1}^n \alpha_i w_i \in \mathbb{R}^n \mid \sum_{i=1}^n \alpha_i w_i \bullet w_i = 0 \ \forall i = 1, \dots, m \right\} \\ &= \left\{ \sum_{i=1}^n \alpha_i w_i \in \mathbb{R}^n \mid \alpha_i w_i \bullet w_i = \alpha_i = 0 \ \forall i = 1, \dots, m \right\} = \\ &= \langle w_{m+1}, \dots, w_n \rangle \end{aligned}$$

Quindi, dato un sottospazio T di $\mathbb{R}^n$, ogni vettore v di $\mathbb{R}^n$ si può decomporre in modo unico come somma $v = v_{//} + v_\perp$ ove $v_{//} = \alpha_1 w_1 + \dots + \alpha_m w_m \in T$ (detto componente di v parallela a T) e $v_\perp = \alpha_{m+1} w_{m+1} + \dots + \alpha_n w_n \in T^\perp$ (detto componente di v ortogonale a T).

Grazie alla Proposizione 5.3.4, possiamo definire un funzione $p_T : \mathbb{R}^n \rightarrow T$ che ad ogni vettore v associa la componente di v parallela a T , $v \mapsto v_{//} = p_T(v)$. Tale funzione si dice **proiezione ortogonale** su T .

Osservazione 5.3.5 Come si calcola la proiezione ortogonale di un vettore $v \in \mathbb{R}^n$ su un sottospazio T di $\mathbb{R}^n$?

Sia $\{u_1, \dots, u_m\}$ una base ortonormale di T . Allora

$$v_{//} = p_U(v) = (v \bullet u_1)u_1 + \dots + (v \bullet u_m)u_m. \quad (5.2)$$

Per dimostrare la formula (5.2) procediamo nel modo seguente: completiamo $\{u_1, \dots, u_m\}$ in una base ortonormale $\mathcal{B} = \{u_1, \dots, u_m, u_{m+1}, \dots, u_n\}$

di $\mathbb{R}^n$ (come abbiamo fatto nella dimostrazione della Proposizione 5.3.4). In particolare abbiamo già notato che i vettori $\{u_{m+1}, \dots, u_n\}$ individuano una base di $T^\perp$. Essendo $\mathcal{B}$ una base di $\mathbb{R}^n$, il vettore v si scrive (in modo unico) come combinazione lineare dei vettori di $\mathcal{B}$:

$$v = \alpha_1 u_1 + \dots + \alpha_n u_n.$$

Per calcolare i coefficienti u_i osserviamo che:

$$v \bullet u_i = \alpha_1 (u_1 \bullet u_i) + \dots + \alpha_n (u_n \bullet u_i).$$

Ma, essendo $\mathcal{B}$ una base ortonormale, i prodotti scalari $u_k \bullet u_i$ sono nulli per ogni $k \neq i$, e $u_i \bullet u_i = 1$, quindi

$$v \bullet u_i = \alpha_i,$$

cioè:

$$v = (v \bullet u_1)u_1 + \dots + (v \bullet u_n)u_n.$$

Come già osservato nella dimostrazione della Proposizione 5.3.4), il vettore $(v \bullet u_1)u_1 + \dots + (v \bullet u_m)u_m$ appartiene al sottospazio T e il vettore $(v \bullet u_{m+1})u_{m+1} + \dots + (v \bullet u_n)u_n$ appartiene al sottospazio $T^\perp$ e da questo segue la formula (5.2).

Osservazione 5.3.6 Finora abbiamo fornito la descrizione qualitativa dell'ortogonale di un sottospazio T , ma non abbiamo calcolato esplicitamente chi sia lo spazio $T^\perp$. Adesso vogliamo fare qualche esempio esplicito. Possiamo lavorare direttamente in coordinate. Dato, ad esempio, $T = \langle (1, 2, 3) \rangle$, vogliamo determinare $T^\perp$. Pur non sapendo ancora chi sia $T^\perp$, sappiamo che $\dim(T) + \dim(T^\perp) = 3$, dunque $\dim T^\perp = 2$. I vettori $(x_1, x_2, x_3) \in T^\perp$ devono soddisfare la condizione $t \bullet (x_1, x_2, x_3) = 0$ per ogni $t \in T$, in particolare (ed è sufficiente) per ogni vettore t di una base di T . Si ha dunque:

$$(1, 2, 3) \bullet (x_1, x_2, x_3) = x_1 + 2x_2 + 3x_3 = 0.$$

Abbiamo trovato un sistema di rango 1 in 3 incognite: sappiamo che lo spazio delle soluzioni ha dimensione 2, esattamente come volevamo! In altre parole abbiamo descritto $T^\perp$ con un'equazione cartesiana. Troviamone ora una descrizione con equazioni parametriche (risolvendo il sistema) e da questa descrizione otterremo un insieme di generatori. Risolviamo il sistema.

$$\begin{cases} x_1 = -2\lambda - 3\mu \\ x_2 = \lambda \\ x_3 = \mu \end{cases}$$

Bastano due soluzioni linearmente indipendenti per determinare un insieme di generatori, cioè basta scegliere prima $(\lambda, \mu) = (1, 0)$ e poi $(\lambda, \mu) = (0, 1)$; da cui otteniamo: $T^\perp = \langle (-2, 1, 0), (-3, 0, 1) \rangle$.

Se, ancora più difficile, $T = \langle (1, 2, 0), (0, -1, 1) \rangle$, allora $T^\perp$ è dato dai vettori (x_1, x_2, x_3) tali che $(1, 2, 0) \bullet (x_1, x_2, x_3) = 0$ e $(0, -1, 1) \bullet (x_1, x_2, x_3) = 0$, quindi $T^\perp$ è l'insieme delle soluzioni del sistema:

$$\begin{cases} x_1 + 2x_2 = 0 \\ -x_2 + x_3 = 0. \end{cases}$$

Abbiamo ottenuto un sistema omogeneo di rango 2 in 3 incognite che quindi avrà come insieme delle soluzioni uno spazio vettoriale di dimensione 1. Con qualche calcolo si ottiene $T^\perp = \langle (-2, 1, 1) \rangle$.

Supponiamo ora di voler determinare la decomposizione del vettore $(1, 0, 0)$ come somma di un vettore di T e di uno di $T^\perp$. Innanzitutto osserviamo che una base ortonormale di $T^\perp$ è data dal versore $\frac{1}{\sqrt{6}}(-2, 1, 1)$ il che ci consente di calcolare immediatamente la componente di $(1, 0, 0)$ ortogonale a T , i.e. la componente di $(1, 0, 0)$ parallela a $T^\perp$: $((1, 0, 0) \bullet \frac{1}{\sqrt{6}}(-2, 1, 1)) \frac{1}{\sqrt{6}}(-2, 1, 1) = -\frac{2}{\sqrt{6}} \frac{1}{\sqrt{6}}(-2, 1, 1) = (\frac{2}{3}, -\frac{1}{3}, -\frac{1}{3})$. Allora la componente di $(1, 0, 0)$ parallela a T è data da $(1, 0, 0) - (\frac{2}{3}, -\frac{1}{3}, -\frac{1}{3}) = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. Quindi la decomposizione del vettore $(1, 0, 0)$ è: $(1, 0, 0) = (1, 0, 0)_{//} + (1, 0, 0)_{\perp} = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}) + (\frac{2}{3}, -\frac{1}{3}, -\frac{1}{3})$.

Nota. Dato un qualsiasi sottospazio T di $\mathbb{R}^n$ vale la relazione $(T^\perp)^\perp = T$.

Osservazione 5.3.7 Quando due vettori sono ortogonali tra loro sono automaticamente linearmente indipendenti. Infatti se prendiamo uno spazio T e prendiamo un vettore v ad esso ortogonale, allora v non può essere scritto come combinazione lineare di vettori di T ed è quindi linearmente indipendente dai vettori di T . In un certo senso l'essere ortogonale a T è il miglior modo per un vettore di essere linearmente indipendente da T , non avendo esso alcuna componente lungo T .

Esempio 5.3.8 Vediamo ora come il procedimento di Gram-Schmidt fornisca anche un metodo per calcolare le proiezioni ortogonali di un vettore su un dato sottospazio. In $\mathbb{R}^3$ sia dato il sottospazio $T = \langle (1, 2, 3) \rangle$. Vogliamo determinare le componenti di $v = (0, 1, 4)$ parallela e ortogonale a T . Supponiamo allora di avere una base ortonormale o_1, o_2, o_3 di $\mathbb{R}^3$ con $\langle o_1 \rangle = \langle (1, 2, 3) \rangle$ e $\langle (1, 2, 3) \rangle^\perp = \langle o_2, o_3 \rangle$, allora $v - (v \bullet o_1)o_1$ appartiene al sottospazio $\langle (1, 2, 3) \rangle^\perp$. In questo modo otteniamo la decomposizione $v = (v \bullet o_1)o_1 + (v - (v \bullet o_1)o_1)$ in

cui $(v \bullet o_1)o_1 \in \langle (1, 2, 3) \rangle$ e $v - (v \bullet o_1)o_1 \in \langle (1, 2, 3) \rangle^\perp$. Adesso passiamo ai calcoli: o_1 è semplicemente una base ortonormale di uno spazio di dimensione 1, cioè la normalizzazione di $(1, 2, 3)$: $o_1 = (\frac{1}{\sqrt{14}}, \frac{2}{\sqrt{14}}, \frac{3}{\sqrt{14}}) = \frac{1}{\sqrt{14}}(1, 2, 3)$. Quindi la componente di $(0, 1, 4)$ lungo $\langle (1, 2, 3) \rangle$ nella decomposizione ortogonale è $((0, 1, 4) \bullet \frac{1}{\sqrt{14}}(1, 2, 3))o_1 = (\frac{1}{\sqrt{14}}(0, 1, 4) \bullet (1, 2, 3))o_1 = \sqrt{14}o_1$.

Esempio 5.3.9 In $\mathbb{R}^4$ dotato del prodotto scalare usuale, siano dati i vettori linearmente indipendenti

$$v_1 = (1, 2, 0, -1), v_2 = (0, 2, 1, 1), v_3 = (1, 1, 1, 1).$$

Determinare una base ortonormale w_1, w_2, w_3 di $V = \langle v_1, v_2, v_3 \rangle$. Determinare $V^\perp$ e descrivere la proiezione ortogonale di $(1, 2, 0, 0)$ su V .

Svolgimento. Applichiamo il metodo di Gram-Schmidt. Normalizziamo v_1 : la sua norma è $\sqrt{1+4+1} = \sqrt{6}$. Quindi $w_1 = (\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, 0, -\frac{1}{\sqrt{6}})$. Consideriamo poi $v_2 - (v_2 \bullet w_1)w_1 = (0, 2, 1, 1) - ((0, 2, 1, 1) \bullet (\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, 0, -\frac{1}{\sqrt{6}}))(\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, 0, -\frac{1}{\sqrt{6}}) = (0, 2, 1, 1) - \frac{3}{\sqrt{6}}(\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, 0, -\frac{1}{\sqrt{6}}) = (0, 2, 1, 1) - (\frac{1}{2}, 1, 0, -\frac{1}{2}) = (-\frac{1}{2}, 1, 1, \frac{3}{2})$. La sua norma è $\sqrt{\frac{1}{4} + 1 + 1 + \frac{9}{4}} = \sqrt{\frac{9}{2}}$. Ne segue che $w_2 = \frac{\sqrt{2}}{3}(-\frac{1}{2}, 1, 1, \frac{3}{2})$. Infine l'ultimo vettore di cui abbiamo bisogno è il vettore $(1, 1, 1, 1) - ((1, 1, 1, 1) \bullet (\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, 0, -\frac{1}{\sqrt{6}}))(\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, 0, -\frac{1}{\sqrt{6}}) - ((1, 1, 1, 1) \bullet \frac{\sqrt{2}}{3}(-\frac{1}{2}, 1, 1, \frac{3}{2}))\frac{\sqrt{2}}{3}(-\frac{1}{2}, 1, 1, \frac{3}{2}) = (1, 1, 1, 1) - \frac{2}{\sqrt{6}}(\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, 0, -\frac{1}{\sqrt{6}}) - \frac{2}{3}(-\frac{1}{2}, 1, 1, \frac{3}{2}) = (1, 1, 1, 1) - (\frac{1}{3}, \frac{2}{3}, 0, -\frac{1}{3}) - (-\frac{1}{3}, \frac{2}{3}, \frac{2}{3}, 1) = (1, -\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$, la cui norma è $\sqrt{1 + 3\frac{1}{9}} = \frac{2}{\sqrt{3}}$. Il terzo vettore della base ortonormale è allora $w_3 = \frac{\sqrt{3}}{2}(1, -\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. Per trovare adesso $V^\perp$ occorre determinare i vettori (x, y, z, t) che siano ortogonali ai vettori di una base di V : $(x, y, z, t) \bullet (1, 2, 0, -1) = (x, y, z, t) \bullet (0, 2, 1, 1) = (x, y, z, t) \bullet (1, 1, 1, 1) = 0$ Si ottiene il sistema

$$\begin{cases} x + 2y - t = 0 \\ 2y + z + t = 0 \\ x + y + z + t = 0 \end{cases}$$

che è un sistema di rango 3 (le righe della matrice associata al sistema sono le coordinate dei vettori v_1, v_2, v_3 che sono linearmente indipendenti), pertanto lo spazio delle sue soluzioni ha dimensione $4 - 3 = 1$. Abbiamo: $\langle (1, 1, -5, 3) \rangle = V^\perp$. La norma del vettore $(1, 1, -5, 3)$ è 6. Dobbiamo ora determinare la proiezione ortogonale di $(1, 2, 0, 0)$ su V . Conviene calcolare la sua componente lungo $V^\perp$ cioè $((1, 2, 0, 0) \bullet (\frac{1}{6}, \frac{1}{6}, -\frac{5}{6}, \frac{1}{2}))(\frac{1}{6}, \frac{1}{6}, -\frac{5}{6}, \frac{1}{2}) =$

$(\frac{1}{12}, \frac{1}{12}, -\frac{5}{12}, \frac{1}{4})$. Questa è la componente di $(1, 2, 0, 0)$ ortogonale a V . La proiezione ortogonale di $(1, 2, 0, 0)$ su V è allora

$$(1, 2, 0, 0) - \left(\frac{1}{12}, \frac{1}{12}, -\frac{5}{12}, \frac{1}{4} \right) = \left(\frac{11}{12}, \frac{23}{12}, \frac{5}{12}, -\frac{1}{4} \right).$$

5.4 Esercizi svolti

Esercizio 5.4.1 Determinare una base ortonormale del sottospazio W di $\mathbb{R}^3$ generato dai vettori $v_1 = (-1, 1, -1)$, $v_2 = (2, 1, 4)$, $v_3 = (4, -1, 6)$.

Svolgimento. Ci chiediamo innanzitutto che dimensione abbia W : i vettori v_1 e v_2 sono linearmente indipendenti e per stabilire se il vettore v_3 è loro combinazione lineare, calcoliamo il seguente rango:

$$rg \begin{pmatrix} -1 & 1 & -1 \\ 2 & 1 & 4 \\ 4 & -1 & 6 \end{pmatrix} = 2.$$

Dunque una base di W è data dai vettori v_1 e v_2 .

I vettori v_1 e v_2 non sono ortogonali: $(-1, 1, -1) \bullet (2, 1, 4) = -2 + 1 - 4 = -5$. Utilizziamo allora il procedimento di ortonormalizzazione di Gram-Schmidt per ottenere, a partire dalla base $\{v_1, v_2\}$, una base ortonormale di W . Per prima cosa normalizziamo il vettore v_1 e poniamo $w_1 = \frac{v_1}{\|v_1\|} = \frac{1}{\sqrt{3}}(-1, 1, -1)$. Costruiamo ora il vettore $w'_2 = v_2 - (v_2 \bullet w_1)w_1 = (2, 1, 4) - \frac{1}{3}((2, 1, 4) \bullet (-1, 1, -1))(-1, 1, -1) = (2, 1, 4) + \frac{5}{3}(-1, 1, -1) = (\frac{1}{3}, \frac{8}{3}, \frac{7}{3})$.

Non resta che normalizzare il vettore w'_2 : abbiamo $\|w'_2\| = \frac{\sqrt{114}}{3}$, pertanto poniamo $w_2 = \frac{w'_2}{\|w'_2\|} = (\frac{1}{\sqrt{114}}, \frac{8}{\sqrt{114}}, \frac{7}{\sqrt{114}})$. L'insieme $\{w_1, w_2\}$ è dunque una base ortonormale di W .

Esercizio 5.4.2 Sia $U = \{(x, y, z) \in \mathbb{R}^3 \mid 2x + y = 0\}$. Determinare una base ortonormale di U .

Svolgimento. L'insieme U è l'insieme di tutti e soli gli elementi di $\mathbb{R}^3$ della forma $(x, -2x, z)$ con $x, z \in \mathbb{R}$. Pertanto una base di U è data dai vettori $v_1 = (1, -2, 0)$, $v_2 = (0, 0, 1)$. Osserviamo che i vettori v_1 e v_2 sono ortogonali: $v_1 \bullet v_2 = (1, -2, 0) \bullet (0, 0, 1) = 0$. Allora per ottenere una base ortonormale di U basta normalizzare i vettori v_1, v_2 : otteniamo così la base ortonormale $w_1 = \frac{1}{\sqrt{5}}(1, -2, 0)$, $w_2 = (0, 0, 1)$.

Esercizio 5.4.3 Siano $A = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}$, $z = (1, 0, -2)$. Determinare l'equazione del sottospazio U di $\mathbb{R}^3$ definito da:

$$U = \{u \in \mathbb{R}^3 : z \bullet Au = 0\}.$$

Svolgimento. Si tratta semplicemente di riscrivere la condizione $z \bullet Au = 0$ in coordinate. Perciò, indicate con (x, y, z) le coordinate del vettore u , si ha:

$$z \bullet Au = (1, 0, -2) \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = (1, 0, -2) \begin{pmatrix} x - y \\ -x + 2y - z \\ -y + z \end{pmatrix} = x + y - 2z = 0.$$

L'equazione del sottospazio U è dunque $x + y - 2z = 0$.

Esercizio 5.4.4 Sia $T = \{(x, y, z) \in \mathbb{R}^3 \mid x + y = 0 = 2y + z\}$.

- 1) Determinare $T^\perp$.
- 2) Determinare la proiezione ortogonale del vettore $u = (1, 0, -1)$ su T .

Svolgimento.

- 1) L'insieme T è un sottospazio di $\mathbb{R}^3$ di dimensione 1. Si ha infatti:

$$T = \{(-y, y, -2y) \mid y \in \mathbb{R}\} = \langle (1, -1, 2) \rangle.$$

Per definizione l'ortogonale di T è l'insieme di tutti e soli i vettori di $\mathbb{R}^3$ ortogonali ad ogni elemento di T , pertanto

$$T^\perp = \{(x, y, z) \mid (x, y, z) \bullet (1, -1, 2) = 0\} = \{(x, y, z) \mid x - y + 2z = 0\}.$$

Naturalmente la dimensione di $T^\perp$ è due, dal momento che $\dim(\mathbb{R}^3) = \dim(T) + \dim(T^\perp)$. Una base di $T^\perp$ è data dai vettori $(1, 1, 0)$ e $(0, 2, 1)$.

- 2) Ogni vettore $v \in \mathbb{R}^3$ si scrive, in modo unico, come somma di un vettore v_T di T e di un vettore $v^\perp$ di $T^\perp$: $v = v_T + v^\perp$. La componente v_T è la proiezione ortogonale del vettore v sul sottospazio T .

La proiezione ortogonale del vettore u su T è dunque:

$$u_T = \frac{u \bullet (1, -1, 2)}{(1, -1, 2) \bullet (1, -1, 2)} (1, -1, 2) = \frac{(1, 0, -1) \bullet (1, -1, 2)}{6} (1, -1, 2) = -\frac{1}{6} (1, -1, 2).$$

Esercizio 5.4.5 Determinare la proiezione ortogonale del vettore $u = (1, 0, 0)$ sul sottospazio V di $\mathbb{R}^3$ generato dai vettori $v_1 = (-2, 1, 0)$, $v_2 = (1, 1, 1)$.

Svolgimento. Sappiamo che il vettore $u = u_{//} + u^\perp$ con $u_{//} \in V$ e $u^\perp \in V^\perp$. Dobbiamo determinare $u_{//}$ cioè la proiezione ortogonale di u su V . Possiamo procedere in 2 modi equivalenti:

1) Determiniamo una base ortonormale di V utilizzando il metodo di ortonormalizzazione di Gram-Schmidt: poniamo

$$w_1 = \frac{v_1}{\|v_1\|} = \frac{1}{\sqrt{5}}(-2, 1, 0),$$

$$w'_2 = v_2 - (v_2 \bullet w_1)w_1 = (1, 1, 1) + \frac{1}{5}(-2, 1, 0) = (\frac{3}{5}, \frac{6}{5}, 1)$$

e, infine, $w_2 = \frac{w'_2}{\|w'_2\|} = \frac{1}{\sqrt{70}}(3, 6, 5)$. I vettori w_1 e w_2 costituiscono una base ortonormale di V . Allora la proiezione ortogonale di u su V è data da:

$$u_{//} = ((1, 0, 0) \bullet w_1)w_1 + ((1, 0, 0) \bullet w_2)w_2 = -\frac{2}{5}(-2, 1, 0) + \frac{3}{70}(3, 6, 5) = (\frac{13}{14}, -\frac{1}{7}, \frac{3}{14}).$$

2) Poiché si ha: $u_{//} = u - u^\perp$ possiamo determinare $u^\perp$ e calcolare $u_{//}$ come differenza di u e $u^\perp$. Dal momento che V ha dimensione 2, $V^\perp$ ha dimensione 1, pertanto calcolare $u^\perp$, cioè la proiezione ortogonale del vettore u su $V^\perp$, richiede certamente meno calcoli rispetto a quelli utilizzati in 1).

Occorre tuttavia determinare $V^\perp$:

$$V^\perp = \{(x, y, z) \in \mathbb{R}^3 \mid (x, y, z) \bullet (-2, 1, 0) = 0 = (x, y, z) \bullet (1, 1, 1)\} = \{(x, y, z) \mid -2x + y = 0 = x + y + z\} = \{(x, 2x, -3x) \mid x \in \mathbb{R}\} = \langle (1, 2, -3) \rangle.$$

Così una base ortonormale di $V^\perp$ è data dal vettore $z = \frac{1}{\sqrt{14}}(1, 2, -3)$. Abbiamo dunque

$$u^\perp = (u \bullet z)z = \frac{1}{14}(1, 2, -3)$$

e

$$u_{//} = u - u^\perp = (1, 0, 0) - \frac{1}{14}(1, 2, -3) = (\frac{13}{14}, -\frac{1}{7}, \frac{3}{14}).$$

5.5 Esercizi proposti

Esercizio 5.5.1 Sia $V = \{(x, y, z, t) \in \mathbb{R}^4 \mid x + y - z + t = 0\}$.

1. Determinare una base ortonormale $\mathcal{B}$ di V .
2. Completare la base $\mathcal{B}$ in una base di $\mathbb{R}^4$.
3. Completare la base $\mathcal{B}$ in una base ortonormale di $\mathbb{R}^4$.

Esercizio 5.5.2 Sia $S = \{(x, y, z) \in \mathbb{R}^3 \mid 2x - y - z = 0\}$.

1. Determinare una base ortonormale di S .
2. Determinare $S^\perp$.
3. Determinare, se possibile, un vettore $v \in \mathbb{R}^3$ la cui proiezione ortogonale su S sia $(1, 1, 0)$.
4. Determinare, se possibile, un vettore $w \in \mathbb{R}^3$ la cui proiezione ortogonale su S sia $(1, 2, 0)$ e la cui proiezione ortogonale su $S^\perp$ sia $(4, -2, -2)$.
5. Determinare, se possibile, un vettore $z \in \mathbb{R}^3$ la cui proiezione ortogonale su S sia $(0, 0, 0)$ e la cui proiezione ortogonale sul sottospazio $W = \langle (1, 1, 1) \rangle$ di $\mathbb{R}^3$ sia $(1, 1, 1)$.

Esercizio 5.5.3 Sia $S = \{(x, y, z) \in \mathbb{R}^3 \mid 2x + 3y + 5z = 0\}$ e sia $L : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ la funzione che ad ogni vettore v di $\mathbb{R}^3$ associa la sua proiezione ortogonale su S .

1. Si determinino i vettori v di $\mathbb{R}^3$ tali che $L(v) = (0, 0, 0)$.
2. Si determinino i vettori w di $\mathbb{R}^3$ tali che $w = L(v)$ per qualche $v \in \mathbb{R}^3$.
3. Esistono vettori z di $\mathbb{R}^3$ tali che $L(z) = z$?

Lezione 6

Spazi metrici

Nella Lezione 4 ci siamo occupati di intersezione e parallelismo di piani e rette nel piano e nello spazio tridimensionale. Queste proprietà si chiamano proprietà affini. In questa lezione introdurremo alcune considerazioni di carattere metrico, cioè legate alla nozione di distanza. Procediamo con ordine: nella Lezione 5 abbiamo introdotto il concetto di norma in $\mathbb{R}^n$. In questo modo siamo in grado di associare ad un vettore v un numero reale non negativo che chiamiamo la norma di v . A questo punto possiamo definire la distanza tra due punti P e Q come la norma del vettore $P - Q$. Cerchiamo di essere più precisi.

6.1 Spazio Euclideo

Definizione 6.1.1 *Chiameremo spazio euclideo n -dimensionale il dato di:*

- *un insieme di punti X che identificheremo con $\mathbb{R}^n$;*
- *lo spazio vettoriale $\mathbb{R}^n$ dotato del prodotto scalare euclideo;*
- *un'azione di $\mathbb{R}^n$ su X cioè una funzione:*

$$\begin{array}{ccc} X \times \mathbb{R}^n & \longrightarrow & X \\ (P, v) & \longmapsto & Q = P + v, \end{array}$$

essendo $Q = P + v$ il punto che si raggiunge applicando il vettore v nel punto P .

In questo modo si ha, ad esempio, che la distanza tra $P_1 = (1, 3)$ e $P_2 = (1, 1)$ è data dalla norma del vettore $P_1 - P_2$ (che coinciderà con la norma del suo opposto $P_2 - P_1$): $\|P_1 - P_2\| = 2$. Nei prossimi paragrafi lavoreremo sul piano o sullo spazio euclideo tridimensionale. Le nozioni trattate in uno dei due ambienti saranno facilmente adattabili all'altro.

Osservazione 6.1.2 Il fatto di aver definito il piano o lo spazio euclideo come *usuale*, è ovviamente legato al fatto di aver usato sullo spazio vettoriale associato al piano o allo spazio tridimensionale il prodotto scalare usuale.

6.2 Perpendicolarità

Parlando di prodotti scalari abbiamo introdotto la nozione di ortogonalità tra due vettori. Possiamo quindi introdurre la nozione di sottovarietà ortogonali.

Definizione 6.2.1 Due rette r_1 e r_2 sono ortogonali se i rispettivi vettori direttori v_1 e v_2 sono ortogonali;

una retta r è ortogonale ad un piano π se il vettore direttore v di r è ortogonale allo spazio direttore V di π : $v \in V^\perp$;

due piani π_1 e π_2 , di giacitura V_1 e V_2 , rispettivamente, sono ortogonali se $V_1^\perp$ è ortogonale a $V_2^\perp$.

A partire da questa definizione si possono fare interessanti osservazioni legate alle dimensioni. Questo dipende dal fatto che la dimensione di V_1 (risp. di V_2) determina la dimensione di $V_1^\perp$ (risp. di $V_2^\perp$).

Esempio 6.2.2 Nello spazio euclideo tridimensionale un piano π ha equazione $ax + by + cz + d = 0$ e il suo spazio direttore è il sottospazio di $\mathbb{R}^3$ di dimensione 2 dato da $\langle v, w \rangle$ ove $v, w \in \mathbb{R}^3$ e le loro coordinate soddisfanno l'equazione $ax + by + cz = 0$. Del resto questa equazione si può pure leggere come $(a, b, c) \bullet (x, y, z) = 0$, allora dire che i vettori v e w soddisfanno l'equazione $ax + by + cz = 0$ equivale a dire che $(a, b, c) \bullet v = (a, b, c) \bullet w = 0$ cioè $\langle (a, b, c) \rangle = \langle v, w \rangle^\perp$. Ne deduciamo che, dato il piano π di equazione $ax + by + cz + d = 0$ il vettore (a, b, c) è ortogonale a π (intendendo ortogonale alla sua giacitura). Se P è un punto del piano π allora la retta per P perpendicolare a π è $P + \langle (a, b, c) \rangle$.

Esempio 6.2.3 Nello spazio euclideo tridimensionale consideriamo la retta $r : (1, 2, 3) + \langle(1, 1, 1)\rangle$. Troviamo una retta ad essa perpendicolare. Una retta perpendicolare a r è data allora, per definizione, da una retta del tipo $P + \langle(\alpha, \beta, \gamma)\rangle$ con $(\alpha, \beta, \gamma) \bullet (1, 1, 1) = 0$. Notiamo che $\dim\langle(1, 1, 1)\rangle^\perp = 2$. Per determinare $\langle(1, 1, 1)\rangle^\perp$ dobbiamo trovare le soluzioni dell'equazione $(1, 1, 1) \bullet (x, y, z) = 0$, cioè $x + y + z = 0$. Si ha: $\langle(1, 1, 1)\rangle^\perp = \langle(1, -1, 0), (0, 1, -1)\rangle$. Una retta per P perpendicolare alla retta r è dunque una retta per P la cui giacitura appartiene allo spazio ora trovato cioè una retta della forma $P + \langle(\alpha, \beta - \alpha, -\beta)\rangle$ per qualche $\alpha, \beta \in \mathbb{R}$, $(\alpha, \beta) \neq (0, 0)$. In particolare esistono infinite rette per P perpendicolari alla retta r , mentre il piano per P ortogonale a r è unico. Infatti tale piano ha giacitura individuata da $\langle(1, 1, 1)\rangle^\perp$ ed è quindi il piano $P + \langle(1, 1, 1)\rangle^\perp$. Un qualsiasi piano perpendicolare alla retta data ha automaticamente giacitura fissata ed è quindi semplicemente individuato dal passaggio per un punto. Imporre che un punto $P_0 = (x_0, y_0, z_0)$ appartenga ad un piano π di equazione $ax + by + cz + d = 0$ equivale a richiedere che (x_0, y_0, z_0) sia soluzione particolare dell'equazione di π . L'equazione di un qualsiasi piano ortogonale alla retta r è $x + y + z + d = 0$ al variare di $d \in \mathbb{R}$. Si noti che l'insieme di piani di equazione $x + y + z + d = 0$, al variare di $d \in \mathbb{R}$, è il fascio di piani paralleli al piano $x + y + z = 0$.

Esercizio 6.2.4 Nel piano euclideo si descriva l'asse del segmento di estremi $A = (1, 2)$ e $B = (3, -4)$.

Svolgimento. L'asse del segmento di estremi A e B è per definizione il luogo dei punti equidistanti dai due punti dati A e B . Si vede subito che tale insieme è dato dalla retta che passa per il punto medio del segmento di vertici A e B ed è ortogonale alla direzione da esso individuata (N.B. Che cosa vuol dire: “si vede subito”? In effetti occorre mostrare che ogni punto che sia equidistante da A e B sta effettivamente su tale retta e che ogni punto di questa retta è equidistante da A e B . Per questo si possono utilizzare le proprietà dei triangoli isosceli). Determiniamo questa retta. Il punto medio del segmento di estremi A e B è il punto $A + \frac{1}{2}(B - A)$, ove con $\frac{1}{2}(B - A)$ intendiamo il prodotto del vettore $B - A$ per lo scalare $\frac{1}{2}$. Quello che stiamo facendo è sommare ad A “metà” del vettore che occorrerebbe per ottenere B . Ne segue che il punto medio M ha coordinate $(1, 2) + \frac{1}{2}(3 - 1, -2 - 4) = (2, -1)$. La direzione individuata dal segmento (cioè la giacitura della retta individuata dai due punti) è $\langle(2, -6)\rangle = \langle(1, -3)\rangle$, pertanto l'equazione della retta ad esso ortogonale sarà del tipo $x - 3y + c = 0$. Imponendo il passaggio per il

punto medio M si ottiene $2 + 3 + c = 0$ cioè $c = -5$; l'asse del segmento di estremi A e B è pertanto la retta di equazione $x + 3y - 5 = 0$.

Si noti che, in generale, dati due punti $A = (x_1, y_1)$ e $B = (x_2, y_2)$ del piano euclideo, il loro punto medio ha coordinate $(\frac{x_1+x_2}{2}, \frac{y_1+y_2}{2})$.

Esercizio 6.2.5 Vogliamo adesso svolgere un esercizio analogo all'esercizio 6.2.4 nel caso tridimensionale. Nello spazio euclideo tridimensionale si descriva l'insieme dei punti equidistanti da $A = (1, 2, 0)$ e $B = (3, -4, 2)$.

Svolgimento. In questo caso il luogo dei punti è dato da un piano ortogonale alla direzione individuata dal segmento e passante per il suo punto medio (anche questo andrebbe dimostrato...). Per il punto medio abbiamo, come al solito, $M = A + \frac{1}{2}(B - A)$. In coordinate si ha: $M = (1, 2, 0) + \frac{1}{2}(3 - 1, -4 - 2, 2)$, cosicché: $M = (2, -1, 1)$. Ora dobbiamo determinare il piano ortogonale alla direzione individuata dal segmento, cioè la direzione del vettore $A - B = (-2, 6, -2)$. La giacitura di un piano ortogonale a tale direzione è $x - 3y + z = 0$, quindi i piani ad essa paralleli (cioè aventi tale giacitura) hanno equazione del tipo: $x - 3y + z + d = 0$; il passaggio per M impone $2 + 3 + 1 + d = 0$, cioè $d = -6$. Otteniamo così il piano cercato: $x - 3y + z - 6 = 0$.

Esercizio 6.2.6 Nello spazio euclideo tridimensionale consideriamo i tre punti $A = (1, 2, 3)$, $B = (0, -1, 3)$ e $C = (-2, 3, 6)$. Si determini (1) il baricentro del triangolo di vertici A, B, C ; (2) il piano individuato da questi tre punti.

Svolgimento. Per quanto riguarda (1), si ricordi che il baricentro è il luogo d'incontro delle mediane del triangolo cioè delle rette che passano per un vertice del triangolo e per il punto medio del lato opposto ad esso. Poiché le tre rette concorrono in uno stesso punto, basterà trovare l'intersezione tra due di esse. Consideriamo allora il vertice A e il lato opposto ad esso cioè il segmento individuato da B e C , il cui punto medio è, per quanto visto nell'esercizio precedente, il punto $M_{BC} = (\frac{0-2}{2}, \frac{-1+3}{2}, \frac{3+6}{2}) = (-1, 1, \frac{9}{2})$. La prima mediana è allora la retta $A + \langle A - M_{BC} \rangle$, cioè: $(1, 2, 3) + \langle (1 - (-1), 2 - 1, 3 - \frac{9}{2}) \rangle = (1, 2, 3) + \langle (4, 2, -3) \rangle$, dunque tutti i punti della retta per A e per il punto medio di B e C hanno coordinate $(1 + 4\beta, 2 + 2\beta, 3 - 3\beta)$ al variare di β nell'insieme dei numeri reali. Consideriamo poi la retta per B che passa per il punto medio del segmento di estremi A e C : il punto medio è $M_{AC} = (-\frac{1}{2}, \frac{5}{2}, \frac{9}{2})$ e la retta ha equazione parametrica $(0, -1, 3) + \langle (-\frac{1}{2}, \frac{5}{2} + 1, \frac{9}{2} - 3) \rangle = (0, -1, 3) + \langle (-1, 7, 3) \rangle$, i suoi punti sono quindi della

forma $(-\gamma, -1 + 7\gamma, 3 + 3\gamma)$, al variare di γ in $\mathbb{R}$. Il baricentro del triangolo di vertici A, B, C è il punto di incontro delle due rette trovate, quindi si debbono trovare i numeri reali β e γ per cui $(1 + 4\beta, 2 + 2\beta, 3 - 3\beta) = (-\gamma, -1 + 7\gamma, 3 + 3\gamma)$. Otteniamo il sistema

$$\begin{cases} -\gamma = 1 + 4\beta \\ 2 + 2\beta = -1 + 7\gamma \\ 3 - 3\beta = 3 + 3\gamma \end{cases}$$

che ha come soluzione la coppia $(\beta, \gamma) = (-\frac{1}{3}, \frac{1}{3})$. Il punto di intersezione delle mediane è pertanto il punto $(-\frac{1}{3}, \frac{4}{3}, 4)$. (In generale, dati i punti $P_1 = (x_1, y_1, z_1)$, $P_2 = (x_2, y_2, z_2)$, $P_3 = (x_3, y_3, z_3)$, il baricentro del triangolo di vertici P_1, P_2, P_3 è il punto di coordinate $(\frac{x_1+x_2+x_3}{3}, \frac{y_1+y_2+y_3}{3}, \frac{z_1+z_2+z_3}{3})$: per la dimostrazione si ripeta il conto precedente mantenendo le coordinate dei punti generiche. Attenzione... cosa succede se i punti sono allineati?)

Passiamo ora a (2): per determinare il piano individuato dai punti A, B, C dobbiamo determinare la sua giacitura ed un suo punto. Di punti ne abbiamo molti (almeno 3...): A, B o C . La giacitura deve contenere i vettori $B - A$ e $C - A$: infatti sommati ad A questi vettori danno, rispettivamente, i punti B e C , che sono punti del piano che cerchiamo. Ora, $C - A = (-3, 1, 3)$ e $B - A = (-1, -3, 0)$: i due vettori sono linearmente indipendenti e quindi lo spazio da essi generato ha dimensione 2, ne consegue che essi individuano la giacitura del piano (che è uno spazio vettoriale di dimensione 2). Risposta: il piano individuato dai punti A, B, C è il piano $(1, 2, 3) + \langle (-3, 1, 3), (-1, -3, 0) \rangle$. I suoi punti sono tutti e soli della forma $(1 - 3s - t, 2 + s - 3t, 3 + 3s)$ al variare di s, t nei reali.

6.3 Distanze tra sottovarietà lineari

Vogliamo ora mostrare come sia possibile definire la distanza tra due sottovarietà lineari.

In questo paragrafo affronteremo il problema di calcolare la distanza tra due sottovarietà lineari del piano e dello spazio euclideo. Siano $\mathcal{S}_1 = P_1 + V_1$ e $\mathcal{S}_2 = P_2 + V_2$ due sottovarietà lineari con P_1, P_2 punti dello spazio euclideo (risp. piano euclideo) e V_1, V_2 sottospazi vettoriali di $\mathbb{R}^3$ (risp. $\mathbb{R}^2$). Come possiamo definire la distanza tra questi due insiemi? Diamo la seguente definizione:

$$Dist(\mathcal{S}_1, \mathcal{S}_2) = \min_{P \in \mathcal{S}_1, Q \in \mathcal{S}_2} (\|P - Q\|)$$

cioè la minima tra le norme di tutti i vettori che congiungono un punto che appartiene alla prima sottovarietà e uno che appartiene alla seconda. Se da una parte questa definizione è facilmente comprensibile (significa che la distanza tra due isole è la minima tra le distanze di tutti i loro punti), dall'altra ricordiamo che $\mathcal{S}_1$ e $\mathcal{S}_2$ non sono insiemi finiti di punti e quindi non è scontato che il minimo in questione esista! Perché, piuttosto, nella definizione non abbiamo usato semplicemente l'estremo inferiore? Facciamo alcune osservazioni. Lavoriamo nello spazio euclideo ma segnaliamo che tutte le considerazioni si applicano anche al caso di dimensione 2.

Osservazione 6.3.1 Date le sottovarietà lineari dello spazio euclideo tridimensionale: $\mathcal{S}_1 = P_1 + V_1$ e $\mathcal{S}_2 = P_2 + V_2$, se $\mathcal{S}_1 \cap \mathcal{S}_2 \neq \emptyset$, allora $\text{Dist}(\mathcal{S}_1, \mathcal{S}_2) = 0$.

Per definizione la distanza fra due sottovarietà $\mathcal{S}_1$ e $\mathcal{S}_2$ è nulla se esistono un punto $P \in \mathcal{S}_1$ ed un punto $Q \in \mathcal{S}_2$ tali che $\|P - Q\| = 0$, i.e., tali che il vettore $P - Q$ coincida col vettore nullo o, equivalentemente, tali che P coincida con Q . Tali punti esistono se $\mathcal{S}_1 \cap \mathcal{S}_2 \neq \emptyset$.

Osservazione 6.3.2 Si tratta del classico Teorema di Pitagora rivisto in modo sapiente. Siano v_1 e v_2 vettori di $\mathbb{R}^3$ fra loro ortogonali: $v_1 \bullet v_2 = 0$. Allora $\|v_1\|^2 + \|v_2\|^2 = \|v_1 + v_2\|^2$. Cioè il quadrato della norma della somma dei vettori è la somma dei quadrati delle norme dei singoli vettori. La dimostrazione consiste in un semplice calcolo:

$$\|v_1 + v_2\|^2 = (v_1 + v_2) \bullet (v_1 + v_2) = v_1 \bullet v_1 + 2v_1 \bullet v_2 + v_2 \bullet v_2 = v_1 \bullet v_1 + v_2 \bullet v_2 = \|v_1\|^2 + \|v_2\|^2.$$

Esempio 6.3.3 Distanza di un punto da una retta nel piano euclideo. Consideriamo nel piano una retta r di equazione $ax + by + c = 0$ e un punto P_0 di coordinate (x_0, y_0) . La giacitura della retta è $\langle(-b, a)\rangle$ e lo spazio ad essa ortogonale è $\langle(-b, a)\rangle^\perp = \langle(a, b)\rangle$. Consideriamo un punto $P_1 = (x_1, y_1)$ della retta r , cioè un punto le cui coordinate soddisfino l'equazione della retta: $ax_1 + by_1 + c = 0$. Prendiamo il vettore che congiunge P_1 e $P_0 = (x_0, y_0)$: $P_0 - P_1 = (x_0 - x_1, y_0 - y_1)$ e determiniamone la componente parallela alla giacitura della retta e la componente perpendicolare: $P_0 - P_1 = (P_0 - P_1)_{//} + (P_0 - P_1)_\perp$ con $(P_0 - P_1)_{//} \in \langle(-b, a)\rangle$ e $(P_0 - P_1)_\perp \in \langle(-b, a)\rangle^\perp$. Il vettore $(P_0 - P_1)_{//}$ appartiene alla giacitura della retta r : ne consegue che il punto $Q = P_1 + (P_0 - P_1)_{//}$ appartiene a r . Osserviamo che $P_0 - Q = (P_0 - P_1)_\perp$ e quindi è un vettore ortogonale alla giacitura della retta. Diciamo che Q è il punto che realizza la minima

distanza tra P_0 e la retta r . Ogni altro punto di r si può scrivere come $Q + w$ ove $w \in \langle (-b, a) \rangle$; ora $P_0 - (Q + w) = (P_0 - Q) + (-w)$ è somma di due vettori ortogonali. Così $\|P_0 - (Q + w)\|^2 = \|P_0 - Q\|^2 + \|-w\|^2 \geq \|P_0 - Q\|^2$ (si veda l'Osservazione 6.3.2). Quindi qualsiasi punto della retta r dista da P_0 più di Q . La distanza di P_0 da r è allora data dalla norma del vettore $P_0 - Q$ che è la norma della componente ortogonale del vettore $P_0 - P_1$ rispetto alla giacitura della retta r . Per ottenere la componente ortogonale del vettore $P_0 - P_1$ normalizziamo il vettore (a, b) : $\frac{1}{\sqrt{a^2+b^2}}(a, b)$ e calcoliamo:

$$((P_0 - P_1) \bullet \frac{1}{\sqrt{a^2+b^2}}(a, b)) \frac{1}{\sqrt{a^2+b^2}}(a, b).$$

La sua norma è allora data da

$$\begin{aligned} |(x_0 - x_1, y_0 - y_1) \bullet \frac{1}{\sqrt{a^2+b^2}}(a, b)| &= \frac{|ax_0 + by_0 - ax_1 - by_1|}{\sqrt{a^2+b^2}} = \\ &= \frac{|ax_0 + by_0 + c|}{\sqrt{a^2+b^2}} = \text{Dist}(P_0, r). \end{aligned}$$

Abbiamo così fornito una formula generale per calcolare la distanza di un punto da una retta nel piano affine. Concludiamo questo esempio notando che la distanza tra due rette parallele r_1 e r_2 è data dalla distanza di un qualsiasi punto di r_1 da r_2 e quindi questo caso si riconduce all'esempio appena svolto.

Esempio 6.3.4 Distanza di un punto da un piano. Ci proponiamo di calcolare la distanza di un punto $P = (x_0, y_0, z_0)$ dello spazio euclideo tridimensionale da un piano π di equazione $ax + by + cz + d = 0$. Tale piano sarà individuato da un punto $P_1 = (x_1, y_1, z_1)$ e dal suo spazio direttore $\langle s, t \rangle$ di dimensione 2, con $s, t \in \mathbb{R}^3$, $s \bullet (a, b, c) = t \bullet (a, b, c) = 0$. Se vogliamo che $\text{Dist}(P, \pi) \neq 0$ occorre che $P - P_1 \notin \langle t, s \rangle$. Ne deriva che $\langle P - P_1, s, t \rangle = \mathbb{R}^3$. Ogni vettore $\mathbf{x}$ di $\mathbb{R}^3$ può quindi essere scritto come $\mathbf{x} = \mathbf{x}_{//} + \mathbf{x}_{\perp}$ con $\mathbf{x}_{//} \in \langle s, t \rangle$ e $\mathbf{x}_{\perp} \in \langle s, t \rangle^{\perp}$. Dire che $\mathbf{x}_{//} \in \langle s, t \rangle$ significa dire che $\mathbf{x}$ è un vettore della giacitura del piano π . Possiamo decomporre anche il vettore $P - P_1$: $P - P_1 = (P - P_1)_{//} + (P - P_1)_{\perp}$ con $(P - P_1)_{//} \in \langle s, t \rangle$. Ora prendiamo il punto $P_1 + (P - P_1)_{//}$: questo è un punto del piano π poiché P_1 è un punto del piano e il vettore $(P - P_1)_{//}$ appartiene alla sua giacitura; poniamo $Q = P_1 + (P - P_1)_{//}$. Osserviamo che $P - Q = (P - P_1)_{\perp}$ e quindi è un vettore ortogonale alla giacitura del piano. Affermiamo che Q è il punto di

minima distanza tra P e π . Infatti ogni altro punto di π si può scrivere come $Q + w$ con $w \in \langle s, t \rangle$. Allora $P - (Q + w) = (P - Q) + (-w)$ è somma di due vettori ortogonali. Così $\|P - (Q + w)\|^2 = \|P - Q\|^2 + \|-w\|^2 \geq \|P - Q\|^2$. Quindi qualsiasi punto di π diverso da Q dista da P più di Q . Come abbiamo trovato Q ? Q è il punto di intersezione della retta passante per P ortogonale al piano π con il piano π stesso: infatti il vettore $P - Q$ è ortogonale al piano! Quello che abbiamo fatto in questo esempio è stato prendere un vettore $P - P_1$, che a priori non era perpendicolare al piano, e sottrarre da esso la sua componente parallela al piano per ottenere un vettore perpendicolare a π . Grazie agli argomenti trattati nel paragrafo 5.2 possiamo allora dare una formula diretta per il calcolo della distanza tra il punto $P = (x_0, y_0, z_0)$ e il piano π di equazione $ax + by + cz + d = 0$: dobbiamo prendere un punto P_1 di π , $P_1 = (x_1, y_1, z_1)$, e la componente del vettore $P - P_1$ ortogonale alla giacitura del piano. La direzione ortogonale alla giacitura del piano π è data dal vettore (a, b, c) , pertanto la componente del vettore $P - P_1$ lungo questa direzione è $(\frac{1}{\sqrt{a^2+b^2+c^2}}(a, b, c) \bullet (x_0 - x_1, y_0 - y_1, z_0 - z_1)) \frac{1}{\sqrt{a^2+b^2+c^2}}(a, b, c)$. La norma del vettore trovato è: $|\frac{1}{\sqrt{a^2+b^2+c^2}}(a, b, c) \bullet (x_0 - x_1, y_0 - y_1, z_0 - z_1)|$, cioè:

$$\frac{|a(x_0 - x_1) + b(y_0 - y_1) + c(z_0 - z_1)|}{\sqrt{a^2 + b^2 + c^2}}.$$

Ma $ax_1 + by_1 + cz_1 + d = 0$ perché P_1 è un punto del piano e le sue coordinate soddisfano l'equazione di π , dunque $ax_1 + by_1 + cz_1 = -d$. Otteniamo così la formula della distanza di un punto $P_0 = (x_0, y_0, z_0)$ dal piano π di equazione $ax + by + cz + d = 0$:

$$d(P_0, \pi) = \frac{|ax_0 + by_0 + cz_0 + d|}{\sqrt{a^2 + b^2 + c^2}}.$$

Esempio 6.3.5 Distanza di un punto da una retta nello spazio euclideo tridimensionale.

Sia $P = (1, 0, 1)$ e sia r la retta di equazioni cartesiane

$$r : \begin{cases} x + y = 0 \\ z = 0. \end{cases}$$

Per determinare la distanza del punto P dalla retta r dobbiamo innanzitutto determinare il piano π passante per P ortogonale alla retta r . Questo piano interseca la retta r in un punto Q . Il punto Q è il punto che realizza il minimo della distanza di P dai punti della retta r (perché? Si noti che il punto P e la retta r individuano un piano). Si ha quindi:

$$\text{dist}(P, r) = \text{dist}(P, Q) = \|P - Q\|.$$

Il vettore direttore della retta r è $v = (1, -1, 0)$ pertanto il piano π ha equazione $x - y = d$ e passa per il punto P , dunque $d = 1$. Intersecando la retta r con il piano π otteniamo il punto $Q = (\frac{1}{2}, -\frac{1}{2}, 0)$, pertanto $\text{dist}(P, r) = \|(\frac{1}{2}, \frac{1}{2}, 1)\| = \sqrt{\frac{3}{2}}$.

Esempio 6.3.6 Distanza tra una retta ed un piano ad essa parallelo.

Consideriamo la retta $r : P + \langle v \rangle = (x_0, y_0, z_0) + \langle (l, m, n) \rangle$ e il piano π di equazione cartesiana $ax + by + cz + d = 0$. Il fatto che la retta sia parallela al piano significa che la giacitura di r è contenuta nella giacitura di π , cioè che $al + bm + cn = 0$. In questo caso tutti i punti della retta hanno la stessa distanza dal piano perciò la distanza tra la retta e il piano si calcola prendendo un qualsiasi punto della retta e determinando la sua distanza dal piano. Si noti che in questo caso i punti di minima distanza, cioè i punti $P \in r$ e $Q \in \pi$ che realizzano la distanza tra la retta e il piano, non sono unici.

Esempio 6.3.7 Distanza tra due rette nello spazio. Abbiamo due casi: o le rette sono sghembe oppure non lo sono.

Siano r_1 e r_2 due rette non sghembe. In questo frangente l'unico caso significativo è quello in cui le rette sono parallele e non coincidenti. Allora saranno del tipo: $r_1 : P + \langle (a, b, c) \rangle$ e $r_2 : Q + \langle (a, b, c) \rangle$ con $P - Q \notin \langle (a, b, c) \rangle$. Preso dunque un punto $P_1 \in r_1$, sia π il piano perpendicolare ad (a, b, c) passante per P_1 : chiamiamo Q_1 il punto di intersezione di π con la retta r_2 . La distanza tra le due rette è allora data da $\|P_1 - Q_1\|$. In questo caso la distanza tra r_1 e r_2 può essere realizzata non solo dai punti evidenziati, ma anche da altri punti nella stessa posizione reciproca.

Caso delle rette sghembe. Siano $r_1 : P + \langle (a, b, c) \rangle$ e $r_2 : Q + \langle (a', b', c') \rangle$ con $P - Q, (a, b, c), (a', b', c')$ vettori linearmente indipendenti e quindi una base di $\mathbb{R}^3$. Prendiamo la componente di $P - Q$ lungo $\langle (a, b, c), (a', b', c') \rangle$: essa si scriverà come $(\alpha a, \alpha b, \alpha c) + (\beta a', \beta b', \beta c')$ per certi $\alpha, \beta \in \mathbb{R}$. Ne segue che i punti $P_1 = P - (\alpha a, \alpha b, \alpha c)$ e $Q_1 = Q + (\beta a', \beta b', \beta c')$ appartengono rispettivamente alle rette r_1 e r_2 e il vettore che li congiunge è $P_1 - Q_1 = (P - (\alpha a, \alpha b, \alpha c)) - (Q + (\beta a', \beta b', \beta c')) = (P - Q) - (\alpha a, \alpha b, \alpha c) - (\beta a', \beta b', \beta c')$ e quindi appartiene a $\langle (a, b, c), (a', b', c') \rangle^\perp$. Se $\tilde{P}$ è un punto di r_1 e $\tilde{Q}$ è un punto di r_2 allora $\tilde{P} = P_1 + v$ con $v \in \langle (a, b, c) \rangle$ e $\tilde{Q} = Q_1 + w$ con $w \in \langle (a', b', c') \rangle$. Ne segue che $\tilde{P} - \tilde{Q} = (P_1 - Q_1) + (v - w)$ e $v - w$ è ortogonale a $P_1 - Q_1$. Si ha pertanto $\|\tilde{P} - \tilde{Q}\|^2 = \|(P_1 - Q_1)\|^2 + \|v - w\|^2 \geq \|(P_1 - Q_1)\|^2$. Quindi la distanza tra un qualsiasi punto di r_1 diverso da P_1 e un qualsiasi punto di

r_2 diverso da Q_1 è maggiore della distanza di P_1 da Q_1 : P_1 e Q_1 sono i punti di minima distanza tra le rette r_1 e r_2 . In questo caso i punti di minima distanza sono unici: si tratta dei punti che realizzano il minimo della distanza di due punti che stanno sulle due sottovarietà. La loro unicità dipende dal fatto che $P_1 - Q_1 \in \langle (a', b', c'), (a, b, c) \rangle^\perp$ e $\dim \langle (a', b', c'), (a, b, c) \rangle^\perp = 1$. Conclusione: date due rette sghembe r_1 e r_2 i punti di minima distanza tra le due rette sono univocamente determinati e sono i punti $P_1 \in r_1$ e $Q_1 \in r_2$ tali che il vettore $P_1 - Q_1$ sia ortogonale allo spazio vettoriale generato dalle giaciture delle due rette (cioè che sia ortogonale ad entrambe).

Esercizio 6.3.8 Esercizio sull'esempio. Nell'esempio precedente abbiamo individuato i punti che realizzano la minima distanza tra due rette sghembe: sono quelli per cui il vettore che li congiunge risulta ortogonale allo spazio vettoriale generato dalle due giaciture. Vogliamo sviluppare tale esempio spiegando un metodo per determinare tali punti. Siano $r_1 : (1, 2, 0) + \langle (1, 0, 0) \rangle$ e $r_2 : (1, 3, 4) + \langle (1, 1, 1) \rangle$. Tali rette sono sghembe (lo si verifichi per esercizio). I punti della prima retta sono della forma $(1+t, 2, 0)$ al variare di $t \in \mathbb{R}$, mentre i punti della seconda sono della forma $(1+r, 3+r, 4+r)$ al variare di $r \in \mathbb{R}$. Il vettore generico che congiunge un punto di r_1 e un punto di r_2 è allora dato da $(t-r, -1-r, -4-r)$: cerchiamo r e t per cui tale vettore sia ortogonale a $(1, 1, 1)$ e $(1, 0, 0)$. Dovrà essere: $(t-r, -1-r, -4-r) \bullet (1, 1, 1) = t-3r-5 = 0$ e $(t-r, -1-r, -4-r) \bullet (1, 0, 0) = t-r = 0$. Otteniamo l'unica soluzione $t = r = -\frac{5}{2}$ (perché è unica? Lo è sempre?). I punti di minima distanza tra r_1 e r_2 sono dunque $P_1 = (-\frac{3}{2}, 2, 0)$ e $P_2 = (-\frac{3}{2}, \frac{1}{2}, \frac{3}{2})$ e la distanza tra le due rette è data da $\|P_1 - P_2\| = \frac{3}{\sqrt{2}}$.

Esercizio 6.3.9 Date due rette sghembe r_1 e r_2 vogliamo determinare la retta r individuata dai punti di minima distanza tra le due rette.

Svolgimento. La retta richiesta si chiama retta di minima distanza tra le due rette r_1 e r_2 . Essa interseca ognuna delle due rette r_1 e r_2 , pertanto è complanare ad entrambe (questo non significa che le tre rette sono complanari altrimenti lo sarebbero pure le due rette di partenza che quindi non sarebbero sghembe). Quindi la retta cercata ha direzione ortogonale sia alla giacitura di r_1 che a quella di r_2 . Consideriamo le rette dell'esempio precedente: la giacitura di r_1 è $\langle (1, 0, 0) \rangle$ e la giacitura di r_2 è $\langle (1, 1, 1) \rangle$. La giacitura di una retta perpendicolare ad entrambe è data da $\langle (1, 0, 0), (1, 1, 1) \rangle^\perp$ e quindi è generata da un vettore non nullo (x, y, z) tale che $(x, y, z) \bullet (1, 0, 0) = 0$ e $(x, y, z) \bullet (1, 1, 1) = 0$. Otteniamo che $\langle (1, 0, 0), (1, 1, 1) \rangle^\perp = \langle (0, 1, -1) \rangle$.

Quindi la giacitura della retta r che passa per i punti di minima distanza di r_1 e r_2 è $\langle(0, 1, -1)\rangle$. Adesso siamo in grado di determinare il piano σ che contiene la retta r_1 e la retta r . Tale piano dovrà essere parallelo alle due rette dal momento che le contiene, quindi la sua giacitura T conterrà le giaciture di r_1 e r : $T \geq \langle(1, 0, 0), (0, 1, -1)\rangle$ e poiché T ha dimensione 2: $T = \langle(1, 0, 0), (0, 1, -1)\rangle$. Del resto σ contiene r_1 e quindi ogni suo punto: in particolare contiene il punto $(1, 2, 0)$. In conclusione il piano σ è: $(1, 2, 0) + \langle(1, 0, 0), (0, 1, -1)\rangle$. Una sua equazione cartesiana è $ax + by + cz + d = 0$ dove $\langle(a, b, c)\rangle = \langle(0, 1, 1)\rangle$. Imponendo che le coordinate del punto $(1, 2, 0)$ soddisfino l'equazione di σ approdiamo alla equazione $y + z - 2 = 0$. Analogamente il piano contenente la retta r_2 e la retta di minima distanza è il piano $(1, 3, 4) + \langle(1, 1, 1), (0, 1, -1)\rangle$ cioè il piano di equazione cartesiana $2x - y - z + 5 = 0$. La retta di minima distanza ha dunque equazioni cartesiane:

$$\begin{cases} 2x - y - z + 5 = 0 \\ y + z = 2. \end{cases}$$

6.4 Il prodotto vettoriale

Questo paragrafo è dedicato alla descrizione di una operazione definita sui vettori di $\mathbb{R}^3$. Tale operazione assocerà ad ogni coppia (ordinata) di vettori un terzo vettore che sarà ortogonale ai primi due. Osserviamo che, dati due vettori linearmente indipendenti (v, w) di $\mathbb{R}^3$, la direzione ortogonale ad entrambi è individuata in quanto $\dim\langle v, w \rangle^\perp = 1$. Per determinare univocamente tale vettore dire che è ortogonale a v e w non è sufficiente: dovremmo indicarne verso (che dipenderà dall'ordine della coppia (v, w)) e norma.

Cominciamo con un esercizio:

Esercizio 6.4.1 Si consideri il piano σ di equazioni parametriche

$$\begin{cases} x = x_0 + \alpha m + \beta m' \\ y = y_0 + \alpha n + \beta n' \\ z = z_0 + \alpha p + \beta p' \end{cases}$$

e si dimostri che la giacitura del piano σ ha equazione cartesiana $ax + by + cz = 0$ con $a = \det \begin{pmatrix} n & n' \\ p & p' \end{pmatrix}$, $b = -\det \begin{pmatrix} m & m' \\ p & p' \end{pmatrix}$, $c = \det \begin{pmatrix} m & m' \\ n & n' \end{pmatrix}$.

Svolgimento. La giacitura del piano σ è il sottospazio di $\mathbb{R}^3$ generato dai vettori (m, n, p) , (m', n', p') . Si tratta pertanto di dimostrare che l'insieme

delle soluzioni dell'equazione $ax + by + cz = 0$ è il sottospazio di $\mathbb{R}^3$ generato dai vettori (m, n, p) , (m', n', p') . Calcolando i determinanti che definiscono a, b e c otteniamo:

$$am + bn + cz = (np' - n'p)m - (mp' - m'p)n + (mn' - m'n)p = 0;$$

analogamente

$$am' + bn' + cp' = (np' - n'p)m' - (mp' - m'p)n' + (mn' - m'n)p' = 0.$$

In particolare il vettore $(\det \begin{pmatrix} n & n' \\ p & p' \end{pmatrix}, -\det \begin{pmatrix} m & m' \\ p & p' \end{pmatrix}, \det \begin{pmatrix} m & m' \\ n & n' \end{pmatrix})$ è ortogonale ad entrambi i vettori (m, n, p) e (m', n', p') e quindi al piano σ .

L'esercizio precedente mostra come sia possibile associare ad una coppia di vettori linearmente indipendenti di $\mathbb{R}^3$ (nell'esempio i vettori che generano la giacitura del piano σ) un terzo vettore ortogonale ad entrambi.

Definizione 6.4.2 Siano $v = (m, n, p)$, $w = (m', n', p')$ vettori di $\mathbb{R}^3$. Si dice *prodotto vettoriale* di v e w il vettore $v \times w = (\det \begin{pmatrix} n & n' \\ p & p' \end{pmatrix}, -\det \begin{pmatrix} m & m' \\ p & p' \end{pmatrix}, \det \begin{pmatrix} m & m' \\ n & n' \end{pmatrix})$.

Osservazione 6.4.3 Indicati con e_1, e_2, e_3 i vettori della base canonica di $\mathbb{R}^3$, il prodotto vettoriale tra i vettori v e w della definizione 6.4.2 viene spesso indicato come segue:

$$v \times w = \det \begin{pmatrix} e_1 & e_2 & e_3 \\ m & n & p \\ m' & n' & p' \end{pmatrix},$$

dove si intende che questo determinante viene sviluppato secondo gli elementi della prima riga.

Osservazione 6.4.4 Proprietà del prodotto vettoriale.

- (i) $v \times w$ è ortogonale sia a v che a w ;
- (ii) $\|v \times w\|$ è l'area del parallelogramma di lati v e w ;
- (iii) $v \times w = 0$ se e solo se i vettori v e w sono linearmente dipendenti;

(iv) $v \times w = -w \times v$ (cioè il prodotto vettoriale non è una operazione commutativa);

(v) l'applicazione

$$\begin{aligned} \times : \mathbb{R}^3 \times \mathbb{R}^3 &\longrightarrow \mathbb{R}^3 \\ (v, w) &\longmapsto v \times w \end{aligned}$$

è bilineare cioè $v \times (w + z) = v \times w + v \times z$, $(w + z) \times v = w \times v + z \times v$,
 $v \times (\alpha w) = \alpha v \times w = (\alpha v) \times w$.

Abbiamo già dimostrato (Esercizio 6.2.4) la proprietà (i). Vediamo le altre:

(ii) Siano $v = (m, n, p)$, $w = (m', n', p')$ e indichiamo con α l'angolo formato dai vettori v e w (Lezione 12, Osservazione ??). Allora l'area del parallelogramma di lati v e w è pari a $\|v\|\|w\|\sin \alpha$. Ora calcoliamo la norma del vettore $v \times w$:

$$\begin{aligned} \|v \times w\|^2 &= (np' - pn')^2 + (mp' - pm')^2 + (mn' - nm')^2 = \\ &= (m^2 + n^2 + p^2)(m'^2 + n'^2 + p'^2) - (mm' + nn' + pp')^2 = \|v\|^2\|w\|^2 - (v \bullet w)^2 = \\ &= \|v\|^2\|w\|^2(1 - \cos^2 \alpha), \end{aligned}$$

e questo prova (ii).

(iii) Abbiamo visto in (ii) che $\|v \times w\| = \|v\|\|w\|\sin \alpha$, essendo α l'angolo compreso tra i vettori v e w . Di conseguenza $\|v \times w\| = 0$ se e solo se $\sin \alpha = 0$ cioè se e solo se i vettori v e w sono allineati. Poiché un vettore è nullo se e solo se è nulla la sua norma, questo conclude l'osservazione (iii).

Infine le proprietà (iv) e (v) sono conseguenze immediate della definizione di prodotto vettoriale.

Esempio 6.4.5 Calcoliamo il prodotto vettoriale dei vettori e_1 ed e_2 della base canonica di $\mathbb{R}^3$:

$$e_1 \times e_2 = \det \begin{pmatrix} e_1 & e_2 & e_3 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} = e_3.$$

Analogamente, $e_2 \times e_3 = e_1$, $e_1 \times e_3 = -e_2$. In particolare osserviamo che il prodotto vettoriale non è associativo. Si ha infatti, ad esempio:

$$(e_1 \times e_2) \times e_2 = e_3 \times e_2 = -e_1$$

ma

$$e_1 \times (e_2 \times e_2) = e_1 \times 0_{\mathbf{R}^3} = 0_{\mathbf{R}^3}.$$

Osservazione 6.4.6 Cerchiamo ora di evidenziare alcune proprietà del prodotto vettoriale.

- (i) Dalla definizione 6.4.2 segue subito che se v_1, v_2, v_3 sono vettori di $\mathbb{R}^3$, vale:

$$\det(v_1 \ v_2 \ v_3) = (v_1 \times v_2) \bullet v_3$$

dove con la scrittura $(v_1 \ v_2 \ v_3)$ intendiamo la matrice che ha sulle colonne le componenti dei vettori v_1, v_2, v_3 , nell'ordine. Infatti siano $v_1 = (m, n, p)$, $v_2 = (m', n', p')$, $v_3 = (a, b, c)$. Allora:

$$\begin{aligned} \det(v_1 \ v_2 \ v_3) &= \det \begin{pmatrix} m & m' & a \\ n & n' & b \\ p & p' & c \end{pmatrix} = \\ &= a \det \begin{pmatrix} n & n' \\ p & p' \end{pmatrix} - b \det \begin{pmatrix} m & m' \\ p & p' \end{pmatrix} + c \det \begin{pmatrix} m & m' \\ n & n' \end{pmatrix} = \\ &= v_3 \bullet (v_1 \times v_2) = (v_1 \times v_2) \bullet v_3. \end{aligned}$$

- (ii) Consideriamo il parallelepipedo P avente come spigoli i vettori v_1, v_2, v_3 e sia β l'angolo formato dai vettori $v_1 \times v_2$ e v_3 . Allora

$$(v_1 \times v_2) \bullet v_3 = \|v_1 \times v_2\| \|v_3\| \cos \beta.$$

Del resto, per l'Osservazione 6.4.4.(ii), $\|v_1 \times v_2\|$ è l'area del parallelogramma di lati v_1, v_2 , cioè l'area di una base b di P, e $h = \|v_3\| \cos \beta$ è la lunghezza dell'altezza di P relativa alla base b . Di conseguenza il prodotto $(v_1 \times v_2) \bullet v_3$ fornisce il volume (orientato) di P. Il prodotto $(v_1 \times v_2) \bullet v_3$ si chiama *prodotto misto* dei vettori v_1, v_2, v_3 .

Osservazione 6.4.7 Abbiamo studiato alcune proprietà del prodotto vettoriale. In particolare abbiamo visto che se partiamo da due vettori v_1, v_2 che sono ortogonali fra loro e di norma pari ad 1, allora il loro prodotto vettoriale $v_3 = v_1 \times v_2$ è ortogonale ad entrambi e la sua norma è l'area del parallelogramma (che in questo caso è un quadrato) di lato 1, dunque $\|v_3\| = 1$. Possiamo allora scrivere la matrice che ha sulle colonne rispettivamente le coordinate di v_1, v_2, v_3 : essendo i tre vettori ortogonali fra loro e di norma eguale ad uno la matrice è una matrice ortogonale e il suo determinante è pari ad 1. In effetti basta pensare che nella formula di (i) dell'osservazione 6.4.6 (mantenendo le stesse notazioni)

$$v_3 = (a, b, c) = \left(\det \begin{pmatrix} n & n' \\ p & p' \end{pmatrix}, -\det \begin{pmatrix} m & m' \\ p & p' \end{pmatrix}, \det \begin{pmatrix} m & m' \\ n & n' \end{pmatrix} \right).$$

Dal momento che il determinante di una matrice ortogonale è uguale ad 1 o a -1 (cf. ??) e dal momento che $\det(v_1 \ v_2 \ v_3) = \|v_1 \times v_2\|^2$ è un numero reale non negativo, ne deduciamo che $\det(v_1 \ v_2 \ v_3) = 1$.

In conclusione il vettore $v_3 = v_1 \times v_2$ è ortogonale a v_1 e a v_2 e il suo verso rende $\det(v_1 \ v_2 \ v_3)$ uguale a 1. In generale una tecnica per determinare il verso del prodotto vettoriale dei vettori v_1 e v_2 (nell'ordine) è quella di prendere v_1 e v_2 come il pollice e l'indice della mano destra e il verso del vettore v_3 sarà indicato dal dito medio.

6.5 Esercizi proposti

Esercizio 6.5.1 Date le rette $r : \begin{cases} x - y - 2 = 0 \\ z = 0 \end{cases}$ ed $s : \begin{cases} y + 2z = -1 \\ x + 1 = 0 \end{cases}$:

1. determinare la posizione reciproca delle rette r ed s ;
2. calcolare la loro distanza;
3. determinare la retta di minima distanza tra r ed s ;
4. determinare, se possibile, una retta incidente sia r che s e passante per il punto $O = (0, 0, 0)$.

Esercizio 6.5.2 Dato il piano π di equazione $2x + y + 2z = -1$, si determini, se possibile, una retta r parallela al piano π che abbia distanza da esso uguale a 3, passante per il punto $P = (2, 0, 2)$ ed ortogonale all'asse y .

Esercizio 6.5.3 Si considerino le rette

$$r : \begin{cases} 2x - y + 3 = 0 \\ x + 2y - 5z = 1 \end{cases} \quad \text{ed} \quad s : \begin{cases} x = 1 + 2t \\ y = -1 + t \\ z = -1 + t \end{cases}$$

Determinare i punti della retta s aventi distanza 1 da r .

Esercizio 6.5.4 Determinare, se possibile, un piano ortogonale al piano π di equazione $x - 2y - z = 1$ e parallelo al vettore $v = (0, 1, -1)$. Un siffatto piano è unico?

Lezione 7

Applicazioni lineari e matrici

7.1 Applicazioni lineari

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

Definizione 7.1.1 Una applicazione $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$ si dice lineare se

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

Questa è la definizione di applicazione lineare. Si noti che dalla definizione segue subito che, per qualunque vettore v di $\mathbb{R}^n$, $L(\mathbf{0}_{\mathbb{R}^n}) = L(0v) = 0L(v) = \mathbf{0}_{\mathbb{R}^m}$. 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 $\mathbb{R}^m$ del vettore $L(v)$.

Osservazione 7.1.2 Sia A una matrice $m \times n$. Allora possiamo definire una funzione lineare

$$f_A : \mathbb{R}^n \rightarrow \mathbb{R}^m$$

associando al vettore $v = (x_1, \dots, x_n)$ il vettore $w = (y_1, \dots, y_m)$ ottenuto mediante il prodotto righe per colonne di A per v :

$$A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}.$$

Chiameremo f_A la funzione lineare associata alla matrice A . Si possono verificare direttamente le proprietà della Definizione 7.1.1.

Definizione 7.1.3 Si chiama *nucleo* di un'applicazione lineare $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, e si indica con $\ker(f)$, il seguente sottoinsieme di $\mathbb{R}^n$:

$$\ker(f) = \{v \in \mathbb{R}^n \mid f(v) = 0_{\mathbb{R}^m}\}.$$

Osservazione 7.1.4 1) Il nucleo di una funzione lineare f è un sottospazio vettoriale del dominio di f .

2) Il nucleo di una funzione lineare f_A associata ad una matrice A è l'insieme delle soluzioni del sistema lineare omogeneo avente A come matrice incompleta.

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

Proposizione 7.1.5 Data una applicazione lineare $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$, L è iniettiva se e solo se $\text{Ker}L$ è il sottospazio banale di $\mathbb{R}^n$.

Definizione 7.1.6 Si chiama *immagine* di un'applicazione lineare $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, e si indica con $\text{Im}(f)$, il seguente sottoinsieme di $\mathbb{R}^m$:

$$\text{Im}(f) = \{w \in \mathbb{R}^m \mid w = f(v), v \in \mathbb{R}^n\}.$$

Osservazione 7.1.7 1) L'immagine di una funzione lineare f è un sottospazio vettoriale del codominio di f .

2) L'immagine di una funzione lineare f_A associata ad una matrice A è generata dai vettori colonna della matrice A . In particolare, il numero di colonne linearmente indipendenti di A fornisce la dimensione dell'immagine di A . Poiché il numero di colonne linearmente indipendenti di una matrice (rango colonne) coincide con il numero di righe linearmente indipendenti della stessa matrice, cioè con il rango di A , si ha $\text{rg}(A) = \dim(\text{Im}f)$.

Ricordiamo che una funzione $f : X \rightarrow Y$ è suriettiva se l'immagine di f coincide con Y . Questo naturalmente resta vero per le funzioni lineari $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$. In questo caso $\text{Im}f$ è un sottospazio vettoriale di $\mathbb{R}^m$ perciò f è suriettiva se e solo se $\dim(\text{Im}f) = m$.

Definizione 7.1.8 Sia $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ una funzione lineare e sia $w \in \mathbb{R}^m$. Si chiama controimmagine di w mediante f e si indica con $f^{-1}(w)$ il seguente sottoinsieme di $\mathbb{R}^n$:

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

Esercizio 7.1.9 Data un'applicazione lineare $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$ e dato un vettore $w \in \mathbb{R}^m$, in che modo possiamo descrivere $L^{-1}(\{w\})$, cioè l'insieme di tutti i vettori di $\mathbb{R}^n$ la cui immagine mediante L è w ?

Soluzione Se $w = \mathbf{0}_{\mathbb{R}^m}$ allora $L^{-1}(\{w\}) = \text{Ker}L$.

Se $w \neq \mathbf{0}_{\mathbb{R}^m}$? Se $w \notin \text{Im}L$ la risposta è facile: $L^{-1}(\{w\}) = \emptyset$.

Se, invece, $w \in \text{Im}L$ esiste sicuramente un elemento $v \in \mathbb{R}^n$ tale che $L(v) = w$. Si ha allora:

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

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

Osservazione 7.1.10 i) Nella costruzione di $L^{-1}(\{w\}) = v + \text{Ker}L$ abbiamo scelto un vettore v tale che $L(v) = w$. Se ne avessimo scelto un altro? Il risultato sarebbe stato lo stesso. Infatti scelto $\bar{v} \neq v$ tale che $L(\bar{v}) = w$, avremmo trovato $L^{-1}(\{w\}) = \bar{v} + \text{Ker}L$, quindi $\bar{v} + \text{Ker}L = v + \text{Ker}L$.

- ii) Che struttura ha $v + \text{Ker}L$? Non possiamo aspettarci molto. Se $v \notin \text{Ker}L$ allora nessuna delle somme $v + k$, $k \in \text{Ker}L$, può essere il vettore nullo, quindi $v + \text{Ker}L$ non è uno spazio vettoriale (se fosse $v + k = \mathbf{0}_V$ si avrebbe $v = -k \in \text{Ker}L$, assurdo). L'altra possibilità è che $v \in \text{Ker}L$, ma allora $v + \text{Ker}L = \text{Ker}L$ è un sottospazio. (Per mostrare l'uguaglianza basta osservare che ogni elemento z di $\text{Ker}L$ si può scrivere come $z = v + (z - v)$ ove $-v \in \text{Ker}L$ per ipotesi e dunque $z - v \in \text{Ker}L$).
- iii) Data una applicazione lineare $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$ l'antiimmagine tramite L di un vettore $w \in \text{Im}L$ è data da un vettore particolare $v \in \mathbb{R}^n$ che soddisfi la condizione $L(v) = w$ e da tutti i vettori che si ottengono sommando v agli elementi di $\text{Ker}L$. Cosicché se L è iniettiva allora $L^{-1}(\{w\})$ è costituito da un solo vettore.
- iv) Se $f_A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ è la funzione lineare associata alla matrice $m \times n$ A , la controimmagine di un vettore $w = (y_1, \dots, y_m)$ mediante f_A è l'insieme delle soluzioni del sistema lineare avente matrice incompleta A e colonna dei termini noti $(y_1, \dots, y_m)$ (scritto in colonna).

7.2 Matrici

Abbiamo visto come le matrici $\mathcal{M}_{m,n}(\mathbb{R})$ rappresentino l'insieme delle applicazioni lineari da $\mathbb{R}^n$ in $\mathbb{R}^m$.

Se componiamo due applicazioni lineari, la funzione composta è ancora lineare? E, nel caso lo sia, che cosa possiamo 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.

7.3 Prodotto righe per colonne

Problema. Consideriamo gli spazi vettoriali $\mathbb{R}^n$, $\mathbb{R}^m$ e $\mathbb{R}^p$. Siano poi $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ e $g : \mathbb{R}^m \rightarrow \mathbb{R}^p$ funzioni lineari associate, rispettivamente, alle matrici $A_f \in \mathcal{M}_{m,n}(\mathbb{R})$ e $A_g \in \mathcal{M}_{p,m}(\mathbb{R})$.

1. La composizione $g \circ f : \mathbb{R}^n \rightarrow \mathbb{R}^p$ 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 $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ e $g: \mathbb{R}^m \rightarrow \mathbb{R}^p$ due applicazioni lineari e sia $g \circ f: \mathbb{R}^n \rightarrow \mathbb{R}^p$ la loro composizione. Per ogni $v \in \mathbb{R}^n$ si ha $(g \circ f)(v) = g(f(v))$ con $f(v) \in \mathbb{R}^m$. Ora, se prendiamo $v_1, v_2 \in \mathbb{R}^n$,

$$\begin{aligned} (g \circ f)(v_1 + v_2) &= g(f(v_1 + v_2)) = g(f(v_1) + f(v_2)) = \\ &= g(f(v_1)) + g(f(v_2)) = (g \circ f)(v_1) + (g \circ f)(v_2) \end{aligned}$$

dove abbiamo usato la linearità di f e g . Sempre per linearità, se $\lambda \in \mathbb{R}$ e $v \in \mathbb{R}^n$ 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? Se

$$A_f = (a_{ij})_{i=1, \dots, m}^{j=1, \dots, n} \quad A_g = (b_{ki})_{k=1, \dots, p}^{i=1, \dots, m},$$

allora

$$A_{g \circ f} = (c_{kj})_{k=1, \dots, p}^{j=1, \dots, n}$$

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 7.3.1 Consideriamo le due matrici

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

Calcoliamo il prodotto AB . La matrice AB rappresenta, secondo l'interpretazione precedente, la composizione di una applicazione da $\mathbb{R}^2$ in $\mathbb{R}^2$ (nelle

notazioni precedenti $B = A_f$) e di una applicazione da $\mathbb{R}^3$ a $\mathbb{R}^2$ ($A = A_g$). 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 $\mathbb{R}^3$ a $\mathbb{R}^2$ con una funzione lineare da $\mathbb{R}^2$ a $\mathbb{R}^2$ (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 7.3.2 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 conoscevamo già il prodotto righe per colonne di due matrici A e B 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 7.3.3 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

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

$$\text{Allora } AB = \begin{pmatrix} 3 & 3 \\ 3 & 3 \end{pmatrix} \text{ e } BA = \begin{pmatrix} 1 & 5 \\ 1 & 5 \end{pmatrix}.$$

Osserviamo che, in modo equivalente, la composizione di applicazioni (lineari) non è un'operazione commutativa. Ad esempio, se consideriamo le

applicazioni lineari $f, g : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ associate alle matrici A e B , naturalmente $f \circ g \neq g \circ f$.

Osservazione 7.3.4 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.

7.4 Matrici invertibili

Proposizione 7.4.1 Se $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ è una funzione lineare invertibile, l'applicazione inversa di f è anch'essa un'applicazione lineare che indicheremo con f^{-1} .

Indichiamo con I_n la cosiddetta matrice identica di ordine n cioè la matrice quadrata di ordine n con elementi diagonali tutti uguali ad 1 ed elementi nulli al di fuori della diagonale principale. Ad esempio:

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Osservazione 7.4.2 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 7.4.3 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 7.4.4 Sia $A \in \mathcal{M}_n(\mathbb{R})$ e sia $f_A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ la funzione lineare associata alla matrice A . Se f_A è invertibile, allora f_A^{-1} è la funzione lineare associata alla matrice A^{-1} .

Quali sono le proprietà delle matrici invertibili? Possiamo fare alcune brevi osservazioni sfruttando l'interpretazione di una matrice quadrata come applicazione lineare o, più precisamente, come endomorfismo. Sia $A \in \mathcal{M}_n(\mathbb{R})$ una matrice invertibile. Allora essa descrive una funzione lineare invertibile da $\mathbb{R}^n$ a $\mathbb{R}^n$. In particolare il rango di A è n e il numero di pivots in una qualunque forma a scalini per righe di A è n . In questa situazione non vi sono molte possibilità dal momento che la matrice ha ordine n : la forma a scalini deve avere tutti i pivots sulla diagonale e ogni elemento della diagonale deve essere un pivot. Questo fa sì che si possa ulteriormente raffinare la forma a scalini. Infatti, dopo opportune operazioni elementari sulle righe, la matrice sarà del tipo

$$\begin{pmatrix} \alpha_1 & * & * & * & \dots & \beta_1 \\ 0 & \alpha_2 & * & * & \dots & \beta_2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \alpha_{n-1} & \beta_{n-1} \\ 0 & 0 & 0 & 0 & 0 & \alpha_n \end{pmatrix}$$

dove $\alpha_i \neq 0$, e dove $*$ e β_j sono numeri reali. A questo punto possiamo continuare a semplificare la matrice con altri cambiamenti di base nel codominio. Ad esempio se sostituiamo la prima riga con la somma della prima riga con l'ultima moltiplicata per $-\frac{\beta_1}{\alpha_n}$ (cosa che può essere fatta poiché $\alpha_n \neq 0$) otterremo una matrice in cui l'elemento di posto $1, n$ (1 è la riga e n la colonna) è uguale a zero. Possiamo ripetere l'operazione per la seconda riga sommandole l'ultima moltiplicata per $-\frac{\beta_2}{\alpha_n}$. In questo modo si trova una matrice in cui gli elementi di posto k, n (k è la riga e n la colonna) con $k \neq n$ sono nulli e quello di posto n, n è α_n . Procedendo in questo modo con le altre colonne si arriva ad una matrice diagonale cioè ad una matrice del tipo:

$$\begin{pmatrix} \alpha_1 & 0 & 0 & 0 & \dots & 0 \\ 0 & \alpha_2 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & \alpha_n \end{pmatrix}.$$

Possiamo fare ancora di meglio: se moltiplichiamo la prima riga per $\frac{1}{\alpha_1}$, la seconda per $\frac{1}{\alpha_2}$ e così via, alla fine otteniamo la matrice identica I_n .

Proposizione 7.4.5 *Una matrice $A \in \mathcal{M}_n(\mathbb{R})$ è invertibile se e solo se mediante operazioni elementari sulle righe può essere ridotta alla matrice identica I_n .*

Dimostrazione. “ $\Rightarrow$ ” Questa implicazione è stata appena vista: se una matrice è invertibile allora esiste una forma a scalini per righe uguale alla matrice identica. “ $\Leftarrow$ ” Supponiamo che una matrice A abbia una forma a scalini per righe uguale alla matrice identica I_n . La riduzione in forma a scalini per righe è prodotta tramite operazioni elementari che non alterano il rango della matrice che quindi è uguale a n . Pertanto l'applicazione associata è invertibile e quindi la matrice è invertibile. **C.V.D.**

Osservazione 7.4.6 Se una matrice $A \in \mathcal{M}_n(\mathbb{R})$ ammette inversa, i.e., se esiste una matrice $B \in \mathcal{M}_n(\mathbb{R})$ tale che $AB = BA = I_n$, allora B è unica. Supponiamo infatti che esista un'altra matrice $C \in \mathcal{M}_n(\mathbb{R})$ tale che $CA = AC = I_n$. Consideriamo allora la matrice CAB . Essa è una matrice quadrata di ordine n e, per la proprietà associativa del prodotto righe per colonne, abbiamo $CAB = C(AB) = (CA)B$. Del resto $AB = CA = I_n$. Dunque abbiamo $CAB = CI_n = I_nB$, cioè $C = B$. Dunque l'inversa di una matrice è unica.

Per riconoscere se una matrice sia invertibile o meno non occorre ridurla alla matrice identica (bastava già una forma a scalini in cui tutti i pivots fossero gli elementi della diagonale). Tuttavia il lavoro fatto non è superfluo: ci fornisce senza ulteriore fatica un metodo per calcolare l'inversa di una matrice (invertibile)! Perché? Quello che abbiamo fatto è stato trovare un modo per ridurre la matrice di partenza A alla matrice identica. In termini molto grossolani è come se avessimo moltiplicato A^{-1} per A . Ma allora se applicassimo le stesse trasformazioni alla matrice identica I_n otterremmo il prodotto di A^{-1} per I_n cioè: $A^{-1}I_n = A^{-1}$. Quindi otterremmo la matrice inversa A^{-1} ! Illustriamo quanto appena detto attraverso un esempio.

Esempio 7.4.7 Prendiamo la matrice

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

cerchiamo di vedere se è invertibile ed eventualmente calcoliamone l'inversa. Procediamo come abbiamo detto prima e consideriamo la matrice che si ottiene formalmente scrivendo le colonne della matrice I_3 a destra di A in un'unica nuova matrice:

$$\left(\begin{array}{ccc|ccc} 2 & 3 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ -1 & 0 & 1 & 0 & 0 & 1 \end{array} \right).$$

Ora applichiamo a tutta la matrice esattamente le stesse trasformazioni per riga che useremmo per ottenere una forma a scalini per righe, diagonale, della matrice A . Cominciamo: scambiamo l'ultima riga di A con la prima:

$$\left(\begin{array}{ccc|ccc} -1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 2 & 3 & 1 & 1 & 0 & 0 \end{array} \right); \quad (7.2)$$

sostituiamo alla terza riga la somma della terza riga con la prima moltiplicata per 2:

$$\left(\begin{array}{ccc|ccc} -1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 3 & 3 & 1 & 0 & 2 \end{array} \right);$$

sostituiamo alla terza riga la somma della terza riga con la seconda moltiplicata per -3 :

$$\left(\begin{array}{ccc|ccc} -1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 3 & 1 & -3 & 2 \end{array} \right);$$

sostituiamo alla prima riga la somma della prima con l'ultima moltiplicata per $-\frac{1}{3}$:

$$\left(\begin{array}{ccc|ccc} -1 & 0 & 0 & -\frac{1}{3} & 1 & \frac{1}{3} \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 3 & 1 & -3 & 2 \end{array} \right);$$

moltiplichiamo la prima riga per -1 e l'ultima per $\frac{1}{3}$:

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{3} & -1 & -\frac{1}{3} \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & \frac{1}{3} & -1 & \frac{2}{3} \end{array} \right).$$

La matrice

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

è dunque la matrice inversa della matrice di partenza. (N.B. Consideriamo 7.2: se moltiplichiamo la matrice A per la matrice che è formata dalle ultime 3 colonne di 7.2, abbiamo

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

cioè la matrice data dalle prime tre colonne di 7.2. Questo significa che la matrice data dalle ultime tre colonne di 7.2 è la matrice che moltiplicata per A dà la matrice data dalle prime tre colonne di 7.2. La stessa considerazione può essere fatta per le successive trasformazioni per riga: di fatto tutte le trasformazioni di righe e colonne possono essere realizzate tramite moltiplicazione per opportune matrici invertibili). Possiamo verificare direttamente, usando il prodotto righe per colonne, che la matrice trovata è effettivamente la matrice inversa della matrice di partenza:

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

Esempio 7.4.8 Consideriamo il seguente sistema nelle incognite x_1, x_2, x_3 :

$$A\mathbf{x} = \begin{pmatrix} 2 & 3 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix}.$$

La matrice $A \in \mathcal{M}_3(\mathbb{R})$ è quella dell'esempio precedente e quindi è invertibile. Il sistema avrà dunque un'unica soluzione data dall'unico vettore di $\mathbb{R}^3$ che è spedito, dall'applicazione rappresentata da A , nel vettore $(1, -2, 0)$, o, equivalentemente, dall'unico vettore immagine di $(1, -2, 0)$ mediante l'applicazione inversa. Poiché conosciamo la matrice dell'applicazione inversa siamo in grado di risolvere il sistema:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} & -1 & -\frac{1}{3} \\ 0 & 1 & 0 \\ \frac{1}{3} & -1 & \frac{2}{3} \end{pmatrix} \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{7}{3} \\ -2 \\ \frac{7}{3} \end{pmatrix}.$$

7.5 Esercizi svolti

Esercizio 7.5.1 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) $f_3 : \mathbb{R}^2 \longrightarrow \mathbb{R}, f_3(x, y) = 2x + 3y.$

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:

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

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.

Esercizio 7.5.2 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 ad f .

- ii) Determinare $\text{Ker } f$ e $\text{Im } f$.

Svolgimento.

- i) La matrice associata a f è la matrice che ha sulle colonne le immagini mediante f dei vettori della base canonica di $\mathbb{R}^3$. 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 è:

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

- ii) Sappiamo che $\text{Im } f$ è generata dai vettori $f(1, 0, 0)$, $f(0, 1, 0)$ e $f(0, 0, 1)$.
Dunque $\text{Im } f = \langle (1, 1, 0), (0, 0, 1) \rangle$.

Abbiamo inoltre $\text{Ker}(f) = \langle (1, -1, 0) \rangle$.

Esercizio 7.5.3 Tra le applicazioni lineari dell'esercizio 7.5.1 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 .

Esercizio 7.5.4 Calcolare i prodotti AB e BA delle matrici

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

Indicata con $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ la funzione associata alla matrice AB , determinare $f(2, -3)$.

Svolgimento. Per calcolare AB e BA dobbiamo semplicemente usare la definizione di prodotto righe per colonne:

$$AB = \begin{pmatrix} 5 & 23 \\ 7 & 11 \end{pmatrix}; \quad BA = \begin{pmatrix} 0 & 1 & -3 & 2 \\ -2 & -3 & -9 & -6 \\ -1 & -2 & -3 & -4 \\ 3 & 11 & -6 & 22 \end{pmatrix}.$$

Dunque, AB è una matrice quadrata di ordine 2 e quindi la matrice associata ad una funzione lineare da $\mathbb{R}^2$ in $\mathbb{R}^2$. BA è una matrice quadrata di ordine 4 ed è pertanto la matrice associata ad una funzione lineare da $\mathbb{R}^4$ in $\mathbb{R}^4$.

Per calcolare l'immagine del vettore $(2, -3)$ mediante f basta calcolare il prodotto righe per colonne della matrice AB per il vettore $(2, -3)$ scritto in colonna.

Esercizio 7.5.5 Determinare la matrice inversa di ciascuna delle seguenti matrici:

$$M_1 = \begin{pmatrix} 1 & 1 \\ -2 & 1 \end{pmatrix}, \quad M_2 = \begin{pmatrix} -1 & 0 & 1 \\ 0 & -1 & 4 \\ 0 & 0 & 2 \end{pmatrix}, \quad M_3 = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -3 \end{pmatrix}$$

Svolgimento. Procediamo come illustrato nell'esempio 7.4.7. Cominciamo dalla matrice M_1 e consideriamo la matrice $\left(\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ -2 & 1 & 0 & 1 \end{array} \right)$.

Sostituiamo alla seconda riga la somma della prima riga moltiplicata per due e della seconda: $\left(\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 0 & 3 & 2 & 1 \end{array} \right)$.

Ora sostituiamo alla prima riga la somma della prima riga moltiplicata per -3 e della seconda: $\left(\begin{array}{cc|cc} -3 & 0 & -1 & 1 \\ 0 & 3 & 2 & 1 \end{array} \right)$.

Infine, moltiplichiamo la prima riga per $-\frac{1}{3}$ e la seconda per $\frac{1}{3}$:

$$\left(\begin{array}{cc|cc} 1 & 0 & \frac{1}{3} & -\frac{1}{3} \\ 0 & 1 & \frac{2}{3} & \frac{1}{3} \end{array} \right).$$

Abbiamo così ottenuto la matrice inversa della matrice M_1 :

$$M_1^{-1} = \begin{pmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{pmatrix}.$$

Analogamente si procede per la matrice M_2 :

$$\left(\begin{array}{ccc|ccc} -1 & 0 & 1 & 1 & 0 & 0 \\ 0 & -1 & 4 & 0 & 1 & 0 \\ 0 & 0 & 2 & 0 & 0 & 1 \end{array} \right) \Rightarrow \left(\begin{array}{ccc|ccc} 2 & 0 & 0 & -2 & 0 & 1 \\ 0 & -1 & 0 & 0 & 1 & -2 \\ 0 & 0 & 2 & 0 & 0 & 1 \end{array} \right) \Rightarrow$$

$$\Rightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 0 & \frac{1}{2} \\ 0 & 1 & 0 & 0 & -1 & 2 \\ 0 & 0 & 1 & 0 & 0 & \frac{1}{2} \end{array} \right).$$

Dunque $M_2^{-1} = \begin{pmatrix} -1 & 0 & \frac{1}{2} \\ 0 & -1 & 2 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}$. Osserviamo che M_2 è una matrice

triangolare superiore e che la sua inversa è anch'essa una matrice triangolare superiore. Questo fatto è vero in generale.

Anche nel caso della matrice M_3 si procede come per qualsiasi altra matrice, ma, essendo essa una matrice diagonale, il numero di passaggi sarà inferiore: la forma diagonale è infatti già di per sé una forma a scalini per righe. Otteniamo dunque, immediatamente,

$$M_3^{-1} = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{3} \end{pmatrix}.$$

In generale, data una matrice diagonale di ordine n con elementi diagonali non nulli $\alpha_1, \alpha_2, \dots, \alpha_n$ (se uno di questi fosse nullo la matrice non sarebbe invertibile!), la sua inversa è la matrice diagonale con elementi diagonali $\alpha_1^{-1}, \alpha_2^{-1}, \dots, \alpha_n^{-1}$.

Esercizio 7.5.6 Sia $h : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ l'applicazione lineare associata alla matrice:

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

Stabilire se h è invertibile. e determinare $h^{-1}(2, 1, 2)$.

Svolgimento. Si verifica facilmente che il rango della matrice H è uguale a 3, pertanto la funzione h è invertibile. A questo punto un modo di risolvere l'esercizio è calcolare la matrice inversa della matrice H . Infatti se h è un endomorfismo invertibile di $\mathbb{R}^n$ e H è la matrice ad esso associata, allora la matrice associata all'endomorfismo inverso h^{-1} è la matrice H^{-1} .

Procediamo col calcolo della matrice H^{-1} :

$$\begin{aligned}
& \left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ -1 & 1 & -2 & 0 & 1 & 0 \\ 0 & -1 & -3 & 0 & 0 & 1 \end{array} \right) \Rightarrow \left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 3 & -2 & 1 & 1 & 0 \\ 0 & -1 & -3 & 0 & 0 & 1 \end{array} \right) \Rightarrow \\
& \left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 3 & -2 & 1 & 1 & 0 \\ 0 & 0 & -11 & 1 & 1 & 3 \end{array} \right) \Rightarrow \left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & -33 & 0 & -9 & -9 & 6 \\ 0 & 0 & -11 & 1 & 1 & 3 \end{array} \right) \Rightarrow \\
& \left(\begin{array}{ccc|ccc} \frac{33}{2} & 0 & 0 & \frac{15}{2} & -9 & 6 \\ 0 & -33 & 0 & -9 & -9 & 6 \\ 0 & 0 & -11 & 1 & 1 & 3 \end{array} \right) \Rightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{5}{11} & -\frac{6}{11} & \frac{4}{11} \\ 0 & 1 & 0 & \frac{3}{11} & \frac{3}{11} & -\frac{2}{11} \\ 0 & 0 & 1 & -\frac{1}{11} & -\frac{1}{11} & -\frac{3}{11} \end{array} \right).
\end{aligned}$$

Abbiamo quindi

$$H^{-1} = \begin{pmatrix} \frac{5}{11} & -\frac{6}{11} & \frac{4}{11} \\ \frac{3}{11} & \frac{3}{11} & -\frac{2}{11} \\ -\frac{1}{11} & -\frac{1}{11} & -\frac{3}{11} \end{pmatrix}.$$

$$\text{Pertanto } h^{-1}(2, 1, 2) = \begin{pmatrix} \frac{5}{11} & -\frac{6}{11} & \frac{4}{11} \\ \frac{3}{11} & \frac{3}{11} & -\frac{2}{11} \\ -\frac{1}{11} & -\frac{1}{11} & -\frac{3}{11} \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} \frac{12}{11} \\ \frac{5}{11} \\ -\frac{9}{11} \end{pmatrix}.$$

7.6 Esercizi proposti

Esercizio 7.6.1 Stabilire quali delle seguenti applicazioni sono lineari:

1. $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $f(x, y) = (x - y + 2, x + y)$;
2. $g : \mathbb{R}^3 \rightarrow \mathbb{R}^2$, $g(x, y, z) = (x^2 - y^2, x + y)$.

Esercizio 7.6.2 Si consideri, al variare di $k \in \mathbb{R}$, l'applicazione $f_k : \mathbb{R}^2 \rightarrow \mathbb{R}^3$, $f_k(x, y) = (x - y, ky, ky)$. Stabilire per quali valori di k l'applicazione f_k è lineare. Per i valori di k trovati:

1. Determinare nucleo e immagine di f_k .
2. Determinare i valori di k tali che il vettore $(1, 0, 0)$ appartenga all'immagine di f_k .

Esercizio 7.6.3 Calcolare l'inversa della matrice

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

Esercizio 7.6.4 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 $AC = I_3$.