Assignment_8

(1) Sliced Score Matching (SSM) loss

步驟 1：從「投影到一維」開始

給定一個隨機方向向量 v（例如單位向量或來自某分布 p(v)），考慮資料在該方向的投影：

$$u = \mathrm{v}^{\mathrm{T}} x$$

對應的邊際密度（投影後的一維密度）記作 $\mathrm{p_v(u)}$，模型的投影 score 定義為

$$\mathrm{S_v}(\mathrm{u};\theta) = \frac{\partial}{\partial \mathrm{u}} \log \mathrm{p}_{\theta,\mathrm{v}}(\mathrm{u})$$

但由鏈式法則，對於原空間的模型 $\mathrm{score}\ \mathrm{S}(\mathrm{x},\theta) = \nabla_{\mathrm{x}} \log \mathrm{p}_\theta(\mathrm{x})$，有

$$\mathrm{S_v}(\mathrm{u};\theta) \ = \ \frac{\partial}{\partial \mathrm{u}} \log \mathrm{p}_\theta(\mathrm{v^T x}) = \mathrm{v^T} \nabla_{\mathrm{x}} \log \mathrm{p}_\theta(\mathrm{x}) = \mathrm{v^T S}(\mathrm{x};\theta)$$

所以投影後的一維 score 就等於 $\mathrm{v^T S}(\mathrm{x};\theta)$。

步驟 2：對一維 score matching 使用 Hyvärinen identity

一維的 score matching loss（對投影變量 u）可寫為（採 Hyvärinen 表示）:

$$\mathrm{L_{SM}^{(1D)}}(\mathrm{v}) = \mathrm{E}_{\mathrm{u}\sim \mathrm{p_v(u)}}[\frac{1}{2}\mathrm{S_v}(\mathrm{u};\theta)^2 + \frac{\partial}{\partial \mathrm{u}}\mathrm{S_v}(\mathrm{u};\theta)]$$

這個等號來自對一維 Hyvärinen identity（用積分分部把原本含未知真實 score 的平方差轉成含 model score 與其一維導數的式子）。

步驟 3：把一維期望換回對原始 x 的期望

注意 $\mathrm{u} = \mathrm{v}^{\mathrm{T}}\mathrm{x}$ 是由 x 決定的，因此可以把對 u 的期望換成對 x 的期望：

$$\mathrm{E}_{\mathrm{u}\sim \mathrm{p}_{\mathrm{v}}(\mathrm{u})}[\cdot] = \mathrm{E}_{\mathrm{x}\sim \mathrm{p}(\mathrm{x})}[\cdot]\ (\textit{內項僅透過}\, v^{\mathrm{T}}x).$$

代入$S_v = v^T$ 並用鏈式法則把 $\frac{\partial}{\partial u}$ 轉成 $v^T\nabla_x$，得到

$$L_{SM}^{(1D)}(v) \;=\; E_{x\sim\, p(x)}!\left[\,\frac{1}{2}\mathrm{big}\big(v^T S(x;\theta)\big)^2 \;+\; v^T\nabla_x\big(v^T S(x;\theta)\big)\right].$$

步驟 4：對方向 v 取期望 → Sliced SM

Sliced score matching 是對多個隨機方向取平均，也就是再對 $v\sim$

$p(v)v\sim p(v)$ 取期望：

$$L_{SSM} = E_{v\sim p(v)}\left[\,L_{SM}^{(1D)}(v)\,\right] = E_{x\sim p(x)}E_{v\sim p(v)}\left[\frac{1}{2}(v^TS)^2 + v^T\nabla_x(v^TS)\right].$$

等價於:

$$L_{SSM} = E_{v\sim p(v)}\left[\,L_{SM}^{(1D)}(v)\,\right] = E_{x\sim p(x)}E_{v\sim p(v)}[(v^TS)^2 + 2v^T\nabla_x(v^TS)].$$

(2) Briefly explain SDE

SDE（Stochastic Differential Equation, 隨機微分方程）

是一種描述系統隨時間演化且受隨機噪聲影響的方程：

$$dx = f(x,t)\,dt + g(x,t)\,dW_t,$$

其中

$f(x,t)$：漂移項 (drift term)，表示平均變化率；

- $g(x,t)$：擴散項 (diffusion term)，表示隨機擾動的強度；
- W_t：Wiener process（布朗運動）。

在 **score-based generative models** 中，SDE 用來模擬加噪與去噪過

程，例如：

- **前向 SDE**：逐漸加入噪聲使資料分布趨近高斯；
- **反向 SDE**：利用學得的 score function 逐步去噪生成資料。

(3) Unanswered question:

如果在 SDE 中加入 Poisson 跳躍過程（Jump process），模型該如何改寫？

在高維系統中，擴散矩陣的相關性 如何影響隨機動態？