Note: Fokker-Planck equation of Langevin Dynamics

Atsushi Nitanda

December 5, 2022

We consider the following Langevin dynamics:

$$\mathrm{d}\theta_t = v(\theta_t)\mathrm{d}t + \sqrt{2}\mathrm{d}W_t, \tag{1}$$

where $v : \mathbb{R}^d \to \mathbb{R}^d$ is a smooth velocity and $(W_t)_{t\geq 0}$ is the standard Brownian motion on $\mathbb{R}^d$ with $W_0 = 0$. We recover Langevin dynamics when $v(\theta) = -\nabla f(\theta)$. Following Vempala and Wibisono (2019), we explain the evolution of the probability distribution $\rho_t(\theta)$ of θ_t can be described by the Fokker-Planck equation in weak sense.

$$\frac{\partial \rho_t}{\partial t} = -\nabla \cdot (\rho_t v) + \Delta \rho_t. \tag{2}$$

Let $\phi : \mathbb{R}^d \to \mathbb{R}$ be any smooth test function. Then, by (1), we get with $\eta > 0$,

$$\begin{aligned}
\theta_{t+\eta} &= \theta_t + \int_t^{t+\eta} v(\theta_s)\mathrm{d}s + \sqrt{2}(W_{t+\eta} - W_t) \\
&= \theta_t + \eta v(\theta_t) + \sqrt{2}(W_{t+\eta} - W_t) + O(\eta^2) \\
&\overset{\mathrm{d}}{=} \theta_t + \eta v(\theta_t) + \sqrt{2\eta} Z + O(\eta^2),
\end{aligned}$$

where $Z \sim \mathcal{N}(0, I)$ is independent of θ_t because $W_{t+\eta} - W_t \sim \mathcal{N}(0, I)$. Therefore,

$$\begin{aligned}
\phi(\theta_{t+\eta}) &\overset{\mathrm{d}}{=} \phi(\theta_t + \eta v(\theta_t) + \sqrt{2\eta} Z + O(\eta^2)) \\
&= \phi(\theta_t) + \eta \nabla\phi(\theta_t)^\top v(\theta_t) + \sqrt{2\eta}\nabla\phi(\theta_t)^\top Z + \eta Z^\top \nabla^2 \phi(\theta_t) Z + O(\eta^{3/2}).
\end{aligned}$$

Noting Z is independent of θ_t, we take the expectation of both sides:

$$\begin{aligned}
\int \phi(\theta)\rho_{t+\eta}(\theta)\mathrm{d}\theta &= \int \phi(\theta)\rho_t(\theta)\mathrm{d}\theta + \eta\mathbb{E}[\nabla\phi(\theta_t)^\top v(\theta_t) + Z^\top \nabla^2\phi(\theta_t) Z] + O(\eta^{3/2}) \\
&= \int \phi(\theta)\rho_t(\theta)\mathrm{d}\theta + \eta\mathbb{E}[\nabla\phi(\theta_t)^\top v(\theta_t)] + \eta\mathbb{E}[\Delta\phi(\theta_t)] + O(\eta^{3/2}),
\end{aligned}$$

where we used $\mathbb{E}[Z^\top \nabla^2 \phi(\theta_t) Z] = \mathbb{E}[\Delta\phi(\theta_t)]$. This implies

$$\begin{aligned}
\int \phi(\theta)\frac{\partial \rho_t}{\partial t}\mathrm{d}\theta &= \int \left(\nabla\phi(\theta)^\top v(\theta) + \Delta\phi(\theta)\right)\rho_t(\theta)\mathrm{d}\theta \\
&= \int \left(\phi(\theta)\nabla\cdot(\rho_t(\theta)v(\theta)) + \phi(\theta)\Delta\rho_t(\theta)\right)\mathrm{d}\theta.
\end{aligned}$$

Hence, we get eq. (2).

References

Vempala, S. and Wibisono, A. (2019). Rapid convergence of the unadjusted langevin algorithm: Isoperimetry suffices. *Advances in neural information processing systems*, 32.