

不完備情報の多段決定問題—公的支出モデルについて

中井 達

九州大学大学院経済学研究院, nakai@en.kyushu-u.ac.jp

1 はじめに

Total positivity of order two (TP_2) は、多段決定問題、とくにベイズ学習を伴う不完備情報マルコフ過程における多段決定問題を考える上で、確率的逐次割り当て問題や、dynamic economy におけジョブ・サーチなどへの応用 (Nakai[10]) が知られているように、重要な役割を果たしている。ここでは、消防活動や警察活動などに見られるような、公的部門における投資問題を、多段決定問題ととらえて解析し、最適政策や住民の満足度や充足度といったものがどのようになるかを考える。とくに、この問題を total positivity of order two (TP_2) の仮定の下で、これらのサービスに対する住民の満足度や充足度を状態と考えたマルコフ過程に従って状態が推移するモデルとして解析する。したがって、これらの状態を改善するために、このような公共のサービスに対してどのように投資すればよいかを求めることになる。

消防活動や警察活動などの公共サービスを考えてみよう。これらの公共サービスにおいて、設備や機器を揃え、スタッフを確保していることと、住民の満足度や充足度の間の関係には、関係はあるものの、スタッフが多いことや、設備が充実していることと、満足度や充足度の間にはずれがあり、直接的な関係を見いだすことは難しい。いっぽう、これらの充足度や満足度は、社会環境や状況などによって変動するものであり、これら公共サービスに対する投資を増やすことによっても変化する。このため、充足度あるいは満足度を状態と考え、この状態は投資額によっても変化するものとし、各年度の予算の範囲でどのように投資すればよいかを決定する多段決定モデルとする。まずはじめに、この過程の状態を直接観測できる場合から始めて、不完備情報の場合まで考える。

2 節と 3 節では、満足度を表す状態がマルコフ過程に従って推移する場合を考え、4 節では部分観測可能なマルコフ過程における学習プロセスを見るために、total positivity of order two (TP_2) との関連を考える。すなわち、状態が何かを直接観測できないが、その状態に関する情報を持っている場合にあたる。この情報は、状態空間上の確率分布で与えられ、ベイズ学習に従うとする。このとき、total positivity of order two (TP_2) の性質を用いて、事前情報と事後情報の関係をはじめ、最適政策や期待利得に関する性質が求められることが知られている。この total positivity of order two (TP_2) については、確率過程に関して Karlin and McGregor [3]、Karlin [2]、Karlin and Rinott [4] などで性質が調べられている。最後に、部分観測可能なマルコフ過程における決定問題として、公共サービスに関する最適投資モデルを考える。

2 逐次支出モデル

消防サービスや警察サービスといった公共サービスに、毎年度の予算の範囲内でこのサービスに予算を投入（支出）することを考えてみよう。これらの公共サービスに対して、実際の設備や施設あるいは人員と満足度や充足度とのあいだには関連があることは確かであるが、かといって設備や施設、人員が多くなったところで、生活環境や状況が変化することによって、これら公共サービスに対する要求が増加し、満足度が低下することもある。この様な状況を表すために、満足度あるいは充足度を状態とし、これらの満足度や充足度が、マルコフ過程に従って変化するとともに、予算を追加して投入することによってこれらの状態の変化を促すことができるモデルを考える。そのため、満足度や充足度を表す状態を s であらわし、便宜的に $s \in [0, 1]$ とするが、状態空間はこれに限る必要はない。この場合、 $s = 1$ であれば公共サービスは住民にとって満足できるものであり、このサービスが充分でなくなれば s が減少するにしたがって、満足度も低下する。

いっぽう、設備や人員を配置して、この公共サービスに対する要求を満たすために、支出した額が x のとき、このサービスに対する満足度あるいは充足度は $S(x)$ となると考える。すなわち、満足度あるいは充足度 $S(x)$ を実現するために必要な投資額 x であり、投資額 x の関数とする。いいかえれば、満足度あるいは充足度が s であり、 $s = S(x)$ ならば、総額 x を投資することによってこの満足度あるいは充足度を実現することができると思う。

ここで、関数 $S(x)$ は x に関して増加関数であり凹関数とするいま、

$$x(s) = \inf\{x | S(x) \geq s, x \geq 0\}$$

とおけば、この関数 $x(s)$ は満足度あるいは充足度 s を実現するために必要な資産を表し、 $x(s)$ は満足度が s であるときの仮想的な資産量ととらえることもできる。つぎに、 $c(s, t)$ を、満足度あるいは充足度が s のとき、この状態を t と変化させるための費用を表す関数とする ($t \geq s$)。

いま、計画期間を n とし、各期ごとの予算の上限を K とすれば、この予算の上限の範囲内で設備や施設あるいは人員を増やすことによって、現在の満足度あるいは充足度 s を上昇させる問題を考える。このとき、最適政策にしたがったときに得られる期待利得を $v_n(s)$ とすれば、最適方程式は

$$v_n(s) = \max_{0 \leq x \leq K} \{-c(s, s + d_s(x)) + v_{n-1}(s + d_s(x))\} \quad (1)$$

となる。ただし、 $v_1(s) = \max_{0 \leq x \leq K} \{-c(s, s + d_s(x)) + u(s + d_s(x))\}$ 、であり、 $d_s(x) = S(x + x(s)) - s$ とする。ここで $d_s(x)$ は、満足度あるいは充足度が s のとき、予算から x だけ追加したときに、満足度あるいは充足度の増加量を表す。また、そのための支出あるいは費用は $-c(s, s + d_s(x))$ である。もし、 $-c(s, s + d_s(x)) = x$ であれば、費用は支出した量に等しい。この関数 $d_s(x)$ に対して、次の性質が成り立つ ($s \leq t$)。

補題 1 $s < t$ ならば $d_s(x) \geq d_t(x)$ である。

補題 2 $s < t$ ならば、任意の $x \geq 0$ に対して、 $s + d_s(x) \leq t + d_t(x)$ である。

例 1 満足度あるいは充足度を表す関数 $S(x)$ が $S(x) = 1 - e^{-x}$ ($x \geq 0$) とすると、満足度あるいは充足度が s のとき、 $x(s)$ は $x = x(s) = -\log(1-s)$ ($0 \leq s \leq 1$) であり、 $\frac{dS(x)}{dx} = e^{-x}$ および $\frac{dx(s)}{ds} = -\frac{1}{(1-s)}$ となる。

ここで、費用関数 $c(s, t)$ は、 t に関して増加かつ凸関数であるとし、 s に関する減少関数と仮定する。また、 $t \rightarrow s$ のとき $c(s, t) \rightarrow 0$ とする。初期条件は $v_0(t) = u(t)$ であり、終端利得 $u(t)$ は t に関して、増加かつ凸関数とする。このとき、帰納法により次の性質が簡単に導かれる。

補題 3 $v_n(s)$ は s に関する非減少関数である。すなわち、 $s \leq t$ ならば $v_n(s) \leq v_n(t)$ である。

補題 4 $v_n(s)$ は n に関して非減少関数である。すなわち、任意の $n \geq 1$ に対して、 $v_n(s) \leq v_{n+1}(s)$ である。

2.1 費用関数 $c(x)$

次に費用関数 $c(x)$ が x のみに依存する場合を考える。この場合の最適方程式は、

$$v_n(s) = \max_{0 \leq x \leq K} \{-c(x) + v_{n-1}(s + d_s(x))\}$$

であり、これまでと同じように $c(x)$ は、 x に関して増加かつ凸関数とし、つぎの仮定を設ける。

仮定 1 $d_s(x)$ は s に関する凹関数である。

前の例では、 $s(x(s+u)+x) - s(x(s)+x) = ue^{-x} = s(x(t+u)+x) - s(x(t)+x)$ であり、仮定 1 を満足する。このときつぎの性質が導かれる。

補題 5 $v_n(s)$ は s に関する凹関数である。

補題 6 計画期間が n で、満足度あるいは充足度が s のとき、最適な支出量を $x_n^*(s)$ とすれば、任意の $s \leq t$ に対して、 $x_n^*(s) \leq x_n^*(t)$ である。

注 1 一般的な費用関数 $c(t, x)$ に対して、 $c(t, x) - c(s, x)$ が任意の $s < t$ に対して、 x に関する減少関数であれば、補題 6 を導くことができる。すなわち、任意の $0 \leq x \leq x^*$ に対して $c(t, x) - c(t, x^*) \leq c(s, x) - c(s, x^*)$ であることから、補題 6 が成り立つ。

補題 7 残りの計画期間が n で、満足度あるいは充足度が s でのとき、最適な支出量を $x_n^*(s)$ とすれば、 $x_{n-1}^*(s) \geq x_n^*(s)$ である。

補題 8 $s < t$ ならば、任意の $n \geq 1$ に対して、 $v_{n-1}(t) - v_{n-1}(s) \geq v_n(t) - v_n(s)$ である。

3 逐次支出モデル: 確率モデル

ここでは、前節の逐次支出モデルにおいて、満足度あるいは充足度をあらわす状態がマルコフ過程にしたがって推移する場合を考える。すなわち、状態が制御できない確率過程にしたがって推移する場合である。いいかえれば、設備や機器、あるいは人員が多くなったとしても、予算内での追加的な支出だけでなく、ある確率過程にしたがって変化し、満足度あるいは充足度が下がる場合も認めることである。ここでは、マルコフ過程にしたがって推移する場合を考える。状態空間をこれまでと同じように $[0, 1]$ とし、推移法則を $(p_s(t))_{0 \leq s \leq 1}$ とする。はじめに、total positive of order two を、つぎのように定義する。

定義 1 集合値関数 $P = (p_s(t))_{s,t \in [0,1]}$ に対して、 $\begin{vmatrix} p_s(u) & p_s(v) \\ p_t(u) & p_t(v) \end{vmatrix} \geq 0$ が、 $s \leq t$ and $u \leq v$ となる任意の s, t, u と v に対して成り立つとき ($s, t, u, v \in [0, 1]$)、この P は total positive of order two あるいは TP_2 の性質を持つという。

このとき推移法則に関して次の仮定を設ける。

仮定 2 推移法則 $(p_s(t))_{0 \leq s \leq 1}$ は TP_2 である。

計画期間が n で、各期ごとの予算額が K とする。このとき、最適に振る舞ったときの満足度あるいは充足度に対する期待利得を $V_n(s)$ とすれば、最適方程式はつぎのようになる。

$$V_n(s) = \max_{x \geq 0} \left\{ -c(x) + \int_0^1 p_{s(x)}(t) V_{n-1}(t) dt \right\} \quad (2)$$

ただし、 $V_1(s) = \max_{x \geq 0} \left\{ -c(x) + \int_0^1 p_{s(x)}(t) u(t) dt \right\}$ であり、 $s(x)$ は、状態が s のとき、 x を追加して支出したときの新たな状態を表す。前節の場合には、 $s(x)$ は $s(x) = s + d_s(x)$ と表すことができる。ここで、 $s(x)$ は、 x に関して増加かつ凹関数であり、 s に関する増加関数とする。

補題 9 集合値関数 $P = (p_s(t))_{s,t \in [0,1]}$ と、 x に関する増加関数 $s(x)$ に対して、任意の s, t, u と v に対して ($u, v \in [0, 1]$)、 $x \leq y$ かつ $u \leq v$ であれば、任意の s ($s \in [0, 1]$) について、 $\begin{vmatrix} p_{s(x)}(u) & p_{s(x)}(v) \\ p_{s(y)}(u) & p_{s(y)}(v) \end{vmatrix} \geq 0$ となる。

つぎに、定義 2 にしたがって、total positivity の性質を用いて確率変数のあいだに半順序を導入する。ここで、確率変数は全順序 \succeq が定義された完備で可分な距離空間上で定義されているとする。

定義 2 2つの確率変数 X と Y が、それぞれ確率密度関数 $f(x)$ と $g(x)$ を持つとする。このとき、 $x \geq y$ を満たす任意の x と y にたいして、 $f(y)g(x) \leq f(x)g(y)$ であるとき、 X は Y より尤度比の意味で大きいといい、 $X \succeq Y$ と表す。

いま、 $\mathcal{F}_{SSD} = \{u \mid u(x) \text{ } x \text{ に関して増加かつ凹関数}\}$ とし、この集合を使って定義 3 により半順序を定義する。

定義 3 2つの確率変数 X と Y が、それぞれ確率密度関数 $f_X(x)$ と $f_Y(x)$ を持つとする。このとき、任意の $u(x)$ に対して \mathcal{F}_{SSD} 、 $E[u(X)] \geq E[u(Y)]$ であるとき $X \geq_{SSD} Y$ とする。

補題 10 の性質が、定義 2 と定義 3 による順序に関して成り立つ。推移法則が TP_2 のとき、補題 11 が Kijima and Ohnishi[5] で示され、このことから補題 12 が導かれる。

補題 10 2つの確率変数 X と Y に対して、 $X \succeq Y$ ならば $X \geq_{SSD} Y$ である。

補題 11 関数 $u(t)$ が、 t に関する増加かつ凹関数であれば、 $\int_0^1 p_s(t)u(t)dt$ もまた s に関する増加関数である。

補題 12 $x < y$ ならば、任意の $u(x) \in \mathcal{F}_{SSD}$ に対して、 $\int_0^1 p_{s(x)}(t)u(t)dt \leq \int_0^1 p_{s(y)}(t)u(t)dt$ である。

補題 11 より、 $V_{n-1}(t)$ が t の増加関数であれば、 $\int_0^1 p_s(t)V_{n-1}(t)dt$ は、 s に関する増加関数である。このことから、 n に関する帰納法を用いてつぎの性質が成り立つ。

補題 13 $V_n(s)$ は n に関する非減少関数である。

補題 14 $V_n(s)$ は、 s に関する非減少関数である。

つぎの例は、この節の仮定を満足する。

例 2

$$p_s(t) = \begin{cases} 0 & 0 \leq t < \frac{s-\alpha}{s+\alpha} \\ \frac{1}{\alpha(s)} & \frac{s-\alpha}{s+\alpha} \leq t \leq \frac{s-\alpha}{s+\alpha} \\ 0 & \frac{s+\alpha}{s+\alpha} < t \leq 1 \end{cases},$$

ここで、 $\alpha(s) = 2\alpha + (s-\alpha) \wedge 0 - (s+\alpha-1) \vee 0$, $a \vee b = \max\{a, b\}$, $a \wedge b = \min\{a, b\}$, $\underline{a} = 0 \vee a$ および $\bar{a} = a \wedge 1$ とする。この $(p_s(t))_{0 \leq s \leq 1}$ は、つぎのような計算から TP_2 となることがわかる。 $s \leq t$ and $u \leq v$ となる任意の s, t, u, v に対して ($s, t, u, v \in [0, 1]$)、 $p_s(u)p_t(v) \geq p_t(u)p_s(v)$ となることは、 $\begin{vmatrix} p_s(u) & p_s(v) \\ p_t(u) & p_t(v) \end{vmatrix} \geq 0$ と等しい。

仮定 3 推移法則 $(p_s(t))_{0 \leq s \leq 1}$ は、 t に関する増加かつ凹関数 $u(t)$ に対して、 $\int_0^1 p_s(t)u(t)dt$ は、 s に関する凹関数である

補題 15 仮定 3 のもとで、 $V_n(s)$ は s に関する凹関数となる。

命題 1 計画期間が n であり、満足度あるいは充足度が s のときの、最適な支出額を $x_n^*(s)$ とすれば、仮定 3 のもとで、任意の $x \leq y$ に対して、 $x_n^*(s) \leq x_n^*(t)$ である。

次の仮定は、性質 2 を示すために必要な仮定である。

仮定 4 推移法則が $(p_s(t))_{0 \leq s \leq 1}$ に対して、つぎの性質を満たす。 t に関する増加かつ凹関数 $u(t)$ に対して、 $\int_0^1 p_s(t)u(t)dt - u(t)$ は、 s に関する減少関数である。

推移法則 $(p_s(t))_{0 \leq t \leq 1}$ が仮定 4 を満たせば、 $\int_0^1 p_{s'}(t)u(t)dt - u(s') \leq \int_0^1 p_s(t)u(t)dt - u(s)$
 または、 $\int_0^1 p_{s'}(t)u(t)dt - \int_0^1 p_s(t)u(t)dt \leq u(s') - u(s)$ となり、このことから補題 16 と補
 題 17 が導かれ、これらの補題から性質 2 が示される。

補題 16 $s < s'$ ならば、任意の $n \geq 1$ に対して、

$$\int_0^1 p_{s'}(t)V_{n-1}(t)dt - \int_0^1 p_s(t)V_{n-1}(t)dt \geq \int_0^1 p_{s'}(t)V_n(t)dt - \int_0^1 p_s(t)V_n(t)dt$$

あるいは、 $0 \leq \int_0^1 p_{s'}(t)(V_n(t) - V_{n-1}(t))dt \leq \int_0^1 p_s(t)(V_n(t) - V_{n-1}(t))dt$ である。

補題 17 $s < s'$ ならば、任意の $n \geq 1$ に対して、

$$V_n(s') - V_n(s) \leq \int_0^1 p_{s'}(t)V_{n-1}(t)dt - \int_0^1 p_s(t)V_{n-1}(t)dt$$

である。

命題 2 計画期間が n で、満足度あるいは充足度が s のときの、最適な支出額を $x_n^*(s)$ とすれ
 ば、任意の $n \geq 1$ に対して、 $x_{n-1}^*(s) \geq x_n^*(s)$ である。

4 部分観測可能なマルコフ過程

状態空間が $[0, 1]$ のマルコフ過程で、推移法則が $P = (p_s(t))_{s, t \in [0, 1]}$ とする。すなわち、
 $p_s = (p_s(t))_{t \in [0, 1]}$ は、任意の $s \in [0, 1]$ に対して、 $[0, 1]$ 上の確率分布を表す。それぞれの状態
 s は、満足度あるいは充足度を表し ($s \in [0, 1]$)、状態に依存する確率変数 Y_s を情報プロセス
 とする。すなわち、それぞれの状態に関する情報を得ることができる情報システムあるいは観
 測過程を考える。さらに、状態は直接観測できず、状態に依存する確率変数 Y_s を通じて情報
 が得られ ($s \in [0, 1]$)、学習プロセスをベイズ学習にしたがって解析することから、仮定 5 を
 設ける。状態 s に対して、確率変数 Y_s は絶対連続で密度関数 $f_s(y)$ を持つとする ($s \in [0, 1]$)。
 この仮定は、Nakai [9] にしたがって一般化でき、多段決定問題へ応用できる (Nakai [6, 7, 8]
 など)。

仮定 5 確率変数 $\{Y_s\}_{s \in [0, 1]}$ に対して、 $s \leq s'$ ならば、 $Y_{s'} \succeq Y_s$ である ($s, s' \in [0, 1]$)。すなわ
 ち、 Y_s は s に関して尤度比の意味で増加する。

仮定 5 において、 $Y_s \succeq Y_{s'}$ のとき、 $x < y$ ならば、 $s \leq s'$ となる任意の s と s' に対して
 $(s, s' \in [0, 1])$ 、 $f_s(y)f_{s'}(x) \leq f_s(x)f_{s'}(y)$ である。このことから、確率変数 Y_s は s の値が小
 さくなるにしたがって、小さな値をとるようになり、状態 0 が一番悪い状態であり、...、状
 態 1 がもっともよい状態となる。推移法則に関する仮定から、現在の状態からより良い状態に
 推移する確率は、現在の状態がよくなるにしたがって増加する。すなわち、それぞれの状態を
 表す s が大きくなれば、より良い状態に推移する確率は大きくなる。

直接観測できない状態に関する情報は、状態空間 $[0, 1]$ 上の確率分布 μ として表し、 S を観測できない状態に関する情報の集合とすれば、

$$S = \left\{ \mu = (\mu(s))_{s \in [0,1]} \mid \int_0^1 \mu(s) ds = 1, \mu(s) \geq 0 (s \in [0,1]) \right\}$$

となる。

S に含まれる情報のあいだに、定義 2 を用いた半順序を定義する。すなわち、 $[0, 1]$ 上の 2 つの確率分布 μ, ν に対して、 $\mu(s')\nu(s) \leq \mu(s)\nu(s')$ が任意の $s, s' (s \leq s', s, s' \in [0, 1])$ について成り立ち、少なくとも 1 つの s と s' の組み合わせについて、 $\mu(s')\nu(s) < \mu(s)\nu(s')$ となるとき、 μ は ν より大きいといい、簡単に $\mu \succ \nu$ と表す。この順序は、半順序であり、この順序もまた total positive of order two、あるいは簡単に TP_2 という。いっぽう、 $p_s = (p_s(u))$ および $p_{s'} = (p_{s'}(u))$ とおけば、 P が仮定 2 を満たすことから、任意の $s, s' (s \leq s', s, s' \in [0, 1])$ に対して、 $p_{s'} \succeq p_s$ となる。この順序関係は部分観測可能なマルコフ過程において一般化でき、詳細は多段決定問題への応用を含めて Nakai [9] にある。このとき、仮定 2 と仮定 5 のもとで、補題 18 が得られる。

補題 18 $\mu \succeq \nu$ ならば、 x に関する非減少な非負関数 $h(x)$ に対して、 $\int_0^\infty h(x) dF_\mu(x) \geq \int_0^\infty h(x) dF_\nu(x)$ となる ($\mu, \nu \in S$)。

この補題において、 $F_\mu(x) = \int_0^1 \mu(s) F_s(x)$ は、weighted distribution function と呼ばれる (De Vylder [1])。また、確率過程の観測できない状態に関して、確率変数 $\{Y_s\}_{s \in [0,1]}$ 情報システム、すなわちこの確率変数を観測することで状態に関して学習を行う。事前情報が μ のとき、まずはじめにこれらの確率変数 $\{Y_s\}_{s \in [0,1]}$ を観測し、ベイズの定理を用いて学習を行う。その後、状態は推移し新しい状態になると考える。もちろん、この順序を変えても同じように解析できる。 y を観測したとき、ベイズの定理にしたがって学習した事後情報を $\mu(y) = (\mu(y, s))_{s \in [0,1]}$ とし、推移法則 P にしたがって状態が推移し、つぎの新しい状態に関する情報を $\overline{\mu(y)} = (\overline{\mu(y, s)})_{s \in [0,1]}$ と表す。

このとき、集合値関数 $h(y, s)$ に対して、定義 4 によって単調性を定義する。

定義 4 任意の $s \in [0, 1]$ と $x \in \mathfrak{R}_+$ に関する非負の集合値関数 $h(x) = (h(x, s))_{s \in [0,1]}$ に対して、任意の s' と $s (s \leq s'$ かつ $s, s' \in [0, 1])$ について、 $x < y$ ならば $h(y) \succeq h(x)$ ($h(x) \succeq h(y)$)、すなわち $h(x, s') h(y, s) \leq h(x, s) h(y, s')$ ($h(x, s') h(y, s) \geq h(x, s) h(y, s')$) となるとする。このとき、関数 $h(x, s)$ を x に関する増加関数 (減少関数) という。

確率変数 $\{Y_s\}_{s \in [0,1]}$ の密度関数 $\{f_s(y) \mid s \in [0, 1]\}$ は、仮定 5 を満たすから、 $f(y) = (f_s(y))_{s \in [0,1]}$ は $f(x) \succeq f(y)$ となる。すなわち、任意の s と $s' (s \leq s'$ および $s, s' \in [0, 1])$ に対して、 $x > y$ ならば $f_s(y) f_{s'}(x) \leq f_s(x) f_{s'}(y)$ となる。したがって、 $f(x)$ は、 x に関する増加関数である。

事前情報 μ と事後情報 $\overline{\mu(x)}$ のあいだには、仮定 2 と仮定 5 のもとで、つぎの基本的な性質が成り立つことが知られている (Nakai [9, 10] など)。

補題 19 $\mu \succ \nu$ ならば、任意の y に対して、 $\mu(y) \succ \nu(y)$ および $\overline{\mu(y)} \succ \overline{\nu(y)}$ である。任意の μ に対して、 $\mu(y)$ と $\overline{\mu(y)}$ は y に関する増加関数である。

補題 19 から、事前情報 μ における順序関係は、 $\mu(y)$ と事後情報 $\overline{\mu(y)}$ に対して保たれることがわかる。さらに、同じ事前情報 μ であれば、観測した値 y が大きくなれば、事後情報 $\overline{\mu(y)}$ もまたよくなる。

5 逐次支出モデル—不完備情報の場合

計画期間を n とし、状態に関する情報は、状態空間上の確率分布 μ として得られているとする。また、予算は最大 K とする。まずはじめに、状態の推移を含まない場合を見る。このとき最適政策にしたがったときの満足度あるいは充足度による期待利得を $V_n(\mu)$ とすれば、最適方程式は

$$V_n(\mu) = \max_{0 \leq x \leq K} \{-c(x) + V_{n-1}(\tilde{\mu}(x))\} \quad (3)$$

となる。ここで、 $V_1(\mu) = \max_{x \geq 0} \left\{ -c(x) + \int_0^1 \tilde{\mu}(t)u(t)dt \right\}$ であり、 $\tilde{\mu}(x) = \int_0^1 \mu(s)p_{s(x)}(t)ds$ は、事前情報が μ のとき、 x を支出した後の状態空間上の分布である。前節における仮定の下で、つぎの性質が得られる。

補題 20 $x > y$ ならば、 $\tilde{\mu}(x) \succeq \tilde{\mu}(y)$ であり、 $\mu \succeq \nu$ ならば、任意の $x(x \geq 0)$ に対して $\tilde{\mu}(x) \succeq \tilde{\nu}(x)$ である。

3 節と同じように、つぎの性質が仮定の下で成り立つ。また、補題 18 から、3 節の仮定の下で、最適政策のもとでの期待利得の単調性が導かれる。

命題 3 $\mu \succeq \nu$ ならば、任意の s に対して $V_n(\mu) \geq V_n(\nu)$ である。

命題 4 任意の $n \geq 1$ のとき、任意の μ に対して $V_n(\mu) \geq V_{n-1}(\mu)$ となる。

最後に、状態が部分観測可能なマルコフ過程にしたがって推移する場合を、4 節の仮定の下で考える。それぞれの満足度あるいは充足度を表す状態 s に対して ($s \in [0, 1]$)、確率変数 Y_s を観測することが、その状態に関する情報システムとする。仮定 2 のもとで、観測できない状態について、ベイズの定理にしたがった学習を考える。4 節で定義した部分観測可能なマルコフ過程に対して、最適に振る舞って得られる満足度あるいは充足度による期待利得を $V_n(\mu)$ とすれば、最適方程式は

$$\begin{aligned} V_n(\mu) &= \int_0^1 V_n(\mu|y)d\mu(y) \\ V_n(\mu|y) &= \max_{x \geq 0} \left\{ -c(x) + V_{n-1}(\overline{\mu(y)}(x)) \right\} \end{aligned} \quad (4)$$

となる。ただし、 $V_1(\mu) = \max_{x \geq 0} \left\{ -c(x) + \int_0^1 \overline{\mu(y)}(x)u(t)dt \right\}$ であり $V_0(\mu) = \int_0^1 u(t)d\mu(t)$ とする。(4) 式において、 $\overline{\mu(y)}$ は、観測した値 y を用いて学習を行った事後情報であり、 $\overline{\mu(y)}(x)$

は追加して x を支出し、そのあと推移法則にしたがって状態が推移した後での状態に関する情報である。性質 3 と性質 4 を導いたと同じように、3 節の仮定の下でつぎの性質が得られる。

命題 5 $V_n(\mu)$ は、 μ に関する非減少関数である。

命題 6 $V_n(\mu)$ は、 n に関する非減少関数である。

参考文献

- [1] F. De Vylder, Duality Theorem for Bounds in Integrals with Applications to Stop Loss Premiums, *Scandinavian Actuarial Journal*, 129–147, (1983).
- [2] S. Karlin, *Total Positivity*, Stanford University Press, Stanford, California, (1968).
- [3] S. Karlin and J. L. McGregor, Classical Diffusion Process and Total Positivity, *Journal of Mathematical Analysis and Applications*, 1, 163–183, (1960).
- [4] S. Karlin and Y. Rinott, Total Positivity Properties of Absolute Value Multinomial Variables with Applications to Confidence Interval Estimates and Related Probabilistic Inequalities, *The Annals of Statistics*, 9, 1035–1049, (1981).
- [5] M. Kijima and M. Ohnishi, Stochastic Orders and Their Applications in Financial Optimization, *Mathematical Methods of Operations Research*, 50, 351–372, (1999).
- [6] T. Nakai, A Sequential Stochastic Assignment Problem in a Partially Observable Markov process, *Mathematics of Operations Research*, 11, 230–240, (1986).
- [7] T. Nakai, An Optimal Selection Problem on a Partially Observable Markov process, In *Stochastic Modelling in Innovative Manufacturing*, Lecture Notes in Economics and Mathematical Systems 445, (Eds. A. H. Christer, S. Osaki and L. C. Thomas), pp. 140–154, Springer-Verlag, Berlin, (1996).
- [8] T. Nakai, An Optimal Assignment Problem for Multiple Objects per Period – Case of a Partially Observable Markov process, *Bulletin of Informatics and Cybernetics*, 31, 23–34, (1999).
- [9] T. Nakai, A Generalization of Multivariate Total Positivity of Order Two with an Application to Bayesian Learning Procedure, *Journal of Information & Optimization Sciences*, 23, 163–176, (2002).
- [10] T. Nakai, Properties of a Job Search Problem on a Partially Observable Markov Chain in a Dynamic Economy, *Computers & Mathematics with Applications*, to appear, (2005).