数理解析研究所講究録 第2114巻 2019年 191-196

繰り返し多目的ゲームのフォーク定理について

島根大学大学院　総合理工学研究科　兼原　眞 (Makoto Kanehara)
Interdisciplinary Graduate School of Science and Engineering, Shimane University
島根大学　総合理工学部　黒岩大史 (Daishi Kuroiwa)
Interdisciplinary Faculty of Science and Engineering, Shimane University

概要

囚人のジレンマは，最も有名な戦略形ゲームの一つである． このゲームではプレイヤーが互いに協力する戦略はナッシュ均衡点にはならないが，無限繰り返しゲームでは，プレイヤーが互いに協力する戦略がナッシュ均衡点となる．これはフォーク定理によって保証されている． 利得関数がベクトル値関数によって与えられるゲームは多目的ゲームと呼ばれている [2]． 多目的ゲームにおけるナッシュ平衡は，[3] [4] 等で研究されている． 本論文では，繰り返し多目的ゲームにおけるイデアルナッシュ均衡点と弱パレートナッシュ均衡点を観測し，繰り返し多目的ゲームのフォーク定理を与える．

1　無限繰り返しゲームとフォーク定理

定義 1.1

$N=\{1,2,\cdots,n\}$ をプレイヤーの集合，空でない集合 A_i をプレイヤー $i\in N$ の戦略集合，$f_i: A_1\times\cdots\times A_n\to\mathbb{R}$ をプレイヤー $i\in N$ の利得関数とする．このとき $G=(N,\{A_i\}_{i\in N},\{f_i\}_{i\in N})$ を戦略形 n 人ゲームという．

定義 1.2

$G=(N,\{A_i\}_{i\in N},\{f_i\}_{i\in N})$ を戦略形 n 人ゲームとする．プレイヤーの戦略の組 $a^*=(a_1^*,\ldots,a_n^*)\in A_1\times\cdots\times A_n$ がナッシュ均衡点とは，

$$^\forall i\in N, {}^\forall a_i\in A_i, f_i(a_i,a_{-i}^*)\leqq f_i(a_i^*,a_{-i}^*)$$

を満たすときをいう．

例 1.1 (囚人のジレンマ)

犯罪組織の 2 人のメンバー A と B が逮捕され，それぞれの囚人は別々に投獄されて取り調べを受けていて互いにコミュニケーションを取ることが出来ない．A と B が互いに相手を裏切り自白すると，二人とも 9 年の懲役刑を受ける．もし A が自白をして B が黙秘をすれば A は罰せられず，B は 10 年の懲役刑を受ける（逆もまた同様）．互いに協力して黙秘をすれば，二人とも 1 年の懲役刑を受ける. この状況を表にすると以下のようになる．

A \ B	協力	裏切り
協力	$(-1,-1)$	$(-10,0)$
裏切り	$(0,-10)$	$(-9,-9)$

この時，(裏切り，裏切り) がナッシュ均衡点となる．

定義 1.3

$G=(N,\{A_i\}_{i\in N},\{f_i\}_{i\in N})$ を戦略形 n 人ゲームとし，$\delta\in(0,1)$ を割引因子とする．

プレイヤー i の戦略集合を $S_i=\{s_i=\{s_i^t\}\mid s_i^1\in A_i, s_i^t: A^{t-1}\to A_i(t\geqq 2)\}$ (ただし，$A^{t-1}=A\times\cdots\times A$($A$ の $t-1$ 回の直積)) とおき，プレイヤー i の利得関数 $F_i:S_1\times\cdots\times S_n\to\mathbb{R}$ を

$$F_i(s)=\sum_{t=1}^{\infty}\delta^{t-1}f_i(\alpha^t(s))$$

とする．ただし，

$$\alpha^t(s)=\begin{cases}(s_1,\ldots,s_n)\in A & (t=1)\\ (s_1^t(\alpha^1(s),\ldots,\alpha^{t-1}(s)),\ldots,s_n^t(\alpha^1(s),\ldots,\alpha^{t-1}(s))) & (t\geq 2),\end{cases}$$

とする．

このとき，$G^\infty(\delta)=(N,\{S_i\}_{i\in N},\{F_i\}_{i\in N})$ を戦略形 n 人ゲーム G の割引因子 δ を持つ無限繰り返しゲームという．

この定義における A^{t-1} は $t-1$ 回目までの履歴と呼ばれる．$G^\infty(\delta)$ も戦略形 n 人ゲームと解釈出来るので，プレイヤーの戦略の組 $s^*=(s_1^*,\ldots,s_n^*)\in S_1\times\cdots\times S_n$ がナッシュ均衡点であるとは，

$$^\forall i\in N, {}^\forall s_i\in S_i, F_i(s_i,s_{-i}^*)\leqq F_i(s_i^*,s_{-i}^*).$$

を満たすときをいう．

例 1.2 (囚人のジレンマの無限繰り返しゲーム)

繰り返しゲーム G^∞ における代表的な戦略として次の四つの戦略を考える．

- All-C : 過去のプレイによらず，常に協力 (cooperate) をとる．
- All-D : 過去のプレイによらず，常に裏切り (denounce) をとる．
- Trigger : 最初は協力をとる．しかし相手が裏切れば，それ以後裏切りをとり続ける．
- TFT(tit for tat) : 最初は協力をとる．以後，相手の前回の行動と同じものをとり続ける．

四つの戦略によるプレイヤーの利得は以下のようになる．

ただし全ての利得は $(1-\delta)$ 倍している．

A \ B	All-C	All-D	Trigger	TFT
All-C	$(-1,-1)$	$(-10,0)$	$(-1,-1)$	$(-1,-1)$
All-D	$(0,-10)$	$(-9,-9)$	$(-9\delta,-10+\delta)$	$(-9\delta,-10+\delta)$
Trigger	$(-1,-1)$	$(-10+\delta,-9\delta)$	$(-1,-1)$	$(-1,-1)$
TFT	$(-1,-1)$	$(-10+\delta,-9\delta)$	$(-1,-1)$	$(-1,-1)$

表より $\delta \geq \dfrac{1}{9}$ のとき，(Trigger, Trigger) がナッシュ均衡点となり得る．従って，互いに協力を選ぶ戦略がナッシュ均衡点になり得る．

次のフォーク定理によりジレンマの解消が保障されている．

定理 1.1 (フォーク定理)

$G^{\infty}(\delta) = (N, \{S_i\}_{i\in N}, \{F_i\}_{i\in N})$ を戦略形ゲーム G の割引因子 δ を持つ無限繰り返しゲームとする．δ と $a = (a_1, \ldots, a_n) \in A_1 \times \cdots \times A_n$ が $\forall i \in N$,

- $\inf\limits_{b_{-i}} \sup\limits_{b_i} f_i(b_i, b_{-i}) < f_i(a)$
- $\dfrac{\sup_{b_i} f_i(b_i, a_{-i}) - f_i(a)}{\sup_{b_i} f_i(b_i, a_{-i}) - \inf_{b_{-i}} \sup_{b_i} f_i(b_i, b_{-i})} \leqq \delta$

を満たすならば $G^{\infty}(\delta) = (N, \{S_i\}_{i\in N}, \{F_i\}_{i\in N})$ のナッシュ均衡点 $s^* = (s_1^*, \ldots, s_n^*) \in S_1 \times \cdots \times S_n$ が存在して，$^{\forall} t \in \mathbb{N}, \alpha^t(s^*) = (a_1, \ldots, a_n)$ が成り立つ．

2　無限繰り返し多目的ゲームとそのフォーク定理

ここでは，[5] で示した無限繰り返し多目的ゲームのフォーク定理を紹介し，その応用例を述べる．まず，目的を複数にした多目的ゲームについて紹介し．無限繰り返し多目的ゲームの定式化をし，多目的ゲームと同様の解を定義する．

多目的ゲームを考えるモチベーションとしては，我々は例えば買い物をする時には価格，品質等の複数の目的重要視する．多目的ゲームを考える動機は実際の問題で最適化を行う場合目的が複数個あるためである．

定義 2.1

$N = \{1, 2, \cdots, n\}$ をプレイヤーの集合，空でない集合 A_i をプレイヤー $i \in N$ の戦略集合，m_i をプレイヤー i の目的の個数，$f_i : A_1 \times \cdots \times A_n \to \mathbb{R}_i^m$ をプレイヤー $i \in N$ の利得関数とする．このとき $G = (N, \{A_i\}_{i\in N}, \{f_i\}_{i\in N})$ を n 人多目的ゲームという．

イデアルナッシュ均衡点と弱パレート均衡点を定義するために次の関係を定義する．

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}, z = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_m \end{pmatrix} \in \mathbb{R}^m$$ とする．このとき

$$y \leq_{\mathbb{R}^m_+} z \overset{\text{def}}{\Longleftrightarrow} y_i \leq z_i, \quad {}^\forall i = 1, 2, \ldots, m$$

$$y <_{\mathbb{R}^m_+} z \overset{\text{def}}{\Longleftrightarrow} y_i < z_i, \quad {}^\forall i = 1, 2, \ldots, m$$

定義 2.2

$G = (N, \{A_i\}_{i \in N}, \{f_i\}_{i \in N})$ を n 人多目的ゲームとし，それぞれのプレイヤーの目的の個数を $m_1, \ldots, m_n$ とする．プレイヤーの戦略の組 $a^* \in \Pi_{i=1}^n A_i$ がイデアルナッシュ均衡点であるとは，

$$\forall i \in N, \forall a_i \in A_i, f_i(a_i, a^*_{-i}) \leq_{\mathbb{R}^{m_i}_+} f_i(a^*_i, a^*_{-i})$$

を満たすときである．

定義 2.3

$G = (N, \{A_i\}_{i \in N}, \{f_i\}_{i \in N})$ を n 人多目的ゲームとし，それぞれのプレイヤーの目的の個数を $m_1, \ldots, m_n$ とする．プレイヤーの戦略の組 $a^* \in \Pi_{i=1}^n A_i$ が弱パレートナッシュ均衡点であるとは，

$$\forall i \in N, \forall a_i \in A_i, f_i(a^*_i, a^*_{-i}) \not<_{\mathbb{R}^{m_i}_+} f_i(a_i, a^*_{-i})$$

を満たすときである．

定義 2.4

$G = (N, \{A_i\}_{i \in N}, \{f_i\}_{i \in N})$ を n 人多目的ゲームとし，それぞれのプレイヤーの目的の個数を $m_1, \ldots, m_n$, $\delta \in (0, 1)$ を割引因子とする．プレイヤー i の戦略集合を

$S_i = \{s_i = \{s^t_i\} \mid s^1_i \in A_i, s^t_i : A^{t-1} \to A_i (t \geqq 2)\}$ (ただし，$A^{t-1} = A \times \cdots \times A$($A$ の $t-1$ 回の直積)) とおき，プレイヤー i の利得関数 $F_i : S_1 \times \cdots \times S_n \to \mathbb{R}^{m_i}$ を

$$F_i(s) = \sum_{t=1}^{\infty} \delta^{t-1} f_i(\alpha^t(s))$$

とする．ただし，

$$\alpha^t(s) = \begin{cases} (s_1, \ldots, s_n) \in A & (t = 1) \\ (s^t_1(\alpha^1(s), \ldots, \alpha^{t-1}(s)), \ldots, s^t_n(\alpha^1(s), \ldots, \alpha^{t-1}(s))) & (t \geq 2), \end{cases}$$

とする．

このとき，$G^\infty(\delta) = (N, \{S_i\}_{i \in N}, \{F_i\}_{i \in N})$ を n 人多目的ゲーム G の割引因子 δ を持つ無限繰り返し多目的ゲームという．

定義 2.5

$G^\infty(\delta) = (N, \{S_i\}_{i \in N}, \{F_i\}_{i \in N})$ を n 人無限繰り返し多目的ゲームとし，それぞれのプレイ

ヤーの目的の個数を $m_1, \ldots, m_n$ とする．プレイヤーの戦略の組 $s^* = (s_1^*, \ldots, s_n^*) \in S_1 \times \cdots \times S_n$ がイデアルナッシュ均衡点であるとは，

$$^\forall i \in N, ^\forall s_i \in S_i, F_i(s_i, s_{-i}^*) \leq_{\mathbb{R}_+^{m_i}} F_i(s_i^*, s_{-i}^*)$$

を満たすときである．

定義 2.6

$G^\infty(\delta) = (N, \{S_i\}_{i\in N}, \{F_i\}_{i\in N})$ を n 人無限繰り返し多目的ゲームとし，それぞれのプレイヤーの目的の個数を $m_1, \ldots, m_n$ とする．プレイヤーの戦略の組 $a^* \in A_1 \times \cdots \times A_n$ が弱パレートナッシュ均衡点であるとは，

$$^\forall i \in N, ^\forall a_i \in A_i, F_i(a_i^*, a_{-i}^*) \not<_{\mathbb{R}_+^{m_i}} F_i(a_i, a_{-i}^*)$$

を満たすときである．

多目的ゲームのフォーク定理として次の 2 つの定理を紹介する．

定理 2.1 ([5])

$G^\infty(\delta) = (N, \{S_i\}_{i\in N}, \{F_i\}_{i\in N})$ をそれぞれのプレイヤーの目的の個数が $m_1, \ldots, m_n$ である n 人多目的ゲーム $G = (N, \{A_i\}_{i\in N}, \{f_i\}_{i\in N})$ の割引因子 $\delta \in (0,1)$ を持つ無限繰り返し多目的ゲームとし，$a = (a_1, \ldots, a_n) \in A_1 \times \cdots \times A_n$ とする．このとき任意の $i \in N$ と $j \in \{1, \ldots, m_i\}$ に対して

- $\inf\limits_{b_{-i}} \sup\limits_{b_i} f_{ij}(b_i, b_{-i}) < f_{ij}(a)$
- $\dfrac{\sup_{b_i} f_{ij}(b_i, a_{-i}) - f_{ij}(a)}{\sup_{b_i} f_{ij}(b_i, a_{-i}) - \inf_{b_{-i}} \sup_{b_i} f_{ij}(b_i, b_{-i})} \leqq \delta$

を満たすならば，$G^\infty(\delta) = (N, \{S_i\}_{i\in N}, \{F_i\}_{i\in N})$ のイデアルナッシュ均衡点 $s^* \in S$ が存在して，任意の $t \in \mathbb{N}$ に対して$\alpha^t(s^*) = (a_1, \ldots, a_n)$ が成り立つ．

定理 2.2 ([5])

$G^\infty(\delta) = (N, \{S_i\}_{i\in N}, \{F_i\}_{i\in N})$ をそれぞれのプレイヤーの目的の個数が $m_1, \ldots, m_n$ である n 人多目的ゲーム $G = (N, \{A_i\}_{i\in N}, \{f_i\}_{i\in N})$ の割引因子 $\delta \in (0,1)$ を持つ無限繰り返し多目的ゲームとし，と $a = (a_1, \ldots, a_n) \in A_1 \times \cdots \times A_n$ とする．ある $i \in N$ と $j \in \{1, \ldots, m_i\}$ が存在して，

- $\inf\limits_{b_{-i}} \sup\limits_{b_i} f_{ij}(b_i, b_{-i}) < f_{ij}(a)$
- $\dfrac{\sup_{b_i} f_{ij}(b_i, a_{-i}) - f_{ij}(a)}{\sup_{b_i} f_{ij}(b_i, a_{-i}) - \inf_{b_{-i}} \sup_{b_i} f_{ij}(b_i, b_{-i})} \leqq \delta$

を満たすならば $G^\infty(\delta) = (N, \{S_i\}_{i\in N}, \{F_i\}_{i\in N})$ の弱パレートナッシュ均衡点 $s^* \in S$ が存在して，$^\forall t \in \mathbb{N}, \alpha^t(s^*) = (a_1, \ldots, a_n)$ が成り立つ．

例 2.1

A 国と B 国は敵対していて，両国は戦争の可能性のある状況にある．

- A 国と B 国が互いに武力行使をしないと国の平和と財政は初期状態から変化しない．
- A 国と B 国が互いに武力行使をすると，国の平和と財政は両方とも初期状態から減少する．
- 片方の国が武力行使をして，もう一方の国が武力行使をしないと，武力行使をした側の国の平和と財政が増加し，武力行使をしなかった側の国の平和と財政が減少する．

このゲームでは利得を $\begin{pmatrix} \text{国の平和} \\ \text{財政} \end{pmatrix}$ としている．

A \ B	武力行使をしない	武力行使をする
武力行使をしない	$\begin{pmatrix} 5 \\ 10 \end{pmatrix} \begin{pmatrix} 5 \\ 10 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 1 \end{pmatrix} \begin{pmatrix} 8 \\ 15 \end{pmatrix}$
武力行使をする	$\begin{pmatrix} 8 \\ 15 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 4 \\ 4 \end{pmatrix} \begin{pmatrix} 4 \\ 4 \end{pmatrix}$

国の平和の指標は 0 から 9 とし (0 が一番低く，9 が一番高い), 財政の単位は 100 万ドルとする．このゲームでは, 武力行使をしないことが協力を意味していて, 武力行使をすることが裏切りを意味している．

(武力行使をしない, 武力行使をしない) は (武力行使をする, 武力行使をする) よりも利得が高いので，定理 3.1 より，無限繰り返し多目的ゲームでは，$\delta \geq \frac{3}{4}$ の時 (武力行使をしない, 武力行使をしない) がイデアルナッシュ均衡点になり，定理 3.2 より $\delta \geq \frac{6}{11}$ では，(武力行使をしない, 武力行使をしない) が弱パレート均衡点となる．

参考文献

[1] John Forbes Nash, Non-cooperative games. Annals of Mathematics 54:195-286 (1951)

[2] David Harold Blackwell, An analog of the minimax theorem for vector payoffs. Pacific Journal of Mathematics 6:1-8 (1956)

[3] Peter Borm, Freek van Megen, Stef Tijs, A perfectness concept for multicriteria games. Math. Methods Oper. Res. 49, 401-412 (1999)

[4] Mark Voorneveld, Sofia Grahn, Martin Durwenberg, Ideal equilibria in noncooperative multicriteria games. Math. Metthods. Oper. Res. 52, 65-77 (2000)

[5] Makoto Kanehara, Daishi Kuroiwa, Folk theorems for repeated multi-objective games, preprint.

[6] 岡田 章, ゲーム理論新版, 有斐閣 (2011)