Homework 1

SID : 201872190

Jaeho, Chang

2019-03-07

Data

$$Y_1, \cdots, Y_n \overset{iid}{\sim} \frac{1}{2}N(-2, 2) + \frac{1}{2}N(6, 4).$$

```
rm(list = ls())
set.seed(1118)
n <- 41
components <- sample(1:2, prob=c(.5, .5), size=n, replace=TR
mus <- c(-2, 6)
sds <- sqrt(c(2, 4))
y.v <- rnorm(n, mean = mus[components], sd=sds[components])
lambda.df <- 10^(-4)

lossf1 <- function(t){t^2}
lossf2 <- function(t){abs(t)}

f <- function(f_y.v, f_mu, f_lossf) {sum(f_lossf(f_y.v - f_m
lassof <- function(f_y.v, f_mu, f_lossf, f_lambda = lambda.d
fplus <- function(x){x*(x > 0)}
```

LS Loss

1. $f(\mu) = \sum_i L(y_i - \mu)$ with $L(t) = t^2$

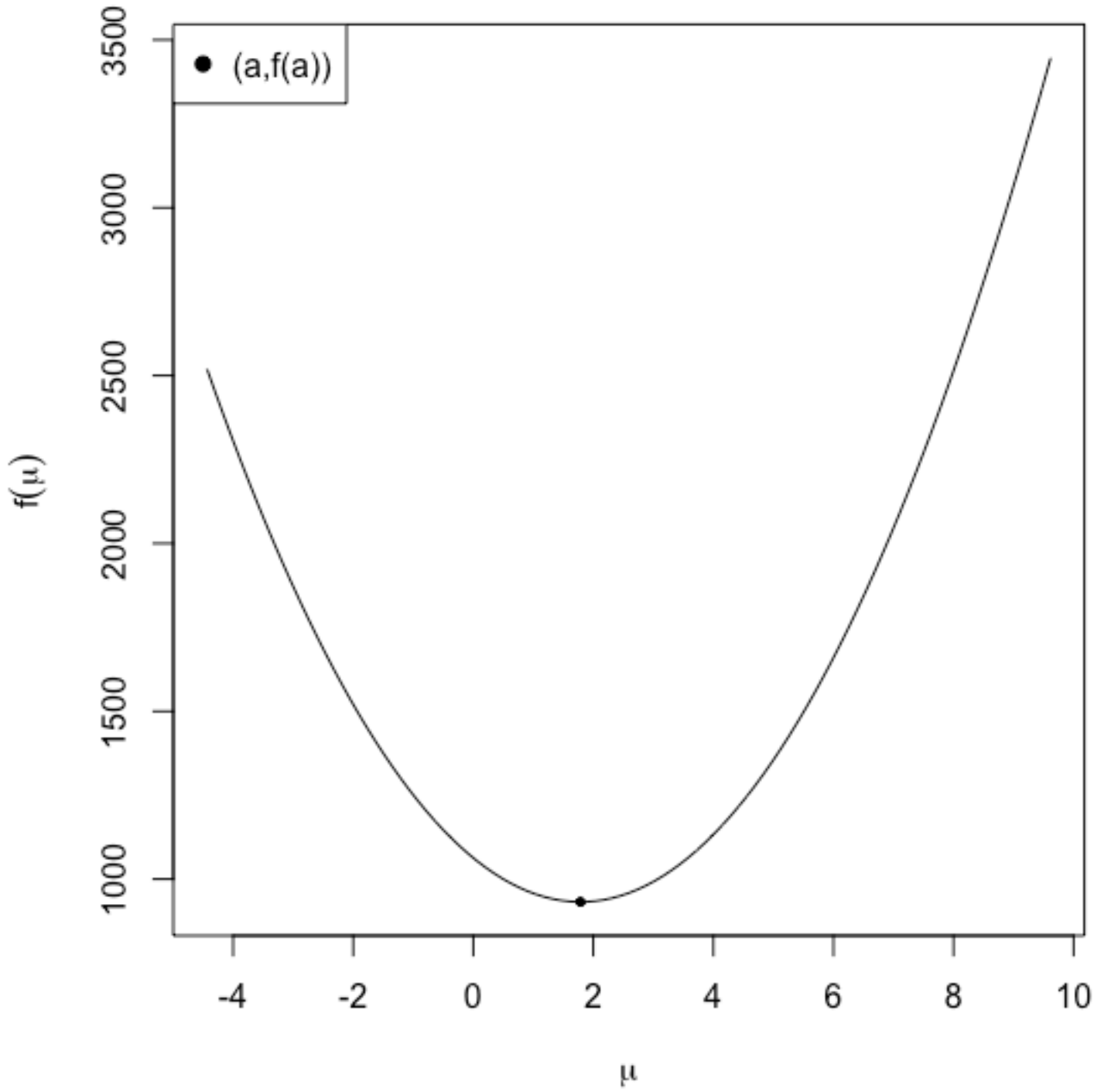
2. Initialize with $a = \min_i y_i$ and $b = \max_i y_i$

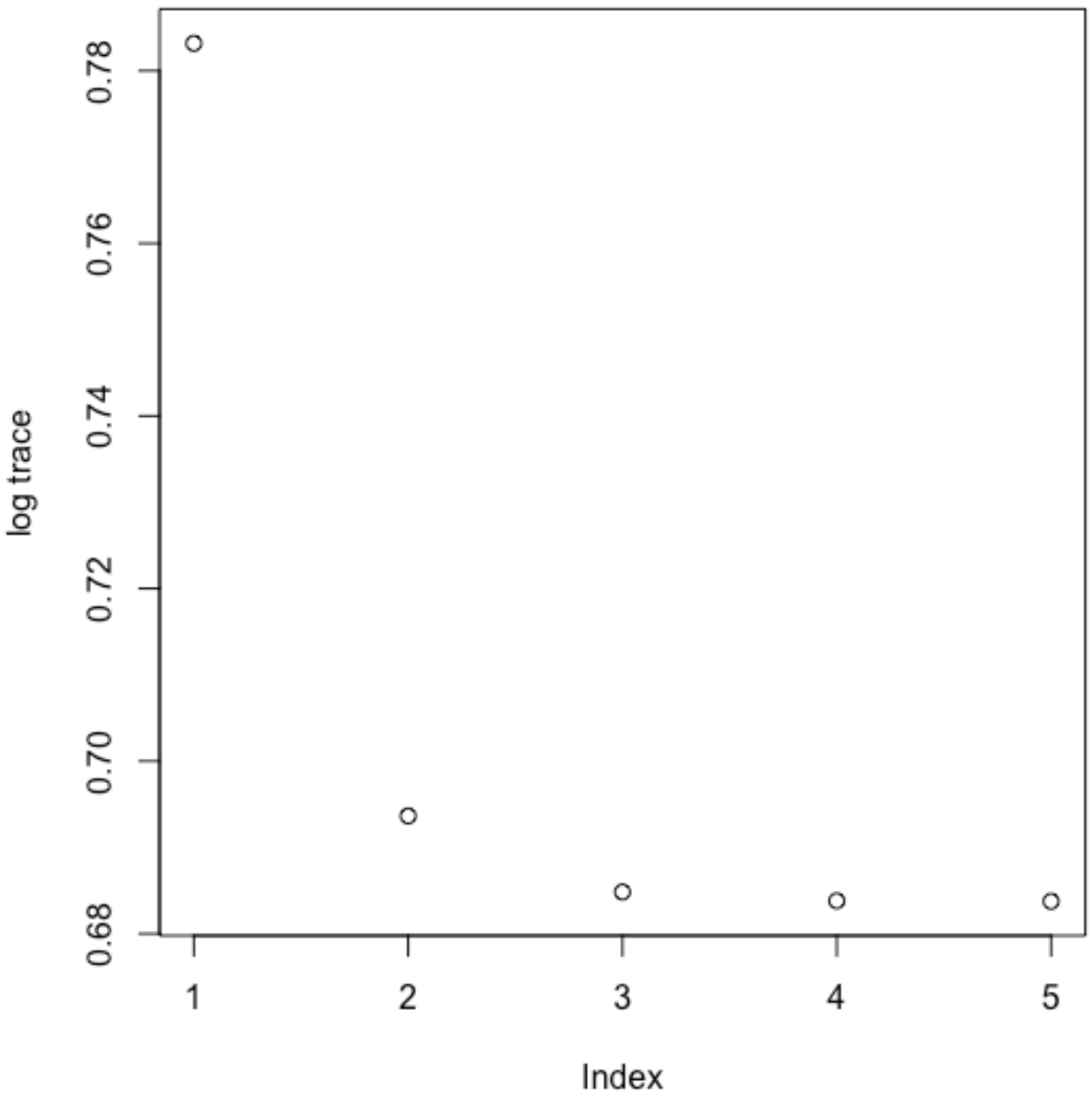
```
grid.search <- function(gs_y.v, gs_lossf, gs_f, gs_true){
  par(mar = c(5, 5, 2, 2))

  # Initial bounds
  a <- min(gs_y.v) ; b <- max(gs_y.v) ; K <- 10 ; eps <- 1e-
  iter.max <- 1e2

  plot(Vectorize(function(t) gs_f(gs_y.v, t, gs_lossf)), a,
       xlab = expression(mu))

  # Step.2 ~ 5
  trace.v <- rep(NA, iter.max)
  for (j in 1:iter.max){
    trace.v[j] <- gs_f(y.v, a, gs_lossf)
    x.v <- seq(a, b, length.out = K)
    f.v <- rep(NA, K)
    for (i in 1:K) {
    f.v[i] <- gs_f(y.v, x.v[i], gs_lossf)
    }
    opt <- which.min(f.v) # i*
    a <- x.v[opt - 1]
    b <- x.v[opt + 1]
    if (b - a < eps) break
  }
  points(c(a, b), c(f(gs_y.v, a, gs_lossf), f(gs_y.v, b, gs_
  legend('topleft', '(a,f(a))', pch = 19)
  plot(log(trace.v[1:j],exp(10)), ylab = 'log trace') # Is t
  return(list(a = a, b = b, mustar = gs_true))
}
grid.search(y.v, lossf1, f, mean(y.v))
```

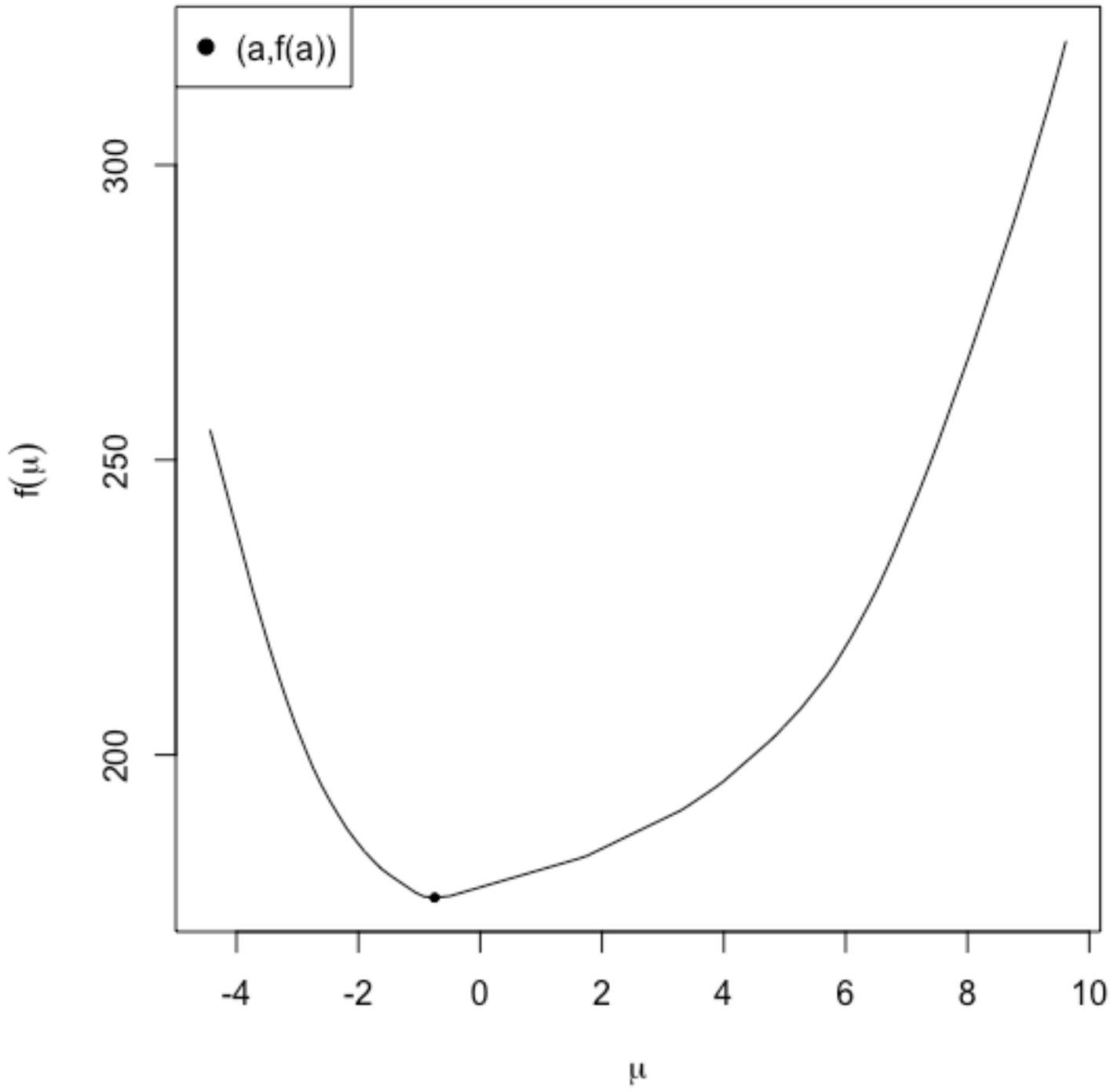

(a,f(a))
f(μ)
3500
3000
2500
2000
1500
1000
-4
-2
0
2
4
6
8
10
μ

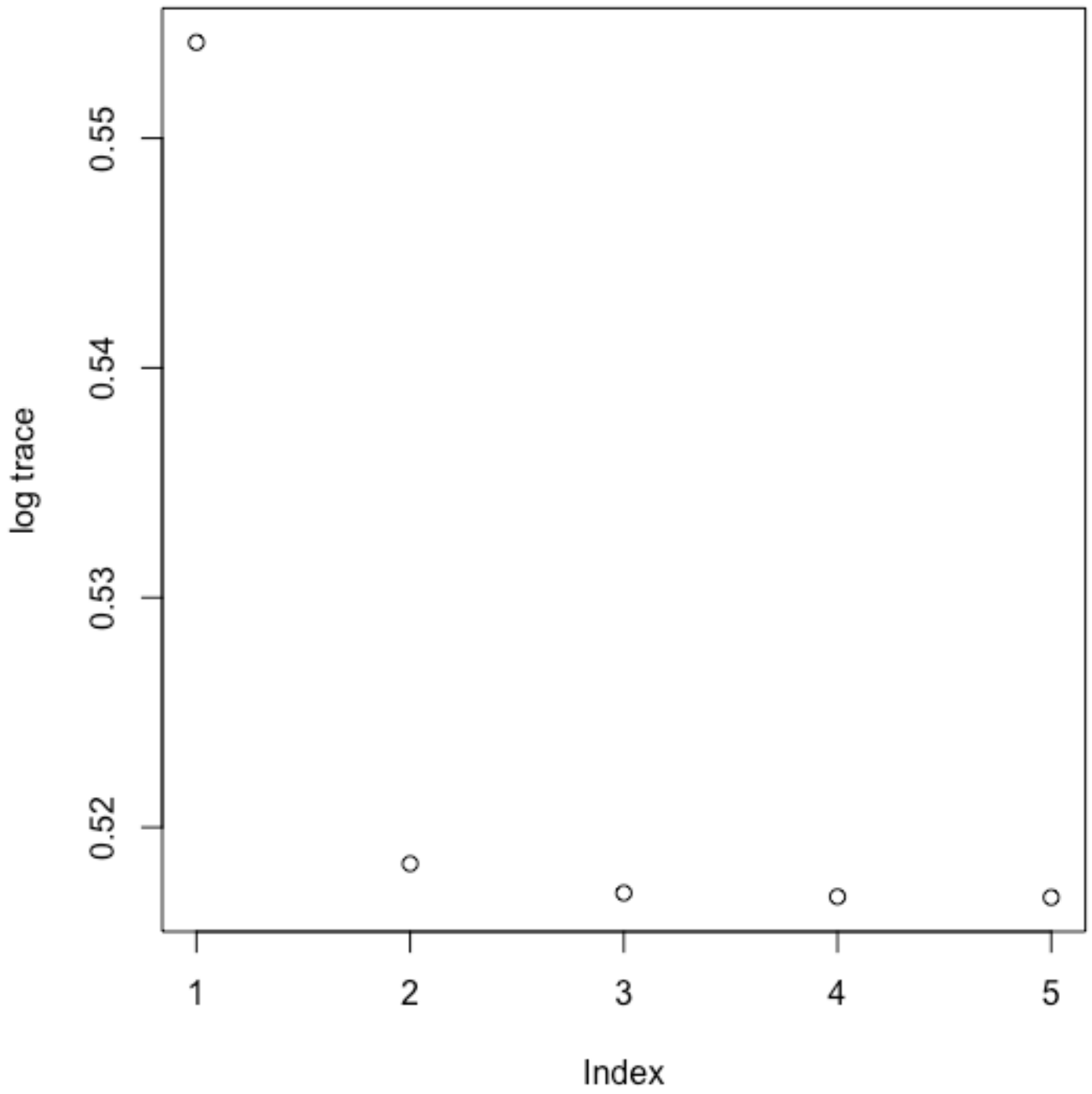



```
## $a
## [1] 1.782829
##
## $b
## [1] 1.790444
##
## $mustar
## [1] 1.787969
```

Absolute deviation Loss

1. $f(\mu) = \sum_i L(y_i - \mu)$ with $L(t) = |t|$
2. Initialize with $a = \min_i y_i$ and $b = \max_i y_i$

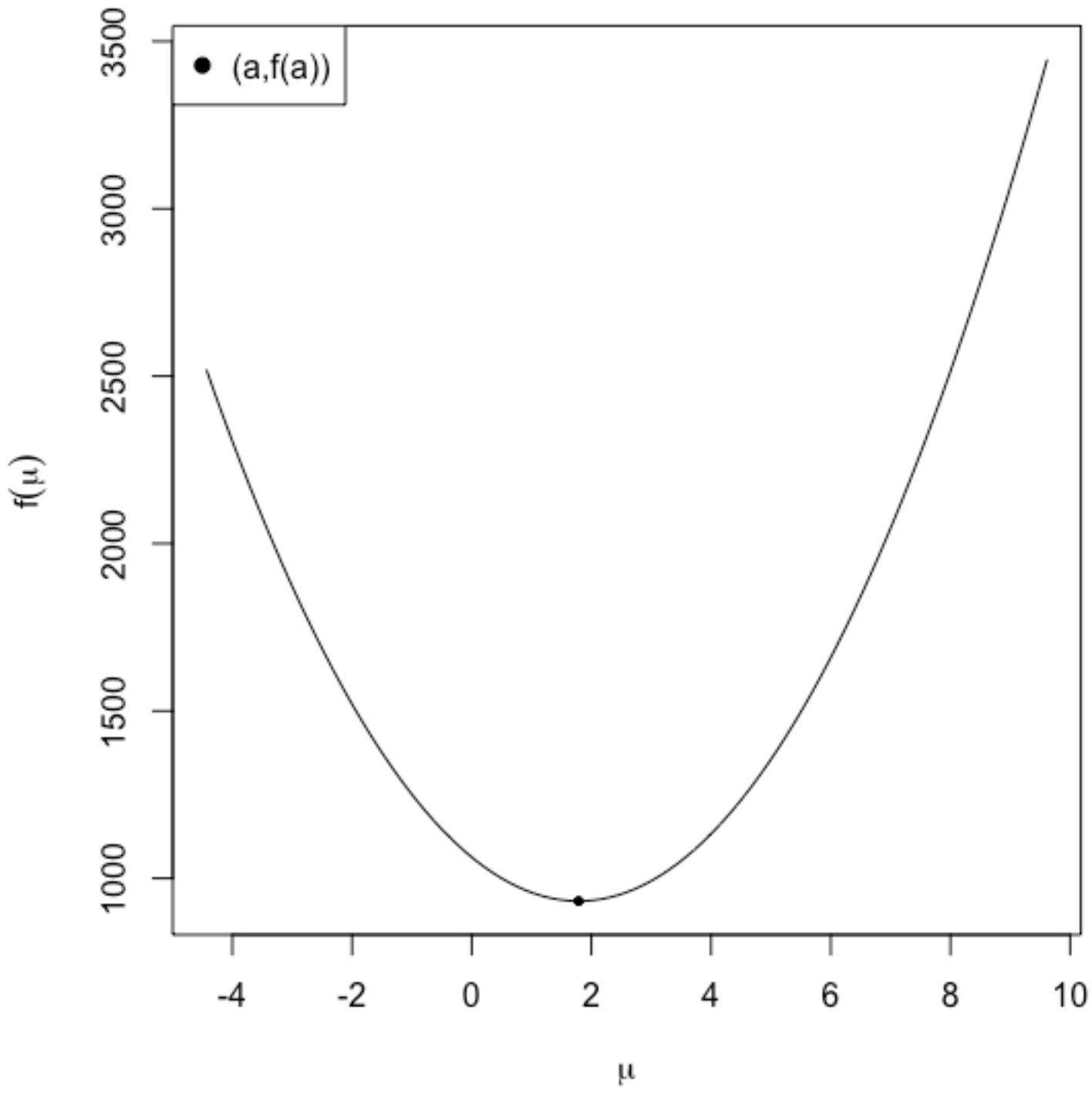


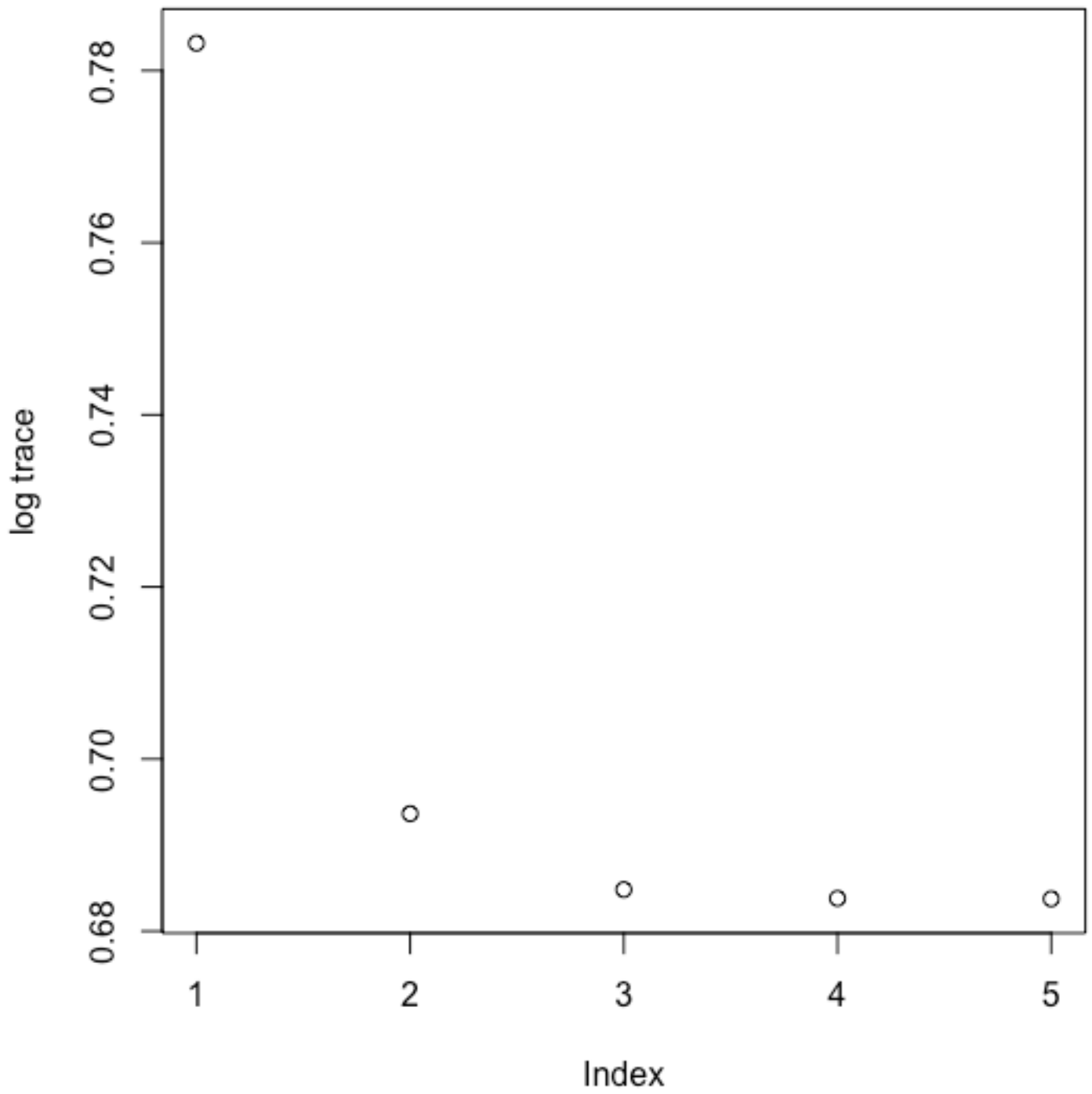



```
## $a
## [1] -0.7513671
##
## $b
## [1] -0.7437519
##
## $mustar
## [1] -0.7463344
```

LASSO

1. $f(\mu) = \sum_i L(y_i - \mu) + \lambda|\mu|$ for squared loss and $\lambda = 10^{-4}$
2. Initialize with $a = \min_i y_i$ and $b = \max_i y_i$
3. Check if $\mu^* = (or \approx) \; sign(\bar{y})(|\bar{y}| - \lambda)_+$






```
## $a
## [1] 1.782829
##
## $b
## [1] 1.790444
##
## $mustar
## [1] 1.787869
```