Note del corso di
ALGEBRA LINEARE CON ELEMENTI DI GEOMETRIA

CL in Intelligenza Artificiale e Data Analytics

Dipartimento di Matematica e Geoscienze

Università degli Studi di Trieste

Prof. Valentina Beorchia

A.A. 2022/2023

9 ottobre 2022

Indice

Capitolo 1

Preliminari

1.1 Relazioni d'equivalenza

Per questa sezione si vedano anche le note del corso propedeutico (Prof. Del Santo).

Una **relazione** in un insieme X è una proprietà che una coppia ordinata di elementi di X può verificare o meno. Per esempio la relazione "$<$" "minore" ha senso negli insiemi numerici $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$; la relazione "$||$" "parallelo" ha senso nell'insieme delle rette del piano, o dei piani dello spazio.

In maniera più formale, una relazione in X è un sottinsieme R del prodotto cartesiano $X \times X$. In tal caso si dirà che x è in relazione R con y se la coppia ordinata $(x, y) \in R$. Si scrive anche xRy.

Per esempio la relazione $<$ in $\mathbb{Z}$ corrisponde al sottinsieme di $\mathbb{Z}\times\mathbb{Z}$: $\{(x, y) \mid x < y\}$. Analogamente la relazione $\leq$ corrisponde al sottinsieme di $\mathbb{Z} \times \mathbb{Z}$: $\{(x, y) \mid x \leq y\}$. La relazione di parallelismo nell'insieme delle rette del piano corrisponde alle coppie di rette (r, r') tali che r, r' sono distinte e parallele oppure sono uguali.

Simboli spesso usati per denotare relazioni sono $\equiv$, $\sim$, $\simeq$, $\cong$, ecc. Un altro esempio di relazione, in $\mathbb{R}$, è il seguente: $x \sim y$ se e solo se $x^2 = y$.

Noi saremo interessati a un tipo particolare di relazioni dette relazioni d'equivalenza.

Definizione 1.1.1 (**Relazione d'equivalenza**). Sia X un insieme e $\sim$ una relazione in X. Si dice che $\sim$ è una **relazione d'equivalenza** se valgono le tre proprietà:

1. riflessiva: per ogni $x \in X$ $x \sim x$;
2. simmetrica: se $x \sim y$ allora $y \sim x$;
3. transitiva: se $x \sim y$ e $y \sim z$ allora $x \sim z$.

Esempi 1.1.2.

1. L'uguaglianza è una relazione d'equivalenza in qualunque insieme X.
2. "Essere congruenti" è una relazione d'equivalenza nell'insieme dei triangoli del piano.
3. $\leq$, $<$ non sono relazioni d'equivalenza.

Il prossimo è un esempio fondamentale. Denotiamo con $\mathbb{N}$ l'insieme dei numeri naturali: $\mathbb{N} = \{0, 1, 2, 3 \ldots\}$.

Definizione 1.1.3 (**Congruenza modulo n**). Si fissi un naturale $n \in \mathbb{N}$. La relazione di congruenza modulo n è la relazione in $\mathbb{Z}$ così definita:

$$x \equiv y \pmod{n} \text{ se e solo se esiste } k \in \mathbb{Z} \text{ tale che } x - y = kn.$$

Si scrive anche $x \equiv_n y$. Si legge "x è congruo a y modulo n.

Proposizione 1.1.4. *La relazione di congruenza modulo n è una relazione d'equivalenza in $\mathbb{Z}$.*

Dimostrazione. 1. $x - x = 0x$ per ogni $x \in \mathbb{Z}$.

2. Se $x \equiv_n y$, si ha $x - y = kn$ per un opportno $k \in \mathbb{Z}$. Allora $y - x = (-k)n$.

3. Se $x \equiv_n y$ e $y \equiv_n z$, esistono $k, h \in \mathbb{Z}$ tali che $x - y = kn$, $y - z = hn$; ma allora $x - z = (x - y) + (y - z) = kn + hn = (k + h)n$, il che prova che $x \equiv_n z$. □

D'ora in poi supporremo sempre $n \geq 2$. La seguente osservazione è importante.

Proposizione 1.1.5. *$x \equiv_n y$ se e solo se x e y hanno lo stesso resto nella divisione per n.*

Dimostrazione. Infatti se x e y hanno lo stesso resto nella divisione per n, si ha: $x = qn+r$, $y = q'n+r$, dove $0 \leq r \leq n - 1$. Ma allora $x - y = (qn + r) - (q'n + r) = (q - q')n$ e perció $x \equiv_n y$.

Viceversa se $x \equiv_n y$, si ha $x = y + kn$. Se r è il resto della divisione di y per n, vale la relazione $y = qn + r$ con $0 \leq r \leq n - 1$; perciò si ha $x = (qn + r) + kn = (q + k)n + r$, dunque r è anche il resto della divisione di x per n. □

Definizione 1.1.6 (**Classi d'equivalenza e insieme quoziente)**. Sia X un insieme in cui è definita una relazione d'equivalenza $\sim$, sia $x \in X$. La **classe d'equivalenza** di x è il sottinsieme di X formato dagli elementi equivalenti a x:

$$[x] = \{y \in X \mid y \sim x\}.$$

Tale insieme si denota anche $[x]_\sim$.

L'insieme delle classi d'equivalenza è detto **insieme quoziente** di X rispetto alla relazione $\sim$ e si indica $X/\sim$.

L'insieme quoziente è un sottinsieme delle insieme delle parti di X, $\mathcal{P}(X)$. Osserviamo che $x \in [x]$ per la proprietà riflessiva. Quindi nessun elemento dell'insieme quoziente $X/\sim$ è l'insieme vuoto $\emptyset$. Inoltre le classi d'equivalenza ricoprono X, ossia X è l'unione delle classi d'equivalenza $[x]$, al variare di $x \in X$.

Definizione 1.1.7 (**Partizione)**. Una **partizione** di un insieme X è un sottinsieme Π dell'insieme delle parti di X che gode delle proprietà:

1. nessun elemento di Π è vuoto;
2. l'unione degli insiemi di Π è uguale a X;
3. se $S, T \in \Pi$, e $S \neq T$ allora $S \cap T = \emptyset$.

Dunque due elementi di una partizione Pi o sono disgiunti o sono uguali.

Proposizione 1.1.8. *L'insieme quoziente $X/\sim$ di una relazione d'equivalenza in X è una partizione di X.*

Dimostrazione. Le prime due proprietà sono già state osservate. Per provare la terza, consideriamo due classi d'equivalenza $[x], [y]$ tali che $[x] \cap [y] \neq \emptyset$. Allora esiste $z \in [x] \cap [y]$, cioè $z \sim x$ e $z \sim y$. Per le proprietà simmetrica e transitiva segue che $x \sim y$. Proviamo che di conseguenza $[x] = [y]$. Infatti, se $u \in [x]$, allora $u \sim x$, ma $x \sim y$, dunque per la proprietà transitiva $u \sim y$ e segue che $u \in [y]$. Abbiamo così provato che $[x] \subset [y]$. L'inclusione opposta è analoga. □

Esempio 1.1.9.

1. L'insieme quoziente $\mathbb{Z}/{\equiv_n}$ si denota $\mathbb{Z}_n$. $\mathbb{Z}_n$ ha n elementi, uno per ciascuno degli n resti della divisione per n: $0, 1, 2, \ldots, n - 1$. Infatti se x ha resto r nella divisione per n, $x = qn + r$ dunque $x \equiv_n r$. Gli elementi di $\mathbb{Z}_n$ si denotano anche $[r]_n$ o $\bar{r}$. Dunque $\mathbb{Z}_n = \{\bar{0}, \bar{1}, \ldots, \overline{n-1}\}$.

Un insieme S si dice finito se esiste un numero naturale n tale che S è in biiezione con l'insieme $\{1, 2, 3, \ldots, n-1, n\}$. Dunque $\mathbb{Z}_n$ è un insieme finito con n elementi.

1.2 Operazioni in $\mathbb{Z}_n$

Sia $n \geq 2$. Nell'insieme $\mathbb{Z}_n$ si possono definire due operazioni, di somma e di prodotto, **indotte** dalle operazioni in $\mathbb{Z}$.

Siano $\bar{x}, \bar{y} \in \mathbb{Z}_n$. Definiamo

$$\bar{x} + \bar{y} = \overline{x + y},$$

$$\bar{x} \cdot \bar{y} = \overline{x \cdot y}.$$

Il prodotto si denota anche semplicemente $\bar{x}\bar{y}$. Queste operazioni di somma e prodotto sono **ben definite**, in quanto non dipendono dai particolari rappresentanti scelti per le due classi. Infatti, sia $\bar{x} = \bar{x}'$ e $\bar{y} = \bar{y}'$. Allora si ha $x' = x + kn, y' = y + hn$, per $k, h \in \mathbb{Z}$ opportuni. Quindi $(x+y)-(x'+y') = (x+y) - (x + kn + y + hn) = -(k+h)n$, da cui segue che $x + y \equiv_n x' + y'$.

Analogamente $xy - x'y' = xy - (x+kn)(y+hn) = -(xh + yk + khn)n$ e perciò $xy \equiv_n x'y'$.

Dalle proprietà della somma in $\mathbb{Z}$ seguono facilmente le proprietà della somma in $\mathbb{Z}_n$:

1. proprietà associativa: $(\bar{x} + \bar{y}) + \bar{z} = \bar{x} + (\bar{y} + \bar{z})$;
2. la classe $\bar{0}$ è l'elemento neutro della somma;
3. $\overline{-x} = -\bar{x}$;
4. proprietà commutativa: $\bar{x} + \bar{y} = \bar{y} + \bar{x}$. Ne segue

Proposizione 1.2.1. $(\mathbb{Z}_n, +)$ *è un gruppo abeliano.*

Analogamente, dalle proprietà del prodotto in $\mathbb{Z}$ segue che valgono le seguenti proprietà del prodotto in $\mathbb{Z}_n$:

1. proprietà associativa: $(\bar{x}\bar{y})\bar{z} = \bar{x}(\bar{y}\bar{z})$;
2. $\bar{1}$ è l'unità del prodotto;
3. proprietà commutativa: $\bar{x}\bar{y} = \bar{y}\bar{x}$;
4. proprietà distributiva: $(\bar{x} + \bar{y})\bar{z} = \bar{x}z + \bar{y}z$.

Capitolo 2

Vettori applicati e loro operazioni

L'algebra lineare e la geometria affine sono due teorie che sono state sviluppate negli ultimi due secoli sul modello dell'algebra dei vettori liberi su una retta, in un piano o in uno spazio fisico, le cui proprietà sono una conseguenza degli assiomi classici della geometria euclidea. L'algebra lineare come strumento si è rivelata molto potente, perché applicabile a svariati contesti diversi, ha permesso di organizzare la geometria dello spazio euclideo in modo più efficiente sia dal punto di vista teorico che dal punto di vista computazionale rispetto all'approccio antico, e perché è alla base della moderna Analisi Funzionale.

È utile, in ogni caso, richiamare e rivedere velocemente la geometria e le operazioni con i vettori applicati e i vettori liberi, che permette di comprendere meglio la scelta delle definizione e degli assiomi dell'algebra lineare, e sarà di aiuto durante tutto il corso per la comprensione degli argomenti più complicati.

Consideriamo un piano o uno spazio "fisico". Un **vettore applicato** è un segmento orientato, ed è assegnato dando un punto di applicazione (il punto iniziale), la sua direzione (la retta su cui giace, detta anche giacitura), il suo modulo (la lunghezza del segmento, che è un numero reale ≥ 0), e il suo verso (uno dei due possibili versi di percorrenza della retta di giacitura). Osserviamo che tra i vettori applicati vi è anche il vettore nullo, identificabile con un punto del piano; per esso la direzione è indeterminata (esso giace su qualunque retta passante per il punto), e così pure il verso.

Un vettore applicato può essere caratterizzato anche come il dato di una coppia di punti, e precisamente: un punto iniziale A, punto di applicazione, ed un punto finale B, e verrà indicato con

$$\vec{AB}.$$

I vettori del tipo $\vec{AA}$, cioè tali che il punto di applicazione coincide con il punto finale, sono detti vettori applicati nulli.

Definizione 2.0.1. La **somma di due vettori applicati** del tipo $\vec{AB}$ e $\vec{BC}$ è per definizione il vettore

$$\vec{AB} + \vec{BC} := \vec{AC}.$$

Osservazione 2.0.1. La somma di un vettore $\vec{AB}$ con un vettore nullo del tipo $\vec{AA}$ oppure $\vec{BB}$ verifica:

$$\vec{AA} + \vec{AB} = \vec{AB}, \qquad \vec{AB} + \vec{BB} = \vec{AB}.$$

Inoltre la somma verifica la proprietà associativa:

$$\vec{AB} + (\vec{BC} + \vec{CD}) = (\vec{AB} + \vec{BC}) + \vec{CD}.$$

Definizione 2.0.2. La **moltiplicazione di un vettore applicato per uno scalare** è definita come segue: per ogni vettore applicato $\vec{AB}$ e per ogni numero reale $a \in \mathbb{R}$, il vettore $a \cdot \vec{AB}$ è quel vettore geometrico che ha

- la stessa direzione di $\vec{AB}$,
- lunghezza pari a quella di $\vec{AB}$ moltiplicata per il valore assoluto $|a|$ di a,
- verso concorde con $\vec{AB}$ se $a > 0$, altrimenti ha verso opposto. Se $a = 0$ si pone $0 \cdot \vec{AB} = \vec{AA}$.

2.1 Vettori liberi (o geometrici)

Due vettori applicati sono detti **equipollenti** se hanno la stessa direzione, la stessa lunghezza e lo stesso verso; in altre parole se giacciono su due rette parallele e se, muovendo una delle due rette parallelamente a se stessa, è possibile sovrapporre i due vettori in modo che i relativi punti iniziali e finali coincidano.

Si osservi che l'equipollenza è una relazione di equivalenza nell'insieme dei vettori applicati. Infatti ci si convince facilmente che valgono le seguenti proprietà: riflessiva, simmetrica e transitiva. Un **vettore libero o geometrico** è una classe di equivalenza di vettori applicati per la relazione di equipollenza (in questo contesto si dice anche classe di equipollenza). Denoteremo tra le parentesi quadre $[\vec{AB}]$ le classi di equipollenza dei vettori applicati ed i corrispondenti vettori liberi con le lettere minuscole:

$$\vec{u} = [\vec{AB}].$$

Il vettore nullo $\vec{0}$ è quel vettore rappresentato da un vettore applicato del tipo $\vec{AA}$.

Denoteremo con $\mathcal{V}^1$, $\mathcal{V}^2$ e $\mathcal{V}^3$ gli insiemi di vettori liberi della retta, del piano e dello spazio rispettivamente.

Osservazione 2.1.1. Fissato un punto O del piano o dello spazio, ogni vettore geometrico $\vec{v}$ 'e rappresentato da un unico vettore applicato $\vec{OP}$. Questa affermazione segue dal quinto postulato della geometria euclidea.

Nell'insieme dei vettori geometrici è possibile definire due operazioni: la somma di vettori e la moltiplicazione per uno scalare reale.

Definizione 2.1.1. La **somma di due vettori geometrici** $\vec{v} = [\vec{AB}]$ e $\vec{u} = [\vec{BC}]$ è così definita:

$$\vec{v} + \vec{u} = [\vec{AB}] + [\vec{BC}] := [\vec{AB} + \vec{BC}] = [\vec{AC}].$$

Osservazione 2.1.2. Si può verificare facilmente che la definizione è **ben posta**, cioè non dipende dai due rappresentanti scelti.

Inoltre, se i vettori $\vec{v}$ e $\vec{u}$ non sono allineati o nulli, la somma si può definire anche tramite la **regola del parallelogramma**:

se $\vec{v} = [\vec{OP}]$ e $\vec{u} = [\vec{OQ}]$, si costruisce il parallelogramma di lati OP ed OQ e si denota con R il quarto vertice di tale parallelogramma. La somma $\vec{v} + \vec{u}$ risulta uguale al vettore geometrico rappresentato da $\vec{OR}$.

Osservazione 2.1.3. L'operazione di somma tra due vettori geometrici soddisfa le seguenti proprietà:

- **Proprietà associativa**: per ogni scelta di una terna di vettori geometrici $\vec{u}, \vec{v}, \vec{w}$, si ha:

$$(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w}).$$

  Sfruttando questa proprietà possiamo, ad esempio, scrivere espressioni come $\vec{u} + \vec{v} + \vec{w}$.

- **Proprietà commutativa**: per ogni coppia di vettori geometrici $\vec{v}$ e $\vec{u}$ si ha

$$\vec{v} + \vec{u} = \vec{u} + \vec{v}.$$

- **Esistenza dell'elemento neutro**: se denotiamo con $\vec{0}$ il vettore nullo, allora per ogni vettore geometrico $\vec{v}$ vale:

$$\vec{v} + \vec{0} = \vec{0} + \vec{v} = \vec{v}.$$

- **Esistenza dell'opposto**: per ogni vettore geometrico $\vec{v}$, esiste il suo opposto, cioè un vettore geometrico $-\vec{v}$, tale che

$$\vec{v} + (-\vec{v}) = (-\vec{v}) + \vec{v} = \vec{0},$$

  dove $\vec{0}$ è il vettore nullo.

  Infatti, si consideri un rappresentante $\vec{AB}$ di $\vec{v}$, e si definisca $-\vec{v}$ come la classe di equipollenza di $\vec{BA}$.

  In seguito, per ogni coppia di vettori $\vec{v}, \vec{w}$, scriveremo semplicemente

$$\vec{v} - \vec{w}$$

  in luogo di

$$\vec{v} + (-\vec{w}).$$

Nell'insieme dei vettori geometrici, oltre alla somma di due vettori, possiamo definire la moltiplicazione per scalari in modo analogo sfruttando la moltiplicazione per scalari con vettori applicati.

Definizione 2.1.2. Per ogni vettore geometrico $\vec{v} = [\vec{AB}]$ e per ogni numero reale $a \in \mathbb{R}$ poniamo

$$a \cdot \vec{v} := [a \cdot \vec{AB}].$$

Anche in questo caso è facile verificare che la definizione è ben posta, cioè non dipende dal rappresentante.

Proprietà dell'operazione di moltiplicazione per scalari: per ogni coppia di vettori geometrici $\vec{v}, \vec{w}$ e per ogni coppia di scalari $a, b \in \mathbb{R}$, si ha:

- $1 \cdot \vec{v} = \vec{v}$;
- $(-1) \cdot \vec{v} = -\vec{v}$;
- $(a + b) \cdot \vec{v} = a \cdot \vec{v} + b \cdot \vec{v}$; $(ab) \cdot \vec{v} = a \cdot (b \cdot \vec{v})$; $a \cdot (\vec{v} + \vec{w}) = a \cdot \vec{v} + a \cdot \vec{w}$.

Sul modello dell'insieme dei vettori geometrici con le due operazioni appena descritte e le loro proprietà introduciamo la seguente definizione di spazio vettoriale. Osserviamo che la nuovo definizione è molto generale e comprende spazi di natura molto diversa.

Definizione 2.1.3. Uno **spazio vettoriale reale** o **$\mathbb{R}$-spazio vettoriale** è un insieme non vuoto V su cui sono definite due operazioni, una somma

$$\begin{aligned} + : V \times V &\to V, \\ (v, w) &\to v + w, \end{aligned}$$

ed un prodotto per scalari

$$\begin{aligned} \cdot : \mathbb{R} \times V &\to V, \\ (a, v) &\to a \cdot v, \end{aligned}$$

in modo che siano soddisfatti i seguenti assiomi, per ogni $u, v, w \in V$, e per ogni $a, b \in \mathbb{R}$:

1. V1: **proprietà associativa**:
$$(u + v) + w = u + (v + w);$$
2. V2: **proprietà commutativa**:
$$u + v = v + u;$$
3. V3: **esistenza del vettore nullo**: esiste $0 \in V$ tale che
$$0 + v = v + 0 = v;$$
4. V4: **esistenza dell'opposto**: per ogni $v \in V$, esiste un vettore $-v \in V$ tale che
$$v + (-v) = (-v) + v = 0;$$
5. V5: **distributiva di $\cdot$ rispetto a $+$**:
$$a \cdot (u + v) = a \cdot u + a \cdot v;$$
6. V6: **distributiva di $\cdot$ rispetto alla somma di $\mathbb{R}$**:
$$(a + b) \cdot v = a \cdot v + b \cdot v;$$
7. V7: $(ab) \cdot v = a \cdot (b \cdot v)$;
8. V8: per ogni $v \in V$ si ha
$$1 \cdot v = v.$$

Notiamo che abbiamo utilizzato lo stesso simbolo 0 per denotare sia il vettore nullo che lo zero come numero reale, per non appesantire la notazione. Inoltre, nel seguito, espressioni del tipo $v + (-w)$ verranno semplificate con $v - w$.

Esempio 2.1.4.

1. L'insieme dei vettori geometrici della retta $\mathcal{V}^1$, del piano $\mathcal{V}^2$ o dello spazio $\mathcal{V}^3$ con le operazioni descritte all'inizio del capitolo è uno spazio vettoriale su $\mathbb{R}$.

2. L'insieme dei numeri reali $\mathbb{R}$ è uno spazio vettoriale su $\mathbb{R}$ con le usuali definizioni di somma tra numeri reali e prodotto tra numeri reali, dove questa volta il prodotto
$$\cdot : \mathbb{R} \times \mathbb{R} \to \mathbb{R}, \qquad (a, b) \to a \cdot b$$
viene considerato come operazione "esterna", nel senso che $\mathbb{R}$ gioca sia il ruolo di insieme degli scalari che il ruolo di spazio dei vettori.

3. il prodotto cartesiano
$$\mathbb{R}\times\mathbb{R}=\mathbb{R}^2:=\left\{\begin{pmatrix} x_1\\ x_2\end{pmatrix}\mid x_1\in\mathbb{R},\ x_2\in\mathbb{R}\right\}$$
con l' operazione di somma così definita:
$$+:\mathbb{R}^2\times\mathbb{R}^2\to\mathbb{R}^2,\qquad \left(\begin{pmatrix} x_1\\ x_2\end{pmatrix},\begin{pmatrix} y_1\\ y_2\end{pmatrix}\right)\to\begin{pmatrix} x_1\\ x_2\end{pmatrix}+\begin{pmatrix} y_1\\ y_2\end{pmatrix}:=\begin{pmatrix} x_1+y_1\\ x_2+y_2\end{pmatrix}$$
e di moltiplicazione per uno scalare:
$$\cdot:\mathbb{R}\times\mathbb{R}^2\to\mathbb{R}^2,\qquad \left(c,\begin{pmatrix} x_1\\ x_2\end{pmatrix}\right)\to c\cdot\begin{pmatrix} x_1\\ x_2\end{pmatrix}:=\begin{pmatrix} c\cdot x_1\\ c\cdot x_2\end{pmatrix}$$
verifica gli assiomi per uno spazio vettoriale reale.

4. Più in generale, per ogni $n\in\mathbb{N}$, il prodotto cartesiano di $\mathbb{R}$ per se stesso n -volte:
$$\mathbb{R}^n=\left\{\begin{pmatrix} x_1\\ x_2\\ \vdots\\ x_n\end{pmatrix}\mid x_1\in\mathbb{R},\ x_2\in\mathbb{R},\dots,x_n\in\mathbb{R}\right\}$$
con l' operazione di somma così definita:
$$+:\mathbb{R}^2\times\mathbb{R}^2\to\mathbb{R}^2,\qquad \left(\begin{pmatrix} x_1\\ x_2\\ \vdots\\ x_n\end{pmatrix},\begin{pmatrix} y_1\\ y_2\\ \vdots\\ y_n\end{pmatrix}\right)\to\begin{pmatrix} x_1\\ x_2\\ \vdots\\ x_n\end{pmatrix}+\begin{pmatrix} y_1\\ y_2\\ \vdots\\ y_n\end{pmatrix}:=\begin{pmatrix} x_1+y_1\\ x_2+y_2\\ \vdots\\ x_n+y_n\end{pmatrix}$$
e di moltiplicazione per uno scalare:
$$\cdot:\mathbb{R}\times\mathbb{R}^2\to\mathbb{R}^2,\qquad \left(c,\begin{pmatrix} x_1\\ x_2\\ \vdots\\ x_n\end{pmatrix}\right)\to c\cdot\begin{pmatrix} x_1\\ x_2\\ \vdots\\ x_n\end{pmatrix}:=\begin{pmatrix} c\cdot x_1\\ c\cdot x_2\\ \vdots\\ c\cdot x_n\end{pmatrix}$$
verifica gli assiomi per uno spazio vettoriale reale.

5. L'insieme delle funzioni reali
$$\mathcal{F}:=\{f|f:\mathbb{R}\to\mathbb{R}\},$$
con l'operazione di somma (definita puntalmente):
$$+:\mathcal{F}\times\mathcal{F}\to\mathcal{F},\qquad (f,g)\to f+g\ :\ (f+g)(r):=f(r)+g(r),$$
e la moltiplicazione per uno scalare
$$\cdot:\mathbb{R}\times\mathcal{F}\to\mathcal{F},\qquad (c,f)\to c\cdot f\ :\ (c\cdot f)(r):=c\cdot f(r)$$
è uno spazio vettoriale reale.

6. L'insieme dei polinomi reali in una indeterminata:
$$\mathbb{R}[x] := \{p(x) \mid p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0,\ a_n, a_{n-1}, \ldots, a_1, a_0 \in \mathbb{R}\}$$
con l'usuale somma tra polinomi e l'usuale prodotto per uno scalare è uno spaizo vettoriale reale.

2.2 Campi

Vogliamo ora generalizzare ancora la nozione di spazio vettoriale, sostituendo a $\mathbb{R}$ un insieme arbitrario, che abbia delle proprietà algebriche simile ad $\mathbb{R}$, cioè che sia un campo.

Definizione 2.2.1. Un **campo** è un insieme non vuoto $\mathbb{K}$ dotato di due operazioni interne, una somma
$$+ : \mathbb{K} \times \mathbb{K} \to \mathbb{K}, \qquad (a, b) \to a + b,$$
ed un prodotto
$$\cdot : \mathbb{K} \times \mathbb{K} \to \mathbb{K}, \qquad (a, b) \to a \cdot b,$$
che in seguito sarà denotato anche con ab, in modo che siano soddisfatti i seguenti assiomi, per ogni $a, b, c \in \mathbb{K}$:

1. K1: **commutatività** $a + b = b + a$, $ab = ba$;
2. K2: **associatività** $(a + b) + c = a + (b + c)$, $a(bc) = (ab)c$;
3. K3: **esistenza dell'elemento neutro** $\exists 0 \in \mathbb{K}$, tale che $a + 0 = a$, inoltre $\exists 1 \in \mathbb{K}$ tale che $a1 = a$;
4. K4: **esistenza dell'opposto e dell'inverso** $\exists -a \in \mathbb{K}$ tale che $a + (-a) = 0$; se $a \neq 0$, $\exists a^{-1} \in \mathbb{K}$ tale che $aa^{-1} = 1$;
5. K5: **distributività di $\cdot$ rispetto ad $+$** $a(b + c) = ab + ac$.

Esempio 2.2.2. Esempi di campi sono i numeri razionali $\mathbb{Q}$, i numeri reali $\mathbb{R}$, i numeri complessi $\mathbb{C}$, le classi $\mathbb{Z}_n$ di congruenza modulo n con n numero primo.

2.3 Esercizi

Sia V uno spazio vettoriale su un campo K. Si dimostrino, usando gli assiomi di spazio vettoriale $V1, V2, \ldots, V8$, le seguenti proprietà:

1. per ogni vettore $\mathbf{v} \in V$ si ha
$$0 \cdot \mathbf{v} = \mathbf{0},$$
dove $0 \in K$ indica l'elemento neutro rispetto alla somma in K, e $\mathbf{0} \in V$ indica il vettore nullo, cioè l'elemento neutro rispetto alla somma in V.

   Infatti, sfruttando il fatto che $0 \in \mathbb{K}$ è l'elemento neutro per la somma e per la $V5$, abbiamo le seguenti uguaglianze:
$$0 \cdot \mathbf{v} = (\mathbf{0} + \mathbf{0}) \cdot \mathbf{v} = \mathbf{0} \cdot \mathbf{v} + \mathbf{0} \cdot \mathbf{v}.$$
   Aggiungendo ora l'opposto $-0 \cdot v$ di $0 \cdot \mathbf{v}$ a destra e sinistra e sfruttando la $V3$, concludiamo che $\mathbf{0} = \mathbf{0} \cdot \mathbf{v}$.

2. per ogni $t \in K$ si ha
$$t \cdot \mathbf{0} = \mathbf{0}.$$

   Infatti, sfruttando le $V3$ e $V5$, deduciamo che
$$a \cdot \mathbf{0} = \mathbf{a} \cdot (\mathbf{0} + \mathbf{0}) = \mathbf{a} \cdot \mathbf{0} + \mathbf{a} \cdot \mathbf{0}.$$

   Aggiungendo $-a \cdot \mathbf{0}$ a destra e sinistra, e sfruttando la $V4$, concludiamo che $\mathbf{0} = \mathbf{a} \cdot \mathbf{0}$.

3. In V esiste un unico vettore nullo.

   Infatti, supponiamo per assurdo che esistano due vettori nulli $\mathbf{0}$ e $\mathbf{0}' \in \mathbf{V}$, cioè tali che
$$\mathbf{v} + \mathbf{0} = \mathbf{0} + \mathbf{v} = \mathbf{v},$$
$$\mathbf{v} + \mathbf{0}' = \mathbf{0}' + \mathbf{v} = \mathbf{v},$$
   per ogni $\mathbf{v} \in \mathbf{V}$. Allora, in particolare, per il vettore $\mathbf{v} = \mathbf{0}$ si avrebbe:
$$\mathbf{0} = \mathbf{0} + \mathbf{0}' = \mathbf{0}'.$$

   Quindi un assurdo.

4. Ogni vettore $\mathbf{v}$ ha un unico opposto.

   Infatti, se $\mathbf{v}$ avesse due opposti, $-\mathbf{v}$ e $-\mathbf{v}'$, allora si avrebbe:
$$-\mathbf{v} = -\mathbf{v} + \mathbf{0} = -\mathbf{v} + (\mathbf{v} + (-\mathbf{v}')) = (-\mathbf{v} + \mathbf{v}) + (-\mathbf{v}') = \mathbf{0} + (-\mathbf{v}') = -\mathbf{v}'.$$

   Che è un assurdo.

5. Per ogni $\mathbf{v} \in \mathbf{V}$ si ha
$$(-1) \cdot \mathbf{v} = -\mathbf{v},$$
   dove $1 \in K$ denota l'elemento neutro per il prodotto in K, -1 il suo opposto e $-\mathbf{v}$ l'opposto di $\mathbf{v}$.

   Infatti,
$$\mathbf{v} + (-\mathbf{1}) \cdot \mathbf{v} = \mathbf{1} \cdot \mathbf{v} + (-\mathbf{1}) \cdot \mathbf{v} = (\mathbf{1} - \mathbf{1}) \cdot \mathbf{v} = \mathbf{0} \cdot \mathbf{v} = \mathbf{0}.$$
   Quindi $(-1) \cdot \mathbf{v}$ è un vettore opposto di $\mathbf{v}$, ma poiché l'opposto è unico, esso coincide con $-\mathbf{v}$.

Capitolo 3

Matrici: prime definizioni

Definizione 3.0.1. Siano $m, n \in \mathbb{N}$ e sia mK un campo. Una **matrice** $m \times n$ **a coefficienti in** $\mathbb{K}$ è una tabella rettangolare di $m \cdot n$ elementi di $\mathbb{K}$ del tipo

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

dove $a_{ij} \in \mathbb{K}$ per ogni $i \in \{1, \dots, m\}$ e per ogni $j \in \{1, \dots, n\}$.

Verrá usate anche la seguente notazione per denotare una matrice:

$$A = (a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}.$$

Per ogni $i = 1, \dots, m$, la **riga i-esima** di A è la matrice $1 \times n$

$$A_{(i)} = (a_{i1}\ a_{i2}\ \dots a_{in}).$$

Per ogni $j = 1, \dots, n$, la **colonna j-esima** di A è la matrice $m \times 1$

$$A^{(j)} = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

L'elemento a_{ij} è detto **elemento di posto** i, j; tale elemento verrà indicato anche con

$$(A)_{ij}.$$

Osserviamo, infine, che i è l'indice di riga e j l'indice di colonna.

Definizione 3.0.2. Se $m = n$, cioè il numero di righe è uguale al numero di colonne, la matrice si dice **matrice quadrata di ordine** n.

Definizione 3.0.3. L'insieme di tutte le matrici $m \times n$ a coefficienti in $\mathbb{K}$ si denota con $M_{m,n}(\mathbb{K})$. L'insieme delle matrici quadrate di ordine n a coefficienti in $\mathbb{K}$ si denota con $M_n(\mathbb{K})$.

Definizione 3.0.4. La **matrice nulla** è la matrice con 0 al posto i, j, per ogni $i \in \{1, \ldots, m\}$ e $j \in \{1, \ldots, n\}$:

$$\mathbf{O} = \begin{pmatrix} 0 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ldots & \vdots \\ \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & \ldots & 0 \end{pmatrix}$$

Esempio 3.0.5. La matrice

$$A = \begin{pmatrix} 2 & 1 & 0 \\ -1 & 13 & 1/4 \end{pmatrix}$$

ha 2 righe e 3 colonne. La prima riga è

$$A_{(1)} = (2 \quad 1 \quad 0),$$

e la seconda riga è

$$A_{(2)} = (-1 \quad 13 \quad 1/4).$$

Le colonne sono, rispettivamente:

$$A^{(1)} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \; A^{(2)} = \begin{pmatrix} 1 \\ 13 \end{pmatrix}, \; A^{(3)} = \begin{pmatrix} 0 \\ 1/4 \end{pmatrix}.$$

Esempio 3.0.6. Le matrici

$$B = \begin{pmatrix} 5 & -3 \\ 2/3 & 0 \end{pmatrix}, \; C = \begin{pmatrix} 12 & 11 & 10 \\ 0 & 9 & 1 \\ 3 & 3 & 3 \end{pmatrix}$$

sono quadrate di ordine 2 e 3, rispettivamente.

Definizione 3.0.7. Siano $A = (a_{ij}), B = (b_{ij}) \in M_{m,n}(\mathbb{K})$, e sia $c \in \mathbb{K}$. La **somma** di A e B si definisce come la matrice $A + B \in M_{m,n}(\mathbb{K})$ il cui elemento di posto i, j è $a_{ij} + b_{ij}$ per ogni $i \in \{1, \ldots, m\}$ e $j \in \{1, \ldots, n\}$; possiamo scrivere:

$$(A + B)_{ij} = (A)_{ij} + (B)_{ij}.$$

Il **prodotto di A per lo scalare** c è la matrice $c \cdot A \in M_{m,n}(\mathbb{K})$ il cui elemento di posto i, j è dato da $c \cdot a_{ij}$, per ogni $i \in \{1, \ldots, m\}$ e $j \in \{1, \ldots, n\}$.

Proposizione 3.0.8. *$M_{m,n}(\mathbb{K})$ con le operazioni di somma e prodotto per scalari definite sopra, è uno spazio vettoriale sul campo $\mathbb{K}$. Il vettore nullo è la matrice nulla, cioè la matrice con $0 \in \mathbb{K}$ al posto i, j, per ogni $i \in \{1, \ldots, m\}$ e $j \in \{1, \ldots, n\}$.*

Dimostrazione. Esercizio. □

Osservazione 3.0.1. Analogamente agli spazi $\mathbb{R}^n$ possiamo definire gli spazi

$$\mathbb{K}^n = \left\{ \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} \mid c_1, c_2, \ldots, c_n \in \mathbb{K} \right\},$$

detto **prodotto cartesiano** del campo $\mathbb{K}$ per se stesso n volte. I suoi elementi sono sequenze ordinate di n scalari di $\mathbb{K}$.

Come nel caso reale, in $\mathbb{K}^n$ possiamo definire una somma e una moltiplicazione per uno scalare, che lo rendono uno spazio vettoriale su $\mathbb{K}$:

$$\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} + \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{pmatrix} = \begin{pmatrix} c_1 + d_1 \\ c_2 + d_2 \\ \vdots \\ c_n + d_n \end{pmatrix},$$

$$a \cdot \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} a\, c_1 \\ a\, c_2 \\ \vdots \\ a\, c_n \end{pmatrix}$$

Osservazione 3.0.2. Possiamo vedere gli elementi di $\mathbb{K}^n$ come matrici $n \times 1$ a coefficienti in $\mathbb{K}$. Sotto questa identificazione, le operazioni di somma e prodotto per scalari definite in $\mathbb{K}^n$ coincidono con quelle di $M_{n,1}(\mathbb{K})$, quindi possiamo identificare

$$\mathbb{K}^n = M_{n,1}(\mathbb{K})$$

come spazi vettoriali.

3.1 Matrici e immagini digitali

Una notevole applicazione della teoria delle matrici è rappresentata dalle immagini digitali. Gli esempi e le osservazioni presenti in queste note sono stati preparati dal prof. Marino Zennaro, professore ordinario di Analisi Numerica presso il DMG, in occasione del suo intervento per il Circolo Matematico nel 2018. Per la presentazione integrale si veda la pagina

https://circolomatematico.units.it/content/23-marzo-2018-matematica-e-immagini-digitali.

Ringrazio il prof. M. Zennaro per la gentile concessione di utilizzo del materiale.

Esempio 3.1.1. Le immagini digitali sono memorizzate sotto forma di una o tre tabelle numeriche rettangolari, ovvero matrici, a seconda che si tratti rispettivamente di un'immagine in scala di grigi o a colori. Le dimensioni delle matrici dipendono dalla risoluzione dell'immagine, ossia dal numero di pixel che la compongono. Una modifica dei numeri contenuti nelle tabelle comporta una modifica corrispondente dell'immagine da esse rappresentata. Sulla base di ciò, formule matematiche più o meno complicate vengono comunemente utilizzate per operare sulle immagini una svariata gamma di trasformazioni.

$i \setminus j$	1	2	3	4	5	6	7	8	9	10
1	255	227	0	171	143	115	87	255	31	3
2	255	0	0	0	143	115	255	255	255	3
3	0	0	199	0	0	255	255	59	255	255
4	0	227	199	171	0	255	87	59	31	255
5	0	227	199	171	0	255	87	59	31	255
6	0	227	199	171	0	255	87	59	31	255
7	0	227	199	171	0	255	87	59	31	255
8	0	0	199	0	0	255	255	59	255	255
9	255	0	0	0	143	115	255	255	255	3
10	255	227	0	171	143	115	87	255	31	3

tabella numerica F ($m = 10$ righe e $n = 10$ colonne)
codifica a 8 bit ($2^8 = 256$ tonalità): da 0 (nero) a 255 (bianco)

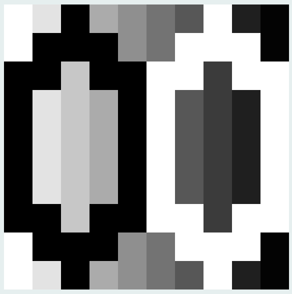


immagine corrispondente (ingrandita 60 volte)

$i \setminus j$	1	2	3	4	5	6	7	8	9	10
1	255 255 255	227 255 255	0 50 50	171 221 221	143 193 193	115 165 65	87 137 37	255 255 205	31 81 0	3 53 0
2	255 255 255	0 50 50	0 50 50	0 50 50	143 193 193	115 165 65	255 255 205	255 255 205	255 255 205	3 50 0
3	0 50 50	0 50 50	199 249 249	0 50 50	0 50 50	255 255 205	255 255 205	59 109 9	255 255 205	255 255 205
4	0 50 50	227 255 255	199 249 249	171 221 221	0 50 50	255 255 205	87 137 37	59 109 9	31 81 0	255 255 205
5	0 50 50	227 255 255	199 249 249	171 221 221	0 50 50	255 255 205	87 137 37	59 109 9	31 81 0	255 255 205
6	0 0 50	227 177 255	199 149 249	171 121 221	0 0 50	255 205 205	87 37 37	59 9 9	31 0 0	255 205 205
7	0 0 50	227 177 255	199 149 249	171 121 221	0 0 50	255 205 205	87 37 37	59 9 9	31 0 0	255 205 205
8	0 0 50	0 0 50	199 149 249	0 0 50	0 0 50	255 205 205	255 205 205	59 9 9	255 205 205	255 205 205
9	255 205 255	0 0 50	0 0 50	0 0 50	143 93 193	115 65 65	255 205 205	255 205 205	255 205 205	3 0 0
10	255 205 255	227 177 255	0 0 50	171 121 221	143 93 193	115 65 65	87 37 37	255 205 205	31 0 0	3 0 0

tabella numerica "tripla" (F_1, F_2, F_3) ($m = 10$ e $n = 10$)
codifica a 8 bit: tre numeri tra 0 e 255 per determinare colore e tonalità

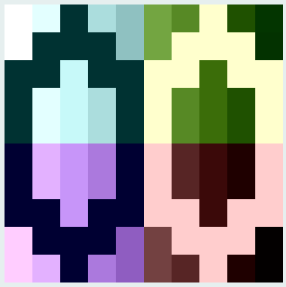


immagine corrispondente (ingrandita 60 volte)

La somma di matrici associate a un immagine e il prodotto per un numero hanno le seguenti interpretazioni.

Esempio 3.1.2. Sia A la matrice di tipo 10×10 associata a un'immagine. Sia B di tipo 10×10 e con tutte le entrate uguali a 255:

$$B = \begin{pmatrix} 255 & 255 & \dots & 255 \\ 255 & 255 & \dots & 255 \\ \vdots & \vdots & \dots & \vdots \\ \vdots & \vdots & \dots & \vdots \\ 255 & 255 & \dots & 255 \end{pmatrix}.$$

Se consideriamo la matrice

$$B + (-1) \cdot A = B - A,$$

otteniamo l'immagine "negativa", in cui i vari colori vengono sostituiti con i loro "opposti". Ad esempio, il bianco al posto del nero ed il nero al posto del bianco.

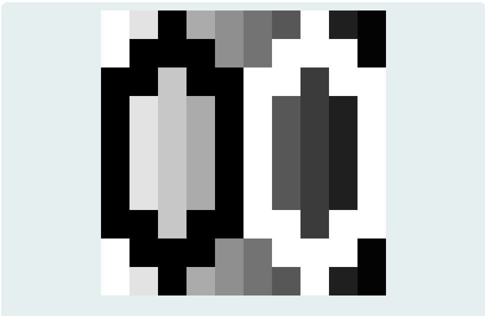
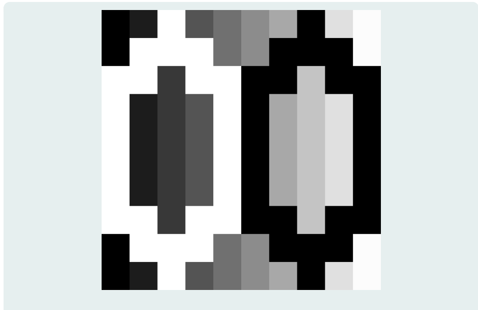


immagine corrispondente (ingrandita 60 volte)



"negativo" dell'immagine in scala di grigi dell'Esempio 1.1

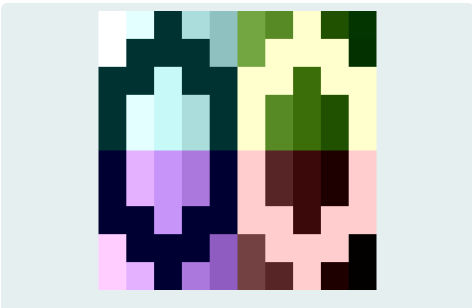
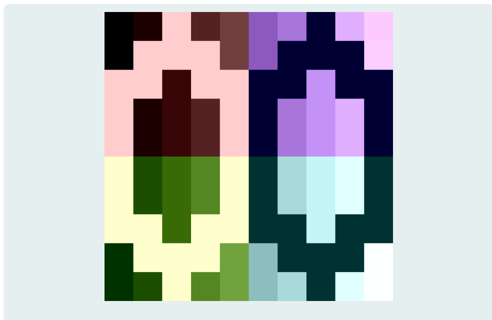
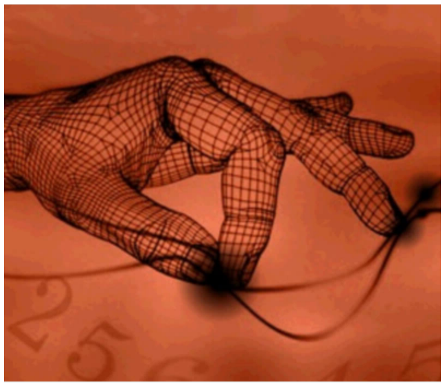


immagine corrispondente (ingrandita 60 volte)



"negativo" dell'immagine a colori dell'Esempio 1.2



Esempio 3.1.3. Scegliamo un opportuno parametro intero $0 < \delta \leq 255$, e sia A la matrice di tipo 10×10 associata a un'immagine. Consideriamo la matrice Δ di tipo 10×10 e con tutte le entrate uguali a δ:

$$\Delta = \begin{pmatrix} \delta & \delta & \dots & \delta \\ \delta & \delta & \dots & \delta \\ \vdots & \vdots & \dots & \vdots \\ \vdots & \vdots & \dots & \vdots \\ \delta & \delta & \dots & \delta \end{pmatrix}.$$

La matrice $A + \Delta$ corrisponde all'operazione di "schiarimento" in maniera uniforme.

Osserviamo che $A + \Delta$ potrebbe avere alcune entrate > 255; in questo caso poniamo tali entrate $= 255$.

Notiamo, infine, che per $\delta = 255$ otteniamo un'immagine completamente bianca.

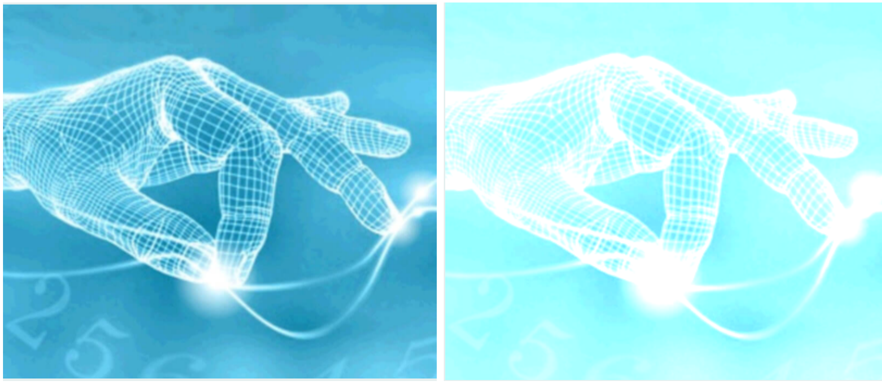


Immagine schiarita con $\delta = 100$

Analogamente possiamo "scurire" l'immagine considerando la matrice $A - \Delta$. In questo caso alcune delle entrate possono risultare < 0. In questo caso le poniamo uguali a 0.

Notiamo che per $\delta = 255$ otteniamo un'immagine completamente nera.

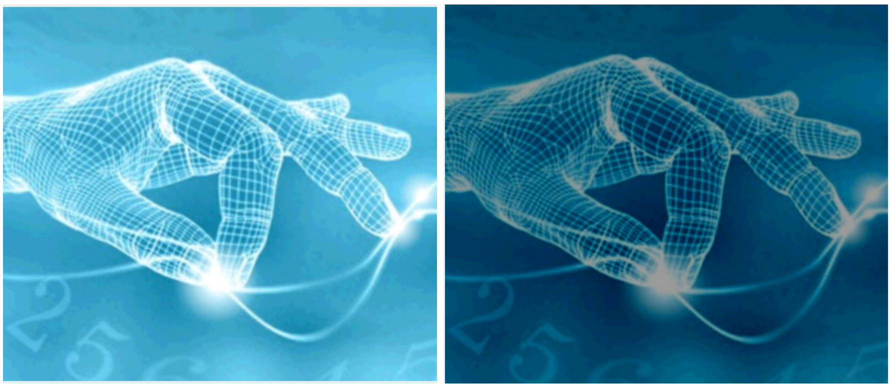


Immagine scurita con $\delta = 100$

Esempio 3.1.4. Partiamo dalla semplice osservazione che

$$255 - z = -(z - 127.5) + 127.5$$

L'applicazione della formula di destra equivale a "spostare l'origine" dell'asse z da 0 a $127.5 = 255/2$.

Nasce quindi l'idea di considerare un generico "fattore di rimodulazione" α e di generalizzare la formula precedente per definire gli elementi di una nuova matrice A' tramite la formula

$$a'_{ij} = \alpha(a_{ij} - 127.5) + 127.5$$

Se si sceglie $\alpha > 1$ si opera una **dilatazione** delle tonalità dei vari colori presenti nell'immagine, che tendono tutte a spostarsi verso quelle estreme (ad esempio, bianco e nero nelle immagini in scala di grigi).

In particolare, per $\alpha \geq 255$ si ottengono matrici A' con elementi tutti uguali a 0 o a 255, quindi immagini con le sole tonalità di colore estreme.

Se si sceglie $\alpha = 1$ non si effettua alcuna modifica alla matrice A e l'immagine rimane invariata.

Se si sceglie invece $0 < \alpha < 1$, si opera una **contrazione** delle tonalità, che tendono tutte a spostarsi verso le due colorazioni grigie medie, corrispondenti a 127 o 128.

In particolare, per $\alpha < 1/127.5$, si ottengono matrici A' con elementi tutti uguali a 127 o a 128, quindi immagini completamente grigie.

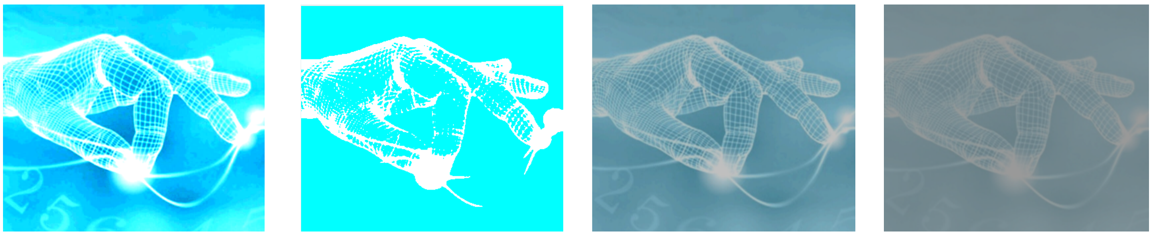


Immagine rimodulata con $\alpha = 2, 255, 0.5, 0.2$, rispettivamente.

Esempio 3.1.5. Terminiamo la rassegna con una sorta di "trucco matematico" che ci permette di "nascondere" un'immagine con matrice G dentro un'altra immagine con matrice F, utilizzando due strumenti, la rimodulazione dei colori e una **combinazione lineare convessa**.

L'idea consiste nel fissare un fattore $\alpha > 0$ abbastanza piccolo (diciamo, ad esempio, 0.05) e nel definire la matrice

$$H = (1 - \alpha) \cdot F + \alpha \cdot G.$$

L'immagine corrispondente ad H 'e molto simile all'immagine F e non si intravede l'immagine G in essa "nascosta". Più piccolo è α, meglio si nasconde G.

La prima immagine contiene la seconda immagine "nascosta" con fattore $\alpha = 0.05$.

Per "estrarre" l'immagine n.2 dall'immagine n.3 basta interpretare la formula precedente come un'equazione nell'incognita G e risolverla (trascurando gli arrotondamenti agli interi nel calcolo di H). Bisogna però disporre anche dell'immagine n.1 e conoscere il fattore α.

Immagine estratta dall'immagine n.3 con fattore $\alpha = 0.05$.

3.2 Matrice trasposta

Definizione 3.2.1. Sia $A \in M_{m,n}(\mathbb{K})$. La **matrice trasposta di** A è la matrice

$$^tA \in M_{n,m}(\mathbb{K})$$

il cui elemento di posto i, j è l'elemento di posto j, i di A, per ogni $i \in \{1, \dots, n\}$ e $j \in \{1, \dots, m\}$, ovvero:

$$(^tA)_{ij} = (A)_{ji}.$$

Proposizione 3.2.2. *Siano* $A, B \in M_{m,n}(\mathbb{K})$. *Allora si ha:*

1. $^t(A + B) = {}^tA + {}^tB$;
2. $^t(^tA) = A$.

Dimostrazione. 1. Dimostriamo che per ogni $i \in \{1, \ldots, m\}$ e per ogni $j \in \{1, \ldots, n\}$, gli elementi di posto i, j di ${}^t(A + B)$ e di ${}^tA + {}^tB$ coincidono; si ha:

$$({}^t(A + B))_{ij} = (A + B)_{ji} = (A)_{ji} + (B)_{ji};$$

inoltre:

$$({}^tA + {}^tB)_{ij} = ({}^tA)_{ij} + ({}^tB)_{ij} = (A)_{ji} + (B)_{ji},$$

quindi l'uguaglianza della tesi vale.

2. Si ha

$$({}^t({}^tA))_{ij} = ({}^tA)_{ji} = (A)_{ij},$$

quindi l'uguaglianza è verificata.

□

Definizione 3.2.3. Sia $A = (a_{ij}) \in M_n(\mathbb{K})$. La **diagonale principale** è quella parte di A composta dagli elementi di posto i, i, per ogni $i = 1, \ldots, n$.

A è detta **diagonale** se $a_{ij} = 0$ per ogni $i \neq j$.

A è **simmetrica** se $A = {}^tA$.

La **matrice unità** $n \times n$è la matrice $\mathbb{I}_n \in M_n(\mathbb{K})$ il cui elemento di posto i, j è dato da 1 se $i = j$ e 0 se $i \neq j$. Useremo spesso il **simbolo di Kronecker** δ_{ij} così definito:

$$\delta_{ij} := \begin{cases} 1 \text{ se } i = j \\ 0 \text{ se } i \neq j \end{cases}$$

La matrice unità ha quindi al posto i, j l'elemento δ_{ij}.

3.3 Il prodotto righe per colonne

Definizione 3.3.1. Siano

$$A = (a_{11}\ a_{12} \ldots a_{1n}) \in M_{1,n}(\mathbb{K}), \quad B = \begin{pmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{n1} \end{pmatrix} \in M_{n,1}(\mathbb{K}).$$

Il **prodotto** $A \cdot B$ è lo scalare definito come segue:

$$A \cdot B = a_{11}b_{11} + a_{12}b_{21} + \cdots + a_{1n}b_{n1} \in \mathbb{K}.$$

Più in generale, se $A \in M_{m,n}(\mathbb{K})$ e $B \in M_{n,p}(\mathbb{K})$, il **prodotto righe per colonne**

$$A \cdot B \in M_{m,p}(\mathbb{K})$$

è quella matrice $m \times p$ il cui elemento di posto i, j è dato dal prodotto della i-esima riga di A e la j-esima colonna di B:

$$(A \cdot B)_{ij} = A_{(i)} \cdot B^{(j)} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}.$$

Notiamo che nel prodotto righe per colonne, il numero delle colonne della matrice a sinistra A deve coincidere con il numero delle righe della matrice a destra B.

Osservazione 3.3.1. Osserviamo che se A e B sono matrici quadrate dello stesso ordine:

$$A, B \in M_n(\mathbb{K}),$$

è sempre possibile moltiplicare $A \cdot B$, ma in generale si ha $A \cdot B \neq B \cdot A$.

Esempio 3.3.2. Siano

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

Si ha:

$$A \cdot B = \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix},$$

mentre

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

Proposizione 3.3.3.

1. *Siano $A, B \in M_{m,n}(\mathbb{K})$, $C, D \in M_{n,p}(\mathbb{K})$, $c \in \mathbb{K}$. Allora valgono le seguenti uguaglianze:*

$$(A + B) \cdot C = A \cdot C + B \cdot C, \quad A \cdot (C + D) = A \cdot C + A \cdot D,$$
$$A \cdot (c \cdot C) = c \cdot (A \cdot C) = (c \cdot A) \cdot C, \quad A \cdot \mathbb{I}_n = \mathbb{I}_m \cdot A = A.$$

2. *Siano $A \in M_{m,n}(\mathbb{K}),\ B \in M_{n,p}(\mathbb{K}),\ C \in M_{p,q}(\mathbb{K})$. Allora vale la seguente uguaglianza:*

$$(A \cdot B) \cdot C = A \cdot (B \cdot C).$$

3. *Siano $A \in M_{m,n}(\mathbb{K}), \quad B \in M_{n,p}(\mathbb{K})$. Allora vale la seguente uguaglianza:*

$${}^t(A \cdot B) = {}^tB \cdot {}^tA.$$

Dimostrazione. Dimostriamo solo l'uguaglianza 3). Si ha

$$({}^t(A \cdot B))_{ij} = (A \cdot B)_{ji} = A_{(j)} \cdot B^{(i)} =$$
$$= a_{j1}b_{1i} + a_{j2}b_{2i} + \cdots + a_{jn}b_{ni}.$$

D'altra parte si ha

$$({}^tB \cdot {}^tA)_{ij} = ({}^tB)_{(i)} \cdot ({}^tA)^{(j)} = b_{1i}a_{j1} + b_{2i}a_{j2} + \cdots + b_{ni}a_{jn} = a_{j1}b_{1i} + a_{j2}b_{2i} + \cdots + a_{jn}b_{ni}.$$

□

3.4 Matrice inversa

Definizione 3.4.1. Una matrice quadrata $A \in M_n(\mathbb{K})$ si dice invertibile, se esiste una matrice quadrata dello stesso ordine $M \in M_n(\mathbb{K})$ tale che sono verificate le seguenti uguaglianze:

$$A \cdot M = M \cdot A = \mathbb{I}_n.$$

Proposizione 3.4.2. *Sia data una matrice quadrata $A \in M_n(\mathbb{K})$.*

1. *Se A è invertibile, allora esiste un' unica matrice $M \in M_n(\mathbb{K})$, tale che $A \cdot M = M \cdot A = \mathbb{I}_n$. Tale M si dice* **matrice inversa di** *A e si indica con A^{-1}.*

2. *Se A è invertibile e $M \in M_n(\mathbb{K})$ è tale che*

$$A \cdot M = \mathbb{I}_n,$$

   allora $M = A^{-1}$.

   Lo stesso vale se $M \cdot A = \mathbb{I}_n$.

3. *Se $A, B \in M_n(\mathbb{K})$ sono invertibili, allora anche $A \cdot B$ è invertibile, e*

$$(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}.$$

Dimostrazione. 1. Se esistesse un' altra $N \in M_n(\mathbb{K})$ tale che $A \cdot N = N \cdot A = \mathbb{I}_n$. Allora sfruttando il fatto che $N \cdot \mathbb{I}_n = N$, che $\mathbb{I}_n = A \cdot M$, la proprietà associativa del prodotto righe per colonne, e che $\mathbb{I}_n \cdot M = M$, abbiamo:

$$N = N \cdot \mathbb{I}_n = N \cdot (A \cdot M) = (N \cdot A) \cdot M = \mathbb{I}_n \cdot M = M,$$

quindi le due matrici sono uguali.

2. Dobbiamo dimostrare che anche $M \cdot A = \mathbb{I}_n$. Ma siccome A è invertibile, esiste la sua inversa A^{-1}. Quindi

$$M \cdot A = \mathbb{I}_n \cdot M \cdot A = (A^{-1} \cdot A) \cdot M \cdot A = A^{-1} \cdot (A \cdot M) \cdot A = A^{-1} \cdot \mathbb{I}_n \cdot A = A^{-1} \cdot A = \mathbb{I}_n.$$

3. È sufficiente osservare che

$$(A \cdot B) \cdot B^{-1} \cdot A^{-1} = A \cdot (B \cdot B^{-1}) \cdot A^{-1} = A \cdot A^{-1} = \mathbb{I}_n,$$

e che $B^{-1} \cdot A^{-1} \cdot (A \cdot B) = \mathbb{I}_n$.

□

Osservazione 3.4.1. Notiamo che $A \neq 0$ non implica che A sia invertibile. Si consideri ad esempio la matrice

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

Osserviamo che A non è invertibile. Infatti, se esistesse

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

tale che $A \cdot M = \mathbb{I}_2$, allora si avrebbe

$$\begin{cases} a + b = & 1 \\ a + b = & 0 \\ c + d = & 0 \\ c + d = & 1, \end{cases}$$

ma il sistema non ha soluzione.

Capitolo 4

Sistemi di equazioni lineari

Siano $m, n \in \mathbb{N} \setminus \{0\}$, e sia $\mathbb{K}$ un campo.

Definizione 4.0.1.

- Un **sistema di m equazioni lineari a coefficienti in $\mathbb{K}$ in n incognite** è un sistema di equazioni della seguente forma:

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

  dove $a_{11}, a_{12}, \ldots, a_{mn} \in \mathbb{K}$ sono i **coefficienti**, $x_1, \ldots, x_n$ sono le **incognite**, n è l' **ordine**, $b_1, \ldots, b_m \in \mathbb{K}$ sono i **termini noti**, del sistema lineare.

- Una **soluzione** del sistema lineare è una n-upla ordinata (vettore colonna)

$$s = \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{pmatrix} \in \mathbb{K}^n$$

  tale che, se si sostituisce ad x_i il valore $s_i \in \mathbb{K}$, per ogni $i = 1, \ldots, n$, le m equazioni sono simultaneamente soddisfatte.

- Il sistema lineare si dice **omogeneo**, rispettivamente **non omogeneo**, se

$$b_1 = b_2 = \cdots = b_m = 0,$$

  rispettivamente se $b_j \neq 0$, per qualche $j = 1, \ldots, m$.

- Il sistema lineare si dice **compatibile** (rispettivamente **incompatibile**), se possiede una soluzione (rispettivamente, se non ha alcuna soluzione).

Osservazione 4.0.1. Ogni sistema lineare omogeneo è compatibile, infatti il vettore nullo

$$0 = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{K}^n$$

è una sua soluzione, detta soluzione banale. Ogni altra soluzione si dice soluzione non banale.

Esempio 4.0.2. 1. Il sistema lineare di ordine 2

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

è incompatibile.

2. Il sistema lineare

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

è compatibile ed ammette l'unica soluzione $s = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$.

3. Il sistema lineare

$$\begin{cases} 2x_1 + 3x_2 & = 1 \\ 4x_1 + 6x_2 & = 2 \end{cases}$$

è compatibile ed ammette infinite soluzioni date da $s = \begin{pmatrix} \frac{1-3t}{2} \\ t \end{pmatrix}$; per ogni valore di $t \in \mathbb{R}$, otteiamo una soluzione del sistema.

Definizione 4.0.3. La **matrice dei coefficienti** del sistema lineare

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

è la matrice A di tipo $m \times n$ a coefficienti in $\mathbb{K}$, il cui elemento di posto ij è il coefficiente a_{ij} del sistema lineare, per ogni $i = 1, \ldots, m$, $j = 1, \ldots, n$, cioè

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

Il **vettore dei termini noti** è il vettore colonna

$$b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Osserviamo che se mettiamo le incognite in colonna

$$X := \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

allora il precedente sistema lineare si può scrivere nella seguente forma:

$$A \cdot X = b,$$

dove A è la matrice dei coefficienti, b il vettore dei termini noti, ed il prodotto è il prodotto righe per colonne.

La **matrice completa** associata al precedente sistema lineare è la matrice

$$(A|b) = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & | & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & | & b_2 \\ \vdots & \vdots & \dots & \vdots & | & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & | & b_m \end{pmatrix}.$$

Teorema 4.0.4. di struttura per le soluzioni dei sistemi lineari omogenei *Sia $A \cdot X = 0$ un sistema lineare omogeneo di ordine n a coefficienti in $\mathbb{K}$. Per ogni coppia di soluzioni*

$$s, s' \in \mathbb{K}^n$$

di $A \cdot X = 0$ e per ogni scalare

$$c \in \mathbb{K},$$

si ha che

$$s + s', \quad cs \in \mathbb{K}^n$$

sono soluzioni di $A \cdot X = 0$.

In particolare, l'insieme delle soluzioni del sistema lineare omogeneo $A \cdot X = 0$ è un sottospazio vettoriale di $\mathbb{K}^n$ (per la definizione di sottospazio vettoriale si veda il prossimo capitolo).

Dimostrazione. Siccome $s, s' \in \mathbb{K}^n$ sono soluzioni del sistema lineare omogeneo in considerazione, si ha che $A \cdot s = 0$ ed $A \cdot s' = 0$. Dobbiamo provare che $A \cdot (s + s') = 0$ e che $A \cdot (cs) = 0$.

Per la proprietà distributiva del prodotto righe per colonne sulla somma di matrici (in questo caso s, s' sono vettori colonna, quindi si possono considerare come matrici di tipo $n \times 1$), si ha che

$$A \cdot (s + s') = A \cdot s + A \cdot s' = 0 + 0 = 0.$$

Segue che $s + s' \in \mathbb{K}^n$ è una soluzione del sistema lineare omogeneo.

Per dimostrare che $A \cdot (cs) = 0$, basta osservare che il prodotto per scalari commuta con il prodotto tra matrici, quindi

$$A \cdot (cs) = cA \cdot s = c0 = 0.$$

□

Teorema 4.0.5. di struttura per le soluzioni dei sistemi lineari *Sia $A \cdot X = b$ un sistema lineare di ordine n, e sia $\widetilde{s} \in \mathbb{K}^n$ una sua soluzione.*

Allora $s \in \mathbb{K}^n$ è soluzione del sistema lineare, se e solo se

$$s = \widetilde{s} + s_0,$$

dove $s_0 \in \mathbb{K}^n$ è una soluzione del sistema lineare omogeneo associato $A \cdot X = 0$.

Dimostrazione. Sia $s \in \mathbb{K}^n$ una soluzione arbitraria del sistema lineare $A \cdot X = b$. Osserviamo che

$$s = \widetilde{s} + (s - \widetilde{s}),$$

quindi basta verificare che $s - \widetilde{s}$ è soluzione del sistema lineare omogeneo associato $A \cdot X = 0$. Sfruttando la proprietà distributiva del prodotto righe per colonne tra matrici si ha

$$A \cdot (s - \widetilde{s}) = A \cdot s + A \cdot (-\widetilde{s}) = A \cdot s - A \cdot (\widetilde{s}) = b - b = 0,$$

dove nella seconda uguaglianza abbiamo sfruttato il fatto che $-\widetilde{s} = (-1)\widetilde{s}$ e la commutatività del prodotto righe per colonne con il prodotto per scalari.

Viceversa, per ogni soluzione $s_0 \in \mathbb{K}^n$ del sistema lineare omogeneo associato, $\widetilde{s} + s_0$ è una soluzione del sistema lineare, poiché

$$A \cdot (\widetilde{s} + s_0) = A \cdot \widetilde{s} + A \cdot s_0 = b + 0 = b.$$

□

4.1 Sistemi lineari con matrici dei coefficienti a scala

Definizione 4.1.1. Sia $A = (a_{ij}) \in M_{m,n}(\mathbb{K})$ e sia $r \in \{0, 1, \ldots, m\}$ il numero delle righe di A diverse dalla riga nulla. La matrice A è detta **a scala** se:

- $r = 0$ (quindi A è la matrice nulla),
- oppure $r > 0$, $A_{(i)} \neq (0\,0 \ldots 0)$ per ogni $i \in \{1, \ldots, r\}$ e, posto

  $$j_i = \min\{j \mid a_{ij} \neq 0\},$$

  si ha che $j_1 < j_2 < \cdots < j_r$.

  Se A è a scala, gli elementi $a_{1j_1}, \ldots, a_{rj_r}$ si dicono i **pivot** di A.

Proposizione 4.1.2. *Sia $A \cdot X = b$ un sistema lineare di ordine n, formato da m equazioni, a coefficienti in $\mathbb{K}$. Supponiamo che $A \in M_{m,n}(\mathbb{K})$ sia una matrice a scala, e sia $r \in 0, \ldots, m$ il numero di righe non nulle di A.*

Allora il sistema lineare $A \cdot X = b$ è compatibile $\iff$

$$b_{r+1} = b_{r+2} = \cdots = b_m = 0.$$

Dimostrazione. Dimostriamo l' implicazione $\Rightarrow$:

Per ipotesi il sistema lineare $A \cdot X = b$ è compatibile, quindi ha una soluzione

$$s = \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{pmatrix}$$

per cui vale $A \cdot s = b$. In particolare, per ogni $i = 1, \ldots, m$, la componente i-esima del vettore $A \cdot s \in \mathbb{K}^n$ è data dal prodotto (righe per colonne) della i-esima riga di A per s, cioè $A_{(i)} \cdot s$.

Sia ora $i > r$; per definizione di r abbiamo che

$$A_{(i)} \cdot s = 0,$$

perció $b_i = 0$.

Implicazione $\Leftarrow$:

Supponiamo che $b_i = 0$ per ogni $i > r$, e dimostriamo che è sempre possibile trovare una soluzione del sistema lineare $A \cdot X = b$. A tale scopo, procediamo risolvendo tutte le equazioni partendo dall' ultima equazione ed andando a ritroso.

Precisamente, le equazioni dalla $(r+1)$-esima alla m-esima sono tutte del tipo $0 = 0$, quindi sono soddisfatte per ogni scelta di $s_1, \ldots, s_n \in \mathbb{K}$,ponendo $x_1 = s_1, \ldots, x_n = s_n$. L' equazione r-esima è della forma seguente:

$$a_{r,j_r} x_{j_r} + a_{r,j_r+1} x_{j_r+1} + \cdots + a_{r,n} x_n = b_r.$$

Siccome $a_{r,j_r} \neq 0$, essendo l' r-esimo pivot di A, possiamo ricavare x_{j_r} in funzione di $x_{j_r+1}, x_{j_r+2}, \ldots, x_n$, ed otteniamo:

$$x_{j_r} = \frac{1}{a_{r,j_r}}(b_r - a_{r,j_r+1} x_{j_r+1} - \cdots - a_{r,n} x_n).$$

Quindi, per ogni $s_1, \ldots, s_{j_r-1} \in \mathbb{K}$ e per ogni $s_{j_r+1}, \ldots, s_n \in \mathbb{K}$, otteniamo una soluzione della equazione r-esima del sistema lineare ponendo

$$x_1 = s_1, \ldots, x_{j_r-1} = s_{j_r-1}, \ldots, x_n = s_n \in \mathbb{K}, x_{j_r} = \frac{1}{a_{r,j_r}}(b_r - a_{r,j_r+1} s_{j_r+1} - \cdots - a_{r,n} s_n).$$

Ora consideriamo l' equazione $(r-1)$-esima. Essa è della forma seguente:

$$a_{r-1,j_{r-1}} x_{j_{r-1}} + a_{r-1,j_{r-1}+1} x_{j_{r-1}+1} + \cdots + a_{r-1,n} x_n = b_{r-1}.$$

Dividendo ambo i membri per $a_{r-1,j_{r-1}}$, che è $\neq 0$, ricaviamo $x_{j_{r-1}}$ in funzione di $x_{j_{r-1}+1}, \ldots, x_n$. Quindi, sostituendo ad x_{j_r} il valore determinanto in precedenza, vediamo che è possibile trovare degli scalari $s_1, s_2, \ldots, s_n \in \mathbb{K}$, tali che $s = \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{pmatrix} \in \mathbb{K}^n$ sia simultaneamente soluzione delle equazioni dalla $(r-1)$-esima alla m-esima del sistema lineare. Procedendo in questo modo vediamo che il sistema lineare $A \cdot X = b$ ha (almeno) una soluzione e quindi è compatibile.

□

4.2 Il metodo di gradinizzazione di Gauß

Definizione 4.2.1. Due sistemi lineari dello stesso ordine sono **equivalenti** se hanno le stesse soluzioni.

Il metodo di Gau_ per stabilire la compatibilità ed eventualmente per trovare le soluzioni di un sistema lineare $A \cdot X = b$, consiste nel trasformare tale sistema in uno ad esso equivalente $\widetilde{A} \cdot X = \widetilde{b}$, con matrice dei coefficienti $\widetilde{A}$ a scala. Quindi si stabilisce la compatibilità di $\widetilde{A} \cdot X = \widetilde{b}$, ed eventualmente se ne trovano le soluzioni, tramite la Proposizione 4.1.2. Per trasformare $A \cdot X = b$ in $\widetilde{A} \cdot X = \widetilde{b}$ si effettuano in modo opportuno le cosiddette **operazioni elementari**.

1. OE1 **Operazione elementare 1 (OE1)**. Questa operazione consiste nello scambio di due equazioni tra di loro. Nella pratica, per risolvere un dato sistema lineare, sarà spesso più comodo effettuare le operazioni elementari direttamente sulla matrice completa $(A \mid b)$. In tal caso la OE1 consiste nello scambio di due righe di $(A \mid b)$ tra di loro.

2. OE2 **Operazione elementare 2 (OE2)**. Questa operazione consiste nella moltiplicazione di ambo i membri di una equazione per uno stesso scalare non nullo. La corrispondente operazione sulla matrice completa $(A \mid b)$ consiste nella moltiplicazione di una sua riga per uno scalare non nullo.

3. OE3 **Operazione elementare 3 (OE3)**. Con questa operazione si sostituisce una equazione con l' equazione che si ottiene sommando ad essa un multiplo di un' altra equazione. La corrispondente operazione sulla matrice completa $(A \mid b)$ consiste nel sostituire una sua riga, ad esempio $(A \mid b)_{(i)}$, con la somma $(A \mid b)_{(i)} + c(A \mid b)_{(j)}$, per qualche $j \neq i$ e $c \in \mathbb{K}$.

Proposizione 4.2.2. *Le operazioni elementari 1, 2 e 3 trasformano un dato sistema lineare in uno ad esso equivalente.*

Dimostrazione. Sia $A \cdot X = b$ un dato sistema lineare di ordine n, con m equazioni, a coefficienti in $\mathbb{K}$. Chiaramente scambiando l' ordine delle sue equazioni, si ottiene un sistema lineare equivalente ad esso. Lo stesso vale se si moltiplica ambo i membri di una equazione per uno scalare non nullo.

Consideriamo quindi l' OE3, e precisamente sostituiamo l' equazione i-esima di $A \cdot X = b$ con l' equazione che si ottiene sommando ad essa c-volte l' equazione j-esima, per qualche $c \in \mathbb{K}$ e per qualche $j \neq i$. Denotiamo con $\widetilde{A} \cdot X = \widetilde{b}$ il sistema lineare che si ottiene in questo modo. Ricordiamo che, per ogni $k = 1, \dots, m$, la k-esima equazione di $A \cdot X = b$ si può scrivere come segue: $A_{(k)} \cdot X = b_k$, dove $A_{(k)}$ è la k-esima riga di A, b_k è la k-esima componente di $b \in \mathbb{K}^m$, ed il prodotto è quello righe per colonne. Analogamente la k-esima equazione di $\widetilde{A} \cdot X = \widetilde{b}$ si scrive come $\widetilde{A}_{(k)} \cdot X = \widetilde{b}_k$, dove

$$\widetilde{A}_{(k)} = \begin{cases} A_{(k)}, & \text{se } k \neq i, \\ A_{(i)} + cA_{(j)}, & \text{se } k = i, \end{cases} \qquad \widetilde{b}_k = \begin{cases} b_k, & \text{se } k \neq i, \\ b_i + cb_j, & \text{se } k = i, \end{cases}$$

Dimostriamo ora che $A \cdot X = b$ e $\widetilde{A} \cdot X = \widetilde{b}$ sono equivalenti. Sia $s \in \mathbb{K}^n$ una soluzione di $A \cdot X = b$; abbiamo quindi

$$A_{(k)} \cdot s = b_k, \ \forall k = 1, \dots, m.$$

Quindi

$$\widetilde{A}_{(k)} \cdot s = \widetilde{b}_k, \ \forall k \neq i,$$

e per $k = i$

$$\widetilde{A}_{(i)} \cdot s = (A_{(i)} + cA_{(j)}) \cdot s = A_{(i)} \cdot s + cA_{(j)} \cdot s = b_i + cb_j = \widetilde{b}_i.$$

Ne segue che s è soluzione di $\widetilde{A} \cdot X = \widetilde{b}$.

Viceversa, se $s \in \mathbb{K}^n$ è una soluzione di $\widetilde{A} \cdot X = \widetilde{b}$, allora

$$\widetilde{A}_{(k)} \cdot s = \widetilde{b}_k, \ \forall k = 1, \dots, m.$$

Quindi

$$A_{(k)} \cdot s = b_k, \ \forall k \neq i,$$

e per $k = i$
$$A_{(i)} \cdot s = (\widetilde{A}_{(i)} - c\widetilde{A}_{(j)}) \cdot s = \widetilde{A}_{(i)} \cdot s - c\widetilde{A}_{(j)} \cdot s = \widetilde{b}_i - c\widetilde{b}_j = b_i.$$
Ne segue che s è soluzione di $A \cdot X = b$, ed i due sistemi lineari sono equivalenti. $\square$

Teorema 4.2.3. *Sia dato un sistema lineare $A \cdot X = b$. È sempre possibile trasformare $A \cdot X = b$, per mezzo delle operazioni elementari, in uno ad esso equivalente, $\widetilde{A} \cdot X = \widetilde{b}$, con matrice dei coefficienti $\widetilde{A}$ a scala.*

Dimostrazione. Sia n il numero di incognite e sia m il numero delle sue equazioni. Possiamo supporre $A \neq 0$, altrimenti A è già a scala. Sia quindi $j \in \{1, \dots, n\}$ il più piccolo indice di colonna tale che $A^{(j)} \neq 0 \in \mathbb{K}^m$, e sia $i \in \{1, \dots, m\}$ un indice di riga tale che $a_{ij} \neq 0 \in \mathbb{K}$.

ALGORITMO DI GAUß

- Scambiando tra di loro la prima con la i-ma riga di $(A|b)$ (operazione OE1), possiamo supporre che sia
$$a_{1j} \neq 0.$$
Abbiamo quindi una matrice del tipo
$$\left(\begin{array}{ccccccccc} 0 & \dots & 0 & a_{1j} & a_{1,j+1} & \dots & a_{1n} & | & b_1 \\ 0 & \dots & 0 & a_{2j} & a_{2,j+1} & \dots & a_{2n} & | & b_2 \\ \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots & | & \vdots \\ 0 & \dots & 0 & a_{mj} & a_{m,j+1} & \dots & a_{mn} & | & b_m \end{array}\right)$$

- Per mezzo di una OE3, sostituiamo le righe di $(A|b)$ dalla seconda alla m-esima con
$$(A|b)_{(k)} - \frac{a_{kj}}{a_{1j}}(A|b)_{(1)}, \ \forall k = 2, \dots, m$$
e troviamo una matrice del tipo
$$\left(\begin{array}{ccccccccc} 0 & \dots & 0 & a_{1j} & a_{1,j+1} & \dots & a_{1n} & | & b_1 \\ 0 & \dots & 0 & 0 & a'_{2,j+1} & \dots & a'_{2n} & | & b'_2 \\ \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots & | & \vdots \\ 0 & \dots & 0 & 0 & a'_{m,j+1} & \dots & a'_{mn} & | & b'_m \end{array}\right)$$

- Applichiamo i punti precedenti alla sottomatrice
$$\left(\begin{array}{ccccccccc} 0 & \dots & 0 & a'_{2,j+1} & a'_{2,j+2} & \dots & a'_{2n} & | & b'_2 \\ \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots & | & \vdots \\ 0 & \dots & 0 & a'_{m,j+1} & a'_{m,j+2} & \dots & a'_{mn} & | & b'_m \end{array}\right).$$

Alla fine del procedimento troviamo una matrice $(\widetilde{A}|\widetilde{b})$ a scala. $\square$

Capitolo 5

Sottospazi vettoriali

Definizione 5.0.1. Sia V uno spazio vettoriale su un campo $\mathbb{K}$. Un sottoinsieme non vuoto $W \subseteq V$ si dice **sottospazio vettoriale** di V se valgono le seguenti condizioni:

- (W1) per ogni $w_1 \in W$ e per ogni $w_2 \in W$ si ha
$$w_1 + w_2 \in W;$$
- (W2) per ogni $W \in W$ e per ogni scalare $a \in \mathbb{K}$ si ha
$$a \cdot w \in W.$$

Osservazione 5.0.1. Osserviamo che un sottospazio vettoriale W è a sua volta uno spazio vettoriale su $\mathbb{K}$ con le operazioni ereditate da V. È facile verificare che valgono gli assiomi $V1, \dots, V8$ di spazio vettoriale.

Esempio 5.0.2. **di sottospazi vettoriali**

1. Il sottoinsieme $W = V$ risulta in modo evidente un sottospazio vettoriale, detto **sottospazio vettoriale improprio**.
2. Il sottoinsieme formato dal solo vettore nullo $\{0\} \subseteq V$ è un sottospazio vettoriale, percheé verifica $W1$ e $W2$, e si chiama **sottospazio vettoriale banale**. Osserviamo che è il più piccolo sottospazio (e anche spazio) vettoriale.
3. Il sottoisieme di $\mathcal{F}(\mathbb{R}, \mathbb{R})$ costituito dalle funzioni **limitate**:
$$\mathcal{L}(\mathbb{R}, \mathbb{R}) := \{f \mid f : \mathbb{R} \to \mathbb{R},\ f \text{ limitata}\} \subset \mathcal{F}(\mathbb{R}, \mathbb{R})$$
è un sottospazio vettoriale. Ricordiamo la definizione: $f : \mathbb{R} \to \mathbb{R}$ si dice limitata se $\exists\ M > 0$, $M \in \mathbb{R}$, tale che $|f(r)| \leq M$ per ogni $r \in \mathbb{R}$.
4. Nello spazio delle funzioni
$$\mathcal{F}(\mathbb{R}, \mathbb{R}) = \{f \mid f : \mathbb{R} \to \mathbb{R}\}$$
il sottoinsieme delle funzioni continue
$$\mathcal{C}^0(\mathbb{R}, \mathbb{R}) := \{f \mid f : \mathbb{R} \to \mathbb{R} \text{ continua}\}$$

è un sottospazio vettoriale perché la somma di funzioni continue e la moltiplicazione di uno scalare per una funzione continua sono ancora funzioni continue.

Analogamente si può verificare facilmente che il sottoinsieme delle funzioni derivabili con derivata continua

$$\mathcal{C}^1(\mathbb{R},\mathbb{R}) := \{f \mid f : \mathbb{R} \to \mathbb{R} \text{ derivabile e } f' \text{ continua}\}$$

è un sottospazio vettoriale.

Inoltre si ha

$$\mathcal{C}^0(\mathbb{R},\mathbb{R}) \supset \mathcal{C}^1(\mathbb{R},\mathbb{R}),$$

e $\mathcal{C}^1(\mathbb{R},\mathbb{R})$ risulta anche sottospazio vettoriale di $\mathcal{C}^0(\mathbb{R},\mathbb{R})$.

5. Consideriamo lo spazio vettoriale $\mathbb{R}[x]$ dei polinomi in una indeterminata a coefficienti in $\mathbb{R}$.

   Il sottoinsieme dei polinomi di grado minore o uguale a un grado fissato $d \in \mathbb{N}$

   $$\mathbb{R}[x]_{\leq d} := \{p(x) \mid \deg p(x) \leq d\}$$

   risulta un sottospazio vettoriale.

Esempio 5.0.3. **Sottoinsiemi che non sono sottospazi vettoriali**

1. La circonferenza

   $$S := \{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mid x_1^2 + (x_2 - 1)^2 = 1\} \subseteq \mathbb{R}^2$$

   non è un sottospazio vettoriale; infatti, ad esempio

   $$\begin{pmatrix} 0 \\ 2 \end{pmatrix} \in S, \text{ ma } -\begin{pmatrix} 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ -2 \end{pmatrix} \notin S.$$

2. In generale, ogni sottoinsieme **limitato** di $\mathbb{R}^2$ e di $\mathbb{R}^n$ in generale non è un sottospazio vettoriale; infatti, la condizione $W2$ implica che i vettori di un sottospazio vettoriale possano assumere "lunghezze" (vedremo la definizione in seguito) arbitrariamente grandi.

3. Le rette del piano che non passano per l'origine non sono sottospazi vettoriali, perché non contengono il vettore nullo.

4. Consideriamo lo spazio vettoriale $\mathbb{R}[x]$ dei polinomi in una indeterminata a coefficienti in $\mathbb{R}$.

   Il sottoinsieme dei polinomi di grado **uguale a un grado fissato** $d \in \mathbb{N}$

   $$\mathbb{R}[x]_d := \{p(x) \mid \deg p(x) = d\}$$

   non risulta un sottospazio vettoriale, perché non verifica la $W1$. Ad esempio, la somma dei due polinomi

   $$x^d - 1, \quad -x^d + 3$$

   non è un polinomio di grado d: $x^d - 1 + (-x^d + 3) = 2$, polinomio costante, di grado zero.

Proposizione 5.0.4. *Sia V uno spazio vettoriale su $\mathbb{K}$, e siano*

$$U \subseteq V, \quad W \subseteq V$$

due suoi sottospazi vettoriali.

Allora l'intersezione

$$U \cap W \subseteq V$$

è ancora un sottospazio vettoriale.

Dimostrazione. Verifichiamo che vale la $W1$ per $U \cap W$:

siano $u_1, u_2 \in U \cap W \subseteq U$; siccome U è sottospazio vettoriale, si ha

$$u_1 + u_2 \in U.$$

Ma abbiamo anche $u_1, u_2 \in U \cap W \subseteq W$; siccome W è sottospazio vettoriale, si ha

$$u_1 + u_2 \in W.$$

Quindi $u_1 + u_2 \in U \cap W$.

Verifichaimo ora la $W2$: sia $u \in U \cap W$ e sia $c \in \mathbb{K}$; essendo $U \cap W \subseteq U$ ed essendo U sottospazio vettoriale, si ha

$$c \cdot u \in U.$$

Analogamente, essendo $U \cap W \subseteq W$ ed essendo W sottospazio vettoriale, si ha

$$c \cdot u \in W.$$

Concludiamo quindi nuovamente che $c \cdot u \in U \cap W$. □

Osservazione 5.0.2. Sia V uno spazio vettoriale su $\mathbb{K}$, e siano

$$U \subseteq V, \quad W \subseteq V$$

due suoi sottospazi vettoriali.

In generale l'unione

$$U \cup W$$

non è un sottospazio vettoriale.

Ci chiediamo allora quale sia il più piccolo sottospazio vettoriale contenente due dati sottospazi. Abbiamo la seguente:

Definizione 5.0.5. Sia V uno spazio vettoriale su $\mathbb{K}$, e siano

$$U \subseteq V, \quad W \subseteq V$$

due suoi sottospazi vettoriali. Il **sottospazio vettoriale somma** $U + W$ è così definito:

$$U + V := \{v \in V \mid \exists u \in U, \exists w \in W, \text{ tali che } v = u + w\},$$

è dato cioè da tutte le possibili somme di vettori di U con vettori di W.

Lemma 5.0.6. *Il sottospazio somma $U + W \subseteq V$ è un sottospazio vettoriale.*

Inoltre, $U + W$ è il più piccolo sottospazio vettoriale contenente $U \cup W$.

Dimostrazione. Esercizio. □

5.1 Combinazioni lineari e sottospazi vettoriali finitamente generati

Definizione 5.1.1. Dati $v_1, \dots, v_k \in V$, una **combinazione lineare** di $v_1, \dots, v_k$ a coefficienti in $\mathbb{K}$ è un vettore del tipo

$$c_1 v_1 + \dots + c_k v_k \in V, \quad c_1, \dots, c_k \in \mathbb{K}.$$

Definizione 5.1.2. Se $v_1, \dots, v_k \in V$ è un numero finito di vettori di V, definiamo

$$\mathrm{Span}(v_1, \dots, v_k) := \{a_1 \cdot v_1 + \dots a_k \cdot v_k \mid a_1, \dots, a_k \in \mathbb{K}\},$$

che risulta un sottospazio vettoriale (verificare per esercizio).

Esempio 5.1.3.

- Sia V uno spazio vettoriale non banale, e sia $v \in V$ un vettore non nullo: $v \neq 0$. Allora

  $$\mathrm{Span}(v) = \{c \cdot v \mid c \in \mathbb{K}\},$$

  consiste cioè di tutti i vettori proporzionali a v. Il sottospazio vettoriale $\mathrm{Span}(v)$ viene chiamato **retta vettoriale**, e il vettore v viene chiamato **generatore**.

- Sia V uno spazio vettoriale non banale e siano $v, w \in V$ due vettori non nulli: $v \neq 0$, $w \neq 0$. Allora

  $$\mathrm{Span}(v, w) = \{a \cdot v + b \cdot w \mid a \in \mathbb{K}, b \in \mathbb{K}\},$$

  consiste cioè di tutte le possibili somme di vettori proporzionali a v e w.

  Nel caso che w non sia proporzionale a v, $\mathrm{Span}(v, w)$ viene chiamato **piano vettoriale**. Nel caso che w sia invece proporzionale a v, si può verificare facilmente che vale

  $$w = c \cdot v \Rightarrow \mathrm{Span}(v, w) = \mathrm{Span}(v) = \mathrm{Span}(w),$$

  quindi si ottiene nuovamente una retta vettoriale.

Osservazione 5.1.1. Osserviamo, in particolare, che i vettori $v_1, \dots, v_k \in \mathrm{Span}(v_1, \dots, v_k)$; infatti, si ha

$$v_1 = 1 \cdot v_1 + 0 \cdot v_2 + \dots + 0 \cdot v_k, v_2 = 0 \cdot v_1 + 1 \cdot v_2 + \dots + 0 \cdot v_k, \dots,$$

$$v_k = 0 \cdot v_1 + 0 \cdot v_2 + \dots + 1 \cdot v_k.$$

Inoltre, per ogni $l < k$, abbiamo

$$\mathrm{Span}(v_1, \dots, v_l) \subseteq \mathrm{Span}(v_1, \dots, v_k).$$

Infatti, ogni $v \in \mathrm{Span}(v_1, \dots, v_l)$ soddisfa

$$v = c_1 \cdot v_1 + c_2 \cdot v_2 + \dots + c_l \cdot v_l,$$

e tale relazione si può riscrivere nella forma:

$$v = c_1 \cdot v_1 + c_2 \cdot v_2 + \dots + c_l \cdot v_l + 0 \cdot v_{l+1} + \dots + 0 \cdot v_k,$$

quindi v è anche combinazione lineare di $v_1, \dots, v_k$.

Notiamo, però, che $l < k$ non implica, in generale,

$$\mathrm{Span}(v_1, \dots, v_l) \subsetneq \mathrm{Span}(v_1, \dots, v_k),$$

come ad esempio nel seguente caso:

$$\mathrm{Span}\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right) = \mathrm{Span}\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 5 \\ 5 \end{pmatrix}\right) \subset \mathbb{R}^2.$$

Esempio 5.1.4.
- Consideriamo i vettori

$$v_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad v_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \in \mathbb{R}^2.$$

Osserviamo che il vettore $w = \begin{pmatrix} 3 \\ 0 \end{pmatrix}$ si puó scrivere come combinazione lineare di v_1 e v_2:

$$w = \begin{pmatrix} 3 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} - 2 \begin{pmatrix} -1 \\ 1 \end{pmatrix} = v_1 - 2v_2,$$

quindi $w \in \mathrm{Span}(v_1, v_2)$.

Sia ora $v \begin{pmatrix} c \\ d \end{pmatrix} \in \mathbb{R}^2$ un vettore qualunque. Ci chiediamo se $v \in \mathrm{Span}(v_1, v_2)$, cioé se esistano due coefficienti λ_1 e λ_2 tali che

$$v = \lambda_1 v_1 + \lambda_2 v_2 = \lambda_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \lambda_2 \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ 2\lambda_1 \end{pmatrix} + \begin{pmatrix} -\lambda_2 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} \lambda_1 - \lambda_2 \\ 2\lambda_1 + \lambda_2 \end{pmatrix}.$$

Per determinare l'esistenza di λ_1 e λ_2 dobbiamo risolvere il sistema lineare

$$\begin{cases} \lambda_1 - \lambda_2 & = c \\ 2\lambda_1 + \lambda_2 & = d. \end{cases}$$

La matrice completa $(A|b)$ associata al sistema lineare é

$$(A|b) = \begin{pmatrix} 1 & -1 & | & c \\ 2 & 1 & | & d \end{pmatrix}.$$

Con l' operazione elementare $(A|b)_{(2)} - 2(A|b)_{(1)}$ riduciamo la matrice a scala e otteniamo:

$$\begin{pmatrix} 1 & -1 & | & c \\ 0 & 3 & | & d-2c \end{pmatrix}.$$

Il sistema lineare ha (un'unica) soluzione

$$\lambda_1 = \frac{d+c}{3}, \qquad \lambda_2 = \frac{d-2c}{3}.$$

Osserviamo che nel caso particolare di prima $c = 3$ e $d = 0$ si ottiene effettivamente $\lambda_1 = 1$ e $\lambda_2 = -2$.

Come conseguenza abbiamo che ogni vettore $v \in \mathbb{R}^2$ verifica $v \in \mathrm{Span}(v_1, v_2)$, quindi

$$\mathbb{R}^2 = \mathrm{Span}(v_1, v_2),$$

cioé v_1 e v_2 sono dei generatori per $\mathbb{R}^2$.

- Consideriamo i due vettori

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \qquad v_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \in \mathbb{R}^3.$$

  In questo caso abbiamo che

$$\mathbb{R}^3 \neq \operatorname{Span}(v_1, v_2).$$

  Infatti, il vettore

$$v = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \notin \operatorname{Span}(v_1, v_2).$$

  Infatti, verifichiamo che non esistono λ_1 e λ_2 tali che

$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \lambda_1 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ \lambda_1 + \lambda_2 \\ \lambda_2 \end{pmatrix}.$$

  La matrice completa associata al sistema lineare è

$$(A|b) = \begin{pmatrix} 1 & 0 & | & 1 \\ 1 & 1 & | & 1 \\ 0 & 1 & | & 1 \end{pmatrix}.$$

  Con l' operazione elementare $(A|b)_{(2)} - (A|b)_{(1)}$ otteniamo

$$(A'|b') = \begin{pmatrix} 1 & 0 & | & 1 \\ 0 & 1 & | & 0 \\ 0 & 1 & | & 1 \end{pmatrix}$$

  e ora con $(A'|b')_{(3)} - (A'|b')_{(2)}$ abbiamo la seguente matrice a scala:

$$(A''|b'') = \begin{pmatrix} 1 & 0 & | & 1 \\ 0 & 1 & | & 0 \\ 0 & 0 & | & 1 \end{pmatrix},$$

  che corrisponde a un sistema lineare incompatibile (senza soluzioni), perché l'ultima riga corrisponde ad un'equazione del tipo $0 = 1$.

Esempio 5.1.5. Vediamo un esempio di due sottospazi vettoriali, la cui unione non è un sottospazio vettoriale: sia $V = \mathbb{R}^2$ e siano

$$U = \operatorname{Span}\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = \{a \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} \mid a \in \mathbb{R}\} = \{\begin{pmatrix} a \\ 0 \end{pmatrix} \mid a \in \mathbb{R}\},$$

e

$$W = \operatorname{Span}\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right) = \{b \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} \mid b \in \mathbb{R}\} = \{\begin{pmatrix} b \\ b \end{pmatrix} \mid b \in \mathbb{R}\}.$$

Osserviamo che i vettori

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \in U \cup W,$$

ma

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \notin U \cup W.$$

Infatti, si ha $\begin{pmatrix} 2 \\ 1 \end{pmatrix} \notin U$ perché non è del tipo $\begin{pmatrix} a \\ 0 \end{pmatrix}$ per alcun $a \in \mathbb{R}$. Inoltre, $\begin{pmatrix} 2 \\ 1 \end{pmatrix} \notin W$ perché non è del tipo $\begin{pmatrix} b \\ b \end{pmatrix}$ per alcun $b \in \mathbb{R}$.

Definizione 5.1.6. Sia V uno spazio vettoriale su un campo $\mathbb{K}$ e siano

$$v_1, \dots, v_k \in V$$

vettori di V. Diremo che $v_1, \dots, v_k$ **generano** V se

$$V = \mathrm{Span}\,(v_1, \dots, v_k),$$

e quindi per ogni vettore $v \in V$ esistono dei coefficienti $\lambda_1, \dots, \lambda_k$ tali che

$$v = \lambda_1\, v_1 + \dots + \lambda_k\, v_k.$$

In questo caso V si dirà **finitamente generato**.

Osservazione 5.1.2. Nel caso in cui $V = \mathbb{K}^n$, possiamo stabilire se k vettori $v_1, \dots, v_k$ sono dei generatori per $\mathbb{K}^n$ studiando un sistema lineare. Più precisamente, se esprimiamo i vettori $v_1, \dots, v_k$ in componenti:

$$v_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix}, \qquad v_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix}, \qquad \dots, \qquad v_k = \begin{pmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{nk} \end{pmatrix},$$

allora un vettore $v = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$ è una loro combinazione lineare se si può scrivere come

$$\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = c_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix} + c_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix} + \dots + c_k \begin{pmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{nk} \end{pmatrix} = \begin{pmatrix} c_1 a_{11} + c_2 a_{12} + \dots + c_k a_{1k} \\ c_1 a_{21} + c_2 a_{22} + \dots + c_k a_{2k} \\ \vdots \\ c_1 a_{n1} + c_2 a_{n2} + \dots + c_k a_{nk} \end{pmatrix} =$$

$$= \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1k} \\ a_{21} & a_{22} & \dots & a_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nk} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{pmatrix}.$$

Se poniamo

$$A := \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1k} \\ a_{21} & a_{22} & \dots & a_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nk} \end{pmatrix},$$

abbiamo che v è combinazione lineare di $v_1, \ldots, v_k$ se e solo se il sistema lineare

$$A \cdot X = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

ammette almeno una soluzione $\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{pmatrix}$, cioè se e solo se $A \cdot X = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$ è compatibile.

5.2 Dipendenza e indipendenza lineare

Definizione 5.2.1. Siano $v_1, \ldots, v_k \in V$; essi si dicono **linearmente dipendenti** se uno di loro si può scrivere come combinazione lineare degli altri.

Esempio 5.2.2. I vettori di $\mathbb{R}^2$

$$v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix},\ v_2 \begin{pmatrix} 1 \\ 2 \end{pmatrix},\ v_3 = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$$

sono linearmente dipendenti, perché $v_3 = 2v_1 + v_2$, ma v_1 e v_2 non sono linearmente dipendenti, e nemmeno v_1 e v_3, così come v_2 e v_3 non sono linearmente dipendenti (verificare).

Definizione 5.2.3. Siano $v_1, \ldots, v_k \in V$; essi si dicono **linearmente indipendenti** se non sono linearmente dipendenti.

La dipendenza e l'indipendenza lineari possono essere caratterizzate nel modo seguente, che può essere scelto come definizione alternativa:

Proposizione 5.2.4. *I vettori $v_1, \ldots, v_k \in V$ sono linearmente dipendenti $\iff$ esiste una loro combinazione lineare con coefficienti non tutti nulli che dia il vettore nullo:*

$$a_1v_1 + a_2v_2 + \cdots + a_kv_k = 0, \quad a_i \text{ non tutti nulli.}$$

Conseguentemente abbiamo:

Proposizione 5.2.5. *I vettori $v_1, \ldots, v_k \in V$ sono linearmente indipendenti $\iff$ l' unica loro combinazione lineare che dia il vettore nullo è quella con tutti i coefficienti nulli:*

$$a_1v_1 + a_2v_2 + \cdots + a_kv_k = 0 \Rightarrow a_1 = a_2 = \cdots = a_k = 0.$$

Osservazione 5.2.1.

1. Nel caso $k = 1$, abbiamo che un vettore v_1 è linearmente dipendente $\iff$ esiste una combinazione lineare

   $$a_1v_1 = 0,$$

   con $a_1 \neq 0$, cioè $\iff$ $v_1 = 0$ è il vettore nullo.

   Conseguentemente, v_1 è linearmente indipendente $\iff$ $v_1 \neq 0$.

2. Due vettori v_1 e v_2 sono linearmente dipendenti $\iff$ esiste $c \in \mathbb{K}$ tale che

$$v_2 = cv_1, \quad \text{oppure } v_1 = cv_2,$$

cioè $\iff$ v_1 e v_2 sono **proporzionali**.

Come conseguenza, due vettori v_1 e v_2 sono linearmente indipendenti $\iff$ non sono proporzionali.

3. Consideriamo dei vettori $v_1, \dots, v_k$ e supponiamo che uno di essi sia nullo:

$$v_i = 0.$$

Allora $v_1, \dots, v_k$ sono linearmente dipendenti; infatti si ha

$$0 \cdot v_1 + \cdots + 1 \cdot v_i + \cdots + 0 \cdot v_k = 0$$

è una loro combinazione lineare con coefficienti non tutti nulli, che dà il vettore nullo.

4. Se tra i vettori $v_1, \dots, v_k$ ce ne sono due uguali

$$v_i = v_j,$$

allora $v_1, \dots, v_k$ sono linearmente dipendenti; infatti, la combinazione lineare

$$0 \cdot v_1 + \cdots + 1 \cdot v_i + \cdots + (-1)v_j + \cdots + 0 \cdot v_k =$$

$$= 0 \cdot v_1 + \cdots + 1 \cdot v_i + \cdots + (-1)v_i + \dots 0 \cdot v_k = 0$$

dà luogo al vettore nullo e non tutti i coefficienti sono nulli.

Osservazione 5.2.2. Nel caso in cui $V = \mathbb{K}^n$, possiamo stabilire se k vettori $v_1, \dots, v_k$ sono linearmente dipendenti studiando un sistema omogeneo di equazioni lineari. Più precisamente, se esprimiamo i vettori $v_1, \dots, v_k$ in componenti:

$$v_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix}, \quad v_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix}, \quad \dots, \quad v_k = \begin{pmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{nk} \end{pmatrix},$$

allora una loro combinazione lineare si può scrivere come

$$c_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix} + c_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix} + \cdots + c_k \begin{pmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{nk} \end{pmatrix} = \begin{pmatrix} c_1a_{11} + c_2a_{12} + \cdots + c_ka_{1k} \\ c_1a_{21} + c_2a_{22} + \cdots + c_ka_{2k} \\ \vdots \\ c_1a_{n1} + c_2a_{n2} + \cdots + c_ka_{nk} \end{pmatrix} =$$

$$= \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1k} \\ a_{21} & a_{22} & \dots & a_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nk} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{pmatrix}.$$

Se poniamo

$$A := \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1k} \\ a_{21} & a_{22} & \dots & a_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nk} \end{pmatrix},$$

abbiamo che $v_1, \dots, v_k$ sono linearmente dipendenti se e solo se il sistema lineare omogeneo

$$A \cdot X = 0$$

ammette una soluzione non banale $\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{pmatrix}$.

Infine, possiamo equivalentemente affermare che $v_1, \dots, v_k$ sono linearmente indipendenti se e solo se il sistema lineare omogeneo

$$A \cdot X = 0$$

ammette solo la soluzione banale (la soluzione nulla).

5.3 Basi

Abbiamo osservato che quando uno dei vettori considerati è combinazione lineare degli altri, esso è irrilevante ai fini dello spazio generato. È quindi naturale cercare di considerare solo insiemi *minimali* di generatori, cioè insiemi in cui togliendo un qualunque vettore, lo spazio generato dai rimanenti è strettamente più piccolo. Queste considerazioni inducono a dare la seguente definizione.

Definizione 5.3.1. Sia V uno spazio vettoriale su un campo $\mathbb{K}$. Un sottoinsieme di vettori

$$\mathcal{B} = \{v_1, \dots, v_n\}$$

si dice **base di** V (finita) se valgono le seguenti:

1. (B1) $v_1, \dots, v_n$ generano V;
2. (B2) $v_1, \dots, v_n$ sono linearmente indipendenti.

Proposizione 5.3.2. *Un sottoinsieme $\{v_1, \dots, v_n\}$ di V è una base di V se e solo se per ogni ettore $v \in V$, esistono e sono unici dei coefficienti $\lambda_1, \dots, \lambda_n$ tali che*

$$v = \lambda_1 v_1 + \dots + \lambda_n v_n.$$

In tal caso gli scalari $\lambda_1, \dots, \lambda_n$ si chiamano **coordinate di** v **nella base** $\{v_1, \dots, v_n\}$.

Dimostrazione. Sia $\{v_1, \dots, v_n\}$ una base di V. Poiché sono dei generatori di V, ogni vettore $v \in V$ si può scrivere come combinazione lineare

$$v = \lambda_1 v_1 + \dots + \lambda_n v_n.$$

Supponiamo che si abbia anche

$$v = \mu_1 v_1 + \cdots + \mu_n v_n.$$

Sottraendo la seconda equazione alla prima otteniamo

$$0 = (\lambda_1 - \mu_1)v_1 + \cdots + (\lambda_n - \mu_n)v_n.$$

Essendo $\{v_1, \ldots, v_n\}$ una base, i vettori sono anche linearmente indipendenti, quindi si ha

$$\lambda_1 = \mu_1, \ldots, \lambda_n = \mu_n,$$

da cui la tesi.

Viceversa, supponiamo che ogni vettore di V si possa esprimere in modo unico nella forma

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

Da ciò segue, in particolare, che i vettori $\{v_1, \ldots, v_n\}$ formano un insieme di generatori di V.

Mostriamo infine che sono linearmenet indipendenti. Consideriamo una loro combinazione lineare che dia il vettore nullo:

$$c_1 v_1 + \cdots + c_n v_n = 0.$$

Siccome il vettore nullo ammette la rappresentazione

$$0 = 0 \cdot v_1 + \cdots + 0 \cdot v_n,$$

per l'ipotesi di unicità della rappresentazione di ogni vettore come combinazione lineare dei vettori $\{v_1, \ldots, v_n\}$, si ha che necessariamente

$$c_1 = 0, \ldots, c_n = 0,$$

quindi $v_1, \ldots, v_n$ sono anche linearmente indipendenti, e formano una base. $\square$

Esempio 5.3.3. Sia $V = \mathbb{K}^n$. Allora si ha una base naturale, detta **base canonica $\mathcal{E}$ di $\mathbb{K}^n$:**

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \; e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \; \ldots, e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Infatti, i vettori di $\mathcal{E}$ formano un insieme di generatori; dato un vettore

$$v = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

arbitrario, esso si esprime come combinazione lineare nel modo seguente:

$$\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = b_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + b_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \cdots + b_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Inoltre, vediamo subito che i vettori di $\mathcal{E}$ sono linearmente indipendenti; infatti, usando il criterio 5.2.2, osserviamo che la matrice A le cui colonne sono formate da $e_1, \ldots, e_n$ corrisponde alla matrice unità $A = \mathbb{I}_n$, e il sistema lineare omogeneo

$$\mathbb{I}_n \cdot X = 0$$

ammette solo la soluzione nulla.

Osservazione 5.3.1. Sottolineiamo nell' esempio precedente che nella base canonica $\mathcal{E}$ di $\mathbb{K}^n$ le **coordinate di un vettore coincidono con le sue componenti**.

Esempio 5.3.4. Una base per lo spazio vettoriale delle matrici $M_{m,n}(\mathbb{K})$ è data da

$$E_{11} := \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & \ldots & 0 \end{pmatrix}, \; E_{12} := \begin{pmatrix} 0 & 1 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & \ldots & 0 \end{pmatrix}, \; \ldots, E_{mn} := \begin{pmatrix} 0 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & \ldots & 1 \end{pmatrix},$$

ovvero dalle matrici E_{ij} che hanno il coefficiente 1 nella posizione i, j e zero altrove.

Infatti, ogni matrice $A = (a_{ij})$ si può scrivere in modo unico come

$$A = a_{11}E_{11} + a_{12}E_{12} + \cdots + a_{mn}E_{mn} = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}E_{ij}.$$

Anche questa base viene detta **base canonica dello spazio delle matrici**.

5.4 Teoremi di estrazione e di completamento

Teorema 5.4.1. Estrazione di una base da un insieme di generatori *Sia V uno spazio vettoriale su un campo $\mathbb{K}$, e siano $v_1, \ldots, v_k$ dei generatori per V. Allora esiste una base $\mathcal{B} \subseteq \{v_1, \ldots, v_k\}$ di V.*

Per la dimostrazione di questo risultato abbiamo bisogno di due Lemmi.

Lemma 5.4.2. *Siano $u_1, \ldots, u_l \in V$ e sia*

$$u \in \operatorname{Span}(u_1, \ldots, u_l).$$

Allora $\operatorname{Span}(u_1, \ldots, u_l) = \operatorname{Span}(u_1, \ldots, u_l, u)$.

Dimostrazione. L' inclusione $\operatorname{Span}(u_1, \ldots, u_l) \subseteq \operatorname{Span}(u_1, \ldots, u_l, u)$ è sempre verificata.

Dimostriamo quindi l'altra inclusione. Sia $w \in \operatorname{Span}(u_1, \ldots, u_l, u)$ un vettore arbitrario. Quindi

$$w = b_1u_1 + \cdots + b_lu_l + bu, \tag{5.4.1}$$

per opportuni coefficienti $b_1, \ldots, b_l, b \in \mathbb{K}$.

Per ipotesi su u possiamo scrivere

$$u = c_1u_1 + \cdots + c_lu_l.$$

Quindi la (5.4.1) può essere riscritta nella forma

$$w = b_1u_1 + \cdots + b_lu_l + b(c_1u_1 + \cdots + c_lu_l) = (b_1 + bc_1)u_1 + \cdots + (b_l + bc_l)u_l,$$

ovvero w si può esprimere anche come combinazione lineare dei vettori $u_1, \ldots, u_l$. Quindi abbiamo dimostrato che $\operatorname{Span}(u_1, \ldots, u_l) \supseteq \operatorname{Span}(u_1, \ldots, u_l, u)$. □

Lemma 5.4.3. *Siamo $u_1, \ldots, u_h \in V$ dei vettori linearmente indipendenti, e sia*

$$u \notin \operatorname{Span}(u_1, \ldots, u_h).$$

Allora $u_1, \ldots, u_h, u$ sono linearmente indipendenti.

Dimostrazione. Supponiamo per assurdo che $u_1, \ldots, u_l, u$ siano linearmente dipendenti e sia

$$c_1 u_1 + \cdots + c_h u_h + cu = 0 \tag{5.4.2}$$

una combinazione lineare con coefficienti non tutti nulli, che dia il vettore nullo.

Osserviamo che se $c = 0$, allora abbiamo una combinazione lineare

$$c_1 u_1 + \cdots + c_h u_h = 0$$

a coefficienti non tutti nulli, e questo è un assurdo perché si contraddice l'ipotesi $u_1, \ldots, u_h$ linearmente indipendenti.

Quindi supponiamo che sia $c \neq 0$; dalla (5.4.2) otteniame

$$u = \frac{c_1}{c} u_1 + \ldots \frac{c_h}{c} u_h,$$

cioè u è combinazione lineare di $u_1, \ldots, u_h$, e questo è un assurdo perché contraddice l'ipotesi $u \notin \operatorname{Span}(u_1, \ldots, u_h)$.

□

Dimostrazione. **del Teorema di Estrazione** Consideriamo il seguente **algoritmo**:

- consideriamo v_1:
  - se $v_1 = 0$, lo scartiamo;
  - se $v_1 \neq 0$, lo scegliamo: $v_1 \in \mathcal{B}$;
- consideriamo v_2:
  - se $v_2 \in \operatorname{Span}(v_1)$, lo scartiamo; per il Lemma 5.4.2, si ha $\operatorname{Span}(v_1) = \operatorname{Span}(v_1, v_2)$;
  - se $v_2 \notin \operatorname{Span}(v_1)$, lo scegliamo: $v_2 \in \mathcal{B}$; per il Lemma 5.4.3, si ha che i vettori finora scelti sono linearmente indipendenti;
- consideriamo v_3:
  - se $v_3 \in \operatorname{Span}(v_1, v_2)$, lo scartiamo; per il Lemma 5.4.2, si ha $\operatorname{Span}(v_1, v_2) = \operatorname{Span}(v_1, v_2, v_3)$
  - se $v_3 \notin \operatorname{Span}(v_1, v_2)$, lo scegliamo: $v_3 \in \mathcal{B}$; per il Lemma 5.4.3, si ha che i vettori finora scelti sono linearmente indipendenti;
- consideriamo v_i:
  - se $v_i \in \operatorname{Span}(v_1, v_2, \ldots, v_{i-1})$, lo scartiamo; per il Lemma 5.4.2, si ha $\operatorname{Span}(v_1, v_2, \ldots, v_{i-1}) = \operatorname{Span}(v_1, v_2, \ldots, v_{i-1}, v_i)$;
  - se $v_2 \notin \operatorname{Span}(v_1, v_2, \ldots, v_{i-1})$, lo scegliamo: $v_i \in \mathcal{B}$; per il Lemma 5.4.3, si ha che i vettori finora scelti sono linearmente indipendenti.

In questo modo, dopo k passi, l'algoritmo si conclude. Per il Lemma 5.4.3, si ha che i vettori scelti sono linearmente indipendenti. Inoltre, per il Lemma 5.4.2 i vettori scartati non modificano lo spazio generato, quindi i vettori scelti formano ancora un insieme di generatori per V.

□

Corollario 5.4.4. *Ogni spazio vettoriale finitamente generato ammette una base.*

Osservazione 5.4.1. Non tutti gli spazi vettoriali sono finitamente generati. Ad esempio lo spazio dei polinomi in una indeterminata a coefficienti reali

$$V = \mathbb{R}[x]$$

non lo è .

Infatti, se per assurdo fosse

$$\mathbb{R}[x] = \mathrm{Span}(P_1(x), \ldots, P_k(x)),$$

ponendo

$$N := \max\{\text{grado } P_i(x) \mid i = 1, \ldots, k\},$$

avremmo che $\mathbb{R}[x]$ contiene solo polinomi di grado $\leq N$.

Teorema 5.4.5. Completamento di vettori linearmente indipendenti a una base *Siano $v_1, \ldots, v_p \in V$ dei vettori linearmente indipendenti e supponiamo che V sia finitamente generato.*

Allora esiste una base $\mathcal{B}$ di V tale che

$$\{v_1, \ldots, v_p\} \subseteq \mathcal{B}.$$

Diremo che abbiamo **completato** *l'insieme $\{v_1, \ldots, v_p\}$ a una base di V.*

Dimostrazione. Essendo V finitamente generato per ipotesi, posiamo scegliere dei generatori $w_1, \ldots, w_r$ per V:

$$V = \mathrm{Span}(w_1, \ldots, w_r).$$

Aggiungendo i vettori $v_1, \ldots, v_p$, l' insieme ottenuto forma ancora un insieme di generatori:

$$V = \mathrm{Span}(v_1, \ldots, v_p, w_1, \ldots, w_r).$$

Applichiamo ora il Teorema di Estrazione di una base da $v_1, \ldots, v_p, w_1, \ldots, w_r$ e osserviamo che nell' algoritmo della dimostrazione, per costruzione i vettori $v_1, \ldots, v_p$ sono scelti. Quindi nella costruzione della base $\mathcal{B}$ si avrà necessariamente

$$\{v_1, \ldots, v_p\} \subseteq \mathcal{B}.$$

□

5.5 Dimensione

In questa sezione vogliamo definire la dimensione di uno spazio vettoriale finitamente generato. Per questo proposito vediamo un risultato che ci permetterà di dedurre che ogni base di uno spazio vettoriale finitamente generato ha lo stesso numero di elementi.

Lemma 5.5.1. di Steinitz *Sia V uno spazio vettoriale e sia $\{v_1, \dots, v_n\}$ una base di V.*
Allora $\forall p > n$ e per ogni scelta di vettori $w_1, \dots, w_p$, essi sono linearmente dipendenti.

Dimostrazione. I vettori $w_1, \dots, w_p$ si scrivono in modo unico come combinazioni lineari dei vettori della base $\{v_1, \dots, v_n\}$:

$$\begin{array}{rl} w_1 = & c_{11}v_1 + c_{21}v_2 + \cdots + c_{n1}v_n, \\ w_2 = & c_{12}v_1 + c_{22}v_2 + \cdots + c_{n2}v_n, \\ \vdots & \vdots \\ w_p = & c_{1p}v_1 + c_{2p}v_2 + \cdots + c_{np}v_n. \end{array}$$

Denotiamo con $\begin{pmatrix} c_{1i} \\ c_{2i} \\ \vdots \\ c_{ni} \end{pmatrix}$ la colonna delle coordinate di w_i.

Una combinazione lineare $a_1w_1+a_2w_2+\cdots+a_pw_p = 0$ dà il vettore nullo se e solo se le coordinate dei vettori w_i soddisfano

$$a_1 \begin{pmatrix} c_{11} \\ c_{21} \\ \vdots \\ c_{n1} \end{pmatrix} + a_2 \begin{pmatrix} c_{12} \\ c_{22} \\ \vdots \\ c_{n2} \end{pmatrix} + \cdots + a_p \begin{pmatrix} c_{1p} \\ c_{2p} \\ \vdots \\ c_{np} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

quindi se e solo se

$$\begin{pmatrix} c_{11} & c_{12} & \dots & c_{1p} \\ c_{21} & c_{22} & \dots & c_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ c_{n1} & c_{n2} & \dots & c_{np} \end{pmatrix} \cdot \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{5.5.1}$$

Quindi $w_1, \dots, w_p$ sono linearmente dipendenti se e solo se il sistema lineare omogeneo

$$\begin{pmatrix} c_{11} & c_{12} & \dots & c_{1p} \\ c_{21} & c_{22} & \dots & c_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ c_{n1} & c_{n2} & \dots & c_{np} \end{pmatrix} \cdot X = 0$$

di n equationi in p incognite ha una soluzione non nulla.

Per l' ipotesi $p > n$, vediamo che la generica soluzione dipende da almeno $p - n \geq 1$ parametri; fissando per tali parametri dei valori non nulli, si ottiene una soluzione non nulla del sistema lineare omogeneo. □

Come conseguenza si ha il seguente importante risultato.

Teorema 5.5.2. *Sia V uno spazio vettoriale. Se $\{v_1, \dots, v_n\}$ e $\{w_1, \dots, w_m\}$ sono due basi di V, allora*

$$n = m.$$

Dimostrazione. Essendo $\{v_1, \dots, v_n\}$ e $w_1, \dots, w_m$ linearmente indipendenti, per il Lemma 5.5.1 di Steinitz si ha

$$m \leq n.$$

Infatti, se si avesse $m > n$, i vettori $w_1, \dots, w_m$ sarebbero linearmente dipendenti.

Analogamente, essendo $\{w_1, \dots, w_m\}$ una base e $v_1, \dots, v_n$ linearmente indipendenti, si ha

$$n \leq m,$$

da cui la tesi.

□

Definizione 5.5.3. Dato uno spazio vettoriale V, definiamo la **dimensione di** V come segue:

- se $V = \{0\}$, poniamo $\dim V := 0$;
- se $V \neq \{0\}$ e V è finitamente generato, poniamo $\dim V :=$ numero di vettori di una sua qualunque base.

Esempio 5.5.4. Per $V = \mathbb{K}^n$ abbiamo visto che c'è la base canonica $\mathcal{E}$, che consta di n vettori, quindi

$$\dim \mathbb{K}^n = n.$$

Esempio 5.5.5. In $V = M_{m,n}(\mathbb{K})$ c'è la base canonica, che consta di $m \cdot n$ vettori, quindi

$$\dim M_{m,n}(\mathbb{K}) = m \cdot n.$$

Esempio 5.5.6. Consideriamo il campo complesso $\mathbb{C}$. Esso è uno spazio vettoriale su $\mathbb{C}$ stesso, e come tale ha dimensione

$$\dim \mathbb{C} = 1.$$

Osserviamo, pero', che $\mathbb{C}$ si può dotare anche di una struttura di spazio vettoriale su $\mathbb{R}$ come segue:

$$+ \; : \mathbb{C} \times \mathbb{C} \to \mathbb{C},$$

dove la somma è quella usuale tra numeri complessi, e la moltiplicazione per uno scalare reale è definita da

$$\cdot \; : \mathbb{R} \times \mathbb{C} \to \mathbb{C}, \quad c \cdot (a + ib) = ac + ibc.$$

Con questa struttura, una base di $\mathbb{C}$ è data da

$$\{1, i\},$$

e quindi

$$\dim_{\mathbb{R}} \mathbb{C} = 2.$$

Vediamo ora che in uno spazio di cui si conosce la dimensione, per individuare una sua base non è necessario verificare sia di avere dei generatori sia l'indipendenza lineare.

Proposizione 5.5.7. *Sia V uno spazio vettoriale di dimensione*

$$\dim V = n.$$

Allora valgono le seguenti:

1. *se $v_1, \dots, v_n$ sono linearmente indipendenti $\Rightarrow$ $v_1, \dots, v_n$ formano una base di V; in particolare, sono anche dei generatori per V.*

2. *se $v_1, \dots, v_n$ sono dei generatori per $V \Rightarrow$ $v_1, \dots, v_n$ formano una base di V, in particolare sono anche linearmente indipendenti.*

Dimostrazione. 1. Siccome $v_1, \dots, v_n$ sono linearmente indipendenti, per il Teorema di Completamento si possono completare a una base di V. Essendo $\dim V = n$, ogni base di V ha esattamente n vettori, quindi non è necessario aggiungere alcun vettore all' insieme $\{v_1, \dots, v_n\}$, che risulta una base. In particolare, $\{v_1, \dots, v_n\}$ è un insieme di generatori.

2. Se $v_1, \dots, v_n$ sono un insieme di generatori per V, per il Teorema di Estrazione da essi si può estrarre una base di V. Essendo $\dim V = n$, ogni base di V ha esattamente n vettori, quindi non è necessario scartare alcun vettore dall' insieme $\{v_1, \dots, v_n\}$, che risulta una base. In particolare, $v_1, \dots, v_n$ sono linearmente indipendenti.

□

Esempio 5.5.8. In particolare, sapendo che $\dim \mathbb{R}^2 = 2$, per trovare una base di $\mathbb{R}^2$ è sufficiente scegliere 2 vettori linearmente indipendenti, cioè 2 vettori non nulli e non proporzionali.

5.6 Dimensione di sottospazi vettoriali

Sia V uno spazio vettoriale di dimensione finita, e sia $W \subseteq V$ un suo sottospazio vettoriale. In questa sezione daremo una limitazione sulla dimensione di W.

Osservazione 5.6.1. Sia $W \subseteq V$ un sottospazio vettoriale di V spazio vettoriale su $\mathbb{K}$, e consideriamo W come spazio vettoriale su $\mathbb{K}$. Se $w_1, \dots, w_k \in W$ sono vettori linearmente indipendenti in $W \Rightarrow$ $w_1, \dots, w_k$ sono linearmente indipendenti anche in V.

Infatti, essendo entrambi W e V spazi vettoriali su $\mathbb{K}$, la condizione di indipendenza lineare come vettori di W o di V è la stessa.

Osservazione 5.6.2. Se V è finitamente generato e $W \subseteq V$ è un suo sottospazio vettoriale $\Rightarrow$ anche W è finitamente generato.

Infatti, sia $\dim V = n$, e fissiamo $w_1, \dots, w_k \in W$ vettori linearmente indipendenti in W. Se $W = \mathrm{Span}(w_1, \dots, w_k)$, allora W è finitamente generato.

Se $W \supsetneq \mathrm{Span}(w_1, \dots, w_k)$, allora possiamo scegliere un vettore

$$w_{k+1} \notin \mathrm{Span}(w_1, \dots, w_k), \quad w_{k+1} \in W.$$

Per il Lemma 5.4.3, i vettori $w_1, \dots, w_k, w_{k+1}$ sono linearmente indipendenti. Ripetiamo il procedimento. Siccome per l' Osservazione 5.6.1 vettori linearmente indipendenti in W sono anche linearmente indipendenti in V, e siccome in V ci sono al più n vettori linearmente indipendenti per il Lemma 5.5.1 di Steinitz, il procedimento termina dopo un numero finito di passi. Quindi troviamo un numero finito di vettori di W che generano W.

Corollario 5.6.1. *Ogni sottospazio vettoriale W di V, spazio vettoriale finitamente generato, è del tipo*

$$W = \mathrm{Span}(w_1, \dots, w_m)$$

per opportui vettori $w_1, \dots, w_m \in W$.

Proposizione 5.6.2. *Sia V uno spazio vettoriale finitamente generato e sia $W \subseteq V$ un sottospazio vettoriale. Allora valgono:*

1. $\dim W \leq \dim V$;
2. $\dim W = \dim V \iff W = V$.

Dimostrazione. 1. Sia $\{w_1, \dots, w_k\}$ una base di W. I vettori $w_1, \dots, w_k$ sono linearmente indipendenti anche in V, per l' Osservazione 5.6.1. Per il Teorema di Completamento, possiamo completare l' insieme $\{w_1, \dots, w_k\}$ a una base $\mathcal{B}$ di V. Quindi si ha

$$\dim W = \#\{w_1, \dots, w_k\} \leq \#\mathcal{B} = n = \dim V.$$

2. Se $W = V$, è chiaro che hanno la stessa dimensione.

   Viceversa, supponiamo $\dim W = \dim V = n$, e fissiamo una base $\{w_1, \dots, w_n\}$ di W. Per l' Osservazione 5.6.1 i vettori $w_1, \dots, w_n$ sono linearmente indipendenti anche in V. Infine, essi formano una base di V per la Proposizione 5.5.7, primo punto. Quindi

$$W = \mathrm{Span}(w_1, \dots, w_n) = V.$$

□

5.7 Formula di Grassmann

In questa sezione vediamo una formula che lega la dimensione di un'intersezione di sottospazi vettoriali con la dimensione del sottospazio somma.

Teorema 5.7.1. Formula di Grassmann *Siano*

$$W_1 \subseteq V, \quad W_2 \subseteq V$$

due sottospazi vettoriali di uno spazio vettoriale finitamente generato V. Allora vale

$$\dim(W_1 \cap W_2) = \dim W_1 + \dim W_2 - \dim(W_1 + W_2). \tag{5.7.1}$$

Dimostrazione. Fissiamo una base di $W_1 \cap W_2$:

$$\mathcal{B}_{W_1 \cap W_2} = \{w_1, \dots, w_r\}, \quad r = \dim W_1 \cap W_2.$$

Essendo $W_1 \cap W_2 \subseteq W_1$ un sottospazio vettoriale, per il Teorema del Completamento, possiamo completare i vettori $w_1, \dots, w_r$ ad una base di W_1:

$$\mathcal{B}_{W_1} = \{w_1, \dots, w_r, v_1, \dots, v_s\}, \quad r + s = \dim W_1.$$

Analogamente, essendo $W_1 \cap W_2 \subseteq W_2$ un sottospazio vettoriale, per il Teorema del Completamento, possiamo completare i vettori $w_1, \dots, w_r$ ad una base di W_2:

$$\mathcal{B}_{W_2} = \{w_1, \dots, w_r, u_1, \dots, u_k\}, \quad r + k = \dim W_2.$$

Vogliamo dimostrare che

$$\dim(W_1 + W_2) = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2) = r + s + r + k - r = r + s + k.$$

A tale scopo affermiamo che

$$\mathcal{B}_{W_1} \cup \mathcal{B}_{W_2} = \{w_1, \dots, w_r, v_1, \dots, v_s, u_1, \dots, u_k\}$$

è una base di $W_1 + W_2$.

Infatti, $\{w_1, \dots, w_r, v_1, \dots, v_s, u_1, \dots, u_k\}$ sono dei generatori per $W_1 + W_2$: sia $w \in W_1 + W_2$; per definizione di spazio somma, w si può scrivere nella forma

$$w = z_1 + z_2, \quad z_1 \in W_1, \quad z_2 \in W_2,$$

per opportuni vettori z_1 e z_2. I due vettori, a loro volta, si possono scrivere come combinazioni lineari delle basi di W_1 e W_2, rispettivamente:

$$z_1 = a_1 w_1 + \dots + a_r w_r + b_1 v_1 + \dots + b_s v_s, \quad z_2 = c_1 w_1 + \dots + c_r w_r + d_1 u_1 + \dots + d_k u_k,$$

e quindi

$$w = (a_1 + c_1) w_1 + \dots + (a_r + c_r) w_r + b_1 v_1 + \dots + b_s v_s + d_1 u_1 + \dots + d_k u_k,$$

cioè ogni vettore di $W_1 + W_2$ è combinazione lineare dei $\{w_1, \dots, w_r, v_1, \dots, v_s, u_1, \dots, u_k\}$.

Mostriamo, infine, che $w_1, \dots, w_r, v_1, \dots, v_s, u_1, \dots, u_k$ sono linearmente indpendenti. Consideri-mao una loro combinazione lineare che dia il vettore nullo:

$$\alpha_1 w_1 + \dots + \alpha_r w_r + \beta_1 v_1 + \dots + \beta_s v_s + \delta_1 u_1 + \dots + \delta_k u_k = 0. \tag{5.7.2}$$

Questa relazione può essere riscritta nella forma

$$\alpha_1 w_1 + \dots + \alpha_r w_r + \beta_1 v_1 + \dots + \beta_s v_s = -\delta_1 u_1 - \dots - \delta_k u_k,$$

quindi il vettore $-\delta_1 u_1 - \dots - \delta_k u_k \in W_2$ è anche combinazione lineare dei vettori della base di W_1, quindi

$$-\delta_1 u_1 - \dots - \delta_k u_k \in W_1 \cap W_2.$$

Come conseguenza può essere scritto come combinazione lineare dei vettori della base $\mathcal{B}_{W_1 \cap W_2}$:

$$-\delta_1 u_1 - \dots - \delta_k u_k = \gamma_1 w_1 + \dots + \gamma_r w_r,$$

da cui

$$\gamma_1 w_1 + \dots + \gamma_r w_r + \delta_1 u_1 + \dots + \delta_k u_k = 0.$$

Quest'ultima è una combinazione lineare dei vettori di $\mathcal{B}_{W_2}$, che sono linearmente indipendenti, quindi

$$\gamma_1 = \dots = \gamma_r = \delta_1 = \dots = \delta_k = 0.$$

Quindi la relazione (5.7.2) diventa

$$\alpha_1 w_1 + \cdots + \alpha_r w_r + \beta_1 v_1 + \cdots + \beta_s v_s = 0,$$

che è una combinazione lineare dei vettori di $\mathcal{B}_{W_1}$, che sono linearmente indipendenti, quindi

$$\alpha_1 = \cdots = \alpha_r = \beta_1 = \cdots = \beta_s = 0.$$

□

Corollario 5.7.2. *Siano*

$$W_1 \subseteq V, \quad W_2 \subseteq V$$

due sottospazi vettoriali di uno spazio vettoriale con $\dim V = n$. *Allora vale*

$$\dim(W_1 \cap W_2) \geq \dim W_1 + \dim W_2 - n. \tag{5.7.3}$$

Dimostrazione. Basta osservare che essendo $W_1 + W_2 \subseteq V$, per la Proposizione 5.6.2, primo punto, si ha

$$\dim W_1 + W_2 \leq n.$$

□

Esempio 5.7.3. In particolare, se W_1 e W_2 sono due piani vettoriali di $\mathbb{R}^3$, abbiamo che

$$\dim(W_1 \cap W_2) \geq 2 + 2 - 3 = 1,$$

cioè due piani vettoriali si intersecano sempre almeno lungo una retta vettoriale.

Capitolo 6

Rango: definizione e prime proprietà

Definizione 6.0.1. Sia $A \in M_{m,n}(\mathbb{K})$ una matrice di tipo $m \times n$ a coefficienti nel campo $\mathbb{K}$. Il **rango** di A è la dimensione del sottospazio vettoriale di $\mathbb{K}^m$ generato dalle colonne di A:

$$\mathrm{rg}(A) := \dim \mathrm{Span}(A^{(1)}, A^{(2)}, \dots, A^{(n)}), \quad \mathrm{Span}(A^{(1)}, A^{(2)}, \dots, A^{(n)}) \subseteq \mathbb{K}^m.$$

In diversi libri di testo, il rango definito qui sopra viene chiamato rango per colonne di A, mentre il rango per righe si definisce come la dimensione del sottospazio vettoriale di $\mathbb{K}^n$ generato dalle righe di A, cioè

$$\dim \mathrm{Span}(A_{(1)}, A_{(2)}, \dots, A_{(m)}), \quad \mathrm{Span}(A_{(1)}, A_{(2)}, \dots, A_{(m)}) \subseteq \mathbb{K}^n.$$

Osserviamo che il rango per righe di A coincide con il rango della matrice trasposta di A.

Osservazione 6.0.1.

1. $\mathrm{rg}(A)$ è uguale al massimo numero di colonne linearmente indipendenti di A.
2. l' inclusione $\mathrm{Span}(A^{(1)}, A^{(2)}, \dots, A^{(n)}) \subseteq \mathbb{K}^m$ implica in particolare che

   $$\mathrm{rg}(A) \leq m.$$

   Inoltre, siccome $\mathrm{Span}(A^{(1)}, A^{(2)}, \dots, A^{(n)})$ è generato da n vettori, si ha anche

   $$\mathrm{rg}(A) \leq n.$$

   Quindi

   $$\mathrm{rg}(A) \leq \min\{m, n\}.$$

Esempio 6.0.2.

1. Se A è la matrice nulla: $A = 0$, allora $\mathrm{rg}(A) = 0$. Viceversa, se $\mathrm{rg}(A) = 0$, allora $A = 0$.
2. $\mathrm{rg}(\mathbb{I}_n) = n$. Infatti le colonne della matrice unità sono i vettori della base canonica di $\mathbb{K}^n$:

   $$(\mathbb{I}_n)^{(1)} = e_1, \dots, (\mathbb{I}_n)^{(n)} = e_n.$$

Proposizione 6.0.3. *Sia $A \in M_{m,n}(\mathbb{K})$, e sia $\widetilde{A}$ una matrice ottenuta da A tramite una sequenza di operazioni elementari. Allora valgono le seguenti affermazioni:*

1. $\mathrm{rg}(A) = \mathrm{rg}(\widetilde{A})$;
2. *se* $\widetilde{A}$ *è a scala,* $\mathrm{rg}(\widetilde{A})$ *è uguale al numero* r *delle righe non nulle di* $\widetilde{A}$.
   Inoltre $A^{(j_1)}, \dots, A^{(j_r)}$ *sono linearmente indipendenti, dove* $\widetilde{a}_{1,j_1}, \dots, \widetilde{a}_{r,j_r}$ *sono i pivot di* $\widetilde{A}$.
3. $\mathrm{rg}({}^tA) = \mathrm{rg}({}^t\widetilde{A})$.
4. $\mathrm{rg}({}^t\widetilde{A}) = \mathrm{rg}(\widetilde{A})$.
5. $\mathrm{rg}(A) = \mathrm{rg}({}^tA)$.

Dimostrazione. 1. Per dimostrare questa affermazione usiamo un risultato che vedremo più avanti, il Teorema di dimensione. Questo teorema afferma che

$$\mathrm{rg}(A) = n - \dim(W), \tag{6.0.1}$$

dove

$$W = \{s \in \mathbb{K}^n | A \cdot s = 0\}$$

è il sottospazio vettoriale di $\mathbb{K}^n$ formato dalle soluzioni del sistema omogeneo di equazioni lineari

$$A \cdot X = 0.$$

Se $\widetilde{A}$ si ottiene da A per mezzo di operazioni elementari, allora il sistema di equazioni lineari

$$\widetilde{A} \cdot X = 0$$

è equivalente ad $A \cdot X = 0$, cioè $\widetilde{W} = W$, dove $\widetilde{W} = \{s \in \mathbb{K}^n | \widetilde{A} \cdot s = 0\}$ è il sottospazio di $\mathbb{K}^n$ delle soluzioni di $\widetilde{A} \cdot X = 0$. Usando la formula (6.0.1) abbiamo quindi:

$$\mathrm{rg}(A) = n - \dim(W) = n - \dim(\widetilde{W}) = \mathrm{rg}(\widetilde{A}).$$

2. Supponiamo ora che $\widetilde{A}$ sia a scala, e sia r il numero delle righe non nulle di $\widetilde{A}$. Sia $\widetilde{a}_{ij}$ l' elemento di posto i, j di $\widetilde{A}$. Siccome $\widetilde{A}$ è a scala, si ha che $a_{ij} = 0$, se $i > r$, quindi $\widetilde{A}^{(j)} \in \mathrm{Span}(e_1, \dots, e_r)$ per ogni $j = 1, \dots, n$, dove $\widetilde{A}^{(j)}$ è la colonna j-esima di $\widetilde{A}$ ed $e_1, \dots, e_r$ sono i primi r vettori della base canonica di $\mathbb{K}^n$. Da questo segue che

$$\mathrm{Span}(\widetilde{A}^{(1)}, \dots, \widetilde{A}^{(n)}) \subseteq \mathrm{Span}(e_1, \dots, e_r),$$

quindi

$$\mathrm{rg}(\widetilde{A}) = \dim \mathrm{Span}(\widetilde{A}^{(1)}, \dots, \widetilde{A}^{(n)}) \leq \dim \mathrm{Span}(e_1, \dots, e_r) = r.$$

Per dimostrare che $\mathrm{rg}(\widetilde{A}) = r$, è sufficiente dimostrare che le colonne

$$\widetilde{A}^{(j_1)}, \dots, \widetilde{A}^{(j_r)}$$

sono linearmente indipendenti, dove $\widetilde{a}_{1j_1}, \widetilde{a}_{2j_2}, \dots, \widetilde{a}_{rj_r}$ sono i pivot di $\widetilde{A}$. A tale scopo, consideriamo una combinazione lineare

$$\lambda_1 \widetilde{A}^{(j_1)} + \dots + \lambda_r \widetilde{A}^{(j_r)} = 0 \tag{6.0.2}$$

e supponiamo che essa dia il vettore nullo. Osserviamo che il membro sinistro della (6.0.2) è un vettore della forma seguente:

$$\begin{pmatrix} \lambda_1 \widetilde{a}_{1j_1} + \lambda_2 \widetilde{a}_{1j_2} + \cdots + \lambda_r \widetilde{a}_{1j_r} \\ \lambda_2 \widetilde{a}_{2j_2} + \ldots \lambda_r \widetilde{a}_{2j_r} \\ \vdots \\ \lambda_r \widetilde{a}_{rj_r} \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{6.0.3}$$

Siccome $\widetilde{a}_{1j_1} \neq 0, \widetilde{a}_{2j_2} \neq 0, \ldots, \widetilde{a}_{rj_r} \neq 0$, il vettore (6.0.3) è nullo se e solo se $\lambda_r = \lambda_{r-1} = \cdots = \lambda_1 = 0$. Concludiamo quindi che le colonne $\widetilde{A}^{(j_1)}, \ldots, \widetilde{A}^{(j_r)}$ sono linearmente indipendenti e $\mathrm{rg}(\widetilde{A}) = r$.

Dimostriamo ora che le colonne $A^{(j_1)}, \ldots, A^{(j_r)}$ di A, sono linearmente indipendenti, dove $j_1, j_2, \ldots, j_r$ sono gli indici di colonna dei pivot $\widetilde{a}_{1j_1}, \widetilde{a}_{2j_2}, \ldots, \widetilde{a}_{rj_r}$ di $\widetilde{A}$.

A tale scopo, consideriamo una combinazione lineare

$$\beta_1 A^{(j_1)} + \cdots + \beta_r A^{(j_r)} = 0 \tag{6.0.4}$$

e supponiamo che essa dia il vettore nullo. Definiamo il vettore $s = \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{pmatrix} \in \mathbb{K}^n$ come segue:

$$s_i := \begin{cases} 0, & \text{se } i \neq j_1, \ldots j_r; \\ \beta_i, & \text{se } i = j_k,\ k = 1, \ldots, r. \end{cases}$$

Con questa scelta abbiamo

$$A \cdot s = \beta_1 A^{(j_1)} + \cdots + \beta_r A^{(j_r)},$$

quindi per la (6.0.4) si ha

$$A \cdot s = 0.$$

Siccome $\widetilde{A}$ è ottenuta da A per mezzo di operazioni elementari, abbiamo che vale anche

$$\widetilde{A} \cdot s = 0.$$

Osserviamo che

$$\widetilde{A} \cdot s = \beta_1 \widetilde{A}^{(j_1)} + \cdots + \beta_r \widetilde{A}^{(j_r)},$$

e siccome $\widetilde{A}^{(j_1)}, \ldots, \widetilde{A}^{(j_r)}$ sono linearmente indipendenti, concludiamo che $\beta_1 = \cdots = \beta_r = 0$, quindi anche $A^{(j_1)}, \ldots, A^{(j_r)}$ sono linearmente indipendenti.

3. È sufficiente dimostrare l' enunciato nel caso in cui $\widetilde{A}$ si ottiene da A per mezzo di una operazione elementare di tipo 1, rispettivamente di tipo 2 o 3. Supponiamo dapprima che A sia ottenuta

da A scambiando la riga i-esima con quella j-esima, per qualche $i \neq j, i, j \in \{1, \ldots, m\}$. Questo corrisponde allo scambio della colonna i-esima con la j-esima di ${}^tA$, quindi

$$({}^t\widetilde{A})^{(k)} = \begin{cases} ({}^tA)^{(k)}, & \text{se } k \neq i, j, \\ ({}^tA)^{(j)}, & \text{se } k = i, \\ ({}^tA)^{(i)}, & \text{se } k = j. \end{cases}$$

Da questo segue immediatamente che

$$\mathrm{Span}(({}^tA)^{(1)}, \ldots, ({}^tA)^{(m)}) = \mathrm{Span}(({}^t\widetilde{A})^{(1)}, \ldots, ({}^t\widetilde{A})^{(m)}),$$

perciò $\mathrm{rg}({}^tA) = \mathrm{rg}({}^t\widetilde{A})$.

Supponiamo ora che $\widetilde{A}$ si ottenga da A per mezzo di una operazione elementare di tipo 3, precisamente sostituendo la riga j-esima con $A_{(j)} + cA_{(i)}$, per qualche $i \neq j, i, j \in \{1, \ldots, m\}$, e $c \in \mathbb{K}$. Allora

$$({}^t\widetilde{A})^{(k)} = \begin{cases} ({}^tA)^{(k)}, & \text{se } k \neq j, \\ ({}^tA)^{(j)} + c({}^tA)^{(i)}, & \text{se } k = j, \end{cases}$$

ed in questo caso abbiamo:

$$\mathrm{Span}(({}^tA)^{(1)}, \ldots, ({}^tA)^{(m)}) = \mathrm{Span}(({}^tA)^{(1)}, \ldots, ({}^tA)^{(j-1)}, ({}^tA)^{(j)} + c({}^tA)^{(i)}, \ldots, ({}^tA)^{(m)}) =$$

$$= \mathrm{Span}(({}^t\widetilde{A})^{(1)}, \ldots, ({}^t\widetilde{A})^{(m)}),$$

perciò $\mathrm{rg}({}^tA) = \mathrm{rg}({}^t\widetilde{A})$.

Si verifica analogamente che se $\widetilde{A}$ si ottiene da A per mezzo di una operazione elementare di tipo 2, allora $\mathrm{rg}({}^tA) = \mathrm{rg}({}^t\widetilde{A})$.

4. Per i punti 1 e 3 appena dimostrati, è sufficiente dimostrare che

$$\mathrm{rg}(\widetilde{A}) = \mathrm{rg}({}^t\widetilde{A}),$$

dove $\widetilde{A}$ è una matrice a scala ottenuta da A per mezzo di operazioni elementari.

Per il punto 2 abbiamo che $\mathrm{rg}(\widetilde{A}) = r$, dove r è il numero delle righe non nulle di A. Dimostriamo quindi che $\mathrm{rg}({}^t\widetilde{A}) = r$.

A tale scopo, sia

$$c_1({}^t\widetilde{A})^{(1)} + \cdots + c_m({}^t\widetilde{A})^{(m)} \in \mathbb{K}^n \tag{6.0.5}$$

una combinazione lineare delle colonne di ${}^t\widetilde{A}$, e supponiamo che sia uguale al vettore nullo. Osserviamo che

$$({}^t\widetilde{A})^{(r+1)} = \cdots = ({}^t\widetilde{A})^{(m)} = 0,$$

quindi possiamo omettere queste colonne nella (6.0.5). Inoltre tale combinazione lineare ha la forma:

$$\begin{pmatrix} 0 \\ \vdots \\ 0 \\ \lambda_1 \widetilde{a}_{1j_1} \\ \vdots \\ \lambda_1 \widetilde{a}_{1j_2} + \lambda_2 \widetilde{a}_{2j_2} \\ \vdots \\ \lambda_1 \widetilde{a}_{1j_r} + \ldots \lambda_r \widetilde{a}_{rj_r} \\ \vdots \end{pmatrix}. \tag{6.0.6}$$

Siccome $\widetilde{a}_{1j_1} \neq 0, \widetilde{a}_{2j_2} \neq 0, \ldots, \widetilde{a}_{rj_r} \neq 0$, il vettore (6.0.6) è nullo se e solo se $\lambda_1 = \lambda_2 = \cdots = \lambda_r = 0$. Concludiamo quindi che le colonne $({}^t\widetilde{A})^{(1)}, \ldots, ({}^t\widetilde{A})^{(r)}$ sono linearmente indipendenti e $\mathrm{rg}({}^t\widetilde{A}) = r$.

□

Teorema 6.0.4. *(Rouché-Capelli). Un sistema lineare $A \cdot X = b$ di m equazioni lineari in n incognite è compatibile $\iff$*

$$\mathrm{rg}(A|b) = \mathrm{rg}(A).$$

In tal caso la sua generica soluzione dipende da $n - r$ parametri, dove $r = \mathrm{rg}(A)$.

La dimostrazione del Teorema di Rouché-Capelli seguirà dal seguente risultato:

Lemma 6.0.5. *Un sistema lineare $A \cdot X = b$ di m equazioni lineari di ordine n è compatibile $\iff$*

$$b \in \mathrm{Span}(A^{(1)}, A^{(2)}, \ldots, A^{(n)}). \tag{6.0.7}$$

Dimostrazione. Abbiamo che $A \cdot X = b$ è compatibile $\iff$ esiste $s = \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{pmatrix} \in \mathbb{K}^n$ tale che $A \cdot s = b \iff$

$$A^{(1)} s_1 + A^{(2)} s_2 + \cdots + A^{(n)} s_n = b \iff b \in \mathrm{Span}(A^{(1)}, A^{(2)}, \ldots, A^{(n)}).$$

□

Dimostrazione. (Teorema di Rouché - Capelli) Per il Lemma 6.0.5 il sistema $A \cdot X = b$ è compatibile se e solo se $b \in \mathrm{Span}(A^{(1)}, A^{(2)}, \ldots, A^{(n)})$. Si può verificare facilmente che tale condizione è soddisfatta se e solo se

$$\mathrm{Span}(A^{(1)}, A^{(2)}, \ldots, A^{(n)}) = \mathrm{Span}(A^{(1)}, A^{(2)}, \ldots, A^{(n)}, b).$$

Siccome in generale vale

$$\mathrm{Span}(A^{(1)}, A^{(2)}, \ldots, A^{(n)}) \subseteq \mathrm{Span}(A^{(1)}, A^{(2)}, \ldots, A^{(n)}, b),$$

i due sottospazi vettoriali sono uguali se e solo se hanno la stessa dimensione, cioè se e solo se $\mathrm{rg}(A) = \mathrm{rg}(A|b)$.

Supponiamo ora che $A \cdot X = b$ sia compatibile. Il teorema di struttura per le soluzioni dei sistemi di equazioni lineari afferma che $S = \{\widetilde{s} + s_0 \mid s_0 \in W\}$, dove $S = \{s \in \mathbb{K}^n | A \cdot s = b\}$ è l' insieme delle soluzioni, $\widetilde{s}$ è una soluzione particolare, e

$$W = \{s_0 \in \mathbb{K}^n | A \cdot s_0 = 0\}$$

è il sottospazio vettoriale di $\mathbb{K}^n$ delle soluzioni del sistema lineare omogeneo associato. Per il Teorema della Dimensione, che vedremo in seguito, $\dim(W) = n - \mathrm{rg}(A) = n - r$. Fissiamo una base $\{v_1, \ldots, v_{n-r}\}$ di W; allora ogni soluzione si scrive nella forma seguente

$$s = \widetilde{s} + c_1 v_1 + \cdots + c_{n-r} v_{n-r},$$

al variare dei parametri $c_1, \ldots, c_{n-r} \in \mathbb{K}$. Quindi le soluzioni di $A \cdot X = b$ dipendono da $n - r$ parametri. □

6.1 Rango e invertibilità

Il seguente teorema afferma che una matrice quadrata è invertibile se e solo se ha rango massimo.

Teorema 6.1.1. *Sia $A \in M_n(\mathbb{K})$. Allora A è invertibile se e solo se* $\mathrm{rg}(A) = n$.

Dimostrazione. Supponiamo che A sia invertibile. Sia $M \in M_n(\mathbb{K})$ la matrice inversa di A, quindi $A \cdot M = \mathbb{I}_n$. Osserviamo che quest ultima equazione si può riscrivere come segue:

$$A \cdot M^{(i)} = e_i, \; i = 1, \ldots, n, \tag{6.1.1}$$

dove e_i è l' i-esimo vettore della base canonica di $\mathbb{K}^n$. Come visto nel Lemma 6.0.5, questo implica che

$$e_1, \ldots, e_n \in \mathrm{Span}(A^{(1)}, A^{(2)}, \ldots, A^{(n)}),$$

quindi

$$\mathbb{K}^n = \mathrm{Span}(e_1, \ldots, e_n) \subseteq \mathrm{Span}(A(1), \ldots, A(n)) \subseteq \mathbb{K}^n,$$

da cui $\mathrm{Span}(A(1), \ldots, A(n)) = \mathbb{K}^n$ e e $\mathrm{rg}(A) = n$.

Viceversa, supponiamo che $\mathrm{rg}(A) = n$, e consideriamo i sistemi lineari

$$A \cdot X = e_i, \; i = 1, \ldots, n. \tag{6.1.2}$$

Osserviamo che per ogni i, il rango della matrice completa $(A \mid e_i)$ soddisfa:

$$n \geq \mathrm{rg}(A \mid e_i) \geq \mathrm{rg}(A) = n,$$

quindi $\mathrm{rg}(A \mid e_i) = \mathrm{rg}(A)$. Per il Teorema di Rouché - Capelli i sistemi (6.1.2) sono compatibili per ogni $i = 1, \ldots, n$, e la generica soluzione dipende da $n - n = 0$ parametri, cioè è unica. Inserendo i vettori soluzione nelle colonne di una matrice si ottiene una $M \in M_n(\mathbb{K})$ tale che

$$A \cdot M = \mathbb{I}_n.$$

Con un ragionamento analogo si può dimostrare che vale anche

$$M \cdot A = \mathbb{I}_n,$$

e quindi A è invertibile e si ha

$$M = A^{-1}.$$

□

6.2 Calcolo della matrice inversa con l' algoritmo di Gauss

Se $A \in M_n(\mathbb{K})$ è invertibile, per calcolare la sua inversa A^{-1} si può procedere come segue. Ricordiamo che A^{-1} è quella matrice $M \in M_n(\mathbb{K})$ tale che $A \cdot M = \mathbb{I}_n$. Quindi, per ogni $i = 1, \ldots, n$, la colonna i-esima $M^{(i)}$ di M è l' unica soluzione del sistema di equazioni lineari

$$A \cdot X = e_i,$$

dove $e_i \in \mathbb{K}^n$ è l' i-esimo vettore della base canonica di $\mathbb{K}^n$. Notiamo che è possibile risolvere questi sistemi lineari simultaneamente, per $i = 1, \ldots, n$, nel seguente modo. Consideriamo la matrice

$$(A \mid \mathbb{I}_n) \in M_{n,2n}(\mathbb{K}).$$

Siccome A è invertibile, si ha $\mathrm{rg}(A) = n$. Non è difficile verificare che è possibile trasformare $(A \mid \mathbb{I}_n)$ tramite una sequenza di operazioni elementari OE1, OE2, OE3, nella matrice

$$(\mathbb{I}_n \mid M).$$

Siccome le operazioni elementari trasformano un sistema di equazioni lineari in uno equivalente, segue che la soluzione di $A \cdot X = e_i$ coincide con la soluzione di $\mathbb{I}_n \cdot X = M^{(i)}$, che è proprio $M^{(i)}$, quindi $M = A^{-1}$.

6.3 Rango e sottomatrici

Definizione 6.3.1. Sia $A = (a_{ij}) \in M_{m,n}(\mathbb{K})$, e siano $p \in \{1, \ldots, m\}$, $q \in \{1, \ldots, n\}$. Siano inoltre $1 \leq i_1 < \cdots < i_p \leq m$, $1 \leq j_1 < \cdots < j_q \leq n$. Con

$$A(i_1, \ldots, i_p \mid j_1, \ldots, j_q)$$

denotiamo la matrice $p \times q$ a coefficienti in $\mathbb{K}$ il cui elemento di posto k, l è dato da

$$a_{i_k, j_l}, \ \forall k = 1, \ldots, p, \forall l = 1, \ldots, q.$$

Ogni matrice del tipo $A(i_1, \ldots, i_p \mid j_1, \ldots, j_q)$ si chiama sottomatrice $p \times q$ di A.

Proposizione 6.3.2. *Sia $A \in M_{m,n}(\mathbb{K})$, e sia B una sottomatrice $p \times q$ di A. Allora* $\mathrm{rg}(B) \leq \mathrm{rg}(A)$.

Dimostrazione. Omessa. □

Teorema 6.3.3. *Sia $A \in M_{m,n}(\mathbb{K})$. Allora*

$$\mathrm{rg}(A) = \max\{\mathrm{rg}(B) | B \text{ sottomatrice quadrata di } A\},$$

e anche

$$\mathrm{rg}(A) = \text{massimo degli ordini delle sottomatrici quadrate ed invertibili di } A.$$

Dimostrazione. Omessa. □

Capitolo 7

Definizione induttiva di determinante

In questo capitolo definiamo il determinante di una matrice quadrata $A \in M_n(\mathbb{K})$ e ne studiamo le principali proprietà, vedremo anche alcune applicazioni al calcolo delle soluzioni dei sistemi di equazioni lineari ed al calcolo della matrice inversa.

Per iniziare consideriamo i casi più semplici $n = 1$ e $n = 2$.

Se $A \in M_1(\mathbb{K})$, una matrice di ordine 1. Allora A è del tipo $A = (a)$ con $a \in \mathbb{K}$ e poniamo il suo determinante per definizione

$$\det A = a.$$

Passando alle matrici di ordine 2, se

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},$$

definiamo

$$\det A := a_{11}\, a_{22} - a_{12}\, a_{21}.$$

Osserviamo che $A \in M_2(\mathbb{K})$ è invertibile se e solo se

$$\det\ A \neq 0.$$

Infatti, consideriamo la matrice

$$B := \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

e osserviamo che vale

$$A \cdot B = \begin{pmatrix} a_{11}\, a_{22} - a_{12}\, a_{21} & 0 \\ 0 & a_{11}\, a_{22} - a_{12}\, a_{21} \end{pmatrix} = \begin{pmatrix} \det\ A & 0 \\ 0 & \det\ A \end{pmatrix} = (\det\ A) \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (7.0.1)$$

Se $\det\ A \neq 0$, la matrice $M := \frac{1}{\det\ A} \cdot B$ verifica

$$A \cdot M = I_2,$$

quindi A è invertibile per definizione e si ha $M = A^{-1}$.

Se $\det\ A = 0$ e A è la matrice nulla, si verifica facilmente che A non è invertibile. Se $\det\ A = 0$ e A non è la matrice nulla, allora anche B non è la matrice nulla e per la relazione (7.0.1) si ha

$$A \cdot B = 0.$$

quindi le colonne di B sono due soluzioni del sistema lineare omogeneo

$$A \cdot X = 0$$

(verificare!), ed essendoci almeno una soluzione non banale, per il Teorema di Rouchè - Capelli, il rango di A verifica

$$\text{rg } A = 1.$$

Siccome una matrice di ordine 2 è invertibile se e solo se il suo rango $\grave{2}$ per un risultato del capitolo precedente, deduciamo che A non ìnvertibile.

Osservazione 7.0.1. **Interpretazione geometrica del determinante di matrici di ordine** 2 Supponiamo pe semplicità che i coefficienti a_{ij} della matrice di ordine 2 che consideriamo siano tutti positivi:

$$a_{ij} > 0,$$

e poniamo $\vec{v} = \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}$, $\vec{w} = \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}$. Allora si può verificare che il modulo

$$|\det\ A| = \text{Area}(P),$$

dove P è il parallelogramma determinato da $\vec{v}$ e $\vec{w}$.

Vogliamo ora definire il determinante di una matrice quadrata qualunque che ne determini l'invertibilità. Abbiamo bisogno di premettere la seguente:

Definizione 7.0.1. Sia $A \in M_n(\mathbb{K})$ una matrice quadrata di ordine n. Denoteremo con

$$A_{ij} \in M_{n-1}(\mathbb{K})$$

la matrice quadrata di ordine $n-1$ ottenuta cancellando la riga i e la colonna j dalla matrice A. Chiameremo A_{ij} **minore** di A.

Definizione 7.0.2. Sia $A \in M_n(\mathbb{K})$. Definiamo il **determinante** di A in modo induttivo:

- se $n = 1$, poniamo $\det\,(a_{11}) = a_{11}$;
- se $n > 1$, poniamo

$$\det\ A := \sum_{i=1}^{n} (-1)^{i+1}\, a_{i1} \det\ A_{i1}. \tag{7.0.2}$$

Esercizio 7.0.3. Si verifichi che nel caso $n = 2$ la definizione appena data coincide con la definizione dell'introduzione.

Esempio 7.0.4. Per la definizione (7.0.2) il determinante di una matrice

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

di ordine 3 è uguale a

$$\det A = a_{11} \det A_{11} - a_{21} \det A_{21} + a_{31} \det A_{31} =$$

$$= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{21}(a_{12}a_{33} - a_{13}a_{32}) + a_{31}(a_{12}a_{23} - a_{13}a_{22}).$$

Enunciamo un importante risultato che riguarda le proprietà del determinante, di cui non daremo la dimostrazione.

Teorema 7.0.5. *Sia $n \geq 1$ e sia $A \in M_n(\mathbb{K})$. Valgono le seguenti proprietà del determinante:*

- (D1) **Multilinearità del determinante** *per ogni i, $1 \leq i \leq n$, se la riga i è somma di due righe:*

$$A_{(i)} = R_1 + R_2,$$

allora

$$\det A = \det \begin{pmatrix} A_{(1)} \\ A_{(2)} \\ \vdots \\ R_1 + R_2 \\ \vdots \\ A_{(n)} \end{pmatrix} = \det \begin{pmatrix} A_{(1)} \\ A_{(2)} \\ \vdots \\ R_1 \\ \vdots \\ A_{(n)} \end{pmatrix} + \det \begin{pmatrix} A_{(1)} \\ A_{(2)} \\ \vdots \\ R_2 \\ \vdots \\ A_{(n)} \end{pmatrix}.$$

Inoltre, se la riga i è prodotto di una riga R per uno scalare $c \in \mathbb{K}$:

$$A_{(i)} = c\, R,$$

allora

$$\det A = \det \begin{pmatrix} A_{(1)} \\ A_{(2)} \\ \vdots \\ c\, R \\ \vdots \\ A_{(n)} \end{pmatrix} = c\, \det \begin{pmatrix} A_{(1)} \\ A_{(2)} \\ \vdots \\ R \\ \vdots \\ A_{(n)} \end{pmatrix}.$$

- (D2) **Alternanza del determinante** *Scambiando due righe di A il determinante cambia segno:*

$$\det \begin{pmatrix} A_{(1)} \\ A_{(2)} \\ \vdots \\ A_{(i)} \\ \vdots \\ A_{(j)} \\ \vdots \\ A_{(n)} \end{pmatrix} = -\det \begin{pmatrix} A_{(1)} \\ A_{(2)} \\ \vdots \\ A_{(j)} \\ \vdots \\ A_{(i)} \\ \vdots \\ A_{(n)} \end{pmatrix}$$

- (D3) **Normalizzazione del determinante** *Se I_n indica la matrice unità, si ha*

$$\det I_n = 1.$$

Teorema 7.0.6. Teorema di caratterizzazione del determinante *Il determinante è l'unica funzione $M_n(\mathbb{K}) \to \mathbb{K}$ che verifichi le proprietà (D1), (D2) e (D3).*

Corollario 7.0.7. *Sia $A \in M_n(\mathbb{K})$. Valgono le seguenti affermazioni:*

1. *Se A ha due righe uguali, allora* $\det A = 0$.
2. *Se A ha una riga nulla, allora* $\det A = 0$.
3. (a) *Se $\widetilde{A}$ è ottenuta da A mediante un numero finito σ di operazioni elementari di tipo OE1 (cioè σ scambi), allora*
   $$\det \widetilde{A} = (-1)^{\sigma} \det A.$$
   (b) *Se $\widetilde{A}$ è ottenuta da A mediante una operazione elementare di tipo OE2, cioè una riga è moltiplicata per uno scalare $c \in \mathbb{K}$, allora*
   $$\det \widetilde{A} = c \det A.$$
   (c) *Se $\widetilde{A}$ è ottenuta da A mediante una operazione elementare di tipo OE3, cioè una riga è dalla somma della riga stessa con un'altra riga, allora*
   $$\det \widetilde{A} = \det A.$$
4. *Se A è una matrice* **triangolare superiore**, *cioè tale che*
   $$a_{ij} = 0 \quad \forall\, i > j, \quad A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1\,n-1} & a_{1n} \\ 0 & a_{22} & \dots & a_{2\,n-1} & a_{2n} \\ 0 & 0 & \dots & a_{3\,n-1} & a_{3n} \\ \vdots & \vdots & \vdots & \vdots & \\ 0 & 0 & \dots & 0 & a_{nn} \end{pmatrix},$$
   allora il determinante è uguale al prodotto degli elementi sulla diagonale:
   $$\det A = a_{11} \cdot a_{22} \cdot \dots \cdot a_{nn}.$$
   In particolare, ciò vale se A è a scala oppure diagonale.

Dimostrazione. 1. Supponiamo che sia $A_{(i)} = A_{(j)}$ per qualche coppia di indici i e j. Allora scambiando la riga i di A con la riga j, la matrice A non cambia. D'altra parte, per la proprietà (D2), uno scambio comporta un cambio di segno nel determinante, quindi
$$\det A = -\det A,$$
da cui segue $\det A = 0$ (se la caratteristica del campo è diversa da 2, ad esempio se $\mathbb{K} = \mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{Z}_3$ ... non vale se $\mathbb{K} = \mathbb{Z}_2$).

2. Se $A_{(i)} = (0\,0\ \dots\ 0)$, allora possiamo scrivere
$$A_{(i)} = 0 \cdot (1\,1\ \dots\ 1),$$
e dalla (D1) si ha $\det A = 0 \cdot \det A' = 0$.

3. I tre enunciati segueno direttamente dalle (D1) e (D2) e dalla proprietà 1) dimostrata sopra.

4. Dimostriamo l'enunciato per induzione su n, l'ordine della matrice.

   Per $n = 1$ l'enunciato è vero.

   Supponiamo che valga per n e dimostriamo che allora vale anche per tutte le matrici triangolari superiori di ordine $n + 1$.

   Sia $A \in M_{n+1}$ triangolare superiore. In particolare abbiamo

$$a_{21} = a_{31} = \cdots = an1 = 0,$$

   quindi la formula per il determinante (7.0.2) diventa semplicemente

$$\det A = a_{11} \ \det A_{11}. \tag{7.0.3}$$

   Osservaimo ora che anche A_{11} è una matrice triangolare superiore, infatti è data da

$$A_{11} = \begin{pmatrix} a_{22} & \dots & a_{2\ n} & a_{2\ n+1} \\ 0 & \dots & a_{3\ n} & a_{3\ n+1} \\ \vdots & \vdots & \vdots & \\ 0 & \dots & 0 & a_{n+1\ n+1} \end{pmatrix} \in M_n(\mathbb{K}).$$

   Quindi possiamo applicare l'ipotesi induttiva a A_{11} che ci fornisce

$$\det A_{11} = a_{22} \cdot a_{33} \cdot \dots \cdot a_{n+1\ n+1},$$

   e sostituendo nell'espressione (7.0.3) troviamo l'enunciato.

□

7.1 Caratterizzazione delle matrici quadrate di rango massimo

Vedremo ora che per le matrici di ordine n qualunque il determinante ci fornisce un criterio per stabilire quando una matrice ha rango massimo, e risulta quindi invertibile.

Teorema 7.1.1. *Sia $A \in M_n(\mathbb{K})$. Allora*

$$\operatorname{rg} A < n \quad \Longleftrightarrow \quad \det A = 0.$$

Equivalentemente si ha:

$$\operatorname{rg} A = n \quad \Longleftrightarrow \quad \det A \neq 0.$$

Dimostrazione. Supponiamo che sia $\operatorname{rg} A < n$. Sia$\widetilde{A}$ una matrice a scala ottenuta di A con un numero finito di operazioni elementari. Allora si ha $\operatorname{rg} A = \operatorname{rg} \widetilde{A}$, ed è facile verificare che la condizione $\operatorname{rg} A < n$ implica che $\widetilde{A}$ ha almeno una riga nulla. Allora per il Corollario 7.0.7,punto 2), si

$$\det \widetilde{A} = 0.$$

D'altra parte, per il Corollario 7.0.7, punti $3.a), b), c)$, abbiamo

$$\det \widetilde{A} = (-1)^{\sigma} c \cdot \det A, \tag{7.1.1}$$

quindi anche $\det A = 0$.

Viceversa, supponiamo $\det A = 0$. Per la formula (7.1.1) si ha che anche $\det \widetilde{A} = 0$. Supponiamo per assurdo che non sia $\operatorname{rg} \widetilde{A} < n$, quindi che sia

$$\operatorname{rg} \widetilde{A} = n.$$

Allora si verifica facilmente che i pivot di $\widetilde{A}$ sono esattamente gli elementi a_{ii} della diagonale. Siccome si ha

$$\det \widetilde{A} = a_{11} \cdot a_{22} \cdot \cdots \cdot a_{nn}$$

per il Corollario 7.0.7, punto 4), essendo tutti gli $a_{ii} \neq 0$ perchè pivot, si avrebbe $\det \widetilde{A} \neq 0$, che è un assurdo.

□

Corollario 7.1.2. *Una matrice $A \in M_n(\mathbb{K})$ è invertibile se e solo se* $\det A \neq 0$.

Dimostrazione. Segue dal Teorema precedente e dal fatto che una matrice quadrata di ordine nè invertibile se e solo se $\operatorname{rg} A = n$.

□

7.2 Regole di Laplace

Il seguente risultato afferma che nell'espressione (7.0.2) per il determinante di una matrice A si può sostituire la prima colonna con una colonna qualsiasi $A^{(k)}$. Questo è particolarmente vantaggioso se A ha una colonna con molti zeri.

Teorema 7.2.1. Sviluppo di Laplace del determinante per colonne *Sia $A \in M_n(\mathbb{K})$. Per ogni indice di colonna $k \in \{1, \dots, n\}$ il determinante di A si può esprimere nella seguente forma:*

$$\det A = \sum_{i=1}^{n} (-1)^{i+k} \, a_{ik} \, \det A_{ik}.$$

Dimostrazione. Cenno: è possibile dimostrare che il termine

$$\sum_{i=1}^{n} (-1)^{i+k} \, a_{ik} \, \det A_{ik}$$

soddisfa le proprietà (D1), (D2), (D3). Per il Teorema di caratterizzazione del determinante ciò implica che coincide con il determinante di$A$. □

Vediamo ora un altro Teorema di Laplace, che afferma che si può calcolare il determinante di una matrice anche a partire da una sua riga fissata.

Teorema 7.2.2. Sviluppo di Laplace del determinante per righe

Sia $A \in M_n(\mathbb{K})$. Per ogni indice di riga $l \in \{1, \dots, n\}$ il determinante di A si può esprimere nella seguente forma:

$$\det A = \sum_{j=1}^{n} (-1)^{l+j} \, a_{lj} \, \det A_{lj}.$$

Corollario 7.2.3. *Sia* $A \in M_n(\mathbb{K})$. *allora si ha*

$$\det A = \overset{t}{\det} A.$$

Dimostrazione. Siccome le colonne di ${}^tA$ corrispondono alle righe di$A$, l'enunciato segue dal Teorema di Sviluppo di Laplace del determinante per righe. □

7.3 Teorema di Binet

Riportiamo il seguente importante teorema, senza dimostrazione.

Teorema 7.3.1. di Binet *Siano* $A, B \in M_n(\mathbb{K})$. *Allora vale*

$$\det(A \cdot B) = \det A \cdot \det B.$$

Corollario 7.3.2. *Sia* $A \in M_n(\mathbb{K})$. *Si ha:*

- $\det(A^n) = (\det A)^n$.
- *Se* $A \in M_n(\mathbb{K})$ *è invertibile, allora si ha*

$$\det(A^{-1}) = \frac{1}{\det A}.$$

Dimostrazione. Il primo punto si dimostra facilmente per induzione. Per il secondo punto, l'inversa di una matrice soddisfa per definizione

$$A^{-1} \cdot A = I_n,$$

da cui otteniamo

$$\det(A^{-1} \cdot A) = \det I_n = 1,$$

e per il Teorema di Binet abbiamo

$$\det(A^{-1} \cdot A) = \det(A^{-1}) \cdot \det A = 1.$$

□

Capitolo 8

Spazi affini

1. Spazi affini

Definizione 1.1. *Sia V uno spazio vettoriale su un campo $\mathbb{K}$. Uno spazio affine su V è un insieme non vuoto $\mathbb{A}$, i cui elementi si dicono i* **punti di** $\mathbb{A}$, *ed una funzione*

$$\sigma : \mathbb{A} \times \mathbb{A} \to V, \qquad \sigma(P, R) = \overrightarrow{PR},$$

tale che valgano i seguenti assiomi:

- (SA1) *per ogni $P \in \mathbb{A}$ e per ogni $v \in V$, $\exists!\ R \in \mathbb{A}$ tale che*
$$v = \overrightarrow{PR};$$
- (SA2) *per ogni terna di punti (non necessariamente distinti) $P, Q, R \in \mathbb{A}$ si ha*
$$\overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{PR}.$$

Proposizione 1.2. *Dai due assiomi di spazio affine si hanno le seguenti proprietà:*

(1) per ogni $P \in \mathbb{A}$ si ha $\overrightarrow{PP} = 0 \in V$;
(2) per ogni $P, R \in \mathbb{A}$, si ha $\overrightarrow{RP} = -\overrightarrow{PR}$;
(3) per ogni punto $P \in \mathbb{A}$, la funzione

$$f_P : \mathbb{A} \to V, \quad f_P(Q) = \overrightarrow{PQ}$$

è una biiezione.

Dimostrazione. (1) Dall'assioma (SA2), ponendo $P = Q = R$, si ha $\overrightarrow{PP} + \overrightarrow{PP} = \overrightarrow{PP}$, e sommando ad entrambi i membri il vettore $-\overrightarrow{PP}$ si ha la tesi.
(2) Dall'assioma (SA2), ponendo $R = P$, si ha $\overrightarrow{PQ} + \overrightarrow{QP} = \overrightarrow{PP} = 0$ per il primo punto, quindi $\overrightarrow{QP}$ è il vettore opposto di $\overrightarrow{PQ}$.
(3) la tesi segue direttamente dall'assioma (SA1).

□

Definizione 1.3. *Sia $\mathbb{A}$ uno spazio affine su uno spazio vettoriale V di dimensione finita. Allora definiamo la* **dimensione di** $\mathbb{A}$ *come*

$$\dim\ \mathbb{A} := \dim\ V.$$

Se

1

- $\dim\ \mathbb{A} = 1$, *allora* $\mathbb{A}$ *si dice* **retta affine***;*
- $\dim\ \mathbb{A} = 2$, *allora* $\mathbb{A}$ *si dice* **piano affine**.

Definizione 1.4. *Se* $V = \mathbb{K}^n$*, possiamo definire una struttura di spazio affine su* $\mathbb{K}^n$ *stesso nel modo seguente:*

$$\sigma : \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}^n, \quad \sigma((p_1, p_2, \ldots, p_n), (q_1, q_2, \ldots, q_n)) := \begin{pmatrix} q_1 - p_1 \\ q_2 - p_2 \\ \vdots \\ q_n - p_n \end{pmatrix} = \overrightarrow{PQ}.$$

Si verifica facilmente che σ *verifica i due assiomi. Lo spazio affine così definito verrà indicato con il simbolo*

$$\mathbb{A}^n_{\mathbb{K}}.$$

Definizione 1.5. *Sia* $\mathbb{A}$ *uno spazio affine su uno spazio vettoriale* V*. Un* **riferimento affine per** $\mathbb{A}$ *è il dato di:*

- *un punto* $O \in \mathbb{A}$*;*
- *una base* $\{v_1, v_2, \ldots, v_n\}$ *di* V*.*

Il punto O *si dice* **origine del riferimento**.

Dato un punto $P \in \mathbb{A}$ *qualsiasi, possiamo definire le* **coordinate di** P *come le coordinate del vettore* $\overrightarrow{OP}$ *nella base* $\{v_1, v_2, \ldots, v_n\}$*, cioè se*

$$\overrightarrow{OP} = a_1 v_1 + a_2 v_2 + \cdots + a_n v_n,$$

definiamo l'n-upla $(a_1, a_2, \ldots, a_n)$ *come coordinate di* P*.*

Esempi 1.6. *Se* $\mathbb{A}$ *è una retta affine, un riferimento affine è dato da un punto* $O \in \mathbb{A}$ *e da un vettore non nullo* $v \in V$*.*

Se $\mathbb{A}$ *è un piano affine, un riferimento affine è dato da un punto* $O \in \mathbb{A}$ *e da due vettori non nullo e non proporzionali* $v_1, v_2 \in V$*.*

Definizione 1.7. *In* $\mathbb{A}^n_{\mathbb{K}}$ *possiamo scegliere il seguente riferimrnto:*

- $O \in \mathbb{A}^n_{\mathbb{K}}$*;*
- *la base canonica di* $\mathbb{K}^n$*:*

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Questo riferimento si dice **riferimento affine canonico di** $\mathbb{A}^n_{\mathbb{K}}$.

2. Sottospazi affini e loro equazioni cartesiane e parametriche

Definizione 2.1. *Sia* $\mathbb{A}$ *spazio affine su* V*. Fissati*

- *un punto* $Q \in \mathbb{A}$*, e*
- *un sottospazio vettoriale* $W \subseteq V$*,*

il **sottospazio affine passante per** Q **e parallelo a** W *è il sootinsieme di* $\mathbb{A}$ *definito*

$$S = \{P \in \mathbb{A} \mid \overrightarrow{QP} \in W\}.$$

Il sottospazio $W \subseteq V$ *si chiama* **giacitura di** S.

Proposizione 2.2. *Dalla definizione si hanno le seguenti proprietà:*

(1) se S *è un sottospazio affine passante per* Q *e di giacitura* W*, allora*

$$Q \in S.$$

(2) Per ogni coppia di punti $P_1, P_2 \in S$*, si ha*

$$\overrightarrow{P_1P_2} \in W.$$

(3) S *ha una struttura di spazio affine su* W.

Definizione 2.3. *Sia* $S \subseteq \mathbb{A}$ *un sottospazio affine di giacitura* W*. Definiamo* $\dim S := \dim W$*.*
Se

$$\dim S = \dim \mathbb{A} - 1,$$

allora S *si dice* **iperpiano** *di* $\mathbb{A}$.

Vediamo ora due modi per descrivere i sottospazi affini, cioè tramite equazioni cartesiane ed equazioni parametriche.

Teorema 2.4. *Sia* $A \cdot x = b$ *un sistema di* m *equazioni lineari di ordine* n *a coefficienti nel campo* $\mathbb{K}$*. Se* $A \cdot x = b$ *è compatibile, allora l'insieme delle sue soluzioni*

$$S = \{s \in \mathbb{K}^n \mid A \cdot s = b\}$$

è un sottospazio affine di $\mathbb{A}^n_{\mathbb{K}}$ *la cui giacitura è il sottospazio vettoriale* W *di* $\mathbb{K}^n$ *formato dalle soluzioni del sistema lineare omogeneo associato:*

$$W = \{s \in \mathbb{K}^n \mid A \cdot s = 0\}.$$

In tal caso $\dim(S) = n - rg(A)$*, e per ogni* $\widetilde{s} \in S$*, il sottoinsieme* S *coincide con il sottospazio affine di* $\mathbb{A}^n_{\mathbb{K}}$ *passante per* $\widetilde{s}$ *e parallelo a* W.

Dimostrazione. Poichè il sistema lineare è compatibile, S non è vuoto. Sia dunque $\widetilde{s}$ una sua soluzione. Per il teorema di struttura dell'insieme delle soluzioni di un sistema lineare, ogni altra soluzione s si può scrivere nella forma

$$s = \widetilde{s} + w, \quad w \in W.$$

Quindi si ha

$$s \in S \iff s - \widetilde{s} \in W \iff \overrightarrow{\widetilde{s}\,s} \in W.$$

Per definizione, S risulta il sottospazio affine di $\mathbb{A}^n_{\mathbb{K}}$ passante per $\widetilde{s}$ e di giacitura W.
Infine, per il Teorema di Dimensione, $\dim W = n - \mathrm{rgA} = \dim S$. □

Definizione 2.5. *Sia* $S \subseteq \mathbb{A}^n_{\mathbb{K}}$ *un sottospazio affine. Un* **sistema di equazioni cartesiane per** S *è un qualunque sistema di equazioni lineari* $Ax = b$*, tale che*

$$S = \{s \in \mathbb{A}^n_{\mathbb{K}} \mid A \cdot s = b\}.$$

Osservazione 2.6. *(1) Un sottospazio affine $S \subseteq \mathbb{A}^n_{\mathbb{K}}$ può essere descritto da diversi sistemi di equazioni Cartesiane. Infatti, se $S = \{s \in \mathbb{A}^n_{\mathbb{K}} \mid A \cdot s = b\}$, allora per ogni sistema di equazioni lineari $\widetilde{A} \cdot x = \widetilde{b}$ equivalente ad $A \cdot x = b$, le equazioni di $\widetilde{A} \cdot x = \widetilde{b}$ sono delle equazioni Cartesiane per S.*

(2) Sia $A \cdot x = b$ un sistema di equazioni cartesiane per il sottospazio affine $S \subseteq \mathbb{A}^n_{\mathbb{K}}$. Allora il punto $(0, 0, \ldots, 0) \in \mathbb{A}^n_{\mathbb{K}}$ appartiene ad S se e solo se $b = 0$, cioè se e solo se il sistema di equazioni lineari $A \cdot x = b$ è omogeneo.

(3) Ogni iperpiano di $\mathbb{A}^n_{\mathbb{K}}$ può essere descritto da una equazione lineare del tipo

$$a_1x_1 + \cdots + a_nx_n = d, \tag{2.1}$$

con $a_1, \ldots, a_n \in \mathbb{K}$ non tutti nulli. Infatti, se $A \cdot x = b$ è un sistema di equazioni cartesiane per S, abbiamo che

$$\dim(S) = n - 1 = n - \mathrm{rg}\,(A).$$

Da questo segue che

$$\mathrm{rg}\,(A) = 1.$$

Siccome il sistema lineare $A \cdot x = b$ 'e compatibile, per il Teorema di Rouché - Capelli si ha

$$\mathrm{rg}\,(A|b) = \mathrm{rg}\,(A) = 1.$$

Sia

$$\widetilde{A} \cdot x = \widetilde{b}$$

un sistema lineare equivalente con $\widetilde{A}$ a scala. Siccome

$$\mathrm{rg}\,(\widetilde{A}|\widetilde{b}) = \mathrm{rg}\,(A|b) = 1,$$

*la matrice $(\widetilde{A}|\widetilde{b})$ ha un'unica riga non nulla, che corrisponde a una equazione lineare del tipo (**??**).*

Vediamo ora come si possono descrivere i sottospazi affini tramite equazioni parametriche.

Sia $S \subseteq \mathbb{A}^n_{\mathbb{K}}$ is sottospazio affine passante per un punto $Q = (q_1, \ldots, q_n) \in \mathbb{A}^n_{\mathbb{K}}$ e di giacitura $W \subseteq \mathbb{K}^n$. Sia

$$\{w_1, \ldots, w_m\}$$

una base del sottospazio vettoriale W. Fissiamo le componenti dei vettori w_i (come elementi di $\mathbb{K}^n$)

$$w_1 = \begin{pmatrix} w_{11} \\ w_{21} \\ \vdots \\ w_{n1} \end{pmatrix}, \quad w_2 = \begin{pmatrix} w_{12} \\ w_{22} \\ \vdots \\ w_{n2} \end{pmatrix}, \ldots, \quad w_m = \begin{pmatrix} w_{1m} \\ w_{2m} \\ \vdots \\ w_{nm} \end{pmatrix}.$$

Allora abbiamo

$$P = (x_1, x_2, \ldots, x_n) \in S \iff \overrightarrow{QP} \in W \iff$$

$$\iff \exists t_1, t_2, \ldots, t_m \in \mathbb{K} \ : \ \overrightarrow{QP} = t_1w_1 + t_2w_2 + \cdots + t_mw_m \iff$$

$$\iff \begin{pmatrix} x_1 - q_1 \\ x_2 - q_2 \\ \vdots \\ x_n - q_n \end{pmatrix} = t_1 \begin{pmatrix} w_{11} \\ w_{21} \\ \vdots \\ w_{n1} \end{pmatrix} + t_2 \begin{pmatrix} w_{12} \\ w_{22} \\ \vdots \\ w_{n2} \end{pmatrix} + \cdots + t_m \begin{pmatrix} w_{1m} \\ w_{2m} \\ \vdots \\ w_{nm} \end{pmatrix} \iff$$

$$\iff \exists t_1, t_2, \dots, t_m \in \mathbb{K} \ : \ \begin{pmatrix} x_1 - q_1 \\ x_2 - q_2 \\ \vdots \\ x_n - q_n \end{pmatrix} = \begin{pmatrix} t_1 w_{11} + t_2 w_{12} + \cdots + t_m w_{1m} \\ t_1 w_{21} + t_2 w_{22} + \cdots + t_m w_{2m} \\ \vdots \\ t_1 w_{n1} + t_2 w_{n2} + \cdots + t_m w_{nm} \end{pmatrix}.$$

Definizione 2.7. *Usando le stesse notazioni di cui sopra, le equazioni*

$$\begin{cases} x_1 = q_1 + t_1 w_{11} + t_2 w_{12} + \cdots + t_m w_{1m} \\ x_2 = q_2 + t_1 w_{21} + t_2 w_{22} + \cdots + t_m w_{2m} \\ \vdots \\ x_n = q_n + t_1 w_{n1} + t_2 w_{n2} + \cdots + t_m w_{nm} \end{cases} \tag{2.2}$$

si dicono **equazioni parametriche di** S.

Osservazione 2.8. *Un sottospazio affine $S \subseteq \mathbb{A}^n_{\mathbb{K}}$ può essere descritto da diversi sistemi di equazioni parametriche.*

3. Parallelismo, incidenza e sottospazi affini sghembi

Definizione 3.1. *Sia $\mathbb{A}$ uno spazio affine su uno spazio vettoriale V, e siano*

$$S_1 \subseteq \mathbb{A}, \quad S_2 \subseteq \mathbb{A}$$

due sottospazi affini di giaciture W_1 e W_2 rispettivamente.

- *S_1 e S_2 si dicono* **paralleli**, *in simboli*

  $$S_1 \parallel S_2,$$

  se $W_1 \subseteq W_2$ oppure $W_2 \subseteq W_1$;
- *se S_1 e S_2 non sono paralleli, allora si dicono* **incidenti** *se*

  $$S_1 \cap S_2 \neq \emptyset,$$

  mentre si dicono **sghembi** *se*

  $$S_1 \cap S_2 = \emptyset.$$

4. Passaggio da equazioni cartesiane ad equazioni parametriche

Sia $S \subseteq \mathbb{A}^n_{\mathbb{K}}$ un sottospazio affine dato dalle equazioni cartesiane $A \cdot x = b$. Per trovare delle equazioni parametriche per S, si risolve il sistema di equazioni lineari $A \cdot x = b$, esprimendo le sue soluzioni in funzione di opportune variabili libere $t_1, \dots, t_m$, che equivale a trovare una soluzione particolare Q di $Ax = b$, ed una base $w_1, \dots, w_m$ dello spazio W delle soluzioni del sistema lineare omogeneo associato $A \cdot x = 0$, in modo che

$$S = Q + W = Q + t_1 w_1 + \ldots + t_m w_m,$$

al variare di $t_1, \dots, t_m \in \mathbb{K}$.

5. Passaggio da equazioni parametriche ad equazioni cartesiane

Sia $S \subseteq \mathbb{A}^n_{\mathbb{K}}$ un sottospazio affine dato dalle equazioni parametriche

$$\begin{cases} x_1 = q_1 + t_1 w_{11} + t_2 w_{12} + \cdots + t_m w_{1m} \\ x_2 = q_2 + t_1 w_{21} + t_2 w_{22} + \cdots + t_m w_{2m} \\ \vdots \\ x_n = q_n + t_1 w_{n1} + t_2 w_{n2} + \cdots + t_m w_{nm} \end{cases}$$

Osserviamo che un punto $P = (x_1, x_2, \dots, x_n) \in \mathbb{A}^n_{\mathbb{K}}$ appartiene ad S se e solo se il seguente sistema lineare delle indeterminate $t_1, t_2, \dots, t_m$ ammette soluzione

$$\overbrace{\begin{pmatrix} w_{11} & \dots & w_{1m} \\ w_{21} & \dots & w_{2m} \\ \vdots & \vdots & \vdots \\ w_{n1} & \dots & w_{nm} \end{pmatrix}}^{A} \cdot \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_m \end{pmatrix} = \overbrace{\begin{pmatrix} x_1 - q_1 \\ x_2 - q_2 \\ \vdots \\ x_n - q_n \end{pmatrix}}^{b}.$$

Ciò è equivalente all'annullamento delle ultime $n - r$ componenti del vettore $\widetilde{b}$, dove $(\widetilde{A}|\widetilde{b})$ è una matrice che si ottiene da $(A|b)$ mediante operazioni elementari e $\widetilde{A}$ è a scala. Osserviamo che $\dim S = m = \operatorname{rg} A$. Quindi, se $\widetilde{b} = \begin{pmatrix} \widetilde{b}_1 \\ \widetilde{b}_2 \\ \vdots \\ \widetilde{b}_n \end{pmatrix}$, delle equazioni cartesiane per S sono date da

$$\begin{cases} \widetilde{b}_{m+1} = 0 \\ \widetilde{b}_{m+2} = 0 \\ \dots \\ \widetilde{b}_n = 0. \end{cases}$$

Capitolo 9

Definizione di applicazione lineare e prime proprietà

Definizione 9.0.1. Siano V e V' due spazi vettoriali su un campo $\mathbb{K}$. Una **applicazione lineare da V in V'** è una funzione $f : V \to V'$ che soddisfa le seguenti proprietà:

- (AL1) **additività:** $f(v_1 + v_2) = f(v_1) + f(v_2)$**, per ogni** $v_1, v_2 \in V$;
- (AL2) **omogeneità:** $f(c \cdot v) = c \cdot f(v)$**, per ogni** $c \in \mathbb{K}$ **e per ogni** $v \in V$.

Una applicazione lineare da V in V:

$$f : V \to V$$

è anche detta un **operatore lineare di** V.

Una applicazione lineare **biettiva** $f : V \to V'$ si dice **isomorfismo**.

Osservazione 9.0.1. Si osservi che nella proprietà (AL1), la somma $v_1 + v_2$ è quella in V , mentre $f(v_1) + f(v_2)$ è la somma dei vettori $f(v_1)$ e $f(v_2)$ in V'. Analogamente, in (AL2), $c \cdot v$ è il prodotto del vettore v per lo scalare c in V, mentre $c \cdot f(v)$ è il prodotto dello scalare c per il vettore $f(v)$ in V'. Per questo motivo si dice anche che una applicazione lineare $f : V \to V'$ **rispetta le operazioni di somma e di prodotto per scalari** di V e V'.

Osservazione 9.0.2. Sia $f : V \to V'$ un'applicazione lineare. Allora si ha:

1. $f(0_V) = 0_{V'}$. Infatti possiamo scrivere

$$0_V = 0 \cdot v, \ 0 \in \mathbb{K}, \ v \in V,$$

e per l'omogeneità (AL2) si ha

$$f(0_V) = f(0 \cdot v) = 0 \cdot f(v) = 0_{V'}.$$

2. Siano $v_1, \dots, v_k$ dei vettori di V. Allora per ogni $v \in \mathrm{Span}(v_1, \dots, v_k)$ si ha $v = \alpha_1 v_1 + \cdots + \alpha_k v_k$ per opportuni $\alpha_1, \dots, \alpha_k$, e per la linearità di f abbiamo

$$f(v) = f(\alpha_1 v_1 + \cdots + \alpha_k v_k) = \alpha_1 f(v_1) + \cdots + \alpha_k f(v_k);$$

quindi $f(v) \in \mathrm{Span}(f(v_1), \ldots, f(v_k))$.

In particolare, se $v_1, \ldots, v_n$ è una BASE d V, per ogni $v \in V$ si ha che

$$f(v) \in \mathrm{Span}(f(v_1), \ldots, f(v_n))).$$

Inoltre, è facile verificare che ogni vettore di $\mathrm{Span}(f(v_1), \ldots, f(v_n)))$ è immagine di un vettore di V. Si ha quindi che le immagini dei vettori di una base di V generano il sottospazio immagine $\mathrm{Im}(f)$:

$$\mathrm{Im}(f) = \mathrm{Span}(f(v_1), \ldots, f(v_n)))$$

3. Indichiamo con

$$\mathcal{L}(V, V') := \{\text{applicazioni lineari } V \to V'\},$$

o anche

$$\mathrm{Hom}(V, V') := \{\text{applicazioni lineari } V \to V'\}.$$

In $\mathcal{L}(V, V')$ possiamo definire una somma di applicazioni e il prodotto per scalari nel modo usuale (puntualmente). Con queste operazioni $\mathcal{L}(V, V')V'$ risulta uno **spazio vettoriale su** $\mathbb{K}$.

Esempio 9.0.2. L'applicazione costante nulla $0 : V \to V'$, che associa ad ogni vettore $v \in V$ il vettore nullo $0_{V'}$ è un'applicazione lineare.

Si osservi che tra le applicazioni costanti, quella nulla è l'unica che risulta lineare.

Esempio 9.0.3. La funzione **identità**

$$\mathrm{Id}_V : V \to V, \quad \mathrm{Id}_V(v) = v$$

è un operatore lineare.

Esempio 9.0.4. Fissata una base $\mathcal{B} = \{v_1, \ldots, v_n\}$ di V, la **funzione coordinate**

$$F_{\mathcal{B}} : V \to \mathbb{K}^n, \quad \text{se } v = c_1 v_1 + \ldots + c_n v_n,$$

$$F_{\mathcal{B}}(v) = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = n - \text{upla delle coordinate di } v \text{ rispetto a } \mathcal{B}$$

è un'applicazione lineare biiettiiva, cioè un isomorfismo.

Esempio 9.0.5. La **proiezione**

$$p_1 : \mathbb{R}^2 \to \mathbb{R}^2, \quad p_1 \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} = \begin{pmatrix} r_1 \\ 0 \end{pmatrix}$$

è un'applicazione lineare.

Esempio 9.0.6. La **rotazione di angolo** α

$$F_\alpha : \mathbb{R}^2 \to \mathbb{R}^2, \quad F_\alpha \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} = \begin{pmatrix} \cos(\alpha) r_1 - \sin(\alpha) r_2 \\ \sin(\alpha) r_1 + \cos(\alpha) r_2 \end{pmatrix}$$

è un'applicazione lineare.

Esempio 9.0.7. Data una matrice qualunque

$$A \in M_{m,n}(\mathbb{K},$$

possiamo definire la seguente applicazione lineare

$$L_A : \mathbb{K}^n \to \mathbb{K}^m, \quad L_A(v) = A \cdot v,$$

dove v è un vettore colonna di $\mathbb{K}^n$ e $A \cdot v$ è la moltiplicazione riga per colonna. La proprietà (AL1) segue dalla proprietà distributiva del prodotto righe per colonne rispetto alla somma, mentre la (AL2) segue dal fatto che la moltiplicazione per scalari commuta con il prodotto righe per colonne.

Esempio 9.0.8. Sia

$$V = \mathcal{C}^1((a,b), \mathbb{R})$$

lo spazio vettoriale delle funzioni derivabili su un intervall0 aperto $(a,b) \subseteq \mathbb{R}$ e con derivata continua, e sia

$$V' = \mathcal{C}^0((a,b), \mathbb{R})$$

lo spazio vettoriale delle funzioni continue sull'intervallo $(a,b) \subseteq \mathbb{R}$.

L' applicazione di **derivazione**

$$D : \mathcal{C}^1((a,b), \mathbb{R}) \to \mathcal{C}^0((a,b), \mathbb{R}), \quad D\, f(t) = f'(t)$$

è lineare.

Esempio 9.0.9. Sia $V = \mathcal{C}^0([a,b], \mathbb{R})$. La funzione

$$I : V \to \mathbb{R}, \quad I(p(x)) = \int_a^b p(x)\, dx$$

è una applicazione lineare.

Teorema 9.0.10. Teorema di struttura per le applicazioni lineari

Siano V e V' due spazi vettoriali sul campo $\mathbb{K}$. Sia $\{v_1, \dots, v_n\}$ una base di V, e siano $w_1, \dots, w_n \in V'$ dei vettori di V' .

Allora esiste un'unica applicazione lineare $f : V \to V'$, tale che

$$f(v_1) = w_1, \quad f(v_2) = w_2, \quad \dots, \quad f(v_n) = w_n.$$

Dimostrazione. **Esistenza** Definiamo una applicazione $f : V \to V'$ come segue: per ogni vettore $v \in V$, esso si esprime in modo unico come combinazione lineare dei vettori della base $\{v_1, \dots, v_n\}$ di V:

$$v = \alpha_1 v_1 + \dots + \alpha_n v_n,$$

dove $\alpha_1, \dots, \alpha_n$ sono le coordinate di v rispetto alla base data; allora definiamo

$$f(v) := \alpha_1 w_1 + \dots + \alpha_n w_n \in V'.$$

Si verifica facilmente che tale f è lineare (esercizio) e, per definizione, si ha che $f(v_i) = w_i$, per ogni $i = 1, \dots, n$.

Unicità Siano f e g due applicazioni lineari che soddisfano la proprietà dell'enunciato $f(v_i) = w_i = g(v_i)$, per ogni $i = 1, \dots, n$. Dimostriamo che $f = g$,cioè che $f(v) = g(v)$ per ogni $v \in V$. A tale scopo sia $v \in V$ un vettore arbitrario. Scriviamo v come combinazione lineare dei vettori della base $\{v_1, \dots, v_n\}$:

$$v = \beta_1 v_1 + \cdots + \beta_n v_n.$$

Sfruttando le proprietà di linearità (AL1) e (AL2) si ha:

$$f(v) = \beta_1 f(v_1) + \cdots + \beta_n f(v_n) = \beta_1 w_1 + \cdots + \beta_n w_n =$$

$$= \beta_1 g(v_1) + \cdots + \beta_n g(v_n) = g(\beta_1 v_1 + \cdots + \beta_n v_n) = g(v).$$

□

9.1 Nucleo e immagine

Definizione 9.1.1. Sia $f : V \to V'$ un'applicazione lineare. Il **nucleo** di f è il sottoinsieme di V

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

L'**immagine** di f è il sottoinsieme di V'

$$\mathrm{Im}(f) = \{w \in V' \mid \exists v \in V : f(v) = w\}.$$

Proposizione 9.1.2. *Sia $f : V \to V'$ un'applicazione lineare. Allora si ha:*

1. $\ker(f)$ *è un sottospazio vettoriale di V.*
2. $\mathrm{Im}(f)$ *è un sottospazio vettoriale diV'.*
3. *f è iniettiva se e solo se* $\ker(f) = \{0\}$.
4. *f è suriettiva se e solo se* $\mathrm{Im}(f) = V'$.

Dimostrazione. 1. Abbiamo già osservato che si ha sempre $0_V \in \ker(f)$. Dimostriamo ora che $\ker(f)$ è chiuso per la somma e per il prodotto per scalari.

Siano $v, \tilde{v} \in \ker(f)$; quindi $f(v) = 0_{V'}$ e $f(\tilde{v}) = 0_{V'}$. Per il vettore somma $v + \tilde{v}$, usando l'additività (AL1) di f, si ha che

$$f(v + \tilde{v}) = f(v) + f(\tilde{v}) = 0_{V'} + 0_{V'} = 0_{V'},$$

quindi anche $v + \tilde{v} \in \ker(f)$. Supponiamo infine $v \in \ker(f)$ e sia $c \in \mathbb{K}$ arbitrario. Per il vettore $c \cdot v$ si ha:

$$f(c \cdot v) = c\, f(v) = c \cdot 0_{V'} = 0_{V'}$$

per l'omogeneità (AL2) di f, quindi $v \cdot v \in \ker(f)$.

2. Siccome $0_{V'} = f(0_V)$, il vettore nullo $0_{V'} \in \mathrm{Im}(f)$. Supponiamo ora che $w, \tilde{w} \in \mathrm{Im}(f)$. Allora esistsono due vettori $v, \tilde{v}\, in V$ tali che
$$w = f(v), \quad \tilde{w} = f(\tilde{v}).$$
Per il vettore somma si ha che
$$w + \tilde{w} = f(v) + f(\tilde{v}) = f(v + \tilde{v})$$
per l'additività (AL1) di f, quindi anche $w + \tilde{w} \in \mathrm{Im}(f)$.
Infine sia $w \in \mathrm{Im}(f)$ e $c \in \mathbb{K}$ arbitrario. Allora esiste un $v \in V$ tale che
$$w = f(v),$$
da cui
$$c \cdot w = c \cdot f(v) = f(c \cdot v)$$
per l'omogeneità (AL2) di f, allora $cw \in \mathrm{Im}(f)$.

3. Supponiamo f iniettiva. Siccome $f(0_V) = 0_{V'}$ e f è iniettiva, non possono esistere altri vettori che abbiano la stessa immagine di 0_V. Quindi $\ker(f) = \{0_V\}$.
Viceversa, supponiamo $\ker(f) = \{0_V\}$, cioè 0_V è l'unico vettore che abbia immagine nulla $0_{V'}$. Siano $v, \tilde{v} \in V$ due vettori diversi:
$$v \neq \tilde{v}.$$
Allora $v - \tilde{v} \neq 0_V$, da cui
$$f(v - \tilde{v}) \neq 0_{V'},$$
per l'ipotesi $\ker(f) = \{0_V\}$. Per la linearità si ha
$$f(v - \tilde{v}) = f(v) - f(\tilde{v}) \neq 0_{V'},$$
cioè
$$f(v) \neq f(\tilde{v}),$$
che è proprio la definizione di iniettività.

4. f è suriettiva se e solo se $\mathrm{Im}(f) = V'$: questa è semplicemente la definizione di suriettività.

□

Definizione 9.1.3. Sia $f : V \to V'$ una applicazione lineare. Il **rango** di f è la dimensione dell'immagine di f e si indica con rg**(f)**:
$$\mathrm{rg}(f) := \dim(\mathrm{Im}(f)).$$

Osservazione 9.1.1. Consideriamo una matrice $A \in M_{m,n}(\mathbb{K})$ e sia
$$L_A : \mathbb{K}^n \to \mathbb{K}^m$$
l'applicazione lineare associata. Per l'Osservazione 9.0.2 (2), se in $\mathbb{K}^n$ fissiamo la base canonica $\{e_1, \ldots, e_n\}$, abbiamo che
$$\mathrm{Im}(L_A) = \mathrm{Span}(L_A(e_1), \ldots, L_A(e_n)).$$

Osserviamo ora che

$$L_A(e_1) = A \cdot \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = A^{(1)}, \ldots, L_A(e_n) = A \cdot \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} = A^{(n)},$$

cioè le immagini dei vettori della base canonica tramite L_A coincidono con le colonne di A. Questo implica che

$$\mathrm{rg}(L_A) \dim \mathrm{Im}(L_A) = \dim \mathrm{Span}(L_A(e_1), \ldots, L_A(e_n)) = \dim \mathrm{Span}(A^{(1)}, \ldots, A^{(n)}) = \mathrm{rg}(A).$$

Notiamo, inoltre, che se $\widetilde{A}$ é ottenuta da A per mezzo di operazioni elementari, allora sappiamo che $\mathrm{rg}(A) = \mathrm{rg}(\widetilde{A})$, ma in generale

$$\mathrm{Im}(L_A) \neq \mathrm{Im}(L_{\widetilde{A}}).$$

Teorema 9.1.4. Teorema di Dimensione *Siano V e V' due spazi vettoriali su un campo $\mathbb{K}$, con V di dimensione finita. Sia $f : V \to V'$ una applicazione lineare. Allora vale la seguente uguaglianza:*

$$\dim(V) = \dim(\ker(f)) + \mathrm{rg}(f).$$

Dimostrazione. Poiché V ha dimensione finita e $\ker(f)$ è un sottospazio vettoriale di V, anche $\ker(f))$ ha dimesnione finita. Fissiamo dunque una base

$$\{v_1, \ldots, v_k\} \quad \text{di } \ker(f),$$

e completiamola a una base di V:

$$\{v_1, \ldots, v_k, u_{k+1}, \ldots, u_n\}.$$

Affermiamo che

$$\{f(u_{k+1}), \ldots, f(u_n)\}$$

è una base di $\mathrm{Im}(f)$ ed osserviamo che da questo segue l'enunciato, perché avremmo

$$\mathrm{rg}(f) = n - k,$$

quindi

$$\dim(V) = n = k + (n-k) = \dim(ker(f)) + \mathrm{rg}(f).$$

Dalla Osservazione 9.0.2 (2) sappiamo che

$$\mathrm{Im}(f) = \mathrm{Span}(f(v_1), \ldots, f(v_k), f(u_{k+1}), \ldots, f(u_n)) = \mathrm{Span}(f(u_{k+1}), \ldots, f(u_n))$$

perché $v_1, \ldots, v_k \in \ker(f)$. Dunque i vettori $\{f(u_{k+1}), \ldots, f(u_n)\}$ sono un insieme di generatori.

Ci rimane da dimostrare che $\{f(u_{k+1}), \ldots, f(u_n)\}$ sono linearmente indipendenti.

Consideriamo una combinazione lineare che dia il vettore nullo:

$$c_{k+1} f(u_{k+1}) + \cdots + c_n f(u_n) = 0_{V'}. \tag{9.1.1}$$

Per linearitá possiamo scrivere

$$0_{V'} = c_{k+1}f(u_{k+1}) + \cdots + c_n f(u_n) = f(c_{k+1}u_{k+1} + \cdots + c_n u_n).$$

Segue che il vettore

$$w = c_{k+1}u_{k+1} + \cdots + c_n u_n \in \ker(f).$$

Come ogni vettore di $\ker(f)$, w si puó esprimere come combinazione lineare dei vettori della base $\{v_1, \dots, v_k\}$ di $\ker(f)$:

$$w = a_1 v_1 + \cdots + a_k v_k.$$

In definitiva abbiamo due espressioni di w come combinazione lineare di vettori linearmente indipendenti:

$$w = c_{k+1}u_{k+1} + \cdots + c_n u_n = a_1 v_1 + \cdots + a_k v_k.$$

Portando le ultime due espressioni allo stesso membro troviamo

$$a_1 v_1 + \cdots + a_k v_k - c_{k+1}u_{k+1} - \cdots - c_n u_n = 0_V,$$

cioè una combinazione lineare dei vettori della base di V che dá il vettore nullo, e per l'indipendenza lineare dei vettori di una base, tutti i coefficienti sono necessariamente nulli:

$$a_1 = 0, \dots, a_k = 0, c_{k+1} = 0, \dots, c_n = 0,$$

in particolare i coefficienti nell'espressione (9.1.1) sono tutti nulli, quindi $f(u_{k+1}), \dots, f(u_n)$ sono linearmente indipendenti.

$\square$

Osservazione 9.1.2. Sia $f : V \to V'$ un' applicazione lineare. Ricordiamo che per ogni vettore $w \in V'$, $f^{-1}(w)$ denota la pre-immagine di w tramite f, cioè

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

Allora si ha che

$$f^{-1}(w) = \tilde{v} + \ker(f),$$

dove $\tilde{v}$ è un qualsiasi elemento di $f^{-1}(w)$. Questo segue, ad esempio, tramite una dimostrazione analoga a quella del Teorema di Struttura per le soluzioni di un sistema lineare (la verifica è lasciata come esercizio). Dal Teorema della dimensione si ha che $\dim(\ker(f)) = \dim(V) - \mathrm{rg}(f)$.

Si osservi che in questo modo, nel caso in cui

$$V = \mathbb{K}^n,\ V' = \mathbb{K}^m,\ f = L_A,$$

per qualche $A \in M_{m,n}(\mathbb{K})$, si ritrova il Teorema di Struttura per le soluzioni di un sistema lineare ed il teorema di Rouché-Capelli.

Corollario 9.1.5. *Siano V e V' due spazi vettoriali di dimensione finita su un campo $\mathbb{K}$. Supponiamo che*

$$\dim(V) = \dim(V').$$

Sia $f : V \to V'$ una applicazione lineare. Allora le seguenti affermazioni sono equivalenti:

1. $\ker(f) = \{0\}$, *cioè f è iniettiva;*

2. $\mathrm{Im}(f) = V'$, *cioè f è suriettiva;*

3. *f è un isomorfismo.*

Dimostrazione. Se $\ker(f) = \{0\}$, allora $\dim(\ker(f)) = 0$. Per il Teorema della Dimensione si ha

$$\dim(\mathrm{Im}(f)) = \dim(V) - \dim(\ker(f)) = \dim(V) - 0 = \dim(V).$$

Per ipotesi $\dim(V) = \dim(V')$, ne segue che $\dim(\mathrm{Im}(f)) = \dim(V')$, e poiché $\mathrm{Im}(f)$ è un sottospazio vettoriale di V', quest'ultima condizione è equivalente a $\mathrm{Im}(f) = V'$. Quindi 1. e 2. sono equivalenti tra di loro.

Supponiamo ora che valga la condizione 2. (oppure la 1.). Per quanto appena dimostrato, vale anche la condizione 1. (rispettivamente la 2.), quindi f è un isomorfismo. Viceversa, se f è un isomorfismo, è iniettiva e suriettiva per definizione, quindi valgono 1. e 2. □

Definizione 9.1.6. Due spazi vettoriali V e V' su un campo $\mathbb{K}$ si dicono isomorfi se esiste un isomorfismo

$$f : V \to V'.$$

In tal caso si scrive $V \cong V'$.

Osservazione 9.1.3. L'isomorfismo è una relazione di equivalenza tra gli spazi vettoriali su uno stesso campo $\mathbb{K}$.

Teorema 9.1.7. *Siano V e V' due spazi vettoriali di dimensione finita su un campo $\mathbb{K}$. Allora $V \cong V'$ se e solo se $\dim(V) = \dim(V')$.*

Dimostrazione. Sia $f : V \to V'$ un isomorfismo. Per il precedente corollario abbiamo $\ker(f) = \{0\}$ ed $\mathrm{Im}(f) = V'$, quindi per il Teorema della Dimensione $\dim(V) = \mathrm{rg}(f) = \dim(V')$.

Viceversa, supponiamo ora che $\dim(V) = \dim(V') = n$. Fissiamo una base $\{v_1, \ldots, v_n\}$ di V ed una base $\{w_1, \ldots, w_n\}$ di V'. Per il Teorema di Struttura per le applicazioni lineari esiste un' unica applicazione lineare $f : V \to V'$ tale che $f(v_i) = w_i$ per ogni $i = 1, \ldots, n$. Essendo $\mathrm{Im}(f) = \mathrm{Span}(f(v_1), \ldots, f(v_n)) = \mathrm{Span}(w_1, \ldots, w_n) = V'$, quindi f è suriettiva. Dal precedente Corollario abbiamo che f è un isomorfismo, quindi $V \cong V'$. □

Corollario 9.1.8. *Se $n \neq m$, gli spazi vettoriali $\mathbb{K}^n$ e $\mathbb{K}^m$ non sono isomorfi.*

9.2 Matrici associate a un'applicazione lineare

Abbiamo visto che ad ogni matrice $A \in M_{m,n}(\mathbb{K})$ corrisponde una applicazione lineare $L_A : \mathbb{K}^n \to \mathbb{K}^m$, con $L_A(v) := A \cdot v$. In questa sezione vedremo che, viceversa, data una applicazione lineare $f : V \to V'$ tra spazi vettoriali di dimensione finita, fissata una base $\mathcal{B}$ di V ed una base $\mathcal{C}$ di V' , possiamo associare una matrice $M^{\mathcal{B}}_{\mathcal{C}}(f)$ (la matrice che rappresenta f rispetto alle basi $\mathcal{B}$ e $\mathcal{C}$), tale che se le coordinate di un vettore $v \in V$ riepsetto alla base $\mathcal{B}$ sono

$$\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix},$$

le coordinate di $f(v) \in V'$ rispetto alla base $\mathcal{C}$ sono date da

$$M^{\mathcal{B}}_{\mathcal{C}}(f) \cdot \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}.$$

Definizione 9.2.1. Siano V e V' due spazi vettoriali di dimensione finita su $\mathbb{K}$. Sia $f : V \to V'$ una applicazione lineare. Siano $\mathcal{B} = \{v_1, \dots, v_n\}$ una base di V e $\mathcal{C} = \{w_1, \dots, w_m\}$ una base di V'. La **matrice che rappresenta f rispetto alle basi $\mathcal{B}$ e $\mathcal{C}$** è definita come segue:

$$M^{\mathcal{B}}_{\mathcal{C}}(f) = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \in M_{m,n}(\mathbb{K}),$$

dove

$$f(v_j) = a_{1j}w_1 + a_{2j}w_2 + \cdots + a_{mj}w_m.$$

In altre parole, **la colonna j-ma di $M^{\mathcal{B}}_{\mathcal{C}}(f)$ è formata dalle coordinate di $f(v_j)$ rispetto alla base $\mathcal{C}$ per ogni** $j = 1, \dots, n$.

Osservazione 9.2.1. Sia $A \in M_{m,n}(\mathbb{K})$ e consideriamo la applicazione lineare $L_A : \mathbb{K}^n \to \mathbb{K}^m$. Siano $\mathcal{E}$ e $\mathcal{E}'$ le basi canoniche di $\mathbb{K}^n$ e $\mathbb{K}^m$, rispettivamente. Allora si ha

$$M^{\mathcal{E}}_{\mathcal{E}'}(L_A) = A.$$

Infatti, se e_j denota il j-mo vettore della base canonica $\mathcal{E}$ di $\mathbb{K}^n$, allora

$$L_A(e_j) = A \cdot e_j = A^{(j)},$$

e le coordinate di $A^{(j)}$ rispetto alla base canonica $\mathcal{E}'$ di $\mathbb{K}^m$ coincidono con gli scalari della colonna stessa.

Osservazione 9.2.2. Sia f la applicazione nulla, cioè l'applicazione

$$f(v) = 0, \quad \forall v \in V.$$

Allora

$$M^{\mathcal{B}}_{\mathcal{C}}(f) = 0$$

è la matrice nulla per ogni scelta di $\mathcal{B}$ e $\mathcal{C}$. Infatti $f(v_j) = 0$ per ogni $j = 1, \dots, n$, e le coordinate del vettore nullo sono tutte nulle in una qualsiasi base $\mathcal{C}$.

Inoltre, vale anche il viceversa: se $M^{\mathcal{B}}_{\mathcal{C}}(f) = 0$ è la matrice nulla, allora

$$f(v_1) = 0, \dots f(v_n) = 0,$$

e per il Teorema di Struttura per Applicazioni Lineari da ciò segue che f è l'applicazione nulla.

Osservazione 9.2.3. Se $V = V'$ e $f = Id_V$, allora

$$M_{\mathcal{B}}^{\mathcal{B}}(Id_V) = I_n,$$

per ogni base $\mathcal{B}$ di V. Infatti se $\mathcal{B} = \{v_1, \dots, v_n\}$ si ha

$$f(v_j) = v_j,$$

e le coordinate di v_j nella base $\mathcal{B}$ sono tutte nulle eccetto la coordinata j-esima che vale 1.

Più in generale, se $V = V'$ e $f = c \cdot Id_V$, per qualche scalare $c \in \mathbb{K}$, allora

$$M_{\mathcal{B}}^{\mathcal{B}}(c \cdot Id_V) = c \cdot I_n,$$

la verifica è lasciata per esercizio. Tale applicazione si chiama **dilatazione**.

Proposizione 9.2.2. *Siano V e V' due spazi vettoriali su $\mathbb{K}$ di dimensione finita, sia $f : V \to V'$ una applicazione lineare, e siano $\mathcal{B} = \{v_1, \dots, v_n\}$ una base di V e $\mathcal{C} = \{w_1, \dots, w_m\}$ una base di V'. Dato un vettore $v \in V$, se*

$$v = \alpha_1 v_1 + \dots + \alpha_n v_n,$$

allora

$$f(v) = \beta_1 w_1 + \dots + \beta_m w_m,$$

dove

$$\begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{pmatrix} = M_{\mathcal{C}}^{\mathcal{B}}(f) \cdot \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}$$

Dimostrazione. L'enunciato segue dalla seguente sequenza di uguaglianze, dove si sfrutta il fatto che f è lineare e la definizione di $M_{\mathcal{C}}^{\mathcal{B}}(f)$:

$$\begin{aligned} f(v) &= f(\alpha_1 v_1 + \dots + \alpha_n v_n) = \\ &= \alpha_1 f(v_1) + \dots \alpha_n f(v_n) = \\ &= \alpha_1(a_{11} w_1 + \dots + a_{m1} w_m) + \dots + \alpha_n(a_{1n} w_1 + \dots + a_{mn} w_m) = \\ &= (\alpha_1 a_{11} + \dots + \alpha_n a_{1n}) w_1 + \dots + (\alpha_1 a_{m1} + \dots + \alpha_n a_{mn}) w_m. \end{aligned}$$

Infine, osserviamo che per ogni $i = 1, \dots, m$, lo scalare

$$\alpha_1 a_{i1} + \dots + \alpha_n a_{in}$$

è proprio il coefficiente i-esimo della matrice colonna

$$M_{\mathcal{C}}^{\mathcal{B}}(f) \cdot \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}.$$

□

Uno dei vantaggi che si ottengono descrivendo una applicazione lineare f per mezzo delle matrici $M_{\mathcal{C}}^{\mathcal{B}}(f)$ è che possiamo usare i risultati concernenti i sistemi lineari per determinare $\mathrm{Im}(f)$ e $\ker(f)$ come segue.

Corollario 9.2.3. *Sia $f : V \to V'$ una applicazione lineare, dove V e V' sono spazi vettoriali di dimensione finita. Siano $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base di V e $\mathcal{C} = \{w_1, \ldots, w_m\}$ una base di V'. Sia $M_{\mathcal{C}}^{\mathcal{B}}(f)$ la matrice che rappresenta f nelle basi $\mathcal{B}$ e $\mathcal{C}$. Allora valgono le seguenti uguaglianze:*

$$\ker(f) = \left\{ v = \alpha_1 v_1 + \cdots + \alpha_n v_n \in V \mid M_{\mathcal{C}}^{\mathcal{B}}(f) \cdot \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} = 0 \right\};$$

$$\mathrm{Im}(f) = \left\{ w = \beta_1 w_1 + \cdots + \beta_m w_m \in W \mid \text{il sistema } M_{\mathcal{C}}^{\mathcal{B}}(f) \cdot X = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{pmatrix} \text{ e compatibile} \right\};$$

$$\mathrm{rg}(f) = \mathrm{rg} M_{\mathcal{C}}^{\mathcal{B}}(f).$$

Teorema 9.2.4. *Siano V e V' due spazi vettoriali di dimensione finita sul campo $\mathbb{K}$. Siano $\mathcal{B} = \{v_1, \ldots, v_n\}$ e $\mathcal{C} = \{w_1, \ldots, w_m\}$ basi di V e V', rispettivamente. Allora l' applicazione*

$$M_{\mathcal{C}}^{\mathcal{B}} : \mathcal{L}(V, V') \to M_{m,n}(\mathbb{K}), \quad f \to M_{\mathcal{C}}^{\mathcal{B}}(f),$$

è un isomorfismo di spazi vettoriali. In particolare, $\mathcal{L}(V, V')$ ha dimensione finita pari a $n \cdot m = \dim(V) \cdot \dim(V')$.

Dimostrazione. Dimostriamo dapprima che $M_{\mathcal{C}}^{\mathcal{B}}$ è lineare. Siano $f, g \in \mathcal{L}(V, V')$ e sia $c \in \mathbb{K}$. Denotiamo con a_{ij} (rispettivamente $b_i j$) l'elemento di posto (i, j) di $M_{\mathcal{C}}^{\mathcal{B}}(f)$ (rispettivamente di $M_{\mathcal{C}}^{\mathcal{B}}(g)$). Dobbiamo verificare che l'elemento di posto (i, j) di $M_{\mathcal{C}}^{\mathcal{B}}(f+g)$ coincide con $a_{ij}+b_{ij}$, per ogni $i = 1, \ldots, m$ e $j = 1, \ldots, n$. Per definizione $[M_{\mathcal{C}}^{\mathcal{B}}(f+g)]_{ij}$ è la coordinata i-esima di $(f+g)(v_j)$ rispetto alla base $\mathcal{C}$. Abbiamo le seguenti uguaglianze:

$$(f+g)(v_j) = f(v_j)+g(v_j) = (a_{1j}w_1+\cdots+a_{mj}w_m)+(b_{1j}w_1+\cdots+b_{mj}w_m) = (a_{1j}+b_{1j})w_1+\cdots+(a_{mj}+b_{mj})w_m.$$

Da questo segue che

$$[M_{\mathcal{C}}^{\mathcal{B}}(f+g)]_{ij} = a_{ij} + b_{ij} = [M_{\mathcal{C}}^{\mathcal{B}}(f)]_{ij} + [M_{\mathcal{C}}^{\mathcal{B}}(g)]_{ij},$$

per ogni $i = 1, \ldots, m$ e $j = 1, \ldots, n$, quindi

$$M_{\mathcal{C}}^{\mathcal{B}}(f+g) = M_{\mathcal{C}}^{\mathcal{B}}(f) + M_{\mathcal{C}}^{\mathcal{B}}(g).$$

Analogamente si dimostra che

$$M_{\mathcal{C}}^{\mathcal{B}}(c \cdot f) = c \cdot M_{\mathcal{C}}^{\mathcal{B}}(f),$$

quindi $M_{\mathcal{C}}^{\mathcal{B}}$ è lineare.

Per dimostrare che $M_{\mathcal{C}}^{\mathcal{B}}$ è iniettiva, è sufficiente provare che

$$\ker(M_{\mathcal{C}}^{\mathcal{B}}) = \{0\}.$$

Sia quindi $f \in \mathcal{L}(V, V')$ tale che $M^{\mathcal{B}}_{\mathcal{C}}(f) = 0 \in M_{m,n}(\mathbb{K})$. Per l'Osservazione 9.2.2 si ha che $f = 0$, da cui segue la tesi.

Per concludere, dimostriamo che $M^{\mathcal{B}}_{\mathcal{C}}$ è suriettiva. Sia quindi $A = (a_{ij}) \in M_{m,n}(\mathbb{K})$. Definiamo $f_A \in \mathcal{L}(V, V')$ come l'unica applicazione lineare tale che

$$f_A(v_j) = a_{1j}w_1 + \cdots + a_{mj}w_m, \qquad \text{per ogni } j = 1, \dots, n.$$

Dalla definizione segue facilmente che $M^{\mathcal{B}}_{\mathcal{C}}(f_A) = A$, quindi $M^{\mathcal{B}}_{\mathcal{C}}$ è suriettiva. □

9.3 Cambiamenti di base

In questa sezione vedremo come cambia la matrice $M^{\mathcal{B}}_{\mathcal{C}}(f)$ al variare delle basi $\mathcal{B}$ e $\mathcal{C}$. Questo seguirà dalla seguente proposizione.

Proposizione 9.3.1. *Siano U, V e W tre spazi vettoriali sul campo $\mathbb{K}$, di dimensione p, n, m, rispettivamente. Siano*

$$\mathcal{D} = \{u_1, \dots, u_p\}, \qquad \mathcal{B} = \{v_1, \dots, v_n\}, \qquad \mathcal{C} = \{w_1, \dots, w_m\}$$

basi di U, V e W, rispettivamente.

Siano $g : U \to V$ e $f : V \to W$ applicazioni lineari. Allora si ha

$$M^{\mathcal{D}}_{\mathcal{C}}(f \circ g) = M^{\mathcal{B}}_{\mathcal{C}}(f) \cdot M^{\mathcal{D}}_{\mathcal{B}}(g).$$

Dimostrazione. Si vedano gli appunti della lezione. □

Corollario 9.3.2. *Sia V uno spazio vettoriale di dimensione n sul campo $\mathbb{K}$. Sia $\mathcal{B}$ una base di V. Allora valgono le seguenti affermazioni.*

1. *Sia $f : V \to V$ un operatore lineare. Allora f è un isomorfismo se e solo se $M^{\mathcal{B}}_{\mathcal{B}}(f)$ è invertibile. In tal caso si ha*
$$M^{\mathcal{B}}_{\mathcal{B}}(f^{-1}) = M^{\mathcal{B}}_{\mathcal{B}}(f)^{-1}.$$
2. *Sia $A \in M_n(\mathbb{K})$. Allora A è invertibile se e solo se $\mathrm{rg}(A) = n$.*

Dimostrazione. 1. Supponiamo che $f : V \to V$ sia un isomorfismo, e sia f^{-1} la funzione inversa di f, che è ancora lineare. Allora valgono le uguaglianze

$$f \circ f^{-1} = f^{-1} \circ f = \mathrm{Id}_V.$$

Abbiamo

$$M^{\mathcal{B}}_{\mathcal{B}}(f \circ f^{-1}) = M^{\mathcal{B}}_{\mathcal{B}}(f^{-1} \circ f) = M^{\mathcal{B}}_{\mathcal{B}}(\mathrm{Id}_V) = \mathbb{I}_n.$$

Dalla Proposizione precedente segue che

$$\mathbb{I}_n = M^{\mathcal{B}}_{\mathcal{B}}(f) \cdot M^{\mathcal{B}}_{\mathcal{B}}(f^{-1}) = M^{\mathcal{B}}_{\mathcal{B}}(f^{-1}) \cdot M^{\mathcal{B}}_{\mathcal{B}}(f)$$

e dalla definizione di matrice invertibile e matrice inversa segue l'enunciato.

Viceversa, se $M^{\mathcal{B}}_{\mathcal{B}}(f)$ è invertibile, allora l'unica soluzione del sistema lineare omogeneo

$$M^{\mathcal{B}}_{\mathcal{B}}(f) \cdot X = 0$$

è $X = (M^{\mathcal{B}}_{\mathcal{B}}(f))^{-1} \cdot 0 = 0$. Da ciò segue che $\ker(f) = \{0\}$, quindi f è iniettivo; essendo un operatore $f : V \to V$, è anche suriettivo, quindi è un isomorfismo.

2. Consideriamo la applicazione lineare

$$L_A : \mathbb{K}^n \to \mathbb{K}^n$$

e ricordiamo che risulta

$$A = M^{\mathcal{E}}_{\mathcal{E}}(L_A),$$

dove $\mathcal{E}$ è la base canonica di $\mathbb{K}^n$. Dal punto 1. si ha che A è invertibile se e solo se L_A è un isomorfismo. D'altra parte, L_A è un isomorfismo se e solo se L_A è suriettiva, cioè se e solo se $\mathrm{rg}(L_A) = n$.

□

Definizione 9.3.3. Sia V uno spazio vettoriale di dimensione n sul campo $\mathbb{K}$. Siano

$$\mathcal{B} = \{v_1, \ldots, v_n\} \quad \text{e} \quad \mathcal{C} = \{w_1, \ldots, w_n\}$$

due basi di V. La **matrice del cambiamento di base da $\mathcal{B}$ a $\mathcal{C}$** è la matrice

$$M^{\mathcal{B}}_{\mathcal{C}}(Id_V) \in M_n(\mathbb{K})$$

che rappresenta la funzione identità $Id_V : V \to V$ rispetto alle basi $\mathcal{B}$ e $\mathcal{C}$.

Osservazione 9.3.1.

1. Per ogni $j = 1, \ldots, n$, siano

$$a_{1j}, \ldots, a_{nj}$$

le coordinate di v_j rispetto alla base $\mathcal{C}$. Allora la colonna j-esima di $M^{\mathcal{B}}_{\mathcal{C}}(Id_V)$ è data da

$$\begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix}.$$

2. Per ogni vettore $v \in V$, se

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}$$

è la colonna delle coordinate di v rispetto alla base $\mathcal{B}$, allora la colonna delle coordinate di v rispetto alla base $\mathcal{C}$ è data da

$$\begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} = M^{\mathcal{B}}_{\mathcal{C}}(Id_V) \cdot \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}.$$

Quindi $M^{\mathcal{B}}_{\mathcal{C}}(Id_V)$ permette di determinare le coordinate di un vettore rispetto alla base $\mathcal{C}$ se sono note le sue coordinate rispetto a $\mathcal{B}$, per questo prende il nome di matrice del cambiamento di base da $\mathcal{B}$ a $\mathcal{C}$.

Proposizione 9.3.4. *La matrice di cambio di base $M^{\mathcal{B}}_{\mathcal{C}}(Id_V)$ è invertibile, e vale la seguente uguaglianza:*

$$M^{\mathcal{B}}_{\mathcal{C}}(Id_V)^{-1} = M^{\mathcal{C}}_{\mathcal{B}}(Id_V).$$

Dimostrazione. Valgono le seguenti identità:

$$\mathbb{I}_n = M^{\mathcal{B}}_{\mathcal{B}}(Id_V), \qquad Id_V = Id_V \circ Id_V, \qquad M^{\mathcal{B}}_{\mathcal{B}}(Id_V \circ Id_V) = M^{\mathcal{C}}_{\mathcal{B}}(Id_V) \cdot M^{\mathcal{B}}_{\mathcal{C}}(Id_V).$$

Da questo segue l'enunciato. □

Come conseguenza otteniamo una formula per passare da una matrice che rappresenta un'applicazione f rispetto a certe basi alla matrice che rappresenta f in altre basi:

Corollario 9.3.5. *Sia $f : V \to V'$ un'applicazione lineare. Siano $\mathcal{B}$ e $\mathcal{B}'$ due basi di V e $\mathcal{C}$ e $\mathcal{C}'$ due basi di V'. Allora vale la relazione*

$$M^{\mathcal{B}'}_{\mathcal{B}'}(f) = M^{\mathcal{B}'}_{\mathcal{B}}(Id_V) \cdot M^{\mathcal{B}}_{\mathcal{C}}(f) \cdot M^{\mathcal{C}}_{\mathcal{C}'}(Id_{V'}), \tag{9.3.1}$$

dove $Id_V : V \to V$ e $Id_{V'} : V' \to V'$ sono le applicazioni identità.

In particolare, se $V = V'$, $\mathcal{C} = \mathcal{B}$ e $\mathcal{B}' = \mathcal{C}'$, si ha

$$M^{\mathcal{B}}_{\mathcal{B}}(f) = \left(M^{\mathcal{B}'}_{\mathcal{B}}(Id_V)\right)^{-1} \cdot M^{\mathcal{B}'}_{\mathcal{B}'}(f) \cdot M^{\mathcal{B}'}_{\mathcal{B}}(Id_V). \tag{9.3.2}$$

Capitolo 10

Diagonalizzazione

Sia V uno spazio vettoriale di dimensione finita n sul campo $\mathbb{K}$, e sia $f : V \to V$ un operatore lineare. Abbiamo visto che, scelta una base $\mathcal{B}$ di V , possiamo rappresentare f per mezzo di una matrice $M^{\mathcal{B}}_{\mathcal{B}}(f)$. Molte proprietà di f possono essere studiate tramite la matrice $M^{\mathcal{B}}_{\mathcal{B}}(f)$. L'obiettivo di questa sezione è di stabilire se, data f, esiste una base di V tale che la matrice che rappresenta f rispetto a tale base sia diagonale.

Ricordiamo che, se $\mathcal{C}$ è un'altra base di V, allora le matrici $M^{\mathcal{C}}_{\mathcal{C}}(f)$ ed $M^{\mathcal{B}}_{\mathcal{B}}(f)$ solo collegate dalla formula:

$$M^{\mathcal{C}}_{\mathcal{C}}(f) = M^{\mathcal{B}}_{\mathcal{C}}(Id_V)^{-1} \cdot M^{\mathcal{B}}_{\mathcal{B}}(f) \cdot M^{\mathcal{B}}_{\mathcal{C}}(Id_V). \tag{10.0.1}$$

Questo motiva la seguente definizione.

Definizione 10.0.1. Due matrici $A, B \in M_n(\mathbb{K})$ sono **simili** se esiste C matrice invertibile tale che

$$B = C^{-1} \cdot A \cdot C.$$

In tal caso si scrive $A \sim B$.

Quindi, se due matrici rappresentano lo stesso operatore rispetto a due basi, allora esse sono simili. Vale anche il viceversa, come afferma il seguente risultato.

Lemma 10.0.2. *Sia $f : V \to V$ un operatore lineare, e sia $A = M^{\mathcal{B}}_{\mathcal{B}}(f) \in M_n(\mathbb{K})$, dove $\mathcal{B} = \{v_1, \dots, v_n\}$ è una base di V. Allora, se $B \in M_n(\mathbb{K})$, si ha che $B \sim A$ se e solo se esiste una base $\mathcal{C}$ di V tale che*

$$B = M^{\mathcal{C}}_{\mathcal{C}}(f).$$

Dimostrazione. Se esiste una base $\mathcal{C}$ di V tale che $B = M^{\mathcal{C}}_{\mathcal{C}}(f)$, allora A e B sono simili per la relazione (10.0.1).

Viceversa, supponiamo che A e B siano simili. Quindi per definizione esiste una matrice invertibile C tale che

$$B = C^{-1} \cdot A \cdot C.$$

Sia c_{ij} l'elemento di posto (i, j) di C. Definiamo una base $\mathcal{C} = \{w_1, \dots, w_n\}$ di V come segue:

$$w_j := c_{1j}v_1 + \cdots + c_{nj}v_n, \qquad j = 1, \dots, n.$$

Con questa costruzione si ha che

$$C = M^{\mathcal{C}}_{\mathcal{B}}(Id_V),$$

quindi, essendo C invertibile, si ha in particolare che gli n vettori $w_1, \dots, w_n$ sono linearmente indipendenti, e siccome $\dim V = n$ sono anche dei generatori. Inoltre possiamo scrivere

$$B = C^{-1} \cdot A \cdot C = M^{\mathcal{B}}_{\mathcal{C}}(Id_V) \cdot M^{\mathcal{B}}_{\mathcal{B}}(f) \cdot M^{\mathcal{C}}_{\mathcal{B}}(Id_V) = M^{\mathcal{C}}_{\mathcal{C}}(f).$$

□

Osservazione 10.0.1. La similitudine è una relazione di equivalenza tra le matrici a coefficienti in $\mathbb{K}$. La verifica è lasciata per esercizio.

Osservazione 10.0.2. La matrice unità $\mathbb{I}_n$ è simile solo a se stessa, la verifica è molto semplice.

Lemma 10.0.3. *Siano $A, B \in M_n(\mathbb{K})$. Se $A \sim B$, allora*

$$\det(A) = \det(B).$$

Dimostrazione. Se $A \sim B$, allora esiste C invertibile tale che

$$B = C^{-1} \cdot A \cdot C.$$

Da ciò segue

$$\det B = \det(C^{-1} \cdot A \cdot C) = \det C^{-1} \cdot \det A \cdot \det C$$

per il Teorema di Binet. Abbiamo visto che

$$\det C^{-1} = \frac{1}{\det C},$$

ed essendo la moltiplicazione tra scalari commutativa, abbiamo

$$\det C^{-1} \cdot \det A \cdot \det C = \det A.$$

□

Osservazione 10.0.3. Si può dimostrare più in generale (esercizio facoltativo) che due matrici simili hanno lo stesso rango.

Osservazione 10.0.4. Notiamo che non vale il viceversa, cioè esistono matrici quadrate con lo stesso determinante, che NON sono simili. Ad esempio, le matrici

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

hanno entrambe determinante 0, ma non sono simili, ad esempio perchè hanno rango diverso:

$$\mathrm{rg} A = 0, \qquad \mathrm{rg} B = 1.$$

Nemmeno le matrici

$$\mathbb{I}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

sono simili, sebbene si abbia $\det \mathbb{I}_2 = 1 = \det B$. Infatti, $\mathbb{I}_2$ è simile solo a se stessa.

Il Lemma 10.0.3 ci permette di definire il determinante di un endomorfismo come segue:

Definizione 10.0.4. Sia V uno spazio vettoriale di dimensione finita su K. Sia f ? End(V). Allora det(f) := det(MB(f)), dove B 'e una base di V . Osserviamo che, per il Lemma 2, det(f) non dipende dalla scelta della base B, quindi la definizione 'e ben posta.

Definizione 10.0.5. Sia V uno spazio vettoriale di dimensione finita n sul campo $\mathbb{K}$ e sia $f : V \to V$ un operatore lineare. Allora f si dice **diagonalizzabile** se esiste una base $\mathcal{B}$ di V tale che $M_{\mathcal{B}}^{\mathcal{B}}(f)$ è diagonale. In tal caso, $\mathcal{B}$ è detta **base di** V **che diagonalizza** f.

Una matrice $A \in M_n(\mathbb{K})$ è detta **diagonalizzabile** se è simile ad una matrice diagonale.

Osservazione 10.0.5. Sia $A \in M_n(\mathbb{K})$. Allora A è diagonalizzabile se e solo se l'operatore $L_A : \mathbb{K}^n \to \mathbb{K}^n$ è diagonalizzabile.

Questo segue dal fatto che $A = M_{\mathcal{E}}^{\mathcal{E}}(L_A)$, dove $\mathcal{E}$ è la base canonica di $\mathbb{K}^n$.

Osservazione 10.0.6. Se $\dim(V) = 1$, allora ogni operatore $f : V \to V$ è diagonalizzabile ed ogni base di V diagonalizza f.

Vedremo che se invece $\dim(V) > 1$, allora esistono operatori che non sono diagonalizzabili, come ad esempio le rotazioni del piano di un angolo $\alpha \neq 0, \pi$.

Il seguente risultato è una riformulazione della Definizione 10.0.5 che motiva le definizioni che seguono.

Lemma 10.0.6. *Sia $f : V \to V$ un operatore lineare. Sia $\mathcal{B} = \{v_1, \dots, v_n\}$ una base di V che diagonalizza f. Allora, per ogni $i = 1, \dots, n$ si ha*

$$f(v_i) = \lambda_i v_i,$$

per opportuni $\lambda_i \in \mathbb{K}$.

Viceversa, se eiste una base $\mathcal{B}$ di V con tali proprietà, allora f è diagonalizzabile e $\mathcal{B}$ diagonalizza f.

Dimostrazione. Per definizione abbiamo che $M_{\mathcal{B}}^{\mathcal{B}}(f)$ è diagonale,quindi del tipo

$$M_{\mathcal{B}}^{\mathcal{B}}(f) = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}.$$

Dalla definizione di matrice associata ad f nella base $\mathcal{B}$ segue che

$$\begin{gathered} f(v_1) = a_{11}v_1 + 0\, v_2 + \dots + 0\, v_n = a_{11}v_1, \\ f(v_2) = 0\, v_1 + a_{22}\, v_2 + \dots + 0\, v_n = a_{22}v_2, \\ \dots \\ f(v_n) = 0\, v_1 + 0\, v_2 + \dots + a_{nn}\, v_n = a_{nn}v_n. \end{gathered}$$

Basta allora porre $\lambda_i = a_{ii}$.

Viceversa, se $\mathcal{B} = \{v_1, \dots, v_n\}$ è una base di V tale che $f(v_j) = \lambda_j v_j$ per ogni $j = 1, \dots, n$, allora

$$M_{\mathcal{B}}^{\mathcal{B}}(f) = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix},$$

che è diagonale, quindi f è diagonalizzabile e $\mathcal{B}$ diagonalizza f. $\square$

Definizione 10.0.7. Sia V uno spazio vettoriale su un campo $\mathbb{K}$. Sia $f : V \to V$ un operatore. Un **autovettore di** f è un vettore $v \in V$ diverso dal vettore nullo, $v \neq 0$, tale che esiste $\lambda \in \mathbb{K}$ per cui vale

$$f(v) = \lambda v.$$

In tal caso, λ si dice **autovalore di** f **relativo all'autovettore** v.

Lo **spettro di** f è l'insieme degli autovalori di f, esso si denota con $Sp(f)$ ed è un sottoinsieme finito del campo $\mathbb{K}$.

Sia $A \in M_n(\mathbb{K})$. Un **autovettore di** A è un vettore $v \in \mathbb{K}^n$ non nullo tale che v sia un autovettore di L_A, cioè tale che $A \cdot v = \lambda v$, per qualche $\lambda \in \mathbb{K}$. In tal caso, λ **è l'autovalore di** A **relativo all'autovettore** v. Lo **spettro di** A si definisce come l'insieme degli autovalori di A e si indica con $Sp(A)$.

Osservazione 10.0.7. Se $f = Id_V$, allora ogni vettore $v \in V\ \{0\}$ è un autovettore di f, con autovalore corrispondente $\lambda = 1$; infatti si ha $f(v) = v = 1 \cdot\ v$ per ogni $v \in V$.

Viceversa, se f è un operatore tale che per ogni $v \in V \setminus \{0\}$ si ha che v è autovettore di f con autovalore 1, allora $f = Id_V$; la verifica è lasciata per esercizio.

Osservazione 10.0.8. Notiamo esplicitamente che un autovalore può essere nullo, e si ha

$$0 \in Sp(f) \iff \ker(f) \neq \{0\}.$$

La verifica è lasciata per esercizio.

La seguente proposizione è una riformulazione del Lemma 10.0.6

Proposizione 10.0.8. *Sia V uno spazio vettoriale sul campo $\mathbb{K}$ e sia $f : V \to V$ un operatore lineare. Allora f è diagonalizzabile se e solo se esiste una base $\mathcal{B}$ di V composta da autovettori.*

La seguente Proposizione illustra le principali proprietà degli autovettori ed autovalori.

Proposizione 10.0.9. *Sia V uno spazio vettoriale su un campo $\mathbb{K}$ e sia f un operatore. Allora valgono le seguenti proprietà.*

1. *Se $v \in V$ è un autovettore di f, allora l'autovalore corrispondente a v è unico.*

2. *Siano $v_1, \ldots, v_m \in V$ autovettori di f con relativi autovalori $\lambda_1, \ldots, \lambda_m \in \mathbb{K}$, rispettivamente. Se*
$$\lambda_i \neq \lambda_j,$$
per ogni $i \neq j$, $i, j \in \{1, \ldots, m\}$, allora $v_1, \ldots, v_m$ sono linearmente indipendenti.

3. *Sia $\lambda \in Sp(f)$ un autovalore di f. Allora l'insieme*
$$V_\lambda(f) := \{v \in V \mid v \text{ autovettore di } f \text{ con autovalore } \lambda\} \cup \{0\} = \ker(f - \lambda \cdot Id_V),$$
è un sottospazio vettoriale di V , chiamato l' **autospazio di** f **relativo all'autovalore** λ.

   In seguito, se non ci sarà pericolo di confusione, $V_\lambda(f)$ verrà denotato con V_λ.

Dimostrazione. (1) Siano λ e μ due autovalori

$$\lambda \neq \mu, \quad \lambda,\ \mu \in Sp(f),$$

relativi all'autovettore v. Allora vale

$$\lambda\, v = f(v) = \mu\, v,$$

da cui

$$0 = \lambda\, v - \mu\, v = (\lambda - \mu)\, v.$$

Siccome v è un autovettore, $v \neq 0$ per definizione, quindi

$$\lambda - \mu = 0,$$

cioè

$$\lambda = \mu.$$

(2) Procediamo per induzione su m. Se $m = 1$, allora v_1 è linearmente indipendente. Infatti per definizione di autovettore $v_1 \neq 0$.

Supponiamo ora che l'enunciato valga per $m-1$ autovettori relativi ad autovalori distinti. Consideriamo una combinazione lineare nulla

$$c_1v_1 + c_2v_2 + \cdots + c_mv_m = 0. \tag{10.0.2}$$

Allora si ha

$$f(c_1v_1 + c_2v_2 + \cdots + c_mv_m) = f(0) = 0; \tag{10.0.3}$$

d'altra parte per la linearità di f ed essendo i v_i autovettori possiamo scrivere

$$f(c_1v_1 + c_2v_2 + \cdots + c_mv_m) = c_1f(v_1) + c_2f(v_2) + \cdots + c_mf(v_m) = c_1\lambda_1v_1 + c_2\lambda_2v_2 + \cdots + c_m\lambda_mv_m.$$

Dalla (10.0.3) otteniamo

$$c_1\lambda_1v_1 + c_2\lambda_2v_2 + \cdots + c_m\lambda_mv_m = 0. \tag{10.0.4}$$

Ora possiamo moltiplicare entrambi i membri della (10.0.2) per λ_m e otteniamo

$$c_1\lambda_mv_1 + c_2\lambda_mv_2 + \cdots + c_m\lambda_mv_m = 0. \tag{10.0.5}$$

Infine, sottraendo la (10.0.5) dalla (10.0.4) otteniamo

$$c_1(\lambda_1 - \lambda_m)v_1 + c_2(\lambda_2 - \lambda_m)v_2 + \cdots + c_{m-1}(\lambda_{m-1} - \lambda_m)v_{m-1} = 0. \tag{10.0.6}$$

Essendo i vettori $v_1, \ldots, v_{m-1}$ autovettori relativi a $m-1$ autovalori distinti, per ipotesi induttiva essi sono linearmente indipendenti. Quindi nella combinazione lineare nulla (10.0.6) tutti i coefficienti

$$c_i(\lambda_i - \lambda_m) = 0$$

sono nulli. Ma osserviamo che per ogni $i \neq m$ si ha $\lambda_i \neq \lambda_m$ per ipotesi, quindi necessariamente

$$c_1 = 0, \ldots, c_{m-1} = 0.$$

Nella (10.0.2) rimane quindi solo il termine

$$c_mv_m = 0.$$

Siccome v_m è un autovettore, è diverso dal vettore nullo, quindi concludiamo che vale anche

$$c_m = 0.$$

(3) Osserviamo che possiamo scrivere

$$V_\lambda(f) = \{v \in V \mid f(v) = v\},$$

perchè $f(0) = 0 = \lambda \cdot 0$, quindi il vettore nulle soddisfa la condizione.

Inoltre, un vettore $v \in V$ soddisfa

$$f(v) = \lambda v$$

se e solo se l'operatore $f - \lambda\, Id_V$ è annullato da v:

$$f(v) = \lambda v \iff f(v) - \lambda v = (f - \lambda\, Id_V)(v) = 0,$$

quindi $v \in V_\lambda(f)$ se e solo se $v \in \ker(f - \lambda\, Id_V)$. In altre parole si ha

$$V_\lambda(f) = \ker(f - \lambda\, Id_V), \tag{10.0.7}$$

che risulta un sottospazio vettoriale perché nucleo di un operatore lineare. □

Corollario 10.0.10. *Sia V uno spazio vettoriale di dimensione finita n su un campo $\mathbb{K}$ e sia $f : V \to V$ un operatore. Allora f ha al più n autovalori distinti.*

Dimostrazione. Siano $\lambda_1, \ldots .\lambda_m \in Sp(f)$ autovalori distinti di f. Siano $v_1, \ldots, v_m$ degli autovettori relativi. Per il punto (2) della precedente Proposizione, $v_1, \ldots, v_m$ sono linearmente indipendenti, quindi

$$m \leq \dim(V) = n.$$

□

10.1 Criteri di diagonalizzabilità

Enunciamo i seguenti risultati senza dimostrazioni, per le quali si rimanda agli appunti delle lezioni oppure ai libri di testo.

Proposizione 10.1.1. *Sia V uno spazio vettoriale di dimensione finita n su un campo $\mathbb{K}$ e sia $f : V \to V$ un operatore. Se $Sp(f) = \{\lambda_1, \ldots, \lambda_k\}$, si ha*

$$\dim V_{\lambda_1}(f) + \dim V_{\lambda_2}(f) + \cdots + \dim V_{\lambda_k}(f) \leq n.$$

Vediamo ora un importante risultato, che ci permetterà di caratterizzare gli operatori diagonalizzabili.

Theorem 10.1.2. Primo criterio di diagonalizzabilità *Sia V uno spazio vettoriale di dimensione finita n su un campo $\mathbb{K}$. Sia $Sp(f) = \{\lambda_1, \ldots, \lambda_k\}$.*

Allora f è digonalizzabile se e solo se vale

$$\dim V_{\lambda_1}(f) + \dim V_{\lambda_2}(f) + \cdots + \dim V_{\lambda_k}(f) = n.$$

10.2 Autovalori come radici del polinomio caratteristico

Vediamo ora come trovare effettivamente gli autovalori di un operatore.

Proposizione 10.2.1. *Sia $\lambda \in \mathbb{K}$ uno scalare. Allora λ è un autovaler di un operatore f se e solose*

$$\lambda \in Sp(f) \iff \ker(f - \lambda\, Id_V) \neq \{0\} \iff \det(f - \lambda\, Id_V) = 0.$$

Osservazione 10.2.1. Osserviamo che la funzione $t \to \det(f - t\, Id_V)$, $\mathbb{K} \to \mathbb{K}[t]$ è una funzione polinomiale; in altre parole, se consideriamo t come una variabile, e calcoliamo $\det(f - t\, Id_V$ formalmente con uno dei metodi visti, otteniamo un polinomio di grado n a coefficienti in $\mathbb{K}$ nella indeterminata t.

Definizione 10.2.2. Sia V uno spazio vettoriale di dimensione finita su un campo $\mathbb{K}$ e sia f un operatore su V. Il **polinomio caratteristico di** f è il polinomio

$$\det(f - t\, Id_V)$$

e si indica con

$$p_f(t).$$

Analogamente, se $A \in M_n(\mathbb{K})$ è una matrice quadrata, il **polinomio caratteristico di** A è il polinomio definito da

$$p_A(t) := \det(A - t\mathbb{I}_n).$$

Osservazione 10.2.2. Se $\dim V = n$, allora $p_f(t)$ è un polinomio di grado n; inoltre, il coefficiente di t^n è $(-1)^n$ e il termine noto è $\det(f)$:

$$p_f(t) = (-1)^n t^n + \cdots + \det(f).$$

Esercizio 10.2.3. Si dimostri che se $A \in M_n(\mathbb{K})$ è una matrice quadrata, allora vale

$$p_A(t) = p_{L_A}(t).$$

Come Corollario della Proposizione 10.2.1 e con l'introduzione del polinomio caratteristico abbiamo il seguente:

Corollario 10.2.4. *Si ha*

$$\lambda \in Sp(f) \iff p_f(\lambda) = 0,$$

cioè **gli autovalori di f coincidono con tutte e sole le radici del polinomio caratteristico di f.**

Una volta trovati gli (eventuali) autovalori di un operatore, vediamo ora uno strumento per determinare la dimensione di ciascun autospazio associato a un autovalore.

Proposizione 10.2.5. *Sia $\lambda \in Sp(f)$. Allora si ha*

$$\dim V_\lambda(f) = \dim V - \mathrm{rg}(f - \lambda\, Id_V).$$

Dimostrazione. Essendo $V_\lambda(f) = \ker(f - \lambda\, Id_V)$ per la relazione (10.0.7), l'enunciato segue dal Teorema di Dimensione per applicazioni lineari. □

Osservazione 10.2.3. Ricordiamo che per determinare il numero

$$\operatorname{rg}(f - \lambda\, Id_V)$$

é sufficiente fissare una base $\mathcal{B}$ qualsiasi di V e calcolare il rango della matrice associata:

$$\operatorname{rg}(f - \lambda\, Id_V) = \operatorname{rg} M^{\mathcal{B}}_{\mathcal{B}}(f - \lambda\, Id_V) = \operatorname{rg}\left(M^{\mathcal{B}}_{\mathcal{B}}(f) - \lambda \mathbb{I}_m\right).$$

Concludiamo il capitolo con un altro criterio per stabilire se un operatore è diagonalizzabile oppure no. Abbiamo prima bisogna di ulteriori definizioni.

Definizione 10.2.6. Sia V uno spazio vettoriale di dimensione n e sia f un operatore su V.

Supponiamo $\lambda \in Sp(f)$. La **molteplicità algebrica di** λ è la molteplicità di λ come radice del polinomio caratteristico $p_f(t)$, e si denota con

$$m_a(\lambda).$$

La **molteplicità geometrica di** λ è la dimensione dell'autospazio $V_\lambda(f)$ relativo a λ, e si denota con

$$m_g(\lambda) = \dim V_\lambda(f).$$

Proposizione 10.2.7. *Valgono sempre le seguenti diseguaglianze:*

$$1 \leq m_g(\lambda \leq m_a(\lambda).$$

Dimostrazione. Omessa (anche a lezione). □

Theorem 10.2.8. Secondo criterio di diagonalizzabilitá *Sia V uno spazio vettoriale di dimensione finita n su un campo $\mathbb{K}$ e sia f un operatore di V.*

Allora f è diagonalizzabile se e solo valgono entrambe le seguenti condizioni:

1. *il polinomio caratteristico $p_f(t)$ si decompone come prodotto di polinomi di grado 1 in t;*
2. *per ogni autovalore $\lambda \in Sp(f)$ si ha*
$$m_g(\lambda) = m_a(\lambda).$$

Vediamo ora una conseguenza immediata del Teorema 10.2.8.

Corollario 10.2.9. *Sia $\dim V = n$. Se f ha n autovalori distinti, allora f è diagonalizzabile.*

Capitolo 11

Prodotti scalari

11.1 Prodotto scalare standard in $\mathbb{R}^n$

Definizione 11.1.1. Il **prodotto scalare standard** su $\mathbb{R}^n$ è la seguente funzione:

$$\cdot : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \quad v = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}, w = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}, \quad v \cdot w := a_1 b_1 + a_2 b_2 + \cdots + a_n b_n.$$

Per ogni vettore $v \in \mathbb{R}^n$ possiamo definire la sua **norma** $\|v\|$ come

$$\|v\| := \sqrt{v \cdot v} = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}.$$

La norma di v si può pensare come la *lunghezza* di v, cioè la distanza dall'origine del punto $(a_1, a_2, \ldots, a_n)$.

Osservazione 11.1.1. Il prodotto scalare standard soddisfa le seguenti proprietà, per ogni scelta di $v, v_1, v_2, w, w_1, w_2 \in \mathbb{R}^n$:

1. **bilinearità**:
   - $(v_1 + v_2) \cdot w = v_1 \cdot w + v_2 \cdot w$,
   - $v \cdot (w_1 + w_2) = v \cdot w_1 + v \cdot w_2$,
   - $(cv) \cdot w = c(v \cdot w) = v \cdot (cw)$ per ogni $c \in \mathbb{R}$
2. **simmetria** $v \cdot w = w \cdot v$,
3. **il prodotto scalare è definito positivo**: $v \cdot v \geq 0$ per ogni $v \in \mathbb{R}^n$, e si ha $v \cdot v = 0 \iff v = 0$.

11.2 Prodotti scalari in spazi vettoriali reali

La nozione di prodotto scalare si può generalizzare a un qualunque spazio vettoriale **reale**; chiameremo prodotto scalare una qualunque funzione che associ a una coppia di vettori un numero reale, e che verifichi le proprietà di bilinearità, simmetria e che sia definita positiva. Più precisamente:

Definizione 11.2.1. Sia V uno spazio vettoriale sul campo $\mathbb{R}$. Una funzione

$$g : V \times V \to \mathbb{R}$$

si dice **prodotto scalare** se soddisfa le seguenti proprietà, per ogni scelta di $v, v_1, v_2, w, w_1, w_2 \in V$:

1. **bilinearità**:
    - $g(v_1 + v_2, w) = g(v_1, w) + g(v_2, w)$,
    - $g(v, w_1 + w_2) = g(v, w_1) + g(v, w_2)$,
    - $g(cv, w) = cg(v, w) = g(v, cw)$ per ogni $c \in \mathbb{R}$
2. **simmetria** $g(v, w) = g(w, v)$,
3. **il prodotto scalare è definito positivo**: $g(v, v) \geq 0$ per ogni $v \in V$, e si ha $g(v, v) = 0 \iff v = 0$.

Esempi 11.2.2. Vediamo ora degli esempi molto diversi dal prodotto scalare standard:

1. sia $V = \mathcal{C}^0([a, b], \mathbb{R})$ lo spazio vettoriale delle funzioni continue su un intervallo chiuso e limitato $[a, b]$, e consideriamo la funzione

    $$g : \mathcal{C}^0([a, b], \mathbb{R}) \times \mathcal{C}^0([a, b], \mathbb{R}) \to \mathbb{R}, \quad g(f, h) := \int_a^b f(x)g(x)dx.$$

    È semplice verificare che g è bilineare e simmetrica; inoltre si ha

    $$g(f, f) = \int_a^b f^2(x)dx \geq 0,$$

    e si ha $g(f, f) = 0$ se e solo se $f = 0$. Quindi g è un prodotto scalare su $\mathcal{C}^0([a, b], \mathbb{R})$.
2. Sia $V = M_2(\mathbb{R})$ lo spazio delle matrici quadrate 2×2; la funzione

    $$g : M_2(\mathbb{R}) \times M_2(\mathbb{R}) \to \mathbb{R}, \quad g(A, B) = a_{11}b_{11} + a_{12}b_{12} + a_{21}b_{21} + a_{22}b_{22}$$

    soddisfa le richieste di un prodotto scalare (esercizio).
    Osservaiamo che è possibile esprimere tale funzione in un'altra forma.
    Introduciamo prima la nozione di **traccia** di una matrice quadrata: se $M \in M_n(\mathbb{K})$, poniamo

    $$\mathrm{Tr}(M) := m_{11} + m_{22} + \cdots + m_{nn}.$$

    Si può verificare facilmente che risulta:

    $$g(A, B) = \mathrm{tr}(A \cdot^t B).$$

    In generale, la funzione

    $$g : M_{m,n}(\mathbb{R}) \times M_{m,n}(\mathbb{R}) \to \mathbb{R}, \quad g(N, P) := \mathrm{Tr}(N \cdot^t P)$$

    è un prodotto scalare sullo spazio delle matrici $m \times n$.

Definizione 11.2.3. Sia g un prodotto scalare su V. La coppia (V, g) si chiama **spazio vettoriale euclideo**.

Per ogni vettore $v \in V$, la **norma** di v è così definita:

$$\|v\| := \sqrt{g(v, v)}.$$

Esempi 11.2.4. 1. Nello spazio $\mathcal{C}^0([a, b], \mathbb{R})$, la norma di una funzione f è data da

$$\|f\| = \sqrt{\int_a^b f^2(x)dx}.$$

2. Nello spazio $M_{m,n}(\mathbb{R})$ la norma di una matrice M è data da

$$\|M\| = \sqrt{\mathrm{Tr}(M \cdot^t M)}.$$

Proposizione 11.2.5. Proprietà della norma *La norma verifica le seguenti:*

- $\forall v \in V$, *si ha* $\|v\| \geq 0$ *e vale* $\|v\| = 0 \iff v = 0$.
- $\|cv\| = |c|\|v\|$, *per ogni* $c \in \mathbb{R}$, *dove* $|c|$ *indica il valore assoluto;*
- **diseguaglianza triangolare***:*
$$\|v + w\| \leq \|v\| + \|w\|,$$
per ogni $v, w \in V$*;*
-
$$|\|v\| - \|w\|| \leq \|v + w\|$$
per ogni $v, w \in V$.

11.3 La diseguaglianza di Cauchy - Schwarz e angolo convesso tra vettori non nulli

Teorema 11.3.1. Diseguaglianza di Cauchy - Schwarz*: Sia* (V, g) *uno spazio vettoriale euclideo. Allora per ogni coppia di vettori* $v, w \in V$ *vale*

$$|g(v, w)| \leq \|v\| \, \|w\|.$$

Inoltre, vale l'uguaglianza se e solo se v *e* w *sono linearmente dipendenti.*

Dimostrazione. Supponiamo dapprima che sia

$$w = 0$$

il vettore nullo. Allora per bilinearità si ha $g(v, 0) = g(v, 0 \cdot v) = 0 \cdot g(v, v) = 0$ e $\|v\| \cdot 0 = 0$, quindi la disuguaglianza è soddisfatta.

Supponiamo ora

$$w \neq 0,$$

e consideriamo il vettore

$$v + tw, \ t \in \mathbb{R}.$$

Per le proprietà della norma si ha

$$\|v + tw\| \geq 0,$$

e vale $\|v + tw\| = 0$ se e solo se $v + tw = 0$, cioè se e solo se v e w sono linearmente dipendenti.

D'altra parte per definizione si ha

$$\|v + tw\|^2 = g(v + tw, v + tw) = g(v, v) + 2tg(v, w) + g(w, w)$$

per bilinearità e simmetria. Possiamo riscrivere l'espressione nella forma

$$\|v\|^2 + 2tg(v, w) + \|w\|^2 \geq 0,$$

e questo vale per ogni $t \in \mathbb{R}$. Il grafico della funzione

$$\varphi(t) := \|v\|^2 + 2tg(v, w) + \|w\|^2$$

è una parabola rivolta verso l'alto, e i suoi valori non sono mai negativi se e solo se il discriminante del polinomio di secondo grado è minore o uguale a zero; quindi

$$4g(v, w)^2 - 4\|v\|^2\|w\|^2 \leq 0,$$

da cui

$$g(v, w)^2 \leq \|v\|^2\|w\|^2,$$

e passando alla radici quadrate

$$|g(v, w)| \leq \|v\|\|w\|.$$

Infine, osserviamo che v eW sono linearmente dipendenti se e solo se $v + cw = 0$ per un opportuno $c \in \mathbb{R}$; ciò equivale ad avere delle uguaglianze in tutte le espressioni considerate.

□

Dal Teorema appena visto deduciamo che se $v, w \in V \setminus \{0\}$ sono entrambi non nulli si ha

$$\left| \frac{g(v, w)}{\|v\|\|w\|} \right| \leq 1.$$

Questo permette di dare la seguente definizione:

Definizione 11.3.2. Siano $v, w \in V \setminus \{0\}$ vettori nulli; l' **angolo convesso tra** v **e** w si definisce come l'angolo $\theta \in [0, \pi]$ tale che

$$\cos\theta = \frac{g(v, w)}{\|v\|\|w\|}.$$

Inoltre, v ew si dicono **ortogonali** se

$$g(v, w) = 0.$$

In simboli scriveremo

$$v \perp w.$$

Dato un sottoinsieme $S \ subsetV$, il **sottoinsieme ortogonale** $S^\perp$ è così definito:

$$S^\perp := \{w \in V | w \perp v, \forall v \in S\}.$$

11.4 Basi ortogonali e ortonormali

Definizione 11.4.1. Sia (V, g) uno spazio vettoriale euclideo, con V di dimensione finita. Una base **ortogonale** di V è una base $\mathcal{B} = \{v_1, \ldots, v_n\}$ tale che

$$v_i \perp v_j, \text{ per ogni } i \neq j.$$

Una **base ortonormale** di V è una base ortogonale $\mathcal{B} = \{v_1, \ldots, v_n\}$ tale che

$$v_i \perp v_j, \text{ per ogni } i \neq j \text{ e } \|v_i\| = 1 \text{ per ogni } i,$$

cioè una base ortogonale formata da **versori**.

Esempi 11.4.2.
1. Per ogni $n \geq 1$, la base canonica $\mathcal{E}$ di $\mathbb{R}^n$ è ortonormale rispetto al prodotto scalare standard.
2. La base $\left\{\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \end{pmatrix}\right\}$ è ortogonale ma non ortonormale; infatti

$$\left\|\begin{pmatrix} 1 \\ 2 \end{pmatrix}\right\| = \left\|\begin{pmatrix} 2 \\ -1 \end{pmatrix}\right\| = \sqrt{5}.$$

Dividendo i due vettori per le loro norme troviamo, invece, una base ortonormale: $\left\{\begin{pmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix}, \begin{pmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{pmatrix}\right.$
.

11.5 Ortonormalizzazione di Gram - Schmidt

Teorema 11.5.1. *Sia (V, g) uno spazio vettoriale euclideo. Siano $w_1, \ldots, w_k \in V$ vettori linearmente indipendenti.*

Allora esiste una base $\{v_1, \ldots, v_k\}$ di $\mathrm{Span}(w_1, \ldots, w_k)$ ortonormale per il prodotto scalare g.

Dimostrazione. Procediamo per induzione su k.

Se $k = 1$, allora definiamo v_1 nel modo seguente:

$$v_1 := \frac{w_1}{\|w_1\|}.$$

$(k - 1 \Rightarrow k)$ Per ipotesi induttiva esistono

$$v_1, \ldots, v_{k-1} \in \mathrm{Span}(w_1, \ldots, w_{k-1})$$

che formano una base ortonormale di $\mathrm{Span}(w_1, \ldots, w_{k-1})$. Sia ora

$$\widetilde{w_k} := g(w_k, v_1)v_1 + \cdots + g(w_k, v_{k-1})v_{k-1},$$

che rappresenta la **proiezione ortogonale di w_k sul sottospazio** $\mathrm{Span}(v_1, \ldots, v_{k-1}) = \mathrm{Span}(w_1, \ldots, w_{k-1})$. Poniamo poi

$$\widetilde{v_k} := w_k - \widetilde{w_k}.$$

Osserviamo che

$$\widetilde{v_k} \neq 0,$$

perché $\widetilde{w_k} \in \mathrm{Span}(v_1, \ldots, v_{k-1})$ e $w_k \notin \mathrm{Span}(v_1, \ldots, v_{k-1})$. Inoltre, osserviamo che il vettore $\widetilde{v_k}$ è ortogonale a ciascun vettor $v_1, \ldots, v_{k-1}$; infatti:

$$\begin{aligned} g(\widetilde{v_k}, v_i) &= g(w_k, v_i) - g(\widetilde{w_k}, v_i) = \\ &= g(w_k, v_i) - g(g(w_k, v_1)v_1 + \cdots + g(w_k, v_{k-1})v_{k-1}, v_i) = \\ &= g(w_k, v_i) - g(w_k, v_i) = 0, \qquad \forall i = 1, \ldots, k-1. \end{aligned}$$

Quindi ponendo

$$v_k := \frac{\widetilde{v_k}}{\|\widetilde{v_k}\|},$$

i vettori $v_1, \ldots, v_k$ risultano a due a due ortogonali e di norma uno, quindi sono linearmente indipendenti in $\mathrm{Span}(w_1, \ldots, w_k)$ e ne formano una base ortonormale.

□

Corollario 11.5.2. *Ogni spazio vettoriale euclideo di dimensione finita ammette una base ortonormale.*

Dimostrazione. Sia $\{w_1, \ldots, w_m\}$ una base di V. Applicando il procedimento di Gram - Schmidt ai vettori $w_1, \ldots, w_n$ otteniamo una base ortonormale di $rmSpan(w_1, \ldots, w_n) = V$. □

Capitolo 12

Operatori simmetrici

12.1 Definizioni e prime proprietà

Definizione 12.1.1. Sia (V, g) uno spazio vettoriale euclideo. Un operatore $f : V \to V$ è detto **autoaggiunto** o **simmetrico** rispetto al prodotto scalare g se vale la seguente uguaglianza per ogni coppia di vettori $v, w \in V$:

$$g(f(v), w) = g(v, f(w)).$$

Gli operatori autoaggiunti vengono chiamati anche simmetrici, poiché sono rappresentati, rispetto a basi ortonormali, da matrici simmetriche. Per dimostrare tale fatto abbiamo bisogno del seguente Lemma.

Lemma 12.1.2. *Sia (V, g) uno spazio vettoriale euclideo e sia $\mathcal{C} = \{u_1, \dots, u_n\}$ una base ortonormale rispetto a $\mathcal{C}$. Allora per ogni coppia di vettori $v, w \in V$, se indichiamo con*

$$\mathbf{x} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}, \qquad \mathbf{y} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

le colonne delle coordinate di v e w nella base $\mathcal{C}$, si ha

$$g(v, w) = {}^t\mathbf{x} \cdot \mathbf{y}.$$

Dimostrazione. Esprimendo i vettori v e w con le coordinate

$$v = a_1u_1 + \dots a_nu_n, \quad w = b_1u_1 + \dots + b_nu_n,$$

e usando la bilinearità e la simmetria di g, otteniamo:

$$\begin{aligned} g(v, w) &= g(a_1u_1 + \dots a_nu_n, b_1u_1 + \dots + b_nu_n) = \\ &= a_1g(u_1, b_1u_1 + \dots + b_nu_n) + \dots + a_ng(u_n, b_1u_1 + \dots + b_nu_n) = \\ &= a_1b_1g(u_1, u_1) + a_1b_2g(u_1, u_2) + \dots + a_nb_ng(u_n, u_n). \end{aligned}$$

Sfruttando ora il fatto che $\mathcal{C}$ è una base ortonormale, e quindi si ha

$$g(u_i, u_i) = 1, \quad g(u_i, u_j) = 0 \; \forall i \neq j,$$

otteniamo

$$g(v, w) = a_1 b_1 + \cdots + a_n b_n,$$

che è uguale a ${}^t\mathbf{x} \cdot \mathbf{y}$.

$\square$

Proposizione 12.1.3. *Sia (V, g) uno spazio vettoriale euclideo di dimensione finita. Sia $f : V \to V$ e sia $\mathcal{C}$ una base ortonormale di (V, g). Allora, f è autoaggiunto rispetto al prodotto scalare g se e solo se la matrice $M^{\mathcal{C}}_{\mathcal{C}}(f)$ che rappresenta f rispetto alla base $\mathcal{B}$ è simmetrica.*

Dimostrazione. Poniamo $A = M^{\mathcal{C}}_{\mathcal{C}}(f)$. Per il Lemma 12.1.2, per ogni coppia $v, w \in V$, se $\mathbf{x}$ e $\mathbf{y}$indicano le colonne delle coordinate dei due vettori nella base $\mathcal{C}$, si ha

$$g(f(v), w) = {}^t(A \cdot \mathbf{x}) \cdot \cdot \mathbf{y} = {}^t\mathbf{x} \cdot {}^tA \cdot \mathbf{y}.$$

Essendo f autoaggiunto, vale anche

$$g(f(v), w) = g(v, f(w)) = {}^t\mathbf{x} \cdot A \cdot \mathbf{y}.$$

Osserviamo ora che se consideriamo i vettori $u_1, \ldots, u_n$ della base $\mathcal{C}$, i loro vettori delle coordinate sono uguali a:

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Ne deduciamo che

$$\begin{aligned} g(f(u_i), u_j) &= ({}^tA)_{ij} = a_{ji} = \\ &= g(u_i, f(u_j)) = (A)_{ij} = a_{ij}, \end{aligned}$$

quindi A è simmetrica.

$\square$

Lemma 12.1.4. 1. *Sia $A \in M_n(\mathbb{R})$ una matrice simmetrica. Allora il polinomio caratteristico $p_A(t) \in \mathbb{R}[t]$ ha n radici reali, non necessariamente distinte.*

2. *Sia (V, g) uno spazio vettoriale euclideo di dimensione n. Se f è simmetrico rispetto a g, allora $P_f(t) \in \mathbb{R}[t]$ ha n radici reali, non necessariamente distinte.*

Dimostrazione. 1. Possiamo considerare $p_A(t) \in \mathbb{R}[t]$ come polinomio a coefficienti complessi: $p_A(t) \in \mathbb{C}[t]$, poiché $\mathbb{R} \subset \mathbb{C}$. Per il Teorema Fondamentale dell'Algebra, $p_A(t)$ ha n radici complesse, contate con le rispettive molteplicità. Dimostriamo ora che ogni radice complessa λ di $p_A(t)$ appartiene ad $\mathbb{R}$.

Consideriamo A come una matrice a coefficienti complessi, e sia

$$L_A : \mathbb{C}^n \to \mathbb{C}^n, \quad L_A(v) = A \cdot v.$$

Siccome λ è una radice del polinomio caratteristico di A, λ è un autovalore di L_A. Sia $v \in \mathbb{C}^n$, $v \neq 0$, un autovettore di L_A corrispondente all'autovalore λ. Vale quindi la seguente uguaglianza:

$$A \cdot v = \lambda v.$$

Inoltre, siccome i coefficienti di A sono numeri reali, abbiamo le seguenti uguaglianze:

$$^t v \cdot A \cdot \overline{v} = {}^t v \cdot (A \cdot \overline{v}) = {}^t v \cdot (\overline{A \cdot v}) = {}^t v \cdot (\overline{\lambda v}) = {}^t v \cdot \overline{\lambda}\overline{v} = \overline{\lambda}\, {}^t v \cdot \overline{v},$$

dove $\overline{v}$ e $\overline{A \cdot v}$ sono i vettori di $\mathbb{C}^n$ ottenuti da v e $A \cdot v \in \mathbb{C}^n$ coniugando tutte le coordinate.
D' altro canto, siccome A è simmetrica, si ha

$$^t v \cdot A \cdot \overline{v} = ({}^t v \cdot A) \cdot \overline{v} = {}^t(A \cdot v) \cdot \overline{v} = {}^t(\lambda v) \cdot \overline{v} = \lambda\, {}^t v \cdot \overline{v}.$$

Quindi $\overline{\lambda}\, {}^t v \cdot \overline{v} = \lambda\, {}^t v \cdot \overline{v}$, da cui $(\overline{\lambda} - \lambda)\, {}^t v \cdot \overline{v} = 0$.
Osserviamo ora che se

$$v = \begin{pmatrix} a_1 + ib_1 \\ a_2 + ib_2 \\ \vdots \\ a_n + ib_n \end{pmatrix}, \qquad \overline{v} = \begin{pmatrix} a_1 - ib_1 \\ a_2 - ib_2 \\ \vdots \\ a_n - ib_n \end{pmatrix},$$

si ha ${}^t v \cdot \overline{v} = a_1^2 + b_1^2 + a_2^2 + b_2^2 + \cdots + a_n^2 + b_n^2$; essendo v un autovettore, almeno una sua compotente è non nulla, e quindi $a_1^2 + b_1^2 + a_2^2 + b_2^2 + \cdots + a_n^2 + b_n^2 > 0$. Ne segue che $\lambda = \overline{\lambda}$, cioè $\lambda \in \mathbb{R}$.

2. Sia $\mathcal{B}$ una base ortonormale di (V, g). Per la Proposizione 12.1.3, la matrice $M_{\mathcal{B}}(f)$ è simmetrica. Il risultato segue ora dal punto precedente e dal fatto che

$$p_f(t) = p_{M_{\mathcal{B}}^{\mathcal{B}}(f)}(t).$$

□

12.2 Teorema spettrale per operatori simmetrici

Teorema 12.2.1. Teorema Spettrale per Operatori Simmetrici *Sia (V, g) uno spazio vettoriale euclideo di dimensione finita n. Sia $f : V \to V$ un operatore simmetrico rispetto al prodotto scalare g. Allora esiste una base ortonormale $\mathcal{B}$ di (V, g) che diagonalizza f.*

Dimostrazione. Procediamo per induzione sulla dimensione n di V . Se $n = 1$, allora ogni base $\{v_1\}$ di V diagonalizza f, quindi in questo caso l' enunciato segue ponendo $\mathcal{B} = \left\{\frac{v_1}{\|v_1\|}\right\}$.

Supponiamo ora che l'enunciato sia vero per spazi vettoriali di dimensione $n - 1$. Siccome f è simmetrico, per il Lemma 12.1.4 esiste un autovalore $\lambda \in \mathbb{R}$ di f. Sia $v \in V$, $v \neq 0$ un autovettore di f con autovalore corrispondente λ. Sia

$$U := \{v\}^{\perp} = \{u \in V \mid g(u, v) = 0\}.$$

il sottospazio vettoriale di V ortogonale a v. Osserviamo che $f(U) \subseteq U$; infatti:

$$g(f(u), v) = g(u, f(v)) = g(u, \lambda v) = \lambda g(u, v),$$

quindi se $g(u, v) = 0$, si ha anche $g(f(u), v) = 0$.

Quindi la restrizione di f ad U definisce un operatore di U, che chiameremo

$$f_U : U \to U$$

Consideriamo ora la restrizione $g_{|U\times U}$ del prodotto scalare g ad $U \times U$. Si verifica facilmente che $g_{|U\times U}$ è ancora un prodotto scalare. Inoltre, completando il versore $\frac{v}{\|v\|}$ ad una base ortonormale $\{\frac{v}{\|v\|}, u_1, \ldots, u_{n-1}\}$ di V, si può verificare che risulta

$$U = \mathrm{Span}(u_1, \ldots, u_{n-1}),$$

quindi

$$\dim(U) = n - 1.$$

Infine, f_U risulta un operatore simmetrico rispetto a $g_{|U\times U}$. Quindi, per ipotesi induttiva, esiste una base $\{v_1, \ldots, v_{n-1}\}$ di U, ortonormale per il prodotto scalare $g_{|U\times U}$, che diagonalizza f_U. Il teorema segue ora definendo

$$v_n := \frac{v}{\|v\|},$$

e $\mathcal{B} := \{v_1, \ldots, v_n\}$.

□

Corollario 12.2.2. *Sia $A \in M_n(\mathbb{R})$ una matrice simmetrica. Allora esiste una matrice ortogonale C tale che $C^{-1} \cdot A \cdot C$ è diagonale, quindi A è diagonalizzabile.*

Dimostrazione. Sia $\mathcal{E}$ la base canonica di $\mathbb{R}^n$. Ricordiamo che $\mathcal{E}$ è ortonormale per il prodotto scalare standard. Siccome $A = M_{\mathcal{E}}(L_A)$, per la Proposizione 12.1.3, L_A è autoaggiunto. Quindi, per il Teorema Spettrale, esiste una base $\mathcal{C}$ di $\mathbb{R}^n$, ortonormale per il prodotto scalare standard, che diagonalizza L_A, cioè tale che $(M^{\mathcal{C}}_{\mathcal{E}}(\mathrm{Id}_{\mathbb{R}^n}))^{-1} \cdot A \cdot M^{\mathcal{E}}_{\mathcal{C}}(\mathrm{Id}_{\mathbb{R}^n})$ è diagonale. Poiché le basi $\mathcal{E}$ e $\mathcal{C}$ sono ortonormali, $M^{\mathcal{E}}_{\mathcal{C}}(\mathrm{Id}_{\mathbb{R}^n})$ è una matrice ortogonale. □

12.3 Procedimento computazionale per la diagonalizzazione di un operatore simmetrico

Vediamo ora un procedimento per trovare, dato un operatore simmetrico f o una matrice simmetrica A, una base ortonormale di (V, g) che diagonalizza f, rispettivamente A.

Riportiamo prima un risultato che sarà utile a tale proposito.

Proposizione 12.3.1. *Sia (V, g) uno spazio vettoriale euclideo. Sia $f : V \to V$ simmetrico rispetto a g. Se $\lambda, \mu \in \mathrm{Sp}(f)$ sono due autovalori distinti, allora gli autospazi corrispondenti V_λ e V_μ sono ortogonali.*

Dimostrazione. Per ogni $v \in V_\lambda$ e per ogni $w \in V_\mu$, valgono le seguenti uguaglianze:

$$\lambda g(v, w) = g(f(v), w) = g(v, f(w)) = \mu g(v, w).$$

Quindi $(\lambda - \mu)g(v, w) = 0$, e siccome $\lambda \neq \mu$, segue che $g(v, w) = 0$, cioè ogni vettore di V_λ è ortogonale ad ogni vettore di V_μ. □

Sia ora (V, g) uno spazio vettoriale euclideo di dimensione finita n, e sia $f : V \to V$ un operatore simmetrico rispetto a g. Per trovare una base ortonormale di (V, g) che diagonalizza f procediamo come segue:

1. Scegliamo una base ortonormale $\mathcal{C}$ di V e determiniamo la matrice $A := M^{\mathcal{C}}_{\mathcal{C}}(f)$ Per la Proposizione 12.1.3, A è simmetrica.

2. Determiniamo il polinomio caratteristico $p_f(t) = \det(A - t\mathbb{I}_n) \in \mathbb{R}[t]$, e le sue radici distinte $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$. Per il Lemma 1 vale la seguente uguaglianza:

$$p_f(t) = (\lambda_1 - t)^{m_a(\lambda_1)} \cdot (\lambda_2 - t)^{m_a(\lambda_2)} \cdot \cdots \cdot (\lambda_k - t)^{m_a(\lambda_k)},$$

   dove $m_a(\lambda_i)$ è la molteplicità algebrica di λ_i, $\forall i = 1, \ldots, k$.

3. Per ogni $i = 1, \ldots, k$, si determina una base $\{w_{i,1}, \ldots, w_{i,m_g(\lambda_i)}\}$ di V_{λ_i}, dove $m_g(\lambda_i)$ è la molteplicità geometrica di λ_i. Osserviamo che, per il teorema spettrale f è diagonalizzabile, quindi per il Secondo criterio di diagonalizzazione si ha $m_g(\lambda_i) = m_a(\lambda_i)$. Osserviamo inoltre che $\{w_{i,1}, \ldots, w_{i,m_g(\lambda_i)}\}$ è data da una base dello spazio delle soluzioni del sistema lineare omogeneo

$$(A - \lambda_i\mathbb{I}_n) \cdot x = 0,$$

   poiché $V_{\lambda_i} = \ker(f - \lambda_i\mathbb{I}_n)$.

4. Per ogni $i = 1, \ldots, k$, applichiamo il procedimento di Gram-Schmidt ai vettori $\{w_{i,1}, \ldots, w_{i,m_g(\lambda_i)}\}$ e troviamo una base $\{v_{i,1}, \ldots, v_{i,m_g(\lambda_i)}\}$ di V_{λ_i} ortonormale rispetto al prodotto scalare g.

5. $\mathcal{B} := \{v_{1,1}, \ldots, v_{k,m_g(\lambda_k)}\}$ è una base ortonormale di (V, g) che diagonalizza f. Osserviamo che per ogni $i \neq j$, i vettori $v_{i,l}$ e $v_{j,m}$ sono ortogonali per la Proposizione 12.3.1.