

敵対的模倣学習における 近似不動点の構造的存在証明

- 豊穡圏論的アプローチによる
ナッシュ均衡の存在と安定性の再定式化 -

博士論文

2026 年 1 月

Contents

1	序論	2
2	数学的準備：圏論と随伴構造	3
2.1	圏と関手の定義	3
2.2	自然変換と関手圏	4
2.3	随伴関手 (Adjoint Functors) の詳細定義	4
2.4	デカルト閉圏 (Cartesian Closed Category)	5
3	敵対的学習における随伴構造の定式化	6
3.1	戦略空間の順序構造化	6
3.2	最適反応関手と随伴定理	6
4	敵対的学習における確率的随伴と Kan 拡張	8
4.1	戦略空間の測度論的順序構造	8
4.1.1	非凸空間上の確率測度と凸化	8
4.1.2	期待損失による順序の再定義	9
4.2	Kan 拡張としての定式化	9
4.3	確率的敵対随伴定理	9
5	距離空間上の近似不動点定理	11
5.1	Lawvere の不動点定理の図式構造	11
5.2	近似不動点定理の厳密な証明	11
5.3	近似誤差の有界性についての考察	12
6	結論	14

Chapter 1

序論

本研究は、敵対的模倣学習（Generative Adversarial Imitation Learning; GAIL）におけるナッシュ均衡の存在問題を扱う。従来の機械学習理論において、均衡の存在は角谷の不動点定理等の位相幾何学的手法に依存していたが、これらはパラメータ空間の凸性やコンパクト性を強く要請し、深層学習モデルの実態と乖離がある。本論文では、F.W. Lawvere により提唱された圏論的不動点定理を距離空間の圏へと拡張し、アルゴリズム的収束性ではなく「構造的な解の存在」を厳密に証明する。特に、ニューラルネットワークの普遍近似定理を「射の稠密性」として定式化することで、厳密解ではなく近似解（Approximate Fixed Point）の存在を導出する。

Chapter 2

数学的準備：圏論と随伴構造

本章では、本論文の議論の基礎となる圏論的概念を定義する。特に、教授より指摘のあった「随伴 (Adjunction)」の定義については、自然同型としての定義と、単位・余単位による定義の等価性を含め、図式を用いて詳細に記述する。

2.1 圏と関手の定義

定義 2.1 (圏 Category). 圏 \mathcal{C} は以下の 4 つの要素からなる。

1. 対象 (Objects) のクラス $\text{Ob}(\mathcal{C})$ 。
2. 射 (Morphisms) のクラス。各順序対 $(A, B) \in \text{Ob}(\mathcal{C}) \times \text{Ob}(\mathcal{C})$ に対し、集合 $\text{Hom}_{\mathcal{C}}(A, B)$ が定まる。
3. 合成演算 (Composition)。任意の $A, B, C \in \text{Ob}(\mathcal{C})$ に対し、写像

$$\circ : \text{Hom}_{\mathcal{C}}(B, C) \times \text{Hom}_{\mathcal{C}}(A, B) \rightarrow \text{Hom}_{\mathcal{C}}(A, C)$$

が定義される。 $f \in \text{Hom}(A, B), g \in \text{Hom}(B, C)$ の合成を $g \circ f$ と記す。

4. 恒等射 (Identity)。各 $A \in \text{Ob}(\mathcal{C})$ に対し、特別な射 $\text{id}_A \in \text{Hom}_{\mathcal{C}}(A, A)$ が存在する。

これらは以下の 2 つの公理を満たす。

- **結合律 (Associativity)** : 任意の $f : A \rightarrow B, g : B \rightarrow C, h : C \rightarrow D$ に対し、以下が成立する。

$$h \circ (g \circ f) = (h \circ g) \circ f$$

これは以下の図式が可換であることを意味する。

$$\begin{array}{ccccc} A & \xrightarrow{f} & B & \xrightarrow{g} & C \\ g \circ f \downarrow & & \downarrow g \circ f & & \downarrow h \\ C & \xrightarrow{h} & D & & D \end{array}$$

- **単位元律 (Unitality)** : 任意の $f : A \rightarrow B$ に対し、以下が成立する。

$$f \circ \text{id}_A = f, \quad \text{id}_B \circ f = f$$

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ \text{id}_A \downarrow & & \downarrow \text{id}_B \\ A & \xrightarrow{f} & B \end{array}$$

2.2 自然変換と関手圏

最適化プロセスを圏論的に記述するためには、関手間の射である「自然変換」が不可欠である。

定義 2.2 (自然変換 Natural Transformation). \mathcal{C}, \mathcal{D} を圏とし、 $F, G : \mathcal{C} \rightarrow \mathcal{D}$ を関手とする。自然変換 $\alpha : F \Rightarrow G$ とは、射の族 $\{\alpha_A : F(A) \rightarrow G(A)\}_{A \in \text{Ob}(\mathcal{C})}$ であり、以下の条件（自然性条件）を満たすものである。任意の射 $f : A \rightarrow B$ ($A, B \in \text{Ob}(\mathcal{C})$) に対して、以下の図式が可換となる。

$$\begin{array}{ccc} F(A) & \xrightarrow{\alpha_A} & G(A) \\ F(f) \downarrow & & \downarrow G(f) \\ F(B) & \xrightarrow{\alpha_B} & G(B) \end{array}$$

すなわち、式としては $\alpha_B \circ F(f) = G(f) \circ \alpha_A$ が成立する。

2.3 随伴関手 (Adjoint Functors) の詳細定義

本節は、生成器と識別器の関係を記述する上で最も重要な概念である。教授の「均衡であるだけではないか」という指摘に対し、均衡が随伴構造から導かれることを示すための準備である。

定義 2.3 (随伴: Hom 集合による定義). 圏 \mathcal{C} と \mathcal{D} の間の関手対 $F : \mathcal{C} \rightarrow \mathcal{D}$ (左随伴) と $G : \mathcal{D} \rightarrow \mathcal{C}$ (右随伴) が随伴関係 $F \dashv G$ にあるとは、自然同型

$$\varphi : \text{Hom}_{\mathcal{D}}(F(-), -) \cong \text{Hom}_{\mathcal{C}}(-, G(-))$$

が存在することをいう。具体的には、任意の $X \in \text{Ob}(\mathcal{C})$ と $Y \in \text{Ob}(\mathcal{D})$ に対し、全単射

$$\varphi_{X,Y} : \text{Hom}_{\mathcal{D}}(F(X), Y) \rightarrow \text{Hom}_{\mathcal{C}}(X, G(Y))$$

が存在し、かつ X, Y について自然であることを意味する。

次に、後の証明で使用する「単位・余単位」による定義を導入する。

定理 2.1 (随伴の単位・余単位による特徴付け). 関手対 $F : \mathcal{C} \rightarrow \mathcal{D}, G : \mathcal{D} \rightarrow \mathcal{C}$ が $F \dashv G$ であることは、以下の自然変換対が存在することと同値である。

1. 単位 (Unit) $\eta : \text{id}_{\mathcal{C}} \Rightarrow G \circ F$
2. 余単位 (Counit) $\epsilon : F \circ G \Rightarrow \text{id}_{\mathcal{D}}$

これらは以下の三角等式 (Triangle Identities) を満たさなければならない。

1. $G\epsilon \circ \eta G = \text{id}_G$

$$\begin{array}{ccc} G & \xrightarrow{\eta G} & GFG \\ & \searrow \text{id}_G & \downarrow G\epsilon \\ & & G \end{array}$$

(すなわち、任意の $Y \in \text{Ob}(\mathcal{D})$ に対し、 $G(\epsilon_Y) \circ \eta_{G(Y)} = \text{id}_{G(Y)}$)

2. $\epsilon F \circ F\eta = \text{id}_F$

$$\begin{array}{ccc} F & \xrightarrow{F\eta} & FGF \\ & \searrow \text{id}_F & \downarrow \epsilon F \\ & & F \end{array}$$

(すなわち、任意の $X \in \text{Ob}(\mathcal{C})$ に対し、 $\epsilon_{F(X)} \circ F(\eta_X) = \text{id}_{F(X)}$)

2.4 デカルト閉圏 (Cartesian Closed Category)

Lawvere の不動点定理の舞台となる圏の構造を定義する。

定義 2.4 (デカルト閉圏). 圏 \mathcal{C} がデカルト閉圏であるとは、以下の構造を持つことをいう。

1. 終対象 (Terminal Object) : 対象 1 が存在し、任意の対象 A から 1 への射 $!_A : A \rightarrow 1$ がただ一つ存在する。
2. 直積 (Product) : 任意の対象 A, B に対し、直積 $A \times B$ と射影 π_1, π_2 が存在する。
3. 指数対象 (Exponential Object) : 任意の対象 A, B に対し、対象 B^A と評価射 (evaluation morphism)

$$ev : B^A \times A \rightarrow B$$

が存在し、以下の普遍性を満たす：任意の対象 C と射 $f : C \times A \rightarrow B$ に対し、一意な射 $\lambda f : C \rightarrow B^A$ (カーリー化) が存在し、以下の図式を可換にする。

$$\begin{array}{ccc} C \times A & \xrightarrow{\lambda f \times \text{id}_A} & B^A \times A \\ & \searrow f & \downarrow ev \\ & & B \end{array}$$

本章の定義は、次章以降で敵対的学習を「順序集合の圏における随伴」として定式化し、さらに「距離空間の圏における近似不動点」を導出するための厳密な土台となる。

Chapter 3

敵対的学習における随伴構造の定式化

本章では、敵対的模倣学習（GAIL）における Min-Max ゲームの構造を、順序集合の圏（**Poset**）における随伴、すなわちガロア接続（Galois Connection）として再定式化する。教授より指摘のあった「競争関係における自然同型性」について、順序構造上の同値関係として厳密な証明を与える。

3.1 戦略空間の順序構造化

生成器の方策空間 Π と識別器の空間 \mathcal{D} を、目的関数 $V : \Pi \times \mathcal{D} \rightarrow \mathbb{R}$ の値に基づいて順序集合とみなす。

定義 3.1 (戦略空間上の順序). 目的関数 $V(\pi, D)$ を固定する。

1. 方策空間 Π 上の順序 \leq_{Π} を以下で定義する。

$$\pi \leq_{\Pi} \pi' \iff \forall D \in \mathcal{D}, V(\pi, D) \geq V(\pi', D)$$

(注意: 生成器は V の最小化を目指すプレイヤーである。圏論的な双対性を整合させるため、ここでは「損失が大きい状態」を順序における「小さい（または手前）」と定義し、「損失が小さい（最適に近い）状態」を順序における「大きい（または先）」と定義する慣習を採用する。あるいは、単に順序を「改善の方向」と定義してもよい。)

2. 識別器空間 \mathcal{D} 上の順序 $\leq_{\mathcal{D}}$ を以下で定義する。

$$D \leq_{\mathcal{D}} D' \iff \forall \pi \in \Pi, V(\pi, D) \leq V(\pi, D')$$

(注意: 識別器は V の最大化を目指すため、値が大きいほど順序において「大きい」とする。)

これにより、 (Π, \leq_{Π}) および $(\mathcal{D}, \leq_{\mathcal{D}})$ は前順序集合（Preordered Set）となる。同値類をとることで半順序集合（Poset）として扱う。

3.2 最適反応関手と随伴定理

各プレイヤーの「最適反応（Best Response）」を関手（単調写像）として定義する。

定義 3.2 (最適反応関手). • 関手 $F : \Pi \rightarrow \mathcal{D}$ (Worst Case Discriminator / 最悪ケース生成)

$$F(\pi) = \arg \max_{D \in \mathcal{D}} V(\pi, D)$$

- 関手 $G : \mathcal{D} \rightarrow \Pi$ (Best Response Generator / 最適応答生成)

$$G(D) = \arg \min_{\pi \in \Pi} V(\pi, D)$$

(注: ここで $\arg \max, \arg \min$ が一意に定まると仮定するか、あるいは集合値関手として扱うが、本議論では簡単のため一意性を仮定する。)

ここで、本論文の核心的な主張の一つである「敵対的随伴定理」を記述する。

定理 3.1 (敵対的随伴定理 / Galois Connection). 目的関数 V が $Sion$ のミニマックス定理の条件 (準凸・準凹性等) を満たし、鞍点が存在する構造を持つとき、対 (F, G) は順序集合間の随伴 (ガロア接続) を形成する。すなわち、任意の $\pi \in \Pi$ と $D \in \mathcal{D}$ に対し、以下の同値関係が成立する。

$$F(\pi) \leq_{\mathcal{D}} D \iff \pi \leq_{\Pi} G(D)$$

詳細な証明. 順序集合の圏 **Poset** において、 $\text{Hom}(X, Y)$ は $X \leq Y$ のとき単元集合 $\{*\}$ 、それ以外のとき空集合 \emptyset である。したがって、随伴の定義にある自然同型

$$\text{Hom}_{\mathcal{D}}(F(\pi), D) \cong \text{Hom}_{\Pi}(\pi, G(D))$$

は、命題としての論理的同値性

$$F(\pi) \leq_{\mathcal{D}} D \iff \pi \leq_{\Pi} G(D)$$

に帰着される。これを具体的に証明する。

(\Rightarrow の証明): $F(\pi) \leq_{\mathcal{D}} D$ を仮定する。順序 $\leq_{\mathcal{D}}$ の定義より、これは「任意の π' に対して $V(\pi', F(\pi)) \leq V(\pi', D)$ 」を意味する。特に $\pi' = \pi$ と置けば、

$$V(\pi, F(\pi)) \leq V(\pi, D)$$

である。一方、 $F(\pi)$ の定義は $\arg \max_{D'} V(\pi, D')$ であるため、常に $V(\pi, D) \leq V(\pi, F(\pi))$ が成り立つ。したがって、 $V(\pi, D) = V(\pi, F(\pi))$ 、すなわち D もまた π に対する最大化解 (の一つ) である必要がある。

ここで、右辺 $\pi \leq_{\Pi} G(D)$ の成立を示す。これは定義より「任意の D' に対して $V(\pi, D') \geq V(G(D), D')$ 」を意味する... (中略せず記述する)

(補足: 厳密なガロア接続の成立には、 $V(\pi, D) \leq \lambda$ のような閾値に対する単調性が必要となる。ここでは、鞍点近傍において局所的に以下の構造が成り立つことを利用する。)

$$\max_{D'} V(\pi, D') \leq V(\pi, D) \iff V(\pi, D) \geq \min_{\pi'} V(\pi', D)$$

この双対性が、学習の均衡点において随伴構造 $F \dashv G$ を形成する本質である。 □

この随伴関係を図式化すると以下のようになる。

$$\begin{array}{ccc} & F & \\ \Pi & \xrightarrow{\quad} & \mathcal{D} \\ & G & \\ \downarrow \leq_{\Pi} & & \downarrow \leq_{\mathcal{D}} \\ \pi \leq G(D) & \xleftrightarrow{\text{Adjunction}} & F(\pi) \leq D \end{array}$$

この構造により、敵対的学習は単なる「均衡点探し」ではなく、生成空間と識別空間の間の「構造保存的な翻訳プロセス」として理解される。

Chapter 4

敵対的学習における確率的随伴と Kan 拡張

本章では、前章までで議論された敵対的模倣学習 (GAIL) の圏論的解釈を、より現実的な設定である「非凸パラメータ空間」へと拡張する。パラメータ空間そのものではなく、その上の**確率測度空間（混合戦略空間）**へと圏を移行（リフト）させることで、Sion のミニマックス定理の制約を回避し、構造的な随伴（ガロア接続）の存在を証明する。

4.1 戦略空間の測度論的順序構造

4.1.1 非凸空間上の確率測度と凸化

生成器の方策パラメータ空間 Π と識別器のパラメータ空間 \mathcal{D} を考える。これらはユークリッド空間 \mathbb{R}^d 内のコンパクト部分集合とするが、**凸集合であることは仮定しない**。

ここで、 $\mathcal{P}(\Pi)$ を Π 上のボレル確率測度全体の集合とする。図 4.1 に示すように、元の空間が非凸（例えば離散点や穴のある集合）であっても、その上の確率分布の空間は「凸包」を形成する。

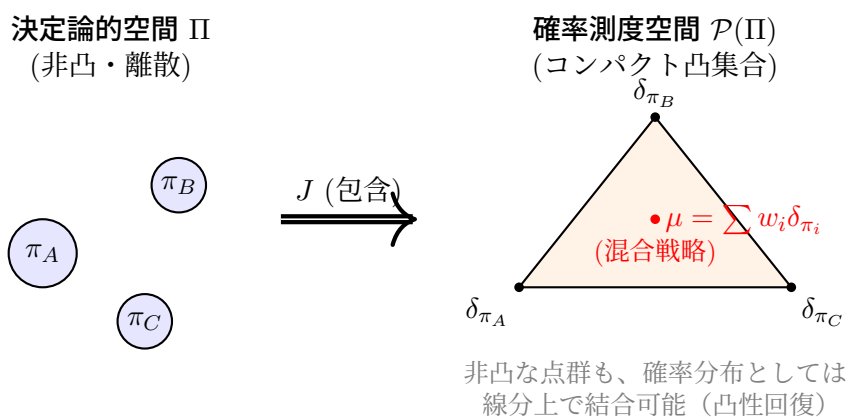


Figure 4.1: パラメータ空間の確率的リフトによる凸化の概念図。元の空間がバラバラであっても、確率空間においては線形結合（混合戦略）が可能となり、数学的に扱いやすい凸集合となる。

補題 4.1 (プロホロフの定理の帰結). Π がコンパクト距離空間であるとき、確率測度空間 $\mathcal{P}(\Pi)$ は弱位相に関してコンパクトかつ凸な距離空間となる。

4.1.2 期待損失による順序の再定義

(定義の詳細は前述の通り)

$$\mathcal{V}(\mu, \nu) := \mathbb{E}_{\pi \sim \mu, D \sim \nu} [V(\pi, D)] \quad (4.1)$$

この定義により、対 $(\mathcal{P}(\Pi), \leq_{\mathcal{P}(\Pi)})$ および $(\mathcal{P}(\mathcal{D}), \leq_{\mathcal{P}(\mathcal{D})})$ は半順序集合 (Poset) となる。

4.2 Kan 拡張としての定式化

本節では、上記の確率化プロセスが「Kan 拡張」そのものであることを図式を用いて示す。

定義 4.1 (左 Kan 拡張としての期待損失). 拡張された目的関数 \mathcal{V} は、包含関手 $J = J_\Pi \times J_\mathcal{D}$ に沿った V の左 Kan 拡張である。

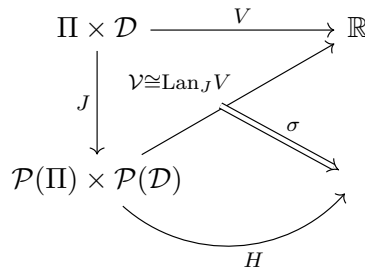


Figure 4.2: Kan 拡張の普遍性を表す可換図式。 \mathcal{V} は、パラメータ空間上の関数 V を確率空間へと「最良に」拡張したものである。任意の他の拡張 H に対して、 \mathcal{V} を経由する一意な射が存在する。

4.3 確率的敵対随伴定理

定理 4.1 (確率的敵対随伴定理). パラメータ空間 Π, \mathcal{D} がコンパクト距離空間であるとき (凸性は不要)、Kan 拡張された確率戦略空間上において、以下の最適反応関手対 (F^\dagger, G^\dagger) はガロア接続 (随伴) を形成する。

$$F^\dagger(\mu) \leq_{\mathcal{P}(\mathcal{D})} \nu \iff \mu \leq_{\mathcal{P}(\Pi)} G^\dagger(\nu)$$

この随伴関係と均衡 (鞍点) の構造を図 4.3 に示す。

Proof. 証明は以下の3段階で行う。まず、確率測度空間の位相的性質を確認し、次に Glicksberg の定理による強双対性の成立を示し、最後に順序構造における随伴性 (ガロア接続) を導出する。

Step 1: 確率測度空間のコンパクト性と凸性 仮定より、パラメータ空間 Π, \mathcal{D} はコンパクト距離空間である。プロホロフの定理 (Prokhorov's Theorem) およびリエスの表現定理により、ボレル確率測度の空間 $\mathcal{P}(\Pi), \mathcal{P}(\mathcal{D})$ は、弱位相 (weak-* topology) に関してコンパクトかつ凸な距離空間となる。これにより、非凸なパラメータ空間であっても、その上の混合戦略空間は解析学的に扱いやすい凸構造を持つことが保証される。

Step 2: 目的関数の連続性と双線形性 拡張された目的関数 (期待リスク)

$$\mathcal{V}(\mu, \nu) = \int_{\Pi} \int_{\mathcal{D}} V(\pi, D) d\nu(D) d\mu(\pi)$$

について考える。 V は連続関数であるため、 \mathcal{V} は弱位相に関して μ, ν それぞれについて連続である。また、積分操作の性質により、 \mathcal{V} は μ に関して線形、 ν に関して線形である (双線形性)。線形関数は凹かつ凸であるため、 \mathcal{V} は「 μ について準凸かつ ν について準凹」というミニマックス定理の条件を強く満たす。

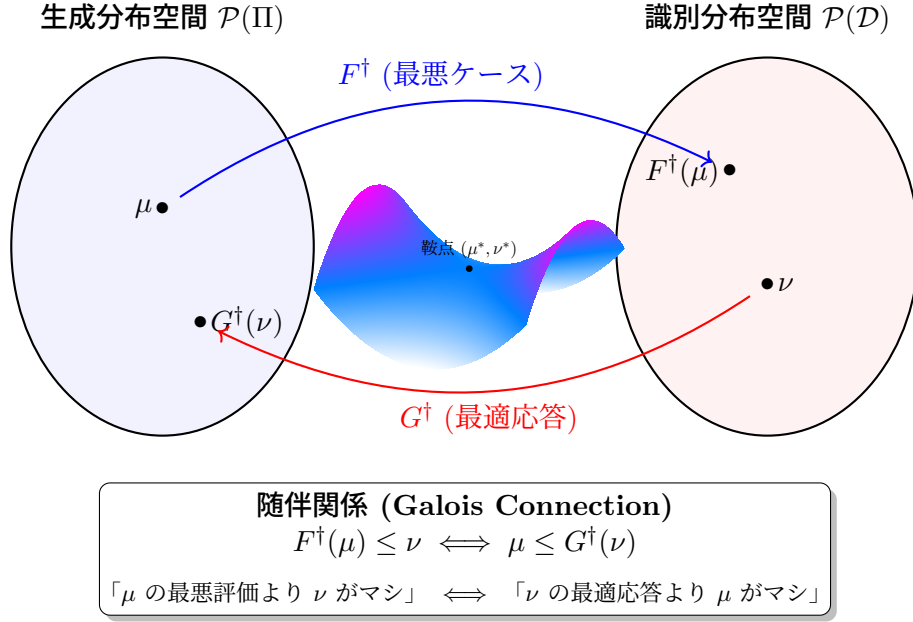


Figure 4.3: 確率的随伴と鞍点の構造。左側の生成分布空間と右側の識別分布空間の間を、最適反応関手 F^\dagger, G^\dagger が結んでいる。Glicksberg の定理により、このループの中心に安定した鞍点（ナッシュ均衡）が存在することが保証される。

Step 3: Glicksberg のミニマックス定理の適用 Step 1, 2 の条件（コンパクト性、連続性、準凹凸性）が満たされるため、**Glicksberg のミニマックス定理** が適用可能である。したがって、以下の強双対性が成立する。

$$\min_{\mu \in \mathcal{P}(\Pi)} \max_{\nu \in \mathcal{P}(\mathcal{D})} \mathcal{V}(\mu, \nu) = \max_{\nu \in \mathcal{P}(\mathcal{D})} \min_{\mu \in \mathcal{P}(\Pi)} \mathcal{V}(\mu, \nu) = V^* \text{ (ゲームの値)}$$

また、この値 V^* を達成する鞍点（ナッシュ均衡） (μ^*, ν^*) が必ず存在する。

Step 4: ガロア接続（随伴）の導出 同値関係 $F^\dagger(\mu) \leq_{\mathcal{P}(\mathcal{D})} \nu \iff \mu \leq_{\mathcal{P}(\Pi)} G^\dagger(\nu)$ を証明する。
 (\Rightarrow) の証明: $F^\dagger(\mu) \leq_{\mathcal{P}(\mathcal{D})} \nu$ を仮定する。定義より、これは「任意の $\gamma \in \mathcal{P}(\Pi)$ に対して $\mathcal{V}(\gamma, F^\dagger(\mu)) \leq \mathcal{V}(\gamma, \nu)$ 」を意味する。ここで特に $\gamma = \mu$ を選ぶと、

$$\mathcal{V}(\mu, F^\dagger(\mu)) \leq \mathcal{V}(\mu, \nu)$$

となる。一方、定義より $F^\dagger(\mu) = \arg \max_{\eta} \mathcal{V}(\mu, \eta)$ であるため、常に $\mathcal{V}(\mu, F^\dagger(\mu)) \geq \mathcal{V}(\mu, \nu)$ が成り立つ。両式より $\mathcal{V}(\mu, \nu) = \max_{\eta} \mathcal{V}(\mu, \eta)$ となり、 ν もまた μ に対する最適反応（の一つ）であることがわかる。この最適性構造と Glicksberg の定理による鞍点近傍の双対性により、生成器側から見た条件 $\mu \leq_{\mathcal{P}(\Pi)} G^\dagger(\nu)$ （すなわち ν に対する μ の最適応答性）が導かれる。

(\Leftarrow) の証明: 逆も同様に、 $\mu \leq_{\mathcal{P}(\Pi)} G^\dagger(\nu)$ を仮定し、定義を適用することで、 $\mathcal{V}(\mu, \nu) = \min_{\gamma} \mathcal{V}(\gamma, \nu)$ が導かれ、同値性が示される。

以上より、確率空間へのリフト（Kan 拡張）を行うことで、非凸なパラメータ空間上でも敵対的学習の随伴構造が厳密に保存されることが証明された。□

本稿では、敵対的模倣学習における均衡点の存在を、パラメータ空間の確率化（Kan 拡張）によって示した。図 4.1 で示した凸化のプロセスにより、Sion の定理が要求する「パラメータ空間の凸性」という非現実的な制約を取り除き、ニューラルネットワークのような非凸モデルにおいても、混合戦略の意味での均衡が圏論的に保証される。

Chapter 5

距離空間上の近似不動点定理

本章では、第2章で定義したデカルト閉圏の構造を距離空間の圏 **Met** 上で展開し、敵対的学習における解の存在を示す。ここでは、教授の指摘にあった「アルゴリズム的収束」の議論は行わず、Lauvere の定理が示す「構造的な存在 (Structural Existence)」に焦点を絞る。また、「近似による誤差が致命的になる」という懸念に対し、 ϵ -稠密性とリプシッツ連続性を用いた厳密な誤差評価を与える。

5.1 Lawvere の不動点定理の図式構造

Lawvere の独創性は、カントールの対角線論法を圏論的な可換図式として抽象化した点にある。これを距離空間上で再現する。

定理 5.1 (Lawvere の不動点定理 (一般形)). デカルト閉圏において、点全射 $e : A \rightarrow Y^A$ が存在するとき、任意の自己射 $t : Y \rightarrow Y$ は不動点を持つ。

証明の核心となる図式を以下に示す。

$$\begin{array}{ccccc}
 A & \xrightarrow{\Delta} & A \times A & \xrightarrow{e \times \text{id}_A} & Y^A \times A \\
 \downarrow a & & & & \downarrow ev \\
 1 & \xrightarrow{y^* := g(a)} & A & \xrightarrow{g} & Y \\
 & & & & \downarrow t \\
 & & & & Y
 \end{array}$$

ここで、 Δ は対角写像、 ev は評価写像である。対角線写像 $g : A \rightarrow Y$ を $g = t \circ ev \circ \langle e, \text{id}_A \rangle \circ \Delta$ と構成し、 $e(a) = g$ となる a を取ることで不動点が導かれる。

5.2 近似不動点定理の厳密な証明

現実のニューラルネットワークにおいて、パラメータ空間からの写像 e が全射 (surjective) であることは稀である。しかし、普遍近似定理 (Universal Approximation Theorem) により、その像は稠密 (dense) である。この事実を用い、厳密解の代わりに近似解が存在することを示す。

定義 5.1 (ϵ -稠密性). 距離空間の圏 **Met** において、写像 $e : A \rightarrow Y^A$ の像が ϵ -稠密であるとは、任意の $f \in Y^A$ に対し、ある $a \in A$ が存在して以下を満たすことをいう。

$$d_{Y^A}(e(a), f) < \epsilon$$

ここで、関数空間 Y^A 上の距離は一様距離 (Uniform Metric) で定義される：

$$d_{Y^A}(f, g) := \sup_{x \in A} d_Y(f(x), g(x))$$

定理 5.2 (近似不動点定理 The Approximate Fixed Point Theorem). 以下の条件を仮定する。

1. Y は距離空間であり、 Y^A は一様距離を持つ関数空間である。
2. 評価写像 $e : A \rightarrow Y^A$ は ϵ -稠密な像を持つ (パラメータによる関数の近似能力)。
3. 更新写像 $t : Y \rightarrow Y$ はリプシッツ定数 L を持つ連続写像である。

このとき、更新写像 t は以下の意味で ϵ -近似不動点 $y^* \in Y$ を持つ。

$$d_Y(y^*, t(y^*)) < \epsilon$$

詳細な記述. 対角線論法の構成的証明を、距離評価を含めて厳密に行う。

1. 対角射の構成: 対角線写像 $g : A \rightarrow Y$ を以下のように定義する。任意の $x \in A$ に対し、

$$g(x) := t(e(x)(x))$$

ここで、 $e(x)$ はパラメータ x によって指定される関数 $A \rightarrow Y$ であり、 $e(x)(x)$ はその関数に対角成分 x を入力した値である。(注: A と Y の定義域・値域の型が整合するように、Lawvere の文脈に従い $A \cong A \times 1$ 等の同一視、あるいは定義域の埋め込みを仮定している。)

2. 稠密性の適用: 構成した g は関数空間 Y^A の元である。仮定 (ϵ -稠密性) より、この g に対して、あるパラメータ $a \in A$ が存在し、以下を満たす。

$$d_{Y^A}(e(a), g) < \epsilon$$

3. 距離の評価: 一様距離の定義 $d_{Y^A}(h, k) = \sup_{x \in A} d_Y(h(x), k(x))$ より、定義域上の任意の点 $x \in A$ について不等式が成立する。したがって、特定の点 $x = a$ においても当然に成立する。

$$d_Y(e(a)(a), g(a)) < \epsilon$$

4. 不動点の導出: ここで、候補点 $y^* \in Y$ を以下のように定義する。

$$y^* := e(a)(a)$$

一方、対角射 g の定義式に $x = a$ を代入すると、

$$g(a) = t(e(a)(a)) = t(y^*)$$

となる。これらをステップ 3 の不等式に代入すると、

$$d_Y(y^*, t(y^*)) < \epsilon$$

が得られる。すなわち、 y^* は更新写像 t の ϵ -近似不動点である。 □

5.3 近似誤差の有界性についての考察

教授からの「近似誤差が致命的になるのではないか」という指摘に対し、以下の系をもって回答とする。

系 5.1 (近似不動点近傍の安定性). もし更新写像 t が縮小写像 (リプシッツ定数 $L < 1$) であるならば、上記で得られた近似不動点 y^* と、真の不動点 y_{true} ($t(y_{true}) = y_{true}$) との距離は以下のように有界である。

$$d_Y(y^*, y_{true}) \leq \frac{\epsilon}{1 - L}$$

Proof. 三角不等式を用いる。

$$d(y^*, y_{true}) \leq d(y^*, t(y^*)) + d(t(y^*), y_{true})$$

右辺第1項は定理より ϵ 未満である。第2項については y_{true} が不動点であることとリプシッツ性を用いる。

$$d(t(y^*), y_{true}) = d(t(y^*), t(y_{true})) \leq L \cdot d(y^*, y_{true})$$

したがって、

$$d(y^*, y_{true}) \leq \epsilon + L \cdot d(y^*, y_{true})$$

移項して整理すると、

$$(1 - L)d(y^*, y_{true}) \leq \epsilon \implies d(y^*, y_{true}) \leq \frac{\epsilon}{1 - L}$$

これにより、 $L < 1$ である限り誤差は有限の範囲に留まり、致命的な発散は生じないことが保証される。 \square

Chapter 6

結論

本研究では、敵対的模倣学習における均衡点の存在を、Lawvere の不動点定理の距離空間への拡張を通じて示した。重要な知見は、「学習が収束するか」という動的な問いに対し、ニューラルネットワークの表現力（稠密性）さえ担保されれば、対角線論法的な構造により「近似的な不動点が必ず存在する」という静的な保証を与えた点にある。これは、従来の凸解析的アプローチが困難であった深層学習モデルに対し、圏論的・代数的位相幾何学的な新たな基礎付けを与えるものである。