数理解析研究所講究録
1043 巻 1998 年 237-240

ゲーム力学における戦略の確率密度の導出

広島大学工学部　奥原 浩之　Koji Okuhara

広島大学工学部　尾崎 俊治　Shunji Osaki

1. はじめに

生物の形質として，大きく親から受け継いだ生得的な形質と，成長の過程での学習による獲得的な形質を考えることができる．これらは，共に外界の環境との相互作用の結果，変質していくものである．形質の変化は，相互作用を通じた競合による淘汰と，構造の不安定化による分岐を利用した分裂により押し進められているものと考えられる．われわれは既に，この概念に従い一部（単純系）の獲得的な形質の変化（学習法）を与える手法を導出している [1]．そこでは，変数間の相互作用が明らかであれば，各変数が収束する確率密度を得ることができる．この手法はニューラルネットワークの学習則 [2] として適用され，構造が変化することにより環境（教師信号）の変化に適応する能力を実現している．

そこで本研究では，生得的な形質の変化を担う遺伝子の表現型に対しても，同様な議論ができることを述べる．ところで，遺伝子をプレイヤーとみなし，その表現型を純あるいは混合戦略と考えれば，表現型の単位あたりの増加率はゲーム力学方程式 [3] に従う．このとき、表現型が Nash 均衡かつ安定な状態である場合は進化的に安定な状態 (Evolutionarily Stable Strategies：以後 ESS)[4] と呼ばれる．そこで，プレイヤーの戦略決定にプレイを行う環境からの外乱が影響を及ぼす場合を考え，ゲームにおける利得行列が与えられれば，戦略の確率密度関数が導出できることを示す。さらに、パートナーシップ・ゲームである場合には ESS の分布が導けることを示す．

2. ゲーム力学と ESS

まず，本研究で考えるゲーム力学を定式化する．集団 X は N 個の表現型 $E_1, E_2, \cdots, E_N$ をもつものとする．表現型 E_i に対する適応度（利得）f_i は集団の状態に依存する．そこで，表現型 E_i の状態を頻度 x_i で与え，集団の状態をベクトルを $\mathbf{x} = [x_1, x_2, \cdots, x_N]^{\mathrm{T}} \in R^N$ で表す．ここで，$x_1 + x_2 + \cdots + x_N = 1$ である．その結果，表現型 E_i を純戦略と見なせば，単体の点 $\mathbf{x} \in S_N$ は混合戦略と考えることができる．このとき，表現型 E_i は端点 $\mathbf{e}_i$ に対応する．ここで，$S_N = \{\mathbf{x} : \sum_i x_i = 1, x_i \geq 0\ (\forall i)\}$ である．また，表現型 E_i を混合戦略と見なして，単体の点 $\mathbf{p} \in S_N$ で定義することも可能である．

表現型 E_i の単位あたりの増加率 $\dot{x}_i/x_i$ が適応度 $f_i(\mathbf{x})$ と集団の平均適応度 $\bar{f}(\mathbf{x}) = \sum_i x_i f_i(\mathbf{x})$ を用いて

$$\frac{\dot{x}_i}{x_i} = f_i(\mathbf{x}) - \bar{f}(\mathbf{x}) \tag{1}$$

で与えられるものとする．いま，適応度 $f_i(\mathbf{x})$ が線形である場合を考える．このとき，表現型 E_i の増加率 $\dot{x}_i$ は

$$\dot{x}_i = x_i\{(\mathbf{A}\mathbf{x})_i - \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}\} \tag{2}$$

となる．ここで，$\mathbf{A} = (a_{ij})$ は $N \times N$ 適応度行列（利得行列）であり，$(\mathbf{A}\mathbf{x})_i$ は $\sum_j a_{ij}x_j$ を表す．本研究では式 (2) をゲーム力学方程式 [3] とよぶ．ここで，3 個の純戦略に 2 個の表現型の状態空間の例を図 1 に示す．

次に，ESS の定義 [4] を述べる．2 つの戦略 $\mathbf{p} \in S_n$ と $\mathbf{q} \in S_n$ で表される表現型 E_1 と E_2 をもつ集団を考える．集団はそれぞれ頻度 $1-\epsilon$ と ϵ でもつ混合種の集団 $\mathbf{m} = \epsilon\mathbf{q} + (1-\epsilon)\mathbf{p}$ であるとする．ESS は表現型 E_1 である集団に表現型 E_2 である集団が少数移入してきても，そこでは増殖することができない状態をいう．つまり，集団混合が $\mathbf{m}$ であるときの戦略 $\mathbf{p}$ に対する利得は，その他全ての戦略 $\mathbf{q}(\neq \mathbf{p})$ に対する利得より大きいこと

$$\mathbf{q}^{\mathrm{T}}\mathbf{A}(\epsilon\mathbf{q} + (1-\epsilon)\mathbf{p}) < \mathbf{p}^{\mathrm{T}}\mathbf{A}(\epsilon\mathbf{q} + (1-\epsilon)\mathbf{p}) \tag{3}$$

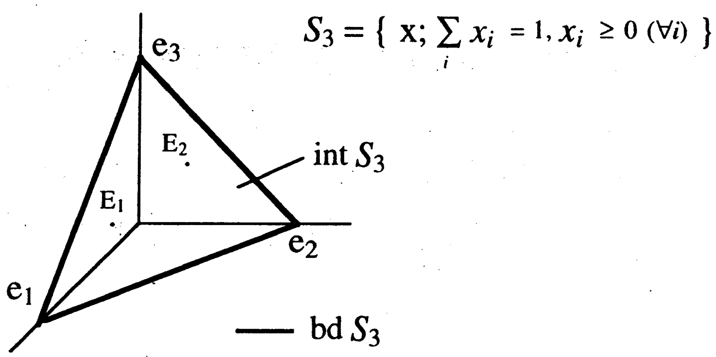

図 1　3 個の純戦略に 2 個の表現型の状態空間の例

が十分小さな$\epsilon > 0$について成り立つ時をいう．これより，戦略 $\mathbf{p}$ が ESS である必要十分条件は次の 2 つの条件

$$(Nash\ 均衡)\quad 全ての戦略\ \mathbf{q}\ に対して, \mathbf{q}^{\mathrm{T}}\mathbf{A}\mathbf{p} \leq \mathbf{p}^{\mathrm{T}}\mathbf{A}\mathbf{p}.$$
$$(安定条件)\quad \mathbf{q}^{\mathrm{T}}\mathbf{A}\mathbf{p} = \mathbf{p}^{\mathrm{T}}\mathbf{A}\mathbf{p}\ ならば\ \mathbf{q}^{\mathrm{T}}\mathbf{A}\mathbf{q} < \mathbf{p}^{\mathrm{T}}\mathbf{A}\mathbf{q}.$$

の成立であることがわかる．Nash 均衡は戦略 $\mathbf{p}$ に対してそれ自身 $\mathbf{p}$ のみならず，代替の戦略 $\mathbf{q}$ が最適な応答となりうることを示している．安定条件は代替の表現型が存在する場合でも，戦略 $\mathbf{p}$ が最適な応答となることを示している．これらをまとめて戦略 $\mathbf{p}$ が ESS である必要十分条件はつぎのように与えられる [5].

定理　戦略 $\mathbf{p} \in S_N$があある近傍全ての戦略 $\mathbf{m}\ (\neq \mathbf{p})$ について $\mathbf{m}^{\mathrm{T}}\mathbf{A}\mathbf{m} < \mathbf{p}^{\mathrm{T}}\mathbf{A}\mathbf{m}$ となれば ESS である．

このことから，戦略 $\mathbf{p} \in intS_N$が ESS であるなら，唯一の ESS であることがわかる．また，戦略 $\mathbf{p} \in bdS_N$が ESS であるなら，複数の ESS(全て bdS_Nに存在) をもつことが可能である．

ここで，ESS の概念を簡単な”タカ・ハト”ゲームで説明する．ハト派同士が出会いプレイすると勝者は利得 G を得，敗者は利得を得なければ失いもしないものとする．ハト派がタカ派に出会うとハト派は逃げることで利得を得ず，タカ派は利得 G を得る．タカ派同士が出会いプレイすると闘争により勝者は利得 G を得るが，敗者は利得 Cを失う．これらを利得行列にまとめると表 1 のようになる．

表 1　”タカ・ハト”ゲームの利得表

	タカ派	ハト派
タカ派	$\frac{G-C}{2}$	G
ハト派	0	$\frac{G}{2}$

集団において各派の頻度が x_1と $x_2(= 1 - x_1)$ であるとする．このとき，各適応度は

$$f_1(\mathbf{x}) = \frac{G-C}{2}x_1 + Gx_2 \tag{4}$$
$$f_2(\mathbf{x}) = \frac{G}{2}x_2 \tag{5}$$

となる．よって各適応度が等しくなる状態 $x_1 = G/C$へ状態が進むと考えられる．定理に従い，進化的に安定な状態を求めると

$$\mathbf{p}^{\mathrm{T}}\mathbf{A}\mathbf{m} - \mathbf{m}^{\mathrm{T}}\mathbf{A}\mathbf{m} = \frac{1}{2C}(G - C\epsilon)^2 \tag{6}$$

が得られ，戦略$\mathbf{p}=[G/C,(C-G)/C]^{\mathrm{T}}$がESSであることがわかる．この戦略は$\mathbf{p}\in intS_N$であることから唯一のESSである．

3. 戦略が従う確率密度の導出

戦略$\mathbf{p}\in S_N$がESSであるならば．戦略$\mathbf{p}$はゲーム力学方程式の漸近安定な休止点であることが知られている[6]．そのため，利得行列$\mathbf{A}$をもつゲームにおいて最適戦略が唯一である場合は，任意の初期状態$\mathbf{x}(0)$を与えゲーム力学方程式に従い状態を遷移させるとESSが導出できる．しかし，複数の最適戦略が存在する場合は，初期状態の与え方によって得られるESSが異なることとなる．そこで，利得行列$\mathbf{A}$が与えられれば，戦略の確率密度が導出できれば有益であると思われる．

まず，ゲーム力学方程式の軌道$\mathbf{x}$はLotka-Volterra方程式

$$\dot{y}_i=g_i(\mathbf{y})=y_i(\mathbf{C}+(\mathbf{B}\mathbf{y})_i) \tag{7}$$

の軌道と位相軌道同値(O.E.D.)であることが示されている[7]．ここで，$\mathbf{y}=[y_1,y_2,\cdots,y_{N-1}]^{\mathrm{T}}\in R^{N-1}$は状態ベクトルである．$\mathbf{B}=(b_{ij})$は$(N-1)\times(N-1)$行列であり，$\mathbf{C}=[c_1,c_2,\cdots,c_{N-1}]^{\mathrm{T}}\in R^{N-1}$である．ただし，

$$y_i=x_i/x_N \tag{8}$$

$$b_{ij}=a_{ij}-a_{Nj} \tag{9}$$

$$c_i=a_{iN}-a_{NN} \tag{10}$$

の関係があるものとする．利得行列$\mathbf{A}$の第N行目は第k行目と可換である．この変換はS_NからR_+^{N-1}への微分可能で可逆な写像となっている．そこで，

$$w_i^2=y_i \tag{11}$$

と変数変換する．変数y_iの定義域は正であるのに対し，w_iの定義域は任意の実数である．そこで，式(5)を

$$\frac{dw_i}{dt}=(\frac{c_i}{2}-\sum_{j=1}^{N-1}\frac{b_{ij}}{2}w_j^2)w_i+q_i \tag{12}$$

とする．ここで，q_iはGauss型白色雑音である．この式を変形されたLotka-Volterra方程式とよぶこととする．このとき，$b_{ij}=b_{ji}$が満たされていれば，式(7)のドリフトはポテンシャル

$$V(\mathbf{w})=-\sum_{i=1}^{N-1}\{\frac{c_i}{4}w_i^2-\sum_{j\neq i}^{N-1}\frac{b_{ij}}{8}w_j^2w_i^2-\frac{b_{ii}}{8}w_i^4\} \tag{13}$$

から導くことができる．今，式(8)を離散化し近似した見本過程として

$$w_i(t+\Delta t)=w_i(t)-\frac{\partial V(\mathbf{w},\phi)}{\partial w_i}\Delta t+\sqrt{Q\Delta t}\,\sigma_i(t) \tag{14}$$

を考えることができる．ただし，$\sigma_i(t)$は独立な確率変数であり，平均0，分散1の正規分布$N(0,1)$に従う．その結果，変数$\mathbf{y}$の確率密度は

$$p_\beta(\mathbf{y})=Z_\beta^{-1}\exp\{-\beta V(\mathbf{y},\phi)\} \tag{15}$$

で得ることができる[8]．ここで，$\beta=2/Q$である．Zは分配関数である．この結果から，変数$\mathbf{y}$を状態ベクトル$\mathbf{x}$へ戻せば利得行列$\mathbf{A}$に対する戦略の従う確率密度が得られることとなる．

特に、パートナーシップ・ゲームのように利得行列$\mathbf{A}$が対称行列である場合には、漸近的に安定な休止点がESSであることから、戦略の従う確率密度はすなわちESSの確率密度となる。

以上の結果をまとめると，N個の純戦略と利得行列$\mathbf{U}$をもつ線形ゲームから最適戦略の従う分布の導出アルゴリズムは次のようになる．

[1] N個の純戦略と利得行列 $\mathbf{U}$ をもつ線形ゲームを考える．

[2] ゲーム力学方程式の適応度行列 $\mathbf{A}$ ($n \times n$ 行列) は利得行列 $\mathbf{U}$ ($N \times N$行列) より $a_{ij} = \mathbf{p}_i \mathbf{U} \mathbf{p}_j$で得られる．

[3] 適応度行列 $\mathbf{A}$ より式 (9),(10) より行列 $\mathbf{B}$ ($(n-1) \times (n-1)$ 行列) とベクトル $\mathbf{C}$ ($\in R^{n-1}$) を得る．その結果，式 (8) の変数変換 ($\mathbf{x} \rightarrow \mathbf{y}$) をとおして Lotka-Volterra 方程式が導かれる．

[4] さらに式 (11) の変数変換 ($\mathbf{y} \rightarrow \mathbf{w}$) により，変形された Lotka-Volterra 方程式が得られる．

[5] ポテンシャルを求めることで確率密度関数が式 (15) のような Gibbs 分布により導出される．

[6] 導出された確率密度関数は変数 $\mathbf{w}$に対するものであるので，頻度 $\mathbf{x}$ に対する確率密度関数を導出する．

4. まとめ

本研究では，プレイヤーの戦略決定にプレイを行う環境からの外乱が影響を及ぼす場合を考え，ゲームにおける利得行列が与えられれば，最適戦略の分布が導出できることを示した．本手法に学習を組み込めば，ゲームの利得行列が未知である場合でもプレイをとおして学習することにより，利得行列の推定が可能であることを示すことが今後の課題である．

参考文献

[1] 佐々木浩二, 奥原浩之, 尾崎俊治, “動径基底関数を複製する競合動径基底関数ネットワークの提案”, 平成 9 年度電気・情報関連学会中国支部連合大会講演論文集, 広島, (October 26-27, 1997).

[2] 奥原浩之, 尾崎俊治, “適者生存型学習則を適用した競合動径基底関数ネットワークの提案”, 電子情報通信学会論文誌, (採録決定済).

[3] P. Taylor, and L. Jonker, “Evolutionarily stable strategies and game dynamics, ”*Math. Biosci.*, **70**, pp. 145-156 (1978).

[4] J. Maynard Smith, “The theory of game and the evolution of animal conflicts,” *J. Theor. Biol.*, **47**, pp. 209-221 (1974).

[5] I. Bomze, “Non-cooperative two-person games in biology: a classification,” *J. Theor. Biol.*, **15**, pp. 31-57 (1986).

[6] E. C. Zeeman, “Population dynamics from game theory.: In global theory of dynamical systems,” *Springer Lecture Notes in Math.*, **819** (1980).

[7] J. Hofbauer, “On the occurrence of limit cycle in the Volterra-Lotka equation, ”*Nonlinear Analysis*, **5**, pp. 1003-1007 (1981).

[8] 奥原浩之, 尾崎俊治,, “一般 Lotka-Volterra 方程式における逆問題の解法”, 電子情報通信学会論文誌, (採録決定済).