

生成敵対的ネットワークの ゲージ理論、作用原理、量子化による幾何学的描像

吉田英樹

2025 年 9 月 26 日

概要

本稿は、生成敵対的ネットワーク (GAN) の学習ダイナミクスを、ゲージ理論、作用原理、そして経路積分による量子化という物理学の概念を用いて統一的に定式化する。ワッサーズタイン多様体を底空間とするファイバー束の枠組みを基礎とし、まず識別器の双対的な役割を余接ベクトル場と束の切断として厳密に定義する。次に、学習プロセスを記述するラグランジアンと散逸関数を定義し、拡張された変分原理を適用することで、勾配フロー方程式が系の運動方程式の適切な極限として現れることを示す。最後に、この枠組みをトポロジーと量子化の概念へと拡張する。GAN ファイバー束の自明なトポロジーが持つ意味を考察し、学習プロセスを決定論的な単一軌道ではなく、可能なすべての学習経路にわたる確率的な重ね合わせとして記述する経路積分量子化を導入する。このアプローチは、GAN の学習における確率性や不安定性を、場の量子論の言語で理解するための新たな理論的視座を提供する。

1 序論

Goodfellow らによって導入された生成敵対的ネットワーク (GAN) は、生成モデリングにおけるパラダイムシフトを象徴する [1]。その広範な経験的成功にもかかわらず、そのダイナミクス、不安定性、そしてモード崩壊のような失敗モードに関する深い理論的理解は、依然として活発な研究分野である。

これまでの理論的分析は、ダイバージェンス最小化との関連 [3] や、積分確率計量 (Integral Probability Metrics) としての定式化 [2] に焦点を当ててきた。これらは静的な目的関数を記述するが、動的な学習プロセス自体を捉えきれていない。本稿は、敵対的学習と微分幾何学の概念との間に橋を架けることで、学習軌道そのものの幾何学的な理解を提供することを目指す。中心的なテーマは、GAN の学習プロセスが、確率分布の無限次元多様体上の幾何学的フローとして厳密に記述できるということである。

本稿では、この視点をファイバー束の言語を用いてさらに厳密化し、物理学の作用原理と量子化の概念へと拡張することで、以下の貢献を提示する。

- 厳密な幾何学的設定:** ワッサーズタイン-2 空間 $(W_2(\mathcal{X}), g_W)$ を、本理論の底空間多様体として形式的に確立する。
- 識別器の双対的役割:** 最適識別器の役割を、余接空間 $T_{P_g}^* W_2(\mathcal{X})$ におけるベクトルを定義するものとして、またそれと等価なものとして、ファイバー束上の「切断」を定める写像として数学的に正確に定義する。
- 勾配フローとしてのダイナミクス:** 生成器のダイナミクスを、この切断から誘導される「敵対的接続」に基づく底空間上の勾配フローとして定式化する。
- 主定理—平坦な接続としての均衡:** GAN の均衡状態が、この接続が「平坦」になる状態と完全に同値であることを証明する。
- 作用原理による導出:** この勾配フローが、系のラグランジアンと散逸関数から、拡張された変分原理を通じて導出されることを示す。
- 量子化への展望:** 経路積分を用いて学習プロセスを量子化し、SGD の確率性や学習の不安定性を場の量

子論の言語で解釈する道筋を示す。

2 数学的準備

2.1 生成敵対的ネットワーク

標準的な GAN の目的関数は、以下のミニマックスゲームとして与えられる [1]。

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim P_{\text{data}}} [\log D(x)] + \mathbb{E}_{z \sim P_z} [\log(1 - D(G(z)))] \quad (1)$$

固定された G に対し、最適識別器は $D_G^*(x) = \frac{P_{\text{data}}(x)}{P_{\text{data}}(x) + P_g(x)}$ で与えられる [1]。これを代入すると、生成器が最小化すべきコスト関数は、イェンゼン・シャノン・ダイバージェンス (JSD) の 2 倍になる [3]。

$$C(G) = \max_D V(D, G) = 2 \cdot D_{\text{JS}}(P_{\text{data}} \| P_g) \quad (2)$$

2.2 底空間：ワッサーズタイン-2 空間

定義 2.1 (ワッサーズタイン-2 空間). \mathcal{X} をコンパクトなリーマン多様体とする。 \mathcal{X} 上の確率測度で 2 次モーメントが有限なものの集合を $\mathcal{P}_2(\mathcal{X})$ とする。ワッサーズタイン-2 距離は以下のように定義される [2]:

$$W_2(\mu, \nu) = \left(\inf_{\gamma \in \Gamma(\mu, \nu)} \int_{\mathcal{X} \times \mathcal{X}} d(x, y)^2 d\gamma(x, y) \right)^{1/2}$$

組 $(\mathcal{P}_2(\mathcal{X}), W_2)$ をワッサーズタイン-2 空間 $\mathcal{W}_2(\mathcal{X})$ と表記する。

Otto の独創的な結果により、 $\mathcal{W}_2(\mathcal{X})$ は形式的に無限次元リーマン多様体と見なせる [4]。

定義 2.2 (接空間とリーマン計量). 測度 $P_g \in \mathcal{W}_2(\mathcal{X})$ における接空間 $T_{P_g} \mathcal{W}_2(\mathcal{X})$ は、スカラー関数の勾配である \mathcal{X} 上のベクトル場の空間と同一視できる。

$$T_{P_g} \mathcal{W}_2(\mathcal{X}) \cong \{v = -\nabla \phi \mid \phi \in C^\infty(\mathcal{X})\}$$

リーマン内積 (オットー計量) は以下で与えられる [4]:

$$g_{P_g}(v_1, v_2) = \int_{\mathcal{X}} \langle v_1(x), v_2(x) \rangle_x dP_g(x)$$

3 GAN ファイバー束の構築と識別器の役割

3.1 GAN ファイバー束の定義

定義 3.1 (GAN ファイバー束). GAN ファイバー束を、以下の要素からなる組 $(E, M, \pi, F, \mathcal{G})$ として定義する。

- 底空間 (*Base Space*): $M = \mathcal{W}_2(\mathcal{X})$
- ファイバー (*Fiber*): $F = C^\infty(\mathcal{X})$ (\mathcal{X} 上の滑らかな実数値関数の空間)
- 全空間 (*Total Space*): $E = M \times F = \mathcal{W}_2(\mathcal{X}) \times C^\infty(\mathcal{X})$
- 射影 (*Projection*): $\pi: E \rightarrow M, \quad \pi(P_g, \phi) = P_g$
- 構造群 (*Structure Group*): 自明な群 $\mathcal{G} = \{\text{id}_F\}$

この束は自明な束であり、ファイバーが底空間上で「ねじれて」いないことを意味する。

3.2 識別器の双対的解釈：余接ベクトルと切断

命題 3.2 (識別器の役割). 最適識別器 $D_G^*(x)$ は、各点 $P_g \in M$ において、二つの等価な数学的対象を定める。

1. 生成器の目的汎関数 $\mathcal{J}'(P_g) = \mathbb{E}_{x \sim P_g} [\log(1 - D_G^*(x))]$ の微分である余接ベクトル $d\mathcal{J}'_{P_g} \in T_{P_g}^* \mathcal{W}_2(\mathcal{X})$ 。
2. GAN ファイバー束 E 上の唯一の切断 $s_D : M \rightarrow E$ 。

Proof. この命題の要点は、最適識別器 D_G^* が一意に定まると、ポテンシャル関数 ϕ_{P_g} も一意に定まり、そのポテンシャル関数が系のすべての幾何学的対象（余接ベクトルと切断）を決定することを示すことにある。

1. まず、任意の生成分布 $P_g \in M$ に対し、最適識別器 $D_G^*(x) = \frac{P_{\text{data}}(x)}{P_{\text{data}}(x) + P_g(x)}$ が一意に定まる。2. 次に、生成器の損失に対応するポテンシャル関数を $\phi_{P_g}(x) = \log(1 - D_G^*(x))$ と定義する。これを具体的に計算すると、

$$\phi_{P_g}(x) = \log \left(1 - \frac{P_{\text{data}}(x)}{P_{\text{data}}(x) + P_g(x)} \right) = \log \left(\frac{P_g(x)}{P_{\text{data}}(x) + P_g(x)} \right)$$

となり、このポテンシャル関数 ϕ_{P_g} も各 P_g に対して一意に定まる。

3. **余接ベクトルの導出:** 汎関数 $\mathcal{J}'(P_g) = \int_{\mathcal{X}} \phi_{P_g}(x) dP_g(x)$ の、任意の接ベクトル $V \in T_{P_g} \mathcal{W}_2(\mathcal{X})$ 方向へのガトー微分 $d\mathcal{J}'_{P_g}(V)$ を考える。これは定義により、 $T_{P_g}^* \mathcal{W}_2(\mathcal{X})$ の元である余接ベクトル $d\mathcal{J}'_{P_g}$ が V に作用することに等しい。この微分は ϕ_{P_g} を用いて計算可能であり、一意に定まる。

4. **切断の導出:** 切断 $s_D : M \rightarrow E$ は、各点 $P_g \in M$ に対し、その上のファイバー F の元を一つ指定する写像である。写像 $P_g \mapsto (P_g, \phi_{P_g}(x))$ は、まさにこの定義を満たす。 ϕ_{P_g} が一意であるため、この切断 s_D も一意に定まる。

以上の通り、ポテンシャル関数 ϕ_{P_g} が中間的な役割を果たし、識別器 D_G^* から余接ベクトル $d\mathcal{J}'_{P_g}$ と切断 s_D が共に一意に定まる。したがって、これら二つの役割は等価である。 \square

4 接続とダイナミクス

定義 4.1 (敵対的接続). 敵対的接続 A_{P_g} を、識別器が定めた余接ベクトル $d\mathcal{J}'_{P_g}$ に対応する勾配ベクトル場として定義する。

$$A_{P_g} := (\text{grad} \mathcal{J}')_{P_g} = \nabla \phi_{P_g}(x) = \nabla \left(\log \frac{P_g(x)}{P_{\text{data}}(x) + P_g(x)} \right)$$

定義 4.2 (GAN 勾配フロー). 生成器の理想化された学習ダイナミクスは、底空間 M 上で、接続の負の方向に沿った勾配フローとして記述される。

$$\frac{dP_t}{dt} = -A_{P_t}$$

これは偏微分方程式 $\frac{\partial P_t}{\partial t} = \nabla \cdot (P_t \cdot A_{P_t})$ と等価である。

5 主定理：平坦な接続としての均衡

定理 5.1 (平坦な接続としての均衡). GAN のダイナミクスが不動点にあるための必要十分条件は、敵対的接続 A_{P_g} がヌルベクトル場であることである。この条件は、 $P_g = P_{\text{data}}$ がほとんど至る所で成立する場合にのみ満たされる。

Proof. この定理を、二つの部分に分けて詳細に証明する。

Part 1: ($P_g = P_{\text{data}} \implies A_{P_g} = 0$). 仮定として、 $P_g(x) = P_{\text{data}}(x)$ が \mathcal{X} 上でほとんど至る所で成立するとする。敵対的接続 A_{P_g} は、ポテンシャル関数 $\phi_{P_g}(x) = \log \frac{P_g(x)}{P_{\text{data}}(x) + P_g(x)}$ の勾配として定義される。こ

のポテンシャル関数に仮定を代入すると、

$$\phi_{P_g}(x) = \log \frac{P_{\text{data}}(x)}{P_{\text{data}}(x) + P_{\text{data}}(x)} = \log \frac{P_{\text{data}}(x)}{2P_{\text{data}}(x)} = \log(1/2) = -\log 2$$

となる。これは x に依存しない定数である。したがって、定数関数の勾配はゼロであるから、

$$A_{P_g}(x) = \nabla(-\log 2) = 0$$

となり、敵対的接続はヌルベクトル場である。勾配フロー $\frac{dP_t}{dt} = -A_{P_t}$ は $\frac{dP_t}{dt} = 0$ となり、この点が不動点であることが示された。

Part 2: ($A_{P_g} = 0 \implies P_g = P_{\text{data}}$). 仮定として、敵対的接続がヌルベクトル場であるとする: $A_{P_g}(x) = 0$ がほとんど至る所で成立する。接続の定義より、これは $\nabla \phi_{P_g}(x) = 0$ を意味する。ベクトル場の勾配がゼロであるならば、そのポテンシャル関数は（定義域の各連結成分上で）定数でなければならない。ある定数を C とすると、

$$\log \frac{P_g(x)}{P_{\text{data}}(x) + P_g(x)} = C$$

この式の両辺の指数をとると、

$$\frac{P_g(x)}{P_{\text{data}}(x) + P_g(x)} = e^C$$

$K = e^C$ と置く。 K は正の定数である。この式を $P_g(x)$ について代数的に解く。

$$\begin{aligned} P_g(x) &= K(P_{\text{data}}(x) + P_g(x)) \\ P_g(x) &= KP_{\text{data}}(x) + KP_g(x) \\ P_g(x) - KP_g(x) &= KP_{\text{data}}(x) \\ P_g(x)(1 - K) &= KP_{\text{data}}(x) \end{aligned}$$

ここで $K = 1$ の場合を考えると、 $0 = P_{\text{data}}(x)$ となり、 P_{data} が確率分布であることに矛盾する。したがって $K \neq 1$ であり、両辺を $1 - K$ で割ることができる。

$$P_g(x) = \frac{K}{1 - K} P_{\text{data}}(x)$$

$\alpha = K/(1 - K)$ は定数であるから、 $P_g(x) = \alpha P_{\text{data}}(x)$ が示された。 P_g と P_{data} は共に確率分布であるため、その全空間での積分は 1 に等しい。

$$\int_{\mathcal{X}} P_g(x) dx = 1 \quad \text{and} \quad \int_{\mathcal{X}} P_{\text{data}}(x) dx = 1$$

この関係を用いると、

$$1 = \int_{\mathcal{X}} P_g(x) dx = \int_{\mathcal{X}} \alpha P_{\text{data}}(x) dx = \alpha \int_{\mathcal{X}} P_{\text{data}}(x) dx = \alpha \cdot 1 = \alpha$$

したがって $\alpha = 1$ でなければならない。これにより、 $P_g(x) = P_{\text{data}}(x)$ がほとんど至る所で成立することが証明された。 \square

6 散逸系としてのダイナミクスと変分原理

標準的な最小作用の原理はエネルギーが保存される系を記述する。しかし、勾配降下法のような最適化プロセスは、ポテンシャルエネルギーが減衰していく「散逸系」である。このような系をラグランジュ形式で厳密に扱うため、レイリーの散逸関数を導入し、変分原理を拡張する。

定義 6.1 (系のラグランジアン). 学習プロセスの保存力部分を記述するラグランジアン \mathcal{L} を、運動エネルギー項とポテンシャルエネルギー項の差として定義する。

$$\mathcal{L}(P_t, \dot{P}_t) := \frac{1}{2} \|\dot{P}_t\|_{g_{P_t}}^2 - \mathcal{J}'(P_t)$$

定義 6.2 (レイリーの散逸関数). 系の散逸（摩擦）効果を記述するため、レイリーの散逸関数 \mathcal{F} を以下のように定義する。

$$\mathcal{F}(P_t, \dot{P}_t) := \frac{1}{2} \gamma \|\dot{P}_t\|_{g_{P_t}}^2$$

ここで $\gamma > 0$ は系の摩擦の大きさを表す正の定数（摩擦係数）である。

定義 6.3 (作用汎関数). 系の作用 S を、ラグランジアンの時間積分として定義する。

$$S[P_t] = \int_0^T \mathcal{L}(P_t, \dot{P}_t) dt$$

定理 6.4 (拡張された変分原理と勾配フロー). ラグランジアン \mathcal{L} と散逸関数 \mathcal{F} で記述される系の運動方程式は、拡張されたラグランジュ方程式 $\frac{\delta S}{\delta P_t} + \frac{\partial \mathcal{F}}{\partial \dot{P}_t} = 0$ から導かれ、

$$\nabla_{\dot{P}_t} \dot{P}_t + \gamma \dot{P}_t = -(\text{grad} \mathcal{J}')_{P_t}$$

となる。さらに、摩擦が極めて大きい過減衰極限 ($\gamma \rightarrow \infty$) において、この運動方程式は勾配フロー方程式に帰着する。

Proof. 散逸系に対する変分原理は、作用 S の変分と、散逸関数から導かれる仮想仕事の和がゼロになることで与えられる。

$$\delta S + \int_0^T \frac{\partial \mathcal{F}}{\partial \dot{P}_t} \delta P_t dt = 0$$

作用 S の第一変分 δS を計算すると、(以前の証明と同様に)

$$\delta S = - \int_0^T g_{P_t}(V, \nabla_t \dot{P}_t + (\text{grad} \mathcal{J}')_{P_t}) dt$$

が得られる。ここで $V = \delta P_t$ は変分ベクトル場である。一方、散逸項の変分は

$$\int_0^T \frac{\partial \mathcal{F}}{\partial \dot{P}_t} V dt = \int_0^T g_{P_t}(\gamma \dot{P}_t, V) dt$$

と書ける。これらを合わせると、任意の変分 V に対して

$$\int_0^T g_{P_t}(V, \nabla_t \dot{P}_t + (\text{grad} \mathcal{J}')_{P_t} + \gamma \dot{P}_t) dt = 0$$

が成立しなければならない。これにより、被積分関数がゼロであることが要求され、以下の運動方程式が得られる。

$$\nabla_t \dot{P}_t + \gamma \dot{P}_t + (\text{grad} \mathcal{J}')_{P_t} = 0 \quad \implies \quad \nabla_{\dot{P}_t} \dot{P}_t + \gamma \dot{P}_t = -(\text{grad} \mathcal{J}')_{P_t}$$

この方程式は、慣性項 ($\nabla_{\dot{P}_t} \dot{P}_t$)、摩擦項 ($\gamma \dot{P}_t$)、そしてポテンシャルからの力 ($-(\text{grad} \mathcal{J}')_{P_t}$) の釣り合いを表している。

次に、勾配降下法の振る舞いに対応する**過減衰極限** ($\gamma \rightarrow \infty$) を考える。この極限では、摩擦が非常に大きいため、慣性は即座に減衰し、速度と加速度は有界に留まると考えられる。運動方程式を \dot{P}_t について整理すると、

$$\dot{P}_t = -\frac{1}{\gamma} (\text{grad} \mathcal{J}')_{P_t} - \frac{1}{\gamma} \nabla_{\dot{P}_t} \dot{P}_t$$

$\gamma \rightarrow \infty$ の極限をとると、右辺第二項はゼロに収束する。

$$\lim_{\gamma \rightarrow \infty} \left(\frac{1}{\gamma} \nabla_{\dot{P}_t} \dot{P}_t \right) = 0$$

したがって、運動方程式は以下のように近似される。

$$\dot{P}_t \approx -\frac{1}{\gamma} (\text{grad} \mathcal{J}')_{P_t}$$

ここで、係数 $1/\gamma$ を学習率 η と解釈すれば、これは勾配フロー方程式そのものである。 \square

7 トポロジーと量子化への展望

7.1 束のトポロジーとその意味

我々の GAN ファイバー束は自明な束であり、学習ランドスケープに大域的なトポロジカルな障害が存在しないことを示唆する。原理的にはどの分布からでも均衡点へと連続的に変形可能である。

7.2 経路積分による量子化

定義 7.1 (経路積分と分配関数). *GAN* の学習プロセスの量子論的な振る舞いを、以下の経路積分によって定義される分配関数 Z によって記述する。

$$Z = \int \mathcal{D}[P_t] \exp(-S[P_t])$$

ここで $\int \mathcal{D}[P_t]$ は、すべての可能な学習経路にわたる汎関数積分を表す。

この定式化は、学習プロセスを、作用が小さい古典的軌道（勾配フロー）の周りに「量子ゆらぎ」を持つ確率的なものとして捉え直す。このゆらぎは SGD における確率性と関連付けられ、学習の不安定性やモード崩壊は、統計物理学における「相転移」のアナロジーで理解できる可能性がある。

8 結論

本稿では、GAN の学習ダイナミクスをファイバー束、作用原理、そして量子化という一貫した幾何学的・物理学的言語で記述した。識別器を切断、学習を接続に沿った勾配フローと見なす描像から出発し、このフローが、散逸を考慮した拡張された変分原理の自然な帰結（過減衰極限）であることを示した。さらに、経路積分による量子化を導入することで、SGD のような確率的な振る舞いを場の量子論の言語で捉える理論的枠組みを提示した。このアプローチは、深層生成モデルの振る舞いを支配する根本的な数学的構造を解明するための、新たな道筋を開くものである。

参考文献

- [1] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. (2014). Generative adversarial nets. In *Advances in Neural Information Processing Systems 27*, pages 2672-2880.
- [2] Martin Arjovsky, Soumith Chintala, and Léon Bottou. (2017). Wasserstein generative adversarial networks. In *Proceedings of the 34th International Conference on Machine Learning*, pages 214-223.

- [3] Sebastian Nowozin, Botond Cseke, and Ryota Tomioka. (2016). f-GAN: Training generative neural samplers using variational divergence minimization. In *Advances in Neural Information Processing Systems 29*, pages 271-279.
- [4] Felix Otto. (2001). The geometry of dissipative evolution equations: the porous medium equation. *Communications in Partial Differential Equations*, 26(1-2):101-174.