Pratical work: EM algorithm

Thomas Roiseux

2022-11-21

Simulation

We first generate:

1. A sample of $n = 100$ observations with a Poisson law using $\lambda = 3$.
2. A sample of $n = 200$ observations with a Poisson law using $\lambda = 15$.
3. A vector of 300 coordinates, which the 100 first are 1 and the others are 2.

```
sample3 <- rpois(100, 3)
print(sample3)
```

```
##   [1] 5 1 2 5 1 6 3 2 3 5 0 2 4 4 7 2 4 2 3 1 3 4 5 5 1 3 3 5 7 1 2 1 5 3 3 2 1
##  [38] 1 3 5 1 3 3 4 0 0 2 5 3 3 2 2 3 2 3 1 2 4 5 1 0 1 1 2 5 4 3 2 0 5 2 4 3 2
##  [75] 0 4 3 3 1 3 4 3 2 7 1 2 4 1 6 4 4 7 4 4 2 3 1 7 5 3
```

```
sample15 <- rpois(200, 15)
print(sample15)
```

```
##   [1] 16 15 17 17 16 24 17 19 12 13 13 13 13 16 14 18 13 15 14 10 16 12 17 12 17
##  [26] 18 12  9 19 15 16 12 21 19 21 19 16 15 14 17 14 17  8 17 14 14 13 12 15 11
##  [51]  9 12 14 14 15 15 13 13 15 19 23 15 16 18 21 10 12 14 18 15 11 20 17 20 14
##  [76] 12 16 15 14 16 20 12 23 15 12 11 19  9 23 19 15 16 17 12  7 17 16 13 11 15
## [101] 15 20 16 18 22 22 21 21  9 20 15 17 22 11 17 13 14 13 18 18  9 15 14 22 18
## [126] 16 10 11 23 18 20 17 20 12 17 14 19  9 14 18 11 15 17 16 15  9 13 13 20  8
## [151] 15 19 17 12 13 18 18  9  8 17 20 15  9 17 20 21 18 18 11 18 13 18 15 18 11
## [176] 14 10 17 27 12  9 13 15 17 10 17 21 19 16 13 15 17 15 16 17 10 17 18 12  7
```

```
v <- c()
for (i in 1:100)
{
  v <- c(v, 1, recursive = TRUE)
}
for (i in 1:200)
{
  v <- c(v, 2, recursive = TRUE)
}
```

Now, we are going to generate a Poisson law using two components:

```
#Settings constants
pi1 <- 0.4
pi2 <- 0.6
lambda1 <- 3
```

```
lambda2 <- 15
sample1 <- rpois(300, lambda1)
sample2 <- rpois(300,lambda2)

mixed_sample = pi1 * sample1 + pi2 * sample2
print(mixed_sample)
```

```
##   [1] 12.8  8.8 12.8  7.2 16.6 10.6  7.8 10.4 11.2  6.8 10.0 12.0  7.8  9.6 10.2
##  [16]  8.8 13.8  9.2 10.0 13.4 11.2  9.4 10.4 12.2 12.6 13.0 13.0 11.8  7.8 15.8
##  [31] 12.2 12.4 11.8 10.2  9.4 10.0  9.0 10.4  8.4 12.2  9.8  9.8 13.6 11.0  7.6
##  [46] 10.4  6.4 10.8 10.0  8.8 11.2  9.4 13.0  7.0  7.2  9.0  8.8  8.8  5.6  7.0
##  [61]  9.2  9.2  9.6 11.6 12.4 13.6 11.0 11.8  8.2 11.6  9.4 14.2  8.2 10.0  5.4
##  [76] 12.0  7.2  8.8  9.8  9.2 14.4 11.2 13.6 12.0  9.6  6.4  8.4  9.2  9.4 11.4
##  [91]  9.2  9.8 14.0  8.2  5.8 11.4 10.0  9.0 12.8 14.6 13.8  7.0  8.4 10.0 15.0
## [106]  9.0 10.8 13.0  7.0 11.4 12.8  5.8 11.6  8.4 10.4 12.0 11.4  8.0 17.4  7.8
## [121] 17.2 12.8  8.4  8.0 13.0  9.4 14.8  6.8  8.2  9.8  8.2  8.0  8.6  4.8 10.2
## [136] 15.0 11.2 10.2 10.6 11.4  6.4 12.4 11.0 12.8  9.0  8.4 11.6 14.4 11.8  9.6
## [151] 11.6  8.8 11.0 10.6 11.2 17.6 10.0 14.2 12.0  6.6  8.8 12.0  7.8 12.0 12.0
## [166] 15.2  9.6 12.8 10.0  9.6 10.6 10.0 12.0  9.2  8.8  8.0 11.6  8.2 15.0 14.4
## [181] 11.4 12.2 12.0 12.8 11.6 10.0  8.8 10.6 13.8  9.4 12.2  7.4  8.0 10.4 14.0
## [196]  7.2 13.6 11.0  7.8  7.4  7.8  4.4  8.4  7.4 12.4 13.8  8.2 12.2 16.2  9.0
## [211] 11.8  8.0 10.4  7.2 10.8 13.6  9.2  8.8 12.0  8.2 11.0  6.0  7.6  9.4  8.2
## [226]  8.6  6.2 10.0  8.4 10.0  9.0  9.8 10.2 12.6  7.2 10.0 12.4  8.0 11.4  9.4
## [241] 13.6 13.8  9.6 17.0 12.0 10.2  9.8  8.2  6.4 13.2 13.6 11.0  9.4  7.6 12.4
## [256] 12.8 10.8 14.0 11.4 13.6  9.2  7.0  8.0  7.8  8.6  9.4  6.4  9.6  5.0  9.0
## [271] 15.2  8.6 10.6 12.6  9.8 16.6 10.0 10.4  8.6  7.4  9.4  6.2 13.0  9.4  8.4
## [286]  9.4 13.0 10.6  8.0 10.4 14.2 11.2  6.0  6.6  9.6  7.2 11.6 11.8 10.2 15.0
```

EM-algorithm

For a K-components Poisson law, the density function is:

$$P(x) = \sum_{i=1}^{k} \pi_i \frac{\mathrm{e}^{-\lambda_i}}{x!} \lambda_i^x$$

We first want the initialization of the EM-algorithm:

```
em_init = function(K){

}
```