日本オペレーションズ・リサーチ学会
2023 年春季研究発表会

ベイズ相関均衡と no-regret dynamics

05000537　国立情報学研究所　*藤井海斗　FUJII Kaito

1.　相関均衡とは

相関均衡とは，仲介者の存在を仮定することで実現する均衡である．本研究では，ベイジアンゲームにおける相関均衡について考える．

まず，相関均衡の説明のために，以下のような交差点におけるゲームを考える（表 1）．交差点に西から入ってくる車をプレイヤー 1，北から入ってくる車をプレイヤー 2 とする．それぞれの車は「進む」と「停まる」という二つの行動をとりうる．どちらも「進む」を選ぶならば，交通事故が発生し，どちらも利得 0 を得る．どちらも「停まる」を選ぶならば，事故は起きないものの，交差点で無意味に停止することになり，どちらも利得 1 を得る．片方が「進む」，もう片方が「停まる」を選ぶ場合，進んだ車は利得 4 を得て，停まった車は利得 3 を得る．

表 1: 交差点のゲーム

	進む	停まる
進む	$0, 0$	$4, 3$
停まる	$3, 4$	$1, 1$

このようなゲームは標準形ゲームと呼ばれるが，標準形ゲームにおける最も有名な均衡のクラスはナッシュ均衡である．ナッシュ均衡とは，各プレイヤーの（確率的な）行動の組であり，どのプレイヤーもそこから逸脱するインセンティブを持たないものを指す．このゲームのナッシュ均衡は 3 つ存在する．一つ目は，プレイヤー 1 が「進む」，プレイヤー 2 が「停まる」を選択するという純粋ナッシュ均衡である．二つ目は，一つ目とプレイヤーを入れ替えた純粋ナッシュ均衡である．三つ目は，どちらのプレイヤーも「進む」と「停まる」を確率 1/2 ずつで選択するという混合ナッシュ均衡である．このように，ナッシュ均衡では各プレイヤーがそれぞれ独立に行動を決定する．

一方で相関均衡では，仲介者が手伝うことで，相関した行動を選択できる．例えば，信号機が仲介者の役割をする場合を考える．ここでは少し抽象化された信号機を想定し，確率 1/2 で東西に赤信号，南北に青信号を表示し，確率 1/2 で反対の表示をすると仮定する．すると，確率 1/2 で一つ目のナッシュ均衡が，確率 1/2 で二つ目のナッシュ均衡が実現する．ここで重要なのは，二つの車の行動が相関していることである．実際に，各プレイヤーの行動は確率的に変化するが，常に相手とは逆の行動をとっている．これはそれぞれが独立に選択した場合には実現できない．このように，信号機のような仲介者が各プレイヤーに行動を推薦することで，プレイヤーの行動のあいだに相関のある均衡が実現できる．

より正確には，相関均衡は以下のように定義される．まず，標準形ゲームの基本的な記法を定義しておく．プレイヤーの集合を $N = \{1, 2, \ldots, n\}$ で表し，各プレイヤー $i \in N$ のとりうる有限個の行動の集合を A_i とする．各プレイヤーの行動の組を $a = (a_1, \ldots, a_n)$ で表し，行動の組全体を $A = A_1 \times \cdots \times A_n$ と書く．また，プレイヤー $i \in N$ 以外の行動の組を $a_{-i} = (a_j)_{j \in N \setminus \{i\}}$ と表記する．各プレイヤーが行動の組 $a \in A$ を選んだときのプレイヤー $i \in N$ の利得を，利得関数 $v_i \colon A \to [0, 1]$ を用いて $v_i(a)$ と書く．すると，相関均衡は，行動の組の分布 $\pi \in \Delta(A)$ であって，各 $i \in N$，各 $\phi \colon A_i \to A_i$ について

$$\mathbb{E}_{a \sim \pi}\left[v_i(a_i, a_{-i})\right] \geq \mathbb{E}_{a \sim \pi}\left[v_i(\phi(a_i), a_{-i})\right]$$

を満たすもののことである．ここで，$\phi \colon A_i \to A_i$ は行動 $a_i \in A_i$ を推薦されたときに $\phi(a_i)$ を選択するような逸脱を表している．相関均衡はナッシュ均衡を含むより広いクラスであり，逆に，相関均衡をプレイヤーごとに独立な分布に限定するとナッシュ均衡になる．

ナッシュ均衡を近似的に計算することの困難性が知られている一方で，近似的な相関均衡は効率的に計算できることが知られている．具体的には，各プレイヤーが同じゲームを何度もプレイするなかで徐々に戦略を改善していく状況をシミュレートすればよい．このとき，各プレイヤーがある種の

リグレットを最小化していれば，経験分布が近似的な相関均衡へと収束することを保証できる [2, 3].

2.　ベイジアンゲーム

本研究では，ベイジアンゲームにおける相関均衡を考える．ベイジアンゲームとは，各プレイヤーが私的情報をもっているような標準形ゲームの拡張である．各プレイヤー $i \in N$ の私的情報はタイプ $\theta_i \in \Theta_i$ と呼ばれる変数で表現される．ここで，Θ_i はプレイヤー i のとりうるタイプ全体の集合であり，$\Theta = \Theta_1 \times \cdots \times \Theta_n$ とする．タイプは，プレイヤーの選好を表す属性のようなものであり，各プレイヤーの利得はタイプによって変化する．例えば，上述の交差点のゲームにおいて，事故を恐れる普通の車と，事故を恐れない危険な車の二種類があるとする．普通のタイプを 0，危険なタイプを 1 とすると，$\Theta_1 = \Theta_2 = \{0, 1\}$ となる．そして，どちらも「進む」を選択するときに危険なタイプは利得 3 を得るが，他の場合の利得は普通の車と同じとする．このとき，利得行列は表 2 のようになる．

表 2: 交差点のベイジアンゲーム

	進む	停まる
進む	$3\theta_1, 3\theta_2$	$4, 3$
停まる	$3, 4$	$1, 1$

このように，タイプによって利得関数が変化するのがベイジアンゲームの特徴である．一般には，ベイジアンゲームにおける利得関数 $v_i \colon \Theta \times A \to [0, 1]$ は，タイプの組 $\theta \in \Theta$ と行動の組 $a \in A$ を受け取って利得の値 $v_i(\theta, a)$ を返す[1].

ベイジアンゲームでは，タイプの組 θ は事前分布 $\mu \in \Delta(\Theta)$ に従って確率的に決まり，この事前分布をあらかじめ全員が知っているとする．例えば，それぞれの車のタイプは確率 1/2 ずつで決まるが，それぞれが実際に知ることができるのは自分のタイプだけである．

3.　コミュニケーション均衡

ベイジアンゲームにおける相関均衡（ベイズ相関均衡）にはいくつかの等価でない定義が知られている [1] が，ここではそのうちコミュニケーション均衡を考える．コミュニケーション均衡は，以下のようなシナリオにおいて妥当な解概念である．まず，仲介者が各プレイヤー $i \in N$ にタイプ $\theta_i \in \Theta_i$ を尋ねる．このとき，プレイヤーは自分の利益のために嘘のタイプを教える可能性がある．次に，仲介者は教えてもらったタイプの組 $\theta \in \Theta$ に基づいて，行動の組 $a \in A$ をランダムに決定する．そして，各プレイヤー $i \in N$ に行動 a_i を推薦する．このとき，プレイヤーは推薦に従わずに別の行動をとる可能性がある．もし，プレイヤーが嘘のタイプを教えるインセンティブがなく，推薦に従わないインセンティブもないとき，このような分布をコミュニケーション均衡という．

より正確には，仲介者が従う分布はタイプの組 θ から行動の組の分布 $\pi(\theta) \in \Delta(A)$ への写像であり，これを $\pi \in \Delta(A)^\Theta$ で表す．プレイヤーが逸脱するインセンティブを持たないための条件は，各プレイヤー $i \in N$，タイプの逸脱 $\psi \colon \Theta_i \to \Theta_i$，推薦された行動からの逸脱 $\phi \colon \Theta_i \times A_i \to A_i$ について，

$$
\begin{aligned}
&\mathbb{E}_{\theta \sim \mu}\left[\mathbb{E}_{a \sim \pi(\theta_i, \theta_{-i})}\left[v_i(\theta; a_i, a_{-i})\right]\right] \\
&\geq \mathbb{E}_{\theta \sim \mu}\left[\mathbb{E}_{a \sim \pi(\psi(\theta_i), \theta_{-i})}\left[v_i(\theta; \phi(\theta_i, a_i), a_{-i})\right]\right]
\end{aligned}
$$

が成り立つことである．本研究では，コミュニケーション均衡に対応したリグレットを定義し，各プレイヤーがこのリグレットを最小化すれば，ダイナミクスの経験分布がコミュニケーション均衡に収束することを証明した．紙幅の都合のため詳細は割愛する．

参考文献

[1] Françoise Forges. Five legitimate definitions of correlated equilibrium in games with incomplete information. *Theory and Decision*, 35: 277–310 (1993).

[2] Dean P. Foster and Rakesh V. Vohra. Calibrated learning and correlated equilibrium. *Games and Economic Behavior*, 21(1-2): 40–55 (1997).

[3] Sergiu Hart, Andreu Mas-Colell. A Simple Adaptive Procedure Leading to Correlated Equilibrium. *Econometrica*, 68(5): 1127–1150 (2000).

[1] このように，利得の値はそのプレイヤーのタイプ θ_i だけでなく，他のプレイヤーも含めたタイプの組 $\theta \in \Theta$ に依存する．これは，タイプがそのプレイヤー本人の属性だけでなく，他のプレイヤーの利得に関わる情報を表現しうるためである．