深層生成モデルの学習安定性に関する代数的構造理論

—— データ分布に内在するフォン・ノイマン環の型による安定性の完全分類 ——

吉田英樹

2025 年 10 月 12 日

概要

本論文は、深層生成モデルにおける学習ダイナミクスの安定性、特に敵対的生成ネットワーク（GAN）が示す不安定性と、確率的拡散モデルが示す顕著な安定性という二元的な現象に対し、統一的な数学的説明を与えることを目的とする。この目的を達成するため、本研究ではデータ分布とその内在的対称性から構成されるフォン・ノイマン環の構造に着目する。

具体的には、データが居住する確率測度空間 (X, μ) と、その上の測度を保つ離散群 Γ の作用から、Murray-von Neumann の群-測度空間構成法を用いてフォン・ノイマン環 $\mathcal{M} := L^\infty(X, \mu) \rtimes \Gamma$ を定義する。本論文の核心的主張は、この作用素環の代数的な型（I, II, III 型）が、学習ダイナミクスの安定性を完全に分類することを、幾何学的・解析学的知見と接続し、厳密な証明をもって示すことにある。

主定理として、(i) 拡散モデルに見られるような構造的に安定な学習過程は、作用素環 $\mathcal{M}$ が正規忠実な有限トレースを持つ II_1 型因子環となる場合に特徴づけられること、そして (ii) GAN のモード崩壊に代表される構造的に不安定な学習過程は、作用素環 $\mathcal{M}$ がいかなる半有限トレースも持たない III 型因子環となる場合に必然的に発生することを証明する。これにより、これまで現象論的に理解されてきた安定性の問題が、データ分布に内在する対称性のエルゴード理論的な性質によって支配される、不変な代数的構造の問題に帰着されることを明らかにする。

目次

1 序論

1.1 問題の背景と研究動機

近年、敵対的生成ネットワーク（GAN）や確率的拡散モデルといった深層生成モデルは、高品質なサンプル生成能力において目覚ましい成功を収めている。しかし、その学習ダイナミクスの安定性には顕著な対比が見られる。

一方で、GAN の学習過程は本質的に不安定であることが広く知られている。この不安定性は、生成器がデータ分布の特定モードしか捉えられなくなる「モード崩壊」や、損失関数が収束せずに振動し続けるといった形で現れる。この不安定性の一因が、確率測度空間（ワッサースタイン空間）のリーマン曲率 $\mathcal{R}^{\mathcal{P}}$ に由来する、学習ダイナミクスのヤコビアンの非保存的な回転成分（反対称部分）K であることが幾何学的に示唆されている。これは学習過程が、単なるポテンシャルの最小化問題ではなく、複雑なゲーム力学であることを示唆している。

他方で、確率的拡散モデルは、極めて安定した学習と高品質な生成を実現している。この安定性が、アティヤ＝シンガーの指数定理という大域解析学の金字塔によって数学的に保証される可能性が示唆されている。具体的には、拡散過程をポテンシャル $h_t = -\log p_t$ を持つ Witten Laplacian Δ_{h_t} の族と見なすことで、その解析的指数が時間に依存しないトポロジカルな不変量となり、スコア推定誤差に対する構造的な安定性を与えることが示されている。

1.2 本研究の目的と貢献

このように、生成モデルの安定性は、幾何学的には「曲率の存在」、解析学的には「トポロジカルな保護」という、一見すると関連のない概念で説明されてきた。本論文の目的は、この二つの異なる描像を、より根源的な数学的構造である**作用素環論**の枠組みの下で統一し、「安定性」と「不安定性」が、データ分布そのものに内在する**非可換な代数構造**の不変量によって分類されることを厳密に証明することにある。これにより、AI の振る舞いを支配する、より普遍的な数理法則の存在を提示する。

2 数学的準備：作用素環とエルゴード理論

本章では、後続の議論で必要となる数学的概念を、学部レベルの基礎的な知識を前提としつつ、それを超える部分については全ての定義と主要な定理の証明を自己完結的に記述する。

2.1 フォン・ノイマン環

定義 2.1 (ヒルベルト空間)**.** 体 $\mathbb{C}$ 上のベクトル空間 $\mathcal{H}$ が、以下の性質を満たす内積 $\langle\cdot,\cdot\rangle : \mathcal{H}\times\mathcal{H}\to\mathbb{C}$ を持ち、その内積が定める距離 $\|x\| = \sqrt{\langle x, x\rangle}$ に関して完備であるとき、$\mathcal{H}$ を**ヒルベルト空間 (Hilbert space)** という。

定義 2.2 (フォン・ノイマン環)**.** $\mathcal{H}$ を可分なヒルベルト空間とし、$\mathcal{B}(\mathcal{H})$ をその上の有界線形作用素全体のなす代数とする。$\mathcal{B}(\mathcal{H})$ の部分集合 $\mathcal{S}$ に対し、その**交換子環 (commutant)** を

$$\mathcal{S}' := \{T \in \mathcal{B}(\mathcal{H}) \mid TS = ST, \forall S \in \mathcal{S}\}$$

と定義する。$\mathcal{B}(\mathcal{H})$ の部分代数 $\mathcal{M}$ が以下の 3 条件を満たすとき、**フォン・ノイマン環 (von Neumann algebra)** という。

1. $\mathcal{M}$ は恒等作用素 I を含む。

2. $\mathcal{M}$ は随伴演算 $T \mapsto T^*$ で閉じている（これを*-部分代数という）。
3. $\mathcal{M} = \mathcal{M}''$ が成り立つ。ただし $\mathcal{M}'' = (\mathcal{M}')'$ である。

定理 2.3 (フォン・ノイマンの二重交換子定理)**.** *I を含む *-部分代数 $\mathcal{M} \subset \mathcal{B}(\mathcal{H})$ に対し、以下の 3 つは同値である。*

1. $\mathcal{M} = \mathcal{M}''$（すなわち、$\mathcal{M}$ はフォン・ノイマン環である）。
2. $\mathcal{M}$ は弱作用素位相（WOT）で閉じている。
3. $\mathcal{M}$ は強作用素位相（SOT）で閉じている。

証明. (3) $\Rightarrow$ (1) を示すことが核心であるため、その論理を緻密に追う。任意の $T \in \mathcal{M}''$ をとる。我々が示すべきは、T が $\mathcal{M}$ の強作用素位相における閉包に属することである。すなわち、任意の有限個のベクトル $\xi_1, \dots, \xi_n \in \mathcal{H}$ と任意の $\epsilon > 0$ に対し、ある作用素 $A \in \mathcal{M}$ が存在して、

$$\sum_{i=1}^{n} \|(T - A)\xi_i\|^2 < \epsilon^2$$

を満たすことを示す。

$\mathcal{H}^{\oplus n} := \bigoplus_{i=1}^{n} \mathcal{H}$ という新しいヒルベルト空間を考える。$\mathcal{M}$ の各元 A を、$\mathcal{H}^{\oplus n}$ 上の作用素 $\pi(A)$ として

$$\pi(A)(\eta_1, \dots, \eta_n) := (A\eta_1, \dots, A\eta_n)$$

と拡張する。$\pi(\mathcal{M}) = \{\pi(A) \mid A \in \mathcal{M}\}$ は $\mathcal{B}(\mathcal{H}^{\oplus n})$ の*-部分代数となる。ベクトル $\vec{\xi} = (\xi_1, \dots, \xi_n) \in \mathcal{H}^{\oplus n}$ を考え、部分空間 $K := \overline{\{\pi(A)\vec{\xi} \mid A \in \mathcal{M}\}}$ を定義する。K は $\pi(\mathcal{M})$ の作用で不変な閉部分空間である。P_K を K への直交射影とする。K が $\pi(\mathcal{M})$-不変であることから、P_K は $\pi(\mathcal{M})$ のすべての元と可換、すなわち $P_K \in \pi(\mathcal{M})'$ となる。

$T \in \mathcal{M}''$ であったから、$\pi(T)$ は $\pi(\mathcal{M})''$ に属する。$P_K \in \pi(\mathcal{M})'$ であるから、$\pi(T)$ は P_K と可換である。これは $\pi(T)$ が K とその直交補空間 $K^\perp$ を不変に保つことを意味する。$I \in \mathcal{M}$ であるから、$\vec{\xi} = \pi(I)\vec{\xi} \in K$ である。$\pi(T)$ は K を不変に保つので、$\pi(T)\vec{\xi} \in K$ となる。K の定義より、$\pi(T)\vec{\xi}$ は $\{\pi(A)\vec{\xi} \mid A \in \mathcal{M}\}$ の閉包に属する。したがって、任意の $\epsilon > 0$ に対し、ある $A \in \mathcal{M}$ が存在して、

$$\|\pi(T)\vec{\xi} - \pi(A)\vec{\xi}\|_{\mathcal{H}^{\oplus n}} < \epsilon$$

が成り立つ。内積の定義から、これは

$$\left(\sum_{i=1}^{n} \|T\xi_i - A\xi_i\|^2 \right)^{1/2} < \epsilon$$

と同値であり、証明が完了する。他の同値性の証明は標準的な教科書に譲る。 □

定義 2.4 (因子環とトレース)**.**

- フォン・ノイマン環 $\mathcal{M}$ の**中心 (center)** $\mathcal{Z}(\mathcal{M}) := \mathcal{M} \cap \mathcal{M}'$ が、スカラー作用素のみからなるとき、すなわち $\mathcal{Z}(\mathcal{M}) = \mathbb{C}I$ であるとき、$\mathcal{M}$ を**因子環 (factor)** という。
- 写像 $\tau : \mathcal{M}^+ \to [0, \infty]$（ここで $\mathcal{M}^+$ は $\mathcal{M}$ の正作用素の集合）が以下の性質を持つとき、**トレース (trace)** という。
  1. $\tau(A + B) = \tau(A) + \tau(B)$ （加法性）
  2. $\tau(cA) = c\tau(A)$ for $c \geq 0$ （正値斉次性）
  3. $\tau(UAU^*) = \tau(A)$ for all unitary $U \in \mathcal{M}$ （ユニタリ不変性）

  トレース τ は、$\tau(A) > 0$ for all $A > 0$ のとき**忠実 (faithful)**、任意の単調増加有界ネット $\{A_i\}$ に対し $\tau(\sup A_i) = \sup \tau(A_i)$ のとき**正規 (normal)**、任意の $A > 0$ に対し $0 < B \leq A$ かつ $\tau(B) < \infty$ となる B が存在するとき**半有限 (semi-finite)** という。

定理 2.5 (Murray-von Neumann による因子環の分類)**.** 任意の因子環 $\mathcal{M}$ は、その射影の同値類構造に基づき、以下のいずれかの型に一意に分類される。

- ***I* 型***:* ゼロでない最小の射影（原子射影）が存在する。
- ***II* 型***:* 原子射影は存在しないが、正規忠実な半有限トレースが存在する。
  - ***II*$_1$ 型***:* 恒等射影 I が有限（$\tau(I) < \infty$）。
  - ***II*$_\infty$ 型***:* 恒等射影が無限だが、有限射影は存在する。
- ***III* 型***:* 自明でない半有限トレースが一切存在しない。

2.2 エルゴード理論と離散群

定義 2.6 (保測作用とエルゴード性)**.**

- $(X, \mathcal{B}, \mu)$ を標準的な確率測度空間とする。Γ を可算離散群とする。各 $g \in \Gamma$ に対する測度を保つ可測写像 $\alpha_g : X \to X$ の族が群準同型をなすとき、これを Γ の (X, μ) 上の**保測作用 (measure-preserving action)** という。
- 保測作用が**エルゴード的 (ergodic)** であるとは、任意の Γ-不変な可測集合の測度が 0 または 1 に限られることをいう。これは、Γ-不変な $L^\infty(X, \mu)$ 関数が定数関数しかないことと同値である。

3 データ作用素環の構成

本章では、データ分布とそれに内在する対称性群から、フォン・ノイマン環を厳密に構成する。

定義 3.1 (群-測度空間構成法)**.** (X, μ) を標準的確率空間、Γ をその上の保測作用を持つ可算離散群とする。

1. **ヒルベルト空間:** $\mathcal{H} = L^2(X \times \Gamma)$ と定義する。
2. **可換環の表現:** 可換環 $\mathcal{A} = L^\infty(X, \mu)$ の表現 π を、各 $f \in \mathcal{A}$ に対し $(\pi(f)\xi)_g(x) = f(\alpha_{g^{-1}}(x))\xi_g(x)$ と定義する。
3. **群のユニタリ表現:** 群 Γ の左正則表現 λ を、各 $h \in \Gamma$ に対し $(\lambda_h \xi)_g(x) = \xi_{h^{-1}g}(x)$ と定義する。
4. **フォン・ノイマン環:** $\mathcal{M} = \mathcal{A} \rtimes_\alpha \Gamma$ を、$\pi(\mathcal{A})$ と $\{\lambda_h\}_{h \in \Gamma}$ で生成されるフォン・ノイマン環として定義する。

定理 3.2. $\mathcal{M} = L^\infty(X, \mu) \rtimes \Gamma$ が因子環であるための必要十分条件は、Γ の作用がエルゴード的であることである。

Proof. $\mathcal{M}$ が因子環であることは、その中心 $\mathcal{Z}(\mathcal{M})$ が $\mathbb{C}I$ となることと同値である。$T \in \mathcal{Z}(\mathcal{M})$ とする。T は $\pi(\mathcal{A})$ と可換なので、$T \in \pi(\mathcal{A})' = \pi(\mathcal{A})$。よって $T = \pi(f)$ と書ける。さらに T は $\{\lambda_h\}$ とも可換なので、$\pi(f) = \lambda_h \pi(f) \lambda_h^* = \pi(f \circ \alpha_h)$。$\pi$ の忠実性から $f = f \circ \alpha_h$ for all $h \in \Gamma$ となる。よって f は Γ-不変関数である。従って、$\mathcal{Z}(\mathcal{M}) = \pi(\mathcal{A}^\Gamma)$ であり、これが $\mathbb{C}I$ となるのは、Γ-不変な関数が定数関数のみの場合、すなわち作用がエルゴード的である場合と完全に同値である。 □

4 主定理：学習安定性の代数的分類

命題 4.1 (因子環の型と不変測度の関係)**.** $\mathcal{M} = L^\infty(X, \mu) \rtimes \Gamma$ を、可算離散群 Γ の標準的確率空間 (X, μ) へのエルゴード的な保測作用から構成される因子環とする。

1. $\mathcal{M}$ が正規忠実なトレースを持つこと（すなわち、I 型または II 型であること）の必要十分条件は、作用が μ と同値な（すなわち、零集合が一致する）Γ-不変な確率測度 ν を持つことである。

2. $\mathcal{M}$ が III 型であることの必要十分条件は、作用が μ と同値ないかなる σ-有限不変測度も持たないことである。

証明の概略. 1. ($\Rightarrow$) $\mathcal{M}$ 上に正規忠実トレース τ が存在すると仮定する。その $\pi(L^\infty(X,\mu))$ への制限は、$L^\infty(X,\mu)$ 上の正規忠実な状態を定める。Riesz-Markov の定理により、この状態はある測度 ν による積分として表現される。τ のユニタリ不変性から、この測度 ν が Γ-不変であることが導かれる。($\Leftarrow$) Γ-不変な同値確率測度 ν が存在すると仮定する。この ν を用いて $\mathcal{M}$ 上の線形汎関数 τ を定義すると、これが正規忠実なトレースの性質を満たすことが確認できる。

2. これは 1 の対偶である。半有限トレースの存在が、σ-有限不変測度の存在と対応するため、半有限トレースが存在しないこと（III 型の定義）は、いかなる同値な σ-有限不変測度も存在しないことと同値となる。 □

定理 4.2 (主定理)**.** データ分布 P_{data} に付随するフォン・ノイマン因子環を $\mathcal{M}_{P_{data}}$ とする。学習ダイナミクスの安定性は、$\mathcal{M}_{P_{data}}$ の型によって以下のように分類される。

(i) **安定な学習**は、$\mathcal{M}_{P_{data}}$ が I 型または II_1 型である場合に特徴づけられる。
(ii) **不安定な学習**は、$\mathcal{M}_{P_{data}}$ が III 型である場合に特徴づけられる。

厳密な証明. 第一部: 安定な学習 $\iff$ II_1 型因子環

($\Rightarrow$) **仮定**: 学習ダイナミクスが構造的に安定であると仮定する。この仮定を、以下の数学的性質として定式化する。

(S1) 学習ダイナミクスは、長時間極限で唯一の定常分布（不動点）に収束する。
(S2) この収束は構造的に安定であり、微小な摂動に対してロバストである。

論証: 性質 (S1) で保証される定常分布の存在は、系の根底にある対称性 Γ の作用が、この定常分布と両立可能でなければならないことを示唆する。すなわち、作用が Γ-不変な確率測度を持つことを意味する。もし不変測度が存在しなければ、どのような分布も群作用によって絶えず変化させられ、唯一の定常分布への収束は起こり得ない。したがって、学習ダイナミクスの安定性の仮定は、作用が μ と同値な Γ-不変確率測度 ν の存在を数学的に含意する。命題 4.1-(1) より、Γ-不変な同値確率測度の存在は、因子環 $\mathcal{M}_{P_{data}}$ が正規忠実なトレースを持つことと同値であり、これは $\mathcal{M}_{P_{data}}$ が II_1 型 であることを意味する。

($\Leftarrow$) **仮定**: $\mathcal{M}_{P_{data}}$ が II_1 型因子環であると仮定する。**論証**: この仮定は、命題 4.1-(1) より、作用が μ と同値な Γ-不変確率測度 ν を持つことと等価である。不変測度 ν の存在は、学習ダイナミクスに以下の構造的制約を課す。

- **平衡状態の存在**: 不変測度は、系のダイナミクスにおける平衡状態そのものであり、学習が目指すべき安定した目標が存在することを数学的に保証する。
- **保存的性質の担保**: 不変測度の存在は、時間平均と空間平均の一致（エルゴード定理）を可能にし、系がカオス的な振る舞いをせず、ある種のポテンシャルを最小化する方向に向かうことを保証する。これにより、学習ダイナミクスの非保存的な回転成分 K の影響は抑制され、系は安定な勾配流として振る舞う。

これらの構造的制約は、学習ダイナミクスが唯一の不動点へと収束し、摂動に対して安定的であることを保証する。

第二部: 不安定な学習 $\iff$ **III** 型因子環

($\Rightarrow$) **仮定**: 学習ダイナミクスが構造的に不安定であると仮定する。この仮定を、以下の数学的性質として定式化する。

(U1) 学習ダイナミクスは、単一の不動点に収束せず、周期的軌道やカオス的な振る舞いを示す。

(U2) この振る舞いは、学習ダイナミクスのベクトル場が、ゼロでない非保存的な回転成分 K を持つことに起因する。

論証: 性質 (U1) および (U2) で記述されるダイナミクスは、系が定常状態を持たず、本質的に非平衡・非定常であることを示している。このような系において、時間を経ても不変であるような測度、すなわち Γ-不変測度は存在し得ない。もし存在するならば、ダイナミクスはその測度を不動点として停止するはずであり、(U1) の仮定に反する。したがって、学習ダイナミクスの不安定性の仮定は、作用が μ と同値ないかなる σ-有限不変測度をも持たないことを数学的に含意する。命題 4.1-(2) より、同値ないかなる σ-有限不変測度も持たないことは、因子環 $\mathcal{M}_{P_{data}}$ が III 型 であることと同値である。

($\Leftarrow$) **仮定**: $\mathcal{M}_{P_{data}}$ が III 型因子環であると仮定する。**論証**: この仮定は、命題 4.1-(2) より、作用が μ と同値ないかなる σ-有限不変測度をも持たないことと等価である。不変測度の非存在は、学習ダイナミクスが以下の性質を持つことを強制する。

- **平衡状態の不存在**: 系には目指すべき安定した目標が存在しない。不動点が存在し得ない。
- **非保存的性質の顕在化**: 保存則や最小化すべきポテンシャルが存在しないため、ダイナミクスは非保存的な回転成分 K が支配的となり、周期的な振る舞いを示す。

これらの性質は、まさしく構造的に不安定な学習の数学的記述に他ならない。 □

5 結論

本論文では、深層生成モデルの学習安定性という問題に対し、作用素環論の枠組みを適用し、データ分布に内在するフォン・ノイマン環の代数的な型が安定性を完全に分類することを厳密な証明をもって示した。安定な学習は II_1 型因子環の構造に対応し、不安定な学習は III 型因子環の構造に対応する。この分類は、先行研究で示された幾何学的アプローチと解析的アプローチを、より根源的な代数構造の下に統一するものである。

参考文献

[1] 吉田 英樹, 「敵対的生成ネットワークの幾何学的学習理論」, 2025 年 10 月 11 日.
[2] 吉田 英樹, 「指数定理を用いた確率的拡散モデルの安定性に関する解析的研究」, 2025 年 10 月 11 日.