Stockholms universitet

Optimisation Algorithms in Statistics I – Autumn 2020
Assignment 4

Frank Miller, frank.miller@stat.su.se,
Department of Statistics, Stockholm University

November 6, 2020

Perform the solutions individually and send your report **until November 27** to me. Try to keep this deadline. However, if you have problems with it, there will be a final deadline on January 25 for all assignments. Please include your name in the filename(s) of your solution file(s).

Problem 4.1

We want to determine the D-optimal design for cubic regression where the independent variable x is allowed to have values between 0 and 10. Four different $x_i \in [0, 10], i = 1, 2, 3, 4$, can be chosen by the experimenter and the proportion of observations done at each x_i is $w_i \geq 0$ with $\sum_{i=1}^4 w_i = 1$. The D-optimal design maximises

$$\det\left(\sum_{i=1}^{4} w_i \mathbf{f}(x_i)\mathbf{f}(x_i)^\top\right), \text{ with } \mathbf{f}(x) = (1, x, x^2, x^3)^\top,$$

under the restrictions mentioned above.

a. Determine a matrix $\mathbf{U}$ and a vector $\mathbf{c}$ such that the constraints can be written in the form $\mathbf{U}\mathbf{y} - \mathbf{c} \geq 0$, where $\mathbf{y}$ is the vector of parameters to be optimised over.

b. Determine the D-optimal design using `constrOptim`. Does the result make sense?

c. Write an `R`-function for a function $\tilde{g}$ where log barriers $\mu \cdot b(\mathbf{y})$ at all constraints are added to the function g (which is to be maximised). The value μ could be a parameter in the function such that you easily can modify it.

d. Choose some reasonable values for μ and compute the optimal value of g using unconstrained optimisation, e.g with `optim`. Hint: Check first how large values of g are to get the μ's roughly right. Report results for a sequence of decreasing μ. Do you obtain similar results as in b. when using small μ?

Problem 4.2

We consider again as in Problem 3.3 the experiment investigating how the growth of garden cress depends on a (potentially) toxic fertilizer. The data is on the homepage in the file `cressdata.txt` (columns: observation number, fertilizer concentration, yield).

We want to estimate now a third-degree polynomial (cubic), again using least squares with L^1 In contrast to the penalized objective function in Problem 3.3, we use now the constrained objective function

$$\text{Minimise } g(\boldsymbol{\beta}) = \|\mathbf{X}\boldsymbol{\beta} - \mathbf{y}\|_2^2 \text{ subject to } \|\tilde{\boldsymbol{\beta}}\|_1 \leq t, \tag{1}$$

where $\mathbf{X}$ is the design matrix with columns 1, fertilizer, fertilizer^2, fertilizer^3, $\tilde{\boldsymbol{\beta}} = (\beta_1, \beta_2, \beta_3)^\top$ is the parameter vector without intercept, $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_3)^\top$ is the complete parameter vector and $\mathbf{y}$ is the yield-data. (We do this time not regularise the intercept, see e.g. Lange (2010), page 310.) The constant $t \geq 0$ is now the regularisation constant. t and λ (in Problem 3.3) are related such that a t in the constrained problem corresponds to an λ in the penalised problem which gives the same solution.

Note that now, $t = \infty$ corresponds to the least squares estimation, where the solution for $\boldsymbol{\beta}$ of the optimisation problem is $(\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top\mathbf{y}$.

a. Write the constraint $\|\tilde{\boldsymbol{\beta}}\|_1 \leq t$ in terms of eight linear constraints $\mathbf{u}_i^\top\boldsymbol{\beta} + c_i \geq 0$ (or as $\mathbf{U}\boldsymbol{\beta} - \mathbf{c} \geq \mathbf{0}$ with a matrix $\mathbf{U}$ with 8 rows).

b. Write an expression for the objective function minus log barriers, $\tilde{g}(\boldsymbol{\beta}) = g(\boldsymbol{\beta}) - \mu \cdot b(\boldsymbol{\beta})$. Determine the gradient of g and of $\tilde{g}$. (Note: You do not have to implement the gradient of $\tilde{g}$.)

c. Compute the Lasso-estimate using `constrOptim` for $t = 1000, 100$ and two other t's. Test two different methods for the inner iteration in `optim`, e.g. Nelder-Mead and BFGS. For the non-Nelder-Mead-method, specify explicitly the gradient when calling `constrOptim`. (Note: here you need the gradient of g, not of $\tilde{g}$.) Check $\|\tilde{\boldsymbol{\beta}}\|_1$ for the solutions: is the solution on the boundary of the set of feasible points or in the inner of the set?