Example of joint model for a longitudinal and a multi-state processes

Loïc Ferrer

January 20, 2016

Contents

Import two databases which contain longitudinal and survival data:

```
load("data.RData")
ls()
```

```
## [1] "data_long" "data_surv"
```

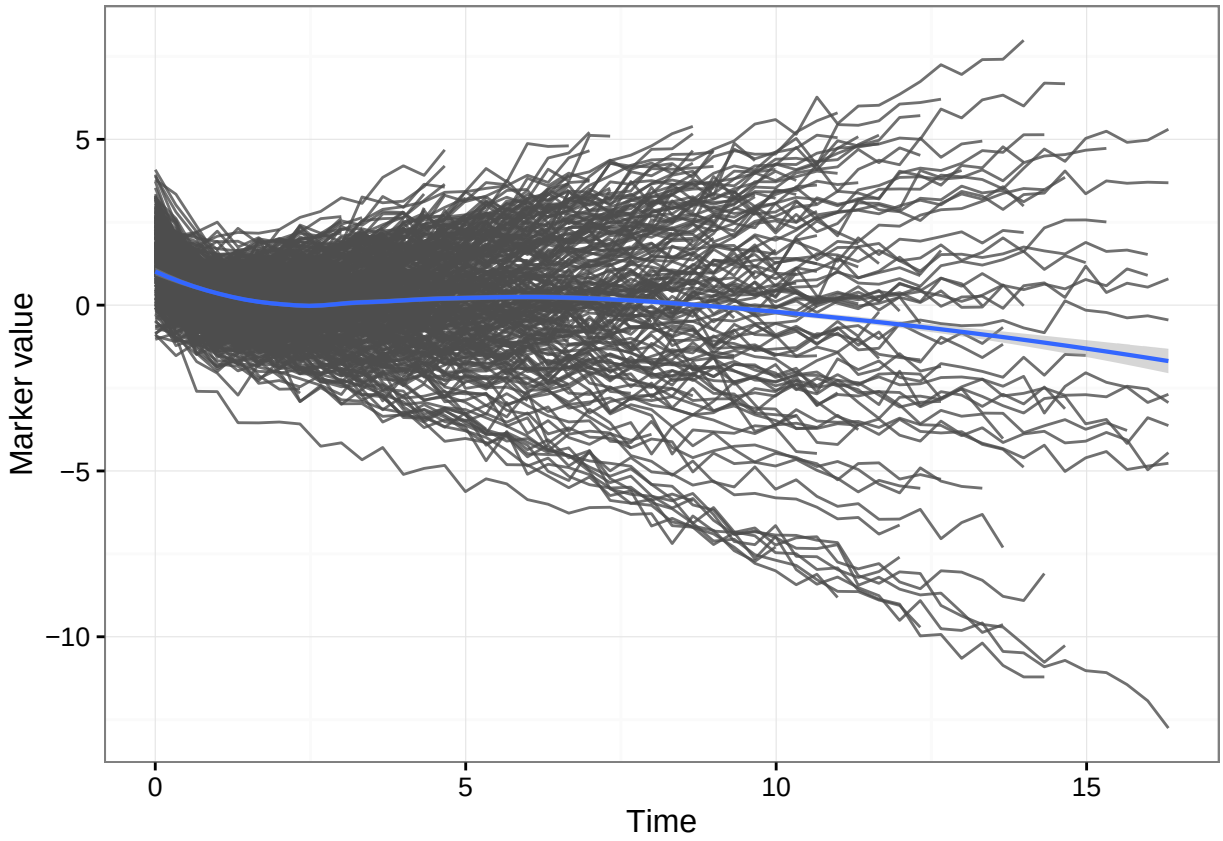
Load the packages and the function to estimate joint multi-state models:

```
library(mstate) # Please use the version 0.2.7
library(JM)
source("JMstateModel.R")
```

Longitudinal sub-part

Plot the individual trajectories of the longitudinal responses:

```
library(ggplot2)
plot_long <- (ggplot(data_long) +
                geom_line(aes(x = times, y = Y, group = id), color = "grey30", alpha = 0.8) +
                stat_smooth(aes(x = times, y = Y), method = "loess", size = 0.75) +
                theme_bw() +
                xlab("Time") +
                ylab("Marker value"))
plot_long
```



Fit the longitudinal responses through a linear mixed model:

```
lmeFit <- lme(fixed = Y ~ (I((1 + times)^(-1.2) - 1) + times) * X,
              data = data_long,
              random = ~ (I((1 + times)^(-1.2) - 1) + times) | id,
              method = "REML",
              control = list(opt = "optim"))
```

Multi-state sub-part

Construct the matrix of transitions:

```
tmat <- matrix(NA, 3, 3)
tmat[1, 2:3] <- 1:2
tmat[2, 3] <- 3
dimnames(tmat) <- list(from = c("State_0", "State_1", "State_2"),
                       to = c("State_0", "State_1", "State_2"))
tmat
```

```
##          to
## from      State_0 State_1 State_2
##   State_0      NA       1       2
##   State_1      NA      NA       3
##   State_2      NA      NA      NA
```

The transition '0 -> 1' is called '1','0 -> 2' is called '2', '1 -> 2' is called '3'.

Define the covariate(s) in the multi-state sub-part:

```
covs <- "X"
```

The *msprep()* function divides the survival database in order to have one line per transition at risk for each subject, with 'Tstart' the entry time in the current state, and 'Tstop' the time of transition or censorship; 'status' denotes if the transition has been performed:

```
data_mstate <- msprep(time = c(NA, "t_State_1", "t_State_2"),
                      status = c(NA, "State_1", "State_2"),
                      data = data_surv,
                      trans = tmat,
                      keep = covs,
                      id = "id")
```

expand.covs() permits to define the covariates for each transition:

```
data_mstate <- expand.covs(data_mstate, covs,
                           append = TRUE, longnames = FALSE)
head(data_mstate)
```

```
## An object of class 'msdata'
##
## Data:
##   id from to trans Tstart     Tstop      time status        X      X.1
## 1  1    1  2     1      0 10.756110 10.756110      0 1.358953 1.358953
## 2  1    1  3     2      0 10.756110 10.756110      1 1.358953 0.000000
## 3  2    1  2     1      0  3.064827  3.064827      0 1.833495 1.833495
## 4  2    1  3     2      0  3.064827  3.064827      1 1.833495 0.000000
## 5  3    1  2     1      0  4.742421  4.742421      0 2.224321 2.224321
## 6  3    1  3     2      0  4.742421  4.742421      0 2.224321 0.000000
##        X.2 X.3
## 1 0.000000   0
## 2 1.358953   0
## 3 0.000000   0
## 4 1.833495   0
## 5 0.000000   0
## 6 2.224321   0
```

The *events()* function indicates the number of observed transitions and their percentages.

```
events(data_mstate)
```

```
## $Frequencies
##          to
## from      State_0 State_1 State_2 no event total entering
##   State_0       0     164     157      179           500
##   State_1       0       0      99       65           164
##   State_2       0       0       0        0             0
##
```

```
## $Proportions
##          to
## from        State_0   State_1   State_2  no event
##   State_0 0.0000000 0.3280000 0.3140000 0.3580000
##   State_1 0.0000000 0.0000000 0.6036585 0.3963415
##   State_2
```

Multi-state model with proportional intensities:

```
coxFit <- coxph(Surv(Tstart, Tstop, status) ~ X.1 + X.2 + X.3 + strata(trans),
                data = data_mstate, method = "breslow", x = TRUE, model = TRUE)
```

Joint multi-state sub-part

Define the derived of the fixed and random parts in the mixed model, and indicate which covariates are kept:

```
dForm <- list(fixed = ~ I((-1.2) * ((1 + times)^(-2.2))) + 1 +
                I((-1.2) * ((1 + times)^(-2.2))):X + X,
              indFixed = c(2:3 ,5:6),
              random = ~ I((-1.2) * ((1 + times)^(-2.2))) + 1,
              indRandom = 2:3)
```

Joint multi-state model with:

- current level and current slope of the biomarker as dependence function,
- cubic B-splines with 1 internal knot to approximate the log-baseline intensities,
- 3 Gauss-Kronrod quadrature points to approximate the integral over time,
- 3 Gauss-Hermite quadrature points in the pseudo-adaptative numerical integration to approximate the integrals over random effects.

```
jointFit_1step_GHk3 <-
  JMstateModel(lmeObject = lmeFit,
               survObject = coxFit,
               timeVar = "times",
               parameterization = "both",
               method = "spline-PH-aGH",
               interFact = list(value = ~ strata(trans) - 1,
                                slope = ~ strata(trans) - 1,
                                data = data_mstate),
               derivForm = dForm,
               Mstate = TRUE,
               data.Mstate = data_mstate,
               ID.Mstate = "id",
               control = list(GHk = 3, lng.in.kn = 1,
                              iter.EM = 150, iter.qN = 600),
               verbose = TRUE)
summary(jointFit_1step_GHk3)
```

```
##
## Call:
```

```
## JMstateModel(lmeObject = lmeFit, survObject = coxFit, timeVar = "times",
##     parameterization = "both", method = "spline-PH-aGH", interFact = list(value = ~strata(trans) -
##         1, slope = ~strata(trans) - 1, data = data_mstate), derivForm = dForm,
##     control = list(GHk = 3, lng.in.kn = 1, iter.EM = 150, iter.qN = 600),
##     Mstate = TRUE, data.Mstate = data_mstate, ID.Mstate = "id",
##     verbose = TRUE)
##
## Data Descriptives:
## Longitudinal Process     Event Process
## Number of Observations: 10065    Number of Events: 420 (84%)
## Number of Groups: 500
##
## Joint Model Summary:
## Longitudinal Process: Linear mixed-effects model
## Event Process: Stratified relative risk model with spline-approximated
##      baseline risk function
## Parameterization: Time-dependent + time-dependent slope
##
##    log.Lik      AIC      BIC
##  -5224.125 10522.25 10678.19
##
## Variance Components:
##                              StdDev    Corr
## (Intercept)                  0.6316  (Intr)  I(+t-1
## I((1 + times)^(-1.2) - 1)  1.3472  0.4666
## times                        0.4062  0.0772  0.5465
## Residual                     0.2750
##
## Coefficients:
## Longitudinal Process
##                                Value Std.Err z-value p-value
## (Intercept)                  -0.2587  0.0901 -2.8707  0.0041
## I((1 + times)^(-1.2) - 1)    0.6244  0.1879  3.3227  0.0009
## times                        -0.1906  0.0348 -5.4783 <0.0001
## X                             0.7928  0.0409 19.3831 <0.0001
## I((1 + times)^(-1.2) - 1):X  1.0632  0.0845 12.5793 <0.0001
## times:X                       0.2593  0.0143 18.0731 <0.0001
##
## Event Process
##                                  Value Std.Err  z-value p-value
## X.1                            -0.2282  0.1460  -1.5627  0.1181
## X.2                             0.0729  0.1426   0.5117  0.6089
## X.3                            -0.3669  0.1639  -2.2378  0.0252
## Assoct:strata(trans)trans=1     0.8848  0.0737  12.0071 <0.0001
## Assoct:strata(trans)trans=2     0.0710  0.0318   2.2298  0.0258
## Assoct:strata(trans)trans=3     0.0836  0.0376   2.2231  0.0262
## Assoct.s:strata(trans)trans=1  0.2968  0.0682   4.3499 <0.0001
## Assoct.s:strata(trans)trans=2 -0.1031  0.0710  -1.4526  0.1463
## Assoct.s:strata(trans)trans=3  0.0846  0.0844   1.0026  0.3160
## bs1(trans=1)                   -7.2582  0.6989 -10.3847 <0.0001
## bs2(trans=1)                   -2.2420  0.5175  -4.3321 <0.0001
## bs3(trans=1)                   -6.1575  0.9362  -6.5774 <0.0001
## bs4(trans=1)                   -7.8106  1.2381  -6.3085 <0.0001
## bs5(trans=1)                   -3.9605  1.6473  -2.4042  0.0162
```

```
## bs1(trans=2)                  -5.3221  0.6278  -8.4775 <0.0001
## bs2(trans=2)                  -2.8321  0.5019  -5.6428 <0.0001
## bs3(trans=2)                  -3.1661  0.7481  -4.2324 <0.0001
## bs4(trans=2)                  -0.9079  0.7417  -1.2242  0.2209
## bs5(trans=2)                   0.0307  0.7274   0.0421  0.9664
## bs1(trans=3)                  -4.0397  4.5368  -0.8904  0.3732
## bs2(trans=3)                  -2.3723  1.1321  -2.0955  0.0361
## bs3(trans=3)                  -3.3503  1.2177  -2.7513  0.0059
## bs4(trans=3)                   0.0499  1.0322   0.0483  0.9615
## bs5(trans=3)                  -0.2859  1.0023  -0.2852  0.7755
##
## Integration:
## method: (pseudo) adaptive Gauss-Hermite
## quadrature points: 3
##
## Optimization:
## Convergence: 0
```

Same joint multi-state model with:

- 9 Gauss-Hermite quadrature points in the pseudo-adaptative numerical integration to approximate the integrals over random effects.

```
jointFit_1step_GHk9 <-
  JMstateModel(lmeObject = lmeFit,
               survObject = coxFit,
               timeVar = "times",
               parameterization = "both",
               method = "spline-PH-aGH",
               interFact = list(value = ~ strata(trans) - 1,
                                slope = ~ strata(trans) - 1,
                                data = data_mstate),
               derivForm = dForm,
               Mstate = TRUE,
               data.Mstate = data_mstate,
               ID.Mstate = "id",
               control = list(GHk = 9, lng.in.kn = 1,
                              iter.EM = 150, iter.qN = 600),
               verbose = TRUE)
summary(jointFit_1step_GHk9)
```

```
##
## Call:
## JMstateModel(lmeObject = lmeFit, survObject = coxFit, timeVar = "times",
##     parameterization = "both", method = "spline-PH-aGH", interFact = list(value = ~strata(trans) -
##         1, slope = ~strata(trans) - 1, data = data_mstate), derivForm = dForm,
##     control = list(GHk = 9, lng.in.kn = 1, iter.EM = 150, iter.qN = 600),
##     Mstate = TRUE, data.Mstate = data_mstate, ID.Mstate = "id",
##     verbose = TRUE)
##
## Data Descriptives:
## Longitudinal Process     Event Process
## Number of Observations: 10065    Number of Events: 420 (84%)
```

```
## Number of Groups: 500
##
## Joint Model Summary:
## Longitudinal Process: Linear mixed-effects model
## Event Process: Stratified relative risk model with spline-approximated
##      baseline risk function
## Parameterization: Time-dependent + time-dependent slope
##
##    log.Lik      AIC      BIC
##  -5223.885 10521.77 10677.71
##
## Variance Components:
##                              StdDev    Corr
## (Intercept)                  0.6312  (Intr)  I(+t-1
## I((1 + times)^(-1.2) - 1)  1.3472  0.4655
## times                        0.4060  0.0762  0.5461
## Residual                     0.2750
##
## Coefficients:
## Longitudinal Process
##                                Value Std.Err z-value p-value
## (Intercept)                  -0.2624  0.0907 -2.8946  0.0038
## I((1 + times)^(-1.2) - 1)    0.6155  0.1994  3.0867  0.0020
## times                        -0.1954  0.0576 -3.3926  0.0007
## X                             0.7949  0.0412 19.3046 <0.0001
## I((1 + times)^(-1.2) - 1):X  1.0707  0.0909 11.7732 <0.0001
## times:X                       0.2639  0.0258 10.2280 <0.0001
##
## Event Process
##                                  Value Std.Err  z-value p-value
## X.1                            -0.2315  0.1463  -1.5827  0.1135
## X.2                             0.0637  0.1428   0.4461  0.6555
## X.3                            -0.3803  0.1644  -2.3134  0.0207
## Assoct:strata(trans)trans=1     0.8833  0.0737  11.9864 <0.0001
## Assoct:strata(trans)trans=2     0.0709  0.0319   2.2229  0.0262
## Assoct:strata(trans)trans=3     0.0880  0.0378   2.3260  0.0200
## Assoct.s:strata(trans)trans=1   0.2974  0.0683   4.3538 <0.0001
## Assoct.s:strata(trans)trans=2 -0.0965  0.0711  -1.3572  0.1747
## Assoct.s:strata(trans)trans=3   0.0855  0.0847   1.0096  0.3127
## bs1(trans=1)                   -7.2101  0.6942 -10.3860 <0.0001
## bs2(trans=1)                   -2.2828  0.5164  -4.4202 <0.0001
## bs3(trans=1)                   -6.0264  0.9334  -6.4567 <0.0001
## bs4(trans=1)                   -7.9873  1.2349  -6.4679 <0.0001
## bs5(trans=1)                   -3.6570  1.5834  -2.3096  0.0209
## bs1(trans=2)                   -5.3150  0.6277  -8.4672 <0.0001
## bs2(trans=2)                   -2.8392  0.5024  -5.6514 <0.0001
## bs3(trans=2)                   -3.1486  0.7482  -4.2082 <0.0001
## bs4(trans=2)                   -0.9266  0.7422  -1.2484  0.2119
## bs5(trans=2)                    0.0456  0.7268   0.0627  0.9500
## bs1(trans=3)                   -4.3460  4.8066  -0.9042  0.3659
## bs2(trans=3)                   -2.4491  1.1426  -2.1434  0.0321
## bs3(trans=3)                   -3.0999  1.2210  -2.5388  0.0111
## bs4(trans=3)                   -0.2039  1.0330  -0.1974  0.8436
## bs5(trans=3)                   -0.1453  0.9878  -0.1471  0.8830
```

```
##
## Integration:
## method: (pseudo) adaptive Gauss-Hermite
## quadrature points: 9
##
## Optimization:
## Convergence: 0
```

To use the multi-step pseudo-adaptive Gauss-Hermite rule, we have to source two functions inspired by JM:

```
source("modified.log.posterior.b2.R")
source("modified.ranef.jointModel.R")
```

Same joint multi-state model with:

- 9–9 Gauss-Hermite quadrature points in the two-step-adaptative numerical integration to approximate the integrals over random effects. We can choose the posterior mode (true definition) or the posterior mean (faster) random effects of the fitted joint model (defined in 'init') to update the quadrature points. Here the mode is used.

```
jointFit_2step_GHk9_9 <-
  JMstateModel(lmeObject = lmeFit,
               survObject = coxFit,
               timeVar = "times",
               parameterization = "both",
               method = "spline-PH-aGH",
               interFact = list(value = ~ strata(trans) - 1,
                                slope = ~ strata(trans) - 1,
                                data = data_mstate),
               derivForm = dForm,
               Mstate = TRUE,
               data.Mstate = data_mstate,
               ID.Mstate = "id",
               control = list(GHk = 9, lng.in.kn = 1),
               init = jointFit_1step_GHk9,
               init.type.ranef = "mode",
               verbose = TRUE)
summary(jointFit_2step_GHk9_9)
```

```
##
## Call:
## JMstateModel(lmeObject = lmeFit, survObject = coxFit, timeVar = "times",
##     parameterization = "both", method = "spline-PH-aGH", interFact = list(value = ~strata(trans) -
##         1, slope = ~strata(trans) - 1, data = data_mstate), derivForm = dForm,
##     init = jointFit_1step_GHk9, control = list(GHk = 9, lng.in.kn = 1),
##     Mstate = TRUE, data.Mstate = data_mstate, ID.Mstate = "id",
##     init.type.ranef = "mode", verbose = TRUE)
##
## Data Descriptives:
## Longitudinal Process     Event Process
## Number of Observations: 10065    Number of Events: 420 (84%)
## Number of Groups: 500
```

```
##
## Joint Model Summary:
## Longitudinal Process: Linear mixed-effects model
## Event Process: Stratified relative risk model with spline-approximated
##      baseline risk function
## Parameterization: Time-dependent + time-dependent slope
##
##    log.Lik      AIC      BIC
##  -5223.877 10521.75 10677.69
##
## Variance Components:
##                             StdDev    Corr
## (Intercept)                 0.6313  (Intr)  I(+t-1
## I((1 + times)^(-1.2) - 1)  1.3472  0.4657
## times                       0.4060  0.0762  0.5461
## Residual                    0.2750
##
## Coefficients:
## Longitudinal Process
##                               Value Std.Err z-value p-value
## (Intercept)                -0.2638  0.0907 -2.9079  0.0036
## I((1 + times)^(-1.2) - 1)   0.6141  0.2002  3.0672  0.0022
## times                      -0.1953  0.0591 -3.3049  0.0010
## X                           0.7955  0.0412 19.2976 <0.0001
## I((1 + times)^(-1.2) - 1):X  1.0713  0.0917 11.6875 <0.0001
## times:X                     0.2639  0.0271  9.7224 <0.0001
##
## Event Process
##                                  Value Std.Err  z-value p-value
## X.1                            -0.2311  0.1462  -1.5805  0.1140
## X.2                             0.0624  0.1428   0.4373  0.6619
## X.3                            -0.3765  0.1644  -2.2897  0.0220
## Assoct:strata(trans)trans=1     0.8833  0.0737  11.9884 <0.0001
## Assoct:strata(trans)trans=2     0.0709  0.0319   2.2235  0.0262
## Assoct:strata(trans)trans=3     0.0868  0.0378   2.2956  0.0217
## Assoct.s:strata(trans)trans=1  0.2969  0.0683   4.3459 <0.0001
## Assoct.s:strata(trans)trans=2 -0.0967  0.0710  -1.3609  0.1736
## Assoct.s:strata(trans)trans=3  0.0853  0.0847   1.0074  0.3138
## bs1(trans=1)                   -7.2081  0.6941 -10.3856 <0.0001
## bs2(trans=1)                   -2.2826  0.5163  -4.4209 <0.0001
## bs3(trans=1)                   -6.0208  0.9332  -6.4517 <0.0001
## bs4(trans=1)                   -7.9946  1.2349  -6.4737 <0.0001
## bs5(trans=1)                   -3.6463  1.5812  -2.3060  0.0211
## bs1(trans=2)                   -5.3193  0.6288  -8.4597 <0.0001
## bs2(trans=2)                   -2.8353  0.5024  -5.6435 <0.0001
## bs3(trans=2)                   -3.1409  0.7482  -4.1979 <0.0001
## bs4(trans=2)                   -0.9254  0.7425  -1.2464  0.2126
## bs5(trans=2)                    0.0430  0.7276   0.0591  0.9529
## bs1(trans=3)                   -4.3599  4.8156  -0.9054  0.3653
## bs2(trans=3)                   -2.4635  1.1435  -2.1544  0.0312
## bs3(trans=3)                   -3.1004  1.2205  -2.5402  0.0111
## bs4(trans=3)                   -0.2006  1.0310  -0.1946  0.8457
## bs5(trans=3)                   -0.1322  0.9835  -0.1344  0.8931
##
```

```
## Integration:
## method: (pseudo) adaptive Gauss-Hermite
## quadrature points: 9
##
## Optimization:
## Convergence: 0
```