ゲーム理論アラカルト
— 確率論の立場から —

河 野 敬 雄
(京都大学大学院理学研究科)

目次

§0. ゲーム感覚 —序にかえて—

ベトナム戦争然り, 湾岸戦争然り, NY の自爆テロに対する対応また然り, 「戦争をゲーム感覚でするな」とは日本のマスコミや識者の言である. ならば, 彼らの感覚は, というと正義の月光仮面か, 報復主義か, はたまた某国への単なる追随か. 相手の選択肢を予想し, 自分の選択肢からもっとも合理的な選択を冷静に計算する, というゲーム感覚で物事を考える方が余程理にかなっていると思うのであるが如何なものであろうか.

「ゲーム理論」というとゲームセンターにあるゲーム (これも立派なゲーム理論の演習問題であるが) で遊ぶような連想があるらしく, どうも評判が悪い. 無論, 戦争は遊びではない. 従って, ゲーム感覚より一層洗練された「ゲーム理論的」にやるべきであろう. 孫子の兵法の真髄といわれる「敵を知り己を知らば百戦危うからず」とはゲーム理論の真髄でもある. 国際政治はしばしば「パワーゲーム」といわれ, これも日本では「パワー」に頼るな, というコンテキストで引用される. 一方ではごく近隣の国に対してゲーム理論的には到底合理的とは思われない「パワー」を前提にしたような感情的な対応をしてはばからないのは一体どうしたことであろうか. 「ゲーム理論」をもっと日常の常識にしなければと密かに思うのであるが, 本講義録がその役に立つかどうかははなはだ心もとない.

本講義録執筆の動機は, 畏友福山克司氏から, 私が平成 8 年に行なった講義録を出版しませんかと勧められたことが発端である. 平成 14 年 1 月に神戸大学において, 「ゲーム理論の初歩: 確率変数を用いた表現」と題して談話会でしゃべらせて頂いた縁もあり, この機会に私の考えを纏めてみたいと思うようになった. なお, 同氏には原稿を仔細にチェックして頂き, 内容の不備を含めて最初の原稿から大幅な改善がなされたのはひとえに同氏の御蔭である. 記して感謝の意を表したい.

今までゲーム理論について他分野の人がいろいろ書いておられるのであるが, 数学者特に確率論を専攻する者としてもう少し別の表現, つまり確率変数を表に出した定式化もあり得るのではないか, という思いを抱き続けていたところであった. 当時に比べて私自身も多少はゲーム理論の知識も増加したし, 以前の内容に必ずしも満足していなかったので, 以前の講義録 (プリント) への若干の加筆修正と新しい節を追加したのが本講義録である. 7 節と 9 節は基本的に数理社会学会において発表した内容である. 8 節は確率論のシンポジウムにおいて発表した内容である. 数学的にも新しい定理 (定理 8.1) を含んでいると信じている. また, 定理 8.2 そのものはゲーム理論の分野でよく知られているが, 数学的に厳密な証明が与えられているようには思われない. 原理的に言って, 確率過程を用いて定式化しなければ数学的証明とは言い難いからである.

著者は確率論を専門とする数学者であり, ゲーム理論の専門家ではない. しかし, ゲーム理論に登場する混合戦略とは数学的には確率変数に他ならず, その場合の「期待効用」とは確率論でいう確率変数の平均に他ならない. とするならば, ゲーム理論をコルモゴロフ流の公理的確率論の立場から完全に定式化して解釈することも可能ではないか (サヴェジ流の主観確率からの解釈ではなく) , と考えたのが本講義録の基本的立場である. ゲーム理論は現在では, 経済学, 心理学, 社会学, 動物行動学 (行動生態学) , 認知科学等の分野ではよく知られている. 従って, 数学的に同じ内容でも異なる状況設定では異なる解釈, 含意を持っている. この講義録では数学書にありがちな無味乾燥な数式の羅列ではなく, 可能な限り, 何らかの意味付けを行なうように心がけたが, 一方では数学的証明として論理的ギャップがないようにしたつもりである. (2003.7.26)

§1. 硬貨投げ vs ジャンケン

西部劇映画では, 2 者択一の選択を迫られた時, 硬貨投げで決めよう, という場面がよく出てくるように思われるが日本ではジャンケンで決めるのではなかろうか. ジャンケンは遊びはもちろん実生活の場面でもしばしば用いられる. 硬貨投げもジャンケンもいずれも必然的理由のない (わからない) 時に, 偶然を利用した公平 (と信じられている) 意志決定法として広く採用されている. それでは, 硬貨投げとジャンケンとはどこがどう違うのであろうか. いずれも偶然 (確率) を用いていることにはかわりないように思われる. しかし, 決定的に異なる部分もある. それは, ジャンケンの場合は自分の手 (グーかチョキかパーか) は自分の意志で決められることである. では, どこが偶然かと言えば相手の手が予測出来ない, ということである. もし, 相手はグーをよく出す人であるという何らかの根拠を情報として持っていれば, 自分はパーを出すのが合理的 (勝とうと思えば) である. ゲーム理論が確率論ではない決定的理由はここにある.

ゲーム理論を初めて理論的に構築したのはよく知られているようにフォン・ノイマン (1903-1957) である. 1928 年彼が弱冠 25 才の時である ([209]) ([1]). しかし, 数学の専門誌に載ったためか, まだ社会的に機が熟していなかったためか, 世上ゲーム理論の出発は経済学者 O.Morgenstern との共著の本「Theory of Games and Economic Behavior」(1944) が出版されてからとされているようである (翻訳: 「ゲームの理論と経済行動」全 5 巻, 東京図書 [38]). この本の影響は大きく, 経済学はもとより社会学, 政治学, 心理学, 統計学, 動物行動学 (行動生態学) 等広範囲の分野に及んでいる. ひところブームをよんだウィーナーのサイバネティックス (1948, [7]) より結局は大きな影響を与えたように思われる. しかしながら, 合理的に判断して合理的に行動する人間を想定する彼らの前提は必ずしも多くの人々の受け入れるところとはなっていない. 特に, アメリカのマクナマラ国防長官がベトナム戦争にゲーム理論とコンピュータを利用して以来どうも日本のインテリには評判が悪い.

彼ら以降のゲーム理論の発展のなかで, 囚人のジレンマの「発見」(くわしくは文献 [36]) は単純な合理性だけでは判断出来ない人間の行動をゲーム理論の枠の中で表現出来るという意味でひとつの大きな成果であった. もう一つの成果はメイナード・スミスによる進化的に安定な戦略 (ESS, Evolutionarily Stable Strategy) という概念の導入である ([80]). ESS 概念の導入により, 生物進化を合理的人間を想定することなく, ある意味で合理的にゲーム理論の言葉で解析することが可能になった. この講義録ではこれらの概念を 2 人 2 状態対称ゲームを中心にして数学的側面から解説する.

ゲーム理論の発展史については [26], [28] にくわしい. ただし, ESS に対する評価はあまり高くないようである.

[1] priority に関しては若干の議論がある. フランスの数学者で確率論の黎明期に活躍した Émile Borel (1871–1956) もゲーム理論をある程度のところまで考えていたようである. [36] の 62 頁および [28] の文献表を参照のこと.

§2. 囚人のジレンマ

人間はなぜ約束を守るのだろうか (約束を守らない人でも, 少なくとも約束を守ることが善であることは認めるであろう.) もちろん, 世界は広いからある社会で当然と思われる価値観が別の社会ではそうでないことはしばしばある. しかし, その場合でも別の価値観は必ず存在している. そして, その価値観に従うことが善であることは認識されている. もし, その価値観を (報復を受けることなく) 破る方がその人にとって得をすることが判っている時, 人々はどうするであろうか. 自分は道徳的に振る舞うとしても相手にもそれを期待出来るであろうか. 不道徳な人が得をする世の中というのが長続きするであろうか. (ここで, 暗黙の内に ESS の概念を念頭に置いている. 6 節を参照せよ).

ここで, 囚人のジレンマと言われるゲームを紹介しよう.

今, A, B 二人の人間が一緒に重大犯罪を行い逮捕されたとする. 証拠は十分ではなく, 二人とも黙秘していれば二人とも微罪で済む可能性がある. しかし二人は別々に取り調べられていて, 検事は次のようにささやく. もし, お前が真実を自白すれば情状酌量してやってもよいぞ. そして彼は考える. 検事は当然相棒にも同じことをささやいているであろう. もし, 自分が黙秘を続けて相棒が検事の誘惑に負けると (最悪の事態を想定) 自分が重大犯罪の主犯にされてしまう. 一方, 自分の方が自白したとすれば, もし相棒が黙秘を続ければ自分は情状酌量されるし, もし相手も自白すれば, 黙秘してひとりだけ主犯にされるよりは軽い刑で済むだろう. いずれにしろ自白する方が得であり, 合理的判断というものである.

以上のような考察をゲーム理論を使って定式化してみる. A, B 二人のプレーヤーは 2 種類の戦略, 協調戦略 (Cooperate, 以下 C と記す) と裏切り戦略 (Defect, 以下 D と記す) からひとつを選ぶものとする. 二人がそれぞれひとつの戦略を選んだ時に得られるそれぞれの利得 (利得の善し悪し, つまり大小の比較は出来ると仮定する. 簡単のために数値で表現するがその値そのものは本質的ではない) は, もし A が戦略 C を選び, B は戦略 D を選んだ時, A は利得 1 を得, B は利得 4 を得るとする. これを, 戦略セット (C, D) に対して利得 (1, 4) と表す. 容易にわかるように戦略セットの組み合わせは $2 \times 2 = 4$ だけある. 以下 (C, C) に対して利得 (3, 3) , (D, C) に対して利得 (4, 1) , (D, D) に対して利得 (2, 2) とする. この関係をよくみると, 自分と相手の戦略を逆にしてみると, 自分の利得と相手の利得が丁度逆になっていることがわかる. つまり, 戦略セットに対して A の利得表 (行列で表す方が理解しやすい) を与えれば, B の利得表は転置行列を考えるか, A と B を読み換えるだけでよい. つまり

(1) 囚人のジレンマ

A の利得表		B の戦略	
		C	D
A の戦略	C	3	1
	D	4	2

によって表現出来る. B の利得表は上の表で A と B を入れ替えればよい. A の利得を表にしてみると言葉ではすぐには分からなかったことが表を眺めるだけで直ちに理解出来る.

つまり, B がどちらを選ぼうとも A にとっては戦略 D を選ぶ方が C を選ぶより有利なことは利得行列の 1 行目と 2 行目の各成分を比較してみれば直ちにわかる.

かくて哀れな囚人は自白に追い込まれるのである. しかし, ちょっと待ってほしい. お互いに戦略 D (裏切り, 自白) を選んだ時の利得は A, B 共に 2 でしかない. しかし, もし双方が戦略 C (協調, 黙秘) を選べば利得は A, B 共に 3 あるではないか. なぜ戦略 C (協調, 黙秘) が選べなかったのであろうか, と彼らはともに後悔の涙を流すことになる. 従って, 囚人のジレンマは to be or not to be 式のハムレットのジレンマとは大いに異なる. 囚人のジレンマは合理的に選択するならばジレンマではない. にもかかわらず深刻な問題として捉えられているのは, ではなぜヒトに限らず多くの動物はこれと同じような状況に置かれても協調を選ぶことが多いのか, という事実をどう説明すればよいかジレンマに立たされるからであろう. 事実, アメリカが唯一の原爆保有国であった時, ソ連に対する先制攻撃を行うことが相当真剣に検討されたとのことである. 協調よりも裏切り (先制攻撃) が圧倒的に利益があり, かつ相手との協調が期待し難いときは裏切りは合理的判断と言える. やがて, 相手も原爆を保有するようになり, ゲームは囚人のジレンマ型から, 次に述べるチキンゲーム型になり, リスクの大きい裏切りより協調がよりお互いの利益になって自分の方から裏切る手は放棄し, 相手からの裏切り (先制攻撃) に備えるようになるのである. 現実と異なるひとつの側面は, 現実にはゲームは繰り返し行われ経験ないし学習によってよりよい戦略を学ぶチャンスがある, という事実である. 繰り返しのあるゲームについては 8 節と 10 節で取り上げる.

次にプレーヤーの数が多くなった場合を考えてみよう. 人間の社会生活を, ひとりひとりがゲームのプレーヤーであると見なして考えると, 個々の人間が最も合理的と考えられる (別の言い方をすれば最も利己的に) 行動をした結果, 最終的に誰もが損をする, という結果に導かれる, ということになりはしないか. このような状況はハーディンによって共有地の悲劇として定式化された ([140]). いま共有地を利用して牧畜を営む村を考える. 共有地の広さが一定であればそこで飼育出来る牛の数には自ずから限度がある. しかし, 各人にとっては牛を一頭でも多く飼育する方が利益は大きいわけだから, 当然牛の数は増え続ける. その結果全ての牛は発育不良になって売値は暴落する, というわけである. 利益はまるまる個人の収入になるが, 損失は皆に拡散される状況はゴミや廃棄物の垂れ流しにもみられる状況である. このような例では人々は, だから社会道徳が必要なのだ, と主張しがちである. では, 別の例を考えてみよう. 社会的に必要ではあるが危険な物 (ゴミ処理場, 原子力発電所, 軍事基地等) をある地域に建設しなければならない, とする. その地域にとって利益より損失 (危険) の方が大きいと判断すれば建設に反対するのが合理的である. この例では利益は全体に広がるが, 損失 (危険) は局所的である, という意味で共有地の悲劇とはプラス・マイナスが逆ではあるが, 構造的には同じ社会的ジレンマということが出来る. このような状況が社会的ジレンマである, という認識が大切である.

なお, 囚人のジレンマの定式化の時,

A の利得表		B の戦略 C	B の戦略 D
A の戦略	C	R	S
	D	T	P

において, $T > R > P > S$ の他に, $R > (T + S)/2$ を仮定することが多い. これは, 裏切りと協調を互いに交互に繰り返した平均利得より, 互いに協調した方が得であることを保証するためである. ここで, T は Temptation (裏切りへの誘惑) , R は Reward (協調への報酬) , P は Punishment (裏切りへの懲罰) , S は Sucker (お人好し) の略である ([3]).

ゲーム理論の定式化では数学的にはプレーヤーの数と選択し得る戦略セットの他に各人の利得が重要な働きをする. 質的な利得, 例えば正義感 (道徳的, 宗教的) , 満足感 (嗜好等) , 罪悪感等も少なくとも大小は比較出来る, と仮定しないと「合理的判断基準」の判定が出来ない. そこで, 普通は理解しやすいように数値化する. ここで, 注意しなければならないことは数値化したからといって, 利得を最大にすることがいつも合理的判断基準とは限らないことである. むしろ, 相手より利得が多い (相手に勝つ) ということが合理的基準である場合がある. その典型はいわゆる持久戦といわれているゲームである. 相手を威嚇し続けて (当然時間に比例するコストがかかる) 根負けした方が退散するとしよう. 戦略は威嚇する時間である. 従って, 戦略の集合は 0 を含む正の半直線である. もし, A が $a \geq 0$ 時間がんばることにすると, $a + \varepsilon$, $(\varepsilon > 0)$ だけがんばるつもりの相手に必ず負けてしまう. この意味で最適な純粋戦略 (—言葉の定義は 4 節参照—) はない. しかし, 勝った相手もその時間だけのコストはかかっているから, 勝ったことによる利得が有限である限り, 戦う利益はないことになる. 実際, 攻められると殻をすてて直ちに退散するやどかりの一種がいるそうである. よそで殻を作りなおす方が有利であれば合理的な判断だというべきであろう. ただし, 父祖伝来の土地を守りたい (縄張りの死守) と考える限り, 例えどんな犠牲をはらっても相手に勝たなければ意味がない.

縄張りの死守とは少し意味あいが違うが持久戦の一種として次のような「ドル・オークション」と呼ばれるゲームを考えてみよう. ([36], 336 頁)

基本的には 1 ドルを競り落とす, という単純なゲームである. ただし, 重大な約束があって, 競り落とされたひとつ前の値段を言った人はその金額を支払わなければならない, というものである. 縄張り争いにやぶれれば被害は甚大である, という感じである. 日本人は直ちに談合するから, このゲームの深刻さがいまいち理解出来ないと思われるが, いま談合しないで真剣に競り勝つことを考えよう. もちろん, 安く競り落とした方が得だから, 最初にあなたは「1 セント」という. 当然, 他の人はみすみす 1 セントで 1 ドルを持って行かれるぐらいならばと「2 セント」で応じる. さあ, あなたはどうするか. ここで引き下がれば, 相手は 98 セントを得るのに自分は 1 セント丸損である. そこで, あなたは「3 セント」と叫ぶ. 今度は相手が考える. ここで, 引き下がるとみすみす 2 セントが丸損である. まだまだ, 勝てば利益があるではないか. そこで相手は「4 セント」と応じる, ... もうお分かりと思うが, 延々と続けてとうとうあなたが 99 セントというところまで来る. これで勝てればとにかく 1 セントだけ利益がでる. ところが, 当然その時相手は 98 セント丸損で

ある. それぐらいならば利益がなくとも 1 ドルで競り落とした方がましだ, と相手は「1 ドル」で応じる. さあ, もうたとえ, あなたが競り勝っても利益はでないことは分かっているが, ここで引き下がると 99 セントは丸損である. それぐらいならば 1 セントの損の方がましである. かくてあなたは「1 ドルと 1 セント」を叫ぶ. 同じことを相手も考えて, 1 ドルの損失よりは 2 セントの損失の方がましだ, と「1 ドル 2 セント」と叫ぶ. さあ, この勝負いつ終わるのであろうか. M. シュービックという人は実際にこのゲームをパーティーの席で行ったらしい. その結果夫婦喧嘩に発展した例もあるとか. 日本人でこのようなゲームに参加する人がいるとは思われないが, 問題は無理矢理ゲームに参加させられた時である (その方が現実性がある). その時はお互いに相手に勝つことが当然の合理的判断である, と考える限り悲劇的である, という意味では囚人のジレンマ・ゲームより深刻である. 囚人のジレンマでは話し合えばよりベターな戦略はあり得た. しかし, 持久戦では最初から戦わないのがベターであるが, それが許されない, となれば直ちに降伏するのがベターかも知れないがそれも受け入れ難いとなればどうする? このようなゲームは, ある時点で工事をする意義に疑義が出ても途中で中止するとそれまでの投資が無駄になる, という論理からさらなる無駄を重ねる結果になるというわが国の公共投資を連想させる.

以上紹介した例はすべて, 判断基準として自分の利得を最大にすることだけを基準とした極めて利己的な個人を仮定している ([2]). しかし, 現実の人間社会を考察する場合, 多くの人は, 人間は他人に協力する能力を持っているし, 利他的行動を取ることが出来ると考えているであろう. ゲーム理論に反感を抱く人の中にはゲーム理論の仮定である, このような利己的人間観に反発をしているケースもあるように思われる. しかし, もし, 利己的基準で判断していない, とするならば, その基準をモデルに取り込んだある種の協力ゲーム理論を構成することは可能である. この点については 7 節で改めて考察したい. しかし, 理論上は非協力ゲームの場合がもっとも基本的と思われるので, 本講義録ではもっぱら非協力ゲーム理論を考察する.

なお, 囚人のジレンマ・ゲームに関する文献は研究論文の範囲内でも文字通り山のようにある. 経済学, 心理学, 社会学, 生態行動学, 認知科学関係等関連するデータベースでまず検索されることをお勧めする.

[2]個人合理性という. ミクロ経済学ではこのような個人を前提として理論を構築する.

§3. 2人2戦略対称ゲームの一般論

話を2人2戦略対称ゲームにもどして, 囚人のジレンマ以外にどのような質的に異なるゲームがあり得るかを考えよう. 結局は利得行列を変えるだけであるから, 例えば利得行列が次のようになっているとする.

(2) チキン (弱虫) ゲーム

A の利得表		B の戦略 C	B の戦略 D
A の戦略	C	3	2
	D	4	1

このゲームは弱虫ゲーム (チキンゲーム) とも呼ばれている. 2人の命知らずの暴走族が, ある距離から互いに向かいあって突進する. そのままつっこむと (戦略セット (D, D)) 正面衝突してお互いに大怪我をする. 相手が先に逃げてくれれば (戦略セット (D, C)) 自分は勝者となり相手を弱虫 (チキン) とののしることが出来て最高である. 同時に逃げた場合は ((C, C)) 引き分けで, (D, D) や (C, D) よりはましであるが, 到底仲間のボスにはなれないであろう. このようなゲームは最初からしないに越したことはない. (C, C) がセカンドベストであるが, 相手から仕掛けられた場合, 自分は C を出すから相手にも C を出すように交渉すると, 相手は D を出した方が自分は得, その上相手にはさらに損をさせることが出来るから, 相手を弱腰とみて交渉を決裂させる可能性が大きい (某国の官房副長官の発想はこれに近いのではないだろうか). かくて最悪の選択 (D, D) に突入するのである. キューバ危機を思い出して貰えばこのことは単なる遊びの問題ではないことがわかる. この時は結局フルシチョフがチキンとなった.

しかし, チキンゲームは良識を持った方が負ける, という危険なゲームだからそのような状況に追い込まれないようにしなければならない. 戦争や自爆テロ, 過激派, 武闘派の発生を歴史的に振り返ると, ある時期から当事者間あるいは内部でチキンゲームが始まり, 後戻り出来なかったことがわかる.

(3) 指導者ゲーム ([34])

A の利得表		B の戦略 C	B の戦略 D
A の戦略	C	2	3
	D	4	1

(4) 英雄ゲーム ([34])

A の利得表		B の戦略 C	B の戦略 D
A の戦略	C	2	4
	D	3	1

(3) と (4) はよく似たゲームであるが, 本質的に別のゲームとみなす. この場合は相手に協調して相手と同じ戦略をとることはお互いに損である (日本人には身につまされる話であ

ろう). 従って, 相手より劣ってもよい覚悟があれば話し合いは可能である. しかし, 相手には負けたくないと思うと (持久戦の話を思い出してほしい) 大変である. 特に (3) は互いに相手より優位に立とうとして, 戦略を選ぶと結果としてお互いに最悪の利得しか得られない (アメリカ型). 一方, (4) は互いに相手に譲ろうとするとかえって最悪の結果を招く (日本型).

(5)

A の利得表		B の戦略 C	B の戦略 D
A の戦略	C	3	4
	D	2	1

(6)

A の利得表		B の戦略 C	B の戦略 D
A の戦略	C	3	4
	D	1	2

(5) と (6) の場合は 1 行目と 2 行目を比べてみると, 相手の選択の如何に関わらず A にとっては戦略 C が望ましく, その時 B が戦略 D をとってくれれば最高にハッピーであるが, 多分 B はとんまでなければ戦略 C を選択するから, 結局 (C, C) でお互いに満足 (すべき) である. 自分は C を選び, 相手には D を選ばせるのが一番よいが, ここではプレーヤーは対等と仮定しているのでこの 2 つのゲームはあまり面白くない. 後で定義するようにこの 2 つの場合は (C, C) が唯一のナッシュ均衡戦略となっている.

文献 [34] では上の (5), (6) を除いた 4 つのケースが本質的に意味のあるゲームとして紹介してあるが, 面白いことに文献 [36] では上記の (3), (4), (5) を面白くない場合として除外してある. (3), (4) においては一方が戦略 C, 他方が戦略 D をとる, という秩序が形成されて, 同じ戦略をとろうとして争う様子を表さないからであろうか. ゲームの解釈にもお国柄が表れる感じがする. 文献 [36] では上記の (6) を行き詰まりゲーム (ただし, C と D の意味を逆にしてある) とし, もう一つ次のゲームを加えて 4 つを基本的としている.

(7) 囚人のジレンマ Part II (鹿狩りゲーム)

A の利得表		B の戦略 C	B の戦略 D
A の戦略	C	4	1
	D	3	2

検事にそそのかされて 2 人ともまんまと自白させられた哀れな囚人は極刑はまぬかれたものの仲良く懲役 10 年を食らって隣り合った独房に収監された. そこで, あらためて話し合って見ると, 2 人で協力すると何とか脱獄出来そうなことが分かった. そこで, 示し合わせて明日脱獄を決行することにした (戦略 (C, C)). ところが, 前夜よく考えてみると万一脱獄に失敗すると刑期は倍増するが, 脱獄犯を密告すると刑期は半減される, という規則を思い出した. だだし, お互いに密告しても証拠がないので取り合ってくれない. 従って, 協力しあえば互いに最高であることは分かっているが, もし相手がセカンドベストを選択

すると, 自分は刑期が倍増するは相手の刑期は逆に半減するはで最悪である. 逆に相手の脱獄を自分が密告すれば自分の刑期は半減される. ならば自分もセカンドベストで満足するとしよう, と考えて両者とも約束の行動を取らなかった (C ＝脱獄, D ＝密告). 囚人は結局まじめに刑期を務めるのがベストなのである.

このゲームの場合, (C, C) と (D, D) という戦略セットがいわゆるナッシュ均衡戦略になっていて, お互いに自分から戦略を変えると自分が不利になる. しかし, (D, D) で安定するのはあまり得ではなく, 何とか話し合って (C, C) に移れないか, ただし, 相手が裏切ると最悪である, はたして相手は本当に合理的選択をしてくれるであろうか, というところに面白さがある. 完全な合理的選択理論に立つ限りこのゲームは面白くない. なお, 文献 [36] では鹿狩りゲームと訳されているが, これはルソーの「人間不平等起源論」の第 2 章に「鹿を捕らえようという場合, 各人は勿論そのためには忠実にその部署を守らなければならぬと感じた. しかし. もし兎がかれらの中のだれかの縄張り内にはいってくると, そのために自分の仲間がその獲物を捕らえ損なうことになろうとも, いっこう頓着しなかった. 」に由来する ([3]).

2 人 2 戦略対称ゲームは 4 つの戦略セットの大小関係によって分類すると $4! = 24$ 通りあるが, 大小関係を全て逆にしても C を D と呼び, D を C と呼び換えれば同じことだから本質的に 12 通りある. (C, C) の利得が (D, D) の利得より大きいと仮定して上記以外の場合を書き下すと以下のようになる.

$$(8) \qquad \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix}, \qquad\qquad (9) \qquad \begin{pmatrix} 4 & 2 \\ 1 & 3 \end{pmatrix}.$$

この 2 つのゲームは (C, C) と (D, D) が共にナッシュ均衡戦略で囚人のジレンマ Part II と違って 2 つめのナッシュ均衡戦略が 2 番目によい利得になっているから, いずれにしろこれ以上欲張る必要はない, という意味であまり争う必要はなさそうである.

しかし, 第 3 者から眺めた場合, 必ずしも平和だ, と言っておれない気はする. 私は (8) を「官僚のジレンマ・ゲーム」と名づけることを提案する. こころは, 全員が一生懸命働くならば最高の利得を得ることはわかっているが, 誰かがさぼったり, へまをすると最悪である. 一方, あまり働かなくても (戦略 D) 利得はそこそこにある. ならば, 働かない方がよい. 公務員のスローガン, 「休まず, 遅れず, 働かず」はその意味で合理的なのである. あるいは, リスク回避的に考えるならばやはり戦略 D がそれなりに合理的である. 官僚は如何なる時にもリスク回避的 (責任回避的) に物事を判断する. または某省のように担当大臣と対立した場合, 全員が大臣に協力するのがベストなのはわかっているが, 自分だけが協力して他の同僚がそうではなかった場合にはひどい目にあうことになる. おいそれと大臣に協力出来ないわけである.

[3] 本田喜代治・平岡昇訳 (岩波文庫, 白 156) 83 頁

(10) $\begin{pmatrix} 4 & 3 \\ 1 & 2 \end{pmatrix}$, (11) $\begin{pmatrix} 4 & 2 \\ 3 & 1 \end{pmatrix}$, (12) $\begin{pmatrix} 4 & 3 \\ 2 & 1 \end{pmatrix}$.

(10), (11), (12) のゲームは (C, C) が唯一のナッシュ均衡戦略でかつ利得が最も高いから極めて安定的に互いに (C, C) をとるであろうからその意味でゲームとしてはもっともつまらないと思われる. しかし, 少し見方を変えて, 非協力という意味ではなく, 自分が謙虚な人柄で, 相手より少ない利益を得ようと思って D を選択すると相手も自分も損をする結果になる. 余計なことを考えずにもっと合理的に考えろ, といいたくなる人がたまにいるのは事実である.

以上の考察はもっぱら自分の利得を最大にするのがよい戦略である, という価値判断で論じてきたが, 合理的判断基準というのは必ずしも利得を最大にすることだけではない. 零和ゲームの時は主にリスクを最小にするのが合理的である, という考え方をする. 対称ゲームの場合にもこの考え方で考察するとどうなるかを考えてみる. 互いに利得が 1 になる可能性のある戦略を避けることにすると, 結果として共に 2 しか得られないゲームは (1) の囚人のジレンマ, (3) の指導者ゲーム, (4) の英雄ゲーム, (7) の囚人のジレンマ Part II, の 4 つである. 結果として共に 3 しか得られないゲームは, (2) のチキンゲーム, (5), (6), (8) の 4 つ, 結果として共に 4 の最高の利得を得られるゲームは (9), (10), (11), (12) の 4 つである. また, 持久戦のようにとにかく相手に勝つ (相手より利得が多い) ということを基準にした場合, 対称ゲームでは相手と同じ戦略をとったのでは必ず引き分けになるから, 実際には実現しないわけであるが, お互いにそう考えて戦略を決めた結果どうなるかというと, 結果として互いに最悪の利得 1 にしか得られないゲームは (2) のチキンゲーム, (3) の指導者ゲーム, (11) の 3 つ, 利得 2 になるゲームは, (1) の囚人のジレンマ, (4) の英雄ゲーム, (7) の囚人のジレンマ Part II, の 3 つ, 利得が 3 になるゲームは (5), (6), (8) の 3 つ, 最高の利得 4 になるハッピーなゲーム (ゲームと言えるかどうか疑わしいが) は, (9), (10), (12) の 3 つである. 3 通りの判断基準で考えても名前のついているゲームはそれなりにジレンマを含んでいて争いの種を含んでいる. 特に (11) は互いに最悪の手を避けようとすると, 結果的に最高の利得を得られて最高にハッピーであるが, ひとたび相手に勝とうとすると, 最悪の結果になる. 極めて中の良かった親友どうしがひとたび競争心を持ったとたんに最悪の結果になる受験生のような関係を象徴しているのかも知れない. このような関係を取り上げたゲーム理論の教科書はないようである.

また, 上記の (7) から (12) は (C, C) という戦略セットが他のすべての戦略セットに対して, A, B 双方にとっても最大の利得であるから, 最大利得を獲得するのが目的のゲームでは争いが起こらない, と考えられている. そのため, 最初から除外して考えているゲーム理論の教科書もあるが, 上でみたように, 相手より相対的に高い利得を得たいとか, 最悪の事態を避けたい, というのであれば話は別である. 同様に, 他のすべての戦略より双方ともに最低の利得しか得られない戦略は最初から考慮しない, というのも問題である. 利得の最大ばかりを考える場合は無視してよい戦略であるが, 最悪な戦略を避けたい, という判断をする場合は, 最悪な戦略が存在しているか, どうかは決定的に重要である.

零和ゲームの場合は, 自分の利得がプラスであるということは相手の利得がマイナスということであるから, 最大の利得を得ようとすることと, 相手に勝とうとすることは同じ

である. この場合は最大利得を得ようと試みるより最悪の事態を避けることを判断基準としている. これに反して, 対称ゲームの場合は相手と同じ戦略をとる限りは同じ利得しか得られないから, 相手に勝つことを重視しないで, 最大利得を得ることを判断基準としている. 現実のゲームではどのタイプかは明確ではないから, むしろ利得優先タイプか, 勝ち負けに拘るタイプか, 最悪の事態を避けようとするタイプかはプレーヤーの性格にも大きく依存するように思われる. 新聞報道によると, アメリカのクリントン大統領は相手に勝つことを判断基準にしているらしい (アメリカ人には多いタイプではあるが, 国際政治をこれでやられるとたまったものではない).

ゲーム理論の面白さは数学的構造は同じでも戦略のイメージを変えるとさまざまな現実場面を想定出来るところにあるように思われる. なお, ゲームの手 (戦略) を交互に出す場合や繰り返し同じゲームを行う場合は同じ利得であっても異なった価値判断が可能であるから, まず同時にイチ, ニのサンで手 (戦略) を示す場合を考えてほしい. あとでくわしく述べるように混合戦略というのは複数の手をある確率に従って出すことをいうので, 確率という以上は繰り返しを前提にしないと意味がない. 日本におけるゲーム理論の第 1 人者である鈴木光男氏は著書の「ゲーム理論入門」 (共立全書 239, (1981), 30 頁) ([25]) において, 「混合戦略はくり返し行われるゲームにのみ有効であるという意見が出されているが, そうではなくて, 1 回限りのゲームにおいても, そうせざるをえない状況においては, とらざるをえない決定方法なのである」と主張しておられるが, 100 年以内に人類が滅亡する確率を議論しても仕方ないように (各人の確信度を 0 と 1 の間の数値で表現する, という心理学的な使い方はあり得るが, それは生起確率とは別である) 1 回きりしか起こらないことが分かっている場合には確率とは何かという, より本質的な問題が生じるから確率というものをあまり形式的に適用すべきではない.

以上, いろいろなそれぞれのタイプのゲームにフィットするような物語を考えたが, 背景になっている判断基準として各プレーヤーの個人合理性, つまり個人の利得のみを最大にすることを価値基準として戦略を決めている, と断定すると社会学的含意あるいは面白さが理解できない場合があることに気づく. 例えば, 囚人のジレンマは完全に個人合理性に立つ限り, 答えは明白でそれ以上議論する余地がない (少なくとも私にとってはジレンマではない). それなのに, 何故ジレンマだ, といって騒ぐのであろうか. それは, 暗黙のうちに人間はそんなに利己的な動物ではない, 実際に社会には自分の利害, 欲得を離れて他人のために尽くす人はいくらでもいるではないか, そのような人々の行動モデルとして, 囚人のジレンマモデルは到底受け入れ難い, ということであろう. あるいはもっと端的に自分はそんなに利己的な人間ではないからこのモデルは到底受け入れ難い, と感じる, というものである. しかし, 「自分はそんなに利己的な人間ではない」という主張は他人から「彼は利己的な人間だ」と思われたくない, という結構利己的な見解ではないだろうか.

それはともかく, 物語とそれから連想される価値判断の背景になっている「基準」をいくつかのタイプに分けて整理してみよう.

タイプ I: 自分の利得のみを最大化する. (利己的個人合理性)

通常, ゲーム理論という場合はこの個人合理性を判断基準として用いている. しかし, 具体的な状況に適用すると何時の間にか少し違う判断基準を暗々裏に持ち込んでいる場合があるように思われるので注意が必要である. その主なものは次のような基準である.

タイプ II: 自分が蒙る可能性のあるリスクを最小にする. (リスク回避型, ミニマックス原理)

タイプ III-a: 相手より相対的に優位に立ちたい. (相対的個人合理性？)

タイプ III-b: 相手より劣位に立ちたくない. (平等志向型, 日本人に多いように思われる. いわゆる足の引っ張り合いとなり, 利己的個人主義よりもかえって始末が悪い場合もある.)

タイプ IV-a:相手との合計利得を最大化する. (全体合理性)

タイプ IV-b:自分の利得と相手の利得の双方を最大化する. (互恵的利他主義)

このタイプ IV - b の場合は, 判断基準が二次元となるので, 選好を決められないケースが出てくる. 数学的には全順序集合にならないから当然である. 日本人には優柔不断な人が多いように感じられるが好意的に解釈するとこのタイプの価値判断をしているのかもしれない.

タイプ V: 相手の利得のみを最大化する. (自己犠牲的)

真に自己犠牲的な選択が有り得るか, という問題は議論し始めるとやっかいである. 一見自己犠牲的と世間が思っている行動, 例えば, 母親の子供に対する愛情, 殉死, 殉教, 自爆テロ等はいずれもタイプ I の立場からの説明が可能である. 今ここで基準にしたのは客観的な (実証可能な) 立場からの批判である. しかし, プレーヤーを本当の個人に限った場合は, 当人が合理的であると思って適用する判断基準を傍から非合理的だ, と言ってみても始まらない. つまり, 合理的基準とは当人が合理的だ, と考える「主観的合理性」なのだろうか. ここに自然科学を科学するときと, 社会科学を科学する時の決定的違いがある. 私自身は, タイプ I の利己的個人合理性のみがあくまで, 客観的に実証し得る判断基準であるべきだ, と考えている. 利己的な合理的個人を想定するのは近代経済学の基本で, 社会学には馴染まない, という根強い批判はもっともであるが, それでは学問にならない (という学問観を持っている). 文系–理系の融合, 統合, 対話を言うならば共通点は多いほど よい. それでは, 主観的合理性をどのように客観的に分析すればよいであろうか. 利己的個人合理主義を批判する多くの論者が持ち出す判断基準は上記のタイプ II から IV のいずれかの立場からの場合が多い. 従って, 以下でもう少し詳しく議論するように, 主観的にタイプ II 以下の判断基準に従った場合は, 利得行列を変換した上でタイプ I の利己的個人合理性の基準を採用して分析すればよいのである.

以上プレーヤー本人の価値基準を 7 通りに分類したが, ゲーム理論では原理的に相手も自分と同じ程度に合理的である, と仮定しているから, 上記のように価値基準のタイプを分けると, 相手のプレーヤーがどのような価値基準の持ち主であるかがわからないと合理的な選択が出来ないことになる. そうすると, 単純に場合の数を計算しても 2 人ゲームで

も (自分と相手は区別しないとして) 価値基準として $7+7\times 6/2=28$ 通りの組合せがる. さらに悪いことに相手のことを確定的に知ることは一般的には不可能であるから, 相手の合理性を疑いだすと大変である. その場合は, 相手がどのタイプであるかを推定する主観確率 (確信度) を導入すればよい (情報不完備ゲームとして定式化することが行われているが本稿では立ち入らない). 相手の合理性を仮定せずに自分の各選択肢毎に相手が自分にとって最悪の手を選ぶと想定してその中での最良の手を選択するのがタイプ II のリスク回避型であり, 相手もそう考える, と仮定するのがミニマックス原理である. ゼロサムゲームではこのタイプが採用されている. しかし, 局所的にゼロサムゲームであるような状況ではタイプ V の自己犠牲的価値基準を誰も絶対に取らない, とは言い切れないであろう. そのような場合はより広い状況設定でゲーム理論を考える必要がある.

補足: 上記の分類では 4 通りの戦略の組み合わせに対する利得はすべて異なると仮定したが, 利得が等しい場合も許すと場合の数はもっと増える. その中でよく知られている現実的なゲームは調整ゲーム (coordination game) と呼ばれているもので, 利得行列が次のように表される.

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

このゲームは $C=$ 協力, $D=$ 裏切り, という意味はなく, 二人の人間が出くわしたときに左に避けるか, 右に避けるか兎に角同じ行動を取れば衝突を避けることが出来るが, そうでないと衝突してしまう, という例で日常でもよく経験することである. 二人, 2 戦略対称ゲームの中では, 囚人のジレンマ・ゲーム, チキンゲームとこの調整ゲームが一番現実感があるように思われる.

§4. ゲーム理論の定式化—標準形—

ここで, 少し一般的な形でゲーム理論を定式化する. この節で解説するのはいわゆる標準形ないしマトリックスゲームと言われている 1 回きりのゲームである. 1 回きりのゲーム, というのは若干妙なネーミングであるが, 要するにゲームの構造, ルール, 全てのプレーヤーの利得関数に関する情報を全てのプレーヤーが共有し, かつ, そのことをまた全てのプレーヤーが知っている (ということを全てのプレーヤーが知っている, ということを...) という情報完備なゲームである. その上に自分も相手も合理的判断をする, と仮定しているから, 例えば囚人のジレンマのようなゲームの場合は, 裏切るのが最良の戦略である, という結論は直ちに導かれる. 従って, なぜこのゲームを「ジレンマ」と呼ぶのか理解できない. 考察の枠組みをこのように限定すると, 論理的には自明な結論しか導かれない. 何故,「ジレンマ」と呼ぶか, という理由は, 実験的に囚人のジレンマ・ゲームをやらせてみると, 被験者が理論が教える通りの選択をするとは限らない場合が多い (つまり「協力」する) , という現実が観察されるためであろう. だから, 人間は合理的な存在ではない, と言ってしまえばみもふたもないわけで, 社会学を学問にするためには, 何とかより高次の合理性でこれらの現象を説明する必要がある. 現実の社会では純粋に 1 回きりのゲームで表現される現象は少なくて, 多くの場合, 暗黙のうちに同じゲームが繰り返されることが想定されている. 従って, より現実感のある理論として繰り返しゲームが提唱されている. 繰り返しゲームに関しては確率過程を用いた定式化をすると数学的にはすっきりする. 従ってこの講義録では 8 節で論じることにする.

この節では一通り伝統的な定式化に従うことにする. 判断基準として, 各自は自分の利得をより多くする戦略の方がベターであると考えている, という利己的個人主義 (タイプ I) を前提にする. かつ, 相手も同じ価値基準を持っているとお互いに信じている, ということも仮定する (共有知識の仮定, 情報完備ゲームという).

n 人のプレーヤーの集合を N とする. $N = \{1, 2, \ldots, n\}$ とする. プレーヤー i は戦略集合 S_i の中から 1 つの戦略 (純粋戦略) を選ぶ. プレーヤー $i \in N$ が純粋戦略 $s_i \in S_i$ を選んだとすると, 全員の手は戦略セット $\vec{s} = (s_1, \ldots, s_n) \in S^{(n)} \equiv S_1 \times \cdots \times S_n$ で表される. ここで, S_i はプレーヤー i が選び得る純粋戦略の全体であって, 後で述べるようにある戦略を確率的に選ぶ (混合戦略という) ことを認めるためには, 数学的には可測空間である必要がある. しかし, 後では位相の概念も必要になることと, 応用上はあまり抽象的な集合を考える必要がないので, 以下の仮定をおく.

仮定 4.1. $\forall i$, S_i は可分完備距離空間である.

このとき, $S^{(n)}$ も直積集合として自然に可分完備距離空間となる. 位相空間の場合, 可測集合としてはボレル可測集合を考える. ゲーム理論の多くの教科書は S_i が有限集合の時しか扱ってないが, 例えば持久戦のように $S_i = [0, +\infty)$ を考える必要があるから, 有限集合だけで定式化するとむしろゲーム理論の本質が見えてこないような気がする. S_i が有限集合の場合は (後には 2 点集合のみを考えるが) このような数学的概念を知っている必

要はないが, S_i が半直線の場合, 本当はより高度の数学的知識 (抽象ルベーグ式積分論) が必要である.

ゲーム理論ではもうひとつの概念, 各戦略セット $\vec{s}$ に対して, プレーヤー $i \in N$ の利得関数 (payoff function) $u_i(\vec{s})$ なるものが定義されている必要がある. さらに, プレーヤー全員は自分に関してはもちろん, プレーヤー全員の利得関数を知っているものと仮定する, 数学的には, 利得関数 u_i は各 i 毎に定義された $S^{(n)}$ から実数への可測関数である. 以下, さらに次の仮定をおく.

仮定 4.2. $\forall i,\ u_i$ は有界連続関数である.

プレーヤー i が選ぶひとつの混合戦略とは S_i に値を取るひとつの確率変数 X_i のことである. ただし, X_i の分布が単位分布の場合, つまり, $\exists s \in S_i,\ \mathrm{Prob}(X_i = s) = 1$ の時は習慣に従って, 純粋戦略 s を取る, という. 論理的には純粋戦略は混合戦略の 1 種である. 正方形は長方形の 1 種であるが, 正方形と分かっている時にわざわざ長方形とは言わないのと同じことである. 従って, 混合戦略セットとは $S^{(n)}$ に値を取る確率変数 $\vec{X} = (X_1, \ldots, X_n)$ のことである. その時のプレーヤー i の利得は確率変数 $u_i(\vec{X})$ となる. こうすると, 利得は 1 回毎にはランダムで確定しないから当然期待値を計算することになる. 期待値は数学的には抽象ルベーグ積分

$$E[u_i(\vec{X})] = \int_{S^{(n)}} u_i(s_1, \ldots, s_n) d\,\mu(s_1, \ldots, s_n)$$

である. ここで μ は $\vec{X}$ の分布である。期待値の現実的意味は, 全員が毎回独立に戦略セット $\vec{X}$ を選んだ場合 (数学的には独立確率変数列 $\vec{X}^{(k)},\ k = 1, 2, \ldots$ で各 $\vec{X}^{(k)}$ の分布が $\vec{X}$ の分布に等しい (i. i. d., independent identically distributed random variables) , 第 T 回までの試行の算術平均 (確率変数) で T を無限に大きくした極限, つまり

$$\lim_{T \to +\infty} \frac{1}{T} \sum_{k=1}^{T} u_i(\vec{X}^{(k)}(\omega)) = E[u_i(\vec{X})]$$

が, よく知られた強大数の法則から確率 1 で成立する, ということである. 従って, 期待値の大小の比較は同じゲームを何度もくり返して得られる平均利得 (算術平均) の大小によって決められる. つまり, 期待値は 1 回で実現するのではなくて, 無限回の試行によって初めて達成されるのである. 1 回毎のゲームにおいては期待値より大きいこともあるし, 小さいこともある. 確率的に考える以上は 1 回きりのゲームからは意味のある結論は得られない.

なお, $\vec{X}$ の各成分の確率変数 $X_1, \ldots, X_n$ は必ずしも独立である必要はない. Nash ([208]) がいう非協力ゲームとは $\vec{X}$ の各成分の確率変数 $X_1, \ldots, X_n$ が独立確率変数からなる場合のことである. プレーヤーどうしがある程度協力しあう状態を表現したければ, 独立性の仮定をゆるめてある種の従属性を仮定すればよい. その一つの試みとして, 本稿では 7 節で「内輪付き合いをするタカ・ハトゲーム」を考察した. 通常, 標準形ゲームという場合は暗黙の内に非協力, つまりプレーヤーの混合戦略は独立確率変数である, ということを仮定してあるが, 展開形ゲーム (本稿では触れない) の場合, 自然に過去に従属する混合戦

略を考察するから, 展開形ゲームが標準形ゲームで表現できる, という場合注意が必要である. 多くの教科書は混合戦略まで含めた対応関係は触れられていない.

確率論でよく知られているように期待値は確率変数の分布のみで決まるから, 期待値のみを議論する場合は分布の関数と考える方が都合がよい. そこで, 次の記号を導入する.

$$\mathcal{P}(S_i) = S_i \text{ 上の確率測度 (分布) の全体}, \quad \mathcal{P}(S^{(n)}) = S^{(n)} \text{ 上の確率測度の全体}$$

と置くと, プレーヤー i の利得の期待値は

$$E[u_i(\vec{X})] = \int_{S^{(n)}} u_i(s_1, \ldots, s_n)\, d\mu(s_1, \ldots, s_n)$$

と書ける. さらに, $X_1, \ldots, X_n$ が独立で各 X_k の分布を $\mu_k \in \mathcal{P}(S_k)$ とすると,

$$E[u_i(\vec{X})] = \int_{S^{(n)}} u_i(s_1, \ldots, s_n)\, d\mu_1(s_1) \ldots \mu_n(s_n)$$

となる. 以下では, 戦略セット $X_1, \ldots, X_n$ は独立確率変数であると仮定する. 従って, 分布としては

$$\mathcal{P}_0(S^{(n)}) = \left\{ \vec{\mu} = \prod_{i=1}^{n} \mu_i,\ (\mu_i; i = 1, \ldots n \text{ の直積測度}) \text{ の全体} \right\} = \prod_{i=1}^{n} \mathcal{P}(S_i),$$

を考えれば十分である.

なお, $\mathcal{P}(S_i)$, $\mathcal{P}(S^{(n)})$ 上には仮定 4. 1. の下でいわゆるプロホーロフ (Prokhorov) の距離が定義されて, 可分完備距離空間となることが知られている (数学辞典第 3 版, 44-G). なお, $\mathcal{P}_0(S^{(n)})$ は $\mathcal{P}(S^{(n)})$ の中の閉集合である.

ここで, $\vec{\mu}_0$, $\vec{\mu}_1$ を戦略セット $\vec{X}$, $\vec{Y}$ の分布とする.

$\mathcal{P}(S^{(n)}) \ni \vec{\mu}_0, \vec{\mu}_1$ に対して,

$$\vec{\mu}_\alpha = \alpha\vec{\mu}_1 + (1-\alpha)\vec{\mu}_0, \quad 0 \le \alpha \le 1$$

で新しい確率測度 $\vec{\mu}_\alpha \in \mathcal{P}(S^{(n)})$ が定義出来るから $\mathcal{P}(S^{(n)})$ は凸集合である.

プレーヤーの数が 2 で戦略集合 S_1, S_2 が有限集合の場合, S_i の元の個数を m_i とすると, 純粋戦略に対するプレーヤー i の利得表は $u_i(j,k)\,;\, s_j \in S_1, t_k \in S_2$ を (j,k) 成分とする $m_1 \times m_2$ 次行列で表され, この行列の性質によってゲームを次のように分類することが出来る (このようなゲームをマトリックス・ゲームと呼んでいる教科書もある).

定義 4.1.

$$\forall j = 1, 2, \ldots, m_1,\ \forall k = 1, 2, \ldots, m_2, \quad \sum_{i=1}^{2} u_i(j,k) = 0$$

が成立するとき, このゲームを 2 人零和ゲームという.

定義 4.2. $S_1 = S_2 = S$ として

$$\forall j, k \in S, \quad u_1(j,k) = u_2(k,j),$$

が成立するとき, このゲームを 2 人対称ゲームという. S は必ずしも有限集合に限る必要はない.

いずれの場合も, プレーヤー 1 の利得行列を与えれば条件からプレーヤー 2 の利得行列が分かるので本質的にはいずれのゲームであるかの仮定とひとつの $m_1 \times m_2$ 次行列または, m 次正方行列を与えればゲームは定まる.

§5. ナッシュ均衡戦略

最初に純粋戦略セットの全体, つまり, $S^{(n)}$ に利得関数を使って次のような同値関係を定義する.

$$\vec{s}=(s_1,\ldots,s_n)\cong\vec{t}=(t_1,\ldots,t_n)\ \text{ if and only if }\ \forall i,\ u_i(s_1,\ldots,s_n)=u_i(t_1,\ldots,t_n).$$

この 2 項関係が同値関係の公理を満たすことは自明であろう. 利得関数の値を問題にする限り, 全員の利得が変わらないような戦略セットは区別する必要がない. この同値関係によって同一視した戦略セットの全体を $S_0^{(n)}$ とする (つまり同値類の全体) . 同値類がひとつしか出来ない場合はゲームにならないから, 以下では常に同値類は二つ以上あると仮定する.

次に, $S_0^{(n)}$ 上に利得関数を使って次のような順序を定義する.

$$\vec{s}=(s_1,\ldots,s_n)\le\vec{t}=(t_1,\ldots,t_n)\ \text{ if and only if }\ \forall i,\ u_i(s_1,\ldots,s_n)\le u_i(t_1,\ldots,t_n).$$

この関係により $S_0^{(n)}$ は半順序集合となる.

定義 5.1. $S_0^{(n)}$ の極大元を パレート最適戦略 (Pareto optimum) という.

パレート最適戦略の直観的意味は, 全員が利得を減らさず, かつ少なくともひとりは利得を増やすような別の戦略セットはもはや存在しない, ということである. 現在, パレート最適戦略に達している時, これ以上自分の利得を増やしたいならば闘争あるのみである. 日本の社会 (特に大学は明白に) は, 多くの場合何もしない現状維持がパレート最適戦略であるから合理的に考えた場合, 全員が賛成する改革案, というものは存在しない.

$S_0^{(n)}$ に最大元が存在する時は, 前節でもみたように, 自分の利得を最大にすることにのみ関心がある場合はもはや争う必要がない. しかし, 相手より多くの利得を得たい, とか最悪の場合に陥りたくない, と考えた時には必ずしも自明なゲームではないから, このようなケースを除外すべきではないであろう (例えば, 3 節の (7) 囚人のジレンマ Part II) .

もうひとつ注意すべきことは, パレート最適戦略はナッシュ均衡戦略と違って, グローバルな概念である. なお, 以下では $S_0^{(n)}$ と $S^{(n)}$ をいちいち区別するのはわずらわしいし, 誤解の恐れもあまりないので $S^{(n)}$ を用い, 必要に応じて同一視を行う.

さて, $\mathcal{P}_0(S^{(n)})\ni\vec{\mu}=(\mu_1,\ldots,\mu_n)$ に対して

$$\begin{aligned}E[u_i(\vec{X})] &= \int_{S^{(n)}} u_i(s_1,\ldots,s_n)\,d\mu_1(s_1)\ldots\mu_n(s_n),\\ &\equiv u_i(\vec{\mu})\equiv u_i(\mu_1,\ldots,\mu_n)\end{aligned}$$

と書く. ここで, ゲーム理論特有の考察方法として, プレーヤー i の立場に立って, 相手全員の戦略は固定して自分だけが戦略を変更した場合に利得がどうなるか, を考察するために以下のような記号を導入する.

$$(\vec{\mu}_{-i}\,;\,\nu_i)=(\mu_1,\ldots,\mu_{i-1},\nu_i,\mu_{i+1},\ldots,\mu_n),$$
$$u_i(\vec{\mu}_{-i}\,;\,\nu_i)=u_i(\mu_1,\ldots,\mu_{i-1},\nu_i,\mu_{i+1},\ldots,\mu_n).$$

確率変数で表現する場合のために以下のような記号を導入する.

$$\begin{aligned}
\mathcal{L}(S_i) &= S_i \text{ に値を取る確率変数 } X_i \text{ の全体 } (i=1,2,\ldots,n),\\
\mathcal{L}(S^{(n)}) &= S^{(n)} \text{ に値を取る確率変数 } \vec{X}=(X_1,\ldots,X_n) \text{ の全体},\\
\mathcal{L}_0(S^{(n)}) &= \{\vec{X}=(X_1,\ldots,X_n)\in\mathcal{L}(S^{(n)}\,;\, X_1,\ldots,X_n \text{ が独立確率変数}\},\\
(\vec{X}_{-i}\,;\,Y_i) &= (X_1,\ldots,X_{i-1},Y_i,X_{i+1}\ldots,X_n)
\end{aligned}$$

混合戦略セットの全体 $\mathcal{P}(S^{(n)})$ や $\mathcal{P}_0(S^{(n)})$ の元に対しても, 平均利得関数 $u_i(\vec{\mu})\,;\, i=1,\ldots,n$ を利用して同値関係と順序を定義することが出来て, 純粋戦略の場合と同様に, 混合戦略に対してもパレート最適戦略が定義出来る. この場合もパレート最適戦略はグローバルな概念だから, ローカルな概念であるナッシュ均衡戦略とは直接関係はない. ただし, 自明な場合, つまり, 最大元 (ゲーム理論では優越戦略という) はナッシュ均衡戦略である.

非協力ゲーム理論の一般論における判断基準は如何に自分にとって利得の大きい戦略を見つけるか, ということである. かつ相手は自分とは独立に同様のことを考えるわけだから (談合は許されない) 結局相手が戦略を変えないという前提で自分にとってよりよい戦略を採らざるを得ない. 互いにそのように考えて, もうよりよい戦略が無くなった時は, これ以上自分の方から戦略を変えると損することはあっても真に得することはない, という状態になるであろう. このような戦略を ナッシュ(Nash) 均衡戦略という. つまり,

定義 5.2. $\vec{\mu}=(\mu_1,\ldots,\mu_n)$ がナッシュ均衡戦略であるとは,

$$\forall i,\ \forall\nu_i\in\mathcal{P}(S_i),\quad u_i(\vec{\mu})\geq u_i(\vec{\mu}_{-i}\,;\,\nu_i)$$

が成立する時をいう.

確率変数を用いて表現すると, $\vec{X}\in\mathcal{L}_0(S^{(n)})$ がナッシュ均衡戦略であるとは,

$$\forall i,\ \forall Y_i\in\mathcal{L}(S_i)\text{ such that }(\vec{X}_{-i},\,Y_i)\in\mathcal{L}_0(S^{(n)}),\ E[u_i(\vec{X})]\geq E[u_i(\vec{X}_{-i}\,;\,Y_i)]$$

が成立する時をいう.

ナッシュ均衡戦略は S_i に適当な仮定を置くと (特に有限集合の時は必ず) 存在するが, (前記の同一視を行っても) 唯一とは限らない. ナッシュ均衡戦略どうしの中でよりよい選択というのは可能性があるわけであるが, ひとつのナッシュ均衡戦略で安定してしまうとナッシュ均衡戦略の定義から自分から戦略を変えると損だから, よりベターな戦略と分かっていても談合をしないで別のナッシュ均衡戦略にどうやってたどりつくかは重要な問題である (囚人のジレンマ Part II の例) . 誰かがたまたま損を承知で別の戦略を試みるとか, とんまな人が損得が分からないまま, ナッシュ均衡戦略以外の戦略をとるとかしないと世の中は安定したままよりよくはならない. つまり, よりよいナッシュ均衡戦略には到達出来ない. ドン・キホーテ的人間の存在価値はあるわけである. 意地悪く言えば相手にドン・キホーテ的戦略をとらせるように仕向ける戦略は全体の利益に適うこともある, ということである.

なお, パレート最適戦略と違って, ナッシュ均衡戦略はローカルな概念であることを示そう.

定理 5.1. $\vec{\mu} = (\mu_1, \ldots, \mu_n)$ がナッシュ均衡戦略であるための必要十分条件は,

$\forall i,\ \exists U_\varepsilon(\mu_i)\,(\mathcal{P}(S_i)$ における μ_i の ε-近傍), such that $u_i(\vec{\mu}) \geq u_i(\vec{\mu}_{-i}\,;\,\nu_i),\ \ \forall \nu_i \in U_\epsilon(\mu_i)$

が成立することである.

証明. 必要性は論理的に自明であるから, 十分であることを証明する. そのために, 対偶を証明する. $\vec{\mu}$ がナッシュ均衡戦略でないと仮定する. 定義から

$$\exists i,\ \exists \nu_i \in \mathcal{P}(S_i),\ \ \text{such that}\ \ u_i(\vec{\mu}) < u_i(\vec{\mu}_{-i}\,;\,\nu_i)$$

が成り立つ. そこで,

$$\nu_{i,\alpha} = (1-\alpha)\,\mu_i + \alpha\,\nu_i \ \ 0 \leq \alpha \leq 1,$$

とおくと, $\mathcal{P}(S_i)$ の中の距離の意味で $\lim_{\alpha \downarrow 0}\ \nu_{i,\alpha} = \mu_i$ かつ, $0 < \forall \alpha \leq 1$ に対して,

$$u_i(\vec{\mu}) < (1-\alpha)u_i(\vec{\mu}) + \alpha u_i(\vec{\mu}_{-i}\,;\,\nu_i) = i(\vec{\mu}_{-i}\,;\,\nu_{i,\alpha})$$

が成り立つ. 従って, 如何なる $\varepsilon > 0$ 近傍をとっても

$$u_i(\vec{\mu}) \geq u_i(\vec{\mu}_{-i}\,;\,\nu_i),\ \ \forall \nu_i \in U_\varepsilon(\mu_i)$$

が成立する, ということはない.

ナッシュ均衡戦略の特徴付けと存在定理.

ナッシュ均衡戦略の存在を保証する定理を証明するために, この節では仮定 4.1. よりもう少し強い次の仮定のもとに考察する.

仮定 5.1. $\forall i,\ S_i$ はコンパクト距離空間である. (コンパクト距離空間は常に可分, 完備である) .

一般に可測空間 $(S, \mathcal{F})$, (ただし, $\forall s \in S$ に対して一点 s は可測集合とする) において, $s \in S$ 上の単位分布を δ_s と書く.

さらに強く $\forall i,\ S_i$ が有限集合であると仮定して考察する. $s \in S_i,\ \vec{\mu} \in \mathcal{P}_0(S^{(n)})$ に対して

$$\xi_{i,s}(\vec{\mu}) = \max\{0,\ u_i(\vec{\mu}_{-i}\,;\,\delta_s) - u_i(\vec{\mu})\},\ \ \xi_i(\vec{\mu}) = \sum_{s \in S_i} \xi_{i,s}(\vec{\mu})$$

とおく. max の定義と S_i が有限集合だから $0 \leq \xi_i(\vec{\mu}) < +\infty$ であることに注意する.

$$\mu_i^* = \frac{\mu_i + \sum_{s\in S_i} \xi_{i,s}(\vec{\mu})\delta_s}{1+\xi_i(\vec{\mu})}, \quad \vec{\mu^*} = (\mu_1^*, \ldots, \mu_n^*) \in \mathcal{P}_0(S^{(n)}),$$

とおいて, $\mathcal{P}_0(S^{(n)})$ から $\mathcal{P}_0(S^{(n)})$ への写像

$$\varphi(\vec{\mu}) = \vec{\mu^*} \tag{5.1}$$

を定義する. $\mathcal{P}_0(S^{(n)})$ は有限次元ユークリッド空間の凸閉集合と同相である. この写像が連続写像であることは容易にわかる. $\vec{\mu}$ がナッシュ均衡戦略であればすべての i に対して $\xi_i(\vec{\mu}) = 0$ であることは明らかである. 逆に, $\mathrm{supp}[\mu_i] \ni \forall s$ に対して, $\xi_{i,s}(\vec{\mu}) > 0$ ということはあり得ないから, $\vec{\mu^*} = \vec{\mu} \Leftrightarrow \forall i,\ \forall s \in S_i,\ \xi_{i,s}(\vec{\mu}) = 0$ である. 従って,

定理 5.2. $S^{(n)}$ が有限集合の時, $\vec{\mu}$ がナッシュ均衡戦略であるための必要十分条件は

$$\varphi(\vec{\mu}) = \vec{\mu}$$

を満たすことである. つまり, ナッシュ均衡戦略とは連続写像 φ の不動点に他ならない.

定理 5.3. $S^{(n)}$ が有限集合の時, ナッシュ均衡戦略は少なくともひとつ存在する.

証明. 有限集合 S_i の要素の個数を n_i とすると, 任意の $\mu_i \in \mathcal{P}(S_i)$ は

$$\mu_i = \sum_{j=1}^{n_i} p_j \delta_j, \quad \sum_{j=1}^{n_i} p_j = 1, \quad \forall j,\ p_j \geq 0,$$

ここで, δ_j は点 $s_j \in S_i$ 上の単位分布で一意に表されるから, $\mathcal{P}(S_i)$ は n_i 次元ユークリッド空間の超平面上の閉凸集合

$$\Big\{(p_1, \ldots, p_{n_i})\,;\, \sum_{j=1}^{n_i} p_j = 1, \quad \forall j,\ p_j \geq 0\Big\}$$

と同相である. 従って, $\mathcal{P}_0(S^{(n)})$ も有限次元閉凸集合と同相で, よく知られたブローウェル (Brower) の不動点定理 ([35]) によって $\mathcal{P}_0$ 上の連続写像 φ には不動点が存在する. なお, 角谷の不動点定理はノイマンのゲーム理論における原論文の証明をすっきりさせるためにブローウェルの不動点定理を拡張したものであるが, ここでは角谷の不動点定理を用いない Nash([208]) の証明に従った.

次に, 一般の $S^{(n)}$ に対して, 仮定 5.1 の下で, $\vec{\mu} \in \mathcal{P}_0(S^{(n)})$ に対して,

$$m_i(\vec{\mu}) \equiv \sup_{s\in S_i} u_i(\vec{\mu}_{-i}\,;\, \delta_s) - u_i(\vec{\mu})$$

とおく. 確率分布による積分の性質から $m_i(\vec{\mu}) \geq 0$ である. さらに仮定5.1 のコンパクト性と u_i の連続性から sup は max に置き換えることができる. つまり, $u_i(\vec{\mu}_{-i}\,;\, \delta_{s_i}) - u_i(\vec{\mu}) = m_i(\vec{\mu})$ を満たすような $s_i = s_i(\vec{\mu}) \in S_i$ が存在する.

なお, 明らかに, $\vec{\mu}$ がナッシュ均衡戦略であることと, $\forall i,\ m_i(\vec{\mu}) = 0$ であることが同値である. 従って, 容易に次の定理が得られる.

定理 5.4. $\mu_i \in \mathcal{P}(S_i)$ に対して, 測度の台 (support) $\mathrm{supp}[\mu_i]$ を, $\mu_i(F) = 1$ を満たす最小の閉集合 F として定義する. また,

$$M_i(\vec{\mu}) \equiv \{s \in S_i\,;\, u_i(\vec{\mu}_{-i}\,;\, \delta_s) = \sup_{t \in S_i} u_i(\vec{\mu}_{-i}\,;\, \delta_t)\}$$

と定義する. 仮定 5.1 の下では $\vec{\mu} \in \mathcal{P}_0(S^{(n)})$ がナッシュ均衡戦略であるための必要十分条件は

$$\forall i\,,\ M_i(\vec{\mu}) \supset \mathrm{supp}[\mu_i]$$

が成り立つことである.

純粋戦略が有限集合でない時, ナッシュ均衡戦略の存在定理の拡張はどこまで可能であろうか. 多くの教科書は有限集合の場合しか扱ってないが, ([53]),142 頁には次の定理が証明なしに述べてある. 数学的にはスタンダードな方法で有限次元の場合から近似することによって証明出来る. まず, 次の補題を証明する.

補題 5.1. 仮定 5.1 の下では $m_i(\vec{\mu})$ は $\mathcal{P}_0(S^{(n)})$ 上で定義された実関数として連続関数である.

証明. 上記で注意したように

$$m_i(\vec{\mu}) = u_i(\vec{\mu}_{-i}\,;\, \delta_{s_i(\vec{\mu})}) - u_i(\vec{\mu})$$

と表される. $\vec{\mu}, \vec{\nu}$ に対して, $m_i(\vec{\mu}) - m_i(\vec{\nu}) \geq 0$ と仮定する (逆向きの不等式が成立つ場合は, 以下の式で $\vec{\mu}$ と $\vec{\nu}$ の役割を逆にして考えればよい) .

$m_i(\vec{\nu})$ の定義より

$$\forall t \in S_i \quad m_i(\vec{\nu}) \geq u_i(\vec{\nu}_{-i}\,;\, \delta_t) - u_i(\vec{\nu})$$

だから, $t = s_i(\vec{\mu})$ とおいて,

$$\begin{aligned} 0 &\leq m_i(\vec{\mu}) - m_i(\vec{\nu}) \\ &\leq (u_i(\vec{\mu}_{-i}\,;\, \delta_{s_i(\vec{\mu})}) - u_i(\vec{\mu})) - (u_i(\vec{\nu}_{-i}\,;\, \delta_{s_i(\vec{\mu})}) - u_i(\vec{\nu})) \\ &\leq |u_i(\vec{\mu}_{-i}\,;\, \delta_{s_i(\vec{\mu})}) - u_i(\vec{\nu}_{-i}\,;\, \delta_{s_i(\vec{\mu})})| + |u_i(\vec{\mu}) - u_i(\vec{\nu})| \end{aligned}$$

なる評価式が得られる. $u_i(\vec{\mu}_{-i}\,;\, \delta_s)$ は $\mathcal{P}_0(S^{(n)}) \times S_i$ 上の連続関数であり, コンパクト性から一様連続である. 従って, 第一項は $\vec{\mu}, \vec{\nu}$ を十分近くとればあらかじめ任意に与えた $\varepsilon > 0$ 以下にできる. $u_i(\vec{\mu})$ も $\vec{\mu}$ に関して連続であるから同様である.

定理 5.5. 仮定 4.2 と仮定 5.1 の下でナッシュ均衡戦略が少なくともひとつ存在する.

証明. 各 S_i が可分空間であるから, $k = 1, 2, \ldots$ に対して有限集合 $\Lambda_{i,1} \subset \Lambda_{i,2} \subset \cdots \subset S_i$ が存在して $\Lambda_i = \cup_{k=1}^{\infty} \Lambda_{i,k}$ が S_i の dense subset であるように出来る. $\Lambda_{i,k}$ 上の確率分布の全体を $\mathcal{P}(\Lambda_{i,k})$ とおくと, $\mathcal{P}(\Lambda_{i,k})$ はもちろん $\mathcal{P}(S_i)$ の部分集合であり, さらに $\cup_{k=1}^{\infty} \mathcal{P}(\Lambda_{i,k})$ は $\mathcal{P}(S_i)$ の dense subset である. $\Lambda_{i,k} = S_i$ とおいて, 定理 5.4 を適用すると,

$$\forall i,\ \exists \vec{\mu}_k \in \mathcal{P}_0(\Lambda_{1,k} \times \cdots \times \Lambda_{n,k}) \text{ such that } m_{i,k}(\vec{\mu}_k) = 0$$

がいえる. ただし, $\vec{\mu} \in \mathcal{P}_0(\Lambda_{1,k} \times \cdots \times \Lambda_{n,k})$ に対して,

$$\begin{aligned} m_{i,k}(\vec{\mu}) &\equiv \sup_{s_i \in \Lambda_{i,k}} u_i(\vec{\mu}_{-i}\,;\, \delta_{s_i}) - u_i(\vec{\mu}) \\ &= u_i(\vec{\mu}_{-i}\,;\, \delta_{s_{i,k}(\vec{\mu})}) - u_i(\vec{\mu}) \end{aligned}$$

で定義した. $s_{i,k} = s_{i,k}(\vec{\mu})$ は前にも注意したように $\vec{\mu}$ に関係して決まる $\vec{\mu}$ の関数である.

ここで, $\mathcal{P}_0(S^{(n)})$ のコンパクト性から, $\mathcal{P}_0(S^{(n)})$ のなかで収束するような $\{\vec{\mu}_k\,;\, k = 1, 2, \ldots, \}$ の部分列 $\{\vec{\mu}_{k_j}\,;\, j = 1, 2, \ldots, \}$ が存在するから,

$$\lim_{j \to \infty} \vec{\mu}_{k_j} = \vec{\nu} \in \mathcal{P}_0(S^{(n)})$$

とおく. $u_i(\vec{\mu}_{-i}\,;\, *)$ と $u_i(\vec{\mu})$ は $\mathcal{P}_0(S^{(n)})$ 上の連続関数であり, 一方, $m_{i,k}$ の定義より,

$$\forall s \in \Lambda_{i,k},\ (k \leq k_j),\quad 0 = m_{i,k_j}(\vec{\mu}_{k_j}) \geq u_i((\vec{\mu}_{k_j})_{-i}\,;\, \delta_s) - u_i(\vec{\mu}_{k_j})$$

だから,

$$\forall i,\ \forall s \in \Lambda_i = \cup_{k=1}^{\infty} \Lambda_{i,k},\quad 0 \geq \lim_{j \to \infty} u_i((\vec{\mu}_{k_j})_{-i}\,;\, \delta_s) - u_i(\vec{\mu}_{k_j}) = u_i(\vec{\nu}_{-i}\,;\, \delta_{s_i}) - u_i(\vec{\nu}).$$

さらに, Λ_i が S_i で dense subset であることと u_i の連続性から,

$$\forall s \in S_i,\quad 0 \geq u_i(\vec{\nu}_{-i}\,;\, \delta_s) - u_i(\vec{\nu})$$

を得る.

以上のことより,

$$\forall i,\quad m_i(\vec{\nu}) = \sup_{s \in S_i} u_i(\vec{\nu}_{-i}\,;\, \delta_s) - u_i(\vec{\nu}) = 0$$

つまり, $\vec{\nu}$ がナッシュ均衡戦略であることがいえた.

定理 5.6. 定理 5.5. と同じことを仮定する. 2 人対称ゲームにおいては, 対称な戦略 $\vec{\mu} = (\mu, \mu)\,,\ \mu \in \mathcal{P}(S)$ でナッシュ均衡戦略となるものが存在する.

証明. 対称ゲームであるから, $u_1(\mu, \nu) = u_2(\nu, \mu)$ が常に成立していることに注意すると, 対称な戦略 $\vec{\mu} = (\mu, \mu)$ がナッシュ均衡戦略であるための必要十分条件は

$$\forall \nu \in \mathcal{P}(S),\quad u_1(\mu, \mu) \geq u_1(\nu, \mu)$$

が成立することである.

$$m(s,\,\mu) = \max\{0,\, u_1(\delta_s, \mu) - u_1(\mu, \mu)\}$$

とおいて, 以下定理 5.3 と定理 5.5 の証明と同様の手順を踏めば証明できる.

上記の定理の証明ではコンパクト性が基本的に効いている. 従って, u_i が有界連続関数であってもコンパクト性を仮定しなかった時, ナッシュ均衡戦略が存在しないような例は容易に構成出来る. 例えば, 2 人対称ゲームにおいて $S = [\,0, +\infty)$ として, $u_1(x, y) = \dfrac{x+y}{1+x+y}$ とすると, $\mu \in \mathcal{P}(S)$ に対して, $\mu_h(A) = \mu(A + h),\ (h > 0)$ で新しい分布 μ_h を定義すると,

$$\begin{aligned} u_1(\mu, \nu) &= \iint_{[\,0,+\infty)\times[\,0,+\infty)} u_1(x, y)\, d\mu(x)\, d\nu(y) \\ &= 1 - \iint_{[\,0,+\infty)\times[\,0,+\infty)} \frac{1}{1+x+y}\, d\mu(x)\, d\nu(y) \\ &< 1 - \iint_{[\,0,+\infty)\times[\,0,+\infty)} \frac{1}{1+x+h+y}\, d\mu(x)\, d\nu(y) \\ &= 1 - \iint_{[\,0,+\infty)\times[\,0,+\infty)} \frac{1}{1+x+y}\, d\mu_h(x)\, d\nu(y) = u_1(\mu_h, \nu) \end{aligned}$$

と出来るから, 如何なる戦略セット (μ, ν) もナッシュ均衡戦略になり得ない.

ナッシュ均衡戦略の例:

2 人 2 戦略対称ゲームのナッシュ均衡戦略を求める.
2 次の正方行列 $A = (a_{i,j})$ をプレーヤー A が戦略 i を, プレーヤー B が戦略 j をとった時の, プレーヤー A の利得行列とする. 各人の戦略は

$$\mu = p_1\delta_1 + p_2\delta_2$$

ここで, $\delta_i\ (i = 1, 2)$ は純粋戦略 i 上の単位分布, と表されるから混合戦略は 2 次元のベクトル $\vec{p} = (p_1, p_2),\ p_1 + p_2 = 1,\ p_1, p_2 \geq 0$ で一意に表される. このようなベクトルの全体を V_2 とおくと, V_2 は区間 $[\,0, 1\,]$ と同一視することが出来る. A, B がそれぞれ, 戦略 $\vec{p}$, $\vec{q}$ をとった時の A, B の利得 $u_{\rm A}$, $u_{\rm B}$ を 2 次元の内積 $(\ ,\ )$ で表すと

$$u_{\rm A}(\vec{p}, \vec{q}) = \sum_{i,j=1}^{2} a_{i,j} p_i q_j = (\vec{p}, A\vec{q}),$$

$$u_{\rm B}(\vec{p}, \vec{q}) = u_{\rm A}(\vec{q}, \vec{p}).$$

従って, A と B の戦略セット $(\vec{p}, \vec{q})$ がナッシュ均衡戦略であるとは, 次の条件を満たすことである.

$$\text{(i)}\ \forall \vec{r} \in V_2,\ u_{\rm A}(\vec{p}, \vec{q}) \geq u_{\rm A}(\vec{r}, \vec{q}),$$

$$\text{(ii) } \forall \vec{r} \in V_2,\ u_{\mathrm{A}}(\vec{q},\vec{p}) \geq u_{\mathrm{A}}(\vec{r},\vec{p}).$$

書き換えると,

$$\text{(i) } \forall \vec{r} \in V_2,\ (\vec{p}-\vec{r}, A\vec{q}) \geq 0,$$

$$\text{(ii) } \forall \vec{r} \in V_2,\ (\vec{q}-\vec{r}, A\vec{p}) \geq 0$$

となる. ここで, $\vec{p}-\vec{r},\ \vec{q}-\vec{r}$ の成分の和は常に 0 になることに注意すると行列 A の各列に列毎に異なってもよい定数を加えても上記の内積の値は変わらないから, $a_{1,2}=a_{2,1}=0$ に帰着させてから計算すると便利である. 以下に 12 通りすべての場合を書き下す.

(1) 囚人のジレンマ

$$A = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

として条件を満たす $(\vec{p},\vec{q})$ を求めればよい. 具体的に計算すると,

$$\text{(i) } 0 \leq \forall r_1 \leq 1,\ (-p_1+r_1)q_1 + (-p_1+r_1)(1-q_1) \geq 0,$$

$$\text{(ii) } 0 \leq \forall r_1 \leq 1,\ (-q_1+r_1)p_1 + (-q_1+r_1)(1-p_1) \geq 0,$$

となる. それぞれ簡単な計算によって

$$\text{(i) } 0 \leq \forall r_1 \leq 1,\ \ r_1 \geq p_1 \ \Rightarrow\ p_1 = 0,$$

$$\text{(ii) } 0 \leq \forall r_1 \leq 1,\ \ r_1 \geq q_1 \ \Rightarrow\ q_1 = 0$$

となる. 従って, (D, D) (裏切り合い) が唯一のナッシュ均衡戦略であることがわかる.

以下同様に (2) から (12) の場合を計算してみる.

$$A = \begin{pmatrix} a_{11} & 0 \\ 0 & a_{22} \end{pmatrix}$$

とおくと, 上記の条件 (i), (ii) は一般に

$$\text{(i) } 0 \leq \forall r_1 \leq 1,\ (p_1-r_1)((a_{11}+a_{22})q_1 - a_{22}) \geq 0,$$

$$\text{(ii) } 0 \leq \forall r_1 \leq 1,\ (q_1-r_1)((a_{11}+a_{22})p_1 - a_{22}) \geq 0,$$

となる.

(2) チキンゲーム

$$a_{11} = -1,\ a_{22} = -1$$

$$\text{(i) } (p_1-r_1)(1-2q_1) \geq 0,\ \ \text{(ii) } (q_1-r_1)(1-2p_1) \geq 0.$$

(i) より

$$(a)\ q_1 = 1/2 \text{ , or } (b)\ q_1 > 1/2 \text{ and } p_1 = 0 \text{ , or } (c)\ q_1 < 1/2 \text{ and } p_1 = 1$$

が得られ, (ii) からは

$$(a')\ p_1 = 1/2\ ,\ \text{or}\ (b')\ p_1 > 1/2\ \text{and}\ q_1 = 0\ ,\ \text{or}\ (c')\ p_1 < 1/2\ \text{and}\ q_1 = 1$$

が得られる. (i) と (ii) が両立するためには (i) の 3 通りの場合と (ii) の 3 通りで 9 通りの組み合わせのうち, 矛盾しない組み合わせを求めると,

$$(p_1 = 1/2, q_1 = 1/2),\ (p_1 = 0, q_1 = 1),\ (p_1 = 1, q_1 = 0)$$

の 3 通りとなる. 従って 3 つのナッシュ均衡戦略が存在するがそのうちの混合戦略セットだけが対称であることに注意しておく.

さらに, 漸近的挙動を調べてみると興味深いことが分かる. 今, $p_1 > 1/2,\ q_1 < 1/2$ の状態にあるとすると, A にとっては, 自分がはますます p_1 を 1 に近づけ, B が q_1 を 0 に近づけてくれる戦略の方が有利になる. 同様に, $p_1 < 1/2,\ q_1 > 1/2$ の状態にあるときは, A はますます p_1 を 0 に近づけ, B は q_1 を 1 に近づける戦略の方が有利になる. 一方, $p_1 > 1/2,\ q_1 > 1/2$ または, $p_1 < 1/2,\ q_1 < 1/2$ にある時は, 双方とも確率を $1/2$ に近づける方がよい. チキンゲームであれば, 一度優劣がつくと, 強い方はますます強気になり, 弱い方はますます弱気にならざるを得ない. 双方とも強気, または双方とも弱気の時だけ, 徐々に対等な状態になり得る. いじめ問題が基本的に本人の主体性の確立にあることが分かるが, 一度優劣がついてからではなかなか, 対等な関係にもどれない, ということもよく理解されるであろう.

(3) 指導者ゲーム

$$a_{11} = -2,\ a_{22} = -2$$

であるから, ナッシュ均衡戦略は (2) の場合と完全に同じであるから,

$$(p_1 = 1/2, q_1 = 1/2),\ (p_1 = 0, q_1 = 1),\ (p_1 = 1, q_1 = 0).$$

(4) 英雄ゲーム

$$a_{11} = -1,\ a_{22} = -3$$

であるから, 基本的には (2) と同じである. 従って,

$$(p_1 = 3/4, q_1 = 3/4),\ (p_1 = 0, q_1 = 1),\ (p_1 = 1, q_1 = 0).$$

(5)

$$a_{11} = 1,\ a_{22} = -3$$

より,

$$\text{(i)}\ (p_1 - r_1)(3 - 2q_1) \geq 0,\quad \text{(ii)}\ (q_1 - r_1)(3 - 2p_1) \geq 0.$$

$(3-2q_1)>0$ より $p_1 \geq r_1$ が任意の r_1 について成立するいう条件より, $p_1=1$ を得る. 同様に $q_1=1$ が得られる. よってナッシュ均衡戦略は,

$$(p_1=1,\ q_1=1)$$

の一組のみ.

(6)

$$a_{11}=2,\ a_{22}=-2$$

条件式は (1) の囚人のジレンマと符号が逆になるだけだから, ナッシュ均衡戦略は

$$(p_1=1,\ q_1=1)$$

の一組のみ.

(7) 囚人のジレンマ Part II

$$a_{11}=1,\ a_{22}=1$$

条件式は (2) のチキンゲームと符号が逆になるだけで同じ. 従って,

$$(a)\ q_1=1/2 \text{ , or } (b)\ q_1>1/2 \text{ and } p_1=1 \text{ , or } (c)\ q_1<1/2 \text{ and } p_1=0$$

と

$$(a')\ p_1=1/2 \text{ , or } (b')\ p_1>1/2 \text{ and } q_1=1 \text{ , or } (c')\ p_1<1/2 \text{ and } q_1=0$$

を得る. ただし, 可能な組み合わせは少し異なり,

$$(p_1=1/2, q_1=1/2),\ (p_1=1, q_1=1),\ (p_1=0, q_1=0)$$

の 3 通りのナッシュ均衡戦略を得る. 注意することは 3 通りすべて対称なナッシュ均衡戦略である.

(8)

$$a_{11}=2,\ a_{22}=2$$

条件式は (7) とまったく同じだから,

$$(p_1=1/2, q_1=1/2),\ (p_1=1, q_1=1),\ (p_1=0, q_1=0)$$

の 3 通りのナッシュ均衡戦略を得る.

(9)

$$a_{11} = 3,\ a_{22} = 1$$

(7), (8) と同様であるが, 混合戦略のナッシュ均衡戦略の確率が異なる.

$$(p_1 = 1/4, q_1 = 1/4),\ (p_1 = 1, q_1 = 1),\ (p_1 = 0, q_1 = 0)$$

の 3 通りのナッシュ均衡戦略を得る.

(10)

$$a_{11} = 3,\ a_{22} = -1$$

条件より $p_1 \geq r_1$ が任意の r_1 について成立するという条件より, $p_1 = 1$ を得る. 同様に $q_1 = 1$ が得られる. よってナッシュ均衡戦略は,

$$(p_1 = 1,\ q_1 = 1)$$

の一組のみ.

(11)

$$a_{11} = 3,\ a_{22} = -1$$

(10) の場合とまったく同じだから, ナッシュ均衡戦略は

$$(p_1 = 1,\ q_1 = 1)$$

の一組のみ.

(12)

$$a_{11} = 2,\ a_{22} = -2$$

(6) の場合とまったく同じだから, ナッシュ均衡戦略は

$$(p_1 = 1,\ q_1 = 1)$$

の一組のみ.

プレーヤー A が戦略 C をとる確率を p_1, プレーヤー B が戦略 C をとる確率を q_1 として, 戦略セットを (p_1, q_1) で表して, 上記の (1) から (12) のタイプのナッシュ均衡戦略を一覧表にまとめるとつぎのようになる.

(1) (0, 0)
(2) (1/2, 1/2), (0, 1), (1, 0)
(3) (1/2, 1/2), (0, 1), (1, 0)
(4) (3/4, 3/4), (0, 1), (1, 0)
(5) (1, 1)
(6) (1, 1)
(7) (1/2, 1/2), (1, 1), (0, 0)
(8) (1/2, 1/2), (1, 1), (0, 0)
(9) (1/4, 1/4), (1, 1), (0, 0)
(10) (1, 1)
(11) (1, 1)
(12) (1, 1)

ナッシュ均衡戦略のタイプ別にゲームの方を分類すると,

(a) ナッシュ均衡戦略が, 対称な純粋戦略ただ一点のみとなるゲーム:
(1), (5), (6), (10), (11), (12),
(さらに, (1) とそれ以外は異なる戦略セットである.)

(b) ナッシュ均衡戦略が, 1 つの対称な混合戦略と, 2 つの純粋非対称な戦略セットからなるゲーム:
(2), (3), (4),

(c) ナッシュ均衡戦略が, 1 つの対称な混合戦略と, 2 つの純粋対称な戦略セットからなるゲーム:
(7), (8), (9).

以上の分類を眺めると, いずれの分類においても (1) 囚人のジレンマ, (2) チキンゲーム, (7) 囚人のジレンマ Part II は構造が異なることが分かる.

2 人 2 戦略対称ゲームの利得表で大小関係を考慮して 12 通りの分類を行ったが, ナッシュ均衡戦略の意味では 3 通りしかない ((a) を 2 つに分けても 4 通り). ナッシュ均衡戦略というのは相手に勝とうとは思わないで, 自分の利得を減らさないことにのみ判断基準を置くからである.

2 人 2 戦略対称ゲームの利得行列で値が一致する場合も含めて一般化して分類すると次のようになる.

$$A = \begin{pmatrix} a_{11} & 0 \\ 0 & a_{22} \end{pmatrix}, \quad a_{11} \geq a_{22}, \quad \text{but not } (a_{11}, a_{22}) \neq (0, 0).$$

とおくと, 条件は一般に

$$\text{(i) } 0 \leq \forall r_1 \leq 1,\ (p_1 - r_1)((a_{11} + a_{22})q_1 - a_{22}) \geq 0,$$

$$\text{(ii) } 0 \leq \forall r_1 \leq 1,\ (q_1 - r_1)((a_{11} + a_{22})p_1 - a_{22}) \geq 0,$$

となり, 次の 5 通りに分けられる.

1. $0 < a_{22} \leq a_{11}$,
2. $a_{22} < 0 < a_{11}$,
3. $a_{22} \leq a_{11} < 0$,
4. $a_{22} = 0 < a_{11}$,
5. $a_{22} < a_{11} = 0$.

上記の条件 (i), (ii) からナッシュ均衡戦略の求め方は具体例についてすでにやっているから, 結果だけをまとめると次のようになる. 前と同様にプレーヤー A が戦略 C をとる確率を p_1, プレーヤー B が戦略 C をとる確率を q_1 として, 戦略セットを (p_1, q_1) で表して, それぞれの場合のナッシュ均衡戦略を記すと,

1. $(\dfrac{a_{22}}{a_{11} + a_{22}}, \dfrac{a_{22}}{a_{11} + a_{22}}),\ (1, 1),\ (0, 0)$,
2. $(1, 1)$,
3. $(\dfrac{a_{22}}{a_{11} + a_{22}}, \dfrac{a_{22}}{a_{11} + a_{22}}),\ (1, 0),\ (0, 1)$,
4. $(1, 1),\ (0, 0)$,
5. $(p_1, 1);\ 0 \leq p_1 \leq 1,\ (1, q_1);\ 0 \leq q_1 \leq 1$.

この分類でみていると, 3 番目のゲームがいろいろな意味で面白い例を多く含んでいるように思われる. 囚人のジレンマ・ゲームは 2 番目のゲームで a_{11} と a_{22} の大小関係を逆にしただけである. 数学的一般化を不用意に行うと具体的な含意が抜け落ちてしまうよい例であろう. 数学者と応用数学者のセンスの違いはこのあたりにあるのかも知れない.

§6. ESS (進化的に安定な戦略) —タカ・ハトゲーム—

ゲーム理論は創設者ノイマンの人柄を反映して, 合理的基準によって合理的判断をする人間を想定しているため, 人間社会一般に適用することには当初から反発と批判があった. しかしながら, ゲームをするのは何も人間に限らない. 動物の生きるための日々の活動は自然に対して, また天敵や, 仲間どうしとの生き残りゲームであると考えることが出来る. 合理的判断基準を最初に設定しなくても, 結果として自分の遺伝子をどれだけ相手より多く残せるか (適応度が高い, という) を基準にすることによって, 人間を含む生物の生き残り戦略をゲーム理論によって解析することの有効性を示したのはメイナード・スミスである ([43]). 彼の解析のキー概念は ESS (Evolutionarily Stable Starategy) という, 対称ナッシュ均衡戦略より少し強い概念である.

ここである 1 種類の生物集団を考える. そして, 仲間どうしで例えば餌をめぐる闘争のゲーム, あるいはメスをめぐるオスどうしの闘争のゲーム (この場合はオスだけを考える) を想像してみよう. 生物が安定して存在して来た以上, 現在彼らが採用している戦略はそれなりに合理的なはずである. それは, どういう意味においてであろうか. 生物の場合, 自分の意志によって現在の戦略を採用しているわけではない, という意味で合理的判断力を持たない生物にゲーム理論を適用するのは不適当のように思われるかも知れない. しかし, 生物の行動が大部分遺伝子によって規定されているとはいえ, なぜそのような遺伝子が生き残ってきたか, という視点に立つと, より適応度の高い戦略を採用した遺伝子は長い年月の自然淘汰のなかでより生き残る確率が高かったであろうと考えられる. そして, 遺伝子は自己複製するとはいえ完全に親と同じ遺伝子を引き継ぐわけではない. その時, 偶然に適応度の高い戦略を採用するような遺伝子が生まれるとその遺伝子はますます集団の中に広がってゆきやがて安定多数になるであろう.

以上の考え方を彼の本に従ってもっとも単純なタカ・ハトゲームについて説明しよう. 数学的には 2 人 2 戦略対称ゲームであるが, ESS 概念はむしろ戦略そのものに対する概念でプレーヤーは補助的な役割を果たすに過ぎない. これに対してナッシュ均衡戦略は各人の利得関数とこみにした概念であることに注意する. 今, 2 種類の純粋戦略, タカ戦略 (H = Hawk) とハト戦略 (D = Dove) およびその混合戦略をとるものとする. この時, H または D の戦略を取る個体 (A とする) が戦略 H または D を取る個体と対戦した時の, A の利得行列が

		相手の戦略 H	相手の戦略 D
A の戦略	H	$\frac{1}{2}(V-C)$	V
	D	0	$\frac{1}{2}V$

で表されるとする. ただし, $V>0$, $C>0$. ここで, 動物行動学 (行動生態学) における利得とは簡単に言って, 自分の子孫を如何に多く残せるか (適応度という) , ということを基準にする. 従って, 必ずしも相手に勝つ必要はない. もちろん, このような単純な考え方では説明出来ないような自己犠牲的と見える行動が動物にも見られるがそのような場合も多くは血縁で繋がっており, 共通の遺伝子を出来るだけ多く子孫に伝える (包括適応度と

いう) ために有利な戦略は何かと考えることによってゲーム理論の枠内で考えることが出来る.

利得表の意味は, 自分がタカ戦略の時, 相手がハト戦略をとってくれると, 自分は利得 $V > 0$ を得るが, 逆に自分が争いを嫌ってハト戦略をとった時, 相手がタカ戦略をとって来ると, 残念ながら利得は何もない. では, 自分もタカ, 相手もタカ, という時はどうなるか. 争いになって結果として共にダメージ $C > 0$ を受ける. 結果については, 互いに対等であると暗に仮定しているので, 勝負は五分五分であるとして, 平均利得 $(V - C)/2$ を得るとする. このあたりは単純化しすぎのきらいはあるが, 確率過程の概念を導入するのは早すぎるので, 1 回毎に $(V - C)/2$ の利得を得る, と考える. また, 互いにハト戦略をとった時は, 共に利益がないこともあるが, ここでは餌を争わないで分け合うと考えて, 互いに $V/2$ の利得がある, とする.

容易に分かるように数学的構造としては $V - C$ の正負, つまり, 得られる利得と争った時のダメージの比較が重要である. 3 節の分類に従うとタカ・ハトゲームは次のように分類される.

(i) $V > C$ の場合. この場合, 戦略セット (H, D) を (D, C) と読み代えると囚人のジレンマ・ゲームと同じ構造になる.

(ii) $V < C$ の場合. この場合は同様の読み代えによって (2) のチキンゲームと同じ構造になる.

今, 自分 (A) が混合戦略 μ , 相手が混合戦略 ν をとった時の A の利得を $u(\mu\,,\nu)$ とおく (従って相手の利得は $u(\nu\,,\mu)$). 今の場合, $\mu\,,\nu$ は 2 点集合 $H\,,D$ 上の確率測度であることから,

$$\mu = p_1\delta_H + p_2\delta_D,\quad p_1 + p_2 = 1,\quad p_1, p_2 \geq 0,$$

$$\nu = q_1\delta_H + q_2\delta_D,\quad q_1 + q_2 = 1,\quad q_1, q_2 \geq 0,$$

と表されるから, $\mu\,,\nu$ の代わりにベクトル $\vec{p} = (p_1\,,p_2)$, $\vec{q} = (q_1\,,q_2)$ で表し, このようなベクトル全体を V_2 と記す. この時,

$$u(\vec{p}, \vec{q}) = \frac{1}{2}(V - C)p_1q_1 + Vp_1q_2 + \frac{1}{2}Vp_2q_2.$$

ここで, 現在ある戦略 $\vec{p}$ をこの生物集団全員が採用しているとする. この戦略が一番よいかどうか生物自身が理解しているわけではないから, たまたま突然変異等で別の戦略 $\vec{q}$ をとる仲間が現れたとする. 戦略 $\vec{p}$ をとっているグループ A に, 戦略 $\vec{q}$ をとるグループ B が侵入してくる, と考える. グループ A とグループ B の数の割合は $1 - \varepsilon$ 対 ε とする. ここで, ε は小さい数とする. つまり, グループ A は多数派, グループ B は少数派の侵入者と考える. 全個体はランダムに遭遇し 2 者間で闘争が行われ, 上に述べた利得表に従って利得を得て, 利得の高いグループの方が数を増し, 利得の低い方はやがて絶滅する, と考える. このように考えると, A グループの個体が A グループの一員と出会う確率は $1 - \varepsilon$, B グループの一員と出会う確率は ε となり, その時戦略 μ を取っている個体の平均利得

$E(\vec{p}, \vec{q})$ は

$$E(\vec{p}, \vec{q}) = (1-\varepsilon)u(\vec{p}, \vec{p}) + \varepsilon u(\vec{p}, \vec{q})$$

となり, 同様に 戦略 ν を取る個体の平均利得 $E(\vec{q}, \vec{p})$ は

$$E(\vec{q}, \vec{p}) = (1-\varepsilon)u(\vec{q}, \vec{p}) + \varepsilon u(\vec{q}, \vec{q})$$

となる. 侵入者がどのような戦略 $\vec{q} \neq \vec{p}$ をとっても

$$\forall \vec{q} \neq \vec{p}\ , E(\vec{p}, \vec{q}) > E(\vec{q}, \vec{p}) \tag{6.1}$$

が成立するときは, 長期的にみて侵入に失敗するであろうと考えられるから, (6.1) を書き換えると,

$$(1-\varepsilon)u(\vec{p}, \vec{p}) + \varepsilon u(\vec{p}, \vec{q}) > (1-\varepsilon)u(\vec{q}, \vec{p}) + \varepsilon u(\vec{q}, \vec{q}) \tag{6.2}$$

を得る. $\vec{p}_\varepsilon = (1-\varepsilon)\vec{p} + \varepsilon\vec{q}$ とおいて, 積分の線形性に注意すると (6.2) は

$$\forall \vec{q} \neq \vec{p},\ u(\vec{p}, \vec{p}_\varepsilon) > u(\vec{q}, \vec{p}_\varepsilon)$$

と同値である. 侵入者は当初は個体数が少ないであろうから, $\varepsilon > 0$ は小さいと考えてよい. この時, 侵入者はやがて消滅して, 多数派は安泰ということになる. このような戦略 $\vec{p}$ を進化的に安定な戦略 (ESS) と呼ぶ. 正確には,

定義 6.1. 戦略 $\vec{p}$ が ESS であるとは,

$$\forall \vec{q} \neq \vec{p},\ \exists \varepsilon_0 > 0\,,\ 0 < \forall \varepsilon < \varepsilon_0,\ (1-\varepsilon)u(\vec{p}, \vec{p}) + \varepsilon u(\vec{p}, \vec{q}) > (1-\varepsilon)u(\vec{q}, \vec{p}) + \varepsilon u(\vec{q}, \vec{q})$$

が成立する時をいう.

定義から容易に分かるように,

定理 6.1. 戦略 $\vec{p}$ が ESS であるための必要十分条件は $\vec{p} \neq \forall \vec{q} \in V_2$ に対して
(i) $u(\vec{p}, \vec{p}) > u(\vec{q}, \vec{p})$
or
(ii) $u(\vec{p}, \vec{p}) = u(\vec{q}, \vec{p})$ and $u(\vec{p}, \vec{q}) > u(\vec{q}, \vec{q})$
が成り立つことである.

証明. $0 < \varepsilon \leq 1$ で表される線分を眺めて見よ.

タカ・ハトゲームについて計算してみると, (i) は

$$u(\vec{p}, \vec{p}) - u(\vec{q}, \vec{p}) = \frac{V-C}{2}(p_1 - q_1)p_1 + V(p_1 - q_1)p_2 + \frac{V}{2}(p_2 - q_2)p_2$$

$$= (p_1 - q_1)(\frac{V}{2} - \frac{C}{2}p_1) > 0. \tag{6.3}$$

となる. ここで, 2 つの場合に分かれる.

(i) $V > C$ の場合, $0 \leq p_1 \leq 1$ だから, $0 \leq \forall q_1 \leq 1$ について, (6.3) が成り立つためには $p_1 = 1$ が必要十分である. つまり, タカ戦略が ESS である. 闘って利得がある場合は闘う戦略を選んだ方 (偶然にしろ) の生物種が生き残って来ている, ということである.

(ii) $0 < V \leq C$ の場合は, $0 \leq \forall q_1 \leq 1$ に対して, (6.3) が成り立つということはない. そこで, 定理 6.1. の (ii) の条件をチェックしてみる. 明らかに, $(\frac{V}{2} - \frac{C}{2}p_1) = 0$ ならば, 前半の条件は成立する. $p_1 = V/C$ として後半の式を計算する.

$$u(\vec{p}, \vec{q}) - u(\vec{q}, \vec{q}) = \frac{(V - Cq_1)^2}{2C} > 0 \quad (\forall q_1 \neq p_1).$$

よって, この場合は, タカ戦略を確率 V/C , ハト戦略を確率 $1 - V/C$ でとるのが ESS であることが分かる.

ここで (ii) の場合に, 多型安定と ESS の違いについて少しふれておく. 今, タカ戦略をとるグループ A が全体の p_1, ハト戦略をとるグループ B が全体の $1 - p_1$ だけいて, ランダムに遭遇してゲームを行いそれぞれの利得を得たとする. A の平均利得は $p_1(V - C)/2 + (1 - p_1)V$ となり, B の平均利得は $(1 - p_1)V/2$ となる. 両者が等しくなる場合は

$$p_1(V - C)/2 + (1 - p_1)V = (1 - p_1)V/2$$

より, $p_1 = V/C$ である. さらに, p_1 が V/C より大きくなると A が不利になり, 小さくなると逆に有利になって, いずれにしろ V/C の割合にもどる方が有利である (多型安定という). 従って, 全員が ESS 混合戦略を採用している, と解釈しても全体の V/C だけの個体がタカ戦略, 残りがハト戦略を採用している, と考えても今の場合は結果は変わらない. しかし, 一般的には両者は異なる概念である. (レプリケーター・ダイナミックスという概念 (ウェイブル ([45], 3 章)) は多型安定を解析する微分方程式系のことである.)

以上の話を一般化するのは容易で前節の記号をそのまま使って, ESS を定義し直すと,

定義 6.2.
戦略 $\mu \in \mathcal{P}(S)$ が ESS であるとは, $\mu \neq \forall \nu \in \mathcal{P}(S)$, $0 < \exists \varepsilon_0$, $0 < \forall \varepsilon < \varepsilon_0$, に対して

$$(1 - \varepsilon)u(\mu\,, \mu) + \varepsilon u(\mu\,, \nu) > (1 - \varepsilon)u(\nu\,, \mu) + \varepsilon u(\nu\,, \nu)$$

が成立する時をいう.

容易に分かるように,

定理 6.2. 戦略 $\mu \in \mathcal{P}(S)$ が ESS であるための必要十分条件は
$\mathcal{P}(S) \ni \forall \nu \neq \mu$,
(i) $u(\mu\,, \mu) > u(\nu\,, \mu)$
or
(ii) $u(\mu\,, \mu) = u(\nu\,, \mu)$ and $u(\mu\,, \nu) > u(\nu\,, \nu)$

が成り立つことである.

注意: 2 人対称ゲームにおいて, 対称な戦略セット (μ, μ) がナッシュ均衡戦略であるとは, 定義から $\mathcal{P}(S) \ni \forall\nu$ に対して $u_1(\mu, \mu) \geq u_1(\nu, \mu)$ が成立することであるから,

$$\text{ESS} \ \Rightarrow \ \text{対称ナッシュ均衡戦略}$$

である. 2 戦略ゲームの場合は, ESS も対称ナッシュ均衡戦略も常に存在するが, 対称ナッシュ均衡戦略が必ずしも ESS であるとは限らない. 例えば, 3 節のゲーム (7), (8), (9) の混合戦略がそうである. 3 戦略ゲームの場合, 対称ナッシュ均衡戦略は定理 5.6. より常に存在するが, ESS は後で示すように必ずしも存在しない.

前節の定理 5.3. に対応する定理として, 次の定理を得る.

定理 6.3. $\mu \in \mathcal{P}(S)$ に対して, 測度の台 (support) $\mathrm{supp}[\mu]$ を $\mu(F) = 1$ を満たす最小の閉集合 F として定義する. また, $\mu \in \mathcal{P}(S)$ に対して,

$$M(\mu) = \{s \in S\,;\, u(\delta_s\,,\, \mu) = \sup_{t \in S} u(\delta_t\,,\, \mu)\}$$

とおく. この時, $\mu \in \mathcal{P}(S)$ が ESS であるための必要十分条件は
(i) $M(\mu) \supset \mathrm{supp}[\mu]$, and
(ii) $\forall\nu \neq \mu$ such that $M(\mu) \supset \mathrm{supp}[\nu]$ に対して $u(\mu - \nu\,,\, \mu - \nu) < 0$
が成り立つことである. (ii) のような性質は一般に conditionally negative definite と言われて数学では馴染みの概念である.

この定理からナッシュ均衡戦略の場合とは異なる重要な系が導かれる.

系 6.1.
(i) μ が ESS で, $M(\mu) = S$ ならば ESS は唯ひとつしか存在しない.
もっと一般的に
(ii) $\mu_1\,,\, \mu_2$ が ESS で, $M(\mu_1) \supset M(\mu_2)$ を満たすならば $\mu_1 = \mu_2$ である.

証明. 定理 6.3. より $M(\mu_1) \supset M(\mu_2) \supset \mathrm{supp}[\mu_2]$. 従って, $u(\mu_1, \mu_1) = u(\mu_2, \mu_1)$. μ_1 が ESS だから $\mu_1 \neq \mu_2$ ならば $u(\mu_1 - \mu_2, \mu_1 - \mu_2) < 0$ である. ところが, 仮定から

$$u(\mu_1 - \mu_2, \mu_1 - \mu_2) = -u(\mu_1 - \mu_2, \mu_2) = u(\mu_2, \mu_2) - u(\mu_1, \mu_2) \geq 0$$

となって矛盾. よって, $\mu_1 = \mu_2$.

ESS 概念は, 多数を占める同一集団に別の戦略をとる少数集団が侵入出来るか, という視点で定式化してあるが, もし多数, 少数という区別をしなけらばどうなるかを考察してみよう. 定義 6.2. を次のように書き変えてみる.

定義 6.3. 戦略 $\mu \in \mathcal{P}(S)$ が global ESS (GESS) であるとは, $\mu \neq \forall \nu \in \mathcal{P}(S)$, $0 < \forall \varepsilon \leq 1$ に対して

$$(1-\varepsilon)u(\mu\,,\,\mu) + \varepsilon u(\mu\,,\,\nu) > (1-\varepsilon)u(\nu\,,\,\mu) + \varepsilon u(\nu\,,\,\nu) \tag{6.4}$$

が成立する時をいう.

この時, 定理 6.2. は次のように変わる.

定理 6.4. 戦略 $\mu \in \mathcal{P}(S)$ が GESS であるための必要十分条件は $\mu \neq \forall \nu \in \mathcal{P}(S)$ に対して,

$$u(\mu\,,\,\nu) > u(\nu\,,\,\nu) \tag{6.5}$$

が成り立つことである.

証明. GESS の定義式において $\varepsilon = 1$ とおけば (6.5) が得られる. 逆に, (6.5) が $\mu \neq \forall \nu \in \mathcal{P}(S)$ に対して成立している, とする. $\nu_\varepsilon \equiv (1-\varepsilon)\mu + \varepsilon\nu$ とおくと, $0 < \forall \varepsilon \leq 1$ に対して, $\nu_\varepsilon \neq \mu$ だから (6.5) 式から

$$0 < u(\mu\,,\,\nu_\varepsilon) - u(\nu_\varepsilon\,,\,\nu_\varepsilon) = \varepsilon(u(\mu\,,\,\nu_\varepsilon) - u(\nu\,,\,\nu_\varepsilon))$$

が $0 < \forall \varepsilon \leq 1$ に対して成立する. これは 定義 6.3 の不等式 (6.4) が成立することにほかならない.

ESS の定義を眺めるともうひとつの問題点に気づく. 定義の中の ε が戦略 ν 毎に依存して決めてよいことである. 連続関数の定義で各点での連続性に対して一様連続性の概念があり, コンパクト性を仮定すると両者が一致する事情と同じ事情があるのではないか, ということが数学に強い人であればすぐに想像される. このことを調べてみよう.

定義 6.4. 戦略 $\mu \in \mathcal{P}(S)$ が uniformly ESS(UESS) であるとは, $0 < \exists \varepsilon_0$, $0 < \forall \varepsilon < \varepsilon_0$, $\mu \neq \forall \nu \in \mathcal{P}(S)$ に対して

$$(1-\varepsilon)u(\mu\,,\,\mu) + \varepsilon u(\mu\,,\,\nu) > (1-\varepsilon)u(\nu\,,\,\mu) + \varepsilon u(\nu\,,\,\nu)$$

が成立する時をいう.

定理 6.5. 純粋戦略の空間 S は有限集合であるとする. この時, μ が ESS ならば UESS である.

証明.

$$f(\mu,\nu,\varepsilon) \equiv u(\mu,(1-\varepsilon)\mu+\varepsilon\nu) - u(\nu,(1-\varepsilon)\mu+\varepsilon\nu) = u(\mu-\nu,\mu) - \varepsilon u(\mu-\nu,\mu-\nu)$$

とおくと, μ が ESS であるための必要十分条件は

$$\mathcal{P}(S) \ni \forall \nu \neq \mu,\ \exists \varepsilon_0 > 0,\ 0 < \forall \varepsilon < \varepsilon_0,\ f(\mu, \nu, \varepsilon) > 0$$

が成立することであり, μ が UESS であるための必要十分条件は

$$\exists \varepsilon_0 > 0,\ 0 < \forall \varepsilon < \varepsilon_0,\ \mathcal{P}(S) \ni \forall \nu \neq \mu,\quad f(\mu, \nu, \varepsilon) > 0$$

が成立することである. 今,

$$M = \max_{s,t \in S} |u(\delta_s, \delta_t)|,\ \lambda = \max_{s \in S} u(\delta_s, \mu),\ \Lambda = \{s \in S\,;\, u(\delta_s, \mu) = \lambda\}$$

とおく. $\mathrm{supp}[\nu] \subset \Lambda$ ならば, $u(\mu - \nu, \mu) = 0$. μ が ESS だから $u(\mu - \nu, \mu - \nu) < 0$. 従って, $\forall \varepsilon > 0$ に対して $f(\mu, \nu, \varepsilon) > 0$ となり問題ない. そこで, $S - \Lambda \neq \emptyset$ とする. S は有限集合だから (コンパクト集合では必ずしも言えない)

$$\lambda' \equiv \max_{s \in S - \Lambda} u(\delta_s, \mu) < \lambda.$$

次に, $\mathrm{supp}[\nu] \not\subset \Lambda$ として,

$$\nu = \nu_1 + \nu_2,\quad \mathrm{supp}[\nu_1] \subset \Lambda,\quad \mathrm{supp}[\nu_2] \subset S - \Lambda,\quad \nu_2(S - \Lambda) = \alpha > 0,\quad \beta = 1 - \alpha$$

とする. 定義から容易に

$$\begin{aligned} u(\mu - \nu, \mu) &= \lambda - \lambda \nu_1(\Lambda) - u(\nu_2, \mu) \geq \alpha(\lambda - \lambda') > 0, \\ u(\mu - \nu, \mu - \nu) &= u(\beta\mu - \nu_1 + \alpha\mu - \nu_2, \beta\mu - \nu_1 + \alpha\mu - \nu_2) \\ &= \beta^2 u(\mu - \nu_1/\beta, \mu - \nu_1/\beta) + u(\beta\mu - \nu_1, \alpha\mu - \nu_2) \\ &\quad + u(\alpha\mu - \nu_2, \beta\mu - \nu_1) + u(\alpha\mu - \nu_2, \alpha\mu - \nu_2). \end{aligned}$$

ここで, μ は ESS だから第 1 項は負である. 従って,

$$u(\mu - \nu, \mu - \nu) < u(\beta\mu - \nu_1, \alpha\mu - \nu_2) + u(\alpha\mu - \nu_2, \beta\mu - \nu_1) + u(\alpha\mu - \nu_2, \alpha\mu - \nu_2) \leq 6\alpha M.$$

故に, $\varepsilon < (\lambda - \lambda')/(6M)$ であれば, $\forall \nu$ に対して, $f(\mu, \nu, \varepsilon) > 0$ が成立するから μ は UESS である.

定理 6.5. は S のコンパクト性だけでは必ずしも成立しないことが, 次の反例からわかる.

反例: $S = [\,0, 1\,]$, 利得関数を $g(s,t) = 1 - |s - t|^{1 - 1/\sqrt{\log 3/t}}$, ただし, $1\Big/\sqrt{\log 1/0} = 0$ と定義する. $g(s,t)$ は $S \times S$ 上の連続関数であり, δ_0 (点 0 上の単位分布) は ESS である (強ナッシュ均衡戦略である—定義 6.5. 参照). 実際,

$$\forall \nu \neq \delta_0,\quad u(\delta_0, \delta_0) = 1 > u(\nu, \delta_0) = \int_0^1 (1 - s)\, d\nu(s).$$

次に, $u(\delta_0-\nu,\delta_0)$ と $u(\delta_0-\nu,\delta_0-\nu)$ を計算する.

$$\begin{aligned} u(\delta_0-\nu,\delta_0) &= \int_0^1 s\,d\nu(s), \\ u(\delta_0-\nu,\delta_0-\nu) &= \int_0^1 s\,d\nu(s) - \int_0^1 \Big(1-t^{1-1/\sqrt{\log 3/t}}\Big)\,d\nu(t) \\ &\quad + \int_0^1\int_0^1 \Big(1-|s-t|^{1-1/\sqrt{\log 3/t}}\Big)\,d\nu(s)\,d\nu(t). \end{aligned}$$

ここで, $s_n\downarrow 0$ を適当に選んで, $\nu=\delta_{s_n}$ とおく. 上の式から

$$u(\delta_0-\delta_{s_n},\delta_0)=s_n,\quad u(\delta_0-\delta_{s_n},\delta_0-\delta_{s_n})=s_n+s_n^{1-1/\sqrt{\log 3/s_n}}.$$

従って,

$$f(\delta_0,\delta_{s_n},\varepsilon)=s_n\Big(1-\varepsilon\Big(1+s_n^{-1/\sqrt{\log 3/s_n}}\Big)\Big).$$

ここで,

$$\varepsilon_n=\frac{1}{1+s_n^{-1/\sqrt{\log 3/s_n}}}$$

とおくと,

$$0<\varepsilon<\varepsilon_n \;\Rightarrow\; f(\delta_0,\delta_{s_n},\varepsilon)>0,$$

$$\varepsilon_n<\varepsilon \;\Rightarrow\; f(\delta_0,\delta_{s_n},\varepsilon)<0.$$

ところが,

$$s_n^{-1/\sqrt{\log 3/s_n}}=e^{(\log 1/s_n)/\sqrt{\log 3/s_n}}\uparrow+\infty$$

だから, $\varepsilon_n\downarrow 0$ となって, δ_0 は UESS ではないことを示している.

ナッシュ均衡戦略にしろ ESS にしろ条件式において等号が成立する場合が面倒である. 従って, 定義式において等号が成立する場合も含めた場合の概念に名前をつけておくと考えやすい. そこで, 次のような定義を導入する.

定義 6.5. 2 人対称ゲームにおいて μ が強ナッシュ均衡戦略であるとは,

$$\mathcal{P}(S)\ni\forall\nu\neq\mu,\quad u(\mu\,,\,\mu)>u(\nu\,,\,\mu)$$

が成立する時をいう.

定理 6.6. S がコンパクト距離空間, 利得関数が $S\times S$ 上の連続関数とする時, 強ナッシュ均衡戦略は singleton (単位分布) によってのみ実現される.

証明.

$$M(\mu)=\{s\in S\,;\,u(\delta_s\,,\,\mu)=\sup_{t\in S}u(\delta_t\,,\,\mu)\}$$

とおくと, 容易に分かるように μ が対称ナッシュ均衡戦略であれば $M(\mu) \supset \mathrm{supp}[\mu]$ である. もし, $M(\mu)$ が 2 点以上を含むと $\mu \neq \nu$ で $M(\mu) \supset \mathrm{supp}[\nu]$ なる戦略 ν が存在して $u(\mu,\mu) = u(\nu,\mu)$ が成立するから矛盾. 一方, $M(\mu)$ は空集合ではあり得ないから $M(\mu)$ は 1 点集合である.

定義 6.6. 戦略 $\mu \in \mathcal{P}(S)$ が被侵入不能戦略であるとは, $\mu \neq \forall\nu \in \mathcal{P}(S)$ に対して

$$\text{(i)}\ \ u(\mu\,,\,\mu) > u(\nu\,,\,\mu)$$

or

$$\text{(ii)}\ u(\mu\,,\,\mu) = u(\nu\,,\,\mu)\,,\ \ \text{and}\ u(\mu\,,\,\nu) \geq u(\nu\,,\,\nu)$$

が成り立つことである.

定義から容易に分かるように

定理 6.7.
(i) 強ナッシュ均衡戦略 $\Rightarrow$ ESS $\Rightarrow$ 被侵入不能戦略 $\Rightarrow$ ナッシュ均衡戦略,
(ii) GESS $\Rightarrow$ UESS $\Rightarrow$ ESS $\Rightarrow$ 被侵入不能戦略 $\Rightarrow$ ナッシュ均衡戦略.

ESS についてもう一つだけ数学的考察を加えておく. ([45], 45 頁参照, 若干変更してある.)

定義 6.7. $\mu \in \mathcal{P}(S)$ が $\nu \in \mathcal{P}(S)$ に対して局所優越であるとは $\nu_\varepsilon = (1-\varepsilon)\mu + \varepsilon\nu$ とおいた時, $1 \geq \varepsilon_0 > 0$ が存在して, 任意の $0 < \varepsilon \leq \varepsilon_0$ について,

$$u(\mu, \nu_\varepsilon) > u(\nu_\varepsilon, \nu_\varepsilon)$$

が成立する時をいう.

各 ν に対して局所優越であることの定義は丁度各点での連続性の定義に似ている. 当然, ある意味で一様に局所優越という概念が定義できる. すなわち,

定義 6.8. $\mu \in \mathcal{P}(S)$ が局所優越であるとは, μ の $\varepsilon(>0)$- 近傍 U_ε が存在して, $\mu \neq \forall\nu \in U_\varepsilon$ にたいして,

$$u(\mu, \nu) > u(\nu, \nu)$$

が成立する時をいう.

定義からただちに,

定理 6.8. $\mu \in \mathcal{P}(S)$ が局所優越ならば, $\mu \neq \forall\nu \in \mathcal{P}(S)$ に対して局所優越である.

定理 6.9. $\mu \in \mathcal{P}(S)$ が任意の $\nu \in \mathcal{P}(S)$ に対して局所優越であるならば, μ は ESS である.

証明. 単純な変形で証明出来る. つまり,

$$u(\mu, \nu_\varepsilon) - u(\nu_\varepsilon, \nu_\varepsilon) = u(\mu, \nu_\varepsilon) - (1-\varepsilon)u(\mu, \nu_\varepsilon) - \varepsilon)u(\nu, \nu_\varepsilon) = \varepsilon(u(\mu, \nu_\varepsilon) - u(\nu, \nu_\varepsilon)) > 0.$$

この式は ESS の定義式に他ならない.

逆の命題はコンパクト性だけでは証明出来ないようであるが, S が有限集合で, 考察する戦略が S 上のすべての確率分布の場合つまり $\mathcal{P}(S)$ の場合には次の定理が成り立つ.

定理 6.10. S を有限集合とする. この時, $\mu \in \mathcal{P}(S)$ に対して, 次の命題は同値である.
(i) 局所優越である,
(ii) $\mu \neq \forall \nu \in \mathcal{P}(S)$ に対して局所優越である,
(iii)ESS である,
(iv) UESS である.

証明. (i)⇒ (ii) ⇒ (iii) ⇒ (iv) はすでに示した. (iv) ⇒ (i) の証明は $\mathcal{P}(S)$ の構造を強く用いる. 仮定から $\varepsilon_0 > 0$ が存在して, 任意の $0 < \varepsilon \leq \varepsilon_0$ に対して,

$$\begin{aligned} 0 < \varepsilon(u(\mu, \nu_\varepsilon) - u(\nu, \nu_\varepsilon)) &= u(\mu, \nu_\varepsilon) - (1-\varepsilon)u(\mu, \nu_\varepsilon) - \varepsilon)u(\nu, \nu_\varepsilon) \\ &= u(\mu, \nu_\varepsilon) - u(\nu_\varepsilon, \nu_\varepsilon) \end{aligned}$$

が成立している. $\mathcal{P}(S)$ はユークリッド空間の単体であるから, μ の近傍 $V(\mu)$ が存在して $\mu \neq \nu \in V(\mu)$ に対して $0 < \varepsilon \leq \varepsilon_0$ と $\lambda \in \mathcal{P}(S)$ が存在して $\nu = \lambda_\varepsilon = (1-\varepsilon)\mu + \varepsilon\lambda$ のように表される. ゆえに (1) 式より,

$$u(\mu, \nu) - u(\nu, \nu) > 0$$

が従う.

ESS はナッシュ均衡戦略に比べて数学的構造が綺麗な感じはするが, 残念ながら存在は保証されない. 2 戦略の場合は一般に ESS は存在するが, 3 戦略になるともう ESS は存在するとは限らなくなる. また, 唯ひとつとも限らない. 定理 5.6 を考慮すると, 対称なナッシュ均衡戦略が必ずしも ESS であるとは限らない.

2 人 3 戦略の場合 ESS が存在しない例をあげておく ([188]).

利得行列を次のように定める.

$$A = \begin{pmatrix} 4 & 1 & 5 \\ 5 & 3 & 0 \\ 1 & 4 & 3 \end{pmatrix}.$$

ここで, 5 節で注意したように各列から列毎に定数ベクトルを引いても結果は変わらないから, 計算が簡単になるように次のように変形しておく.

$$A' = \begin{pmatrix} 3 & -3 & 2 \\ 4 & -1 & -3 \\ 0 & 0 & 0 \end{pmatrix}.$$

まず, singleton (単位分布) が強ナッシュ均衡になり得ないことは対角成分をチェックすれば明らかである. 従って, 定理 6.2. の (ii) を満たすベクトル $\vec{q}$ が存在する. そこで (ii) の後半部分をチェックする. $u(\mu-\nu, \mu-\nu) < 0$ と書き換えられるから, $\vec{p}$ が ESS ならば $\vec{p}$ と異なるある確率ベクトル $\vec{q}$ に対して $(\vec{p}-\vec{q}, A'(\vec{p}-\vec{q})) < 0$ となるはずである. ところが, $p_1 - q_1 = x$, $p_2 - q_2 = y$ とおくと, 簡単な計算より,

$$(\vec{p}-\vec{q}, A'(\vec{p}-\vec{q})) = (x+y)^2 + y^2$$

となって負にはなり得ない. 従って ESS は存在しない.

もちろん, 定理 5.6. より対称なナッシュ均衡戦略は存在する. 定理 5.3. を利用して具体的に求めてみる. まず, singleton (単位分布) がナッシュ均衡になり得ないことは対角成分を見れば明らかであるから, サポートが 2 点の場合 (3 通り) とフルサポート (3 点) の場合についてそれぞれチェックする.

(i) $\vec{p} = \begin{pmatrix} p \\ 1-p \\ 0 \end{pmatrix}$, $0 < p < 1$ の場合.

$$A'\vec{p} = \begin{pmatrix} 6p-3 \\ 5p-1 \\ 0 \end{pmatrix}.$$

従って, 上記の注意から $6p-3 = 5p-1 \geq 0$ でなければならない. ところがこのような確率ベクトルは存在しない.

(ii) $\vec{p} = \begin{pmatrix} p \\ 0 \\ 1-p \end{pmatrix}$, $0 < p < 1$ の場合.

$$A'\vec{p} = \begin{pmatrix} p+2 \\ 7p-3 \\ 0 \end{pmatrix}.$$

従って, 上記の注意から $p+2 = 0$, $7p-3 \leq 0$ でなければならない. ところがこのような確率ベクトルは存在しない.

(iii) $\vec{p} = \begin{pmatrix} 0 \\ p \\ 1-p \end{pmatrix}$, $0 < p < 1$ の場合.

$$A'\vec{p} = \begin{pmatrix} -5p+2 \\ 2p-3 \\ 0 \end{pmatrix}.$$

従って, 上記の注意から $2p-3=0,\ -5p+2\leq 0$ でなければならない. ところがこのような確率ベクトルは存在しない.

(iv) $\vec{p}=\begin{pmatrix} p_1 \\ p_2 \\ 1-p_1-p_2 \end{pmatrix}$, $0<p_1<1,\ 0<p_2<1$ の場合.

$$A'\vec{p}=\begin{pmatrix} p_1-5p_2+2 \\ 7p_1+2p_2-3 \\ 0 \end{pmatrix}.$$

従って, 上記の注意から $p_1-5p_2+2=0,\ 7p_1+2p_2-3=0$ を満たさなければならない. これを解くと $p_1=\dfrac{11}{37},\ p_2=\dfrac{17}{37},\ p_3=\dfrac{9}{37}$ を得る. これが唯一のナッシュ均衡戦略である.

§7. 内輪付き合いをするタカ・ハトゲーム [4]
— 平和主義者の生き残り戦略—

有名なタカ・ハトゲームの利得行列 $U(\cdot\,,\cdot)$ は

	H	D
H	$(V-C)/2$	V
D	0	$V/2$

で表される. $V > C$ の場合は H (タカ) が優越戦略であるから, ハト (平和主義者) は侵入できない. 他方, $V < C$ つまり, タカ戦略をとると獲得利益よりも闘争コストが上回る場合, タカ派は一定の割合 $p = V/C$ だけしか生存できない (多型安定). 逆に言うとハト派が存在し得るのはタカ同士が戦うと互いに傷ついて損をする場合だけである. 3 節の分類に従えば, $V > C$ の場合は囚人のジレンマ, $V < C$ ならばチキンゲームになっている.

ただし, 以上の考察の大前提はランダム・マッチング, すなわち, タカ派対ハト派が $p : (1-p)$ の割合で存在している場合, タカ派に出くわす割合が同じ確率 p であると仮定している.

現実には, ハト (平和主義者) は敢えてタカに接近しようとは思わないであろう. 何らかの情報を得て行動すると考えると, ランダム・マッチングの仮定は現実的とは思われない. 問題はランダム・マッチング (独立性) の仮定をはずした場合にどのように数学的に表現するのが適切か, ということである. 本節では, 必ずしも独立ではない二つの確率変数によって, このような状態を表現することを試みる.

ゲーム理論の分野で知られている Aumann は次のように述べている ([115], p.68).

In describing these phenomena, it is best to view a randomized strategy as a random variable with values in the pure strategy space than as a distribution over pure strategies.

残念ながら彼の定式化は, 本稿が基礎にしているコルモゴロフ流の確率論ではなく, Savage 流の主観確率 ([216]) による定式化に基づいている. 彼の言う確率変数は本質的にはコルモゴロフの公理を満たしている確率変数と同一であるが, どうも解釈の仕方が違うようである.

定式化.

いま, タカ派とハト派の割合が $p : 1-p$ であるような集団から二人のプレーヤーを選ぶ. 各々の属性 (タカかハト) を表す確率変数を X_1, X_2 とする. 2 次元の確率変数 (X_1, X_2) の分布を次のように仮定する.

仮定 7.1 各々の周辺分布は集団の割合と同じである. 式で表現すると,

$$\mathrm{Prob}(X_1 = H) = \mathrm{Prob}(X_2 = H) = p.$$

[4] 第 34 回数理社会学会 (2002.9.15-16, 於岩手県立大学) における研究発表の原稿に加筆

ランダム・マッチングという場合は, X_1 と X_2 が独立な確率変数である, と仮定する場合に相当する. 本節では, ハト派が争いを避けて選択的にハト同士で内輪に付き合う傾向を表現したいわけであるから, 独立性を仮定することは出来ない. 一般に 2 次元の確率変数 (X_1, X_2) の分布は仮定 7.1 だけでは決まらない. あと一つのパラメータを自由に選ぶことが出来る.

独立性を仮定しない場合, 相関係数を考えるのが普通であろうが, 相関係数の定義を思い出してもらうとわかるように, 「相関」という概念は実数値確率変数どうしの間でしか定義できない. ところが, タカ・ハトゲームではとる値は実数ではない. 従って, 「独立でない」, という現象を表現するために「相関係数」は使わないで, 条件付分布にパラメータを導入して独立でない事象を計量的に表現する. そこで, 分布を定めるパラメータ α と β を次のように定める. ただし, 後に述べるようにこの二つのパラメータは独立ではない. つまり, あとで述べるような関数関係がある.

$$\text{(i)} \quad \mathrm{Prob}(X_2 = H / X_1 = D) = (1-\alpha)p,$$

$$\text{(ii)} \quad \mathrm{Prob}(X_2 = H / X_1 = H) = (1+\beta)p.$$

ここで, $\mathrm{Prob}(A/B)$ は事象 B が与えられたときの事象 A の条件付確率を表す. ただし, $\mathrm{Prob}(B) = 0$ のときは $\mathrm{Prob}(A/B) = 0$ と約束する. 条件付確率は「確率」だから, 値は 0 と 1 の間を任意に取り得る. つまり, パラメータ α の範囲は $-(1-p)/p \leq \alpha \leq 1$ である. しかし, 我々の関心は $\alpha \geq 0$ の範囲にあるから正の範囲が p に関係しないのはむしろ都合がよいと考えられる.

このとき, 仮定 7.1 と 上記の (i), (ii) より,

$$\begin{aligned}
\mathrm{Prob}(X_2 = H) &= p \\
&= \mathrm{Prob}(X_2 = H / X_1 = H)\,\mathrm{Prob}(X_1 = H) \\
&\quad + \mathrm{Prob}(X_2 = H / X_1 = D)\,\mathrm{Prob}(X_1 = D) \\
&= (1+\beta)p^2 + (1-\alpha)p(1-p).
\end{aligned}$$

従って, α と β の間には

$$\beta p = \alpha(1-p)$$

なる関係がある. 容易にわかるように,

$$\text{確率変数 } X_1 \text{ と } X_2 \text{ が独立} \iff \alpha = \beta = 0$$

である. $\alpha > 0$ の場合には, ハト派は内輪づき合いをしていると理解することができる. $\alpha = 1$ の場合には, ハト派同士だけで付き合い, タカ派とハト派には交流がないことを意味する.

さて, 以上によって, 分布が定まったから, タカ派の平均利得 $W(H)$ とハト派の平均利得 $W(D)$ を計算する. これらは条件付平均を用いて次のように表される.

$$
\begin{aligned}
W(H) &= E[U(X_1,\, X_2)/X_1 = H] \\
&= U(H,\, H)\,\mathrm{Prob}(X_2 = H/X_1 = H) + U(H,\, D)\,\mathrm{Prob}(X_2 = D/X_1 = H) \\
&= \frac{V-C}{2}(p+\beta p) + V(1-p-\beta p) \\
&= \frac{V}{2}(2-p-\alpha(1-p)) - \frac{C}{2}(p+\alpha(1-p)).
\end{aligned}
$$

一方,

$$
\begin{aligned}
W(D) &= E[U(X_1,\, X_2)/X_1 = D] \\
&= U(D,\, H)\,\mathrm{Prob}(X_2 = H/X_1 = D) + U(D,\, D)\,\mathrm{Prob}(X_2 = D/X_1 = D) \\
&= \frac{V}{2}(1-p+\alpha p)\ .
\end{aligned}
$$

ここで, 両者を比較して $W(H) \geq W(D)$ の条件を求めると,

$$
p \leq \frac{V}{C} - \frac{\alpha}{1-\alpha}
$$

を得る. この式を眺めると, $V > C$ の場合, つまり従来ならばハト派が存在し得ない場合でも $\alpha > 0$ ならば可能性がある, ということが分かる. 定量的に言えば $\alpha > (V-C)/V$ ならば $p < 1$ となり, ハト派が存在し得るという結論が得られる. この場合, タカ (H) とハト (D) の利得が等しくなるのは

$$
p = \frac{V}{C} - \frac{\alpha}{1-\alpha}
$$

のときであるから, この割合は, ある程度の内輪付き合い $\alpha(> (V-C)/V)$ を一定と仮定したとき, ある種の均衡点 (多型安定) である, ということが出来る. 実際, $p < V/C - \alpha/(1-\alpha)$ の場合, 上記の計算から $W(H) > W(D)$ であり, p は増加して均衡点に達しようとする. 逆向きの不等式の場合は減少して均衡点に達する. ただし, この均衡点の定義は Aumann([115], [117]) のいう correlated equilibrium とは異なる. 彼の定義によって我々のタカ・ハトゲームを分析してみると, 奇妙なことに $V/C > 1$ の場合に, 以下でみるように $\alpha > 0$ では均衡戦略が存在しない. $\alpha < 0$ の場合の均衡戦略を解釈するとハト派がむしろ内輪付き合いをしない方が均衡戦略になってしまい, 解釈の点で矛盾を生じる. 数学的にはいろいろな均衡の定義はあり得る可能性はあるが, 具体例に適用してみた場合に現実感がうまく反映していないと数理モデルとしては具合が悪いであろう.

Aumann による correlated equilibrium の定義を確率変数を用いて表現すると以下のようになっている. ([115], [117], [190] 参照)

$N = \{1, 2, \ldots, n\}$ をプレーヤーの集合 ($n \geq 2$), $S_i,\ i \in N$ をプレーヤー i の選択肢の集合 (有限集合) , $S = S_i \times \cdots \times S_n$, $s = (s_i,\, s_{-i}) \in S$, $u_i(s) = u_i(s_i,\, s_{-i})$ をプレーヤー i の pay off functions とする. X_i を S_i に値をとる確率変数, $(X_1, \ldots, X_n)$ を S に値を取

る確率変数とする.

定義 7.1 $(X_1, \ldots, X_n)$ (あるいはこれらの確率変数で決まる n 次元分布が, と言ってもよい) が correlated equilibrium であるとは, $\forall i \in N$, $\forall s, \forall t \in S_i$ に対して,

$$E[u_i(X_i, X_{-i})/X_i = s] \geq E[u_i(t, X_{-i})/X_i = s]$$

が成立するときをいう. ここで, $E[X/Y = s]$ は事象 $\{Y = s\}$ 上での X の条件付平均を表す. ただし, $\mathrm{Prob}(X_i = s) = 0$ の場合, 上式は常に成立しているとみなす.

例. $n = 2$ の対称ゲーム (タカ・ハトゲーム) の場合でかつ (X_1, X_2) の各周辺分布が等しい範囲内で (つまり我々の定式化の範囲内で) 上式の条件を書き直すと

$$\text{(i)} \quad E[U(H, X_2)/X_1 = H] \geq E[U(D, X_2]/X_1 = H],$$

と

$$\text{(ii)} \quad E[U(D, X_2)/X_1 = D] \geq E[U(H, X_2)/X_1 = D]$$

が成り立つことである. 以下, $0 < p < 1$ の範囲の correlated equilibrium を求める. (i) を書き直すと

$$\begin{aligned}&(U(H, H) - U(D, H))\,\mathrm{Prob}(X_2 = H/X_1 = H)\\ &\quad \geq (U(D, D) - U(H, D))\,\mathrm{Prob}(X_2 = D/X_1 = H).\end{aligned}$$

(ii) を書き直すと

$$\begin{aligned}&(U(D, H) - U(H, H))\,\mathrm{Prob}(X_2 = H/X_1 = D)\\ &\quad \geq (U(H, D) - U(D, D))\,\mathrm{Prob}(X_2 = D/X_1 = D).\end{aligned}$$

ここで,

$$\begin{aligned}\mathrm{Prob}(X_2 = H/X_1 = H) &= \alpha + (1-\alpha)p, \quad \mathrm{Prob}(X_2 = D/X_1 = H) = (1-\alpha)(1-p),\\ \mathrm{Prob}(X_2 = H/X_1 = D) &= (1-\alpha)p, \quad \mathrm{Prob}(X_2 = D/X_1 = D) = 1 - (1-\alpha)p\end{aligned}$$

を代入して整理すると, (i) は

$$\frac{V}{C} - \alpha \geq (1-\alpha)p.$$

(ii) は

$$(1-\alpha)p \geq \frac{V}{C}$$

となる. したがって, 上記 (i), (ii) が両立するためには $\alpha \leq 0$ が必要条件となる. $\alpha < 0$ の場合, (α が許される範囲は $-(1-p)/p \leq \alpha \leq 1$ だった) V, C, α が

$$\frac{V}{(1-\alpha)C} < 1$$

を満たしていれば, (特に $V/C > 1$ であっても) correlated equilibrium は α を固定しても連続濃度存在することになる. 前述したように, $\alpha < 0$ の場合はハト派の内輪付き合いとは逆の意味になるから, この範囲で如何なる意味にしろ均衡戦略が存在するのは具合が悪いように思われる.

§8. 繰り返し囚人のジレンマ (ランダム停止時刻を持つ場合) [5]

「囚人のジレンマ」とは, 最も簡単な 2 人対称ゲームであって, 利得行列は次のように表される.

		プレーヤー II	
		C	D
プレーヤー I	C	R	S
	D	T	P

囚人のジレンマ, という場合は $T > R > P > S$ と $2R > T + S$ を仮定する ($2R > T + S$ は仮定しない場合もある). この 2 人対称ゲーム (利得がマトリックスで与えられるからマトリックスゲームともいう) の意味は, プレーヤー I が戦略 C(ooperate) を, プレーヤー II が同じく戦略 C を選択した場合のプレーヤー I の (同じ戦略だからプレーヤー II も同じ) 利得が $u(C, C) = R$, プレーヤー I が戦略 C, プレーヤー II が戦略 D(efect) を選択した場合のプレーヤー I の利得が $u(C, D) = S$, (この時, プレーヤー II の利得は対称だから $u(D, C) = T$), ... であることを示している. 2 人対称ゲームは, プレーヤーの利得というより, 各プレーヤーの戦略全体 (戦略セット, または戦略プロファイルという) に対する, 一方の戦略の利得とみなすこともできる.

繰り返しゲームとはランダムに決められる停止時刻 ζ まで, 各回毎にプレーヤーは手 (C 又は D) をランダムに選ぶゲームであると考えられるからプレーヤー I の一連の手は C または D に値を取る, 抽象確率空間 $(\Omega, \mathcal{F}, \mathrm{Prob})$ 上で定義された確率変数の列

$$\mathbf{X} = (X_1, X_2, \ldots)$$

で表現することが出来る. プレーヤー II についても同様に

$$\mathbf{Y} = (Y_1, Y_2, \ldots)$$

で表す.

記号の準備:

$\mathcal{F}_n(X)$: $X_1, \ldots, X_n$ から生成される sub σ-algebra.

$\mathcal{F}_n(X, Y)$: $X_1, Y_1, \ldots, X_n, Y_n$ から生成される sub σ-algebra.

プレーヤー I の効用 (利得) 関数 u:

$$u(C, C) = R,\ u(D, C) = T,\ u(C, D) = S,\ u(D, D) = P.$$

ζ : 自然数に値を取る確率変数 =ランダム停止時刻 ζ までゲームが続く, と仮定する.

ここで, ゲームのルールと戦略を定義する.

[5] 科研費による研究集会「確率過程とその周辺 (2002.12.9-12.) 於: 慶応大学」における講演原稿に加筆

プレーヤーIの一連の手 (各回にCを出すか, Dを出すか) を確率変数列 $\mathbf{X} = (X_1, X_2, \ldots)$ で表現する. プレーヤー II についても同様に $\mathbf{Y} = (Y_1, Y_2, \ldots)$ とする. 従って, 実際に実現した, あるゲームの n 回までの手は見本関数

$$(X_1(\omega), Y_1(\omega)), \ldots, (X_n(\omega), Y_n(\omega)) \, ; \, \omega \in \Omega$$

に他ならない.

ところで, 各回のゲームは相手の手とは独立であるから (非協力ゲームの仮定) 上記の確率過程は常に次の仮定を満たす範囲で考える.

仮定:
(i) 1 回目の手を表す確率変数 X_1 と Y_1 は独立である.
(ii)(各回の手も独立である=条件付独立) : $\forall n = 1, 2, \ldots$ に対して,

$$\begin{aligned} &\mathrm{Prob}(X_{n+1} = *, \, Y_{n+1} = ** / \mathcal{F}_n(X, Y)) \\ &\quad = \mathrm{Prob}(X_{n+1} = * / \mathcal{F}_n(X, Y)) \times \mathrm{Prob}(Y_{n+1} = ** / \mathcal{F}_n(X, Y)) \end{aligned}$$

が成り立つ. ここで, $*$ と $**$ は C 又は D を表す. また, $\mathrm{Prob}(A/\mathcal{F}')$ は sub σ-algebra $\mathcal{F}'$ に関する, 事象 A の条件付確率.

以上のような定式化によるプレーヤーIの平均利得 $u(\mathbf{X}, \mathbf{Y})$ は次のように与えられる. ζ は $\mathbf{X}, \mathbf{Y}$ と独立であることを必ずしも仮定しない. (どちらか一方のプレーヤーのみが停止時刻を決めることが出来るゲームも考えられているから)

$$u(\mathbf{X}, \mathbf{Y}) = E[\sum_{n=1}^{\zeta} u(X_n, Y_n)].$$

この時, 容易に

$$\begin{aligned} u(\mathbf{X}, \mathbf{Y}) &= \sum_{k=1}^{+\infty} E\left[\sum_{n=1}^{k} u(X_n, Y_n) \, ; \, \zeta = k\right] \\ &= \sum_{n=1}^{+\infty} \sum_{k=n}^{+\infty} E[u(X_n, Y_n) \, ; \, \zeta = k] \\ &= \sum_{n=1}^{+\infty} E[u(X_n, Y_n) \, ; \, \zeta \geq n]. \end{aligned}$$

を得る.

さらに, ζ が $(\mathbf{X}, \mathbf{Y})$ と独立であり, その分布は幾何分布 $G(\delta)$ (i.e. $\mathrm{Prob}(\zeta = n) = (1-\delta)\delta^{n-1} \, ; \, n = 1, 2, \ldots$) であると仮定すると,

$$\begin{aligned} u(\mathbf{X}, \mathbf{Y}) &= \sum_{n=1}^{+\infty} E[u(X_n, Y_n)] \, \mathrm{Prob}(\zeta \geq n) \\ &= \sum_{n=1}^{+\infty} E[u(X_n, Y_n)] \, \delta^{n-1}. \end{aligned}$$

となり, 経済学の分野でよく知られている割引因子による利得と一致する ([163]).

次にプレーヤーの「戦略」を定義する. 一般的に定義する前に, 直観的にも理解できてよく知られている戦略が我々の定式化の下ではどのように表現できるか, を見てみよう.

定義 8.1. 相手の手の如何に係わらず常に D を出す戦略を "all-D" 戦略という (全面裏切り). 意味は明らかであろう.

定義 8.2. $\mathbf{X} = (X_1, X_2, \ldots)$ が "TFT" 戦略 (Tit for Tat, オウム返し戦略, しっぺ返し戦略) であるとは, 相手の任意の戦略 $\mathbf{Y} = (Y_1, Y_2, \ldots)$ に対して
(i) $X_1 = C$ a.s.,
(ii) $\mathrm{Prob}(X_{n+1} = C/\mathcal{F}_n(X, Y) \wedge \{Y_n = C\}) = 1$ a.s.,
(iii) $\mathrm{Prob}(X_{n+1} = D/\mathcal{F}_n(X, Y) \wedge \{Y_n = D\}) = 1$ a.s.
を満たすときをいう. 意味は明らかであろう. C に対して C を返す, という意味ではしっぺ返しというのはちょっと不適当な気がする. なお, (i) において, C を D に置き換えた戦略 (つまり, 最初は D を出す) を "D-TFT" 戦略という.

次に, よく知られた "トリガー" 戦略を説明する. 言葉で説明すると, 最初は C を出し, 相手が C (協力) を出す限り自分も C(協力) を出す. ただし, 相手が D (裏切り) を出した次の回から決して相手を許さず D を出し続ける, という非寛容な戦略である. 確率変数を使った定義をするために確率過程論ではよくやるように, C または D への到達時刻を次のように定義する.

$$\begin{aligned} \tau_C(Y) &= \min\{n \geq 1\,;\, Y_n(\omega) = C\} \\ &= \infty \quad \text{if } \{\ \} = \emptyset. \end{aligned}$$

同様に $\tau_D(Y)$ が定義できる. このように定義すると, マルコフ過程論でよく知られているように $\tau_C(Y)$ は増大する sub σ-algebra $\mathcal{F}_n(Y)\,;\, n = 1, 2, \ldots$ に関してマルコフ時間になっている.

マルコフ時間の定義:
定義 8.3. 自然数または ∞ に値を取る確率変数 τ が, 増大する sub σ-algebra $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots$ に対して "マルコフ時間" であるとは $\forall n = 1, 2, \ldots,\ \{\omega\,;\, \tau(\omega) = n\} \in \mathcal{F}_n$ が成り立つときをいう.

マルコフ時間に関しては, よく知られているように sub σ-algebra $\mathcal{F}_\tau$ が次のように定義される.

$$\mathcal{F}_\tau = \{A \in \mathcal{F}\,;\, \forall n,\ A \cap \{\tau = n\} \in \mathcal{F}_n\}.$$

定義 8.4. $\mathbf{X} = (X_1, X_2, \ldots)$ が "Trigger" 戦略であるとは, 相手の任意の戦略 $\mathbf{Y} = (Y_1, Y_2, \ldots)$ に対して,

(i) $X_1 = C$ a.s.,

(ii) $\mathrm{Prob}(X_{n+1} = D/\mathcal{F}_n(X, Y) \wedge \{\tau_D(Y) \leq n\}) = 1,\ n = 1, 2, \ldots$ a.s.,

(iii) $\mathrm{Prob}(X_{n+1} = C/\mathcal{F}_n(X, Y) \wedge \{\tau_D(Y) > n\}) = 1,\ n = 1, 2, \ldots$ a.s.

が成り立つときをいう.

定義 8.5. さて, 一般にこの繰り返しゲームにおけるプレーヤー I の「戦略」とは確率過程 $X_1, X_2, \ldots$ の分布を次のようなルールで決めておくことである.

(i) 確率変数 X_1 の分布を決定する.

(ii) プレーヤー II の任意の「戦略」と自分の「戦略」の n 回までの実現値

$$(X_1(\omega), \ldots, X_n, Y_1(\omega), \ldots, Y_n(\omega))$$

を知った時の $n+1$ 回目の条件付確率

$$\mathrm{Prob}(X_{n+1} = */\mathcal{F}_n(X, Y))\,;\, n = 1, 2, \ldots$$

を決めておく.

従って, プレーヤー I の「戦略」から決まる確率変数列 $(X_2, X_3, \ldots)$ の分布は相手の手 $(Y_1, Y_2, \ldots)$ にも依存して決まる. つまり, 「戦略」とは一意に決まるひとつの確率過程ではないことに注意する. しかし, 相手も戦略を決めると, 仮定 (ii) より確率過程 $((X_1, Y_1), (X_2, Y_2), \ldots)$ が一意に決まる.

次に, マルコフ過程論で使うシフトオペレーターを真似て, ゲームが k 回まで終わった後の「部分ゲーム」

$$(\theta_k\mathbf{X}, \theta_k\mathbf{Y}) = ((\theta_k X)_1, (\theta_k Y)_1), ((\theta_k X)_2, (\theta_k Y)_2), \ldots)$$

を次のように定義する.

(i) $\mathrm{Prob}((\theta_k X)_1 = *, (\theta_k Y)_1 = **) = \mathrm{Prob}(X_{k+1} = *, Y_{k+1} = **/\mathcal{F}_k(X, Y))$,

(ii) $n = 1, 2, \ldots,$ に対して,

$$\begin{aligned}&\mathrm{Prob}((\theta_k X)_{n+1} = *, (\theta_k Y)_{n+1} = **/\mathcal{F}_n(\theta_k X, \theta_k Y))\\ &= \mathrm{Prob}(X_{k+n+1} = *, Y_{k+n+1} = **/\mathcal{F}_k(X, Y) \vee \mathcal{F}_n(\theta_k X, \theta_k Y)).\end{aligned}$$

我々の仮定 (ii) から, $(\theta_k X)_1$ と $(\theta_k Y)_1$ は独立であり, 以下同様に確率過程

$$(((\theta_k X)_1, (\theta_k Y)_1), ((\theta_k X)_2, (\theta_k Y)_2), \ldots)$$

は仮定 (ii) を満たすことが分かる.

さらに, 増大する sub σ-algebra $\mathcal{F}_n(X,Y)$ に関するマルコフ時間 τ が与えられて, 条件 $\mathrm{Prob}(\tau < +\infty) > 0$ を満たすとき,

(i)

$$\begin{aligned}&\mathrm{Prob}((\theta_\tau X)_1 = *,\, (\theta_\tau Y)_1 = **)\\ &= \mathrm{Prob}(X_{\tau+1} = *,\, Y_{\tau+1} = **/\mathcal{F}_\tau(X,Y) \wedge \{\tau < +\infty\}),\end{aligned}$$

(ii) $n = 1, 2, \ldots,$ に対して,

$$\begin{aligned}&\mathrm{Prob}((\theta_\tau X)_{n+1} = *,\, (\theta_\tau Y)_{n+1} = **/\mathcal{F}_n(\theta_\tau X,\, \theta_\tau Y))\\ &= \mathrm{Prob}(X_{\tau+n+1} = *,\, Y_{\tau+n+1} = **/(\mathcal{F}_\tau(X,Y) \vee \mathcal{F}_n(\theta_\tau X,\, \theta_\tau Y))\end{aligned}$$

で定義される新しい確率過程 $(\theta_\tau \mathbf{X},\, \theta_\tau \mathbf{Y})$ はマルコフ時間の性質から仮定 (i), (ii) を満たしていることが容易に確認できる.

Remark: $\mathbf{X}$ が "Trigger" 戦略ならば $\theta_{\tau_D(Y)}\mathbf{X}$ は "all-D"戦略である.

$\mathbf{X}$ が "TFT" 戦略ならば $\theta_{\tau_D(Y)}\mathbf{X}$ は "D-TFT" 戦略で, $\theta_{\tau_C(Y)}\mathbf{X}$ は "TFT" 戦略である.

定理 8.1. ランダム停止時間 ζ は両プレーヤーのすべての戦略と独立であると仮定する. さらに, その平均 $E[\zeta]$ は有限であるとする. $\mathrm{Prob}(\zeta = n) = \zeta_n$, $\bar{\zeta}_n = \sum_{k=n}^{+\infty} \zeta_k$ とおく.

このとき,

$$(*) \qquad \forall k,\ (R-P) \sum_{n=k+1}^{+\infty} \bar{\zeta}_n \geq (T-R)\bar{\zeta}_k$$

が成り立つならば "Trigger" 戦略は Nash 均衡である. つまり, $\mathbf{X}$, $\mathbf{X}'$ が "Trigger" 戦略であるとするとき,

$$u(\mathbf{X}',\, \mathbf{X}) \geq u(\mathbf{Y},\, \mathbf{X})$$

が任意の戦略 $\mathbf{Y}$ に対して成立する. 逆に, 上の $(*)$ が成り立たないとすると, "Trigger" 戦略は Nash 均衡ではない.

なお, "all-D" 戦略は常に Nash 均衡である. この方面の研究の目的は, (C, C) 状態の時に, 次回に D を出すと長期的に見て損をするような戦略は何か, という観点から研究されている. 従って, "all-D" は "自明"なナッシュ均衡である, ということが出来る.

系. もし, $n_0 = \max\{n\,;\, \zeta_n > 0\} < +\infty$ を満たす n_0 が存在するならば (i.e. 有限回でゲームが終わる) "Trigger" 戦略は Nash 均衡ではない. 実際, 定理 1 において, 左辺は

$$\sum_{n=n_0+1}^{+\infty} \bar{\zeta}_n = 0$$

であるが, 右辺は

$$\bar{\zeta}_{n_0} > 0$$

であって, 定理の条件を満たさない.

例. ランダム停止時間 ζ の分布が幾何分布, つまり, $\zeta_n = (1-\delta)\delta^{n-1}$ であるならば, $\bar{\zeta}_k = \delta^{k-1}$ かつ $\sum_{n=k+1}^{+\infty} \bar{\zeta}_n = \delta^k/(1-\delta)$ だから. 簡単な計算によって $\delta \geq (T-R)/(T-P)$ が得られる. これは既によく知られている不等式である.

定理 8.2. ランダム停止時間 ζ は両プレーヤーのすべての戦略と独立であると仮定する. さらに, その分布は幾何分布であると仮定する. このとき, 不等式

$$\delta \geq \max\{(T-R)/(T-P),\, (T-R)/(R-S)\}$$

が満たされるならば "TFT"戦略は Nash 均衡である.

定理 8.1 の証明. $\mathbf{X}$ $\mathbf{X}'$ を Trigger 戦略とする. このとき, $\bar{\zeta}_n = \sum_{k=n}^{+\infty} \zeta_k$, $\mathrm{Prob}(\tau_D(Y) = n) = \tau_n,\ n = 1, 2, \ldots, \infty$, とおくと, ζ は独立だから,

$$\begin{aligned}
u(\mathbf{X}', \mathbf{X}) &= E\Big[\sum_{n=1}^{\zeta} u(X'_n, X_n)\Big] = R\,E[\,\zeta\,] \\
&= R\sum_{k=1}^{+\infty} E[\,\zeta\,;\, \tau_D(Y) = k\,] + R\,E[\,\zeta\,]\tau_\infty \\
&= R\sum_{k=1}^{+\infty} E[\,\zeta\,]\,\tau_k + R\,E[\,\zeta\,]\tau_\infty \\
&= R\sum_{k=1}^{+\infty}\sum_{n=1}^{+\infty} \bar{\zeta}_n \tau_k + R\,E[\,\zeta\,]\tau_\infty
\end{aligned}$$

を得る. 同様に,

$$\begin{aligned}
u(\mathbf{Y}, \mathbf{X}) &= \sum_{n=1}^{+\infty} \bar{\zeta}_n E[u(Y_n, X_n)] \\
&= \sum_{n=1}^{+\infty} \bar{\zeta}_n E[u(Y_n, X_n)\,;\, \tau_D(Y) > n,] \\
&\quad + \sum_{n=1}^{+\infty} \bar{\zeta}_n E[u(Y_n, X_n)\,;\, \tau_D(Y) = n\,] \\
&\quad + \sum_{n=1}^{+\infty} \bar{\zeta}_n E[u(Y_n, X_n)\,;\, \tau_D(Y) < n\,] \\
&\equiv \mathrm{I} + \mathrm{II} + \mathrm{III}.
\end{aligned}$$

まず, $\tau_D(Y) > n$ ならば, $Y_n = X_n = C$ であることに注意すると,

$$\begin{aligned}
\mathrm{I} &= R\sum_{n=1}^{+\infty} \bar{\zeta}_n\, \mathrm{Prob}(\tau_D(Y) > n\,) \\
&= R\sum_{n=1}^{+\infty} \bar{\zeta}_n \Big(\sum_{k=n+1}^{+\infty} \tau_k + \tau_\infty\Big) = R\sum_{k=2}^{+\infty}\Big(\sum_{n=1}^{k-1} \bar{\zeta}_n\Big)\tau_k + R\,E[\,\zeta\,]\tau_\infty
\end{aligned}$$

を得る. 同様に, $\tau_D(Y) = n$ ならば $Y_n = D$ かつ $X_n = C$ であり, $\tau_D(Y) < n$ ならば $X_n = D$ であることに注意すると,

$$\text{Ⅱ} = T \sum_{n=1}^{+\infty} \bar{\zeta}_n \, \tau_n,$$

かつ,

$$\begin{aligned} \text{Ⅲ} &\leq P \sum_{n=1}^{+\infty} \bar{\zeta}_n \, \mathrm{Prob}(\tau_D(Y) < n\,) \\ &= P \sum_{n=2}^{+\infty} \sum_{k=1}^{n-1} \tau_k = \mathrm{Prob} \sum_{k=1}^{+\infty} \Big(\sum_{n=k+1}^{+\infty} \bar{\zeta}_n \Big) \tau_k \\ (\text{Ⅲ} &= \quad \Leftrightarrow \quad \forall\, n \geq \tau_D(Y),\ Y_n = D). \end{aligned}$$

を得る. 以上の評価式から

$$\begin{aligned} u(\mathbf{Y}, \mathbf{X}) &= \text{Ⅰ} + \text{Ⅱ} + \text{Ⅲ} \\ &\leq R \sum_{k=2}^{+\infty} \Big(\sum_{n=1}^{k-1} \bar{\zeta}_n \Big) \tau_k + R\, E[\,\zeta\,] \tau_\infty + T \sum_{k=1}^{+\infty} \bar{\zeta}_k \tau_k + P \sum_{k=1}^{+\infty} \Big(\sum_{n=k+1}^{+\infty} \bar{\zeta}_n \Big) \tau_k. \end{aligned}$$

が得られる. 従って

$$u(\mathbf{X}', \mathbf{X}) - u(\mathbf{Y}, \mathbf{X}) \geq \sum_{k=1}^{+\infty} \Big(R \sum_{n=k}^{+\infty} \bar{\zeta}_n - T\, \bar{\zeta}_k - P \sum_{n=k+1}^{+\infty} \bar{\zeta}_n \Big) \tau_k.$$

が得られるから, 少々変形して最終的に, 定理の仮定から

$$R \sum_{n=k}^{+\infty} \bar{\zeta}_n - T\, \bar{\zeta}_k - P \sum_{n=k+1}^{+\infty} \bar{\zeta}_n = (R - P) \sum_{n=k+1}^{+\infty} \bar{\zeta}_n - (T - R)\, \bar{\zeta}_k \geq 0.$$

を得る. (前半の証明終わり)

逆に, $\mathbf{X}$ を "Trigger"戦略とし,

$$(R - P) \sum_{k=n_0+1}^{+\infty} \bar{\zeta}_k < (T - R) \sum_{k=n_0}^{+\infty} \zeta_k$$

を満たす自然数 n_0 が存在すると仮定すると, 明らかに $\mathrm{Prob}(\tau_D(Y) = n_0) = 1$ を満たす確率過程 $\mathbf{Y} = (Y_1, Y_2, \ldots)$ を構成することが出来るから, 上記の評価式を考慮すると,

$$u(\mathbf{X}', \mathbf{X}) < u(\mathbf{Y}, \mathbf{X})$$

が得られて, このとき戦略セット $(\mathbf{X}, \mathbf{Y})$ はナッシュ均衡でない. (定理 8.1 の証明終わり)

定理 8.2 の証明のために, 知られている次の Lemma を証明する.

Lemma 8.1. δ が条件 $1 > \delta \geq \max\{(T-R)/(T-P),\, (T-R)/(R-S)\}$ を満たすならば, $n = 2, 3, \ldots$ に対して,

$$R \sum_{k=0}^{n} \delta^k \geq T + P \sum_{k=1}^{n-1} \delta^k + S\,\delta^n.$$

が成り立つ.

証明. まず初めに, 囚人のジレンマというときに付加的に仮定する条件 $2R > T + S$ がここで効いてくることに注意する. すなわち, この仮定によって, $\delta < 1$ で条件を満たすことが可能となる. また, 上記の不等式において, $n = 1$ の場合も, 右辺の第 2 項をゼロと理解することによって成り立つことに注意する.

さて, 上記の不等式の両辺に $1 - \delta$ を掛けて整理すると,

$$\delta \in I = \Big[\max\Big\{ \frac{T-R}{T-P},\, \frac{T-R}{R-S} \Big\},\, 1 \Big)$$

の範囲で

$$f_n(\delta) = -(T-R) + (T-P)\delta + (P-S)\delta^n - (R-S)\delta^{n+1} \geq 0$$

を示せばよいことがわかる.

$$a = \frac{T-R}{T-P},$$

$$\frac{f_n(\delta)}{T-P} \equiv g_n(\delta) = -a + \delta + \frac{P-S}{T-P}\delta^n \Big(1 - \frac{R-S}{P-S}\delta \Big)$$

とおく.

Case I: $1 - \dfrac{R-S}{P-S}\delta \geq 0$ の場合, $\delta \in I$ ならば, $-a + \delta \geq 0$ かつ, 第 3 項も非負だから, $g_n(\delta) \geq 0$ を得る. なお, $T - R$ と $P - S$ の大小関係は決められないことに注意する.

Case II: $1 - \dfrac{R-S}{P-S}\delta < 0$ の場合, $\delta \in I$ において,

$$g_{n+1}(\delta) - g_n(\delta) = \frac{P-S}{T-P}\Big(1 - \frac{R-S}{P-S}\delta \Big)(\delta - 1)\delta^n \geq 0.$$

よって, 帰納的に $\delta \in I$ かつ Case II の範囲で, $0 \leq g_1(\delta) \leq g_2(\delta) \leq \cdots$ を得る. いずれの場合も $\delta \in I$ の範囲で $f_n(\delta) \geq 0$ となる. (Lemma の証明終わり)

定理 8.2 の証明. ゲーム理論の教科書にもこの定理の「証明」は書いてあるが, 多分に直観的で戦略が必ずしも正確に確率過程である場合も含まれているのか疑わしいように思われる.

プレーヤー I も II も "TFT"戦略 $\mathbf{X}$, $\mathbf{X}'$ をとるとすると, 互いの手は常に C となるから, その平均利得は

$$\begin{aligned} u(\mathbf{X}', \mathbf{X}) &= E\Big[\sum_{n=1}^{\zeta} u(X'_n, X_n) \Big] \\ &= \frac{R}{1-\delta} \equiv U. \end{aligned}$$

次に,

$$A = \sup_{\mathbf{Y}} u(\mathbf{Y},\, \mathbf{X})$$

とおく. ここで, supremum はあらゆる抽象確率空間上で定義された, あらゆる戦略 **(Y, X)**, **X**=TFT の範囲で考える. 定義より, 任意の $\varepsilon > 0$ に対して, プレーヤー II の戦略 $\mathbf{Y} = (Y_1, Y_2, \ldots)$ が存在して,

$$A - \varepsilon < E\Big[\sum_{n=1}^{+\infty} u(Y_n,\, X_n)\delta^{n-1}\Big]$$

が成り立つ. **X'**=TFT と **Y** は独立であると仮定して, 次の評価式が得られる.

$$\begin{aligned} u(\mathbf{X}',\, \mathbf{X}) - A + \varepsilon \;&>\; E\Big[\sum_{n=1}^{+\infty}(u(X'_n,\, X_n) - u(Y_n,\, X_n))\delta^{n-1}\,;\, Y_1 = C\Big] \\ &\quad + E\Big[\sum_{n=1}^{+\infty}(u(X'_n,\, X_n) - u(Y_n,\, X_n))\delta^{n-1}\,;\, Y_1 = D\Big] \\ &\equiv\; I_C + I_D. \end{aligned}$$

$$\begin{aligned} I_C \;&=\; E\Big[\sum_{n=1}^{+\infty}(u(X'_n,\, X_n) - u(Y_n,\, X_n))\delta^{n-1}\,;\, Y_1 = C,\, \tau_D(Y) = \infty\Big] \\ &\quad + E\Big[\sum_{n=1}^{+\infty}(u(X'_n,\, X_n) - u(Y_n,\, X_n))\delta^{n-1}\,;\, Y_1 = C,\, \tau_D(Y) < +\infty\Big] \\ &=\; E\Big[\sum_{n=1}^{\tau_D(Y)-1}(u(X'_n,\, X_n) - u(Y_n,\, X_n))\delta^{n-1}\,;\, Y_1 = C,\, \tau_D(Y) < +\infty\Big] \\ &\quad + E\Big[\sum_{n=\tau_D(Y)}^{+\infty}(u(X'_n,\, X_n) - u(Y_n,\, X_n))\delta^{n-1}\,;\, Y_1 = C,\, \tau_D(Y) < +\infty\Big] \\ &=\; E\Big[\delta^{\tau_D(Y)-1}\sum_{n=1}^{+\infty}\begin{pmatrix} u(X'_{n+\tau_D(Y)-1},\, X_{n+\tau_D(Y)-1}) \\ -u(Y_{n+\tau_D(Y)-1},\, X_{n+\tau_D(Y)-1}) \end{pmatrix}\delta^{n-1}\,;\, Y_1 = C,\, \tau_D(Y) < +\infty\Big] \\ &=\; E\Big[\delta^{\tau_D(Y)-1}\sum_{n=1}^{+\infty}\begin{pmatrix} u((\theta_{\tau_D(Y)-1}X')_n,\, (\theta_{\tau_D(Y)-1}X)_n) \\ -u((\theta_{\tau_D(Y)-1}Y)_n,\, (\theta_{\tau_D(Y)-1}X)_n) \end{pmatrix}\delta^{n-1}\,;\, \begin{matrix} Y_1 = C, \\ \tau_D(Y) < +\infty \end{matrix}\Big] \\ &\quad (\theta_{\tau_D(Y)-1}\mathbf{X} \text{ は “TFT”戦略であり, つぎの remark も参照して}) \\ &\geq\; (U - A)E[\,\delta^{\tau_D(Y)-1}\,;\, Y_1 = C,\, \tau_D(Y) < +\infty\,] \end{aligned}$$

を得る.

Remark: $\tau_D(Y) - 1$ はマルコフ時間ではないが, **X** が “TFT”戦略であることと $\tau_D(Y)$ の定義から, 確率過程 $(\theta_{\tau_D(Y)-1}\mathbf{Y},\, \theta_{\tau_D(Y)-1}\mathbf{X})$ が $\{Y_1 = C,\, \tau_D(Y) < +\infty\}$ 上で定義可能で, かつ我々の仮定 (i), (ii) を満たしている. さらに, $\theta_{\tau_D(Y)-1}\mathbf{X}$ は “TFT”戦略となるか

ら, A (定数) の定義から上記の不葬式が得られる.
一方,

$$\begin{aligned}
I_D &= E\Big[\sum_{n=1}^{+\infty}(u(X'_n,\,X_n)-u(Y_n,\,X_n))\delta^{n-1}\,;\,Y_1=D,\,\tau_C(Y)=\infty\Big]\\
&\quad + E\Big[\sum_{n=1}^{+\infty}(u(X'_n,\,X_n)-u(Y_n,\,X_n))\delta^{n-1}\,;\,Y_1=D,\,\tau_C(Y)<+\infty\Big]\\
&= E\Big[\frac{R}{1-\delta}-\Big(T+\frac{\delta P}{1-\delta}\Big)\,;\,Y_1=D,\,\tau_C(Y)=+\infty\Big]\\
&\quad + E\Big[\sum_{n=1}^{+\infty}(u(X'_n,\,X_n)-u(Y_n,\,X_n))\delta^{n-1}\,;\,Y_1=D,\,\tau_C(Y)=2\Big]\\
&\quad + E\Big[\sum_{n=1}^{+\infty}(u(X'_n,\,X_n)-u(Y_n,\,X_n))\delta^{n-1}\,;\,Y_1=D,\,2<\tau_C(Y)<+\infty\Big]\\
&\equiv I_{D,1}+I_{D,2}+I_{D,3}.
\end{aligned}$$

とおくと, δ に対する仮定より

$$I_{D,1}=\frac{\delta(T-P)-(T-R)}{1-\delta}\,\mathrm{Prob}(Y_1=D,\,\tau_C(Y)=\infty)\geq 0.$$

また,

$$\begin{aligned}
I_{D,2} &= (R+\delta\,R-T-\delta\,S)\,\mathrm{Prob}(Y_1=D,\,\tau_C(Y)=2)\\
&\quad + E\Big[\sum_{n=3}^{+\infty}(u(X'_n,\,X_n)-u(Y_n,\,X_n))\delta^{n-1}\,;\,Y_1=D,\,\tau_C(Y)=2\Big]\\
&= (\delta(R-S)-(T-R))\,\mathrm{Prob}(Y_1=D,\,\tau_C(Y)=2)\\
&\quad + E\Big[\delta^2\sum_{n=1}^{+\infty}(u((\theta_2X')_n,\,(\theta_2X)_n)-u((\theta_2Y)_n,\,(\theta_2X)_n))\delta^{n-1};\,\begin{matrix}Y_1=D,\\ \tau_C(Y)=2\end{matrix}\Big]\\
&\quad (\ \delta\geq\frac{(T-R)}{(R-S)}\text{かつ }\theta_2\mathbf{X}\text{ は “TFT”戦略だから, }I_C\text{ の場合と同様にして})\\
&\geq \delta^2(U-A)\,\mathrm{Prob}(Y_1=D,\,\tau_C(Y)=2)
\end{aligned}$$

となる. 最後に,

$$\begin{aligned}
I_{D,3} &= E\Big[\sum_{n=1}^{\tau_C(Y)}(u(X'_n,\,X_n)-u(Y_n,\,X_n))\delta^{n-1}\,;\,Y_1=D,\,2<\tau_C(Y)<+\infty\Big]\\
&\quad + E\Big[\sum_{n=\tau_C(Y)+1}^{+\infty}(u(X'_n,\,X_n)-u(Y_n,\,X_n))\delta^{n-1};\,Y_1=D,\,2<\tau_C(Y)<+\infty\Big]\\
&= E\Big[R\sum_{n=0}^{\tau_C(Y)-1}\delta^n-\Big(T+S\delta^{\tau_C(Y)-1}+P\sum_{n=1}^{\tau_C(Y)-2}\delta^n\Big);\,\begin{matrix}Y_1=D,\\ 2<\tau_C(Y)<+\infty\end{matrix}\Big]\\
&\quad + E\Big[\delta^{\tau_C(Y)}\sum_{n=1}^{+\infty}\begin{pmatrix}u((\theta_{\tau_C(Y)}X')_n,\,(\theta_{\tau_C(Y)}X)_n)\\ -u((\theta_{\tau_C(Y)}Y)_n,\,(\theta_{\tau_C(Y)}X)_n)\end{pmatrix}\delta^{n-1}\,;\,\begin{matrix}Y_1=D,\\ 2<\tau_C(Y)<+\infty\end{matrix}\Big]\\
&\text{(Lemma 1 を考慮して, さらに }\theta_{\tau_C(Y)}\mathbf{X}\text{ が “TFT”戦略だから }I_C\text{ の場合と同様にして)}\\
&\geq (U-A)E[\delta^{\tau_C(Y)}\,;\,Y_1=D,\,2<\tau_C(Y)<+\infty].
\end{aligned}$$

を得る. 以上を総合して,

$$U - A + \varepsilon \quad \geq \quad (U - A)\begin{pmatrix} E[\,\delta^{\tau_D(Y)-1}\,;\, Y_1 = C,\, \tau_D(Y) < +\infty\,] \\ + E[\,\delta^{\tau_C(Y)}\,;\, Y_1 = D,\, \tau_C(Y) < +\infty\,] \end{pmatrix}.$$

ところが, $\tau_D(Y) \geq 2$ on $\{Y_1 = C\}$ かつ $\tau_C(Y) \geq 2$ on $\{Y_1 = D\}$ だから

$$E[\delta^{\tau_D(Y)-1}\,;\, Y_1 = C,\, \tau_D(Y) < +\infty] + E[\delta^{\tau_C(Y)}\,;\, Y_1 = D,\, \tau_C(Y) < +\infty] \leq \delta < 1$$

である. したがって, $U - A + \epsilon \geq \delta(U - A)$ が得られ, $\epsilon > 0$ は任意だから, 結論

$$U \geq A$$

を得る. (定理 8.2 の証明終わり)

§9 Trigger v.s. T.F.T.:サンクションとしての有効性[6]

アクセルロッド ([3]) 以来, T.F.T (オウム返し, しっぺ返し) 戦略が最良の戦略であるかのごとき印象が世間的に広まっているが, あらゆる戦略に対して優越しているわけではなく, あくまで平均的に強い, という意味である. 本節において, ある限られた状況下においてではあるが, 理論的に言って, Trigger 戦略の方が T.F.T. 戦略より明らかに有利であることを証明しよう. このことは例えば, 村社会のように閉鎖的, 安定な社会にあってはTrigger 戦略を「村八分にすること」, と解釈すると日本社会において「村八分」がサンクションとして如何に強力な戦略であるか, ということの理論的説明でもある.

「村八分」と聞いて人は何を連想するであろうか. 大辞林で「村八分」を引いてみると,
(1) 江戸時代以来, 村落で行われた制裁の一. 規約違反などにより村の秩序を乱した者やその家族に対して, 村民全部が申し合わせて絶交するもの. 俗に, 葬式と火災の二つの場合を例外とするからという.
(2) 仲間はずれにすること.
とある. 一方, Yahoo で「村八分 法」で検索すると 2,600 件ヒットする. (ちなみに,「村八分」だけで検索すると 11,800 件ヒットするが, 同名の音楽グループがあるせいらしく,「村八分 音楽」で検索すると 2,750 件ヒットする).
現在「村八分」, という言葉を用いる場合, 多くは上記 (2) の意味で使われており, さほど深刻な状況ではないようである. 本来の意味については江戸時代, あるいはせいぜい明治期まで日本の農村で行われていた, 近代市民社会ではあってはならない「封建的で前近代的」な遅れた社会慣習である, という理解のされ方が普通ではないだろうか ([7]). 実際, 京都大学附属図書館ではキーワードとして「村八分」で検索しても「該当する書誌はありません」と返される. いろいろ検索するうちに, 実は「村八分」が明治期以降の犯罪のひとつとして取り上げられていることが判明した ([8]). 現実に, 現代でも深刻な例は後をたたないようで, 同じく Yahoo で検索してみると, 例えば兵庫県佐用町水根部落における人権侵害事件裁判を知ることが出来る.
このように,「村八分」という概念は単なる比喩以上の意味で現代の日本社会にそれなりに機能していると考えられる. ならば, それを単に人権を侵害する悪習である, と切って捨てるのではなく, 何故日本社会にかくも深く浸透しているのであろうか, と考えてみる価値はあるのではないだろうか.

ところで,「村八分」された場合, それはいつ解除されるのであろうか. どうも決まったルール, 了解事項はないように思われる. 仲介者を通していわゆる「詫びを入れる」 (全面的に恭順の意を表する, 無条件降伏) という言い方はあるが, 一旦「村八分」にされると, その程度で (と言ってこれ以上の方法は難しいが) 許されるのであろうか ([9]). ここまで考

[6]第 35 回数理社会学会 (2003.3.15-16 於大分大学. 一般講演「「村八分」の進化ゲーム論的考察—Trigger v.s. T.F.T.:サンクションとしての有効性— の原稿に加筆
[7]平凡社大百科事典: むらはちぶ (村八分) (福田アジオ)
[8](1993) 磯川全次・田村 勇・畠山 篤共著『犯罪の民俗学』批評社, 第八章「村八分とその心意伝承」(1993).
[9]平凡社大百科事典: むらはちぶ (村八分) (福田アジオ) 「期間が定められるということがほとんどなかっ

察し, ゲーム理論とをつきあわせて考えると, 「村八分」とは 「Trigger」戦略ではないだろうか. という連想に辿り着く. 日本人の意識の中に根強く「村八分」という概念が存在する以上, 進化ゲーム理論的に考えると, 「Trigger」戦略が他の戦略に比べて何らかの意味で長期的に見て, 有効なサンクションでなくてはならない.

本研究では, 「村八分」に対する批判として対比される近代的刑法, つまり, 「罪」に応じて予め決められた「代価」を支払えば, それ以上のサンクションを科さない, 科してはいけない (人権侵害である) という立場を 「TFT」戦略 (オウム返し, 協力には協力で, 裏切りには裏切りで対応する) とみなして, 「割引率のある繰り返し囚人のジレンマ」の枠組みを用いて, 両者の損得を比較する. その結果, 「村八分」が実はそれなりに合理性があることを数理モデルによって示す.

日本近代史の研究者田中によると ([10]), 「 (江戸時代の) 多くの村の秩序は, 封建社会を維持するために幕府や奉行所がつくったものではなく, 村人が自分たちの生活を成り立たせるため, 自らの意志でつくり出したものであった」 (129 頁) , あるいは, 「村人がとりきめた秩序 (拘束) は, 封建的秩序でもなければ身分的秩序でもなく, 生きていくために必要な秩序で, それが生産の上で大切な役割を果たすかぎり, 現代でも必要なものであることが納得できる」 (131 頁 - 132 頁) と述べている. この見方が正しいとするならば, 「村八分」のような村社会の「掟」は社会的秩序発生の生きた教材として進化ゲーム理論的に扱うことはまったく自然な成り行きである.

ゲーム理論的に考えたときに, 戦略としての「村八分」がはたして Trigger 戦略と言えるかどうかを考えてみたい.

福田 ([11]) によると, 「村として違法行為と考えられることに対しては制裁処分が行われた. いま紹介したものだけでも, 閉門, 罰金, 除名, 追放が登場している. 」とある. 言葉から類推すると, 罰金は一時的制裁 (TFT 戦略) であり, 追放は半永久的制裁 (Trigger 戦略) であるように思われる. 閉門と除名についてはいつ解除されるのか, 解除するルールが存在していたかどうかによって, 判断は分かれる. 言葉の印象からは罰金の方が追放に比べて軽い「制裁」ではないか, と思われる. つまり, TFT 戦略の方が Trigger 戦略に比してサンクションとしての効果は小さい, という予見が成り立つ.

では, 「村八分」はどのように位置付けられるであろうか. 福田 (同) によると, 「天明八年 (1788) の越前国大野郡大月村の「相定申証文之事」で, 5 箇条の規程を記した末尾に, 「此上, 相背候者有之ば, 可為村八分候」としている. 」とあるところを見ると制裁の最後の手段であったのかも知れない. ということは最強の「制裁」であったとも考えられる. とするならば, 一旦「村八分」が発動された場合に福田が主張するごとく短期間でこの制裁が終了したとは考えにくい. つまり, 本研究でモデル化しているように Trigger 戦略を対応されるのが妥当である, と考えられる.

「村八分」が村の「掟」として生活の中で次第に形作られてきたと考えられるので, その形態や, 刑罰としての軽重は当然江戸期を通じても地域差, 時間差はあると思われるが, 明治以降「村八分」という用語が定着し, もっぱら「仲間はずし」にすることという, 一定

たことは, 制裁がそれほど長時間でなかったことを示している」, とあるが, 根拠のない推測のように思われる.

[10](2000) 田中圭一 『百姓の江戸時代』ちくま新書 270(2000).

[11]塚本 学 編, 『日本の近世 第 8 巻 村の生活文化』 中央公論社 (1992), 「3. 村の共同と秩序 (福田アジオ)」, 113 頁

の集団内の私的制裁として恐れられ, 時には近代的法意識からは逸脱した違法な制裁として人権問題に発展するほど日本社会に根強く残る「規範」であることを考えると, もっと科学的, 合理的分析方法による研究があってもよいのではないか, と考えたのが本研究の動機である.

河上 ([13], 5 頁) が「自発的遵守, ないし無自覚な習熟が, 法の妥当を保障しているのであり, その源泉は歴史社会の基層にある」と主張するように「生ける法は国家権力の強制によってその妥当性を担保するという必要はない. 」 (同書第 1 編第 2 章「生ける法の研究—エールリッヒの法思想」 (飯野靖夫) 40 頁) , ここで権力による強制が担保されていない「法」とは, 1) 社会集団の成員に共有されること, 2) 個体の行動が規制されること, 3) 集団のメンバーが他個体にもこの規制がかかることを期待すること, 4) 逸脱を排除する何らかの機構が働くこと, (同書第 4 編第 2 章 241 頁「チンパンジー社会における規範と社会意識」 (黒田末壽) と考えるのが妥当であろう. なお, 彼は「多数の成員が逸脱者に向ける嫌悪が有効な懲罰になることは, 民族誌を紐解くまでもなく, 私たちの日常にありふれた事柄である. 」 (同 240 頁) とあるが, 江戸時代のように, 相当に整備された制裁としての「村八分」は他の未開社会, 中世封建社会, 近代的市民社会を通じて日本独特である, という印象を持つ (これらの報告例, 研究報告, 解説書を見出すことが出来なかった). 彼自身, 日本の「村八分」や「仲間はずし」をイメージしたのではないだろうか (チンパンジーの社会でも. 「多数の成員が逸脱者に向ける嫌悪が有効な懲罰になり得る」ことがよく知られている.)

ゲーム理論による定式化

非協力 (裏切り, 「D」で表す) が常に短期的には有利であっても, 長い付き合いの過程では協力 (「C」で表す) を選ぶ方が有利となる場合もあることは, 割引率 δ を持つ繰り返し囚人のジレンマ・ゲームによって表現できることが知られている ([163]). 農耕を基盤とする農村社会においては, まさに割引率を持つ繰り返し囚人のジレンマ・ゲームが行われている, と考えるのは自然な定式化であろう. この枠内で「村八分」という戦略を考えると, ある一人のプレーヤーが一度 D を出した場合, 他の全員が永久に D を出すこと (この場合の D は「非協力」というよりは「制裁」という意味になるが) , つまり Trigger 戦略である, と捉えることが出来る. 「村八分」を封建的で遅れた, 非合理な習慣とみなす人たちの戦略は, このようなゲーム理論的枠組みでは,「TFT」戦略とみなすことが出来るであろう. つまり,「村八分」がある種の合理性を持っていることを示すには, ゲーム理論的枠組みの中で, Trigger 戦略が TFT 戦略より有利であることを示せばよい. なお,「Trigger」戦略と 「TFT」 戦略はこの両者のみでは, たとえ両者が入り混じっていても見かけ上の表現形態は常に C が実現して区別できない. また,「all-C」や「all-D」戦略に対する利得も両者で区別がない.

では, Trigger 戦略と TFT 戦略の優劣が分かるような相手の戦略は何であろうか. 本研究では, パラメータ $(0 < \varepsilon < 1)$ で表される「日和見」戦略のみを考察することにする. ここで, 「日和見」戦略は実際の表現形態によって二通りに分かれる. 一つは 「all-C 的日

和見」戦略, つまり, 原則として all-C 戦略であるが, 時々 (ランダムに) D を出してみる戦略である. もう一つは 「all-D 的日和見」戦略, つまり, 原則として all-D 戦略であるが時々 (ランダムに) C を出してみる戦略である. この両者は数学的構造としては同じである, とみなせる. つまり, 「日和見」戦略とは, 繰り返し囚人のジレンマ・ゲームの n 回目の手を確率変数 Y_n で表した時, 確率変数列 $\mathbf{Y} = (Y_1, Y_2, \ldots)$ が *i.i.d.*(独立同分布を持つ確率変数列) として表現できる. $\mathrm{Prob}(Y_n = C) = \varepsilon$ とおくと, ε が 1 に近いときが「all-C 的日和見戦略」であり, 0 に近いときが 「all-D 的日和見戦略」である, とみなすことが出来る.

通常のように, 1 回の囚人のジレンマ・ゲームの戦略セットが (C, D) の時の C の利得を $u(C, D) = S$ とおく. 以下同様に $u(C, C) = R,\ u(D, C) = T,\ u(D, D) = P$ とする. 囚人のジレンマだから, $T > R > P > S$ と $2R > T + S$ を仮定する. 繰り返しゲームの利得計算にあたってはいわゆる割引率 δ を考慮した無限和を用いる. 混合戦略であるから, 平均利得で比較する.

具体的には, プレーヤー I の手を表す確率変数列 $\mathbf{X} = (X_1, X_2, \ldots)$ とプレーヤー II の手を表す確率変数列 $\mathbf{Y} = (Y_1, Y_2, \ldots)$ が与えられた時, プレーヤー I の利得 $U(\mathbf{X}, \mathbf{Y})$ は

$$U(\mathbf{X}, \mathbf{Y}) = \sum_{n=1}^{+\infty} E[u(X_n, Y_n)]\delta^{n-1}$$

で表される.

Trigger 戦略者 $\mathbf{X_1}$ が日和見戦略者 $\mathbf{Y}$ と対戦した時の利得 $U(\mathbf{X_1}, \mathbf{Y})$ と TFT 戦略者 $\mathbf{X_2}$ が日和見戦略者 $\mathbf{Y}$ と対戦した時の利得 $U(\mathbf{X_2}, \mathbf{Y})$ を比較してみよう.

日和見戦略者が初めて D を出す時刻 (確率変数, マルコフ時間である) $\tau_D(Y)$ および, 初めて D を出した後に初めて C を出す時刻 (確率変数) $\tau_C^*(Y)$ は次のように定義できる.

$$\tau_D(Y) = \min\{n \geq 1\,;\, Y_n = D\},\quad \tau_C^*(Y) = \min\{n > \tau_D(Y)\,;\, Y_n = C\}.$$

これらの概念を利用すると, 確率過程論のスタンダードなテクニックによって次のような結果を得る. (証明は後述)

$$U(\mathbf{X_1}, \mathbf{Y}) = \frac{\varepsilon}{1-\varepsilon\delta}R + \frac{(1-\varepsilon)}{1-\varepsilon\delta}S + \frac{\varepsilon(1-\varepsilon)\delta}{(1-\delta)(1-\varepsilon\delta)}T + \frac{(1-\varepsilon)^2\delta}{(1-\delta)(1-\varepsilon\delta)}P, \tag{9.1}$$

$$U(\mathbf{X_2}, \mathbf{Y}) = \frac{\varepsilon(1-\delta+\varepsilon\delta))}{1-\delta}R + \frac{(1-\varepsilon)(1-\delta+\varepsilon\delta)}{1-\delta}S + \frac{\varepsilon(1-\varepsilon)\delta}{1-\delta}T + \frac{(1-\varepsilon)^2\delta}{1-\delta}P. \tag{9.2}$$

これらの結果より,

$$U(\mathbf{X_1}, \mathbf{Y}) - U(\mathbf{X_2}, \mathbf{Y}) = \frac{(1-\varepsilon)\varepsilon^2\delta^2}{(1-\varepsilon\delta)(1-\delta)}(T-R) + \frac{(1-\varepsilon)^2\varepsilon\delta^2}{(1-\varepsilon\delta)(1-\delta)}(P-S) > 0$$

が得られる. つまり, δ と ε の値に拠らずに常に Trigger 戦略の方が TFT 戦略に比べて日和見戦略に対して有利である. このことは「村八分」がサンクションとして近代的法治主義より有効であることを意味しているのではないだろうか.

次に, メイナード・スミス ([43]) に従って, 進化ゲーム論的に考察してみよう. 「村八分」にしたグループとされたグループがランダムマッチングをするとも思えないが, 兎に角パラメータ $0 < \alpha < 1$ を導入して, Trigger 戦略者の平均利得 $U(\text{Trigger})$ を

$$U(\text{Trigger}) = (1-\alpha)U(\mathbf{X_1}, \mathbf{X_1}) + \alpha U(\mathbf{X_1}, \mathbf{Y})$$

で定義する. 同様に日和見戦略者の平均利得 $U(\text{日和見})$ を

$$U(\text{日和見}) = (1-\alpha)U(\mathbf{Y}, \mathbf{X_1}) + \alpha U(\mathbf{Y}, \mathbf{Y})$$

で定義する.

$$\begin{aligned} U(\text{Trigger}) - U(\text{日和見}) &= (1-\alpha)(U(\mathbf{X_1}, \mathbf{X_1}) - U(\mathbf{Y}, \mathbf{X_1})) \\ &\quad + \alpha(U(\mathbf{X_1}, \mathbf{Y}) - U(\mathbf{Y}, \mathbf{Y})) \\ &= U(\mathbf{X_1}, \mathbf{X_1}) - U(\mathbf{Y}, \mathbf{X_1}) \\ &\quad + \alpha(U(\mathbf{X_1}, \mathbf{Y}) + U(\mathbf{Y}, \mathbf{X_1}) - U(\mathbf{X_1}\,\mathbf{X_1}) - U(\mathbf{Y}, \mathbf{Y})) \\ &\equiv f(\alpha) \text{ とおく.} \end{aligned}$$

よく知られているように, $\delta \geq (T-R)/(T-P)$ のとき, Trigger 戦略はナッシュ均衡戦略であるから, $f(0) \geq 0$ である. 一方,

$$\begin{aligned} f(1) &= U(\mathbf{X_1}, \mathbf{Y}) - U(\mathbf{Y}, \mathbf{Y}) \\ &= -\frac{\varepsilon(1-\varepsilon)(1-\delta-\varepsilon\delta)}{(1-\delta)(1-\varepsilon\delta)}(T-R) - \frac{(1-\varepsilon)^2(1-\delta-\varepsilon\delta)}{(1-\delta)(1-\varepsilon\delta)}(P-S). \end{aligned}$$

であるから,

$$f(1) > 0 \Longleftrightarrow \frac{1-\delta}{\delta} < \varepsilon$$

が得られる. δ が 1 に近く, かつ ε が 1 に近い (all-C 的日和見) に対しては, $f(\alpha)$ が線分であることから, $0 < \alpha < 1$ のすべての値に対して $f(\alpha) > 0$ となり, Trigger 戦略 (村八分) が有利である. これにより, 「村八分」という戦略がいかに強力なサンクションであるかがわかる.

これに対して近代法的法治主義である TFT 戦略を同様に解析すると,

$$f(1) = -\varepsilon(1-\varepsilon)(T-R) - (1-\varepsilon)^2(P-S) < 0$$

となるから, たとえ δ が十分 1 に近くても (この場合は TFT もナッシュ均衡戦略であり, $f(0) \geq 0$ であるが) ϵ の値の如何に関わらず, α が 1 に近いと, つまり, 日和見戦略者の割合が高くなると, TFT 戦略がサンクションとして機能しなくなることがわかる. (最近の日本社会が連想されるのではないだろうか.)

discussion

千葉正士は ([12]) は「現代社会における未開のサンクション (primitive sanction) の問題が提起される. 現代人は組織的なかんずく法的サンクションをサンクションの典型ないし大部分とかんがえて, 現代社会にも実に多種多様な非限定的・インフォーマルないし社会的・道徳的・宗教的なサンクションが, やはり時には法的サンクションよりも強力であることをとかく忘れる. . . . そうしてみると, サンクションは, 未開社会のそれが現代社会のそれと区別して問題とされるよりも, 未開社会と現代社会と双方を通じて, 未開のサンクションと文明化したサンクション (civilized sanction) との連続群が問題にされ, その諸形態・諸機能が分析されるべきではなかろうか. 」と述べている. 江戸時代の村の「掟」の発生は, 未開社会と現代社会における「規範」を連続的に捉えるためのキーになるのではなかろうか.

石部雅亮 ([13]) を拾い読みしてみると, 「ヴェーバーは, 習俗 (Sitte), 習律 (Konvention) と法との間はきわめて流動的であり相互に移行しあうものととらえられている. 習俗というのは, 習慣や無反省の模倣に基づく行為の事実上の規則性である. このような段階から, やがてそれは諒解 (Einverständenisse) に基づく共同社会行為に進み, 一定の習慣となった行為が拘束力をもつという観念が成立するにいたる. 」「習俗, 習律から法への移行は, 法規範の成立の道の一つであり, とくにそこには伝統の力が強く作用しているとみられる. . . . ヴェーバーにとって, 法形成の主体は個人であって, 民族のような超個人的有機体ではない. . . . このように個人の創意から新しい内容の共同社会行為, したがって諒解が成立するが, それが永続的なものとなりうるかどうかは, また外的な生活条件に適合しているかどうかにかかっている. 」とある. ここには意識されていないにしろ, 法規範の成立, 発展が共同体個々人をプレーヤーとする進化ゲーム論的に分析, 理解し得るという可能性を示唆しているように思われる.

それでは, 何故に「村八分」はサンクションとして強力なのであろうか. 仲間はずれにする者と, された者が互いに「制裁する」, 「制裁された」, という共通の意識を持っていなければ「村八分」は「制裁」として機能しないであろう. ここで, 土居健郎の「甘えの構造」がひとつの示唆を与えているように思われる. 土居は「「甘え」とは一体感を求めることである」([14]) と述べている. 一体感を求めようとする相手を拒絶すること, あるいは拒絶されるのではないだろうか, と思い, 思わせること, は双方にとって, 「制裁された」, 「制裁する」という共通了解が生れる基盤であるように思われる. 結局, 「村八分」と「甘えの構造」は表裏一体となって相互に補完し強め合っているのではないだろうか. であるからこそ, 「村八分」と「甘え」とは日本人の社会慣習の中に根強く残っているのである.

「甘え」は母子関係を基礎にしていることは明白であるから, 万国共通どの民族においても観察されるであろうという考え方もあるが, 日本における社会関係の中での「甘え」はやはり一種独特のものがある.

---

[12]川島武宣編 『法社会学講座 9, 歴史・文化と法 1』岩波書店「第 3 節 (千葉正士) 未開社会におけるサンクションの諸形態」(1973), 48 頁 - 49 頁.

[13]川島武宣編 『法社会学講座 7, 社会と法 1』岩波書店「第 2 節 (石部雅亮) ヴエーバーの理論」(1973), 85 頁ー 86 頁.

[14]大塚久雄, 川島武宣, 土居健郎: 『「甘え」と社会科学』弘文堂選書 (1976), 55 頁.

たとえば, 斎藤 ([15]) によると, 友人関係において, 相互に何かをしてあげた, してもらった (一種の「甘え」) の認識は日米で相当に違う, との実証研究がある. 彼によると (同, 2 頁) , 日本の学生は「お互いがお互いの行動をモニターし, それに合わせた行動を心がける結果として, 関係性内の人々は自分たちの関係性について共通した認識を有する」のに対してアメリカ文化では (日本文化と比較した場合) , 互いの行動をモニターする程度は低く, 従って, 関係性内の人々が自分たちの関係性について共通認識を有する傾向は小さくなる」と述べている.

式 (9.1) と式 (9.2) の証明.

$\mathbf{Y} = (Y_1, Y_2, \ldots)$ は独立同分布確率変数列であるから, $\tau_D(Y)$ の分布は幾何分布

$$\mathrm{Prob}(\tau_D(Y) = k) = (1-\varepsilon)\varepsilon^{k-1}, \quad k = 1, 2, \ldots.$$

である. いま, $\mathbf{X_1}$ を Trigger 戦略, $\mathbf{Y}$ を日和見戦略とするとき,

$$\begin{aligned} U(\mathbf{X_1}, \mathbf{Y}) &= E\Big[\sum_{n=1}^{\infty} u(X_n, Y_n)\delta^{n-1}\Big] \\ &= E\Big[\sum_{n=1}^{\tau_D(Y)-1} u(X_n, Y_n)\delta^{n-1}\Big] + E[\,u(X_{\tau_D(Y)}, Y_{\tau_D(Y)})\delta^{\tau_D(Y)-1}\,] \\ &\quad + E\Big[\sum_{n=\tau_D(Y)+1}^{\infty} u(X_n, Y_n)\delta^{n-1}\Big] \\ &\equiv \mathrm{I} + \mathrm{II} + \mathrm{III} \end{aligned}$$

とおく. ここで,

$$E[\delta^{\tau_D(Y)-1}] = \sum_{k=1}^{\infty} \delta^{k-1}(1-\varepsilon)\varepsilon^{k-1} = \frac{1-\varepsilon}{1-\varepsilon\delta}$$

であるから, Trigger 戦略の定義と $\tau_D(Y)$ の定義 (意味) から

$$\begin{aligned} \mathrm{I} &= R\,E\Big[\sum_{n=1}^{\tau_D(Y)-1} \delta^{n-1}\Big] = R\frac{1-E[\,\delta^{\tau_D(Y)-1}\,]}{1-\delta} = \frac{R}{1-\delta}\Big(1 - \frac{1-\varepsilon}{1-\varepsilon\delta}\Big) \\ &= \frac{\varepsilon R}{1-\varepsilon\delta}. \\ \mathrm{II} &= SE[\,\delta^{\tau_D(Y)-1}\,] = \frac{1-\varepsilon}{1-\varepsilon\delta}S. \\ \mathrm{III} &= E\Big[\sum_{\tau_D(Y)+1}^{\infty} u(X_n, Y_n)\delta^{n-1}\Big] = \sum_{k=1}^{\infty} E\Big[\sum_{n=k+1}^{\infty} u(X_n, Y_n)\delta^{n-1}\,;\, \tau_D(Y) = k\Big] \\ &= \sum_{k=1}^{\infty}\sum_{n=k+1}^{\infty} (\varepsilon T + (1-\varepsilon)P)\delta^{n-1}(1-\varepsilon)\varepsilon^{k-1} \\ &= \frac{\delta(1-\varepsilon)}{(1-\delta)(1-\varepsilon\delta)}(\varepsilon T + (1-\varepsilon)P). \end{aligned}$$

[15]斎藤耕太, 「文化と互恵性」平成 15 年度 (2003) 人間・環境学研究科修士論文

が得られる. これらを合わせて, (9.1) 式

$$\begin{aligned} U(\mathbf{X_1}, \mathbf{Y}) &= \mathrm{I} + \mathrm{II} + \mathrm{III} \\ &= \frac{\varepsilon}{1-\varepsilon\delta} R + \frac{(1-\varepsilon)}{1-\varepsilon\delta} S + \frac{\varepsilon(1-\varepsilon)\delta}{(1-\delta)(1-\varepsilon\delta)} T + \frac{(1-\varepsilon)^2\delta}{(1-\delta)(1-\varepsilon\delta)} P \end{aligned}$$

を得る.

TFT 戦略に対する同様の計算は少々難しい. 本質的に,「マルコフ性」よりも強い性質である「強マルコフ性」を用いるからである. ただし, 今の場合は離散時間だから丁寧に計算して行けばよい. まず, $(Y_1, Y_2, \ldots)$ が *i.i.d.* だから,

$$\mathrm{Prob}(\tau_C^*(Y) = n / \tau_D(Y) = k) = \varepsilon(1-\varepsilon)^{n-k-1}$$

が得られる. ここで, $\mathrm{Prob}(A/B)$ は事象 B の下での, 事象 A の条件付確率を表す.

これによって,

$$\begin{aligned} E[\delta^{\tau_C^*(Y)-1}] &= \sum_{k=1}^{\infty} E[\delta^{\tau_C^*(Y)-1} / \tau_D(Y) = k](1-\varepsilon)\varepsilon^{k-1} \\ &= \sum_{k=1}^{\infty} \sum_{n=k+1}^{\infty} \delta^{n-1}\varepsilon(1-\varepsilon)^{n-k-1}(1-\varepsilon)\varepsilon^{k-1} \\ &= \sum_{k=1}^{\infty} \sum_{m=1}^{\infty} \delta^{m+k-1}\varepsilon^k(1-\varepsilon)^m \\ &= \frac{\varepsilon\delta}{1-\varepsilon\delta} \sum_{m=1}^{\infty} \delta^{m-1}(1-\varepsilon)^m \\ &= \frac{\delta\varepsilon(1-\varepsilon)}{(1-\varepsilon\delta)(1-\delta+\varepsilon\delta)} \end{aligned}$$

が得られる. 従って, いま, $\mathbf{X_2}$ を TFT 戦略, $\mathbf{Y}$ を日和見戦略とするとき,

$$\begin{aligned} U(\mathbf{X_2}, \mathbf{Y}) &= E\Big[\sum_{n=1}^{\infty} u(X_n, Y_n)\delta^{n-1}\Big] \\ &= E\Big[\sum_{n=1}^{\tau_D(Y)-1} u(X_n, Y_n)\delta^{n-1}\Big] + E[\,u(X_{\tau_D(Y)}, Y_{\tau_D(Y)})\delta^{\tau_D(Y)-1}\,] \\ &\quad + E\Big[\sum_{n=\tau_D(Y)+1}^{\tau_C^*(Y)-1} u(X_n, Y_n)\delta^{n-1}\Big] + E[\,u(X_{\tau_C^*(Y)}, Y_{\tau_C^*(Y)})\delta^{\tau_C^*(Y)-1}\,] \\ &\quad + E\Big[\sum_{n=\tau_C^*(Y)+1}^{\infty} u(X_n, Y_n)\delta^{n-1}\Big] \\ &\equiv \mathrm{I} + \mathrm{II} + \mathrm{III} + \mathrm{IV} + \mathrm{V} \end{aligned}$$

を個別に計算する. TFT 戦略の定義と $\tau_D(Y)$, $\tau_C^*(Y)$ の定義 (意味) から I, II は Trigger 戦略の場合と同じであるから,

$$\mathrm{I} = \frac{\varepsilon\, R}{1-\varepsilon\delta}, \qquad \mathrm{II} = \frac{1-\varepsilon}{1-\varepsilon\delta} S.$$

Ⅲ, Ⅳ については,

$$\begin{aligned}
\text{Ⅲ} &= P\,E\Big[\sum_{\tau_D(Y)+1}^{\tau_C^*(Y)-1}\delta^{n-1}\Big] \\
&= \frac{P}{1-\delta}E[\,\delta^{\tau_D(Y)}-\delta^{\tau_C^*(Y)-1}\,] \\
&= \frac{P}{1-\delta}\Big(\frac{(1-\varepsilon)\delta}{1-\varepsilon\delta}-\frac{\varepsilon(1-\varepsilon)\delta}{(1-\varepsilon\delta)(1-\delta+\varepsilon\delta)}\Big) \\
&= \frac{(1-\varepsilon)^2\delta}{(1-\varepsilon\delta)(1-\delta+\varepsilon\delta)}P,
\end{aligned}$$

$$\text{Ⅳ} = T\,E[\,\delta^{\tau_C^*(Y)-1}\,] = \frac{(1-\varepsilon)\varepsilon\delta}{(1-\varepsilon\delta)(1-\delta+\varepsilon\delta)}T.$$

V については少々注意が必要である. Trigger の場合とは決定的に異なる変形が必要である. つまり, マルコフ時間 $\tau_C^*(Y)$ だけシフトした後に始まる部分ゲームの概念が必要である. 厳密な定義については [197] を参照されたい. (直観的に理解することは可能)

$$\begin{aligned}
\text{V} &= E\Big[\sum_{n=\tau_C^*(Y)+1}^{\infty}u(X_n,\,Y_n)\delta^{n-1}\Big] \\
&= E\Big[\sum_{k=1}^{\infty}u(X_{\tau_C^*(Y)+k},\,Y_{\tau_C^*(Y)+k})\delta^{\tau_C^*(Y)+k-1}\Big] \\
&= E\Big[\,\delta^{\tau_C^*(Y)}\sum_{k=1}^{\infty}u((\theta_{\tau_C^*(Y)}X)_k,\,(\theta_{\tau_C^*(Y)}Y)_k)\delta^{k-1}\Big] \\
&= U(\mathbf{X_2},\,\mathbf{Y})E[\,\delta^{\tau_C^*(Y)}\,] \\
&= U(\mathbf{X_2},\,\mathbf{Y})\frac{\varepsilon(1-\varepsilon)\delta^2}{(1-\varepsilon\delta)(1-\delta+\varepsilon\delta)}
\end{aligned}$$

を得る. 従って,

$$U(\mathbf{X_2},\,\mathbf{Y}) = \text{I}+\text{Ⅱ}+\text{Ⅲ}+\text{Ⅳ}+\text{V}$$

において, V の中に $U(\mathbf{X_2},\,\mathbf{Y})$ を含むから, 上記の方程式を $U(\mathbf{X_2},\,\mathbf{Y})$ について解くと (9.2) 式が得られる.

$(\mathbf{Y},\,\mathbf{Y}')$ を二つの独立同分布の見極め戦略とした場合, $U(\mathbf{Y},\,\mathbf{Y}')$ の計算は容易で

$$U(\mathbf{Y},\,\mathbf{Y}') = \frac{\varepsilon^2}{1-\delta}R+\frac{\varepsilon(1-\varepsilon)}{1-\delta}T+\frac{\varepsilon(1-\varepsilon)}{1-\delta}S+\frac{(1-\varepsilon)^2}{1-\delta}P$$

となる.

§10. 繰り返し囚人のジレンマ (マルコフ連鎖として)

混合戦略のところですでにふれたように, 1 回のゲームで混合戦略をとる, ということの意味は混合戦略セットをひとつの確率変数とみなしたとき, 独立同分布な確率変数列に関する強大数の法則より, 多数回の試行の算術平均が確率 1 で 1 回の混合戦略セットの平均に収束する, という確率論における有名な定理によって基礎づけられている. しかしながら, くり返しゲームにおいては, 各回の戦略セット (確率変数) が独立ではなく, 例えば前回の相手の戦略によって次回の自分の戦略を変える, ということを定式化しようとすればより一層明白に確率過程の概念が必要であることが理解されるであろう.

「確率過程」とは確率分布の時間的変化を考察する, ということだけではない. 例えば, 平均 0, 分散 $t > 0$ が時間を表すとして, 正規分布

$$p(t,x) = \frac{1}{\sqrt{2\pi t}} \exp(-\frac{x^2}{2t})$$

を考えてみよう. 容易に分かるようにグラフは, 時間 t が小さいときは山が高くなり, t が 0 に近づく時はいわゆるディラックのデルタ関数に測度の意味で近づくが, ディラックのデルタ関数は普通の意味の関数ではないので, 関数の意味では原点を除いて 0 に各点収束する. 一方, t が無限大に近づくときは山はどんどんなだらかになって, 直線上の一様分布に近づくが全測度は 1 に保たれたままなので, 測度の意味では収束しないが, 恒等的に 0 という関数に一様収束する. $p(t,x)$ の確率分布としての興味はこの程度である. なお, 確率論では, 個別の現象は確率的にしか理解できないから, 確率過程は全体的平均量しか表現していない, と思われがちであるが, それは正しくない.

確率過程 $\{X(t,\omega)\,;\, 0 \le t < +\infty, \omega \in \Omega\}$ において $\omega =$ 根元事象を固定して $X(t,\omega)$ を t の関数として考えた時, sample (見本, 標本) または sample function (見本関数, 標本関数) といい, これが個別の行動を表している. そして, 例えば t を無限大にした時に, 確率 1 で標本が一定の性質を持つということが証明された場合は個別の状態は全体の性質でもあることになる (ミクロ・マクロの相互関係) . ただし, モデルによっては見本関数が何を表すか明確でない場合もある. 多くのモデルでは各 1 回の実験データの系列が見本関数である. つまり, **確率モデル**とは, 確率変数 (random variables) で何かを表現すること. 確率変数はコルモゴロフの公理系によって, 数学的に厳密に定義された概念である (主観確率ではない). すなわち, ある確率空間 (予め適当に定めておく) Ω から状態空間 (State Space) S への可測写像 (measurable map) である. マルコフ連鎖を考える場合は可測性の議論は必要ないので省略する. 写像であるから, $\Omega \ni \omega$ に対して, $S \ni X(\omega)$ が定義されている. この $X(\omega)$ を X の実現値という, 普通の関数の場合 f と $f(x)$ を混同して使うが厳密には f は関数自身を表し, $f(x)$ はこの関数の, 点 x に於ける値 (この場合, 実現値とは言わないが) を表す. 確率変数 X と $X(\omega)$ の関係は普通の関数 f と $f(x)$ の関係と同じことである.

注意: いろいろな事象の確率は確率変数の分布と見なせるが, 確率分布を定めるだけでは現象を表現したことにならない. 分布は同じでも確率変数としては違い得る.

マルコフ連鎖を定義するために, 最小限度の数学的準備をしておく.

定義 10.1. 確率過程 (stochastic process)($\{X_t; t \in T\}$, または $\{X(t)\,;\, t \in T\}$ で表す.) とは, 空でない任意の集合 T をパラメーター空間として持つ確率変数の族 (family) である. つまり, 各 $t \in T$ を固定する毎に, X_t または $X(t)$ が確率変数であることを意味する. 特に $T =$ 自然数ないし整数の時には, 確率変数列という. T が 1 次元ユークリッド空間 (の一部) でない時は, 確率場 (stochastic field) という言葉も使う.

マルコフ連鎖 (Markov Chain) についての original 論文は
A. A. Markov; Extension of the law of large numbers to dependent variables. (in Russian) Izv. Fiz-Mat. Obsc. pri Kazansk. Univ. (2 Ser.)vol. 15, no. 4(1906), pp. 135-156.
— ; Investigations of general experiments connected by a Markov chain. Zap. Akad. Nauk Fix. Mat. Otdel. VIII Ser. 25-3(1910). (in Russian) Collected Works, Izd. Akad. Nauk SSSR, 1951, pp. 465-509.

注意: マルコフ過程 (Markov process) とマルコフ連鎖 (Markov chain) の違い.
(1) 数学辞典第 3 版 1176 頁:「状態空間 S が可算集合であるような Markov 過程を Markov 連鎖と呼ぶ. 」
(2)J. L. Doob; Stochastic processes. Wiley series in Probability and Mathematical Statistics (1953). Chapiter V. Sect. 1. Markov chains – definitions. "A Markov chain is defined as a Markov process whose random variables can only assume values in a certain finite or denumerably infinite set. "

2, 3 の確率論の有名な専門家に尋ねたところ, 彼らの間でも必ずしも定義は一定していない. 時間が離散な場合をマルコフ連鎖と呼ぶ人は専門家の中にも多い.

有限マルコフ連鎖 (finite Markov chain) の「有限」は状態空間 S が有限集合の場合に用いる. 以下簡単のために状態空間 S が有限集合の場合を考える.

$$S = \{s_1, s_2, \ldots, s_N\}$$

とする. 数学上は可算無限集合であっても定義の部分はそのまま成立する. しかし, 以下に述べる定理については, 有限集合の場合と無限集合の場合は大いに異なる場合もある.

定義 10.2. 事象 B が与えられた時の, 事象 A の条件付き確率

$$\mathrm{Prob}(A/B) = \frac{\mathrm{Prob}(A \bigcap B)}{\mathrm{Prob}(B)}.$$

定義 10.3. $\{X_n\,;\, n = 0, 1, 2, \ldots\}$ が独立確率変数列であるとは, 任意の $n = 1, 2, 3, \ldots$ 任意の $s_0, \ldots, s_n$, に対して,

$$\mathrm{Prob}(X_n = s_n / X_0 = s_0, \ldots, X_{n-1} = s_{n-1}) = \mathrm{Prob}(X_n = s_n)$$

が成立する時をいう.

定義 10.4. $\{X_n\,;\,n=0,1,2,\ldots\}$ がマルコフ連鎖であるとは, 任意の $n=1,2,3,\ldots$ 任意の $s_0,\ldots,s_n$ に対して,

(10.1) $\quad \mathrm{Prob}(X_n=s_n/X_0=s_0,\ldots,X_{n-1}=s_{n-1})=\mathrm{Prob}(X_n=s_n/X_{n-1}=s_{n-1})$

が成立する時をいう. (10.1) をマルコフ性 (Markov property) という.

定義 10.5.

$$\mathrm{Prob}(X_n=j/X_{n-1}=i)=P_{i,j}^{(n)},\quad P^{(n)}=(P_{i,j}^{(n)}),\quad n=1,2,\ldots$$

$P^{(n)}$ をステップ $n-1$ における推移確率行列 (transition probability matrix) という. 定義から, $P^{(n)}$ は, 状態空間 S の要素の数を N とすると, N 次正方行列である.

定理 10.1. マルコフ連鎖は $n=1,2,\ldots$ に対する推移確率行列 $P^{(n)}$ と, 初期分布

$$\mathrm{Prob}(X_0=i)=\mu_i\quad i=s_1,\ldots,s_N$$

によって一意に定まる. (確率空間が一意に定まる, という意味ではない. 確率法則, すなわち任意の有限次元結合分布が一致する場合に 2 つの確率過程は同じとみなす.)

注意: 有限状態空間を持つ任意の確率変数列 $\{X_n\,;n=0,1,2,\ldots\}$ に対して,

$$\mathrm{Prob}(X_n=j/X_{n-1}=i)=P_{i,j}^{(n)},\quad P^{(n)}=(P_{i,j}^{(n)})\quad n=1,2,\ldots$$

を定義することは出来るが, これらと初期分布とだけでは確率法則は定まらない. 3 次元以上の結合分布に関する情報がないからである. マルコフ性を仮定した場合は, 3 次元以上の結合分布が 2 次元結合分布から決定出来ることを意味する. ちなみに, 独立確率変数列の確率法則は 1 次元分布から決定できる.

定義 10.6. $P^{(n)}$ が n に依存しない時, 時間的に一様なマルコフ連鎖という. 多くの教科書では, いちいち断らなければ時間的一様性ははじめから仮定することが多い.

さて, くり返しのあるゲームを我々の立場で定式化すると次のようになる. 第 k 回目のゲームにおける戦略セットを $\vec{X_k}=(X_{1,k},\ldots,X_{n,k})$ とする (n はプレーヤーの数). ここで, 任意有限次元の確率法則を定めると確率過程 $\{\vec{X_k}\,;\,k=1,2,\ldots\}$ が定まって, くり返しのあるゲームがひとつ定式化されたことになる. 現実のゲームはその実現値である. 特に, 前回の相手の戦略だけを考慮に入れて次回の戦略を決めるのであれば, この確率過程は離散時間のマルコフ過程となる. さらに, 各自の戦略集合が有限集合であれば, マルコフ連鎖であり, その規則が時間に依存しなければ, 時間的に一様なマルコフ連鎖となる. 時間的に一様なマルコフ連鎖については数学的によく研究されている. ここで必要なことだけを述べておく.

定理 10.2.
(i) 有限状態を持つ既約エルゴード的マルコフ連鎖は唯一の定常分布 $\vec{\mu} \in \mathcal{P}(S^{(n)})$ を持つ.
(ii) この時,

$$\forall i\,,\ \lim_{T\to+\infty}\frac{1}{T}u_i(\vec{X}_T) = \int_{S^{(n)}} u_i(s_1,\dots,s_n)\,d\vec{\mu}(s_1,\dots,s_n)$$

が確率 1 で成り立つ (強大数の法則).

マルコフ連鎖で定式化出来る, よく知られたくり返しゲームは Tit for Tat (TFT) と呼ばれる戦略である. これはしっぺ返し戦略またはオウム返し戦略と訳されている. 言葉通り, 相手が協力してくれれば, 次回は自分も協力し, 相手が裏切れば, 次回は自分も裏切り戦略を選ぶのである. ただし, 最初は無条件に協力を選ぶ (初期分布). アクセルロッドはコンピュータシミュレーションによってさまざまな戦略の中で, TFT 戦略が平均的に極めて優れた戦略であることを示して以来一躍有名になった (戦略としては以前からよく知られていた). ただ, この戦略の難点はノイズに弱いことである. つまり, 相手も TFT 戦略を採用しているとして, 相手の協力戦略を誤って裏切りと誤解すると, 次回は裏切り返し, それを見た相手も次次回裏切る, ...つまり, お互いに裏切り続ける, という悲惨な結果になる. 小さな誤解が取り返しのつかない悲劇を生むことになる点, 徹底した TFT はあまりおすすめ出来ない. イスラエルとパレスティナの関係を見ればわかるであろう. その改良版として TF2T つまり 2 回続けて裏切られたら, 次回から裏切り返すが 1 回なら許す, という戦略等々確率変数としての独立性の仮定をはずすといろいろな可能性が出てくる. TFT は定義からマルコフ連鎖であるが, TF2T は 2 重マルコフと呼ばれる確率過程になる. 2 戦略の場合, M. Nowak ([84]) が時間的に一様なマルコフ連鎖 (つまり, 前回相手が協力してくれたか裏切ったかによって次回自分が協力する確率を変える) の場合の囚人のジレンマのくり返しゲームについて詳細に分析している. 以下に概略を紹介しよう.

純粋戦略は $S = (C, D)$ の 2 点集合である. $Z_n = (X_n, Y_n)\,;\, n = 0, 1, \dots$ は $S \times S$ の値を取る時間的に一様なマルコフ連鎖とする. 従って, 推移確率行列は 4 次の行列で表される. 各状態は $(C,C), (C,D), (D,C), (D,D)$ の順に並んでいるものとする. ここで, 例えば $Z_n(\omega) = (C, D)$ は根元事象 ω (実現値) において, n 回目にプレーヤー A が戦略 C を出し, プレーヤー B が戦略 D を出したことを意味する. この時推移確率行列 P が次式で与えられているとする.

$$P = \begin{pmatrix} pp' & p(1-p') & (1-p)p' & (1-p)(1-p') \\ qp' & q(1-p') & (1-q)p' & (1-q)(1-p') \\ pq' & p(1-q') & (1-p)q' & (1-p)(1-q') \\ qq' & q(1-q') & (1-q)q' & (1-q)(1-q') \end{pmatrix}.$$

ここで, p または q はプレーヤー B が直前にそれぞれ C または D を出した時, プレーヤ A が C を出す条件付き確率である (D を出す条件付き確率はそれぞれ $(1-p)$ および $(1-q)$). プレーヤー B に対しても同様にプレーヤ A の直前の手によって C を出す条件付き確率をそれぞれ p' または q' とする. その上で互いに独立に戦略を決める. 従って, 例えば第 1 行第 1 列成分は直前に A, B 共に C だった時, 次に A, B 共に C を出す条件

付き確率を表している. また, 初期分布は次のような場合に限る. すなわち, A, B は互いに独立に C をそれぞれ確率 y, y' で出すものとする. 従って, 独立性から初期分布は

$$\pi_0 = (yy', y(1-y'), (1-y)y', (1-y)(1-y'))$$

と表される. このとき, プレーヤー A の戦略とは 3 次元ベクトル $(y, p, q),\ 0 \leq y \leq 1,\ 0 \leq p \leq 1,\ 0 \leq q \leq 1$ をひとつ決めることである. 例えば上で述べた TFT は $(1, 1, 0)$ で表される. また, 相手の手の如何に関わらず最初から最後まで C を出す戦略は $(1, 1, 1)$ で表される (この戦略を All C と記す). また, 相手の手の如何に関わらず最初から最後まで D を出す戦略は $(0, 0, 0)$ で表される (この戦略を All D と記す). 残念ながら TF2T は単純マルコフ過程ではないから, この定式化では表されない.

では, 利得関数はどう表されるのであろうか. 戦略 C, D に対しては囚人のジレンマを仮定しているから, 利得関数は

$$f(C,C) = R,\ f(C,D) = S,\ f(D,C) = T,\ f(D,D) = P,$$
$$T > R > P > S,\ R > (S+T)/2$$

である. しかし, プレーヤー A とプレーヤー B の戦略は独立ではないし, 1 回毎に分布は異なる. 従って, プレーヤー A が戦略 $\vec{a} = (y, p, q)$ を取り, プレーヤー B が戦略 $\vec{b} = (y', p', q')$ を取った時のプレーヤー A の利得 $u(\vec{a}, \vec{b})$ は最終的に無限に時間をかけて安定した (もし実現し得るならば) 時の利得とする, すなわち,

$$u(\vec{a}, \vec{b}) = \lim_{n\to} E_{\pi_0}[f(X_n, Y_n)] \tag{10.2}$$

である. ここで, $E_{\pi_0}[\cdot]$ は初期分布 π_0 と推移確率行列によって一意に決まったマルコフ連鎖 $Z_n\,;\, n = 0, 1, 2, \ldots$ の分布による平均である. この極限が存在しない時は, マルコフ過程が周期的な場合であるから, いわゆる算術平均をとれば (非周期的な場合も含めて) 極限が存在して定義可能である. すなわち,

$$u(\vec{a}, \vec{b}) = \lim_{n\to} \frac{1}{n} \sum_{k=1}^{n} E_{\pi_0}[f(X_k, Y_k)]. \tag{10.3}$$

で定義する.

有限マルコフ連鎖の一般論より, このマルコフ連鎖が既約 (従って正再帰) , 非周期的であれば (10.2) の極限は, 初期分布に依存せず推移確率行列によってのみ一意に決まる定常分布による積分で表される. すなわち,

$$r = p - q,\quad r' = p' - q',\quad s = \frac{q'r + q}{1 - rr'},\quad s' = \frac{qr' + q'}{1 - rr'}$$

とおくと,

(i) Case I; $|rr'| < 1$. この時, マルコフ過程は既約, 非周期的であって, 定常分布は

$$\pi = (ss', s(1-s'), (1-s)s', (1-s)(1-s'))$$

である. 従って,

$$\begin{aligned} u(\vec{a},\vec{b}) &= Rss' + Ss(1-s') + T(1-s)s' + P(1-s)(1-s') \\ &= (R-S-T+P)ss' - (P-S)s + (T-P)s' + P. \end{aligned}$$

となって, 当然初期分布のパラメーター y, y' には依存しない.
(ii) Case II; $r = r' = 1$. この時, マルコフ過程は状態 (C,C) と (D,D) がそれぞれ trap で, (C,D), (D,C) の 2 点がひとつの再帰類を作り, しかも周期は 2 である. 従って, (10.3) は

$$\begin{aligned} u(\vec{a},\vec{b}) &= Ryy' + \frac{S+T}{2}(y(1-y') + y'(1-y)) + P(1-y)(1-y') \\ &= (R-S-T+P)yy' + \frac{S+T-2P}{2}(y+y') + P. \end{aligned}$$

(iii) Case III; $r = r' = -1$. この時は, 再帰類が (C,C), (D,D) と (C,D), (C,D) の 2 つできて, いずれも周期は 2 である.

$$u(\vec{a},\vec{b}) = \frac{R+P}{2}(yy' + (1-y)(1-y')) + Sy(1-y') + T(1-y)y'.$$

(iv) Case IV; $r = 1$, $r' = -1$. この時は, (C,C) から出発したマルコフ連鎖は $(C,C) \Rightarrow (C,D) \Rightarrow (D,D) \Rightarrow (D,C) \Rightarrow (C,C)$ と順に回って行くから, 周期は 4 でかつ

$$u(\vec{a},\vec{b}) = \frac{R+S+T+P}{4}$$

となる. Case II と III の場合だけは初期分布に依存する.

上記の枠組みでもパラメーター (R,S,T,P,y,p,q) は結構多いので多くの場合分けが必要であるが, Nowak の結論のいくつかを以下に紹介しておく.
(1) ESS は存在しない.
(2) TFT($\vec{a} = (1,1,0)$) は被侵入不能戦略である. ただし, 戦略 $\vec{a} = (y,1,0)$, $y < 1$ はナッシュ均衡戦略ではない. (相手を信頼するならば最初から信頼した方がよい.)
(3) 被侵入不能戦略であるための必要条件は $p = 1$ である (協力に対しては必ず協力で返すこと) . ただし, q (裏切られても協力するお人好し) は度が過ぎると (一定値を越えると) より非協力な戦略が侵入可能.

以上の結果は上記の枠組みの中で可能なすべての戦略の中で議論したが, 最初から取り得る戦略に制限を設けるとどうなるであろうか. ノイズによって, 確実には協力, 非協力が返せない状況を考える. すなわち p, q に次のような制限を設ける.

$$\varepsilon > 0 \text{ を固定して } \ \varepsilon \le p \le 1-\varepsilon, \ \ \varepsilon \le q \le 1-\varepsilon$$

とする. この時の結論の一部を紹介する.
(1) この枠内での All D=$(0,\varepsilon,\varepsilon)$ は唯一の強ナッシュ均衡戦略である (従って, ESS).
(2) この枠内の TFT =$(1,1-2\varepsilon,\varepsilon)$ は $\dfrac{\varepsilon}{2\varepsilon(1-r)} < q$ を満たす戦略に支配される. ここで, プレーヤー i の戦略 $\nu_i \in \mathcal{P}(S_i)$ が戦略 $\nu_i' \in \mathcal{P}(S_i)$ を (弱) 支配するとは,

$$\forall \vec{\mu}, \ u_i(\vec{\mu}_{-i}\,;\,\nu_i) \ge u_i(\vec{\mu}_{-i}\,;\,\nu_i')$$

が成立すること, ただし, 少なくともひとつの戦略に対しては真に不等号が成立する時をいう.

少し悲観的結論ではある. 完全なる悪人はいないにしても性善説は所詮幻想に過ぎないのかもしれない. しかし, いくつかの結論は R, S, T, P の値に依存するので, これらの値をうまく決めてやるともう少しましな結論が導かれる. 希望を失ってはいけない.

繰り返し囚人のジレンマ・ゲームに関連する文献は [68], [69], [89], [182], [183], [196], [217] 等文字通り山のようにあるが, ここで定式化した方法や目的, 解釈が異なるので詳細は省略する.

以下, 参考文献として最初にあげてあるものは主に, ゲーム理論, 進化ゲームに関する一般的入門書ないし多少専門的な教科書およびゲーム理論的視点を生かした一般教養書で, ゲーム理論を知っているとより興味がもてると感じた書物をあげてある (もちろん, 筆者の目にとまったものに限られているが). その後にに分野別に纏めた文献は筆者が目にしたゲーム理論関係の文献を本文とは一応無関係に集めた. ゲーム理論が多方面の分野に応用されている様子がこの文献表からも感じられるであろう.

参考文献

[1] 青木昌彦, 奥野正寛編: 経済システムの比較制度分析, 東京大学出版会 (1996).

[2] 青木昌彦: 比較制度分析にむけて, 瀧澤弘和・谷口和弘訳, NTT 出版 (2001).

[3] R. アクセルロッド: つきあい方の科学, 松田裕之訳, HBJ 出版局, (1987).

[4] アビナッシュ・ディキシット, バリー・ネイルバフ: 戦略的思考とは何か—エール大学式理論の発想法—菅野 隆, 嶋津祐一訳 (1991) TBS ブリタニカ Avinash K. Dixit & Barry J. Nalebuff: Thinking Strategically: The Competitive Edge in Business, Politics and Everyday Life. (1991).

[5] アマルティア・セン: 集合的選択と社会的厚生, 志田基与師監訳, 勁草書房 (2000). Amartya Sen: Collective Choice and Social Welfare, Holden-Day, Inc. (1970).

[6] 石原英樹, 金井雅之: 進化的意思決定 シリーズ意思決定の科学 5, 朝倉書店 (2002).

[7] N. ウィーナー: サイバネティクス, 岩波 (1962)) Wiener, N.: Cybernetics : or, control and communication in the animal and the machine, J. Wiley, (1947),

[8] 海野道郎, 盛山和夫編: 秩序問題と社会的ジレンマ ハーベスト社 (1991).

[9] 遠藤利彦: 喜怒哀楽の起源—情動の進化論・文化論—, 岩波科学ライブラリー 41 (1996).

[10] 太田勝造: 法律, 社会科学の理論とモデル 7, 東京大学出版会 (2000).

[11] 岡田 章: ゲーム理論, (1996) 有斐閣.

[12] オーマン, R. J.: ゲーム理論の基礎, 丸山 徹・立石寛訳, 勁草書房 (1991). Aumann, R. J.: Lecture on Game Theory. (1989)

[13] 河上倫逸編: ゆらぎの法律学 規範の基層とそのダイナミズム, 風行社 (1997).

[14] 木村邦博: 大集団のジレンマ—集合行為と集団規模の数理— ミネルヴァ書房 (2002).

[15] R. ギボンズ: 経済学のためのゲーム理論入門, 福岡正夫・須田伸一訳, 創文社 (1995). R. Gibbons: Game Theory for Applied Economists. Princeton Univ. Press (1992).

[16] J. R. クレブス, N. B. ディビス: 行動生態学 (原著第 2 版) 山岸 哲, 巌佐 庸 共訳 蒼樹書房 (1991). Krebs, J. R. & Davies, N. B.: Introduction to Behavioural Ecology (2nd ed.) (1987).

[17] J. R. クレブス, N. B. ディビス: 進化からみた行動生態学 山岸 哲, 巌佐 庸 監訳 蒼樹書房 (1994). Krebs, J. R. & Davies, N. B.: Behavioural Ecology–An Evolutionary Approach (third ed.) (1984).

[18] 美しさをめぐる進化論—容貌の社会生物学 蔵 琢也 勁草書房 (1993).

[19] 佐伯 胖: 「きめ方」の論理ー社会的決定理論への招待, 東京大学出版会 (1980).

[20] 佐伯 胖, 亀田達也編著: 進化ゲームとその展開, 共立出版 (2002).

[21] 佐藤嘉倫: 意図的社会変動の理論 合理的選択理論による分析, 東京大学出版会 (1998).

[22] スーザン・ブラックモア: ミーム・マシーンとしての私, 上下 垂水雄二訳 草思社 (2000). Susan Blackmore, The Meme Machine. (1999)

[23] 鈴木光男編著: ゲーム理論の展開, 東京図書 (1973).

[24] 鈴木光男, 中邨健二郎: 社会システム—ゲーム論的アプローチ, 共立出版 (1976).

[25] 鈴木光男: ゲーム理論入門, 共立全書 239 (1981).

[26] 鈴木光男: 新ゲーム理論, 勁草書房 (1994).

[27] 鈴木光男, 武藤滋夫: 協力ゲームの理論 UP 応用数学選書 11. 東京大学出版会 (1985).

[28] 鈴木光男: ゲーム理論の世界, 勁草書房 (1999).

[29] 東京大学公開講座: ゲーム—駆け引きの世界—: 東京大学出版会 (1999).

[30] デズモンド・モリス: 舞い上がったサル: 中村保男訳 飛鳥新社 (1996). Desmond Morris The Human Animal, (1994)

[31] 永久壽夫: ゲーム理論の政治経済学ー選挙制度と防衛政策ー PHP 研究所 (1995).

[32] 中山幹夫: はじめてのゲーム理論 有斐閣ブックス (1997)

[33] 西田利貞: 人間性はどこから来たか 京都大学学術出版会 (1999)

[34] 西山賢一: 勝つためのゲームの理論—適応戦略とは何か—, Blue Backs, (1986).

[35] 野口 広: 不動点定理, 共立出版 (数学ワンポイント双書 25) (1979).

[36] W. パウンドストン: 囚人のジレンマ, 松浦俊輔他訳, 青土社 (1995). Theory of Games) (1959), 13-42.

[37] 長谷川壽一, 長谷川眞理子: 進化と人間行動, 東京大学出版会 (2000).

[38] フォン・ノイマン, モルゲンシュテルン: ゲームの理論と経済行動, 全 5 巻, 銀林 浩他訳, 東京図書 (1972), Theory of Games and Economic Behavior (1944).

[39] マイケル・テーラー: 協力の可能性—協力, 国家, アナーキー, 松原 望訳, 木鐸社 (1995). M. Taylor: The Possibility of Cooperation. Cambridge Univ. Press (1987).

[40] マット・カートミル: 人はなぜ殺すか—狩猟仮説と動物観の文明史 内田亮子訳 新曜社 (1995). Matt Cartmill, A View to a Death in the Morning–Hunting and Nature through History. (1993).

[41] マット・リドレー: 徳の起源—他人をおもいやる遺伝子—, 岸 由二監修 翔泳社 (2000), Mait Rindley, The origins of virtue, (1996)

[42] 松原 望: 息決定の基礎, シリーズ意思決定の科学 1, 朝倉書店 (2001).

[43] J. メイナード・スミス: 進化とゲーム理論—闘争の論理—, 寺本英, 他訳, 産業図書, (1985).

[44] 山岸俊男: 社会的ジレンマのしくみ, セレクション社会心理学15, サイエンス社 (1990).

[45] J. ウェイブル: 進化ゲームの理論 大和瀬達二監訳 オフィス カノウチ (1998) Weibull, J. W.: Evolutionary Game Theory, MIT Press (1995).

[46] Fudenberg, D. & Levine, D. K.: The theory of learning in games. MIT press (1998).

[47] Fudenberg, D. & Tirole, J.: Game Theory. (1991) MIT Press.

[48] Greenberg, J.: The Theory of Social Situations. An Alternative Game-Theoretic Approach. Cambridge Univ. Press (1990).

[49] Hardin, R.: Collective Action. Resources for the Future, Inc. (1982).

[50] Luce, R. D. & Raiffa, H.: Games and Decisions. Introduction and Critical Survey. Wiley (1957).

[51] Maitra, A. P. & Suddlerth, W. D.: Discrete Gambling and Stochastic Games. Applications of Mathematics, vol. 32 (1996). Springer.

[52] Myerson, R. B.: Game Theory: Analysis of Conflict. Harvard University Press. (1991)

[53] Petrosjan, L. A. and Zenkevich, N. A.: Game Theory. World Scientific, Series on Optimization vol. 3 (1996). Springer.

[54] Schelling, T. C.: The Strategy of Conflict. Harvard Univ. Press (1960).

[55] Roth, A. Sotomayor, M. A. O.: Two-sided matching. A study in game-theoretic modeling and analysis. (1990). Cambridge Univ. Press.

[56] Young, H. P.: Individual Strategy and Social Structure. An Evolutionary Theory of Institutions. (1998). Princeton Univ. Press.

以上はゲーム理論に関する専門書ないし入門書, 解説書あるいはゲーム理論的視点から書かれた教養書の類である. ゲーム理論に関しては, 動物行動学 (行動生態学) , 進化生物学, 経済学, 心理学, 認知科学, 数学関係の雑誌に多くの文献がある. 以下の文献は, 論文を読んでいて引用されている文献を京都大学の中央図書館, 理学部, 農学部, 経済学部, 経済研究所, 教育学部で実際入手して手元にある論文をリストアップしたものである. ゲーム理論に関する論文は文字通り山のようにある. 分野が違うと原理的には (数学的には) 同じことをしているのかもしれないが感覚が異なり, 互いに知らない, という可能性もあるように感じられる. 引用文献を遡っていってもなかなか収束しない, という印象を持った. 当該分野での応用上は意味があるが理論としては思いつきに類する定式化である可能性も否定できないように思われる.

動物行動学 (行動生態学) , 進化生物学関係

[57] Axelrod, R. & Hamilton, W. D.: The evolution of cooperation. Science, vol. 211 (1981), 1390-1396.

[58] Bishop, D. T. & Cannings, C.: The war of attrition with random rewards. J. Theor. Biol. vol. 74 (1978), 377-388.

[59] Bishop, D. T. & Cannings, C.: A generalized war of attrition. J. Theor. Biol. vol. 70 (1978), 85-124.

[60] Boyd, R. & Lorberbaum, J. P.: No pure strategy is evolutionarily stable in the repeated Prisoner's Dilemma game. Nature Vol. 327 (1987), 58-59.

[61] Boyd, R. & Richerson, P. J.: Group selection among alternative evolutionarily stable strategies. J. Theor. Biol. vol. 145 (1990), 331-342.

[62] Crawford, V. P.: Nash equilibrium and evolutionary stability in large and finite, population "playing the field" models. J. Theor. Biol. vol. 145 (1990), 83-94.

[63] Brown, J. S.: A theory for the evolutionary game. Theor. Popul. Biol. vol. 31 (1987), 140-166.

[64] Cressman, R.: Evolutionarily stable sets in symmetric extensive two-person games. Math. Biosciences. vol. 108 (1992), 179-201.

[65] Cressman, R.: Strong stability and density-dependent evolutionarily stable , strategies. J. Theor. Biol. vol. 145 (1990), 319-330.

[66] Curry, G. L. & Feldman, R. M.: Foundations of stochastic development. J. Theor. Biol. vol. 74 (1978), 397-410.

[67] Eshel, I.: On the handicap principle - A critical defence. J. Theor. Biol. vol. 70 (1978), 245-250. Letters to the Editor,

[68] Feldman, M. W. and Thomas, E. A. C.: Behavior-dependent contexts for repeated plays of the prisoner's dilemma II: Dynamical aspects of the evolution of cooperation. J. Theor. Biol. vol. 128 (1987), 297-315.

[69] Foster, D.: Stochastic evolutionary game dynamics. Theoretical Population Biology, 1990, vol. 38, 219-232.

[70] Harley, C. B.: Learning the evolutionarily stable strategy. J. Theor. Biol. vol. 89 (1981), 611-633.

[71] Hines, W. G. S. & Maynard Smith, J.: Games between relatives. J. Theor. Biol. vol. 79 (1979), 19-30.

[72] Hines, W. G. S.: Evolutionary stable strategies: A review of basic theory. Theor. Popul. Biol. vol. 31 (1987), 195-272.

[73] Lombardo, M. P.: Mutual restraint in tree swallows: A test of the TIT FOR TAT model of, reciprocity. Science, vol. 227 (1985), 1363-1365.

[74] Losert, V. & Akin, E.: Dynamics of games and genes: Discrete versus continuous time. J. Mah. Biol. vol. 17 (1983), 241-251.

[75] Macnair, M. R.: An ESS for the sex ratio in animals, with particular reference to the, social hymenoptera. J. Theor. Biol. vol. 70 (1978), 449-459.

[76] Matsuno, K.: Evolution of dissipative system: A theoretical basis of Margalef's, principle on ecosystem. J. Theor. Biol. vol. 70 (1978), 23-31.

[77] May, R. M.: The evolution of cooperation. Nature vol. 292 (1981), 291-292.

[78] May, R. M.: More evolution of cooperation. Nature vol. 327 (1987), 15-17.

[79] Maynard Smith, J.: The handicap principle - A comment. J. Theor. Biol. vol. 70 (1978), 251-252.

[80] Maynard Smith, J.: The theory of games and the evolution of animal conflicts. J. Theor. Bilol. vol. 47 (1974), 209-221.

[81] Milinski, M.: TIT FOR TAT in sticklebacks and the evolution of cooperation. Nature, vol. 325 (1987), 433-435.

[82] Mowbray, M.: Evolutionarily stable strategy distributions for the repeated Prisoner's Dilemma. J. Theor. Bilol. vol. 187 (1997), 223-229.

[83] Nishimura, K. & Stephens, D. W.: Iterated Prisoner's Dilemma: Pay-off variance. J. Theor. Bio6l. vol. 188 (1997), 1-10.

[84] Nowak, M.: Stochastic strategies in the prisoner's dilemma. Theor. Population Biology, vol. 38 (1990), 93-112.

[85] Riley, J. G.: Evolutionary equilibrium strategies. J. Theor. Biol. vol. 76 (1979), 109-123.

[86] Selten, R.: A note on evolutionarily stable strategies in asymmetric animal conflicts. J. Theor. Biol. vol. 84 (1980), 93-101.

[87] Taylor, P. D. & Jonker, L. B.: Evolutionarily stable strategies and game dynamics. Math. Biosci. vol. 40 (1978), 1 45-156.

[88] Thomas, B.: On evolutionarily stable sets. J. Math. Biol. vol. 22 (1985), 105-115.

[89] Vickers, G. T.: On the definition of an evolutionarily stable strategy. J. Theor. Biol. vol. 129 (1987), 349-353.

[90] Vickers, G. T. & Cannings, C.: Patterns of ESS's I. J. Theor. Biol. vol. 132 (1988), 387-408.

[91] Vickers, G. T. & Cannings, C.: Patterns of ESS's II. J. Theor. Biol. vol. 132 (1988), 409-420.

[92] Zeeman, E. C.: Dynamics of the evolution of animal conflicts. J. Theor. Biol. vol. 89 (1981), 249-270.

経済学, 社会学, 心理学, 認知科学関係

[93] 海野道郎: 社会学におけるゲーム理論的アプローチ. 心理学評論, Vol. 32, No. 3 (1989), 296-311.

[94] 奥野 (藤原) 正寛, 松井彰彦: 文化の接触と進化. 経済研究. Vol. 46, No. 2 (1995), 97-114.

[95] 清成透子, 山岸俊男: コミットメント形成による部外者に対する信頼の低下. Japanese Journal of Experimental Social Psychology, vol. 36, No. 1 (1996), 56-67.

[96] 鈴木光男: 多極化時代の同盟関係ー 4 人ゲームによる分析ー. 数理科学 1968 年 1 月号, 12-20.

[97] 鈴木光男: 紛争の論理と日本人. 中央公論 1971 年 8 月号, 294-311.

[98] 鈴木光男: 組織と分配のゲーム理論. 土屋守章・富永健一編「企業行動とコンフリクトー企業行動コンフリクト報告 2」, 日本経済新聞社 (1972), 61-88.

[99] 鈴木光男, 中邨健二郎: 社会的意志決定と coalition power. 季刊 理論経済学, Vol. XXIII, No. 3 (1972), 3-14.

[100] 鈴木光男: 非協力ゲーム理論の性格. 数理科学 1983 年 11 月号 76-82, 12 月号, 77-83.

[101] 鈴木光男: 費用分担ゲームの解. 数理科学 1984 年 10 月号, 63-68.

[102] 鈴木光男: ゲーム理論への招待. 経済学セミナー. 1986 年 4 月号, 112-116, 5 月号, 97-101, 6 月号, 110-114, 7 月号, 104-109, 10 月号, 121-126, 11 月号, 121-126, 12 月号, 138-142, 1987 年 2 月号, 102-108, 3 月号, 113-118.

[103] 高橋伸幸, 山岸俊男: 利他的行動の社会関係的基盤. Japanese Journal of Experimental Social Psychology, vol. 36, No. 1 (1996), 1-11.

[104] 戸田正直: 茸喰いロボットの実験. 数理科学 1968 年 1 月号, 22-29.

[105] 永田素彦, 矢守克也: コンフリクト状況のマクロ構造分析. Japanese Journal of Experimental Social Psychology, vol. 35, No. 2 (1995), 164-177.

[106] 中邨健二郎: 一般協力 3 人ゲームの解. 数理科学 1971 年 1 月号 32-40.

[107] 林 直保子: 繰り返しのない囚人のジレンマの解決と信頼感の役割, Japanese Journal of Psychology, vol. 66, No. 3 (1995), 184-190.

[108] 森 久美子: 囚人のジレンマにおける社会的価値志向性と利得構造認知. Japanese Journal of Experimental Social Psychology, vol. 38, No. 1 (1998), 48-62.

[109] 山岸俊男: 社会的ジレンマ研究の主要な理論的アプローチ. 心理学評論, Vol. 32, No. 3 (1989), 262-294.

[110] 渡部 幹, 林 直保子, 寺井 滋, 山岸俊男: 互酬性の期待にもとづく 1 回限りの囚人のジレンマにおける協力行動. Japanese Journal of Experimental Social Psychology, vol. 36, No. 2 (1996), 183-196.

[111] 山岸俊男, 山岸みどり, 高橋伸幸, 林 直保子, 渡部 幹: 信頼とコミットメント形成— 実験研究— Japanese Journal of Experimental Social Psychology, vol. 35, No. 1 (1995), 23-34.

[112] 渡部 幹, 林 直保子, 寺井 滋, 山岸俊男: 互酬性の期待にもとづく 1 回限りの囚人のジレンマにおける協力行動. Japanese Journal of Experimental Social Psychology, vol. 36, No. 2 (1996), 183-196.

[113] 渡邊席子, 山岸俊男: “フォールス・コンセンサス”が “フォースル (誤り)”でなくなるときー 1 回限り囚人のジレンマを用いた実験研究ー, Japanese Journal of Psychology, vol. 67, No. 6 (1997), 421-428.

[114] 渡辺昭夫: 沖縄返還ー変貌する日米同盟. 中央公論 1971 年 8 月特大号, 100-112.

[115] Aumann, R. J.: Subjectivity and correlation in randomized strategies. Journal of Mathematical Economics vol. 1 (1974), 67-96.

[116] Aumann, R. J.: Survey of repeated games. Essays in Games Theory and Mathematical Economics in Honor of Oskar Morgenstern. (1981), pp. 11-42.

[117] Aumann, R. J.: Correlated equilibrium as an expression of Bayesian rationality. Econometrica, vol. 55, no. 1 (1987), 1-18.

[118] Axelrod, R.: The emergence of cooperation among egoists. The American Political Science Review, Vol. 75 (1981), 306-318.

[119] Banks, J. S. & Sobel, J.: Equilibrium selection in signaling games. Econometrica, vol. 55, No. 3 (1987), 647-661.

[120] Binmore, K. G. & Samuelson, L.: Drift. European Economic Rev. vol. 38 (1994), 859-867.

[121] Binmore, K. G.: Evolutionary stability in repeated games played by finite automata. J. Economic Theory, 1992, vol. 57, 278-305.

[122] Blume, A. Kim, Y-G. & Sobel, J.: Evolutionary stability in games of communication. Games and Economic Behavior, vol. 5 (1993), 547-575.

[123] Bomze, I. M.: Non-cooperative two-person games in biology: A classification. International J. Game Theory, vol. 15, issue 1 (1986), 31-57.

[124] Boylan, R. T.: Laws of large numbers for dynamical systems with randomly matched, individuals. J. Economic Theory (1992), vol. 57, 473-504.

[125] Brams, S. J. & Wittman, D.: Nonmyopic Equilibria in 2×2 games. Conflict Management and Peace Science, vol. 6 (1981), 39-62.

[126] Cabrales, A. & Sobel, J.: On the limit points of discrete selection dynamics. J. Economic Theory, vol. 57 (1992), 407-419.

[127] Canning, D.: Average behavior in learning models. J. Economic Theory, vol. 57 (1992), 442-472.

[128] Damme, E. van.: Stable equilibria and forward induction. J. Economic Theory, vol. 48 (1989), 476-496.

[129] Damme, E. V.: Evolutionary game theory. European Economic Rev. vol. 38 (1994), 847-858.

[130] Delel, E. & Scotchmer, S.: On the evolution of optimizing behavior. J. Economic Theory, vol. 57 (1992), 392-406.

[131] Dov Monderer: Fictitious play property for games with identical interests. J. of Economic Theory, vol. 68 (1996), 258-265.

[132] Farrell, J.: Meaning and credibility in Cheap-talk games. Games and Economic Behavior, vol. 5 (1993), 514-531.

[133] Friedman, D.: Evolutionary games in economics. Econometrica, vol. 59, no. 3 (1991), 637-666.

[134] Fudenberg, D. & Keeps, D. M.: Learning mixed equilibria. Games and Economic Behavior, vol. 5 (1993), 320-367.

[135] Fudenberg, D. & Maskin, E.: The folk theorem in repeated games with discounting or with incomplete information. Econometrica, vol. 54 (1986), 533-554.

[136] Fudenberg, D. & Levine, D. K.: Consistency and cautious fictitious play. J. of Economic Dynamics and Control. vol. 19 (1995), 1065-1089.

[137] Gilboa, I. & Matsui, A.: Social stability and equilibrium. Econometrica, vol. 59, no. 3 (1991), 859-867.

[138] Grandmont, J-M. & Hildenbrand, W.: Stochastic processes of temporary equilibria. J. Mathematical Economics 1 (1974), 247-277.

[139] Grote, J.: A global theory of games I. J. Mathematical Economics 1 (1974), 223-235.

[140] Hardin, G.: The tragedy of the commons. Science, vol. 162 (1968), 1243-1248.

[141] Hsiao, Chih-Ru: Shapley value for multichoice cooperative games, I. Games and Economic Behavior, vol. 5 (1993), 240-256.

[142] Jordan, J. S.: Three problems in learning mixed-strategy Nash equilibria. Games and Economic Behavior, vol. 5 (1993), 368-386.

[143] Kandori, M. Mailath, G. & Rob, R.: Learning, mutation and long run equilibria in games. Econometrica, vol. 61, no. 1 (1993), 29-56.

[144] Kohlberg, E. & Mertens, J-F.: On the strategic stability of equilibria. Econometrica, vol. 54, no. 5 (1986), 1003-1037.

[145] Kreps, D. M. & Wilson, R.: Sequential equilibria. Econometrica, vol. 50, no. 4 (1982), 863-894.

[146] Macy, M. W.: Learning theory and the logic of critical mass. American Sociological Review. vol. 55 (1990), 809-826.

[147] Macy, M. W.: Chains of cooperation: Threshold effects in collective action. American Sociological Review. vol. 56 (1991), 730-747.

[148] Macy, M. W.: Learning to cooperate: Stochastic and tacit collusion in social exchange. American J. of Sociology, vol. 97, No. 3 (1991), 808-843.

[149] Macy, M. W.: Backward-looking social control. American Sociological Review. vol. 58 (1993), 819-836.

[150] Mailath, G. J.: Introduction: Symposium on evolutionary game theory. J. Economic Theory, vol. 57 (1992), 259-277.

[151] Maruta, T.: A note on stochastic equilibrium selection. 大阪府立大学紀要 vol. 42 (1998), 65-81.

[152] Matsui, A.: Cheap-talk and cooperation in a society. J. Economic Theory, vol. 54 (1991), 245-258.

[153] Matsui, A.: Best response dynamics and socially stable strategies. J. Economic Theory, vol. 57 (1992), 343-362.

[154] McClintock, C. G. , Nuttin, Jr. J. M. & McNeel, S. P.: Sociometric choice, visual presence, and game-playing behavior. Behavioral Science, vol. 15 (1970), 124-131.

[155] McKelvey, R. D. & Palfrey, T. R.: An experimental study of the centipede game. Econometrica, vol. 60, no. 4 (1992), 803-836.

[156] Miyasawa, K.: On the convergence of the learning process in a 2×2 non-zero-sum two person game. Economic Research Program Research Memorandom No. 33 (1961). Princeton University Economic Research Program.

[157] Muliere, P.: A note on stochastic dominance and inequality measures. J. Economic Theory, vol. 49 (1989), 314-323.

[158] Neuefeind, W.: A stochastic bargaining process for n-person games. J. Mathematical Economics 1 (1974), 175-191.

[159] Samuelson, L.: Limit evolutionarily stable strategies in two-player, normal form games. Games and Economic Behavior, vol. 3 (1991), 110-128.

[160] Samuelson, L.: Evolutionary stability in asymmetric games. J. Economic Theory, vol. 57 (1992), 363-391.

[161] Samuelson, P. A.: St. Petersburg paradoxes: Defanged, dissected, and historically described. J. Economic Literature, vol. 15 (1977), 24-55.

[162] Selten, R.: Evolution, learning and economic behavior. Games and Economic Behavior, vol. 3 (1991), 3-24.

[163] Shubik, M.: Game theory, behavior, and the paradox of the Prisoner's Dilemma: three solutions. J. Conflict Resolution, vol. 14 (1970), 181-194.

[164] Sloth, B.: The theory of voting and equilibria in noncooperative games. Games and Economic Behavior, vol. 5 (1993), 152-169.

[165] Stahl, D. O.: Evolution of smart$_n$ players. Games and Economic Behavior, vol. 5 (1993), 604-617.

[166] Svensson, L-G.: Nash implementation of competitive equilibria in a model with , indivisible goods. Econometrica, vol. 59, no. 3 (1991), 869-877.

[167] Swinkels, J. M.: Evolutionary stability with equilibrium entrants. J. Economic Theory, vol. 57 (1992), 306-332.

[168] Swinkels, J. M.: Evolution and strategic stability: From Maynard Smith to Kohlberg and, Mertens. J. Economic Theory, vol. 57 (1992), 333-342.

[169] Tan, T. C-C.: The Bayesian foundations of solution concepts of games. J. Economic Theory, vol. 45 (1988), 370-391.

[170] Wärneryd, K.: Cheap talk, coordination, and evolutionary stability. Games and Economic Behavior, vol. 5 (1993), 532-546.

[171] Weibull, J. W.: The 'as if' approach to game theory: Three positive results and four, obstacles. European Economic Rev. vol. 38 (1994), 868-881.

[172] Weil, P.: Confidence and the real value of money in an overlapping generations economy. The Quarterly J. of Economics. vol. CII, Issue 1 (1987), 1-22.

[173] Young, H. P. The evolution of conventions. Econometrica, vol. 61, No. 1 (1993), 57-84.

数学関係: 最初の 4 冊はノイマン・モルゲンシュテルン (1944) の直後, 数学の分野で研究されたゲーム理論に関する基本的結果が収められている (表紙が赤っぽいからか赤本と称されているらしい.).

[174] Contributions to the theory of games. vol. I: Eds. Kuhn, H. W. & Tucker, A. W. Annals of Mathematics Studies No. 24. (1950), Princeton Univ. Press.

[175] Contributions to the theory of games. vol. II: Eds. Kuhn, H. W. & Tucker, A. W. Annals of Mathematics Studies No. 28. (1953), Princeton Univ. Press.

[176] Contributions to the theory of games. vol. III: Eds. Dresher, M. , Tucker, A. W. & Wolfe, P. Annals of Mathematics Studies No. 39. (1957), Princeton Univ. Press.

[177] Contributions to the theory of games. vol. IV: Eds. Tucker, A. W. & Luce, R. D. Annals of Mathematics Studies No. 40. (1959), Princeton Univ. Press.

[178] Advances in game theory. Eds. Dresher, M. Sharpley, L. S. & Tucker, A. W. Annals of Mathematics Studies No. 52. (1964), Princeton Univ. Press.

[179] Akin, E.: Exponential families andgame dynamics. Can. J. Math. vol. 34 (1982), no. 2, 374-405.

[180] Andre, C.: Two-person stopping games. Advances in Appl. Prob. vol. 7 (1975), 244-245.

[181] Bishop, D. T. & Cannings, C.: Models of animal conflict. Advances in Applied Prob. vol. 8, no. 4 (1976), 616-621.

[182] Blackwell, D. & Ferguson, T. S.: The big match. Ann. Math. Stat. vol. 39, no. 1 (1968), 159-163.

[183] Blad, M. C.: A dynamic analysis of the repeated prisoner's dilemma game. International J. Game Theory, vol. 15, Iss. 2 (1986), 83-99.

[184] Boyd, R. & Lorberbaum, J. P.: No pure strategy is evolutionarily stable in the repeated prisoner's, dilemma game. Nature, vol. 327 (1987), 58-59.

[185] Dubins, L, Maitra, A. , Purves, R. & Sudderth, W.: Measurable, nonleavable gambling problems. Israel J. Mathematics, Vol. 67, No. 3 (1989), 257-271.

[186] Federgruen, A.: On N-person stochastic games with denumerable state space. Adv. Appl. Prob. 10 (1978), 452-471.

[187] Gillette, D.: Stochastic games with zero stop probabilities. Annals of Mathematics Studies. vol. 39 (1957), 179-187.

[188] Haigh, J.: Game theory and evolution. Advances in Appl. Prob. vol. 7 (1975), 8-11.

[189] Harsanyi, J. C.: Games with randomly disturbed payoffs: A new rationale for mixed-strategy equilibrium points. International J. of Game Theory. vol. 2 (1973), 1-23.

[190] Hart, S. & Schmeidler, D.: Existence of correlated equilibria. Mathematics of Operations Reserach. Vol. 14, No. 1 (1989), 18-25.

[191] Hines, W. G. S.: Strategy stability in complex populations. J. Appl. Prob. vol. 17 (1980), 600-610.

[192] Hines, W. G. S.: Strategy stability in complex randomly mating diploid populations. J. Appl. Prob. vol. 19 (1982), 653-659.

[193] Hines, W. G. S.: Mutations, perturbations and evolutionarily stable strategies. J. Appl. Prob. vol. 19 (1982), 204-209.

[194] Kakutani, S.: A generalization of Brouwer's fixed point theorem. Duke Math. J. vol. 8 (1941), 457-459.

[195] Kamerud, D. B.: Repeated matrix games. J. Appl. Prob. vol. 16 (1979), 830-842.

[196] Kohlberg, E.: Repeated games with absorbing states. Ann. Stat. vol. 2 (1974), no. 4, 724-738.

[197] Kôno, N.: Nash equilibria of randomly stopped repeated Prisoner's Dilemma. ICM2002GTA Proceedings Volume. Eds. H. Gao and al. Qingdao Pub. House, Qingdao, (2002), 363-367.

[198] Kuhn, H. W.: Extensive games and the problem of information. Ann. Math. Studies, vol. 28 (1953), 193-216. Contributions to the Theory of, Games I.

[199] Lakshmivarahan, S. & Narendra, K. S.: Learning algorithmes for two-person zero-sum stochastic games with incomplete information: A unified approach. SIAM J. Control and Optimization, vol. 20, No. 4 (1982), 541-552.

[200] Lehrer, E.: Internal correlation in repeated games. International J. of Game Theory, vol. 19 (1991), 431-456.

[201] Maitra, A. & Parthasarathy, T.: On stochastic games. J. of Optimization Theory and Applications. vol. 5, No. 4 (1970), 289-300.

[202] Maitra, A. , Purves, R. & Sudderth, W.: Leavable gambling problems with unbounded utilities. Transactions of the American Math. Soc. vol. 320, No. 2 (1990), 543-567.

[203] Maitra, A. & Sudderth, W.: Borel stochastic games with limsup payoff. Ann. of Prob. vol. 21, No. 2 (1993), 861-885.

[204] Melolidakis, C.: On stochastic games with lack of information on one side. International J. of Game Theory. vol. 18 (1989), 1-29.

[205] Mills, H. D.: Marginal values of matrix games and linear programs. Ann. Math. Studies, vol. 38 (1956), 183-193. Contributions to the Theory of, Games II.

[206] Nachbar, J. H.: "Evolutionary" selection dynamics in games: Convergence and limit properties. International J. of Game Theory, vol. 19 (1990), 59-89.

[207] Nash, J. F. Jr.: Equilibrium points in N-person games. Proc. N. A. S. vol. 36 (1950), 48-49.

[208] Nash, J.: Non-cooperative games, Ann. Math. vol. 54 (1951), no. 2, 286-295.

[209] Neumann, J. v.: Zur Theorie der Gesellschaftsspiele. Mathematische Annalen, band 100 (1928), 295-320. (On the theory of games of strategy, translated by S. Bargmann, Contributions of the, Theory of Games, vol. IV, Ann. Math. Studies, vol. 40 (1959), 13-42)

[210] Ohtsubo, Y.: A nonzero-sum extension of Dynkin's stopping problem. Mathematics of Operations Research, vol. 12, No. 2 (1987), 277-296.

[211] Owen, G. & Shapley, L. S.: Optimal location of candidates in ideological space. International J. of Game Theory, vol. 18 (1989), 339-356.

[212] Parthasarathy, T.: Discounted, positive, and noncooperative stochastic games. International J. of Game Theory, vol. 2 (1973), 25-37.

[213] Parthasarathy, T. & Sinha, S.: Existence of stationary equilibrium strategies in non-zero sum discounted stochastic games with uncountable state space and state-independent transitions. International J. of Game Theory, vol. 18 (1989), 189-194.

[214] Ravindran, G. & Szajowski, K.: Non-zero sum game with priority as Dynkin's game. Math. Japonica vol. 37, No. 3 (1992), 401-413.

[215] Robinson, J.: An iterative method of solving a game. Ann. Math. vol. 54, no. 2 (1951), 296-301.

[216] Savage, L. J.: The foundations of statistics. Wiley publications in statistics . Mathematical statistics (1954).

[217] Shapley, L. S.: Stochastic games. Proc. Nat. Acad. Sci. U. S. A. vol. 39 (1953), 1095-1100.

[218] Sennott, L. I.: Nonzero-sum stochastic games with unbounded coasts: Discounted and average cost cases. ZOR-Mathematical Methods of Operations Research, vol. 40, Issue 2 (1994), 145-162.

[219] Sober, M. J.: Noncooperative stochastic games. Annals of Mathematical Statistics, vol. 42, No. 6 (1971), 1930-1935.

[220] Sorin, S.: Some results on the existence of Nash equilibria for non-zero sum games with incomplete information. International J. Game Theory, vol. 12 (1983), Iss. 4, 193-205.

[221] Sorin, S.: Asymptotic properties of a non-zero sum stochastic game. International J. Game Theory, vol. 15 (1986), Iss. 2, 101-107.

[222] Suzuki, A. & Muto, S.: Farsighted stability in prisoner's dilemma. J. of the Operations Research Society of Japan. vol. 43, no. 2 (2000), 249-265.

[223] Taylor, P. D.: Evolutionarily stable strategies with two types of player. J. Appl. Prob. vol. 16 (1979), 76-83.

[224] Zamir, S.: On the notion of value for games with infinitely many stages. Ann. Stat. vol. 1 (1973), no. 4, 791-796.

[225] 「確率ゲーム理論とその周辺」: 数理科学講究録 741. 京都大学数理解析研究所 (1991. 1).