Le pagine che seguono presentano una sintesi estrema di numerosi argomenti utilizzabile per presentazioni rapide e concentrate.

ELEMENTI DI GEOMETRIA

Prof. M. Spera

Lezione VIII

Complementi di algebra lineare

Sinossi

Teorema di Sylvester } Parte I

Teorema spettrale Parte II

PARTE I

☆ Richiami sulle forme quadratiche

$(V, \langle\,|\,\rangle)$ spazio euclideo (reale) $\dim V = n$

Sia $A : V \to V$ un operatore lineare (un endomorfismo, o omomorfismo di V)

A è detto simmetrico se, $\forall\, x, y \in V$, si ha:

$$\boxed{\langle x \,|\, A y \rangle = \langle A x \,|\, y \rangle} \quad (:= \langle x \,|\, A^t y \rangle)$$

ovvero, si confonde con l'operatore lineare trasposto A^t.

In termini di una base ortonormale di V, sia questa $\{e_1, e_2, \dots, e_n\}$, la condizione precedente equivale a

$$\langle e_i \,|\, A e_j \rangle = \langle A e_i \,|\, e_j \rangle \quad (:= \langle e_i \,|\, A^t e_j \rangle)$$

a_{ij} : elemento ij della matrice corrispondente ad A $\qquad \langle e_j \,|\, A e_i \rangle = a_{ji} \qquad a^t_{ij}$

$$\Rightarrow \quad \boxed{a_{ij} = a_{ji}} \quad (:= a^t_{ij})$$

$\Rightarrow$ ☆ la matrice (a_{ij}) associata ad A è simmetrica.

☆ un vettore $v \in V$, $v \neq 0$ si dice autovettore di A, corrispondente all'autovalore $\lambda \in \mathbb{R}$ se $\quad A v = \lambda v$

gli autovalori si ottengono risolvendo l'equazione

$$\boxed{\det(A - \lambda I) = 0}$$

ovvero, sono le radici del polinomio caratteristico di A.

In generale $\lambda \in \mathbb{C}$, ma se A è simmetrica, tutti gli autovalori sono reali e sono n, se contati con la dovuta molteplicità, per il teorema fondamentale dell'algebra.

Un operatore simmetrico è (ortogonalmente) diagonalizzabile, ovvero esiste una base ortonormale costituita da autovettori di A.

[Si osservi che autovettori corrispondenti ad autovalori distinti sono ortogonali:

$$A v_1 = \lambda_1 v_1 \; ; \; A v_2 = \lambda_2 v_2 \quad \Rightarrow$$

$$\langle v_2 | A v_1 \rangle = \lambda_1 \langle v_2 | v_1 \rangle$$

$$\| $$

$$\langle v_1 | A v_2 \rangle = \lambda_2 \langle v_1 | v_2 \rangle$$

simmetria $\Rightarrow (\lambda_1 - \lambda_2) \langle v_1 | v_2 \rangle = 0 \Rightarrow \langle v_1 | v_2 \rangle = 0$.]

Consideriamo la forma quadratica associata ad A, ovvero la funzione

$Q_A : V \to \mathbb{R}$ data da

$$\boxed{Q_A(v) = \langle v | A v \rangle} \; (= \langle A v | v \rangle) =$$

$$= \boxed{\sum_{\substack{i=1 \\ j=1}}^{n} a_{ij} v_i v_j} \, , \qquad (a_{ji} = a_{ij})$$

$$\left(\text{se } v = \sum_{i=1}^{n} v_i e_i \, , \quad (\{e_i\} \text{ ortonormale}) \right)$$

[Limiteremo le nostre ulteriori considerazioni specialmente al caso $n = 2$.

Se $n = 2$: $A \to (a_{ij}) \to \begin{pmatrix} a & b \\ b & d \end{pmatrix}$]

In termini di una base ortonormale di autovettori, è

$$Q_A(v) = \sum_{i=1}^{n} \lambda_i v_i^2$$

per $n = 2$ è $\boxed{Q_A(v) = \lambda_1 v_1^2 + \lambda_2 v_2^2}$

☆ gli autovalori di A (simmetrica) sono riscattabili da una "caratterizzazione variazionale", importantissima nelle applicazioni.

La dimostriamo nel caso $n = 2$.
(la ritroveremo poi con la teoria dei moltiplicatori di Lagrange)

Vincoliamo Q_A sulla circonferenza unitaria

$$S^1 = \{ v \in V \; / \; v_1^2 + v_2^2 = 1 \}$$

(Scegliamo una volta per tutte una base ortonormale di autovettori).

Sia $\lambda_1 \leq \lambda_2$. Se $\lambda_1 = \lambda_2$ è $Q_A(v) =$
$= \lambda_1$ (= costante).

Sia $\lambda_1 \le \lambda_2$, $\|v\|=1$

$$Q_A(v) = \lambda_1 v_1^2 + \lambda_2 v_2^2 =$$
$$= \lambda_1 (v_1^2 + v_2^2) + (\lambda_2 - \lambda_1) v_2^2 =$$
$$= \lambda_1 + (\lambda_2 - \lambda_1) v_2^2$$

$$\Rightarrow \boxed{\lambda_1 = \min_{\|v\|=1} Q_A(v)} \quad \left(= \min_{v \neq 0} \frac{Q_A(v)}{\|v\|^2} \right)$$

quoziente di Rayleigh

Analogamente

$$\boxed{\lambda_2 = \max_{\|v\|=1} Q_A(v)} \quad \left(= \max_{v \neq 0} \frac{Q_A(v)}{\|v\|^2} \right)$$

☆ Definizioni Q_A è detta

equiv: $Q_A(v) > 0$ e $= 0$ sse $v = 0$ ... ecc

* definita positiva se $Q_A(v) > 0 \quad \forall v \neq 0$
* definita negativa se $Q_A(v) < 0 \quad \forall v \neq 0$
* indefinita se assume sia valori positivi che negativi

In tali casi Q_A si dice non degenere

Ovviamente
Q_A è def. pos. $\Leftrightarrow 0 < \lambda_1 \le \lambda_2$
Q_A è def. neg. $\Leftrightarrow \lambda_1 \le \lambda_2 < 0$
Q_A è indef. $\Leftrightarrow \lambda_1 < 0 < \lambda_2$

Se $0 = \lambda_1 < \lambda_2$ Q_A si dice semidefinita positiva
Se $\lambda_1 < \lambda_2 = 0$ Q_A = semidef. negativa
(Se $\lambda_1 = \lambda_2 = 0$ $Q_A \equiv 0$)

☆ Teorema Sia Q_A la forma quadratica associata a $A: V \to V$ (dim $V = 2$)

$A \leftrightarrow \begin{pmatrix} a & b \\ b & d \end{pmatrix}$ rispetto ad una certa base ortonormale

1) Q_A è definita positiva $\Leftrightarrow$ $\det A = ad - b^2 > 0$ e $a > 0$ (oppure $d > 0$)
$\lambda_1 > 0,\ \lambda_2 > 0$ [e ciò implica $d > 0$; risp. $a > 0$]

2) Q_A è definita negativa $\Leftrightarrow$ $\det A > 0$ e $a < 0$ (oppure $d < 0$)
$\lambda_1 < 0,\ \lambda_2 < 0$ e ciò implica $d < 0$; risp. $a < 0$

3) Q_A è indefinita $\Leftrightarrow$ $\det A < 0$
$\lambda_1 < 0,\ \lambda_2 > 0$

1ª dim. $\det A = \lambda_1 \lambda_2$
$\operatorname{tr} A = \lambda_1 + \lambda_2 = a + d$
(traccia)

3) è ovvia

Dimostriamo 1) (2) è analoga)

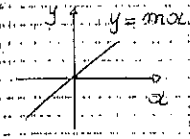
$\Rightarrow$ $\lambda_1 > 0,\ \lambda_2 > 0 \Rightarrow \det A > 0$ e $a + d > 0$
$\Rightarrow ad > b^2$ e $a + d > 0 \Rightarrow ad > 0$ e $a + d > 0$
$\Rightarrow a > 0$ e $d > 0$

$\Leftarrow$ $\det A > 0$ e $a > 0 \Rightarrow ad > b^2$ e $a > 0$
$\Rightarrow ad > 0$ e $a > 0 \Rightarrow d > 0 \Rightarrow$
$a + d = \lambda_1 + \lambda_2 > 0$ e $\lambda_1 \lambda_2 > 0$
$\Rightarrow \lambda_1 > 0,\ \lambda_2 > 0$

Accenniamo ad una 2ª dim. basata sulla def. diretta:

poniamo $v_1 = x$ $v_2 = y$

$$\boxed{Q_A(v) = ax^2 + 2bxy + dy^2}$$



$= (\text{se } y = mx) =$

$$= x^2\left[a + 2bm + dm^2\right]$$

$v \neq 0 \Rightarrow x \neq 0$ (a meno che ...)

Esaminiamo il polinomio

$$\boxed{dm^2 + 2bm + a}$$

$$\frac{\Delta}{4} = b^2 - ad$$

$\begin{cases} z = ax^2 + 2bxy + dy^2 \\ y = mx \end{cases}$ piano per l'asse z

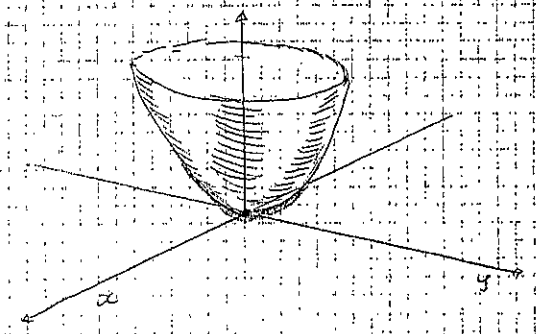
$\Rightarrow$ le nostre conclusioni seguono facilmente ... $\square$

Interpretazione geometrica

$$\boxed{z = Q_A(v) = \lambda_1 x^2 + \lambda_2 y^2}$$

Q_A def. positiva ($\Leftrightarrow \lambda_1 \leq 0, \lambda_2 \leq 0$)

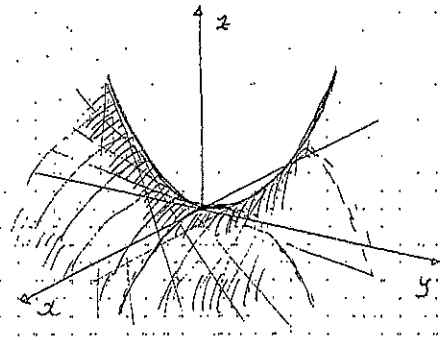
$\Rightarrow$ paraboloide ellittico



Q_A def. negativa ... grafico concavo

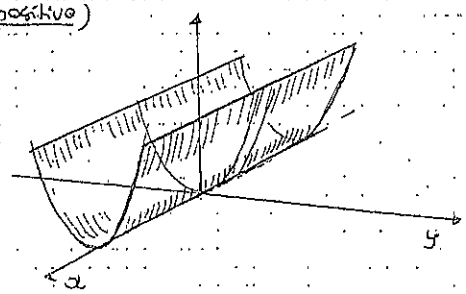
Q_A indefinita $\lambda_1 < 0$ $\lambda_2 > 0$

$\Rightarrow$ paraboloide iperbolico ("sella")



Q_A semidefinita : cilindro parabolico

(caso positivo)



☆ Sia $A = \begin{pmatrix} a & b \\ b & d \end{pmatrix}$ simmetrica

Allora i suoi autovalori sono <u>reali</u>.

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

$$0 = \begin{vmatrix} \lambda - a & -b \\ -b & \lambda - d \end{vmatrix} = (\lambda - a)(\lambda - d) - b^2 = 0$$

$$\Rightarrow \lambda^2 - (a+d)\lambda + ad - b^2 = 0$$

$$\Rightarrow \Delta = (a+d)^2 - 4ad + 4b^2 =$$

$$= (a-d)^2 + 4b^2 \geq 0$$

$\square$

Nota: sul tracciamento delle curve di livello di una forma quadratica

$$q_A : \mathbb{R}^2 \to \mathbb{R}$$

$$q_A(v) = v^t A v$$

Siano λ_1, λ_2 gli autovalori di A

Assumiamo $r(A) = 2$, sicchè la segnatura è $(2,0)$ def. pos
$(0,2)$ def. neg
$(1,1)$ indef.

Data una base di autovettori di A: $A e_i' = \lambda_i e_i'$ $\quad i = 1,2$
ortonormale

è

$$q_A(\underset{\substack{\| \\ x'e_1' + y'e_2'}}{v}) = \lambda_1 x'^2 + \lambda_2 y'^2 \qquad \text{(forma can. metrica)}$$

Sia, in ogni caso $\lambda_1 < \lambda_2$

$\boxed{1^\circ \text{ Caso}}$ si ponga $0 < \lambda_1 < \lambda_2$ (tipo $(2,0)$) def. positiva

una discussione analoga ha luogo nel caso $(0,2)$

Curve di livello: ellissi

$$\lambda_1 x'^2 + \lambda_2 y'^2 = c \geq 0$$

poniamo $c > 0$ (per $c = 0$ si trova l'origine)

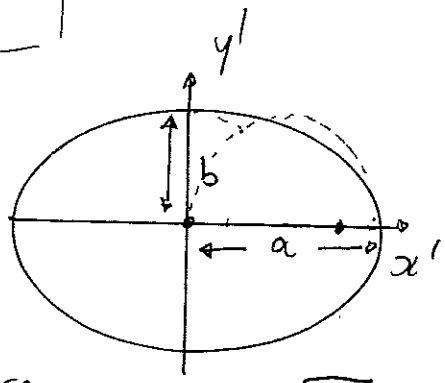
$$\lambda_1 x'^2 + \lambda_2 y'^2 = c \qquad c>0$$

$\Leftrightarrow$

$$\boxed{\frac{x'^2}{\left(\sqrt{\frac{c}{\lambda_1}}\right)^2} + \frac{y'^2}{\left(\sqrt{\frac{c}{\lambda_2}}\right)^2} = 1}$$

cf l'eq. canonica

curve di livello



$$a = \sqrt{\frac{c}{\lambda_1}} > b = \sqrt{\frac{c}{\lambda_2}}$$

semiasse maggiore semiasse minore

$c = \sqrt{a^2-b^2}$

Si osservi che $\dfrac{a}{b} = \sqrt{\dfrac{\lambda_2}{\lambda_1}}$

$\boxed{2^\circ \text{ Caso}}$

poniamo:

$$\lambda_1 < 0 < \lambda_2$$

$$\lambda_1 x'^2 + \lambda_2 y'^2 = c$$

Sia $c > 0$. Poniamo successivamente

$$\overset{>0}{\lambda_2} y'^2 - \overset{>0}{(-\lambda_1)} x'^2 = c$$

$$\boxed{\frac{y'^2}{\Big(\underbrace{\sqrt{\frac{c}{\lambda_2}}}_{a}\Big)^2} - \frac{x'^2}{\Big(\underbrace{\sqrt{\frac{c}{-\lambda_1}}}_{b}\Big)^2} = 1}$$

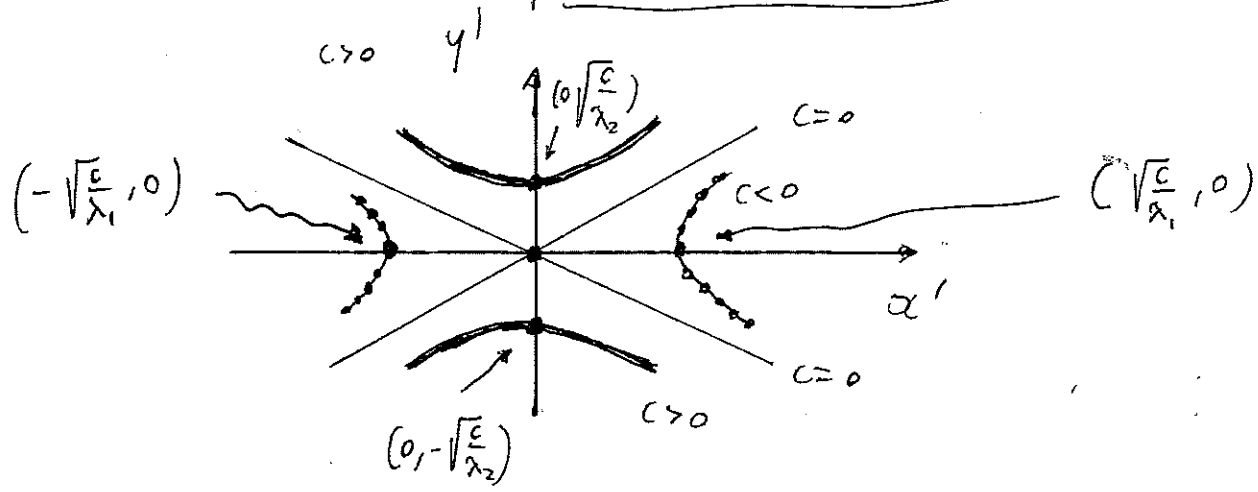
nel grafico:

Asse y' : asse focale
(ovvero $x'=0$)

Se $c=0$ $\quad \lambda_2 y'^2 - (-\lambda_1) x'^2 = 0$

asintoti :
(cono isotropo di q_A..)

$$\boxed{y' = \pm\sqrt{\frac{-\lambda_1}{\lambda_2}}\, x'} \qquad y' =$$

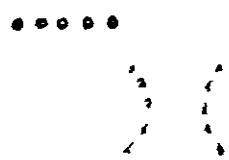


Se $c<0$ poniamo

$$(-\lambda_1)x'^2 - \lambda_2 y'^2 = -c \qquad -c>0$$

$$\Leftrightarrow \quad \boxed{\frac{x'^2}{\left(\sqrt{\frac{c}{\lambda_1}}\right)^2} - \frac{y'^2}{\left(\sqrt{\frac{-c}{\lambda_2}}\right)^2} = 1}$$

nel grafico:



asse focale : $y'=0$ $\qquad$ asintoti : $y' = \pm\sqrt{-\frac{\lambda_1}{\lambda_2}}\, x'$

(gli stessi....)

- $\mathbb{R}^7$

$$T: \begin{pmatrix} 0 & & & & & & \\ & 0 & 1 & & & & \\ & 0 & 0 & & & & \\ & & & 0 & 1 & 0 & \\ & & & 0 & 1 & 1 & \\ & & & 0 & 0 & 0 & \\ & & & & & & 0 \end{pmatrix} \qquad p_T(\lambda) = (-\lambda)^7 = -\lambda^7$$

$m_a(0) = 7 \qquad m_g(0) = 3$ (verificare)

(è un esempio di operatore nilpotente: $T^k = 0$ per qualche k. Qui, il più piccolo k è $= 4$ verificare)

- Siano P_1, P_2 proiezioni (in V, $\dim V = n$)

si ha $P_1 \sim P_2 \iff \rho(P_1) = \rho(P_2)$

Dim. una proiezione è certamente diagonalizzabile; in una base adattata è

$$m(P) = \left(\begin{array}{c|c} I_\rho & \\ \hline & 0 \end{array}\right) \qquad \rho = \rho(P)$$

$$p_c^P(\lambda) = (1-\lambda)^{\rho(P)} (-\lambda)^{n-\rho(P)} = (-1)^n \lambda^n (1-\lambda)^\rho,$$

da cui la conclusione.

Forme bilineari

A titolo di motivazione si veda la trattazione del prodotto scalare elementare del capitolo precedente

Sia dato (V, K), $\dim V = n \geq 1$,

$K = \mathbb{C}, \mathbb{R}$ ($\mathbb{Q}$)

Def. Una forma bilineare su V è una funzione

$$b: V \times V \to K$$
$$(v, w) \mapsto b(v, w).$$

tale che, $\forall v_i, w_i \in V$, $\forall \alpha_i, \beta_i \in K$ si abbia

$$\begin{cases} b(\alpha_1 v_1 + \alpha_2 v_2, w) = \alpha_1 b(v_1, w) + \alpha_2 b(v_2, w) \\ b(v, \beta_1 w_1 + \beta_2 w_2) = \beta_1 b(v, w_1) + \beta_2 b(v, w_2) \end{cases}$$

(ovvero, b è lineare in entrambi gli argomenti)

Oggetti siffatti esistono: basta prendere $b(v, w) = 0$ $\forall v, w$ (forma bil. NULLA)

L'insieme delle forme bilineari, $T^2(V^*)$, acquista in modo naturale una struttura di spazio vettoriale su K, qualora si definisca ($\forall b_1, b_2 \in T^2(V^*)$, $\beta_1, \beta_2 \in K$

$$(\beta_1 b_1 + \beta_2 b_2)(v, w) := \beta_1 b_1(v, w) + \beta_2 b_2(v, w)$$

da definirsi — specifica l'azione su v e w arbitrari — operazioni in K

b è detta <u>simmetrica</u> se vale

inoltre

$$(**) \qquad b(w,v) = b(v,w) \qquad \forall v,w \in V.$$

<u>antisimmetrica</u> se $(**)$ è sostituita da

$$(***) \qquad b(w,v) = -b(v,w).$$

Notazioni: forme bil. simmetriche: $S^2(V^*)$

= = antisimmetriche $\Lambda^2(V^*)$

(sono entrambi s.spazi vettoriali di $T^2(V^*)$)

Fissata ora una base $e = (e_1 \dots e_n)$ in V, risulta chiaro che b è determinata dai valori $b(e_i, e_j)$ $i,j = 1 \dots n$

Posto infatti $v = \sum_{i=1}^n x_i e_i$ $\quad w = \sum_{j=1}^n y_j e_j$

si ha:

$$b(v,w) = b\left(\sum_i x_i e_i, \sum_j y_j e_j\right) = \text{(bilinearità)}$$

$$\sum_{i,j} x_i y_j \, b(e_i, e_j).$$

Poniamo allora $A = (a_{ij}) \in M_n(K)$

con $a_{ij} = b(e_i, e_j)$,

sicché

$$b(v,w) = \sum_{i,j} a_{ij} x_i y_j = \sum_{i=1}^n x_i \left(\sum_{j=1}^n a_{ij} y_j\right) \text{ ovvero}$$

$$\boxed{b(v,w) = x^t A y} \quad \text{ove} \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

$$\underset{1\times n}{x^t} \; \underset{n\times n}{A} \; \underset{n\times 1}{y}$$

$$x^t = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix} \qquad A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & & & \\ a_{n1} & \dots & \dots & a_{nn} \end{bmatrix} \qquad y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

Si dice che $A \in M_n(K)$ è la <u>matrice rappresentativa</u> della b rispetto alla base $e = (e_1 \dots e_n)$.

Viceversa, data $A \in M_n(K)$ e fissata una base $e = (e_1 \dots e_n)$ in V, si può costruire in modo ovvio una forma bilineare (basta porre $b(e_i, e_j) := a_{ij}$)

la cui matrice corrispondente (procedendo a ritroso), sempre rispetto alla base e, sarà chiaramente A.

È chiaro altresì che b è <u>simmetrica</u>

$\Leftrightarrow A = A^t \quad (a_{ij} = a_{ji})$, e <u>antisimmetrica</u>

$\Leftrightarrow A = -A^t \quad (a_{ij} = -a_{ji})$ (ciò sarà vero per una qualsiasi matrice rappresentativa, v. oltre)

Osserviamo esplicitamente che, data $A \in M_n(K)$, essa può direttamente interpretarsi come forma bilineare $b \equiv b_A$ rispetto alla base canonica di K^n (di cui, ovviamente, è la matrice corrispondente); ciò è analogo a quanto facevamo con gli omomorfismi.

◇ ◇ ◇

Esaminiamo ora l'effetto di un cambiamento di base su $T^2(V^*)$:

Sia $M : e \longmapsto e' \qquad M \in GL(V)$

$$\left(M = \underset{\substack{\| \\ m_{ee'}(I)}}{M_{ee'}} = \underset{\substack{\| \\ m_{e'e}(I)}}{M_{e'e}}^{-1} \right)$$

È, ricordiamo $\quad x = Mx'$.

[le colonne di M dànno le componenti dei vettori della nuova base e' rispetto alla vecchia base e e pertanto le vecchie coordinate (x) in funzione delle nuove (x')].

Successivamente, si ha:

$$b(v, w) = x^t A y = (Mx')^t A (My') =$$

$$= x'^t (M^t A M) y' \equiv x'^t B y'$$

ove si sia posto

$$B = M^t A M \qquad \underline{\underline{M \in GL(V)}}$$

◇ <u>Def.</u>

Due matrici quadrate legate dalla (4) si dicono <u>congruenti</u>; si scrive

$$A \approx B$$

Si verifica subito che $\approx$ è una relazione di equivalenza (da <u>non confondersi</u> con la similitudine!)

$A \approx A$: basta prendere $M = I$

$A \approx B \Rightarrow B \approx A$: se $B = M^t A M$, è

$$(M^t)^{-1} B M^{-1} = A \Rightarrow A = (M^{-1})^t B M^{-1}$$

$A \approx B$, $B \approx C$ $\Rightarrow$ $A \approx C$:

infatti da $B = M^t A M$,

$C = N^t B N$ $(M, N \in GL(V))$

si ha $C = N^t M^t A M N = (MN)^t A (MN)$.

◇ Def. Si definisce rango di b (not: $r(b)$)

come $r(b) := r(A)$

A matrice rappresentativa qualsiasi;

la def. è ben posta in virtù della ☆:

$$r(B) = r(\underbrace{M^t A M}_{\in GL(V)}) = r(A)$$

Pertanto $A \approx B$ $\Leftrightarrow$ esse corrispondono ad una stessa b rispetto a basi diverse

Se $r(b) = \max\ (= n)$ b è detta non degenere.

Analogamente a quanto si è visto per la similitudine, si può introdurre il concetto di congruenza ad un livello più astratto.

Sia $M \in GL(V)$

Poniamo, per ogni $b \in T^2(V^*)$

$$b_M(v, w) := b(Mv, Mw)$$

(facciamo agire $GL(V)$ su $T^2(V^*)$).

Diciamo allora che $b_1 \approx b_2$ (b₁ congruente a b₂) se $\exists\ M \in GL(V)$ tale che $(b_1)_M = b_2$

Vale a dire: $\forall v, w \in V$: $b_1(Mv, Mw) = b_2(v, w)$

(di nuovo si ha una relazione di equivalenza)

Concentriamoci ora su $S^2(V^*)$

(forme bilineari simmetriche)

Vogliamo determinarne un invariante completo per congruenza

(ed esaminare la stessa questione per le matrici simmetriche)

Osservazione: ricordiamo che poiché la simmetria è un concetto inerente a b, è chiaro che ogni matrice rappresentativa di b risulterà simmetrica a sua volta!

$$b: \begin{cases} A & \text{base } e \\ B & \text{base } e' \end{cases}$$

$$B = M^t A M$$

Sia $A = A^t$. È allora (a titolo di esercizio verificiamolo anche così)

$$B^t = (M^t A M)^t = M^t A^t (M^t)^t =$$
$$= M^t A^t M = M^t A M \quad (A = A^t)$$
$$= B$$

In altre parole, vogliamo descrivere completamente la classe di congruenza di una matrice biunivocetrica o di una forma bilineare simmetrica

◇ Def. Data $b \in S^2(V^*)$, due vettori $v, w \in V$ sono detti b-ortogonali (o semplicemente ortogonali, ove non vi sia tema di confusione) se

$$b(v,w) = b(w,v) = 0$$

Può accadere che $b(v,v) = 0$ anche se $v \neq \underline{0}$ (per $v = \underline{0}$ è ovvio)

◇ Un tale vettore è detto isotropo (per b)

(tale concetto ricoprirà un ruolo fondamentale per ciò che segue e, in seguito, sarà cruciale la sua utilizzazione per definire intrinsecamente coniche e quadriche..

◇ Dunque: $v \neq \underline{0}$ è isotropo se $b(v,v) = 0$

Dato un sottoinsieme $S \subseteq V$, si dice b-ortogonale di S, e si denota con $S^{\perp_b}$ (si dice anche "S b-ortogonale, o S ortogonale" (se non c'è pericolo di confusione)) il sottoinsieme

$$S^{\perp_b} = \{ w \in V \mid b(v,w) = b(w,v) = 0 \ \forall v \in S \}$$

ovvero, $S^{\perp_b}$ consiste di tutti i vettori

b-ortogonali a tutti i vettori di S.

Altra terminologia: $S^{\perp_b}$ spazio polare di S; tale terminologia ha origine nella geometria proiettiva

Si verifica subito che $S^{\perp_b}$ è un sottospazio vettoriale di V, anche se S non lo è:

$$S^{\perp_b} \leq V$$

Infatti: siano $w_1, w_2 \in S^{\perp_b}$:

$$b(w_1, v) = b(w_2, v) = 0 \quad \forall\, v \in S$$

Si considera $\alpha_1 w_1 + \alpha_2 w_2$ ($\alpha_i \in K$ arbitrari)

Si ha immediatamente

$$b(\alpha_1 w_1 + \alpha_2 w_2, v) = \alpha_1 \overbrace{b(w_1, v)}^{0} + \alpha_2 \overbrace{b(w_2, v)}^{0}$$
$$= 0 \quad \forall\, v \in S$$

sicché $\alpha_1 w_1 + \alpha_2 w_2 \in S^{\perp_b}$ □

Osservazione: $v \in V$ è isotropo $\Leftrightarrow$ $\langle v \rangle \leq \langle v \rangle^{\perp_b}$. (eq: $\{v\} \leq \{v\}^{\perp_b}$)

(ciò è importante)

(ovvero: $\langle v \rangle$ è contenuto nel suo spazio polare)

◇ Def. Sia data $b \in S^2(V^*)$; una base $e = (e_1 \dots e_n)$ di V è detta b-ortogonale (o, semplicemente, ortogonale, ove ciò non generi confusione), o diagonalizzante se consiste di vettori mutuamente b-ortogonali:

$$b(e_i, e_j) = \beta_i \delta_{ij} \quad \left(= \begin{matrix} \beta_i & \text{se } i=j \\ 0 & \text{se } i \neq j \end{matrix} \right)$$

per $\beta_i \in K$, ovvero

$$A = \begin{pmatrix} \beta_1 & & 0 \\ & \ddots & \\ 0 & & \beta_n \end{pmatrix}$$

(i β_i possono essere in parte (o tutti) nulli

☆ equivalentemente la sua matrice rappresentativa corrispondente è diagonale.

⚠ Non stiamo parlando in alcun modo di autovettori e autovalori: gli oggetti in gioco sono forme bilineari simmetriche, non endomorfismi.

Vogliamo far vedere che, per i K da noi usati, è sempre possibile determinare una base diagonalizzante; in altre parole si ha il

Teorema Sia (V, K) $(K = \mathbb{R}, \mathbb{C}, \mathbb{Q})$ $\dim V \geq 1$.

i) Data $b \in S^2(V^*)$ (forma bilineare simmetrica), esiste (*) una base b-ortogonale (diagonalizzante); ovvero, ogni $b \in S^2(V^*)$ è diagonalizzabile

ii) In termini matriciali, ogni matrice simmetrica $A \in M_n(K)$ $(A = A^t)$ è congruente ad una matrice diagonale

(*) Non è unica.

Premettiamo alla dimostrazione due lemmi

Lemma 1 Sia $b \in S^2(V^*)$ non identicamente nulla. Allora essa ammette almeno un vettore non isotropo (si può anche dire anisotropo)

Dim. Per ipotesi esistono $v, w \in V$ (necessariamente $\neq 0$) tali che $b(v, w) \neq 0$.

Consideriamo i tre vettori v, w, $v + w$.

Tra questi ve ne è almeno uno non isotropo.

Infatti, se, per fissare le idee, v e w fossero isotropi, risulterebbe

$$b(v+w, v+w) = \underbrace{b(v,v)}_{=0} + 2\,b(v,w) + \underbrace{b(w,w)}_{=0}$$

(simmetria)

$$= 2\,b(v,w) \neq 0$$

Sicché $v + w$ sarebbe comunque non isotropo.

Lemma 2 Sia $b \in S^2(V^*)$ non identicamente nulla e sia v $(\neq \underline{0})$ un vettore non isotropo $(b(v,v) \neq 0)$. Allora

$$V = \langle v \rangle \oplus \langle v \rangle^{\perp}$$

(ovvero $\langle v \rangle$ e il suo spazio polare $\langle v \rangle^{\perp}$ sono in somma diretta e generano V.

Dim. È chiaro, in virtù della definizione di vettore isotropo, che se $w \in \langle v \rangle \cap \langle v \rangle^{\perp}$, è necessariamente $w = \underline{0}$:

infatti, se $w \neq \underline{0}$, è $w = \lambda v$, $\lambda \neq 0$,

e $b(w,w) = \lambda^2 b(v,v) = 0 \Rightarrow b(v,v) = 0$

contro l'ipotesi. I due sottospazi sono pertanto in somma diretta

D'altra parte è anche, $\forall\, w \in V$,

$$w = \underbrace{w - \frac{b(v,w)}{\underbrace{b(v,v)}_{\neq 0}} v}_{w_1} + \underbrace{\frac{b(v,w)}{b(v,v)} v}_{\in \langle v \rangle}$$

ma $w_1 := w - \dfrac{b(v,w)}{b(v,v)} v \in \langle v \rangle^{\perp}$

Infatti $b(w_1, \xi v) =$

$$= \xi\, b(w_1, v) = \xi\, b\left(w - \frac{b(v,w)}{b(v,v)} v, v\right) =$$

$$= \xi \left[b(v,w) - \frac{b(v,w)\, b(v,v)}{b(v,v)} \right] = 0 \qquad \forall \xi \in K$$

I due sottospazi, pertanto, generano V.

Ciò conclude la dimostrazione. □

Dimostrazione del teorema

La dimostrazione si fa per induzione su $n = \dim V$, e basta supporre b non identicamente nulla (in tal caso ogni base va bene); inoltre $n \geq 1$ (altrimenti non vi è nulla da dire.

Per $n = 1$ il teorema è vero. (verificarlo!).

Supponiamolo vero per $n-1$, dimostriamolo vero per n; concluderemo allora in base al principio di induzione.

Sia allora e_1 $(\neq \underline{0})$ un vettore non isotropo (Lemma 1). Allora è

$$V = \langle e_1 \rangle \oplus \langle e_1 \rangle^{\perp}$$

(Lemma 2).

Restringendo b a $\langle e_1 \rangle^{\perp}$ (che ha dimensione $n-1$) e applicando l'ipotesi induttiva, si trova una base $(e_2, e_3 \dots e_n)$ di $\langle e_1 \rangle^{\perp}$

che ivi diagonalizza b:

$$b(e_i, e_j) = \beta_i \delta_{ij} \qquad \begin{array}{l} i = 2, \dots n \\ j = 2, \dots n \end{array}$$

notare

Ma, d'altra parte (per definizione) è

$$b(e_1, e_j) = 0 \qquad j = 2, \dots n$$

Perciò, posto $\beta_1 := b(e_1, e_1)$ $(\neq 0)$

si ha, in definitiva

$$b(e_i, e_j) = \beta_i \delta_{ij} \qquad \begin{array}{l} i = 1 \dots n \\ j = 1 \dots n \end{array}$$

ovvero, $e = (e_1, e_2 \dots e_n)$ diagonalizza b.

$$A = \left(\begin{array}{c|ccc} \beta_1 & 0 & 0 \cdots & 0 \\ \hline 0 & \beta_2 & & 0 \\ 0 & & \ddots & \\ \vdots & 0 & & \beta_n \\ 0 & & & \end{array}\right)$$

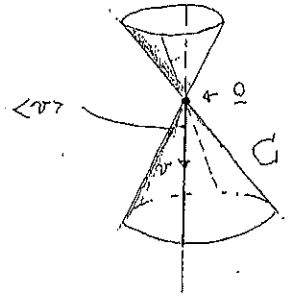
★ Ripetiamo, la stessa dimostrazione ci dice che $A \in M_n(K)$, $A = A^t$ (A simmetrica) è diagonalizzabile nel senso della congruenza (in seguito faremo vedere di più): $\exists\ M \in GL_n(K)$ tale che $M^t A M = D$, con D diagonale ($a_{ij} = 0$ per $i \neq j$)

Commenti i) La dimostrazione precedente è semplice e rapida, ma i due lemmi da cui scaturisce appaiono per il momento poco motivati: vedremo poi il loro contenuto geometrico.

Per il momento osserviamo che l'insieme (*) dei vettori isotropi (e del vettore nullo) non è uno spazio vettoriale, in generale, ma, per la sua struttura geometrica, da enucleare in futuro, è tuttavia importante: esso è detto ☆ cono isotropo.

I suoi sottoinsiemi della forma $\langle v \rangle$, con v vettore isotropo sono detti generatrici del cono isotropo: dal punto di vista affine si tratta infatti di rette per l'origine, $\underline{o}$, di V. [Proiettivamente, lette come punti, daranno vita alle coniche, quadriche e, in generale, ipersuperfiche]

Osserviamo che tale cono può ridursi a $\{\underline{o}\}$.



V

(*)

$C = q^{-1}(0)$

($q(v) = b(v,v)$ .. v. oltre)

ii) Data $e' = (e'_1 \dots e'_n)$ base b-ortogonale.

e $D = \begin{pmatrix} \beta_1 & & 0 \\ & \ddots & \\ 0 & & \beta_n \end{pmatrix}$

se $A = (a_{ij})$ è la matrice rappresentativa di b rispetto ad una base $e = (e_1 \dots e_n)$,

è $M^t A M = D$

con M matrice dell'applicazione

(*) $e_i \longmapsto e'_i$

rispetto a e: $M = m_{ee}(*)$

$\left(M = m_{ee'}(I) = m_{e'e}(I)^{-1} \right)$

$x = M x'$ vecchie coordinate in funzione delle nuove)

iii) Ribadiamo, e la cosa verrà ulteriormente chiarita in seguito, dopo i dovuti approfondimenti, che la congruenza è un concetto diverso dalla similitudine.

Osserviamo, a mo' di chiarimento, che, ad esempio, il polinomio caratteristico di una matrice A (prendiamola simmetrica, reale ma il risultato è generale), P_C^A

non è invariante per congruenza

$$\left(\det(M^t B M) = \det(M^t) \det B \det M = (\det M)^2 \det B \neq \det B \text{, in generale} \right)$$

Quello che si può però dire è ($A = A^t$ reale) che i segni delle radici (che vedremo, sono reali, nel caso simmetrico reale) non cambiano.

Ma approfondiremo a suo tempo la questione.

$$(*) \quad \underbrace{\det(M^t A M - \lambda I)}_{P_C^{M^t A M}} \neq \underbrace{\det(\cdot A \cdot - \lambda I)}_{P_C^A}$$

vedere gli esempi alla fine della dispensa

(a meno che $M^t M = I \ldots$)

Forme quadratiche ($K \doteq \mathbb{C}, \mathbb{R}, \mathbb{Q} \ldots$)

Sia (V, K), $\dim V = n \geq 1$; $b \in S^2(V^*)$.

Consideriamo la funzione

$$q : V \to K$$
$$v \mapsto q(v) := b(v, v)$$

Essa è detta forma quadratica associata a b.

(matricialmente: $q(x) = x^t A x = \sum_{i,j} a_{ij} x_i x_j$).

b può essere ricostruita da q nel modo seguente ("polarizzando" q)

$$(*) \quad b(v, w) = \frac{1}{2} \left[q(v+w) - q(v) - q(w) \right]$$

IDENTITÀ DI POLARIZZAZIONE ← origine geometrica del termine, che chiariremo

(verifica diretta $b(v+w, v+w) = \ldots = b(w,w) + b(v,v) + 2\, b(v,w) \ldots$)

Concretamente, data una forma quadratica in k^n

$$q(x) = \sum_{i,j} a'_{ij} x_i x_j \qquad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

(ovvero, un polinomio omogeneo di 2° grado nelle variabili $x_1 \dots x_n$ $\in k$)

Si può sempre fare in modo che

$$q(x) = \sum_{i,j} a_{ij} x_i x_j = x^t A x$$

con A simmetrica ($A = A^t$):

basta prendere $a_{ij} := \frac{1}{2}(a'_{ij} + a'_{ji})$

(basta fare questo per $i \neq j$)

Esempio: $a'_{11} x_1^2 + a'_{12} x_1 x_2 + a'_{21} x_2 x_1 +$

$+ a'_{22} x_2^2 =$

$= a'_{11} x_1^2 + (a'_{12} + a'_{21}) x_1 x_2 + a'_{22} x_2^2 =$

$= a'_{11} x_1^2 + 2 a_{ij} x_1 x_2 + a'_{22} x_2^2$

$= a_{11} x_1^2 + a_{12} x_1 x_2 + \underset{\substack{\parallel \\ a_{21}}}{a_{12}} x_2 x_1 + a_{22} x_2^2$

$$= (x_1 \ x_2) \begin{pmatrix} a_{11} & a_{12} \\ \underset{\substack{\parallel \\ a_{12}}}{a_{21}} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

5.79

Assumeremo ciò tacitamente, in seguito.

Dalla forma quadratica

$$q(v) = x^t A x \quad , \quad \underline{\underline{A = A^t}}$$

(se $v = \sum_{i=1}^{n} x_i e_i$, $e = (e_1 \dots e_n)$ base di V)

si risale alla forma bilineare simmetrica corrispondente (forma polare corrispondente alla forma quadratica data) ponendo semplicemente:

$$b(v, w) = x^t A y$$

($v = \sum_i x_i e_i$, $w = \sum y_i e_i$)

(questa operazione è detta polarizzazione)

Di conseguenza, possiamo lavorare indipendentemente con forme bilineari simmetriche o con le loro forme quadratiche associate

$$\begin{tikzcd} b \arrow[r, bend left, "v = w"] & q \arrow[l, bend left, "\text{polarizzazione}"] \end{tikzcd}$$

Possiamo allora parlare di diagonalizzazione di una forma bilineare o di una forma quadratica indifferentemente

5.80

Dunque, su K ($= \mathbb{C}, \mathbb{R}, \mathbb{Q} \dots$)
ogni forma quadratica può essere diagonalizzata: in una base b-ortogonale opportuna, la matrice rappresentativa diviene

$$D = \begin{pmatrix} \beta_1 & & 0 \\ & \ddots & \\ 0 & & \beta_n \end{pmatrix}$$

Si osservi che vi saranno, in ogni caso, $n-r$ valori $\beta_i = 0$ ($r = r(b) = r(A)$, A matrice rappresentativa qualsiasi)

(r è invariante per congruenza)

Nota: il cono isotropo relativo a b è $q^{-1}(0)$.

◇ ◇ ◇

Vogliamo ora esaminare in dettaglio i casi $K = \mathbb{C}$, $K = \mathbb{R}$; per arrivare a risultati più precisi, e trovare invarianti completi per congruenza.

Tralasceremo il caso $K = \mathbb{Q}$, poiché la teoria corrispondente è molto più difficile, ed esula dai nostri scopi (ciò porterebbe al teorema di Hasse-Minkowski)

$\boxed{K = \mathbb{C}}$ (lavoriamo con le forme quadratiche)

Mostriamo che la base diagonalizzante (b-ortogonale) si può scegliere in modo che

$$A = \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right) \qquad r = r(A) = r(b) \equiv r(q)$$

Infatti, sia $f = (f_1 \dots f_n)$ una base b-ortogonale e $v = \sum_{i=1}^{n} x_i f_i$.

Allora

$$q(v) = \sum_{i=1}^{n} \beta_i x_i^2 = \sum_{i=1}^{r} \beta_i x_i^2$$

(possiamo sempre fare in modo che $\beta_i = 0$ per $i = r+1 \dots n$)

⚠ I β_i non sono invarianti per congruenza (mutando base possono variare ed è proprio questa possibilità che sfrutteremo), ma in ogni caso ce ne sono sempre $n-r$ uguali a zero.

Sia ora α_i tale che $\beta_i = \alpha_i^2$, $i=1,\dots r$.

(è necessariamente $\alpha_i \neq 0$)

Otteniamo, successivamente:

$$q(v) = \sum_{i=1}^{r} \alpha_i^2 x_i^2 = \sum_{i=1}^{r} \underbrace{(\alpha_i x_i)}_{y_i}^2 = \sum_{i=1}^{r} y_i^2$$

Che corrisponde al cambiamento di base seguente:

$$v = \sum_{i=1}^{n} x_i f_i = \sum_{i=1}^{r} x_i f_i + \sum_{i=r+1}^{n} x_i f_i =$$

$$= \sum_{i=1}^{r} \overbrace{\alpha_i x_i}^{y_i} \underbrace{\left(\frac{f_i}{\alpha_i}\right)}_{e_i} + \sum_{i=r+1}^{n} x_i f_i \qquad (\alpha_i \neq 0,\ i=1..r)$$

ossia, si passa da f ad e, con

$$e = \left(\underset{\substack{\equiv \\ e_1}}{\frac{f_1}{\alpha_1}} \cdots \underset{\substack{\equiv \\ e_r}}{\frac{f_r}{\alpha_r}},\ \underset{\substack{\equiv \\ e_{r+1}}}{f_{r+1}}, \dots \underset{\substack{\equiv \\ e_n}}{f_n} \right)$$

$$y_i = \alpha_i x_i, \quad i=1,\dots r$$
$$y_i = x_i, \quad i=r+1,\dots,n$$

<u>Conclusione</u>: Se $K = \mathbb{C}$, il rango r è un <u>invariante completo</u> per congruenza: ogni $A = A^t \in M_n(\mathbb{C})$ è congruente a $D = \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right)$, con $r = r(A)$

$\boxed{K = \mathbb{R}}$ vi sono, rispetto a $\mathbb{C}$, due complicazioni.

1) Se $\beta_i < 0$, non si può risolvere in $\mathbb{R}$ l'equazione $\beta_i = \alpha_i^2$.

A ciò poniamo rimedio così: sia, per $\beta_i < 0$, scelto α_i in modo che:

$$\underbrace{-\beta_i}_{>0} = \alpha_i^2$$

[* di fatto, si legge, come sempre $\mathbb{R} \subseteq \mathbb{C}$]

Se ora p è il numero di $\beta_i > 0$ troviamo, in analogia al caso precedente:

$$q(v) = \sum_{i=1}^{p} y_i^2 - \sum_{i=p+1}^{r} y_i^2$$

e A diventa

$$\left(\begin{array}{c|c|c} I_p & & 0 \\ \hline & -I_{r-p} & \\ \hline 0 & & \boxed{0} \end{array}\right)$$

2) Sorge a questo punto l'ulteriore problema: p, e, conseguentemente, $r-p$, rimane invariato qualora si

passi ad una nuova base b-ortogonale con lo stesso procedimento ?

La risposta è sì ed è fornita dal

* <u>teorema di inerzia di Sylvester</u>.

mostriamo dunque tale invarianza :

sia

$$q(v) = \sum_{i=1}^{p} y_i^2 - \sum_{i=p+1}^{n} y_i^2 =$$

$$= \sum_{i=1}^{t} z_i^2 - \sum_{i=t+1}^{n} z_i^2$$

ove $v = \sum y_i e_i = \sum z_i b_i$

$e = (e_1, \dots e_n)$, $b = (b_1 \dots b_n)$ basi diagonalizzanti opportune.

Sia per assurdo $t < p$.

Troviamo successivamente, posto

$S = \langle e_1 \dots e_p \rangle$, $T = \langle b_{t+1} \dots b_n \rangle$

$\dim S = p$, $\dim T = n-t$.

e $\dim S + \dim T = p + n - t > n$

$(p > t)$

ma è pure

$\dim(S+T) = \dim(\langle e_1 \dots e_p, b_{t+1} \dots b_n \rangle) \leq n$

$(S+T \leq V)$

Pertanto (Grassmann)

$$n < \dim S + \dim T = \dim(S+T) + \dim(S \cap T)$$

$$\leq n + \dim(S \cap T)$$

cioè $n < n + \dim(S \cap T)$.

ossia $\dim(S \cap T) \geq 1$

Consideriamo allora $v \neq \underline{0}$, $v \in S \cap T$.

Si ha

$$q(v) = \sum_{i=1}^{p} y_i^2 > 0 \qquad v \in S, \ v \neq \underline{0}$$

e, nello stesso tempo

$$q(v) = -\sum_{i=t+1}^{n} z_i^2 \leq 0 \qquad v \in T, \ v \neq \underline{0}$$

$$= -\sum_{i=t+1}^{n} z_i^2$$

il che è assurdo.

Analogamente escludiamo che $p < t$ e pertanto $\boxed{p = t}$, che conclude la dimostrazione.

◇ ◇ ◇

Def. Poniamo

$$(\underbrace{p}_{\text{indice di positività}}, \underbrace{\overset{\text{rango}}{r} - p}_{\text{indice di negatività}}) : \text{ segnatura di } q$$

(v. anche poco oltre)

$(\nu = n - r$: indice di nullità$)$

Dunque, in $\mathbb{R}$, ogni forma quadratica è determinata, a meno di congruenze, dalla sua segnatura: in altre parole, la segnatura è un invariante completo per congruenza, nel caso reale

Riassumendo

	$\mathbb{C}$	$\mathbb{R}$
Invariante completo per congruenza di forme bilineari (quadratiche) e matrici simmetriche	r rango	$(p, r-p)$ segnatura
		...

Sia nel caso reale che in quello complesso, le forme canoniche

$$\sum_{i=1}^{r} x_i^2 \qquad \sum_{i=1}^{p} x_i^2 - \sum_{i=p+1}^{r} x_i^2$$

si dicono affini;

nel caso reale si parla anche di forma di Sylvester di una data forma quadratica (o forma bilineare simmetrica):

la parola "forma" compare dunque in due significati diversi, che comunque non si confondono. Sempre nel caso reale, una base corrispondente alla forma di Sylvester è anche detta base di Sylvester.

Perciò, ricapitolando, a livello matriciale, se:

$\boxed{K = \mathbb{C}}$ $\quad A, B \in M_n(\mathbb{C})$ sono $(A \approx B)$ congruenti
$A = A^t, \quad B = B^t$

$\iff \quad r(A) = r(B)$ (rango: in questo caso è un invariante completo)

(poiché, per il teorema di diagonalizzazione (per congruenza) ex 1 risultano entrambe congruenti a $D = \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right)$)

Se $D = M^t A M$, $D = N^t B N$ è $M^t A M = N^t B N$

$$A = (M^t)^{-1} N^t B N M^{-1} = \underbrace{(NM^{-1})^t}_{P^t} B \underbrace{NM^{-1}}_{P}$$

$P \in GL_n(\mathbb{C})$

$\boxed{K = \mathbb{R}}$ $\quad A, B \in M_n(\mathbb{R})$
$A = A^t, \quad B = B^t$

$A \approx B \iff \begin{pmatrix} p(A), & r(A) - p(A) \\ p(B), & r(B) - p(B) \end{pmatrix} =$ (stessa segnatura)

poiché, per il teorema di Sylvester, essi risultano entrambe congruenti a $D = \left(\begin{array}{c|c|c} I_p & 0 & 0 \\ \hline 0 & I_{r-p} & 0 \\ \hline 0 & 0 & 0 \end{array}\right)$

Di nuovo, se
$D = M^t A M = N^t B N$, è
$$A = \underbrace{(NM^{-1})^t}_{P^t} B \underbrace{(NM^{-1})}_{P}$$

p indice di positività
$r-p$ indice di negatività
$n-r$ indice di nullità

Schema riassuntivo

sulla diagonalizzabilità delle forme quadratiche, ovvero, matrici simmetriche (nel senso della congruenza)

q forma quadratica
b bilineare simmetrica

ex K (non identicamente nulla) $(= \mathbb{C}, \mathbb{R}, \mathbb{Q} \ldots)$

1. $\exists\, v \in V$ non isotropo
2. Per $v \in V$ non isotropo $V = \langle v \rangle \oplus \langle v \rangle^{\perp_b}$
3. induzione su n

strumenti

$\Rightarrow$ q (b) diagonalizzabile (per congruenza) $\quad D = \begin{pmatrix} \beta_1 & & 0 \\ & \ddots & \\ 0 & & \beta_n \end{pmatrix}$

$\delta_{ij} \cdot \beta_i = b(e_i, e_j)$

In particolare:

$\boxed{K = \mathbb{C}}$

si può giungere a $D = \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right)$

$\Rightarrow$ r (rango) invariante completo (★)

$K = \mathbb{Q}$ difficile: teorema di Hasse-Minkowski

Idea di base: in $\mathbb{C}$, se $\beta \neq 0$, $\exists\, \alpha \in \mathbb{C}$ t.c. $\alpha^2 = \beta$ (se $\beta = 0$ è $\alpha = 0$..)

$\boxed{K = \mathbb{R}}$

si giunge a $D = \left(\begin{array}{c|c|c} I_p & & 0 \\ \hline & -I_{r-p} & \\ \hline 0 & & 0 \end{array}\right)$ (*)

$p = \#i : \beta_i > 0$
$r - p = \#i : \beta_i < 0$

Idea: $\beta > 0 \quad \exists\, \alpha \in \mathbb{R} : \alpha^2 = \beta$
$\beta < 0 \quad \exists\, \alpha \in \mathbb{R} : \alpha^2 = -\beta$

★ fatto cruciale: l'inerzia, ovvero l'indipendenza di $(p, r-p)$ della base diagonalizzante (che fornisce D nella forma indicata (*)) (Sylvester)

segnatura $(p, r-p)$ invariante completo

Terminologia per $K = \mathbb{R}$

una forma quadratica reale q è detta

rango r $(\leq n = \dim V)$

- definita positiva se $q(v) > 0 \quad \forall v \neq \underline{0}$

  forma canonica (affine): $\sum_{i=1}^{n} x_i^2$ (basta esplicitare la suddetta condizione in una base di Sylvester).

  segnatura: $(n, 0)$

  La corrispondente forma bilineare è detta prodotto scalare (reale)

- semidefinita positiva, se $q(v) \geq 0 \quad \forall v \in V$

  forma canonica (affine) $\sum_{i=1}^{r} x_i^2$

  segnatura $(r, 0)$

- definita negativa, semidef. negativa (definizioni analoghe)

  segnatura: $(0, n)$ $\quad (0, r)$

  forme canoniche: $-\sum_{i=1}^{n} x_i^2$ $\quad -\sum_{i=1}^{r} x_i^2$

- indefinita negli altri casi

  forma canonica: $\sum_{i=1}^{p} x_i^2 - \sum_{i=p+1}^{r} x_i^2$

  segnatura: $(p, r-p)$

In generale, q (e b) è detta non degenere se $r = n$ (nelle applicazioni geometriche, analitiche e fisico-matematiche è questo il caso più interessante.

⚠ Spesso, con la parola segnatura si denota l'intero $s = \underbrace{p}_{\text{ind. positività}} - \underbrace{(r-p)}_{\text{ind. negatività}} = 2p - r$

Si osservi che allora $(p, r-p)$ si può ricostruire da r e s:

$$s = 2p - r \qquad p = \tfrac{1}{2}(s+r)$$

$$r - p = r - \tfrac{1}{2}(s+r) = \tfrac{1}{2}(r-s)$$

Dunque, $q_1 \approx q_2 \iff r_1 = r_2$ e $s_1 = s_2$

congruenti (uso lo stesso simbolo..)

In base alla terminologia introdotta, possiamo affermare che

$$p = \sup \left\{ \dim W_+ \;\middle|\; \begin{array}{l} W_+ \leq V, \\ q|_{W_+} \text{ def. positiva} \end{array} \right\}$$

indice di positività

$$n - p = \sup \left\{ \dim W_- \;\middle|\; \begin{array}{l} W_+ \leq V, \\ q|_{W_-} \text{ def. negativa} \end{array} \right\}$$

indice di negatività

$$\nu = n - r \qquad \sup \left\{ \dim W_0 \;\middle|\; \begin{array}{l} W_0 \leq V, \\ q|_{W_0} \text{ id. nulla} \end{array} \right\}$$

indice di nullità

(vedremo in seguito che, come conseguenza del teorema spettrale (Disp. 6), si avrà anche che

p = numero di autovalori di $A = A^t$ positivi contati con la loro molteplicità

$r - p$ = numero di autovalori negativi contati con la loro molteplicità

$\nu = n - r$ = numero di autovalori nulli contati con la loro molteplicità.)

- Criterio di Cartesio

Data una forma quadratica $q: V \to \mathbb{R}$ essa è definita positiva se e solo se, detto p_c^A il polinomio caratteristico associato ad una sua qualsiasi matrice rappresentativa, nessuno dei suoi coefficienti si annulla e i loro segni si alternano.

È definita negativa se e solo se essi hanno tutti lo stesso segno

Dim. Per semplicità trattiamo solo il caso $n = 3$, fornendo poi alcuni cenni per il caso generale (il caso $n = 2$ è semplicissimo).

$$p_c^A(\lambda) = \prod_{k=1}^{3} (\lambda_k - \lambda) =$$

$$(\lambda_1 - \lambda)(\lambda_2 - \lambda)(\lambda_3 - \lambda) =$$

$$(-\lambda)^3 \cdot 1 + (-\lambda)^2 \overbrace{[\lambda_1 + \lambda_2 + \lambda_3]}^{\beta_2} +$$

$$+ (-\lambda) \underbrace{(\lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_1\lambda_3)}_{\beta_1} + \underbrace{\lambda_1\lambda_2\lambda_3}_{\beta_0}$$

$$\equiv \sum_{j=0}^{3} (-\lambda)^j \beta_j = \sum_{j=0}^{3} \lambda^j \underbrace{(-1)^j \beta_j}_{\alpha_j} \equiv \sum_{j=0}^{3} \alpha_j \lambda^j$$

q è definita positiva $\Leftrightarrow$ $\lambda_j > 0$, $j = 1,2,3$.

Se ciò accade, è chiaro che $\beta_j > 0$ $\forall j$

$\Rightarrow$ gli α_j $(\neq 0)$ hanno segni alterni.

Viceversa, se ciò accade ($\alpha_j \neq 0$ $\forall j$ e con segni alterni) i β_j sono tutti > 0.

Allora gli autovalori sono $\neq 0$

altrimenti $\beta_0 = \lambda_1 \lambda_2 \lambda_3 = 0$

Se ora un autovalore (ex. λ_1) fosse negativo

si avrebbe $\beta_0 = \lambda_1 \lambda_2 \lambda_3 < 0$ e ciò è assurdo,

e così pure se tre autovalori fossero negativi

se due autovalori fossero negativi (ex: $\lambda_1 < 0$ $\lambda_2 < 0$)

si avrebbe, da $\beta_2 > 0$, $\lambda_3 > -(\lambda_1 + \lambda_2)$

e, da $\beta_1 > 0$,

$$\underbrace{\lambda_1 \lambda_2}_{>0} + (\lambda_1 + \lambda_2)\lambda_3 > 0$$

$$\Rightarrow \quad \lambda_1 \lambda_2 > -(\lambda_1 + \lambda_2)\lambda_3 > (\lambda_1 + \lambda_2)^2$$

$$\Rightarrow \quad \lambda_1 \lambda_2 > \lambda_1^2 + \lambda_2^2 + 2\lambda_1\lambda_2$$

$$\Rightarrow \quad 0 > \lambda_1^2 + \lambda_2^2 + \lambda_1\lambda_2 > 0 \quad \text{assurdo.}$$

Dunque $\lambda_j > 0$ $\forall j$, sicché q è definita positiva

Il ragionamento è analogo nell'altro caso...

In generale è

$$P_C^A(\lambda) = \prod_{k=1}^{n} (\lambda_k - \lambda) =$$

$$\equiv \sum_{j=0}^{n} (-\lambda)^{n-j} S_j^{(n)}$$

funzione simmetrica elementare in $\lambda_1 \dots \lambda_n$

sono polinomi omogenei di grado j in $\lambda_1 \dots \lambda_n$

Esempi:

$n = 2$

discutere il criterio di Sylvester .. vedremo che alla fine del Cap. 6

$$\begin{cases} S_2^{(2)} = \lambda_1 \lambda_2 = \det A \\ S_1^{(2)} = \lambda_1 + \lambda_2 = \operatorname{tr} A \\ S_0^{(2)} = 1 \end{cases}$$

$n = 3$

(il caso esaminato)

$$\begin{cases} S_3^{(3)} = \lambda_1 \lambda_2 \lambda_3 = \det A \\ S_2^{(3)} = \lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3 \\ S_1^{(3)} = \lambda_1 + \lambda_2 + \lambda_3 = \operatorname{tr} A \\ S_0^{(3)} = 1 \end{cases}$$

ed è facile generalizzare le idee della dimostrazione data per $n = 3$ ....

Criterio di Sylvester

Sia $A = A^t \in M_n(\mathbb{R})$, detta q_A la forma quadratica corrispondente $q_A(x) = x^t A x$, $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, si chiama che q_A è definita positiva se $A = M^t M$, $M \in GL_n(\mathbb{R})$

(ciò è immediato dal teorema di Sylvester)

Possiamo accertare ciò anche tramite il seguente criterio di Sylvester, che enunciamo soltanto

◇ Criterio di Sylvester: i) Se i minori principali(*) di A sono tutti positivi, allora q_A è definita positiva

(*) ottenuti orlando via via a partire da a_{11} →

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

$d_1 = a_{11} > 0$

$d_2 = a_{11}a_{22} - a_{12}a_{21} > 0$

$\vdots$

$d_n = |A| > 0$

Se invece $(-1)^k d_k > 0$ $k = 1 \dots n$

allora q_A è def. negativa

◆ Un'osservazione su $\sim$ vs $\approx$
similitudine congruenza

Sia $I_n = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$

vista come operatore (endomorfismo) in $\mathbb{R}^n$:

$[I_n]_{\sim} = \{I_n\}$ ← classe di similitudine

vista come forma bilineare (quadratica)

$[I_n]_{\approx} = \{M^t M \ / \ M \in GL_n(\mathbb{R})\}$ ← classe di congruenza

$\equiv$ prodotti scalari (segnatura $(n, 0)$)

◇ Nel caso $n=2$, $r=2$

è facile concludere che (*), se $A = \begin{pmatrix} a & b \\ b & d \end{pmatrix}$

(*) (il criterio di Sylvester, di Cartesio, o con un argomento elementare diretto (v. anche fine cap. 6))

- q_A def. positiva (tipo $(2,0)$)

$\iff$ $a > 0$ (o, eq. $d > 0$)

e $ad - b^2 = |A| > 0$

- q_A def. negativa (tipo $(0,2)$) $\iff$

$a < 0$ e $|A| > 0$

(o, eq. $d < 0$)

- q_A indefinita (tipo $(1,1)$)

$\iff$ $|A| < 0$

Ciò è importante nell'ambito della classificazione dei punti critici delle funzioni di due variabili

Inciso:

Ma il tutto è generale: (nel caso non degenere)

la matrice Hessiana calcolata in un punto critico (*) (è simmetrica)

fornisce indicazioni sulla natura dello stesso (max o min locale, o pto di sella etc.) attraverso la sua segnatura

$f = f(\underbrace{x_1 \dots x_n}_{x})$ $\quad \mathbb{R}^2$

Se $\frac{\partial f}{\partial x_i}(x_0) = 0$, si studia $\left(\frac{\partial^2 f}{\partial x_i \partial x_j}\Big|_{x_0} \right)$ matrice Hessiana

pto critico

è una matrice simmetrica (Lemma di Schwarz) ..

◇ Esempi (con qualche anticipazione)

$n = 2$

$v = \begin{pmatrix} x \\ y \end{pmatrix}$ $\quad x_1 = x \quad x_2 = y$ $\quad z$: variabile ausiliaria

$\boxed{(2,0)}$ $\quad z = q(v) = x^2 + y^2$ $\quad$ paraboloide ellittico

Consideriamo

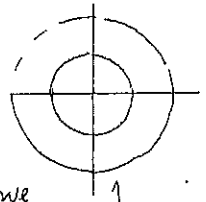
$$\begin{cases} z = x^2 + y^2 \\ z = c \end{cases}$$ ← piani // al piano (x,y)

$x^2 + y^2 = c$ ← nel piano: circonferenze (☆) di centro l'origine e raggio $\sqrt{c}$, se $c \geq 0$

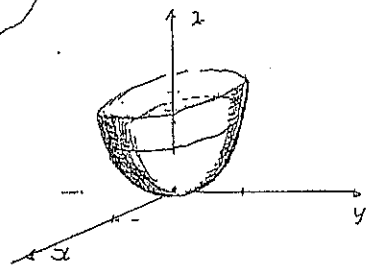
altrimenti senza pti reali

nello spazio: cilindro (ellittico)

$z = c$



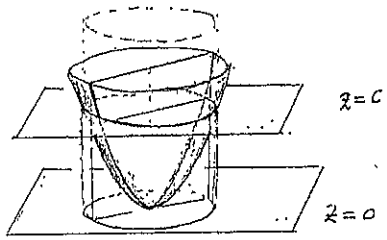
☆ curve di livello



sezioni con i piani coordinati: parabole

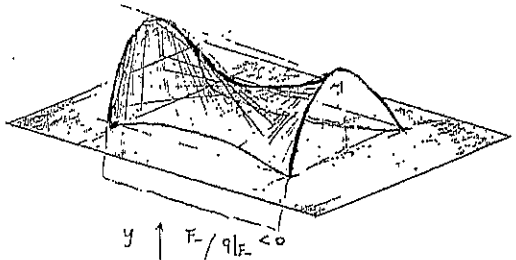
(e con un generico piano del fascio di asse z)

(☆) Utilizziamo liberamente la geometria elementare, impiegando anche nozioni metriche, a priori non necessarie



$\boxed{(0,2)}$ ... analogo ..

$\boxed{(1,1)}$ $\quad z = x^2 - y^2$ $\quad$ paraboloide iperbolico (sella)



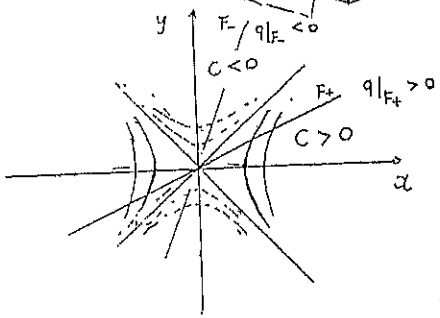
$$\begin{cases} z = x^2 - y^2 \\ z = c \end{cases}$$

: nel piano

$x^2 - y^2 = c$

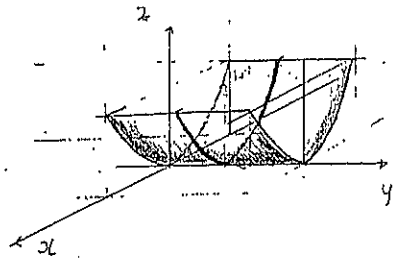
curve di livello : iperboli

sezioni col fascio di piani per l'asse z : parabole



$\boxed{(1,0)}$ $\quad z = x^2$ $\quad$ ([(0,1) .. analogo])

cilindro parabolico



Esempio In $\mathbb{R}^3$ sia

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \qquad r(A) = 3$$

$$v = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \qquad w = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}$$

$$b_A(v, w) = v^t A w = (x \ y \ z) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} =$$

$$= xx' + yy' - zz'$$

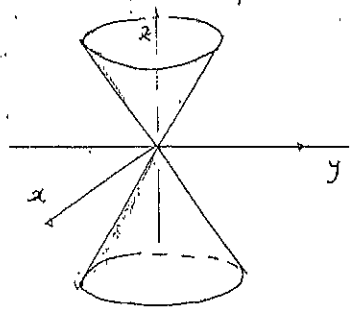
$$q_A(v) = x^2 + y^2 - z^2 \quad \text{(già in forma canonica)}$$

Segnatura $(2, 1)$

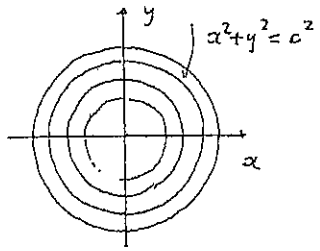
Cono isotropo (è effettivamente un cono!)

$$q_A^{-1}(0) = \{ v \in V \ / \ q(v) = 0 \}$$

$$x^2 + y^2 - z^2 = 0$$



pianta



$$\begin{cases} x^2 + y^2 = z^2 \\ z = c \end{cases} \Rightarrow \begin{cases} x^2 + y^2 = c^2 \\ z = c \end{cases}$$

curve di livello: circonferenze con centro in O e raggio $= |c|$

Consideriamo $v = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$

Si tratta di un vettore isotropo:

$$q_A(v) = 1 - 1 = 0 \qquad \langle v \rangle = \begin{cases} x = \alpha \\ y = 0 \\ z = \alpha \end{cases}$$

Determiniamo $\langle v \rangle^{\perp_A}$

Si ha:

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

ovvero (anche in modo diretto

$$x \cdot 1 + y \cdot 0 - z \cdot 1 =$$

$$= x - z = 0 \qquad \begin{cases} x = \alpha \\ y = y \\ z = \alpha \end{cases}$$

$$\Rightarrow \langle v \rangle^{\perp_A} = \langle \underbrace{\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}}_{\text{base}} \rangle$$

Si noti che, come è giusto che sia $\langle v \rangle \leq \langle v \rangle^{\perp_A}$

* Geometricamente $\langle v \rangle^{\perp_A}$ è il piano tangente al cono isotropo lungo la generatrice $\langle v \rangle$ [e non è finita qui ... rimandiamo alla teoria proiettiva delle coniche]

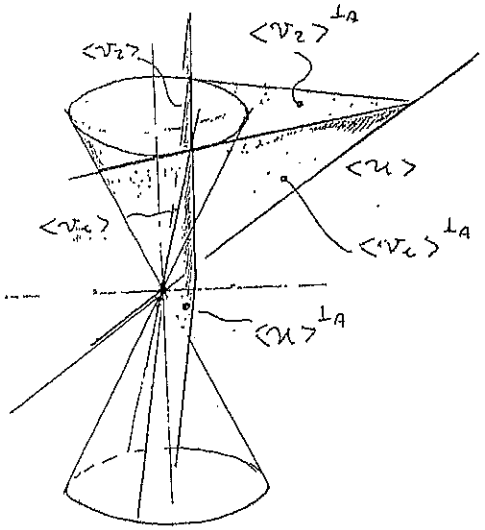
v. figura

Si verifichino, a titolo di esercizio, i seguenti fatti

1. Si determini la coppia di rette intersezione del cono con il piano $\langle u \rangle^{\perp_A}$ (esse costituiscono due generatrici del cono) $\langle v_1 \rangle$, $\langle v_2 \rangle$ per certi vettori (non univocamente determinati) v_1, v_2, entrambi $\neq \underline{0}$

2. Si determinino $\langle v_1 \rangle^{\perp_A}$ e $\langle v_2 \rangle^{\perp_A}$

3. Si determini infine $\langle v_1 \rangle^{\perp_A} \cap \langle v_2 \rangle^{\perp_A}$ e si verifichi che essa è la retta $\langle u \rangle$ (cf. la figura)



Una comprensione più profonda si avrà tramite la geometria proiettiva

Eliminazione di Gauss simmetrica (Diagonalizzazione per congruenza di una matrice simmetrica) $\equiv$ metodo del completamento dei quadrati

Sia $A \in M_n(K)$ $\qquad K = \mathbb{Q}, \mathbb{R}, \mathbb{C}..$

Formiamo $\boxed{(A \mid I)}$

si arriva ad una matrice diagonale; poi si può andare oltre utilizzando metodi ad hoc per $\mathbb{C}$ e $\mathbb{R}$.

I. Se ora $\boxed{a_{11} \neq 0}$, applichiamo opportune operazioni di riga e colonna in modo da ottenere

$$\begin{array}{c} \\ n \end{array} \left\{ \begin{array}{c} \overbrace{\qquad\qquad}^{n} \\ \left(\begin{array}{c|ccc} a_{11} & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \boxed{B} & \\ 0 & & & \end{array}\right. \end{array} \right. \begin{array}{c} \overbrace{\qquad\qquad}^{n} \\ \left.\begin{array}{ccc} * & * & * \\ \vdots & & \\ * & * & * \end{array}\right) \end{array}$$

Si noti che le operazioni di riga coinvolgono tutta $(A \mid I)$, ma quelle di colonna non toccano la parte destra

II. Se $a_{11} = 0$, ma c'è un elemento diagonale $a_{ii} \neq 0$, si scambino le righe R_1 e R_i e le colonne C_1 e C_i in modo da portarlo in prima posizione: ci si riconduce al caso I.

III. Se tutti gli elementi diagonali sono

nulli, si sceglie un $a_{ij} \neq 0$

(se non ci è nessun elemento siffatto, la matrice è nulla, e qualunque base la diagonalizza)

Aggiungiamo ora la riga j-esima alla i-esima e così la colonna j-esima alla colonna i-esima:

$$R_j + R_i \mapsto R_i \qquad C_j + C_i \mapsto C_i$$

si pone $R_j + R_i$ al posto di R_i = =

Si ottiene allora $2a_{ij}$ al posto i-esimo della diagonale. Si può allora applicare $\frac{1}{2}$ e quindi I, o direttamente I se $i=1$

In definitiva, ci si riconduce sempre a I, e a questo punto si itera il procedimento, lavorando su B.

Si arriva allora a (D, Q)
↑ diagonale

Dimostriamo che, posto $P = Q^t$, è

$$\boxed{D = P^t A P}$$

Sia E_i : operazione elementare di riga;
" F_i = " di colonna:
è subito $F_i = E_i^t$

le operazioni di riga moltiplicano a sinistra, quelle di colonna moltiplicano a destra

Si ha

$$Q = E_k \dots E_2 E_1 I = E_k \dots E_1$$

per opportune operazioni di riga

(le operazioni di colonna non toccano la parte destra).

Si ha invece:

$$\begin{aligned} D &= E_k \dots E_1 A F_1 F_2 \dots F_k \\ &= E_k \dots E_1 A E_1^t E_2^t \dots E_k^t \\ &= \underbrace{(E_k \dots E_1)}_{Q} A \underbrace{(E_k \dots E_1)^t}_{Q^t} \end{aligned}$$

Posto $Q^t = P$

$$= P^t A P$$

Tale metodo è equivalente al metodo del completamento dei quadrati (è in questa forma che venne usato da Gauss; come metodo è antichissimo...)

Inciso

Il metodo di soluzione di un'eq. di 2° grado è esattamente di questo tipo:

$$ax^2 + bx + c = 0 \qquad a \neq 0$$

poniamo $a > 0$, non si perde in generalità

$$x^2 + \frac{b}{a}x + \frac{c}{a} = 0$$

$$x^2 + 2\frac{b}{2a} \cdot a\,x + \frac{b^2}{4a^2} - \frac{b^2}{4a^2} + \frac{c}{a} = 0$$

$$\left(x + \frac{b}{2a}\right)^2 + \frac{4ac - b^2}{4a^2} = 0$$

$$\left(x + \frac{b}{2a}\right)^2 = \frac{\Delta}{4a^2}$$

$$x = -\frac{b}{2a} \pm \frac{\sqrt{\Delta}}{2a} = \frac{-b \pm \sqrt{\Delta}}{2a}$$

◇ ◇ ◇

Una notazione più corretta sarebbe (per avere la massima generalità) operando nel campo complesso

$$x = \frac{-b + \sqrt{\Delta}}{2a}$$

$\sqrt{\ }$ già contiene le due radici

• Esempio

$$A = \left(\begin{array}{cc|c} 1 & 1 & 0 \\ 1 & 2 & 0 \\ \hline 0 & 0 & 3 \end{array}\right)$$

Basta lavorare con $A_L = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$

$$\left(\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{array}\right) \xrightarrow{\text{righe}} \left(\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 0 & 1 & -1 & 1 \end{array}\right)$$

$$\xrightarrow{\text{colonne}} \left(\begin{array}{cc|cc} 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 1 \end{array}\right)$$

In definitiva

$$\underbrace{\left(\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \hline 0 & 0 & 3 \end{array}\right)}_{D} \quad \overbrace{\left(\begin{array}{cc|c} 1 & 0 & 0 \\ -1 & 1 & 0 \\ \hline 0 & 0 & 1 \end{array}\right)}^{Q}$$

Segnatura: $(3, 0)$

Verifica:

$$Q A Q^t = \left(\begin{array}{cc|c} 1 & 0 & 0 \\ -1 & 1 & 0 \\ \hline 0 & 0 & 1 \end{array}\right)\left(\begin{array}{cc|c} 1 & 1 & 0 \\ 1 & 2 & 0 \\ \hline 0 & 0 & 3 \end{array}\right)\left(\begin{array}{cc|c} 1 & -1 & 0 \\ 0 & 1 & 0 \\ \hline 0 & 0 & 1 \end{array}\right)$$

basta calcolare $\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} =$

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

$$\Rightarrow \quad Q A Q^t = \left(\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \hline 0 & 0 & 3 \end{array}\right) = D$$

Esaminiamo lo stesso esempio in altro modo (completamento dei quadrati)

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \qquad X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

$$X^t A X =$$

$$= x^2 + 2xy + 2y^2 + 3z^2 =$$

$$= x^2 + 2xy + y^2 + y^2 + 3z^2 =$$

$$= (x+y)^2 + y^2 + 3z^2$$

Poniamo

$$x' = x + y \qquad x = x' - y'$$
$$y' = y \qquad y = y'$$
$$z' = z \qquad z = z'$$

$$X^t A X =$$

$$= x'^2 + y'^2 + 3z'^2 = X'^t D X'$$

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix} \qquad X' = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}$$

$$P = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$X = P X'$$

proseguendo .. $= x''^2 + y''^2 + z''^2$

$$x' = x'' \qquad y' = y'' \qquad z' = \frac{z''}{\sqrt{3}} \quad \Rightarrow \quad \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ \frac{1}{\sqrt{3}} \end{pmatrix} \right)$$

base di Sylvester

$$x = x'' - y'' \qquad y = y'' \qquad z = \frac{z''}{\sqrt{3}}$$

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{\sqrt{3}} \end{pmatrix} \cdot \begin{pmatrix} x'' \\ y'' \\ z'' \end{pmatrix}$$

Osserviamo, in questo semplice esempio, che

$$P_C^D(\lambda) = (1-\lambda)^2 (3-\lambda)$$

$$P_C^A(\lambda) = (3-\lambda) \cdot \begin{vmatrix} 1-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} =$$

$$= (3-\lambda) \{ (1-\lambda)(2-\lambda) - 1 \} =$$

$$= (3-\lambda) \{ \lambda^2 - 3\lambda + 2 - 1 \} =$$

$$= (3-\lambda) \{ \lambda^2 - 3\lambda + 1 \}$$

$$\lambda^2 - 3\lambda + 1 = 0$$

$$\lambda = \frac{3 \pm \sqrt{9-4}}{2} = \frac{3 \pm \sqrt{5}}{2} = \begin{cases} \lambda_+ \\ \lambda_- \end{cases}$$

entrambe positive

$$= (3-\lambda)(\lambda - \lambda_+)(\lambda - \lambda_-)$$

Le radici (a parte 3) sono diverse $P_C^D \neq P_C^A$ $(\Rightarrow A \not\sim D)$

ma $A \approx D$ $(q_A \approx q_D)$

Segnatura di entrambe : $(3, 0)$

(corrispondono pertanto a prodotti scalari)

Notare che i segni delle radici sono rimasti invariati

Se però poniamo (per es.)

$$D_1 = \begin{pmatrix} \lambda_+ & 0 & 0 \\ 0 & \lambda_- & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

è certamente $A \sim D_1$ (◇)

e, nello stesso tempo $A \approx D_1$

(e, ovviamente, è $D \approx D_1$, $D \not\sim D_1$)

(◇) A è diagonalizzabile (nel senso della similitudine) poiché è simmetrica (v. oltre); prescindendo da questo, basta osservare che i suoi autovalori sono distinti.

È chiaro che, per def., $P_C^{D_1} = P_C^{A}$

e dunque $A \sim D_1$.

- Consideriamo in $\mathbb{R}^3$ $\quad v = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$

$$q(v) = x^2 + 2xy + 4yz + y^2 + z^2$$

Si ha

$$q(v) = \underbrace{x^2 + 2xy + y^2}_{(x+y)^2} + z^2 + 4yz =$$

$$= (x+y)^2 + z^2 + 2(2y)\cdot z + 4y^2 - 4y^2$$

$$= (x+y)^2 + (z+2y)^2 - 4y^2$$

poniamo per esempio
$$\begin{cases} x' = x+y \\ y' = 2y \\ z' = z+2y \end{cases}$$
ovvero

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 2 & 1 \end{pmatrix}}_{\text{non singolare}} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \Rightarrow q(v) = x'^2 - y'^2 + z'^2$$

$\Rightarrow$ segnatura $(2, 1)$

Risolvere lo stesso esercizio in termini puramente matriciali

5.114

Consideriamo

$$q(X) = xy \qquad X = \begin{pmatrix} x \\ y \end{pmatrix}$$

$$A = \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix} \qquad (\text{notare : } a_{11} = a_{22} = 0)$$

Basta porre $x = x' + y'$, $y = x' - y'$ (per es.)

$$X' = \begin{pmatrix} x' \\ y' \end{pmatrix}$$

$$q(X') = (x'+y')(x'-y')$$
$$= x'^2 - y'^2$$

$$X = MX' \qquad M = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

$\Rightarrow$ segnatura $= (1, 1)$

Anche qui, lavorare matricialmente.

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

$$q(v) = x^2 + 2xy + 2xz + 3yz$$

$\left(v = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \right)$. Si ha, successivamente

$$q(v) = x^2 + 2xy + 2xz + 2yz + yz$$
$$= (x+y+z)^2 - y^2 - z^2 + yz$$
$$= (x+y+z)^2 - y^2 - z^2 + 2y\frac{z}{2} =$$
$$= (x+y+z)^2 - \left(y - \frac{z}{2}\right)^2 + \frac{z^2}{4} - z^2 =$$
$$= (x+y+z)^2 - \left(y - \frac{z}{2}\right)^2 - \frac{3}{4}z^2$$
$$= \underbrace{(x+y+z)}_{x'}{}^2 - \underbrace{\left(y - \frac{z}{2}\right)}_{y'}{}^2 - \underbrace{\left(\frac{\sqrt{3}}{2}z\right)}_{z'}{}^2$$
$$= x'^2 - y'^2 - z'^2$$

($\Rightarrow$ segnatura : $(1, 2)$)

Procedere anche matricialmente ....

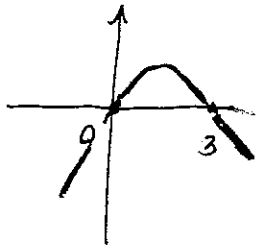
Nel capitolo 9 esamineremo altri metodi di diagonalizzazione per congruenza.

◆ Studiare la forma quadratica $q_{A_\alpha}: \mathbb{R}^2 \to \mathbb{R}$ rappresentata, al variare di $\alpha \in \mathbb{R}$, dalla seguente matrice

$$A_\alpha = \begin{pmatrix} 1 & \alpha \\ \alpha & 3\alpha \end{pmatrix}$$

Si ha $\det A_\alpha = 3\alpha - \alpha^2 = (3-\alpha)\alpha$

$\operatorname{tr} A_\alpha = 1 + 3\alpha$



★ $\det A_\alpha > 0$ per $0 < \alpha < 3$

Da $1 > 0$ si trova q_{A_α} def. positiva segnatura $(2,0)$

(è un prod. scalare)

★ $\det A_\alpha < 0$ per $\alpha < 0$ e $\alpha > 3$

$\Rightarrow$ q_{A_α} è in tal caso indefinita $(1,1)$ (segnatura)

★ $\det A_\alpha = 0$ per $\alpha = 0$ e $\alpha = 3$

$\boxed{\alpha = 0}$ $A_0 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ $\Rightarrow$ semid. positiva $(1,0)$

$\boxed{\alpha = 3}$ si ha $\det A_3 = 0$ e $\operatorname{tr} A_3 = 1 + 3\cdot 3 = 10 > 0$

$\Rightarrow$ è semidefinita positiva.

◆ Variante: studiare $A_\alpha = \begin{pmatrix} -1 & \alpha \\ \alpha & 3\alpha \end{pmatrix}$

PARTE II

Quadro sinottico

$K=\mathbb{R}$	$K=\mathbb{C}$
$(V, \langle\,|\,\rangle)$	$(V, \langle\,|\,\rangle)$
• spazio euclideo	• spazio hermitiano
$\langle\,|\,\rangle$ forma bilineare simmetrica definita positiva	$\langle\,|\,\rangle$ forma sesquilineare simmetrica (in senso hermitiano) def. positiva
• $T \in End(V)$	• $T \in End(V)$
• T^t operatore trasposto (o aggiunto)	• T^* aggiunto (coniugato hermitiano)
$\langle v \,|\, T^t w\rangle := \langle Tv \,|\, w\rangle$	$\langle v \,|\, T^* w\rangle := \langle Tv \,|\, w\rangle$
• in una base ortonormale $e = (e_1 \dots e_n)$	
$m_{ee}(T^t) = m_{ee}(T)^t$	$m_{ee}(T^*) = \overline{m_{ee}(T)}^t = m_{ee}(T)^t$
• operatore simmetrico: $T = T^t$	• operatore hermitiano $T = T^*$
$\langle v \,|\, Tw\rangle = \langle Tv \,|\, w\rangle$	$\langle v \,|\, Tw\rangle = \langle Tv \,|\, w\rangle$
$m_{ee}(T) = m_{ee}(T)^t$, $A = A^t$ matrice simmetrica	$m_{ee}(T) = \overline{m_{ee}(T)}^t$, $A = A^*$ matrice hermitiana o autoaggiunta
• gruppo ortogonale $O(V) = \{ O : \langle Ov \,|\, Ow\rangle = \langle v \,|\, w\rangle \}$	• gruppo unitario $U(V) = \{ U : \langle Uv \,|\, Uw\rangle = \langle v \,|\, w\rangle \}$
(◇) $O^{-1} = O^t$ come operatore	(◇) $U^{-1} = U^*$

• matricialmente (base ortonormale) stesse condizioni

Interpretazione delle (◇): le righe (e le colonne) di O e U costituiscono basi ortonormali in $\mathbb{R}^n$, $\mathbb{C}^n$ rispetto ai rispettivi prodotti scalari standard

Fondamentale è, a questo punto, il

◈ Teorema spettrale

Sia $(V, \langle\,|\,\rangle)$ uno spazio hermitiano

Sia $T \in End(V)$, $T = T^*$ operatore hermitiano.

Valgono allora i seguenti fatti:

i) $\emptyset \neq \sigma(T) \subseteq \mathbb{R}$

ii) Siano $\lambda, \mu \in \sigma(T)$, $\lambda \neq \mu$

Allora $V_\lambda \perp V_\mu$ (ortogonale (o perpendicolare))

questo è il contenuto del teorema spettrale propriamente detto

iii) T è unitariamente diagonalizzabile ovvero esiste una base ortonormale $e = (e_1 \dots e_n)$ costituita da autovettori di T:

$$T e_i = \lambda_i e_i$$
$$\langle e_i \,|\, e_j\rangle = \delta_{ij}$$

equivalentemente: $V = \bigoplus_{\lambda_k \text{ distinti}} V_{\lambda_k}$ somma diretta ortogonale

Dunque, data una qualsiasi matrice $m_{ff}(T)$ con $f = (f_1 \dots f_n)$ ortonormale (pertanto $m_{ff}(T) = m_{ff}(T)^*$

esiste $U \in U(n)$ tale che

$$U^{-1} m_{ff}(T) \cdot U = m_{ee}(T) = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

si prende $U = M_{fe} = m_{fe}(I)$

entrambe ortonormali

$$\underbrace{m_{ef}(I)}_{U^{-1}} \, m_{ff}(T) \, \underbrace{m_{fe}(I)}_{U} = m_{ee}(T)$$

$$m_{ef}(I) = m_{fe}(I)^{-1}$$

Teorema spettrale II

(versione matriciale)

una matrice hermitiana $A = A^* \in M_n(\mathbb{C})$ ha tutti gli autovalori reali, e risulta unitariamente diagonalizzabile, ovvero esiste $U \in U(n)$ tale che

$$\underbrace{U^{-1}}_{U^*} A \, U = D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

autovalori di A

(di nuovo, autospazi corrispondenti ad autovalori distinti sono ortogonali)

☆ Questa versione del teorema segue subito dall'altra interpretando $A = A^*$ come endomorfismo $A: \mathbb{C}^n \to \mathbb{C}^n$ hermitiano, con riferimento al prodotto scalare standard, che rende la base canonica ortonormale.

Prima di dimostrare il teorema spettrale, notiamo che esso rimane vero per un operatore simmetrico $T = T^t \in End(V)$, con $(V, \langle | \rangle)$ spazio euclideo:

In concreto una matrice simmetrica $A = A^t \in M_n(\mathbb{R})$ può sempre essere interpretata come una matrice hermitiana $A = A^* \in M_n(\mathbb{C})$

(ciò semplicemente perché, essendo reale, l'operazione di coniugazione hermitiana si riduce alla trasposizione. Inoltre letto il polinomio caratteristico come polinomio complesso a coefficienti reali, esso ha radici reali (vedi oltre))

Teorema spettrale: caso euclideo

Dunque: $T = T^t \in End(V)$, $(V, \langle | \rangle)$ spazio euclideo ha tutti gli autovalori reali e risulta ortogonalmente diagonalizzabile, ovvero esiste una base ortonormale di autovettori

$$T e_i = \lambda_i e_i \qquad \langle e_i | e_j \rangle = \delta_{ij}$$

A livello matriciale, previa interpretazione di $A = A^t \in M_n(\mathbb{R})$ come operatore simmetrico nello spazio euclideo $(\mathbb{R}^n, \langle | \rangle_{standard})$ (in cui la base canonica è ortonormale), si ha che una matrice simmetrica ha tutti gli autovalori reali e risulta essere ortogonalmente diagonalizzabile

ovvero $\exists \; O \in O(n)$ tale che

$$\underset{\overset{\shortparallel}{O^t}}{O^{-1}} A O = D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

autovalori di A

$M_n(\mathbb{R})$

Osserviamo ancora che il teorema spettrale vale per gli operatori unitari ($U U^* = U^* U = I$)

(in questo caso $\sigma(U) \subset \{ z \in \mathbb{C} \mid |z| = 1 \}$ — vedi oltre — la circonferenza di centro $O = (0,0)$ e raggio 1)

e, più in generale, per gli operatori normali, ovvero gli $N \in End(V)$ ($(V, \langle | \rangle)$ spazio hermitiano) tali che $N^* N = N N^*$

(ovvero, quelli che commutano col loro aggiunto)

Si noti che tale classe comprende gli operatori hermitiani e gli unitari.

In generale lo spettro $\sigma(N)$ è un sottoinsieme (finito) del piano complesso.

Teorema spettrale

quadro sinottico

- versione operatoriale

$(V, \langle \,	\, \rangle)$ spazio euclideo	$(V, \langle \,	\, \rangle)$ spazio hermitiano
$T = T^t$ (T simmetrico)	$T = T^*$ (T hermitiano)
ortogonalmente diagonalizzabile	unitariamente diagonalizzabile

$\exists$ base ortonormale di autovettori

- versione matriciale

$A = A^t \in M_n(\mathbb{R})$	$A = A^* \in M_n(\mathbb{C})$
ortogonalmente diagonalizzabile	unitariamente diagonalizzabile
(ortogonalmente simile ad una matrice diagonale)	(unitariamente simile ad una matrice diagonale)
$D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$	$D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$
con λ_i autovalori di A ($\in \mathbb{R}$)	λ_i autovalori di A ($\in \mathbb{R}$)
$O^{-1} A O = O^t A O = D$	$U^{-1} A U = U^* A U = D$
$O^{-1} = O^t$	$U^{-1} = U^*$

- versione operatoriale (e matriciale) equivalente:

$$V = \bigoplus_{\lambda_k \text{ distinti}} V_{\lambda_k}$$

Somma diretta e ortogonale

Conseguenza del teorema spettrale:

per gli operatori hermitiani (più in generale, normali), P_G è un invariante completo per similitudine: ciò semplicemente perché tali operatori sono, in particolare, diagonalizzabili.

Dimostrazione del teorema spettrale

i) $\sigma(T)$ è non vuoto per il teorema fondamentale dell'algebra.

Inoltre, da $\lambda \in \sigma(T)$

$$Tv = \lambda v \qquad v \neq \underline{0} \text{ autovettore corrispondente a } \lambda$$

Si ha subito

$$\langle v | Tv \rangle = \langle v | \lambda v \rangle = \lambda \underbrace{\langle v | v \rangle}_{>0}$$

$$(T=T^*) \quad \| $$

$$\langle \underbrace{Tv}_{\lambda v} | v \rangle = \bar{\lambda} \underbrace{\langle v | v \rangle}_{>0} \quad \Rightarrow \quad \lambda = \bar{\lambda}$$

ovvero $\lambda \in \mathbb{R}$

ii) Siano $\underset{Tv=\lambda v}{v \in V_\lambda}$, $\underset{Tw=\mu w}{w \in V_\mu}$

Allora

$$\langle v | Tw \rangle = \langle v | \mu w \rangle = \mu \langle v | w \rangle$$

$$(T=T^*) \quad \|$$

$$\langle Tv | w \rangle = \langle \lambda v | w \rangle = \bar{\lambda} \langle v | w \rangle = \lambda \langle v | w \rangle \quad (\lambda \in \mathbb{R})$$

$$\Rightarrow (\mu - \lambda) \langle v | w \rangle = 0 \quad \Rightarrow \quad (\text{essendo } \mu \neq \lambda)$$

$$\langle v | w \rangle = 0$$

Procediamo per induzione su $n = \dim V$

(si utilizza una forma del principio di induzione equivalente a quella "tradizionale")

Per $n=1$ il teorema è banalmente vero.

Supponiamo vero il teorema per $\dim V \leq n-1$

Sia ora V_λ un autospazio, $\lambda \in \sigma(T)$

Si ha $V = V_\lambda \oplus V_\lambda^\perp$
$(\dim V_\lambda^\perp \leq n-1)$

osserviamo che V_λ e $V_\lambda^\perp$ sono entrambi T-invarianti (eq. T-stabili) ovvero

$$T V_\lambda \subseteq V_\lambda \quad \text{(ciò è chiaro)}$$

e

$$T V_\lambda^\perp \subseteq V_\lambda^\perp$$

Sia infatti $w \in V_\lambda^\perp$; facciamo vedere che $Tw \in V_\lambda^\perp$

$\forall v \in V_\lambda$ si ha

$$\langle v | Tw \rangle \underset{(T=T^*)}{=} \langle Tv | w \rangle = \bar{\lambda} \langle v | w \rangle = \lambda \langle v | w \rangle = 0 \quad \text{(per ipotesi)}$$

ovvero $Tw \perp V_\lambda$
$(Tw \in V_\lambda^\perp)$

ora $T|_{V_\lambda^\perp} : V_\lambda^\perp \to V_\lambda^\perp$

è unitariamente diagonalizzabile poiché $\dim V_\lambda^\perp \leq n-1$ (ipotesi induttiva)

inoltre $T|_{V_\lambda} : V_\lambda \to V_\lambda$

è ovviamente unitariamente diagonalizzabile (è già diagonale ...) . Pertanto T è unitariamente diagonalizzabile

$$m_{bb}(T) = \begin{array}{c} k\left\{\vphantom{\begin{matrix}\lambda\\ \ddots\\ \lambda\end{matrix}}\right. \\ \\ \end{array} \left(\begin{array}{c|c} \overbrace{\begin{matrix}\lambda & & \\ & \ddots & \\ & & \lambda\end{matrix}}^{k} & 0 \\ \hline 0 & \begin{matrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_{n-k}\end{matrix} \end{array}\right)$$

$V_\lambda = \langle e_1 \dots e_k \rangle$ base ortonormale di autovettori

$k = m_g(\lambda)$

$V_\lambda^\perp = \langle f_1 \dots f_{n-k} \rangle$ base ortonormale diagonalizzante

base adattata alla decomposizione $V = V_\lambda \oplus V_\lambda^\perp$

Sia $b = (e_1 \dots e_k, f_1 \dots f_{n-k})$: è la base ortonormale di autovettori diagonalizzante

Ovviamente, $m_a(\lambda) = m_g(\lambda) \quad \forall \lambda \in \sigma(T)$

Sul teorema spettrale (versione reale)

Se si rinuncia ad inquadrare il teorema nell'ambito naturale, che è quello complesso, conviene riformulare così la prima condizione

Teorema spettrale (caso euclideo)

Sia $T \in End(V)$, $(V, \langle\,|\,\rangle)$ euclideo, T simmetrico. Allora

i) $P_c^T(\lambda) = \prod (\lambda_i - \lambda)$ $\qquad \lambda_i, \lambda \in \mathbb{R}$

(è la condiz. di triangolabilità ...)

ii) esiste una base ortonormale di V costituita da autovettori di T.

L'operatore di proiezione ortogonale

$$P_U : V \longrightarrow V$$
$$v \longmapsto v_U \; (= P_U(v))$$

è un operatore simmetrico:

Se $e = (\underbrace{e_1 \dots e_k}_{\text{base orton. di } U}, \underbrace{e_1^\perp \dots e_{n-k}^\perp}_{\text{base orton. di } U^\perp})$

e è una base orton. di $V = U \oplus U^\perp$

è $M_{ee}(P_U) = \left(\begin{array}{c|c} \begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix} & \\ \hline & \begin{matrix} 0 & & \\ & \ddots & \\ & & 0 \end{matrix} \end{array}\right) = \left(\begin{array}{c|c} I_k & 0 \\ \hline 0 & 0 \end{array}\right)$

(che è simmetrica)

Ancora sul teorema spettrale (formulazione equivalente)

$$\underset{\substack{\text{endom.}\\ \text{simmetrico}}}{T} = \sum_{\lambda_i \text{ distinti}} \lambda_i \, P_{V^T_{\lambda_i}}$$

operatore di proiezione ortogonale su $V^T_{\lambda_i}$

(questa è la versione che si generalizza in dimensione infinita (la teoria è dovuta a J. von Neumann)

$$\mathrm{Mee}(P_{V^T_{\lambda_i}}) = \begin{pmatrix} 0 & & & & & 0 \\ & \ddots & & & & \\ & & \boxed{I_{\dim(\lambda_i)}} & & & \\ & & & 0 & & \\ & & & & \ddots & \\ 0 & & & & & 0 \end{pmatrix}$$

$$e = (\underbrace{\cdots}_{\substack{\text{base orton.}\\ \text{di } V^T_{\lambda_1}}} \underbrace{\cdots}_{\substack{\text{base orton.}\\ \text{di } V^T_{\lambda_2}}} \cdots \underbrace{\cdots}_{\substack{\text{base orton}\\ \text{di } V^T_{\lambda_r}}})$$

$\lambda_1 \dots \lambda_r$ distinti

$$\mathrm{Mee}(T) = \begin{pmatrix} \boxed{\begin{matrix}\lambda_1 & 0 \\ 0 & \ddots \lambda_1\end{matrix}} & & 0 \\ & \ddots & \\ 0 & & \boxed{\begin{matrix}\lambda_r & 0 \\ 0 & \ddots \lambda_r\end{matrix}} \end{pmatrix} = \sum \lambda_i \, \mathrm{Mee}(P_{V^T_{\lambda_i}})$$

Notare che $P_{V^T_{\lambda_i}} P_{V^T_{\lambda_j}} = P_{V^T_{\lambda_j}} P_{V^T_{\lambda_i}} = P_{V^T_{\lambda_j}} \delta_{ij}$

Questa è la forma migliore per procedere al cosiddetto calcolo funzionale (operatoriale e matriciale) senza entrare in dettaglio (si veda un esempio poco oltre), diciamo che il calcolo funzionale permette di definire funzioni di un operatore (simmetrico, nella fattispecie)

$$f(T) := \sum_{\lambda_i \text{ distinti}} f(\lambda_i) \, P_{V^T_{\lambda_i}}$$

Teorema spettrale e forme quadratiche

Il teorema spettrale testè dimostrato ci dice che, se $T = T^t \in \mathrm{End}(V)$ $(K = \mathbb{R})$
(T simmetrico)

allora p_c^T caratterizza completamente la sua classe di similitudine, ovvero

p_c è un invariante completo per similitudine per operatori simmetrici (non in generale!)

Equivalentemente, lo spettro e la relativa molteplicità costituiscono un invariante completo (per operatori simmetrici)

Ciò è vero per una matrice simmetrica $A = A^t \in M_n(\mathbb{R})$, poiché la si può sempre interpretare come operatore simmetrico (relativo al prodotto scalare standard in $\mathbb{R}^n$)

Ora, dato T simmetrico, consideriamo la forma quadratica associata a T, q_T, così definita:

$$q_T(v) = \langle v \mid Tv \rangle$$

prodotto scalare in V, spazio euclideo

Fissata ora una base ortonormale $e = (e_1 \dots e_n)$ qualsiasi, e posto, al solito $v = \sum_i x_i e_i$, $a_{ij} = \langle e_i | T e_j \rangle$

e $A = (a_{ij})$,

$$q_T(v) \doteq x^t A x \qquad (a_{ij} = a_{ji})$$

Ora, in base al teorema spettrale, esiste una base ortonormale $e' = (e_1' \dots e_n')$ di autovettori di T: $Te_i' = \lambda_i e_i'$, ottenibile da e tramite un elemento $O \in O(V)$: esplicitamente

$O^{-1} = O^t$ $\qquad x = O x'$

(cf. disp. 3: cambio di base)

Si ha allora

$$q_T(v) = x^t A x = (Ox')^t A (Ox') =$$

$$= x'^t O^t A O x' \equiv x'^t \underset{\text{diagonale}}{D} x' = \sum_{i=1}^{n} \lambda_i x_i'^2$$

ossia

$$(\#) \quad \boxed{O^t A O = D} \qquad D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

autovalori

ma è pure da $O^t = O^{-1}$

$$(\#\#) \quad \boxed{O^{-1} A O = D}$$

Pertanto A e D risultano nello stesso tempo simili e congruenti.

Data ora $A \in M_n(\mathbb{R})$, $A = A^t$ e interpretandola direttamente come operatore simmetrico rispetto al prod. scalare standard troviamo che (ripetiamo)

Teorema Ogni matrice simmetrica $A = A^t \in M_n(\mathbb{R})$ è ortogonalmente simile ad una matrice diagonale, nel senso che $\exists\ O \in O(n, \mathbb{R})$ $(O^t = O^{-1})$ tale che $O^t A O = O^{-1} A O = D$, D diagonale

$$D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

Sulla diagonale principale compaiono gli autovalori di A (in un certo ordine) contati con la loro molteplicità.

Si dice anche: una matrice quadrata reale e simmetrica è ortogonalmente simile ad una matrice diagonale.

È anche vero che: una matrice quadrata reale e simmetrica è ortogonalmente congruente ad una matrice diagonale.

Osserviamo che, viceversa, partendo da una matrice diagonale $D \in M_n(\mathbb{R})$ (ovviamente simmetrica), si ha, posto

$$A := O^t D O \quad , \text{ con } \quad O \in O(n, \mathbb{R}),$$

che $A^t = (O^t D O)^t = O^t D^t (O^t)^t = O^t D O = A$

Si dice anche che $A = A^t$ è ortogonalmente diagonalizzabile.

Si noti che, in termini di forme quadratiche (e, di congruenza), prima sapevamo solo che $A = A^t$ era congruente ad una matrice diagonale:

$$M^t A M = D_1 \qquad D_1 \text{ diagonale}$$

ora sappiamo che M si può scegliere ortogonale: $M^t = M^{-1}$, sicché si ha anche similitudine con una matrice diagonale D_2 (in generale diversa!)

$$\underset{\overset{\shortparallel}{M^t}}{M^{-1}} A M = D_2$$

Pertanto, a livello di forme quadratiche, data q $(K = \mathbb{R})$, ha senso parlare di una forma canonica affine (fornita dal teorema di Sylvester) e di una forma canonica metrica (fornita dal teorema spettrale)

ovviamente, quest'ultima, rispetto ad un prodotto scalare di riferimento (q può sempre essere in tal caso interpretata come $q = q_T$, con $T = T^t$)

$$q \begin{cases} \longrightarrow \text{forma can. affine} \\ \longrightarrow \text{forma canonica metrica} \end{cases}$$

(rispetto a $\langle\, | \,\rangle$ fissato in V)

Concretamente, data $A = A^t$, letta come op. simmetrico, e considerata

$$q_A(v) \doteq \langle x | A x \rangle \quad (= x^t A x)$$

prod. scalare standard

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

coordinate di v rispetto alla base canonica

Si ha, rispetto ad una base ortonormale $e' = (e'_1 \dots e'_n)$ di autovettori di A: $A e'_i = \lambda_i e'_i$ (e relative coordinate x')

$$q_A(v) = \sum_{i=1}^{n} \lambda_i x_i'^2$$

$x' = \begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix}$

coordinate relative ad una base ortonormale di autovettori corrispondenti

autovalori di A.

è evidente allora che la forma canonica metrica determina quella affine, ma non viceversa: due forme quadratiche metricamente equivalenti sono anche affinemente equivalenti, ma non viceversa, in generale

A livello di matrici simmetriche; matrici simmetriche simili sono anche ortogonalmente simili, (e dunque ortogonalmente congruenti), dunque congruenti. Il viceversa è falso, in generale.

Corollario. Data una matrice simmetrica $A = A^t \in M_n(\mathbb{R})$, $(r = r(A))$

si ha

p = numero di autovalori positivi (contati con la loro molteplicità)
indice di positività

$r - p$ = numero di autovalori negativi (contati con la loro molteplicità)
indice di negatività

$\nu = n - r$ = numero di autovalori nulli ($\dim \operatorname{Ker} A$) (contati con la loro molteplicità)
indice di nullità

Dim. Immediata dalla discussione precedente e dal teorema di Sylvester.
Ricordiamo che, comunque, per determinare la segnatura di A, non è necessario ricorrere a tale criterio (ovvero, al calcolo del polinomio caratteristico).

Schema ricapitolativo (II)

$$A = A^t \in M_n(\mathbb{R})$$

matrice simmetrica

può essere diagonalizzata per

- congruenza
- ortogonale congruenza ≡ similitudine

congruenza:

- ☆ forma di Sylvester (forma canonica affine) a priori non entrano in gioco autovalori e autovettori
- ✧ invariante completo:
  - ☆ Segnatura

ortogonale congruenza ≡ similitudine:

- ☆ forma canonica metrica (relativa al prodotto scalare standard in $\mathbb{R}^n$)
  - Sono coinvolti in modo essenziale autovalori e autovettori
- ✧ invariante completo: spettro, con la relativa molteplicità ovvero: polinomio caratteristico P_c^T

È chiaro che il secondo modo è più fine del primo: conoscendo P_c^A, conosciamo i segni degli autovalori (e gli autovalori nulli) e quindi la segnatura.

Per calcolare quest'ultima non è necessario determinare $\sigma(A)$ e la relativa molteplicità, ovvero P_c^T

☆ Tabella riassuntiva (I)

$T = T^t$ endom. simmetrica $\longleftrightarrow$ A, B simmetriche

A e B corrispondono allo stesso endom. simmetrico espresso rispetto a basi ortonormali diverse

$A \sim B$ (teor. spettrale) $\Leftrightarrow$ $A \overset{\circ}{\sim} B$

$A \approx B$ $\Leftarrow$ $A \overset{\circ}{\approx} B$

corrispondono alla stessa q rispetto a basi diverse

Ex: $A = \begin{pmatrix} 1 & 0 \\ 0 & -4 \end{pmatrix}$ $B = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}$ segnatura $(2,1)$, $n=4$ $r=3$

$A \approx B$ ma $A \not\sim B$

similitudine $\sim$ ortogonale congruenza $\overset{\circ}{\approx}$ congruenza

$\sim \equiv \overset{\circ}{\sim}$ ≡ ortogonale congruenza

$\approx$ equivalenza metrica per forme quadratiche

q_A, q_B forme quadratiche $\longleftrightarrow$ q_T

b_A, b_B forme bil. associate

$\langle \cdot | T \cdot \rangle$

metricamente equivalenti

(stessa forma con. metrica) corrisp. allo stesso T

$q_A \overset{\circ}{\approx} q_B$

$q_A \approx q_B$ $\Leftarrow$

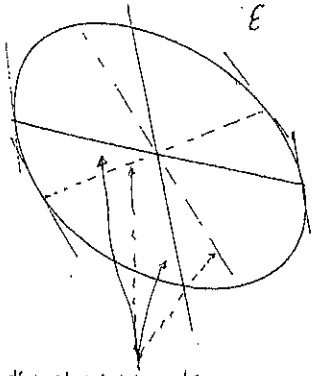
(stessa forma can. affine)

$\Leftrightarrow$ teorema di Sylvester

$(P_A, n_A - P_A) = (P_B, n_B - P_B)$ (nec. e suff. $n_A = n_B$)

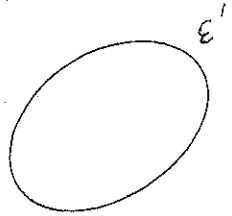
La differenza tra i due tipi di diagonalizzazione avrà una vivida interpretazione geometrica

Esempio (anticipando un po')

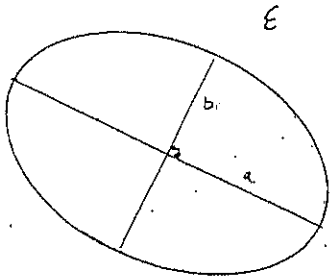


diametri coniugati

affinemente vi sarà, per esempio, una sola ellisse ecc.

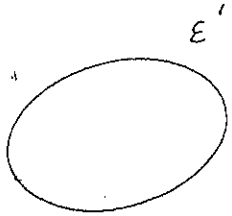


$\mathcal{E}$ ed $\mathcal{E}'$ sono affinemente la stessa conica



assi: diametri coniugati ortogonali (univocamente determinati) eccettuato il caso delle circonferenze..

metricamente, vi saranno per esempio, infinite ellissi (individuate dalla lunghezza dei loro semiassi



$\mathcal{E}$ ed $\mathcal{E}'$ non sono metricamente equivalenti

- Esempio. Sia $\underset{T}{\overset{\shortparallel}{A}} : \mathbb{R}^2 \to \mathbb{R}^2$

$(\mathbb{R}^2, \langle\ |\ \rangle_{\text{standard}})$ $\quad T \equiv A = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$ $\quad v = \begin{pmatrix} x \\ y \end{pmatrix}$

$$q(v) = (x \ \ y) \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} =$$

$$= x^2 + 4xy + 3y^2 \qquad (= \langle v | A v \rangle)$$

Determiniamo la forma canonica affine (di Sylvester) di A direttamente, col metodo di Gauss del completamento dei quadrati:

$$x^2 + 4xy + 3y^2 = x^2 + 4xy + 4y^2 - y^2 =$$

$$= (x + 2y)^2 - y^2 \equiv x'^2 - y'^2$$

(ove si ponga $\begin{array}{l} x' = x + 2y \\ y' = y \end{array}$

ovvero $\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$)

$\Rightarrow$ segnatura: $(1, 1)$

Determiniamone la forma canonica metrica (relativamente al prodotto scalare di riferimento), A è letto come operatore simmetrico

Calcoliamo allora

$$p_C^A(\lambda) = \begin{vmatrix} 1-\lambda & 2 \\ 2 & 3-\lambda \end{vmatrix} = (1-\lambda)(3-\lambda) - 4 =$$

$$= \lambda^2 - 4\lambda + 3 - 4 = \lambda^2 - 4\lambda - 1$$

equazione caratteristica: $\lambda^2 - 4\lambda - 1 = 0$

$$\lambda = 2 \pm \sqrt{4+1} = 2 \pm \sqrt{5} \equiv \lambda_\pm$$

forma canonica metrica:

$$\underset{\substack{\vee \\ 0}}{\lambda_+} \xi^2 + \underset{\substack{\wedge \\ 0}}{\lambda_-} \eta^2$$

($\Rightarrow$ ancora segnatura $= (1,1)$)

Determiniamo una base ortonormale di autovettori e $O \in O(2)$ tale che

$$\underset{\substack{\| \\ O^{-1}}}{O^t} A O = D = \begin{pmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{pmatrix}$$

Calcoliamo gli autospazi $V_{\lambda_\pm}$

basta determinare V_{λ_+}: si ha necessariamente $V = V_{\lambda_+} \oplus V_{\lambda_+}^\perp = V_{\lambda_+} \oplus V_{\lambda_-}$

$$(A - \lambda_+) v = \underline{0}$$

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

$\Rightarrow$ (basta una sola equazione)

$$(1-\lambda_+)x + 2y = 0$$

$$y = \frac{\lambda_+ - 1}{2} x = \underbrace{\frac{1+\sqrt{5}}{2}}_{\varphi} x$$

(sezione aurea..)

Si ha: $V_{\lambda_+} = \left\langle \begin{pmatrix} 1 \\ \varphi \end{pmatrix} \right\rangle$

Costruiamo una base ortonormale

$$\left\| \begin{pmatrix} 1 \\ \varphi \end{pmatrix} \right\| = \sqrt{1+\varphi^2} = \sqrt{1 + \frac{1+5+2\sqrt{5}}{4}} =$$

$$= \sqrt{\frac{10+2\sqrt{5}}{4}} = \sqrt{\frac{5+\sqrt{5}}{2}} \equiv \beta$$

$V_{\lambda_+} = \langle e_1' \rangle$ con $e_1' = \frac{1}{\beta} \begin{pmatrix} 1 \\ \varphi \end{pmatrix}$

$(\|e_1'\| = 1)$

Si noti che ponendo $e_1'' = \frac{e_1'}{\sqrt{\lambda_+}}$, $e_2'' = \frac{e_2'}{\sqrt{-\lambda_-}}$

otteniamo una nuova base di Sylvester.

Le curve di livello di q sono iperboli aventi gli stessi asintoti (anche la coppia di asintoti va considerata.) : questi ultimi costituiscono il <u>cono isotropo</u> $q^{-1}(0)$

a..) $= \{ v \in \mathbb{R}^2 \mid q(v) = 0 \}$; determiniamolo:

$$(*) \quad x^2 + 4xy + 3y^2 = 0$$

Poniamo $y = mx \quad x \neq 0$

(attenzione, potremmo, ma qui non accadrà, prendere $x=0$, cioè l'asse y..) ; si ha subito, sostituendo in $(*)$: $3m^2 + 4m + 1 = 0$

che fornisce immediatamente $m = \begin{cases} -\frac{1}{3} \\ -1 \end{cases}$

$$\Rightarrow \quad q^{-1}(0) = \underset{\langle \binom{1}{-\frac{1}{3}} \rangle}{\overset{W_1}{\parallel}} \cup \underset{\langle \binom{1}{-1} \rangle}{\overset{W_2}{\parallel}}$$

$$V_{\lambda_-} = V_{\lambda_+}^{\perp} = \langle e_2' \rangle$$

con $e_2' \underset{\text{per esempio}}{=} = \frac{1}{\beta} \begin{pmatrix} \varphi \\ -1 \end{pmatrix}$

Poniamo Ω : $e_1 \longmapsto e_1'$
$e_2 \longmapsto e_2'$

$$\Omega = \begin{pmatrix} \frac{1}{\beta} & \frac{\varphi}{\beta} \\ \frac{\varphi}{\beta} & -\frac{1}{\beta} \end{pmatrix} = M_{ee'} = m_{ee'}(I)$$

$T \equiv A$

$D = m_{e'e'}(T)$

$A = m_{ee}(T)$

Si ha

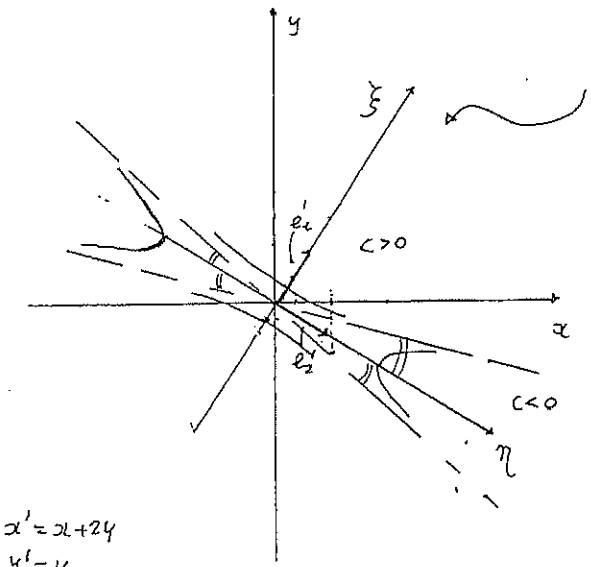
$$\begin{tikzcd} \mathbb{R}^2 \arrow[r, "D"] \arrow[d, "\Omega"'] & \mathbb{R}^2 \\ \mathbb{R}^2 \arrow[r, "A"] & \mathbb{R}^2 \arrow[u, "\Omega^{-1} = \Omega^t"'] \end{tikzcd}$$

$$D = m_{e'e'}(T) = \underset{\substack{\Omega^{-1} \\ \parallel \\ \Omega^t}}{m_{e'e}(I)} \; \underset{A}{m_{ee}(T)} \; \underset{\Omega}{m_{ee'}(I)}$$

$$D = \Omega^t A \Omega \qquad \Omega = O$$

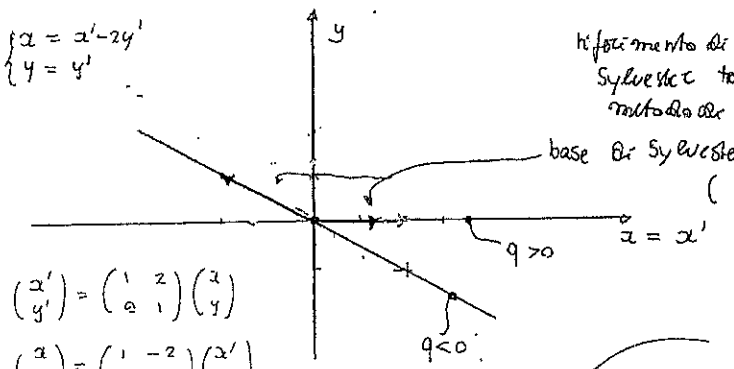
[Si verifichi anche direttamente che $\Omega D = A \Omega$..]

Si noti che, essendo $\det \mathcal{R} = -1 < 0$ le basi (e_1, e_2) e (e_1', e_2') non sono equiorientate



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

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



Riferimento di Sylvester trovato col metodo di Gauss

base di Sylvester $\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} -2 \\ 1 \end{pmatrix} \right)$

è la base del nuovo riferimento

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

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

$m_{e'e}(I) = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ $\quad m_{ee'}(I) = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}$

Diamo infine un altro procedimento per determinare una base diagonalizzante per congruenza: di fatto è il metodo utilizzato nella dimostrazione del teorema.

Partiamo da un vettore non isotropo: (per cambiare) prendiamo per es. $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$

Determiniamo $\langle v_1 \rangle^{\perp_A}$: risolviamo
(è una retta per l'origine)

$$(x \quad y) \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0 \quad \Rightarrow$$

$$(x \quad y) \begin{pmatrix} 2 \\ 3 \end{pmatrix} = 0 \quad \Rightarrow \quad 2x + 3y = 0$$

$y = -\frac{2}{3}x$. Prendiamo, per esempio $v_2 = \begin{pmatrix} 3 \\ -2 \end{pmatrix} \in \langle v_1 \rangle^{\perp_A}$; allora (v_1, v_2) è una base diagonalizzante (per congruenza).

Ricalcoliamo la segnatura.

$$v_1^t A v_2 = (0 \quad 1) \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = (0 \quad 1) \begin{pmatrix} 2 \\ 3 \end{pmatrix} =$$

$= 3 > 0$;

$$v_2^t A v_2 = (3 \quad -2)\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}\begin{pmatrix} 3 \\ -2 \end{pmatrix} =$$

stato

$$= (3 \quad -2)\begin{pmatrix} 3-4 \\ 6-6 \end{pmatrix} = (3 \quad -2)\begin{pmatrix} -1 \\ 0 \end{pmatrix} =$$

$= -3 < 0$ $\Rightarrow$ ancora (ovviamente!) troviamo $(1,1)$. Una base di Sylvester è allora fornita da $(\tilde{v}_1, \tilde{v}_2)$,

con $\tilde{v}_1 = \dfrac{v_1}{\sqrt{3}}$, $\tilde{v}_2 = \dfrac{v_2}{\sqrt{3}}$

$$\left(\tilde{v}_1^t A \tilde{v}_1 = 1 \ , \ \tilde{v}_2^t A \tilde{v}_2 = -1 \right)$$

INCISO

Sia dato $T = T^t \in End(V)$ in $(V, \langle | \rangle)$ spazio euclideo; consideriamo q_T, forma quadratica ad esso associata, rispetto a $\langle | \rangle$

$$q_T(v) = \langle v | Tv \rangle$$

T è detto operatore <u>positivo</u> se q_T è definita positiva. In base al teorema di Sylvester, si avrà pure T positivo $\Leftrightarrow$ $T = M^t M$, per $M \in GL(V)$. $\Leftrightarrow$: i suoi autovalori sono positivi

$$q_T(v) = \langle v | M^t M v \rangle = \langle Mv | Mv \rangle$$

Risulta pertanto definito un nuovo prodotto scalare

$$\langle v | w \rangle_M := \langle Mv | Mw \rangle \qquad M \in GL(V)$$

Di fatto, tutti i prodotti scalari si possono costruire così a partire da uno dato.

Stesso discorso a livello di matrici:

$A = A^t \in M_n(\mathbb{R})$ è <u>positiva</u> (ossia, dà luogo ad un prodotto scalare $\Leftrightarrow$ $A = M^t M$, con $M \in GL_n(\mathbb{R})$ $\Leftrightarrow$ A ha autovalori tutti positivi.

Si ricordi poi il criterio di Sylvester (vedi disp. 5)

* calcolo funzionale; esempi

Sia $A \in M_n(\mathbb{R})$ diagonalizzabile:

$\exists\ M \in GL_n(\mathbb{R})$ tale che

$$M^{-1} A M = D \quad \Rightarrow \quad M D M^{-1} = A$$

Calcoliamo A^2 si ha

$$A^2 = M D \underbrace{M^{-1} \cdot M}_{= I} D M^{-1} = M D^2 M^{-1}$$

Induttivamente si trova

$$A^k = M D^k M^{-1}$$

Per un polinomio generico $\quad p(D) = \begin{pmatrix} p(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & p(\lambda_n) \end{pmatrix}$

$$p(A) = M\, p(D)\, M^{-1}$$

Ciò suggerisce di porre, per una funzione "generica"

$$f(A) := M f(D) M^{-1}$$

es: $f = \exp$

$$\exp A := M \underbrace{\exp D}_{= \begin{pmatrix} e^{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_n} \end{pmatrix}} M^{-1}$$

INCISO Si noti che $\exp A$ può essere definita per ogni matrice direttamente

$$\exp A = \sum_{k=0}^{\infty} \frac{A^k}{k!}$$

Se A è diagonalizzabile tale definizione riproduce il risultato precedente

(La serie converge, a patto di una norma su $M_n(\mathbb{R})$ con la proprietà $\|AB\| \le \|A\|\,\|B\|$

Sicché $\|A^k\| \le \|A\|^k$

Ad esempio $\|A\| = \sup_{\|v\| \le 1} \|Av\|$ "norma del sup" o norma "C^*".

poiché vale $\|A^t A\| = \|A\|^2$

o la Norma Hilbert-Schmidt, indotta dal prodotto scalare

$$\langle A | B \rangle := -\mathrm{Tr}\, A^t B$$

$$\|A\|_{HS} = \sqrt{\langle A | A \rangle} = \sqrt{\mathrm{Tr}\, A^t A}$$

Tutto ciò è importantissimo per varie applicazioni Eq. differenziali, meccanica quantistica ecc.

Esercizio:

Calcolare $\sin A$, con

$$A = \begin{pmatrix} \frac{3}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{3}{2} \end{pmatrix}$$

Sol. La matrice A è diagonalizzabile (è simmetrica: di più, è ortogonalmente diagonalizzabile)

Mostrare che $\lambda = 1, 2$ e che

$$M = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \text{ è tale che}$$

$$M \underset{\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}}{\underset{\parallel}{D}} \underset{M^t}{\underset{\parallel}{M^{-1}}} = A$$

pertanto

$$\sin A = \ldots = M \cdot \sin D \cdot M^{-1} =$$

$$= M \begin{pmatrix} \sin 1 & 0 \\ 0 & \sin 2 \end{pmatrix} M^{-1} =$$

$$\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \sin 1 & 0 \\ 0 & \sin 2 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} =$$

$$= \begin{pmatrix} \frac{\sin 1}{\sqrt{2}} & \frac{\sin 2}{\sqrt{2}} \\ -\frac{\sin 1}{\sqrt{2}} & \frac{\sin 2}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{\sin 1 + \sin 2}{2} & \frac{-\sin 1 + \sin 2}{2} \\ \frac{-\sin 1 + \sin 2}{2} & \frac{\sin 1 + \sin 2}{2} \end{pmatrix}$$

Operativamente, è utile la seguente osservazione.

Data una base qualsiasi, la si può sempre dichiarare ortonormale, dando vita con ciò ad un nuovo prodotto scalare.

Sia, per ex. in $\mathbb{R}^n$, il prodotto scalare standard $\langle x \mid y \rangle = x^t y$

Data una nuova base e le relative coordinate con apici $x', y' \ldots$ costruiamo un nuovo prodotto scalare $\langle\langle \mid \rangle\rangle$ così:

$$\langle\langle x' \mid y' \rangle\rangle = x'^t y'$$

Ma, da $x' = Mx$ $\quad y' = My$

segue

$$x'^t y' = x^t M^t M y \quad , \text{ pertanto}$$

$$\langle\langle x, y \rangle\rangle = x^t M^t M y = \langle Mx \mid My \rangle$$

in accordo con la discussione precedente.

FINE DELL'INCISO

- Esempio

Torniamo all'operatore $T: \mathbb{R}^2 \to \mathbb{R}^2$

tale che $m_{e'e'}(T) = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$

$e_1' = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad e_2' = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

Si era osservato che T non era un operatore simmetrico. Tuttavia, avendo, rispetto ad una certa base, matrice rappresentativa simmetrica, risulta certamente diagonalizzabile.

⚠ Il punto è che non è ortogonalmente diagonalizzabile rispetto al prodotto scalare fissato (quello standard), poiché una matrice ortogonale muta basi ortonormali in basi ortonormali. Nulla vieta che si cambi il prodotto scalare, dichiarando la base e' ortonormale: rispetto a questo nuovo prodotto scalare, certamente T risulterà ortogonalmente diagonalizzabile.

Tuttavia ricordiamo, che, data direttamente una matrice simmetrica $A = A^t \in M_n(\mathbb{R})$, essa risulta ortogonalmente diagonalizzabile poiché si può interpretare canonicamente come endomorfismo simmetrico (rispetto alla base canonica, ortonormale rispetto al prodotto scalare standard).

Non ha alcun senso chiedersi se essa possa provenire da un endomorfismo non simmetrico (a meno che questo non sia esplicitamente dato).

☆ Il principio di Rayleigh

Sia dato uno spazio euclideo $(V, \langle | \rangle)$, $\dim V = n \geq 1$ e $T = T^t$ simmetrico

Sia $e = (e_1 \dots e_n)$ una base <u>ortonormale</u> di autovettori di T:

$$T e_i = \lambda_i e_i$$

Consideriamo, per $v \neq \Omega$, il

<u>quoziente di Rayleigh</u>: $\dfrac{\langle v | Tv \rangle}{\langle v | v \rangle}$

e studiamolo come funzione di $v \in V$:

Di fatto, possiamo restringerci alla

<u>Sfera unitaria</u> $S_n(V) := \{ v \in V \mid \langle v | v \rangle = 1 \}$

Sicché si ha:

$$S_n(V) \ni v \longmapsto \langle v | Tv \rangle = q_T(v)$$

Posto $v = \sum_{i=1}^n x_i e_i$ si ha, in $\mathbb{R}^n$

$$\underset{\{x : \sum_{i=1}^n x_i^2 = 1\}}{\underset{\parallel}{S_n(\mathbb{R}^n)}} \ni x \longmapsto \sum_{i=1}^n \lambda_i x_i^2$$

Per semplificare la discussione supponiamo che i λ_i siano distinti e che

$$\lambda_1 < \lambda_2 < \dots < \lambda_n$$

Fissiamo λ_j. Si ha, successivamente

$$\sum_{i=1}^n \lambda_i x_i^2 = \lambda_j \underbrace{\sum_{i=1}^n x_i^2}_{\substack{\parallel \\ 1}} + \sum_{i=1}^{j-1} (\lambda_i - \lambda_j) x_i^2 + \sum_{i=j+1}^n (\lambda_i - \lambda_j) x_i^2$$

ovvero

$$\sum_{i=1}^n \lambda_i x_i^2 = \lambda_j + \underbrace{\sum_{i=1}^{j-1} (\lambda_i - \lambda_j) x_i^2}_{\substack{\leq \\ 0}} + \underbrace{\sum_{i=j+1}^n (\lambda_i - \lambda_j) x_i^2}_{\substack{\geq \\ 0}}$$

Se ora poniamo $(*)$ $x_i = 0$ per $i = 1, 2 \dots j-1$

ovvero $v \perp \langle e_1 \dots e_{j-1} \rangle$

otteniamo

$$\sum_{i=1}^n \lambda_i x_i^2 = \sum_{i=j}^n \lambda_i x_i^2 = \lambda_j + \underbrace{\sum_{i=j+1}^n \overbrace{(\lambda_i - \lambda_j)}^{>0} x_i^2}_{\substack{\geq \\ 0}}$$

espressione che è minima $(= \lambda_j)$

$\iff$ $x_i = 0$ per $i = j+1 \dots n$, che,

insieme a (*), dà

$$x_j^2 = 1 \implies x_j = \pm 1$$

In altre parole, abbiamo provato che

Caratterizzazione estremale degli autovalori

$$\boxed{\begin{aligned} \lambda_j &= \min_{\substack{\|v\|=1 \\ v \in \langle e_j \dots e_n \rangle \\ \| \\ \langle e_1 \dots e_{j-1} \rangle^\perp}} q_T(v) \\ &\text{in particolare} \\ \lambda_1 &= \min_{\|v\|=1} q_T(v) \\ \lambda_n &= \max_{\|v\|=1} q_T(v) \end{aligned}}$$

** Principio di Rayleigh

ovvero gli autovettori emergono come punti critici vincolati di q_T (sulla sfera unitaria): gli autovalori emergono come valori critici

Il tutto si vedrebbe in modo ancora più elegante tramite la teoria dei moltiplicatori di Lagrange

Discutiamo il caso particolare $n = 2$

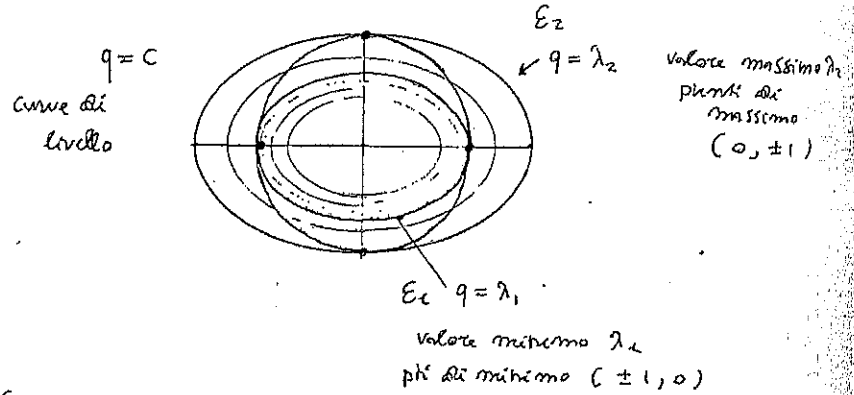
Cominciamo con $\quad 0 < \lambda_1 < \lambda_2 \qquad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \equiv \begin{pmatrix} x \\ y \end{pmatrix}$

$$q(v) = \lambda_1 x^2 + \lambda_2 y^2 =$$

$$= \lambda_1 (x^2+y^2) + (\lambda_2 - \lambda_1) y^2 = \lambda_1 + \overbrace{(\lambda_2 - \lambda_1)}^{>0} y^2$$

$$= \lambda_2 (x^2+y^2) + (\lambda_1 - \lambda_2) x^2 = \lambda_2 + \underbrace{(\lambda_1 - \lambda_2)}_{<0} x^2$$

ove $x^2 + y^2 = 1$



$\mathcal{E}_1$:

$$\lambda_1 x^2 + \lambda_2 y^2 = \lambda_1 \qquad x^2 + \frac{y^2}{\left[\sqrt{\frac{\lambda_1}{\lambda_2}}\right]^2} = 1$$

$a = 1$
$b = \sqrt{\frac{\lambda_1}{\lambda_2}} < 1$

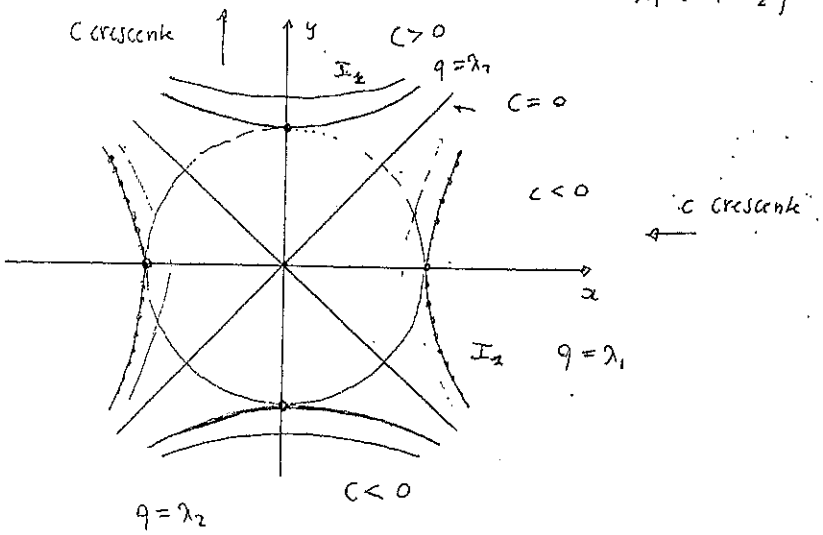
$\mathcal{E}_2$: $\lambda_1 x^2 + \lambda_2 y^2 = \lambda_2$

$$\frac{x^2}{\left[\sqrt{\frac{\lambda_2}{\lambda_1}}\right]^2} + y^2 = 1$$

$$a = \sqrt{\frac{\lambda_2}{\lambda_1}} > 1$$

$n = 2 \qquad \lambda_1 < 0 < \lambda_2$

$q(v) = \lambda_1 x^2 + \lambda_2 y^2$



$I_1 \qquad \lambda_1 x^2 + \lambda_2 y^2 = \lambda_1 \qquad x^2 + \dfrac{y^2}{\frac{\lambda_1}{\lambda_2}} = 1$

$$x^2 - \frac{y^2}{\left[\sqrt{\frac{|\lambda_1|}{\lambda_2}}\right]^2} = 1$$

$I_2 \qquad \lambda_1 x^2 + \lambda_2 y^2 = \lambda_2 \qquad \dfrac{x^2}{\frac{\lambda_2}{\lambda_1}} + y^2 = 1$

$$-\frac{x^2}{\left[\sqrt{\frac{\lambda_2}{|\lambda_1|}}\right]^2} + y^2 = 1$$

In dimensione 2 si può procedere anche così:

Sia $\begin{cases} x = \cos\varphi \\ y = \sin\varphi \end{cases} \qquad \varphi \in [0, 2\pi)$

$x^2 + y^2 = \cos^2\varphi + \sin^2\varphi = 1$

consideriamo la funzione

$$Q(\varphi) := \lambda_1 \underbrace{\cos^2\varphi}_{x^2} + \lambda_2 \underbrace{\sin^2\varphi}_{y^2}$$

Determiniamone i punti e i valori critici.

$$Q'(\varphi) = -2\lambda_1 \cos\varphi \sin\varphi + 2\lambda_2 \sin\varphi \cos\varphi$$
$$= 2(\lambda_2 - \lambda_1) \sin\varphi \cos\varphi =$$
$$= (\lambda_2 - \lambda_1) \sin 2\varphi$$
$$Q''(\varphi) = 2(\lambda_2 - \lambda_1) \cos 2\varphi$$

Si ha $Q'(\varphi) = 0$ per $\varphi = 0, \frac{\pi}{2}, \pi, \frac{3}{2}\pi$

$0, \pi$ pti di min. relativo (e assoluto) $\rightsquigarrow (\pm 1, 0)$
valore minimo: λ_1

$\frac{\pi}{2}, \frac{3}{2}\pi$ pti di max. relativo (e assoluto)
valore massimo: λ_2 $\rightsquigarrow (0, \pm 1)$

In accordo, ovviamente, con quanto trovato precedentemente.