ワンステップアップセミナー　関東学院大学 KGU 関内メディアセンター　　2012/8/29

ゲーム理論の基礎と応用

大浦宏邦

１　ゲーム理論とは

自分の行動が他者の利得に影響を与え、他者の行動が自分の利得に影響を与える状況（社会的相互依存状況。ゲーム状況ともいう）で、どのような行動がとられるかを考察する理論体系がゲーム理論である。プレーヤー（相互作用の参加者）、戦略（行動の選択肢）、利得（結果の良し悪しの評価）をゲームの 3 要素という。

ある戦略の組み合わせがあって、誰もがその状態から戦略を変えても利得を増やすことができないときに、その戦略の組み合わせはナッシュ均衡であるという。ナッシュ均衡はゲーム理論の静学的な解としてよく用いられる。

またプレーヤーがゲームの結果によって戦略を修正していくことで、ある状態に落ち着くとき、その状態を漸近安定であるという。漸近安定はゲーム理論の動学的な解としてよく用いられる。

このセミナーでは、家事の分担ゲームと新規参入ゲームを題材として、静学分析から認知レベルダイナミクスの分析方法まで学ぶことにしよう。

２　家事の分担ゲーム

掃除や洗濯などの家事を行うと衛生上、心理上の利益がある代わりに時間や手間などのコストがかかる。家事を夫婦で分担する場合を想定して誰が家事のコストを負担することになるのかを考えてみよう。

家事をすることで得られる便益をｂ、家事のコストをｃとする。夫婦がともに家事を行う場合は家事のコストが折半されるので夫婦ともにｂ－c/2 の利得を得ることができる。夫婦いずれかが家事を行う場合、家事を行う側の利得はｂ－ｃ、行わない側の利得はｂとなる。夫婦いずれもが家事を行わない場合はコストが発生しない代わりに利益も得られないのでともに０の利得となる。

ｂ＝６、ｃ＝４として利得表を書くと次のようになる。

夫　＼　妻	家事する	家事しない
家事する	４、４	２、６
家事しない	６、２	０、０

【演習】

ペアを組んだ相手と 10 回家事分担ゲームを行う、結果を記録表に記入する。ゲームはじゃんけんの要領で同時に手を出し合い、指１本を「家事する」、指２本を「家事しない」と考えて記録表に戦略と利得を記入する。

10 回プレーができたら、今度は毎回相手を変えて同様に 10 回プレーを行い、結果を比較してみる。

［静学分析］

（夫：家事する、妻：家事しない）と（夫：家事しない、妻：家事する）が純粋戦略ナッシュ均衡、（夫：確率 1/2 で家事する、妻：確率 1/2 で家事する）が混合戦略ナッシュ均衡となる。

［試行錯誤ダイナミクス］

同じ相手と繰り返しプレーをする場合は、試行錯誤ダイナミクスや最適反応ダイナミクスで分析できる。たとえば、時刻 t に夫が家事をする確率を $x(t)$、妻が家事をする確率を $y(t)$とすると試行錯誤ダイナミクスは

$$E[\Delta x] = x(1-x)(u_{11}-u_{12})/\text{分母}$$

$$E[\Delta y] = y(1-y)(u_{21}-u_{22})/\text{分母}$$

となる。（Ｅは期待値を示す。u_{11} は夫が家事をするときの利得、u_{12} は夫が家事をしないときの利得、u_{21}、u_{22} は妻の利得を示す）

$u_{11}=4y+2(1-y)=2y+2$　　$u_{12}=6y$　　なので

$u_{11}-u_{12}=-4y+2$

これより

$$E[\Delta x] = x(1-x)(-4y+2)/\text{分母}$$

$$= -4x(1-x)(y-1/2)/\text{分母} \quad -①$$

同様に

$$E[\Delta y] = -4y(1-y)(x-1/2)/\text{分母} \quad -②$$

①より、$y>1/2$ のとき $E[\Delta x]<0$　　$y<1/2$ のとき $E[\Delta x]>0$

②より、$x>1/2$ のとき $E[\Delta y]<0$　　$x<1/2$ のとき $E[\Delta y]>0$

これより（ｘ、ｙ）＝（１、０）と（０、１）が漸近安定であることが分かる。メンバーの固定した試行錯誤ダイナミクスでは夫または妻が家事をする状態、つまり上記の純粋ナッシュ均衡（手を変えると損をする strict ナッシュ均衡でもある）の状態になって安定する

ことが予想される。夫と妻のどちらが家事をするようになるかは初期状態に依存する。

なお、試行錯誤ダイナミクスは夫が家事をする傾向性 $P_{11}(t)$ と家事をしない傾向性 $P_{12}(t)$ についてのダイナミクス

$$P_{11}(t+1)=(1-\phi)P_{11}(t)+R_1(t)$$

$$P_{12}(t+1)=(1-\phi)P_{12}(t)+R_2(t)$$

と、家事をする確率は家事をする傾向性に比例するという仮定

$$x(t)=P_{11}(t)/(P_{11}(t)+P_{12}(t))$$

とから導くことができる。ϕ（０～１までの値）は忘却パラメーターであり $R_i(t)$ は時刻 t に選択肢 i に与えられる強化の大きさである。

［ランダムマッチングモデル］

相手を次々に変えてプレーする状況は集団型のランダムマッチングモデルで近似できる。「家事しない」の集団に「家事する」が侵入した場合、もとの戦略の平均利得が０なのに対し侵入戦略の利得は２となる。模倣ダイナミクスを考えると利得の高い戦略が模倣されやすいので「家事しない」には「家事する」が侵入可能と考えられる。

「家事する」の集団に「家事しない」が侵入する場合はもとの戦略の平均利得が４で侵入戦略の利得は６である。したがって、「家事する」に「家事しない」は侵入可能となる。いずれの戦略も安定戦略にはならないので「家事する」と「家事しない」が入り混じった状態が続くと予想される。

街角でたまたま出会った人が自分より利得が高いときに、その人の戦略をまねるというもっとも簡単な模倣ダイナミクスモデルを考えてみよう。「家事する」の割合が x、「家事しない」の割合が $1-x$ のときに、あるプレーで利得４を得る人の割合は x^2、利得２を得る人の割合は $x(1-x)$、利得６を得る人の割合は $x(1-x)$、利得０を得る人の割合は $(1-x)^2$ である。

「家事する」人が「家事しない」に戦略を変えるのは利得４や２の「家事する」人が利得６の「家事しない」人に会った場合である。「家事する」人の割合は x で、利得６の人の割合は $x(1-x)$ なので、戦略の見直しをする人がこのケースに該当する確率は $x^2(1-x)$ である。この確率で「家事する」からの流出が生じる。

「家事しない」人が「家事する」に戦略を変えるのは利得０の「家事しない」人が利得４や２の「家事する」人に会った場合である。利得０の人の割合は $(1-x)^2$ で「家事する」人の割合は x なので、戦略の見直しをする人がこのケースに該当する確率は $x(1-x)^2$ である。この確率で「家事する」への流入が生じる。

単位時間に出会いがおきる確率を r とすると、以上の考察より

$$\begin{aligned} dx/dt &= r\,(x(1-x)^2-x^2(1-x)) \\ &= r\,x(1-x)(1-2x) \\ &= -2r\,x(1-x)(x-1/2) \end{aligned}$$

となる。これより

$0 < x < 1/2$　のとき　$dx/dt > 0$　なので x は増加

$1/2 < x < 1$　のとき　$dx/dt < 0$　なので x は減少

となり、$x = 0$ と 1 は不安定均衡点、$x = 1/2$ が漸近安定点と分かる。

この場合は「家事する」の割合が 1/2 で混合ナッシュ均衡に相当する状態が安定となったが、利得の大きさに依存して戦略見直しの確率や相手を模倣する確率が変わる場合には混合ナッシュ均衡が安定状態となるとは限らない。

また、プレーヤーに「夫」や「妻」といった両者に識別可能なラベルが貼られている場合にはラベルを利用した条件付戦略が可能になる。「夫」の集団の中での戦略頻度と「妻」の集団の中での戦略頻度の変化を考えると「夫の集団は全員家事をし、妻の集団は全員を家事しない」状態と「夫の集団は全員家事をしないが、妻の集団は全員を家事する」状態が模倣ダイナミクスや試行錯誤ダイナミクスの漸近安定状態となる。

［利得の測定］

利得を構成するコストとベネフィットは、就労に要する時間と就労で得られる賃金のように一般に単位が異なっている。進化ゲーム理論を生物学で用いる場合には「闘争のコスト」や「餌獲得の利益」を適応度（子供の数の期待値）の増減に与える効果と考えることで次元をそろえることができる。

試行錯誤ダイナミクスの場合、利得は行動選択肢のとりやすさに与えられる強化の大きさと考えられる。この場合、「強化の大きさ」がコストとベネフィットを統合した次元となる。模倣ダイナミクスの場合は「模倣のされやすさ」が利得に対応するが、他者からはベネフィットは見えやすいがコストは一般に見えにくいため、ベネフィットに重み付けられた確率で模倣がなされる可能性がある。ベネフィットに着目して模倣を行った後に、試行錯誤でコストを体感することでさらに戦略が修正されることが現実のプロセスかもしれない。

最適反応ダイナミクスでは、コストとベネフィットの両方を考慮した「状態の評価」に基づいて最適な選択肢を選ぶことが想定される。無差別曲線の考え方を応用すると状態の評価値を推定することができる。

【演習】　次の状態の「望ましさ」を 10 点満点で評価してみる

１時間掃除をして部屋がきれいに片付いている状態　（　　　）

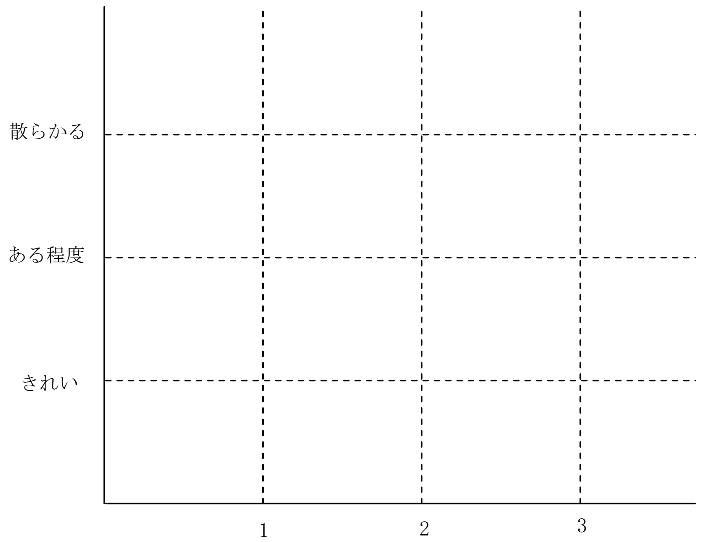
３時間掃除をして部屋が散らかっている状態　（　　　）

１時間掃除をして部屋が散らかっている状態　（　　　）

３時間掃除をして部屋がきれいに片付いている状態　（　　　）

２時間掃除をして部屋がある程度片付いている状態　（　　　）

評価値を図に記入して等高線（無差別曲線）を引いてみる。



同様の方法で男性の平均値や女性の平均値を用いて無差別曲線を引くことで、任意の状態に対する男女の評価値（利得）を推定することもできる。

３　新規参入ゲーム

老舗のＢチェーンが店舗を展開している地域に新興のＡチェーンが進出する場合をモデル化したゲームを考えてみよう。

まずＡチェーンはＢチェーンが展開している地域に進出するかしないかを意思決定する必要がある。進出しない場合はＢチェーンはその地域で利益ｂをあげつづけることができる。Ａチェーンは利益０のままである。

Ａチェーンが進出してきた場合、Ｂチェーンは対抗措置をとるかとらないかを決める必要がある。対抗措置をとらなければＡチェーンはｒのシェアを確保しＢチェーンのシェアは１－ｒに落ちるものとする。

Ｂチェーンが値下げや宣伝強化で対抗措置をとるとＡチェーンは進出を取りやめるかもしれないし、同様の対抗措置をとるかもしれない。両者が対抗措置をとると両者の利益が減ることになる。結果としてＡチェーンが撤退して争いが終結する可能性と、Ｂチェーンが妥協してＡチェーンがｒのシェアを確保する可能性がある。ここでは簡単のためにＡチェーンの撤退とＢチェーンの妥協のいずれかの結果が確率 1/2 で生じ、その間、両者に対抗措置のコストｃが発生するものとしよう。

ｂ＝４、ｃ＝３、ｒ＝1/2　としてゲームの木を描くと次のようになる。

　　　　　　参入　　　　対抗
　　　　Ａ　→　Ｂ　→　（－２、０）
非参入↓　　非対抗↓
（０、４）　（２、２）

【演習】

Ａチェーンと Ｂチェーンに分かれて各自 10 回ずつ対戦する（指１本が参入または対抗、指２本が非参入または非対抗）。結果を記録表に記入する。

Ａチェーンと Ｂチェーンを入れ替えて再び 10 回ずつ対戦する。１回目と２回目の間にチェーン内で作戦会議を開いてもよいものとする。

［静学分析］

後ろ向き推論法によるとＢは「非対抗」の方が得で、Ｂが「非対抗」ならＡは「参入」が得になる。両者が後ろ向き推論法を用いて意思決定するならばＡが参入しＢが非対抗をとると考えられる。

また、ゲームの木を一般形の利得表に書き直すと次のようになる。

Ａ ＼ Ｂ	対　抗	非対抗
参　入	－２、０	２、２
非参入	０、４	０、４

これより（参入、非対抗）が strict ナッシュ均衡、（非参入、対抗）が strict ではないナッシュ均衡になる。前者を第一均衡、後者を第二均衡と呼ぶことにするが、第一均衡は後ろ向き推論法の予想と一致する。第二均衡は「Ｂが対抗することを予期して、Ａが参入を見送る」状態に相当する。Ｂにとって対抗措置はコストがかかるため実際にはとりにくい点に注目して、第二均衡は「根拠のない脅しによる均衡」と呼ばれることもある。

［試行錯誤ダイナミクス］

時刻ｔに確率ｘ(t)でＡが「参入」をとり、確率ｙ(t)でＢが「対抗」をとるとすると、試行錯誤ダイナミクスは

$$
\begin{aligned}
E[\Delta x] &= x(1-x)(u_{11}-u_{12})/\text{分母} \\
&= x(1-x)(-4y+2)/\text{分母} \\
&= -4x(1-x)(y-1/2)/\text{分母} \quad -③
\end{aligned}
$$

$$E[\Delta y] = xy(1-y)(u_{21}-u_{22})/\text{分母}$$
$$= -2xy(1-y)/\text{分母} \qquad -④$$

となる。ｙが変化するのはＢに手番が回るときなので、Ｂに手番が回る確率ｘがＥ［Δｙ］にはかけられていることに注意しよう。

③より　$y > 1/2$ のとき　$E[\Delta x] < 0$

　　　　$y < 1/2$ のとき　$E[\Delta x] > 0$

となる。

また④より　$x > 0$ のとき　$E[\Delta y] < 0$

　　　　　　$x = 0$ のとき　$E[\Delta y] = 0$

これより（ｘ、ｙ）＝（１、０）が漸近安定になることがわかる。Ａが常に参入しＢが常に対抗しない状態、すなわち第一均衡が漸近安定である。

またｘ＝０、ｙ＞1/2 を満たす状態も安定である。この状態からモデル外の要因で少し離れた状態に変化してもそのまま遠くに離れていくことはない。ただ、もとの状態に返るわけではなく、もとの状態の近くに帰ってくる。このような安定のことをリャプノフ安定という。

このリャプノフ安定の状態群はＢが確率 1/2 以上で対抗し、Ａが参入を差し控える状態で第二均衡に相当する。Ａ、Ｂが共に試行錯誤する場合は、第二均衡はリャプノフ安定となり「根拠がない脅し」と言われるほどには根拠がないわけではないことがわかる。

［最適反応ダイナミクス］

次にＡ、Ｂが相手の行動を予測し、この予測に対して最適反応する場合を考えてみよう。実際に両者がとった行動の情報を用いて予測を修正していくことを想定するとこの場合にもダイナミクスが生じる。これを最適反応ダイナミクスという。このダイナミクスでは「相手の行動パターンについての認知」が変化していくので認知ダイナミクスと呼ばれることもある。

「時刻ｔにＢがどれぐらいの確率で対抗をとるか」についてＡが持つ認知を $y_a(t)$ とする。また、「時刻ｔにＡがどれぐらいの確率で参入してくるか」についてＢがもつ認知を $x_b(t)$ としよう。ＡやＢが相手の実際の行動からあとで述べる方法で認知を更新していくとき、認知レベルのダイナミクスは

$$E[\Delta x_b] = (x - x_b)/\text{分母}$$
$$E[\Delta y_a] = x(y - y_a)/\text{分母}$$

となる。y_a が更新されるのはＢに手番が回ったときだけなので、$E[\Delta y_a]$ にはＢに手番

が回る確率ｘがかけられている。

$y_a > 1/2$ のときＡは 1/2 より大きい確率でＢが対抗してくると考えているので参入を差し控える。このときｘ＝０である。したがって

$$E[\Delta x_b] = -x_b/\text{分母} < 0$$

なので、$y_a > 1/2$ のとき x_b は減少する。

$y_a < 1/2$ のときはＢの対抗確率が低いとＡは思っているのでＡは参入を選択する。このときｘ＝１なので

$$E[\Delta x_b] = (1 - x_b)/\text{分母} > 0$$

つまり、$y_a < 1/2$ のとき x_b は増加する。

ＢはＡが参入してくると必ず対抗しないのでｙは実は０である。したがって

$$E[\Delta y_a] = -x y_a/\text{分母}$$

となり

$$x > 0 \quad \text{なら} \quad E[\Delta y_a] < 0$$

$$x = 0 \quad \text{なら} \quad E[\Delta y_a] = 0$$

となる。

これより、$(x_b, y_a) = (1, 0)$ がまず漸近安定と分かる。ＡはＢが常に対抗しないと考え、ＢはＡが常に参入してくると思っている状態が漸近安定である。このときＡは実際に常に参入を行い、Ｂは常に対抗しないのでこれは第一均衡に相当する状態である。

他方、$x_b = 0$、$y_a > 1/2$ を満たす状態群も安定でこれはリャプノフ安定となる。ＡはＢが 1/2 より大きい確率で対抗すると思っているために参入を行わず、それゆえＢがＡが参入してくる確率を０と認識している状態、第二均衡に相当する状態は最適反応ダイナミクスでもリャプノフ安定であることが明らかとなった。

最適反応ダイナミクスはプレーヤーに相当程度の合理的な思考を要求するダイナミクスであるが、この場合でも第二均衡がリャプノフ安定となるということはプレーヤーが賢くないために「根拠のない脅し」が有効になっているわけではないことを示唆すると言えよう。

［最適反応ダイナミクスの導出］

「時刻ｔにＡがどれぐらい参入をとりやすいか」についてＢが持つ認知を $P_{ba1}(t)$、「時刻ｔにＡがどれぐらい非参入をとりやすいか」についてＢが持つ認知を $P_{ba2}(t)$ とする。

この認知がＡの実際の行動によって次のように更新されるものとする。

$$P_{ba1}(t+1) = (1-\phi)P_{ba1}(t) + \delta_a(t)$$

$$P_{ba2}(t+1) = (1-\phi)P_{ba2}(t) + (1-\delta_a(t))$$

ここで ϕ $(0 < \phi < 1)$ は忘却のパラメーター、$\delta_a(t)$ をＡが時刻ｔに実際に参入をとった

ときに１、非参入をとったときに０をとる値である。

「時刻ｔにＡがどれぐらいの確率で参入をとるか」についてＢが持つ認知 $x_b(t)$ が

$$x_b(t)=P_{ba1}(t)/(P_{ba1}(t)+P_{ba2}(t))$$

とすると忘却のある仮想プレーのセッティングとなる。このとき

$$\begin{aligned}\Delta x_b&=x_b(t+1)-x_b(t)\\&=P_{ba1}(t+1)/(P_{ba1}(t+1)+P_{ba2}(t+1))-x_b(t)\\&=((1-x_b(t))P_{ba1}(t+1)-x_b(t)P_{ba2}(t+1))/(P_{ba1}(t+1)+P_{ba2}(t+1))\end{aligned}$$

ここで

$$分子=(1-x_b(t))((1-\phi)P_{ba1}(t)+\delta_a(t))-x_b(t)((1-\phi)P_{ba2}(t)+(1-\delta_a(t)))$$

時刻がそろったので添え字ｔを省略すると

$$\begin{aligned}分子&=(1-x_b)((1-\phi)P_{ba1}+\delta_a)-x_b((1-\phi)P_{ba2}+(1-\delta_a))\\&=(1-x_b)(1-\phi)P_{ba1}+(1-x_b)\delta_a-x_b(1-\phi)P_{ba2}-x_b(1-\delta_a)\\&=(1-x_b)(1-\phi)x_b(P_{ba1}+P_{ba2})+\delta_a-x_b\delta_a\\&\quad-x_b(1-\phi)(1-x_b)(P_{ba1}+P_{ba2})-x_b+x_b\delta_a\\&=\delta_a-x_b\end{aligned}$$

よって

$$\Delta x_b=(\delta_a-x_b)/分母$$

δ_a は確率ｘで１となる（Ａは確率ｘで参入する）ので

$$E[\Delta x_b]=(x-x_b)/分母$$

y_a については、$P_{ab1}(t)$ と $P_{ab2}(t)$ の更新がＢに手番が回ったときになされると仮定すると同様に

$$E[\Delta y_a]=x(y-y_a)/分母$$

となる。

★認知レベルのダイナミクスが $E[\Delta x_b]=(x-x_b)/分母$ の形となるマイクロプロセスは上であげたもの以外にもあると思われる。どのようなマイクロプロセスのクラスならこの形になるのかを調べるのも興味深い課題である。

［試行錯誤と最適反応の対戦］

Ａが試行錯誤、Ｂが最適反応をとる場合には、何がおこるのであろうか。

このときｘが試行錯誤ダイナミクスにしたがって変化し、ｘｂが最適反応ダイナミクスにしたがって変化するので

$$E[\Delta x]=-4x(1-x)(y-1/2)/分母$$

$$E[\Delta x_b]=(x-x_b)/分母$$

である。これより$x_b > x$のときにx_bは減少、$x_b > x$のときにx_bは増加する。また、Bは最適反応をとるので常に$y = 0$である。したがって$0 < x < 1$の範囲でxは増加する。

以上より、この場合は$x = 1$、$x_b = 1$が漸近安定状態で他に安定状態は存在しない。$x = 1$、$y = 0$は第一均衡にあたるのでAが試行錯誤、Bが最適反応の場合は第一均衡のみが安定で第二均衡は実現しないことがわかる。

【練習】

Aが最適反応、Bが試行錯誤の場合について考察してみる。

【解答】

ダイナミクス方程式は

$$E[\Delta y_a] = x(y - y_a)/\text{分母}$$

$$E[\Delta y] = -2xy(1 - y)/\text{分母}$$

である。

Aは最適反応をとるので

$y_a > 1/2$　のとき　$x = 0$　　　　$y_a < 1/2$　のとき　$x = 1$

となる。よって

$y_a > 1/2$ のときは　　　$E[\Delta y_a] = 0$、$E[\Delta y] = 0$

$y_a < 1/2$ のときは　　　$y_a > y$　のとき　y_a減少

　　　　　　　　　　　　$y_a < y$　のとき　y_a増加

また$0 < y < 1$の範囲でy減少

これより$y_a = 0$（なので$x = 1$）、$y = 0$が漸近安定となり第一均衡に相当する。また、$y_a > 1/2$の状態はリャプノフ安定になり、これが第二均衡に相当する。

一般に最適反応は試行錯誤より「賢い」アルゴリズムと考えられるが、このゲームの場合は「賢い」方が損な結果を招きやすいことがわかる。

今回はチキンゲーム系のゲームを中心にみてきた。チキンゲームは資源の争奪や支配・服従、権力といった現象の背後に存在するゲームである。この状況における認知ダイナミクスを考えることで予期の持つ効果をモデル化することもできる。

他方、権力はサンクションの提供を通じて社会的ジレンマの回避に貢献する場合もある。今回は社会的ジレンマには直接触れなかったが、権力が単なる搾取に留まる場合とジレンマ回避に貢献することでパレート改善をもたらす場合との違いの考察など、関連する興味深い研究テーマは沢山あると考えられる。