1 Richiami di algebra lineare

Definizione 1.1 (matrici e vettori) *Una* matrice A *e' un insieme di numeri* A_{hk}*,* $h = 1, ..., m$*,* $k = 1, ..., n$*, ordinati in base alla coppia di indici* h *e* k *nel modo seguente*

$$\begin{bmatrix} A_{11} & . & . & . & A_{1n} \\ A_{21} & . & . & . & A_{2n} \\ . & . & . & . & . \\ . & . & . & . & . \\ A_{m1} & . & . & . & A_{mn} \end{bmatrix}$$

m *ed* n *si dicono dimensioni della matrice e la matrice si dice* $m \times n$*. Una matrice* $m \times 1$ *si dice anche* vettore colonna *e una matrice* $1 \times n$ vettore riga.

Un insieme di relazioni affini

$$\begin{gathered} y_1 = A_{11}x_1 + A_{12}x_2 + \ ... \ + A_{1n}x_n + b_1 \\ y_2 = A_{21}x_1 + A_{22}x_2 + \ ... \ + A_{2n}x_n + b_2 \\ .... \\ y_m = A_{m1}x_1 + A_{m2}x_2 + \ ... \ + A_{mn}x_n + b_m \end{gathered}$$

si scrive, in forma matriciale,

$$\begin{bmatrix} y_1 \\ . \\ . \\ . \\ . \\ y_m \end{bmatrix} = \begin{bmatrix} A_{11} & . & . & . & A_{1n} \\ A_{21} & . & . & . & A_{2n} \\ . & . & . & . & . \\ . & . & . & . & . \\ A_{m1} & . & . & . & A_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ . \\ . \\ . \\ . \\ x_n \end{bmatrix} + \begin{bmatrix} b_1 \\ . \\ . \\ . \\ . \\ b_m \end{bmatrix}$$

cioe', in forma matriciale compatta,

$$y = Ax$$

Definizione 1.2 (prodotto righe per colonne) *Data una matrice* A $m \times p$ *e una matrice* B $p \times n$ *il loro prodotto righe per colonne e' la matrice* $m \times n$ *data da*

$$AB = \begin{bmatrix} (A_{11}B_{11} + A_{12}B_{21} + ... + A_{1p}B_{p1}) & (A_{11}B_{12} + A_{12}B_{22} + ... + A_{1p}B_{p2}) & . & . & . & (A_{11}B_{1n} + \\ (A_{21}B_{11} + A_{22}B_{21} + ... + A_{2p}B_{p1}) & (A_{21}B_{12} + A_{22}B_{22} + ... + A_{2p}B_{p2}) & . & . & . & (A_{21}B_{1n} + \\ . & . & . & . & . & \\ . & . & . & . & . & \\ . & . & . & . & . & \\ (A_{m1}B_{11} + A_{m2}B_{21} + ... + A_{mp}B_{p1}) & (A_{m1}B_{12} + A_{m2}B_{22} + ... + A_{mp}B_{p2}) & . & . & . & (A_{m1}B_{1n} + \end{bmatrix}$$

Se A e B sono matrici quadrate (cioe' una matrici $n \times n$) sono definiti sia AB che BA ma, in generale,

$$AB \neq BA$$

Definizione 1.3 (matrice inversa: definizione) *La matrice*

$$I = \begin{bmatrix} 1 & 0 & . & . & 0 \\ 0 & 1 & 0 & . & 0 \\ . & . & . & . & . \\ . & . & . & . & 0 \\ 0 & . & . & 0 & 1 \end{bmatrix}$$

si dice matrice identita'. *Data una matrice quadrata (cioe' una matrice $n \times n$), A, la sua matrice inversa (se esiste) e' la matrice A^{-1} tale che*

$$AA^{-1} = I = A^{-1}A$$

Definizione 1.4 (matrice trasposta) *La* matrice trasposta A^T *e' la matrice ottenuta trasformando le righe di A in colonne, cioe'*

$$A^T = \begin{bmatrix} A_{11} & . & . & . & A_{m1} \\ A_{12} & . & . & . & A_{m2} \\ . & . & . & . & . \\ . & . & . & . & . \\ A_{1m} & . & . & . & A_{mn} \end{bmatrix}$$

Definizione 1.5 (determinante) *Il determinante di una matrice 2×2 $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ e' dato da*

$$det(A) = A_{11}A_{22} - A_{12}A_{21}$$

e si indica anche con

$$det(A) = \begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix}$$

Il determinante di una matrice 3×3 $A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}$ e' dato da

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

Il determinante di una matrice 4×4 $A = \begin{bmatrix} A_{11} & . & . & A_{14} \\ A_{21} & . & . & A_{24} \\ . & . & . & . \\ A_{41} & . & . & A_{44} \end{bmatrix}$ *e' dato da*

$$det(A) = A_{11} \begin{vmatrix} A_{22} & A_{23} & A_{24} \\ A_{32} & A_{33} & A_{34} \\ A_{42} & A_{43} & A_{44} \end{vmatrix} - A_{21} \begin{vmatrix} A_{12} & A_{13} & A_{14} \\ A_{32} & A_{33} & A_{34} \\ A_{42} & A_{43} & A_{44} \end{vmatrix} + A_{31} \begin{vmatrix} A_{12} & A_{13} & A_{14} \\ A_{22} & A_{23} & A_{24} \\ A_{42} & A_{43} & A_{44} \end{vmatrix} - A_{41} \begin{vmatrix} A_{12} & A_{13} & A_{14} \\ A_{22} & A_{23} & A_{24} \\ A_{32} & A_{33} & A_{34} \end{vmatrix}$$

Analogamente (per induzione) si definisce il determinante di una matrice $n \times n$ *con* n *qualunque.*

Proposizione 1.1 (matrice inversa: esistenza e calcolo) *La matrice inversa* A^{-1} *esiste se e solo* $det(A) \neq 0$*. In tal caso e' data da*

$$A^{-1} = \frac{1}{det(A)} \begin{bmatrix} \alpha_{11} & -\alpha_{12} & . & . & (-1)^{1+n}\alpha_{1n} \\ -\alpha_{21} & \alpha_{22} & . & . & (-1)^{2+n}\alpha_{2n} \\ . & . & . & . & . \\ . & . & . & . & . \\ (-1)^{m+1}\alpha_{m1} & . & . & . & (-1)^{m+n}\alpha_{mn} \end{bmatrix}$$

dove α_{hk} *e' il determinante della matrice che si ottiene da* A^T *cancellando la riga* h *e la riga* k*.*

Definizione 1.6 (rango) *Il rango di una matrice* $m \times n$*,* A*, e' il massimo intero* r *per cui esiste una sottomatrice di* A *di dimensioni* $r \times r$

$$\begin{bmatrix} A_{i_1 j_1} & . & . & . & A_{i_1 j_r} \\ A_{i_2 j_1} & . & . & . & A_{i_2 j_r} \\ . & . & . & . & . \\ . & . & . & . & . \\ A_{i_r j_1} & . & . & . & A_{i_r j_r} \end{bmatrix}$$

con determinante non nullo.

Definizione 1.7 (matrice simmetrica) *Una matrice si dice simmetrica se*

$$A = A^T$$

Definizione 1.8 (matrice definita positiva) *Una matrice simmetrica* A 2×2 *si dice definita positiva se*

$$\begin{aligned} & x^T A x \\ &= \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{12} & A_{22} \end{bmatrix} \begin{bmatrix} x \\ \\ y \end{bmatrix} \\ &= A_{11}x^2 + A_{22}y^2 + 2A_{12}xy > 0 \end{aligned}$$

per ogni $(x, y) \neq (0, 0)$.

In generale una matrice simmetrica A $n \times n$ *si dice definita positiva se*

$$
\begin{aligned}
& x^T A x \\
&= \begin{bmatrix} x_1 & . & . & . & . & x_n \end{bmatrix} \begin{bmatrix} A_{11} & . & . & . & A_{1n} \\ A_{12} & . & . & . & A_{2n} \\ . & . & . & . & . \\ . & . & . & . & . \\ A_{1n} & . & . & . & A_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ . \\ . \\ . \\ . \\ x_n \end{bmatrix} \\
&= A_{11}x_1^2 + A_{22}x_2^2 + \ ... \ + 2A_{12}x_1x_2 + 2A_{13}x_1x_3 + \ ... + 2A_{23}x_2x_3 + \ ... \ + 2A_{(n-1)n}x_{n-1}x_n \\
&> 0
\end{aligned}
$$

per ogni $x \in \mathbf{R}^n$ *le cui componenti non siano tutte nulle.*

2 Coppie di variabili aleatorie congiuntamente Gaussiane (o normali)

Definizione 2.1 *Due variabili aleatorie* X *e* Y *si dicono congiuntamente Gaussiane (o normali) se la loro densita' congiunta e' data da*

$$
\begin{aligned}
f(x, y) &= K\text{exp}\left\{-\tfrac{1}{2}\begin{bmatrix} (x-\mu_1) & (y-\mu_2) \end{bmatrix} C^{-1} \begin{bmatrix} (x-\mu_1) \\ (y-\mu_2) \end{bmatrix}\right\} \\
&= K\text{exp}\left\{-\tfrac{1}{2}\left((C^{-1})_{11}(x-\mu_1)^2 + (C^{-1})_{22}(y-\mu_2)^2 + 2\,(C^{-1})_{12}(x-\mu_1)(y-\mu_2)\right\}\right.
\end{aligned}
$$

dove K *e' la costante di normalizzazione,* μ_1 *e* μ_2 *sono numeri reali e* C^{-1} *e' una matrice simmetrica definita positiva.* μ_1 *e* μ_2 *risultano essere rispettivamente* $\mathbf{E}[X]$ *ed* $\mathbf{E}[Y]$ *e* C *risulta essere la matrice di covarianza di* X *e* Y*:*

$$
C = \begin{bmatrix} \mathbf{Var}[X] & \mathbf{Cov}[X, Y] \\ \mathbf{Cov}[X, Y] & \mathbf{Var}[Y] \end{bmatrix}
$$

Osservazione 2.1 (Gaussiane indipendenti) *Se* X *e* Y *sono ciascuna Gaussiana,* X *con media* μ_1 *e varianza* σ_1^2*,* Y *con media* μ_2 *e varianza* σ_2^2*, e sono indipendenti, allora sono congiuntamente Gaussiane. Infatti in questo caso*

$$
f(x, y) = f_X(x)f_Y(y) = K\text{exp}\left\{-\frac{1}{2}\left(\frac{(x-\mu_1^2)}{\sigma_1^2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right)\right\}
$$

D'altro canto

$$
C = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}, \quad C^{-1} = \begin{bmatrix} \frac{1}{\sigma_1^2} & 0 \\ 0 & \frac{1}{\sigma_2^2} \end{bmatrix}, \quad \begin{bmatrix} (x-\mu_1) & (y-\mu_2 \end{bmatrix}) C^{-1} \begin{bmatrix} (x-\mu_1) \\ (y-\mu_2) \end{bmatrix} = \frac{(x-\mu_1^2)}{\sigma_1^2} + \frac{(y-\mu_2)^2}{\sigma_2^2}
$$

e quindi $f(x, y)$ *e' una densita' congiuntamente Gaussiana.*

Esempio 2.1 *Siano X e Y variabili aleatorie Gaussiane di medie rispettivamente 0 e -3 e matrice di covarianza*

$$\begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$$

Allora

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

e quindi la densita' di probabilita' congiunta di X e Y e'

$$f_{X,Y}(x,y) = K\mathrm{e}^{-\frac{2x^2+2xy+2(y+3)^2}{2}}$$

Proposizione 2.1 (proprieta' delle Gaussiane) *Se due variabili X e Y sono congiuntamente Gaussiane:*

a) le marginali di X e di Y sono Gaussiane;

b) se la coppia (U,V) e' una trasformazione affine della coppia (X,Y), cioe'

$$U = A_{11}X + A_{12}Y + b_1$$

$$V = A_{21}X + A_{22}Y + b_2$$

e

$$det(A) \neq 0$$

allora U e V sono congiuntamente Gaussiane; in particolare, per il punto a), U e V sono ciascuna una variabile aleatoria Gaussiana;

c) se X e Y sono scorrelate (cioe' $\mathbf{Cov}[X,Y] = 0$) allora sono indipendenti.

Osservazione 2.2 *Se*

$$U = A_{11}X + A_{12}Y + b_1$$

possiamo sempre accoppiare a questa relazione la relazione

$$Y = Y$$

cosicche' la coppia di variabili aleatorie (U,Y) sia una trasformazione affine di (X,Y) (la matrice A sara' $\begin{bmatrix} A_{11} & A_{12} \\ 0 & 1 \end{bmatrix}$). Quindi la coppia (U,Y) sara' congiuntamente Gaussiana e U sara' Gaussiana. Percio' una qualunque combinazione affine U di due variabili aleatorie congiuntamente Gaussiane (in particolare indipendenti) e' sempre una variabile aleatoria Gaussiana.

3 n-ple di variabili aleatorie congiuntamente Gaussiane (o normali)

Tutto quanto detto per le coppie di variabili aleatorie congiuntamente Gaussiane rimane valido per un qualunque numero n di variabili aleatorie X_1, ..., X_n.

Definizione 3.1 *n variabili aleatorie X_1, ..., X_n si dicono congiuntamente Gaussiane (o normali) se la loro densita' congiunta e' data da*

$$f(x_1,...,x_n) = K\exp\left\{ -\frac{1}{2}\left[\begin{array}{ccccc} (x_1-\mu_1) & . & . & . & (x_n-\mu_n) \end{array} \right] C^{-1} \left[\begin{array}{c} (x-\mu_1) \\ . \\ . \\ . \\ (x_n-\mu_n) \end{array} \right] \right\}$$

dove K e' la costante di normalizzazione, μ_1, ... , μ_n sono numeri reali e C e' una matrice definita positiva. μ_1, ... , μ_n risultano essere le medie e C risulta essere la matrice di covarianza.

Osservazione 3.1 *Se X_1, ..., X_n sono variabili aleatorie indipendenti, ciascuna Gaussiana, allora sono congiuntamente Gaussiane.*

Proposizione 3.1 (proprieta' delle Gaussiane) *Se X_1, ..., X_n sono congiuntamente Gaussiane:*

a) Per ogni sottofamiglia $\{i_1, i_2, ..., i_k\}$ di $\{1, ..., n\}$, $X_{i_1}, ..., X_{i_k}$ sono congiuntamente Gaussiane; in particolare ciascuna variabile X_i, $i = 1, ..., n$, e' Gaussiana;

b) se $(Y_1, ..., Y_m)$, $m \leq n$, e' una trasformazione affine di $(X_1, ..., X_n)$, cioe'

$$\left[\begin{array}{c} Y_1 \\ . \\ . \\ . \\ . \\ Y_m \end{array} \right] = \left[\begin{array}{ccccc} A_{11} & . & . & . & A_{1n} \\ A_{21} & . & . & . & A_{2n} \\ . & . & . & . & . \\ . & . & . & . & . \\ A_{m1} & . & . & . & A_{mn} \end{array} \right] \left[\begin{array}{c} X_1 \\ . \\ . \\ . \\ . \\ X_n \end{array} \right] + \left[\begin{array}{c} b_1 \\ . \\ . \\ . \\ . \\ b_m \end{array} \right]$$

e il rango di A e' m, allora Y_1, ..., Y_m sono congiuntamente Gaussiane; in particolare, per il punto a), ciascuna e' una variabile aleatoria Gaussiana;

c) se X_i e X_j sono scorrelate (cioe' $\mathbf{Cov}[X_i, X_j] = 0$) allora sono indipendenti.

Osservazione 3.2 *Se*

$$U = A_{11}X_1 + A_{12}X_2 + \ ... \ + A_{1n}X_n + b_1$$

possiamo sempre accoppiare a questa relazione le relazioni

$$X_2 = X_2, \quad X_3 = X_3, \quad ... \quad X_n = X_n$$

cosicche' $(U, X_2, ..., X_n)$ sara' una trasformazione affine di $(X_1, ..., X_n)$ (la matrice A sara' $\begin{bmatrix} A_{11} & . & . & . & A_{1n} \\ 0 & 1 & 0 & . & 0 \\ . & . & . & . & . \\ . & . & . & . & 0 \\ 0 & . & . & 0 & 1 \end{bmatrix}$*). Quindi U, X_2, ..., X_n saranno congiuntamente Gaussiane e U sara' Gaussiana. Percio' una qualunque combinazione affine U di variabili aleatorie congiuntamente Gaussiane (in particolare indipendenti) e' sempre una variabile aleatoria Gaussiana.*

Esempio 3.1 Sia (X, Y, Z) una terna di v.a. con densita' congiunta Gaussiana, medie nulle e matrice di covarianza

$$\begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix}$$

e siano

$$\begin{array}{l} U = X - Z + 5 \\ V = Y - 2X \\ W = Y + Z \end{array} .$$

(U, V, W) e' una trasformazione affine di (X, Y, Z) percio' sappiamo che (U, V, W) ha densita' congiunta gaussiana, e ci e' sufficiente calcolare le medie e la matrice di covarianza di (U, V, W). Essendo le medie di X, Y, Z nulle, il vettore le medie di U, V, W sono $5, 0, 0$. Per calcolare la matrice di covarianza, basta calcolare le varianze e covarianze di U, V e W, utilizzando le varianze e covarianze di X, Y e Z, che possiamo leggere nella matrice di covarianza:

$$\begin{aligned} \mathbf{Var}[U] &= \mathbf{Var}[X] + \mathbf{Var}[-Z] + 2\mathbf{Cov}[X, -Z] \\ &= \mathbf{Var}[X] + \mathbf{Var}[Z] - 2\mathbf{Cov}[X, Z] = 1 + 2 = 3 \end{aligned}$$

$$\begin{aligned} \mathbf{Var}[V] &= \mathbf{Var}[Y] + \mathbf{Var}[-2X] + 2\mathbf{Cov}[Y, -2X] \\ &= \mathbf{Var}[Y] + 4\mathbf{Var}[X] - 4\mathbf{Cov}[Y, X] = 2 + 4 + 4 = 10 \end{aligned}$$

$$\begin{aligned} \mathbf{Var}[W] &= \mathbf{Var}[Y] + \mathbf{Var}[Z] + 2\mathbf{Cov}[Y, Z] \\ &= 2 + 2 + 2 = 6 \end{aligned}$$

$$\begin{aligned} \mathbf{Cov}[U, V] &= \mathbf{Cov}[X, Y] + \mathbf{Cov}[X, -2X] + \mathbf{Cov}[-Z, Y] + \mathbf{Cov}[-Z, -2X] \\ &= \mathbf{Cov}[X, Y] - 2\mathbf{Var}[X] - \mathbf{Cov}[Z, Y] + 2\mathbf{Cov}[Z, X] \\ &= -1 - 2 - 1 = -4 \end{aligned}$$

$$\begin{aligned}\mathbf{Cov}[U,W] &= \mathbf{Cov}[X,Y]+\mathbf{Cov}[X,Z]+\mathbf{Cov}[-Z,Y]+\mathbf{Cov}[-Z,Z]\\ &= \mathbf{Cov}[X,Y]+\mathbf{Cov}[X,Z]-\mathbf{Cov}[Z,Y]-\mathbf{Var}[Z]\\ &= -1-1-2=-4\end{aligned}$$

$$\begin{aligned}\mathbf{Cov}[V,W] &= \mathbf{Cov}[Y,Y]+\mathbf{Cov}[Y,Z]+\mathbf{Cov}[-2X,Y]+\mathbf{Cov}[-2X,Z]\\ &= \mathbf{Var}[Y]+\mathbf{Cov}[Y,Z]-2\mathbf{Cov}[X,Y]-2\mathbf{Cov}[X,Z]\\ &= 2+1+2=5\end{aligned}$$

Troviamo percio' che la matrice di covarianza di (U,V,W) e'

$$\begin{bmatrix} 3 & -4 & -4\\ -4 & 10 & 5\\ -4 & 5 & 6\end{bmatrix}$$

Esempio 3.2 Sia (X_1,X_2,X_3,X_4) un vettore gaussiano di media 0 (cioe' il vettore delle medie e' il vettore nullo) e matrice di covarianza

$$\begin{bmatrix} 1 & 0 & 1 & 0\\ 0 & 4 & 0 & 2\\ 1 & 0 & 3 & 0\\ 0 & 2 & 0 & 10\end{bmatrix}$$

Dalla matrice di covarianza vediamo che X_1 e' indipendente da X_2 e da X_4 e che X_3 e' indipendente da X_2 e da X_4.

Sia

$$U = X_1 - 0.1\,X_2 - 2X_3 + 0.5X_4 + 5$$

Allora U e' Gaussiana di media 5 e varianza $\mathbf{Var}[U] = 35.34$