Class 7- Polynomial Regression



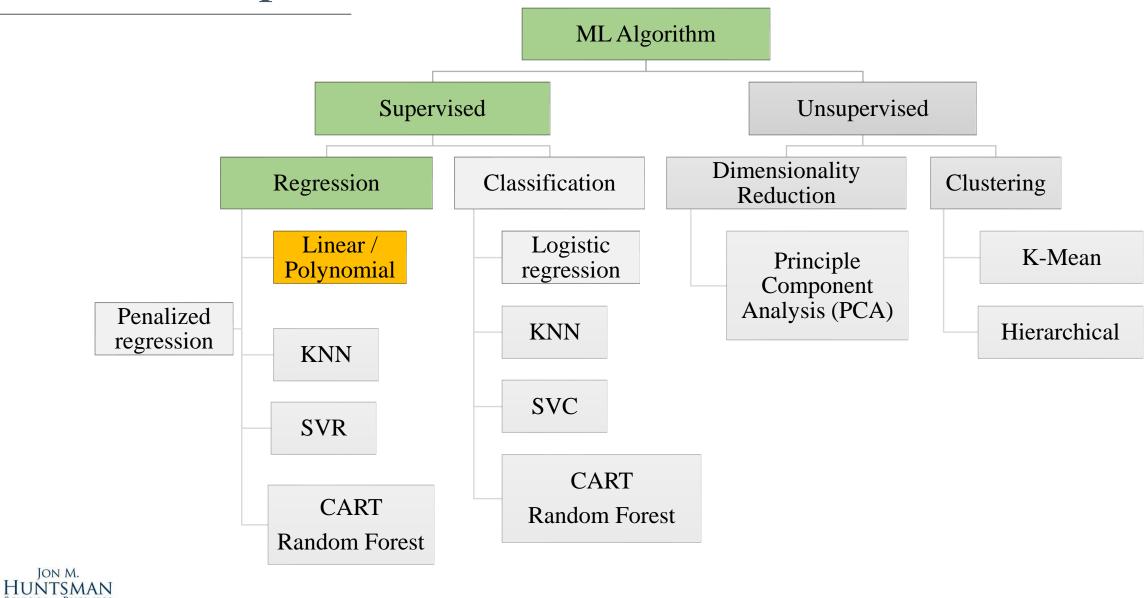
Prof. Pedram Jahangiry





UtahStateUniversity

Road map



Prof. Pedram Jahangiry



Polynomial regression model

The polynomial regression model is a <u>special case</u> of multiple linear regression models!

- Create new variables $X_1 = X$, $X_2 = X^2$, ... etc and then treat as multiple linear regression.
- Not really interested in the coefficients; more interested in the fitted function!

$$\hat{f}(X) = f_{w,b}(X) = b + w_1 x + w_2 x^2 + \dots + w_d x^d$$

- W is a d-dimensional vector of parameters
- b is a real number
- d is the polynomial degree of the model (we either fix the d at some reasonably low value, else use **cross-validation** to choose d



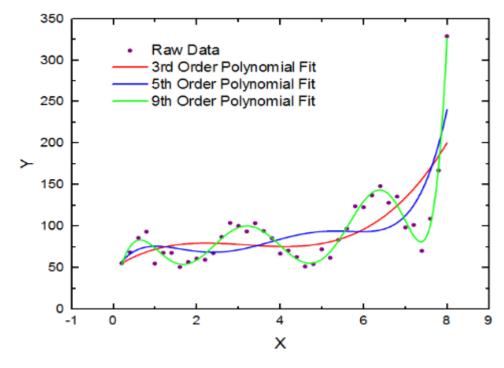


The optimization problem

The optimization problem is defined as:

$$Min_{w,b} MSE = \frac{1}{N} \sum_{i=1}^{N} (y_i - \widehat{y}_i)^2 = \frac{1}{N} \sum_{i=1}^{N} (y_i - f_{w,b}(X_i))^2$$

- We use the same **loss function** as in linear regression!
- The solution to this optimization problem is w^* and b^*
- Now we can make predictions!





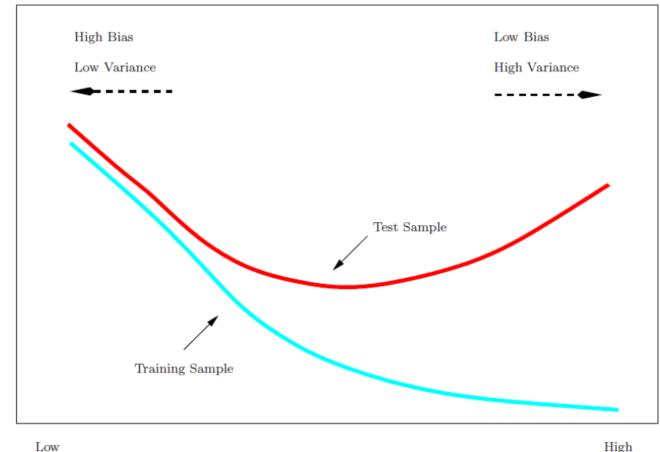


Polynomial Regression Evaluation Metrics

$$MSE = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y})^2$$

$$RMSE = \sqrt{MSE}$$

Prediction Error



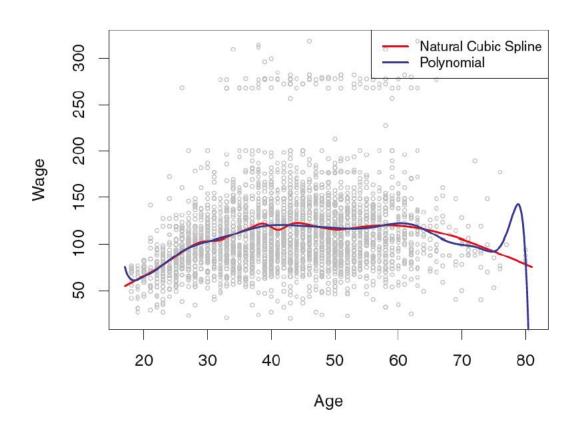
High

Model Complexity



Caveats!

- In general quadratic loss functions are sensitive to outliers
- Polynomials have notorious tail behavior
- Polynomials are global fit!
- Solution: piecewise polynomial, splines and local regressions.







Piecewise polynomials and splines!

