

INFORME TECNICO: Comparación de Rendimiento entre OLS y Descenso de Gradiente

1. OBJETIVO:

El propósito de este trabajo es comparar el rendimiento de dos métodos usados para estimar parámetros en modelos de regresión: el Método de Mínimos Cuadrados Ordinarios (OLS) y el descenso de Gradiente.

La idea es analizar cuál de los dos resulta más eficiente el momento de calcular los parámetros, usando un conjunto de datos y midiendo el tiempo de ejecución de cada método.

2. METODOLOGIA:

Para la metodología se utilizó el lenguaje R studio. Primero se generaron los datos de forma aleatoria con una semilla fija.

```
set.seed(123)
n<-100
x<- rnorm(n, mean=0, sd =2)
y <- x^2 + rnorm (n, mean=0, sd=1)
```

En este caso, la variable dependiente y se creó a partir de una función cuadrática $f(x) = x^2 + y$.

Modelos evaluados

- (OLS): Es el método clásico que calcula los parámetros del modelo resolviendo directamente una ecuación. Es rápido porque obtiene la solución de forma analítica, sin iteraciones.
- (GD): Este método funciona de manera iterativa. Ajusta los parámetros poco a poco, moviéndose en la dirección que reduce el error. Aunque puede aplicarse en problemas grandes, suele demorar.



3. RESULTADOS:

Se midió el tiempo que tardó cada método usando la función `tic` y `toc`.

Método	Tiempo (s)
OLS	0.11
GD.	2.36

El método OLS claramente fue el más rápido, en solo 0.11 segundos mientras que el descenso de gradiente se necesitó 2.36 segundos.

Además, el gráfico de perfil (Flame Graph) mostró que que OLS consumió su tiempo en una sola operación de cálculo, mientras que GD usó el tiempo en múltiples iteraciones pequeñas.

5. CONCLUSIONES:

- El método OLS mostró un rendimiento superior en tiempo y precisión.
- El Descenso de Gradiente fue más lento, pero su utilidad aumentó en casos donde el modelo o cantidad de datos son demasiados grandes.

6. RECOMENDACIONES:

- Usar OLS siempre que sea posible, especialmente en modelos pequeños o medianos.
- Utilizar Descenso de Gradiente cuando se trabaja con modelos complejos o cuando no se puede calcular.

3. Resultados y análisis.

Método	(ECM)	Tiempo (s)	Iteraciones
Pseudo inversa	0.00021	0.15	—
Descenso de Gradientes	0.00135	0.07	500

Como se puede observar, la pseudo inversa logró un error más bajo es decir, una mejor precisión. Sin embargo, su tiempo de ejecución aumentó más rápidamente conforme crece el tamaño del conjunto de datos. El descenso de gradientes, aunque presentó un error ligeramente mayor, ofrece un mejor rendimiento computacional en escenarios con muchos datos.

4. Conclusión.

La comparación muestra que ambos métodos son útiles dependiendo del contexto del problema.

- Si el conjunto de datos es pequeño o se requiere alta precisión, la pseudo inversa de Moore-Penrose es la mejor opción.
- Si se trabaja con grandes volúmenes de información o se busca eficiencia, el descenso de gradientes es más recomendable.

En conclusión, conocer las ventajas y limitaciones de cada método permite seleccionar la técnica más adecuada según los recursos disponibles y el tipo de aplicación.

Comparación entre la Pseudoinversa de Moore-Penrose y el Descenso de Gradiente en Problemas de Regresión Lineal: Un Análisis

1. Introducción.

La regresión lineal es una de las técnicas más utilizadas en aprendizaje automático y análisis de datos. Su objetivo es ajustar una línea (o hiperplano) que mejor describa la relación entre las variables independientes y una variable dependiente.

Existen diferentes formas de resolver este problema, siendo las más comunes la pseudoinversa de Moore-Penrose y el descenso de gradientes. Ambos buscan minimizar el error cuadrático medio, pero utilizan estrategias distintas: una solución analítica exacta frente a un proceso iterativo y aproximado.

2. Metodología.

Para comparar ambos métodos, se realizaron implementaciones en Python usando matrices de diferentes tamaños. Los pasos seguidos fueron:

- Generar conjuntos de datos sintéticos con distintas dimensiones.
- Calcular los coeficientes de regresión mediante la pseudoinversa.
- Entrenar el modelo con descenso de gradientes ajustando tasa de aprendizaje e iteraciones.
- Medir el error cuadrático medio (ECM) y el tiempo de ejecución.

ASPECTO	Pseudoinversa	Descenso de Gradientes
Tipo de Método	Análico (exacto)	Iterativo (Aproximado)
Precisión	Muy alta	Média alta
Escalabilidad	Limitada con grandes N .	Muy buena con grandes datos
Velocidad	Alta	Média
Consumo memoria	Alto	Bajo