Error estimates for Feynman-Kac semi-groups.

IHP - Young Researchers' Seminar

Grégoire Ferré – Gabriel Stoltz

CERMICS - ENPC, INRIA Paris & Labex BEZOUT

Thursday, June 22nd, 2017

École des Ponts ParisTech

www.enpc.fr

Outline

1. Problem and motivation

2. Error estimates on ergodic averages

3. Error estimates for Feynman-Kac semi-groups

1. Problem and motivation

Large deviations

- Consider an ergodic dynamics on the torus with invariant measure ν:

$$dX_t = b(X_t)dt + \sqrt{2\beta^{-1}}dB_t.$$

- Idea of large deviations:

$$\mathbb{P}\left[\frac{1}{t}\int_0^t W(X_s)ds = a\right] \asymp \mathrm{e}^{-tI(a)},$$

where I is the **rate function**.

Large deviations

- Consider an ergodic dynamics on the torus with invariant measure ν:

$$dX_t = b(X_t)dt + \sqrt{2\beta^{-1}}dB_t.$$

- Idea of large deviations:

$$\mathbb{P}\left[\frac{1}{t}\int_0^t W(X_s)ds = a\right] \asymp \mathrm{e}^{-tI(a)},$$

where I is the **rate function.**

- **Donsker-Varadhan** [1975]: if one sets

$$\lambda(k) := \sup_{a\in\mathbb{R}} \{ka - I(a)\},$$

then $\lambda(k)$ is the largest eigenvalue of $\mathcal{L} + kW$ where $\mathcal{L}$ is the generator of (X_t).

Feynman-Kac semi-groups

- Consider (λ, ν_W) the principal eigenvalue and eigenfunction of $\mathcal{L}^\dagger + W$. **Feynman-Kac** formula gives

$$\lambda = \lim_{t\to\infty} \frac{1}{t} \log \mathbb{E}\left[e^{\int_0^t W(X_s)ds} \right].$$

- We then consider, for an observable φ,

$$\Phi_t(\mu)(\varphi) = \frac{\mathbb{E}_\mu\left[\varphi(X_t) e^{\int_0^t W(X_s)ds}\right]}{\mathbb{E}_\mu\left[e^{\int_0^t W(X_s)ds}\right]} \xrightarrow[t\to\infty]{} \int_{\mathcal{D}} \varphi \, d\nu_W.$$

Feynman-Kac semi-groups

- Consider (λ, ν_W) the principal eigenvalue and eigenfunction of $\mathcal{L}^\dagger + W$. **Feynman-Kac** formula gives

$$\lambda = \lim_{t\to\infty} \frac{1}{t} \log \mathbb{E}\left[e^{\int_0^t W(X_s)ds} \right].$$

- We then consider, for an observable φ,

$$\Phi_t(\mu)(\varphi) = \frac{\mathbb{E}_\mu\left[\varphi(X_t) e^{\int_0^t W(X_s)ds} \right]}{\mathbb{E}_\mu\left[e^{\int_0^t W(X_s)ds} \right]} \xrightarrow[t\to\infty]{} \int_{\mathcal{D}} \varphi\, d\nu_W.$$

- Problem related to Markov processes conditioned on large deviations (Chetrite & Touchette – Tsobgni-Nyawo & Touchette – Angeletti & Touchette).
- Discretization and error estimates for such semi-groups ?

Problem: error estimates

- We discretize the process (X_t) into a Markov chain (x_n) with evolution operator $Q_{\Delta t}\varphi(x) = \mathbb{E}(\varphi(x_{n+1})|x_n = x)$. Consider for example the estimator:

$$\Phi_{\Delta t,n}(\mu)(\varphi) = \frac{\mathbb{E}_\mu\left[\varphi(x_n)\mathrm{e}^{\Delta t \sum_{i=0}^{n-1} W(x_i)}\right]}{\mathbb{E}_\mu\left[\mathrm{e}^{\Delta t \sum_{i=0}^{n-1} W(x_i)}\right]} \xrightarrow[n\to\infty]{} \int_{\mathcal{D}} \varphi \, d\nu_{W,\Delta t}.$$

- Natural question: is $\nu_{W,\Delta t}$ close to ν_W? Can we find $C > 0$, $p \geqslant 1$ s.t.

$$\left| \int_{\mathcal{D}} \varphi \, d\nu_{W,\Delta t} - \int_{\mathcal{D}} \varphi \, d\nu_W \right| \leqslant C \Delta t^p,$$

depending on the numerical scheme?

2. Error estimates on ergodic averages

Error estimates on the time step bias

- Ergodic average:

$$\frac{1}{N_{\text{iter}}} \sum_{n=0}^{N_{\text{iter}}-1} \varphi(x_n) \xrightarrow[n\to\infty]{} \int_{\mathcal{D}} \varphi \, d\nu_{\Delta t}.$$

- Systematic error (Talay & Tubaro – Mattingly, Stuart & Tretyakov – Debussche & Faou – Abdulle, Vilmart, Konstantinos & Zygalakis – Leimkuhler, Matthews & Stoltz ...)

Theorem

There exists $p \geqslant 1$ and f regular such that for all φ

$$\int_{\mathcal{D}} \varphi \, d\nu_{\Delta t} = \int_{\mathcal{D}} \varphi \, d\nu + \Delta t^p \int_{\mathcal{D}} \varphi f \, d\nu + O(\Delta t^{p+1})$$

- The function f is solution to the Poisson equation of the form $\mathcal{L}^* f = g$, where g and p depend on the numerical scheme.

Sketch of proof

Idea: $Q_{\Delta t}$ approximates the exact flow $\mathrm{e}^{t\mathcal{L}}$ over a time step Δt. We assume that:
$Q_{\Delta t}\varphi = \varphi + \Delta t \mathcal{A}_1\varphi + \ldots + \Delta t^p \mathcal{A}_p\varphi + \Delta t^{p+1}\mathcal{A}_{p+1}\varphi + O(\Delta t^{p+2})$
The proof relies on the stationarity equations,

$$\int_{\mathcal{D}} P_t\varphi d\nu = \int_{\mathcal{D}} \varphi d\nu, \quad \int_{\mathcal{D}} Q_{\Delta t}\varphi d\nu_{\Delta t} = \int_{\mathcal{D}} \varphi d\nu_{\Delta t}.$$

Idea: expand $\nu_{\Delta t}$ as a function of ν, *i.e.* search f s.t.

$$\int_{\mathcal{D}} (Q_{\Delta t}\varphi)(1 + \Delta t^p f) d\nu = \int_{\mathcal{D}} \varphi(1 + \Delta t^p f) d\nu + \Delta t^{p+2} R_{\Delta t}(\varphi).$$

Then, if $\int_{\mathcal{D}} \mathcal{A}_k\varphi d\nu = 0$ for $k = 1, \ldots, p$ we have at order $p+1$

$$\int_{\mathcal{D}} \Big[\mathcal{A}_{p+1}\varphi + (\mathcal{A}_1\varphi) f\Big] d\nu = 0, \text{ so } f = -(\mathcal{A}_1^*)^{-1}\mathcal{A}_{p+1}\mathbb{1}.$$

3. Error estimates for Feynman-Kac semi-groups

A double discretization

- We want to discretize the flow

$$\mathbb{E}\left[\varphi(X_t)\mathrm{e}^{\int_0^t W(X_s)\,ds}\right] = \mathrm{e}^{t(\mathcal{L}+W)}\varphi$$

  over a time step Δt.

- We need to discretize both the process *and* the integral. We denote $Q_{\Delta t}^W$ the approximated flow, for example

$$(Q_{\Delta t}^W\varphi)(x) = \mathrm{e}^{\Delta t W(x)}(Q_{\Delta t}\varphi)(x), \quad (Q_{\Delta t}^W\varphi)(x) = \mathrm{e}^{\frac{\Delta t}{2}W(x)}(Q_{\Delta t}\varphi\mathrm{e}^{\frac{\Delta t}{2}W})(x).$$

- Same strategy for the long time error, show that

$$Q_{\Delta t}^W \approx \mathrm{e}^{\Delta t(\mathcal{L}+W)}$$

  with an expansion in Δt.

Main result

We assume that:

$$Q_{\Delta t}^{W}\varphi = \varphi + \Delta t \mathcal{A}_1^W \varphi + \ldots + \Delta t^p \mathcal{A}_p^W \varphi + \Delta t^{p+1} \mathcal{A}_{p+1}^W \varphi + O(\Delta t^{p+2}).$$

Theorem: error estimate on the invariant measure

Under «mild» assumptions, if for $k = 1, \ldots, p$ there exists $a_k \in \mathbb{R}$ s.t.

$$\forall\, \varphi \in \mathcal{C}^\infty, \quad \int_{\mathcal{D}} \mathcal{A}_k^W \varphi \, d\nu_W = a_k \int_{\mathcal{D}} \varphi \, d\nu_W,$$

then

$$\int_{\mathcal{D}} \varphi \, d\nu_{W,\Delta t} = \int_{\mathcal{D}} \varphi \, d\nu_W + \Delta t^p \int_{\mathcal{D}} \varphi f \, d\nu_W + O(\Delta t^{p+1})$$

Sketch of proof

Again, start from stationarity equations:

$$\int_{\mathcal{D}} \mathrm{e}^{t(\mathcal{L}+W)} \varphi \, d\nu_W = \mathrm{e}^{t\lambda} \int_{\mathcal{D}} \varphi \, d\nu_W,$$

and

$$\int_{\mathcal{D}} Q_{\Delta t}^{W} \varphi \, d\nu_{W,\Delta t} = \underbrace{\left(\int_{\mathcal{D}} Q_{\Delta t}^{W}(\mathbb{1}) \, d\nu_{W,\Delta t} \right)}_{\text{creation of probability: } \mathrm{e}^{\Delta t \lambda_{\Delta t}}} \left(\int_{\mathcal{D}} \varphi \, d\nu_{W,\Delta t} \right).$$

Sketch of proof

Again, start from stationarity equations:

$$\int_{\mathcal{D}} e^{t(\mathcal{L}+W)}\varphi d\nu_W = e^{t\lambda} \int_{\mathcal{D}} \varphi d\nu_W,$$

and

$$\int_{\mathcal{D}} Q_{\Delta t}^W \varphi d\nu_{W,\Delta t} = \underbrace{\left(\int_{\mathcal{D}} Q_{\Delta t}^W(\mathbb{1}) d\nu_{W,\Delta t}\right)}_{\text{creation of probability: } e^{\Delta t \lambda_{\Delta t}}} \left(\int_{\mathcal{D}} \varphi d\nu_{W,\Delta t}\right).$$

Next steps:

- identify f solution to a Poisson equation s. t. $\nu_{W,\Delta t} \approx (1 + \Delta t^p f)\nu_W$,
- build an approximate eigenvector $\hat{h}_{W,\Delta t}$ s.t. $Q_{\Delta t}^W \hat{h}_{W,\Delta t} \approx e^{\Delta t \lambda_{\Delta t}} \hat{h}_{W,\Delta t}$,
- a priori estimate on the creation of probability $e^{\Delta t \lambda_{\Delta t}}$.
- technical details.

Useful corollary

- We are mainly interested in the eigenvalue λ. We define:

$$\mathrm{e}^{\Delta t \lambda_{\Delta t}} := \int_{\mathcal{D}} Q_{\Delta t}^{W}(\mathbb{1}) d\nu_{W,\Delta t}.$$

Useful corollary

- We are mainly interested in the eigenvalue λ. We define:

$$\mathrm{e}^{\Delta t \lambda_{\Delta t}} := \int_{\mathcal{D}} Q_{\Delta t}^{W}(\mathbb{1}) d\nu_{W,\Delta t}.$$

- Error estimate on $\lambda_{\Delta t}$

Eigenvalue as a partition function

If $Q_{\Delta t}^{W}$ is consistent at order p, then

$$\lambda_{\Delta t} := \frac{1}{\Delta t} \log \left[\int_{\mathcal{D}} Q_{\Delta t}^{W}(\mathbb{1}) d\nu_{W,\Delta t} \right] = \lambda + \Delta t^{p} C + O(\Delta t^{p+1}),$$

where $C \in \mathbb{R}$ is a constant depending on f.

Consequences

Applications to statistical physics and Diffusion Monte Carlo:

- If the dynamics is discretized with a second order scheme $Q_{\Delta t}$, then the splitting

$$Q_{\Delta t}^{W} = \mathrm{e}^{\frac{\Delta t}{2} W}\left(Q_{\Delta t}(\mathrm{e}^{\frac{\Delta t}{2} W}\cdot)\right)$$

  provides a second-order discretization of the Feynman-Kac semi-group.
- For Diffusion Monte Carlo, the dynamics (X_t) is a brownian motion, so the flow $Q_{\Delta t}$ is exact. The above splitting is immediately second order.

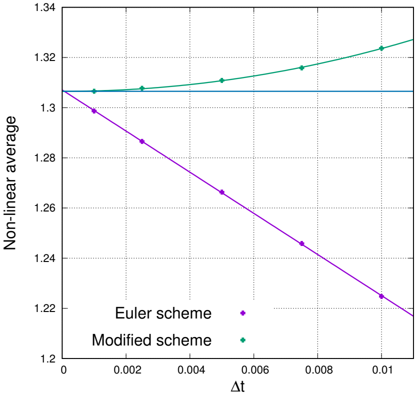
Statistical approximation

- Path average over a set of replicas $(x_n^m)_{m=1}^M$,

$$\Phi_{\Delta t,n}(\varphi) \approx \frac{\frac{1}{M}\sum_{m=1}^{M} \varphi(x_n^m) \mathrm{e}^{\sum_{k=0}^{n-1} W(x_k^m)\Delta t}}{\frac{1}{M}\sum_{m=1}^{M} \mathrm{e}^{\sum_{k=0}^{n-1} W(x_k^m)\Delta t}}$$

- Important variance of the exponential weights.
- Population dynamics over the M replicas of the systems. For each step:
    1. evolve the replicas with kernel $Q_{\Delta t}$,
    2. compute probabilities $(p_m)_{m=1}^M$ from the weights $\mathrm{e}^{\Delta t W}$,
    3. kill or clone each replica with probability p_m and resize the population,
    4. average φ over the replicas.

Overdamped Langevin dynamics on a one dimensional torus.



Estimation of $\lambda_{\Delta t}$ for:

- $dX_t = -V'(X_t)dt + dB_t$,
- $V(x) = cos(2\pi x)$,
- $W = |V|^2$,
- Euler-Maruyama scheme and 2nd order modified scheme,
- comparison to Galerkin discretization.

We indeed observe first and second order convergence.

Conclusion & Tracks

Conclusion

Summary:

- a numerical large deviation problem,
- discretization of SDE′s and error on the invariant measure,
- new error estimates on the invariant measure of Feynman-Kac semi-groups,
- alternative representation of the principal eigenvalue.

Perspectives

- convergence of Feynman-Kac semi-groups for an unbounded state space,
- adaptative scheme, variance reduction.

Conclusion & Tracks

Conclusion

Summary:

- a numerical large deviation problem,
- discretization of SDE′s and error on the invariant measure,
- new error estimates on the invariant measure of Feynman-Kac semi-groups,
- alternative representation of the principal eigenvalue.

Perspectives

- convergence of Feynman-Kac semi-groups for an unbounded state space,
- adaptative scheme, variance reduction.