University of Toronto

Coursework & Homework Demonstration

YiFei Gu

Q1

(a) Want to show the joint density of $g(u, x) = \frac{u}{|\mathcal{C}_h|}$; Given $f(u, v)$, let $v/u = x$, then by change of variable, the Jacobian of which is:

$$\mathcal{J}(u, x) = \det \begin{vmatrix} \frac{\partial u}{\partial u} & \frac{\partial u}{\partial x} \\ \frac{\partial v}{\partial u} & \frac{\partial v}{\partial x} \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ x & u \end{vmatrix} = u$$

And thus the joint density is then:

$$g(u, x) = f(u, v)u = \frac{u}{|\mathcal{C}_h|} \ , \ 0 \leq u \leq \sqrt{h(x)}$$

The marginal density is given by:

$$f_X(x) = \int_{-\infty}^{\sqrt{h(x)}} \frac{u}{|\mathcal{C}_h|} \ du = \frac{h(x)}{2|\mathcal{C}_h|}$$

And thus the density of X is $\gamma h(x)$ where $\gamma = \frac{1}{|2\mathcal{C}_h|}$

(b) Let $v/u = x$, Note that $(u, v) \in \mathcal{C}_h$ with $0 \leq u \leq \sqrt{h(x)}$, then u must within the range of $[0, \max_x \sqrt{h(x)}]$
Since $u \leq \sqrt{h(x)}$, then $1 \leq \frac{1}{u}\sqrt{h(x)}$ and $v = ux \leq x\sqrt{h(x)}$.
If $v \geq 0$, then

$$v = ux \leq \max_x x\sqrt{h(x)} = v_+$$

And if $v \leq 0$, then

$$v = ux \geq \min_x x\sqrt{h(x)} = v_-$$

And thus, $(u, v) \in \mathcal{C}_h$ must lie in a rectangle $[0, u_+] \times [v_-, v_+]$

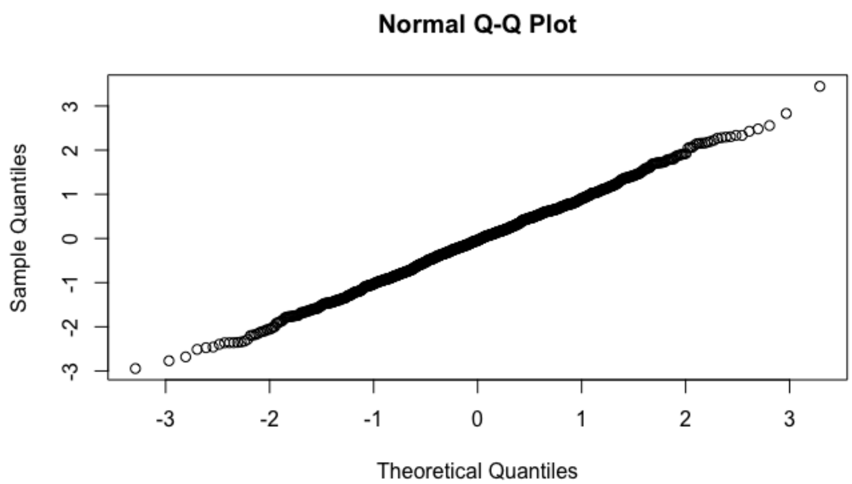
(c) The rejection rate is given by $\frac{|C_h|}{|\mathcal{D}_h|}$, with $h(x) = \exp(-x^2/2)$. Then:

$$\mathcal{C}_h = \frac{1}{2} \int_{-\infty}^{\infty} h(x) \ d(x) = \frac{1}{2} \int_{-\infty}^{\infty} \exp(x^2/2) \ d(x) = \sqrt{\pi/2}$$

Note that $|\mathcal{D}_h| = |\sqrt{2/e}| + \sqrt{2/e} = 2\sqrt{2/e}$. Thus, the theoretical rejection rate is $\frac{|C_h|}{|\mathcal{D}_h|} = \frac{\sqrt{\pi/2}}{2\sqrt{2/e}} \approx 0.7306$.

```
> rnm <- function(n) {
+     b <- sqrt(2/exp(1))
+     r <- 0
+     x <- NULL
+     for (i in 1:n) {
+         reject <- T
+         while (reject) {
+             u <- runif(1)
+             v <- runif(1,-b,b)
```

```
+ if (u<=exp(-(v/u)^2/4)) {
+     x <- c(x, v/u)
+     reject <- F
+ }
+ else r <- r +1
+ }
+ }
+ accept.rate <- n/(n+r)
+ qqnorm(x)
+ print(accept.rate)
+ }
> rnm(100000)
[1] 0.7267442
```



Q2

(a) Note that since i and n are arbitrary, it suffices to show that, for all i, if y_i is linear, then $\hat{\theta}_i = y_i$. With y_i linear implies $y_i = a \times i + b$, then the objective function becomes:

$$\sum_{i=1}^{n}(y_i - \theta_i)^2 + \lambda \sum_{i=2}^{n-1} (a \times (i+1) + b - 2(a \times i + b) + a \times (i-1) + b) = \sum_{i=1}^{n}(y_i - \theta_i)^2$$

Thus the above equation is minimized with $\hat{\theta}_i = y_i$, with objective function being 0.

(b) Let $||\mathbf{y}^* - \mathbf{X}\theta||^2 = \sum_{i=1}^{n}(y_i - \theta_i)^2 + \lambda \sum_{i=2}^{n-1}(\theta_{i+1} - 2(\theta_i) + \theta_{i-1})^2$, this gives

$$||\mathbf{y}^* - \mathbf{X}\theta||^2 = \begin{bmatrix} y_1 - \theta_1 \\ y_2 - \theta_2 \\ ... \\ y_n - \theta_n \\ ... \\ \lambda(\theta_3 - 2\theta_2 + \theta_1) \\ \lambda(\theta_4 - 2\theta_3 + \theta_2) \\ ... \\ \lambda(\theta_n - 2\theta_{n-1} + \theta_{n-2}) \end{bmatrix}$$

This gives:

$$\mathbf{y}^* = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ ... \\ y_n \\ 0 \\ 0 \\ ... \\ 0 \end{bmatrix} \in \mathbb{R}^{n+n-1-2+1} = \mathbb{R}^{2n-2}, \mathbf{X} = \begin{bmatrix} 1 & 0 & 0 & ... & ... & 0 \\ 0 & 1 & 0 & ... & ... & 0 \\ 0 & 0 & 1 & ... & ... & 0 \\ 0 & ... & ... & ... & ... & 0 \\ 0 & 0 & 0 & ... & ... & 1 \\ -\lambda & 2\lambda & \lambda & ... & ... & 0 \\ 0 & -\lambda & 2\lambda & \lambda & ... & 0 \\ ... & ... & ... & ... & ... & ... \\ 0 & 0 & ... & -\lambda & 2\lambda & \lambda \end{bmatrix} \in \mathbb{R}^{2n-2\times n}, \theta = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ . \\ . \\ . \\ \theta_{n-1} \\ \theta_n \end{bmatrix} \in \mathbb{R}^n$$

(c) Want to show the objective function is non-increasing from one iteration to the next.
Assume in the ith iteration, p was chosen out of n, denoted w, with its complement $\bar{w}$.
For the $i+1$th iteration, let the subset chosen denotes $w+k$, with its complement $\bar{w}-k$. Note that during the ith iteration in step 2, $\hat{\theta}_w$ was defined to minimize the objective function, with respect to θ_w.
In the $i+1$th iteration, $\hat{\theta}_{w+k}$ is defined for minimization, however, since $\hat{\theta}_w \cap \hat{\theta}_{w+k} \neq \emptyset$, the minimization takes non increasing steps in each iteration as a subset of previous minimization process is contained. Thus the function is non-increasing.

(d)
```
> HP <- function(x,lambda,p=20,niter=200) {
+     n <- length(x)
+     a <- c(1,-2,1)
+     aa <- c(a,rep(0,n-2))
+     aaa <- c(rep(aa,n-3),a)
+     mat <- matrix(aaa,ncol=n,byrow=T)
+     mat <- rbind(diag(rep(1,n)),sqrt(lambda)*mat)
+     xhat <- x
+     x <- c(x,rep(0,n-2))
+     sumofsquares <- NULL
+     for (i in 1:niter) {
+         w <- sort(sample(c(1:n),size=p))
+         xx <- mat[,w]
+         y <- x - mat[,-w]%*%xhat[-w]
+         r <- lsfit(xx,y,intercept=F)
+         xhat[w] <- r$coef
+         sumofsquares <- c(sumofsquares,sum(r$residuals^2))
+     }
+     r <- list(xhat=xhat,ss=sumofsquares)
+     r
```

```
+ }
> r1 <- HP(data,lambda=2000,p=50,niter=1000)
> r2 <- HP(data,lambda=2000,p=40,niter=1000)
> r2 <- HP(data,lambda=2000,p=30,niter=1000)
> r50 <- HP(data,lambda=2000,p=50,niter=1000)
> r40 <- HP(data,lambda=2000,p=40,niter=1000)
> r30 <- HP(data,lambda=2000,p=30,niter=1000)
> r10 <- HP(data,lambda=2000,p=10,niter=1000)
> plot(r10$ss)
> plot(r30$ss)
> plot(r40$ss)
> plot(r50$ss)
```

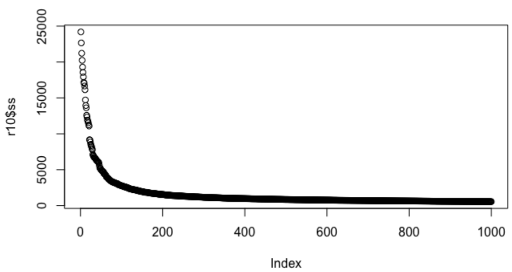


Figure 1: p = 10

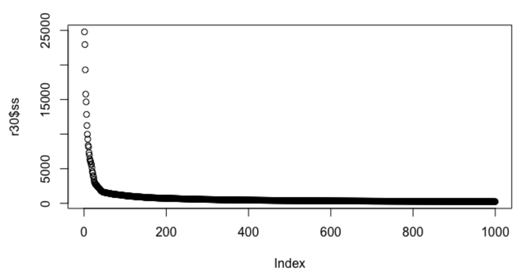


Figure 2: p = 30

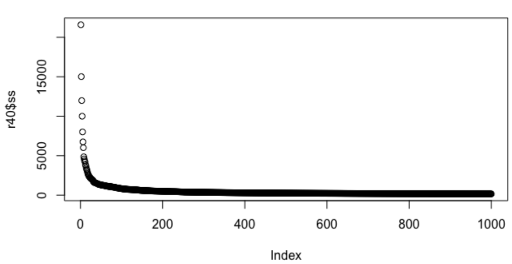


Figure 3: p = 40

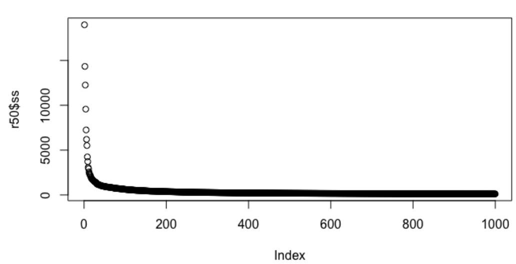


Figure 4: p = 50

As shown above, the objective function decreases more drastically as value of p increases. The rate of convergence is positively correlated with value of p.