コーディネーションを望むプレイヤーによる繰返しゲーム

—非定常マルコフチェーンによるモデル化—

01402911 小樽商科大学　行方 常幸　NAMEKATA Tsuneyuki

1. はじめに

ゲーム理論において登場するプレイヤーは合理的(fully rational)に行動すると仮定されている場合と、部分的に合理的(boundedly rational)であると仮定されている場合がある。また、完全には合理的ではない後者のプレイヤーが繰返しゲームを行う場合、各期において、相手のプレイヤーが今までに取った手の相対頻度から相手の手を予想し、その期の自分の期待利得を最大にするように自分の手を決定すると仮定される場合が多い[1]。

本報告では完全には合理的ではないプレイヤーが各期に自分の手を決める時に、その期の利得の最大化を試みるのではなく、自分と相手の手のコーディネーションを望むと仮定する。このようなコーディネーションを望むプレイヤーからなる無限母集団を考え、2 つの手を持つステージゲームを繰り返すと想定し、この繰返しゲームを非定常マルコフチェーンによりモデル化する。

2. モデル

例として表 1で与えられた囚人のジレンマゲームを繰返す、次のようなプレイヤーを想定してみよう。(C,C)は両プレイヤーにとって最も望ましい。しかし、D は C を支配しているので両プレイヤーとも D を取ろうとする動機が存在する。(C,D)や(D,C)は片方のプレイヤーに得で、もう一方のプレイヤーに損だから、以後コンフリクトが増大するから最も望ましくない。(D,D)は Nash 均衡だから(C,D)や(D,C)よりは望ましい。すなわち、プレイヤーは(C,C)や(D,D)のような自分と相手の手のコーディネーションを望む。

表 1 囚人のジレンマ

		プレイヤー 2	
		C	D
プレイヤー 1	C	3,3	0,6
	D	6,0	1,1

さて、上記のようなプレイヤーからなる無限母集団が表 1で与えられているゲームを繰返し行うと仮定し、そのモデル化として次のような非定常マルコフチェーンを考える。各プレイヤーは $C_1,C_2,\ldots,C_m,D_{m+1},\ldots,D_{m+n}$ の $m+n$ 種類の状態のうち 1 つを取ると仮定する。状態 C_i $(i=1,\ldots,m)$のプレイヤーは手 C を取り、状態 D_i $(i=m+1,\ldots,m+n)$のプレイヤーは手 D を取る。プレイヤーの状態遷移は次のようである。まず、確率αで状態は変わらない。残りの確率 $1-\alpha$で、状態 C_i のプレイヤーは相手が手 C を取れば状態 C_{i-1} に遷移し、相手が手 D を取れば状態 C_{i+1} に遷移する。また、状態 D_i のプレイヤーは相手が手 C を取れば状態 D_{i-1} に遷移し、相手が手 D を取れば状態 D_{i+1} に遷移する。ただし、$C_0=C_1$, $C_{m+1}=D_{m+1}$, $D_m= C_m$, $D_{m+n+1}=D_{m+n}$ と解釈する。各プレイヤーは無限母集団内の任意のプレイヤーとランダムに対戦すると仮定する。無限母集団内におけるプレイヤーの状態 $C_1,C_2,\ldots,C_m,D_{m+1},\ldots,D_{m+n}$ の第 t 期 $(t=0,1,\ldots)$での分布を $z(t)=(z_1(t),\cdots,z_{m+n}(t))\left(\sum_{k=1}^{m+n} z_k(t)=1,\ z_k(t)\geq 0\right)$ とすれば、状態遷移図は図 1のようになり、$(z(t))_{t=0,1,\ldots}$ の満たす再帰関係式は(1)で与えられる。ただし、$p(t)=\sum_{k=1}^{m} z_k(t)$は t 期に手 C

© 日本オペレーションズ・リサーチ学会. 無断複写・複製・転載を禁ず.

を取るプレイヤーの割合である。

$$
(1)\quad \begin{cases} z_1(t+1)=\alpha z_1(t)+(1-\alpha)p(t)\bigl(z_1(t)+z_2(t)\bigr) \\ z_k(t+1)=\alpha z_k(t)+(1-\alpha)\bigl(p(t)z_{k+1}(t)+(1-p(t))z_{k-1}(t)\bigr)\ (k=2,\ldots,m+n-1) \\ z_{m+n}(t+1)=\alpha z_{m+n}(t)+(1-\alpha)(1-p(t))\bigl(z_{m+n}(t)+z_{m+n-1}(t)\bigr) \end{cases}
$$

また、$p(t)$の再帰関係式は次式(2)で与えられる。

$$
(2)\quad p(t+1)=p(t)+(1-\alpha)p(t)z_{m+1}(t)-(1-\alpha)(1-p(t))z_m(t)
$$

この非定常マルコフチェーンの定常分布$\bigl((z_1,z_2,\cdots,z_{m+n}),p\bigr)$は次の方程式(3)で与えられる。

$$
(3)\quad \begin{cases} z_1=p(z_1+z_2) \\ z_k=pz_{k+1}+(1-p)z_{k-1}\ (k=2,\ldots,m+n-1) \\ z_{m+n}=(1-p)(z_{m+n}+z_{m+n-1}) \\ z_1+\cdots+z_{m+n}=1 \\ p=\sum_{i=1}^{m} z_i \end{cases}
$$

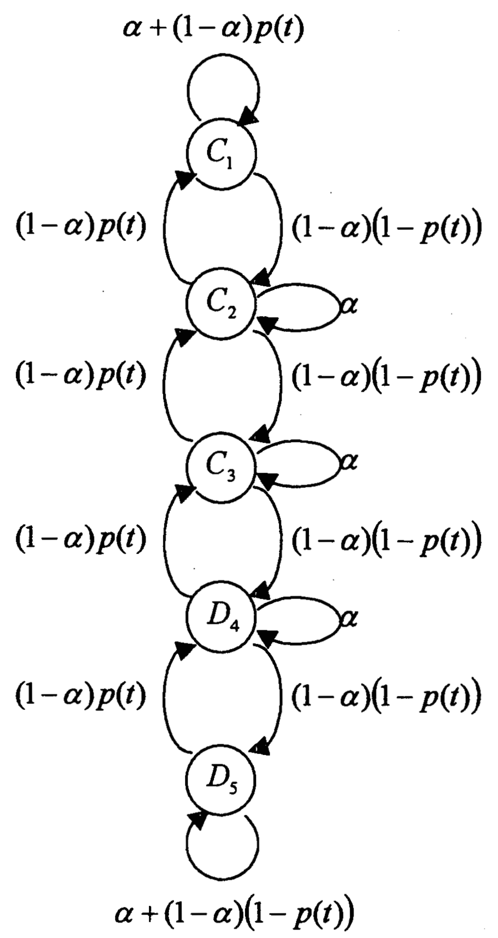

図 1

これは容易に求めることができ、(1)$\bigl((0,\cdots,0,1),0\bigr)$と(2)$\bigl((1,0,\cdots,0),1\bigr)$と(3)$\bigl((\tilde{z}_1,\tilde{z}_2,\cdots,\tilde{z}_{m+n}),\tilde{p}\bigr)$の3つとなる。初期分布を特別な形に限定すれば次のような漸近的な結果が得られる。

「$m>n\geq 2$と仮定する。任意の$r>0$に対して、$z_k=z_1r^{k-1}\ (k=1,\ldots,m+n), z_1=\left(1+r+\cdots+r^{m+n-1}\right)^{-1}$となる$z=(z_1,\ldots,z_{m+n})$を初期分布とする。また、$\tilde{r}$を$1+r+\cdots+r^{m-2}=r^m\left(1+r+\cdots+r^{n-2}\right)$の唯一解とする。($\tilde{r}>1$となる。)この時、次が成り立つ:

1. $0<r<\tilde{r}$ならば$z(t)$は定常分布(2)に漸近する。
2. $r=\tilde{r}$ならばzは定常分布(3)である。
3. $\tilde{r}<r$ならば$z(t)$は定常分布(1)に漸近する。

ただし、$\tilde{p}=\dfrac{1+\tilde{r}+\cdots+\tilde{r}^{m-1}}{1+\tilde{r}+\cdots+\tilde{r}^{n+m-1}}$である。」

3. 考察

数値例としてm=3,n=2 で前節で取り上げたタイプ以外の初期分布$z=(0.372,0,0,0,0.628)$を取り上げる(α=0)。40 期目あたりで$(1,0,0,0,0)$に漸近する。外部から見れば、初め 37.2%の人が手 C を取っているだけだったが、時間が経過するにつれてほとんどの人が手 C を取っていることになる。相手も自分と同じ手 C を取って欲しいと望む人の割合が相対的に少なくても(0.372<0.628)D を取ろうとする誘惑にかなり対抗できるなら(m=3>2=n)、協調行動が実現する。

参考文献

[1] Binmore, Ken. *Fun and Games*. Chapter 9. D.C. Heath and Company, 1992.

© 日本オペレーションズ・リサーチ学会. 無断複写・複製・転載を禁ず.