Capturing rare events with the heterogeneous multiscale method

David Kelly

Eric Vanden-Eijnden

Courant Institute
New York University
New York NY
www.dtbkelly.com

October 1, 2016

SIAM MPE 16, *Philadelphia, PE*

Fast-slow systems

Fast slow SDEs:

$$\frac{dX^\varepsilon}{dt} = f(X^\varepsilon, Y^\varepsilon)$$
$$\frac{dY^\varepsilon}{dt} = \varepsilon^{-1} g(X^\varepsilon, Y^\varepsilon) + \varepsilon^{-1/2}\sigma(X^\varepsilon, Y^\varepsilon)\frac{dW}{dt}$$

where $\varepsilon \ll 1$.

Let Y_x be 'virtual fast process' with frozen x:

$$\frac{dY_x}{dt} = g(x, Y_x) + \sigma(x, Y_x)\frac{dW}{dt}$$

Assume that Y_x has an ergodic invariant measure μ_x and is sufficiently mixing.

Averaging

The slow variables satisfy an **averaging principle**

$$X^{\varepsilon} \to_{a.s.} \overline{X} \quad \text{where} \quad \frac{d\overline{X}}{dt} = F(\overline{X})$$

and $F(x) = \int f(x,y)\mu_x(dy)$.

A simple metastable example

Suppose $\mu > 0$ and

$$\frac{dX^\varepsilon}{dt} = Y^\varepsilon - (X^\varepsilon)^3$$
$$dY^\varepsilon = \frac{\theta}{\varepsilon}(\mu X^\varepsilon - Y^\varepsilon)dt + \frac{\sigma}{\sqrt{\varepsilon}}dW$$

This has averaged equation $\frac{d\overline{X}}{dt} = \mu\overline{X} - \overline{X}^3$. Symmetric double-well potential w/equilibria at $\pm\sqrt{\mu}$ and saddle at origin.

When $\varepsilon \ll 1$, the long time behavior of X^ε will be qualitatively different to the averaged system. The system exhibits hopping between wells due to fluctuations from the average.

The central limit theorem describes **small fluctuations** about the average.

If we let $Z^\varepsilon = \varepsilon^{-1/2}(X^\varepsilon - \overline{X})$ then one can show $Z^\varepsilon \to_w \overline{Z}$ where

$$d\overline{Z} = B_0(\overline{X})\overline{Z}dt + \eta(\overline{X})dV$$

where V is a std Brownian motion and

$$\begin{aligned} B_0(x) &= \int \nabla_x f(x,y)\mu_x(dy) \\ &\quad + \int_0^\infty \int \nabla_y \mathbf{E}_y(\tilde{f}(x, Y_x(\tau)))\nabla_x b(x,y)\mu_x(dy)d\tau \\ \eta(x)\eta^T(x) &= \int_0^\infty \mathbf{E}\tilde{f}(x, Y_x(\tau))\tilde{f}(x, Y_x(0)))^T d\tau \end{aligned}$$

where $\tilde{f}(x,y) = f(x,y) - F(x)$.

Suppose X^ε satisfies a **large deviations principle**:

$$\lim_{\varepsilon \to 0} \varepsilon \log \mathbf{P}(X^\varepsilon \in \Gamma) = - \inf_{\gamma \in \Gamma} \mathcal{S}_{[0,T]}(\gamma)$$

for a set Γ of continuous paths $\gamma : [0, T] \to \mathbb{R}^d$ in the slow state space.

A large deviation principle quantifies many important features of $O(1)$ fluctuations in metastable systems.

For instance, suppose that $D \subset \mathbb{R}^d$ is open w/ smooth boundary ∂D, and x^* is an asymptotically stable equilibrium for the averaged system $\frac{d\overline{X}}{dt} = F(\overline{X})$.

Define the transition time $\tau^\varepsilon = \inf\{t > 0 : X^\varepsilon \notin D\}$. Define the *quasi-potential*

$$\mathcal{V}(x, y) = \inf_{T>0} \inf_{\gamma(0)=x, \gamma(T)=y} \mathcal{S}_{[0,T]}(\gamma)$$

Then the mean first passage/exit time is given by

$$\lim_{\varepsilon \to 0} \varepsilon \log \mathbf{E} \tau^\varepsilon = \inf_{y \in \partial D} \mathcal{V}(x, y)$$

For FS systems, **Varadhan's Lemma** (reverse) tells us the following:

Let $u(t,x) = \lim_{\varepsilon \to 0} \varepsilon \log \mathbf{E}_x \exp(\varepsilon \varphi(X^\varepsilon(t)))$. If u satisfies the Hamilton-Jacobi equation

$$\partial_t u = \mathcal{H}(x, \nabla u) \quad , \quad u(0,x) = \varphi$$

for suitable class of φ, then X^ε satisfies an LDP with rate function

$$\mathcal{S}_{[0,T]}(\gamma) = \int_0^T \mathcal{L}(\gamma(s), \dot{\gamma}(s)) ds$$

where $\mathcal{L}$ is the Lagrangian associated with the Hamiltonian $\mathcal{H}$

$$\mathcal{L}(x, \beta) = \sup_\theta (\theta \cdot \beta - \mathcal{H}(x, \theta)) \ .$$

Moral of the story: we can identify LDPs via the associated HJ equation.

Heterogeneous multi scale method for FS systems

A simple numerical scheme for the slow variables $x_n^\varepsilon \approx X^\varepsilon(n\Delta t)$ when $\varepsilon \ll 1$:

$$x_{n+1}^\varepsilon = x_n^\varepsilon + \int_{n\Delta t}^{(n+1)\Delta t} f(x_n^\varepsilon, Y_{x_n^\varepsilon}^\varepsilon(s))ds$$

Then approximate the integral by simulating the virtual fast process on mesh size $\delta t \ll \Delta t$

$$\int_{n\Delta t}^{(n+1)\Delta t} f(x_n^\varepsilon, Y_{x_n^\varepsilon}^\varepsilon(s))ds \approx \sum_{j=0}^{N-1} f(x_n^\varepsilon, y_{n,j}^\varepsilon)\delta t$$

where $N\delta t = \Delta t$ and (for instance) is given by Euler-Maruyama

$$y_{n,j+1}^\varepsilon = y_{n,j}^\varepsilon + \varepsilon^{-1} g(x_n^\varepsilon, y_{n,j}^\varepsilon)\delta t + \varepsilon^{-1/2}\sigma(x_n^\varepsilon, y_{n,j}^\varepsilon)\sqrt{\delta t}\xi_{n,j}$$

for $j = 0, \dots, N-1$.

Speeding up the method

The key observation of HMM is that one does not need the virtual process Y_x^ε over the whole window $[n\Delta t, (n+1)\Delta t)$, but only over a fraction of it $[n\Delta t, (n+1/\lambda)\Delta t]$ for some $\lambda \geq 1$.

By the ergodic theorem

$$\frac{1}{\Delta t}\int_{n\Delta t}^{(n+1)\Delta t} f(x_n^\varepsilon, Y_{x_n^\varepsilon}^\varepsilon(s))ds \approx F(x_n^\varepsilon) \approx \frac{\lambda}{\Delta t}\int_{n\Delta t}^{(n+1/\lambda)\Delta t} f(x_n^\varepsilon, Y_{x_n^\varepsilon}^\varepsilon(s))ds$$

provided that $\Delta t/\varepsilon$ and $\Delta t/(\varepsilon\lambda)$ are larger than the mixing time for Y_x.

HMM summary

The update $x_n^\varepsilon \mapsto x_{n+1}^\varepsilon$ works in two steps

1 - Micro step: Compute an approximation $F_{n,\lambda}(x_n^\varepsilon)$ of the integral

$$\frac{\lambda}{\Delta t} \int_{n\Delta t}^{(n+1/\lambda)\Delta t} f(x_n^\varepsilon, Y_{x_n^\varepsilon}^\varepsilon(s)) ds$$

by simulating the virtual fast process $Y_{x_n^\varepsilon}^\varepsilon$ over the window $[n\Delta t, (n+1/\lambda)\Delta t)$. Requires $\delta t \ll \Delta t$, $\delta t \ll \varepsilon$ and $\Delta t/(\varepsilon\lambda)$ larger than mixing time.

2 - Macro step: $x_{n+1}^\varepsilon = x_n^\varepsilon + F_{n,\lambda}(x_n^\varepsilon)\Delta t$

We know that HMM is **consistent with the averaging principle**. That is, as $\varepsilon \to 0$ the sequence x_n^ε defined by HMM converges to

$$\overline{x}_{n+1} = \overline{x}_n + F(\overline{x}_n)\Delta t$$

which is a consistent numerical method for the averaged equation $\frac{d\overline{X}}{dt} = F(\overline{X})$.

What about fluctuations?

1 - Let $z_n^\varepsilon = \varepsilon^{-1/2}(x_n^\varepsilon - \overline{x}_n)$. Does z_n^ε converge to a numerical scheme for $\overline{Z}$ as $\varepsilon \to 0$?

2 - Let $u_{n,\lambda}(x) = \lim_{\varepsilon\to 0} \varepsilon \log \mathbf{E}_x \exp(\varepsilon^{-1}\varphi(x_n^\varepsilon))$. Is $u_{n,\lambda}$ a numerical method for the true HJ equation?

HMM Fluctuations are inflated by λ

As $\varepsilon \to 0$, z_n^ε converges to $\overline{z}_n$, which is a numerical scheme for the SDE

$$d\overline{Z}_\lambda = B_0(\overline{X})\overline{Z}_\lambda dt + \sqrt{\lambda}\eta(\overline{X})dV \ .$$

Moreover, we find that $u_{\lambda,n}(x)$ is a numerical method for the HJ equation

$$\partial_t u_\lambda = \frac{1}{\lambda}\mathcal{H}(x, \lambda \nabla u_\lambda)$$

where $\mathcal{H}$ is the true Hamiltonian for X^ε.In particular, the quasi-potential is $\mathcal{V}_\lambda(x,y) = \lambda^{-1}\mathcal{V}(x,y)$. It follows that mean first passage times will shrink

$$\mathbf{E}\tau_\varepsilon \asymp \exp\left(\frac{1}{\varepsilon\lambda}\mathcal{V}(x^*, \partial D)\right)$$

Why the inflation?

In the HMM approximation, with $\lambda \in \mathbb{Z}$, we are essentially replacing

$$\int_{n\Delta t}^{(n+1/\lambda)\Delta t} f(x, Y_x^\varepsilon(s))ds + \cdots + \int_{(n+(\lambda-1)/\lambda)\Delta t}^{(n+1)\Delta t} f(x, Y_x^\varepsilon(s))ds$$

with

$$\int_{n\Delta t}^{(n+1/\lambda)\Delta t} f(x, Y_x^\varepsilon(s))ds + \cdots + \int_{n\Delta t}^{(n+1/\lambda)\Delta t} f(x, Y_x^\varepsilon(s))ds$$

ie. Replace sum of λ weakly correlated random variables with $\lambda \times$ first random variable. Clearly this inflates the variance.

Parallel HMM

There is a simple way to fix the problem. The update $x_n^\varepsilon \mapsto x_{n+1}^\varepsilon$ works in two steps

1 - λ **parallel micro steps**: Compute an approximation $F_{n,\lambda}(x_n^\varepsilon)$ of the integral

$$\sum_{k=1}^{\lambda} \frac{1}{\Delta t} \int_{n\Delta t}^{(n+1/\lambda)\Delta t} f(x_n^\varepsilon, Y_{x_n^\varepsilon,k}^\varepsilon(s))ds$$

by simulating λ independent copies of the virtual fast processes $Y_{x_n^\varepsilon,k}^\varepsilon$ for $k = 1, \ldots, \lambda$ over the window $[n\Delta t, (n+1/\lambda)\Delta t)$.

2 - **Macro step**: $x_{n+1}^\varepsilon = x_n^\varepsilon + F_{n,\lambda}(x_n^\varepsilon)\Delta t$

Parallel HMM

- Since the virtual fast processes are independent, they can be simulated in parallel. This is a kind of parallel-in-time method.

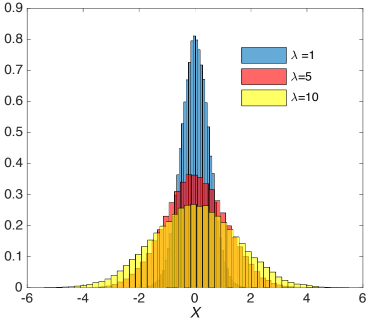
- We can show that this method is in fact consistent with X^ε at both the level of small fluctuations and large deviations.

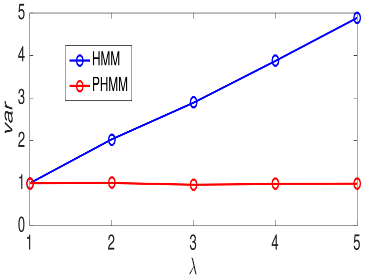
Small fluctuations example I

Suppose $\mu < 1$ and

$$\frac{dX^\varepsilon}{dt} = Y^\varepsilon - X^\varepsilon$$
$$dY^\varepsilon = \frac{\theta}{\varepsilon}(\mu X^\varepsilon - Y^\varepsilon)dt + \frac{\sigma}{\sqrt{\varepsilon}}dW$$

This has averaged equation $\frac{d\overline{X}}{dt} = (\mu - 1)\overline{X}$.





Large deviations example

Suppose $\mu > 0$ and

$$\frac{dX^\varepsilon}{dt} = Y^\varepsilon - (X^\varepsilon)^3$$

$$dY^\varepsilon = \frac{\theta}{\varepsilon}(\mu X^\varepsilon - Y^\varepsilon)dt + \frac{\sigma}{\sqrt{\varepsilon}}dW$$

This has averaged equation $\frac{d\overline{X}}{dt} = \mu\overline{X} - \overline{X}^3$.

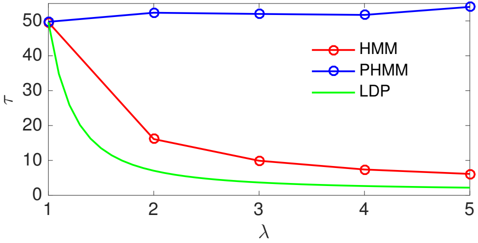


Figure: Mean first passage time

References

D. Kelly, E. Vanden-Eijnden. *Capturing rare events with the heterogeneous multiscale method.* **arXiv** (2016).

All my slides are on my website (www.dtbkelly.com) **Thank you**!