Some exercices on R

Sophie DONNET

1 EM algorithm

After the presentation given on mixture models and EM algorithm, give a try at the package mixtools on the dataset pikeraw.

```r
library('mixdist')
library('mixtools')
data('pikeraw',package='mixdist')
y = pikeraw[,1]
```

1. Propose several initialisations for the parameters. and run the algorithm for each iteration

```r
y = pikeraw[,1];
g = pikeraw[,2]; m = v = p = rep(0,5);
for (k in 1:5){
  m[k] = mean(y[g==k]);
  v[k] = var(y[g==k]);
  p[k] = sum(g==k)/length(y)}
EM1 = normalmixEM(y,lambda = p,mu = m,sigma = sqrt(v),maxit = 10000)
EM2 = normalmixEM(y,lambda = rep(1/5,5),mu = c(25,35,45,55,55),sigma =sqrt(rep(3,5)),maxit = 10000)
```

Try with a random initialization.

2. For each output, plot the likelihood along the iterations `EM1$all.loglik`. Is it always increasing? Why? Is the limit always the same?

3. Compare the different parameter estimations?

4. Having a look at `EM1$posterior`, how would you estimate the class indices $Z_i$? Compare them with the true groups given in $g$. Comment.
2 Fixed point theorem and application

Finding the root of a function $g$ e.g. the point where $x = g(x)$ is a classical problem in mathematics. Sometimes, it is possible to reformulate the problem as finding $x$ such that $f(x) = x$. The solution to this problem is called fixed point.

Given a function $f$ defined on the real numbers with real values and given a point $x_0$, $x_0$ in the domain of $f$, the fixed point iteration is

$$x_{n+1} = f(x_n), \ n = 0, 1, 2, \ldots$$

which gives rise to the sequence $x_0, x_1, x_2, \ldots$ which is hoped to converge to a point $x$. If $f$ defined on the real line with real values is Lipschitz continuous with Lipschitz constant $L < 1$, then this function has precisely one fixed point, and the fixed-point iteration converges towards that fixed point for any initial guess $x_0$.

The algorithm corresponding to this method is written as:

1. Choose a starting value $x_0$;
2. Calculate $x_n = f(x_{n-1})$ for $n = 1, 2, \ldots$.
3. Repeat step 2 until $|x_n - x_{n-1}| < \epsilon$ or $\frac{|x_n - x_{n-1}|}{|x_{n-1}|} < \epsilon$.

Exercice

We want to find the solution of

$$\frac{1 - (1 + i)^{-10}}{i} = 8.21$$

1. Write a first algorithm in R using while. Write a function containing this algorithm with $x_0$ as input argument.
2. Look for help on the web about repeat. Write a second function with this second program using repeat.
3. Look at the help for the uniroot function in the R help. Propose a third method to solve the equation using this function. Write a third corresponding function.
4. We now want to compare the 3 methods in term of computational time. To do that, we will use the microbenchmark package.

```R
> install.packages("microbenchmark")
> library(microbenchmark)
> x0 <- 0.5;
> res <- microbenchmark(func.one(x0), func.two(x0),func.three(x0),times=1000)
> autoplot(res)
```