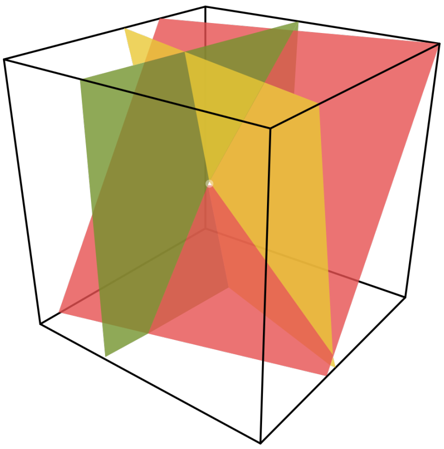


Geometria e Algebra lineare

Riassunto da: ""

corso A
Università degli studi di Torino, Torino
settembre 2023

Indice

1 Sistemi di equazioni lineari

Prima di parlare di sistemi definiamo cosa si intende per *equazione lineare*: un'equazione lineare è un'uguaglianza del tipo

$$a_1 x_1 + a_2 x_2 ... a_n x_n = b$$

espressa nelle incognite x_1, x_2,... x_n. Un'equazinone di questo genere ha come soluzione una n-upla di numeri reali che sostituiti al posto delle incognite rende vera l'uguaglianza (la *risoluzione* dell'equazione consiste nel trovare questa n-upla).

esempio

Definiamo un'equazione $** : x_1 - x_2 + 2x_3 = 4$ con $x_1 = x_2 - 2x_3 + 4$

L'insieme delle soluzioni di $*$ lo indichiamo con $S_{(*)}$ $S_{(*)} = \left\{ \left(x_2 - 2x_3 + 4x_2 x_3 \right) : x_2, x_3 \in \mathbb{R} \right\}$ In cui x_2 e x_3 sono i parametri liberi che variano.

esempio

Utilizzando un'altra equazione $** : 2x - 3y = 0$

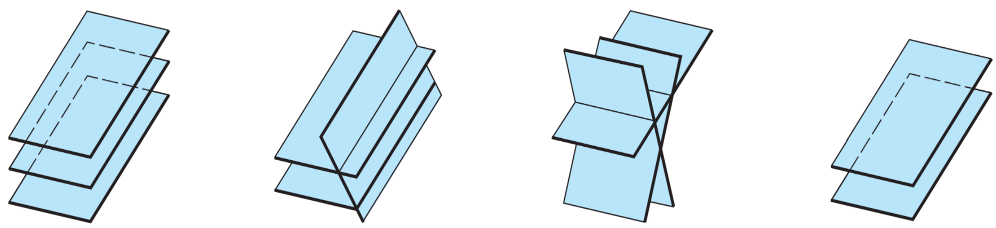
L'insieme delle sue soluzioni sarà: $S_{(*)} = \left\{ \left(xy \right) : x, y \in \mathbb{R} \right\}$ oppure tramite un parametro t per il quale $\left\{ x = ty = \frac{2}{3}t, \quad t \in \mathbb{R} \right.$ $\quad S_{(*)} = \left\{ (t, \frac{2}{3}t) : t \in \mathbb{R} \right\}$

Definiamo invece un *sistema lineare* di r equazioni lineari in n incognite $x_1, x_2 \dots x_n$ una struttura del tipo:

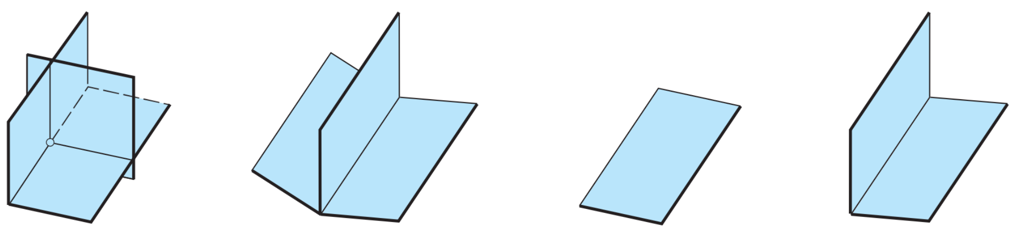
$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{2,1}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots a_{r1}x_1 + a_{r2}x_2 + \cdots + a_{rn}x_n = b_r \end{cases}$$

i coefficienti sono espressi nella forma a_{ij} per agevolarne il riconoscimento all'interno del sistema. Il pedice i indica l'indice di riga, il pedice j è l'indice di colonna. I termini noti b presentandosi una sola volta per riga hanno solo l'indice di riga. Se i termini noti sono tutti nulli il sistema si dirà ***omogeneo***

Diremo soluzione del sistema una n-upla di numeri reali $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ che risolve ciascuna delle equazioni del sistema. Il sistema si dice ***compatibile*** se ammette soluzioni (altrimenti ***incompatibile***). Due sistemi sono ***equivalenti*** se hanno lo stesso insieme di soluzioni.



Nessuna soluzione



Una soluzione

∞ siluzioni
(1 par. libero)

∞ siluzioni
(3 par. liberi)

∞ siluzioni
(1 par. libero)

Teorema di Rouchè-Capelli

Un sistema lineare $AX = B$ con $A \in \mathbb{R}^{r,n}$, $X \in \mathbb{R}^{n,p}$, $B \in \mathbb{R}^{r,p}$ è compatibile se e solo se il rango della matrice dei coefficienti coincide con il rango della matrice completa. Se il sistema è compatibile, le soluzioni dipendono da $p \cdot (n - \text{rk}A)$ parametri liberi.

$$\text{rk}\left(A|\overline{b}\right) = \text{rk}(A) \quad \Rightarrow \quad \text{sistema compatibile}$$

Poiché si opera solo sui coefficienti e non sulle incognite, i calcoli su essi risultano facilitati tramite l'utilizzo di tabelle (matrici). Un sistema quindi, nella sua forma matriciale (completa perché contiene anche i termini noti) il sistema si presenta così:

$$\left(A|\overline{b}\right) = \left(\begin{array}{ccc|c} a_{1,1} & \cdots & a_{1,n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{r,1} & \cdots & a_{r,n} & b_r \end{array}\right)$$

In questa forma la matrice è scomponibile e riscrivibile come il prodotto scalare tra il vettore dei coefficienti $\overline{a}$ e il vettore delle incognite $\overline{x}$:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{r1} & a_{r2} & \cdots & a_{rn} \end{pmatrix} \begin{pmatrix} x_1 x_2 \\ \vdots x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_r \end{pmatrix}$$

1.1 Sistemi omogenei

Particolarità dei sistemi omogenei è il fatto che operando sulle righe, non si va ad alterare sulla colonna di zeri. Ogni sistema omogeneo ricade in due possibili scenari:

1. Ha solo la soluzione banale;
2. Ha altre infinite soluzioni oltre quella banale.

Teorema: parametri liberi

Se un sistema lineare omogeneo ha n incognite e nell sua forma ridotta la sua matrice completa ha $\text{rank}A = n$ (nessuna riga nulla), allora

$$\text{parametri liberi } = n - \text{rank}A.$$

Se conosco una soluzione x_0 di Σ, sommandoci una qualsiasi soluzione del suo sistema associato Σ_0 Un sistema lineare è compatibile se e solo se il rango della matrice dei coefficienti coincide con il rango della matrice completa.

Diciamo di avere un sistema molto semplice a un'equazione è il suo sistema omogeneo associato:

$$\Sigma : 3x - y = 5 \qquad \to y = 3x - 5$$

$$\Sigma_0 = 3x - y = 0 \qquad \to y = 3x.$$

Le soluzioni di Σ saranno:

$$S(\Sigma) : \left\{ \begin{pmatrix} x \\ 3x - 5 \end{pmatrix} : x \in \mathbb{R} \right\}$$

di Σ_0 invece:

$$S(\Sigma_0) : \left\{ x \begin{pmatrix} 1 \\ 3 \end{pmatrix} : x \in \mathbb{R} \right\}$$

Le soluzioni di Σ possono essere riscritte come una soluzione particolare (prendiamo quella con $x = 0$) sommata a tutte le soluzioni del sistema omogeneo associato:

$$S(\Sigma) : \begin{pmatrix} 0 \\ -5 \end{pmatrix} + x \cdot \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

dimostrazione

Iniziamo provando che la somma di due soluzioni di un sistema omogeneo rimane una soluzione:

$$A\overline{x} = \overline{0} \qquad (\Sigma_0)$$

$$A\overline{y} = \overline{0} \qquad (\Sigma_0)$$

$$\overline{x}, \overline{y} \in S(\Sigma_0)$$

$$A\left(\overline{x} + \overline{y}\right) = A\overline{x} + A\overline{y} = \overline{0} + \overline{0} = \overline{0}$$

$$\Rightarrow \quad \overline{x} + \overline{y} \in S(\Sigma_0)$$

Ovvio anche che $\lambda\overline{x}$ o $\lambda\overline{y}$ entrambi $= 0 \quad \forall \lambda \in \mathbb{R}$.
Proviamo poi che la somma di due soluzioni di Σ *non è mai* soluzione di Σ

$$A\overline{x} = \overline{b} \qquad (\Sigma)$$

$$A\overline{y} = \overline{b} \qquad (\Sigma)$$

$$\overline{x}, \overline{y} \in S(\Sigma)$$

$$A\left(\overline{x} + \overline{y}\right) = A\overline{x} + A\overline{y} = \overline{b} + \overline{b} = 2\overline{b}$$

Infine proviamo che una soluzione di Σ+ una qualsiasi soluzione di Σ_0 è sempre una soluzione di Σ:

$$\overline{x} \in S(\Sigma), \quad \overline{y} \in S(\Sigma_0) \quad \Rightarrow \quad A\left(\overline{x} + \overline{y}\right) = A\overline{x} + A\overline{y} = \overline{b} + \overline{0}$$

$$= \overline{b}$$

2 Matrici

Definiamo una matrice di r righe e n colonne con $a_{ij} \in \mathbb{R}; i = 1,\dots,r; j = 1,\dots,n$ definita nello spazio $\mathbb{R}^{rn}$ in questo modo:

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{r,1} & a_{r,2} & \cdots & a_{r,n} \end{bmatrix}$$

Esistono matrici **quadrate** se hanno stesso numero di righe e di colonne ($\mathbb{R}^{nn}$), **diagonali** se tutti gli elementi sono zeri tranne quelli sulla diagonale maggiore (matrice *unità* se la diagonale contiene solo 1), **nulle** se tutti gli elementi sono zeri, **riga** se hanno una riga sola, **colonna** se hanno una colonna sola.

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

Definizione: matrice ridotta

Una matrice si dice **ridotta** se in ogni sua riga *non nulla* esiste un elemento sotto al quale ci sono solo zeri, questo elemento viene chiamato *pivot*. Se una matrice è ridotta chiameremo *rango della matrice* il numero di righe non nulle in essa: $rk(A) \leq min\{r, n\}$.

Definizione: matrice a scala

Chiameremo *primo pivot* il primo pivot nella prima riga partendo da sinistra. Una matrice si dice **a scala se *è ridotta*** e se la riga R_i è tutta fatta di zeri e, quindi, anche la riga R_j per ogni $j > i$; ovvero:

$$R_i = \overline{0} \quad \Rightarrow \quad R_j = \overline{0} \quad \forall j > i$$

Se la riga $R_i \neq \overline{0}$ il *primo pivot* di R_i è strettamente a destra del primo pivot di R_{i-1}.

Proposizione Il rango di una matrice A non supera mai il minimo fra il numero di righe e di colonne.

$$rk(a) \leq min\{r, n\}$$

dimostrazione

Sia B' la riduzione a scala di B.
Sappiamo che r_2 di B' comincia con almeno uno zero, r_3 con almeno 2 zeri, r_4 con almeno 3 zeri e così via.

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ 0 & 0 & a_{33} & a_{3n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{j1} & a_{j2} & \dots & a_{jn} \\ 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\Rightarrow \quad R_{n+1} \text{è tutta di zeri} \quad \Rightarrow \quad R_j \text{è tutta di zeri} \forall j \geq n+1$$

Se il rango è il numero di righe non nulle, e le righe dalla $n+1$ in poi sono tutte nulle, allora il rango è sicuramente minore o uguale a n. ■

2.1 Operazioni tra matrici

E' ammessa la somma tra matrici dello stesso ordine $\mathbb{R}^{rn}$ e sono ammesse la proprietà commutativa, associativa, esistenza dellelemento neutro ed esistenza dellopposto. Il prodotto $A \cdot B$ è ammesso se il numero di righe della prima è uguale al numero di colonne della seconda, ovvero se $A \in \mathbb{R}^{rn}$ e $B \in \mathbb{R}^{np}$. Per il prodotto sono valide la proprietà associativa, la distributiva del prodotto rispetto alla somma.

Prodotto come combinazione lineare

Un altro modo per descrivere il prodotto tra matrici è come *combinazione lineare di vettori colonna*:

$$\begin{bmatrix} -1 & 3 & 2 \\ 1 & 2 & -3 \\ 2 & 1 & -2 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ -9 \\ -3 \end{bmatrix}$$

può anche essere scritto come:

$$2\begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} - 1\begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} + 3\begin{bmatrix} 2 \\ -3 \\ -2 \end{bmatrix} = \begin{bmatrix} 1 \\ -9 \\ -3 \end{bmatrix}$$

La trasposta di una matrice

Definizione: matrice trasposta

Data una matrice $A \in \mathbb{R}^{r,n}$, si dice trasposta di A (${}^tA$) la matrice che si ottiene scambiano le righe con le colonne di A: Se $A = (a_{ij})$, ${}^tA = (a_{ji})$.

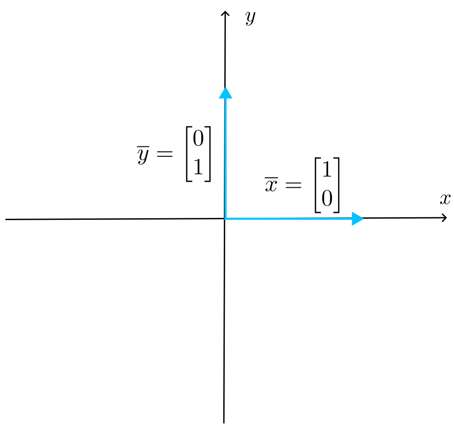
$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \end{bmatrix} \qquad {}^tA = \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 3 & 2 \end{bmatrix}$$

Proprietà delle matrici trasposte:

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

Interpretazione geometrica delle matrici

Dopo aver parlato di vettori, approfondiamo il concetto di *matrice* e il suo comportamento come *spazio vettoriale*. Più in particolare vedremo il prodotto tra matrici come trasformazioni lineari dello spazio vettoriale (linearmente perché nessuna linea viene curvata e l'origine rimane fissata). Questa interpretazione di una matrice rende i conti più facili ed intuitivi.
Diciamo di avere due vettori giacenti sui due assi:

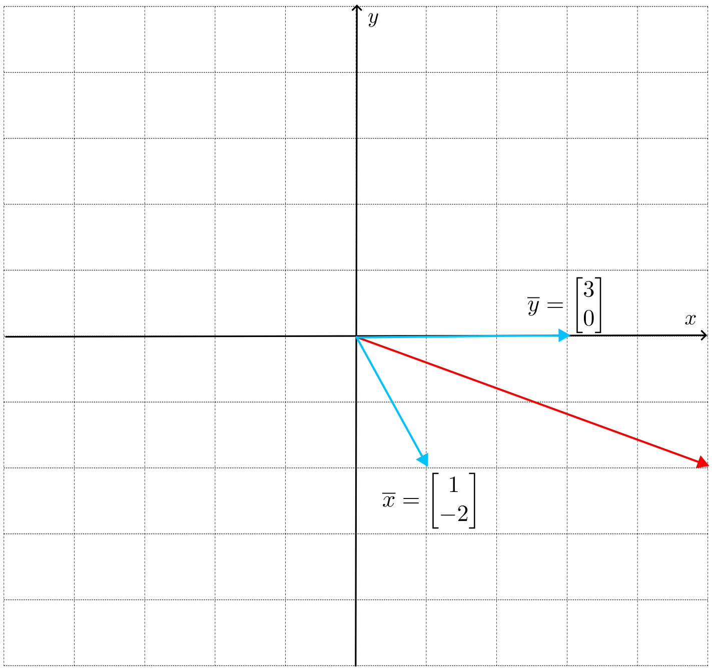


Diciamo ora di avere una matrice A che descrive dove questi due vettori cadono a seguito della trasformazione da essa descritta; la matrice A basta per descrivere dove cadrà ogni vettore (x,y).

$$x = \begin{bmatrix} 1 & 3 \\ 2 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 3 \\ -2 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = x\begin{bmatrix} 1 \\ -2 \end{bmatrix} + y\begin{bmatrix} 3 \\ 0 \end{bmatrix}$$

Se, come abbiamo detto prima, le linee non vengono deformate, il vettore che nel primo grafico sarebbe stato (1,1), nel secondo è intuitivo pensare che ora sia (4,-2); Il prodotto tra le matrici lo conferma.



Possiamo arrivare alla conclusione che ogni matrice può essere interpretata come una trasformazione dello spazio, a prescindere dall'ordine della matrice.

Prodotto come composizione di trasformazioni Se applichiamo più trasformazioni consecutive, quindi tramite più matrici, interpretiamo la composizione di queste trasformazioni come il prodotto tra le matrici.

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

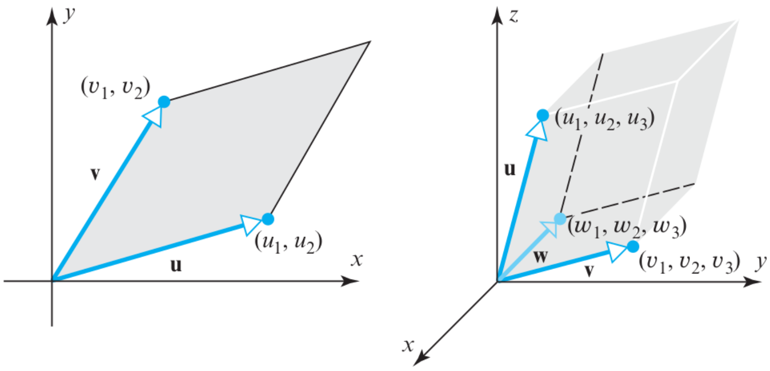
L'ordine di composizione va letto da destra verso sinistra: viene eseguita prima la blu, poi la rossa (cosa tipica delle notazione delle funzioni: $f(g(x))$).
Pensando in questi termini, prodotto come composizione di trasformazioni, comprendiamo perché $AB \neq BA$; e la proprietà associativa diventa chiara e logica: $A(BC) = (AB)C$ in quanto l'ordine delle trasformazioni rimane invariato.

2.2 Determinante

Interpretazione geometrica del determinante

Il determinante di una matrice geometricamente rappresenta il fattore di "stretching" di un'area, o volume tridimensionale o n-dimensionale (qualunque cosa sia).

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \rightarrow det A = 3$$



Calcolo del determinante

Il determinante di una matrice è una funzione $det : \mathbb{R}^{nn} \Longrightarrow \mathbb{R}$ che verifica queste due proprietà:

1. Se a è un numero reale, ossia una matrice di ordine uno quadrata, allora $det(a) = a$.
2. Se A $= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ allora $det(A) = a_{11}a_{22} - a_{12}a_{21}$

Dallo sviluppo del caso due notiamo che compaiono due addendi ciascuno dei quali è il prodotto di due fattori. I due fattori nei due prodotti di iniziano entrambi uno con un 1 e uno con un 2 per poi seguire con le *permutazioni di* $(1,2) : (1,2),(2,1)$. Perché il secondo prodotto ha un $-$ davanti? Il segno è dettato dalla *parità* della permutazione, ovvero: per arrivare alla coppia $(1,2)$ si devono attuare degli scambi, se il numero di scambi è pari, il segno non cambia, se gli scambi sono dispari, il segno cambia.

esempio

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \rightarrow det(A) = a_{11}a_{22}a_{33} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33}$$

$$= \sum_{\sigma} \epsilon(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} a_{3\sigma(3)}$$

Definizione: determinante

Il determinante di una matrice quadrata $A = (a_{ij})$ di ordine n è dato da:

$$\sum_{\sigma} \epsilon(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \dots a_{n\sigma(n)}$$

dove σ è una qualsiasi permutazione dei numeri $1,2,\dots,n$ e $\epsilon(\sigma)$ è il suo segno.

Come cambia il determinante dopo le 3 mosse?

1. $det^t(A) = det(A)$;
2. Se A' si ottiene scambiando due righe o due colonne di A, allora $det(A') = -det(A)$;

3. Se faccio moltiplico una riga per un numero reale λ allora $det(A^1) = \lambda^n det(A)$;
4. Se addiziono a una riga un multiplo di un'altra riga il determinante non cambia;
5. Una matrice con due righe o colonne uguali ha determinante nullo;
   (a) data la matrice $A \in \mathbb{R}^{n,n}$, $rk(A) = n \longleftrightarrow det(A) \neq 0$
   (b) analogamente $rk(A) < n \longleftrightarrow det(A) = 0$
6. $det(A+B) \neq det(A) + det(B) \quad \forall A, B \in \mathbb{R}^{n,n}$
7. **Teorema di Binet**: $det(AB) = det(A)det(B) \quad \forall A, B \in \mathbb{R}^{n,n}$;
8. $det(A^{-1}) = (det(A))^{-1}$;

dimostrazione determinante (2)

È conseguenza della definizione di determinante e del fatto che lo scambio di due righe comporta il cambiamento di segno di ciascuna permutazione. Per esempio, nel caso della matrice quadrata di ordine 2 si ha:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

Se scambio due righe:

$$\begin{vmatrix} a_{21} & a_{22} \\ a_{11} & a_{12} \end{vmatrix} = -a_{22}a_{11} + a_{21}a_{12}$$

■

dimostrazione determinante (5a)

Operando su una matrice A e la rendo A' a scala triangolare superiore. So allora che $det(A') = \lambda \cdot det(A)$ e $det(A) \neq 0 \Leftrightarrow det(A') \neq 0$. Quindi $a'_{ij} \neq 0 \quad \forall j \in \{1, \ldots, n\}$.
Cioè A' ha n righe non nulle, quindi $rk(A') = n$. ■

dimostrazione determinante (8)

Per il teorema di Binet:

$$AA^{-1} = I \qquad A^{-1} = \frac{I}{A} \qquad det(A^{-1}) = det(\frac{I}{A}) = \frac{1}{det(A)}$$

■

osservazioni su matrici inverse Se il determinante di una matrice è uguale a zero, significa che la matrice è non invertibile. Questo perché il determinante descrive, anche se non esplicitamente, il numero di soluzioni del sistema di equazioni associato.
Il determinante è infatti strettamente legato al **rango**: se il rango, ovvero il numero di righe non nulle, di una matrice $A \in \mathbb{R}^{r,n}$ è minore di n sappiamo che il determinante vale 0 e che di conseguenza il sistema *non può avere una singola soluzione*. Infatti se la matrice ha una riga nulla o più, le soluzioni saranno infinite e legate a uno o più parametri liberi.
Se il sistema associato alla matrice A non ha una singola soluzione è chiaro come non possa esistere una matrice A' inversa che soddisfi:

$$AA^{-1} = I \quad A^{-1} = \frac{I}{A}$$

L'equazione ha infatti una sola soluzione se e solo se A fosse unicamente definita.

2.3 Teoremi di Laplace

I teoremi di Laplace permettono di semplificare i conti nel calcolo del determinante di una matrice $n \times n$ a conti di un determinante $(n-1) \times (n-1)$. I conti vengono semplificati perché si procede a scegliere un elemento a_{ij} nella matrice (vedremo perché di solito è uno in una riga o colonna con tanti zeri) , "nascondendo" tutti gli elementi della riga e colonna del nostro candidato e andremo a calcolare il determinante della matrice "rimanente", questo determinante lo chiameremo **minore** di a_{ij} e lo indichiamo con M_{ij}. Ora serve definire il **cofattore**; il cofattore di a_{ij} è il numero A_{ij} definito dalla formula:

$$A_{ij} = (-1)^{ij} \cdot M_{ij}$$

Vediamo come il fattore $(-1)^{ij}$ da segno positivo o negativo se la posizione di a_{ij} è pari o dispari (a_{11} è pari, a_{12} è dispari...).

Primo Teorema di Laplace

Fissata la riga i-esima, il determinante di una matrice quadrata $A \in \mathbb{R}^{n,n}$ è dato dalla somma di tutti i prodotti tra gli elementi della riga e i rispettivi cofattori (questo metodo funziona anche con le colonne) :

$$\det(A) = \sum_{j=1}^{n} a_{ij} A_{ij}$$

Secondo Teorema di Laplace

In una matrice quadrata $A \in \mathbb{R}^{n,n}$ la somma dei prodotti tra gli elementi di una riga (o colonna) e i cofattori *di una riga parallela* è zero:

$$\begin{aligned} 0 &= \sum_{k=1}^{n} a_{ik} A_{jk} \\ &= \sum_{h=1}^{n} a_{hi} A_{hj} \qquad i \neq j \end{aligned}$$

verifica

È conseguenza evidente della proprietà (2) del determinante secondo la quale *se scambio due righe o colonne a una matrice allora il suo determinante cambia di segno.* Si puó interpretare come lo sviluppo del determinante di un matrice in cui, nel primo caso, la riga j-esima coincide con la riga i-esima e nel secondo caso, la colonna j-esima coincide con la riga i-esima.

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

Se per esempio scegliamo di moltiplicare gli elementi della prima riga per i complementi della seconda abbiamo:

$$\begin{aligned} & a\begin{vmatrix} b & c \\ h & i \end{vmatrix} - b\begin{vmatrix} a & c \\ g & i \end{vmatrix} + c\begin{vmatrix} a & b \\ g & h \end{vmatrix} \\ &= abi - ach - bai + bcg + cah - cbg = 0 \end{aligned}$$

2.4 Matrici inverse

Definizione: matrice invertibile

Una matrice $A \in \mathbb{R}^{n,n}$ si dice invertibile se $\exists$ una matrice X tale che $AX = XA = I$.

Proprietà generali delle matrici inverse

1. Se esiste una matrice inversa allora questa è univocamente determinata e la chiamo A^{-1};
2. $(AB)^{-1} = B^{-1}A^{-1}$;
3. Se A, B sono invertibili non è detto che lo sia $A + B$;
4. $({}^tA)^{-1} = {}^t(A^{-1})$.

dimostrazione (1)

Supponiamo che X e X' soddisfino:

$$XA = I = AX$$

$$X'A = I = AX'$$

$$XAX = \begin{matrix} (XA)X' = IX' = X' \\ X(AX') = XI = X \end{matrix} \rightarrow XA = AX'$$

Abbiamo dimostrato che se esiste una X inversa a sinistra per A ed esiste una X' inversa a esa per A, allora $X = X'$ e quindi A è invertibile e X è la sua inversa.

dimostrazione (2)

Vedo se la candidata ad inversa $B^{-1}A^{-1}$ soddisfa le proprietà richieste:

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

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

■

Teorema: matrice inversa

Sia A una matrice quadrata di ordine n, se $\det(A) \neq 0$ allora esiste linversa di A ed è:

$$A^{-1} = \frac{1}{\det(A)}\text{adj}(A)$$

dimostrazione

Dai teoremi di Laplace so che:

$$\sum_{j=1}^{n} a_{ij}A_{kj} = \begin{cases} \det A & \text{se} \quad i = k \\ 0 & \text{se} \quad i \neq k \end{cases}$$

ovvero che la somma dei prodotti tra tutti gli elementi di una riga di A e i rispettivi cofattori è uguale o a 0 o al determinante di A.
Ovvero che il prodotto tra la matrice A e la trasposta della matrice dei cofattori di A (adj$A$) si può scrivere come:

$$A \cdot \text{adj}A = \begin{bmatrix} \det A & 0 & \dots & 0 \\ 0 & \det A & \dots & 0 \\ \vdots & 0 & \ddots & 0 \end{bmatrix} = \det A \cdot I$$

Possiamo notare quindi che dopo le opportune operazioni ci si riconduce alla formula iniziale. ■

Calcolo della matrice inversa 1

Il primo metodo consiste nello svolgimento di un'equazione matriciale:

$$AX = I$$

Che si risolve come:

$$(A|I) = \left[\begin{array}{cccc|cccc} a_{11} & a_{12} & \dots & a_{1n} & 1 & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2n} & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} & 0 & 0 & \dots & 1 \end{array}\right]$$

Calcolo della matrice inversa 2

Possiamo calcolare la matrice inversa anche a partire dalla nozione di determinante dopo aver parlato dei teoremi di Laplace.

Definizione: matrice aggiunta

Si dice **matrice aggiunta** di A la trasposta della matrice contentente i *cofattori* di A:

$$\text{Adj}(A)_{ij} = [A_{ij}]$$

Per esempio:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ -1 & 2 & 5 \\ 0 & 1 & 2 \end{bmatrix} \qquad \text{Adj}(A) = \begin{bmatrix} -1 & -1 & 4 \\ 2 & 2 & -8 \\ -1 & -1 & 4 \end{bmatrix}$$

I teoremi di Laplace più la matrice adiacente ci permettono di determinare in modo esplicito la formula dell'inversa.

2.5 Teorema di Cramer

Subito dopo aver descritto un nuovo modo per calcolare la matrice inversa vediamo come può tornare utile nella risoluzione di sistemi lineari con n incognite e n equazioni.

$$A\overline{x} = \overline{b}$$

$$\overline{x} = A^{-1}\overline{b}$$

$$\overline{x} = \frac{1}{\det(A)}\text{adj}(A)\cdot\overline{b}$$

$$\overline{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \frac{1}{\det(A)} \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \dots & A_{nn} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \frac{1}{\det(A)} \begin{bmatrix} A_{11}b_1 & A_{12}b_2 & \dots & A_{1n}b_n \\ A_{21}b_1 & A_{22}b_2 & \dots & A_{2n}b_n \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1}b_1 & A_{n2}b_2 & \dots & A_{nn}b_n \end{bmatrix}$$

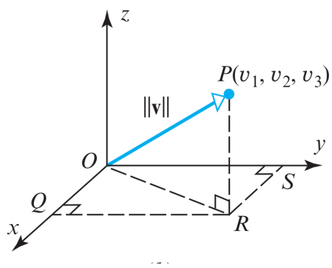
da cui:

$$\begin{aligned} x_i &= \frac{1}{\det(A)}(b_1A_{2i} + b_2A_{2i} + \cdots + b_nAni) \\ &= \frac{1}{\det(A)}\det(\overline{a}_1|\overline{a}_2\dots|\overline{b}|\dots|\overline{a}_n) \end{aligned}$$

3 Prodotto scalare e vettoriale

Prima di parlare di prodotto scalare è necessario introdurre due concetti fondamentali:

1. **Lunghezza** di un vettore che d'ora in poi chiameremo *norma*;
2. **Angolo** tra due vettori.

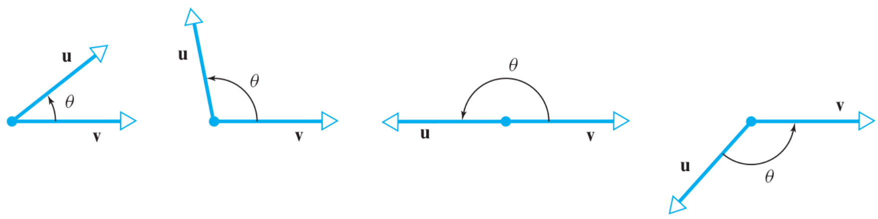


La norma del vettore è definita dalla formula:

$$\|v\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}.$$

Vedremo come saranno di estrema importanza i vettori di norma 1, o vettori unitari. Per esempio in V_3 i vettori unitari sono i, j, k. In generale per *normalizzare* un vettore basta dividerlo per la sua lunghezza, quindi per la sua norma:

$$u = \frac{v}{\|v\|}.$$



Per quanto riguarda l'angolo tra due vettori invece prenderemo in considerazione solo la parte compresa tra 0 e π.

Ora che sappiamo cosa sono norma di un vettore e angolo tra vettori possiamo parlare di **prodotto scalare**. E' infatti necessario introdurre un'operazione moltiplicativa "utile" per vettori in $\mathbb{R}^2$ e in $\mathbb{R}^3$.

Definizione: prodotto scalare

Il prodotto scalare (in V_3) $x \cdot y$ di due vettori x e y in V_3 è la funzione:

$$\cdot : \quad V_3 \times V_3 \quad \longrightarrow \quad \mathbb{R}.$$

così definita:

$$x \cdot y = \|x\| \|y\| \cos \hat{xy}.$$

Dalla definizione troviamo altre due espressioni di norma e angolo (notare come ora il concetto di angolo sia esteso a tutto $\mathbb{R}^n$):

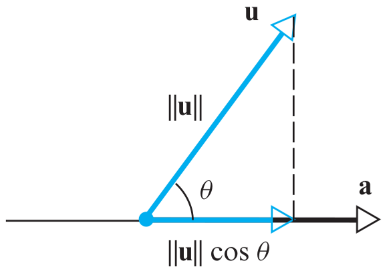
$$\|v\| = \sqrt{v \cdot v} \qquad \cos\theta = \frac{u \cdot v}{\|u\| \|v\|}.$$

3.1 Proprietà del prodotto scalare

- $x \cdot y = y \cdot x \quad \forall x, y \in V_3$;
- $(x + y) \cdot z = x \cdot z + y \cdot z$;
- $(\lambda x) \cdot z = \lambda x \cdot z = x \cdot (\lambda z)$;
- $x \cdot x \geq 0 \qquad = 0 \Longleftrightarrow x = 0$.

3.2 Interpretazione geometrica del prodotto scalare

Il prodotto scalare $\|a\|\|u\|\cos\theta$ non è altro che il prodotto della lunghezza di uno dei due vettori ($\|a\|$) per la proiezione ortogonale con segno dell'altro sul primo($\|u\|\cos\theta$).



Teorema: vettore proiezione ortogonale

Dati due vettori u e a non nulli il vettore proiezione ortogonale di u su a è:

$$p = \frac{u \cdot a}{\|a\|^2} a.$$

dimostrazione

Il mio obiettivo è quello di scrivere la proiezione di u su a in questa forma:

$$p = * \frac{a}{\|a\|}.$$

dove "*" indica la lunghezza della proiezione. Guardando la figura in alto sappiamo che la proiezione $\|p\| = \|u\|\cos\theta$. Quindi:

$$p = \|u\|\cos\theta \frac{a}{\|a\|}.$$

Per elimiare il coseno di theta risaliamo alla formula di prodotto scalare:

$$u \cdot a = \|u\|\|a\|\cos\theta \quad \longrightarrow \quad \|u\|\cos\theta = \frac{u \cdot a}{\|a\|}.$$

Quindi:

$$p = \frac{u \cdot a}{\|a\|^2} a.$$

Teorema di Pitagora generalizzato

Dati u e v vettori ortogonali tra loro in $\mathbb{R}^n$ con prodotto standard, allora

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2.$$

la dimostrazione è molto semplice, il termine $2u \cdot v$ vale zero.

3.3 Prodotto vettoriale

Definizione: prodotto vettoriale

Il prodotto vettoriale (in V_3) $x \cdot y$ di due vettori x e y in V_3 è la funzione:

$$\wedge : \quad V_3 \times V_3, \quad \longrightarrow \quad (x, y) \mapsto x \wedge y.$$

così definita:

$$x \wedge y = \|x \wedge y\| \|x\| \|y\| \sin \hat{xy}.$$

Il verso del vettore risultante dal prodotto vettoriale ha il verso che segue la *regola della mano destra*. Possiamo anche scrivere il prodotto scalare tramite lo sviluppo di determinanti in questo modo:

$$u \wedge v = \begin{vmatrix} i & j & k \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}$$

$$= \left(\begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix} i, - \begin{vmatrix} u_1 & u_3 \\ v_1 & v_3 \end{vmatrix} j, + \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix} k \right).$$

Proprietà del prodotto vettoriale

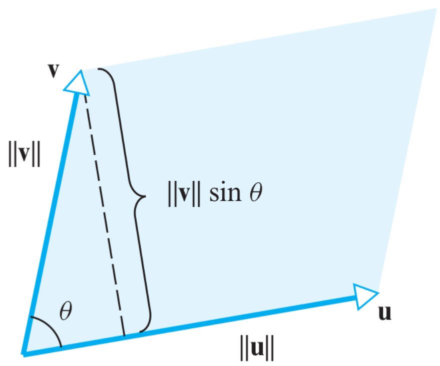
- $u \wedge v = -(v \wedge u)$;
- $u \wedge (v + w) = u \wedge v + u \wedge w$;
- $k(u \wedge v) = ku \wedge v = u \wedge kv$;
- $u \wedge u = 0$

Interpretazione geometrica del prodotto vettoriale

Dati u e v vettori in uno spazio tridimensionale, dall'identità di Lagrange sappaiamo che:

$$\begin{aligned} \|u \wedge v\|^2 &= \|u\|^2 \|v\|^2 - (u \cdot v)^2 \\ &= -\|u\|^2 \|v\|^2 \cos\theta^2 \\ &= \|u\|^2 \|v\|^2 \left(1 - \cos\theta^2\right) \\ &= \|u\|^2 \|v\|^2 \sin\theta^2 \end{aligned}$$

Da questo possiamo notare che il prodotto vettoriale può essere inteso anche come area del parallelogramma che ha come lati i due vettori.



3.4 Prodotto misto

Definizione: Prodotto misto

Dati due vettori u e v, allora
$$u \cdot (v \wedge w)\,.$$
è chiamato prodotto misto di u v e w.

Geometricamente il prodotto misto rappresenta $\frac{1}{6}$ del volume del tetraedro formato dai tre vettori.

4 Spazi Vettoriali

Definizione: Spazio Vettoriale

Si definisce **spazio vettoriale** sul campo dei numeri reali $\mathbb{R}$ un certo insieme V nel quale sono definite le seguenti operazioni:

1. somma $+: V \times V \longrightarrow V$.
2. prodotto $\cdot: \mathbb{R} \times V \longrightarrow V$.

Un gruppo $(V, \times)$ si dice *commutativo* (o Abeliano) se $v \times y = y \times x \quad \forall x, y \in V$

Proprietà della somma

1. commutativa: $x + y = y + x, \quad \forall x, y \in V$
2. associativa: $(x + y) + z = x + (y + z), \quad \forall x, y \in V$
3. esistenza dell'elemento neutro: $\exists 0 \in V : 0 + x = x + 0, \quad \forall x, y \in V$
4. esistenza dell'opposto: $\forall x \in V \exists -x \in V : x + (-x) = (-x) + x = 0$

Proprietà del prodotto

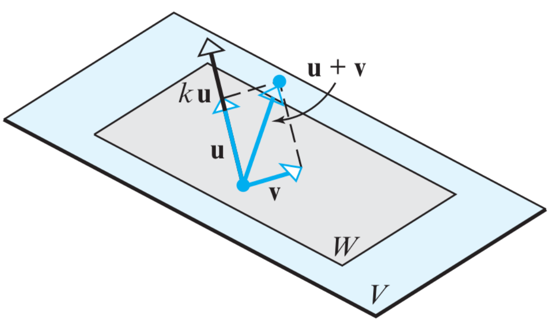
1. (diciamo) distributiva: $\lambda(x + y) = \lambda x + \lambda y, \quad \forall x, y \in V, \forall \lambda \in \mathbb{R}$
2. $(\lambda + \mu) \cdot \overline{x} = \lambda \overline{x} + \mu \overline{x}, \quad \forall x, y \in V, \forall \lambda \in \mathbb{R}$
3. $(\lambda \cdot \mu)\overline{x} = \lambda(\mu \cdot \overline{x}), \quad \forall x, y \in V$
4. $1 \cdot \overline{x} = \overline{x}, \quad \forall x, y \in V$

4.1 Sottospazi vettoriali

Definizione: sottospazio vettoriale

Sia V uno spazio vettoriale reale, $W \subseteq V$ è un sottospazio vettoriale di V se W è uno spazio vettoriale rispetto alle stesse operazioni di V, quindi rispetto alle operazioni di *somma* e *prodotto*.

- Se ho 2 elementi $\overline{x}, \overline{y} \in W \quad \Rightarrow \quad \overline{x} + \overline{y} \in W$;
- Se ho 2 elementi $\overline{x} \in W, \lambda \in \mathbb{R} \quad \Rightarrow \quad \lambda \overline{x} \in W$.



In figura vediamo come u e v siano contenuti in W ma la loro somma no.

Esempio fondamentale di sottospazio vettoriale

L'insieme delle soluzioni di un sistema lineare omogeneo di m equazioni in n incognite è un sottospazio vettoriale di $\mathbb{R}^n$. L'insieme delle soluzioni di

$$AX = O \qquad A \in \mathbb{R}^{m,n}, X \in \mathbb{R}^{n,1}, O \in \mathbb{R}^{m,1}$$

è detto *nullspace* e coincide con l'insieme

$$N(A) = \left\{X \in \mathbb{R}^n | AX = O\right\}$$

Dati X_1 e $X_2 \in N(A)$ si deve dimostrare che

$$\lambda X_1 \mu X_2 \in N(A) \quad \forall \lambda, \mu \in \mathbb{R}$$

che sviluppando

$$A\left(\lambda X_1 + \mu X_2\right) = \lambda A X_1 + \mu A X_2 = O.$$

+ **Somma di due sottospazi**

La somma di due sottospazi è il più piccolo sottospazio contenente l'unione dei due si esprime come l'insieme di tutti i vettori ottenuti dalla somma di vettori appartenenti ai sottospazi sommati.

$$W_1 + W_2 = \left\{\overline{x} \in V : \quad \overline{x} = \overline{y} + \overline{z} \qquad y \in W_1, \quad z \in W_2\right\}.$$

Quindi $W_1 + W_2$ contiene W_1 e contiene W_2 e $W_1 + W_2$ è sottospazio.

dimostrazione

Possiamo dire che è sottospazio se la somma tra ogni vettore ricade in esso così come il prodotto tra ogni vettore e uno scalare.
Prendiamo due vettori $\overline{x}$ e $\overline{y}$ entrambi $\in W_1 + W_2$:

$$\overline{x} = \overline{x}_1 + \overline{x}_2 \quad \overline{x}_1 \in W_1, \quad \overline{x}_2 \in W_2.$$

$$\overline{y} = \overline{y}_1 + \overline{y}_2 \quad \overline{y}_1 \in W_1, \quad \overline{y}_2 \in W_2.$$

$$\begin{aligned}(+) \Rightarrow \quad \overline{x} + \overline{y} &= \overline{x}_1 + \overline{x}_2 + \overline{y}_1 + \overline{y}_2 \\ &= \left(\overline{x}_1 + \overline{y}_1\right) + \left(\overline{x}_2 + \overline{y}_2\right) \quad \Rightarrow \quad \in W_1 + W_2\end{aligned}$$

$$\begin{aligned}(\cdot) \Rightarrow \quad \lambda\left(\overline{x} + \overline{y}\right) &= \lambda \overline{x}_1 + \lambda \overline{x}_2 + \lambda \overline{y}_1 + \lambda \overline{y}_2 \\ &= \lambda\left(\overline{x}_1 + \overline{y}_1\right) + \lambda\left(\overline{x}_2 + \overline{y}_2\right) \quad \Rightarrow \quad \in W_1 + W_2\end{aligned}$$

Definizione: Somma diretta

$(V, +, \cdot)$ spazio vettoriale, $W_1, W_2 \leq V$. Diciamo che la somma $W_1 + W_2$ è una somma diretta se ogni $\overline{x} \in W_1 + W_2$ si scrive in modo unico come $\overline{x}_1 + \overline{x}_2$ con $\overline{x}_1 \in W_1$ e $\overline{x}_2 \in W_2$.
La somma verrà scritta come:

$$W_1 \oplus W_2 = V.$$

Proposizione In V spazio vettoriale, $W_1, W_2 \leq V$. Allora:

$$W_1 \quad \text{e} \quad W_2 \quad \text{sono in somma diretta} \quad \Longleftrightarrow \quad W_1 \cap W_2 = \{\overline{0}\}.$$

dimostrazione

$$\Leftarrow \quad \text{se esiste un x con almeno 2 decomposizioni} \quad \Rightarrow \quad \exists \overline{z} \in W_1 \cap W_2, \quad \overline{z} \neq \overline{0}.$$

Prendo allora tale $\overline{x}$ che si scompone in due coordinate x e in due y:

$$\overline{x}_1 + \overline{x}_2 = \overline{x} = \overline{y}_1 + \overline{y}_2, \quad \overline{x}_1 \neq \overline{x}_2, \quad \overline{y}_1 \neq \overline{y}_2.$$

$$\overline{x}_1 + \overline{x}_2 = \overline{y}_1 + \overline{y}_2.$$

$$\overline{y}_1 - \overline{x}_1 = \overline{x}_2 - \overline{y}_2 \quad = \overline{z}.$$

Il vettore $\overline{z}$ è contenuto sia in W_1 sia in W_2, quindi $W_1 \cap W_2 \neq \overline{0}$.
Contronominale: se $W_1 \cap W_2 \neq \{\overline{0}\} \quad \Rightarrow \quad$ non ho unicità di scrittura

$$\Rightarrow \quad \text{Se} \quad \overline{z} \in W_1 \cap W_2, \quad \overline{z} \neq 0.$$

$$x \in W_1 + W_2.$$

$$\overline{x} = \overline{x}_1 + \overline{z} + \overline{x}_2 - \overline{z}.$$

■

∩ Intersezione di due sottospazi

L'intersezione di due sottospazi vettoriali W_1 e W_2 contiene tutti i vettori contenuti sia in W_1 sia in W_2.

Teorema: l'intersezione è sottospazio Immediata conseguenza delle definizioni di sottospazio vettoriale e di intersezione. Se W_1 e W_2 sono sottospazi allora lo deve essere anche $W_1 \cap W_2$.

4.2 Combinazione lineare

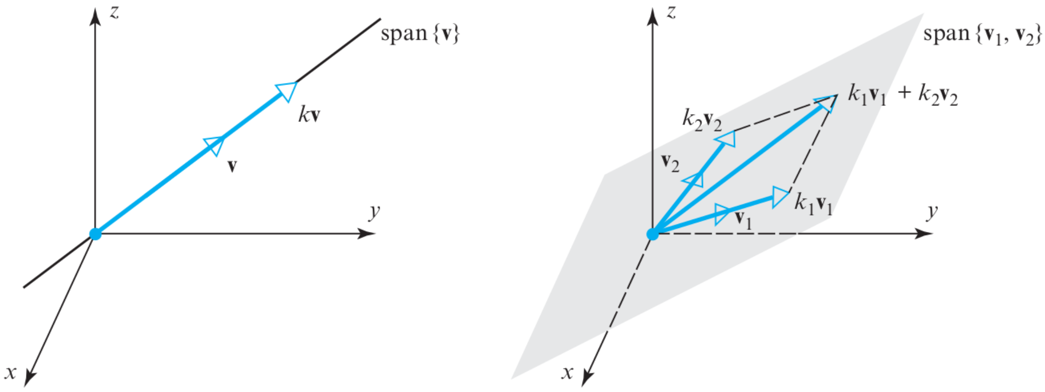
Definizione: combinazione lineare

Dato V spazio vettoriale, $\overline{v}_1, \dots, \overline{v}_n \in V$, una composizione lineare di $\overline{v}_1, \dots, \overline{v}_n$ è una strutura del tipo $\lambda_1 \overline{v}_1, \dots, \lambda_n \overline{v}_n$ con ogni $\lambda \in \mathbb{R}$.

Ogni sottospazio possiamo dire essere generato da combinazioni lineari dei vettori che lo generano.

Teorema: somma più piccolo sottospazio

Se $S = \{w_1, w_2, \dots, w_r\}$ è un insieme di vettori non nullo contenuto in V, allora:
L'insieme $W = \mathscr{L}((w_1), (w_2), \dots, (w_r))$, è il **più piccolo** sottospazio di V che contiene tutti i vettori di S.
In questo caso si dice che W **è generato** da S.



E' importante riconoscere che gli insiemi di generatori *non sono unici*. Per esempio qualsiasi vettore non nullo sulla linea in figura sarebbe generatore di v. Sono generatori tutte le combinazioni lineari di un insieme di generatori.

4.3 Indipendenza lineare

Diciamo di avere uno spazio xy con vettori standard i e j. Ogni vettore in xy può essere espresso in modo unico come combinazione lineare di i e j. Supponiamo ora di introdurre una terza coordinata w a 45 gradi tra gli assi x e y.

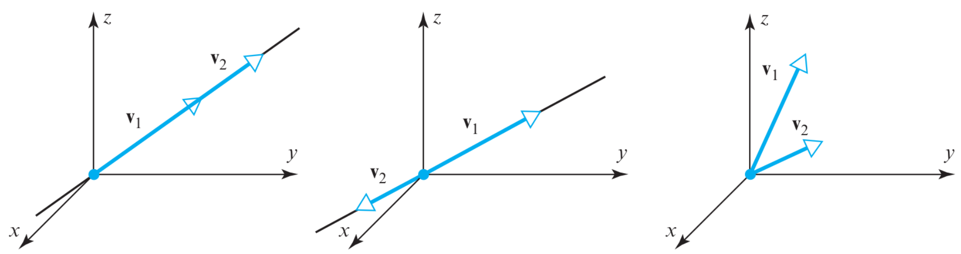
Questo terzo asse risulta essere totalmente superfluo poiché lui stesso può essere espresso come combinazione lineare di i e j per cui non aiuta a descrivere nessun vettore sul piano.

Definizione: Vettori linearmente indipendenti

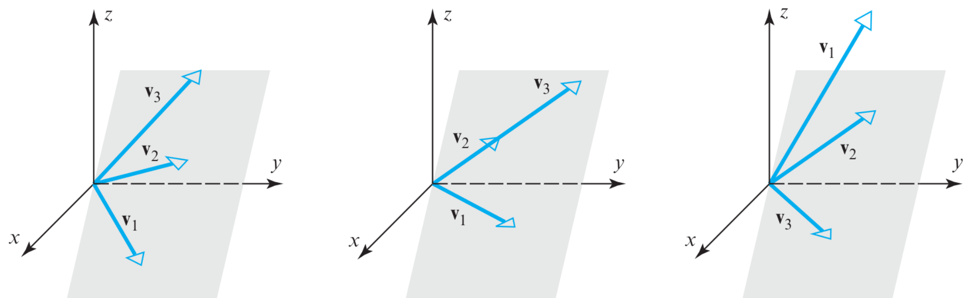
Dato un insieme di vettori $V = \{v_1, v_2, \ldots, v_r\}$, questo è detto *linearmente indipendente* se nessun vettore in V può essere espresso come combinazione lineare di uno degli altri. Quindi se e solo se:

$$k_1 v_1 + k_2 v_2 + \cdots + k_r v_r = 0.$$

è risolto solo da $k_1 = k_2 = \cdots = k_r = 0$.
Un insieme di vettori linearmente indipendenti si dice **libero**.



l.d. *l.i.* *l.i.*



l.d. *l.d.* *l.i.*

4.4 Basi

Ora che sappiamo cosa sono un insieme di generatori e un insieme di vettori linearmente indipendenti possiamo definire il concetto di base.

Definizione: Base

Viene chiamata base di V un insieme di vettori che **genera** V e al contempo è linearmente indipendente.

Teorema: unicità di scrittura

Data una base $\mathscr{B} = \{v_1 + v_2 + \cdots + v_n\}$ ogni vettore $v \in V$ può essere espresso nella forma

$$v = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n$$

in un solo modo.

dimostrazione

Diciamo che esista un'altra combinazione lineare che esprime v; abbiamo:

$$v = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n$$

$$v = k_1 v_1 + k_2 v_2 + \cdots + k_n v_n$$

allora una scrittura meno l'altra deve essere uguale zero

$$c_1 v_1 + c_2 v_2 + \cdots + c_n v_n - k_1 v_1 + k_2 v_2 + \cdots + k_n v_n = 0$$

$$(c_1 - k_1)\, v_1 + (c_2 - k_2)\, v_2 + \cdots + (c_n - k_n)\, v_n = 0$$

da ciò risulta che

$$c_1 - k_1 = c_2 - k_2 = \cdots = c_n - k_n = 0$$

$$c_i = k_i \quad \forall\, i \in 1, \ldots, n.$$

Lemma di Steinitz

Dato V finitamente generato e una sua base $\mathscr{B} = \bar{v}_1, \ldots, \bar{v}_n$:

$$V = \mathscr{L}(\bar{v}_1, \ldots, \bar{v}_n)$$

Prendiamo un insieme libero $\mathscr{I}$ contenuto in V:

$$\mathscr{I} = \{\bar{w}_1, \ldots, \bar{w}_p\} \quad \text{libero} \subseteqq V \qquad \Longrightarrow \qquad p \le n$$

dimostrazione

Prendiamo un vettore $\bar{w} \in V$ non nullo.

$$\bar{w}_1 = \lambda_1 \bar{v}_1 + \cdots + \lambda_n \bar{v}_n$$

visto che il vettore è non nullo posso dire senza perdita di generalità che uno dei λ è diverso da zero. Divido allora tutto per un lambda, diciamo λ_1 così da esplicitare $\bar{v}_1$:

$$\bar{v}_1 = \frac{1}{\lambda_1} \bar{w}_1 - \frac{\lambda_2}{\lambda_1} \bar{v}_2 - \cdots - \frac{\lambda_n}{\lambda_1} \bar{v}_n$$

$$\implies V = \mathscr{L}(\bar{w}_1, \bar{v}_2, \ldots, \bar{v}_n)$$

Posso ripetere l'operazione con $\bar{w}_2$ perché so che è impossibile che tutti i $\lambda_2, \ldots, \lambda_n$ siano nulli (se no si avrebbe $\bar{w}_2 = \lambda \bar{w}_1$ e $\mathscr{I}$ sarebbe libero):

$$\bar{v}_2 = \frac{1}{\lambda_2} \bar{w}_2 - \frac{\lambda_1}{\lambda_2} \bar{w}_1 - \frac{\lambda_2}{\lambda_3} \bar{v}_3 - \cdots - \frac{\lambda_n}{\lambda_2} \bar{v}_n$$

$$\implies V = \mathscr{L}(\bar{w}_1, \bar{w}_2, \bar{v}_3, \ldots, \bar{v}_n)$$

Iterando il processo posso concludere in due modi:

1. Metto tutti i $\bar{w}_j$, possibile solo se $p \leq n$;
2. Arrivo a scrivere una base di V del tipo $\mathscr{L}\left(\bar{w}_1, \ldots, \bar{w}_p\right)$ con $p > n$. Ciò è possibile se

$$\bar{w}_{n+1} \in V \quad \implies \quad \bar{w}_{n+1} = \lambda_1 \bar{w}_1 + \cdots + \lambda_n \bar{w}_n$$

assurdo poiché $\{\bar{w}_1 + \cdots + \bar{w}_{n+1}\} \subsetneqq \mathscr{I}$ che ha dimensione $p \leq n$.

Formula di Grassman

Siano W_1 e W_2 due sottospazi vettoriali di uno spazio vettoriale V, allora:

$$\dim(W_1 + W_2) = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2)$$

dimostrazione

Iniziamo prendendo una base per $W_1 \cap W_2$: $\mathscr{B} = (a_1, \ldots, a_k)$.
Costruisco una base per ognuno dei sottospazi:

$$\mathscr{C} = (a_1, \ldots, a_k, b_{k+1}, \ldots, b_l) \qquad \text{base di } W_1$$

$$\mathscr{D} = \left(a_1, \ldots, a_k, c_{k+1}, \ldots, c_p\right) \qquad \text{base di } W_1$$

La tesi consiste nel dimostrare che

$$\mathscr{E} = \left(a_1, \ldots, a_k, b_{k+1}, \ldots, b_l, c_{k+1}, \ldots, c_p\right)$$

è una base di $W_1 + W_2$.
So di per certo che $\mathscr{E}$ è un insieme di generatori (ho messo solo generatori), devo però dimostrare che è *libero*, ovvero che:

$$\alpha a_1 + \cdots + \alpha a_k + \beta b_{k+1} + \cdots + \beta b_l + \gamma c_{k+1} + \cdots + \gamma c_p = 0 \tag{1}$$

Portiamo a destra dell'uguale la parte appartenente a W_2 e chiamiamola c:

$$\alpha a_1, \cdots + \alpha a_k + \beta b_{k+1} + \cdots + \beta b_l = -\gamma c_{k+1} - \cdots - \gamma c_p = c$$

Da questa equazione vediamo che c è contenuto sia in W_1 sia in W_2, allora si può scrivere come combinazione lineare dei generatori di $W_1 \cap W_2$:

$$c = \lambda_1 a_1 + \cdots + \lambda_k a_k$$

Quindi ora sappiamo che

$$c = \begin{matrix} \lambda_1 a_1 + \cdots + \lambda_k a_k \\ \alpha a_1 + \cdots + \alpha a_k + \beta b_{k+1} + \cdots + \beta b_l \end{matrix}$$

quindi possiamo riscrivere la (1) in questo modo:

$$\lambda_1 a_1 + \cdots + \lambda_k a_k = \alpha a_1 + \cdots + \alpha a_k + \beta b_{k+1} + \cdots + \beta b_l$$

da cui

$$(\lambda_1 - \alpha) a_1 + \cdots + (\lambda_k - a_k) - \beta b_{k+1} - \cdots - \beta b_l = 0$$

ma i vettori che compaiono sono i vettori della base $\mathscr{C}$ da cui segue che

$$\beta_{k+1} = \cdots = \beta$$

4.5 Cambiamento di base

In moltissimi casi risulta più comodo esprimere un vettore rispetto a una base diversa da quella di partenza.
Se cambiamo da una base $\mathscr{B}$ a una $\mathscr{B}'$ come saranno correlate le componenti $[v]_B$ e $[v]_{B'}$?
Le vecchie coordinate sono legate alle nuove dalla seguente equazione:

$$[v]_B = M_B^{B'}\,[v]_{B'}$$

$M_B^{B'}$ è detta **matrice del cambiamento di base** e ha nelle colonne i vettori della base di partenza rispetto la base di arrivo. Quindi la matrice di cambiamento di base da $\mathscr{B} = v_1, \ldots, v_n$ a $\mathscr{B}' = e_1, \ldots, e_n$ sarà:

$$M_B^{B'} = \left[\ [v'_1]_B \ \middle|\ \ldots\ \middle|\ [v'_n]_B\ \right]$$

per convincersene

Diciamo di avere base di partenza e base di arrivo:

$$\mathscr{B} = \{\bar{v}_1, \bar{v}_2\} \qquad \mathscr{B}' = \{\bar{v}_1', \bar{v}_2'\}$$

Sappiamo che i vettori della nuova base possono essere scritti come combinazione lineare dei ettori della vecchia base: a partire da $\mathscr{B}$ definiamo i vettori di $\mathscr{B}'$.

$$v_1' = a\bar{v}_1 + b\bar{v}_2$$

$$v_2' = c\bar{v}_1 + d\bar{v}_2$$

La base di arrivo può essere riscritta come $\mathscr{B}' = \{(a\bar{v}_1 + b\bar{v}_2), (c\bar{v}_1 + d\bar{v}_2)\}$.
Un vettore rispetto a $\mathscr{B}'$ si può riscrivere come combinazione lineare dei vettori della base:

$$[v]_{\mathscr{B}'} = \begin{bmatrix} k_1 \\ k_2 \end{bmatrix}$$

$$\begin{aligned} v &= k_1(a\bar{v}_1 + b\bar{v}_2) + k_2(c\bar{v}_1 + d\bar{v}_2) \\ &= (k_1 a + k_2 c)\,\bar{v}_1 + (k_1 b + k_2 d)\,\bar{v}_2 \end{aligned}$$

Adesso possiamo riscrivere il vettore rispetto alla base di partenza $\mathscr{B}$

$$\begin{aligned} [v]_{\mathscr{B}} &= \begin{bmatrix} k_1 a & k_2 c \\ k_1 b & k_2 d \end{bmatrix} \\ &= \begin{bmatrix} a & c \\ b & d \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} \end{aligned}$$

Abbiamo ritrovato la forma

$$[v]_{\mathscr{B}} = M_B^{B'}\,[v]_{\mathscr{B}'}$$

Metodo per trovare la matrice di cambiamento di base

Per trovare la matrice di cambiamento di base si può utilizzare un metodo simile a quello impiegato per trovare la matrice inversa: si scrive la matrice orlata con a sinistra la base $\mathscr{B}$ e a destra la base $\mathscr{B}'$. Quindi si riduce per righe fino a quando la matrice non ha la forma

$$\left[\begin{array}{c|c} I & M_B^{B'} \end{array}\right]$$

4.6 Spazio delle righe, delle colonne, Nullspace

Prendiamo in esame la matrice

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{r,1} & a_{r,2} & \cdots & a_{r,n} \end{bmatrix}$$

Chiamiamo spazio delle colonne l'insieme dei vettori colonna nella matrice A e spazio delle righe l'insieme dei vettori riga in A. Il **nullspace** di A invece (introdotto nel capitolo sui sistemi lineari) ricordiamo esseere la solzione dell'equazione $AX = 0$.

C(A) e nullspace Che relazione intercorre tra lo spazio delle colonne e il nullspace? Scriviamo l'equazione $Ax = b$ per colonne:

$$Ax = x_1 c_1 + x_2 c_2 + \cdots + x_n c_n = b$$

da questa scrittura notiamo che b *può essere scritto come combinazione lineare delle colonne di* A. Allora $b \in C(A)$, b fa parte dello spazio delle colonne di A.

Teorema del rango

Data $f : V \to W$, con $\dim V = n$, allora possiamo affermare in modo equivalente che:

$$\dim Ker(f) + \dim Im(f) = \dim(V)$$

$$N(A) + C(A) = n$$

Teorema di nullità + rango

Il teorema afferma che

$$R(A) \oplus^{\perp} N(A) = \mathbb{R}^n$$

ovvero che lo spazio delle righe è in somma diretta e ortogonale con il *nullspace* della matrice A.
Questo ci mostra anche:

$$\dim N(A) + \mathrm{rk} A = n$$

Ovvero che la dimensione nel nullspace equivale al numero di righe nulle e quindi anche al numero di parametri liberi.

dimostrazione

Sappiamo che il nullspace è l'insieme dei vettori $\bar{x}$ che soddisfano

$$A\bar{x} = 0$$

Se riscriviamo A evidenziandone le **righe** notiamo che

$$\begin{pmatrix} R_1 \\ \vdots \\ R_r \end{pmatrix} \bar{x} = \bar{0} \quad \to \quad R_j \cdot \bar{x} = \bar{0}$$

Quindi $\bar{x}$ è ortogonale allo spazio delle righe (rispetto al prodotto scalare standard).

4.7 Rank

Possiamo definire il rango di una matrice in tre modi diversi.

1. Numero di righe non nulle della matrice ridotta a scala;
2. Dimensione dello spazio delle righe o dello spazio delle colonne;
3. Massimo ordine di un minore non nullo della matrice.

5 Applicazioni Lineari

Le applicazioni lineari sono particolari tipi di *funzioni* che preservano la struttura di spazio vettoriale.

$$f : V \longrightarrow W \qquad V, W \text{ sottospazi vettoriali}$$

Definizione: Applicazione lineare

Dati due spazi vettoriali reali V, W, si dice applicazione lineare o **omomorfismo** o trasformazione lineare da V in W una funzione

$$f : V \longrightarrow W.$$

che verifica le seguenti proprietà:

$$f\left(\bar{0}_v\right) = \bar{0}_W$$
$$f\left(x + y\right) = f(x) + f\left(y\right)$$
$$f\left(\lambda x + \mu y\right) = \lambda f(x) + \mu f\left(y\right).$$

per ogni x e y in V e per ogni λ e μ in $\mathbb{R}$.

Teorema fondamentale delle applicazioni lineari

Dati V e W spazi vettoriali,

$$\mathscr{B} = \{\overline{v}_1, \ldots, \overline{v}_n\} \text{ base di } V$$
$$\mathscr{C} = \{\overline{a}_1, \ldots, \overline{a}_n\} \text{ insieme di vettori in } W.$$

Allora esiste ed è **unica** l'applicazione lineare

$$f : V \longrightarrow W \text{ t.c.}$$

$$f\left(\overline{v}_i\right) = \overline{a}_i \quad \forall i \in \{1, \ldots, n\}.$$

In altre parole per assegnare un'applicazione lineare tra due spazi vettoirali V e W, di cui almeno V di dimenzione finita, è sufficiente conoscere le immagini dei vettori di una vase di V.

dimostrazione

$$\overline{x} \in V.$$
$$\overline{x} = x_1\overline{v}_1 + x_2\overline{v}_2 + \cdots + x_n\overline{v}_n.$$

$$\begin{aligned} f\left(\overline{x}\right) &= f\left(x_1\overline{v}_1 + x_2\overline{v}_2 + \cdots + x_n\overline{v}_n\right) \\ &= x_1 f\left(\overline{v}_1\right) + x_2 f\left(\overline{v}_2\right) + \cdots + x_n f\left(\overline{v}_n\right) \\ &= x_1\overline{a}_1 + x_2\overline{a}_2 + \cdots + x_n\overline{a}_n \end{aligned}$$

Viceversa se definiamo f dicendo che

$$f\left(\overline{x}\right) = x_1\overline{a}_1 + x_2\overline{a}_2 + \cdots + x_n\overline{a}_n.$$

Allora

- f è lineare: $\quad f\left(\lambda\overline{x} + \mu\overline{y}\right) = \lambda f\left(\overline{x}\right) + \mu f\left(\overline{y}\right)$.
- $f\left(\overline{v}_i\right) = \overline{a}_i \quad \forall i$

Quindi definire un'applicazione lineare tra due spazi vettoriali equivale a *conoscere le immagini degli elementi di una base del dominio.*

Teorema: Suriettivià

Data l'applicazione lineare $f : V \to W$, f si dice suriettiva se

$$\mathrm{im} f = W$$

o se

$$\mathrm{rank}(A) = \dim W$$

Una funzione è suriettiva se il sottospazio immagine corrisponde con il codominio, per questo non può esistere un'isola, no dai, non può esistere una funzione suriettiva del tipo $f : \mathbb{R}^2 \to \mathbb{R}^3$ (si ha che $\mathrm{Im} f$ ha dimensione massima 2, quindi non potrà mai generare tutto $\mathbb{R}^3$).

Teorema: Iniettività

Data l'applicazione lineare $f : V \to W$, f si dice suriettiva *se l'immagine di ogni insieme libero di V è un insieme libero di W.*

Oppure se

$$\ker f = \{\bar{0}_V\}$$

Mandando insiemi liberi in insiemi liberi, l'unico vettore che può essere mappato in 0 è solo $\{\bar{0}\}$.

Isomorfismo e automorfismo

Con $\dim V = \dim W$ $\quad f$ iniettiva se e solo se f suriettiva.

1. f è un **isomorfismo** $\iff \ker f = \{\bar{0}_V\}$
2. f è un **isomorfismo** $\iff \mathrm{Im} f = W$

Se invece prendiamo un endomorfismo $f : V \to V$

1. f è un **automorfismo** $\iff \ker f = \{\bar{0}_V\}$
2. f è un **automorfismo** $\iff \mathrm{Im} f = V$

5.1 Controimmagine

Teorema fondamentale per comprendere meglio alcuni concetti legati al kernel:

Definizione: controimmagine

Sia $f : V \to W$ e sia $\mathcal{K}$ sottospazio vettoriale di W, allora

$$f^{-1}(\mathcal{K}) = \left\{\bar{x} \in V : f(x) \in \mathcal{K}\right\}$$

Teorema: f^{-1} è sottospazio

La controimmagine $f^{-1}(\mathcal{K})$ di un sottospazio vettoriale $\mathcal{K}$ di W è un sottospazio vettoriale di V.

dimostrazione

Dalla definizione di sottospazio vettoriale bisogna dimostrare che *per ogni* $\bar{x}_1, \bar{x}_2 \in f^{-1}(\mathscr{K})$ e per ogni $\lambda, \mu \in \mathbb{R}$ si ha:

$$\lambda \bar{x}_1 + \mu \bar{x}_2 \in f^{-1}(\mathscr{K})$$

Quindi dobbiamo dimostrare che

$$f\left(\lambda \bar{x}_1 + \mu \bar{x}_2\right) \in \mathscr{K}$$

Per linearità segue

$$f\left(\bar{x}_1 + \mu \bar{x}_2\right) = \lambda f(\bar{x}_1) + \mu f(\bar{x}_2)$$

Poiché $\mathscr{K}$ è sottospazio vettoriale l'ultima uguaglianza vale ed è contenuta in esso; allora abbiamo dimostrato la tesi.

Nucleo

Definizione: Nucleo

Chiamiamo nucleo di un'applicazione lineare $f : V \to W$ il sottospazio vettoriale di V controimmagine del sottospazio vettoriale $\{\bar{0}_W\}$ e si indica con:

$$\ker f = \left\{\bar{x} \in V \,|\, f(x) = \bar{0}_W\right\}$$

Il fatto che il kernel sia un sottospazio deriva proprio dal teorema precedente.

6 Autovettori, autovalori

6.1 Automorfismi e sottospazi invarianti

Definizione: Automorfismo

Un endomorfismo **anche biettivo** (quindi isomorfismo) lo chiamiamo **automorfismo**. Dato $f : V \to V$ posso associare a f la matrice rappresentativa $M^{\mathscr{B}}(f)$.

$$f \text{ automorfismo} \iff M^{\mathscr{B}}(f) \text{ invertibile} \iff \det M^{\mathscr{B}}(f) \neq 0$$

Definizione: Sottospazio invariante

Dato un sottospazio W e un vettore $w \in W$ il sottospazio si dice invariante se $f(w) \subseteq W$.
Un sottospazio invariante è, per esempio, un *autospazio*; infatti i vettori di un autospazio sono multipli delle loro immagini: $f(\overline{x}) = \lambda \overline{x}$.

6.2 Autovettori e autovalori

Definizione: Autovettori

Data l'endomorfismo $f : V \to V$ un vettore $\overline{x} \in V \quad \overline{x} \neq \overline{0}$ si dice **atuovettore** se

$$f(\overline{x}) = \lambda \overline{x}$$

.

- Lo scalare λ è l'**autovalore** associato a $\overline{x}$.
- V_λ si dice **autospazio** associato a λ ed è l'insieme di tutti gli autovettori $\{\overline{x} \in V : f(x) = \lambda \overline{x}\}$.

Quindi chiamiamo autovettori tutti quei vettori che vengono mandati da una certa funzione f in multipli di loro stessi. Ora dobbiamo occuparci di come trovare questi autovettori e i corrispondenti autovalori.
Diciamo di avere A matrice associata all'endomorfismo $f : V \to V$, per trovare gli atuovalori impostiamo la seguente uguaglianza:

$$\begin{aligned} A\overline{x} &= \lambda \overline{x} \\ A\overline{x} - \lambda \overline{x} &= \overline{0} \\ (A - \lambda I)\overline{x} &= \overline{0} \end{aligned}$$

L'uguaglianza è rappresentata da un sistema omogeneo che ha soluzioni non banali solo se il rango della matrice $A - \lambda I$ non è massimo, quindi solo se il determinante è uguale a 0.

$$\det(A - \lambda I) = 0$$

Il polinomio risultante dall'equazione viene chiamato *polinomio caratteristico* di A: $P_A(\lambda)$. Le radici del polinomio caratteristico sono i nostri autovalori:

$$\det(A - \lambda I) = \begin{vmatrix} a_{11} - \lambda & & & \\ & a_{22} - \lambda & & \\ & & \ddots & \\ & & & a_{nn} - \lambda \end{vmatrix} = \text{polinomio di grado } n \text{ in } \lambda$$

$$= (a_{11}-\lambda)(a_{22}-\lambda)\dots(a_{nn}-\lambda)$$
$$= \lambda^n + c_1\lambda^{n-1} + \dots + c_n.$$

Dato un certo polinomio caratteristico, è chiamata *molteplicità algebrica* il numero di volte che un certo λ_0 annulla il polinomio: $\mathrm{m_a}(\lambda_0)$.

Un altro punto di vista Abbiamo detto che per trovare gli autovalori dobbiamo risolvere la sequante equazione:

$$(A-\lambda I)\overline{x} = \overline{0}$$

Dal capitolo sul determinante di una matrice sappiamo che il determinante rappresenta un fattore di stretching; nel nostro caso serve che un vettore non nullo ($\overline{x}$) venga mandato in zero dalla matrice $A-\lambda I$, l'unico caso in cui questo è possibile è quando il determinante di tale matrice è uguale a zero.

Autospazi Vale la pena soffermarsi su cosa sono, come si trovano e su alcune proprietà degli autospazi. Per trovare un autospazio V_{λ_0} associato a un autovalore λ_0 si calcola il nullspace della matrice $A-\lambda_0 I$:

$$V_{\lambda_0} = \mathrm{N}(A-\lambda_0 I)$$

Il numero di generatori del nullspace, ovvero la sua dimensione, viene chiamata *molteplicità geometrica*: $\mathrm{m_g}(\lambda_0)$

1. Ogni vettore $\overline{x} \in V_{\lambda_1},\dots,V_{\lambda_k}$ si scrive in modo unico come $\overline{x}_1+,\dots,+\overline{x}_k$ con $x_j \in V_{\lambda_j}$, ovvero gli autospazi associati agli autovalori di un certo polinomio caratteristico sono in **somma diretta**.

   Questo implica
   $$\dim\left(V_{\lambda_1}\oplus,\dots,\oplus V_{\lambda_k}\right) = \dim V_{\lambda_1} + \dots + \dim V_{\lambda_k}.$$

2. La molteplicità geometrica di un autospazio è minore o uguale alla molteplicità algebrica:
   $$\mathrm{m_g}(\lambda_0) \le \mathrm{m_a}(\lambda_0).$$

dimostrazione

Diciamo di avere un autospazio V_{λ_0} di dimensione k ($k = \mathrm{m_g}(\lambda_0)$) sottospazio vettoriale di V di dimensione n.
Prendo una base di $V_{\lambda_0}\mathscr{B}' = \{\overline{v}_1,\dots,\overline{v}_k\}$ e la completo ad una base di V:

$$\mathscr{B} = \{\overline{v}_1,\dots,\overline{v}_k,\overline{v}_{k+1},\dots,\overline{v}_n\}$$

La matrice associata sarà:

$$M^{\mathscr{B}}(f) = \begin{bmatrix} \lambda_0 & 0 & & \\ 0 & \lambda_0 & & \\ \vdots & 0 & & \\ 0 & 0 & & \end{bmatrix}$$

Sappiamo che nelle prime k colonne c'è $\lambda_0 I$, nelle rimanenti $n-k$ colonne invece, non sappiamo dire cosa ci sia. Rimaniamo con una matrice del genere:

$$M^{\mathscr{B}}(f) = \begin{bmatrix} \lambda_o I_k & B \\ 0 & C \end{bmatrix} \qquad B \in \mathbb{R}^{k,n-k}, \quad C \in \mathbb{R}^{n-k,n-k}$$

Ora tenendo a mente che l'obiettivo è dimostrare che la molteplicità geometrica di λ_0 sia minore o uguale alla molteplicità geometrica, ovvero la molteplicità algebrica (il numero di radici reali del polinomio caratteristico) sia almeno k, andiamo a calcolare i nostri lambda.

$$\det\left(M^{\mathscr{B}}(f) - \lambda I\right) = \begin{vmatrix} \lambda_o I_k - \lambda I_k & B \\ 0 & C-\lambda I \end{vmatrix} = 0$$
$$= (\lambda_0-\lambda)^k (C-\lambda I) = 0$$

L'ultima uguaglianza conferma che la molteplcità algebrica di λ_o sia almeno uguale a k.

Teorema Una matrice quadrata associata a una funzione f è invertiile se e solo se $\lambda = 0$ non è un suo autovalore.
Se ci riflettiamo, se eiste un $\lambda_0 = 0$, vuol dire che la funzione manda un vettore non nullo in un vettore nullo:

$$f\left(\overline{x}\right) = 0\overline{x}.$$

Questo ci dice che la funzione è *non-iniettiva* e che di conseguenza non può essere biettiva, condizione necessaria perché una funzione sia invertibile.

7 Diagonalizzazione

Scopo principale della diagonalizzazione sarà quello di trovare basi di soli autovettori. Il processo è schematizzato in questo modo:

$$\begin{tikzcd}
(V_{B'}) \arrow[r, "f"] \arrow[d, "\mathrm{id}"'] & (V_{B'}) \\
(V_B) \arrow[r, "f"] & (V_B) \arrow[u, "\mathrm{id}"']
\end{tikzcd}$$

Data una matrice scritta rispetto alla base B' la si riscrive rispetto alla base B, si applica la trasformazione lineare f e si ritorna alla base B'.

$$M_{B'}^{B'}\left(f\right) = M_{B'}^{B}\left(\mathrm{id}\right) M_{B}^{B}\left(f\right) M_{B}^{B'}\left(\mathrm{id}\right).$$

Quindi se chiamiamo

$$A = M_B^B\left(f\right) \qquad A' = M_{B'}^{B'}\left(f\right) \qquad P = M_B^{B'}\left(\mathrm{id}\right).$$

possiamo riscrivere la precedente come:

$$A' = P^{-1}AP.$$

Diremo che A' è **simile** a A. Le matrici simili condividono molte caratteristiche:

1. A e $P^{-1}AP$ hanno stesso determinante;

$$\begin{aligned}
\det\left(A'\right) &= \det\left(P^{-1}AP\right) \\
&= \det\left(P^{-1}\right)\det\left(A\right)\det\left(P\right) \\
&= \frac{1}{\det\left(P\right)}\det\left(A\right)\det\left(P\right) \\
&= \det\left(A\right)
\end{aligned}$$

2. A invertibile $\iff$ A' invertibile;

3. A e A' hanno stesso *rango*, *nullspace*, *polinomio caratteristico*.

Allo scopo di diagonalizzare una matrice le matrici simili sono essenziali, infatti, una matrice quadrata A si dice **diagonalizzabile** se è simile a un'altra matrice diagonale.

7.1 Criteri di diagonalizzabilità

$$f : V \longrightarrow V \text{ endomorfismo.} \quad \text{Sono equivalenti:}$$

1. f diagonalizzabile ;

2. $P_f\left(\lambda\right)$ ha solo radici reali e per ogni λ_j si ha $\mathrm{m_g}\left(\lambda_j\right) = \mathrm{m_a}\left(\lambda_j\right)$;

3. Se gli autovalori sono tutti distinti, $\dim\left(V_{\lambda_1}\right) + \cdots + \dim\left(V_{\lambda_k}\right) = \dim\left(V\right)$;

4. Se gli autovalori sono tutti distinti, $V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_k}$.

dimostrazione $1 \Rightarrow 2$

Sia f diagonalizzabile e $\mathscr{B}$ base di autovettori.

$$M^B(f) = \begin{bmatrix} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_1 & & & \\ & & & \lambda_2 & & \\ & & & & \ddots & \\ & & & & & \lambda_k \end{bmatrix} \begin{array}{l} \left.\begin{array}{l} \\ \\ \end{array}\right\} m_1 \\ \left.\begin{array}{l} \\ \end{array}\right\} m_2 \\ \left.\begin{array}{l} \\ \end{array}\right\} m_k \end{array}$$

$$P_f(\lambda) = \det\left(M^B(f) - \lambda I\right) \qquad = (\lambda_1 - \lambda)^{m_1} (\lambda_2 - \lambda)^{m_2} \dots (\lambda_k - \lambda)^{m_k}$$

Il polinomio ha solo radici reali; il massimo grado del polinomio lo si trova sommando gli esponenti:

$$m_1 + m_2 + \cdots + m_k = \dim V$$

Infatti il numero di colonne della matrice è uguale alla dimensione di V (essendo le colonne sicuramente linearmente indipendenti) ed è uguale al numero di righe, uguale a $m_1 + \cdots + m_k$.

Ora dobbiamo dimostrare che per ogni λ_j si ha $\mathrm{m_g}(\lambda_j) = \mathrm{m_a}(\lambda_j)$.
Troviamo il sottospazio associato a λ_1:

$$V_{\lambda_1} = \mathrm{N}\left(M^B(f) - \lambda_1 I\right)$$

$$= \begin{bmatrix} 0 & & & & & & & & & \\ & \ddots & & & & & & & & \\ & & 0 & & & & & & & \\ & & & \lambda_2 - \lambda_1 & & & & & & \\ & & & & \ddots & & & & & \\ & & & & & \lambda_2 - \lambda_1 & & & & \\ & & & & & & \ddots & & & \\ & & & & & & & \lambda_k - \lambda_1 & & \\ & & & & & & & & \ddots & \\ & & & & & & & & & \lambda_k - \lambda_1 \end{bmatrix} \begin{array}{l} \left.\begin{array}{l} \\ \\ \\ \end{array}\right\} m_1 \\ \left.\begin{array}{l} \\ \\ \\ \end{array}\right\} m_2 \\ \left.\begin{array}{l} \\ \\ \\ \\ \end{array}\right\} m_k \end{array}$$

Le prime righe (contrassegnate da m_1) sono nulle, quindi il rank della matrice è $\mathrm{rk} M^B(f) = \dim V - m_1$. Quindi la dimensione del nullspace è

$$\dim V - \operatorname{rank}(A - \lambda_1 I) = m_1$$

$$\Longrightarrow \mathrm{m_g}(\lambda_1) = \mathrm{m_a}(\lambda_1) = m_1$$

dimostrazione $2 \Rightarrow 3$

Sappiamo che il polinomio caratteristico si scompone in equazioni lineari:

$$P_f(\lambda) = (\lambda_1 - \lambda)^{m_1} \dots (\lambda_k - \lambda)^{m_k}.$$

Dalla dimostrazione precedente sappiamo che la molteplicità algebrica è uguale a quella geometrica. Ricordando che $\mathrm{m_g}(\lambda_j) = \dim\left(V_{\lambda_j}\right)$ e che $m_j = \mathrm{m_a}(\lambda_j)$

$$m_1 + m_2 + \cdots + m_k = \dim V$$

$$\mathrm{m_g}(\lambda_j) = \mathrm{m_a}(\lambda_j)$$

$$\mathrm{m_a}(\lambda_1) + \cdots + \mathrm{m_a}(\lambda_k) = \dim V$$

$$\Longrightarrow \dim\left(V_{\lambda_1}\right) + \cdots + \dim\left(V_{\lambda_k}\right) = \dim V.$$

dimostrazione: $3 \Rightarrow 4$

$$\dim\left(V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_k}\right) = \dim\left(V_{\lambda_1}\right) + \cdots + \dim V_{\lambda_k} = \dim V$$

Sappiamo infatti che $V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_k}$ è un sottospazio di V con la stessa dimensione.

dimostrazione $4 \Rightarrow 1$

So che V si decompone in somma diretta di autospazi

$$V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_k}$$

Ora prendo una base per ogni autospazio

$$\begin{array}{ccc} B_1 & \text{base per} & V_{\lambda_1} \\ \vdots & & \vdots \\ B_k & \text{base per} & V_{\lambda_k} \end{array}$$

Sappiamo che B base di V si ottiene unendo tutte le basi $B_1, \ldots, B_k$ degli autospazi:

$$B = B_1 \cup \cdots \cup B_k$$

B è fatta da autovettori, quindi f diagonalizzabile.

7.2 Endomorfismi autoaggiunti

Definizione: Endomorfismo autoaggiunto

Chiamiamo **autoaggiunto** l'endomorfismo nel qual vale:

$$f\left(\overline{x}\right) \cdot \overline{y} = \overline{x} f\left(\overline{y}\right) \qquad \forall \overline{x}, \overline{y} \in V$$

Ora che abbiamo studiato le applicazioni lineari che hanno come matrice associata una matrice simmetrica, introduciamo un teorema fondamentale.

Teorema spettrale

Dato $(V, \cdot)$ euclideo e $f \in \mathrm{End}(V)$, allora:

$$f \text{ autoaggiunto} \iff \exists \mathscr{B} \text{ ortonormale di autovettori}$$

In particolare:

$$f \text{ autoaggiunto} \quad \Rightarrow \quad f \text{ diagonalizzabile.}$$

dimostrazione

$\Longleftarrow$ Diretta conseguenza della proposizione: se B ortonormale di autovettori, $M^B\left(f\right)$ è diagonale (quindi simmetrica).

$\Longrightarrow$ Dimostriamo che esiste una base ortonormale di autovettori in tre punti:

1. Voglio solo radici reali;
2. Mostro che gli autospazi sono ortogonali fra loro;
3. Mostro che $V_{\lambda_1} \oplus^{\perp} \cdots \oplus^{\perp} V_{\lambda_k} = V$.

1) *f autoaggiunto $\Longrightarrow P_f(\lambda)$ ha solo radici reali.*
Il polinomio $P_f(\lambda)$ avrà n radici complesse ($\lambda \in \mathbb{C}$). Sappiamo che ogni radice complessa viene accoppiata con il suo coniugato, per questo, faremo vedere che vale

$$\lambda = \bar{\lambda}$$

Troviamo gli autovalori:

$$\det(A - \lambda I) = 0$$

$$\exists x \in \mathbb{C}^n \quad \text{t.c.} \quad AX = \lambda X$$

Trovo lambda coniugato:

$$\bar{A X} = \overline{\lambda X} \to A\overline{X} = \bar{\lambda}\overline{X}$$

Ho trovato che $\overline{X}$ è autovettore di A con autovalore $\overline{\lambda}$. Adesso mostriamo che $\lambda = \overline{\lambda}$:

$${}^t\bar{X}AX = {}^t\bar{X}\lambda X = \lambda\, {}^t\bar{X}X$$

$$({}^t\bar{X}A)X = {}^t(A\bar{X})X = \bar{\lambda}\, {}^t\bar{X}X$$

poiché siamo partiti da due espressioni uguali la loro differenza deve essere uguale a zero:

$$\lambda\, {}^t\bar{X}X - \bar{\lambda}\, {}^t\bar{X}X = 0 \to (\lambda - \bar{\lambda})({}^t\bar{X}X) = 0$$

poiché

$${}^t\bar{X}X = (\bar{x}_1 \dots \bar{x}_n) \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = |x_1|^2 + \cdots + |x_n|^2 > 0$$

sappiamo che

$$(\lambda - \bar{\lambda})({}^t\bar{X}X) = 0 \Longleftrightarrow \lambda = \bar{\lambda}$$

2) *Gli autospazi sono ortogonali fra loro.*
Prendo due vettori appartenenti a due autospazi diversi:

$$\bar{x} \in V_\lambda, \bar{y} \in V_\mu \qquad \lambda \neq \mu$$

$$f(\bar{x}) \cdot \bar{y} = \bar{x} \cdot f(\bar{y})$$
$$\lambda\bar{x} \cdot \bar{y} = \bar{x} \cdot \mu\bar{y}$$

$$(\lambda - \mu)\bar{x} \cdot \bar{y} = 0 \Longleftrightarrow \bar{x} \cdot \bar{y} = 0$$

3) *Gli autospazi sono in somma diretta-ortogonale.*
Per dimostrare che la somma degli autospazi coincide con tutto V è necessario mostrare W insieme degli autospazi è $= V$, e quindi che $W^{\perp} = \{\bar{0}\}$. Procederemo in questo ordine:

1. Iniziamo col mostrare che $W^{\perp}$ è un sottospazio invariante;
2. Studiando la restrizione di f a $W^{\perp}$ mostriamo che in esso non possono esserci autovettori.

3.1) Se $W \leq V$ è invariante per f, allora lo è anche $W^{\perp}$. Voglio far vedere che se $\bar{x} \in W^{\perp}$ allora $f(\bar{x}) \in W^{\perp}$. Diciamo di avere un vettore $\bar{w} \in W$:

$$f(\bar{x}) \cdot \bar{w} = \bar{x} \cdot f(\bar{w}) = 0$$

quindi $f(\bar{x}) \in W^{\perp}$ e $W^{\perp}$ è sottospazio invariante.

3.2) Ora supponiamo che $W \neq V$ e che di conseguenza $W^{\perp} \neq \{\bar{0}\}$. Studiamo la restrizione di f a $W^{\perp}$:

$$f : W^{\perp} \to W^{\perp}$$

Visto che vale

$$f(\bar{x}) \cdot \bar{y} = \bar{x} \cdot f(\bar{y}) \qquad \forall \bar{x}, \bar{y} \in V$$

$\forall \bar{x}, \bar{y} \in W^{\perp} \quad f|_{W^{\perp}}$ è **autoaggiunto** e possiamo usare le dimostrazioni precedenti. Per lo *step 1* infatti $P_{f|_{W^{\perp}}}(\lambda)$ ha solo radici reali e quindi

$$\exists \bar{x} \in W^{\perp}, \exists \lambda \in \mathbb{R} \quad \text{t.c.} \quad f(\bar{x}) = \lambda \bar{x}$$

Questo ci dice che esistono autovettori $\bar{x} \in W^{\perp}$ e che quindi $W^{\perp}$ deve essere contenuto in W insieme che contiene tutti gli autospazi, vero solo se $W^{\perp} = \bar{0}$ e quindi $W = V$.

Corollario Sia una matrice A simmetrica, A è sempre diagonalizzabile. Esiste allora una matrice D diagonale e una P invertibile tale che:

$$D = P^{-1} A P$$

E' inoltre possibile individuare una matrice ortogonale Q tale che:

$$D = Q^{-1} A Q = {}^{t}Q A Q.$$

Dato l'endomorfismo $f : V \longrightarrow V$ con $(V, \cdot)$ euclideo, sono equivalenti:

1. f autoaggiunto;
2. $\forall B$ ortonormale, $M^{B}(f)$ è simmetrica.
3. $\exists B$ ortonormale t.c. $M^{B}(f)$ simmetrica.

dimostrazione $1 \Rightarrow 2$

Prendo $B = \{\overline{v}_1, \ldots, \overline{v}_n\}$ base ortonormale e due vettori $\overline{x}$ e $\overline{y} \in V$.

Scritti come vettori colonna abbiamo X, Y, AX e AY rispettivamente per x, y, $f(x)$ e $f(y)$.
Possiamo riscrivere (poiché il prodotto è standard)

$$f(\overline{x}) \cdot \overline{y} = \overline{x} f(\overline{y})$$

come

$${}^{t}(AX)\, Y = {}^{t}X A Y$$

$${}^{t}X\, {}^{t}A Y = {}^{t}X A Y.$$

dalla quale risulta che $A = {}^{t}A$.

8 Forme bilineari

Definizione: Forma bilineare

Una forma bilineare su uno spazio vettoriale V è una funzione

$$\phi : V \times V \to \mathbb{R}$$

per cui valgono le seguenti proprietà:

1. $\phi(\bar{x} + \bar{x}', \bar{y}) = \phi(\bar{x}, \bar{y}) + \phi(\bar{x}', \bar{y})$
2. $\phi(\bar{x}, \bar{y}' + \bar{y}') = \phi(\bar{x}, \bar{y}) + \phi(\bar{x}, \bar{y}')$
3. $\phi(\lambda \bar{x}, \bar{y}) = \lambda \phi(\bar{x}, \bar{y})$

Studieremo principalmente le forme bilineari simmetriche, ovvero quelle che soddisfano

$$\phi(\bar{x}, \bar{y}) = \phi(\bar{y}, \bar{x})$$

Ogni prodotto scalare è un esempio di FBS e poiché ci si propone di studiare opportune generalizzazione di prodotto scalare, prenderemo in considerazione solo FBS.

Da notare come sia possibile definire in modo naturale na struttura di spazio vettoriale su $\mathbb{R}$ sull'insieme delle forme bilineari simmetriche; possiamo infatti definire l'operazione di somma:

$$\phi_1 + \phi_2(\bar{x}, \bar{y}) = \phi_1(\bar{x}, \bar{y}) + \phi_2(\bar{x}, \bar{y}) \qquad \bar{x}, \bar{y} \in V$$

e l'operazione di prodotto per uno scalare:

$$(\lambda\phi)(\bar{x}, \bar{y}) = \lambda\phi(\bar{x}, \bar{y}) \qquad \bar{x}, \bar{y} \in V, \lambda \in \mathbb{R}$$

8.1 Matrice associata alla FBS

Definiamo una base di V $\mathscr{B} = (\bar{v}_1, \dots, \bar{v}_n)$ e due vettori $\bar{x}, \bar{y} \in V$, è possibile esprimere $\phi(\bar{x}, \bar{y})$ in termini dei vettori della base:

$$\phi(\bar{x}, \bar{y}) = \sum_{i,j=1}^{n} x_i y_j \phi(\bar{v}_i, \bar{v}_j)$$

La matrice associata sarà quindi

$$M^{\mathscr{B}}\phi = \begin{bmatrix} \phi(\bar{v}_1, \bar{v}_1) & \phi(\bar{v}_1, \bar{v}_2) & \dots & \phi(\bar{v}_1, \bar{v}_n) \\ \phi(\bar{v}_2, \bar{v}_1) & \phi(\bar{v}_2, \bar{v}_2) & \dots & \phi(\bar{v}_2, \bar{v}_n) \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ \phi(\bar{v}_n, \bar{v}_1) & \phi(\bar{v}_n, \bar{v}_2) & \dots & \phi(\bar{v}_n, \bar{v}_n) \end{bmatrix} = A$$

e vale

$$\phi(\bar{x}, \bar{y}) = {}^{t}XAY$$

8.2 Forma quadratica

Una volta definito un nuovo tipo di prodotto scalare serve un nuovo concetto di *norma* di un vettore $\|\bar{x}\| = \bar{x} \cdot \bar{x}$.

Definizione: Forma quadratica

Data la forma bilineare $\phi \in B_s(V, \mathbb{R})$, l'applicazione:

$$Q : V \to \mathbb{R}, \qquad \bar{x} \mapsto Q(\bar{x}) = \phi(\bar{x}, \bar{x})$$

si dice forma quadrata associata alla forma bilineare associata.

E' facile notare che le forme quadratiche non sono applicazioni lineari:

$$Q(\lambda\bar{x}) = \phi(\lambda\bar{x}, \lambda\bar{x}) = \lambda^2 Q(\bar{x})$$

$$Q(\bar{x}+\bar{y}) = Q(\bar{x}) + 2\phi(\bar{x},\bar{y}) + Q(\bar{y})$$

Definizione: Nucleo di una forma bilineare

Il nucleo di una forma bilineare è il sottospazio contenente tutti i vettori *ortogonali a tutto lo spazio* rispetto a ϕ, ovvero che annullano la forma bilineare:

$$\phi(\bar{x},\bar{y}) = 0 \quad \forall \bar{y} \in V$$

Quindi

$$\ker\phi = \left\{\bar{x} \in V | \phi(\bar{x},\bar{y}) = 0 \quad \forall \bar{y} \in V\right\}$$

Definizione: Cono isotropo

Chiamiamo cono isotropo l'insieme dei vettori *ortogonali a loro stessi* rispetto a ϕ, ovvero che annullano la forma quadratica:

$$\phi(\bar{x},\bar{x}) = Q(x) = 0$$

Quindi

$$\mathscr{I}_\phi = \left\{\bar{x} \in V | Q(\bar{x}) = 0\right\}$$

Notiamo che se un vettore è ϕ-ortogonale a tutto lo spazio allora è anche ϕ-ortogonale anche a se stesso; quindi il kernel di una certa forma bilineare simmetrica è sempre contenuto nel cono isotropo.

Classificazione forma quadratica

1. **Semidefinita** positiva (negativa) se $Q(\bar{x}) \geq 0 (\leq 0)$ ma esistono vettori che annullano la forma quadratica.
2. **Definita** positiva (negativa) se $Q(\bar{x}) \geq 0 (\leq 0), \quad Q(\bar{x}) = 0 \Longleftrightarrow \bar{x} = 0$
3. **Indefinita** se $Q(\bar{x})$ ha segno variabile al variare dei vettori di V.

Diciamo che una forma bilineare è *degenere* se il suo nucleo non è nullo.

Una fbs definita positiva non può mai essere degenere, infatti:

$$Q(\bar{x}) = 0 \Longleftrightarrow \bar{x} = 0$$

quindi ha nucleo nullo.

8.3 Cauchy-Schwarz e Minkowski

Ora dobbiamo introdurre due disuguaglianze fondamentali per comprendere alcuni concetti chiave:

Disuguaglianza di Cauchy-Schwarz

Se una forma bilineare simmetria è definita positiva, allora vale

$$|\phi(\bar{x},\bar{y})| \leq \sqrt{\phi(\bar{x},\bar{x})}\sqrt{\phi(\bar{y},\bar{y})}$$

e quindi

$$\left[\phi(\bar{x},\bar{y})\right]^2 \leq Q(\bar{x})Q(\bar{y}), \quad \forall \bar{x},\bar{y} \in V$$

dimostrazione

$$Q(\lambda\bar{x}+\bar{y}) = \lambda^2 Q(\bar{x}) + 2\lambda\phi(\bar{x},\bar{y}) + Q(\bar{y})$$

Essendo $\phi(\bar{x},\bar{y})$ definita positiva si ha che $Q(x) \geq 0$. Quindi il discriminante ridotto $\left(\frac{b}{2}\right)^2 - ac$ del trinomio deve essere negativo (se lo immagina come una parabola sempre con ordinata positiva con concavità verso l'alto, questa non deve può avere intersezioni con l'asse x):

$$\left[\phi(\bar{x},\bar{y})\right]^2 \leq Q(\bar{x})Q(\bar{y})$$

Si noti che se $\bar{x}$ è un vettore isotropo, quindi vale $Q(\bar{x}) = 0$, allora

$$\left[\phi(\bar{x},\bar{y})\right]^2 \leq 0 \cdot Q(\bar{y}) \quad \rightarrow \quad \phi(\bar{x},\bar{y}) = 0$$

che mostra che rispetto a forme bilineari simmetriche semidefinite positive *se un vettore è isotropo, allora è anche nel nucleo*. Questo è esteso alle definite positive per le quali vale $Q(\bar{x}) = \phi(\bar{x},\bar{y}) = 0$, il nucleo e il cono isotropo coincidono nel vettore nullo.

Poiché il kernel è un sottospazio vettoriale, se il cono isotropo coincide con esso anch'esso lo è. L'unico caso in cui il cono isotropo non è uno spazio vettoriale e se riferito a una forma bilineare simmetrica *indefinita*.

Diciamo di avere due vettori $\bar{x}, \bar{y}$ isotropi:

$$Q(\bar{x}) = 0, \quad Q(\bar{y}) = 0$$

$$Q(\bar{x}+\bar{y}) = Q(\bar{x}) + 2\phi(\bar{x},\bar{y}) + Q(\bar{y}) = 0 + 2\phi(\bar{x},\bar{y}) + 0$$

non so se il termine in ϕ vale zero, quindi il cono isotropo non è chiuso per la somma; lo sarebbe se $\bar{x}$ o $\bar{y}$ appartenessero al nucleo di ϕ.

Disuguaglianza di Minkowsky

$$\sqrt{Q(\bar{x}+\bar{y})} \leq \sqrt{Q(\bar{x})} + \sqrt{Q(\bar{y})}$$

8.4 Teorema di esistenza di una forma canonica

Data la forma bilineare simmetrica $\phi: V \times V \to \mathbb{R}$, l'obiettivo è quello di definire un buon prodotto scalare in modo che l'endomorfismo definito da

$$f(\bar{x}) \cdot \bar{y} = \phi(\bar{x},\bar{y})$$

sia **autoaggiunto**. Se l'endomorfismo è autoaggiunto posso utilizzare il teorema spettrale per diagonalizzarne la matrice associata (la stessa associata a ϕ).

Prendo una base $\mathscr{B} = \{\bar{v}_1 \dots \bar{v}_n\}$ e definisco due vettori nello spazio

$$\bar{x} = x_1\bar{v}_1 + \dots + x_n\bar{v}_n$$

$$\bar{y} = y_1\bar{v}_1 + \dots + y_n\bar{v}_n$$

Definisco ora un prodotto scalare:

$$\bar{x} \cdot \bar{y} = x_1y_1 + x_2y_2 + \dots + x_ny_n$$

Rispetto a questo prodotto scalare la nostra **base è ortonormale**.

Definisco ora l'endomorfismo $f: V \to V$ determinato univocamente da $f(\bar{v}_i)$. Trovo le componenti di $f(\bar{v}_i)$ facendo $f(\bar{v}_i) \cdot \bar{v}_j$ che abbiamo detto essere uguale a $\phi(\bar{v}_i, \bar{v}_j)$. Si ha quindi:

$$f(\bar{v}_i) \cdot \bar{v}_j = \phi(\bar{v}_i, \bar{v}_j) = {}^tXAY$$

Le componenti le esplicitiamo come:

$$f(\bar{v}_1) = \phi(\bar{v}_1, \bar{v}_1)\,\bar{v}_1 + \cdots + \phi(\bar{v}_1, \bar{v}_n)\,\bar{v}_n$$

$$f(\bar{v}_2) = \phi(\bar{v}_2, \bar{v}_1)\,\bar{v}_1 + \cdots + \phi(\bar{v}_2, \bar{v}_n)\,\bar{v}_n$$

$$\vdots$$

Abbiamo quindi specificato f in modo unico e vale

$$f(\bar{x}) \cdot \bar{y} = \phi(\bar{x}, \bar{y}) \quad \forall \bar{x}, \bar{y}$$

con

$$M^{\mathscr{B}}(f) = {}^{t}A = A \quad \Longrightarrow \quad f \text{ autoaggiunto} \quad \Longrightarrow \quad \exists \mathscr{B}' \text{ ortonormale di autovettori}$$