

オンライン・コンテンツ 6.1: 繰り返しゲームでの部分ゲーム完全均衡

図斎 大

本稿は、浅古泰史・図斎大・森谷文利『活かすゲーム理論』有斐閣の第6章までを読んだ読者を対象としたオンライン・コンテンツです。この文書でも引き続きモデル 6.3 (終わりのない仕事) を考えます。モデル 6.3 自体の説明や利得表は本文を参照してください。

3 節で無限繰り返しゲームにおいて、部分ゲーム完全均衡を用いた解法に関して議論するのは「オンライン・コンテンツで行う」と予告していました。それを、ここでは深く掘り下げていきます。実は一回逸脱原理が効いてくるのも部分ゲーム完全均衡ですし、また繰り返しゲームに限らない、「部分ゲーム完全」という言葉の意味がよく見えてきます。本格的に繰り返しゲーム、あるいはゲーム理論における「均衡」を理解したい方は頑張ってみてください。

1. ここでの「均衡」とは何なのか

第6章の 2.1 項で述べたように繰り返しゲームは逐次手番ゲームです。このことは繰り返しが有限でも無限でも変わりません。我々は無限繰り返しゲームを第3節で解くときに、ナッシュ均衡を用いていたわけですが、本来なら逐次手番ゲームは後ろ向き帰納法で解くのが自然ですね。しかし、それは不可能です。なぜなら、後ろ向き帰納法なら、帰納法を始めるための起点として、ゲームの末尾、つまり終わりが無いといけません。でも、無限繰り返しゲームの状況はまさに「終わり」がないのですから、後ろ向き帰納法で解きようがありません。後ろ向き帰納法で解けないということは、部分ゲーム完全均衡は用いることはできないことを意味するのでしょうか。

第5章では後ろ向き帰納法を用いることができるゲームのみを考えてきたので、それを後ろ向き帰納法を用いて部分ゲーム完全均衡を求めました。第5章の最後で均衡経路だけでなく、均衡経路では起こらない仮想的な状況でも利得を最大化しているかどうか、ナッシュ均衡より強い、部分ゲーム完全均衡の特徴だと述べました。後ろ向き帰納法を忘れて、この特徴自体から部分ゲーム完全均衡を定義してみましょう。一度そのように定義してしまえば、後ろ向き帰納法が使えないときでも、部分ゲーム完全均衡を用いて分析することができるようになります。

まず、「部分ゲーム完全」というのは、どの「部分」のゲームを切り出しても「完全」に均衡という意味です。「部分」ということは、ゲームの途中からの戦略の変更、そして「完全」ということはどんな途中経過のパターン(行動の履歴)に対してもということです。つまり、どのような途中経過のパターンに対しても、それ以降について相手の戦略が変わらないとしたときに、やはり自分の利得を改善するような他の戦略がない。そのようなどこも変えようのない戦略をどのプレーヤーも取っているのが、**部分ゲーム完全均衡**という意味です。

実際に使えるように理解するために、まず「部分」という意味をきちんと定義しましょう。図 6-2 のような半展開形では、ステージゲームがゲームの木のようにつらなっていますね。どこでもよいのでステージゲームを一つ選んで、そこから先の部分をすべて切り出しましょう。この切り出したものも、また半展開形のゲームになっています。「部分ゲーム」とはこのように切り出したゲームを指します。もちろん最後のステージゲームのどれか一つで切り出してもよいです。また、用語の問題ですが、何も切り落とさずに元のゲーム全体というのも「部分ゲーム」には含みます。元々の最初のステージゲームで切ったふりをしているとでも思えばよいでしょう。

この「部分ゲーム」で大事なものは、まず、この切り出したステージゲームの頂点（新たな初期点）においては、元のゲームにおけるこのステージまでの途中経過をプレイヤーがみな知っているということです。また切り落とすのは頂点にしたステージゲームよりも前のところだけで、そこから先はすべて残すということは、意思決定において、きちんと先のゲームについて考慮することです。だからこそ、元の全体ゲームに関していえば、各プレイヤーの戦略で行動を切り替える条件となりうるような途中経過をいろいろと考えたうえで最終的にどんな結果になるかを我々は分析したいのですが、この「色々な途中経過」というのはいろいろなステージゲームを頂点することに対応し、そしてそこから先はすべて入れるということで「最終的に」ということを反映しているわけです。つまり、「部分ゲーム」というのは、当てずっぽうにゲームをチョキチョキ切るのではなく、このように元のゲームの分析をするという目的に合うように切ったものなのです。半展開形では、各ステージゲームから切ることとわかりやすいのですが、展開形一般についてはちょっとややこしくなります。なので、それは上級の教科書を見ていただきたいのですが、結構ややこしくて面食らうかもしれません。ただこの説明を念頭にいれておくとよいでしょう。

第5章では「後ろ向き帰納法の解は部分ゲーム完全均衡である」といっていましたが、それはもちろん正しいです。ただ、実は「後ろ向き帰納法で解けるとしたら」という前提が隠されていました。無限繰り返しゲームは後ろ向き帰納法では解けません。しかし、部分ゲームというのは、切った後を見ているというだけですから、終わりがあるかどうかは問題にしません。なので、部分ゲーム完全均衡という概念は無限繰り返しゲームでも当てはめられ、それゆえ後ろ向き帰納法で解けないときでも、部分ゲーム完全均衡は存在します。

2. 空脅しの懸念、再び

単なる「ナッシュ均衡」と比べると、先の下線部の条件「どのような途中経過のパターンに対しても」が入るだけ、部分ゲーム完全均衡は厳しい条件を課しています。つまり、均衡戦略に従っていれば実際には起こらないような、均衡経路の外の途中経過でも利得の改善の余地がないことを、部分ゲーム完全均衡では要求しています。この点は本文の第 6 章3節の分析ではまだ触れていませ

んでした。モデル 6.3 において、両社ともトリガー戦略を取る時を思い出しましょう。均衡経路では 2 社とも「高」をつけ続けるので、その「外」ということはいずれかが「低」を一度でも取ったということです。つまりトリガー戦略においては、既に協力が崩れてこれからは未来永劫の非協力になります。均衡を部分ゲーム完全にするということは、ここで本当に「低」をつけて、協力しなかったことを罰するののかということを問うています。後ろ向き帰納法の解としてですが、第5章で部分ゲーム完全均衡を学んだ時には、ナッシュ均衡では脅しが本当に実行されるものか空脅しなのかを見極めるためでした。それはここでも同じです。

しかし互いがトリガー戦略をとる均衡については心配する必要はありません。相手がトリガー戦略を取るならば、こちらも「低」を取るしかないですね。つまり、このような均衡経路外の部分ゲームでも利得の改善の余地がないわけです。均衡経路の部分ゲームについては既に本文の第 3.2 項で $d \geq 0.5$ ならナッシュ均衡、つまりやはり利得の改善の余地がないことを確かめています。まとめると、 $d \geq 0.5$ が満たされていれば、どの部分ゲームでも利得の改善の余地がない、すなわち互いがトリガー戦略をとるというのは部分ゲーム完全均衡だと言えます。

先のトリガー戦略で「低価格」を取る引き金（トリガー）は、「どこかでだれかが非協力＝低価格を一度でも付けた」ということでした。この「誰か」というのは、実は自分自身も含みます。つまり、真暗はまだ高価格を取ってくれていた中で低価格を付けたのが真光だったとしても、真光もまた（この低価格を付けたのはうっかりだったとして、それ以降は）トリガー戦略に従うのなら、この真光自身がその直後から「低価格」を付け始めるのです。自噴の念に駆られたとも逆ギレとも言え唖然とはしますが、このおかげで真暗も躊躇なく「低価格」を付け始められるのです。なぜなら、真光がトリガー戦略に従う限りは「低価格」から変わることはないので、真暗が少しでも利得を高くするには「低価格」をつけるしかないのですから。なので、トリガー戦略に関しては、部分ゲーム完全でもナッシュ均衡から新たな条件を課すわけでもなく、割引因子が $d \geq 0.5$ であればナッシュ均衡だけでなく部分ゲーム完全均衡になっています。

部分ゲーム完全均衡による違いを見るために、しっぺ返し戦略を考えましょう。モデル 6.3 においては、しっぺ返し戦略というのは、最初のステージでは「高」価格をつけて、それからは直前のステージで相手の取ったのと同じ価格をつけるということを意味しています。今後のために、どんなときに「高」あるいは「低」を選ぶのかを特定させて、この戦略を記述しておきましょう。

しっぺ返し戦略の記述：

- | | | |
|-------------------------|---|--------|
| (O) 現在のステージが最初のステージである | ⇒ | 「高」を選ぶ |
| (H) 直前のステージで、相手が「高」を選んだ | ⇒ | 「高」を選ぶ |

(L) 直前のステージで、相手が「低」を選んだ $\Rightarrow$ 「低」を選ぶ

このしっぺ返し戦略を2社とも取るのが部分ゲーム完全均衡になるのかを確認しましょう。つまり、いろいろな途中経過のパターンを考え、それぞれで相手がしっぺ返し戦略に従うと想定した下で、自分もしっぺ返し戦略を取り続けるかを検証します。

トリガー戦略だと、協力が一度崩れると、ずっと非協力というステージゲームのナッシュ均衡が続くので、逸脱として考える戦略も簡単でした。それで、本文で説明したように、繰り返しゲームにおける、目先の裏切りの果実と将来の非協力による罰との比較という核となる考え方にずっと入れました。しかし、しっぺ返し戦略についてはずっと非協力というわけではないので、逸脱しようとして最初に協力しないとしても後をどうするかがそんなに自明ではなさそうです。罰が空脅しになってないかも気になるでしょうから、部分ゲーム完全均衡を使うことになります。トリガー戦略を超えて、一般に一回逸脱原理が成り立つのは部分ゲーム完全均衡に関してです。そして、協力が一度崩れた後でもまだいろいろな選択があり得る時こそ、一回逸脱原理は分析の強い武器となります。

まずしっぺ返し戦略に慣れるために最初のステージから考えてみましょう。これを今後のためにケース0と呼んでおきます。トリガー戦略と同様に、もしも両社ともにしっぺ返し戦略に従うままなら、共に「高」をつけ続けるので、ずっと(高, 高)という結果が続き、そして両社ともに利得60を得続けます。つまり、最初のステージでしっぺ返し戦略を保った時の真光の沿う利得の現在価値は、本文の式(1)と同じように、

$$60 + 60d + 60d^2 + \dots = \frac{60}{1-d} \quad (\text{A1})$$

となります。他方で、真光がしっぺ返し戦略から逸脱したとしましょう。一回逸脱原理によると、「この最初のステージでのみ逸脱して、次のステージからはまたしっぺ返し戦略に戻る」という戦略を考えれば十分です。この戦略の下で選ぶ行動は何でしょう。まず、最初のステージでは「高」から逸脱するので、「低」を取るようになります。真暗は「高」をつけているのですから、(低, 高)という結果によって、真光は裏切りの利得80を得ます。次のステージはどうなるでしょう。真暗はしっぺ返しとして、直前に真光の取った「低」を取ります。他方で、真光もしっぺ返し戦略に戻るのです。ならば、真倉が直前に取った行動、つまり「高」を取るようになります。なので、(高, 低)という結果になり、真光はむしろ裏切られることになり利得は0となります。この更に次のステージ、つまり3年目はどうなるでしょう。両社それぞれがしっぺ返し戦略を取るなら、真光が今度は「低」を、そして真暗が「高」を選ぶことになりますね。つまり、また(低, 高)という結果によって、真光は裏切りの利得80を得ます。もうお気づきでしょう。この一回逸脱の下では(低, 高)と(高, 低)が往復して繰り返すこととなります。真光の利得は80と0の間を行き来するので、総利得の現在価値は

$$\begin{aligned}
& 80 + 0d + 80d^2 + 0d^3 + 80d^4 + 0d^5 + \dots \\
&= (80 + 0d) + (80 + 0d)d^2 + (80 + 0d)d^4 + \dots \\
&= (80 + 0d) \times \frac{1}{1 - d^2} = \frac{80 + 0d}{1 - d^2} \tag{A2}
\end{aligned}$$

となります。2行目の変形がやや技巧的ですが、2年を「1期」とし、奇数年の利得80と偶数年の(割引)利得0dをひとかたまりの「各期の利得」とみなし、それを2年おきで繰り返すので、 d^2 を1期分の割引因子としているのだと解釈できます。3行目ではそのような解釈で現在価値の公式を適用しています。ともかく、しっぺ返し戦略をこの最初のステージで最適な戦略として取るためには、(A1)の利得が(A2)の利得を下回らない、つまり

$$\frac{60}{1 - d} \geq \frac{80 + 0d}{1 - d^2}, \quad \text{すなわち } d \geq \frac{1}{3} \tag{A3}$$

をdが満たさないといけません。この後者の式を導くにあたっては、前者の両辺に $1 - d^2$ を掛けたいので、 $1 - d^2 = (1 - d)(1 + d)$ だということを思い出して、 $60(1 + d) = 60 + 60d \geq 80 + 0d$ 、つまり $60d \geq 20$ を得ました。(実はこの $60 + 60d \geq 80 + 0d$ という式は、2年をひとかたまりとして見たときに、A1とA2の利得の流列におけるこのひとかたまりでの割引利得の比較になっています。)まとめると、割引因子dが1/3以上なら、最初のステージでの一回逸脱を考えると、両社ともにしっぺ返し戦略から逸脱しないということになります。

しかし展開形でゲームを考えるということは、ゲームの進行中に行動を随時変えていくのを認めるということです。なので、部分ゲーム完全均衡では、ゲームのスタート時でなく、途中経過においても戦略を変えるインセンティブがないかを見ないといけません。「途中経過」、すなわち各プレイヤーの行動の履歴というのはいろいろありえます。無限繰り返しゲームなら、そもそものステージの数からして無限なので、行動の履歴(部分ゲーム)も無限にありえます。なので、部分ゲーム完全均衡を分析するにあたっては、様々な行動の履歴を場合分けするのが大事になります。ではどのように場合分けしましょうか。目的は均衡の候補になっている戦略、ここではしっぺ返し戦略から逸脱する、つまり異なる行動を選びたいかということです。となると、このしっぺ返し戦略での場合分けを活用すればよいのです。すなわち、「しっぺ返し戦略の記述」として書き下したもののから以下のように場合(ケース)分けします。(以下の行動の組では、前者が「自分」=真光、後者が真暗の行動とします。)

ケース HH:	直前のステージが(高, 高).	ケース HL:	直前のステージが(高, 低).
ケース LH:	直前のステージが(低, 高).	ケース LL:	直前のステージが(低, 低).

「しっぺ返し戦略の記述」の中では、直前のステージでの「相手」の行動だけで場合分けをしていましたが、ここでは両社の行動、つまり「自分」も含めて場合分けしています。それはなぜでしょう？こ

れから今のステージで行動を変えるインセンティブを分析したいのです。インセンティブ、つまり利得が上がるかどうかは、相手が今のステージ以降どんな行動を取るか次第です。なので、相手がどんな行動を今のステージで取るかを予想しなければなりません。なので、相手の今のステージでの行動を決める条件、つまりしっぺ返し戦略の下では、「自分の直前のステージでの行動」も、この場合分けに含めるのです。実際、上のように両社の行動の組で場合分けすると、しっぺ返し戦略に従ったときの各社の行動は以下のように特定できます。

ケース HH $\Rightarrow$ しっぺ返し戦略に両社が従うなら、今は（高，高）。

ケース HL $\Rightarrow$ しっぺ返し戦略に両社が従うなら、今は（低，高）。

ケース LH $\Rightarrow$ しっぺ返し戦略に両社が従うなら、今は（高，低）。

ケース LL $\Rightarrow$ しっぺ返し戦略に両社が従うなら、今は（低，低）。

このそれぞれの場合で、「自分」である真光が今のステージでの行動を変えるインセンティブがないかをチェックしていきます。これから先の議論で、ちょっとひっかかるかもしれないのでここで注意しておきますが、途中経過が「均衡経路では起こりえないような」ものになったのがなぜなのかということは考えません。過去のどこかで誰かが均衡戦略からずれたはずですが、それがどうして起こったかは見逃します。ともかくそういうことも「あり得る」と考えて、そのような「仮想的」なケースでどうふるまうかを検討するのが部分ゲーム完全の考え方です。つまり、そのように起きてしまった過去は受け入れたとしたときに、そのような過去の後に何をすべきかについても「戦略」には指定されているわけですが、元々の戦略に指定されたとおりに従うのかどうかを、部分ゲーム完全均衡では見ていきます。またこれから先では、「今」だけでなく、その次、さらに次と先を見ていきます。この時に最初に「今」として見ていたところからどれだけ先のことが、ごちゃごちゃになるといけないので、この「今」というのを T 年目としておきましょう。つまり、この T 年目に真光が改めてしっぺ返し戦略に従うか、それか一回逸脱原理に基づいてこの T 年目の行動は変えてみるかを検討しているということです。

面白い（そしてもちろん鍵となる）ケースとして、ケース HL、つまり直前の $T-1$ 年目に真光は高価格をとったのに、真暗が低価格をつけた状況を考えます。ケース HL で、両社がしっぺ返し戦略に従うのなら、今、この T 年目には真光は「低」を、真暗は「高」をつけて、（低，高）という結果になりますね。すると今度はケース LH になって、次の $T+1$ 年目には（高，低）という結果、つまりケース HL に戻ります。すなわち、 T 年目以降は（低，高） $\rightarrow$ （高，低） $\rightarrow$ （低，高） $\rightarrow$ （高，低） $\rightarrow \dots$ と、（低，高）と（高，低）を行き来することになります。真光の利得は（低，高）では 80、（高，低）では 0 となり、この 2 つの利得を往復することになります。この利得の流列は (A2) と同じですね。なのでその総利得の現在価値は、(A2) で計算した通りです。(A2) で求めた $80/(1-d^2)$ がまた (HL) でしっぺ返し戦略に従うときの利得です。

他方で、T年目で真光がしっぺ返し戦略から逸脱したとします。つまり、一回逸脱原理によると、真光がしっぺ返し戦略に従って「低」をつけるのではなく、「高」をつけるということです。真暗は「高」をつけたままなので、(高, 高)という結果になりますね。すると今度はケースHH になって、次のT+1 年目には(高, 高)という結果、つまり(ケースHH)が続きます。すなわち、T年目で逸脱すると、その年以降は(高, 高)がずっと続き、両社ともに利得 60 を得続けることになります。ということは、(A1)と同じ利得の流列で、(A1)で計算した通りの総利得の現在価値を真光は逸脱で得るのです。(A1)で求めた $60/(1-d)$ が今度はケースHLでしっぺ返し戦略から逸脱するときの利得です。

とすると(A2)と(A1)を比べるから(A3)と同じかと思うかもしれませんが。ちょっと待ってください。このケースHLでは(A2)がしっぺ返し戦略に従う利得、(A1)が逸脱する利得です。これはケース0と真逆です。つまり、このケースHLで真光はしっぺ返し戦略に従う条件は

$$\frac{80}{1-d^2} \geq \frac{60}{1-d}, \quad \text{すなわち } d \leq \frac{1}{3} \quad (\text{A4})$$

となって、(A3)と逆の不等号を得ました。

部分ゲーム完全均衡ではどのケースでもしっぺ返し戦略に従うインセンティブがないといけません。つまりケース0については(A3)での $d \geq 1/3$ という条件、ケース HL については(A4)での $d \leq 1/3$ という条件のいずれもが成立しないといけません。ともに満たすのはギリギリの $d = 1/3$ だけです。なのでたまたま割引因子がちょうど $d = 1/3$ という魔法の数字ぴったりでないとだめということです。(このようにちょうどの数字でないといけなのは、ナイフの刃(knife edge)の上ぴったりでボールかなんかを載せるくらい現実的には難しいということで、knife edge と言われます。)

しかも、実はケース LL とケース LH を比べると同様に真逆の不等号条件を得ます。ケース LL で両社がしっぺ返し戦略に従うならずっと(低, 低)が続き、利得40を両社は得続けます。このケースで一回逸脱するなら自分のみ「低」ではなく「高」を取るのですから(高, 低)としてケース LH へと移ります。その次は(低, 高)となりケース HL へ、そしてまた(高, 低)でケース LH へ戻るので、自分の利得については0と80の往復が続きます。しかしこのパターンはケース LH からスタートしてしっぺ返し戦略に従うのでも実現しますね。他方で、ケース LH で逸脱するなら、むしろ(低, 低)が続き、ケース LL でしっぺ返し戦略に従うのと同じパターンになります。ということで、この二つのケースでもまた、従うときと逸脱するときとでちょうど真逆の利得のパターンになるので、真逆の不等号が従う条件になってしまうというわけです。そしてこの利得の設定では $d = 1$ というのが境目になります。(また2年をひとかたまりで見ると、 $40 + 40d$ と $0 + 80d$ を比べることになります。)となると、しっぺ返

し戦略の組が部分ゲーム完全均衡になるためには、 $d = 1/3$ と $d = 1$ をともに d は満たさないといけません。 $1/3$ と 1 は異なる数字なので無理ですね。

つまり、しっぺ返し戦略の組は(我々の利得の設定ならば)どんな割引因子 d でも部分ゲーム完全均衡にはなり得ないということになります。これはまさにしっぺ返し戦略が「空脅し」になっているということです。しっぺ返しては自分が「低」をつけて裏切ると次の期に相手がしっぺ返しして「低」をつけてきます。相手がそのようにしっぺ返しをしてくることを前提にすれば、自分が将来を軽んじないかぎり、たしかにその相手のしっぺ返しが「低」をつけさせない「脅し」として効くというのが、(A3)が示唆するものです。しかし、もしも均衡経路と違って「低」をつけてきたときに、つまりケース LH で、本当にしっぺ返しをするのかどうかというのが部分ゲーム完全均衡で問われているところです。このとき律儀にしっぺ返しをするということは、(低, 高)と(高, 低)の往復、いわばビンタの応酬という「罰」をしっぺ返しする側も受け入れなければなりません。しかし、ここで許してあげれば、また(高, 高)へ戻れるのです。それでもしっぺ返しをするのかというのが(A4)の条件になるわけです。そしてここでは、しっぺ返しからの逸脱というのが、単に(高, 高)の繰り返しという均衡経路へびったり戻ることになってしまうので、(A4)は「ビンタの応酬での利得 $\geq$ 均衡経路での利得」、(A3)は「均衡経路での利得 $\geq$ ビンタの応酬での利得」という同じものを逆に比べているだけになり、両方を満たすのは等号をぴったり満たすナイフの刃になってしまうわけです。

ところで、ケース HH で一回逸脱だとビンタの応酬になりますが、二回逸脱して、最初に「低」をとるだけでなく、次も(「低」に対して)「低」を取り、そしてからしっぺ返し戦略に戻るとしましょう。そうすると、この二期目に両者が低をとることになるので、しっぺ返しに戻ってもずっと両者が低をとることになります。それで(低, 高) $\rightarrow$ (低, 低) $\rightarrow$ (低, 低) $\rightarrow \dots$ と、2期目以降はずっと(低, 低)が続くことになります。このときの利得和は $80 + 40d + 40d^2 + \dots = 80 + 40d/(1 - d)$ となります。実はこの二回逸脱を防ぐには $60/(1 - d) \geq 80 + 40d/(1 - d)$ 、つまり $d \geq 1/2$ が必要となり、(A3)の $d \geq 1/3$ では足りません。ただ部分ゲーム完全均衡を考えると、(低, 高)のあとに、しっぺ返し戦略に戻って「高」をつけるのではなく、「低」へと逸脱するというのはケース LH で考えていたことでした。そのケースで逸脱しないのなら、ケース HH で一回逸脱してケース LH に入った後に更に逸脱することはない、つまりケース HH で二回逸脱することはないと言えます。つまり、部分ゲーム完全均衡ではすべてのケースを抑えているからこそ、逸脱のいろんなパターンも、それぞれのケースの一回逸脱で十分捉えられているのです。以上が、一回逸脱原理が部分ゲーム完全均衡で成立する論理です。

まとめると、しっぺ返し戦略は「脅し」としては強力でも、実際に「脅し」をするインセンティブはないので、本当に「脅した通りの罰を実行するのか」というとそんなことはなくて「空脅し」でしかない、というわけです。なので、脅してくる側のインセンティブまで先読みするなら、「しっぺ返し」は有効で

はありません。このちょっと素直には思いつかない微妙なところはナッシュ均衡だけでは出てこない、部分ゲーム完全均衡だからこそ明らかになるところです。以上の手順をまとめましょう。

ゲームの攻略法

無限繰り返しゲームでの部分ゲーム完全均衡は以下のように分析できる。

Step 0. 均衡の候補となる戦略を各プレーヤーについて、どんな途中経過（部分ゲーム）でもひとつの行動を定めるように、曖昧さを残さずに書く。

Step 1. 部分ゲームの場合分けをする：候補の戦略において取る行動を切り替える条件を見る。その条件を基に、途中経過を場合分け。

Step 2. 各部分ゲームで、候補の戦略の組がその部分ゲームでナッシュ均衡になっていることを確かめる：1で考えた「場合」のそれぞれについて、各プレーヤーを「自分」として考える。相手が戦略を候補のものから変えない、またこの「場合」に行きついたという過去の途中経過は変えようがないという前提の下で

a. 自分も戦略を候補のものから全く変えないとき

b. 自分は最初の行動を改めて利得が良くなるように変え、その後からの行動は元の候補の戦略に従う（一回逸脱）とき

のそれぞれについて、各プレーヤーの選ぶ行動がどうなっていくかを見してみる。それから「自分」の利得の割引現在価値を計算し、前者 a が後者 b よりも良くなるための割引因子の条件を求める。

Step 3. このように求めた条件をすべての場合、すべてのプレーヤーで満たすような割引因子では、候補の戦略の組が部分ゲーム完全均衡になっている。

注意）相手の「戦略」を変えないというのは、行動を選ぶ切り替え条件とその条件の下での選択する行動を変えないということ。つまり、b で「自分」の戦略を変えたときには、結果としてそれが相手の戦略での切り替え条件にひっかかって、相手が行動は変えるということはある。

ちなみに、しっぺ返し戦略は本文でも触れた政治学者ロバート・アクセルロッドによるシミュレーションコンテストで優勝したことから、「協力」を支える戦略として有名になりました。ただ理論的に考えると部分ゲーム完全均衡ではないという大きな問題があるわけです。「脅し」によって協力が実現するという考えは社会の制度・慣習を考える上で重要な示唆ではあります。それが「しっぺ返し」という特定の戦略で済むのか、このアクセルロッドの結果がどこまで信用に足るものなのかというところは批判的に検証されています（Binmore, 1998）。他方で部分ゲーム完全均衡自体も、どこまで現

実の人間が「先読み」できるのかという点で現代のゲーム理論では是々非々で見直されています(川越, 2020)。またシミュレーションという点でもアクセルロッドが考えたよりも豊かなプレイヤー間の情報・コミュニケーションの構造を入れて, 社会生物学系の進化ゲーム理論で研究が進められています(中丸, 2020)。

参考文献

川越敏司(2020)『行動ゲーム理論(第2版)』NTT出版。

中丸麻由子(2020)『社会の仕組みを信用から理解する: 協力進化の数理』共立出版。

Binmore, K. (1998), "Book review: The Complexity of Cooperation by Robert Axelrod" *Journal of Artificial Societies and Social Simulation* 1(1), <https://jasss.soc.surrey.ac.uk/1/1/review1.html>, 山形浩生訳「アクセルロッド『対立と協調の科学』書評: 「しっぺ返し」はそんなにすごいものではありません」
<https://cruel.org/candybox/axelrodhype.html> (両サイトともに 2023 年 2 月 24 日アクセス)。

練習問題

本文の練習問題 6.2(i)を振り返ろう。この問いでの「厳罰戦略」がお互いに取りるのが, 部分ゲーム完全均衡となるための d の範囲を求めて, それを一回逸脱原理を用いて証明せよ。

解答

厳罰戦略の定義で i) のケースの部分ゲームは, 結局は厳罰戦略に従うなら「高」を取ることにしている場合ということで, 本文練習問題の解答例での議論でおさえている。つまり, このケース i) で厳罰戦略に従う条件は $d \geq 2/5$ となる。

ii) のケースを追加的に考えないといけない。両者ともに厳罰戦略に従うなら, 2社は (低, 低) $\rightarrow$ (高, 高) $\rightarrow$ (高, 高) $\rightarrow \dots$ と 2 期目に (高, 高) に戻り, 総利得の現在価値は

$$10 + 60d + 60d^2 + 60d^3 + \dots$$

となる。一回逸脱原理に従い, 全日が「1 期目に (低に対する最適反応の) 中をとり, 2 期目以降は (i) の厳罰戦略に従う」という戦略に自分だけ変えた場合の, 総利得の現在価値を求める。この場合, 2社は (中, 低) $\rightarrow$ (低, 低) $\rightarrow$ (高, 高) $\rightarrow$ (高, 高) $\rightarrow \dots$ と 3 期目に (高, 高) に戻る。従って, 厳罰戦略から逸脱した場合の総利得の現在価値は

$$20 + 10d + 60d^2 + 60d^3 + \dots$$

となる。やはり差は最初の2期目だけなので、このケースで厳罰戦略から逸脱しない条件は

$$80 + 10d \geq 20 + 10d$$

となる。これも解くと $d \geq 1/5$ となる。

部分ゲーム完全均衡のためには、この両方のケースで厳罰戦略から逸脱しない条件を満たさないといけない、つまり $d \geq 2/5$ と $d \geq 1/5$ の両方を満たさないといけない。前者さえ満たしていれば両方を満たし、厳罰戦略が部分ゲーム完全均衡となる。つまり、 $d \geq 2/5$ がこの問いで求める d の範囲である。

追記) 前者のケース i) の条件はナッシュ均衡の条件だったが、こちらのほうが後者よりも厳しい条件なので、部分ゲーム完全均衡とナッシュ均衡の条件が変わらない。しかしこれは、たまたま今の利得の設定ゆえ、特に(中, 高)と(高, 高)の利得の差 $80 - 60 = 20$ が(中, 低)と(低, 低)の利得の差 $20 - 10 = 10$ よりも高いという設定のためである。もしも逆なら、部分ゲーム完全均衡の条件のほうがナッシュ均衡の条件よりも厳しくなる。(しかし、以上の議論でわかるように、部分ゲーム完全均衡で考える様々な部分ゲームでの条件の中にナッシュ均衡で考えるケースを必ず含むので、前者の条件が後者の条件よりも緩くなるということはない。)