ゲーム理論
――初学者のための基礎と無線リソース制御への応用・記述言語としての集合論――

山本高至　Koji Yamamoto　京都大学

1 はじめに

ゲーム理論は，利害関係のある意思決定主体間の相互作用を定式化し，議論するための数学である．本稿のメインターゲットとして想定したのは，研究室に配属される学生のような，ゲーム理論はもちろん，集合論等の記述に不慣れな方である．そもそもゲーム理論自体がどういうものなのか，どういう初歩的応用があるのかを知りたい方のために，**2.** でゲーム理論の数学面での基礎を，最適化理論との共通点・相違点に注目して説明する．次いで，**3.** で無線リソース制御への初歩的な応用例を述べる．このような応用も重要であるが，ゲーム理論はそれまでになかったパラダイム－物の見方・思考の枠組み－や記述言語を提供することも重要である．この点を **4.** で説明する．

そこから一歩進んで，ゲーム理論などの数学を勉強しようと思った方，あるいは勉強せざるを得なくなった方をターゲットとして，筆者の経験からヒントを２点提供したい．１点目として，「多くの大学生は数学が言語であるという認識に失敗したせいでつまずいている」[1] という指摘に代表されるように，数学には大学受験数学などに見られる計算方法という面だけではなく，記述言語という面があることを認識する必要がある．**5.** では，ゲーム理論自体が記述言語であることに加えて，ゲーム理論の記述言語として特に集合論が用いられていることを紹介する．２点目は筆者の意見として，記述言語としての数学を習得するには，数学の中での関連する分野－例えばゲーム理論であれば最適化理論－との共通点・相違点を意識すると理解しやすいのではないか，ということである．**2.** のゲーム理論の説明は，この観点も配慮している．また，**6.** では確率，線形代数，強化学習との関連を指摘する．

2 数学としてのゲーム理論

2.1　最適化問題と戦略形ゲーム

読者がよりなじみ深いであろう最適化問題をまず導入する．後述の戦略形二人ゲームと対比させるため，二つの決定変数 x_1，x_2 を持つ目的関数 f を考える．目的関数 f に関する最大化問題は，次式で表される．

$$\max_{(x_1, x_2)} f(x_1, x_2) \tag{1}$$

一方，戦略形二人ゲームは，以下の二つの最適化問題の組合せと捉えて差し支えない．

$$\begin{cases} \max_{x_1} f_1(x_1, x_2) \\ \max_{x_2} f_2(x_1, x_2) \end{cases} \tag{2}$$

ただし，ゲーム理論で一般的かつ望ましい表現に関しては，**5.** を参照されたい．

戦略形ゲーム（式（2））は，プレーヤ（主体）１が利得関数 f_1 を最大化する戦略 x_1 を，プレーヤ２が利得関数 f_2 を最大化する戦略 x_2 をそれぞれ決定する状況を表現できている．なお，ゲーム理論における戦略と利得関数という用語は，最適化理論における決定変数と目的関数に対応している．表１に最適化理論とゲーム理論の用語の対応をまとめる．

利得関数は両プレーヤの戦略 x_1，x_2 に依存する関数である一方，各プレーヤは各々の戦略を表す決定変数しか変更できないという点が，戦略形ゲーム（式（2））と最適化問題（式（1））との決定的な違いである．この際，プレーヤ２の戦略 x_2 に依存して，プレーヤ１の最適な戦略－利得関数 f_1 が最大となる戦略－が変わるという意味での相互作用が，戦略形ゲームでは定式化されている．

最適化問題の特別なクラスとして線形計画問題や凸計画問題などがあるのと同様，戦略形ゲームの特別なクラスとしてポテンシャルゲームなどがある．また，式

©電子情報通信学会 2019

表 1　最適化理論とゲーム理論の用語の対応

最適化理論	ゲーム理論
最適化問題	（戦略形）ゲーム
（対応する用語は存在しない）	プレーヤ
決定変数	戦略
制約条件・実行可能領域	戦略集合
目的関数	利得関数
最適解	ナッシュ均衡

(1) とは異なる表現の最適化問題として最適制御問題などがあるのと同様，戦略形ゲームとは異なるゲームとして展開形ゲームや進化ゲームなどがある．更には，ゲーム理論の応用としてマッチング理論やオークション理論などがある．

2.2　最適解とナッシュ均衡

最適化問題（式 (1)）の最適解は次式を満たす (x^*_1, x^*_2) である．

$$f(x^*_1, x^*_2) \geq f(x_1, x_2), \forall(x_1, x_2) \qquad (3)$$

一方，戦略形ゲーム（式 (2)）における解概念はナッシュ均衡と呼ばれる．まず与えられた x_2 に対して，f_1 が最大となる x_1 を最適応答と呼ぶ．次に，互いに最適応答となっている戦略組，具体的には次式を同時に満たす戦略組 $(x^\star_1, x^\star_2)$ としてナッシュ均衡が定義される．

$$\begin{cases} f_1(x^\star_1, x^\star_2) \geq f_1(x_1, x^\star_2), \forall x_1 \\ f_1(x^\star_1, x^\star_2) \geq f_2(x^\star_1, x_2), \forall x_2 \end{cases} \qquad (4)$$

ナッシュ均衡は，その状況にいれば，各プレーヤが自分だけ戦略を変更しても利得が向上せず，自分だけが戦略を変更する動機を持たないという意味がある．

3　無線リソース制御へのゲーム理論の応用

本章では，送信電力制御とチャネル割当という無線リソース制御の基本となる制御に関するゲーム理論の応用例について，基礎的な事項に限って紹介する．なお，無線リソース制御が必要となる理由は，複数の通信の間の干渉の存在と言える．筆者は干渉が相互作用であることに着目し，相互作用を定式化可能なゲーム理論を応用した研究を行ってきた[(2)]．

議論したい問題を戦略形ゲーム（式 (2)）をはじめとしたゲームとして定式化できれば，ゲーム理論を使って議論することが可能である．ただ，ゲーム理論で扱われるゲームは 2. で述べたように多岐にわたるため，個々の事例は各論的にならざるを得ないことを指摘しておきたい．

全体像を知りたいという読者には，ゲーム理論の無線通信応用を網羅的に述べた書籍[(3)]を紹介したい．しかしながら，余りに内容が多く，恐らく初学者には向いていないと思われる．筆者としては，まずは本章で説明するような基本となる幾つかのゲームを，納得できるまで理解することから始め，ボトムアップ的に全体像を感じて頂くのが望ましいのではないかと思う．

3.1　送信電力制御

伝送速度を最大化する送信電力制御について，最適化理論とゲーム理論の両方の側面から議論する．信号対干渉雑音電力比に対する伝送速度の関数として，シャノンの通信路容量の式を用いる．二つの送受信局組が同時送信をしている際，単位帯域当りの伝送速度の和は次式で表される[(4, §3.6)]．

$$f(p_1, p_2) = f_1(p_1, p_2) + f_2(p_1, p_2) \qquad (5)$$

$$f_1(p_1, p_2) = \log_2\left(1+\frac{G_{11}p_1}{G_{12}p_2+N}\right) \qquad (6)$$

$$f_2(p_1, p_2) = \log_2\left(1+\frac{G_{22}p_2}{G_{21}p_1+N}\right) \qquad (7)$$

ただし，p_i は送信局 i の送信電力，G_{ji} は送信局 i と受信局 j の間のチャネル利得，N は雑音電力を表す．

$f(p_1, p_2)$ は上に凸な関数でないため，式 (5) を目的関数とした最大化問題に対しては，非線形計画法で代表的な凸計画法のような単純な解法が存在するわけではない．このことは，符号反転した $-f(p_1, p_2)$ のヘッセ行列が，任意の (p_1, p_2) で半正定値とはならないことから確認される[(4, §3.6)]．そこで，何らかの解が得られるよう，問題を簡単化する．例えば，帯域を複数のチャネルに分割し，各チャネルでは単一の送受信局組のみ通信ができると限定する．この条件下で伝送速度の和を最大化するチャネル割当・送信電力同時設定問題は，凸計画問題となる[(4, §3.6)]．

一方，送受信局組 i の利得関数を $f_i(p_1, p_2)$ とした戦略形ゲームはどうなるだろうか．ここでは，送信電力 p_i の最大値を P_i とする．まず送受信局組 i の最適応答は，$p^\star_i = P_i$ である．なぜなら，$f_i(p_1, p_2)$ は p_i に関する単調増加関数だからである．これは，相手のプレーヤの戦略によらないことに注意されたい．したがって，ナッシュ均衡は，全ての送受信局組が最大電力を設定している状況である．これはゲーム理論を使わずとも明らかな結果ではあるが，ゲーム理論の枠組みで把握し直してみ

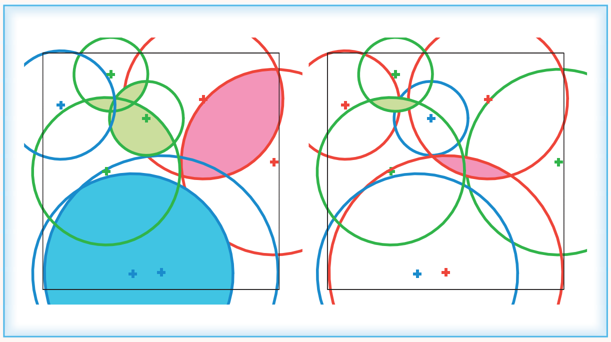

図１　オーバラップを低減するチャネル選択の前（左）と後（右）の例．円が送信電力に依存して異なるカバレージを，色がチャネルを表しており，チャネル選択により塗りつぶされたオーバラップが小さくなっていることが分かる．

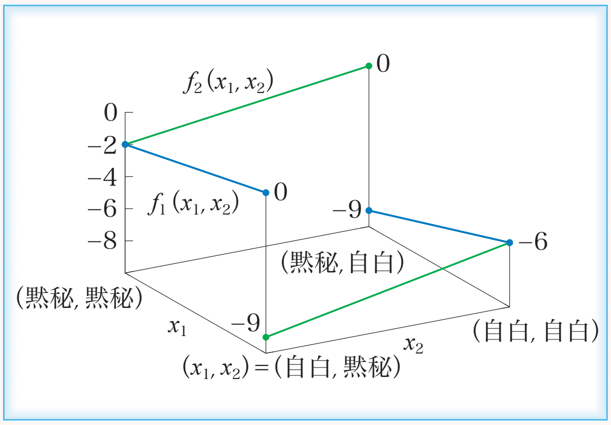

図２　囚人のジレンマ

ると，理解がより深まるのではなかろうか．

3.2　チャネル割当

一般に，分散的にチャネル選択を行うと，選択されるチャネルが収束しないことがある．これはサービスの観点からは望ましくない．収束しないことは，ゲーム理論の観点から言えば，ナッシュ均衡が存在していないことに対応する．ナッシュ均衡の存在が保証されている戦略形ゲームのクラスとして，ポテンシャルゲーム[(5), (6)]がある．ポテンシャルゲームとは，戦略の変更に関して，

$$f_1(x_1, x_2) - f_1(x_1', x_2) = \phi(x_1, x_2) - \phi(x_1', x_2) \tag{8}$$

$$f_2(x_1, x_2) - f_2(x_1, x_2') = \phi(x_1, x_2) - \phi(x_1, x_2') \tag{9}$$

を満たす関数ϕが存在する戦略形ゲームである．ポテンシャルゲームにおいては，各プレーヤの最適応答により，関数ϕが単調に増加する．戦略の選択肢数が有限であれば，最適応答に伴うϕの増加は有限回数で収束する．収束した点は，ナッシュ均衡である．

筆者らが提案した無線 LAN 向けの分散チャネル選択手法[(7), (8)]では，分散制御の際の収束性を保証するため，ポテンシャルゲームとなるように利得関数を設計している．例えば，無線 LAN のアクセスポイントのカバレージオーバラップを低減する分散チャネル選択手法[(7)]では，カバレージオーバラップ（を符号反転したもの）を利得関数とするゲームがポテンシャルゲームになることを証明している．図１はその結果例である．チャネル選択を含めたポテンシャルゲームの無線通信応用の詳細に関しては，サーベイ論文[(9)]や解説[(10)]を参照されたい．

4　パラダイムとしてのゲーム理論

３. で述べたような応用例があるという面とは別に，主体間の相互作用が定式化できるという，それまでになかったパラダイムを提供する点で，ゲーム理論は重要である．

有名な例であるが，ゲーム理論によって初めて定式化された主体間の相互作用として，囚人のジレンマ[(11)]を紹介する．まず，次の状況を考える．捕まった二人の共犯者は，二人とも黙秘していれば刑期２年だが，自白した側は司法取引により刑期なしに減刑される一方，黙秘を続けている側は刑期９年となる．双方が自白した場合は刑期６年となる．

図２を用いて説明する．図２の縦軸は刑期の符号反転であり，大きい方が望ましい．プレーヤiの戦略x_iは黙秘，自白のいずれかであり，例えばf_i（黙秘，黙秘）$=-2$は，二人のプレーヤが共に黙秘した場合，両プレーヤの刑期が２年であることを表している．

ここでナッシュ均衡になっているのは，$x_1 = x_2 =$ 自白の場合のみであり，双方が刑期６年である．ほかの状況では，少なくとも片方のプレーヤが戦略を変更することで自らの利得の向上（刑期の短縮）が可能である．すなわち，$x_1 = x_2 =$ 黙秘であれば，プレーヤ２は戦略x_2を自白に変更すれば，利得を-2から０に向上できる．また，$x_1 =$ 自白，$x_2 =$ 黙秘であれば，プレーヤ２は戦略x_2を自白に変更すれば，利得を-9から-6に向上できる．

一方，二人にとって望ましいのは，双方が刑期２年の，双方が黙秘の場合である．すなわち，戦略形ゲームの範囲においては，二人にとって望ましい結果には至らない．

ただ，これは単なる思考実験にすぎないとも言える．なぜこの結果が，特に経済学において大きなインパクト

を持つに至ったのだろうか．その答えを知るためには，それ以前のパラダイムを知る必要がある．以下は文献（12）の引用である．

「伝統的に経済学では（中略）各人が自己利益のみを考えて合理的に行動すれば，社会的に望ましい結果が得られるという，アダム・スミスの「見えざる手」を信じる傾向が強かったと言える．しかし，囚人のジレンマが教えるのは，それと反対のことである．（中略）そこで生じているメカニズムが正確に定式化されたことは，非常に大きなインパクトとなった．」

ゲーム理論が現代経済学に広範囲に影響を与えていることに関しても文献（12）を参照頂きたい．筆者自身の体験として，ゲーム理論を理解していたことで，多様な現代経済学，例えば行動経済学がどういうものなのかを大まかに理解することが容易であった．更には，行動経済学を一つの手段とした様々な論考，例えば無謀と分かっていた日米開戦がなぜ防げなかったのか[(13)]などといった社会科学の最新の論考にも触れることができている．

ゲーム理論がパラダイムを変えたという，別の説明を試みよう．自分が何か行動を起こすのではなく，「自分以外の人が変わるべきだ」という文句を聞くことがある．しかし，囚人のジレンマの例で言うと，黙秘したプレーヤが，自白したプレーヤに対し「お前さえ自白しなければ，刑期が9年になることはなかった」と文句を言っているようなものではないだろうか．そうであれば，文句を言うことは，問題を解決するという目的に対して合理的な行動とは言えない．このような問題を解決したければ，戦略形ゲーム－すなわちその構成要素であるプレーヤ，戦略，利得－を自分の変更可能な範囲で変更し，結果としての均衡を動かすしかない．

その上で，ガンジーの有名な言葉“Be the change that you want to see in the world.”「世界に変化を望むのであれば，私たち自らがその変化を起こさなければならない」[(14)]を考えてみよう．この言葉は，数多くある偉人の格言の一つに過ぎないとも捉えることはできる．しかし，ゲーム理論を理解してから読むと，真理としか思えないのではなかろうか．その意味で，ゲーム理論を理解すると，ものの見方が変わると言えるのではないだろうか．

5　ゲーム理論とその記述言語

以上ではできる限り，「難しい」数学表現を使わずに説明した．しかしながら，ゲーム理論の専門書や論文を読もうとすると，どうしても「難しい」数学表現は不可避である．なぜ「難しい」と感じるのかを理解しておけば，万一そのような表現に出会ったときに，何を勉強すればよいかを見つける手掛かりになるだろう．**5.1** では，戦略形ゲームの定義を述べる．また，ここで用いられる集合論の用語を **5.2** で説明する．

5.1　戦略形ゲームの定義

ゲーム理論やその応用を扱った書籍や論文では，戦略形ゲームやナッシュ均衡は以下のように定義される．

定義 1　プレーヤの「添字集合」[*1]を $\mathcal{I}$，プレーヤ $i \in \mathcal{I}$ の戦略集合を $\mathcal{X}_i$，利得関数を $f_i : \Pi_{j\in\mathcal{I}}\ \mathcal{X}_j \to \mathbb{R}$ とする．プレーヤの添字集合，戦略集合，利得関数の「順序組」

$$(\mathcal{I}, (\mathcal{X}_i)_{i\in\mathcal{I}}, (f_i)_{i\in\mathcal{I}}) \tag{10}$$

を戦略形ゲームと呼ぶ．

定義 2　戦略形ゲーム $(\mathcal{I}, (\mathcal{X}_i)_{i\in\mathcal{I}}, (f_i)_{i\in\mathcal{I}})$ において，$x_{-i} \in \mathcal{X}_{-i} := \Pi_{j\in\mathcal{I}\setminus\{i\}}\ \mathcal{X}_j$ に対するプレーヤ i の最適応答とは，次式で定義される「対応」である．

$$\mathrm{BR}_i(x_{-i}) = \underset{x_i \in \mathcal{X}_i}{\arg\max}\, f_i(x_i, x_{-i}) \tag{11}$$

定義 3　全てのプレーヤの戦略が，ほかのプレーヤの戦略に対して最適応答となっている戦略組を，ナッシュ均衡と呼ぶ．

定義1は戦略形二人ゲーム（式（2））を抽象化したものである．しかしながら，数学－特に集合論－に慣れていない場合，ゲーム理論以前につまずいてしまうのではなかろうか．

ただ，論文では定義1が好まれるし，望ましい．この理由は，必要な全ての構成要素が記述されていることや，どのような戦略形ゲームにおいてもナッシュ均衡を一意に定義できることであろう．

5.2　ゲーム理論の記述言語としての集合論

定義1，定義2の読解に必要な，言語としての集合論を復習しよう．集合の概念や用語は，「現代数学を語るための基礎的な言語という性格をもつ」[(15)]．本節の多くは，集合論の教科書[(15)]による．また，多くの数学記号は国際規格 ISO 80000-2:2009[(16)]に従っている．必要に応じて読者自身で調べられるよう，英語表記や文献における記載箇所も併記する．

まず，関数と，その一般化としての対応（correspon-

＊1　鍵括弧の用語は集合論の用語であり，5.2で定義する．

dence）を定義する．二つの集合$\mathcal{A}$と$\mathcal{B}$を考える．$\mathcal{A}$から$\mathcal{B}$への関数（あるいは写像）とは，$\mathcal{A}$の任意の元aに対して，$\mathcal{B}$の一つの元を定める規則である．これを$f: \mathcal{A} \to \mathcal{B}$と表記する(15, §1.5.E) (16, #11.3)．また，$\mathcal{A}$から$\mathcal{B}$への対応とは，$\mathcal{A}$の任意の元aに対して，$\mathcal{B}$の部分集合を定める規則である(15, §1.5.B)．対応は集合値関数（set-valued function）と呼ばれることもある．

5.2.1　戦略形二人ゲーム

プレーヤ数が2の場合を考える．戦略集合$\mathcal{X}_i$は，戦略x_iの取り得る値の集合を表している．すなわち，$x_i \in \mathcal{X}_i$である．**3.1**の送信電力制御であれば$\mathcal{X}_i = [0, P_i]$であり，**4.**の囚人のジレンマであれば$\mathcal{X}_i = \{$黙秘，自白$\}$である．

戦略の順序対（ordered pair）(x_1, x_2)を考える(15, §1.3.A)，(16, #5.14)．この順序対全体の集合は，$\mathcal{X}_1$と$\mathcal{X}_2$の直積（Cartesian product）と呼ばれ，$\mathcal{X}_1 \times \mathcal{X}_2$と表記される(15, §1.3.A)，(16, #5.16)．最適化理論で言えば，変数(x_1, x_2)の実行可能領域を表している．内包的記法(15, §1.1.B)，(16, #5.4)を用いれば，以下のとおりである．

$$\mathcal{X}_1 \times \mathcal{X}_2 := \{(x_1, x_2) \mid x_1 \in \mathcal{X}_1 \wedge x_2 \in \mathcal{X}_2\} \qquad (12)$$

$\wedge$は論理積（logical conjunction）(16, #4.1)，$:=$は右辺による左辺の定義(16, #7.3)を表す．

プレーヤiの利得関数f_iは，$\mathcal{X}_1 \times \mathcal{X}_2$から実数全体の集合$\mathbb{R}$ (15, §1.1.A)，(16, #11.3)への関数であり，$f_i: \mathcal{X}_1 \times \mathcal{X}_2 \to \mathbb{R}$と表記される．

プレーヤの集合$\{1, 2\}$，戦略集合の順序対$(\mathcal{X}_1, \mathcal{X}_2)$，利得関数の順序対$(f_1, f_2)$の三つが，戦略形二人ゲームを構成する必要十分な要素である．そこで，これらの順序組$(\{1, 2\}, (\mathcal{X}_1, \mathcal{X}_2), (f_1, f_2))$を，戦略形二人ゲームと呼ぶ．これは，グラフ理論において，頂点集合$\mathcal{V}$と辺集合$\mathcal{E}$の順序$(\mathcal{V}, \mathcal{E})$をグラフと呼ぶことに通じるところがある．順序組は次項で説明する．

5.2.2　任意のプレーヤ数の戦略形ゲーム

プレーヤ数がIの場合を考える．戦略の順序対の一般化である戦略の順序組（ordered tuple）＊2 $(x_1, x_2, \ldots x_I)$は$(x_i)_{i \in \mathcal{I}}$，あるいは略して$(x_i)$と表記される(16, #5.15)．ここで，$\mathcal{I} := \{1, 2, \ldots, I\}$を添字集合（index set）と呼ぶ(15, §1.5.B)．同様に戦略集合の順序組は$(\mathcal{X}_i)_{i \in \mathcal{I}}$と表記される．$(x_i)_{i \in \mathcal{I}}$全体の集合を$(\mathcal{X}_i)_{i \in \mathcal{I}}$の直積と呼び，これを$\Pi_{j \in \mathcal{I}} \mathcal{X}_j$と表記する(15, §1.5.D)，(16, #5.17)．プレーヤiの利得関数は$f_i: \Pi_{j \in \mathcal{I}} \mathcal{X}_j \to \mathbb{R}$である．

プレーヤの添字集合$\mathcal{I}$，戦略集合の順序組$(\mathcal{X}_i)_{i \in \mathcal{I}}$，利得関数の順序組$(f_i)_{i \in \mathcal{I}}$の三つが，戦略形ゲームを構成する必要十分な要素である．これらの順序組$(\mathcal{I}, (\mathcal{X}_i)_{i \in \mathcal{I}}, (f_i)_{i \in \mathcal{I}})$を，戦略形ゲームと呼ぶ．

ゲーム理論特有の表記として，プレーヤi以外の戦略の順序組$(x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_I)$を$x_{-i}$と表記する．$x_{-i}$全体の集合は，直積$\mathcal{X}_{-i} := \Pi_{j \in \mathcal{I} \setminus \{i\}} \mathcal{X}_j$であり，$x_{-i} \in \mathcal{X}_{-i}$と書ける．ここで，記号$\setminus$は集合の差，すなわち二つの集合$\mathcal{A}$，$\mathcal{B}$に関して$\mathcal{A} \setminus \mathcal{B}$は$\mathcal{A}$から集合$\mathcal{B}$の要素を除外した$\mathcal{A} \setminus \mathcal{B} = \{x \mid x \in \mathcal{A} \wedge x \notin \mathcal{B}\}$である(16, #5.13)＊3．

定義2において最適応答BR_iを関数ではなく，対応と呼ぶのは，利得関数が最大となる戦略は一つとは限らないためである．そもそも，arg maxが対応である＊4．次のarg maxを使わない表現をすれば，関数ではなく対応であることが明示的になるだろう．

$$\mathrm{BR}_i(x_{-i}) := \{x_i^\star \in \mathcal{X}_i \mid f_i(x_i^\star, x_{-i}) \geq f_i(x_i, x_{-i}), \forall x_i \in \mathcal{X}_i\} \qquad (13)$$

6　他の数学との関連

本章では，ゲーム理論と，確率や線形代数，強化学習との代表的な関連を指摘する．

6.1　混合戦略と確率

戦略形ゲーム（式(2)）における戦略集合を，離散添字集合$\mathcal{X}_1 = \{1, \ldots, m\}$，$\mathcal{X}_2 = \{1, \ldots, n\}$とする．これまでのように戦略集合の中から一つの戦略を選ぶほかに，戦略を確率的に決定することも考えられる．これらを区別するために，前者を純戦略，後者を混合戦略と呼ぶ．プレーヤ1の混合戦略集合は，純戦略集合X_1上の確率分布の集合

$$\Delta(\mathcal{X}_i) := \left\{ (p_1, p_2, \ldots, p_m) \in \mathbb{R}_{\geq}^m \;\middle|\; \sum_{i=1}^{m} p_i = 1 \right\} \qquad (14)$$

である．$\mathbb{R}_{\geq}$は非負実数全体の集合を表す(15, §1.1.A)，(16, #6.4)．

＊2　文献(15, §1.5.B-C)では，元の順序組を一般化した概念として「添字付けられた元の族（indexed family of elements）」，集合の順序組を一般化した概念として「添字付けられた集合族（indexed family of sets）」が定義されているが，ゲーム理論の論文の多くを読むにあたっては，順序組で差し支えないと思われる．順序の定義については文献(15, §3.1.A)も参照されたい．

＊3　文献(15, §1.2.C)では，記号\の代わりに−が使われている．記号\は二項演算子であり，TEXでは\setminusを用いるべきである．\backslashでは他の二項演算子と同様の空白が入らない．

＊4　関数として扱われることもよくある．

表２　強化学習とゲーム

強化学習問題	戦略形ゲーム
行動価値関数 $Q : \mathcal{A} \times \mathcal{S} \to \mathbb{R}$	利得関数 $f_i : \mathcal{X}_i \times \mathcal{X}_{-i} \to \mathbb{R}$
グリーディ方策 $\arg\max_{\alpha} Q(\alpha, s)$	最適応答 $\arg\max_{x_i} f_i(x_i, x_{-i})$

6.2　線形代数・最適化理論

$a_{ij} := f_1(i, j)$ と書き換える．混合戦略（$p_1, \ldots, p_m$）$\in \Delta(\mathcal{X}_1)$，（$q_1, \ldots, q_m$）$\in \Delta(\mathcal{X}_2)$ に対し，プレーヤ１の利得の期待値は次式のとおり行列表現できる．

$$\begin{bmatrix} p_1 & \cdots & p_m \end{bmatrix} \begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mn} \end{bmatrix} \begin{bmatrix} q_1 \\ \vdots \\ q_n \end{bmatrix} \tag{15}$$

2. で述べたとおり，ゲーム理論は最適化理論と密接なつながりがある．また，線形代数，幾何学，線形計画法，ゲーム理論（ゼロサムゲーム．式（2）において $f_2(x_1, x_2) = -f_1(x_1, x_2)$ であるもの）のつながりに関しては文献（17，§8）に直観的な説明がある．

6.3　強化学習

機械学習の一分野である強化学習において，基本的な方法であるQ学習とゲームとの共通点・相違点を議論する．単一主体の強化学習問題[18]では，取り得る状態の集合$\mathcal{S}$と，取り得る行動の集合$\mathcal{A}$を既知として，観測された状態に対して各行動がどのような報酬をもたらすかという未知の行動価値関数 $Q : \mathcal{A} \times \mathcal{S} \to \mathbb{R}$ [*5] を学習する．

表２に強化学習問題と戦略形ゲームの対応を例示する．戦略形ゲームの利得関数と比較すると，行動集合$\mathcal{A}$が戦略集合$\mathcal{X}_i$に，状態集合$\mathcal{S}$が他プレーヤの戦略集合$\mathcal{X}_{-i}$に対応していると捉えられる．利得関数は既知と想定されるが，行動価値関数は未知であり，このために学習が必要である．

単一主体の強化学習問題は，マルコフ決定過程－現在の状態と選択された行動に対して次の状態が確率的に決まる状況－における学習問題である．一方，複数の主体が存在する場合のマルコフ決定過程を，マルコフゲームと呼ぶ．逆に言えば，単一主体のマルコフゲームがマルコフ決定過程である．マルコフゲームや，より一般の確率ゲームにおける均衡や学習アルゴリズム，無線通信への応用については，文献（19）にまとめられている．

＊5　通常は$\mathcal{S} \times \mathcal{A}$という順で書くが，戦略形ゲームとの対比を意図して$\mathcal{A} \times \mathcal{S}$という順にしている．

7　おわりに：ゲーム理論は役に立つのか

筆者がゲーム理論を無線通信に応用した発表をすると，講演後の立ち話や懇親会の場でよく受ける質問があった．「ゲーム理論は役に立つの？」という漠然とした質問である．本小特集の趣旨「社会で活躍する」≒「役に立つ」という関連があり，議論する．

この質問の明確な記述を試みると，二つのパターンがありそうである．まず思い至ったのは，「ゲーム理論は，最適化問題の解を求めるために使えるのか？」というものである．この場合は，「最適化問題の解を求めるには最適化理論を使って下さい．ゲーム理論はゲームという最適化問題とは違う問題設定をしたときに使えます．」が答えとなろう．

もう一つのパターンは「ゲーム理論は無線通信システムの課題を解決してくれるのか？」と漠然と聞いているものである．これについては，数学理論の間の関係が役に立つであろう．すなわち，**2.** で述べたとおりの最適化問題とゲームの関連を説明した後で，ゲーム理論を最適化理論に置き換えた「最適化理論は無線通信システムの課題を解決するのか？」，もっと具体的に言えば「無線通信システムの課題は最適化問題として定式化できるのか？」という逆質問をすればよさそうである．答えは，「定式化できる課題も，そうでない課題もある」というだけである．定式化できる場合には，更にその中に「解が求まる」場合があり，更にその中に「実用上意味のある解がある」≒「役に立つ」場合もあろう．一方，それ以外の「役に立たない」場合もあろう．いずれにせよ，解けないことが分かれば，それ以上無駄骨を折ることもなくなり，その意味で意味はあると思うが，論文にはならないことが多いであろう．

最近は「機械学習は役に立つの？」という素朴な疑問をよく聞く．今後も似たような疑問が出てくるだろう．機械学習も定義次第ではありそうだが，数学にすぎない範囲の機械学習に関しては同様に考えれば，答えは自分で見いだせるであろう．

謝辞　有益な意見を頂いた校閲御担当者，研究室学生諸君に感謝する．特に神矢翔太郎君，香田優介君からは，集合論に関する有益な意見を頂いた．図１は，神矢翔太郎君に提供頂いた．

■ 文献

（1）　新井紀子，“言語としての数学，”数理科学，no.575，pp.11-16，May 2011．
（2）　山本高至，“ゲーム理論の無線通信への応用，”信学誌，vol.95，no.12，pp.1089-1093，Dec. 2012．
（3）　Z.Han，D.Niyato，W.Saad，T.Başar，and A.

Hjorungnes, Game Theory in Wireless and Communication Networks: Theory, Models, and Applications, Cambridge Univ. Pr., 2011.

(4) 三瓶政一，阪口　啓（監），無線分散ネットワーク，電子情報通信学会，May 2011.

(5) D.Monderer and L.S.Shapley, "Potential games," Games Econ. Behav., vol.14, no.1, pp.124-143, May 1996, DOI:10.1006/game.1996.0044

(6) 宇井貴志，"ポテンシャルゲームと離散凹性，"第 17 回 RAMP シンポジウム論文集，pp.89-105，2005.

(7) S.Kamiya, K.Yamamoto, T.Nishio, M. Morikura, and T.Sugihara, "Spatial co-channel overlap mitigation through channel assignment in dense WLAN: Potential game approach," IEICE Trans.Commun., vol.E100-B, no.7, pp.1094-1104, July 2017,DOI:10.1587/transcom.2016SCP0013

(8) B.Yin, S.Kamiya, K.Yamamoto, T.Nishio, M. Morikura, and H.Abeysekera, "Mitigating throughput starvation in dense WLANs through potential game-based channel selection," IEICE Trans. Fundamentals, vol.E100-A, no.11, pp.2341-2350, Nov. 2017, DOI:10.1587/transfun.E100.A.2341

(9) K. Yamamoto, "A comprehensive survey of potential game approaches to wireless networks," IEICE Trans. Commun., vol.E98-B, no.9, pp.1804-1823, Sept. 2015, DOI:10.1587/transcom.E98.B.1804

(10) 山本高至，"無線通信とポテンシャルゲーム，"計測制御，vol.55，no.11，pp.996-1001，Nov. 2016.

(11) 岡田　章，ゲーム理論新版，有斐閣，2011.

(12) 瀧澤弘和，現代経済学 - ゲーム理論・行動経済学・制度論，中央公論新社，2018.

(13) 牧野邦昭，経済学者たちの日米開戦 - 秋丸機関「幻の報告書」の謎を解く，新潮社，2018.

(14) スティーブン・R・コヴィー，7 つの習慣，キングベアー出版，1996.

(15) 松坂和夫，集合・位相入門，岩波書店，1968.

(16) ISO 80000-2:2009 (E), "Quantities and units - part 2: Mathematical signs and symbols to be used in the natural sciences and technology," Feb. 2009.

(17) G.Strang, Linear Algebra and Its Applications, Academic Pr., 1976.

(18) C. Szepesvári, Algorithms for Reinforcement Leaning, Morgan and Claypool Pub., 2010.

(19) S.Lasaulce and H.Tembine, Game Theory and Learning for Wireless Networks: Fundamentals and Applications, Academic Pr., 2011.

山本高至（正員：シニア会員）

2005 京大大学院情報学研究科博士課程了．同年同研究科助手．2011 同准教授．2008-09 スウェーデン王立工科大客員研究員．博士（情報学，京大）．無線リソース制御，ゲーム理論，確率幾何解析に関する研究に従事．2011 本会論文賞，2012 IEEE 関西支部 Gold 賞，2016 本会末松安晴賞各受賞．IEEE 会員．