Notas sobre Regresión Lineal y Logística

Carlos

April 24, 2025

Contents

1 Regresión Logística: Log-verosimilitud y derivadas

La log-verosimilitud para la regresión logística está dada por:

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^{N} y^{(i)} \log \sigma(\boldsymbol{\theta}^T \boldsymbol{x}^{(i)}) + (1 - y^{(i)}) \log\big(1 - \sigma(\boldsymbol{\theta}^T \boldsymbol{x}^{(i)})\big)$$

Antes de calcular la derivada, recordamos:

$$\sigma(z) = \frac{1}{1 + e^{-z}} \quad \text{y} \quad \frac{d}{dz}\sigma(z) = \sigma(z)(1 - \sigma(z))$$

La derivada de la log-verosimilitud con respecto a θ_j es:

$$\frac{\partial \ell(\boldsymbol{\theta})}{\partial \theta_j} = \sum_{i=1}^{N} \left[y^{(i)} - \sigma(\boldsymbol{\theta}^T \boldsymbol{x}^{(i)}) \right] x_j^{(i)}$$

2 Derivación recursiva del gradiente

Esta fórmula se utiliza para construir el vector gradiente necesario en algoritmos como descenso del gradiente:

$$\nabla_{\boldsymbol{\theta}} \ell(\boldsymbol{\theta}) = \sum_{i=1}^{N} \left[y^{(i)} - \sigma(\boldsymbol{\theta}^T \boldsymbol{x}^{(i)}) \right] \boldsymbol{x}^{(i)}$$

3 Regresión Lineal

3.1 Historia

Sir Francis Galton sugirió el concepto de regresión lineal en 1894.

3.2 Modelo de Regresión Lineal

Regresión lineal simple:

$$y = \beta_0 + \beta_1 x + \xi$$

Regresión lineal multivariada:

$$y = \sum_{i=1}^{d} \beta_i x_i, \quad x_0 = 1$$

En forma matricial:

$$\boldsymbol{\beta} = (\boldsymbol{X}^T \boldsymbol{X})^{-1} \boldsymbol{X}^T \boldsymbol{y}$$

3.3 Regresión Polinomial

El modelo polinomial tiene la forma:

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \cdots + \beta_n x^n + \xi$$

Usamos mínimos cuadrados para estimar los coeficientes minimizando:

$$\arg\min_{\beta_0,\beta_1} \sum_i [y_i - (\beta_0 + \beta_1 x_i)]^2$$

4 Estimaciones con modelo lineal centrado

Si usamos:

$$y_i = \beta_0^* + \beta_1(x_i - \bar{x}) + \varepsilon_i$$

entonces se obtiene:

$$\beta_0 = \bar{y}, \quad \beta_1 = \frac{S_{xy}}{S_{xx}}$$

5 Evaluación del modelo lineal

5.1 F-test

$$F = \frac{\sum(\hat{y}_i - \bar{y})^2/(m-1)}{\sum(y_i - \hat{y}_i)^2/(n-m)} \sim F_{(m-1,n-m)}$$

5.2 t-test

$$t_j = \frac{\hat{\beta}_j}{\sqrt{C_{jj}\frac{\sum(y_i-\hat{y}_i)^2}{n-m}}}$$

6 Fundamentos de la Regresión Logística

Sea $\boldsymbol{X} \in \mathbb{R}^{n\times d}$, y vector binario:

$$P(y_i = 1) = p_i, \quad \mathbb{E}[y_i] = p_i = \boldsymbol{x}_i^T\boldsymbol{\beta}, \quad \mathrm{Var}(y_i) = p_i(1-p_i)$$

7 Función logística y función logit

$$p_i = \frac{e^{\boldsymbol{x}_i^T\boldsymbol{\beta}}}{1 + e^{\boldsymbol{x}_i^T\boldsymbol{\beta}}} \quad \text{y} \quad \eta_i = \log\left(\frac{p_i}{1-p_i}\right) = \boldsymbol{x}_i^T\boldsymbol{\beta}$$

8 Log-verosimilitud y derivadas

$$\ell(\boldsymbol{\beta}) = \sum \left[y_i \log(p_i) + (1 - y_i) \log(1 - p_i) \right]$$

$$\nabla_{\boldsymbol{\beta}} \ell(\boldsymbol{\beta}) = \boldsymbol{X}^T (\boldsymbol{y} - \boldsymbol{p})$$

$$\nabla^2_{\boldsymbol{\beta}} \ell(\boldsymbol{\beta}) = -\boldsymbol{X}^T \boldsymbol{V} \boldsymbol{X}, \quad \text{donde } \boldsymbol{V} = \text{diag}(p_i(1 - p_i))$$

9 Regularización

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

10 Desviación

$$DEV(\hat{\boldsymbol{\beta}}) = -2 \ln L(\hat{\boldsymbol{\beta}})$$

11 Newton-Raphson

$$\boldsymbol{\beta}^{(c+1)} = \left(\boldsymbol{X}^T \boldsymbol{V} \boldsymbol{X} + \lambda \boldsymbol{I} \right)^{-1} \boldsymbol{X}^T \boldsymbol{V} \boldsymbol{Z}^{(c)}$$

12 WLS y función cuadrática

La estimación WLS consiste en minimizar:

$$\frac{1}{2} \boldsymbol{\beta}^T \left(\boldsymbol{X}^T \boldsymbol{V} \boldsymbol{X} + \lambda \boldsymbol{I} \right) \boldsymbol{\beta} - \boldsymbol{\beta}^T (\boldsymbol{X}^T \boldsymbol{V} \boldsymbol{Z}^{(c)})$$