Geometría de la muestra

viernes, 10 de febrero de 2023 7:08 a. m.

$$\mu = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_p \end{bmatrix} \qquad \Sigma = \begin{bmatrix} \sigma_{11} & & \text{cov} \\ & \ddots & \\ \text{cov} & & \sigma_{pp} \end{bmatrix}$$

Obs

Sea $\mathbb{X}$ un vector aleatori.

la distribución de $\mathbb{X}$ se puede describir mediante la distribución conjunta de sus entradas

$$\underline{f_{X_1,\dots,X_p}(x_1,\dots,x_p)} = f_{\mathbb{X}}(\mathbb{X})$$

Recordatorio:

X_i, X_k son independientes, la función de distribución conjunta es:

$$F_{X_i X_k}(x_i, x_k) = P(X_i \le x_i)\,P(X_k \le x_k)$$

$$= \underline{F_{X_i}(x_i) \cdot F_{X_k}(x_k)}$$

también se cumple para la densidad

Generalizando:

si son p v.a. continuas, son independientes si.

$$f_{X_1,\dots,X_p}(x_1,\dots,x_p) = \prod_{i=1}^{p} f_{X_i}(x_i)$$

$$X_i, X_k \text{ indep} \Rightarrow \text{cov}(X_i, X_k) = 0$$

No necesariamente en el caso contrario.

Notación: Sea $\mathbb{X}$ vector aleatorio Denotamos

$$\mu_{\mathbb{X}} = \mu = E(\mathbb{X})$$

$$\Sigma = E((\mathbb{X}-\mu)(\mathbb{X}-\mu)') \longrightarrow$$ Los elementos de Σ son las p varianzas en la diagonal y las $\frac{p(p-1)}{2}$ covarianzas.

Definimos la matriz de correlación poblacional como

$$\rho = \begin{bmatrix} \rho_{11} & \cdots & \rho_{1p} \\ & \ddots & \\ \rho_{p1} & \cdots & \rho_{pp} \end{bmatrix} \longrightarrow \underbrace{\rho_{11} = \cdots = \rho_{pp} = 1}_{\text{diagonal.}}$$

Matriz desv. estándar

$$V^{1/2} = \begin{bmatrix} \sqrt{\sigma_{11}} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\sigma_{pp}} \end{bmatrix}$$

1) $V^{1/2}\rho V^{1/2} = \Sigma \rightarrow$ Tarea demostrarlo

2) $\rho = \left(V^{1/2}\right)^{-1}\Sigma\left(V^{1/2}\right)^{-1}$

Se pueden definir particiones de vectores y matrices

$$\mathbb{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_q \\ \hdashline X_{q+1} \\ \vdots \\ X_p \end{pmatrix} = \begin{pmatrix} \mathbb{X}^{(1)} \\ \hdashline \mathbb{X}^{(2)} \end{pmatrix}$$

$$\mu = E(\mathbb{X}) = \begin{pmatrix} E(\mathbb{X}^{(1)}) \\ \hdashline E(\mathbb{X}^{(2)}) \end{pmatrix} = \begin{pmatrix} \mu^{(1)} \\ \hdashline \mu^{(2)} \end{pmatrix}$$

$$\Sigma = \left(\begin{array}{c:c} \Sigma_{11} & \Sigma_{12} \\ \hdashline \Sigma_{21} & \Sigma_{22} \end{array}\right)$$

$$\Sigma_{12} \neq \Sigma_{21}$$

$$\Sigma_{12}' = \Sigma_{21}$$

Combinaciones lineales de vectores:

$Y = aX + b \qquad a, b \in \mathbb{R} \qquad X$ v.a.

$$E(Y) = a\,E(X) + b$$

$$Var(Y) = a^2\,Var(X)$$

$a, b, c \in \mathbb{R} \qquad X, Y, Z$ v.a.

$$cov(aX, bY) = ab\,cov(X, Y)$$

$$cov(X, Y + c) = cov(X, Y)$$

$$cov(X, Y + Z) = cov(X, Y) + cov(X, Z)$$

$$c = \begin{pmatrix} a \\ b \end{pmatrix} \qquad \mathbb{X} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$$

$$E(c'\mathbb{X}) = c' E(\mathbb{X}) = c'\mu$$

$$Var(c'\mathbb{X}) = c' \Sigma c$$

$\bar{\mathbb{X}} \to$ para estimar μ

$S_n \to$ par estimar Σ

↳ sesgado

$S \to$ para estimar Σ

↳ insesgado

Geometría de la muestra

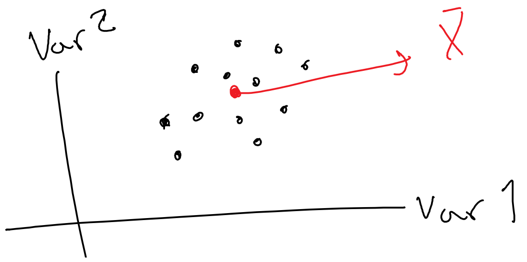
Asumimos que la muestra es aleatoria y las mediciones diferentes no relacionadas.

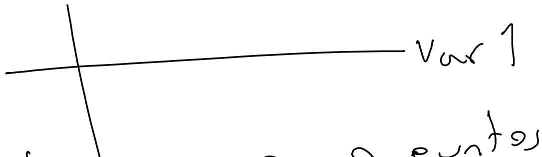
$$\mathbb{X} = \begin{pmatrix} X_{11} & \cdots & X_{1p} \\ \vdots & \ddots & \vdots \\ X_{n1} & \cdots & X_{np} \end{pmatrix} = \begin{pmatrix} \mathbb{X}'_1 \\ \vdots \\ \mathbb{X}'_n \end{pmatrix}$$

→ los valores de las p-variables en la obs 1.

- Cada fila es una observación

n vectores en $\mathbb{R}^p$

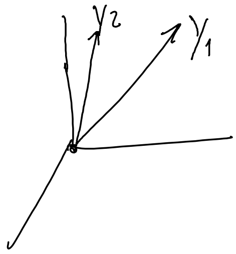


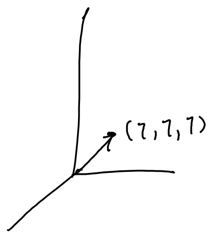


Se puede visualizar como n puntos en $\mathbb{R}^p$ ($p \leq 3$)
n mediciones de las p variables
$\bar{X} \to$ centro de gravedad de la nube de puntos

También se puede representar como p vectores en $\mathbb{R}^n$

$$\mathbb{X} = \begin{pmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & & \\ x_{n1} & \cdots & x_{np} \end{pmatrix} = \left(\mathbb{Y}_1 \,|\, \mathbb{Y}_2 \,|\, \cdots \,|\, \mathbb{Y}_p \right)$$



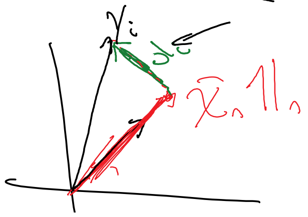


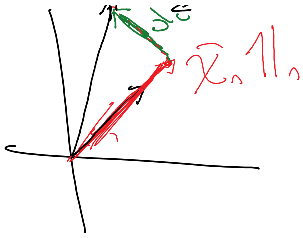
Definimos $\mathbb{1}_n$ como $\mathbb{1}_n = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$

La proyección ortogonal de $\mathbb{Y}_i$ sobre $\mathbb{1}_n$

$$= \underbrace{\left(\frac{x_{1i} + x_{2i} + \cdots + x_{ni}}{n} \right)} \mathbb{1}_n$$

$$= \bar{x}_i \mathbb{1}_n$$





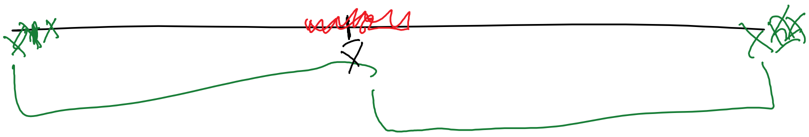
$$d_i = y_i - \bar{x}_i 1_n = \begin{pmatrix} x_{1i} - \bar{x}_i \\ \vdots \\ x_{ni} - \bar{x}_i \end{pmatrix}$$

$$L_{d_i} = \sqrt{d_i' d_i}$$

$$L_{d_i}^2 = d_i' d_i = \sum_{j=1}^{n} (x_{ji} - \bar{x}_i)^2$$

$$\frac{L_{d_i}^2}{n} = S_{ii}$$

$$L_{d_i}^2 = n \cdot S_{ii}$$



Vectores largos representan varianzas grandes,
cortos representan varianzas pequeñas.