オンライン・コンテンツ4.3: 色々な進化動学と進化的安定状態

図斎 大

本稿は，浅古泰史・図斎大・森谷文利『活かすゲーム理論』有斐閣の第4章までを読んだ読者を対象としたオンライン・コンテンツです。

本書では進化動学の代表として最適反応動学を取り扱いました。均衡を前提とせずに，進化動学のように調整過程から社会を分析する分野を進化ゲーム理論と言います。日本語に限らず入門レベルのゲーム理論全般の教科書で進化ゲーム理論を紹介するときには，これまでは複製子動学(replicator dynamic)という進化動学，あるいは進化的安定状態(evolutionary stable state, ESS)というナッシュ均衡を精緻化した均衡概念が用いられていました。これらは「進化」を考える元となった生物学でのゲーム理論で代表的なものです。このオンライン・コンテンツでは複製子動学や ESS を最適反応動学と繋げながら，それらの生物学的な意味や，進化動学の広さを紹介します。また従来の教え方と接合するよう，本書の例を用いて複製子動学や ESS に関する演習問題も与えます。2.2 節では偏微分と線形代数（負値定符号）を用いますが，それ以外は初等的な計算にとどめています。

1.いろいろな進化動学

本書の第4章 3.2 節では「進化動学」を，人々が実際にゲームをプレーし，その時々の状況を観察しながら学習しながら戦略を変えていく過程を総称したものだと定義しました。そして，無数にプレーヤーがいる「ポピュレーションゲーム」において各プレーヤーにバラバラに選択変更の機会が訪れるという仮定の下で，各時点における状況を観察したうえで，各プレーヤーが自身の選択をどのように変えるのかを設定することで，進化動学のモデルが具体化されるのだと説明しました。

本書で取り上げた「最適反応動学」(best response dynamics, BRD)では，プレーヤーは戦略改訂の機会を得たら，その時点でもっとも利得の大きい戦略へと移るのでした。この背後には，現在の戦略分布に基づく期待利得を正しく把握していること，また利得を改善する機会を見過ごすことがないという前提があります。しかし，プレーヤーが観察できる情報に制限があったり，個人の認識や行動の遂行にズレがあったり，単純な利得に基づく合理的判断を躊躇するということもありえます。そのような様々な意思決定のあり方の下で，ゲームのプレーがどのように動いていくのか，そして果たして（ナッシュ均衡のような通常の）均衡概念が適応的な学習の末に達成されるかといったことも，進化動学の枠組みで統一的に考えられます。

以下で最適反応動学だけでない他の進化動学で考えられている意思決定の仕方を見てみます。その前の準備として一言。いろいろな意思決定を考えると，戦略改訂の機会を得て，何らかの移るべき（利得がより高い）戦略の候補が念頭にあったとしても，必ずしもそれに移らないということもありえます。つまり，その候補に移るのが確率的である，と。この移りやすさを**スイッチング・レート(switching rate)**という数値で表します。たとえばレートが0.8といったら，長さt の時間の間に確率

$0.8t$でその候補へと移ること，レート1.5なら確率$1.5t$で移ることを意味します。（この時間の長さtが十分短いという前提で，このレート自体は1を超えていてもよいです。）プレーヤーが得ている情報に応じてスイッチング・レートをどう設定するかということに，とどのつまりは意思決定の仕方というのはまとめられます[1]。

1.1. 色々な進化動学の例

緩和された最適反応：戦略変更を躊躇するとき

通常の最適反応動学と同様に，プレーヤーはきちんと現時点の最適反応が正しくわかっている状況を考えます。それでも，もしも最適反応に切り替えたところで利得がそんなに改善しないなら，プレーヤーは面倒くさがって戦略を変更しないかもしれません。まったく変更しないというのも極端なので，利得の改善が小さいほど最適反応へのスイッチングレートが小さくなるというくらいにしておきましょう[2]。このように最適反応へのスイッチング・レートが利得差とともに増える進化動学を**緩和された最適反応動学**(tempered best response dynamic; Zusai, 2018)と呼び，略して **tBRD** と称します。

> **緩和された最適反応動学（tBRD）**
>
> 戦略改訂の機会を得たプレーヤーは
>
> 1. 最適反応：現時点の社会全体での戦略分布を観察し，それに対する最適反応を見つける。
> 2. スイッチ：現在の戦略から最適反応に切り替えることによる利得の改善の幅が大きくなるほど，より高いスイッチングレートで最適反応へと戦略を変える。

[1] 数学的には意思決定の仕方は，プレーヤーの得ている情報から個々の戦略間のスイッチングレートを決める関数にまとめられるということです。Sandholm (2010)は，このように関数で表された意思決定のルールを改訂プロトコル(revision protocol)と呼びました。さりげない点ですが，そこに焦点を置くことで，それまでてんでバラバラだった進化ゲーム理論を統一的に分析することが容易になりました（図斎，2021）。ここでは tBRD と成功の模倣 IoS というふたつだけを紹介しますが，他の代表的な進化動学にも関心があれば，Sandholm (2010)で学ばれるのを強くお勧めします。

[2] これは経済学的には戦略を変えるためのスイッチングコストが確率的に発生するのだと解釈できます。最適反応に変えることによる利得の改善の幅がこのスイッチングコストが小さいなら，戦略改訂の機会をやり過ごします。利得の改善の幅が大きくなるほどスイッチングコストを超えやすくなるので，戦略を実際に改訂しやすくなり，スイッチングレートが高くなるということになります。

このオンライン・コンテンツでは，スイッチングレートは利得の改善の幅，すなわち [最適反応からの利得]−[現在の利得]と等しくなるように決まるとしましょう。つまり，現在の戦略をi，移る先の候補の戦略をjとすると，

$$(\text{tBRD})\begin{cases} j\text{が最適反応なら} & [i \to j\text{のスイッチングレート}] = [j\text{からの利得}] - [i\text{からの利得}]; \\ j\text{が最適反応でないなら} & [i \to j\text{のスイッチングレート}] = 0 \end{cases}$$

というのがここで考える tBRD です。

模倣：どんなゲームをプレイしているのかもわからないとき

現時点での利得や正しい戦略分布はおろか，そもそもどんなゲームをプレイしているのかもわからないという状況を考えましょう。このような状況で，社会の中からランダムにサンプルとなる人を取ってきて，その人たちの戦略が自らの戦略よりも良さそうなら，それを真似するという進化動学全般を**模倣動学**(imitative dynamics; Schlag, 1998)と呼びます。

成功を模倣する動学（Imitation of Success; Hofbauer, 1995）

戦略改訂の機会を得たプレーヤーは

1. サンプリング：まず，社会の中からランダムにサンプルとなる人を一人取り，その人の戦略と利得を観察する。
2. 模倣： この観察した利得が大きいほど高いスイッチングレートで，観察したサンプルの人の戦略へと切り替える。

どれだけ真似をするかというスイッチングレートの設定はいろいろ考えられますが，真似する戦略の利得と等しくするというのが最初に思いつくかもしれません[3]。このオンライン・コンテンツでも基本的にそのようにしたいところなのですが，少し修正をします。というのも，もしも利得が負なら，それに等しくさせたスイッチングレートも負になります。スイッチングレートに時間を掛けてスイッチする確率を得ますが，その確率が負だと意味が通りませんね。なので，利得が負になりえることも鑑みて，利得にゲタを履かせた，つまり定数を足したものでスイッチングレートを設定するというのが，この動学

[3] 成功を模倣する動学（IoS）でのスイッチングレートは元の戦略に依存しません。しかし，逆に元の戦略が低いほど，取り合えず（真似する戦略の利得を気にせず）誰か他の人のに移る（Björnerstedt & Weibull, 1996）とか，tBRD のように利得差に応じてスイッチングレートを決める（Schlag, 1998）とかもありえます。しかしこれらはどれも戦略分布の遷移ということでは，成功の模倣 (IoS)と同じもの（複製子動学）に帰着されます。

での普通の設定です。この定数を負の利得を打ち消すくらい十分大きく設定することで，スイッチングレートが負にならないようにします。この定数は確率としてのスイッチングレートの解釈のために導入していますが，定数（そしてその大きさ）が最終的に我々が見たい戦略分布の遷移には影響しないことを後で確認します。

以上をまとめましょう。現在の戦略をi，移る先の候補の戦略をjとすると，

(IoS) [$i \to j$のスイッチングレート] =[社会の中でjを取る人の割合]×[戦略jの利得+ある定数]

というのがここで考える成功の模倣動学です。実はこのスイッチングレートから，冒頭で生物学で代表的な進化動学として触れた複製子動学が導かれるのを次の 1.2 節でみます。

1.2. 数値例

以上のような意思決定ルールの下で戦略分布がどう変わっていくのかを見るために，ちょっと計算をしてみましょう。そのために本書第4章でのモデル 4.1「PoyPoy 入れてる?」で，特に$(p,q) = (0.5,0.1)$という状況を考えます。3.3 節を振り返ると，最適反応の(p,q)は$(0,1)$なので通常の最適反応動学なら$(0.5,0.1)$から$(0,1)$へと真っすぐ向かっていきます。本書の図 4-3 でもそのような矢印，つまりベクトルが描かれていますね。このように動学が向かっていく方向を表すベクトルを遷移ベクトル(transition vector)と呼びます。数学的には(p,q)が時間と共に変わっていくということは，(p,q)が時間tの関数であるということです。そして，この向かう方向というのは(p,q)の時間tに対する微分$(dp/dt, dq/dt)$で決まります。慣習的には時間微分を，ドットを元の変数の記号の上につけることで表します。なので，遷移ベクトル$(dp/dt, dq/dt)$を$(\dot{p}, \dot{q})$と表します。つまり，通常の最適反応動学での$(0.5,0.1)$からの遷移ベクトルは

$$\begin{pmatrix}\dot{p}\\ \dot{q}\end{pmatrix} = \underset{\text{最適反応}}{\begin{pmatrix}0\\ 1\end{pmatrix}} - \underset{\text{現在の状態}}{\begin{pmatrix}0.5\\ 0.1\end{pmatrix}} = \begin{pmatrix}-0.5\\ 0.9\end{pmatrix}$$

と書けます。他の動学ではどうでしょう?

tBRD

tBRD では，最適反応へのスイッチングレートが利得の改善幅に比例します。この意思決定ルールのためには，何が最適反応かだけでなく，それがもたらす利得がいくらかという情報も必要です。なので，$(p,q) = (0.5,0.1)$の下での，各プレーヤーの各戦略の利得を以下のように計算しておきます。

「使う」の期待利得　　否の期待利得

お客さん: $0.1 \times 3 + 0.9 \times (-1) = -0.6$ $0.1 \times 0 + 0.9 \times 1 = 0.9$

お店: $0.5 \times 3 + 0.5 \times (-1) = 1.$ $0.5 \times 0 + 0.5 \times 1 = 0.5.$

お客さんたちのなかで$1 - p = 0.5$の人たちは既に最適反応である「否」を選んでいるので，戦略を変えません。残る$p = 0.5$の人たちが「使う」を取っています。†BRD では必ずしも最適反応である「否」に移るわけではありません。「使う」から「否」に変えることによる利得の改善幅は$0.9 - (-0.6) = 1.5$なので，(†BRD) 式に従うと1.5がスイッチングレートになります。つまり，この現状では，時間×1.5の確率で，残された$p = 0.5$の人々が徐々に「使う」から「否」に移っていき，その分だけお客さんの中で「使う」割合pが減っていきます。すなわち，

$$\dot{p} = -0.5 \times (0.9 - (-0.6)) = -0.5 \times 1.5 = -0.75$$

となります。

お店の方では「使う」が最適反応なので，$q = 0.1$のお店は戦略を変えません。他方で$1 - q = 0.9$のお店は「否」を取っており，「使う」に変えることで利得が$1 - 0.5 = 0.5$だけ改善します。従って，スイッチングレート1で，この$1 - q = 0.9$の人々が「否」から「使う」へと移っていき，その分だけお客さんの中で「使う」割合pが増えていきます。すなわち，

$$\dot{q} = 0.9 \times (1 - 0.5) = 0.9 \times 0.5 = 0.45$$

となります。従って，†BRD の下での遷移ベクトルは$\begin{pmatrix}\dot{p}\\ \dot{q}\end{pmatrix} = \begin{pmatrix}-0.75\\ 0.45\end{pmatrix}$となります。$p$が減るのに対して$q$がどれだけ増えるのかを，$\dot{q}$の$\dot{p}$に対する比$\dot{q}/\dot{p}$として求めると，$0.45/(-0.75) = -0.6$となります。つまり，$p$が減るのの 0.6 倍のスピードで$q$が（負，つまり減るのと反対に）増えていきます。標準の最適反応動学だとこの比は$0.9/(-0.5) = -1.8$なので，q が増えるスピードはpが減るのの 1.8 倍です。†BRD だと比が小さくなっていますが，それはお店側の方で最適反応へ切り替えるにことによる利得の改善幅が小さいこと(お店 0.5:お客さん 1.5)を反映しています。

成功の模倣

模倣動学では，サンプルを一人ランダムに取ってきて，確率的にその人の戦略へと移ります。「成功の模倣」(IoS)という設定では，この確率はサンプルの人の利得(+ゲタ)に比例するとしたのでした。ただ，確率が負になると意味が通らない一方で，このゲームでは−1という利得も起こりえるので，以下では，スイッチングレートを[サンプルの人の利得 + 1]としましょう。それとここでは社会の中にお客さんとお店という二つの異なるプレーヤーの役がありますが，サンプルは同じ役から取ってく

るものとします。つまり，お店は他のお店を模倣し，お客さんをサンプルに取らないということです。ちなみに，$(p,q)=(0.5,0.1)$だと，お客さんに関しては「使う」も「否」もシェアが 0.5 ずつで，0.5 という数字を見た時に，どちらのシェアの意味か分かりづらくなってしまうので，以下ではお店の方に注目します。

お店の中で現状は$q=0.1$の割合で PoyPoy を使っているのですから，確率 0.1 で PoyPoy を使っている店をサンプルすることになります。そしてその利得は1なので，「使う」店をサンプルした後のスイッチングレートは$1+1=2$です。これで実際に戦略を変えている店はどれだけいるでしょう。元々「使う」を選んでいた店は，サンプルした店を模倣すると言ってもこれまでと同じ「使う」なので変わりませんね。なので，新たに「使う」へとスイッチしているのは，これまで「否」を取っていた店たちです。店の中で「否」のシェアは$1-q=0.9$なので，単位時間当たり$0.9\times0.1\times(1+1)=0.18$の確率で，「否」から「使う」へと切り替える店が出ます。

他方で，「否」を選んでいるお店をサンプルする確率は「否」のシェアだけの$1-q=0.9$です。そして「否」の利得は0.5なので，このサンプルでのスイッチングレートは$0.5+1=1.5$です。そして，新たに「否」を取ることになるのはこれまで「使う」を取ってきた店で，そのシェアは$q=0.1$なので，単位時間当たり$0.1\times0.9\times(0.5+1)=0.135$の確率で，「使う」から「否」へと切り替える店が出てきます。つまり，お店の中で「使う」を取る割合は，単位時間当たりのレートで差し引き$0.18-0.135=0.045$だけ変わっていきます。つまり，

	いま自分が「使う」	「否」をサンプル	「否」へスイッチする		いま自分が「否」	「使う」をサンプル	「使う」へスイッチする
$\dot{q}=$	0.9	× 0.1	×(1+1)	−	0.1	× 0.9	×(0.5+1)

$$=0.18-0.135=0.045$$

となります。お客さんに関しても同様に

$$\dot{p}=0.5\times0.5\times(-0.6+1)-0.5\times0.5\times(0.9+1)=-0.375$$

となります。まとめると，成功の模倣の下での遷移ベクトルは$\begin{pmatrix}\dot{p}\\ \dot{q}\end{pmatrix}=\begin{pmatrix}-0.375\\ 0.045\end{pmatrix}$となります。

ちょっと数式で遊んでみると，$\dot{q}$の式は

$$\dot{q}=0.9\times\{0.1\times(1+1)-0.1\times(0.5+1)\}=0.9\times\{(1-0.9)\times1-0.1\times0.5\}$$

と書き直せます。( ) の中にあったゲタの＋1 が打ち消し合ってますね。更に$q=0.9$だったので，

$$\frac{\dot{q}}{q}=(1-0.9)\times1-0.1\times0.5=1-(0.9\times1+0.1\times0.5)$$

となります。左辺の $\dot{q}/q$は，現在のシェアqに対するシェアの増加$\dot{q}$の比なので，シェアの成長率と言えます。右辺の最初の1は「使う」利得で，後の$0.9 \times 1 + 0.1 \times 0.5$は社会全体での平均のお店の利得になっています。すなわち，右辺は「使う」ことの，社会全体の平均と比較した，相対的な利得と言えます。従って，この式は

「使う」戦略のシェアの成長率 ＝「使う」戦略の相対的な利得

として解釈できます。実は，これは「成功の模倣」に基づく進化動学全般で成り立つ式です。このように，その戦略の相対利得がそのまま各戦略のシェアの成長率となる進化動学を**複製子動学** (replicator dynamic)と言います。

複製子動学は生物学でよくつかわれる動学です。「戦略」を生物の種類だとすると，ゲームで表すのは生存競争の環境，そして進化動学は生物種の淘汰の過程を描いています。ただこれまでだと，どの戦略を取るかというのは，より高い利得を取ろうという個々人の意思決定の結果として現れるのでした。これは経済学・社会科学的な観点を反映しています。他方で，生存競争の中でどの生物種が多く残るかは，個々の選択というよりは，利得が低い，つまり現在の状況に適応していないものが淘汰されることの結果と言えるでしょう。なので，生物学的な解釈においては，「利得」というのは意思決定の指標ではなく，その生物種が結果としてどれだけ多く子孫を残せるかを表していると考え，「適応度」と呼んでいます。この子孫を残すということ自体が，次の世代，つまり直近の将来でどれだけその生物種＝戦略が増えるかということを意味しているので，生物学においては直接，上述の式のように，つまり

各生物種（戦略）のシェアの成長率 ＝その生物種（戦略）の相対的な適応度（利得）

として進化動学の式を立てることが自然になります。

1.3. まとめ

モデル 4.1 について他の点でも同様の計算で求めた遷移ベクトルを各点ごとに図示したのが，以下の図 1 になります。（ただしベクトルの長さは揃えて，代わりに元の遷移ベクトルの長さを色で表しています。黒，青から，遷移ベクトルが長くなるにつれ（つまりスイッチングレートが高くなるにつれ），赤，黄色と変わっています。）

図 1a. 標準の最適反応動学

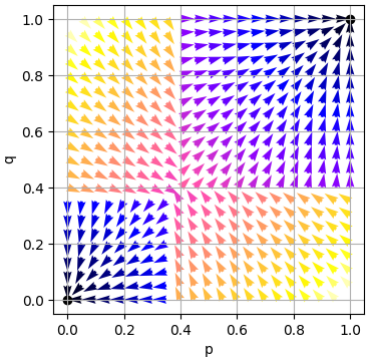


図 1b. tBRD

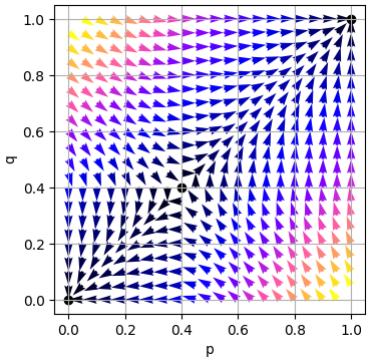
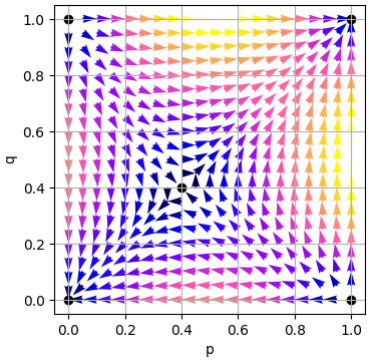


図 1c. 成功の模倣



このような遷移ベクトル$(\dot{p},\dot{q})$を出発点となる初期状態のものから積み重ねていくことで（数学的には積分することで），それぞれの進化動学で，どのように戦略分布が変わっていくかがわかります。といったものの，最適反応動学と違って各点ごとに少しづつ方向が変わるので，各点ごとに遷移ベクトルを計算しないといけません。手計算だと無理そうですね。この図を描くにあたっては，先のような遷移ベクトルの計算をコンピュータにさせています。つまり，遷移ベクトルを一般的に求める式を立てて，その計算をコンピュータにさせるというのがひとつの策です。あるいは，そもそもコンピュータの中で，仮想的にたくさんのプレーヤーが住みゲームをプレイする世界を作る，つまりモデルをシミュレーションさせることも，ビデオゲームを作るような要領でできます。実際，進化動学では上述の通り，各プレーヤーがどのように意思決定するかというのを明示するので，それをプログラミングのコードにすることでシミュレーションができます。本書のオンライン・コンテンツ 4.4 でも，上図を描くために使った Python というプログラミング言語で書いたコードと共に，そのようなシミュレーションのコードをお見せします。そうしたコンピュータシミュレーションによる計算的手法は社会生物学（ブックリストで挙げた中丸，2020 など）や進化人類学（田村，2023）ではポピュラーです。他方で，経済理論では数学的に厳密・精緻に組み立てられたロジックのつながりに関心があるので，数式自体を分析し，面白い性質を見つけて，それを数学的に証明するというのが主流です。このためには上

述のような計算を微分方程式の形にするのですが，それは進化動学の本格的な教科書の決定版と言える Sandholm (2010)に譲ります[4]。

ともかく先の計算や上の図を振り返ると，やはり異なる進化動学ごとに微妙に違うようですね。しかし，どれでも(p,q)がナッシュ均衡である$(0,0),(0.4,0.4),(1,1)$のいずれかちょうどならば，遷移ベクトルが延びていません。実際，tBRD や成功の模倣に関して，上述のような計算をすると $\dot{p}$も $\dot{q}$もゼロになっていることが確かめられます。（演習問題とします。また，このオンライン・コンテンツのおまけとして，上述の進化動学について遷移ベクトルを計算するための Excel ファイルを提供するので，それでもぱっと調べられます。）つまり，ナッシュ均衡は定常状態（上図では黒点）になっています[5]。このことは，tBRD や模倣動学のように実際の利得に従うような進化動学では一般的に成り立ちます。安定性に関してもやはり最適反応動学と同じように$(p,q)=(0,0)$や$(1,1)$が安定で，$(p,q)=(0.4,0.4)$は不安定というのは，tBRD でも成功の模倣の動学でも言えそうです。ここではこの遷移ベクトルをそれぞれで計算して安定性を図から判断していますが，どうせ共通しているのなら，なにか動学に依らず，ゲーム自体から安定性を言えるといいですね。それが次に学ぶ進化的安定のアイデアです。

2. 進化的安定

均衡の安定性を判断するときには，均衡からちょっとズレた状態をまず考えます。進化動学ではそこから均衡へと戻っていくかを，その時々の状態の遷移を積み重ねて見ています。しかし，そこまでしっかり見届けなくても，均衡からのズレが損になるならば，人々はいずれ均衡へと戻っていくでしょう。もちろん人々がインセンティブ，つまり利得を正しく認識できなかったり，それに背くような選択を試みるのなら，そんなに単純にはいかないかもしれません。しかし正しいインセンティブを妨げるものがなければ，きっかり最適反応を取らずともちょっとくらいそれをやる過ごすことがあったり，意思決定の仕方の細かい違いには左右されずに，戻るインセンティブがあるかどうかというゲーム自体の構造から，その均衡へと戻っていくと言えるでしょう。

---

[4]微分方程式にするということは戦略分布が時間の経過の中では連続的にしか変わらないということを含意しています。これを正当化するために無数にプレーヤーがいるということ，そしてそれらはバラバラのランダムなタイミングで戦略改訂を行うという状況の設定が効いてきます。

[5]模倣に基づく進化動学では，どれかの戦略を取る割合がまったくのゼロになってる状態では，その戦略をサンプルでとることはありません。それゆえ，その状態が実はナッシュ均衡でなかったとしても，定常状態になります。図 1.c でも$(p,q)=(0,1),(1,0)$はナッシュ均衡でないのに定常状態（黒点）になっていますね。

あるいは生物学でのアナロジーで言えば，利得というのがその生物種の適応度という意味を持っているのでした。複製子動学ではその適応度がちょうど成長率になるのですが，そんなにぴったりでなくても，今の状況により適応している種がより栄えやすいというのが素直な進化でしょう。ならば，均衡の手前で，ある種が均衡状態でのシェアまでまだ伸びきっていないとしても，その種の適応度が他よりも高ければ，いずれこの均衡状態に行きつくだろうと言えます。このように，均衡からズレていたときに，均衡状態と比べてシェアが小さい戦略こそ利得が他と比べて高くなることをもって，動学をたどらずとも，安定的だと言い切ってしまうのが「進化的安定」という考え方です。

安定性という言葉は本来は，状態がズレたときに元へ戻ることを意味しています。進化動学での安定性は実際にズレからどのように動いていくのかを考えて，本当に元に戻るのかを確認しました。この「進化的安定」ではどのように実際に戻るのかまでは考えません。その分だけ「ゆるい」概念なのですが，それゆえ直感的ですし，また特定の動学に左右されにくいだろうと言えます[6]。なので，この進化的安定による分析もゲーム理論ではよく行われます。

緩いというのは本来の動学的な安定性とのつながりにおいてであって，「進化的安定」という概念自体はきちんと数学的に定義されます。なので，進化的安定による分析をするときには，その定義をきちんと満たしているかを厳密に確かめることになります。ただその定義で案外と細かいところを詰めるために数学的表現が複雑になるので，以下では，均衡でただ一つの戦略がみんなに取られているときと，選択しうる戦略がすべて多かれ少なかれ（ゼロでない）正のシェアで取られているときとを分けて（またそれ以外の状況は捨てて），それぞれで進化的安定を定義します。

2.1. 単一戦略均衡での進化的安定

まずただ一つの戦略だけがみんなに取られている均衡での進化的安定を定義します。このような均衡をここでは単一戦略均衡(monomorphic equilibrium)と呼ぶことにしましょう[7]。ともかく均衡ということは，この均衡ちょうどにあるときには，このみんなが取っている戦略より高い利得を付ける戦略がないということです。この均衡でみんなに取られている戦略を戦略 I としましょう。ここで，均衡からのズレとして，一部のプレーヤーたちがたまたま戦略 I から戦略 2 へと移ったとしましょう。この

[6] この概念自体で言えるのはあくまで「だろう」という予想であって，論理的にはこの「緩い」概念でも十分にいろいろな動学で安定性をもたらすことを厳密に証明する必要はあります。それがこの節の最後に挙げる一連の論文(Hofbauer & Sandholm, 2009 など)が行っていることです。

[7] この節では本書と同じく，各人が純粋戦略を選ぶポピュレーションゲームで，また利得が戦略分布に対する期待利得として定まる（つまり各戦略の利得が戦略分布の関数として線形になる）状況を念頭に置きます。ただ単一戦略均衡での進化的安定性は，この前者の仮定を外しても，つまり各人が混合戦略を選ぶものだとしても同様の形で定義されます。

ズレの後に，戦略2のほうが戦略1よりも利得が低くなってしまったならば，このプレーヤーたちがある程度インセンティブに従って動く限りは戦略 2 から戦略 1 へと戻っていくでしょう。これがここでの進化的安定性です。

安定だというためには，ズレの大きさはしぼっても（つまり局所的でも）よいけれども，どんな方向のズレでも，元の状態に戻らないといけないのでした。なので，このズレた先が戦略2だけでなくて，戦略1の他のどの戦略に関しても言える必要があります。またズレの大きさを明示しておくとわかりやすいので，ϵで表しましょう。戦略1から他のいずれかの戦略へとズレた人のシェアがϵということです。以上を踏まえて，単一戦略均衡からズレた時に元に戻ってくれるであろうインセンティブの条件を書いてみましょう。一般的に表すために，単一戦略均衡で取っていた戦略を1としていたのをiとおきましょう。まず，ϵだけ戦略iから他の戦略へとズレることを考えます。このズレた先の戦略をjで表します。このときに，戦略iの利得の方が戦略jよりも高い利得になるのなら，この戦略jへとズレてしまった人たちも元の戦略iへと戻っていくでしょう。で，安定性のためには，このjというズレた先の戦略はどんなものでも，またこのズレの大きさϵはどんなに大きくてもとは言わなくても，特定の大きさだけでなく「十分に」小さければどんなでも許されないといけません。（もしも中程度のズレで言えて，しかしすごく小さいズレで言えなくなるというのなら，徐々に戻っていくときに，ズレが小さくなっていくのですから，戻る途中でダメになってしまいますね。）この「十分に」小さくというのは，「ある程度」の大きさよりも小さくなればと言い換えられます。ここでは，この「ある程度」を 上に横線を付けた$\bar{\epsilon}$で表します。すると，以上のインセンティブによる安定性の条件，つまり「進化的安定」の条件は以下のように数学的に表せます。

定義：単一戦略均衡の進化的安定性

みんなが戦略iをとっている単一戦略均衡が進化的安定であるとは，
「ある程度」許せるズレの大きさとなる$\bar{\epsilon}$というのが（ゼロでなく）あって，どんな他の戦略$j \neq i$に関しても，ズレの大きさϵが$\bar{\epsilon}$より小さいなら，シェアϵだけのプレーヤーが戦略iから戦略jへと移った状態において，

[ズレる前の戦略iの利得]> [ズレた後の戦略jの利得]　　（★）

となることである[8]。

[8] 次の 2.2 節で導入する記号を使って，もっと数学っぽく表現するなら，$\exists \bar{\epsilon} > 0 \, \forall j \neq i \, [\epsilon < \bar{\epsilon} \Rightarrow \pi_i((1-\epsilon)\boldsymbol{e}^i + \epsilon \boldsymbol{e}^j) > \pi_j((1-\epsilon)\boldsymbol{e}^i + \epsilon \boldsymbol{e}^j)\,]$と，この定義の条件は表されます。ここで$\boldsymbol{e}^i$は戦略$i$のみ1で他は0をつけるベクトルで，$\boldsymbol{e}^j$も同様に戦略$j$のみ1をつけるベクトルとします。従って，$(1-\epsilon)\boldsymbol{e}^i + \epsilon \boldsymbol{e}^j$は$\epsilon$だけ戦略$j$を，$1-\epsilon$は戦略$i$を取っているという戦略分布です。

ちなみに，進化的安定な均衡自体は**進化的安定状態(ESS, evolutionary stable state)**と一般的に呼びますが，特に単一戦略均衡の ESS については，そこで取られている行動（純粋戦略）のことを，**進化的安定戦略**(evolutionary stable strategy)と呼ぶこともあります。

この定義は均衡そのものでは，そこで取られている戦略が他の戦略よりも利得が強に（つまり，等しいことなく）高いなら，つまり（★）が均衡においてきちんと>で満たされるなら，成り立ちます[9]。そのような単一（純粋）戦略均衡を強均衡（strict equilibrium）と呼びます[10]。本書のモデル 4.1 での$(p,q)=(0,0),(1,1)$はそれぞれ強均衡です。強均衡でないならどういうときでしょう？それは均衡で利得が等しくなる，いわば「惜しい」戦略が他にあるという状況です。進化的安定であるとは，この惜しい戦略にはズレが均衡で取られている戦略に比べて不利に働くということです。このような状況は次に考える複数の戦略が混ざる均衡のほうが似ているので，「不利に働く」ということのきちんとした意味は次節で説明します。

2.2. 完全に混ざっている均衡での進化的安定

今度はどの戦略も大きかれ小さかれ正のシェアで取られているという，つまり**完全に混ざっている**(completely mixed)均衡を考えます。まず簡単にプレーヤーが取り得る戦略が1，2の二つしかないとしましょう。そこで我々が考えるのは戦略 1 も 2 もともに正のシェアで取られている均衡です。ここで，ϵだけのシェアにあたるプレーヤーたちがたまたま 1 から 2 へとズレたとします。進化的安定はこのプレーヤーたちが戻るインセンティブがあるということです。

それは本質的には，また★式が（$i=1,j=2$で）成立するということです。しかし戦略1，2も共に均衡で取られていたことから，もう少し条件をわかりやすく，特に「ある程度のズレ」なんてことも言わずに，均衡点そのものでの利得関数の振る舞いから言えます。まず，戦略 1,2 もこの均衡でともに誰かが取っているということは，どちらの戦略も利得が等しくないといけません。しかしズレた後に★式が成り立つということは，元の戦略 1 と比べて移った先の戦略 2 の利得が相対的に下がるということです。これが，ズレが戦略 1 よりも戦略 2 に不利に働くということの意味です。完全に混ざ

[9] ただし，各戦略の利得が戦略分布の関数として連続であるということが前提です。連続性の下では，ある状態（点）で強い不等号>が成立していれば，あまりそこから離れていない（十分に近い）限りは同じ強い不等号が成立します。それゆえ，強均衡で★がその均衡でさえ満たされていれば，その均衡周辺でも★が成立すると言えます。

[10] ある戦略が強均衡で使われているというのはそれが強支配戦略であることより弱い条件です。今日支配戦略と違って，均衡から離れた状態ではこの戦略より利得の高い戦略があるというのも許容されます。

っている均衡が進化的安定であるには，均衡からの人々の選ぶ戦略のズレが，移る元の戦略と比べて移った先の戦略に不利になるようそれぞれの戦略の利得を変化させればよいのです。

数学的に「均衡からのズレが移る元の戦略と比べて移った先の戦略に不利に働く」というのを表しましょう。これはズレが移った先に不利に働くとは，各戦略の均衡からのシェアの増減と利得の増減が負の相関になっているということです。シェアと利得の負の相関というのは，平均的にみると均衡よりもシェアが高くなっている戦略は利得が均衡よりも減るということを意味します。つまり，均衡からよりもシェアが大きくなってしまっていることにペナルティーがかけられているような感じです。

ゲームにおいては利得は人々が取る選択，特にポピュレーションゲームではプレーヤーたちの中の戦略分布によって決まるのでした。これをきちんというためにもう少し数学っぽい表現を使いましょう。各戦略iの利得をπ_i，そして社会の中での戦略（戦略）の分布を$\boldsymbol{x}$とおくと，π_iが$\boldsymbol{x}$の関数だということです[11]。ここで戦略分布というのは各戦略のシェアを集めたものですから，$\boldsymbol{x} = (x_1, \dots, x_S)$と表されます。均衡は$\boldsymbol{x}^* = (x_1^*, \dots, x_S^*)$とします。この均衡からの戦略$i$のシェア$x_i$と利得$\pi_i$の変化量を$\Delta x_i = x_i - x_i^*, \Delta\pi_i = \pi_i(\boldsymbol{x}) - \pi_i(\boldsymbol{x}^*)$,と表すと，それらが負の相関を持つというのは$\sum_{i\in\mathcal{S}} \Delta x_i \Delta\pi_i < 0$ということです。$x_i$をすべての戦略で集めて$\boldsymbol{x}$にしたように，$\Delta x_i$を$\Delta\boldsymbol{x} = (\Delta x_1, \dots, \Delta x_S)$と，$\Delta\pi_i$を$\Delta\boldsymbol{\pi} = (\Delta\pi_1, \dots, \Delta\pi_S)$と並べましょう。これらはベクトルなので，相関$\sum_{i\in\mathcal{S}} \Delta x_i \Delta\pi_i$はベクトルの内積$\Delta\boldsymbol{x} \cdot \Delta\boldsymbol{\pi}$であり，負の相関というのはこの内積が負ということです。

各戦略iについてπ_iが微分可能だとしましょう。π_iを$\boldsymbol{x}$に関して微分し均衡$\boldsymbol{x}^*$で評価した$\frac{d\pi_i}{dx}(\boldsymbol{x}^*)$を用いると，戦略$i$の利得の変化$\Delta\pi_i$は$\frac{d\pi_i}{dx}(\boldsymbol{x}^*)\Delta x = \sum_{j\in\mathcal{S}} \frac{\partial\pi_i}{\partial x_j}(\boldsymbol{x}^*)\,\Delta x_j$として近似できます[12]。そして，ヤコビアン$\frac{d\boldsymbol{\pi}}{dx}(\boldsymbol{x}^*)$を用いると，その利得の変化を全ての戦略でまとめたベクトル$\Delta\boldsymbol{\pi} = (\Delta\pi_1, \Delta\pi_2, \dots)$は$\frac{d\boldsymbol{\pi}}{dx}(\boldsymbol{x}^*)\Delta\boldsymbol{x}$として近似できます[13]。従って，負の相関というのは$\Delta\boldsymbol{x} \cdot \frac{d\boldsymbol{\pi}}{dx}(\boldsymbol{x}^*)\Delta\boldsymbol{x} < 0$ということになり

---

[11] 以下では取り得る戦略を{戦略 1,戦略 2}の2つとは絞らずに，{戦略 1,…,戦略S}とS個あるとして，この取りうる戦略の集合を筆文字の$\mathcal{S}$で表し，つまり$\mathcal{S} = \{1, \dots, S\}$とします。また複数の数を並べたベクトルということをはっきりさせるため，太字でベクトルを表します。このとき並べるのは（この文中ではスペースの節約のため横に書きますが）縦に並べてるとします。つまり太字は列ベクトルです。また前節とは異なり，微分による特徴づけを明らかにするために，利得の戦略分布に対する線形性は仮定しません。

[12] $\frac{d\pi_i}{dx}(\boldsymbol{x}^*) := \left(\frac{\partial\pi_i}{\partial x_1}(\boldsymbol{x}^*), \dots, \frac{\partial\pi_i}{\partial x_S}(\boldsymbol{x}^*)\right)$は，$\pi_i$を$x$の各成分，つまり各戦略のシェア$x_1, \dots, x_S$に関して偏微分して$\boldsymbol{x}^*$で評価したものを横に並べた行ベクトルです。

[13] ここで$\frac{d\boldsymbol{\pi}}{dx}(\boldsymbol{x}^*) := \left(\frac{d\pi_1}{dx}(\boldsymbol{x}^*), \frac{d\pi_2}{dx}(\boldsymbol{x}^*), \dots,\right)$は，行ベクトルとなる$\frac{d\pi_i}{dx}(\boldsymbol{x}^*)$を各行動の利得関数$\pi_i$で得て，それをすべての行動で（縦に）並べたものです。これを$\boldsymbol{\pi}$の$\boldsymbol{x}$に関するヤコビアン（ヤコブ行列，

ます。従って，進化的安定というのは，$\Delta \boldsymbol{x} \cdot \frac{d\boldsymbol{\pi}}{d\boldsymbol{x}}(\boldsymbol{x}^*)\Delta \boldsymbol{x} < 0$が任意の（戦略分布として動きうる）方向の$\Delta \boldsymbol{x}$で成立するという条件になります。これは線形代数の用語を使うと，$\frac{d\boldsymbol{\pi}}{d\boldsymbol{x}}(\boldsymbol{x}^*)$が負値定符号になっているということです[14]。

定義：完全に混ざっている均衡の進化的安定性

どの戦略も正のシェアをもつ，つまり完全に混ざっている均衡x^*が進化的安定であるとは，
均衡x^*で評価した利得関数πの戦略分布xに関するヘッシアン$\frac{d\pi}{dx}(x^*)$が負値定符号となることである。

この負値定符号の条件が満たされていれば，上述の動学や広く「インセンティブに整合的な」動学で，実際に均衡が安定的（ズレても戻ってくる）ことが知られています。(Hofbauer & Sandholm, 2009; Mertikopoulos & Sandholm, 2018; Zusai, 2022.) 本書のように，利得が単に戦略分布に対する期待利得，つまり[相手の取りうる各戦略のシェア]×[利得表でのその戦略に対する自分の利得]という積を相手の戦略全てに関して足し合わせたものなら，ヘッシアン$\frac{d\boldsymbol{\pi}}{d\boldsymbol{x}}(\boldsymbol{x}^*)$は単に利得表での自分の利得に関する利得行列です。

このモデル 4.1 では各プレーヤーの利得行列は負値定符号ではありません（むしろ正値定符号です）。実際，お店の方もお客さんの方も利得行列$\boldsymbol{\Pi}$と相手方の戦略分布を表す$\boldsymbol{x}'$を

$$\boldsymbol{\Pi} = \begin{pmatrix} 3 & -1 \\ 0 & 1 \end{pmatrix}, \boldsymbol{x}' = \begin{pmatrix} \text{相手方の中で「使う」シェア} \\ \text{相手方の中で「否」のシェア} \end{pmatrix}$$

として，利得を$\boldsymbol{\Pi x}'$と書けます。そして任意の2次元ベクトル$\boldsymbol{v} = (v_1, v_2)$に対して$\boldsymbol{v} \cdot \boldsymbol{\Pi v} = 3v_1^2 - v_1 v_2 + v_2^2 = 0.5(v_1 - v_2)^2 + 2.5v_1^2 + 0.5v_2^2$となり，$\boldsymbol{v}$が零ベクトル$(0,0)$でない限り$\boldsymbol{v} \cdot \boldsymbol{\Pi v} > 0$なので，この$\boldsymbol{\Pi}$は正値定符号です。従って，$(p, q) = (0.4, 0.4)$は進化的安定ではありません。

Jacobian）といいます。細かいところですが，$\frac{d\boldsymbol{\pi}}{d\boldsymbol{x}}(\boldsymbol{x}^*)$は$S \times S$行列であり，$\Delta \boldsymbol{x} = \boldsymbol{x} - \boldsymbol{x}^*$は列ベクトルなので$S \times 1$行列とみなせるので，$\frac{d\boldsymbol{\pi}}{d\boldsymbol{x}}(\boldsymbol{x}^*)\Delta \boldsymbol{x}$は行列の掛け算で$S \times 1$行列，つまり行ベクトルになります。

[14]負値定符号の意味を簡単に説明しておきましょう。$\mathbf{M}$を$n \times n$の行列とします。同じn次元のベクトル$\boldsymbol{v}$にこの行列をかけたもの（つまり行列によって線形変換した）もの$\mathbf{M}\boldsymbol{v}$と，変換前のベクトル$\boldsymbol{v}$自体との内積は$\boldsymbol{v} \cdot \mathbf{M}\boldsymbol{v}$と書けます。このベクトル$v$が零ベクトル$(0,0,\dots,0)$でない限り，この内積が常に負である，つまり$\boldsymbol{v} \cdot \mathbf{M}\boldsymbol{v} < 0$なら，この行列$\mathbf{M}$は負値定符号(negative definite)であるといいます。これが反対に$\boldsymbol{v} \cdot \mathbf{M}\boldsymbol{v} > 0$なら$\mathbf{M}$は正値定符号です。行列の固有値を習っているなら，負値定符号であるためにはすべての固有値が負であることが必要十分であることは計算で判別するために覚えておくとよいでしょう。

細かいところを詰めると，この二つの状況の間に，いくつかの戦略のシェアは正だけど，他にシェアがゼロの戦略もあるという均衡も考えられます。また，ここではポピュレーションゲームとして各プレーヤーが純粋戦略を取るとき（混合戦略はそれを集計した戦略分布）を考えました。進化的安定に関しては，各プレーヤーが混合戦略をとれるときのものもポピュラーです。こうしたときも上述のふたつの定義での条件が肝とはなりますが，これを組み合わせ，また微妙な調整を必要とします。そこまで踏み込むのは沼にはまるので，もしも気になったら，様々な状況での進化的安定の概念について簡潔にまとめている Cressman (2010)を参照してください。

演習問題

1． 本書の他の例において，各均衡が進化的安定になっているかを判定してみよ。また，tBRD と成功の模倣の動学，それぞれの遷移ベクトルを均衡点で求めて，それが零ベクトルになっていることを確認せよ。また少し離れたところで遷移ベクトルを計算することで，その均衡点がその動学で安定的かどうかの見当をつけてみよ。

略答

手荷物検査（表 4-2）で唯一の均衡$(p, q) = (0.2, 0.75)$は進化的安定。渋滞（モデル 4-2）も唯一の均衡$p = 2/3$は進化的安定。動学的安定性もこれらの例では最適反応動学と同じ。

参考文献

Björnerstedt, J. and Weibull, J.W. (1996) “Nash equilibrium and evolution by imitation.” In: K.J. Arrow, E. Colombatto, M. Perlman, C. Schmidt (Eds.), *The Rational Foundations of Economic Behaviour*, St. Martin’s Press, pp. 155-171.

Cressman, R. (2010). “Learning and Evolution in Games: ESS.” In: Durlauf, S.N. & Blume, L.E. (eds.) *The New Palgrave Dictionary of Economics*, Ed.2, Palgrave Macmillan.

Sandholm, W.H. (2010) *Population Games and Evolutionary Dynamics*, MIT Press.

Hofbauer, J. (1995) “Imitation dynamics for games.” mimeo, University of Vienna.

Hofbauer, J. and Sandholm, W.H. (2009) “Stable games and their dynamics,” *Journal of Economic Theory*, 144, 1665-1693.

Mertikopoulos, P. and Sandholm, W. H. (2018): “Riemannian game dynamics,” *Journal of Economic Theory,* 177, 315 - 364.

Schlag, K.H. (1998) “Why imitate, and if so, how? A boundedly rational approach to multi-armed bandits.” *Journal of Economic Theory,* 78,130-156.

Zusai, D. (2018) “Tempered best response dynamics.” *International Journal of Game Theory* 47 (1), pp.1–34.

Zusai, D. (2022) “Net gains in evolutionary dynamics: A unifying and intuitive approach to dynamic stability.” https://arxiv.org/abs/1805.04898

田村光平(2023)『つながりの人類史』PHP 研究所。

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