Siano X e Y due v.a. a valori discreti in $\mathbb{R}$. Per ogni $a, b \in \mathbb{R}$ consideriamo la v.a. $Y - aX - b$. Il momento di ordine due della variabile aleatoria $Y - aX - b$ è

$$E[(Y - aX - b)^2].$$

Si tratta di una funzione che dipende da a e b, cioè

$$\mathbb{R}^2 \ni (a, b) \longmapsto E[(Y - aX - b)^2] \in \mathbb{R}$$

è una funzione da $\mathbb{R}^2$ in $\mathbb{R}$.

<u>Il problema che vogliamo risolvere è il seguente:</u>

"determinare, se esiste, $(\bar{a}, \bar{b}) \in \mathbb{R}^2$ t.c.

$$\min_{(a,b) \in \mathbb{R}^2} E[(Y - aX - b)^2] = E[(Y - \bar{a}X - \bar{b})^2] \text{"}.$$

Se la coppia $(\bar{a}, \bar{b}) \in \mathbb{R}^2$ che soddisfa il problema precedente esiste chiameremo $Y = \bar{a}X + \bar{b}$ retta di regressione.

Si tratta di risolvere un problema di minimo libero. Lo vogliamo risolvere mettendo in evidenza la speranza matematica e la varianza rispettivamente di X e Y.

Per questa ragione scriviamo la funzione $E[(Y - aX - b)^2]$ mettendo in evidenza le precedenti grandezze.

Quindi

somma zero

$$E[(Y-aX-b)^2] = E[(Y - E[Y] - a(X-E[X]) + E[Y] - aE[X] - b)^2]$$

somma zero

$$= E\Big[(Y-E[Y])^2 + a^2(X-E[X])^2 + B^2 - 2(Y-E[Y])a(X-E[X]) + 2(Y-E[Y])B - 2a(X-E[X])B\Big],$$

dove $B = E[Y] - a[X] - b$; quindi per la linearità di E risulta:

$$E[(Y-aX-b)^2] = E[(Y-E[Y])^2] + E[a^2(X-E[X])^2] + E[B^2] - 2E[(Y-E[Y])a(X-E[X])] + 2E[(Y-E[Y])B] - 2E[a(X-E[X])B]$$

$$= Var(Y) + a^2 Var(X) + B^2 - 2a E[(Y-E[Y])(X-E[X])] + 2B E[(Y-E[Y])] - 2aB E[(X-E[X])]$$

$$= Var(Y) + a^2 Var(X) + B^2 - 2a\, Cov(X,Y);$$

perché $E[(Y-E[Y])(X-E[X])] = E[XY - E[Y]X - E[X]Y + E[X]E[Y]]$

$$= E[XY] - E[Y]E[X] - \cancel{E[X]E[Y]} + \cancel{E[X]E[Y]}$$

$$= E[XY] - E[X]E[Y] \overset{df}{=} Cov(X,Y),$$ inoltre

$$E[(Y-E[Y])] = E[Y] - E[Y] = 0 \quad e \quad E[X-E[X]] = E[X] - E[X] = 0$$

Troviamo i punti critici della funzione

$$(a,b) \xmapsto{f} \underbrace{Var(Y)}_{\text{numero}} + a^2 \underbrace{Var(X)}_{\text{numero}} + (\underbrace{E[Y]}_{\text{numero}} - aE[X] - b)^2 - 2a \underbrace{Cov(X,Y)}_{\text{numero}}$$

$\nabla f(a,b) = \left(\frac{\partial f}{\partial a}(a,b), \frac{\partial f}{\partial b}(a,b)\right)$. I punti critici sono soluzioni di

$$\begin{cases} \frac{\partial f}{\partial a}(a,b) = 0 \\ \frac{\partial f}{\partial b}(a,b) = 0 \end{cases} \Longleftrightarrow \begin{cases} 2a\,Var(X) - 2E[X](E[Y] - aE[X] - b) - 2Cov(X,Y) = 0 \\ -2(E[Y] - aE[X] - b) = 0 \end{cases}$$

$$\Longleftrightarrow \begin{cases} 2a\,Var(X) - 2Cov(X,Y) = 0 \\ b = E[Y] - aE[X] \end{cases} \Longleftrightarrow \begin{cases} a = \frac{Cov(X,Y)}{Var(X)} \\ b = E[Y] - \frac{Cov(X,Y)}{Var(X)} E[X] \end{cases}$$

In questo caso occorre supporre $Var(X) \neq 0$.

Calcoliamo anche la matrice Hessiana di f

$$Hf(a,b) = \begin{bmatrix} \frac{\partial^2 f(a,b)}{\partial a^2}, & \frac{\partial^2 f(a,b)}{\partial a \partial b} \\ \frac{\partial^2 f(a,b)}{\partial a \partial b}, & \frac{\partial^2 f(a,b)}{\partial b^2} \end{bmatrix} = \begin{bmatrix} 2Var(X) + 2(E[X])^2, & 2E[X] \\ 2E[X], & 2 \end{bmatrix}$$

Calcoliamo gli autovalori

$$\det(Hf(a,b) - \lambda I) = 0 \Longleftrightarrow \det\begin{pmatrix} 2Var(X) + 2(E[X])^2 - \lambda, & 2E[X] \\ 2E[X], & 2-\lambda \end{pmatrix} = 0$$

$$\Longleftrightarrow (2Var(X) + 2(E[X])^2 - \lambda)(2-\lambda) - 4(E[X])^2 = 0 \Longleftrightarrow$$

$$4Var(X) + \cancel{4(E[X])^2} - 2\lambda - 2\lambda Var(X) - 2\lambda(E[X])^2 + \lambda^2 - \cancel{4(E[X])^2} = 0 \Longleftrightarrow$$

$$\lambda^2 - 2(1 + Var(X) + E[X]^2)\lambda + 4Var(X) = 0$$

$$\frac{\Delta}{4} = (1 + Var(X) + E[X]^2)^2 - 4Var(X)$$
$$= 1 + Var(X)^2 + E[X]^4 + \underline{2Var(X)} + 2E[X]^2 + 2Var(X)E[X]^2 - \underline{4Var(X)}$$

$$= 1 + Var(X)^2 + E[X]^4 - 2Var(X) + 2E[X]^2 + 2Var(X)E[X]^2$$

ho effettuato una minorazione.

$$\geq \quad 1 + Var(X)^2 + E[X]^4 - 2Var(X) - 2E[X]^2 + 2Var(X)E[X]^2$$

$$= \left(1 - Var(X) - E[X]^2\right)^2 \geq 0.$$ Quindi le radici sono entrambe reali. Inoltre,

$$\lambda_1 = 1 + Var(X) + E[X]^2 + \sqrt{\frac{\Delta}{4}} > 0 \quad \text{e}$$

$$\lambda_2 = 1 + Var(X) + E[X]^2 - \sqrt{\frac{\Delta}{4}} > 1 + Var(X) + E[X]^2 - \sqrt{\left(1 + Var(X) + E[X]^2\right)^2}$$

$$\geq 0$$

Ovvero entrambi gli autovalori sono positivi.

Pertanto la funzione è convessa e in $\left(\frac{Cov(X,Y)}{Var(X)}, E[Y] - \frac{Cov(X,Y)}{Var(X)}E[X]\right)$, f realizza un minimo assoluto, cioè per

$$(\bar{a}, \bar{b}) = \left(\frac{Cov(X,Y)}{Var(X)}, E[Y] - \frac{Cov(X,Y)}{Var(X)} E[X]\right) \text{ si ha}$$

$$\min_{(a,b) \in \mathbb{R}^2} E[(Y - aX - b)^2] = E[(Y - \bar{a}X - \bar{b})^2]$$

e la retta di regressione è quindi:

$$y = \bar{a}X + \bar{b}.$$

Domanda: come si poteva concludere più rapidamente che $Hf(a,b)$ è semidefinita positiva per ogni $(a,b) \in \mathbb{R}^2$?

Domanda: anche calcolando direttamente gli autovalori come si poteva concludere che entrambi gli autovalori sono positivi o nulli senza scrivere la formula risolutiva?