UNIVERSIDAD GALILEO

Postgrado en Análisis y Predicción de Datos

Curso: Algoritmos en la Ciencia de Datos Horario: lunes: 18:00 – 21:00

Tutor: PhD. Alberth Alvarado

**Laboratorio No. 2**

|  |
| --- |
| Edgar Geovany Ocaña Orozco – 24010004 |
| Andrea María Hernández Marroquín – 240110074 |

Guatemala, 15 de septiembre de 2024

**2. SECCIÓN DOCUMENTAL DE EXPERIMENTACIÓN**

**1) Solución Cerrada**

**Tabla

Descripción generada automáticamente**

**2) GD**

**Interfaz de usuario gráfica, Gráfico

Descripción generada automáticamente**

**Tabla

Descripción generada automáticamente**

**Imagen que contiene Interfaz de usuario gráfica

Descripción generada automáticamente**

**Tabla

Descripción generada automáticamente**

**Texto

Descripción generada automáticamente**

**Gráfico

Descripción generada automáticamente**

**Interfaz de usuario gráfica, Tabla, Excel

Descripción generada automáticamente**

**Interfaz de usuario gráfica, Aplicación, Tabla, Excel

Descripción generada automáticamente**

**Tabla

Descripción generada automáticamente**

**3. SECCIÓN DOCUMENTAL DE CONCLUSION**

**1) ¿Por qué en la práctica no se utiliza el método de solución cerrada?**

La solución cerrada, aunque teóricamente perfecta porque da la respuesta exacta en una operación, tiene limitaciones significativas en la práctica, especialmente en contextos de datos grandes:

1. Complejidad Computacional
2. Requerimientos de Memoria
3. Sensibilidad a Ruido y Outliers

**Comente cómo el step size afecta la convergencia del algoritmo GD y con cuál step size constante obtuvo el “mejor” resultado.**

El tamaño del paso de 0.0005 pareció proporcionar un buen equilibrio entre velocidad de convergencia y estabilidad, alcanzando valores bajos de la función objetivo sin mostrar signos de divergencia incontrolable.

**Comente cómo el step size afecta la convergencia del algoritmo SGD y con cuál step size constante obtuvo el “mejor” resultado.**

El step size de 0.005, aunque variable, parece ofrecer el mejor equilibrio entre velocidad y estabilidad, pero con una recomendación de ser ligeramente reducido para disminuir la variabilidad y las oscilaciones observadas. Es vital realizar más pruebas con tasas ligeramente menores para optimizar la convergencia.

Implementar técnicas adicionales como decaimiento adaptativo del step size o métodos de optimización avanzados (como Adam o RMSprop) que ajustan el step size de manera adaptativa podría ser beneficioso para controlar las oscilaciones y mejorar la estabilidad y velocidad de convergencia del SGD.

**Comente cómo el step size afecta la convergencia del algoritmo MBGD, puede experimentar con otro step sizes para justificar su respuesta. ¿Con cuál step size constante obtuvo el “mejor” resultado?**

Basado en los datos proporcionados, el step size de 0.0005 parece ofrecer el mejor equilibrio entre estabilidad y progreso hacia la convergencia, aunque para aplicaciones donde el tiempo de convergencia es crítico, un ajuste fino del step size en torno a 0.005 podría explorarse con precaución para mejorar la eficiencia sin comprometer demasiado la estabilidad.

Sería beneficioso implementar técnicas como **decaimiento del step size** o **ajustes adaptativos del step size** para manejar mejor la variabilidad y la convergencia, especialmente en el caso de step sizes mayores como 0.005.

**Compare el desempeño de cada uno de los métodos implementados.**

**Solución Cerrada:** Tiene un valor de función objetivo comparativamente bajo y un error en xxx de 0.16, pero no requiere iteraciones, lo que la hace ideal desde el punto de vista de la convergencia inmediata. Sin embargo, como discutido previamente, puede ser inviable para grandes dimensiones o datos correlacionados debido a problemas de escalabilidad y precisión.

**GD:** Varias entradas con diferentes tasas de aprendizaje. En general, parece que mientras más baja es la tasa de aprendizaje, más estables son los resultados, pero a costa de requerir más iteraciones para converger.

**MBGD:** Muestra una amplia gama de resultados, dependiendo del tamaño del lote y la tasa de aprendizaje. En general, los tamaños de lote más pequeños y las tasas de aprendizaje más **bajas** parecen ofrecer un equilibrio entre estabilidad y eficiencia, aunque algunas configuraciones muestran una alta variabilidad en la función objetivo.

**SGD:** Presenta una variabilidad considerable en el valor de f(x)f(x)f(x), lo que es típico para este método debido a su naturaleza estocástica. Las tasas de aprendizaje más altas parecen causar una mayor variabilidad e incluso divergencia en algunos casos.