Computazione per l'interazione naturale: fondamenti probabilistici (2)

Corso di Interazione uomo-macchina II

Prof. Giuseppe Boccignone

Dipartimento di Informatica
Università di Milano

boccignone@di.unimi.it
http://boccignone.di.unimi.it/IUM2_2014.html

Estensione a due variabili casuali //distribuzioni cumulative congiunte e marginali

- Date due variabili casuali continue X, Y , la distribuzione di probabilità cumulativa congiunta è definita come

$$F_{XY}(x, y) = P(X \leq x, Y \leq y)$$

- Valgono le seguenti

1. $0 \leq F_{XY}(x, y) \leq 1$
2. $\lim_{x,y \to \infty} F_{XY}(x, y) = 1$
3. $\lim_{x,y \to -\infty} F_{XY}(x, y) = 0$

- Data la distribuzione cumulativa di probabilità congiunta $F_{XY}(x, y)$ le distribuzioni cumulative marginali di probabilità F_X e F_Y sono definite come

$$F_X(x) = \lim_{y \to \infty} F_{XY}(x, y)$$

$$F_Y(y) = \lim_{x \to \infty} F_{XY}(x, y)$$

Estensione a due variabili casuali //distribuzioni di probabilità congiunte e marginali

- Date due variabili casuali discrete X, Y

$$p_{XY}(x,y) = P(X = x, Y = y)$$

- valgono le seguenti

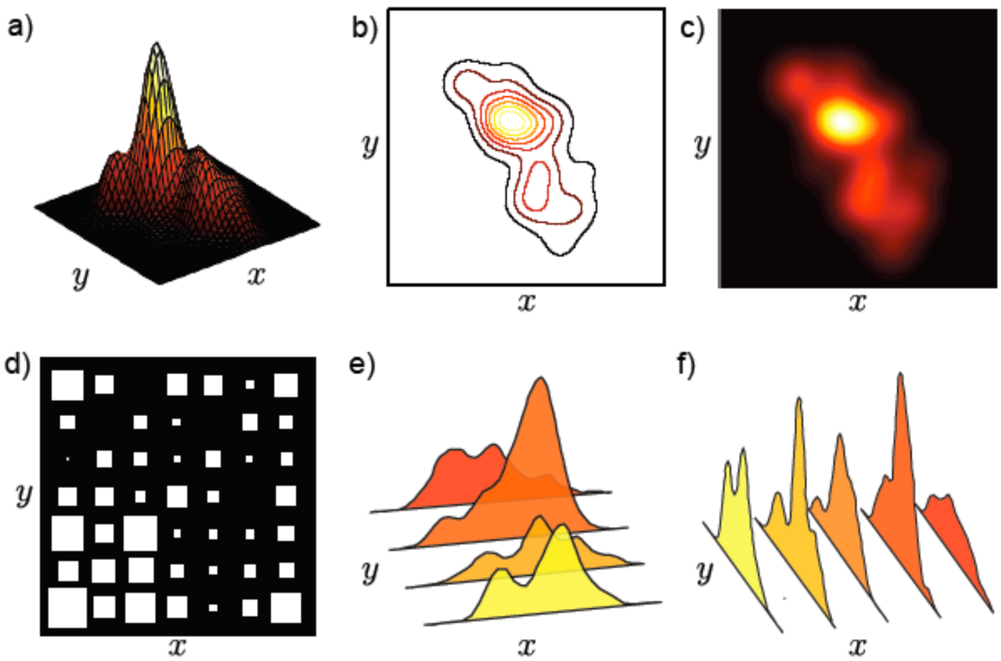
1. $0 \leq p_{XY}(x,y) \leq 1$
2. $\sum_{x \in V(X)} \sum_{y \in V(Y)} p_{XY}(x,y) = 1$

- Data la distribuzione di probabilità p_{XY} (x, y) le distribuzioni marginali di probabilità p_X e p_Y sono definite come

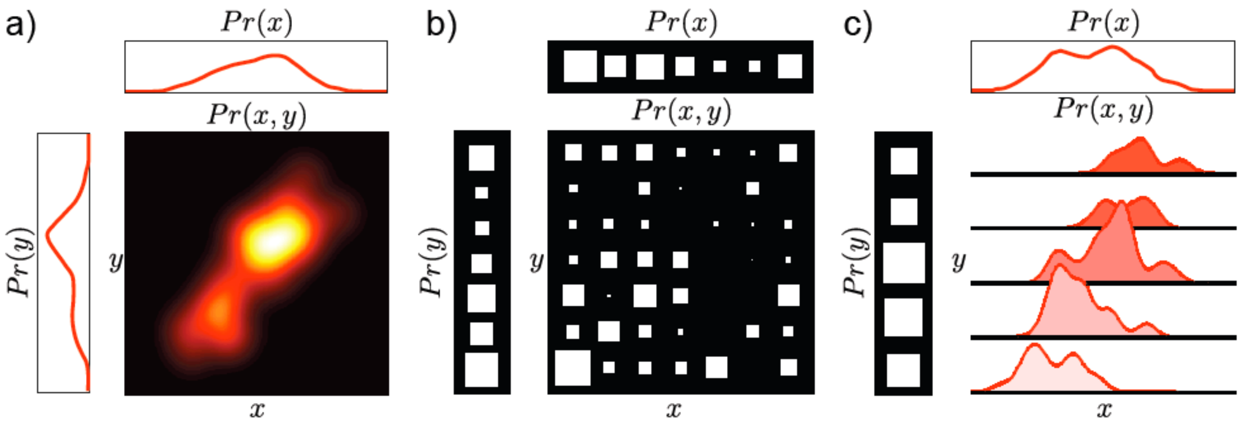
$$p_X(x) = \sum_{y \in V(Y)} p_{XY}(x,y)$$

$$p_Y(y) = \sum_{x \in V(Y)} p_{XY}(x,y)$$

Estensione a due variabili casuali //distribuzioni di probabilità congiunte e marginali



Estensione a due variabili casuali
//distribuzioni di probabilità congiunte e marginali



Estensione a due variabili casuali
//densità di probabilità congiunte e marginali

- Se $F_{XY}(x, y)$ è dovunque derivabile rispetto sia a x che a y, allora la densità di probabilità congiunta è definita come

$$f_{XY}(x,y) = \frac{\partial^2 F_{XY}(x,y)}{\partial x \partial y}$$

- vale

$$\int\int_{x \in A} f_{XY}(x,y)dxdy = P((X,Y) \in A)$$

- Data la densità di probabilità cumulativa $f_{XY}(x, y)$ le densità marginali di probabilità f_X e f_Y sono definite come

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)dy$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)dx$$

Estensione a due variabili casuali //probabilità condizionate

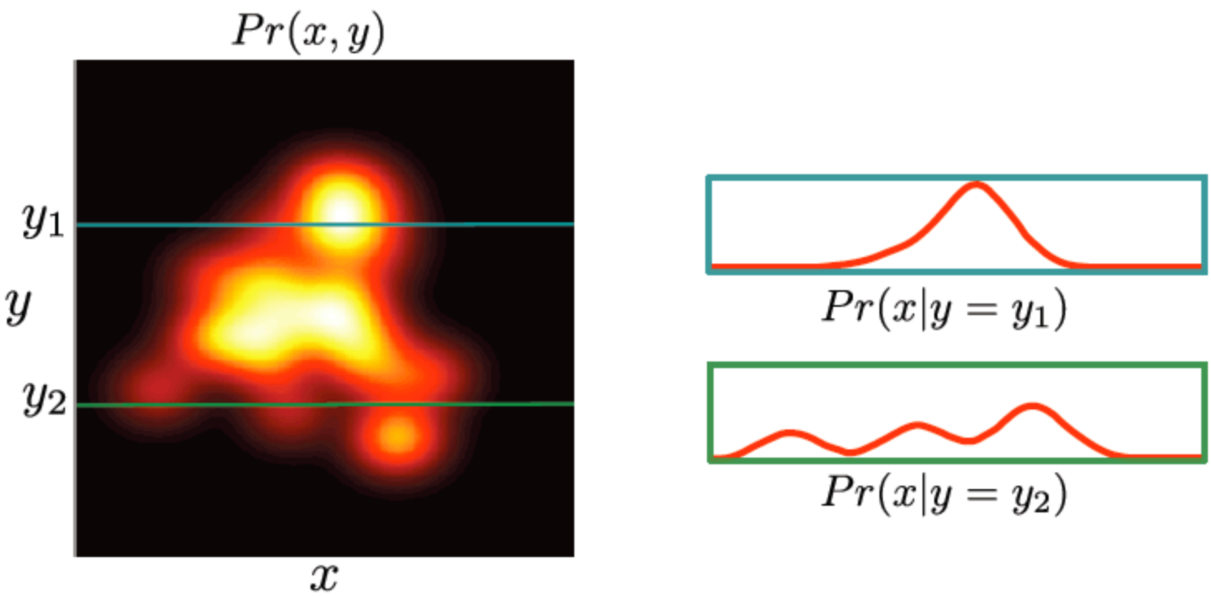
- Date due variabili casuali X, Y discrete , la probabilità condizionata di X rispetto a Y è definita come

$$p_{X|Y}(x \mid y) = \frac{p_{XY}(x,y)}{p_Y(y)}$$

- X, Y continue

$$f_{X|Y}(x \mid y) = \frac{f_{XY}(x,y)}{f_Y(y)}$$

Estensione a due variabili casuali //probabilità condizionate



Estensione a due variabili casuali //Teorema di Bayes

- X, Y discrete

$$p_{X|Y}(x \mid y) = \frac{p_{XY}(x,y)}{p_Y(y)} = \frac{p_{Y|X}(y \mid x)p_X(x)}{\sum_{z \in V(X)} p_{Y|X}(y \mid z)p_X(z)}$$

- X, Y continue

$$f_{X|Y}(x \mid y) = \frac{f_{XY}(x,y)}{f_Y(y)} = \frac{f_{Y|X}(y \mid x)f_X(x)}{\int_{-\infty}^{\infty} f_{Y|X}(y \mid z)f_X(z)dz}$$

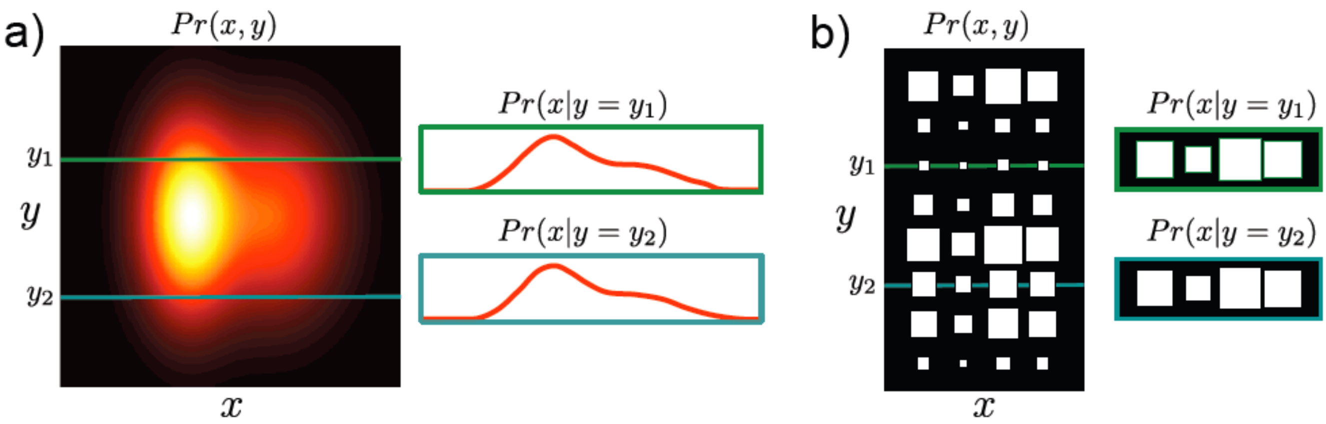
Estensione a due variabili casuali //indipendenza

- Due variabili casuali sono indipendenti se

$$F_{XY}(x,y) = F_X(x)F_Y(y)$$

- ovvero
  1. $p_{XY}(x,y) = p_X(x)p_Y(y)$ per X, Y discrete e per tutti i valori $x \in V(X), y \in V(Y)$
  2. $p_{X|Y}(x \mid y) = p_X(x)$, per X, Y discrete e per $x \in V(X)$ e $p_Y(y) \neq 0$
  3. $f_{XY}(x,y) = f_X(x)f_Y(y)$ per X, Y continue e per tutti i valori $x, y \in \mathbb{R}$
  4. $f_{X|Y}(x \mid y) = f_X(x)$, per X, Y continue e per $x \in \mathbb{R}$ e $f_Y(y) \neq 0$

Estensione a due variabili casuali //indipendenza



Estensione a due variabili casuali //valore atteso

- Date due variabili casuali X, Y discrete e una funzione $g : \mathbb{R}^2 \mapsto \mathbb{R}$ il valore atteso di g è

$$E[g(X,Y)] = \sum_{x \in V(X)} \sum_{y \in V(Y)} g(x,y) p_{XY}(x,y)$$

- X, Y continue

$$E[g(X,Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f_{XY}(x,y) dx dy$$

- Valgono le seguenti

1. $E[f(X,Y) + g(X,Y)] = E[f(X,Y)] + E[g(X,Y)]$
2. Se X e Y sono indipendenti, $E[f(X)g(Y)] = E[f(X)]E[g(Y)]$

Estensione a due variabili casuali //covarianze

- Date due variabili casuali X, Y , la loro covarianza è definita

$$Cov[X,Y] = E[(X - E[X])(Y - E[Y])]$$

- come nel caso della varianza:

$$\begin{aligned} Cov[X,Y] &= E[(X - E[X])(Y - E[Y])] \\ &= E[XY - XE[Y] - YE[X] + E[X]E[Y]] \\ &= E[XY] - E[X]E[Y] - E[Y]E[X] + E[E[X]E[Y]] \\ &= E[XY] - E[X]E[Y] \end{aligned}$$

- Valgono inoltre:

1. $Var[X+Y] = Var[X] + Var[Y] + 2Cov[X,Y]$
2. Se X e Y sono indipendenti, $Cov[X,Y] = 0$

Estensione a n variabili casuali

- Date n variabili causali $X_1, X_2, \dots, X_n$

$$F_{X_1,X_2,\dots,X_n}(x_1, x_2, \dots, x_n) = P(X_1 \le x_1, X_2 \le x_2, \dots, X_n \le x_n)$$

- valgono

1. $0 \le F_{X_1,X_2,\dots,X_n}(x_1, x_2, \dots, x_n) \le 1$
2. $\lim_{x_1 \to \infty} \lim_{x_2 \to \infty} \dots \lim_{x_n \to \infty} F_{X_1,X_2,\dots,X_n}(x_1, x_2, \dots, x_n) = 1$

- Per ogni X_i è definita la marginale

$$F_{X_i}(x_i) = \lim_{x_i \to \infty} F_{X_1,X_2,\dots,X_n}(x_1, x_2, \dots, x_n)$$

Estensione a n variabili casuali //caso discreto

- Date n variabili casuali discrete, la distribuzione di probabilità congiunta

$$p_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots,x_n) = P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n)$$

- per la quale valgono:

1. $0 \leq p_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots,x_n) \leq 1$
2. $\sum_{x_1 \in V(X_1)} \sum_{x_2 \in V(X_2)} \cdots \sum_{x_n \in V(X_n)} p_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots,x_n) = 1$

- La distribuzione di probabilità marginale è

$$p_{X_i}(x_i) = \sum_{x_1 \in V(X_1)} \cdots \sum_{x_{i-1} \in V(X_{i-1})} \sum_{x_{i+1} \in V(X_{i+1})} \cdots \sum_{x_n \in V(X_n)} p_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots,x_n)$$

Estensione a n variabili casuali //caso continuo

- Se la F è derivabile, la densità di probabilità congiunta è definita

$$f_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots,x_n) = \frac{\partial^n F_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots,x_n)}{\partial x_1 \partial x_2 \ldots \partial x_n}$$

- quindi

$$\int \cdots \int_{x \in A} f_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots,x_n) dx_1 \ldots dx_n = P((x_1,x_2,\ldots,x_n) \in A)$$

- La densità marginale di probabilità è

$$f_{X_i}(x_i) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots,x_n) dx_1 \ldots dx_n$$

Estensione a n variabili casuali //regola del prodotto

- Dalla definizione di probabilità condizionate:

$$
\begin{aligned}
f(x_1, x_2, \ldots, x_n) &= f(x_n \mid x_1, x_2, \ldots, x_{n-1}) f(x_1, x_2, \ldots, x_{n-1}) \\
&= f(x_n \mid x_1, x_2, \ldots, x_{n-1}) f(x_{n-1} \mid x_1, x_2, \ldots, x_{n-2}) f(x_1, x_2, \ldots, x_{n-2}) \\
&= \cdots \\
&= f(x_1) \prod_{i=2}^{n} f(x_i \mid x_1, \ldots, x_{i-1})
\end{aligned}
$$

Estensione a n variabili casuali //valore atteso

- Date una distribuzione f(x1, x2, . . . , xn) e una funzione $g : \mathbb{R}^n \mapsto \mathbb{R}$

$$E_f[g(x_1, x_2, \ldots, x_n)] = \int_{\mathbf{R}^n} g(x_1, x_2, \ldots, x_n) f(x_1, x_2, \ldots, x_n) dx_1 dx_2 \ldots dx_n$$

Estensione a n variabili casuali //covarianza

- Dato un insieme X1,X2, . . . ,Xn di variabili casuali, è possibile definire il vettore aleatorio

$$\mathbf{x} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} \qquad \mathbf{x} \in \mathbb{R}^n$$

- Anche una funzione $g : \mathbb{R}^n \mapsto \mathbb{R}^m$ è rappresentabile come vettore

$$g(\mathbf{x}) = \begin{bmatrix} g_1(\mathbf{x})) \\ g_2(\mathbf{x})) \\ \vdots \\ g_m(\mathbf{x}) \end{bmatrix} \qquad \mathbf{x} \in \mathbb{R}^n$$

Estensione a n variabili casuali //covarianza

- Dato il vettore $\mathbf{x} \in \mathbb{R}^n$ la sua matrice di covarianza è la matrice n x n tale che, per ogni coppia i, j

$$\Sigma_{i,j} = Cov[X_i, X_j].$$

- ovvero:

$$\begin{aligned}
\mathbf{\Sigma} &= \begin{bmatrix} Cov[X_1, X_1] & \cdots & Cov[X_1, X_n] \\ \vdots & \ddots & \vdots \\ Cov[X_n, X_1] & \cdots & Cov[X_n, X_n] \end{bmatrix} \\
&= \begin{bmatrix} Var[X_1] & \cdots & Cov[X_1, X_n] \\ \vdots & \ddots & \vdots \\ Cov[X_n, X_1] & \cdots & Var[X_n] \end{bmatrix} \\
&= \begin{bmatrix} E[X_1^2] - E[X_1]^2 & \cdots & E[X_1 X_n] - E[X_1]E[X_n] \\ \vdots & \ddots & \vdots \\ E[X_n X_1] - E[X_n]E[X_1] & \cdots & E[X_n^2] - E[X_n]E[X_n] \end{bmatrix} \\
&= E[\mathbf{x}\mathbf{x}^T] - E[\mathbf{x}]E[\mathbf{x}]^T \\
&= E[(\mathbf{x} - E[\mathbf{x}])(\mathbf{x} - E[\mathbf{x}])^T]
\end{aligned}$$

Estensione a n variabili casuali //covarianza

- $\boldsymbol{\Sigma}$ è matrice simmetrica:
  - per definizione di covarianza $Cov[X_i, X_j] = Cov[X_j, X_i]$
- $\boldsymbol{\Sigma}$ è definita positiva:
  - per qualunque vettore **z:**

$$\begin{aligned}
\mathbf{z}^T\boldsymbol{\Sigma}\mathbf{z} &= \sum_{i=1}^{n}\sum_{j=1}^{n} \Sigma_{ij} z_i z_j \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} Cov[X_i, X_j] z_i z_j \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} E[(X_i - E[X_i])(X_j - E[X_j])] z_i z_j \\
&= E\left[\sum_{i=1}^{n}\sum_{j=1}^{n} (X_i - E[X_i])(X_j - E[X_j]) z_i z_j\right]
\end{aligned}$$

  - ma

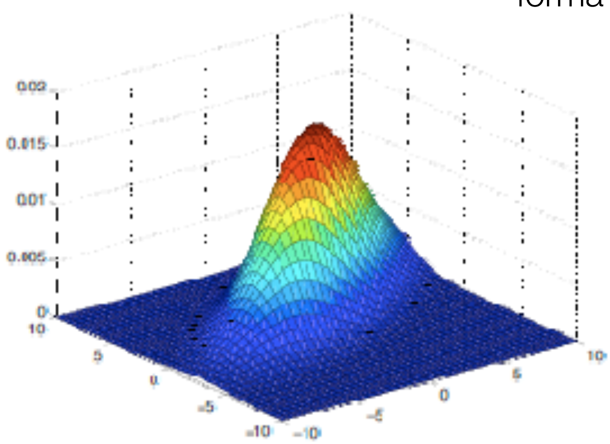
$$\sum_{i=1}^{n}\sum_{j=1}^{n} (X_i - E[X_i])(X_j - E[X_j]) z_i z_j = (\overline{\mathbf{x}}^T\mathbf{z})^2 \geq 0$$

Estensione a n variabili casuali //il caso Gaussiano

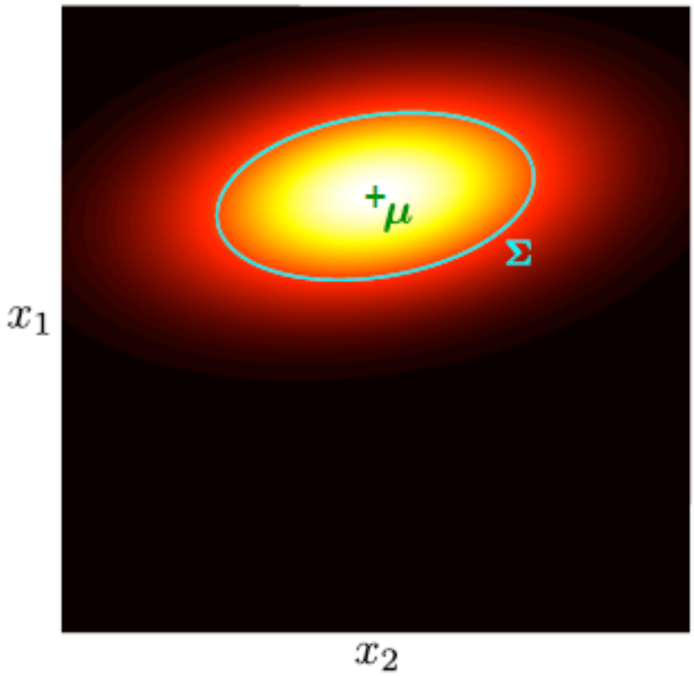
- Un vettore aleatorio $\mathbf{x} \in \mathbb{R}^n$ ha una distribuzione normale di media $\boldsymbol{\mu} \in \mathbb{R}^n$ e matrice di covarianza $\boldsymbol{\Sigma}$ di dimensione n x n

$$f(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{n/2}|\boldsymbol{\Sigma}|^{1/2}} exp\left(-\frac{1}{2}\boxed{(\mathbf{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})}\right)$$

forma quadratica di **x**

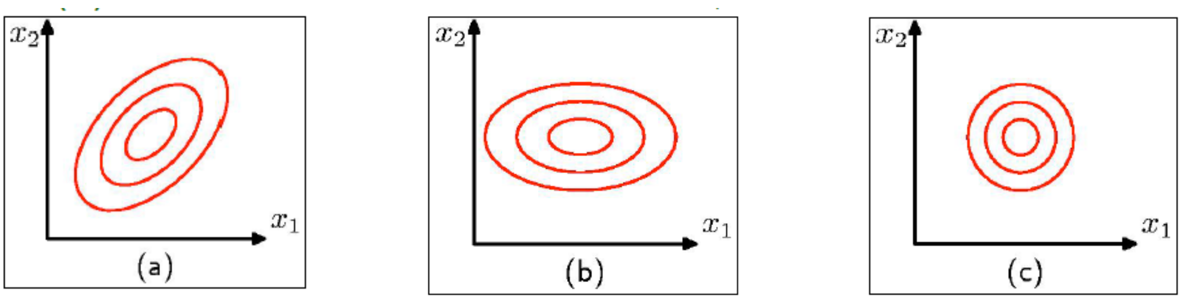


Estensione a n variabili casuali
//il caso Gaussiano



Estensione a n variabili casuali
//il caso Gaussiano

- I contorni di densità costante sono determinati dalla matrice di covarianza
  - Per esempio in 2D

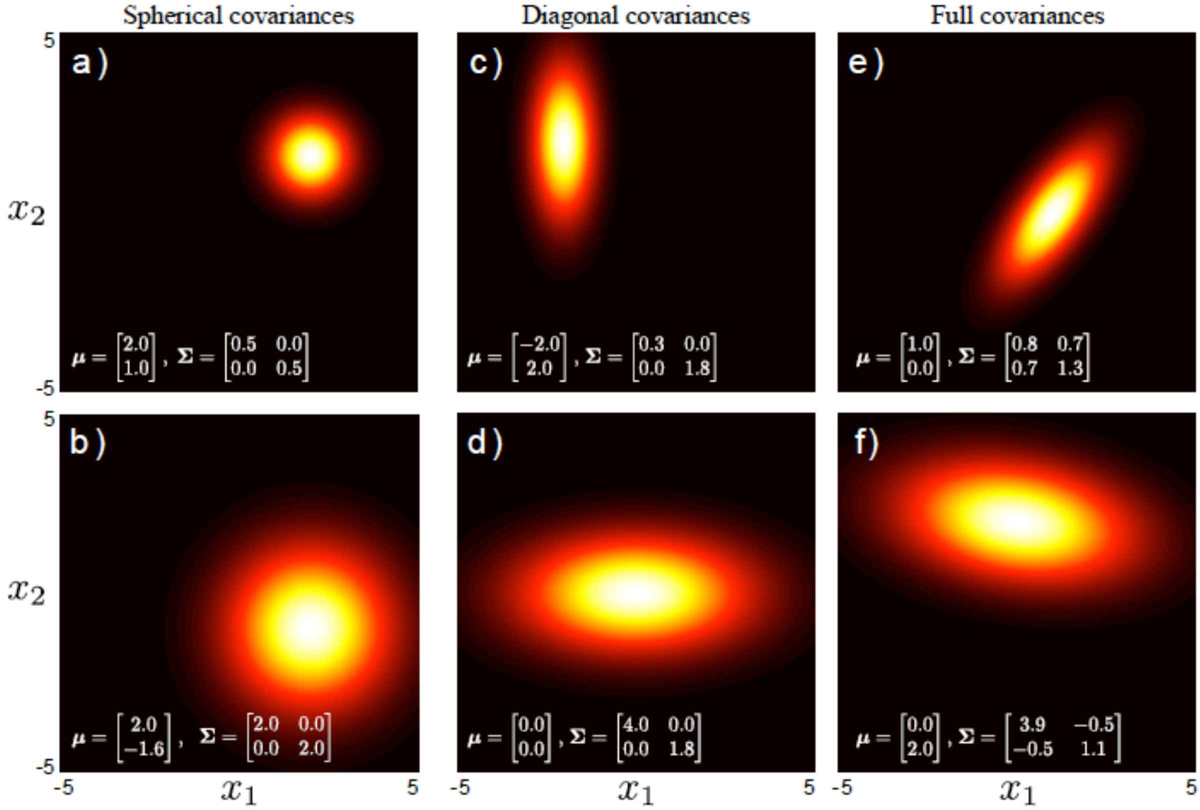


Forma generale

Forma diagonale
(allineata con gli assi)

Forma isotropa
(matrice identità)

Estensione a n variabili casuali
//il caso Gaussiano



Distribuzione Normale multivariata:
//il caso multivariato

- Consideriamo il caso della matrice di covarianza diagonale

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \qquad \boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{bmatrix} \qquad \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_n^2 \end{bmatrix}$$

- allora

$$|\boldsymbol{\Sigma}| = \sigma_1^2 \sigma_n^2 \ldots \sigma_n^2 \qquad \boldsymbol{\Sigma}^{-1} = \begin{bmatrix} \frac{1}{\sigma_1^2} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sigma_2^2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sigma_n^2} \end{bmatrix}$$

- ovvero

$$(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) = \sum_{i=1}^{n} \frac{(x_i - \mu_i)^2}{\sigma_i^2}$$

$$f(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{1}{2}\frac{(x_i - \mu_i)^2}{\sigma_i^2}\right)$$

Distribuzione multivariata = prodotto di n distribuzioni gaussiane univariate, ognuna relativa ad una coordinata xi

Distribuzione Normale multivariata:
//il caso multivariato: geometria della Gaussiana

- $\boldsymbol{\Sigma}$ soddisfa l'equazione agli autovettori $\Sigma u_i = \lambda_i u_i$
- $\boldsymbol{\Sigma}$ reale e simmetrica: i suoi autovalori i sono tutti reali e possiamo scegliere un insieme di autovettori ortonormali $\mathbf{u}_i^T \mathbf{u}_j = 1$ se $i = j$

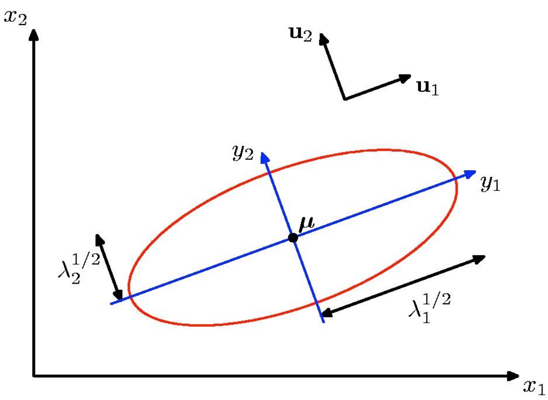
$$\boldsymbol{\Sigma} = \sum_{i=1}^{n} \lambda_i \mathbf{u}_i \mathbf{u}_i^T \qquad \boldsymbol{\Sigma}^{-1} = \sum_{i=1}^{n} \frac{1}{\lambda_i} \mathbf{u}_i \mathbf{u}_i^T$$

- Sostituendo all'interno della forma quadratica

$$(\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu}) = \sum_{i=1}^{n} \frac{\left(\mathbf{u}_i^T (\mathbf{x}-\boldsymbol{\mu})\right)^2}{\lambda_i}$$

- Indicando con $y_i = \mathbf{u}_i^T (\mathbf{x}-\boldsymbol{\mu})$ la proiezione di

  $\mathbf{x}-\boldsymbol{\mu}$ lungo la direzione di $\mathbf{u}_i$



$$(\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu}) = \sum_{i=1}^{n} \frac{y_i^2}{\lambda_i} \quad \Rightarrow \quad f(\mathbf{x};\boldsymbol{\mu},\boldsymbol{\Sigma}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\lambda_i}} \exp\left(-\frac{1}{2}\frac{y_i^2}{\lambda_i}\right)$$

Distribuzione Normale multivariata:
//medie e covarianze

- Valgono le seguenti come nel caso univariato

$$\int_{\mathbf{R}^n} f(\mathbf{x}|\boldsymbol{\mu},\boldsymbol{\Sigma}) d\mathbf{x} = 1$$

$$\int_{\mathbf{R}^n} x_i f(\mathbf{x}|\boldsymbol{\mu},\boldsymbol{\Sigma}) d\mathbf{x} = \mu_i$$

$$\int_{\mathbf{R}^n} (x_i-\mu_i)(x_j-\mu_j) f(\mathbf{x}|\boldsymbol{\mu},\boldsymbol{\Sigma}) d\mathbf{x} = \Sigma_{ij}$$

Distribuzione Normale multivariata: //somma di variabili gaussiane indipendenti

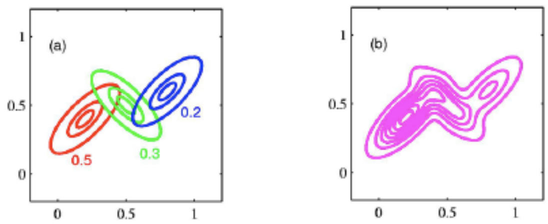
- Siano $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$

$$\mathbf{x} \sim Normal(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$
$$\mathbf{y} \sim Normal(\boldsymbol{\mu}', \boldsymbol{\Sigma}')$$

- La somma $\mathbf{z} = \mathbf{x} + \mathbf{y}$ è ancora distribuita normalmente

$$\mathbf{z} \sim Normal(\overline{\boldsymbol{\mu}}, \overline{\boldsymbol{\Sigma}}) \qquad \overline{\boldsymbol{\mu}} = \boldsymbol{\mu} + \boldsymbol{\mu}' \qquad \overline{\boldsymbol{\Sigma}} = \boldsymbol{\Sigma} + \boldsymbol{\Sigma}'$$

- Si noti che la distribuzione di z = x + y è cosa diversa dalla somma delle distribuzioni di x e y (che in generale non è gaussiana).



Distribuzione Normale multivariata: //congiunta, marginale e condizionata di v.a.g.

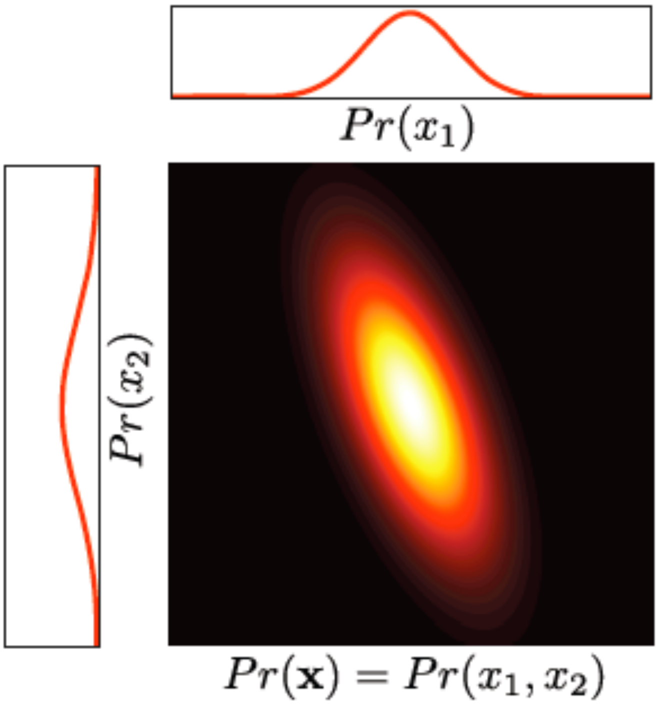
- Sia $\mathbf{x} \sim Normal(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ con

1. $\mathbf{x}_A \in \mathbb{R}^m$ e $\mathbf{x}_B \in \mathbb{R}^{n-m}$ tali che
$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_A \\ \mathbf{x}_B \end{bmatrix}$$
2. $\boldsymbol{\mu}_A \in \mathbb{R}^m$ e $\boldsymbol{\mu}_B \in \mathbb{R}^{n-m}$ tali che
$$\boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_A \\ \boldsymbol{\mu}_B \end{bmatrix}$$
3. $\boldsymbol{\Sigma}_{AA} \in \mathbb{R}^{m \times m}$, $\boldsymbol{\Sigma}_{AB} \in \mathbb{R}^{m \times (n-m)}$, $\boldsymbol{\Sigma}_{BA} \in \mathbb{R}^{(n-m) \times m}$ e $\boldsymbol{\Sigma}_{BB} \in \mathbb{R}^{(n-m) \times (n-m)}$ tali che
$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{AA} & \boldsymbol{\Sigma}_{AB} \\ \boldsymbol{\Sigma}_{BA} & \boldsymbol{\Sigma}_{BB} \end{bmatrix}$$

- La distribuzione di x è la congiunta

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_A \\ \mathbf{x}_B \end{bmatrix} \sim Normal\left(\begin{bmatrix} \boldsymbol{\mu}_A \\ \boldsymbol{\mu}_B \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Sigma}_{AA} & \boldsymbol{\Sigma}_{AB} \\ \boldsymbol{\Sigma}_{BA} & \boldsymbol{\Sigma}_{BB} \end{bmatrix} \right)$$

Distribuzione Normale multivariata:
//congiunta, marginale e condizionata di v.a.g.



Distribuzione Normale multivariata:
//congiunta, marginale e condizionata di v.a.g.

- La distribuzione di x è la congiunta

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_A \\ \mathbf{x}_B \end{bmatrix} \sim Normal\left(\begin{bmatrix} \boldsymbol{\mu}_A \\ \boldsymbol{\mu}_B \end{bmatrix}, \begin{bmatrix} \Sigma_{AA} & \Sigma_{AB} \\ \Sigma_{BA} & \Sigma_{BB} \end{bmatrix} \right)$$

- Marginali:

$$p(\mathbf{x}_A) = \int_{\mathbf{x}_B \in \mathbf{R}^{n-m}} p(\mathbf{x}|\boldsymbol{\mu}, \Sigma) d\mathbf{x}_B = \int_{\mathbf{x}_B \in \mathbf{R}^{n-m}} p(\mathbf{x}_A, \mathbf{x}_B|\boldsymbol{\mu}, \Sigma) d\mathbf{x}_B \qquad \mathbf{x}_A \sim Normal(\boldsymbol{\mu}_A, \Sigma_{AA})$$

$$p(\mathbf{x}_B) = \int_{\mathbf{x}_A \in \mathbf{R}^{m}} p(\mathbf{x}|\boldsymbol{\mu}, \Sigma) d\mathbf{x}_A = \int_{\mathbf{x}_A \in \mathbf{R}^{m}} p(\mathbf{x}_A, \mathbf{x}_B|\boldsymbol{\mu}, \Sigma) d\mathbf{x}_A \qquad \mathbf{x}_B \sim Normal(\boldsymbol{\mu}_B, \Sigma_{BB})$$

- Condizionate

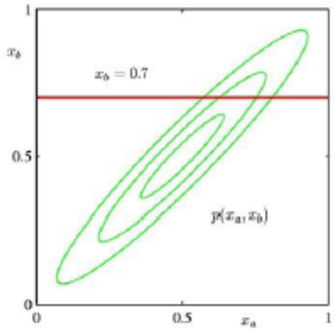
$$p(\mathbf{x}_A|x_B) = \frac{p(\mathbf{x}_A, \mathbf{x}_B|\boldsymbol{\mu}, \Sigma)}{\int_{\mathbf{x}_A \in \mathbf{R}^m} p(\mathbf{x}|\boldsymbol{\mu}, \Sigma) d\mathbf{x}_A}$$

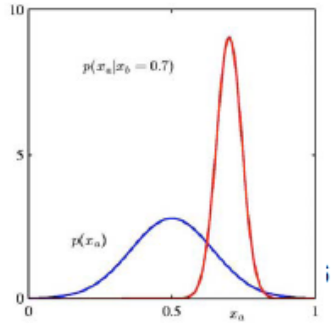
$$\mathbf{x}_A|\mathbf{x}_B \sim Normal(\boldsymbol{\mu}_A + \Sigma_{AB}\Sigma_{BB}^{-1}(\mathbf{x}_B - \boldsymbol{\mu}_B), \Sigma_{AA} - \Sigma_{AB}\Sigma_{BB}^{-1}\Sigma_{BA})$$

$$p(\mathbf{x}_B|x_A) = \frac{p(\mathbf{x}_A, \mathbf{x}_B|\boldsymbol{\mu}, \Sigma)}{\int_{\mathbf{x}_B \in \mathbf{R}^{n-m}} p(\mathbf{x}|\boldsymbol{\mu}, \Sigma) d\mathbf{x}_B}$$

$$\mathbf{x}_B|\mathbf{x}_A \sim Normal(\boldsymbol{\mu}_B + \Sigma_{BA}\Sigma_{AA}^{-1}(\mathbf{x}_A - \boldsymbol{\mu}_A), \Sigma_{BB} - \Sigma_{BA}\Sigma_{AA}^{-1}\Sigma_{AB})$$

Distribuzione Normale multivariata: //congiunta, marginale e condizionata di v.a.g.





- La distribuzione di x è la congiunta

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_A \\ \mathbf{x}_B \end{bmatrix} \sim Normal\left(\begin{bmatrix} \boldsymbol{\mu}_A \\ \boldsymbol{\mu}_B \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Sigma}_{AA} & \boldsymbol{\Sigma}_{AB} \\ \boldsymbol{\Sigma}_{BA} & \boldsymbol{\Sigma}_{BB} \end{bmatrix} \right)$$

- Marginali:

$$\mathbf{x}_A \sim Normal(\boldsymbol{\mu}_A, \boldsymbol{\Sigma}_{AA})$$

$$\mathbf{x}_B \sim Normal(\boldsymbol{\mu}_B, \boldsymbol{\Sigma}_{BB})$$

- Condizionate

$$\mathbf{x}_A|\mathbf{x}_B \sim Normal(\boldsymbol{\mu}_A + \boldsymbol{\Sigma}_{AB}\boldsymbol{\Sigma}_{BB}^{-1}(\mathbf{x}_B - \boldsymbol{\mu}_B), \boldsymbol{\Sigma}_{AA} - \boldsymbol{\Sigma}_{AB}\boldsymbol{\Sigma}_{BB}^{-1}\boldsymbol{\Sigma}_{BA})$$

$$\mathbf{x}_B|\mathbf{x}_A \sim Normal(\boldsymbol{\mu}_B + \boldsymbol{\Sigma}_{BA}\boldsymbol{\Sigma}_{AA}^{-1}(\mathbf{x}_A - \boldsymbol{\mu}_A), \boldsymbol{\Sigma}_{BB} - \boldsymbol{\Sigma}_{BA}\boldsymbol{\Sigma}_{AA}^{-1}\boldsymbol{\Sigma}_{AB})$$

Distribuzione Normale multivariata: //formula di Bayes

- Siano

$$\mathbf{x} \sim Normal(\boldsymbol{\mu}, \boldsymbol{\Sigma}_1)$$

$$\mathbf{y}|\mathbf{x} \sim Normal(\mathbf{A}\mathbf{x} + \mathbf{b}, \boldsymbol{\Sigma}_2)$$

- La probabilità a posteriori vale:

$$\mathbf{x}|\mathbf{y} \sim Normal(\overline{\boldsymbol{\mu}}, \overline{\boldsymbol{\Sigma}}) \qquad \begin{aligned} \overline{\boldsymbol{\mu}} &= (\boldsymbol{\Sigma}_1^{-1} + \mathbf{A}^T\boldsymbol{\Sigma}_2^{-1}\mathbf{A})^{-1}(\mathbf{A}^T\boldsymbol{\Sigma}_2^{-1}(\mathbf{y} - \mathbf{b}) + \boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}) \\ \overline{\boldsymbol{\Sigma}} &= (\boldsymbol{\Sigma}_1^{-1} + \mathbf{A}^T\boldsymbol{\Sigma}_2^{-1}\mathbf{A})^{-1} \end{aligned}$$

- Inoltre:

$$\mathbf{y} \sim Normal(\overline{\overline{\boldsymbol{\mu}}}, \overline{\overline{\boldsymbol{\Sigma}}}) \qquad \begin{aligned} \overline{\overline{\boldsymbol{\mu}}} &= \mathbf{A}\boldsymbol{\mu} + \mathbf{b} \\ \overline{\overline{\boldsymbol{\Sigma}}} &= \boldsymbol{\Sigma}_2 + \mathbf{A}\boldsymbol{\Sigma}_1\mathbf{A}^T \end{aligned}$$