A3

Yuhan Hu

2019/11/15

Contents

Q1 **2**

a. 2

b. 3

c. 4

Q2 **6**

a. 6

b. 7

Q1

a.

we want to minimize $Var(V^TAV) = E[(V^TAV)^2] - tr(A)^2$ where $tr(A)^2$ can be regarded as constant here. So to minimize $Var(V^TAV)$ we want to minimize $E[(V^TAV)^2] = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n} a_{ij}a_{kl}E(V_iV_jV_kV_l)$ (As given in Hint)

Given V_i is iid with mean 0 and variance 1

Case1. $i = j = k = l$, $E(V_iV_jV_kV_l) = E(V_i)E(V_j)E(V_k)E(V_l) = 0 \times 0 \times 0 \times 0 = 0$

Case2. $i = j = k \neq l$,$E(V_iV_jV_kV_l) = E(V_i^3)E(V_l) = E(V_i^3) \times 0 = 0$

Case3. $i = j$,$k = l$ but $i, j \neq k, l$,$E(V_iV_jV_kV_l) = E(V_i^2)E(V_k^2)$ since $Var(V_i) = E(V_i^2) - E(V_i)^2, Var(V_i) = 1, E(V_i) = 0$,we have $E(V_i^2) = 1$; thus $E(V_i^2)E(V_k^2) = 1$

Case4. $i \neq j \neq k \neq l$, $E(V_iV_jV_kV_l) = E(V_i^4)$

All possible cases are listed above, so we know that for $E[(V^TAV)^2] = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n} a_{ij}a_{kl}E(V_iV_jV_kV_l)$, only case 3 and case 4 affect the result where effect of case 3 can be represents by a constant term.

Therefore, we have $E[(V^TAV)^2] = \sum_{i=1}^{n} a_{ii}^2 E(V_i^4) + constant$

We want to minimize $\sum_{i=1}^{n} a_{ii}^2 E(V_i^4)$ in order to minimize $Var(V^TAV)$

$Var(V_i^2) = E((V_i^2)^2 - E^2(V_i^2) = E(V_i^4) - E^2(V_i^2)$

$E(V_i^2) = 1$ is shown above, so $E^2(V_i^2) = 1^2 = 1$

so $E(V_i^4) = 1 + Var(V_i^2)$

Now we want to minimize $Var(V_i^2)$

$Var(V_i^2)$ is minimized when V_i^2 is constant. Given the form of V_i from question, we have V_i^2 always equals 1, which is a constant.

Therefore, $Var(V^TAV)$ is minimized.

b.

Given the H matrix $H = X(X^TX)^{-1}X^T = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix}$

so we have

$H \begin{bmatrix} V \\ 0 \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \begin{bmatrix} V \\ 0 \end{bmatrix} = V \begin{bmatrix} H_{11} \\ H_{21} \end{bmatrix} + 0 \begin{bmatrix} H_{12} \\ H_{22} \end{bmatrix} = \begin{bmatrix} H_{11}V \\ H_{21}V \end{bmatrix}$

Subsititute V by $H_{11}^{k-1}V$ in the above equation, we have $H \begin{bmatrix} H_{11}^{k-1}V \\ 0 \end{bmatrix} = \begin{bmatrix} H_{11}^{k}V \\ H_{21}H_{11}^{k-1}V \end{bmatrix}$

c.

```
leverage <- function(x1,x2,w,r=10,m=100) {
  qrx1 <- qr(x1)
  qrx2 <- qr(x2)
  n <- nrow(x1)
  lev1 <- NULL
  lev2 <- NULL
  for (i in 1:m) {
    v <- ifelse(runif(n)>0.5,1,-1)
    v[-w] <- 0
    v01 <- qr.fitted(qrx1,v)
    v02 <- qr.fitted(qrx2,v)
    f1 <- v01
    f2 <- v02
    for (j in 2:r) {
      v01[-w] <- 0
      v02[-w] <- 0
      v01 <- qr.fitted(qrx1,v01)
      v02 <- qr.fitted(qrx2,v02)
      f1 <- f1 + v01/j
      f2 <- f2 + v02/j
      }
    lev1 <- c(lev1,sum(v*f1))
    lev2 <- c(lev2,sum(v*f2))
    }
  se1 <- exp(-mean(lev1))*sd(lev1)/sqrt(m)
  se2 <- exp(-mean(lev2))*sd(lev2)/sqrt(m)
  diff <- 0.5*(exp(-mean(lev1))+exp(-mean(lev2)))*
    sd(lev1-lev2)/sqrt(m)
  lev1 <- 1 - exp(-mean(lev1))
  lev2 <- 1 - exp(-mean(lev2))
  r <- list(lev=c(lev1,lev2),std.err=c(se1,se2,diff))
  r
  }

x <- c(1:1000)/1000
X1 <- 1
for (i in 1:10) {
  X1 <- cbind(X1,cos(2*i*pi*x),sin(2*i*pi*x))}

library(splines) # loads the library of functions to compute B-splines

X2 <- cbind(1,bs(x,df=10))

for (j in 0:19){
  s <- 50*j+1
  e <- 50*(j+1)
  print(leverage(X1,X2,c(s:e),m=1000,r=15)$lev)
}
```

From the output, for first design, leverage does not change much when data points are approaching extreme points.For the second design, leverage decrease at the begining then increase when we are close to those extreme data points.

Q2

a.

Given X is gamma distribution with mean $E(X) = \alpha/\lambda$ and variance $Var(X) = \alpha/\lambda^2$

so $\lambda = \frac{E(X)}{Var(X)}$ and $\alpha = \frac{E^2(X)}{Var(x)}$

Therefore, $\hat{\lambda} = \frac{\bar{x}}{s^2}, \hat{\alpha} = \frac{\bar{x}^2}{s^2}$

b.

$L = \Pi_{i=1}^{n} f(x_i) = \frac{\lambda^{n\alpha}(X_i \dots X_n)^{\alpha-1} e^{-\lambda \sum X_i}}{[\Gamma(\alpha)]^n}$

$l = \ln L = n\alpha \ln \lambda + (\alpha - 1)\sum_{i=1}^{n} \ln X_i - \lambda \sum_{i=1}^{n} X_i - n \ln \Gamma(\alpha)$

$\frac{\partial l}{\partial \alpha} = n \ln \lambda + \sum_{i=1}^{n} \ln X_i - n\frac{\Gamma'(\alpha)}{\Gamma(\alpha)}$

Let $\frac{\Gamma'(\alpha)}{\Gamma(\alpha)} = \phi(\alpha)$

so $\frac{\partial l}{\partial \alpha} == \sum_{i=1}^{n} \ln X_i + n(\ln \lambda - \phi(\alpha))$

$\frac{\partial l}{\partial \lambda} = \frac{n\alpha}{\lambda} - \sum_{i=1}^{n} X_i$

Fisher matrix is $-Hessian = -\begin{bmatrix} \frac{\partial^2 l}{\partial^2 \alpha} & \frac{\partial^2 l}{\partial \alpha \partial \lambda} \\ \frac{\partial^2 l}{\partial \lambda \partial \alpha} & \frac{\partial^2 l}{\partial^2 \lambda} \end{bmatrix} = \begin{bmatrix} n\phi'(\alpha) & -\frac{n}{\lambda} \\ -\frac{n}{\lambda} & \frac{n\alpha}{\lambda^2} \end{bmatrix}$

```r
q2b <- function(x,eps=1.e-8,max.iter=50) {
  n <- length(x)
  x1 <- mean(x)^2/var(x)
  x2 <- mean(x)/var(x)
  theta <- c(x1,x2)
  score1 <- sum(log(x)) + n*(log(x2) - digamma(x1))
  score2 <- n*x1/x2-sum(x)
  score <- c(score1,score2)
  iter <- 1
  while (max(abs(score))>eps && iter<=max.iter) {

    info.11 <- n*trigamma(x1)
    info.12 <- -n/x2
    info.22 <- n*x1/x2^2
    info <- matrix(c(info.11,info.12,info.12,info.22),ncol=2)

    theta <- theta + solve(info,score)
    x1 <- theta[1]
    x2 <- theta[2]
    iter <- iter + 1

    score1 <- sum(log(x)) + n*(log(x2) - digamma(x1))
    score2 <- n*x1/x2-sum(x)
    score <- c(score1,score2)
    }
  if (max(abs(score))>eps) print("No convergence")
  else {
    print(paste("Number of iterations =",iter-1))

    info.11 <- n*trigamma(x1)
    info.12 <- -n/x2
    info.22 <- n*x1/x2^2
    info <- matrix(c(info.11,info.12,info.12,info.22),ncol=2)
    r <- list(alpha=x1,lambda=x2,info=info,varcovar=solve(info))
    r
    }
}
```

```r
x <- rgamma(100,shape=0.5)
r <- q2b(x)
r

x <- rgamma(100,shape=10)
r <- q2b(x)
r

x <- rgamma(100,shape=0.01)
r <- q2b(x)
r
```