ゲームの逆解析：行動データに基づく利得の推定法

中国地方建設局　○正会員　井上慎也　　鳥取大学工学部　正会員　喜多秀行
鳥取大学工学部　正会員　福山敬　　鳥取大学工学部　正会員　谷本圭志

1. 緒言

個々人の意思決定が相互に依存する現象の分析にはゲーム理論が有用であるが，なかには利得が不明な現象も存在する．このとき，行動結果から利得を推定する方法があれば分析を行うことが可能であるが，著者の知る限りこのような方法は存在しない．そこで，本研究ではゲーム的状況を対象とし，行動結果や行動に影響する要因から利得を推定する方法論の構築を行う．

2. 基本モデルの構築

2.1 対象とするゲーム

利得を推定する方法論の構築にあたり，まずそれらの基本となるゲームを定義する．過度の複雑さを避け議論の見通しをよくするため幾つかの条件を導入する．行動主体であるプレイヤーは$k(=1,\cdots,N)$人存在し，各プレイヤーはゲーム的状況に直面したとき，それぞれいくつかの戦略$S_i^k \in \boldsymbol{S}^k (i=1,\cdots,I)$をとることができ，その中から自分の利得$U_i^k \in \boldsymbol{U}^k$が最大となる最適戦略$S^{k^*}$をとる．各プレイヤーの利得$U_i^k$は他プレイヤーの戦略$\boldsymbol{S}^{-k}(=\{\boldsymbol{S}^1,\cdots,\boldsymbol{S}^{k-1},\boldsymbol{S}^{k+1},\cdots,\boldsymbol{S}^N\})$との組み合わせで求まり，$U_i^k=U_i^k(S_i^k,\boldsymbol{S}^{-k})$で表す．また，各プレイヤーはお互いの行動に関して，事前に打ち合わせをしたり，取り決めを行うことはできない．各プレイヤーの意思決定は同時であると考える．各プレイヤーはそれぞれ，相手プレイヤー，相手の行動規範，相手の利得，相手がおかれている環境を知っている．

2.2 利得の推定法

全てのプレイヤーが最適戦略S^{k^*}をとり，かつ各プレイヤーのとる戦略が互いに相手の戦略の最適戦略となるとき，最適戦略の組を均衡解$\boldsymbol{E}$という．このときの均衡解$\boldsymbol{E}$を以下のように表す．

$$\begin{aligned}\boldsymbol{E} &= \boldsymbol{E}(\boldsymbol{U}) \\ &= \left\{S^{1^*},S^{2^*},\cdots,S^{N^*}\right\} \end{aligned} \tag{1}$$

各プレイヤーは均衡解$\boldsymbol{E}$によって決められた行動$\boldsymbol{A}=\{\boldsymbol{A}^k\}$をとる．

$$\boldsymbol{A}^k=\boldsymbol{A}^k(\boldsymbol{E}(\boldsymbol{U})) \tag{2}$$

利得$\boldsymbol{U}$のうち，分析者にとって観測可能な要因による確定項を$\boldsymbol{V}$，観測不可能な要因に起因する誤差項を$\boldsymbol{\epsilon}$とし，その線形性を仮定すると，

$$\boldsymbol{U}=\boldsymbol{V}+\boldsymbol{\epsilon} \tag{3}$$

で表すことができる．確定項$\boldsymbol{V}$は，パラメータ$\boldsymbol{\alpha}$と行動に影響する要因$\boldsymbol{X}$の線形結合として以下のように定義する．

$$\boldsymbol{V} = \boldsymbol{V}(\boldsymbol{\alpha},\boldsymbol{X})$$

本研究では，この確定項のパラメータ$\boldsymbol{\alpha}$を推定することを試みる．まず，各プレイヤーの利得の確定項が$\boldsymbol{V}$という条件の下で行動結果の組$A_i=\{A_i^1,\boldsymbol{A}^{-1}\}(\boldsymbol{A}^{-1}=\{\boldsymbol{A}^2,\cdots,\boldsymbol{A}^N\})$が生起する確率$p_i$を以下のように表す．

$$\begin{aligned} p_i &= prob(A_i|\boldsymbol{V}) \\ &= prob(A_i|\boldsymbol{V}(\boldsymbol{\alpha},\boldsymbol{X})) \end{aligned} \tag{4}$$

観測されるデータは，各プレイヤーの行動結果の組A_iと行動に影響する要因$\boldsymbol{X}$である．観測データ数をM個とすると，行動結果の組A_iが$m(=1,\cdots,M)$回同時に起こる同時確率，すなわち，尤度Lは以下の尤度関数で与えられる．

$$L=\prod_{i\in I}\prod_{m\in M}p_i^m \tag{5}$$

尤度関数を最大とするパラメータ$\hat{\boldsymbol{\alpha}}$は次式で与えられる．

$$\hat{\boldsymbol{\alpha}}=\arg\{\max_{\boldsymbol{\alpha}} L\} \tag{6}$$

推定されたパラメータの検定には尤度比ρ^2を用いる．

3. ゲームモデル

3.1 モデルの定式化

以下では，簡単のため2人ゲームを考える．P1,P2をプレイヤーとし，P1の行動は$A_1,A_2(\in \boldsymbol{A}^1)$であり，行動$A_1$をとる確率を$p$とする．一方，P2の行動を$A_3,A_4(\in \boldsymbol{A}^2)$とし，行動$A_3$をとる確率を$q$とする．$p$および$q$は混合戦略まで許容したものであり，0以上1以下の値をとる．各プレイヤーは得られる利得に対して期待利得でその結果を評価する．このとき，利得行列を図-1のように表す．ここで，P1,P2の利得をU_{ij}^kで表す．U_{ij}^kのiはP1が行動A_1をとるとき1，行動A_2をとるとき2，jはP2が行動A_3をとるとき1，行動A_4をとるとき2，kはP1のとき1，P2のとき2である．

P2

		A_3 (q)	A_4 (1-q)
P1	A_1 (p)	(U^1_{11}, U^2_{11})	(U^1_{12}, U^2_{12})
	A_2 (1-p)	(U^1_{21}, U^2_{21})	(U^1_{22}, U^2_{22})

図-1：利得行列

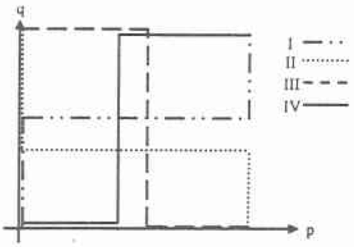

図-2：各プレイヤーの各最適反応曲線

	q_1 $(\bar{q} < 0)$	q_2 $(0 < \bar{q} < 1)$	q_3 $(\bar{q} > 1)$
p_1 $(\bar{p} < 0)$	(1,0)	(0,0)	(0,0)
p_2 $(0 < \bar{p} < 1)$	(1,0)	$(\bar{p},\bar{q})$	(0,1)
p_3 $(\bar{p} > 1)$	(1,1)	(1,1)	(0,1)

図-3：卍型の行動結果

3.2 行動結果の導出

各プレイヤーの期待利得を定式化し，各プレイヤーの期待利得の最大化よりP1,P2の最適反応を示し，行動結果を導出する．P1の期待利得は以下の式で求めることができる．

$$\begin{aligned} U(p,q) &= p\left\{qU^1_{11} + (1-q)U^1_{12}\right\} \\ &+ (1-p)\left\{qU^1_{21} + (1-q)U^1_{22}\right\} \end{aligned} \tag{7}$$

P1の最適反応を求める為，P1の期待利得式(7)を自身の行動で偏微分することにより求めることができる．

$$\hat{q} = q(U^1_{11} - U^1_{21} - U^1_{12} + U^1_{22}) + (U^1_{12} - U^1_{22}) \tag{8}$$

式(8)が正のときP1はA_1をとり，負のときは戦略A_2をとる．ここで，式(8)がゼロとなるときのqを$\bar{q}$とすると，

$$\bar{q} = \frac{U^1_{22} - U^1_{12}}{(U^1_{22} - U^1_{12}) + (U^1_{11} - U^1_{21})} \tag{9}$$

となる．このとき，$\bar{q}$の分母の正負の判定により2つの最適反応曲線(図-2中のI,II)が書ける．P2についても同様に行う(図-2中のIII,IV)．各プレイヤーの最適反応曲線をそれぞれ組み合わせることで，4種類の利得構造がでてくる(図-2)．例えば図-2のIとIIIの組み合わせが卍型の利得構造を示す．各利得構造が生起する確率を$\xi^i(\alpha,x)$とおく．各利得構造において，$\bar{p}$,$\bar{q}$を0以下，0以上1以下，1以上と変化させることで，各利得構造における行動結果を導出できる．ここで，各$\bar{p},\bar{q}$の状態が生起する確率を$p_1 \sim p_3$,$q_1 \sim q_3$とする．卍型における行動結果を図-3，行動結果の生起確率を図-4に示す．従って，行動の組の生起確率$\psi(\alpha,x)$は以下の式で与えられる．

$$\begin{aligned} \psi(\alpha,x) &= \sum_i(\text{各利得構造の生起確率}\xi^i(\alpha,x)) \\ &\times\ (\text{各利得構造における行動の組の生起確率}) \end{aligned} \tag{10}$$

$p_3q_1 + p_3q_2$ $+ p_2q_2 \times \bar{p}\bar{q}$	$p_1q_1 + p_2q_1$ $+ p_2q_2 \times \bar{p}(1-\bar{q})$
$p_2q_3 + p_3q_3$ $+ p_2q_2 \times (1-\bar{p})\bar{q}$	$p_1q_2 + p_1q_2$ $+ p_2q_2 \times (1-\bar{p})(1-\bar{q})$

図-4：卍型の行動結果の生起確率

P2 (B地域の住民)

P1 (A地域の住民)	地域Bで会合を行う	地域Aで会合を行う
地域Aで会合を行う	(ϵ_1, ϵ_5)	$(E_A+\epsilon_2, E_A-C_{BA}+\epsilon_6)$
地域Bで会合を行う	$(E_B-C_{AB}+\epsilon_3, E_B+\epsilon_7)$	$(-R-C_{AB}+\epsilon_4, -R-C_{BA}+\epsilon_8)$

$\epsilon_1 \sim \epsilon_8$　：観測者に未知の誤差項
E_A, E_B　：地域AまたはBで会合が行われたときの利得
C_{AB}　：$P1$が地域Bに行ったときの移動コスト
C_{BA}　：$P2$が地域Aに行ったときの移動コスト
$-R$　：すれ違いが生じることの不効用(禁止的に大きな値)

図-5：数値計算における利得行列

表-1 推定結果

	x_A	x_B	C_{AB}	C_{BA}	α_1	$\hat{\alpha}_1$	ρ^2
ケース1	[0,10]	[0,10]	[-20,0]	[-20,0]	5.0	4.816	0.988
ケース2	[0,20]	[0,20]	[-30,0]	[-30,0]	3.0	2.896	0.992
ケース3	[0,90]	[0,90]	[-80,0]	[-80,0]	8.0	8.210	0.999

4. 数値計算

上記モデルの有用性を吟味するために数値計算を行った．利得行列を図-5のように設定する．このとき，行動結果は$(A,B),(A,A),(B,B)$と複数行動$(A,A)(B,B)$である．また，複数行動$(A,A)(B,B)$が選ばれたとき，各総利得の比較を行い，総利得の大きい方の地域で会合が行われ，このプレイヤーの行動を分析者も知っている．ここで，$E_A = \alpha_1 x_A, E_B = \alpha_1 x_B, x_A$は地域$A$の会場数,$x_B$は地域$B$の会場数として，$\alpha_1$を推定する．計算手順は，観測データを作るために全ての変数に値を与え，そのときの利得の大小関係から行動結果を導出する．具体的には表-1のように設定した．いずれのケースにおいても推定値$\hat{\alpha}_1$は良い値となる．

5. 結言

本研究では，行動結果と行動に影響する要因から利得を推定する方法論を構築し，2×2ゲームにおける利得推定法を提案した．今後の課題はn人ゲームへの拡張である．