DIZZYSNEWINFEC: EFFICIENT DETERMINISTIC/STOCHASTIC SIMULATIONS IN R FOR A METAPOPULATION BY USING SIR/SEIR MODELS

TRAN THI CAM GIANG

Abstract. Predicting the potential spread of an infectious disease is still a difficult problem for scientists. It requires much more than simple connecting subpopulations in a metapopulation and takes into account many factors about the pathogen and the affected subpopulation. Therefore, this 'dizzysNewInfec' package allows us to simulate dynamics of an infectious disease through subpopulations by using the SIR/SEIR models and by implementing the direct algorithm of Gillespie in 1977 and the adaptive tau leaping to approximate the trajectory of a continuous-time stochastic process. Consequently, result returned is biological data in time horizon about the disease dynamic, we can perform analysis on this biological data. This vignette presents a few examples of SIR/SEIR applied to biological problems.

1. Introduction

Fundamentally, Kermack-McKendrick gave the first epidemic model to provide a mathematical description of the kinetic transmission of an infectious disease in an unstructured subpopulation. According to this model, today we have known well the SIR/SEIR deterministic epidemic models. This is the two basic models very popularly used by scientists. However, Keeling2008 [**?**] show that all the deterministic models are essentially fixed "clockwork" systems with the same starting conditions, exactly the same trajectory is always observed. It isn't right for dynamics of real pathogens in the real-world. So stochastic models are created and concerned with approximating or mimicking the random or probabilistic element from the deterministic models. Moreover, when the quantities in a system are small enough and extinction is probable to occur, then stochastic effects become critical to take into account. This is reason, in the 'dizzysNewInfec' package, it permits us to obtain the dynamics of the deterministic and the approximate dynamics of the stochastic epidemic models.

Based on the stochastic models, their processes are in Markov process, it means that the future state of the process, conditional on the present state, is independent of the past. In the case, our package focus on simulating dynamics from a continuous-time Markov process for which the transition rates are constants, isn't a function of time. We use the exact algorithm of Gillespie in 1977 and the approximate algorithm described as the "adaptive tau-leaping algorithm". With these two algorithms, each has its private advantages and its private disadvantages. For the exact algorithm, it give us a really exact approach of simulating population-based time-to-event through two step with many iterations of 1) searching the time of next event by an exponentially distributed function and 2) searching the nature of next event. This Gillespie's solution becomes too slow and impractical as any one transition rate grows large. Hence, approximate models are born instead of the Gillespie's solution, they are concerned with larger transition rates and with increasing simulation speed while still maintaining reasonable accuracy. The "adaptive tau-leaping algorithm" known as an approximate method reduces the number of iterations by treating transition rates as constant over time periods for which this approximation leads to little error [**?**].

The **dizzysNewYann** package in R implements both the exact solution and the approximate solution for the SIR and SEIR models by integrating the R package and the C++ implementation. We can choose one of the two solutions to simulate when the number of subpopulations in a metapopulation increases. We use C++ to perform the algorithms, and R to create interfaces. Therefore, new implementation is much faster than any pure R implementation. Moreover, we introduce a new interpretation of the probabilistic derivation of multi-population epidemic model.

2. Methods

In this section, first we will talk about the deterministic model, the stochastic model of the SEIR model. Then, we will have transformation the SEIR model into the SIR model through the usage of the two algorithms. We hope that the models and the algorithms should be well understood before obtainning simulation results.

2.1. **Deterministic model:** To describe infectious diseases in a in a spatial context, we consider a metapopulation of n sub-populations. In subpopulation i of size N_i, disease dynamics can be deterministically described by the following set of differential equations:

$$\begin{eqnarray}\frac{dS_{i}}{dt}& = &\muN_{i} - \lambda_{i}S_{i} - \muS_{i}\end{eqnarray}$$