Optimización analítica: otro ejemplo simple

- Estimar los parámetros† $\boldsymbol{\Theta} \equiv (\mu, \sigma)$ de una gaussiana univariada (en $\mathbb{R}^1$):

$$p(x; \boldsymbol{\Theta}) \stackrel{\text{def}}{=} p(x;\ \mu, \sigma) \ = \ \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

- Logaritmo de la verosimilitud de una muestra $S = \{x_1, \ldots, x_N\}$:

$$\begin{aligned} q_S(\boldsymbol{\Theta}) \ \equiv \ L_S(\mu, \sigma) \ &= \ \log \prod_{n=1}^{N} p(x_n; \mu, \sigma) \ = \ \sum_{n=1}^{N} \log p(x_n; \mu, \sigma) \\ &= \ N \ \log \frac{1}{\sigma\sqrt{2\pi}} - \frac{1}{2\sigma^2} \sum_{n=1}^{N} (x_n - \mu)^2 \end{aligned}$$

- Para una estimación de máxima verosimilitud basta hacer $\boldsymbol{\nabla} L_S(\boldsymbol{\Theta}) = \mathbf{0}$. En nuestro caso unidimensional (ejercicio):

$$\frac{\partial L_S(\mu, \sigma)}{\partial \mu} = 0 \ \Rightarrow \ \hat{\mu} \ = \ \frac{1}{N} \sum_{n=1}^{N} x_n \, ; \qquad \frac{\partial L_S(\mu, \sigma)}{\partial \sigma} = 0 \ \Rightarrow \ \hat{\sigma}^2 = \frac{1}{N} \sum_{n=1}^{N} (x_n - \hat{\mu})^2$$

† media y desviación típica

$$p(x;\Theta) \overset{def}{=} p(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)}$$

$$q_s(\Theta) \equiv L_s(\mu,\sigma) = \log \prod_{n=1}^{N} p(x_n;\mu,\sigma) = \sum_{n=1}^{N} \log p(x_n;\mu,\sigma)$$

$$L_s(\mu,\sigma) = \log\left(\prod_{n=1}^{N}\left(\frac{1}{\sigma\sqrt{2\pi}} e^{\left(\frac{-1}{2\sigma^2}\cdot(x-\mu)^2\right)}\right)\right)$$

$$= \log\left(\left(\frac{1}{\sigma\sqrt{2\pi}}\right)^N \cdot e^{\left(\frac{-1}{2\sigma^2}\cdot\sum_{n=1}^{N}(x_n-\mu)^2\right)}\right)$$

$$= \log\left(\frac{1}{\sigma\sqrt{2\pi}}\right)^N + \log\left(e^{\left(-\frac{1}{2\sigma^2}\cdot\sum_{n=1}^{N}(x_n-\mu)^2\right)}\right)$$

$$= N\cdot\log\left(\frac{1}{\sigma\sqrt{2\pi}}\right) - \frac{1}{2\sigma^2}\cdot\sum_{n=1}^{N}(x_n-\mu)^2$$

$$= N\log 1 - N\log(\sigma\sqrt{2\pi}) - \frac{1}{2\sigma^2}\cdot\sum_{n=1}^{N}(x_n-\mu)^2$$

$$\frac{\partial L_s(\mu,\sigma)}{\partial\mu} = -\frac{1}{2\sigma^2}\cdot\sum_{n=1}^{N} 2(x_n-\mu) = 0$$

$$\Rightarrow -\frac{2}{2\sigma^2}\cdot\sum_{n=1}^{N}(x_n-\mu) = 0$$

$$\Rightarrow -\frac{1}{\sigma^2}\sum_{n=1}^{N} x_n + \frac{1}{\sigma^2}\sum_{n=1}^{N}\mu = 0$$

$$\Rightarrow \frac{\sigma^2}{\sigma^2}\sum_{n=1}^{N} x_n = N\mu$$

$$\Rightarrow \boxed{\mu = \frac{\sum_{n=1}^{N} x_n}{N}}$$

$$L_s(\mu,\sigma) = N\log 1 - N\log(\sigma\sqrt{2\pi}) - \frac{1}{2\sigma^2}\cdot\sum_{n=1}^{N}(x_n-\mu)^2$$

$$= N\log 1 - N\log(\sigma^2)^{\frac{1}{2}} - N\log(2\pi)^{\frac{1}{2}} - \frac{1}{2\sigma^2}\cdot\sum_{n=1}^{N}(x_n-\mu)^2$$

$$= N\log 1 - \frac{N}{2}\log(\sigma^2) - \frac{N}{2}\log(2\pi) - \frac{1}{2\sigma^2}\sum_{n=1}^{N}(x_n-\mu)^2$$

$$\frac{\partial}{\partial\sigma^2}\left(N\log 1 - \frac{N}{2}\log(\sigma^2) - \frac{N}{2}\log(2\pi) - \frac{1}{2\sigma^2}\sum_{n=1}^{N}(x_n-\mu)^2\right) =$$

$$= 0 - \frac{N}{2\sigma^2} - 0 + \left(\frac{1}{2(\sigma^2)^2}\right)\left(\sum_{n=1}^{N}(x_n-\mu)^2\right)$$

$$= \underbrace{\frac{-N}{2(\sigma^2)^2}}_{\text{Nunca será 0}}\left(\underbrace{\sigma^2 - \frac{1}{N}\sum_{n=1}^{N}(x_n-\mu)^2}_{\text{Debe ser 0}}\right) = 0$$

$$\sigma^2 - \frac{1}{N}\sum_{n=1}^{N}(x_n-\mu)^2 = 0$$

$$\boxed{\sigma^2 = \frac{1}{N}\sum_{n=1}^{N}(x_n-\mu)^2}$$