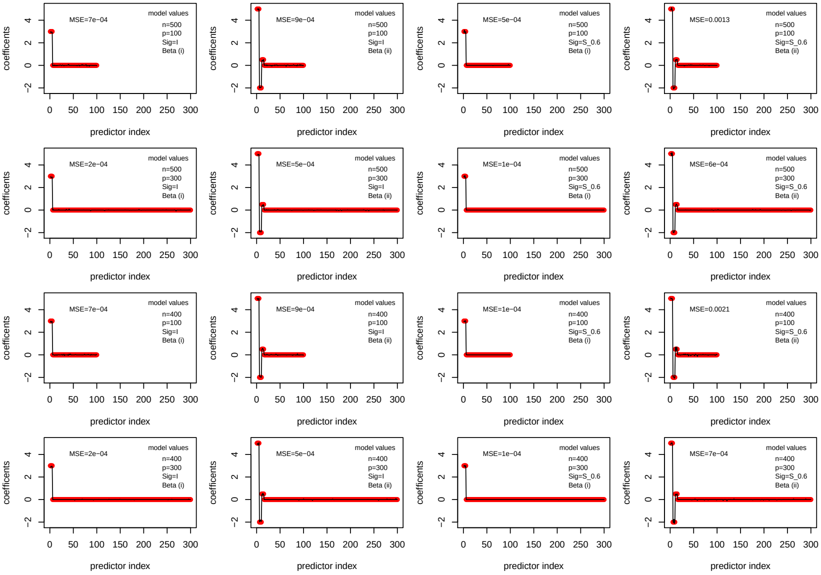
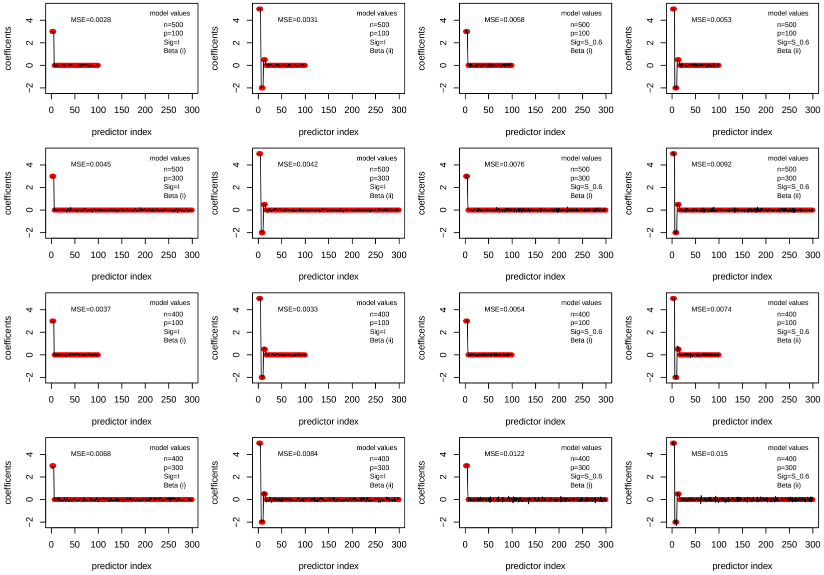
AMS 268 - HW #1

Raquel Barata

LASSO


coefficents
predictor index
MSE=7e−04
model values
n=500
p=100
Sig=I
Beta (i)
MSE=9e−04
model values
n=500
p=100
Sig=I
Beta (ii)
MSE=5e−04
model values
n=500
p=100
Sig=S_0.6
Beta (i)
MSE=0.0013
model values
n=500
p=100
Sig=S_0.6
Beta (ii)
MSE=2e−04
model values
n=500
p=300
Sig=I
Beta (i)
MSE=5e−04
model values
n=500
p=300
Sig=I
Beta (ii)
MSE=1e−04
model values
n=500
p=300
Sig=S_0.6
Beta (i)
MSE=6e−04
model values
n=500
p=300
Sig=S_0.6
Beta (ii)
MSE=7e−04
model values
n=400
p=100
Sig=I
Beta (i)
MSE=9e−04
model values
n=400
p=100
Sig=I
Beta (ii)
MSE=1e−04
model values
n=400
p=100
Sig=S_0.6
Beta (i)
MSE=0.0021
model values
n=400
p=100
Sig=S_0.6
Beta (ii)
MSE=2e−04
model values
n=400
p=300
Sig=I
Beta (i)
MSE=5e−04
model values
n=400
p=300
Sig=I
Beta (ii)
MSE=1e−04
model values
n=400
p=300
Sig=S_0.6
Beta (i)
MSE=7e−04
model values
n=400
p=300
Sig=S_0.6
Beta (ii)

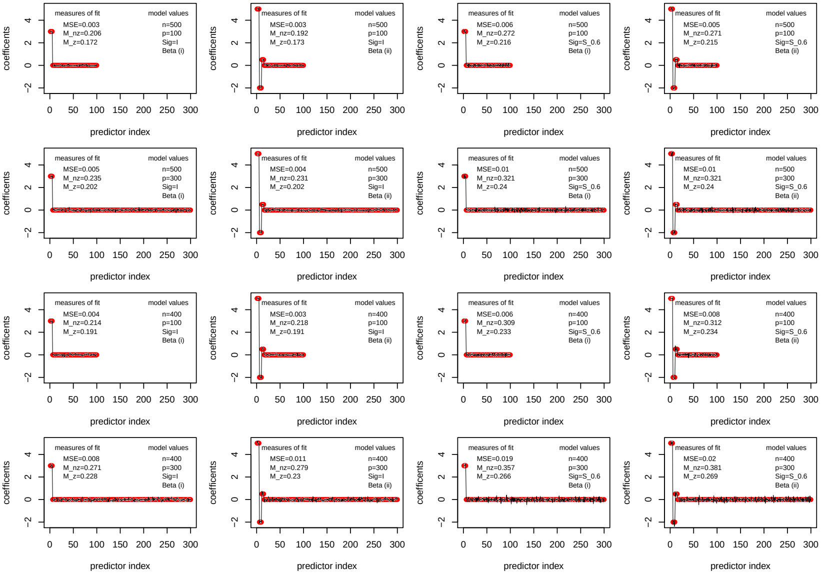
Ridge Regression



LASSO vs. Ridge Regression

- True values in red with estimates $\tilde{\beta}$ in black.
- λ_j chosen to minimize error.
- MSE $= \frac{1}{p}\sum_{i=1}^{p}(\beta - \tilde{\beta})^2$, where $\tilde{\beta}$ is the estimate for β, are unanimously smaller when data is fit using LASSO for all 16 combinations.
- Ridge regression coefficents $\tilde{\beta}_j$ are farther from true values when $\beta_j = 0$.

Spike and Slab



Spike and Slab

- True values β_j in red, $E(\beta_j|y)$ in black, 95% posterior intervals in grey.
- Spike and slab code written using the model in the notes with fixed $v_0 = 0.1^2$ and $v_1 = 10^2$.
- MSE $= \frac{1}{p}\sum_{i=1}^{p}(\beta - E(\beta_j|y))^2$.
- $M_{nz} = mean(L_j : \beta_j^0 \neq 0)$ and $M_z = mean(L_j : \beta_j^0 = 0)$ where L_j is the length of the 95% posterior interval for β_j.
- Variable selection is more straight forward for Spike and Slab than LASSO as posterior samples provide marginal inclusion probabilities $Pr(\gamma_j = 1|y, X)$ for coefficient β_j.

Posterior Prediction for $n = 500$, $p = 100$, $\Sigma = S_{0.6}$ and $\beta^{(i)}$

- 95% posterior predicive intervals for 50 new predictors in grey, $y_{est,i}$ in black, and true values $y_{pred,i}$ in red.
- $\text{MSPE} = \frac{1}{50}\sum_{i=1}^{50}(y_{pred,i} - y_{est,i})^2$.