

Práctica de Laboratorio

Tópicos de Inteligencia Artificial ...

CComp9-1

Alexander Giuliano Pinto De la Gala
alexander.pinto@ucsp.edu.pe

Universidad Católica San Pablo

1 Introducción

El presente Informe describe la experiencia y resultados de la primera práctica de Laboratorio del curso de Tópicos de Inteligencia Artificial referida a Regresión Lineal.

La práctica consiste en la implementación de diversas rutinas para un modelo de Regresión Lineal de tres conjuntos de datos.

Para realizar un aprendizaje supervisado se decide una representación una hipótesis h en una computadora. Por lo tanto podemos aproximar y como una función lineal de x :

$$h_{\theta}(x) = \theta^T x = \theta_0 x_0 + \theta_1 x_1 + \dots + \theta_n x_n \quad (1)$$

donde $\theta_0, \theta_1, \theta_2, \dots, \theta_n$ son los parámetros, también llamados pesos, los cuales parametrizan el espacio de la función lineal mapeando de X a Y . Esta función puede ser simplificada así:

$$h(x) = \sum_{i=0}^d \theta_i x_i = \theta^T x \quad (2)$$

La función de costo a minimizar está definida por:

$$J(\theta) = \frac{1}{2m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2 \quad (3)$$

El algoritmo de Gradiente descendiente se representa como:

Repetir:

$$\theta_j = \theta_j - \alpha \frac{1}{m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)}) x_j^{(i)} \quad (4)$$

donde: α es la tasa de aprendizaje.

Se hace también la implementación de la resolución analítica de θ dada por la Ecuación Normal:

$$\theta = (X^T X)^{-1} X^T y \quad (5)$$

La métrica utilizada será el Error Cuadrático Medio definido por:

$$MSE = \frac{1}{n} \sum (y - \hat{y})^2 \quad (6)$$

donde:

n es el número de elementos a comparar

y los valores y reales del conjunto de datos

\hat{y} los valores y calculados a partir de nuestros modelos

2 Implementación

La implementación¹ fue realizada en Python 3.6, con uso de bibliotecas como numpy y pandas.

La data utilizada corresponde a tres *datasets* "Precio Viviendas" (`ex1data2(Home.1f)`), "Temperatura Océano" (`oceanosimple`) y "Consumo de Petróleo" (`petrolconsumption`).

La data fue preprocesada, se dividió cada conjunto de manera aleatoria creando dos subconjuntos. El primero denominado "train" con un 70% de la data del conjunto la cual servirá para el entrenamiento del modelo. El segundo subconjunto se denomina "test" con el 30% de la data restante, la cual servirá para las pruebas y validación de nuestro modelo.

La data fue normalizada siguiente una por medio de la puntuación estándar definida por:

$$\frac{X - \mu}{\sigma} \quad (7)$$

donde:

μ es la media aritmética

σ es la desviación estándar del conjunto de datos.

3 Experimentos y Resultados

3.1 Experimento 1

Consiste en el cálculo del Error Cuadrático MSE de la regresión lineal usando la Ecuación Normal para los conjuntos de entrenamiento y prueba de "Consumo de Petróleo", "Precio Viviendas" y "Temperatura Océano".

La tabla 1 resume los resultados. Se puede apreciar que el error de entrenamiento y prueba son similares para cada dataset, lo que indica que no tenemos un problema de *overfitting*. Solo para el caso del dataset Precio Viviendas el error del conjunto de prueba es mayor al del entrenamiento.

¹ El código de la implementación se encuentra disponible en <https://github.com/giulianodelagala/RegresionLineal>

Table 1: Error Cuadrático de la Regresión Lineal

Conjunto	Precio Viviendas	Temperatura Océano	Consumo de Petróleo
Entrenamiento	0.265738	0.158686	0.333132
Prueba	0.329316	0.139772	0.308714

3.2 Experimento 2

Consiste en la búsqueda de los mejores parámetros de entrenamiento para los conjuntos "Consumo de Petróleo", "Precio Viviendas" y "Temperatura Océano" usando Regresión Lineal con Gradiente Descendiente. Se debe variar los parámetros de entrenamiento según la tabla 2:

Table 2: Parámetros para el Experimento 2

Parámetros	Valores
Tasa de Aprendizaje	0.01, 0.05, 0.1, 0.2, 0.3, 0.4
Iteraciones	[500,3500] en incrementos de 500

La tabla 3 resumen los resultados del experimento. Cada fila representa la variación de la tasa de aprendizaje, y las columnas la variación del número de iteraciones.

De la tabla 3 se puede apreciar que la convergencia se produce luego de la iteración número 500, para casi todas las tasas de aprendizaje. Sólo esto no se cumple cuando la tasa de aprendizaje es muy pequeña (0.01), para este caso se necesita un mayor número de iteraciones para que el error se estabilice, en los datasets Precio de Viviendas y Temperatura Océano se tiene valores estables luego de 1500 iteraciones y en el dataset Consumo de Petróleo al menos 3000 iteraciones. Luego de analizar la tabla una tasa de aprendizaje de 0.10 para un número de iteraciones igual a 500, son parámetros adecuados.

3.3 Experimento 3

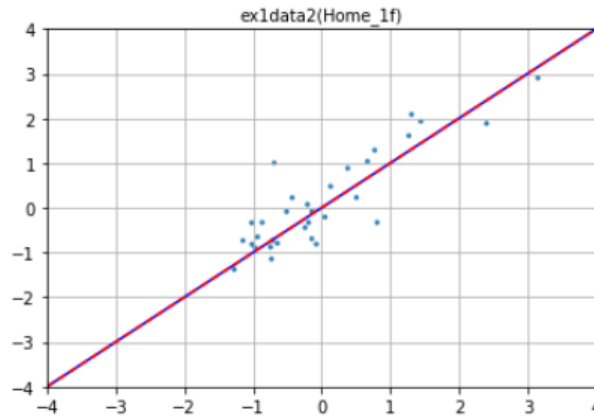
Para el conjunto "Precio Viviendas", se debe plotear los ejemplos de entrenamiento y graficar las rectas que forman θ obtenidas al usar la Regresión Lineal con la Ecuación Normal y el Gradiente Descendiente.

Se tiene como vector $\theta = [0.11919, 0.908149]$, como resultados tanto de la Ecuación Normal como de la Gradiente Descendiente, por lo que las rectas se superponen, como se puede apreciar en el gráfico 1.

Table 3: Error Cuadrático de la Regresión Lineal para diferentes parámetros

Precio de Viviendas							
	500	1000	1500	2000	2500	3000	3500
0.01	0.327403	0.329307	0.329316	0.329316	0.329316	0.329316	0.329316
0.05	0.329316	0.329316	0.329316	0.329316	0.329316	0.329316	0.329316
0.10	0.329316	0.329316	0.329316	0.329316	0.329316	0.329316	0.329316
0.20	0.329316	0.329316	0.329316	0.329316	0.329316	0.329316	0.329316
0.30	0.329316	0.329316	0.329316	0.329316	0.329316	0.329316	0.329316
0.40	0.329316	0.329316	0.329316	0.329316	0.329316	0.329316	0.329316
Temperatura Océano							
	500	1000	1500	2000	2500	3000	3500
0.01	0.139973	0.139774	0.139772	0.139772	0.139772	0.139772	0.139772
0.05	0.139772	0.139772	0.139772	0.139772	0.139772	0.139772	0.139772
0.10	0.139772	0.139772	0.139772	0.139772	0.139772	0.139772	0.139772
0.20	0.139772	0.139772	0.139772	0.139772	0.139772	0.139772	0.139772
0.30	0.139772	0.139772	0.139772	0.139772	0.139772	0.139772	0.139772
0.40	0.139772	0.139772	0.139772	0.139772	0.139772	0.139772	0.139772
Consumo de Petróleo							
	500	1000	1500	2000	2500	3000	3500
0.01	0.311688	0.308750	0.308708	0.308712	0.308713	0.308714	0.308714
0.05	0.308713	0.308714	0.308714	0.308714	0.308714	0.308714	0.308714
0.10	0.308714	0.308714	0.308714	0.308714	0.308714	0.308714	0.308714
0.20	0.308714	0.308714	0.308714	0.308714	0.308714	0.308714	0.308714
0.30	0.308714	0.308714	0.308714	0.308714	0.308714	0.308714	0.308714
0.40	0.308714	0.308714	0.308714	0.308714	0.308714	0.308714	0.308714

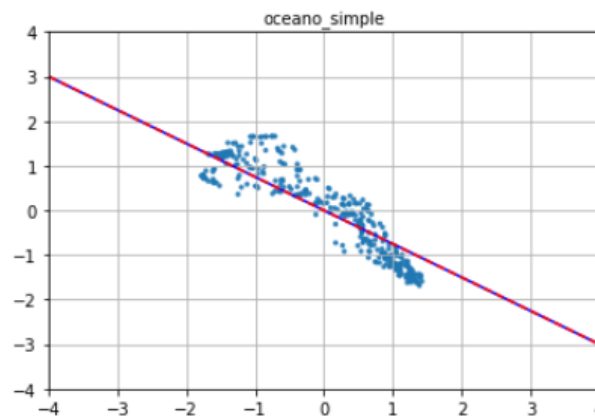
Fig.1: Ejemplos de entrenamiento y rectas de Regresión Lineal y Gradiente Descendiente conjunto "Precio Vivienda"



La superposición de las rectas indica que el Gradiente Descendiente ha hallado una respuesta adecuada. Se puede apreciar visualmente que la recta representa bien al conjunto de datos.

El mismo comportamiento se puede observar para el conjunto "Temperatura Océano" (Figura 2), esto como parte adicional fuera de experimento.

Fig. 2: Ejemplos de entrenamiento y rectas de Regresión Lineal y Gradiente Descendiente conjunto "Temperatura Océano"



3.4 Experimento 4

Para los conjuntos "Consumo de Petróleo", "Precio Viviendas", "Temperatura Océano", muestre tres gráficas de la función de costo del conjunto de entrenamiento y conjunto de prueba de la Regresión Lineal.

Los resultados para el experimento se muestran de la figura 3 a la figura 5

Las gráficas muestran el comportamiento de la función costo para los tres datasets, considerando una tasa de aprendizaje de 0.05 y 500 iteraciones.

Como se pueden apreciar de las figuras, la Curva de la Función Costo tienen un comportamiento esperado, por lo que podemos asegurar que la implementación es correcta ya que estamos convergiendo correctamente. Se puede apreciar que el costo se estabiliza en todos los casos antes de llegar a las 100 iteraciones, por lo que se podría disminuir el número de iteraciones y reducir el costo computacional.

Tanto las curvas de Train como de Test para todos los conjuntos tienen el mismo comportamiento, y se llega a estabilizar a valores similares a los hallados en el experimento 1, por lo que es otra manera de comprobar que nuestro modelo de Regresión Lineal con Gradiente Descendiente tiene un comportamiento esperado.

Las conclusiones del experimento 1 también aplican en este caso, en cuanto las curvas de Train y Test tienen comportamiento y estabilización similar.

Fig. 3: Función de Costo del conjunto "Precio Viviendas"

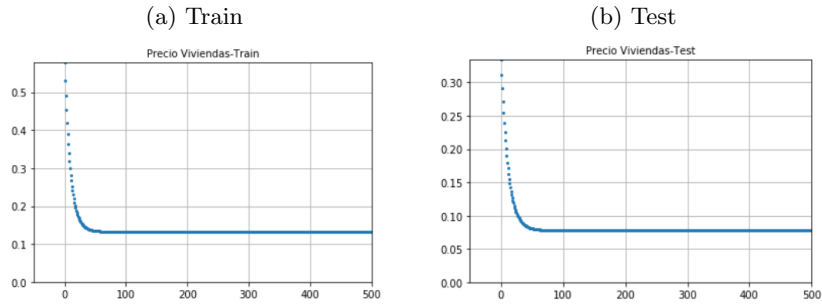


Fig. 4: Función de Costo del conjunto "Temperatura Océano"

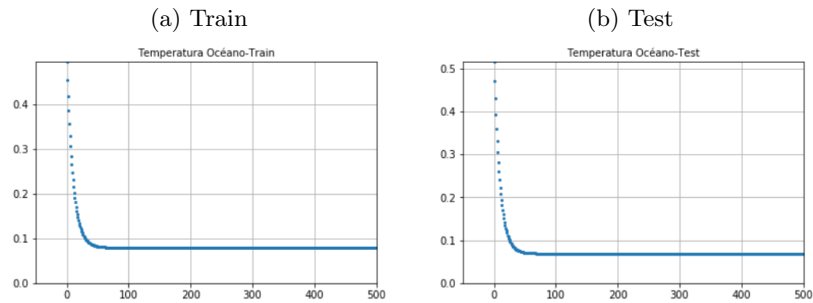
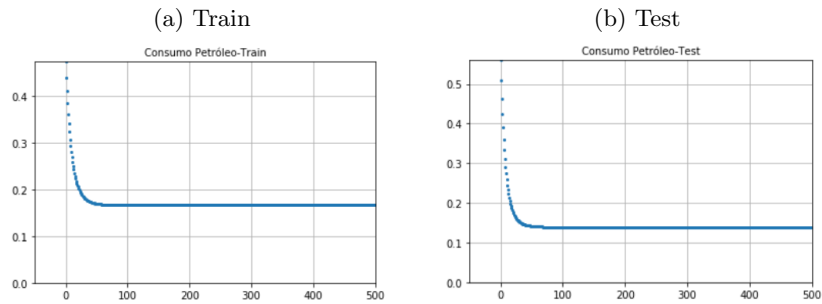


Fig. 5: Función de Costo del conjunto "Consumo de Petróleo"



4 Conclusiones

La presente ha descrito los experimentos de la Práctica de Laboratorio 1 del Curso de Tópicos de Inteligencia Artificial. Los diversos experimentos han demostrado que se consigue un comportamiento esperado para las implementa-

ciones de las funciones de Regresión Lineal. La comparación de resultados que ofrece la Ecuación Normal y el método de Gradiente Descendiente muestran que se ha logrado una implementación adecuada. Las gráficas de ploteo de entrenamiento y la recta conseguido por Regresión Lineal, muestran que la recta describe de manera esperada cada conjunto. Las gráficas de la Función de Costo para la Gradiente Descendiente, señalan que la convergencia se produce en relativamente pocas iteraciones, por lo que se podría reducir el costo computacional al disminuir este valor.

References

1. Andrew Ng, CS229 Lecture Notes