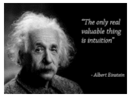
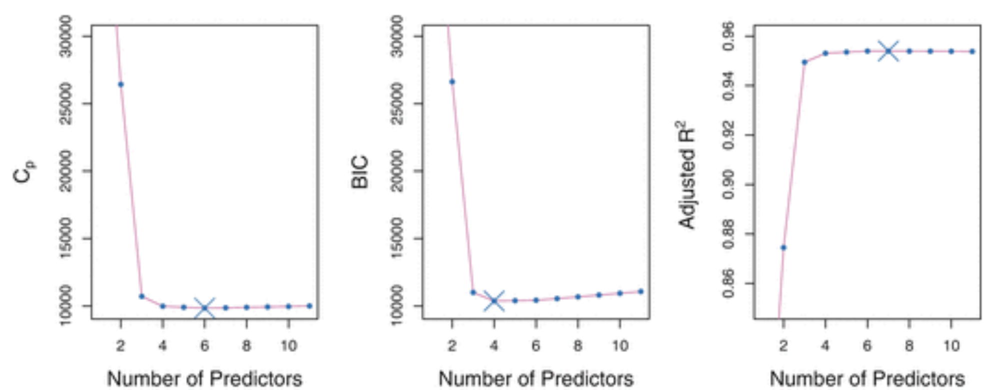
Error Measures



Error Measures and Model Size

- The **MSE** (all models) and **R^2** (some models) are good measures of **model fit** and individual **model quality** → **low MSE** and **high R^2**
- However, the **MSE** goes **down** and the **R^2** goes **up** as more variables are added to the model, so these are **not** so **useful** to **compare** models
- In addition, the **training MSE** tends to **underestimates** the **test MSE**, particularly as the model increases in **size** and **complexity**
- So, is it **worth** the added model **complexity** to improve the **MSE**?
- There are some measurement methods that **adjust** for the number of **variables** in a model: Mallow's **Cp,** Akaike Information Criterion **(AIC),** Bayesian Information Criterion **(BIC)** and **Adjusted R^2** (already covered)

Plots of Cp, BIC and Adjusted R2 against the number of variables for the Credit data in the ISLR package



Popular Error Measures

- **Mallow's Cp** error measure is a **"constant"** (thus the C) that **adjusts** the **MSE** by applying a **penalty** for the number of **variables** **"p"** (thus the "p") **variance σ** in the **errors**.
- The **Training Cp** is generally a good **estimator** of the **Test MSE**. As with most estimators, the **Cp** works better with larger samples →

$$Cp = MSE + model\ complexity\ penalty = MSE + \frac{2p\sigma^2}{n}$$

- Akaike Information Crierion (**AIC**) and Bayesian Information Criterion (**BIC**) are similar measures but the formulas are based on **log-likelihood** function of the model and therefore these formulas change depending on the estimation method. In sum:
  - ➢ The **true measure** of model error is still the **MSE**
  - ➢ **Cp**, **AIC** and **BIC** are **not** very useful to assess an **individual model's quality**
  - ➢ But they are very **useful** when **comparing** the quality of two or more models
  - ➢ The **lower** the Cp, AIC and BIC, the higher the **model quality**

KOGOD SCHOOL
of
BUSINESS