

Assignment 3

Create an R function named `L1bestsub` that performs best subset median regression. The function should be based on the following input:

- a column vector y (response)
- a matrix X of predictors (predictors should be in columns)
- k : $k = 1$ should base final model selection on criterion 2.3 whereas $k = 2$ should correspond to equation 2.4 from the JASA paper that has been uploaded in Blackboard.

Your function should report: a) the predictors that are included in the final median regression based on the chosen BIC criterion (which depends on the value of k) and b) coefficient estimates, the corresponding confidence intervals and the BIC of the final model.

Assignment 3

- Compare your function versus the analogous least-squares-based function presented in the lecture slides, using synthetic data: specifically, apply `rnorm` to create 3 X matrices with 10 columns and $n = 25, 50, 100$ and 2 responses, y_1, y_2 , per X matrix.
- Responses should be linear functions of the first 5 columns of X with coefficients `c(0.3,0.25,0.2,0.15,0.1)` with an intercept equal to 0.5. y_1 should be based on standard normal errors whereas y_2 should be based on Laplace errors with parameters 0 and 1 for location and scale, respectively.
- Discuss your findings; create a few figures that convey useful info wrt your result.

Assignment 3

- E-mail your responses by noon, Tuesday October 20.
- Your code should be included in a pdf (with comments that will help you presenting it in class), but also attached in a separate R script.
- Use the following file names for your pdf and script:
LASTNAME_FIRSTNAME_ASUID_ASSIGNMENTNUMBER
- Prepare your pdfs carefully; each week some of you will present their work.