Statistical Computing: Homework 3

Due on April 24 (Sunday) 23:30

You may take 2 problems among 3 as your choice. You may get extra points for taking all 3 problems.

1. Consider the Poisson regression problem in Lecture 6 (notes p12):

$$Y_i \sim \text{Poi}(\lambda_i), \quad i = 1, 2, ..., n; \quad \log \lambda_i = \beta_0 + \beta_1 x_i.$$

The log-likelihood function of (β_0, β_1) satisfies

$$\begin{aligned}\ell(\beta_0, \beta_1) &= \sum_{i=1}^{n} \log\left(\frac{\lambda_i^{y_i}\, e^{-\lambda_i}}{y_i!}\right) = \sum_{i=1}^{n} \left\{\log\left(e^{y_i(\beta_0+\beta_1 x_i)}\, e^{-e^{\beta_0+\beta_1 x_i}}\right) - \log(y_i!)\right\} \\ &= (\text{constant}) + \sum_{i=1}^{n} y_i(\beta_0 + \beta_1 x_i) - \sum_{i=1}^{n} e^{\beta_0+\beta_1 x_i},\end{aligned}$$

where the (constant) does not depend on (β_0, β_1).

Using DataA, solve the MLE of (β_0, β_1) using the following methods.

(a) newton method

(b) gradient descent

<u>DataA (extract from Bikeshare data):</u>

Y_i is the number of bikers at ith day; x_i is the temperature (rescaled) at ith day ($i = 1, 2, ..., 200$). Detail variable descriptions can be found in "Bikeshare {ISLR2}".

2. Consider the following normal model:

$$\begin{aligned}&Y_i \sim N(\mu_i, \sigma_i^2), \quad i = 1, 2, ..., n; \\ &\mu_i = \beta_0 + \beta_1 x_i, \quad \sigma_i^2 = e^{\alpha_0+\alpha_1 x_i}, \quad \boldsymbol{\alpha} = (\alpha_0, \alpha_1)' \in R^2,\ \boldsymbol{\beta} = (\beta_0, \beta_1)' \in R^2.\end{aligned}$$

Using DataB, solve the MLE of $(\boldsymbol{\beta}, \boldsymbol{\alpha})$ based on the (block) coordinate descent method. (You may treat $\boldsymbol{\beta}$ as a block and $\boldsymbol{\alpha}$ as a block.)

<u>DataB (extract from Boston Housing data):</u>

$Y_i = \log(\text{medv}_i)$ for the ith suburb region of Boston; and $x_i = \text{lstat}_i$. Detail variable descriptions can be found in "Boston {ISLR2}".

3. Consider an alternative fused lasso problem:

$$\min_{\boldsymbol{\beta}} \left\{ \frac{1}{2} \sum_{t=1}^{n} (y_t - \beta_t)^2 + \tau \sum_{t=3}^{n} |\beta_t - 2\beta_{t-1} + \beta_{t-2}| \right\}, \quad \boldsymbol{\beta} = (\beta_1, ..., \beta_n)',$$

which can be reformulated as

$$\min_{\boldsymbol{\beta}} \left\{ \frac{1}{2} \|\boldsymbol{y} - \boldsymbol{\beta}\|^2 + \tau \sum_{t=3}^{n} |\delta_t| \right\},$$

where

$$\delta_t \equiv \beta_t - 2\beta_{t-1} + \beta_{t-2} = \big(0, ..., 0, 1, -2, \underbrace{1}_{t\text{-th}}, 0, ..., 0\big)\boldsymbol{\beta}, \quad t = 3, ..., n;$$

$$\underbrace{\boldsymbol{\delta}}_{(n-2)\times 1} \equiv (\delta_3, \delta_4, ..., \delta_n)' = \underbrace{D}_{(n-2)\times n} \boldsymbol{\beta}, \quad D = \begin{bmatrix} 1 & -2 & 1 & \cdots & 0 \\ 0 & 1 & -2 & 1 & \cdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & -2 & 1 \end{bmatrix}.$$

Using ADMM to solve $\{\beta_t\}$ for DataC.

<u>DataC:</u> $\{(t, y_t) : t = 1, 2, ..., 100\}$ (same data shown in R Lab7-2 Example 5)