

敵対的生成ネットワークにおけるゲーム理論に関する一考察

Generative Adversarial Network with Game Theory

吉 川 満
KIKKAWA Mitsuru

要 旨

本稿では、深層学習の分野で現在注目されている、敵対的生成ネットワークをゲーム理論の立場で取り上げた。そこで、まず敵対的生成ネットワークとゲーム理論における等価性を示した。次に、敵対的生成ネットワークにおいては、学習の安定性は重要な課題の一つであり、ゲーム理論の研究を利用することにより、安定性を予想することができる。

Abstract

This paper examines the Generative Adversarial Network, which is developing in Deep Learning, with game theory. We first showed equivalence in the Generative Adversarial Network and game theory under some conditions. Second, in the Generative Adversarial Network, the stability of learning is one of unresolved issues, and the local stability of the equilibrium can be predicted with game theory.

キーワード：敵対的生成ネットワーク、進化ゲーム理論、Replicator方程式、ゼロサムゲーム、Lyapunov関数

Keywords：Generative Adversarial Network, Evolutionary Game Theory, Replicator equation, Zero-Sum Game, Lyapunov function

1. はじめに

近年、人工知能が注目されている。その中には、ゲーム理論の考えが利用されているものもある。例えば、将棋や囲碁などゲームなどでは、ミニマックス法が利用されている。これらのゲームは、ゼロサムゲームであることから、自分は自分の点数を最大化(MAX)する手を指し、相手はこちらの得点を最小化(MIN)する手を指すということが成り立つとすると、最善の手が決まることから、相手の行動の予測に用いられている。また、Leibo, et al. [13] では、ビデオゲームで囚人のジレンマ環境を作り出し、そこにマルチエージェントを置いて深層強化学習をしたら、協力関係が生成されるという。これら以外にも、様々なゲーム理論と関連する研究が発表されている。

特に本稿では、敵対的生成ネットワーク(Generative Adversarial Network, 以下、GAN)を取り上げ、ゲーム理論として、捉え直す。これは Goodfellow, et al. [4]によって最初に提案され、現在非常に注目を集めている分野で、次々に関連研究が発表されている¹。この GAN は実務的には学習されず、画像が生成されないことがあり、学習の安定化が問題が課題となっている。試行錯誤が必要であるため、理論的に最適化により(局所)Nash 均衡に到達するよう保証する研究が行われている²。

また、Finn, et al. [2], Ho and Ermon [9]では、目標となる行動から模倣学習という逆強化学習と似ていることから、逆強化学習とこの GAN と損失関数の等価性を示している。

その他、計量経済学関連の研究として、Igami [11]では、Bonanza や AlphaGo を動学的離散選択問題と

して、(ミクロ)計量経済分析等で用いられている代表的な推定法から捉え直している研究もある。

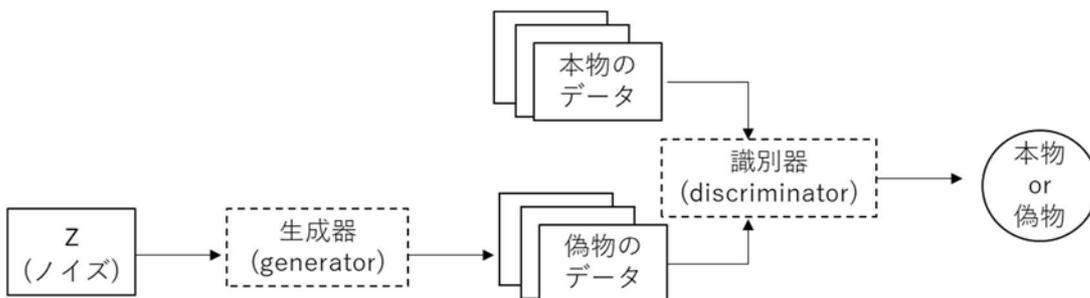
本稿は、GAN をゲーム理論として捉え直し、ゲーム理論との等価性を論じた。次のように構成されている。第2節で GAN をまとめ、第3節で、この GAN をゲーム理論から捉え直し、得られた数学的結果について記し、第4節で、まとめ、残された課題を記す。

2. 敵対的生成ネットワーク (GAN)

2.1 理論モデル

GAN は「教師なし機械学習」で使用される人工知能アルゴリズムの一種であり、ゼロサムゲームの考え方で互いに競合する2つのニューラルネットワーク (生成器ネットワーク(generator)と識別器ネットワーク(discriminator)) によって構成されている。例えば、画像生成を目的とするなら生成器が画像を生成・出力し、識別器側がその正否(本物か偽物)を判定する(図1)。生成器は本物と非常に良く似た画像を生成して、識別器側を騙すように学習し、識別器はより正確に識別しようと学習していく。最終的には、識別器側が本物の画像と人工的に作成された偽物の画像とを区別できなくなるまで、このプロセスは繰り返される。そのため最終的には、訓練データと生成データを見分けることができなくなるため、識別器の正答率は 50% になる。このように2つのネットワークが相反した目的のもとに学習するため敵対的と呼称される所以である。

図1 GAN の概念図



¹ 例えば、これに関連する研究は GAN Zoo (<https://deephunt.in/the-gan-zoo-79597dc8c347>)としてまとめられており、この図からも一連の研究が増加していることが分かる。また、深層学習の研究分野における権威の1人である Yann LeCun は、GAN はこの10年の機械学習で最も面白いアイデアであるとも述べている(“GANs are the most interesting idea in the last 10 years in ML”).

² 例えば、Heusel, et al. [7], Kodali, et al. [13], Nagarajan, et al. [15]などがある。

次に, Goodfellow, et al. [4]の理論的な内容の詳細について, 簡潔にまとめる。ここで, G は生成器(generator), D は識別器(discriminator), x は訓練データ, z はノイズとする。生成器 G はノイズ z を入力としてデータを生成する。生成器 G の訓練手順は, 識別器 D が誤っている確率を最大にすることであり, $D(x)$ は, そのデータが訓練データである確率を表す。識別器 D は訓練データと生成データに対して正しくラベル付けを行う確率を最大化しようとする。一方, 生成器 G は $\log(1 - D(x))$ を最小化しようとする。これらをまとめて以下のように表現している。

$$\min_G \max_D V(D, G) = E_{x \sim p_{data}(x)} [\log D(x)] + E_{z \sim p_z(z)} [\log(1 - D(G(z)))] \quad (1)$$

識別器 D が上手に分類できるようになると, $D(x)$ が大きくなり, $\log D(x)$ が大きくなる。また, 偽物だとばれて $D(G(z))$ は小さくなるため, $\log(1 - D(G(z)))$ は大きくなる。一方, 生成器 G が訓練データに似ているものを生成できるようになると, 識別器 D が上手に分類できなくなるため $D(G(z))$ は大きくなり, $\log(1 - D(G(z)))$ は小さくなるという構造になっている。

また, この(1)の関数は, 交差エントロピーであり, 生成器と識別器の持つ確率分布関数の違い・差がどの程度あるのかを示す。

学習時は, 次のアルゴリズムのように識別器と生成器を交互に更新していく。

図 2 アルゴリズム 1 ミニバッチ³確率的勾配降下法 (Goodfellow, et al. [4])

- (1) ミニバッチサイズ m 個のノイズ $\{z^{(1)}, \dots, z^{(m)}\}$ を $p_z(z)$ から取り出す(生成する)
- (2) ミニバッチサイズ m 個のサンプル $\{x^{(1)}, \dots, x^{(m)}\}$ をデータ生成分布 $p_{data}(x)$ から取り出す
- (3) 下記式の θ_d における確率的勾配を上るように識別器 D を更新する

$$\nabla_{\theta_d} \frac{1}{m} \sum_{i=1}^m [\log D(x^{(i)}) + \log(1 - D(G(z^{(i)})))]$$

- (4) 上記までを k 回繰り返す
- (5) ミニバッチサイズ m 個のノイズ $\{z^{(1)}, \dots, z^{(m)}\}$ を $p_z(z)$ から取り出す
- (6) 下記式の θ_g における確率的勾配を下るように生成器 G を更新する

$$\nabla_{\theta_g} \frac{1}{m} \sum_{i=1}^m \log(1 - D(G(z^{(i)})))$$

- (7) ここまで全てを, 訓練回数分だけ繰り返す

識別器 D を十分な回数(k 回)更新した上で生成器 G を 1 回更新することで, 常に識別器が新しい G の状態に適用できるように学習を進める

また, Goodfellow, et al. [4]では次のことを証明している。

³ ミニバッチ学習法は, 1 度に複数の学習サンプルを用いて学習を行う手法であり, 深層学習においてよく利用されている。

命題 1 (Goodfellow, et al. [4]) 固定した生成器 G において、最適な識別器 D は、次のようである。

$$D_G^*(x) = \frac{p_{data}(x)}{p_{data}(x) + p_g(x)}$$

ここで、 $C(G)$ を次のように置く。

$$\begin{aligned} C(G) &= \max_D V(G, D) = E_{x \sim p_{data}} [\log D_G^*(x)] + E_{z \sim p_z} [\log (1 - D_G^*(G(z)))] \\ &= E_{x \sim p_{data}} [\log D_G^*(x)] + E_{x \sim p_g} [\log (1 - D_G^*(x))] \\ &= E_{x \sim p_{data}} \left[\log \frac{p_{data}(x)}{p_{data}(x) + p_g(x)} \right] + E_{x \sim p_g} \left[\log \frac{p_g(x)}{p_{data}(x) + p_g(x)} \right] \end{aligned}$$

定理 1 (Goodfellow, et al. [4]) 仮想的訓練基準 $C(G)$ の大域最小値は、 $p_g = p_{data}$ のとき、得られる。また、その時の $C(G)$ の値は $-\log 4$ である。

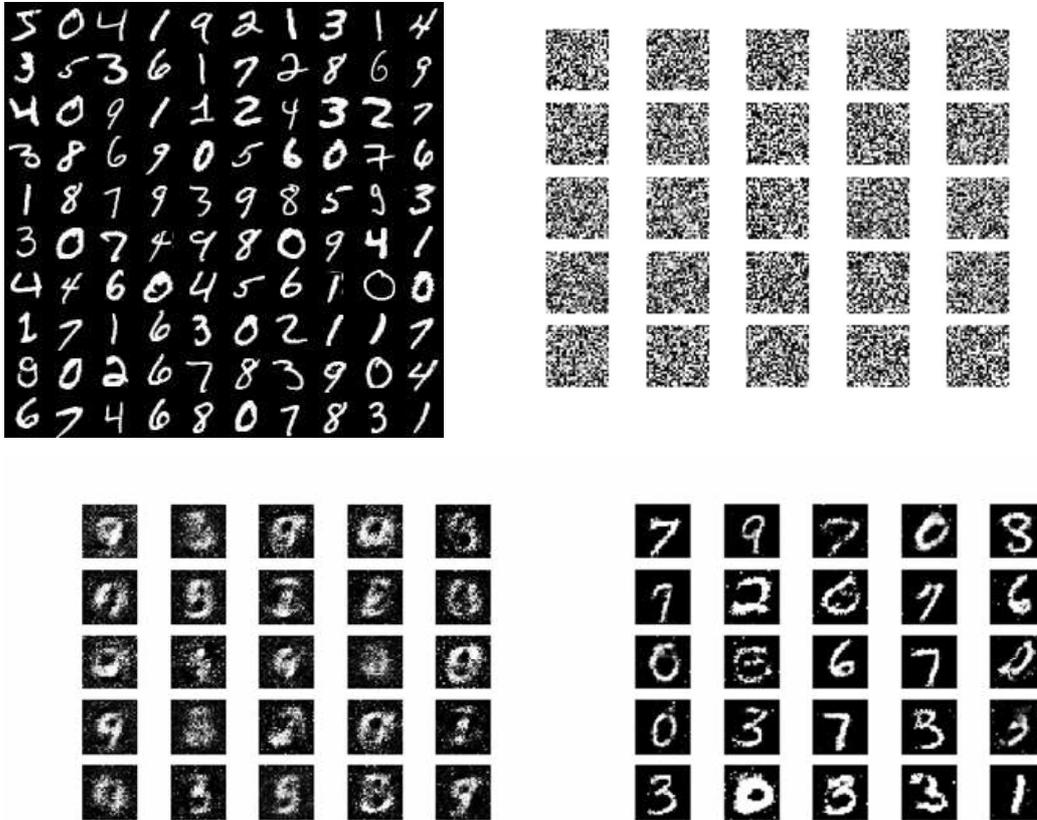
命題 2 (Goodfellow, et al. [4]) 生成器 G と識別器 D に十分な容量があり、アルゴリズム 1 の各ステップで、識別器はその与えられた生成器 G のもとで、最適に到達することを許され、基準を改善するように更新され、条件、 $E_{x \sim p_{data}} [\log D_G^*(x)] + E_{x \sim p_g} [\log (1 - D_G^*(x))]$ を満たすとき、 p_g は p_{data} に収束する。

2.2 実験結果

Goodfellow, et al. [4]と同様に MNIST のデータを GAN で実装し、画像を生成させた⁴。図 3 にあるように、学習ステップが増加するにつれて、数字が鮮明に作成されていることが分かる。

⁴ MNIST のデータは、機械学習の分野でよく利用されている手書き文字の認識のためのサンプルデータで、0-9 のいずれかの数字が学習用、テスト用それぞれで 60,000 枚、10,000 枚含まれている。実験で使用したコード・プログラムは、Python、ニューラルネットワークライブラリー Keras を用いて作成された。GitHub (<https://github.com/mitsurukikkawa/Python/tree/master/GAN>)にあり、このコードは、Keras GAN (<https://github.com/eriklindernoren/Keras-GAN/tree/master/gan>)をベースにしている。

図3 実験結果 (上段左: 学習用データ, 右: 0ステップ, 下段左: 1,000ステップ後, 右: 30,000ステップ後)



3. ゲーム理論からのアプローチ

本節では、GAN をゲーム理論、特に進化ゲーム理論の立場で捉える。GAN の理論モデルは生成器 G と識別器 D の間には、ゲーム的状况にあり、上述のようにプレイヤーの目的は生成器 G は識別器 D を騙すように x を生成し、識別器 D は生成器 G に騙されないように正しく識別するようにと正反対であるため、直感的に 2 人ゼロサムゲームと捉えることができる⁵。

そこで本稿では、進化ゲーム理論を用いて、純粋戦略の数が 2 つの場合の非対称 2 人ゲームとして捉え、GAN とゲーム理論との等価性について示す。

3.1 命題 1

Goodfellow, et al. [4] の命題 1 は、 $V(G, D)$ を最大化しており、最適識別器 D に関しては、 $p_g(x), p_{data}(x)$

⁵ Goodfellow, et al. [3] には、利得関数 $v(\theta^{(g)}, \theta^{(d)})$, $-v(\theta^{(g)}, \theta^{(d)})$ が記載されているが、元論文 Goodfellow, et al. [4] には、明示的な利得表の記述はない。

をゲーム理論における利得と捉えると、混合戦略の Nash 均衡の確率に関するものと等価である⁶。ゲーム理論においては、各プレイヤーの最適反応戦略を導出(各プレイヤーの期待利得の最大化を行うこと)し、全てのプレイヤーの最適反応戦略を取り合う戦略の組を(Nash)均衡と呼ぶ。この方法で行くと、命題 1 と同様のものが導出される。

3.2 定理 1

Goodfellow, et al. [4]の定理 1 は、交差エントロピー $C(G)$ の最大値に関するものである。この $C(G)$ は、進化ゲーム理論においては次の命題にあるように Lyapunov 関数、第 1 積分に対応する。特に本稿では、GAN はプレイヤータイプが 2 種類、戦略の数が 2 つでゼロサムゲーム(利得表 1)に対応するので、このゲームの Replicator 方程式は、次のようになる。これの導出に関しては、付録 A を参照されたい。

$$\dot{y} = y(1-y)\{a - (a+b)x\}, \quad \dot{x} = x(1-x)\{-b + (a+b)y\} \quad (2)$$

ただし y をプレイヤー 1 が戦略 1 をとる確率、 x をプレイヤー 2 が戦略 2 をとる確率とする。

利得表 1 正規化したゼロサムゲームの利得表

	戦略 1	戦略 2
戦略 1	a, -a	0, 0
戦略 2	0, 0	b, -b

そこで Lyapunov 関数・第 1 積分を Hofbauer and Sigmund [10], 大浦[17], 吉川[12] などにあるような方法で導出すると、次が得られる。導出に関しては付録 A を参照されたい。

$$H = -2a \log x - 2b \log(1-x) \quad (3)$$

(1)式における $E_{x \sim p_{data}(x)}[\log D(x)]$, $E_{z \sim p_z(z)}[\log(1-D(z))]$ は、それぞれ期待利得 $-2a \log x$, $-2b \log(1-x)$ に対応し、GAN の評価関数(1)と、 a, b が正の値を持つとき等価となる⁷。上記をまとめると、次の命題となる。

命題 3 Replicator 方程式を用いた進化ゲーム理論におけるゼロサムゲーム(利得表 1)の大域関数、例えば一般的な Lyapunov 関数は(3)式のように導出することができる。また、GAN の評価関数(1)と、 a, b が正の値を持つとき等価となる。

注 1 $a = b = \frac{1}{2}$ とする時、式(3)が最大値をとるのは、確率 x が $\frac{1}{2}$ の時であり、またその時の値は $-\log 4$ である。

⁶ 通常のゲーム理論において、プレイヤー 1 がある戦略を採用する確率とプレイヤー 2 がある戦略を採用するという確率の 2 変数の枠組みであった。しかし本稿のモデルでは、プレイヤー 1(生成器)とプレイヤー 2(識別器)の戦略を採用する確率は、生成器が識別器を騙す確率が高ければ、正しく識別する確率は小さくなる。つまり、ある一方のプレイヤーにおける確率が多ければ、もう一方は少なくなるといった相反するといった $y = 1 - x$ という関係にある。

⁷ ここでは x は確率であるため、 $\log x$ はマイナスの値を持つ。

る。

この進化ゲーム理論は動的な理論体系であるため、どの戦略の組が Nash 均衡となるのかという均衡選択に関する研究が豊富にある。そのため GAN において、この進化ゲーム理論の研究を活用すると、GAN において均衡選択、安定性について予測することが可能となる。

命題 4 このゲームにおける内点解(混合戦略)の局所安定性は鞍点であり、不安定である。

証明 付録 B

注 2 (Akin and Losert [1], Hirsch, et al. [8], 岡田 [17])

非協力ゲーム理論において、ゼロサムゲームにおける混合戦略(内点解)の Nash 均衡は鞍点であり、Lyapunov 安定である。

Replicator 方程式は、非線形微分方程式であり、純粋戦略が 3 つ以上の場合には、様々な挙動を示すことが知られている。

例 1 Taylor and Jonker [22]は純粋戦略の数が 3 つの場合、混合戦略が漸近安定となるゲームを示した。また、Skyrms [21]は戦略が 4 つある場合には、ストレンジアトラクター(Hopf 分岐を通じて生まれる)が生じ得ることを示している。

例 1 から進化ゲーム理論においては、内点解が安定な場合が存在するため、GAN を拡張し、条件によっては、安定的な内点解が存在すると推測することができる。

3.3 命題 2

Goodfellow, et al. [4]の命題 2 はアルゴリズム 1 の収束性について記述している。ゲーム理論においては、上記の Replicator 方程式の他、フィクティシヤス・プレイ、最適反応動学など様々な学習方法が使われている。収束スピードが異なるが、学習プロセスに関して、Nash 均衡が異なるということはなく、基本的には同じ挙動を示している⁸。

4. まとめ

以上のように、ゲーム理論の立場から GAN を考察することにより、次のことが分かった。

(1) 命題 3 により、GAN の評価関数は、ある条件の下で、進化ゲーム理論における Lyapunov 関数と等

⁸ 例えば、Sandholm[19]などを参照されたい。

価である。

(2) 進化ゲーム理論に関する研究を利用することで、GAN における学習の安定性に予測することが可能となる。

本稿における残された課題として、ゲーム理論においてある条件の下では、安定な内点解が存在するため、実際に GAN における安定的な学習を行うような実験結果を示すことである。

付録 A Lyapunov 関数の導出

まず個体数の変化を取り扱う Lotka-Volterra 方程式から Replicator 方程式を導出し、次に Lyapunov 関数の導出を行う。

あるプレイヤーの集団があり、ゲーム的な状況において、プレイヤーの戦略 i の頻度を p_i 、戦略 i を採用する確率を x_i 、戦略を採用する確率の分布をベクトル $x = (x_1, x_2, \dots, x_n)$, $n < \infty$ とする。

各プレイヤーは、ある戦略を採用し、その結果相対的により利得が得られるのであれば、その戦略を採用する頻度が増加すると仮定する。戦略の頻度 p_i の変化率を r_i とすると、このことは次のように表すことができる。

$$\dot{p}_i = r_i p_i, \quad i = 1, \dots, n$$

ここで r_i を戦略に無関係な成分と戦略に依存する部分に分解する。戦略に無関係な部分を r 、戦略に依存する部分を u_i とする。 u_i は一般に戦略の分布 x にも依存するので、 $u_i(x)$ と書くことができ、 $u_i(x)$ はゲーム理論という「利得」に対応する。そのため、ゲームにおいて各戦略の頻度 p_i ($i = 1, \dots, n$) の変化法則は次のように書き換えることができる。

$$\dot{p}_i = r_i (r + u_i(x)) p_i, \quad i = 1, \dots, n \quad (4)$$

この式は各戦略の頻度の変化を表したものであり、戦略の頻度の変化率 \dot{p}_i は自然増加率 r に各戦略の利得 $u_i(x)$ を加えたものとなる。

これを各戦略が採用される確率、 $x_i = \frac{p_i}{p}$ に置き換える。ただし、 $p = \sum_{i=1}^n p_i$ であり、十分大きな数とする。恒等式 $p x_i = p_i$ の両辺の時間微分をとると、 $\dot{p}_i = \dot{x}_i p + x_i \dot{p}$ が得られる。また、 $\dot{p} = \sum_{j=1}^n \dot{p}_j$ であるので、これと(4)式から整理すると、次が得られる。これは Replicator 方程式と呼ばれている。

$$\dot{x}_i = (u_i(x) - \sum_{j=1}^n u_j x_j) x_i, \quad i, j = 1, \dots, n \quad (5)$$

戦略の採用確率の変化率 $\frac{\dot{x}_i}{x_i}$ は、その時点での各戦略の採用確率 x_i の下で、その戦略者が得る利得 $u(x) := \sum_{j=1}^n u_j x_j$ と全ての戦略における平均利得 $u(x)$ との差によって定義される。この式から平均利得より高い利得が得られる戦略の採用確率は増加し、平均より低い利得しか得られない戦略の採用確率は減少する。また、増加や減少の度合い・スピードは、各戦略の利得の平均からの乖離度に比例する。つまり、平均より高い利得が得られる戦略はより急速に増加することとなる。

次に本稿では、生成器、識別器というプレイヤー種類が 2 つ、戦略の数が 2 つの非対称 2 人ゲームを考えている。特に利得表 A1 で表現される一般的なゲーム状況を表す利得表の場合、このゲームの Replicator 方程式は、次のようになる。

利得表 A1 非対称 2 人ゲーム

	戦略 1	戦略 2
戦略 1	f_1, g_1	f_3, g_3
戦略 2	f_2, g_2	f_4, g_4

$$\dot{y} = y(1-y)\{f_1 - f_2 + x(f_3 - f_4 - f_1 + f_2)\}$$

$$\dot{x} = x(1-x)\{g_4 - g_2 + y(g_3 - g_4 - g_1 + g_2)\}$$

ただし, y をプレイヤー 1 が戦略 1 をとる確率, x をプレイヤー 2 が戦略 2 をとる確率とする。ここで, $f_1 - f_2 = a, g_4 - g_2 = d, f_4 - f_3 = c, g_1 - g_3 = b$ と置くと, 先ほどの Replicator 方程式は次のようになり, この一般的な非対称 2 人ゲームは利得表 A2 に記すことができる⁹。

利得表 A2 正規化した利得表

	戦略 1	戦略 2
戦略 1	a, b	0, 0
戦略 2	0, 0	c, d

$$\dot{y} = y(1-y)\{a - (a+c)x\},$$

$$\dot{x} = x(1-x)\{d - (b+d)y\}$$
 (6)

そこで Lyapunov 関数・第 1 積分を Hofbauer[10], 吉川[12], 大浦[18]などにあるように導出すると, 次が得られる。

$$H(x,y) = \log \frac{y^d(1-y)^b}{x^a(1-x)^c} = \log x^{-a}y^d + \log(1-x)^{-c}(1-y)^b$$
 (7)

特に, 本稿では, $x+y=1, a=-b, c=-d$ (ゼロサムゲーム, 利得表 1) が満たすゲームを考えているので, 次のように変形することができる。

$$H = -2a \log x - 2c \log(1-x)$$
 (8)

(1)式における $E_{x \sim p_{data}(x)}[\log D(x)], E_{z \sim p_z(z)}[\log(1-D(z))]$ は, それぞれ期待利得 $-2a \log x, -2c \log(1-x)$ に対応し, GAN の評価関数(1)と, a, c が正の値を持つとき等価となる。

付録 B 局所安定性

Replicator 方程式(2)の(Nash)均衡点の局所安定性を調べる。(2)において, $\dot{x} = 0, \dot{y} = 0$ となる (y, x) の組みを (y^*, x^*) と置くと, 均衡点は次の 5 点である。

$$(y^*, x^*) = (0,0), (0,1), (1,0), (1,1), \left(\frac{b}{a+b}, \frac{a}{a+b}\right)$$

⁹ 一般的な非対称 2 人ゲーム(利得表 A2)を a, b, c, d の符号からゲームを分類すると, 次のように 4 つに分類分けすることができる (Weilbull [23])。

- (I) 非ジレンマ, 囚人のジレンマ型ゲーム($ac < 0, bd < 0$) 純粋戦略の Nash 均衡は 1 つ存在する
- (II) コーディネーション型ゲーム($a < 0, b > 0, c > 0, d > 0$) 純粋戦略の Nash 均衡は 2 つ $(y^*, x^*) = (0,1), (1,0)$ 存在する
- (III) チキン型ゲーム($a < 0, b < 0, c < 0, d < 0$) 純粋戦略の Nash 均衡は 2 つ $(y^*, x^*) = (0,0), (1,1)$ 存在する
- (IV) マッチング・ペニー型ゲーム($ab < 0, cd < 0, ac > 0, bd < 0$) 純粋戦略の Nash 均衡は存在しない

ただし、内点解が存在するためには、 $0 \leq \frac{b}{a+b} \leq 1$, $0 \leq \frac{a}{a+b} \leq 1$ を満たす必要である。

次に内点解・混合戦略の局所安定性を考える。(2)の2行2列のJacobi行列 $J(y, x)$ は、次のようになる。

このJacobi行列に内点解の値を代入し、固有値を求めること局所安定性は分かる。

$$J(y, x) = \begin{pmatrix} \frac{\partial y}{\partial x} & \frac{\partial y}{\partial y} \\ \frac{\partial x}{\partial x} & \frac{\partial x}{\partial y} \end{pmatrix} = \begin{pmatrix} -(a+b)y(1-y) & (1-2y)\{a - (a+b)x\} \\ (1-2x)\{-b + (a+b)y\} & (a+b)x(1-x) \end{pmatrix} \quad (9)$$

内点均衡、混合戦略の均衡 $(y^*, x^*) = (\frac{b}{a+b}, \frac{a}{a+b})$ の局所安定性を考える。このときのJacobi行列は、次のようになる。

$$J\left(\frac{b}{a+b}, \frac{a}{a+b}\right) = \begin{pmatrix} -\frac{ab}{a+b} & 0 \\ 0 & \frac{ab}{a+b} \end{pmatrix}$$

$(\lambda - \frac{ab}{a+b})(\lambda + \frac{ab}{a+b}) = 0$ からこのJacobi行列の固有値は、 $\lambda = \pm \frac{ab}{a+b}$ となり、鞍点となる。

注3 f-GAN (Nowozin, et al. [16])では、Jeansen-Shannon divergence・交差エントロピーから Kullback-Leibler divergence・相対エントロピーに目的関数を変更し、内点解が鞍点である条件を導出している。ゲーム理論においても、この Kullback-Leibler divergence・相対エントロピーを Lyapunov 関数として取り扱うものもある(Sandholm [19])。

参考文献

- [1] Akin, Ethan and Viktor Losert (1984) "Evolutionary dynamics of zero-sum games," *Journal of Mathematical Biology*, Vol. 20, pp. 231-258
- [2] Finn, Chelsea, Paul Christiano, Pieter Abbeel and Sergey Levine (2016) "A Connection Between Generative Adversarial Networks, Inverse Reinforcement Learning, and Energy-Based Models," <https://arxiv.org/abs/1611.03852>
- [3] Goodfellow, Ian, Yoshua Bengio and Aaron Courville, *Deep Learning*, MIT Press, 2016 (訳「深層学習」KADOKAWA)
- [4] Goodfellow, Ian J., Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville and Yoshua Bengio (2014) "Generative Adversarial Nets," <https://arxiv.org/abs/1406.2661>
- [5] Goodfellow, Ian J. (2016) "NIPS 2016 Tutorial: Generative Adversarial Networks," <https://arxiv.org/abs/1701.00160>
- [6] Hartford, Jason, James R. Wright and Kevin Leyton-Brown (2016) "Deep Learning for Predicting Human Strategic Behavior," *Advances in Neural Information Processing Systems 29 (NIPS 2016)*
- [7] Heusel, Martin, Hubert Ramsauer, Thomas Unterthiner, Bernhard Nessler and Sepp Hochreiter (2017) "GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium," <https://arxiv.org/abs/1706.08500>
- [8] Hirsch, Morris W., Stephen Smale and Robert L.Devaney, *Differential Equations, Dynamical*

- Systems and an Introduction to Chaos*, 3rd edition, Academic Press, 2012 (訳「力学系入門 第3版」共立出版)
- [9] Ho, Jonathan and Stefano Ermon (2016) “Generative Adversarial Imitation Learning,” <https://arxiv.org/abs/1606.03476>
- [10] Hofbauer, Josef and Karl Sigmund, *Evolutionary Games and Population Dynamics*, Cambridge University Press, 1998 (訳「進化ゲームと微分方程式」現代数学社)
- [11] Igami, Mitsuru (2017) “Artificial Intelligence as Structural Estimation: Economic Interpretations of Deep Blue, Bonanza, and AlphaGo,” <https://arxiv.org/abs/1710.10967>
- [12] 吉川満 (2007) 「非対称2人ゲームにおける漸近安定な均衡の発生とその変化」進化経済学論集 第11集
- [13] Kodali, Naveen, Jacob Abernethy, James Hays and Zsolt Kira (2017) “On Convergence and Stability of GANs,” <https://arxiv.org/abs/1705.07215>
- [14] Leibo, Joel Z., Vinicius Zambaldi, Marc Lanctot, Janusz Marecki and Thore Graepel (2017) “Multi-agent Reinforcement Learning in Sequential Social Dilemmas,” Proceeding AAMAS '17 Proceedings of the 16th Conference on Autonomous Agents and Multiagent Systems, pp. 464-473 <https://arxiv.org/abs/1702.03037>
- [15] Nagarajan, Vaishnavh and J. Zico Kolter (2017) “Gradient descent GAN optimization is locally stable,” <https://arxiv.org/abs/1706.04156>
- [16] Nowozin, Sebastian, Botond Cseke and Ryota Tomioka (2016) “f-GAN: Training Generative Neural Samplers using Variational Divergence Minimization,” <https://arxiv.org/abs/1606.00709>
- [17] 岡田章「ゲーム理論 新版」有斐閣, 2011年
- [18] 大浦宏邦「社会科学者のための進化ゲーム理論」勁草書房, 2008年
- [19] Sandholm, William H., *Population Games and Evolutionary Dynamics*, MIT Press, 2010
- [20] Schuurmans, Dale and Martin A. Zinkevich (2016) “Deep Learning Games,” Advances in Neural Information Processing Systems 29 (NIPS 2016)
- [21] Skyrms, Brian(1992) “Chaos in game dynamics,” Journal of Logic, Language, and Information, Vol.1, pp. 111-130
- [22] Taylor, Peter D. and Leo B. Jonker (1978) “Evolutionary stable strategies and game dynamics,” Mathematical Biosciences, Vol. 40, pp. 145-156
- [23] Weibull, Jörgen W., *Evolutionary Game Theory*, MIT Press, 1995 (訳「進化ゲームの理論」オフィスカノウチ)