LINEAR REGRESSION

modelo lineal que busca minimizar la suma de los residuos al cuadrado

Modelo:

$$\min_w \| Xw - y \|_2^2 \rightarrow \text{Norma } 2$$

El objetivo es encontrar los parámetros w que minimizan el residuo

LASSO

Modelo lineal que reduce el número de variables y por ende el número de parámetros a estimar.

modelo:

$$\min_w \frac{1}{2n_{muestras}} \| Xw - y \|_2^2 + \alpha \|w\|_1$$

Expresión que penaliza el vector w es decir se vuelve cero para aquellos w no relevantes.

bandimension

Elastic Net

Es un modelo de regresión lineal
que regulariza con los parámetros L_1 y L_2

modelo:

$$\frac{1}{2N_{muestras}} \cdot \|Y - Xw\|_2^2 + \alpha * L_1 * \|w\|_1$$

$$+ 0.5 * \alpha * (1 - L_1) * \|w\|_2^2$$

Es una combinación del modelo Lasso y Ridge.

Kernel Ridge

Proyecta los datos a un Espacio más pequeño y sobre ese nuevo espacio se minimiza:

$$\ell_{ridge}(w) = \frac{1}{2} \sum_{i=1}^{N} (y_i - \phi(x_i)^T w)^2 + \frac{\lambda}{2} \sum_{j=1}^{p} \|w\|_2$$

ϕ Es el parámetro a encontrar para minimizar la suma de cuadrados de los residuos.

SGD

modelo que permite el localizar el mínimo en una función de forma iterativa.

La función de costo es J que reduce la diferencia entre el modelo y la salida

$$\theta_{t+1} = \theta_t - \eta \frac{\partial J}{\partial \theta_t}$$

tasa de aprendizaje

La derivada de todo el set de entrenamiento

Conjunto parámetros seleccionados,

Todo el set de entrenamiento.

Bayesian Ridge Regression

$$P(w|\lambda) = \mathcal{N}(w|0, \lambda^{-1} I_p)$$

Es una modelo de regresión definido en términos probabilísticos

Usa los parámetros α λ

↓

Que se obtienen por max verosimilitud