IDENTIFICAZIONE SISTEMI INCERTI

RICHIAMI VARIABILI ALEATORIE



Definiamo la variabile aleatoria vettoriale

$$X = [X_1 \dots X_n]^T$$

una variabile descritta univocamente dalla densità di probabilità (PDF), cioè una funzione

$$f_X : \mathbb{R}^n \to \mathbb{R} \geq 0 \quad t.c. \quad \forall \$ \subset \mathbb{R}^n \to P(X \in \$) = \int_{\$} f_X(x)dx$$

Deve essere inoltre rispettata la condizione di normalizzazione

$$P(X \in \mathbb{R}^n) = \int_{\mathbb{R}^n} f_X(x)\,dx = \int f_X(x)\,dx = 1$$

A livello di notazione scriviamo

$$X \sim f_X(\cdot)$$

Alternativamente utilizziamo la distribuzione cumulativa (CDF), definita formalmente come

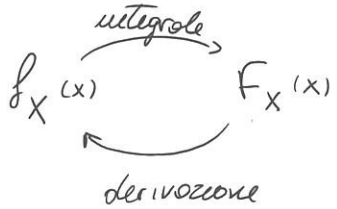
$$F_X : \mathbb{R}^n \to \mathbb{R} \geq 0$$



$$F_X(x) = P(X < x) = P(X_1 < x_1, \dots X_n < x_n) =$$

$$= \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \dots \int_{-\infty}^{x_n} f_X(x)\,dx_1 \dots dx_n$$

Abbiamo quindi due modi equivalenti per definire completamente le variabili aleatorie



Vediamo ora alcuni operatori

- valore atteso $\mathbb{E}[\cdot]$, applicabile alla generica funzione $g(X)$ della v.a.

$$X \sim f_X(x)$$

$$\mathbb{E}[g(X)] = \int g(x) f_X(x)\, dx$$

- Proprietà: linearità
- Ci serve per definire i momenti di una v.a.:
  1) I ordine (media) → dato $g(X) = X \rightarrow m_X = \mathbb{E}[X] = \int x f_X(x)\, dx$

  2) II ordine (matrice di covarianza) → $[g(X) = (X - m_X)(X - m_X)^T]$
  Definiamo quindi la matrice di covarianza come
  $\Sigma_X \in \mathbb{R}^{n \times n} : \Sigma_X = \mathbb{E}[(X - m_X)(X - m_X)^T] =$

$$= \int (x - m_X)(x - m_X)^T f_X(x)\, dx$$

- È bene osservare che gli elementi di m_X sono le medie degli elementi di X

$$m_i = \mathbb{E}[X_i]$$

- Gli elementi di Σ_X sono invece:
  1) sulla diagonale abbiamo la varianza di X_i

$$\Sigma_{ii} = \mathrm{Var}(X_i) = \mathbb{E}[(x_i - m_i)^2] \quad \forall i = 1, \dots, n$$

  2) fuori diagonale abbiamo invece la covarianza di X_i X_j

$$\Sigma_{ij} = \mathrm{cov}(X_i, X_j) = \mathbb{E}[(X_i - m_i)(X_j - m_j)]$$

  3) Vale che $\Sigma_{ij} = \Sigma_{ji} \rightarrow \Sigma_X = \Sigma_X^T$

- Σ_X è semi-definita positiva → $v^T \Sigma_X v \geq 0 \rightarrow \Sigma_X \Sigma_X^T \geq 0 \rightarrow$ tutti gli autovalori sono reali non negativi → $\lambda_i \in \mathbb{R} \geq 0 \quad \forall i$

Supponendo di definire

$$X \in \mathbb{R}^n \quad Y \in \mathbb{R}^p$$

e una PDF congiunta

$$f_{XY}(x,y)$$

Se andiamo a definire

$$Z = \begin{bmatrix} X^T & Y^T \end{bmatrix} \text{ t.c } \forall \$ \subset \mathbb{R}^{n+p} \quad \boxed{P(Z \in \$) = \int_{\$} f_{XY}(x,y)\,dx\,dy}$$

variabile congiunta

Possiamo ottenere le marginali

$$f_X(x) = \int_{\mathbb{R}^p} f_{XY}(x,y)\,dy$$

$$f_Y(y) = \int_{\mathbb{R}^n} f_{XY}(x,y)\,dx$$

Definiamo ora i momenti mercanti

- matrice di cross-covarianza

$$\Sigma_{XY} = \mathbb{E}\left[(X - m_X)(Y - m_Y)\right] \in \mathbb{R}^{n \times p}$$

e vale $\Sigma_{XY} = \Sigma_{YX}^T$. Inoltre se $\Sigma_{XY} = \emptyset$ le due v.a. si dicono non correlate

Parliamo di distribuzione condizionata (cioè distribuzione di X condizionata al fatto che Y abbia assunto il valore y) e la indichiamo con:

$f_{X|Y}(x|y)$ è quindi una funzione di x in quanto y non è una variabile

e rispetto ad X deve valere la condizione di normalizzazione

$$\int f_{X|Y}(x|y)\,dx = 1$$

È bene notare che questo distribuzione è legata alla congiunta da una relazione di proporzionalità

$$f_{X|Y}(x|y) \propto f_{XY}(x,y)$$

Se integro ambo i membri rispetto a x ottengo

$$\int f_{X|Y}(x|y)\, dx = 1$$

$$\int f_{XY}(x,y)\, dx = f_Y(y)$$



Da cui possiamo banalmente ricavare che

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

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

Da cui ricaviamo la formula di Bayes

$$f_{X|Y}(x|y) = \frac{f_{Y|X}(y|x) f_X(x)}{\underbrace{f_Y(y)}_{\text{termine di normalizzazione}}}$$

$$= \frac{f_{Y|X}(y|x) f_X(x)}{\int f_{Y|X}(y|x) f_X(x) dx}$$

Definiamo in particolare

$f_{X|Y}(x|y)$ PDF a posteriori

$f_X(x)$ PDF a priori

$f_{Y|X}(y|x)$ funzione di verosimiglianza



Parliamo adesso di indipendenza: due v.a. X e Y sono indipendenti se la PDF congiunta la possiamo scrivere come prodotto delle marginali

$$f_{XY}(x,y) = f_X(x) f_Y(y)$$

e quindi

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

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



indipendenza $\Rightarrow$ non correlazione
~~$\Leftarrow$~~ il viceversa non è vero

Inoltre se X e Y sono v.a. indipendenti, $g(X)$ e $h(Y)$ sono anch'esse indipendenti

Funzioni di v.a.

$$X \sim f_X(x) \qquad Y = g(X)$$

supponendo g invertibile, definiamo $h(\cdot) = g^{-1}(\cdot)$
Possiamo quindi scrivere

$$f_Y(y) = f_X(h(y)) \left| \det \frac{\partial h}{\partial y}(y) \right|$$

Caso particolare di funzione di una v.a.



$$Y = h(X) + V, \qquad X \sim f_X(\cdot) \quad V \sim f_V(\cdot)$$

con X, V indipendenti

Questo che abbiamo scritto è sostanzialmente il modello di un sensore in cui V è un rumore additivo. Andiamo a scrivere

$$f_{Y|X}(y|x) = f_V(y - h(x))$$

$$f_{X,Y}(x,y) = f_V(y - h(x)) \cdot f_X$$

$$f_Y(y) = \int f_V(y - h(x)) \cdot f_X(x)\, dx$$