

Practica 1 Regresion Lineal

JoseLu

March 24, 2022

Abstract

The abstract text goes here.

1 Introduction

2 1. Ejercicio sobre la Busqueda Iterativa de Optimos

2.1 Ejercicio1: Algoritmo de Gradiente Descendiente

2.2 Ejercicio2: Uso Gradiente Descendiente

Primero que nada necesitamos definir la funcion, en este caso $E(u, v)$ y tambien necesitaremos las derivadas parciales de esa misma funcion, esto lo haremos para calcular el gradiente de esta funcion

Las iteraciones que tardara este algoritmo en encontrar el minimo depender de mas factores que solo la precision, ademas dependera del learning rate, este valor influye en la longitud de avance del algoritmo de descenso

Las coordenadas en donde el algoritmo encuentra el minimo dependeran del punto inicial, en mi caso las coordendas serian $(x_f, y_f) = ()$

2.3 Ejercicio3:

La funcion sobre la que trabajaremos ahora es $f(x, y) = x^2 + 2y^2 + 2\sin(2\pi x)\sin(y\pi)$ y tendremos que utilizar el gradiente descente para minimizar la funcion. En este caso tenemos un punto inicial $(x_0 = -1, y_0 = 1)$, un learning rate $\eta = 0.01$ y un maximo de 50 iteraciones, en este caso no nos dan un valor sobre el cual calcular el error, esto se puede deber a que la funcion tiene partes en las que es negativa por lo que no se puede aplicar esta tecnica.

Tras ejecutar las 50 iteraciones obtendremos la siguiente tabla.

Coordenadas					
	(-0.5, -0.5)	(1, 1)	(2.1,-2.1)	(-3, 3)	(-2, 2)
Valor obtenido					

Table 1: Coches disponibles

2.4 Error Cuadratico Medio (MSE)

Sabemos que nuestra funcion de MSE es $\frac{1}{n}(h(x) - y)^2$ pero como sabemos que nuestra funcion $h(x)$ es $h_w(x) = w^T \cdot x$

2.5 Gradiente para el SGD

En el caso de SGD necesitamos las derivadas parciales del error, en este caso seria $\frac{\partial E_{in}(w)}{\partial w_j} = \frac{2}{M} \sum_{n=1}^M x_{nj}(h(x_n) - y_n)$

2.6 sgd

La diferencia entre el Gradiente Descendiente y el Gradiente Descendiente Estocastico es que utilizamos unos subconjuntos de los datos llamados minibatches (los cuales deben ser disjuntos) y seguidamente en vez de iterar sobre todo el conjunto de datos lo hacemos sobre el minibatch y seguimos calculando w de la misma forma

3 Conclusion

Write your conclusion here.