Convergencia de los parámetros theta\_0 y theta\_1

El descenso de gradiente logra aproximarse muy bien a la solución analítica porque la función de costo es convexa en el caso de la regresión lineal simple, lo que significa que el algoritmo siempre converge a un mínimo global (si está bien configurado).

Para obtener resultados más precisos basta con probar con diferentes tasas de aprendizaje de alfa y métodos como el descenso de gradiente adaptativo (Adam, RMSprop) para mejorar la convergencia en problemas más complejos.

Comparación con implementación de bucles separados

Se realizó una prueba donde theta\_0 y theta\_1 se actualizan en bucles separados en lugar de simultáneamente. En este caso, los valores finales fueron incorrectos. Cuando los parámetros se actualizan por separado, se pierde la sincronización en la corrección de la dirección del gradiente. El segundo parámetro se ajusta utilizando el nuevo valor del primero (ya modificado), en lugar de usar los valores originales. Esto genera una dinámica de actualización inestable y puede impedir la convergencia o llevar a soluciones erróneas por lo que tenemos que asegurarnos que las actualizaciones se hagan simultáneamente dentro de cada iteración del descenso de gradiente o explorar otros métodos como el método de Newton, que actualiza los parámetros de manera más eficiente.

Evaluación de la historia de costos (cost function)\*\*

La función de costo theta muestra una reducción progresiva, lo cual indica que el modelo está aprendiendo adecuadamente.La reducción progresiva confirma que el algoritmo está minimizando el error cuadrático medio (MSE) correctamente. Sin embargo, la velocidad de descenso depende de la tasa de aprendizaje de alfa.

Como soluciones alternativas se puede investigar si el algoritmo podría converger más rápido usando tasas de aprendizaje adaptativa, como en AdaGrad o Adam.

Generalización a otros problemas

Se aplicó la regresión logistica en un conjunto de datos de diabetes a pesar de tener buenos resultados en el modelo podriamos investigar qué tan bien se comporta el modelo en términos de colinealidad de variables, normalización y estandarización