Project and Exercises

Course “**Mathematical strategies for complex stochastic dynamics**”,

Summer 2025, FUB

June 13, 2025

1 Project

Select one of the following topics. Implement the code based on the relevant notebooks provided in the practice course. Conduct numerical experiments and study a few questions/issues related to the method and the numerical results. Write a report that includes the following parts:

1. A short introduction to the topic.
2. Description of the adopted method.
3. Details of numerical experiments (e.g. dataset, neural network architecture, training parameters such as learning rate and number of epochs...) and discussions of numerical observations.
4. Conclusion (e.g. capability, complexity of the method) and outlook (e.g. potential challenge for high-dimensional and large datasets)

A few questions are provided, but you are encouraged to consider other questions related to the method and numerical experiments that you find interesting.

To submit, please send by email the zip file containing both the report and code.

Topic 1. Train a diffusion model on one of the provided datasets in $\mathbb{R}^2$.

Below are a few questions to be answered.

1. Is there a way to quantitatively evaluate the quality of the numerical result (e.g. a number/score that measures the closeness between the generated samples and the training dataset)?
2. Compare training results with fewer or more training epochs. How many training epochs are required to obtain a good result?
3. Compare training results with different neural network architectures. What will happen when the neural network is linear (i.e. no nonlinear activation) or the sizes of layers are very small?
4. Compare training results with different choices of $\eta_{\min}, \eta_{max}$.
5. Compare the training results with different batch-sizes. How does the training (for example, computational time) depend on batch-sizes?

Topic 2. Train a flow-matching generative model on one of the provided datasets and study the following questions.

1. How is the quality of the numerical result?
2. Compare training results with fewer or more training epochs. How many training epochs are required to obtain a good result?
3. Compare training results with different neural network architectures. What will happen when the sizes of neural network layers are small or large?
4. Besides the linear interpolation $X_t = (1-t)X_0 + tX_1$, also implement the code with a different interpolant (and change the loss objective accordingly; see the lecture note), for example,
$$X_t = \big(1 - \sin(\frac{t\pi}{2})\big)X_0 + \sin\big(\frac{t\pi}{2}\big)X_1\,, \tag{1}$$
Compare the results to the results obtained using linear interpolation.
5. Optionally, compare flow-matching models to diffusion models (in terms of approaches, computational cost, quality of results).

2 Exercises

1. Consider the ODE in $\mathbb{R}^d$
$$\begin{aligned} \frac{dx(t)}{dt} =& f(x(t))\,, \quad t \in [0,T]\,, \\ x(0) =& x, \quad x \in \mathbb{R}^d\,. \end{aligned} \tag{2}$$
Assume that $x(0)$ is a random variable whose probability density is $p_0(x)$. Let $p(x,t)$ be the probability density of $x(t)$ at time $t \in [0,T]$.
   (a) Show that
   $$\frac{d}{dt}p(x(t),t) = -\mathrm{div} f(x(t))\, p(x(t),t), \quad t \in [0,T]\,.$$
   (b) Show that
   $$p(x(t),t) = p_0(x)\mathrm{e}^{-\int_0^t (\mathrm{div} f)(x(s))ds}.$$

   Hint: Recall that $p(x,t)$ solves the continuity equation
   $$\begin{aligned} &\frac{\partial p}{\partial t} + \mathrm{div}(fp) = 0\,, \quad \forall\, t \in (0,T] \\ &p(x,0) = p_0\,. \end{aligned}$$
2. Prove the identity
$$\int_0^t B_s^2 dB_s = \frac{1}{3}B_t^3 - \int_0^t B_s ds \tag{3}$$
Hint: Apply Ito's formula to B_t^3 and then use it to show that both sides satisfy the same SDE (and have the same value zero at $t = 0$).

3. Let B_t be a one-dimensional Brownian motion. Consider the SDE for $X_t = (X_t^1, X_t^2) \in \mathbb{R}^2$:

$$\begin{cases} dX_t^1 = -\frac{1}{2}X_t^1 dt - X_t^2 dB_t \\ dX_t^2 = -\frac{1}{2}X_t^2 dt + X_t^1 dB_t \end{cases} \tag{4}$$

   (a) Show that $|X_t|^2$ is constant, i.e. $|X_t| = |X_0|$ for all $t \geq 0$.

   (b) Show that $X_t = (\cos(B_t), \sin(B_t))$ solves (4).

   Hint: Apply Ito's formula.

4. Consider the Brownian dynamics

$$dX_t = -\nabla V(X_t)\, dt + \sqrt{2\beta^{-1}} dB_t\,, \quad t > 0 \tag{5}$$

   where $\beta > 0$ is a constant.

   (a) Write down the expression of the infinitesimal generator $\mathcal{L}$ of (5).

   (b) Let $\pi(x) = \frac{1}{Z}\mathrm{e}^{-\beta V(x)}$ be the density of Boltzmann distribution, where Z is a normalizing constant. Show that π is invariant under the dynamics (5).

   (c) Show that $\mathcal{L}$ is a self-adjoint operator with respect to the weighted inner product $\langle g_1, g_2 \rangle_\pi = \int_{\mathbb{R}^d} g_1 g_2 \pi dx$.

   Hint: See the lecture note.

5. Consider the Ornstein-Unlenbeck (OU) process

$$dX_t = -\kappa X_t\, dt + \sqrt{2\beta^{-1}} dB_t\,, \quad t > 0\,, \tag{6}$$

   where $\kappa, \beta > 0$, and X_0 is fixed.

   (a) Show that

$$X_t = \mathrm{e}^{-\kappa t} X_0 + \sqrt{2\beta^{-1}} \int_0^t \mathrm{e}^{-\kappa(t-s)} dB_s, \quad t > 0\,. \tag{7}$$

   (b) Compute the mean and variance of X_t.

   Hint: Apply Ito's formula to $\mathrm{e}^{\kappa t} X_t$ and then integrate to get (7). Use (7) and properties of Ito's integral (Ito isometry) to compute the mean and variance.

6. Consider the Brownian dynamics X_t in (5). Recall that the semigroup T_t is defined as $(T_t f)(x) = \mathbb{E}(f(X_t)|X_0 = x)$ for a test function. Denote $\mathbb{E}_\pi(f) = \int_{\mathbb{R}^d} f(x)\pi(x)dx$. Assume that its invariant density satisfies the so-called Poincáre inequality. That is, there exists a positive constant $\rho > 0$ such that the inequality

$$\mathbb{E}_\pi\left[(f(x) - \mathbb{E}_\pi f)^2\right] \leq \frac{1}{\rho}\mathbb{E}_\pi |\nabla f(x)|^2 \tag{8}$$

holds for all C^1-smooth function $f : \mathbb{R}^d \to \mathbb{R}$. Show that

$$\mathbb{E}_\pi\Big[\big((T_t f)(x) - \mathbb{E}_\pi(T_t f)\big)^2\Big] \le \mathrm{e}^{-\frac{2\rho t}{\beta}} \mathbb{E}_\pi\big[(f(x) - \mathbb{E}_\pi f)^2\big], \tag{9}$$

for any $t \ge 0$.

Hint: Derive a differential inequality for the left hand side of (9), by taking derivative and using

(a) the identity $\frac{d(T_t f)}{dt} = \mathcal{L} T_t f$,

(b) the identity $\mathbb{E}_\pi(f\mathcal{L}g) = -\frac{1}{\beta}\mathbb{E}_\pi(\nabla f \cdot \nabla g)$ for test functions f and g,

(c) the fact that $\mathbb{E}_\pi(\mathcal{L}T_t f) = 0$ (which is implied by (b)),

(d) the Poincáre inequality (8).

Integrate the inequality to conclude the proof.

7. Consider the Langevin dynamics

$$\begin{aligned} dQ_t =& P_t dt \\ dP_t =& -\nabla V(Q_t)\, dt - \gamma P_t dt + \sqrt{2\gamma\beta^{-1}} dB_t\,. \end{aligned} \tag{10}$$

Let $\pi(q,p) = Z_1^{-1}\mathrm{e}^{-\beta\left(V(q)+\frac{|p|^2}{2}\right)}$ be the invariant density of (10), where Z_1 is a normalizing constant. For two test functions $f, g : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$, denote by $\langle f, g\rangle_\pi$ the weighted scalar product with respect to π, i.e.

$$\langle f, g\rangle_\pi = \int_{\mathbb{R}^d\times\mathbb{R}^d} f(q,p)g(q,p)\pi(q,p)dqdp. \tag{11}$$

Let $\mathcal{L}^*$ be the adjoint of $\mathcal{L}$ with respect to (11), such that

$$\langle \mathcal{L} f, g\rangle_\pi = \langle f, \mathcal{L}^* g\rangle_\pi\,, \tag{12}$$

for two general test functions f, g. Let $\mathcal{T}$ denote the flip of momentum p. That is, given $f : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$, we have $(\mathcal{T} f)(q,p) = f(q,-p)$.

Show that

(a) $\mathcal{L}^* = \mathcal{T}\mathcal{L}\mathcal{T}$.

(b) $\langle \mathcal{L}\mathcal{T} f, g\rangle_\pi = \langle f, \mathcal{L}\mathcal{T} g\rangle_\pi$.

Hint:

(a) Write down the expression of the generator $\mathcal{L}$ (see lecture notes).

(b) To prove (a), derive the expression of $\mathcal{L}^*$ using the definition (12).

(c) Use (a) and (12) to prove (b).

8. Let $\xi : \mathbb{R}^d \to \mathbb{R}$ be a C^2-smooth map, which satisfies

(a) $\nabla\xi(x)$ is non-zero for all $x \in \mathbb{R}^d$.

(b) The image of ξ is $\mathbb{R}$, that is, for any $z \in \mathbb{R}$, there exist $x \in \mathbb{R}^d$ such that $\xi(x) = z$.

Let $f : \mathbb{R}^d \to \mathbb{R}$ be a C^2-smooth function. Define the function $\widetilde{f} : \mathbb{R} \to \mathbb{R}$ as

$$\widetilde{f}(z) = \int_{\mathbb{R}^d} f(x)\delta(\xi(x) - z)dx\,, \quad \forall z \in \mathbb{R}\,. \tag{13}$$

Derive an expression for $\widetilde{f}'(z)$ (the derivative of $\widetilde{f}$ with respect to z).

Hint:

(a) Compute the integral $\int_{\mathbb{R}} \widetilde{f}'(z)g(z)dz$ for a test function g, using integration by parts.

(b) Use the identity (the propertity of δ function; see lecture note)

$$\int_{\mathbb{R}} \widetilde{f}(z)g(z)dz = \int_{\mathbb{R}^d} f(x)g(\xi(x))dx\,.$$

(c) Use the fact (by chain rule):

$$\nabla g(\xi(x)) = g'(\xi(x))\nabla\xi(x) \Longrightarrow g'(\xi(x)) = \frac{\nabla g(\xi(x)) \cdot \nabla\xi(x)}{|\nabla\xi(x)|^2}\,.$$