



繰り返しN人囚人のジレンマゲームに関する一考察

メタデータ	言語: Japanese 出版者: 公開日: 2008-05-21 キーワード (Ja): キーワード (En): 作成者: 若林, 高明 メールアドレス: 所属:
URL	https://doi.org/10.32150/00005675

繰り返しN人囚人のジレンマゲームに関する一考察

若林 高明

北海道教育大学旭川校情報科学

A Study of Iterated N-person Prisoners' Dilemma Games

WAKABAYASHI Taka'aki

Information Science, Asahikawa Campus, Hokkaido University of Education, Asahikawa 007-8621

ABSTRACT

Iterated N-person Prisoners' Dilemma Games (INPD) are required for investigations of collective behavior. We conduct a numerical experiment of Iterated N-person Prisoner's Dilemma Games by computer simulation in this paper. We present that cooperative behavior can be established under a certain environment.

1 はじめに

公共財の相互作用や集合行動の問題の中での個人の選好は、一般的には囚人のジレンマゲーム (Prisoners' Dilemma Game)^[1]における選好になるとされる。囚人ジレンマゲームは本来、二人ゲームであるが、現実の世界の協力に関する興味深い問題の多くは二人以上の当事者を含んでいる。そのため、二人ゲームよりもN人ゲーム ($N \geq 3$) の分析により注意が向けられるべきである。本稿では、N人ゲームの場合^[2]を扱う。いずれの場合も、1回限りのゲームにおいては、各プレイヤーが取り得る戦略である「協調」(C)と「裏切り」(D)のうち、「裏切り」が唯一の優越されない戦略である。即ち、どのプレイヤーも、他のプレイヤーがどうしようと「協調」を取るより「裏切り」を取る方が高い利得を獲得する。そのため、主な考察の対象は、有限回または加算無限回のゲームが繰り返し行われる場合である。

繰り返し二人囚人のジレンマゲームに関しては、著名なAxelrodのトーナメント^[3]を始めとする多数の考察がなされ、繰り返しN人囚人ジレンマゲームにおいてもTaylor^[4]などにより考察がなされているが、二人ゲーム程ではない。N人囚人ジレンマゲームにおいては、各プレイヤーについて、戦略「裏切り」が「協調」に優越する一方で、どのプレイヤーも、すべてが裏切るという結果 (D, D, \dots, D) よりも、すべてが協力するという結果 (C, C, \dots, C) を選好する。従って、主な興味の対象は、繰り返しゲームを通じてプレイヤー相互の協調的行動が形成されるか否か、形成されるとすればどのような条件の下でそれがなされるのかということである。本稿では、繰り返しゲームを期に区切って、各プレイヤーが混合戦略を取り、期ご

とに各プレイヤーの戦略が変更されうる条件下でのゲームのコンピュータシミュレーションを通じて、プレイヤー相互の協調的行動の形成が行われるか否かの検出を試みる。

2 繰り返しN人囚人ジレンマゲームの定式化

2.1 N人囚人ジレンマゲーム

本項では、Taylor (1987)^[4]に基づいて、N人囚人ジレンマゲーム (NPD) の定式化を述べる。

ゲームに参加する個人をプレイヤーと呼ぶ。N人のプレイヤーは、戦略C (協調) とD (裏切り) を選択できる。戦略の選択は全員が同時に行う。即ち、それぞれ他のプレイヤーの戦略を知らぬままに戦略を選択しなければならない。プレイヤー*i*に着目するとして、*i*がCを選び、*i*以外の v ($v < N$) 人のプレイヤーがCを選んだときのゲームの利得を $f(v)$ とし、*i*がDを選び、*i*以外の v 人のプレイヤーがCを選んだときのそれを $g(v)$ とする。 $f(v)$ と $g(v)$ は全てのプレイヤーにとって同じであるとする。

利得関数 f と g に関して以下の三つが仮定される。

1. $v \geq 0$ の各値に対して $g(v) > f(v)$
2. $f(N-1) > g(0)$
3. すべての $v > 0$ に対して $g(v) > f(v)$

上記の条件1.~3.を満たす利得関数 f 、 g の一例のグラフを図1に示す。

2.2 繰り返しN人囚人ジレンマゲーム

繰り返しゲームとはゲームの有限回または加算無限回の連続であり、連続している各ゲームを繰り返しゲームの要素ゲームという。INPDは、N人囚人ジレンマゲームを要素ゲームとする繰り返しゲームである。

繰り返しゲームを通じて、プレイヤー間の協調的行動が現れる、即ち、協調的戦略を取るプレイヤーの比率が高い状態が出現することが望ましい。非協調的戦略を取るプレイヤーの比率が高くなってしまうと、全プレイヤーの利得の合計が減少するからである。

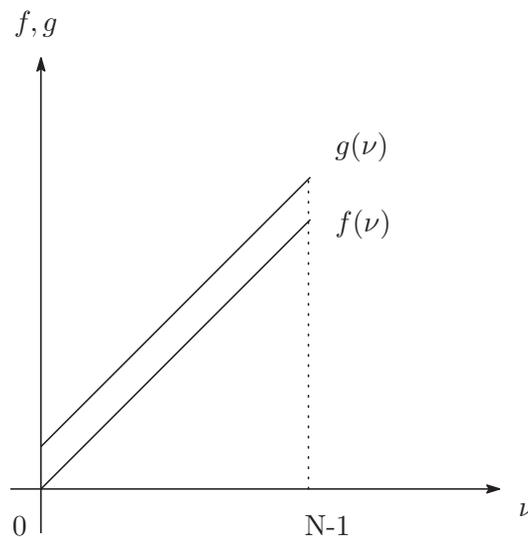


図1. 利得関数の一例

3 シミュレーションの概要

2. 2項で定式化したINPDのシミュレーションを行う。以下にその概要を述べる。ゲームの繰り返しの回数は有限回とする。

時点 $t=0$ においてゲームが開始される。以後、1時点ごとに要素ゲームが行われ、要素ゲーム m 回を1期とし、各期の始めに戦略の見直しを行う。

繰り返しゲームにおいて、各プレイヤーが各期ごとに取りうる戦略を以下とする。各プレイヤーは、繰り返しゲームを通じて混合戦略を取り、戦略の確率ベクトルに基づき、乱数発生により各期首において下記の戦略のうち一つを選択する。

B_n : 条件付協調戦略。繰り返し二人囚人ジレンマゲームにおけるしっぺ返し (TIT-FOR-TAT) 戦略を N 人の場合に一般化したものである。最初のゲームでは無条件に C を取り、2回目以降のゲームでは、前回の要素ゲームで全プレイヤーのうち自身を除く n ($n < N$) 人が C を取ったとき、かつそのときに限り C を取る。

C^∞ : 全面協調戦略。常に C を取る。

D^∞ : 全面裏切り戦略。常に D を取る。

各プレイヤーは、上記の戦略を各期に一つだけ選択できる。

各プレイヤーが取りうる上記の戦略 $B_n/C^\infty/D^\infty$ をそれぞれ、戦略 $0, 1, 2$ とし、あるプレイヤーが第 t 期において各戦略を p_0^t, p_1^t, p_2^t ($p_0^t + p_1^t + p_2^t = 1$) の確率で取るものとする。これらの確率は、前期の結果および次期の期待が考慮されて遷移するものとする。但し、最初の期である第0期に取る戦略は、あらかじめ設定された比率に基づいて乱数発生により決定されるものとする。

第1期以降は、以下の方法により戦略 j ($j = 0, 1, 2$) を取る確率を決定する。一般に、人間や動物は過去にうまくいった行動を繰り返し、うまくいかなかった行動を忌避する。これをシミュレーションに取り入れるため、過去の利得と次期の戦略選択に反映させる方法を採用する。加えて、将来の利得に対する期待をも次期の戦略選択に反映させることとする。

第 t 期 ($t > 0$) に各戦略 j ($j = 0, 1, 2$) を取る確率は、前期 (第 $t-1$ 期) の利得に基づく報酬 R_j^{t-1} と第 t 期の利得に対する期待 E_j^t が反映されて第 t 期の期首に更新されるものとする。但し、 R_j^{t-1} は、第 $t-1$ 期に戦略 j ($j = 0, 1, 2$) を取ったときの1要素ゲーム当たりの利得の平均値があらかじめ設定された期待水準の値を上回ったとき、かつそのときに限り、第 t 期に戦略 j が採用される確率 p_j^t の計算に反映されるものとする。

本シミュレーションにおいては、簡潔を旨とするため、必要な値の計算は簡略な方法を用いる。

期待利得は以下により求める。

次期の戦略が B_n の場合：

第 t 期の利得の期待値は、平均的見積り e_{avr} 、悲観的見積り e_p および楽観的見積り e_o を計算し、これらより、以下の3点見積りを計算することにより求める。

$$E_0^t = E^t(B_n) = \frac{e_p + 4e_{avr} + e_o}{6} \quad (1)$$

平均的な見積りは、前期の協調者数の平均値 (nc^{t-1}) と同数が協調すると考えて行う。着目しているプレイヤー自身が協調か裏切りのどちらであったかによって、利得関数における自身を除く協調者数が nc^{t-1} と $nc^{t-1} - 1$ になる場合があるが、これによる利得関数の値の差が利得に与える影響は小さいと考えて、以

下では前回の自身の手に拘わらず協調者数を nc^{t-1} として期待利得の見積りを行う。

$$e_{avr} = \begin{cases} f(nc^{t-1}) \sum_{i=1}^m \delta^i, & nc^{t-1} \geq n \\ \delta f(nc^{t-1}) + g(nc^{t-1}) \sum_{i=2}^m \delta^i, & nc^{t-1} < n \end{cases} \quad (2)$$

ここで、 δ は未来係数であり、1 時点先の要素ゲームの利得が現在の δ ($0 < \delta < 1$) 倍の価値に割り引かれるということである。

悲観的見積りにおいては、自身以外の全員が裏切ると考える。従って、自身の手は 1 回目から C 。2 回目以降が D となるので、見積りは以下の式で与えられる。

$$e_p = \delta f(0) + g(0) \sum_{i=2}^m \delta^i \quad (3)$$

楽観的見積りにおいては、自身以外の全員が協調すると考える。見積りは以下の式で与えられる。

$$e_o = f(N-1) \sum_{i=1}^m \delta^i \quad (4)$$

第 t 期の戦略が C^∞ の場合。

$$E_1^t = E^t(C^\infty) = \frac{(f(0) + 4f(nc^{t-1}) + f(N-1)) \sum_{i=1}^m \delta^i}{6} \quad (5)$$

第 t 期の戦略が D^∞ の場合。

$$E_2^t = E^t(D^\infty) = \frac{(g(0) + 4g(nc^{t-1}) + g(N-1)) \sum_{i=1}^m \delta^i}{6} \quad (6)$$

第 t 期の各戦略の強度ベクトル cw^t の各成分は

$$cw_j^t = \begin{cases} cw_j^{t-1} + r_j^{t-1} + e_j^t & \text{第 } t-1 \text{ 期の戦略が } j \text{ かつ第 } t-1 \text{ 期の得点が期待水準値以上の場合} \\ cw_j^{t-1} + e_j^t & \text{上記以外の場合} \end{cases} \quad (7)$$

で与えられる。但し、初期強度ベクトル cw^0 の各成分はあらかじめ設定されるものとする。但し、上記の r_j^{t-1} は前期 (第 $t-1$ 期) の戦略 j による利得の 1 要素ゲームあたりの平均値であり、 e_j^t は、上記で計算した $E^t(j)$ ($j = 0, 1, 2$) を 1 要素ゲームあたりの得点に基準化したものである。各戦略を取る確率は、強度ベクトルの成分を、その和が 1 となるように基準化することにより求められる。

$$p_j^t = \frac{cw_j^t}{\sum_{k=0}^2 cw_k^t} \quad (j = 0, 1, 2) \quad (8)$$

4 シミュレーションの条件および結果

以下の条件でシミュレーションを行った。

・実行環境

OS : SUSE Linux 10.0, Compiler : GNU C compiler version 4.02,

CPU : Dual Xeon 3.0GHz, Memory : 2GB。

・パラメータ

プレイヤー数 N : 7, 戦略 B_n における n : 3, 1期当たりの要素ゲーム数 m : 5,

要素ゲームの繰り返し回数 T : 100, 期数 T/m : 20,

利得関数 : $f(v) = v, g(v) = v + 1$ ($v = 0, 1, \dots, 6$),

利得の期待水準 : 4, 未来係数 δ : 0.9,

戦略の確率ベクトルの初期値 (p_0^0, p_1^0, p_2^0) : (0.1, 0.1, 0.8)。

以下では、戦略 C^∞ と B_n を協調的戦略と呼ぶ。最初は非協調的戦略を取るプレイヤーの比率が高い状態（各戦略の確率ベクトル $(p_0^0, p_1^0, p_2^0) = (0.1, 0.1, 0.8)$ ）から出発して、繰り返しゲームの間に、協調的戦略を取るプレイヤーの比率が高い状態に転換できるか否かに着目する。図2に、50回のシミュレーションを実行したときの「プレイヤー数×期数」に占める協調的戦略数の比率の平均値の推移を示す。横軸は期数を、縦軸は協調的戦略数の比率を示す。平均的には、期数を重ねるに従って、協調的戦略数の比率が増加し、繰り返しゲーム全体を通じて取られる戦略の過半数が協調的戦略になっていることが分かる。

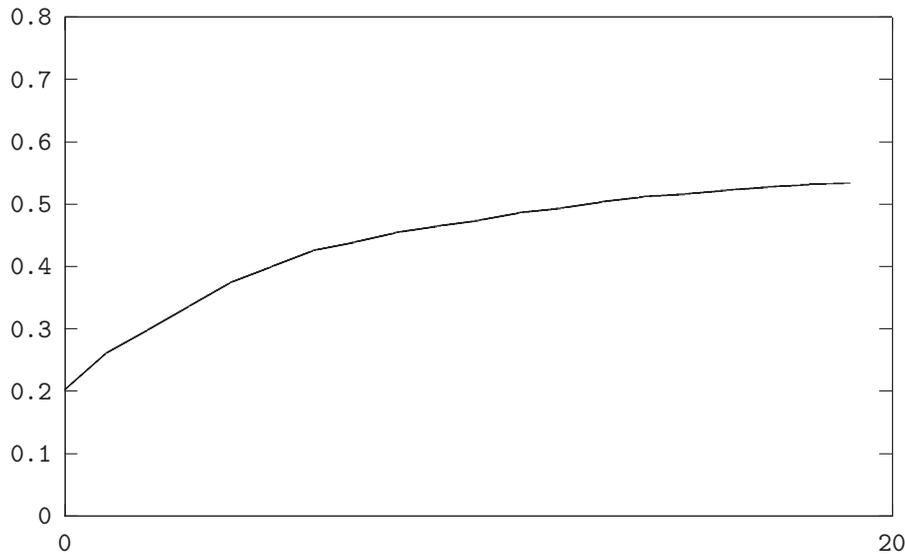


図2. 協調的戦略の割合の平均値の推移

5 おわりに

本稿においては、繰り返しN人囚人ジレンマゲーム (INPD) のシミュレーションを行い、繰り返しゲームを期に区切って、各プレイヤーが混合戦略を取り、期ごとに戦略の変更を可能とする条件下で、各期の戦

略比率の決定に過去の利得に基づく報酬と将来の利得に対する期待を反映させることにより、協調的戦略を取る比率を増加させることが出来るということを確認した。今後の展望としては、戦略の拡張やプレイヤー毎のパラメータの変化などを取り入れた、より精緻なシミュレーションを行った場合の協調的な状態をもたらす条件を探索することなどが挙げられる。

参考文献

- [1] Rapoport A. and Chammah A.M., *Prisoners' Dilemma*, 270pp., University of Michigan Press (1965)
- [2] Hamburger H., *N-person Prisoners' Dilemma*, *Journal of Mathematical Sociology* **3**, pp.27-48(1973)
- [3] Axelrod R., *The Evolution of Cooperation*, 256pp., Basic Books (1984)
=ロバート・アクセルロッド著, 松田裕之訳, つきあい方の科学, 255pp., HBJ 出版局 (1987)
- [4] Taylor M., *Possibility of Cooperation ; Studies in Rationality and Social Change*, 220pp., Cambridge University Press (1987)
=マイケル・テーラー著, 松原望訳, 協力の可能性, 258pp., 木鐸社 (1995)

(旭川校准教授)