数理解析研究所講究録
第 1589 巻 2008 年 6-15

評価と関連した不完備情報の多段決定問題について

九州大学・経済学研究院　中井　達 (Tōru Nakai)
Faculty of Economics,
Kyushu University

1　不完備情報のマルコフ過程と評価

民間企業とは異なり、自治体などの公的部門での評価では、量的なものだけでなく質的な評価が重要である。そのため、公的部門における活動サイクルは、Hedley[2] のように、インプット → アウトプット → アウトカムとして捉えることが多い。この中で、アウトカムは、得られた生産物やサービスと目的あるいは目標といった基準との関係で考えられ、質的な評価が求められ、このアウトカムをもとにつぎの期の決定が行われる。したがって、多段決定問題と考えることができる。

消防活動などの公共サービスに対する支出を、毎年度の予算の範囲内で行うことを考える。これらのサービスに対して、実際の設備や施設あるいは人員と、このサービスに対して満足するかというこ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、生活環境や経済状況などが変化することで、これらのサービスに対する要求が増加し、満足を感じている住民の割合が低下することもある。そこで、生産物やサービスに対して満足を感じている住民の割合をアウトカムの１つの指標ととらえ、この指標は確率的に推移する状態によっても変化するとする。また、予算を追加して支出することで、状態を変化させることができ、その結果アウトカムの指標である住民の割合の変化を促すことができるとする。このようなアウトカムに関連する観測値をもとに決定を行う、多段決定問題を考える。また、外的要因を確率過程によって表し、状態が決定だけでなく確率過程の推移によっても変化する問題を考える。とくに、この確率過程としてマルコフ過程を仮定する。さらに、この状態を直接観測できない不完備情報の決定問題を考える。

2　アウトカムにもとづく最適支出モデル

2.1　支出の逐次決定モデル

このモデルを解析するために状態空間が $[0,\infty)$ あるいは $(-\infty,\infty)$ のマルコフ過程を考え、この状態とアウトカムの指標である対象とするサービスに対して満足を感じている住民の割合との関係を、$[0,\infty)$ あるいは $(-\infty,\infty)$ 上の確率変数の分布関数 $\Phi(x)$ を用いて表す。すなわち、マルコフ過程の状態が $s\in[0,\infty)$ のとき、対象とするサービスに対して満足を感じている住民の割合が $\Phi(s)$ である。このように、$[0,\infty)$ あるいは $(-\infty,\infty)$ を状態空間とするモデルとして解析し、$\Phi(s)=1$ であれば対象とする

サービスに住民すべてが満足していると考えられ、この s が減少するにしたがって、満足している住民の割合も減少することになる。

状態を s とするとき、この状態が確率的に推移しない場合について考える。このとき、対象とするサービスに対満足を感じている住民の割合は、この状態に応じて定まる。

いま、状態が s のとき、各期ごとの予算の範囲内で x を支出する。そのときの支出に伴う費用を $c(x)$ とし、その結果として状態は s と支出額 x の関数として $\sigma(s,x)=s(x)$ とする。ここでは、記号を簡単にするために $\sigma(s,x)$ の代わりに $s(x)$ と表す。また、費用関数が $c(x)=x$ であれば費用と支出額は等しい場合である。

はじめに、$s(x)$ に関する条件のために、2 変数関数 $g(x,s)$ に関するつぎの定義を導入する (Ross[9])。

定義 1 2 変数関数 $g(x,s)$ が、$x<y$ および $s<t$ となる x,y と s,t に対して

$$g(y,t)+g(x,s)\leq g(x,t)+g(y,s)$$

となるとき、この関数を submodular という。

このとき、$c(x)$ と $s(x)$ に対してつぎの仮定をもうける。

仮定 1 $s(x)$ は、s と x の 2 変数関数とみたとき、submodular である。すなわち、$x<y$ および $s<t$ のとき

$$\sigma(t,y)-\sigma(t,x)\leq\sigma(s,y)-\sigma(s,x) \tag{1}$$

あるいは

$$t(y)-t(x)\leq s(y)-s(x)$$

となる。また、$c(x)$ は、x に関して増加かつ凸関数とし、$s(x)$ は、x に関して (単調) 増加かつ凹関数であり、s に関する (単調) 増加関数とする。また、$c(0)=0$ であり $s(0)=s$ とする。

もし、$\sigma(s,x)=s+d(x)$ であれば、(1) 式を満足する。ここでは、不完備情報のマルコフ過程における決定問題を考えるために、$\sigma(s,x)=s+d(x)$ と仮定する。また、初期状態が s のときの利得を $u(s)$ とし、$u(s)$ は、s に関して増加な凹関数とする。

3　確率的な多段最適支出モデル

前節では、状態 s は外部の状況に影響されず、新たに支出することで、変化させるモデルを考えた。つぎに、この状態がマルコフ過程にしたがって確率的に推移する。いいかえれば、設備や人員を増やすために、予算内での追加的な支出を行うだけでなく、ある確率過程にしたがって状態が変化し、それに伴ってアウトカムの指標が下がることも認めるモデルである。

状態空間をこれまで同様に $[0,\infty)$ とし、状態の推移法則を $(p_s(t))_{0\leq s\leq 1}$ とする。以下の議論は、状態空間が $(-\infty,\infty)$ であっても、同様に考えることができる。

3.1 確率的順序関係とその性質

はじめに、ここで用いる確率的順序関係を導入する。ここで用いるものは、LRD、FSD、SSD である。これらの記号と定義は、Kijima and Ohnishi[3] にしたがう。

T 1 確率密度関数 $f_X(x)$ と $f_Y(x)$ を持つ 2 つの確率変数 X と Y に対して、$x \geq y$ となる任意の x と y に対して、$f_X(y)f_Y(x) \leq f_X(x)f_Y(y)$ であるとき、X は Y より尤度比の意味で大きいといい、$X \geq_{LRD} Y$ あるいは $X \succeq Y$ と表す。

つぎに、関数の2つの集合を $\mathcal{F}_{FSD} = \{u \mid u(x)$ は、x に関する増加関数 $\}$、$\mathcal{F}_{SSD} = \{u \mid u(x)$ は、x に関する増加かつ凹関数 $\}$ とし、この集合を使って定義 2 と 3 により確率変数のあいだに半順序を定義する。

T 2 確率密度関数 $f_X(x)$ と $f_Y(x)$ を持つ 2 つの確率変数 X と Y が、$u(x) \in \mathcal{F}_{FSD}$ となる任意の $u(x)$ に対して、$E[u(X)] \geq E[u(Y)]$ であるとき $X \geq_{FSD} Y$ とする。

T 3 確率密度関数 $f_X(x)$ と $f_Y(x)$ を持つ 2 つの確率変数 X と Y が、$u(x) \in \mathcal{F}_{SSD}$ となる任意の $u(x)$ に対して、$E[u(X)] \geq E[u(Y)]$ であるとき $X \geq_{SSD} Y$ とする。

これらの性質を用いて導入される確率変数のあいだの順序が半順序であることは、簡単に示すことができる。さらに、これらの順序関係に関して、補題 1 が成り立つ。

補題 1 2 つの確率変数 X と Y に対して、$X \geq_{LRD} Y$ ならば $X \geq_{FSD} Y$ であり、$X \geq_{FSD} Y$ ならば $X \geq_{SSD} Y$ である。

3.2 マルコフ過程の推移法則

つぎにマルコフ過程の推移法則 $(p_s(t))_{0 \leq s \leq 1}$ を考える。いま、2 つの確率変数 S_s, S_t をそれぞれ状態が s および t のとき、推移法則に従って推移したあとの状態を表す確率変数とする。また、2 つの確率変数 $S_{s(x)}, S_{s(y)}$ は、それぞれ状態が s と t のとき x を追加して支出したときの推移後の状態を表す確率変数であり、仮定 1 より $x < y$ ならば、$s(x) < s(y)$ となっている。このとき、このマルコフ過程の性質を確率的な順序関係で定義する。このとき、つぎの仮定をおく。

仮定 2 推移法則 $(p_s(t))_{0 \leq s \leq 1}$ に対して、$s < t$ ならば、$S_t \geq_{LRD} S_s$ とする。

まず、性質 2 のもとで、Kijima and Ohnishi[3] から、つぎの性質が成り立つ。

補題 2 $s < s'$ ならば $S_{s'} \geq_{SSD} S_s$ とする。このとき、s に関して増加かつ凹関数 $u(s)$ に対して、$\int_0^\infty p_s(t)u(t)dt \leq \int_0^\infty p_{s'}(t)u(t)dt$ である。

$s < s'$ のとき $S_{s'} \geq_{LRD} S_s$ ならば、$S_{s(y)} \geq_{SSD} S_{s(x)}$ だから、補題 2 から補題 3 が導かれる。

補題 3 $s < s'$ ならば $S_{s(y)} \geq_{LRD} S_{s(x)}$ とする。このとき、

$$\int_0^\infty p_{s(x)}(t)u(t)dt \leq \int_0^\infty p_{s(y)}(t)u(t)dt$$

である。

さらに、仮定1より $x < y$ ならば、$s(x) < s(y)$ だから、仮定2のもとで $S_{s(y)} \geq_{LRD}$ $S_{s(x)}$ である。また、補題1から、s に関して増加かつ凹関数 $u(s)$ に対して補題2が成り立ち、s に関する増加関数 $u(s)$ に対して補題3が成り立つ。ところで、$s < s'$ ならば $S_{s'} \geq_{LRD} S_s$ あるいは $S_{s'} \succeq S_s$ であることを推移法則に当てはめれば、つぎのようになる。ここで、確率変数は全順序 $\geq$ 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 推移法則 $\boldsymbol{P} = (p_s(t))_{s,t\in[0,\infty)}$ は、$s \leq t$ および $u \leq v$ となる任意の s,t,u と v に対して $(s,t,u,v \in [0,\infty))$、$\begin{vmatrix} p_s(u) & p_s(v) \\ p_t(u) & p_t(v) \end{vmatrix} \geq 0$ となる。

集合値関数 $\boldsymbol{P} = (p_s(t))_{s,t\in[0,\infty)}$ が、このような性質を持つとき、この $\boldsymbol{P}$ は TP_2(total positive of order two) の性質を持つという。この性質は、ベイズ学習を伴う多段決定問題を考える上で重要な役割を果たしている (Nakai[7] など)。

さらに、$x < y$ ならば、$S_{s(y)} \geq_{LRD} S_{s(x)}$ となることは、つぎのように表せる。

補題 4 推移法則 $\boldsymbol{P} = (p_s(t))_{s,t\in[0,\infty)}$ と関数 $s(x)$ を考える。このとき、任意の s,t,u と v に対して $(u,v \in [0,\infty))$、$x \leq y$ かつ $u \leq v$ であれば、任意の $s\,(s \in [0,\infty))$ 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。

3.3 逐次決定モデル

計画期間が n で、各期ごとの予算額の上限が K とする。このとき、最適に振る舞ったときの状態に対する期待利得を $V_n(s)$ とすれば、状態がマルコフ過程にしたがって推移するから、最適方程式はつぎのようになる。

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

ただし、$V_1(s) = \max_{0\leq x\leq K}\left\{-c(x) + \int_0^\infty p_{s(x)}(t)u(t)dt\right\}$ である。

補題 5 $V_n(s)$ は、s に関する非減少関数である。すなわち、$s < s'$ ならば、$V_n(s) \geq V_n(s')$ である。

性質 1 計画期間が n であり、状態が s のときの、最適な支出額を $x_n^*(s)$ とする。このとき、$s \leq s'$ ならば、$x_n^*(s) \leq x_n^*(s')$ である。

性質 2 計画期間が n で、状態が 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。

ところで、最適政策にしたがったときの最適値 $V_n(s)$ の n に関する単調性について考える。基本的に、公的サービスに対する支出は、将来の満足度や充足度による期待効用が現時点に比べて悪くなったとしても、これらのサービスを打ち切ることはできず、続けて行う必要がある。したがって、満足度や充足度を表す状態の関数として表される効用と、推移法則によっては、$V_n(s)$ は n に関して増加することもあれば、減少することも考えられる。このことは、帰納法を用いれば、$n=1$ のときの性質により $V_n(s)$ の n に関する単調性が定まる。ところで、$u(s)$ が s に関する凸関数で、$E[S_{s(0)}] \geq s$ であれば、イェンセン (Jensen) の不等式より、$V_1(s) \geq V_0(s)$ となる。したがって、$V_n(s)$ は n に関する非減少関数となる。この場合は、追加の支出をしなくとも、期待効用は現在の充足度や満足度による効用より大きくなる場合となっている。このことは、公的なサービスは状態が良くなる傾向にあっても、あるいは悪くなる傾向を持つにしても、いずれの場合にもサービスは続けて行かなくてはならず、これが通常の最適停止問題などと異なっている点である。

4 部分観測可能なマルコフ過程と学習プロセス

4.1 部分観測可能なマルコフ過程と情報

状態空間を $[0,\infty)$ とするマルコフ過程で、推移確率を $(p_s(t))_{s,t\in[0,\infty)}$ とすれば、$\boldsymbol{p}_s = (p_s(t))_{t\in[0,\infty)}$ は状態空間 $[0,\infty)$ の任意の状態 $s \in [0,\infty)$ に対して、状態空間上の確率分布となっている。以下では状態を直接観測できない部分観測可能なマルコフ過程における多段決定問題を考える。

状態に関する情報は、状態空間 $[0,\infty)$ 上の確率分布 $\boldsymbol{\mu}$ として表し、$\mathcal{S}$ を状態に関する情報全体の集合とすれば、

$$\mathcal{S} = \left\{ \boldsymbol{\mu} = (\mu(s))_{s\in[0,\infty)} \,\middle|\, \int_0^1 \mu(s)ds = 1, \mu(s) \geq 0\,(s \in [0,\infty)) \right\}$$

となる。

$\mathcal{S}$ に含まれる情報のあいだに、定義 1 を用いた半順序を定義する。すなわち、$[0,\infty)$ 上の 2 つの確率分布 $\boldsymbol{\mu}, \boldsymbol{\nu}$ に対して、$\mu(s')\nu(s) \leq \mu(s)\nu(s')$ が任意の s, s' ($s \leq s'$, $s, s' \in [0,\infty)$) 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、$\boldsymbol{\mu}$ は $\boldsymbol{\nu}$ より大きいといい、簡単に $\boldsymbol{\mu} \succ \boldsymbol{\nu}$ と表す。いっぽう、$\boldsymbol{p}_s = (p_s(u))$ および $\boldsymbol{p}_{s'} = (p_{s'}(u))$ とおけば、$\boldsymbol{P}$ が仮定 2 を満たすことから、任意の s, s' ($s \leq s'$, $s, s' \in [0,\infty)$) に対して、$\boldsymbol{p}_{s'} \succeq \boldsymbol{p}_s$ 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る (Nakai [5])。

補題 6 $\boldsymbol{\mu} \succeq \boldsymbol{\nu}$ ならば ($\boldsymbol{\mu}, \boldsymbol{\nu} \in \mathcal{S}$)、$x$ に関する非減少な非負関数 $h(x)$ に対して、$\displaystyle\int_0^\infty h(x)dF_{\boldsymbol{\mu}}(x) \geq \int_0^\infty h(x)dF_{\boldsymbol{\nu}}(x)$ となる。$F_{\boldsymbol{\mu}}(x) = \displaystyle\int_0^1 \mu(s)F_s(x)$ とする。

事前情報 $\boldsymbol{\mu}$ に対して、

$$\overline{\mu}(s) = \int_{-\infty}^{\infty} \mu(t)p_t(s)ds \qquad (3)$$

を、マルコフ過程の推移法則にしたがって推移したあとの状態に関する事後情報とする。この $\overline{\mu} = (\overline{\mu}(t))_{t\in(-\infty,\infty)}$ に関して、つぎの性質が成り立つ (Nakai[5] など)。

補題 7 $\mu \succ \nu$ ならば $\overline{\mu} \succ \overline{\nu}$ である。

4.2 学習プロセス

状態 s 対して、この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確率変数 Y_s を通して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。また、学習プロセスはベイズ学習にしたがって解析することから、仮定 3 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 ($s \in [0,\infty)$)。この仮定は、Nakai [5] にしたがって一般化でき、多段決定問題へ応用できる (Nakai [4] など)。また、学習をベイズの定理にしたがって行うことから、推移法則 $(p_{s(x)})(t))_{0\le s\le 1}$ が TP_2 の性質を持つと仮定して議論する。

仮定 3 確率変数 $\{Y_s\}_{s\in[0,\infty)}$ に対して、$s \le s'$ ならば、$Y_{s'} \succeq Y_s$ である ($s, s' \in [0,\infty)$)。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 3 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のである。

確率過程の状態に関して、確率変数 $\{Y_s\}_{s\in[0,\infty)}$ を観測することによって、状態に関して学習を行う。事前情報が μ のとき、はじめにこれらの確率変数 $\{Y_s\}_{s\in[0,\infty)}$ 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 $\mu(y) = (\mu(y,s))_{s\in[0,\infty)}$ とすれば、

$$\mu(y)(s) = \frac{\mu(s) f_s(y)}{\int_{-\infty}^{\infty} \mu(s) f_s(y) ds}. \tag{4}$$

である。その後で推移法則 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 $\overline{\mu(y)} = (\overline{\mu(y,s)})$ とする。ここで、

$$\overline{\mu(y)}(s) = \int_{-\infty}^{\infty} \mu(y)(t) p_t(s) dt. \tag{5}$$

である。

事前情報 μ と事後情報 $\overline{\mu(x)}$ のあいだには、つぎの基本的な性質が成り立つ (Nakai [5] など)。

補題 8 $\mu \succ \nu$ 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。

この性質は、Nakai[5] のように一般化でき、不完備情報のマルコフ過程における決定問題への応用は、Nakai[4] などにある。

4.3 Gradually Condition

不完備情報の多段決定問題を考えるために、いくつかの準備をする。ここで考えたモデルでは、決定がつぎの期の状態に影響することからも、これらの吟味が必要である。状態に関する事前情報が $\boldsymbol{\mu}$ のとき、支出を x としたという条件の下での状態空間上の確率分布を $\boldsymbol{\mu}_x = (\mu_x(s))$ とすれば、仮定から $\mu_x(s) = \mu(s - d(x))$ となっている。つぎに、事前情報が $\boldsymbol{\mu}$ のとき、支出を x としたという条件の下で状態が推移し、つぎの期における状態空間上の確率分布を $\overline{\mu_x} = (\overline{\mu_x}(s))$ とすれば、つぎのようになる。

$$\boldsymbol{\mu}_x(t) = \int_0^\infty \mu_x(s)p_s(t)ds = \int_0^\infty \mu(s)p_{s(x)}(t)ds. \tag{6}$$

ここで、$s(0) = s$ だから、$\overline{\mu} = \int_0^\infty \mu(s)p_s(t)ds = \boldsymbol{\mu}_0$ である。

定義 3 $\mathcal{S}$ に含まれる状態空間上の確率分布 $\boldsymbol{\mu}$ が $s < t, s' < t'$ と $s - s' = t - t' = c < 0$ を満たす任意の $s < s', t < t'$ に対して、$\dfrac{\mu(s)}{\mu(s')} \geq \dfrac{\mu(t)}{\mu(t')}$ となるとき、この $\boldsymbol{\mu}$ は *gradually condition* を満足するという。

また、$\boldsymbol{\mu}$ が gradually condition を満足するとき、$\boldsymbol{\mu}_x$ もまた、gradually condition を満足する。状態空間上の正規分布 $\mu(s) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\dfrac{(s-a)^2}{2\sigma^2}}$ はこの性質を満たす。

$\mathcal{S}$ に含まれる状態に関する情報 $\boldsymbol{\mu}$ に対して、$\overline{\mu}(t)$ を (3) 式で定義された推移後の状態に関する事後情報とする。推移確率に関してつぎの仮定 4 をおく。

仮定 4 任意の $s < s'$, $t \leq t'$ および $u < v$ となる s, s', t, t', u, v に対して

$$p_u(s)p_v(t') - p_u(t)p_v(s') \geq p_v(s)p_u(t') - p_v(t)p_u(s')$$

とする。すなわち、$\begin{vmatrix} p_u(s) & p_u(t) \\ p_v(s') & p_v(t') \end{vmatrix} \geq \begin{vmatrix} p_v(s) & p_v(t) \\ p_u(s') & p_u(t') \end{vmatrix}$ である。

補題 9 状態全体の集合 $\mathcal{S}$ に含まれる確率分布 $\boldsymbol{\mu}$ が *gradually condition* を満足するならば、$\overline{\mu}$ もまた *gradually condition* を満足する。

補題 10 $\boldsymbol{\mu}$ が *gradually condition* を満足するならば、$\overline{\mu_x}$ もまた *gradually condition* を満足する。

推移法則 $p_v(s) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(s-v)^2}{2\sigma^2}}$ は、仮定 4 の条件を満足する。事後情報 $\overline{\boldsymbol{\mu}(y)}$ が、gradually condition を満足するかどうかを調べるために仮定 5 をおく。

仮定 5 確率変数 Y_s の密度関数 $f_s(y)$ は、$t - s = t' - s' > 0$ となる $s < s'$ と $t < t'$ に対して、$\dfrac{f_s(y)}{f_{s'}(y)} \geq \dfrac{f_t(y)}{f_{t'}(y)}$ となる $(s \in (-\infty, \infty))$。

補題 11 状態全体の集合 $\mathcal{S}$ に含まれる確率分布 $\boldsymbol{\mu}$ が *gradually condition* を満足するならば、任意の y に対して $\boldsymbol{\mu}(y)$ もまた *gradually condition* を満足する。

補題 12 μ が *gradually condition* を満足するならば、任意の y に対して、$\overline{\mu(y)}$ もまた *gradually condition* を満足する。

$f_s(y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-s)^2}{2\sigma^2}}$ とすれば、仮定 5 を満たす。

4.4 単調性

S に含まれる状態空間上の確率分布 μ に関して、事前情報を μ としたときの事後分布を表す記号をまとめておくことにしよう。

μ: 事前分布

$\overline{\mu}$: (3) 式で定義される、状態が推移したあとでの確率分布

μ_x: x を支出すると決定したあとでの状態空間上の確率分布

$\mu(y)$: 情報プロセスから情報として y が得られたとき、(4) 式で定義されるベイズの定理にしたがって学習を行ったあとでの事後情報

$\overline{\mu(y)}$: 事前情報を $\mu(y)$ としたとき、(5) 式で定義される推移法則 $\boldsymbol{P}$ にしたがって状態が推移したあとの確率分布

$\overline{\mu_x}$: 事前情報が μ のとき、x を支出すると決定したあとで、(7) 式で定義される推移法則 $\boldsymbol{P}$ にしたがって状態が推移したあとの確率分布

$\overline{\mu(y)_x}$: 事前情報が $\mu(y)$ のとき、x を支出すると決定したあとで、(8) 式で定義される推移法則 $\boldsymbol{P}$ にしたがって状態が推移したあとの確率分布

状態に関する事前情報が μ のとき、

$$\overline{\mu_x}(s) = \int_{-\infty}^{\infty} \mu(t) p_{t(x)}(s) dt \tag{7}$$

は、x を支出すると決定したあとでの状態空間上の確率分布である。

ここでは、学習と決定、推移の順序をつぎのように考える。すなわち、事前情報が μ のとき、はじめに情報プロセスを観測し、この情報をもとにベイズの定理を用いて $\mu(y)$ と学習を行う。その後、支出額 x を決定し、推移法則 $\boldsymbol{P}$ にしたがって状態が推移し、新しい状態になると考える。その結果、推移後の新しい状態に関する情報は、

$$\overline{\mu(y)_x}(s) = \int_{-\infty}^{\infty} \mu(y)(t) p_{t(x)}(s) dt \tag{8}$$

とすれば、$\overline{\mu(y)_x} = (\overline{\mu(y)_x}(s))$ となる。

補題 13 状態全体の集合 S に含まれる確率分布 μ が *gradually condition* を満足するとき、$x > x'$ ならば、$\mu_x \succeq \mu_{x'}$ である。ただし、$\mu_x = (\mu_{s(x)})$ とする。

補題 14 状態全体の集合 S に含まれる確率分布 μ と ν が *gradually condition* を満足するとき、$\mu \succeq \nu$ ならば、任意の $x(\geq 0)$ に対して、$\mu_x \succeq \nu_x$ である。

補題 15 状態全体の集合 S に含まれる確率分布 μ と ν が *gradually condition* を満足するとき、$\mu \succeq \nu$ ならば、任意の $x(\geq 0)$ に対して $\overline{\mu_x} \succeq \overline{\nu_x}$ かつ $\overline{\mu(y)_x} \succeq \overline{\nu(y)_x}$ である。

補題 16 状態全体の集合 S に含まれる確率分布 μ が *gradually condition* を満足するとき、$y > y'$ ならば $\overline{\mu(y)_x} \succeq \overline{\mu(y')_x}$ である。

補題 17 μ が *gradually condition* を満足するとき、$x > x'$ ならば、$\overline{\mu(y)_x} \succeq \overline{\mu(y)_{x'}}$ である。

4.5 不完備情報の確率的な多段最適支出モデル

最後に、状態がマルコフ過程にしたがって推移し、その状態を直接知ることができない場合の逐次支出モデルを考えることにしよう。状態に関する情報は、情報プロセスを通して得られる。したがって、このモデルは、4 節の部分観測可能なマルコフ過程での逐次決定問題として定式化できる。

このような部分観測可能なマルコフ過程での逐次決定問題において、状態に関する情報は、状態空間上の確率分布として表され、情報プロセスから得られた観測値をもとにベイズの定理にしたがって学習を行う。また、4 節の部分観測可能なマルコフ過程において、それぞれの状態 s $(s \in [0, \infty))$ に対して、確率変数 Y_s を観測過程とし、この値を観測することが情報プロセスである。仮定 2 のもとで、これらの確率変数 Y を観測することで情報を獲得し、その情報をもとにベイズの定理に基づいた学習プロセスによって、情報を改良する。状態に関する情報が μ で、計画期間が n のとき、最適政策にしたがって得られる総期待利得を $\widetilde{V}_n(\mu)$ とすれば、最適性の原理より、つぎのような再帰方程式が得られる。

$$
\begin{aligned}
\widetilde{V}_n(\mu) &= \int_{-\infty}^{\infty} \widetilde{V}_n(\mu|y)d\mu(y) \\
\widetilde{V}_n(\mu|y) &= \max_{0\leq x\leq K}\left\{-c(x) + \widetilde{V}_{n-1}(\overline{\mu(y)_x})\right\} \qquad (9)
\end{aligned}
$$

ここで、$\widetilde{V}_0(\mu) = \displaystyle\int_{-\infty}^{\infty} u(t)d\mu(t)$ とする。(9) 式において、$\mu(y)$ を情報プロセスから得られた値 y をもとに、情報を改良したあとの状態に関する情報とする。 すなわち、事前情報が μ のとき、まず始めに情報プロセスから観測値 y を観測し、状態に関する情報をベイズの定理にしたがって $\mu(y)$ と改良するのである。そのあと、決定を x としたあとで、状態が s であれば、推移法則 $(p_{s(x)}(t))_{0\leq s\leq 1}$ にしたがって状態が推移する。こうして、この確率過程は新しい状態となり、この新しい状態に関する情報は (8) 式のように、$\overline{\mu(y)_x}$ となる。これは、学習したあと 1 期間経過後の状態空間上の確率分布である。そのあとで、最適政策にしたがって得られる残り計画期間での総期待利得は $\widetilde{V}_{n-1}(\overline{\mu(y)_x})$ となる。したがって、n に関する帰納法を用いれば、2 節の仮定の下でつぎの性質が得られる。.

性質 3 状態全体の集合 $\mathcal{S}$ に含まれる確率分布 μ と ν が *gradually condition* を満足するとき、$\mu \succeq \nu$ ならば、$\widetilde{V}_n(\mu) \geq \widetilde{V}_n(\nu)$ である。

$\mu \succ \nu$ であれば、$u(t)$ が t の非減少な非負関数なので、補題6より $\widetilde{V}_0(\mu) \geq \widetilde{V}_0(\nu)$ である。また、任意の情報 y に対して、補題8から、$\mu(y) \succ \nu(y)$ である。さらに、補題14から、任意の決定 x に対して、$\overline{\mu(y)_x} \succeq \overline{\nu(y)_x}$ となる。これらの事後情報に関する単調性から、任意の決定 x と観測値 y に対して、$\mu \succ \nu$ ならば、$\overline{\mu(y)_x}(t) \succeq \overline{\nu(y)_x}(t)$ であり、n に関する帰納法によって性質 3 を示すことができる。

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] Hedley, T. P. (1998), "Measuring Public Sector Effectiveness Using Private Sector Methods", *Public Productivity & Management Review*, **21** (3), 251–258.

[3] M. Kijima and M. Ohnishi, Stochastic Orders and Their Applications in Financial Optimization, *Mathematical Methods of Operations Research*, **50**, 351–372, (1999).

[4] 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).

[5] 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).

[6] T. Nakai, Economy, Efficiency and Effectiveness, In *Policy Analysis in the Era of Globalization and Localization* (Eds. Research Project Group for Policy Evaluation in Kyushu University), Kyushu University Press, 165–193, 2006.

[7] T. Nakai, Properties of a Job Search Problem on a Partially Observable Markov Chain in a Dynamic Economy, *Computers & Mathematics with Applications*, vol. 51, 189–198, 2006.

[8] T. Nakai, A Sequential Expenditure Problem for Public Sector Based on the Outcome, *Recent Advances in Stochastic Operations Research* (Eds. T. Dohi, S. Osaki and K. Sawaki), World Scientific Publishing, 277–295, 2007.

[9] S. M. Ross, *Stochastic Processes*, John-Wiley and Sons, New York, New York, 1983.