HW5 - Exercise5

JooChul Lee

12 October 2018

Contents

1 Follow the lecture notes to verify the validity of the provided E- and M-steps.

$$
\begin{aligned}
Q(\mathbf{\Psi} \mid \mathbf{\Psi}^{(k)}) &= E[l_n^c(\mathbf{\Psi})|\mathbf{x}, \mathbf{y}, \mathbf{\Psi}^{(k)}] && (1)\\
&= \sum_z P(z|\mathbf{x}, \mathbf{y}, \mathbf{\Psi}^{(k)}) l_n^c(\mathbf{\Psi}) && (2)\\
&= \sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=0}^{1} P(z_{ij} = k|\mathbf{x}, \mathbf{y}, \mathbf{\Psi}^{(k)}) k \log\left\{\pi_j \phi(y_i - \mathbf{x}_i^\top \boldsymbol{\beta}_j; 0, \sigma^2)\right\} && (3)\\
&= \sum_{i=1}^{n}\sum_{j=1}^{m} P(z_{ij} = 1|\mathbf{x}, \mathbf{y}, \mathbf{\Psi}^{(k)}) \log\left\{\pi_j \phi(y_i - \mathbf{x}_i^\top \boldsymbol{\beta}_j; 0, \sigma^2)\right\} && (4)\\
&= \sum_{i=1}^{n}\sum_{j=1}^{m} E[z_{ij}|\mathbf{x}, \mathbf{y}, \mathbf{\Psi}^{(k)}] \log\left\{\pi_j \phi(y_i - \mathbf{x}_i^\top \boldsymbol{\beta}_j; 0, \sigma^2)\right\} && (5)\\
&= \sum_{i=1}^{n}\sum_{j=1}^{m} \frac{P(z_{ij} = 1, \mathbf{x}_i, y_i|\mathbf{\Psi}^{(k)})}{P(\mathbf{x}_i, y_i|\mathbf{\Psi}^{(k)})} \log\left\{\pi_j \phi(y_i - \mathbf{x}_i^\top \boldsymbol{\beta}_j; 0, \sigma^2)\right\} && (6)\\
&= \sum_{i=1}^{n}\sum_{j=1}^{m} \frac{\pi_j^{(k)} \phi(y_i - \mathbf{x}_i^\top \boldsymbol{\beta}_j^{(k)}; 0, \sigma^{2^{(k)}})}{\sum_{j=1}^{m} \pi_j^{(k)} \phi(y_i - \mathbf{x}_i^\top \boldsymbol{\beta}_j^{(k)}; 0, \sigma^{2^{(k)}})} \log\left\{\pi_j \phi(y_i - \mathbf{x}_i^\top \boldsymbol{\beta}_j; 0, \sigma^2)\right\} && (7)\\
&= \sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}^{(k+1)} \left\{\log \pi_j + \log \phi(y_i - \mathbf{x}_i^\top \boldsymbol{\beta}_j; 0, \sigma^2)\right\} && (8)
\end{aligned}
$$

Now, Maximize $Q(\mathbf{\Psi} \mid \mathbf{\Psi}^{(k)})$.

For the $\pi_j^{(k+1)}$ and $j = 1, \ldots, m-1$,

$$
\begin{aligned}
\frac{\partial}{\partial \pi_j} Q(\mathbf{\Psi} \mid \mathbf{\Psi}^{(k)}) &= \frac{\partial}{\partial \pi_j} (\sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij}^{(k+1)} \log \pi_j + \mathbf{C}), (\mathbf{C}\ \textit{is not related to}\ \pi_j) & (9)\\
&= \frac{\partial}{\partial \pi_j} \{\sum_{i=1}^{n} \sum_{j=1}^{m-1} p_{ij}^{(k+1)} \log \pi_j + \sum_{i=1}^{n} p_{im}^{(k+1)} \log(1 - \pi_1 - \ldots - \pi_{m-1})\} & (10)\\
&= \frac{\sum_{i=1}^{n} p_{ij}^{(k+1)}}{\pi_j} - \frac{\sum_{i=1}^{n} p_{im}^{(k+1)}}{\pi_m} & (11)\\
&= 0 & (12)
\end{aligned}
$$

From (11) and (12),

$$
\frac{\sum_{i=1}^{n} p_{ij}^{(k+1)}}{\pi_j} = \frac{\sum_{i=1}^{n} p_{im}^{(k+1)}}{\pi_m} \tag{13}
$$

Then,

$$
\begin{aligned}
\sum_{j=1}^{m} \sum_{i=1}^{n} p_{ij}^{(k+1)} &= \sum_{j=1}^{m} \pi_j \frac{\sum_{i=1}^{n} p_{im}^{(k+1)}}{\pi_m} & (14)\\
&= \frac{\sum_{i=1}^{n} p_{im}^{(k+1)}}{\pi_m} & (15)
\end{aligned}
$$

Also, we know that

$$
\begin{aligned}
\sum_{j=1}^{m} \sum_{i=1}^{n} p_{ij}^{(k+1)} &= \sum_{i=1}^{n} 1 & (16)\\
&= n & (17)
\end{aligned}
$$

Then from (14-17),

$$
\frac{\sum_{i=1}^{n} p_{im}^{(k+1)}}{\pi_m} = n \tag{18}
$$

From (13) and (18),

$$
\frac{\sum_{i=1}^{n} p_{ij}^{(k+1)}}{\pi_j} = n \tag{19}
$$

Thus, from (18) and (19),

$$
\pi_j^{(k+1)} = \frac{\sum_{i=1}^{n} p_{ij}^{(k+1)}}{n}\ \textit{for}\ j = 1, \ldots, m
$$

For the $\beta_j^{(k+1)}$,

$$
\begin{aligned}
\frac{\partial}{\partial \beta_j} Q(\boldsymbol{\Psi} \mid \boldsymbol{\Psi}^{(k)}) &= \frac{\partial}{\partial \beta_j}(-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}^{(k+1)} \frac{(y_i - \mathbf{x}_i' \beta_j)^2}{\sigma^2} + \mathbf{C}), (\mathbf{C}\ \textit{is not related to}\ \beta_j) \tag{20} \\
&= -\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}^{(k+1)} \mathbf{x}_i \frac{(y_i - \mathbf{x}_i' \beta_j)}{\sigma^2} \tag{21} \\
&= 0 \tag{22}
\end{aligned}
$$

From (21) and (22),

$$
\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}^{(k+1)} \mathbf{x}_i (y_i - \mathbf{x}_i' \beta_j) = 0 \tag{23}
$$

Thus,for $j = 1, \ldots, m$

$$
\beta_j^{(k+1)} = \left(\sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i' p_{ij}^{(k+1)} \right)^{-1} \left(\sum_{i=1}^{n} \mathbf{x}_i p_{ij}^{(k+1)} y_i \right)
$$

For the $\sigma^{2^{(k+1)}}$,

$$
\begin{aligned}
\frac{\partial}{\partial \sigma^2} Q(\boldsymbol{\Psi} \mid \boldsymbol{\Psi}^{(k)}) &= \frac{\partial}{\partial \sigma^2}(-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}^{(k+1)} log\sigma^2 - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}^{(k+1)} \frac{(y_i - \mathbf{x}_i' \beta_j^{(k+1)})^2}{\sigma^2} + \mathbf{C}), (\mathbf{C}\ \textit{is not related to}\ \sigma^2) \tag{24} \\
&= -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}^{(k+1)} \frac{1}{\sigma^2} + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}^{(k+1)} \frac{(y_i - \mathbf{x}_i' \beta_j^{(k+1)})^2}{\sigma^4} \tag{25} \\
&= 0 \tag{26}
\end{aligned}
$$

From (25) and (26),

$$
\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}^{(k+1)} \frac{1}{\sigma^2} = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}^{(k+1)} \frac{(y_i - \mathbf{x}_i' \beta_j^{(k+1)})^2}{\sigma^4} \tag{27}
$$

From (27), since $\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}^{(k+1)} = n$,

$$
\sigma^{2^{(k+1)}} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}^{(k+1)} (y_i - \mathbf{x}_i' \beta_j^{(k+1)})^2}{n}
$$

2 Algorithm in R with a function regmix_em

```r
regmix_em = function(y, xmat, pi.init , beta.init, sigma.init,
                     control = list(maxit = 500, tol = 1e-5))
{
   current.pi = pi.init; new.beta<- current.beta <- beta.init; current.sigma = sigma.init
   xmat = as.matrix(xmat)
   p = matrix(0,length(y),length(pi.init))
   for(k in 1:control$maxit)
   {
      for(i in 1:length(pi.init))
         p[,i] = current.pi[i] * dnorm(y - xmat %*% current.beta[,i],0,current.sigma)
      new.p =  p/rowSums(p)
      new.pi = colMeans(new.p)
      for(j in 1:length(pi.init))
         new.beta[,j] = solve(t(xmat * new.p[,j]) %*% xmat ) %*% t(xmat * new.p[,j] ) %*% y
      new.sigma = sqrt(sum(new.p * (y  %*% t(rep(1, length(pi.init)))
                                    - xmat %*% new.beta)^2)/length(y))
      if( sqrt(sum((current.pi - new.pi)^2) + sum((current.beta - new.beta)^2) +
         sum((current.sigma - new.sigma)^2)) < control$tol ) break
      current.pi = new.pi; current.beta = new.beta; current.sigma = new.sigma
   }
   return(list(pi = current.pi, beta = current.beta, sigma = current.sigma, iter = k))
}
```

#Generate data with the following and estimate the parameters

```
regmix_sim <- function(n, pi, beta, sigma) {
    K <- ncol(beta)
    p <- NROW(beta)
    xmat <- matrix(rnorm(n * p), n, p) # normal covaraites
    error <- matrix(rnorm(n * K, sd = sigma), n, K)
    ymat <- xmat %*% beta + error # n by K matrix
    ind <- t(rmultinom(n, size = 1, prob = pi))
    y <- rowSums(ymat * ind)
    data.frame(y, xmat)
}
n <- 400
pi <- c(.3, .4, .3)
bet <- matrix(c( 1,  1,  1,
                 -1, -1, -1), 2, 3)
sig <- 1
set.seed(1205)
dat <- regmix_sim(n, pi, bet, sig)
regmix_em(y = dat[,1], xmat = dat[,-1], pi.init = pi / pi / length(pi),
          beta.init = bet * 0, sigma.init = sig/sig, control = list(maxit = 500, tol = 1e-5))
```

```
## $pi
## [1] 0.3333333 0.3333333 0.3333333
##
## $beta
##            [,1]       [,2]       [,3]
## [1,]  0.3335660  0.3335660  0.3335660
## [2,] -0.4754645 -0.4754645 -0.4754645
##
## $sigma
## [1] 1.732492
##
## $iter
## [1] 2
```

When the given initial values from the question is used, the estimated betas are the same. And the estimated πs are same and the estimated sigma is 1.73249. The number of iteration is 2.

```r
regmix_em(y = dat[,1], xmat = dat[,-1], pi.init = pi / pi / length(pi),
          beta.init = matrix(-2:3,2,3), sigma.init = sig/sig,
          control = list(maxit = 500, tol = 1e-5))
```

```
## $pi
## [1] 0.3454126 0.3858304 0.2687570
##
## $beta
##            [,1]      [,2]       [,3]
## [1,] -0.9136624 0.8796664  0.9912273
## [2,] -1.1990377 0.9341813 -1.2424797
##
## $sigma
## [1] 1.023598
##
## $iter
## [1] 63
```

I also tried it with the different initial values for betas. So, the estimated values are diffrent from the above. The number of the iteration is 63.