ゲーム理論

第2回 戦略形ゲーム

佐賀大学大学院 工学系研究科
知能情報システム学専攻
上田　俊
Email: sgrueda@cc.saga-u.ac.jp
https://sites.google.com/view/sgrueda/in-japanese

アウトライン

- 戦略形ゲーム
  - ゲームの要素
  - 支配戦略と支配戦略均衡
  - 囚人のジレンマ
- 最適反応とナッシュ均衡
- 混合戦略
  - 2人ゼロ和ゲームとミニマックス定理
  - ナッシュ均衡の計算方法

新聞社の競争

- ライバル関係にある2つの新聞社 (旭日新聞, 東都日報) が翌日の1面記事を経済記事にするか, スポーツ記事にするか悩んでいる.
- 80%の人は1面記事が経済ニュースなら買い, 20%の人はスポーツニュースなら買う.
- (新聞社目線) どの記事を1面に載せるべきか？
- (ゲーム理論目線) 翌朝の2つの新聞の一面記事は経済かスポーツか予想したい.

利得表

- 2人戦略形ゲームは利得表を用いて表現できる.

旭日 \ 東都	経済	スポーツ
経済	(40, 40)	(80, 20)
	[illegible]0)	(10, 10)

第1プレイヤーの戦略
(選択可能な行動)

第1プレイヤーの利得

第2プレイヤーの利得

戦略形ゲームの定義

- 戦略形ゲーム (game in strategic form)
  - $G = (N, \{S_i\}_{i \in N}, \{f_i\}_{i \in N})$

  - $N = \{1, \dots, n\}$: プレイヤーの集合
  - S_i はプレイヤー i の選択可能な行動あるいは戦略の集合
  - f_i は直積集合 $S = S_1 \times \cdots \times S_n$ 上の実数値関数であり，プレイヤー i の利得関数を表す．
- 標準形ゲーム (game in normal form) とも

ゲームの流れ

- すべてのプレイヤー $1, \dots, n$ は他のプレイヤーの選択を知らずにそれぞれの戦略 $s_1 \in S_1$, …, $s_n \in S_n$ を選択する.
- その結果, プレイヤー i は利得 $f_i(s_1, \dots, s_n)$ を得る.
- プレイヤーの目的は自己の利得の最大化である.
- ゲームのプレイにおいてゲームの各要素はすべてのプレイヤーの共有知識 (common knowledge) とする.

ゲームの分析

- 旭日新聞の立場に立って，どの戦略をとるべきか考える．
- 東都が経済 ⇒ 経済
- 東都がスポーツ ⇒ 経済
- つまり，東都がどちらの戦略を取っても経済ニュースを1面に掲載することが最適

旭日＼東都	経済	スポーツ
経済	(40, 40)	(80, 20)
スポーツ	(20, 80)	(10, 10)

支配戦略 (dominant strategy)

- 相手の取る戦略に関わらず，得られる利得が最大となる戦略のこと

- プレイヤー i の2つの戦略 s_i と t_i に対して，戦略 s_i が戦略 t_i を支配する (dominate) とは，他の $n-1$ 人のプレイヤーが持つすべての戦略の組 $s_{-i} \in S_i \times \cdots \times S_{i-1} \times S_{i+1} \times \cdots \times S_n$ に対して，$f_i(s_i, s_{-i}) > f_i(t_i, s_{-i})$ が成立することである.

支配戦略均衡

- すべてのプレイヤーが支配戦略を持つとき，その組合せを支配戦略均衡と呼ぶ．
- 常に存在するとは限らない．
- 人が遊んで面白いと思うゲームには，普通支配戦略はない．

旭日＼東都	経済	スポーツ
経済	$\boxed{(40, 40)}$	(80, 20)
スポーツ	(20, 80)	(10, 10)

支配戦略均衡

囚人のジレンマ (1/2)

- 重大な犯罪を犯した2人が個別に取り調べを受けている.
  - 証拠が不足しており，容疑者の自白がなければ逮捕できない.
  - 別件の軽微な犯罪の証拠は揃っている.
- 検察は自白が欲しいため，司法取引を持ち掛ける.
  - 両方が黙秘の場合，別件容疑だけのため，1年の懲役
  - 両方が自白した場合，両方に8年の懲役
  - 片方が黙秘，片方が自白の場合
    - 黙秘した方はすべての罪を被り10年の懲役
    - 自白した方は司法取引により3か月の拘留のみ

	黙秘	自白
黙秘	(1年, 1年)	(10年, 3ヵ月)
自白	(3ヵ月, 10年)	(8年, 8年)

囚人のジレンマ (2/2)

- (自白, 自白) の支配戦略均衡が存在する.
  - 相手が黙秘する場合, 1年 > 3ヵ月なので自白する.
  - 相手が自白する場合, 10年 > 8年なので自白する.
- 2人にとって, 最も良い結果は(黙秘, 黙秘)
  - パレート最適な結果という.
- なぜこのゲームが注目されているのか？
  - 社会状況における個人合理性(自分の利得の追及) ≠ 全体合理性 (全体の利得の追及)

	黙秘	自白
黙秘	(1年, 1年)	(10年, 3ヵ月)
自白	(3ヵ月, 10年)	(8年, 8年)

支配戦略均衡

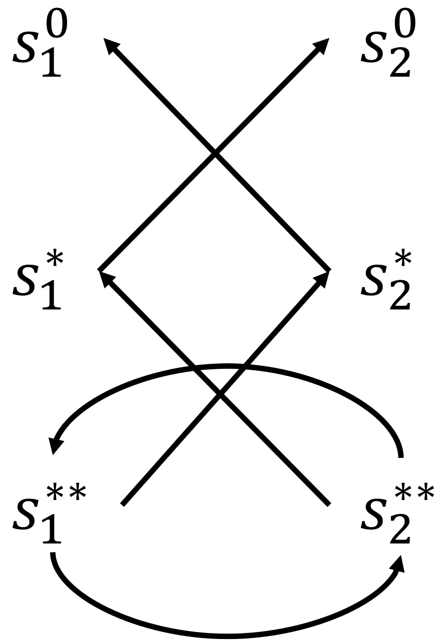
最適反応

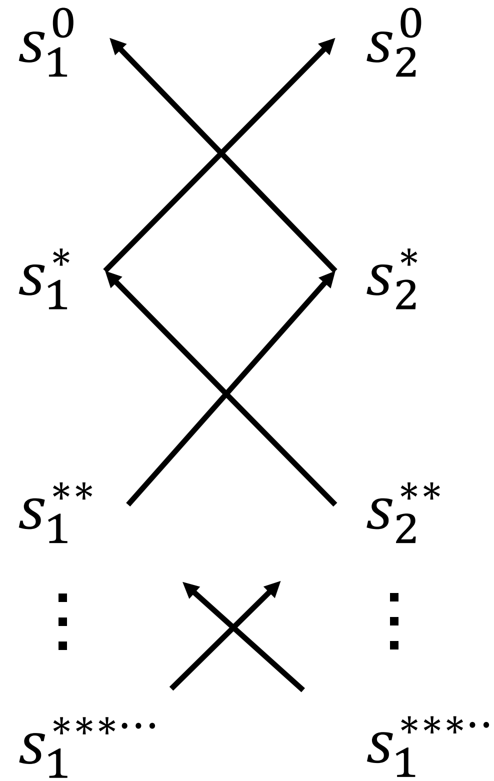
- プレイヤー i の戦略 $s_i \in S_i$ が他の $n-1$ 人のプレイヤーの戦略の組 $s_{-i} = (s_1, \cdots, s_{i-1}, s_{i+1},$

ナッシュ均衡

- 戦略形 n 人ゲーム G において，プレイヤーの戦略の組 s^* がナッシュ均衡点 (Nash equilibrium point) であるとは，すべてのプレイヤー $i(=1,\cdots,n)$ に対して戦略 s_i^* が他のプレイヤーの戦略の組 s_{-i}^* に対する最適反応であるときをいう．

推論と戦略決定の連鎖





推論が停止する.
(s_1^{**}, s_2^{**}) がナッシュ均衡点

推論が停止しない…

硬貨合わせゲーム

- 2人 (P1, P2) がそれぞれ硬貨の表か裏を選択する.
  - 違う面を選択したら, P1の勝ち. P2がP1に100円を支払う.
  - 同じ面を選択したら, P2の勝ち. P1がP2に100円を支払う.

P1 \ P2	表	裏
表	(-1, 1)	(1, -1)
裏	(1, -1)	(-1, 1)

混合戦略

- 確率的に行動を選択する戦略を混合戦略 (mixed strategy) と呼ぶ.
  - 行動 S_i 上の確率分布 q_i が戦略となる.
  - 利得の期待値の最大化を行う.
  - 最適反応，均衡点等は期待利得に関して同様に定義される.
- これまでのように確定的に行動を選択する戦略を純粋戦略 (pure strategy) と呼ぶ.

ゲームの混合拡大

- 戦略形ゲーム $G = (N, \{S_i\}_{i \in N}, \{f_i\}_{i \in N})$ の混合拡大 (mixed extension)
  - $G^* = (N, \{Q_i\}_{i \in N}, \{F_i\}_{i \in N})$
  - $N = \{1, \dots, n\}$: プレイヤーの集合
  - Q_i は S_i 上の確率分布の全体である. S_i 上の確率分布 q_i をプレイヤー i の混合戦略という.
  - F_i は直積集合 $Q = Q_1 \times \cdots \times Q_n$ 上の実数値関数で, 次のように定義される.

$$F_i(q_i, \cdots, q_n) = \sum_{s_1 \in S_1} \cdots \sum_{s_n \in S_n} \left\{ \prod_{j=1}^{n} q_j(s_j) \right\} f_i(s_1, \cdots, s_n)$$

  - ただし, $q_j(s_j)$ は混合戦略 q_j が純粋戦略 s_j に付与する確率を表す. $F_i(q_i, \cdots, q_n)$ をプレイヤー i の期待利得関数 (expected payoff function) という.

混合戦略の例

- P2が常に表を選択するとき，P1が表1/2, 裏1/2の混合戦略をとる.
  - P1の期待利得は$(-1) \times {}^{1}/_{2} + 1 \times {}^{1}/_{2} = 0$
- このゲームのナッシュ均衡は互いに1/2の確率で表・裏を選ぶ (混合) 戦略の組.
  - 純粋戦略同士の組ではナッシュ均衡は存在しない.

P1 \ P2	表	裏
表	(-1, 1)	(1, -1)
裏	(1, -1)	(-1, 1)

均衡点の存在

- 戦略形 n 人ゲーム $G^* = (N, \{Q_i\}_{i\in N}, \{F_i\}_{i\in N})$ において，混合戦略の範囲で少なくとも1つの均衡点が存在する．
  - 角谷の不動点定理 (Kakutani, 1941) を用いて証明できる．
- 2人ゲームのナッシュ均衡計算問題はPPAD完全である．(Chen and Deng, 2006)
  - PPAD完全な問題を解く多項式時間アルゴリズムは発見されていない．
  - ただし，2×2 (2人2行動) ゲームであれば容易に計算できる．

ゼロ和ゲーム

- すべてのプレイヤーの利得の和が常に0であるゲーム
  - $\sum_{i=1}^{n} f_i(s_1, \cdots, s_n) = 0$
- 2人ゼロ和ゲーム
  - (プレイヤー1の利得) = − (プレイヤー2の利得)
  - 硬貨合わせゲームも2人ゼロ和ゲーム
  - じゃんけんも2人ゼロ和ゲーム

マックスミニ戦略とミニマックス戦略

- $\min_{q_2 \in Q_2} F(q_1^*, q_2) = \max_{q_1 \in Q_1} \min_{q_2 \in Q_2} F(q_1, q_2)$ を満たす戦略 q_1^* をプレイヤー1のマックスミニ戦略 (maxmini strategy) と呼び，右辺の値をマックスミニ値という.
  - 最小の利得を最大化した戦略
- $\max_{q_1 \in Q_1} F(q_1, q_2^*) = \min_{q_2 \in Q_2} \max_{q_1 \in Q_1} F(q_1, q_2)$ を満たす戦略 q_2^* をプレイヤー2のミニマックス戦略 (minimax strategy) と呼び，右辺の値をミニマックス値という.

ミニマックス定理

- ゼロ和2人ゲームにおいて，以下が成り立つ (ミニマックス定理)：
$$\max_{q_1 \in Q_1} \min_{q_2 \in Q_2} F(q_1, q_2) = \min_{q_2 \in Q_2} \max_{q_1 \in Q_1} F(q_1, q_2)$$

- マックスミニ戦略とミニマックス戦略の組 (q_1^*, q_2^*) はゼロ和2人ゲームのナッシュ均衡点となっている．

精巧堂 vs. 便乗工房

- 右のゲームのナッシュ均衡を求める.
- 精巧堂の混合戦略
  - ゴジラ: q_1
  - モスラ: $1 - q_1$
- 便乗工房の混合戦略
  - ゴジラ: q_2
  - モスラ: $1 - q_2$

精巧堂 \ 便乗工房	ゴジラ	モスラ
ゴジラ	(120, 120)	(216, 24)
モスラ	(192, 48)	(96, 96)

精巧堂の期待利得

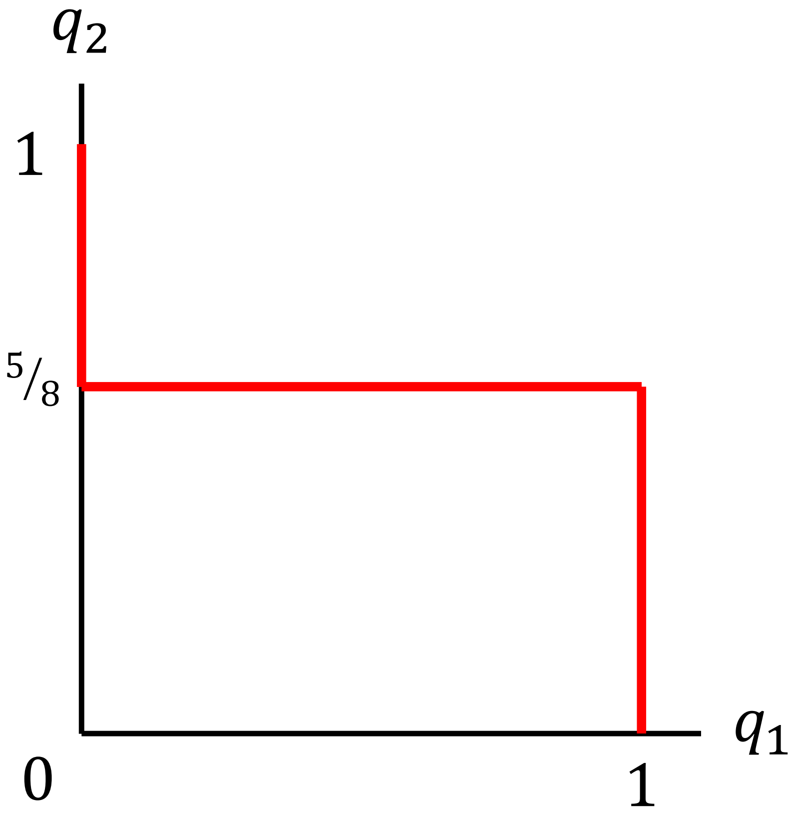
- 精巧堂の期待利得を求める.
- ゴジラを選択した場合
  - $120 \times q_2 + 216 \times (1 - q_2) = -96q_2 + 216$
- モスラを選択した場合
  - $192 \times q_2 + 96 \times (1 - q_2) = 96q_2 + 96$

精巧堂＼便乗工房	ゴジラ q_2	モスラ $1 - q_2$
ゴジラ q_1	(120, 120)	(216, 24)
モスラ $1 - q_1$	(192, 48)	(96, 96)

精巧堂の最適反応グラフ

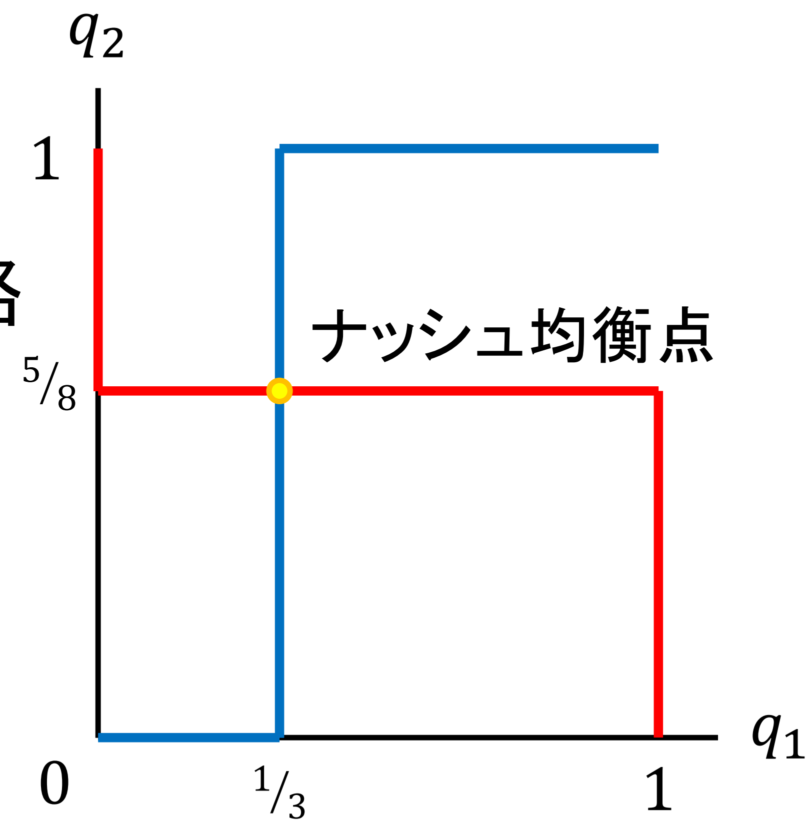
- ゴジラ: $-96q_2 + 216$
- モスラ: $96q_2 + 96$

- 精巧堂の最適反応戦略
  - $q_2 < {}^5/_8$ のとき, $q_1 = 1$
  - $q_2 = {}^5/_8$ のとき, 任意のq_1
  - $q_2 > {}^5/_8$ のとき, $q_1 = 0$



便乗工房の最適反応グラフ

- ゴジラ: $72q_1 + 48$
- モスラ: $-72q_1 + 96$
- 便乗工房の最適反応戦略
  - $q_1 < {}^{1}/_{3}$ のとき，$q_2 = 0$
  - $q_1 = {}^{1}/_{3}$ のとき，任意のq_2
  - $q_1 > {}^{1}/_{3}$ のとき，$q_2 = 1$
- 交点がナッシュ均衡



まとめ

- 戦略形ゲーム
- 支配戦略
  - 相手の取る戦略に関わらず，得られる利得が最大となる戦略
  - その戦略の組による均衡を支配戦略均衡と呼ぶ.
- ナッシュ均衡
  - 互いに最適反応になっている戦略の組
  - (混合戦略まで拡張した場合) すべてのゲームにナッシュ均衡点が少なくとも1つ存在する.