

マルチエージェント実験経済学の最近の動向

公立はこだて未来大学
川越 敏司

実験経済学におけるシミュレーション研究

研究の方向性の変化
 定性的記述から定量的記述へ
 法則性発見からデータ検証へ
 学習理論のモデル化
 完全合理性から限定合理性までを含む一般的な学習モデルの構築
 強化学習とbelief-based学習の統合化
 推論の深さの違いや先読みを考慮

法則性発見からデータ検証へ

法則性発見型シミュレーション
 モデルのパラメータを様々に変更して創発特性を探る
 トーナメント形式による優位な戦略の発見
 データ検証型シミュレーション
 最尤法によるモデルのパラメータ推測
 ブートストラップ・シミュレーションによるモデル間の予測精度比較

ゲーム理論実験における数々のアノマリー

交渉ゲームにおける公平分割への逸脱
 公共財供給における過剰供給
 オークションにおけるビッド上昇バイアス
 混合戦略からの逸脱
 利得の変換に関する不変性の逸脱

利得の変換に関する不変性

利得の正1次変換に関してナッシュ均衡は不変である
 しかし、実験では戦略的状況を変えない利得の絶対的大きさの変更が被験者の行動に影響を与える

コイン合わせゲーム(1)

	L	R	
T	9, 0	0, 1	ゲーム1
B	0, 1	1, 0	

	L	R	
T	9, 0	0, 4	ゲーム2
B	0, 4	1, 0	

	L	R	
T	36, 0	0, 4	ゲーム3
B	0, 4	4, 0	

コイン合わせゲーム(2)

T	0.500
L	0.100

ナッシュ均衡

T	0.462
L	0.296

ゲーム1

T	0.459
L	0.275

ゲーム2

T	0.404
L	0.321

ゲーム3

コイン合わせゲーム(3)

ゲーム3において、次のような先読みプロセスが見られた

- まず1はTを高い頻度で選ぶ
 - はそれに対しRを選んで対抗する
 - そこで1はTを選ぶことが危険と考えBを選ぶようになる
 - それで2はLを選んで対抗する
- 先読みを考慮する必要

旅行者のジレンマ(1)

人が同時に $[2, 10]$ の価格 P, Q を選択
 ならば、ともに P の利益
 ならば、それぞれ $P+R, P-R$ の利益
 支配される戦略の逐次消去によって得られる唯一のナッシュ均衡は $P=Q=2$ となる
 および $R=5$ の場合を比較実験

旅行者のジレンマ(2)

	8	9	10
8	8, 8	10, 6	10, 6
9	6, 10	9, 9	11, 7
10	6, 10	7, 11	10, 10

R=2の場合

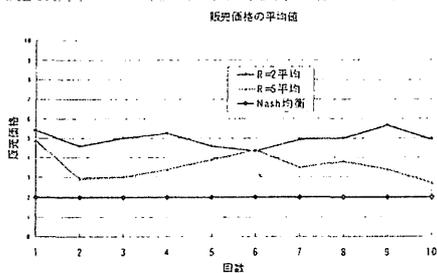
戦略9は10を支配

	8	9
8	8, 8	10, 6
9	6, 10	9, 9

戦略8は9を支配

	8
8	8, 8

旅行者のジレンマ(3)



旅行者のジレンマ(4)

ナッシュ均衡
 支配される戦略を8回消去
 段階の思考ステップが必要
 と $R=5$ とで平均価格に有意な差
 思考ステップの差異を考慮する必要

学習に関するゲーム理論

強化学習

例: レプリケータ・ダイナミクス
学習

例: クールノー学習, Fictitious Play 学習
強化学習と Belief-Based 学習の統合化

完全合理性と限定合理性を含む一般モデル

13

強化学習(1)

プレイヤー i の戦略 j : S_i^j
 t 期にプレイヤー i が選んだ戦略: $S_i(t)$
 プレイヤー i の利得: $u_i(S_i^j, S_{-i})$
 t 期のプレイヤー i の戦略 j の強化子: $R_i^j(t)$
 更新ルール:

$$R_i^j(t) = \begin{cases} \phi \cdot R_i^j(t-1) + u_i(S_i^j, S_{-i}(t)) & S_i^j = S_i(t) \\ \phi \cdot R_i^j(t-1) & S_i^j \neq S_i(t) \end{cases}$$

14

強化学習(2)

線形確率ルール

$$P_i^j(t+1) = \frac{R_i^j(t)}{\sum_k R_i^k(t)}$$

ロジット確率ルール

$$P_i^j(t+1) = \frac{e^{\lambda \cdot R_i^j(t)}}{\sum_k e^{\lambda \cdot R_i^k(t)}}$$

15

Fictitious Play 学習(1)

プレイヤー i 以外の戦略 k : S_{-i}^k
 戦略 S_{-i}^k の度数更新ルール:

$$F_{-i}^k(t) = F_{-i}^k(t-1) + \begin{cases} 1 & S_{-i}(t) = S_{-i}^k \\ 0 & S_{-i}(t) \neq S_{-i}^k \end{cases}$$

t 期の戦略 S_{-i}^k の頻度に関する信念:

$$B_{-i}^k(t) = \frac{F_{-i}^k(t)}{\sum_j F_{-i}^j(t)}$$

16

Fictitious Play 学習(2)

t 期の i の戦略 j の期待効用: $E_i^j(t)$

$$E_i^j(t) = \sum_k u_i(S_i^j, S_{-i}^k) \cdot B_{-i}^k(t)$$

ロジット確率ルール

$$P_i^j(t+1) = \frac{e^{\lambda \cdot E_i^j(t)}}{\sum_k e^{\lambda \cdot E_i^k(t)}}$$

17

Quantal Response Equilibria

Caltech の McKelvey and Palfrey (1995) の研究

プレイヤー i が戦略 j を選ぶ確率: P_i^j
 i が戦略 j を選ぶときの期待効用: π_i^j
 次式の不動点を QRE という。

$$P_i^j = \frac{e^{\lambda \cdot \pi_i^j}}{\sum_k e^{\lambda \cdot \pi_i^k}}$$

18

QREの性質

パラメータ λ は合理性の程度を表す
 $\lambda = 0$ ならばランダム選択
 $\lambda \rightarrow \infty$ ならばナッシュ均衡
 は必ず存在する
 によって複数均衡から均衡選択可能
 静的な均衡モデルだが説明力が高い
 動的な学習モデルへの拡張も可能

19

EWA Learning(1)

のColin Camerer(1999)による研究
 3つのパラメータ
 誘因に関する成長率: κ
 誘引に関する割引率: ϕ
 実現利得と仮想利得との重み付け: δ
 強化学習とBelief-Based学習の統合化
 強化学習: $\delta = 0, \kappa = 1$
 - Fictitious Play学習: $\delta = 1, \kappa = 0, \phi = 1$

20

EWA Learning(2)

誘因更新ルール

$$A_i^j(t) = \frac{\phi \cdot N(t-1) \cdot A_i^j(t-1)}{N(t)} + \frac{[\delta + (1-\delta) \cdot I(s_i^j, s_i(t))] \cdot u_i(s_i^j, s_i(t))}{N(t)}$$

21

EWA Learning(3)

経験に関する重み付け更新ルール

$$N(t) = (1-\kappa) \cdot \phi \cdot N(t-1) + 1$$

ロジット確率ルール

$$P_i^j(t+1) = \frac{e^{\lambda \cdot A_i^j(t)}}{\sum_k e^{\lambda \cdot A_i^k(t)}}$$

22

EWAにおける思考ステップ

ステップ思考者は、他者が0からK-1ステップの思考をすると考える

ステップ思考者の分布

$$f(K) = \frac{e^{-\tau K}}{K!}$$

ステップ思考者の誘因

$$A_i^j(K) = \sum_h u_i(s_i^j, s_{-i}^h) \cdot \left\{ \frac{\sum_{c=0}^{K-1} f(c)}{\sum_{d=0}^{K-1} f(d)} \cdot P_{-i}^h(c) \right\}$$

EWAにおける先読み(1)

先読みプレイヤー: s

- 適応的プレイヤー: a
 - 社会にはsが α 、aが $1-\alpha$ の割合で存在
 は、 α に関する予測値 α をもとに行動
- $\alpha < \alpha$
 先読みプレイヤーの自信過剰
- $\alpha > \alpha$
 先読みプレイヤーの過大評価

24

EWAにおける先読み(2)

α はEWAに従い学習する
の学習ルール

$$A_i^j(s, t) = \sum_k B_{-i}^k(t+1) \cdot u_i(s_i^j, s_{-i}^k)$$

$$B_{-i}^k(t+1) = \alpha' P_{-i}^k(s, t+1) + (1-\alpha') P_{-i}^k(a, t+1)$$

$$P_i^j(t+1) = \frac{e^{\lambda \cdot A_i^j(s, t)}}{\sum_k e^{\lambda \cdot A_i^k(s, t)}}$$

25

学習理論における課題

時間スケール

人間の被験者が経験できる時間スケールで
学習できるモデルの構築

パラメータの不変性

同じゲームに対してデータによって異なるパ
ラメータの推定値を生み出さない

なるべくパラメータの少ないモデルを構築

思考ステップや先読みの考慮

26

参考文献

“
” *Japanese Economic Review*, 47,
186-209

- C. F. Camerer et al. (2001) “Behavioral
Game Theory: Thinking, Learning, and
Teaching”, mimeo

27