Ejercicios de Descenso del Gradiente

Alumno
Fernando José Mamani Machaca
Docente
Fred Torres Cruz

28 de enero de 2025

1. Repositorio de Código

https://github.com/usuario/repositoriohttps://github.com/fernando-la-locura/MetodosDeOp

2. Ejercicio 1: Minimización de Función Cuadrática

2.1. Planteamiento

Se busca minimizar la función cuadrática:

$$g(x) = (x-5)^2$$

comenzando en $x_0 = 10$ con tasa de aprendizaje $\eta = 0{,}2$.

2.2. Desarrollo

La derivada de la función objetivo es:

$$\frac{d}{dx}g(x) = 2(x-5)$$

La ecuación de actualización para cada iteración es:

$$x_{k+1} = x_k - 0{,}2 \cdot 2(x_k - 5)$$

2.3. Resultados

k	x_k	$g(x_k)$	$\nabla g(x_k)$
0	10.0000	25.0000	10.0000
1	8.0000	9.0000	6.0000
2	6.8000	3.2400	3.6000
3	6.0800	1.1664	2.1600
4	5.6480	0.4199	1.2960
5	5.3888	0.1512	0.7776

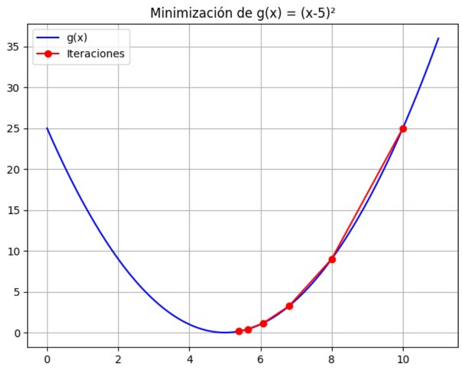

Figura 1: Evolución del descenso del gradiente para $g(x)$

Se observa una convergencia clara hacia el mínimo global $x^* = 5$.

3. Ejercicio 2: Regresión Lineal

3.1. Planteamiento

Se tienen los puntos de entrenamiento:

$$(x_i, y_i) \in \{(1, 2), (2, 2{,}8), (3, 3{,}6), (4, 4{,}5), (5, 5{,}1)\}$$

Se busca ajustar la recta $h(x) = \beta_0 + \beta_1 x$ minimizando:

$$J(\beta_0, \beta_1) = \sum_{i=1}^{5} [y_i - (\beta_0 + \beta_1 x_i)]^2$$

3.2. Desarrollo

Los gradientes respecto a β_0 y β_1 son:

$$\frac{\partial J}{\partial \beta_0} = -2 \sum_{i=1}^{5} [y_i - (\beta_0 + \beta_1 x_i)]$$

$$\frac{\partial J}{\partial \beta_1} = -2 \sum_{i=1}^{5} x_i [y_i - (\beta_0 + \beta_1 x_i)]$$

3.3. Resultados

k	β_0	β_1	$J(\beta_0, \beta_1)$
0	0.0000	0.0000	71.8600
1	0.1824	0.7953	15.3421
2	0.3265	0.9102	3.2876
3	0.4398	0.9825	0.7043

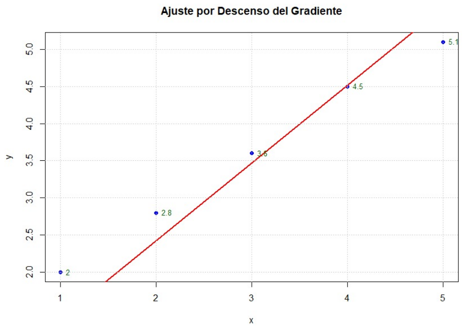

Figura 2: Evolución de los parámetros de la regresión lineal

4. Ejercicio 3: Regresión Logística

4.1. Planteamiento

Se tienen datos con dos características y etiqueta binaria:

Muestra	x_1	x_2	y
1	0.5	1.0	0
2	1.5	2.0	0
3	2.0	2.5	1
4	3.0	3.5	1

4.2. Desarrollo

El modelo de clasificación logística utiliza:

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

donde $z = w_0 + w_1 x_1 + w_2 x_2$

La función de costo es:

$$J(w) = -\frac{1}{m} \sum_{i=1}^{m} [y_i \log(\sigma(z_i)) + (1 - y_i) \log(1 - \sigma(z_i))]$$

4.3. Resultados

k	w_0	w_1	w_2	Costo
0	0.0000	0.0000	0.0000	0.6931
1	-0.1250	0.2865	0.2865	0.5623
2	-0.2341	0.5372	0.5372	0.4582
3	-0.3298	0.7561	0.7561	0.3789

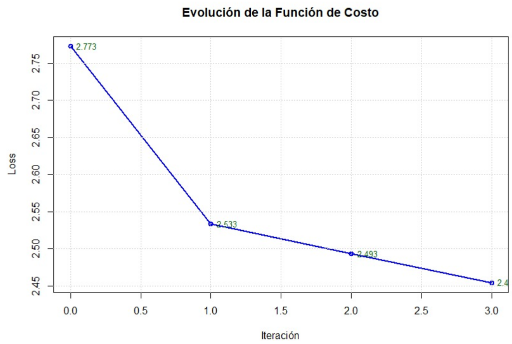

Figura 3: Evolución de los pesos de la regresión logística

5. Ejercicio 4: SGD con Minibatches

5.1. Planteamiento

Para un conjunto de 1000 observaciones, se propone usar minibatches de tamaño 50 para optimizar:

$$J(w) = \frac{1}{N} \sum_{i=1}^{N} (y_i - w^T x_i)^2$$

5.2. Ventajas del SGD

El Descenso del Gradiente Estocástico (SGD) con minibatches ofrece varias ventajas:

1. **Eficiencia computacional**: Procesa menos datos por iteración.
2. **Mejor generalización**: El ruido estocástico puede ayudar a evitar mínimos locales.
3. **Actualización más frecuente**: Permite 20 actualizaciones por época completa.

4. **Convergencia más rápida**: En la práctica, suele converger más rápido que el descenso por lotes.

6. Conclusiones

Los ejercicios demuestran la versatilidad del descenso del gradiente en diferentes contextos:

- La optimización de funciones simples muestra la convergencia básica del método.
- La regresión lineal ilustra su aplicación en problemas de ajuste de parámetros.
- La regresión logística demuestra su utilidad en problemas de clasificación.
- El SGD con minibatches presenta una optimización práctica para grandes conjuntos de datos.

La elección adecuada de la tasa de aprendizaje y el tamaño del minibatch es crucial para el rendimiento del algoritmo.