部分観測可能なマルコフ過程における決定問題について

01402656　九州大学大学院経済学研究院　中井　達 NAKAI Tōru

1　はじめに

部分観測可能なマルコフ過程での多段決定問題を解析することを考える。そのために、MTP_2(multivariate total positivity of order two) を一般化し、このマルコフ過程の状態に関するベイズ学習の性質を求める。このマルコフ過程の状態空間を完備で可分な全順序が定義された距離空間とし、すべての情報は状態空間上の確率測度で表されるとする。Nakai [2] などでは、部分観測可能なマルコフ連鎖での学習プロセスと多段決定問題の関係について解析し、最適選択問題に関連して MTP_2 の場合に解析を行ったきたが、. その場合にも非減少関数の期待値に関して順序を保つ性質を持ち、これは [1], [4] などの FKG 不等式と共通な性質である。ここでは、MTP_2 を一般化した順序を定義し、命題 1 で同様の性質を持つことがわかる。また、ここでは学習プロセスをベイズの定理を用いたことから、この性質を用いた仮定の下で事前情報と事後情報との関係も求められる。

2　一般化した MTP_2

$\mathcal{S}$ を完備で可分な全順序 $\leq$ が定義された距離空間で、$\mathcal{B}$ をボレル集合とする可測空間を $(\mathcal{S},\mathcal{B})$ とする。直積空間 $\mathcal{S}^n=\prod_{i=1}^{n}\mathcal{S}$ に自然な半順序を仮定し、直積ボレル集合を $\mathcal{B}^n=\prod_{i=1}^{n}\mathcal{B}$ とする ($\mathcal{S}^1=\mathcal{S},\mathcal{B}^1=\mathcal{B}$)。直積空間 $(\mathcal{S}^n,\mathcal{B}^n)$ に対し、$P(\mathcal{S}^n)$ を直積空間上の確率測度全体の集合とする。

定義 1 $\mathcal{S}^n$ の部分集合 $A^n=\prod_{i=1}^n A_i$ と $B^n=\prod_{i=1}^n B_i (A_i,B_i\subset\mathcal{S},A_i\cap B_i=\emptyset(i=1,\cdots,n))$ に対し、$A_i\prec B_i$ ならば $A_i\vee B_i=B_i$ および $A_i\wedge B_i=A_i$ とする。このとき、A^n,B^n に対して、$\vee$ と $\wedge$ を $A^n\vee B^n=\prod_{i=1}^n A_i\vee B_i, A^n\wedge B^n=\prod_{i=1}^n A_i\wedge B_i$ とする。

定義 2 $(\mathcal{S}^n,\mathcal{B}^n)$ 上の確率測度 μ_n,ν_n が、ボレル集合 $A^n=\prod_{i=1}^n A_i, B^n=\prod_{i=1}^n B_i$ で $A_i,B_i\subset\mathcal{S},A_i\cap B_i=\emptyset(i=1,\cdots,n)$ となるものに対して、$\mu_n(A^n\vee B^n)\nu_n(A^n\wedge B^n)-\mu_n(A^n)\nu_n(B^n)\geq 0$ であり、少なくとも 1 つの組み合わせ A^n,B^n に対して $\mu_n(A^n\vee B^n)\nu_n(A^n\wedge B^n)>\mu_n(A^n)\nu_n(B^n)$ のとき、$\mu_n\succ\nu_n$ とする。任意の $A^n\in\mathcal{B}^n$ に対し、確率 1 で $\mu_n(A^n)=\nu_n(A^n)$ のとき $\mu_n=\nu_n$ とする。$\mu_n=\nu_n$ または $\mu_n\succ\nu_n$ のとき、$\mu_n\succeq\nu_n$ とする。

確率測度が絶対連続であれば、この定義は MTP_2 に等しく、この関係を一般化した MTP_2 という (GMTP_2)。

補題 1 $\mu_n,\nu_n\in P(\mathcal{S}^n)$ で、$\mu_n\succeq\nu_n$ とする。μ_{n-1},ν_{n-1} を μ_n,ν_n の周辺測度とすれば $(\mu_{n-1}(A^{n-1})=\mu_n(A^{n-1}\times\mathcal{S}),\nu_{n-1}(B^{n-1})=\nu_n(B^{n-1}\times\mathcal{S})$、$\mu_{n-1}\succeq\nu_{n-1}$ である。

いま、k 変数関数 $\varphi:\mathcal{R}_+^k\to\mathcal{R}_+$ が、$\boldsymbol{x}\prec\boldsymbol{y}$ に対して $\varphi(\boldsymbol{x})\leq\varphi(\boldsymbol{y})$ $(\varphi(\boldsymbol{x})\geq\varphi(\boldsymbol{y}))$ のとき、この関数を $\boldsymbol{x}$ の非減少 (非増加) 関数という。

命題 1 $\mu_n,\nu_n\in P(\mathcal{S}^n)$ で、$\mu_n\succeq\nu_n$ とするとき、$h:\mathcal{S}^n\to\mathcal{R}_+$ を有界な非減少関数とすれば $\int_{\mathcal{S}^n}h(s)d\mu_n(s)\geq\int_{\mathcal{S}^n}h(s)d\nu_n(s)$ である。

命題 1 より、GMTP_2 が非減少関数の期待値に関して順序を保つ性質を持ち、これは FKG 不等式 ([1],[4] など) と共通のものである。

3　部分観測可能なマルコフ過程

$\mathcal{S}$ を部分観測可能なマルコフ過程の状態空間とし、$(\mathcal{S},\mathcal{B})$ を可測空間とする。$P(\cdot|s)$ $(s\in\mathcal{S})$ を推移法則とし、任意の $B\in\mathcal{B}$ に対して $p(\cdot|s)$

を確率測度とすれば $P(B|s)=\int_B dp(t|s)$ である $(s\in\mathcal{S})$。また、任意の状態 s に対して、期待値有限の k 次確率変数 $\boldsymbol{X}_s$ を考え、この確率変数の標本値から状態についての情報を得る。この確率変数の確率分布を $\Pr(\boldsymbol{X}_s\in C)=\boldsymbol{F}_k(C|s)=\int_C d\boldsymbol{F}_k(x|s)$ とする $(s\in\mathcal{S}, C\in\mathcal{X}^k=\prod_{i=1}^{k}\mathcal{X}$。ここで、$\mathcal{X}$ を $\mathcal{R}_+$ のボレル集合とする。状態についての情報は $\mathcal{S}$ 上の確率測度で表され、$\boldsymbol{P}(\mathcal{S})$ を情報の集合とする。

標本値 $\boldsymbol{x}=(x_1,\cdots,x_k)$ $(\in\mathcal{R}_+^k)$ と事前情報 $\boldsymbol{\mu}$ に対して、事後情報は存在しベイズの方法で学習する。事前情報が $\boldsymbol{\mu}$ のとき、推移法則に従って状態が推移し、事後確率は $\overline{\mu}(B)$ となる。ここで、標本値 $\boldsymbol{x}\in\mathcal{R}_+^k$ を知って、事後情報を $\overline{\boldsymbol{\mu}(\boldsymbol{x})}$ とする。

つぎに、$\mathcal{R}_+^k=\prod_{i=1}^{k}\mathcal{R}_+, \mathcal{X}^k=\prod_{i=1}^{k}\mathcal{X}$ とする。ただし $\mathcal{X}$ を $\mathcal{R}_+$ のボレル集合とする。

仮定 1 $\boldsymbol{X}^k=\prod_{i=1}^k X_i, \boldsymbol{Y}^k=\prod_{i=1}^k Y_i$ を互いに素なボレル集合で、$X_i\cap Y_i=\emptyset\,(i=1,\cdots,k)$ とし $(\boldsymbol{X}^k,\boldsymbol{Y}^k\in\mathcal{X}^n)$、任意の $s,t\in\mathcal{S}$ に対して、$\boldsymbol{F}_k(\boldsymbol{X}^k\vee\boldsymbol{Y}^k|s\vee t)\boldsymbol{F}_k(\boldsymbol{X}^k\wedge\boldsymbol{Y}^k|s\wedge t)\geq\boldsymbol{F}_k(\boldsymbol{X}^k|s)\boldsymbol{F}_k(\boldsymbol{Y}^k|t)$ である。

仮定 2 $s\prec t$ のとき $s,t\in\mathcal{S}$ 、$A\prec B$ $(A,B\in\mathcal{B})$ ならば $\boldsymbol{P}(A|s)\boldsymbol{P}(B|t)\geq\boldsymbol{P}(B|s)\boldsymbol{P}(A|t)$ である。

仮定 1 と 2 のもとで、事前情報と事後情報のあいだには次の性質を持つ。

定理 1 $\boldsymbol{x}\prec\boldsymbol{y}$ ならば、$\overline{\boldsymbol{\mu}(\boldsymbol{x})}\succeq\overline{\boldsymbol{\mu}(\boldsymbol{y})}$ である $(\boldsymbol{\mu}\in\boldsymbol{P}(\mathcal{S}))$。$\boldsymbol{\mu}\succeq\boldsymbol{\nu}$ ならば、$\overline{\boldsymbol{\mu}(\boldsymbol{x})}\succeq\overline{\boldsymbol{\nu}(\boldsymbol{x})}$ である $(\boldsymbol{x}\in\mathcal{R}_+^k)$。

$\boldsymbol{P}(B|s)$ と $\boldsymbol{F}_k(\cdot|s)$ が絶対連続ならば、Nakai[2, 3] で求められている。つぎに、$\varphi(\boldsymbol{x})$ が $\boldsymbol{x}$ の非減少関数のとき、仮定 1 から $\Phi(s)=\mathrm{E}[\varphi(\boldsymbol{X}_s)]$ は s の非減少関数である。$\mathrm{E}\boldsymbol{\mu}[\varphi(\boldsymbol{X})]=\int_{\mathcal{S}}\left\{\int_{\mathcal{R}_+^k}\varphi(\boldsymbol{x})d\boldsymbol{F}(\boldsymbol{x}|s)\right\}d\boldsymbol{\mu}(s)$ だから、命題 1 より次の性質が成り立つ。

補題 2 $\boldsymbol{\mu}\succeq\boldsymbol{\nu}\,(\boldsymbol{\mu},\boldsymbol{\nu}\in\boldsymbol{P}(\mathcal{S}))$ のとき、$\boldsymbol{x}$ の非減少関数 $\varphi(\cdot)$ に対して、$\mathrm{E}\boldsymbol{\mu}[\varphi(\boldsymbol{X})]\geq\mathrm{E}\boldsymbol{\nu}[\varphi(\boldsymbol{X})]$ である。

4 部分観測可能なマルコフ過程での多段決定問題

最後に、この部分観測可能なマルコフ過程での最適停止問題を考える。状態についての事前情報を $\boldsymbol{\mu}$ とし、標本値 $\boldsymbol{x}\in\mathcal{R}_+^k$ を得たとき、利得 $\varphi(\boldsymbol{x})$ を得て停止するかどうかを決める。停止しなければ、事後情報を求め、次の値を観測する。この $\varphi(\boldsymbol{x})$ は $\boldsymbol{x}$ の非減少関数とすし、期待利得を最大にすることが目的である。事前情報が $\boldsymbol{\mu}$ のとき、$v_n(\boldsymbol{\mu})$ を n 帰還問題の最適値とすれば、次の最適方程式を満足する。

$$v_n(\boldsymbol{\mu})=\mathrm{E}\boldsymbol{\mu}[\max\{\varphi(\boldsymbol{X}),v_{n-1}(\overline{\boldsymbol{\mu}(\boldsymbol{X})})\}],\quad(1)$$

ただし、$v_n(\boldsymbol{\mu}|\boldsymbol{x}):=\max\{\varphi(\boldsymbol{x}),v_{n-1}(\overline{\boldsymbol{\mu}(\boldsymbol{x})})\}$ とする。定理 1 より次の性質が成り立つ。

命題 1 $v_n(\boldsymbol{\mu}|\boldsymbol{x})$ は $\boldsymbol{\mu}$ と $\boldsymbol{x}$ の非減少関数であり、$v_n(\boldsymbol{\mu})$ は $\boldsymbol{\mu}$ の非減少関数である。

いま、$S_n(\boldsymbol{\mu})=\{\boldsymbol{x}|\varphi(\boldsymbol{x})\geq v_{n-1}(\overline{\boldsymbol{\mu}(\boldsymbol{x})})\}$ とすれば、最適政策を決定する停止領域となる。

命題 2 $\boldsymbol{\mu}\succeq\boldsymbol{\nu}$ ならば、$S_n(\boldsymbol{\mu})\subset S_n(\boldsymbol{\nu})$ である。

参考文献

[1] C. M. Fortuin, P. W. Kasteleyn and J. Ginibre, Correlation Inequalities on Some Partailly Ordered Sets, *Comm. on Math. Phys.*, **22**, 89–103, 1971.

[2] T. Nakai, A Sequential Stochastic Assignment Problem in a Partially Observable Markov Chain, *Math. of Oper. Res.*, **11**, 230–240, 1986.

[3] T. Nakai, An Optimal Selection Problem on a Partially Observable Markov Chain, *Stochastic Modelling in Innovative Manufacturing*, Lec. Notes in Econ. and Math. Sys. **445**, 140–154, Springer, 1996.

[4] C. J. Preston, A Generalization of the FKG Inequaliteis, *Comm. on Math. Phys.*, **36**, 233–241, 1974.