MCMC for continuous time switching models

Florian Stimberg
Fakultät Elektrotechnik und Informatik
Technische Universität Berlin
Berlin, Germany
flostim@cs.tu-berlin.de

Andreas Ruttor
Fakultät Elektrotechnik und Informatik
Technische Universität Berlin
Berlin, Germany
andreas.ruttor@tu-berlin.de

Manfred Opper
Fakultät Elektrotechnik und Informatik
Technische Universität Berlin
Berlin, Germany
opperm@cs.tu-berlin.de

Models involving both continuous *and* discrete state processes are of great interest, especially in systems biology[1]. Our model consist of an Ornstein-Uhlenbeck process $x(t)$, which is defined by the stochastic differential equation

$$dx = (A\mu(t) + b - \lambda x)dt + \sigma dW,$$

where $\mu(t) \in \{0, 1\}$ is a telegraph process with transition rates f_+ and f_-. This model can be used to describe transcriptional regulation in cells. The protein concentration is represented by x, while μ denotes whether the transcription factor is active. While protein concentration can be observed fairly reliable through measuring the concentration of the corresponding mRNA, the activity of the transcription factor is hard to detect. Therefore, the μ-process is unobserved in our model, but we get observations D from the x-process, which are corrupted by Gaussian noise with variance s^2.

Given these observations we want to infer the posterior distribution of $\mathbf{x}$, $\mu_{0:T}$, as well as the model parameters $\Theta = \{A, b, \lambda, \sigma, s, f_+, f_-\}$. For this purpose we use a Gibbs sampler, which alternates between sampling $\mu_{0:T}$ conditioned on $\mathbf{x}$ and Θ, $\mathbf{x}$ given $\mu_{0:T}$ and Θ, and each parameter in Θ conditioned on all the other parameters, $\mathbf{x}$, and $\mu_{0:T}$.

A path of $\mu_{0:T}$ is sampled in continuous time using a Metropolis-Hastings step. For drawing a proposal of $\mu_{0:T}$ we compute the posterior transition rates approximately assuming the transition rates are piecewise constant. We then use Gillespie's algorithm to sample a new path for $\mu_{0:T}$. This proposed path is accepted according to the Metropolis-Hastings ratio, which can be expressed exactly in terms of the jump times[2].

Since $p(x(t + \Delta t)|x(t), D, \mu_{0:T}, \Theta)$ is Gaussian, we can easily draw samples from the x-process at arbitrary points in time. The parameters A, b, s, and the transition rates f_+, f_- can be sampled directly, while we apply a Metropolis-Hastings step to sample λ and σ.

It is important to highlight that our sampling scheme gives exact results independent of the time discretization used. The only influence of the discretization is that smaller time steps lead to higher acceptance rates for the Metropolis-Hastings steps. The method presented here can be extended to more realistic scenarios of multiple proteins and transcription factors.

[1] Guido Sanguinetti, Andreas Ruttor, Manfred Opper, and Cedric Archambeau (2009) Switching regulatory models of cellular stress response. *Bioinformatics*, 25,1280-1286.

[2] Darren J. Wilkinson (2006) *Stochastic Modelling for Systems Biology*. Chapman & Hall / CRC