

Assignment 10

1.

Consider the stochastic differential equation (SDE):

$$dx_t = f(x_t, t)dt + g(x_t, t)dW_t,$$

and let $p(x, t)$ be the marginal density of x_t . The Fokker-Plank equation associated with this SDE is

$$\frac{\partial}{\partial t} p(x, t) = -\frac{\partial}{\partial x} [f(x, t)p(x, t)] + \frac{1}{2} \frac{\partial^2}{\partial x^2} [g^2(x, t)p(x, t)].$$

This can be rewritten in a divergence (continuity) form:

$$\frac{\partial}{\partial t} p(x, t) = -\frac{\partial}{\partial x} \left[f(x, t)p(x, t) - \frac{1}{2} \frac{\partial}{\partial x} [g^2(x, t)p(x, t)] \right].$$

For a time-dependent density $p(x, t)$, we define the probability flow $J(x, t)$ (also called the probability current) as the quantity satisfying the continuity equation:

$$\frac{\partial}{\partial t} p(x, t) = -\frac{\partial}{\partial x} J(x, t).$$

Comparing with the divergence form, we identify:

$$J(x, t) = f(x, t)p(x, t) - \frac{1}{2} \frac{\partial}{\partial x} [g^2(x, t)p(x, t)].$$

We can define the probability-flow velocity field $v(x, t)$ such that

$$v(x, t) := \frac{J(x, t)}{p(x, t)}.$$

Substituting $J(x, t)$:

$$\begin{aligned} v(x, t) &= \frac{J(x, t)}{p(x, t)} = f(x, t) - \frac{1}{2} \frac{\partial}{\partial x} \frac{[g^2(x, t)p(x, t)]}{p(x, t)} \\ &= f(x, t) - \frac{1}{2} \frac{\left[\frac{\partial}{\partial x} g^2(x, t) \right] p(x, t) + g^2(x, t) \left[\frac{\partial}{\partial x} p(x, t) \right]}{p(x, t)} \\ &= f(x, t) - \frac{1}{2} \frac{\partial}{\partial x} g^2(x, t) - \frac{1}{2} g^2(x, t) \left[\frac{\partial}{\partial x} \log p(x, t) \right] \end{aligned}$$

If the flow velocity of the particles is $v(x, t)$, then the deterministic trajectories (streamlines) satisfy the ordinary differential equation (ODE):

$$\frac{dx_t}{dt} = v(x_t, t),$$

or equivalently,

$$dx_t = v(x_t, t) dt = \left[f(x_t, t) - \frac{1}{2} \frac{\partial}{\partial x} g^2(x_t, t) - \frac{g^2(x_t, t)}{2} \frac{\partial}{\partial x} \log p(x_t, t) \right] dt.$$

2. AI 的未來與機器學習的基石

1. AI 的未來能力

我認為最具科學意義、而目前 AI 還完全做不到的能力是「AI 能自主提出新的數學定理並給出可正式驗證的證明」。

現在的模型頂多能模仿證明的語氣，但在嚴格邏輯、抽象概念或跨領域結構上仍非常脆弱。因此，若 AI 要做到「真正的數學創造」至少需要三項能力：

第一，能從大量定理與證明中提取結構性模式。

AI 應該能分析形式化資料庫與自然語言論文，理解定理與引理的依賴關係、常見證明策略，並建立類似數學家心中的「概念結構圖」。

第二，AI 需要能自己提出猜想（Conjecture Generation）。

例如找出群或環中新的恆等式、提出新的拓樸不變量，或發現某些矩陣群的特殊性質。這些必須是「真正的新命題」，而不是教材內容或細微變形。

第三，生成可被 proof assistant 驗證的證明。

這點是區分「像證明」與「真正證明」的關鍵。如果 AI 能在 Lean、Coq 或 Isabelle 上寫出可逐步檢查的證明，就表示它具備真正的邏輯能力，而不只是語言生成。

數學本身是許多科學的基礎，因此如果 AI 能創造全新的數學定理，將可能影響量子資訊、演算法、密碼學等多個領域，帶來深層的理論性發展。

2. 涉及的機器學習類型

AI 如果要達到這個能力，它必須具備兩個核心能力：

- 理解數學結構（抽象概念、模式、定理之間的關係）
- 在巨大證明空間中探索有效證明路徑

其中「非監督式學習」負責第一項，「強化學習」負責第二項。

為什麼需要「非監督式學習」？

- 數學資料沒有標籤（沒人會標示「這個引理重要」、「這裡要用某技巧」）。
- AI 必須靠自己從定理、證明、定義中找出「抽象結構」與「概念關係」。
- 能讓 AI 自動建立數學的 Internal Representation（概念空間）。

為什麼需要「強化學習」？

- 證明是一個巨大的搜尋問題：每一步推理都是一個 Action。
- Proof Assistant 會回饋每一步是否有效。
- RL 利用這些回饋來引導 AI 找到成功的證明路徑。

該任務中的「資料來源」是什麼？

- 形式化證明庫（Lean mathlib、Coq libraries）
- 數學論文、教科書
- 定義/定理的依賴圖（Proof Graph）
- 證明的語法樹（AST）

該任務中的「目標訊號」是什麼？

- 非監督式學習：從資料本身萃取的數學結構、定理的依賴關係等。
- 強化學習：Proof Assistant 提供的回饋（合法/不合法、是否接近完成）

是否存在環境互動？

有。環境就是 Proof Assistant (Lean、Coq)，AI 每一步推理都會得到回饋，形成完整的 RL 互動循環，這讓 AI 能透過試錯學習證明策略。

3. 第一步的「模型化」

模型問題：讓 AI 在一個簡單的代數結構（例如群或環）中，自動找出一個未被記載的恆等式，並使用 Proof Assistant 完成形式化證明。

如何代表理想中的最終能力？

這個問題雖然比真正「提出新定理」簡單，但流程完全相同：
AI 必須理解結構 → 生成猜想 → 搜尋證明 → 通過形式化驗證。
因此它可以被視為未來能力的縮小版測試。

它的可測試性（如何知道模型成功？）

模型是否成功非常容易判斷：

- 產生的恆等式能被 Lean/Coq 嚴格驗證
- 不是教材中已有的例子（非平凡性）
- 在不同例子中都成立（可泛化）

需要哪些數學或機器學習工具？

- 數學：基本代數（群、環）、形式化語言（Lean/Coq）
- 機器學習：非監督式學習（建立概念表示）、強化學習（探索證明路徑）、
符號推理模型（生成證明步驟）、資訊理論工具（判斷恆等式的簡潔性）