Problem 1 — Sliced Score Matching (SSM)

Goal: Show that

$$L_{\text{SSM}}(\theta) = \mathbb{E}*x, v!\left[|v^\top S(x;\theta)|^2 + 2, v^\top \nabla_x (v^\top S(x;\theta))\right],$$

where $S(x;\theta) = \nabla_x \log p*\theta(x)$

Definition

Start from the *sliced Fisher divergence* :

$$L_{\text{SSM}}(\theta) = \tfrac{1}{2}\mathbb{E}_{x,v}!\left(v^\top \nabla_x \log p(x) - v^\top S(x;\theta)\right)^2,$$

with #v# a random direction

Expand & Drop Constants

$$L_{\text{SSM}} \doteq \mathbb{E}_{x,v}!\left[\tfrac{1}{2}(v^\top S)^2 - (v^\top \nabla_x \log p)(v^\top S)\right].$$

First term of $p(x)$ is constant → omit

Integration by Parts

$$\mathbb{E}*p(x)[(v^\top \nabla_x \log p), g(x)] = -\mathbb{E}*p(x)[v^\top \nabla_x g(x)],$$

where $g(x) = v^\top S(x;\theta)$
Substitute:

$$L_{\text{SSM}} \doteq \mathbb{E}_{x,v}!\big[\tfrac{1}{2}(v^\top S)^2 + v^\top \nabla_x (v^\top S)\big].$$

Multiply by 2 and ignore constants → final form as in assignment.

References

- Hyvärinen (2005), *Score Matching*
- Yang Song & S. Ermon (2019), *Sliced Score Matching* arXiv:1905.07088

Problem 2 — Stochastic Differential Equations (SDE)

Definition

A stochastic differential equation describes dynamics under deterministic **drift** and stochastic **diffusion**:

$$dx_t = f(x_t, t), dt + G(x_t, t), dW_t, \qquad x(0) = x_0.$$

- $x_t \in \mathbb{R}^d$: state variable
- $f(x_t, t)$: drift term
- $G(x_t, t)$: diffusion matrix
- W_t: standard Brownian motion

Itô form:

$$x_t = x_0 + \int_0^t f(x_s, s), ds + \int_0^t G(x_s, s), dW_s.$$

Special cases

- *Pure drift : $dx_t = f(x_t, t)dt$ → deterministic ODE
- *Pure diffusion : $dx_t = G(x_t, t)dW_t$ → Brownian-like motion

Wiener Process W_t

A standard Brownian motion satisfies:

1. $W_0 = 0$
2. Stationary Gaussian increments $W_{t+\Delta} - W_t \sim \mathcal{N}(0, \Delta I)$
3. Independent increments
4. Continuous paths with probability 1

Moments:

$$\mathbb{E}[W_t] = 0, \quad \mathrm{Var}(W_t) = t$$

Noise $h(t) = \frac{dW_t}{dt}$ is formally its derivative

Euler–Maruyama Approximation

Partition ([0,T]) into (N) steps (\Delta t=T/N):

$$X_{n+1} = X_n + f(X_n, t_n)\Delta t + G(X_n, t_n)\Delta W_n,$$

where $\Delta W_n \sim \mathcal{N}(0, \Delta t, I)$

Problem 3 — Questions

- How does the SSM approximation behave in high-dimensional manifolds?