Optimización analítica: otro ejemplo algo menos simple

Ejercicio

Estimar los parámetros $\mathbf{\Theta} \equiv (\mu_1, \mu_2)$ de una gaussiana bivariada (en $\mathbb{R}^2$), en la que la matriz de covarianza

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}$$

es conocida:

$$p(x_1, x_2; \mu_1, \mu_2) = A \cdot \exp\left(-B\left[\frac{(x_1-\mu_1)^2}{\sigma_1^2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2} - \frac{2\sigma_{12}(x_1-\mu_1)(x_2-\mu_2)}{(\sigma_1\sigma_2)^2}\right]\right)$$

donde

$$A = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-(\sigma_{12}/\sigma_1\sigma_2)^2}}, \qquad B = \frac{1}{2(1-(\sigma_{12}/\sigma_1\sigma_2)^2)}$$

Otro ejercicio algo más complejo: Asumir que Σ tampoco es conocida; es decir estimar: $\mathbf{\Theta} \equiv (\mu_1, \mu_2, \sigma_1, \sigma_2, \sigma_{12})$

3.1. (p.3.9, 0.5 puntos) Obtener los estimadores de máxima verosimiltud del vector media de una gaussiana bi-variada, cuya matriz de covarianza es fija y conocida partir de una muestra de de vectores bi-dimensionales x_1,x_2,...,x_N.

La matriz de covarianza es conocida: $\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}$

$$p(x_1, x_2, \mu_1, \mu_2) = \underbrace{\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\left(\frac{\sigma_{12}}{\sigma_1\sigma_2}\right)^2}}}_{A} \cdot e^{\left(-\underbrace{\frac{1}{2\left(1-\left(\frac{\sigma_{12}}{\sigma_1\sigma_2}\right)^2\right)}}_{B} \cdot \underbrace{\left[\frac{(x_1-\mu_1)^2}{\sigma_1^2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2} - \frac{2\sigma_{12}(x_1-\mu_1)(x_2-\mu_2)}{(\sigma_1\sigma_2)^2}\right]}_{C}\right)}$$

$$L_s(\mu_1, \mu_2) \Rightarrow \log\left(\prod_{n=1}^{N} (p(x_n; \mu_1, \mu_2)\right) \Rightarrow \log\left(\prod_{n=1}^{N} (A \cdot e^{(-BC)})\right)$$

$$\Rightarrow \log\left(A^N \cdot e^{\left(-B \cdot \underbrace{\sum_{n=1}^{N}\left(\frac{(x_{n1}-\mu_1)^2}{\sigma_1^2} + \frac{(x_{n2}-\mu)^2}{\sigma_2^2} - \frac{2\sigma_{12}(x_{n1}-\mu_1)(x_{n2}-\mu_2)}{(\sigma_1\sigma_2)^2}\right)}_{C'}\right)}\right)$$

$$\Rightarrow \log(A^N) + \log(e^{-BC'}) \Rightarrow \log(A^N) - BC' \Rightarrow \log\left(\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\left(\frac{\sigma_{12}}{\sigma_1\sigma_2}\right)^2}}\right)^n - BC'$$

$$\Rightarrow N\log A - BC' \Rightarrow N\log 1 - N\log\left(2\pi\sigma_1\sigma_2\sqrt{1-\left(\frac{\sigma_{12}}{\sigma_1\sigma_2}\right)^2}\right. - BC'$$

$$\Rightarrow -N\log\left(2\pi\sigma_1\sigma_2\sqrt{1-\left(\frac{\sigma_{12}}{\sigma_1\sigma_2}\right)^2}\right. - \frac{1}{2\left(1-\left(\frac{\sigma_{12}}{\sigma_1\sigma_2}\right)^2\right)} \cdot \left(\sum_{n=1}^{N}\left(\frac{(x_{n1}-\mu_1)^2}{\sigma_1^2} + \frac{(x_{n2}-\mu_2)^2}{\sigma_2^2} - \frac{2\sigma_{12}(x_{n1}-\mu_1)(x_{n2}-\mu_2)}{(\sigma_1\sigma_2)^2}\right)\right)$$

$\boxed{\text{Derivada parcial en función } \mu_1}$

$$\frac{\partial Ls(\mu_1,\mu_2)}{\partial \mu_1} = 0 - \frac{1}{2\left(1-\left(\frac{\sigma_{12}}{\sigma_1\sigma_2}\right)\right)} \cdot \left(\frac{1}{\sigma_1^2} \cdot \sum_{n=1}^{N} 2\cdot(x_{n1}-\mu_1) + 0 - \frac{2\sigma_{12}}{(\sigma_1\sigma_2)^2} \sum_{n=1}^{N} -(x_{n2}-\mu_2) \right) = 0$$

$$\Rightarrow -\frac{1}{2\left(1-\left(\frac{\sigma_{12}}{\sigma_1\sigma_2}\right)\right)} \cdot \left(\frac{2}{\sigma_1^2}\right) \cdot \underbrace{\left(\sum_{n=1}^{N} (x_{n1}-\mu_1) + \frac{\sigma_{12}}{\sigma_2^2} \sum_{n=1}^{N} (x_{n2}-\mu_2) \right)}_{\text{Debe ser } 0} = 0$$

nunca serán 0

$$\Rightarrow \sum_{n=1}^{N} (x_{n1}-\mu_1) + \frac{\sigma_{12}}{\sigma_2^2} \sum_{n=1}^{N} (x_{n2}-\mu_2)) = 0 \Rightarrow \underbrace{\sum_{n=1}^{N} x_{n1}}_{\substack{\text{lo llamaremos} \\ S_1}} - \sum_{n=1}^{N} \mu_1 + \frac{\sigma_{12}}{\sigma_2^2} \underbrace{\sum_{n=1}^{N} x_{n2}}_{\substack{\text{lo llamaremos} \\ S_2}} - \frac{\sigma_{12}}{\sigma_2^2} \sum_{n=1}^{N} \mu_2 = 0$$

$$\Rightarrow \boxed{S_1 - N\mu_1 + \frac{\sigma_{12}}{\sigma_2^2} S_2 - \frac{\sigma_{12}}{\sigma_2^2} N\mu_2 = 0}$$

$\boxed{\text{Derivada parcial en función de } \mu_2}$

$$\frac{\partial L_s(\mu_1,\mu_2)}{\partial \mu_2} = 0 - \frac{1}{2\left(1-\left(\frac{\sigma_{12}}{\sigma_1\sigma_2}\right)\right)} \cdot \left(\frac{1}{\sigma_2^2} \sum_{n=1}^{N} 2(x_{n2}-\mu_2) + \frac{2\sigma_{12}}{(\sigma_1\sigma_2)^2} \sum_{n=1}^{N} (x_{n1}-\mu_1) \right) = 0$$

$$\Rightarrow \underbrace{- \frac{1}{2\left(1-\left(\frac{\sigma_{12}}{\sigma_1\sigma_2}\right)\right)} \cdot \left(\frac{1}{\sigma_2^2}\right)}_{\text{Nunca serán 0}} \cdot \underbrace{\left(\sum_{n=1}^{N} (x_{n2}-\mu_2) + \frac{\sigma_{12}}{\sigma_1^2} \sum_{n=1}^{N} (x_{n1}-\mu_1) \right)}_{\text{Debe ser 0}} = 0$$

$$\Rightarrow \sum_{n=1}^{N} (x_{n2}-\mu_2) + \frac{\sigma_{12}}{\sigma_1^2} \sum_{n=1}^{N} (x_{n1}-\mu_1)) = 0 \Rightarrow \underbrace{\sum_{n=1}^{N} x_{n2}}_{\text{lo llamamos } S_2} - \sum_{n=1}^{N} \mu_2 + \frac{\sigma_{12}}{\sigma_1^2} \underbrace{\sum_{n=1}^{N} x_{n1}}_{\text{lo llamamos } S_1} - \frac{\sigma_{12}}{\sigma_1^2} \sum_{n=1}^{N} \mu_1 = 0$$

$$\Rightarrow \boxed{S_2 - N\mu_2 + \frac{\sigma_{12}}{\sigma_1^2} S_1 - \frac{\sigma_{12}}{\sigma_1^2} N\mu_1 = 0}$$

SISTEMAS DE ECUACIONES CON LOS RESULTADOS

$$\begin{cases} S_1 - N\mu_1 + \dfrac{\sigma_{12}}{\sigma_2^2} S_2 - \dfrac{\sigma_{12}}{\sigma_2^2} N\mu_2 = 0 \\[2ex] S_2 - N\mu_2 + \dfrac{\sigma_{12}}{\sigma_1^2} S_1 - \dfrac{\sigma_{12}}{\sigma_1^2} N\mu_1 = 0 \end{cases}$$

Despejamos μ_1 en la primera ecuación

$$\mu_1 = \frac{\sigma_{12}}{\sigma_2^2 N} S_2 - \frac{\sigma_{12} N\mu_2}{\sigma_2^2 N} + \frac{S_1}{N}$$

Sustituimos μ_1 en la segunda ecuación

$$S_2 - N\mu_2 + \frac{\sigma_{12}}{\sigma_1^2} S_1 - \left(\frac{\sigma_{12}}{\sigma_1^2} N\right)\left(\frac{\sigma_{12} S_2}{\sigma_2^2 N} - \frac{\sigma_{12} N\mu_2}{\sigma_2^2 N} + \frac{S_1}{N}\right) = 0$$

SIMPLIFICAMOS

$$S_2 - N\mu_2 + \cancel{\frac{\sigma_{12}}{\sigma_1^2} S_1} - \left(\frac{\sigma_{12}}{\sigma_1^2} \cancel{N}\right)\left(\frac{\sigma_{12} S_2}{\sigma_2^2 \cancel{N}} - \frac{\sigma_{12} N\mu_2}{\sigma_2^2 \cancel{N}} + \frac{\cancel{S_1}}{\cancel{N}}\right) = 0$$

$$S_2 - N\mu_2 - \left(\frac{\sigma_{12}^2}{\sigma_1^2 \sigma_2^2}\right)\left(S_2 - \mu_2\right) = 0 \Longrightarrow \underbrace{\left(S_2 - N\mu_2\right)}_{\text{debe ser } 0} \underbrace{\left(1 - \frac{\sigma_{12}^2}{\sigma_1^2 \sigma_2^2}\right)}_{\text{En general no será nunca } 0} = 0$$

$$S_2 - N\mu_2 = 0 \qquad \mu_2 = \frac{S_2}{N} \Longrightarrow \mu_2 = \frac{\sum_{n=1}^{N} x_{n_2}}{N}$$

SISTEMA DE ECUACIONES 2

Habiendo obtenido que $\mu_2 = \frac{S_2}{N}$ en la segunda ecuacion, sustituimos μ_2 en la primera ecuación:

$$S_1 - N\mu_1 + \frac{\sigma_{12}}{\sigma_2^2} S_2 - \frac{\sigma_{12}}{\sigma_2^2} N\mu_2 = 0 \Rightarrow S_1 - N\mu_1 + \frac{\sigma_{12}}{\sigma_2^2} S_2 - \frac{\sigma_{12} N}{\sigma_2^2} \cdot \frac{S_2}{N} = 0$$

$\boxed{\text{Simplificamos}}$

$$\Rightarrow S_1 - N\mu_1 + \frac{\sigma_{12}}{\sigma_2^2} S_2 - \frac{\sigma_{12} \cancel{N} S_2}{\sigma_2^2 \cancel{N}} = 0 \Rightarrow S_1 - N\mu_1 = 0$$

$$\Rightarrow \mu_1 = \frac{S_1}{N} \Rightarrow \frac{\sum_{n=1}^{N} x_{n1}}{N}$$

Resultados

$$\mu_1 = \frac{\sum_{n=1}^{N} x_{n1}}{N} \qquad \mu_2 = \frac{\sum_{n=1}^{N} x_{n2}}{N}$$