

繰り返し社会的ジレンマゲームにおける意思決定モデルの構築

水野景子^a 清水裕士^b

要約

社会的ジレンマゲーム (social dilemma game: SDG) での意思決定を繰り返し行くと、初期には高い協力率が次第に低下する推移が見られることが多くの研究で観察されている (e. g. Davis, 1993; Issac and Walker, 1998). 本研究では、社会的価値志向性 (social value orientation: SVO) モデルと学習モデルを組み合わせて導出された新たなモデルにより、繰り返し SDG における協力率の推移の説明を試みた。実験で得たデータに対するベイズ統計モデリングを用いた実験で得たデータに対するモデル比較の結果、他者の協力に対する期待 (信念) と収益率 (SDG の利得構造) をそれぞれ学習するモデルの当てはまりが最もよいことが示された。モデルから、プレイヤーが SDG の利得構造を学習して非協力を選び始めると他者の協力に対する期待が下がり、他者の協力に対する期待が 50%を下回ると、自身のもつ利得の不平等を回避する傾向によってさらに非協力を選ぶようになる、というメカニズムがあることが示唆される。

JEL 分類番号 : C64, C51, C52

キーワード : 社会的ジレンマ, 社会的価値志向性, 社会選好, ベイズ統計モデリング

^a 関西学院大学社会学研究科 k.mizuno@kwansei.ac.jp

^b 関西学院大学社会学部 hsimizu@kwansei.ac.jp

1. 問題

1. 1. 繰り返し社会的ジレンマゲームにおける協力率の推移

皆が自分の得になるように振る舞った結果として全体の不利益が生まれる状況を社会的ジレンマ (social dilemma) という。社会的ジレンマ状況では他者が協力するか否かに関係なく、協力しないほうが自身の得になる。社会的ジレンマの利得構造を持つ実験課題を社会的ジレンマゲーム (social dilemma game: SDG)¹ という。SDG での意思決定を繰り返し行くと、初期には高い協力率が次第に低下する推移が見られることが多くの研究で観察されている (e. g. Davis, 1993; Issac and Walker, 1998)。

全体的な協力率の高さや協力率の個人差については、社会的価値志向性² (social value orientation) によって説明される。SVO は、自分だけでなく他者の利得を考慮するような選好であり、社会的ジレンマでの協力の個人差を規定する重要な要因である (レビューとして Au and Kwong, 2004; Balliet et al. ,2009; Bogaert et al. ,2008)。しかし、社会的価値志向性は協力の個人差を表す理論であるため、初期の協力の高さやその個人差を説明することはできない。

協力率が低下する推移が見られることの説明として、学習仮説 (e.g. Palfrey and Prisbrey, 1997) と条件つき協力仮説 (e.g. Andreoni, 1995; Fischbacher, Gächter and Fehr, 2001) が提案されてきた。学習仮説では、協力率の減衰を利得構造の学習によって説明する。すなわち、ゲームの初期には参加者が社会的ジレンマの利得構造を十分に理解しておらず「混乱」によって協力を選択しているが、ゲームを繰り返すにつれて協力しないほうが自身の得になることを学習すると考える。一方、条件つき協力仮説では、協力率の減衰そのものを戦略であると考え、他者の協力の程度に応じて自身も協力する戦略に基づいて行動するが、他者より多く協力することを避けるため、全体として次第に協力率が減る。のちに、他者の協力の程度だけでなく自身の直近の行動によっても行動が変わる傾向が見られることが分かり、気まぐれな条件つき協力 (moody conditional cooperation) と呼ばれている。

近年、気まぐれな条件つき協力傾向の背後にあるメカニズムとして強化学習 (reinforcement learning) があるという議論がある (Ezaki and Horita and Takezawa and Masuda, 2016)。気まぐれな条件つき協力のもつ二つの特徴、すなわち他者の行動に応じて自身の行動を変える特徴と自身が直前にとった行動によって次の行動が変わる特徴は、自身の得られた利得によって各選択肢の価値を更新する強化学習モデルとよく似ている。実

¹ 囚人のジレンマゲーム (prisoner's dilemma game) や公共財ゲーム (public goods game) などがあるが、本稿ではそれらをまとめて社会的ジレンマゲームと呼ぶ。

² 社会選好 (social preference) と同義である。

験データにおいても、強化学習モデルが気まぐれな条件つき協力行動を説明したことが示されている (Horita and Takezawa and Inukai and Kita and Masuda, 2017).

1.2. 本研究の提案モデル

そこで本研究では、SVO モデルと学習モデルを組み合わせた新たなモデルにより、繰り返しの社会的ジレンマゲームにおける協力率の推移の説明を試みる。SVO モデルとして、武藤 (2006) モデル

$$U_i(x_i, x_j) = (1 - a_i)x_i + a_i \frac{1}{n-1} \sum_{j \neq i}^n x_j - e_i \frac{1}{n-1} \sum_{j \neq i}^n |x_i - x_j| \quad (1)$$

を用いる。 x_i は自身の利得、 x_j は他者の利得、 n はプレイヤー人数、 a_i は利他性、 e_i は平等性を表す。このモデルは数理的に望ましい性質を持ち (武藤, 2009)、経験的にもデータによく当てはまる (水野・清水, 2020)。また、Fehr and Schmidt (1999) モデルと数理的に等しいことが示されている (武藤, 2006)。

N 人の SDG で協力、もしくは非協力が選択されるメカニズムを考える。このとき、プレイヤーが得られる利得は、 $N - 1$ 人のプレイヤーのうち何人が協力したか (Y_k とする) による。何人が協力するかの確率を $P(Y_k | \pi_i)$ とする。 π_i は、プレイヤーのもつ、他者の協力に対する期待 (信念) を表す。 π_i は $n - 1$ 人のプレイヤーに対して共通で、独立であるとする s 。

本モデルでは、ソフトマックス行動選択 (softmax action selection) に基づき、協力 (C) と非協力 (D) の効用差から協力が選択される確率が生成されると考える。(1) に基づくと期待効用差、すなわち非協りに比べて協力を選びやすい傾向は

$$EU_i^C - EU_i^D = \sum_k^{n-1} P(Y_k | \pi_i) U_i(x_i^C, x_j^C) - \sum_k^{n-1} P(Y_k | \pi_i) U_i(x_i^D, x_j^D) \quad (2)$$

となる。(2) の x_i^C と x_j^D にそれぞれ、

$$x_i^C = rC(Y_k + 1) \quad (3)$$

$$x_i^D = rCY_k \quad (4)$$

を代入 (r は一人当たりの限界収益率 MPCR; 以下収益率、 C は協力にかかるコスト) し、 $P(Y_k | \pi_i)$ を試行数 $n - 1$ の二項分布であると仮定して整理すると、

$$EU_i^C - EU_i^D = C\{(r - 1) + a_i + 2e_i(\pi_i - 0.5)\} \quad (5)$$

となる。収益率と利他性は協力に対して線形に正の効果を持ち、平等性と他者の協力に対する期待 (0~1 の範囲をとる確率) はそれらの交互作用によって協りに影響することが表現されている。

式(1)のモデルのどのパラメータが学習されているのか確かめるために、4つのモデルを用

意した。(5)の π_i (他者の協力に対する期待)が学習される信念学習モデル, (5) r (収益率) がゲームを繰り返すごとに学習され, 正しい値に近づく収益率学習モデル, そのどちらも学習される信念・収益率学習モデル, 自身の得られた利益に応じて協力, 非協力の効用をそれぞれ更新する強化学習モデル (ベンチマークモデル) である.

2. 実験データによるモデルの実証

上記のモデルのうちどのモデルが最もよく行動データにフィットするか確かめるため, 二度の実験を行ったのち, ベイズ統計モデリングを用いて各モデルの情報量規準 (負の対数周辺尤度) を算出した. 実験での参加者の課題は, 社会選好のパラメータである利他性と平等性の個人ごとの値を推定するための課題 (水野・清水, 2020) の二つであった.

2.1. 実験 1

社会心理学の授業を受講する学生を対象とした実験を行った. 実験には, 4 人の参加者で相互作用するゲーム (一般的な SDG) を行うと教示するソーシャル条件と, コンピュータ相手に一人でゲームを行うと教示するギャンブル条件を設けた. 両条件の利得構造は同じで, 教示のみが異なった. 両条件とも実際には参加者同士での相互作用はしておらず, 実験者側で予め決められたフィードバックを参加者全員に対して行った.

SDG において参加者は, ゲームのピリオド開始時に与えられた 20 トークン (仮想通貨) を使ってグループに貢献するかしないか選択した. 全員がグループに貢献した金額は 1.6 倍されて 4 人の参加者に分配されると教示された. ピリオドは 20 回繰り返された. 参加者が課題内で得たトークンは, 1 トークン 1 円の電子版 Amazon ギフトカードとして実際に支払われた.

結果を表 1 および図 1 に示した. 情報量規準 (負の対数周辺尤度) によるモデル比較を行った結果, ソーシャル条件においては信念・収益率学習モデルの当てはまりが最もよく, 一方ギャンブル条件では, 強化学習モデルの当てはまりが最もよかった.

2.2. 実験 2

実験 2 では 20 ピリオドの繰り返しのあと次第に他者の協力率が上昇するフィードバックを行い, 計 40 回の SDG を行った. 実験条件は実験 1 と同様であった. 参加者が課題内で得たトークンは, 1 トークン 0.5 円の電子版 Amazon ギフトカードとして実際に支払われた.

実験 2 でも, ソーシャル条件において信念・収益率学習モデルの当てはまりが最もよいことが示された (表 1 および図 1).

表 1 モデル比較の結果 (数値は負の対数周辺尤度)

	実験1		実験2	
	ソーシャル条件	ギャンブル条件	ソーシャル条件	ギャンブル条件
信念学習	736.69	682.90	1264.94	1565.55
収益率学習	723.13	639.34	1293.83	1465.73
信念・収益率学習	720.28	674.24	1213.76	1544.00
強化学習	729.00	631.28	1299.33	1409.47

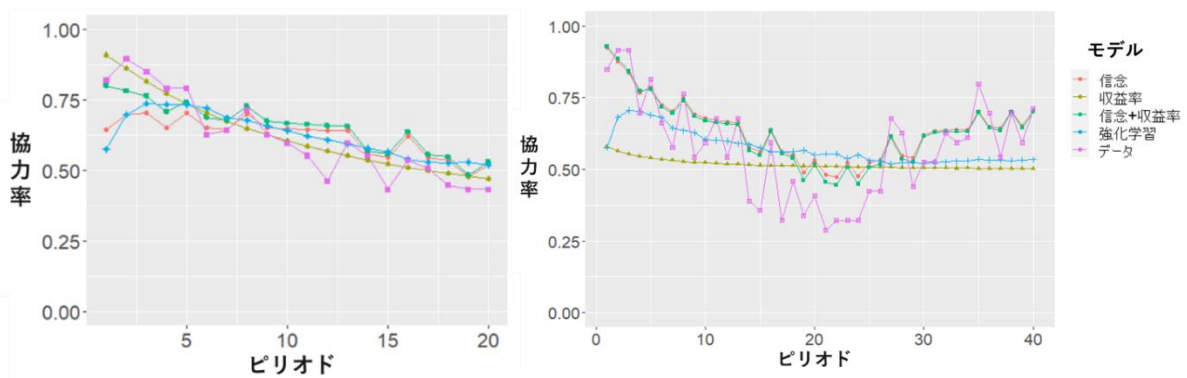


図 1 事後予測チェック³の結果 (左が実験 1, 右が実験 2)

3. 考察

本研究では、導出したモデル (5) に基づく学習モデル三つとベンチマークモデルとして用意した強化学習モデルのなかで、他者の協力に対する期待 (信念) と収益率 (SDG の利得構造) の両方を学習するモデルが最もよく実験データを説明した。このことから、プレイヤーが SDG の利得構造を学習して非協力を選び始めると他者の協力に対する期待が下がり、他者の協力に対する期待が 50%を切ると、自身のもつ利得の不平等を回避する傾向によってさらに非協力を選ぶようになる、というメカニズムが示唆される。

本来であればSDGでは他者の行動に関係なく非協力を選んだほうが自身の得になるのにも関わらず社会的価値志向性があることによって他者の協力への期待が協りに影響する本研究のモデルは、社会心理学における目標/期待理論 (Pruitt & Kimmel, 1977) と整合する。

本研究の意義として、協力・非協力の意思決定に関する解像度の高い数理モデルを構築したことによる協力研究の発展可能性、そして現実の社会問題への介入可能性が挙げられる。本研究のモデルは、罰制度やコミュニケーションの機会がある状況でどの要素が変化した

³ モデルに完全に当てはまったときにどのようなデータになるかをシミュレーションしたもの。

結果として協力率が上がっているのか、集団間の SDG ではどの要素が変化するのかといった様々な研究の基礎となり得る。

引用文献

- Andreoni, J. 1995. Cooperation in public-goods experiments: kindness or confusion?.
The American Economic Review, 85, 891-904.
- Davis, D.D., and Holt, C. A. 1993. Experimental economics. Princeton University Press, Princeton, NJ.
- Dawes, R. M. 1980. Social dilemmas. Annual review of psychology, 31, 169-193.
- Ezaki T, Horita Y, Takezawa M, Masuda N. 2016. Reinforcement learning explains conditional cooperation and its moody cousin. PLoS computational biology, 12, 1-13.
- Fehr, E., and Schmidt, K. M. 1999. A theory of fairness, competition, and cooperation. The quarterly journal of economics, 114, 817-868.
- Horita Y, Takezawa M, Inukai K, Kita T, Masuda N. 2017. Reinforcement learning accounts for moody conditional cooperation behavior: experimental results. Scientific reports, 7, 39275.
- Fischbacher, U., Gächter, S., and Fehr, E. 2001. Are people conditionally cooperative? Evidence from a public goods experiment. Economics letters, 71, 397-404.
- Isaac, R. M., and Walker, J. M. 1988. Group size effects in public goods provision: The voluntary contributions mechanism. The Quarterly Journal of Economics, 103, 179-199.
- 水野景子・清水裕士. 2020. 不平等忌避モデルの経験的検討: ベイズ統計モデリングによるモデル比較. KG 社会学批評, 9, 41-51.
- 武藤正義 2006. 多様な社会的動機の基礎理論. 理論と方法, 21, 63-76.
- 武藤正義 2009. 二者関係における社会的動機の公理論的導出. 理論と方法, 24, 301-316.
- Palfrey, T. R., and Prisbrey, J. E. 1997. Anomalous behavior in public goods experiments: How much and Why?. The American Economic Review, 87, 829-846.
- Pruitt, D. G., & Kimmel, M. J. 1977. Twenty years of experimental gaming: Critique, synthesis, and suggestions for the future. Annual review of psychology, 28, 363-392.