

Resumen

El artículo "Comparing the Moore-Penrose and Gradient Descent for solving linear Regression Problems: A Performance Analysis" compara dos métodos fundamentales para resolver regresiones lineales: la pseudoinversa de Moore-Penrose y el descenso del gradiente. El objetivo principal es determinar en qué condiciones cada método resulta más eficiente, preciso y estable al estimar los coeficientes de un modelo lineal.

La regresión lineal busca ajustar una función lineal que minimice la suma de los errores al cuadrado. Este problema puede resolverse mediante una solución analítica exacta (la pseudoinversa) o una solución iterativa (descenso del gradiente). La pseudoinversa proporciona el resultado óptimo directamente incluso cuando las matrices no son invertibles aunque puede volverse costosa y sensible numéricamente en matrices muy grandes o mal condicionadas. Por su parte, el descenso del gradiente ajusta los parámetros progresivamente a través de múltiples iteraciones dependiendo de la tasa de aprendizaje, la condición numérica del conjunto de datos y los criterios de parada.

Para comparar ambos enfoques, el autor realizó experimentos con datos sintéticos y conjuntos reales. En los datos sintéticos, se variaron el número de observaciones, el número de características y el nivel de condicionamiento de las matrices. Se midieron tres métricas: tiempo de ejecución, error cuadrático medio (MSE) y número de iteraciones. Los resultados mostraron que la pseudoinversa fue consistentemente más rápida y precisa, mientras que el descenso del gradiente requirió mucho más tiempo y menos precisión.

En conjuntos reales, se observaron las mismas tendencias: la pseudoinversa alcanzó soluciones exactas y estables, mientras que el descenso del gradiente dependió fuertemente de la normalización de las variables y de la correcta elección de los hiperparámetros.

En conclusión, es importante el condicionamiento de los datos como factor determinante del rendimiento del descenso del gradiente y propone como líneas futuras de investigación la incorporación de métodos regularizados como Ridge o Lasso, el uso de optimizadores avanzados como SGD, Adam y el diseño de enfoques híbridos que combinen la exactitud de la pseudoinversa con la escalabilidad del descenso del gradiente.

Comparación entre pseudoinversa y descenso del gradiente

Evaluaremos el rendimiento, precisión y consumo de memoria de los métodos, pseudoinversa y descenso de la gradiente.

Se utilizó un script en R para generar matrices sintéticas y vectores de respuesta siguiendo el modelo lineal con 10 variables. Los métodos implementados fueron:

- Pseudoinversa de Moore-Penrose
- Descenso del gradiente por lote Batch GD
- Descenso del gradiente estocástico SGD

se midieron tres métricas principales:

- tiempo de ejecución
- error cuadrático medio MSE
- uso aproximado de memoria

Resultados

escenario	método	tiempo	MSE	memoria
$n=1000, d=10$	pseudoinversa	0.0012s	0.0099	10 MB
	gradiente BGD	0.11s	0.0102	12 MB
$n=100000, d=10$	pseudoinversa	Falla	-	> 80 GB
	gradiente SGD	18.4s	0.0113	150 MB

Interpretación:

en caso pequeño $n=1000$ la pseudoinversa fue extraordinariamente rápida y precisa. el descenso del gradiente alcanzó una precisión similar pero tardó más que la pseudoinversa.

en el caso grande $n=100000$ la pseudoinversa se vuelve inviable por su enorme consumo de memoria. sus operaciones requieren almacenar más de 80 GB de memoria, en cambio el descenso del gradiente logró manejar el volumen sin tanta memoria y convergiendo en segundos.