

Class Notes

Statistical Computing & Machine Learning

Class 16

Review bias and variance

Example, a roughly quadratic cloud of points. Better to fit it with a constant, straight line, a quadratic?

- Depends on the amount of data.
- What if you have only $n=3$ or 4?

Constant will have the least sampling variation but the most bias.

Quadratic will have the least bias but the most sampling variation.

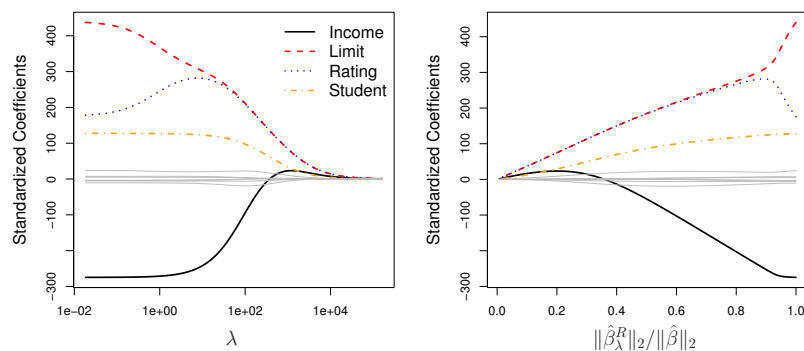
Shrinkage idea:

Make a linear combination of the constant model with the “full” model.

Methods

Ridge regression

ISLR Figure 6.4.



LASSO

ISLR Figure 6.7.

