

UNIVERSIDAD DE GUADALAJARA



CENTRO UNIVERSITARIO DE CIENCIAS
EXACTAS E INGENIERÍA
DEPARTAMENTO DE
CIENCIAS
COMPUTACIONALES

Seminario de Solución de Problemas de
Sistemas Basados en Conocimiento

Práctica No. 5

Regresión Polinomial Simple

Nombre: Hurtado González Edgar Arturo
Código: 212597894

Introducción

La regresión polinomial simple es una técnica utilizada en estadísticas para modelar la relación entre una variable independiente x y una variable dependiente y mediante un polinomio. A diferencia de la regresión lineal, que utiliza una función lineal para modelar los datos, la regresión polinomial permite ajustar curvas más complejas al conjunto de datos al incluir términos de mayor grado.

En esencia, en lugar de ajustar una línea recta a los datos, la regresión polinomial ajusta una curva polinómica. Por ejemplo, un modelo de regresión polinomial de segundo grado tiene la forma $y = b_0 + b_1x + b_2x^2$, donde b_0 , b_1 y b_2 son coeficientes que se estiman a partir de los datos.

Este enfoque puede capturar relaciones no lineales entre variables, lo que lo hace útil cuando la relación entre x e y no es lineal. Sin embargo, a medida que se aumenta el grado del polinomio, el modelo puede volverse demasiado

flexible y sobreajustarse a los datos de entrenamiento, lo que lleva a un mal rendimiento al predecir nuevos datos (sobreajuste). En estos casos, es importante encontrar un equilibrio entre la complejidad del modelo y su capacidad para generalizar a nuevos datos.

La regresión polinomial es una extensión de la regresión lineal simple que permite modelar relaciones no lineales entre variables. En lugar de ajustar una línea recta a los datos, se ajusta un polinomio de grado n a los puntos de datos.

La fórmula general de la regresión polinomial de grado n es: $y = b_0 + b_1x + b_2x^2 + \dots + b_nx^n$. Aquí, y es la variable dependiente, x es la variable independiente, b_0, b_1, \dots, b_n son los coeficientes que se estiman a partir de los datos y n es el grado del polinomio.

Este método puede ser útil cuando la relación entre las variables no puede describirse adecuadamente con una línea recta. Sin embargo, un punto clave es el balance entre la complejidad del modelo y la capacidad de generalización.

- *Sobreajuste:* A medida que se aumenta el grado del polinomio, el modelo se vuelve más flexible y puede ajustarse mejor a los datos de entrenamiento. Sin embargo, también existe un mayor riesgo de sobreajuste, donde el modelo se adapta demasiado a las peculiaridades de los datos de entrenamiento y no puede generalizar bien a nuevos datos.
- *Selección del grado del polinomio:* En la práctica, se busca un equilibrio. Se pueden utilizar técnicas como la validación cruzada o métodos de información (como el criterio de información de Akaike o el criterio de información bayesiano) para seleccionar el grado del polinomio que mejor se ajuste a los datos y tenga una buena capacidad predictiva en nuevos datos.
- *Interpretación:* Mientras que la regresión lineal simple ofrece coeficientes directamente interpretables (la pendiente y la intersección), la regresión polinomial puede ser más compleja de interpretar debido a la presencia de múltiples términos de grado superior.

Es crucial tener en cuenta que, aunque la regresión polinomial puede ser útil para modelar relaciones no lineales, no es la solución para todos los problemas. En ocasiones, otras técnicas como modelos no lineales específicos o transformaciones de variables pueden ser más adecuadas para capturar la relación entre las variables.

La regresión polinomial es una herramienta poderosa para modelar relaciones no lineales entre variables. Permite capturar patrones más complejos en los datos, ofreciendo flexibilidad para ajustarse a curvas más complejas que una simple línea recta.

Desarrollo

El código implementado en esta ocasión realiza una regresión polinómica de grado 4 para predecir la velocidad máxima de un coche en función de su precio. Te lo explico paso a paso:

1. *Carga de datos:* Lee un archivo CSV que contiene datos de dos columnas: el precio del coche y su velocidad máxima.
2. *Funciones de regresión:*
 - `polynomial_regression`: Calcula la regresión polinómica de grado especificado utilizando el método de mínimos cuadrados.
3. *Cálculo del error:* Define una función para calcular el error cuadrático medio entre los valores reales y predichos.
4. *Validación cruzada Leave One Out:*
 - Divide el conjunto de datos en entrenamiento y prueba, iterando a través de cada punto de datos como conjunto de prueba una vez.
 - Para cada iteración, entrena el modelo de regresión polinómica de grado 4 con el conjunto de entrenamiento y calcula el error utilizando el conjunto de prueba.
5. *Búsqueda del mejor modelo:*
 - Almacena los errores calculados para cada iteración de Leave One Out.
 - Encuentra el modelo con el error cuadrático medio más bajo y guarda sus coeficientes.
6. *Resultados:*
 - Imprime el error cuadrático medio promedio y los coeficientes del mejor modelo encontrado.
7. *Visualización:*
 - Grafica los puntos originales (precio del coche vs. velocidad máxima) y la línea de regresión polinómica de grado 4 obtenida con los mejores coeficientes encontrados.

Conclusión

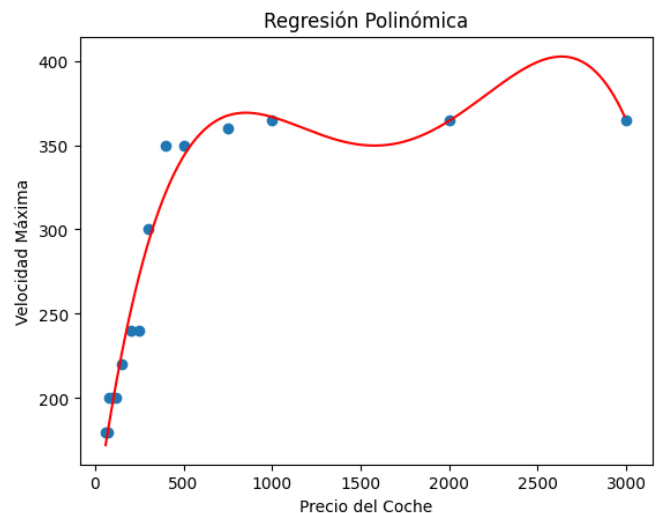
En conclusión, la regresión polinómica es una técnica valiosa, pero su aplicación requiere precaución y un entendimiento claro de cómo seleccionar el grado del polinomio, controlar el sobreajuste y evaluar la capacidad de generalización del modelo. Considerar otras alternativas y técnicas también es importante para obtener resultados precisos y significativos.

El código realiza una regresión polinómica de grado 4 para predecir la velocidad máxima de un coche en función de su precio, usando validación cruzada Leave One Out para encontrar el mejor modelo y luego visualiza los resultados obtenidos.

Resultados

MSE Promedio: 516951.13969465654

Coefficientes del Mejor Modelo: [1.29594714e+02, 7.58590837e-01, -8.29258846e-04, 3.61276950e-07, -5.34744206e-11].



Referencias:

colaboradores de Wikipedia. (2023a, octubre 21). *Regresión polinomial*. Wikipedia, la enciclopedia libre. https://es.wikipedia.org/wiki/Regresi%C3%B3n_polinomial

Gonzalez, L. (2022, 7 septiembre). *Regresión polinomial – teoría*. Aprende IA. <https://aprendeia.com/algoritmo-regresion-polinomial-machine-learning/>

Statologos. (2021, 7 mayo). *Introducción a la regresión polinomial*. <https://statologos.com/regresion-polinomial/>