Reescritura de la funcion objetivo

Dada la funcion objetivo original:

$$E(w_0, w_1, \ldots, w_D) = \sum_{i=1}^{N} \left(y_i - \left(\sum_{d=1}^{D} w_d \cdot x_{id} + w_0 \right) \right)^2 + \lambda \sum_{d=1}^{D} w_d^2 + \lambda w_0^2$$

1. Demostracion de la reescritura

Para demostrar que se puede expresar como:

$$E(\hat{w}) = \|X\hat{w} - y\|^2 + \lambda \|\hat{w}\|^2$$

Demostracion:
Primero, definimos la matriz de diseo extendida (X) como:

$$X = \begin{bmatrix} 1 & x_{11} & x_{12} & \ldots & x_{1D} \\ 1 & x_{21} & x_{22} & \ldots & x_{2D} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{N1} & x_{N2} & \ldots & x_{ND} \end{bmatrix}$$

donde (x_{ij}) es el elemento (j)-esimo de la observacion (i)-esima.
Definimos el vector de pesos extendido ($\hat{w}$) como:

$$\hat{w} = \begin{bmatrix} w_0 \\ w_1 \\ w_2 \\ \vdots \\ w_D \end{bmatrix}$$

La funcion objetivo original se puede expresar como:

$$E(w_0, w_1, \ldots, w_D) = \|X\hat{w} - y\|^2 + \lambda \|\hat{w}\|^2$$

donde ($\| \cdot \|$) denota la norma euclidiana.

2. Solucion optima para ($\hat{w}$)

Para encontrar el minimo de $E(\hat{w})$, derivamos con respecto a ($\hat{w}$) y establecemos la derivada igual a cero:

$$\frac{\partial E(\hat{w})}{\partial \hat{w}} = 2X^T(X\hat{w} - y) + 2\lambda \hat{w} = 0$$

Resolviendo para ($\hat{w}$):

$$\hat{w} = (X^T X + \lambda I)^{-1} X^T y$$

donde (I) es la matriz identidad de dimensiones ($D + 1 \times D + 1$).

3. Hessiano de $(E(\hat{w}))$ es definido positivo (p.d.)

El Hessiano de $(E(\hat{w}))$ es:

$$H = 2X^T X + 2\lambda I$$

Para demostrar que es p.d., debemos verificar que todos sus valores propios son positivos. Dado que $(X^T X)$ es simetrica y no negativa, y $(\lambda > 0)$, podemos concluir que (H) es p.d.