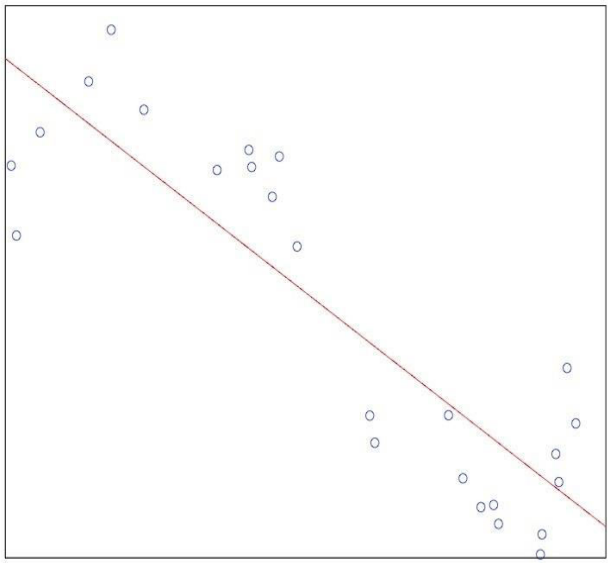
Problem Definition

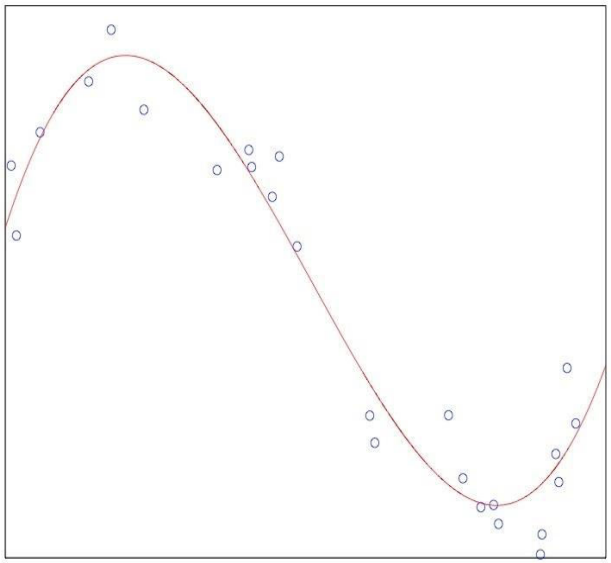
- Bias Variance trade-off

Problem Definition

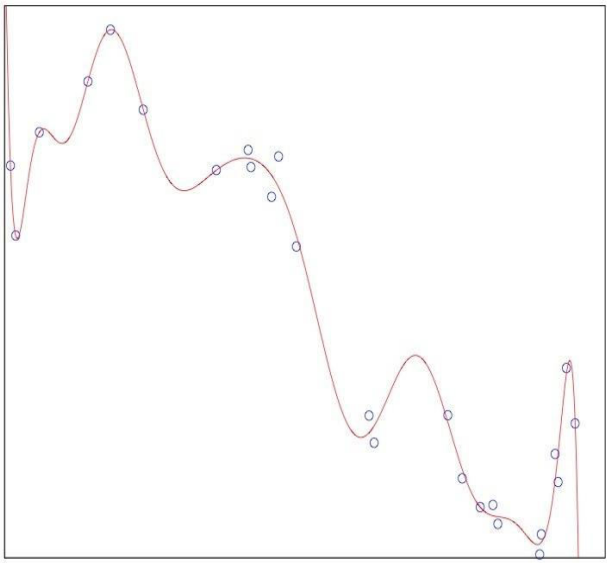
- Bias Variance trade-off



underfit
(degree = 1)



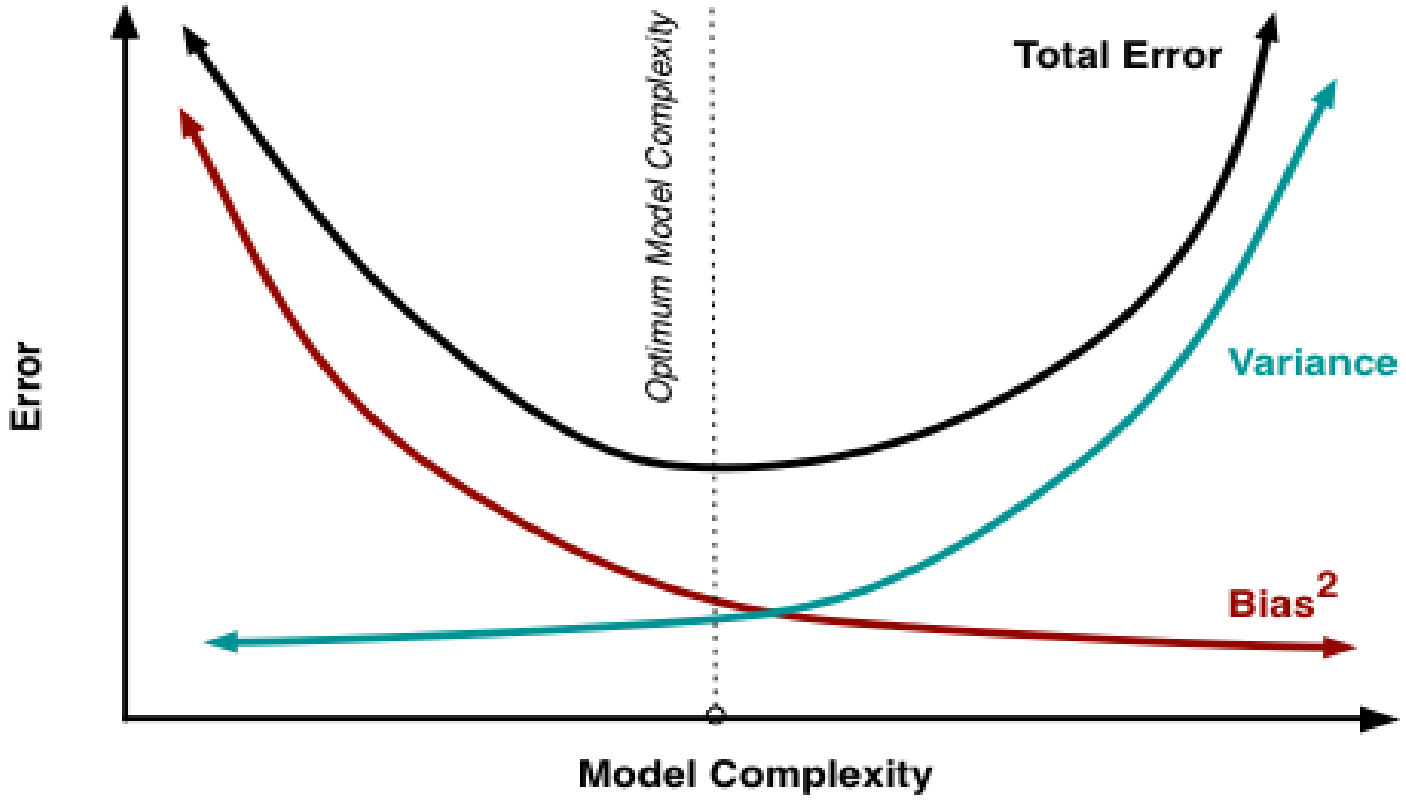
ideal fit
(degree = 3)



overfit
(degree = 20)

Problem Definition

- Bias Variance trade-off



Problem Definition

- Score function for model selection

Problem Definition

- Score function for model selection

score(model) = error(model) + penalty-function(model)

Akaike Information Criterion

Akaike Information Criterion

- Maximum likelihood case

$$AIC = -2\log \mathcal{L}(\hat{\theta}|y) + 2k$$

bias variance

- $\mathcal{L}$: likelihood function.
- $\hat{\theta}$: maximum likelihood estimate of θ.
- k: number of estimated parameters (including the variance).
- y is the random sample from the density function $f(x)$.

Bayesian Information Criterion

Bayesian Information Criterion

$$\mathrm{BIC} = -2 \cdot \ln L + k \ln(n)$$

- L: likelihood function
- k: number of parameters
- n: number of observations