SAINT PETERSBURG STATE UNIVERSITY

Faculty of Applied Mathematics and Control Processes

Mathematical Game Theory and Statistical Decisions Department

Applied Statistics in R

Laboratory work № 5

Professor: Parilina Elena M.

Student: Orlov Ivan M., 19.M09-пу

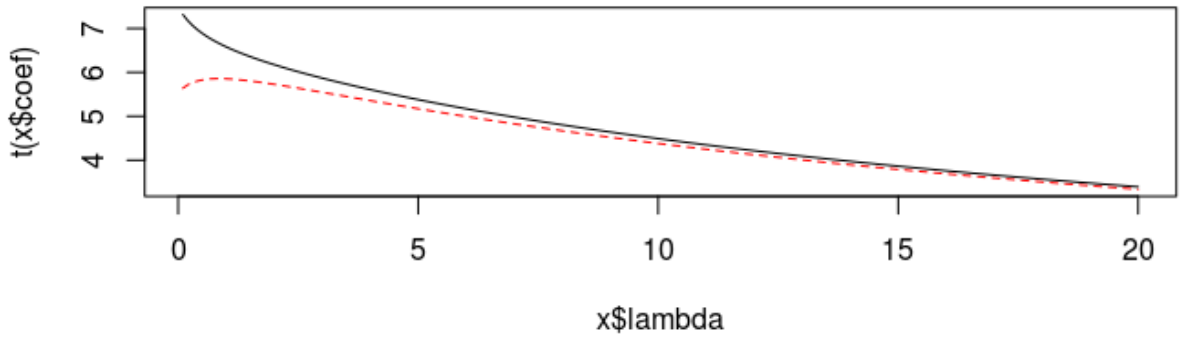
Saint Petersburg

2020

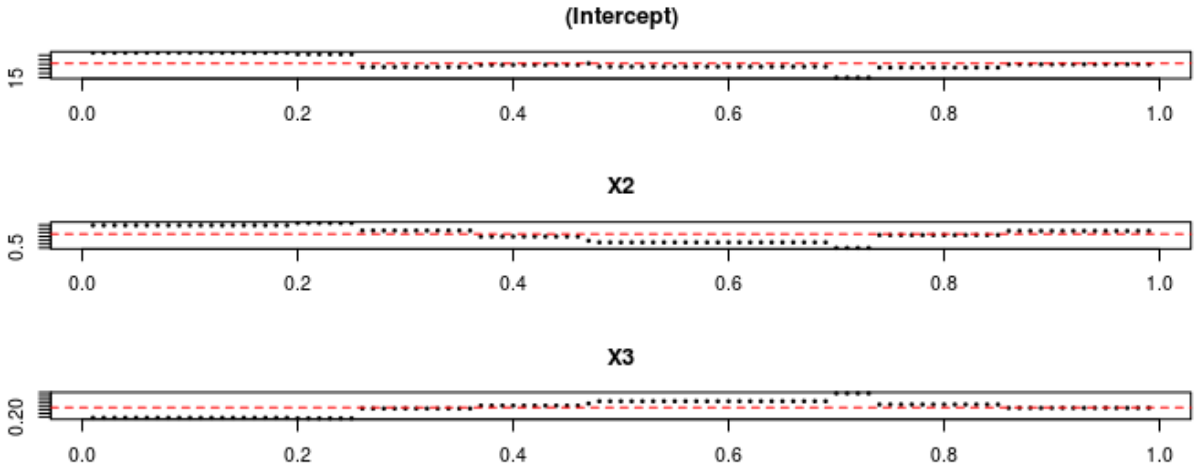
blood model is:

Model\Coef	Intercept	X2	X3
Linear	30.994	0.861	0.335
Ridge 0.1	30.916	0.842	0.342
Ridge 20	86.31	0.39	0.202
Quantile 0.01	42.858	1.088	0.188
Quantile 0.99	29.91	0.94	0.328

Ridge:



Quantile:

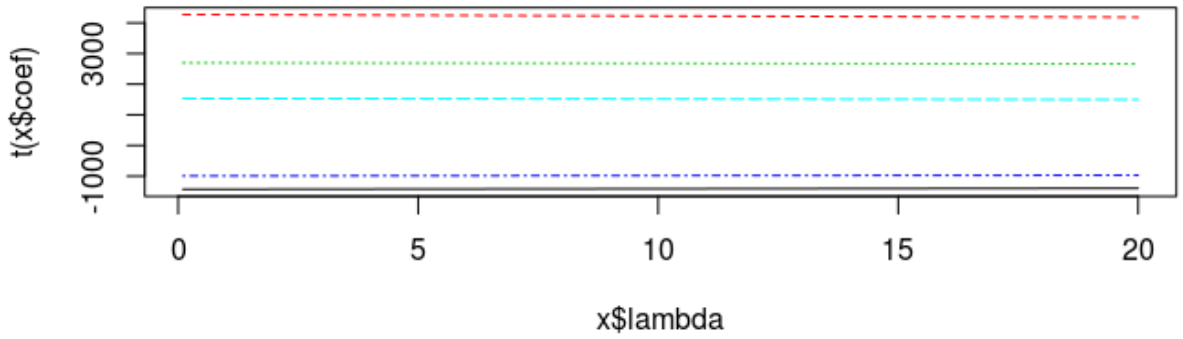


Linear, ridge with lambda 0.1 and quantile with tau 0.99 regressions provide similar results while both ridge 20 and quantile 0.01 regressions significantly deviate.

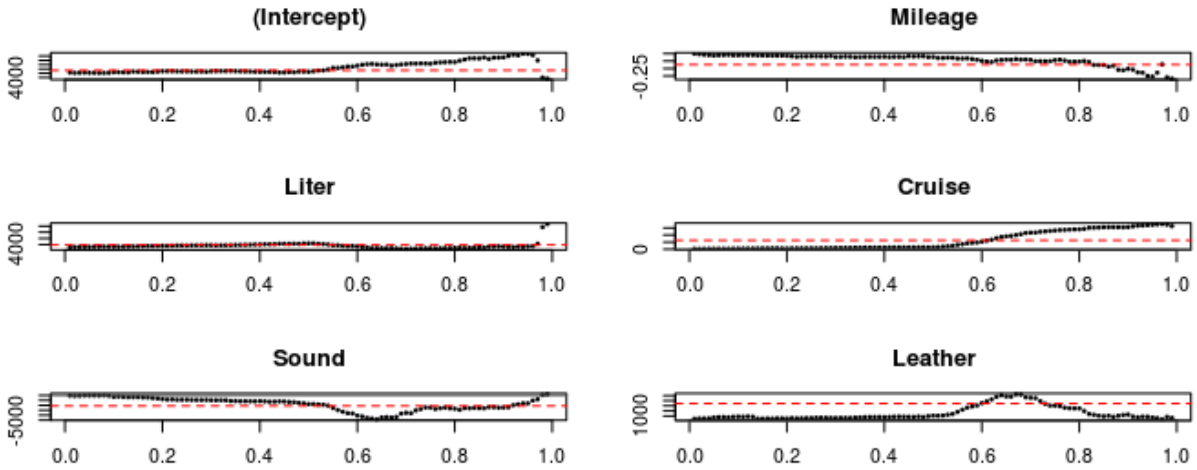
kuiper model is:

Model\Coef	Intercept	Mileage	Liter	Cruise	Sound	Leather
Linear	7318.267	-0.174	3864.31	6250.63	-2126.427	3435.814
Ridge 0.1	7319.482	-0.174	3863.915	6250.219	-2126.183	3435.401
Ridge 20	7557.283	-0.17	3786.924	6168.781	-2078.944	3355.197
Quantile 0.01	6294.795	-0.101	3054.918	3221.007	292.22	448.127
Quantile 0.99	3465.071	-0.268	10654.581	16275.624	199.629	642.492

Ridge:



Quantile:

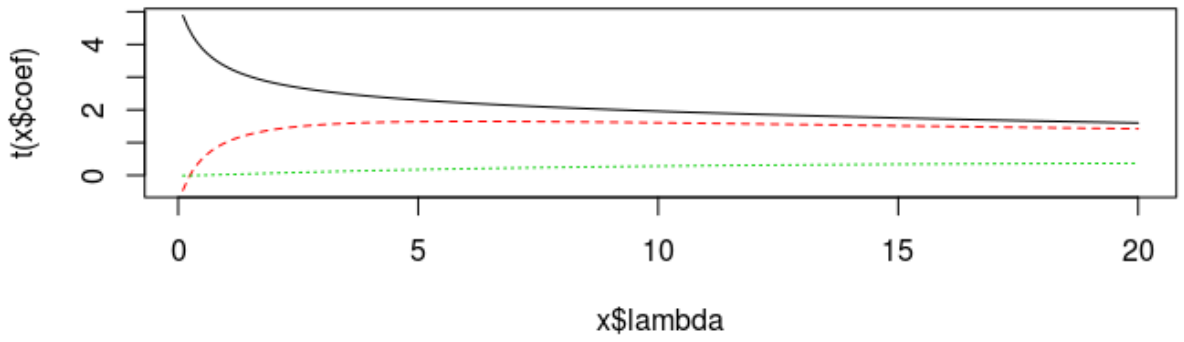


Linear and ridge (both lambdas 0.1 and 20) regressions provide similar results while both quantile regressions significantly deviate.

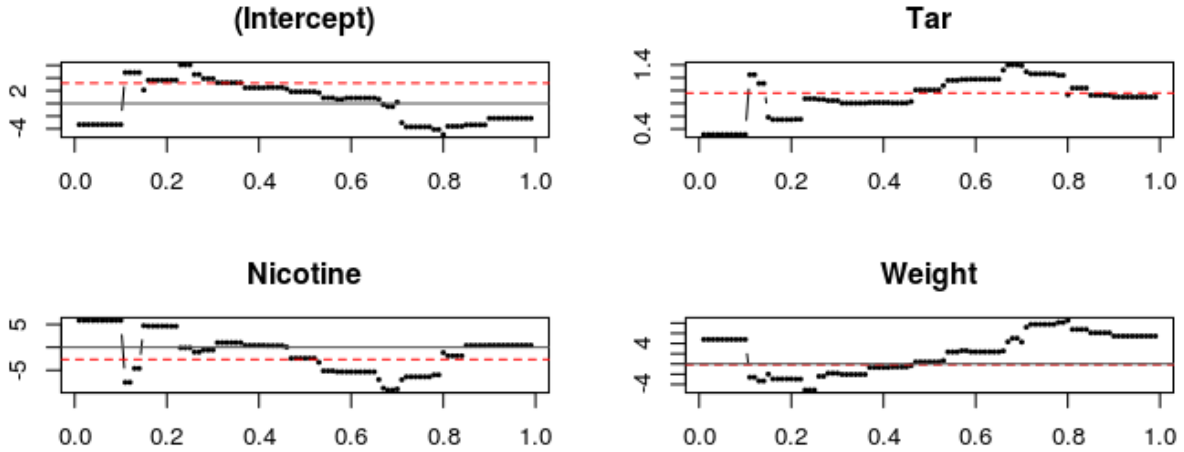
cigarettes model is:

Model\Coef	Intercept	Tar	Nicotine	Weight
Linear	3.202	0.963	-2.632	-0.131
Ridge 0.1	3.072	0.879	-1.328	-0.128
Ridge 20	1.286	0.288	4.089	4.263
Quantile 0.01	-3.339	0.314	5.856	4.794
Quantile 0.99	-2.345	0.899	0.44	5.45

Ridge:



Quantile:



Linear and ridge with ambda 0.1 and regressions provide results that are closest to be called similar while every other regression deviates significantly from any other.

```
library(quantreg)
library(readxl)
library(lmtest)
library(MASS)

blood <- read_excel("Datasets/blood.xlsx")

kuiper <- read_excel("Datasets/kuiper.xls")

cigarettes <- read.delim("Datasets/cigarettes.dat.txt", header=FALSE)
names(cigarettes) = c("Name", "Tar", "Nicotine", "Weight", "CO")

research_reg = function(data, formula, reg_name){
  print(reg_name)
  res = 0
  if (reg_name == 'linear'){
    res = lm(formula = formula, data = data)
  }
  else if (reg_name == 'ridge') {
    res = lm.ridge(formula = formula, lambda = seq(0.1, 20, by=0.1), data = data)
  }
  else if (reg_name == 'quantile') {
    res = rq(formula = formula, tau = 1:99/100, data = data)
  }

  print(res)
  plot(res)
}



formulas = c(X1 ~ X2 + X3,
             Price ~ Mileage + Liter + Cruise + Sound + Leather,
             CO ~ Tar + Nicotine + Weight)
datas = list(blood, kuiper, cigarettes)

#for (f in formulas)
#  pairs(f, main = 'Simple Scatterplot Matrix')

regs = c('linear', 'ridge', 'quantile')

for (i in seq(1,3))
  for (r in regs)
    research_reg(data = datas[[i]], formula = formulas[[i]], reg = r)
```