Lecture 1: Introduction

2025–04–23

Content of this course:

1. Basic of stochastic processes (4 weeks)

   Langevin dynamics, Markov chains, generators, Fokker-Planck equation, convergence to equilibrium, Ito's formula

2. Model reduction for stochastic dynamics (4 weeks)

   averaging, collective variables, effective dynamics, Markov state modeling

3. Machine learning techniques (4 weeks)

   stochastic gradient descent, autoencoders, PDE eigenvalue problems by deep learning, diffusion models, continuous normalizing flow, flow-matching

Chapter 1

Introduction

1.1 Useful results from probability theory

1.1.1 Gaussian random variables

The probability density function of Gaussian random variables (RVs) X in $\mathbb{R}$ with mean x_0 and variance σ^2 is

$$p(x) = (2\pi\sigma^2)^{-\frac{1}{2}} \mathrm{e}^{-\frac{(x-x_0)^2}{2\sigma^2}}, \quad x \in \mathbb{R}. \tag{1.1}$$

We often write $X \sim \mathcal{N}(x_0, \sigma^2)$.

Proposition 1. Assume $X \sim \mathcal{N}(0, \sigma_1^2)$.

1. $\mathbb{E}(X) = \mathbb{E}(X^3) = 0$, $\mathbb{E}(X^2) = \sigma_1^2$, $\mathbb{E}(X^4) = 3\sigma_1^4$.
2. $\lambda X \sim \mathcal{N}(0, \lambda^2\sigma_1^2)$, for $\lambda > 0$.
3. Assume that $Y \sim \mathcal{N}(0, \sigma_2^2)$ and Y is independent of X. Then, we have $\lambda_1 X + \lambda_2 Y \sim \mathcal{N}(0, \lambda_1^2\sigma_1^2 + \lambda_2^2\sigma_2^2)$.

Similarly, for Gaussian RVs X in $\mathbb{R}^d$ with mean x_0 and co-variance $\Sigma \in \mathbb{R}^{d\times d}$, the probability density is

$$p(x) = (2\pi)^{-\frac{d}{2}} (\det \Sigma)^{-\frac{1}{2}} \mathrm{e}^{-\frac{1}{2}(x-x_0)^\top \Sigma^{-1}(x-x_0)}, \quad x \in \mathbb{R}^d. \tag{1.2}$$

1.2 ODEs

Let's recall some facts about ODEs. Consider ODE in $\mathbb{R}^d$

$$\begin{aligned} \frac{dx(t)}{dt} =& f(x(t)), \quad t \in [0, T], \\ x(0) =& x, \quad x \in \mathbb{R}^d, \end{aligned} \tag{1.3}$$

where the vector field $f : \mathbb{R}^d \to \mathbb{R}^d$ is Lipschitz.

Known results:

- Existence and uniqueness of the solution to (1.3) for finite time.

- The solution to (1.3) is C^1-smooth.

Given the initial state x, the system (1.3) is deterministic. Let us introduce the operator

$$(\mathcal{L}g)(x) = f(x) \cdot \nabla g(x), \qquad \forall g \in C^1(\mathbb{R}^d), \quad x \in \mathbb{R}^d . \tag{1.4}$$

Lemma 1. For any C^1 function $g : \mathbb{R}^d \times [0, T] \to \mathbb{R}$, we have

$$\frac{dg(x(t), t)}{dt} = \Big(\frac{\partial g}{\partial t} + \mathcal{L}g\Big)(x(t), t) \tag{1.5}$$

Proof. The proof is straightforward using chain rule, ODE (1.3), and (1.4). □

Numerical scheme
Let $\Delta t = \frac{T}{N}$. The explicit Euler scheme is:

$$x_{n+1} = x_n + f(x_n)\Delta t\,, n = 0, 1, 2, \cdots . \tag{1.6}$$

There are also more advanced schemes such as implicit schemes and Runge-Kutta methods, which have better accuracy and stability.

1.2.1 Gradient system

Consider the case where $f(x) = -\nabla V(x)$ for some smooth potential function $V : \mathbb{R}^d \to \mathbb{R}$. The ODE (1.3) becomes

$$\frac{dx(t)}{dt} = -\nabla V(x(t))\,, \tag{1.7}$$

which is often called a gradient flow. In this case, we have $\mathcal{L}^{\mathrm{grad}} g = -\nabla V \cdot \nabla g$.

Lemma 2. V is non-increasing under the ODE flow (1.7).

Proof.
$$\frac{dV(x(t))}{dt} = \nabla V(x(t)) \cdot \frac{dx(t)}{dt} = -|\nabla V(x(t))|^2 \le 0\,, \tag{1.8}$$
therefore $V(x(t))$ is non-increasing. □

Lemma 2 shows that V is a Lyapunov function of ODE (1.7). It suggests that $x(t)$ will approach to local minima of V as $t \to \infty$.

Example. Let $d = 1$.

1. quadratic potential: $V(x) = \frac{1}{2}x^2$. Hence, $f(x) = -x$ and ODE (1.7) becomes
$$\frac{dx(t)}{dt} = -x(t)\,, \tag{1.9}$$
whose solution is $x(t) = \mathrm{e}^{-t}x(0)$, which converges to 0 for all initial value $x(0)$.

2. double well potential: $V(x) = \frac{1}{4}(x^2-1)^2$. We have $f(x) = -x(x^2-1)$. For all initial values other than 0, the flow converges to one of the two local minima $x_{\text{left}} = -1$ and $x_{\text{right}} = 1$.

1.2.2 Hamiltonian system

Assume that the ODE (1.3) describes the equation of a physical system of m particles in $\mathbb{R}^3$. Denote by q and p the coordinates and momentum of these particles, where $q, p \in \mathbb{R}^{3m}$, and let $x = (q, p) \in \mathbb{R}^d$, with $d = 6m$. Let $H : \mathbb{R}^{3m} \times \mathbb{R}^{3m} \to \mathbb{R}$ be a C^1 smooth function.

$$\begin{aligned} \frac{dq(t)}{dt} &= \nabla_p H \\ \frac{dp(t)}{dt} &= -\nabla_q H \,. \end{aligned} \tag{1.10}$$

In this case, we have $\mathcal{L}^{\text{Ham}} g = \nabla_p H \cdot \nabla_q g - \nabla_q H \cdot \nabla_p g$.

Lemma 3. H is a conserved quantity under the ODE (1.10).

Proof.

$$\frac{dH(q(t), p(t))}{dt} = \mathcal{L}^{\text{Ham}} H = \nabla_q H \cdot \nabla_p H - \nabla_p H \cdot \nabla_q H = 0\,,$$

which shows that H is conserved. □

In particular, choose

$$H(q, p) = V(q) + \frac{1}{2} p^\top M^{-1} p\,. \tag{1.11}$$

where $M \in \mathbb{R}^{3m \times 3m}$ is a constant positive definite mass matrix. Then, (1.10) becomes

$$\begin{aligned} \frac{dq(t)}{dt} &= M^{-1} p(t) \\ \frac{dp(t)}{dt} &= -\nabla_q V(q(t))\,. \end{aligned} \tag{1.12}$$