Ejemplo: aplicación de la técnica de multiplicadores de Lagrange

- Lagrangiana: $\Lambda(p_1, p_2, p_3, \beta) = 4\log p_1 + 2\log p_2 + \log p_3 + \beta\,(1 - p_1 - p_2 - p_3)$
- Soluciones óptimas en función del multiplicador de Lagrange:

$$\left.\begin{array}{l} \dfrac{\partial \Lambda}{\partial p_1} = \dfrac{4}{p_1} - \beta = 0 \\ \dfrac{\partial \Lambda}{\partial p_2} = \dfrac{2}{p_2} - \beta = 0 \\ \dfrac{\partial \Lambda}{\partial p_3} = \dfrac{1}{p_3} - \beta = 0 \end{array}\right\} \qquad \begin{array}{l} p_1^\star(\beta) = \dfrac{4}{\beta} \\ p_2^\star(\beta) = \dfrac{2}{\beta} \\ p_3^\star(\beta) = \dfrac{1}{\beta} \end{array}$$

- Función dual de Lagrange:

$$\Lambda_D(\beta) \;=\; 4\log\frac{4}{\beta} + 2\log\frac{2}{\beta} + \log\frac{1}{\beta} + \beta(1 - \frac{4}{\beta} - \frac{2}{\beta} - \frac{1}{\beta}) \;=\; \beta - 7\log\beta - 7 + 10\log 2$$

- Valor óptimo del multiplicador de Lagrange: $\dfrac{d\Lambda_D}{d\beta} = 1 - \dfrac{7}{\beta} = 0 \;\Rightarrow\; \beta^\star = 7$
- Solución final: $p_1^\star = p_1^\star(\beta^\star) = \dfrac{4}{7} \qquad p_2^\star = p_2^\star(\beta^\star) = \dfrac{2}{7} \qquad p_3^\star = p_3^\star(\beta^\star) = \dfrac{1}{7}$

EJERCICIO: Demostrar que en cualquier problema de clasificación en C clases, la estimación de máxima verosimilitud de la probabilidad a priori de cada clase c, $1 \leq c \leq C$, es $\hat{p}_c = n_c/N$, donde $N = \sum_c n_c$ es el número total de datos observados y n_c es el número de datos de la clase c.

3.5. (p.3.18, 0.5 puntos) Ejercicio al pie de la página indicada: estimación por máxima verosimilitud de las probabilidades a priori de un clasificador genérico en C clases.

Modelo: $P(c=c) = p_c$

Verosimilitud y logarítmo de la verosimilitud:

$$P(S|\Theta) = \prod_{c=1}^{C} \prod_{n=1}^{N_c} p_c$$

$$q_S(\Theta) = L_S(\Theta) = \log P(S|\Theta) = \sum_{c=1}^{C} N_c \cdot \log p_c$$

Estimación de máxima verosimilitud

$$\Theta^* = \arg\max_{\Theta} L_S(\Theta) = \arg\max_{\sum_{c=1}^{C} p_c = 1} \left(\sum_{c=1}^{C} N_c \cdot \log p_c \right)$$

$$\boxed{\text{Aplicación de la técnica de multiplicadores de Lagrange}}$$

Lagrangiana

$$\Lambda(p_1 \cdots p_n, \beta) = \sum_{c=1}^{C} N_c \cdot \log p_c + \beta\left(1 - \sum_{c=1}^{C} p_c\right)$$

Soluciones óptimas en función del multiplicador de Lagrange

$$\left.\frac{\partial \Lambda}{\partial p_c} = \frac{N_c}{p_c} - \beta \right\} \quad p_c^*(\beta) = \frac{N_c}{\beta}$$

Función dual de Lagrange

$$\Lambda_D(\beta) = \sum_{c=1}^{C} N_c \cdot \log \frac{N_c}{\beta} + \beta \left(1 - \sum_{c=1}^{C} \frac{N_c}{\beta} \right)$$

Valor óptimo del multiplicador de Lagrange

$$\frac{d\Lambda_D}{d\beta}\left(\sum_{c=1}^{C} N_c \cdot \log \frac{N_c}{\beta} + \beta \left(1 - \sum_{c=1}^{C} \frac{N_c}{\beta} \right) \right) \Rightarrow$$

$$\Rightarrow \frac{d\Lambda_D}{d\beta}\left(\sum_{c=1}^{C} N_c \cdot \log N_c - \sum_{c=1}^{C} N_c \cdot \log \beta + \beta - \frac{\beta}{\beta} \sum_{c=1}^{C} N_c \right)$$

$$\Rightarrow \quad 0 - \left(\sum_{c=1}^{C} N_c \right) \cdot \frac{1}{\beta} + 1 - 0$$

$$\Rightarrow \quad -\frac{N}{\beta} = -1$$

$$\Rightarrow \quad N = \beta$$

Solución final

$$p_n^*(\beta^*) = \frac{N_c}{N}$$