Resumen Matemático Ampliado: textitLogistic Regression in Data Analysis: An Overview (Maalouf, 2011)

Carlos – Compilado con ayuda de ChatGPT

Modelo de regresión logística

Para una variable binaria $y_i \in \{0, 1\}$, el modelo de regresión logística predice la probabilidad:

$$p_i = \frac{1}{1 + e^{-x_i\beta}}$$

Función de verosimilitud

$$L(\beta) = \prod_{i=1}^{n} p_i^{y_i}(1 - p_i)^{1-y_i}$$

Log-verosimilitud

$$\ell(\beta) = \sum_{i=1}^{n} [y_i \ln(p_i) + (1 - y_i) \ln(1 - p_i)]$$

Derivadas

Gradiente:

$$\nabla_\beta \ell(\beta) = X^\top (y - p)$$

Hessiano (segunda derivada):

$$\nabla_\beta^2 \ell(\beta) = -X^\top V X, \quad V = \text{diag}(p_i(1 - p_i))$$

Regularización Ridge (L2)

Se añade un término penalizado:

$$\ell_\lambda(\beta) = \ell(\beta) - \frac{\lambda}{2}\|\beta\|^2$$

Gradiente regularizado:

$$\nabla_\beta \ell_\lambda(\beta) = X^\top (y - p) - \lambda\beta$$

Hessiano regularizado:

$$\nabla^2_\beta \ell_\lambda(\beta) = -X^\top V X - \lambda I$$

TR-IRLS (Trust Region IRLS)

Actualización generalizada del paso de Newton:

$$\beta^{(k+1)} = \beta^{(k)} + s$$

Donde s es solución de:

$$(X^\top V X + \lambda I)s = X^\top V z - \lambda\beta$$

Eventos raros y correcciones

Ajuste del intercepto (King & Zeng)

$$\tilde{\beta}_0 = \hat{\beta}_0 - \ln\left[\left(\frac{1-\tau}{\tau}\right)\left(\frac{\hat{y}}{1-\hat{y}}\right)\right]$$

Muestreo estratificado

Peso para observación i:

$$w_i = \frac{Q_i}{H_i}$$

Verosimilitud ponderada:

$$\ell_w(\beta) = \sum_{i=1}^{n} w_i \ln\left(\frac{e^{x_i\beta}}{1+e^{x_i\beta}}\right)$$

Corrección de Firth

Se basa en el ajuste de penalización de tipo Jeffreys:

$$\ell^*(\beta) = \ell(\beta) + \frac{1}{2}\log|I(\beta)|$$

Regla de decisión

$$\hat{y}_i = \begin{cases} 1 & \text{si } p_i \geq c \\ \\ 0 & \text{si } p_i < c \end{cases} \quad (\text{usualmente } c = 0.5)$$