mathematik
fakultät8

Simulationstechnik A
Prof. Dr. Nicole Radde November 2, 2023
sebastian.hoepfl@isa.uni-stuttgart.de

Exercise 2 - Gillespie's Algorithm and τ-Leaping

For all our exercises, we will consider the following system consisting of substrate S_1, enzyme S_2, complex S_3 and product S_4:

$$S_1 + S_2 \xrightarrow{c_1} S_3$$
$$S_3 \xrightarrow{c_2} S_1 + S_2$$
$$S_3 \xrightarrow{c_3} S_4 + S_2$$

It is often referred to as Michaelis Menten Kinetics and will serve as a manageable system to train techniques and algorithms from the lecture.

1. **Random variables used in Gillespie's Stochastic Simulation Algorithm**

   Gillespie's Stochastic Simulation Algorithm relies on two important random variables. The first determines the type of reaction happening next while the second one specifies the duration of that reaction. Suppose there is access to a uniformly distributed random variable $\mathcal{U}(0,1)$ that returns values between 0 and 1. Find a transformation for $\mathcal{U}(0,1)$ to obtain the following random variables:

   a) A random variable $\mathcal{I}$ that returns the index for the next reaction where the probability $\mathcal{P}$ of reaction r happening is proportional to its propensity $a_r(X(t))$. The corresponding probability density is therefore given by

   $$\mathcal{P}_{\mathcal{I}}(\mathcal{I} = j) = \frac{a_j(X(t))}{\sum_{i=1}^{3} a_i(X(t))} \qquad j \in \{1,2,3\} \tag{1}$$

   A classic approach to derive the transformation is to invert the cumulative probability density function (CPDF) of the random variable of interest. For a discrete random variable, this is a step function. Since it is tedious to derive closed form, one often uses the look-up-method. The index j of the next reaction can thereby be obtained:

   1 Draw $\xi_1 \sim \mathcal{U}(0,1)$

   2 Set j to be the smallest integer satisfying $\sum_{k=1}^{j} a_k(X(t)) > \xi_1 \sum_{i=1}^{3} a_i(X(t))$

   For Michaelis-Menten this can be implemented using a few *if*-statements. In larger settings, however, an hierarchical approach might be the better choice.

b) An exponentially distributed random variable $\mathcal{E}(\lambda)$ with $\lambda = \sum_{i=1}^{3} a_i(X(t))$. The probability density of $\mathcal{E}$ is then given by

$$\mathcal{P}_{\mathcal{E}}(\mathcal{E} = t) = \lambda e^{-\lambda t} \qquad t \in \mathbb{R}_+ \tag{2}$$

One first has to derive the CPDF. Integrating the probability density function leads to

$$F_{\mathcal{E}}(t) = c - e^{-\lambda t} \tag{3}$$

and c needs to be 1 in order to construct a valid, normalized probability density function. In this case, inverting the CPDF is possible:

$$\begin{aligned} u &= F_{\mathcal{E}}(t) \\ u &= 1 - e^{-\lambda t} \\ -\frac{ln(1-u)}{\lambda} &= t \end{aligned} \tag{4}$$

By instantiating u with ξ_2, one obtains a sample from $\mathcal{E}(\lambda)$.

Hint: If a transformation in closed form is not possible, you can also provide an algorithm returning the solution.

2. **A closer look at τ-Leaping**

While Gillespie's Stochastic Simulation Algorithm approximates the solution of the underlying CME with arbitrary high precision (given enough sample paths), τ-Leaping uses an approximation preventing it from providing the exact solution.

a) Which condition needs to be fulfilled in order to keep the resulting error as small as possible?

The propensities should be nearly constant over time. This can be achieved by having very few reactions during one time increment. But the specific propensity function is of course also essential: While some propensities depend only linearly on the number of educts, others are quadratic or even of higher polynomial degree.

b) What kind of distribution is used in τ-Leaping to determine how often a certain reaction j is triggered during the next time interval $[t, t+\tau]$?

If the propensities would be exactly constant over time, we would obtain a counting process for which the probability of one reaction of type j occurring during the next infinitesimal time interval dt can be described as $dt \cdot a_j(X(t))$. One can show that the number of reactions occurring in $[t, t+\tau]$ then is Poisson distributed with $\lambda = \tau a_j(X(t))$. (Proof in SPM script)

3. **Implementing Gillespie's Stochastic Simulation Algorithm**

a) In comparison to the CME in exercise 1, what kind of boundary treatment is needed? Implement your solution.

The only needed boundary treatment is for the final state of a trace. Then, the time is set to the whole experiment duration.

b) Define the two random numbers ξ_1 and ξ_2 needed in every step of Gillespies's Algorithm according to (1) and (2).

c) Implement the condition under which each of the 3 possible reactions is triggered.

d) Describe the qualitative difference of results obtained using CME and Gillespie's Algorithm, respectively. Therefore also vary the number of traces.

4. **Implementing τ-Leaping**

a) Define the numbers of triggered reactions for each reaction type.

b) What kind of boundary treatment is needed? Implement your solution. One has to make sure that there are enough educts available for all reactions.

c) What happens if the number of time steps is reduced to very low numbers?

The quantitative error increases, but there are no instabilities as one could expect.