Alternative Penalized Regression Methods

1 Lasso, SCAD and MCP

In this exercise we would like to get familiar with some alternative penalized methods which are presented in the course (see **Chapter 4.4**). After having read the documentation of the R package *ncvreg* and installed it, perform the following steps:

(a) Fix the generating vector $\boldsymbol{\beta} = (4, 2, -4, -2, 0, 0, \ldots, 0)$ and set the seed equal to 11 (i.e. set.seed(11)).

(b) Generate from a MVN (multivariate normal) a matrix $\mathbf{X_{n*p}}$ with $n = 200$ and $p = 1000$. You can choose the location vector as you wish but set the scale matrix with an autoregressive form $\boldsymbol{\Sigma} = [\sigma_{lm}]_{l,m=1,\ldots,p}$ with $\sigma_{lm} = \rho^{|l-m|}$.

(c) For each $\boldsymbol{\rho} = [0\ 0.2\ 0.5]$ generate $\hat{\mathbf{y}}$ thanks to the relation $\mathbf{y} = \mathbf{X_{n*p}}\,\boldsymbol{\beta} + \boldsymbol{\epsilon}$ where ϵ_i is a standard normal. Suppose for simplicity that the errors are uncorrelated.

(d) Compare the solution paths (graphically as a function of λ) for the lasso, SCAD and MCP by fixing several values for γ (choose e.g. $\gamma = (1.5, 2, 3, 3.7, 5)$) for each value of ρ indicated at point c.