Penalized Regression with Various Design Matrix Correlation Structures

Hao Chai

Department of Statistics and Actuarial Science
University of Iowa

October 22, 2013

Problem Review

The goal is to minimize

$$||\mathbf{y} - X\boldsymbol{\beta}||_2^2 + \sum_{i=1}^{p} \rho(|\beta_i|; \lambda).$$

Want to compare the performance of LASSO and MCP under different correlation structures.
Set $\boldsymbol{\beta} = (3, 1, -1, 3, \mathbf{0})$, $n = 100$, $p = 500$ and signal to noise ratio $s/n = 1.5$. There are 200 simulated data sets.

Problem Review

The goal is to minimize

$$||\mathbf{y} - X\boldsymbol{\beta}||_2^2 + \sum_{i=1}^{p} \rho(|\beta_i|; \lambda).$$

Want to compare the performance of LASSO and MCP under different correlation structures.
Set $\boldsymbol{\beta} = (3, 1, -1, 3, \mathbf{0})$, $n = 100$, $p = 500$ and signal to noise ratio $s/n = 1.5$. There are 200 simulated data sets.

Three Correlation Structures

cov0: The columns of X are independent, i.e. $\mathbf{x}_i$ and $\mathbf{x}_j$ are independent if $i \neq j$.

cov1: $corr(\mathbf{x}_i, \mathbf{x}_j) = 0.8^{|i-j|}$, for $i, j = 1, 2, \ldots, p$.

cov2: For $i = 1, 2, \ldots, p$, $\mathbf{u}_i = \begin{cases} N(\mathbf{0}, I), & \text{with probability 0.1;} \\ 0, & \text{with probability 0.9.} \end{cases}$

For $i = 1, 2, \ldots, p$, $\mathbf{v}_i \overset{i.i.d.}{\sim} N(\mathbf{0}, I)$. Then for $i = 1, 2, \ldots, p$,

$$\mathbf{x}_i = \mathbf{v}_i + a \cdot \sum_{j=-3}^{3} \mathbf{u}_{i+j},$$

with $\mathbf{u}_{-2} = \mathbf{u}_{-1} = \mathbf{u}_0 = \mathbf{u}_{n+1} = \mathbf{u}_{n+2} = \mathbf{u}_{n+3} = \mathbf{0}$ for the brevity of notations. In this case, $a = 2$.

Selection Error Rates for LASSO and MCP

	cov0	cov1	cov2
Ltype1	0.029(0.028)	0.036(0.023)	0(0.001)
Ltype2	0.249(0.195)	0.475(0.079)	0.684(0.111)
LFDR	0.717(0.197)	0.862(0.077)	0.048(0.144)
LFIR	0.002(0.002)	0.004(0.001)	0.005(0.001)
Mtype1	0.006(0.007)	0.003(0.005)	0.005(0.006)
Mtype2	0.324(0.186)	0.498(0.025)	0.469(0.139)
MFDR	0.386(0.276)	0.245(0.293)	0.383(0.296)
MFIR	0.003(0.001)	0.004(0)	0.004(0.001)

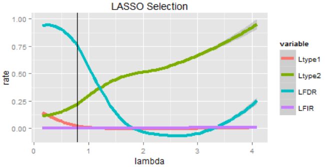
The initial letter “L" stands for LASSO and “M" stands for MCP. “type1", “type2", “FDR" and “FIR" are type I error rate, type II error rate, the false discovery rate and the false insignificant rate respectively.

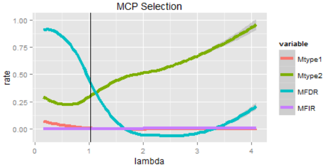
Estimation Error Rates for LASSO and MCP

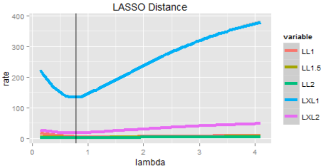
	cov0	cov1	cov2
LL1	5.804(2.499)	6.233(1.965)	5.2(0.492)
LL2	2.019(0.324)	2.097(0.28)	3.193(0.314)
LXL1	145.357(26.384)	115.978(21.723)	269.357(21.373)
LXL2	18.646(3.318)	14.482(2.684)	33.954(2.797)
ML1	3.291(1.031)	3.062(0.934)	4.494(1.602)
ML2	1.56(0.308)	1.586(0.203)	2.23(0.644)
MXL1	111.632(20.756)	62.65(16.676)	192.853(49.974)
MXL2	14.347(2.772)	7.822(2.064)	24.107(6.386)

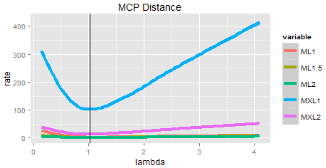
The initial letter "L" stands for LASSO and "M" stands for MCP. $L1 = ||\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}||_1$, $L2 = ||\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}||_2$, $XL1 = ||X(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})||_1$ and $XL2 = ||X(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})||_2$.

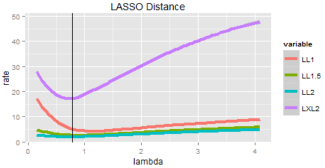
Covariance Structure 0

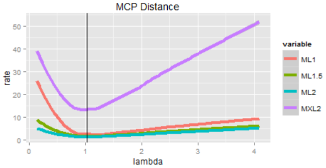




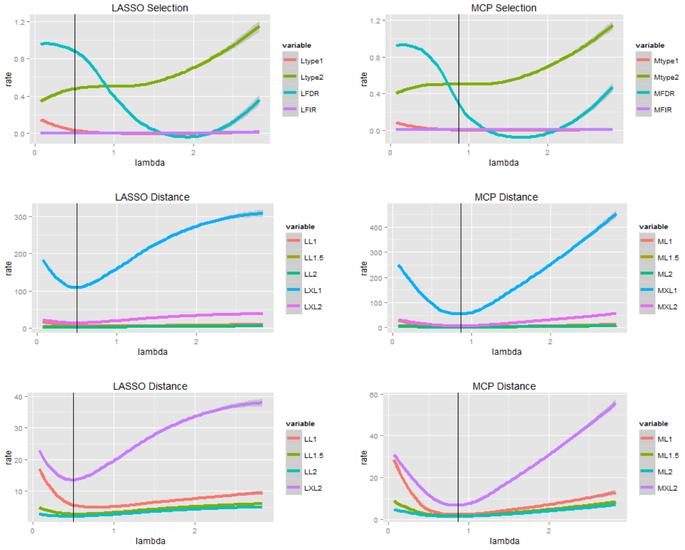




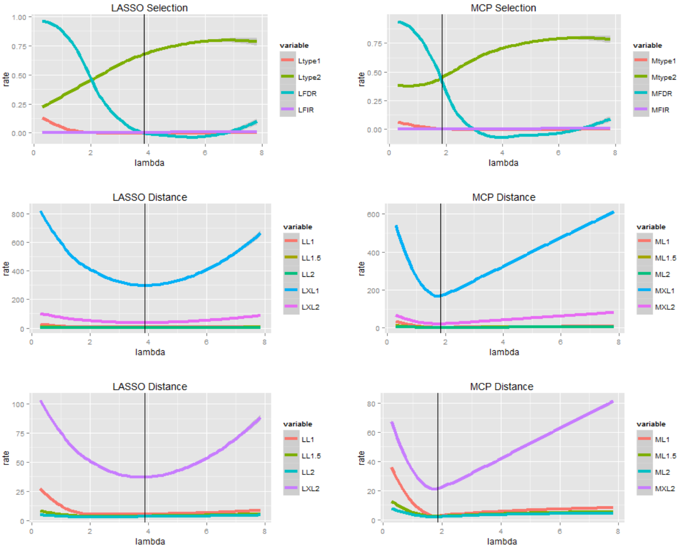




Covariance Structure 1



Covariance Structure 2



Thank you!