Chapter 6 in ISL: Regularization

DJM

17 April 2018

Regularization

- Another way to control bias and variance is through ~~regularization~~ or ~~shrinkage~~.
- Rather than selecting a few predictors that seem reasonable, maybe trying a few combinations, use them all.
- I mean ~~ALL~~.
- But, make your estimates of β "smaller"

Some optimization terms

- An optimization problem has 2 components:
  1. The "Objective function": e.g. $\frac{1}{n}\sum_i (y_i - x_i'\beta)^2$.
  2. The "constraint": e.g. "fewer than 5 non-zero entries in β".
- A constrained minimization problem is written

$$\min_{\beta} f(\beta) \text{ subject to} C(\beta)$$

- $f(\beta)$ is the objective function
- $C(\beta)$ is the constraint

Regularization

One way to do this for regression is to solve (say):

$$\min_{\beta} \frac{1}{n}\sum_i (y_i - x_i'\beta)^2$$
$$\text{s.t.} \sum_j \beta_j^2 < t$$

for some $t > 0$.

- This is called "ridge regression".
- The ~~minimizer~~ of this problem is called $\widehat{\beta}_{\text{ridge},t}$