smfsb - Stochastic Modelling for Systems Biology, second edition

Darren Wilkinson

May 28, 2018

1 Overview

The `smfsb` package provides all of the R code associated with the book, Wilkinson (2011). The book should therefore be regarded as the main source of documentation regarding the code. However, there should be sufficient documentation here in order to get started with using the software. In particular, owners of Wilkinson (2006) should find it relatively straightforward to get to grips with this new R package. Note that most of this code is intended primarily to be pedagogic. It is not intended to be especially efficient or robust, and therefore is not recommended for "production environments". Almost all of the code is pure R code, intended to be inspected from the R command line. In order to keep the code short, clean and easily understood, there is almost no argument checking or other boilerplate code. This is not a bug, so please don't report it as such. Much of the code is computationally intensive, and would be speeded up by porting to C. Again, I haven't done this in order to keep the code as simple and easy to understand as possible.

See the web home page for Wilkinson (2011) for further details:
`http://www.staff.ncl.ac.uk/d.j.wilkinson/smfsb/2e/`

2 Installation

It is hoped that by the time the book is in print, the package will be available from CRAN, and it should therefore be possible to install using

```
install.packages("smfsb")
```

from any machine with an internet connection.

The package is being maintained on R-Forge, and so it should always be possible to install the very latest nightly build from the R command prompt with

```
install.packages("smfsb",repos="http://r-forge.r-project.org")
```

Once installed, the package can be loaded ready for use with

```
library(smfsb)
```

3 Accessing documentation

I have tried to ensure that the package and all associated functions and datasets are properly documented with runnable examples. So,

```
help(package="smfsb")
```

will give a brief overview of the package and a complete list of all functions. The list of vignettes associated with the package can be obtained with

```
vignette(package="smfsb")
```

At the time of writing, *this* vignette is the only one available, and can be accessed from the R command line with

```
vignette("smfsb",package="smfsb")
```

Help on functions can be obtained using the usual R mechanisms. For example, help on the function `StepGillespie` can be obtained with

```
?StepGillespie
```

and the associated example can be run with

```
example(StepGillespie)
```

A list of demos associated with the package can be obtained with

```
demo(package="smfsb")
```

A list of data sets associated with the package can be obtained with

```
data(package="smfsb")
```

For example, the small table, `mytable` from the introduction to R in Chapter 4 can by loaded with

```
data(mytable)
```

After running this command, the data frame `mytable` will be accessible, and can be examined by typing

```
mytable
```

at the R command prompt.

4 Simulation of stochastic kinetic models

The main purpose of this package is to provide a collection of tools for building and simulating stochastic kinetic models. This can be illustrated using a simple Lotka–Volterra predator–prey system. First, consider the prey, X_1 and the predator X_2 as a stochastic network as

$$\begin{aligned} X_1 &\longrightarrow 2X_1 \\ X_1 + X_2 &\longrightarrow 2X_2 \\ X_2 &\longrightarrow \emptyset. \end{aligned}$$

The first "reaction" represents predator reproduction, the second predator–prey interaction and the third predator death. We can write this in tabular form as

Pre		Post		Hazard
1	0	2	0	$\theta_1 x_1$
1	1	0	2	$\theta_2 x_1 x_2$
0	1	0	0	$\theta_3 x_2$

This can be encoded in R as a stochastic Petri net (SPN) using

```
# SPN for the Lotka-Volterra system
LV=list()
LV$Pre=matrix(c(1,0,1,1,0,1),ncol=2,byrow=TRUE)
LV$Post=matrix(c(2,0,0,2,0,0),ncol=2,byrow=TRUE)
LV$h=function(x,t,th=c(th1=1,th2=0.005,th3=0.6))
{
 with(as.list(c(x,th)),{
         return(c(th1*x1, th2*x1*x2, th3*x2 ))
        })
}
```

which could be created directly by executing

```
data(spnModels)
```

Functions for simulating from the transition kernel of the Markov process defined by the SPN can be created easily by passing the SPN object into the appropriate constructor. For example, if simulation using the Gillespie algorithm is required, a simulation function can be created with

```
stepLV=StepGillespie(LV)
```

This function can then be used to advance the state of the process. For example, to simulate the state of the process at time 1, given an initial condition of $X_1 = 50$, $X_2 = 100$ at time 0, use

```
stepLV(c(x1=50,x2=100),0,1)
```

Alternatively, to simulate a realisation of the process on a regular time grid over the interval $[0, 100]$ in steps of 0.1 time units, use

```
out = simTs(c(x1=50,x2=100),0,100,0.1,stepLV)
```

This returns an R time series object which can be plotted directly. See the help and runnable example for the function `StepGillespie` for further details.

5 Inference for stochastic kinetic models from time course data

Estimating the parameters of stochastic kinetic models using noisy time course measurements on some aspect of the system state is a very important problem. Wilkinson (2011) takes a Bayesian approach to the problem, using particle MCMC methodology. For this, a key aspect is the use of a particle filter to compute an unbiased estimate of marginal likelihood. This is accomplished using

the function `pfMLLik`. Once a method is available for generating unbiased estimates for the marginal likelihood, this may be embedded into a fairly standard marginal Metropolis–Hastings algorithm for parameter estimation. See the help and runnable example for `pfMLLik` for further details, along with the particle MCMC demo, which can by run using `demo(PMCMC)`.

6 References

Wilkinson, D. J. (2006) *Stochastic Modelling for Systems Biology*, Chapman & Hall/CRC Press.

Wilkinson, D. J. (2011) *Stochastic Modelling for Systems Biology*, second edition, Chapman & Hall/CRC Press.