Homework of Dataminning, CH7

Zexian Wang, Student ID 15420151152805

2017-04-16

Q5

(a)

$\hat{g_2}$ will have smaller trainning RSS, because it can be regarded as a higher order polynomial due to the order of the derivative penalty function.

(b)

$\hat{g_1}$ will have smaller test RSS because $\hat{g_2}$ may overfit.

(c)

When $\lambda = 0$, $\hat{g_2} = \hat{g_1}$, so their training RSS and test RSS are the same.

Q11

(a)

```
set.seed(1)
n = 100
X1 <- rnorm(n)
X2 <- rnorm(n)
eps <- rnorm(100,mean=0,sd=1)
beta <- c(1,2,3)
Y <- cbind(rep(1,n),X1,X2)%*%beta + eps
```

(b)

```
beta1 <- 5
```

(c)

```
a <- Y - beta1*X1
beta2 <- lm(a~X2)$coef[2]
```

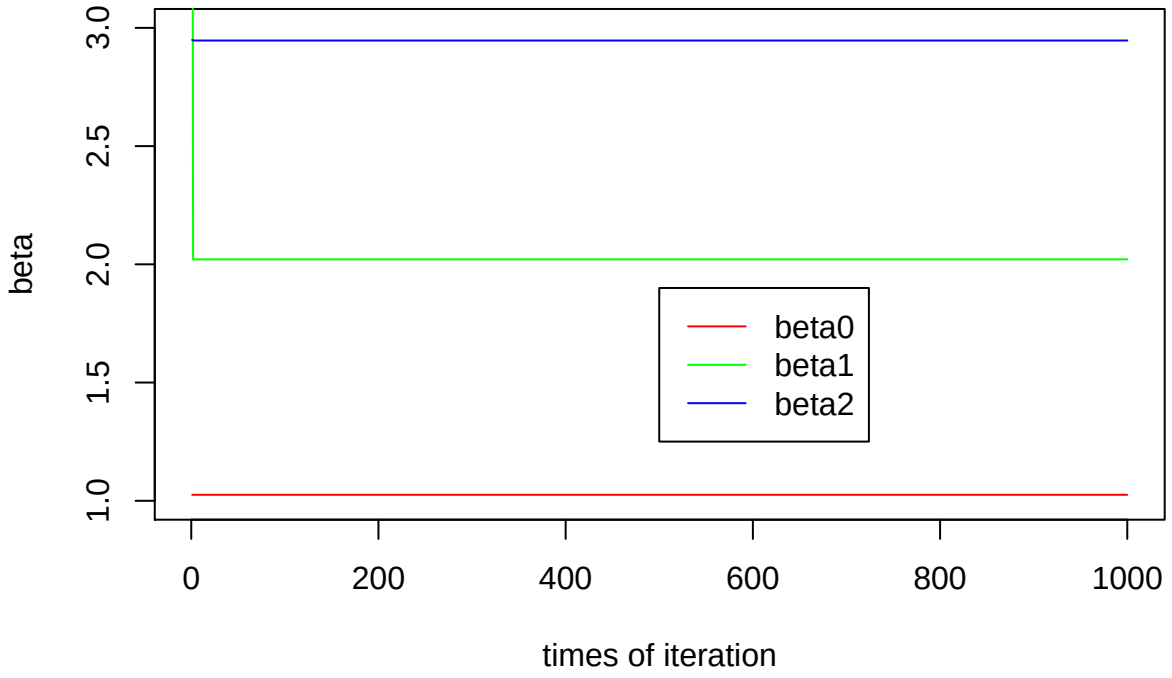
(d)

```
a <- Y - beta2*X2
beta1 <- lm(a~X1)$coef[2]
```

(e)

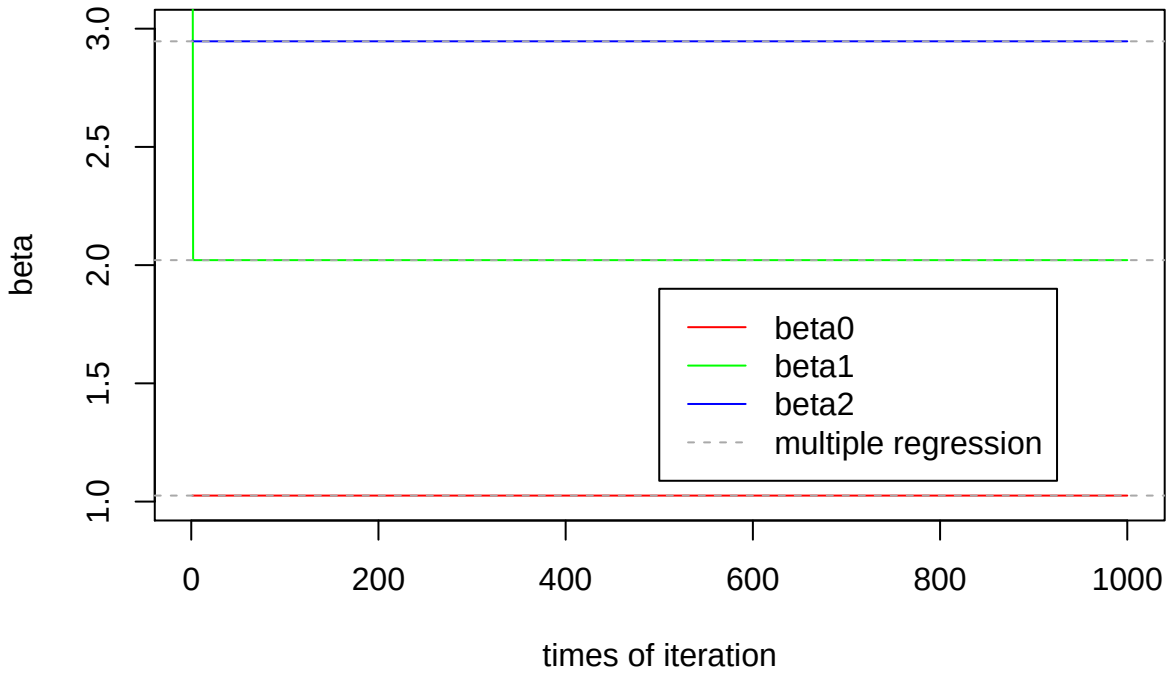
```
iter  <- 1000
beta0 <- rep(NA_real_, iter)
beta1 <- rep(NA_real_, iter)
beta2 <- rep(NA_real_, iter)
beta1[1] <- 5

for (i in 1:iter){
  a <- Y - beta1[i]*X1
```

```
  beta2[i] <- lm(a~X2)$coef[2]
  a <- Y - beta2[i]*X2
  fit <- lm(a~X1)
  beta0[i] <- fit$coef[1]
  if(i<1000){
    beta1[i+1] <- fit$coef[2]
  }
}
plot(1:iter, beta0, type = "l", xlab = "times of iteration", ylab = "beta", col = 'red', ylim = c(1,3))
lines(1:iter, beta1, col = "green")
lines(1:iter, beta2, col = "blue")
legend(500,1.9,c("beta0","beta1","beta2"),lty = c(1,1),col=c("red","green","blue"))
```



(f)

```
mulfit <- lm(Y~X1+X2)
plot(1:iter, beta0, type = "l", xlab = "times of iteration", ylab = "beta", col = 'red', ylim = c(1,3))
lines(1:iter, beta1, col = "green")
lines(1:iter, beta2, col = "blue")
abline(h=mulfit$coef[1], lty=2, col = "darkgrey")
abline(h=mulfit$coef[2], lty=2, col = "darkgrey")
abline(h=mulfit$coef[3], lty=2, col = "darkgrey")
legend(500,1.9,c("beta0","beta1","beta2", "multiple regression"),lty = c(1,1,1,2),col=c("red","green","
```



The estimated multiple regression coefficients match exactly with the coefficients obtained using backfitting.

(g)

Only one iteration is enough when the relationship between Y and X is linear.