Score Matching and SDE

Chia Yu Ou

October 2025

1. Deriving the Sliced Score Matching (SSM) Loss Expression

We aim to show that the Sliced Score Matching (SSM) loss can be written as:

$$L_{\text{SSM}} = \mathbb{E}_{x\sim p(x)}\mathbb{E}_{v\sim p(v)}\left[\left\|v^T S(x;\theta)\right\|^2 + 2v^T\nabla_x\left(v^T S(x;\theta)\right)\right]$$

We start from the Implicit Score Matching (ISM) loss:

$$L_{\text{ISM}}(\theta) = \mathbb{E}_{x\sim p(x)}\left[\|S_\theta(x)\|^2 + 2\cdot\mathrm{tr}\left(\nabla_x S_\theta(x)\right)\right]$$

Now, recall the trace identity using isotropic Gaussian vectors $v \sim \mathcal{N}(0, I)$:

$$\mathbb{E}_{v\sim\mathcal{N}(0,I)}[v^T A v] = \mathrm{tr}(A)$$

Therefore, the trace term can be rewritten as:

$$\mathrm{tr}\left(\nabla_x S_\theta(x)\right) = \mathbb{E}_{v\sim\mathcal{N}(0,I)}\left[v^T\nabla_x S_\theta(x) v\right]$$

The first norm term can also be written as:

$$\|S_\theta(x)\|^2 = \mathbb{E}_{v\sim\mathcal{N}(0,I)}\left[\|v^T S_\theta(x)\|^2\right]$$

Combining both terms, we obtain:

$$L_{\text{SSM}}(\theta) = \mathbb{E}_{x\sim p(x)}\mathbb{E}_{v\sim\mathcal{N}(0,I)}\left[\|v^T S_\theta(x)\|^2 + 2v^T\nabla_x(v^T S_\theta(x))\right]$$

This completes the derivation.

2. Briefly Explain SDE (Stochastic Differential Equation)

A standard ordinary differential equation (ODE) is of the form:

$$\frac{dy(t)}{dt} = f(t, y(t))$$

An SDE extends this by adding a stochastic term:

$$dy(t) = f(t, y(t))\,dt + g(t, y(t))\,dW(t)$$

where $W(t)$ is a Wiener process (Brownian motion), representing randomness over time.

Key Points:

- An SDE models a system whose change over time is affected not only by deterministic dynamics but also by stochastic noise.
- The solution to an SDE is a **stochastic process**, meaning that the outcome at each time step is a probability distribution.
- Analogous to how we solve ODEs numerically using Euler's method, SDEs can be solved using the Euler–Maruyama method.

Intuition:

SDEs are widely used in fields such as physics, finance, and generative modeling, where systems evolve with both predictable and unpredictable components. In score-based generative models, SDEs allow us to define continuous-time diffusion processes for sampling.