Polynomial Regression Model

Polynomial regression is a type of regression analysis where the relationship between the independent variable (X) and the dependent variable (Y) is modeled as an nth degree polynomial. It is an extension of simple linear regression, which assumes a linear relationship between X and Y. Polynomial regression allows for more flexible modeling, accommodating curves and non-linear patterns in the data.

The general form of a polynomial regression model is:

$$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \beta_3 X^3 + \ldots + \beta_{\mathrm{n}} X^{\mathrm{n}} + \varepsilon$$

where Y is the dependent variable, X is the independent variable, $\beta_0, \beta_1, \beta_2, \ldots, \beta_n$ are the coefficients of the polynomial terms, n is the degree of the polynomial, and ε is the error term.

To fit a polynomial regression model, you need to determine the degree of the polynomial (n) and estimate the coefficients ($\beta_0, \beta_1, \beta_2, \ldots, \beta_n$) using a method like least squares regression. The least squares method minimizes the sum of squared differences between the actual Y values and the predicted Y values from the polynomial regression model.

Once you have estimated the coefficients, you can use the polynomial regression model to make predictions for new values of X by plugging them into the equation.

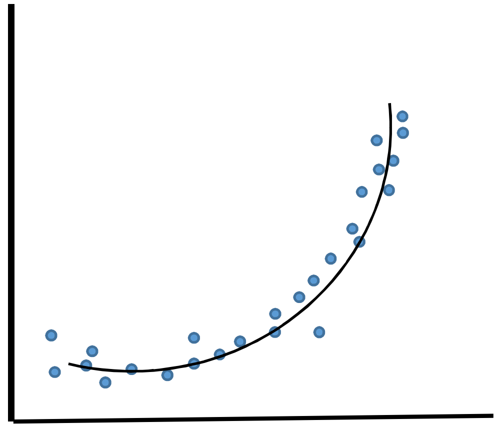
It's important to note that while polynomial regression can capture non-linear relationships, higher degree polynomials can lead to overfitting. Overfitting occurs when the model fits the training data too closely, resulting in poor generalization to new data. Therefore, it's essential to select an appropriate degree for the polynomial that balances the fit to the data and the risk of overfitting.

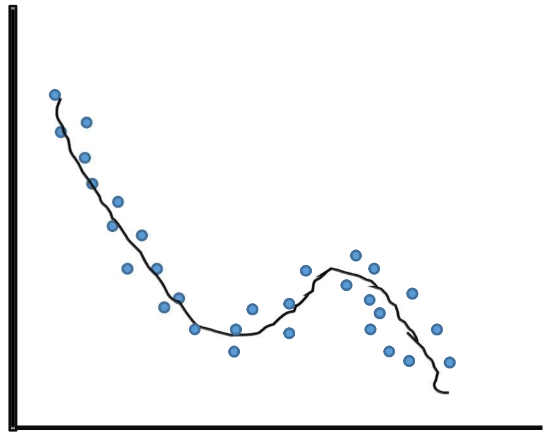
$$Y = \beta_0 + \beta_1 X$$

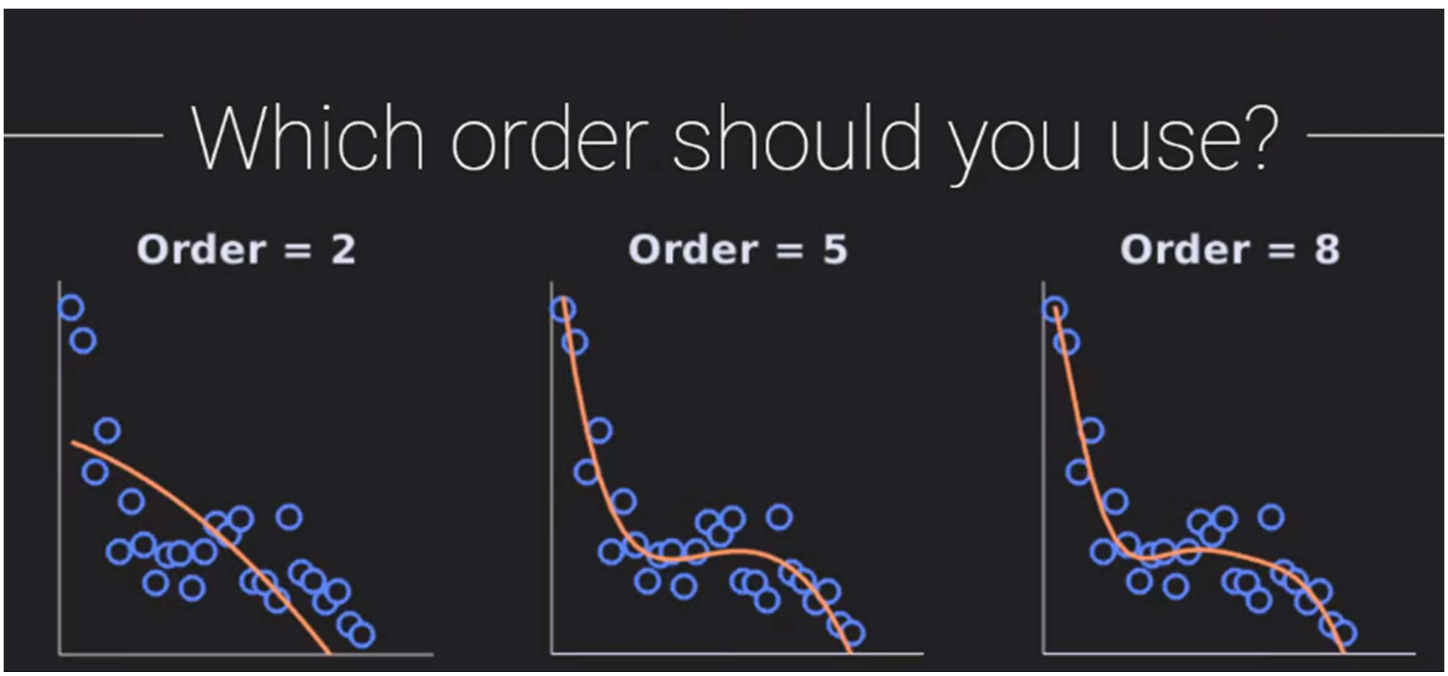
$$Y = \beta^0 X^0 + \beta_1 X + \beta_2 X^2 + \varepsilon$$

$$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \beta_3 X^3 + \ldots + \beta_{\mathrm{n}} X^{\mathrm{n}} + \varepsilon$$

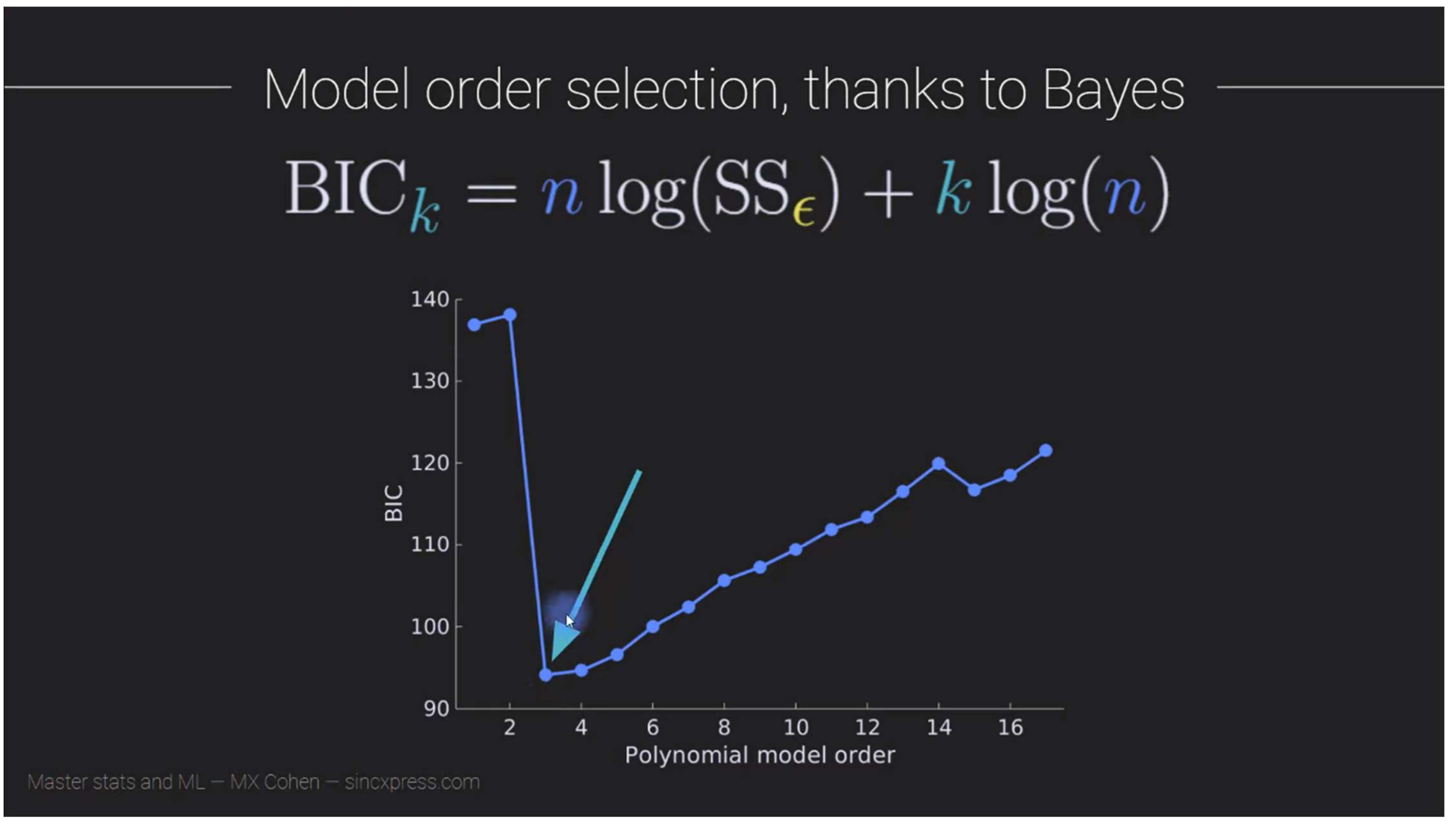
The coefficients are all linear, this is a standard linear model






Which order should you use?
Order = 2
Order = 5
Order = 8


Order = 11
Order = 14
Order = 17


Model order selection, thanks to Bayes
$\mathrm{BIC}_k = n\log(\mathrm{SS}_\epsilon) + k\log(n)$
BIC
140
130
120
110
100
90
2
4
6
8
10
12
14
16
Polynomial model order
Master stats and ML — MX Cohen — sincxpress.com