Shrinking models

Christopher A. Sims
Princeton University
sims@princeton.edu

March 15, 2018

LASSO

Our model through several of these methods:

$$y_i = X_i\beta + \varepsilon_i$$

X_i is of high dimension, maybe bigger than number of observations. LASSO: Choose β to minimize

$$\sum (y_i - X_i\beta)^2 + \lambda \sum_j |\beta_j|$$

- This objective function is equivalent to the log likelihood for a model with independent Laplace priors on the β_j's.

Ridge

- content

Comparing ridge and LASSO with orthogonal regressors

- content

Bayesian Model Averaging

- content

Likelihood-based factor analysis, principal components

Comparing PC and FA to preceding methods

- Unlike the preceding approaches, this one shrinks dimensionality before looking at the dependent variable.
- Simply looking for the best linear predictor takes us back to OLS.
- But, inspired by neural networks, one could impose sparsity on α in

$$y = X\alpha\beta + \varepsilon$$

Many instruments

- content