確率過程としての時間発展的なゲーム理論 (Evolving Game Theories as Stochastic Process)

Mitsuru KIKKAWA　（大学院研究員)
mitsurukikkawa@hotmail.co.jp

THE LATEST VERSION IS AVAILABLE AT
http://kikkawa.cyber-ninja.jp/index.htm

分量が多いため Handoutなし

OUTLINE

1. Introduction (Motivation, Purpose)
2. Related Literatures and Review
3. Repeated Game Theory (Random Stopping)
4. Stochastic Evolutionary Game Theory
   4-1. Markov Chain
   4-2. Perfect Memory (Martingale)
5. Summary (Future works)

1. INTRODUCTION

OUR CONTRIBUTIONS

- 時間発展的なゲーム理論(繰り返しゲーム、進化ゲーム)を確率論の枠組みで定式化し、その性質を利用し、分析した。その過程で数理としては「積に関するWaldの等式」を得た。

MOTIVATIONS

- This Talk is consist of the project for Game Theory as Probability Theory (Normal Form Game , Statistical Mechanics, etc.).
- **合理的な人間ならば、何度も同じゲーム(繰り返しゲーム)を行っている途中でこのゲームを認識するであろう。**
- **ではそれを説明する数理は？**

→確率の概念を用いて記述。

2. Related Literatures and Review

MAJOR REFERENCES

- **河野敬雄** (2003): **「進化ゲームアラカルト – 確率論の立場から -」**『*Rokko Lectures in Mathematics*』, 13.

→**確率論の立場**(Kolmogorov**の公理系**)**で様々なモデルを分析。この論文の考えを継承し、研究した。**

Full paper is available at

http://www.math.kobe-u.ac.jp/publications/rlm13.pdf

- Nowak, Martin A. (1990): "Stochastic Strategies in the Prisoner's Dilemma," *Theoretical Population Biology*, Vol.38, pp.93-112.

→**囚人のジレンマゲームを**Markov Chain **として記述し、一般的な性質を導出。ノイズを導入し、**ESS**の存在を証明。河野敬雄** (2003)Ch.10 **繰り返し囚人のジレンマ**(**マルコフ連鎖として**)

Martingale

Repeated Game:

- Fudenberg, Drew and Levine, David K. (1992):*The Review of Economic Studies*,Vol.59, No. 3, pp. 561-579.
- Mailath, G. J. and Samuelson, L.(2006): *Repeated Games and Reputations - Long-Run Relationships*, Oxford University Press.

Stochastic Game:

- Aumann, Robert J. and Maschler, Michael B. with the collaboration of Richard B. Stearns (1995): *Repeated Games with Incomplete Information*, MIT Press.
- Sorin, Sylvain (2002): *A First Course on Zero-Sum Repeated Games*, Springer-Verlag.

→これらの先行研究とは異なり、最も基本的なゲームがマルチンゲールであるかどうかを調べた。

Probability

- 高校の数学で「確率」は次のように定義されている.
- **Laplace流確率**「ある標本空間に含まれる基本事象の個数がN個であり, どの基本事象の起こることも同程度に確からしいものとする。事象Aがその中に a個の基本事象を含んでいるならば、Aの確率P(A)はa/N とする.」
- ここで試行を無限の場合を考えると、それぞれの事象の**確率は0** となる！？
- そこでKolmogorovの公理系(集合論の立場)の導入。

定義 $(\Omega, \mathcal{F})$ を可測空間とするとき, $\mathcal{F}$ を定義域とする実数値関数Pが次の性質を満たすならば, Pは $(\Omega, \mathcal{F})$上の**確率**, あるいは, **確率測度**であるという.

(a) $0 \leqq P(A) \leqq 1, \ \forall A \in \mathcal{F}$ (b) $P(\Omega) = 1$

(c) $\{A_n\}$が互いに素な$\mathcal{F}$の元の列であれば, すなわち, $A_n \in \mathcal{F}, \forall n \geqq 1, A_n \cap A_k = \varphi, n \neq k$ ならば, $P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$ が成り立つ.

Is mixed strategy a random variable or a probability distribution ?

- Most of the game theory's textbooks are adopted the mixed strategy as a probability distribution over pure strategies.

しかし、無限回繰り返しゲーム理論の場合、確率が定義できるのか?（Aumann (1989))。

- **Random variable** **Ω から** R^d **の中への** *F* **可測な関数**

  Ex) **サイコロを振る時出る目の数、くじを引く時の賞金など。**

- **Probability distribution** **確率変数**X**のとる値** xi **と**X**が** xi **をとる確率** pi **との対応関係を示したもの。**

  $Px(E) := P(X^{-1}(E)), \forall E \in B.$

→Mixed Strategy is a Random Variable の方がいいのでは？

Method

- Step1) This game is a **Martingale** or not.

→Yes

- Step 2) Apply **Optimal Stopping Theorem** (Stopping time is finite)
- Step 3) Derive **Wald 's equation**
- Step 4) Derive a **Nash Equilibrium**.

- We apply this Method to **Repeated Game** (Section 3) and **Stochastic Evolutionary Game** (Section 4).

Martingale に関する準備

Definition. A sequence $X=\{X_n; n \geqq 0\}$ is a **martingale** with respect to the sequence $\mathcal{F}=\{\mathcal{F}_n; n \geqq 0\}$ if, for all $n \geqq 0$,

(i) $E(|X_n|)<\infty$

(ii) $\{\mathcal{F}_n\}$ is filtration, $\{X_n\}$ is adapted for $\{\mathcal{F}_n\}$.

(iii) $E(X_{n+1} \mid \mathcal{F}_n) = X_n$ a.s. , $\forall\ n \geqq 0$

一種の
合理的
期待

In (iii), if "$=$" can be replaced "$\leqq$", we call the pair $\{X_n, \mathcal{F}_n ; n \geqq 0\}$ a supermartingale, and "$\geqq$", we call the pair $\{X_n, \mathcal{F}_n ; n \geqq 0\}$ a submartingale.

Proposition. Markov Chain is a Martingale.

定理 (**マルチンゲールの収束定理**) $\{X_n, \mathcal{F}_n ; n \geqq 1\}$ **が劣マルチンゲールで** $\sup_n E(X^+_n) < \infty$ **を満たしているとすると**, $n \to \infty$ **のとき**, $\{X_n\}$ **は** $E(|X|) < \infty$ **を満たす極限**X**に概収束する**.

Definition. A random variable n taking values in $\{0,1,2,\ldots\} \cup \{\infty\}$ is called a **stopping time**, or **Markov times** with respect to the filtration $\{\mathcal{F}_n \mid n \geqq 0\}$ if $\{\omega \mid N(\omega) = n\} \in \mathcal{F}_n, \forall n \geqq 0$.

定理 (**任意抽出定理**) $(X_k, \mathcal{F}_k)$**を劣マルチンゲール**, T_k, $k=1,2,\ldots$ **を**$\mathcal{F}_n$- **停止時間とする**. T_k **は有界**$T_k \leqq m_k$, **かつ**, **増大** $T_k < T_{k+1}$, $k=1,2,\ldots$ **とする**. Y_k**を**

$$Y_{k(\omega)} \equiv X_{T_{k(\omega)}}(\omega),\ k=1,2,\ldots$$

と定義すれば, (Y_k, F_{T_k})**も劣マルチンゲールである**.

補題1 $\mathcal{F}_n$-**停止時間** S **と** T **が** S≦T **を満たせば**, $\mathcal{F}_S \subset \mathcal{F}_T$ **である**.

補題2 (X_k, F_k) **を劣マルチンゲール**, S, T **を** $\mathcal{F}_n$-**停止時間とする**.

$$S \leqq T \leqq m \text{ (定数)のとき}, \ E(X_T \mid \mathcal{F}_S) \geqq X_S, \text{ a.s.}$$

Theorem. Let $(Y, \mathcal{F})$ be a martingale, and let T be a stopping time. Then $E(Y_T) = E(Y_0)$ if the following holds:

a) $P(T<\infty) = 1$, $E(T) < \infty$

b) there exists a constant c such that

$$E(|Y_{n+1} - Y_n| \mid \mathcal{F}_n) \leqq c, \text{ for all } \forall n < T.$$

- **Wald's equation.** Let $X_1, X_2, \ldots$ be independent identically distributed random variables with fintie mean μ , and let $S_n = \sum_{i=1}^{n} X_i$. It is easy to see that $Y_n = S_n - n\mu$ constitutes a martingale with respect to the filtration $\{F_n\}$ where $F_n = \sigma(Y_1, Y_2, \ldots, Y_n)$. Now

$$E(\,|Y_{n+1} - Y_n| \mid F_n) = E\,|X_{n+1} - \mu|$$
$$= E\,|X_1 - \mu| < \infty.$$

We deduce from Optimal Stopping theorem that $E(Y_T) = E(Y_0) = 0$ for any stopping time T with finite mean, implying that

$$\underline{E(S_T) = \mu\, E(T).}$$

展開形ゲームに関する事項

定義1 **プレイヤー行動戦略** i **は独立な確率測度の族の** $(b_i(u_i))_{u_i \in U}$, ただし $b_i(u_i)$ は行動の集合 $A(u_i)$ 上の確率測度である。特に各情報集合 $u_i \in U$, そして行動 $a \in A(u_i)$ のとき, $b_i(u_i)(a)$ と定義する。

定義2 $\Gamma = (K, P, p, U, h)$**を展開形ゲームとする. ゲームの木**K**の部分木**K' **に対して,**

Γのすべての情報集合は K' **の手番と** K' **以外の手番を同時に含むことはない**

ならば、ゲームΓの各構成要素を部分木K' **に制限することにより部分木**K' **をもつ展開形ゲームを定義できる。このようなゲームを、ゲームΓの部分ゲームという. 便宜上, ゲーム自身もΓの**1**つの部分ゲームと見なす。**

定義3　ゲームΓの部分ゲームΓ' とΓ'における行動戦略の組 $b'=(b'_1,\ldots,b'_n)$に対して, 部分ゲームΓ'全体をプレイヤーの期待利得ベクトル$\hat{H}^s(b')=(\hat{H}^s_1(b'),\ldots,\hat{H}^s_n(b'))$で置き換えてできるゲームを, 部分ゲームΓ'と行動戦略の組 b' によるゲームΓの**縮約ゲーム**といい, $T(\Gamma \mid \Gamma', b)$ と表す.

定義4　展開形ゲームΓの部分ゲームΓ' がプレイヤーの行動戦略の組 $b=(b_1,\ldots,b_n)$ によって**到達可能**であるとは, 行動の戦略の組 bの下でΓ' の初期点o' が正の確率 $p(o' \mid b)>0$ を持つことである.

定義5 展開形ゲームにおいてすべてのプレイヤー i (=1,…,n) のすべての情報集合 $u \in U_i$ がただ1つの手番からなるとき, ゲームΓ は**完全情報ゲーム**であるという.

定理1 (Kuhn [15]) ゲームの木が有限の長さを持つ完全情報ゲームでは, 純戦略による均衡点が存在する。ただし、ゲームの木の長さとは1つのプレイに含まれる手番の最大数のことである.

REVIEW: Replicator Equation

REPLICATOR EQ. $\dot{x}_i = x_i\left((Ax)_i - x\cdot Ax\right), i=1,\cdots,n.$

ある戦略 i を取ったときの自分の利得が平均利得よりも大きい場合には, その戦略を取る確率が高くなる(学習・模倣)。またゲームをしている周りの主体がその戦略を取る確率が高いほどその増加率も高くなる(外部性の存在), ということを示している.

また利得が高ければ、その戦略をとる人が多いという仮定(利得の単調性)を仮定すれば、選択ダイナミクスから一意に導出される。

特に戦略が2つのとき

$$\dot{x} = x(1-x)\{b-(a+b)x\} \quad \cdots (*)$$

(注意) Potential Function で考えている. 2

主な分類：

(I) 非ジレンマ型: a > 0. b < 0, ESS :1つ
(II) 囚人のジレンマ：a < 0. b > 0, ESS :1つ
(III) コーディネーション型: a>0,b>0, ESS 2つ
(IV)タカ＝ハト型: a<0,b < 0, ESS 1つ(混合戦略)

1

	戦略1	戦略2
戦略1	a,a	0,0
戦略2	0,0	b,b

利得表

3. Repeated Game Theory (Random Stopping)

I made it.

１．ランダム停止時刻を持つ繰り返しゲーム理論を確率過程として捉え、マルチンゲールであるかどうかを調べた。

２．ある条件の下、このゲームはマルチンゲールであった。この性質を利用し、繰り返し囚人のジレンマゲームにおいて、Pareto最適な協調行動を採用する条件を導出した。

KEYWORDS

- Repeated Game, Martingale, Random Stopping Time

MODEL

- **通常の繰り返しゲーム理論にランダム停止時間、いつ終わるか分からない場合。**

例）一方のプレイヤーがドロップアウトする場合や外生的な要因でゲームが終了する場合。

仮定　すべてのプレイヤーはゲームの停止までの利得和を最大にする。

定義 戦略型ゲームG**のランダム停止時刻**$\zeta(\leqq\infty)$ **をもつランダム停止時間を持つ繰り返しゲーム** $G^\wedge\zeta$ **とは,**

$$G^\wedge\zeta = (N, \{Si\}i\in N, \{Fi\}i\in N)$$

で定義される.

定義 このゲーム$G^\wedge\zeta$におけるプレイヤー1の「**戦略**」とは確率過程 $X_1, X_2, \ldots$ の分布が次のようなルールで帰納的に定まることである.

(i) 確率変数 X_1 の分布を決定する

(ii) プレイヤー2の任意の「戦略」と自分の「戦略」のn回までの実現値

$$(X_1(\omega), \ldots , X_n(\omega), Y_1(\omega), \ldots , Y_n(\omega))$$

を知った時の n+1 回目の条件付確率

$$\mathrm{Prob}\ (X_{n+1} = * \mid \mathcal{F}_n(X,Y)),\ n=1,2, \ldots$$

を決める.

- プレイヤーiは行動の組の列 $a(s) = \{a^t(s)\}_{t=1}^{\zeta}$ を平均利得和は次のように与えられる.

$$F_i(s) = E\left[\sum_{t=1}^{\zeta} f_i\left(a^t(s)\right)\right]$$

によって評価する. $F_i(s)$ を戦略の組 $s=(s_1,\dots,s_n)$ に対する i の利得と呼ぶ。このとき

$$u(X,Y) = \sum_{k=1}^{+\infty} E\left[\sum_{n=1}^{k} u(X_n,Y_n);\zeta = k\right] = \sum_{k=1}^{+\infty}\sum_{k=n}^{+\infty} E\left[u(X_n,Y_n);\zeta = k\right]$$

$$= \sum_{n=1}^{+\infty} E\left[u(X_n,Y_n);\zeta \geq n\right].$$

を得る。

Example

- 停止時間ζが確率変数列(X,Y) と独立であり, その分布は幾何分布 G(δ) (Prob(ζ=n)=(1-δ)δ^(n-1), n=1,2,…) であると仮定すると,

$$u(X,Y)=\sum_{n=1}^{\infty}E\left[u(X_n,Y_n)\right]\mathrm{Pr}ob(\zeta\geq n)=\sum_{n=1}^{\infty}E\left[u(X_n,Y_n)\right]\delta^{n-1}$$

となる.

→この効用関数は繰り返しゲーム理論で使われている割引因子δによる割引利得和と等しい。

Method

- Step1) This game is a **Martingale** or not.

→Yes

- Step 2) Apply **Optimal Stopping Theorem** (Stopping time is finite)
- Step 3) Derive **Wald 's equation**
- Step 4) Derive a **Nash Equilibrium**.

- We apply this Method to **Repeated Game** (Section 3) and **Stochastic Evolutionary Game** (Section 4).

命題　このゲームは停止時刻が幾何分布に従い、有限の場合, 期待利得が0の場合マルチンゲールである。また無限の場合マルチンゲールである.

Proof : $S_n = u_1 + \delta u_2 + \cdots + \delta^{n-1} u_n$　とおく. ただし仮定から 0<δ<1, n はランダム停止時刻である. さらにここで

$S'_n = \sum_{i=1}^{n} \left(\delta^{i-1} u_i - \mu \right), \ \mu = \delta^{n-1} E(u_n)$　とおくと,

$$E(S'_n \mid F_{n-1}) = E(S'_{n-1} \mid F_{n-1}) + E(\delta^{n-1} u_n - \mu \mid F_{n-1}) = S'_{n-1}$$

これより S'n はマルチンゲールである。よってSn がマルチンゲールであるためには, (i) E(δ^{n-1} un) が0 , (i) を満たさない場合は, (ii) n → ∞の場合である。つまり停止時刻が有限の場合、期待利得が0の場合マルチンゲールである。また停止時刻が無限の場合マルチンゲールとなることが分かる。(i) の場合は期待利得の戦略的同等性から期待利得が0となるものが存在することが分かる。

(証 終)

Method

- Step1) This game is a <u>**Martingale**</u> or not.

→Yes

- Step 2) Apply <u>**Optimal Stopping Theorem**</u> (Stopping time is finite)
- Step 3) Derive <u>**Wald 's equation**</u>
- Step 4) Derive a <u>**Nash Equilibrium**</u>.

- We apply this Method to **Repeated Game** (Section 3) and **Stochastic Evolutionary Game** (Section 4).

Wald の方程式の導出

任意抽出定理を用いると、

$$E(S'_N) = E(u_1 - \mu)$$

を得るが、これは0に等しい。ここで S'n=Sn-nμ に注意すると、次の関係式を得る.

$$E\left(\sum_{i=1}^{N} \delta^{i-1} u_i\right) = E(N)\mu.$$

この関係式から、有限の停止時刻までの利得和は平均の停止時刻と期待利得の平均との積で定まるので、各利得の大きさから Nash 均衡が定まる。

Method

- Step1) This game is a ***Martingale*** or not.

→Yes

- Step 2) Apply **Optimal Stopping Theorem** (Stopping time is finite)
- Step 3) Derive **Wald 's equation**
- Step 4) Derive a **Nash Equilibrium**.

- We apply this Method to **Repeated Game** (Section 3) and **Stochastic Evolutionary Game** (Section 4).

EXAMPLE (Prisoner’s Dilemma Game)

- 利得表1

	戦略C	戦略D
戦略C	R, R	S,T
戦略D	T,S	P,P

利得表1’

	戦略C	戦略D
戦略C	R’,R’	S’,T’
戦略D	T’,S’	P’,P’

ただしT>R>P>S, 利得表1’は停止時間(有限)までの平均期待利得を表している。つまり $R'=\sum_{i=1}^{n}\delta^{n-1}R$,T’, S’, P’も同様に定義する。

- Pareto　最適な協調行動(C,C)を実現するためにはNash均衡条件から, R’≧T’かつS’ ≧ P’という条件が必要である。これから $1\geq p^{*}\geq\frac{T}{R+T}\geq 0$ かつ $1\geq p^{*}\geq\frac{S}{S+P}\geq 0.$

よって $1\geq p^{*}\geq\max\left\{\frac{T}{R+T},\frac{S}{S+P},0\right\}.$ を得る。

ただし p*:~~停止時間まで戦略~~Cを取る平均確率.

またこのゲームがマルチンゲールであるための条件

p*^2(R’+S’)+(1-p*)^2(T’+P’)=0を満たしている必要がある.

この結果が示唆していること

- 有限回繰り返しゲーム理論は通常非協調な戦略がNash均衡。

→停止時刻まで足し合わせ、最適な行動の経路を導出。

- これに対して、Friedman(1985)はε均衡 (εの範囲内でNash均衡よりも良い戦略がある、一種の認識のズレ)を導入し、協調の可能性を示した。
- 我々のものは、途中でこのゲームを認識し、平均的な利得を計算し、戦略を決定するので、その最適からのズレから、協調行動の可能性を示唆している。

4. Stochastic Evolutionary Game Theory

I made it.

１．展開形ゲームをマルコフ連鎖と捉え、定常分布の諸性質を調べた(Nowak (1990))。

２．このモデルを拡張させ、過去全ての構造に影響がある場合ゲーム理論を定式化した。ここでもマルチンゲールであるかどうか調べ、このゲームはマルチンゲールであることが分かった。その過程で、積に関するWaldの方程式を導出した。この性質を利用し、繰り返し囚人のジレンマゲームにおいて、Pareto 最適な協調行動を採用する条件を導出した。

KEYWORDS

- Stochastic Evolutionary Game, Markov Chain, Extensitve Game , Martingale,

Markov Chain

- Markov Chain ＝ game in extensive form
- Behavioral strategy (but we don't analyze the relation between a mixed strategy and behavioral strategy)
- $(\Omega,\mathcal{F},P)$ **を確率空間とし**, Ω**は空間**, $\mathcal{F}$**は空間**Ω**の部分集合の族**, P **は** $(\Omega,\mathcal{F})$ **上の確率とする。**

- **展開形ゲーム：** $\Gamma = (K, P, p, U, h)$
- K:**ゲームの木**, P: **プレイヤーの分割**, p: **偶然手番の確率分布族であり**, **確率変数**, U : **情報分割**, h: **利得関数**
- **このとき**, $\mathcal{F}_n \in U$, $n \geqq 0$ **となる**.

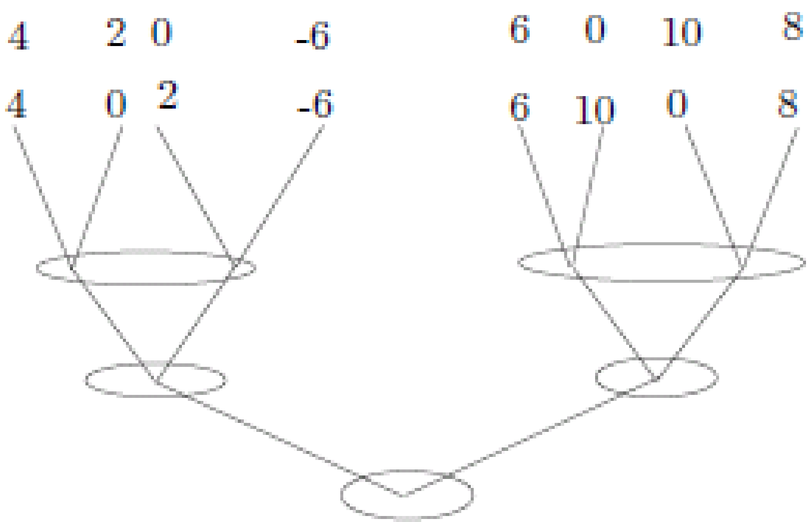

図 1. 展開形ゲームの例

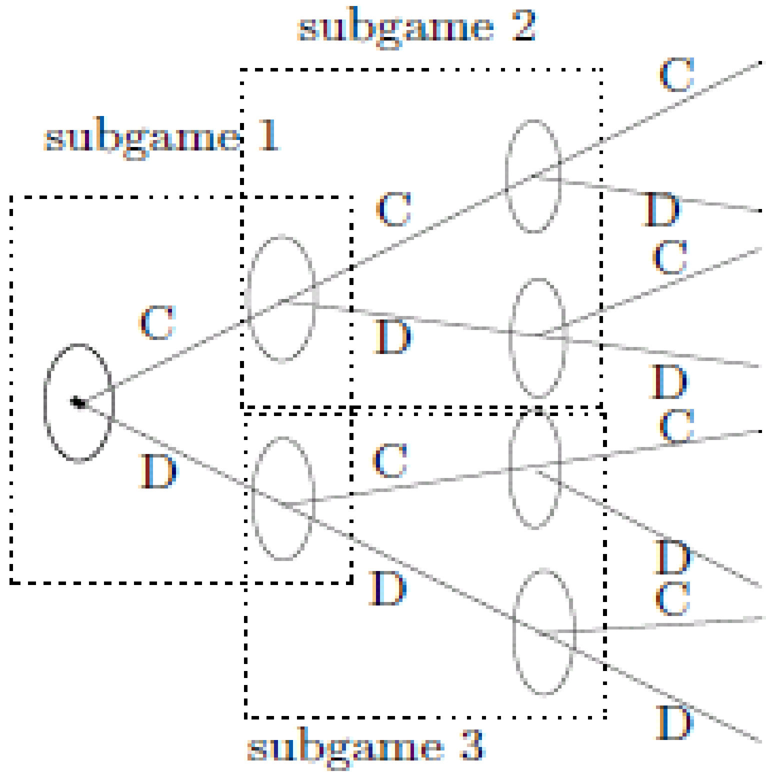

図 2. Markov 連鎖としてのゲーム

確率過程に関する事項

定義6 **確率過程** $\{X(t,\omega);\ 0 \leqq t < +\infty,\ \omega \in \Omega\}$ とは、空でない任意の集合Tをパラメーター空間として持つ確率空間$(\Omega, \mathcal{F}, P)$上で定義された確率変数の族である.
よって $t \in T$ を固定する毎に, $X(t)$ が確率変数であることを意味する。特にTが自然数ないし整数の時には, 確率変数列という.

定義7 事象Bが与えられた時の, 事象Aの**条件付き確率**とは次を満たすことをいう.

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

仮定 1 状態空間Sが有限集合, つまり戦略の数が有限の場合を考える。よって

$$S = \{s_1, s_2, \dots, s_N\}$$

とする。

定義8 $\{X_n; n=0,1,2,\dots\}$ が**独立確率変数列**であるとは、任意の $n=1,2,3,\dots$, 任意の状態 $s_0, \dots, s_n$ に対して

$$P(X_n=s_n \mid X_0=s_0, \dots, X_{n-1}=s_{n-1}) = P(X_n=s_n)$$

が成り立つときをいう.

定義9 $\{X_n; n=0,1,2,\dots\}$ が**Markov連鎖**であるとは,任意の $n=1,2,3,\dots$, 任意の状態 $s_0, \dots, s_n$ に対して

$$P(X_n=s_n \mid X_0=s_0, \dots, X_{n-1}=s_{n-1}) = P(X_n=s_n \mid X_{n-1}=s_{n-1}) \tag{2.1}$$

が成り立つときをいう. また(2.1) を**Markov性**という.

定義10

$P(X_n=j \mid X_{n-1}=i) = P^{\wedge}(n)_{ij}$, $P^{\wedge}(n)=(P^{\wedge}(n)_{ij})$, $n=1,2,\ldots, N$

$P^{\wedge}(n)$ をステップ n-1 における**推移確率行列**という.

よって$P^{\wedge}(n)$ は状態空間Sの要素の数をNとすると, N次正方行列である.

定理2 Markov連鎖は $n=1,2,\ldots$ に対する推移確率行列 $P^{\wedge}(n)$と, 初期分布

$$P(X_0=i)=\mu_i,\ i \in S$$

によって一意に定まる.

定義11 $P^{\wedge}(n)$ がn に依存しないとき, **時間的に一様なMarkov連鎖**という.

定義12　初期分布μ_iのとき, 次を満たす確率分布μ_jは Markov連鎖の**定常分布**と呼ばれる.

$$\mu_j = \sum_{i \in S} \mu_i P^{(n)}, \qquad j \in S$$

まとめ：

第k回目のゲームにおける戦略セットを $\vec{X}_k = \left(X_{1,k}, \cdots, X_{m,k}\right)$ とする(ただしmはゲームをしているプレイヤーの数). ここで任意の有限次元の確率法則を定めると確率過程 $\left\{\vec{X}_k, k = 1,2,\cdots\right\}$ が定まって、繰り返しのあるゲームが定式化されたことになる。特に前回の対戦の戦略のみを考慮に入れて次回の戦略を決めるのであれば、この確率過程は離散時間のMarkov過程であり、展開形ゲームとなる。さらに各自の戦略の集合が有限集合であれば、Markov連鎖でありその規則が時間に依存しなければ、時間的に一様なMarkov連鎖となる。

定理3 (大数の強法則)

(i) **有限状態を持つ既約エルゴード的Markov連鎖は唯一の定常分布** $\vec{\mu} \in P(S^{(n)})$**を持つ。**

(ii) **このとき**

$$\lim_{T \to +\infty} \frac{1}{T} u_i\left(\vec{X}_T\right) = \int_{S^{(n)}} u_i(s_1, \cdots, s_n) d\vec{\mu}(s_1, \cdots, s_n), \forall i,$$

が確率1で成り立つ。

まとめ：

非協力ゲーム：ゲームの木が有限の長さを持つ時、純戦略による均衡の存在(定理1)。

進化ゲーム：ゲームの木が無限の長さを持つ時、定常分布の存在(定理3)。

EQUILIBRIUM

定義13 **戦略** $\vec{X} \in L_0(S^{(n)})$ **が****Nash均衡****であるとは**

$$E\left[u_i\left(\vec{X}\right)\right] \geq E\left[u_i\left(\vec{X}_{-i}, Y_i\right)\right], \forall i, \forall Y_i \in L(S_i), (\vec{X}_{-i}, Y_i) \in L(S^{(n)})$$

が成り立つときをいう.

定義14 **戦略** $\vec{X} \in L_0(S^{(n)})$ **が進化的に安定な戦略とは**,

$$(1-\varepsilon)E\left[u_i(\vec{X})\right] + \varepsilon E\left[u_i(\vec{X}, Y_i)\right] > (1-\varepsilon)E\left[u_i(\vec{X}_{-i}, Y_i)\right] + \varepsilon E\left[u_i(\vec{Y})\right]$$

$$\forall \vec{Y} \neq \vec{X}, \exists \varepsilon_0 > 0, 0 < \forall \varepsilon < \varepsilon_0$$

が成り立つときをいう.

記号： L(Si)**を**Si**に値を取る確率変数**Xi**の全体**(i=1,…n)

L(S^(n))**を**S^(n)**に値を取る確率変数** $\vec{X} = (X_1, \cdots, X_n)$ の全体,

L0(S^(n))**を** $\left\{\vec{X} = (X_1, \cdots, X_n) \in L(S^{(n)}), X_1, \cdots, X_n \text{が独立確率変数}\right\}$,

$\left(\vec{X}_{-i}, Y_i\right)$を$(X_1, \cdots, X_{i-1}, Y_i, X_{i+1}, \cdots, X_n)$,

$\left(\vec{X}_i, Y_i\right)$を$(Y_1, \cdots, Y_{i-1}, X_i, Y_{i+1}, \cdots, Y_n)$,

戦略が2つの場合

- **設定**: **プレイヤー：**I、II。**純粋戦略の集合**S＝{C,D}, Z_n＝(X_n, Y_n); n＝0,1,… **は** S×S **の値を取る時間的に一様な** Markov**連鎖とする。**
- **このときの推移確率行列**P(**各状態は**(C,C),(C,D),(D,C),(D,D)**の順に並んでいる。例** $Z_n(\Omega)$＝(C,D) **は根元事象**Ω**において**, n**回目にプレイヤー**I**が戦略**C**を出し、プレイヤー**II**が戦略**D**を出したことを意味する。**

$$P=\begin{pmatrix} pp' & p(1-p) & (1-p)p' & (1-p)(1-p') \\ qp' & q(1-p') & (1-q)p' & (1-q)(1-p') \\ pq' & p(1-q') & (1-p)q' & (1-p)(1-q') \\ qq' & q(1-q') & (1-q)q' & (1-q)(1-q') \end{pmatrix}$$

- p**または**q**はプレイヤー**II**が直前にそれぞれ**C**または**D**を出した時、プレイヤー**I**が**C**を出す条件付き確率である。**

Column
に着目

…合

…ヤ…II。…略の集合…,D},
Zn=(Xn, … 0, … S×S … を取る時 … に一様な
Markov… す

- このと … 移 … 行列P … 様は(C,C … ,(D,C),
(D,D)の … る。 … =(C,D) … 事象Ω
におい … 目 … イヤ … 戦略Cを … プレイ
ヤーIIが戦略Dを出したことを意味する。

$$P=\begin{pmatrix} pp' & p(1-p') & (1-p)p' & (1-p)(1-p') \\ qp' & q(1-p') & (1-q)p' & (1-q)(1-p') \\ pq' & p(1-q') & (1-p)q' & (1-p)(1-q') \\ qq' & q(1-q') & (1-q)q' & (1-q)(1-q') \end{pmatrix}$$

- pまたはqはプレイヤーIIが直前にそれぞれCまたはDを出した時、プレイヤーIがCを出す条件付き確率である。

…が2つの場合

Rowに着目

前期の両プレイヤーの戦略(C,C)

前期の各プレイヤーの戦略(C,D)

前期の各プレイヤーの戦略(D,C)

前期の両プレイヤーの戦略(D,D)

$$\begin{pmatrix} \ldots p' & p(1-p) & (1-p)p' & (1-p)(1-p') \\ \ldots ' & q(1-p') & (1-q)p' & (1-q)(1-p') \\ pq' & p(1-q') & (1-p)q' & (1\ldots \\ qq' & q(1-q') & (1-q)q' & (1\ldots \end{pmatrix}$$

- pまたはqはプレイ… それぞれCまたはDを出した時、… を出す条件付き確率である。

EXAMPLE

2つの場合

イヤー：I、II。純粋戦略の集合S={C,D},
$Z_n=(X_n,Y_n)$; n=0,1,… はS×Sの値を取る時間的に一様なMarkov連鎖とする。

プレイヤーI, II とも
前期の行動が戦略C
であり、今期もプレイ
ヤーI, II共に戦略C

- このと 各状態は(C,C),(C,D),(D,C),(D,D)の 例 $Z_n(\Omega)$=(C,D) は根元事象Ωにおい ーIが戦略Cを出し、プレイヤーIIが戦略Dを出したことを意味する。

$$
P=\begin{pmatrix}
pp' & p(1-p) & (1-p)p' & (1-p)(1-p') \\
qp' & q(1-p') & (1-q)p' & (1-q)(1-p') \\
pq' & p(1-q') & (1-p)q' & (1-p)(1-q') \\
qq' & q(1-q') & (1-q)q' & (1-q)(1-q')
\end{pmatrix}
$$

- pまたはqはプレイヤーIIが直前にそれぞれCまたはDを出した時、プレイヤーIがCを出す条件付き確率である。

EXAMPLE 2つの場合

- イヤー：I、II。純粋戦略の集合S={C,D}, $Z_n=(X_n,Y_n)$; n=0,1,… はS×Sの値を取る時間的に一様なMarkov連鎖とする。
- このときの推移確率行列P(各状態は(C,C),(C,D),(D,C),(D,D)の順に並んでいる。例 $Z_n(\Omega)=(C,D)$ は根元事象Ωにおいて, n回目にプレイヤーIが戦略Cを出し、プレイヤーII とを意味する。

> プレイヤーI, II とも前期の行動が戦略Dであり、今期もプレイヤーI, II共に戦略C

$$
P = \begin{pmatrix} & & -p)p' & (1-p)(1-p') \\ & & -q)p' & (1-q)(1-p') \\ & p(1-q') & (1-p)q' & (1-p)(1-q') \\ qq' & q(1-q') & (1-q)q' & (1-q)(1-q') \end{pmatrix}
$$

- pまたはqはプレイヤーIIが直前にそれぞれCまたはDを出した時、プレイヤーIがCを出す条件付き確率である。

EXAMPLE

2つの場合

イヤー：I、II。純粋戦略の集合S={C,D}, $Z_n=(X_n, Y_n)$; n=0,1,… はS×Sの値を取る時間的に一様なMarkov連鎖とする。

- このときの推移確率行列P(各状態は(C,C),(C,D),(D,C),(D,D)の順に並んでい　　　　　　　　　根元事象Ω において, n回目にプ　　　　　　　　　出し、プレイヤーIIが戦略Dを出し

前期各プレイヤーの戦略(C,D)であり、今期の各プレイヤー戦略は(D,C)

$$P=\begin{pmatrix} pp' & p(1-p') & (1-p)p' & (1-p)(1-p') \\ qp' & q(1-p') & (1-q)p' & (1-q)(1-p') \\ pq' & p(1-q') & (1-p)q' & (1-p)(1-q') \\ qq' & q(1-q') & (1-q)q' & (1-q)(1-q') \end{pmatrix}$$

- pまたはqはプレイヤーIIが直前にそれぞれCまたはDを出した時、プレイヤーIがCを出す条件付き確率である。

- **初期分布**: {C,C }, {C,D }, {D,C }, {D,D}

$$\pi_0 = \left(yy', y(1-y'), (1-y)y', (1-y)(1-y')\right)$$

y, y': **プレイヤー**I, II **は互いに独立に**C**をそれぞれ出す確率。**

- **利得関数**: f(C,C)=A, f(C,D)=0, f(D,C)=0, f(D,D)=B
- **利得表**1

	戦略C	戦略D
戦略C	A, A	0,0
戦略D	0,0	B,B

対称2**人ゲーム**

- **極限あり**

$$u(\vec{a},\vec{b}) = \lim_{n\to\infty} E_{\pi_0}[f(X_n, Y_n)]$$

極限なし

$$u(\vec{a},\vec{b}) = \lim_{n\to\infty} \frac{1}{n} E_{\pi_0}[f(X_n, Y_n)]$$

命題1 (Nowak[18]を変更)

(i) $|rr'|<1$ **の時**, Markov**過程は既約**, **非周期的であって、定常分布は** $\pi=\left(ss',s(1-s'),(1-s)s',(1-s)(1-s')\right)$**に収束する**.

このとき利得関数は次の値に収束する。

$$u(\vec{a},\vec{b})=Ass'+B(1-s)(1-s')$$

(ii) $r=r'=1$ **の時**, Markov**過程は状態**(C,C)**と**(D,D)**がそれぞれ定常分布であり**, (C,D), (D,C) **の**2**点がひとつの再帰類を作り、しかも周期は**2**である**. $u(\vec{a},\vec{b})=Ayy'+B(1-y)(1-y')$

(iii) $r=r'=-1$ **の時**, Markov**過程は再帰類が**(C,C),(D,D) **と**(C,D),(D,C)**の**2**つできて、いずれも周期は**2**である。このとき**1**周期辺りの期待利得は次の値に収束する。**

$$u(\vec{a},\vec{b})=\frac{A+B}{2}\left(yy'+(1-y)(1-y')\right).$$

(iv) $r=1, r=-1$ **の時**, (C,C) **から出発した**Markov**過程は**(C,C) ⇒ (C,D) ⇒ (D,D) ⇒ (D,C) **と順に回っていき**, **周期は**4**である**. **このとき**1**周期辺りの期待利得は次の値に収束する**.

$$u(\vec{a},\vec{b})=\frac{A+B}{4}.$$

命題2 (i) $|rr'| < 1$ ($\Leftrightarrow$ $r, r' \in (-1,1)$) を満たすとき, このゲームの定常分布は $A, B < 0$ のときESSとなる。

(ii) $r = r' = 1$ を満たすとき、このゲームの定常分布, 戦略の組 (C,C), (D,D) は $A > 0$, $B<0$ のとき (C,C) がESSとなり, $A<0$, $B>0$ のとき (D,D) がESSとなる、また $A, B > 0$ のとき (C,C), (D,D) それぞれESSとなる。ただしこの場合どちらの定常分布となるかは初期分布 y, y' による.

ノイズの導入：当初から取りうる戦略に制限を設ける。

$\varepsilon > 0$ を固定して, $\varepsilon \leqq p \leqq 1 - \varepsilon$, $\varepsilon \leqq q \leqq 1 - \varepsilon$,

命題3 $|rr'| < 1$ ($\Leftrightarrow$ $r, r' \in (-1,1)$) を満たすとき, このゲームにおいて, それぞれのゲームの型における定常分布はESSとなる。

RiSK DOMINANCE

定義15 (Harsanyi and Selten [7]) **コーディネーションゲームにおいて**, Nash **均衡** E_1 **が** Nash**均衡**E_2 **をリスク支配であるとは**, $A > B$ **のときをいう。また** E_2 **が** E_1 **をリスク支配であるとは**, $A < B$ **のときをいう**.

	戦略1	戦略2
戦略1	4, 4	0,3
戦略2	3 ,0	2, 2

系1 **このゲームにおいて**, $A, B > 0$ (**コーディネーション型**)**のとき**, **リスク支配されている戦略も定常分布となりえる**.

4-2. Perfect Memory (Martingale)

GAME WITH PERFECT RECALL

定義18 展開形ゲームΓが**完全記憶ゲーム**であるとは，すべてのプレイヤー i (i=1,…, n) のすべての情報集合 u, v∈ Ui に対して，もしvのある手番 y が u から枝 c によって到達可能ならば，v のすべての手番が同じ枝 c によって u から到達可能であることである．

- 完全記憶ゲームではすべてのプレイヤーは各手番において，

(1) 過去の自分の手番でのすべての選択，および

(2) 過去の自分の手番で利用可能であったすべての情報

を記憶している．完全情報ゲームは，完全記憶ゲームの特別な場合である．

- (3.1)　**確率変数列** $\{X_n; n=0,1,\ldots\}$, **任意の** $n=1,2,\ldots$, **任意の** $s_0,\ldots, s_n$ **に対して**,

$$P(X_n=s_n \mid X_0=s_0, \ldots, X_{n-1} = s_{n-1})$$

Method

- Step1) This game is a ***Martingale*** or not.

→Yes

- Step 2) Apply **Optimal Stopping Theorem** (Stopping time is finite)
- Step 3) Derive **Wald 's equation**
- Step 4) Derive a **Nash Equilibrium**.

- We apply this Method to **Repeated Game** (Section 3) and **Stochastic Evolutionary Game** (Section 4).

Main Theorem

定理 (**積に関する**Wald**の方程式**) X_n **は確率変数とする**. T **が** $E(T)<\infty$**を満たす停止時間**, μ**は**X_n**の平均であり**, **有限の値であるとき**,

$$E(X_1 \times \dots \times X_T) = \mu^{E(T)}$$

が成り立つ.

補題3 (**積のマルチンゲールに関する角谷の定理**) $X_1, X_2, \dots$ **を非負独立確率変数列で**, **各々の平均は**1**であるとする**. $M_0=1$**と定義し**, $n\in N$**に対して** $M_n := X_1 X_2 \dots X_n$ **とする**. **このとき** M**は非負マルチンゲールであり**, $M_\infty := \lim M_n$ **が** a.s. **に存在する**. **そして以下の**(i)-(v) **は同値である**.

(i) $E(M_\infty) = 1$, (ii) L^1**の意味で** $M_n \to M_\infty$、

(iii) (M) **は一様可積分**(UI), (iv) $\prod a_n>0$, **ただし** $0<a_n := E(X_n^{1/2}) \leqq 1$, (v) $\Sigma(1-a_n)<\infty$

もし上の5**つのどれか**1**つ成り立たないときには**,

$$P(M_\infty=0)=1 \text{ である}.$$

- **Proof**：$X_1, X_2, \ldots$ を独立な非負確率変数の列とし、$E(X_k) = \mu$, $\forall k$ とし，$M_0=\mu$, $F_0=\{\varphi, \Omega\}$ として，$M_n=X_1X_2\ldots X_n$, $F_n = \sigma(X_1,X_2,\ldots,X_n)$ と定義する．さらに $Y_n = M_n - \mu^T$ と置く．

  まずこの Y_n がそれぞれ filtration $\{F_n\}$ についてマルチンゲールであることを示す。ここでは M_n がマルチンゲールであれば、Y_n もマルチンゲールであるので、M_n のマルチンゲールであることを示す。このとき $n \geqq 1$ に対して、

$$E(M_n \mid F_{n-1})=E(M_{n-1}X_n \mid F_{n-1}) = M_{n-1} E(X_n|F_{n-1})=M_{n-1}E(X_n) = \mu M_{n-1}$$

を a.s. の意味で得る．よって $\mu \leqq 1$ のとき優マルチンゲールであり，$\mu=1$ のときマルチンゲールである。また上式の両辺を μ^n で割ると，$E(M_n/\mu^n \mid F_{n-1}) = M_{n-1}/\mu^{n-1}$ となる．よって　M_n/μ^n はマルチンゲールである．

  次にこれが任意の停止定理の条件を満たすことを示す。ここで補題3を用いると次の条件が成り立つことが分かる。

$$E(|Y_{n+1} - Y_n| \mid F_n) = E(M_n (1-X_n))< \infty$$

よって定理6より，$E(Y_T)=E(Y_0)=\mu$, $\forall T$ となる。これから

$E(Y_n) = E(M_n-\mu^T) = (X_1-\mu) = 0$.　よって $E(S_n-\mu^T)=0$,
$E(X_1 \times \cdot \cdot \cdot \times X_n) =\mu^{E(T)}$.　　　　　　　　（証終）

Method

- Step1) This game is a **Martingale** or not.

→Yes

- Step 2) Apply **Optimal Stopping Theorem** (Stopping time is finite)
- Step 3) Derive **Wald 's equation**
- Step 4) Derive a **Nash Equilibrium**.

- We apply this Method to **Repeated Game** (Section 3) and **Stochastic Evolutionary Game** (Section 4).

命題5　この完全記憶があるゲームはマルチンゲールである。

- ノイズ

$$P(X_n = s_n + \varepsilon \mid X_0 = s_0, \ldots, X_{n-1} = s_{n-1})$$

命題6　平均が0であるような無相関のノイズがある完全記憶があるゲームはマルチンゲールであり, 平均が0でないような相関のあるノイズがある完全記憶があるゲームはマルチンゲールではない。

Method

- Step1) This game is a **Martingale** or not.

→Yes

- Step 2) Apply **Optimal Stopping Theorem** (Stopping time is finite)
- Step 3) Derive **Wald 's equation**
- Step 4) Derive a **Nash Equilibrium**.

- We apply this Method to **Repeated Game** (Section 3) and **Stochastic Evolutionary Game** (Section 4).

EXAMPLE (Prisoner's Dilemma Game)

- **利得表**1

	戦略C	戦略D
戦略C	R, R	S,T
戦略D	T,S	P,P

利得表1'

	戦略C	戦略D
戦略C	R',R'	S',T'
戦略D	T',S'	P',P'

ただしT>R>P>S, **利得表**1' **は停止時間**(**有限**)**までの平均期待利得を表している。**

- Pareto **最適な協調行動**(C,C) **を実現するためには** Nash **均衡条件から**, R'≧T' **かつ** S' ≧ P' **という条件が必要である。これから** $1 \geq p^* \geq \frac{T}{R+T} \geq 0$ かつ $1 \geq p^* \geq \frac{S}{S+P} \geq 0.$

よって

$$1 \geq p^* \geq \max\left\{\frac{T}{R+T}, \frac{S}{S+P}, 0\right\}.$$

を得る。p*:**停止時間まで戦略**C**を取る平均確率**.

この結果が示唆していること(再)

- **有限回繰り返しゲーム理論は通常非協調な戦略がNash均衡。**

→停止時刻まで足し合わせ、最適な行動の経路を導出。

- **これに対して、Friedman(1985)はε均衡(εの範囲内でNash均衡よりも良い戦略がある、一種の認識のズレ)を導入し、協調の可能性を示した。**
- **我々のものは、途中でこのゲームを認識し、平均的な利得を計算し、戦略を決定するので、その最適からのズレから、協調行動の可能性を示唆している。**

5. Summary

- 時間発展的なゲーム理論(繰り返しゲーム理論、確率的進化ゲーム理論)を確率論の枠組みで捉え、分析対象としているゲームがマルチンゲールであるならば、例にあるように分析が簡単になることがある。
- 具体的には、セクション3の繰り返しゲーム理論では過去の行動がフィルトレーションとなり、マルチンゲールとなる場合がある。セクション4の進化ゲーム理論では、展開形ゲーム理論はMarkov連鎖として捉えることができた。これを過去の全ての行動に依存するモデルに拡張し、このゲームはマルチンゲールであった。
- 有限回繰り返し囚人のジレンマゲームにおいては、平均を考えていたので、これの最適値からズレが協調の可能性を示している。一種の限定合理性。

FUTURE WORKS

- **後半部分は、既存の確率的進化ゲーム理論** (Kandori, *et al*. (1993), Young (1993) **など**)**との関連を調べる。**
- **記憶期間の問題**

 1**期前か、完全記憶かの**2**者択一だった。それを**N**期間の場合を考える** (Sabourian 1998)**。**

→**この問題は繰り返し囚人のジレンマゲームにおいて協調行動の発生がシミュレーションによる分析によって知られている。**

FAQ.

- Q1) 後半の研究は、展開形ゲームの行動戦略についての研究であり、進化ゲーム理論とは関係ないのでは？

A1) 展開形ゲーム理論はBackward InductionでNash均衡を求めますが、今回の報告は、 Backward Inductionによる意思決定を行っていません。このゲームの定常分布を調べました。ただ、進化的な枠組み(ある戦略の適応度が高ければ、次期のある戦略を採用する人が多くなる)を組み入れていないので、進化ゲーム理論というのは言いすぎかもしれません。ただ先行文献であるNowak(1990)はこのモデルを進化ゲームとしている。またKandor, *et al.* (1993) やYoung(1993) はよく摂動ありのマルコフ連鎖のゲームと言われているので、このゲームを確率的進化ゲーム理論としました。

REFERENCES

- Abreu, Dilip (1988): Econometrica, **56**, 383.
- Aumann, R. J. (1974):JME, **1**, 67.
- Aumann, R.J. (1989), *Lectures on game theory,* Westview Press.
- Aumann, R. J. and Sorin, S. (1989): GEB, **1**, 5.
- Aumann, R. J. and Maschler, M. B. with the collaboration of R. B. Stearns (1995): *Repeated Games with Incomplete Information*, MIT Press.
- Cole, H.L. and Kocherlakota, N. R. (2005): GEB, **53**, 59.
- Doob, J.L. (1953): *Stochastic Processes*,John Wiley & Sons, Inc.
- Dynkin, E.B. (1969):Soviet Mathe. Dokl.,**10**, 270.
- Ellison, G. (1993):Econometrica, **61**, 1047.

- Friedman, J. W. (1985): JET, **35**, 390.
- Fudenberg, D. and Levine, D. K. (1992): RES, **59**, 561.
- Harsanyi, J. and Selten, R. (1988): *A General Theory of Equilibrium Selection in Games*, MIT Press.
- Hofbauer, J.and Sigmund, K. (1998): *Evolutionary Game and Population Dynamics*, Cambridge U P.
- **今井晴雄**, **岡田章** (2002): **「ゲーム理論の新展開」勁草書房**.
- Jensen, M., Sloth, B., and Whitta-J, Hans J. (2005): Economic Theory, **25**, 171.
- Kakutani, S.(1948): The Annals of Mathematics, 2nd Ser., **49**, 214.
- Kandori, M., Mailath, G. J. and Rob, R. (1993):Econometrica, **61**, 29.

- **吉川満** (2008): **『北海道大学数学講究録』 ,126,** 173.
- **河野敬雄** (2003): 『Rokko Lectures in Mathematics』 **13**.
- Kuhn, H.W. (1953): in H.Kuhn and A.Tucker (eds.), *Contributions to the Theory of Games*, Vol.II, Princetion UP, 193.
- Lehrer, Ehud (1988): JET, **46,** 130.
- Mailath, G. J. and Samuelson, L.(2006): *Repeated Games and Reputations - Long-Run Relationships*, Oxford U P.
- Matsui, A. and Matsuyama, K. (1995): JET **65**,415.
- Nowak, M. A. (1990): TPB, **38**, 93.
- **岡田章** (1996): **「ゲーム理論」有斐閣**, 1996**年**.
- Rapoport, A. and Chammah, A. M. , *Prisoner's dilemma: A Study in Conflict and Cooperation*, The University of Michigan Press, 1965.
- Sabourian, H.(1998): JME, **30**, 1.

- Selten, R. (1975): IJGT,**4**, 25.
- Sorin, S. (2002): *A First Course on Zero-Sum Repeated Games*, Springer-Verlag.
- Di Tillio, A. 2004): Mimeo.
- Williams, David (1991): *Probability with Martingales*, Cambridge U P.
- Young, H.P.(1993): Econometrica, **61**, 57.

Acknowledgements