Title

Course

Christian Oppegård Moen

DD MM YYYY

Contents

```r
source("./additionalFiles/probAhelp.R")
source("./additionalFiles/probAdata.R")
figPath = "./Figures/"
```

2.

In this problem we want to implement the EM-algorithm. We have found the conditional expectation $Q(\lambda_0, \lambda_1) = Q(\lambda_0, \lambda_1 | \lambda_0^{(t)}, \lambda_1^{(t)})$. This corresponds to the E-step in the EM algorithm. In the M-step of the algorithm is to determine

$$(\lambda_0^{(t+1)}, \lambda_1^{(t+1)}) = \text{argmax} \;\; Q(\lambda_0, \lambda_1).$$

This can be found be finding the partial derivates and $Q(\lambda_0, \lambda_1)$ and set them equal to zero.

$$\frac{\partial}{\partial \lambda_0} Q(\lambda_0, \lambda_1) = \frac{n}{\lambda_0} - \sum_{i=1}^{n} \left(u_i z_i + (1 - u_i) \left(\frac{1}{\lambda_0^{(t)}} - \frac{z_i}{e^{\lambda_0^{(t)} z_i} - 1} \right) \right) = 0$$

$$\frac{\partial}{\partial \lambda_1} Q(\lambda_0, \lambda_1) = \frac{n}{\lambda_1} - \sum_{i=1}^{n} \left((1 - u_i) z_i + u_i \left(\frac{1}{\lambda_1^{(t)}} - \frac{z_i}{e^{\lambda_1^{(t)} z_i} - 1} \right) \right) = 0$$

We solve these two equations for λ_0 and λ_1 respectively. This gives the M-step

$$\lambda_0^{(t+1)} = n / \sum_{i=1}^{n} \left(u_i z_i + (1 - u_i) \left(\frac{1}{\lambda_0^{(t)}} - \frac{z_i}{e^{\lambda_0^{(t)} z_i} - 1} \right) \right)$$

$$\lambda_1^{(t+1)} = n / \sum_{i=1}^{n} \left((1 - u_i) z_i + u_i \left(\frac{1}{\lambda_1^{(t)}} - \frac{z_i}{e^{\lambda_1^{(t)} z_i} - 1} \right) \right)$$

Let $\lambda^{(t)} = (\lambda_0^{(t)}, \lambda_1^{(t)})$. We want to implement the EM-algorithm and we use the convergence criterion

$$d(x^{(t+1)}, x^t) = ||\lambda^{(t+1)} - \lambda^{(t)}||_2 < \epsilon.$$

The function below returns the conditional expectation, that is the E-step of the EM algorithm.

```
cond_expectation <- function(lambda0, lambda1, lambda0t, lambda1t, u, z) {
    n = length(u)
    exp = n * (log(lambda0) + log(lambda1)) - (lambda0 * sum(u * z + (1 - u) * (1/lambda0t -
        (z)/(exp(lambda0t * z) - 1)))) - (lambda1 * sum(u * z + (1 - u) * (1/lambda1t -
        (z)/(exp(lambda1t * z) - 1))))
    return(exp)
}
```

Under is a function that implement M-step.

```
M_step <- function(lam0, lam1, u, z) {
    n = 199
    lambda0next = n/sum(u * z + (1 - u) * (1/lam0 - z/(exp(lam0 * z) - 1)))
    lambda1next = n/sum((1 - u) * z + u * (1/lam1 - z/(exp(lam1 * z) - 1)))

    return(c(lambda0next, lambda1next))
}
```

Under the EM algorithm is implemented.

```
EM_algorithm <- function(lambda, u, z, epsilon = 1e-14) {
    lambda0 = lambda[1]
    lambda1 = lambda[2]
    lambda = c(lambda0, lambda1)
    list0 <- c()
    list1 <- c()
    for (i in 1:300) {
        lambda0t = M_step(lambda0, lambda1, u, z)[1]
        lambda1t = M_step(lambda0, lambda1, u, z)[2]
        lambdat = c(lambda0t, lambda1t)
        list0 <- c(list0, lambda0t)
        list1 <- c(list1, lambda1t)
        norm = norm(lambdat - lambda, type = "2")
        lambda0 = lambda0t
        lambda1 = lambda1t
        lambda = c(lambda0t, lambda1t)
        if (norm < epsilon) {
            break
        }
    }
    return(list(lambdas0 = list0, lambdas1 = list1))
}

# The estimated MLEs of lambda0 and lambda1
lambdas <- EM_algorithm(c(2.5, 5), u, z)
lambdas0 = lambdas$lambdas0
lambdas1 = lambdas$lambdas1
MLE_lambda0 = lambdas0[length(lambdas0)]
MLE_lambda1 = lambdas1[length(lambdas1)]
```

```
MLE_lambda0
```

```
## [1] 3.463089
```

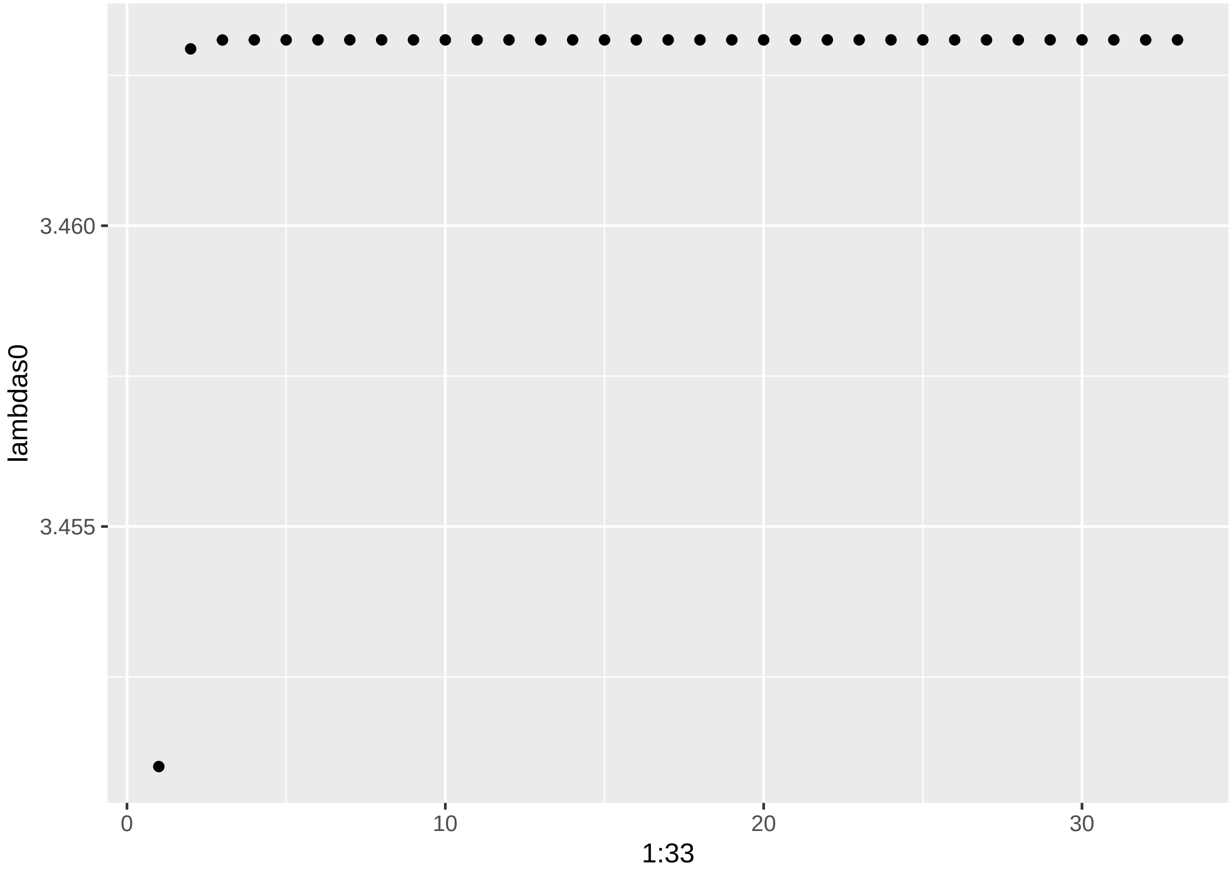
```
MLE_lambda1
```

```
## [1] 9.334877
```

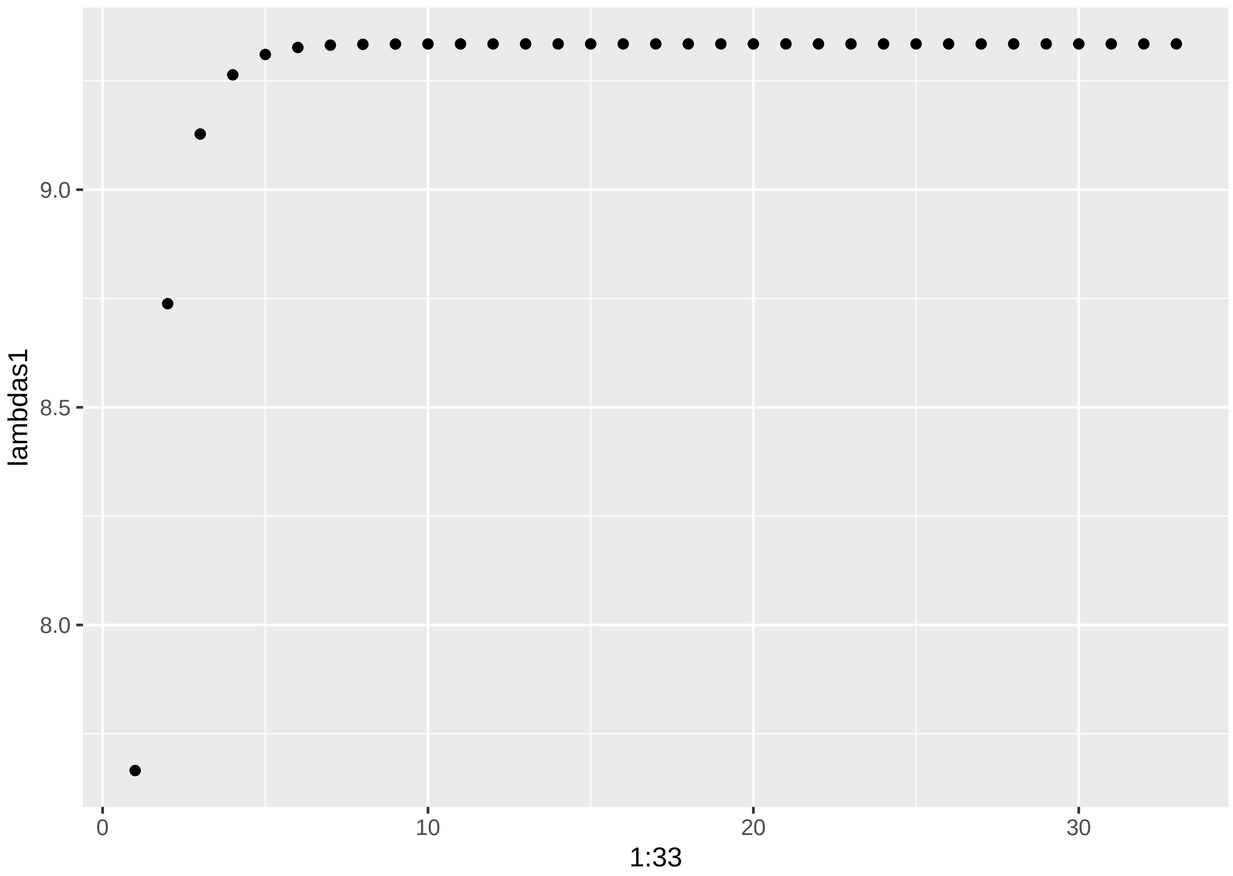
The maaximum likelihood estimates for λ_0 is 3.463089 and 9.3348769 for λ_1.

We also want to visualize the covergence.

```
library(ggplot2)
lambdas = data.frame(lambdas)
ggplot(data = lambdas) + geom_point(mapping = aes(x = 1:33, y = lambdas0))
```



```
ggplot(data = lambdas) + geom_point(mapping = aes(x = 1:33, y = lambdas1))
```



4.

We want to find an analytical formula for $f_{Z_i,U_i}(z_i,u_i|\lambda_0,\lambda_1)$.

$$
\begin{aligned}
f_{Z_i,U_i}(z_i,u_i|\lambda_0,\lambda_1) &= P(\max(X_i,Y_i)=z, I(X_i \geq Y_i)=u_i|\lambda_0,\lambda_1) \\
&= u_i P(\max(X_i,Y_i)=z_i, X)
\end{aligned}
$$

4.

We want to find an analytical formula of $f_{Z_i,U_i}(z_i,u_i|\lambda_0,\lambda_1)$. We start by looking at the case where $u_i=0$, and thus $z_i=y_i$. The cdf is given by

$$
F_{Z_i}(z_i|u_i=0) = P(Y_i \leq z_i|X_i \leq y_i) = \int_0^{z_i}\int_0^{y_i} f_{Y_i}(y_i|\lambda_1) f_{X_i}(x_i|\lambda_0) dx_i dy_i
$$

$$
= \int_0^{z_i}\int_0^{y_i} \lambda_1 \exp(-\lambda_1 y_i)\lambda_0 \exp(-\lambda_0 x_i) = \int_0^{z_i} \lambda_1 \exp(-\lambda_1 y_i)(1-\exp(-\lambda_0 y_i))dy_i
$$

$$
= -\lambda_1 \cdot \frac{\exp(-\lambda_1 z_i - \lambda_0 z_i) - 1}{-\lambda_1 - \lambda_0} - \exp(-\lambda_1 z_i) + 1
$$

$$
\implies f(z_i|u_i=0) = \frac{\mathrm{d}F_{Z_i}(z_i|u_i=0)}{\mathrm{d}z_i} = \exp(-\lambda_1 z_i)\lambda_1(1-\exp(-\lambda_0 z_i))
$$

For $u_i = 1$, we have

$$F_{Z_i}(z_i|u_i=1) = P(X_i \leq z_i, Y_i \leq x_i) = \int_0^{z_i}\int_0^{x_i} f_{X_i}(x_i|\lambda_0) f_{Y_i}(y_i|\lambda_1)\mathrm{dy}_i\mathrm{dx}_i$$
$$= -\lambda_0 \frac{exp(-z_i\lambda_0 - z_i\lambda_1) - 1}{-\lambda_0 - \lambda_1} + 1$$
$$\implies f(z_i|u_i=1) = \frac{\mathrm{d}F_{Z_i}(z_i|u_i=1)}{\mathrm{d}z_i} = \exp(-\lambda_0 z_i)\lambda_0(1-\exp(-\lambda_1 z_i))$$

The likelihood is given by

$$L(\lambda_0, \lambda_1|\mathbf{z},\mathbf{u}) = \prod_{i=0}^{n} f_{Z_i,U_i}(z_i, u_i|\lambda_0,\lambda_1)$$

where

$$f_{Z_i,U_i}(z_i,u_i|\lambda_0,\lambda_1) = \begin{cases} \lambda_1 e^{-\lambda_1 z_i}(1-e^{-\lambda_0 z_i}), & u_i = 0 \\ \lambda_0 e^{-\lambda_0 z_i}(1-e^{-\lambda_1 z_i}), & u_i = 1. \end{cases}$$

The log likelihood is therefore given by

$$\begin{aligned} l(\lambda_0,\lambda_1|\mathbf{z},\mathbf{u}) &= \sum_{i:u_i=0}\left(\ln(\lambda_1) - \lambda_1 z_i + \ln(1-e^{-\lambda_0 z_i})\right) + \sum_{i:u_i=1}\left(\ln(\lambda_0) - \lambda_0 z_i + \ln(1-e^{-\lambda_1 z_i})\right) \\ &= n_0\ln(\lambda_1) + n_1\ln(\lambda_0) + \sum_{i:u_i=0}\left(\ln(1-e^{-\lambda_0 z_i}) - \lambda_1 z_i\right) + \sum_{i:u_i=1}\left(\ln(1-e^{-\lambda_1 z_i}) - \lambda_0 z_i\right), \end{aligned}$$

where $n_0 = \sum_{i=1}^n \mathrm{I}(u_i = 0)$ and $n_1 = \sum_{i=1}^n \mathrm{I}(u_i = 1)$.

To find the maximum likelihood estimators $\hat{\lambda}_0$ and $\hat{\lambda}_1$ we need to solve,

$$\frac{\partial l(\lambda_0,\lambda_1|\mathbf{z},\mathbf{u})}{\partial\lambda_0} = \frac{n_1}{\lambda_0} + \sum_{i:u_i=0}\frac{z_i e^{\lambda_0 z_i}}{e^{\lambda_0 z_i}-1} - \sum_{i:u_i=1} z_i = 0$$

and

$$\frac{\partial l(\lambda_0,\lambda_1|\mathbf{z},\mathbf{u})}{\partial\lambda_1} = \frac{n_0}{\lambda_1} + \sum_{i:u_i=1}\frac{z_i e^{\lambda_0 z_i}}{e^{\lambda_0 z_i}-1} - \sum_{i:u_i=0} z_i = 0.$$

To be sure it is possible to solve the optimization problem, we find the Hessian

$$\nabla^2 l(\lambda_0,\lambda_1|\mathbf{z},\mathbf{u}) = \begin{bmatrix} -\frac{n_1}{\lambda_0^2} - \sum_{i:u_i=0}\frac{z_i^2 e^{\lambda_0 z_i}}{(e^{\lambda_0 z_i}-1)^2} & 0 \\ 0 & -\frac{n_0}{\lambda_1^2} - \sum_{i:u_i=1}\frac{z_i^2 e^{\lambda_1 z_i}}{(e^{\lambda_1 z_i}-1)^2} \end{bmatrix}.$$

Since the Hessian is negative definite, there is only one maxima and it is possible to solve the optimization problem. We have solved the problem numerically with the `optim` function in R.

```
# Find the z-values where u=0 and where u=1

z_0 = z[u == 0]
z_1 = z[u == 1]

l <- function(lambdas) {
    ll <- sum(log(lambdas[2] * exp(-lambdas[2] * z_0) * (1 - exp(-lambdas[1] * z_0)))) +
        sum(log(lambdas[1] * exp(-lambdas[1] * z_1) * (1 - exp(-lambdas[2] * z_1))))
    return(-ll)
}

lambdas = optim(par = c(1, 1), fn = l)$par
```

The maximum likelihood estimator of λ_0 is 3.4627636 and the maximum likelihood estimator of λ_1 is 9.3349621.

The difference from the values obtained for the EM algorithm is very small. An advantage of using this approach compared is that it is less computationally expensive. The EM algorithm can be slow.