Algebra lineare, anno 2012/2013

27 dicembre 2012

Indice

Sommario

Queste note rappresentano un sunto delle lezioni svolte in classe. Degli esercizi svolti in classe si riporta il testo, ma non sempre lo svolgimento nei dettagli.

Un punto di riferimento più completo sono le dispense del modulo di Algebra (corso di laurea in Scienze statistiche, Bologna) della Prof. Guidotti, reperibili al link `http://campus.unibo.it/87355/`.

1. Lo spazio $\mathbb{R}^n$

1.1. Operazioni tra vettori e combinazioni lineari

Definiamo per $n = 1, 2, \dots,$ lo spazio $\mathbb{R}^n$ come segue: (i)

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

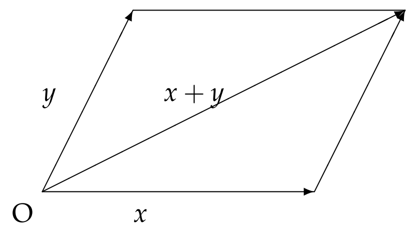
Gli elementi di $\mathbb{R}^n$ si chiamano *punti* o *vettori*. Scritto $x = (x_1, \dots, x_n)$, i numeri $x_1, \dots, x_n$ si chiamano le *componenti* del vettore x.

Introduciamo le seguenti operazioni tra vettori:

(A) Somma tra vettori: poniamo, dati $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n) \in \mathbb{R}^n$

$$x + y := (x_1 + y_1, \dots, x_n + y_n). \tag{1.1}$$

La sommma tra vettori può anche essere interpretata tramite la cosiddetta "regola del parallelogramma".



Osserviamo le seguenti proprietà della somma: se indichiamo con $0 = 0_n = (0, \dots, 0) \in \mathbb{R}^n$ il *vettore nullo*, allora risulta

$$x + 0 = x \quad \text{per ogni } x \in \mathbb{R}^n.$$

Infine, come per i numeri, valgono le proprietà commutativa e associativa:

$$\begin{aligned} &v + w = w + v \\ &u + (v + w) = (u + v) + w \end{aligned}$$

per ogni $u, v, w \in \mathbb{R}^n$.

(B) Prodotto di un vettore con uno scalare: dati $u = (u_1, \dots, u_n)$ e $\lambda \in \mathbb{R}$, poniamo

$$\lambda u := (\lambda u_1, \dots, \lambda u_n). \tag{1.2}$$

Valgono le proprietà immediate

$$\begin{gathered} (\alpha + \beta)v = \alpha v + \beta v \\ \alpha(v + w) = \alpha v + \alpha w \\ \alpha(\beta v) = (\alpha\beta)v \\ 1v = v \end{gathered}$$

per ogni $v, w \in \mathbb{R}^n$ e $\alpha, \beta \in \mathbb{R}$.

Osserviamo che il vettore *opposto* di v, che è $-v := (-v_1, \dots, -v_n)$, soddisfa

$$-v + v = v + (-v) = 0.$$

Definizione 1.1 (combinazione lineare). *Dati dei vettori* $v_1, \dots, v_p \in \mathbb{R}^n$*, una combinazione lineare dei detti vettori è un vettore della forma*

$$v = \alpha_1 v_1 + \cdots + \alpha_p v_p = \sum_{k=1}^{p} \alpha_k v_k,$$

dove $\alpha_1, \dots, \alpha_p \in \mathbb{R}$.

I numeri reali $\alpha_1, \dots, \alpha_p$ si chiamano *pesi* o *coefficienti* della combinazione lineare (brevemente c.l.).

Esercizio 1.2 (svolto in classe). *Dati i vettori* $v_1 = (1, 0, 1)$ *e* $v_2 = (0, 1, 2)$*, dire se il vettore* $w = (2, -1, 0)$ *può essere scritto come combinazione lineare di* v_1 *e* v_2*. Rispondere alla stessa domanda per il vetttore* $u = (2, -1, 1)$.

Esercizio 1.3 (svolto in classe). *Dati i vettori coordinati in* $\mathbb{R}^3$,

$$e_1 = (1, 0, 0), \quad e_2 = (0, 1, 0) \qquad e_3 = (0, 0, 1),$$

verificare che ogni vettore $u = (u_1, u_2, u_3) \in \mathbb{R}^3$ *si può scrivere come conbinazione lineare dei vettori* e_1, e_2, e_3.

1.2. Dipendenza lineare

Iniziamo osservando che data una qualsiasi famiglia $v_1, \dots, v_p$ di vettori in $\mathbb{R}^n$, è sempre vero che

$$0v_1 + 0v_2 + \cdots + 0v_p = 0_n,$$

cioè la combinazione lineare con tutti pesi nulli dà il vettore nullo. Tale combinazione lineare si chiama a volte combinazione lineare banale. Ci interessa discutere se il vettore nullo possa essere scritto come combinazione lineare *non banale* dei vettori della famiglia $v_1, \dots, v_p$.

Definizione 1.4. *Dei vettori* $v_1, \dots, v_p \in \mathbb{R}^n$ *si dicono* linearmente dipendenti *se esistono dei numeri* $\lambda_1, \dots, \lambda_p$ ***non tutti nulli*** *tali che risulti*

$$\sum_{k=1}^{p} \lambda_k v_k = 0. \tag{1.3}$$

Se invece l'unica scelta di $\lambda_1, \dots, \lambda_p$ *che assicura* (1.3) *è la scelta banale, allora i vettori si dicono* linearmente indipendenti.

Esempio 1.5. *Verificato in classe che i vettori* $v_1 = (1, 0, 1)$, $v_2 = (0, 1, 1)$ *e* $v_3 = (1, 1, 2)$ *sono dipendenti.*

Esempio 1.6. *Discussione di alcune situazioni semplificate:* (ii)

(i) *Un singolo vettore* $v \in \mathbb{R}^n$ *è dipendente se e solo se* $v = 0_n$.

(ii) *Se una famiglia* $v_1, \dots, v_p$ *contiene il vettore nullo (ad esempio* $v_1 = 0_n$*), allora tale famiglia è dipendente. Infatti vale*

$$1v_1 + 0v_2 + \cdots + 0v_p = 0_n$$

e la combinazione lineare non è banale.

(iii) *La famiglia dei* versori coordinati e_1, e_2, e_3 *in* $\mathbb{R}^3$ *è indipendente. Infatti,*

$$\lambda_1 e_1 + \lambda_2 e_2 + \lambda_3 e_3 = 0_3 \quad \Rightarrow \quad (\lambda_1, \lambda_2, \lambda_3) = 0 \quad \Rightarrow \quad \lambda_1 = \lambda_2 = \lambda_3 = 0.$$

Lo stesso si può provare in $\mathbb{R}^n$ *per la famiglia dei versori coordinati*

$$e_1 = (1, 0, \dots, 0), \qquad e_2 = (0, 1, 0, \dots, 0), \qquad e_n = (0, 0, \dots, 1) \in \mathbb{R}^n.$$

(iv) *Due vettori v_1, v_2 in $\mathbb{R}^n$ sono dipendenti se e solo se esiste $\lambda \in \mathbb{R}$ tale che $v_2 = \lambda v_1$, oppure esiste $\mu \in \mathbb{R}$ tale che $v_1 = \mu v_2$.*

Proposizione 1.7. *Valgono le seguenti due affermazioni.*

- *Se una famiglia di vettori $v_1, \dots, v_p$ in $\mathbb{R}^n$ è dipendente, allora ogni famiglia $v_1, \dots, v_p, w_1, \dots, w_q$, con $w_1, \dots, w_q \in \mathbb{R}^n$ e $q = 0, 1, 2, \dots$ è dipendente.*
- *Se una famiglia $v_1, \dots, v_p$ è indipendente, allroa ogni famiglia che si ottiene da essa eliminando dei vettori è indipendente.*

Dimostrazione. Proviamo la prima affermazione. Siano $v_1, \dots, v_p$ dipendenti e siano $w_1, \dots, w_q \in \mathbb{R}^n$ dei vettori. Allora esistono pesi non tutti nulli $\alpha_1, \dots, \alpha_p$ tali che

$$\alpha_1 v_1 + \cdots + \alpha_p v_p = \sum_{k=1}^{p} \alpha_k v_k = 0.$$

Ma allora anche

$$\sum_{k=1}^{p} \alpha_k v_k + \sum_{j=1}^{p} 0 w_j = 0.$$

Quest'ultima combinazione lineare è non banale perché almeno uno degli α_k è non zero. Quindi abbiamo provato che la famiglia allargata $v_1, \dots, v_p, w_1, \dots, w_q$ è dipendente.

Proviamo ora il secondo punto. Siano $v_1, \dots, v_p$ indipendenti. Dobbiamo provare che ogni sottofamiglia di $v_1, \dots, v_p$ è indipendente. Per semplicità di notazioni, supponiamo che tale sottofamiglia sia costituita dai primi s vettori $v_1, \dots, v_s$, dove $s < p$. Siano $\alpha_1, \dots, \alpha_s$ dei pesi per cui $\sum_{k=1}^{s} \alpha_k v_k = 0$. Questo inplica che

$$\sum_{k=1}^{s} \alpha_k v_k + \sum_{k=s+1}^{p} 0 v_k = 0.$$

Dunque abbiamo una combinazione lineare di $v_1, \dots, v_s, v_{s+1}, \dots, v_p$ che dà il vettore nullo. Ma allora tutti i suoi pesi devono essere nulli. In particolare $\alpha_1, \dots, \alpha_s$ sono tutti nulli. □

Proposizione 1.8 (Caratterizzazione della dipendenza lineare)**.** *Una famiglia di vettori $v_1, \dots, v_p$ è linearmente dipendente se e solo se esiste almento uno tra essi che si scrive come c.l. degli altri: cioè, se esiste $j \in \{1, \dots, p\}$ tale che*

$$v_j = \lambda_1 v_1 + \cdots + \lambda_{j-1} v_{j-1} + \lambda_{j+1} v_{j+1} + \cdots + \lambda_n v_n = \sum_{k \neq j} \lambda_k v_k,$$

per opportuni scalari λ_k.

Dimostrazione. Siano $v_1, \dots, v_p$ dipendenti. Allora esistono $\alpha_1, \dots, \alpha_p$ non tutti nulli tali che $\sum_{k=1}^{p} \alpha_k v_k = 0$. Sia j tale che $\alpha_j \neq 0$. Allora sarà

$$\alpha_j v_j + \sum_{k \neq j} \alpha_k v_k = 0 \quad \Rightarrow v_j = \sum_{k \neq j} \Big(-\frac{\alpha_k}{\alpha_j} \Big) v_k.$$

Abbiamo dunque scritto il vettore α_j come combinazione lineare dei rimanenti.

Viceversa, supponiamo che data la famiglia $v_1, \dots, v_p$, risulti per almeno un j

$$v_j = \sum_{k \neq j} \lambda_k v_k$$

con una opportuna scelta di coefficienti λ_k. Allora, ovviamente,

$$v_j - \sum_{k \neq j} \lambda_k v_k = 0_n.$$

Abbiamo trovato una c.l. dei vettori $v_1, \dots, v_p$ che vale il vettore nullo e cha ha un coefficiente pari ad 1. Quindi i vettori $v_1, \dots, v_p$ sono dipendenti. □

Esempio 1.9. *Consideriamo* $v_1 = (1,1,0)$, $v_2 = (0,0,1)$ *e* $v_3 = (2,2,0)$. *Poichè la famiglia dei due vettori* v_1, v_3 *è dipendente (sono proporzionali), allora a maggior ragione la famiglia "allargata"* v_1, v_2, v_3 *è dipendente (è il contenuto della Proposizione 1.7). Osserviamo che si può scrivere facilmente*

$$v_3 = 2v_1 + 0v_2.$$

Quindi v_3 *è c.l. dei rimanenti vettori* v_1, v_2. *Quindi la Proposizione 1.8 vale, ad esempio, con* $j = 3$. *Osserviamo d'altra parte che non possiamo scrivere*

$$v_2 = \lambda_1 v_1 + \lambda_3 v_3$$

per nessuna scelta di λ_1, λ_2.

1.3. Sottospazi

Definizione 1.10 (sottospazio di $\mathbb{R}^n$)**.** *Sia* $V \subset \mathbb{R}^n$ *un sottoinsieme non vuoto. Si dice che* V *è un sottospazio di* $\mathbb{R}^n$ *se valgono le seguenti due proprietà:*

$$v + w \in V \quad \textit{per ogni possibile scelta di } v \textit{ e } w \in V; \tag{a}$$

$$\lambda v \in V \quad \textit{per ogni possibile scelta di } v \in V \textit{ e } \lambda \in \mathbb{R}. \tag{b}$$

Un sottospazio è quindi "chiuso" rispetto alle operazioni di somma e prodotto con uno scalare. Osserviamo che se V è un sottospazio, allora deve essere $0 \in V$.

Esempio 1.11. *Ecco alcuni esempi discussi in classe:*
- *(a) Casi banali:* $V = \{0\} \subset \mathbb{R}^n$. $V = \mathbb{R}^n$.
- *(b)* $V_1 = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 = 2x_1\}$ *è un sottospazio;*
- *(c)* $V_2 = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 = x_1^2\}$ *non è un sottospazio;*
- *(d)* $V_3 = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 = 1\}$ *non è un sottospazio;*

Un modo per introdurre sottospazi di $\mathbb{R}^n$ è generandoli attraverso combinazioni lineari.

Definizione 1.12 (spazio generato da un'insieme di vettori)**.** *Dati dei vettori* $v_1, \dots, v_p \in \mathbb{R}^n$, *poniamo*

$$\operatorname{span}\{v_1, \dots, v_p\} = \{\lambda_1 v_1 + \cdots + \lambda_p v_p : \lambda_1, \dots, \lambda_p \in \mathbb{R}\} = \Big\{ \sum_{k=1}^{p} \lambda_k v_k : \lambda_k \in \mathbb{R}\ \forall k \Big\}.$$

Proposizione 1.13. *Se* $v_1, \dots, v_p \subset \mathbb{R}^n$, *allora l'insieme* $\operatorname{span}\{v_1, \dots, v_p\}$ *è un sottospazio.*

Dimostrazione. Se $v, w \in V$, allora $v = \sum_{k=1}^{p} \lambda_k v_k$ e $w = \sum_{k=1}^{p} \mu_k v_k$, per opportune scelte dei pesi. Ma allora $v + w = \sum_{k=1}^{p} (\lambda_k + \mu_k) v_k \in \operatorname{span}\{v_1, \dots, v_p\}$, come desiderato.

Per verificare la seconda proprietà, basta osservare che se $v = \sum_{k=1}^{p} \lambda_k v_k$ e $\alpha \in \mathbb{R}$, allora $\alpha v = \sum_{k=1}^{p} (\alpha \lambda_k) v_k \in \operatorname{span}\{v_1, \dots, v_p\}$. □

Esempio 1.14. *Descrizione del sottospazio*

$$\operatorname{span}\{e_1, e_2\} = \{(\lambda_1, \lambda_2, 0) \in \mathbb{R}^3 : \lambda_1, \lambda_2 \in \mathbb{R}\} \subset \mathbb{R}^3.$$

Si tratta del piano nello spazio tridimensionale fatto dai punti con terza coordinata nulla. Esso coincide anche con il piano nello spazio $\mathbb{R}^3$ *che contiene i tre punti* $0_n, e_1, e_2$.

1.4. Basi di sottospazi

Breve discussione dei due esempi segenti di sottospazi vettoriali: (iii)
(a) $V_1 = \text{span}\{(1,1,0),(2,2,0)\} \subset \mathbb{R}^3$.
(b) $V_2 = \text{span}\{(0,1,0),(0,0,1)\} \subset \mathbb{R}^3$.
Osserviamo V_1 è generato anche dal solo vettore $(1,1,0)$ e che i vettori $(1,1,0),(2,2,0)$ sono dipendenti. Nel secondo caso, invece, i generatori $(0,1,0)$ e $(0,0,1)$ risultano indipendenti.

Definizione 1.15 (base di un sottospazio $V \subset \mathbb{R}^n$)**.** *Sia $V \subset \mathbb{R}^n$ un sottospazio. Una famiglia $v_1,\dots,v_p$ di vettori giacenti in V si dice* base *di V se:*
(1) risulta $\text{span}\{v_1,\dots,v_p\} = V$*;*
(2) i vettori $v_1,\dots,v_p$ sono linearmente indipendenti.

Esempio 1.16 (base canonica)**.** *I vettori*

$$e_1 = (1,0,\dots,0), \quad e_2 = (0,1,\dots,0), \quad \dots, \quad e_n = (0,\dots,0,1) \in \mathbb{R}^n$$

sono una base di $\mathbb{R}^n$. Tale base si chiama base canonica.

Esempio 1.17. *La famiglia $v_1 = (1,1)$, $v_2 = (0,1)$ è una base di $\mathbb{R}^2$. Poiché anche la famiglia e_1, e_2 è una base di $\mathbb{R}^2$, concludiamo che possono esserci più basi di un sottospazio V.*

Esercizio 1.18. *(svolti in classe)*
(1) Vedere se i vettori $v_1 = (1,1,0,1)$, $v_2 = (0,1,1,0)$ e $v_3 = (1,0,1,1)$ sono una base di $\text{span}\{v_1,v_2,v_3\}$.
(2) Stessa domande per i vettori $v_1 = (1,0,0)$, $v_2 = (0,1,1)$ e $v_3 = (0,0,0)$.

Ora ci apprestiamo a vedere che il numero di elementi che compongono due diverse basi di un sottospazio V è sempre lo stesso. Precisamente vale quanto segue:

Teorema 1.19. *Sia $V \subset \mathbb{R}^n$ è un sottospazio e sia $v_1,\dots,v_p$ una base di V. Allora ogni famiglia $w_1,\dots,w_q \in V$ di vettori con $q > p$ è linearmente dipendente.*

Dimostrazione. Tanto per cominciare osserviamo che è sufficiente provare il teorema nel caso in cui $q = p+1$. Infatti, se $w_1,\dots,w_p,w_{p+1}$ sono dipendenti e $q > p+1$, allora anche la famiglia $w_1,\dots,w_p,w_{p+1},w_{p+2},\dots,w_q$ è dipendente. Lo dice la Propisizione 1.7.

Proviamo il teorema nel caso modello già significativo in cui $p = 2$ e $q = 3$. La generalizzazione al caso qualsiasi si ottiene con lo stesso ragionamento.

Siano v_1, v_2 vettori di una base di V. Siano $w_1,w_2,w_3 \in V$. Dobbiamo provare che w_1,w_2,w_3 sono dipendenti. Possiamo scrivere

$$\begin{cases} w_1 = a_1v_1 + a_2v_2 \\ w_2 = b_1v_1 + b_2v_2 \\ w_3 = c_1v_1 + c_2v_2 \end{cases}$$

con coefficienti a_j, b_j, c_j reali. Guardiamo la prima equazione. Almeno uno tra a_1 e a_2 deve essere $\neq 0$ (altrimenti sarebbe $w_1 = 0$ e quindi la famiglia w_1,w_2,w_3 sarebbe dipendente). Possiamo assumere ad esempio $a_1 \neq 0$ (il caso $a_2 \neq 0$ si discute in modo analogo analogo). Quindi la prima equazione dice che $v_1 \in \text{span}\{w_1,v_2\}$. Ma allora, dalla seconda e terza equazione si deduce che

$$w_2 \in \text{span}\{w_1,v_2\} \quad \text{e} \quad w_3 \in \text{span}\{w_1,v_2\}.$$

In altre parole, per coefficienti opportuni $\lambda_1,\lambda_2,\mu_1,\mu_2 \in \mathbb{R}$ vale

$$\begin{cases} w_2 = \lambda_1w_1 + \lambda_2v_2 \\ w_3 = \mu_1w_1 + \mu_2v_2 \end{cases}$$

A questo punto, se $\lambda_2 = 0$, abbiamo finito (la prima equazione diventa una relazione di dipendenza per la famiglia w_1, w_2, w_3). Se invece $\lambda_2 \neq 0$, allora possiamo ricavare v_2 dalla prima equazione e scrivere $v_2 = k_1 w_1 + k_2 w_2$. Inserendo nella seconda equazione troviamo una relazione di lineare dipendenza tra i vettori w_1, w_2, w_3. □

Ecco una conseguenza immediata di quanto appena provato:

Corollario 1.20. *Due basi $v_1, \ldots, v_p$ e $w_1, \ldots, w_q$ di uno stesso sottospazio $V \subset \mathbb{R}^n$ hanno necessariamente lo stesso numero di elementi: $p = q$.*

A questo punto ha senso la seguente definizione

Definizione 1.21 (Dimensione di un sottospazio)**.** *La dimensione di un sottospazio $V \subset \mathbb{R}^n$ è il numero di elementi di una qualsiasi sua base.*

Osservazione 1.22. *Valgono i seguenti due fatti.*

(a) La dimensione di $\mathbb{R}^n$ è n. Infatti, la famiglia $e_1, \ldots, e_n$ è una base di $\mathbb{R}^n$. (Verificato in classe).

(b) In $\mathbb{R}^n$ non si possono trovare più di n vettori indipendenti. Segue dal Teorema 1.19. Ad esempio, senza nessun calcolo si può affermare che i tre vettori $v_1 = (1,2)$, $v_2 = (1,0)$ e $v_3 = (3,1) \in \mathbb{R}^2$ sono dipendenti.

Osservazione 1.23. *Se $v_1, \ldots, v_p$ sono una base di V sottospazio di $\mathbb{R}^n$, allora per ogni $v \in V$ la scelta dei numeri $\lambda_1, \ldots, \lambda_p \in \mathbb{R}$ che realizzano l'uguaglianza $v = \sum_{j=1}^p \lambda_j v_j$ è unica. Infatti, se fosse*

$$v = \sum_{j=1}^{p} \lambda_j v_j \quad e \quad v = \sum_{j=1}^{p} \mu_j v_j,$$

sottraendo membro a menbro le due uguaglianze, si otterrebbe

$$\sum_{j=1}^{p} (\lambda_j - \mu_j) v_j = 0_n.$$

Ma, poiché i vettori $v_1, \ldots, v_p$ sono indipendenti, questo implica che $\lambda_j = \mu_j$ per ogni j.

Linguaggio. *Se $v_1, \ldots, v_p$ sono una base di V sottospazio di $\mathbb{R}^n$ e se $z \in V$ si scrive come $z = \sum_j \lambda_j v_j$, allora i numeri $\lambda_1, \ldots, \lambda_p$ si chiamano* coordinate *del vettore z rispetto alla base $v_1, \ldots, v_p$.*

Esempio 1.24 (Discusso in classe)**.** *Siano $v_1 = (1,1,0)$, $v_2 = (0,1,0)$ e $v_3 = (1,0,2)$. Verificare che i vettori v_1, v_2, v_3 sono indipendenti. Si verifica (compito a casa) che sono una base di $\mathbb{R}^3$ (cioè che ogni vettore z può essere scritto come c.l. dei detti vettori). Trovare le coordinate del vettore $z = (4,0,4) \in \mathbb{R}^3$ nella base assegnata.*

1.5. Esercizi per casa

(a) Verificare che i vettori

$$e_1 = (1,0,\ldots,0), \quad e_2 = (0,1,\ldots,0), \quad \ldots, \quad e_n = (0,\ldots,0,1) \in \mathbb{R}^n$$

sono indipendenti.

(b) Per le seguenti famiglie di vettori, si dica se sono dipendenti o indipendenti:

$$w_1 = (1,1,0), \quad w_2 = (0,1,2), \quad w_3 = (1,1,1)$$

$$v_1 = (1,-1,0), \quad v_2 = (1,3,-1), \quad v_3 = (5,3,-2)$$
$$u_1 = (6,2,3,4), \quad u_2 = (0,5,3,-1), \quad u_3 = (0,0,7,-2)$$

(c) Dire quale dei seguenti insiemi è un sottospazio:

$$V_1 = \{(x,0) \in \mathbb{R}^2 : x \geq 0\}$$
$$V_2 = \{(x,y) \in \mathbb{R}^2 : x + y = 3\}$$
$$V_3 = \{(x_1,\dots,x_n) \in \mathbb{R}^n : \sum_{k=1}^{n} b_k x_k = 0\} \quad \text{con } b_1,\dots,b_n \in \mathbb{R} \text{ parametri fissati}$$
$$V_4 = \{(x_1,x_2,x_3,x_4) \in \mathbb{R}^4 : \ x_1 + 2x_3 + x_4 = 0 \quad \text{e} \quad x_3 - x_1 + x_2 = 0\}$$
$$V_5 = \{0,1,-1,2,-2,3,-3,\dots\} = \{\text{l'insieme dei numeri interi}\}$$

(d) Sono assegnati i vettori in $\mathbb{R}^4$ $v_1 = (1,1,0,0)$, $v_2 = (0,1,1,-1)$, $v_3 = (2,0,-3,1)$. Tali vettori sono una base di $\mathbb{R}^4$? Sono una base di $\operatorname{span}\{v_1,v_2,v_3\}$? Il vettore $w = (3,0,-4,2)$ sta nello $\operatorname{span}\{v_1,v_2,v_3\}$? In caso affermativo, dire quali sono i pesi da attribuire a una c.l. di v_1,v_2,v_3 affinché essa uguagli w.

Osserviamo anche il seguente fatto (iv)

Proposizione 1.25. *Se V è un sottospazio di $\mathbb{R}^n$ di dimensione p, allora una famiglia di p vettori linearmente indipendenti è necessariamente una base di V.*

Dimostrazione. Ragioniamo per assurdo. Se $v_1,\dots,v_p$ non fossero una base di V, questo significherebbe che l'inclusione

$$\operatorname{span}\{v_1,\dots,v_p\} \subset V$$

è stretta. Cioè che esiste $z \in V$ ma

$$z \notin \operatorname{span}\{v_1,\dots,v_p\}. \tag{1.4}$$

Facciamo vedere che questo ci porta alla conclusione che la famiglia $v_1,\dots,v_p,z$ è indipendente e quindi a contraddire il Teorema 1.19.

Analizziamo la c.l.

$$\lambda_1 v_1 + \dots + \lambda_p v_p + \mu z = 0 \tag{1.5}$$

L'uguaglianza (1.5) non può valere con $\mu \neq 0$, perché altrimenti sarebbe $z \in \operatorname{span}\{v_1,\dots,v_p\}$, in contraddizioone con (1.4). Allora deve essere $\mu = 0$. Pertanto, dalla (1.5) e dalla lineare indipendenza di $v_1,\dots,v_p$, otteniamo che $\lambda_1 = \dots = \lambda_p = 0$. Abbiamo quindi scoperto che la famiglia di $p+1$ vettori $v_1,\dots,v_p,w$ è linearmente indipendente e contenuta in V. Questo contraddice, come già detto, il Teorema 1.19. Dunque deve essere $\operatorname{span}\{v_1,v_2,\dots,v_p\} = V$ e quindi i vettori $v_1,\dots,v_p$ sono una base di V. □

Esempio 1.26. *Il rusultato precedente si applica ad esempio nelle seguenti situazioni.*

- *I vettori $v_1 = (1,1)$ e $v_2 = (1,2)$ sono una base di $\mathbb{R}^2$ perché sono indipendenti in $\mathbb{R}^2$ che ha dimensione* 2.
- *Una famiglia di n vettori $v_1,\dots,v_n \in \mathbb{R}^n$ sono una base si $\mathbb{R}^n$ se sono indipendenti.*

2. Struttura euclidea di $\mathbb{R}^n$

2.1. Prodotto scalare, vettori ortogonali e norma

Introduciamo il prodotto scalare euclideo in $\mathbb{R}^n$.

Definizione 2.1 (prodotto scalare canonico in $\mathbb{R}^n$). *Dati $x = (x_1,\dots,x_n)$ e $y = (y_1,\dots,y_n) \in \mathbb{R}^n$, il numero reale*

$$\langle x,y \rangle := \sum_{j=1}^{n} x_j y_j = x_1 y_1 + \dots + x_n y_n$$

si chiama prodotto scalare *di x ed y.*

Il prodotto scalare è un'operazione binaria che a partire da due vettori restituisce un numero reale come risultato. Esso possiede le seguenti proprietà: per ogni $x, y, z \in \mathbb{R}^n$ e $\lambda \in \mathbb{R}$

$$\langle x, y\rangle = \langle y, x\rangle \tag{P1}$$
$$\langle x+y, z\rangle = \langle x, z\rangle + \langle y, z\rangle \tag{P2}$$
$$\langle \lambda x, y\rangle = \lambda\langle x, y\rangle \tag{P3}$$
$$\langle x, x\rangle \geq 0 \quad \text{e} \tag{P4}$$
$$\langle x, x\rangle = 0 \quad \text{se e solo se } x = 0 := (0, \dots, 0) \in \mathbb{R}^n. \tag{P4bis}$$

La proprietà (P1) esprime la *simmetria* del prodotto scalare. Le proprietà (P2) e (P3) assieme si chiamano *linearità* rispetto al primo argomento. Utilizzando la simmetria e la linearità rispetto al primo argomento è facile riconoscere che la proprietà di linearità vale anche rispetto al secondo argomento. Cioè

$$\begin{aligned} &\langle x, y+z\rangle = \langle x, y\rangle + \langle x, z\rangle \quad \text{e} \\ &\langle x, \lambda y\rangle = \lambda\langle x, y\rangle \qquad\qquad \text{per ogni } x, y, z \in \mathbb{R}^n \quad \lambda \in \mathbb{R}. \end{aligned}$$

Esempio 2.2. *Calcolare* $\langle (1,2), (-1,3)\rangle$ *in* $\mathbb{R}^2$ *e* $\langle x, (\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n})\rangle$ *in* $\mathbb{R}^n$.

Definizione 2.3 (vettori ortogonali o perpendicolari)**.** *Due vettori x ed y in $\mathbb{R}^n$ si dicono ortogonali se* $\langle x, y\rangle = 0$.

Esempio 2.4 (discussi in classe)**.**
- *In $\mathbb{R}^n$ il vettore nullo 0_n e un qualsiasi altro vettore sono ortogonali.*
- *In $\mathbb{R}^2$ il vettore $u = (a, b)$ è perpendicolare al vettore $v = (-b, a)$.*
- *In $\mathbb{R}^3$ due qualsiasi dei vettori coordinati e_1, e_2, e_3 sono ortogonali tra loro. vale dunque*

$$\langle e_1, e_2\rangle = \langle e_1, e_3\rangle = \langle e_2, e_3\rangle = 0$$

- *Analogamente, i vettori della base canonica di $\mathbb{R}^n$ soddisfano*

$$\langle e_j, e_k\rangle = 0 \quad \textit{per ogni } j, k = 1, \dots, n \textit{ con } j \neq k,$$

Linguaggio. *Se x e y sono vettori ortogonali, si dice anche che x* è ortogonale a *y oppure equivalentemente che y* è ortogonale a *x. A volte si usa anche la scrittura $x \perp y$ oppure $y \perp x$.*

Definizione 2.5 (norma euclidea)**.** *Dato $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, definiamo il numero reale $\|x\| \in \mathbb{R}$, che chiameremo* norma di *x nel modo seguente:*

$$\|x\| := \sqrt{\langle x, x\rangle} = \sqrt{\sum_j x_j^2} = \sqrt{x_1^2 + \cdots + x_n^2}.$$

Proprietà della norma: si deducono da quelle del prodotto scalare.

$$\|\lambda x\| = |\lambda|\,\|x\| \tag{N1}$$
$$\|x\| \geq 0 \quad \text{e} \tag{N2}$$
$$\|x\| = 0 \quad \text{se e solo se } x = 0 := (0, \dots, 0) \in \mathbb{R}^n. \tag{N2bis}$$
$$\|x+y\| \leq \|x\| + \|y\| \quad \text{per ogni } x, y \in \mathbb{R}^n. \tag{N3}$$

La proprietà (N3) si chiama *disuguaglianza triangolare*.

Dato un vettore $x \in \mathbb{R}^n$, $x \neq 0$, allora $\|x\| > 0$. Il nuovo vettore ((v))

$$\frac{1}{\|x\|}x = \frac{x}{\|x\|}$$

ha norma uno e punta nella stessa direzione di x. Infatti

$$\left\| \frac{1}{\|x\|} x \right\| = \left| \frac{1}{\|x\|} \right| \|x\| = 1$$

Tale vettore si chiama *normalizzato* di x. Ad esempio, normalizziamo $(1,2)$.

$$\frac{(1,2)}{\|(1,2)\|} = \Big(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}} \Big).$$

Esercizio 2.6 (Svolto in classe)**.** *Verificare che per ogni $x \in \mathbb{R}^n$, $y \in \mathbb{R}^n$ vale la formula*

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x, y \rangle \tag{2.1}$$

usando le proprietà del prodotto scalare.

La formula (2.1) generalizza la formula per il quadrato di un binomio al caso vettoriale. Essa implica immediatamente il "Teorema di Pitagora"

Teorema 2.7 (Teorema di Pitagora)**.** *Se $x \perp y$, allora vale*

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2 .$$

Discusso in classe il legame con il Teorema di Pitagora classico della geometria piana.

Esercizio 2.8. *Verificare che se $v_1, v_2, v_3 \in \mathbb{R}^n$ sono ortogonali a coppie, allora*

$$\|v_1 + v_2 + v_3\|^2 = \|v_1\|^2 + \|v_2\|^2 + \|v_3\|^2$$

Una formula analoga vale con p vettori $v_1, \dots, v_p$ ortogonali a coppie in $\mathbb{R}^n$.

Definizione 2.9 (distanza euclidea tra due vettori)**.** *Dati x ed $y \in \mathbb{R}^n$, la distanza $d(x,y)$ tra x ed y è il numero*

$$d(x,y) := \|x-y\| = \sqrt{\sum_{k=1}^{n} (x_k - y_k)^2}.$$

Dalle proprietà della norma deduciamo che $d(x,y) \geq 0$ e che $d(x,y) = 0$ se e solo se $x = y$.

2.2. Proiezione di un vettore su un sottospazio unidimensionale

Dato $a \in \mathbb{R}^n$, vettore non nullo, prendiamo il sottospazio unidimensionale

$$V = \operatorname{span}\{a\} = \{\lambda a : \lambda \in \mathbb{R}\}.$$

Dato un vettore qualsiasi $b \in \mathbb{R}^n$, vogliamo trovare il vettore di V che ha distanza minima da b. Il problema è banale se $b \in V$ (basta scegliere b stesso). È evidente però che prendendo un punto "a caso" in $\mathbb{R}^n$ non possiamo aspettarci che esso appartenga a V.

Per risolvere il problema posto, seguendo la formula della Definizione 2.9, dobbiamo trovare il valore di λ per cui è minimo

$$d(b, \lambda a) = \|b - \lambda a\| .$$

Osserviamo che un numero λ è punto di minimo per la funzione $\lambda \mapsto d(b, \lambda a)$ se e solo se tale numero è punto di minimo per la funzione $\lambda \mapsto d(b, \lambda a)^2$. Per comodità conviene lavorare sulla funzione con il quadrato per poter usare la formula (2.1). Poniamo

$$g(\lambda) := \|b - \lambda a\|^2 = \|b\|^2 - 2\lambda \langle b, a \rangle + \lambda^2 \|a\|^2 .$$

La funzione è un polinomio di grado due. Derivando troviamo $g'(\lambda) = -2\langle b, a\rangle + 2\lambda \|a\|^2$, che si annulla nell'unico punto
$$\lambda^* = \frac{\langle b, a\rangle}{\|a\|^2}.$$
Il punto λ^* è di minimo perché la derivata seconda $g''(\lambda^*) = 2\|a\|^2$ è positiva. Quindi il vettore
$$\lambda^* a = \frac{\langle b, a\rangle}{\|a\|^2} a$$
è la soluzione del problema. Osserviamo che il grafico di g è una parabola con la concavità verso l'alto. Il punto λ^* è l'unico punto di minimo di g.

Abbiamo quindi dimostrato il seguente teorema

Teorema 2.10. *Se $a \neq 0$, $b \in \mathbb{R}^n$ e indichiamo con $V = \operatorname{span}\{a\} \subset \mathbb{R}^n$ il sottospazio unidimensionale contenente a, allora il vettore*
$$\frac{\langle b, a\rangle}{\|a\|^2} a$$
è il vettore di V che ha minima distanza da b. Tale vettore è unico. [1]

Linguaggio. *Il vettore trovato nel teorema appena enunciato si chiama* proiezione di b su V. *Indicheremo tale vettore con $\pi_V b$:*
$$\pi_V b = \frac{\langle b, a\rangle}{\|a\|^2} a.$$

Esercizio 2.11. *Svolti in classe i seguenti esercizi*
- *Calcolare la proiezione $\pi_V(1,2)$ del vettore $(1,2)$ su $V = \operatorname{span}\{(0,1)\}$.*
- *Calcolare la proiezione $\pi_V(\frac{1}{2}, \frac{3}{2})$ del vettore $(\frac{1}{2}, \frac{3}{2})$ su $V = \operatorname{span}\{(2,1)\}$.*

Facendo delle figure corrispondenti all'esercizio appena fatto si osserva che in entrambi i casi il vettore $q = b - \pi_V b$ è ortogonale al vettore a. Tale fatto vale in generale, come afferma la seguente proposizione:

Proposizione 2.12. *Sia $a \neq 0$ in $\mathbb{R}^n$ e sia $V = \operatorname{span}\{a\} \subset \mathbb{R}^n$. Allora il vettore proiezione $\pi_V b$ soddisfa*
$$(b - \pi_V b) \perp a \tag{2.2}$$

Tale proprietà di ortogonalità giustifica la denominazione "proiezione ortogonale" di b su V per il vettore $\pi_V b$.

Dimostrazione. Per verificare l'affermazione basta calcolare, usando le proprietà (P2) e (P3), il prodotto scalare
$$\left\langle b - \frac{\langle b, a\rangle}{\|a\|^2}, a \right\rangle = \langle b, a\rangle - \frac{\langle b, a\rangle}{\|a\|^2}\langle a, a\rangle = 0,$$
come desiderato. □

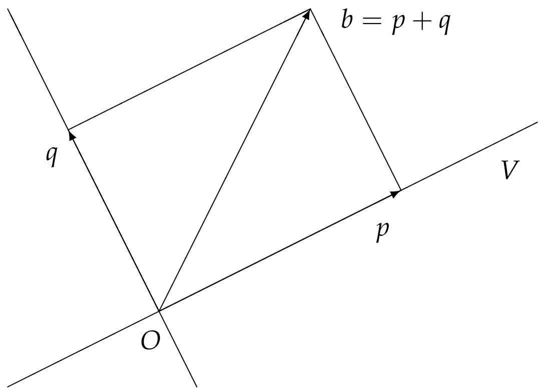
Osservazione 2.13. *La procedura appena discussa ci permette di scrivere, assegnati il sottospazio unidimensionale $V = \operatorname{span}\{a\}$ e il vettore $b \in \mathbb{R}^n$, una decomposizione del tipo*
$$b = p + q \tag{2.3}$$

[1] Osserviamo che se cambiamo a con un vettore $a' = \mu a$, con $\mu \neq 0$, il sottospazio V non cambia e – come ci si può aspettare – la proiezione non cambia: $\frac{\langle b, a\rangle}{\|a\|^2} a = \frac{\langle b, a'\rangle}{\|a'\|^2} a'$

con

$$\begin{cases} p \in \operatorname{span}\{a\} \\ q \perp a \end{cases} \tag{2.4}$$

dove $p = \pi_V b$ *e* $q = b - \pi_V b$. *Osserviamo che una decomposizione del tipo* (2.3) *con le proprietà aggiuntive* (2.4) *è unica. Infatti se fosse* $b = p' + q'$ *e* p', q' *soddisfacessero* (2.4), *allora sarebbe* $p' - p = q' - q$. *Ma* $p - p' \in \operatorname{span}\{a\}$, *mentre* $(q - q') \perp a$. *In definitiva,* $\|p' - p\|^2 = \langle p' - p, q' - q\rangle = 0$ *e pertanto* $p - p' = 0 = q - q'$.



Esercizio 2.14 (svolto in classe)**.** *Calcolare la proiezione di* $(1, 0, 1)$ *sulla retta* $V = \operatorname{span}\{(3, 4, 0)\}$.

Osservazione 2.15. *Se* $V = \operatorname{span}\{a\}$ *con* $a \neq 0$, *un vettore* $b \in \mathbb{R}^n$ *soddisfa* $b = \pi_V b$ *se e solo se* (vi)
$b \in V$. *Infatti, da una parte, se* $b \in \operatorname{span}\{a\}$, *allora l'unico punto di minima distanza di* b *da* V *è* $\pi_V b = b$ *stesso che ha distanza zero. Viceversa, se assumiamo che* $b = \pi_V b$, *allora*

$$b = \pi_V b \quad \Rightarrow \quad b = \frac{\langle b, a\rangle}{\|a\|^2} a \quad \Rightarrow \quad b \in \operatorname{span}\{a\} = V.$$

Proposizione 2.16. *Se* $V = \operatorname{span}\{a\} \subseteq \mathbb{R}^n$ *con* $a \neq 0$, *allora vale*

$$\|\pi_V b\| \leq \|b\| \tag{2.5}$$

e vale l'uguaglianza in (2.5) *se e solo se* $b \in V$.

La proposizione afferma in sostanza che applicando l'operazione di proiezione su un sottospazio a un vettore $b \in \mathbb{R}^n$, la lunghezza del vettore proiettato non è mai maggiore di quella del vettore di partenza.

Dimostrazione. Basta ricordare la proprietà di ortogonalità dei due addendi nella decomposizione $b = \pi_V b + (b - \pi_V b)$ e usare il teorema di Pitagora:

$$\|b\|^2 = \|\pi_V b\|^2 + \|b - \pi_V b\|^2 \geq \|\pi_V b\|^2.$$

L'uguaglianza vale se e solo se $\|b - \pi_V b\| = 0$, cioeè se e solo se $b = \pi_V b$ che equivale a $b \in V$. □

Una consegenza importante (di fatto una riscrittura in termini differenti) della stima (2.5) è la disuguaglianza seguente:

Proposizione 2.17 (disuguaglianza di Cauchy–Schwarz)**.**

$$|\langle a, b\rangle| \leq \|a\| \, \|b\| \tag{2.6}$$

per ogni $a, b \in \mathbb{R}^n$. *Inoltre nella disuguaglianza* (2.6) *vale l'uguaglianza se e solo se* a *e* b *sono linearmente dipendenti (cioè collineari).*

Dimostrazione. Se $a = 0$ non c'è nulla da dimostrare. Se $a \neq 0$, riscriviamo (2.5):

$$\|\pi_V b\| \leq \|b\| \quad \Leftrightarrow \quad \left\| \frac{\langle b, a\rangle}{\|a\|^2} a \right\| \leq \|b\| \quad \Leftrightarrow \quad \frac{|\langle b, a\rangle|}{\|a\|^2} \|a\| \leq \|b\|,$$

che dà immediatamente (2.6). Vale l'uguaglianza se e solo se $b \in \operatorname{span}\{a\}$. □

Una conseguenza della disuguaglianza di Cauchy–Schwarz è la disuguaglianza triangolare:

Proposizione 2.18 (disuguaglianza triangolare). *Vale per ogni $a, b \in \mathbb{R}^n$*

$$\|a + b\| \leq \|a\| + \|b\|$$

Dimostrazione. Partiamo dalla formula del quadrato di un binomio (versione vettoriale) e usiamo Cauchy–Schwarz al momento opportuno

$$\begin{aligned} \|a + b\|^2 &= \|a\|^2 + \|b\|^2 + 2\langle a, b\rangle \leq \|a\|^2 + \|b\|^2 + 2|\langle a, b\rangle| \\ &\leq \|a\|^2 + \|b\|^2 + 2\,\|a\|\,\|b\| = (\|a\| + \|b\|)^2 \end{aligned}$$

e la prova si conclude facendo la radice quadrata. □

2.3. Famiglie e basi ortonormali

Definizione 2.19. *Una famiglia di vettori $v_1, \dots, v_p$ in $\mathbb{R}^n$ si dice* ortogonale *se $\langle v_j, v_k\rangle = 0$ per ogni $j \neq k$. Se i vettori hanno anche tutti norma unitaria, $\|v_j\| = 1$ per ogni j, alloira la famiglia si dice* ortonormale*.*

Esempio 2.20. *visti in classe:*
(1) La famiglia $v_1 = (1, 0, 0)$, $v_2 = (0, 3, 4)$ è ortogonale. Normalizzando v_2 diventa ortonormale.
(2) Una qualsiasi sotto famiglia di $e_1, \dots, e_n$ in $\mathbb{R}^n$ è ortonormale.
(3) I vettori $v_1 = (1, 1)$ e $v_2 = (-2, 2)$ sono ortogonali. Normalizzarli per renderli ortonormali.

Proposizione 2.21. *Vettori ortonormali sono indipendenti.*

Dimostrazione. Prendiamo $v_1, \dots, v_p \in \mathbb{R}^n$ ortonormali. Scriviamo $\sum_{j=1}^p \lambda_j v_j = 0$. Moltiplicando questa uguaglianza per v_k con $k \in \{1, \dots, p\}$ otteniamo

$$\left\langle v_k, \sum_{j=1}^p \lambda_j v_j \right\rangle = \langle v_k, 0\rangle = 0.$$

Ma a sinistra, usando la linearità del prodotto scalare rispetto al secondo argomento, possiamo scrivere

$$\left\langle v_k, \sum_{j=1}^p \lambda_j v_j \right\rangle = \sum_{j=1}^p \lambda_j \langle v_k, v_j\rangle = \lambda_k.$$

Dunque otteniamo che $\lambda_k = 0$. Ripetendo per ogni k si ottiene la lineare indipendenza dei vettori in questione. □

Definizione 2.22 (base ortonormale). *Se $V \subset \mathbb{R}^n$ è un sottospazio, una famiglia ortonormale che sia anche una base di V si chiama* base ortonormale *di V.*

Esempio 2.23.
- *La base canonica $e_1, \dots, e_n$ è una base ortonormale di $\mathbb{R}^n$.*
- *La famiglia $v_1 = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ e $v_2 = (-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ è una base ortonormale di $\mathbb{R}^2$.*

Proposizione 2.24. *Se $v_1, \dots, v_p$ è una base ortonormale di V, allora per ogni $x \in V$ abbiamo lo sviluppo (con scelta dei coefficienti unica)*
$$x = \sum_{j=1}^{p} \langle x, v_j \rangle v_j$$

Dimostrazione. Basta prendere $x \in V$. Poiché $v_1, \dots, v_p$ è una base di V, esistono unici dei coefficienti c_j per i quali vale $x = \sum_{j=1}^{p} c_j v_j$. Se fissiamo $k \in \{1, \dots, p\}$ e moltiplichiamo l'uguaglianza appena scritta per il vettore v_k, troviamo
$$\langle v_k, x \rangle = \Big\langle v_k, \sum_{j=1}^{p} c_j v_j \Big\rangle = \sum_{j=1}^{p} c_j \langle v_k, v_j \rangle = c_k.$$
Dunque abiamo scoperto che $c_k = \langle v_k, x \rangle$, per ogni $k = 1, \dots, p$, come volevamo. □

Esercizio 2.25. *(Svolto in classe.)*
(a) *Scrivere i coefficienti che appaiono quando si sviluppa il vettore $(3, 4)$ come combinazione dei vettori $v_1 = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ e $v_2 = (-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$.*
(b) *Verificare che i vettori*
$$v_1 = \Big(\frac{1}{2}, 0, \frac{\sqrt{3}}{2}\Big) \quad v_2 = (0, 1, 0) \quad v_3 = \Big(-\frac{\sqrt{3}}{2}, 0, \frac{1}{2}\Big)$$
sono una base ortonormale di R^3. Trovare le coordinate in tale base del vettore $w = (1, 2, 1)$.
(c) *Sia $V = \{x \in \mathbb{R}^3 : x_1 + x_2 + x_3 = 0\}$. Verificare che i vettori*
$$v_1 = \Big(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}\Big) \quad v_2 = \Big(\frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}\Big)$$
sono una base ortonormale di V. Scrivere le coordinate in tale base del vettore $w = (-2, 3, -1)$.

2.4. Costruzione di basi ortonormali. Procedimento di Gram–Schmidt

Descriviamo ora il metoro di Gram–Schmidt, che, a partire da una base qualsiasi $a_1, \dots, a_p$ di $V = \mathrm{span}\{a_1, \dots, a_p\}$, permette di ottenere una base ortonormale $v_1, \dots, v_p$ di V. Descriviamo tale procedimento dapprima in dimensione bassa.

Sottospazi di dimensione $p = 2$. Se V è un sottospazio di dimensione 2 generato dai vettori a_1 e a_2, $V = \mathrm{span}\{a_1, a_2\}$, allora la nuova famiglia
$$\begin{cases} v_1 = a_1 \\ v_2 = a_2 - \dfrac{\langle a_2, v_1 \rangle}{\|v_1\|^2} v_1 \end{cases}$$
è una base ortonormale di V. Nella costruzione, non abbiamo fatto altro che sottrarre al vettore a_2 la sua proiezione sul sottospazio generato da v_1. Sappiamo dalla lezione precedente che v_2 cosí costruito è ortogonale a v_1. [2] Fatti questi passaggi, è sufficiente normalizzare i vettori v_1 e v_2 per ottenere una famiglia $v_1' = \dfrac{v_1}{\|v_1\|}$, $v_2' = \dfrac{v_2}{\|v_2\|}$ che è una base ortonormale di V.

Esempio 2.26 (svolto in classe)**.** *Applicazione del metodo alla famiglia $v_1 = (1, 1)$ e $v_2 = (1, 2)$.*

Caso di sottospazi di dimensione $p = 3$. Se $V = \mathrm{span}\{a_1, a_2, a_3\}$ è un sottospazio di $\mathbb{R}^n$ di (vii)

[2] Notiamo che v_2 non può essere 0 perché altrimenti sarebbe $a_2 \in \mathrm{span}\{a_1\}$.

dimensione 3 e a_1, a_2, a_3 sono una base, allora si può ottenere una base ortogonale definendo

$$\begin{cases} v_1 = a_1 \\ v_2 = a_2 - \dfrac{\langle a_2, v_1 \rangle}{\|v_1\|^2} v_1 \\ v_3 = a_3 - \dfrac{\langle a_3, v_1 \rangle}{\|v_1\|^2} v_1 - \dfrac{\langle a_3, v_2 \rangle}{\|v_2\|^2} v_2 \end{cases}$$

Abbiamo ottenuto il vettore v_3 sottraendo da a_3 le sue proiezioni su a_1 e su a_2. È facile verificare che v_3 è perpendicolare a entrambi v_1 e v_2. La famiglia v_1, v_2, v_3 è dunque una famiglia ortogonale e ciascuno dei tre vettori trovati è non nullo (questo segue dal fatto che a_1, a_2, a_3 sono indipendenti). Possiamo dunque normalizzare i tre vettori e ottenere infine una base ortonormale di V.

Il procedimento per famiglie con più di tre vettori dovrebbe essere a questo punto chiaro. Partendo da $a_1, \dots, a_p$ indipendenti in $\mathbb{R}^n$, otteniamo una famiglia ortogonale che genera $\operatorname{span}\{a_1, \dots, a_p\}$ ponendo

$$\begin{cases} v_1 = a_1 \\ v_2 = a_2 - \dfrac{\langle a_2, v_1 \rangle}{\|v_1\|^2} v_1 \\ v_3 = a_3 - \dfrac{\langle a_3, v_1 \rangle}{\|v_1\|^2} v_1 - \dfrac{\langle a_3, v_2 \rangle}{\|v_2\|^2} v_2 \\ \vdots \\ v_p = a_p - \displaystyle\sum_{k=1}^{p-1} \dfrac{\langle a_p, v_k \rangle}{\|v_k\|^2} v_k \end{cases}$$

Esercizio 2.27 (svolto in classe)**.** *Applicare il processo di ortogonalizzazione alla famiglia* $a_1 = (1, 1, 0)$, $a_2 = (0, 1, 0)$, $a_3 = (0, -1, 2)$ *in* $\mathbb{R}^3$.

Definizione 2.28 (sottospazio ortogonale)**.** *Se* $V \subset \mathbb{R}^n$, *allora il sottospazio*

$$V^\perp = \{x \in \mathbb{R}^n : \langle x, v \rangle = 0 \; \forall \, v \in V\}$$

si chiama sottospazioo ortogonale di V.

Di interesse particolare è il caso in cui l'insieme V di partenza è anch'esso un sottospazio.

Se conosciamo un insieme di generatori di V, anche non indipendenti, allora si può trovare $V^\perp$ usando la seguente proposizione.

Proposizione 2.29. *Se* $V = \operatorname{span}\{a_1, \dots, a_p\} \subset \mathbb{R}^n$ *è un sottospazio, allora*

$$V^\perp = \{x \in \mathbb{R}^n : \langle x, a_1 \rangle = \cdots = \langle x, a_p \rangle = 0\}. \tag{2.7}$$

Dimostrazione. Sia $x \in V^\perp$. Allora, in particolare $\langle x, a_j \rangle = 0$ per ogni j. Questo prova l'inclusione $\subseteq$ in (2.7). Ora proviamo $\supseteq$. Sia x tale che $\langle x, a_j \rangle = 0$ per ogni $j = 1, \dots, p$. Dobbiamo provare che $x \in V^\perp$, cioè che $x \perp v$ per ogni $v \in \operatorname{span}\{a_1, \dots, a_p\}$. Prendiamo ora $v = \sum_j \lambda_j a_j \in \operatorname{span}\{a_1, \dots, a_p\}$. Allora, per la linearità del prodotto scalare rispetto al secondo argomento,

$$\Big\langle x, \sum_j \lambda_j a_j \Big\rangle = \sum_j \lambda_j \langle x, a_j \rangle = 0,$$

perché ciascun $\langle x, a_j \rangle = 0$. $\square$

Esercizio 2.30 (In classe)**.** *Sia* $V = \mathrm{span}\{(1,1,0,0,0),(0,0,-1,2,1)\}$. *Descrivere* $V^\perp$.

Osservazione 2.31 (unicità della decomposizione in somma ortogonale)**.** *Notiamo che se* $V \subset \mathbb{R}^n$ *è un sottospazio, allora per ogni vettore* $x \in \mathbb{R}^n$ *la decomposizione*

$$x = p + q \quad \text{con } p \in V \text{ e } q \in V^\perp \tag{2.8}$$

è unica. Infatti, se ci fosse una seconda decomposizione $x = p' + q'$ *con le stesse proprietà, sarebbe* $x = p + q = p' + q'$ *e quindi* $p - p' = q' - q$. *Allora*

$$\|p - p'\|^2 = \langle p - p', p - p' \rangle = \langle p - p', q' - q \rangle = 0,$$

perché $p - p' \in V$ *e* $q' - q \in V^\perp$ *sono perpendicolari. In definitiva,* $p - p' = 0$ *e quindi anche* $q' - q = 0$.

2.5. Proiezione ortogonale su un sottospazio di dimensione qualsiasi

Mostriamo qui come le proprietà della proiezione su sottospazi unidimensionali si estendono a sottospazi qualsiasi.

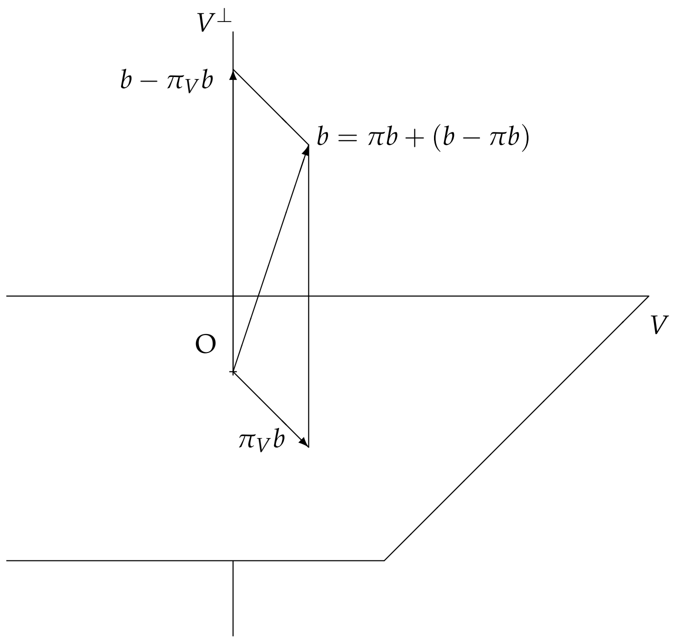
Teorema 2.32. *Dato un sottospazio* V *si* $\mathbb{R}^n$, *per ogni* $b \in \mathbb{R}^n$ *esiste un unico vettore* $\pi_V b \in V$ *tale che*

$$\|b - \pi_V b\| = \min_{p \in V} \|v - p\| .$$

Tale vettore possiede la seguente proprietà: se decomponiamo $b = \pi_V b + (b - \pi_V b)$, *allora risulta*

$$b - \pi_V b \in V^\perp.$$

Si può visualizzare la situazione tramite la seguente figura.



Dimostrazione del Teorema 2.32. Consideriamo una base ortonormale di V, $v_1, \ldots, v_d$. Ogni vettore $p \in S$ si scrive nella forma $p = \sum_{k=1}^d \lambda_k v_k$. Allora

$$\|p - b\|^2 = \Big\| \sum_{k=1}^d \lambda_k v_k - b \Big\|^2 = \sum_{k=1}^d \lambda_k^2 + \|b\|^2 - 2 \sum_{k=1}^d \lambda_k \langle v_k, b \rangle$$

Abbiamo usato il Teorema di Pitagora per famiglie di d vettori (Esercizio 2.8), la linearità del prodotto scalare rispetto ai suoi argomenti e il fatto che i vettori v_j hanno norma unitaria. Possiamo anche riscrivere "completando il quadrato"

$$\|p-b\|^2 = \sum_{k=1}^{d} \lambda_k^2 + \|b\|^2 - 2\sum_{k=1}^{d} \lambda_k \langle v_k, b\rangle = \|b\|^2 + \sum_{k=1}^{d} \{\lambda_k - \langle v_k, b\rangle\}^2 - \sum_k \langle v_k, b\rangle^2.$$

Dobbiamo trovare la scelta (o le scelte) di $\lambda_1, \dots, \lambda_p$ tali che la quantitàs appena scritta è minima. Poiché il primo e il terzo addendo non dipendono dai λ_j, analizziamo solo il termine centrale. Ciascuno degli addendi $\{\dots\}^2$ è minimo se e solo se $\lambda_k = \langle v_k, b\rangle$. Quindi otteniamo la formula utile

$$\pi_V b = \sum_{k=1}^{d} \langle b, v_k\rangle v_k.$$

Non resta che verificare la proprietà di ortogonalità: preso un qualsiasi v_j, abbiamo

$$\langle b - \pi_V b, v_j\rangle = \left\langle b - \sum_{k=1}^{d} \langle b, v_k\rangle v_k, v_j\right\rangle = 0,$$

usando la linearità del prodotto scalare e il fatto che $\langle v_k, v_j\rangle$ è sempre zero, a meno che $k = j$ (in tal caso vale 1). A questo punto la Proposizione 2.29 ci dice che $b - \pi_V b \perp \operatorname{span}\{v_1, \dots, v_d\}$, come desideravamo $\square$

Esercizio 2.33 (svolto in classe). *Calcolare la proiezione di $b = (1, 2, 1)$ sul sottospazio*

$$V = \operatorname{span}\Big\{\Big(-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\Big), \Big(\frac{1}{2}, \frac{1}{\sqrt{2}}, \frac{1}{2}\Big)\Big\}.$$

A casa: scrivere la proiezione di un vettore (b_1, b_2, b_3) generico sul medesimo sottospazio.

Osserviamo infine che se V è un sottospazio di $\mathbb{R}^n$, allora

$$\dim(V) + \dim(V^\perp) = n$$

Non dimostriamo questa affermazione [3]

2.6. Esercizi per casa

1. Dato $a = (1, 1, 0, 2)$ e $V = \operatorname{span}\{a\}$, trovare $\pi_V(1, -1, 3, 0)$.
2. Dato $a = (1, -2, 1)$ e posto $V = \operatorname{span}\{a\}$, calcolare per $x = (x_1, x_2, x_3) \in \mathbb{R}^3$ generico la proiezione $\pi_V x$.
3. Prendere in $\mathbb{R}^n$ due vettori coordinati e_j ed e_k con $j \neq k$. Quanto vale $\|e_j + e_k\|$? Posto $V = \operatorname{span}\{e_j + e_k\}$, calcolare $\pi_V x$ per $x \in \mathbb{R}^n$ qualsiasi.
4. Sia $V = \operatorname{span}\{(1, 1, 1), (2, 0, 2)\}$. Trovare una base ortonormale di V.
5. Sia $V = \operatorname{span}\{(1, 2, 0, 3), (0, 0, 1, 1), (1, 2, -1, 2)\}$. Trovare una base ortonormale di V.
6. Sia $V = \{x \in \mathbb{R}^3 : x_1 + 2x_2 - x_3 = 0\}$. Individuare una base di V e applicare se necessario il procedimento di ortonormalizzazione per ottenere una base ortonormale di V.

[3] Per verificarla basta considerare delle basi ortonormali $v_1, \dots, v_p$ di V e $w_1, \dots, w_q$ di $V^\perp$ e riconoscere che la famiglia $v_1, \dots, v_p, w_1, \dots, w_q$ è una base di $\mathbb{R}^n$. Con considerazioni simili si può riconoscere che, se V è un sottospazio, allora $(V^\perp)^\perp = V$.

7. Descrivere il sottospazio $(\text{span}\{(1,2,1)\})^{\perp}$.
8. Trovare una base di

$$V = \{x \in \mathbb{R}^4 : x_1 + x_3 = 0 \quad \text{e} \quad x_2 + 2x_3 - x_4 = 0\} \subset \mathbb{R}^4$$

9. Trovare una base di $(\text{span}\{(1,0,2,0),(1,1,-1,2)\})^{\perp}$.

3. Matrici

Definiamo (viii)

$$\mathbb{R}^{m\times n} := \left\{ A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} : a_{jk} \in \mathbb{R} \ \forall j = 1,\dots,m \quad k = 1,\dots,n \right\} \tag{3.1}$$

L'elemento di posto j,k della matrice A è il numero a_{jk} Una matrice in $\mathbb{R}^{m\times 1}$ si chiama *vettore colonna.* Le colonne di A sono gli m vettori colonna

$$\begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix}, \begin{bmatrix} a_{12} \\ \vdots \\ a_{m2} \end{bmatrix}, \dots, \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix} \in \mathbb{R}^{m\times 1}$$

Una matrice in $R^{1\times n}$ si chiama *vettore riga.* Le *righe* di A sono

$$[a_{11},\dots,a_{1n}],\dots,[a_{21},\dots,a_{2n}],\dots,[a_{m1},\dots,a_{mn}] \in \mathbb{R}^{1\times n}.$$

Sulle matrici possiamo effettuare operazioni di somma e di prodotto con scalare, esattamente come sui vettori. Se A e $B \in \mathbb{R}^{m\times n}$ e se $\lambda \in \mathbb{R}$, allora

$$(A+B)_{jk} = a_{jk} + b_{jk}$$
$$(\lambda A)_{jk} = \lambda a_{jk}$$

Svolti in classe: calcolo di somme e prodotti con scalare su alcune matrici numeriche. Osserviamo anche le uguaglianze $A + (-A) = 0_{m\times n}$ e $A + 0_{m\times n} = A$ per ogni $A \in \mathbb{R}^{m\times n}$. Qui $0_{m\times n}$ indica la matrice i cui elementi sono tutti zero e $-A = (-1)A$.

Attenzione: due matrici A e B si possono sommare solo se hanno lo stesso numero di righe e di colonne.

3.1. Prodotto di matrici

Date $A \in \mathbb{R}^{m\times n}$ e $B \in \mathbb{R}^{n\times p}$, possiamo definire il prodotto $AB \in \mathbb{R}^{m\times p}$, il cui elemento j,k ha la forma

$$(AB)_{jk} := \sum_{\ell=1}^{n} a_{j\ell} b_{\ell k} \tag{3.2}$$

Osservamo che $(AB)_{jk}$ è il prodotto scalare tra la j-esima riga di A e la k-esima colonna di B.

Esercizio 3.1. *Svolti in classe i calcoli di alcuni prodotti di matrici.*

Osservazione 3.2. *Il prodotto AB tra matrici A e B non si può fare comunque siano A e B. Occorre che il numero di colonne di A sia pari al numero di righe di B. Ad esempio, date $A = \left(\begin{smallmatrix} 1 & 2 \\ 0 & 1 \end{smallmatrix}\right)$ e $B = \left(\begin{smallmatrix} 1 & 2 & 3 \\ 0 & 1 & 0 \end{smallmatrix}\right)$ calcolare AB. Si provi (...) a calcolare BA.*

Osservazione 3.3. *Il prodotto tra matrici possiede la proprietà associativa:*

$$(AB)C = A(BC) \quad \forall A \in \mathbb{R}^{m\times n}\ b \in \mathbb{R}^{n\times p}\ C \in \mathbb{R}^{p\times q}.$$

In vista di tale proprietà non è equivoca la scrittura ABC.

Inoltre vale la proprietà distributiva

$$\begin{aligned} A(B+B') &= AB + AB' \quad \forall A \in \mathbb{R}^{m\times n} \quad B, B' \in \mathbb{R}^{n\times p} \\ (A+A')B &= AB + A'B \quad \forall A, A' \in \mathbb{R}^{m\times n} \quad B \in \mathbb{R}^{n\times p} \end{aligned}$$

Verifichiamo a titolo di esempio la proprietà associativa. Siano A, B, C tre matrici per cui i prodotti AB e BC sono definiti. Allora, se indichiamo con M_{jk} l'elemento di posto j, k di una matrice M, abbiamo

$$\begin{aligned} ((AB)C)_{jk} &= \sum_\ell (AB)_{j\ell} C_{\ell k} = \sum_\ell \Big(\sum_s A_{js} B_{s\ell}\Big) C_{\ell k} = \sum_s \sum_\ell A_{js} B_{s\ell} C_{\ell k} \\ &= \sum_s A_{js} \sum_\ell B_{s\ell} C_{\ell k} = \sum_s A_{js} (BC)_{sk} = (A(BC))_{jk} \end{aligned}$$

Osservazione 3.4 (Mancanza della proprietà commutativa)**.** *In generale non è detto che se il prodotto AB è definito lo sia anche il prodotto BA. Anche nel caso particolare delle matrici quadrate* ***non è vero*** *che il prodotto tra matrici possiede la proprietà commutativa. Per convincersene fare la prova con le matrici*

$$A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \qquad e \quad B = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}.$$

Linguaggio. *Una matrice si dice* quadrata *se ha lo stesso numero di righe e di colonne. $\mathbb{R}^{n\times n}$ è l'insieme delle matrici quadrate. La* diagonale *di una matrice quadrata $A \in \mathbb{R}^{n\times n}$ è costituita dagli n elementi $a_{11}, \dots, a_{nn}$.*

Definizione 3.5 (matrice identità)**.** *Definiamo la matrice $I_n \in \mathbb{R}^{n\times n}$ come segue*

$$I_n := \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

La matrice ha la seguente proprietà: se $A \in \mathbb{R}^{p\times q}$, allora

$$AI_q = A = I_p A.$$

Verificato in classe che $AI_q = A$.

Definizione 3.6 (matrice trasposta)**.** *Data $A \in \mathbb{R}^{n\times m}$, definiamo $A^T \in \mathbb{R}^{m\times n}$ come segue:*

$$(A^T)_{jk} := A_{kj}$$

In sostanza l'operazione di trasposizione scambia righe e colonne di una matrice. Ad esempio $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$. Vale ovviamente $(A^T)^T = A$, per ogni matrice A.

Esercizio 3.7. *Verificato in classe che $(AB)^T = B^T A^T$.* [4] (ix)

[4]Un'altra proprietà utile, non discussa in classe, descrive il comportamento dell'operazione di trasposizione nei prodotti scalari. Precisamente, se $A \in \mathbb{R}^{m\times n}$, allora, pensando i vettori di $\mathbb{R}^n$ con vettori colonna,

$$\langle Ax, y\rangle = \langle x, A^T y\rangle \quad \text{per ogni } x \in \mathbb{R}^{n\times 1} = \mathbb{R}^n \text{ e } y \in \mathbb{R}^{m\times 1} = \mathbb{R}^m. \tag{3.3}$$

Definizione 3.8 (matrice simmetrica). *Una matrice quadrata $A \in \mathbb{R}^{n\times n}$ si dice simmetrica se $A = A^T$, o se, in altre parole $a_{jk} = a_{kj}$ per ogni scelta di j, k.*

Esercizio 3.9 (in classe). *Verificare che $A^T A$ è una matrice simmetrica, comunque sia $A \in \mathbb{R}^{p\times q}$. Verificare anche che se B è quadrata, allora $B + B^T$ è simmetrica.*

3.2. Matrice inversa

Definiamo ora la matrice inversa di una matrice quadrata $A \in R^{n\times n}$.

Definizione 3.10 (matrice inversa). *Sia $A \in R^{n\times n}$. Una matrice $B \in \mathbb{R}^{n\times n}$ si chiama* inversa *di A se vale*

$$AB = BA = I_n. \tag{3.4}$$

Osservazione 3.11 (unicità dell'inversa). *Una matrice $A \in \mathbb{R}^{n\times n}$ ha al massimo una inversa Infatti, se B e $C \in \mathbb{R}^{n\times n}$ fossero entrambe inverse di A, allora sarebbe $AB = I_n$ e $CA = I_n$. Moltiplicando a sinistra la prima uguaglianza per C, otterremmo*

$$CAB = CI_n = C \quad \Rightarrow B = C,$$

dove abbiamo usato al membro di sinistra il fatto che $CA = I_n$.

Se A ammette inversa, allora si dice che A è *invertibile* e la sua inversa si indica con A^{-1}.

Esempio 3.12. *Calcolo dell'inversa di $A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$*

È ovvio che la matrice $A = 0_{n\times n}$ non ha inversa. È però importante osservare che ci sono matrici non nulle che non ammettono inversa.

Esempio 3.13. *Discusso in classe il fatto che la matrice $A = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$ non è invertibile. Più concretamente questo significa che se prendiamo una qualsiasi matrice $B = \begin{pmatrix} x & y \\ z & t \end{pmatrix}$, nessuna scelta delle incognite x, y, z, t rende vera l'uguaglianza $\begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} x & y \\ z & t \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Discussione in classe.*

Esempio 3.14. *Una matrice* diagonale $A = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & \lambda_n \end{pmatrix}$ *è invertibile se e solo se tutti gli gli n numeri $\lambda_1, \dots, \lambda_n$ sono diversi da zero. In tal caso vale $A^{-1} = \begin{pmatrix} \lambda_1^{-1} & 0 & \dots & 0 \\ 0 & \lambda_2^{-1} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & \lambda_n^{-1} \end{pmatrix}$*

Osservazione 3.15. *La matrice inversa possiede le seguenti proprietà immediate: se A e $B \in \mathbb{R}^{n\times n}$ sono invertibili, allora AB è invertibile e vale*

$$(AB)^{-1} = B^{-1}A^{-1}$$

Inoltre per ogni A invertibile, vale $(A^{-1})^{-1} = A$.

Un esempio interessante, che evidenzia gli aspetti nuovi del prodotto di matrici rispetto al prodotto tra numeri è il seguente.

Esempio 3.16. *Se* $M = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$ *e* $N = \begin{pmatrix} 2 & 3 \\ -2 & -3 \end{pmatrix}$, *allora vale* $MN = 0$. *Questo evidenzia il fatto che non vale la legge di annullamento del prodotto tra matrici. Le matrici* M *ed* N *hanno prodotto nullo senza che nessuno dei due fattori sia la matrice nulla.*

L'esempio precedente suggerisce anche che, lavorando con le matrici, ci vuole cautela nel fare "semplificazioni" di fattori nei prodotti. Le matrici invertibili sono quelle che si possono "semplificare" nei prodotti. Precisamente:

Proposizione 3.17. *Se* $A \in \mathbb{R}^{n\times n}$ *è invertibile e se* $B, C \in \mathbb{R}^{n\times p}$, *allora*

$$AB = AC \quad \Leftrightarrow \quad B = C. \tag{3.5}$$

Dimostrazione. L'implicazione $\Leftarrow$ è ovvia. Quella $\Rightarrow$ si ottiene moltiplicando a sinistra per A^{-1} l'uguaglianza $AB = AC$. □

Con la scelta $A = M$, $B = N$ e $C = 0$, dove M, N sono quelle dell'esempio 3.16, si vede che (3.5) è falsa.

4. Sistemi lineari

Iniziamo ad analizzare la seguente equazione matriciale. Sia $A \in \mathbb{R}^{m\times n}$, $x \in \mathbb{R}^{n\times 1}$ un vettore colonna e $b \in \mathbb{R}^{m\times q}$ un altro vettore colonna. Ha senso dal punto di vista delle dimensioni delle matrici, scrivere l'uguaglianza

$$Ax = b \tag{S1}$$

Pensiamo che A e b siano dei dati assegnati, mentre x è un'incognita. Cosí stiamo impostando un'equazione matriciale di primo grado. In maniera più esplicita, se scriviamo [5]

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \in \mathbb{R}^{m\times n} \quad x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^{n\times 1}, \quad b = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

e esplicitiamo il prodotto, otteniamo

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{cases} \tag{S2}$$

Questo è il modello di *sistema lineare di m equazioni in n incognite*. I dati del problema sono i numeri a_{jk} e b_i, mentre le incognite sono $x_1, \ldots, x_n$.

Se partiamo dal sistema (S2), esso può essere scritto in modo compatto nella forma (S1).

Linguaggio. *Le matrici* $A \in \mathbb{R}^{m\times n}$ *e* $b \in \mathbb{R}^{m\times 1}$, *che individuano univocamente il sistema* (S2), *affiancate assieme, danno luogo a una nuova matrice* $[A|b]$, *che ha la forma*

$$[A|b] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right] \in \mathbb{R}^{m\times(n+1)}$$

e che si chiama matrice completa *del sistema. La matrice* A *si chiama invece* matrice incompleta.

[5] Avvertenza: usualmente, nel contesto del calcolo matriciale, in vettore $x \in \mathbb{R}^n$ si identifica con il vettore colonna $x \in \mathbb{R}^{n\times 1}$.

Esercizio 4.1 (in classe)**.** *Scrivere il sistema corrispondente all'equazione metriciale $Ax = b$ individuata dai dati $A = \left(\begin{smallmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{smallmatrix}\right)$ e $b = \left(\begin{smallmatrix} 7 \\ 8 \end{smallmatrix}\right)$*

Scrivere viceversa le matrici A, b che individuano il sistema di tre equazioni in quattro incognite

$$\begin{cases} x_1 + x_2 + 2x_3 = 1 \\ x_2 + x_4 = 4 \\ x_3 - 2x_4 = 5 \end{cases}$$

Esempio 4.2 (discussione dei sistemi con una sola equazione)**.** *Prima di addentrarci nella discussione dei sistemi in generale, guardiamo il caso modello delle equazioni scalari in n incognite.*

Consideriamo l'equazione

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b.$$

In questo caso l'approccio con le matrici non è necessario. Distinguiamo alcune situazioni:

- *Caso 1. Vale $a_1 = \cdots = a_n = b = 0$. In tal caso ogni scelta di $x_1, \ldots, x_n$ è soluzione dell'equazione. Ci sono dunque infinite soluzioni e si dice che l'equazione è* indeterminata.
- *Caso 2. Vale $a_1 = \cdots = a_n = 0$ ma $b \neq 0$. In tal caso non c'è nessuna soluzione. L'equazione si dice* impossibile.
- *Caso 3. Almeno uno dei coefficienti $a_1, \ldots, a_n$ è diverso da zero. Ad esempio $a_1 \neq 0$. Allora l'equazione si riscrive nella forma*

$$x_1 = -\frac{1}{a_1}\sum_{k=2}^{n} a_kx_k + \frac{b}{a_1}.$$

Ogni vettore della forma $x = \left(-\frac{1}{a_1}\sum_{k=2}^{n} a_kp_k + \frac{b}{a_1}, p_2, \ldots, p_n\right)$ è soluzione del sistema. Dunque vale quanto segue.
  - *Se $n \geq 2$ le soluzioni sono infinite e possono essere "parametrizzate" da $n-1$ parametri che abbiamo indicato con $p_2, \ldots, p_n$. Per ogni scelta di tali parametri si ottiene una soluzione dell'equazione. L'equazione è dunque indeterminata.*
  - *Se invece $n = 1$, l'equazione è della forma $a_1x = b$ e c è un'unica soluzione $x = -\frac{b}{a_1} \in \mathbb{R}$, come noto dalla scuola superiore. L'equazione si dice in tal caso* determinata.

4.1. Sistemi a scala

I seguenti tre sistemi sono degli esempi di sistemi a scala: (x)

$$\begin{cases} x_1 + x_2 = 3 \\ x_2 = 2 \end{cases} \qquad \begin{cases} x_1 + x_3 + x_4 = 1 \\ x_2 - 2x_3 = 0 \\ x_3 - 2x_4 = 0 \end{cases} \qquad \text{e} \qquad \begin{cases} x_1 + x_2 - x_3 + 2x_4 + x_6 = 2 \\ x_2 + x_5 = 2 \\ x_4 + 3x_6 = 0 \\ x_5 + x_6 = 1. \end{cases}$$

In ciascuno di essi, l'incognita di indice piú basso che appare nella riga j-esima ha indice inferiore a quella dell'incognita di indice piú basso che compare nella $(j+1)$-esima. Ad esempio il secondo di tali sistemi si risolve "dal basso in alto" come segue:

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

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

In definitiva, tutte le variabili si esprimono in termini di x_4, che rimane indeterminata. Se indichiamo $x_4 = p$, allora per ogni valore di $p \in \mathbb{R}$ troveremo una soluzione che ha la forma

$$x = (-1 - 3p, 4p + 4, 2p + 2, p).$$

Nel terzo sistema, le sei incognite si esprimono tutte in termini di due parametri liberi (svolto in classe).

Guardiamo le matrici dei sistemi sopra scritti:

$$\left[\begin{array}{cc|c} 1 & 1 & 3 \\ 0 & 1 & 2 \end{array}\right] \qquad \left[\begin{array}{cccc|c} 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & -2 & 0 & 0 \\ 0 & 0 & 1 & -2 & 0 \end{array}\right] \quad \text{e} \quad \left[\begin{array}{cccccc|c} 1 & 1 & -1 & 2 & 0 & 1 & 2 \\ 0 & 1 & 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{array}\right]$$

Questi soo tre esempi di matrici a scala. Le matrici associate ai sistemi a scala si chiamano *matrici a scala* e sono definite come segue:

Definizione 4.3 (pivot e matrice a scala)**.** *Se $A \in \mathbb{R}^{m\times n}$, il* pivot *della riga $j-esima$, $[a_{j1}, \ldots, a_{jn}]$ è l'elemento non nullo che sta piú a sinistra nella riga. Una riga identicamente nulla non ha pivot. Una matrice è a scala se:*

- *tutte le righe nulle sono in fondo;*
- *il pivot di ogni riga è strettamente a sinistra del pivot della riga sottostante.*

Esempio 4.4. *Dire quale delle seguenti matrici è a scala:*

$$\begin{bmatrix} 0 & 1 & 2 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{bmatrix}, \qquad \begin{bmatrix} 1 & 2 \\ 0 & 0 \\ 0 & 3 \end{bmatrix}, \qquad \begin{bmatrix} 0 & 1 & 2 \\ 0 & 3 & 4 \end{bmatrix}.$$

4.2. Operazioni ammissibili per portare un sistema alla forma triangolare

Il nostro proposito è quello di mostrare come un sistema qualsiasi si possa trasformare in un sistema a scala lasciando invariato l'insieme delle soluzioni. Partendo dal sistema $Ax = b$ troveremo un nuovo sistema a scala $Ux = c$ con U della stessa taglia di A tale che x risolve il primo se e solo se risolve il secondo.

Conviene scrivere il sistema (S2) nella forma

$$\begin{cases} r_1 x = b_1 \\ \vdots \\ r_m x = b_m \end{cases} \tag{S3}$$

dove $r_j = (a_{j1}, \ldots, a_{jn}) \in \mathbb{R}^{1\times n}$ è la j-esima riga della matrice A.

Proposizione 4.5 (operazioni elementari e soluzioni di un sistema)**.** *Una qualsiasi delle seguenti tre* operazioni elementari *lascia invariato l'insieme delle soluzioni del sistema.*

(a) Moltiplicare una equazione per un numero $\lambda \neq 0$.

(b) Scambiare due equazioni tra loro

(c) Sostituire la coppia di equazioni

$$\begin{cases} r_j x = b_j \\ r_k x = b_k \end{cases} \qquad (*)$$

con la nuova coppia

$$\begin{cases} r_j x = b_j \\ r_k - \lambda r_j x = b_k - \lambda b_j \end{cases} \qquad (**)$$

dove $\lambda \in \mathbb{R}$ e lasciare invariate tutte le altre.

Dimostrazione. Essendo (a) e (b) immediate, verifichiamo solo (c). Se x risolve $(*)$, allora possiamo moltiplicare la prima equazione per λ e sottrarre tale uguaglianza membro a membro dalla seconda equazione. Si ottiene quindi che x soddisfa $(**)$.

In modo analogo si riconosce che se x soddisfa $(**)$, allora soddisfa anche $(*)$. □

4.3. Operazioni ammissibili viste sulle matrici e algoritmo di Gauss

Consideriamo la matrice completa del sistema lineare $Ax = b$:

$$[A|b] = \left[\begin{array}{c|c} r_1 & b_1 \\ \vdots & \vdots \\ r_m & b_m \end{array}\right] =: \begin{bmatrix} R_1 \\ \vdots \\ R_m \end{bmatrix} \in \mathbb{R}^{m\times(n+1)}. \tag{4.1}$$

È facile osservare che le tre operazioni (a), (b) e (c) sulle equazioni del sistema corrispondono alle seguenti tre operazioni sulla matrice $[A|b]$

(i) Moltiplicare una riga per un numero $\lambda \neq 0$.
(ii) Scambiare due righe tra loro
(iii) Prese due righe R_k ed R_j e un numero $\lambda \in \mathbb{R}$, sostituire la riga k-esima, R_k, con la riga $R_k - \lambda R_j$.

A volte è comodo sostituire l'operazione (iii) con la seguente, un po' piú generale, che si ottiene applicando in sequenza la (i) e la (iii):

(iii)′ presi $\mu \neq 0$ e $\lambda \in \mathbb{R}$, sostituire la riga R_k con la riga $\mu R_k - \lambda R_j$,

L'operazione (iii)′ si ottiene dapprima moltiplicando la riga R_k per $\mu \neq 0$ e poi sottraendo ad essa λR_j. **Attenzione: nell'operazione (iii)′ è cruciale che sia $\mu \neq 0$.** Altrimenti tutta l'informazione contenuta nell'equazione k-esima si perde.

Applicando in modo opportuno le operazioni elementari dette sopra, è possibile trasformare la matrice $[A|b]$ di un qualsiasi sistema alla forma "a scala". Tale procedimento si chiama *algoritmo di Gauss*. Lo descriveremo attraverso alcuni esempi:

Esempio 4.6 (svolto in classe)**.** *Risolvere, usando l'algoritmo di Gauss sulle matrici, i sistemi*

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

Risolviamo in queste note a titolo di esempio il secondo sistema, partendo dalla matrice completa

$$\left[\begin{array}{ccc|c} 0 & 2 & 1 & 1 \\ 1 & 1 & -2 & 0 \\ 0 & 1 & -1 & 1 \end{array}\right] \overset{\text{scambio } R_2 \text{ e } R_1}{\simeq} \left[\begin{array}{ccc|c} 1 & 1 & -2 & 0 \\ 0 & 2 & 1 & 1 \\ 0 & 1 & -1 & 1 \end{array}\right] \overset{R_3 \mapsto 2R_3 - R_2}{\simeq} \left[\begin{array}{ccc|c} 1 & 1 & -2 & 0 \\ 0 & 2 & 1 & 1 \\ 0 & 0 & -3 & 1 \end{array}\right]$$

Abbiamo indicato con $\simeq$ i passaggi che separano due matrici che si ottengono l'una dall'altra attraverso una operazione elementare sulle righe. Sopra tale simbolo è indicata sinteticamente l'operazione (ad esempio $R_3 \mapsto 2R_3 - R_2$ significa che stiamo sostituendo la terza riga R_3 con la riga $2R_3 - R_2$). A questo punto il sistema è diventato a scala della forma seguente

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

e può essere risolto esplicitamente per sostituzione "all'indietro". La soluzione è $(-\frac{4}{3}, \frac{2}{3}, -\frac{1}{3})$. Verificare che le coordinate di questo vettore soddisfano tanto il sistema di partenza quanto quello a scala apppena ottenuto.

(xi)

Esempio 4.7. *Risolvere i sistemi*

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

dove $t \in \mathbb{R}$ é un paramtero assegnato.

Iniziamo dal primo e passiamo alle matrici:

$$\left[\begin{array}{cc|c} 1 & 2 & 3 \\ 3 & -1 & 1 \\ 2 & -3 & 0 \end{array}\right] \overset{R_2 \mapsto R_2 - 3R_1}{\simeq} \left[\begin{array}{cc|c} 1 & 2 & 3 \\ 0 & -7 & -8 \\ 2 & -3 & 0 \end{array}\right] \overset{R_3 \mapsto R_3 - 2R_1}{\simeq} \left[\begin{array}{cc|c} 1 & 2 & 3 \\ 0 & -7 & -8 \\ 0 & -7 & -6 \end{array}\right] \overset{R_3 \mapsto R_3 - R_2}{\simeq} \left[\begin{array}{cc|c} 1 & 2 & 3 \\ 0 & -7 & -8 \\ 0 & 0 & 2 \end{array}\right]$$

Se scriviamo il sistema corrispondente, notiamo che la terza equazione ha la forma $0 = 2$. Dunque il sistema non ha soluzioni (sistema impossibile).

Ora risolviamo il secondo, sempre usando l'algoritmo di Gauss:

$$\left[\begin{array}{ccc|c} 2 & -1 & 3 & 0 \\ 1 & 0 & 1 & 2 \\ 1 & -1 & 2 & t \end{array}\right] \overset{R_2 \mapsto 2R_2 - R_1}{\simeq} \left[\begin{array}{ccc|c} 2 & -1 & 3 & 0 \\ 0 & 1 & -1 & 4 \\ 1 & -1 & 2 & t \end{array}\right] \overset{R_3 \mapsto 2R_3 - R_1}{\simeq} \left[\begin{array}{ccc|c} 2 & -1 & 3 & 0 \\ 0 & 1 & -1 & 4 \\ 0 & -1 & 1 & 2t \end{array}\right]$$

$$\overset{R_3 \mapsto R_3 + R_2}{\simeq} \left[\begin{array}{ccc|c} 2 & -1 & 3 & 0 \\ 0 & 1 & -1 & 4 \\ 0 & 0 & 0 & 2t+4 \end{array}\right]$$

Il sistema corrispondente è

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

Se $t \neq -2$, allora la terza equazione non è verificata per nessun x e il sistema è impossibile. Se invece $t = -2$, allora possiamo risolvere all'indietro e esprimere tutto in termini di un parametro libero $x_4 = p$. Le soluzioni saranno allora del tipo $(2 - p, 4 + p, p)$, con $p \in \mathbb{R}$. Tali soluzioni risolvono sia il sistema originale che quello a scala. Provare per credere. Il sistema ha infinite soluzioni e si dice indeterminato.

Risolviamo infine l'ultimo sistema:

$$\left[\begin{array}{cccc|c} 2 & 1 & 3 & 0 & 0 \\ 1 & 2 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 2 & 2 \end{array}\right] \overset{R_2 \mapsto 2R_2 - R_1}{\simeq} \left[\begin{array}{cccc|c} 2 & 1 & 3 & 0 & 0 \\ 0 & 3 & -1 & 0 & 2 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 2 & 2 \end{array}\right] \overset{R_3 \mapsto 2R_3 - R_1}{\simeq} \left[\begin{array}{cccc|c} 2 & 1 & 3 & 0 & 0 \\ 0 & 3 & -1 & 0 & 2 \\ 0 & -1 & -1 & 2 & 0 \\ 1 & 0 & 1 & 2 & 2 \end{array}\right]$$

$$\overset{R_4 \mapsto 2R_4 - R_1}{\simeq} \left[\begin{array}{cccc|c} 2 & 1 & 3 & 0 & 0 \\ 0 & 3 & -1 & 0 & 2 \\ 0 & -1 & -1 & 2 & 0 \\ 0 & -1 & -1 & 4 & 4 \end{array}\right] \overset{R_3 \mapsto 3R_3 + R_2}{\simeq} \left[\begin{array}{cccc|c} 2 & 1 & 3 & 0 & 0 \\ 0 & 3 & -1 & 0 & 2 \\ 0 & 0 & -4 & 6 & 2 \\ 0 & -1 & -1 & 4 & 4 \end{array}\right]$$

$$\overset{R_4 \mapsto 3R_4 + R_2}{\simeq} \left[\begin{array}{cccc|c} 2 & 1 & 3 & 0 & 0 \\ 0 & 3 & -1 & 0 & 2 \\ 0 & 0 & -4 & 6 & 2 \\ 0 & 0 & -4 & 12 & 14 \end{array}\right] \overset{R_4 \mapsto R_4 - R_3}{\simeq} \left[\begin{array}{cccc|c} 2 & 1 & 3 & 0 & 0 \\ 0 & 3 & -1 & 0 & 2 \\ 0 & 0 & -4 & 6 & 2 \\ 0 & 0 & 0 & 6 & 12 \end{array}\right]$$

che corrisponde al sistema

$$\begin{cases} 2x_1 + x_2 + 3x_3 = 0 \\ 3x_2 - x_3 = 2 \\ -4x_3 + 6x_4 = 2 \\ 6x_4 = 12 \end{cases}$$

Risolvando questo ultimo sistema all'indietro otteniamo l'unica soluzione $(x_1, x_2, x_3, x_4) = (-\frac{9}{2}, \frac{3}{2}, \frac{5}{2}, 2)$, che risolve anche il sistema iniziale.

Osservazione 4.8. *Abbiamo visto negli esempi – ed è vero in generale – che un sistema lineare $Ax = b$ può avere nessuna, una o infinite soluzioni. Nessun'altra evenienza può verificarsi. Dal punto di vista del linguaggio, il sistema $Ax = b$ si dice:*
- determinato *se ha una unica soluzione;*
- indeterminato*, se ammette infinite soluzioni;*
- impossibile*, se non ammette nessuna soluzione.*

Osservazione 4.9. *Se partiamo da una matrice A osservata attraverso le sue righe, $A = \begin{bmatrix} R_1 \\ \vdots \\ R_m \end{bmatrix}$ e applichiamo l'algoritmo di Gauss, arriviamo a una nuova matrice, a scala $\begin{bmatrix} S_1 \\ \vdots \\ S_m \end{bmatrix}$. Tale matrice sarà fatta da un certo numero di righe non nulle (diciamo $r \leq m$) ed eventualmente qualche riga nulla: $S_{r+1} = \cdots = S_m = 0$*

$$\begin{bmatrix} S_1 \\ \vdots \\ S_m \end{bmatrix} = \begin{bmatrix} S_1 \\ \vdots \\ S_r \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Allora succedono due cose:
- *vale* $\operatorname{span}\{R_1, \dots, R_m\} = \operatorname{span}\{S_1, \dots, S_m\} = \operatorname{span}\{S_1, \dots, S_r\}$*;*
- *le righe non nulle della matrice a scala ottenuta sono indipendenti*

In altre parole, possiamo affermare che l'algoritmo di Gauss ci permette di trovare una base dello spazio $\operatorname{span}\{R_1, \dots, R_m\}$*.*

Esercizio 4.10. *(in classe) Usando l'osservazione appena fatta, trovare una base del sottospazio di $\mathbb{R}^3$* $\operatorname{span}\{(1, 2, 1), (2, 1, 2), (3, 3, 3,)\}$*.*

4.4. Sistemi omogenei e spazio nullo di una matrice

Un sistema $Ax = b$ si dice omogeneo se $b = 0$, cioè se si scrive nella forma

$$Ax = 0.$$

Questo tipo di sistema ha sempre una soluzione almeno: il vettore nullo, che viene chiamata in gergo la "soluzione banale". L'insieme delle soluzioni di un sistema omogeneo è un sottospazio, che si chiama spazio nullo della matrice A.

Definizione 4.11 (Spazio nullo di $A \in \mathbb{R}^{m \times n}$)**.** *Data $A \in \mathbb{R}^{m \times n}$, lo spazio nullo $\mathcal{N}(A)$ della matrice A è l'insieme delle soluzioni del sistema omogeneo $Ax = 0$:*

$$\mathcal{N}(A) := \{x \in \mathbb{R}^{n \times 1} : Ax = 0\}. \tag{4.2}$$

È immediato verificare che $\mathcal{N}(A)$ è un sosttospazio.

4.5. Sistemi lineari visti attraverso le colonne di A e soluzioni ai minimi quadrati

Sia $A \in \mathbb{R}^{n \times p}$ e $b \in \mathbb{R}^{n \times 1}$. Il prodotto Ax si scrive nella forma: (xii)

$$Ax = [a_1, \dots, a_p]x = a_1 x_1 + \cdots + a_p x_p. \tag{4.3}$$

Se pensiamo $x_1, \dots, x_p$ come pesi, allora tale prodotto è una combinazione lineare delle colonne di A, cioè di $a_1, \dots, a_p$. [6] Tenendo in mente questo punto di vista, osserviamo che si può riformulare la nozione di lineare indipendenza con l'aiuto della notazione matriciale:

[6] Approfittiamo per menzionare che, nel contesto del calcolo matriciale, si tende spesso ad identificare lo spazio $\mathbb{R}^n$ con $\mathbb{R}^{n \times 1}$. Cioè i vettori di $\mathbb{R}^n$ li pensiamo usualmente come vettori colonna.

Proposizione 4.12. *Dei vettori* $a_1, \dots, a_p \in \mathbb{R}^n$ *sono indipendenti se e solo se il sistema lineare omogeneo* $Ax = 0$ *la cui matrice* A *ha tali vettori come colonne ammette soluzioni non banali.*

Dimostrazione. I vettori sono dipendenti se e solo se $a_1 x_1 + \cdots + a_p x_p = 0$ per almeno una scelta di $x \neq 0$, il vettore nullo. Questo significa affermare che il sistema $Ax = 0$ ha soluzioni non banali. □

Esercizio 4.13 (svolti in classe)**.** *Verificare per quali valori di* $p \in \mathbb{R}$ *i vettori* $(1, 2, 3)$, $(2, 1, 3)$ *e* $(1, 3, p)$ *sono indipendenti.*

Contunuando la disacussione sul prodotto $[a_1, \dots, a_p]x$, osserviamo che, al variare di $x \in \mathbb{R}^p$, tale prodotto descrive tutto lo spazio generato dalle colonne di A.

$$\{Ax : x \in \mathbb{R}^p\} = \{a_1 x_1 + \cdots + a_p x_p : x \in \mathbb{R}^p\} = \operatorname{span}\{a_1, \dots, a_p\}.$$

Tale spazio si chiama *spazio delle colonne di* A e si indica a volte con $\mathcal{C}(A)$. Ma allora si può osservare che, assegnato un dato $b \in \mathbb{R}^n$, il sistema $Ax = b$ ammette soluzione se e solo se $b \in \operatorname{span}\{a_1, \dots, a_p\}$. Cioè se b giace nello spazio delle colonne di A.

Se succede che il sistema non ha soluzione, allora introduciamo una nozione di soluzione approssimata che chiameremo soluzione ai minimi quadrati. Tale (o tali) soluzioni si ottengono sostituendo il "bersaglio irraggiungibile" $b \notin \operatorname{span}\{a_1, \dots, a_p\}$ con la sua proiezione ortogonale sullo spazio delle colonne, cioè scegliondo quel valore di x che rende il vettore Ax il piú vicino possibile a b.

Ecco una definizione formale

Definizione 4.14 (soluzione ai minimi quadrati)**.** *Consideriamo il sistema* $Ax = b$, *dove* $A = [a_1, \dots, a_p] \in \mathbb{R}^{n \times p}$ *e* $b \in \mathbb{R}^n$. *Un vettore* $x^* \in \mathbb{R}^p$ *è soluzione ai minimi quadrati del sistema in questione se risulta*

$$Ax^* = \pi_V b,$$

dove $\pi_V b$ *è la proiezione ortogonale del vettore* b *sul sottospazio* $V = \operatorname{span}\{a_1, \dots, a_p\}$.

Osservazione 4.15. *Osserviamo le seguenti cose riguardanti la nozione di soluzione ai minimi quadrati di un sistama* $Ax = b$:

(1) *Se* $b \in \operatorname{span}\{a_1, \dots, a_p\}$, *allora il sistema* $Ax = b$ *ha soluzione. Ma in tal caso vale* $\pi_V b = b$. *Quindi parlare di soluzione ai minimi quadrati è lo stesso che parlare di soluzione esatta.*

(2) *Per le proprietà (ampiamente discusse) delle proiezioni, se* x^* *è una soluzione ai minimi quadrati di un sistema* $Ax = b$, *allora:*

- *il "bersaglio"* Ax^* *è, tra tutti i punti di* $\operatorname{span}\{a_1, \dots, a_p\}$, *quello piú vicino a* b *(tradotto in formule*
$$\|b - Ax^*\| \le \|b - Ax\| \quad \text{per ogni } x \in \mathbb{R}^p);$$
- *il vettore* Ax^* *è l'unico vettore in* $V = \operatorname{span}\{a_1, \dots, a_p\}$ *per il quale vale la proprietà di ortogonalità*
$$(b - Ax^*) \perp V \tag{4.4}$$

Esempio 4.16. *Guardiamo il sistema*

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

Se $b_3 \neq 0$ *il sistema è impossibile. Ma* $\pi_V b = \begin{bmatrix} b_1 \\ b_2 \\ 0 \end{bmatrix}$. *Dunque la soluzione ai minimi quadrati è* $x_1 = b_1$, $x_2 = b_2$.

Per trovare la soluzione ai minimi quadrati di un sistema si usa il seguente teorema:

Teorema 4.17 (Teorema sulle soluzioni ai minimi quadrati). *Consideriamo il sistema*

$$Ax = b \tag{4.5}$$

, dove $A = [a_1, \dots, a_p] \in \mathbb{R}^{n\times p}$ *e* $b \in \mathbb{R}^n$. *Allora:*

(i) esiste sempre una soluzione ai minimi quadrati

(ii) un vettore $x^* \in \mathbb{R}^p$ *è soluzione ai minimi quadrati di* (4.5) *se e solo se esso soddisfa*

$$A^T A x^* = A^T b \tag{4.6}$$

(iii) la soluzione di (4.6) *è unica se le colonne di* A *sono indipendenti.*

Le equazioni (4.6) si chiamano *equazioni normali* associate al sistema $Ax = b$.

Osservazione 4.18 (Non discussa in classe). *Nel caso in cui le equazioni normali abbiano piú di una soluzione, cioè – come dice il punto (iii) – se le colonne di* A *sono dipendenti, si sceglie come soluzione ai minimi quadrati quella di norma minima tra tutte le soluzioni di* (4.6).

Dimostrazione. La parte (i) è ovvia, perché abbiamo già fatto, studiando le proiezioni, la fatica di dimostrare che la proiezione di b su un sottospazio V esiste (e ce n'è una sola).

La parte (ii) segue dal seguente fatto: sappiamo che x^* è soluzione ai minimi quadrati se e solo se $Ax^* = \pi_V b$. Ma, poiché $\pi_V b$ è l'unico punto di V che soddisfa $b - \pi_V b \perp V$, ciò equivale alla condizione

$$(b - Ax^*) \perp V \quad \Leftrightarrow \quad (b - Ax^*) \perp \operatorname{span}\{a_1, \dots, a_p\}$$

$$\Leftrightarrow \quad \begin{cases} \langle a_1, b - Ax^* \rangle = 0 \\ \vdots \\ \langle a_p, b - Ax^* \rangle = 0 \end{cases} \quad \Leftrightarrow \quad \begin{cases} a_1^T (b - Ax^*) = 0 \\ \vdots \\ a_p^T (b - Ax^*) = 0 \end{cases} \quad \Leftrightarrow \quad A^T(b - Ax^*) = 0,$$

che è esattamente (4.6), cioè il sistema delle equazioni normali. [7]

Il punto (iii) non è stato discusso in classe.

Per gli studenti interessati: supponiamo che le colonne di A sono indipendenti e mostriamo che $A^T A x = A^T b$ ha soluzione unica, comunque sia il dato b. Se ci fosse piú di una soluzione, ad esempio due soluzioni x e $x' \in \mathbb{R}^p$, risulterebbe $Ax = Ax' = b$ e quindi $A(x - x') = 0$. Di conseguenza, $A^T A(x - x') = 0$. Ma allora

$$\langle A^T A(x - x'), (x - x') \rangle = 0 \quad \Rightarrow \quad \langle A(x - x'), A(x - x') \rangle = \|A(x - x')\|^2 = 0 \quad \Rightarrow \quad A(x - x') = 0.$$

Pertanto $x - x' = 0$, perché le colonne di A sono indipendenti. Nei passaggi precedenti abbiamo usato la proprietà (3.3) della matrice trasposta. □

4.6. Esercizi per casa

(a) Sia

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

Risolvere il sistema $Ax = b$ con $b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \in \mathbb{R}^{3\times 1}$ dato assegnato. Indicare con u, v e $w \in \mathbb{R}^{3\times 1}$ i tre vettori soluzione corrispondenti ai dati $b = e_1$, $b = e_2$ e $b = e_3$. Costruire la matrice

$$M = [u, v, w]$$

[7] Di fatto il calcolo appena effettuato dimostra la seguente proprietà di ortogonalità: se A è una matrice qualsiasi e $\mathcal{C}(A)$ e $\mathcal{N}(A)$ indicano rispettivamente lo spazio delle colonne e lo spazio nullo di A, allora vale

$$\mathcal{C}(A)^\perp = \mathcal{N}(A^T). \tag{4.7}$$

che ha come colonne i tre vettori trovati. Verificare infine che $AM = I_3$ e che $MA = I_3$. Cioè che M è l'inversa di A.

(b) Con la stessa tecnica dell'esercizio precedente, trovare l'inversa della matrice

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

(c) Risolvere ai minimi quadrati il sistema inconsistente

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

(d) Verificare se i vettori $(1,2,1,3)$, $(0,1,0,1)$, $(1,1,3,2)$ e $(-1,2,3,1)$ sono dipendenti o indipendenti. (Si può usare la Proposizione 4.12).

(e) Dire, trovandone una base, quanto vale la dimensione del sottospazio

$$V = \operatorname{span}\{(1,2,1,1),(3,1,-1,0),(1,1,q,p)\}$$

a seconda dei valori assunti dai parametri p e q. (Si può usare l'Osservazione 4.9).

(f) Consideriamo il sistema

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

Trovare, se esistono, le soluzioni. Se non esistono soluzioni trovare la soluzione ai minimi quadrati.

(g) Risolvere il sistema omogeneo

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

Trovare una base e la dimensione del sottospazio costituito dalle soluzioni di tale sistema.

4.7. Sistemi quadrati e matrice inversa

Facciamo ora un paio di considerazioni su sistemi lineari nei quali ci sia un numero di equazioni pari al numero di incognite. In tal caso la matrice $A \in \mathbb{R}^{n\times n}$ è quadrata. (xiii)

Teorema 4.19. *Se $A \in \mathbb{R}^{n\times n}$ è invertibile, allora ogni sistema lineare*

$$Ax = b$$

ammette un'unica soluzione

Dimostrazione. Supponiamo innanzitutto che ci sia una soluzioe x di $Ax = b$. Allora, moltiplicando per A^{-1} a sinistra otteniamo

$$A^{-1}Ax = A^{-1}b \quad \Rightarrow \quad x = A^{-1}b.$$

Quindi, se c'è una soluzione, essa deve avere la forma $A^{-1}b$. Ora verifichiamo che il vettore $A^{-1}b$ è soluzione. Basta sostituire tale vettore nel sistema: $A(A^{-1}b) = I_n b = b$, come desiderato. Dunque $x = A^{-1}b$ è l'unica soluzione del sistema. □

Vale anche il viceversa del teorema precedente.

Teorema 4.20. *Sia $A \in \mathbb{R}^{n\times n}$ e assumiamo che tutti i sistemi lineari $Ax = b$ ammettano almeno una soluzione. Allora A è invertibile. Cioè esiste $U \in \mathbb{R}^{n\times n}$ tale che $AU = UA = I_n$.*

Dimostrazione. Per ogni $j = 1, \ldots, n$, indichiamo con $u_j \in \mathbb{R}^n$ una soluzione del sistema lineare $Au_j = e_j$. Tale soluzione esiste per ipotesi. Allora la matrice $U = [u_1, \ldots, u_n]$ soddisfa

$$AU = A[u_1, \ldots, u_n] = [Au_1, \ldots, Au_n] = [e_1, \ldots, e_n] = I_n.$$

Questo prova metà della tesi affermata.

Non dimostriamo il fatto che la matrice U soddisfa anche $UA = I_n$. I passi di tale dimostrazione sono schematizzati sotto per gli eventuali studenti interessati.

Passo 1. Poiché per ipotesi il sistema $Ax = b$ ha soluzione per ogni $b \in \mathbb{R}^n$, lo spazio delle colonne di A è tutto $\mathbb{R}^n$: cioè $\mathcal{C}(A) = \mathbb{R}^n$. Dunque le colonne $a_1, \ldots, a_n$ di A sono indipendenti.

Passo 2. Verifichiamo che il sistema omogeneo $Ax = 0$ ha solo la soluzione banale. Infatti, preso un qualsiasi vettore $x = \sum_{j=1}^n x_j e_j \in \mathbb{R}^n$, se vale

$$A \sum x_j e_j = \sum x_j A e_j = \sum x_j a_j = 0,$$

allora $x_j = 0$ per ogni j, perché le colonne di A sono indipendenti. Questo, in formule, significa che $\mathcal{N}(A) = (0)$.

Passo 3. Poiché $\mathcal{N}(A)$ è il sottospazio nullo, allora $(\mathcal{N}(A))^\perp = \mathbb{R}^n$. Se applichiamo ora la formula (4.7) alla matrice A^T, risulta $\mathcal{C}(A^T) = (\mathcal{N}(A))^\perp = \mathbb{R}^n$. Quindi, il sistema $A^T y = c$ ha soluzione per ogni dato $c \in \mathbb{R}^n$. e dunque esiste $V \in \mathbb{R}^{n\times n}$ tale che $A^T V = I$. QUindi $V^T A = I_n$. La matrice V^T si comporta come inversa di A dal lato sinistro.

Passo 4. Abbiamo provato che esistono due matrici quadrate U, V tali cha $AU = V^T A = I_n$. Per l'unicità dell'inversa (Osservazione 3.11) questo prova che $V^T = U = A^{-1}$. □

5. Determinanti

Il determinante di una matrice $A = [a] \in \mathbb{R}^{1\times 1}$ è il numero $\det[a] := a \in \mathbb{R}$.

5.1. Determinanti di matrici 2×2 e regola di Cramer

Consideriamo una matrice

$$A = [a, b] = \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}$$

dove a e $b \in \mathbb{R}^{2\times 1}$ sono le colonne. Definiamo

$$\det(A) := a_1 b_2 - a_2 b_1$$

Notiamo subito le due proprietà immediate:

$$\det(I_2) = 1$$
$$\det(A^T) = \det(A)$$

Ora osserviamo alcune ulteriori proprietà significative del determinante che prendono una forma chiara se parametrizziamo la matrice attraverso le sue due colonne:

$$A = [a, b] \qquad \text{con } a \text{ e } b \in \mathbb{R}^{2\times 1}.$$

Proposizione 5.1 (proprietà del determinante come funzione delle colonne)**.** *Osserviamo il determinante come funzione delle colonne. Esso soddisfa per ogni $a, a', a'', b \in \mathbb{R}^2$ e $\lambda \in \mathbb{R}$:*

(1) $\det[a' + a'', b] = \det[a', b] + \det[a'', b]$
(2) $\det[\lambda a, b] = \lambda \det[a, b]$
(3) $\det[b, a] = -\det[a, b]$

Le proprietà (1) e (2) assieme esprimono la *linearità rispetto al primo argomento.* La proprietà (3) si chiama *proprietà antisimmetrica.*

Dalle proprietà precedenti si deducone subito che

$$\det[a, a] = 0 \quad \text{e} \tag{5.1}$$

$$\det[a, \lambda' b' + \lambda'' b''] = \lambda' \det[a, b'] + \lambda'' \det[a, b''], \tag{5.2}$$

per ogni $a, b', b'' \in \mathbb{R}^2$ e $\lambda', \lambda'' \in \mathbb{R}$. Dunque vale la linearità nel secondo argomento.

Consideriamo il sistema di due equazioni in due incognite

$$ax + by = p \tag{5.3}$$

dove $a, b, p \in \mathbb{R}^{2\times 1}$. Discuteremo nelle prossime lezioni il seguente teorema:

Teorema 5.2. *Se $\det[a, b] \neq 0$, allora, comunque sia il dato $p \in \mathbb{R}^2$, il sistema* (5.3) *ha un'unica soluzione.*

La regola di Cramer permette di esprimere la soluzione del sistema (5.3) attraverso una formula esplicita.

Teorema 5.3 (regola di Cramer)**.** *Se $[a, b] \in \mathbb{R}^{2\times 2}$ e se $\det[a, b] \neq 0$, allora, dato $p \in \mathbb{R}^2$, la soluzione del sistema $ax + by = p$ è*

$$x = \frac{\det[p.b]}{\det[a, b]} \qquad y = \frac{\det[a, p]}{\det[a, b]}. \tag{5.4}$$

Dimostrazione. Siano x e y i due numeri che rendono vera l'uguaglianza $ax + by = p$. Allora, se usiamo la linearità del determinante rispetto alla prima colonna, troviamo

$$\det[p, b] = \det[ax + by, p] = \det[ax, p] + \det[by, p] = x \det[a, p] + y \det[b, b] = x \det[a, p].$$

Abbiamo anche usato il fatto che il determinante di una matrice con due colonne uguali è nullo. Confrontando il primo e l'ultimo termine otteniamo

$$x = \frac{\det[p.b]}{\det[a, b]},$$

come desideravamo. Il valore di y si ricava con lo stesso ragionamento, partendo dall'uguaglianza $\det[a, p] = \det[a, ax + by]$ e utilizzando la linearità del determinante nel secondo argomento. □

Esempio 5.4 (in classe)**.** *Risolvere con la regola di Cramer il sistema*

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

Esempio 5.5 (Svolto in classe)**.** *Presa una matrice con determinante non nullo $A = \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}$, scrivere la forma esplicita della soluzione del sistema*

$$\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} p_1 \\ p_2 \end{bmatrix}$$

con dato p arbitrario. Scegliendo le soluzioni u e v corrispondenti rispettivamente a $p = e_1$ e $p = e_2$ scrivere la forma della matrice inversa di A. Fatti i conti si ottiene

$$A^{-1} = \frac{1}{a_1 b_2 - a_2 b_1} \begin{bmatrix} b_2 & -b_1 \\ -a_2 & a_1 \end{bmatrix} \tag{5.5}$$

5.2. Determinanti di matrici 3×3

Consideriamo la matrice (xiv)

$$A = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} \in \mathbb{R}^{3\times 3}.$$

Definiamo il determinante di A come il risultato comune di uno qualsiasi delle seguenti espressioni

$$\begin{aligned}
\det A &= a_1 \det \begin{bmatrix} b_2 & c_2 \\ b_3 & c_3 \end{bmatrix} - a_2 \det \begin{bmatrix} b_1 & c_1 \\ b_3 & c_3 \end{bmatrix} + a_3 \det \begin{bmatrix} b_1 & c_1 \\ b_2 & c_2 \end{bmatrix} \\
&= -b_1 \det \begin{bmatrix} a_2 & c_2 \\ a_3 & c_3 \end{bmatrix} + b_2 \det \begin{bmatrix} a_1 & c_1 \\ a_3 & c_3 \end{bmatrix} - b_3 \det \begin{bmatrix} a_1 & c_1 \\ a_2 & c_2 \end{bmatrix} \\
&= c_1 \det \begin{bmatrix} a_2 & b_2 \\ a_3 & b_3 \end{bmatrix} - c_2 \det \begin{bmatrix} a_1 & b_1 \\ a_3 & b_3 \end{bmatrix} + c_3 \det \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}
\end{aligned}$$

Tutte queste espressioni coincidono e producono il risultato:

$$\det A = a_1 b_2 c_3 + b_1 c_2 a_3 + c_1 a_2 b_3 - c_1 b_2 a_2 - b_1 a_2 c_3 - a_1 c_2 b_3$$

Le tre diverse scritture esibite si chiamano *sviluppi in cofattori* o *sviluppi di Laplace*, rispetto alla prima, seconda e terza colonna, del determinante di A.

Esempio 5.6. *Calcolo di*

$$\det \begin{bmatrix} 1 & 2 & -1 \\ 1 & 0 & 0 \\ 2 & -1 & 3 \end{bmatrix}.$$

Ovviamente vale $\det(I_3) = 1$. È anche immediato verificare (fatto in classe) che

$$\det(A^T) = \det A$$

Alla luce di tale uguaglianza ciascun sviluppo in cofattori rispetto alle colonne della matrice A^T diventa uno sviluppo rispetto alle righe di A. Abbiamo dunque:

$$\begin{aligned}
\det(A) &= a_1 \det \begin{bmatrix} b_2 & c_2 \\ b_3 & c_3 \end{bmatrix} - b_1 \det \begin{bmatrix} a_2 & c_2 \\ a_3 & c_3 \end{bmatrix} + c_1 \det \begin{bmatrix} a_2 & b_2 \\ a_3 & b_3 \end{bmatrix} \\
&= -a_2 \det \begin{bmatrix} b_1 & c_1 \\ b_3 & c_3 \end{bmatrix} + b_2 \det \begin{bmatrix} a_1 & c_1 \\ a_3 & c_3 \end{bmatrix} - c_2 \det \begin{bmatrix} a_1 & b_1 \\ a_3 & b_3 \end{bmatrix}. \\
&= a_3 \det \begin{bmatrix} b_1 & c_1 \\ b_2 & c_2 \end{bmatrix} - b_3 \det \begin{bmatrix} a_1 & c_1 \\ a_2 & c_2 \end{bmatrix} + c_3 \det \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}
\end{aligned}$$

Quelli appena scritti sono i tre *sviluppi in cofattori* rispetto alla prima, seconda e terza riga della matrice A. Osserviamo che ogni sviluppo in cofattori rispetta la seguente “scacchiera” di segni

$$\begin{bmatrix} + & - & + \\ - & + & - \\ + & - & + \end{bmatrix}.$$

Le proprietà fondamentali del determinante come funzione delle colonne sono le seguenti

Proposizione 5.7. *Per ogni* $a, b, c \in \mathbb{R}^{3\times 1}$ *e* $\lambda \in \mathbb{R}$, *risulta*

$$\det[\lambda a, b, c] = \lambda \det[a, b, c] \tag{5.6}$$
$$\det[a + a', b, c] = \det[a, b, c] + \det[a', b, c] \tag{5.7}$$
$$\det[a, b, c] = -\det[b, a, c] = \det[c, a, b] = -\det[c, b, a]. \tag{5.8}$$

Dimostrazione. La verifica della proposizione è un semplice calcolo. □

Osservazione 5.8. *(1) Usando le priprietà di linearità rispetto al primo argomento e di antisimmetria, è facile vedere che la linearità vale anche in tutti gli altri argomenti. Ad esempio*

$$\det[a, \lambda b + \lambda' b', c] = \lambda \det[a, b, c] + \lambda' \det[a, b', c]$$

per ogni $a, b, b', c \in \mathbb{R}^3$ *e* $\lambda\lambda' \in \mathbb{R}$. *Lo stesso vale rispetto al terzo argomento.*

(2) Se due colonne di A sono uguali, allora il determinante è zero. Infatti, usando l'antisimmetria,

$$\det[a, a, b] = -\det[a, a, b] \quad \Rightarrow \quad \det[a, a, b] = 0.$$

Lo stesso vale se altre due colonne sono uguali.

(3) Se le colonne $a, b, c \in \mathbb{R}^3$ *sono dipendenti, allora* $\det[a, b, c] = 0$. *Infatti, se i tre vettori colonna sono dipendenti, uno di essi è combinazione degli altri. Assumiamo ad esempio che sia* $c = \lambda a + \mu b$, *per* $\lambda, \mu \in \mathbb{R}$ *opportuni. Allora sarà*

$$\det[a, b, c] = \det[a, b, \lambda a + \mu b] = \lambda \det[a, b, a] + \mu \det[a, b, b] = 0.$$

Dalle proprietà appena discusse, si ricava la regola di Cramer. Assumiamo il fatto seguente (vedere discussione nella prossima lezione): se $\det A \neq 0$, allora la soluzione di $Ax = b$ esiste ed è unica per ogni $b \in \mathbb{R}^3$.

Teorema 5.9 (Regola di Cramer). *Se* $A = [a_1, a_2, a_3] \in \mathbb{R}^{3\times 3}$ *e se* $\det(A) \neq 0$, *allora per ogni dato* $p \in \mathbb{R}^3$ *la soluzione del sistema* $Ax = p$ *ha le seguenti componenti:*

$$x_1 = \frac{\det[p, a_2, a_3]}{\det[a_1, a_2, a_3]} \quad x_2 = \frac{\det[a_1, p, a_3]}{\det[a_1, a_2, a_3]} \quad x_3 = \frac{\det[a_1, a_2, p]}{\det[a_1, a_2, a_3]}$$

La dimostrazione della regola di Cramer è analoga a quella della regola di Cramer in dimensione due e viene lasciata al lettore come esercizio.

Esempio 5.10. *(In classe.) Risolvere con il metodo di Cramer il sistema*

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

Esempio 5.11. *Determinante di matrici triangolari superiori: vale*

$$\det \begin{bmatrix} a_1 & b_1 & c_1 \\ 0 & b_2 & c_2 \\ 0 & 0 & c_3 \end{bmatrix} = a_1 b_2 c_3. \tag{5.9}$$

Osserviamo tale determinante è non nullo se e solo se tutti i tre elementi sulla diagonale sono diversi da zero.

5.3. Determinanti e proprietà dei sistemi lineari

Il nostro proposito ora è quello di dimostrare il seguente teorema.

Teorema 5.12. *Sia $A \in \mathbb{R}^{3\times 3}$. Allora:*
- *se $\det A \neq 0$, allora per ogni scelta di $b \in \mathbb{R}^3$ il sistema $Ax = b$ è determinato;*
- *se $\det A = 0$, allora il sistema $Ax = b$ non è mai determinato; è invece indeterminato o impossibile, a seconda della scelta di b.*

Preliminarmente alla dimostrazione di questo teorema, osserviamo che poiché $\det A = \det A^T$, allora le proprietà di linearità e antisimmetria valgono anche rispetto alle righe: precisamente,

$$\det \begin{bmatrix} \lambda r_1 + \lambda' r_1' \\ r_2 \\ r_3 \end{bmatrix} = \lambda \det \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} + \lambda' \det \begin{bmatrix} r_1' \\ r_2 \\ r_3 \end{bmatrix} \tag{5.10}$$

per ogni $r_1, r_1', r_2, r_3 \in \mathbb{R}^{1\times 3}$, vettori riga e $\lambda, \lambda' \in \mathbb{R}$. Analoghe proprietà di linearità valgono rispetto alle altre righe. Inoltre

$$\det \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} = -\det \begin{bmatrix} r_2 \\ r_1 \\ r_3 \end{bmatrix} = \det \begin{bmatrix} r_3 \\ r_1 \\ r_2 \end{bmatrix} = -\det \begin{bmatrix} r_3 \\ r_2 \\ r_1 \end{bmatrix}. \tag{5.11}$$

Cioè il determinante cambia segno, a fronte di uno scambio di due righe.

Ricordiamo che le operazioni elementari per trasformare un sistema qualunque in un sistema a scala sono le seguenti (vedere pagina 24):

(i) Moltiplicare una riga per un numero $\lambda \neq 0$.
(ii) Scambiare due righe tra loro
(iii) Prese due righe R_k ed R_j e un numero $\lambda \in \mathbb{R}$, sostituire la riga k-esima, R_k, con la riga $R_k - \lambda R_j$.

Quello che dicono le proprietà (5.10) e (5.11) è che, se applichiamo a un sistema quadrato un'operazione elementare

$$\begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \quad \longrightarrow \quad \text{operazione (i), (ii) o (iii)} \quad \longrightarrow \quad \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$$

allora vale

$$\det \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} \neq 0 \quad \Leftrightarrow \quad \det \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} \neq 0$$

In particolare: se applichiamo (i), allora il determinante cambia segno; se applichiamo (ii), allora il determinante viene moltiplicato per $\lambda \neq 0$; se applichiamo (iii), allora il determinante resta invariato. Verifichiamo quest'ultima affermazione, supponendo ad esempio di aver sostituito r_2 con $r_2 + \lambda r_1$: allora

$$\det \begin{bmatrix} r_1 \\ r_2 + \lambda r_1 \\ r_3 \end{bmatrix} = \det \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} + \lambda \det \begin{bmatrix} r_1 \\ r_1 \\ r_3 \end{bmatrix} = \det \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix}.$$

In definitiva, applicando l'algoritmo di Gauss trasformiamo il sistema $Ax = b$ in un nuovo sistema

$$Sx = c \in \mathbb{R}^3 \tag{5.12}$$

con le seguenti proprietà:

- Il sistema finale (5.12) è a scala;
- $x \in \mathbb{R}^3$ è soluzione del sistema (5.12) se e solo se risolve il sistema iniziale;
- $\det A \neq 0$ se e solo se $\det S \neq 0$.

Alla luce di queste considerazioni, risulta chiaro che è sufficiente fare la dimostrazione del Teorema 5.18 per un sistema a scala.

Dimostrazione del teorema 5.18. Il generico sistema a scala quadrato in tre variabili ha matrice incompleta di una delle forme seguenti:

$$\begin{bmatrix} a & * & * \\ 0 & b & * \\ 0 & 0 & c \end{bmatrix}, \quad \begin{bmatrix} a & * & * \\ 0 & b & * \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} a & * & * \\ 0 & 0 & b \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} a & * & * \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$\begin{bmatrix} 0 & a & * \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & a \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

dove a, b, c sono numeri non nulli (i pivot), mentre i caratteri jolly sono numeri reali qualsiasi. Solo la prima delle sette matrici modello scritte ha determinante non zero (matrice non singolare). Le altre hanno tutte determinante nullo (matrici singolari).

È immediato vedere che il sistema $\left[\begin{smallmatrix} a & * & * \\ 0 & b & * \\ 0 & 0 & c \end{smallmatrix}\right] \left[\begin{smallmatrix} x \\ y \\ z \end{smallmatrix}\right] = \left[\begin{smallmatrix} p \\ q \\ r \end{smallmatrix}\right]$ ha una soluzione unica. Questo si vede per sostituzione all'indietro usando il fatto che i tre pivot a, b, c sono non nulli.

Guardiamo ora uno degli altri sei sistemi modello: ad esempio

$$\begin{bmatrix} a & * & * \\ 0 & b & * \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} p \\ q \\ r \end{bmatrix}.$$

Ci sono due possibilità: se $r \neq 0$, allora il sistema non ha nessuna soluzione. Se invece $r = 0$, allora nella seconda equazione possiamo ricavare la y in termini della z e, sostituendo all'indietro, anche la x in termini della z. In definitiva ci sono infinite soluzioni. Abbiamo dunque scoperto che il sistema è impossibile o indeterminato. Non è determinato per nessuna scelta del dato nel membro di destra.

Una discussione analoga vale per tutti gli altri casi, ma viene lasciata allo studente. □

Osserviamo che, sebbene la dimostrazione sia stata presentata in dimensione tre, il Teorema 5.18 vale anche per sistemi $n \times n$.

Osservazione 5.13. *Notiamo che il determinante di una somma di matrici* non *è la somma dei determinanti. Ad esempio, per una matrice* 3×3 *scritta nella forma* $A = [a, b, c]$*, abbiamo* (xv)

$$\det(2A) = \det[2a, 2b, 2c] = 2^3 \det[a, b, c] = 8 \det A.$$

Quindi in particolare ***non è vero*** *che* $\det(A + B) = \det(A) + \det(B)$.

Il determinante si comporta bene invece rispetto al prodotto. Vale la formula di Binet–Cauchy

$$\det(AB) = \det(A)\det(B) \tag{5.13}$$

per ogni coppia A, B *di matrici quadrate.*

5.4. Una formula per la matrice inversa

Il metodo piú semplice per trovare la matrice inversa di una assegnata matrice A utilizza l'algoritmo di Gauss (o qualche suo raffinamento). Ora discutiamo brevemente una scrittura esplicita della matrice inversa di una matrice 3×3 che utilizza la regola di Cramer. La discussione che segue si generalizza al caso $n \times n$.

Partiamo da una matrice generica 3×3 della forma

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

La *matrice dei cofattori* di A è la seguente matrice 3×3

$$\operatorname{cof} A := \begin{bmatrix} \det \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} & -\det \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} & \det \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \\ -\det \begin{bmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{bmatrix} & \det \begin{bmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{bmatrix} & -\det \begin{bmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{bmatrix} \\ \det \begin{bmatrix} a_{12} & a_{13} \\ a_{22} & a_{33} \end{bmatrix} & -\det \begin{bmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{bmatrix} & \det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \end{bmatrix} \tag{5.14}$$

Esempio 5.14 (in classe). *Calcolo di*

$$\operatorname{cof} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & -1 \\ 1 & 1 & 0 \end{bmatrix} = \det \begin{bmatrix} \det \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} & -\det \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} & \det \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \\ -\det \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix} & \det \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} & -\det \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \\ \det \begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix} & -\det \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix} & \det \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 1 & -1 & -1 \\ 1 & -1 & 1 \\ -3 & 1 & 1 \end{bmatrix}.$$

Tornando alla matrice dei cofattori scritta in (5.14), osserviamo, ad esempio, che

$$(\operatorname{cof} A)_{12} = -\det \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} = \det \begin{bmatrix} a_{11} & 1 & a_{13} \\ a_{21} & 0 & a_{23} \\ a_{31} & 0 & a_{33} \end{bmatrix} = \det[a_1, e_1, a_3],$$

cioè il determinante della matrice che si ottiene sostituendo nella seconda colonna il primo versore coordinato. Qui abbiamo parametrizzato la matrice tramite le colonne a_1, a_2, a_3. Notiamo che davanti al determinante 3×3 non c'è il segno meno.

In generale, l'elemento $(\operatorname{cof} A)_{jk}$ si ottiene come determinante della matrice 3×3 ottenuta sostituendo la colonna k-esima di A con il versore j-esimo:

$$(\operatorname{cof} A)_{jk} = \det[\dots, a_{k-1}, e_j, a_{k+1}, \dots].$$

Questo vale anche per matrici di dimensione qualsiasi. Piú esplicitamente risulta, per $A = [a_1, a_2, a_3] \in \mathbb{R}^{3\times 3}$:

$$\operatorname{cof} A = \begin{bmatrix} \det[e_1, a_2, a_3] & \det[a_1, e_1, a_3] & \det[a_1, a_2, e_1] \\ \det[e_2, a_2, a_3] & \det[a_1, e_2, a_3] & \det[a_1, a_2, e_2] \\ \det[e_3, a_2, a_3] & \det[a_1, e_3, a_3] & \det[a_1, a_2, e_3] \end{bmatrix}. \tag{5.15}$$

Ora consideriamo il sistema $Ax = b$ associato alla matrice $A = [a_1, a_2, a_3]$ con $b \in \mathbb{R}^3$ dato qualsiasi. con $b \in \mathbb{R}^3$ dato assegnato e $\det[a_1, a_2, a_3] \neq 0$. Se la matrice A ha determinante non nullo, allora la soluzione di tale sistema può essere espressa come segue tramite la regola di Cramer:

$$x = \frac{1}{\det A}\begin{bmatrix}\det[b, a_2, a_3]\\ \det[a_1, b, a_3]\\ \det[a_1, a_2, b]\end{bmatrix}.$$

Sappiamo anche (Esercizi (a) e (b) a pag. 28 e dimostrazione del Teorema 4.20) che se consideriamo le tre particolari soluzioni $u, v, w \in \mathbb{R}^3$ dei sistemi $Au = e_1$, $Av = e_2$ e $Aw = e_3$, allora la matrice $[u, v, w] \in \mathbb{R}^{3\times 3}$ è l'inversa di A. Scriviamo tale inversa:

$$A^{-1} = \frac{1}{\det A}[u, v, w] = \frac{1}{\det A}\begin{bmatrix}\det[e_1, a_2, a_3] & \det[e_2, a_2, a_3] & \det[e_3, a_2, a_3]\\ \det[a_1, e_1, a_3] & \det[a_1, e_2, a_3] & \det[a_1, e_3, a_3]\\ \det[a_1, a_2, e_1] & \det[a_1, a_2.e_2] & \det[a_1, a_2, e_3]\end{bmatrix}.$$

La matrice appena a scritta contiene la trasposta di $\operatorname{cof} A$. Dunuque:

Teorema 5.15 (formula per l'inversa). *Se A è quadrata e $\det A \neq 0$, allora*

$$A^{-1} = \frac{1}{\det A}(\operatorname{cof} A)^T. \tag{5.16}$$

Esempio 5.16. *Calcolare l'inversa di $A = \begin{bmatrix}1 & 2 & 1\\ 0 & 1 & -1\\ 1 & 1 & 0\end{bmatrix}$. La matrice A è quella dell'Esempio 5.14. Prendiamo da tale esempio il calcolo di* $\operatorname{cof} A$ *Risulta*

$$\det A = 1\det\begin{bmatrix}1 & -1\\ 1 & 0\end{bmatrix} + 1\det\begin{bmatrix}2 & 1\\ 1 & -1\end{bmatrix} = -2$$

Quindi

$$A^{-1} = \frac{1}{\det A}(\operatorname{cof} A)^T = -\frac{1}{2}\begin{bmatrix}1 & -1 & -1\\ 1 & -1 & 1\\ -3 & 1 & 1\end{bmatrix}^T = \cdots = \begin{bmatrix}-\frac{1}{2} & -\frac{1}{2} & \frac{3}{2}\\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2}\\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2}\end{bmatrix}$$

Verificare la correttezza del conto calcolando per esercizio AA^{-1} e $A^{-1}A$.

5.5. Determinanti di matrici $n \times n$ e loro proprietà

Non è difficile indovinare come potrà essere definito il determinante di una matrice quadrata di taglia piú grande di 3. Consideriamo $A \in \mathbb{R}^{n\times n}$,

$$A = \begin{bmatrix}a_{11} & \dots & a_{1n}\\ \vdots & \ddots & \vdots\\ a_{n1} & \dots & a_{nn}\end{bmatrix}$$

e definiamo il determinante di una matrice $n \times n$ come il risultato di uno degli sviluppi equivalenti:

$$\det A = \sum_{k=1}^{n} a_{jk}(\operatorname{cof} A)_{jk} = \sum_{k=1}^{n} a_{kj}(\operatorname{cof} A)_{kj}$$

con $j = 1, \dots, n$. Ci sono dunque $2n$ sviluppi in cofattori (o sviluppi di Laplace): n per righe e altri n per colonne. Qui il numero $(\operatorname{cof} A)_{jk}$ si ottiene moltiplicando $(-1)^{j+k}$ per il determinante della matrice ottenuta cancellando la j-esima riga e la k-esima colonna dalla matrice A.

Ad esempio

$$\det \begin{bmatrix} 1 & 5 & 9 & 13 \\ 2 & 6 & 10 & 14 \\ 3 & 7 & 11 & 15 \\ 4 & 8 & 12 & 16 \end{bmatrix} = -5 \det \begin{bmatrix} 2 & 10 & 14 \\ 3 & 11 & 15 \\ 4 & 12 & 16 \end{bmatrix} + 6 \det \begin{bmatrix} 1 & 9 & 13 \\ 3 & 11 & 15 \\ 4 & 12 & 16 \end{bmatrix}$$
$$-7 \det \begin{bmatrix} 1 & 9 & 13 \\ 2 & 10 & 14 \\ 4 & 12 & 16 \end{bmatrix} + 8 \det \begin{bmatrix} 1 & 9 & 13 \\ 2 & 10 & 14 \\ 3 & 11 & 15 \end{bmatrix} = \cdots$$

Descriviamo alcune proprietà del determinante seguendo la stessa linea seguita per i casi si dimensione bassa. Innanzitutto, un qualsiasi sviluppo di Laplace mostra che $\det(I_n) = 1$. Inoltre si simostra anche che

$$\det A^T = \det A.$$

La verifica di quest'affermazione non è però semplice come nei casi 2×2 e 3×3 già discussi.

Esempio 5.17. *Se T è triangolare superiore, cioè*

$$T = \begin{bmatrix} t_{11} & t_{12} & \dots & t_{1n} \\ 0 & t_{22} & \dots & t_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & t_{nn} \end{bmatrix} \tag{5.17}$$

allora vale

$$\det T = t_{11} t_{22} \cdots t_{nn}. \tag{5.18}$$

Ore elenchiamo le proprietà del determinante come funzione delle colonne. Scriviamo $A = [a_1, \dots, a_n] \in \mathbb{R}^{n \times n}$, con $a_1, \dots, a_n$ colonne di A. Allora si dimostra che valgono le seguenti proprietà:

(A) $\det[\lambda a_1 + \lambda' a'_1, a_2, , \dots, a_n] = \lambda \det[a_1, \dots, a_n] + \lambda' \det[a'_1, \dots, a_n]$, per ogni scelta di vettori $a_1, a'_1, a_2, \dots, a_n \in \mathbb{R}^n$ e $\lambda.\lambda' \in \mathbb{R}$

(B) Il determinante cambia segno se si scambiano due colonne di A

Unendo le proprietà (A) di linearità rispetto al primo argomento e (B) di antisimmetria, si deduce che vale la linearità rispetto a un qualsiasi altro argomento. Inoltre il determinante è zero se due colonne sono uguali. Piú generalmente,

$$\det A = 0 \quad \text{se le colonne di } A \text{ sono linearmente dipenedenti.}$$

Inoltre, il Teorema 5.18 vale per sistemi qualsiasi:

Teorema 5.18. *Sia $A \in \mathbb{R}^{n \times n}$. Allora:*

- *se $\det A \neq 0$, allora per ogni scelta di $b \in \mathbb{R}^n$ il sistema $Ax = b$ è determinato e (regola di Cramer), scritta $A = [a_1, \dots, a_n]$, la componente j-esima della soluzione ha la forma*

$$x_j = \frac{\det[a_1, \dots, a_{j-1}, b, a_{j+1}, \dots, a_n]}{\det(A)};$$

- *se $\det A = 0$, allora il sistema $Ax = b$ non è mai determinato; è indeterminato o impossibile, a seconda della scelta di b.*

Esempio 5.19 (svolto in classe)**.** *Calcolo dei coeficienti della retta di regressione relativa a una famiglia di punti del piano $(a_1, b_1), \dots, (a_n, b_n)$.*

6. Esercizi per casa

(a) Risolvere con l'algoritmo di Gauss e con la regola di Cramer i seguenti sistemi quadrati:

$$\begin{cases} z + 2y = p \\ 3x + 5y = q \end{cases} \quad \text{con } p, q \in \mathbb{R} \quad \text{e} \quad \begin{cases} x + 2y + z = 1 \\ -x + 3z = 2 \\ x + y = -1 \end{cases}$$

(b) Per quali $q \in \mathbb{R}$ il sistema

$$\begin{cases} x = -1 \\ 2x + y = 3 \\ -x + y = 0 \\ y = q \end{cases}$$

ha soluzione? Risolvere il sistema per tale valore e risolverlo ai minimi quadrati per $q = 0$.

(c) Trovare una base e la dimensione dello spazio delle soluzioni del sistema omogeneo

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

(d) Dire se sono linearmente dipendenti o no le famiglie di vettori

$$u_1 = (1, 2, 3),\ u_2 = (0, 1, 0),\ u_3 = (-2, 1, 0),\ u_4 = (0, 1, -1)$$

e

$$v_1 = (1, 1, 3, 1),\ v_2 = (0, 1, 2, 0),\ v_3 = (-1, 2, 3, 2),\ v_4 = (2, 0, 2, 3).$$

7. Avvertenze conclusive e istruzioni per il compito

1. Questo file non verrà piú modificato. In caso di necessità una eventuale lista di errata apparirà al link `http://www.dm.unibo.it/~morbidel/linear_algebra_errata_12_12.pdf`.

2. La prova scritta conterrà alcuni esercizi su entrambi i moduli del corso. **Non** si possono usare libri, appunti, né dispositivi elettronici.