HW4(HW3 continued)

Getong Zhong (gz2338) & Ruoqiao Li (rl3288)

2023-02-13

1

```r
set.seed(1111)
mvn_gen <- function(n, mu, sigma, factorization = "Cholesky") {
  d <- length(mu)
  Z <- matrix(rnorm(n*d), nrow = n, ncol = d)
  if(factorization == "Cholesky") { Q <- chol(sigma) } else {
    if(factorization == "Spectral") {
      ev <- eigen(sigma)
      lambda <- ev$values
      P <- ev$vectors
      Q <- P %*% diag(sqrt(lambda)) %*% t(P) } else {
        stop("Arg factorization must be 'Cholesky' or 'Spectral'.")
      } }
  mu <- matrix(mu, nrow = d, ncol = 1)
  J = matrix(1, nrow = n, ncol = 1)
  X <- Z %*% Q + J %*% t(mu)
  return(data.frame(X))
}

sig <- matrix(c(1,0,0,0,0.2,0,0,0,0,0,
                0,1,0,0,0,0.9,0,0,0,0,
                0,0,1,0,0,0,0,0.2,0,0,
                0,0,0,1,0,0,0,0,0.9,0,
                0.2,0,0,0,1,0,0,0,0,0,
                0,0.9,0,0,0,1,0,0,0,0,
                0,0,0,0,0,0,1,0,0,0,
                0,0,0.2,0,0,0,0,1,0,0,
                0,0,0,0.9,0,0,0,0,1,0,
                0,0,0,0,0,0,0,0,0,1), 10, 10, byrow = TRUE)

dat_gen <- function(x=500){
  dat10 <- mvn_gen(n = x, mu = c(0, 0, 0, 0, 0, 0, 0, 0, 0, 0),
                   sigma = sig, factorization = "Cholesky")
  X11 <- c()
  for (i in 1:x){
    X11[i] <- (1 + exp(-(0 + dat10$X1[i] * 0.8 + dat10$X2[i]
    * (-0.25) + dat10$X3[i] * 0.6 + dat10$X4[i] * (-0.4)  +
      dat10$X5[i]* (-0.8) + dat10$X6[i] * (-0.5) +
      dat10$X7[i] * 0.7 +(-0.25) * dat10$X2[i] * dat10$X2[i])))^(-1)
```

```
  }
  dich <- runif(x)
  A <- c()
  for (j in 1:x){
    if (dich[j] < X11[j]) {
      A[j] <- 1
    } else if (dich[j] > X11[j]){
      A[j] <- 0
    }
  }
  Y <- c()
  for (k in 1:x){
    Y[k] <- 3.85 + 0.3 * dat10$X1[k] - 0.36 * dat10$X2[k]
    - 0.73 * dat10$X3[k] - 0.2 * dat10$X4[k] + 0.71 * dat10$X8[k]
    - 0.19 * dat10$X9[k] + 0.26 * dat10$X10[k] - 0.4 * A[k]
  }
  M <- cbind(dat10, X11, A, Y)
  return(M)
}
```

2

```
R = 1000
set.seed(1111)
dat_list <- replicate(n = R,
                      expr = dat_gen(),
                      simplify = FALSE)
```

3

```
x = 0
for(i in 1:1000){
  x = x + mean(dat_list[[i]]$X11*dat_list[[i]]$A)
}
true_average_treatment_effect <- x/1000
true_average_treatment_effect
```

```
## [1] 0.2923422
```

4

```
set.seed(100)
med_1 <- function(dat, method = 1){
  if (method == 1){
    return(mean(dat[dat$A == 1, ]$Y) - mean(dat[dat$A == 0, ]$Y))
  }else{
    mod <- lm(Y ~ . - X11 , data = dat)
```

```
    return(mod$coefficients["A"])
  }
}
med_1(dat = dat_gen(), 1)
```

```
## [1] 0.3102015
```

```
med_1(dat = dat_gen(), 2)
```

```
##            A
## 2.487573e-17
```

5

```
mse1 <- function(vec) {
  return(mean((med_1(vec) - true_average_treatment_effect)^2))
}

mse2 <- function(vec) {
  return(mean((med_1(vec,2) - true_average_treatment_effect)^2))
}
data.frame("method 1"=c(1,1),  "method 2"= c(2,2))
```

```
##   method.1 method.2
## 1        1        2
## 2        1        2
```

```
matrix <- data.frame("method 1"=c(mean(sapply(X = dat_list,FUN = mse1)),
                            mean(sapply(X = dat_list,FUN = med_1, simplify = TRUE) - true_average
                 "method 2" = c(mean(sapply(X = dat_list,FUN = mse2, simplify = TRUE)),
                            mean(sapply(X = dat_list,FUN = med_1, method = 2, simplify = TRUE))-
rownames(matrix) <- c("MSE", "Bias")
matrix
```

```
##          method.1    method.2
## MSE  0.001558053  0.08546397
## Bias 0.001448240 -0.29234222
```

6

One challenge that our group encountered was to figure out the true average treatment effect, which we finally computed as the mean of the production of X11 and A. Another challenge was to figure out the average treatment effect by using OLS regression. The result was negative after the first run and we realized that when using the OLS regression the variable X11 should be omitted and the result is no longer negative. From the MSE and bias matrix we can now that method 2 seems like have lower bias but higher mse, this is also known as a bias-variance trade-off. In this case, we would choose method 2 over method 1 since both the MSE are not big but method 1 has a higher bias value.