R squared and adjusted R squared

R squared

$$\sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 + (\hat{y}_i - \bar{y})^2$$

$$\text{Total S.S.} = \text{Residual S.S} + \text{Model S.S}$$

$$R^2 = 1 - \frac{\text{Residual S.S}}{\text{Total S.S}}$$

Adjusted R squared: penalize the addition of extraneous predictors

$$\bar{R}^2 = 1 - (1 - R^2)\frac{n-1}{n-p-1}$$

$$= R^2 - (1 - R^2)\frac{p}{n-p-1}$$

Where:

p = number of explanatory variables

N = sample size

AIC & BIC

Both AIC and BIC penalizes the complexity of the model

$$AIC = nlog(R.S.S/n) + 2k$$

$$BIC = nlog(R.S.S/n) + plog(n)$$

where:

n = sample size

k,p = number of free parameters to be estimated

Boxcox: choosing the optimal lambda for transformation

$$t_\lambda(y) = \begin{cases} \frac{y^\lambda - 1}{\lambda} & \lambda \neq 0 \\ log(y) & \lambda = 0 \end{cases}$$