Optimal importance sampling using stochastic control

Brandenburg University of Technology Cottbus - Senftenberg

Lorenz Richter, Carsten Hartmann

Freie Universität Berlin

The setting

Goal: Compute $\mathbb{E}_\pi[f(X)], f:\mathbb{R}^d \to \mathbb{R}$. Model π as stationary distribution of a stochastic process $X_t \in \mathbb{R}^d$ associated with path measure $\mathbb{P}$ and described by

$$\mathrm{d}X_t = \nabla \log \pi(X_t)\mathrm{d}t + \sqrt{2}\mathrm{d}W_t$$

and compute $\mathbb{E}_\mathbb{P}[f(X_T)]$.

Objective: Deal with variance and convergence speed.

Idea: Consider the controlled process associated with path measure $\mathbb{Q}^u$

$$\mathrm{d}X_t^u = \left(\nabla \log \pi(X_t^u) + \sqrt{2}u(X_t^u, t)\right)\mathrm{d}t + \sqrt{2}\mathrm{d}W_t$$

and do importance sampling in path space: $\mathbb{E}_{\mathbb{Q}^u}\left[f(X_T^u)\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{Q}^u}\right]$.

Outlook: Exploit ergodicity to find a good control for infinite times:

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T f(X_t)\mathrm{d}t = \mathbb{E}_\pi[f(X)].$$

Optimal change of measure

Donsker-Varadhan: Consider $W(X_{0:T}) = \int_0^T g(X_t,t)\mathrm{d}t + h(X_T)$. The duality between sampling and an optimal change of measure

$$\gamma(x,t) := -\log \mathbb{E}_\mathbb{P}[\exp(-W)|X_t = x] = \inf_{\mathbb{Q}^u \ll \mathbb{P}}\{\mathbb{E}_{\mathbb{Q}^u}[W] + \mathrm{KL}(\mathbb{Q}^u\|\mathbb{P})\}$$

brings a zero-variance estimator, i.e. $\mathrm{Var}_{\mathbb{Q}^*}\left(\exp(-W)\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{Q}^*}\right) = 0.$

→ $\gamma(x,t)$ is the value function of a control problem and fulfills the HJB equation

→ the control costs are $J(u) = \mathbb{E}\left[\int_0^T \left(g(X_t^u,t) + \frac{1}{2}|u(X_t^u,t)|^2\right)\mathrm{d}t + h(X_T^u)\right]$

→ the optimal change of measure reads $\mathrm{d}\mathbb{Q}^* = \frac{e^{-W}}{\mathbb{E}[e^{-W}]}\mathrm{d}\mathbb{P}$

→ the optimal control is $u^*(x,t) = -\sqrt{2}\nabla_x\gamma(x,t)$

→ $\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{Q}^u}$ is computed via Girsanov's theorem

Computing the optimal change of measure

Gradient descent

Parametrize the control (i.e. the change of measure) in ansatz functions φ_i:

$$u(x,t) = -\sqrt{2}\sum_{i=1}^m \alpha_i(t)\varphi_i(x).$$

Compute the minimization of the costs $J(u)$ with a gradient descent in α, i.e.

$$\alpha^{k+1} = \alpha^k - \eta_k\nabla_\alpha\hat{J}(u(\alpha^k))$$

Different estimators of the gradient scale differently with the time horizon T:

- $G_{\text{finite differences}} \propto T$
- $G_{\text{centered likelihood ratio}} \propto T^2$
- $G_{\text{likelihood ratio}} \propto T^3$

Cross-entropy method

It holds $J(u) = J(u^*) + \mathrm{KL}(\mathbb{Q}^u\|\mathbb{Q}^*)$, however we minimize $\mathrm{KL}(\mathbb{Q}^*\|\mathbb{Q}^u)$ since this is feasible.
Again parametrize $u(\alpha)$ in ansatz functions. We then need to minimize the cross-entropy functional for which we get a necessary optimality condition in form of the linear equation $S\alpha = b$, which we solve iteratively.

Approximate policy iteration

Successive linearization of HJB equation: the cost functional $J(u)$ solves a PDE of the form

$$A(u)J(u) = l(x,u),$$

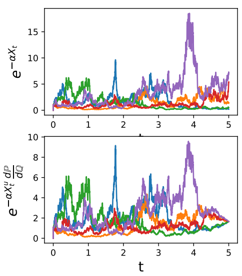
where $\gamma(x,t) = \min_u J(u;x,t)$. Start with an initial guess of the control policy and iterate

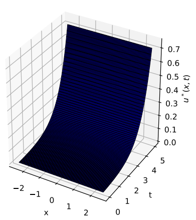
$$u_{k+1}(x,t) = -\sqrt{2}\nabla_x J(u_k;x,t),$$

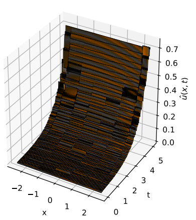
yielding a fixed point iteration in the non-optimal control u.

Numerical simulations

- Sample $\mathbb{E}[\exp(-\alpha X_T)]$ with $\pi = \mathcal{N}(0,1)$ and $u^*(x,t) = -\sqrt{2}\alpha e^{t-T}$, $\hat{u}$ determined by gradient decent







- normalized variances with $\alpha = 1, T = 5, \mathbb{E}[\exp(-\alpha X_T)] = 1.649$:

Δt	vanilla	with u^*	with $\hat{u}$
10^{-2}	4.64	1.70×10^{-4}	5.69×10^{-2}
10^{-3}	4.02	1.69×10^{-7}	3.21×10^{-2}
10^{-4}	4.55	1.73×10^{-8}	6.45×10^{-2}

- Sample $\mathbb{E}[X_T]$ with $\pi = \mathcal{N}(b,1)$, i.e. consider $h(x) = -\log x$, w.r.t.

$$\mathrm{d}X_s = (-X_s + b)\mathrm{d}s + \sqrt{2}\mathrm{d}W_s, X_t = x$$

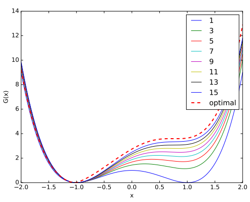
– then $X_T \sim \mathcal{N}(b + (x-b)e^{t-T}, 1 - e^{2(t-T)})$

$$u^*(x,t) = \frac{\sqrt{2}e^{t-T}}{b + (x-b)e^{t-T}}$$

- **Molecular dynamics**: compute transition probabilities, i.e. first hitting times

– consider $g = 1, h = 0$, i.e. we sample hitting times $\mathbb{E}[e^{-\tau}]$

– optimal change of measure can be interpreted as a tilting of the potential guiding the dynamics



References

[1] Hartmann, C., Richter, L., Schütte, C., & Zhang, W. (2017). Variational characterization of free energy: theory and algorithms. Entropy, 19(11), 626.
[2] Richter, L. (2016). Efficient statistical estimation using stochastic control and optimization. Master's thesis, Freie Universität Berlin.
[3] Fleming, W.H., Soner, H.M. (2006). Controlled Markov processes and viscosity solutions. Springer, New York, NY, USA.
[4] Dai Pra, P., Meneghini, L., Runggaldier, W.J. (1996) Connections between stochastic control and dynamic games. Math. Control Signals Syst, 9, 303-326.