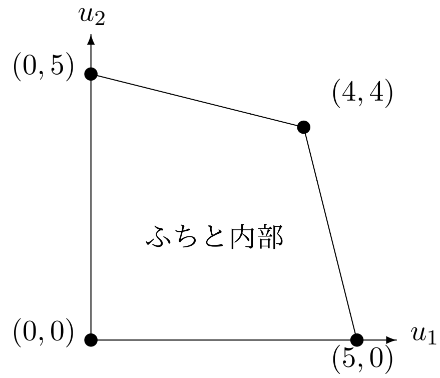
2019年度　ゲームの理論a 演習第3回解答

Takako Fujiwara-Greve

1. (a) 以下の四辺形のふちと内部である。



   (b) 以下の純戦略の組み合わせ (s_1^*, s_2^*) を考える。
   s_1^* : 1期目はAを行い、2期目は1期目が(A,L)のときはBを、そうでないときはCをする。
   s_2^* : 1期目はLを行い、2期目は1期目が(A,L)のときはMを、そうでないときはRをする。

   (s_1^*, s_2^*) が部分ゲーム完全均衡であることを証明する。
   2期目：歴史が(A,L)のときは(B,M)が行われ、これはGのナッシュ均衡である。
   その他の歴史のときは(C,R)が行われ、これもGのナッシュ均衡である。
   以上から2期目の期初から始まる全ての部分ゲームにおいて (s_1^*, s_2^*) をその部分ゲームに制限した戦略の組み合わせはナッシュ均衡であることが示された。
   あとは、1期目の期初から始まる部分ゲーム、すなわち G^2 そのものでもナッシュ均衡であることを示せばよい。

   s_2^* を所与として、Player 1のいろいろな戦略を考える。
   s_1^* に従うと、G^2 の総利得は $4+3=7$。
   1期目にBをすると2期目はPlayer 2はRをしてくるので、最大でも $5+1=6$ しか得られない。
   1期目にCをしても、2期目はPlayer 2はRをしてくるので、最大でも $0+1=1$ しか得られない。
   したがって s_1^* は s_2^* に対する最適反応である。

   同様に s_1^* を所与として、Player 2のいろいろな戦略を考える。
   s_2^* に従うと、G^2 の総利得は $4+3=7$。
   1期目にMをすると2期目はPlayer1はCをしてくるので、最大でも $0+1=1$ しか得られない。
   1期目にRをしても、2期目はPlayer1はCをしてくるので、最大でも $5+1=6$ しか得られない。
   したがって s_2^* は s_1^* に対する最適反応である。 □

2. (a) $x>0$ であるから、どちらのプレイヤーについても純戦略XはYに支配されている。したがってこれは囚人のジレンマであり、混合戦略を含めただ一つのナッシュ均衡は(Y,Y)である。

(b) グリム・トリガー戦略から任意の歴史の後で one-step deviation を行い、その後また s^* に従ったとしても、ずっと s^* に従う方が長期利得が低くならない、ということを証明すればよい。

1回でも誰かがYを行なった歴史の後：
この歴史の後の s^*（の継続戦略）はずっと G' のナッシュ均衡を行うということであるから、この継続戦略から1期間だけ他の行動に変えて戻っても総利得は高まらない。
（Xを正の確率で行うことになるが、そのときの継続利得は最大で $0+\delta\cdot 1+\delta^2\cdot 1+\cdots$ となる。なぜなら相手はずっとYをするからである。これに対し、自分も s^* にしたがってずっとYを行えば $1+\delta\cdot 1+\delta^2\cdot 1+\cdots$ が得られる。）

これまでの歴史が (X,X) だけからなる場合、あるいは第1期の場合：
1期間だけYを行うことを考える。すると歴史は「1回でも誰かがYを行なったもの」に変化するので次期からはお互いずっとYをすることになる。したがってこのような one-step deviation から得られる総利得は

$$(3+x)+\delta\cdot 1+\delta^2\cdot 1+\cdots=(3+x)+\frac{\delta}{1-\delta}$$

である。これに対し、s^* にずっと従うと

$$3+\delta\cdot 3+\delta^2\cdot 3+\cdots=\frac{3}{1-\delta}$$

が得られる。
また、1期間だけYとXを正の確率で行うとすると $p[(3+x)+\delta\cdot 1+\delta^2\cdot 1+\cdots]+(1-p)[3+\delta\cdot 3+\delta^2\cdot 3+\cdots]$ という長期利得になるので、Yを確実に行うような one-step deviation の総利得が s^* の総利得より高くなければ、こちらの one-step deviation の総利得も s^* の総利得より高くならない。
したがって、

$$\frac{3}{1-\delta}\geqq(3+x)+\frac{\delta}{1-\delta}$$

であれば、s^* が最適である。これは

$$\delta\geqq\frac{x}{2+x}$$

と同値である。したがって、δ の下限は $\frac{x}{2+x}$ である。

(c) $\frac{x}{2+x}$ を x で微分すると

$$\frac{d}{dx}\left[\frac{x}{2+x}\right]=\frac{2}{(2+x)^2}>0$$

であるから、δ の下限は x の増加関数である。ゲーム理論的な理由は、x が増えるということは「目的行動の組み合わせ」(X,X) から1期間だけでも逸脱するときの利得の増分が増えるということなので、より逸脱しやすいゲームであるということであり、そのため、将来をより重視している状況（すなわち、より高い δ) でないと逸脱を防ぐことができないということである。

3. (a) 期待利得を計算すると以下のようになる。

A \ B	S	H
(S,S')	4, 4	0.4, 1
(S,H')	1.6, 0.8	1.2, 1
(H,S')	3.4, 3.6	0.2, 1
(H,H')	1, 0.4	1, 1

(b) 純戦略の（ベイジアン）ナッシュ均衡は2つあって ((S,S'), S) と ((S,H'), H) である。