TD MAD - EM Algorithm

Piseth KHENG, Borachhun YOU

24 October 2022

Simulation

1. Simulate a sample of $n = 100$ observations of Poisson distribution with parameter $\lambda = 3$

```r
poisson3 <- rpois(n=100, lambda=3)
```

2. Simulate a sample of $n = 200$ observations of Poisson distribution with parameter $\lambda = 15$

```r
poisson15 <- rpois(n=200, lambda=15)
```

3. Create a vector of 100 "1" followed by 200 "2"

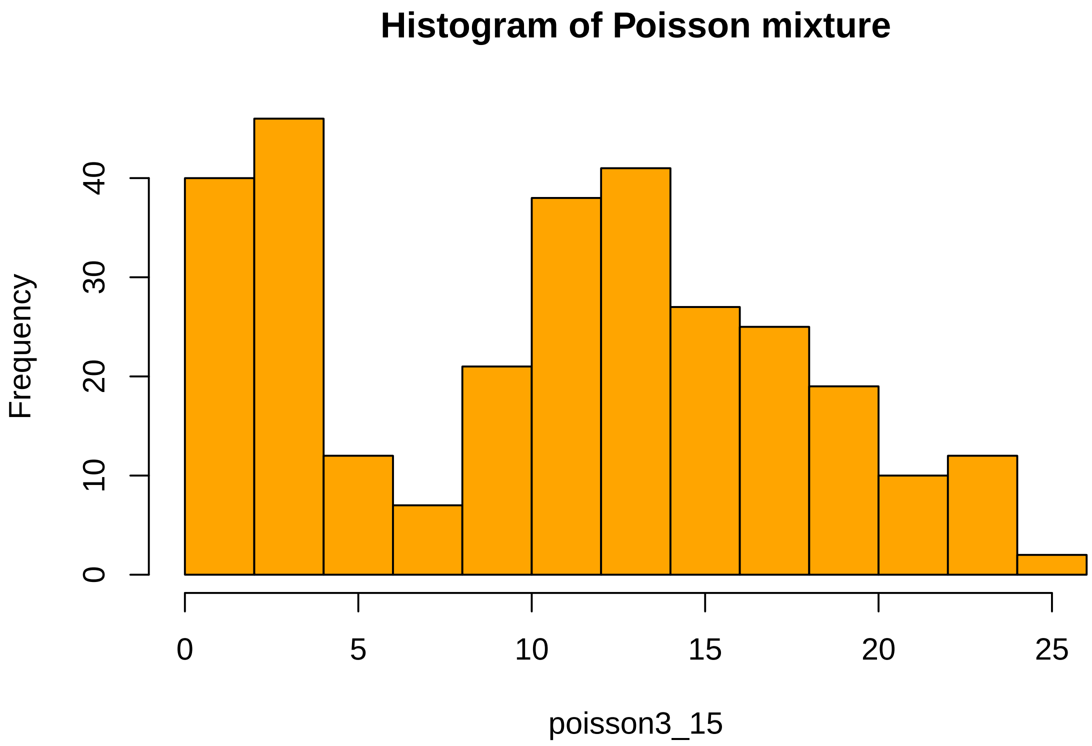
```r
true_class <- rep(c(1,2), c(100,200))
```

4. Simulate a mixture of two Poisson distribution

$$\mathbb{P}(x) = \pi_1 \frac{e^{-\lambda_1}}{x!} \lambda_1^x + \pi_2 \frac{e^{-\lambda_2}}{x!} \lambda_2^x$$

with $\lambda_1 = 3$ and $\lambda_2 = 15$, $\pi_1 = \frac{1}{3}$.

```r
poisson3_15 <- c(poisson3, poisson15)
hist(poisson3_15, col="orange", main="Histogram of Poisson mixture")
```



EM algorithm for a Poisson mixture model of K components

Poisson distribution of K components:

$$\mathbb{P}(x) = \sum_{k=1}^{K} \pi_k \mathbb{P}_k(x) = \sum_{k=1}^{K} \pi_k \frac{e^{-\lambda_k}}{x!} \lambda_k^x$$

We have the observations $X = (x_1, ..., x_n)$, and the missing data $Z = (z_1, ..., z_n)$.

The objective is to find the parameters:

$$\theta = \{\pi_1, ..., \pi_K, \lambda_1, ..., \lambda_K\}$$

1. Initialization

We take $\pi_k = \frac{1}{K}$ and $\lambda_k = k, \forall k \in [\![1, K]\!]$

```
EM_init_poisson <- function(K) {
  pi <- rep(1/K, times=K)
  lambda <- 1:K
  return(c(pi, lambda))
}
```

2. E step

At iteration q, E step computes $Q(\theta|\theta^{(q)})$

$$
\begin{aligned}
Q(\theta|\theta^{(q)}) &= \mathbb{E}_{Z|X;\theta^{(q)}}\left[\log \mathbb{P}(X, Z; \theta)\right] = \mathbb{E}_{Z|X;\theta^{(q)}}\left[\log \prod_{i=1}^{n} \mathbb{P}(x_i, z_i; \theta)\right] \\
&= \sum_{i=1}^{n} \mathbb{E}_{z_i|x_i;\theta^{(q)}}\left[\log\left(\mathbb{P}(z_i) \times \mathbb{P}(x_i|z_i;\theta)\right)\right]
\end{aligned}
$$

Let $z_{ik} = 1_{\{z_i=k\}}$

$$
\begin{aligned}
Q(\theta|\theta^{(q)}) &= \sum_{i=1}^{n} \mathbb{E}_{z_i|x_i;\theta^{(q)}}\left[\sum_{k=1}^{K} z_{ik} \log\left(\mathbb{P}(z_i = k) \times \mathbb{P}(x_i|z_i = k;\theta)\right)\right] \\
&= \sum_{i=1}^{n} \sum_{k=1}^{K} \mathbb{P}(z_i = k|x_i;\theta^{(q)}) \log\left(\mathbb{P}(z_i = k) \times \mathbb{P}(x_i|z_i = k;\theta)\right) \\
&= \sum_{i=1}^{n} \sum_{k=1}^{K} \mathbb{P}(z_i = k|x_i;\theta^{(q)}) \log\left(\pi_k \times \mathbb{P}_k(x_i;\lambda_k)\right)
\end{aligned}
$$

Let

$$
t_{ik}^{(q)} = \mathbb{P}(z_i = k|x_i;\theta^{(q)}) = \frac{\pi_k^{(q)} \mathbb{P}_k(x_i;\lambda_k^{(q)})}{\sum_{m=1}^{K} \pi_m^{(q)} \mathbb{P}_m(x_i;\lambda_m^{(q)})}
$$

At each iteration q, we only need to compute $t_{ik}^{(q)}, i \in [\![1, n]\!], k \in [\![1, K]\!]$

```
# theta : c(pi_k, lambda_k)
E_step_poisson <- function(X, K, theta) {
  pi <- theta[1:K]
  lambda <- theta[(K+1):(2*K)]
  t <- matrix(0, nrow=length(X), ncol=K)
  for (i in 1:length(X)) {
    for (k in 1:K) {
      t[i,k] <- (pi[k] * dpois(X[i],lambda[k])) / sum(pi * dpois(X[i],lambda))
    }
  }
  return(t)
}
```

3. M step

At iteration q, M step computes

$$
\theta^{(q+1)} = \arg\max_{\theta} Q(\theta|\theta^{(q)})
$$

$$
\begin{aligned}
\frac{\partial}{\partial \lambda_1} Q(\theta|\theta^{(q)}) &= \frac{\partial}{\partial \lambda_1} \sum_{i=1}^{n} \sum_{k=1}^{K} t_{ik}^{(q)} \log\left(\pi_k \times \frac{e^{-\lambda_k}}{x_i!} \lambda_k^{x_i}\right) \\
&= \frac{\partial}{\partial \lambda_1} \sum_{i=1}^{n} t_{i1}^{(q)} \left(\log \pi_1 - \lambda_1 - \log(x_i!) + x_i \log \lambda_1\right) \\
&= \sum_{i=1}^{n} t_{i1}^{(q)} \left(-1 + \frac{x_i}{\lambda_1}\right) = 0
\end{aligned}
$$

$$\lambda_1^{(q+1)} = \frac{\sum_{i=1}^n t_{i1}^{(q)} x_i}{\sum_{i=1}^n t_{i1}^{(q)}}$$

We calculate the derivative the same way for the other λ $(\lambda_2, ...\lambda_K)$. We get:

$$\lambda_k^{(q+1)} = \frac{\sum_{i=1}^n t_{ik}^{(q)} x_i}{\sum_{i=1}^n t_{ik}^{(q)}},\ k \in [\![1, K]\!]$$

Since there is a constraint $\sum_{k=1}^K \pi_k = 1$, we consider Lagrangian:

$$\mathcal{L}(\theta, \alpha) = Q(\theta|\theta^{(q)}) + \alpha \left(\sum_{k=1}^K \pi_k - 1 \right)$$

$$\frac{\partial}{\partial \pi_1} \mathcal{L}(\theta, \alpha) = \frac{\partial}{\partial \pi_1} \sum_{i=1}^n t_{i1}^{(q)} \log \left(\pi_1 \times \frac{e^{-\lambda_1}}{x_i!} \lambda_1^{x_i} \right) + \alpha = \sum_{i=1}^n \frac{t_{i1}^{(q)}}{\pi_1} + \alpha = 0$$

$$\pi_1 = -\frac{\sum_{i=1}^n t_{i1}^{(q)}}{\alpha}$$

We calculate the derivative the same way for the other π $(\pi_2, ...\pi_K)$. We get:

$$\pi_k = -\frac{\sum_{i=1}^n t_{ik}^{(q)}}{\alpha},\ k \in [\![1, K]\!]$$

$$\frac{\partial}{\partial \alpha} \mathcal{L}(\theta, \alpha) = \sum_{k=1}^K \pi_k - 1 = 0$$

$$\alpha = -\sum_{i=1}^n \sum_{k=1}^K t_{ik}^{(q)}$$

$$\pi_k^{(q+1)} = \frac{\sum_{i=1}^n t_{ik}^{(q)}}{\sum_{i=1}^n \sum_{k=1}^K t_{ik}^{(q)}} = \frac{\sum_{i=1}^n t_{ik}^{(q)}}{n},\ k \in [\![1, K]\!]$$

```r
# t : result of E step
M_step_poisson <- function(X, K, t) {
  lambda <- sapply(1:K, function(k) sum(t[,k]*X) / sum(t[,k]))
  pi <- sapply(1:K, function(k) sum(t[,k]) / length(X))
  return(c(pi, lambda))
}
```

4. Testing the programmed EM algorithm on the simulated data:

We take the stopping condition of the algorithm:

$$\frac{||\theta^{(q)} - \theta^{(q+1)}||_2^2}{||\theta^{(q)}||_2^2} < \epsilon$$

where $\epsilon = 10^{-6}$

```r
X <- poisson3_15
K <- 2

# initialization
theta <- EM_init_poisson(K)

repeat {

  # E step
  t <- E_step_poisson(X, K, theta)

  # M step
  new_theta <- M_step_poisson(X, K, t)

  # stopping condition
  if (sum((theta - new_theta)^2) / sum(theta^2) < 1e-6) {
    break
  }

  theta <- new_theta
}

print(paste(c("pi_1 =", "pi_2 =", "lambda_1 =", "lambda_2 ="), new_theta))
```

```
## [1] "pi_1 = 0.332209467133731"    "pi_2 = 0.667790532866269"
## [3] "lambda_1 = 2.97165388250249" "lambda_2 = 14.9939059546283"
```