Stochastic Simulation: Lecture 10

Prof. Mike Giles

Oxford University Mathematical Institute

Stochastic chemical reactions

The SSA algorithm (and other equivalent methods) computes each reaction one by one – exact but very costly

"Tau-leaping" is equivalent to the Euler-Maruyama method for SDEs – the rates λ_k are frozen at the start of the timestep, so for each timestep of size h just need a sample from a Poisson distribution $Poiss(\lambda_k\, h)$ to obtain the number of reactions in that timestep.

i.e. for piecewise constant $\lambda(s)$,

$$Y\left(\int_0^{(n+1)\,h} \lambda(s)\,\mathrm{d}s\right) - Y\left(\int_0^{n\,h} \lambda(s)\,\mathrm{d}s\right) \sim Poiss(\lambda\, h)$$

Stochastic chemical reactions

Anderson & Higham (2012) developed (and analysed) a very elegant and efficient multilevel version of this algorithm – big savings because finest level usually has 1000's of timesteps.

Key challenge: how to couple coarse and fine path simulations?

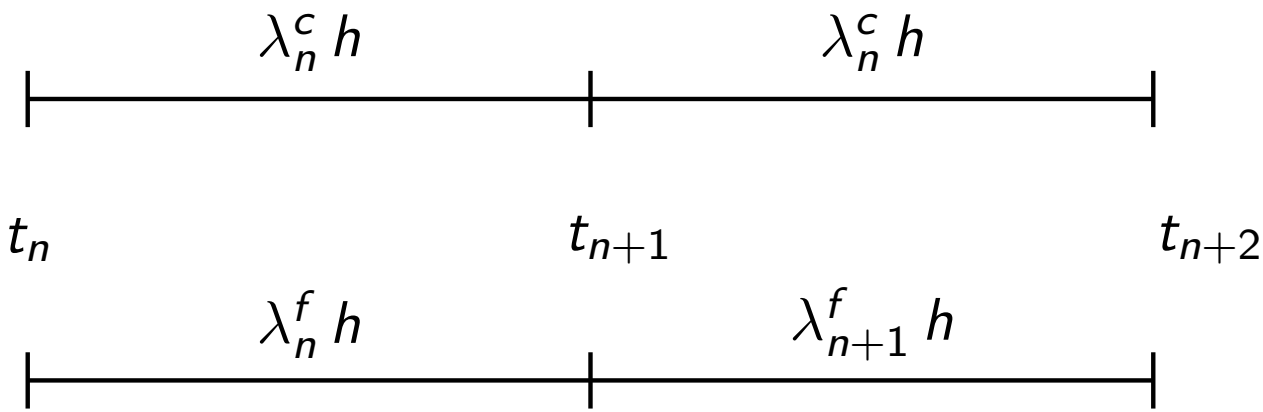
Crucial observation: for $t_1, t_2 \geq 0$

$$Poiss(t_1) + Poiss(t_2) \stackrel{d}{=} Poiss(t_1 + t_2)$$

Stochastic chemical reactions

Solution (for uniform timesteps with refinement factor of 2)

- simulate the Poisson variable on the coarse timestep as the sum of two fine timestep Poisson variables
- couple the fine path and coarse path Poisson variables by using common variable based on smaller of two rates



If $\lambda_n^f < \lambda_n^c$, use $Poiss(\lambda_n^c\, h) \sim Poiss(\lambda_n^f\, h) + Poiss((\lambda_n^c - \lambda_n^f)\, h)$
If $\lambda_n^c < \lambda_n^f$, use $Poiss(\lambda_n^f\, h) \sim Poiss(\lambda_n^c\, h) + Poiss((\lambda_n^f - \lambda_n^c)\, h)$

Tau-leaping MLMC algorithm

Input: fine timestep h, final time $T = N\,h$, refinement factor M, initial states $\widehat{X}^f = \widehat{X}^c = X$

for $n = 1, N$ **do**
  for each k **do**
    compute λ_k^f, and also λ_k^c **if** $\mathrm{mod}(n-1, M) = 0$
    $R_{1,k} := Poiss(\min(\lambda_k^f, \lambda_k^c)h)$
    $R_{2,k} := Poiss(|\lambda_k^f - \lambda_k^c|h)$
    $\widehat{X}^f := \widehat{X}^f + (R_{1,k} + 1_{\lambda_k^f > \lambda_k^c} R_{2,k})\ \zeta_K$
    $\widehat{X}^c := \widehat{X}^c + (R_{1,k} + 1_{\lambda_k^c > \lambda_k^f} R_{2,k})\ \zeta_K$
  end for
end for

Numerical analysis

Anderson & Higham also analysed the variance and proved that

$$\mathbb{E}[\,\|\widehat{X}^f - \widehat{X}^c\|^2] = O(h).$$

Since the cost is $O(h^{-1})$ this is very similar to the Euler-Maruyama method applied to SDEs, and the overall complexity is $O(\varepsilon^{-2}|\log\varepsilon|^2)$ for ε RMS error (independent of the total number of reactions performed).

Alternative couplings

The Anderson/Higham coupling is very elegant and effective, but not the only possibility

The key thing is to

- make sure the telescoping sum is respected so you are estimating the same $\mathbb{E}[P_\ell]$ in both $\mathbb{E}[P_{\ell+1} - P_\ell]$ and $\mathbb{E}[P_\ell - P_{\ell-1}]$
- try to minimise the variance $\mathbb{V}[P_\ell - P_{\ell-1}]$

Alternative couplings: I

Go back to original formulation of tau-leaping

$$\widehat{X}(t_n) = \widehat{X}(0) + \sum_k R_k(t_n)\ \zeta_k$$

where

$$R_k(t) = Y_k\left(\int_0^t \lambda_k(\widehat{X}(s))\,\mathrm{d}s\right)$$

and

$$\widehat{X}(s) = \widehat{X}(t_m),\ \text{for } s \in [t_m, t_{m+1}),$$

and use the same underlying unit-rate Poisson process Y_k for both fine and coarse paths.

Alternative couplings: I

The implementation might be quite tricky – I don't think anyone has tried this.

Might require "Binomial bridge" conditional sampling:
given $Y(t_1)$ and $Y(t_2)$, then for any $t_1 < t < t_2$ we know that

$$Y(t) - Y(t_1) \sim B\left(Y(t_2) - Y(t_1), (t - t_1)/(t_2 - t_1)\right)$$

where $B(n, p)$ returns value k with probability $\binom{n}{k} p^k (1-p)^k$

Alternative couplings: II

For a scalar random variable X with a cumulative ditribution function (CDF)

$$C(x) = \mathbb{P}(X < x)$$

can generate samples X using $X = C^{-1}(U)$ where U is a uniform random variable on $(0, 1)$.

For the coupling, we need two Poisson variates $Poiss(\lambda_1 h)$ and $Poiss(\lambda_2 h)$ for the same fine timestep. Hence, use the same U, and invert slightly different Poisson CDF's to obtain the random variables.

Alternative couplings: II

Input: fine timestep h, final time $T = N\,h$, initial states $\widehat{X}^f = \widehat{X}^c = X$

for $n = 1, N$ **do**
  for each k **do**
    compute λ_k^f, and also λ_k^c **if** $\mathrm{mod}(n-1, M) = 0$
    generate uniform r.v. U
    $R_k^f := C_{Poiss}^{-1}(\lambda_k^f\, h, U)$
    $R_k^c := C_{Poiss}^{-1}(\lambda_k^c\, h, U)$
    $\widehat{X}^f := \widehat{X}^f + R_k^f\, \zeta_K$
    $\widehat{X}^c := \widehat{X}^c + R_k^c\, \zeta_K$
  end for
end for

MLMC

Previously, I have said that the usual MLMC estimator is

$$Y_\ell = \widehat{P}_\ell(\omega) - \widehat{P}_{\ell-1}(\omega)$$

for the same $\omega \in \Omega$.

That works in many situations, but sometimes (as in this case) it is not clear what it means to have the same ω for both $\widehat{P}_\ell$ and $\widehat{P}_{\ell-1}$.

What we really need is

- $\mathbb{E}[Y_\ell] = \mathbb{E}[\widehat{P}_\ell] - \mathbb{E}[\widehat{P}_{\ell-1}]$
- $\mathbb{V}[Y_\ell] \ll 1$

MLMC

So a more general definition is

$$\widehat{Y}_\ell = \widehat{P}_\ell(\mu) - \widehat{P}_{\ell-1}(\nu)$$

where the joint distribution (μ, ν) has the correct marginals for μ and ν.

This links to the Wasserstein metric measuring the distance between two probability measures μ, ν:

$$W_p(\mu, \nu) \equiv (\inf \mathbb{E}[d(X, Y)^p])^{1/p}$$

where the inf is over all couplings such that X and Y have marginals μ and ν respectively. In 1D, an optimal coupling is given by

$$X = C_\mu^{-1}(U), \;\; Y = C_\nu^{-1}(U),$$

where $U \sim U(0, 1)$ is a unit interval uniform r.v.

Extra bits

Once the timestep is reduced down to a size for which there are very few reactions per timestep, it makes sense to switch to SSA.

Anderson & Higham (2012) came up with a very nice way to couple the finest tau-leaping level to an SSA treatment, so the final algorithm is unbiased.

The key idea is the "coarse" path uses tau-leaping, and the "fine" path uses the exact updating of the rates λ, and each reaction k can be split it into two reactions:

- one with rate $\min(\lambda^f, \lambda^c)$
- one with rate $|\lambda^f - \lambda^c|$

then use either Direct Method or Next Reaction Method for coupled simulation.

This leads to an $O(\varepsilon^{-2})$ complexity overall, with only a $(\log N)^2$ dependence on the number of reactions per path.

Extra bits

Model reduction: some biochemical reaction networks are very complex – can use a simpler approximate model (e.g. based on some forward-backward reactions being in equilibrium) as an additional "level"

Anderson & Higham (2012) also give an example of this.

Extra bits – adaptation

Adaptive time-stepping:

- Can be helpful to improve accuracy, especially when there is a fast initial transient.
- MLMC treatment essentially the same as for SDEs.

Adaptive treatment of reactions:

- some handled by SSA, some by tau-leaping, perhaps even some as Langevin SDEs
- This has been explored by Moraes *et al* (2016)

Extra bits – level 0 c.v.

Moraes *et al* (2016) also introduced an interesting control variate for the very coarsest tau-leaping level.

Start from

$$X(t) = X(0) + \sum_k Y_k \left(\int_0^t \lambda_k(X(s))\,\mathrm{d}s \right) \zeta_k,$$

replace Y_k by identity, since $\mathbb{E}[Y_k(s)] = s$, to get

$$Z(t) = X(0) + \sum_k \left(\int_0^t \lambda_k(Z(s))\,\mathrm{d}s \right) \zeta_k, \quad \Longrightarrow \dot{Z} = \sum_k \lambda_k \zeta_k$$

and then we have the approximation

$$\widetilde{X}(t) = X(0) + \sum_k Y_k \left(\int_0^t \lambda_k(Z(s))\,\mathrm{d}s \right) \zeta_k.$$

Extra bits – level 0 c.v.

Defining

$$K = \int_0^T \lambda_k(Z(s))\,\mathrm{d}s$$

then

$$\mathbb{E}[\widetilde{X}(T)] = X(0) + \sum_k K\,\zeta_k$$

and for any polynomial $f(X)$ can compute $\mathbb{E}[f(\widetilde{X}(T))]$.

$\widetilde{X}(T)$ can then be simulated using the same Y_k as the coarsest level tau-leaping $\widehat{X}$ simulation.

Extra bits

Probably good opportunities for further research in this area

Maybe explore applications in stochastic event simulation for Operational Research?

Also interesting challenges in writing generic high performance software

Key references

D.F. Anderson, D.J. Higham. "Multi-level Monte Carlo for continuous time Markov chains, with applications in biochemical kinetics". SIAM Multiscale Modelling and Simulation, 10(1):146-179, 2012.

D.F. Anderson, D.J. Higham, Y. Sun. "Complexity of multilevel Monte Carlo tau-leaping". SIAM Journal on Numerical Analysis, 52(6):3106-3127, 2014.

A. Moraes, R. Tempone, P. Vilanova. "A multilevel adaptive reaction-splitting simulation method for stochastic reaction networks". SIAM Journal on Scientific Computing, 38(4):A2091-A2117, 2016.