繰り返しゲームにおけるサブゲーム完全トリガー戦略均衡と割引パラメーターの決定

前　野　冨士生

目　　次

1．はじめに

非協力繰り返しゲームの戦略で問題となるのは，プレイヤーが協力的な戦略（必ずしもナッシュ均衡戦略とは限らない）から逸脱した場合に，課せられる罰則（脅し）が信頼にたるものかどうかということにある。罰則が信頼できるものであることがすべてのプレイヤーに周知されておれば逸脱をしないのが得策であるが，信頼にたるものでないときは逸脱して短期の高利得を獲得するのが望ましい場合もある。たとえば価格競争を行っている２人のプレイヤーの囚人のジレンマゲームを考える。協力的な行動をすればジレンマ解に陥らないが，一方が逸脱して安い価格をつけるという戦略をとると，利得増加につながる場合は，そのプレイヤーは逸脱するかもしれない。しかしそこに信頼のある脅し（credible threat）があれば協力的な行動をとることになる。ただし１回限りのゲームであれば，お互いに安全な行動をとった結果，ジレンマ解となる。しかし繰り返しゲームになると，その構造の定式化は複雑になるが，プレイヤーにとっては協力的な行動をとる方が得策となる。本稿ではＪ．Ｗ．フリードマン[1]に従って繰り返しゲームで重要な役割を演じる割引パラメーターを検討し１つの試論を提示する。さらにこの割引パラメーターが大きくなると均衡解も多くなるが，フリードマンによれば，その多くの均衡解から極小数の均衡解を導出している。この点についてもコメントを行う。

2．戦略の定式化

以下の議論に必要な記号の定式化を行う。

$N=\{1, \cdots, n\}$ はプレイヤーの集合

S_i はプレイヤー i の戦略空間

$S=S_1 X \cdots X S_n$ は個別戦略空間の直積でゲームの戦略空間

$s_i \in S_i$ はプレイヤー i の戦略

$s=(s_1, \cdots, s_n)$ は戦略の組

$U_i : S \rightarrow R$，プレイヤー i の利得関数

$s \backslash t_i=(s_1, \cdots, s_{i-1}, t_i, s_{i+1}, \cdots, s_n)$ は s_i を t_i に替えたスカラー値

$\beta_i=\frac{1}{1+r_i}$，i の割引パラメーター，ただし r_i は割引率

$\beta=(\beta_1, \cdots \beta_n)$ すべての $i \in N$ に対して $\beta_i \in (0, 1]$

$h_t\ (s_0, s_1, \cdots, s_{t-1})$ ゲームの歴史

$h_t \in S^t = X_{\tau=0}^{t-1} S^{\tau}$

ただし時点０の歴史の集合は空であり S^0 とする。繰り返しゲームにおけるプレイヤー i の t 期の戦略は

$\rho_i^t(h_t) : S^t \rightarrow S_i$[2]　$i \in N$

と定義すると戦略の組 $\rho=(\rho^0, \rho^1, \cdots)$[3] は経路 $K(\rho)=(K_0(\rho), K_1(\rho) \cdots)$

を作る。この経路は

$$K_0(\rho)=\rho^0$$

$$K_t(\rho)=\rho^t(K_0(\rho),\ \cdots,\ K_{t-1}(\rho))$$

として決定される。この戦略の組と経路の関係を図1の簡単な戦略形を用いて述べる。

		プレイヤー2 s_{21}	s_{22}
プレイヤー1	s_{11}	2,2	0,5
	s_{12}	5,0	1,1

図 1

このゲームを2回繰り返す場合を考える。1回目は4通りの組み合わせがあり，2回ではそのおのおのに対して4通りあるから合計16通りの組み合わせがある。たとえば，一回目に（s_{11}, s_{21}）を選択し，2回目に（s_{12}, s_{22}）が選択されると，2回の戦略でできる経路は（(s_{11}, s_{21}), (s_{12}, s_{22})）である。従って戦略が決まると経路が確定できる。このゲームでも2期間で16通りの経路が考えられる[4)]。ゲームに参加するプレイヤーが2人以上で，それぞれのプレイヤーが2個以上の戦略をもち，T（T＝∞）期間を考えた場合，経路の定式化は複雑になる。そこで以下ではある経路から別の経路へ移る場合のルールを決め，少数の経路のみで戦略を表わす場合に限定する。このようにして作られた戦略の組を単純戦略の組と呼ぶ[5)]。記号により定義すると，

もし，$h_t=(K_0^0,\ \cdots K_{t-1}^0)$ なら，$\rho^t(h_t)=K_t^0$ で K^0が現行経路である。

もし，$s_t=\rho^t(h_t)=K_\tau^i$ なら，現行経路は変化せず，$s_{t+1}=\rho^{t+1}(h_t,\ K_\tau^i)=K_{\tau+1}^i$

となる。

一般に単純戦略の組は $n+1$ 個の経路によって特徴づけられる。第1の経路 K^0が始めにとられ，そこからの逸脱がおこるまで続く。残りの $K^1, \cdots K^n$ は罰則経路であり，K^i はプレイヤー i の罰則経路である。単純戦略の組の下でのルールは，つねに現行経路が存在し，もしどのプレイヤーも現行経路から逸脱しないなら，現行経路を永久にとりつづけることである。もしプレイヤー iが逸脱するなら，このとき，プレイヤー iの罰則経路が現行経路となり，その後 K^iからの逸脱がおこらない限り永久に K^iをとりつづける。K^iが現行経路で，もしプレイヤー iが逸脱するなら，罰則経路が現行経路となり，その経路の始めから出発することになる[6)]。

このように経路が戦略の組によって定義されるとすると，プレイヤー iの繰り返しゲームの利得関数は（2．1）式で与えられる。

$$\Pi_i(K)=\sum_{t=0}^{\infty}\beta_i^t U_i(K_t(\rho)) \qquad (2.\ 1)$$

さらに繰り返しゲームを $\Gamma=(N,\ S,\ U,\ \beta)$ とし，(N, S, U) は Γ の1期ゲームまたは段階ゲームとする。このとき次の定理が成立する[7)]。

定理1　割引をともなう繰り返しゲーム Γ で単純戦略の組 $\rho^k=(K^0,\ K^1,\ \cdots K^n)$ がサブゲーム完全均衡であるための必要十分条件は，すべての $s_i\in S_i$, $j\in\{0\}\cup N$ と $t\geq 0$, $i\in N$ に対して（2．2）式が成立することである。

$$\sum_{\tau=t}^{\infty}\beta_i^{\tau-t}U^i(K_\tau^i)\geq U_i(K_t^j\backslash s_i)+\beta_i\Pi_i(K^i) \qquad (2.\ 2)$$

（2．2）式の左辺は逸脱のない場合で，右辺第1項はプレイヤー iが1期目に逸脱し，第2項は罰則経路を表わす。ただし第2項は第1項と（2．1）式との関係で言えば

$$\frac{\beta^0 U_i}{1-\beta_i}-\beta^0 U_i=\frac{U_i-U_i(1-\beta_i)}{1-\beta_i}=\frac{\beta_i U_i}{1-\beta_i}$$

となり，$\Pi_i=U_i/1-\beta_i$ である。

罰則経路は $K^i=K^j$ で誰が逸脱しても同じであり，また K^0 はコンスタントな経路あるいは協力的な経路である。すなわち定理1はプレイヤー i による ρ_i^k からの逸脱はサブゲームで彼の利得を増加できない条件を示している。

3．トリガー戦略と簡単な例題

次に上述の単純戦略の組との関係でトリガー戦略について述べる。

今，s^*, s^cという戦略を考え，$U(s^*)>U(s^c)$

であるとき，プレイヤー i は時点１で戦略 s^* を選びその後もいづれかのプレイヤーがそれ以前に s^c を選んでいないかぎり，s^* を選びつづける。しかし，いずれかのプレイヤーがそれ以前に s^c を選んだときにはそれ以後は s^c を選びつづける。このようにトリガー戦略とはプレイヤーの誰かが協力しなくなるまでは協力し，誰かが協力しなくなったとき，非協力の引き金（トリガー）が引かれる[8)]。このことをたとえば図２の囚人のジレンマゲームで言うなら次のようになる。

		プレイヤー２	
		s_2^*	s_2^c
プレイヤー１	s_1^*	3,3	−2,5
	s_1^c	5,−2	−1,−1

図　2

このゲームを繰り返し行うとき，プレイヤー１，プレイヤー２とも s^* を選択すると２人とも３という利得を得つづけるが，一方が逸脱して $s_i^c(i=1,\ 2)$ を選択すると，逸脱したプレイヤーは，$(N,\ S,\ U)$ で５の利得を得ることになる。しかしそれ以後は両プレイヤーとも－１の利得となるのである。ちなみに $(s_1^c,\ s_2^c)$ は $(N,\ S,\ U)$ でのナッシュ均衡である。

このようにトリガー戦略とはプレイヤー間の暗黙の取り決めに従う戦略であり，トリガー戦略でのルールはすべてのプレイヤーが周知しているのである。このトリガー戦略を次のように２つに分けて考える。

〔A〕　グリム・トリガー戦略
$(s^*,\ s^c\cdots)$ あるいは $(s^*,\ s^c,\ \infty)$

〔B〕　有限復帰トリガー戦略
$(s^*, s^c, \cdots, s^*, \cdots)$ あるいは (s^*, s^c, T)

すなわち，〔A〕のグリムトリガー戦略では，初期に s_i^* を選択し，その後のすべての期間で s_i^* を選択する限り，s_i^* が選択されるが，s_i^* から逸脱すれば s_i^c が選択されつづける。これに対して〔B〕の有限復帰トリガー戦略は，初期に s_i^*，その後もすべての期間で s^* であるなら，s_i^* が選択されつづけるが，もし任意の時点 t で s^* から逸脱するなら，s_i^c が $t+1,\ \cdots t+T$ に対して選択され，$t+T+1$ においてプレイヤー i は s_i^* に復帰する。特に s^c が選択されているときに，ある時点 t において，s^c から逸脱すると，罰則のサイクルが再び始まるとする。

以下でトリガー戦略について興味ある例を示す[9)]。

５人の貸別荘のオーナーがいて，s_{it} をシーズン t に所有者 i によって選択される賃貸料とする。各オーナーの利得関数は（３．１）式とする。

$$U_i(s_t)=180s_{it}-6s_{it}^2+s_{it}\sum_{j=1}^{5}s_{jt}\qquad i=1,\ \cdots 5 \tag{3. 1}$$

各夏期が別々の非協力ゲームとみなされるとき，（３．１）式より

$$\partial U_i / \partial s_{it}=0$$

これより，$s_{it}=30$ となり，各オーナー i の利得は4500となる。

各夏期にオーナーが協力的行動をとるトリガー戦略を用いた場合，$s_{it}=s_{jt}$ となるから（３．１）式は

$$U_i(s)=180s-6s^2+s(5s)$$

$$\frac{dU_i}{ds}=0$$

で $s_{it}=90$ であり，利得は8100となる。ただし，ある期 t で $s=90$ 以外の選択がされると，プレイヤー i は，$s_{i,t+1}=30$ を了解しているものとする。

$s_{it}=30$ の賃貸料ではその別荘地は安いが故に高い占有率となり，$s_{it}=90$ の賃貸料は割高を感じて占有率は低くなると考えられる。

またトリガー戦略では，あるオーナーは他のオーナーの高い価格を利用して，自分の１期間の高い利益を得ることが可能となる。他のオーナーが $s_{it}=90$ のとき，オーカー i は次のように利得関数を利用する。すなわち，

$$U_i(s_t)=180s_{it}-6s_{it}^2+s_{it}(4\times 90)$$

$$\frac{du_i}{ds_{it}}=0$$

これより，$s_{it}=54$であり，利得は14580となる。

このように1期間の利得のみを考えて逸脱すれば6480（＝14580－8100）の余分の利得を得ることができる。ここで，オーナー i を除くオーナーがトリガー戦略を用いているとき，オーナー i の戦略を考えてみる。もしオーナー i もトリガー戦略を用いるなら，毎期8100の利得を受けとり，割引パラメーターを考慮すれば，無限回の繰り返しゲームでの利得は$8100/(1-\beta_i)$となる。しかし，もしオーナーが第1期に逸脱するなら，1期に14580を獲得し，その後は $s_{it}=30$となり利得は4500である。この場合，割引された利得は

$$14580+4500\beta_i/(1-\beta_i)$$

である。これより

$$8100/1-\beta_i \gtreqless 14580+4500\beta_i/1-\beta_i$$

で，$\beta_i>\frac{9}{14}$なら，トリガー戦略が有利であり，$\beta_i<9/14$なら，逸脱する方が有利であり，$\beta_i=\frac{9}{14}$なら，両者は無差別であることがわかる。

このようにトリガー戦略で重要な役割を果たすのは，信頼できる脅しと割引パラメーターである。したがってトリガー戦略均衡は，自己強制的協定（self-entorcing agreements）と考えられる。この脅しという強制メカニズムがあることで，各プレイヤーは逸脱すると利得が小さくなり，非協力均衡戦略をプレイすることになる。しかし，例題でもみたように，割引パラメーターの値によっては逸脱した方が得になる場合もある。そこで次に，割引パラメーターについて検討する。

4．割引パラメーターとトリガー戦略

まず，グリムトリガー戦略均衡が存在する場合を検討する。ここで

$$r_i(s)=max_{s_i'\in s_i}U_i(s\backslash s_i') \qquad (4.\ 1)$$

と定義する。すなわち，$r_i(s)$ は (N, S, U) の s に関する最適反応利得であると定義する。このとき，次の定理が成立する[10]。

定理2　Γ を繰り返しゲームとする。$s^c\in S$ を (M, S, U) の均衡点であるとし，$(s^*, s^c)\in S\times S$をグリムトリガー戦略の組とする。

$$\beta_i\geq\frac{r_i(s^*)-U_i(s^*)}{r_i(s^*)-U_i(s^c)}\quad i\in N \qquad (4.\ 2)$$

は (s^*, s^c) が Γ のサブゲーム完全均衡点であるための必要十分条件である。

証明　（2，2）式は

$$\frac{U_i(s^*)}{1-\beta_i}\geq r_i(s^*)+\frac{\beta_iU_i(s^c)}{1-\beta_i} \qquad (4.\ 3)$$

であり，（4．3）式は（4．2）式と同値である。

証明終り。

もし，$U_i(s^*)\leq U_i(s^c)$であれば $\beta_i\in(0, 1)$ に対して，（4．3）式は成立しない。

証明

$U_i(s^*)\leq U_i(s^c)$の両辺にマイナスをかける

$$-U_i(s^*)\geq -U_i(s^c)$$

両辺に$r_i(s^*)$を加える。

$$r_i(s^*)-U_i(s^*)\geq r_i(s^*)-U_i(s^c)$$

両辺を $r_i(s^*)-U_i(s^c)>0$ でわる

$$\frac{r_i(s^*)-U_i(s^*)}{r_i(s^*)-U_i(s^c)}\geq\frac{r_i(s^*)-U_i(s^*)}{r_i(s^*)-U_i(s^c)}$$

$$\therefore\quad \beta_i\geq\frac{r_i(s^*)-U_i(s^*)}{r_i(s^*)-U_i(s^c)}\geq 1$$

となり成立しない　証明終り。

系1　定理2の条件の下で，もし $U(s^*)>U(s^c)$であれば，$\beta_i\in(0, 1)$，$i\in N$が存在して，(s^*, s^c) はトリガー戦略非協力均衡である。

証明

$r_i(s^*)\geq U_i(s^*)>U_i(s^c)$より

$$\beta_i^* = \frac{r_i(s^*) - U_i(s^*)}{r_i(s^*) - U_i(s^c)} \in (0,\ 1),\ i \in N \quad (4.\ 4)$$

となる。$\beta_i \in [\beta_i^*,\ 1)$，$i \in N$ を選ぶことで（4．2）が成立する。　証明終り。

系 2，定理 2 の下で，$U(s^*) > U(s^c)$ であれば割引パラメーター $\beta \in (0,\ 1)^n$ と有限期間 T（$=\infty$）が存在して有限復帰トリガー戦略，$(s^*,\ s^c,\ T)$ はサブゲーム完全均衡である。

証明：

$\beta_i \in (\beta_i^*,\ 1)$ をすべてのプレイヤーに対して選択する。均衡では次式が成立する。

$$\frac{U_i(s^*)}{1-\beta_i} \geq r_i(s^*) + \frac{\beta_i U_i(s^c)(1-\beta_i^T)}{1-\beta_i} + \frac{\beta_i^{T+1} U_i(s^*)}{1-\beta_i} \quad (4.\ 5)$$

（4．5）は（4．6）と同値である[11]。

$$\beta_i \geq \frac{r_i(s^*) - U_i(s^*) + \beta_i^{T+1}(U_i(s^*) - U_i(s^c))}{r_i(s^*) - U_i(s^c)} \quad (4.\ 6)$$

（4．5）の右辺第 1 項は逸脱による利得，第 2 項は 2 期から T 期までの罰則経路上での割引利得を表わす。これは次のようになる。2 期から T 期までの割引利得を S_n とすると，

$$S_n = \beta_i U_i(s^c) + \beta^2 U_i(s^c) + \cdots + \beta^T U_i(s^c)$$

$$\beta S_n = \beta_i^2 U_i(s^c) + \beta_i^3 U_i(s^c) + \cdots + \beta_i^{t+1} U_i(s^c)$$

$$\therefore\ S_n(\beta - 1) = \beta^{T+1} U_i(s^c) - \beta U_i(s^c)$$

$$\therefore\ S_n = \frac{\beta^{T+1} U_i(s^c) - \beta U_i(s^c)}{1-\beta}$$

第 3 項は復帰後の合計割引利得を表わし，

$$\beta_i^{T+1} U_i(s^*) + \beta^{T+2} U_i(s^*) + \cdots$$

である。

左辺は逸脱がない場合の割引利得である。

証明終り。

ところで，（4．2）式と（4．6）式の β_i の値を比較すると明らかなように，$U_i(s^*) \geq U_i(s^c)$ が成立しておれば，有限復帰トリガー戦略の β_i の値が大きくなっている。このことは，有限復帰トリガー戦略均衡の割引パラメーターが，グリムトリガー戦略均衡のそれに比べて制約がきついことになる。しかし有限復帰時点の T を T^* とすれば，T^* が十分大きくなれば（4．2）式と（4．6）式の β_i は等しくなる。したがって T^* をどこで決めるかが問題となる。

5．β_i 再論

これまでは，フリードマンによる 1 期目の逸脱を問題にしたグリムトリガー戦略と有限復帰トリガー戦略の場合の β_i の決定であるが，以下では〔A′〕1 期目に逸脱するのではなく，T 期に逸脱して T＋1 期以後罰則経路には入るグリムトリガー戦略の場合，〔B′〕1 期目に逸脱するのではなく，T 期に逸脱して T＋1 期から T′ まで罰則経路で T′＋1 から復帰する有限復帰トリガー戦略の場合を検討する。図 3 参照。

〔A′〕，〔B′〕の場合の β_i の決定を試みる

〔A′〕のケース

$(s^*,\ s^c)$ がサブゲーム完全均衡となる $\beta_{it} \in (\beta_i^*,\ 1)$ の条件は（4．3）式より

$$\frac{U_i(s^*)}{1-\beta_i} \geq \frac{1-\beta_i^T}{1-\beta_i} U_i(s^*) + \beta_i^T r_i(s^*) + \frac{U_i(s^c)\,\beta_i^{T+1}}{1-\beta_i} \quad (5.\ 1)$$

である。

すなわち，（5．1）式の右辺第 1 項は 0 ～ T－1 期まで逸脱なしの戦略 s^* をとった利得であり，第 2 項は T 期で逸脱によって得た利得であり，第 3 項は T＋1 で罰則経路に入った利得である。ただし 0 ～T－1 期までの合計利得 S は

$$S = (1 + \beta_i + \cdots + \beta_i^{T-1})\,U_i(s^*)$$

$$\beta_i S = (\beta_i + \beta_i^2 + \cdots + \beta_i^{T})\,U_i(s^*)$$

$$\therefore\ S = \frac{1-\beta_i^T}{1-\beta_i} U_i(s^*)$$

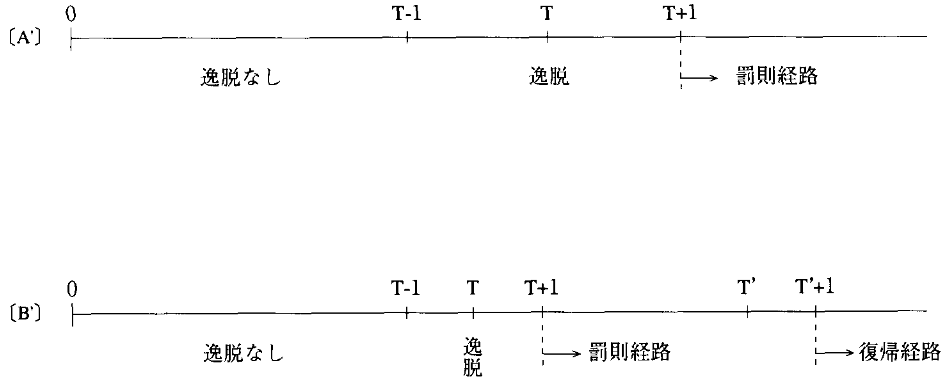

図3

（5．1）式の両辺に $1-\beta_i$ をかけて整理すると，

$$U_i(s^*) \geq U_i(s^*) - \beta_i^T U_i(s^*) + \beta_i^T r_i(s^*) - \beta_i^{T+1} r_i(s^*) + U_i(s^c)\,\beta_i^{T+1}$$

$$\therefore\quad \beta_i^{T+1}(r_i(s^*) - U_i(s^c)) \geq (r_i(s^*) - U_i(s^*))\,\beta_i^T$$

両辺を β_i^T で割ると

$$\beta_i \geq \frac{r_i(s^*) - U_i(s^*)}{r_i(s^*) - U_i(s^c)}$$

この式は1期目に逸脱して，その後グリムトリガー戦略のケースである（4．2）式と同じ結果である。したがって，次の命題が成立する。

命題1　グリムトリガー均衡戦略は，1期目に逸脱して，その後罰則経路を続ける場合と，T－1期までは逸脱せず，T期に逸脱し，その後罰則経路に入る場合，β_i の範囲は同一である。

〔B'〕のケース

有限復帰均衡戦略を決める β_i の条件は（4．5）式より，（5．2）式を満たす β_i である。

$$\frac{U_i(s^*)}{1-\beta_i} \geq \frac{1-\beta_i^T}{1-\beta_i} U_i(s^*) + \beta_i^T r_i(s^*) + \frac{\beta_i^{T+1} U_i(s^c)\,(1-\beta^{T-T})}{1-\beta_i} + \frac{\beta_i^{T+1}}{1-\beta_i} U_i(s^*) \qquad (5.\ 2)$$

（5．2）式の右辺第1項と第2項は（5．1）式の右辺第1項と第2項に同じである。第3項は罰則経路がT＋1～T'期までとなるので，その合計利得は

$$S = (\beta_i^{T+1} + \beta_i^{T+2} + \cdots + \beta_i^{T'})\,U_i(s^c)$$

$$\beta_i S = (\beta_i^{T+2} + \beta_i^{T+3} + \cdots + \beta_i^{T'+1})\,U_i(s^c)$$

$$\therefore\quad S = \frac{(\beta_i^{T'+1} - \beta_i^{T+1})\,U_i(s^c)}{\beta_i - 1}$$

第4項はT'＋1で復帰する利得の合計である。

（5．2）の両辺に（$1-\beta_i$）をかけて整理すると，

$$U_i(s^*) \geq U_i(s^*) - \beta_i^T U_i(s^*) + \beta_i^T r_i(s^*) - \beta^{T+1} r_i(s^*) + \beta_i^{T+1} U_i(s^c) - \beta_i^{T'+1} U_i(s^c) + \beta_i^{T'+1} U_i(s^*)$$

$$\therefore\quad \beta^{T+1}(r_i(s^*) - U_i(s^c)) \geq \beta_i^T(r_i(s^*) - U_i(s^*)) + \beta^{T'+1}(U_i(s^*) - U_i(s^c))$$

両辺を β_i^T で割ると

$$\beta_i \geq \frac{r_i(s^*) - U_i(s^*) + \beta_i^{T'-T+1}(U_i(s^*) - U_i(s^c))}{r_i(s^*) - U_i(s^c)} \qquad (5.\ 3)$$

（4．6）式と（5．3）式の違いは割引パラメーターの期間の差である。従って次の命題が成立する。

命題2　有限復帰トリガー均衡戦略，〔B〕と〔B'〕の β_i の範囲は異なる。T＋1＞T'－T＋1より，$0<\beta_i<1$ であるから，T期に逸脱する β_i の値が大きくなる。すなわち，制約は後者〔B'〕がきつくなる。

6．小数均衡選択についてのコメント

以上のように，トリガー戦略均衡における β_i を満たす範囲の（s^*，s^c）には多くの均衡戦略が存在する。次にその均衡から小数の均衡を選び出す方法を検討する[12]。

図4は2人のプレイヤーの例で，s^* と s^c の関係を示している。

$U(s^c)$ の右上方で $R(\Gamma)$ の領域における任意の利得は適切な β_i が与えられたときの $U(s^*)$ をあらわす。したがって β_i の値に依存して数多くの潜在的均衡が存在する。

$\psi(\Gamma)$ を Γ のトリガー戦略（s^*，s^c）の集合とする。（s^*，s^c）$\in \psi(\Gamma)$ は s^c が（N, S, U）の均衡で，（4．2）式を満足することである。すなわち，β_i に関しては，グリムトリガー均衡戦略を満足している。今，$\psi(\Gamma)$ の局所的に効率的な要素を $\psi^*(\Gamma)$ で表わすと，$\varphi^*(\Gamma)$ は支配されない均衡を表わす[13]。式で示すと，

$$\psi^*(\Gamma) = \{(s^*,\ s^c) \in \varphi(\Gamma) \mid U' > U(s^*) \text{ は } U' \notin R(\Gamma) \text{ である}\}$$

局所効率的利得ベクトル $\{U(s^*) \in R(\Gamma) \mid s^*,\ s^c) \in \psi^*(\Gamma)\}$ は繰り返しゲームの効率的フロンティアの下方に位置する場合がある。図4の $U(s^*)$ は局所的効率であるが大域的効率性を必ずしも意味しない[14]。

もし（s^*，s^c）$\in \psi(\Gamma)$ で $U(s^*)$ が利得可能フンロンティア上にあれば，利得可能フロンティアが凹であるときは，繰り返しゲームの利得はパレート最適である。この場合，$U(s^*)$ は $\psi(\Gamma)$ の凸包の境界上にあることである（図4の点線部分も含める）。

このように繰り返しゲームでは，協力的な戦略となる可能性が大きくなり凸包を作ることが可能となる。このとき $U(s^*)$ が凸包の境界にあれば $U(s^*)$ は大域的に効率的である。ただし凸包の境界上に $U(s^*)$ がくる均衡戦略に限定すると支配されない戦略ではあるが，均衡点の数は少くなる。

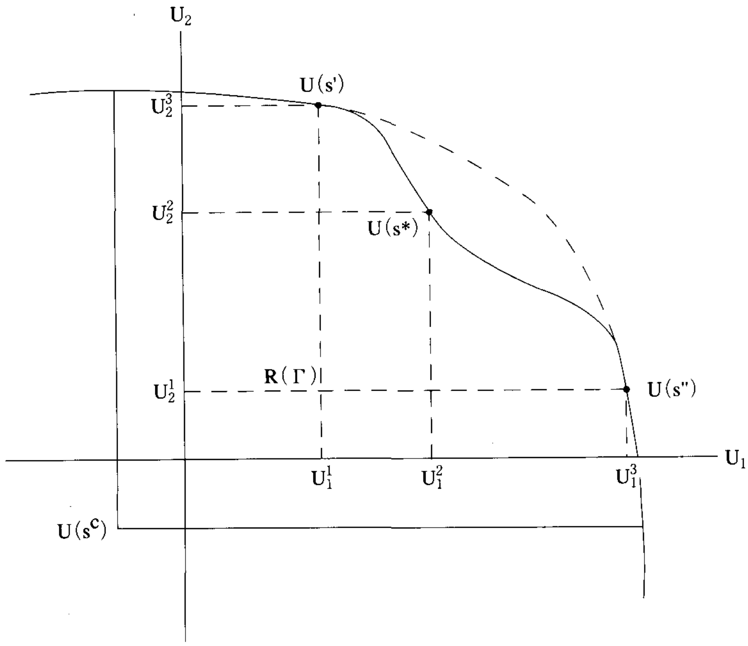

図4

このように大域的に効率的なトリガー戦略に限定しても割引パラメーター β_i が1に近いなら，多くの均衡点が存在することは明らかである。この多くの均衡点から極小数の均衡点を取り出すことは，フリードマン（1971）で試みられ，その基本的な考え方はバランスのとれた誘惑均衡の理論である。フリードマンによれば，（4．4）式の β^*_i がすべての i に対して同じ値をもつ s^* を選択することとしている。

式で表わせば

$$\frac{r_i(s^*)-U_i(s^*)}{U_i(s^*)-U_i(s^c)}=\frac{r_j(s^*)-U_j(s^*)}{U_j(s^*)-U_j(s^c)} \qquad (6.\ 1)$$

（6．1）式の分子はトリガー戦略からの逸脱による1期間の利得であり，分母は逸脱による罰則からの期間当り利得である。この比は従って，逸脱しょうとする誘惑（temptation）を測ると考えられ，この比が大きければ誘惑が大きいことを示している。均等な誘惑均衡（balanced temptation equilibrium）ではすべてのプレイヤーの誘惑は等しくなる[15]。

ところで，$\beta_i \in [\beta^*_i,\ 1)$ より $\beta^*_i=\beta_i$ と考えると

$$\frac{1}{\beta_i}=\frac{r_i(s^*)-U_i(s^c)}{r_i(s^*)-U_i(s^*)}$$

$$\therefore\ \frac{1}{\beta_i}-1$$

$$=\frac{r_i(s^*)-U_i(s^c)-(r_i(s^*)-U_i(s^*))}{r_i(s^*)-U_i(s^*)}$$

$$\frac{1}{\frac{1}{\beta_i}-1}=\frac{\beta_i}{1-\beta_i}=\frac{r_i(s^*)-U_i(s^*)}{U_i(s^*)-U_i(s^c)} \qquad (6.\ 2)$$

すなわち（6．1）式は（6．2）式で表わすことができる。（6．2）式で均衡条件を表わす（4．3）式の右辺に代入すると，

$$r_i(s^*)+\frac{r_i(s^*)-U_i(s^*)}{U_i(s^*)-U_i(s^c)}\,U_i(s^c)$$

$$=\frac{U_i(s^*)\,\{r_i(s^*)-U_i(s^c)\}}{U_i(s^*)-U_i(s^c)} \qquad (6.\ 3)$$

他方，

$$1-\beta_i=1-\frac{r_i(s^*)-U_i(s^c)}{r_i(s^*)-U_i(s^c)}$$

$$=\frac{U_i(s^*)-U_i(s^c)}{r_i(s^*)-U_i(s^c)}$$

であるから，（4．3）の左辺は

$$\frac{U_i(s^*)}{1-\beta_i}=\frac{\{r_i(s^*)-U_i(s^c)\}\,U_i(s^*)}{U_i(s^*)-U_i(s^c)} \qquad (6.\ 4)$$

となり（6．3），（6．4）式から，（4．3）式の両辺は等しくなる。つまり（4．4）式の β^*_i に対して（6．1）式が成立すれば，逸脱しても有利にならず，特定の β^*_i をみたすという意味では均衡は小数にしぼられる。ところで（6．2）式より，β_i が大きいほど，逸脱の魅力は大きくなることを示しながら，さらに大きい $\beta_i>\beta^*_i$ をもってくることで（4．2）式が成立する。すなわち（4．2）式は逸脱しない β_i の範囲を示している。このことは指摘すべき1つの問題といわねばならない。

7．むすび

本稿では，J．W．フリードマンに従ってグリムトリガー戦略と有限復帰トリガー戦略について検討した。その際，フリードマンは，第1期に逸脱した場合を検討しているが，本稿ではT期に逸脱した場合を考え，それぞれのトリガー戦略での命題を導出した。そこで明らかになったことは割引パラメーター β_i の値が，グリムトリガー戦略では1期に逸脱した場合とT期に逸脱した場合は同じになるが，有限復帰トリガー戦略では，β_i の値が異なることが示された。

さらに，β_i の制約範囲はグリムトリガー戦略の方が大きいことが示された。β_i の値が大きくなるほど均衡戦略の数も多くなるが，この多数の均衡戦略から極小数の均衡を導出する試みをフリードマンはおこなっている。この点についても1つの問題点を指摘した。

注

1）Friedman. J.W ［5］

2）便宜のため，ρ_i^0 は S^0から S_iへの関数であるとすると，ρ_i^0 は S_iの任意の要素となる。

3）
$$\rho = \begin{pmatrix} \rho_1 \\ \cdot \\ \cdot \\ \cdot \\ \rho_n \end{pmatrix} = \begin{pmatrix} \rho_1^0 \cdots \rho_1^t \cdots \\ \cdot \quad\quad \cdot \\ \cdot \quad\quad \cdot \\ \cdot \quad\quad \cdot \\ \rho_n^0 \cdots \rho_n^t \cdots \end{pmatrix}$$
$$(\rho^0 \cdots \rho^t \cdots)$$

4）村田［8］

5）Friedman ［5］ p. 121

6）複数のプレイヤーが同時に逸脱する場合，min (i, j, k) が罰則経路となる。すなわち番号の若いプレイヤーの経路が罰則経路となるが，プレイヤー間の共謀はないので，この問題は重要とはならない。

7）Friedman ［5］。証明はp. 122参照

8）Gibbons, R. ［6］，邦訳　p. 95，奥野，銘村著［7］p. 240

9）Friedman ［5］ pp. 125-126

10) *Ibid.*, p. 127

11) （4．5）式より
$U_i(s^*) \geq r_i(s^*)(1-\beta_i)+\beta_i U_i(s^c)(1-\beta_i^T)+\beta_i^{T+1}U_i(s^*)$ より
$\beta_i(r_i(s^*)-U_i(s^c) \geq r_i(s^*)-U_i(s^*)+\beta_i^{t+1}(U_i(s^*)-U_i(s^c))$
より，（4．6）となる。

12) Friedman ［3］，Aumann (1959，1961) は非協力繰り返しゲームでの協力的な結果について最初の体系的研究がみられる。

13) 多くの均衡のうち，極小数の均衡を選ぶのは，支配される均衡を除くことが必要である。

14) Friedman ［5］ 図4で奇数番目に戦略 s' を偶数番目に戦略 s'' を選択することで利得 $U(s^*)$ より増加させることを示している。図4で $U(s')$は，プレイヤー I は U_1^1 の利得を，プレイヤー2は U_2^2 の利得を得るものとする。$U(s'')$も同様である。(s', s'') の選択はトリガー戦略であり，このときプレイヤーが (s', s'') の戦略にそむくなら，s^c になるという条件で。

15) この均衡の存在の正確な条件はFriedman［3］，［4］8章

参考文献

〔1〕Aumann, R,J. 1959. "Acceptable points in General Cooperalive n-person Games," In A, W, Tucker and R.D. Luce, ed., *Contributions to the Theory of Games,* vol. IV. Princeton Uncversity press.

〔2〕――. 1961. "The Core of a Cooperative Game without Side Payments," *Transactions of the American Mathematics Society* 98 : 539-52.

〔3〕Friendman, J,W. 1971. "A Non-Cooperative Equilibrium for Supergames," *Review of Economic Studies* 38 : 1-12.

〔4〕――. 1977. *Oligopoly and Theory of Games.* Amsterdam : North-Holland.

〔5〕――. 1990. *Game Theory with Applications to Economics.*, Oxford University Press.

〔6〕Gibbons, R. *Game Theory for Applied Economists.* Princeton University Press, 1992.
木村憲二訳『応用経済学のためのゲーム理論入門』マグロウヒル，1992年。

〔7〕奥野正寛，鈴村興太郎『モダン・エコノミックス 2　ミクロ経済学Ⅱ』岩波書店，1988年。

〔8〕村田省三『経済のゲーム分析』牧野書店，1992年。

〔付記〕

本稿は，1995年度阪南大学産業経済研究所共同研究「ゲーム理論と統計」の成果報告の一部である。

(1995年12月5日受理)