

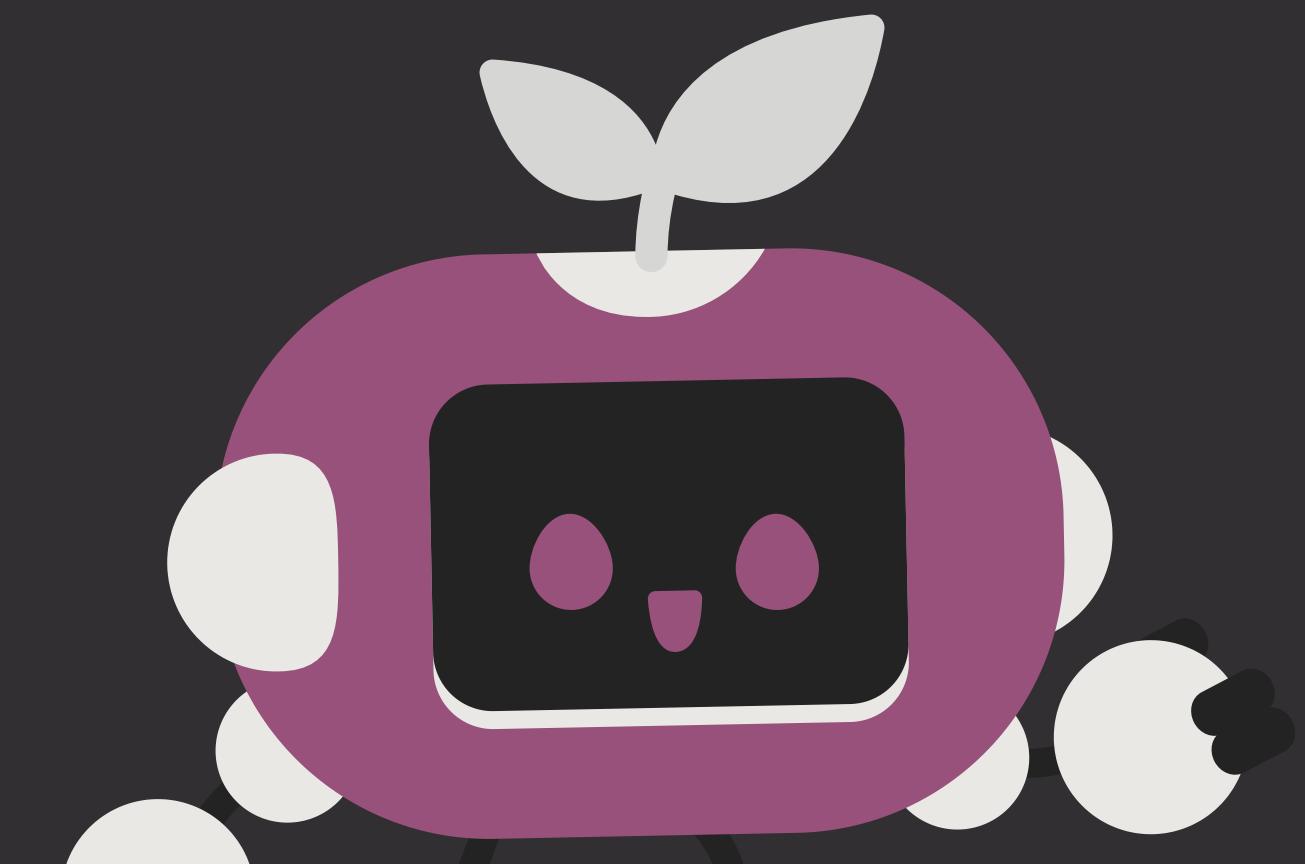


ESCUELA POLITÉCNICA NACIONAL
FIS
MÉTODOS NUMÉRICOS

DESCENSO DE GRADIENTE

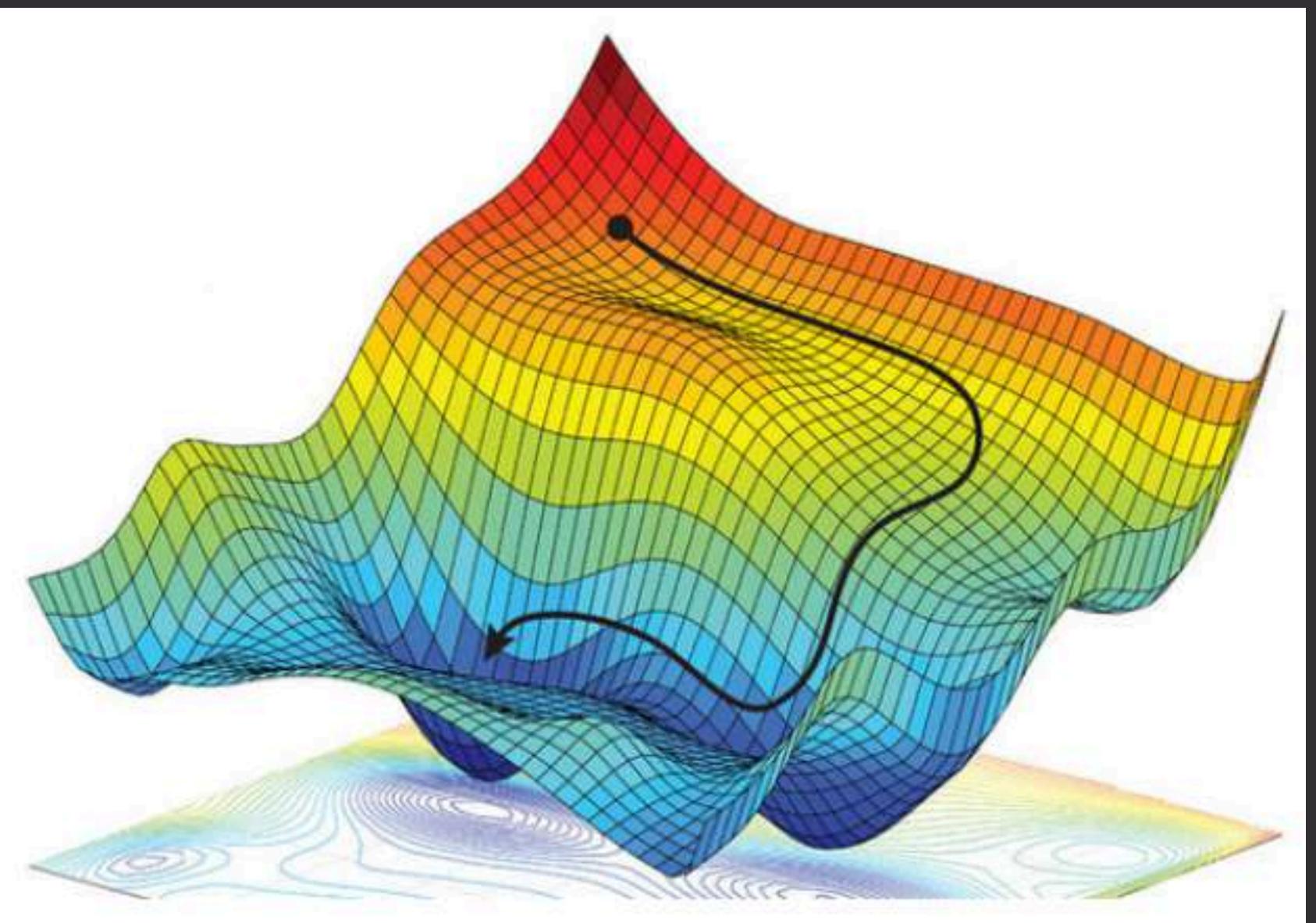
Integrantes:

- Lenin Amangandi
- Tamara Benavidez
- Camila Caicedo



INTRODUCCIÓN

El descenso del gradiente es un método de optimización de primer orden, ya que toma las primeras derivadas de la función de coste. Permite saber cómo ajustar los parámetros de la red de tal forma que minimice su desviación a la salida.



FORMULACIÓN MATEMÁTICA DEL ALGORITMO

1. ¿Qué problema se quiere resolver?

$$\min_{\theta} J(\theta)$$

Donde:

- θ : parámetros del modelo.
- $J(\theta)$: función de costo o pérdida

2. Derivada

$$\frac{dJ}{d\theta}$$

- (+) → la función está subiendo
- (-) → la función está bajando
- Su magnitud → qué tan empinada es la pendiente

3. Regla básica del descenso de gradiente

$$\theta_{nuevo} = \theta_{viejo} - \alpha \frac{dJ}{d\theta}$$

Donde:

- θ_{nuevo} : valor actual
- α : *learning rate* (tamaño del paso)
- $\frac{dJ}{d\theta}$: pendiente

FORMULACIÓN MATEMÁTICA DEL ALGORITMO

4. Extensión a múltiples dimensiones

$$\theta = \begin{pmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_n \end{pmatrix}$$

$$J(\theta_1, \theta_2, \dots, \theta_n)$$

5. El gradiente

$$\nabla J(\theta) = \begin{pmatrix} \frac{\partial J}{\partial \theta_1} \\ \frac{\partial J}{\partial \theta_2} \\ \vdots \\ \frac{\partial J}{\partial \theta_n} \end{pmatrix}$$

6. Regla general del descenso de gradiente

$$\theta^{(k+1)} = \theta^{(k)} - \alpha \nabla J(\theta^{(k)})$$

Donde:

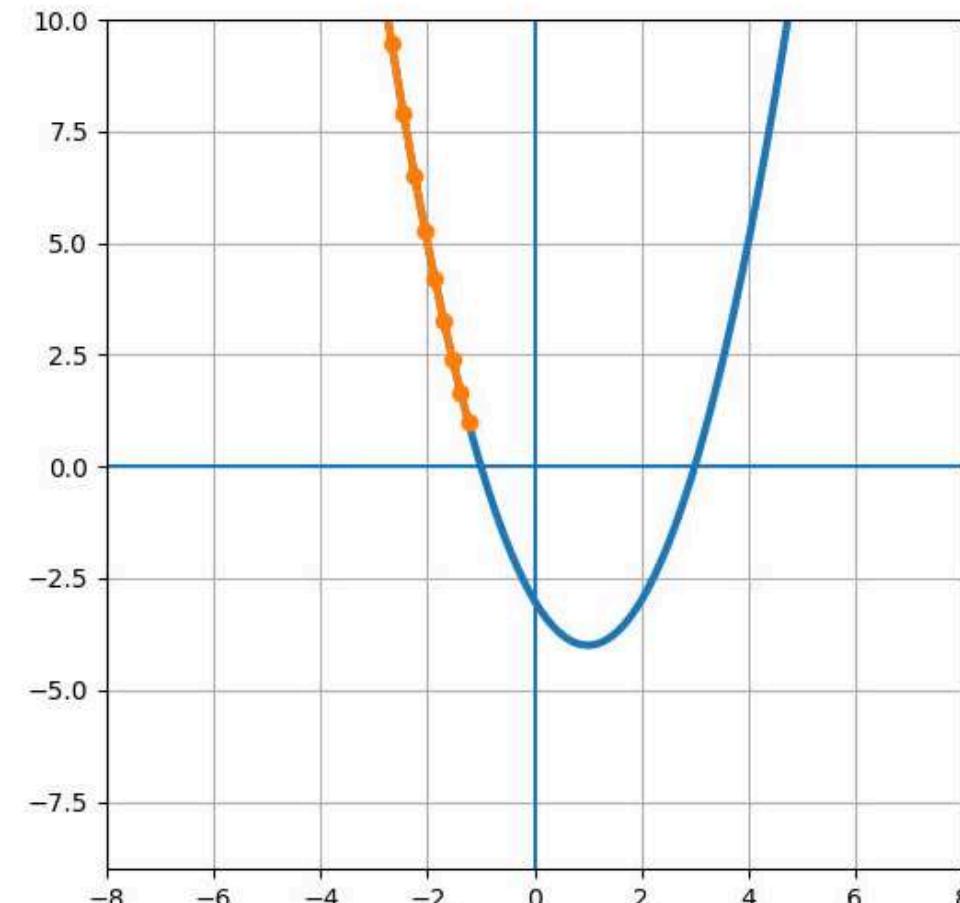
- k : iteración
- $\theta^{(k)}$: parámetros actuales
- ∇J : gradiente en ese punto

FORMULACIÓN MATEMÁTICA DEL ALGORITMO

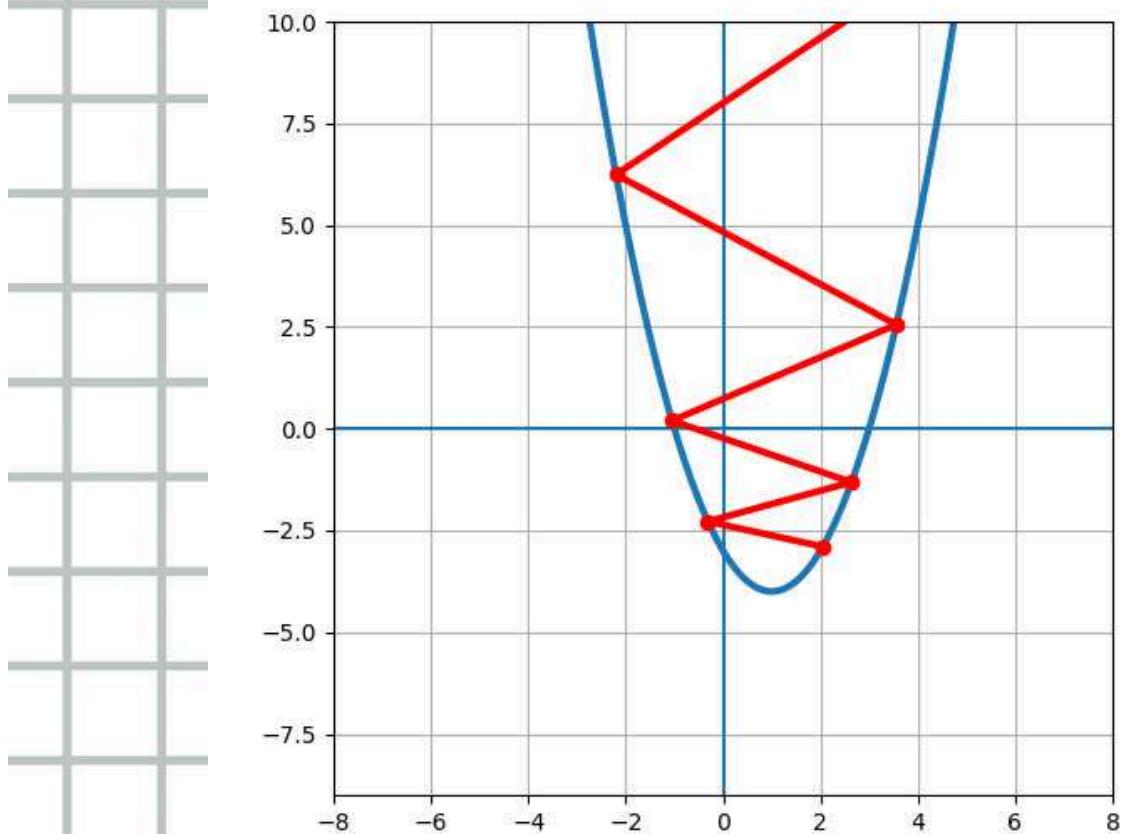


7. Constante de Tasa de Aprendizaje

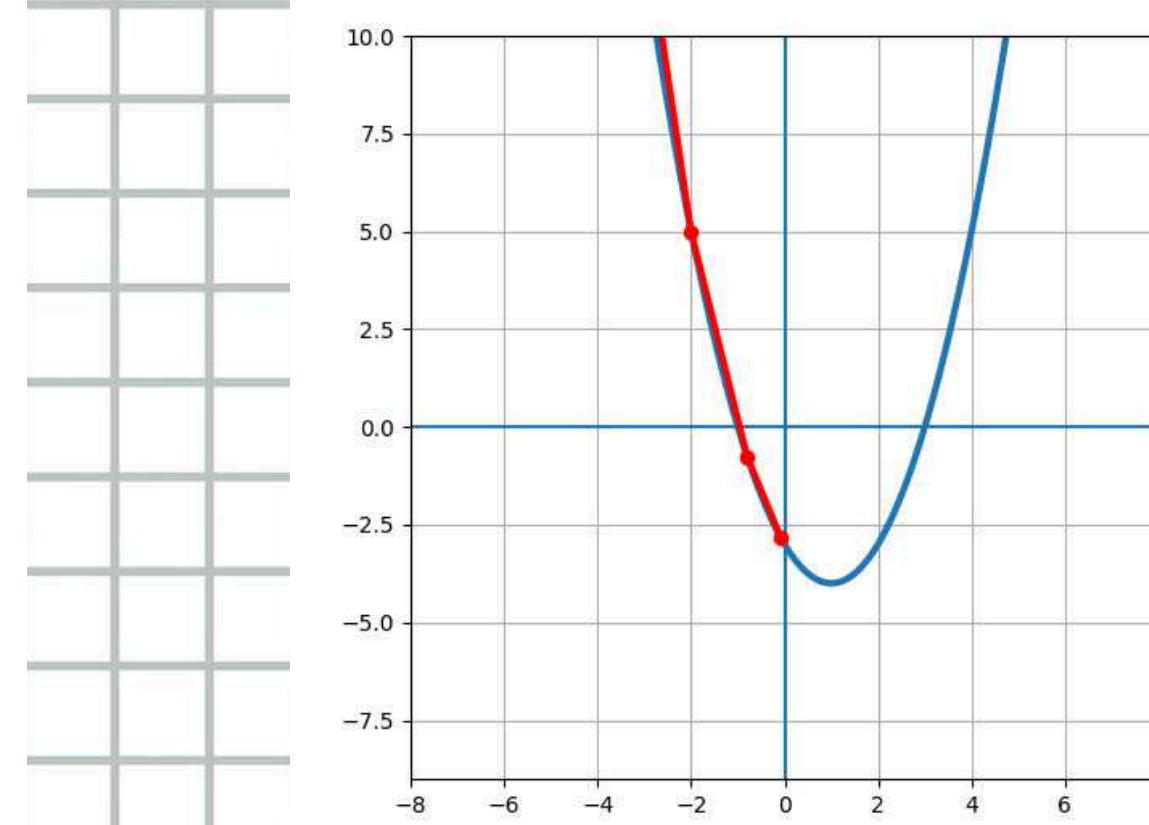
7.1 Corta



7.2 Grande



7.3 Óptima



COMPARACIÓN CON EL MÉTODO DE MÍNIMOS CUADRADOS



- El descenso de gradiente busca la solución paso a paso, mínimos cuadrados la obtiene directamente.
- El descenso de gradiente es más práctico para problemas grandes. Mínimos cuadrados se vuelve más costoso por el cálculo de la matriz inversa.
- El descenso de gradiente requiere ajustes previos, como la tasa de aprendizaje y número de iteraciones, mínimos cuadrados no.
- Mínimos cuadrados es exacto; descenso de gradiente es aproximado pero controlable.
- El descenso de gradiente es más intuitivo y visual, mínimos cuadrados es más abstracto y matemático.
- El descenso de gradiente utiliza menos memoria, mientras que mínimos cuadrados exige más recursos.

VENTAJAS

Se puede observar cómo disminuye el error en cada paso, lo que facilita el análisis del comportamiento del modelo y detectar problemas durante el entrenamiento.

Se puede detener cuando el error deja de mejorar, lo que puede ahorrar tiempo de cómputo sin necesidad de llegar a una solución exacta.

Se puede entrenarse de manera incremental, actualizando el modelo al llegar nuevos datos, sin rehacer todo el cálculo.

DESVANTAJAS

Depende fuertemente de la inicialización, si los valores iniciales son poco adecuados puede ralentizar la convergencia o llevar a soluciones poco eficientes.

En funciones más complejas, el algoritmo puede quedarse en una solución que no es la mejor posible.

El ruido puede afectar la trayectoria de descenso y producir actualizaciones poco estables.

CONCLUSIONES

- Se logró formular matemáticamente el algoritmo de descenso de gradiente, identificándolo como un método iterativo de primer orden que utiliza el gradiente de la función objetivo para aproximar soluciones de problemas de optimización.
- Se pudo comparar el descenso de gradiente con el método de mínimos cuadrados, destacando que el primero obtiene soluciones aproximadas de manera iterativa, mientras que el segundo entrega una solución exacta en problemas lineales.
- Se logró identificar las ventajas y limitaciones de ambos métodos, resaltando la eficiencia del descenso de gradiente en problemas grandes y las dificultades computacionales del método de mínimos cuadrados.

QUIZ

2/7

¿Qué información utiliza el descenso de gradiente para avanzar?

La matriz inversa

El valor de la función

El gradiente de la función

La segunda derivada

[https://wayground.com/admin
/quiz/695aebcf52303a8272926](https://wayground.com/admin/quiz/695aebcf52303a8272926)
746



GRACIAS

