1. Show that the sliced score matching (SSM) loss can also be written as

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

已知 $L_{ISM} = E_{x\sim p(x)}\{\|S(x;\theta)\|^2 + 2\nabla_x\cdot S(x;\theta)\}$. $x\in\mathbb{R}^n$

其中 $\nabla_x\cdot S(x;\theta) = \mathrm{div}\, S(x;\theta)$ 在每次訓練更新參數時,要計算 n 個導數,太貴

因此改用只沿著隨機方向 v 去求導,每次更新只須計算 1 個方向導數

Hutchinson Trace Estimation :　　計算矩陣 vv^T 每個 component 的期望值

Let $v\in\mathbb{R}^n$ be a random vector such that $E[vv^T]=I_n$,

then $\mathrm{tr}(A) = E_{v\sim p(v)}(v^T A v)$

Taking $A=\nabla_x S(x;\theta)$, then by Hutchinson Trace Estimation,

$$\mathrm{tr}(\nabla_x S(x;\theta)) = E_{v\sim p(v)}(v^T\nabla_x S(x;\theta)\, v) \quad \text{if } v \text{ satisfies } E[vv^T]=I_n$$

$$\because\ \mathrm{tr}(\nabla_x S(x;\theta)) = \mathrm{div}\, S(x;\theta) = \nabla_x\cdot S(x;\theta)$$

and $$E_{v\sim p(v)}[v^T\nabla_x S(x;\theta)^T v] = E_{v\sim p(v)}[v^T\nabla_x(v^T S(x;\theta))]$$

$$\begin{aligned}\because\ \|S(x;\theta)\|^2 &= S(x;\theta)^T S(x;\theta)\\ &= S(x;\theta)^T E_{v\sim p(v)}[vv^T]\, S(x;\theta)\\ &= E_{v\sim p(v)}[S(x;\theta)^T vv^T S(x;\theta)]\\ &= E_{v\sim p(v)}[\|v^T S(x;\theta)\|^2]\end{aligned}$$

Hence, $$\begin{aligned}L_{SSM} &= E_{x\sim p(x)}\{\|S(x;\theta)\|^2 + 2\nabla_x\cdot S(x;\theta)\}\\ &= E_{x\sim p(x)}E_{v\sim p(v)}\{\|v^T S(x;\theta)\|^2 + 2v^T\nabla_x(v^T S(x;\theta))\}\end{aligned}$$

2. Briefly explain SDE.

SDE:

$$dX_t = f(X_t,t)\,dt + G(X_t,t)\,dW_t$$

是一個符號化的表示（如同ODE中用 $dx=f(x,t)dt$ 去描述 $\frac{dx}{dt}=f(x,t)$，並不是嚴謹的定義），真正嚴格的定義是

$$X_t = X_0 + \overbrace{\int_0^t f(X_s,s)\,ds}^{\text{固定的函數}} + \overbrace{\int_0^t g(X_s,s)\,dW_s}^{\text{是一隨機過程}}$$

其中第一個積分是 Lebesgue integral，第二個積分是 Itô integral (stochastic integral 的一種)

在SDE中，

- X_t 表示一隨機過程
  (隨機過程應是要寫成 $\{X_t\}_{t\in T}$，是以時間為 index 的隨機變數所成的集合。但就如在ODE中，方程 $\frac{dx}{dt}=f(x,t)$ 的解 x 我們會用在 t 時間的值 $x(t)$ 來描述；因此，在SDE中也直接用 X_t 來代表整個過程)

- dt 表示微小時間的增量

- dX_t 表示 X_t 在微小時間 dt 內的隨機增量
  (此 d 並不是求導，而是根據 Itô integral 定義出的隨機增量符號)

- W_t 表示布朗運動 (Wiener process)，也是一種隨機過程，可理解為隨機擾動隨時間累積的結果，滿足

  $$W_0 = 0$$
  $$W_t - W_s \sim N(0, t-s)$$

  等性質

• dW_t 表示 W_t 在微小時間 dt 內的隨機增量

Remark：SDE 的解 X_t 是一個隨機過程

$$X : \Omega \times T \longrightarrow \mathbb{R},$$

$$X(\omega, t) := X_t(\omega)$$

並不是一個隨機變數(函數)

當 t is fixed，X_t 是隨機過程中的一個隨機變數

當樣本 ω is fixed，$X_t(\omega)$ 是一個時間的函數

透過

$$X_t = X_0 + \int_0^t f(X_s, s)\,ds + \underbrace{\int_0^t g(X_s, s)\,dW_s}_{\text{是一隨機過程}}$$

可看出隨機過程 X_t 是初值 X_0 + 固定的值 $\int_0^t f(X_s, s)\,ds$ + Wiener process

不同的隨機樣本 ω，則會有不同的布朗運動路徑 $W_t(\omega)$，因此則會對應到不同的解曲線 $X_t(\omega)$ (ω 固定，t 在變)

∴ SDE 找出的不是單一的函數，而是一族隨機生成的時間函數

$$\{\, t \longmapsto X_t(\omega) \mid \omega \in \Omega \,\}$$

3. Unanswered Questions

   There are unanswered questions from the lecture, and there are likely more questions we haven't covered.

- Take a moment to think about these questions.
- Write down the ones you find important, confusing, or interesting.
- You do **not** need to answer them—just state them clearly.

Q1：什麼是 Itô integral？

Q2：對於上課說的不同的初值 x_0 在某固定時間 t 時的值所符合的 pdf 是什麼意義？（還是不清楚）
（似乎與隨機動力系統有關？）

Q3：SDE 背後完整的理論架構是什麼？