

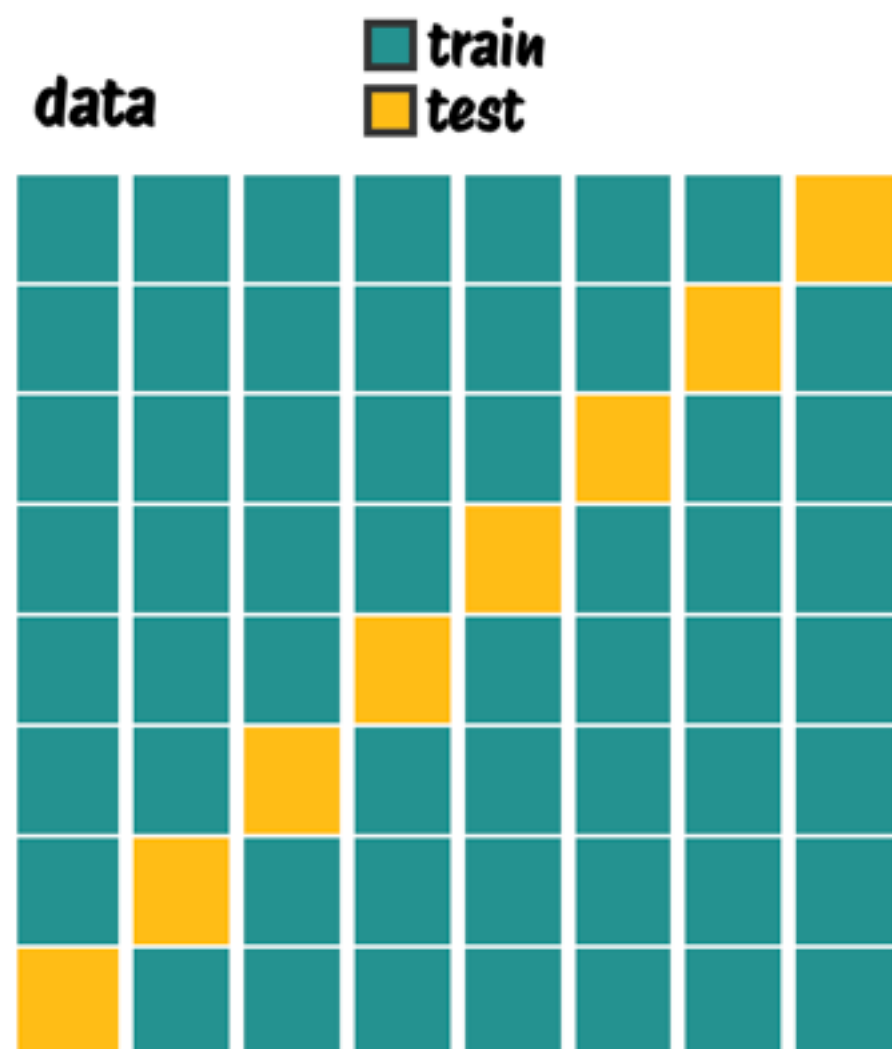
Model Selection Criteria: Validation by Prediction Error

Last week: how to use cross validation to choose a set of predictors by directly estimate prediction error using cross-validation techniques

$$\text{e.g. } \text{MSE} = \frac{\text{RSS}}{n}$$

$$\text{RMSE} = \sqrt{\frac{\text{RSS}}{n}}$$

$$R^2 = 1 - \frac{\text{RSS}}{\text{TSS}}$$



Now: indirectly estimating test performance using an approximation

Model Selection Criteria

Four ways to estimate test performance using an approximation

Full model has p predictors

RSS is the residual sum of squares for model with d predictors

$\hat{\sigma}^2 = \text{RSS}_p / (n - p - 1)$ is an estimate of the error variance for full model

1. Mallow's C_p criterion:

For a given model with d (out of the p available) predictors

$$C_p = \frac{1}{n} (\text{RSS} + 2d\hat{\sigma}^2)$$

we are penalizing models of higher dimensionality (larger d , greater penalty)

\implies choose the model which has **minimum** C_p