

協力の条件 ウェブ付録 目次

第6章 ウェブ付録

付録	レプリケータ方程式の導き方	2
----	---------------	---

第8章 ウェブ付録

付録1	N 人囚人のジレンマにおける同類性志向の一般的定式化	7
付録2	図8.2においてC選択が合理的であるための 確率 α と確率 β の条件	9
付録3	オストラム実験における個人レベルの合理的な選択と そのジレンマ性	12

第9章 ウェブ付録

付録	命題9.1 警告—謝罪 TFT 戦略がナッシュ均衡 であることの証明	14
----	---------------------------------------	----

第10章 ウェブ付録

付録1	図10.1のシグナリングゲームにおいて均衡をなす 戦略の組を導き出す方法	18
付録2	Aの戦略を離散解に限定することについて	19
付録3	シグナリングゲームとしての壟断戦における均衡の求め方	20
付録4	分離均衡と一括均衡	25

第11章 ウェブ付録

付録	ナッシュ均衡と部分ゲーム完全均衡	26
----	------------------	----

第6章 ウェブ付録

付録 レプリケータ方程式の導き方

レプリケータ方程式の導き出し方は、微分方程式を用いて説明されることが多い。しかし、じつはもっと簡単に導き出すことができる。ここでは、大浦(2008: 90-95)を参考にしながら、よりわかりやすく説明しよう。

まず、戦略を A (アリス), B (ボブ), C (キャサリン) の3つとする。

集団におけるそれぞれの個体数を, h_A, h_B, h_C とする。総個体数を h_T とする。

集団におけるそれぞれの比率を, x_A, x_B, x_C とする。

集団におけるそれぞれの戦略の「平均利得」を U_A, U_B, U_C とする。また、集団全体の平均利得を $\bar{U}$ とする。

導き出したいことは、このとき、次のレプリケータ方程式が成立することである。

$$\begin{aligned}dx_A/dt &= (U_A - \bar{U})x_A \\dx_B/dt &= (U_B - \bar{U})x_B \\dx_C/dt &= (U_C - \bar{U})x_C\end{aligned}\tag{6.15}$$

以下、どのような前提と論理でこの方程式が導かれるかを説明する。

■ 粗再生産率と純再生産率

戦略の増え方は、まず、粗再生産率で規定されている。粗再生産率とは、期初の1個の個体が、ある一定時間たった期末において、平均的に何個の個体が増えているかの倍数を表す。たとえば、一定時間ののち、個体数がもとの個体数に対して2倍だと粗再生産率は2。個体数が半分になると、粗再生産率は0.5である。粗再生産率は常に ≥ 0 と仮定する。

なお、時間の単位としては、この「一定時間」を1とする。

この、一定時間におけるそれぞれの戦略の粗再生産率を, w_A, w_B, w_C とする。

粗再生産率から1を引いたもの、すなわち「粗再生産率-1」を純再生産率

と呼び、記号 u で表す。すなわち、

$$u_A = w_A - 1$$

$$u_B = w_B - 1$$

$$u_C = w_C - 1$$

純再生産率は、一定時間が経過したのち、次の期首では前期のもとになった個体は退出するとして、個体数の変化の倍率を示す指標になっている。たとえば、粗再生産率が2のとき、その一定時間において純粹に増加した個体数の比率は1である。

なお、以上の諸変数のうち、個体数と比率は時間 t の関数とみなされる。つまり、時間の経過とともに変化する。その変化のしかたがレプリケータ・ダイナミクスである。

■ 各戦略における個体数の変化のしかた

ここで、「一定時間」での個体数の変化を考える。たとえば戦略アリスは、もとの個体数が h_A 。その個体数に対して粗再生産率が w_A 。したがって、一定時間の経過において個体数は $w_A \times h_A$ に変化する。他方、同じ時間幅において旧世代は退出する。したがって、

$$\text{一定時間における個体数の変化} = w_A \times h_A - h_A \quad (6.16)$$

である。

次に、ある微少な時間幅 Δt を考え、その時間幅における個体数の変化を Δh_A で表す（ Δ はデルタと読む）。したがって、

$$\Delta h_A = (w_A \times h_A - h_A) \Delta t \quad (6.17)$$

となる。

つまり、ある時間幅 Δt における個体数の変化 Δh_A は、一定時間における個体数の変化数をもとに、その時間幅に比例すると考えられる。

上の式は、

$$\Delta h_A = (w_A - 1) h_A \Delta t \quad (6.18)$$

と書ける。ここで、純再生産率を用いると、

$$\Delta h_A = u_A h_A \Delta t$$

この両辺を Δt で割って、

$$\frac{\Delta h_A}{\Delta t} = u_A h_A$$

この式は、 Δt がどんなに小さくても成立する。したがって、極限をとれば、

$$\frac{dh_A}{dt} = u_A h_A \quad (6.19a)$$

という微分の式が得られる。他の戦略に関しても同じ式が成立するので

$$\frac{dh_B}{dt} = u_B h_B \quad (6.19b)$$

$$\frac{dh_C}{dt} = u_C h_C \quad (6.19c)$$

となる。

■ レプリケータ方程式の導出

以上の準備作業を踏まえ、時間に伴う戦略比率の変化率を表す式を導出する。

まず、それぞれの戦略の個体数は、総個体数 h_T に比率を乗じたものだから、たとえばアリスについては

$$h_A = x_A h_T \quad (6.20)$$

という関係が成立している。ここで、 $h_T = h_A + h_B + h_C$ だから、

$$h_A = x_A (h_A + h_B + h_C) \quad (6.21)$$

この式の右辺と左辺それぞれについて、時間 t の導関数を求める。積の微分を適用すると、結果は、

$$dh_A/dt = dx_A/dt (h_A + h_B + h_C) + x_A (dh_A/dt + dh_B/dt + dh_C/dt)$$

ここで右辺第2項を移項し、かつ $h_T = h_A + h_B + h_C$ を考慮すれば、

$$dx_A/dt(h_T) = dh_A/dt - x_A(dh_A/dt + dh_B/dt + dh_C/dt)$$

この両辺を全個体数 h_T で割り、(6.19) の各式を用いると、

$$dx_A/dt = u_A x_A - x_A(u_A x_A + u_B x_B + u_C x_C) \quad (6.22)$$

となる。

この右辺第2項の () 内を、

$$\bar{u} = u_A x_A + u_B x_B + u_C x_C \quad (6.23)$$

とおく。 $\bar{u}$ は、純再生産率 u_A , u_B , および u_C のウェイトづけ平均値である。そうすると、(6.22) 式は、次のようになる。

$$\begin{aligned} dx_A/dt &= u_A x_A - \bar{u} x_A \\ &= (u_A - \bar{u}) x_A \end{aligned} \quad (6.24)$$

ここで、純再生産率と期待利得についての基本的な仮定を導入する。すなわち、再生産率（進化適応度）は相互作用において獲得する平均利得の大きさに比例するというもとの前提に基づいて、「それぞれの純再生産率は、戦略の期待利得（平均利得）と一次関数の関係にある」とするのである。

よって、各戦略の期待利得を大文字の U_A , U_B , U_C とすれば、各純再生産率は、

$$\begin{aligned} u_A &= aU_A + b \\ u_B &= aU_B + b \\ u_C &= aU_C + b \end{aligned}$$

となっている。ただし、 $a > 0$ で、 b は何らかの定数。また、集団全体の平均利得を $\bar{U}$ とすれば、(6.23) 式は、

$$\bar{u} = a\bar{U} + b$$

となる。（なお、この一次式の仮定について注釈しておこう。純再生産率は、定義より最小値が -1 で、最大値は不特定である。他方、平均利得はさまざまなモデルや現実があるところから、数値にとくに制限はない。モデルによっては、正の値だけのこともあ

るし、マイナスの値をとることもある。ここでは、平均利得の尺度がどうあれ、適当な係数 a と b をもってくれば、上のような対応が成立しているということだけが仮定されているのである。)

さて、こうした純再生産率と期待利得との関係式を (6.24) 式に代入すれば、定数 b が消えて、次式をうる。

$$dx_A/dt = a(U_A - \bar{U})x_A \quad (6.25)$$

他の戦略についても同様に導くことができる。これが求めるべきレプリケータ方程式の一手手前の式である。

最後に、比例定数 a は具体的な時間と再生産率を想定するときは重要だが、抽象的に「各戦略比率の時間的な変化の傾向」を見る場合には、 $a=1$ と想定してもかまわない。よって、結論として、冒頭のレプリケータ方程式 ((6.15) 式) が導かれる。

第8章 ウェブ付録

付録1 N 人囚人のジレンマにおける同類性志向の一般的定式化

ここでは、 N 人囚人のジレンマという状況のなかで、相互作用における同類性志向がどのように組み込まれるか、そしてその際の選択性指標などはどのように定式化できるかを説明する。

まず、集団全体はそれぞれ N 人からなる小集団に分かれていて、各小集団のなかで、人びとは本文 (8.4) 式で表される N 人囚人のジレンマをプレイするとする。また、人びとは C (協力) を選択する C タイプと D (非協力) を選択する D タイプとに分かれていると想定する。集団全体における C タイプの個体の比率を p とする。そのうえで、小集団は C タイプの比率の異なる2つのグループに分かれていて、一方のグループにおいては N 人のなかにおける C タイプの比率は p_1 、他方のグループでは p_2 とする。ただし、 $p_2 < p < p_1$ とする。この差が、ランダムとは異なる同類性志向を可能にする。

ここで、全体比率 p と、 p_1 および p_2 を所与とすると、それぞれの比率をもつ N 人小集団の割合は次のように自動的に定まっている。すなわち、今、全小集団のなかで p_1 の比率をもつ小集団の割合を q とする (たとえば、小集団の数が10で、そのうち3つが比率 p_1 をもっているとすれば、 $q=0.3$)。そうすると p_2 の比率をもつ小集団の割合は $1-q$ となり、この q は、

$$qp_1 + (1-q)p_2 = p \quad (8.13)$$

を満たさなければならない。

ここからタイプ別の相互作用確率を求めていく。まず、任意の C タイプの個人が小集団において相互作用するとき、その相手が C タイプである確率を p_c とする。これは「任意の C タイプ個人をとったとき、その人が属している小集団における C タイプ比率」の平均に等しく、概念的に、

$$\begin{aligned} & p_1 \times \text{「} p_1 \text{ の小集団に属す } C \text{ タイプ個人の比率」} \\ & + p_2 \times \text{「} p_2 \text{ の小集団に属す } C \text{ タイプ個人の比率」} \end{aligned}$$

で求められる。

そこでまず「C タイプ個人のうち、 p_1 の小集団に属す比率」を求めよう。それは次のようにして求まる。

比率 p_1 の任意の小集団には C タイプが Np_1 人いる。総人口を T とすれば、 N 人小集団の数は T/N である。したがって、 p_1 の小集団に属す C タイプ個人の総数は、 $Np_1 \times q \times T/N$ 。割合としては、全体を pT で除して、

$$C \text{ タイプの個人のうち、} p_1 \text{ の小集団に属している割合} = qp_1/p$$

これが求める比率である。同様にして、 p_2 についても、

$$C \text{ タイプの個人のうち、} p_2 \text{ の小集団に属している割合} = (1-q)p_2/p$$

と求められる。

よって、C タイプの個人にとっての小集団における C タイプの平均的な比率を p_c とすると、

$$p_c = p_1 \times qp_1/p + p_2 \times (1-q)p_2/p \quad (8.14)$$

となる。

次に、やはり任意の D タイプの個人が相互作用するとき、相手が C タイプである確率を p_D とし、これを求める。

D タイプ比率 p_1 の小集団には D タイプが $N(1-p_1)$ 人いる。したがって、 p_1 の小集団に属す D タイプ個人の総数は、 $N(1-p_1) \times q \times T/N$ 。割合としては、全体を $(1-p)T$ で除して、

$$D \text{ タイプの個人のうち、} p_1 \text{ の小集団に属している割合} = q(1-p_1)/(1-p)$$

となる。同様にして、

$$p_2 \text{ の小集団に属している割合} = (1-q)(1-p_2)/(1-p)$$

となる。よって、D タイプの個人にとっての小集団における C タイプの平均的な比率 p_D は、

$$p_D = p_1 \times q(1-p_1)/(1-p) + p_2 \times (1-q)(1-p_2)/(1-p) \quad (8.15)$$

と求められる。

このように、 p_C と p_D とが決まれば、選択性の指標 r が事後的に決まる。それは、たとえば

$$r = p_C - p_D$$

で求めることができる。そして、 r が決まれば、(8.7) の2式から、C タイプの平均利得と D タイプの平均利得がそれぞれ求められ、それによって進化ゲーム的状况での動向が決まるのである。

付録2 図 8.2 において C 選択が合理的であるための確率 α と確率 β の条件

本文で示したように、情報がなくときに C 選択が合理的であるための条件は、(8.10) 式で与えられる。この式は、

$$\alpha\beta \geq (2c - b)/b \quad (8.16)$$

と変換できる。この右辺は、 N 人4人のジレンマの利得に関する基本条件である (8.5) の不等式のもとで、0 から 1 のあいだの値をとる。この式が満たされるかどうかは、2つの確率 α と β の値による。以下、まずこれらの確率を特定化することにしよう。

(1) β について。確率 β は「自分が先番でかつ C を選択したという条件のときに、相手もまた C を選択する確率」である。このとき相手は図 8.2 の B のノードにいる。しかし、相手は自身が B にいることがわからない。わかっているのはたんに「情報が来ていない」ということだけである。先番かどうかもわからない。情報が来ていないのは、自身が先番だからかもしれないし、あるいは自身は後番で相手（今の自分）が C を選択したのかもしれない。

つまり、相手の B の位置は、自分の A の位置と同等である。したがって、確率 β は、相手にとって「情報がなくときに C を選択する確率」に等しい。

ここで、「2人は同じように合理的だ」と仮定する。そうすると、確率 β は、(8.16) 式が成立する確率に等しい。つまり、「(8.16) の不等式が成立している」確率 = β である。これは次のような再帰的な関係式になる。

$$Pr(\alpha\beta \geq (2c - b)/b) = \beta \quad (8.17)$$

(2) α について。 α は「情報がないうきに、自分が先番である確率」で、これは条件付き確率 $Pr(\text{先番} \mid \text{情報がないう})$ である。この確率は、ベイズの定理より

$$\alpha = Pr(\text{先番} \mid \text{情報がないう}) = \frac{Pr(\text{情報がないう} \mid \text{先番}) \times Pr(\text{先番})}{Pr(\text{情報がないう})} \quad (8.18)$$

である。ここで、 $Pr(\text{情報がないう} \mid \text{先番})$ は、確率 1 であり、 $Pr(\text{先番})$ は $1/2$ である。また「情報がないう」という事象は、「自分が先番であるか、もしくは相手が先番でかつ相手が C を選択したとき」に起こる。そして「相手が先番でかつ相手が C を選択する」という事象の確率は、「相手が先番」の確率 $\times$ 「相手が先番であるという条件のもとで、相手が C を選択する」確率である。

ここで、「相手が先番であるという条件のもとで、相手が C を選択する」確率を、先ほどの β の場合と同じように、自分にとっての確率と同等だと考えるとする。すなわち、この確率は「先番であるときに、C を選択する」確率である。これはまさに CD 戦略を選択する確率に等しい。そしてこれは、(8.16) の不等式が成立する確率であり、先ほど (8.17) において β に等しいと想定した。そうすると、

$$\begin{aligned} Pr(\text{情報がないう}) &= Pr(\text{自分が先番}) + Pr(\text{相手が先番でかつ相手が C を選択}) \\ &= 1/2 + 1/2 \times \beta = (1 + \beta)/2 \end{aligned}$$

となる。この式を (8.18) に代入して、

$$\alpha = Pr(\text{先番} \mid \text{情報がないう}) = 1/(1 + \beta) \quad (8.19)$$

最後に、これを (8.17) の α に代入すると、

$$Pr(\beta/(1 + \beta) \geq (2c - b)/b) = \beta \quad (8.20)$$

という再帰的な式が導かれる。

以上で、2つの確率の特定化は完了する。結果として、「情報がないうきに相手が C を選択する確率」である β を軸とする (8.20) 式が導かれた。ここで改めて確認すると、この式は「CD 戦略をとることが合理的である確率」が

β そのものだというを表している。そして重要なことは、左辺の不等式に当該の確率そのものが入っていることだ。そのことは、これは一種の方程式で、これを満たす β はその「解」になることを意味する。

では、どのような β が解になるか。まず、すでに述べたようにこの不等式の右辺は、0 から 1 のあいだの値をとるが、そこにさらに「CD 戦略の組がナッシュ均衡となる条件 (8.8) 式もしくは $N=2$ のときの $b \geq 4c/3$ を考慮すると、

$$0 < (2c - b)/b \leq 1/2 \quad (8.21)$$

が導かれる。

これに対して、不等式 (8.20) の左辺の $\beta/(1+\beta)$ を β の関数として見ると、これはやや上に凸の単調増加関数で、最小値は $\beta=0$ のとき $\beta/(1+\beta)=0$ 、そして最大値は $\beta=1$ のときに $\beta/(1+\beta)=1/2$ である。

ここで、不等式 (8.20) の右辺を $g = (2c - b)/b$ と置く。 g はゲームにおいてあらかじめ与えられている b と c によって決まるある定数で、 β には依存しない。

このとき、(8.20) の再帰式を満たす β の値は 2 つしかない。それは、 $\beta=1$ と $\beta=0$ である。 $\beta=1$ のとき、左辺の不等式は $1/2 \geq g$ となるが、これは $(4c/3 \leq b < 2c)$ の範囲内の任意の g に対して満たされる。したがって、 $Pr(1/2 \geq g) = 1 = \beta$ 。また $\beta=0$ のとき、左辺の不等式は $0 \geq g$ となるが、 g は 0 より大きいので (8.21) 式) 満たされない。よって、 $Pr(0 \geq g) = 0 = \beta$ 。

この 2 つの β に関する解は、ともに内的に整合的である。 $\beta=1$ は「図 8.2 の B のノードで相手が C を選択する確率は 1 である」ことを意味し、それは同時に「自分が CD 戦略をとることが合理的である確率が 1 である」ことを意味している。他方、 $\beta=0$ は「相手が C を選択する確率は 0 で、したがって、自分が CD 戦略をとることが合理的である確率も 0 である」ことを意味する。このどちらも「合理的」なのである。

この合理性は相互依存的である。自分の選択の合理性は相手が同じように合理的であることに依存している。相手もそうだ。そして、この点を超えて、2 つの解のどちらかを他よりも合理的だと判断できる論理は存在しない。

このことは、相手が CD 戦略をとっていることが確実ではないとき、自分の CD 戦略の採用が合理的だとは言えない、ことを意味するのである。

表 8.3 単独で投資を増やしたときの収益等の変化

	もとの集合的最適点	個人投資を 1 人だけ 5.0 に増加
個人投資量	4.5	5 *
トークンからの収益	102.5	100 *
全体投資量 x	36	36.5
全体投資からの収益 $F(x)$	504	506.44
個人の分け前率	0.125	0.137 *
個人収益 u_i	165.5	169.4 *
全体収益	1324	1323.94

* は、投資を増やした個人についての値

付録 3 オストラム実験における個人レベルの合理的な選択とそのジレンマ性

本文中で説明したように、各人の個人投資量が等しく増加するとき、個人の収益が最大となるのは $x_i = 4.5$ のときである。このとき、個人収益の合計としての集団全体の収益も最大化される。

しかし、個人にとっては、他者の投資量がこのレベルのとき、単独で投資量を増やせば個人の収益は増大する。たとえば、他の全員が 4.5 の投資のとき、自分だけ 5.0 に増やしたとしよう。表 8.3 に、その結果、個人収益および全体収益等がどう変化するかを示した。全体の投資量は 36.5 に増大し、そこからの全体の収益 $F(x)$ は 506.44 で、36 のときの 504 を若干上回る。また、個人の取り分比率が、 $1/8$ の 0.125 から 0.137 に増加する。これにより、5.0 に増やした個人の収益は、169.4 に増加する。しかし、全体の収益は、わずかではあるが、もとの 1324 から 1323.94 へ減少するのである。

それでは、個人レベルではどこまで投資を増やしていくことが利益になるのだろうか。具体的には、それは個人レベルの収益関数である式 (8.11) を個人の投資量 x_i の関数として最大化するという選択となる。それは、式 (8.11) を x_i で偏微分して 0 と置いた式、

$$-w + \frac{x_i}{x}(a - 2bx) + (a - bx)\left(1 - \frac{x_i}{x}\right) = 0 \quad (8.22)$$

を満たす x_i の値として決まる。

この x_i の解は、他者の投資量すなわち全体の投資量 x に依存している。ここで、この実験構造ではすべての被験者には同一の利得構造が与えられていることを考慮すれば、全員が同じ x_i をとっていてかつその x_i が (8.22) を満たすならば、その x_i の組はナッシュ均衡をもたらすことがわかる。なぜなら、全員にとって、 x_i はその近傍で個人収益の極大値をもたらす投資量であり、そこからの一方的な逸脱は収益の低下を招くからである。

したがって、ナッシュ均衡での x_i は x/n を満たし、そこでは限界収益率は 0 となる。これを用いて (8.22) 式を解くと、均衡個人投資量は、

$$x_i^* = (a - w) / b(n + 1)$$

となる。本文中と同じように、 $a=23$, $b=0.25$, $w=5$, $n=8$ とすれば、 $x_i^*=8$ となり、合計の x は 64 である。この値は、全員にとって最適な投資量 36 よりもかなり大きく、このとき、個人当たりの収益は最適な場合よりも低下して 141 である。つまり、個人レベルで合理的なナッシュ均衡での投資量は集団的に最適なときの投資量よりも大きいが、逆に収益は低下しているのである（なお、このナッシュ均衡点以外に、同一の x_i ではない投資量の組でナッシュ均衡をもたらすようなものは存在しない。オストラムらはこの証明を示していないが、簡単にチェックすることができる）。

このように、この実験の構造では、集団的に最適な点からこのナッシュ均衡の点までのあいだは、 N 人囚人のジレンマの状況が生じている。これが、共有地の悲劇の構造である。

第9章 ウェブ付録

付録

命題 9.1 警告—謝罪 TFT 戦略がナッシュ均衡であることの証明

証 明

本文に記したように、警告—謝罪 TFT 戦略は次のような方針からなっている。

- i) ノイズを除いて、自分からは（相手がDを出さない限り）Dを出さない。
- ii) 相手のDに対しては、次のiii)の場合を除いて、次の回にDで返す。
- iii) ただし、自分がノイズとしてDを出し、その結果として次の回に相手が出した場合には、DではなくCで返す。

2人のプレイヤーがともに警告—謝罪 TFT 戦略を採用していて、プレイヤーAがノイズを発したとき、その手の連鎖と利得の連鎖は次のようになっている。

A	DCCC……	yzxx……
B	CDCC……	zyxx……

ナッシュ均衡の証明には、この連鎖からの逸脱が、利得を増やさないことを示せばいい。

ここからの逸脱とは、Aがノイズを発したあとに、あるいは同時に、2人のどちらかが警告—謝罪 TFT から（ノイズとしてではなく、意図的に）逸脱した手を出すことである。これには基本的に次の4つの異なる可能性が存在する。

- (a) 第1手目でAがノイズを発するときに、同時に、Bの第1手目がCではなくD（Bの逸脱）。
- (b) 第2手目での逸脱：Aの第2手目がCではなくD（Aの逸脱）。
- (c) 第2手目での逸脱：Bの第2手目がDではなくC（Bの逸脱）。
- (d) 第3手目での逸脱：Aの第3手目がCではなくD（Aの逸脱）。

これら以降、つまり上の手の連鎖でBが第3手目にDを出すという逸脱か

ら先は、警告―謝罪 TFT における新しいノイズの発生とみなしうるので、以上の(a)～(d)ですべての場合が網羅されている。

警告―謝罪 TFT からの逸脱が利得を増やすかどうかは、これらすべての場合について、相手は警告―謝罪 TFT に従っているという前提のもとで、逸脱した場合の利得の流れによる総利得（割引因子で計算したもの）がもとのそれを上回るかどうかを調べればいい。

以下、それぞれの場合について、調べよう。

(a) B が逸脱して、第 1 手目に D を出す場合

このとき、A の第 2 手目は TFT に従って D。B は逸脱しているので、その第 2 手目は C も D もどちらもありうる。そのそれぞれについて検討する必要があるので、この場合は、さらに 2 つに分かれる。。

(a-1) まず、B の第 2 手目が D のとき。このとき、手の連鎖は次のようになる。

A DDC……

B DD……

A の第 3 手目は、「B の第 2 手目の D は自分のミスに対する加罰なので、警告―謝罪 TFT 戦略の論理に従って、それへの応答は C」となっている。

B の第 3 手目は、やはり C でも D でもありうる。

ここで、B の第 3 手目以降に関して、次の 3 通りがありうる。

(ア)ずっと D

(イ) m 回 ($m \geq 1$) D を続けて、C に戻す

(ウ)第 3 手目で C

(ア)のとき、A は TFT の論理に従って第 4 手目以降、D で対抗する。したがって、B の第 3 手目以降の利得の流れは、 $ywvw\cdots$ となる。これは、第 5 章でトリガー戦略のところで見たように、もとの $xxxx\cdots$ の流れよりも（割引因子に関する条件のもとで）利得が高くはない。また、はじめの 2 手については、もとの zy が wv に変わるが、これも命題 9.1 の条件(4)のもとで zy の流れのほうが高いか等しい。よって、(ア)の逸脱は利得を増加させない。

(イ)のとき、B の第 3 手目以降の利得の流れは、 $yw\cdots wz$ となる (w が m 回続く)。これについても、やはり第 5 章の TFT 戦略がナッシュ均衡をなすとい

う命題 5.2 の証明の「場合 2」において、この利得の流れがもとの x だけが続く場合と比べて等しいか低い利得しかもたらさないことが証明されている。

(ウ)のときは、単純にもとの第 1 手目からの利得の流れ zyx と逸脱したときの wvx とを比べればいい。そして条件(4)より、前者の利得のほうが高いか等しいことがわかる (いったん両プレイヤーの手がともに C になると、それ以降に関する利得の流れの比較は、それ以降に生じる逸脱について行えばいい)。

(a-2) 次に、B の第 2 手目が C のとき。このとき、手の連鎖は次のようになる。

A DDC……

B DC……

このあと、B の第 3 手目が C になるか D になるかで、利得の流れは異なってくる。まず、C のときは簡単で、そのときはたんに最初の 2 手の流れを比較すればいい。そうすると、B のもとの利得 zy が逸脱では wz になっているが、これも条件(4)と $y-z > y-w$ から $z+ay > w+az$ なので、もとの利得のほうが高い。

D のときは、ふたたび、「そのあと、ずっと D が続くか」、それとも「 m 回 D が続いて C に戻るか」によって異なるが、どちらの場合も、(a-1) の(ア)もしくは(イ)での検討結果を適用でき、それと先ほどの「第 3 手目が C」のときの利得計算とを組み合わせると、結果として逸脱の利得が低いことが導かれる。

(b) A の第 2 手目が C ではなく D (A の逸脱)

このとき、もとの警告—謝罪 TFT を維持している B の第 3 手目は D になるが、逸脱している A の第 3 手目は、C でも D でもありうる。

A DD……

B CDD……

ここでも場合に分けて考えよう。

(b-1) 第 3 手目が D。このときは、「ずっと D」かもしくは「 m 回 D を出して C に戻る」かのどちらかで、前者はトリガー戦略での逸脱、後者は TFT 戦略での逸脱と同じになり、どちらも逸脱したときの利得は逸脱しない場合よりも高くはない。

(b-2) 第 3 手目が C。このとき、まず第 3 手目までの A の利得の流れを逸

脱のときと逸脱しないときとで比較すると、条件(3)より、逸脱しないほうが利得が高いか等しい。次に、第4手目以降に関しては、もしもどこかでDを出せば、それはTFTにおける逸脱と同じことになり、やはり利得は増えない。Cを出し続ければ、第4手目以降に関しては利得は変わらない。いずれにしても、はじめの3手で逸脱の利得が高くないことから、全体として逸脱の利得が高くないことがわかる。

(c) Bの第2手目がDではなくC (Bの逸脱)

A DCC……

B CC……

ここでは第2手までで、Bの利得はもとよりも低下している。Bが第3手目以降にどこかでDを出しても、それは利得の増加をもたらさない。よって、全体としてBの利得は低下する。

(d) Aの第3手目がCではなくD (Aの逸脱)

A DCD……

B CDCD……

ここでは、Aの利得の比較は、逸脱した第3手目以降からの利得の流れを見ればいい。第3手目以降は、通常のTFT戦略における逸脱と同等であり、したがってそこでの逸脱は割引因子の条件のもとで利得を高めることにはならない。

以上で、すべての場合について、「警告—謝罪TFT」戦略から逸脱した場合の利得の変化を調べたことになる。結果として、どの場合も、逸脱しても利得はもとの利得よりも増加することはない。

よって、「警告—謝罪TFT」戦略の組はナッシュ均衡である。

(証明終わり)

第 10 章 ウェブ付録

付録 1 図 10.1 のシグナリングゲームにおいて均衡をなす戦略の組を導き出す方法

まず、均衡を求めていくステップを概略的に見ておこう。これを理解するうえでは、プレイヤーにとって何がゲームの構造として所与で、何が合理的に選択すべきものかを識別しておく必要がある。図 10.1 のモデルに即して言えば、所与のものと選択で決まるものに加えて、信念として設定されるものがあり、結果として次の 3 種類に分かれる。

(a) ゲームの構造として所与のもの：ゲームの樹状構造，A のタイプの種類，A が選択するシグナルの種類，プレイヤーたちの選択肢，自然の確率 p ，利得を構成する b ， c ，割引因子 α ，およびそれらによって構成される利得。

(b) プレイヤーの選択として決まるもの：(i) プレイヤー A が選択する戦略。タイプが決まったあとにどのシグナルを発信するかという条件付き確率 $P_A(S_j | t_i)$ のセット。具体的には図 10.1 の α と β の値。(ii) プレイヤー B にとっての戦略。A からのシグナルに対して自分の行動選択を割り当てる行動戦略関数 $f(S_j)$ 。

(c) 意図的な選択ではないが、プレイヤーが抱く信念：B の信念。すなわち、シグナル S_j の観測を所与としたときの、A のタイプの条件付き確率 $P_B(t_i | S_j)$ についての B の信念。

さて、これらを踏まえたうえで、整合性と逐次合理性とを満たす完全ベイジアン均衡となる戦略の組を求めるステップは次の 7 つからなる。

- (1) 最初に、図 10.1 のモデルでは、均衡を満たす A の行動戦略、つまりタイプに応じてどのシグナルを発信するかに関する条件付き確率 $P_A(S_j | t_i)$ は、0 か 1 の値だけをとるものとするという前提をおく（この前提の理由については、本付録の 2 「A の戦略を離散解に限定することについて」で説明している）。その前提のもとでは、A の行動戦略を (α, β) のベクトルで表すとき、可能な行動戦略としては $(1, 1)$ ， $(1, 0)$ ， $(0, 1)$ ，そして $(0, 0)$ の 4 つしか存在しない。

このあとは、ここでリストアップされた4個のAの戦略のそれぞれについて分析する。

- (2) Aの戦略が与えられると、本文の(10.2)式から、シグナルS1の発信によるタイプ1の事後確率 q_1 とシグナルS2の発信によるタイプ1の事後確率 q_2 がそれぞれ導かれる。
- (3) その事後確率から、整合性条件により、Bの信念 r_1 と r_2 の値を確定する。
- (4) Aのタイプの確率に関するBの信念が決まると、そのもとでBの期待利得を最大化するBの行動戦略 $f(S_j)$ を導出することができる。今の場合、Bの行動戦略は「S1のときにCかDか」と「S2のときにCかDか」の組で表され、全部で4種類ある。これらの戦略を、CC、CD、DC、DDと表す。たとえばCCは、S1のときにCを選択し、S2のときもCを選択するという戦略を表している。
- (5) 以上によって導かれるBの行動戦略について、当初のAの行動戦略によるAの利得を求める。
- (6) その利得をもとに、Aの当初の行動戦略が、それによって導かれるBの行動戦略に対する「最適応答戦略」になっているか否かを調べる。
- (7) もしもAの当初の戦略が最適応答戦略になっていれば、Aの戦略とBの行動戦略との組は完全ベイジアン均衡である。さもないければ、完全ベイジアン均衡ではないことがわかる。

これらの各ステップの詳細はこのあとの付録3「シグナリングゲームとしての壟断戦における均衡の求め方」で説明するが、その前に、Aの戦略をなす確率を0と1に限定する理由について説明しておきたい。

付録2 Aの戦略を離散解に限定することについて

シグナリングゲームにおいて、Aの行動戦略である「タイプ集合 T のある要素 t_i が出現したときにシグナル S_j を選択する条件付き確率」 $P_A(S_j | t_i)$ は、理論上は連続量として想定される。しかし、実際の行動選択を考えると、自らのタイプを知ったあとのAは、実際にどちらかのシグナルを選択しなければならない。選択そのものは、本来的には離散的、すなわち0か1かである。

ゲーム理論一般では混合戦略についても検討することが多いので、この条件付き確率についても連続量という設定がありえる（現実との接合としては、プレイヤーはサイコロのようにランダムに手を発生させる装置を使って離散的な行動を選択するのだとする）。ただ、シグナリングゲームの場合は、純粋戦略に限定して議論していることが多い。この場合の純粋戦略とは、Aのシグナルの選択が0か1かで離散的であること、そしてBの行動戦略 $f(S_j)$ もCかDかの離散的な選択であることを意味している。

本書でこれまで扱っているゲームのモデルは基本的に純粋戦略のものに限定してきているので、ここでも均衡解に関しては純粋戦略の範囲内に限定して分析することとする。したがって、Aの（タイプが判明する前の事前の）行動戦略 $P_A(S_j | t_i)$ （つまり、 α と β ）と、CかDかを選ぶBの行動戦略 $f(S_j)$ とは離散的であると想定する。

その前提のもとでは、Aの行動戦略を (α, β) のベクトルで表すと、可能な行動戦略として存在するのは $(1, 1)$ 、 $(1, 0)$ 、 $(0, 1)$ 、そして $(0, 0)$ の4つとなる。また、Bの行動戦略は、たとえばS1を観測したときにCを選択し、S2を観測したときにDを選択する戦略を「CD」と表記するとすれば、CC、CD、DC、およびDDの4つの戦略になる。

付録3 シグナリングゲームとしての壟断戦における均衡の求め方

図10.1のモデルにおいて、整合性と逐次合理性とを満たす完全ベイジアン均衡を求めることにしよう。ここでは、付録1のリストに挙げた7つのステップについて、完全ベイジアン均衡をなす戦略の組を導くプロセスを説明する。

(1) Aの戦略を離散解に限定する

上の付録2で述べたように、ここでは純粋戦略に限定して考察するので、Aの行動戦略である条件付き確率 $P_A(S_j | t_i)$ は、確率ではあるけれども、値は0か1かに限定する。その結果、Aの行動戦略は—— (α, β) のベクトルで表すと—— $(1, 1)$ 、 $(1, 0)$ 、 $(0, 1)$ 、そして $(0, 0)$ の4つとなる。

以下では、これら4つのAの戦略それぞれについて分析する。

(2) Aのタイプの事後確率

(10.2)の2つの式から、Aの戦略のそれぞれについて、シグナルの発信を情報とするAのタイプの事後確率を求める。結果は次の通り。

- i) A の戦略が $(\alpha=1, \beta=1)$ のとき。 q_1 は $q_1=p$ と定まる (p は自然の確率)。 q_2 は分母と分子の両方が 0 になるので定まらない。これを $q_2=undef$ と記す。
- ii) A の戦略が $(1, 0)$ のとき。 $q_1=1, q_2=0$ と定まる。
- iii) A の戦略が $(0, 1)$ のとき。 $q_1=0, q_2=1$ と定まる。
- iv) A の戦略が $(0, 0)$ のとき。 $q_1=undef, q_2=p$ 。

(3) 事後確率と整合的な B の信念

以上の事後確率と整合的な B の信念は、次のように決まる。

- i) A の戦略が $(\alpha=1, \beta=1)$ のとき。 $r_1=p$ は決まる。他方、 r_2 は、事後確率が $undef$ なので、任意の値が整合性条件を満たす。つまり、 q_2 に関する B の信念 r_2 は何であってもいい。ただし、その値によって B の行動戦略が決まり、A の戦略 $(1, 1)$ がそれに対する最適応答戦略になっているかを、調べなければならない ($\rightarrow(6)$)。
- ii) A の戦略が $(1, 0)$ のとき。 $r_1=1, r_2=0$ 。
- iii) A の戦略が $(0, 1)$ のとき。 $r_1=0, r_2=1$ 。
- iv) A の戦略が $(0, 0)$ のとき、 r_1 は何でも可。他方、 $r_2=p$ 。

(4) B の行動戦略

B は、A がどのタイプのときにどちらのシグナルを発信するかの条件付き確率についての信念の値が本文の (10.1) 式の右辺よりも大きいか等しければ C を選択し、さもなければ D を選択する。この分岐点の値を r^* とおけば、それは

$$r^* = \frac{c(1-a)}{b-ac}$$

である。この分岐点の値 r^* は、利得のパラメータ b, c と割引因子 a との値によってあらかじめ決まっている。そして、TFT 戦略同士がナッシュ均衡となる条件である $a \geq b/c$ のもとで、 $0 < r^* < 1$ となっている。つまり、一般的にこの分岐点は 0 と 1 のあいだにある。

以上を踏まえて、A のそれぞれの戦略について、B の戦略がどうなるかを見てみよう。

- i) A の戦略が $(\alpha=1, \beta=1)$ のとき。整合性条件より $r_1=p$ なので、S1 の発信に対して、自然の確率 $p \geq r^*$ であれば C で、さもなければ D。他方、 r_2 は B の信念として何であってもいい。よって、ケースが次の 4 つ

に分かれる。

(ア) $p \geq r^*$, $r_2 \geq r^*$ のとき。B の戦略は CC。

(イ) $p \geq r^*$, $r_2 < r^*$ のとき。B の戦略は CD。

(ウ) $p < r^*$, $r_2 \geq r^*$ のとき。B の戦略は DC。

(エ) $p < r^*$, $r_2 < r^*$ のとき。B の戦略は DD。

ii) A の戦略が (1, 0) のとき。 $r_1 = 1$, $r_2 = 0$ なので、戦略は一義的に CD と決まる。

iii) A の戦略が (0, 1) のとき。 $r_1 = 0$, $r_2 = 1$ なので、戦略は一義的に DC と決まる。

iv) A の戦略が (0, 0) のとき、 r_1 は何でも可で $r_2 = p$ なので、 i) のときと同じように次の 4 つの場合に分かれる。

(ア) $r_1 \geq r^*$, $p \geq r^*$ のとき。B の戦略は CC。

(イ) $r_1 \geq r^*$, $p < r^*$ のとき。B の戦略は CD。

(ウ) $r_1 < r^*$, $p \geq r^*$ のとき。B の戦略は DC。

(エ) $r_1 < r^*$, $p < r^*$ のとき。B の戦略は DD。

(5) B の行動戦略がもたらす A の利得

以上のように B の行動戦略が定まると、最終的に到達する帰結がわかる。それは図 10.1 の○で囲ったもののどれかである。そこから、A の戦略のそれぞれについて、A の利得を求めることができる。

i) A の戦略が ($\alpha = 1$, $\beta = 1$) のとき。A はどのタイプでもシグナルは S1。

(ア) B の戦略が CC のとき。A がタイプ 1 のときは、利得は $(b-c)/(1-a)$ 。他方、タイプ 2 のときは、A の利得は b 。

(イ) B の戦略が CD のとき。A はどのタイプであってもシグナル S1 を出すので、B は C を選択する。A のタイプごとの利得は(ア)と同じ。

(ウ) B の戦略が DC のとき。A のタイプがどちらであっても、利得は 0。

(エ) B の戦略が DD のとき。A のタイプがどちらであっても、利得は 0。

ii) A の戦略が (1, 0) のとき。B の戦略は CD。A がタイプ 1 のとき、シグナルは S1。B は C で、A の利得は $(b-c)/(1-a)$ 。タイプ 2 のときは、シグナルは S2。B は D を選択し、利得は 0。

iii) A の戦略が (0, 1) のとき。B の戦略は DC。A がタイプ 1 のとき、シ

グナルは S2。B は C を選択し、利得は $(b-c)/(1-a)$ 。タイプ 2 のときは、シグナルは S1。B は D を選択し、利得は 0。

iv) A の戦略が (0, 0) のとき。A はどちらのタイプであっても S2 を出す。

(ア) B の戦略が CC のとき。A はどちらのタイプであっても S2 を出し、B は C を選択。よって A の利得は、タイプ 1 のとき $(b-c)/(1-a)$ 、タイプ 2 のとき、 b 。

(イ) B の戦略は CD。A はどちらのタイプであっても S2 を出し、B は D を選択。よって A の利得は、どちらのタイプでも 0。

(ウ) B の戦略は DC。B は C を選択するので、A の利得はタイプ 1 のとき $(b-c)/(1-a)$ 、タイプ 2 のとき、 b 。

(エ) B の戦略は DD。A の利得は、どちらのタイプでも 0。

(6) B の戦略を所与として、A が戦略を変更したとき、A の利得は増えるか

均衡では、各プレイヤーの戦略が相手の戦略に対する最適応答戦略になっていなければならない。(5)までのステップにおいて、A の戦略のそれぞれについて、それに対する B の最適な戦略を導き出してきた。しかし逆に、この B の戦略に対して、もとの A の戦略が最適応答戦略になっているかをチェックしなければならない。それぞれの組について、A が別の戦略をとっても利得が増えないかどうかを調べると、結果は次のようになる。

i) A の戦略が $(\alpha=1, \beta=1)$ のとき

(ア) $p \geq r^*$, $r_2 \geq r^*$ で、B の戦略は CC のとき。A の利得は、タイプ 1 のとき $(b-c)/(1-a)$ 、タイプ 2 のとき b 。このとき、A は戦略を変更しても利得は増えない。

(イ) $p \geq r^*$, $r_2 < r^*$ で、B の戦略は CD。A の利得は(ア)と同じで、このときも、A は戦略を変更しても利得は増えない。

(ウ) $p < r^*$, $r_2 \geq r^*$ で、B の戦略は DC。A の利得はタイプにかかわらず 0。ここで B の戦略はそのまま、A が戦略を $(1, 0)$ に変えると、タイプ 1 のときは利得が 0 で変わらないが、タイプ 2 のときは S2 を出すので、B は C を選択し、利得が b に増える。また $(0, 1)$ に変えても、タイプ 1 のとき、S2 を出し、B は C で利得は $(b-c)/(1-a)$ 。よって戦略の変更により、利得を増やすことができる。

- (エ) $p < r^*$, $r_2 < r^*$ で、B の戦略は DD。A のタイプがどちらであっても、利得は 0。だが、戦略を変えても利得は増えない。
- ii) A の戦略が (1, 0) のとき。 $r_1 = 1$, $r_2 = 0$ なので、B の戦略は CD。A がタイプ 1 のとき、利得は $(b-c)/(1-a)$ 。タイプ 2 のときは、0。このとき、B の戦略はそのまま、A が戦略を (1, 1) に変えると、タイプ 2 でも A は S2 を出し、B は C を選択するので、A の利得は b に増える。
- iii) A の戦略が (0, 1) のとき。 $r_1 = 0$, $r_2 = 1$ なので、B の戦略は DC。A がタイプ 1 のとき利得は $(b-c)/(1-a)$ 。タイプ 2 のときは、利得は 0。ここでは A が戦略をたとえば (0, 0) に変えると、タイプ 2 のとき S2 を出し、B は C で、利得は b に増える。
- iv) A の戦略が (0, 0) のとき。
- (ア) $r_1 \geq r^*$, $p \geq r^*$ のとき。B の戦略は CC。A の利得は、タイプ 1 のとき $(b-c)/(1-a)$, タイプ 2 のとき b で、戦略を変えても利得は増えない。
- (イ) $r_1 \geq r^*$, $p < r^*$ のとき。B の戦略は CD。A はどちらのタイプであっても S2 を出し、B は D を選択。よって A の利得は、どちらのタイプでも 0。A は戦略をたとえば (1, 1) に変えれば、タイプ 1 のとき S1 を出し、B は C で利得は $(b-c)/(1-a)$ に増える。
- (ウ) $r_1 < r^*$, $p \geq r^*$ のとき。B の戦略は DC。A の利得はタイプ 1 のとき $(b-c)/(1-a)$, タイプ 2 のとき b 。戦略を変えても利得は増えない。
- (エ) $r_1 < r^*$, $p < r^*$ のとき。B の戦略は DD。A の利得は、どちらのタイプでも 0。A が戦略を変えても、利得は増えない。
- (7) 以上より、戦略を変えても利得が増えないケースは、i) について(ア)(イ)(エ), iv) について(ア)(ウ)(エ)の、計 6 個となる。この結果を、自然の確率 p と B の信念の値の分岐点 r^* との大小関係をもとに整理すると、「均衡となる戦略の組」は次の通りである。

① $p \geq r^*$ のときには、次の 4 個の均衡が存在する。

- イ) A の戦略は (1, 1), B の戦略は CC。
- ロ) A の戦略は (1, 1), B の戦略は CD。
- ハ) A の戦略は (0, 0), B の戦略は CC。
- ニ) A の戦略は (0, 0), B の戦略は DC。

② $p < r^*$ のときには、次の2個の均衡が存在する。

ホ) A の戦略は (1, 1), B の戦略は DD。

ヘ) A の戦略は (0, 0), B の戦略は DD。

付録4 分離均衡と一括均衡

シグナリングゲームでは、一般的に均衡を「分離均衡 (separating equilibria)」と「一括均衡 (pooling equilibria)」とに分けて論じることが多い。一括均衡とは、シグナルの送信者が「タイプが異なっても同じシグナルを発信する」という性質の戦略（一括戦略と呼ぶ）をとるときに成立する均衡であり、分離均衡とは、送信者が「タイプによって発信するシグナルを変える」という戦略（分離戦略と呼ぶ）をとるときに成立する均衡である。

図 10.1 の場合には、一括均衡は A の戦略が (1, 1) か (0, 0) の場合であり、分離均衡は (1, 0) か (0, 1) の場合である。そして、付録3で見たように、結論としては一括均衡しか存在しない。

一括均衡は、「タイプが何であれ、同一のシグナルを発信する」ということで、「異なるシグナルが発信されて、タイプに関する情報が得られる」可能性がないということである。実際、結論として同定されたすべての均衡では、B にわかる「A のタイプ」の事後確率は、結局、自然の確率 p と同じになる。B はもともと p の値は知っているという想定だから、これはシグナルが何の意味ももたなかったことを意味する。

なお、図 10.1 のモデルで分離均衡が存在しない理由は、次の通り。付録3のステップ(6)に示したように、A が分離戦略をとると (ii) と iii) のとき、A がタイプ2のときに発信するシグナルに対して B は D を選択することになる。しかし、A にとってはタイプ2のとき、B が C を選択してくれたほうが望ましい。したがって、A には戦略を変更するインセンティブが存在するのである。

第 11 章 ウェブ付録

付録 ナッシュ均衡と部分ゲーム完全均衡

ゲーム理論における均衡の概念としては「ナッシュ均衡」がよく知られているが、やや複雑なゲームになると、ナッシュ均衡をなす戦略の組の数が多くなりすぎたり、直感的には「合理的」とみなし難いものも存在するようになるという問題がある。そうした問題を整理し、ナッシュ均衡のなかでもより厳密に「合理的な均衡の概念」としてふさわしいものとして打ち立てられたのが、Selten (1975) の「部分ゲーム完全均衡」の概念である。この均衡概念を、単純化した最後通牒ゲームを用いて説明しよう。

図 11.5 は、手持ちが 10 ドル、提案できる額を 1 ドルと 5 ドルという 2 つに限定したものである。

提案者 A の選択肢は「1 ドルを提案する = U : *Unfair*」と「5 ドルを提案する = F : *Fair*」の 2 つである。A にとっての戦略はこの 2 つである。他方、応答者 B は、B1 と B2 の 2 つのノードで受け入れ (a =accept) か拒否 (r =reject) かを選択する。B にとっての戦略は、B1 と B2 のそれぞれにおける選択の組み合わせで構成される。たとえば、B1 でも B2 でも「受け入れ」を選択する、というのが 1 つの戦略となる。こうした戦略は全部で 4 つある。図 11.5 の記号を用いると、それらは $a1a2$, $a1r2$, $r1a2$, $r1r2$ である。こうして、A における 2 つの戦略と B における 4 つの戦略とからなる 8 個の組のうち、相互に相手の戦略に対する最適応答戦略となっているものがナッシュ均衡になる。

これは、表 11.6 のように、それぞれの組に対応する利得を提示すれば、わかりやすい。

ここにはナッシュ均衡をなす戦略の組は 3 個あり、その利得が楕円で囲まれている。 U - $a1a2$ の組がナッシュ均衡をなすことはほぼ自明だ。他方、 U - $a1r2$ の組は、B2 のノードでは B は拒否を選択するという戦略であるが、A が U を選択するという戦略との組み合わせでは、ゲームが到達するのは B1 のノードであり、そこでは $a1$ が選択されて (9, 1) の利得が結果するのである。

F - $r1a2$ の組では、A が公正な提案 F をなし、B1 のノードで B は $a2$ つまり「受け入れる」を選択して、利得は (5, 5) となる。この戦略の組もまたナッ

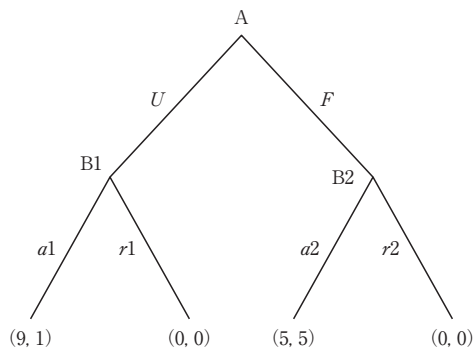


図 11.5 単純化された最後通牒ゲーム

シュ均衡である。それは、表 11.6 を見ればわかるように、お互いに相手の戦略を所与とすると、戦略を変更しても利得は上がらないかもしくは下がってしまうからである。

しかし、これら 3 つのうち、均衡として直感的にわかりやすいのは、 $U-a1a2$ だけだ。そしてこれだけが部分ゲーム完全均衡となる。

それに対して、たとえば、 $F-r1a2$ の組は、次の点で均衡というにはすぐわないところがある。というのは、A は、B が合理的である限り、 U を選択すれば利得 9 を得られるとわかるので、 F を選択するはずがないからである。つまりこの戦略の組は、プレイヤーの合理性を前提にすれば実現するはずのない組なのである。また、 $U-a1r2$ の組も、実際には到達するはずのない B2 のノードでの B の非合理的な選択 $r2$ を含んでいる。

そうすると、 $U-a1r2$ や $F-r1a2$ の組のような戦略の組を「真」の均衡概念からは除外することができるように、より厳格な均衡の概念を想定する必要がある。それが「部分ゲーム完全均衡」である。それは通常、次のような概念として提示されている。

部分ゲーム完全均衡の一般的な定義 展開型ゲームにおいてナッシュ均衡をなす戦略の組が部分ゲーム完全均衡であるとは、その戦略の組がすべての部分ゲームにおいてナッシュ均衡であることである。

表 11.6 戦略の組ごとの利得とナッシュ均衡

		B の戦略			
A の戦略		$a1a2$	$a1r2$	$r1a2$	$r1r2$
U		9, 1	9, 1	0, 0	0, 0
F		5, 5	0, 0	5, 5	0, 0

部分ゲームとは、展開型ゲームを樹木のように見たとき、それを構成する枝のことをいう。ただし、全体も 1 つの部分ゲームとみなす。たとえば図 11.5 のゲームでは、3 つの部分ゲームがある。1 つは全体ゲームそのもの。それに加えて、B1 から始まる部分ゲームと B2 から始まる部分ゲームである（複雑なゲームの場合、部分ゲームについてより厳密な定義が必要だが、ここでは省略する）。

もっとも、上の定義の文章では、「戦略の組が、すべての部分ゲームにおいてナッシュ均衡である」とはどういうことかわかり難い。たとえば B1 から始まる部分ゲームを考えてみよう。その部分ゲームに関して、たとえば U - $a1r2$ という組がナッシュ均衡であるか、あるいはそうではないとはいったいどういうことか。

じつは、上の定義については、次の文を補っておかなければならない。

上の定義への但し書き ただし、部分ゲームにおいては、各戦略は当該の部分ゲームに制限された戦略のことを意味する。

「部分ゲームに制限された戦略」とは、次の意味である。もともとの戦略は全体ゲームにおいて定義されているので、部分ゲームについてはその戦略を構成する要素のすべてが関わるわけではない。図 11.5 の場合に、B1 からの部分ゲームと B2 からの部分ゲームにおいては、A の戦略はまったく関係ない。プレイヤーは B だけだ。そして、B の戦略についても、たとえば $a1a2$ という戦略は、B1 からの部分ゲームにおいて関わるのは $a1$ だけで $a2$ は関係がない。したがって $a1a2$ という戦略は、「B1 からの部分ゲームに制限」されるとたんに「 $a1$ 」になる。同じように、 $r1a2$ は B1 からの部分ゲームに制限されるとたんに $r1$ となる。

表 11.7 部分ゲームに制限された戦略の組

もとの戦略の組	B1 からの部分ゲーム	B2 からの部分ゲーム
$U-a1a2$	$a1$	$a2$
$U-a1r2$	$a1$	$r2$
$F-r1a2$	$r1$	$a2$

これが、「部分ゲームに制限された戦略」という概念だ（より厳密には、たとえばグレーヴァ〔2011: 71〕などを参照）。

次に戦略の組だが、それは「組」といっても、今の場合、B1 からの部分ゲームと B2 からの部分ゲームに関しては A の選択は無関係なのだから、B の戦略だけから構成されることになる（もちろん、部分ゲームでも 2 人のプレイヤーの選択が含まれることもある）。そして B の戦略も今見たように、かなり単純化される。つまり、B1 では各戦略はたんに受け入れる（ $a1$ ）か拒否する（ $r1$ ）かの選択に縮約される。B2 からの部分ゲームに関しても同様である。

以上より、表 11.6 において「ナッシュ均衡」であった 3 つの戦略の組は、部分ゲームに制限すると、このように 1 人のプレイヤー（今の場合は B）にとつての単純な選択に帰着する。この場合、「戦略が部分ゲームにおいてもナッシュ均衡」というときの「ナッシュ均衡」とは、結局は、そのプレイヤーにとっての最適戦略を意味することになる。

表 11.7 をみると、もとの 3 つのナッシュ均衡をなす戦略の組のうち、「すべての部分ゲームにおいてナッシュ均衡」という条件を満たすのは、最初の $U-a1a2$ だけである。なぜなら、 $U-a1r2$ は B2 からの部分ゲームにおいてナッシュ均衡ではないし、 $F-r1a2$ は B1 からの部分ゲームにおいてナッシュ均衡ではないからである（ともに、0 の利得しかもたらさず、他の選択肢に劣る）。

以上が、ナッシュ均衡の概念を基盤に、より厳格な均衡としての「部分ゲーム完全均衡」である。

なお、部分ゲーム完全均衡がすべての部分ゲームでナッシュ均衡だということは、最小の部分ゲームから出発してだんだんとそれを含むより大きな部分ゲームについてナッシュ均衡となる戦略の組をたどっていけば、結果として後退帰納法と同じ道筋を見いだすことになる。グレーヴァ（2011: 71）によれば、有

限の長さの完全情報の展開型ゲームでは、後退帰納法の解は部分ゲーム完全均衡をなす戦略の組をもたらす。