数理解析研究所講究録
第 798 巻 1992 年 191-201

尤度比を用いた順序と多段決定問題について

中井　達

九州大学経済学部

1 Introduction

不完備情報のマルコフ決定過程における optimal policy およびその optimal policy に従って最適に振る舞ったときに得られる値について考える。ここでは Markov chain の状態に関して不完備な情報しか与えられておらず、prior information は Markov chain の状態空間の上に定義された確率ベクトルとして表されるものと考え、Markov chain の状態と information process から観測できる値を表す確率変数との関係は既知とする。Markov chain の状態に関する学習はベイズの方法に従って行う場合について考える。

まず始めに、Markov chain の状態空間上の確率ベクトル全体に likelihood ratio の考えを用いた順序、すなわち likelihood ratio ordering を導入し、いくつかの仮定のもとで、事前分布と事後分布の間の基本的な関係を考える。それらの性質は Sequential Decision Problem を考える上で基本的なものである。(Nakai [4]・[5]・[6])

次に、shifted likelihood ratio ordering と呼ばれる順序について考える。この順序に関してもいくつかの仮定のもとで、likelihood ratio ordering の場合と同様に、prior information と posterior information の間に基本的な関係が成り立つ。

最後に、ここで考える二つの順序のもとで、これらの順序と複数回停止する事が出来る Optimal Stopping Problem との関係にについて考える。

2 Formulation

次のような information process を考える。$\{Y_t; t = 1, 2, ...\}$ を状態空間が $\{1, 2, 3, ...\}$ である Markov chain とし、Markov chain の transition probability は既知であるとし、それを $\mathbf{P} = (p_{ij}), \{i, j = 1, 2, ...\}$ であるとする。いま $\mathcal{S}$を $\mathcal{S} = \{\mathcal{P} | \mathcal{P} = (p_1, p_2, ...), p_i \geq 0, \sum_{i=1}^{\infty} p_i = 1\}$ によって定義されているとする。このとき、この Markov chain の状態がどのようなものであるかについての情報は、全て$\mathcal{S}$に含まれる確率分布 $\mathcal{P}$によって表されるものと考える。

Markov chain の状態に関する情報は、information process $\{X_{it}; i, t = 1, 2, ...\}$ を通して得られるものと考える。ここで X_{it}は時点 t での Markov chain の状態に依存した値をと

る実数値確率変数であり、その確率分布関数は decision-maker に対し既知であるものと考え、$\Pr\{X_{it} \le x | Y_t = k\} = F_k(x), x \in R, k \in \{1, 2, ...\}$, とする。($i, t = 1, 2, ...$)。ここで、逐次に観測することの出来る確率変数は $X_{i1}, X_{i2}, ...$ であるとする。

この部分観測が可能な Markov chain の上での次のような Optimal Stopping Problem を取り上げ、その optimal policy と求めた optimal policy に従って得られる値と学習の過程との関係について考える。このとき Markov chain の状態に関する事前の情報がこの Markov chain の状態空間上の確率分布 $\mathcal{P}(\in \mathcal{S})$ によって表され、残りの決定期間が N であり、そのうち k 回停止する事ができるという状態を、ここで考える Optimal Stopping Problem の状態と言い、$(\mathcal{P}, k, N)$ により表す。この問題の状態が $(\mathcal{P}, k, N)$ であるとき、この問題を $P_N^k(\mathcal{P})$ であらわし、問題 $P_N^k(\mathcal{P})$ において最適に振る舞って得られる total expected reward を $v_N^k(\mathcal{P})$ とする。

問題の状態が $(\mathcal{P}, k, N)$ であるとき、decision-maker は Markov chain の状態に依存する確率変数 $\{X_t\}$ の実現値 x を information process より観測し、値 x をもとにして、decision-maker は利得 $u(\mathcal{P}, x)$ を得て停止するか、それとも見逃して次の期に進むか決定を行う。もし、値 x を観測したこの期で停止すれば、decision-maker は利得 $u(\mathcal{P}, x)$ を得、次の期に進む。そして次の期の状態は $(\overline{T(\mathcal{P}, x)}, k-1, N-1)$ となる。ここで $\overline{T(\mathcal{P}, x)}$ は、information process の確率変数 $\{X_t\}$ の実現値 x を観測したとき、その値をもとにして Markov chain の状態についてベイズの定理に従って学習を行ったときの事後分布を表し、(1) 式と (2) 式で求められる。もし、観測値 x に対して、この値を見送り次の期に進む事にすれば、利得を得る事はなく、停止した場合と同じように Markov chain の状態に関して学習を行い次の期に進み、問題の状態は $(\overline{T(\mathcal{P}, x)}, k, N)$ となる。$N = k$ ならば、常に利得 $u(\mathcal{P}, x)$ を各期毎に得る。

このとき、Markov chain の状態に関して information process から観測値 x を得たときの事後分布は、ベイズの定理に従って学習を行うから、まず観測値 $x \in R = (-0, \infty)$ に対して Markov chain の状態に関する事後確率分布が全ての $j = 1, 2, ...$ に対して、次のように求められる。

$$\begin{cases} T_j(\mathcal{P}, x) &= \dfrac{p_j f_j(x)}{\sum_{i=1}^{n} p_i f_i(x)} \\ T(\mathcal{P}, x) &= (T_1(\mathcal{P}, x), T_2(\mathcal{P}, x), ...) \end{cases} \tag{1}$$

次に、遷移行列 $\mathbf{P}$ に従って状態が遷移するから、次の期の始めの Markov chain の状態に関する prior information $\overline{T(\mathcal{P}, x)}$ は、

$$\begin{cases} \overline{T_j(\mathcal{P}, x)} &= \sum_{i=1}^{n} T_i(\mathcal{P}, x) p_{ij}, \\ \overline{T(\mathcal{P}, x)} &= (\overline{T_1(\mathcal{P}, x)}, \overline{T_2(\mathcal{P}, x)}, ...) \end{cases} \tag{2}$$

のように求められる。

3 Likelihood Ratio Ordering

この節では、$\mathcal{S}$に含まれる Markov chain の状態空間上の確率分布に対して、likelihood ratio を用いて次のような順序を導入する事を考える。

Definition 1 $\mathcal{S}$に含まれる任意の$\mathcal{P}$と$\mathcal{Q}$に対して、$\mathcal{P} >_l \mathcal{Q}$であるとは、全ての$i$と$j$（$j \leq i, i, j = 1, 2, ...$）に対して

$$p_i q_j \leq p_j q_i, \text{ すなわち } \begin{vmatrix} q_i & q_j \\ p_i & p_j \end{vmatrix} \geq 0 \tag{3}$$

が成立し、少なくとも一つのiとjの組み合わせに対して、$p_i q_j < p_j q_i$, が成り立つ場合を言う。もし、$p_i = q_i$、が任意の$i = 1, 2, ...$に対して成り立つとき$\mathcal{P} =_l \mathcal{Q}$であるとする。また、$\mathcal{P} \geq_l \mathcal{Q}$であるとは、$\mathcal{P} =_l \mathcal{Q}$かつ、$\mathcal{P} >_l \mathcal{Q}$が成り立つことを言う。

上で定義した順序を likelihood ratio ordering と呼び、この順序に関して以下のような性質が成り立つ。

Lemma 1 定義*1*において定義した、$\mathcal{S}$における順序は半順序である。

Lemma 2 $\mathcal{S}$に含まれる任意の$\mathcal{P}$と$\mathcal{Q}$に対して、$\mathcal{P} \geq_l \mathcal{Q}$であるならば、次の不等式を満足するような$1 < k$が存在する。$\dfrac{p_k}{q_k} > 1 \geq \dfrac{p_{k+1}}{q_{k+1}}$

Lemma 3 $\mathcal{S}$に含まれる任意の$\mathcal{P}$に対して、$(1, 0, ..., 0) \geq_l \mathcal{P}$かつ$\mathcal{P} \geq_l (0, ..., 0, 1)$である。

次に、この節で定義した likelihood ratio による順序すなわち likelihood ratio ordering のもとでの決定問題を考えるために、ここで確率分布関数と確率遷移行列$\mathbf{P}$に関して、次のような4つの仮定を設ける。

Assumption 1 Markov chain の状態がkであるとき、この状態kに依存した確率変数の条件付き期待値は有界であり、それをμ_kで表す。また、確率分布関数$F_k(x)$は確率密度関数$f_k(x)$を持つとし、$dF_k(x) = f_k(x)\,dx$とする。

Assumption 2 もし、$j \leq i\,(i, j = 1, 2, ..., n)$であれば$x \leq y$に対して

$$f_j(x) f_i(y) \leq f_j(y) f_i(x) \text{ すなわち } \begin{vmatrix} f_i(x) & f_i(y) \\ f_j(x) & f_j(y) \end{vmatrix} \geq 0 \tag{4}$$

を満足する。

Assumption 3 $1 \leq m < k,$ であれば、$x_{mk} = \sup\{\, x | f_m(x) \leq f_k(x)\},$ となるx_{mk}で、次の条件を満足するものが存在する。

(1) $x > x_{mk}$ ならば $f_m(k) > f_k(x) > f_{k+1}(x) > .....$

(2) $x \leq x_{mk}$ ならば $f_k(x) \geq f_m(x) \geq f_{m-1}(x) \geq ..... \geq f_1(x)$

Assumption 4 任意の $i, j (i \geq j, i, j = 1, 2, ...)$ に対して、

$$p_{mj} p_{ni} \geq p_{nj} p_{mi} \text{ すなわち } \begin{vmatrix} p_{ni} & p_{nj} \\ p_{mi} & p_{mj} \end{vmatrix} \geq 0 \tag{5}$$

が　$m \leq n (m, n = 1, 2, ...)$ を満足する全ての m と n に対して成り立つ。

ここで、(1) 式および (2) 式で与えられた posterior information の性質を、Assumption 1 から Assumption 4 のもとで考えてみる。まず次のような性質が成り立つことがわかる。

Lemma 4 $x \leq y$を満たす任意の x と yに対して $T(\mathcal{P}, x) \leq_l T(\mathcal{P}, y)$ が全ての $\mathcal{P} \in S_n$に対して成り立つ

Proposition 1 $x \leq y$を満たす任意の x と yに対して$\overline{T(\mathcal{P}, x)} \leq_l \overline{T(\mathcal{P}, y)}$ が全ての $\mathcal{P} \in S_n$ に対して成り立つ

Lemma 5 $\mathcal{S}$に含まれる任意の $\mathcal{P}, \mathcal{Q}$ に対して、$\mathcal{P} \geq_l \mathcal{Q}$ であるならば、$T(\mathcal{P}, x) \geq_l T(\mathcal{Q}, x)$ が全ての $x \in \mathbf{R}$ に対して成り立つ。

Proposition 2 $\mathcal{S}$に含まれる任意の $\mathcal{P}, \mathcal{Q}$ に対して、$\mathcal{P} \geq_l \mathcal{Q}$ であるならば、$\overline{T(\mathcal{P}, x)} \geq_l \overline{T(\mathcal{Q}, x)}$ が全ての $x \in \mathbf{R}$ に対して成り立つ。

Lemma 6 今、$\{f_i(x)\}_{i=1,2,\ldots,n}$を、Assumption 3 の性質を満足するような、確率密度関数の有限列と考える。($n \geq 2$) 次に、$\{a_i\}_{i=1,2,\ldots,n}$を、実数の列とし、$\sum_{i=1}^{n} a_i = 0$ を満足するものとする。

このとき、適当な $1 \leq k < n$ が存在し、$a_i \geq 0, \quad j \leq k$, および $a_j \leq 0, \quad j > k$ であるとする。

以上の仮定のもとで、いま $g(x) = \sum_{i=1}^{n} a_i f_i(x)$ とおけば、

$$\int_0^\infty h(x) g(x) dx \geq 0$$

が成り立つ。

Corollary 1 今、$\{f_i(x)\}_{i=1,2,.....}$を、Assumption 3 の性質を満足するような、確率密度関数の列と考える。次に、$\{a_i\}_{i=1,2,.....}$を、実数の列とし、$\sum_{i=1}^{\infty} a_i = 0$ を満足するものとする。

このとき、適当な $1 \leq k < n$ が存在し、$a_i \geq 0, \quad j \leq k$, および $a_j \leq 0, \quad j > k$ であるとする。以上の仮定のもとで、いま $g(x) = \sum_{i=1}^{\infty} a_i f_i(x)$ とおけば、

$$\int_0^\infty h(x) g(x) dx \geq 0$$

が成り立つ。

Lemma 7 関数 $u(\mathcal{P})$ が $\mathcal{P}$に関して増加かつ convex な関数であれば

$$u^*(\mathcal{P}) = \int_0^\infty u(\overline{T(\mathcal{P}, x)}) dF_{\mathcal{P}}(x)$$

もまた $\mathcal{P}$に関して増加かつ convex な関数である。

4 Shifted Likelihood Ratio Ordering

状態空間$\mathcal{S}$上の確率ベクトル全体に、前節で定義した likelihood ratio ordering とは異なる新しい順序を導入し、同様の問題を改めて考える。まず、次の順序を考える。

Definition 2 $\mathcal{S}$に含まれる二つの確率ベクトル$\mathcal{P}$と$\mathcal{Q}$に対して、$\mathcal{P} >_s \mathcal{Q}$であるとは、任意の整数$i$と$j$($i \leq j$)に対して、$a \geq 0$ならば

$$p_j q_{i+a} \geq p_i q_{j+a} \text{ すなわち } \begin{vmatrix} q_{i+a} & q_{j+a} \\ p_i & p_j \end{vmatrix} \geq 0 \tag{6}$$

が成り立ち、少なくとも一つのiとjの組に対して、$p_j q_{i+a} > p_i q_{j+a}$が成り立つ事をいう。また、全ての$i$に対して、$p_i = q_i$であるとき$\mathcal{P} =_s \mathcal{Q}$ということにする。次に、$\mathcal{P} \geq_s \mathcal{Q}$であるとは、$\mathcal{P} >_s \mathcal{Q}$または、$\mathcal{P} =_s \mathcal{Q}$が成り立つ場合をいう。最後に、全ての$\mathcal{P} \in \mathcal{S}$に対して、$\mathcal{P} \geq_s \mathcal{P}$であるとする。

この定義2によって定められた順序を、shifted likelihood ratio ordering という。また、$k = 0$ならば、この定義は likelihood ratio ordering に等しくここで定義した順序は likelihood ratio ordering の条件を満足する。この順序については、Shanthikumar and Yao [8] および Keilson and Sumita [3] において、いろいろな性質が調べられている。この節では、この順序に対しても第3節で考えたものと同様の問題を改めて考える。第3節で考えたようないくつかの仮定のもとで、この順序の性質に関して考える。また、ここでは、Assumption 1 と Assumption 3 は、先の likelihood ratio ordering と同様に成り立つものと考える。また、$\mathcal{P} \geq_s \mathcal{P}$であることを仮定したことから、$\mathcal{P}$が PF_2 (Polya frequency of order 2) である。ここに、PF_2であるとは、この場合では$\dfrac{p_{i-1}}{p_i}$がiに関して増加する関数である事を表す。この PF_2については、Shanthikumar and Yao [8] 等にも述べられている。

Assumption 5 もし、$j < i\,(i, j = 1, 2, ...)$であれば$x \leq y$に対して、$a \geq 0$ならば

$$f_j(y) f_{i+a}(x) \geq f_i(y) f_{j+a}(x) \text{ すなわち } \begin{vmatrix} f_{i+a}(x) & f_{j+a}(x) \\ f_i(y) & f_j(y) \end{vmatrix} \geq 0 \tag{7}$$

を満足する。

Assumption 6 任意の$i, j\,(i \geq j, i, j = 1, 2, ...)$に対して、$k \geq 0$ならば

$$p_{n\,i+k} p_{m\,j} \geq p_{n\,j+k} p_{m\,i} \text{ すなわち } \begin{vmatrix} p_{n\,i+k} & p_{n\,j+k} \\ p_{m\,i} & p_{m\,j} \end{vmatrix} \geq 0 \tag{8}$$

が　$m \leq n\,(m, n = 1, 2, ...)$を満足する全ての$m$と$n$に対して成り立つ。

Assumption 7 任意の$i, j(i \geq j, i, j = 1, 2, ...)$に対して、$a \geq 0$ならば

$$\begin{vmatrix} p_{n\,i+a} & p_{n\,j+a} \\ p_{m\,i} & p_{m\,j} \end{vmatrix} = -\begin{vmatrix} p_{m\,i+a} & p_{m\,j+a} \\ p_{n\,i} & p_{n\,j} \end{vmatrix} \geq 0 \tag{9}$$

が　$m \leq n\,(m, n = 1, 2, ...)$を満足する全ての$m$と$n$に対して成り立つ。

Remark 1 確率変数 $\{X_i\}$ の確率密度関数に関して、これらの確率変数に shifted likelihood ratio ordering を導入すれば、もし、$j < i\,(i, j = 1, 2, ...)$ と $x \le y$ に対して、$a \ge 0$ ならば

$$f_j(x)f_i(y+a) \le f_j(y)f_i(x+a) \text{ すなわち } \begin{vmatrix} f_i(x+a) & f_i(y+a) \\ f_j(x) & f_j(y) \end{vmatrix} \ge 0 \tag{10}$$

としなければならないが、ここでは事後分布の性質を考える上で Assumption 5 を用いる。

ここで定義した shifted likelihood ratio ordering について、Assumption 3 を仮定した事により likelihood ratio ordering の場合に成立した期待値に関する基本的な二つの Lemma 、すなわち Lemma 6 と Lemma 7 が、この場合にも成り立つ事がいえる。従って、shifted likelihood ratio ordering に関しても likelihood ratio ordering の場合と同じように Lemma 6 と Lemma 7 を用いる事が出来る。

Lemma 8 定義 *1* において定義した、$\mathcal{S}$における順序は半順序である。

Lemma 9 shifted likelihood ratio ordering に対して、次のような性質を持つ。$\mathcal{S}$に含まれる任意の $\mathcal{P}$と $\mathcal{Q}$ に対して、$\mathcal{P} \ge_s \mathcal{Q}$ が成り立つならば、次の不等式を満足するような $1 < k$ が存在する。

$$\frac{p_k}{q_k} > 1 \ge \frac{p_{k+1}}{q_{k+1}}.$$

$\mathcal{P}$上の関数 $u(\mathcal{P})$ が、shifted likelihood rayio ordering に関して増加関数であるとは、likelihood ratio ordering の場合と同じように、$\mathcal{P} \ge_s \mathcal{Q}$ となる $\mathcal{P}$と $\mathcal{Q}$ に関して $u(\mathcal{P}) \ge u(\mathcal{Q})$ を満足するときを言う事にする。

ここでは、likelihood ratio ordering の場合と同じように reward function $u(\mathcal{P}, x)$ は、$\mathcal{P}$ に関して非減少、convex な関数であり非負で有限な値を取る関数とする。また、x に関して増加関数であるとする。

Lemma 10 $x \le y$を満たす任意の x と yに対して $T(\mathcal{P}, x) \le_s T(\mathcal{P}, y)$ が全ての $\mathcal{P} \in S_n$に対して成り立つ

Proposition 3 $x \le y$を満たす任意の x と yに対して$\overline{T(\mathcal{P}, x)} \le_s \overline{T(\mathcal{P}, y)}$ が全ての $\mathcal{P} \in S_n$ に対して成り立つ

Lemma 11 $\mathcal{S}$に含まれる任意の$\mathcal{P}, \mathcal{Q}$ に対して、$\mathcal{P} \ge_s \mathcal{Q}$であるならば、$T(\mathcal{P}, x) \ge_s T(\mathcal{Q}, x)$ が全ての $x \in \mathbf{R}$ に対して成り立つ。

Proposition 4 $\mathcal{S}$に含まれる任意の $\mathcal{P}, \mathcal{Q}$ に対して、$\mathcal{P} \ge_s \mathcal{Q}$ であるならば、$\overline{T(\mathcal{P}, x)} \ge_s \overline{T(\mathcal{Q}, x)}$ が全ての $x \in \mathbf{R}$ に対して成り立つ。

この節でみてきたように shifted likolihood ratio ordering においても、likelihood ratio ordering の場合と同様に適当な仮定を設ければ、Proposition 3 および Proposition 4 という事後確率分布に関する基本的な性質が成り立つから、この場合においても likelihood ratio ordering の場合と同様に Sequential Decision Problem に応用する事が出来る。

5 Partially Obervable Optimal Stopping Problem

不完備情報の Sequential Decision Problem の中で、Optimal Stopping Problem $P_N^k(\mathcal{P})$ について考える。likelihood ratio ordering に関する結果は、適当な仮定のもとで shifted likelihood ratio ordering においても同じように得られるから、この節で述べる問題に関する性質はそれぞれの順序に対応する仮定のもとで、同じように求めることが出来る。以下では、likelihood ratio ordering の場合に関した結果について述べる。likelihood ratio ordering の場合に関する Sequential Decision Problem については、Nakai[4] および [5] において特別な場合について述べられている。

この問題は Ross [7] で扱われている問題と同様に定式化でき、optimal policy のもとで最適に振る舞って得られる値 $v_N^k(\mathcal{P})$ は、次の recursive function を満足する。

$$\begin{aligned} v_N^k(\mathcal{P}) &= E[v_N^k(\mathcal{P}|X)] \\ &= \int_0^\infty v_N^k(\mathcal{P}|x)dF_\mathcal{P}(x) \qquad (11) \\ v_N^k(\mathcal{P}|x) &= \max\{u(\mathcal{P},x) + v_{N-1}^{k-1}(\overline{T(\mathcal{P},x)}),\ v_{N-1}^k(\overline{T(\mathcal{P},x)})\} \qquad (12) \end{aligned}$$

また、boundary condition は次の式で表される。

$$\begin{aligned} v_N^N(\mathcal{P}) &= E_\mathcal{P}[u(\mathcal{P},X)] + \int_0^\infty v_{N-1}^{N-1}(\overline{T(\mathcal{P},x)})dF_\mathcal{P}(x), \\ v_1^1(\mathcal{P}) &= E_\mathcal{P}[u(\mathcal{P},X)], \quad v_N^0(\mathcal{P}) = 0 \end{aligned}$$

ここで、$F_\mathcal{P}(x) = \sum_{i=1}^n p_i F_i(x)$ とする。この関数は weighted distribution function として知られているものである。(de Vylder [2])

複数回停止可能な Optimal Stopping Problem の optimal policy 、及び optimal policy に従ったとき得られる total expected reward の性質について考える。以下の Proposition にこれらの性質が述べられており、これらの結果は kに関する帰納法によって証明する事が出来る。すなわち、すべての Nに対して、$k-1$ およびそれ以下の値に対して、次の 4 つの Proposition が成り立つ事を仮定し、kの場合に Nに関する帰納法によって次に述べる 4 つの Proposition を証明する。

いま、ここで、全ての $k = 1, 2, ..., N$に対して、関数 $w_N^k(\mathcal{P})$ を

$$\begin{aligned} w_N^k(\mathcal{P}) &= v_N^k(\mathcal{P}) - v_N^{k-1}(\mathcal{P}) \qquad (13) \\ w_N^k(\mathcal{P}) &= E_\mathcal{P}[w_N^k(\mathcal{P}|X)] \\ w_N^k(\mathcal{P}|x) &= v_N^k(\mathcal{P}|x) - v_N^{k-1}(\mathcal{P}|x) \qquad (14) \\ w_N^1(\mathcal{P}|x) &= v_N^1(\mathcal{P}|x), \quad w_N^1(\mathcal{P}) = v_N^1(\mathcal{P}) \end{aligned}$$

によって定義する。このとき次の性質が成り立つ。

Proposition 5 関数 $v_N^k(\mathcal{P})$ は、$\mathcal{P}$に関して増加かつ convex な関数であり、$w_N^k(\mathcal{P})$ は $\mathcal{P}$に関して非負かつ増加な関数である。また、関数 $v_N^k(\mathcal{P}|x)$ および、$w_N^k(\mathcal{P}|x)$ は、x に関して非負かつ増加な関数である。

Proposition 6 $N > 1$ かつ、$N \geq k \geq 1$ であれば

$$v_N^k(\mathcal{P}) \geq v_{N-1}^k(\mathcal{P}), \quad w_N^k(\mathcal{P}) \geq w_{N-1}^k(\mathcal{P}) \tag{15}$$

$$v_N^k(\mathcal{P}|x) \geq v_{N-1}^k(\mathcal{P}|x), \quad w_N^k(\mathcal{P}|x) \geq w_{N-1}^k(\mathcal{P}|x) \tag{16}$$

が全ての $\mathcal{P}$ に対して成立する。

Proposition 7 $N \geq 1$ かつ、$n \geq k \geq 1$ であれば

$$w_N^k(\mathcal{P}) \leq w_N^{k-1}(\mathcal{P}) \leq .... \leq w_N^2(\mathcal{P}) \leq v_N^1(\mathcal{P})$$

$$w_N^k(\mathcal{P}|x) \leq w_N^{k-1}(\mathcal{P}|x) \leq .... \leq w_N^2(\mathcal{P}|x) \leq v_N^1(\mathcal{P}|x)$$

が全ての $\mathcal{P}$ に対して成立する。

まず始めに二つの集合

$$\begin{cases} S_{N+1}^k(\mathcal{P}) & = \{x \mid u(\mathcal{P}, x) \geq \bar{w}_N^k(\mathcal{P}|x)\} \\ C_{N+1}^k(\mathcal{P}) & = \{x \mid u(\mathcal{P}, x) < \bar{w}_N^k(\mathcal{P}|x)\} \end{cases} \tag{17}$$

を、上のように定義する。ただし、$N \geq 1$ および $N \geq k \geq 1$ に対して

$$\bar{w}_N^k(\mathcal{P}|x) = w_N^k(\overline{T(\mathcal{P}, x)})$$

とする。このとき、集合 $S_{N+1}^k(\mathcal{P})$ は、問題 $P_{N+1}^k(\mathcal{P})$ に対する stopping region を表し、$C_{N+1}^1(\mathcal{P})$ は continuance region を表している。従って、(12)式は次のように書き改める事が出来る。

$$v_N^k(\mathcal{P}|x) = (u(\mathcal{P}, x) + \bar{v}_{N-1}^{k-1}(\mathcal{P}|x))I_{S_N^k(\mathcal{P})} + \bar{v}_{N-1}^k(\mathcal{P}|x)I_{C_N^k(\mathcal{P})}, \tag{18}$$

ここで、関数 I_A は、集合 $A \in \mathcal{S}$ の Indicator function であるとし、全ての $N \geq 1$ および $N \geq k \geq 1$ に対して

$$\bar{v}_N^k(\mathcal{P}|x) = v_N^k(\overline{T(\mathcal{P}, x)}) \tag{19}$$

とおく。

Proposition 8 $N \geq 1$ および、$N \geq k \geq 1$ であれば

$$\text{(a)} \quad S_{N+1}^k(\mathcal{P}) \subset S_N^k(\mathcal{P}),$$

$$\text{(b)} \quad S_{N+1}^k(\mathcal{P}) \supset S_{N+1}^{k-1}(\mathcal{P})$$

が成り立つ。

Proposition 9 関数 $u(\mathcal{P}, x)$ が $\mathcal{P}$ に関して独立である、すなわち $u(\mathcal{P}, x) = u(x)$ と表されるとき、任意の $\mathcal{P}$、$\mathcal{Q}$ に対して $\mathcal{P} \geq_l \mathcal{Q}$ であれば $S_{N+1}^k(\mathcal{P}) \subset S_{N+1}^k(\mathcal{Q})$ である。

6 Optimal Policy および Optimal Value

複数回停止することの出来る Optimal Stopping Problem の optimal policy を決定づける二つの集合すなわち、stopping region $S_N^k(\mathcal{P})$ および continuance region $C_N^k(\mathcal{P})$ の性質について考えた。ここでは、optimal policy の性質と optimal policy のもとで最適に振る舞ったときに得られる total expected reward の値について考える。

そのために始めに次のような二つの non-negative measurable function $u(x)$ および $v(x)$ に対して、二つの関数 $U_F(u(x), g(x))$ および $V_F(u(x), g(x))$ を次のように定義する。

$$U_F(u(x), g(x)) = \int_0^\infty (u(x) - g(x))^+ dF(x) \tag{20}$$

$$V_F(u(x), g(x)) = \int_0^\infty g(x)dF(x) + U_F(u(x), g(x)) \tag{21}$$

ただし、$h(x)^+ = \max\{h(x), 0\}$ とする。これらの関数は DeGroot[1] などで定義されたよく知られた関数

$$T_F(z) = \int_0^\infty (x - z)dF(x), \text{および} \quad S_F(z) = z + T_F(z)$$

の一般化と考える事が出来るものである。すなわち、もし $u(x) = x$ および $g(x) = z$とすれば、

$$U_F(u(x), g(x)) = T_F(z) \text{ かつ } V_F(u(x), g(x)) = S_F(z)$$

となる。

これらの関数を用いて、non-negative function の列 $\{g_{N,i}(\mathcal{P})\}$ ($\mathcal{P} \in \mathcal{S}$かつ $1 \leq i \leq N$) を次のように帰納的に定義する。

$$g_{N,i}(\mathcal{P}) = V_{F_\mathcal{P}}(u(\mathcal{P}, x), g_{N-1,i}(\overline{T(\mathcal{P}, x)}) - U_{F_\mathcal{P}}(u(\mathcal{P}, x), g_{N-1,i-1}(\overline{T(\mathcal{P}, x)})) \tag{22}$$

$$g_{N,0}(\mathcal{P}) = \infty \quad \text{かつ} \quad g_{N,N+1}(\mathcal{P}) = 0\,(N \geq 0)$$

ここで、次のような集合を考える。

$$S_{N,i}(\mathcal{P}) = \{x | g_{N-1,i}(\overline{T(\mathcal{P}, x)}) \leq u(\mathcal{P}, x) < g_{N-1,i-1}(\overline{T(\mathcal{P}, x)})\} \tag{23}$$

$$U_{N,i}(\mathcal{P}) = \bigcup_{j=1}^{i-1} S_{N,j}(\mathcal{P}) \quad \text{かつ} \quad L_{N,i}(\mathcal{P}) = \mathbf{R}_+ - U_{N,i+1}(\mathcal{P})$$

ただし、ここに $U_{N,1}(\mathcal{P}) = L_{N,N}(\mathcal{P}) = \phi$かつ $U_{N,N+1}(\mathcal{P}) = \mathbf{R}_+$とする。

一方

$$g_{1,1}(\mathcal{P}) = \sum_{j=1}^m p_j \int_0^\infty u(\mathcal{P})\, dF_j(x)$$

であるから、(22) 式によって生成した関数列は定義できる。この関数列に関して次のような性質が成り立つ。

Proposition 10 関数 $g_{N,i}(\mathcal{P})$ は、任意の整数 Nおよび i に対して、$\mathcal{P}$に関して増加する関数である。

Proposition 11 関数 $g_{N,i}(\mathcal{P})$ は、任意の $\mathcal{P} \in \mathcal{S}$ および整数 N に対して、i に関して減少する関数である。

Proposition 12 関数 $g_{N,i}(\mathcal{P})$ は、任意の $\mathcal{P} \in \mathcal{S}$ および整数 i に対して、N に関して増加する関数である。

Corollary 2 三つの集合 $S_{N+1,i}(\mathcal{P})$ 、$U_{N+1,i}(\mathcal{P})$ および $L_{N+1,i}(\mathcal{P})$ は、互いに素であり、また $S_{N+1,i}(\mathcal{P}) \cup U_{N+1,i}(\mathcal{P}) \cup L_{N+1,i}(\mathcal{P}) = R_+$ である。

Corollary 3 いま、関数 $h_{N+1,i}(\mathcal{P}|x)$ を、次のように定義する。

$$\begin{aligned} h_{N+1,i}(\mathcal{P}|x) &= g_{N,i-1}(\overline{T(\mathcal{P},x)})I_{U_{N+1,i}(\mathcal{P})}(x) + u(\mathcal{P},x)I_{S_{N+1,i}(\mathcal{P})(x)} \\ &\quad + g_{N,i}(\overline{T(\mathcal{P},x)})I_{L_{N+1,i}(\mathcal{P})}(x) \end{aligned} \tag{24}$$

このとき、関数 $g_{N+1,i}(\mathcal{P})$ は、次のように表せる。

$$g_{N+1,i}(\mathcal{P}) = \int_0^\infty h_{N+1,i}(\mathcal{P}|x)\, dF_{\mathcal{P}}(x) \tag{25}$$

すなわち、関数 $h_{N+1,i}(\mathcal{P}|x)$ は、関数 $g_{N+1,i}(\mathcal{P})$ の integrand である。

Proposition 13 任意の $\mathcal{P} \in \mathcal{S}$ と $i\,(1 \geq i \geq N)$ に対して

$$U_{N+1,i}(\mathcal{P}) \subset U_{N,i}(\mathcal{P})$$

である。

Proposition 14 $u(\mathcal{P},x) = u(x)$ である、すなわち関数 $u(\mathcal{P},x)$ が $\mathcal{P}$ に依存しないとき、もし、$\mathcal{P} \geq_l \mathcal{Q}$ ($\mathcal{P},\mathcal{Q} \in \mathcal{S}$) であり、$1 \geq i \geq N+1$ であれば、

$$U_{N+1,i}(\mathcal{P}) \subset U_{N+1,i}(\mathcal{Q})$$

である。

Proposition 15 問題の状態が $(N,k,\mathcal{P})$ である複数回停止が可能な optimal stopping problem $(\mathcal{P})$ で、optimal policy および optimal policy のもとで最適に振る舞って得られる total expected reward $v_N^k(\mathcal{P})$ は、次のように示される。

（１） $v_N^k(\mathcal{P}) = \sum_{i=1}^{k} g_{N,i}(\mathcal{P})$

（２） もし、information process から得られる観測値の値 x が、$x \in S_{N,i}(\mathcal{P})$ であれば、optimal policy は、この期に停止して値 x を採用し次の期に進むことであり、そうでなければ、この値を見送り次の期に進むことである。

Remark 2 Proposition 15 で、問題 $P_N^k(\mathcal{P})$ において、optimal policy に従ったときに得られる total expected reward が、$v_N^k(\mathcal{P}) = \sum_{i=1}^{k} g_{N,i}(\mathcal{P})$ で表されることから、$g_{N,i}(\mathcal{P})$ は問題 $P_N^{k-1}(\mathcal{P})$ で、停止することが出来る機会が 1 回増えることによって decision-maker の最適に振る舞ったときの期待利得の増加分を表していると考えられる。すなわち、問題 $P_N^k(\mathcal{P})$ における、最後につけ加えた停止する機会の寄与を表すと考えることもできる。

参考文献

[1] DeGroot, M. H., *Optimal Statistical Decisions*, McGraw-Hill, New York, New York, 1970.

[2] De Vylder, F., Duality Theorem for Bounds in Integrals with Applications to Stop Loss Premiums, *Scandinavian Actuarial Journal*, pp. 129-147, 1983.

[3] Keilson, J., Sumita, U., Uniform Stochastic Ordering and Related Inequalities, *The Canadian Journal of Statistics*, vol. 10, pp.181-198, 1982.

[4] Nakai, T., Optimal Stopping Problem in a Finite State Partailly Observable Markov Chain, *Journal of Information & Optimization Sciences*, vol. 2, pp. 159-176, 1983.

[5] Nakai, T., The Problem of Optimal Stopping in a Partially Observable Markov Chain, *Journal of Optimization Theory and Applications*, vol. 45, pp. 425-442, 1985.

[6] Nakai, T., A Sequential Stochastic Assignment Problem in a Partially Observable Markov Chain, *Mathematics of Operations Research*, vol. 11, pp. 230-240, 1986.

[7] Ross, S. M., *Applied Probability with Optimization Applications*, Holden-Day, San Fransisco, California, 1970.

[8] Shanthikumar, J. G., Yao, D. D., The Preservation of Likelihood Ratio Ordering under Convolution, *Stochastic Processes and their Applications*, vol. 23, pp. 259-267, 1986.