INTERPRETAZIONE GEOMETRICA DEL VALORE DI ATTESA CONDIZIONATO

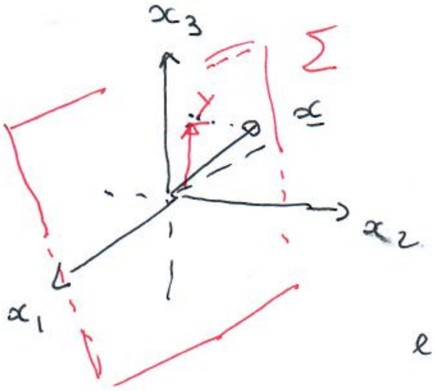
Esercizio (algebra lineare)

Sia $\mathbb{R}^3$ come spazio euclideo

$$\Sigma = \{ \underline{x} \in \mathbb{R}^3 : \quad x_1 + x_2 + x_3 = 0 \}$$ sottospazio vettoriale

Dato $\underline{x} \in \mathbb{R}^3$ trovare $\underline{y} \in \Sigma$

t.c. $d(\underline{x}, \underline{y})$ sia minima



OSS

$$\Sigma = \{ \underline{x} \in \mathbb{R}^3 : \quad \underline{x} \cdot (1,1,1) = 0 \}$$

evidentemente

~~attendente~~ $\underline{y}$ ottimale $= \Pi_\Sigma \underline{x}$ = proiezione ortogonale di $\underline{x}$ su Σ

$$\mathbb{R}^3 = \Sigma \oplus \Sigma^\perp ; \quad \{ \underline{z} \in \mathbb{R}^3 : (\underline{z}, \underline{x}) = 0 \quad \forall \underline{x} \in \mathbb{R}^3 \}$$

ovvero se $\underline{x} \in \mathbb{R}^3$ $\exists!$ $\underline{y} \in \Sigma$

t.c.

$$\underline{x} = \underline{y} + \underbrace{(\underline{x} - \underline{y})}_{\in \Sigma^\perp} \quad [(\underline{y}, \underline{x} - \underline{y}) = 0]$$

questo $\underline{y}$ è la proiezione ortogonale di $\underline{x}$ su Σ

Per calcolare $\underline{y}$ mi organizzo una base ortonormale di Σ

$\underline{e}_1 = \frac{1}{\sqrt{2}}(1\ 0\ -1)$ $\qquad$ $\underline{e}_2 = \frac{1}{\sqrt{6}}(1\ -2\ 1)$

$(\underline{e}_1, \underline{e}_2) = 0$, $\underline{e}_1, \underline{e}_2 \in \Sigma$ $\quad$ [$\underline{e}_1, \underline{e}_2$ sono $\perp$ a $(1,1,1)$]

$\underline{n} = \frac{1}{\sqrt{3}}(1,1,1)$ base per $\Sigma^\perp$

o ~~Vettor~~ $\underline{e}_1, \underline{e}_2, \underline{n}$ è base ortonormale di $\mathbb{R}^3$

$$\underline{x} = \underbrace{(\underline{e}_1, \underline{x})\,\underline{e}_1 + (\underline{e}_2, \underline{x})\,\underline{e}_2}_{\underset{\pi_\Sigma \underline{x}}{\parallel}} + \underbrace{(\underline{n}, \underline{x})\,\underline{n}}_{\in\,\Sigma^\perp}$$

(6) $(\Omega, \mathcal{F}, \mathbb{P})$ $\quad$ $|\Omega| < +\infty$ $\quad$ SPAZIO di PROBABILITÀ

~~diciam~~

$L^2(\Omega, \mathbb{P}) = \{ X : \Omega \to \mathbb{R} \} = \{$ variabili aleatorie $\}$

è spazio euclideo con il prodotto scalare

$$(X, Y) := \mathbb{E}(X \cdot Y) = \sum_{\omega \in \Omega} X(\omega) Y(\omega)\, \mathbb{P}(\{\omega\})$$

Dato la v.a. $Y \in L^2(\Omega, \mathbb{P})$

vogliamo trovare la sua migliore approssimazione (nella geometria di $L^2(\Omega, \mathbb{P})$) nei contesti seguenti

(1) Migliore approssimazione di Y con le variabili aleatorie certe (= costanti)

$$\Sigma = \{ X : \Omega \to \mathbb{R} \quad X \text{ v.a. certa} \} = \mathbb{R} \cdot 1$$

= sottospazio vettoriale di $L^2(\Omega, \mathbb{P})$

$$\| 1 \|^2 = (1, 1) = \mathbb{E}(1 \cdot 1) = 1$$

quindi

migliore approssimazione di Y $= (Y, 1)\, 1 = \mathbb{E}(Y \cdot 1)\, 1 = \mathbb{E}(Y)$

Fisso ora una v.a. $X \in L^2(\Omega, \mathbb{P})$ (Penso di aver osservato X e di voler predire Y)
X non certa

(2) Migliore approssimazione di Y tra le funzioni affini di X

$$\Sigma = \{ \alpha X + \beta, \quad \alpha, \beta \in \mathbb{R} \}$$

= spazio vettoriale generato da $\{1, X\}$

Mi organizzo una base ortonormale in Σ

$1, \ \dfrac{X - \mathbb{E}(X)}{\sqrt{\mathbb{V}(X)}}$ è ortonormale

$\hat{Y}$ = migliore approx. di Y tra le funzioni affini di X

$$= \mathbb{E}(Y) \cdot 1 + \left(Y, \frac{X - \mathbb{E}(X)}{\sqrt{\mathbb{V}(X)}} \right) \frac{X - \mathbb{E}(X)}{\sqrt{\mathbb{V}(X)}}$$

$$= \mathbb{E}(Y) + \frac{1}{\mathbb{V}(X)} \, \mathbb{E}\big(Y(X-\mathbb{E}(X))\big)\,(X-\mathbb{E}(X))$$

$$= \mathbb{E}(Y) + \mathrm{cov}(X,Y)\,\frac{X-\mathbb{E}(X)}{\mathbb{V}(X)}$$

questo ci fornisce l'interpretazione geometrica di $\mathrm{cov}(\cdot,\cdot)$

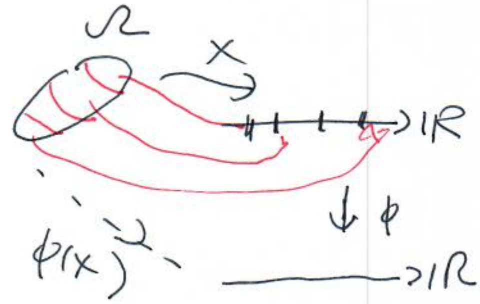
(3) Migliore approssimazione di Y tra le funzioni di X

interpretazione: misuro X e voglio predirre Y
questa è la scelta migliore

$$\Sigma = \{\, \phi(X), \quad \phi: \mathbb{R} \to \mathbb{R} \,\}$$

sottospazio vettoriale di $L^2(\Omega, \mathbb{P})$

oss. $Z \in \Sigma$ sse Z è costante sugli atomi della partizione indotta da X



$$\hat{Y} = \Pi_{\Sigma} Y = \mathbb{E}(Y|X)$$

ovvero

$\mathbb{E}(Y|X)$ è la migliore approssimazione di Y (ai sensi di $L^2(\Omega, \mathbb{P})$) tra le funzioni di X

basta verificare che

$$\mathbb{E}(Y|X) \perp Y - \mathbb{E}(Y|X)$$

$$\Updownarrow$$

$$\mathbb{E}\left(\mathbb{E}(Y|X) \cdot [Y - \mathbb{E}(Y|X)] \right) = 0$$

$$\Updownarrow \qquad \mathbb{E}(\mathbb{E}(\cdot|X)) = \mathbb{E}(\cdot)$$

$$\mathbb{E}\left(\mathbb{E}\left(\mathbb{E}(Y|X) [Y - \mathbb{E}(Y|X)] \mid X \right) \right) = 0$$

è vero

$$= \mathbb{E}(Y - \mathbb{E}(Y|X) | X) = 0$$

infatti

X, Y v.a e $\phi : \mathbb{R} \to \mathbb{R}$

$$\mathbb{E}(\phi(X) Y | X) = \phi(X) \, \mathbb{E}(Y|X)$$

ai fini di $\mathbb{E}(\cdot|X)$ è costante

per definizione

$\mathbb{E}(Y|X)$ è costante sugli atomi della partizione indotta da X

$$\mathbb{E}(\phi(X) Y | X) = \sum_{x \in Im(X)} 1\!\!1_{X=x} \, \mathbb{E}(\phi(X) Y | X = x)$$

$$= \sum_{x \in Im(X)} 1\!\!1_{X=x} \, \phi(x) \, \mathbb{E}(Y|X=x)$$

$$= \phi(X) \, \mathbb{E}(Y|X)$$

10. VARIABILI ALEATORIE CONTINUE

In molte situazioni, sia teoriche che applicate, appaiono naturalmente variabili aleatorie continue, cioè con $Im(X)$ più che numerabile.

- generatore di numeri casuali
- misure di grandezze continue
- ...

Evidentemente se $x \in \mathbb{R}$ $\qquad \mathbb{P}(X = x) = 0$

ma per identificare la distribuzione di X osserviamo che

$$\mathbb{P}(X \in [x, x + \Delta x]) \propto \Delta x$$

proporzionale a

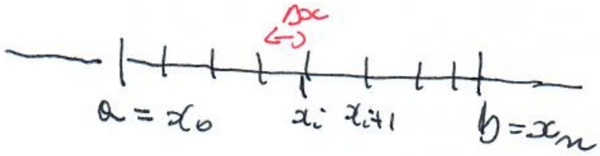
Introduciamo quindi la densità di probabilità

$f_X : \mathbb{R} \to [0, +\infty)$ per cui

$$\mathbb{P}(X \in [x, x + \Delta x]) = f_X(x) \Delta x + o(\Delta x)$$

infinitesimo di ordine superiore

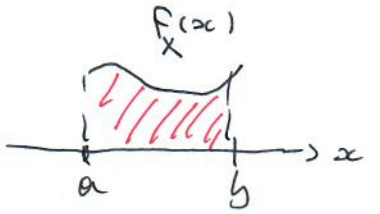
In tal modo se $a < b$



$$\mathbb{P}(a \le X < b) \simeq \sum_{i=0}^{n-1} \mathbb{P}(x_i \le X < x_{i+1}) =$$

$$= \sum_{i=0}^{n-1} \left[f_X(x_i)\,(x_{i+1}-x_i) + o(x_{i+1}-x_i) \right]$$

$$\xrightarrow[n\to\infty]{} \int_a^b f_X(x)\,dx$$

$\mathbb{P}(a \leq X < b) = \text{Area}\Big($ 

$\Big)$

$$= \int_a^b f_X(x)\,dx$$

La distribuzione di una v.a. continua è univocamente caratterizzata dalla densità di probabilità f_X

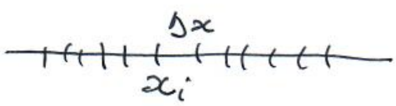
$f_X : \mathbb{R} \to [0,+\infty)$ (può capitare $f_X(x) > 1$ per qualche x, non è una probabilità)

con la condizione di normalizzazione

$1 = \mathbb{P}(-\infty < X < +\infty) = \int_{-\infty}^{+\infty} f_X(x)\,dx = \text{Area}\Big($ 

$\Big)$

- Valore di attesa di v.a. continue



$$\mathbb{E}(X) \simeq \sum_i x_i\, \mathbb{P}(x_i \leq X < x_i + \Delta x)$$

$$= \sum_i x_i \left[f_X(x_i)\,\Delta x + o(\Delta x) \right] \longrightarrow \int_{-\infty}^{+\infty} x\, f_X(x)\,dx$$

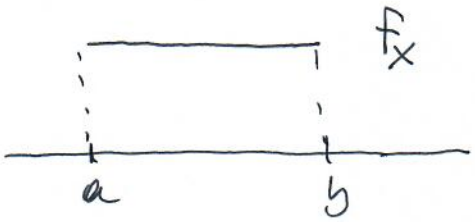
• Variabile aleatoria uniforme

Siano ~~dati~~ $a < b$ $\qquad X \sim \mathrm{UNIF}(a,b)$

se ogni punto di (a,b) ha la stessa probabilità

(attenzione vuole: prob $= 0$ comunque)

Nel senso che la densità di probabilità è costante in (a,b)



imponendo la normalizzazione

$$f_X = \begin{cases} \dfrac{1}{b-a} & x \in (a,b) \\ 0 & x \notin (a,b) \end{cases}$$

$$\mathbb{E}(X) = \text{punto medio} = \frac{a+b}{2} \quad \text{per simmetria}$$

verifico

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} x\, f_X(x)\, dx = \frac{1}{b-a}\int_a^b x \, dx$$

$$= \frac{1}{2}\,\frac{1}{b-a}\, x^2 \Big|_a^b = \frac{1}{2}\,\frac{b^2-a^2}{b-a} = \frac{1}{2}(a+b)$$

$$\mathbb{V}(X) = \mathbb{E}\left([X-\mathbb{E}(X)]^2\right) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$$

$$\mathbb{E}(X^2) = \int_{-\infty}^{+\infty} x^2 f_X(x)\, dx = \frac{1}{b-a}\int_a^b x^2 \, dx$$

$$= \frac{1}{3}\,\frac{1}{b-a}\, x^3 \Big|_a^b = \frac{1}{3}\,\frac{b^3-a^3}{b-a} = \frac{1}{3}(b^2+ab+a^2)$$

quindi

$$\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \frac{1}{3}(b^2+ab+a^2) - \left(\frac{a+b}{2}\right)^2$$

$$= \frac{1}{3}(b^2+ab+a^2) - \frac{1}{4}(a^2+b^2+2ab)$$

$$= \frac{1}{12}(a^2+b^2-2ab) = \frac{1}{12}(b-a)^2$$

$(b-a)^2$ si poteva ottenere per "omogeneità"

se X si misura in metri a,b si misurano in metri

f_X " " " $\frac{1}{\text{metri}}$ (in modo che la probabilità sia indipendente dall'unità di misura

X^2 " " " $(\text{metri})^2$

Inoltre per invarianza per traslazioni

$(\mathbb{V}(X) = \mathbb{V}(X+c)$ se $c \in \mathbb{R})$ necessariamente $b-a$

$\frac{1}{12}$ dipende invece dal calcolo

- V.A. INDIPENDENTI. Una delle definizioni nel caso discreto è adatta anche al caso di v.a. continue

X e Y sono v.a. indipendenti sse

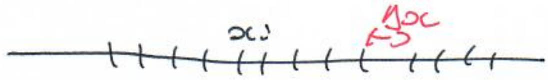
$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\,\mathbb{P}(Y \in B) \quad \forall A, B \subset \mathbb{R}$$

- Somma di v.a. continue indipendenti

Se X, Y indipendenti e $Z = X+Y$

$$\mathbb{P}(Z=z) = \sum_{x} \mathbb{P}(X=x)\, \mathbb{P}(Y=z-x)$$

Vogliamo ricavare la formula analoga per v.a. continue. Trovare cioè la densità di Z in termini delle densità di X e Y.



$$\mathbb{P}(Z \in [z, z+\Delta z])$$

$$= \sum_{i} \mathbb{P}(X \in [x_i, x_{i+1}),\ X+Y \in [z, z+\Delta z))$$

$$\simeq \sum_{i} \mathbb{P}(X_i \in [x_i, x_i+\Delta x),\ Y \in [z-x_i, z-x_i+\Delta z))$$

indip

$$= \sum_{i} \left(f_X(x_i)\,\Delta x + o(\Delta x)\right)\left(f_Y(z-x_i)\,\Delta z + o(\Delta z)\right)$$

$$= \Delta z \int_{-\infty}^{+\infty} f_X(x)\, f_Y(z-x)\, dx + o(\Delta z)$$

ovvero

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(x)\, f_Y(z-x)\, dx$$

Salvo che ci sono le densità invece delle probabilità e integrale invece di somma è la stessa formula del caso discreto.

- Esercizio.

Sia X v.a continua con densità f_X

$\phi: \mathbb{R} \to \mathbb{R}$ $\nearrow$ e $Y = \phi(X)$ $\qquad f_Y = ?$

- I^a soluzione: usare le regole note per il cambio di variabili negli integrali (per voi)

- II^a soluzione

$$\mathbb{P}(Y \in [y, y+\Delta y]) = \mathbb{P}(\phi(X) \in [y, y+\Delta y])$$

ϕ è invertibile

$$= \mathbb{P}(X \in [\phi^{-1}(y), \underbrace{\phi^{-1}(y+\Delta y)}_{= \phi^{-1}(y) + (\phi^{-1})'(y)\Delta y + o(\Delta y)}])$$

$$= f_X(\phi^{-1}(y))\,(\phi^{-1})'(y)\,\Delta y + o(\Delta y)$$

da cui

$$f_Y(y) = f_X(\phi^{-1}(y)) \frac{1}{\phi'(\phi^{-1}(y))}$$

OSS. se ϕ non è monotona bisogna ricondursi a intervalli in cui lo è

EX X v.a. con densità f_X

$Y = X^2$ v.a. con densità $f_Y = ?$