Moving to saemix 3.0

Emmanuelle

05/10/2020

TODO list

- **DONE** kompare files with 2.4 to take into account changes made for the CRAN compiler
- **DONE** error models
  - combined 2 versus combined 1
- **TODO** bugfixes
  - **DONE** simulated annealing
  - **DONE** ypred/ppred (changed in documentation, online help, and definitions in code)
- **TODO** testthat functions
  - SaemixData class
    - **DONE** class structure, reading data
    - validate names
    - covariates
  - SaemixModel class
    - **DONE** class structure, defining model
    - **DONE** validate covariance model
  - SaemixObject class
    - **DONE** class structure
    - computation of BIC, AIC (Maud)
    - summary function
  - plots
  - computations ?
  - display options
    - displayProgress=FALSE by default ? **DONE**
    - verbose=FALSE ? **DONE**
    - or indicate how to suppress all messages ? (maybe not a good idea)
- examples
  - **DONE** continuous response models
  - **DONE** binary data
  - **DONE** categorical data, TTE data
    - **DONE*** add a section example of use with donttest to the online page
  - **DONE** RTTE data, count data (maybe count data=salamander data ⇒ RAPI data)
- documentation
  - error models
  - **DONE** new models from Belhal for responses defined by their likelihood
  - new references
  - new defaults: no plot for conditional distribution
- **TODO** next version
  - interface to mkin ?
  - error messages to solve better ?
    - unable to find starting parameter values
    - add FAQ as to how to diagnose this

        * maybe need a predict function for these models
    - bookdown: https://github.com/saemixdevelopment/saemix_bookdown
        * case-studies for categorical data and RTTE data
            · simulate using R instead of mlx
        * move online or keep pdf/tex version ?
            · if move completely online, need to transform the LaTeX guide into a bookdown
            · **DONE** copy on iame github
- **DONE** dependencies
    - ggplot2 in Depends for 3.0
    - npde 3.2:
        * Suggests for ggplot2
        * @importFrom statements in roxygen documentation for all objects from ggplot2 that are needed (ex: geom_bar,...)

Comparing 2.4 and 3.0

Function list

2.4

aaa_generics.R func_diagnostics.R func_FIM.R main_estep.R main.R SaemixObject.R compute_LL.R func_distcond.R func_plots.R main_initialiseMainAlgo.R SaemixData.R SaemixRes.R func_aux.R func_estimParam.R func_simulations.R main_mstep.R SaemixModel.R zzz.R

3.0

aaa_generics.R func_aux.R func_estimParam.R func_stepwise.R main.R SaemixRes.R backward.R func_compare.R func_FIM.R main_estep.R SaemixData.R stepwise.R compute_LL.R func_distcond_noplot.R func_plots.R main_initialiseMainAlgo.R SaemixModel.R zzz.R forward.R func_distcond.R func_simulations.R main_mstep.R SaemixObject.R

functions in 3.0 not in 2.4

- covariate selection algorithm (Maud)
    - forward.R, backward.R, stepwise.R, func_stepwise.R, func_compare.R
    - func_diagnostics.R (? Maud or something else ?)
- version without plots of func_distcond.R

Function comparison

See **changes2.4to3.0.ods** for side-by-side comparison of functions tracing the changes between 2.4 and 3.0, as well as the changes carried over from 2.4 immediately.

Loading

aaa_generics.R TODO

- check generic definition of read
- check aliases for some internal functions

zzz.R

- check if the current version works or if we need to add date

Classes

SaemixData

- check object validation
  - create testthat functions
- harmonise output messages across the package
  - no messages to be output to stdin by default
- new slot automatic in 3.0
  - normally allows automatic recognition and filling in arguments not given
  - ☐ TODO: testthat function
  - also test validate.names()
- use is(x, "data.frame") instead of testing class(x)=="data.frame" (class may have more than 1 value)

SaemixModel

- main change in 3.0: added discrete response with a modelType argument
  - for joint models, will need some tweak (maybe as a vector of response types ?)
- new function validate_covariance.model
  - ☐ TODO: change name to validate.covariance.model
  - check function (add testthat) and add documentation

SaemixRes

- ☑ check definition of ypred and ppred
  - ypred should be f(theta_MAP) and ppred should be E(f(theta))
  - checked and corrected

SaemixObject

- maybe change name saemix.simul to saemix.simulate
- ☐ TODO: logLik.saemix, AIC.SaemixObject, BIC.SaemixObject
  - Maud made changes in the 2.4
  - but Johannes also made some changes in 3.0
  - ☑ TODO: add Johannes's changes to the 2.4 version **in fact** already in 2.4 but
- options
  - ☐ TODO: see CRAN for the rules (like for 2.4, in compute_LL.R)

Computational functions

compute_LL.R

- check if alias ggq.mlx exists

func_aux.R

- modified combined error model
  - ☐ TODO: add to documentation
- conditional distribution function split into 2 functions according to nature of the model
  - ☐ TODO: add to documentation
- check computation of compute.LLy

func_FIM.R

- ☐ TODO: secure the code for discrete data models (only compute LL by linearisation ? do not use altogether ?)
- ☐ TODO: check cat and replace with message

  - check name of option to print out messages (maybe name of option changed)

Main algorithm

main_initialiseMainAlgo.R

- ☐ TODO: check cat and replace with message
  - check name of option to print out messages (maybe name of option changed)

main_estep.R

- check computation of compute.LLy now completely replacing the computation of U.y
- 4th kernel added (Laplacian kernel)
  - ☐ TODO: add default option (0 iterations)
  - ☐ TODO: add to documentation and recommendations

main_mstep.R

- check computation of sigma and influence of SA (nbiter.saemix changed to nbiter.sa)
  - ☐ TODO: check
- ☐ TODO: add Lucie's changes +++
  - ask Lucie to check afterwards

main.R

- ☐ TODO: check cat and replace with message
  - check name of option to print out messages (maybe name of option changed)

Parameters, Simulations

func_distcond.R

- commented out the plots for the moment
  - ☐ TODO: add option to output the graphs

func_estimParam.R

- renamed the function to predict newdata as in 2.4 (also has an Roxygen documentation now)
  - copied the old file in newCode
  - ☐ TODO: check that the code within is similar

func_simulations.R

- renamed simul.saemix in saemix.simul in 3.0 for consistency (all functions start with saemix and not end with it)
  - ☐ TODO: add to documentation and CHANGELOG

Plots

- ☐ TODO: check cat and replace with message
  - check name of option to print out messages (maybe name of option changed)

New functions by Maud

BIC criterion

- definition and computation included in SaemixObject (see above)
  - ☐ TODO: add to documentation and CHANGELOG

Algorithm

- create a notebook to test the algorithm and integrate it
- add test example
  - ☐ TODO: add to documentation

HMM

☐ TODO: check status with Maud

Some bugs from TODO.txt - check if solved

- createSaemixObject.initial Currently, calling the createSaemixObject.initial function with a model parameter that isn't 'structural' will fail with the following error :

Error in (function (cl, name, valueClass) : assignment of an object of class "NULL" is not valid for @'respar' in an object of class "SaemixRes"; is(value, "numeric") is not TRUE

CHANGE fix.seed to be FALSE by default otherwise this can really mess up a simulation study...

- censored responses (12/05/21, Chris) bug

2. `If the column in the data supplied for name.cens is not named "cens", saemix() throws an error.`