戦略的行動と経済取引（ゲーム理論入門）

4. 囚人のジレンマと繰り返しゲーム

「囚人のジレンマ」ゲーム：復習

- 2人のプレイヤーによる戦略形ゲーム
  - 各プレイヤーの戦略：「協調（cooperation）」or「裏切り（defection）」
    - 協調＝「黙秘」、裏切り＝「自白」
  - 各プレイヤーの利得：

		プレイヤー2	
		協調（C）	裏切り（D）
プレイヤー1	協調（C）	4，4	0，5
	裏切り（D）	5，0	1，1

- u_1(D,C)>u_1(C,C)>u_1(D,D)>u_1(C,D)となっている

「囚人のジレンマ」ゲーム：復習

- 各プレイヤーにとって、Dは強支配戦略
  - 相手がCを選んだときとDを選んだときのいずれのケースにおいても、自分にとってはDを選ぶ方がCを選ぶよりも望ましい
- ⇒ ナッシュ均衡は (D,D)
- 2人のプレイヤーがともにCを選べば、両者の利得はナッシュ均衡 における各プレイヤーの利得よりも大
- ⇒「2人とも協調」という、より望ましい選択肢があるにも関わらず、「2人とも裏切り」という望ましくない結果が均衡では選択されてしまう

繰り返しゲーム

- 繰り返しゲーム（repeated game）
  - 同一の戦略形ゲームが繰り返しプレイされるゲーム
    - 有限回繰り返しゲーム：プレイの回数が有限
    - 無限回繰り返しゲーム：プレイの回数が無限
- 元のゲームをGで表す
  - 「成分ゲーム（component game）」or「ステージゲーム（stage game）」という
- ⇒ T回繰り返しゲームをG(T)、無限回繰り返しゲームをG(∞) で表す

繰り返しゲームにおける戦略と利得

- 繰り返しゲームにおける各プレイヤーの戦略：
  - 過去のプレイに依存して毎回の行動を決定する行動プログラム
    - 毎回の行動≠戦略
- プレイヤーの集合を N={1,…,n} とし、t回目のプレイにおけるプレイヤーiの行動を$a_i^t \in A_i$で表す
- ⇒ t回目までのプレイの履歴（history）：$h^t=(a^1,\dots,a^t)$
  - $a^s=(a_1^s,\dots,a_n^s) \in A$：s回目のプレイで選択された行動の組
    - 「囚人のジレンマ」ゲームでは、毎回の行動は (C,C), (C,D), (D,C), (D,D) のいずれか
- ⇒ t+1回目のゲームにおけるプレイヤーiの行動：t回目までのゲームの可能な履歴 $h^t \in A^t$に対して、行動a_i^{t+1}を対応させる関数 $a_i^{t+1}=f_i^{t+1}(h^t)$
- プレイヤーiの戦略：この関数を集めたもの $(f_i^1, f_i^2, \dots, f_i^{t+1}, \dots)$

5

繰り返しゲームにおける戦略と利得

- 繰り返しゲームにおける各プレイヤーの戦略が与えられる ⇒ 毎回のゲームの利得が決定
- 将来の利得：ゲームの開始時点の現在価値で評価
  - 通常、将来実際に得られる利得よりも割り引いて評価する
    - 例：100年後に100万円を手にすることができる ⇒ 何ももらえないのと同じ
- t回目のゲームでの利得を v_t とする
- ⇒ 利得の現在価値＝$\delta^{t-1} v_t$
  - 0≦δ<1：将来利得に対する割引因子（discount factor）

6

繰り返しゲームにおける戦略と利得

- 仮定：各プレイヤーは利得の現在価値合計（discounted sum of present value）を最大にするように行動 $$U=\sum_{t=1}^{\infty}\delta^{t-1}v_t$$
  - (1-δ) をかけた「平均利得（average payoff）」を考えることもある
  - 無限回繰り返しゲームの場合： $$U=(1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}v_t$$

7

「囚人のジレンマ」ゲームの無限回繰り返しゲーム

- 成分ゲーム：「囚人のジレンマ（prisoner's dilemma）」ゲーム

		プレイヤー2	
		協調（C）	裏切り（D）
プレイヤー1	協調（C）	4，4	0，5
	裏切り（D）	5，0	1，1

- PDゲームが無限回繰り返しプレイされると仮定
- 各プレイヤーは割引利得和を最大にするように戦略を選ぶ
  - 各プレイヤーの割引因子：$\delta \in [0,1)$

8

無限回繰り返しPDゲーム：戦略

- 以下の3つの戦略を考える：
  - all-C
    - 過去のプレイの結果とは無関係に、常にC
  - all-D
    - 過去のプレイの結果とは無関係に、常にD
  - トリガー（trigger）戦略
    - 最初はCをとる
    - 以後、双方がCをとる限りCをとるが、一方が一度でもDをとれば、その後は永遠にDをとり続ける

9

無限回繰り返しPDゲーム：利得

- プレイヤー2が「all-C」のとき、プレイヤー1が
  - 「all-C」のとき：毎回のゲームで (C,C)
  - ⇒ プレイヤー1の割引利得和： $U_1 = 4 + 4\delta + 4\delta^2 + \cdots = \dfrac{4}{1-\delta}$
  - 「all-D」のとき：毎回のゲームで (D,C)
  - ⇒ プレイヤー1の割引利得和： $U_1 = 5 + 5\delta + 5\delta^2 + \cdots = \dfrac{5}{1-\delta}$
  - 「トリガー」のとき：毎回のゲームで (C,C)
  - ⇒ プレイヤー1の割引利得和： $U_1 = 4/(1-\delta)$

10

無限回繰り返しPDゲーム：利得

- プレイヤー2が「all-D」のとき、プレイヤー1が
  - 「all-C」のとき：毎回のゲームで (C,D)
  - ⇒ プレイヤー1の割引利得和： $U_1 = 0 + 0\cdot\delta + 0\cdot\delta^2 + \cdots = 0$
  - 「all-D」のとき：毎回のゲームで (D,D)
  - ⇒ プレイヤー1の割引利得和： $U_1 = 1 + 1\cdot\delta + 1\cdot\delta^2 + \cdots = \dfrac{1}{1-\delta}$
  - 「トリガー」のとき：最初 (C,D)、その後永遠に (D,D)
  - ⇒ プレイヤー1の割引利得和： $U_1 = 0 + 1\cdot\delta + 1\cdot\delta^2 + \cdots = \dfrac{\delta}{1-\delta}$

11

無限回繰り返しPDゲーム：利得

- プレイヤー2が「トリガー」のとき、プレイヤー1が
  - 「all-C」のとき：毎回のゲームで (C,C)
  - ⇒ プレイヤー1の割引利得和： $U_1 = 4/(1-\delta)$
  - 「all-D」のとき：最初 (D,C)、その後永遠に (D,D)
  - ⇒ プレイヤー1の割引利得和： $U_1 = 5 + 1\cdot\delta + 1\cdot\delta^2 + \cdots = \dfrac{5-4\delta}{1-\delta}$
  - 「トリガー」のとき：毎回のゲームで (C,C)
  - ⇒ プレイヤー1の割引利得和： $U_1 = 4/(1-\delta)$

12

無限回繰り返しPDゲームの利得行列

- 利得行列（平均利得で表現）：

		プレイヤー2		
		all-C	all-D	TR
プレイヤー1	all-C	4，4	0，5	4，4
	all-D	5，0	1，1	5-4δ, δ
	TR	4，4	δ, 5-4δ	4，4

13

無限回繰り返しPDゲームのナッシュ均衡

- 戦略の組 (all-D, all-D) は、いかなるδ∈(0,1) に対してもナッシュ均衡
  - 毎回、元のPDゲームのナッシュ均衡 (D,D) をプレイし続ける

		プレイヤー2		
		all-C	all-D	TR
プレイヤー1	all-C	4，4	0，5	4，4
	all-D	5，0	1，1	5-4δ, δ
	TR	4，4	δ, 5-4δ	4，4

14

無限回繰り返しPDゲームのナッシュ均衡

- 「毎回 (C,C) が繰り返される状態」が無限回繰り返しPDゲームのナッシュ均衡として達成されることは可能か？
- 次のケースでは、「毎回 (C,C) が繰り返される状態」が達成されている
  - 2人のプレイヤーがともにトリガー戦略の組 (TR, TR) を採用
- 割引因子が十分大きければ、トリガー戦略の組 (TR, TR) がナッシュ均衡で達成可能
  - 将来を十分高く評価

15

トリガー戦略による協調の達成

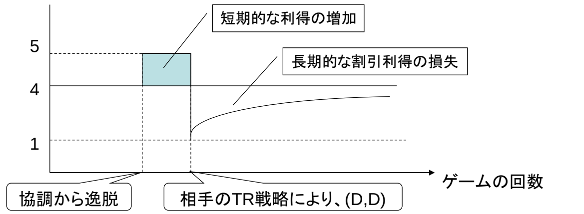
- トリガー戦略の組 (TR, TR) がナッシュ均衡であるためには：
  - プレイヤー2（1）がTRを選んでいるときに、プレイヤー1（2）はTRを選ぶことで最大の割引利得和を得ることができる
- 相手がTRを選んだときに、自分がTR以外を選んだときの利得：
  - all-Cは4 ⇒ TRと同じ利得
  - all-Dを選んだときは5－4δ
- ⇒ 4≧5－4δならば、TRは相手のTRに対する最適反応戦略
- ⇒δ≧1/4

		プレイヤー2		
		all-C	all-D	TR
プレイヤー1	all-C	4，4	0，5	4，4
	all-D	5，0	1，1	5-4δ, δ
	TR	4，4	δ, 5-4δ	4，4

16

トリガー戦略による協調の達成：解釈

- 相手がTRを選んでいるときに、自分はall-Dに変更することで、
  - 初回は利得の増加： $5-4=1$
  - しかし、その後はずっと (D,D) がプレイされるので、利得の損失の割引和： $(4-1)\delta+(4-1)\delta^2+\cdots=3\delta/(1-\delta)$
- ⇒ 短期的な利得の増加≦長期的な割引利得の損失ならば、all-Dに戦略を変更する誘因は存在しない



17

他の戦略による協調の達成

- しっぺ返し（tit for tat）戦略：
  - 最初はCをとる
  - 以後、相手の前回の行動と同じものをとる
- ⇒ 割引因子がある水準以上ならば、毎回のゲームで (C,C) が達成される
  - 詳しくは、岡田章『ゲーム理論・入門』（有斐閣アルマ）、岡田章『ゲーム理論』（有斐閣）を参照
- 協調が達成されるための条件は、トリガー戦略よりも厳しい：
  $\delta_{TFT} \geq \delta_{TR}$, δ_i: 戦略i の下で協調が達成されるための割引因子の下限
  - TFT戦略：相手が裏切った場合、自分は次回だけ裏切る
  - ⇒ TR戦略に比べて、弱い制裁（punishment）のルール
  - ⇒ 裏切ったプレイヤーが被る長期的損失はTR戦略のケースに比べて低い

18

フォーク定理

- 無限回繰り返しゲーム：ナッシュ均衡は複数存在
  - (all-D, all-D)
  - 割引因子が十分大きければ、(TR, TR) もナッシュ均衡
- フォーク定理（folk theorem）：
  - 割引因子が1に十分近ければ、すべての個人合理的な利得の組み合わせは、無限回繰り返しゲームのナッシュ均衡として達成される
    - 個人合理的：相手がどんな戦略を選んだとしても、ある利得水準以上を最低限確保できる

19

有限回繰り返しゲーム

- 繰り返しPDゲームで協調が達成されることにおいて、ゲームが「無限回」繰り返されるという仮定が重要
- ゲームが有限回で終了すると仮定
- ⇒ 終点が存在するので、バックワード・インダクションで解を求めることが可能
- ⇒ (C,C) は達成されない
- T回繰り返しPDゲームを考える
  - T回目（最終回）のゲームにおいては、各プレイヤーはどんな履歴の下でも行動Dを選択するのが最適
  - ⇒ T回目のゲームにおける行動は確定しているので、T-1回目のゲームが最終回 ⇒ T回目と同様、各プレイヤーは行動Dを選択
  - ⇒ ･･･ ⇒ 1回目のゲームにおいて、各プレイヤーは行動Dを選択

20