```
#####Problem 1
library("cluster")
library("kernlab")
reduced_voting_record2005 <- read.delim("~/Dropbox/School/Statistics/Stat 154 Spring 2014/HW2/
reduced_voting_record2005.txt", header=F)
reduced_house_party2005 <- na.omit(t(read.delim("~/Dropbox/School/Statistics/Stat 154 Spring 2014/
HW2/reduced_house_party2005.txt", header=F)))
reduced_voting_record2005$V670 = reduced_house_party2005
colorsRep = 1:401
colorsRep[reduced_house_party2005==1] = "blue"
colorsRep[reduced_house_party2005==0] = "red"
colorsRep[reduced_house_party2005==2] = "grey"
party_votes = matrix(nrow=3, ncol=669)
row.names(party_votes) = c("dem", "rep", "indep")
for(i in 1:669){
  party_votes[1, i] = 2*sum(reduced_voting_record2005[,i][reduced_voting_record2005$V670==1])
  party_votes[2, i] = 2*sum(reduced_voting_record2005[,i][reduced_voting_record2005$V670==0])
  party_votes[3, i] = 2*sum(reduced_voting_record2005[,i][reduced_voting_record2005$V670==2])
}
Prob1a = hclust(dist(party_votes))
plot(Prob1a)
Problb = prcomp(reduced_voting_record2005)
plot(Problb$x, col = colorsRep)
Prob1c = kpca(as.matrix(reduced_voting_record2005))
plot(rotated(Prob1c), col=colorsRep)
Prob1d=specc(as.matrix(reduced_voting_record2005), centers = 3)
plot(Prob1d, col = colorsRep)
####Problem 2
library(FNN)
N=10
error1 = rnorm(N, 0, sqrt(1))
Problem2.1 = function(N, k, sigma, test = 2*pi/100*(1:100), error = error1){
  i = 1:N
  x = (i - 1/2)/N*2*pi
  f = cos(10*x)+2
  y = f + error
  results = list()
  for(i in (1:length(test))){
    firstset = list()
    firstset[[1]] = knn.reg(train = x, test = test[i], y=y, k = k)$pred
    firstset[[2]] = x
    firstset[[3]] = y
    results[[i]] = firstset
  }
  return(results)
}

firstk =c()
secondk = c()
thirdk = c()
P2.1 = Problem2.1(10, 1, 0.1)
P2.2 = Problem2.1(10, 3, 0.1)
P2.3 = Problem2.1(10, 10, 0.1)
for(i in 1:100){
  firstk = c(firstk, P2.1[[i]][[1]])
```

```
  secondk= c(secondk, P2.2[[i]][[1]])
  thirdk = c(thirdk, P2.3[[i]][[1]])
}

plot(x = 2*pi/100*(1:100), firstk, type="l")
plot(x = 2*pi/100*(1:100), secondk, type="l")
plot(x = 2*pi/100*(1:100), thirdk, type="l")
###Problem 2.1.1
EPE.X= function(k, f){
  i = 1:N
  x = (i - 1/2)/N*2*pi
    Bias.x = (f(pi, n = 1)-1/k*sum(f(x[order(abs(X[i]-X))[1:k]])))^2
    Var.x = 1/k
    EPE= 1 + Bias.x + Var.x
    Bias[i] = Bias.x
    Var = c(Var, Var.x)
  final.df = data.frame(Bias, Var, EPE)
  return(final.df)
}
N=10
X = runif(N, 0, 2*pi)
f0 = function(x, n = N){
  cos(10*x)+2+ rnorm(n, 0, sqrt(1))}



i = 1:N
x = (i - 1/2)/N*2*pi
y = f0(x)
y1 = lm(y ~ 1)
y1form.X = function(x, n=N){
  y1$coefficients[1]+ rnorm(n, 0, sqrt(1))}

y2 = lm(y ~ 1 + x)
y2form.X = function(x, n=N){
  y2$coefficients[2]*x+y2$coefficients[1]+ rnorm(n, 0, sqrt(1))}

y3 = lm(y ~ 1 + x + I(x^2))
y3form.X = function(x, n=N){
  y3$coefficients[3]*x^2+y3$coefficients[2]*x+y3$coefficients[1] + rnorm(n, 0, sqrt(1))
}
EPE.X(1, f0)
EPE.X(3, f0)
EPE.X(10, f0)

EPE.X(1, y1form.X)
EPE.X(3, y1form.X)
EPE.X(10, y1form.X)

EPE.X(1, y2form.X)
EPE.X(3, y2form.X)
EPE.X(10, y2form.X)

EPE.X(1, y3form.X)
EPE.X(3, y3form.X)
```

```
EPE.X(10, y3form.X)

y = f0(X)
y1 = lm(y ~ 1)
y1form.X = function(x){
  y1$coefficients[1]+ rnorm(N, 0, sqrt(1))}

y2 = lm(y ~ 1 + X)
y2form.X = function(x){
  y2$coefficients[2]*x+y2$coefficients[1]+ rnorm(N, 0, sqrt(1))}

y3 = lm(y ~ 1 + X + I(X^2))
y3form.X = function(x){
  y3$coefficients[3]*x^2+y3$coefficients[2]*x+y3$coefficients[1]
}
y = y3form.X(X)

EPE.X.X= function(X, k, f){
  Bias = c()
  Var = c()
  EPE = c()
  for(i in 1:length(X)){
    Bias.x = (f(X[i]) - 1/k*sum(f(
         X[order(abs(X[i]-X))[1:k]])))^2
    Var.x = 1/k
    EPE[i] = 1 + Bias.x + Var.x
    Bias[i] = Bias.x
    Var = c(Var, Var.x)
}
return(data.frame(EPE, Bias, Var))
}
EPE.X.X(X, 1, f0)
sapply(EPE.X.X(X, 1, f0), mean)
EPE.X.X(X, 3, f0)
sapply(EPE.X.X(X, 3, f0), mean)
EPE.X.X(X, 10, f0)
sapply(EPE.X.X(X, 10, f0), mean)

sapply(EPE.X.X(X, 1, y1form.X), mean)
sapply(EPE.X.X(X, 3, y1form.X), mean)
sapply(EPE.X.X(X, 10, y1form.X), mean)

sapply(EPE.X.X(X, 1, y2form.X), mean)
sapply(EPE.X.X(X, 3, y2form.X), mean)
sapply(EPE.X.X(X, 10, y2form.X), mean)

sapply(EPE.X.X(X, 1, y3form.X), mean)
sapply(EPE.X.X(X, 3, y3form.X), mean)
sapply(EPE.X.X(X, 10, y3form.X), mean)


###For EPE(pi)), where X is unformily distributed from 0 to 2*pi, the linear function gives the lowest
output;
####however, the quadratic is only marginally higher.
```

```
###For E(EPE(X)), where X is unformily distributed from 0 to 2*pi, the quadratic function gives the lowest
output;
####however, the linear is only marginally higher.
###Problem 2.2
N=100
error1 = rnorm(N, 0, sqrt(1))
Problem2.1 = function(N, k, sigma, test = 2*pi/100*(1:100), error = error1){
  i = 1:N
  x = (i - 1/2)/N*2*pi
  f = cos(10*x)+2
  y = f + error
  results = list()
  for(i in (1:length(test))){
    firstset = list()
    firstset[[1]] = knn.reg(train = x, test = test[i], y=y, k = k)$pred
    firstset[[2]] = x
    firstset[[3]] = y
    results[[i]] = firstset
  }
  return(results)
}

firstk =c()
secondk = c()
thirdk = c()
P2.1 = Problem2.1(10, 1, 0.1)
P2.2 = Problem2.1(10, 3, 0.1)
P2.3 = Problem2.1(10, 10, 0.1)
for(i in 1:100){
  firstk = c(firstk, P2.1[[i]][[1]])
  secondk= c(secondk, P2.2[[i]][[1]])
  thirdk = c(thirdk, P2.3[[i]][[1]])
}

plot(x = 2*pi/100*(1:100), firstk, type="l")
plot(x = 2*pi/100*(1:100), secondk, type="l")
plot(x = 2*pi/100*(1:100), thirdk, type="l")
###Problem 2.1.1
EPE.X= function(k, f){
  i = 1:N
  x = (i - 1/2)/N*2*pi
  Bias.x = (f(pi, n = 1)-1/k*sum(f(x[order(abs(X[i]-X))[1:k]])))^2
  Var.x = 1/k
  EPE= 1 + Bias.x + Var.x
  Bias[i] = Bias.x
  Var = c(Var, Var.x)
  final.df = list(Bias, Var, EPE)
  return(final.df)
}
N=100
X = runif(N, 0, 2*pi)
f0 = function(x, n = N){
  cos(10*x)+2+ rnorm(n, 0, sqrt(1))}
```

```
i = 1:N
x = (i - 1/2)/N*2*pi
y = f0(x)
y1 = lm(y ~ 1)
y1form.X = function(x, n=N){
  y1$coefficients[1]+ rnorm(n, 0, sqrt(1))}

y2 = lm(y ~ 1 + x)
y2form.X = function(x, n=N){
  y2$coefficients[2]*x+y2$coefficients[1]+ rnorm(n, 0, sqrt(1))}

y3 = lm(y ~ 1 + x + I(x^2))
y3form.X = function(x, n=N){
  y3$coefficients[3]*x^2+y3$coefficients[2]*x+y3$coefficients[1] + rnorm(n, 0, sqrt(1))
}
EPE.X(1, f0)[1,]
EPE.X(3, f0)[1,]
EPE.X(10, f0)[1,]
EPE.X(20, f0)[1,]
EPE.X(50, f0)[1,]


EPE.X(1, y3form.X)[1,]
EPE.X(3, y3form.X)[1,]
EPE.X(10, y3form.X)[1,]
EPE.X(20, y3form.X)[1,]
EPE.X(50, y3form.X)[1,]


y6 = lm(y ~ 1 + x + I(x^2)+I(x^3)+I(x^4)+I(x^5))
y6form.X = function(x, n=N){
  sum(as.numeric(y6$coefficients)*(rep(x, 6))^(0:5))
}
EPE.X(1, y6form.X)
EPE.X(3, y6form.X)
EPE.X(10, y6form.X)
EPE.X(20, y6form.X)
EPE.X(50, y6form.X)

y = f0(X)
y1 = lm(y ~ 1)
y1form.X = function(x){
  y1$coefficients[1]+ rnorm(N, 0, sqrt(1))}

y2 = lm(y ~ 1 + X)
y2form.X = function(x){
  y2$coefficients[2]*x+y2$coefficients[1]+ rnorm(N, 0, sqrt(1))}

y3 = lm(y ~ 1 + X + I(X^2))
y3form.X = function(x){
  y3$coefficients[3]*x^2+y3$coefficients[2]*x+y3$coefficients[1]
}
y = y3form.X(X)
```

```
y6 = lm(y ~ 1 + X + I(X^2)+I(X^3)+I(X^4)+I(X^5))
y6form.X = function(x, n=N){
  sum(as.numeric(y6$coefficients)*(rep(x, 6))^(0:5))
}

EPE.X.X= function(X, k, f){
  Bias = c()
  Var = c()
  EPE = c()
  for(i in 1:length(X)){
    Bias.x = (f(X[i]) - 1/k*sum(f(
      X[order(abs(X[i]-X))[2:(k+1)]])))^2
    Var.x = 1/k
    EPE[i] = 1 + Bias.x + Var.x
    Bias[i] = Bias.x
    Var = c(Var, Var.x)
  }
  return(data.frame(EPE, Bias, Var))
}
EPE.X.X(X, 1, f0)
sapply(EPE.X.X(X, 1, f0), mean)
EPE.X.X(X, 3, f0)
sapply(EPE.X.X(X, 3, f0), mean)
EPE.X.X(X, 10, f0)
sapply(EPE.X.X(X, 10, f0), mean)

sapply(EPE.X.X(X, 1, y1form.X), mean)
sapply(EPE.X.X(X, 3, y1form.X), mean)
sapply(EPE.X.X(X, 10, y1form.X), mean)

sapply(EPE.X.X(X, 1, y2form.X), mean)
sapply(EPE.X.X(X, 3, y2form.X), mean)
sapply(EPE.X.X(X, 10, y2form.X), mean)

sapply(EPE.X.X(X, 1, y3form.X), mean)
sapply(EPE.X.X(X, 3, y3form.X), mean)
sapply(EPE.X.X(X, 10, y3form.X), mean)

sapply(EPE.X.X(X, 1, y6form.X), mean)
sapply(EPE.X.X(X, 3, y6form.X), mean)
sapply(EPE.X.X(X, 10, y6form.X), mean)
sapply(EPE.X.X(X, 20, y6form.X), mean)
sapply(EPE.X.X(X, 50, y6form.X), mean)


###Probem2.2.2

EPE.X.X= function(X, k, f){
  Bias = c()
  Var = c()
  EPE = c()
  for(i in 1:length(X)){
    Bias.x = (
      f(X[i]) -
        1/k*sum(f(
```

```
        X[order(abs(X[i]-X))[1:k]]
        ))
           )^2
    Var.x = 1/k
    EPE[i] = 1 + Bias.x + Var.x
    Bias[i] = Bias.x
    Var = c(Var, Var.x)
}
return(data.frame(Bias, Var, EPE))
}


###Problem 3
N = 50
x = runif(5, 0, 2*pi)
f0 = function(x, k){
  summing = c()
  for(i in 1:4){
    summing[i] = sum(sin(sqrt(k)*x[i]))+sum(cos(x[i]*x[(i+1)]))
}
return(summing)
}
y = f0(x, 1)

y1 = lm(y ~ 1 + x[1:4])
y1form.X = function(x, n = N){
  y1$coefficients[2]*x+ y1$coefficients[1]}
EPE.X.X(x, 1, y1form.X)

y = f0(x, 4)

y1 = lm(y ~ 1 + x)
y1form.X = function(x){
  y1$coefficients[2]*x+ y1$coefficients[1]}
EPE.X.X(x, 4, y1form.X)

y = f0(x, 5)

y1 = lm(y ~ 1 + x)
y1form.X = function(x){
  y1$coefficients[2]*x+ y1$coefficients[1]}

EPE.X.X= function(X, k, f){
  Bias = c()
  Var = c()
  EPE = c()
  for(i in 1:length(X)){
    Bias.x = (f(X[i])-1/k*sum(f(x[order(abs(X[i]-X))[1:k]])))^2
    Var.x = 1/k
    EPE[i] = 1 + Bias.x + Var.x
    Bias[i] = Bias.x
    Var = c(Var, Var.x)
  }
  final.df = data.frame(Bias, Var, EPE)
  return(final.df)
```

```
}

EPE.X.X(x, 5, y1form.X)
####Problem 4
Problem4 = data.frame(v1=rnorm(100, 0, sqrt(1)))
for( i in 2:1000){
  Problem4[,i]=rnorm(100, 0, sqrt(1))
}
i0 = sample(x=1:1000, 1)
Xi0= Problem4[,i0]
Xj= Problem4[,-c(i0)]

diff.Xi0.Xj = c()
for(i in 1:999){
  diff.Xi0.Xj[i] = sqrt(sum((Xi0-Xj[,i])^2))
}
1/sqrt(100)*min(diff.Xi0.Xj)
1/sqrt(100)*max(diff.Xi0.Xj)
```