12/4

Lecture 24: Diffusion models

Derivation of Fokker-Planck:

Consider SDE

$$dX_t = v_t(X_t)\,dt + \sqrt{2}\,dB_t$$

Let q_t be the density of X_t.

w.t.s.

$$\frac{\partial q_t}{\partial t} = -\operatorname{div}(q_t \cdot v_t) + \Delta q_t \qquad (\ast)$$

Take any smooth "test function" $\phi : \mathbb{R}^d \to \mathbb{R}$.

$$\frac{\partial}{\partial t} \mathbb{E}_{x \sim q_t}[\phi(x)] = \frac{\partial}{\partial t} \int \phi(x)\, q_t(x)\, dx$$

$$= \int \phi(x) \frac{\partial q_t}{\partial t}(x)\, dx \qquad \text{(LHS of } (\ast) \text{ integrated against } \phi\text{)}$$

w.t.s. this $= \int \phi(x) \left\{ -\operatorname{div}(q_t \cdot v_t) + \Delta q_t \right\}(x)\, dx$

We can also write

$$\frac{\partial}{\partial t} \mathbb{E}_{x \sim q_{r_t}} \left[\phi(x)\right] = \lim_{h \to 0} \frac{\mathbb{E}_{x \sim q_{r_{t+h}}}[\phi(x)] - \mathbb{E}_{x \sim q_{r_t}}[\phi(x)]}{h} \qquad (\star\star)$$

For $X_{t+h} \sim q_{r_{t+h}}$,

$$X_{t+h} = X_t + \int_t^{t+h} v_t(X_s)\, ds + \sqrt{2h} \cdot g \quad , \quad g \sim N(0, I_d)$$

$$= X_t + h\, v_t(X_t) + \sqrt{2h} \cdot g + O(h^2)$$

so by Taylor expansion,

$$\begin{aligned} \phi(X_{t+h}) = \phi(X_t) &+ \langle h\, v_t(X_t), \nabla \phi(X_t) \rangle \\ &+ \langle \sqrt{2h}\, g, \nabla \phi(X_t) \rangle \\ &+ \tfrac{1}{2} 2h\, g^T \nabla^2 \phi(x_t)\, g \\ &+ O(h^{3/2}) \end{aligned}$$

note: for the Gaussian term, we expand to the second derivative!

and thus

$$\mathbb{E}[\phi(X_{t+h})] = \mathbb{E}[\phi(X_t)] + h\,\mathbb{E}\big(\langle v_t(X_t), \nabla\phi(X_t)\rangle\big) + h\,\mathbb{E}\big[\underbrace{\mathrm{Tr}\,\nabla^2\phi(X_t)}_{\Delta\phi(X_t)}\big] + O(h^{3/2}),$$

So $(\spadesuit)$ yields

$$\frac{\partial}{\partial t}\mathbb{E}_{x\sim q_t}[\phi(x)] = \mathbb{E}\Big(\underbrace{\langle v_t(X_t), \nabla\phi(X_t)\rangle}_{\mathrm{I}} + \underbrace{\Delta\phi(X_t)}_{\mathrm{II}}\Big)$$

I: $\int \langle q_t(x)\, v_t(x), \nabla\phi(x)\rangle\, dx$

$\overset{\text{(int. by parts)}}{=} \int \boxed{-\mathrm{div}(q_t \cdot v_t)}(x)\cdot\phi(x)\, dx$

II: $\int q_t(x)\,\Delta\phi(x)\, dx$

self-adjointness of Laplacian, a.k.a integration by parts

$= \int \phi(x)\, \boxed{\Delta q_t(x)}\, dx$ $\square$

Can also write Fokker-Planck as

$$\frac{\partial}{\partial t} q_t = \operatorname{div}\left(q_t \left(\nabla \ln q_t - v_t\right)\right)$$

using $\quad \Delta q_t = \operatorname{div}\left(q_t \nabla \ln q_t\right)$

Can conclude that forward and reverse processes' Fokker-Planck eq.'s are time reversals of each other:

F.P. for forward SDE ($dx_t = -x_t\,dt + \sqrt{2}\,dB_t$):

$$\frac{\partial q_t}{\partial t}(x_t) = \boxed{\operatorname{div}\left(q_t \cdot \left(\nabla \ln q_t + x_t\right)\right)}$$

F.P. for reverse SDE ($dx^{\leftarrow}_t = \left\{x^{\leftarrow}_t + 2\nabla \ln q^{\leftarrow}_t(x^{\leftarrow}_t)\right\}dt + \sqrt{2}\,dB_t$)

$$\frac{\partial q^{\leftarrow}_t}{\partial t} = \operatorname{div}\left(q^{\leftarrow}_t \cdot \left(\nabla \ln q^{\leftarrow}_t - \left[x^{\leftarrow}_t + 2\nabla \ln q^{\leftarrow}_t\right]\right)\right)$$

$$= \boxed{-\operatorname{div}\left(q^{\leftarrow}_t \cdot \left(\nabla \ln q^{\leftarrow}_t + x^{\leftarrow}_t\right)\right)}$$

same up to sign, i.e. time reversal

Heuristic proof for Girsanov's:

$$dX_t = b_t\,dt + \sqrt{2}\,dB_t \qquad (1)$$

$$dX_t = b'_t\,dt + \sqrt{2}\,dB_t \qquad (2)$$

Consider discrete-time approx, i.e.

$$\hat{X}_{(k+1)h} \leftarrow \hat{X}_{kh} + h\cdot b_{kh}(\hat{X}_{kh}) + \sqrt{2h}\, g_{kh}, \quad g_{kh} \sim N(0, I_d)$$

$$\hat{X}_{(k+1)h} \leftarrow \hat{X}_{kh} + h\cdot b'_{kh}(\hat{X}_{kh}) + \sqrt{2h}\, g_{kh}$$

Likelihood of observing trajectory

$$(\hat{X}_0, \hat{X}_h, \hat{X}_{2h}, \ldots\ldots \hat{X}_{Nh})$$

under (1) vs (2):

$$(1): \quad \prod_{k=0}^{N-1} \exp\left(-\frac{\|\hat{X}_{(k+1)h} - \hat{X}_{kh} - h\cdot b_{kh}(\hat{X}_{kh})\|^2}{4h}\right)$$

$$(2): \quad \prod_{k=0}^{N-1} \exp\left(-\frac{\|\hat{X}_{(k+1)h} - \hat{X}_{kh} - h\cdot b'_{kh}(\hat{X}_{kh})\|^2}{4h}\right)$$

$$\frac{(1)}{(2)} = \prod_{k=0}^{N-1} \exp\left(-\frac{1}{4h}\left[\|b_{kh}(\hat{x}_{kh})\|^2 \cdot h^2 - \|b'_{kh}(\hat{x}_{kh})\|^2 \cdot h^2 - 2h\langle \underbrace{\hat{x}_{(k+1)h} - \hat{x}_{kh}}_{\substack{\text{under } (2), \\ \text{this } = h \cdot b'_{kh}(\hat{x}_{kh}) + \sqrt{2h}\, g_{kh}}}, b_{kh}(\hat{x}_{kh}) - b'_{kh}(\hat{x}_{kh})\rangle\right]\right)$$

$$= \prod_{k=0}^{N-1} \exp\left(-\frac{1}{4h}\left[-\|b_{kh}(\hat{x}_{kh}) - b'_{kh}(\hat{x}_{kh})\|^2 \cdot h^2 + h \cdot 2\sqrt{2}\,\langle \sqrt{h}\, g_{kh}, b_{kh}(\hat{x}_{kh}) - b'_{kh}(\hat{x}_{kh})\rangle\right]\right)$$

$$= \exp\left(\frac{-1}{4}\sum_{k=0}^{N-1} h\,\|b_{kh}(\hat{x}_{kh}) - b'_{kh}(\hat{x}_{kh})\|^2 + \frac{1}{\sqrt{2}}\sum_{k=0}^{N-1}\langle \underbrace{\sqrt{h}\, g_{kh}}_{\text{"}dB_{kh}\text{"}}, b_{kh}(\hat{x}_{kh}) - b'_{kh}(\hat{x}_{kh})\rangle\right)$$

$$\xrightarrow{h \to 0} \exp\left(-\frac{1}{4}\int_0^T \|b_t - b'_t\|^2\, dt + \frac{1}{\sqrt{2}}\int_0^T \langle dB_t, b_t - b'_t\rangle\right) \qquad \square$$

Movement bound for reverse process:

need to bound:

1) $\mathbb{E}\|X_r\|^2$

2). $\mathbb{E}_{q_{T-r}}\|\nabla \ln q_{T-r}(X_r)\|^2$

For 1), we have

$$X_r = e^{-r} X_0 + \sqrt{1-e^{-2r}}\, g \quad \text{for } g \sim N(0, I),$$

So $\mathbb{E}\|X_r\|^2 \lesssim \mathbb{E}\|X_0\|^2 + \mathbb{E}\|g\|^2$

$$= m_2^2 + d$$

For 2, we have

$$\mathbb{E}\|\nabla \ln q_{T-r}(X_r)\|^2$$

$$= \int \langle \nabla \ln q_{T-r}(x),\ q_{T-r}(x) \cdot \nabla \ln q_{T-r}(x) \rangle dx$$

$$= \int \langle \nabla q_{T-r}(x),\ \nabla \ln q_{T-r}(x) \rangle\, dx$$

(int. by parts)

$$= -\int q_{T-r}(x) \cdot \mathrm{div}\left(\nabla \ln q_{T-r}(x)\right) dx$$

$$= -\int q_{T-r}(x) \cdot \underbrace{\mathrm{Tr}\Big(\underbrace{\nabla^2 \ln q_{T-r}(x)}_{\|\cdot\|_{op} \le L}\Big)}_{\le Ld},$$

so $\mathbb{E}\,\|\nabla \ln q_{T-r}(x_r)\|^2 \le Ld$