状態の推移を伴う提携行動に関する考察

中国地方整備局　正会員　○石本裕亮
鳥取大学工学部　正会員　谷本圭志
鳥取大学工学部　正会員　喜多秀行

1．はじめに

我々が直面する様々な社会問題を解決するに当たって，主体間の提携が不可欠となっている．提携形成メカニズムは従来「提携形成問題」として主にゲーム理論を用いて分析されてきた．そこでは，ゲームの利得や費用が主体の選択する行動には影響されない静的な状況を仮定してモデル化している．しかし，主体の行動がゲームの環境（状態）を変える場合がある．例えば，環境問題はその現形である．この場合，主体の行動と状態の間には相互依存の関係がある．

本研究では，この関係を明示した「提携形成問題」を二人確率ゲームを用いてモデル化するとともに，利得構造が提携形成行動に及ぼす影響について分析する．

2．モデル分析

2.1　二人確率ゲーム

プレイヤー（主体）の集合を $N=\{1,2\}$ で表し，任意のプレイヤーを $i(\in N)$ で表す．二人確率ゲームは要素 $\{S,\Theta_1,\Theta_2,q,\pi_1,\pi_2,\beta\}$ で定義される．ゲームは無限回繰り返され，その過程においていくつかの異なる状態 $s(\in S)$ が生起する．

プレイヤーが行動を選択することにより次の二点が起こる．まず，プレイヤー i は形成された提携の下で事業を行い，それにより瞬間の利得（instantaneous payoff）$\pi_i(s,\theta_1,\theta_2)$ を得る．ただし，$\theta_i(\in\Theta_i)$ はプレイヤー i の行動である．次いで，確率過程 $q(\bullet|s,\theta_1,\theta_2)$ に従って次期の状態 s' へと推移する．推移確率は今期の状態とゲームの結果，つまりプレイヤーの行動の組の条件付確率で与えられる．これらのプロセスが無限回続くため，プレイヤーは瞬間の利得ではなく，総期待割引利得を最大化する．

以下では，状態が二つ（$S=\{1,2\}$）であるゲームを対象にする．プレイヤーは形成された提携の下である事業を実施することで利得を得る．プレイヤーが利用可能な事業の方式（以後，「方式」と呼ぶ）の集合を $K=\{X,Y\}$ とする．プレイヤーの行動は「方式 X で相手と提携を形成する（X）」，「方式 Y で相手と提携を形成する（Y）」のいずれかを表明することとする．すなわち，$\Theta_i=K(\forall i\in N)$ である．プレイヤーが同じ方式 $k(\in K)$ を選択したとき方式 k の下で提携が形成され，異なる場合，単独提携が形成される．状態 s の下でのゲームにおいて方式 X で提携が形成された場合にプレイヤー1が得られる瞬間の利得を a_s, 方式 Y で提携が形成された場合のそれを c_s, 単独提携が形成された場合のそれを b_s とする．$p_{ss'},q_{ss'},r_{ss'}$ は状態の推移確率である．また $g_1(s)$ は初期の状態が s のときにプレイヤー1がゲームを無限回プレイすることによって得る総期待割引利得である．すると，ゲームの標準形表現は表1のようになる．なお，β は割引因子である．

表1　状態 s の下でのゲームの利得行列
（プレイヤー1の利得のみを示している）

		プレイヤー2	
		X	Y
プレイヤー1	X	$a_s+\beta\sum_{s'=1}^{2}q_{ss'}\cdot g_1(s')$	$b_s+\beta\sum_{s'=1}^{2}p_{ss'}\cdot g_1(s')$
	Y	$b_s+\beta\sum_{s'=1}^{2}p_{ss'}\cdot g_1(s')$	$c_s+\beta\sum_{s'=1}^{2}r_{ss'}\cdot g_1(s')$

2.2　ナッシュ均衡解

各プレイヤーの戦略は，各状態での各行動に割り当てる確率である．総期待割引利得が次式を満たす場合，プレイヤー1の戦略 Σ_1^* は最適応答である．

$$I_1\left(\Sigma_1^*,\Sigma_2\right)(s)\geq I_1\left(\Sigma_1,\Sigma_2\right)(s)\ \left(\forall\Sigma_1,\forall s\in S\right)\quad(2.1)$$

ただし，I_i は初期の状態が s と観測され，その後無限回繰り返す過程において得られるプレイヤー i の総期待割引利得である．全てのプレイヤーの戦略が

表２　(X, X)が最適応答であるための条件式

初期の状態が1のとき	条件式１	$\frac{1-q_{22}}{D_1}\cdot a_1+\frac{1}{D_1}\cdot q_{12}\cdot a_2$	≧	$\frac{1}{D_2}\cdot p_{12}\cdot a_2+\frac{1-q_{22}}{D_2}\cdot b_1$
	条件式２	$\frac{1-q_{22}}{D_1}\cdot a_1+\frac{1}{D_1}\cdot q_{12}\cdot a_2$	≧	$\frac{1-p_{22}}{D_3}\cdot a_1+\frac{1}{D_3}\cdot q_{12}\cdot b_2$
	条件式３	$\frac{1-q_{22}}{D_1}\cdot a_1+\frac{1}{D_1}\cdot q_{12}\cdot a_2$	≧	$\frac{1-p_{22}}{D_4}\cdot b_1+\frac{1}{D_4}\cdot p_{12}\cdot b_2$
参照		$D_1=(1-q_{11})(1-q_{22})-q_{12}\cdot q_{21}$		
		$D_2=(1-p_{11})(1-q_{22})-p_{12}\cdot q_{21}$		
		$D_3=(1-q_{11})(1-p_{22})-q_{12}\cdot p_{21}$		
		$D_4=(1-p_{11})(1-p_{22})-p_{12}\cdot p_{21}$		

＊$\beta q_{ss'}$を$q_{ss'}$と表している．他も同様．

他のプレイヤーの戦略の組に対する最適応答であるとき，この最適戦略の組がナッシュ均衡解である．

2.3　最適応答の条件の導出

状態１，２の下であるプレイヤーの行動がそれぞれk_1，k_2であることをベクトル（k_1，k_2）で表す．ここでプレイヤーが純粋戦略をとると仮定し，状態１においてプレイヤー2 の戦略が（X, X）であるとの条件の下で，プレイヤー1 の戦略（X, X）が最適応答となるための条件を表２に整理した．表２の条件式１はプレイヤー1 が戦略（X, X）をとった場合に得られる総期待割引利得が戦略（Y, X）とった場合に得られる総期待割引利得を上回っていることを示している．条件２は戦略（X, Y）をとった場合，条件式３は戦略（Y, Y）をとった場合に得られる総期待割引利得との大小関係を示している．なお，D_1～D_4は推移確率により構成される行列式であるが，正であることに留意を要する．状態２における条件式やその他の戦略の組についても同様に条件を求めることができる．このように，推移確率の条件によって均衡解が決定されることが分かる．

3.　数値実験

各結果においてプレイヤーが得る瞬間の利得の大小関係を「利得構造」と呼ぶ．利得構造を特徴づけるパラメータをδで表す．表３は実験において設定している利得構造を示している．表４は，δが 0.0 のケースで，プレイヤーが短期的な利得(瞬間の利得)の最大化を目的とするとき，（X, X）と（Y, Y）が均衡上の行動であり，長期的な場合，（X, X）のみがそれであることを示している．これは主体が短期的な視点しか持ち合わせていない場合の行動は長期的視点からは必ずしも合理的ではないことを示している．具体的に「どのような場合にそのようなことが生じるか」は表２に示す解析的な検討結果によって判定可能である．

表３　シミュレーションで用いる利得構造

状態１

		プレイヤー2 X	プレイヤー2 Y
プレイヤー1	X	1−δ, 2	0, 0
	Y	0, 0	2−δ, 1

状態２

		プレイヤー2 X	プレイヤー2 Y
プレイヤー1	X	1, 2	0, 0
	Y	0, 0	−1, 1

表４　数値実験の結果(プレイヤー1 の戦略)

δ		0.0	0.8	1.6	2.4
短期	状態１	(X, X) (Y, Y)	(X, X) (Y, Y)	(Y, Y)	－
	状態２	(X, X)	(X, X)	(X, X)	－
長期	状態１	(X, X)	(Y, Y)	(Y, Y)	－
	状態２	(X, X)	(X, X)	(X, X)	－

4.　今後の課題

状態やプレイヤーの行動の数に関する一般化が今後の課題である．