

Note del corso

GEOMETRIA E ALGEBRA LINEARE

082747 - Ingegneria dell'Automazione, Elettrica, Elettronica ed Informatica

Davide Frapporti

Anno accademico 2025-2026

DIPARTIMENTO DI MATEMATICA, POLITECNICO DI MILANO
Email: `davide.frapporti@polimi.it`

© 2025 Davide Frapporti - Tutti i diritti sono riservati

Indice

0	Introduzione	1
0.1	Vettori geometrici nel piano e nello spazio	1
0.2	Rette e piani	4
1	Matrici e sistemi lineari	6
1.1	Matrici	7
1.1.1	Operazioni fra matrici	8
1.2	Sistemi lineari	12
1.2.1	Matrici a scala	14
1.2.2	Operazioni elementari sulle righe e metodo di eliminazione di Gauss	15
1.2.3	Teoremi di Rouché-Capelli e di struttura per le soluzioni di un sistema lineare	18
1.2.4	Matrice inversa	21
1.3	Determinante	24
1.3.1	Interpretazione geometrica del determinante	25
1.3.2	Determinante e operazioni elementari sulle righe	26
2	Spazi vettoriali	29
2.1	Spazi vettoriali	29
2.2	Sottospazi vettoriali	30
2.3	Combinazioni lineari	31
2.3.1	Generatori	32
2.3.2	Indipendenza lineare	32
2.4	Basi e dimensione	35
2.4.1	Un'osservazione sul rango	39
2.5	Equazioni cartesiane e parametriche	39
2.6	Intersezione e somma di sottospazi vettoriali	40
2.6.1	Formula di Grassmann	42
2.7	Sottospazi affini	44
2.7.1	Sottospazi affini di $\mathbb{R}^3$ e loro posizioni reciproche	47
3	Applicazioni lineari	52
3.1	Applicazioni lineari: definizione e proprietà base	53
3.1.1	L'insieme delle mappe lineari è uno spazio vettoriale	55
3.2	Suriettività ed iniettività delle mappe lineari	55
3.2.1	Suriettività	56
3.2.2	Iniettività	57

3.2.3	Biettività	57
3.3	Teorema nullità più rango e sue conseguenze	58
3.4	Come definire una mappa lineare?	60
3.4.1	Matrice rappresentativa	62
3.4.2	Matrice di cambio di base	64
4	Endomorfismi e diagonalizzazione	66
4.1	Autovettori ed autovalori	67
4.2	Polinomio caratteristico	68
4.3	Molteplicità	70
4.4	Sintesi del processo di diagonalizzazione	74
4.5	Complementi: matrice compagna	76
5	Spazi euclidei	77
5.1	Forme bilineari	77
5.2	Prodotti scalari	79
5.2.1	Distanza	82
5.2.2	Angolo tra vettori	82
5.3	Ortogonalità	83
5.3.1	Basi ortogonali e basi ortonormali	83
5.3.2	Ortonormalizzazione di Gram-Schmidt	85
5.3.3	Complemento ortogonale e proiezioni ortogonali	86
5.4	Complementi: Prodotto vettoriale	91
5.5	Sottospazi affini e distanze	92
5.6	Endomorfismi di spazi euclidei	93
5.6.1	Matrici ortogonali	93
5.6.2	Isometrie e trasformazioni affini	95
5.6.3	Endomorfismi simmetrici	96
5.7	Forme quadratiche	100
5.7.1	Segno di una forma quadratica	102
5.7.2	Segnatura e classificazione delle forme quadratiche	102
5.7.3	Criterio di Cartesio	106
5.7.4	Complementi: metodo dei minori di nord-ovest	106
6	Quadriche e loro classificazione	108
6.1	Ipersuperfici quadriche	108
6.1.1	Centro di una quadrica	110
6.2	Classificazione metrica ed affine	111
6.2.1	Quadriche a centro	112
6.2.2	Quadriche non a centro	114
6.2.3	Sintesi del processo di classificazione	118
6.3	Simmetrie di quadriche	118
6.3.1	Simmetrie assiali	118
6.3.2	Simmetrie di rotazione	120
6.4	Nomenclatura di coniche e quadriche in $\mathbb{R}^3$	121

Capitolo 0

Introduzione

In questo capitolo preliminare, inizieremo con una discussione informale su vettori liberi e vettori applicati in $\mathbb{R}^2$ e $\mathbb{R}^3$, che probabilmente il lettore ha già incontrato alle scuole superiori. Saremo approssimativi e non del tutto precisi: l'intenzione è quella di fornire il succo di alcuni concetti che saranno ampliati e generalizzati in queste note.

0.1 Vettori geometrici nel piano e nello spazio

Definizione 0.1. Un *vettore applicato* è un segmento con un verso dato, quindi è determinato dal un punto iniziale A e da un punto finale B . Sarà indicato con $\overrightarrow{AB}$.

Un vettore applicato $\overrightarrow{AB}$ descrive un *movimento assoluto* dal punto A al punto B . In fisica i vettori applicati sono largamente usati, per esempio per indicare una forza applicata ad un corpo, o la velocità di un oggetto, etc.

Osservazione. Un vettore applicato $\overrightarrow{AB}$ è determinato da 4 dati:

1. il punto di applicazione A ;
2. la direzione (della retta passante per A e B);
3. il verso (seguendo il quale ci muoviamo lungo la retta, da A a B);
4. il modulo (la lunghezza del segmento di estremi A e B).

Osservazione. Il vettore applicato $\overrightarrow{BA}$ ha la stessa direzione e lo stesso modulo di $\overrightarrow{AB}$, ma ha verso opposto.

Sia V_{app} l'insieme dei vettori applicati, e su V_{app} definiamo la seguente relazione: $\overrightarrow{AB}$ è in relazione con $\overrightarrow{CD}$ se i due vettori hanno la stessa direzione, lo stesso verso e lo stesso modulo.

Questa relazione gode delle proprietà riflessiva, simmetrica e transitiva ed è pertanto una relazione di equivalenza. La classe di equivalenza di $\overrightarrow{AB}$ è data da tutti i vettori che descrivono lo stesso spostamento, anche se sono applicati in altri punti dello spazio.

In sostanza, dato un vettore applicato $\overrightarrow{AB}$ ci stiamo “dimenticando” del punto di applicazione, e quindi stiamo considerando un *movimento relativo*: ci interessa la differenza tra punto di arrivo e punto di partenza. Questo è descritto da un vettore libero.

Definizione 0.2. Un *vettore libero* $\vec{v}$ è determinato da direzione, verso e modulo.

Ogni vettore applicato $\overrightarrow{AB}$ induce un vettore libero, prendendo la sua classe di equivalenza $[\overrightarrow{AB}]$. D'altra parte, ogni vettore libero è una classe di equivalenza di vettori applicati e quindi corrisponde ad infiniti vettori applicati, uno per ogni possibile punto di applicazione.

Operazioni con i vettori liberi

Andiamo ora a definire 2 operazioni con vettori liberi. Per questo lo “zero” sarà di fondamentale importanza.

Definizione 0.3. Il *vettore nullo* $\vec{0}$ è l'unico vettore di lunghezza nulla. Non ha direzione né verso. È il vettore libero associato a $\overrightarrow{AA}$: $\vec{0} = [\overrightarrow{AA}]$.

Prodotto per uno scalare. La prima operazione che consideriamo è la moltiplicazione di un vettore libero per un numero reale $c \in \mathbb{R}$, che chiamiamo *scalare*.

Sia $\vec{v}$ un vettore libero e sia $c \in \mathbb{R}$.

Se $c = 0$, allora $c \cdot \vec{v} = \vec{0}$.

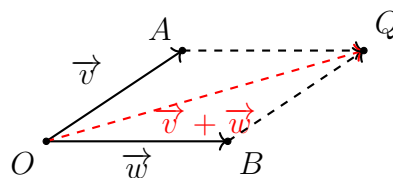
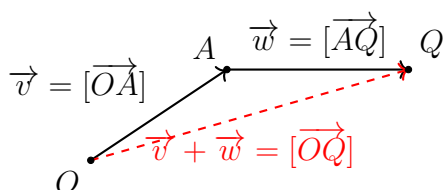
Se $c \neq 0$, allora $c \cdot \vec{v}$ è il vettore libero avente

- (i) la stessa direzione di $\vec{v}$;
- (ii) lo stesso verso di $\vec{v}$, se $c > 0$, e il verso opposto se $c < 0$;
- (iii) lunghezza uguale alla lunghezza di $\vec{v}$ moltiplicata per $|c|$.

Si noti che $c \cdot \vec{v}$ si ottiene contraendo/dilatando ed eventualmente capovolgendo il vettore $\vec{v}$.

Somma di vettori. La somma di due vettori liberi è determinata con la regola del parallelogramma (vedi figura). Dati i vettori liberi $\vec{v}$ e $\vec{w}$, la loro somma $\vec{v} + \vec{w}$ si ottiene nel modo seguente (immagine a sinistra): si applica $\vec{v}$ nel punto O e poi si applica $\vec{w}$ nella punta di $\vec{v}$. Sia Q il punto dove si trova la punta di $\vec{w}$, allora $\vec{v} + \vec{w} = [\overrightarrow{OQ}]$.

In alternativa (immagine a destra), si possono applicare $\vec{v}$ e $\vec{w}$ nello stesso punto O e considerare il parallelogramma avente $\vec{v}$ e $\vec{w}$ come due lati consecutivi. Allora $\vec{v} + \vec{w}$ è la diagonale del parallelogramma.



Queste due operazioni soddisfano le seguenti proprietà.

Proprietà. Siano $\vec{v}$, $\vec{w}$ e $\vec{u}$ vettori liberi e siano c, d numeri reali, allora

- la somma è associativa: $(\vec{v} + \vec{w}) + \vec{u} = \vec{v} + (\vec{w} + \vec{u})$;
- la somma è commutativa: $\vec{v} + \vec{w} = \vec{w} + \vec{v}$;
- la somma ha un elemento neutro: $\vec{v} + \vec{0} = \vec{0} + \vec{v} = \vec{v}$;

- ogni vettore ha un inverso additivo: $-\overrightarrow{AB} = \overrightarrow{BA}$;
- il prodotto con uno scalare è associativo: $(cd) \cdot \vec{v} = c \cdot (d \cdot \vec{v})$;
- il prodotto con uno scalare ha un elemento neutro: $1 \cdot \vec{v} = \vec{v}$;
- distributività: $(c + d) \cdot \vec{v} = c \cdot \vec{v} + d \cdot \vec{v}$.
- distributività: $c \cdot (\vec{v} + \vec{w}) = c \cdot \vec{v} + c \cdot \vec{w}$.

Sistemi di riferimento

Per descrivere un punto nel piano o nello spazio tridimensionale possiamo usare le “coordinate”.

Seguendo quanto il lettore ha studiato alle scuole superiori, consideriamo qui il piano $\mathbb{R}^2$ e lo spazio tridimensionale $\mathbb{R}^3$ con il *sistema di riferimento cartesiano*.

In $\mathbb{R}^2$ il sistema di coordinate cartesiano $(O, \vec{i}, \vec{j})$ è descritto da: un punto O (origine), un vettore libero $\vec{i}$ di lunghezza 1 che descrive l’“asse x ”, e un vettore libero $\vec{j}$ di lunghezza 1 che descrive l’“asse y ” (ottenuto da $\vec{i}$ con una rotazione di $\frac{\pi}{2}$ in senso antiorario).

In $\mathbb{R}^3$ il sistema di coordinate cartesiano $(O, \vec{i}, \vec{j}, \vec{k})$ è descritto da: un punto O (origine), un vettore libero $\vec{i}$ di lunghezza 1 che descrive l’“asse x ”, un vettore libero $\vec{j}$ di lunghezza 1 che descrive l’“asse y ” (ottenuto da $\vec{i}$ con una rotazione di $\frac{\pi}{2}$ in senso antiorario), e un vettore libero $\vec{k}$ di lunghezza 1 che descrive l’“asse z ” (ottenuto da $\vec{i}$ e $\vec{j}$ con la regola della mano destra).

Riassumendo, abbiamo fissato l’origine O e due/tre assi ortogonali orientati e abbiamo scelto un modo per misurare le lunghezze su questi assi. Quindi la posizione di ciascun punto nel piano $\mathbb{R}^2$ o nello spazio tridimensionale $\mathbb{R}^3$ è completamente determinata prendendo le proiezioni ortogonali sugli assi. Possiamo fare lo stesso con i vettori, quindi possiamo identificare

1. il punto P ;
2. il vettore applicato $\overrightarrow{OP}$;
3. e il vettore libero $[\overrightarrow{OP}]$.

Pertanto, in $\mathbb{R}^3$ (e analogamente in $\mathbb{R}^2$) ogni vettore libero $\vec{v} = [\overrightarrow{OP}]$ è descritto attraverso le coordinate (x_P, y_P, z_P) del punto P e può essere espresso come

$$\vec{v} = x_P \cdot \vec{i} + y_P \cdot \vec{j} + z_P \cdot \vec{k}.$$

Utilizzando le coordinate, è possibile eseguire facilmente operazioni sui vettori, ad esempio sia $\vec{v} = 2\vec{i} + 3\vec{j} - 2\vec{k}$ e $\vec{w} = 3\vec{i} - 3\vec{j} + 3\vec{k}$, allora

$$\begin{aligned}\vec{v} + \vec{w} &= 5\vec{i} + 0\vec{j} + \vec{k}; \\ 2\vec{v} &= 4\vec{i} + 6\vec{j} - 4\vec{k}.\end{aligned}$$

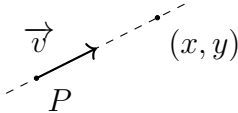
0.2 Rette e piani

Rette

Come si può descrivere una retta in $\mathbb{R}^2$ o in $\mathbb{R}^3$?

Dando la direzione della retta e un suo punto, quindi dando un punto $P = (x_P, y_P)$ ed un vettore libero $\vec{v} = (v_x, v_y)$.

La traslazione (cioè lo spostamento) da P ad un punto (x, y) sulla retta è una dilatazione/contrazione di $\vec{v}$, cioè $(x - x_P, y - y_P) = t(v_x, v_y)$ per $t \in \mathbb{R}$:

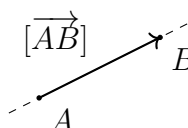
$$\begin{cases} x = x_P + tv_x \\ y = y_P + tv_y \end{cases}$$


Analogamente, in $\mathbb{R}^3$ la retta data dal punto $P = (x_P, y_P, z_P)$ e dal vettore libero $\vec{v} = (v_x, v_y, v_z)$ è descritta dalle equazioni

$$\begin{cases} x = x_P + tv_x \\ y = y_P + tv_y \\ z = z_P + tv_z \end{cases}$$

Osservazione. Un modo alternativo per dare una retta, consiste nel dare 2 punti distinti sulla retta: “per due punti distinti passa una sola retta”.

Dati due punti $A = (x_A, y_A, z_A)$ e $B = (x_B, y_B, z_B)$, essi descrivono un vettore libero $[\vec{AB}]$ e considerando il punto A (o B) basta applicare il ragionamento precedente.



$$\begin{cases} x = x_A + t(x_B - x_A) \\ y = y_A + t(y_B - y_A) \\ z = z_A + t(z_B - z_A) \end{cases}$$

Le equazioni viste dipendono da un parametro ($t \in \mathbb{R}$) e sono delle *equazioni parametriche* della retta.

Attenzione! Le equazioni parametriche di una retta non sono uniche! Infatti esse dipendono sia dalla scelta di un suo punto, sia da un vettore non nullo che descrive la direzione della retta. Sia i punti della retta, sia i vettori che descrivono la sua direzione sono infiniti e quindi c'è una infinità di modi per rappresentare la stessa retta.

Esplicitando il parametro t e sostituendolo nelle altre equazioni, otteniamo delle *equazioni cartesiane* della retta, ad esempio:

$$\begin{cases} x = 1 + 2t \\ y = 2 - 5t \\ z = 1 + t \end{cases} \rightarrow \begin{cases} t = z - 1 \\ x = 1 + 2(z - 1) \\ y = 2 - 5(z - 1) \end{cases} \rightarrow \begin{cases} x - 2z = -1 \\ y + 5z = 7 \end{cases}$$

In generale le equazioni cartesiane della retta sono:

$$\begin{array}{cc} \text{in } \mathbb{R}^2 & \text{in } \mathbb{R}^3 \\ ax + by = c & \begin{cases} a_1x + b_1y + c_1z = d_1 \\ a_2x + b_2y + c_2z = d_2 \end{cases} \end{array}$$

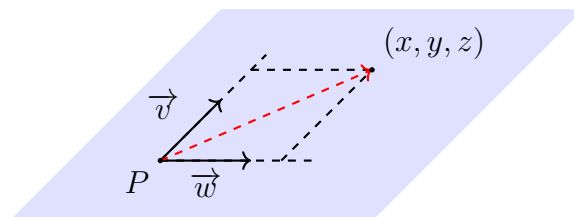
Attenzione! Anche le equazioni cartesiane di una retta non sono uniche.

Piani (in $\mathbb{R}^3$)

Similmente al caso delle rette, un piano in $\mathbb{R}^3$ può essere descritto dando un suo punto e due sue direzioni distinte, quindi dando un punto $P = (x_P, y_P, z_P)$ e due vettori liberi $\vec{v} = (v_x, v_y, v_z)$ e $\vec{w} = (w_x, w_y, w_z)$ che non sono uno multiplo dell'altro.

La traslazione (cioè lo spostamento) da P ad un punto (x, y, z) sul piano si scompone come una dilatazione/contrazione di $\vec{v}$ più una dilatazione/contrazione di $\vec{w}$, cioè $(x - x_P, y - y_P, z - z_P) = t(v_x, v_y, v_z) + s(w_x, w_y, w_z)$ per $t, s \in \mathbb{R}$:

$$\begin{cases} x = x_P + tv_x + sw_x \\ y = y_P + tv_y + sw_y \\ z = z_P + tv_z + sw_z \end{cases}$$



Osservazione. Un modo alternativo per dare un piano, consiste nel dare 3 punti distinti e non collineari, cioè che non giacciono su una stessa retta (ma formano un triangolo)

Dati tre punti $A = (x_A, y_A, z_A)$, $B = (x_B, y_B, z_B)$ e $C = (x_C, y_C, z_C)$, essi descrivono due vettori liberi $[\vec{AB}]$ e $[\vec{AC}]$ e considerando il punto A basta applicare il ragionamento precedente.

$$\begin{cases} x = x_A + t(x_B - x_A) + s(x_C - x_A) \\ y = y_A + t(y_B - y_A) + s(y_C - y_A) \\ z = z_A + t(z_B - z_A) + s(z_C - z_A) \end{cases}$$

Il fatto che i 3 punti siano non collineari corrisponde al fatto che i due vettori liberi $[\vec{AB}]$ e $[\vec{AC}]$ hanno direzioni diverse.

Queste equazioni dipendono da due parametri ($t, s \in \mathbb{R}$) e sono delle *equazioni parametriche* del piano (come per le rette queste *non* sono univocamente determinate!).

Come nel caso delle rette possiamo esplicitare il parametro t in una delle equazioni e sostituendolo nelle altre due, e poi ripetere il ragionamento con s . Otteniamo in questo modo una descrizione in forma algebrica del piano, cioè un'*equazione cartesiana* della piano, ad esempio:

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

In generale, l'equazione cartesiana del piano è della forma $ax + by + cz = d$ (anche in questo caso non c'è un unico modo di descrivere algebricamente un piano).

Un'equazione di questo tipo, polinomiale di grado 1, è detta *lineare*. Le equazioni ed i sistemi lineari avranno un ruolo centrale in questo corso e saranno introdotti nel prossimo capitolo.

Capitolo 1

Matrici e sistemi lineari

In questo capitolo introduciamo, descriviamo e manipoliamo due famiglie di oggetti che ci accompagneranno per tutto il corso: le matrici e i sistemi lineari.

In ogni definizione, teorema, esempio, esercizio, ecc. di questo corso dobbiamo dichiarare con quali “numeri” stiamo lavorando, introduciamo il concetto di *campo*.

Definizione 1.1. Un *campo* $\mathbb{K}$ è un insieme dotato di due operazioni: $+: \mathbb{K} \times \mathbb{K} \rightarrow \mathbb{K}$ (somma) e $\cdot: \mathbb{K} \times \mathbb{K} \rightarrow \mathbb{K}$ (prodotto), che soddisfano le seguenti proprietà:

- C1) $\forall a, b, c \in \mathbb{K}: (a + b) + c = a + (b + c)$ (associatività di $+$)
- C2) $\forall a, b \in \mathbb{K}: a + b = b + a$ (commutatività di $+$)
- C3) $\exists 0 \in \mathbb{K}: a + 0 = 0 + a = a$ (elemento neutro di $+$)
- C4) $\forall a \in \mathbb{K}, \exists b \in \mathbb{K}: a + b = b + a = 0$ (elemento opposto: $b = -a$)
- C5) $\forall a, b, c \in \mathbb{K}: (a \cdot b) \cdot c = a \cdot (b \cdot c)$ (associatività di $\cdot$)
- C6) $\forall a, b \in \mathbb{K}: a \cdot b = b \cdot a$ (commutatività di $\cdot$)
- C7) $\exists 1 \in \mathbb{K}: a \cdot 1 = 1 \cdot a = a$ (elemento neutro di $\cdot$)
- C8) $\forall a \in \mathbb{K}, a \neq 0, \exists b \in \mathbb{K}: a \cdot b = b \cdot a = 1$ (elemento inverso: $b = a^{-1}$)
- C9) $\forall a, b \in \mathbb{K}: a \cdot (b + c) = a \cdot b + a \cdot c$ (distributività della moltiplicazione sulla somma)
- C10) $0 \neq 1$.

Gli elementi di un campo $\mathbb{K}$ sono chiamati *scalari*.

Esempio. i) I numeri naturali $\mathbb{N} = \{0, 1, 2, \dots\}$ con la consueta somma e moltiplicazione non formano un campo, poiché non esistono inversi additivi.

ii) I numeri interi $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$ con la consueta somma e moltiplicazione non formano un campo, poiché non esistono inversi moltiplicativi.

Esempio. I numeri razionali $\mathbb{Q}$ (“frazioni”) e i numeri reali $\mathbb{R}$ con la consueta somma e moltiplicazione formano un campo.

ii) Anche i numeri complessi $\mathbb{C}$ introdotti nel corso “Analisi matematica 1” formano un campo, e abbiamo una catena naturale di inclusioni: $\mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$.

Esempio. Esistono campi con un numero finito di elementi! Ad esempio $\mathbb{K} = \{0, 1\}$ con somma e prodotto dati dalle seguenti tabelle:

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

1.1 Matrici

Definizione 1.2. Siano dati $m, n \in \mathbb{Z}$ positivi e sia $\mathbb{K}$ un campo.

Una *matrice* $m \times n$ (o *matrice* di tipo (m, n)) a coefficienti in $\mathbb{K}$ è una tabella di elementi di $\mathbb{K}$ disposti su m righe e n colonne. Tali numeri sono detti *entrate*, *coefficienti*, o *componenti* della matrice. L'insieme di tutte le matrici $m \times n$ a coefficienti in $\mathbb{K}$ si indica con $\mathcal{M}_{\mathbb{K}}(m, n)$.

Notazione. Altre notazioni comuni per l'insieme di tutte le matrici $m \times n$ a coefficienti in $\mathbb{K}$ sono $\mathcal{M}_{m \times n}(\mathbb{K})$, $\text{Mat}_{\mathbb{K}}(m, n)$ o $\text{Mat}(m \times n, \mathbb{K})$.

Esempio.

$$A = \begin{pmatrix} -1 & 0 & 2 & \pi \\ 0 & \sqrt{2} & -2/3 & 6 \end{pmatrix} \in \mathcal{M}_{\mathbb{R}}(2, 4) \quad B = \begin{pmatrix} 1 & \pi \\ 0 & 0 \\ -3\sqrt{19} & 2 \end{pmatrix} \in \mathcal{M}_{\mathbb{R}}(3, 2) \quad C = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \in \mathcal{M}_{\mathbb{R}}(2, 2)$$

$$D = \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix} \in \mathcal{M}_{\mathbb{R}}(3, 1) \quad E = \begin{pmatrix} 1 & 2 & -9 & 0 \end{pmatrix} \in \mathcal{M}_{\mathbb{R}}(1, 4)$$

Terminologia. Vi sono alcuni casi particolari che vale la pena evidenziare:

- $m = n = 1$. $A = (a_{1,1})$, quindi abbiamo una biiezione naturale $\mathcal{M}_{\mathbb{K}}(1, 1) \cong \mathbb{K}$.
- $m = n$. Una matrice di tipo (n, n) è chiamata *matrice quadrata*. L'insieme delle matrici quadrate $n \times n$ a coefficienti in $\mathbb{K}$ si indica con $\mathcal{M}_{\mathbb{K}}(n)$.
- $m = 1, n > 1$. $A = (a_{1,1} \ a_{1,2} \ \dots \ a_{1,n})$ è chiamata *vettore riga* (di lunghezza n).

Un vettore riga dipende da n scelte di elementi in $\mathbb{K}$, quindi $\mathcal{M}_{\mathbb{K}}(1, n)$ è naturalmente in biiezione con il *prodotto cartesiano*

$$\mathbb{K}^n = \underbrace{\mathbb{K} \times \mathbb{K} \times \dots \times \mathbb{K}}_{n\text{-volte}} = \{(k_1, k_2, \dots, k_n) : k_i \in \mathbb{K}\}$$

- $m > 1, n = 1$. $A = \begin{pmatrix} a_{1,1} \\ a_{2,1} \\ \vdots \\ a_{m,1} \end{pmatrix}$ è chiamato *vettore colonna* (di altezza m).

Come per i vettori riga, abbiamo una biiezione naturale tra l'insieme dei vettori colonna $\mathcal{M}_{\mathbb{K}}(m, 1)$ e $\mathbb{K}^m$, poiché ogni vettore colonna dipende da m elementi di $\mathbb{K}$.

- la matrice $O \in \mathcal{M}_{\mathbb{K}}(m, n)$ avente tutte le entrate nulle viene detta *matrice nulla*.

In generale, per rappresentare una matrice A , utilizziamo la seguente notazione:

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix} = (a_{i,j})$$

Quindi $a_{i,j}$ è l'elemento sulla riga i -esima e sulla colonna j -esima. Questo elemento può anche essere indicato con $(A)_{i,j}$.

1.1.1 Operazioni fra matrici

Uguaglianza

Siano $A \in \mathcal{M}_{\mathbb{K}}(m, n)$ e $B \in \mathcal{M}_{\mathbb{K}}(p, q)$ due matrici. Per essere uguali, A e B devono essere dello stesso tipo: $m = p$ e $n = q$, ed avere le stesse entrate: $a_{i,j} = b_{i,j} \quad \forall i = 1, \dots, m, \forall j = 1, \dots, n$.

Somma di matrici

Siano $A, B \in \mathcal{M}_{\mathbb{K}}(m, n)$ matrici *dello stesso tipo*, allora definiamo la loro somma $A + B$ come la matrice in $\mathcal{M}_{\mathbb{K}}(m, n)$, ottenuta sommando componente per componente gli elementi di A e B :

$$\begin{aligned} + : \mathcal{M}_{\mathbb{K}}(m, n) \times \mathcal{M}_{\mathbb{K}}(m, n) &\longrightarrow \mathcal{M}_{\mathbb{K}}(m, n) \\ (A, B) &\longmapsto (A + B), (A + B)_{ij} = a_{i,j} + b_{i,j} \end{aligned}$$

Esempio.

$$\mathbb{K} = \mathbb{Q}, m = 2, n = 3, \quad \begin{pmatrix} 1 & 3 & 2 \\ 0 & 2 & -4 \end{pmatrix} + \begin{pmatrix} -1 & 3 & 6 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 6 & 8 \\ 1 & 2 & -4 \end{pmatrix}$$

Osservazione. La somma è definita solo per matrici dello stesso tipo.

È facile verificare che la somma di matrici eredita tutte le proprietà della somma in $\mathbb{K}$:

Proprietà. La somma di matrici soddisfa le seguenti proprietà.

- *Associatività:* $\forall A, B, C \in \mathcal{M}_{\mathbb{K}}(m, n): (A + B) + C = A + (B + C)$.
- *Commutatività:* $\forall A, B \in \mathcal{M}_{\mathbb{K}}(m, n): A + B = B + A$.
- *La matrice nulla è l'unico elemento neutro per la somma, cioè è l'unica matrice tale che $O + A = A + O = A$ per ogni $A \in \mathcal{M}_{\mathbb{K}}(m, n)$*
- *L'opposto di $A = (a_{i,j}) \in \mathcal{M}_{\mathbb{K}}(m, n)$ è la matrice $A' = (a'_{i,j}) \in \mathcal{M}_{\mathbb{K}}(m, n)$, con $A + A' = O$, ovvero $a'_{i,j} = -a_{i,j}$.*

Prodotto per scalare

Sia $A \in \mathcal{M}_{\mathbb{K}}(m, n)$ una matrice e $k \in \mathbb{K}$ uno scalare. Moltiplicando ogni entrata di A per lo scalare k otteniamo una nuova matrice $kA \in \mathcal{M}_{\mathbb{K}}(m, n)$:

$$\begin{aligned} \cdot : \mathbb{K} \times \mathcal{M}_{\mathbb{K}}(m, n) &\longrightarrow \mathcal{M}_{\mathbb{K}}(m, n) \\ (k, A) &\longmapsto kA, (kA)_{i,l} = ka_{i,j} \end{aligned}$$

Esempio.

$$\mathbb{K} = \mathbb{Q}, m = 2, n = 3, \quad -2 \cdot \begin{pmatrix} 1 & 3 & 2 \\ 0 & 2 & -4 \end{pmatrix} = \begin{pmatrix} -2 & -6 & -4 \\ 0 & -4 & 8 \end{pmatrix}$$

È facile verificare le seguenti proprietà.

Proprietà. *Il prodotto per scalare soddisfa le seguenti proprietà.*

- *Associatività:* $\forall k_1, k_2 \in \mathbb{K}, \forall A \in \mathcal{M}_{\mathbb{K}}(m, n): k_1 \cdot (k_2 \cdot A) = (k_1 \cdot k_2) \cdot A$.
- *1 $\in \mathbb{K}$ è l'elemento neutro:* $1 \cdot A = A$.
- *Distributività:* $\forall k_1, k_2 \in \mathbb{K}, \forall A \in \mathcal{M}_{\mathbb{K}}(m, n): (k_1 + k_2) \cdot A = k_1 \cdot A + k_2 \cdot A$
- *Distributività:* $\forall k \in \mathbb{K}, \forall A, B \in \mathcal{M}_{\mathbb{K}}(m, n): k \cdot (A + B) = k \cdot A + k \cdot B$.
- *Legge di annullamento del prodotto per scalare:* per ogni $A \in \mathcal{M}_{\mathbb{K}}(m, n)$ e $k \in \mathbb{K}$, vale $kA = O$ se e solo se $k = 0$ oppure $A = O$.

Osservazione. Si noti che $A + (-1 \cdot A) = O$, quindi l'inverso additivo di A è $(-1 \cdot A) = -A$.

Matrice trasposta

Data una matrice A , possiamo definire un'altra matrice scambiando righe e colonne di A .

Definizione 1.3. Sia $A \in \mathcal{M}_{\mathbb{K}}(m, n)$ una matrice. La *trasposta di A* è la matrice $B \in \mathcal{M}_{\mathbb{K}}(n, m)$ definita da $b_{i,j} = a_{j,i}$, ed è indicata con A^T .

$$\begin{aligned} T : \mathcal{M}_{\mathbb{K}}(m, n) &\longrightarrow \mathcal{M}_{\mathbb{K}}(n, m) \\ A &\longmapsto A^T, (A^T)_{i,j} = a_{j,i} \end{aligned}$$

Esempio. Sia $A = \begin{pmatrix} 1 & 3 & 2 \\ 0 & 2 & -4 \end{pmatrix} \in \mathcal{M}_{\mathbb{Q}}(2, 3)$, allora $A^T = \begin{pmatrix} 1 & 0 \\ 3 & 2 \\ 2 & -4 \end{pmatrix} \in \mathcal{M}_{\mathbb{Q}}(3, 2)$.

Dalle definizioni date seguono direttamente le seguenti proprietà.

Proprietà. *Valgono le seguenti proprietà:*

- $\forall A \in \mathcal{M}_{\mathbb{K}}(m, n): (A^T)^T = A$;
- $\forall A, B \in \mathcal{M}_{\mathbb{K}}(m, n): (A + B)^T = A^T + B^T$;
- $\forall A \in \mathcal{M}_{\mathbb{K}}(m, n), k \in \mathbb{K}: (kA)^T = k(A^T)$;

Moltiplicazione di matrici (righe per colonne)

Concludiamo questa sezione introducendo una nuova operazione detta prodotto tra matrici, o anche prodotto righe per colonne.

Definiremo questa operazione in 2 passaggi: definiremo prima la moltiplicazione riga-colonna, cioè il prodotto di un vettore riga per un vettore colonna; passeremo poi a dare la definizione generale.

Siano $A = (a_{1,1} \dots a_{1,p}) \in \mathcal{M}_{\mathbb{K}}(1, p)$ un vettore riga e $B = \begin{pmatrix} b_{1,1} \\ \vdots \\ b_{p,1} \end{pmatrix} \in \mathcal{M}_{\mathbb{K}}(p, 1)$ un vettore colonna. Il prodotto delle matrici A e B (nell'ordine indicato!) è la matrice di $\mathcal{M}_{\mathbb{K}}(1, 1) = \mathbb{K}$, indicata con AB e definita da

$$AB = (a_{1,1} \dots a_{1,p}) \begin{pmatrix} b_{1,1} \\ \vdots \\ b_{p,1} \end{pmatrix} = (a_{1,1} \cdot b_{1,1} + a_{1,2} \cdot b_{2,1} + \dots + a_{1,p} \cdot b_{p,1}) = \left(\sum_{l=1}^p (a_{1,l} \cdot b_{l,1}) \right).$$

Esempio. ($\mathbb{K} = \mathbb{Q}$, $p = 3$)

$$(1, 2, 3) \cdot \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} = (0 \cdot 1 + 2 \cdot 1 + 3 \cdot (-1)) = (-1)$$

Questa operazione è chiamata *moltiplicazione riga-colonna*. In generale, per determinare $A \cdot B$ è necessario eseguire l'operazione mn volte.

Siano $A \in \mathcal{M}_{\mathbb{K}}(m, p)$ e $B \in \mathcal{M}_{\mathbb{K}}(p, n)$ (attenzione al tipo delle matrici!), allora il loro prodotto è definito come

$$\begin{aligned} \cdot : \mathcal{M}_{\mathbb{K}}(m, p) \times \mathcal{M}_{\mathbb{K}}(p, n) &\longrightarrow \mathcal{M}_{\mathbb{K}}(m, n), \\ (A, B) &\longmapsto A \cdot B \end{aligned}$$

dove

$$(A \cdot B)_{i,j} = a_{i,1} \cdot b_{1,j} + a_{i,2} \cdot b_{2,j} + \dots + a_{i,p} \cdot b_{p,j} = \sum_{l=1}^p (a_{i,l} \cdot b_{l,j})$$

Quindi l'elemento $(A \cdot B)_{i,j}$ si ottiene moltiplicando la i -esima riga di A con la j -esima colonna di B .

Osservazione. Si noti che non richiediamo che A e B siano dello stesso tipo, ma è necessario che il numero di colonne di A coincida con il numero di righe di B , altrimenti AB non è definito. La matrice AB ha tante righe quante A e tante colonne quante B .

Esempio. i) Siano $A \in \mathcal{M}_{\mathbb{Q}}(2, 3)$ e $B \in \mathcal{M}_{\mathbb{Q}}(3, 2)$ le matrici

$$A = \begin{pmatrix} 1 & 3 & 2 \\ 0 & 2 & -4 \end{pmatrix}, \quad B = \begin{pmatrix} -1 & 1 \\ 3 & 0 \\ 6 & 0 \end{pmatrix}$$

Per calcolare $C = AB$ dobbiamo eseguire 4 moltiplicazioni riga-colonna, ad esempio $c_{1,2}$ si ottiene dalla moltiplicazione riga-colonna della 1ª riga di A con la 2ª colonna di B :

$$\begin{pmatrix} 1 & 3 & 2 \\ 0 & 2 & -4 \end{pmatrix} \cdot \begin{pmatrix} -1 & 1 \\ 3 & 0 \\ 6 & 0 \end{pmatrix} = \begin{pmatrix} 1 \cdot (-1) + 3 \cdot 3 + 2 \cdot 6 & 1 \cdot 1 + 3 \cdot 0 + 2 \cdot 0 \\ 0 \cdot (-1) + 2 \cdot 3 + (-4) \cdot 6 & 0 \cdot 1 + 2 \cdot 0 + (-4) \cdot 0 \end{pmatrix} = \begin{pmatrix} 20 & 1 \\ -18 & 0 \end{pmatrix}$$

ii) Sia $A \in \mathcal{M}_{\mathbb{Q}}(2, 3)$ come sopra e sia $B' = \begin{pmatrix} -1 & 3 & 6 \\ 1 & 0 & 0 \end{pmatrix} \in \mathcal{M}_{\mathbb{Q}}(2, 3)$, allora il prodotto $A \cdot B'$ non è definito, poiché A ha 3 colonne, ma B' ha solo 2 righe.

Definizione 1.4. La *matrice identità* I_n è la matrice quadrata $I_n \in \mathcal{M}_{\mathbb{K}}(n, n)$ avente:

$$(I_n)_{i,j} = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases} \quad I_n = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \ddots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

In una matrice quadrata, la posizione con indice di riga uguale all'indice di colonna forma la *diagonale*.

Proprietà. Il prodotto righe per colonne soddisfa le seguenti proprietà:

- *Associatività:* $\forall A \in \mathcal{M}_{\mathbb{K}}(m, n), B \in \mathcal{M}_{\mathbb{K}}(n, p), C \in \mathcal{M}_{\mathbb{K}}(p, q)$ vale $(AB)C = A(BC) \in \mathcal{M}_{\mathbb{K}}(m, q)$.
- *Elemento neutro:* $\forall A \in \mathcal{M}_{\mathbb{K}}(m, n)$, vale $A \cdot I_n = A$ e $I_m \cdot A = A$.
- *Distributività:* $\forall A \in \mathcal{M}_{\mathbb{K}}(m, n), B_1, B_2 \in \mathcal{M}_{\mathbb{K}}(n, p)$ vale $A(B_1 + B_2) = AB_1 + AB_2$. Inoltre, $\forall A_1, A_2 \in \mathcal{M}_{\mathbb{K}}(m, n), B \in \mathcal{M}_{\mathbb{K}}(n, p)$ vale $(A_1 + A_2)B = A_1B + A_2B$.
- *Compatibilità col prodotto per scalare:* $\forall A \in \mathcal{M}_{\mathbb{K}}(m, n), B \in \mathcal{M}_{\mathbb{K}}(n, p), k \in \mathbb{K}$ vale $k(AB) = (kA)B = A(kB)$.
- *Compatibilità con la trasposta:* $\forall A \in \mathcal{M}_{\mathbb{K}}(m, p), B \in \mathcal{M}_{\mathbb{K}}(p, n)$: $(AB)^T = B^T A^T$.

Dimostrazione. Per dimostrare le uguaglianze tra matrici usiamo la definizione. In tutte le uguaglianze da dimostrare i tipi delle matrici a destra e a sinistra dell'uguale coincidono, quindi dobbiamo solo controllare che le entrate sulla i -esima riga e j -esima colonna coincidano.

Associatività. Siano $D := AB$ e $E := BC$:

$$d_{i,j} = (AB)_{i,j} = \sum_{l=1}^n (a_{i,l} \cdot b_{l,j}) \in \mathcal{M}_{\mathbb{K}}(m, p), \quad e_{i,j} = (BC)_{i,j} = \sum_{l=1}^p (b_{i,l} \cdot c_{l,j}) \in \mathcal{M}_{\mathbb{K}}(n, q),$$

quindi

$$((AB)C)_{i,j} = \sum_{s=1}^p (d_{i,s} \cdot c_{s,j}) = \sum_{s=1}^p \sum_{l=1}^n (a_{i,l} \cdot b_{l,s}) \cdot c_{s,j} = \sum_{l=1}^n a_{i,l} \left(\sum_{s=1}^p b_{l,s} \cdot c_{s,j} \right) = \sum_{l=1}^n a_{i,l} e_{l,j} = (A(BC))_{i,j}$$

Elemento neutro. $(AI_n)_{i,j} = \sum_{k=1}^n a_{i,k} e_{k,j} = a_{i,j} \cdot e_{j,j} = a_{i,j}$; analogamente $(I_m A)_{i,j} = a_{i,j}$.

La distributività e la compatibilità col prodotto per scalare seguono dalla definizione, e sono lasciate al lettore per esercizio.

Dimostriamo infine la compatibilità con la trasposta: $(AB)_{i,j} = \sum_{l=1}^p a_{i,l} b_{l,j}$, quindi

$$((AB)^T)_{i,j} = (AB)_{j,i} = \sum_{l=1}^p a_{j,l} b_{l,i} = \sum_{l=1}^p b_{l,i} a_{j,l} = \sum_{l=1}^p (B^T)_{i,l} (A^T)_{l,j} = (B^T A^T)_{i,j}.$$

□

E la *commutatività*? In generale, il prodotto di matrici non è commutativo!

- i) Siano $A \in \mathcal{M}_{\mathbb{K}}(m, p)$ e $B \in \mathcal{M}_{\mathbb{K}}(p, n)$. $AB \in \mathcal{M}_{\mathbb{K}}(m, n)$, ma BA non è nemmeno definito se $m \neq n$.
- ii) Supponiamo quindi che $m = n$, quindi sia AB che BA sono definiti: $AB \in \mathcal{M}_{\mathbb{K}}(m)$ e $BA \in \mathcal{M}_{\mathbb{K}}(p)$. Se $p \neq m$, le due matrici sono di tipo diverso, quindi sono diverse.
- iii) Supponiamo ulteriormente che $m = n = p$, quindi $AB, BA \in \mathcal{M}_{\mathbb{K}}(n)$. Anche in questo caso AB e BA possono essere diverse, ad esempio $\mathbb{K} = \mathbb{Q}$, $m = n = p = 2$:

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \text{ allora } AB = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = BA$$

- iv) Si noti che in alcuni casi è possibile che $AB = BA$, ad esempio $\mathbb{K} = \mathbb{Q}$, $m = n = p = 2$:

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \text{ allora } AB = BA = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Osservazione. L'ultimo esempio mostra che il prodotto righe per colonne *non soddisfa* la legge di annullamento del prodotto: è possibile che il prodotto di 2 matrici diverse da zero sia la matrice nulla O , un fenomeno che non può verificarsi per gli elementi di un campo, ad esempio per i numeri reali.

Per una matrice non quadrata $B \in \mathcal{M}_{\mathbb{K}}(m, n)$, $m \neq n$ il prodotto $B \cdot B$ non è definito. Sia invece $A \in \mathcal{M}_{\mathbb{K}}(n)$ una matrice quadrata, allora $A \cdot A \in \mathcal{M}_{\mathbb{K}}(n)$ e possiamo iterare il prodotto, quindi possiamo definire la potenza p -esima di A per ogni intero $p \geq 0$:

$$A^0 := I_n, A^1 := A, A^2 := A \cdot A, A^3 = A^2 \cdot A = A \cdot A \cdot A, \dots, A^p = A^{p-1} A$$

Osservazione. La non commutatività del prodotto di matrici *non permette* di usare i “prodotti notevoli” nel fare i calcoli. Ad esempio, in generale si ha che

$$(A + B)^2 = (A + B) \cdot (A + B) = A \cdot A + A \cdot B + B \cdot A + B \cdot B \neq A^2 + 2AB + B^2$$

Per convincersene, basta prendere le matrici al punto iii) precedente.

1.2 Sistemi lineari

Definizione 1.5. Siano m, n numeri interi positivi. Sia $A \in \mathcal{M}_{\mathbb{K}}(m, n)$ una matrice e sia $b \in \mathcal{M}_{\mathbb{K}}(m, 1)$ un vettore colonna (di altezza m).

Un *sistema lineare* (o sistema di equazioni lineari) è un'equazione della forma $Ax = b$, dove $x \in \mathcal{M}_{\mathbb{K}}(n, 1)$:

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \quad (1.1)$$

La matrice A è chiamata *matrice dei coefficienti* e il vettore colonna b è chiamato *colonna dei termini noti*.

Riscrivendo l'equazione (1.1), otteniamo un sistema di m equazioni lineari in n variabili:

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \dots + a_{1,n}x_n &= b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \dots + a_{2,n}x_n &= b_2 \\ \vdots &\vdots \\ a_{m,1}x_1 + a_{m,2}x_2 + \dots + a_{m,n}x_n &= b_m \end{cases}$$

Notazione. i) Ogni riga è un'equazione polinomiale di grado 1, tali equazioni sono dette lineari. Abbiamo visto nel Capitolo 0, che le equazioni che descrivono rette e piani in $\mathbb{R}^2$ e $\mathbb{R}^3$ sono lineari. In particolare, una retta in $\mathbb{R}^3$ è descritta da un sistema lineare $Ax = b$ con $A \in \mathcal{M}_{\mathbb{R}}(2, 3)$, $b \in \mathcal{M}_{\mathbb{R}}(2, 1)$.

ii) Un sistema lineare $Ax = b$ è solitamente rappresentato attraverso la sua *matrice completa* $(A|b) \in \mathcal{M}_{\mathbb{K}}(m, n+1)$, le cui prime n colonne sono quelle di A e l'ultima è b :

$$(A|b) = \left(\begin{array}{cccc|c} a_{1,1} & a_{1,2} & \dots & a_{1,n} & b_1 \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} & b_n \end{array} \right)$$

Esempio. La matrice completa ($\mathbb{K} = \mathbb{Q}$)

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 2 & -1 & 1 & 1 \end{array} \right) \text{ rappresenta il sistema lineare } \begin{cases} x_1 + x_2 + x_3 = 0 \\ 2x_1 - x_2 + x_3 = 1 \end{cases}$$

Definizione 1.6. Una *soluzione* del sistema lineare $Ax = b$, è un vettore colonna $s \in \mathcal{M}_{\mathbb{K}}(n, 1)$ le cui componenti $s_1, s_2, \dots, s_n \in \mathbb{K}$ soddisfano simultaneamente tutte le m equazioni.

Indichiamo con $\text{Sol}(A|b)$ l'insieme di tutte le soluzioni di $Ax = b$:

$$\text{Sol}(A|b) = \{s \in \mathcal{M}_{\mathbb{K}}(n, 1) \text{ tali che } As = b\} \subseteq \mathcal{M}_{\mathbb{K}}(n, 1).$$

Si dice che il sistema $Ax = b$ è *risolubile* se ammette almeno una soluzione: $\text{Sol}(A|b) \neq \emptyset$.

Esempio. Sia $m = n = 1$, in modo che il sistema lineare sia costituito da una singola equazione in una variabile: $ax = b$, con $a, b \in \mathbb{K}$.

Se $a \neq 0$, possiamo dividere entrambi i lati per a e otteniamo un'unica soluzione: $x = \frac{b}{a}$; ad esempio $3x = 4$ ha un'unica soluzione: $x = \frac{4}{3}$.

Se $a = 0$, l'equazione si riduce a $0x = b$ e abbiamo 2 sottocasi:

i) se $b = 0$, allora abbiamo un'equazione della forma " $0 = 0$ ", quindi qualsiasi $k \in \mathbb{K}$ è una soluzione;

ii) se $b \neq 0$, allora abbiamo un'equazione della forma " $0 = 1$ ", che non ha alcuna soluzione.

Esempio. Sia $m = n = 2$ ($\mathbb{K} = \mathbb{R}$) e si consideri il sistema lineare

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

La seconda equazione è equivalente a $x = y$, sostituendo questa nella prima equazione otteniamo $2x = 2$, quindi $x = 1, y = 1$ è l'unica soluzione del sistema lineare.

Consideriamo ora i sistemi lineari

$$I) \begin{cases} x + y = 2 \\ x + y = 1 \end{cases} \quad II) \begin{cases} x + y = 2 \\ 3x + 3y = 6 \end{cases}$$

Il sistema lineare I) non ha soluzione poiché $2 \neq 1$. Il sistema lineare II) si riduce alla singola equazione $x + y = 2$ poiché la seconda equazione è 3 volte la prima. Quindi qualsiasi coppia $(s, 2 - s)$, $s \in \mathbb{K}$ è una soluzione.

Primo obiettivo del corso: capire se un dato sistema lineare ha soluzioni e, in caso affermativo, come trovarle e/o descriverle.

1.2.1 Matrici a scala

Definizione 1.7. Una matrice $A \in \mathcal{M}_{\mathbb{K}}(m, n)$ è detta *in forma a scala* (o solo *a scala*) se:

- ogni riga non nulla (cioè con almeno un'entrata diversa da zero) inizia con più zeri rispetto alla precedente;
- tutte le righe nulle (cioè con soli zeri) si trovano sotto tutte le righe non nulle.

Un sistema lineare $Ax = b$ è *in forma a scala* (o *a scala*) se la sua matrice completa $(A|b)$ è a scala.

Definizione 1.8. Sia $A \in \mathcal{M}_{\mathbb{K}}(m, n)$ una matrice a scala. La prima voce diversa da zero (da sinistra) in ogni riga non nulla è chiamata *pivot*.

Osservazione. Ogni pivot è posizionato su una colonna più a destra della colonna del pivot “precedente”; quindi i pivot di A sono in posizioni $a_{i,j}$ con $j \geq i$.

Esempio. La matrice nulla e la matrice identità sono a scala (su ogni campo $\mathbb{K}$).

Le matrici ($\mathbb{K} = \mathbb{R}$)

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

sono in forma a scala. I pivot della prima matrice sono 2, 4 e 3; nella seconda matrice i pivot sono 1 e 4.

Le seguenti matrici ($\mathbb{K} = \mathbb{R}$) non sono in forma a scala:

$$\begin{pmatrix} 2 & 2 & 3 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 4 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Esempio. Consideriamo i seguenti sistemi lineari (su $\mathbb{K} = \mathbb{R}$) in forma a scala

$$i) (A|b) = \left(\begin{array}{ccc|c} 1 & -1 & 3 & 5 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 1 & 7 \end{array} \right) \longleftrightarrow \begin{cases} x_1 - x_2 + 3x_3 = 5 \\ x_2 + 2x_3 = 2 \\ x_3 = 7 \end{cases}$$

È facile risolverlo: l'ultima equazione restituisce $x_3 = 7$, lo sostituiamo nella seconda equazione e otteniamo $x_2 = 2 - 2 \cdot 7 = -12$; analogamente $x_1 = -28$, quindi $\text{Sol}(A|b) = \{(-28, -12, 7)^T\}$.

$$ii) (A|b) = \left(\begin{array}{ccc|c} 1 & -1 & 3 & 5 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & 7 \end{array} \right) \longleftrightarrow \begin{cases} x_1 - x_2 + 3x_3 = 5 \\ x_2 + 2x_3 = 2 \\ 0 = 7 \end{cases}$$

Non ha soluzioni, perché $0 = 7$ è impossibile.

$$iii) (A|b) = \left(\begin{array}{ccc|c} 1 & -1 & 3 & 5 \\ 0 & 1 & 2 & 2 \end{array} \right) \longleftrightarrow \begin{cases} x_1 - x_2 + 3x_3 = 5 \\ x_2 + 2x_3 = 2 \end{cases}$$

Nell'ultima equazione compaiono 2 variabili, quindi esprimiamo x_2 in termini di x_3 : $x_2 = 2 - 2x_3$, sostituendo nella prima equazione, otteniamo $x_1 = 5 + x_2 - 3x_3 = 7 - 5x_3$, quindi abbiamo una variabile libera (x_3) e

$$\text{Sol}(A|b) = \left\{ \begin{pmatrix} 7-5t \\ 2-2t \\ t \end{pmatrix} \mid t \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} 7 \\ 2 \\ 0 \end{pmatrix} + \begin{pmatrix} -5 \\ -2 \\ 1 \end{pmatrix} t \mid t \in \mathbb{R} \right\} = \begin{pmatrix} 7 \\ 2 \\ 0 \end{pmatrix} + \begin{pmatrix} -5 \\ -2 \\ 1 \end{pmatrix} \mathbb{R}$$

$$iv) (A|b) = \left(\begin{array}{ccc|c} 1 & 1 & -1 & 1 \end{array} \right) \longleftrightarrow x_1 + x_2 - x_3 = 1$$

C'è solo un'equazione con 3 variabili, quindi esprimiamo una di esse, diciamo x_3 , in termini delle altre due: $x_3 = x_1 + x_2 - 1$, quindi abbiamo due variabili libere (x_1 e x_2) e

$$\text{Sol}(A|b) = \left\{ \begin{pmatrix} s \\ t \\ s+t-1 \end{pmatrix} \mid t, s \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} t + \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} s \mid t, s \in \mathbb{R} \right\} = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \mathbb{R} + \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \mathbb{R}$$

$$v) (A|b) = \left(\begin{array}{ccc|c} 1 & 0 & -1 & 1 \\ 0 & 0 & 1 & 2 \end{array} \right) \longleftrightarrow \begin{cases} x_1 - x_3 = 1 \\ x_3 = 2 \end{cases}$$

La seconda equazione restituisce $x_3 = 2$, lo sostituiamo nella prima e otteniamo $x_1 = 1 + 2 = -3$, e non ci sono vincoli su x_2 , che è quindi una variabile libera: $\text{Sol}(A|b) = \{(3, s, 2)^T \mid s \in \mathbb{R}\}$

$$\text{Sol}(A|b) = \left\{ \begin{pmatrix} 3 \\ s \\ 2 \end{pmatrix} \mid s \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} 3 \\ 0 \\ 2 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} t \mid t \in \mathbb{R} \right\} = \begin{pmatrix} 3 \\ 0 \\ 2 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \mathbb{R}$$

Dato un sistema lineare a scala, possiamo facilmente verificare se è risolubile e, in caso affermativo, determinare tutte le soluzioni mediante sostituzione (come sopra).

Quindi il nostro obiettivo è ora quello di trasformare un sistema lineare in uno *equivalente* (cioè avente lo stesso insieme di soluzioni), che sia in forma a scala. Lo faremo utilizzando le *operazioni elementari sulle righe*.

1.2.2 Operazioni elementari sulle righe e metodo di eliminazione di Gauss

Sia $A \in \mathcal{M}_{\mathbb{K}}(m, n)$ una matrice e siano $R_1, \dots, R_m$ le sue righe. Le *operazioni elementari sulle righe* sono:

1. scambiare due righe ($R_i \leftrightarrow R_j$);
2. moltiplicare una riga per uno scalare non nullo $a \in \mathbb{K}, k \neq 0$ ($R_i \rightarrow kR_i$);
3. aggiungere ad una riga un multiplo (anche nullo) di un'altra riga: ($R_i \rightarrow R_i + kR_j$, $k \in \mathbb{K}$).

Osservazione. Considerando le operazioni corrispondenti sulle equazioni di un sistema lineare, è facile dimostrare che le operazioni elementari sulle righe *non modificano* l'insieme delle soluzioni di un sistema lineare. In altre parole, se la matrice $(U|b')$ è ottenuta da $(A|b)$ attraverso una sequenza di operazioni elementari sulle righe, allora $\text{Sol}(A|b) = \text{Sol}(U|b')$.

Il *metodo di eliminazione di Gauss*, in breve MEG, è anche detto algoritmo di Gauss, e prende come input una matrice $A \in \mathcal{M}_{\mathbb{K}}(m, n)$ e restituisce come output una matrice $U \in \mathcal{M}_{\mathbb{K}}(m, n)$ in forma a scala, ottenuta da A tramite una sequenza di operazioni elementari sulle righe. L'idea è quella di “ripulire” le colonne una alla volta, da sinistra verso destra.

Metodo di eliminazione di Gauss (MEG).

Input: $A \in \mathcal{M}_{\mathbb{K}}(m, n)$.

Passo 1. Se A è in forma a scala, allora Output: $U = A$.

Passo 2. Sia j l'indice della prima colonna diversa da zero e sia i un indice di riga tale che $a_{i,j} \neq 0$. Si scambino la prima riga con la riga i -esima riga per ottenere la matrice B :

$$A = \begin{pmatrix} 0 & \dots & 0 & * & \dots & * \\ \vdots & & \vdots & & & \\ 0 & \dots & 0 & a_{i,j} & & \\ \vdots & & \vdots & & & \\ 0 & \dots & 0 & * & \dots & * \end{pmatrix} \xrightarrow{R_1 \leftrightarrow R_i} B = \begin{pmatrix} 0 & \dots & 0 & p_1 & \dots & * \\ \vdots & & \vdots & & & \\ 0 & \dots & 0 & b_{i,j} & & \\ \vdots & & \vdots & & & \\ 0 & \dots & 0 & * & \dots & * \end{pmatrix}$$

Si osservi che $p_1 = a_{i,j} \neq 0$.

Passo 3. Per ogni riga $l \neq 1$ tale che $b_{l,j} \neq 0$, applichiamo la mossa: $R_l \rightarrow R_l + \left(-\frac{b_{l,j}}{p_1}\right) R_1$,

in modo che $b_{l,j}$ diventi $b_{l,j} + \left(-\frac{b_{l,j}}{p_1}\right) p_1 = 0$:

$$B = \begin{pmatrix} 0 & \dots & 0 & p_1 & \dots & * \\ \vdots & & \vdots & & & \\ 0 & \dots & 0 & b_{l,j} & & \\ \vdots & & \vdots & & & \\ 0 & \dots & 0 & * & \dots & * \end{pmatrix} \xrightarrow{R_l \rightarrow R_l + \left(-\frac{b_{l,j}}{p_1}\right) R_1} C = \begin{pmatrix} 0 & \dots & 0 & p_1 & * & \dots & * \\ \vdots & & \vdots & 0 & & & \\ \vdots & & \vdots & 0 & & & \\ 0 & \dots & 0 & 0 & \boxed{A'} & & \end{pmatrix}$$

Passo 4. “Riavviare” l'algoritmo con $A' \in \mathcal{M}_{\mathbb{K}}(m-1, n')$ ($n' < n$).

Osservazione. i) L'algoritmo termina dopo al più m iterazioni.

ii) Esistono diversi modi per portare una matrice A in forma a scala, quindi l'output non è unico!

Esempio. Si consideri la matrice $M = \begin{pmatrix} 0 & -3 & -5 & 1 \\ 2 & -4 & 2 & 6 \\ 1 & -1 & 3 & 5 \end{pmatrix} \in \mathcal{M}_{\mathbb{R}}(3, 4)$. Utilizziamo l'operazione elementare sulle righe ridurla a scala:

$$\begin{pmatrix} 0 & -3 & -5 & 1 \\ 2 & -4 & 2 & 6 \\ 1 & -1 & 3 & 5 \end{pmatrix} \xrightarrow{R_1 \leftrightarrow R_3} \begin{pmatrix} 1 & -1 & 3 & 5 \\ 2 & -4 & 2 & 6 \\ 0 & -3 & -5 & 1 \end{pmatrix} \xrightarrow{R_2 \leftrightarrow R_2 - 2R_1} \begin{pmatrix} 1 & -1 & 3 & 5 \\ 0 & -2 & -4 & -4 \\ 0 & -3 & -5 & 1 \end{pmatrix}$$

$$\xrightarrow{R_2 \leftrightarrow -\frac{1}{2}R_2} \begin{pmatrix} 1 & -1 & 3 & 5 \\ 0 & 1 & 2 & 2 \\ 0 & -3 & -5 & 1 \end{pmatrix} \xrightarrow{R_3 \leftrightarrow R_3 + 3R_2} \begin{pmatrix} 1 & -1 & 3 & 5 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 1 & 7 \end{pmatrix}$$

Interpretando M come la matrice aumentata di un sistema lineare, otteniamo:

$$\begin{cases} 3x_2 - 5x_3 = 1 \\ 2x_1 - 4x_2 + 2x_3 = 6 \\ x_1 - x_2 + 3x_3 = 5 \end{cases} \text{ è equivalente a } \begin{cases} x_1 - x_2 + 3x_3 = 5 \\ x_2 + 2x_3 = 2 \\ x_3 = 7 \end{cases}$$

L'**algoritmo di Gauss-Jordan** è un perfezionamento del metodo di eliminazione di Gauss. Una volta portata la matrice in forma a scala, si eseguono ulteriori operazioni elementari sulle righe per "*ripulire*" le colonne dei pivot, da destra verso sinistra. Più precisamente, per ogni pivot applichiamo la mossa 2. per trasformare ogni pivot in 1, e lo stesso metodo del Passo 3. del MEG per trasformare in zeri le entrate "sopra" il pivot.

Il risultato finale sarà una *in forma a scala ridotta*, cioè una matrice a scala tale che ogni pivot è uguale a 1 e ogni pivot è l'unico elemento diverso da zero nella sua *colonna*.

Vediamo due esempi ($\mathbb{K} = \mathbb{R}$).

Esempio. i) Continuiamo l'esempio precedente:

$$\left(\begin{array}{ccc|c} 1 & -1 & 3 & 5 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 1 & 7 \end{array} \right) \xrightarrow{\begin{matrix} R_1 \rightarrow R_1 - 3R_3 \\ R_2 \rightarrow R_2 - 2R_3 \end{matrix}} \left(\begin{array}{ccc|c} 1 & -1 & 0 & -16 \\ 0 & 1 & 0 & -12 \\ 0 & 0 & 1 & 7 \end{array} \right) \xrightarrow{R_1 \rightarrow R_1 + R_3} \left(\begin{array}{ccc|c} 1 & 0 & 0 & -28 \\ 0 & 1 & 0 & -12 \\ 0 & 0 & 1 & 7 \end{array} \right)$$

Interpretando la matrice iniziale come la matrice completa di un sistema lineare, dalla forma finale si vede subito che l'insieme delle soluzioni è $\{(-28, -12, 7)^T\}$.

ii) Consideriamo il sistema lineare

$$\begin{cases} 2x_1 - 6x_2 + 2x_3 + 2x_4 = 6 \\ -3x_1 + 9x_2 + \quad \quad + 3x_4 = -9 \end{cases}.$$

La sua matrice completa si riduce a:

$$\left(\begin{array}{cccc|c} 2 & -6 & 2 & 2 & 6 \\ -3 & 9 & 0 & 3 & -9 \end{array} \right) \xrightarrow{\begin{matrix} R_1 \rightarrow \frac{1}{2}R_1 \\ R_2 \rightarrow \frac{1}{3}R_2 \end{matrix}} \left(\begin{array}{cccc|c} 1 & -3 & 1 & 1 & 3 \\ -1 & 3 & 0 & 1 & -3 \end{array} \right) \xrightarrow{R_2 \rightarrow R_2 + R_1} \left(\begin{array}{cccc|c} 1 & -3 & 1 & 1 & 3 \\ 0 & 0 & 1 & 2 & 0 \end{array} \right)$$

$$\xrightarrow{R_1 \rightarrow R_1 - R_2} \left(\begin{array}{cccc|c} 1 & -3 & 0 & -1 & 3 \\ 0 & 0 & 1 & 2 & 0 \end{array} \right) \text{ che corrisponde a } \begin{cases} x_1 - 3x_2 - x_4 = 3 \\ x_3 + 2x_4 = 0 \end{cases}$$

Definizione 1.9. Le variabili corrispondenti ai pivot sono chiamate *variabili dipendenti*; le restanti sono chiamate *variabili indipendenti* o *variabili libere*.

Nel sistema lineare

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

le variabili x_2, x_4 sono variabili libere, mentre x_1, x_3 sono variabili dipendenti e possono essere espresse in termini delle variabili libere: $x_1 = 3 + 3x_2 + x_4$ (1^a equazione), $x_3 = -2x_4$ (2^a equazione):

$$\text{Sol} = \left\{ \begin{pmatrix} 3 + 3s + t \\ s \\ -2t \\ t \end{pmatrix} \mid t, s \in \mathbb{R} \right\} = \begin{pmatrix} 3 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ -2 \\ 1 \end{pmatrix} \mathbb{K} + \begin{pmatrix} 3 \\ 1 \\ 0 \\ 0 \end{pmatrix} \mathbb{K}$$

Osservazione. L'algoritmo di Gauss(-Jordan) permette di risolvere un sistema lineare esprimendo le variabili dipendenti in termini delle variabili libere.

Definizione 1.10. Sia $A \in \mathcal{M}_{\mathbb{K}}(m, n)$ una matrice e sia $U \in \mathcal{M}_{\mathbb{K}}(m, n)$ una riduzione a scala di A . Il *rango* di A è il numero di pivot di U e viene indicato con $\text{rg}(A)$.

Si osservi che il numero di pivot di U è pari al numero di righe non-nulle di U .

Esempio. La matrice nulla $O \in \mathcal{M}_{\mathbb{K}}(m, n)$ ha rango $\text{rg}(O) = 0$. La matrice identità $I_n \in \mathcal{M}_{\mathbb{K}}(n)$ ha rango $\text{rg}(I_n) = n$.

La matrice $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{pmatrix}$ si riduce a $A' = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, quindi $\text{rg}(A) = 1$.

La matrice $B = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ si riduce a $B' = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$, quindi $\text{rg}(B) = 2$.

Osservazione. i) Esistono diverse possibilità per portare una matrice A in forma a scala. Più avanti nel corso dimostreremo che *tutte le riduzioni a scala di A hanno lo stesso numero di pivot*. In altre parole, il rango è indipendente dalla riduzione a scala.

ii) Le operazioni elementari sulle righe preservano il rango.

iii) Poiché i pivot si trovano in colonne e righe diverse, per $A \in \mathcal{M}_{\mathbb{K}}(m, n)$, vale $0 \leq \text{rg}(A) \leq m$ e $0 \leq \text{rg}(A) \leq n$, ovvero $0 \leq \text{rg}(A) \leq \min\{m, n\}$.

1.2.3 Il Teorema di Rouché-Capelli e il Teorema di struttura per le soluzioni di un sistema lineare

Esempio. Consideriamo il sistema lineare

$$(A|b) = \left(\begin{array}{cccc|c} 1 & 0 & 2 & 3 & 4 \\ 0 & 1 & 2 & -1 & 5 \\ 0 & -1 & -2 & 1 & \alpha \end{array} \right) \quad \text{dove } \alpha \in \mathbb{K} \text{ è un parametro.}$$

Applicando la mossa $R_3 \rightarrow R_3 + R_2$, otteniamo la matrice

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

Se $5 + \alpha = 0$, l'ultima equazione diventa “ $0 = 0$ ” e abbiamo delle soluzioni: $x_1 = 4 - 2x_3 - 3x_4$, $x_2 = 5 - 2x_3 + x_4$:

$$\text{Sol}(A|b) = \left\{ \begin{pmatrix} 4 - 2s - 3t \\ 5 - 2s + t \\ s \\ t \end{pmatrix} \mid s, t \in \mathbb{K} \right\} = \left\{ \begin{pmatrix} 4 \\ 5 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} -2 \\ -2 \\ 1 \\ 0 \end{pmatrix} s + \begin{pmatrix} -3 \\ 1 \\ 0 \\ 1 \end{pmatrix} t \mid s, t \in \mathbb{K} \right\}$$

Se $5 + \alpha \neq 0$, l'ultima equazione ha la forma “ $0 = 1$ ”, quindi non esiste alcuna soluzione.

Teorema 1.11 (Teorema di Rouché-Capelli). *Sia $Ax = b$ un sistema lineare con $A \in \mathcal{M}_{\mathbb{K}}(m, n)$ e $b \in \mathcal{M}_{\mathbb{K}}(m, 1)$.*

1. *Il sistema lineare $Ax = b$ è risolubile se e solo se $\text{rg}(A) = \text{rg}(A|b)$.*
2. *Se $\text{rg}(A) = \text{rg}(A|b) = r$, allora esistono $n - r + 1$ vettori colonna $v, w_1, \dots, w_{n-r} \in \mathcal{M}_{\mathbb{K}}(n, 1)$, tali che $\text{Sol}(A|b) = \{v + t_1 w_1 + \dots + t_{n-r} w_{n-r} : t_i \in \mathbb{K}\}$.*
3. *In particolare, $Ax = b$ ha un'unica soluzione se e solo se $\text{rg}(A) = \text{rg}(A|b) = n$*

Dimostrazione. Applicando il MEG a $(A|b)$ otteniamo una matrice $(U|b')$ in forma a scala tale che $\text{Sol}(A|b) = \text{Sol}(U|b')$. Ci sono 2 possibilità per $(U|b')$: o tutti i pivot sono a sinistra (1), oppure c'è un pivot a destra (2):

$$(1) \left(\begin{array}{cccc|ccc} 0 & \dots & 0 & p_1 & & & \\ 0 & \dots & 0 & 0 & p_2 & & \\ 0 & \dots & 0 & 0 & 0 & \ddots & \\ 0 & \dots & 0 & 0 & 0 & \dots & p_r & \dots \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \end{array} \right) \quad (2) \left(\begin{array}{cccc|cccc} 0 & \dots & 0 & p_1 & & & & \\ 0 & \dots & 0 & 0 & p_2 & & & \\ 0 & \dots & 0 & 0 & 0 & \ddots & & \\ 0 & \dots & 0 & 0 & 0 & \dots & p_r & \dots \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \end{array} \right)$$

dove i p_j sono i pivot, in particolare $p_j \neq 0$.

Nel caso (2), la riga $(r + 1)$ -esima richiede “ $0 = 1$ ”, quindi il sistema lineare non è risolubile e $\text{rg}(A|b) = \text{rg}(A) + 1$: $\text{rg}(A) < \text{rg}(A|b)$.

Nel caso (1), $\text{rg}(A) = \text{rg}(A|b)$ e nessuna riga è della forma “ $0 = 1$ ”, quindi il sistema lineare è risolubile mediante sostituzione: ci sono $r = \text{rg}(A) = \text{rg}(A|b)$ equazioni diverse da zero, che ci consentono di esprimere le (r) variabili dipendenti (corrispondenti ai pivot) in termini delle $n - r$ variabili indipendenti, quindi le soluzioni dipendono da $n - r$ parametri. $\square$

Esempio. Si consideri il sistema lineare $Ax = b$

$$\left(\begin{array}{cccc|c} -2 & 2 & -1 & 3 & -3 \\ 3 & -3 & 6 & 0 & 0 \\ 1 & -1 & 5 & 3 & -3 \end{array} \right)$$

Dopo aver applicato l'algoritmo di Gauss-Jordan, otteniamo il sistema lineare equivalente

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

Quindi $\text{rg}(A) = \text{rg}(A|b) = 2$ e $n = 4$: ci sono 2 variabili dipendenti x_1, x_3 e 2 variabili libere: x_2 e x_4 :

$$\text{Sol}(A|b) = \left\{ \begin{pmatrix} 2+s+2t \\ s \\ -1-t \\ t \end{pmatrix} \mid s, t \in \mathbb{K} \right\} = \left\{ \begin{pmatrix} 2 \\ 0 \\ -1 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} s + \begin{pmatrix} 2 \\ 0 \\ -1 \\ 1 \end{pmatrix} t \mid s, t \in \mathbb{K} \right\}$$

Osservazione. a) La parte parametrica della soluzione non dipende dalla colonna dei termini noti, ma solo dalla matrice dei coefficienti. Più precisamente questa è $\text{Sol}(A|0)$, e.g. nell'ultimo esempio vale:

$$\left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} s + \begin{pmatrix} 2 \\ 0 \\ -1 \\ 1 \end{pmatrix} t \mid s, t \in \mathbb{K} \right\} = \text{Sol}(A|0)$$

b) La colonna dei termini noti contribuisce alla parte costante/fissa della soluzione (e.g. $\begin{pmatrix} 2 \\ 0 \\ -1 \\ 0 \end{pmatrix}$ nell'ultimo esempio).

Definizione 1.12. Un sistema lineare della forma $Ax = 0$ (ovvero i termini noti sono tutti pari a zero) è chiamato *sistema lineare omogeneo*.

Osservazione. Un sistema lineare omogeneo ha sempre una soluzione, ovvero $x = (0, \dots, 0)^T$, chiamata soluzione banale.

Teorema 1.13 (Teorema di struttura per le soluzioni di un sistema lineare). *Sia $Ax = b$ un sistema lineare risolubile e sia $v_P \in \text{Sol}(A|b)$ una sua soluzione particolare: $Av_P = b$. Allora*

$$\text{Sol}(A|b) = v_P + \text{Sol}(A|0).$$

In altre parole: per ogni soluzione $v_0 \in \text{Sol}(A|0)$ del sistema omogeneo associato, il vettore colonna $v_P + v_0$ è una soluzione di $\text{Sol}(A|b)$; viceversa ogni soluzione di $Ax = b$ è della forma $v_0 + v_P$, dove $v_0 \in \text{Sol}(A|0)$.

Dimostrazione. Sia $v_0 \in \text{Sol}(A|0)$, allora $A(v_P + v_0) = Av_P + Av_0 = b + 0 = b$, so $v_P + v_0 \in \text{Sol}(A|b)$.

Viceversa, sia v una soluzione di $Ax = b$, allora $v - v_P \in \text{Sol}(A|0)$, infatti $A(v - v_P) = Av - Av_P = b - b = 0$. $\square$

Conseguenze del Teorema di Rouché-Capelli

Concludiamo questa sezione con due conseguenze del Teorema di Rouché-Capelli.

Lemma 1.14. *Sia $A \in \mathcal{M}_{\mathbb{K}}(m, n)$ una matrice con $m < n$. Allora, per ogni $b \in \mathcal{M}_{\mathbb{K}}(m, 1)$, il sistema lineare $Ax = b$ non può avere un'unica soluzione.*

Dimostrazione. Per il teorema di Rouché-Capelli, $Ax = b$ ha un'unica soluzione se e solo se $n = \text{rg}(A|b) = \text{rg}(A)$, ma $\text{rg}(A) \leq \min\{m, n\} = m < n$. $\square$

Teorema 1.15 (Teorema di Cramer). *Sia $A \in \mathcal{M}_{\mathbb{K}}(n)$ una matrice quadrata. Allora $\text{rg}(A) = n$ se e solo se il sistema lineare $Ax = b$ ha un'unica soluzione per una (e quindi ogni) scelta di $b \in \mathcal{M}_{\mathbb{K}}(n, 1)$.*

Dimostrazione. $\Rightarrow$] Supponiamo che $\text{rg}(A) = n$ e consideriamo il sistema lineare $Ax = b$ per $b \in \mathcal{M}_{\mathbb{K}}(n, 1)$ arbitrario. Vogliamo dimostrare che $Ax = b$ ha una soluzione unica.

Consideriamo la matrice completa $(A|b)$ e la riduciamo a scala utilizzando l'algoritmo di Gauss, ottenendo la matrice $(U|b')$. Poiché $\text{rg}(A) = \text{rg}(U) = n$ e U è una matrice quadrata di tipo $n \times n$, i pivot di U appaiono sulla diagonale:

$$(A|b) \xrightarrow{\text{Gauss}} (U|b') = \left(\begin{array}{ccccc|c} p_1 & * & * & * & * & * \\ 0 & p_2 & & & & \\ 0 & 0 & \ddots & & & \\ \vdots & & & \ddots & & \\ 0 & 0 & \dots & & p_n & * \end{array} \right)$$

Quindi $\text{rg}(A|b) = \text{rg}(U|b') = n$ e per il teorema di Rouché-Capelli $Ax = b$ ha un'unica soluzione.

$\Leftarrow$] Supponiamo che il sistema lineare $Ax = b$ abbia un'unica soluzione per un $b \in \mathcal{M}_{\mathbb{K}}(n, 1)$. Allora, per il teorema di Rouché-Capelli vale $\text{rg}(A|b) = \text{rg}(A) = n$. $\square$

Osservazione. Si noti che per verificare se una matrice quadrata $A \in \mathcal{M}_{\mathbb{K}}(n)$ ha $\text{rg}(A) = n$, è sufficiente trovare un $b \in \mathcal{M}_{\mathbb{K}}(n, 1)$ tale che $Ax = b$ abbia un'unica soluzione, ad esempio si verifica se $Ax = 0$ ha un'unica soluzione.

Viceversa, se per un $b \in \mathcal{M}_{\mathbb{K}}(n, 1)$ (e.g. $b = 0$ va bene!) il sistema lineare $Ax = b$ ha più soluzioni, allora $\text{rg}(A) < n$.

1.2.4 Matrice inversa

Per ogni $k \in \mathbb{K}, k \neq 0$, esiste $k' \in \mathbb{K}$ tale che $k \cdot k' = k' \cdot k = 1$, k' è l'inverso di k .

Vale lo stesso per le matrici?

Definizione 1.16. Una matrice quadrata $A \in \mathcal{M}_{\mathbb{K}}(n)$ è *invertibile* se esiste una matrice $B \in \mathcal{M}_{\mathbb{K}}(n)$ tale che $AB = BA = I_n$.

Se esiste una matrice B di questo tipo, essa è unica e viene chiamata *matrice inversa* di A e indicata con A^{-1} .

Osservazione. Se $A \in \mathcal{M}_{\mathbb{K}}(m, n)$ non è una matrice quadrata ($m \neq n$), allora A non è invertibile (questo fatto verrà dimostrato nel proseguimento di queste note).

Esempio. Sia $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \in \mathcal{M}_{\mathbb{R}}(2)$. Vogliamo determinare se A è invertibile, quindi cerchiamo una matrice $X \in \mathcal{M}_{\mathbb{R}}(2)$ tale che $AX = I_2$, e poi verifichiamo che $XA = I_2$:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} x_{1,1} & x_{1,2} \\ x_{2,1} & x_{2,2} \end{pmatrix} = \begin{pmatrix} x_{1,1} + 2x_{2,1} & x_{1,2} + 2x_{2,2} \\ 3x_{1,1} + 4x_{2,1} & 3x_{1,2} + 4x_{2,2} \end{pmatrix} \longleftrightarrow \begin{cases} x_{1,1} + 2x_{2,1} = 1 \\ 3x_{1,1} + 4x_{2,1} = 0 \\ x_{1,2} + 2x_{2,2} = 0 \\ 3x_{1,2} + 4x_{2,2} = 1 \end{cases}$$

Il sistema lineare è equivalente a due sistemi lineari (uno nelle variabili $x_{1,1}, x_{2,1}$ e uno nelle variabili $x_{1,2}, x_{2,2}$), entrambi aventi A come matrice dei coefficienti:

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} x_{1,1} \\ x_{2,1} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} x_{1,2} \\ x_{2,2} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Usiamo l'algoritmo di Gauss-Jordan per risolverli contemporaneamente!

$$(A|I_2) = \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{array} \right) \xrightarrow{R_3 \rightarrow R_3 - 3R_1} \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -2 & -3 & 1 \end{array} \right) \xrightarrow{\begin{matrix} R_1 \rightarrow R_1 + R_2 \\ R_2 \rightarrow -\frac{1}{2}R_2 \end{matrix}} \left(\begin{array}{cc|cc} 1 & 0 & -2 & 1 \\ 0 & 1 & 3/2 & -1/2 \end{array} \right)$$

Quindi $B = \begin{pmatrix} -2 & 1 \\ 3/2 & -1/2 \end{pmatrix}$ soddisfa $AB = I_2$, ed è facile verificare che vale anche $BA = I_2$:
 A è invertibile e $A^{-1} = B$.

Teorema 1.17. Sia $A \in \mathcal{M}_{\mathbb{K}}(n)$ una matrice quadrata. Le seguenti sono equivalenti:

1. A è invertibile;
2. $\text{rg}(A) = n$;
3. per ogni scelta di $b \in \mathcal{M}_{\mathbb{K}}(n, 1)$, il sistema lineare $Ax = b$ ha un'unica soluzione.

Dimostrazione. Dal teorema di Cramer (Teorema 1.15) sappiamo che 2. e 3. sono equivalenti, quindi restano da dimostrare le due implicazioni: $1 \Rightarrow 3$ e $2 \Rightarrow 1$.

1. $\Rightarrow$ 3.] Supponiamo che A sia invertibile e consideriamo il sistema lineare $Ax = b$, per un $b \in \mathcal{M}_{\mathbb{K}}(n, 1)$. Allora $x = A^{-1}b$ è una soluzione: $A(A^{-1}b) = (AA^{-1})b = I_n b = b$; e ogni soluzione ha questa forma: $Ax = b \Rightarrow A^{-1}Ax = A^{-1}b \Rightarrow x = A^{-1}b$. Quindi $x = A^{-1}b$ è l'unica soluzione.

2. $\Rightarrow$ 1.] Il rango di A è n , quindi in base al teorema di Cramer possiamo risolvere qualsiasi sistema lineare $Ax = b$ con $b \in \mathcal{M}_{\mathbb{K}}(n, 1)$, e la soluzione è unica. Consideriamo gli n sistemi lineari:

$$A \begin{pmatrix} x_{1,1} \\ x_{2,1} \\ \vdots \\ x_{n,1} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad A \begin{pmatrix} x_{1,2} \\ x_{2,2} \\ \vdots \\ x_{n,2} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad A \begin{pmatrix} x_{1,n} \\ x_{2,n} \\ \vdots \\ x_{n,n} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

e siano $b_1, b_2, \dots, b_n \in \mathcal{M}_{\mathbb{K}}(n, 1)$ le loro soluzioni. Definiamo $B \in \mathcal{M}_{\mathbb{K}}(n, n)$ come la matrice le cui colonne sono $b_1, b_2, \dots, b_n$. Quindi, per costruzione, B soddisfa $AB = I_n$; e dobbiamo dimostrare che $BA = I_n$.

Consideriamo ora il sistema lineare $Bx = 0$. Applicando A ad entrambi i lati otteniamo $A(Bx) = A(0) = 0 \Rightarrow I_n x = 0 \Rightarrow x = 0$. Quindi $Bx = 0$ ha un'unica soluzione e per l'equivalenza $2. \Leftrightarrow 3.$ (Teorema di Cramer) $\text{rg}(B) = n$.

Ripetendo il ragionamento fatto per A con la matrice B , troviamo una matrice C tale che $BC = I_n = AB$, allora

$$A = AI_n = A(BC) = (AB)C = I_n C = C$$

In altre parole B è invertibile, $A = C$, e quindi anche A è invertibile. $\square$

In pratica, data una matrice quadrata $A \in \mathcal{M}_{\mathbb{K}}(n)$, per determinare se è invertibile e trovare la sua inversa, dobbiamo risolvere n sistemi lineari in n variabili, tutti aventi A come matrice dei coefficienti, quindi è più conveniente risolverli simultaneamente utilizzando l'algoritmo di Gauss-Jordan (come negli esempi). Visivamente, ci sono 2 casi:

i) $(A|I_n) \xrightarrow{\text{Gauss-Jordan}} (I_n|B)$ allora A è invertibile e $A^{-1} = B$

ii) $(A|I_n) \xrightarrow{\text{Gauss-Jordan}} (A'|*)$ dove A' ha una riga nulla; allora A non è invertibile

Esempio. Determiniamo l'inversa (se esiste) delle seguenti matrici:

$$i) A = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 6 \\ 3 & 0 & 2 \end{pmatrix} \in \mathcal{M}_{\mathbb{R}}(3), \quad ii) M = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} \in \mathcal{M}_{\mathbb{R}}(2)$$

$$i) \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 2 & 1 & 6 & 0 & 1 & 0 \\ 3 & 0 & 2 & 0 & 0 & 1 \end{array} \right) \xrightarrow{\begin{array}{l} R_2 \rightarrow R_2 - 2R_1 \\ R_3 \rightarrow R_3 - 3R_1 \end{array}} \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 4 & -2 & 1 & 0 \\ 0 & 0 & -1 & -3 & 0 & 1 \end{array} \right) \xrightarrow{\begin{array}{l} R_1 \rightarrow R_1 + R_3 \\ R_2 \rightarrow R_2 + 4R_3 \\ R_3 \rightarrow -R_3 \end{array}} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -2 & 0 & 1 \\ 0 & 1 & 0 & -14 & 1 & 4 \\ 0 & 0 & 1 & 3 & 0 & -1 \end{array} \right)$$

Quindi $C = \begin{pmatrix} -2 & 0 & 1 \\ -14 & 1 & 4 \\ 3 & 0 & -1 \end{pmatrix}$ soddisfa $AB = I_3 = BA = I_3$: A è invertibile e $A^{-1} = B$.

$$ii) \left(\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 2 & 2 & 0 & 1 \end{array} \right) \xrightarrow{R_2 \rightarrow R_2 - 2R_1} \left(\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 0 & 0 & -2 & 1 \end{array} \right)$$

e non possiamo continuare, perché il sistema lineare non ha soluzioni. Questo significa che M non è invertibile, infatti $\text{rg}(M) = 1 < 2$.

Concludiamo osservando come si comporta l'inverso rispetto al prodotto righe per colonne e alla trasposizione.

Proposizione 1.18. Siano $A, B \in \mathcal{M}_{\mathbb{K}}(n)$ matrici invertibili, allora

i) AB è invertibile e $(AB)^{-1} = B^{-1}A^{-1}$.

ii) A^T è invertibile e $(A^T)^{-1} = (A^{-1})^T$.

Dimostrazione. i) $(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AI_n A^{-1} = AA^{-1} = I_n$.

Analogamente, $(B^{-1}A^{-1})(AB) = I_n$.

ii) $A^T(A^{-1})^T = (A^{-1}A)^T = I_n^T = I_n$.

Analogamente, $(A^{-1})^T A^T = (AA^{-1})^T = I_n^T = I_n$. $\square$

1.3 Determinante

In questa sezione vediamo uno strumento che permette, attraverso semplici calcoli fatti solo da somme e prodotti di scalari, di rispondere ad una sola ma decisiva domanda: quand'è che una matrice quadrata $A \in \mathcal{M}_{\mathbb{K}}(n)$ ha rango massimo $\text{rg}(A) = n$?

Questo sarà utile nel prossimo capitolo per testare la (in-)dipendenza lineare dei vettori.

Definizione 1.19. Il *determinante* è una funzione $\det : \mathcal{M}_{\mathbb{K}}(n) \rightarrow \mathbb{K}$ definita ricorsivamente:

$$(n = 1) \quad \det(a_{1,1}) = a_{1,1};$$

$$(n \geq 2) \quad \text{espansione di Laplace: } \det(A) = \sum_{j=1}^n (-1)^{i+j} a_{i,j} \det(\hat{A}_{i,j}), \text{ dove } i \text{ è un indice riga fissato e } \hat{A}_{i,j} \in \mathcal{M}_{\mathbb{K}}(n-1) \text{ è la matrice ottenuta da } A \text{ rimuovendo la riga } i\text{-esima e la colonna } j\text{-esima.}$$

Esempio. $\mathbb{K} = \mathbb{R}, i = 1$:

$$\det \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = (-1)^{1+1} 1 \det \begin{pmatrix} 5 & 6 \\ 8 & 9 \end{pmatrix} + (-1)^{1+2} 2 \det \begin{pmatrix} 4 & 6 \\ 7 & 9 \end{pmatrix} + (-1)^{1+3} 3 \det \begin{pmatrix} 4 & 5 \\ 7 & 8 \end{pmatrix}$$

Fatto 1.20. *i) Il determinante di una matrice $A \in \mathcal{M}_{\mathbb{K}}(n)$ non dipende dallo sviluppo scelto.*

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

Da ii) segue che alternativamente si può considerare lo sviluppo di Laplace lungo una colonna fissata (la j -esima): $\det(A) = \sum_{i=1}^n (-1)^{i+j} a_{i,j} \det(\hat{A}_{i,j})$, dove (come sopra) $\hat{A}_{i,j} \in \mathcal{M}_{\mathbb{K}}(n-1)$ è la matrice ottenuta da A rimuovendo la riga i -esima e la colonna j -esima.

Esempio. Calcoliamo il determinante di una matrice arbitraria di tipo 2×2 , sviluppando lungo la prima colonna:

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = (-1)^{1+1} a \det(d) + (-1)^{1+2} c \det(b) = ad - bc.$$

Esempio. Calcoliamo il determinante della matrice $A = \begin{pmatrix} 4 & 0 & 3 \\ 1 & 2 & 0 \\ 7 & 3 & 0 \end{pmatrix} \in \mathcal{M}_{\mathbb{R}}(3, 3)$. Sviluppiamo lungo la seconda colonna:

$$\begin{aligned} \det \begin{pmatrix} 4 & 0 & 3 \\ 1 & 2 & 0 \\ 7 & 3 & 0 \end{pmatrix} &= (-1)^{2+1} 0 \det \hat{A}_{2,1} + (-1)^{2+2} 2 \det \hat{A}_{2,2} + (-1)^{2+3} 3 \det \hat{A}_{2,3} \\ &= 2 \det \begin{pmatrix} 4 & 3 \\ 7 & 0 \end{pmatrix} - 3 \det \begin{pmatrix} 4 & 3 \\ 1 & 0 \end{pmatrix} = 2(4 \cdot 0 - 7 \cdot 3) - 3(4 \cdot 0 - 1 \cdot 3) = -42 + 9 = -33 \end{aligned}$$

Osservazione. Se $a_{i,j} = 0$, allora non è necessario calcolare il determinante della matrice $\hat{A}_{i,j}$, quindi è opportuno selezionare la riga/colonna con il numero più alto di zeri.

Esempio. Espandiamo la matrice A lungo la terza colonna:

$$\det \begin{pmatrix} 4 & 0 & 3 \\ 1 & 2 & 0 \\ 7 & 3 & 0 \end{pmatrix} = (-1)^{3+1} 3 \det \hat{A}_{3,1} + (-1)^{3+2} 0 \det \hat{A}_{3,2} + (-1)^{3+3} 0 \det \hat{A}_{3,3} = 3 \det \begin{pmatrix} 1 & 2 \\ 7 & 3 \end{pmatrix} = -33$$

Vediamo ora alcuni casi particolari:

Definizione 1.21. A è una matrice *triangolare superiore* se $a_{ij} = 0$ per ogni $j < i$.

A è una matrice *triangolare inferiore* se $a_{ij} = 0$ per ogni $j > i$.

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

Esempio. $\begin{pmatrix} 0 & 1 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix}$ è triangolare superiore, $\begin{pmatrix} 1 & 0 & 0 \\ 2 & 0 & 0 \\ 4 & 0 & -1 \end{pmatrix}$ è triangolare inferiore $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ è diagonale. La matrice identità I_n è una matrice diagonale e ogni matrice a scala è triangolare superiore.

Lemma 1.22. Sia $A \in \mathcal{M}_{\mathbb{K}}(n)$ matrice triangolare. Allora $\det A = a_{1,1} \cdot a_{2,2} \cdots a_{n,n}$; in particolare, $\det(I_n) = 1$.

Dimostrazione. Sviluppando ripetutamente lungo la prima colonna/riga otteniamo $\det A = a_{1,1} \det \hat{A}_{1,1} = a_{1,1} \cdot a_{2,2} \cdots a_{n,n}$. $\square$

Teorema 1.23 (Formula di Binet).

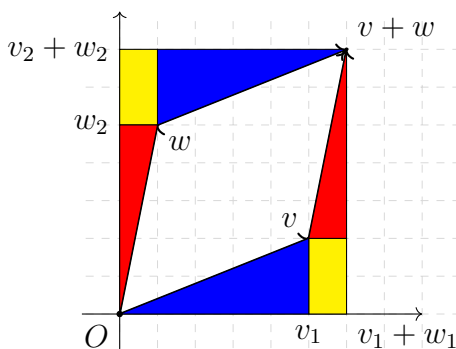
Siano $A, B \in \mathcal{M}_{\mathbb{K}}(n)$ matrici quadrate, allora $\det(AB) = \det(A) \cdot \det(B)$.

Corollario 1.24. Sia $A \in \mathcal{M}_{\mathbb{K}}(n)$ una matrice invertibile, allora $\det(A^{-1}) = \frac{1}{\det A}$.

Dimostrazione. $1 = \det(I_n) = \det(AA^{-1}) = \det(A) \cdot \det(A^{-1})$, so $\det(A^{-1}) = \frac{1}{\det A}$. $\square$

1.3.1 Interpretazione geometrica del determinante

Siano $v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ e $w = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$ due vettori liberi di $\mathbb{R}^2$. Consideriamo il parallelogramma definito da v e w e ne calcoliamo l'area (vedi Figura).



Il rettangolo di vertici opposti O e $v + w$ ha area $(v_1 + w_1)(v_2 + w_2)$ e può essere diviso in 7 parti (figura): il parallelogramma (bianco); 2 rettangoli congruenti (in giallo, di area $w_1 v_2$); e 2 coppie di triangoli congruenti (in rosso e blu, di area $w_1 w_2 / 2$ e $v_1 v_2 / 2$).

Quindi l'area del parallelogramma bianco è data da

$$\begin{aligned} \text{Area} &= (v_1 + w_1)(v_2 + w_2) - 2(w_1 v_2) - 2(w_1 w_2 / 2) - 2(v_1 v_2 / 2) \\ &= (v_1 v_2 + w_1 v_2 + v_1 w_2 + w_1 w_2) - (v_1 v_2 + 2w_1 v_2 + w_1 w_2) \\ &= v_1 w_2 - w_1 v_2 = \det \begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \end{pmatrix} = -\det \begin{pmatrix} w_1 & v_1 \\ w_2 & v_2 \end{pmatrix} \end{aligned}$$

Quindi il determinante della matrice $A \in \mathcal{M}_{\mathbb{R}}(2)$ di colonne v e w è l'area orientata (cioè con segno) del parallelogramma: se per andare da v a w si ruota in senso antiorario (come in figura) allora $\det A = \text{Area}$, se invece si ruota in senso orario, allora $\det A = -\text{Area}$, infatti $\det \begin{pmatrix} w_1 & v_1 \\ w_2 & v_2 \end{pmatrix} = -\text{Area}$ perchè (in figura) per andare da w a v si ruota in senso orario.

Analogamente, per $A \in \mathcal{M}_{\mathbb{R}}(3)$, il suo determinante $\det(A)$ è il volume con segno del parallelepipedo ρ di lati le colonne c_1, c_2, c_3 di A . Più precisamente $\det A = \text{Vol}(\rho)$ se i vettori c_1, c_2, c_3 (in quest'ordine!) sono orientati secondo la regola della mano destra; altrimenti vale $\det A = -\text{Vol}(\rho)$. In generale abbiamo la seguente

Proprietà. Sia $A \in \mathcal{M}_{\mathbb{R}}(n)$ e siano $c_1, \dots, c_n \in \mathbb{R}^n$ le sue colonne. Allora il determinante di A è il volume orientato del politopo di lati $c_1, \dots, c_n \in \mathbb{R}^n$.

1.3.2 Determinante e operazioni elementari sulle righe

Fatto 1.25. Sia $A \in \mathcal{M}_{\mathbb{K}}(n)$ una matrice quadrata, e siano $C_1, \dots, C_n \in \mathcal{M}_{\mathbb{K}}(n, 1)$ le sue colonne: $A = (C_1, \dots, C_n)$. Allora valgono le seguenti proprietà.

1. Il determinante è multilineare, cioè: tenendo fisse tutte le colonne meno una, la funzione determinante $\det$ è lineare (vedi Capitolo 3), ovvero additiva, e rispetta il prodotto per scalare; cioè per ogni $i = 1, \dots, n$, $D \in \mathcal{M}_{\mathbb{K}}(n, 1)$ e $\lambda \in \mathbb{K}$ valgono:

$$a) \det(C_1, \dots, C_{i-1}, C_i + D, C_{i+1}, \dots, C_n) = \det(C_1, \dots, C_{i-1}, C_i, C_{i+1}, \dots, C_n) + \det(C_1, \dots, C_{i-1}, D, C_{i+1}, \dots, C_n)$$

$$b) \det(C_1, \dots, C_{i-1}, \lambda C_i, C_{i+1}, \dots, C_n) = \lambda \det(C_1, \dots, C_{i-1}, C_i, C_{i+1}, \dots, C_n)$$

2. Se la matrice ha due colonne uguali allora $\det A = 0$, cioè se esistono $i < j \in \{1, \dots, n\}$ tali che $C_i = C_j$ allora

$$\det(C_1, \dots, C_i, \dots, C_j = C_i, \dots, C_n) = 0$$

Dal fatto che $\det(A) = \det(A^T)$, e quindi che quanto detto per le colonne vale anche per le righe, otteniamo le seguenti proprietà.

Proprietà. Sia $A \in \mathcal{M}_{\mathbb{K}}(n)$ una matrice quadrata. Le operazioni elementari sulle righe influiscono sul determinante come segue:

- scambiare due righe ($i \neq j$): $A \xrightarrow{R_i \leftrightarrow R_j} B : \det B = -\det A$.
- moltiplicare una riga per uno scalare $k \in \mathbb{K}$: $A \xrightarrow{R_i \rightarrow kR_i} B : \det B = k \det A$.
- sommare ad una riga un multiplo di un'altra riga: $A \xrightarrow{R_i \rightarrow R_i + kR_j} B : \det B = \det A$.

La prima proprietà dice che il determinante è *alternante*.

Dimostrazione. Visto che $\det(A) = \det(A^T)$, dimostriamo le analoghe proprietà sulle colonne, usando la stessa notazione del Fatto 1.25.

La prima proprietà segue dalla seguente catena di uguaglianze, in cui vengono usate 1.a) e 2. del Fatto 1.25:

$$\begin{aligned} 0 &= \det(C_1, \dots, C_i + C_j, \dots, C_i + C_j, \dots, C_n) \\ &= \det(C_1, \dots, C_i, \dots, C_i + C_j, \dots, C_n) + \det(C_1, \dots, C_j, \dots, C_i + C_j, \dots, C_n) \\ &= \det(C_1, \dots, C_i, \dots, C_i, \dots, C_n) + \det(C_1, \dots, C_i, \dots, C_j, \dots, C_n) + \\ &\quad \det(C_1, \dots, C_j, \dots, C_i, \dots, C_n) + \det(C_1, \dots, C_j, \dots, C_j, \dots, C_n) \\ &= 0 + \det(C_1, \dots, C_i, \dots, C_j, \dots, C_n) + \det(C_1, \dots, C_j, \dots, C_i, \dots, C_n) + 0 \end{aligned}$$

La seconda proprietà è 1.b) del Fatto 1.25.

La terza proprietà segue dalla seguente catena di uguaglianze, in cui viene usata 1. del Fatto 1.25:

$$\begin{aligned}\det(C_1, \dots, C_i + kC_j, \dots, C_j, \dots, C_n) &= \det(C_1, \dots, C_i, \dots, C_j, \dots, C_n) + \det(C_1, \dots, kC_j, \dots, C_j, \dots, C_n) \\ &= \det(C_1, \dots, C_i, \dots, C_j, \dots, C_n) + k \det(C_1, \dots, C_j, \dots, C_j, \dots, C_n) \\ &= \det(C_1, \dots, C_i, \dots, C_j, \dots, C_n) + 0 \quad \square\end{aligned}$$

Osservazione. Si noti che dalla seconda proprietà segue che per ogni $A \in \mathcal{M}_{\mathbb{K}}(n)$ e $\lambda \in \mathbb{R}$ vale $\det(\lambda A) = \lambda^n \det(A)$.

Esempio. Siano $A, B \in \mathcal{M}_{\mathbb{Q}}(3)$ matrici invertibili con $\det A = -2$, $\det B = -3$. Allora

$$\det\left(\frac{1}{2}B^{-6}A^2B^5\right) = \left(\frac{1}{2}\right)^3 \frac{(\det A)^2}{\det B} = -\frac{1}{6}.$$

Esempio. Calcoliamo il determinante della matrice $A = \begin{pmatrix} 0 & 0 & -3 & 5 \\ 1 & 8 & 6 & 7 \\ 2 & 0 & 4 & 7 \\ 3 & 0 & 6 & 9 \end{pmatrix} \in \mathcal{M}_{\mathbb{R}}(4)$. Iniziamo sviluppando lungo la seconda colonna:

$$\begin{aligned}\det A &= 8 \det \begin{pmatrix} 0 & -3 & 5 \\ 2 & 4 & 7 \\ 3 & 6 & 9 \end{pmatrix} \xrightarrow{R_3 \rightarrow \frac{1}{3}R_3} 24 \det \begin{pmatrix} 0 & -3 & 5 \\ 2 & 4 & 7 \\ 1 & 2 & 3 \end{pmatrix} \xrightarrow{R_2 \rightarrow R_2 - 2R_3} 24 \det \begin{pmatrix} 0 & -3 & 5 \\ 0 & 0 & 1 \\ 1 & 2 & 3 \end{pmatrix} \\ &= -24 \det \begin{pmatrix} 0 & -3 \\ 1 & 2 \end{pmatrix} = -72\end{aligned}$$

Concludiamo questo capitolo dando una caratterizzazione delle matrici invertibili in termini di determinante.

Teorema 1.26 (Caratterizzazione delle matrici invertibili). *Sia $A \in \mathcal{M}_{\mathbb{K}}(n)$ una matrice quadrata. Le seguenti sono equivalenti:*

1. A è invertibile;
2. $\text{rg}(A) = n$;
3. per ogni scelta di $b \in \mathcal{M}_{\mathbb{K}}(n, 1)$, il sistema lineare $Ax = b$ ha un'unica soluzione.
4. $\det(A) \neq 0$.

Dimostrazione. Dal Teorema 1.17, sappiamo che $1. \Leftrightarrow 2. \Leftrightarrow 3.$, quindi ci resta solo da dimostrare che $1. \Leftrightarrow 4.$.

1. $\Rightarrow$ 4.] Come nella dimostrazione del Corollario 1.24, se A è invertibile, allora $\det(A) \cdot \det(A^{-1}) = 1$, quindi $\det(A) \neq 0$.

4. $\Rightarrow$ 1.] Supponiamo ora che $\det(A) \neq 0$. Applicando il MEG trasformiamo A in una matrice U ridotta a scala con $\det(U) = c \det(A)$, con $c \in \mathbb{K}$, $c \neq 0$, quindi $\det U \neq 0$. Essendo ridotta a scala, U è anche triangolare superiore: $0 \neq \det(U) = u_{1,1} \cdots u_{n,n}$, quindi i pivot appaiono sulla diagonale (si ricordi che i pivot sono sempre in posizione $u_{i,j}$ con $j \geq i$):

$$U = \begin{pmatrix} u_{1,1} & * & * & \cdots & * \\ 0 & u_{2,2} & * & \cdots & * \\ 0 & 0 & \ddots & & \\ \vdots & \vdots & & \ddots & \\ 0 & 0 & \cdots & & u_{n,n} \end{pmatrix}$$

Otteniamo che $n = \text{rg}(U) = \text{rg}(A)$, quindi A è invertibile. $\square$

Più in generale vale il seguente teorema, che lega il rango di una matrice (non necessariamente quadrata!) con i determinanti delle sue sottomatrici quadrate.

Non dimostreremo questo risultato, ma invitiamo il lettore a dimostrarlo lui stesso, dopo aver letto il prossimo capitolo.

Notazione. Sia $A \in \mathcal{M}_{\mathbb{K}}(m, n)$ una matrice, e siano $I \subseteq \{1, \dots, m\}$ e $J \subseteq \{1, \dots, n\}$ due insiemi di cardinalità k (cioè stiamo scegliendo k righe e k colonne di A). Denotiamo con $A_{IJ} \in \mathcal{M}_{\mathbb{K}}(k)$ la sottomatrice ottenuta considerando solo gli elementi su queste k righe e k colonne:

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & -1 \\ 2 & 5 & 8 & 7 \end{pmatrix} \xrightarrow{I = \{1, 2\} \atop J = \{2, 4\}} A_{IJ} = \begin{pmatrix} 2 & 4 \\ 1 & -1 \end{pmatrix}$$

Teorema 1.27 (Teorema degli orlati). *Sia $A \in \mathcal{M}_{\mathbb{K}}(m, n)$, allora $\text{rg}(A) = r$ se e solo se:*

- *esistono $I \subseteq \{1, \dots, m\}$ e $J \subseteq \{1, \dots, n\}$ due insiemi di cardinalità r tali che $\det A_{IJ} \neq 0$.*
- *una qualunque scelta di $r+1$ righe $I' \subseteq \{1, \dots, m\}$, e $p+1$ colonne $J' \subseteq \{1, \dots, n\}$, con $I \subset I'$ e $J \subset J'$, soddisfa $\det A_{I'J'} = 0$.*

Nel secondo punto stiamo aggiungendo un “orlo” alla matrice A_{IJ} , la stiamo “orlando” aggiungendo una riga ed una colonna, da qui il nome del teorema.

Esempio. Nell'esempio precedente la matrice A ha rango $\text{rg}(A) = 2$, infatti la terza riga si ottiene sommando alla seconda riga il doppio della prima: $R_3 = 2R_1 + R_2$.

Verifichiamolo usando il criterio degli orlati: $\det \begin{pmatrix} 2 & 4 \\ 1 & -1 \end{pmatrix} = -6 \neq 0$, mentre i determinanti delle due matrici che orlano A_{IJ} sono entrambi nulli:

$$\begin{pmatrix} 1 & 2 & 4 \\ 0 & 1 & -1 \\ 2 & 5 & 7 \end{pmatrix} = 1 \cdot 12 + 2 \cdot (-6) = 0, \quad \begin{pmatrix} 2 & 3 & 4 \\ 1 & 2 & -1 \\ 5 & 8 & 7 \end{pmatrix} = 2 \cdot (22) - 1 \cdot (-11) + 5 \cdot (-11) = 0.$$

Corollario 1.28. *Sia $A \in \mathcal{M}_{\mathbb{K}}(m, n)$. Se esistono $I \subseteq \{1, \dots, m\}$ e $J \subseteq \{1, \dots, n\}$ due insiemi di cardinalità k tali che $\det A_{IJ} \neq 0$, allora $\text{rg}(A) \geq k$.*

Capitolo 2

Spazi vettoriali

Nel primo capitolo abbiamo discusso l'aspetto algebrico degli oggetti che abbiamo introdotto: matrici, sistemi lineari, soluzioni e determinanti.

In questo capitolo inizieremo a considerarli da un punto di vista geometrico che generalizza ciò che abbiamo visto nell'Introduzione e spiegherà alcune scelte fatte nel capitolo precedente.

2.1 Spazi vettoriali

Definizione 2.1. Uno *spazio vettoriale* su un campo $\mathbb{K}$ è un insieme V dotato di due operazioni interne: una *somma* $+: V \times V \rightarrow V$ e un *prodotto per scalare* $\cdot: \mathbb{K} \times V \rightarrow V$ che soddisfano

SV1) $\forall u, v, w \in V: (u + v) + w = u + (v + w)$ ($+$ è associativa)

SV2) $\forall v, w \in V: v + w = w + v$ ($+$ è commutativa)

SV3) $\exists 0 \in V: v + 0 = 0 + v = v \forall v \in V$ (il *vettore nullo* è l'elemento neutro di $+$)

SV4) $\forall v \in V, \exists v' \in V: v + v' = v' + v = 0$ (esistenza dell'opposto)

SV5) $\forall c, d \in \mathbb{K}, v \in V: (c \cdot d) \cdot v = c \cdot (d \cdot v)$ ($\cdot$ è associativo)

SV6) $\forall v \in V: 1 \cdot v = v$ ($1 \in \mathbb{K}$ è l'elemento neutro di $\cdot$)

SV7) $\forall c, d \in \mathbb{K}, v \in V: (c + d) \cdot v = c \cdot v + d \cdot v$ (distributività)

SV8) $\forall c \in \mathbb{K}, v, w \in V: c \cdot (v + w) = c \cdot v + c \cdot w$ (distributività)

Gli elementi di uno spazio vettoriale sono chiamati *vettori*.

Notazione. A volte si dice anche che V è un $\mathbb{K}$ -spazio vettoriale.

Esempio.

- I vettori liberi in $\mathbb{R}^2$ e $\mathbb{R}^3$ formano uno spazio vettoriale su $\mathbb{R}$ (con le operazioni viste nell'Introduzione).
- $\mathcal{M}_{\mathbb{K}}(m, n)$ è uno spazio vettoriale su $\mathbb{K}$ (con le operazioni definite nel Capitolo 1); in particolare lo è $\mathcal{M}_{\mathbb{K}}(m, 1) = \mathbb{K}^m$.
- L'insieme delle funzioni reali continue $\mathcal{C}^0(\mathbb{R})$ con le operazioni $(f + g)(x) = f(x) + g(x)$, $(\lambda f)(x) = \lambda f(x)$, è uno spazio vettoriale su $\mathbb{R}$.

- L'insieme dei polinomi $\mathbb{K}[t] = \{\sum_{k=0}^n a_k t^k\}$ a coefficienti in $\mathbb{K}$ è uno spazio vettoriale su $\mathbb{K}$ rispetto alle operazioni usuali:

$$\left(\sum_{k=0}^{n_1} a_k t^k\right) + \left(\sum_{k=0}^{n_2} b_k t^k\right) = \sum_{k=0}^{\max(n_1, n_2)} (a_k + b_k) t^k, \quad \lambda \left(\sum_{k=0}^n a_k t^k\right) = \sum_{k=0}^n (\lambda a_k) t^k$$

- L'insieme delle soluzioni di un sistema lineare omogeneo $\text{Sol}(A|0)$, $A \in \mathcal{M}_{\mathbb{K}}(m, n)$ è un $\mathbb{K}$ -spazio vettoriale rispetto alle operazioni di somma e prodotto per scalare per le matrici: siano $v_1, v_2 \in \text{Sol}(A|0)$ e sia $\lambda \in \mathbb{K}$, allora

$$A(v_1 + v_2) = Av_1 + Av_2 = 0 + 0 = 0 \quad A(\lambda v_1) = \lambda Av_1 = \lambda 0 = 0$$

Notazione. Ci sono due zeri in circolazione: lo scalare zero e il vettore nullo. Se non è chiaro dal contesto, risolveremo questa ambiguità indicando con $0_{\mathbb{K}}$ lo scalare zero e con 0_V il vettore nullo di V .

Osservazione. i) Il vettore nullo è unico. Siano $0, 0' \in V$ due vettori che soddisfano SV3, allora $0 = 0 + 0' = 0'$.

ii) L'opposto di $v \in V$ è unico. Siano v_1, v_2 due vettori che soddisfano SV4: $v + v_1 = v + v_2 = 0$, allora $v_1 = v_1 + 0 = v_1 + (v + v_2) = (v_1 + v) + v_2 = 0 + v_2 = v_2$.

iii) $0_{\mathbb{K}} \cdot v = 0_V$. Infatti, $0_{\mathbb{K}} \cdot v = (0_{\mathbb{K}} + 0_{\mathbb{K}}) \cdot v = 0_{\mathbb{K}} \cdot v + 0_{\mathbb{K}} \cdot v$, da cui segue $0_V = 0_{\mathbb{K}} \cdot v$.

iv) $\lambda 0_V = 0_V$. Infatti, $\lambda 0_V = \lambda(0_V + 0_V) = \lambda 0_V + \lambda 0_V$, da cui segue $\lambda 0_V = 0_V$.

v) Vale la legge di annullamento del prodotto: $\lambda \cdot v = 0_V \Rightarrow \lambda = 0_{\mathbb{K}}$ o $v = 0_V$. Sia $\lambda \neq 0$, allora $0_V = \lambda^{-1} 0_V = \lambda^{-1}(\lambda v) = 1 \cdot v = v$.

vi) L'opposto di $v \in V$ è $v' = (-1) \cdot v = -v$. Infatti, $(v + (-1)v) \stackrel{SV7}{=} (1 + (-1))v = 0_{\mathbb{K}} \cdot v = 0_V$.

2.2 Sottospazi vettoriali

Definizione 2.2. Sia V uno spazio vettoriale sul campo $\mathbb{K}$. Un sottoinsieme $W \subset V$ è un *sottospazio vettoriale* (e si scrive $W \triangleleft V$) se valgono le seguenti tre condizioni:

S1) $0_V \in W$;

S2) $\forall w_1, w_2 \in W$ vale $w_1 + w_2 \in W$ (W è chiuso rispetto alla somma);

S3) $\forall w \in W$ e $\lambda \in \mathbb{K}$ vale $\lambda \cdot w \in W$ (W è chiuso rispetto al prodotto per scalare).

Osservazione. Queste tre condizioni ci assicurano che W è uno spazio vettoriale su $\mathbb{K}$ con le operazioni ereditate/indotte da V .

Esempio.

- $W = \{0_V\} \triangleleft V$;
- $V \triangleleft V$.
- L'insieme delle funzioni differenziabili reali $\mathcal{C}^1(\mathbb{R})$ è un sottospazio vettoriale di $\mathcal{C}^0(\mathbb{R})$.
- L'insieme dei polinomi $\mathbb{K}[t]_{\leq d} = \{p(t) \in \mathbb{K}[t] \mid \deg(p) \leq d\}$ di grado $\leq d$ è un sottospazio vettoriale di $\mathbb{K}[t]$. Più precisamente:

$$\mathbb{K}[t]_{\leq d} \triangleleft \mathbb{K}[t]_{\leq d+n} \triangleleft \mathbb{K}[t], \quad \text{for } k \in \mathbb{N}.$$

- $\{(x, y, z)^T \in \mathbb{R}^2 \mid x \geq 0, y \geq 0\}$ e $\{(x, y, z)^T \in \mathbb{R}^2 \mid xy \geq 0\}$ non sono sottospazi vettoriali di $\mathbb{R}^2$.
- I sottospazi vettoriali di $\mathbb{R}^2$ sono: il sottospazio banale $\{0_{\mathbb{R}^2}\}$, $\mathbb{R}^2$ stesso, e le rette passanti per l'origine $0_{\mathbb{R}^2} = (0, 0)^T$.
- I sottospazi vettoriali di $\mathbb{R}^3$ sono: il sottospazio banale $\{0_{\mathbb{R}^3}\}$, $\mathbb{R}^3$ stesso, le rette passanti per l'origine $0_{\mathbb{R}^3} = (0, 0, 0)^T$ e i piani passanti per l'origine $0_{\mathbb{R}^3}$.
- L'insieme delle soluzioni di un sistema lineare omogeneo $\text{Sol}(A|0)$, $A \in \mathcal{M}_{\mathbb{K}}(m, n)$ è un sottospazio vettoriale di $\mathcal{M}_{\mathbb{K}}(n, 1)$, per esempio $W = \{(x, y, z)^T \in \mathbb{R}^3 \mid x + 2y - z = 0\} \triangleleft \mathbb{R}^3$.
- L'insieme delle soluzioni di un sistema lineare non omogeneo $\text{Sol}(A|b)$, $b \neq 0$ non è un sottospazio vettoriale di $V = \mathcal{M}_{\mathbb{K}}(n, 1)$, visto che $b \neq A \cdot 0_V = 0$, per esempio $W = \{(x, y, z)^T \in \mathbb{R}^3 \mid x + 2y - z = 1\} \not\triangleleft \mathbb{R}^3$ (questo è un sottospazio *affine*, ne parleremo alla fine del capito).

2.3 Combinazioni lineari

Definizione 2.3. Sia V uno spazio vettoriale sul campo $\mathbb{K}$, e sia $\{v_1, \dots, v_n\} \subset V$ un sottoinsieme finito di vettori.

Una *combinazione lineare* di $\{v_1, \dots, v_n\}$ è un elemento di V della forma

$$\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n \quad \text{con } \lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{K}.$$

L'insieme di tutte le combinazioni lineari di $\{v_1, \dots, v_n\}$ è chiamato *span* di $\{v_1, \dots, v_n\}$:

$$\text{Span}(v_1, \dots, v_n) = \{\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n \mid \lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{K}\}.$$

Proposizione 2.4. Sia V uno spazio vettoriale $\mathbb{K}$ e sia $\{v_1, \dots, v_n\} \subset V$ un sottoinsieme finito di vettori. Allora $\text{Span}(v_1, \dots, v_n)$ è un sottospazio vettoriale di V .

Dimostrazione. S1) Scegliamo $\lambda_1 = \lambda_2 = \dots = \lambda_n = 0_{\mathbb{K}}$: $0_{\mathbb{K}}v_1 + 0_{\mathbb{K}}v_2 + \dots + 0_{\mathbb{K}}v_n = 0_V \in \text{Span}(v_1, \dots, v_n)$.

S2) Siano $u = \sum_{i=1}^n \lambda_i v_i$ e $w = \sum_{i=1}^n \mu_i v_i$, allora $u+w = \sum_{i=1}^n (\lambda_i + \mu_i) v_i \in \text{Span}(v_1, \dots, v_n)$.

S3) Sia $u = \sum_{i=1}^n \lambda_i v_i$ e sia $\gamma \in \mathbb{K}$, allora $\gamma u = \sum_{i=1}^n (\gamma \lambda_i) v_i \in \text{Span}(v_1, \dots, v_n)$. $\square$

Esempio. 1. $\mathbb{K}[t]_{\leq d} = \{p(t) \in \mathbb{K}[t] \mid \deg(p) \leq d\} = \text{Span}(1, t, \dots, t^d)$.

2. In $\mathbb{K}^3 (= \mathcal{M}_{\mathbb{K}}(3, 1))$ sia $S = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}$, allora

$$\text{Span}(S) = \left\{ \lambda_1 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \mid \lambda_1, \lambda_2 \in \mathbb{K} \right\} = \left\{ \begin{pmatrix} \lambda_1 + \lambda_2 \\ \lambda_1 \\ \lambda_2 \end{pmatrix} \mid \lambda_1, \lambda_2 \in \mathbb{K} \right\}.$$

3. Analogamente, sia $S = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$, allora

$$\text{Span}(S) = \left\{ \begin{pmatrix} \lambda_1 + \lambda_3 \\ \lambda_2 + 2\lambda_3 \\ -\lambda_3 + \lambda_4 \end{pmatrix} \mid \lambda_1, \lambda_2, \lambda_3, \lambda_4 \in \mathbb{K} \right\}.$$

Osservazione. Si noti che un sistema lineare $Ax = b$ è risolvibile se e solo se b è una combinazione lineare delle colonne di A ! In altre parole, $Ax = b$ è risolvibile se e solo se $b \in \text{Span}(c_1, \dots, c_n)$, dove $c_i \in \mathcal{M}_{\mathbb{K}}(m, 1)$ sono le colonne di $A \in \mathcal{M}_{\mathbb{K}}(m, n)$.

Per esempio ($n = 3$): $\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \in \text{Span} \left(\begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right)$, infatti, il sistema lineare $\begin{pmatrix} 2 & 1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ è risolubile ($\lambda_1 = 1, \lambda_2 = -1$).

2.3.1 Generatori

Definizione 2.5. Sia V uno spazio vettoriale sul campo $\mathbb{K}$.

Si dice che l'insieme $\{v_1, \dots, v_n\} \subset V$ genera V (o che $\{v_1, \dots, v_n\}$ è un *insieme di generatori* di V , o che $v_1, \dots, v_n$ generano V) se $\text{Span}(v_1, \dots, v_n) = V$; cioè se ogni elemento in V si può scrivere come combinazione lineare di $\{v_1, \dots, v_n\}$.

Osservazione. i) Essere generatori di V significa che $\{v_1, \dots, v_n\}$ sono sufficienti per descrivere V .

ii) Se $\{v_1, \dots, v_n\} \subseteq \{v_1, \dots, v_n, v_{n+1}, \dots, v_s\}$ e $\text{Span}(v_1, \dots, v_n) = V$, allora $\text{Span}(v_1, \dots, v_n, \dots, v_s) = V$.

Esempio. Continuiamo gli esempi di prima.

1. $\{1, t, \dots, t^d\}$ genera $\mathbb{K}[t]_{\leq d}$.
2. $\left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}$ non genera $\mathbb{K}^3$, e.g. il sistema lineare

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

non è risolubile.

3. $\left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right\}$ è un insieme di generatori di $\mathbb{K}^3$, infatti il sistema lineare

$$\left(\begin{array}{cccc|c} 1 & 0 & 1 & 0 & b_1 \\ 0 & 1 & 2 & 0 & b_2 \\ 0 & 0 & -1 & 1 & b_3 \end{array} \right)$$

è risolubile per ogni $(b_1, b_2, b_3)^T \in \mathbb{K}^3$.

Osservazione. Quanti elementi deve avere un insieme di generatori di $\mathbb{K}^m$?

È necessario che la matrice “corrispondente” (con m righe) abbia m pivot a sinistra (in modo che il sistema lineare sia risolubile), quindi occorrono almeno m colonne: in altre parole, un insieme di generatori di $\mathbb{K}^m$ ha almeno m elementi!

2.3.2 Indipendenza lineare

Esempio. i) $T = \{1, t, t^2, t^3\}$ è un insieme di generatori di $\mathbb{K}[t]_{\leq 3}$, e $p(t) = (1+t)^3$ può essere scritto in modo univoco come combinazione lineare dei vettori in T : $(1+t)^3 = 1 + 3t + 3t^2 + t^3$.

ii) $S = \left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right\}$ è un insieme di generatori di $\mathbb{K}^3$, e ogni $v \in \mathbb{K}^3$ (e.g. $\begin{pmatrix} 2 \\ -3 \\ 5 \end{pmatrix}$) può essere scritto in diversi modi come combinazione lineare dei vettori in S . Infatti per il teorema di Rouché-Capelli l'insieme delle soluzioni del sistema lineare

$$\left(\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 2 \\ 0 & 1 & 2 & 0 & -3 \\ 0 & 0 & -1 & 1 & 5 \end{array} \right)$$

dipende da un parametro: $\text{Sol} = \left\{ \begin{pmatrix} 2 \\ -3 \\ 0 \\ 5 \end{pmatrix} + \begin{pmatrix} -t \\ -2t \\ t \\ t \end{pmatrix} \mid t \in \mathbb{K} \right\}$.

Vogliamo eliminare questa molteplicità di rappresentazioni ed ottenere una rappresentazione univoca!

Osservazione. In $\mathbb{K}^m$, se consideriamo più di m vettori, non possiamo avere una rappresentazione univoca (se esiste): infatti, la matrice “corrispondente” (con m righe) avrebbe più colonne che righe (vedi Lemma 1.14).

In altre parole, nell'esempio precedente l'ambiguità deriva dal fatto che abbiamo troppe variabili, cioè troppi vettori. Quindi l'idea è quella di rimuovere dei vettori.

Esempio. Continuiamo l'esempio precedente e consideriamo la rappresentazione di $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Risolviamo il corrispondente sistema lineare omogeneo

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

Per il Teorema 1.13, le soluzioni del sistema lineare omogeneo sono $\left\{ \begin{pmatrix} -t \\ -2t \\ t \\ t \end{pmatrix} \mid t \in \mathbb{K} \right\}$ e prendendo $t = 1$ otteniamo $\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 2\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} - 1\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \in \text{Span}\left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right)$, quindi possiamo rimuovere $\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$ da S senza cambiare lo span: $\text{Span}(S) = \text{Span}\left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right)$.

Definizione 2.6. Sia V uno spazio vettoriale sul campo $\mathbb{K}$. Un sottoinsieme finito $\{v_1, \dots, v_n\} \subset V$ è un *insieme di vettori linearmente indipendenti* se $\lambda_1 \cdot v_1 + \dots + \lambda_n \cdot v_n = 0_V \Leftrightarrow \lambda_1 = \dots = \lambda_n = 0_{\mathbb{K}}$.

Si dice che $\{v_1, \dots, v_n\} \subset V$ è un *insieme di vettori linearmente dipendenti*, se non è un insieme di vettori linearmente indipendenti, cioè se esistono scalari non tutti nulli $\lambda_1, \dots, \lambda_n$ tali che $\lambda_1 \cdot v_1 + \dots + \lambda_n \cdot v_n = 0_V$.

Osservazione. a) Essere linearmente indipendenti significa che non abbiamo informazioni ridondanti.

b) Se $S = \{v_1, \dots, v_n\}$ è un insieme di vettori linearmente indipendenti, allora qualsiasi sottoinsieme di S è un insieme di vettori linearmente indipendenti.

c) Due vettori $v_1, v_2 \in V$ sono linearmente dipendenti se e solo se uno è multiplo dell'altro

d) $\{0_V, v_2, \dots, v_n\} \subset V$ è un insieme di vettori linearmente dipendenti.

e) Se $\{v_1, \dots, v_n\} \subset V$ è un insieme di vettori linearmente dipendenti, allora $\{v_1, \dots, v_n, v_{n+1}, \dots, v_s\} \subset V$ è un insieme di vettori linearmente dipendenti.

Esempio. i) $\{1, t, \dots, t^d\}$ è un insieme di vettori linearmente indipendenti di $\mathbb{K}[t]_{\leq d}$.

ii) $S = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$ è un insieme di vettori linearmente dipendenti, infatti

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

iii) $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ sono linearmente indipendenti, infatti il sistema lineare omogeneo

$$\left(\begin{array}{cc|c} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{array} \right)$$

ha $(0, 0)^T$ come unica soluzione.

Proposizione 2.7 (“di scarto”/ “di aggiunta”). Sia V uno spazio vettoriale sul campo $\mathbb{K}$ e sia $\{v_1, \dots, v_n\} \subset V$ un sottoinsieme finito.

1. $\{v_1, \dots, v_n\}$ è un insieme di vettori linearmente dipendenti se e solo se uno dei vettori è una combinazione lineare degli altri.
2. Se $\{v_1, \dots, v_n\}$ è un insieme di vettori linearmente indipendenti, allora $v_1, \dots, v_n, v_{n+1}$ sono linearmente dipendenti se e solo se $v_{n+1} \in \text{Span}(v_1, \dots, v_n)$.

Dimostrazione. 1. $\Rightarrow$] Se $v_1, \dots, v_n$ sono linearmente dipendenti, esistono $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ non tutti pari a zero (ad esempio $\lambda_j \neq 0_{\mathbb{K}}$), tali che

$$\sum_{i=1}^n \lambda_i v_i = 0_V \implies v_j = \sum_{i=1, i \neq j}^n -\left(\frac{\lambda_i}{\lambda_j}\right) v_i$$

$$\Leftarrow] \text{ Se } v_k = \sum_{i=1, i \neq k}^n \alpha_i v_i \ (\alpha_i \in \mathbb{K}), \text{ allora } \sum_{i=1}^{k-1} \alpha_i v_i + (-1)v_k + \sum_{i=k+1}^n \alpha_i v_i = 0.$$

2. Esercizio. □

Teorema 2.8. Sia V uno spazio vettoriale sul campo $\mathbb{K}$, sia $\{v_1, \dots, v_n\} \subset V$ un insieme di vettori linearmente indipendenti e $\{w_1, \dots, w_m\} \subset V$ un insieme di generatori. Allora $n \leq m$.

Dimostrazione. Dimostriamo il teorema per assurdo. Supponiamo quindi che le ipotesi valgano e inoltre che $m < n$, e cerchiamo una contraddizione.

Poiché $v_1 \in V = \text{Span}(w_1, \dots, w_m)$, esistono scalari $\lambda_1, \dots, \lambda_m \in \mathbb{K}$ tali che $v_1 = \lambda_1 w_1 + \dots + \lambda_m w_m$. Poiché $v_1 \neq 0_V$, non tutti i λ_i sono zero, senza perdita di generalità possiamo supporre che $\lambda_1 \neq 0_{\mathbb{K}}$, quindi

$$\lambda_1 w_1 = v_1 - \lambda_2 w_2 - \dots - \lambda_m w_m \implies w_1 = \lambda_1^{-1} v_1 - \sum_{i=2}^m (\lambda_1^{-1} \lambda_i) w_i$$

Quindi $V = \text{Span}(w_1, w_2, \dots, w_m) = \text{Span}(v_1, w_2, \dots, w_m)$.

Consideriamo ora $v_2 \in V = \text{Span}(v_1, w_2, \dots, w_m)$: esistono scalari $\mu_1, \mu_2, \dots, \mu_m \in \mathbb{K}$, tali che $v_2 = \mu_1 v_1 + \mu_2 w_2 + \dots + \mu_m w_m$. Poiché v_1, v_2 sono linearmente indipendenti, non tutti $\mu_2, \dots, \mu_m$ sono zero e senza perdita di generalità possiamo assumere $\mu_2 \neq 0_{\mathbb{K}}$, quindi

$$w_2 = -(\mu_2^{-1} \mu_1) v_1 + \mu_2^{-1} v_2 - \sum_{i=3}^m (\mu_2^{-1} \mu_i) w_i$$

Quindi $V = \text{Span}(v_1, w_2, w_3, \dots, w_m) = \text{Span}(v_1, v_2, w_3, \dots, w_m)$.

Ripetendo questo processo m -volte ($m < n$), otteniamo $V = \text{Span}(w_1, \dots, w_m) = \text{Span}(v_1, \dots, v_m)$, quindi $v_{m+1} \in \text{Span}(v_1, \dots, v_m)$, contraddicendo il fatto che $v_1, \dots, v_m, v_{m+1}$ sono linearmente indipendenti. □

Spazi finitamente generati

Definizione 2.9. Uno spazio vettoriale V è *finitamente generato* se esiste un sottoinsieme finito $\{v_1, \dots, v_n\} \subset V$ tale che $V = \text{Span}(v_1, \dots, v_n)$.

Esempio. • $\mathbb{K}[t]_{\leq d} = \text{Span}(1, t, \dots, t^d)$ è finitamente generato.

- $\mathbb{K}^n$, $\mathcal{M}_{\mathbb{K}}(m, n)$ sono finitamente generati.

$$\mathbb{K}^n = \text{Span} \left(\begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \right)$$

$$\mathcal{M}_{\mathbb{K}}(m, n) = \text{Span} \left(\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & 0 & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix} \right)$$

- Se V è finitamente generato, allora ogni suo sottospazio vettoriale $W \triangleleft V$ è finitamente generato.

Esempio. • $\mathbb{K}[t]$ non è finitamente generato: le combinazioni lineari che possiamo ottenere da un insieme finito di polinomi non possono avere un grado arbitrariamente elevato.

- $\mathcal{C}^0(\mathbb{R})$ non è finitamente generato.

Buone notizie! In questo corso consideriamo solo spazi vettoriali finitamente generati. D'ora in poi, *ogni spazio vettoriale che consideriamo sarà implicitamente assunto come finitamente generato.*

2.4 Basi e dimensione

Definizione 2.10. Sia V uno spazio vettoriale sul campo $\mathbb{K}$.

Una *base* di V è un insieme ordinato $\mathcal{B} = \{v_1, \dots, v_n\} \subset V$ di vettori linearmente indipendenti che generano V .

Esempio. • $\mathcal{E} = \left\{ \underbrace{\begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}}_{e_1}, \underbrace{\begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}}_{e_2}, \dots, \underbrace{\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}}_{e_n} \right\}$ è una base di $\mathbb{K}^n$ chiamata *base canonica*.

- Analogamente,

$$\mathcal{E} = \left\{ \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & 0 & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix} \right\}$$

è una base di $\mathcal{M}_{\mathbb{K}}(m, n)$ chiamata *base canonica*.

- Per ogni $\alpha \in \mathbb{K}$ l'insieme $\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} \alpha \\ 1 \end{pmatrix} \right\}$ è una base di $\mathbb{K}^2$.
- $\mathcal{E} = \{1, t, \dots, t^d\}$ è una base di $\mathbb{K}[t]_{\leq d}$ chiamata *base canonica*.

Essere una base significa che $\{v_1, \dots, v_n\}$ sono sufficienti per descrivere V , e non ci sono ambiguità. Infatti:

Proposizione 2.11 (Esistenza ed unicità delle coordinate). *Sia V uno spazio vettoriale sul campo $\mathbb{K}$, e sia $\mathcal{B} = \{v_1, \dots, v_n\} \subset V$ una base di V . Allora ogni elemento di V può essere scritto in modo unico come combinazione lineare di $\{v_1, \dots, v_n\}$.*

Dimostrazione. $V = \text{Span}(v_1, \dots, v_n)$, quindi per ogni $v \in V$ esistono $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ tali che $v = \lambda_1 v_1 + \dots + \lambda_n v_n$.

Per dimostrare l'unicità, scegliamo due rappresentazioni di v :

$$v = \lambda_1 v_1 + \dots + \lambda_n v_n, \quad v = \mu_1 v_1 + \dots + \mu_n v_n.$$

e consideriamo la loro differenza: $0_V = v - v = (\lambda_1 - \mu_1)v_1 + \dots + (\lambda_n - \mu_n)v_n$; ma $v_1, \dots, v_n$ sono linearmente indipendenti, quindi $\lambda_1 - \mu_1 = 0_{\mathbb{K}}, \dots, \lambda_n - \mu_n = 0_{\mathbb{K}}$, ovvero $\lambda_1 = \mu_1, \dots, \lambda_n = \mu_n$. $\square$

Definizione 2.12. Sia V uno spazio vettoriale sul campo $\mathbb{K}$, $\mathcal{B} = \{v_1, \dots, v_n\} \subset V$ una base di V , tale che per ogni $v \in V$ esistono scalari unici $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ tali che

$$v = \sum_{i=1}^n \lambda_i v_i.$$

Gli scalari $\lambda_1, \dots, \lambda_n$ sono le *coordinate* di v rispetto alla base $\mathcal{B}$ e sono indicati con $(v)_{\mathcal{B}} = (\lambda_1, \dots, \lambda_n)^T \in \mathbb{K}^n$.

Osservazione. Sia V uno spazio vettoriale sul campo $\mathbb{K}$, e sia $\mathcal{B} = \{v_1, \dots, v_n\} \subset V$ una base di V . Siano $v_1, v_2 \in V$ e $k_1, k_2 \in \mathbb{K}$, allora

$$(k_1 v_1 + k_2 v_2)_{\mathcal{B}} = k_1 (v_1)_{\mathcal{B}} + k_2 (v_2)_{\mathcal{B}}$$

Quindi passando alle coordinate, possiamo descrivere un vettore in V tramite un vettore in $\mathbb{K}^n$ ($n = \dim V$).

Cosa hanno in comune due basi di uno spazio vettoriale? Il numero di elementi!

Teorema 2.13 (Teorema della dimensione). *Sia V uno spazio vettoriale sul campo $\mathbb{K}$, e siano $\mathcal{B} = \{b_1, \dots, b_n\} \subset V$ e $\mathcal{C} = \{c_1, \dots, c_m\} \subset V$ basi di V . Allora $m = n$.*

Dimostrazione. Utilizzando il fatto che $b_1, \dots, b_n$ sono linearmente indipendenti e che $c_1, \dots, c_m$ sono generatori, otteniamo $n \leq m$ dal Teorema 2.8. Invertendo le proprietà: $b_1, \dots, b_n$ sono generatori e che $c_1, \dots, c_m$ sono linearmente indipendenti, otteniamo $m \leq n$ dal Teorema 2.8. $\square$

Definizione 2.14. Sia V uno spazio vettoriale (finitamente generato) sul campo $\mathbb{K}$. La *dimensione* di V è il numero di elementi di una base di V ed è indicata con $\dim V$.

Esempio. • $\dim\{0\} = 0$, infatti $\mathcal{B} = \emptyset$ è una base di $\{0\}$.

- $\dim \mathbb{K}^n = n$.
- $\dim \mathcal{M}_{\mathbb{K}}(m, n) = m \cdot n$.
- $\dim \mathbb{K}[t]_{\leq d} = d + 1$.
- Sia $A \in \mathcal{M}_{\mathbb{K}}(m, n)$, allora $\text{Sol}(A|0) \triangleleft \mathbb{K}^n = \mathcal{M}_{\mathbb{K}}(n, 1)$ è un sottospazio vettoriale. Il Teorema di Rouché-Capelli ci dice che $\text{Sol}(A|0) = \{\lambda_1 v_1 + \cdots \lambda_{n-r} v_{n-r} \mid \lambda_i \in \mathbb{K}\} = \text{Span}(v_1, \dots, v_{n-r})$ dove $r = \text{rg}(A)$, e garantisce che non ci siano generatori ridondanti, ovvero che $v_1, \dots, v_{n-r}$ siano linearmente indipendenti, quindi $\dim \text{Sol}(A|0) = n - \text{rg}(A)$.

Teorema 2.15. *Sia V uno spazio vettoriale sul campo $\mathbb{K}$ e sia $W \triangleleft V$ un suo sottospazio. Allora i) $\dim W \leq \dim V$, e inoltre ii) $\dim W = \dim V \Leftrightarrow W = V$.*

Dimostrazione. i) Sia $\{w_1, \dots, w_s\} \subset W \subset V$ una base di W , in particolare un insieme di vettori linearmente indipendenti in V . Per il Teorema 2.8 $\dim W = s \leq \dim V$.

Per ii). Se $W = V$, chiaramente $\dim W = \dim V$.

Ora, sia $\dim W = \dim V = n$ e supponiamo che esista $v \in V$, $v \notin W$. Aggiungendo v a una base $\{w_1, \dots, w_n\}$ di W , otteniamo $n + 1$ linearmente indipendenti in V .

Quindi $v \notin \text{Span}(w_1, \dots, w_n)$, quindi $\{w_1, \dots, w_n, v\}$ sono linearmente indipendenti in V (per il punto 2 della Proposizione 2.7), contraddizione. $\square$

Teorema 2.16 (Teorema della base). *Sia V un $\mathbb{K}$ -spazio vettoriale (finitamente generato!). Allora valgono le seguenti affermazioni.*

1. *Se $\dim V = n$ e $\{w_1, \dots, w_m\}$ è un insieme di generatori di V , allora $m \geq n$.*

Inoltre, se $m = n$, allora $\{w_1, \dots, w_m\}$ è una base, mentre se $m > n$ possiamo rimuovere $m - n$ vettori in modo tale che quelli rimanenti siano una base di V .

“Da ogni insieme di generatori possiamo estrarre una base.”

2. *Se $\dim V = n$ e siano $\{v_1, \dots, v_p\}$ vettori linearmente indipendenti, allora $p \leq n$.*

Inoltre, se $p = n$, allora $\{v_1, \dots, v_p\}$ è una base, mentre se $p < n$ esistono vettori $v_{p+1}, \dots, v_n$ tali che $\{v_1, \dots, v_p, v_{p+1}, \dots, v_n\}$ è una base di V .

“Ogni insieme di vettori linearmente indipendenti può essere completato fino a formare una base.”

Dimostrazione. 1. La prima affermazione segue dal Teorema 2.8.

Se $w_1, \dots, w_m$ sono linearmente dipendenti, in base al punto 1 della Proposizione 2.7 uno dei vettori (ad esempio w_m) è una combinazione lineare degli altri, quindi potremmo escluderlo e avere comunque un insieme di generatori.

Ora, se $m = n$, allora $w_1, \dots, w_m$ non possono essere linearmente dipendenti, altrimenti potremmo escludere un vettore e ottenere un insieme di generatori di V di cardinalità $n - 1$, il che è impossibile.

D'altra parte, se $m > n$, allora $w_1, \dots, w_m$ sono linearmente dipendenti per il Teorema 2.8. Per la Proposizione 2.7 possiamo scartare un vettore ed avere ancora un sistema di generatori con $m - 1$ elementi. Applicando questa procedura $m - n$ volte, possiamo escludere $m - n$ vettori e ottenere una base di V .

2. La prima affermazione segue dal Teorema 2.8.

Per la seconda, se $p = n$, scegliamo $W = \text{Span}(v_1, \dots, v_p)$ e applichiamo il Teorema 2.16. Se $p < n$, allora $\text{Span}(v_1, \dots, v_p) \neq V$, quindi esiste $v_{p+1} \notin \text{Span}(v_1, \dots, v_p)$. Per la Proposizione 2.7 $\{v_1, \dots, v_p, v_{p+1}\}$ sono linearmente indipendenti. Ripetendo questo passaggio $n - p$ volte, si ottiene infine una base di V . $\square$

Dalla precedente proposizione deduciamo le seguenti importanti osservazioni.

Osservazione. i) Ogni spazio vettoriale (finitamente generato) ha una base!

ii) In uno spazio vettoriale V di dimensione n , un insieme di n vettori $S = \{v_1, \dots, v_n\}$ è un insieme di generatori se e solo se è un insieme di vettori linearmente indipendenti. Quindi, in questo caso particolare, per verificare se S è una base, è sufficiente verificare una sola delle due condizioni.

Esempio. • $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ sono linearmente indipendenti in $\mathbb{K}^3$, ma non sono generatori, perché $2 < 3 = \dim \mathbb{K}^3$. Possiamo completare questo insieme a base di $\mathbb{K}^3$, per esempio $\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$

- Sia $S := \{1 + t, t + t^2, 1 + t^2\} \subset \mathbb{K}[t]_{\leq 2}$. Sono generatori? Sono linearmente indipendenti? $\dim \mathbb{K}[t]_{\leq 2} = 3$, possiamo rispondere simultaneamente alle 2 domande.

Controlliamo se $1 + t, t + t^2, 1 + t^2$ sono linearmente indipendenti:

$$0 = \lambda_1(1+t) + \lambda_2(t+t^2) + \lambda_3(1+t^2) = (\lambda_1 + \lambda_3) + (\lambda_1 + \lambda_2)t + (\lambda_2 + \lambda_3)t^2 \Leftrightarrow \begin{cases} \lambda_1 + \lambda_3 = 0 \\ \lambda_1 + \lambda_2 = 0 \\ \lambda_2 + \lambda_3 = 0 \end{cases}$$

Quindi $1 + t, t + t^2, 1 + t^2$ sono linearmente indipendenti se e solo se il sistema lineare ha una soluzione unica, cioè se e solo se la sua matrice dei coefficienti ha determinante diverso da zero:

$$\det \underbrace{\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}}_A = \det \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} + \det \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = 1 + 1 \neq 0.$$

Quindi $\{1 + t, t + t^2, 1 + t^2\}$ è una base di $\mathbb{K}[t]_{\leq 2}$.

Si noti che le colonne di A sono le coordinate di $1 + t, t + t^2, 1 + t^2$ rispetto alla base canonica $\{1, t, t^2\}$ di $\mathbb{K}[t]_{\leq 2}$.

Osservazione. i) Come già osservato, prendendo le coordinate, possiamo descrivere un vettore in V tramite un vettore in $\mathbb{K}^n$ ($n = \dim V$). Quindi possiamo “tradurre un problema su V in un problema su $\mathbb{K}^n$ ”. Nel prossimo capitolo renderemo questo concetto più preciso e concreto.

ii) In $\mathbb{K}^n$, per verificare se n vettori formano una base, è sufficiente verificare se sono linearmente indipendenti, ovvero se il sistema lineare omogeneo “corrispondente” ha una soluzione unica. La sua matrice dei coefficienti A è una matrice quadrata con gli n vettori come colonne, quindi il sistema lineare ha una soluzione unica se e solo se A è invertibile.

In altre parole, $\{v_1, \dots, v_n\} \subset \mathbb{K}^n$ è una base di $\mathbb{K}^n$ se e solo se la matrice $A = (v_1 | \dots | v_n)$ (le v_i sono le colonne) ha rango massimo (cioè $\det A \neq 0$).

2.4.1 Un'osservazione sul rango

Definizione 2.17. Sia $A \in \mathcal{M}_{\mathbb{K}}(m, n)$ una matrice. Siano $R_1, \dots, R_m \in \mathcal{M}_{\mathbb{K}}(1, n)$ le sue righe e siano $C_1, \dots, C_n \in \mathcal{M}_{\mathbb{K}}(m, 1)$ le sue colonne.

Lo *spazio delle righe* di A è $\text{Row}(A) = \text{Span}(R_1, \dots, R_m) \triangleleft \mathbb{K}^n$.

Lo *spazio delle colonne* di A è $\text{Col}(A) = \text{Span}(C_1, \dots, C_n) \triangleleft \mathbb{K}^m$.

Il *nucleo* di A è $\ker A = \text{Sol}(A|0) = \{v \in \mathbb{K}^n | Av = 0\}$.

Esempio. Sia $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \in \mathcal{M}_{\mathbb{K}}(2, 3)$, e $B = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 0 & 2 \end{pmatrix} \in \mathcal{M}_{\mathbb{K}}(2, 3)$ allora

$$\text{Row}(A) = \text{Span}((1, 2, 3), (4, 5, 6)) \triangleleft \mathbb{K}^3, \quad \text{Col}(A) = \text{Span}\left(\begin{pmatrix} 1 \\ 4 \end{pmatrix}, \begin{pmatrix} 2 \\ 5 \end{pmatrix}, \begin{pmatrix} 3 \\ 6 \end{pmatrix}\right) \triangleleft \mathbb{K}^2.$$

$$\text{Row}(B) = \text{Span}((1, 0, 1), (2, 0, 2)) \triangleleft \mathbb{K}^3, \quad \text{Col}(B) = \text{Span}\left(\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}\right) \triangleleft \mathbb{K}^2.$$

Si osservi che $\dim \text{Row}(A) = \dim \text{Col}(A) = \text{rg}(A) = 2$ e $\dim \text{Row}(B) = \dim \text{Col}(B) = \text{rg}(B) = 1$.

Teorema 2.18. Sia $A \in \mathcal{M}_{\mathbb{K}}(m, n)$, allora

$$\text{rg}(A) = \dim \text{Row}(A) = \dim \text{Col}(A) = \text{rg}(A^T)$$

Idea: le operazioni elementari sulle righe non modificano $\text{Row}(A)$, quindi se U è in forma a scala ed è ottenuta da A utilizzando l'algoritmo di Gauss, abbiamo $\text{Row}(A) = \text{Row}(U)$, ma $\dim \text{Row}(U)$ coincide con il numero di pivot, ovvero il numero di righe diverse da zero di U (sono linearmente indipendenti!).

Attenzione: In generale $\text{Col}(A) \neq \text{Col}(U)$, ma $\dim \text{Col}(A) = \dim \text{Col}(U) = \text{numero di pivot}$.

Proposizione 2.19. Sia $A \in \mathcal{M}_{\mathbb{K}}(m, n)$ una matrice di rango $\text{rg}(A) = r$, e sia U una sua riduzione a scala, con $q_1, \dots, q_r$ gli indici delle colonne di U contenenti i pivot. Allora le colonne $c_{q_1}, \dots, c_{q_r}$ di A formano una base di $\text{Col}(A)$.

Osservazione. Il Teorema 2.18 ci fornisce una nuova interpretazione/dimostrazione della parte 1. del Teorema 1.11 (Rouché-Capelli). Infatti,

$$Ax = b \text{ è risolvibile} \Leftrightarrow b \in \text{Col}(A) = \text{Span}(C_1, \dots, C_n) \Leftrightarrow \text{Span}(C_1, \dots, C_n) \stackrel{(*)}{=} \text{Span}(C_1, \dots, C_n, b)$$

Ma $\text{Span}(C_1, \dots, C_n) \triangleleft \text{Span}(C_1, \dots, C_n, b)$, quindi l'uguaglianza $(*)$ vale se e solo se

$$\dim \text{Span}(C_1, \dots, C_n) = \dim \text{Span}(C_1, \dots, C_n, b) \Leftrightarrow \dim \text{Col}(A) = \dim \text{Col}(A|b) \Leftrightarrow \text{rg}(A) = \text{rg}(A|b).$$

2.5 Equazioni cartesiane e parametriche

Come visto finora, esistono due modi per definire un sottospazio vettoriale $W \triangleleft \mathbb{K}^n$:

- implicitamente, come soluzioni di un sistema lineare omogeneo, cioè come nucleo di una matrice; si tratta di *equazioni cartesiane*;
- esplicitamente, come span di alcuni vettori; si tratta di *equazioni parametriche*.

Osservazione. i) Se $\dim W = r$, allora W può essere espresso utilizzando r vettori, quindi r parametri; oppure tramite un sistema lineare omogeneo con $n - r$ equazioni (Teorema di Rouché-Capelli).

ii) Potremmo utilizzare più vettori o equazioni, ma non sarebbero linearmente indipendenti, ovvero avremmo informazioni ridondanti.

Cartesiane $\rightarrow$ parametriche) Date delle equazioni cartesiane, cioè un sistema lineare omogeneo, risolvendo il sistema si ottengono delle equazioni parametriche.

Parametriche $\rightarrow$ cartesiane) Date le equazioni parametriche $W = \text{Span}(w_1, \dots, w_r) \triangleleft \mathbb{K}^n$, possiamo interpretarle come $W = \text{Col}(A)$, (dove $A \in \mathcal{M}_{\mathbb{K}}(n, r)$ ha $w_1, \dots, w_r$ come colonne). Dall'osservazione dopo il Teorema 2.18, vediamo che un vettore $v = (x_1, \dots, x_n)^T \in \mathbb{K}^n$ appartiene a W se e solo se $v \in \text{Col}(A)$ se e solo se $\text{rg}(A) = \text{rg}(A|v)$. Imponendo questa condizione otteniamo equazioni cartesiane di W .

Esempio. Sia $W = \text{Span}((2, -2, 1, 0)^T, (1, 0, -1, 1)^T) \triangleleft \mathbb{R}^4$, allora $(x_1, x_2, x_3, x_4)^T \in W$ se e solo se la matrice

$$\left(\begin{array}{cc|c} 2 & 1 & x_1 \\ -2 & 0 & x_2 \\ 1 & -1 & x_3 \\ 0 & 1 & x_4 \end{array} \right)$$

ha rango 2. Usando le operazioni elementari sulle righe (che non cambiano il rango) portiamo la matrice nella seguente forma:

$$\left(\begin{array}{cc|c} 1 & -1 & x_3 \\ 0 & 1 & x_4 \\ 0 & 0 & x_1 - 2x_3 - 3x_4 \\ 0 & 0 & x_2 + 2x_3 + 2x_4 \end{array} \right) \text{ che ha rango 2 se e solo se } v \text{ è soluzione di } \begin{cases} x_1 - 2x_3 - 3x_4 = 0 \\ x_2 + 2x_3 + 2x_4 = 0 \end{cases}$$

Queste sono equazioni cartesiane di W .

Osservazione. In alternativa, date le equazioni parametriche, possiamo recuperare le equazioni cartesiane eliminando i parametri, come fatto nell'Introduzione.

2.6 Intersezione e somma di sottospazi vettoriali

Lemma 2.20. Sia V un $\mathbb{K}$ -spazio vettoriale e siano $U, W \triangleleft V$ sottospazi vettoriali di V . Allora $U \cap W$ è un sottospazio vettoriale di V .

Dimostrazione. S1) $0_V \in U$ e $0_V \in W$, quindi $0_V \in U \cap W$.

S2) Siano $v_1, v_2 \in U \cap W$, allora $v_1 + v_2 \in U$ (poiché $U \triangleleft V$) e $v_1 + v_2 \in W$ (poiché $W \triangleleft V$), quindi $v_1 + v_2 \in U \cap W$.

S3) Sia $v \in U \cap W$ e $\lambda \in \mathbb{K}$, allora $\lambda v \in U$ (poiché $U \triangleleft V$) e $\lambda v \in W$ (poiché $W \triangleleft V$), quindi $\lambda v \in U \cap W$. $\square$

Esempio. Siano $W = \{(x, y, z)^T \in \mathbb{R}^3 \mid x + 2y - z = 0\} \triangleleft \mathbb{R}^3$ e $U = \{(x, y, z)^T \in \mathbb{R}^3 \mid -y + 3z = 0\} \triangleleft \mathbb{R}^3$, allora

$$U \cap W = \left\{ (x, y, z)^T \in \mathbb{R}^3 \mid \begin{cases} x + 2y - z = 0 \\ -y + 3z = 0 \end{cases} \right\} = \left\{ \begin{pmatrix} -5 \\ -3 \\ 1 \end{pmatrix} t \mid t \in \mathbb{R} \right\}.$$

E l'unione?

Osservazione. In generale l'unione di due sottospazi *non* è unsottospazio. Per esempio, consideriamo $U = \{(x, y)^T \in \mathbb{R}^2 \mid x + y = 0\} \triangleleft \mathbb{R}^2$ e $W = \{(x, y)^T \in \mathbb{R}^2 \mid x - y = 0\} \triangleleft \mathbb{R}^2$ sottospazi di $\mathbb{R}^2$. Allora $(1, 1)$ e $(1, -1)$ appartengono all'unione $U \cup W$, ma la loro somma no: $(1, 1) + (1, -1) = (2, 0) \notin U \cup W$.

Definizione 2.21. Sia V un $\mathbb{K}$ -spazio vettoriale, e siano $U, W \triangleleft V$ suoi sottospazi. Definiamo la loro *somma* $U + W$ come

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

“ $U + W$ contiene tutti i vettori di V che possono essere scritti come somma di un vettore in U ed un vettore in W ”.

Lemma 2.22. Sia V un $\mathbb{K}$ -spazio vettoriale e siano $U, W \triangleleft V$ sottospazi vettoriali di V . Allora $U + W$ è un sottospazio vettoriale di V .

Dimostrazione. S1) $0_V = \underbrace{0_V}_{\in U} + \underbrace{0_V}_{\in W} \in U + W$.

S2) Siano $v_1 = u_1 + w_1, v_2 = u_2 + w_2 \in U + W$, con $u_1, u_2 \in U, w_1, w_2 \in W$. Allora, $v_1 + v_2 = \underbrace{(u_1 + u_2)}_{\in U} + \underbrace{(w_1 + w_2)}_{\in W} \in U + W$.

S3) Sia $\lambda \in \mathbb{K}$ e sia $v = u + w \in U \cap W$ con $u \in U, w \in W$. Allora, $\lambda v = \underbrace{(\lambda u)}_{\in U} + \underbrace{(\lambda w)}_{\in W} \in U + W$. □

Osservazione. i) $U \cap W \triangleleft U \triangleleft U + W$ e analogamente $U \cap W \triangleleft W \triangleleft U + W$.

ii) Se $M \triangleleft V$ e $U, W \subseteq M$, allora $U + W \subseteq M$; cioè $U + W$ è il più piccolo sottospazio di V che contiene sia U che W .

Infatti, sia $u + w \in U + W$ ($u \in U, w \in W$). Visto che M contiene sia U che W , si ha che $u + w \in M$ (dato che $M \triangleleft V$).

Esempio. a) Siano $U = \{(x, y)^T \in \mathbb{R}^2 \mid x + y = 0\} \triangleleft \mathbb{R}^2$ e $W = \{(x, y)^T \in \mathbb{R}^2 \mid x - y = 0\} \triangleleft \mathbb{R}^2$, allora $U + W = \mathbb{R}^2$. Infatti, ogni $(a, b) \in \mathbb{R}^2$ può essere scritto come $(a, b) = \left(\frac{a-b}{2}, \frac{b-a}{2}\right) + \left(\frac{a+b}{2}, \frac{a+b}{2}\right) \in U + W$.

b) Sia $U = \left\{\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} t \mid t \in \mathbb{R}\right\} \triangleleft \mathbb{R}^3$ e $W = \left\{\begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix} s \mid s \in \mathbb{R}\right\} \triangleleft \mathbb{R}^3$, allora

$$U + W = \left\{\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} t + \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix} s \mid t, s \in \mathbb{R}\right\} \triangleleft \mathbb{R}^3.$$

Definizione 2.23. Sia V un $\mathbb{K}$ -spazio vettoriale e siano $U, W \triangleleft V$ sottospazi. La somma $U + W$ si dice *diretta* e si scrive $U \oplus W$ se $U \cap W = \{0_V\}$.

Esempio. Nei due esempi precedenti la somma è diretta, in particolare in a) $U \oplus W = \mathbb{R}^2$.

Osservazione. La condizione $U \cap W = \{0_V\}$ è equivalente a dire che tutti i vettori di $U + W$ possono essere scritti in modo unico come somma di un vettore in U e un vettore in W . La dimostrazione di questo fatto è lasciata per esercizio al lettore.

2.6.1 Formula di Grassmann

Teorema 2.24 (Formula di Grassmann). *Sia V uno spazio vettoriale $\mathbb{K}$ e siano $U, W \triangleleft V$ sottospazi vettoriali. Allora*

$$\dim U + \dim W = \dim(U \cap W) + \dim(U + W)$$

Dimostrazione. Poniamo $\dim U = p$, $\dim W = q$ e $\dim(U \cap W) = r$. Dal Teorema 2.16 abbiamo che $r \leq p$ e $r \leq q$, e dobbiamo dimostrare che $\dim(U + W) = p + q - r$.

Sia $\mathcal{B} = \{b_1, \dots, b_r\}$ una base di $U \cap W$ e completiamola fino a ottenere una base di U $\{b_1, \dots, b_r, c_{r+1}, \dots, c_p\}$ mediante l'aggiunta di $p - r$ vettori. Completiamo inoltre $\mathcal{B}$ ad una base di W $\{b_1, \dots, b_r, d_{r+1}, \dots, d_q\}$ mediante l'aggiunta di $q - r$ vettori.

Mostreremo che $S = \{b_1, \dots, b_r, c_{r+1}, \dots, c_p, d_{r+1}, \dots, d_q\}$ è una base di $U + W$, quindi $\dim(U + W) = r + (p - r) + (q - r) = p + q - r$.

Generatori. Sia $v \in U + W$, allora esistono $u \in U, w \in W$ tali che $v = u + w$. Ma quindi esistono scalari tali che

$$\begin{aligned} u &= \alpha_1 b_1 + \dots + \alpha_r b_r + \gamma_{r+1} c_{r+1} + \dots + \gamma_p c_p, \\ w &= \beta_1 b_1 + \dots + \beta_r b_r + \delta_{r+1} d_{r+1} + \dots + \delta_q d_q \end{aligned}$$

e quindi $v = (\alpha_1 + \beta_1)b_1 + \dots + (\alpha_r + \beta_r)b_r + \gamma_{r+1}c_{r+1} + \dots + \gamma_p c_p + \delta_{r+1}d_{r+1} + \dots + \delta_q d_q$.

Sono linearmente indipendenti: scegliamo $p + q - r$ scalari tali che

$$0_V = \lambda_1 b_1 + \dots + \lambda_r b_r + \eta_{r+1} c_{r+1} + \dots + \eta_p c_p + \tau_{r+1} d_{r+1} + \dots + \tau_q d_q$$

e definiamo $u := \lambda_1 b_1 + \dots + \lambda_r b_r + \eta_{r+1} c_{r+1} + \dots + \eta_p c_p \in U$, in modo che $-u = \tau_{r+1} d_{r+1} + \dots + \tau_q d_q \in W$ e per S3) $u \in W$. Pertanto, $u \in U \cap W$ ed esistono r scalari $\mu_1, \dots, \mu_r$ tali che $u = \mu_1 b_1 + \dots + \mu_r b_r$. Da ciò segue che

$$\mu_1 b_1 + \dots + \mu_r b_r + \tau_{r+1} d_{r+1} + \dots + \tau_q d_q = u - u = 0_V.$$

Essendo $b_1, \dots, b_r, d_{r+1}, \dots, d_q$ linearmente indipendenti, otteniamo $\mu_1 = \dots = \mu_r = \tau_{r+1} = \dots = \tau_q = 0_{\mathbb{K}}$. In particolare,

$$0_V = u = \lambda_1 b_1 + \dots + \lambda_r b_r + \eta_{r+1} c_{r+1} + \dots + \eta_p c_p$$

e essendo $b_1, \dots, b_r, c_{r+1}, \dots, c_p$ linearmente indipendenti, otteniamo $\lambda_1 = \dots = \lambda_r = \eta_{r+1} = \dots = \eta_p = 0_{\mathbb{K}}$. $\square$

Una conseguenza immediata della formula di Grassmann è la seguente affermazione.

Corollario 2.25. *Sia V uno spazio vettoriale $\mathbb{K}$ e siano $U, W \triangleleft V$ sottospazi vettoriali. Allora la somma $U + W$ è diretta se e solo se $\dim(U \cap W) = 0$ se e solo se $\dim U + \dim W = \dim(U + W)$.*

In particolare, sia $\mathcal{B}_U$ una base di U e $\mathcal{B}_W$ una base di W , allora la somma $U + W$ è diretta se e solo se $\mathcal{B}_U \cup \mathcal{B}_W$ è una base di $U + W$.

Se la somma non è diretta, $\mathcal{B}_U \cup \mathcal{B}_W$ genera $U + W$, ma per ottenere una base, dobbiamo eliminare alcuni vettori

Esempio. In $\mathbb{R}^4$ consideriamo il sottospazio subspaces

$$U = \left\{ (x, y, z, w)^T \in \mathbb{R}^4 \mid \begin{cases} x + y - z = 0 \\ x - y + w = 0 \end{cases} \right\}, \quad W = \text{Span} \left(\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right).$$

Vogliamo determinare una base (e la dimensione) di U , W , $U \cap W$, $U + W$ e determinare se la somma è diretta.

Notiamo che $U = \text{Sol}(A|0)$, dove $A = \begin{pmatrix} 1 & 1 & -1 & 0 \\ 1 & -1 & 0 & 1 \end{pmatrix}$, che si riduce in scala a $\begin{pmatrix} 1 & 1 & -1 & 0 \\ 0 & -2 & 1 & 1 \end{pmatrix}$ (applicando $R_2 \rightarrow R_2 - R_1$), così che

$$U = \left\{ \begin{pmatrix} (s-t)/2 \\ (s+t)/2 \\ s \\ t \end{pmatrix} \mid s, t \in \mathbb{R} \right\} = \text{Span} \left(\begin{pmatrix} 1 \\ 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 0 \\ 2 \end{pmatrix} \right) \Rightarrow \dim U = 2$$

Per trovare una base di W , scriviamo i generatori di W come colonne di una matrice e la riduciamo. In base alla Proposizione 2.19, le colonne corrispondenti ai pivot *nella matrice originale* sono una base per W :

$$\begin{pmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow{R_2 \rightarrow R_2 - R_1} \begin{pmatrix} 1 & 1 & 2 \\ 0 & -1 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow{R_3 \rightarrow R_3 + R_2} \begin{pmatrix} 1 & 1 & 2 \\ 0 & -1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\text{Quindi } W = \text{Span} \left(\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right) \text{ e } \dim W = 2.$$

(*) Consideriamo ora $U + W$. La somma $U + W$ contiene tutti i vettori che possono essere scritti come somma di un vettore in U e un vettore in W , quindi è generata da una base di U insieme a una base di W :

$$U + W = \text{Span} \left(\begin{pmatrix} 1 \\ 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right)$$

Procedendo come sopra, otteniamo che quei quattro vettori sono linearmente indipendenti, quindi $\dim(U + W) = 4$. Secondo la formula di Grassmann $\dim(U \cap W) = 0$, ovvero $U \cap W = \{0_V\}$.

Alternativa a ().* Potremmo aver iniziato considerando $U \cap W$. Un vettore in $W = \text{Span} \left(\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right)$ è della forma $\begin{pmatrix} a \\ 0 \\ b \\ 0 \end{pmatrix}$ per $a, b \in \mathbb{R}$. Tale vettore appartiene a $U \cap W$ se soddisfa le condizioni per essere in U :

$$\begin{cases} a + (a - b) - b = 0 \\ a - (a - b) + 0 = 0 \end{cases} \Leftrightarrow \begin{cases} a = b \\ b = 0 \end{cases}$$

Quindi esiste un'unica possibilità: $(a, b) = (0, 0)$. In altre parole, $U \cap W = \{0_V\}$, quindi $\dim(U \cap W) = 0$. Secondo la formula di Grassmann $\dim(U + W) = 4$, ovvero $U \oplus W = \mathbb{R}^4$.

2.7 Sottospazi affini

Nelle sezioni precedenti ci siamo concentrati sugli spazi vettoriali e sui sottospazi vettoriali. In particolare abbiamo visto che qualsiasi sottospazio vettoriale di $\mathbb{K}^n$ può essere scritto come spazio delle soluzioni di un sistema lineare omogeneo $Ax = 0$ per una certa matrice $A \in \mathcal{M}_{\mathbb{K}}(m, n)$.

Cosa succede per sistemi lineari qualsiasi?

Definizione 2.26. Sia $Ax = b$ un sistema lineare risolubile, con $A \in \mathcal{M}_{\mathbb{K}}(m, n)$ e $b \in \mathbb{K}^m$. L'insieme delle soluzioni $S := \text{Sol}(A|b)$ è un *sottospazio affine* di $\mathbb{K}^n$.

Il sottospazio vettoriale corrispondente $S_0 := \ker(A)$ si chiama la *giacitura* di S e la *dimensione* di un sottospazio affine è la dimensione della sua giacitura: $\dim S := \dim S_0$. Gli elementi di un sottospazio affine sono chiamati *punti*.

Definizione 2.27. Sia S un *sottospazio affine* di $\mathbb{K}^n$.

S è una *retta* se $\dim S = 1$.

S è un *piano* se $\dim S = 2$.

S è un *iperpiano* se $\dim S = n - 1$.

Esempio. Ogni sottospazio vettoriale di $\mathbb{K}^n$ è anche un sottospazio affine di $\mathbb{K}^n$; in particolare $\mathbb{K}^n$ è un sottospazio affine di $\mathbb{K}^n$ (basta considerare $A = O$ e $b = 0$).

Un punto di coordinate $P = (p_1, \dots, p_n)^T \in \mathbb{K}^n$ è un sottospazio affine di $\mathbb{K}^n$: è la soluzione del sistema lineare $Ax = b$, con $A = I_n$ e $b = (p_1, \dots, p_n)^T$.

Per evitare banalità, d'ora in poi supporremo che $\text{rg}(A) \neq 0$.

Osservazione. 1) Si noti che se $m > \text{rg}(A)$, allora possiamo applicare il MEG per ridurre $(A|b)$ in forma a scala, e scartare le $m - \text{rg}(A)$ equazioni corrispondenti alle righe nulle. Quindi, possiamo sempre supporre che $m = \text{rg}(A) = \text{rg}(A|b) \leq n$.

2) Sia $Ax = b$ un sistema lineare risolubile e sia $P \in \text{Sol}(A|b)$ una soluzione. Ricordiamo che, per il Teorema 1.13, vale $\text{Sol}(A|b) = P + \text{Sol}(A|0)$; in particolare, se $P_1, P_2 \in \text{Sol}(A|b)$ allora $P_1 - P_2 \in \text{Sol}(A|0) = \ker(A)$.

Esempio (Sottospazi affini di $\mathbb{R}^2$). Siano $A \in \mathcal{M}_{\mathbb{R}}(m, 2)$ e $b \in \mathbb{R}^m$ con $0 < m = \text{rg}(A) = \text{rg}(A|b) \leq 2$. Sia $S := \text{Sol}(A|b)$ e sia $S_0 := \text{Sol}(A|0) = \ker(A)$ la sua giacitura.

Caso $m = 1$. $Ax = b$ è costituito da una singola equazione: $S = \{(x, y)^T \in \mathbb{R}^2 \mid ax + by = c\}$, quindi $\dim S = 1$: S è una *retta*.

Caso $m = 2$. $Ax = b$ è costituito da due equazioni (indipendenti), $\dim S = 0$ e S è un punto:

$$S = \left\{ (x, y)^T \in \mathbb{R}^2 \mid \begin{cases} a_1x + b_1y = c_1 \\ a_2x + b_2y = c_2 \end{cases} \right\}.$$

Come per i sottospazi vettoriali, risolvendo il sistema lineare $Ax = b$, possiamo descrivere il sottospazio affine $S := \text{Sol}(A|b)$ attraverso *equazioni parametriche*. Per il Teorema 1.13 (vedi anche il Teorema 1.11 (Rouché-Capelli)) vale

$$S = \text{Sol}(A|b) = P + \text{Sol}(A|0) = P + \text{Span}(w_1, \dots, w_{n-r})$$

dove $r = \text{rg}(A) = \text{rg}(A|b)$, $P \in \mathbb{K}^n$ è una soluzione di $Ax = b$ e $\{w_1, \dots, w_{n-r}\}$ è una base della giacitura $S_0 = \ker(A)$. In particolare, $\dim S = \dim S_0 = n - \text{rg}(A)$.

Come si fa il viceversa? Cioè date equazioni parametriche $T = \{P + w_1 t_1 + \dots + w_r t_r\} = P + \text{Span}(w_1, \dots, w_r)$ come passiamo a equazioni cartesiane?

Interpretando $W = \text{Span}(w_1, \dots, w_r) \triangleleft \mathbb{K}^n$ come $W = \text{Col}(M)$, (dove $M \in \mathcal{M}_{\mathbb{R}}(n, r)$ ha $w_1, \dots, w_r$ come colonne), vediamo che un punto $Q \in \mathbb{K}^n$ appartiene a T se e solo se $Q - P \in \text{Span}(w_1, \dots, w_r) = W$, se e solo se $\text{rg}(M) = \text{rg}(M|Q - P)$. Riducendo $(M|Q - P)$ con le operazioni elementari sulle righe e imponendo questa condizione otteniamo equazioni cartesiane di T , come nei seguenti esempi.

Esempio. • Sia $T = \begin{pmatrix} 1 \\ -2 \\ -3 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \mathbb{R} + \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \mathbb{R}$, allora $(x, y, z)^T \in T$ se e solo se la matrice

$$\left(\begin{array}{cc|c} 1 & 0 & x-1 \\ 0 & 1 & y+2 \\ 1 & 1 & z+3 \end{array} \right)$$

ha rango 2. Usando operazioni elementari sulle righe, portiamo la matrice nella seguente forma

$$\left(\begin{array}{cc|c} 1 & 0 & x-1 \\ 0 & 1 & y+2 \\ 0 & 0 & z+3-x+1-y-2 \end{array} \right)$$

che ha rango 2 se e solo se $(x, y, z)^T$ è soluzione di $x + y - z = 2$: questa è un'equazione cartesiana di T .

• Sia $T = \begin{pmatrix} 0 \\ 4 \\ -3 \end{pmatrix} + \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \mathbb{R}$, allora $(x, y, z)^T \in T$ se e solo se la seguente matrice ha rango 1:

$$\left(\begin{array}{c|c} 1 & x \\ 2 & y-4 \\ 3 & z+3 \end{array} \right) \longrightarrow \left(\begin{array}{c|c} 1 & x \\ 0 & y-4-2x \\ 0 & z+3-3x \end{array} \right) \text{ quindi } T = \left\{ (x, y, z)^T \in \mathbb{R}^3 \mid \begin{cases} 2x - y = -4 \\ 3x - z = 3 \end{cases} \right\}$$

Osservazione. In alternativa, date le equazioni parametriche, possiamo recuperare le equazioni cartesiane eliminando i parametri, come fatto nell'Introduzione.

Oppure, ricavando prima equazioni cartesiane per la giacitura W di T : $W = \ker(A)$ e poi porre la colonna dei termini noti $b \in \mathcal{M}_{\mathbb{K}}(m, 1)$ pari a $b = A \cdot P$.

Esempio. Come nell'esempio precedente sia $T = \begin{pmatrix} 0 \\ 4 \\ -3 \end{pmatrix} + \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \mathbb{R}$, allora W si ottiene

imponendo $1 = \text{rg} \begin{pmatrix} 1 & x \\ 2 & y \\ 3 & z \end{pmatrix} = \text{rg} \begin{pmatrix} 1 & x \\ 0 & 2x-y \\ 0 & 3x-z \end{pmatrix}$, quindi $W = \ker \begin{pmatrix} 2 & -1 & 0 \\ 3 & 0 & -1 \end{pmatrix}$; e infine

$$b = \begin{pmatrix} 2 & -1 & 0 \\ 3 & 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 4 \\ -3 \end{pmatrix} = \begin{pmatrix} -4 \\ 3 \end{pmatrix}; \quad T = \text{Sol} \left(\begin{array}{ccc|c} 2 & -1 & 0 & -4 \\ 3 & 0 & -1 & 3 \end{array} \right).$$

Retta passante per due punti & piano passante per tre punti

Ricordiamo brevemente alcune cose già viste nel Capitolo 0.

Una retta è determinata in modo univoco da 2 punti (distinti) P, Q su di essa. Infatti, $Q - P$ determina la direzione della retta e il punto P (o Q) ci dà il punto di applicazione: le equazioni parametriche della retta che passa per P e Q sono: $P + (Q - P)t$, $t \in \mathbb{K}$.

Esempio. La retta $l \subset \mathbb{R}^3$ che passa per $P = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ e $Q = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}$ è

$$l = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix} \mathbb{R} \quad \longrightarrow \quad l = \left\{ (x, y, z)^T \in \mathbb{R}^3 \mid \begin{cases} x - y = -1 \\ x - z = -2 \end{cases} \right\}$$

Analogamente, un piano è determinato in modo univoco da 3 punti *non collineari* (cioè non sulla stessa linea) P, Q, R su di esso. Infatti, i vettori $Q - P$ e $R - P$ determinano la giacitura del piano e il punto P (o Q o R) ci dà il punto di applicazione: le equazioni parametriche del piano che passa per P, Q e R sono: $P + (Q - P)t_1 + (R - P)t_2$, $t_1, t_2 \in \mathbb{R}$.

Esempio. Il piano $\pi \subset \mathbb{R}^3$ che passa per $P = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, $Q = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$ e $R = \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}$ è

$$\pi = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ -3 \\ 1 \end{pmatrix} \mathbb{R} + \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix} \mathbb{R} \quad \longrightarrow \quad \pi = \{ (x, y, z)^T \in \mathbb{R}^3 \mid 7x - y - 3z = 6 \}$$

Esempio. I punti $P = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, $Q = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}$ e $R = \begin{pmatrix} 7 \\ 8 \\ 9 \end{pmatrix}$ sono collineari, infatti considerando $P + (Q - P)t_1 + (R - P)t_2$, $t_1, t_2 \in \mathbb{R}$ otteniamo una retta:

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix} \mathbb{R} + \begin{pmatrix} 6 \\ 6 \\ 6 \end{pmatrix} \mathbb{R} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \mathbb{R}$$

Posizioni reciproche di sottospazi

Definizione 2.28. Siano $S_1 = \text{Sol}(A_1|b_1)$ e $S_2 = \text{Sol}(A_2|b_2)$ sottospazi affini di $\mathbb{K}^n$ e siano $S_{1,0} = \ker(A_1)$ e $S_{2,0} = \ker(A_2)$ le loro giaciture.

I sottospazi affini S_1 e S_2 sono *paralleli* se $S_{1,0} \subseteq S_{2,0}$ oppure $S_{2,0} \subseteq S_{1,0}$.

I sottospazi affini S_1 e S_2 sono *incidenti* se hanno punti in comune $S_1 \cap S_2 \neq \emptyset$ e non sono uno contenuto nell'altro: $S_1 \not\subseteq S_2$ o $S_2 \not\subseteq S_1$.

Infine S_1 e S_2 si dicono *sghebbi* se non hanno punti in comune $S_1 \cap S_2 = \emptyset$ e le loro giaciture sono in somma diretta: $S_{1,0} \cap S_{2,0} = \{0\}$.

Esempio.

- In $\mathbb{R}^2$ le rette $r_1 := \{x + y = 2\}$ e $r_2 := \{2x + 2y = -1\}$ sono incidenti, infatti $r_{1,0} = r_{2,0} = \{x + y = 0\}$.
- In $\mathbb{R}^2$ le rette $r_1 := \{x + y = 2\}$ e $r_2 := \{x - y = 0\}$ sono incidenti ($r_1 \cap r_2 = \{(1, 1)^T\}$).
- In $\mathbb{R}^3$ il piano $\pi := \{x - 2y + z = 1\}$ e la retta $r := \{x - z = 2, x - y = -1\}$ sono paralleli, infatti $r_0 = \text{Span}((1, 1, 1)^T) \subseteq \pi_0$.
- In $\mathbb{R}^3$, la retta $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \mathbb{R}$ e il piano $\begin{pmatrix} 4 \\ -1 \\ 7 \end{pmatrix} + \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} \mathbb{R} + \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \mathbb{R}$ sono paralleli, infatti $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \in \text{Span} \left(\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right)$.

- In $\mathbb{R}^3$ le rette $r_1 = \begin{cases} x - y = -1 \\ x - z = -2 \end{cases}$ e $r_2 = \begin{pmatrix} 0 \\ 4 \\ -3 \end{pmatrix} + \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \mathbb{R}$ sono sghembe. Infatti, inserendo le coordinate di un punto di r_2 nelle equazioni di r_1 otteniamo

$$\begin{cases} (0+t) - (4+2t) = -1 \\ (0+t) - (-3+3t) = -2 \end{cases} \leftrightarrow \begin{cases} -t = -5 \\ -2t = 1 \end{cases}$$

che non ha soluzioni, i.e. r_1 e r_2 non hanno punti comuni. Inoltre le giaciture sono in somma diretta, infatti: $r_{1,0} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \mathbb{R}$ mentre $r_{2,0} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$.

- In $\mathbb{R}^4$ i piani $\pi_1 := \{(x, y, z, w)^T \in \mathbb{R}^4 \mid x = y = 0\}$ e $\pi_2 := \{(x, y, z, w)^T \in \mathbb{R}^4 \mid x - 1 = w = 0\}$ non sono né incidenti, né paralleli, né sghembi.

Lemma 2.29. *Siano $S_1 = \text{Sol}(A_1|b_1)$ e $S_2 = \text{Sol}(A_2|b_2)$ sottospazi affini paralleli di $\mathbb{K}^n$. Allora, o $S_1 \cap S_2 = \emptyset$ o $S_1 \subseteq S_2$ o $S_2 \subseteq S_1$.*

Dimostrazione. Se $S_1 \cap S_2 = \emptyset$ siamo a posto.

In caso contrario, prendiamo $P \in S_1 \cap S_2$, e supponiamo $S_{1,0} \subseteq S_{2,0}$, allora per ogni $Q \in S_1$ vale $P - Q \in S_{1,0} \subseteq S_{2,0}$, cioè $Q \in S_2$; in altre parole $S_1 \subseteq S_2$.

Analogamente, se $S_1 \cap S_2 \neq \emptyset$ e $S_{2,0} \subseteq S_{1,0}$, allora $S_2 \subseteq S_1$. $\square$

2.7.1 Sottospazi affini di $\mathbb{R}^3$ e loro posizioni reciproche

Siano $A \in \mathcal{M}_{\mathbb{R}}(m, 3)$ e $b \in \mathbb{R}^m$ con $0 < m = \text{rg}(A) = \text{rg}(A|b) \leq 3$. Sia $S := \text{Sol}(A|b)$ e sia $S_0 := \text{Sol}(A|0) = \ker(A)$ la sua giacitura.

Caso $m = 3$. $Ax = b$ è costituito da tre equazioni (indipendenti), quindi $\dim S = 0$: S è un *punto*.

Caso $m = 2$. $Ax = b$ è costituito da due equazioni (indipendenti), quindi $\dim S = 1$ e S è una *retta*:

$$S = \{(x, y, z)^T \in \mathbb{R}^3 \mid \begin{cases} \alpha_1 x + \beta_1 y + \gamma_1 z = \delta_1 \\ \alpha_2 x + \beta_2 y + \gamma_2 z = \delta_2 \end{cases}\}.$$

Caso $m = 1$. $Ax = b$ è definito da una singola equazione: $S = \{(x, y, z)^T \in \mathbb{R}^3 \mid \alpha x + \beta y + \gamma z = \delta\}$, quindi $\dim S = 2$: S è un *piano*.

Osservazione. Per descrivere un piano in $\mathbb{R}^3$ è sufficiente dare il vettore $(a, b, c)^T$ dei coefficienti di una sua equazione. Tale vettore è unico a meno di moltiplicarlo con un uno scalare non nullo, ed è chiamato *vettore normale* (o *vettore direttore*) al piano.

Nel Capitolo 5 vedremo che geometricamente questo vettore è “ortogonale/perpendicolare” al piano.

Vediamo ora le posizioni reciproche di due sottospazi affini di $\mathbb{R}^3$.

Un punto P può appartenere o meno ad un sottospazio affine $S = \text{Sol}(A|b)$. Più precisamente, $P \in S \Leftrightarrow A \cdot P = b$.

Retta - retta

Siano $l_1 = \text{Sol}(A_1|b_1)$ e $l_2 = \text{Sol}(A_2|b_2)$ due rette in $\mathbb{R}^3$. La prima cosa da verificare è se si incontrano o meno, quindi dobbiamo capire se il sistema lineare

$$\begin{pmatrix} A_1 \\ A_2 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

ha soluzioni. La matrice $A := \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} \in \mathcal{M}_{\mathbb{R}}(4, 3)$ ha rango $\text{rg}(A) = 2$ o 3 , mentre la matrice completa $(A|b) = \begin{pmatrix} A_1 & | & b_1 \\ A_2 & | & b_2 \end{pmatrix} \in \mathcal{M}_{\mathbb{R}}(4, 4)$ ha $\text{rg}(A|b) = \text{rg}(A)$ o $\text{rg}(A|b) = \text{rg}(A) + 1$. Quindi ci sono 4 possibilità per la posizione reciproca di l_1 e l_2 .

Caso 1: $\text{rg}(A) = \text{rg}(A|b) = 2$: cioè i sistemi lineari $A_1x = b_1$ e $A_2x = b_2$ sono equivalenti, quindi $l_1 = l_2$: sono *rette coincidenti*.

Caso 2: $\text{rg}(A) = 2$ e $\text{rg}(A|b) = 3$: per il Teorema di Rouché-Capelli (1.11) le due rette non hanno punti comuni, ma stessa giacitura ($\dim \ker(A) = 1 = \dim \ker(A_1) = \dim \ker(A_2)$ implica $\ker(A_1) = \ker(A_2)$), quindi l_1 e l_2 sono *rette parallele*.

Esempio. Le rette $r_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\mathbb{R}$ e $r_2 = \begin{pmatrix} 4 \\ -1 \\ 7 \end{pmatrix} + \begin{pmatrix} -2 \\ -2 \\ -2 \end{pmatrix}\mathbb{R}$ sono parallele, visto che $\text{Span}\left(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\right) = \text{Span}\left(\begin{pmatrix} -2 \\ -2 \\ -2 \end{pmatrix}\right)$, ma non coincidenti poiché $\begin{pmatrix} 4 \\ -1 \\ 7 \end{pmatrix} \notin r_1$.

Caso 3: $\text{rg}(A) = 3 = \text{rg}(A|b)$: per il Teorema di Rouché-Capelli (1.11) le due rette hanno un unico punto comune: l_1 e l_2 sono *rette incidenti*.

Caso 4: $\text{rg}(A) = 3$ e $\text{rg}(A|b) = 4$: in base al Teorema di Rouché-Capelli (1.11), le due rette non hanno punti comuni e hanno direzioni diverse $\ker(A_1) \neq \ker(A_2)$: l_1 e l_2 sono *rette sghembe*.

Esempio. • $\begin{cases} x - y = -1 \\ x - z = -2 \end{cases}$ e $\begin{cases} -2x + y + z = 3 \\ y - z = -1 \end{cases}$ sono rette coincidenti, infatti

$$\text{rg} \left(\begin{pmatrix} 1 & -1 & 0 & | & -1 \\ 1 & 0 & -1 & | & -2 \\ -2 & 1 & 1 & | & 3 \\ 0 & 1 & -1 & | & -1 \end{pmatrix} \right) = \text{rg} \left(\begin{pmatrix} 1 & -1 & 0 & | & -1 \\ 0 & 1 & -1 & | & -1 \\ 0 & -1 & 1 & | & 1 \\ 0 & 1 & -1 & | & -1 \end{pmatrix} \right) = 2$$

• $\begin{cases} x - y = -1 \\ x - z = -2 \end{cases}$ e $\begin{cases} -2x + y + z = 4 \\ y - z = -2 \end{cases}$ sono rette parallele, infatti

$$\left(\begin{pmatrix} 1 & -1 & 0 & | & -1 \\ 1 & 0 & -1 & | & -2 \\ -2 & 1 & 1 & | & 4 \\ 0 & 1 & -1 & | & -2 \end{pmatrix} \right) \rightarrow \left(\begin{pmatrix} 1 & -1 & 0 & | & -1 \\ 0 & 1 & -1 & | & -1 \\ 0 & -1 & 1 & | & 2 \\ 0 & 1 & -1 & | & -2 \end{pmatrix} \right) \rightarrow \left(\begin{pmatrix} 1 & -1 & 0 & | & -1 \\ 0 & 1 & -1 & | & -1 \\ 0 & 0 & 0 & | & 1 \\ 0 & 0 & 0 & | & 0 \end{pmatrix} \right)$$

• $l_1 = \begin{cases} x - y = -1 \\ x - z = -2 \end{cases}$ e $l_2 = \begin{pmatrix} 0 \\ -1 \\ -2 \end{pmatrix} + \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \mathbb{R}$ sono rette incidenti, infatti sostituendo le equazioni di l_2 in quelle di l_1 otteniamo

$$\begin{cases} (0+t) - (-1+2t) = -1 \\ (0+t) - (-2+3t) = -2 \end{cases} \leftrightarrow \begin{cases} -t = -2 \\ -2t = -4 \end{cases} \leftrightarrow t = 2$$

un'unica soluzione. Ciò significa che le rette l_1 e l_2 hanno un unico punto comune $(2, 3, 4)^T$ (ovvero sono rette incidenti).

• $l_1 = \begin{cases} x - y = -1 \\ x - z = -2 \end{cases}$ e $l_2 = \begin{pmatrix} 0 \\ 4 \\ -3 \end{pmatrix} + \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \mathbb{R}$ sono rette sghembe, infatti sostituendo le equazioni di l_2 in quelle di l_1 otteniamo

$$\begin{cases} (0+t) - (4+2t) = -1 \\ (0+t) - (-3+3t) = -2 \end{cases} \leftrightarrow \begin{cases} -t = -5 \\ -2t = 1 \end{cases}$$

che non ha soluzioni, ovvero l_1 e l_2 non hanno punti comuni. Inoltre, le loro giaciture sono diverse, infatti $l_{1,0} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \mathbb{R}$ non contiene $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$.

Osservazione. a) Quindi se le rette l_1 ed l_2 hanno infiniti punti in comune, sono coincidenti. Se hanno un solo punto in comune sono incidenti.

Se non hanno punti in comune sono o parallele o sghembe, e questo si verifica confrontando le loro giaciture.

b) Se due rette hanno la stessa giacitura, allora o sono coincidenti, o sono parallele (non coincidenti).

Se invece le giaciture sono diverse, allora le rette o sono incidenti o sono sghembe.

c) Avendo equazioni parametriche delle rette $l_1 = P_1 + v_1 \mathbb{R}$ e $l_2 = P_2 + v_2 \mathbb{R}$, possiamo leggere i quattro casi come segue.

Caso 1, *rette coincidenti*: $\text{rg}(v_1|v_2) = 1 = \text{rg}(v_1|v_2|P_2 - P_1)$.

Caso 2, *rette parallele* (non coincidenti): $\text{rg}(v_1|v_2) = 1$, $\text{rg}(v_1|v_2|P_2 - P_1) = 2$.

Caso 3, *rette incidenti*: $\text{rg}(v_1|v_2) = 2 = \text{rg}(v_1|v_2|P_2 - P_1)$.

Caso 4, *rette sghembe*: $\text{rg}(v_1|v_2) = 2$, $\text{rg}(v_1|v_2|P_2 - P_1) = 3$.

Infatti, $\text{rg}(v_1|v_2) = 1$ significa che le rette hanno la stessa direzione, mentre $\text{rg}(v_1|v_2) = 2$ significa che le rette hanno direzioni diverse.

E se, aggiungendo la colonna $P_2 - P_1$, il rango aumenta, significa che il vettore $P_2 - P_1$ è indipendente da $\{v_1, v_2\}$, cosicché le traslazioni di $v_1 \mathbb{R}$ per P_1 e la traslazione di $v_2 \mathbb{R}$ per P_2 separano le rette.

Esempio. • Le rette $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \mathbb{R}$ e $\begin{pmatrix} 4 \\ -1 \\ -2 \end{pmatrix} + \begin{pmatrix} -2 \\ -2 \\ -2 \end{pmatrix} \mathbb{R}$ hanno la stessa direzione, quindi sono coincidenti o parallele. Sono parallele poiché $\begin{pmatrix} 4 \\ -1 \\ -2 \end{pmatrix} - \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 3 \\ -3 \\ -5 \end{pmatrix} \notin \text{Span} \left(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right)$.

• Le rette $\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} + \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \mathbb{R}$ e $\begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} + \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \mathbb{R}$ hanno direzioni diverse, quindi sono o incidenti o sghembe. Poiché $\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} - \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ -4 \\ 1 \end{pmatrix} \in \text{Span} \left(\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \right)$, le rette sono incidenti.

Retta - piano

Sia $l = \text{Sol}(A_1|b_1)$ una retta e $\pi = \text{Sol}(A_2|b_2)$ un piano.

La matrice $A := \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} \in \mathcal{M}_{\mathbb{R}}(3, 3)$ ha rango $\text{rg}(A) = 2$ o 3 , mentre la matrice completa $(A|b) = \begin{pmatrix} A_1 & b_1 \\ A_2 & b_2 \end{pmatrix} \in \mathcal{M}_{\mathbb{R}}(3, 4)$ ha $\text{rg}(A) \leq \text{rg}(A|b) \leq 3$, quindi ci sono 3 casi.

Caso 1: $\text{rg}(A) = \text{rg}(A|b) = 2$: ciò significa che la retta e il piano hanno infiniti punti in comune: la retta è contenuta in π : $l \subset \pi$.

Per essere più precisi, ciò significa che l'equazione cartesiana che definisce il piano è una combinazione lineare delle equazioni che definiscono la retta. Più precisamente, sia $\begin{cases} \alpha_1 x + \beta_1 y + \gamma_1 z = \delta_1 \\ \alpha_2 x + \beta_2 y + \gamma_2 z = \delta_2 \end{cases}$ un'equazione cartesiana della retta, allora esistono $\lambda, \mu \in \mathbb{R}$ (non entrambi zero) tali che π è dato dall'equazione:

$$\lambda(\alpha_1 x + \beta_1 y + \gamma_1 z - \delta_1) + \mu(\alpha_2 x + \beta_2 y + \gamma_2 z - \delta_2) = 0 \quad (2.1)$$

Definizione 2.30. Lasciando variare λ, μ in $\mathbb{R}$ (non entrambi nulli!), l'equazione (2.1) descrive tutti i piani contenenti la retta $l = \begin{cases} \alpha_1 x + \beta_1 y + \gamma_1 z = \delta_1 \\ \alpha_2 x + \beta_2 y + \gamma_2 z = \delta_2 \end{cases}$. Questo è chiamato *fascio di piani* passanti per l .

Caso 2: $\text{rg}(A) = 2$ e $\text{rg}(A|b) = 3$: ciò significa che la retta e il piano non hanno punti comuni e che la giacitura della retta $\ker(A_1)$ è contenuta nella giacitura del piano $\ker(A_2)$, quindi la retta e il piano sono *paralleli* e $l \not\subset \pi$.

Caso 3: $\text{rg}(A) = 3 = \text{rg}(A|b)$: la retta e il piano hanno un unico punto comune, quindi la retta e il piano sono *incidenti*.

Osservazione. Analogamente al caso retta-retta, avendo equazioni parametriche della retta $l = P_1 + v_1\mathbb{R}$ e del piano $\pi = P_2 + v_2\mathbb{R} + v_3\mathbb{R}$, possiamo distinguere i tre casi:

Caso 1, $l \subset \pi$: $\text{rg}(v_1|v_2|v_3) = 2 = \text{rg}(v_1|v_2|v_3|P_2 - P_1)$.

Caso 2, *retta e piano sono paralleli* (e $l \not\subset \pi$): $\text{rg}(v_1|v_2|v_3) = 2$ e $\text{rg}(v_1|v_2|v_3|P_2 - P_1) = 3$.

Caso 3, *la retta e il piano sono incidenti*: $\text{rg}(v_1|v_2|v_3) = 3$.

Esempio. • La retta $l = \begin{cases} x - y = -1 \\ x - z = -2 \end{cases}$ è contenuta nel piano $\pi = \{x + y - 2z = -3\}$, infatti

$$\text{rg} \left(\begin{array}{ccc|c} 1 & -1 & 0 & -1 \\ 1 & 0 & -1 & -2 \\ 1 & 1 & -2 & -3 \end{array} \right) = \text{rg} \left(\begin{array}{ccc|c} 1 & -1 & 0 & -1 \\ 0 & 1 & -1 & -1 \\ 0 & 2 & -2 & -2 \end{array} \right) = 2$$

• La retta $l = \begin{cases} x - y = -1 \\ x - z = -2 \end{cases}$ e il piano $\pi = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}\mathbb{R} + \begin{pmatrix} 1 \\ -3 \\ -3 \end{pmatrix}\mathbb{R}$ sono incidenti,

infatti sostituendo le equazioni di π in quelle di l otteniamo un'unica soluzione (per s, t)

$$\begin{cases} (1 + s + t) - (0 + 2s + t) = -1 \\ (1 + s + t) - (1 + 3s - 3t) = -2 \end{cases} \Leftrightarrow \begin{cases} s = 2 \\ -2s + 4t = -2 \end{cases} \Leftrightarrow \begin{cases} s = 2 \\ t = \frac{1}{2} \end{cases}$$

• La retta $l = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}\mathbb{R}$ e il piano $\pi = x + y - z = 10$ sono paralleli, infatti sostituendo le equazioni di l in quelle di π otteniamo un'equazione impossibile: $10 = (1 + t) + (0 + 2t) - (1 + 3t) = 2$.

Piano - piano

Siano $\pi_1 = \text{Sol}(A_1|b_1)$ e $\pi_2 = \text{Sol}(A_2|b_2)$ piani.

La matrice $A := \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} \in \mathcal{M}_{\mathbb{R}}(2, 3)$ ha rango $\text{rg}(A) = 1$ o 2 , mentre la matrice completa $(A|b) = \begin{pmatrix} A_1 & b_1 \\ A_2 & b_2 \end{pmatrix} \in \mathcal{M}_{\mathbb{R}}(2, 4)$ ha $\text{rg}(A) \leq \text{rg}(A|b) \leq 2$, quindi ci sono 3 casi.

Caso 1: $\text{rg}(A) = \text{rg}(A|b) = 1$: ciò significa che i sistemi lineari $A_1x = b_1$ e $A_2x = b_2$ sono equivalenti, quindi $\pi_1 = \pi_2$: sono *piani coincidenti*.

Caso 2: $\text{rg}(A) = 1$ e $\text{rg}(A|b) = 2$: per il Teorema di Rouché-Capelli (1.11) i piani non hanno punti comuni, ma hanno la stessa direzione ($\dim \ker(A) = 2 = \dim \ker(A_1) = \dim \ker(A_2)$ implica $\ker(A_1) = \ker(A_2)$) quindi π_1 e π_2 sono *piani paralleli*.

Per essere più precisi, entrambi i sistemi lineari che definiscono i piani consistono in una singola equazione, e il fatto che $\text{rg}(A) = 1$ significa che le equazioni sono della forma:

$$\pi_1 : \alpha x + \beta y + \gamma z = \delta_1 \quad \pi_2 : \alpha x + \beta y + \gamma z = \delta_2$$

Quindi se $\delta_1 = \delta_2$ i piani coincidono, altrimenti sono paralleli.

Definizione 2.31. Sia π il piano dell'equazione

$$\alpha x + \beta y + \gamma z = \delta.$$

Lasciando variare δ in $\mathbb{R}$, descriviamo tutti i piani paralleli a π e abbiamo un *fascio di piani paralleli* o *fascio improprio di piani*.

Caso 3: $\text{rg}(A) = 2 = \text{rg}(A|b)$: in base al Teorema di Rouché-Capelli (1.11) hanno una retta comune: π_1 e π_2 sono *piani incidenti*.

Osservazione. Come sopra, avendo le equazioni parametriche dei piani $\pi_1 = P_1 + v_1\mathbb{R} + v_2\mathbb{R}$ e $\pi_2 = P_2 + w_1\mathbb{R} + w_2\mathbb{R}$, possiamo leggere i tre casi come segue:

Caso 1, *piani coincidenti*: $\text{rg}(v_1|v_2|w_1|w_2) = 2 = \text{rg}(v_1|v_2|w_1|w_2|P_2 - P_1)$.

Caso 2, *piani paralleli* (non coincidenti): $\text{rg}(v_1|v_2|w_1|w_2) = 2$ e $\text{rg}(v_1|v_2|w_1|w_2|P_2 - P_1) = 3$.

Caso 3, *piani incidenti*: $\text{rg}(v_1|v_2|w_1|w_2) = 3$.

Capitolo 3

Applicazioni lineari

Nel capitolo precedente abbiamo introdotto gli spazi vettoriali. Qui introdurremo le funzioni tra spazi vettoriali, che ne preservano la struttura, cioè le applicazioni (o mappe o funzioni) lineari.

Prima di introdurre le applicazioni lineari e le loro proprietà, ricordiamo brevemente alcune definizioni generali.

Definizione 3.1. Sia $f : X \rightarrow Y$ una funzione.

La funzione f è *suriettiva* se “ogni elemento nel codominio può essere raggiunto”:

$$\forall y \in Y \exists x \in X : y = f(x).$$

La funzione f è *iniettiva* se “a elementi diversi nel dominio corrispondono elementi diversi nel codominio”:

$$\forall x_1, x_2 \in X : f(x_1) = f(x_2) \Rightarrow x_1 = x_2$$

La funzione f è *biiettiva* se è sia suriettiva che iniettiva:

$$\forall y \in Y \exists! x \in X : y = f(x)$$

“Ogni elemento nel codominio è raggiunto da un solo elemento del dominio”.

Se $f : X \rightarrow Y$ è biiettiva (o biunivoca), è *invertibile*, infatti ad ogni $y \in Y$ è associato a un unico $x \in X$. Quindi, $y \mapsto x$, dove x è l'unico elemento di X mappato su y , definisce la mappa inversa $f^{-1} : Y \rightarrow X$. Essa soddisfa:

$$f \circ f^{-1} = id_Y, \quad f^{-1} \circ f = id_X.$$

Esempio. La mappa $\mathbb{R} \rightarrow [-1, 1], t \mapsto \sin(t)$ è suriettiva, ma non iniettiva.

La mappa $\mathbb{R} \rightarrow \mathbb{R}^2, t \mapsto (2t, -t)$ è iniettiva, ma non suriettiva.

La mappa $\mathbb{R} \rightarrow \mathbb{R}, t \mapsto 3t$ è sia suriettiva che iniettiva, quindi è biunivoca.

La mappa $\mathbb{R} \rightarrow \mathbb{R}^2, t \mapsto (\cos(t), \sin(t))$ non è né suriettiva né iniettiva.

In questo corso ci concentreremo sulle funzioni tra spazi vettoriali che preservano la struttura dello spazio vettoriale, in particolare quelle che mappano sottospazi vettoriali su sottospazi vettoriali come nel secondo e terzo esempio, ma non nell'ultimo, poiché quest'ultimo mappa la retta reale sul cerchio unitario.

3.1 Applicazioni lineari: definizione e proprietà base

Definizione 3.2. Siano V, W spazi vettoriali sullo stesso campo $\mathbb{K}$.

Una funzione $f : V \rightarrow W$ è un'applicazione lineare se

- L1) $f(v_1 + v_2) = f(v_1) + f(v_2)$ per tutti $v_1, v_2 \in V$ (additività)
- L2) $f(\lambda \cdot v) = \lambda \cdot f(v)$ per tutti $v \in V, \lambda \in \mathbb{K}$ (omogeneità)

Esempio.

- La mappa nulla $V \rightarrow W, v \mapsto 0_W$ è lineare.
- La mappa identità $id_V : V \rightarrow V, v \mapsto v$ è lineare.
- Sia $A \in \mathcal{M}_{\mathbb{K}}(m, n)$. La mappa $\mathcal{L}_A : \mathbb{K}^n \rightarrow \mathbb{K}^m, v \mapsto Av$ è lineare:

$$\mathcal{L}_A(v_1 + v_2) = A(v_1 + v_2) = Av_1 + Av_2 = \mathcal{L}_A(v_1) + \mathcal{L}_A(v_2), \quad \mathcal{L}_A(\lambda \cdot v) = A(\lambda \cdot v) = \lambda Av = \lambda \mathcal{L}_A(v)$$

- La trasposizione $\mathcal{M}_{\mathbb{K}}(m, n) \rightarrow \mathcal{M}_{\mathbb{K}}(n, m), A \mapsto A^T$ è lineare.
- La “derivata” $\mathbb{K}[t]_{\leq d} \rightarrow \mathbb{K}[t]_{\leq d}, \sum_{n=0}^d a_n t^n \mapsto \sum_{n=1}^d n a_n t^{n-1}$ è lineare.
- La *traccia* di una matrice quadrata $\text{tr} : \mathcal{M}_{\mathbb{K}}(n) \rightarrow \mathbb{K}, A \mapsto \text{tr}(A) = a_{1,1} + a_{2,2} + \cdots + a_{n,n}$ (somma delle entrate sulla diagonale) è una mappa lineare.
- Per $n > 1$, il determinante $\det : \mathcal{M}_{\mathbb{K}}(n) \rightarrow \mathbb{K}, A \mapsto \det(A)$ non è una mappa lineare: $\det(cA) = c^n \det A$.
- $f : \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto x^2 + y^2$ non è una mappa lineare, per esempio $f(2 \cdot (1, 1)) = 8 \neq 4 = 2 \cdot f(1, 1)$.

Osservazione. Le condizioni L1, L2 possono essere verificate simultaneamente dimostrando che

$$f(\lambda v_1 + \mu v_2) = \lambda f(v_1) + \mu f(v_2) \quad \forall v_1, v_2 \in V, \lambda, \mu \in \mathbb{K}.$$

Lemma 3.3. Sia $f : V \rightarrow W$ una mappa lineare tra $\mathbb{K}$ -spazi vettoriali $\mathbb{K}$.

Allora $f(0_V) = 0_W$.

Dimostrazione. $f(0_V) = f(0_{\mathbb{K}} \cdot v) = 0_{\mathbb{K}} \cdot f(v) = 0_W$. □

In particolare, se $f(0_V) \neq 0_W$, allora $f : V \rightarrow W$ non è una mappa lineare, ad esempio $e^0 = 1 \neq 0$, quindi $\exp : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto e^x$ non è lineare.

Attenzione, questa condizione è necessaria, ma non sufficiente per la linearità: $f : \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto x^2 + y^2$ non è lineare, ma $f(0, 0) = 0$.

Proposizione 3.4. Siano V, W, U spazi vettoriali su $\mathbb{K}$ e siano $f : V \rightarrow W, g : W \rightarrow U$ applicazioni lineari. Allora

1. La composizione $g \circ f : V \rightarrow U, v \mapsto g(f(v))$ è lineare.
2. Se $f : V \rightarrow W$ è invertibile (cioè una funzione biunivoca), allora $f^{-1} : W \rightarrow V$ è lineare.

Dimostrazione. 1. $g(f(\lambda v_1 + \mu v_2)) = g(\lambda f(v_1) + \mu f(v_2)) = \lambda g(f(v_1)) + \mu g(f(v_2))$.

2. Dobbiamo dimostrare che $f^{-1}(\lambda w_1 + \mu w_2) = \lambda f^{-1}(w_1) + \mu f^{-1}(w_2)$. Definiamo $v_1 = f^{-1}(w_1)$ e $v_2 = f^{-1}(w_2)$, così che $w_1 = f(v_1)$ e $w_2 = f(v_2)$. Sfruttando $f^{-1} \circ f = id_V$ e la linearità di f otteniamo

$$f^{-1}(\lambda w_1 + \mu w_2) = f^{-1}(\lambda f(v_1) + \mu f(v_2)) = f^{-1}(f(\lambda v_1 + \mu v_2)) = \lambda v_1 + \mu v_2 = \lambda f^{-1}(w_1) + \mu f^{-1}(w_2).$$

□

Caso speciale. 1. Siano $A \in \mathcal{M}_{\mathbb{K}}(m, p)$ e $B \in \mathcal{M}_{\mathbb{K}}(p, n)$. Allora $\mathcal{L}_A : \mathbb{K}^p \rightarrow \mathbb{K}^m$ e $\mathcal{L}_B : \mathbb{K}^n \rightarrow \mathbb{K}^p$, quindi le possiamo comporre:

$$\begin{array}{ccccc} \mathcal{L}_A \circ \mathcal{L}_B : & \mathbb{K}^n & \longrightarrow & \mathbb{K}^p & \longrightarrow & \mathbb{K}^m \\ & v & \mapsto & Bv & & \\ & & & w & \mapsto & Aw \end{array}$$

cioè $\mathcal{L}_A \circ \mathcal{L}_B(v) = A(Bv) = (AB)v = \mathcal{L}_{AB}(v)$, e quindi $\mathcal{L}_A \circ \mathcal{L}_B = \mathcal{L}_{AB}$ (si osservi che $AB \in \mathcal{M}_{\mathbb{K}}(m, n)$).

2. Se ipotizziamo che $A \in \mathcal{M}_{\mathbb{K}}(n)$ sia invertibile (ovvero $\det A \neq 0$), allora $\mathcal{L}_A$ è biunivoca/invertibile $(\mathcal{L}_A)^{-1} = \mathcal{L}_{A^{-1}}$.

“Le mappe lineari si comportano bene con i sottospazi vettoriali”.

Lemma 3.5. Siano V, W spazi vettoriali $\mathbb{K}$ e sia $f : V \rightarrow W$ una mappa lineare.

1. Se $U \triangleleft V$, allora $f(U) = \{w \in W \mid \exists u \in U \text{ s.t. } w = f(u)\} \triangleleft W$.
2. Se $Z \triangleleft W$, allora $f^{-1}(Z) = \{v \in V \mid f(v) \in Z\} \triangleleft V$.

Dimostrazione. Esercizio.

□

Consideriamo ora due casi che saranno rilevanti in questo corso.

Definizione 3.6. Siano V, W $\mathbb{K}$ -spazi vettoriali e sia $f : V \rightarrow W$ un'applicazione lineare. Il *nucleo* (o *kernel*) di f è

$$\ker(f) = f^{-1}(\{0_W\}) = \{v \in V \mid f(v) = 0_W\} \triangleleft V.$$

L'*immagine* di f è

$$\text{Im}(f) = f(V) = \{w \in W \mid \exists v \in V \text{ s.t. } w = f(v)\} \triangleleft W.$$

Osservazione. Sia $A \in \mathcal{M}_{\mathbb{K}}(m, n)$. Allora $\mathcal{L}_A : \mathbb{K}^n \rightarrow \mathbb{K}^m$, e

$$\ker(\mathcal{L}_A) = \{v \in \mathbb{K}^n \mid \mathcal{L}_A(v) = 0\} = \{v \in \mathbb{K}^n \mid Av = 0\} = \text{Sol}(A|0) = \ker(A)$$

mentre l'immagine di A è $\text{Im}(A) := \text{Im}(\mathcal{L}_A)$. Quindi dato $b \in \mathbb{K}^m$, si ha che

$$b \in \text{Im}(A) \Leftrightarrow \exists v \in \mathbb{K}^n \text{ s.t. } Av = b \Leftrightarrow Ax = b \text{ è risolubile} \Leftrightarrow b \in \text{Col}(A)$$

cioè $\text{Im}(A) = \text{Col}(A)$, in particolare $\dim \text{Im}(A) = \text{rg}(A)$.

In analogia, definiamo il rango di una mappa lineare.

Definizione 3.7. Siano V, W spazi vettoriali sul campo $\mathbb{K}$ e sia $f : V \rightarrow W$ un'applicazione lineare. Il *rank* di f è $\text{rg}(f) := \dim \text{Im}(f)$.

Esempio. Sia $f : \mathbb{K}[t]_{\leq 2} \rightarrow \mathbb{K}[t]_{\leq 1}, p(t) \mapsto tp''(t) - p'(t)$.

Verifichiamo innanzitutto che f è lineare utilizzando il fatto che la “derivata” è lineare:

$$\begin{aligned} f(\lambda p(t) + \mu q(t)) &= t(\lambda p(t) + \mu q(t))'' - (\lambda p(t) + \mu q(t))' = t(\lambda p''(t) + \mu q''(t)) - (\lambda p'(t) + \mu q'(t)) \\ &= \lambda(tp''(t) - p'(t)) + \mu(tq''(t) - q'(t)) = \lambda f(p(t)) + \mu f(q(t)) \end{aligned}$$

Sia $p(t) = a_0 + a_1t + a_2t^2$, allora $f(p(t)) = 2a_2t - (a_1 + 2a_2t) = -a_1$, quindi

$$\ker(f) = \{p(t) \in \mathbb{K}[t]_{\leq 2} \mid tp''(t) - p'(t) = 0\} = \{a_0 + a_2t^2 \mid a_0, a_2 \in \mathbb{K}\} = \text{Span}(1, t^2)$$

$$\text{Im}(f) = \{-a_1 \mid a_1 \in \mathbb{K}\} = \text{Span}(1).$$

3.1.1 L'insieme delle mappe lineari è uno spazio vettoriale

Siano V, W spazi vettoriali sul campo $\mathbb{K}$.

L'insieme $F(V, W) = \{f : V \rightarrow W\}$ delle funzioni da V a W può essere equipaggiato con 2 operazioni interne come segue. Siano $f, g \in F(V, W)$ e $\lambda \in \mathbb{K}$, allora

$$\begin{aligned} f + g : V &\rightarrow W, (f + g)(v) = f(v) + g(v) \\ \lambda f : V &\rightarrow W, (\lambda f)(v) = \lambda(f(v)) \end{aligned}$$

Si può vedere che $F(V, W)$ con queste 2 operazioni è uno spazio vettoriale (su $\mathbb{K}$).

Definizione 3.8. Siano V, W spazi vettoriali sul campo $\mathbb{K}$. L'insieme delle applicazioni lineari $f : V \rightarrow W$ verrà indicato con $\text{Hom}(V, W) = \{f : V \rightarrow W \mid f \text{ è lineare}\}$.

Se $V = W$, allora un'applicazione lineare $f : V \rightarrow V$ è chiamata *endomorfismo*, e $\text{End}(V) := \text{Hom}(V, V)$.

Lemma 3.9. $\text{Hom}(V, W)$ è un sottospazio vettoriale di $F(V, W)$.

Dimostrazione. Esercizio. □

3.2 Suriettività ed iniettività delle mappe lineari

Proposizione 3.10. Siano V, W $\mathbb{K}$ -spazi vettoriali e $f \in \text{Hom}(V, W)$.

Se $V = \text{Span}(v_1, \dots, v_n)$, allora $\text{Span}(f(v_1), \dots, f(v_n)) = \text{Im}(f)$.

In particolare, $\dim \text{Im}(f) \leq \dim V$.

Dimostrazione. Sia $w \in \text{Im}(f)$, quindi esiste $v \in V$ tale che $f(v) = w$. Da $V = \text{Span}(v_1, \dots, v_n)$ deduciamo che esistono $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ tali che $v = \lambda_1v_1 + \dots + \lambda_nv_n$. Quindi

$$w = f(v) = f(\lambda_1v_1 + \dots + \lambda_nv_n) = \lambda_1f(v_1) + \dots + \lambda_nf(v_n) \in \text{Span}(f(v_1), \dots, f(v_n)).$$

In particolare, se $\{v_1, \dots, v_n\}$ è una base di V , otteniamo $\dim \text{Im}(f) \leq n = \dim V$. □

Osservazione. Quindi un'applicazione lineare può solo “trasformare” uno spazio vettoriale in uno di dimensione uguale o inferiore, ma non in uno di dimensione maggiore; per esempio non può “trasformare” una retta in un piano.

3.2.1 Suriettività

Teorema 3.11 (Criterio di suriettività per le mappe lineari). *Siano V, W spazi vettoriali sul campo $\mathbb{K}$ e sia $f : V \rightarrow W$ una mappa lineare. Allora f è suriettiva se e solo se $\text{rg}(f) = \dim W$.*

Dimostrazione. f è suriettiva se e solo se $\text{Im}(f) = W$. Poiché $\text{Im}(f) \triangleleft W$, l'affermazione segue dal Teorema 2.16. $\square$

“Le mappe lineari suriettive conservano la proprietà di essere generatori.”

Lemma 3.12. *Siano V, W spazi vettoriali sul campo $\mathbb{K}$, $f \in \text{Hom}(V, W)$ e $\{v_1, \dots, v_n\}$ un insieme di generatori di V .*

Allora f è suriettiva, se e solo se $\text{Span}(f(v_1), \dots, f(v_n)) = W$.

Dimostrazione. Dalla Proposizione 3.10 sappiamo che $\text{Span}(f(v_1), \dots, f(v_n)) = \text{Im}(f)$. $\square$

Esempio. i) Sia $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \in \mathcal{M}_{\mathbb{R}}(2, 3)$. La mappa $\mathcal{L}_A : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ è suriettiva? $\{e_1, e_2, e_3\}$ generano $\mathbb{R}^3$, quindi dalla proposizione precedente sappiamo che

$$\text{Im}(\mathcal{L}_A) = \text{Span}(\mathcal{L}_A(e_1), \mathcal{L}_A(e_2), \mathcal{L}_A(e_3)) = \text{Span}\left(\begin{pmatrix} 1 \\ 4 \end{pmatrix}, \begin{pmatrix} 2 \\ 5 \end{pmatrix}, \begin{pmatrix} 3 \\ 6 \end{pmatrix}\right) = \text{Col}(A) = \mathbb{R}^2$$

Quindi $\mathcal{L}_A$ è suriettiva.

ii) Consideriamo $B = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \in \mathcal{M}_{\mathbb{R}}(3, 2)$. La mappa $\mathcal{L}_B : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ è suriettiva?

Come sopra, sappiamo che

$$\text{Im}(\mathcal{L}_B) = \text{Span}(\mathcal{L}_B(e_1), \mathcal{L}_B(e_2)) = \text{Span}\left(\begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix}, \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix}\right) = \text{Col}(B)$$

Da $\dim \text{Col}(B) = \text{rg}(B) = 2 < 3 = \dim \mathbb{R}^3$, segue che la mappa $\mathcal{L}_B$ non è suriettiva.

iii) Sia $C = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{pmatrix} \in \mathcal{M}_{\mathbb{R}}(2, 3)$. La mappa $\mathcal{L}_C : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ è suriettiva?

Come sopra

$$\text{Im}(\mathcal{L}_C) = \text{Span}(\mathcal{L}_C(e_1), \mathcal{L}_C(e_2), \mathcal{L}_C(e_3)) = \text{Span}\left(\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 4 \end{pmatrix}, \begin{pmatrix} 3 \\ 6 \end{pmatrix}\right) = \text{Span}\left(\begin{pmatrix} 1 \\ 2 \end{pmatrix}\right) \neq \mathbb{R}^2$$

Quindi $\mathcal{L}_C$ non è suriettiva.

Osservazione. Per una matrice $A \in \mathcal{M}_{\mathbb{K}}(m, n)$, la mappa $\mathcal{L}_A : \mathbb{K}^n \rightarrow \mathbb{K}^m$ è suriettiva se e solo se $\dim \text{Im } A = m$, ovvero se e solo se $m = \dim \text{Col}(A) = \text{rg}(A)$. Ma $\text{rg}(A) \leq \min(m, n)$, quindi $m \leq n$ è una condizione necessaria (ma non sufficiente) affinché $\mathcal{L}_A$ sia suriettiva.

3.2.2 Iniettività

Teorema 3.13 (Criterio di iniettività per le mappe lineari). *Siano V, W spazi vettoriali sul campo $\mathbb{K}$ e $f \in \text{Hom}(V, W)$. Allora f è iniettiva se e solo se $\ker(f) = \{0_V\}$.*

Dimostrazione. $\Rightarrow$] Supponiamo che f sia iniettiva e che $v \in \ker f$, allora $f(v) = 0_W = f(0_V)$. Poiché f è iniettiva, deduciamo che $v = 0_V$, quindi $\ker(f) = \{0_V\}$.

$\Leftarrow$] Supponiamo che $\ker(f) = \{0_V\}$, e che $v_1, v_2 \in V$ siano vettori con $f(v_1) = f(v_2)$, allora $0_W = f(v_1) - f(v_2) = f(v_1 - v_2)$, quindi $v_1 - v_2 \in \ker(f) = \{0_V\}$. Quindi $v_1 = v_2$. $\square$

“Le mappe lineari iniettive conservano la proprietà di essere linearmente indipendenti.”

Proposizione 3.14. *Siano V, W $\mathbb{K}$ -spazi vettoriali e sia $f : V \rightarrow W$ un'applicazione lineare iniettiva.*

Se $\{v_1, \dots, v_n\} \subset V$ sono linearmente indipendenti, allora $\{f(v_1), \dots, f(v_n)\} \subset W$ sono linearmente indipendenti.

Dimostrazione. Scriviamo 0_W come combinazione lineare di $\{f(v_1), \dots, f(v_n)\}$:

$$0_W = \lambda_1 f(v_1) + \dots + \lambda_n f(v_n) = f(\lambda_1 v_1 + \dots + \lambda_n v_n) \implies \lambda_1 v_1 + \dots + \lambda_n v_n \in \ker(f) = \{0_V\}$$

quindi $\lambda_1 v_1 + \dots + \lambda_n v_n = 0_V$. Utilizzando il fatto che $\{v_1, \dots, v_n\} \subset V$ sono linearmente indipendenti, otteniamo $\lambda_1 = \dots = \lambda_n = 0_{\mathbb{K}}$. $\square$

Esempio. i) Sia $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \in \mathcal{M}_{\mathbb{R}}(2, 3)$. La mappa $\mathcal{L}_A : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ è iniettiva se e solo se $0 = \dim \ker(A)$. Per il teorema di Rouché-Capelli, le soluzioni del sistema lineare $Ax = 0$ sono uno spazio vettoriale di dimensione $3 - \text{rg}(A) = 3 - 2 = 1$, quindi $\mathcal{L}_A$ non è iniettiva.

ii) Consideriamo $B = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \in \mathcal{M}_{\mathbb{R}}(3, 2)$. Le soluzioni del sistema lineare $Bx = 0$

formano uno spazio vettoriale di dimensione $2 - \text{rg}(B) = 2 - 2 = 0$, quindi $\mathcal{L}_B : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ è iniettiva.

iii) Consideriamo $D = \begin{pmatrix} 1 & 2 \\ 2 & 4 \\ 3 & 6 \end{pmatrix} \in \mathcal{M}_{\mathbb{R}}(3, 2)$. Le soluzioni del sistema lineare $Dx = 0$

formano uno spazio vettoriale di dimensione $2 - \text{rg}(D) = 2 - 1 = 1$, quindi $\mathcal{L}_D : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ non è iniettiva.

Osservazione. Per una matrice $A \in \mathcal{M}_{\mathbb{K}}(m, n)$, la mappa $\mathcal{L}_A : \mathbb{K}^n \rightarrow \mathbb{K}^m$ è iniettiva se e solo se $\dim \ker A = 0$. Per il Teorema di Rouché-Capelli deve valere $n = \text{rg}(A)$. Ma $\text{rg}(A) \leq \min(m, n)$, quindi $n \leq m$ è una condizione necessaria (ma non sufficiente) affinché $\mathcal{L}_A$ sia iniettiva.

3.2.3 Biettività

Terminologia. Una mappa lineare biettiva (o biunivoca) $f : V \rightarrow W$ è detta *isomorfismo*, e si dice che V e W sono isomorfi.

Una funzione è biettiva se è sia suriettiva che iniettiva, quindi un isomorfismo preserva sia la proprietà di essere generatori che la proprietà di essere linearmente indipendenti. In altre parole: “le mappe lineari biettive preservano le basi”.

Proposizione 3.15. *Siano V, W $\mathbb{K}$ -spazi vettoriali e sia $f : V \rightarrow W$ un'applicazione lineare biunivoca.*

Se $\{v_1, \dots, v_n\}$ è una base di V , allora $\{f(v_1), \dots, f(v_n)\}$ è una base di W , in particolare $\dim V = \dim W$.

Dimostrazione. Segue da Proposizione 3.10 e Proposizione 3.14. □

Corollario 3.16. *Sia $A \in \mathcal{M}_{\mathbb{K}}(m, n)$ tale che $\mathcal{L}_A$ è biunivoca, allora $m = n$.*

Osservazione. Sia $A \in \mathcal{M}_{\mathbb{K}}(m, n)$ una matrice. Se $m \neq n$, allora $\mathcal{L}_A : \mathbb{K}^n \rightarrow \mathbb{K}^m$ non è invertibile, e quindi anche A non è invertibile. Ecco perché nella Sezione 1.2.4 ci siamo limitati alle *matrici quadrate*.

3.3 Teorema nullità più rango e sue conseguenze

Teorema 3.17 (Teorema nullità più rango). *Siano V, W spazi vettoriali su un campo $\mathbb{K}$ e sia $f : V \rightarrow W$ un'applicazione lineare. Allora*

$$\dim \ker f + \dim \operatorname{Im}(f) = \dim V.$$

Dimostrazione. Sia $\mathcal{B} = \{b_1, \dots, b_p\}$ una base di $\ker f$ (ricordiamo che $\mathcal{B} = \emptyset$ se $\ker f = \{0\}$) e completiamola base di V :

$$\mathcal{B}' = \{b_1, \dots, b_p, b_{p+1}, \dots, b_n\}.$$

Dalla Proposizione 3.10, sappiamo che

$$\begin{aligned} \operatorname{Im}(f) &= \operatorname{Span}(f(b_1), \dots, f(b_p), f(b_{p+1}), \dots, f(b_n)) = \operatorname{Span}(0_W, \dots, 0_W, f(b_{p+1}), \dots, f(b_n)) \\ &= \operatorname{Span}(f(b_{p+1}), \dots, f(b_n)). \end{aligned}$$

Quindi $\{f(b_{p+1}), \dots, f(b_n)\}$ genera $\operatorname{Im}(f)$. Se dimostriamo che sono linearmente indipendenti, ne formano una base, quindi $\dim \operatorname{Im}(f) = n - p$ e abbiamo finito.

Siano quindi $\lambda_{p+1}, \dots, \lambda_n \in \mathbb{K}$ degli scalari tali che

$$0_W = \lambda_{p+1}f(b_{p+1}) + \dots + \lambda_n f(b_n) = f(\lambda_{p+1}b_{p+1} + \dots + \lambda_n b_n)$$

quindi $\lambda_{p+1}b_{p+1} + \dots + \lambda_n b_n \in \ker(f)$.

Ma $\ker(f) = \operatorname{Span}(b_1, \dots, b_p)$, quindi esistono $\lambda_1, \dots, \lambda_p \in \mathbb{K}$ tali che

$$\lambda_{p+1}b_{p+1} + \dots + \lambda_n b_n = \lambda_1 b_1 + \dots + \lambda_p b_p$$

Otteniamo $\lambda_1 b_1 + \dots + \lambda_p b_p - \lambda_{p+1}b_{p+1} - \dots - \lambda_n b_n = 0_V$, ma $\mathcal{B}'$ è un insieme di vettori linearmente indipendenti, quindi $\lambda_1 = \dots = \lambda_p = \lambda_{p+1} = \dots = \lambda_n = 0_{\mathbb{K}}$. □

Esempio. Sia $f : \mathbb{K}[t]_{\leq 2} \rightarrow \mathbb{K}[t]_{\leq 1}, p(t) \mapsto tp''(t) - p'(t)$ (vedi Esercizio a pagina 55).

Abbiamo visto che $\ker(f) = \operatorname{Span}(1, t^2)$ e $\operatorname{Im}(f) = \operatorname{Span}(1)$, quindi $\dim \ker(f) = 2$ e $\dim \operatorname{Im}(f) = 1$.

Osservazione. Sia $A \in \mathcal{M}_{\mathbb{K}}(m, n)$, allora il Teorema 3.17 applicato a $\mathcal{L}_A : \mathbb{K}^n \rightarrow \mathbb{K}^m$ ci dice

$$\dim \text{Sol}(A|0) = \dim \ker(\mathcal{L}_A) = n - \dim \text{Im}(A) = n - \text{rg}(A),$$

quindi lo spazio delle soluzioni di un sistema lineare risolubile dipende da $n - \text{rg}(A)$ parametri, come già sapevamo dalla parte 2. del Teorema 1.11 (Rouché-Capelli).

Osservazione. Sia $f \in \text{Hom}(V, W)$, il teorema di rango-nullità ha le seguenti conseguenze (già dedotte in precedenza per $\mathcal{L}_A$, $A \in \mathcal{M}_{\mathbb{K}}(m, n)$).

1. Se f è suriettiva, allora $\dim \text{Im}(f) = \dim W$, quindi $\dim W = \dim V - \dim \ker f \leq \dim V$. In altre parole, se $\dim V < \dim W$, allora f non può essere suriettiva.
2. Se f è iniettiva, allora $\dim \ker(f) = 0$, quindi $\dim V = \dim \text{Im}(f) \leq \dim W$. In altre parole, se $\dim V > \dim W$, allora f non può essere iniettiva.
3. Se f è biunivoca, allora $\dim V = \dim W$. In altre parole, se $\dim V \neq \dim W$, allora f non può essere biettiva (vedi anche Proposizione 3.15).

Proposizione 3.18. *Siano V e W spazi vettoriali sul campo $\mathbb{K}$ della stessa dimensione $\dim V = \dim W$ e sia $f : V \rightarrow W$ una mappa lineare.*

Allora f è iniettiva se e solo se f è suriettiva se e solo se f è biettiva.

Dimostrazione. Sia f iniettiva, allora $\dim \ker(f) = 0 \Rightarrow \dim W = \dim V \stackrel{TNR}{=} \dim \text{Im}(f)$ e quindi f è suriettiva.

Sia f suriettiva, allora $\dim V = \dim W = \dim \text{Im}(f) \Leftrightarrow \dim \ker(f) \stackrel{TNR}{=} \dim V - \dim \text{Im}(f) = 0$ e quindi f è iniettiva. $\square$

Ne segue un'altra caratterizzazione delle matrici invertibili:

Corollario 3.19. *Sia $A \in \mathcal{M}_{\mathbb{K}}(n)$ allora A è invertibile se e solo se $\mathcal{L}_A$ è biettiva.*

Dimostrazione. $\mathcal{L}_A$ è biettiva $\Leftrightarrow \mathcal{L}_A$ è suriettiva $\Leftrightarrow \text{rg}(A) = n \Leftrightarrow A$ è invertibile. $\square$

Esempio. $\diamond$ Sia $f : \mathbb{K}[t]_{\leq 3} \rightarrow \mathcal{M}_{\mathbb{K}}(2, 2)$ la mappa lineare data da

$$f(p(t)) = \begin{pmatrix} p(0) & p'(0) \\ p''(0) & p'''(0) \end{pmatrix} \quad \text{i.e.} \quad f(a_0 + a_1t + a_2t^2 + a_3t^3) = \begin{pmatrix} a_0 & a_1 \\ 2a_2 & 6a_3 \end{pmatrix}$$

$a_0 + a_1t + a_2t^2 + a_3t^3 \in \ker(f) \Leftrightarrow a_0 = a_1 = 2a_2 = 6a_3 = 0$, i.e. $\ker(f) = \{0\}$, quindi f è iniettiva.

Poiché $\dim \mathbb{K}[t]_{\leq 3} = \dim \mathcal{M}_{\mathbb{K}}(2, 2) = 4$, f è anche suriettiva, quindi è un isomorfismo.

$\diamond$ Sia $g : \mathbb{K}[t]_{\leq 3} \rightarrow \mathcal{M}_{\mathbb{K}}(2, 2)$ la funzione lineare definita da $g(p(t)) = \begin{pmatrix} p(0) & p'(0) \\ -p(0) & -p''(0) \end{pmatrix}$ i.e.

$$g(a_0 + a_1t + a_2t^2 + a_3t^3) = \begin{pmatrix} a_0 & a_1 \\ -a_0 & -2a_2 \end{pmatrix}$$

Otteniamo $a_0 + a_1t + a_2t^2 + a_3t^3 \in \ker(g) \Leftrightarrow a_0 = -a_0 = a_1 = -2a_2$, i.e. $\ker(g) = \{a_3t^3 \mid a_3 \in \mathbb{K}\} = \text{Span}(t^3)$. Quindi g non è iniettiva, e quindi nemmeno suriettiva perché $\dim \mathbb{K}[t]_{\leq 3} = \dim \mathcal{M}_{\mathbb{K}}(2, 2) = 4$.

Notiamo che $\dim \ker(g) = 1$, quindi $\operatorname{rg}(g) = 4 - 1 = 3$, infatti $\operatorname{Im}(g) = \operatorname{Span}(g(1), g(t), g(t^2), g(t^3)) = \operatorname{Span}\left(\begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & -2 \end{pmatrix}\right)$.

◇ Sia $h : \mathbb{K}[t]_{\leq 2} \rightarrow \mathbb{K}^2$ la mappa lineare data da $h(p(t)) = \begin{pmatrix} p(1) \\ p(-1) \end{pmatrix}$ i.e.

$$h(a_0 + a_1 t + a_2 t^2) = \begin{pmatrix} a_0 + a_1 + a_2 \\ a_0 - a_1 + a_2 \end{pmatrix}$$

Otteniamo $a_0 + a_1 t + a_2 t^2 \in \ker(h) \Leftrightarrow a_0 + a_1 + a_2 = a_0 - a_1 + a_2 = 0_{\mathbb{K}}$, i.e. $\ker(h) = \{a_0 + a_1 t + a_2 t^2 \mid a_1 = 0_{\mathbb{K}}, a_2 = -a_0\} = \operatorname{Span}(1 - t^2)$, quindi h non è iniettiva, ma $\operatorname{rg}(h) = 3 - 1 = 2 = \dim \mathbb{K}^2$, e quindi h è suriettiva.

3.4 Come definire una mappa lineare?

Siano V, W spazi vettoriali sul campo $\mathbb{K}$, come possiamo costruire una mappa lineare $f : V \rightarrow W$? Il prossimo risultato ci dice che è sufficiente decidere il comportamento di f rispetto ad una base di V .

Teorema 3.20 (Teorema di interpolazione). *Siano V, W spazi vettoriali sul campo $\mathbb{K}$. Sia $\mathcal{B} = \{b_1, \dots, b_n\}$ una base di V e siano $w_1, \dots, w_n$ vettori di W . Allora esiste un'unica funzione lineare $f : V \rightarrow W$ tale che $f(v_i) = w_i$ per ogni $i = 1, \dots, n$.*

Quindi per dare una mappa lineare è sufficiente dare l'immagine dei vettori di una base: "Una mappa lineare è completamente determinata dal suo comportamento su una base".

Dimostrazione. Per prima cosa vediamo come si definisce f , poi dimostreremo l'unicità. Sia $v \in V$ e consideriamo le sue coordinate rispetto alla base $\mathcal{B}$: $(v)_{\mathcal{B}} = (\lambda_1, \dots, \lambda_n)^T \in \mathbb{K}^n$, cioè $v = \lambda_1 b_1 + \dots + \lambda_n b_n$. Definiamo

$$f : V \rightarrow W, \quad f(v) := \sum_{i=1}^n \lambda_i w_i$$

Chiaramente vale $f(v_i) = w_i$ per ogni $i = 1, \dots, n$ e verifichiamo che tale funzione è lineare. Siano $v, u \in V$, $(u)_{\mathcal{B}} = (\mu_1, \dots, \mu_n)^T$, $(v)_{\mathcal{B}} = (\lambda_1, \dots, \lambda_n)^T$, le loro coordinate rispetto alla base $\mathcal{B}$, e $h, k \in \mathbb{K}$. Allora $(h \cdot u + k \cdot v)_{\mathcal{B}} = (h\mu_1 + k\lambda_1, \dots, h\mu_n + k\lambda_n)^T$, quindi

$$f(h \cdot u + k \cdot v) = f\left(\sum_i (h\mu_i + k\lambda_i) b_i\right) = \sum_i (h\mu_i + k\lambda_i) w_i \stackrel{S2-S3}{=} h \sum_i (\mu_i) w_i + k \sum_i (\lambda_i) w_i = hf(u) + kf(v)$$

Verifichiamo ora che f è l'unica funzione in $\operatorname{Hom}(V, W)$ con le proprietà richieste. Supponiamo quindi che $g \in \operatorname{Hom}(V, W)$ soddisfi $g(v_i) = w_i$ per ogni $i = 1, \dots, n$. Sia $v \in V$, e siano $(v)_{\mathcal{B}} = (\lambda_1, \dots, \lambda_n)^T$, le sue coordinate rispetto alla base $\mathcal{B}$, allora

$$f(v) - g(v) = f\left(\sum_i \lambda_i v_i\right) - g\left(\sum_i \lambda_i v_i\right) = \sum_i \lambda_i w_i - \sum_i \lambda_i w_i = 0_W$$

Quindi $f = g$. □

Esempio. Sia $f : \mathbb{Q}^2 \rightarrow \mathbb{Q}[t]_{\leq 2}$ la mappa lineare definita da

$$f \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1 + t, \quad f \begin{pmatrix} 1 \\ 1 \end{pmatrix} = t^2 - t$$

allora $f(\frac{1}{3}) = f(-2(\frac{1}{0}) + 3(\frac{1}{1})) = -2(1+t) + 3(t^2-t) = -2 - 5t + 3t^2$.

Esempio. Sia $g : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ la mappa lineare definita tramite

$$g(e_1) = \begin{pmatrix} 1 \\ 6 \end{pmatrix}, \quad g(e_2) = \begin{pmatrix} -2 \\ 0 \end{pmatrix}, \quad g(e_3) = \begin{pmatrix} 3 \\ -2 \end{pmatrix}$$

allora $g \begin{pmatrix} x \\ y \\ z \end{pmatrix} = g(xe_1 + ye_2 + ze_3) = x \begin{pmatrix} 1 \\ 6 \end{pmatrix} + y \begin{pmatrix} -2 \\ 0 \end{pmatrix} + z \begin{pmatrix} 3 \\ -2 \end{pmatrix} = \begin{pmatrix} 1 & -2 & 3 \\ 6 & 0 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$.

Questa costruzione si generalizza a qualsiasi mappa lineare $\mathbb{K}^n \rightarrow \mathbb{K}^m$, in altre parole:

Lemma 3.21. *Ogni mappa lineare $f : \mathbb{K}^n \rightarrow \mathbb{K}^m$ è della forma $f = \mathcal{L}_A$ per una certa $A \in \mathcal{M}_{\mathbb{K}}(m, n)$.*

Dimostrazione. Basta prendere come colonne di A i vettori $f(e_1), \dots, f(e_n)$, dove $e_1, \dots, e_n$ è la base canonica di $\mathbb{K}^n$, infatti:

$$f((x_1, \dots, x_n)^T) = f(\sum_i x_i e_i) = \sum_i x_i f(e_i) = A \cdot (x_1, \dots, x_n)^T \quad \square$$

Vediamo ora come una mappa lineare tra spazi vettoriali arbitrari può essere codificata/descritta tramite una matrice.

Esempio. Sia $f : \mathbb{K}[t]_{\leq 3} \rightarrow \mathcal{M}_{\mathbb{K}}(2, 2)$ la mappa lineare

$$f(a_0 + a_1 t + a_2 t^2 + a_3 t^3) = \begin{pmatrix} a_0 & a_1 \\ 2a_2 & 6a_3 \end{pmatrix}$$

Qui un ruolo centrale è svolto dai coefficienti a_0, a_1, a_2, a_3 , e vorremmo sottolinearne l'importanza.

Il primo passo consiste nel collegare qualsiasi spazio vettoriale con un $\mathbb{K}^n$ adeguato.

Teorema 3.22 (Teorema di isomorfismo con $\mathbb{K}^n$). *Sia V uno spazio vettoriale $\mathbb{K}$ di dimensione $\dim V = n$. Allora esiste un isomorfismo $f : V \rightarrow \mathbb{K}^n$.*

Dimostrazione. Sia $\mathcal{B} = \{b_1, \dots, b_n\}$ una base di V e si definisca la mappa lineare $f : V \rightarrow \mathbb{K}^n$ tramite $f(b_j) = e_j$, dove $\{e_1, \dots, e_n\}$ è la base canonica di $\mathbb{K}^n$.

Otteniamo immediatamente che $\text{Im}(f) = \text{Span}(f(b_1), \dots, f(b_n)) = \text{Span}(e_1, \dots, e_n) = \mathbb{K}^n$, quindi f è suriettiva. Ma $\dim V = n = \dim \mathbb{K}^n$, quindi f è biunivoca. $\square$

Corollario 3.23. *Siano V, W spazi vettoriali sul campo $\mathbb{K}$. Essi sono isomorfi ($V \cong W$) se e solo se $\dim V = \dim W$.*

Dimostrazione. $\Rightarrow$] Dimostrato nella Proposizione 3.15.

$\Leftarrow$] Per il Teorema 3.22, esistono isomorfismi $f : V \rightarrow \mathbb{K}^n$ e $g : V \rightarrow \mathbb{K}^n$; quindi $g^{-1} \circ f : V \rightarrow W$ è un isomorfismo. $\square$

Osservazione. La dimostrazione del Teorema 3.22, ci dice anche come costruire un'isomorfismo $f : V \rightarrow \mathbb{K}^n$. Data una base $\mathcal{B}$ di V , mappiamo ogni vettore $v \in V$ sulle sue coordinate rispetto a $\mathcal{B}$: $f(v) = (v)_{\mathcal{B}}$.

Abbiamo quindi un dizionario per passare da V a $\mathbb{K}^n$ e viceversa.

Definizione 3.24. Sia V uno spazio vettoriale sul campo $\mathbb{K}$ e sia $\mathcal{B} = \{b_1, \dots, b_n\}$ una base di V , e sia $\mathcal{E} = \{e_1, \dots, e_n\}$ la base canonica di $\mathbb{K}^n$.

La mappa lineare $X_{\mathcal{B}} : V \rightarrow \mathbb{K}^n, b_i \mapsto e_i$ è chiamata *mappa delle coordinate* di $\mathcal{B}$.

La mappa lineare $P_{\mathcal{B}} : \mathbb{K}^n \rightarrow V, e_i \mapsto b_i$ è chiamata *mappa di parametrizzazione* di $\mathcal{B}$.

Sia $v = \lambda_1 b_1 + \dots + \lambda_n b_n \in V$, allora $X_{\mathcal{B}}(v) = \lambda_1 e_1 + \dots + \lambda_n e_n = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \in \mathbb{K}^n$, quindi $X_{\mathcal{B}}$ estrapola le coordinate di v rispetto alla base $\mathcal{B}$.

Si noti che $P_{\mathcal{B}} = X_{\mathcal{B}}^{-1}$, quindi $P_{\mathcal{B}} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \alpha_1 b_1 + \dots + \alpha_n b_n$.

Ora abbiamo tutti gli ingredienti per codificare una mappa lineare in una matrice.

3.4.1 Matrice rappresentativa

Siano V, W spazi vettoriali $\mathbb{K}$, $f : V \rightarrow W$ una mappa lineare., $\mathcal{B} = \{b_1, \dots, b_n\}$ una base di V e $\mathcal{C} = \{c_1, \dots, c_m\}$ una base di W .

Consideriamo la mappa $X_{\mathcal{C}} \circ f \circ P_{\mathcal{B}} : \mathbb{K}^n \rightarrow \mathbb{K}^m$. Per il Lemma 3.21, essa coincide con la mappa $\mathcal{L}_A$ per una matrice $A \in \mathcal{M}_{\mathbb{K}}(m, n)$, e vogliamo determinare questa matrice A .

$$X_{\mathcal{C}} \circ f \circ P_{\mathcal{B}} = \mathcal{L}_A : \mathbb{K}^n \rightarrow \mathbb{K}^m, \quad \begin{array}{ccc} V & \xrightarrow{f} & W \\ P_{\mathcal{B}} \uparrow & & \downarrow X_{\mathcal{C}} \\ \mathbb{K}^n & \xrightarrow{\mathcal{L}_A} & \mathbb{K}^m \end{array} \quad A = ?$$

La mappa $X_{\mathcal{C}} \circ f \circ P_{\mathcal{B}} = \mathcal{L}_A : \mathbb{K}^n \rightarrow \mathbb{K}^m$ è determinata dal suo comportamento rispetto alla base canonica $\{e_1, \dots, e_n\}$ di $\mathbb{K}^n$:

$$A \cdot e_j = \mathcal{L}_A(e_j) = X_{\mathcal{C}}(f(P_{\mathcal{B}}(e_j))) = X_{\mathcal{C}}(f(b_j)) = (f(b_j))_{\mathcal{C}} \iff A = \left((f(b_1))_{\mathcal{C}} \mid \dots \mid (f(b_n))_{\mathcal{C}} \right)$$

In parole: la colonna j -esima di A è data dalle coordinate di $f(b_j)$ rispetto alla base $\mathcal{C}$.

Definizione 3.25. Siano V, W spazi vettoriali sul campo $\mathbb{K}$, $f : V \rightarrow W$ una mappa lineare, $\mathcal{B} = \{b_1, \dots, b_n\}$ una base di V e $\mathcal{C} = \{c_1, \dots, c_m\}$ una base di W .

La *matrice rappresentativa* di f rispetto a $\mathcal{B}$ e $\mathcal{C}$ è l'unica matrice $\mathcal{M}_{\mathcal{C}}^{\mathcal{B}}(f) = A \in \mathcal{M}_{\mathbb{K}}(m, n)$ tale che

$$\mathcal{L}_A = X_{\mathcal{C}} \circ f \circ P_{\mathcal{B}},$$

cioè la cui j -esima colonna è data dalle coordinate di $f(b_j)$ rispetto alla base $\mathcal{C}$.

Notazione La matrice rappresentativa di f rispetto a $\mathcal{B}$ e $\mathcal{C}$ è anche chiamata *matrice di trasformazione* o *matrice che rappresenta f rispetto a $\mathcal{B}$ e $\mathcal{C}$* .

Esempio. $\diamond$ Per una matrice $A \in \mathcal{M}_{\mathbb{K}}(m, n)$, abbiamo $\mathcal{M}_{\mathcal{E}'}^{\mathcal{E}}(\mathcal{L}_A) = A$, dove $\mathcal{E}$ e $\mathcal{E}'$ sono le basi canoniche di $\mathbb{K}^n$ e di $\mathbb{K}^m$.

$\diamond$ Consideriamo $V = \mathbb{K}[t]_{\leq 2}$ e $W = \mathcal{M}_{\mathbb{K}}(2, 2)$, con le basi $\mathcal{B} = \{1, t, t^2\}$ e $\mathcal{C} = \left\{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right\}$ rispettivamente. Sia $f : \mathbb{K}[t]_{\leq 2} \rightarrow \mathcal{M}_{\mathbb{K}}(2, 2)$ la mappa lineare

$$f(a_0 + a_1 t + a_2 t^2) = \begin{pmatrix} a_0 & a_0 + a_1 + a_2 \\ a_0 - a_1 + a_2 & 0 \end{pmatrix}$$

Allora

$$\begin{aligned} f(1) &= \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} = 1 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + 1 \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + 1 \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + 0 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \\ f(t) &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = 0 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + 1 \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + (-1) \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + 0 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \\ f(t^2) &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = 0 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + 1 \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + 1 \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + 0 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \end{aligned} \implies \mathcal{M}_{\mathcal{C}}^{\mathcal{B}}(f) = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & -1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

$\diamond$ Consideriamo gli spazi vettoriali $\mathbb{K}[t]_{\leq 2}$ e $\mathbb{K}^2$, con basi $\mathcal{B} = \{1, 1 + t, 1 + t^2\}$ e $\mathcal{C} = \{e_1, e_1 + e_2\}$ rispettivamente. Sia $h : \mathbb{K}[t]_{\leq 2} \rightarrow \mathbb{K}^2$ la mappa lineare

$$h(a_0 + a_1 t + a_2 t^2) = \begin{pmatrix} a_0 + a_1 + a_2 \\ a_1 - 2a_2 \end{pmatrix}$$

Allora

$$\begin{aligned} h(1) &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1 \cdot e_1 + 0 \cdot (e_1 + e_2), \\ h(1 + t) &= \begin{pmatrix} 2 \\ 1 \end{pmatrix} = 1 \cdot e_1 + 1 \cdot (e_1 + e_2), \\ h(1 + t^2) &= \begin{pmatrix} 2 \\ -2 \end{pmatrix} = 4 \cdot e_1 - 2 \cdot (e_1 + e_2) \end{aligned} \implies \mathcal{M}_{\mathcal{C}}^{\mathcal{B}}(h) = \begin{pmatrix} 1 & 1 & 4 \\ 0 & 1 & -2 \end{pmatrix}$$

L'utilità principale della matrice di trasformazione è quella di fornirci un dizionario molto efficace per tradurre problemi su spazi vettoriali astratti in problemi su $\mathbb{K}^n$.

Fatto 3.26. Siano V, W spazi vettoriali sul campo $\mathbb{K}$, $\mathcal{B} = \{b_1, \dots, b_n\}$ una base di V e $\mathcal{C} = \{c_1, \dots, c_m\}$ una base di W . La funzione

$$\begin{aligned} \text{Hom}(V, W) &\longrightarrow \mathcal{M}_{\mathbb{K}}(m, n) \\ f &\longmapsto \mathcal{M}_{\mathcal{C}}^{\mathcal{B}}(f) \end{aligned}$$

è un isomorfismo di spazi vettoriali.

In particolare $\dim \text{Hom}(V, W) = mn$, e per ogni $f, g \in \text{Hom}(V, W)$, $\lambda \in \mathbb{K}$ vale

$$\mathcal{M}_{\mathcal{C}}^{\mathcal{B}}(f + g) = \mathcal{M}_{\mathcal{C}}^{\mathcal{B}}(f) + \mathcal{M}_{\mathcal{C}}^{\mathcal{B}}(g) \quad \mathcal{M}_{\mathcal{C}}^{\mathcal{B}}(\lambda f) = \lambda \mathcal{M}_{\mathcal{C}}^{\mathcal{B}}(f)$$

Come spiegato nelle proposizioni seguenti, tramite la matrice di trasformazione possiamo: i) tradurre un problema su una mappa lineare in un problema matriciale; ii) ridurre calcoli complicati ad operazioni matriciali.

Proposizione 3.27. Siano V, W $\mathbb{K}$ -spazi vettoriali, con basi rispettivamente $\mathcal{B} = \{b_1, \dots, b_n\}$ e $\mathcal{C} = \{c_1, \dots, c_m\}$. Sia $f : V \rightarrow W$ una mappa lineare e sia $A := \mathcal{M}_{\mathcal{C}}^{\mathcal{B}}(f)$ la matrice rappresentativa corrispondente. Se $v = \sum_{j=1}^n \lambda_j b_j$ e $f(v) = \sum_{i=1}^m \mu_i c_i$, allora

$$A \cdot \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_m \end{pmatrix} \quad \text{cioè} \quad A \cdot (v)_{\mathcal{B}} = (f(v))_{\mathcal{C}}$$

In particolare, $v \in \ker f \Leftrightarrow \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} \in \ker A$ e $\dim \operatorname{Im}(f) = \operatorname{rg}(f) = \operatorname{rg}(A)$.

Dimostrazione. Per definizione di A abbiamo $A \cdot X_{\mathcal{B}}(v) = X_{\mathcal{C}}(f(v)) = (f(v))_{\mathcal{C}}$, e il risultato segue dalla definizione di mappa delle coordinate. $\square$

Abbiamo già visto una ragione alla base della definizione del prodotto righe per colonne, ovvero tradurre i sistemi lineari in matrici. Vediamo ora altre motivazioni.

Proposizione 3.28. Siano V, W, U spazi vettoriali $\mathbb{K}$, con basi rispettivamente $\mathcal{B}, \mathcal{C}$ e $\mathcal{D}$.
1) Siano $f : V \rightarrow W$ e $g : W \rightarrow U$ applicazioni lineari, allora

$$\mathcal{M}_{\mathcal{D}}^{\mathcal{B}}(g \circ f) = \mathcal{M}_{\mathcal{D}}^{\mathcal{C}}(g) \cdot \mathcal{M}_{\mathcal{C}}^{\mathcal{B}}(f).$$

2) Sia $f : V \rightarrow W$ una mappa lineare invertibile, allora

$$(\mathcal{M}_{\mathcal{C}}^{\mathcal{B}}(f))^{-1} = \mathcal{M}_{\mathcal{B}}^{\mathcal{C}}(f^{-1}).$$

3) Siano $\mathcal{B}', \mathcal{C}'$ basi rispettivamente di V e W e sia $f : V \rightarrow W$ una mappa lineare, allora

$$\mathcal{M}_{\mathcal{C}'}^{\mathcal{B}'}(f) = \mathcal{M}_{\mathcal{C}'}^{\mathcal{C}}(id_W) \cdot \mathcal{M}_{\mathcal{C}}^{\mathcal{B}}(f) \cdot \mathcal{M}_{\mathcal{B}}^{\mathcal{B}'}(id_V).$$

Dimostrazione. La dimostrazione di 1) è un calcolo che utilizza le varie definizioni, simile alla dimostrazione della Proposizione 3.27.

2) e 3) seguono da 1). $\square$

3.4.2 Matrici di cambio di base

Definizione 3.29. Sia V uno spazio vettoriale $\mathbb{K}$ e siano $\mathcal{B}, \mathcal{B}'$ basi di V . La matrice

$$\mathcal{M}_{\mathcal{B}'}^{\mathcal{B}}(id_V),$$

è chiamata *matrice di cambio di base* o *matrice di transizione*.

Si osservi che $\mathcal{M}_{\mathcal{B}'}^{\mathcal{B}}(id_V) = \mathcal{M}_{\mathcal{B}}^{\mathcal{B}'}(id_V)^{-1}$.

Sottolineiamo che le colonne di $\mathcal{M}_{\mathcal{B}'}^{\mathcal{B}}(id_V)$ sono le coordinate dei vettori in $\mathcal{B}$ rispetto alla base $\mathcal{B}'$, e vale $(v)_{\mathcal{B}'} = \mathcal{M}_{\mathcal{B}'}^{\mathcal{B}}(id_V) \cdot (v)_{\mathcal{B}}$.

Esempio. Consideriamo $\mathbb{R}^2$ con basi $\mathcal{B} = \{e_1, e_2\}$ e $\mathcal{B}' = \{e_1 - 2e_2, 2e_1 - 3e_2\}$. Allora

$$e_1 = -3(e_1 - 2e_2) + 2(2e_1 - 3e_2), \quad e_2 = -2(e_1 - 2e_2) + 1(2e_1 - 3e_2),$$

Quindi

$$\mathcal{M}_{\mathcal{B}'}^{\mathcal{B}}(id_{\mathbb{R}^2}) = \begin{pmatrix} -3 & -2 \\ 2 & 1 \end{pmatrix}$$

Quindi, ad esempio, sia $v \in \mathbb{R}^2$ il vettore con coordinate $(v)_{\mathcal{B}} = (4, -7)^T$ rispetto a $\mathcal{B}$, allora v rispetto a $\mathcal{B}'$ ha coordinate

$$(v)_{\mathcal{B}'} = \mathcal{M}_{\mathcal{B}'}^{\mathcal{B}}(id_{\mathbb{R}^2}) \cdot \begin{pmatrix} 4 \\ -7 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Se nel punto 3) della Proposizione 3.28 prendiamo $V = W$ (cioè $f \in \text{Hom}(V, V) = \text{End}(V)$, $\mathcal{C} = \mathcal{B}$ e $\mathcal{C}' = \mathcal{B}'$) otteniamo la formula del *cambio di base*:

$$\begin{aligned} \mathcal{M}_{\mathcal{B}'}^{\mathcal{B}'}(f) &= \mathcal{M}_{\mathcal{B}'}^{\mathcal{B}}(id_V) \cdot \mathcal{M}_{\mathcal{B}}^{\mathcal{B}'}(f) \cdot \mathcal{M}_{\mathcal{B}}^{\mathcal{B}'}(id_V) \\ \mathcal{M}_{\mathcal{B}'}^{\mathcal{B}'}(f) &= \mathcal{M}_{\mathcal{B}}^{\mathcal{B}'}(id_V)^{-1} \cdot \mathcal{M}_{\mathcal{B}}^{\mathcal{B}'}(f) \cdot \mathcal{M}_{\mathcal{B}}^{\mathcal{B}'}(id_V) \end{aligned}$$

Definizione 3.30. Siano $A, B \in \mathcal{M}_{\mathbb{K}}(n)$ matrici quadrate. Diciamo che A e B sono *simili* (o *coniugate*), se esiste una matrice $P \in \mathcal{M}_{\mathbb{K}}(n)$ invertibile, tale che $B = P^{-1}AP$.

Osservazione. Due matrici $A, B \in \mathcal{M}_{\mathbb{K}}(n)$ simili rappresentano la stessa applicazione lineare rispetto a basi diverse (la matrice P dà il cambio di base).

In particolare, matrici simili $A, B \in \mathcal{M}_{\mathbb{K}}(n)$ hanno lo stesso rango $\text{rg}(A) = \text{rg}(B)$, lo stesso determinante $\det(B) = \det(P^{-1}AP) = \det(A)$ (per Binet) e la stessa traccia: $\text{tr}(A) = \text{tr}(B)$ (Esercizio).

Capitolo 4

Endomorfismi e diagonalizzazione

In questo capitolo ci concentreremo su una particolare famiglia di mappe lineari, gli endomorfismi: mappe lineari da uno spazio vettoriale a se stesso.

Definizione 4.1. Sia V uno spazio vettoriale su $\mathbb{K}$. Una mappa lineare $f : V \rightarrow V$ è detta *endomorfismo*. $\text{End}(V) := \text{Hom}(V, V)$ denota l'insieme degli endomorfismi di V .

Dal Fatto 3.26 segue che, se $\dim V = n$, allora $\text{End}(V)$ è uno spazio vettoriale di dimensione n^2 ; e che fissata una base $\mathcal{B}$ di V , si ha un isomorfismo

$$\begin{aligned}\text{End}(V) &\longrightarrow \mathcal{M}_{\mathbb{K}}(n) \\ f &\longmapsto \mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(f)\end{aligned}$$

Il nostro obiettivo è quello di trovare una base $\mathcal{B}$ di V , per la quale $\mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(f)$ si presenti in una forma particolarmente maneggevole e il significato geometrico di f sia chiaro.

Esempio. Consideriamo lo spazio vettoriale $\mathbb{R}^2$ e la mappa lineare $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ data dalla riflessione rispetto a una retta $l \subset \mathbb{R}^2$ di inclinazione θ che passa per $0_{\mathbb{R}^2}$.

Consideriamo la base canonica $\mathcal{E} = \{e_1, e_2\}$ di $\mathbb{R}^2$ e scriviamo la matrice rappresentativa $\mathcal{M}_{\mathcal{E}}^{\mathcal{E}}(f)$. Considerazioni geometriche e trigonometriche danno:

$$f(e_1) = \begin{pmatrix} \cos(2\theta) \\ \sin(2\theta) \end{pmatrix}, \quad f(e_2) = \begin{pmatrix} \sin(2\theta) \\ -\cos(2\theta) \end{pmatrix} \implies \mathcal{M}_{\mathcal{E}}^{\mathcal{E}}(f) = \begin{pmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{pmatrix} = A.$$

Consideriamo ora una base $\mathcal{B} = \{b_1, b_2\}$ di $\mathbb{R}^2$, che colga la geometria del problema: b_1 giace sulla retta l e b_2 è ortogonale a l , ad esempio $b_1 = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}$ e $b_2 = \begin{pmatrix} -\sin(\theta) \\ \cos(\theta) \end{pmatrix}$. La matrice rappresentativa $\mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(f)$ è:

$$f(b_1) = b_1, f(b_2) = -b_2 \implies \mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(f) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = D.$$

Sia P la matrice di cambio di base

$$P := \mathcal{M}_{\mathcal{E}}^{\mathcal{B}}(id_{\mathbb{R}^2}) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

Secondo la Proposizione 3.28, il cambiamento di base influisce sulla matrice rappresentativa, come segue:

$$\mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(f) = \mathcal{M}_{\mathcal{B}}^{\mathcal{E}}(id_V) \mathcal{M}_{\mathcal{E}}^{\mathcal{E}}(f) \mathcal{M}_{\mathcal{E}}^{\mathcal{B}}(id_V) \quad \text{ovvero } D = P^{-1}AP.$$

Domanda. Dato un endomorfismo $f : V \rightarrow V$, è possibile trovare una base $\mathcal{B}$ di V tale che la matrice $\mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(f)$ rappresentativa di f rispetto a tale base sia diagonale?

In altre parole, data una matrice $A \in \mathcal{M}_{\mathbb{K}}(n)$, esistono una matrice diagonale $D \in \mathcal{M}_{\mathbb{K}}(n)$ e una matrice invertibile $P \in \mathcal{M}_{\mathbb{K}}(n)$ tali che $D = P^{-1}AP$?

Notazione. Quando dovremo scrivere una matrice contenente molti zeri, useremo la convenzione di omettere le entrate nulle.

Indichiamo una matrice diagonale $\begin{pmatrix} \delta_1 & & \\ & \delta_2 & \\ & & \ddots \\ & & & \delta_n \end{pmatrix}$ con la notazione abbreviata $\text{diag}(\delta_1, \dots, \delta_n)$.

4.1 Autovettori ed autovalori

Definizione 4.2. Sia V uno spazio vettoriale $\mathbb{K}$ e sia $f : V \rightarrow V$ un endomorfismo.

Uno scalare $\lambda \in \mathbb{K}$ è un *autovalore* di f , se esiste un vettore non nullo $v \in V$, $v \neq 0_V$ tale che $f(v) = \lambda \cdot v$.

Un vettore non nullo $v \in V$, $v \neq 0_V$ è un *autovettore* di f se esiste uno scalare $\lambda \in \mathbb{K}$ tale che $f(v) = \lambda \cdot v$; in questo caso, λ è detto *autovalore* di v .

Lo spettro di f è l'insieme di tutti gli autovalori di f : $S(f) = \{\lambda \in \mathbb{K} \mid \lambda \text{ è un autovalore di } f\}$.

Esempio. 1) Nell'esempio della pagina precedente, i vettori b_1 e b_2 sono autovettori con autovalori rispettivamente pari a 1 e -1.

2) Sia $A = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 2 & 3 & 1 \end{pmatrix} \in \mathcal{M}_{\mathbb{Q}}(3, 3)$. L'endomorfismo $\mathcal{L}_A : \mathbb{Q}^3 \rightarrow \mathbb{Q}^3$ ha autovettore

$u = (1, 1, 1)^T$ di autovalore 6, infatti $A \cdot u = (6, 6, 6) = 6 \cdot u$.

3) Sia $f : V \rightarrow V$ un endomorfismo e sia $v \in \ker(f)$, $v \neq 0$, allora $f(v) = 0_V = 0_{\mathbb{K}} \cdot v$, quindi $0_{\mathbb{K}}$ è un autovalore di f .

Definizione 4.3. Sia V uno spazio vettoriale $\mathbb{K}$ e sia $f \in \text{End}(V)$. Diciamo che f è *diagonalizzabile* se esiste una base di V costituita da autovettori di f .

Osservazione. Il significato della Definizione 4.3 è che un endomorfismo è *rappresentabile in forma diagonale* precisamente quando è diagonalizzabile, infatti se $\mathcal{B} = \{b_1, \dots, b_n\}$ è una base di V costituita da autovettori di $f \in \text{End}(V)$: $f(b_i) = \lambda_i b_i$, allora

$$\mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(f) = \begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots \\ & & & \lambda_n \end{pmatrix} = \text{diag}(\lambda_1, \dots, \lambda_n)$$

Viceversa, se $\mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(f)$ è diagonale, allora tutti gli elementi di $\mathcal{B}$ sono autovettori.

Alla luce di questa osservazione e delle formule del cambio di base (si veda la Sezione 3.4.2), otteniamo che un endomorfismo $f \in \text{End}(V)$ è diagonalizzabile se e solo se la matrice $A := \mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(f)$ rappresentativa di f rispetto ad una base $\mathcal{B}$ di V è simile ad una matrice diagonale.

Infatti, se $\mathcal{C} = \{c_1, \dots, c_n\}$ è una base di V costituita da autovettori, allora basta prendere $D := \mathcal{M}_{\mathcal{C}}^{\mathcal{C}}(f)$ e $P := \mathcal{M}_{\mathcal{B}}^{\mathcal{C}}(id_V)$, e per la Proposizione 3.28 vale $P^{-1}AP = D$.

Viceversa, se esistono una matrice diagonale $D = \text{diag}(\delta_1, \dots, \delta_n) \in \mathcal{M}_{\mathbb{K}}(n)$ e una matrice invertibile $P \in \mathcal{M}_{\mathbb{K}}(n)$ tali che $D = P^{-1}AP$, allora D rappresenta f rispetto alla base $\mathcal{C} = \{c_1, \dots, c_n\}$, determinata da $P = \mathcal{M}_{\mathcal{B}}^{\mathcal{C}}(id)$. Per definizione $f(c_i) = \delta_i c_i$: c_i è un autovettore.

Abbiamo quindi il seguente criterio.

Teorema 4.4 (1° criterio di diagonalizzazione). *Sia V un $\mathbb{K}$ -spazio vettoriale di dimensione $\dim V = n$, $f \in \text{End}(V)$.*

Allora, l'endomorfismo f è diagonalizzabile se e solo se la matrice $A := \mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(f)$, rappresentativa di f rispetto a una base $\mathcal{B}$ di V , è simile ad una matrice diagonale: esistono una matrice diagonale $D \in \mathcal{M}_{\mathbb{K}}(n)$ e una matrice invertibile $P \in \mathcal{M}_{\mathbb{K}}(n)$ tali che $D = P^{-1}AP$.

Osservazione. a) Da 3) dell'esempio precedente, si ottiene che un endomorfismo $f : V \rightarrow V$ non è iniettivo se e solo se $\ker(f) \neq \{0_V\}$ se e solo se $0_{\mathbb{K}}$ è un autovalore di f .

b) In generale, uno scalare $\lambda \in \mathbb{K}$ è un autovalore di f se e solo se $\ker(\lambda \cdot id_V - f) \neq \{0_V\}$, infatti vogliamo che esista un $v \in V$ non nullo tale che

$$f(v) = \lambda v \Leftrightarrow \lambda v - f(v) = 0_V \Leftrightarrow (\lambda \cdot id_V - f)v = 0_V.$$

c) Benché il concetto di diagonalizzabilità sia definito relativamente ad un endomorfismo, è consuetudine parlare di diagonalizzazione, autovettori ed autovalori della matrice $A \in \mathcal{M}_{\mathbb{K}}(n)$ riferendosi alla diagonalizzazione, agli autovalori e agli autovettori dell'endomorfismo associato $\mathcal{L}_A : \mathbb{K}^n \rightarrow \mathbb{K}^n$ (si ricorda che $A = \mathcal{M}_{\mathcal{E}}^{\mathcal{E}}(\mathcal{L}_A)$).

4.2 Polinomio caratteristico

Come si trovano gli autovalori e gli autovettori di un endomorfismo?

Siano V un $\mathbb{K}$ -spazio vettoriale di dimensione $\dim V = n$, $f : V \rightarrow V$ un endomorfismo e $A := \mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(f)$ la sua matrice rappresentativa rispetto alla base $\mathcal{B}$ di V .

Abbiamo la seguente catena di equivalenze:

$$\begin{aligned} \lambda \in \mathbb{K} \text{ è un } \underline{\text{autovalore}} \text{ di } f &\iff \ker(\lambda \cdot id_V - f) \neq \{0_V\} \\ &\iff \ker(\lambda \cdot I_n - A) \neq \{0_V\} \\ &\iff \text{rg}(\lambda \cdot I_n - A) < n \\ &\iff \det(\lambda \cdot I_n - A) = 0 \end{aligned}$$

Quindi, per trovare gli autovalori, cerchiamo gli scalari che danno un sistema lineare omogeneo con più di una soluzione.

Gli autovettori relativi a un autovalore sono le soluzioni non banali di questo sistema lineare omogeneo.

Esempio. Consideriamo la mappa lineare $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ data dalla riflessione rispetto alla retta $l \subset \mathbb{R}^2$ di inclinazione $\theta = \frac{\pi}{6}$ che passa per $0_{\mathbb{R}^2}$. La sua matrice rappresentativa rispetto alla base canonica di $\mathbb{R}^2$ è:

$$A = \mathcal{M}_{\mathcal{E}}^{\mathcal{E}}(f) = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}$$

Gli autovalori di f soddisfano

$$0 = \det(\lambda I_n - A) = \det \begin{pmatrix} \lambda - \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \lambda + \frac{1}{2} \end{pmatrix} = (\lambda - \frac{1}{2})(\lambda + \frac{1}{2}) - \frac{\sqrt{3}}{2} \cdot \frac{\sqrt{3}}{2} = \lambda^2 - 1 = (\lambda - 1)(\lambda + 1)$$

cioè $S(f) = \{-1, 1\}$.

Ora, per determinare gli autovettori relativi all'autovalore $\lambda = 1$, dobbiamo calcolare $\ker(1 \cdot I_n - A)$:

$$\ker \begin{pmatrix} 1 - \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & 1 + \frac{1}{2} \end{pmatrix} = \ker \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{3}{2} \end{pmatrix} = \ker \begin{pmatrix} 1 & -\sqrt{3} \\ 0 & 0 \end{pmatrix} = \text{Span} \left(\begin{pmatrix} \sqrt{3} \\ 1 \end{pmatrix} \right)$$

Similmente, gli autovettori relativi all'autovalore $\lambda = -1$ si ottengono calcolando $\ker(-1 \cdot I_n - A)$:

$$\ker \begin{pmatrix} -1 - \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -1 + \frac{1}{2} \end{pmatrix} = \ker \begin{pmatrix} -\frac{3}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} = \ker \begin{pmatrix} \sqrt{3} & 1 \\ 0 & 0 \end{pmatrix} = \text{Span} \left(\begin{pmatrix} 1 \\ -\sqrt{3} \end{pmatrix} \right)$$

Abbiamo ottenuto 2 autovettori $\left(\begin{pmatrix} \sqrt{3} \\ 1 \end{pmatrix} \right)$ e $\left(\begin{pmatrix} 1 \\ -\sqrt{3} \end{pmatrix} \right)$ che sono linearmente indipendenti, quindi abbiamo una base di $\mathbb{R}^2$: f è diagonalizzabile (come già sapevamo).

Il nostro obiettivo ora è quello di replicare questa procedura in generale e trovare una strategia per determinare se un endomorfismo è diagonalizzabile e, in tal caso, come trovare una base di autovettori.

Definizione 4.5. Sia $A \in \mathcal{M}_{\mathbb{K}}(n)$ una matrice quadrata. Il *polinomio caratteristico* di A è il polinomio $p_A(t) = \det(tI_n - A) \in \mathbb{K}[t]$.

Il polinomio caratteristico di un endomorfismo $f : V \rightarrow V$ è il polinomio $p_f(t) = \det(tI_n - A) \in \mathbb{K}[t]$, per qualsiasi matrice rappresentativa $A := \mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(f)$ di f .

Si noti che effettivamente $p_f(t)$ non dipende dalla scelta della base $\mathcal{B}$ di V .

Lemma 4.6. Siano $A, B \in \mathcal{M}_{\mathbb{K}}(n)$ matrici simili, allora $p_A(t) = p_B(t)$.

Dimostrazione. Sia $P \in \mathcal{M}_{\mathbb{K}}(n)$ una matrice invertibile tale che $B = P^{-1}AP$. Allora, per il teorema di Binet e le proprietà del prodotto righe per colonne, abbiamo

$$\begin{aligned} p_B(t) &= \det(tI_n - B) = \det(tP^{-1}I_nP - P^{-1}AP) = \det(P^{-1}(tI_n - A)P) \\ &= \det(P^{-1}) \det(tI_n - A) \det P = \det(tI_n - A) = p_A(t) \quad \square \end{aligned}$$

Il polinomio caratteristico gode delle seguenti proprietà.

Proprietà. Sia V uno spazio vettoriale $\mathbb{K}$ e sia $f \in \text{End}(V)$.

- Gli autovalori di f sono le radici in $\mathbb{K}$ del suo polinomio caratteristico $p_f(t) \in \mathbb{K}[t]$.
- Il polinomio caratteristico $p_f(t)$ ha grado n .
- In particolare, lo spettro $S(f)$ ha cardinalità al più n .

Osservazione. L'ultima proprietà segue dal fatto che un polinomio $p(t) \in \mathbb{K}[t]$ di grado n , ha al più n radici in $\mathbb{K}$ (contate con molteplicità, si veda la prossima sezione).

Esempio. a) Il polinomio caratteristico di una matrice $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_{\mathbb{K}}(2)$ è

$$\det \begin{pmatrix} t-a & -b \\ -c & t-d \end{pmatrix} = t^2 - (a+d)t + (ad-bc)$$

b) In generale, il polinomio caratteristico di una matrice $A \in \mathcal{M}_{\mathbb{K}}(n)$ è della forma

$$p_A(t) = t^n - \operatorname{tr}(A)t^{n-1} + a_{n-2}t^{n-2} + \cdots + (-1)^n \det A$$

dove $\operatorname{tr}(A) = a_{1,1} + a_{2,2} + \cdots + a_{n,n}$ è la *traccia* di A .

Osservazione. 1) Alcuni autori definiscono il polinomio caratteristico come $\det(A - tI_n)$. Questa definizione differisce dalla nostra definizione $p_A(t) = \det(tI_n - A)$ per un segno $(-1)^n$. Mantiene la proprietà di avere come radici gli autovalori di A ; tuttavia la nostra definizione fornisce un polinomio *monico**, mentre con la definizione alternativa è monico solo quando n è pari.

2) Nel problema della diagonalizzazione, il campo $\mathbb{K}$ gioca un ruolo importante!

La catena naturale di inclusioni $\mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$ produce le inclusioni

$$\mathcal{M}_{\mathbb{Q}}(n, n) \subset \mathcal{M}_{\mathbb{R}}(n, n) \subset \mathcal{M}_{\mathbb{C}}(n, n)$$

È possibile che una matrice sia diagonalizzabile su $\mathbb{C}$ ma non su $\mathbb{R}$, e che una matrice sia diagonalizzabile su $\mathbb{R}$ ma non su $\mathbb{Q}$.

Esempio. c) Il polinomio caratteristico di $M = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ è $t^2 + 1$.

Su $\mathbb{R}$ il polinomio $t^2 + 1$ non ha radici, quindi M non ha autovalori e non può essere diagonalizzabile.

D'altra parte, se $\mathbb{K} = \mathbb{C}$, allora $t^2 + 1$ ha 2 radici: $\pm i$ (i è l'unità immaginaria: $i^2 = -1$) e vedremo che M è diagonalizzabile su $\mathbb{C}$.

d) Il polinomio caratteristico di $N = \begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix}$ è $t^2 - 2$, che non ha radici in $\mathbb{Q}$, quindi N non può essere diagonalizzabile su $\mathbb{Q}$. Ma $t^2 - 2$ ha 2 radici reali: $\pm\sqrt{2}$ e vedremo che N è diagonalizzabile su $\mathbb{R}$.

4.3 Molteplicità

Sia $\lambda \in \mathbb{K}$ un autovalore dell'endomorfismo $f : V \rightarrow V$. Gli autovettori relativi all'autovalore λ (insieme a 0_V) sono le soluzioni di un certo sistema lineare omogeneo, quindi formano un sottospazio vettoriale di V .

Definizione 4.7. Sia V uno spazio vettoriale sul campo $\mathbb{K}$ e sia $\lambda \in \mathbb{K}$ un autovalore dell'endomorfismo $f : V \rightarrow V$.

- L'*autospatio* associato a λ è il sottospazio vettoriale $V_\lambda := \ker(\lambda \cdot id_V - f)$.
- La *molteplicità geometrica* di λ è $\operatorname{mg}(\lambda) := \dim V_\lambda = \dim V - \operatorname{rg}(\lambda \cdot id_V - f)$.
- La *molteplicità algebrica* di λ è $\operatorname{ma}(\lambda) := \max\{k \in \mathbb{N} \mid (t - \lambda)^k \text{ divide } p_f(t)\}$.

La molteplicità algebrica di un autovalore λ conta quante volte λ è radice di $p_f(t)$, ovvero "a quale potenza $(t - \lambda)$ appare in $p_f(t)$ ".

*monico significa che il coefficiente del termine di grado più alto è 1.

Esempio. Sia $p_1(x) = (x-1)^3(x+2)$, allora f ha 2 autovalori reali: 1 e -2 , con molteplicità algebriche $\text{ma}(1) = 3$, $\text{ma}(-2) = 1$.

Sia $p_2(x) = (x+1)^4(x^2+1)^2$. Come polinomio su $\mathbb{Q}$ o $\mathbb{R}$, p_2 ha solo un'unica radice: -1 con $\text{ma}(-1) = 4$.

Su $\mathbb{C}$ invece, p_2 si scompone completamente $p_2(t) = (x+1)^4(x-i)^2(x+i)^2$, quindi ha 3 radici complesse: -1 , i , $-i$ con $\text{ma}(-1) = 4$, $\text{ma}(i) = \text{ma}(-i) = 2$.

Ogni polinomio reale $p(t) \in \mathbb{R}[t]$ di grado 2 può avere 0 radici (e.g. $t^2 + 3$), 1 radice (e.g. $t^2 + 2t + 1$), o 2 radici (e.g. $t^2 - 3t + 2$), ma mai più di 2 radici, infatti vale il seguente risultato, che il lettore conosce dalle scuole superiori e si dimostra usando la divisione tra polinomi.

Teorema 4.8. Sia $p(t) \in \mathbb{K}[t]$ un polinomio di grado n , e siano $x_1, \dots, x_k \in \mathbb{K}$ tutte le sue radici in $\mathbb{K}$. Allora $\sum_i \text{ma}(x_i) \leq n$.

Su $\mathbb{C}$, vale un risultato più forte (che non dimostriamo)

Teorema 4.9 (Teorema fondamentale dell'algebra). Ogni polinomio complesso non nullo $p(t) \in \mathbb{C}[t]$ di grado $n \in \mathbb{N}$ ha esattamente n radici, contate con molteplicità.

In altre parole, se $\{z_1, \dots, z_r\} \subset \mathbb{C}$ sono le radici complesse di $p(t)$, allora $\sum_{i=1}^r \text{ma}(z_i) = n$.

Proposizione 4.10. Sia V un $\mathbb{K}$ -spazio vettoriale e sia $\lambda \in \mathbb{K}$ un autovalore dell'endomorfismo $f : V \rightarrow V$. Allora

$$1 \leq \text{mg}(\lambda) \leq \text{ma}(\lambda) \leq \dim V$$

Dimostrazione. La prima e l'ultima disuguaglianza seguono dalle rispettive definizioni.

Per quella centrale, sia $\{b_1, \dots, b_p\}$ una base di $V_\lambda \triangleleft V$ e completiamola fino a ottenere una base $\mathcal{B} = \{b_1, \dots, b_p, b_{p+1}, \dots, b_n\}$ di V . La matrice rappresentativa $A = \mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(f)$ è

$$A = \begin{pmatrix} \lambda & 0 & \dots & 0 & a_{1,p+1} & \dots & a_{1,n} \\ 0 & \lambda & \dots & 0 & a_{2,p+1} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & 0 & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda & a_{p,p+1} & \dots & a_{p,n} \\ 0 & 0 & \dots & 0 & a_{p+1,p+1} & \dots & a_{p+1,n} \\ \vdots & \vdots & \ddots & 0 & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & a_{n,p+1} & \dots & a_{n,n} \end{pmatrix}$$

Quindi il polinomio caratteristico è $p_f(t) = \det(tI_n - A) =$

$$\det \begin{pmatrix} t - \lambda & 0 & \dots & 0 & -a_{1,p+1} & \dots & -a_{1,n} \\ 0 & t - \lambda & \dots & 0 & -a_{2,p+1} & \dots & -a_{2,n} \\ \vdots & \vdots & \ddots & 0 & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & t - \lambda & a_{p,p+1} & \dots & -a_{p,n} \\ 0 & 0 & \dots & 0 & t - a_{p+1,p+1} & \dots & -a_{p+1,n} \\ \vdots & \vdots & \ddots & 0 & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & -a_{n,p+1} & \dots & t - a_{n,n} \end{pmatrix} = (t - \lambda)^p \begin{pmatrix} t - a_{p+1,p+1} & \dots & -a_{p+1,n} \\ \vdots & \ddots & \vdots \\ -a_{n,p+1} & \dots & t - a_{n,n} \end{pmatrix}$$

da cui segue $\text{mg}(\lambda) = p \leq \text{ma}(\lambda)$. □

Definizione 4.11. Sia V uno spazio vettoriale su $\mathbb{K}$, $f \in \text{End}(V)$ e sia $\lambda \in \mathbb{K}$ un autovalore di f .

L'autovalore λ è detto *regolare* se $\text{ma}(\lambda) = \text{mg}(\lambda)$.

L'autovalore λ è detto *semplice* se $\text{ma}(\lambda) = 1$.

Si noti che se $\text{ma}(\lambda) = 1$, allora $\text{mg}(\lambda) = 1$; quindi un autovalore semplice è automaticamente regolare.

Proposizione 4.12. Sia V uno spazio vettoriale su $\mathbb{K}$ e siano $\lambda_1, \lambda_2 \in \mathbb{K}, \lambda_1 \neq \lambda_2$ autovalori distinti dell'endomorfismo $f : V \rightarrow V$. Allora $V_{\lambda_1} \cap V_{\lambda_2} = \{0_V\}$.

Dimostrazione. Sia $v \in V_{\lambda_1} \cap V_{\lambda_2}$. Da $v \in V_{\lambda_1}$, otteniamo $f(v) = \lambda_1 v$; e da $v \in V_{\lambda_2}$, otteniamo $f(v) = \lambda_2 v$. Pertanto, $\lambda_1 v = f(v) = \lambda_2 v$, quindi $(\lambda_1 - \lambda_2)v = 0_V$. Poiché $\lambda_1 - \lambda_2 \neq 0_{\mathbb{K}}$, dalla legge di annullamento del prodotto otteniamo $v = 0_V$. $\square$

In particolare, due vettori propri con autovalori distinti sono linearmente indipendenti. Più in generale, “gli autospazi contengono informazioni indipendenti!”.

Proposizione 4.13. Siano V un $\mathbb{K}$ -spazio vettoriale, $f \in \text{End}(V)$ e $\lambda_1, \dots, \lambda_k \in \mathbb{K}$ autovalori distinti ($\lambda_i \neq \lambda_j$ per $i \neq j$) di f . Per $i = 1, \dots, k$, sia v_i un autovettore in V_{λ_i} . Allora $\{v_1, \dots, v_k\}$ è un insieme di vettori linearmente indipendenti.

Dimostrazione. Per induzione su k .

Il caso base $k = 2$ è risolto nella Proposizione 4.12.

Supponiamo ora che $\{v_1, \dots, v_{k-1}\}$ sia un insieme di vettori linearmente indipendenti e dimostriamo che $v_1, \dots, v_k$ sono linearmente indipendenti. Sia $0_V = \mu_1 v_1 + \dots + \mu_{k-1} v_{k-1} + \mu_k v_k$ (*). Applicando f a (*) otteniamo

$$\begin{aligned} 0_V &= f(0_V) = f(\mu_1 v_1 + \dots + \mu_{k-1} v_{k-1} + \mu_k v_k) = \mu_1 f(v_1) + \dots + \mu_{k-1} f(v_{k-1}) + \mu_k f(v_k) \\ &= \mu_1 \lambda_1 v_1 + \dots + \mu_{k-1} \lambda_{k-1} v_{k-1} + \mu_k \lambda_k v_k \end{aligned}$$

e moltiplicando (*) per λ_k otteniamo

$$0_V = \mu_1 \lambda_k v_1 + \dots + \mu_{k-1} \lambda_k v_{k-1} + \mu_k \lambda_k v_k$$

Sottraendo queste ultime 2 combinazioni lineari otteniamo

$$0_V = 0_V - 0_V = \mu_1 (\lambda_1 - \lambda_k) v_1 + \dots + \mu_{k-1} (\lambda_{k-1} - \lambda_k) v_{k-1}$$

Poiché $\{v_1, \dots, v_{k-1}\}$ è un insieme di vettori linearmente indipendenti, otteniamo $\mu_1 (\lambda_1 - \lambda_k) = \dots = \mu_{k-1} (\lambda_{k-1} - \lambda_k) = 0_{\mathbb{K}}$; ma i λ_i sono a due a due diversi, quindi $\mu_1 = \dots = \mu_{k-1} = 0_{\mathbb{K}}$ e ci rimane $\mu_k v_k = 0_V$ da cui concludiamo che $\mu_k = 0_{\mathbb{K}}$. $\square$

Diagonalizzare un endomorfismo $f : V \rightarrow V$ significa che possiamo trovare una base di V costituita da autovettori di f . Ma gli autovettori si dividono in autospazi differenti, quindi per diagonalizzare f dobbiamo trovare una base $\mathcal{B}_{\lambda_i}$ per ogni autospazio V_{λ_i} e avere $\dim V = \sum_{i=1}^r \text{mg}(\lambda_i)$. Infatti,

$$\mathcal{B} = \mathcal{B}_{\lambda_1} \cup \dots \cup \mathcal{B}_{\lambda_r}$$

è un insieme di vettori linearmente indipendenti (segue dalla Proposizione 4.13), e per generare V , abbiamo bisogno che $\mathcal{B}$ abbia $n = \dim V$ elementi.

Si noti che

$$\begin{array}{ccccccc} \text{ma}(\lambda_1) & + & \text{ma}(\lambda_2) & + & \cdots & + & \text{ma}(\lambda_r) & \leq \text{grad } p_f = n = \dim V \\ \vee & & \vee & & & & \vee & \\ \text{mg}(\lambda_1) & + & \text{mg}(\lambda_2) & + & \cdots & + & \text{mg}(\lambda_r) \end{array}$$

quindi, affinché $\text{mg}(\lambda_1) + \text{mg}(\lambda_2) + \cdots + \text{mg}(\lambda_r) = \dim V$, è necessario che tutti i $\leq$ siano uguaglianze.

Riassumendo, abbiamo dedotto il seguente importante criterio.

Teorema 4.14 (2° criterio di diagonalizzazione). *Sia V uno spazio vettoriale di dimensione $\dim V = n$ sul campo $\mathbb{K}$, e sia $f : V \rightarrow V$ un endomorfismo. Sia $S(f) = \{\lambda_1, \dots, \lambda_r\}$ l'insieme degli autovalori di f . Allora le seguenti affermazioni sono equivalenti:*

1. f è diagonalizzabile.
2. $\sum_{i=1}^r \text{ma}(\lambda_i) = n$ e per ogni $\text{ma}(\lambda_i) = \text{mg}(\lambda_i)$ per ogni $i = 1, \dots, k$.
3. $\sum_{i=1}^r \text{mg}(\lambda_i) = n$.

Corollario 4.15. *Sia V uno spazio vettoriale sul campo $\mathbb{K}$ di dimensione $\dim V = n$, e sia $f : V \rightarrow V$ un endomorfismo. Se p_f ha tutte le sue radici in $\mathbb{K}$ e ogni autovalore è semplice, allora f è diagonalizzabile.*

Dimostrazione. p_f ha tutte le sue radici in $\mathbb{K}$ significa che $\sum_{\lambda \in S(f)} \text{ma}(\lambda) = n$.

Ogni autovalore è semplice significa che $1 \leq \text{mg}(\lambda) \leq \text{ma}(\lambda) = 1$, quindi $\text{ma}(\lambda) = \text{mg}(\lambda)$ per ogni $\lambda \in S(f)$. $\square$

Esempio. a) Il polinomio caratteristico di $M = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ è $t^2 + 1$, quindi M non è diagonalizzabile su $\mathbb{R}$ (come già visto).

Come matrice su $\mathbb{C}$, M ha due autovalori: $\pm i$, entrambi semplici, quindi M è diagonalizzabile su $\mathbb{C}$.

b) Analogamente, il polinomio caratteristico di $N = \begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix}$ è $t^2 - 2$, quindi N non è diagonalizzabile su $\mathbb{Q}$, ma è diagonalizzabile su $\mathbb{R}$.

Esempio. Il polinomio caratteristico di $B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \mathcal{M}_{\mathbb{R}}(2, 2)$ è $(t - 1)^2$, quindi ha un unico autovalore 1 con molteplicità algebrica $\text{ma}(1) = 2$.

Calcoliamo ora l'autospazio corrispondente:

$$V_1 = \ker(I_2 - B) = \ker \begin{pmatrix} 1-1 & -1 \\ 0 & 1-1 \end{pmatrix} = \ker \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} = \text{Span}\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) \Rightarrow g_1 = 1 < 2$$

Quindi, B non è diagonalizzabile (né su $\mathbb{R}$ né su $\mathbb{C}$).

4.4 Sintesi del processo di diagonalizzazione

Input Un endomorfismo $f : V \rightarrow V$, dove $\dim V = n$.

Passo 0 Scegliere una base $\mathcal{B}$ di V e scrivere $A = \mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(f) \in \mathcal{M}_{\mathbb{K}}(n)$.

Passo 1 Determinare il polinomio caratteristico $p_A(t) = \det(tI_n - A)$ di A .

Passo 2 Scomporre $p_A(t)$ su $\mathbb{K}$, ovvero determinare gli autovalori $\lambda_1, \dots, \lambda_r$ di A e le loro molteplicità algebriche $\text{ma}(\lambda_1), \dots, \text{ma}(\lambda_r)$

Controllo 1 Se $\sum_{i=1}^r a_{\lambda_i} < n$, allora **Output**: f non è diagonalizzabile.

Passo 3 Per ogni $\lambda_i \in S(f)$ trovare una base $\mathcal{C}_{\lambda_i}$ dell'autospazio V_{λ_i} e determinarne la molteplicità geometrica $\text{mg}(\lambda_i)$.

Controllo 2 Se $\text{mg}(\lambda_i) < \text{ma}(\lambda_i)$ per almeno un $i \in \{1, \dots, r\}$, allora **Output**: f non è diagonalizzabile.

Output f è diagonalizzabile: $\mathcal{C} = \mathcal{C}_{\lambda_1} \cup \dots \cup \mathcal{C}_{\lambda_r}$ è una base di V costituita dai vettori propri di f , e inoltre:

$$\begin{array}{ccccccc} \mathcal{M}_{\mathcal{C}}^{\mathcal{C}}(f) & = & \mathcal{M}_{\mathcal{B}}^{\mathcal{C}}(id_V) & \cdot & \mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(f) & \cdot & \mathcal{M}_{\mathcal{C}}^{\mathcal{B}}(id_V) \\ \parallel & & \parallel & & \parallel & & \parallel \\ D & = & P^{-1} & \cdot & A & \cdot & P \end{array}$$

dove $D = \text{diag}(\lambda_1, \dots, \lambda_r)$ e le colonne di P sono i vettori di $\mathcal{C}_{\lambda_1} \cup \dots \cup \mathcal{C}_{\lambda_r}$ (rispettando l'ordine!).

Esempio. • La matrice $A = \begin{pmatrix} 3 & -2 & -1 \\ 0 & 0 & 1 \\ 2 & -2 & -1 \end{pmatrix}$ ha polinomio caratteristico

$$\begin{aligned} p_A(t) &= \det \begin{pmatrix} t-3 & 2 & 1 \\ 0 & t & -1 \\ -2 & 2 & t+1 \end{pmatrix} = t \det \begin{pmatrix} t-3 & 1 \\ -2 & t+1 \end{pmatrix} + \det \begin{pmatrix} t-3 & 2 \\ -2 & 2 \end{pmatrix} \\ &= t(t^2 - 2t - 1) + 2t - 2 = t^3 - 2t^2 + t - 2 = (t^2 + 1)(t - 2) \end{aligned}$$

A non è diagonalizzabile su $\mathbb{R}$, poiché 2 è l'unica radice reale di $p_A(t)$ e $\text{ma}(2) = 1$.

D'altra parte, su $\mathbb{C}$ la matrice A ha 3 radici distinte, ciascuna con molteplicità algebrica uno, quindi su $\mathbb{C}$ la matrice A è diagonalizzabile (Corollario 4.15).

• La matrice $A = \begin{pmatrix} 3 & -1 & -2 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix}$ ha polinomio caratteristico

$$\begin{aligned} p_A(t) &= \det \begin{pmatrix} t-3 & 1 & 2 \\ -1 & t & 1 \\ 0 & -1 & t-1 \end{pmatrix} = (t-3) \det \begin{pmatrix} t & 1 \\ -1 & t-1 \end{pmatrix} + \det \begin{pmatrix} 1 & 2 \\ -1 & t-1 \end{pmatrix} \\ &= (t-3)(t^2 - t + 1) + (t+1) = t^3 - 4t^2 + 5t - 2 = (t-1)^2(t-2) \end{aligned}$$

Su $\mathbb{R}$ (e anche su $\mathbb{C}$) la matrice A ha 2 radici: 1, 2 con $\text{ma}(1) = 2$ e $\text{ma}(2) = 1$: $1+2 = 3$ quindi il controllo 1 è superato.

Determiniamo ora le molteplicità geometriche dei 2 autovalori: poiché 2 è un autovalore semplice, sappiamo che $\text{mg}(2) = 1$, quindi procediamo calcolando $\text{mg}(1) = \dim V_1$:

$$V_1 = \ker(1I_3 - A) = \ker \begin{pmatrix} -2 & 1 & 2 \\ -1 & 1 & 1 \\ 0 & -1 & 0 \end{pmatrix} = \ker \begin{pmatrix} 1 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & -2 & 0 \end{pmatrix} = \ker \begin{pmatrix} 1 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \text{Span} \left(\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right)$$

quindi $\text{mg}(1) = 1 < 2 = \text{ma}(1)$. Pertanto, A non è diagonalizzabile né su $\mathbb{R}$ né su $\mathbb{C}$.

• La matrice $A = \begin{pmatrix} 4 & -3 & -3 \\ 1 & 0 & -1 \\ 1 & -1 & 0 \end{pmatrix}$ ha polinomio caratteristico

$$\begin{aligned} p_A(t) &= \det \begin{pmatrix} t-4 & 3 & 3 \\ -1 & t & 1 \\ -1 & 1 & t \end{pmatrix} = -\det \begin{pmatrix} 3 & 3 \\ t & 1 \end{pmatrix} + (-1) \begin{pmatrix} t-4 & 3 \\ -1 & 1 \end{pmatrix} + t \begin{pmatrix} t-4 & 3 \\ -1 & t \end{pmatrix} \\ &= -(3-3t) - (t-1) + t(t^2-4t+3) = t^3 - 4t^2 + 5t - 2 = (t-1)^2(t-2) \end{aligned}$$

Il polinomio caratteristico è lo stesso dell'esempio precedente, quindi procediamo direttamente al calcolo di $\text{mg}(1) = \dim V_1$:

$$V_1 = \ker(1I_3 - A) = \ker \begin{pmatrix} -3 & 3 & 3 \\ -1 & 1 & 1 \\ -1 & 1 & 1 \end{pmatrix} = \ker \begin{pmatrix} 1 & -1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \text{Span} \left(\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right)$$

quindi $\text{mg}(1) = \dim V_1 = 2 = 2 = \text{ma}(1)$. Quindi, A è diagonalizzabile sia su $\mathbb{R}$ che su $\mathbb{C}$.

Per determinare P ci manca una base di V_2 , quindi la determiniamo:

$$V_2 = \ker(2I_3 - A) = \ker \begin{pmatrix} -2 & 3 & 3 \\ -1 & 2 & 1 \\ -1 & 1 & 2 \end{pmatrix} = \ker \begin{pmatrix} 1 & -1 & -2 \\ -1 & 2 & 1 \\ -2 & 3 & 3 \end{pmatrix} = \ker \begin{pmatrix} 1 & -1 & -2 \\ 0 & 1 & -1 \\ 0 & 1 & -1 \end{pmatrix} = \text{Span} \left(\begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} \right)$$

Sia $P := \begin{pmatrix} 1 & 1 & 3 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$ e $D := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$, allora $D = P^{-1}AP$.

4.5 Complementi: matrice compagna

Lo studente curioso potrebbe porsi la seguente

Domanda. “Qualsiasi polinomio *monico* $p(t) = t^n + a_{n-1}t^{n-1} + \dots + a_1t + a_0 \in \mathbb{K}[t]$ è il polinomio caratteristico di una matrice $A \in \mathcal{M}_{\mathbb{K}}(n)$?”.

In questa sezione rispondiamo positivamente a tale domanda.

Definizione 4.16. La *matrice compagna* del polinomio monico $p(t) = t^n + a_{n-1}t^{n-1} + \dots + a_1t + a_0 \in \mathbb{K}[t]$ è la matrice quadrata $C(p) \in \mathcal{M}_{\mathbb{K}}(n)$ definita come

$$C(p) = \begin{pmatrix} 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & \dots & 0 & -a_1 \\ 0 & 1 & \dots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & -a_{n-1} \end{pmatrix}$$

Esempio. • La matrice compagna di $t^2 - 2$ è $\begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix}$.

• La matrice compagna di $t^3 - 2t^2 + 1$ è $\begin{pmatrix} 0 & 0 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 2 \end{pmatrix}$.

Proposizione 4.17. Il polinomio caratteristico di $C(p)$ è $p(t)$.

Dimostrazione. Il polinomio caratteristico di $C(p)$ è $\det(tI_n - C(p)) =$

$$\det \begin{pmatrix} t & 0 & \dots & 0 & a_0 \\ -1 & t & \dots & 0 & a_1 \\ 0 & -1 & \dots & 0 & a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & -1 & t+a_{n-1} \end{pmatrix} = (-1)^{n+1} a_0 \underbrace{\det \begin{pmatrix} -1 & t & \dots & 0 \\ 0 & -1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & -1 \end{pmatrix}}_{=U} + t \underbrace{\det \begin{pmatrix} t & 0 & \dots & 0 & a_1 \\ -1 & t & \dots & 0 & a_2 \\ 0 & -1 & \dots & 0 & a_3 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & -1 & t+a_{n-1} \end{pmatrix}}_{=M}$$

Poiché il determinante della matrice triangolare superiore $U \in \mathcal{M}_{\mathbb{K}}(n-1)$ è $\det U = (-1)^{n-1}$, otteniamo $\det(tI_n - C(p)) = a_0 + t \det(M)$. Visto che $M \in \mathcal{M}_{\mathbb{K}}(n-1)$ è della stessa forma di $tI_n - C(p)$, possiamo ripetere il ragionamento precedente espandendo il determinante di $M \in \mathcal{M}_{\mathbb{K}}(n-1)$ lungo la prima riga, e quindi $\det(tI_n - C(p)) = a_0 + a_1t + t^2 \det(N)$, dove $N \in \mathcal{M}_{\mathbb{K}}(n-2)$ è ottenuta da $C(p)$ rimuovendo le prime due righe e colonne.

Ripetendo questo processo otteniamo

$$\begin{aligned} \det(tI_n - C(p)) &= a_0 + a_1t + a_2t^2 + \dots + a_{n-3}t^{n-3} + t^{n-2} \det \begin{pmatrix} t & a_{n-2} \\ -1 & t + a_{n+1} \end{pmatrix} \\ &= a_0 + a_1t + \dots + t^{n-2}(a_{n+2} + ta_{n+1} + t^2) = p(t) \end{aligned} \quad \square$$

Esempio. • Vogliamo una matrice $M \in \mathcal{M}_{\mathbb{R}}(3)$ il cui polinomio caratteristico abbia tre radici reali, ma solo una razionale, ad esempio $p(t) = (t-1)(t^2-2) = t^3 - t^2 - 2t + 2$.

Scegliamo $M = C(p) = \begin{pmatrix} 0 & 0 & -2 \\ -1 & 0 & 2 \\ 0 & -1 & 1 \end{pmatrix}$

• Vogliamo una matrice $N \in \mathcal{M}_{\mathbb{R}}(3)$ il cui polinomio caratteristico abbia una sola radice reale, ad esempio $q(t) = (t+3)(t^2+1) = t^3 + 3t^2 + t + 3$.

Scegliamo $N = C(q) = \begin{pmatrix} 0 & 0 & -3 \\ -1 & 0 & -1 \\ 0 & -1 & -3 \end{pmatrix}$

Capitolo 5

Spazi euclidei

In questo capitolo discutiamo le proprietà geometriche degli spazi vettoriali reali relative alla loro struttura *euclidea*. Introduciamo il concetto di *prodotto scalare*: un “prodotto di due vettori”, che ci permetterà di definire angoli e lunghezze, l’ortogonalità e le proiezioni ortogonali. Vedremo anche come misurare distanze tra sottospazi (affini), e come la struttura euclidea influenzi la diagonalizzazione di endomorfismi.

5.1 Forme bilineari

Definizione 5.1. Sia V uno spazio vettoriale. Una *forma bilineare* su V è una funzione $f : V \times V \rightarrow \mathbb{K}$ che soddisfa le seguenti condizioni:

- B1) $f(v_1 + v_2, w) = f(v_1, w) + f(v_2, w)$ per tutti $v_1, v_2, w \in V$
- B2) $f(\lambda \cdot v, w) = \lambda \cdot f(v, w)$ per tutti gli $v, w \in V, \lambda \in \mathbb{K}$
- B3) $f(v, w_1 + w_2) = f(v, w_1) + f(v, w_2)$ per tutti $v, w_1, w_2, w \in V$
- B4) $f(v, \lambda \cdot w) = \lambda \cdot f(v, w)$ per tutti gli $v, w \in V, \lambda \in \mathbb{K}$

Bilineare significa *lineare in entrambe le componenti*, e valgono

$$f\left(\sum_i a_i v_i, w\right) = \sum_i a_i f(v_i, w), \quad f\left(v, \sum_j b_j w_j\right) = \sum_j b_j f(v, w_j)$$

L’insieme di tutte le forme bilineari su V viene indicato con $\text{Bil}(V)$ ed è un sottospazio vettoriale di $F(V \times V, \mathbb{K})$ (vedi Sezione 3.1.1).

Lemma 5.2. Sia $f \in \text{Bil}(V)$, allora $f(0_V, w) = f(v, 0_V) = 0_{\mathbb{K}}$ per ogni $v, w \in V$.

Dimostrazione. $f(0_V, w) = f(0_{\mathbb{K}}v, w) = 0_{\mathbb{K}}f(v, w) = 0_{\mathbb{K}}$; similmente $f(w, 0_V) = 0_{\mathbb{K}}$. $\square$

Esempio. • Su $V = \mathbb{R}^2$, sia $f\left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}\right) = x_1x_2 - y_1y_2$.

• Su $V = \mathbb{R}^2$, sia $g\left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}\right) = x_1x_2 - x_1y_2 + y_1y_2$.

• Sia $A \in \mathcal{M}_{\mathbb{K}}(n)$, allora A definisce una forma bilineare su $\mathbb{K}^n$: $f(v, w) = v^T A w$.

Per esempio:

$$f\left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}\right) = x_1x_2 - y_1y_2 = (x_1, y_1) \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$$

$$g\left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}\right) = x_1x_2 - x_1y_2 + y_1y_2 = (x_1, y_1) \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$$

Quindi ogni matrice quadrata definisce una forma bilineare. Come per le applicazioni lineari, vale anche il viceversa una volta scelta una base $\mathcal{B} = \{b_1, \dots, b_n\}$ di V .

Definizione 5.3. Sia $\mathcal{B} = \{b_1, \dots, b_n\}$ una base di V , e sia $f \in \text{Bil}(V)$.

La *matrice di Gram di f rispetto a $\mathcal{B}$* (o semplicemente matrice di f rispetto a $\mathcal{B}$) è la matrice

$$A = \mathcal{M}_{\mathcal{B}}(f) \in \mathcal{M}_{\mathbb{K}}(n), \quad a_{i,j} := f(b_i, b_j).$$

Proposizione 5.4. Sia $\mathcal{B} = \{b_1, \dots, b_n\}$ una base di V , e sia $f \in \text{Bil}(V)$.

Allora $f(v, w) = (v)_{\mathcal{B}}^T \cdot \mathcal{M}_{\mathcal{B}}(f) \cdot (w)_{\mathcal{B}}$

Dimostrazione. Se $v = \sum_{i=1}^n x_i b_i$ e $w = \sum_{j=1}^n y_j b_j$, allora

$$\begin{aligned} f(v, w) &= f\left(\sum_{i=1}^n x_i b_i, w\right) = \sum_{i=1}^n x_i f(b_i, w) = \sum_{i=1}^n x_i f(b_i, \sum_{j=1}^n y_j b_j) = \sum_{i=1}^n x_i \sum_{j=1}^n y_j f(b_i, b_j) \\ &= \sum_{i,j=1}^n x_i f(b_i, b_j) y_j = \sum_{i,j=1}^n x_i a_{i,j} y_j = (x_1, \dots, x_n) A \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \quad \square \end{aligned}$$

Esempio. Sia $V = \mathbb{R}[t]_{\leq 1}$ e sia $h : V \times V \rightarrow \mathbb{R}$, $h(p(t), q(t)) = p(0)q(0) + 3p(1)q(1)$.

È facile vedere che $h \in \text{Bil}(V)$. La sua matrice di Gram rispetto alla base canonica $\mathcal{E} = \{1, t\}$ è

$$\mathcal{M}_{\mathcal{E}}(h) = \begin{pmatrix} h(1, 1) & h(1, t) \\ h(t, 1) & h(t, t) \end{pmatrix} = \begin{pmatrix} 4 & 3 \\ 3 & 3 \end{pmatrix}$$

La matrice di Gram di h rispetto alla base $\mathcal{B} = \{t, 1-t\}$ è

$$\mathcal{M}_{\mathcal{B}}(h) = \begin{pmatrix} h(t, t) & h(t, 1-t) \\ h(1-t, t) & h(1-t, 1-t) \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}$$

Siano $\mathcal{B}$ e $\mathcal{C}$ basi di V , $f \in \text{Bil}(V)$ e sia $P := \mathcal{M}_{\mathcal{B}}^{\mathcal{C}}(id_V)$ la matrice di cambio di base. Visto che per ogni $v \in V$ vale $(v)_{\mathcal{B}} = \mathcal{M}_{\mathcal{B}}^{\mathcal{C}}(id_V) \cdot (v)_{\mathcal{C}}$, allora le matrici di Gram $A := \mathcal{M}_{\mathcal{B}}(f)$ e $B := \mathcal{M}_{\mathcal{C}}(f)$ sono legate come segue:

$$B = P^T \cdot A \cdot P$$

Infatti,

$$(B)_{i,j} = f(c_i, c_j) = f\left(\sum_{s=1}^n p_{s,i} b_s, \sum_{t=1}^n p_{t,j} b_t\right) = \sum_{s,t=1}^n p_{s,i} f(b_s, b_t) p_{t,j} = \sum_{s,t=1}^n (P^T)_{i,s} (A)_{s,t} p_{t,j} = (P^T A P)_{i,j}$$

Definizione 5.5. Due matrici $A, B \in \mathcal{M}_{\mathbb{K}}(n)$ si dicono *congruenti* se esiste una matrice invertibile $P \in \mathcal{M}_{\mathbb{K}}(n)$ tali che $B = P^T \cdot A \cdot P$, cioè se A e B rappresentano la stessa forma bilineare.

Definizione 5.6. Una forma bilineare $f \in \text{Bil}(V)$ si dice *simmetrica* se $f(v, w) = f(w, v)$ per ogni $v, w \in V$.

Una matrice $A \in \mathcal{M}_{\mathbb{K}}(n)$ è *simmetrica* se $A^T = A$.

Esempio. Le forme bilineari $f \in \text{Bil}(\mathbb{R}^2)$ e $h \in \text{Bil}(\mathbb{R}[t]_{\leq 1})$ degli esempi precedenti sono simmetriche, mentre $g \in \text{Bil}(\mathbb{R}^2)$ non lo è: $g((1, 0)^T, (0, 1)^T) = -1 \neq 0 = g((0, 1)^T, (1, 0)^T)$.

Osservazione. Sia $A \in \mathcal{M}_{\mathbb{K}}(n)$ una matrice simmetrica: $A^T = A$, e sia $P \in \mathcal{M}_{\mathbb{K}}(n)$ una matrice invertibile, allora

$$(P^T A P)^T = P^T A^T (P^T)^T = P^T A P$$

cioè ogni matrice congruente ad una matrice simmetrica è a sua volta simmetrica.

Proposizione 5.7. Sia $f \in \text{Bil}(V)$; allora le seguenti sono equivalenti:

1. f è simmetrica;
2. la matrice di Gram $\mathcal{M}_{\mathcal{B}}(f)$ è simmetrica per una (e quindi ogni) base $\mathcal{B}$ di V .

Dimostrazione. 1. $\Rightarrow$ 2.] Sia $\mathcal{B} = \{b_1, \dots, b_n\}$ una base di V , allora $a_{i,j} = f(b_i, b_j) = f(b_j, b_i) = a_{j,i}$.

2. $\Rightarrow$ 1.] Sia $\mathcal{B} = \{b_1, \dots, b_n\}$ una base di V , e supponiamo che $A := \mathcal{M}_{\mathcal{B}}(f)$ sia simmetrica. Visto che ogni matrice 1×1 è simmetrica abbiamo che

$$f(v, w) = ((v)_{\mathcal{B}})^T A (w)_{\mathcal{B}} = (((v)_{\mathcal{B}})^T A (w)_{\mathcal{B}})^T = ((w)_{\mathcal{B}})^T A (v)_{\mathcal{B}} = f(w, v). \quad \square$$

Sospendiamo per il momento la discussione sulle forme bilineari (la riprenderemo più avanti) e ci focalizziamo su una classe particolare di forme bilineari: i *prodotti scalari*.

Attenzione! Benché alcuni dei risultati che seguiranno valgano su un campo $\mathbb{K}$ qualsiasi, nella prossima sezione abbiamo bisogno di un campo numerico *ordinato* (cioè dobbiamo saper dire se uno scalare è positivo o meno), come $\mathbb{Q}$ o $\mathbb{R}$, ma non $\mathbb{C}$ o il campo con due elementi descritto nel Capitolo 1.

Su $\mathbb{R}$ abbiamo la possibilità di estrarre la radice quadrata dei numeri positivi (su $\mathbb{Q}$ questo non si può fare), quindi *d'ora in poi fino alla fine di queste note lavoreremo esclusivamente sul campo reale $\mathbb{K} = \mathbb{R}$* .

5.2 Prodotti scalari

Definizione 5.8. Sia V uno spazio vettoriale reale. Una forma bilineare simmetrica $f \in \text{Bil}(V)$ è un *prodotto scalare* se f è *definita positiva*, cioè $f(v, v) \geq 0$ per ogni $v \in V$ e $f(v, v) = 0 \Leftrightarrow v = 0_V$.

Uno spazio vettoriale reale V equipaggiato con un prodotto scalare, si chiama *spazio euclideo*.

Notazione. Se $f \in \text{Bil}(V)$ è un prodotto scalare scriviamo $\langle v, w \rangle$ invece di $f(v, w)$.

Esempio. Su $\mathbb{R}^n$ la forma bilineare data da

$$\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}, \quad \langle v, w \rangle = \left\langle \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}, \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \right\rangle = v_1 w_1 + \cdots + v_n w_n.$$

è un prodotto scalare detto *prodotto scalare standard* (o canonico), e la sua matrice di Gram rispetto alla base canonica è la matrice identità I_n .

$$\text{Ad esempio } (n=3): \left\langle \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \right\rangle = 0 \cdot 1 + 2 \cdot 1 + 3 \cdot (-1) = -1.$$

$$\left\langle \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \right\rangle = 1^2 + 2^2 + 3^2 = 14, \quad \left\langle \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, -2 \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} + \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \right\rangle = -2(-1) + 14 = 16$$

Esempio. • La forma bilineare $f \in \text{Bil}(\mathbb{R}^3)$ data dalla matrice $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$ è simmetrica ($A = A^T$) ed è definita positiva:

$$f(v, v) = \begin{pmatrix} v_1 & v_2 & v_3 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = v_1^2 + 2v_1v_2 + 2v_2^2 + 3v_3^2 = (v_1 + v_2)^2 + v_2^2 + 3v_3^2 \geq 0$$

e $f(v, v) = 0$ se e solo se $v_1 + v_2 = v_2 = v_3 = 0$, cioè se $v = 0_V$.

• La forma bilineare $g \in \text{Bil}(\mathbb{R}^4)$ data dalla matrice $D = \text{diag}(-1, 1, 1, 1)$ è simmetrica ($D = D^T$) ma non è definita positiva: $f(e_1, e_1) = -1 < 0$.

Esempio. • Sia $V = \mathcal{M}_{\mathbb{R}}(m, n)$. La forma bilineare $\langle A, B \rangle = \text{tr}(A^T B)$ è un prodotto scalare.

• Sia $V = \mathbb{R}[t]_{\leq 1}$. La forma bilineare $\langle p(t), q(t) \rangle = p(0)q(0) + 3p(1)q(1)$ è un prodotto scalare, infatti per $p(t) = a + bt$ vale $\langle p(t), p(t) \rangle = a^2 + 3(a + b)^2 \geq 0$ e vale $= 0$ se e solo se $a = a + b = 0$, cioè $p(t)$ è il polinomio nullo.

Vediamo ora come misurare la “lunghezza” di un vettore.

Definizione 5.9. Sia $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo. La *norma* di un vettore $v \in V$ è $\|v\| = \sqrt{\langle v, v \rangle}$.

Un vettore $v \in V$ è un *vettore unitario* (o versore) se $\|v\| = 1$.

Si noti che, se $v \in V, v \neq 0$, allora $\frac{v}{\|v\|}$ è un vettore unitario.

Esempio. In $V = \mathbb{R}^3$ con il prodotto scalare standard: $\|(1, 2, 3)^T\| = \sqrt{14}$ quindi $(\frac{1}{\sqrt{14}}, \frac{2}{\sqrt{14}}, \frac{3}{\sqrt{14}})^T$ è un vettore unitario.

Proposizione 5.10. Siano $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo, e $v, w \in V$. Allora

1. $\|v\| \geq 0$ e $\|v\| = 0 \Leftrightarrow v = 0$;

2. $||\lambda \cdot v|| = |\lambda| \cdot ||v||$ per tutti i $\lambda \in \mathbb{R}$;
 3. $\langle v, w \rangle = \frac{||v+w||^2 - ||v||^2 - ||w||^2}{2}$ (Formula di polarizzazione).

Dimostrazione. Le proprietà 1. e 2. derivano dalla definizione di norma.

3. Espandendo $||v+w||^2$ sfruttando la bilinearità del prodotto scalare otteniamo:

$$||v+w||^2 = \langle v+w, v+w \rangle = \langle v, v \rangle + 2\langle v, w \rangle + \langle w, w \rangle = ||v||^2 + ||w||^2 + 2\langle v, w \rangle$$

e la tesi segue immediatamente. $\square$

Teorema 5.11 (Disuguaglianza di Cauchy-Schwarz). *Sia $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo. Allora, per ogni $v, w \in \mathbb{R}^n$ vale*

$$|\langle v, w \rangle| \leq ||v|| \cdot ||w|| \quad (5.1)$$

Inoltre, in (5.1) vale l'uguaglianza se e solo se v e w sono linearmente dipendenti.

Dimostrazione. Se $w = 0$, allora $|\langle v, w \rangle| = 0 \leq ||v|| \cdot ||w||$; quindi supponiamo che $w \neq 0$.

I vettori v, w sono linearmente dipendenti se e solo se esiste $\lambda \in \mathbb{R}$ tale che $v = \lambda w$, ovvero $||v - \lambda w||^2 = 0$. Per utilizzare questa osservazione, introduciamo la funzione ausiliaria $g : \mathbb{R} \rightarrow \mathbb{R}$ definita tramite

$$g(t) := ||v - tw||^2 = \langle v - tw, v - tw \rangle = \langle v, v \rangle - 2t\langle v, w \rangle + t^2\langle w, w \rangle$$

Quindi, $g(t) = t^2||w||^2 - 2t\langle v, w \rangle + ||v||^2$ è un polinomio di grado 2 e $g(t) \geq 0$ per ogni $t \in \mathbb{R}$; inoltre, esiste $\lambda \in \mathbb{R}$ tale che $g(\lambda) = 0$ se e solo se v e w sono linearmente dipendenti. Pertanto, il discriminante di g

$$\Delta = 4\langle v, w \rangle^2 - 4||w||^2||v||^2 \leq 0 \Leftrightarrow \langle v, w \rangle^2 \leq ||w||^2||v||^2$$

e $\langle v, w \rangle^2 = ||w||^2||v||^2$ se e solo se v e w sono linearmente dipendenti.

Poiché $||w|| \cdot ||v|| \geq 0$, l'enunciato segue da un'estrazione di radice quadrata. $\square$

La disuguaglianza di Cauchy-Schwarz ha una serie di importanti conseguenze:

Lemma 5.12 (Disuguaglianza triangolare). *Sia $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo. Allora per ogni $v, w \in V$ vale*

$$||v+w|| \leq ||v|| + ||w||$$

Dimostrazione. Trattandosi di una disuguaglianza tra numeri reali positivi, la sua validità non cambia se eleviamo entrambi i membri al quadrato.

Combinando la formula di polarizzazione con la disuguaglianza di Cauchy-Schwarz otteniamo

$$||v+w||^2 = ||v||^2 + ||w||^2 + 2\langle v, w \rangle \leq ||v||^2 + ||w||^2 + 2||v|| \cdot ||w|| = (||v|| + ||w||)^2 \quad \square$$

5.2.1 Distanza

Su uno spazio euclideo $(V, \langle \cdot, \cdot \rangle)$, la norma ci permette di definire la *distanza* tra due vettori, cioè una funzione $d : V \times V \rightarrow \mathbb{R}$, $d(v, w) = \|v - w\|$ che gode delle seguenti

Proprietà.

- $d(v, w) \geq 0$, e $d(v, w) = 0 \Leftrightarrow v = w$
- $d(v, w) = d(w, v)$
- $d(v, z) \leq d(v, w) + d(w, z)$ (*disuguaglianza triangolare della distanza*)

5.2.2 Angolo tra vettori

Siano $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo e $v, w \in V$ vettori diversi da zero: $\|v\|, \|w\| > 0$. Dalla disuguaglianza di Cauchy-Schwarz otteniamo:

$$-1 \leq \frac{\langle v, w \rangle}{\|w\| \|v\|} \leq 1$$

Quindi possiamo definire l'*angolo* tra v, w come segue:

Definizione 5.13. Siano $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo e $v, w \in V$ vettori non nulli: $\|v\|, \|w\| > 0$. L'*angolo* $\theta_{v,w}$ tra v e w è l'unico $\theta \in [0, \pi]$ tale che

$$\cos \theta = \frac{\langle v, w \rangle}{\|w\| \cdot \|v\|}.$$

Osservazione. Da questa definizione deduciamo che $\langle v, w \rangle = \|v\| \cdot \|w\| \cdot \cos \theta_{v,w}$, che coincide con la definizione di prodotto scalare che il lettore probabilmente ha già visto in fisica.

Esempio. • In $\mathbb{R}^3$ equipaggiato con il prodotto scalare standard, consideriamo $v = (1, 2, 3)^T$ e $w = (0, 1, -1)^T$. L'angolo tra loro è

$$\arccos \left(\frac{-1}{\sqrt{14}\sqrt{2}} \right) = 1,76\dots$$

• In $V = \mathbb{R}[t]_{\leq 1}$ con il prodotto scalare $\langle p(t), q(t) \rangle = p(0)q(0) + 3p(1)q(1)$, l'angolo tra $v = t$ e $w = t - 1$ è $\frac{\pi}{2}$, infatti $\langle v, w \rangle = 0$.

Alla luce di questa definizione, la formula di polarizzazione può essere riscritta in una forma diversa che prende il nome di *formula di Carnot* o *teorema del coseno*, che è una generalizzazione del celebre *teorema di Pitagora*

Proposizione 5.14 (Formula di Carnot e Teorema di Pitagora). *Siano $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo, e $v, w \in V$, allora*

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2 + 2\|v\| \cdot \|w\| \cos(\theta_{v,w})$$

In particolare se $\langle v, w \rangle = 0$ allora $\|v + w\|^2 = \|v\|^2 + \|w\|^2$.

5.3 Ortogonalità

Definizione 5.15. Sia $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo. Due vettori $v, w \in V$ sono *ortogonali* se $\langle v, w \rangle = 0$ (si scrive $v \perp w$).

Due sottospazi vettoriali $U, W \triangleleft V$ sono *ortogonali* se $\langle u, w \rangle = 0 \ \forall u \in U, w \in W$.

Osservazione. i) Il vettore nullo è ortogonale a qualsiasi altro vettore: $0_V \perp v, \forall v \in V$.

ii) Il vettore nullo è l'unico vettore ortogonale a se stesso: $v \perp v \Leftrightarrow \langle v, v \rangle = 0 \Leftrightarrow v = 0_V$.

Esempio. In $V = \mathbb{R}[t]_{\leq 1}$ con il prodotto scalare $\langle p(t), q(t) \rangle = p(0)q(0) + 3p(1)q(1)$, vogliamo trovare tutti i vettori ortogonali a $p(t) = t - 1$. Passando a coordinate rispetto alla base $\mathcal{E} = \{1, t\}$ vogliamo $a, b \in \mathbb{R}$ tali che $0 = (a, b) \begin{pmatrix} 4 & 3 \\ 3 & 3 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \end{pmatrix} = (a, b) \begin{pmatrix} -1 \\ 0 \end{pmatrix}$, cioè (a, b) è un multiplo di $(0, 1)$: i vettori ortogonali a $p(t) = t - 1$ sono $\text{Span}(t)$.

Nella Sezione 2.7, abbiamo introdotto i sottospazi affini di $\mathbb{K}^n$ (sono le soluzioni di un sistema lineare risolubile), e abbiamo visto che si ottengono prendendo un sottospazio $W \triangleleft \mathbb{K}^n$ e traslandolo di un vettore v : $S = v + W$. Nessuno ci vieta di ripetere la stessa procedura per uno spazio vettoriale qualsiasi, ad esempio in $\mathbb{R}[t]_{\leq 3}$ prendiamo $S = t + \text{Span}(1, 1 + t^2 - t^3)$ e la sua giacitura è il sottospazio vettoriale associato: $S_0 = \text{Span}(1, 1 + t^2 - t^3)$.

Definizione 5.16. Sia $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo. Due sottospazi affini $S_1, S_2 \triangleleft V$ sono *ortogonali*, se le loro giaciture sono sottospazi vettoriali ortogonali.

Esempio. Consideriamo $\mathbb{R}^3$ con il prodotto scalare standard.

• Le rette $r_1 = \begin{pmatrix} 4 \\ 1 \\ 2 \end{pmatrix} + \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \mathbb{R}$ e $r_2 = \begin{pmatrix} 7 \\ -1 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \mathbb{R}$ sono ortogonali, infatti $\langle \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \rangle = 0$.

• Sia r la retta di equazione $r = \begin{cases} x + y = 10 \\ 2y + 2z = -17 \end{cases}$, allora ogni piano di equazione parametrica $P + \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \mathbb{R} + \begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix} \mathbb{R}$ ($P \in \mathbb{R}^3$) è ortogonale a r , infatti la giacitura della retta è $r_0 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, e quindi ogni piano di equazione $x - y + z = d$ ($d \in \mathbb{R}$) è ortogonale a r .

• Sia π il piano di equazione $\pi = x - 2y + 5z = 20$, allora ogni retta di equazione parametrica $P + \begin{pmatrix} 1 \\ -2 \\ 5 \end{pmatrix} \mathbb{R}$ ($P \in \mathbb{R}^3$) è ortogonale a π .

Osservazione. In $\mathbb{R}^3$ con il prodotto scalare standard, sia $\pi := \{\alpha x + \beta y + \gamma z = \delta\} \subset \mathbb{R}^3$ un piano, il vettore $(\alpha, \beta, \gamma)^T$ è ortogonale a π ed è il *vettore normale* del piano (o vettore direttore). È unico a meno di riscalamento.

5.3.1 Basi ortogonali e basi ortonormali

Definizione 5.17. Sia $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo e sia $W \triangleleft V$ un sottospazio vettoriale con base $\mathcal{B} = \{b_1, \dots, b_p\}$.

$\mathcal{B}$ è una *base ortogonale* di W se $\langle b_i, b_j \rangle = 0 \ \forall i \neq j$.

$\mathcal{B}$ è una *base ortonormale* di W se $\langle b_i, b_j \rangle = 0 \ \forall i \neq j$ e $\|b_i\| = 1 \ \forall i$.

Esempio. • La base canonica $\mathcal{E}$ di $\mathbb{R}^n$ è una base ortonormale di $\mathbb{R}^n$ per il prodotto scalare standard.

• La base $\{(1, 1)^T, (1, -1)^T\}$ di $\mathbb{R}^2$ è una base ortogonale, ma non è una base ortonormale, per il prodotto scalare standard. Per ottenere una base ortonormale, dobbiamo riscalare ogni vettore: $\|(1, 1)^T\| = \|(1, -1)^T\| = \sqrt{2}$, quindi $\left\{\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)^T, \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)^T\right\}$ è una base ortonormale di $\mathbb{R}^2$.

• La base $\{t, t-1\}$ di $\mathbb{R}[t]_{\leq 1}$ è una base ortogonale, ma non è una base ortonormale, per il prodotto scalare $\langle p(t), q(t) \rangle = p(0)q(0) + 3p(1)q(1)$. Dato che $\|t\|^2 = 3$ e $\|t-1\|^2 = 1$, allora $\left\{\frac{t}{\sqrt{3}}, t-1\right\}$ è una base ortonormale.

Teorema 5.18. Sia $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo e sia $\mathcal{B}_W = \{b_1, \dots, b_p\}$ una base ortogonale di $W \triangleleft \mathbb{R}^n$. Allora per ogni $v \in W$ abbiamo

$$v = \frac{\langle v, b_1 \rangle}{\|b_1\|^2} b_1 + \dots + \frac{\langle v, b_p \rangle}{\|b_p\|^2} b_p \quad (5.2)$$

Se $\mathcal{B}_W = \{b_1, \dots, b_p\}$ è una base ortonormale di W , allora $v = \langle v, b_1 \rangle b_1 + \dots + \langle v, b_p \rangle b_p$.

Definizione 5.19. Il coefficiente $\lambda_j = \frac{\langle v, b_j \rangle}{\|b_j\|^2}$ è il *coefficiente di Fourier* di v rispetto al vettore b_j .

Dimostrazione. $\mathcal{B}_W$ è una base di W , quindi per ogni $v \in W$ esistono (unici) scalari $\lambda_j \in \mathbb{R}$ tali che $v = \lambda_1 b_1 + \dots + \lambda_p b_p$. Calcolando $\langle v, b_j \rangle$ otteniamo

$$\langle v, b_j \rangle = \left\langle \sum_{i=1}^p \lambda_i b_i, b_j \right\rangle = \sum_{i=1}^p \lambda_i \langle b_i, b_j \rangle = \lambda_j \langle b_j, b_j \rangle \implies \lambda_j = \frac{\langle v, b_j \rangle}{\|b_j\|^2} \quad \square$$

Esempio. • La base $\{(1, -2)^T, (2, 1)^T\}$ è una base ortogonale di $\mathbb{R}^2$ con il prodotto scalare standard, e $v = (4, 3)^T$ si scompone come:

$$\begin{pmatrix} 4 \\ 3 \end{pmatrix} = \frac{\langle (4, 3)^T, (1, -2)^T \rangle}{\|(1, -2)^T\|^2} \begin{pmatrix} 1 \\ -2 \end{pmatrix} + \frac{\langle (4, 3)^T, (2, 1)^T \rangle}{\|(2, 1)^T\|^2} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = -\frac{2}{5} \begin{pmatrix} 1 \\ -2 \end{pmatrix} + \frac{11}{5} \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

• La base $\{t, t-1\}$ di $\mathbb{R}[t]_{\leq 1}$ è una base ortogonale, per il prodotto scalare $\langle p(t), q(t) \rangle = p(0)q(0) + 3p(1)q(1)$. Allora $1+t$ si scompone come:

$$1+t = \frac{\langle 1+t, t \rangle}{\|t\|^2} t + \frac{\langle 1+t, t-1 \rangle}{\|t-1\|^2} (t-1) = \frac{6}{3} t + \frac{-1}{1} (t-1) = 2t - 1(t-1)$$

Osservazione. 1) Con un calcolo simile a quello appena visto si dimostra che se $\{v_1, \dots, v_s\} \subset V$ sono vettori a due a due ortogonali e $v_i \neq 0_V$, allora $\{v_1, \dots, v_s\}$ è un insieme di vettori linearmente indipendenti.

$$0_V = \sum_{j=1}^s x_j v_j \implies 0 = \langle 0_V, v_i \rangle = \left\langle \sum_{j=1}^s x_j v_j, v_i \right\rangle = \sum_{j=1}^s x_j \langle v_j, v_i \rangle = x_i \|v_i\|^2 \implies x_i = 0$$

2) Sia $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo e sia $\mathcal{B} = \{b_1, \dots, b_n\}$ una sua base ortogonale. Allora la matrice di Gram $\mathcal{M}_{\mathcal{B}}(\langle \cdot, \cdot \rangle)$ è diagonale.

Se inoltre $\mathcal{B}$ è ortonormale, allora $\mathcal{M}_{\mathcal{B}}(\langle \cdot, \cdot \rangle) = I_n$.

Quindi usando una base $\mathcal{B}$ ortonormale, si può calcolare il prodotto scalare usando le coordinate: $\langle v, w \rangle = ((v)_{\mathcal{B}})^T \cdot (w)_{\mathcal{B}} = \langle (v)_{\mathcal{B}}, (w)_{\mathcal{B}} \rangle_S$, dove $\langle \cdot, \cdot \rangle_S$ denota il prodotto scalare standard di $\mathbb{R}^n$.

Vediamo ora un processo per costruire una base ortogonale (ortonormale) di un sottospazio vettoriale $W \triangleleft V$ partendo da una base $\mathcal{B} = \{b_1, \dots, b_p\}$ di W .

5.3.2 Ortonormalizzazione di Gram-Schmidt

Siano $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo e $W \triangleleft V$ un suo sottospazio vettoriale. Il processo di ortonormalizzazione di Gram-Schmidt ci assicura che W ha sempre una base ortonormale, inoltre la dimostrazione è algoritmica, cioè ci da un metodo concreto per calcolare una base ortonormale di W .

Teorema 5.20 (Ortonormalizzazione di Gram-Schmidt). *Siano $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo, e sia $\mathcal{B} = \{b_1, \dots, b_p\}$ una base di $W \triangleleft V$. Allora, esistono una base ortogonale $\mathcal{C} = \{c_1, \dots, c_p\}$ e una base ortonormale $\mathcal{D} = \{d_1, \dots, d_p\}$ di W , tali che*

$$\text{Span}(b_1, \dots, b_j) = \text{Span}(c_1, \dots, c_j) = \text{Span}(d_1, \dots, d_j) \quad \forall j = 1, \dots, p$$

Attenzione: ad ogni passo non modifichiamo lo span dei primi j vettori!

Dimostrazione. Per costruire una base ortogonale $\mathcal{C}$ procediamo ricorsivamente.

Iniziamo fissando $c_1 := b_1$ e poi procediamo in modo ricorsivo per calcolare c_j usando i vettori $c_1, \dots, c_{j-1}$ appena trovati:

- $c_1 := b_1$;
- $c_2 := b_2 - \frac{\langle b_2, c_1 \rangle}{\|c_1\|^2} c_1$
- $c_3 := b_3 - \frac{\langle b_3, c_1 \rangle}{\|c_1\|^2} c_1 - \frac{\langle b_3, c_2 \rangle}{\|c_2\|^2} c_2$
- $\dots c_j := b_j - \sum_{k=1}^{j-1} \frac{\langle b_j, c_k \rangle}{\|c_k\|^2} c_k \dots$

Infine, per determinare una base ortonormale $\mathcal{D}$ di W , normalizziamo ogni c_j : $d_j := \frac{c_j}{\|c_j\|}$.

Dobbiamo ora dimostrare che le basi trovate hanno le proprietà richieste.

Verifichiamo per induzione che ad ogni passo non modifichiamo lo span dei primi j vettori. Il passo iniziale è banalmente vero.

Supponiamo ora che $\text{Span}(b_1, \dots, b_{j-1}) = \text{Span}(c_1, \dots, c_{j-1})$, allora

$$\begin{aligned} \text{Span}(c_1, \dots, c_{j-1}, c_j) &= \text{Span}(c_1, \dots, c_{j-1}, b_j - \sum_{k=1}^{j-1} \frac{\langle b_j, c_k \rangle}{\|c_k\|^2} c_k) = \text{Span}(c_1, \dots, c_{j-1}, b_j) \\ &= \text{Span}(b_1, \dots, b_{j-1}, b_j) \end{aligned}$$

Verifichiamo ora che la base $\mathcal{C}$ ottenuta tramite il processo di Gram-Schmidt sia effettivamente una base ortogonale. Lo facciamo per induzione, supponendo che ad ogni passo i vettori $c_1, \dots, c_{j-1}$ siano a due a due ortogonali e dimostrando che il “nuovo vettore” c_j è ortogonale a ciascuno dei precedenti:

$$\langle c_i, c_j \rangle = \langle c_i, b_j - \sum_{k=1}^{j-1} \frac{\langle b_j, c_k \rangle}{\|c_k\|^2} c_k \rangle = \langle c_i, b_j \rangle - \sum_{k=1}^{j-1} \frac{\langle b_j, c_k \rangle}{\|c_k\|^2} \langle c_i, c_k \rangle = \langle c_i, b_j \rangle - \frac{\langle b_j, c_i \rangle}{\|c_i\|^2} \langle c_i, c_i \rangle = 0$$

□

Osservazione. Ogni base ortogonale (rispettivamente ortonormale) di W può essere completata in una base ortogonale (rispettivamente ortonormale) di V .

Esempio. In $\mathbb{R}^4$ equipaggiato con il prodotto scalare standard consideriamo la base $\left\{b_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, b_2 = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 1 \end{pmatrix}, b_3 = \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, b_4 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}\right\}$ e applichiamo il processo di Gram-Schmidt:

$$\begin{aligned} c_1 &= b_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \\ c_2 &= b_2 - \frac{\langle b_2, c_1 \rangle}{\|c_1\|^2} c_1 = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 1 \end{pmatrix} - \frac{2}{2} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \end{pmatrix} \\ c_3 &= b_3 - \frac{\langle b_3, c_1 \rangle}{\|c_1\|^2} c_1 - \frac{\langle b_3, c_2 \rangle}{\|c_2\|^2} c_2 = \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} - \frac{-2}{4} \begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -1/2 \\ 0 \\ 1/2 \end{pmatrix} \\ c_4 &= b_4 - \frac{\langle b_4, c_1 \rangle}{\|c_1\|^2} c_1 - \frac{\langle b_4, c_2 \rangle}{\|c_2\|^2} c_2 - \frac{\langle b_4, c_3 \rangle}{\|c_3\|^2} c_3 = b_4 - 0c_1 - 0c_2 - 0c_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \end{aligned}$$

5.3.3 Complemento ortogonale e proiezioni ortogonali

Analizziamo il significato geometrico del passo $c_2 := b_2 - \frac{\langle b_2, c_1 \rangle}{\|c_1\|^2} c_1$.

Sia θ l'angolo tra c_1 e b_2 , in modo che

$$\langle b_2, c_1 \rangle = \|b_2\| \cdot \|c_1\| \cdot \cos \theta \implies c_2 := b_2 - \frac{\langle b_2, c_1 \rangle}{\|c_1\|^2} c_1 = b_2 - \underbrace{\frac{c_1}{\|c_1\|} \cdot \|b_2\| \cdot \cos \theta}_w$$

Il vettore w è un multiplo di c_1 di lunghezza $\|b_2\| \cdot \cos \theta$, quindi stiamo rimuovendo da b_2 la sua componente che giace nella direzione di c_1 e ci rimane un vettore ortogonale a c_1 . Idea geometrica: w è la “proiezione ortogonale” di b_2 sulla direzione di c_1 , quindi c_2 è un vettore ortogonale a c_1 .

Vediamo ora come formalizzare e generalizzare questa idea.

Definizione 5.21. Siano $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo e $W \triangleleft V$ un suo sottospazio vettoriale. Il *complemento ortogonale* di W è

$$W^\perp := \{v \in V \mid \langle v, w \rangle = 0 \ \forall w \in W\}.$$

Osservazione. i) $W^\perp \triangleleft V$.

ii) Se $\mathcal{B} = \{b_1, \dots, b_p\}$ è una base di W , per verificare se $v \in W^\perp$, basta controllare che $\langle v, b_j \rangle = 0$ per ogni $j = 1, \dots, p$.

Esempio. • In $\mathbb{R}^3$ ^(†) consideriamo $W = \text{Span}((1, 1, 1)^T)$, allora $W^\perp = \{(x, y, z)^T \in \mathbb{R}^3 \mid \langle (x, y, z)^T, (1, 1, 1)^T \rangle = 0\} = \{(x, y, z)^T \in \mathbb{R}^3 \mid x + y + z = 0\}$.

• In $\mathbb{R}^4$ ^(†) consideriamo $W = \text{Span}((1, -1, 0, 0)^T, (1, 0, 1, 1)^T)$, allora $W^\perp = \{(x, y, w, z)^T \in \mathbb{R}^4 \mid \langle (x, y, w, z)^T, (1, -1, 0, 0)^T \rangle = \langle (x, y, w, z)^T, (1, 0, 1, 1)^T \rangle = 0\}$, cioè

$$W^\perp = \{(x, y, z)^T \in \mathbb{R}^4 \mid x - y = x + w + z\} = \ker \begin{pmatrix} 1 & -1 & 0 & 0 \\ 1 & 0 & 1 & 1 \end{pmatrix} = \text{Span} \left(\begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -1 \\ 0 \end{pmatrix} \right)$$

[†]In $\mathbb{R}^n$, se non specifichiamo il prodotto scalare, intendiamo il prodotto scalare standard.

• In V con base $\mathcal{B} = \{v_1, v_2\}$ consideriamo il prodotto scalare con matrice di Gram $S := \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ e sia $W = \text{Span}(v_1 - 2v_2)$.

Visto che $(v_1 - 2v_2)_{\mathcal{B}} = (1, -2)^T$, abbiamo che $W^\perp = \{v \in V \mid \langle v, v_1 - 2v_2 \rangle = 0\} = \{(v)_{\mathcal{B}} \in \mathbb{R}^2 \mid (1, -2) \cdot S \cdot (v)_{\mathcal{B}} = 0\} = \{(v)_{\mathcal{B}} \in \mathbb{R}^2 \mid (-1, -3) \cdot (v)_{\mathcal{B}} = 0\} = \text{Span}(3v_1 - v_2)$.

• Consideriamo ora $K = \{(x, y, z)^T \in \mathbb{R}^3 \mid x + 2y + 3z = 0, 4x + 5y + 6z = 0\}$. Allora un vettore in K è ortogonale a $(1, 2, 3)^T$ e a $(4, 5, 6)^T$, quindi $K^\perp = \text{Span}((1, 2, 3)^T, (4, 5, 6)^T)$.

Osservazione. In $V = \mathbb{R}^n$ con il prodotto scalare standard, dato un sottospazio W , per determinare $W^\perp$ si può procedere come segue.

• Se $W = \text{Sol}(A|0)$ (equazioni cartesiane), allora $W^\perp = \text{Row}(A) = \text{Span}(R_1^T, \dots, R_m^T)$, dove R_i sono le righe di A , infatti ogni elemento di W è soluzione di $Ax = 0$, e quindi è ortogonale a tutte le righe di A .

• Viceversa, se $W = \text{Span}(v_1, \dots, v_p)$ (equazioni parametriche), allora $W^\perp = \text{Sol}(M|0)$, dove M è la matrice che ha $v_1^T, \dots, v_p^T$ come righe.

Proposizione 5.22. *Siano $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo e $W \triangleleft V$ un suo sottospazio vettoriale. Allora*

$$W \oplus W^\perp = V$$

In particolare, $\dim W + \dim W^\perp = \dim V$.

Dimostrazione. Sia $u \in W \cap W^\perp$, allora $\langle u, u \rangle = 0$. Dunque $u = 0$, e la somma $W + W^\perp$ è diretta e dobbiamo dimostrare che è l'intero V .

Sia $v \in V$ e prendiamo una base ortogonale $\mathcal{B} = \{b_1, \dots, b_p\}$ di W e definiamo

$$w := \sum_{k=1}^p \frac{\langle v, b_k \rangle}{\|b_k\|^2} b_k \in W \quad (5.3)$$

allora $v = w + (v - w)$ e ci resta solo da verificare che $v - w \in W^\perp$:

$$\langle b_j, v - w \rangle = \langle b_j, v - \sum_{k=1}^p \frac{\langle v, b_k \rangle}{\|b_k\|^2} b_k \rangle = \langle b_j, v \rangle - \sum_{k=1}^p \frac{\langle v, b_k \rangle}{\|b_k\|^2} \langle b_j, b_k \rangle = \langle b_j, v \rangle - \frac{\langle v, b_j \rangle}{\|b_j\|^2} \langle b_j, b_j \rangle = 0$$

□

Osservazione. i) Si noti l'analogia tra le formule (5.2), (5.3) e le formule usate nell'ortonormalizzazione di Gram-Schmidt.

ii) La somma $V = W \oplus W^\perp$ è diretta, quindi ogni $v \in V$ si decompone in modo *unico* come somma di un elemento $w \in W$ ed di un elemento $u \in W^\perp$ ($v = w + u$).

La formula (5.3) ci dà un modo esplicito per determinare w , e poichè w è unico non dipende dalla base ortogonale di W scelta.

Definizione 5.23. Sia $W \triangleleft V$ un sottospazio vettoriale dello spazio euclideo $(V, \langle \cdot, \cdot \rangle)$ e sia $\mathcal{B}_W = \{b_1, \dots, b_p\}$ una base ortogonale di W . La *proiezione ortogonale* su W è la mappa lineare:

$$\begin{aligned} \pi_W : V &\longrightarrow V \\ v &\longmapsto \sum_{k=1}^p \frac{\langle v, b_k \rangle}{\|b_k\|^2} b_k \end{aligned}$$

Osservazione. i) $W^\perp = \ker(\pi_W)$ e $(W^\perp)^\perp = W$.

ii) Visto che $V = W \oplus W^\perp$, allora vale $v = \pi_W(v) + \pi_{W^\perp}(v)$, cioè $\pi_{W^\perp} = id_V - \pi_W$.

iii) La proiezione ortogonale $\pi_W(v)$ di v è il vettore di W che meglio approssima v , cioè che minimizza la distanza da v :

$$d(v, \pi_W(v)) < d(v, w) \quad \forall w \in W, w \neq \pi_W(v)$$

in altre parole,

$$\min\{d(v, w) : w \in W\} = \|v - \pi_W(v)\| = \|\pi_{W^\perp}(v)\| \quad (5.4)$$

Infatti, siano $u_1 := v - \pi_W(v) \in W^\perp$ e $u_2 := v - w$, allora $u_3 := \pi_W(v) - w \in W$ e per Pitagora $\|u_2\|^2 = \|u_1\|^2 + \|u_3\|^2 > \|u_1\|^2$ (visto che $\|u_3\| > 0$). Estruendo la radice segue la tesi.

Esempio. • Ogni passo del processo di Gram-Schmidt può essere riformulato come $c_j := b_j - \pi_{H_j}(b_j)$, dove $H_j = \text{Span}(c_1, \dots, c_{j-1})$.

• In $\mathbb{R}^3$, si considerino $v = (2, 3, 0)^T \in \mathbb{R}^3$ e $H = \text{Span}((1, 0, 1)^T, (0, 1, 0)^T) \triangleleft \mathbb{R}^3$.

$(1, 0, 1)^T$ e $(0, 1, 0)^T$ sono ortogonali, quindi la proiezione ortogonale di v su H è:

$$\pi_H(v) = \frac{\langle (2, 3, 0)^T, (1, 0, 1)^T \rangle}{\|(1, 0, 1)^T\|^2} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \frac{\langle (2, 3, 0)^T, (0, 1, 0)^T \rangle}{\|(0, 1, 0)^T\|^2} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix}$$

Esempio. • In $\mathbb{R}[t]_{\leq 1}$ con il prodotto scalare $\langle p(t), q(t) \rangle = p(0)q(0) + 3p(1)q(1)$, consideriamo $W = \text{Span}(2 - t)$, allora

$$\pi_W(a + bt) = \frac{\langle a + bt, 2 - t \rangle}{\|2 - t\|^2} (2 - t) = \frac{2a + 3(a + b)}{4 + 3} (2 - t) = \frac{5a + 3b}{7} (2 - t)$$

• In $V = \mathbb{R}^3$, consideriamo $H = \text{Span}((1, 2, 3)^T)$, quindi

$$\pi_H((x, y, z)^T) = \frac{\langle (x, y, z)^T, (1, 2, 3)^T \rangle}{\|(1, 2, 3)^T\|^2} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \frac{x + 2y + 3z}{14} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

e $H^\perp = \ker(\pi_H) = \{(x, y, z)^T \in \mathbb{R}^3 \mid x + 2y + 3z = 0\}$.

Proprietà. La proiezione ortogonale $\pi_W : V \rightarrow V$ soddisfa le seguenti proprietà:

1. $\pi_W(v) \in W$ per ogni $v \in V$;
2. $\pi_W(u) = u \Leftrightarrow u \in W$; in particolare $\pi_W \circ \pi_W = \pi_W$ (π_W è idempotente);
3. $v - \pi_W(v) \in W^\perp$.

Osservazione. Queste 3 proprietà sono *caratterizzanti* della proiezione ortogonale, cioè $\pi_W : V \rightarrow V$ è l'unico endomorfismo di V che le soddisfa simultaneamente.

Infatti, completiamo $\mathcal{B}_W$ a $\mathcal{B}' = \{b_1, \dots, b_p, b_{p+1}, \dots, b_n\}$ base ortogonale di V , e supponiamo che $f \in \text{End}(V)$ soddisfi le 3 proprietà, allora

$$f(b_1) = b_1, \dots, f(b_p) = b_p, f(b_{p+1}) = 0_V, \dots, f(b_n) = 0_V \quad (5.5)$$

Siccome π_W ha lo stesso comportamento sulla base $\mathcal{B}'$ usando il teorema di interpolazione (Teorema 3.20) otteniamo che $f = \pi_W$.

Matrici rappresentative di proiezioni ortogonali

Concludiamo questa sezione descrivendo le matrici rappresentative dell'endomorfismo $\pi_W : V \rightarrow V$, la proiezione ortogonale di V sul sottospazio $W \triangleleft V$.

Sia $\mathcal{B}_W = \{b_1, \dots, b_p\}$ una base ortogonale di W e completiamola a $\mathcal{B} = \{b_1, \dots, b_p, b_{p+1}, \dots, b_n\}$ base ortogonale di V . L'equazione (5.5) ci dice che $\mathcal{B}$ è una base di autovettori per π_W , più precisamente $\{b_1, \dots, b_p\}$ sono una base dell'autospazio relativo all'autovalore 1, mentre $\{b_{p+1}, \dots, b_n\}$ sono una base dell'autospazio relativo all'autovalore 0. In particolare, π_W è diagonalizzabile e la sua matrice rappresentativa rispetto a $\mathcal{B}$ è

$$\mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(\pi_W) = \text{diag}(\underbrace{1, \dots, 1}_{p\text{-volte}}, \underbrace{0, \dots, 0}_{(n-p)\text{-volte}}) = D$$

Quindi presa una base $\mathcal{C}$ di V , la formula del cambio di base ci dice

$$\mathcal{M}_{\mathcal{C}}^{\mathcal{C}}(\pi_W) = P^{-1}DP, \quad \text{dove } P = \mathcal{M}_{\mathcal{B}}^{\mathcal{C}}(id_V)$$

C'è un'altra formula che descrive $\mathcal{M}_{\mathcal{C}}^{\mathcal{C}}(\pi_W)$ e non richiede di conoscere una base ortogonale di V .

Teorema 5.24. *Sia $W \triangleleft V$ un sottospazio vettoriale dello spazio euclideo $(V, \langle \cdot, \cdot \rangle)$ e sia $\mathcal{B}_W = \{b_1, \dots, b_p\}$ una base ortonormale di W e sia $\mathcal{C}$ una base di V . Allora*

1. $A = \mathcal{M}_{\mathcal{C}}^{\mathcal{C}}(\pi_W) = \sum_{i=1}^p \underbrace{(b_i)_{\mathcal{C}} \cdot ((b_i)_{\mathcal{C}})^T}_{\in \mathcal{M}_{\mathbb{R}}(n)} \cdot S$, dove S è la matrice di Gram $S = \mathcal{M}_{\mathcal{C}}(\langle \cdot, \cdot \rangle)$.
2. $\mathcal{M}_{\mathcal{C}}^{\mathcal{C}}(\pi_{W^\perp}) = I_n - A$.

Dimostrazione. 2. segue dal fatto che $\pi_{W^\perp} = id_V - \pi_W$.

1. Dobbiamo trovare la matrice A tale che $A \cdot (v)_{\mathcal{C}} = (\pi_W(v))_{\mathcal{C}}$ (Proposizione 3.27). Calcoliamo dunque $(\pi_W(v))_{\mathcal{C}}$:

$$(\pi_W(v))_{\mathcal{C}} = \left(\sum_{k=1}^p \langle v, b_k \rangle b_k \right)_{\mathcal{C}} = \sum_{k=1}^p \langle b_k, v \rangle (b_k)_{\mathcal{C}} = \sum_{k=1}^p (b_k)_{\mathcal{C}} \cdot ((b_k)_{\mathcal{C}})^T \cdot S \cdot (v)_{\mathcal{C}} = \left(\sum_{k=1}^p (b_k)_{\mathcal{C}} \cdot ((b_k)_{\mathcal{C}})^T \cdot S \right) \cdot (v)_{\mathcal{C}}$$

dove, nelle ultime due uguaglianze abbiamo usato le proprietà del prodotto righe per colonne. $\square$

Corollario 5.25. *Se nel teorema precedente si prende $\mathcal{C}$ base ortonormale di V , allora $S = I_n$ e quindi $A = \mathcal{M}_{\mathcal{C}}^{\mathcal{C}}(\pi_W) = \sum_{i=1}^p (b_i)_{\mathcal{C}} \cdot ((b_i)_{\mathcal{C}})^T$.*

Esempio. • In $\mathbb{R}^4$ con il prodotto scalare standard (quindi la base canonica $\mathcal{E}$ è una base ortonormale) consideriamo $W = \text{Span} \left(w_1 := \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, w_2 := \begin{pmatrix} 1 \\ 2 \\ 0 \\ 0 \end{pmatrix} \right)$.

Usando Gram-Schmidt determiniamo una base ortonormale di W :

$$c_1 := w_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \quad c_2 := w_2 - \frac{\langle c_1, w_2 \rangle}{\|c_1\|^2} c_1 = \begin{pmatrix} 1 \\ 2 \\ -2 \\ 0 \end{pmatrix} - \frac{3}{3} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ -2 \\ -1 \end{pmatrix}$$

Quindi $\{d_1 := \frac{1}{\sqrt{3}}(1, 1, 0, 1)^T, d_2 := \frac{1}{\sqrt{6}}(0, 1, -2, -1)^T\}$ è una base ortonormale di W e

$$\begin{aligned}\mathcal{M}_{\mathcal{E}}^{\varepsilon}(\pi_W) &= \sum_{i=1}^p d_i^T \cdot d_i = \frac{1}{3} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 & 1 \end{pmatrix} + \frac{1}{6} \begin{pmatrix} 0 \\ 1 \\ -2 \\ -1 \end{pmatrix} \begin{pmatrix} 0 & 1 & -2 & -1 \end{pmatrix} \\ &= \frac{1}{3} \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 \end{pmatrix} + \frac{1}{6} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -2 & -1 \\ 0 & -2 & 4 & 2 \\ 0 & -1 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1/3 & 1/3 & 0 & 1/3 \\ 1/3 & 1/2 & -1/3 & 1/6 \\ 0 & -1/3 & 2/3 & 1/3 \\ 1/3 & 1/6 & 1/3 & 1/2 \end{pmatrix}\end{aligned}$$

Controllo:

$$\pi_W(x) = \frac{\langle x, c_1 \rangle}{\|c_1\|^2} c_1 + \frac{\langle x, c_2 \rangle}{\|c_2\|^2} c_2 = \frac{x_1 + x_2 + x_4}{3} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} + \frac{x_2 - 2x_3 - x_4}{6} \begin{pmatrix} 0 \\ 1 \\ -2 \\ -1 \end{pmatrix}$$

$$\text{Quindi, } \mathcal{M}_{\mathcal{E}}^{\varepsilon}(\pi_W) = \begin{pmatrix} \left| \pi_W(e_1) \right| & \left| \pi_W(e_2) \right| & \left| \pi_W(e_3) \right| & \left| \pi_W(e_4) \right| \\ \left| \pi_W(e_1) \right| & \left| \pi_W(e_2) \right| & \left| \pi_W(e_3) \right| & \left| \pi_W(e_4) \right| \\ \left| \pi_W(e_1) \right| & \left| \pi_W(e_2) \right| & \left| \pi_W(e_3) \right| & \left| \pi_W(e_4) \right| \\ \left| \pi_W(e_1) \right| & \left| \pi_W(e_2) \right| & \left| \pi_W(e_3) \right| & \left| \pi_W(e_4) \right| \end{pmatrix} = \begin{pmatrix} 1/3 & 1/3 & 0 & 1/3 \\ 1/3 & 1/2 & -1/3 & 1/6 \\ 0 & -1/3 & 2/3 & 1/3 \\ 1/3 & 1/6 & 1/3 & 1/2 \end{pmatrix}$$

Determiniamo ora la proiezione ortogonale di $v := (1, -1, 1, -1)^T$ su W e su $W^{\perp}$:

$$\pi_W(v) = \mathcal{M}_{\mathcal{E}}^{\varepsilon}(\pi_W) \cdot \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1/3 \\ -2/3 \\ 2/3 \\ 0 \end{pmatrix} \quad \text{e} \quad \pi_{W^{\perp}}(v) = v - \pi_W(v) = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} - \begin{pmatrix} -1/3 \\ -2/3 \\ 2/3 \\ 0 \end{pmatrix} = \begin{pmatrix} 4/3 \\ -1/3 \\ 1/3 \\ -1 \end{pmatrix}$$

• In $\mathbb{R}[t]_{\leq 1}$ con il prodotto scalare $\langle p(t), q(t) \rangle = p(0)q(0) + 3p(1)q(1)$, consideriamo la base $\mathcal{B} = \{t, t-1\}$ che è ortogonale, ma non ortonormale e $\mathcal{M}_{\mathcal{B}}(\langle \cdot, \cdot \rangle) = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}$. Inoltre, sia $W = \text{Span}(1-2t)$, allora $\|1-2t\|^2 = 4$ e $(1-2t)_{\mathcal{B}} = (-1, -1)^T$. Quindi

$$\mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(\pi_W) = \frac{1}{4} \begin{pmatrix} -1 \\ -1 \end{pmatrix} \cdot \begin{pmatrix} -1 & -1 \end{pmatrix} \cdot \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} = \frac{1}{4} \begin{pmatrix} 3 & 1 \\ 3 & 1 \end{pmatrix}$$

• In $\mathbb{R}^2$ con il prodotto scalare dato dalla matrice $\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ consideriamo $W = \text{Span}(\begin{pmatrix} 1 \\ -2 \end{pmatrix})$.

Allora $\|(\begin{pmatrix} 1 \\ -2 \end{pmatrix})\| = (1, -2)(\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix})(\begin{pmatrix} 1 \\ -2 \end{pmatrix}) = 5$, quindi $w := \frac{1}{\sqrt{5}}(\begin{pmatrix} 1 \\ -2 \end{pmatrix})$ forma una base ortonormale di W :

$$\mathcal{M}_{\mathcal{E}}^{\varepsilon}(\pi_W) = \frac{1}{5} \begin{pmatrix} 1 \\ -2 \end{pmatrix} \cdot \begin{pmatrix} 1 & -2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 1 & -2 \\ -2 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} -1 & -3 \\ 2 & 6 \end{pmatrix}$$

Determiniamo ora la proiezione ortogonale di $v := (1, 1)^T$ su W e su $W^{\perp}$:

$$\pi_W(v) = \mathcal{M}_{\mathcal{E}}^{\varepsilon}(\pi_W) \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} -4/5 \\ 8/5 \end{pmatrix} \quad \text{e} \quad \pi_{W^{\perp}}(v) = v - \pi_W(v) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} - \begin{pmatrix} -4/5 \\ 8/5 \end{pmatrix} = \begin{pmatrix} 9/5 \\ -3/5 \end{pmatrix}$$

5.4 Complementi: Prodotto vettoriale

Nello spazio euclideo $\mathbb{R}^3$ equipaggiato con il prodotto scalare standard è possibile definire un altro prodotto tra vettori, il cui risultato è un altro vettore: il *prodotto vettoriale* (o prodotto esterno)

Definizione 5.26. Siano $v, w \in \mathbb{R}^3$. Il loro *prodotto vettoriale* è il vettore

$$v \times w = \|v\| \cdot \|w\| \cdot \sin(\theta_{v,w}) \cdot n$$

dove $\theta_{v,w}$ è l'angolo tra v e w , e n è il versore ortogonale a v e w , e determinato dalla regola della mano destra (v : pollice, w : indice, n : medio). Si osservi che per definizione $\theta_{v,w} \in [0, \pi]$, e quindi $\sin \theta_{v,w} \geq 0$.

Sia $\mathcal{E} = \{e_1, e_2, e_3\}$ la base canonica di $\mathbb{R}^3$. C'è una formula esplicita per determinare il prodotto vettoriale come “determinante misto” (i.e. la matrice contiene sia scalari che vettori)

Siano $v = (v_1, v_2, v_3)^T, w = (w_1, w_2, w_3)^T \in \mathbb{R}^3$, allora

$$v \times w = \underline{\det} \begin{pmatrix} e_1 & v_1 & w_1 \\ e_2 & v_2 & w_2 \\ e_3 & v_3 & w_3 \end{pmatrix} = \left(\det \begin{pmatrix} v_2 & w_2 \\ v_3 & w_3 \end{pmatrix}, -\det \begin{pmatrix} v_1 & w_1 \\ v_3 & w_3 \end{pmatrix}, \det \begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \end{pmatrix} \right)^T \quad (5.6)$$

dove il primo “determinante” si espande *esclusivamente* lungo la prima colonna.

Esempio.

$$\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \times \begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix} = \det \begin{pmatrix} e_1 & 1 & 0 \\ e_2 & 1 & 2 \\ e_3 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 2 - 0 \\ -(2 - 0) \\ 2 - 0 \end{pmatrix} = \begin{pmatrix} 2 \\ -2 \\ 2 \end{pmatrix}$$

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \times \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} = \det \begin{pmatrix} e_1 & 1 & 4 \\ e_2 & 2 & 5 \\ e_3 & 3 & 6 \end{pmatrix} = \begin{pmatrix} 12 - 15 \\ -(6 - 12) \\ 5 - 8 \end{pmatrix} = \begin{pmatrix} -3 \\ 6 \\ -3 \end{pmatrix}$$

Proprietà. Siano $v, w \in \mathbb{R}^3$, allora

1. $v \times w$ è ortogonale sia a v che a w , e quindi all sottospazio generato da loro
2. $\|v \times w\| = \|v\| \cdot \|w\| \cdot \sin(\theta)$ è l'area del parallelogramma di vertici $0, v, w, v + w$.
3. $v \times w = 0$ se e solo se v, w sono linearmente dipendenti.
4. La mappa $\times : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^3$ è bilineare (= lineare in entrambe le componenti), e alternante, cioè $v \times w = -w \times v$.

Osservazione. i) Il prodotto vettoriale non è associativo; per esempio

$$e_1 \times (e_2 \times e_2) = e_1 \times 0 = 0, \text{ but } (e_1 \times e_2) \times e_2 = e_3 \times e_2 = -e_1.$$

ii) Siano $u = (u_1, u_2, u_3)^T, v = (v_1, v_2, v_3)^T, w = (w_1, w_2, w_3)^T \in \mathbb{R}^3$. Dalla definizione di prodotto scalare standard e prodotto vettoriale segue:

$$\langle u, v \times w \rangle = \det \begin{pmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{pmatrix}.$$

5.5 Sottospazi affini e distanze

Come osservato a pagina 83, un sottospazio affine S di uno spazio vettoriale V si ottiene prendendo un sottospazio vettoriale $W \triangleleft V$ e traslandolo di un vettore v : $S = v + W$. Ad esempio in $\mathbb{R}[t]_{\leq 3}$ prendiamo $S = t + \text{Span}(1, 1 + t^2 - t^3)$ e la sua giacitura è il sottospazio vettoriale associato: $S_0 = \text{Span}(1, 1 + t^2 - t^3)$.

Definizione 5.27. Siano $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo e $P, Q \in V$ due punti. Allora $d(P, Q) = \|P - Q\| = \|\overrightarrow{PQ}\|$.

Siano $S_1, S_2 \subset V$ sottospazi affini, allora

$$d(S_1, S_2) := \inf\{d(P, Q) \mid P \in S_1, Q \in S_2\} \stackrel{(*)}{=} \min\{d(P, Q) \mid P \in S_1, Q \in S_2\}.$$

In particolare, se S_1 e S_2 sono incidenti, allora $d(S_1, S_2) = 0$.

Osservazione. (*) La distanza tra due punti è una funzione a valori reali inferiormente limitata e in generale, non è detto che esista un minimo, ma sicuramente esiste il suo estremo inferiore. Però dimostreremo tra poco che il valore dell'estremo inferiore viene assunto dalla funzione distanza e quindi è effettivamente un minimo.

Proposizione 5.28. Siano $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo e $S_1 = P_1 + S_{1,0}$, $S_2 = P_2 + S_{2,0}$ due sottospazi affini. Allora

$$d(S_1, S_2) = \|\pi_{W^\perp}(P_2 - P_1)\|, \quad \text{dove} \quad W = S_{1,0} + S_{2,0}.$$

Dimostrazione. Per definizione, $d(S_1, S_2) = \inf\{d(P_1 + v_1, P_2 + v_2) : v_1 \in S_{1,0}, v_2 \in S_{2,0}\}$. Da $d(P_1 + v_1, P_2 + v_2) = \|P_2 - P_1 + v_2 - v_1\| = d(P_2 - P_1, v_2 - v_1)$ segue $d(S_1, S_2) = \inf\{d(P_2 - P_1, w) : w \in W\} = \min\{d(P_2 - P_1, w) : w \in W\} = \|\pi_{W^\perp}(P_2 - P_1)\|$, dove l'ultima uguaglianza è l'equazione (5.4). $\square$

Corollario 5.29. In $\mathbb{R}^n$ con il prodotto scalare standard, si consideri un iperpiano di equazione $\pi = \{a_1x_1 + \dots + a_nx_n = b\}$, e un punto $P = (y_1, \dots, y_n)^T \in \mathbb{R}^n$, allora

$$d(P, \pi) = \frac{|a_1y_1 + \dots + a_ny_n - b|}{\sqrt{a_1^2 + \dots + a_n^2}}$$

Dimostrazione. Sia ora $Q = (z_1, \dots, z_n)^T \in \pi$. La giacitura di P è il sottospazio banale, mentre $\pi_0 = \{a_1x_1 + \dots + a_nx_n = 0\}$, quindi dobbiamo considerare la proiezione ortogonale di $P - Q$ su $\pi_0^\perp = \text{Span}((a_1, \dots, a_n)^T)$:

$$\begin{aligned} d(P, \pi) &= \left\| \frac{\langle (y_1 - z_1, \dots, y_n - z_n)^T, (a_1, \dots, a_n)^T \rangle}{\|(a_1, \dots, a_n)^T\|^2} (a_1, \dots, a_n)^T \right\| \\ &= \frac{|a_1(y_1 - z_1) + \dots + a_n(y_n - z_n)|}{\|(a_1, \dots, a_n)^T\|^2} \|(a_1, \dots, a_n)^T\| = \frac{|a_1y_1 + \dots + a_ny_n - b|}{\|(a_1, \dots, a_n)^T\|} \quad \square \end{aligned}$$

Esempio. Sia $P = (1, 2, 3)^T \in \mathbb{R}^3$, e sia $\pi \subset \mathbb{R}^3$ il piano di equazione $x - 3y + z = 10$, e $r \subset \mathbb{R}^3$ la retta $r = \begin{cases} x - 3y + z = 4 \\ x - y = 1 \end{cases}$.

La distanza $d(P, \pi)$ è

$$d(P, \pi) = \frac{|1 \cdot 1 + (-3) \cdot 2 + 1 \cdot 3 - 10|}{\sqrt{1^2 + (-3)^2 + 1^2}} = \frac{12}{\sqrt{11}}.$$

La retta r ha equazione parametrica $r = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \mathbb{R}$. Calcoliamo la proiezione ortogonale di $P - \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}$ su $\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \mathbb{R}$, la giacitura di r : $\frac{\langle (1, 1, 2)^T, (1, 3, 2)^T \rangle}{\|(1, 1, 2)^T\|^2} \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} = \frac{8}{6} \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$ quindi la distanza $d(P, r)$ è

$$d(P, r) = \|\pi_{r^\perp} \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}\| = \|\begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} - \frac{4}{3} \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}\| = \|\begin{pmatrix} -1/3 \\ 5/3 \\ -2/3 \end{pmatrix}\| = \sqrt{\frac{1}{9} + \frac{25}{9} + \frac{4}{9}} = \sqrt{\frac{10}{3}}$$

Esempio. Vediamo ora un paio di esempi di distanza tra 2 rette in $\mathbb{R}^3$:

• Le rette $l_1 = \begin{cases} x - y = -1 \\ x - z = -2 \end{cases}$ e $l_2 = \begin{cases} -2x + y + z = 4 \\ y - z = -2 \end{cases}$ sono parallele: $l_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \mathbb{R}$, $l_2 = \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \mathbb{R}$, con giacitura $W = \text{Span} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. La loro distanza è

$$d(l_1, l_2) = \|\pi_{W^\perp} \begin{pmatrix} 0 - (-1) \\ 1 - 0 \\ 1 - 2 \end{pmatrix}\| = \|\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - \pi_W \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}\| = \|\begin{pmatrix} 1/3 \\ 1/3 \\ -2/3 \end{pmatrix}\| = \sqrt{\frac{2}{3}}$$

• Le rette $l_1 = \begin{cases} x - y = -1 \\ x - z = -2 \end{cases}$ e $l_3 = \begin{pmatrix} 0 \\ 4 \\ -3 \end{pmatrix} + \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \mathbb{R}$ sono sghembe, e le loro giaciture sono rispettivamente $\text{Span}((1, 1, 1)^T)$ e $\text{Span}((1, 2, 3)^T)$. Per calcolarne la distanza dobbiamo quindi trovare una base di $\text{Span}((1, 1, 1)^T, (1, 2, 3)^T)^\perp = \ker \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix} = \ker \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix} = \text{Span}((1, -2, 1)^T)$. Quindi

$$d(l_1, l_3) = \left\| \frac{\langle (1, -2, 1)^T, (0, -3, 5)^T \rangle}{\|(1, -2, 1)^T\|^2} \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \right\| = \frac{|\langle (1, -2, 1)^T, (0, -3, 5)^T \rangle|}{\|(1, -2, 1)^T\|} = \frac{11}{\sqrt{6}}$$

5.6 Endomorfismi di spazi euclidei

In questa sezione ci occupiamo di studiare gli endomorfismi di uno spazio euclideo.

5.6.1 Matrici ortogonali

Definizione 5.30. Una matrice $A \in \mathcal{M}_{\mathbb{R}}(n)$ è *simmetrica* se $A^T = A$.

Invece, A è *ortogonale* se $A^T = A^{-1}$ (in particolare, A è invertibile).

Esempio. Le seguenti matrici sono ortogonali:

$$I_n, \quad \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}, \quad \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

Osservazione. i) Se $A, B \in \mathcal{M}_{\mathbb{R}}(n)$ sono ortogonali, allora anche AB , A^{-1} e A^T sono ortogonali.

ii) Se A è ortogonale, allora $1 = \det(AA^T) = \det(A)^2$, quindi $\det(A) = \pm 1$.

Lemma 5.31. *Una matrice $A \in \mathcal{M}_{\mathbb{R}}(n)$ è ortogonale se e solo se le sue colonne (o righe) formano una base ortonormale di $\mathbb{R}^n$, rispetto al prodotto scalare standard.*

Dimostrazione. La condizione $A^T = A^{-1}$ si riscrive come $A^T A = I_n$, e le righe di A^T sono esattamente le colonne di A . Se c_k denota la k -esima colonna di A e $\langle \cdot, \cdot \rangle$ è il prodotto scalare standard di $\mathbb{R}^n$, allora vale che:

$$A^T \cdot A = I_n \iff (A^T \cdot A)_{i,j} = c_i^T \cdot c_j = \langle c_i, c_j \rangle = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases} \quad \square$$

Lemma 5.32. *Sia $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo e siano $\mathcal{B}$ e $\mathcal{C}$ due basi ortonormali, allora la matrice di cambio di base $A = \mathcal{M}_{\mathcal{C}}^{\mathcal{B}}(id_V)$ è una matrice ortogonale.*

Dimostrazione. Sia $\{b_1, \dots, b_n\}$ la base ortonormale $\mathcal{B}$, allora la k -esima colonna c_k di A è data dalle coordinate di b_k rispetto alla base $\mathcal{C}$: $c_k = (b_k)_{\mathcal{C}}$.

Visto che $\mathcal{C}$ è una base ortonormale, la matrice di Gram $\mathcal{M}_{\mathcal{C}}(\langle \cdot, \cdot \rangle)$ è l'identità e quindi

$$\langle b_i, b_j \rangle = ((b_i)_{\mathcal{C}})^T \cdot \mathcal{M}_{\mathcal{C}}(\langle \cdot, \cdot \rangle) \cdot (b_j)_{\mathcal{C}} = ((b_i)_{\mathcal{C}})^T \cdot (b_j)_{\mathcal{C}} = pss(c_i, c_j) = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases}$$

dove $pss(c_i, c_j)$ denota il prodotto scalare standard delle colonne c_i e c_j di A . $\square$

Le matrici ortogonali di ordine 2 hanno solo due possibili forme

Proposizione 5.33. *Sia $A \in \mathcal{M}_{\mathbb{R}}(2)$ una matrice ortogonale, allora esiste $\theta \in \mathbb{R}$ tale che*

$$A = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}, \quad \text{oppure} \quad A = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

Dimostrazione. Le colonne di A formano una base ortonormale di $\mathbb{R}^2$, rispetto al prodotto scalare standard.

Sia $(a, b)^T$ la prima colonna, allora $1 = \|(a, b)^T\| = a^2 + b^2$, esiste quindi $\theta \in \mathbb{R}$ tale che $a = \cos(\theta)$ e $b = \sin(\theta)$.

Visto che $(a, b)^T \perp (b, -a)^T$ e che la seconda colonna $(c, d)^T$ di A è un vettore ortogonale alla prima colonna, abbiamo che: $(c, d) = k(b, -a)$ per un qualche $k \in \mathbb{R}$, ma $1 = \|(c, d)^T\| = |k| \cdot \|(b, -a)\| = |k|$, cioè $(c, d)^T = \pm 1(\sin(\theta), -\cos(\theta))^T$. $\square$

Osservazione. La prima matrice l'abbiamo incontrata all'inizio del Capitolo 4 e rappresenta la *riflessione* rispetto alla retta di inclinazione $\theta/2$. La seconda matrice invece descrive la *rotazione* di angolo θ , infatti e_1 viene mandato in $(\cos(\theta), \sin(\theta))^T$ e e_2 viene mandato in $(-\sin(\theta), \cos(\theta))^T$.

Esempio. La matrice $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ descrive la riflessione rispetto alla retta di inclinazione $\pi/4$.

La matrice $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ descrive una rotazione di angolo $\pi/2$ (in senso antiorario).

5.6.2 Isometrie e trasformazioni affini

Abbiamo visto che le matrici ortogonali di $\mathcal{M}_{\mathbb{R}}(2)$ rappresentano rotazioni, e riflessioni, che sono trasformazioni che preservano lunghezze, distanze e angoli. Queste sono casi particolari di *isometrie lineari*: endomorfismi che preservano il prodotto scalare.

Definizione 5.34. Sia $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo. Un endomorfismo $f \in \text{End}(V)$ è una *isometria lineare* se $\langle f(v), f(w) \rangle = \langle v, w \rangle$ per ogni $v, w \in V$.

Le isometrie lineari sono trasformazioni “rigide” dello spazio, infatti soddisfano le seguenti

Proprietà. Sia $f \in \text{End}(V)$ un’isometria lineare. Allora

- f preserva le norme: $\|f(v)\| = \|v\|$ per ogni $v \in V$.
- f preserva gli angoli: $\theta_{f(v), f(w)} = \theta_{v, w}$ per ogni $v, w \in V$.
- f è iniettiva (e quindi biettiva)

Dimostrazione. Le prime due proprietà seguono dalla definizione di norma e di angolo. Per la terza, sia $v \in \ker(f)$, allora $\|v\| = \|f(v)\| = \|0_V\| = 0 \Rightarrow v = 0_V$ e si conclude usando il Teorema 3.13. $\square$

Proposizione 5.35. Siano $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo, $\mathcal{B}$ una base ortonormale di V e $f \in \text{End}(V)$.

Allora f è un’isometria lineare se e solo se la matrice rappresentativa $A := \mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(f)$ è ortogonale ($AA^T = I_n$).

Dimostrazione. Visto che $\mathcal{B}$ è ortonormale la matrice di Gram $\mathcal{M}_{\mathcal{B}}(\langle \cdot, \cdot \rangle)$ è l’identità e quindi

$$\langle f(v), f(w) \rangle = ((f(v))_{\mathcal{B}})^T \cdot (f(w))_{\mathcal{B}} = (A \cdot (v)_{\mathcal{B}})^T \cdot A \cdot (w)_{\mathcal{B}} = ((v)_{\mathcal{B}})^T \cdot (A^T A) \cdot (w)_{\mathcal{B}}$$

Analogamente $\langle v, w \rangle = ((v)_{\mathcal{B}})^T \cdot (w)_{\mathcal{B}}$.

Quindi f è un’isometria lineare se e solo se $x^T \cdot (A^T A) \cdot y = x^T \cdot y$ per ogni $x, y \in \mathbb{R}^n$.

In particolare, se A è ortogonale, allora f un’isometria lineare. Viceversa, se f è un’isometria lineare allora $(A^T A)_{i,j} = e_i^T \cdot A^T A \cdot e_j = e_i^T \cdot e_j = (I_n)_{i,j}$ per ogni i, j , cioè $A^T A = I_n$ ($n = \dim V$). $\square$

Osservazione. Posso riconoscere un’isometria lineare dalla matrice rappresentativa *solo* usando una base ortonormale!

In generale, per una base $\mathcal{B}$ qualsiasi vale: f è un’isometria lineare se e solo se $A^T B A = B$, dove $A := \mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(f)$ e $B = \mathcal{M}_{\mathcal{B}}(\langle \cdot, \cdot \rangle)$ è la matrice di Gram di $\langle \cdot, \cdot \rangle$.

Esistono altre trasformazioni rigide di uno spazio euclideo che non sono lineari, ad esempio le traslazioni: $T_{v_0} : V \rightarrow V, v \mapsto v + v_0$ (traslazione per $v_0 \in V$).

Definizione 5.36. Sia V uno spazio vettoriale. Una *trasformazione affine* è una funzione $T : V \rightarrow V$ tale che $T(v) = f(v) + v_0$, dove $v_0 \in V$ e $f \in \text{End}(V)$ è un isomorfismo.

Moralmente, “trasformazione affine = isomorfismo lineare + traslazione”.

Esempio. $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2, T\left(\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right) = \begin{pmatrix} 2 & -1 \\ 0 & 4 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 1 \\ -3 \end{pmatrix}$ è una trasformazione affine.

Osservazione. In generale una trasformazione affine “deforma” lo spazio, non preserva angoli e lunghezze; però manda sottospazi affini in sottospazi affini della stessa dimensione (rette in rette, piani in piani, etc.) e preserva le posizioni reciproche (incidenza, parallelismo, etc.)

Definizione 5.37. Sia $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo. Una trasformazione affine $T : V \rightarrow V$ è un *isometria* se $\|v - w\| = \|T(v) - T(w)\|$ per ogni $v, w \in V$.

Esempio. • Ogni traslazione è un’isometria.

- La mappa $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dell’esempio precedente non è un’isometria.
- $L : \mathbb{R}^2 \rightarrow \mathbb{R}^2, L\left(\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right) = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 2 \\ -3 \end{pmatrix}$ è un’isometria.

Teorema 5.38. Sia $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo. Una trasformazione affine $T : V \rightarrow V$ è un *isometria* se e solo se esiste un’isometria lineare $f \in \text{End}(V)$ e un $v_0 \in V$ tali che $T(v) = f(v) + v_0$ per ogni $v \in V$.

Moralmente, “isometria = isometria lineare + traslazione”, e preserva angoli, lunghezze, distanze, etc.

Dimostrazione. Se $f \in \text{End}(V)$ è un’isometria lineare, allora $\|v - w\| = \|f(v) - f(w)\| = \|f(v) - v_0 + v_0 - f(w)\| = \|T(v) - T(w)\|$ per ogni $v, w \in V$.

Viceversa, supponiamo che $T(v) = f(v) + v_0$, dove $v_0 \in V$ e $f \in \text{End}(V)$, sia un’isometria, cioè

$$\|v - w\| = \|T(v) - T(w)\| = \|f(v) - v_0 - (f(w) - v_0)\| = \|f(v) - f(w)\| \quad (5.7)$$

per ogni $v, w \in V$ e dobbiamo dimostrare che $\langle f(v), f(w) \rangle = \langle v, w \rangle$ per ogni $v, w \in V$.

Prendendo $w = 0_V$ in (5.7), abbiamo che $\|v\| = \|f(v)\|$ per ogni $v \in V$. Ora, per ogni $v, w \in V$ vale

$$\|f(v)\|^2 - 2\langle f(v), f(w) \rangle + \|f(w)\|^2 = \|f(v) - f(w)\|^2 = \|f(v - w)\|^2 = \|v - w\|^2 = \|v\|^2 - 2\langle v, w \rangle + \|w\|^2$$

da cui segue che $\langle f(v), f(w) \rangle = \langle v, w \rangle$ per ogni $v, w \in V$. $\square$

5.6.3 Endomorfismi simmetrici

Definizione 5.39. Siano $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo e $f \in \text{End}(V)$ un endomorfismo. Si dice che f è *simmetrico* se, per ogni $v, w \in V$: $\langle f(v), w \rangle = \langle v, f(w) \rangle$.

Esempio. La proiezione ortogonale su un sottospazio $W \triangleleft V$ è un endomorfismo simmetrico: sia $\{b_1, \dots, b_p\}$ una base ortonormale di W , allora $\pi_W(v) = \sum_{i=1}^p \langle v, b_i \rangle b_i$:

$$\langle \pi_W(v), w \rangle = \left\langle \sum_{i=1}^p \langle v, b_i \rangle b_i, w \right\rangle = \sum_{i=1}^p \langle v, b_i \rangle \langle b_i, w \rangle = \langle v, \left\langle \sum_{i=1}^p w, b_i \right\rangle b_i \rangle = \langle v, \pi_W(w) \rangle$$

Proposizione 5.40. Siano $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo, $\mathcal{B}$ una base ortonormale di V e $f \in \text{End}(V)$.

Allora f è simmetrico se e solo se la matrice rappresentativa $A := \mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(f)$ è simmetrica ($A = A^T$).

Dimostrazione. Visto che $\mathcal{B}$ è ortonormale la matrice di Gram $\mathcal{M}_{\mathcal{B}}(\langle \cdot, \cdot \rangle)$ è l'identità e quindi

$$\langle f(v), w \rangle = ((f(v))_{\mathcal{B}})^T \cdot (w)_{\mathcal{B}} = (A \cdot (v)_{\mathcal{B}})^T \cdot (w)_{\mathcal{B}} = ((v)_{\mathcal{B}})^T \cdot A^T \cdot (w)_{\mathcal{B}}$$

Analogamente

$$\langle v, f(w) \rangle = ((v)_{\mathcal{B}})^T \cdot (f(w))_{\mathcal{B}} = ((v)_{\mathcal{B}})^T \cdot A \cdot (w)_{\mathcal{B}}$$

Quindi f è simmetrico se e solo se $x^T \cdot A^T \cdot y = x^T \cdot A \cdot y$ per ogni $x, y \in \mathbb{R}^n$.

In particolare, se A è simmetrica, allora f è simmetrico. Viceversa, se f è simmetrico allora $(A^T)_{i,j} = e_i^T \cdot A^T \cdot e_j = e_i^T \cdot A \cdot e_j = (A)_{i,j}$ per ogni i, j , cioè $A = A^T$. $\square$

Osservazione. Posso riconoscere un endomorfismo simmetrico dalla matrice rappresentativa *solo* usando una base ortonormale!

In generale, per una base $\mathcal{B}$ qualsiasi vale: f è simmetrico se e solo se $A^T B = B A$, dove $A := \mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(f)$ e $B = \mathcal{M}_{\mathcal{B}}(\langle \cdot, \cdot \rangle)$ è la matrice di Gram di $\langle \cdot, \cdot \rangle$.

Esempio. La proiezione ortogonale su un sottospazio $W \triangleleft V$ è un endomorfismo simmetrico. Inoltre, se $\mathcal{B} = \{b_1, \dots, b_p, b_{p+1}, \dots, b_n\}$ è una base ortonormale di V tale che $\{b_1, \dots, b_p\}$ è una base (ortonormale) di W , abbiamo che

$$\mathcal{M}_{\mathcal{B}}^{\mathcal{B}}(\pi_W) = \text{diag}(\underbrace{1, \dots, 1}_{p\text{-volte}}, \underbrace{0, \dots, 0}_{(n-p)\text{-volte}}) = D$$

cioè $\mathcal{B}$ è una base ortogonale di V costituita da autovettori di π_W .

Definizione 5.41. Siano $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo e $f \in \text{End}(V)$ un endomorfismo. Si dice che f è *ortogonalmente diagonalizzabile* se esiste una base ortogonale di V costituita da autovettori di f .

Sia $\mathcal{B} = \{b_1, \dots, b_n\}$ una base ortogonale di V costituita da autovettori di f e si consideri la corrispondente base ortonormale $\mathcal{C} = \{c_1 = \frac{b_1}{\|b_1\|}, \dots, c_n = \frac{b_n}{\|b_n\|}\}$.

La base $\mathcal{C}$ è una base di autovettori per f , quindi $\mathcal{M}_{\mathcal{C}}^{\mathcal{C}}(f)$ è una matrice diagonale, in particolare è simmetrica. In altre parole ogni endomorfismo ortogonalmente diagonalizzabile è simmetrico.

Vale anche il viceversa!

Teorema 5.42 (Teorema spettrale). *Siano $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo e $f \in \text{End}(V)$ un endomorfismo.*

Allora f è simmetrico se e solo se f è ortogonalmente diagonalizzabile.

In termini matriciali, la definizione e il teorema spettrale si rileggono come segue.

Una matrice $A \in \mathcal{M}_{\mathbb{R}}(n)$ è *ortogonalmente diagonalizzabile* se esiste una base $\mathcal{B} = \{b_1, \dots, b_n\}$ ortogonale di $\mathbb{R}^n$ (rispetto al prodotto scalare standard) costituita da autovettori di A . In questo caso, la corrispondente base ortonormale $\mathcal{C}$ è anche una base di autovettori per A .

Dal primo teorema di diagonalizzazione A è simile alla matrice diagonale $D = \mathcal{M}_{\mathcal{C}}^{\mathcal{C}}(\mathcal{L}_A)$ tramite la matrice di cambio di base $P = \mathcal{M}_{\mathcal{C}}^{\mathcal{B}}(id_{\mathbb{R}^n})$: $P^{-1}AP = D$. Inoltre P è la matrice di cambio di base per due basi ortonormali, e dal Lemma 5.32 abbiamo che P è una matrice ortogonale.

Riformulando: $A \in \mathcal{M}_{\mathbb{R}}(n)$ ortogonalmente diagonalizzabile significa che esistono matrici $P, D \in \mathcal{M}_{\mathbb{R}}(n)$, P ortogonale, D invertibile tali che $D = P^{-1}AP = P^TAP$, ovvero $A = PDP^T$. Considerando la trasposta di entrambi i lati, otteniamo

$$A^T = (PDP^T)^T = (P^T)^T D^T P^T = PDP^T = A$$

quindi una matrice ortogonalmente diagonalizzabile è simmetrica.

E il Teorema spettrale ci dice che vale anche il viceversa: $A \in \mathcal{M}_{\mathbb{R}}(n)$ è ortogonalmente diagonalizzabile se e solo se A è simmetrica.

La dimostrazione del teorema spettrale esula dallo scopo di queste note, ma possiamo capire come costruire P e D partendo da una matrice simmetrica A .

L'idea è quella di utilizzare il processo di diagonalizzazione "classico" (vedi Sezione 4.4) e adattarlo.

Input Una matrice simmetrica $A \in \mathcal{M}_{\mathbb{R}}(n)$

Passo 1 Determinare il polinomio caratteristico $p_A(t) = \det(tI_n - A)$ di A e i suoi autovalori $\lambda_1, \dots, \lambda_r$.

Passo 2 Per ogni λ_i trovare una base $\mathcal{B}_{\lambda_i}$ dello spazio proprio V_{λ_i} .

Si noti che i controlli 1 e 2 non sono necessari, poiché A è diagonalizzabile dal teorema spettrale: quindi dobbiamo trovare n radici reali (contate con molteplicità) e ogni radice ha una molteplicità geometrica uguale a quella algebrica.

Quindi abbiamo ottenuto un **risultato intermedio**: $\mathcal{B} = \mathcal{B}_{\lambda_1} \cup \dots \cup \mathcal{B}_{\lambda_r}$ è una base di $\mathbb{R}^n$ composta da autovettori di A , ma in generale non è ortogonale. Tuttavia, l'ortogonalità tra autovettori di autospazi distinti è garantita dal seguente risultato.

Lemma 5.43. *Siano v, w autovettori della matrice simmetrica $A \in \mathcal{M}_{\mathbb{R}}(n)$ appartenenti ad autospazi distinti: $v \in V_{\lambda}$, $w \in V_{\mu}$ con $\lambda \neq \mu$. Allora $\langle v, w \rangle = 0$.*

Dimostrazione. Consideriamo $\langle v, Aw \rangle = \langle v, \mu w \rangle = \mu \langle v, w \rangle$. D'altra parte $\langle v, Aw \rangle = \langle Av, w \rangle = \langle \lambda v, w \rangle = \lambda \langle v, w \rangle$. Ne consegue che $(\lambda - \mu) \langle v, w \rangle = 0$, ma $\lambda \neq \mu$, quindi $\langle v, w \rangle = 0$. $\square$

Per garantire che gli autovettori di una base di uno stesso autospazio siano tra loro ortogonali, utilizziamo il processo di Gram-Schmidt:

Passo 3 Utilizzare il processo di Gram-Schmidt su ciascuna base $\mathcal{B}_{\lambda_i}$, per ottenere una base ortonormale $\mathcal{B}_{\lambda_i}^{on}$ dell'autospazio V_{λ_i} .

Output $\mathcal{B}^{on} = \mathcal{B}_{\lambda_1}^{on} \cup \dots \cup \mathcal{B}_{\lambda_r}^{on}$ è una base ortonormale di $\mathbb{R}^n$ costituita da autovettori di A . Prendendo P con colonne costituite dai vettori di $\mathcal{B}^{on}$, si ottiene che $P^{-1}AP = P^TAP$ è una matrice diagonale.

Esempio. a) La matrice $A = \begin{pmatrix} 1 & 3 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \in \mathcal{M}_{\mathbb{R}}(3)$ è simmetrica, quindi ortogonalmente diagonalizzabile. Determiniamo una base ortogonale di $\mathbb{R}^3$ formata da autovettori di A .

Il polinomio caratteristico di A è

$$p_A(t) = \det \begin{pmatrix} t-1 & -3 & 0 \\ -3 & t-1 & 0 \\ 0 & 0 & t+2 \end{pmatrix} = (t+2)((t-1)^2-9) = (t+2)(t^2-2t-8) = (t+2)^2(t-4)$$

Determiniamo ora una base per entrambi gli autospazi:

$$V_4 = \ker \begin{pmatrix} 3 & -3 & 0 \\ -3 & 3 & 0 \\ 0 & 0 & 6 \end{pmatrix} = \ker \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \text{Span} \left(\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right)$$

$$V_{-2} = \ker \begin{pmatrix} -3 & -3 & 0 \\ -3 & -3 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \ker \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \text{Span} \left(\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right)$$

Poiché $\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ sono ortogonali, per ottenere una base ortonormale basta normalizzare i 3 vettori:

$$\left\{ \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

è una base ortonormale di $\mathbb{R}^3$ costituita da autovettori di A e

$$P^T A P = \begin{pmatrix} 4 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{pmatrix} \quad \text{dove} \quad P = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

b) La matrice $A = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix} \in \mathcal{M}_{\mathbb{R}}(3)$ è simmetrica, quindi è ortogonalmente

diagonalizzabile. Il suo polinomio caratteristico è

$$p_A(t) = \det \begin{pmatrix} t-2 & 1 & 1 \\ 1 & t-2 & 1 \\ 1 & 1 & t-2 \end{pmatrix} = (t+2)((t-2)^2-1)-(t-2-1)+(1-(t-2)) = t^3-6t^2+9t = t(t-3)^2$$

Determiniamo ora una base per entrambi gli autospazi:

$$V_0 = \ker \begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix} = \ker \begin{pmatrix} 1 & 1 & -2 \\ 0 & -3 & 3 \\ 0 & 3 & -3 \end{pmatrix} = \ker \begin{pmatrix} 1 & 1 & -2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix} = \text{Span} \left(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right)$$

$$V_3 = \ker \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} = \ker \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \text{Span} \left(\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right)$$

Poiché $\langle \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \rangle = 1$, questi due vettori non sono ortogonali, quindi applichiamo il processo di Gram-Schmidt per ottenere una base ortogonale di V_3 :

$$c_1 := \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \quad c_2 := \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} - \frac{\langle (1, 0, -1)^T, (1, -1, 0)^T \rangle}{\|(1, -1, 0)^T\|^2} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ -1 \end{pmatrix}$$

Dopo la normalizzazione, otteniamo una base ortonormale di $\mathbb{R}^3$ costituita da autovettori di A :

$$\left\{ \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}, \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ \frac{\sqrt{2}}{\sqrt{3}} \end{pmatrix} \right\}$$

e inoltre vale

$$P^T A P = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix} \quad \text{dove} \quad P = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{\sqrt{2}}{\sqrt{3}} \end{pmatrix}$$

Teorema 5.44 (Decomposizione spettrale). *Siano $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo e $f \in \text{End}(V)$ un endomorfismo simmetrico, e siano $\{\lambda_1, \dots, \lambda_k\}$ i suoi autovalori. Inoltre sia $\pi_j : V \rightarrow V$ la proiezione ortogonale sull'autospazio V_{λ_j} relativo a λ_j .*

Allora, per ogni v in V vale

$$f(v) = \sum_{j=1}^k \lambda_j \pi_j(v), \quad \text{cioè} \quad \mathcal{M}_B^B(f) = \sum_{j=1}^k \lambda_j \mathcal{M}_B^B(\pi_j)$$

Esempio. Consideriamo l'esempio precedente, parte a). $A = \begin{pmatrix} 1 & 3 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \in \mathcal{M}_{\mathbb{R}}(3)$ è ortogonalmente diagonalizzabile, e gli autospazi sono

$$V_4 = \text{Span} \left(b_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right), \quad V_{-2} = \text{Span} \left(b_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, b_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right)$$

e $\{b_1, b_2, b_3\}$ è una base ortogonale di autovettori. Allora

$$\begin{aligned} A &= \mathcal{M}_{\mathcal{E}}^{\mathcal{E}}(\mathcal{L}_A) = 4\mathcal{M}_{\mathcal{E}}^{\mathcal{E}}(\pi_4) - 2\mathcal{M}_{\mathcal{E}}^{\mathcal{E}}(\pi_{-2}) \\ &= 4\frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \cdot (1 \ 1 \ 0) - 2 \left(\frac{1}{2} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} (1 \ -1 \ 0) + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} (0 \ 0 \ 1) \right) \\ &= 4 \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 0 \end{pmatrix} - 2 \begin{pmatrix} 1/2 & -1/2 & 0 \\ -1/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

5.7 Forme quadratiche

Sospendiamo ora la discussione sui prodotti scalari e torniamo a considerare forme bilineari simmetriche. In quest'ultima sezione introduciamo le *forme quadratiche* e vediamo come classificarle. Questo ci servirà nel prossimo capitolo per studiare e classificare coniche e quadriche.

Definizione 5.45. Sia V uno spazio vettoriale reale e sia $f : V \times V \rightarrow \mathbb{R}$ una forma bilineare simmetrica.

La *forma quadratica* associata a $f \in \text{Bil}(V)$ è

$$q_f : V \rightarrow \mathbb{R}, \quad q(v) := f(v, v).$$

Esempio. • Su $V = \mathbb{R}^2$, sia $f \left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \right) = x_1 x_2 - y_1 y_2$, allora la forma quadratica associata è $q_f \left(\begin{pmatrix} x \\ y \end{pmatrix} \right) = x^2 - y^2$.

• Sia $f \in \text{Bil}(\mathbb{R}^2)$, data dalla matrice $\begin{pmatrix} a & b \\ b & c \end{pmatrix}$ allora la forma quadratica associata è $q_f \left(\begin{pmatrix} x \\ y \end{pmatrix} \right) = ax^2 + 2bxy + cy^2$.

• Su V consideriamo la forma bilineare nulla: $n : V \times V \rightarrow \mathbb{R}, n(v, w) = 0$, allora la forma quadratica associata è la *forma quadratica nulla*: $q_n : V \rightarrow \mathbb{R}, q_n(v) = 0$.

• Sia $V = \mathbb{R}[t]_{\leq 1}$ e sia $h : V \times V \rightarrow \mathbb{R}, h(p(t), q(t)) = p(0)q(0) + 3p(1)q(1)$, allora la forma quadratica associata è $q_h(p(t)) = p(0)^2 + 3p(1)^2$.

• Dato un prodotto scalare, la forma quadratica associata è il quadrato della norma.

Proprietà. Siano V uno spazio vettoriale reale, $f : V \times V \rightarrow \mathbb{R}$ una forma bilineare simmetrica e $q_f : V \rightarrow \mathbb{R}$ la forma quadratica, associata, allora:

- $q_f(0_V) = 0$;
- $q_f(\lambda v) = \lambda^2 q_f(v)$.
- $f(v, w) = \frac{q_f(v+w) - q_f(v) - q_f(w)}{2}$ (*Formula di polarizzazione*)

In particolare, l'ultima proprietà ci dice che data una forma quadratica $q : V \rightarrow \mathbb{R}$, possiamo ricavare la forma bilineare simmetrica associata (cf. Proposizione 5.10).

Osservazione. Se la forma quadratica $q_f : V \rightarrow \mathbb{R}$ è definita su $V = \mathbb{R}^n$, allora assume la forma

$$q_f \left(\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \right) = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \end{pmatrix} \cdot A \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i=1}^n a_{ii} x_i^2 + 2 \sum_{i < j} a_{ij} x_i x_j$$

dove $A = \mathcal{M}_{\mathcal{E}}(f)$. In particolare q_f è data tramite un polinomio *omogeneo di grado 2* (cioè privo di monomi di grado 1 o 0), oppure dal polinomio nullo.

Quindi, data una forma quadratica $q : V \rightarrow \mathbb{R}$ e fissata una base $\mathcal{B}$ di V , posso rappresentare q tramite un polinomio *omogeneo* di grado 2 o tramite il polinomio nullo (se q è la forma quadratica nulla).

Esempio. • In $\mathbb{R}^1$, $q(x) = 2x^2$ è una forma quadratica e la matrice associata è $A = (2)$.

• In $\mathbb{R}^3$, $q((x_1, x_2, x_3)^T) = x_1^2 + x_2^2 - 3x_3^2 - 4x_1 x_2 + 8x_1 x_3 + 2x_2 x_3$ è una forma quadratica e la matrice associata è:

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

• In $\mathbb{R}^2$, $q((x_1, x_2)^T) = x_1^2 - 3x_2^2 + 4x_1$ non è una forma quadratica.

5.7.1 Segno di una forma quadratica

Definizione 5.46 (Segno di una forma quadratica). Siano V uno spazio vettoriale, $f \in \text{Bil}(V)$ simmetrica e sia q_f la forma quadratica associata a f .

Allora q_f (e anche f) si dice

- *definita positiva*, se $q_f(v) > 0$ per ogni $v \in V$, $v \neq 0_V$ (cioè f è un prodotto scalare);
- *definita negativa*, se $q_f(v) < 0$ per ogni $v \in V$, $v \neq 0_V$;
- *semi-definita positiva*, se $q_f(v) \geq 0$ per ogni $v \in V$;
- *semi-definita negativa*, se $q_f(v) \leq 0$ per ogni $v \in V$;
- *indefinita* altrimenti, cioè se esistono $v_1, v_2 \in V$ tali che $q_f(v_1) < 0 < q_f(v_2)$.

Esempio. • Ogni prodotto scalare definisce una forma quadratica definita positiva.

- in $\mathbb{R}^2$, $q(x_1, x_2) = x_1^2 + x_2^2$ è definita positiva.
- in $\mathbb{R}^2$, $q(x_1, x_2) = -x_1^2 - 3x_2^2$ è definita negativa.
- in $\mathbb{R}^2$, $q(x_1, x_2) = 2x_1^2$ è semi-definita positiva, ma non definita positiva ($q(0, 1) = 0$).
- in $\mathbb{R}^2$, $q(x_1, x_2) = x_1^2 - 4x_2^2$ è indefinita ($q(1, 0) = 1$, $q(0, 1) = -4$).

5.7.2 Segnatura e classificazione delle forme quadratiche

Se $\mathcal{C}$ è una base di V , definiamo la matrice $\mathcal{M}_{\mathcal{C}}(q_f)$ di una forma quadratica come la matrice di Gram della sua forma bilineare simmetrica f associata: $\mathcal{M}_{\mathcal{C}}(q_f) := \mathcal{M}_{\mathcal{C}}(f)$.

Nella Sezione 5.1 abbiamo visto che se cambiamo coordinate e consideriamo la base $\mathcal{B}$, allora $A := \mathcal{M}_{\mathcal{C}}(f)$ e $B := \mathcal{M}_{\mathcal{B}}(f)$ sono congruenti, cioè

$$A = P^T \cdot B \cdot P$$

dove P è una matrice invertibile, ovvero $P = \mathcal{M}_{\mathcal{B}}^{\mathcal{C}}(id_V)$.

Il nostro obiettivo è trovare una base $\mathcal{B}$ di V tale che la matrice $\mathcal{M}_{\mathcal{B}}(q_f)$ sia il più semplice possibile.

Definizione 5.47. Dati interi non negativi $r, s, t \in \mathbb{Z}_{\geq 0}$ definiamo $I_{r,s,t} \in \mathcal{M}_{\mathbb{R}}(r+s+t)$ come la matrice diagonale:

$$I_{r,s,t} := \text{diag}(\underbrace{1, \dots, 1}_r, \underbrace{-1, \dots, -1}_s, \underbrace{0, \dots, 0}_t)$$

Teorema 5.48. Sia V uno spazio vettoriale e sia $f \in \text{Bil}(V)$ simmetrica. Allora esiste una base $\mathcal{B} = \{b_1, \dots, b_n\}$ di V in cui

$$f(v_i, v_j) = 0 \text{ se } i \neq j \quad e \quad f(v_i, v_i) \in \{-1, 0, 1\}.$$

Cioè $\mathcal{M}_{\mathcal{B}}(f) = I_{r,s,n-r+s}$ per qualche r, s . In altre parole, ogni forma bilineare simmetrica è diagonalizzabile.

Dimostrazione. Sia $\mathcal{C}$ una base di V e consideriamo la matrice di Gram $A = \mathcal{M}_{\mathcal{C}}(f)$. Poichè A è simmetrica (Proposizione 5.7), per il teorema spettrale in forma matriciale, esiste una matrice ortogonale $P \in \mathcal{M}_{\mathbb{R}}(n)$ tale che

$$P^{-1}AP = P^TAP = D = \text{diag}(\lambda_1, \dots, \lambda_p, \underbrace{0, \dots, 0}_{n-p})$$

dove $\lambda_1, \dots, \lambda_p$ sono gli autovalori non nulli di A : $p = \text{rg}(A)$ e $n - p = \dim \ker A$. Equivalentemente, esiste una base $\mathcal{D} = \{w_1, \dots, w_n\}$ tale che $P = \mathcal{M}_{\mathcal{C}}^{\mathcal{D}}(id_V)$ e $D = \mathcal{M}_{\mathcal{D}}(f)$, cioè:

$$\begin{aligned} f(w_i, w_i) &= \lambda_i & \text{se } i &= 1, \dots, p \\ f(w_i, w_i) &= 0 & \text{se } i &\geq p+1 \\ f(w_i, w_j) &= 0 & \text{se } j &\neq i \end{aligned}$$

Ciascun $\lambda_i = f(w_i, w_i)$ è o positivo o negativo. A meno di riordinare gli elementi della base $\mathcal{D}$, possiamo supporre che

$$\begin{aligned} f(w_i, w_i) &= \lambda_i > 0 & \text{se } i &= 1, \dots, r \\ f(w_i, w_i) &= \lambda_i < 0 & \text{se } i &= r+1, \dots, p \end{aligned}$$

Se ora “normalizziamo” i vettori w_i , cioè consideriamo $\mathcal{B} = \{v_1, \dots, v_n\}$ dove $v_i := \frac{w_i}{\sqrt{|\lambda_i|}}$ per $i = 1, \dots, p$ e $v_j = w_j$ per $j = p+1, \dots, n$, otteniamo una base tale che

$$\mathcal{M}_{\mathcal{B}}(f) = \text{diag}(\underbrace{1, \dots, 1}_r, \underbrace{-1, \dots, -1}_s, \underbrace{0, \dots, 0}_{n-(r+s)}) = I_{r,s,n-(r+s)} \quad \square$$

Esempio. Prendiamo la forma quadratica $q_f : \mathbb{R}^4 \rightarrow \mathbb{R}$, $q_f((x_1, x_2, x_3, x_4)^T) = 4x_1x_2 + x_3^2$, allora la matrice di Gram rispetto alla base canonica della forma bilineare associata è:

$$\mathcal{M}_{\mathcal{E}}(f) = \begin{pmatrix} 0 & 2 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Il suo polinomio caratteristico è $p(t) = t(t-1)(t^2-4)$, che ha 4 radici distinte, quindi $\mathcal{M}_{\mathcal{E}}(f)$ è diagonalizzabile, ed è simile a

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

quindi esiste una base $\mathcal{B}$ di $\mathbb{R}^4$ tale che

$$\mathcal{M}_{\mathcal{B}}(f) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = I_{2,1,1}$$

e in questa base, la forma quadratica associata è $q_f((y_1, y_2, y_3, y_4)^T) = y_1^2 + y_2^2 - y_3^2$,

Osservazione. i) La dimostrazione del teorema ci dà un metodo esplicito per trovare la base $\mathcal{B}$ e la matrice diagonale $\mathcal{M}_{\mathcal{B}}(f)$.

ii) Il numero r di “1” che compiono in $\mathcal{M}_{\mathcal{B}}(f)$ coincide con il numero di autovalori positivi di A (contati con molteplicità). Analogamente, il numero s di “-1” che compiono in $\mathcal{M}_{\mathcal{B}}(f)$ coincide con il numero di autovalori negativi di A , mentre $n - (r + s) = \dim \ker A$.

Definizione 5.49. Siano V uno spazio vettoriale e $f \in \text{Bil}(V)$ simmetrica. Data una base $\mathcal{B}$ di V , il numero di autovalori positivi/negativi/nulli (contati con molteplicità) di $\mathcal{M}_{\mathcal{B}}(f)$ si chiama *indice di positività/negatività/nullità*.

Questi 3 valori sono ben definiti, cioè *indipendenti* dalla scelta della base $\mathcal{B}$, come vedremo tra poco (Teorema 5.51).

Definizione 5.50. Siano V uno spazio vettoriale e $q_f : V \rightarrow \mathbb{R}$ una forma quadratica. La *segnatura* di q_f è la terna (r, s, t) , dove r è l'indice di positività di f , s , è l'indice di negatività di f , e $t = \dim V - (r + s)$ è il suo indice di nullità.

Cioè la segnatura di q_f è (r, s, t) se e solo se $\mathcal{M}_{\mathcal{B}}(f)$ è congruente a $I_{r,s,t}$, e la matrice $I_{r,s,n-(r+s)}$ è la *forma canonica di Sylvester* di q_f .

Esempio. La segnatura di $q_f : \mathbb{R}^4 \rightarrow \mathbb{R}$, $q_f((x_1, x_2, x_3, x_4)^T) = 4x_1x_2 + x_3^2$ è $(2, 1, 1)$.

Teorema 5.51 (legge di inerzia (di Sylvester)). *Siano V uno spazio vettoriale di dimensione $\dim V = n$, $f \in \text{Bil}(V)$ simmetrica, e $\mathcal{B}$ e $\mathcal{C}$ basi di V tali che $\mathcal{M}_{\mathcal{B}}(f) = I_{r,s,n-(r+s)}$ e $\mathcal{M}_{\mathcal{C}}(f) = I_{r',s',n-(r'+s')}$. Allora $r = r'$ e $s = s'$.*

Dimostrazione. Visto che $\mathcal{M}_{\mathcal{B}}(f)$ e $\mathcal{M}_{\mathcal{C}}(f)$ sono congruenti hanno lo stesso rango, e quindi $r + s = r' + s'$.

Siano $\mathcal{B} = \{b_1, \dots, b_n\}$ e $\mathcal{C} = \{c_1, \dots, c_n\}$, allora $\{b_1, \dots, b_r, c_{r'+1}, \dots, c_n\}$ sono linearmente indipendenti. Infatti, supponiamo di avere il vettore nullo come combinazione lineare dei vettori citati:

$$0_V = \sum_{i=1}^r x_i b_i + \sum_{j=r'+1}^n y_j c_j$$

e poniamo $v := \sum_{i=1}^r x_i b_i$ ottenendo di conseguenza $-v = \sum_{j=r'+1}^n y_j c_j$. Ora

$$f(v, v) = \sum_{i,k=1}^r x_i x_k f(b_i, b_k) = \sum_{i,k=1}^r x_i^2 \geq 0$$

e l'uguaglianza si verifica solamente se $x_i = 0$ per ogni $i = 1, \dots, r$; inoltre

$$f(-v, -v) = \sum_{j,k=r'+1}^n y_j y_k f(c_j, c_k) = \sum_{j,k=r'+1}^{r'+s'} -y_j^2 \leq 0$$

Siccome $f(-v, -v) = (-1)^2 f(v, v) = f(v, v)$, confrontando le due equazioni precedenti, si ottiene che entrambe le disuguaglianze sono delle uguaglianze. In particolare $x_i = 0$ per ogni $i = 1, \dots, r$ e dunque $v = 0_V$. Di conseguenza anche $-v = 0_V$ ed essendo $\{c_{r'+1}, \dots, c_n\}$ linearmente indipendenti otteniamo che anche $y_j = 0$ per ogni $j = r' + 1, \dots, n$.

Quindi gli $r + n - r'$ vettori $\{b_1, \dots, b_r, c_{r'+1}, \dots, c_n\}$ sono linearmente indipendenti e dunque $r + n - r' \leq n \Leftrightarrow r \leq r'$.

Analogamente, $\{c_1, \dots, c_{r'}, b_{r+1}, \dots, b_n\}$ sono linearmente indipendenti e quindi $r' \leq r$. Da cui otteniamo $r = r'$ e $s = s'$. $\square$

Corollario 5.52. *Due matrici simmetriche $A, B \in \mathcal{M}_{\mathbb{R}}(n)$ sono congruenti se e solo se hanno la stessa segnatura.*

Classificare una forma quadratica significa determinare la sua forma canonica di Sylvester.

Osservazione. i) Ogni forma quadratica su V può essere rappresentata come

$$q_f \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i=1}^r x_i^2 - \sum_{j=r+1}^{r+s} x_j^2,$$

dove r è l'indice di positività, s quello di negatività e $n - (r + s)$ è l'indice di nullità.

ii) La segnatura di una forma quadratica su V con $\dim V = n$ ci dice anche il suo segno:

Segnatura	Segno
$(n, 0, 0)$	definita positiva
$(0, n, 0)$	definita negativa
$(*, 0, *)$	semi-definita positiva
$(0, *, *)$	semi-definita negativa
$(*, *, 0)$	indefinita
$(*, *, *)$	

dove $*$ indica un numero naturale strettamente positivo.

Abbiamo quindi la *classificazione delle forme quadratiche di $\mathbb{R}^2$* :

Proposizione 5.53. *Le forme quadratiche di $\mathbb{R}^2$ sono, a meno di congruenza, le seguenti:*

$$x^2 + y^2, \quad x^2 - y^2, \quad -x^2 - y^2, \quad x^2, \quad -x^2, \quad 0;$$

di signature $(2, 0, 0)$, $(1, 1, 0)$, $(0, 2, 0)$, $(1, 0, 1)$, $(0, 1, 1)$, e $(0, 0, 2)$.

Le forme quadratiche di $\mathbb{R}^2$ sono della forma

$$q_f \left(\begin{pmatrix} x \\ y \end{pmatrix} \right) = (x, y) \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = ax^2 + 2bxy + cy^2$$

E la segnatura è data dai segni degli autovalori λ_1, λ_2 di A , radici del polinomio caratteristico $\det(\lambda I_n - A) = \lambda^2 - \text{tr}(A)\lambda + \det A$, dove $\text{tr}(A) = a + c$ e $\det A = ac - b^2$.

Poiché $\lambda_1 \lambda_2 = \det A$ e $\lambda_1 + \lambda_2 = \text{tr}(A)$, si ottiene:

- se $\det A > 0$ e $\text{tr}(A) > 0$, allora q è definita positiva;
- se $\det A > 0$ e $\text{tr}(A) < 0$, allora q è definita negativa;
- se $\det A = 0$ e $\text{tr}(A) > 0$, allora q è semi-definita positiva;
- se $\det A = 0$ e $\text{tr}(A) < 0$, allora q è semi-definita negativa;
- se $\det A = 0$ e $\text{tr}(A) = 0$, allora q è la forma quadratica nulla;
- se $\det A < 0$, allora q è indefinita.

Attenzione! In dimensione maggiore non è sufficiente considerare la positività/negatività di traccia e determinante per stabilire il segno di una forma quadratica. Ad esempio, le matrici

$$\begin{pmatrix} 1 & & \\ & 3 & \\ & & 1 \end{pmatrix}, \quad \text{e} \quad \begin{pmatrix} -1 & & \\ & 3 & \\ & & -1 \end{pmatrix}$$

hanno entrambe determinante e traccia positivi, ma la prima è definita positiva, mentre la seconda è indefinita.

5.7.3 Criterio di Cartesio

Per stabilire la segnatura e il segno di una forma quadratica dobbiamo quindi stabilire il segno degli autovalori di una sua matrice rappresentativa, cioè il segno delle radici del suo polinomio caratteristico.

Per fare questo può essere utile il criterio di Cartesio, che ricordiamo brevemente.

Teorema 5.54 (Criterio di Cartesio). *Sia $p(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_d x^d \in \mathbb{R}[x]$ un polinomio reale di grado n tale che tutte le sue radici siano reali. Allora*

1. *0 è radice di $p(x)$ se e solo se $d \leq 1$ e $\text{ma}(0) = d$.*
2. *Dato l'insieme ordinato dei coefficienti $v := (a_n, a_{n-1}, \dots, a_d)$, allora $p(x)$ ha tante radici reali positive (contate con molteplicità) quante sono le variazioni di segno degli elementi contigui dell'insieme v (gli eventuali zeri non si contano).*

In particolare, se una matrice è simmetrica, allora è diagonalizzabile e quindi tutti i suoi autovalori sono reali e possiamo applicare il criterio di Cartesio.

Esempio. • Se $p_B(x) = x^4 - 2x^3 - 11x^2 + 28x - 16$, allora l'insieme ordinato dei coefficienti è $v := (1, -2, -11, 28, -16)$ e ci sono 3 variazioni di segno $1 \rightarrow -2, -11 \rightarrow 28 \rightarrow -16$: la segnatura è $(3, 1, 0)$.

• Se $p_B(x) = x^4 - 2x^3 - 12x^2 + 16$, allora l'insieme ordinato dei coefficienti è $(1, -2, -12, 0, 16)$ che si riduce a $v := (1, -2, -12, 16)$ rimuovendo lo 0. Ci sono 2 variazioni di segno, quindi la segnatura è $(2, 2, 0)$.

• Sia $A = \begin{pmatrix} 1 & k & 1 \\ k & -1 & 2 \\ 1 & 2 & 0 \end{pmatrix}$, $k \in \mathbb{R}$. Determinare per quali $k \in \mathbb{R}$ la matrice A determina

un prodotto scalare (cioè per quali $k \in \mathbb{R}$ la forma quadratica corrispondente è definita positiva).

Il polinomio caratteristico di A è $p_A(x) = \det(xI_3 - A) = x^3 - (k^2 + 6)x + 4k - 3$. L'insieme ordinato dei coefficienti è $(1, -(k^2 + 6), 4k - 3)$. Visto che $1 > 0$ e $-(k^2 + 6) < 0$, applicando Cartesio otteniamo che la segnatura di A è: i) $(1, 2, 0)$ se $4k - 3 < 0$; ii) $(2, 1, 0)$ se $4k - 3 > 0$; iii) $(1, 1, 1)$ se $4k - 3 = 0$.

5.7.4 Complementi: metodo dei minori di nord-ovest

Vediamo ora un metodo per stabilire se una forma quadratica è definita positiva o negativa, calcolando solo i determinanti di opportune sottomatrici.

Definizione 5.55. Sia $A \in \mathcal{M}_{\mathbb{R}}(n)$ e definiamo l' s -esimo minore di nord-ovest di A come la matrice ${}^sA \in \mathcal{M}_{\mathbb{R}}(s)$ ottenuta da A rimuovendo le ultime $n - s$ righe e le ultime $n - s$ colonne

Proposizione 5.56 (Metodo dei minori di nord-ovest). Sia $A \in \mathcal{M}_{\mathbb{R}}(n)$ una matrice rappresentativa della forma quadratica $q : V \rightarrow \mathbb{R}$, allora

1. q è definita positiva se e solo se $\det({}^sA) > 0$ per ogni s ;
2. q è definita negativa se e solo se $(-1)^s \det({}^sA) > 0$ per ogni s .

Esempio. • La matrice $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$ descrive un prodotto scalare, infatti $\det({}^1A) = \det(1) = 1 > 0$, $\det({}^2A) = \det \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = 1 > 0$ e anche $\det({}^3A) = \det A = 3 > 0$.

• Sia $A = \begin{pmatrix} 1 & k & 1 \\ k & -1 & 2 \\ 1 & 2 & 0 \end{pmatrix}$, $k \in \mathbb{R}$. Allora $\det({}^1A) = \det(1) = 1 > 0$ e $\det({}^2A) = \det \begin{pmatrix} 1 & k \\ k & -1 \end{pmatrix} = -(1 + k^2) < 0$, quindi A non è né definita positiva, né definita negativa.

Capitolo 6

Quadriche e loro classificazione

In questo capitolo studiamo le (ipersuperfici) quadriche in $\mathbb{R}^n$, cioè le ipersuperfici descritte da un polinomio di grado 2 in n variabili.

In particolare vedremo come classificarle metricamente ed affinemente e come determinare le loro simmetrie.

6.1 Ipersuperfici quadriche

Definizione 6.1. Una (ipersuperficie) quadrica di $\mathbb{R}^n$ è l'insieme dei punti di $\mathbb{R}^n$ che soddisfano un'equazione polinomiale di grado 2:

$$Q := \left\{ (x_1, \dots, x_n) \in \mathbb{R}^n \mid q(x) = \sum_{i,j=1}^n a_{ij} x_i x_j + 2 \sum_{i=1}^n b_i x_i + c = 0 \right\}$$

Quindi abbiamo una forma quadratica descritta dalla matrice $A \in \mathcal{M}_{\mathbb{R}}(n)$ con $(A_{i,j}) = a_{i,j}$ (in particolare $A = A^T$), una forma lineare descritta dal vettore $b = (b_1, \dots, b_n)^T \in \mathbb{R}^n$ e il termine costante $c \in \mathbb{R}$.

Notazione. Se $n = 2$ una (ipersuperficie) quadrica è detta *conica*.

Esempio. • In $\mathbb{R}^2$, le seguenti equazioni descrivono delle coniche:

$$x_1^2 + 2x_2^2 = 3, \quad x_1 x_2 = 1, \quad x_1^2 - 2x_2 = 0$$

Nel primo caso abbiamo un'*ellisse*, nel secondo un'*iperbole* e nel terzo una *parabola*.

• Le seguenti equazioni descrivono delle quadriche in $\mathbb{R}^3$

$$x_1^2 + x_2^2 + x_3^2 = 1, \quad x_1^2 + x_2^2 = 1$$

Nel primo caso abbiamo una *sfera*, mentre nel secondo un cilindro.

Rappresentazioni matriciali.

Data l'equazione di una quadrica $q(x) = \sum_{i,j=1}^n a_{ij} x_i x_j + 2 \sum_{i=1}^n b_i x_i + c$, questa può essere riscritta come:

$$q(x) = x^T A x + 2b^T x + c \quad \text{dove } (A_{i,j}) = a_{i,j}, \quad b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

Alternativamente, si può descrivere $q(x)$ con un'unica matrice:

$$q(x) = \sum_{i,j=1}^n a_{ij} x_i x_j + 2 \sum_{i=1}^n b_i x_i + c = \begin{pmatrix} x_1 & \dots & x_n & 1 \end{pmatrix} \left(\begin{array}{c|c} A & b \\ \hline b^T & c \end{array} \right) \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ 1 \end{pmatrix} = x^T A' x$$

La matrice $A' \in \mathcal{M}_{\mathbb{R}}(n+1)$ è la *matrice completa* della quadrica.

Esempio. La conica $x_1^2 + 2x_2^2 - 4x_1x_2 + x_1 - 6x_2 + 5 = 0$ si può riscrivere come:

$$0 = x_1^2 + 2x_2^2 - 4x_1x_2 + x_1 - 6x_2 + 5 = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + 2 \cdot \begin{pmatrix} \frac{1}{2} & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + 5$$

Oppure come

$$0 = x_1^2 + 2x_2^2 - 4x_1x_2 + x_1 - 6x_2 + 5 = \begin{pmatrix} x_1 & x_2 & 1 \end{pmatrix} \begin{pmatrix} 1 & -2 & \frac{1}{2} \\ -2 & 1 & -3 \\ \frac{1}{2} & -3 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ 1 \end{pmatrix}$$

Definizione 6.2. Una quadrica è *degenere* se la sua matrice completa non ha rango massimo, cioè se $\det(A') = 0$.

Esempio. • Le matrici complete delle coniche descritte in precedenza sono:

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

quindi sono tutte coniche non degeneri.

Le seguenti equazioni invece descrivono delle coniche degeneri:

$$0 = x_1^2 - 1 \leftrightarrow \begin{pmatrix} 1 & & \\ & 0 & \\ & & -1 \end{pmatrix}, \quad 0 = x_1^2 + 2x_1x_2 + x_2^2 - 2x_1 - 2x_2 + 1 \leftrightarrow \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix}.$$

Nel primo caso abbiamo una coppia di *rette parallele*, mentre nel secondo una *retta doppia* (di equazione $x_1 + x_2 - 1 = 0$).

Domanda. Data una quadrica esiste una “trasformazione di coordinate” in cui l’equazione che la descrive diventi più semplice? Se sì, come trovarla?

In altre parole ci stiamo chiedendo se possiamo “cambiare il sistema di riferimento” in modo da semplificare la quadrica.

Esempio. Consideriamo la conica $0 = 3x_1^2 + 2x_1x_2 + 3x_2^2 - 1 = x^T A x - 1$ con $A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$.

La matrice A è simmetrica e quindi ortogonalmente diagonalizzabile, cioè esiste una base ortonormale $\mathcal{B}$ di autovalori per A : A ha autovalori 4, 2 e autovettori corrispondenti sono $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})^T$ e $(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})^T$.

Quindi

$$P^T A P = D = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix} \quad \text{con} \quad P = \mathcal{M}_{\mathcal{E}}^{\mathcal{B}}(id_{\mathbb{R}^2}) = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

e applicando il cambio di coordinate $x = Py$, la conica diventa

$$0 = (Py)^T A (Py) - 1 = y^T P^T A P y - 1 = y^T D y - 1 \Leftrightarrow 4y_1^2 + 2y_2^2 = 1$$

la conica è un'ellisse.

Nelle prossime pagine risponderemo positivamente alla domanda e inoltre mostreremo che ogni quadrica può essere ricondotta ad una *forma canonica*, cioè vedremo come “classificare” le quadriche.

Ci sono “2 tipi” di trasformazioni affini: le isometrie (che sono rigide: preservano angoli, lunghezze, distanze, etc.; vedi Definizione 5.37) e le restanti (che non sono rigide e “deformano” lo spazio, non preservando angoli e lunghezze). Pertanto possiamo classificare le quadriche in 2 modi.

Definizione 6.3. Sia $Q \subseteq \mathbb{R}^n$ una quadrica. Un processo di *semplificazione* dell'equazione di Q (detto anche processo di *classificazione*) si dice

- *metrico* se le trasformazioni ammesse sono solo isometrie, cioè matrici ortogonali (isometrie lineari) e traslazioni;
- *affine*, se le trasformazioni ammesse sono matrici invertibili e traslazioni.

Esempio. Nell'esempio precedente abbiamo semplificato *metricamente* la conica, infatti abbiamo fatto un unico cambio di coordinate dato da una matrice ortogonale.

Se ora facciamo un'ulteriore cambio di coordinate $z_1 = 2y_1$, $z_2 = \sqrt{2}y_2$, cioè $z = \begin{pmatrix} 2 \\ \sqrt{2} \end{pmatrix} y$, la conica diventa $z_1^2 + z_2^2 = 1$ (circonferenza unitaria). L'ultimo cambio di coordinate non è dato da una matrice ortogonale quindi abbiamo una semplificazione *affine*.

6.1.1 Centro di una quadrica

La quadrica $Q := \{q(x) = x^T A x + 2b^T x + c = 0\}$ ha un centro di simmetria in 0 precisamente quando $q(x) = 0 \Leftrightarrow q(-x) = 0$ per ogni $x \in \mathbb{R}^n$. In particolare, abbiamo che $q(x) = q(-x)$, cioè

$$x^T A x + 2b^T x + c = x^T A x - 2b^T x + c \iff b^T x = 0, \quad \forall x \in \mathbb{R}^n$$

cioè $b = 0$.

Quindi Q ha una simmetria rispetto a 0 se e solo se $b = 0$, cioè non ho la parte lineare.

Strategia. Per capire se una quadrica ha un centro di simmetria in v_0 , posso applicare la traslazione $T : \mathbb{R}^n \rightarrow \mathbb{R}^n, v \mapsto v - v_0$ che porta v_0 nell'origine e vedere se la parte lineare si annulla (e quindi ho un centro di simmetria in v_0) oppure no.

Nelle nuove coordinate $y = x - v_0$ (cioè $x = y + v_0$) la quadrica diventa

$$\begin{aligned} 0 &= x^T A x + 2b^T x + c = (y + v_0)^T A (y + v_0) + 2b^T (y + v_0) + c \\ &= y^T A y + y^T A v_0 + v_0^T A y + v_0^T A v_0 + 2b^T y + 2b^T v_0 + c \end{aligned}$$

Notiamo che $y^T Av_0 = v_0^T Ay$; infatti $y^T Av_0 = (y^T Av_0)^T = v_0^T A^T y = v_0^T Ay$, dove la prima uguaglianza vale perché stiamo trasponendo una matrice di ordine 1, e la seconda vale perché $A = A^T$. Quindi otteniamo

$$0 = y^T Ay + 2(v_0^T A + b^T)y + \underbrace{2b^T v_0 + v_0^T Av_0 + c}_{c'} \stackrel{A=A^T}{=} y^T Ay + 2(\underbrace{Av_0 + b}_{b'})^T y + c' \quad (6.1)$$

Quindi per eliminare la parte lineare posso traslare per una soluzione v_0 del sistema lineare $Ax = -b$.

Definizione 6.4. Una quadrica $Q := \{q(x) = x^T Ax + 2b^T x + c = 0\}$ si dice *a centro* se ha un *centro di simmetria*, cioè se l'equazione (dei centri) $Ax = -b$ ha almeno una soluzione. L'insieme soluzione $\text{Sol}(A| -b)$ è detto l'*insieme dei centri di simmetria*.

Osservazione. Se A ha rango massimo $\text{rg}(A) = n$, allora è invertibile e quindi $\text{Sol}(A| -b)$ ha un unico centro di simmetria.

Esempio. • La quadrica $x^T \begin{pmatrix} 1 & 1 & -1 \\ 1 & 2 & 0 \\ -1 & 0 & -1 \end{pmatrix} x + 2 \begin{pmatrix} 1 & -1 & 0 \end{pmatrix} x + 1 = 0$ ha un unico centro,

infatti $\left(\begin{array}{ccc|c} 1 & 1 & -1 & -1 \\ 1 & 2 & 0 & 1 \\ -1 & 0 & -1 & 0 \end{array} \right)$ ha un'unica soluzione: $\text{Sol}(A| -b) = (-1, 1, 1)^T$.

• La quadrica $x^T \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} x + 2 \begin{pmatrix} -3 & -3 & -3 \end{pmatrix} x + 11 = 0$ ha infiniti centri, infatti

$\left(\begin{array}{ccc|c} 1 & 1 & 0 & 3 \\ 1 & 1 & 0 & 3 \\ 0 & 0 & 2 & 3 \end{array} \right)$ ha infinite soluzioni: $\text{Sol}(A| -b) = \{(3 - t, t, \frac{3}{2})^T \mid t \in \mathbb{R}\}$.

• La quadrica $x^T \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} x + 2 \begin{pmatrix} -1 & 2 & 3 \end{pmatrix} x + 1 = 0$ non ha centro, infatti

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

• Vedremo in seguito che le tre quadriche sono rispettivamente un *iperboloide iperbolico*, un *cilindro ellittico* ed un *paraboloide iperbolico*.

6.2 Classificazione metrica ed affine

Nel processo di classificazione dobbiamo distinguere tra le quadriche con un centro di simmetria (per esempio una sfera o un cilindro) e quelle prive di centro di simmetria (per esempio una parabola), perché nel primo caso possiamo eliminare la parte lineare con una traslazione, mentre nel secondo no.

6.2.1 Quadriche a centro

Sia $Q := \{q(x) = x^T A x + 2b^T x + c = 0\}$ una quadrica a centro, cioè $\text{Sol}(A| - b)$ è non è vuoto e sia quindi $v_0 \in \text{Sol}(A| - b)$ un centro di simmetria.

◇ Applicando la traslazione $T : \mathbb{R}^n \rightarrow \mathbb{R}^n, v \mapsto v - v_0$, otteniamo nuove coordinate $y = T(x) = x - v_0$, nelle quali la quadrica diventa (si veda (6.1))

$$0 = q(y + v_0) = y^T A y + c' \quad \text{dove} \quad c' = 2b^T v_0 + \underbrace{v_0^T A v_0}_{-b^T} + c = b^T v_0 + c$$

◇ Ora semplifichiamo la forma quadratica data da A (seguiamo la stessa idea della dimostrazione del Teorema 5.48).

Visto che A è simmetrica, per il teorema spettrale esiste una matrice ortogonale $P \in \mathcal{M}_{\mathbb{R}}(n)$ tale che

$$P^{-1} A P = P^T A P = D = \text{diag}(\lambda_1, \dots, \lambda_n).$$

Più precisamente $P = \mathcal{M}_{\mathcal{E}}^{\mathcal{C}}(id_{\mathbb{R}^n})$, dove $\mathcal{C}$ è una base ortonormale di autovettori di A .

Applicando quindi l'isometria lineare $L : \mathbb{R}^n \rightarrow \mathbb{R}^n, v \mapsto P^{-1}v = P^T v$ otteniamo nuove coordinate $z = L(y) = P^T y$ (cioè $y = Pz$) nelle quali la quadrica diventa

$$0 = q(Pz) = (Pz)^T A (Pz) + c' = z^T P^T A P z + c' = z^T D z + c'$$

e in queste coordinate l'equazione è molto semplice, infatti compaiono solo termini quadratici e il termine costante: non ci sono termini lineari e termini misti di grado 2 (e.g. come $x_1 x_2$): la matrice completa della quadrica è

$$\left(\begin{array}{ccc|c} \lambda_1 & & & 0 \\ & \ddots & & \vdots \\ & & \lambda_n & 0 \\ \hline 0 & \dots & 0 & c' \end{array} \right)$$

Definizione 6.5. La *forma canonica metrica* di una quadrica a centro è

$$\begin{cases} \lambda_1 z_1^2 + \dots + \lambda_p z_p^2 = 0 & \text{se } c' = 0; \\ \frac{\lambda_1}{c'} z_1^2 + \dots + \frac{\lambda_p}{c'} z_p^2 + 1 = 0 & \text{se } c' \neq 0; \end{cases} \quad (6.2)$$

dove $\lambda_1, \dots, \lambda_p$ sono gli autovalori non nulli di A .

Se ora accettiamo di applicare una trasformazione di coordinate non ortogonale, possiamo semplificare ulteriormente l'equazione (similmente a quanto fatto nella dimostrazione del Teorema 5.48) e abbiamo la *classificazione affine*.

◇ Applicando la dilatazione/compressione $\mathcal{L}_S : \mathbb{R}^n \rightarrow \mathbb{R}^n, v \mapsto S v$ data dalla matrice

$$S = \left(\begin{array}{ccc|c} \alpha_1 & & & \\ & \ddots & & \\ & & \alpha_p & \\ \hline & & & 1 \\ & & & & \ddots & \\ & & & & & 1 \end{array} \right) \quad \text{dove} \quad \alpha_j = \begin{cases} \sqrt{|\lambda_j|} & \text{se } c' = 0 \\ \sqrt{\left| \frac{\lambda_j}{c'} \right|} & \text{se } c' \neq 0 \end{cases}$$

otteniamo nuove coordinate $w = S(z)$, cioè $z_i = \alpha_i^{-1} w_i$ se $i = 1, \dots, p$, mentre $z_i = w_i$ se $i = p + 1, \dots, n$. Sostituendo nella corrispondente equazione (6.2) otteniamo

$$\begin{aligned} 0 &= \sum_i \lambda_i \left(\frac{1}{\sqrt{|\lambda_i|}} w_i \right)^2 = \sum_i \operatorname{sgn}(\lambda_i) w_i^2 & \text{se } c' = 0 \\ 0 &= 1 + \sum_i \frac{\lambda_i}{c'} \left(\sqrt{\left| \frac{c'}{\lambda_i} \right|} w_i \right)^2 = 1 + \sum_i \operatorname{sgn} \left(\frac{\lambda_i}{c'} \right) w_i^2 & \text{se } c' \neq 0 \end{aligned}$$

dove $\operatorname{sgn} : \mathbb{R} \rightarrow \mathbb{R}$ è la funzione segno: $\operatorname{sgn}(x) = \begin{cases} -1 & \text{se } x < 0 \\ 0 & \text{se } x = 0 \\ 1 & \text{se } x > 0 \end{cases}$.

Osservazione. Moralmemente, rimpiazziamo tutti i coefficienti nella forma canonica metrica con i rispettivi segni.

Definizione 6.6. La *forma canonica affine* di una quadrica a centro è

$$\begin{cases} w_1^2 + \dots + w_r^2 - w_{r+1}^2 - \dots - w_{r+s}^2 = 0 & \text{se } c' = 0; \\ w_1^2 + \dots + w_r^2 - w_{r+1}^2 - \dots - w_{r+s}^2 + 1 = 0 & \text{se } c' \neq 0; \end{cases}$$

dove $(r, s, n - (r + s))$ è la segnatura della forma quadratica associata ad A .

Esempio. • Consideriamo la conica $C := \{3x_1^2 + 2x_1x_2 + 3x_2 + 2x_1 + 6x_2 + 1 = 0\}$.

In questo caso $A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$, $b = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$ e $c = 1$. La conica C è a centro, infatti il sistema lineare $Ax = -b$ è risolubile:

$$\left(\begin{array}{cc|c} 3 & 1 & -1 \\ 1 & 3 & -3 \end{array} \right) \xrightarrow{MEG} \left(\begin{array}{cc|c} 3 & 0 & 0 \\ 0 & 1 & -1 \end{array} \right)$$

e $v_0 := (0, -1)^T$ è l'unico centro della conica. Ponendo $y = x - v_0$, cioè $x = y + v_0$ otteniamo l'equazione:

$$3y_1^2 + 2y_1y_2 + 3y_2 - 2 = 0$$

infatti $c' = b^T v_0 + c = (1, 3) \cdot (0, -1)^T + 1 = -2$.

A questo punto dobbiamo diagonalizzare ortogonalmente la matrice A , ma basta calcolarne il polinomio caratteristico:

$$p_A(t) = t^2 - 6t + 8 = (t - 4)(t - 2)$$

Gli autovalori sono 2 e 4, quindi la forma canonica metrica e la forma canonica affin e di C sono

$$2z_1^2 + 4z_2^2 - 2 = 0, \quad w_1^2 + w_2^2 - 1 = 0 \quad (\text{ellisse}).$$

• La quadrica $x^T \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} x + 2 \begin{pmatrix} -3 & -3 & -3 \end{pmatrix} x + 11 = 0$ ha infiniti centri: $\operatorname{Sol}(A| -$

$b) = \{(3 - t, t, \frac{3}{2})^T \mid t \in \mathbb{R}\}$. Sia $v_0 = (3, 0, \frac{3}{2})^T$ un centro, ponendo $y = x - v_0$ otteniamo l'equazione

$$y_1^2 + 2y_1y_2 + y_2^2 + 2y_3^2 - \frac{1}{2} = 0$$

infatti $c' = (-3, -3, -3)v_0 + c = -9 - 9/2 + 11 = -1/2$.

Il polinomio caratteristico di A è $p_A(t) = \det(tI_3 - A) = t(t-2)^2$ e gli autovalori sono 0 e 2 (doppio) e quindi la forma canonica metrica della quadrica è

$$2z_1^2 + 2z_2^2 - \frac{1}{2} = 0,$$

mentre la sua forma canonica affine è $w_1^2 + w_2^2 - 1 = 0$ (*cilindro ellittico*).

6.2.2 Quadriche non a centro

Sia $Q := \{q(x) = x^T A x + 2b^T x + c = 0\}$ una quadrica priva di centro, cioè $\text{Sol}(A| - b)$ non ha soluzioni, cioè $-b \notin \text{Col}(A) \Leftrightarrow b \notin \text{Col}(A)$. In particolare $\text{rg}(A) < \text{rg}(A| - b)$, e quindi $\text{rg}(A) < n$.

Però posso risolvere il sistema lineare $Ax = -b$ in modo approssimato, cioè considerare il vettore $\tilde{b} \in \text{Col}(A)$ che meglio approssima b . In altre parole, la proiezione ortogonale di b sul sottospazio $W = \text{Col}(A)$: $\tilde{b} := \pi_W(b) \in W = \text{Col}(A)$. In questo modo il sistema lineare $Ax = -\tilde{b} = \pi_W(-b)$ ha sempre soluzione.

Lemma 6.7. Sia $A \in \mathcal{M}_{\mathbb{R}}(n)$ una matrice simmetrica: $A = A^T$, e sia $W = \text{Col}(A)$. Allora:

i) $\ker A = \ker A^2 = W^\perp$;

ii) dato $b \in \mathbb{R}^n$ sia $\tilde{b} := \pi_W(b)$, allora $\text{Sol}(A| - \tilde{b}) = \text{Sol}(A^2| - Ab)$.

Dimostrazione. i) Sia $v \in \ker A$, allora $A \cdot v = 0$. Dunque $A^2 \cdot v = A \cdot (A \cdot v) = A \cdot 0 = 0$, cioè $v \in \ker A^2$.

Viceversa, sia $v \in \ker A^2$, allora $\|A \cdot v\|^2 = (A \cdot v)^T \cdot (A \cdot v) = v^T \cdot A^T \cdot A \cdot v \stackrel{A=A^T}{=} v^T \cdot A^2 \cdot v = 0$, cioè $A \cdot v = 0$: $v \in \ker A$.

Inoltre, $\text{Col}(A)^\perp \stackrel{A=A^T}{=} \text{Row}(A)^\perp = \{x \in \mathbb{R}^n : A \cdot x = 0\} = \ker A$.

ii) Scriviamo $b = b' + \tilde{b}$ con $b' = \pi_{W^\perp}(b)$; in particolare $Ab' = 0$ visto che $W^\perp = \ker A$.

Sia $v \in \text{Sol}(A| - \tilde{b})$, cioè $Av = -\tilde{b}$, allora $A^2v = -Ab' - A\tilde{b} = -A(b' + \tilde{b})$ e quindi $A^2v = -Ab$. Dunque $\text{Sol}(A| - \tilde{b}) \subseteq \text{Sol}(A^2| - Ab)$, ma questi due sottospazi affini hanno la stessa giacitura ($\ker A = \ker A^2$) e quindi sono uguali. $\square$

Definizione 6.8. Sia $Q := \{q(x) = x^T A x + 2b^T x + c = 0\}$ una quadrica, e sia $\tilde{b} := \pi_W(b)$ con $W = \text{Col}(A)$. L'asse della quadrica è definito come

$$\mathcal{A}_Q = \text{Sol}(A| - \tilde{b}) = \text{Sol}(A^2| - Ab).$$

Osservazione. Per quanto osservato l'asse di una quadrica esiste sempre.

Se la quadrica è a centro, l'asse coincide con l'insieme dei centri di simmetria.

Se la quadrica è priva di centro, il sistema lineare $Ax = -b$ non ha soluzioni e quindi $\text{rg}(A) < n$, e dunque $\text{Sol}(A| - \tilde{b}) = \text{Sol}(A^2| - Ab)$ ha un numero infinito di soluzioni.

Definizione 6.9. Sia $Q \subset \mathbb{R}^n$ una quadrica priva di centro. Chiamiamo *vertice* di Q un qualunque punto $v_0 \in Q \cap \mathcal{A}_Q$.

Esempio. • La conica $x_1^2 - x_2 = x^T \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} x + 2 \begin{pmatrix} -0 & -\frac{1}{2} \end{pmatrix} x = 0$ è una parabola e non ha centro, determiniamo il suo asse:

$$\mathcal{A} = \text{Sol}(A^2 | -Ab) = \text{Sol}(A^2 | 0) = \ker A^2 = \ker A = \text{Span}((0, 1)^T).$$

Quindi il vertice è $v_0 = (0, 0)^T$.

• La quadrica $x^T \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} x + 2 \begin{pmatrix} -1 & 2 & 3 \end{pmatrix} x + 1 = 0$ non ha centro, calcoliamo il suo asse:

$$(A^2 | -Ab) = \left(\begin{array}{ccc|c} 2 & 2 & 0 & -1 \\ 2 & 2 & 0 & -1 \\ 0 & 0 & 4 & 6 \end{array} \right) \implies \mathcal{A} = \text{Sol}(A^2 | -Ab) = \left\{ \left(-\frac{1}{2} - t, t, \frac{3}{2} \right)^T \mid t \in \mathbb{R} \right\}$$

Calcoliamo i vertici: inserendo le coordinate di un punto di $\mathcal{A}$ nell'equazione della quadrica otteniamo $6t + \frac{27}{4} = 0$, quindi la quadrica ha un unico vertice $v_0 := (\frac{5}{8}, -\frac{9}{8}, \frac{3}{2})^T$.

Ora procediamo come nel caso delle quadriche a centro ed applichiamo una serie di trasformazioni affini per semplificare l'equazione della quadrica.

Sia quindi $Q := \{q(x) = x^T A x + 2b^T x + c = 0\}$ una quadrica priva di centro, e sia $v_0 \in \mathcal{A}_Q$ un vertice Q , cioè $Av_0 = -\pi_W(b)$ con $W = \text{Col}(A)$ e $q(v_0) = v_0^T A v_0 + 2b^T v_0 + c = 0$
Attenzione: si può dimostrare che un vertice esiste sempre: $Q \cap \mathcal{A}_Q \neq \emptyset$ (Esercizio).

◊ Applicando la traslazione $T : \mathbb{R}^n \rightarrow \mathbb{R}^n, v \mapsto v - v_0$, otteniamo nuove coordinate $y = T(x) = x - v_0$, nelle quali la quadrica diventa (si veda (6.1))

$$0 = q(y + v_0) = y^T A y + 2 \underbrace{(v_0^T A + b^T)}_{b'^T} y + \underbrace{2b^T v_0 + v_0^T A v_0 + c}_{c' = q(v_0) = 0}$$

dove $b' = (Av_0 + b) = b - \pi_W(b) = \pi_{\ker A}(b)$.

◊ Ora semplifichiamo la forma quadratica data da A usando il teorema spettrale (seguiamo la stessa idea della dimostrazione del Teorema 5.48).

Visto che A è simmetrica, esiste una base ortonormale $\mathcal{C}$ di autovettori per A , e posso imporre che $\frac{b'}{\|b'\|} \in \ker A$ appartenga alla base (si veda pagina 98):

$$\mathcal{C} := \{b_1, \dots, b_p, \underbrace{\frac{b'}{\|b'\|}, b_{p+2}, \dots, b_n}_{\text{base ortonormale di } V_0 = \ker A}\}$$

dove $b_1, \dots, b_p$ sono autovettori relativi agli autovalori non-nulli $\lambda_1, \dots, \lambda_p$ di A : $p = \text{rg}(A)$.

Sia $P \in \mathcal{M}_{\mathbb{R}}(n)$ la matrice ortogonale di cambio di base $P = \mathcal{M}_{\mathcal{C}}^{\mathcal{C}}(id_{\mathbb{R}^n})$, allora

$$P^{-1}AP = P^TAP = D = \text{diag}(\lambda_1, \dots, \lambda_p, 0, \dots, 0).$$

Applicando l'isometria lineare $L : \mathbb{R}^n \rightarrow \mathbb{R}^n, v \mapsto P^{-1}v = P^T v$ otteniamo nuove coordinate $z = L(y) = P^T y$ (cioè $y = Pz$) nelle quali la quadrica diventa

$$0 = q(Pz) = (Pz)^T A (Pz) + 2(b')^T (Pz)' = z^T P^T A P z + 2(P^T b')^T z$$

dove $P^T b' = P^{-1} b'$ sono le coordinate di b' nella base $\mathcal{C}$, cioè $P^T b' = (\underbrace{0, \dots, 0}_p, \|b'\|, 0, \dots, 0)^T$.

In queste coordinate l'equazione è molto semplice, infatti compaiono solo termini quadratici, ed un unico termine di grado 1: non ci sono ulteriori termini lineari, o termini misti di grado 2, o termini noti. La matrice completa della quadrica è

$$\left(\begin{array}{cccccc|c} \lambda_1 & & & & & & 0 \\ & \ddots & & & & & \vdots \\ & & \lambda_p & & & & 0 \\ & & & 0 & & & \|b'\| \\ & & & & \ddots & & 0 \\ & & & & & 0 & \vdots \\ 0 & \dots & 0 & \|b'\| & 0 & \dots & 0 \end{array} \right)$$

Definizione 6.10. La *forma canonica metrica* di una quadrica priva di centro $x^T A x + 2b^T x + c = 0$ è

$$\lambda_1 z_1^2 + \dots + \lambda_p z_p^2 + 2\|b'\| z_{p+1} = 0 \quad (6.3)$$

dove $\lambda_1, \dots, \lambda_p$ sono gli autovalori non nulli di A , e $b' = \pi_{\ker A}(b)$.

Come nel caso delle quadriche a centro, se ora accettiamo di applicare una trasformazione di coordinate non ortogonale, possiamo semplificare ulteriormente l'equazione ed ottenere la *classificazione affine*.

◊ Moralmemente, rimpiazziamo tutti i coefficienti nella forma canonica metrica con i rispettivi segni. Formalmente, applicando la dilatazione/compressione $\mathcal{L}_S : \mathbb{R}^n \rightarrow \mathbb{R}^n, v \mapsto S v$ data dalla matrice

$$S = \left(\begin{array}{cccc|ccc} \sqrt{|\lambda_1|} & & & & & & \\ & \ddots & & & & & \\ & & \sqrt{|\lambda_p|} & & & & \\ & & & \|b'\| & & & \\ \hline & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{array} \right)$$

otteniamo nuove coordinate $w = S(z)$, cioè $z_i = \frac{w_i}{\sqrt{|\lambda_i|}}$ se $i = 1, \dots, p$, $z_{p+1} = \frac{w_{p+1}}{\|b'\|}$,

mentre $z_i = w_i$ se $i = p+2, \dots, n$. Sostituendo nell'equazione (6.3) otteniamo

$$0 = \sum_{i=1}^p \lambda_i \left(\frac{w_i}{\sqrt{|\lambda_i|}} \right)^2 + 2\|b'\| \frac{w_{p+1}}{\|b'\|} = \sum_i \text{sgn}(\lambda_i) w_i^2 + 2w_{p+1}$$

dove $\text{sgn} : \mathbb{R} \rightarrow \mathbb{R}$ è la funzione segno.

Osservazione. Se nella matrice S sostituiamo $s_{p+1,p+1} = \|b'\|$ con $s_{p+1,p+1} = -\|b'\|$, allora il cambio di coordinate ci restituisce l'equazione

$$0 = \sum_i \operatorname{sgn}(\lambda_i) w_i^2 - 2w_{p+1}$$

Quindi, indipendentemente dai segni di $\lambda_1, \dots, \lambda_p$, possiamo sempre assumere che il coefficiente di w_{p+1} sia positivo.

Definizione 6.11. La *forma canonica affine* di una quadrica priva di centro è

$$w_1^2 + \dots + w_r^2 - w_{r+1}^2 - \dots - w_{r+s}^2 + 2w_{r+s+1} = 0$$

dove $(r, s, n - (r + s))$ è la segnatura della forma quadratica associata ad A .

Esempio. • La quadrica $Q : \{x^T \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} x + 2 \begin{pmatrix} -1 & 2 & 3 \end{pmatrix} x + 1 = 0\}$ non ha centro,

e il suo asse è $\mathcal{A} = \operatorname{Sol}(A^2 | -Ab) = \{(-\frac{1}{2} - t, t, \frac{3}{2})^T \mid t \in \mathbb{R}\}$. La matrice A ha rango 2 e una base di $\ker A$ è $\{(1, -1, 0)^T\}$; calcoliamo quindi

$$b' = \pi_{\ker A}(b) = \frac{\langle (-1, 2, 3)^T, (1, -1, 0)^T \rangle}{\|(1, -1, 0)^T\|^2} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = -\frac{3}{2} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad \|b'\| = \frac{3}{\sqrt{2}}$$

A questo punto dobbiamo diagonalizzare ortogonalmente la matrice A e basta calcolarne il polinomio caratteristico:

$$p_A(t) = (t + 2)(t^2 - 2t) = t(t - 2)(t + 2)$$

Gli autovalori sono 0, -2 e 2, quindi la forma canonica metrica di Q è

$$2z_1^2 - 2z_2^2 + 3\sqrt{2}z_3 = 0$$

e la sua forma canonica affine è $w_1^2 - w_2^2 + 2w_3 = 0$ (*paraboloide iperbolico*).

• La quadrica $Q : \{x^T \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} x + 2 \begin{pmatrix} -2 & -2 & -2 \end{pmatrix} x + 4 = 0\}$ non ha centro,

infatti $\left(\begin{array}{ccc|c} 1 & -1 & 0 & 2 \\ -1 & 1 & 0 & 2 \\ 0 & 0 & 0 & 2 \end{array} \right)$ non ha soluzione. La matrice A ha rango 2 e una base di $\ker A$

è $\{(1, 1, 0)^T\}$; calcoliamo quindi

$$b' = \pi_{\ker A}(b) = \frac{\langle (2, 2, 2)^T, (1, 1, 0)^T \rangle}{\|(1, 1, 0)^T\|^2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \|b'\| = 2\sqrt{2}$$

A questo punto dobbiamo diagonalizzare ortogonalmente la matrice A e basta calcolarne il polinomio caratteristico:

$$p_A(t) = t(t^2 - 2t) = t^2(t - 2)$$

Gli autovalori sono 0 (doppio) e 2, quindi la forma canonica metrica di Q è

$$2z_1^2 + 4\sqrt{2}z_2 = 0$$

e la sua forma canonica affine è $w_1^2 + 2w_2 = 0$ (*cilindro parabolico*).

6.2.3 Sintesi del processo di classificazione

Input: $Q := \{q(x) = x^T A x + 2b^T x + c = 0\}$

Domanda: Q è a centro? Cioè $Ax = -b$ ha soluzione?

Sì

No

1. Calcolare un centro $v_0 \in \text{Sol}(A| - b)$ e determinare $c' = b^T v_0 + c$.
2. Calcolare $\lambda_1, \dots, \lambda_p \neq 0$ autovalori non-nulli di A .
3. La forma canonica metrica è

1. Calcolare $b' = \pi_{\ker A}(b)$ e $\|b'\|$.
2. Calcolare $\lambda_1, \dots, \lambda_p \neq 0$ autovalori non-nulli di A .
3. La forma canonica metrica è

$$\begin{cases} \lambda_1 z_1^2 + \dots + \lambda_p z_p^2 = 0 & \text{se } c' = 0; \\ \frac{\lambda_1}{c'} z_1^2 + \dots + \frac{\lambda_p}{c'} z_p^2 + 1 = 0 & \text{se } c' \neq 0; \end{cases} \quad \lambda_1 z_1^2 + \dots + \lambda_p z_p^2 + 2\|b'\| z_{p+1} = 0$$

4. Per le forme canoniche affini, sostituire tutti i coefficienti con i rispettivi segni.

Esempio. Consideriamo la quadrica $Q : \{x^T \begin{pmatrix} 1 & 1 & -1 \\ 1 & 2 & 0 \\ -1 & 0 & -1 \end{pmatrix} x + 2 \begin{pmatrix} 1 & -1 & -0 \end{pmatrix} x + 1 = 0\}$.

Il polinomio caratteristico di A è $p_A(t) = \det(tI_3 - A) = t^3 - 2t^2 - 3t + 2$ e per il criterio di Cartesio ha 2 radici positive ed una negativa; quindi la segnatura di A è $(2, 1, 0)$.

Ora vediamo se Q è a centro: $\left(\begin{array}{ccc|c} 1 & 1 & -1 & -1 \\ 1 & 2 & 0 & 1 \\ -1 & 0 & -1 & 0 \end{array} \right)$ ha un'unica soluzione $v_0 = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$,

da cui $c' = b^T v_0 + c = -2 + 1 = -1 < 0$.

Per cui la forma canonica affine di Q è $x_1^2 - x_2^2 - x_3^2 + 1 = 0$ (*iperboloide iperbolico*, ad una falda).

6.3 Simmetrie di quadriche

Concludiamo questo capitolo e queste note con una breve discussione sulle simmetrie di quadriche.

Definizione 6.12. Data una quadrica $Q \subset \mathbb{R}^n$, una *simmetria* di Q è un'isometria $\sigma : \mathbb{R}^n \rightarrow \mathbb{R}^n$ tale che $\sigma(Q) = Q$, cioè $\sigma(v) \in Q \Leftrightarrow v \in Q$.

6.3.1 Simmetrie assiali

Sia $S := P_0 + W \subset \mathbb{R}^n$ un sottospazio affine ($P_0 \in \mathbb{R}^n$, $W \triangleleft \mathbb{R}^n$), e definiamo la “riflessione rispetto ad S ”:

$$\begin{aligned} \sigma_S : \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ v &\longmapsto v - 2\pi_{W^\perp}(v - P_0) \end{aligned}$$

Nel caso in cui $S = P_0$ sia solo un punto, allora la mappa assume la forma: $\sigma_{P_0}(v) = v - 2(v - P_0) = 2P_0 - v$.

Definizione 6.13. In particolare, $Q \subset \mathbb{R}^n$ è *simmetrica rispetto al sottospazio affine* $S := P_0 + W \subset \mathbb{R}^n$ se $v \in Q \Leftrightarrow \sigma_S(v) \in Q$.

Esempio. • Sia $Q \subset \mathbb{R}^n$ una quadrica a centro. Se v_0 è un centro di simmetria, allora Q è simmetrica rispetto a v_0 : $v \in Q \Leftrightarrow 2v_0 - v \in Q$.

- La quadrica $x^T \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} x + 2 \begin{pmatrix} -3 & -3 & -3 \end{pmatrix} x + 11 = 0$ ha infiniti centri $S :=$

$\text{Sol}(A|-b) = \{(3-t, t, \frac{3}{2})^T \mid t \in \mathbb{R}\}$, ed è simmetrica rispetto a ciascun centro. Inoltre Q è simmetrica rispetto alla la retta dei centri S . Infatti $\sigma_S((x, y, z)^T) = (3-x, 3-y, 3-z)^T$ e sostituendo $(3-x, 3-y, 3-z)^T$ nell'equazione della quadrica, si riottiene la stessa equazione, quindi $v \in Q \Leftrightarrow \sigma_S(v) \in Q$.

Osservazione. Una quadrica Q è sempre simmetrica rispetto al suo asse $\mathcal{A}_Q$.

Sia $Q = \{x \in \mathbb{R}^n \mid x^T A x + b^T x + c = 0\} \subset \mathbb{R}^n$ una quadrica *a centro*, e sia $\mathcal{B} = \{v_1, \dots, v_n\}$ una base ortonormale di autovettori di A . Questa base descrive 2 tipi di simmetrie:

- rispetto alla retta $v_0 + v_j \mathbb{R}$ (dove v_0 è un centro);
- rispetto all'iperpiano H_j passante per un centro v_0 ed ortogonale a v_j :

$$H_j = \{x \in \mathbb{R}^n \mid \langle x - v_0, v_j \rangle = 0\} = v_0 + \text{Span}(v_1, \dots, v_{j-1}, v_{j+1}, \dots, v_n)$$

Lasciamo come esercizio per il lettore la verifica che la quadrica Q è simmetrica rispetto a questi sottospazi affini.

Sia ora $Q = \{x \in \mathbb{R}^n \mid x^T A x + b^T x + c = 0\} \subset \mathbb{R}^n$ una quadrica *priva di centro*, e sia $\mathcal{B} = \{v_1, \dots, v_p, v_{p+1}, \dots, v_n\}$ una base ortonormale di autovettori di A , dove $v_1, \dots, v_p$ hanno autovalori non nulli e $p = \text{rg}(A)$. Sia inoltre $v_0 \in Q \cap \mathcal{A}_Q$ un vertice della quadrica. Allora posso definire p simmetrie, una per ciascun iperpiano:

$$H_j = \{x \in \mathbb{R}^n \mid \langle x - v_0, v_j \rangle = 0\} = v_0 + \text{Span}(v_1, \dots, v_{j-1}, v_{j+1}, \dots, v_p, \dots, v_n), \quad j \in \{1, \dots, p\}.$$

Si osservi che ciascuno di questi iperpiani contiene l'asse della quadrica $\mathcal{A}_Q$.

Anche in questo caso lasciamo come esercizio per il lettore la verifica che la quadrica Q è simmetrica rispetto a questi iperpiani.

Esempio. • La quadrica $x^T \begin{pmatrix} 37 & 20 & 2 \\ 20 & 49 & 22 \\ 2 & 22 & 58 \end{pmatrix} x - 9 = 0$ ha un unico centro: $0 \in \mathbb{R}^3$.

La matrice $A = \begin{pmatrix} 37 & 20 & 2 \\ 20 & 49 & 22 \\ 2 & 22 & 58 \end{pmatrix}$ ha 3 autovalori distinti: 81, 45, 18; e gli autospazi sono

$$V_{81} = \text{Span}((1, 2, 2)^T), V_{45} = \text{Span}((-2, -1, 2)^T) \text{ e } V_{18} = \text{Span}((2, -2, 1)^T).$$

Abbiamo quindi 6 simmetrie: 3 rispetto alle 3 rette V_λ , $\lambda \in \{18, 45, 81\}$ e 3 rispetto ai 3 piani generati da due dei tre autovettori menzionati, e.g. $H_1 := \text{Span}((1, 2, 2)^T, (-2, -1, 2)^T)$.

- La quadrica $Q : \{x^T \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} x + 2 \begin{pmatrix} -1 & 2 & 3 \end{pmatrix} x + 1 = 0\}$ non ha centro, e il suo

asse è $\mathcal{A} = \text{Sol}(A^2 | -Ab) = \{(-\frac{1}{2} - t, t, \frac{3}{2})^T \mid t \in \mathbb{R}\}$.

Come visto in precedenza, gli autovalori di A sono 0, -2 e 2 , e gli autospazi corrispondenti sono $V_0 = \text{Span}((-1, 1, 0)^T)$, $V_{-2} = \text{Span}((0, 0, 1)^T)$ e $V_2 = \text{Span}((1, 1, 0)^T)$.

Possiamo quindi definire 2 simmetrie, una per ciascun piano π_λ ($\lambda \in \{-2, 2\}$) ortogonale a V_λ e passante per un punto $v_0 \in \mathcal{A}_Q$ (e.g. $v_0 = (-\frac{1}{2}, 0, \frac{3}{2})^T$):

$$\pi_2 : x + y = -\frac{1}{2} \quad \pi_{-2} : z = \frac{3}{2}$$

6.3.2 Simmetrie di rotazione

Sia $v_0 + W = S \triangleleft \mathbb{R}^n$ un sottospazio affine di dimensione $\dim W \geq 2$, allora in questo caso possiamo “ruotare” W attorno all’asse $v_0 + W^\perp$ tramite una matrice ortogonale $M \in \mathcal{M}_{\mathbb{R}}(n)$ tale che $Mw = w$ per $w \in W^\perp$. Più precisamente, consideriamo

$$R : \mathbb{R}^n \rightarrow \mathbb{R}^n, v \mapsto v_0 + M(v - v_0)$$

Sia $w \in \mathbb{R}^n$ e consideriamo $w_1 := \pi_W(w)$ e $w_2 := \pi_{W^\perp}(w)$, quindi $w = w_1 + w_2$. Applichiamo ora R al punto $v_0 + w$:

$$R(v_0 + w_1 + w_2) = v_0 + M(w_1 + w_2) = v_0 + Mw_1 + Mw_2 = v_0 + Mw_1 + w_2$$

In altre parole, R ruota la parte in W , e tiene fermi i punti di $v_0 + W^\perp$ che è l’*asse di rotazione*.

Lasciamo come esercizio per il lettore la verifica che l’isometria R è una simmetria della quadrica Q .

Sia $Q = \{x \in \mathbb{R}^n \mid x^T A x + b^T x + c = 0\} \subset \mathbb{R}^n$ una quadrica e supponiamo che A abbia un autovalore λ con $\text{ma}(\lambda) > 1$, allora Q è una *quadrica di rotazione*, visto che possiamo “ruotare” l’autospazio relativo a λ tramite una rotazione ($v_0 \in \mathcal{A}_Q$)

$$R : \mathbb{R}^n \rightarrow \mathbb{R}^n, v \mapsto v_0 + M(v - v_0)$$

dove $M \in \mathcal{M}_{\mathbb{R}}(n)$ è una matrice ortogonale tale che $Mv = v$ per $v \in V_\lambda^\perp$. Quindi, R ruota la quadrica Q attorno all’*asse di rotazione* $v_0 + V_\lambda^\perp$.

Vediamo cosa succede nel caso $n = 3$. Sia $Q = \{x \in \mathbb{R}^3 \mid x^T A x + b^T x + c = 0\} \subset \mathbb{R}^3$ una quadrica e supponiamo che A abbia un autovalore λ con $\text{ma}(\lambda) > 1$ e supponiamo per semplicità $\lambda \neq 0$. Ci sono 2 possibilità per la forma canonica metrica di Q :

i) $\lambda x_1^2 + \lambda x_2^2 + \lambda_3 x_3^2 + c' = 0$ se è a centro; ii) $\lambda x_1^2 + \lambda x_2^2 + 2||b'||x_3 = 0$ se non è a centro.

Dividendo per λ e riordinando, otteniamo

i) $x_1^2 + x_2^2 = -\frac{c'}{\lambda} - \frac{\lambda_3}{\lambda} x_3^2 = 0$; ii) $x_1^2 + x_2^2 = -\frac{2||b'||}{\lambda} x_3 = 0$;

Intersecando la quadrica con piani affini $x_3 = \alpha \in \mathbb{R}$ possiamo ottenere o l’insieme vuoto o una circonferenza.

Esempio. • La quadrica $Q : x^T \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} x + 2 \begin{pmatrix} -3 & -3 & -3 \end{pmatrix} x + 11 = 0$ ha infiniti centri: $\text{Sol}(A| - b) = \{(3 - t, t, \frac{3}{2})^T \mid t \in \mathbb{R}\} = \mathcal{A}_Q$.

Abbiamo visto che A gli autovalori di A sono 0 e 2 (doppio), quindi l’asse di rotazione è $v_0 + V_2^\perp$, dove $v_0 \in \mathcal{A}_Q$.

D’altra parte A è simmetrica quindi $V_2 \perp V_0$ e quindi l’asse di rotazione $v_0 + V_0$ coincide con l’asse della quadrica.

- La quadrica $Q : x^T \begin{pmatrix} 50 & 16 & -32 \\ 16 & 26 & -16 \\ -32 & -16 & 50 \end{pmatrix} x - 36 = 0$ è a centro, e la matrice $A =$

$\begin{pmatrix} 50 & 16 & -32 \\ 16 & 26 & -16 \\ -32 & -16 & 50 \end{pmatrix}$ ha 2 autovalori distinti: 18 (doppio) e 90 e gli autospazi corrispondenti sono $V_{18} = \text{Span}((1, 2, 2), (2, -2, 1))$, $V_{90} = \text{Span}((-2, -1, 2)^T)$.

Quindi Q ha una simmetria di rotazione attorno all'asse $v_0 + V_{18}^\perp = v_0 + V_{90}$, dove v_0 è un centro; in questo caso $v_0 = 0$.

6.4 Nomenclatura di coniche e quadriche in $\mathbb{R}^3$

Per quanto visto, ogni quadrica ha una forma canonica metrica/affine.

Nelle due tabelle seguenti elenchiamo tutte le possibilità per le forme canoniche di coniche e quadriche in $\mathbb{R}^3$. Usiamo la convenzione che gli autovalori λ_j riportati sono non nulli, mentre i restanti sono nulli. Inoltre, l'etichetta (D) significa che la conica/quadrica è degenera.

Coniche

Forma canonica metrica	Forma canonica affine	Tipo affine	A centro?
$\lambda_1 x_1^2 + \lambda_2 x_2^2 + 1 = 0$	$\lambda_1, \lambda_2 > 0 \quad x_1^2 + x_2^2 + 1 = 0$	$\emptyset$ (ellisse immaginaria)	Sì
	$\lambda_1, \lambda_2 < 0 \quad x_1^2 + x_2^2 - 1 = 0$	ellisse	Sì
	$\lambda_1 > 0, \lambda_2 < 0$ $x_1^2 - x_2^2 + 1 = 0$	iperbole	Sì
(D) $\lambda_1 x_1^2 + \lambda_2 x_2^2 = 0$	$\lambda_1 \lambda_2 > 0 \quad x_1^2 + x_2^2 = 0$	punto $(0, 0)$ (rette immaginarie incidenti $x_1 = \pm i x_2$)	Sì
	$\lambda_1 \lambda_2 < 0 \quad x_1^2 - x_2^2 = 0$	rette incidenti $x_1 = \pm x_2$	Sì
(D) $\lambda_1 x_1^2 + 1 = 0$	$\lambda_1 > 0 \quad x_1^2 + 1 = 0$	$\emptyset$ (rette immaginarie parallele $x_1 = \pm i$)	Sì
	$\lambda_1 < 0 \quad x_1^2 - 1 = 0$	rette parallele $x_1 = \pm 1$	Sì
(D) $x_1^2 = 0$	$x_1^2 = 0$	retta $x_1 = 0$ (retta doppia)	Sì
$\lambda_1 x_1^2 + 2\lambda_2 x_2 = 0$	$x_1^2 + 2x_2 = 0$	parabola	No

Quadriche in $\mathbb{R}^3$

Forma canonica metrica	Forma canonica affine	Tipo affine	A centro?
$\lambda_1 x_1^2 + \lambda_2 x_2^2 + \lambda_3 x_3^2 + 1 = 0$	$\lambda_1, \lambda_2, \lambda_3 > 0$ $x_1^2 + x_2^2 + x_3 + 1 = 0$	$\emptyset$ (ellissoide immaginario)	Sì
	$\lambda_1, \lambda_2, \lambda_3 < 0$ $x_1^2 + x_2^2 + x_3^2 - 1 = 0$	ellissoide	Sì
	$\lambda_1, \lambda_2 > 0, \lambda_3 < 0$ $x_1^2 + x_2^2 - x_3^2 + 1 = 0$	iperboloide a due falde (iperboloide ellittico)	Sì
	$\lambda_1 > 0, \lambda_2, \lambda_3 < 0$ $x_1^2 - x_2^2 - x_3^2 + 1 = 0$	iperboloide ad una falda (iperboloide iperbolico)	Sì
(D) $\lambda_1 x_1^2 + \lambda_2 x_2^2 + \lambda_3 x_3 = 0$	$\lambda_1, \lambda_2, \lambda_3 > 0$ o $\lambda_1, \lambda_2, \lambda_3 < 0$ $x_1^2 + x_2^2 + x_3^2 = 0$	punto $(0, 0, 0)$	Sì
	$\lambda_1, \lambda_2 > 0, \lambda_3 < 0$ o $\lambda_1 > 0, \lambda_2, \lambda_3 < 0$ $x_1^2 + x_2^2 = x_3^2$	cono (a due falde)	Sì
(D) $\lambda_1 x_1^2 + \lambda_2 x_2^2 + 1 = 0$	$\lambda_1, \lambda_2 > 0$ $x_1^2 + x_2^2 + 1 = 0$	$\emptyset$ (cilindro immaginario)	Sì
	$\lambda_1, \lambda_2 < 0$ $x_1^2 + x_2^2 - 1 = 0$	cilindro ellittico	Sì
	$\lambda_1 > 0 > \lambda_2$ $x_1^2 - x_2^2 + 1 = 0$	cilindro iperbolico	Sì
(D) $\lambda_1 x_1^2 + \lambda_2 x_2^2 = 0$	$\lambda_1 \lambda_2 > 0$ $x_1^2 + x_2^2 = 0$	retta doppia $\{(0, 0, \alpha)^T\}$	Sì
	$\lambda_1 \lambda_2 < 0$ $x_1^2 - x_2^2 = 0$	piani incidenti $\{(\alpha, \alpha, \beta)^T\}, \{(\alpha, -\alpha, \beta)^T\}$	Sì
(D) $\lambda_1 x_1^2 + 1 = 0$	$\lambda_1 > 0$ $x_1^2 + 1 = 0$	$\emptyset$ (piani immaginari paralleli $x_1 = \pm i$)	Sì
	$\lambda_1 < 0$ $x_1^2 = 1$	piani paralleli $x_1 = \pm 1$	Sì
(D) $\lambda_1 x_1^2 = 0$	$x_1^2 = 0$	piano doppio $x_1 = 0$	Sì
$\lambda_1 x_1^2 + \lambda_2 x_2^2 + 2 b' x_3 = 0$	$\lambda_1 \lambda_2 > 0$ $x_1^2 + x_2^2 + 2x_3 = 0$	paraboloide ellittico	No
	$\lambda_1 \lambda_2 < 0$ $x_1^2 - x_2^2 + 2x_3 = 0$	paraboloide iperbolico	No
(D) $\lambda_1 x_1^2 + 2 b' x_2 = 0$	$x_1^2 + 2x_2 = 0$	cilindro parabolico	No