

03.2-Regresion Lineal

November 18, 2019

0.1 ### Regresión lineal

Resumen:

Hipótesis:

0.1.1

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

Parámetros:

0.1.2

$$\theta_0, \theta_1$$

Función de costo:

0.1.3

$$J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^n (h_{\theta}(x_i) - y_i)^2$$

Objetivo:

0.1.4

$$\min_{\theta_0, \theta_1} J(\theta_0, \theta_1)$$

Descenso del gradiente A continuación aprenderemos un algoritmo llamado descenso del gradiente, el cuál utilizaremos para optimizar nuestra función de costo y hallar los parámetros correctos para nuestra hipótesis.

Resulta que el algoritmo del descenso del gradiente es un algoritmo más general y se utiliza no sólo en la regresión lineal. De hecho, se utiliza en y todo los algoritmos de machine learning.

Este es el planteamiento del problema:

Tenemos alguna función: $J(\theta_0, \theta_1)$

Gradient descent algorithm

```
repeat until convergence {  
   $\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta_0, \theta_1)$    (for  $j = 0$  and  $j = 1$ )  
}
```

Correct: Simultaneous update

```
temp0 :=  $\theta_0 - \alpha \frac{\partial}{\partial \theta_0} J(\theta_0, \theta_1)$   
temp1 :=  $\theta_1 - \alpha \frac{\partial}{\partial \theta_1} J(\theta_0, \theta_1)$   
 $\theta_0 :=$  temp0  
 $\theta_1 :=$  temp1
```

Queremos: $\min_{\theta_0, \theta_1} J(\theta_0, \theta_1)$

- Iniciamos con valores aleatorios o cualquier valor específico para θ_0 y θ_1
- Haremos pequeños cambios a θ_0 y θ_1 para reducir $J(\theta_0, \theta_1)$ hasta que obtengamos un mínimo global o tal vez un mínimo local de la función de costo.

Algoritmo α es un número llamado learning rate o índice de aprendizaje, y lo que hace es básicamente, controlar que tan grande es un paso que damos con el descenso del gradiente. El término de la derivada valuada en uno de nuestros parámetros lo que hace es darnos el valor de la pendiente de la recta tangente a nuestra función de costo, esto básicamente lo que nos permite es saber si debemos aumentar o disminuir el valor del parámetros que estamos evaluando y en cuánto debemos modificarlo para encontrar el valor que estamos buscando que minimiza nuestra función.

No haremos énfasis en el cálculo de las derivadas de nuestra función de costo simplemente vamos a tomarlas y sustituirlas en nuestro algoritmo.

Finalmente estamos listos para implementar nuestro modelo de regresión lineal y ajustar sus parámetros utilizando el algoritmo del descenso del gradiente.

Gradient descent algorithm

repeat until convergence {

$$\theta_0 := \theta_0 - \alpha \frac{1}{m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})$$

$$\theta_1 := \theta_1 - \alpha \frac{1}{m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)}) \cdot x^{(i)}$$

}