Polynomial Regression

Polynomial regression is a regression technique that allows the relationship between dependent variable(y) and independent variable(x) as modeled in the nth degree polynomial allowing a curved fit instead of straight line in linear regression.

formula: $\mathbf{y= b_0+b_1x_1+ b_2x_1^2+ b_2x_1^3+ \ldots\ldots b_nx_1}$

Need for polynomial Regression

1. In real world dataset,The relationship between variables are not linear,Therefore the polynomial regression helps to capture the complex patterns in the data thus providing a better fit than a simple linear model.

2.Relationship between variables can be better presented in curves than straight.Hence provide a more accurate prediction.

3.By adjusting the degree of polynomial helps in controlling the complexity of the model.Higher degree polynomial can fit the data more closely,but having a risk of overfitting the data.

Types of polynomial Regression

Linear: $ax + b = 0$,if degree=1

quadratic: $ax^2+bx + c = 0$, if degree=2

cubic:$ax^3 + bx^2 + cx = 0$,degree=3

It will continue on the basis of degree.

Equation of the polynomial Regression Model

Simple linear regression equation:y=$b_0 + b_1x$ (1)

Multiple linear regression equation:$y = b_0 + b_1x + b_2x_2 + b_3x_3 +\ldots + b_nx_n$ (2)

Polynomial Regression Equation: $y = b_0 + b_1x + b_2x^2 + b_3x^3 + ... + b_nx^n$ (3)

By comparing the above three equations it can be understood that all the equation are polynomial equation but vary by degree of variables.

The simple and multiple linear equations are polynomial equations with single degree,if we add degree to the equation it will convert to a polynomial linear equation.

Overfitting Vs Underfitting in Polynomial Regression

As the degree of the polynomial increases, the model may fit the training data too closely, capturing noise and outliers. This can lead to poor generalization to new, unseen data.

Linear analysis of a dataset may result in underfitting, where the model is too simplistic and fails to capture the true complexity of the data. Whereas the polynomial regression is able to correct underfitting by allowing flexible fits.

To prevent overfitting, add more training samples. This helps the algorithm avoid learning noise and promotes better generalization.

How to find the right degree of equation?

Inorder to find the right degree of equation to avoid underfitting and overfitting:

1.Forward selection:This method increases the degree until it is significant enough to define the best possible model.

2.Backward selection:This method decreases the degree until it is significant enough to define the best possible model.

Cost Function

Cost Function measures the performance of a machine learning model for given data.

Calculates error between predicted and expected values, presenting a single real number.

Cost function:Average error of n-sample data,used for entire training data.

Loss function:error of individual data points,used for one training set.

Cost function:$y = \frac{1}{n}\sum_{i=1}^{n}\left(y_{pred} - y_{i}\right)^{2}$

Polynomial Regression Reduces costs in the cost function by providing a curvilinear shape to the regression line.Enhances fitting for complex data, especially with higher-order polynomials.

The ideal Cost Function value is 0 or close to 0, indicating optimal model performance.

To approach the ideal Cost Function, Gradient Descent is applied.Gradient Descent updates weights iteratively, minimizing errors and enhancing model accuracy.

Gradient Descent

Gradient descent is an optimization algorithm used to find the values of parameters (coefficients) of a function that minimizes a cost function (cost).

The values of slope (m) and slope-intercept (b) will be set to 0 at the start of the function, and the learning rate (α) will be introduced.

The learning rate (α) is set to an extremely low number,

The learning rate is a tuning parameter in an optimization algorithm that determines the step size at each iteration to attain a minimum of a cost function.

update slope and intercept using the calculated derivatives and the learning rate.

Iterative process involves the following equations:

m=m−α⋅derivative of m

$b = b - \alpha \cdot \text{derivative of } b$

The process of updating the values of m and b continues until the cost function reaches the ideal value of 0 or close to 0.

The values of m and b now will be the optimum value to describe the best fit line.