MA 471: Lab Assignment 03

Due on Monday, August 21, 2017

A. K. Dey

Sai Teja Talasila
140123040

Problem 1

R Code :

```
rmixnormal <- function(n,p,mu1,s1,mu2,s2)
{
      res = vector(,n)
      for(i in 1:n)
      {
            u = runif(1)
            if(u<p)
                  res[i] = rnorm(1,mu1,s1)
            else
                  res[i] = rnorm(1,mu2,s2)
      }
      return(res)
}
dmixnorm <- function(x,p,mu1,s1,mu2,s2)
{
      return(p*dnorm(x,mu1,s1) + (1-p)*dnorm(x,mu2,s2))
}
dmynorm <- function(X,mu,s)
{
      res = vector(,length(X))
      for(i in 1:length(X))
      {
            res[i] = dnorm(X[i],mu,s)
      }
      return(res)
}
EMmixnorm <- function(X,maxiter=1000)
{
      n = length(X)
      p = runif(1)
      mu1 = rnorm(1,0,1)
      mu2 = rnorm(1,0,1)
      s1 = rexp(1,1)
      s2 = rexp(1,1)
      m = vector(,n)
      for(t in 1:maxiter)
      {
            m = (p*dmynorm(X,mu1,s1))/(p*dmynorm(X,mu1,s1) + (1-p)*dmynorm(X,mu2,s2))
            m[is.na(m)] = 0.5

            p = mean(m)
            mu1 = (m%*%X)/sum(m)
            mu2 = ((1-m)%*%X)/sum(1-m)
            s1 = (m%*%((X-mu1)^2))/sum(m)
            s1 = sqrt(s1)
            s2 = ((1-m)%*%((X-mu2)^2))/sum(1-m)
            s2 = sqrt(s2)
      }
      return(c(p,mu1,s1,mu2,s2))
```

```
}
X = rmixnormal(200,0.4,0,1,0,5)
v = EMmixnorm(X,1000)
params = c("p","mu1","s1","mu2","s2")
cat("Estimates\n")
for(i in 1:5)
{
    cat(params[i]," = ",v[i],"\n")
}
p = v[1]
mu1 = v[2]
s1 = v[3]
mu2 = v[4]
s2 = v[5]
L = sum(log(dmixnorm(X,0.4,0,1,0,5)))
cat("Log loglihood : ",L)
```

Explanation :

$$p^{(t+1)} = \frac{1}{n}\sum_{i=1}^{n} m_i^1$$

$$\mu_j^{(t+1)} = \frac{\boldsymbol{m}_j.\boldsymbol{x}}{\mathbf{1}.\boldsymbol{m}_j} \qquad \forall j = 1,2$$

$$\sigma_j^{(t+1)} = \frac{\boldsymbol{m}_j.\boldsymbol{s}_j^{(t+1)}}{\mathbf{1}.\boldsymbol{m}_j} \qquad \forall j = 1,2$$

where

$$\boldsymbol{s}_{i,j}^{(t+1)} = (x_i - \mu_j^{(t+1)})^2$$

$$m_i = \frac{p^{(t)}N(x_i|\mu_1^{(t)},(\sigma_1^{(t)})^2)}{p^{(t)}N(x_i|\mu_1^{(t)},(\sigma_1^{(t)})^2) + (1-p^{(t)})N(x_i|\mu_2^{(t)},(\sigma_2^{(t)})^2)}$$

```
Estimates
p  =  0.599388
mu1  =  -0.08259709
s1  =  5.052732
mu2  =  -0.08925698
s2  =  0.9425193
```

Problem 2

R Code :

```r
y = rnorm(2000,0,sqrt(5))
beta = 0.5*sum(y^2)+0.5
alpha = (length(y)+5)/2
cat("MAP estimate : ")
cat(beta/(alpha+1),"\n")
cat("Bayesian estimate : ")
cat(beta/(alpha-1))
```

Explanation :

$$y_i \sim N(0,\sigma^2) \qquad \forall i = 1,\ldots,n$$

$$\text{posterior}(\sigma^2) \propto \frac{1}{(2\pi)^{n/2}\sigma^n} e^{-\frac{\sum_{i=1}^{n} y_i^2}{2\sigma^2}} .(\sigma^2)^{-\frac{5}{2}-1} e^{-\frac{1}{2\sigma^2}}$$

$$\text{posterior}(\sigma^2) \propto e^{\frac{-((\sum_{i=1}^{n} y_i^2)+1)}{2\sigma^2}} .(\sigma^2)^{-(\frac{5}{2}+\frac{n}{2})-1}$$

Comparing it with the pdf of Inverse Gamma distribution,

$$f(x;\alpha,\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{-\alpha-1} e^{(\frac{-\beta}{x})}$$

we get,

$$\alpha = \frac{n+5}{2}$$

$$\beta = \frac{1}{2}(\sum_{i=1}^{n} y_i^2 + 1)$$

```
MAP estimate : 5.098528
Bayesian estimate : 5.10871
```