数理解析研究所講究録
第1559巻 2007年 67-81

評価に関連するマルコフ過程での最適化問題について

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

1　アウトカムにもとずく評価と決定

民間企業では数値的に表される指標として、収益あるいは利潤によって経営が良好であるかどうかを判断することは可能である。しかし、自治体や公団などの公的部門で、民間企業などと同じような規準で評価すれば、これらの部門における結果としてのサービスや利益を十分に評価することはできない。そのため、自治体などの公的部門におけるマネジメント・サイクルあるいは活動サイクルを、Hedley[2] にあるように

$$\text{インプット} \to \text{アウトプット} \to \text{アウトカム}$$

として捉える、このサイクルをもとに評価が考えられている。公的部門におけるマネジメント・サイクルにおいてはインプットとして費やした資源をもとに活動し、アウトプットとして生産物やサービスをうむ。その結果として、これらの生産物が、アウトカムとして考える基準あるいは期待したものとなっているかを判断するのである。したがって、インプットとアウトプットとの関係は、費やされた資源とその結果として得られた生産物の比として考えることができる。それに対して、アウトプットとアウトカムは、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、生産物が目的あるいは目標に到達しているかどうかで判断あるいは評価を行うのである。

このように、インプットとアウトプットの関係は比較的簡単に説明できるものが多いが、アウトプットとアウトカムの関係については、目的あるいは目標に達しているかどうかをどのように評価するかが問題となっている。

このようなインプット・アウトプット・アウトカムによるマネジメントサイクルの評価においては、経済性 (economy)、効率性 (efficiency)、有効性 (effectiveness) の基準で評価するシステムが基本的な考え方である。ここで、効率性はインプットとアウトプットのあいだの関係を評価するものであり、有効性はアウトプットとアウトカムのあいだを評価するものである。さらに、効率性には期待されるアウトプットを実現するという目的で、インプットを抑えるという面と、与えられたインプットのなか

で、アウトプットを大きくするという面を持っている。しかし、有効性はこのような評価においては重要であるが、アウトカムに関わることもあり、その方法について確立されているとはいえない。

ところで、有効性はアウトプットとアウトカムの関係を見ようとするものではあるが、必ずしもアウトプットとアウトカムの関係のみで評価できるものではない。インプットとアウトプットのあいだには関連はあるものの、直接的な関係を限定することができないので、マネジメントサイクルにおいてアウトプットをインプットとアウトプットのあいだに入れることで、その関連を表すした。しかし、依然としてインプットとアウトカムの関係を規定することは難しい。

多くの場合には、インプットとアウトプットについては、比較的明らかな数値で表されることが多い。しかし、アウトカムを観測し、それを評価するために数値化することは困難な場合が多い。その理由としては、つぎのようなものが考えられる。(1) アウトカムを数値的に表して、目的や目標の評価基準を計測することは困難である。(2) アウトカムとインプットの、直接的な関連性が明確ではない。(3) 施策や政策を実行するための費用から、結果としてえられる特定のサービスのために用いられた費用を区別することが難しい。(4) アウトカムを評価するにあたって、政策やプログラムがその活動や結果に関連していることはわかっても、その政策やプログラムのその結果に対する寄与の度合いが明確ではない。

このようにアウトカムは評価において重要な要素であるにもかかわらず、アウトカムは数値化することが難しい。ここでは、アウトカムを考慮した決定問題を考えるために、生産物あるいはサービスに対して満足していると思う住民の割合をその基準ととらえ、多段決定問題としてモデル化することを試みる。

2　アウトカムにもとづいた支出の逐次決定モデル

2.1　アウトカムと確率過程

消防活動や警察活動といった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、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、あるいは充足していると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ものとする。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。

2.2 支出の逐次決定モデル

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[14])。

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。このときには、支出が同じであっても、プロセスの状態が異なれば、対称となるサービスに対して満足している住民の割合の改善度は異なる。このことからも、$\sigma(s,x)=s+d(x)$ と仮定しても問題ない。

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で、状態 s を変化させて、アウトカムの指標である対象とするサービスに満足を感じている住民の割合を上げることができる。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

$$v_n(s)=\max_{0\leq x\leq K}\{-c(x)+v_{n-1}(s(x))\} \tag{2}$$

となる。ただし、$v_1(s)=\max_{0\leq x\leq K}\{-c(x)+u(s(x))\}$ である。初期条件は $v_0(s)=u(s)$ であり、$u(s)$ は、s に関して増加な凹 (concave) 関数とする。

3 公共部門に対する支出の逐次決定モデル: 確率モデル

前節では、アウトカムの１つの指標と考えた満足を感じている住民の割合を状態空間上の確率分布で表し、その状態 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。しかし、一般的には予算からの支出とは異なる、社会状況や経済状態などの外部から要因の影響を受けてこの状態が変化し、それに伴って対象とするサービスに対して満足を感じる住民の割合は変化する。したがって、この状態が確率的に推移するものとし、ここではマルコフ過程にしたがうとする。いいかえれば、設備や機器、あるいは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。

3.0.1 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、確率変数のあいだに導入する。ここで用いるものは、LRD(likelihood ratio order)、FSD(first order stochastic dominance)、SSD(second order stochastic dominance) である。これらの記号と定義は、Kijima and Ohnishi[3] にしたがうことにする。

定義 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 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。

この定義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。

つぎに、関数の 2 つの集合

$$\begin{aligned} \mathcal{F}_{FSD} &= \{u \mid u(x) \text{ は、} x \text{ に関する増加関数 }\} \\ \mathcal{F}_{SSD} &= \{u \mid u(x) \text{ は、} x \text{ に関する増加かつ凹関数 }\} \end{aligned}$$

とし、この集合を使って定義 3 と 4 により確率変数のあいだに半順序を定義する。

定義 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、$u(x) \in \mathcal{F}_{FSD}$ 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。

定義 4 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 3 つの定義 (定義 2、定義 4 と定義 3) による順序関係に関して、補題 1 が成り立つ。すなわち、定義 2 による順序関係は、定義 4 による順序関係より強く、定義 4 による順序関係は、定義 3 による順序関係より強いことがわかる。

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。

3.0.2 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 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。

はじめに、$s < s'$ ならば $S_{s'} \geq_{SSD} S_s$ であることを仮定すれば、定義 3 から

補題 2 $s < s'$ ならば、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。

すなわち、関数 $u(t)$ が、t に関する増加かつ凹関数であれば、$\int_0^\infty p_s(t)u(t)dt$ もまた s に関する増加関数である。この性質と、$s < s'$ ならば $S_{s'} \geq_{SSD} S_s$ 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 $x < y$ ならば、s に関する増加関数 $u(s)$ に対して、

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

である。

つぎに、$s < s'$ ならば $S_{s'} \geq_{FSD} S_s$ であることを仮定すれば、定義 3 から

補題 4 $s < s'$ ならば、s に関する増加関数 $u(s)$ 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。

この性質と、$s < s'$ ならば $S_t \geq_{FSD} S_{s'}$ ならば、$S_{s(y)} \geq_{FSD} S_{s(x)}$ だから、補題 4 から補題 5 が導かれる。

補題 5 $x < y$ ならば、s に関する増加関数 $u(s)$ に対して、

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

である。

さらに、$s < s'$ ならば $S_{s'} \geq_{LRD} S_s$ あるいは $S_{s'} \succeq S_s$ であることを仮定すれば、仮定 1 より $x < y$ ならば、$s(x) < s(y)$ だから、$S_{s(y)} \geq_{LRD} S_{s(x)}$ 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 5 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ものとする。

定義 5 推移法則 $\boldsymbol{P} = (p_s(t))_{s,t\in[0,\infty)}$ は、$s \le t$ および $u \le v$ となる任意の s, t, u と v に対して $(s, t, u, v \in [0, \infty))$、$\begin{vmatrix} p_s(u) & p_s(v) \\ p_t(u) & p_t(v) \end{vmatrix} \ge 0$ 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 TP_2(total positivity of order two) 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[12] など) が知られているように、重要な役割を果たしている。

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

補題 6 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 $(u, v \in [0, \infty))$、$x \le y$ かつ $u \le v$ であれば、任意の $s\ (s \in [0, \infty))$ について、$\begin{vmatrix} p_{s(x)}(u) & p_{s(x)}(v) \\ p_{s(y)}(u) & p_{s(y)}(v) \end{vmatrix} \ge 0$ となる。

このとき、このマルコフ過程の推移法則に対して、つぎの仮定をおく。

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

3.1 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。

$$V_n(s) = \max_{0\le x\le K} \left\{ -c(x) + \int_0^\infty p_{s(x)}(t) V_{n-1}(t) dt \right\} \tag{3}$$

ただし、

$$V_1(s) = \max_{0\le x\le K} \left\{ -c(x) + \int_0^\infty p_{s(x)}(t) u(t) dt \right\}$$

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関数で前節の条件を満たすものである。

補題 7 $V_n(s)$ は、s に関する非減少関数である。すなわち、$s < s'$ ならば、$V_n(s) \ge V_n(s')$ 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、$s \le s'$ ならば、$x_n^*(s) \le x_n^*(s')$ 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。ところで、任意の s に対して $V_{n-1}(s) \leq V_{n-2}(s)$ ならば、$\int_0^\infty p_{s(x)}(t)V_{n-1}(t)dt \leq \int_0^\infty p_{s(x)}(t)V_{n-2}(t)dt$ となるので、

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

より、$V_n(s) \leq V_{n-1}(s)$ となることがわかる。反対に、任意の s に対して $V_{n-1}(s) \geq V_{n-2}(s)$ ならば、$V_n(s) \geq V_{n-1}(s)$ となる。したがって、帰納法を用いれば、$n = 1$ のときの性質によって、$V_n(s)$ の n に関する単調性が定まる。すなわち、$n = 1$ のときは、$V_1(s) = \max_{0 \leq x \leq K} \left\{ -c(x) + \int_0^\infty p_{s(x)}(t)u(t)dt \right\}$ であり、$V_0(s) = u(s)$ だから、$V_1(s) \geq V_0(s)$ であれば $V_n(s)$ は n に関する非減少関数であり、$V_1(s) \leq V_0(s)$ であれば $V_n(s)$ は n に関する非増加関数となることがわかる。

ところで、$u(s)$ が s に関する凸関数のときにはどうなるだろうか。いま、状態が s のとき、追加して x の支出による推移後の状態を表す確率変数 $S_{s(x)}$ に対して、$E[S_{s(0)}] \geq s$ であれば、イェンセン (Jensen) の不等式より $\int_0^\infty p_s(t)u(t)dt \geq u(s)$ となるので、

$$V_1(s) \geq -c(0) + \int_0^\infty p_{s(0)}(t)u(t)dt = \int_0^\infty p_s(t)u(t)dt \geq u(s) = V_0(s)$$

より、$V_1(s) \geq V_0(s)$ となることがわかる。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。こらまでとは異なり、以下では状態を直接観測できないとする。すなわち、部分観測可能なマルコフ連鎖における多段決定問題を考えることにする。

直接観測できない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、$\mathcal{S}$ を観測できない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 2 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。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、$\boldsymbol{p}_s = (p_s(u))$ 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、詳細は多段決定問題への応用を含めて Nakai [7] にある。このとき、仮定 2 と仮定 3 のもとで、補題 8 が得られる。

補題 8 $\boldsymbol{\mu} \succeq \boldsymbol{\nu}$ 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。

この補題において、$F_{\boldsymbol{\mu}}(x) = \displaystyle\int_0^1 \mu(s)F_s(x)$ は、weighted distribution function と呼ばれる (De Vylder [1])。

4.2 学習プロセス

それぞれの状態 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。さらに、マルコフ過程の状態を直接には観測できず、状態に依存する確率変数 Y_s を通じて情報が得られ $(s\in[0,\infty))$、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

[7] 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, 5, 6] など)。また、ここでは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を進める。

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 において、$Y_s \succeq Y_{s'}$ のとき、$x < y$ ならば、$s \le s'$ となる任意の s と s' に対して $(s, s' \in [0,\infty))$、$f_s(y)f_{s'}(x) \le f_s(x)f_{s'}(y)$ 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のである。

確率過程の観測できない状態に関して、確率変数 $\{Y_s\}_{s\in[0,\infty)}$ を情報システムとする。すなわち、この確率変数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 $\boldsymbol{\mu}$ のとき、まず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 $\boldsymbol{\mu}(y) = (\mu(y,s))_{s\in[0,\infty)}$ とし、その後で推移法則 $\boldsymbol{P}$ 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 $\overline{\boldsymbol{\mu}(y)} = (\overline{\mu(y,s)})$ となる。ここで、

$$\overline{\mu(y,s)} = \frac{\mu(s)f_s(y)}{\displaystyle\int_0^\infty \mu(s)f_s(y)ds}$$

である。

このとき、集合値関数 $\boldsymbol{h}(y,s)$ に対して、定義 6 によって単調性を定義する。

定義 6 任意の $s \in [0,\infty)$ と $x \in \Re_+$ に関する非負の集合値関数 $\boldsymbol{h}(x) = (h(x,s))_{s\in[0,\infty)}$ に対して、任意の s' と s $(s \le s'$ かつ $s, s' \in [0,\infty))$ について、$x < y$ ならば $\boldsymbol{h}(y) \succeq \boldsymbol{h}(x)$ $(\boldsymbol{h}(x) \succeq \boldsymbol{h}(y))$ とする。すなわち $h(x,s')\,h(y,s) \le h(x,s)\,h(y,s')$ $(h(x,s')\,h(y,s) \ge h(x,s)\,h(y,s'))$ である。このとき、関数 $\boldsymbol{h}(x,s)$ を x に関する増加関数 (減少関数) という。

確率変数 $\{Y_s\}_{s\in[0,\infty)}$ の密度関数 $\{f_s(y) \mid s \in [0,\infty)\}$ は、仮定 3 を満たすから、$\boldsymbol{f}(y) = (f_s(y))_{s\in[0,\infty)}$ は $\boldsymbol{f}(x) \succeq \boldsymbol{f}(y)$ となる。すなわち、任意の s と s' $(s \le s'$ および

$s, s' \in [0, \infty))$ 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、$\boldsymbol{f}(x)$ は、x に関する増加関数である。

事前情報 $\boldsymbol{\mu}$ と事後情報 $\overline{\boldsymbol{\mu}(x)}$ のあいだには、仮定 2 と仮定 3 のもとで、つぎの基本的な性質が成り立つ (Nakai [7] など)。

補題 9 $\boldsymbol{\mu} \succ \boldsymbol{\nu}$ ならば、任意の y に対して、$\boldsymbol{\mu}(y) \succ \boldsymbol{\nu}(y)$ および $\overline{\boldsymbol{\mu}(y)} \succ \overline{\boldsymbol{\nu}(y)}$ である。任意の $\boldsymbol{\mu}$ に対して、$\boldsymbol{\mu}(y)$ と $\overline{\boldsymbol{\mu}(y)}$ は y に関する増加関数である。

補題 9 から、事前情報 $\boldsymbol{\mu}$ における順序関係は、$\boldsymbol{\mu}(y)$ と事後情報 $\overline{\boldsymbol{\mu}(y)}$ に対して保たれることがわかる。さらに、同じ事前情報 $\boldsymbol{\mu}$ であれば、観測した値 y が大きくなれば、事後情報 $\overline{\boldsymbol{\mu}(y)}$ もまたよくなる。

4.3 Gradually Condition

状態について、不完備な情報しか与えられていないときの最適決定問題を考えるために、いくつかの準備をする。ここで考えた支出モデルでは、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。まず、以下の議論では 2.2 節にあるように、$\sigma(s, x) = s + d(x)$ とする。ここで、

$$\boldsymbol{\mu}_x(t) = \int_0^\infty \mu_x(s)p_s(t)ds = \int_0^\infty \mu(s)p_{s(x)}(t)ds. \tag{4}$$

とおく。これは、事前情報が $\boldsymbol{\mu}$ のとき、追加して x を支出したときの、状態空間上の事後分布を表している。ここで、$s(0) = s$ だから、$\overline{\boldsymbol{\mu}} = \int_0^\infty \mu(s)p_s(t)ds = \boldsymbol{\mu}_0$ である。

状態全体の集合 $\mathcal{S}$ に含まれる確率分布 $\boldsymbol{\mu}$ が

> $s < t, s' < t'$ と $s - s' = t - t' = c < 0$ を満たす任意の $s < s'$, $t \leq t'$ に対して、
>
> $$\frac{\mu(s)}{\mu(s')} \geq \frac{\mu(t)}{\mu(t')}$$

の性質を満たすとき、この $\boldsymbol{\mu}$ は gradually condition を満足するということにする。

例 1 状態空間上の正規分布 $\mu(s) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(s-a)^2}{2\sigma^2}}$ はこの性質を満足する。

補題 10 状態全体の集合 $\mathcal{S}$ に含まれる確率分布 $\boldsymbol{\mu}$ が *gradually condition* を満足するとき、$x > y$ ならば、$\boldsymbol{\mu}_x \succeq \boldsymbol{\mu}_y$ である。ただし、$\boldsymbol{\mu}_x = (\mu_{s(x)})$ とする。

補題 11 状態全体の集合 $\mathcal{S}$ に含まれる確率分布 $\boldsymbol{\mu}$ と $\boldsymbol{\nu}$ が *gradually condition* を満足するとき、$\boldsymbol{\mu} \succeq \boldsymbol{\nu}$ ならば、任意の $x (\geq 0)$ に対して、$\boldsymbol{\mu}_x \succeq \boldsymbol{\nu}_x$ である。

ここで、

$$\overline{\mu}(t) = \int_0^\infty \mu(s)p_s(t)ds$$

とおく。つぎの性質を導くため、推移法則に関してつぎの仮定をおく。

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

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

とする。すなわち、

$$\begin{vmatrix} p_u(s) & p_u(t) \\ p_v(s') & p_v(t') \end{vmatrix} \ge \begin{vmatrix} p_v(s) & p_v(t) \\ p_u(s') & p_u(t') \end{vmatrix}$$

である。

補題 12 状態全体の集合 $\mathcal{S}$ に含まれる確率分布 μ が *gradually condition* を満足するならば、$\overline{\mu}$ もまた *gradually condition* を満足する。

例 2 正規分布による推移法則 $p_v(s) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(s-v)^2}{2\sigma^2}}$ は、仮定 4 の条件を満足する。

確率変数 Y_s の密度関数 $f_s(y)$ が $(s \in [0, \infty))$、任意の $s < s', t < t'$ で $t - s = t' - s' > 0$ となる s, s', t, t' に対して、性質

$$\frac{f_s(y)}{f_{s'}(y)} \ge \frac{f_t(y)}{f_{t'}(y)}$$

が成り立つと仮定した。このことから、情報プロセスからの観測値 y が得られたときの事後情報 $\overline{\mu(y)} = (\overline{\mu(y, s)})$ すなわち、

$$\overline{\mu(y, s)} = \frac{\mu(s)f_s(y)}{\int_0^\infty \mu(s)f_s(y)ds}$$

は、つぎの性質を持つ。

補題 13 状態全体の集合 $\mathcal{S}$ に含まれる確率分布 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。

implies ここで、観測できない状態に関する情報が μ で、追加して支出した額が x のときの、状態空間上の確率分布を

$$\widetilde{\mu(x)} = \overline{\mu_x}(t) = \int_0^1 \mu(s)p_{s(x)}(t)ds \tag{5}$$

とおく。このとき、これまでの性質をまとめれば、つぎの補題が得られる。

補題 14 状態全体の集合 S に含まれる確率分布 $\boldsymbol{\mu}$ が *gradually condition* を満足するならば、$\widetilde{\boldsymbol{\mu}(x)}$ もまた *gradually condition* を満足する。

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

補題 6 から、任意の x に対して推移法則 $(p_{s(x))}(t))_{0\leq s\leq 1}$ が TP_2 であるから、これまでに議論してきた仮定の下で、つぎの性質が成り立つ。

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

4.4 逐次支出モデル–不完備情報の場合

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、観測できない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。観測できない状態に関する情報が $\boldsymbol{\mu}$ 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 $\widetilde{V}_n(\boldsymbol{\mu})$ 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。

$$
\begin{aligned}
\widetilde{V}_n(\boldsymbol{\mu}) &= \int_0^\infty \widetilde{V}_n(\boldsymbol{\mu}|y)d\boldsymbol{\mu}(y) \\
\widetilde{V}_n(\boldsymbol{\mu}|y) &= \max_{x\geq 0}\left\{-c(x) + \widetilde{V}_{n-1}(\boldsymbol{\mu}\widetilde{(y)}(x))\right\} \qquad (6)
\end{aligned}
$$

ここで、$\widetilde{V}_0(\boldsymbol{\mu}) = \int_0^1 u(t)d\mu(t)$ とする。(6) 式において、$\boldsymbol{\mu}(y)$ を情報プロセスから得られた値 y をもとに、情報を改良し多あとの状態に関する情報とする。 すなわち、事前情報が $\boldsymbol{\mu}$ 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 $\boldsymbol{\mu}(y)$ 改良するのである。追加した支出 x を行ったあとで、観測できない状態が s のとき、推移法則 $(p_{s(x)}(t))_{0\leq s\leq 1}$ にしたがって 1 期間移動する。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 (5) 式のように、$\mu\widetilde{(y)}(x)$ 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 $\tilde{V}_{n-1}(\mu\widetilde{(y)}(x))$ となる。よって、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、$\tilde{V}_n(\mu) \geq \tilde{V}_n(\nu)$ である。

もし、$\mu \succ \nu$ であれば、補題 9 より任意の観測値 y に対して、$\mu(y) \succ \nu(y)$ であり、補題 11 から、任意の支出額 x に対して、$\widetilde{\mu(x)} \succeq \widetilde{\nu(x)}$ となっている。これらの事後情報に関する単調性から、つぎのことがいえる。すなわち、任意の支出額 x と観測値 y に対して、$\mu \succ \nu$ ならば、$\mu\widetilde{(y)}(x) \succeq \nu\widetilde{(y)}(x)$ であり、このことから性質 3 が n に関する帰納法によって示すことができる。

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, A Sequential Stochastic Assignment Problem in a Partially Observable Markov process, *Mathematics of Operations Research*, **11**, 230–240, (1986).

[5] 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).

[6] 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).

[7] 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).

[8] 中井　達, 効率性と有効性—評価と最適化,『政策分析 2003–政策・制度への歴史的接近の視軸から』(九州大学大学院経済学研究院政策評価研究会編), 九州大学出版会,　275–301, 2003.

[9] 中井　達, 評価技法と政策評価について,『政策分析 2004–国際化・分権化時代の日本経済の存立基盤–』（九州大学大学院経済学研究院政策評価研究会編）, 九州大学出版会, 405–448, 2004.

[10] 中井　達,『政策評価–費用便益分析から包絡分析法まで–』, ミネルヴァ書房, 京都, 2005.

[11] T. Nakai, Economy, Efficiency and Effectiveness, *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.

[12] T. Nakai, Properties of a Job Search Problem on a Partially Observable Markov Chain in a Dynamic Economy, *Computers & Mathematics with Applications*, The Special Issue on The Second Euro-Japanese Workshop on Stochastic Risk Modelling for Finance, Insurance, Production and Reliability, vol. 51, 189–198, 2006.

[13] 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.

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