EM ALGORITHM FOR A MIXTURE OF POISSONS

Throughout we will adopt the following notations:

(1) X will be a vector (of integers) of size N, representing the dataset,
(2) K will be the number of components,
(3) $\underline{\pi} = (\pi_1, ..., \pi_n)$ will be a vector of numbers in $[0, 1]$ that sum to 1,
(4) $\underline{\lambda} = (\lambda_1, ..., \lambda_n)$ will be a vector of positive numbers. The l-th component will be the parameter of the l-th Poisson distribution, and
(5) $p : \mathbb{R}^K_{>0} \to \mathbb{R}^{NK}$ will be a function that sends a vector $\underline{\lambda} = (\lambda_1, ..., \lambda_k)$ to the matrix of size $N \times K$ that at the component (i, j) is

$$p(\underline{\lambda})[i, l] := \frac{\lambda_l^{X[i]} e^{-\lambda_l}}{(X[i])!}.$$

Namely, $p(\underline{\lambda})$ is a matrix whose entry in position i, l is the Poisson distribution with parameter λ_l evaluated at the datapoint $X[i]$.

1. Expected log likelyhood

This is $\sum_{i=1}^N \frac{\log(P(X[i]))}{N}$. Namely, if the parameters are $\underline{\lambda}$ and $\underline{\pi}$:

$$\sum_{i=1}^N \frac{\log(P(X[i]))}{N} = \sum_{i=1}^N \frac{\log(\sum_{l=1}^K p(\underline{\lambda})[i, l] \cdot \underline{\pi}[l])}{N}.$$

Again observe that the argument of logarithm is the i-th component of the matrix multiplication $p(\underline{\lambda}) \cdot \underline{\pi}$.

Safety check: recall that the expected log-likelyhood increases at each step of the EM algorithm.

2. E-Step

We compute, for every i, l, the probability $P(Z_l|X[i])$ assuming the parameters $\underline{\lambda} = (\lambda_1, ..., \lambda_K)$ and the probabilities of the components $\underline{\pi} = (\pi_1, ..., \pi_K)$.

So we define q to be a matrix of shape $N \times K$ whose entry (i, l) is

$$q[i, l] = P(Z_l|X[i]) = \frac{p(\underline{\lambda})[i, l]\pi_l}{\sum_j p(\underline{\lambda})[i, j]\pi_j}$$

3. M-Step for $\underline{\pi}$

This does not depend on the distribution we are using, it just depends on the matrix q of the E-step. After some algebra, the updated vector $\pi_{\text{new}} = (\pi_{\text{new},1}, ..., \pi_{\text{new},K})$ has l-th component

$$\pi_{\text{new}}[l] = \frac{\sum_{i=1}^N q[i, l]}{\sum_{i,j} q[i, j]}$$

4. M-Step for $\underline{\lambda}$

As before, q is the matrix of the E-step. Then after some algebra:

$$\lambda_{\text{new}}[l] = \frac{\sum_{i=1}^{N} q[i,l] X[i]}{\sum_{i=1}^{N} q[i,l]}.$$

The numerator is the l-th component of the matrix multiplication $X \cdot q$.