1. Applicazioni lineari simmetriche.

Consideriamo lo spazio $\mathbb{R}^n$ col prodotto scalare canonico $X \cdot Y = {}^tXY = x_1y_1 + \ldots + x_ny_n$.

Definizione. Un'applicazione lineare $L_A\colon \mathbb{R}^n \to \mathbb{R}^n$, $\quad X \mapsto L_A(X) = AX$ si dice *simmetrica* se

$$L_A(X) \cdot Y = X \cdot L_A(X) \quad \text{ossia} \quad AX \cdot Y = X \cdot AY \qquad \text{per ogni } \ X, Y \in \mathbb{R}^n. \qquad (*)$$

Osserviamo che $AX \cdot Y = {}^tX\,{}^tAY$ e $X \cdot AY = {}^tX\,AY$, da cui segue che l'applicazione L_A è simmetrica se e solo se la matrice A è simmetrica, cioè soddisfa ${}^tA = A$.

Esercizio 1.1. Sia $L_A\colon \mathbb{R}^n \to \mathbb{R}^n$, $\quad X \mapsto AX$ un'applicazione simmetrica. Sia U un sottospazio di $\mathbb{R}^n$ e sia $U^\perp$ il suo complemento ortogonale. Se $L_A(U) \subset U$, allora anche $L_A(U^\perp) \subset U^\perp$.

Dim. Sia w un arbitrario elemento di $U^\perp$. Per far vedere che $L_A(w)$ appartiene ancora a $U^\perp$, verifichiamo che $L_A(w) \cdot u = 0$ per ogni $u \in U$. Poiché per la (*) $L_A(w) \cdot u = w \cdot L_A(u)$ e per ipotesi $w \in U^\perp$ e $L_A(u) \in U$, segue che $L_A(w) \cdot u = 0$. In particolare $L_A(w) \in U^\perp$ come richiesto.

Per le matrici simmetriche reali vale il seguente teorema di diagonalizzazione.

Teorema spettrale. *Sia A una matrice simmetrica reale $n \times n$.*

(i) Il polinomio caratteristico di A ha n radici reali (contate con la loro molteplicità); in altre parole, A ha n autovalori reali.

(ii) Sia λ un autovalore di A di molteplicità algebrica k e sia V_λ l'autospazio corrispondente. Allora $\dim V_\lambda = k$.

(ii) Autospazi relativi ad autovalori distinti sono ortogonali.

Dim. (i) La dimostrazione di questo fatto sarà ottenuta come caso particolare del teorema spettrale per matrici hermitiane. Nel frattempo lo verifichiamo nel caso di una matrice simmetrica $A = \begin{pmatrix} a & b \\ b & d \end{pmatrix}$ di ordine 2: il polinomio caratteristico di A risulta $P_\lambda(A) = \lambda^2 - (a+c)\lambda + (ac - b^2)$. Poiché il discriminante del polinomio $\Delta = (a-c)^2 + b^2$ è non negativo, le sue radici sono necessariamente reali.

(ii) Sia $L_A\colon \mathbb{R}^n \to \mathbb{R}^n$, $X \mapsto AX$. Siano $\lambda_1, \ldots, \lambda_k$ gli autovalori di L_A e siano $V_{\lambda_1}, \ldots, V_{\lambda_k}$ gli autospazi corrispondenti. Consideriamo il seguente sottospazio di $\mathbb{R}^n$

$$U = V_{\lambda_1} \oplus \ldots \oplus V_{\lambda_k} = \{X \in \mathbb{R}^n \mid \exists \lambda \in \mathbb{R} : \ AX = \lambda X\}.$$

È chiaro dalla definizione di U che $L_A(U) \subset U$. Dall'Esercizio 1.1 segue che $L_A(U^\perp) \subset U^\perp$; pertanto la restrizione ad $U^\perp$ di L_A definisce un'applicazione lineare *simmetrica* $L_A|_{U^\perp}\colon U^\perp \to U^\perp$. Notare che il prodotto scalare ristretto a $U^\perp$ rimane definito positivo. Per il punto (i), questa applicazione ha autovalori reali, per cui esistono $\sigma \in \mathbb{R}$ e un vettore $w \in U^\perp$ tali che $L_A(w) = \sigma w$. Questo contraddice la definizione di U, e implica $U^\perp = \{0\}$. In particolare, $\mathbb{R}^n = V_{\lambda_1} \oplus \ldots \oplus V_{\lambda_k}$ e

ogni autospazio di A ha dimensione massima, uguale alla molteplicità algebrica dell'autovalore corrispondente.
(iii) Siano λ e μ autovalori distinti di A e siano V_λ e V_μ i rispettivi autospazi. Facciamo vedere che elementi arbitrari $v \in V_\lambda$ e $w \in V_\mu$ sono ortogonali fra loro. Dalla (*) si ha

$$Av \cdot w = (\lambda v) \cdot w = \lambda(v \cdot w) = v \cdot Aw = v \cdot (\mu w) = \mu(v \cdot w),$$

da cui segue che $(\lambda - \mu)\ (v \cdot w) = 0$. Poiché $\lambda \neq \mu$, necessariamente vale $v \cdot w = 0$, cioè $v \perp w$ come richiesto.

Direttamente da fatti (i)(ii)(iii) segue che

(iv) Esiste una base ortonormale di $\mathbb{R}^n$ formata da autovettori di A.

(v) La matrice A è diagonalizzabile mediante una matrice ortogonale, ossia esiste una matrice ortogonale M tale che

$$M^{-1}AM = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 & 0 \\ 0 & \lambda_2 & \dots & 0 & 0 \\ 0 & 0 & \dots & \dots & 0 \\ 0 & 0 & \dots & \dots & \lambda_n \end{pmatrix}, \tag{3}$$

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

Se

$$\{\begin{pmatrix} v_{11} \\ \vdots \\ v_{n1} \end{pmatrix}, \dots, \begin{pmatrix} v_{1n} \\ \vdots \\ v_{nn} \end{pmatrix}\}$$

è una qualunque base ortonormale di $\mathbb{R}^n$ formata da autovettori di A (di autovalori $\lambda_1, \dots, \lambda_n$ rispettivamente), la matrice ortogonale

$$M = \begin{pmatrix} v_{11} & \dots & v_{1n} \\ \vdots & \vdots & \vdots \\ v_{n1} & \dots & v_{nn} \end{pmatrix}$$

soddisfa la relazione (3).

Osservazione. *Una matrice reale A è diagonalizzabile mediante una matrice ortogonale se e solo se è simmetrica.*

Dim. Una direzione è data dal punto (v) del teorema precedente. Supponiamo adesso che esista una matrice ortogonale M tale che $M^{-1}AM = D$, con D matrice diagonale. Poiché $M^{-1} = {}^tM$, questo equivale a dire che $A = MDM^{-1} = M\,D\,{}^tM$. Calcolando la trasposta di A, troviamo

$${}^tA = {}^t(M\,D\,{}^tM) = {}^t({}^tM)\,{}^tD\,{}^tM = M\,D\,{}^tM = A.$$

Dunque A è simmetrica, come richiesto.

Definizione. Due matrici A e B si dicono *congruenti* se esiste una matrice invertibile N tale che $B = {}^tN\,A\,N$.

Si verifica facilmente che la relazione di congruenza è una relazione di equivalenza fra matrici. In particolare se A è congruente a B e B è congruente a C, allora anche A è congruente a C. Come conseguenza del Teorema di Sylvester qui di seguito risulta che *due matrici simmetriche A e B sono congruenti se e solo se hanno lo stesso numero di autovalori positivi, negativi e nulli.*

Teorema (Sylvester). *Sia A una matrice simmetrica. Allora esiste una matrice invertibile N tale che*

$$
{}^{t}N\,A\,N = \begin{pmatrix} I_p & O & O \\ O & -I_q & O \\ O & O & O_r \end{pmatrix},
$$

dove p è uguale al numero di autovalori positivi di A, q è uguale al numero di autovalori negativi di A, ed r è uguale al numero di autovalori nulli di A.

Dim. Dal teorema spettrale segue che esiste una matrice ortogonale M tale che

$$
M^{-1}\,A\,M = {}^{t}M\,A\,M = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 & 0 \\ 0 & \lambda_2 & \dots & 0 & 0 \\ 0 & 0 & \dots & \dots & 0 \\ 0 & 0 & \dots & \dots & \lambda_n \end{pmatrix},
$$

dove $\lambda_1, \dots, \lambda_n$ sono gli autovalori di A. Supponiamo adesso che $\lambda_1, \dots, \lambda_p$ siano positivi, $\lambda_{p+1}, \dots, \lambda_{p+q}$ siano negativi e $\lambda_{p+q+1}, \dots, \lambda_{p+q+r}$ siano nulli (con $p+q+r = n$). Consideriamo adesso la matrice

$$
H = \begin{pmatrix} \frac{1}{|\lambda_1|^{1/2}} & 0 & \dots & 0 & 0 \\ 0 & \frac{1}{|\lambda_2|^{1/2}} & \dots & 0 & 0 \\ 0 & 0 & \dots & \dots & 0 \\ 0 & 0 & \dots & \dots & I_r \end{pmatrix}
$$

È immediato verificare che

$$
{}^{t}H\,{}^{t}M\,A\,M H = {}^{t}(MH)\,A\,M H = \begin{pmatrix} I_p & O & O \\ O & -I_q & O \\ O & O & O_r \end{pmatrix}.
$$

Dunque la matrice $N = MH$ è la matrice invertibile cercata.

2. Forme quadratiche reali.

Una forma quadratica reale in n variabili $x_1, \dots, x_n$ è un polinomio omogeneo di secondo grado a coefficienti reali

$$F\colon \mathbb{R}^n \longrightarrow \mathbb{R}, \quad F(x_1, \dots, x_n) = a_{11}x_1 + \ldots + a_{nn}x_n^2 + \sum_{1 \le i < j \le n} 2a_{ij}x_i x_j, \quad a_{ij} \in \mathbb{R}.$$

Esempi di forme quadratiche:

(a) $F\colon \mathbb{R} \to \mathbb{R}, \quad F(x_1) = 4x_1^2$;

(b) $F\colon \mathbb{R}^2 \to \mathbb{R}, \quad F(x_1, x_2) = 2x_1^2 - x_1x_2$;

(c) $F\colon \mathbb{R}^2 \to \mathbb{R}, \quad F(x_1, x_2, x_3) = x_1^2 - x_1x_2 + 3x_2x_3 + 3x_3^3$.

(d) Sia $X \cdot Y$ il prodotto scalare canonico in $\mathbb{R}^n$. Il quadrato della norma di un vettore $X = (x_1, \dots, x_n) \in \mathbb{R}^n$ è una forma quadratica nelle sue coordinate $\|X\|^2 = x_1^2 + \ldots + x_n^2$.

(e) Sia $f\colon \mathbb{R}^n \to \mathbb{R}$ una funzione e sia $a \in \mathbb{R}^n$ un punto critico di f (un punto in cui si annullano tutte le derivate parziali di f). L'Hessiano di f in a determina una forma quadratica data da

$$H(v_1, \dots, v_n) = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a) v_i v_j, \qquad v_i = (x_i - a_i).$$

La natura del punto critico dipende dal segno della funzione H al variare di $(v_1, \dots, v_n) \neq (0, \dots, 0)$: si tratta di un minimo locale (in senso stretto) se $H(v_1, \dots, v_n) > 0$, per ogni $(v_1, \dots, v_n) \neq (0, \dots, 0)$, di un massimo locale (in senso stretto) se $H(v_1, \dots, v_n) < 0$, per ogni $(v_1, \dots, v_n) \neq (0, \dots, 0)$, né di un massimo né di un minimo se $H(v_1, \dots, v_n)$ assume sia valori positivi che negativi al variare di $(v_1, \dots, v_n) \neq (0, \dots, 0)$.

(f) Sia data l'equazione di secondo grado

$$\mathcal{E}: \quad ax_1^2 + bx_2^2 + cx_1x_2 + dx_1 + ex_2 + f = 0$$

in $\mathbb{R}^2$. La forma quadratica ad essa associata $Q(x_1, x_2) = ax_1^2 + bx_2^2 + cx_1x_2$ determina il tipo di luogo geometrico definito da $\mathcal{E}$. Ad esempio, se $ax_1^2 + bx_2^2 + cx_1x_2 > 0$, per ogni $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, tale luogo geometrico è un'ellisse, un punto o l'insieme vuoto. Dalla classificazione delle forme quadratiche in due variabili segue la classificazione delle coniche del piano; dalla classificazione delle forme quadratiche in tre variabili segue la classificazione delle quadriche dello spazio.

Per studiare una forma quadratica F e *determinare un sistema di coordinate rispetto al quale la sua espressione sia più semplice possibile*, la scriviamo in forma matriciale come

$$F(X) = {}^tXAX,$$

dove

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n \quad \text{ed} \quad A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{12} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{pmatrix}$$

è la matrice simmetrica con coefficienti

$$a_{ii} = \mathrm{coeff}(x_i^2),\ i = 1, \dots, n, \qquad a_{ij} = a_{ji} = \frac{1}{2}\mathrm{coeff}(x_i x_j),\ 1 \le i < j \le n.$$

Esempio. La forma quadratica $F(x_1, x_2) = x_1^2 + x_2^2 + 3x_1x_2$ si scrive come $F(X) = {}^tX\,AX$, con $A = \begin{pmatrix} 1 & 3/2 \\ 3/2 & 1 \end{pmatrix}$. La forma quadratica $F(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 + 3x_1x_2 + 2x_1x_3$ si scrive come $F(X) = {}^tX\,AX$, con $A = \begin{pmatrix} 1 & 3/2 & 1 \\ 3/2 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$.

Consideriamo adesso il cambiamento di coordinate in $\mathbb{R}^n$, dato da

$$X = MY, \tag{1}$$

dove M è una matrice invertibile $n \times n$. Sostituendo la relazione (1) nell'espressione della forma quadratica $F(X) = {}^tXAX$ troviamo

$${}^tXAX = {}^t(MY)A(MY) = {}^tY({}^tMAM)Y. \tag{2}$$

Ciò significa che nelle coordinate $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ la matrice simmetrica associata alla *stessa* forma quadratica è data da

$${}^tMAM.$$

La relazione (2) dice che

Due matrici simmetriche A e B definiscono la stessa forma quadratica (rispetto a sistemi di coordinate diversi) se e solo se sono congruenti.

Osservazione. Se il cambiamento di coordinate $X = MY$ è dato da una matrice ortogonale (caratterizzata dalla relazione ${}^tMM = I_n$), allora

$${}^tMAM = M^{-1}AM, \tag{3}$$

e le matrici simmetriche A e ${}^tMAM$ sono anche *coniugate.*

Per le proprietà di diagonalizzabilità delle matrici simmetriche, la relazione (3) ci permette di determinare un sistema di coordinate rispetto al quale la forma quadratica $F(X) = {}^tXAX$ non ha termini misti. Si tratta di coordinate indotte da una base ortonormale di $\mathbb{R}^n$ formata da autovettori di A.

Teorema 1. *Sia $F(X) = {}^tXAX$ una forma quadratica reale, dove A è una matrice simmetrica. Esiste un cambiamento di coordinate $X = MY$, dato da una matrice ortogonale M, che trasforma F in un'espressione del tipo*

$$\lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_n y_n^2, \tag{4}$$

dove i coefficienti $\lambda_1, \ldots, \lambda_n$ sono gli autovalori di A.

Dim. Dai risultati della sezione precedente, data una matrice simmetrica A esiste una matrice ortogonale M tale che

$$M^{-1}AM = D = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 & 0 \\ 0 & \lambda_2 & \ldots & 0 & 0 \\ 0 & 0 & \ldots & \ldots & 0 \\ 0 & 0 & \ldots & \ldots & \lambda_n \end{pmatrix},$$

dove $\lambda_1, \ldots, \lambda_n$ sono gli autovalori di A. Precisamente M è la matrice del cambiamento di base da una base ortonormale $\mathcal{B}$ di $\mathbb{R}^n$ formata da autovettori di A (di autovalori $\lambda_1, \ldots, \lambda_n$, rispettivamente) alla base canonica:

$$M = C_{\mathcal{B},Can} = \begin{pmatrix} v_{11} & \ldots & v_{1n} \\ \vdots & \vdots & \vdots \\ v_{n1} & \ldots & v_{nn} \end{pmatrix}, \quad \mathcal{B} = \{ \begin{pmatrix} v_{11} \\ \vdots \\ v_{n1} \end{pmatrix}, \ldots, \begin{pmatrix} v_{1n} \\ \vdots \\ v_{nn} \end{pmatrix} \}.$$

Se $X = MY$ è il cambiamento di coordinate corrispondente, dalla (3) segue che

$$F(Y) = {}^tY{}^tMAMY = {}^tYM^{-1}AMY = {}^tYDY = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_n y_n^2,$$

come richiesto.

Osservazione. L'espressione (4) è unica a meno dell'ordine dei coefficienti $\lambda_1, \ldots, \lambda_n$ e si chiama *forma canonica metrica* della forma quadratica. La proposizione precedente classifica le forme quadratiche a meno di cambiamenti di coordinate ortogonali.

Esercizio. Se X è autovettore di A relativo all'autovalore λ, allora $F(X) = \lambda\|X\|^2$.
Dim. Se X è autovettore di A relativo all'autovalore λ, allora $F(X) = {}^tXAX = {}^tX\lambda X = \lambda\|X\|^2$.

Corollario. *Siano λ_1 e λ_n rispettivamente il minimo e il massimo autovalore di A. Allora il minimo ed il massimo della funzione $F(X) = {}^tXAX$ sull'insieme*

$$S = \{X \in \mathbb{R}^n \mid \|X\| = 1\}$$

sono λ_1 e λ_n. Se X_0 è un autovettore unitario di autovalore λ_1, allora $F(X_0) = \lambda_1$. Analogamente se Y_0 è un autovettore unitario di autovalore λ_n, allora $F(Y_0) = \lambda_n$.

Dim. Siano λ_1 e λ_n rispettivamente il minimo e il massimo autovalore di A. Dimostriamo innanzitutto che

$$\lambda_1\|X\|^2 \leq F(X) \leq \lambda_n\|X\|^2, \quad \text{per ogni } X \in \mathbb{R}^n.$$

Dall'equazione (4) e dal fatto che $\|X\|^2 = \|MY\|^2 = \|Y\|^2$, abbiamo le stime richieste

$$F(X) = F(MY) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_n y_n^2 \leq \lambda_n\|Y\|^2 = \lambda_n\|X\|^2;$$

$$F(X) = F(MY) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_n y_n^2 \geq \lambda_1\|Y\|^2 = \lambda_1\|X\|^2.$$

La parte rimanente dell'enunciato segue direttamente dall'esercizio precedente.

Consideriamo adesso cambiamenti di coordinate $X = NY$, dove N è una matrice invertibile *qualunque.* In questo caso l'espressione di una forma quadratica può essere ulteriormente semplificata e portata nella cosidetta *forma canonica affine.*

Teorema 2. *Sia $F(X) = {}^tXAX$ una forma quadratica reale, dove A è una matrice simmetrica. Esiste un cambiamento di coordinate $X = NY$, dato da una matrice invertibile N, che trasforma F in un'espressione del tipo*

$$\varepsilon_1 z_1^2 + \varepsilon_2 z_2^2 + \ldots + \varepsilon_n z_n^2, \tag{5}$$

dove i coefficienti $\varepsilon_1, \ldots, \varepsilon_n$ sono uguali a $1, -1, 0$. Precisamente ci sono tanti coefficienti uguali ad 1 quanti sono gli autovalori positivi di A, tanti coefficienti uguali a -1 quanti sono gli autovalori negativi di A e tanti coefficienti uguali a 0 quanti sono gli autovalori nulli di A.

Dim. La proposizione è una conseguenza diretta della relazione (2) e del teorema di Sylvester.

In molte applicazioni si richiede determinare il *segno* di una forma quadratica al variare di $X = (x_1, \ldots, x_n)$ in $\mathbb{R}^n$, con $X \neq (0, \ldots, 0)$ (osserviamo che per l'omogeneità $F(0, \ldots, 0) = 0$).

Definizione.

- Una forma quadratica F si dice *definita positiva* se $F(X) > 0$, per ogni $X \neq O$; si dice *semidefinita positiva* se $F(X) \geq 0$, per ogni $X \neq O$.
- Una forma quadratica F si dice *definita negativa* se $F(X) < 0$, per ogni $X \neq O$; si dice *semidefinita negativa* se $F(X) \leq 0$, per ogni $X \neq O$.
- Una forma quadratica F si dice *indefinita* se, al variare di $X \neq O$, assume sia valori positivi che valori negativi.

Osservazione. Dai teoremi 1 e 2 segue che il segno di una forma quadratica F dipende esclusivamente dal il segno e dalla nullità degli autovalori di una qualunque matrice simmetrica ad essa associata:

- Una forma quadratica F è *definita positiva* (risp. *definita negativa*) se e solo se tutti gli autovalori di A sono positivi (risp. negativi).
- Una forma quadratica F è *semidefinita positiva* (risp. *semidefinita negativa*) se e solo se tutti gli autovalori di A sono non negativi (risp. non positivi).
- Una forma quadratica F è *indefinita* se e solo se A ha sia autovalori positivi che negativi.

Il seguente criterio ci permette di determinare il segno degli autovalori non nulli di una matrice simmetrica, anche senza calcolarli esplicitamente.

Criterio di Cartesio. *Sia $P(\lambda) = a_n\lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0$ un polinomio di grado n in λ le cui radici sono tutte* reali *e non nulle (se il polinomio ha k radici nulle, possiamo ridurci al caso di un polinomio con radici non nulle mettendo in evidenza il fattore λ^k). Sia $a_n, a_{n-1}, \ldots, a_1, a_0$ la successione dei coefficienti di P. Allora il numero di radici positive di P è uguale al numero di variazioni di segno nel passare da a_n al coefficiente non nullo successivo, e cosi' via.*