

R Notebook

Problem 2

a.

```
# set.seed(12345)
# y <- seq(from=1, to=100, length.out = 1e8) + rnorm(1e8)
# z<-0
# for(i in 1:length(y)){
#   z<-z+(y[i]-mean(y))^2
# }
# z
```

b.

```
set.seed(12345)
y <- seq(from=1, to=100, length.out = 1e8) + rnorm(1e8)
sum((y-mean(y))^2)
```

```
## [1] 81774703375
```

We need to spend a lots of time to run the part a, however, part b only cost 10.5 seconds.

Problem 3

```
set.seed(1256)
theta <- as.matrix(c(1,2),nrow=2)
X <- cbind(1,rep(1:10,10))
h <- X%%theta+rnorm(100,0,0.2)
```

Problem 4

```
beta<-function(X,Y){
  beta<-inv((t(X)%*%X))%*%t(X)%*%Y
  return(beta)
}
```

I have written a function to compute beta of a linear regression model. Inputting X and Y value, we can get beta.

Problem5

```
# set.seed(12456)
# G <- matrix(sample(c(0,0.5,1),size=16000,replace=T),ncol=10)
# R <- cor(G) # R: 10 * 10 correlation matrix of G
# C <- kronecker(R, diag(1600)) # C is a 16000 * 16000 block diagonal matrix
# id <- sample(1:16000,size=932,replace=F)
# q <- sample(c(0,0.5,1),size=15068,replace=T) # vector of length 15068
# A <- C[id, -id] # matrix of dimension 932 * 15068
# B <- C[-id, -id] # matrix of dimension 15068 * 15068
# p <- runif(932,0,1)
# r <- runif(15068,0,1)
```

```
# C<-NULL  
# y<-p+A*%*inv(B)*%(q-r)
```