Notas de Regresión Lineal Univariada
Hojas 01 a 10

Carlos Leonardo

Hoja 01: Introducción y función de costo

Modelo lineal univariado:

$$h_\theta(x) = \theta_0 + \theta_1 x$$

Función de costo:

$$J(\theta_0, \theta_1) = \frac{1}{2N} \sum_{i=1}^{N} \left(h_\theta(x^{(i)}) - y^{(i)} \right)^2$$

Objetivo: minimizar $J(\theta_0, \theta_1)$

Hoja 02: Derivadas parciales

Derivada de J respecto a θ_0:

$$\frac{\partial J}{\partial \theta_0} = \frac{1}{2N} \sum_{i=1}^{N} 2(h_\theta(x^{(i)}) - y^{(i)}) = \frac{1}{N} \sum_{i=1}^{N} (h_\theta(x^{(i)}) - y^{(i)})$$

Derivada respecto a θ_1:

$$\frac{\partial J}{\partial \theta_1} = \frac{1}{N} \sum_{i=1}^{N} (h_\theta(x^{(i)}) - y^{(i)}) x^{(i)}$$

Hoja 03: Gradiente y vectorización

Descenso por gradiente:

$$\theta_j := \theta_j - \alpha \cdot \frac{\partial J}{\partial \theta_j}$$

Matriz de diseño:

$$X = \begin{bmatrix} 1 & x^{(1)} \\ 1 & x^{(2)} \\ \vdots & \vdots \\ 1 & x^{(N)} \end{bmatrix}, \quad y = \begin{bmatrix} y^{(1)} \\ y^{(2)} \\ \vdots \\ y^{(N)} \end{bmatrix}$$

Predicciones:

$$h_\theta = X\theta$$

Costo vectorial:

$$J(\theta) = \frac{1}{2N} \|X\theta - y\|^2$$

Hoja 04: Gradiente vectorial

$$\nabla_\theta J = \frac{1}{N} X^T (X\theta - y)$$

$$\theta := \theta - \alpha \cdot \frac{1}{N} X^T (X\theta - y)$$

Hoja 05: Mínimos cuadrados

Minimizar:

$$J(\theta) = \frac{1}{2N} (X\theta - y)^T (X\theta - y)$$

$$\nabla_\theta J = \frac{1}{N} X^T (X\theta - y) \Rightarrow X^T X\theta = X^T y \Rightarrow \theta = (X^T X)^{-1} X^T y$$

Hoja 06: Desarrollo de la derivada

Expandimos:

$$(X\theta - y)^T (X\theta - y) = \theta^T X^T X\theta - 2y^T X\theta + y^T y$$

Derivando:

$$\nabla_\theta J = \frac{1}{2N} (2X^T X\theta - 2X^T y) = \frac{1}{N} X^T (X\theta - y)$$

Hoja 07: Geometría de la función

- $J(\theta)$ es convexa - Tiene un mínimo global - El gradiente apunta hacia la dirección de mayor incremento

Hoja 08: Ejemplo de iteración

$$\theta := \theta - \alpha \cdot \nabla_\theta J$$

Para un punto específico, se puede usar:

$$\theta_0 := \theta_0 - \alpha(h_\theta(x^{(i)}) - y^{(i)})$$

$$\theta_1 := \theta_1 - \alpha(h_\theta(x^{(i)}) - y^{(i)})x^{(i)}$$

Hoja 09: Interpretación geométrica

La dirección del gradiente indica cómo ajustar los parámetros para minimizar el error. El mínimo se alcanza cuando $\nabla_\theta J = 0$

Hoja 10: Gradiente completo en forma matricial

$$J(\theta) = \frac{1}{2N}(X\theta - y)^T(X\theta - y) \quad \Rightarrow \quad \nabla_\theta J = \frac{1}{N}X^T(X\theta - y)$$