連続行為空間上の進化ゲームならびに関連した平均場ゲーム

(Evolutionary game on continuous action space and related mean field game)

北陸先端科学技術大学院大学　吉岡秀和

Japan Advanced Institute of Science and Technology, Hidekazu Yoshioka

同志社大学　辻村元男

Doshisha University, Motoh Tsujimura

1. はじめに

個体や集団による意思決定は，人間社会に留まらない多様な生物社会において見られる普遍的な行為である．人間社会における例としては，環境に配慮しつつも持続的な成長を目指す企業経営 (Xu et al., 2024) やパンデミック下における市街地封鎖やワクチン接種の政策決定 (Prieur et al., 2024) 等が挙げられる．人間以外の生物社会における例としては，鳥類や哺乳類の大移動である「渡り」の時期の集団意思決定 (Kao et al., 2024) やアリの採餌行動 (Cristín, 2024) 等が挙げられる．

集団内においてある個体の意思決定が他個体の意思決定に左右される状況を数理的に記述でき，さらに，こうした行為の帰結として生じる集団レベルでの行動をも記述できる理論は動的ゲーム理論と呼ばれる．その守備範囲は極めて広範であるが，少なくとも以下に述べる 2 種類の枠組みが存在する．進化ゲーム理論 (Sandholm, 2020) と平均場ゲーム理論である (Lasry and Lions, 2007)．本稿では，前者が後者のある極限で現れることを論じる．すなわち，後者の方が前者より複雑であるという立場をとる．

進化ゲーム理論では，意思決定を規定する行為の空間 (以下，表記の簡単化のために単に空間と呼ぶ) を領域として，力学系，例えば差分方程式や微分方程式に基づいて集団内の行為の確率分布 (以下，行為分布と呼ぶ) の時間発展を記述する．物理学において現れる熱方程式や移流方程式のように，時間発展は前向き，すなわち現在から未来に向けてである．この事からも示唆されるように，一般に進化ゲーム理論では現在の集団内の状態に基づいて未来の状態が逐次的に意思決定されていく．

一方，平均場ゲーム理論は最適制御理論と深く関連しており，とりわけ動的計画原理との関りが深い．最適制御理論における動的計画原理は，未来のあるべき姿を実現するために今何を行うべきであるかを導くことを目的としている．すなわち，未来から現在へ，時間後ろ向きに問題を解いていく．ある目的地に所定時間時刻までに到着するには一体何時に家を出れば良いか，という状況がもっとも単純な動的計画原理に相当する．一般の最適制御理論では，動的計画方程式 (ハミルトン・ヤコビ・ベルマン方程式とも呼ばれる) を時間後ろ向きに求解することで，最適制御則，ならびにその制御下におけるシステムの振る舞いを分析する．

平均場ゲームは，形式的にはシステムや評価関数が行為分布に陽的に依存した動的計画原理であるとみなせる (Lasry and Lions, 2007)．後に具体例を示すが，この場合には動的計画方程式

を求解するだけでは最適制御則は得られず，したがって制御下でシステムの振る舞いもわからない．制御下における行為分布を規定する法則，すなわち進化ゲームの支配方程式に該当する式も同時に解く必要がある．進化ゲームは時間前向きの行為分布の時間発展を記述する方程式に基づく一方，これと同時に動的計画方程式を解く平均場ゲームでは，互いに時間逆向きに発展していく方程式からなるシステムを取り扱う必要性が生じる．この点が，平均場ゲームの数学・技術的な難しさを生んでいる．本稿では，特定の型の進化ゲームに着目し，これがある平均場ゲームの極限として現れることを論じる．

本稿は Yoshioka (2024a) に基づき，細部の計算や設定を省いたうえで，以下のあらましに沿う．本稿では，連続な空間上における進化ゲームの代表的なモデルとしてレプリケータ動力学に着目する．ただし，$t \geq 0$ を時刻，空間を 1 次元閉区間 $\Omega = [0,1]$ とし，時刻 t，空間内の位置 x における行為分布の確率密度関数 (以下，密度と呼ぶ) を $p_t(x)$ と書く：

$$\begin{aligned} \underbrace{\frac{\mathrm{d}}{\mathrm{d}t} p_t(x)}_{\text{密度の時間変化}} &= \underbrace{\left(\int_{y \in \Omega} \rho(y, x, p_t)\, p_t(y)\, \mathrm{d}y \right) p_t(x)}_{\text{確率の流入}} - \underbrace{\left(\int_{y \in \Omega} \rho(x, y, p_t)\, p_t(y)\, \mathrm{d}y \right) p_t(x)}_{\text{確率の流出}} \\ &= \underbrace{\int_{y \in \Omega} \left(\rho(y, x, p_t) - \rho(x, y, p_t) \right) p_t(y)\, \mathrm{d}y}_{\text{変化率}} \times \underbrace{p_t(x)}_{\text{密度}} \end{aligned}, \quad t > 0,\ x \in \Omega . \quad (1)$$

各変数や各項の意味は次節で述べるが，ここで重要であるのは，式(1)が 1 行目のように確率の収支式，すなわちフォッカー・プランク方程式 (例えば，Example 16 in Bect (2010)) として理解できる点である．また，同式 2 行目のように右辺が変化率と密度の積として分解される点も重要である．

フォッカー・プランク方程式は確率過程の確率密度関数の時間発展を記述し，とくに式(1)の型は跳躍が駆動する確率微分方程式に付随することがわかる．このことを動機に，その確率微分方程式を与えるような最適制御問題はどのようなものであるかを考察する．まさに，制御工学における，所与の最適制御則を与える評価関数を見出す「逆最適制御」 (例えば，Komaee, 2023) に他ならない．レプリケータ動力学の場合には，幸運にも評価関数が陽的に導かれる．このような異なるモデル同士のつながりの探索は，それ自身に意味があるのみならず，既往研究と比較して柔軟で適用範囲が広いモデルを発見するための糸口になるかもしれない．

進化ゲームと平均場ゲームを関連付ける方法は無数にあると考えられる．そこで，本稿では各ゲームの背景を鑑みた議論を採用する．すなわち，進化ゲームは未来のことを考えないが平均場ゲームは未来のことを考える，という差異に着想を得て，前者が後者の割引無限大の極限になるようにする．割引率が大きいほど未来よりも現在の状態に重きを置いた制御則が得られることから，「割引率が無限に大きい＝未来のことは一切考えない」という風に解釈できる．この方法は確かに形式的には上手くいくが，本稿執筆時点であくまで形式的に過ぎず，数学的な厳密性 (一体どういうノルムでの収束なのか，収束するとしてその速度はどの程度か等) は補完できていない．ただし，数値計算による見積もりはなされている (Yoshioka, 2024a).

レプリケータ動力学ならびにその亜種である進化ゲームの支配方程式については，詳細な解

析がなされているために (Mendoza-Palacios and Hernández-Lerma, 2024)，それら自体は数学的に目新しい対象ではない．しかし，平均場ゲームの着眼点からレプリケータ動力学を眺めた文献は見当たらなかった．割引無限大の極限に基づきレプリケータ動力学とフォッカー・プランク方程式を関連付ける研究は存在するが (Bardi and Cardaliaguet, 2021; Bertucci et al., 2019)，進化ゲームはあまり取り扱われていない．

さいごに，本稿は新規的な研究アイデアを示すものではなく，Yoshioka (2024a) に基づいて進化ゲームと平均場ゲームの繋がりを和文で概説するものである．本稿を通して，進化ゲームと平均場ゲームの双方についての議論が進むことを期待している．

2. レプリケータ動力学

本稿の表記法は，進化ゲームに関する文献（例えば，Chapter 1.3 in Mendoza-Palacios and Hernández-Lerma や Cheung et al. (2014)）に基づく．前述のように，集団内の個人の行動の領域が 1 次元のコンパクト空間であると仮定し，一般性を失わずに $\Omega=[0,1]$ とおき，Ω のボレル集合体を $\mathcal{B}$ と書く．また，Ω 上の有限符号付き測度の空間を $\mathcal{M}(\Omega)$，そのノルムを全変動ノルム $\|\mu\|=\sup_g\left|\int_\Omega g(x)\mu(\mathrm{d}x)\right|$ として与える．ここで，$\sup_g$ は Ω 上のすべての可測関数に対するものである．同様に，Ω 上の確率測度の空間を $\mathcal{P}(\Omega)$ と書く．このとき，包含関係 $\mathcal{P}(\Omega)\subset\mathcal{M}(\Omega)$ が成り立つ．また，時間依存する測度 $\mu_t\in\mathcal{M}(\Omega)$ の時間微分を $\frac{\mathrm{d}\mu_t}{\mathrm{d}t}$ と書き，ノルムの意味で理解する（$\lim_{\varepsilon\to0}\left\|\frac{\mathrm{d}\mu_t}{\mathrm{d}t}-\frac{\mu_{t+\varepsilon}-\mu_t}{\varepsilon}\right\|=0$）．とくに指定がない限り，$A\in\mathcal{B}$ は任意のボレル可測集合である．

2.1 動力学と効用

元来，連続空間上のレプリケータ動力学は以下で与えられる無限次元常微分方程式であり，行為分布の時間発展を記述するものである：

$$\underbrace{\frac{\mathrm{d}}{\mathrm{d}t}\mu_t(A)}_{\text{確率の時間変化}}=\int_{x\in A}\left(\underbrace{U(x,\mu_t)-\int_{y\in\Omega}U(y,\mu_t)\mu_t(\mathrm{d}y)}_{\text{自分の効用−効用の期待値}}\right)\mu_t(\mathrm{d}x),\quad t>0. \tag{2}$$

ただし，初期条件 $\mu_0\in\mathcal{P}(\Omega)$ を課す．ここで，$U:\Omega\times\mathcal{P}(\Omega)\to\mathbb{R}$ は効用，すなわち行為分布によって最適化されるべき関数である．ナッシュ均衡の意味での最大化が考えられることが多いが，本稿では深く触れない．例えば，有界な可測関数 $f:\Omega\times\Omega\to\mathbb{R}$ を用いて，効用を

$$U(x,\mu)=\int_\Omega f(x,y)\mu(\mathrm{d}y) \tag{3}$$

と与える場合が最も単純である．

技術的には，レプリケータ動力学の Well-posedness を論じるためには効用を $U:\Omega\times\mathcal{M}(\Omega)\to\mathbb{R}$ としたうえで下記の様な条件を課す必要があるが（例えば，Mendoza-Palacios and Hernández-Lerma や Cheung et al. (2014)），本稿では深く触れない．重要であるのは，効用をどのような関数

形で与えても良いわけではないという点である：すべての $x, y \in \Omega$， $\mu, \nu \in \mathcal{M}_2(\Omega)$ に対してある定数 $K > 0$ があり，以下が成り立つ.

$$\left|U(x,\mu)\right| \le K \quad \text{(有界性)} \tag{4}$$

および

$$\left|U(x,\mu) - U(x,\nu)\right| \le K\|\mu - \nu\| \quad \text{(リプシッツの連続性).} \tag{5}$$

ここで，集合 $\mathcal{M}_2(\Omega)$ は全変動ノルムが 2 以下 ($\|\mu\| \le 2$) となる要素のみからなる $\mathcal{M}(\Omega)$ の部分空間である．この“2”という数字は本質的ではなく，1 より大きい実数，例えば 1.1 等で代用することができる.

2.2 動力学の書き直し

式(2)をそのまま眺めているだけでは，先に進むことができない．式(2)は確率分布の時間発展を記述するはずであるから, 何らかの保存則として書き直されるべきである. 天下り的であるが, 式(2)は以下のように書き直すことが可能である (Remark 1 of Chaung (2024))：

$$\frac{\mathrm{d}}{\mathrm{d}t}\mu_t(A) = \int_{x\in A}\int_{y\in\Omega} \rho(y,x,\mu_t)\,\mu_t(\mathrm{d}y)\,\mu_t(\mathrm{d}x) - \int_{x\in A}\int_{y\in\Omega} \rho(x,y,\mu_t)\,\mu_t(\mathrm{d}y)\,\mu_t(\mathrm{d}x). \tag{6}$$

ここで， ρ は以下のジャンプ率である：

$$\rho(x,y,\mu) = \max\left\{U(y,\mu) - U(x,\mu), 0\right\}. \tag{7}$$

式(6)への書き直しは，以下のように確認できる：

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\mu_t(A) &= \int_{x\in A}\int_{y\in\Omega} \max\left\{U(x,\mu_t) - U(y,\mu_t), 0\right\}\mu_t(\mathrm{d}y)\,\mu_t(\mathrm{d}x) \\
&\quad - \int_{x\in A}\int_{y\in\Omega} \max\left\{U(y,\mu_t) - U(x,\mu_t), 0\right\}\mu_t(\mathrm{d}y)\,\mu_t(\mathrm{d}x) \\
&= \int_{x\in A}\int_{y\in\Omega} \Big(\max\left\{U(x,\mu_t) - U(y,\mu_t), 0\right\} - \max\left\{U(y,\mu_t) - U(x,\mu_t), 0\right\}\Big)\mu_t(\mathrm{d}y)\,\mu_t(\mathrm{d}x). \\
&= \int_{x\in A}\int_{y\in\Omega} \big(U(x,\mu_t) - U(y,\mu_t)\big)\,\mu_t(\mathrm{d}y)\,\mu_t(\mathrm{d}x) \\
&= \int_{x\in A}\Big(U(x,\mu_t) - \int_{y\in\Omega} U(y,\mu_t)\,\mu_t(\mathrm{d}y)\Big)\mu_t(\mathrm{d}x)
\end{aligned} \tag{8}$$

ここでさらに測度 μ には密度 p が存在する ($p_t(x)\mathrm{d}x = \mu_t(\mathrm{d}x)$) と仮定すれば，式(1)が得られる.

以上の結果は既に先行研究で知られていたが，次節からレプリケータ動力学の新しい側面に触れる.

2.3 動力学の背後にある確率微分方程式

以下では，確率過程 $X = (X_t)_{t\ge 0}$ により集団内におけるある個体 (以下，代表個体と呼ぶ) の行為の時間発展を記述する. 集団には無限に多くの個体が属しており，各個体が同じ形の確率微分方程式 (同一の確率微分方程式，ではないことに注意) に基づいて行為を逐次的に更新していくと仮定する.

式(1)は，後述する特殊な場合，以下の確率微分方程式に付随するフォッカー・プランク方程式とみなせる：

$$X_t = \begin{cases} \mathrm{d}L_t & \left(\text{at jump time of } L\right) \\ X_{t-} & \left(\text{otherwise}\right) \end{cases}, \quad t > 0 . \tag{9}$$

ただし，初期時刻 0 に初期条件 X_0 を与える必要がある．ここで，“$t-$”は，時間について左制限を取ることを意味し，$L = \left(L_t\right)_{t\geq 0}$ はレヴィ測度 $u_t\left(z\right)$ を持つ複合ポアソン過程である．ただし，$u = \left(u_t\right)_{t\geq 0}$ は，Ω から $\left[0,+\infty\right)$ の可予測な写像であり，形式的には $u_t\left(z\right)/\int_\Omega u_t\left(z\right)\mathrm{d}z$ がジャンプサイズの確率密度関数をあらわす (ただし，分母が 0 以外の場合)．このとき，記法の濫用を許して，もしマルコフ制御 (Øksendal and Sulem, 2019) のように $u_t\left(z\right) = p_t\left(z\right)\rho\left(X_t, z, p_t\right)$ と置ければ，確かに式(1)が確率微分方程式(9)に付随するフォッカー・プランク方程式とみなせる．また，ジャンプサイズが Ω 内にあることから，確率過程 X の値域は Ω である．

確率微分方程式(9)では，L の各ジャンプ時刻において X が Ω 内の位置をランダムに更新する．ただし，一様分布のように完全にランダムではなく，更新位置はジャンプ時刻直前の X やその密度 p に依存する．そのため，式(9)は通常の意味での確率微分方程式ではなく，ジャンプが密度 p に依存する，より数学的に難しい対象であろう．また，式(9)レプリケータ動力学を確率微分方程式の観点から再解釈できるという重大な発見を示している．次節ではこの点をさらに追及し，「最適」なジャンプ率 ρ として与えるような平均場ゲームはどのようなものであるかを探る．

3. 平均場ゲーム

3.1 評価関数と最適性方程式

本節ではまず，レプリケータ動力学と関連する平均場ゲームを定式化する．平均場ゲームは評価関数やダイナミックスが密度に陽的に依存する最適制御問題と見なせることに注意する．確率微分方程式(9)のジャンプ強度 u を制御変数として，固定の有限期間 $\left(0,T\right)$ における以下の最適制御問題を考える．ここで，$t\in\left[0,T\right]$ かつ $x\in\Omega$ である：

$$\Phi\left(t,x\right) = \sup_u \mathbb{E}\left[\int_t^T \underbrace{e^{-\delta\left(s-t\right)}}_{\text{割引}}\left(\underbrace{\delta U\left(X_s,\mu_s\right)}_{\text{効用}} - \underbrace{\int_\Omega c\left(z,u_s\right)\mathrm{d}z}_{\text{コスト}}\right)\mathrm{d}s \,\middle|\, X_t = x\right]. \tag{10}$$

ただし，上限 $\sup_u$ は $u_t\left(z\right) = p_t\left(z\right)\rho\left(X_t, z, p_t\right)$ に限らないジャンプ強度についてとる．ここで，$T > 0$ は固定された終端時刻であり，$\delta > 0$ は割引率である．式(10)の左辺は値関数と呼ばれ，代表個体が時刻 t に行為 $x\in\Omega$ を選択していた場合に，以降の時間に得られる正味の効用である．式(10)の右辺は，効用から行為を変更することによるコスト c を差し引き，なおかつ割引の補正がかかった量の期待値である．

本稿における割引の役割は重要である．式(10)の右辺からわかるように，割引 $e^{-\delta\left(s-t\right)}$ は現在時刻から時間が経過するにつれて減少する指数関数である．割引率 δ が大きくなるほど指数関数の減衰が強くなり，将来的に発生し得る効用やコストをより鑑みない最適制御が導かれること

になる．また，式(10)の右辺の効用の項に割引率δがかかっているが，このスケーリングが平均場ゲームから進化ゲームを導くために重要である．ごく形式的な計算ではあるが，

$$\int_t^T \delta e^{-\delta(s-t)} U\left(X_s, \mu_s\right) \mathrm{d}s \to U\left(x, \mu_t\right) \quad \text{as} \quad \delta \to +\infty \tag{11}$$

と捉える．係数$\delta e^{-\delta s}$が強度δの指数分布にしたがい，その$\delta \to +\infty$の極限が原点に集中するディラックのデルタとなるためである．さいごに，実際に右辺の最大化を実現するuを最適制御と呼び，$u = u^*$と書く．

さて，密度$p_t(x)\mathrm{d}x = \mu_t(\mathrm{d}x)$について，動的計画原理 (例えば，Gomes and Saúde (2014)) に基づくと，pは以下のフォッカー・プランク方程式に支配される：

$$\frac{\mathrm{d}}{\mathrm{d}t} p_t(x) = p_t(x) \int_{y \in \Omega} \left\{ u_t^*(y, x) - u_t^*(x, y) \right\} p_t(y) \mathrm{d}y \,. \tag{12}$$

ただし，$0 < t \le T$かつ$x \in \Omega$である．また，値関数Φは以下の動的計画方程式に支配される：

$$-\frac{\partial}{\partial t}\Phi(t,x) = -\delta\Phi(t,x) + \sup_u \left\{ \int_\Omega \Big(u(z)\big(\Phi(t,z) - \Phi(t,x)\big) - c\big(z, u(z)\big) \Big) \mathrm{d}z \right\} + \delta U\left(x, p_t\right). \tag{13}$$

ただし，$0 \le t < T$かつ$x \in \Omega$である．式(12)には適切な初期条件として何らかの密度p_0，式(13)には終端条件$\Phi(T, \cdot) = 0$が課される．ここで，式(13)の上限は，可測関数$u : \Omega \to [0, +\infty)$に対してとられ，$u^*$はその最大化限であるとする：

$$u_t^*(x, \cdot) = \arg\max_{u(\cdot)} \left\{ \int_\Omega \Big(u(z)\big(\Phi(t,z) - \Phi(t,x)\big) - c\big(z, u(z)\big) \Big) \mathrm{d}z \right\}. \tag{14}$$

式(12)と式(1)を比較すると，$u_t^*(x, y)$に$\rho(x, y, p_t)\, p_t(y)$が対応することがわかる．

3.2 逆最適制御の応用

洞察(11)に基づくと，値関数Φは効用そのものではないが，効用「のような量」であると理解することはできる．これを念頭に置きながら，逆最適制御に基づいてコストcを設計する．とくに，次式を満足するようなcを目指す：

$$c'^{(-1)}\left(z, \Phi(t,z) - \Phi(t,x)\right) = p(z) \max\left\{ \Phi(t,z) - \Phi(t,x), 0 \right\}. \tag{15}$$

ここで，$c'^{(-1)}$はc'の第二番目の引数についての逆関数である．式(15)においてΦをUで置き換えれば，左辺が$p(z)\rho(x, z, p)$となり，当初の目的に近づく．以下，cの第二引数が非負であることに注意する．

直感的には，式(14)の右辺の最大化限は以下の一階微分の条件を満足するという要請が自然であろう：

$$\frac{\partial}{\partial u} c(z, u) = c'(z, u) = \Phi(t,z) - \Phi(t,x) \,. \tag{16}$$

ここで，$c'(z, u)$は第二番目の引数に対する偏微分である．このとき，もし$\Phi(t,z) - \Phi(t,x) \ge 0$であれば次式を得る：

$$u^* = c'^{(-1)}\left(z, \Phi(t,z) - \Phi(t,x)\right). \tag{17}$$

式(15)を仮定すると，$\Phi(t,z) - \Phi(t,x) \geq 0$ を仮定しているために

$$c'^{(-1)}\left(z, \Phi(t,z) - \Phi(t,x)\right) = p(z)\left(\Phi(t,z) - \Phi(t,x)\right) \tag{18}$$

を得る．もし，さらに $p(z) > 0$ であれば，以下の c が式(18)を導くことがわかる：

$$c(u,z) = \frac{u^2}{2p(z)}. \tag{19}$$

実際に計算すると

$$\frac{\partial}{\partial u} c(z,u) = \frac{u}{p(z)} \text{(一階微分)} \quad \text{かつ} \quad \frac{\partial^2}{\partial u^2} c(z,u) = \frac{1}{p(z)} > 0 \text{ (二階微分が正)} \tag{20}$$

であるため，式(14)右辺の最大化限が $u^* = p(z)\left(\Phi(t,z) - \Phi(t,x)\right)$ として得られる．

さて，$\Phi(t,z) - \Phi(t,x) < 0$ の場合も考慮し，式(19)を以下のように拡張する：

$$c(u,z) = \begin{cases} \dfrac{u^2}{2p(z)} & \left(u \geq 0,\ p(z) > 0\right) \\ +\infty & \left(u > 0,\ p(z) = 0\right) \end{cases}. \tag{21}$$

コスト(21)の意味は以下のように考えられる．まず，分母にジャンプ後，すなわち行為を更新した後の密度 $p(z)$ が入っていることから，このコストは，行為を更新した後における密度が小さいほど，意思決定を更新するためのコストが高く，その逆も同様であることを意味する．また，誰も用いていない行為（$p(z) = 0$）への更新は生じ得ない．以上から，より密度が高い行為を模倣するとより少ないコストしか発生しないことになり，周囲の意思決定の状態を鑑みながら自身の意思決定をしていくという進化ゲーム論的な考え方，とりわけレプリケータ動力学の背景と合致している．

さて，式(21)の c を用いると

$$\sup_u \left\{ \int_\Omega \left(u(z)\left(\Phi(z) - \Phi(x)\right) - c\left(z, u(z)\right)\right) \mathrm{d}z \right\} = \frac{1}{2} \int_\Omega p(z) \max\left\{\Phi(t,z) - \Phi(t,x), 0\right\}^2 \mathrm{d}z \tag{22}$$

を得るために，動的計画方程式は次式のようになる：

$$-\frac{\partial}{\partial t}\Phi(t,x) = -\delta\Phi(t,x) + \frac{1}{2}\int_\Omega p_t(z) \max\left\{\Phi(t,z) - \Phi(t,x), 0\right\}^2 \mathrm{d}z + \delta U\left(x, p_t\right). \tag{23}$$

さらに，式(12)は

$$\frac{\mathrm{d}}{\mathrm{d}t} p_t(x) = p_t(x)\left(\Phi(t,x) - \int_{y\in\Omega} \Phi(t,y)\, p_t(y)\, \mathrm{d}y\right), \tag{24}$$

また，最適制御 u^* は状態 X と値関数 Φ を用いて以下のようになる：

$$u_t^* = p_{t-}(\cdot) \max\left\{\Phi(t-,\cdot) - \Phi\left(t-, X_{t-}\right), 0\right\}. \tag{25}$$

3.3 割引無限大極限

以上の話をまとめると，逆最適制御に基づいて設計されたコスト(21)に基づく平均場ゲームの最適性方程式系は，時間前向きに解くべき式(24)ならびに時間後ろ向きに解くべき式(23)から構成される．さて，式(23)を以下のように変形する：

$$\Phi(t,x)-U(x,p_t)=\frac{1}{\delta}\left\{\frac{\partial}{\partial t}\Phi(t,x)+\frac{1}{2}\int_\Omega p_t(z)\max\left\{\Phi(t,z)-\Phi(t,x),0\right\}^2\mathrm{d}z\right\}. \tag{26}$$

もし，右辺の一番外側にあるカッコ{ }内の各項が$\delta\to+\infty$の極限において有限な値に留まるのであれば，この極限のもとで$\Phi(t,x)=U(x,p_t)$が導かれる．このとき，式(24)は確かにレプリケータ動力学になっている．では，この極限はどのような場合に正当化されるであろうか．完全な回答は得られていないが，少なくとも終端時刻Tにおいて極限$\Phi(T,x)=U(x,p_T)$の成立は期待できない．これは，動的計画方程式において終端条件$\Phi(T,\cdot)=0$が課されているためである．また，式(26)右辺の第一項の方が第二項よりも正則性が低いことが想定されるために，正当化がより難しい可能性がある．

また，時間依存の密度$p=(p_t)_{t\geq 0}$を所与の関数であるとすると，動的計画方程式(23)はハミルトニアン$H:\Omega\times\mathcal{P}(\Omega)\times\mathbb{R}\times B(\Omega)\to\mathbb{R}$を使用して以下のように書き直す．ここで，$B(\Omega)$は$\Omega$から$\mathbb{R}$への有界関数である：

$$-\frac{\partial}{\partial t}\Phi(t,x)=H\left(x,p_t,\Phi(t,x),\Phi(t,\cdot)-\Phi(t,x)\right), \tag{27}$$

ここで，

$$H(x,p_t,\alpha,\beta)=-\delta\alpha+\frac{1}{2}\int_\Omega p_t(z)\max\left\{\beta(t,z)-\alpha,0\right\}^2\mathrm{d}z+\delta U(x,p_t) \tag{28}$$

である．このとき，ハミルトニアンHは，それぞれ第二，第三の引数に対してそれぞれ非増加および非減少である．これは非局所縮退楕円性として知られる条件に相当し，粘性解の意味で動的計画方程式を論じるに際して重要な性質である (例えば，Barles Imbert (2008)や Rodríguez-Paredes and Topp (2023)).

以上から，効用Uと密度pが十分に高い正則性を有し，さらに値関数Φについての事前評価が得られていれば，pを所与とした動的計画方程式について粘性解の意味で論じることができると考えられる．しかし，本稿の平均場ゲームではpとΦを切り離して考えることができない．ここが数学的に真に難しい部分であろう．

3.4 その他の議論

進化ゲームとりわけレプリケータ動力学と平均場ゲームには，割引率という時間の逆数の尺度を通した繋がりがあることがわかった．では，この結果が応用学問分野に与える示唆にはどのようなものがあり得るだろうか．

まず，効用関数が適切にモデル化されているという前提に基づくが，多様な意思決定の時間スケールに対応したゲーム理論を導くことができると考えられる．例えば，Sustainable Development Goals や Green Transformation など，持続可能性に関わる諸概念が世にあふれており，各団体や企

業はこれらに沿った行動計画や経営を求められている．本稿の枠組みによって，割引率が小さく遠い将来まで見据えた団体と割引率が大きく近視眼的な団体が取るべき意思決定や，その帰結を分析することが可能になると考えられる．例えば，我が国の各地域の観光産業では，多くの観光客を望む一方，過度の観光客の受け入れが地域の環境や景観，遺産などの破壊に至る，いわゆるオーバーツーリズムを引き起こす可能性がある．受け入れ観光客数に重点を置き割引率が大きい意思決定は，まさにオーバーツーリズムを引き起こす可能性があるのではないかと考えている．

4. おわりに

本稿では，レプリケータ動力学と繋がる平均場ゲームについて論じた．両者を関連付ける中心的なアイデアは，以下の①-②であった．

① レプリケータ動力学をフォッカー・プランク方程式とみなしたうえで，その背後に存在する確率過程を見出すこと．
② 平均場ゲームにおける評価関数を逆最適制御の考え方に基づいて設計し，割引無限大極限においてレプリケータ動力学が現れるようにすること．

実は，上記の①-②いずれにおいても，レプリケータ動力学を他の進化ゲームの支配方程式，例えばロジット動力学や射影動力学に置き換えても，ある程度同様の理論展開が可能であることがわかってきた (Yoshioka, 2024b-c)．どちらの場合も，逆最適制御によって望ましい評価関数が閉形式で求まり，工学的には大変に都合が良い．ただし，射影動力学は非平滑 (微分方程式の解が時間方向について局所的にくの字型に折れ曲がるさまを想像すれば良い) であるため，特定の正則化が必要である．各動力学の性質が互いに異なり，したがって異なる平均場ゲームを導くであろう点も重要である．例えば，レプリケータ動力学では初期時刻に使用されていない行為はその後の時刻でも使用されない (式(1)の右辺が 2 行目のように分解できることによる)．一方，ロジット動力学では初期時刻に使用されていない行為であってもそのすぐ後，ごくわずかかもしれないが使用される．射影動力学では，行為の使用・不使用が時間依存で可逆的に変化し得る．

本稿で論じた方法は，進化ゲームと平均場ゲームを繋ぐ唯一の方法ではないと考えている．例えば，モデル予測制御の考え方に基づいて両者を関連付ける理論 (Degond et al. (2017)や Yoshioka and Tsujimura (2024) など) がある．しかし，この場合には，平均場ゲームからは進化ゲームとは異なる別の方程式が導かれるようである．両ゲーム理論についても未解決課題が山積しており，それらの解決が待たれる．例えば，拡散ノイズがある進化ゲームの知見が少ない：

$$\frac{\mathrm{d}}{\mathrm{d}t}p_t(x)=p_t(x)\int_{y\in\Omega}\big(\rho(y,x,p_t)-\rho(x,y,p_t)\big)p_t(y)\mathrm{d}y+Dx(1-x)\Delta p_t(x),\quad t>0,\ x\in\Omega. \tag{29}$$

ここで，Δ は一次元のラプラシアンであり，$D>0$ は拡散係数である．式(29)の拡散項は退化楕円型をしているが，これは行為が空間 Ω から外にはみ出ないようにするためである．この拡散項は，最も単純な多項式過程であるヤコビ過程に付随している．

さいごに，本稿のアプローチは，「ある対象を複数の観点から解釈する」ことに基づく．ゲーム理論に限らず，このような考え方が大事である場面は数えきれないであろうと考えられる．

謝辞

本研究は JSPS 科研費 (番号 22K14441，22H02456) ならびに 2024 年度日本生命財団環境問題研究助成・若手・奨励研究 (番号 24), JST さきがけ (番号 JPMJPR24KE) の支援を受けた．また，この研究は 2024 年度京都大学数理解析研究所共同研究（公開型)「ファイナンスの数理解析とその応用」の成果を含む．

引用文献

Xu, Y., Fei, C., & Fei, W. (2024). Optimal dynamic contract with Knightian uncertainty of ESG rating. International Journal of General Systems, 53(3), 381-402. https://doi.org/10.1080/03081079.2023.2282988

Prieur, F., Ruan, W., & Zou, B. (2024). Optimal lockdown and vaccination policies to contain the spread of a mutating infectious disease. Economic Theory, 77(1), 75-126. https://doi.org/10.1007/s00199-023-01537-6

Kao, A. B., Banerjee, S. C., Francisco, F. A., & Berdahl, A. M. (2024). Timing decisions as the next frontier for collective intelligence. Trends in Ecology & Evolution, Online published. https://doi.org/10.1016/j.tree.2024.06.003

Cristín, J., Fernández-López, P., Lloret-Cabot, R., Genovart, M., Méndez, V., Bartumeus, F., & Campos, D. (2024). Spatiotemporal organization of ant foraging from a complex systems perspective. Scientific Reports, 14(1), 12801. https://doi.org/10.1038/s41598-024-63307-1

Sandholm, W. H. (2020). Evolutionary game theory. In: Sotomayor, M., Pérez-Castrillo, D., Castiglione, F. (eds) Complex Social and Behavioral Systems. Encyclopedia of Complexity and Systems Science Series. Springer, New York, 573-608. https://doi.org/10.1007/978-1-0716-0368-0_188

Lasry, J. M., & Lions, P. L. (2007). Mean field games. Japanese journal of mathematics, 2(1), 229-260. https://doi.org/10.1007/s11537-007-0657-8

Yoshioka, H. (2024). Generalized replicator dynamics based on mean-field pairwise comparison dynamic. Preprint. https://arxiv.org/abs/2407.20751

Bect, J. (2010). A unifying formulation of the Fokker–Planck–Kolmogorov equation for general stochastic hybrid systems. Nonlinear Analysis: Hybrid Systems, 4(2), 357-370. https://doi.org/10.1016/j.nahs.2009.07.008

Komaee, A. (2023). Dynamic Gain Adaptation in Linear Quadratic Regulators. IEEE Transactions on Automatic Control. 69(8), 5094-5108. https://dx.doi.org/10.1109/TAC.2023.3344551

Mendoza-Palacios, S., & Hernández-Lerma, O. (2024). Evolutionary games and the replicator dynamics. Elements in Evolutionary Economics. Cambridge University Press.

Cheung, M. W. (2014). Pairwise comparison dynamics for games with continuous strategy space. Journal

of Economic Theory, 153, 344-375. https://doi.org/10.1016/j.jet.2014.07.001

Bardi, M., & Cardaliaguet, P. (2021). Convergence of some mean field games systems to aggregation and flocking models. Nonlinear Analysis, 204, 112199. https://doi.org/10.1016/j.na.2020.112199

Bertucci, C., Lasry, J. M., & Lions, P. L. (2019). Some remarks on mean field games. Communications in Partial Differential Equations, 44(3), 205-227. https://doi.org/10.1080/03605302.2018.1542438

Øksendal, B., & Sulem, A. (2019). Applied Stochastic Control of Jump Diffusions. Springer, Cham.

Gomes, D. A., & Saúde, J. (2014). Mean field games models—a brief survey. Dynamic Games and Applications, 4, 110-154. https://doi.org/10.1007/s13235-013-0099-2

Barles, G., & Imbert, C. (2008). Second-order elliptic integro-differential equations: viscosity solutions' theory revisited. In Annales de l'IHP Analyse non linéaire, Vol. 25, No. 3, pp. 567-585. https://doi.org/10.1016/j.anihpc.2007.02.007

Rodríguez-Paredes, A., & Topp, E. (2023). Rates of convergence in periodic homogenization of nonlocal Hamilton–Jacobi–Bellman equations. ESAIM: Control, Optimisation and Calculus of Variations, 29, 43. https://doi.org/10.1051/cocv/2023038

Yoshioka, H. (2024b). Computational analysis on a linkage between generalized logit dynamic and discounted mean field game. Preprint. https://arxiv.org/abs/2405.15180

Yoshioka, H. (2024c). Numerical analysis of the projection dynamics and their associated mean field control. Dynamic Games and Applications. https://doi.org/10.1007/s13235-024-00598-z

Degond, P., Herty, M., & Liu, J. G. (2017). Mean-field games and model predictive control. Communications in Mathematical Sciences, 15(5), 1403-1422. https://dx.doi.org/10.4310/CMS.2017.v15.n5.a9

Yoshioka, H., & Tsujimura, M. (2024). Generalized logit dynamic and its connection to a mean field game: theoretical and computational investigations focusing on resource management. Dynamic Games and Applications, Published online. https://doi.org/10.1007/s13235-024-00569-4