ゲーム表現の拡張と繰り返しゲーム Repeated Game

-- 合理的思考の技術 7 --

小林憲正

Department of Value and Decision Science (VALDES)
Tokyo Institute of Technology

May 28, 2012

Outline

1. 混合戦略 Mixed Strategy とナッシュ均衡の存在定理
2. 同時決定を含む拡張と部分ゲーム完全均衡 SPNE
3. 繰り返しゲームとフォーク定理

（純粋戦略の）ナッシュ均衡が存在しない例

1 / 2	グー	チョキ	パー
グー	0, 0	1, -1	-1, 1
チョキ	-1, 1	0, 0	1, -1
パー	1, -1	-1, 1	0, 0

ナッシュ均衡の存在定理

Theorem (ナッシュ均衡の存在定理 [6])

以下の条件を満たす標準形ゲーム $\langle N, A, u \rangle$ はナッシュ均衡を持つ。
$\forall i \in N$:

- A_i がユークリッド空間 Euclidean space の非空 nonempty でコンパクト compact かつ凸 convex な部分集合
- u_i が連続 continuous かつ A_i 上で擬凹 quasi-concave

Corollary (混合戦略のナッシュ均衡の存在定理)

任意の有限標準形ゲームは混合戦略ナッシュ均衡を持つ。

Example (じゃんけん)

二人じゃんけんの混合戦略ナッシュ均衡は唯一存在し、両者が $1/3$ ずつの確率で、それぞれグー、チョキ、パーを出す。

ゲームの混合拡張 [8]

Definition (混合拡張 Mixed Extension)

標準形ゲーム $\langle N, A, u \rangle$ の混合拡張とは、次の標準形ゲーム $\langle N, \times_{j \in N} \Delta(A_j), U : \times_{j \in N} \Delta(A_j) \to \Re^N \rangle$:

- $\Delta(A_i)$ はプレーヤー$i \in N$ の混合戦略 mixed strategy の集合
- $U_i(\alpha) = \sum_{a \in A} \Pi_{j \in N} \alpha_j(a_j) \cdot u_i(a)$ はくじ $\alpha \in \times_{j \in N} \Delta(A_j)$ の期待効用

確率 1 でいずれかの行動 $a_i \in A_i$ を選択する戦略を純粋戦略 pure strategy という。

混合戦略ナッシュ均衡の性質・問題点

Proposition (サポートの全てが最適応答)

混合戦略の組 $\alpha^* = (\alpha_i^*)_{i \in N} \in \times_{j \in N} \Delta(A_j)$ がナッシュ均衡である iff $\forall i \in N$ について、$\forall a_i \in$ supp α_i^* が α_{-i}^* に対して最適応答となっている

つまり、特に、正確に α_i^* の確率分布に従ってプレーする強いインセンティブがあるわけではない。

Example (じゃんけん)

二人ジャンケンにおいて、相手がグー、チョキ、パーを等確率でだしてくれば、自分はなにを出しても同じ利得が得られる。

Outline

1. 混合戦略 Mixed Strategy とナッシュ均衡の存在定理
2. 同時決定を含む拡張と部分ゲーム完全均衡 SPNE
3. 繰り返しゲームとフォーク定理

完全情報ゲームの定式化

Definition (完全情報ゲーム)

$\langle N, \mathcal{K}, P, u \rangle$:

- N プレーヤーの集合
- $\mathcal{K} = \langle H, A \rangle$ 根付き木 rooted tree:
  - H ノード node(履歴 history) の集合
    最終履歴 terminal history の集合を $Z(\subset H)$ と書く
  - A 枝 arc(行動 actions) の集合
- $P : H \setminus Z \to N$ 各決定ノードに、そこで決定するプレーヤーを対応付けるプレーヤー関数,
- $u_i : Z \to \Re$ プレーヤー$i \in N$ の効用関数.

同時決定を含む拡張 [8]

Definition (同時決定を含む完全情報ゲーム)

（拡張部分のみ提示）$\Gamma = \langle N, \mathcal{K}, P, u \rangle$:

- $P : H \setminus Z \to \mathcal{P}(N)$ 各履歴 $h \in H$ に、その直後に行動するプレーヤーの部分集合を対応させる関数,
- $A(h) = \times_{i \in P(h)} A_i(h)$ 各履歴 h 直後の行動の集合上のプレーヤー分割.

Definition (戦略 Strategy)

同時決定を含む完全情報ゲーム $\Gamma = \langle N, \mathcal{K}, P, u \rangle$ におけるプレーヤー $i \in N$ の（純粋）戦略は、全ての $i \in P(h)$ を満たす最終でない履歴 $h \in H \setminus Z$ に $A_i(h)$ 上の行動を対応付ける関数。

ノード・枝と履歴・行動

完全情報ゲームの木では、決定ノードと履歴、枝と行動が丁度一対一対応する。

- 木の根に対応して、空の履歴 null history（初期履歴 initial history）$\varnothing \in H$ という記号を用いる
- 履歴を空の履歴からの行動プロファイルの列で表すことができる e.g.) $h = (a^1, ..., a^T)$
- 同様に、履歴の合成を考えることができる e.g.) $h = (h', h'')$

部分ゲーム

Definition (部分ゲーム Subgame)

同時決定を含む完全情報ゲーム $\Gamma = \langle N, \mathcal{K}, P, u \rangle$ の履歴 $h \in H$ 後の部分ゲームとは、Γ の構造を全て保存したゲーム $\Gamma(h) = \langle N, \mathcal{K}|_h, P|_h, u|_h \rangle$:

- $\forall h' \in H|_h, (h, h') \in H$
- $\forall h' \in H|_h, P|_h(h') = P(h, h')$
- $\forall h' \in Z|_h, u_i|_h(h') = u_i(h, h')$

Definition (戦略の制限 Restriction)

$\Gamma = \langle N, \mathcal{K}, P, u \rangle$ におけるプレーヤー$i \in N$ の戦略 s_i の部分ゲーム $\Gamma(h)$ 上への制限 $s_i|_h$ は以下を満たす

$$\forall h' \in H|_h, s_i|_h(h') = s_i(h, h')$$

部分ゲーム完全均衡 SPNE

Definition (部分ゲーム完全均衡 SPNE)

同時決定を含む完全情報ゲーム $\Gamma = \langle N, \mathcal{K}, P, u \rangle$ における戦略プロファイル s^* が部分ゲーム完全均衡であるとは、$\forall h \in H$ の部分ゲーム $\Gamma(h)$ において、s^* がナッシュ均衡となっていることである。

Outline

1. 混合戦略 Mixed Strategy とナッシュ均衡の存在定理
2. 同時決定を含む拡張と部分ゲーム完全均衡 SPNE
3. 繰り返しゲームとフォーク定理

繰り返しゲーム

Definition (繰り返しゲーム)

- 同じゲーム(これを成分ゲーム constituent game (stage game) という) $G = \langle N, A, u \rangle$ のプレーを繰り返す。
- 最終的に得られる利得は、各期に得られる利得の総和を割り引いたもの(割引因子 discount factor $\delta_i \in [0, 1)$ で割り引くバージョンもある)

$$\sum_{t=1}^{T} u_i(a^t)/T \qquad (\sum_{t=1}^{T} \delta_i^t u_i(a^t))$$

- 繰り返しゲームでは、履歴は $h = (a^1, ..., a^t) \in A^t \in H$ と表すことができる。

繰り返しゲームとヒステリシス Hysterisis

- 繰り返しゲームでは、履歴は過去のプレーのみによって特徴付けられる。つまり、同じ t 期であれば、その期に属する任意の履歴 $h \in A^t$ でプレーするゲームの物理的環境は不変。
- ゲーム理論と標準的な自然科学理論との大きな差の一つは、前者では一般に履歴に大きく依存せずに状態を記述することが難しいことである。

達成可能性

Definition (達成可能 Feasible な利得プロファイル)

$G = \langle N, A, u \rangle$ 上で達成可能な利得プロファイル $v \in \Re^N$ の集合は、A 上の結果で得られる利得プロファイルの凸結合、すなわち

$$\{\sum_{a \in A} \alpha_a u(a) | (\forall a \in A, \alpha_a \geq 0) \wedge \sum_{a \in A} \alpha_a = 1\}$$

公的混合 Public Randomizing

この集合上の任意の利得プロファイルは、適当な回数の割合で A 上の結果を混ぜてプレーすることで合意することにより、無限に近く近似的に達成可能。

個人合理性

Definition (ミニマックス Minmax 利得)

$$\underline{v}_i = \min_{a_{-i} \in A_{-i}} \max_{a_i \in A_i} u_i(a)$$

プレーヤーi に対する最も厳しい制裁から得られる i の利得。

Definition (個人合理性 Individual Rationality)

$u \in \Re^N$

- $u \geq \underline{v}$ 個人合理的 individually rational
- $u \gg \underline{v}$ 狭義個人合理的 strictly individually rational

達成可能かつ個人合理的な利得プロファイルの集合を V と表すことがある。

トリガー戦略としっぺ返し

Definition (トリガー戦略)

$v \in V$ をプレーヤー間で達成を目指す利得の合意内容とする。以下の戦略をトリガー戦略と呼ぶ。

- 他のプレーヤーが合意内容に従ってプレーしている間は、自分も合意内容に従う
- 他のプレーヤー誰かひとり j が逸脱した場合、j に対して minmax の制裁を課すプレーをとり続ける。

似たようなプレーの仕方に「しっぺ返し tit for tat(目には目を、歯には歯を)」がある。これは、相手の戦略を一期遅れで模倣する戦略。

フォーク定理 Folk Theorem

Theorem (完全フォーク定理 [5])

無限回繰り返しゲームでは、個人合理的かつ達成可能な任意の利得プロファイルがあるナッシュ均衡で達成可能。

Proof.

トリガー戦略の組はナッシュ均衡。 □

無限回繰り返しゲームの SPNE でも同様の性質があるが、制裁の設計がトリガー戦略に比べて格段に難しくなる。

合意内容の設計とコミュニケーション

無限回繰り返しゲームにおける合意内容の設計・調整にコミュニケーションが大変有効。

実例

Example (囚人のジレンマ)

	C	D
C	$3,3$	$1,6$
D	$6,1$	$2,2$

(C, D) と (D, C) を適当な割合でプレーすることを繰り返すのが良さそう。
(C, C) は一回交代で交互に (C, D) と (D, C) をプレーする繰り返しにパレート支配される。

Example (チキン)

	ハト	タカ
ハト	$3,3$	$1,4$
タカ	$4,1$	$0,0$

今度は、例えば (ハト, ハト) を繰り返すのが良い。
しかし、(タカ, ハト) を無限回繰り返すのも依然として SPNE。
(例　俺様夫と従順な妻)

有限回繰り返しゲームと部分ゲーム完全均衡

Theorem (有限回繰り返しゲームと部分ゲーム完全均衡)

有限回繰り返しゲームでは、成分ゲームのナッシュ均衡が一つしか存在しない場合、その均衡をすべての期にプレーする戦略プロファイルのみが部分ゲーム完全均衡となる。

成分ゲームに均衡が二つ以上存在する場合は、パレート支配される均衡のプレーを脅しに用いることにより、最終期のみ裏切るような、フォーク定理と似たようなプレーが均衡となり得る。[4]

望ましい小さな非合理性としての ϵ-均衡

Definition (ϵ-均衡)

$\forall \epsilon > 0$ に対して、戦略プロファイル s^* が ϵ-均衡であるとは、$\forall s_i \in S_i$,

$$u_i(s_i, s_i^*) - u_i(s^*) \leq \epsilon$$

- 割引因子が十分に大きなプレーヤー同士が十分長期にわたってプレーする場合、ϵ を限りなくゼロに近くしても、フォーク定理と似たような結果が得られる。[3, 2]
- 似たような分析として、最終期 T に関する微小な共有知識の破れを仮定しただけでも、フォーク定理と似た結果が得られることが知られている。[7]

相関均衡点 Correlated Equilibrium

Definition (相関均衡点(ややこしいので形式的定義は省略) [1])

ナッシュ均衡のような安定性を有するような、不確実性ノードの状態によってプレーヤーがプレーを変化させるような行動の決め方の組

Example

晴れたら外でデート、雨が降ったら室内でデート。

- 公的混合の考え方を一回ゲームに適用したのが、相関均衡点の考え方。
- 混合戦略の一般化と考えることができる。

[1] Robert J. Aumann.
Correlated equilibrium as an expression of bayesian rationality.
Econometrica, 55(1):1--18, 1987.

[2] Jean-Pierre Benoit and Vijay Krishna.
Finitely repeated games.
Econometrica, 53(4):905--922, 1985.

[3] Jean-Pierre Benoit and Vijay Krishna.
Nash equilibria of finitely repeated games.
International Journal of Game Theory, 16(3):197--204, 1987.
10.1007/BF01756291.

[4] James W. Friedman.
Cooperative equilibria in finite horizon noncooperative supergames.
Journal of Economic Theory, 35(2):390 -- 398, 1985.

[5] Drew Fudenberg and Eric Maskin.
The folk theorem in repeated games with discounting or with incomplete information.
Econometrica, 54(3):533--554, 1986.

[6] John Nash.
Equilibrium points in n-person games.
In Proceedings of the National Academy of Sciences, volume 36, pages 48--49, 1950.

[7] Abraham Neyman.
Cooperation in repeated games when the number of stages is not commonly known.
Econometrica, 67(1):45--64, 1999.

[8] Martin J. Osborne and Ariel Rubinstein.
A Course in Game Theory.
MIT Press, Cambridge, 1994.