Assignment 8

2025 年 11 月 25 日

1 Sliced Score Matching (SSM) 損失函數推導

我們希望證明 Sliced Score Matching (SSM) 損失函數可以寫成以下形式：

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

其中，SSM 損失是由 Implicit Score Matching (ISM) 損失改寫而來：

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

$S(x;\theta) = \nabla_x \log p(x;\theta)$ 是我們希望訓練的模型分數函數。我們將 L_{ISM} 的兩部分分別利用 Hutchinson's Trace Estimator 進行轉換。

1.1 第一部分：$\|S(x;\theta)\|^2$ 的轉換

假設投影向量 v 服從標準高斯分佈 $v \sim \mathcal{N}(0, I)$，則 $\mathbb{E}_{v\sim p(v)}[vv^\top] = I$。因此，對於範數項 (Norm Term) $||S(x;\theta)||^2$：

$$\begin{aligned}
\|S(x;\theta)\|^2 &= S(x;\theta)^\top S(x;\theta) \\
&= S(x;\theta)^\top \left(\mathbb{E}_{v\sim p(v)}[vv^\top]\right) S(x;\theta) \quad (\text{因爲}\mathbb{E}[vv^\top] = I) \\
&= \mathbb{E}_{v\sim p(v)}\left[S(x;\theta)^\top vv^\top S(x;\theta)\right] \\
&= \mathbb{E}_{v\sim p(v)}\left[(v^\top S(x;\theta))^\top (v^\top S(x;\theta))\right] \\
&= \mathbb{E}_{v\sim p(v)}\left[\|v^\top S(x;\theta)\|^2\right]
\end{aligned}$$

將期望 $\mathbb{E}_{x\sim p(x)}$ 帶回，得到 ISM 損失的第一部分轉換結果：

$$\mathbb{E}_{x\sim p(x)}\|S(x;\theta)\|^2 = \mathbb{E}_{x\sim p(x)}\mathbb{E}_{v\sim p(v)}\|v^\top S(x;\theta)\|^2$$

1.2 第二部分：$2\nabla_x \cdot S(x;\theta)$ 的轉換

第二部分是散度項 $\nabla_x \cdot S(x;\theta)$，其定義爲 $S(x;\theta)$ 的 Jacobian 矩陣 $\nabla_x S(x;\theta)$ 的跡 (Trace)：

$$\nabla_x \cdot S(x;\theta) = \sum_{i=1}^{d} \frac{\partial S_i}{\partial x_i} = \operatorname{trace}\big(\nabla_x S(x;\theta)\big)$$

利用 Hutchinson's Trace Estimator，散度可以被估計爲：

$$\nabla_x \cdot S(x;\theta) = \mathrm{trace}\big(\nabla_x S(x;\theta)\big) = \mathbb{E}_{v\sim p(v)}\left[v^\top \big(\nabla_x S(x;\theta)\big) v\right]$$

接下來，我們利用向量微積分的性質：

$$v^\top \big(\nabla_x S(x;\theta)\big) v = v^\top \nabla_x \big(v^\top S(x;\theta)\big)$$

因此，散度可以寫爲：

$$\nabla_x \cdot S(x;\theta) = \mathbb{E}_{v\sim p(v)}\left[v^\top \nabla_x \big(v^\top S(x;\theta)\big)\right]$$

將期望 $\mathbb{E}_{x\sim p(x)}$ 帶回，得到 ISM 損失的第二部分轉換結果：

$$\mathbb{E}_{x\sim p(x)}\big[2\nabla_x \cdot S(x;\theta)\big] = \mathbb{E}_{x\sim p(x)}\mathbb{E}_{v\sim p(v)}\big[2\, v^\top \nabla_x \big(v^\top S(x;\theta)\big)\big]$$

1.3 最終結論

將第一部分與第二部分的結果相加，即得到 SSM 損失的等價形式：

$$L_{\mathrm{SSM}} = \mathbb{E}_{x\sim p(x)}\mathbb{E}_{v\sim p(v)}\Big[\|v^\top S(x;\theta)\|^2 + 2\, v^\top \nabla_x \big(v^\top S(x;\theta)\big)\Big].$$

2 隨機微分方程式 (SDE) 簡介

2.1 基本概念

隨機微分方程式 (Stochastic Differential Equation, SDE) 是在傳統微分方程中加入隨機噪聲項的一種方程。它用於描述受不確定性或隨機波動影響的動態系統。

2.2 一般形式

SDE 的一般形式爲：

$$dx_t = \underbrace{f(x_t,t)}_{\text{漂移項 (Drift Term)}}\, dt + \underbrace{G(x_t,t)}_{\text{擴散項 (Diffusion Term)}}\, dW_t, \quad x(0) = x_0,$$

其中：

- x_t：** 系統狀態 ** (隨機過程)，代表 SDE 的解。
- $f(x_t,t)$：** 漂移項 **，描述系統的 ** 確定性趨勢 ** 或平均方向。
- $G(x_t,t)$：** 擴散項 **，控制隨機噪聲的強度和形式。
- dW_t：** 維納過程 (Wiener Process)** 的微分，代表 ** 隨機噪音 **，也稱爲布朗運動 (Brownian Motion)。

SDE 的求解通常需要使用 ** 伊藤積分 (Itô Calculus)**。

2.3 實例：幾何布朗運動 (Geometric Brownian Motion)

股價的隨機運動 (Geometric Brownian Motion) 是一個常見的 SDE 實例：

$$dS_t = \mu S_t dt + \sigma S_t dW_t, \quad S(0) = S_0$$

其中：

- μ：股價的平均增長率 (drift coefficient)。
- σ：股價的波動率 (volatility, diffusion coefficient)。

如果忽略隨機性 (即 $\sigma = 0$)，股價將按照確定性指數增長 $S_t = S_0 e^{\mu t}$。加入隨機性後，股價的軌跡變得不可預測，但在趨勢 μ 周圍隨機波動。

3 待討論問題

1. 噪音影響長期趨勢的例子：大部分 SDE 的長期趨勢會受到漂移項 ($f(x_t, t)$) 影響，有沒有會受到擴散項 ($G(x_t, t)$ 帶來的噪音) 影響的例子？
2. 布朗運動對不同 SDE 系統的長期影響差異：布朗運動 W_t 對於不同結構的 SDE 系統，其長期行爲 (如穩定性、極限分佈等) 有什麼差異？