テクニカルノート

線形な利得構造を有する多人数ゲームにおける離散戦略と連続戦略のダイナミクスに関する考察

谷　本　　潤[†]

ゲームの時間展開がレプリケータダイナミクスに従う，線形な利得構造を有する連続戦略ゲーム（協調（Cooperation，C）を 1，裏切り（Defect，D）を 0 とすると $[0,1]$ の実数値を戦略としてとりうる）の帰結は，離散戦略ゲーム（C もしくは D の離散値しかとりえない）のそれと一致することを，2 人ゲームおよび多人数ゲームの場合に分けて演繹的に検討した．

A Study on a Difference of Dynamics between Discrete and Continuous Strategies of a Multi Player Game Having Linear Payoff Structure

Jun Tanimoto[†]

A deductive analysis concerning replicator dynamics proved that a continuous strategy game (in which a player chooses an arbitrary real number between $[0,1]$ as a cooperative fraction) has the same equilibrium as a discrete strategy game (in which a player chooses only C or D), which has the same linear payoff structure as a continuous strategy game. The deduction is shown for two-player and multi-player games.

1. 緒　　言

進化ゲーム理論におけるジレンマゲーム研究は，利己的な判断に基づいて相互作用を行うマルチエージェント系の情報処理問題への応用という観点から重要な課題である（たとえば，ジレンマゲームの代表的 Archetype である Prisoner's Dilemma は学習エージェントのプロトタイプ的問題としてよく用いられる）．

ところで進化ゲームのダイナミクスは Replicator Dynamics（以下，RD）により考量されることが一般的である．多くのゲームではプレーヤの手は，協調（Cooperation，以下 C）か裏切り（Defection，以下 D）かの離散的 2 戦略（以下，離散戦略）で規定されている．しかし，現実社会におけるゲーム的状況下では，C を 1，D を 0 で表すとすれば，プレーヤ手（戦略）は $[0,1]$ の実数値をとりうる（以下，連続戦略）と仮定すべきであろう．

Doebeli ら[2)]，Killingback ら[3)] は 2 人ゲームで連続戦略を許した場合の RD の概念を提示している．離散戦略における RD は，各離散戦略（C か D か）の母集団における戦略分布の時間変化は当該戦略をとった場合の利得期待値がその時点の戦略分布により決まる全平均利得期待値に対してどれだけ大きいかに比例して推移することを表式したものである．連続戦略では，戦略そのものが連続実数値をとるから「戦略分布」が定義できなくなるが，Doebeli らは，均衡点では resident に対して異なる連続戦略を有する mutant が利得上のメリットをあげえない，との条件から RD を導出している[2)]．彼らは 2 人連続戦略を例に論証を展開しており，多人数ゲームにおいても同様の演繹が成り立つとしているものの，詳細は示されていない．

定性的推論としては，連続戦略ゲームの帰結（ダイナミクスが定常に至った状態）と離散戦略ゲームのそれとは基本的には同様になると考えられる．なぜなら，連続戦略とは，あるプレーヤに特定してみれば離散的な純戦略を時間方向に確率的に繰り出す混合戦略と見なすことができるからである．しかし，たとえば，Izquierdo ら[4)] が遺伝的アルゴリズムを戦略適応に用いたシミュレーションで連続戦略と離散戦略ではその定常時の戦略分布が大きく異なると報告しているよう

† 九州大学大学院総合理工学研究院
Interdisciplinary Graduate School of Engineering Sciences, Kyushu University

表 1 2 人連続戦略ゲームのゲーム構造
Table 1 Game structure of a continuous strategy, 2-player game.

Conditions	Further condition	$x^* = \frac{-S+P}{-S+R+P-T}$	$\left.\frac{\partial D_c(x_r)}{\partial x_r}\right\|_{x_r=x^*} = -S+R+P-T$	Game structure
$P>S$ and $T>R$	$P-S>T-R$	$1<x^*$	Plus	D-dominate (PD)
	$P-S<T-R$	$x^*<0$	Minus	
$P>S$ and $T<R$		$0<x^*<1$	Plus	Bi-stable (Stag Hunt)
$P<S$ and $T>R$		$0<x^*<1$	Minus	Polymorphic (Chicken)
$P<S$ and $T<R$	$P-S>T-R$	$x^*<0$	Plus	C-dominate (Trivial)
	$P-S<T-R$	$1<x^*$	Minus	

に，このことはダイナミクスを何に依拠したかによって必ずしも自明でない．

本論では，RD に従う，線形な利得構造を有する連続戦略ゲームにおけるダイナミクスの帰結と離散戦略ゲームのそれとが一致することを演繹により確認する．

連続戦略ゲームに関する既往研究としては以下があげられる．

Scheuring[5] は Doebeli らの結果をふまえて，連続型 2 人囚人ジレンマゲーム（Prisoner's Dilemma，以下 PD）ではプレーヤが空間構造を有する場合に協調が創発することを演繹および数値実験により示している．Huck ら[6] は連続戦略 2 人および N 人 PD で協調を創発させる頑強な適応ルール GLAD を提示しているが，これは離散戦略ゲームにおける PAVLOV 同様の "Win-stay&Lose-shift" 戦略[7] にほかならない．Taylor[8] は連続戦略 2 人 PD に注目し，2 者が繰り返し相手の前手（$[0,1]$ の協調度）に応じて自分の戦略を変化させる交渉ゲームの均衡について議論している．Chong ら[9] は 2 戦略ゲームを拡張した多段戦略ゲーム 2 人 PD について議論しており，戦略の多段性が協調創発に影響すると主張している．

以上の既往研究では戦略適応のメカニズムはそれぞれ異なる．しかし，最も基本的なダイナミクスとして RD により進化ゲームを考えた場合，離散戦略ゲームと連続戦略ゲームのゲーム帰結とに差異があるのか否かを把握しておくことは，離散戦略ゲームを拡張して連続戦略ゲームを考察するうえで有意な基礎を供するものと考える．

一般に離散戦略ゲームの RD を議論する場合，集団中で C をとるプレーヤの割合 x（戦略分布）の定常値を解析対象とする．連続戦略ゲームの RD では集団中の任意のプレーヤがとる実数値の戦略 x（C を 1 とする）が定常時でどうなるかを解析する．両ゲームで x の持つ厳密な意味は異なるが，ゲームが定常に至る x により達成される集団の平均的な協調率は一致する．本論では，「ゲーム帰結の一致」をこのような意味で用い，あえて両ゲームで x を混用して表記することにする．

2. 2 人連続戦略ゲーム

2 人連続戦略ゲームにおいてプレーヤがとる戦略を $x \in [0,1]$ で表す．離散戦略 2×2 ゲームにおける C，D はそれぞれ $x=1$，$x=0$ で表されることになる．離散戦略 2×2 ゲームにおける自他の戦略組 (C,C)，(C,D)，(D,C)，(D,D) による利得を R，S，T，P で表し（ただし，$R>P$），連続戦略時はこれらを線形内挿して利得を付与する構造を仮定する場合，連続戦略 x_f が連続戦略 x_o と対戦したときに得る利得 $P(x_f, x_o)$ は，

$$P(x_f,x_o) = P + (S-P)x_f + (-P+T)x_o + (-S+R+P-T)x_f x_o. \quad (1)$$

連続戦略ゲームにおけるダイナミクスは，resident（その連続戦略を x_r で表す）に対して異なる連続戦略を有する mutant（その連続戦略を x_m で表す）が得る利得上のメリットに応じて，侵入を受けた resident は自らの連続戦略を推移させることにより表される[2]．すなわち，

$$\dot{x} = \left.\frac{\partial[P(x_m,x_r) - P(x_r,x_r)]}{\partial x_m}\right|_{x_m=x_r}. \quad (2)$$

均衡点では $\dot{x}=0$ より，均衡点における連続戦略 x^* を得る．また，$\left.\frac{\partial[P(x_m,x_r)-P(x_r,x_r)]}{\partial x_m}\right|_{x_m=x_r} \equiv D_c(x_r)$ とおくと，$\left.\frac{\partial D_c(x_r)}{\partial x_r}\right|_{x_r=x^*}$ の正負を調べることにより均衡点の湧出（$\left.\frac{\partial D_c(x_r)}{\partial x_r}\right|_{x_r=x^*} > 0$），吸引（$\left.\frac{\partial D_c(x_r)}{\partial x_r}\right|_{x_r=x^*} < 0$）を知ることができる．2 人連続戦略ゲームでは式 (1) より，

$$D_c(x_r) = (S-P) + (-S+R+P-T)x_r \quad (3)$$

を得るから，均衡点およびゲーム構造は**表 1** のようにまとめることができる．戦略の定義域 $[0,1]$ と x^* の

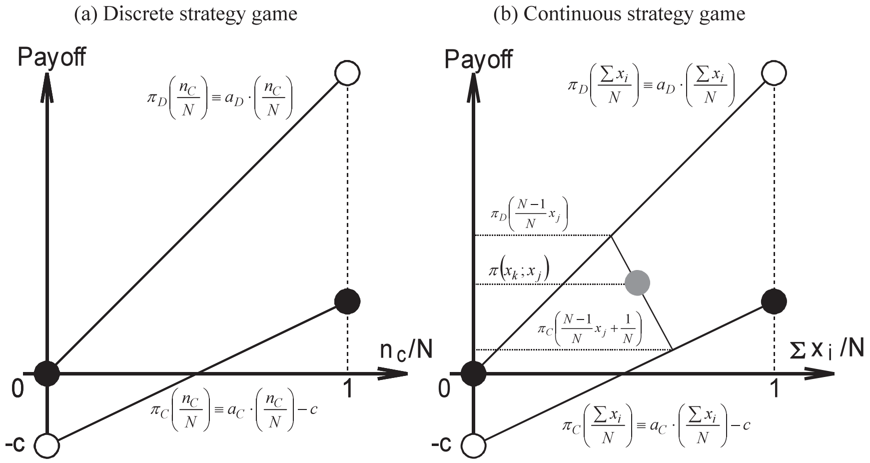

図 1 (a) 離散戦略，(b) 連続戦略の多人数ゲームの利得関数．グレープロットは focal の戦略 x_k，その他 $N-1$ 人の opponents の平均戦略が x_j の場合の focal 利得．○と●は定義上その点を「含まず」，「含んで」を意味する

Fig. 1 Payoff functions of Multi-player game for (a) Discrete strategy game and (b) Continuous strategy game. The gray open circle indicates a focal player's payoff ($\pi(x_k;x_j)$) whose strategy is x_k, while average of other $N-1$ opponents is x_j. Both open and close circles indicate that the definition contains and does not contain the point, respectively.

位置，また均衡点が湧出か吸引かを合わせて考えることにより，ゲームの帰結が D-dominate，Bi-stable，Polymorphic，C-dominate のどれになるかを判断できる．たとえば，$1 < x^*$ でこの均衡点が湧出であれば，戦略定義域 $[0,1]$ では x は単調に減少し，ゲームの帰結は $x=0$ の D-dominate となる．表より 2 人連続戦略ゲームのゲームの帰結を含むゲーム構造は 2 人離散戦略ゲーム（2×2 ゲーム）のそれと一致することが分かる．

3. 多人数ゲーム

3.1 離散戦略ゲーム

線形な利得関数を考える．すなわち，N 人のプレーヤ中の協調者数を n_C で表すとき，協調 C と裏切り D の利得を

$$\pi_C\left(\frac{n_C}{N}\right) = a_C \cdot \left(\frac{n_C}{N}\right) - c \tag{4a}$$

$$\pi_D\left(\frac{n_C}{N}\right) = a_D \cdot \left(\frac{n_C}{N}\right) \tag{4b}$$

で与える（かくて線形利得関数は 3 つのパラメータ a_C，a_D，c により一般的に表された．関数形を**図 1** (a) に示す．いわゆる，Public Goods Game では $a_C = a_D = b > 0$，$c > 0$ なる 2 パラメータで利得構造が表される）．いま，C の戦略分布を x で表し，C および D 戦略をとるプレーヤの期待利得をそれぞれ f_C，f_D で，全プレーヤの期待利得を $\overline{f}$ で表記することにすれば，RD は定義より，

$$\frac{\dot{x}}{x} = f_C - \overline{f} \;\Leftrightarrow\; \dot{x} = x(1-x)(f_C - f_D) \tag{5}$$

となる．N 人ゲームであることに留意すると，f_C，f_D は，

$$f_C = \sum_{j=0}^{N} \binom{N-1}{j} \cdot x^j \cdot (1-x)^{N-1-j} \cdot \pi_C\left(\frac{j+1}{N}\right) \tag{6a}$$

$$f_D = \sum_{j=0}^{N} \binom{N-1}{j} \cdot x^j \cdot (1-x)^{N-1-j} \cdot \pi_D\left(\frac{j}{N}\right) \tag{6b}$$

で与えられる[10]．式 (5) より均衡点は 0，1 および $f_C = f_D$ を与える x^* である．式 (6) に式 (4) を代入すると，

$$x^* = \frac{c \cdot N - a_C}{(N-1)(a_C - a_D)} \tag{7}$$

を得る．均衡点の性質は式 (5) で $\dot{x} = 0$ とした x に関する 3 次式により表される．ここで $D_d(x) = x(1-x)(f_C - f_D)$ とおく．均衡点は 3 つあるので，

$$\left.\frac{\partial D_d(x)}{\partial x}\right|_{x=0} = -c + \frac{a_C}{N} \tag{8}$$

$$\left.\frac{\partial D_d(x)}{\partial x}\right|_{x=1} = -a_C + a_D + c - \frac{a_D}{N} \tag{9}$$

表 2 多人数連続戦略ゲームと離散戦略ゲームとの構造比較

Table 2 Game structure of a continuous strategy, multi player game compared with a discrete strategy game.

Conditions	Further condition	x^*	Discrete strategy $\left.\frac{\partial D_d(x)}{\partial x}\right\|_{x=0}$	Discrete strategy $\left.\frac{\partial D_d(x)}{\partial x}\right\|_{x=1}$	Discrete strategy $\left.\frac{\partial D_d(x)}{\partial x}\right\|_{x=x^*}$	Continuous strategy $\left.\frac{\partial D_c(x_r)}{\partial x_r}\right\|_{x_r=x^*}$	Game structure
$a_C > a_D$	$c \cdot N > a_C$ and $\frac{c \cdot N - a_C}{(N-1)(a_C - a_D)} > 1$	$1 < x^*$	Minus	Plus	Minus	Plus	D-dominate (PD)
	$c \cdot N > a_C$ and $0 < \frac{c \cdot N - a_C}{(N-1)(a_C - a_D)} < 1$	$0 < x^* < 1$	Minus	Minus	Plus		Bi-stable (Stag Hunt)
	$c \cdot N < a_C$	$x^* < 0$	Plus	Minus	Minus		C-dominate (Trivial)
$a_C < a_D$	$c \cdot N < a_C$ and $\frac{a_C - c \cdot N}{(N-1)(a_D - a_C)} > 1$	$1 < x^*$	Plus	Minus	Plus	Minus	C-dominate (Trivial)
	$c \cdot N < a_C$ and $0 < \frac{a_C - c \cdot N}{(N-1)(a_D - a_C)} < 1$	$0 < x^* < 1$	Plus	Plus	Minus		Polymorphic (Chicken)
	$c \cdot N > a_C$	$x^* < 0$	Minus	Plus	Plus		D-dominate (PD)

$$\left.\frac{\partial D_d(x)}{\partial x}\right|_{x=x^*} = -\frac{(a_C - c \cdot N)\big(a_D - a_D \cdot N + (a_C - c)N\big)}{(a_C - a_D)N(N-1)} \tag{10}$$

の正負により判定される各均衡点における湧出，吸引と，$x^* = \frac{c \cdot N - a_C}{(N-1)(a_C - a_D)}$ が区間 $[0,1]$ に対してどこに位置するかが分かれば，ゲーム構造は定まる．

3.2 連続戦略ゲーム

自分を除く $N-1$ 人の連続戦略の平均値を x_j，自らの連続戦略を x_k で表すと，自らが得る利得 $\pi(x_k; x_j)$ は，

$$\pi(x_k; x_j) = x_k \cdot \pi_C\left(\frac{N-1}{N} x_j + \frac{1}{N}\right) + (1 - x_k) \cdot \pi_D\left(\frac{N-1}{N} x_j\right) \tag{11}$$

となる（図 1 (b)）．

resident（その連続戦略を x_r で表す）に対して異なる連続戦略を有する mutant（その連続戦略を x_m で表す）が侵入を試みる状況を想定すると，2 人連続戦略ゲームにおける式 (2) は以下のようになる．

$$\dot{x} = \left.\frac{\partial[\pi(x_m, x_r) - \pi(x_r, x_r)]}{\partial x_m}\right|_{x_m = x_r}. \tag{12}$$

同様に $\left.\frac{\partial[\pi(x_m,x_r) - \pi(x_r,x_r)]}{\partial x_m}\right|_{x_m=x_r} \equiv D_c(x_r)$ とおくと，式 (4)，(11)，(12) より，

$$D_c(x_r) = \frac{N-1}{N}(a_C - a_D) \cdot x_r - \frac{1}{N} a_C - c. \tag{13}$$

よって，$D(x_r) = 0$ を解くと，式 (7) と同一の均衡点 x^* を得る．さらに，式 (13) より，

$$\left.\frac{\partial D_c(x_r)}{\partial x_r}\right|_{x_r = x^*} = \frac{N-1}{N}(a_C - a_D) \tag{14}$$

を得，均衡点の湧出，吸引が定まる（式 (14) と式 (10) は異なることに留意する）．均衡点 x^* が区間 $[0,1]$ に対してどこに位置するか，その湧出，吸引が分かれば，ゲーム構造は定まる．

3.3 離散戦略ゲームと連続戦略ゲームの比較

表 2 にゲーム構造パラメータの条件別に，離散戦略ゲーム，連続戦略ゲームの各均衡点における湧出，均衡をまとめて示した．同一のゲーム構造を持つ離散戦略ゲームと連続戦略ゲームでは，ダイナミクス帰結は一致することが分かる．

4. 結　論

線形な利得構造を有する連続戦略ゲームのダイナミクスの帰結と離散戦略ゲームのそれとが一致することを 2 人ゲームおよび多人数ゲームについて演繹により確認した．

参 考 文 献

1) 蜷川　繁：繰り返し囚人のジレンマゲームにおける 1/f ゆらぎ，情報処理学会論文誌，Vol.44, No.10, pp.2514–2517 (2003).
2) Doebeli, M., Hauert, C. and Killingback, T.: The Evolutionary Origin of Cooperators and Defectors, *SCIENCE*, Vol.306, pp.859–862 (2004).
3) Killingback, T. and Doebeli, M: The Continuous Prisoner's Dilemma and the Evolution of Cooperation through Reciprocal Altruism with Variable Investment, *American Nauralist*, Vol.160, pp.421–437 (2002).
4) Izquierdo, S.S. and Izquierdo, L.R.: On the structural robustness of evolutionary models of cooperation, LNCS 4224, pp.172–182 (2006).
5) Scheuring, I.: The iterated continuous prisoner's dilemma game cannot explain the evolution of interspecific mutualism in unstructured populations, *Journal of Theoretical Biology*, Vol.232, pp.99–104 (2005).
6) Huck, S., Normann, H.T. and Oechssler, J.: GLAD: a simple adaptive strategy that yields cooperation in dilemma games, *Physica D*, Vol.200, pp.133–138 (2005).
7) Posch, M.: Win-stay, lose-shift strategies for repeated games: memory length, aspiration levels and noise, *Journal of Theoretical Biology*, Vol.198, pp.183–195 (1999).
8) Taylor, P.D.: Stability in negotiation games and the emergence of cooperation, *Proc. Royal Society B*, Vol.271, pp.669–674 (2004).
9) Chong, S.Y. and Yao, X.: Behavioral Diversity, Choices and Noise in the Iterated Prisoner's Dilemma, *IEEE Trans. Evolutionary Computation*, Vol.9, No.6, pp.540–551 (2005).
10) Hauert, C., Michor, F., Nowak, M. and Doebeli, M.: Synergy and discounting of cooperation an social dilemmas, *Journal of Theoretical Biology*, Vol.239, pp.195–202 (2006).

(平成 19 年 2 月 8 日受付)
(平成 19 年 4 月 6 日採録)

谷本　　潤（正会員）

1965 年生．早稲田大学理工学部卒業，同大学院理工学研究科修了．工学博士．東京都立大学工学部・助手，早稲田大学理工学総合研究センター・講師，九州大学大学院総合理工学研究科・講師，助教授を経て，2003 年より教授．専門は都市建築環境工学，人間–環境–社会システム工学．空気調和衛生工学会賞，日本建築学会奨励賞，日本建築学会学会賞（論文）．『ハンディブック建築』（オーム社，2005 年）等．小説家，水彩画家．第 67 回コスモス文学新人賞奨励賞，第 76 回コスモス文学新人賞，第 19 回国民文化祭美術展入選，第 37〜39 回福岡市美術展入選等．