跡見学園女子大学マネジメント学部紀要　第4号　（2006年3月15日）

繰り返しゲームにおける均衡特定化と実験の展望

Survey on Equilibrium Characterization and Experimentation of Repeated Games

木　村　友　二*，丹　野　忠　晋**

1　はじめに

長期的な相互依存関係にある経済主体間の協調はどのように生成されるのか？この問いに対してゲーム理論的なアプローチは長年の繰り返しゲームの研究で答えてきた．完全情報の下では将来の利得を評価する際に用いられる割引因子が十分1に近いのならば協力が達成される可能性があるというのが第1の答えである．現実にはプレイヤー間の行動を必ずしも完全に観察することができないのは自明であろう．そのような場合には第1の条件に加えてプレイヤー間の行動がある種の情報を通じて統計的に識別可能であるという新たな条件が付加される．このような長期的な協力の可能性を追求するとともに将来の利得を現在の利得と同様に評価していった究極の場合に効率的な利得がどの程度まで均衡として実現できるかを近年のゲーム理論は研究している．本論の第1の目的はこのような場合の均衡を特定するアルゴリズムを展望することにある．第2にどのような場合でも協調を達成するには割引因子が十分に大きくて無限回の繰り返しが要求されている．このような割引因子の大きさと繰り返し回数の多さと協調の程度は実験的にどのように支持されるかを概観するのが目的である．

主に Fudenberg and Tirole (1992) に掲載されている繰り返しゲームの基本的な命題を原論文やその後の発展を取り入れることによって見通しの良い証明にすることを目指している．具体的には continuous at infinity や unimprovability などの性質や詳しい証明を付ける．ダイナミック・プログラミングの繰り返しゲームへの応用と発展に焦点を絞るが，上の2の補題にはあまり詳しい証明は教科書レベルではない．そして，例や基本的な均衡の定義から始まって Fudenberg and Levine (1994) の均衡利得特定化定理を証明することが目標である．次

に Fudenberg, Levine and Maskin (1994) のフォーク定理の紹介と最近の成果である Tanno (2000) の結論を紹介する．さらに，繰り返しゲームの応用は産業組織理論や企業理論が主であったが，Ljungquist and Sargent (2000) に見られるように最近はマクロ経済学にもその応用範囲が広がっている．このような分野の論文を読むための必須な知識を手短にまとめよう．理論的に予測可能な行動はどのように実験で確認できるのかを Dal Bó (2004) 論文を中心に見ていく．

2　ステージゲームの例―共同生産ゲーム

基本となるゲームは，不完全公共情報の無限回繰り返しゲームである．ここでは不完全モニタリングの代表例である企業のモニタリング問題を取り上げよう．2 人のプレイヤーがいる．すなわちプレイヤー集合は $N=\{1,2\}$ とする．各々は働く (Work) かあるいは怠ける (Shirk) の選択肢を持っていて，行動集合を $A_i=\{w,s\}$ としよう．プレイヤーが観察できる生産高は，高い生産 $(y=12)$ か低い生産 $(y=0)$ である．公共シグナルの集合を $Y=\{0,12\}$ とする．ここで不確実な生産の実現確率は以下のように表されるとする．

$$\pi(12|w,w)=2/3, \pi(12|w,s)=\pi(12|s,w)=1/3, \pi(12|w,w)=0$$

生産に成功すれば半分の産出量を分かち合い労働の不効用が発生するとしよう．プレイヤーの利得は次のようになる．

$$r_i(a_i,y):=\begin{cases}\frac{y}{2}-3 & \text{if } a_i=w\\ \frac{y}{2} & \text{if } a_i=s\end{cases}$$

生産の実現確率を用いて期待利得はアクションにのみ依存することになる．例えば，両者が働いた場合 (w,w) は次のような利得を得ることになる．

$$g_i(w,w)=(12/2-3)\cdot 2/3+(0/2-3)\cdot 1/3=1$$

そうして表 1 のように利得表を書くことができる．明らかにこのゲームは囚人のジレンマゲームと同じ構造であることが見て取れる．その実現可能な利得ベクトルの集合を表すと図 1 のようになる．このようなゲームを無限回繰り返したとき，割引因子が十分に 1 に近いときの均衡利得を求めることが我々の 1 つの目標である．

表 1　共同生産ゲームの利得表

	w	s
w	1,　1	−1,　2
s	2,　−1	0,　0

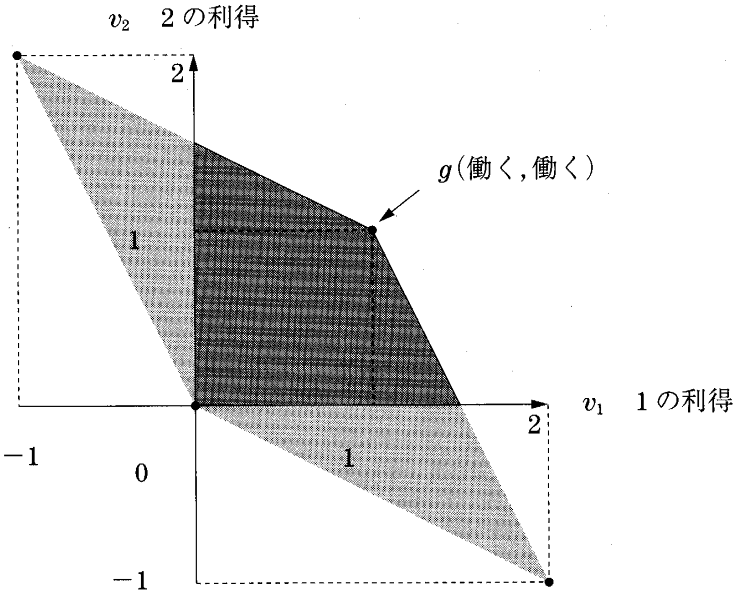


図1　共同生産ゲームの実現可能利得集合

3．ステージゲーム

一応，後半の Tanno (2000) の定理の解説のために短期 Short-run (SR) プレイヤーを含めて定式化するが，後半部分まで本質的には Short-run プレイヤーがいなくても結果は変わらない．

- プレイヤー集合:
  $N := \{1, 2, \ldots, n\}\ (n \geq 2), \quad SR := \{n+1, n+2, \ldots, n+l\}, \quad (l \geq 0)$
- プレイヤー集合： $N' = N \cup SR$
- プレイヤー i の純行動: $a_i \in A_i$, $\#A_i = m_i < \infty$
- 純粋プロファイル: $a \in A := \times_{i \in N} A_i$
- Long-run (LR) player の純粋プロファイル: A_{LR}
- i の混合行動: $\alpha_i \in \mathcal{A}_i$
- プロファイル: $\alpha \in \mathcal{A} := \times_{i \in N} \mathcal{A}_i$
- LR と SR のプロファイル全体: $\mathcal{A}_{LR}$, $\mathcal{A}_{SR}$
- i を除くプロファイル: $a_{-i} \in A_{-i}$, $\alpha_{-i} \in \mathcal{A}_{-i}$
- public outcome: $y \in Y$, $\#Y := m < \infty$
- i の利得関数: $r_i : A_i \times Y \to \mathbb{R}$
- a の下での y の確率: $\pi(y|a)$, $\pi(y|a_i, a_{-i})$
- a の下での i の期待利得: $g_i(a) := \sum_{y \in Y} \pi(y \mid a) r_i(a_i, y)$
- α の下での i の期待利得:

$g_i(\alpha) := \sum_{y \in Y} \sum_{a \in A} \prod_{i \in N'} \alpha_i(a_i)\pi(y \mid a) r_i(a_i, y)$

- a の下での利得ベクトル(1):　$\boldsymbol{g}(a) = (g_1(a), \ldots, g_n(a))^{\mathrm{T}}$
- $i \in N$ のミニマックス値:

$\underline{v}_i := \min_{\alpha_{-i}} \max_{a_i} g_i(a_i, \alpha_{-i})$.

- ミニマックス利得ベクトル:

$\underline{\boldsymbol{v}} := (\underline{v}_1, \ldots, \underline{v}_n)^{\mathrm{T}}$

- 実行可能利得ベクトル集合(2):

$V := \mathrm{co}\, \boldsymbol{g}(A)$

- 実行可能かつ個人合理的利得集合:

$V^* := V \cap \{\boldsymbol{v} \in \mathbb{R}^n \mid \boldsymbol{v} \geq \underline{\boldsymbol{v}}\}$.

- Short-run player の混合最適対応:

$B : \mathcal{A}_{LR} \to \mathcal{A}_{SR}$

- Short-run player の純粋最適対応:

$\hat{B} : A_{LR} \to \mathcal{A}_{SR}$,　graph $\hat{B}$ を $\hat{B}$ のグラフとする.

- 実行可能，かつ誘因整合的利得集合:

$V' := \mathrm{co}\, \boldsymbol{g}(\mathrm{graph}\, \hat{B})$

- 実行可能，個人合理的，かつ (short-run player) 誘因整合的利得集合:

$V^{**} := V' \cap \{\boldsymbol{v} \in \mathbb{R}^n \mid \boldsymbol{v} \geq \underline{\boldsymbol{v}}\}$

4　繰り返しゲーム

繰り返しゲームにおいて LR プレイヤーと SR プレイヤーの間の関係は図 2 に表されている．LR は無限に生きる個人である．それには過去の自分の取ってきた履歴や公共シグナルの履歴を持っている．第 t 期に入ると SR プレイヤーが生まれてくる．すべてのプレイヤーが自分の行動を取った後にそれに依存する形で公共シグナルが実現する．それと自分の取った行動から自分の利得が実現する．利得を受け取ると SR プレイヤーは死滅して新しい SR が取って代わる．

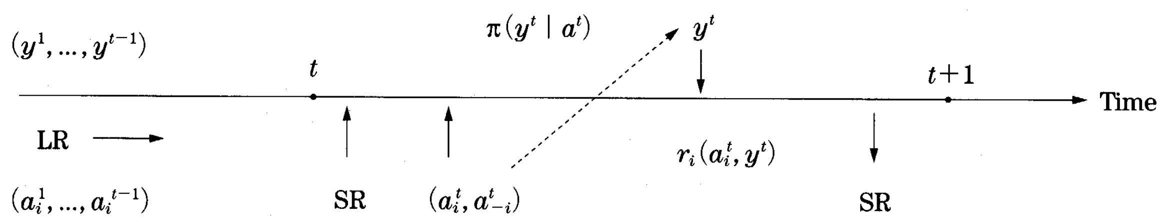


図 2　シグナルと利得の実現

4．1　繰り返しゲームの利得

各プレイヤーは共通の割引因子 $\delta \in (0,1)$ を有している．ある繰り返しゲームの戦略プロファイル $(\alpha^t)_{t=1}^{\infty}$ の下での i の繰り返しゲームの平均割引期待利得は次のように表現できる．

$$(1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}g_i(\alpha^t)$$

このように $(1-\delta)$ を掛けることで利得を平均化してステージゲームの利得と比較が可能になる．なぜならば， $S := \sum_{t=1}^{\infty}\delta^{t-1}g_i(\alpha^t)$ と置くと，平均化された利得 s は

$$S = s + \delta s + \delta\ 2s + \cdots = \frac{s}{1-\delta}$$

となるからである．プレイヤーが長期的な利益を重視した究極的な状態である δ が 1 に収束した場合がとても興味がある．

定理 4.1 (Abel's Theorem) 任意の点列 $(a_t)_{t=1}^{\infty}$ に対して，

$$\lim_{\delta\to 1}(1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}a_t = \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}a_t.$$

ただし, どちらかの極限が存在するとする．　□

このアーベルの定理の証明は Hardy, G. H. (1949, Theorem 55,p.108) にある．右辺は必ずしも収束しない．さらに Myerson (1991.p.315) が指摘しているように $(a_t)_{t=1}^{\infty}$ が有界でも収束は確保できない．右辺が収束しないときは,

$$\liminf_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}a_t$$

がその代わりを務める．Radner (1985) が用いているように下極限はうまく定義できて $(a_t)_{t=1}^{\infty}$ が有界ならばそれは有限の値を取る．

4．2　履歴

繰り返しゲームはどのステージも同じ構造を持っている．そして，あるステージ以降のゲーム全体はまた無限回繰り返しゲームになる．では時点の違うステージゲームは何が違うのか？それは今までにプレイされてきた履歴すなわちヒストリーが違う．ここで記法を用意しよう．

- 初期 public history:　 $h(0)$,　 $H^0 := \{h(0)\}$
- 時点 t の public history 全体:
  $\mathcal{H}(t) := \{(h(t-1), y^t) \mid h(t-1) \in \mathcal{H}(t-1) \text{ and } y^t \in Y\}$
- 時点 t のある public history :

$h^t = (h(0), y^1, \ldots, y^t) \in \mathcal{H}(t)$

- プレイヤー i の初期 private history:　$h_i(0)$,　$H_i(0) := \{h_i(0)\}$
- 時点 t のプレイヤー i の private history 全体:
  $\mathcal{H}_i(t) := \{(h_i(t-1), \alpha_i^t) \mid h_i(t-1) \in \mathcal{H}_i(t-1) \text{ and } \alpha_i^t \in \mathcal{A}_i\}$
- 時点 t のプレイヤー i のある private history :
  $h_i(t) = (h(0), \alpha_i^1, \ldots, \alpha_i^t) \in \mathcal{H}_i(t)$
- 時点 t の private history 全体:
  $\mathcal{H}_p(t) := \times_{i \in N'} \mathcal{H}_i(t)$
- 時点 t の private history プロファイル:
  $h_p(t) := (h_i(t))_{i \in N'} \in \mathcal{H}_p(t)$

4．3　戦略

繰り返しゲームの戦略とは各期においてその期のどんな履歴に対して自分の行動を定めた行動計画を意味している．履歴とステージゲームの行動から戦略は以下のように定義できる．

- 時点 t の $i \in N$ の戦略:
  $\sigma_i^t : \mathcal{H}(t-1) \times \mathcal{H}_i(t-1) \to \mathcal{A}_i$
- 時点 t の $j \in SR$ の戦略:
  $\sigma_j^t : \mathcal{H}(t-1) \to \mathcal{A}_j$
- i の戦略, 時点 t の戦略プロファイル, 戦略プロファイル:
  $\sigma_i := (\sigma_i^t)_{t=1}^{\infty}$,　その全体 $\mathcal{S}_i$
  $\sigma^t := (\sigma_i^t)_{i \in N'}$,　その全体 $\mathcal{S}^t$
  $\sigma := (\sigma_i)_{i \in N'}$,　その全体 $\mathcal{S}$

このゲームで重要なのは公共戦略と呼ばれる戦略のクラスである．

定義 4.1 時点 t の i の **public 戦略**を次のように定義する．

$$\forall\ h_i(t-1) \in \mathcal{H}_i(t-1)\ \forall\ h_i(t-1) \in \mathcal{H}_i(t-1)$$
$$\sigma_i^t(h(t-1), h_i(t-1)) = \sigma_i^t(h(t-1), h_i'(t-1)) \qquad \square$$

つまりこの期のアクションは public ヒストリーのみによって定まる戦略が public 戦略である．ここで，σ_i^t が public 戦略のとき簡単に $\sigma_i^t : \mathcal{H}(t-1) \to \mathcal{A}_i$ と書くことにしよう．さらに正しくは

$$\sigma_i^t(h(t-1)) = \alpha^t \quad および \quad \sigma_i^t(h(t-1))(a) = \alpha^t(a)$$

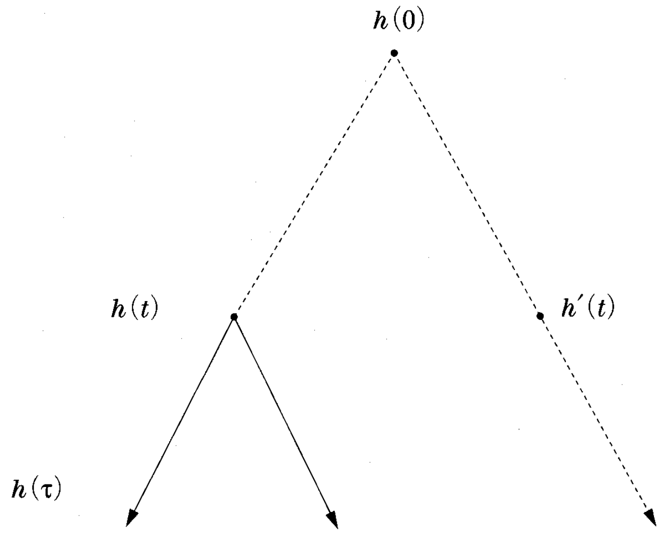


図 3　履歴

であるが記法を省略して $\sigma_i^t(a|h(t-1)) = \alpha^t(a)$ のように書く．

この繰り返しゲームを G としよう．図 3 のようにすべてのプレイヤーが public 戦略を取ったならば各 public シグナルが実現した後のゲームをサブゲームと見なすことができる．そのようなサブゲームを $G(h(t))$ を以下に定義する．ゲーム全体は $G(h(0))$ と同一視できる．

4．4　履歴上の確率

各プレイヤーが純粋戦略をとっていたとしても実現される public シグナルは確率的に変動する．ヒストリー上の確率測度は以下のように帰納的に定義できる．最初に

$$\mu^0(h(0)|\sigma) := 1 \qquad (t=0)$$

および $t \geq 1$ 期で $\mu^{t-1}(h(t-1)|\sigma)$ が定義できているとして，

$$\mu^t(h(t)|\sigma) := \mu^{t-1}(h(t-1)|\sigma) \cdot \sum_{a \in A} \prod_{i \in N'} \sigma_i^t(a_i|h(t-1))\pi(y^t|a)$$

と再帰的に定義をする．ただし，ここで $h(t) = (h(t-1), y^t)$ である．つまり，戦略 σ によってヒストリー $\mathcal{H}(t)$ 上の確率 $\mu^t(\cdot|\sigma)$ が誘導される．もちろん，すべての時期に対して $(\mathcal{H}(t), \mu^t(\cdot|\sigma))_{t=1}^{\infty}$ が定義される．

4．5　例の再考

ここで前の例のヒストリーを考えて見よう．次のような戦略を考えよう．

s_1 : いつも s (サボリ) を選ぶ．

s_2 : 最初は w (真面目) で，直前の生産が 0 であれば w を選ぶ.

そうでなければ s を選ぶ.

この戦略でのある public history と private history は以下のようになる.

$$h(3) = (h(0), 0, 12, 0) \in H(3)$$
$$h_1(3) = (h(0), s, s, s) \in H_1(3)$$
$$h_2(3) = (h(0), w, w, s) \in H_2(3)$$

この時の $h(3)$ が実現する確率 $\mu^3(h(3)|s_1, s_2)$ を求めてみよう.

$$\pi(y^1 = 0|s, w) = 2/3, \pi(y^2 = 12|s, w) = 1/3, \pi(y^3 = 0|s, s) = 1$$

だから $\mu^3(h(3)|s_1, s_2) = 2/3 \cdot 1/3 \cdot 1 = 2/9$ となることが感得できる.

4．6　ヒストリー上の確率に関する期待値と利得

ステージゲームでの混合戦略を許した戦略 σ を用いたとき t 期の利得の期待値は以下のようになる.

$$\mathbb{E}\left[g_i(\sigma^t(h(t-1)))|\sigma\right] := \sum_{h(t-1)\in H(t-1)} g_i(\sigma^t(h(t-1))) \cdot \mu^{t-1}(h(t-1)|\sigma)$$

ただし $\mathbb{E}[\,\cdot\,|\sigma]$ は σ から誘導される public history 上の確率測度に関する期待値とする．戦略 σ が明らかなときは省略して $\mathbb{E}[\,\cdot\,]$ と書く．共通の割引因子 $\delta \in (0,1)$ によって Public 戦略プロファイル σ での i の利得は

$$u_i(\sigma;\delta) := (1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}\mathbb{E}\left[g_i(\sigma^t(h(t-1)))\right]$$

と定義される.

4．7　サブゲームの戦略と利得

次にサブゲームの戦略と利得を定義しよう．時点 $\tau > t$ に対してヒストリー $h(\tau)$ は $h(t)$ から到達可能でるとは,

ある列 $(y^{t+1}, \ldots, y^{\tau})$ に対して $h(\tau) = (h(t), y^{t+1}, \ldots, y^{\tau})$

が満たされることである．Public ヒストリー $h(t)$ から到達可能な $\tau\ (>t)$ 期のヒストリー全体を $H_{h(t)}(\tau)$ としよう．戦略 σ の t 期の public ヒストリー $h(t)$ における制限 $\sigma|_{h(t)}$ とは以下の条件満たす戦略プロファイルである．

$$\sigma^\tau|_{h(t)} : H_{h(t)}(\tau) \to \mathcal{A}, \quad \sigma|_{h(t)} = (\sigma^\tau|_{h(t)})_{\tau=t+1}^{\infty}$$
$$\tau = t+1, \quad \sigma^{t+1}|_{h(t)}(h(t)) = \sigma^{t+1}(h(t))$$
$$\tau \geq t+2, \quad \sigma^\tau|_{h(t)}(h(\tau-1)) = \sigma^\tau(h(\tau-1)), \quad h(\tau-1) \in H_{h(t)}(\tau-1)$$

そうすると $\sigma|_{h(t)}$ はサブゲーム $G(h(t))$ の戦略である．ヒストリー $h(t)$ が initial history となる．ここで $\sigma|_{h(t)}$ の全体を $\mathcal{S}(h(t))$ とする．

制限戦略 $\sigma|_{h(t)}$ は $h(t)$ から到達可能な public history 上の確率を定める．

$$\tau = t, \qquad \mu^t(h(t)|\sigma|_{h(t)}) := 1$$
$$\tau \geq t+1, \qquad \mu^\tau(h(\tau)|\sigma|_{h(t)}) :=$$
$$\mu^{\tau-1}(h(\tau-1)|\sigma|_{h(t)}) \cdot \sum_{a\in A}\prod_{i\in N'} \sigma_i^\tau|_{h(t)}(a_i|h(\tau-1))\pi(y^\tau|a).$$

ただし，$h(\tau-1) \in H_{h(t)}(\tau-1)$ かつ $h(\tau) = (h(\tau-1), y^\tau)$ である．

サブゲーム $G(h(t))$ で戦略 $\sigma|_{h(t)}$ を用いたとき $\tau\ (>t)$ 期の利得の期待値は

$$\mathbb{E}\left[g_i(\sigma^\tau|_{h(t)}(h(\tau-1)))\right] := \sum_{h(\tau-1)\in H_{h(t)}(\tau-1)} g_i(\sigma^\tau|_{h(t)}(h(\tau-1))) \cdot \mu^{\tau-1}(h(\tau-1)|\sigma|_{h(t)})$$

となる．繰り返しゲーム G と戦略 σ に対してサブゲーム $G(h(t))$ での制限された戦略 $\sigma|_{h(t)}$ での t 期を基準にした利得は以下のようになる．

$$u_i(\sigma|_{h(t)};\delta) := (1-\delta)\sum_{\tau=1}^{\infty}\delta^{\tau-1}\mathbb{E}\left[g_i(\sigma^{t+\tau}|_{h(t)}(h(t+\tau-1)))\right]$$

4．8　完全公共均衡

次に均衡概念を定義しよう．我々が用いる均衡概念は完全公共均衡 (Perfect Public Equilibrium) と呼ばれる．

定義 4.2 (Perfect Public Equilibrium) 戦略 $\hat{\sigma} \in \mathcal{S}$ が **Perfect public** 均衡 (PPE) であるとは以下の条件を満たすことである．戦略 $\hat{\sigma}$ は public 戦略であり，どんな時点 t 期のどんなサブゲーム $G(h(t))$ に対しても $\hat{\sigma}$ から導かれるヒストリー $h(t)$ で制限された戦略 $\hat{\sigma}|_{h(t)}$ は Nash 均衡を形成する． □

記号で表現すれば,

$$\forall\, t \in \mathbb{N}\ \forall\, h(t) \in H(t)\ \forall\, i \in N'\ \forall\, \sigma_i|_{h(t)} \in \mathcal{S}_i(h(t))$$
$$u_i(\hat{\sigma}|_{h(t)};\delta) \geq u_i(\sigma_i|_{h(t)}, \hat{\sigma}_{-i}|_{h(t)};\delta)$$

となる．割引因子 δ の下での PPE 戦略プロファイル全体を $\mathcal{S}^*(\delta)$ としよう．そうすると $\mathcal{S}^*(\delta) \neq \emptyset$ が分かる．Perfect information ケースでは，$\mathcal{S}$ と $\mathcal{S}^*(\delta)$ は適当な位相を入れるとコンパクトになることが良く知られている(3)．位相は様々あるが Börges (1989) の方法が一番自然である．一般ケースでもコンパクトになるだろう(4)．

一般には戦略集合の性質や利得関数の性質を吟味する必要があるが，多数の均衡が出現するのが普通でありできるだけ効率的な利得ベクトルを見出すのが主要なテーマである繰り返しゲーム理論では均衡利得に注目する．割引因子 δ の下での均衡利得ベクトル集合を

$$E(\delta) := \left\{ \boldsymbol{v} \in \mathbb{R}^n \mid \boldsymbol{v} = (u_1(\sigma^*;\delta), \ldots, u_n(\sigma^*;\delta))^{\mathrm{T}} \text{ and } \sigma^* \in \mathcal{S}^*(\delta) \right\}.$$

と定義する．この $E(\delta)$ はコンパクトである(5)．利得集合の割引因子に関する**単調性**が成立する(6)．

$$E(\delta) \subset E(\delta') \quad \text{for } \delta \leq \delta' < 1.$$

この単調性を用いて極限均衡利得集合を次のように定義しよう．

$$\lim_{\delta\to 1} E(\delta) := \bigcup_{\delta\in[0,1)} E(\delta).$$

この単調性により,

$$\boldsymbol{v} \in \lim_{\delta\to 1} E(\delta) \iff \exists\, \underline{\delta} < 1\ \forall\, \delta \in [\underline{\delta}, 1)\ \boldsymbol{v} \in E(\delta).$$

これで利得の収束性を云々する必要はなくなった．

4．9　色々な性質

他の均衡に関する性質をあげよう．まず最初に既に述べたようにあらゆる割引因子に対して $E(\delta) \neq \emptyset$ が成り立つ．ある均衡戦略 $\sigma^* \in \mathcal{S}^*(\delta)$ から i のみが public ではない戦略に逸脱したとしても σ^*_{-i} が public 戦略なので，逸脱による利益はない．PPE では各情報集合上の信念は均衡になんの役割も果たさない．つまり，どんな情報集合上のノードの確率も同じ利得上の確率を生み出すからである．Public 戦略を用いている限り public シグナル上の確率が変動する場合にのみに着目すればよい．また，PPE $\Longrightarrow$ Sequential Equilibrium (SE) および PPE $\Longrightarrow$

Perfect Bayesian Equilibrium (PBE) が成り立つ．PPE と SE が等しくなる条件は FL94 の Theorem 5.2 にある．$E(\delta)$ は定常である．つまり，どのようなサブゲーム $G(h(t))$ でも均衡利得集合は一定である．PPE は再帰構造を持つ．どんな PPE のステージゲームアクションもそれ以降の continuation game である PPE を生み出す．しかし，PBE はそうではない．以上のような理由で真部分ゲームはないにもかかわらずサブゲーム $G(h(t))$ を定義することができるのである．

5 One-Stage-Deviation-Principle

この原理はダイナミック・プログラミングにおいて Howard (1960) が初めて導出した．この原理は繰り返しゲームにおいても大変有用な原理である．

定義 5.1 (One-Stage-Deviation Principle) 戦略 $\hat{\sigma}$ が *One-Stage-Deviation Principle* あるいは *Unimprovability* を満たすとは，任意の時点 t, ヒストリー $h(t)$ に対して，次の条件を満たすプレイヤー i と戦略 $\sigma_i|_{h(t)}$ が存在しないことである．ヒストリー $h(t)$ が到達した下で，

$$\left.\begin{array}{l}\sigma_i^\tau|_{h(t)} \neq \hat{\sigma}^\tau|_{h(t)} \text{ for } \tau = t+1 \\ \sigma_i^\tau|_{h(t)} = \hat{\sigma}^\tau|_{h(t)} \text{ for } \tau \geq t+2\end{array}\right\} \implies u_i(\sigma_i|_{h(t)}, \hat{\sigma}_{-i}|_{h(t)}; \delta) > u_i(\hat{\sigma}|_{h(t)}; \delta). \quad \square$$

この定義は一回だけ逸脱して元に戻っても利益を得ることはないという意味であるが，定義の中の否定を中に入れてその同値な表現を考えてみると分かりやすい．

補題 5.1 (OSDP と同値) $\hat{\sigma}$ が *One-Stage-Deviation Principle* を満たすことの必要十分条件は，

$\forall\, t \in \mathbb{N} \;\; \forall\, h(t) \in H(t) \;\; \forall\, i \in N' \;\; \forall\, \sigma_i|_{h(t)} \in \mathcal{S}_i(h(t))$

$$\left.\begin{array}{l}\sigma_i^\tau|_{h(t)} \neq \hat{\sigma}^\tau|_{h(t)} \text{ for } \tau = t+1 \\ \sigma_i^\tau|_{h(t)} = \hat{\sigma}^\tau|_{h(t)} \text{ for } \tau \geq t+2\end{array}\right\} \implies u_i(\hat{\sigma}|_{h(t)}) \geq u_i(\sigma_i|_{h(t)}, \hat{\sigma}_{-i}|_{h(t)}).$$

であることである． $\square$

次の定理は dynamic programming の最適性原理あるいはベルマン原理と同値である．

定理 5.1 (Howard (1960)) σ^* がサブゲーム完全均衡である必要十分条件は σ^* が *OSDP* を満たす事である． $\square$

5．1　定理 5.1 の証明

[SPN $\Longrightarrow$ OSDP] は明らか．これから十分性 [OSDP $\Longrightarrow$ SPN] を示そう．

補題 5.2 $\hat{\sigma}$ が *OSDP* ならば有限回の逸脱の列によって *improve* できない． □

補題 5.2 の証明:

一般性を失うことなく，純粋戦略に絞ろう．背理法を用いて，$\hat{s}$ は OSDP だが有限回の逸脱により improve されるとする．ある期 t のサブゲーム $G(h(t))$ が存在して，あるプレイヤー i のある戦略 s_i はサブゲーム $G(h(t))$ で $\hat{s}_{-i}$ に対して $\hat{s}_i$ よりもより良い反応である．

$$u_i(s_i|_{h(t)}, \hat{s}_{-i}|_{h(t)}) > u_i(\hat{s}|_{h(t)}) \tag{1}$$

一般性を失うことなく $t+1$ 期で最初の deviation が起こるとする．

$$s_i^{t+1}|_{h(t)} \neq \hat{s}^{t+1}|_{h(t)}$$

T 期を逸脱が起こる最大の時期とする．

$$s_i^T|_{h(t)} \neq \hat{s}_i^T|_{h(t)} \quad s_i^\tau|_{h(t)} = \hat{s}_i^\tau|_{h(t)} \text{ for } \tau \geq T+1.$$

有限回の逸脱なので $T < \infty$.

まず, $t+1 = T$ ならば，最初で最後の逸脱，つまり一回しか逸脱はない．OSDP より (1) は矛盾する．

次に，逸脱が2回起こったとする．この時，$t+2 \leq T$ である．ここで新たに次の戦略 $\tilde{s}_i$ を考えよう：T 期以前は s_i と一致し，T 期を含めた T 期以降は $\hat{s}_i$ と一致する．

$$\tilde{s}_i^\tau|_{h(t)} := s_i^\tau|_{h(t)} \text{ for } \tau \leq T-1, \quad \tilde{s}_i^\tau|_{h(t)} := \hat{s}_i^\tau|_{h(t)} \text{ for } \tau \geq T.$$

T は逸脱の最後の時点であるので，戦略 s_i は $T+1$ 期以降は $\hat{s}_i$ と一致していることから，OSDP から s_i は $h(t)$ から到達可能な任意のサブゲーム $G(h(T-1))$ で $\hat{s}_{-i}$ に対して高々 $\tilde{s}_i$ と同じくらい良い反応である．

$$u_i(s_i|_{h(T-1)}, \hat{s}|_{h(T-1)}) \leq u_i(\tilde{s}_i|_{h(T-1)}, \hat{s}_{-i}|_{h(T-1)}).$$

戦略 $\tilde{s}_i$ は T 期以前は s_i と一致しているので，s_i はサブゲーム $G(h(t))$ で $\hat{s}_{-i}$ に対して高々 $\tilde{s}_i$ と同じくらい良い反応である．

$$u_i(s_i|_{h(t)}, \hat{s}_{-i}|_{h(t)}) \leq u_i(\tilde{s}_i|_{h(t)}, \hat{s}_{-i}|_{h(t)}). \tag{2}$$

この時 $\tilde{s}_i$ は t+1 期のみ $\hat{s}_i$ と異なっているので，OSDP より

$$u_i(\tilde{s}_i|_{h(t)}, \hat{s}_{-i}|_{h(t)}) \leq u_i(\hat{s}|_{h(t)}). \tag{3}$$

(2) より (1) に矛盾する．これで 2 回逸脱が起こっても improve できないことが分かった．ここで，

$$k \text{ 回の逸脱で improve できない．} \tag{4}$$

と仮定する．$t+1$ から T まで $k+1$ 回の逸脱があるとする．そのとき，$T \geq t+k+1$ で $t+1$ から $T-1$ まで k 回の逸脱がある．同様に $\tilde{s}_i$ を構成すると，(2) が成り立つ．一方，$\tilde{s}_i$ は k 回の逸脱しているので，(4) より (3) が成り立つ．(2) より (1) と矛盾する． ■

6 Continuous at Infinity

遠いかなたのステージゲームの利得が全体のゲームに対する影響はとても小さくなければならない．そうでなければ利得列の収束性が満たされない．このことはダイナミックな経済モデルでも同じ事が言える[7]．OSDP が一般の逸脱で成立するためには *continuous at infinity* という条件が必要である．

定義 6.1 (Continuous at Infinity) ゲーム G は *continuous at infinity(CI)* を満たすとは次の条件が成立することである．

$$\forall\, \varepsilon > 0\; \exists\, t \in \mathbb{N} \;\; \forall\, s \in S\; \forall\, \hat{s} \in S \;\; \forall\, i \in N$$
$$(s^{\tau})_{\tau=1}^{t} = (\hat{s}^{\tau})_{\tau=1}^{t} \implies |u_i(s) - u_i(\hat{s})| < \varepsilon$$ □

補題 6.1 繰り返しゲームは *continuous at infinity* をもつ． □

補題 6.1 の証明:
実際，少し省略して CI は以下の条件と等しい．

$$\forall\, \varepsilon > 0\; \exists\, t \in \mathbb{N} \;\; \forall\, h(t) \in H(t) \;\; \forall\, s \in S\; \forall\, \hat{s} \in S \;\; \forall\, i \in N$$
$$(s^{\tau})_{\tau=1}^{t} = (\hat{s}^{\tau})_{\tau=1}^{t} \implies |\delta^t u_i(s|_{h(t)}) - \delta^t u_i(\hat{s}|_{h(t)})| < \varepsilon.$$

ここで，

$$L := \max_{i \in N} \max_{a \in A} |g_i(a)|$$

と置く．そうすると，

$$|\delta^t u_i(s|_{h(t)}) - \delta^t u_i(\hat{s}|_{h(t)})| \leq \delta^t \cdot 2L < 3L\delta^t =: z$$

になる．右辺は，

$$t = \frac{\log z - \log 3L}{\log \delta}$$

となるので，t を次のように選択すればよい．

$$t := \max\{\lceil \frac{\log \varepsilon - \log 3L}{\log \delta} \rceil, 1\}. \qquad \blacksquare$$

6．1　定理 5.1 の十分性の後半の証明

$\Longleftarrow$ を背理法で証明する．補題 5.2 より $\hat{s}$ が OSDP ならば有限の逸脱の列によって improve できないことが分かっている．$\hat{s}$ は SPN ではないとして，これに矛盾することを示そう．SPN ではないのである時点 t のサブゲーム $G(h(t))$ が存在してあるプレイヤー i は $\hat{s}_i$ とは異なる戦略 s_i を用いて自分の効用を improve することができる．

$$\varepsilon := \delta^t u_i(s_i|_{h(t)}, \hat{s}_{-i}|_{h(t)}) - \delta^t u_i(\hat{s}|_{h(t)}) > 0 \tag{5}$$

このとき，次のような戦略 $\tilde{s}_i$ を考えよう：ある時点 $\tau > t$ までは，$\tilde{s}_i$ は s_i と一致し，τ 以後は $\hat{s}_i$ と一致する．補題 6.1 の CI により，ある時点 τ が十分大きければ，次が成り立つ．どんな $h(\tau) \in H_{h(t)}(\tau)$ に対しても

$$\delta^\tau u_i(s_i|_{h(\tau)}, \hat{s}_{-i}|_{h(\tau)}) - \delta^\tau u_i(\tilde{s}_i|_{h(\tau)}, \hat{s}_{-i}|_{h(\tau)}) < \varepsilon/2 \tag{6}$$

(5) から (6) を引くとどんな $h(\tau) \in H_{h(t)}(\tau)$ に対しても

$$\delta^t u_i(s_i|_{h(t)}, \hat{s}_{-i}|_{h(t)}) - \delta^\tau u_i(s_i|_{h(\tau)}, \hat{s}_{-i}|_{h(\tau)}) + \delta^\tau u_i(\tilde{s}_i|_{h(\tau)}, \hat{s}_{-i}|_{h(\tau)}) - \delta^t u_i(\hat{s}|_{h(t)}) > \varepsilon/2$$

これは，ある時点 $\tau > t$ が存在して，1 から t までは $\hat{s}_i$，$t+1$ から τ は s_i と一致し，$\tau+1$ 以後は $\hat{s}_i$ と一致する戦略 $\tilde{s}_i$ は $\hat{s}_i$ を少なくとも $\varepsilon/2$ だけ improve できることを意味する．すなわち，

$$\delta^t u_i(\tilde{s}_i|_{h(t)}, \hat{s}_{-i}|_{h(t)}) - \delta^t u_i(\hat{s}|_{h(t)}) > \varepsilon/2$$

これは有限回の逸脱をもつ $\tilde{s}_i$ が $\hat{s}_i$ を改善しているので矛盾である．$\blacksquare$

OSDP は Sequential Equilibrium の Sequential Rationality とも関係がある．実際，Hendon, Jacobsen, and Sloth (1996) [16] は OSDP と SR は同値であることを示した．しかも，その同値性は Kreps and Wilson が均衡の条件として提示した第 2 の条件 Consistency よりももっと弱い条件の下で成り立っている．

7 Enforceability

ここでは均衡利得の特定化定理が証明される．そのためにはいくつかの概念を知ることが重要である．特に強制可能性 (Enforceability) が鍵となる考え方である．

7．1 Dynamic programming decomposition の直感的説明

ダイナミック・プログラミング・アプローチを直感的に理解するために利得の流列を分解して見よう．アクションの列 $(\alpha^t)_{t=1}^{\infty}$ の下での i の繰り返しゲームの利得 v_i は，

$$
\begin{aligned}
v_i &= (1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}g_i(\alpha^t) = (1-\delta)g_i(\alpha^1) + (1-\delta)\delta g_i(\alpha^2) + (1-\delta)\delta^2 g_i(\alpha^3) + \cdots \\
&= (1-\delta)g_i(\alpha^1) + \delta(1-\delta)\left(g_i(\alpha^2) + \delta g_i(\alpha^3) + \cdots\right) \\
&= (1-\delta)g_i(\alpha^1) + \delta(1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}g_i(\alpha^{t+1}) \\
&= (1-\delta)g_i(\alpha^1) + \delta w_i(1)
\end{aligned}
$$

ただし，$w_i(1) = (1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}g_i(\alpha^{t+1})$ である．同様に t 期の利得は

$$w_i(t-1) = (1-\delta)g_i(\alpha^t) + \delta w_i(t)$$

と書くことができる．ただし，$w_i(t) = (1-\delta)\sum_{\tau=1}^{\infty}\delta^{\tau-1}g_i(\alpha^{t+\tau})$ である．この $w_i(t)$ を continuation payoff という．ここで $(\alpha^t)_{t=1}^{\infty}$ の下で利得ベクトルを

$$\boldsymbol{v} = (1-\delta)\boldsymbol{g}(\alpha^1) + \delta\boldsymbol{w}(1)$$

と書いたとき各プレイヤーは各 t 期に α^t から逸脱するインセンティブがなければ，今期の利得 $\boldsymbol{g}(\alpha^t)$ と来期以降の利得 $\boldsymbol{w}(t)$ に分解できる．この $\boldsymbol{w}$ を public outcome y に依存させて構成しよう．

7．2 Enforceability and Self-decomposability

定義 7.1 (Enforceability) $\delta \in (0,1)$, $W \subset \mathbb{R}^n$, $\boldsymbol{v} \in \mathbb{R}^n$ は所与とする．戦略プロファイル α は $\boldsymbol{v}$, W, δ に関して *enforceable (E)* であるとは

$$\alpha \in \operatorname{graph} B \ \textit{and}\ \exists\, \boldsymbol{w} : Y \to W\ \forall i \in N$$

For $a_i \in \operatorname{support}\alpha_i$,

$$v_i = (1-\delta)g_i(a_i,\ \alpha_{-i}) + \delta\sum_{y\in Y}\pi(y|a_i,\ \alpha_{-i})w_i(y) \tag{7a}$$

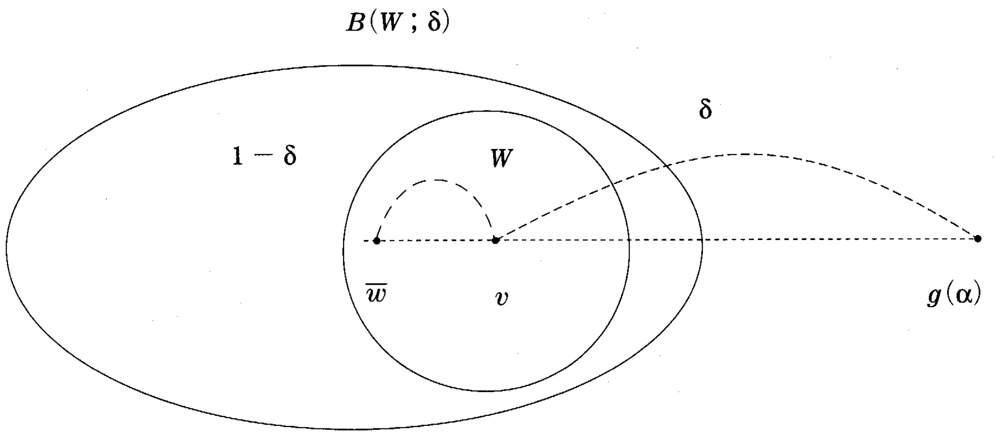


図 4　Self-decomposability

$$\textit{For } a_i' \notin \operatorname{support} \alpha_i,$$

$$v_i \geq (1-\delta) g_i(a_i',\ \alpha_{-i}) + \delta \sum_{y \in Y} \pi(y|a_i',\ \alpha_{-i}) w_i(y) \tag{7b}$$

を満たすことである．ただし $\boldsymbol{w}(y) = (w_1(y), \ldots, w_n(y))^{\mathrm{T}}$ とする． □

行動プロファイル α が $\boldsymbol{v}$, W, δ に関して enforceable であることが成り立つとき，ある変数に注目するため言い換えを行うことがある．第 1 は，利得ベクトル $\boldsymbol{v}$ は α, W, δ に関して *decomposable* という．第 2 の言い換えは，$\boldsymbol{v}$ and δ に関して $\boldsymbol{w}(y)$ は α を *enforce* するというものである．また，continuation payoff の取りうる領域が任意である場合，すなわち $W = \mathbb{R}^n$ であるとき，単に α が $\boldsymbol{v}$ と δ に関して enforceable であるこという．

定義 7.2 (Decomposable set) $W \subset \mathbb{R}^n$ と $\delta \in (0,1)$ が与えられているとする．

$$B(W;\delta) := \{\boldsymbol{v} \in \mathbb{R}^n \mid \exists\, \alpha \in \mathcal{A}\ \alpha\ \textit{is enforceable w.r.t.}\ \boldsymbol{v}, W, \delta\}$$

この集合 $B(W;\delta)$ を *decomposable set* と呼ぶ． □

定義 7.3 (Self-decomposability) 割引因子 δ が与えられているとする．W は *self-decomposable* であるとは $W \subset B(W;\delta)$ を満たすことである． □

8　Abreu, Pearce, and Stacchetti の定理

定理 8.1 (Abreu, Pearce, and Stacchetti (1990)) W が有界ならば

$$W \subset B(W;\delta) \Longrightarrow W \subset E(\delta)$$

が成り立つ． □

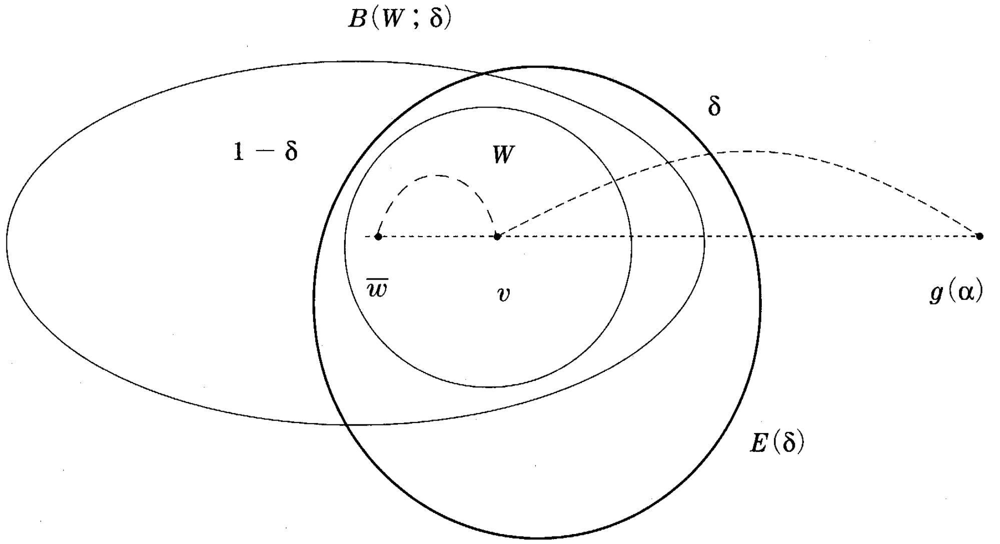


図 5　Abreu, Pearce, and Stacchetti の定理

例：ナッシュ均衡.　ステージゲーム Nash 均衡を α^e として，これを無限回繰り返す戦略の利得は均衡利得であることを示す．$\boldsymbol{v}^e := \boldsymbol{g}(\alpha^e)$ とする．定理 8.1 より $W := \{\boldsymbol{v}^e\}$ は self-decomposable であることを示せばよい．$\boldsymbol{v}^e \in B(W;\delta)$ であることを示そう．Enforceability の定義において，

$$\boldsymbol{w}(y) = \boldsymbol{v}^e \text{ for all } y \in Y$$

でなければならない．$B(W;\delta)$ の定義で存在を示さないといけない α を α^e とする．すると (7a) は，

$$(1-\delta)g_i(\alpha^e) + \delta w_i(y) = (1-\delta)g_i(\alpha^e) + \delta g_i(\alpha^e) = g_i(\alpha^e) = v_i^e$$

(7b) は α^e がステージゲーム Nash 均衡である事から直ちに満たされる．よって，$\boldsymbol{v}^e \in B(W;\delta)$ が示された．

8．1　定理 8.1 の証明

集合 W は有界で self-decomposable (SD) である，すなわち

$$W \subset B(W;\delta) \tag{8}$$

ならば $W \subset E(\delta)$ であることを示す．それは，(8) を満たしている任意の $\boldsymbol{v} \in W$ に対して $\boldsymbol{v} \in E(\delta)$ であることを示す事と同じである．そのとき，ある戦略 σ が存在して，

1. σ は public 戦略である

2. 定理 1 より σ は OSDP を満たす

3. σ は $\boldsymbol{v}$ を生成する

ことを示せばよい．

任意に $\boldsymbol{v} \in W$ を取る．SD 条件 (8) から $\boldsymbol{v} \in B(W;\delta)$ であるので，ある action α^1 が存在する．また E 条件から $\boldsymbol{w}^1 : Y \to W$ が存在して (7a) と (7b) を満たす．この時

$$\sigma^1(h(0)) := \alpha^1$$

とする．

$t = 1$ の public outcome が y^1 であったとする．この時，$\boldsymbol{w}^1(y^1) \in W$ である．新たに $\boldsymbol{v}^1 := \boldsymbol{w}^1(y^1)$ と置こう．$\boldsymbol{v}^1 \in W$ と SD 条件 (8) からある action $\alpha^2_{\boldsymbol{w}^1(y^1)}$ と E 条件から $\boldsymbol{w}^2_{\boldsymbol{w}^1(y^1)} : Y \to W$ が存在して (7a) と (7b) を満たす．この時

$$\sigma^2(h(0), y^1) := \alpha^2_{\boldsymbol{w}^1(y^1)}$$

とおこう．

以下同様に，ある history $h(t-1)$, あるアクション $\alpha^t_{h(t-1)}$, continuation payoff $\boldsymbol{w}^t_{h(t-1)}$ が定まっているときに，t 期に y^t が実現したとする．その時，$h(t) := (h(t-1), y^t)$ とおこう．SD 条件 $\boldsymbol{w}^t_{h(t-1)}(y^t) \in B(W;\delta)$ から存在を保証される $\alpha^{t+1}_{h(t)}$ に対して，

$$\sigma^{t+1}(h(t)) := \alpha^{t+1}_{h(t)}$$

と定義していく．E 条件から新たな continuation payoff $\boldsymbol{w}^{t+1}_{h(t)}$ が存在して，$\boldsymbol{w}^{t+1}_{h(t)}(y^{t+1})$ と $\alpha^{t+1}_{h(t)}$ に関して (7a) と (7b) が成り立つ．こうして生成される戦略を σ とする．

(1) 定義より明らかに σ は public 戦略である．(2) 任意の history $h(t-1)$ におけるアクション $\sigma^t(h(t-1)) = \alpha^t$ は $h(t-1)$ が到達可能で来期以降の continuation payoff $\boldsymbol{w}^t(y^t)$ の下で (7a) と (7b) を満たしている．

$$\begin{aligned}
&\forall\, a_i \in \text{support}\,\alpha^t_i\ \forall a'_i \notin \text{support}\,\alpha^t_i \\
&\quad (1-\delta)g_i(a_i,\ \alpha^t_{-i}) + \delta \sum_{y^t \in Y} \pi(y^t|a_i,\ \alpha^t_{-i}) w^t_i(y^t) \\
&\quad \geq (1-\delta)g_i(a'_i,\ \alpha^t_{-i}) + \delta \sum_{y^t \in Y} \pi(y^t|a'_i,\ \alpha^t_{-i}) w^t_i(y^t)
\end{aligned}$$

これは α^t から一回だけ逸脱して $t+1$ 期から σ に復帰するインセンティブはないことを示している．よって OSDP が満たされている．

(3) $u_i(\sigma;\delta) = v_i$ を示そう．σ の定義により，

$$v_i = (1-\delta)g_i(\alpha^1) + \delta \sum_{y^1 \in Y} \pi(y^1|\alpha^1) w_i^1(y^1)$$

$\mu^0(h(0)|\sigma) = 1$ と $h(1) = (h(0), y^1)$ に注意すると，

$$\begin{aligned}
&= (1-\delta)\mathbb{E}\left[g_i(\sigma^1(h(0)))\right] + \delta\mathbb{E}\left[w_i^1(h(1))\right] \\
&= (1-\delta)\mathbb{E}\left[g_i(\sigma^1(h(0)))\right] + \delta\mathbb{E}\left[(1-\delta)g_i(\alpha^2_{h(1)}) + \delta \sum_{y^2 \in Y} \pi(y^2|\alpha^2_{h(1)}) w_i^2(y^2; h(1))\right]
\end{aligned}$$

$h(2) = (h(0), y^1, y^2)$ に注意すると，

$$\begin{aligned}
&= (1-\delta)\mathbb{E}\left[g_i(\sigma^1(h(0)))\right] + \delta(1-\delta)\mathbb{E}\left[g_i(\sigma^2(h(1)))\right] + \delta^2\mathbb{E}\left[w_i^2(h(2))\right] = \cdots \\
&= (1-\delta)\sum_{t=1}^{\tau} \delta^{t-1}\mathbb{E}\left[g_i(\sigma^t(h(t-1)))\sigma\right] + \delta^\tau \mathbb{E}\left[w_i^\tau(h(\tau))\right].
\end{aligned}$$

ここで $\delta \in (0,1)$ で W は有界なので，

$$\delta^\tau \mathbb{E}\left[w_i^\tau(h(\tau))\right] \to 0 \quad (\tau \to \infty).$$

よって，

$$v_i = \lim_{\tau\to\infty} (1-\delta)\sum_{t=1}^{\tau} \delta^{t-1}\mathbb{E}\left[g_i(\sigma^t(h(t-1)))\right] = u_i(\sigma;\delta).$$

であることが証明された． ∎

つぎに W 自体が均衡利得の場合には次の定理が成り立つ．

定理 8.2 $E(\delta) = B(E(\delta);\delta)$ □

Proof: $E(\delta) \supset B(E(\delta);\delta)$ の証明．任意のベクトル $\boldsymbol{v} \in B(E(\delta);\delta)$ をとる．すると，$E(\delta)$, $\boldsymbol{v}$ δ に関して enforceable アクション α が存在する．これを 1 期目のアクションとする．Enforceability より continuation payoff $\boldsymbol{w}^1(y^1)$ が存在して $\boldsymbol{w}^1(y^1) \in E(\delta)$ が成り立つ．その時に条件から σ_{y^1} が存在して $\boldsymbol{u}(\sigma_{y^1};\delta) = \boldsymbol{w}^1(y^1)$ かつ $\sigma_{y^1} \in \mathcal{S}^*(\delta)$ となる．新たに，戦略を $\sigma' := (\alpha, (\sigma_{y^1})_{y^1 \in Y})$ とおけば σ' は PPE で $\boldsymbol{v}$ を生成する．

$E(\delta) \subset B(E(\delta);\delta)$ の証明．任意のベクトル $\boldsymbol{v} \in E(\delta)$ をとる．$\boldsymbol{v}$ を生成する PPE 戦略 σ が存在する．特に 1 期の戦略 σ^1 と 2 期以降戦略 $(\sigma^t)_{t=2}^{\infty}$ から作られる continuation payoff $\boldsymbol{w}^1(y)$ に対して逸脱しても利益にならない．$h(1) = (h(0), y^1)$ であることを思い出そう．

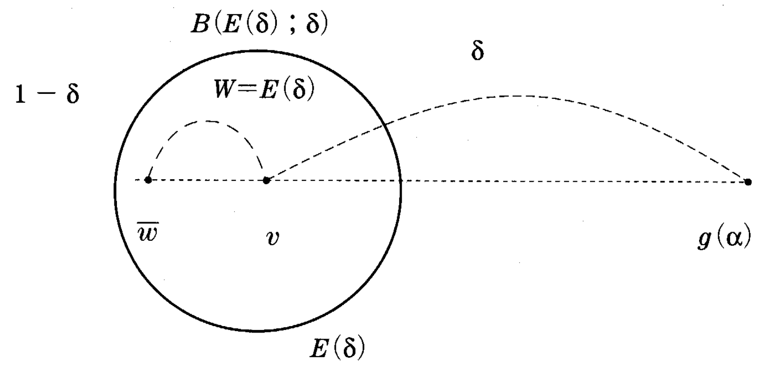


図 6　均衡利得集合の性質

$$
\begin{aligned}
v_i &= (1-\delta)g_i(\sigma^1) + (1-\delta)\sum_{\tau=1}^{\infty}\delta^{\tau-1}\mathbb{E}\left[g_i(\sigma^{\tau+1}(h(\tau)))\right] \\
&= (1-\delta)g_i(\sigma^1) + (1-\delta)\sum_{y^1\in Y}\pi(y^1|\sigma^1)\sum_{\tau=1}^{\infty}\delta^{\tau-1}\mathbb{E}\left[g_i(\sigma^{\tau+1}|_{h(1)}(h(\tau)))\right] \\
&= (1-\delta)g_i(\sigma^1) + (1-\delta)\sum_{y^1\in Y}\pi(y^1|\sigma^1)u_i(\sigma|_{h(1)};\delta).
\end{aligned}
$$

ここで $w_i^1(y^1) := u_i(\sigma|_{h(1)};\delta)$ と定義すれば，サブゲーム完全性より $\boldsymbol{w}^1(y^1)$ は $E(\delta)$ に属さなければならない． ■

この定理が意味することは，写像 $B(\cdot;\delta) : W \mapsto B(W;\delta)$ を考えると $E(\delta)$ は最大の不動点(集合) であるということである．

8．2　Locally Self-decomposability

Abreu, Pearce, and Stacchetti の Self-decomposability 条件をより弱めたのが Fudenberg, Levine, and Maskin (1994) の locally self-decomposability である．それを紹介しよう．

定義 8.1 $W \subset \mathbb{R}^n$ は *locally self-decomposable* であるとは

$$\forall\, \boldsymbol{v} \in W \;\; \exists\, \delta_{\boldsymbol{v}} \in (0,1) \;\; \exists\, \boldsymbol{v} \text{ のある開近傍 } U_{\boldsymbol{v}}$$

$$U_{\boldsymbol{v}} \cap W \subset B(W;\delta_{\boldsymbol{v}}).$$

を満たすことである． □

定理 8.3 (Fudenberg, Levine, and Maskin (1994)) $W \subset \mathbb{R}^n$ はコンパクト，凸，locally self-decomposable であるならば，

$$\exists\, \delta' < 1 \;\forall\, \delta \in [\delta', 1)\; W \subset E(\delta).$$

が成り立つ． □

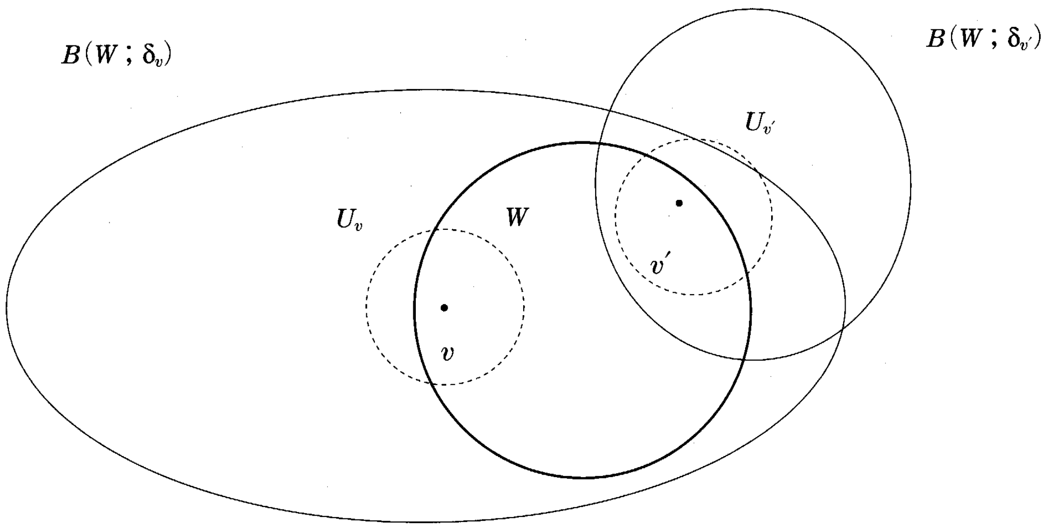


図 7　Locally Self-decomposability

Proof:　集合 W は locally self-decomposable なので

$$\begin{aligned} &\exists\, W \text{ の開被覆 } \{U_\lambda\}_{\lambda\in\Lambda}\ \exists \text{付随する割引因子 } \{\delta_{U_\lambda}\} \\ &\forall \boldsymbol{v} \in U_\lambda \cap W \exists\, \alpha\ \exists\, \boldsymbol{w}(\cdot;\delta_{U_\lambda}) : Y \to W \\ &\text{(7a) と (7b) が } \delta_{U_\lambda} \text{ の下で成り立つ} \end{aligned} \tag{9}$$

集合 W はコンパクトなので, $\{U_\lambda\}_{\lambda\in\Lambda}$ からその有限被覆 $\{U_{\lambda'}\}_{\lambda'\in\Lambda'}$ を取り出せる．ここで次のように最大の割引因子を設定する．

$$\delta' := \max\{\delta_{U_{\lambda'}} \mid \lambda' \in \Lambda'\}.$$

この $\delta_{U_{\lambda'}}$ の数は有限なので, $\delta' < 1$ になる．ここで

$$\forall\, \delta \in (\delta', 1)\ W \subset B(W;\delta)$$

を導こう．任意に $\delta \in (\delta', 1)$ と $\boldsymbol{v} \in W$ を取る．(8) から $\boldsymbol{w}(y;\delta_{U_{\lambda'}})$ は $\boldsymbol{v} \in U_{\lambda'} \cap W$ と $\delta_{U_{\lambda'}}$ に関して α を enforce する．$\boldsymbol{v}$ は α, δ, およびつぎの $\boldsymbol{w}(y;\delta)$ に関して decomposable であることを示す．ここで $\underline{\delta} := \delta_{U_{\lambda'}}$ とおいて

$$\boldsymbol{w}(y;\delta) := \frac{\delta - \underline{\delta}}{\delta(1-\underline{\delta})}\boldsymbol{v} + \frac{\underline{\delta}(1-\delta)}{\delta(1-\underline{\delta})}\boldsymbol{w}(y;\underline{\delta}).$$

とする．この self-decomposability を確かめる．ベクトル $\boldsymbol{v}, \boldsymbol{w}(y;\underline{\delta}) \in W$ であり W は凸であるから

$$\frac{\delta - \underline{\delta}}{\delta(1-\underline{\delta})} > 0, \quad \frac{\underline{\delta}(1-\delta)}{\delta(1-\underline{\delta})} > 0, \quad \frac{\delta - \underline{\delta}}{\delta(1-\underline{\delta})} + \frac{\underline{\delta}(1-\delta)}{\delta(1-\underline{\delta})} = \frac{\delta(1-\underline{\delta})}{\delta(1-\underline{\delta})} = 1,$$

が成り立つので $\boldsymbol{w}(y;\delta) \in W$ が分かる．次に Enforceability に関しては次のように導くことができる．

$$
\begin{aligned}
&(1-\delta)g_i(\alpha) + \delta \sum_{y \in Y} \pi(y|\alpha) w_i(y;\delta) \\
&= (1-\delta)g_i(\alpha) + \delta \sum_{y \in Y} \pi(y|\alpha) \left(\frac{\delta - \underline{\delta}}{\delta(1-\underline{\delta})} v_i + \frac{\underline{\delta}(1-\delta)}{\delta(1-\underline{\delta})} w_i(y;\underline{\delta}) \right) \\
&= (1-\delta)g_i(\alpha) + \frac{\delta - \underline{\delta}}{1-\underline{\delta}} v_i + \frac{\underline{\delta}(1-\delta)}{1-\underline{\delta}} \sum_{y \in Y} \pi(y|\alpha) w_i(y;\underline{\delta}) \\
&= (1-\delta)g_i(\alpha) + \frac{\delta - \underline{\delta}}{1-\underline{\delta}} v_i + \frac{1-\delta}{1-\underline{\delta}} (v_i - (1-\underline{\delta}) g_i(\alpha)) \\
&= (1-\delta)g_i(\alpha) + \frac{\delta - \underline{\delta}}{1-\underline{\delta}} v_i + \frac{1-\delta}{1-\underline{\delta}} v_i - (1-\delta) g_i(\alpha) = v_i.
\end{aligned}
$$

同様に $a_i \notin \operatorname{support} \alpha_i$ に関しても α と $\boldsymbol{w}(y;\delta)$ は δ の下で (7a) と (7b) を満たすことを示すことができる．集合 W はコンパクトなので W は有界である．よって定理 8.3 が証明された． ∎

9　線型計画問題

均衡利得を求める場合に自己分解性が重要な概念であることを先に示した．さらにこのような性質を持つ集合を具体的に求める方法を紹介しよう．

行動プロファイル $\alpha \in \mathcal{A}$ と方向ベクトル

$$\boldsymbol{\lambda} \in \Lambda := \{\boldsymbol{\lambda} \in \mathbb{R}^n \mid \|\boldsymbol{\lambda}\| = 1\}$$

を所与とする．次のような線形計画問題 LP#1 を考えよう．

$$
\begin{aligned}
&k^*(\alpha, \boldsymbol{\lambda}) := \\
&\max_{\boldsymbol{v}, \boldsymbol{w}(y)} \langle \boldsymbol{\lambda}, \boldsymbol{v} \rangle \quad \text{subject to}
\end{aligned}
\tag{10}
$$

$$v_i = (1-\delta) g_i(a_i,\ \alpha_{-i}) + \delta \sum_{y \in Y} \pi(y|a_i,\ \alpha_{-i}) w_i(y) \tag{10a}$$

$$\text{for } a_i \in \operatorname{support} \alpha_i \text{ for } i \in LR,$$

$$v_i \geq (1-\delta) g_i(a_i,\ \alpha_{-i}) + \delta \sum_{y \in Y} \pi(y|a_i,\ \alpha_{-i}) w_i(y) \tag{10b}$$

$$\text{for } a_i \notin \operatorname{support} \alpha_i \text{ for } i \in LR,$$

$$\langle \boldsymbol{\lambda}, \boldsymbol{v} \rangle \geq \langle \boldsymbol{\lambda}, \boldsymbol{w}(y) \rangle \quad \text{for all } y \in Y. \tag{10c}$$

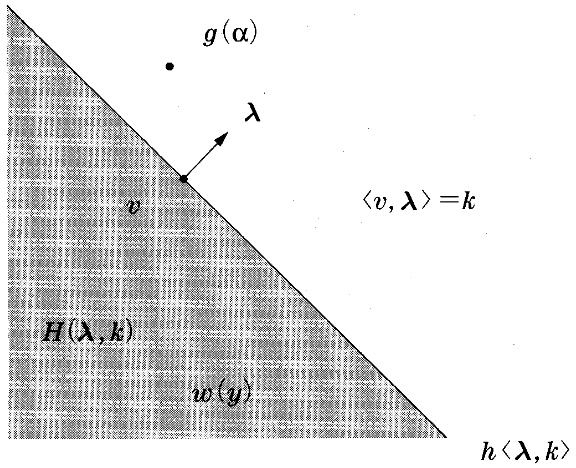


図8　半空間

半空間と超平面を以下のように定義する．

$$H(\boldsymbol{\lambda}, k) := \{\boldsymbol{v} \in \mathbb{R}^n \mid \langle \boldsymbol{\lambda}, \boldsymbol{v} \rangle \leq k\}, \quad h(\boldsymbol{\lambda}, k) := \{\boldsymbol{v} \in \mathbb{R}^n \mid \langle \boldsymbol{\lambda}, \boldsymbol{v} \rangle = k\}$$

この問題は以下のような性質を持っている．

- LP#1 は feasible solution をもたないときは，$k^*(\alpha, \boldsymbol{\lambda}) := -\infty$.
- $k^*(\alpha, \boldsymbol{\lambda})$ は δ に依存しない．
- $k^*(\alpha, \boldsymbol{\lambda}) \leq \langle \boldsymbol{\lambda}, \boldsymbol{g}(\alpha) \rangle$.
- $k^*(\alpha, \boldsymbol{\lambda}) = \langle \boldsymbol{\lambda}, \boldsymbol{g}(\alpha) \rangle \iff \boldsymbol{w}(y) \in h(\boldsymbol{\lambda}, k^*(\alpha, \boldsymbol{\lambda}))$
- α がステージゲーム Nash 均衡であれば，$k^*(\alpha, \boldsymbol{\lambda}) = \langle \boldsymbol{\lambda}, \boldsymbol{g}(\alpha) \rangle$.

次に最大スコア Maximum score を以下に定義する．

$$k^*(\boldsymbol{\lambda}) := \sup_{\alpha \in \mathrm{graph}\, B} k^*(\alpha, \boldsymbol{\lambda}).$$

さらに方向 $\boldsymbol{\lambda}$ の最大半空間を以下に定義する．

$$H^*(\boldsymbol{\lambda}) := \{\boldsymbol{v} \in \mathbb{R}^n \mid \langle \boldsymbol{\lambda}, \boldsymbol{v} \rangle \leq k^*(\boldsymbol{\lambda})\}.$$

方向 $\boldsymbol{\lambda}$ の最大超平面は $h^*(\boldsymbol{\lambda}) := h(\boldsymbol{\lambda}, k^*(\boldsymbol{\lambda}))$ を満たしている．明らかに $h^*(\boldsymbol{\lambda}) \cap V' \neq \emptyset$ が成り立つ．ノルムが 1 でよい訳は次のように分かる．方向 $\boldsymbol{\lambda}' := t\boldsymbol{\lambda}$ for some $t > 0$ を考えよう．(10c) より最適解は $\boldsymbol{\lambda}$ と同じである．半空間の定義により 2 つの最大半空間は等しい．

最後に集合 Q を以下に定義する．

$$Q := \bigcap_{\boldsymbol{\lambda} \in \mathbb{R}^n} H^*(\boldsymbol{\lambda}).$$

明らかに $Q \subset V'$ である．そして，このアルゴリズムは次の定理によって正当化される．

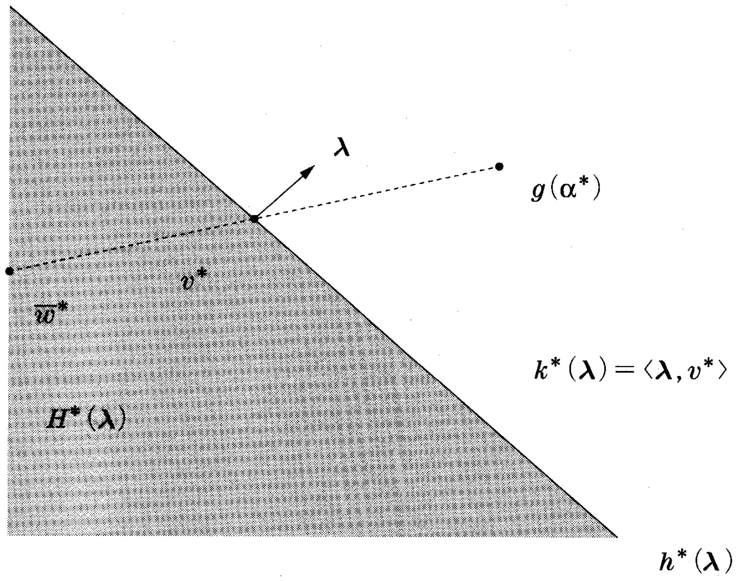


図9　最大半空間

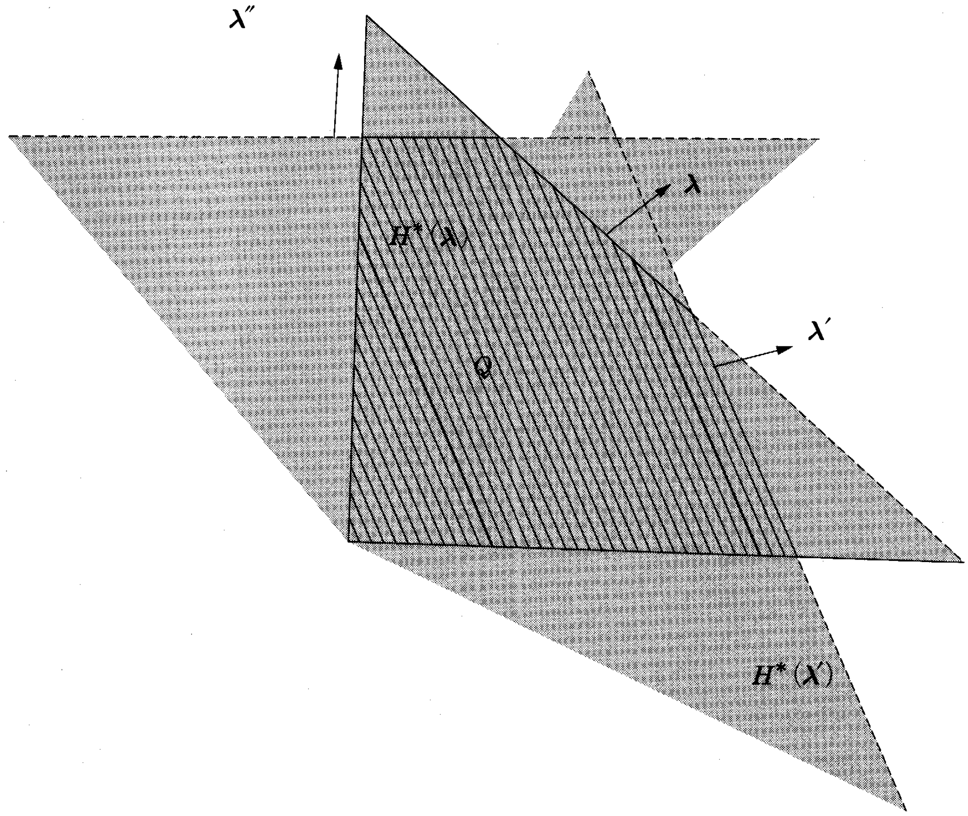


図10　集合 Q

定理 9.1 (Fudenberg and Levine(1994)) $\dim Q = n$ ならば,

$$\lim_{\delta\to 1} E(\delta) = Q$$

が成り立つ. ∎

10 Sufficient Rank

前節では均衡特定化アルゴリズムを紹介した．これらの手法は必ずしもフォーク定理が成り立たなくても均衡利得を定めることを可能にしている．それではフォーク定理が成り立つ場合はどのような条件が必要であろうか？この節では鍵概念となる Sufficient Rank (十分階数条件) を紹介しよう．

i の情報行列を $\Pi_i := (\pi_i(y_i|a_i))_{m_i \times \mu_i}$ とする．α_{-i} の下での $m_i \times 1$ 利得ベクトルを次のように定義する．

$$\boldsymbol{g}_i(\alpha_{-i}) := \begin{pmatrix} g_i(a_i^1, \alpha_{-i}) \\ g_i(a_i^2, \alpha_{-i}) \\ \vdots \\ g_i(a_i^{m_i}, \alpha_{-i}) \end{pmatrix}$$

定義 10.1 (sufficient rank FL (1994)) 戦略プロファイル α が *sufficient rank (SR)* 条件を満たすとは以下の条件が成り立つことである．

$$\forall i \in N \ \forall \boldsymbol{x} \in \mathbb{R}^{m_i} \boldsymbol{x}^{\mathrm{T}} \Pi_i = \boldsymbol{0}^{\mathrm{T}} \Longrightarrow \langle \boldsymbol{x}, \boldsymbol{g}_i(\alpha_{-i}) \rangle = 0$$

SR が満たされれば i のどんなアクションも彼の public outcome のみに依存した継続利得 continuation payoff によって今期の利得と継続利得の和によって無差別にすることができる．あるステージゲーム Nash 均衡を α^e とする．$\boldsymbol{g}(\alpha^e)$ と $\boldsymbol{g}(\alpha^e)$ をパレート支配するパレート効率的かつ実行可能な利得との凸包を V^0 とする．

定理 10.1 (FLM (1994) のフォーク定理) $\dim V^0 = n$ のとき，$\lim_{\delta \to 1} E(\delta) \supset V^0$ である． ∎

11 Weak Sufficient Rank

前節ではフォーク定理を導く Sufficient Rank (十分階数条件) とそれを用いたフォーク定理を紹介した．この節ではその条件よりももっと弱い Weak Sufficient Rank (弱十分階数条件) を導入して同様にフォーク定理を提示する．

$\boldsymbol{g}_i(\alpha_{-i})$ に対して α_i の下で正の確率を付与するアクションに対応する要素を上へ, 0 の確率を付与する要素を下へ持ってくる置換を施す:

$$\boldsymbol{g}_i(\alpha_{-i}) = \begin{pmatrix} \boldsymbol{g}_i^p(\alpha_{-i}) \\ \boldsymbol{g}_i^0(\alpha_{-i}) \end{pmatrix}$$

$\boldsymbol{g}_i^p(\alpha_{-i})$ は正の確率を付与する利得ベクトル, $\boldsymbol{g}_i^0(\alpha_{-i})$ は確率 0 のベクトルである. $\#\,\mathrm{support}\,\alpha_i = p_i$ ならば前者は p_i 列ベクトル, 後者は $m_i - p_i$ 列ベクトルである. $\boldsymbol{g}_i(\alpha_{-i})$ と同じ置換を Π_i にも施して, 小行列 Π_i^p, Π_i^0 を作る:

$$\Pi_i = \begin{pmatrix} \Pi_i^p \\ \Pi_i^0 \end{pmatrix}$$

Π_i^p は $p_i \times \mu_i$ 行列, Π_i^0 は $(m_i - p_i) \times \mu_i$ 行列である.

定義 11.1 (Weak Sufficient Rank (Tanno (2000))) 戦略プロファイル α は *weak sufficient rank condition (WSR)* を満たすとは次の条件

$$\forall\, i \in N\ \forall\, \boldsymbol{x} \in \mathbb{R}^{p_i}\ \forall\, \boldsymbol{y} \in \mathbb{R}^{m_i - p_i}$$

$$\boldsymbol{x}^{\mathrm{T}}\Pi_i^p + \boldsymbol{y}^{\mathrm{T}}\Pi_i^0 = \boldsymbol{0}^{\mathrm{T}}\ \textit{and}\ \boldsymbol{y} \geq \boldsymbol{0} \Longrightarrow \langle \boldsymbol{x}, \boldsymbol{g}_i^p(\alpha_{-i})\rangle + \langle \boldsymbol{y}, \boldsymbol{g}_i^0(\alpha_{-i})\rangle \leq 0.$$

が満たされることである. □

あきらかに SR ならば WSR であることが分かる.

定理 11.1 *E* と *WSR* は同値である.

補題 11.1 パレート効率的なプロファイルは *WSR* を満たす.

定義 11.2 (decomposability of the origin (DO)) 戦略プロファイル α は原点を分解できる *decomposability of the origin (DO)* とは次の条件が満たされることである.

$$\forall\, i \in N\ \ \exists\, \boldsymbol{h}_i \in \mathbb{R}^{\mu_i}$$

$$\Pi_i^p \boldsymbol{h}_i + \boldsymbol{g}_i^p(\alpha_{-i}) = \boldsymbol{0} \quad \textit{and} \quad \Pi_i^0 \boldsymbol{h}_i + \boldsymbol{g}_i^0(\alpha_{-i}) \leq \boldsymbol{0}$$

補題 11.2 *WSR* と *DO* は同値である.

定義 11.3 (ecomposability of any payoff (DA)) 戦略プロファイル α は任意の利得を分解できる *decomposability of any payoff (DA)* とは以下の条件が成り立つことである.

$$\forall\, i \in N\ \ \forall\, v_i \in \mathbb{R}\ \forall\, \delta \in (0,1]\ \ \exists\, w_i : Y_i \to \mathbb{R}$$

$$v_i = (1-\delta) g_i(a_i,\ \alpha_{-i}) + \delta \sum_{y_i \in Y_i} \pi_i(y_i|a_i) w_i(y_i)$$

$$\text{for } a_i \in \mathrm{support}\,\alpha_i$$

$$v_i \geq (1-\delta) g_i(a_i,\ \alpha_{-i}) + \delta \sum_{y_i \in Y_i} \pi_i(y_i|a_i) w_i(y_i)$$

$$\text{for } a_i \notin \mathrm{support}\,\alpha_i$$

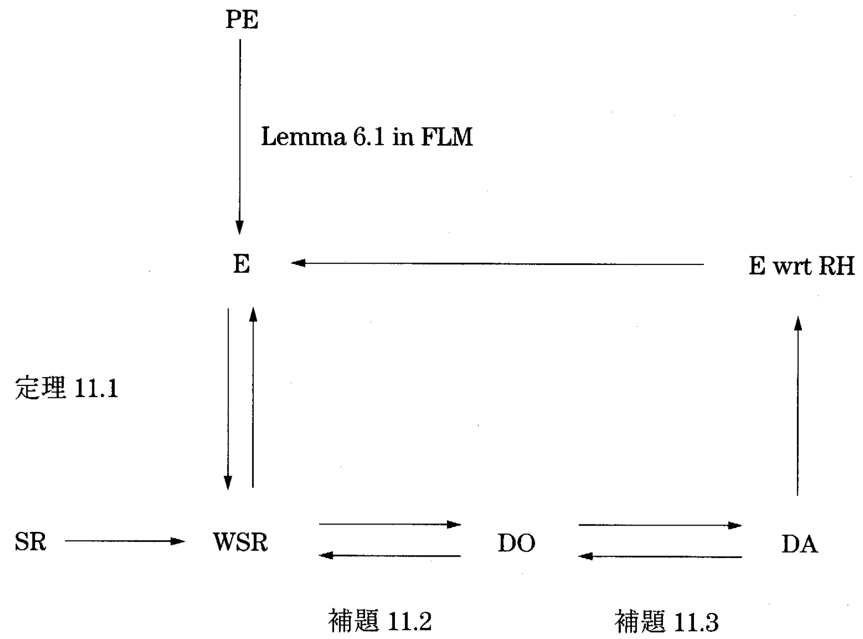


図 11　執行性と同値な条件：PE: パレート効率的, E: 執行性, WSR: 弱十分階数, SR: 十分階数, DO: 原点の分解性, DA: 任意の利得の分解性, E wrt RH: 正則超平面に関する執行性.

補題 11.3 *WSR と DA は同値である.*

Tanno (2000) は Weak sufficient rank 条件の下でフォーク定理を証明した.

- **条件 1**: すべての extreme profiles $a^e \in A^{\mathrm{ex}}$ は WSR 条件を満たす.
- **条件 2**: すべてのミニマックスプロファイル $\underline{\alpha}^i$ は WSR 条件を満たす.

定理 11.2 (フォーク定理 (Theorem 6.1, p.15)) 条件 *1* と *2* の下で

$$\lim_{\delta \to 1} E(\delta) \supset V^*$$ ■

11．1　例 (Investment Game)

Tanno (2000) が検討した Investment Game を見てみよう.

- $n = 2$, プレイヤー 1 は生産を行う．プレイヤー 2 は 1 に投資をする (i) かしないか (q) を決断.
- 投資をしなければ 1 の利得は 0, 2 の利得は 1.
- $A_1 = \{\mathrm{g}, \mathrm{c}, \mathrm{a}\}$.
- $a_1 = \mathrm{g}, \mathrm{c}$ のとき生産は不確実: 確率 2/3 で産出 6, 確率 1/3 で産出 0.
- $a_1 = \mathrm{g}$ のとき 1 はすべての産出量 (6) を得る.
- $a_1 = \mathrm{c}$ のとき 1 は半分の産出量 (3) を得, 残り (3) を 2 に支払う.

- $a_1 = \mathrm{a}$ のとき 1 はコスト (3) を負担して生産を確実にすることができる.
- 費用を除いた産出量 (3) をすべて 2 に配当する.
- 1 は 2 の行動をすべて観察できる.
- 2 は 1 の配当支払いのみを観察できる.
- 1 の public outcome は $\underline{y} = 0, \overline{y} = 3$.

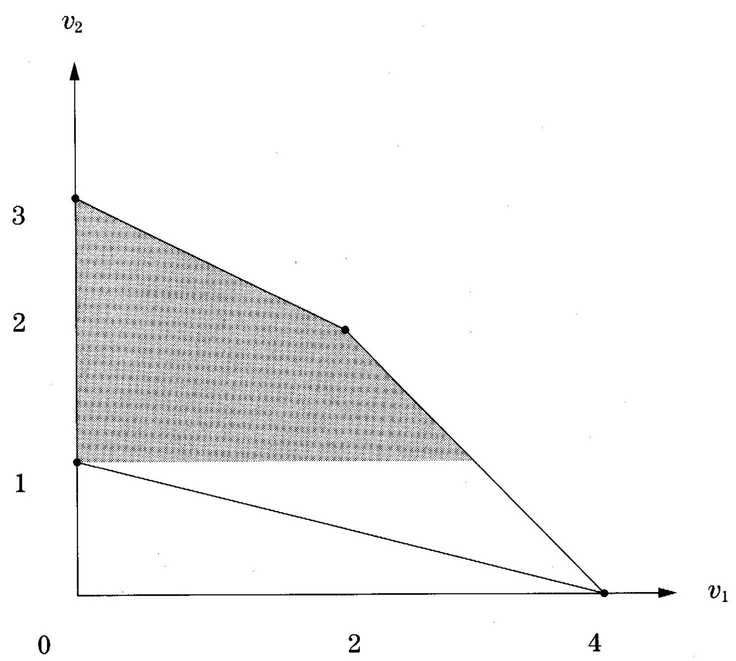
$$\Pi_1 = \begin{array}{c} \\ \mathrm{g} \\ \mathrm{c} \\ \mathrm{a} \end{array} \begin{array}{c} \begin{array}{cc} \underline{y} & \overline{y} \end{array} \\ \begin{pmatrix} 1 & 0 \\ \frac{1}{3} & \frac{2}{3} \\ 0 & 1 \end{pmatrix} \end{array}$$

$r_1(a_1, y_1, \mathrm{i})$		
	$\underline{y}$	$\overline{y}$
g	4	0
c	0	3
a	−3	0

$r_2(\underline{y}, \mathrm{i}) = 0$

$r_2(\overline{y}, \mathrm{i}) = 3$

	i		q	
g	4	0	0	1
c	2	2	0	1
a	0	3	0	1



- $a = (\mathrm{g}, \mathrm{i})$ は weak sufficinet rank を満たすが sufficient rank は満たさない.

例その 2

- $a = (\mathrm{g}, \mathrm{i})$ は sufficient rank を満たさない.

$$\Pi_1 = \begin{pmatrix} 1 & 0 \\ \frac{1}{3} & \frac{2}{3} \\ 0 & 1 \end{pmatrix} \quad \boldsymbol{g}_1(a_2^1) = \begin{pmatrix} 4 \\ 2 \\ 0 \end{pmatrix}$$

$$\boldsymbol{x}^{\mathrm{T}} \begin{pmatrix} 1 & 0 \\ \frac{1}{3} & \frac{2}{3} \\ 0 & 1 \end{pmatrix} = \mathbf{0}^{\mathrm{T}} \Rightarrow \boldsymbol{x} \begin{pmatrix} 4 \\ 2 \\ 0 \end{pmatrix} \neq 0$$

- $a = (\mathrm{g}, \mathrm{i})$ は weak sufficient rank を満たす.
- g と c のみに正の確率を付与する場合:

$$\boldsymbol{x}^{\mathrm{T}} \begin{pmatrix} 1 & 0 \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix} + y(0\ 1) = \mathbf{0}^{\mathrm{T}} \text{ and}$$

$$y \geq 0 \Rightarrow \langle \boldsymbol{x}, (4\ 2)^{\mathrm{T}} \rangle + \langle y, 0 \rangle \leq 0.$$

- c と a のみに正の確率を付与する場合:

$$\boldsymbol{x}^{\mathrm{T}} \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ 0 & 1 \end{pmatrix} + y(1\ 0) = \mathbf{0}^{\mathrm{T}} \text{ and}$$

$$y \geq 0 \Rightarrow \langle \boldsymbol{x}, (2\ 0)^{\mathrm{T}} \rangle + \langle y, 4 \rangle \leq 0.$$

12 実験

カルテルを経済学的にモデル化する際に用いられるゲーム理論は繰り返しゲームの手法であり，そこで得られる理論的帰結は通常現実の経済からのデータによって検証することは難しい．しかし，近年の実験経済学の発展によって，この繰り返しゲームの理論から導かれる帰結を，コントロールされた実験室において被験者に対して実験をうまく設計することによって，実証的な検証を行う研究が行われるようになってきている．この点についての過去の実験研究を，Dal Bó（2004）の記述に沿って概観する．実験研究では，有限繰り返しゲームにおいても協力関係が発生するとの研究もあるが，ここでは基本的に無限繰り返しゲームの理論予測の検証する研究について言及し，特に，割引率の取り扱いという観点から，各研究の簡単な内容の紹介をする．

Murnighan and Roth (1983) と Roth and Murnighan (1978) は無限繰り返し囚人のジレンマゲームの分析を目的としており，どちらの論文も割引率をゲームの確率的終了によって表現し実験を行っている．これらの研究では，平均的に，協力の戦略を取ることが均衡結果の時はそれが均衡結果ではないトリートメントよりも多く協力が観察された．そして，ゲームの停止確率を変化させることによって，協力の発生する割合の比較の検討をしており，結果は表 2 のとおりである．さらに，Roth and Murnighan (1978) では，表 2 にみられるように，より高いゲームの継続確率において，ゲームの一回目における協力するプレーヤーの数が大きくなっ

ていた．対照的に，Murnighan and Roth (1983) では，より高いゲームの継続確率がゲームの最初の回においてより多くの協力関係を示しているわけではないことが示された．

表 2　繰り返しゲームの 1 回目において観察された「協力」の戦略をとった割合（%）- ゲームの継続確率

ゲームの継続確率	0.105	0.5	0.895
Roth and Murnighan (1978)	19.2	9.75	36.36
Murnighan and Roth (1983)	17.83	37.48	29.07

従って，割引因子が大きくなることによって，ここでは必ずしも協調の割合が高くなっているというわけではない．しかし，これらの研究においては，被験者への報酬の与え方がプレーヤーの利得に比例する形で与えられているわけではなく，実験参加に対する固定的報酬として与えられており，ゲームのプレーと報酬が結びついておらず，自己の効用に従ったプレーを実験において行っているとは考えられない．また，どちらの実験においても，被験者はお互いに被験者同士でプレーするのではなく，機械的にトリガー戦略やしっぺ返し戦略を用いる実験者とプレーしている．これらの点が現実的ではなく問題があり，結果が必ずしも繰り返しゲームの状況を適切に表現した結果となっているわけではない．Feinberg and Husted (1993) は，複占モデルから導かれた利得によって，ステージゲームを囚人のジレンマゲームに置き換えて実験を行っている．確率的にゲームが終了するが，割引率はステージゲームからステージゲームに移る際の一定割合の利得の減少によって表現されている．そして，この減少割合を変化させることによって，割引率に応じて被験者が協力の戦略を取る割合の比較を行っている．結果は，割引因子の上昇によって協力の割合の上昇が観察されたが，その割合は想定していたものよりは低いものであった．しかし，被験者に支払われる報酬が，かなり低いものであったことから報酬から導かれるインセンティブによってゲームのプレーが行われたかどうかという点について疑問が残る．

Palfrey and Rosenthal (1994) は，公共財供給の設定において，プレーヤーの貢献費用に不完備情報が存在するときの繰り返しゲームの実験結果と 1 回のプレーの場合の実験結果を比較をしている．繰り返しゲームの設定においては，少なくとも 20 回のゲームが行われた後に，継続確率が 0.9 によってゲームが繰り返されるように設計されている．結果は，繰返しによって，1 回のゲームのケースよりも協調する割合が 29 % から 40 % に増えたが，完全に協力が実現すようになるということからは遠いものであった．このゲームにおいては，ゲームの設定が複雑であったことから，このような結果になったものと考えられ，通常の繰り返しゲームの想定において予測される協力が実現するというものは，より単純な環境において成立するものと考え

られる．

Engle-Warnick and Slonim (2001) は，ステージゲームが信頼ゲームである繰り返しゲームを，割引因子をゲームの確率的停止によって表現することによって実験を行っている．信頼ゲームとは，一方のプレーヤーが初期保有所得を持っており，それをもう一方のプレーヤーに渡すかどうかを決め，その所得を受け取ったプレーヤーはその所得だけでなく，追加的余剰を獲得する．そして，後者のプレーヤーがその全てかもしくはいくらかを元のプレーヤーに戻すかどうかという意思決定をするというものである．この論文の目的は，繰り返しゲームのおけるプレーヤーが実際に採用した戦略を調べることであり，観察された被験者の行動からプレーヤーの戦略を推測するという研究を行っている．特に，この論文では，被験者に対して取っている戦略を顕示させるような意思決定をさせることなく，被験者が取った行動から，戦略については有限オートマトンを仮定するなどの方法を用いて，行動から直接戦略を推測するという方法を取っているという特徴がある．結果として，0.8 の確率でステージゲームが終了するという設定において，トリガー戦略が採用されているということが観察された．

Duffy and Ochs(2004) は，囚人のジレンマゲームにおいて，被験者間の固定された組み合わせによって行われる繰り返しゲームと，確率的組み合わせによるゲームのプレーの比較を行っている．割引因子はゲームの確率的終了によって表現されており，ゲームの継続確率が 0.9 の下において，確率的組み合わせにおいては 6 %という低い割合の協力が観察されたのに対して，固定組み合わせの繰返しのケースには 55 %という高い割合の協力が観察された．確率的組み合わせのゲームにおいても協力が均衡として維持されるとの理論研究があるが，それに反する結果が得られたと考えられる．

Aoyagi and Fréchette (2004) は，不完全公的モニタリングのケースについて公的シグナルのノイズと協力の割合の関係を調べるために，囚人のジレンマにノイズを加えた繰り返しゲームの実験を行っている．不完全公的モニタリングとは，例えば，寡占市場において，各企業は数量競争をしているが，相手企業の実際に販売した数量を観察することはできず，相手企業の数量に対するノイズを持った指標である市場価格のみが観察できるような状況を意味するものである．理論モデルの予測においては，このような不完全公的モニタリングの状況においては，十分高い割引因子の下で，各プレーヤーが比較的単純な戦略である閾値戦略を取ることによって，協力の戦略を取ることが均衡となり，その均衡の最大利得は各プレーヤーの戦略に対する不完全モニタリングによるノイズが大きくなればなるほど減少することが示されている．ここで閾値戦略とは，観察できる公的シグナルの大きさがある閾値以上の時には協力を続け，もしその閾値より公的シグナルが下回ったなら，逸脱の戦略を取り罰則のフェーズとなるが，公的シグナルがまたその閾値を上回ったならば数期後には再び協力の戦略に戻るというものである．割引因子は 0.9 の確率的なステージゲームの終了によって表現されている．ノイズが低くなる

と協力する傾向にあるという理論での予測に合致した結果が実験において観察されたことが示されている．

そして，Dal Bó (2004) は，無限繰り返しゲームのインプリケーションである，囚人のジレンマゲームにおいて繰り返しのプレーが，プレーヤーが自分が逸脱の戦略を取ったならば将来相手プレーヤーから罰則として逸脱の戦略を取られるという恐れから協力を維持するという帰結を実験によって実証的に検証している．これまでの実験研究において，理論的には均衡として協力の戦略が取られることがない有限繰り返しゲームにおいても実際には被験者の間で協力の戦略が取られることが観察されており，その有限の繰り返し回数が増えれば協力する割合が増えることも知られている．従って，仮に無限繰り返しゲームの設定で実験を行って協力の行動が被験者に見られたとしても，それが上記のような無限繰り返しゲームの将来への認識からもたらされたものか，有限繰り返しゲームにおいても観察されるような効果によるものかを識別しなければ，無限繰り返しゲームの理論予測を正確に実証的に検証したことにはならない．そこで，この論文では，一方で，ゲームの確率的停止によって割引因子が表現された無限繰り返し囚人のジレンマの実験を行い，もう一方で，その確率的停止によって割引率が表現された無限繰り返しゲームの実験における期待プレー回数の有限繰り返し囚人のジレンマゲームの実験を行い，無限繰り返しゲームの結果を有限繰り返しゲームの結果でコントロールすることによって無限繰り返しゲームの理論予測の実証的検証を行っている．つまり，無限繰り返しゲームの設定の実験において割引因子の変化による協力割合の変化の結果から，有限繰り返しゲームによって観察される繰り返し期間の延長によって生じる協力割合の増加の効果をコントロールし取り除くことによって，無限繰り返しゲームの想定の下での純粋な割引率の変化の効果の分析を行った．その結果，割引因子の上昇による協力の増加の傾向は高いものであることが分かり，無限繰り返しゲームの理論研究での帰結を以前の研究よりも支持するような結果が得られた．

従って，無限繰り返しゲームの実証的な研究をするために実験を設計する際には，理論的に重要なパラメーターである割引因子は，ゲームの確率的終了によって表現することが可能である．しかし，理論的予測では観察されないが実験においては有限繰り返しゲームにおいて被験者の協力の戦略を取ることが観察される．そのため，今後，カルテル行動の記述として無限繰り返しゲームの設定を用いて実験研究を行う際には，この効果が存在することに留意しながら，その設計を行う必要があるものと思われる．

注

＊　独立行政法人経済産業研究所研究スタッフ，競争政策研究センター客員研究員

＊＊　跡見学園女子大学マネジメント学部専任講師，競争政策研究センター客員研究員．この研究に対して跡見学園女子大学から特別研究助成費の資金援助が与えられた．

(1) この $\boldsymbol{g}$ を $\mathcal{A}$ から $\mathbb{R}^n$ への写像として考える.

(2) ただし，ある集合 M に対して $\mathrm{co}\,M$ は M の凸包を表わす.

(3) Fudenberg and Levine (1983), Harris (1985), Börges (1989) を見よ.

(4) この時，利得集合はコンパクトなので利得関数はを上手く連続にする位相を見つければよい.

(5) Abreu, Pearce, and Stacchetti (1990, p.1051) を見よ．彼らの系ではなく普通に点列を用いての証明もできる.

(6) Abreu, Pearce, and Stacchetti (1990, p.1053) を見よ.

(7) ゲームでは Fudenberg and Levine (1983) が基本文献である.

参考文献

[1] Abreu, D., D. Pearce, and E. Stacchetti (1986), "Optimal Cartel Monitoring with Imperfect Information." *Journal of Economic Theory*, **39**, 251–269.

[2] Abreu, D., D. Pearce, and E. Stacchetti (1990), "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring." *Econometrica*, **58**, 1041–1063

[3] Aoyagi, M. and Fr?chette, G. (2004): "Collusion in Repeated Games with Imperfect Public Monitoring," mimeo.

[4] Dal Bó, Pedro (2004): "Cooperation Under the Shadow of the Future: Experimental Evidence from Infinitely Repeated Games," mimeo.

[5] Börges, T. (1989), "Perfect Equilibrium Histories of finite and Infinite Horizon Games." *Journal of Economic Theory*, **47**, 218–227.

[6] Duffy, J. and Ochs, J. (2004): "Cooperative Behavior and the Frequency of Social Interaction,"mimeo.

[7] Engle-Warnick, J. and Slonim, R.L. (2001): "Inferring Repeated-Game Strategies from Actions: Evidence from Trust Game Experiments,"mimeo.

[8] Feinberg, R.M. and Husted, T.A. (1993): "An Experimental Test of Discount-Rate Effects on Collusive Behaviour in Duopoly Markets," *Journal of Industrial Economics*, 41.2: 153-60.

[9] Fudenberg, D. and D. Levine (1983), "Subgame-perfect Equilibria of Finite and Infinite Horizon Games." *Journal of Economic Theory*, **31**, 251–268.

[10] Fudenberg, D. and D. Levine (1994), "Efficiency and Observability with Long-run and Short-run Players." *Journal of Economic Theory*, **62**, 103–135.

[11] Fudenberg, D., D. Levine, and E. Maskin (1994), "The Folk Theorem with Imperfect Public Information." *Econometrica*, **62**, 997–1039.

[12] Fudenberg, D. and J. Tirole (1992), *Game Theory.* MIT Press.

[13] Hardy, G. H. (1949), *Divergent Series*, Clarendon Press.

[14] Harris, C. (1985), " A Characterization of the Perfect Equilibria of Infinite Horizon Games." *Journal of Economic Theory*, **37**, 99–127.

[15] Halmos, P.R. (1950), *Measure Theory.* Springer-Verlag, New York.

[16] Hendon, E., H.J.Jacobsen, and B.Sloth (1996), "The One-Shot-Deviation Principle for Sequential Rationality." *Games and Economic Behavior*, **12**, 274–282.

[17] Lars Ljungquist and Thomas J. Sargent (2000): *Recursive Macroeconomic Theory*, MIT Press.

[18] Myerson, R.B. (1991), *Game Theory: Analysis Conflict.* Harvard University Press.

[19] Murnighan, J.K. and Roth A.E. (1983): "Expecting Continued Play in Prisoner's Dilemma Games," *Journal of Conflict Resolution*, 27.2: 279-300.

[20] Palfrey, T.R. and Rosenthal, H. (1994): "Repeated Play, Cooperation and Coordination: An Experimental Study," *Review of Economic Studies*, 61.3: 545-65.

[21] Radner, R., (1985), "Repeated Principal-Agent Games with Discounting," *Econometrica*, **53**, no.5, 1173–1198.

[22] Roth A.E. and Murnighan, J.K. (1978): "Equilibrium Behavior and Repeated Play of the Prisoner's Dilemma," *Journal of Mathematical Psychology*, 17: 189-97.

[23] Tanno, T. (2000), "The Equilibrium Characterization with Long-run and Short-run Players." mimeo.