Note di Algebra Lineare

II Modulo

Luca Barbieri Viale

http://users.unimi.it/barbieri/

Dipartimento di Matematica
Università degli Studi di Milano

Prerequisiti II Modulo

Assumo che il lettore abbia una certa domestichezza con matrici, sistemi lineari e spazi vettoriali. Assumo noti i rudimenti di analisi reale e complessa, radici di equazioni poliniomali e loro molteplicità, e "teorema fondamentale dell'algebra" che richiamo brevemente.

Numeri Complessi

Ricordiamo che se $z \in \mathbb{C}$ è un numero complesso questo può essere rappresentato da una coppia di numeri reali (a, b) mediante l'espressione $z = a + ib$ dove $i^2 = -1$. Questo identifica $\mathbb{C}$ con un $\mathbb{R}$-spazio vettoriale di dimensione 2 ma bisogna sempre ricordare che $\mathbb{C}$ stesso è un campo con il prodotto definito dalla moltiplicazione di numeri complessi $z \cdot z' = aa' - bb' + (ab' + a'b)i$ qui $z' = a' + ib'$. Una rappresentazione grafica di $\mathbb{C}$ è anche possibile mediante la forma trigonometrica $z = \rho(\cos(\theta) + i\sin(\theta))$ dove $\rho = |z|$ e θ è un argomento di z. Inoltre $z = \rho e^{i\theta}$ ed in tale modo si vede che $z^n = \rho^n(\cos(n\theta) + i\sin(n\theta))$, formule utili per risolvere alcune equazioni.[1] In generale si ha:

Teorema *Dato un polinomio $P(t) \in \mathbb{C}[t]$ di grado ≥ 1 esiste sempre un numero $z \in \mathbb{C}$ tale che $P(z) = 0$.*

Questo è detto teorema fondamentale dell'algebra, la sua dimostrazione non è elementare. Si deduce però facilmente che

$$P(t) = c(t - z_1)^{m_1} \cdots (t - z_r)^{m_r}$$

dove $c \in \mathbb{C}$ è una costante, i numeri complessi $z_1, \ldots, z_r$ sono appunto le radici di $P(t) = 0$ ed i numeri naturali $m_1, \ldots m_r$ le corrispondenti molteplicità tali che $m_1 + \cdots + m_r =$ grado di $P(t)$. Inoltre, se $P(t) \in \mathbb{R}[t]$ ha coefficienti reali, ovvero $P(t) = a_0 + a_1 t + \cdots + a_n t^n$ con $a_k \in \mathbb{R}$, allora

- $z = a + ib$ è radice se e solo se $\overline{z} = a - ib$ è radice, e
- se z ha molteplicità m anche $\overline{z}$ ha molteplicità m.

Infatti, $\overline{P(z)} = \overline{a_0 + a_1 z + \cdots + a_n z^n} = \overline{a_0} + \overline{a_1 z} + \cdots + \overline{a_n z^n} = a_0 + a_1\overline{z} + \cdots + a_n\overline{z}^n = P(\overline{z})$, poichè $\overline{a_k} = a_k$ in quanto $a_k \in \mathbb{R}$.

[1] Le equazioni $X^n = \alpha$ si risolvono infatti scrivendo $\alpha \neq 0$ in forma trigonometrica: hanno sempre n-radici distinte che si dispongono su i vertici di un poligono regolare ad n-lati.

Programma II Modulo

In questa seconda parte analizziamo, innanzitutto, il concetto di ortogonalità in spazi vettoriali, mediante l'introduzione di prodotti scalari. Tale concetto è relativo ma sufficientemente naturale da permettere la costruzione di basi ortonormali mediante un semplice algoritmo. Presentiamo inoltre lo studio degl'operatori lineari e delle loro matrici associate, provvedendo un criterio affinché tali matrici possano risultare più semplici possibile (ovvero diagonali!). Questo permetterà di ricavare informazioni da particolari simmetrie delle matrici e trattare le forme quadratiche anch'esse mediante opportune matrici.

1 Algoritmo di Gram-Schmidt

Desideriamo ora indagare le proprietà metriche di spazi vettoriali reali trasportando quanto possibile dell'intuizione geometrica di $\mathcal{V}_2$ e $\mathcal{V}_3$. Come abbiamo visto per i vettori geometrici, si possono determinare sottospazi ortogonali, mediante il prodotto scalare, e definire proiezioni e riflessioni in modo naturale. Vogliamo dunque isolare il concetto di prodotto scalare, in generale, per spazi vettoriali qualunque (anche non di tipo finito!) e definire proiezioni e riflessioni di vettori su sottospazi. Innanzitutto, però, ritorniamo alla somma e intersezione di sottospazi.

Decomposizioni di spazi vettoriali

Sussiste la seguente relazione tra le dimensioni dell'intersezione e della somma di due sottospazi.

Teorema 1.1 (Grassmann) *Se $V_0, V_1 \subset V$ sono sottospazi di tipo finito si ha*

$$\dim(V_0 + V_1) = \dim(V_0) + \dim(V_1) - \dim(V_0 \cap V_1)$$

Dimostrazione: Sia $\dim(V_0 \cap V_1) = t$ e sia $w_1, \dots, w_t$ una base di $V_0 \cap V_1$. Completiamo $w_1, \dots, w_t$ a base di V_0 sia $w_1, \dots, w_t, u_{t+1}, \dots, u_r$ quindi $\dim(V_0) = r$; completiamo a base di V_1 sia $w_1, \dots, w_t, v_{t+1}, \dots, v_s$ quindi $\dim(V_1) = s$. Basta ora verificare che la successione

$$w_1, \dots, w_t, u_{t+1}, \dots, u_r, v_{t+1}, \dots, v_s$$

è una base di $V_0 + V_1$. Infatti, chiaramente

$$V_0 + V_1 = L\{w_1, \dots, w_t, u_{t+1}, \dots, u_r, v_{t+1}, \dots, v_s\}$$

e $w_1, \dots, w_t, u_{t+1}, \dots, u_r, v_{t+1}, \dots, v_s$ sono $r + s - t$ vettori linearmente indipendenti (perchè!?). $\odot$

Da questa formula segue che se $V_0 \cap V_1 = 0$ allora $\dim(V_0 + V_1) = \dim(V_0) + \dim(V_1)$. Vogliamo generalizzare a più sottospazi queste proprietà.

$\boxed{\odot}$ Attenzione! Possiamo avere tre sottospazi V_0, V_1, V_2 tali che $V_i \neq V_j$ e $V_i \cap V_j = 0$ per $i \neq j$ ma $\dim(V_0 + V_1 + V_2) \neq \dim(V_0) + \dim(V_1) + \dim(V_2)$.

Somma diretta

Siano $V_0, V_1, \dots, V_r \subset V$ sottospazi. Sono fatti equivalenti

- $\dim(V_0 + \cdots + V_r) = \dim(V_0) + \cdots + \dim(V_r)$
- scelte basi di $V_0, V_1, \dots, V_r$ la successione di tutti i vettori delle singole basi è una base di $V_0 + \cdots + V_r$
- $V_i \cap (V_0 + \cdots \widehat{V_i} \cdots + V_r) = 0$ per ogni $0 \leq i \leq r$, dove $\widehat{V_i}$ indica che si omette V_i nella somma.

Se vale una qualunque condizione equivalente qui sopra esposta ogni vettore dello spazio somma $v \in V_0 + \cdots + V_r$ si scrive in modo unico come somma di vettori $v_0 + \cdots + v_r$ dove $v_i \in V_i$ per $0 \leq i \leq r$. Diremo allora che tale somma è $\boxed{\textit{diretta}}$. In particolare, scriveremo

$$V = V_0 \oplus \cdots \oplus V_r$$

per indicare che V stesso è somma diretta di sottospazi $V_0, V_1, \dots, V_r \subset V$. Tale scrittura di V come somma diretta di suoi sottospazi viene anche detta $\boxed{\textit{decomposizione}}$ di V in somma diretta. Se ad esempio V ha base $v_1, \dots, v_n$ ciascun singolo vettore v_i genera un sottospazio V_i e V si decompone nella somma diretta dei V_i.

Prodotto scalare

Al fine d'indagare le proprietà metriche di spazi vettoriali reali (e loro corrispondenti decomposizioni) sia V un $\mathbb{R}$-spazio vettoriale (non necessariamente di tipo finito) e consideriamo un'operazione $(u, v) \mapsto \langle u, v\rangle \in \mathbb{R}$ ovvero un'applicazione

$$\langle\ ,\ \rangle : V \times V \to \mathbb{R}$$

che sia formalmente identica al prodotto scalare tra vettori geometrici.

Definizione 1.2 *Diciamo* $\boxed{\text{prodotto scalare}}$ *una operazione* $\langle\ ,\ \rangle$ *su un* $\mathbb{R}$-*spazio vettoriale* V *tale che sia*

Simmetrica $\langle u, v\rangle = \langle v, u\rangle$ *per ogni* $u, v \in V$*;*

Bilineare $\langle \lambda u + \mu v, w\rangle = \lambda\langle u, w\rangle + \mu\langle v, w\rangle$ *per ogni* $u, v, w \in V$ *e* $\lambda, \mu \in \mathbb{R}$*;*

Positiva $\langle v, v\rangle \geq 0$ *per ogni* $v \in V$ *e* $\langle v, v\rangle = 0 \Leftrightarrow v = 0$*.*

Uno spazio vettoriale V con un prodotto scalare[2] viene anche detto prehilbertiano. Dalla positività del prodotto scalare si ottiene una $\boxed{\textit{norma}}$ su V ponendo

$$\|v\| \stackrel{\text{def}}{=} \sqrt{\langle v, v\rangle}$$

tale che

- $\|v\| = 0 \Leftrightarrow v = 0$
- $\|\lambda v\| = |\lambda|\|v\|$
- $\|u + v\|^2 + \|u - v\|^2 = 2\|u\|^2 + 2\|v\|^2$ che viene detta identità del parallelogramma, e infine
- $4\langle u, v\rangle = \|u + v\|^2 - \|u - v\|^2$

Questa ultima formula permette di ricavare il prodotto scalare dalla norma.

Teorema 1.3 *Sia* V *uno spazio vettoriale con un prodotto scalare. Per ogni* $u, v \in V$ *si hanno le seguenti diseguaglianze*

[2] Nel caso di $\mathbb{C}$-spazi vettoriali la proprietà di simmetria deve esser però modificata con la simmetria hermitiana $\langle u, v\rangle = \overline{\langle v, u\rangle}$ e la teoria risulta simile.

Diseguaglianza di Cauchy-Schwarz:

$$|\langle u, v\rangle| \leq \|u\|\|v\|$$

inoltre $|\langle u, v\rangle| = \|u\|\|v\| \Leftrightarrow$ *u e v sono linearmente dipendenti;*

Diseguaglianza triangolare:

$$\|u+v\| \leq \|u\| + \|v\|$$

Dimostrazione: Siccome $\langle u, 0\rangle = \langle 0, u\rangle = 0$ possiamo supporre $v \neq 0$ per dimostrare la prima diseguaglianza. Se $u = \lambda v$ allora si ha

$$|\langle \lambda v, v\rangle| = |\lambda|\langle v, v\rangle = |\lambda|\|v\|^2 = \|\lambda v\|\|v\|$$

In generale, sia $x \in \mathbb{R}$ un qualunque numero reale

$$0 \leq \|u + xv\|^2 = \langle u + xv, u + xv\rangle = \langle u, u\rangle + 2\langle u, v\rangle x + \langle v, v\rangle x^2 \stackrel{\text{def}}{=} p(x)$$

Il polinomio di secondo grado $p(x)$ risulta sempre positivo o nullo e quindi $\langle u, v\rangle^2 - \|u\|^2\|v\|^2 \leq 0$. Infine, se $|\langle u, v\rangle| = \|u\|\|v\|$ allora il polinomio $p(x)$ è un quadrato che si annulla per qualche numero reale $x = \lambda$ e dunque $\|u + \lambda v\| = 0$ da cui segue che $u + \lambda v = 0$. La diseguaglianza triangolare segue facilmente da Cauchy-Schwarz

$$\|u+v\|^2 \leq \|u\|^2 + 2|\langle u, v\rangle| + \|v\|^2 \leq \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 = (\|u\| + \|v\|)^2$$

$\odot$

Dalla diseguaglianza di Cauchy-Schwarz, se $u, v \neq 0$, segue che

$$-1 \leq \frac{\langle u, v\rangle}{\|u\|\|v\|} \leq 1$$

e quindi possiamo definire l'angolo $\widehat{uv}$ tra due vettori non nulli in modo che

$$u \perp v \Leftrightarrow \langle u, v\rangle = 0 \Leftrightarrow \widehat{uv} = \frac{\pi}{2} \tag{1}$$

In questo caso possiamo dire che u e v sono $\boxed{\textit{ortogonali}}$ osservando che questa nozione dipende dalla scelta di un prodotto scalare.

Lemma 1.4 *Se vettori $v_1, \dots v_r$ non nulli sono due a due ortogonali allora sono linearmente indipendenti.*

Dimostrazione: Infatti, supponiamo $\langle v_i, v_j \rangle = 0$ per $i \neq j$, e sia $\lambda_1 v_1 + \cdots + \lambda_r v_r = 0$. Allora $\langle \lambda_1 v_1 + \cdots \lambda_r v_r, v_k \rangle = \lambda_k \langle v_k, v_k \rangle = 0$ e quindi $\lambda_k = 0$ per ogni k siccome $v_k \neq 0$. $\odot$

Definizione 1.5 *Sia V uno spazio vettoriale con un prodotto scalare. Diremo che una successione di vettori $v_1, \dots v_n$ di V è* $\boxed{\text{ortonormale}}$ *se $\langle v_i, v_j \rangle = \delta_{ij}$. Se tale successione genera V ovvero $V = L\{v_1, \dots v_n\}$, diremo che V possiede una* $\boxed{\text{base ortonormale}}$.

Fissato un prodotto scalare si ottiene anche una $\boxed{\textit{distanza}}$ ponendo

$$d(u, v) \stackrel{\text{def}}{=} \|u - v\|$$

tale che

- $d(u, v) \geq 0$ e $d(u, v) = 0 \Leftrightarrow u = v$
- $d(u, v) = d(v, u)$
- $d(u, v) \leq d(u, w) + d(w, v)$ anche detta disuguaglianza triangolare.

Uno spazio vettoriale dotato di una distanza con le proprietà sopra elencate si dice $\boxed{\textit{metrico}}$. Analogamente uno spazio V si dice $\boxed{\textit{normato}}$ se dotato di norma nel senso che si ha un'applicazione

$$\| \ \| : V \to \mathbb{R}^+$$

non necessariamente indotta da un prodotto scalare, tale che $\|v\| = 0 \Leftrightarrow v = 0$, $\|\lambda v\| = |\lambda| \|v\|$ e $\|u + v\| \leq \|u\| + \|v\|$.

Abbiamo osservato che un prodotto scalare garantisce una norma e quindi una distanza ma, in generale, non viceversa.[3]

[3] Uno spazio metrico diventa uno spazio normato ponendo $\|u\| \stackrel{\text{def}}{=} d(u, 0)$ se e solo se la distanza soddisfa inoltre le seguenti $d(\lambda u, \lambda v) = |\lambda| d(u, v)$ e $d(u, v) = d(u + w, v + w)$. Una norma risulta indotta da un prodotto scalare ogni volta che soddisfa l'identità del parallelogramma sopra esposta.

Gli spazi euclidei $\mathbb{R}^n$

Vogliamo ora definire un prodotto scalare su $\mathbb{R}^n$ che generalizza il prodotto scalare tra vettori geometrici. Dato $x \in \mathbb{R}^n$ sia $x = (x_1, \dots, x_n)$ e sia $y \in \mathbb{R}^n$

$$y^T = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

Poniamo

$$\langle x, y \rangle \stackrel{\text{def}}{=} xy^T = x_1 y_1 + \cdots + x_n y_n$$

anche inteso come prodotto righe per colonne. Si ha che $\langle x, y \rangle = xy^T = yx^T = \langle y, x \rangle$. Inoltre $\langle x, \lambda y \rangle = \lambda xy^T = \lambda \langle x, y \rangle$ e $\langle x, y + y' \rangle = x(y + y')^T = x(y^T + y'^T) = xy^T + xy'^T = \langle x, y \rangle + \langle x, y' \rangle$. Infine $\langle x, x \rangle = xx^T = x_1^2 + \cdots + x_n^2$ da cui si ha

$$\| x \| \stackrel{\text{def}}{=} \sqrt{x_1^2 + \cdots + x_n^2}$$

e la distanza $\boxed{\textit{euclidea}}$ indotta. Si noti che la base canonica di $\mathbb{R}^n$ è ortonormale. Infine, si ha la seguente:

Teorema 1.6 (Formula di Aggiunzione) *Sia $A \in M_{m,n}(\mathbb{R})$. Allora per ogni $x \in \mathbb{R}^n$ ed $y \in \mathbb{R}^m$ (considerati come vettori colonna!)*

$$\langle Ax, y \rangle = \langle x, A^T y \rangle$$

Dimostrazione: Osserviamo che considerando i vettori colonna si ha $\langle a, b \rangle = a^T b$ e quindi

$$\langle Ax, y \rangle = (Ax)^T y = x^T A^T y = \langle x, A^T y \rangle.$$

$\odot$

Gli spazi di funzioni

Considerando funzioni integrabili è possibile definire, integrando, prodotti scalari tra queste. Ad esempio sia $\mathcal{C}^0([a, b])$ lo spazio delle funzioni reali continue su un intervallo chiuso $[a, b] \subset \mathbb{R}$. Definiamo un'operazione che risulta (verificando!) un prodotto scalare

$$\langle f, g \rangle \stackrel{\text{def}}{=} \int_a^b f(t)g(t)dt$$

mediante l'integrale definito. Si ottiene anche una norma

$$\| f\| \stackrel{\text{def}}{=} \Big(\int_a^b f(t)^2 dt \Big)^{1/2}$$

In particolare, lo spazio dei polinomi $\mathbb{R}[t] \subset \mathcal{C}^0([-1,1])$ si considera normato

$$\| p\| \stackrel{\text{def}}{=} \Big(\int_{-1}^1 p(t)^2 dt \Big)^{1/2}$$

e si ha

$$\langle t^r, t^s \rangle \stackrel{\text{def}}{=} \int_{-1}^1 t^{r+s} dt = \frac{1}{r+s+1} - \frac{(-1)^{r+s+1}}{r+s+1} = \begin{cases} \frac{2}{r+s+1} & \text{se r+s è pari} \\ 0 & \text{se r+s è dispari} \end{cases}$$

Ad esempio, $\| 1\| = \sqrt{2}$ e 1 è ortogonale ad t.

Decomposizioni ortogonali

Sia dunque V un $\mathbb{R}$-spazio vettoriale dotato di prodotto scalare $\langle\ ,\ \rangle$ e sia $W \subset V$ un sottospazio. Definiamo

$$W^\perp \stackrel{\text{def}}{=} \{v \in V / \langle v, w \rangle = 0\ \forall w \in W\}$$

che risulta ancora un sottospazio di V. Si noti che se $W = L\{v_1, \ldots, v_r\}$ si ha che $v \in W^\perp \Leftrightarrow \langle v, v_i \rangle = 0$ per ogni $i = 1, \ldots r$.

Abbiamo sempre che $W \cap W^\perp = 0$, infatti se $v \in W \cap W^\perp$ allora $\langle v, v \rangle = \| v\|^2 = 0 \Leftrightarrow v = 0$. Vogliamo inoltre mostrare che $W + W^\perp = V$ e quindi che si ha la $\boxed{\textit{decomposizione ortogonale}}$

$$W \oplus W^\perp = V$$

Se V è di tipo finito dal teorema di Grassman si ha che

$$\dim(W^\perp) = \dim(V) - \dim(W)$$

ed infine $(W^\perp)^\perp = W$.

Proiezioni e riflessioni

Supponiamo $W = L\{v_1\}$ con $\| v_1\| = 1$. Analogamente ai vettori geometrici definiamo la proiezione di un qualunque vettore $v \in V$ su W mediante la formula

$$P_W(v) \stackrel{\text{def}}{=} \langle v, v_1\rangle v_1$$

e quindi si ha che $v - P_W(v) \in W^\perp$. Infatti, $\langle v - P_W(v), v_1\rangle = \langle v, v_1\rangle - \langle v, v_1\rangle\| v_1\|^2$. In generale se $W = L\{v_1, \dots, v_r\}$ dove la successione $v_1, \dots, v_r$ costituisce una base ortonormale di W definiamo la $\boxed{\textit{proiezione ortogonale}}$ su W come

$$P_W(v) \stackrel{\text{def}}{=} \langle v, v_1\rangle v_1 + \cdots \langle v, v_r\rangle v_r$$

Analogamente abbiamo che $v - P_W(v) \in W^\perp$ siccome $\langle v_i, v_j\rangle = \delta_{ij}$. Dunque un qualunque vettore v si scrive come somma di $P_W(v) \in W$ e $v - P_W(v) \in W^\perp$ e si ha che $P_W(v) = 0 \Leftrightarrow v \in W^\perp$ e $P_W(v) = v \Leftrightarrow v \in W$. Definiamo, analogamente, la $\boxed{\textit{riflessione}}$ lungo W mediante

$$R_W(v) \stackrel{\text{def}}{=} 2P_W(v) - v$$

Algoritmo di ortonormalizzazione

Vediamo ora come costruire basi ortonormali. Questa costruzione, nota come algoritmo di Gram-Schmidt, non è altro che l'iterazione del procedimento di proiezione. Sia $W = L\{w_1, \dots, w_r\}$ dove adesso $w_1, \dots, w_r$ è una qualunque base di W. Vogliamo, trovare una base ortonormale di W ovvero $v_1, \dots, v_r$ che generano W e tali che $\langle v_i, v_j\rangle = \delta_{ij}$. Denotiamo $W_i \stackrel{\text{def}}{=} L\{w_1, \dots w_i\}$ per ogni $i = 1, \dots, r$ fissato e sia

$$\text{vers}(u) \stackrel{\text{def}}{=} \frac{u}{\| u\|}$$

per ogni $u \in V$. Poniamo

- $v_1 \stackrel{\text{def}}{=} \text{vers}(w_1)$
- $v_2 \stackrel{\text{def}}{=} \text{vers}(w_2 - P_{W_1}(w_2))$ da cui risulta $W_2 = L\{v_1, v_2\}$;
- $v_3 \stackrel{\text{def}}{=} \text{vers}(w_3 - P_{W_2}(w_3))$

  $\vdots$

- $v_{k+1} \overset{\text{def}}{=} \text{vers}(w_{k+1} - P_{W_k}(w_{k+1}))$ da cui risulta $W_{k+1} = L\{v_1, \dots, v_{k+1}\}$ ed il processo termina[4] quando $k + 1 = r$.

Si noti che la successione $v_1, \dots, v_r$ è ortonormale per costruzione e si ha

$$w_{k+1} - P_{W_k}(w_{k+1}) = w_{k+1} - \langle w_{k+1}, v_1\rangle v_1 - \cdots - \langle w_{k+1}, v_k\rangle v_k$$

dalla definizione di proiezione ortogonale.

In conclusione:

- se $W = V$ l'algoritmo permette di ricavare una base ortonormale da una qualunque base di V
- se $W \subset V$ l'algoritmo permette di ricavare una base ortonormale da una qualunque base di W e quindi scrivere ogni vettore $v \in V$ come somma della sua proiezione ortogonale $P_W(v) \in W$ e del vettore $v - P_W(v) \in W^\perp$ dimostrando che $V = W \oplus W^\perp$.

[4]Notare che il procedimento si può comunque applicare anche a successioni infinite di vettori.

2 Operatori lineari

Un $\boxed{\textit{operatore lineare}}$ tra due k-spazi vettoriali V e V' è un'applicazione $\varphi : V \to V'$ tale che comunque dati vettori $u, v \in V$ e scalari $\lambda, \mu \in k$ si ha

$$\varphi(\lambda u + \mu v) = \lambda\varphi(u) + \mu\varphi(v).$$

Ad esempio, sia $V = \mathbb{R}[X]$: l'applicazione che associa ad un polinomio la sua derivata di ordine n fissato, $\varphi(p) \stackrel{\text{def}}{=} p^{(n)}$, è un'importante paradigmatico operatore lineare, $\varphi : V \to V$. In questo caso però la spazio vettoriale V non ha dimensione finita e quindi il suo studio risulta maggiormente involuto. Affronteremo lo studio di tali operatori solo per spazi vettoriali di tipo finito.

Omomorfismi, Isomorfismi, Endomorfismi, Automorfismi

Gli operatori lineari vengono anche detti $\boxed{\textit{omomorfismi}}$ in altri testi applicazioni lineari, ed infine trasformazioni lineari. Un'operatore lineare $\varphi : V \to V'$ è invertibile o anche un $\boxed{\textit{isomorfismo}}$ se esiste $\varphi^{-1} : V' \to V$ lineare. Un'operatore lineare $\varphi : V \to V$ di uno spazio in sé stesso è anche detto $\boxed{\textit{endomorfismo}}$ se inoltre φ è invertibile viene detto $\boxed{\textit{automorfismo}}$.

Denotiamo $\mathrm{Hom}_k(V, V')$ l'insieme di tutti gli operatori lineari da V in V'. Questo e' ancora un k-spazio vettoriale con le ovvie operazioni di somma e prodotto per uno scalare definiti per gli spazi di funzioni. Denotiamo $\mathrm{End}_k(V) \stackrel{\text{def}}{=} \mathrm{Hom}_k(V, V)$. Dati $\varphi, \psi \in \mathrm{End}_k(V)$ la composizione $\varphi\psi$ è ancora un endomorfismo e munisce $\mathrm{End}_k(V)$ di un prodotto associativo, distributivo rispetto alla somma, con identità l'applicazione identica.

Matrici associate

Fissata una matrice $M \in M_{m,n}(k)$ consideriamo l'operazione

$$X \mapsto MX$$

che associa ad un vettore colonna con n-righe un vettore colonna con m-righe. Pensiamo che il vettore colonna X rappresenta coordinate di un vettore v di uno spazio vettoriale V rispetto ad una base B (fissata) e MX rappresenta le coordinate di un vettore di uno spazio V' rispetto ad un'altra base B' (anch'essa fissata): tale operatore da V in V' risulterà dipendere dalla scelta delle basi B e B'. Inversamente, vedremo che ogni operatore lineare da V in

V' si può descrivere mediante una matrice come sopra, avendo scelto delle basi B e B'.

Operatori da $\mathbb{R}^n$ in $\mathbb{R}^m$

In particolare, sia $A \in M_{m,n}(\mathbb{R})$ e $x \in \mathbb{R}^n$ una n-upla di numeri reali; indichiamo $X \overset{\text{def}}{=} x^T$ e sia $(AX)^T \overset{\text{def}}{=} y \in \mathbb{R}^m$. Si ottiene un'applicazione

$$T_A : \mathbb{R}^n \to \mathbb{R}^m \qquad\qquad x \mapsto T_A(x) \overset{\text{def}}{=} y$$

Osserviamo che:

- $A(X + X') = AX + AX'$ ovvero $T_A(x + x') = T_A(x) + T_A(x')$
- $A(\lambda X) = \lambda(AX)$ ovvero $T_A(\lambda x) = \lambda T_A(x)$ (per ogni $\lambda \in \mathbb{R}$)

e dunque si ha $T_A(\lambda x + \lambda' x') = \lambda T_A(x) + \lambda' T_A(x')$ ovvero che l'applicazione T_A è un operatore lineare. Inoltre, fissato un vettore colonna b, l'insieme delle soluzioni del sistema $S : AX = b$, $\mathrm{Sol}(S) = \{X/AX = b\} = T_A^{-1}(b)$ s'identifica con la controimmagine del vettore $b^T \in \mathbb{R}^m$ ed in particolare il sottospazio delle soluzioni del sistema omogeneo è $T_A^{-1}(0) = \{x \in \mathbb{R}^n/T_A(x) = 0\}$.

Sia $A = (a_{ij})$ la matrice considerata. Osserviamo che una n-upla $x \in \mathbb{R}^n$ ovvero $x = (x_1, \dots, x_n)$ viene trasformata mediante T_A in una m-upla

$$T_A(x) = (a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n, \dots, a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n)$$

semplicemente dalla definizione di prodotto righe per colonne. Quindi si ha che le colonne della matrice A sono i trasformati dei vettori $e_1, \dots, e_n$ della base canonica di $\mathbb{R}^n$ infatti $T_A(1, 0, \dots, 0) = (a_{11}, a_{21}, \dots, a_{m1})$, etc.

Se inversamente supponiamo di avere un'applicazione lineare $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ otteniamo una matrice M_φ associata a φ scrivendo i vettori $\varphi(e_1), \dots, \varphi(e_n)$ in colonna. Si noti che M_φ ha ancora m-righe ed n-colonne. Siccome $x = x_1e_1 + \cdots + x_ne_n$ e φ è lineare si ha che

$$\varphi(x) = x_1\varphi(e_1) + \cdots + x_n\varphi(e_n) = T_{M_\varphi}(x).$$

Denotiamo $\mathrm{Hom}\,(\mathbb{R}^n, \mathbb{R}^m)$ l'$\mathbb{R}$-spazio vettoriale di tutti gli operatori lineari tra $\mathbb{R}^n$ ed $\mathbb{R}^m$. Abbiamo dunque che

$$\varphi \mapsto M_\varphi : \mathrm{Hom}\,(\mathbb{R}^n, \mathbb{R}^m) \to M_{m,n}(\mathbb{R}) \tag{2}$$

è un'applicazione bigettiva con inversa

$$A \mapsto T_A : M_{m,n}(\mathbb{R}) \to \mathrm{Hom}\,(\mathbb{R}^n, \mathbb{R}^m).$$

Osserviamo che tali operazioni sono anche lineari, ad esempio $M_{\varphi+\psi} = M_\varphi + M_\psi$, e dunque definiscono un isomorfismo

$$\mathrm{Hom}\,(\mathbb{R}^n, \mathbb{R}^m) \stackrel{\cong}{\to} M_{m,n}(\mathbb{R})$$

Operatori associati ad una matrice

In generale, sia V un k-spazio vettoriale con una base $B : v_1, \ldots, v_n$ fissata e sia V' un'altro k-spazio vettoriale con un'altra base $B' : v'_1, \ldots, v'_m$ anch'essa fissata. Sia $A \in M_{m,n}(k)$ una matrice dove si noti che $n = \dim(V) =$ il numero di colonne ed $m = \dim(V') =$ il numero di righe di A. Vogliamo costruire un'applicazione lineare $T_A : V \to V'$.

Analogamente a quanto detto sopra siano $x = (x_1, \ldots, x_n)$ le coordinate di un vettore $v \in V$ ovvero il vettore $v = x_1 v_1 + \cdots + x_n v_n$ ed il vettore colonna $v_B \stackrel{\text{def}}{=} x^T$. Moltiplicando v_B per A otteniamo un vettore colonna Av_B che possiamo pensare come coordinate $y = (y_1, \ldots, y_m)$ del vettore trasformato $T_A(v) \stackrel{\text{def}}{=} y_1 v'_1 + \cdots + y_m v'_m$ ovvero $y^T \stackrel{\text{def}}{=} T_A(v)_{B'}$. In questo modo risulta $v \mapsto T_A(v) : V \to V'$ ottenuto da $v_B \mapsto Av_B$. In particolare, si ha che $T_A(v_i)_{B'}$ è la i-esima colonna della matrice A in quanto $v_{iB} = e_i^T$ e $Ae_i^T = i$-esima colonna della matrice A. Quindi

$$A = (T_A(v_1)_{B'} | \cdots | T_A(v_n)_{B'})$$

Osserviamo che tale operatore $T_A : V \to V'$ può cambiare con la scelta delle basi: infatti, cambiando la base B' di V' il vettore $T_A(v)$ cambia, inoltre, lo stesso vettore $v \in V$ ha coordinate diverse rispetto a basi B diverse di V, moltiplicando per A possiamo ottenere risultati diversi e quindi diversi vettori in V', anche rispetto alla stessa base B'.

Matrici associate ad un operatore

Se adesso consideriamo un operatore lineare $\varphi : V \to V'$ tra k-spazi vettoriali con basi $B : v_1, \ldots, v_n$ e $B' : v'_1, \ldots, v'_m$ fissate. Siano $\varphi(v_1)_{B'}, \ldots, \varphi(v_n)_{B'}$ i vettori colonna che rappresentano le coordinate rispetto a B' dei trasformati

$\varphi(v_1), \ldots, \varphi(v_n)$ dei vettori della base B. Questa è la matrice delle coordinate di tali vettori $\varphi(v_1), \ldots, \varphi(v_n)$ e si dice $\boxed{\textit{matrice associata}}$

$$M_\varphi^{BB'} \stackrel{\text{def}}{=} (\varphi(v_1)_{B'} | \cdots | \varphi(v_n)_{B'})$$

Ovviamente tale matrice cambia al variare di B e B'. Osserviamo che l'operatore associato a $M_\varphi^{BB'}$ è φ stesso ovvero $T_{M_\varphi^{BB'}} = \varphi$. Infatti, $T_{M_\varphi^{BB'}}(v_i)$ ha proprio le coordinate di $\varphi(v_i)$ rispetto a B' ed esiste un'unica applicazione lineare determinata[5] dai trasformati di una base.

Teorema 2.1 *Siano V e V' due k-spazi vettoriali con basi B e B' fissate. Sia* $\dim(V) = n$ *e* $\dim(V') = m$. *Esiste una corrispondenza biunivoca tra operatori lineari e matrici*

$$\varphi \mapsto M_\varphi^{BB'} : \mathrm{Hom}_k(V, V') \xrightarrow{\simeq} M_{m,n}(k)$$

tale che:

$$M_{\varphi+\psi}^{BB'} = M_\varphi^{BB'} + M_\psi^{BB'}$$

e per ogni $\lambda \in k$

$$M_{\lambda\varphi}^{BB'} = \lambda M_\varphi^{BB'}$$

Nucleo ed immagine

Sia $\varphi : V \to V'$ un operatore lineare. Definiamo il $\boxed{\textit{nucleo}}$ di φ

$$\mathrm{Ker}\,(\varphi) \stackrel{\text{def}}{=} \{v \in V / \varphi(v) = 0\}$$

Si vede facilmente che è un sottospazio di V. Inoltre si ha che $\varphi : V \to V'$ è iniettiva come applicazione se e solo se $\mathrm{Ker}\,(\varphi) = 0$. Infatti, $\varphi(u) = \varphi(v) \Leftrightarrow \varphi(u - v) = 0$, per $u, v \in V$.

Chiamiamo $\boxed{\textit{immagine}}$ di φ l'immagine di φ come applicazione

$$\mathrm{Im}\,(\varphi) \stackrel{\text{def}}{=} \{\varphi(v) / v \in V\}$$

è anch'essa un sottospazio. Se $v_1, \ldots, v_n$ generano V abbiamo che

$$\mathrm{Im}\,(\varphi) = L\{\varphi(v_1), \ldots, \varphi(v_n)\}$$

[5]Si estende per linearità!

Infatti, $V = L\{v_1, \ldots, v_n\}$ ed essendo φ lineare trasforma ogni combinazione lineare di $v_1, \ldots, v_n$ in una combinazione lineare di $\varphi(v_1), \ldots, \varphi(v_n)$.

Sia $v_1, \ldots, v_n$ una qualunque successione di vettori linearmente indipendenti che non appartengono al nucleo, ovvero $L\{v_1, \ldots, v_n\} \cap \mathrm{Ker}\,(\varphi) = 0$, allora $\varphi(v_1), \ldots, \varphi(v_n)$ sono linearmente indipendenti. Infatti, se $\lambda_1\varphi(v_1) + \cdots + \lambda_n\varphi(v_n) = 0$ allora $\varphi(\lambda_1 v_1 + \cdots + \lambda_n v_n) = 0$ quindi $\lambda_1 v_1 + \cdots + \lambda_n v_n \in \mathrm{Ker}\,(\varphi)$ ma $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$ per ipotesi, dunque $\lambda_1 = \cdots = \lambda_n = 0$ poichè $v_1, \ldots, v_n$ sono linearmente indipendenti.

Dunque, se $\mathrm{Ker}\,(\varphi) = 0$ e $B : v_1, \ldots, v_n$ è una base di V allora i trasformati $\varphi(v_1), \ldots, \varphi(v_n)$ costituiscono una base di $\mathrm{Im}\,(\varphi)$. In generale si ha:

Teorema 2.2 *Sia $\varphi : V \to V'$ un operatore lineare tar k-spazi vettoriali e supponiamo che V sia di tipo finito. Si ha*

$$\dim(V) = \dim(\mathrm{Ker}\,(\varphi)) + \dim(\mathrm{Im}\,(\varphi))$$

Dimostrazione: Sia $v_1, \ldots, v_k$ una base di $\mathrm{Ker}\,(\varphi)$ che completiamo a base $v_1, \ldots, v_k, v_{k+1}, \ldots, v_n$ di V. Siccome $V = L\{v_1, \ldots, v_n\}$ si ha che l'immagine $\mathrm{Im}\,(\varphi) = L\{\varphi(v_1), \ldots, \varphi(v_k), \varphi(v_{k+1}), \ldots, \varphi(v_n)\}$ dove $\varphi(v_1) = \cdots = \varphi(v_k) = 0$ per ipotesi. Dunque $\mathrm{Im}\,(\varphi) = L\{\varphi(v_{k+1}), \ldots, \varphi(v_n)\}$ e questi rimanenti vettori $\varphi(v_{k+1}), \ldots, \varphi(v_n)$ sono linearmente indipendenti. Infatti, $V = \mathrm{Ker}\,(\varphi) \oplus L\{v_{k+1}, \ldots, v_n\}$ per costruzione e dunque $L\{v_{k+1}, \ldots, v_n\} \cap \mathrm{Ker}\,(\varphi) = 0$ ed abbiamo già osservato che questo implica la lineare indipendenza dei trasformati $\varphi(v_{k+1}), \ldots, \varphi(v_n)$. Dunque $\dim(\mathrm{Im}\,(\varphi)) = n - k$ dove $\dim(\mathrm{Ker}\,(\varphi)) = k$ e $\dim(V) = n$. $\odot$

Se anche V' ha dimensione finita, siano B e B' basi di V e V', rispettivamente, e sia $M_\varphi^{BB'}$ una qualunque matrice associata a φ. Abbiamo sempre che

$$\rho(M_\varphi^{BB'}) = \dim(\mathrm{Im}\,(\varphi)) \tag{3}$$

infatti sulle colonne di $M_\varphi^{BB'}$ ci sono le coordinate rispetto a B' di vettori che generano $\mathrm{Im}\,(\varphi)$. Inoltre, il sistema omogeneo $M_\varphi^{BB'} X = 0$ ha come soluzioni le coordinate di vettori che generano $\mathrm{Ker}\,(\varphi)$. Coerentemente con il Teorema di Rouché-Capelli si ha

$$n - \rho(M_\varphi^{BB'}) = \dim(\mathrm{Ker}\,(\varphi)) \qquad n = \dim(V) \tag{4}$$

Le precedenti formule sono utili per calcolare esplicitamente la dimensione di nucleo ed immagine.

Sia $\varphi : V \to V'$ un operatore lineare, da quanto osservato segue che:

- se $\mathrm{Ker}\,(\varphi) = 0$ (ovvero φ iniettiva) allora $\dim(V) \leq \dim(V')$;
- se $\mathrm{Im}\,(\varphi) = V'$ (ovvero φ surgettiva) allora $\dim(V) \geq \dim(V')$;
- se $\mathrm{Ker}\,(\varphi) = 0$ e $\mathrm{Im}\,(\varphi) = V'$ (ovvero φ bigettiva) allora si ha necessariamente che $\dim(V) = \dim(V')$.

Infine, se $\dim(V) = \dim(V')$ allora si ha:

- φ bigettiva $\Leftrightarrow$ φ iniettiva $\Leftrightarrow$ φ surgettiva.

Dimensione di $\mathrm{Hom}\,(V, V')$

Consideriamo gli $\mathbb{R}$-spazi vettoriali $\mathrm{Hom}\,(\mathbb{R}^n, \mathbb{R}^m)$ e $M_{m,n}(\mathbb{R})$. Sia

$$\phi : \mathrm{Hom}\,(\mathbb{R}^n, \mathbb{R}^m) \to M_{m,n}(\mathbb{R})$$

l'operatore lineare che associa ad ogni $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ la matrice associata M_φ rispetto alle basi canoniche. Abbiamo visto che ϕ è un'applicazione lineare bigettiva dunque possiamo dedurre che

$$\dim(\mathrm{Hom}\,(\mathbb{R}^n, \mathbb{R}^m)) = \dim(M_{m,n}(\mathbb{R})) = mn$$

Analogamente per k-spazi vettoriali V e V' con dimensioni n ed m si ha che $\dim(\mathrm{Hom}\,(V, V')) = \dim(M_{m,n}(k)) = mn$.

Matrici di passaggio

Vogliamo ora capire come cambia la matrice associata al variare delle basi. Studiamo a questo fine la matrice associata alla composizione di due operatori $\varphi : V \to V'$ e $\psi : V' \to V''$. Innanzi tutto consideriamo il caso in cui $V = \mathbb{R}^n$, $V' = \mathbb{R}^m$ e $V'' = \mathbb{R}^p$, con matrici associate $A = M_\varphi$ e $B = M_\psi$ rispetto alle corrispondenti basi canoniche. L'applicazione $X \mapsto BAX$ associata alla matrice prodotto ovvero

$$T_{BA} : \mathbb{R}^n \to \mathbb{R}^p$$

risulta essere la composta $T_B T_A$. Inoltre

$$T_{M_\psi} T_{M_\varphi}(x) = T_{M_\psi}(\varphi(x)) = \psi(\varphi(x)) = \psi\varphi(x)$$

per ogni $x \in \mathbb{R}^n$, e dunque $M_{\psi\varphi} = M_\psi M_\varphi$. In generale si ha:

Teorema 2.3 *Siano $\varphi : V \to V'$ e $\psi : V' \to V''$ due operatori lineari tra k-spazi vettoriali di tipo finito. Siano $M_\varphi^{BB'}$ e $M_\psi^{B'B''}$ matrici associate rispetto a basi B, B' e B'' rispettivamente di V, V' e V''. Si ha che*

$$M_{\psi\varphi}^{BB''} = M_\psi^{B'B''} M_\varphi^{BB'}$$

è la matrice associata a $\psi\varphi : V \to V''$ rispetto alle suddette basi.

Dimostrazione: Basta osservare che

$$M_\psi^{B'B''} M_\varphi^{BB'} v_B = M_\psi^{B'B''} \varphi(v)_{B'} = \psi(\varphi(v))_{B''}$$

e dall'unicità della matrice associata (fissate le basi!) si ha l'uguaglianza. ⊙

Operatori invertibili

Consideriamo un isomorfismo $\varphi : V \to V'$. Osserviamo che, in particolare, nel precedente teorema, se $V'' = V$, $B = B''$ e $\psi = \varphi^{-1}$, allora $\dim(V) = \dim(V')$ e $M_\varphi^{BB'}$ è quadrata ed invertibile. Inoltre $\varphi^{-1}\varphi = id$ è l'applicazione identica di V e quindi

$$I = M_{id}^{BB} = M_{\varphi^{-1}\varphi}^{BB} = M_{\varphi^{-1}}^{B'B} M_\varphi^{BB'}$$

Vale anche il viceversa, e abbiamo che:

- $\varphi : V \to V'$ è un isomorfismo $\Leftrightarrow$ scelte basi B e B' la matrice $M_\varphi^{BB'}$ è invertibile e si ha
$$M_{\varphi^{-1}}^{B'B} = (M_\varphi^{BB'})^{-1}$$

Ricordiamo che l'applicazione $\varphi : V \to k^n$ con $n = \dim(V)$ e che associa $v \mapsto v_B$, ad un vettore le sue coordinate rispetto ad una base B di V, è un esempio importante d'isomorfismo. Se denotiamo con K la base canonica di k^n si vede che $M_\varphi^{BK} = I$, dunque anche $M_{\varphi^{-1}}^{KB} = I$.

Cambio base

Consideriamo ora l'applicazione identica $id : V \to V$ e la sua matrice associata M_{id}^{BC} rispetto a basi diverse B e C di V. Se $B \neq C$ questa matrice $M_{id}^{BC} \neq I$ non è l'identità. Infatti, ha sulle colonne le coordinate dei vettori di B rispetto a C. Siccome id è chiaramente invertibile con inversa sé stessa, si ha che

$$(M_{id}^{BC})^{-1} = M_{id}^{CB}$$

Chiamiamo queste matrici invertibili $\boxed{\textit{matrici di passaggio}}$ da B a C e viceversa.

Applicando quanto detto sopra, a basi B, C di V e B', C' di V', la matrice $M_{\varphi}^{BB'}$ associata ad un qualunque operatore lineare $\varphi : V \to V'$ risulta dalla composizione di

$$V \xrightarrow{M_{id}^{BC}} V \xrightarrow{M_{\varphi}^{CC'}} V' \xrightarrow{M_{id}^{C'B'}} V'$$

ovvero otteniamo la formula

$$M_{\varphi}^{BB'} = M_{id}^{C'B'} M_{\varphi}^{CC'} M_{id}^{BC}$$

Matrici simili

In particolare, consideriamo $\varphi : V \to V$ ovvero $\varphi \in \mathrm{End}_k(V)$. Siano B e C due basi diverse di V. Dalla precedente formula dove $V = V'$, $B = B'$ e $C = C'$ otteniamo

$$M_{\varphi}^{BB} = M_{id}^{CB} M_{\varphi}^{CC} M_{id}^{BC}$$

Da quanto già osservato $(M_{id}^{CB})^{-1} = M_{id}^{BC}$ e quindi abbiamo trovato la relazione che intercorre tra le matrici associate a basi diverse.

Data una matrice invertibile P questa si può sempre pensare come una matrice $M_{id}^{K'K}$ di passaggio dalla base canonica K di k^n ad un'altra base K' di k^n considerando le colonne di P come coordinate dei vettori della nuova base K' rispetto a K. Analogamente, data una base B, la matrice P determina una matrice M_{id}^{CB} dove la nuova base C si ricava dalle colonne di P (= coordinate dei vettori che definiscono C rispetto a B).

Definizione 2.4 *Diciamo che due matrici $A, A' \in M_{n,n}(k)$ sono* $\boxed{\text{simili}}$ *se sono associate allo stesso operatore ovvero se esiste una matrice invertibile P detta di* $\boxed{\text{passaggio}}$ *tale che*

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

e denotiamo con $A \simeq A'$ tale relazione.

Si noti che la relazione di similitudine è riflessiva, simmetrica e transitiva ovvero è una relazione di equivalenza. La simmetria segue dal fatto che se $A' = P^{-1}AP$ allora $A = PA'P^{-1}$ e ponendo $Q = P^{-1}$ si ha $A = Q^{-1}A'Q$. La transitività segue considerando il prodotto delle matrici di passaggio. Osserviamo che $A \simeq A'$ implica $\rho(A) = \rho(A')$ ed, inoltre, $\det(A) = \det(A')$.

Diagonalizzazione

Siamo ora volti a trovare la matrice associata più semplice possibile fissato un operatore $\varphi \in \mathrm{End}_k(V)$ al variare delle basi di V ovvero:

Problema 2.5 *Dato $\varphi \in \mathrm{End}_k(V)$ stabilire se o quando esiste una base B di V tale che M_φ^{BB} sia una matrice diagonale.*

Questo significa che per una matrice diagonale

$$\Delta = \begin{pmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

si ha $\Delta = M_\varphi^{BB}$ e $B : v_1, \dots, v_n$ è la base. Allora, dalla definizione di matrice associata si ha necessariamente

$$\varphi(v_1) = \lambda_1 v_1,\ \varphi(v_2) = \lambda_2 v_2,\ \dots\ ,\ \varphi(v_n) = \lambda_n v_n$$

dove gli scalari $\lambda_1, \dots, \lambda_n \in k$ possono essere anche in parte uguali o nulli.

Esempio

Sia $V = \mathbb{R}^2$ e sia $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ abbiamo che $\varphi = T_A : \mathbb{R}^2 \to \mathbb{R}^2$ è definito da $T_A(x, y) = A(x, y)^T = (y, 0)$. Cerchiamo una matrice Δ simile ad A ovvero cerchiamo una base $B : v_1, v_2$ di $\mathbb{R}^2$ e dei numeri reali $\lambda_1, \lambda_2 \in \mathbb{R}$ tali che $T_A(v_1) = \lambda_1 v_1,\ T_A(v_2) = \lambda_2 v_2$. Sia $v = (x, y)$ dunque

$$T_A(v) = (y, 0) = \lambda(x, y) \ \Leftrightarrow\ \lambda x = y \text{ e } \lambda y = 0.$$

Se $\lambda \neq 0$ si ha $v = (0,0)$. Se $\lambda = 0$ si ha $v = (x,0)$. Dunque per trovare $v_1, v_2 \in \{v = (x,y)/T_A(v) = \lambda v\}$ si deve avere $v_1, v_2 \in \operatorname{Ker}(T_A)$ ma allora v_1 e v_2 sono sempre linearmente dipendenti! Infatti, $\dim \operatorname{Ker}(T_A) = 1$. Quindi, non esiste nessuna base $B : v_1, v_2$ di $\mathbb{R}^2$ tale che M_φ^{BB} sia una matrice diagonale.

Autovalori ed autovettori

Diremo che $\lambda \in k$ è un $\boxed{\textit{autovalore}}$ per $\varphi : V \to V$ se esiste un vettore non nullo $v \in V$ tale che

$$\varphi(v) = \lambda v \qquad v \neq 0$$

ed in tal caso diciamo che v è un $\boxed{\textit{autovettore}}$ di φ associato a λ.

Definizione 2.6 *Diremo che un operatore lineare $\varphi : V \to V$ è* $\boxed{\text{semplice}}$ *o una sua matrice associata A è* $\boxed{\text{diagonalizzabile}}$ *se vale una delle seguenti condizioni equivalenti:*

- *A è simile ad una matrice diagonale Δ*
- *esiste una* base *di V costituita da autovettori per φ.*

Infatti, se esiste una base di V di autovettori $B : v_1, \ldots, v_n$ per φ la matrice M_φ^{BB} associata a tale base è già diagonale; se $A \simeq \Delta$ esiste una matrice P invertibile tale che $P^{-1}AP = \Delta$ sia diagonale e quindi P corrisponde al cambio base necessario per passare da una base qualunque di V ad una base di autovettori di V.

Autospazi

Sia $\varphi : V \to V$ e sia $\lambda : V \to V$ l'operatore $v \mapsto \lambda v$ ottenuto dalla moltiplicazione per uno scalare $\lambda \in k$ in V. L'insieme

$$V_\lambda \stackrel{\text{def}}{=} \{v \in V/\varphi(v) = \lambda v\}$$

è un sottospazio di V. Infatti, $\varphi(v) = \lambda v$ se e solo se $(\varphi - \lambda)(v) = 0$ dunque

$$V_\lambda = \operatorname{Ker}(\varphi - \lambda)$$

che viene detto $\boxed{\textit{autospazio}}$ relativo all'autovalore $\lambda \in V$. In effetti,

$$V_\lambda \neq 0 \Leftrightarrow \lambda \text{ è autovalore.}$$

Si ha che

$$\dim(V_\lambda) = \dim(\mathrm{Ker}\,(\varphi - \lambda)) = \dim(V) - \rho(M_\varphi^{BB} - \lambda I)$$

per una qualunque base B di V (infatti $M_{\varphi-\lambda}^{BB} = M_\varphi^{BB} - \lambda I$). Dunque

$$\lambda \text{ è autovalore} \Leftrightarrow \dim(V_\lambda) \geq 1 \Leftrightarrow \det(M_\varphi^{BB} - \lambda I) = 0.$$

Autospazi associati ad autovalori distinti

Diciamo che $\lambda_1, \ldots, \lambda_r$ sono $\boxed{\textit{autovalori distinti}}$ se $\lambda_i \neq \lambda_j$ per $i \neq j$. Osserviamo che se $\lambda_1, \ldots, \lambda_r$ sono autovalori distinti e $v_1, \ldots, v_r$ sono corrispondenti autovettori allora $v_1, \ldots, v_r$ sono linearmente indipendenti. Infatti, per induzione, supponiamo $\alpha_1 v_1 + \cdots + \alpha_r v_r = 0$ allora applicando φ si ha $\alpha_1\varphi(v_1) + \cdots + \alpha_r\varphi(v_r) = \alpha_1\lambda_1 v_1 + \cdots + \alpha_r\lambda_r v_r = 0$. Quindi

$$\alpha_1\lambda_1 v_1 + \cdots + \alpha_r\lambda_r v_r - \lambda_r(\alpha_1 v_1 + \cdots + \alpha_r v_r) = 0$$

da cui si ottiene

$$\alpha_1(\lambda_1 - \lambda_r)v_1 + \cdots + \alpha_{r-1}(\lambda_{r-1} - \lambda_r)v_{r-1} = 0$$

e per ipotesi induttiva $\alpha_i(\lambda_i - \lambda_r) = 0$ per $i = 1, \ldots, r-1$. Siccome $\lambda_i \neq \lambda_r$ si ha che $\alpha_i = 0$ per $i = 1, \ldots, r-1$, dunque $\alpha_r v_r = 0$ ma $v_r \neq 0$ dunque $\alpha_r = 0$. Si ha dunque che se abbiamo $\lambda_1, \ldots, \lambda_r$ autovalori distinti l'unione di basi degli autospazi corrispondenti $V_{\lambda_1}, \ldots, V_{\lambda_r}$ è base dello spazio somma $V_{\lambda_1} + \cdots + V_{\lambda_r}$, che risulta una somma diretta. Dunque, esiste una base di V fatta da autovettori quando $V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_r}$ ovvero

$$\varphi \text{ è semplice} \Leftrightarrow \dim(V_{\lambda_1}) + \cdots + \dim(V_{\lambda_r}) = \dim(V).$$

In particolare, se φ ha un numero di autovalori distinti pari alla dimensione di V allora $\dim(V_{\lambda_i}) = 1$ per $i = 1, \ldots, r$ con $r = \dim(V)$. Ovvero

- $\lambda_1, \ldots, \lambda_n$ autovalori distinti con $n = \dim(V) \Rightarrow \varphi$ è semplice.

Questa condizione è sufficiente ma non necessaria! Infatti, la moltiplicazione per uno scalare $\lambda : V \to V$ ha un'unico autovalore ed un unico autospazio $V_\lambda = V$ e quindi tutti i vettori (non nulli) sono autovettori, una qualunque base è una base di autovettori.

Polinomio caratteristico

Da quanto detto sopra, il problema di stabilire se φ è semplice ovvero se una sua matrice associata A è diagonalizzabile si riconduce al calcolo degl'autovalori. A questo fine definiamo il $\boxed{\textit{polinomio caratteristico}}$ come segue

$$P_A(t) \stackrel{\text{def}}{=} \det(A - tI) \qquad P_A(t) \in k[t]$$

avendo già osservato che λ è autovalore $\Leftrightarrow P_A(\lambda) = 0$. Osserviamo che se $A \simeq A'$ ovvero $A' = P^{-1}AP$ allora

$$P_{A'}(t) = \det(A' - tI) = \det(P^{-1}AP - tP^{-1}P) = \det(P^{-1}(A - tI)P) = P_A(t)$$

Definizione 2.7 *Sia $\varphi \in \mathrm{End}_k(V)$ definiamo $P_\varphi(t) \stackrel{\text{def}}{=} P_A(t)$ per una* qualunque *matrice A associata a φ.*

Se φ fosse semplice: $A \simeq \Delta$ e la matrice diagonale Δ ha polinomio caratteristico

$$P_\Delta(t) = (\lambda_1 - t)^{m_1} \cdots (\lambda_r - t)^{m_r}$$

dove $\lambda_1, \ldots, \lambda_r \in k$ sono gli autovalori ed $m_1, \ldots, m_r$ le corrispondenti molteplicità come radici di $P_\Delta(t) = 0$. Inoltre, chiaramente

$$m_i = \dim(V_{\lambda_i}) = \dim(V) - \rho(\Delta - \lambda_i I)$$

per $i = 1, \ldots, r$ ed $m_1 + \cdots + m_r =$ grado di $P_\Delta(t) = \dim(V)$. In conclusione, abbiamo trovato condizioni necessarie per φ semplice ovvero

- $P_\varphi(t) = 0$ deve avere tutte le radici $\lambda_1, \ldots, \lambda_r$ in k e
- $m_i = \dim(V_{\lambda_i})$ per ciascun $i = 1, \ldots, r$ dove m_i è la molteplicità di λ_i come radice di $P_\varphi(t) = 0$.

Ma le condizioni trovate sono anche sufficienti! Infatti, il seguente risultato risponde al Problema 2.5 da cui siamo partiti.

Teorema 2.8 (Criterio di diagonalizzabilità) *Sia $\varphi \in \mathrm{End}_k(V)$ e sia $P_\varphi(t)$ il suo polinomio caratteristico. Siano $\lambda_1, \ldots, \lambda_r$ le radici di $P_\varphi(t) = 0$ e siano $m_1, \ldots, m_r$ le corrispondenti molteplicità. Allora φ è semplice* se e solo se *$\lambda_1, \ldots, \lambda_r \in k$ ed $m_1 = \dim(V_{\lambda_1}), \ldots, m_r = \dim(V_{\lambda_r})$.*

Dimostrazione: Basta osservare che $\dim(V) = n =$ grado di $P_\varphi(t)$ e che si ha $m_1 + \cdots + m_r = n$. Se φ è semplice allora esiste una base B di autovettori tale che $M_\varphi^{BB} = \Delta$ è diagonale e $P_\varphi(t) = P_\Delta(t)$ soddisfa le condizioni. Viceversa, abbiamo visto che $\dim(V_{\lambda_1}) + \ldots + \dim(V_{\lambda_r}) \leq n$ e vale l'eguaglianza $\Leftrightarrow$ φ è semplice. Siccome $\dim(V_{\lambda_i}) = m_i$ per ogni $i = 1, \ldots, r$ si ha la tesi. $\odot$

Ricordiamo che se φ è un operatore lineare qualunque, anche se $k = \mathbb{C}$, nel qual caso $P_\varphi(t) = 0$ ha sempre soluzioni, la condizione che $m_i = \dim(V_{\lambda_i})$ può non essere verificata. In generale abbiamo:

Lemma 2.9 *Sia $\varphi \in \mathrm{End}_k(V)$. Sia $\lambda_i \in k$ autovalore ed m_i la sua molteplicità come radice di $P_\varphi(t) = 0$. Sia V_{λ_i} il corripondente autospazio allora*

$$1 \leq \dim(V_{\lambda_i}) \leq m_i$$

Dimostrazione: Infatti, sia $v_1, \ldots, v_r$ una base di V_{λ_i} che completiamo a base B di V mediante vettori $v_{r+1}, \ldots, v_n$. La matrice associata $A = M_\varphi^{BB}$ allora risulta

$$A = \begin{pmatrix} \lambda_i & 0 & 0 & \star & \ldots & \star \\ 0 & \lambda_i & 0 & \star & \ldots & \star \\ \ldots & \ldots & \ldots & \star & \ldots & \star \\ \ldots & \ldots & \lambda_i & \star & \ldots & \star \\ \ldots & \ldots & \ldots & \star & \ldots & \star \\ 0 & 0 & 0 & \star & \ldots & \star \end{pmatrix}$$

dove le prime r-colonne sono le coordinate dei trasformati degl'autovettori $v_1, \ldots, v_r$ e le restanti $n - r$-colonne denotano coordinate dei trasformati dei successivi $v_{r+1}, \ldots, v_n$ vettori di B. Il polinomio caratteristico è

$$P_\varphi(t) = P_A(t) = (\lambda_i - t)^r Q(t)$$

dove $Q(t)$ è un polinomio che potrebbe aver ancora λ_i come radice. Dunque $r \leq m_i$ come affermato. $\odot$

Un argomento analogo al precedente permette di ricavare la $\boxed{\textit{forma triangolare}}$ delle matrici come segue:

Teorema 2.10 *Sia $A \in M_{n,n}(k)$ una matrice tale che $P_A(t) = 0$ ha tutte le soluzioni in k, ad esempio $k = \mathbb{C}$. Allora A è simile ad una matrice triangolare superiore.*

Dimostrazione: Sia $T_A = \varphi \in \operatorname{End}_k(k^n)$ ed A la sua matrice associata rispetto alla base canonica. Per induzione sulla dimensione n supponiamo vera la tesi per $n-1$ con $n \geq 2$ (per $n = 1$ è banale). Sia v un autovettore con autovalore $\lambda \in k$ ovvero $\varphi(v) = \lambda v$ e sia B la base che si ottiene completando v ad una base di k^n. La matrice associata $A' = M_\varphi^{BB}$ allora risulta

$$A' = \left(\begin{array}{c|ccc} \lambda & \star & \dots & \star \\ \hline 0 & \star & \dots & \star \\ \vdots & \star & A_\lambda & \star \\ 0 & \star & \dots & \star \end{array}\right)$$

dove A_λ è la sottomatrice $n-1 \times n-1$ complementare dell'entrata λ e A' è simile ad A. Siccome $P_{A_\lambda}(t) = P_A(t)/(\lambda - t)$ quindi A_λ è simile ad una triangolare superiore per ipotesi induttiva. Dunque esistono P e Q tali che $A' = P^{-1}AP$ ed $U = Q^{-1}A_\lambda Q$ sia triangolare superiore. La matrice

$$P' = \left(\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & Q & \\ 0 & & & \end{array}\right)$$

è dunque tale che $(P')^{-1}A'P' = (P')^{-1}P^{-1}APP'$ sia la seguente matrice triangolare superiore

$$\left(\begin{array}{c|ccc} \lambda & \star & \dots & \star \\ \hline 0 & & & \\ \vdots & & U & \\ 0 & & & \end{array}\right)$$

$\odot$

Osserviamo che questa matrice triangolare è la matrice più semplice possibile che si può sempre associare ad un operatore lineare complesso qualunque ovvero se $k = \mathbb{C}$. In generale, gli operatori razionali o reali non possono esser rappresentati da matrici triangolari in quanto il polinomio caratteristico ha ben volentieri poche radici in $k = \mathbb{Q}$ o $\mathbb{R}$ (ad esempio!?).

Operatori autoaggiunti

Abbiamo stabilito un criterio per decidere se la matrice associata A ad un operatore $\varphi \in \mathrm{End}_k(V)$ possa essere diagonalizzabile ovvero per l'esistenza di una matrice di passaggio P tale che $P^{-1}AP$ sia diagonale.

D'altra parte, se $k = \mathbb{R}$ e V è anche munito di un prodotto scalare $\langle \ , \ \rangle$ abbiamo visto che, mediante l'algoritmo di Gram-Schmidt, possiamo sempre trovare basi ortonormali di V.

Sia $V = \mathbb{R}^n$ lo spazio euclideo con prodotto scalare ordinario $\langle \ , \ \rangle$ e sia $T_A = \varphi : V \to V$ un operatore lineare semplice dove $A \in M_{n,n}(\mathbb{R})$ è la matrice associata rispetto alla base canonica. Desideriamo ora stabilire quando esiste una base ortonormale B di V costituita da autovettori per φ. Infatti, il seguente esempio mostra che, anche se esistono basi di autovettori, queste basi possono non essere ortonormali.

Esempio

Sia $V = \mathbb{R}^2$. Dati numeri reali λ_1, λ_2 tali che $\lambda_1 \neq \lambda_2$ prendiamo una base v_1, v_2 di V costituita da vettori non ortogonali ovvero $\langle v_1, v_2\rangle \neq 0$. L'operatore $\varphi : V \to V$ semplice definito da $\varphi(v_1) = \lambda_1 v_1$ e $\varphi(v_2) = \lambda_2 v_2$ è tale che $V_{\lambda_1} = L\{v_1\}$ e $V_{\lambda_2} = L\{v_2\}$ ma $V_{\lambda_1}^{\perp} \neq V_{\lambda_2}$.

Matrici ortogonali

Supponiamo che esista una base ortonormale di autovettori. La matrice di passaggio, P tale che $P^{-1}AP = \Delta$ sia diagonale, è la matrice delle coordinate (rispetto alla base canonica K di $\mathbb{R}^n$) di una base B di $\mathbb{R}^n$ costituita da tali autovettori per $T_A : \mathbb{R}^n \to \mathbb{R}^n$. Ovvero $P = M_{id}^{BK}$ e la composta di

$$\mathbb{R}^n \xrightarrow{P} \mathbb{R}^n \xrightarrow{A} \mathbb{R}^n \xrightarrow{P^{-1}} \mathbb{R}^n$$

è $T_\Delta : \mathbb{R}^n \to \mathbb{R}^n$. La matrice $P^{-1} = M_{id}^{KB}$ si calcola facilmente in questo caso $P^{-1} = P^T$. Infatti sia $B : v_1, \ldots, v_n$ tale base, quindi P ha sulle colonne i vettori $v_1, \ldots, v_n$ mentre P^T ha sulle righe i vettori $v_1, \ldots, v_n$ dunque

$$P^T P = (\langle v_i, v_j\rangle) = (\delta_{ij}) = I$$

In generale, diremo che:

- una matrice P è una $\boxed{\textit{matrice ortogonale}}$ $\Leftrightarrow P^{-1} = P^T$

- una matrice A è $\boxed{\textit{ortogonalmente simile}}$ ad una matrice A' se esiste una matrice ortogonale P tale che $A' = P^T AP$ (dove $P^T = P^{-1}$).

Siccome il prodotto di matrici ortogonali è ortogonale si vede che questa è una relazione di equivalenza.

Se A è ortogonalmente simile ad una matrice diagonale $\Delta = P^T AP$ la matrice P rappresenta automaticamente una base ortonormale di autovettori per $T_A : \mathbb{R}^n \to \mathbb{R}^n$.

Le matrici P ortogonali corrispondono a cambi base ortonormali di $\mathbb{R}^n$ ed inoltre si ha che $\det(P) = \pm 1$. Le matrici ortogonali con determinante 1 sono anche dette di rotazione (perché!?).

Matrici simmetriche

Cerchiamo una condizione necessaria affinché $A \in M_{n,n}(\mathbb{R})$ sia ortogonalmente simile ad una matrice diagonale. Osserviamo che se A è ortogonalmente simile ad una matrice A' si ha che $A' = P^T AP \Leftrightarrow A = PA'P^T$ e quindi A è simmetrica $\Leftrightarrow$ A' simmetrica! Infatti, se $A^T = A$ allora

$$(A')^T = (P^T AP)^T = P^T A^T (P^T)^T = A'$$

ed analogamente per A'. Se $\Delta = P^T AP$ con P ortogonale e Δ diagonale, quindi simmetrica, allora anche A è simmetrica. Ecco la condizione necessaria cercata

- A ortogonalmente simile ad una matrice diagonale $\Rightarrow$ A simmetrica.

Equivalentemente: consideriamo l'operatore $\boxed{\textit{aggiunto}}$ $T_A^* : \mathbb{R}^n \to \mathbb{R}^n$ definito dalla matrice trasposta ovvero $T_A^* \stackrel{\text{def}}{=} T_{A^T}$. Ricordiamo che dal Teorema 1.6 ovvero dalla formula di aggiunzione abbiamo che

$$\langle T_A(x), y\rangle = \langle x, T_A^*(y)\rangle$$

per ogni $x, y \in \mathbb{R}^n$. Abbiamo che l'operatore $T_A : \mathbb{R}^n \to \mathbb{R}^n$ coincide con $T_A^* : \mathbb{R}^n \to \mathbb{R}^n$ $\Leftrightarrow$ $A = A^T$.

Definizione 2.11 *Diciamo che $T_A : \mathbb{R}^n \to \mathbb{R}^n$ è* $\boxed{\text{autoaggiunto}}$ *se $T_A^* = T_A$ ovvero se $A = A^T$.*

Dunque, se $\mathbb{R}^n$ possiede una base ortonormale di autovettori per $T_A : \mathbb{R}^n \to \mathbb{R}^n$ allora T_A è autoaggiunto. Supponiamo dunque A simmetrica e studiamo l'operatore autoaggiunto $T_A : \mathbb{R}^n \to \mathbb{R}^n$ associato.

Teorema spettrale

Sia $V = \mathbb{R}^n$ lo spazio euclideo. Sia $\varphi = T_A : \mathbb{R}^n \to \mathbb{R}^n$ un operatore lineare autoaggiunto. Siano $\lambda_i, \lambda_j \in \mathbb{R}$ due autovalori distinti $\lambda_i \neq \lambda_j$ con autovettori $v_i \in V_{\lambda_i}$ e $v_j \in V_{\lambda_j}$. Siccome $T_A = T_A^*$ dalla formula di aggiunzione $\langle T_A(v_i), v_j\rangle = \langle v_i, T_A(v_j)\rangle$ segue che $\langle \lambda_i v_i, v_j\rangle = \langle v_i, \lambda_j v_j\rangle$ e quindi si ha $\lambda_i\langle v_i, v_j\rangle = \lambda_j\langle v_i, v_j\rangle$ ovvero $(\lambda_i - \lambda_j)\langle v_i, v_j\rangle = 0$ da cui si ottiene $\langle v_i, v_j\rangle = 0$. Abbiamo dunque che

- due autospazi $V_{\lambda_i} \neq V_{\lambda_j}$ associati ad autovalori distinti $\lambda_i \neq \lambda_j$ sono ortogonali $V_{\lambda_i} \perp V_{\lambda_j}$.

Sia $m_i =$ la molteplicità di $\lambda_i \in \mathbb{R}$ come radice dell'equazione caratteristica $P_A(t) = 0$. Applichiamo il ragionamento visto nel Lemma 2.9 per dimostrare che $\dim(V_{\lambda_i}) = m_i$. Sia $\dim(V_{\lambda_i}) = r$ e sia $v_1, \ldots, v_r$ una base ortonormale di V_{λ_i} che completiamo a base ortonormale B di $\mathbb{R}^n$ mediante vettori $v_{r+1}, \ldots, v_n$ applicando l'algoritmo di Gram-Schmidt. La matrice associata $A' = M_\varphi^{BB}$ è ortogonalmente simile ad A e quindi anch'essa simmetrica. Allora risulta

$$A' = \left(\begin{array}{ccc|ccc} \lambda_i & \ldots & 0 & 0 & \ldots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \ldots & \lambda_i & 0 & \ldots & 0 \\ \hline 0 & \ldots & 0 & \star & \ldots & \star \\ \vdots & & \vdots & \vdots & A_\star & \vdots \\ 0 & \ldots & 0 & \star & \ldots & \star \end{array}\right)$$

infatti le prime r-colonne sono le coordinate dei trasformati degl'autovettori $v_1, \ldots, v_r$ e per simmetria annullano le prime r righe nell'entrate diverse dalla diagonale. Il polinomio caratteristico è $P_A(t) = P_{A'}(t) = (\lambda_i - t)^r Q(t)$ dove $Q(t) = P_{A_\star}(t)$ è un polinomio che non può aver ancora λ_i come radice. Infatti, in questo caso ci sarebbe un'altro autovettore con autovalore λ_i linearmente indipendente da $v_1, \ldots, v_r$ contro l'ipotesi. Quindi $r = m_i$ ovvero abbiamo dimostrato che

- $\dim(V_{\lambda_i}) = m_i =$ la molteplicità di $\lambda_i \in \mathbb{R}$ come radice di $P_A(t) = 0$.

Per completare lo studio di un operatore lineare autoaggiunto $T_A : \mathbb{R}^n \to \mathbb{R}^n$ siamo dunque ricondotti allo studio dei suoi autovalori ovvero del suo $\boxed{\textit{spettro}}$

$$\mathrm{Sp}\,(A) \stackrel{\text{def}}{=} \{\lambda \in \mathbb{C} / P_A(\lambda) = 0\}.$$

Teorema 2.12 (Spettrale) *Ogni matrice simmetrica $A \in M_{n,n}(\mathbb{R})$ è ortogonalmente simile ad una matrice diagonale ovvero il corrispondente operatore lineare autoaggiunto $T_A : \mathbb{R}^n \to \mathbb{R}^n$ è semplice.*

Dimostrazione: Consideriamo $A \in M_{n,n}(\mathbb{C})$ come matrice di numeri complessi e quindi $T_A : \mathbb{C}^n \to \mathbb{C}^n$ il corrispondente operatore. Sia $\lambda \in \mathrm{Sp}(A)$ ovvero $\lambda \in \mathbb{C}$ una radice di $P_A(\lambda) = 0$ e sia $z = (z_1, \dots, z_n) \in \mathbb{C}^n$ un autovettore. Dunque se $X = z^T$ si ha

$$AX = \lambda X$$

e prendendo i complessi coniugati[6] si ottiene

$$A\,\overline{X} = \overline{\lambda}\,\overline{X}$$

in quanto A è reale, avendo posto $\overline{X} = \overline{z}^T$. Il prodotto di matrici $\overline{X}^T AX \in \mathbb{C}$ risulta uno scalare che si calcola in due modi:

- $\overline{X}^T AX = \overline{X}^T (AX) = \overline{X}^T \lambda X = \lambda\,\overline{X}^T X$
- $\overline{X}^T AX = (\overline{X}^T A)X = (A\overline{X})^T X = \overline{\lambda}\,\overline{X}^T X$

poiché A è simmetrica. Quindi $(\lambda - \overline{\lambda})\,\overline{X}^T X = 0$ e osservando che $\overline{X}^T X = \overline{z}z^T = \overline{z_1}z_1 + \dots + \overline{z_n}z_n = |z_1|^2 + \dots + |z_n|^2 > 0$ in quanto $z \neq 0$ si conclude che $\lambda = \overline{\lambda}$ ovvero $\lambda \in \mathbb{R}$. Quindi A ha spettro reale $\mathrm{Sp}(A) \subset \mathbb{R}$ ovvero gli autovalori son tutti reali. Per quanto già osservato, la tesi segue dal criterio di diagonalizzabilità, Teorema 2.8. $\odot$

Caso complesso

Osserviamo che il precedente argomento suggerisce che si possa definire un'operazione $\langle\ ,\ \rangle : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ ponendo

$$\langle x, y\rangle \stackrel{\mathrm{def}}{=} x_1\overline{y_1} + \dots + x_n\overline{y_n}$$

tale che $\langle z, z\rangle \in \mathbb{R}$ e $\langle z, z\rangle \geq 0$, dunque una norma $\|z\| \stackrel{\mathrm{def}}{=} \sqrt{\langle z, z\rangle}$. Tale operazione (detta prodotto hermitiano) non è però simmetrica ma si ha

$$\langle x, y\rangle = \overline{\langle y, x\rangle}$$

[6] Il coniugio $z \mapsto \overline{z} : \mathbb{C}^n \to \mathbb{C}^n$ è un automorfismo $\mathbb{R}$-lineare di $\mathbb{C}^n$.

detta simmetria hermitiana. In questo caso, data una matrice $A = (a_{ij}) \in M_{n,n}(\mathbb{C})$ si considera la trasposta hermitiana $A^H = (\overline{a}_{ji})$. L'analogo complesso di una matrice simmetrica reale è una matrice hermitiana ovvero tale che $A = A^H$ (notare che una matrice diagonale è hermitiana $\Leftrightarrow$ è reale!). L'analogo complesso delle matrici ortogonali sono le matrici unitarie U tali che $U^{-1} = U^H$.

Si dimostra analogamente una formula di aggiunzione $\langle Ax, y\rangle = \langle x, A^H y\rangle$ dove $x, y \in \mathbb{C}^n$ sono intesi come vettori colonna. Se $A = A^H$ e $\lambda \in \mathrm{Sp}\,(A)$ con autovettore z allora

$$\lambda\|z\|^2 = \langle Az, z\rangle = \langle z, A^H z\rangle = \langle z, \lambda z\rangle = \overline{\lambda}\|z\|^2$$

e dunque nuovamente $\lambda \in \mathbb{R}$. L'operatore autoaggiunto $T_A : \mathbb{C}^n \to \mathbb{C}^n$ associato ad una matrice hermitiana $A = A^H$ è semplice ed esiste una base ortonormale di $\mathbb{C}^n$ costituita da autovettori per T_A. Tale base costituisce le colonne di una matrice unitaria U tale che $U^H AU = \Delta$ sia diagonale reale.

Infine, il Teorema spettrale reale è naturalmente deducibile dal Teorema spettrale complesso, infatti, il prodotto hermitiano ristretto ad $\mathbb{R}^n$ è il prodotto scalare ordinario.

Proiettori e riflettori

Sia $V \subseteq \mathbb{R}^n$ un sottospazio. Abbiamo definito la proiezione ortogonale di un vettore $x \in \mathbb{R}^n$ mediante una base ortonormale $v_1, \dots, v_r$ di V come

$$P_V(x) = \langle x, v_1\rangle v_1 + \cdots \langle x, v_r\rangle v_r$$

Completando a base ortonormale $B : v_1, \dots, v_r, v_{r+1}, \dots, v_n$ di $\mathbb{R}^n$ abbiamo che $V = L\{v_1, \dots, v_r\}$ e $V^\perp = L\{v_{r+1}, \dots, v_n\}$ e possiamo considerare l'operatore lineare $P_V : \mathbb{R}^n \to \mathbb{R}^n$ che ha autovalori 0 ed 1 ed autospazi $V_0 = \mathrm{Ker}\,(P_V) = V^\perp$ e $V_1 = \mathrm{Im}\,(P_V) = V$. La matrice diagonale $\Delta = M_{P_V}^{BB}$ associata alla base B di autovettori è

$$\Delta = \begin{pmatrix} 1 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & & \vdots \\ 0 & \dots & 1 & \dots & 0 \\ \vdots & & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 0 \end{pmatrix}$$

Dunque P_V è autoaggiunto, $P_V^2 = P_V$ e queste proprietà determinano P_V univocamente ovvero è l'unico operatore con autospazi $V_0 = V^\perp$ e $V_1 = V$ detto $\boxed{\textit{proiettore}}$ di $\mathbb{R}^n$ associato al sottospazio V.

Desideriamo ora calcolare la matrice A associata rispetto alla base canonica. Tale matrice simmetrica risulta da

$$\mathbb{R}^n \xrightarrow{P^T} \mathbb{R}^n \xrightarrow{\Delta} \mathbb{R}^n \xrightarrow{P} \mathbb{R}^n$$

ovvero dal prodotto $P\Delta P^T$ dove $P = M_{id}^{BK}$ è la matrice delle coordinate della base B rispetto alla base canonica di $\mathbb{R}^n$. Consideriamo la sottomatrice M di P delle prime r-colonne ovvero la matrice $M \in M_{n,r}(\mathbb{R})$ delle coordinate dei vettori $v_1, \ldots, v_r$ rispetto alla base canonica. Abbiamo che

$$A = P\Delta P^T = MM^T$$

in quanto la sottomatrice $r \times r$ di Δ individuata dalle prime r-righe è l'identica e le ultime $n - r$ righe e colonne di Δ sono nulle. Evidentemente $M^T M = I$ in quanto $v_1, \ldots, v_r$ sono ortonormali.

Analogamente si dice $\boxed{\textit{riflettore}}$ di $\mathbb{R}^n$ associato al sottospazio V l'operatore indotto dalla riflessione

$$R_V(x) = 2P_V(x) - x$$

con matrice associata rispetto alla base canonica $2MM^T - I$ che è simmetrica e ortogonale. Abbiamo che R_V è autoaggiunto, $R_V^2 = id$ e queste proprietà determinano R_V univocamente ovvero è l'unico operatore con autospazi $V_{-1} = V^\perp$ e $V_1 = V$.

Decomposizione spettrale

Sia $A \in M_{n,n}(\mathbb{R})$ una matrice simmetrica e sia $T_A : \mathbb{R}^n \to \mathbb{R}^n$ l'operatore autoaggiunto associato. Il Teorema spettrale afferma che T_A ha autovalori reali siano $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$ tutti distinti, eventualmente con molteplicità pari alla dimensione del corrispondente autospazio. Si ha una decomposizione

$$\mathbb{R}^n = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_r}$$

tale che l'operatore T_A ristretto ad i singoli autospazi agisce come la moltiplicazione $\lambda_i : V_{\lambda_i} \to V_{\lambda_i}$. Se scriviamo $x \in \mathbb{R}^n$ come somma di autovettori $v_i \in V_{\lambda_i}$, sia $x = v_1 + \cdots + v_r$ si ottiene

$$T_A(x) = \lambda_1 v_1 + \cdots + \lambda_r v_r$$

Consideriamo i proiettori associati ad i singoli autospazi

$$P_{V_{\lambda_i}} : \mathbb{R}^n \to \mathbb{R}^n$$

e siano P_i le corrispondenti matrici associate rispetto alla base canonica. Si ha che $\lambda_i P_{V_{\lambda_i}}$ ristretto a V_{λ_i} è proprio $\lambda_i : V_{\lambda_i} \to V_{\lambda_i}$ e $V_{\lambda_i}^{\perp}$ è il suo nucleo (ovvero la somma diretta dei restanti autospazi). In conclusione otteniamo che $P_{V_{\lambda_i}}(x) = v_i$ e si ha

$$T_A(x) = \lambda_1 P_{V_{\lambda_1}}(x) + \cdots + \lambda_r P_{V_{\lambda_r}}(x)$$

da cui segue la $\boxed{\textit{decomposizione spettrale}}$ della matrice simmetrica

$$A = \lambda_1 P_1 + \cdots + \lambda_r P_r$$

rispetto alle matrici $P_1, \ldots, P_r$ associate ai proiettori. Osserviamo, infine che per i corrispondenti riflettori $R_i = 2P_i - I$ e si ha inoltre che

$$2A - \kappa I = \lambda_1 R_1 + \cdots + \lambda_r R_r$$

dove $\kappa \stackrel{\text{def}}{=} \lambda_1 + \cdots + \lambda_r$. Quest'ultima costante $\kappa \in \mathbb{R}$ è legata alla $\boxed{\textit{traccia}}$ della matrice $A = (a_{ij})$ ovvero $\mathrm{Tr}\,(A) \stackrel{\text{def}}{=} a_{11} + \cdots + a_{nn}$. Infatti, si ha sempre[7] che

$$P_A(t) = (-1)^n t^n + \mathrm{Tr}\,(A)(-1)^{n-1} t^{n-1} + \cdots + \det(A)$$

e quindi la traccia non cambia nella diagonalizzazione ovvero

$$\mathrm{Tr}\,(A) = m_1 \lambda_1 + \cdots + m_r \lambda_r$$

dove $m_1, \ldots, m_r$ sono le molteplictà degl'autovalori come radici di $P_A(t) = 0$. In particolare $\kappa = \mathrm{Tr}\,(A)$ se $r = n$.

[7] Anche se A non è simmetrica!?

A Forme quadratiche

Una $\boxed{\textit{forma quadratica}}$ è una funzione $Q : \mathbb{R}^n \to \mathbb{R}$ definita da un polinomio omogeneo di secondo grado. Desideriamo studiare tali forme quadratiche, innanzi tutto, decidere quando sono positive, negative o nulle, suggerendo come trattare l'equazioni definite da queste.

Ad esempio, abbiamo che in due variabili le forme quadratiche $Q : \mathbb{R}^2 \to \mathbb{R}$ si scrivono

$$Q(x, y) = ax^2 + bxy + cy^2$$

rispetto a coordinate (x, y) per la base canonica di $\mathbb{R}^2$ e costanti $a, b, c \in \mathbb{R}$. Supponiamo, per esempio, $a \neq 0$ e completiamo i quadrati

$$Q(x, y) = a\Big(x^2 + \frac{b}{a}xy + \frac{c}{a}y^2\Big) = a\Big(x + \frac{b}{2a}y\Big)^2 + \frac{4ac - b^2}{4a}\, y^2$$

Mediante le nuove coordinate

$$\begin{cases} X = x + \frac{b}{2a}y \\ Y = y \end{cases}$$

si ottiene che

$$Q(x, y) = Q(X, Y) = \alpha X^2 + \beta Y^2$$

dove

$$\begin{cases} \alpha = a \\ \beta = \frac{4ac-b^2}{4a} \end{cases}$$

Considerando la matrice simmetrica

$$A = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}$$

Si vede che $\det(A)/a = \beta$. Se $a > 0$ si ha:

- se $\det(A) > 0$ allora $Q(x, y) > 0$ per ogni $(x, y) \in \mathbb{R}^2$
- se $\det(A) = 0$ allora $Q(x, y) = aX^2 \geq 0$ e si annulla per qualche $(x, y) \neq (0, 0) \in \mathbb{R}^2$
- se $\det(A) < 0$ allora $Q(x, y) > 0$ ed anche $Q(x, y) < 0$ per qualche $(x, y) \in \mathbb{R}^2$.

Analogamente se $a < 0$ possiamo decidere il segno di $Q(x, y)$ dal segno di $\det(A)$, nelle nuove coordinate (X, Y) è facile studiare $Q(x, y)$. In questo modo però, se $b \neq 0$, il cambio di coordinate

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} 1 & b/2a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

non è svolto mediante una matrice ortogonale. Quanto osservato permette di concludere a ritroso che la forma quadratica

$$Q(X, Y) = \begin{pmatrix} X & Y \end{pmatrix} \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} = \alpha X^2 + \beta Y^2$$

si può anche riscrivere

$$Q(x, y) = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 1 & 0 \\ b/2a & 1 \end{pmatrix} \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} \begin{pmatrix} 1 & b/2a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} =$$

$$= \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x & y \end{pmatrix} A \begin{pmatrix} x \\ y \end{pmatrix}$$

ma in questo caso A non è simile alla matrice diagonale $\Delta = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}$.

A.1 Forme bilineari

Data una matrice $A \in M_{n,n}(k)$ otteniamo sempre una $\boxed{\textit{forma bilineare}}$ ovvero k-lineare in entrambe le variabili

$$(x, x') \mapsto x^T A x' : k^n \times k^n \to k$$

dove x, x' sono vettori colonna ad n-entrate in k. Se $x = u_B$ e $x' = v_B$ sono coordinate di vettori di un k-spazio vettoriale V rispetto ad una sua base B fissata, mediante l'isomorfismo $v \mapsto v_B : V \xrightarrow{\cong} k^n$ otteniamo una forma bilineare[8] $\langle\ ,\ \rangle : V \times V \to k$ ovvero

$$\langle u, v \rangle \stackrel{\text{def}}{=} u_B^T A v_B : V \times V \to k$$

[8]Utilizziamo la stessa notazione del prodotto scalare per analogia ma non si ottiene necessariamente un prodotto scalare!

Viceversa, data una forma bilineare $\langle\ ,\ \rangle : V \times V \to k$, scelta una base $B : v_1, \ldots, v_n$ di V ponendo

$$\langle v_i, v_j \rangle \stackrel{\text{def}}{=} a_{ij}$$

otteniamo una matrice $A = (a_{ij})$ detta matrice associata alla forma bilineare e si ha

$$\langle u, v \rangle = u_B^T A v_B$$

Osserviamo che A è simmetrica se e solo se la forma bilineare è simmetrica ovvero $\langle u, v \rangle = \langle v, u \rangle$ per ogni $u, v \in V$.

A.1.1 Matrici congruenti

Se cambiamo base, la matrice cambia nel modo seguente. Siano B e C basi di V tali che A ed A' siano le matrici associate alla forma bilineare $\langle\ ,\ \rangle : V \times V \to k$ rispetto alla base B e C rispettivamente. Allora si ha che esiste una matrice invertibile P di passaggio tale che $Pu_C = u_B$, ovvero $u_B^T = u_C^T P^T$, e inoltre $Pv_C = v_B$ quindi

$$\langle u, v \rangle = u_B^T A v_B = u_C^T P^T A P v_C = u_C^T A' v_C$$

ovvero

$$A' = P^T A P$$

in quanto entrambe queste matrici sono associate alla forma bilineare rispetto alla base C e quindi coincidono. Diciamo che due matrici A ed A' sono $\boxed{\textit{congruenti}}$ se esiste una matrice invertibile P tale che $A' = P^T A P$. Si verifica facilmente che matrici congruenti a matrici simmetriche sono simmetriche e che la relazione di congruenza è una relazione di equivalenza.

Abbiamo che A ed A' ortogonalmente simili $\Rightarrow$ A ed A' simili e congruenti ma non viceversa. Infine, matrici congruenti non hanno, in generale, lo stesso polinomio caratteristico.

A.1.2 Dualità

La relazione di congruenza preserva la caratteristica ovvero $\rho(A) = \rho(A')$ poiché P è invertibile. Si può definire dunque il rango della forma bilineare e si dice che una forma bilineare è $\boxed{\textit{non degenere}}$ se ha rango massimo. Si definisce lo spazio vettoriale $\boxed{\textit{duale}}$

$$V^\vee \stackrel{\text{def}}{=} \mathrm{Hom}_k(V, k)$$

e si hanno due applicazioni lineari

$$\delta : u \mapsto \langle u, - \rangle : V \to V^\vee \qquad \delta' : v \mapsto \langle -, v \rangle : V \to V^\vee$$

tali che $\delta = \delta'$ se e solo se la forma bilineare è simmetrica. Si può vedere che δ é un isomorfismo $\Leftrightarrow$ δ' é un isomorfismo $\Leftrightarrow$ $\langle\ ,\ \rangle$ è non degenere.

A.2 Diagonalizzazione di forme quadratiche

Un polinomio omogeneo $Q(x_1, \ldots, x_n)$ di secondo grado nelle variabili $x_1, \ldots, x_n$ si scrive

$$Q(x_1, \ldots, x_n) = \sum_{i,j=1}^{n} a_{ij} x_i x_j$$

Se $i \neq j$ il termine a_{ij} compare 2 volte: supponendo $a_{ij} = a_{ji}$ il monomio corrispondente risulterebbe $2a_{ij}x_ix_j$. Possiamo sempre formare una matrice simmetrica $A = (a_{ij})$ tale che

$$Q(x_1, \ldots, x_n) = \begin{pmatrix} x_1 \ldots x_n \end{pmatrix} A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Data una matrice $A \in M_{n,n}(k)$ ed un k-spazio vettoriale V otteniamo sempre una $\boxed{\textit{forma quadratica su } V}$ definita mediante una sua base B ponendo

$$v \mapsto v_B^T A v_B : V \to k$$

Osserviamo che matrici congruenti danno origine alla stessa forma quadratica. Sia $A' = P^T AP$ e sia B' un'altra base di V determinata da P. Allora si ha che $v_B = Pv_{B'}$ e $v_B^T = v_{B'}^T P^T$ quindi

$$v_B^T A v_B = v_{B'}^T P^T A P v_{B'} = v_{B'}^T A' v_{B'}$$

Inoltre se A non è simmetrica la matrice $1/2(A + A^T)$ è simmetrica e produce la stessa forma quadratica.

A.2.1 Forma quadratica associata

Sia $\langle\ ,\ \rangle : V \times V \to V$ una forma bilineare simmetrica. Ponendo

$$Q(v) \overset{\text{def}}{=} \langle v, v \rangle$$

si ottiene una forma quadratica su V detta forma quadratica associata alla forma bilineare. Si ha

- $Q(\lambda v) = \lambda^2 Q(v)$
- $2\langle u, v\rangle = Q(u+v) - Q(u) - Q(v)$.

Infatti,

$$Q(u+v) - Q(u) - Q(v) = \langle u+v, u+v\rangle - \langle u, u\rangle - \langle v, v\rangle = 2\langle u, v\rangle$$

dalla bilinearità di $\langle\ ,\ \rangle$. Dunque una forma quadratica individua univocamente la forma bilineare simmetrica cui è associata. Altrimenti detto ogni matrice simmetrica congruente a quella definita dal polinomio $Q(x_1, \ldots, x_n)$, una volta scelta una base di V, individua simultaneamente Q e $\langle\ ,\ \rangle$.

A.2.2 Forme quadratiche reali

Supponiamo quindi $V = \mathbb{R}^n$ e sia $Q : \mathbb{R}^n \to \mathbb{R}$ una forma quadratica reale. Dunque $Q : \mathbb{R}^n \to \mathbb{R}$ s'interpreta mediante una matrice simmetrica associata A rispetto alla base canonica $Q(x) = x^T Ax$ e questa è la forma quadratica associata alla forma bilineare $\langle x, x'\rangle = x^T Ax'$.

Possiamo allora cambiare base di $\mathbb{R}^n$ in modo che A sia ortogonalmente simile ad una matrice diagonale, per il Teorema spettrale. Otteniamo una matrice ortogonale P tale che $P^T AP = \Delta$ e quindi A è simile e congruente a Δ. Si ha

$$\begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} = P^T \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \qquad \Leftrightarrow \qquad P \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

in quanto $P^T = P^{-1}$ e trasponendo

$$\begin{pmatrix} x_1 \ldots x_n \end{pmatrix} = \begin{pmatrix} X_1 \ldots X_n \end{pmatrix} P^T \qquad \Leftrightarrow \qquad \begin{pmatrix} x_1 \ldots x_n \end{pmatrix} P = \begin{pmatrix} X_1 \ldots X_n \end{pmatrix}$$

dunque sostituendo

$$Q(x_1, \ldots, x_n) = \begin{pmatrix} x_1 \ldots x_n \end{pmatrix} A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} X_1 \ldots X_n \end{pmatrix} P^T AP \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} =$$

$$= \begin{pmatrix} X_1 \dots X_n \end{pmatrix} \Delta \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} = \lambda_1 X_1^2 + \dots + \lambda_n X_2^2$$

dove $\lambda_1, \dots, \lambda_n$ sono gli autovalori di A eventualmente ripetuti.

A.2.3 Diagonalizzazione in generale

Sia V un k-spazio vettoriale $\dim(V) = n$. Sia $\langle\ ,\ \rangle : V \times V \to V$ una forma bilineare simmetrica. Sia $A \in M_{n,n}(k)$ una matrice simmetrica che rappresenta $\langle\ ,\ \rangle$ o la corrispondente forma quadratica, rispetto ad una base. Non è difficile dimostrare che A è sempre congruente ad una matrice diagonale ma tale argomento esula dal contesto qui trattato.

Osserviamo comunque che, mediante la riduzione di Gauss, possiamo, ad esempio, ridurre A per righe fino ad una matrice triangolare inferiore QA dove Q è prodotto di matrici elementari e quindi invertibile. Se nella riduzione non si operano scambi di righe o moltiplicazioni di righe per scalari, siccome A è simmetrica, le stesse operazioni per colonne permettono di ridurre QA ad una matrice diagonale QAQ^T. Questo significa che esiste $P = Q^T$ invertibile tale che $\Delta = P^T AP$ sia diagonale.

Quanto detto permette inoltre di ricavare la $\boxed{\textit{forma di Sylvester}}$ di una matrice simmetrica reale che tiene conto solo dei segni dei termini sulla diagonale di Δ. Ricordiamo però che in questo caso $P_\Delta(t) \neq P_A(t)$ in generale.

A.3 Classificazione delle forme quadratiche

Data una forma quadratica $Q : \mathbb{R}^n \to \mathbb{R}$ diciamo che

- Q è $\boxed{\textit{definita positiva}}$ se $Q(x) > 0$ per ogni $x \neq 0$
- Q è $\boxed{\textit{semidefinita positiva}}$ se $Q(x) \geq 0$ per ogni $x \neq 0$ ed esiste $x \neq 0$ tale che $Q(x) = 0$
- Q è $\boxed{\textit{definita negativa}}$ se $Q(x) < 0$ per ogni $x \neq 0$
- Q è $\boxed{\textit{semidefinita negativa}}$ se $Q(x) \leq 0$ per ogni $x \neq 0$ ed esiste $x \neq 0$ tale che $Q(x) = 0$
- Q è $\boxed{\textit{indefinita}}$ se $Q(x) > 0$ e $Q(x') < 0$ per qualche $x, x' \neq 0$.

Analoga terminologia per una matrice simmetrica $A \in M_{n,n}(\mathbb{R})$ corrispondente a Q osservando che tali proprietà sono dunque preservate per congruenza. Inoltre, dal Teorema spettrale segue che la positività o negatività delle matrici dipende solo dal segno degl'autovalori.

A.3.1 Prodotto scalare

Un prodotto scalare è una forma bilineare, simmetrica e definita positiva. Una matrice simmetrica A definita positiva definisce anche una forma bilineare simmetrica $\langle\ ,\ \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ tale che $Q(x) = \langle x, x\rangle > 0$ per $x \neq 0$ e quindi $\langle\ ,\ \rangle$ risulta un prodotto scalare. In particolare, per $A = I$ si ottiene il prodotto scalare ordinario in $\mathbb{R}^n$ rispetto alla base canonica ovvero $Q(x) = x^T x$. Altrimenti, per l'algoritmo di ortonormalizzazione di Gram-Schmidt, esiste una base ortonormale $B : v_1, \ldots, v_n$ di $\mathbb{R}^n$ rispetto a questo prodotto scalare ($v_1, \ldots, v_n$ potrebbero però non essere ortogonali rispetto al prodotto scalare ordinario!). Quindi, la matrice associata rispetto a questa base B risulta $(\langle v_i, v_j\rangle) = (\delta_{ij}) = I$ ed è congruente ad A ovvero esiste una matrice invertibile P tale che $A = P^T I P$. Sono fatti equivalenti:

- $\langle\ ,\ \rangle$ è un prodotto scalare su $\mathbb{R}^n$
- Q è definita positiva
- $A = P^T P$ per una matrice invertibile P

Quest'ultima affermazione significa infatti che A è congruente all'identità ed in particolare A è definita positiva.

A.3.2 Esempio: coniche e quadriche

Ritornando al punto di partenza sia $Q(x, y) = ax^2 + bxy + cy^2$ rispetto a coordinate (x, y) per la base canonica di $\mathbb{R}^2$ e costanti $a, b, c \in \mathbb{R}$. La matrice simmetrica associata

$$A = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}$$

ha polinomio caratteristico $P_A(t) = t^2 - (a + c)t + \det(A)$ ed ha sempre radici reali gli autovalori λ_1 e λ_2. Esiste P ortogonale tale che $P^T AP = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \stackrel{\text{def}}{=} \Delta$. Il cambio di coordinate $\begin{pmatrix} x \\ y \end{pmatrix} = P \begin{pmatrix} X \\ Y \end{pmatrix}$ permette di scrivere $Q(x,y) = \begin{pmatrix} x & y \end{pmatrix} A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} X & Y \end{pmatrix} \Delta \begin{pmatrix} X \\ Y \end{pmatrix} = \lambda_1 X^2 + \lambda_2 Y^2$. Si ha che

$$\begin{cases} \lambda_1 + \lambda_2 = a + c \\ \lambda_1 \lambda_2 = ac - b^2/4 = \det(A) \end{cases}$$

La positività di $Q(x,y)$ è quindi determinata da $\lambda_1, \lambda_2 > 0 \iff \det(A) > 0$ ed $a > 0$. Uno tra λ_1, λ_2 è nullo $\iff \det(A) = 0$. Uno tra λ_1, λ_2 è negativo $\iff \det(A) < 0$. Infine la matrice ortogonale P ha come colonne autovettori ortonormali (x_1, y_1) e (x_2, y_2) e può sempre essere scelta di rotazione ovvero

$$P = \begin{pmatrix} x_1 = \cos(\theta) & x_2 = -\sin(\theta) \\ y_1 = \sin(\theta) & y_2 = \cos(\theta) \end{pmatrix}$$

eventualmente cambiando segno ad un autovettore.

Sia $f(x,y) = Q(x,y) + L(x,y) + \Gamma$ dove $L(x,y) = \ell_1 x + \ell_2 y$ è lineare con $\ell_1, \ell_2 \in \mathbb{R}$ e $\Gamma \in \mathbb{R}$ è una costante. Il luogo geometrico definito da $f(x,y) = 0$ è detto $\boxed{\textit{conica}}$ nel piano. Mediante il cambiamento di coordinate suddetto si ottiene che $f(X,Y) = Q(X,Y) + L(X,Y) + \Gamma$ dove ora la forma quadratica è diagonale. Se $f(X,Y) = 0$ non degenera ad una coppia di rette o un punto, ed entrambi λ_1, λ_2 sono non nulli ovvero $\det(A) \neq 0$, completando i quadrati[9] si ottiene un'ellisse[10] se $\det(A) > 0$ o un'iperbole se $\det(A) < 0$, altrimenti, per $\det(A) = 0$ si ha una parabola.

Infine, se $f(x,y,z) = Q(x,y,z) + L(x,y,z) + \Gamma$ dove $L(x,y,z)$ è lineare e $\Gamma \in \mathbb{R}$ analogamente otteniamo una forma diagonale $Q(X,Y,Z) = \lambda_1 X^2 + \lambda_2 Y^2 + \lambda_3 Z^3$ mediante un cambiamento di coordinate ortogonale. Il luogo dello spazio definito da $f(x,y,z) = 0$ è detto $\boxed{\textit{quadrica}}$ e si classifica mediante gli autovalori. Le quadriche non degeneri a piani, coni o cilindri sono: ellissoide, iperboloide ad una o due falde, paraboloide ellittico ed iperbolico, detto sella.

[9] Il completamento dei quadrati corrisponde ad una traslazione di $\mathbb{R}^2$ ovvero ad un ulteriore cambiamento di coordinate $\chi = X - X_0$ ed $\Upsilon = Y - Y_0$ che porta il centro delle nuove coordinate nel punto (X_0, Y_0).

[10] L'ellisse potrebbe anche avere solo punti immaginari come $X^2 + Y^2 = -1$.