Mariia Okuneva, M.Sc.

Data Mining
Home Assignment 3

Consider following random variables

$$u \sim N(0,1) \qquad x_1 \sim N(0,1) \qquad x_2 \sim N(0,1) \qquad x_3 \sim N(0,1)$$

and a regression model

$$y = 1 + 0.5\, x_1^2 + u.$$

Your task is to write an `R` script that contains the following parts. First, download the script template *HA3_ yourname.R* from OLAT.

1. Generate samples of size $n = 100$ of all random variables, including y.
2. Estimate the following model using ridge regression

   $$y = \beta_0 + \beta_1 x_1^2 + \beta_2 x_2^2 + \beta_3 x_3^2 + u.$$

   You are allowed to use functions `optim` or `nlm` for this purpose. Program the estimation procedure from scratch without using integrated ridge regression methods.
3. In order to pick the optimal shrinkage intensity, perform cross-validation with an order of your choosing. You may use anything you need from the previous home assignment.
4. Comment on your results.

Remarks: Write comments for everything you do. Codes that are not written using the template and/or that return error messages will not be evaluated.
Submission: Submit your scripts via email to *mokuneva[at]stat-econ.uni-kiel.de* until the end of July 11th (until 00:00:00, July 11th)