

ある不完備情報の多段決定モデルについて

中井 達

九州大学経済学部

1 不完備情報の動的決定モデル

不完備情報の決定モデルを考えると、情報は状態空間上の確率分布で与えられることが多く、これらの確率分布を比較する必要がある。これらの情報を比較するためには、情報の間に何らかの順序を導入することは自然である。Nakai [6, 7, 9] などにおいて、 TP_2 (totally positivity of order two) として知られている尤度比を用いた順序を導入し、この順序と多段決定モデルにおける最適政策やその政策にしたがったときに得られる値の性質について考えた。

第2節では、部分観測可能なマルコフ連鎖を考え、尤度比を用いた順序関係を導入する。さらに、それぞれの時点で、そこでは多変量の確率変数を観測することで、情報を得るものとする。ここでは、学習のプロセスとしては、ベイズの定理に従うものとし、事前情報と事後情報の関係について考える。特にこれらの関係には、基本的な関係が得られることがわかる。Brown and Solomon においても、同様の順序関係を [1] で扱い、Nakai [6, 7, 9, 11] で得られた結果をまとめておく。したがって、ここではそれぞれの時点で観測する多変量の確率変数は互いに独立とは限らず、 MTP_2 (multi variate totally positivity of order two) と呼ばれる性質を考える。また、Holley [2], Kemperman [5], Preston [12], Karlin and Rinott [3, 4] など得られた性質を利用する。

第3節では、複数回選択できる最適停止問題を、この枠組みの中で解析し、不完備情報の決定モデルに関連した性質を求める。さらに、最適政策と、その政策にしたがって得られる値の性質について考える。

2 部分観測可能なマルコフ決定モデル

まずはじめに、尤度比を用いて確率分布の間に順序を導入する。

定義 1 いま、 X と Y を非負な多変量確率変数とし、それらの確率密度関数をそれぞれ f_X および f_Y とする。いま、

$$f_X(y)f_Y(x) \leq f_X(x \vee y)f_Y(x \wedge y)$$

のとき、確率変数 X は、確率変数 Y より尤度比の意味で大きいと言い、 $X \geq_l Y$ と表す。

ここで、 $x \wedge y = (\min(x_1, y_1), \dots, \min(x_k, y_k))$ および $x \vee y = (\max(x_1, y_1), \dots, \max(x_k, y_k))$ とする。この順序関係を MTP_2 (multivariate totally positivity of order two) という。この順序は、 TP_2 の性質の一般化である。いま、 $S = \left\{ \Phi \mid \Phi = (\phi_0, \phi_1, \phi_2, \dots), \phi_s \geq 0, \sum_{s=0}^{\infty} \phi_s = 1 \right\}$ を、非負整数の集合 $\{0, 1, 2, \dots\}$ 上の確率分布全体を表し、この集合に尤度比を用いて、順序を導入する。

定義 2 いま、 Φ と Ψ を、 S に含まれる 2 つの確率分布とする。すべての 2 つの整数の組 i と j ($i \leq j, i, j = 0, 1, 2, \dots$) に対して、

$$\phi_j \psi_i \geq \phi_i \psi_j, \text{ i.e., } \begin{vmatrix} \phi_j & \psi_j \\ \phi_i & \psi_i \end{vmatrix} \geq 0 \quad (1)$$

とし、少なくとも 1 つの組み合わせ i と j に対して、 $\phi_j \psi_i > \phi_i \psi_j$ とする。このとき、 $\Phi >_l \Psi$ と表す。また、すべての $i = 0, 1, 2, \dots$ に対して、 $\phi_i = \psi_i$ のとき、 $\Phi =_l \Psi$ とする。さらに、 $\Phi =_l \Psi$ かつ $\Phi >_l \Psi$ のとき、 $\Phi \geq_l \Psi$ とする。

補題 1 定義 2 で導入した順序は半順序である。

つぎに、部分観測可能なマルコフ連鎖を考え、このマルコフ連鎖の状態を直接に知ることができないとする。ここでは、情報過程を通して、状態に関する情報を得ることができる。Nakai [6, 7, 11] において、いくつかの基本的な性質が得られている。また、部分観測可能なマルコフ連鎖での多段決定モデルについて、[9] で考えられている。

いま、 $\{0, 1, 2, \dots\}$ を、このマルコフ連鎖の状態全体を表す集合とし、 $P = (p_{i,j})_{i,j=0,1,2,\dots}$ をその推移確率行列とする。ここで、これらの状態に依存する確率変数から得られる標本値を得て、この状態に関する情報を得る。したがって、これらの確率変数を観測するプロセスが、情報過程になる。このマルコフ連鎖の状態が i のとき、この状態に依存する、非負の k 次元多変量確率変数を X_i とする。さらに、この絶対連続な確率変数の、分布関数を

$$\Pr(X_s \leq x \mid Y_n = i) = F_i(x) \quad (x \in R^k, i \in \{0, 1, 2, \dots\}, n \in \{0, 1, 2, \dots\}), \quad (2)$$

とし、その確率密度関数を $f_i(x)$ とする。ここで、 Y_n は時点 n でのこのマルコフ連鎖の状態を表す確率変数とする。

また、このマルコフ連鎖の状態に関する情報は、状態空間上の確率分布 $\Phi = (\phi_0, \phi_1, \phi_2, \dots)$ で表されているとする。集合 S に対し、定義 2 で考えた半順序を仮定する。さらに、すべての標本値 x と事前情報 Φ に対して、事後情報は存在し、その学習プロセスは、ベイズの定理によるものとする。

したがって、状態についての事前情報が Φ のとき、推移確率行列 P にしたがって、マルコフ連鎖の状態が推移するから、状態に関する情報は

$$\begin{cases} \bar{\phi}_j = \sum_{i=0}^{\infty} \phi_i p_{ij}, \\ \bar{\Phi} = (\bar{\phi}_0, \bar{\phi}_1, \bar{\phi}_2, \dots) \end{cases} \quad (3)$$

となる。つぎに、標本値 $x \left(\in R_+^k = (0, \infty)^k \right)$ を得てから、状態に関する情報をベイズの定理にしたがって修正し、 $T(\bar{\Phi}|x)$ となる。すなわち、任意の $j = 0, 1, 2, \dots$ に対して

$$\begin{cases} T_j(\bar{\Phi}|x) = \frac{\bar{\phi}_j f_j(x)}{\sum_{i=0}^{\infty} \bar{\phi}_i f_i(x)} \\ T(\bar{\Phi}|x) = (T_0(\bar{\Phi}|x), T_1(\bar{\Phi}|x), T_2(\bar{\Phi}|x), \dots) \end{cases} \quad (4)$$

となる。つぎに、事前情報と事後情報の関係を 3 つの仮定の下で考察する。

仮定 1 マルコフ連鎖の状態が i のとき、条件付き期待値 $\mu_i = E[X|Y = i]$ は有界である。ここで、 Y をマルコフ連鎖の状態を表す確率変数とする。

仮定 2 マルコフ連鎖の状態が i のとき、 X^i を k 次元の多変量確率変数とし ($i = 0, 1, 2, \dots$)、その確率密度関数を $f_i(x)$ とする。このとき、

$$f_{i \wedge j}(x \wedge y) f_{i \vee j}(x \vee y) \geq f_j(y) f_i(x) \quad (5)$$

とする。

仮定 3 状態空間が $\{0, 1, 2, 3, \dots\}$ のマルコフ連鎖で、その推移確率行列を $(p_{ij})_{i,j=0,1,2,3,\dots}$ とし、 TP_2 の性質を持つとする。すなわち、任意の i と j ($i \geq j, i, j = 0, 1, 2, \dots$) にたいして、 $p_{mj} p_{ni} \geq p_{nj} p_{mi}$ である。ただし、 $m \leq n$ ($m, n = 1, 2, \dots$) とする。

これらの性質をみたすとき、この部分観測可能なマルコフ連鎖は MTP_2 の性質を持つという。

このとき、仮定 1 と 2, 3 のもとで、(3) 式と (4) 式で定義された事後情報に関する基本的な性質が成り立つ。

部分観測可能なマルコフ連鎖の状態について、それぞれの期で多変量の確率変数から得られる標本値を用いて情報を得る。また、すべての情報は状態空間上の確率分布 Φ によって与えられるとする ($\Phi \in S$)。それぞれの期で、これらの標本値をもとに、状態に関する情報を改良する。いま、 $\Phi \in S$ を状態に関する事前情報とする。 k 次元の多変量確率変数からの標本値 $\{x_m\}_{m=1,\dots,k}$ に対して、ベイズの定理によって情報を $T(\bar{\Phi}|x)$ とする。

はじめに、マルコフ連鎖の状態が推移確率行列 P にしたがって推移し、状態についての情報は、(3) 式によって $\bar{\Phi}$ となる。つぎに、標本値を知って情報を (4) 式から $T(\bar{\Phi}|x)$ と学習する。ここで、 $x = (x_1, \dots, x_k)$ とする。

定理 1 すべての $\Phi \in \mathcal{S}$ に対し、 $x \prec y$ ならば、 $T(\bar{\Phi}|x) \leq_i T(\bar{\Phi}|y)$ である。

補題 2 任意の Φ と $\Psi \in \mathcal{S}$ に対し、 $\Phi \geq_i \Psi$ ならば、 $\bar{\Phi} \geq_i \bar{\Psi}$ である。

定理 2 任意の $x \in R^k$ で $\Phi \geq_i \Psi$ とすれば $T(\bar{\Phi}|x) \geq_i T(\bar{\Psi}|x)$ である。

これらの性質は、Nakai [10] で用いたものと同様の手法を用いて示すことができる。つぎに、 k 次元の多変量確率変数からの標本のあいだに順序を定義する。

定義 3 いま、 $x = (x_1, \dots, x_k)$ と $y = (y_1, \dots, y_k) \in R^k$ を、 k 次元の多変量確率変数 $X = (X_1, \dots, X_k)$ からの二つの標本とする。ここで、 $x_i \leq y_i$ ($i = 1, 2, \dots, k$) のとき、 x は y より小さいといい、 $x \prec y$ と表す。

また、状態に関する情報全体の集合 $\mathcal{S}$ に、定義 2 で考えた順序を仮定する。Nakai [7, 8, 9, 10, 11] などでのこの順序について考察されている。また、[7, 8] では、状態の数が有限の場合を、[9] では、可算個の場合を扱っている。また、それぞれの期で観測できる標本が、 k 個の独立な確率変数の場合にはつぎの方法で表すことができる。

いま、 $x_{(1)}, \dots, x_{(k)}$ を、 k 個の独立な確率変数から得られる標本 $x_1, \dots, x_k$ 順序統計量とする ($x_{(1)} \geq \dots \geq x_{(k)}$)。ここでは、便宜上、標本を値の大きなものから小さいものへ並べる。このとき、 k 個の標本の間につぎの順序を導入し、つぎの仮定を設ければ良い。

定義 4 2 つの k 個の標本 $x, y \in R^k$ にたいして、 $x_{(i)} \leq y_{(i)}$ ($i = 1, 2, \dots, k$) のとき、 $x \prec y$ とする。

補題 3 任意の x と y に対し、 $x \prec y$ ならば $f_j(y)f_i(x) \geq f_i(y)f_j(x)$, $i < j$ ($i, j = 1, 2, \dots$) である。

Nakai [10] で、この場合に事前情報と事後情報のあいだに成り立つ 3 つの性質が得られている。また、 $n = 1$ のときについては、[7, 8] に詳しい。

標本 x が得られたとき、事後情報はつぎの性質を持つ。

定理 3 すべての $\Phi \in \mathcal{S}$ に対して、 $T(\bar{\Phi}|x)$ は、 MTP_2 の性質を持つ。

定理 1 と 2 は、この場合にも成り立つ。 $x \prec y$ から $x \wedge y = x$ および $x \vee y = y$ が成り立つから、つぎの性質は明かである。

補題 4 $f_i(x)$ が仮定 2 を満たすものとする。いま、 $x \prec y$ ならば、

$$f_j(y)f_i(x) \geq f_i(y)f_j(x) \quad (6)$$

が、任意の $i < j$ ($i, j = 1, 2, \dots$) に対して成り立つ。すなわち、すべての $i = 0, 1, 2, \dots$ に対して $f_i(x)$ は x に関して MTP_2 の性質を持つ。

この補題から、定理 1 で得られたのと同様の性質が導かれる。また、 $X_1, \dots, X_k$ が互いに独立なときには、補題 3 より $x \prec y$ かつ $i < j$ ($i, j = 1, 2, \dots$) となるすべての x と y に対して (6) 式が導かれる。このことは、 MTP_2 の性質に他ならない。一方、(6) 式と、 x ($i = 0, 1, 2, \dots$) に対して $f_i(x)$ が MTP_2 の性質を持つことを仮定する。このとき、簡単な計算から補題 5 が得られる。

補題 5 任意の $i < j$ ($i, j = 1, 2, \dots$) に対して、(6) 式を仮定する。また、 $f_i(x)$ は x ($i = 0, 1, 2, \dots$) に関して MTP_2 の性質を持つとする。このとき、仮定 2 が成り立つ。

補題 6 $X_1, \dots, X_k$ が独立のとき、(6) 式から仮定 2 が導かれる。

定義 5 k 変数関数 $\varphi: R^n \rightarrow R$ が、 $x \prec y$ のとき $\varphi(x) \leq \varphi(y)$ ($\varphi(x) \geq \varphi(y)$) ならば、この関数は x に関して非減少 (非増加) 関数という。

Holley [2], Kemperman [5], Preston [12], Karlin and Rinott [3, 4] において、性質 1 と 2 が得られている。

性質 1 $f_i(x)$ を R^k 上の確率密度関数とし、

$$f_1(x \wedge y) f_2(x \vee y) \geq f_1(x) f_2(y) \quad (7)$$

とする。このとき、

$$\int \varphi(x) f_1(x) dx \leq \int \varphi(x) f_2(x) dx \quad (8)$$

が、 x に関して増加する、非負可測関数 $\varphi(x)$ に対して成り立つ。

補題 7

$$f_i(x \wedge y) f_j(x \vee y) \geq f_j(y) f_i(x)$$

とする。ただし、 $x, y \in R^k$ かつ $i \leq j$ ($i, j = 1, 2, 3, \dots$) である。このとき、

$$\int \varphi(x) f_i(x) dx \leq \int \varphi(x) f_j(x) dx$$

が、 x に関して増加する任意の関数 $\varphi(\cdot)$ に対して成り立つ。

補題 8 S に含まれるすべての Φ と Ψ に対して、 $\Phi \geq \Psi$ ならば

$$E_{\Psi}[\varphi(X)] = \sum_{i=0}^{\infty} \phi_i \int \varphi(x) f_i(x) dx \leq \sum_{i=0}^{\infty} \psi_i \int \varphi(x) f_i(x) dx = E_{\Phi}[\varphi(X)]$$

が、 x に関して増加する任意の関数 $\varphi(\cdot)$ に対して成り立つ。

性質 2 $f(x_1, \dots, x_k)$ を、 k 次の多変量確率変数 $X = (X_1, \dots, X_k)$ の同時密度関数とする。この関数が MTP_2 の性質を持てば、周辺密度関数 $f(x_{i_1}, \dots, x_{i_m})$ もまた、 MTP_2 の性質を持つ ($i_1 < \dots < i_m, \{i_1, \dots, i_m\} \subset \{1, 2, \dots, k\}$)。

補題 9 $f(x_1, \dots, x_k)$ を k 次の多変量確率変数 $X = (X_1, \dots, X_k)$ の同時確率密度関数とする。この関数が、 MTP_2 の性質を持てば、周辺密度関数 $f(x_m)$ は TP_2 の性質を持つ ($m = 1, 2, \dots, k$)。

3 部分観測可能なマルコフ連鎖での最適停止問題

前節までで扱った部分観測可能なマルコフ連鎖での n 期間の最適停止問題を考える。それぞれの期では、 k 次の多変量確率変数から得られる標本 $x = (x_1, \dots, x_k)$ を観測して、この期で停止するかどうかを決定する。この確率変数は、部分観測可能なマルコフ連鎖の状態に依存し、状態に関する情報はすべて状態空間上の確率分布で表されているものとする。このとき、停止すれば標本の大きさに依存する利得 $\varphi(x)$ を得ることができる。停止しなければ、この標本から状態についての情報を得て、つぎの期に進み新たな標本を観測する。利得関数 $\varphi(x)$ は x に関して増加する関数とする。例えば、 $\varphi(x) = \max_{1 \leq i \leq k} x_i$ とすれば、この関数は x に関して増加することは簡単にわかる。

このマルコフ連鎖の状態についての情報が Φ のとき、 n 期間のあいだ最適政策にしたがったときに得られる総期待利得を $v_n(\Phi)$ とすれば、最適性の原理より、つぎの最適方程式が得られる。

$$v_n(\Phi) = E_{\Phi}[v_n(\Phi|X)] \quad (9)$$

$$v_n(\Phi|x) = \max \{ \varphi(x), v_{n-1}(T(\bar{\Phi}, x)) \}. \quad (10)$$

これらの関数と、前節で得られた性質よりつぎの結果が得られる。

補題 10 $v_n(\Phi)$ は Φ に関して増加する。すなわち、 $\Phi \leq \Psi$ ならば $v_n(\Phi) \leq v_n(\Psi)$ である。

補題 11 $v_n(\Phi|x)$ は Φ と x に関して増加する。

これらの性質は n に関する帰納法を用いて示される。 $n = 1$ の場合は明かである。これらの性質を $n - 1$ 以下の場合に成り立つとする。定理 2 より、 $\Phi \geq \Psi$ ならば、 $T(\bar{\Phi}|x) \geq T(\bar{\Psi}|x)$ である。このことから、 $v_n(\Phi|x)$ は Φ について増加する関数である。一方、補題 4 より定理 1 が成り立つから、 $x \prec y$ ならば $T(\bar{\Phi}|x) \leq T(\bar{\Phi}|y)$ である。帰納法の仮定より、 $v_n(\Phi|x)$ は x に関する増加関数であることがわかる。補題 8 と 11 より、補題 10 が得られる。

つぎに、 R^k に含まれる領域 $S_n(\Phi)$ を $S_n(\Phi) = \{x | \varphi(x) \geq v_{n-1}(T(\bar{\Phi}, x))\}$ とする。この領域は、この問題の停止領域を示し、したがって最適政策を定めるものである。この領域に関して、つぎの性質が得られる。

補題 12 領域 $S_n(\Phi)$ に対して、 $\Phi \leq \Psi$ ならば $S_n(\Psi) \subset S_n(\Phi)$ である。

補題 11 より、 $v_n(\Phi|x)$ は Φ に関して増加することがわかる。さらに、 $\Phi \leq \Psi$ ならば、 $\varphi(x) \geq v_{n-1}(T(\bar{\Psi}, x)) \geq v_{n-1}(T(\bar{\Phi}, x))$ も示される。これらのことからこの補題が導かれる。

つぎに、Nakai [7, 8, 10] で考えられた問題と同様の、最適選択問題として知られる多段決定問題について考える。この問題は、 n 期間の決定モデルで、それぞれの期で、 k 次の多変量確率変数から得られる標本値を観測し、それらの中から m 個を選択して総期待利得を最大にする問題である。部分観測可能なマルコフ連鎖の状態に関する情報が Φ のとき、 $v_n^m(\Phi)$ を最適政策にしたがったときに得られる総期待利得とする。このとき、最適性の原理よりつぎの再帰方程式が得られる。

$$v_n^m(\Phi) = E_{\Phi}[v_n^m(\Phi|X)] \quad (11)$$

$$v_n^m(\Phi|x) = \max \{ \varphi(x) + v_{n-1}^{m-1}(T(\bar{\Phi}, x)), v_{n-1}^m(T(\bar{\Phi}, x)) \}. \quad (12)$$

これらの再帰方程式と、前節で得られた性質を用いてつぎの結果が得られる。

補題 13 $v_n^m(\Phi|x)$ は m に関する増加関数である。

補題 14 $v_n^m(\Phi)$ は Φ に関する増加関数である。すなわち、 $\Phi \leq \Psi$ ならば $v_n^m(\Phi) \leq v_n^m(\Psi)$ である。

補題 15 $v_n^m(\Phi|x)$ は x に関する増加関数である。

補題 13 は定義から明らかであり、残りの性質は n に関する帰納法で示すことが出来る。つぎに、関数 $h_n^m(\Phi|x)$ をつぎのように定義する。

$$h_n^m(\Phi|x) = v_{n-1}^m(T(\bar{\Phi}, x)) - v_{n-1}^{m-1}(T(\bar{\Phi}, x)) \quad (13)$$

このとき、つぎの性質が得られる。

補題 16 $h_n^m(\Phi|x)$ は m に関する非増加関数である。

また、 R^k に領域 $S_n^m(\Phi)$ を $S_n^m(\Phi) = \{x | \varphi(x) \geq h_n^m(\Phi|x)\}$ で定義する。この領域は、前にも述べたと同じように、この最適選択問題の選択領域を示し、この問題の最適政策を表す。この領域に関して、つぎの性質が成り立つ。

補題 17 領域 $S_n^m(\Phi)$ に対して、 $\Phi \leq \Psi$ ならば、 $S_n^m(\Psi) \subset S_n^m(\Phi)$ である。

性質 3 最適政策にしたがったときに得られる総期待利得 $v_n^m(T(\bar{\Phi}, x))$ はつぎの関係式を満足する。

$$v_n^m(T(\bar{\Phi}, x)) = \sum_{i=1}^m h_n^i(\Phi|x)$$

参考文献

- [1] M. Brown and H. Solomon, Optimal Issuing Policies under Stochastic Field Lives, *Journal of Applied Probability*, vol. 10, 761–768, 1973.
- [2] R. Holley, Remarks on the FKG Inequalities, *Communications in Mathematical Physics*, vol. 36, pp. 227–231, 1974.
- [3] S. Karlin and Y. Rinott, Class of Orderings of Measures and Related Correlation Inequalities I : Multivariate Totally Positive Distributions, *Journal of Multivariate Analysis*, vol. 10, 467–498, 1980.
- [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*, vol. 9, 1035–1049, 1981.
- [5] J. H. B. Kemperman, On the FKG-Inequality for Measures on a Partially Ordered Space, *Indagationes Mathematicae*, vol. 39, pp. 313–331, 1977.
- [6] T. Nakai, Optimal Stopping Problem in a Finite State Partially Observable Markov Chain, *Journal of Information & Optimization Sciences*, vol. 4, 159–176, 1983.
- [7] T. Nakai, The Problem of Optimal Stopping in a Partially Observable Markov Chain, *Journal of Optimization Theory and Applications*, vol. 45, 425–442, 1985.
- [8] T. Nakai, A Sequential Stochastic Assignment Problem in a Partially Observable Markov Chain, *Mathematics of Operations Research*, vol. 11, 230–240, 1986.
- [9] T. Nakai, A Stochastic Ordering and Related Sequential Decision Problems, *Journal of Information & Optimization Sciences*, vol. 11, 49–65, 1990.
- [10] T. Nakai, A Partially Observable Decision Problem under a Shifted Likelihood Ratio Ordering *Proceedings of the Australia–Japan Workshop on Stochastic Models in Engineering, Technology and Management* (Eds. S. Osaki and D. N. Pra Murthy), World Scientific Publishing, 413–422, 1993.
- [11] T. Nakai, A Learning Procedure for Sequential Decision Problem with Multiple Observations per Period, *Proceedings for 'The First International Conference on Operational and Quantitative Management' Jaipur*, vol. 1, 127–134, 1997.
- [12] C. J. Preston, A Generalization of the FKG Inequalities, *Communications in Mathematical Physics*, vol. 36, pp. 233–241, 1974.