数理解析研究所講究録 1383 巻 2004 年 87-102

多段決定問題と Total Positivity

中井　達 (Tōru Nakai)

九州大学大学院経済学研究院
(Faculty of Economics, Kyushu University)

1 はじめに

TP_2(total positive of order two) は、多段決定問題の性質を解析する上で基本的な性質であり、確率的逐次割り当て問題、最適選択問題、ジョブ・サーチ、取り替え問題など、多くの分野で応用されている。とくに、不完備情報の多段決定問題を、部分観測可能なマルコフ過程における決定問題としてモデル化するとき、学習プロセスをベイズの定理を用いて行うベイズ学習にしたがうときには、重要な役割を果たすことは知られている (Albright[1],Monahan[8],Ohnishi Kawai and Mine[17],Nakai[9, 10, 11, 12, 13, 14, 15] など)。そのために、Lippman and MacCall [7] で扱われた動的な経済 (dynamic economy) におけるジョブ・サーチについてみることにする。このジョブ・サーチは最適停止問題の一つであり、状態はあるマルコフ連鎖にしたがって推移し、状態が何であるかを知ることができる完全情報モデルであった。このモデルにおいては、仕事から得られる賃金などが、経済の状態に依存すると考え、経済の状態はいくつかのクラスに分けて考られている動的な経済のモデルである。このようなジョブ・サーチにおいては、マルコフ連鎖の推移法則と、状態に依存する賃金を表す確率変数に関して、ある条件の下で、その最適政策は取りうる賃金の集合を reservation wage(留保賃金) と呼ばれる値によって分割される 2 つの互いに素な集合によって決まることが知られている。しかし、一般的にはこの経済の状態がこれらのクラスのどこに属しているかは、未知であることが多く、このようなジョブ・サーチをモデル化するときには、Lippman and MacCall で考えられた仮定では充分ではなく、TP_2 の性質を用いた新たな仮定の下で、最適政策やその政策にしたがったときに得られる期待利得などの性質を求めることができる。一般的に、不完備情報のジョブ・サーチでは、完全情報のモデルのような reservation wage(留保賃金) が存在するとは言えない。ここでは、不完備情報のモデルとして部分観測可能なマルコフ過程におけるジョブ・サーチとして定式化し、このような未知の状態に関する情報については、事前に知ることができるものとする。このような部分観測可能なマルコフ過程の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で表され、学習過程はベイズの定理にしたがうものとすれば、TP_2 の性質を用いた仮定の下で、事前情報、事後情報、最適政策やマルコフ過程の状態の推移に関連する性質の関係をみることができる。

2 節では動的な経済におけるジョブ・サーチで、Lippman and MacCall で扱われたような状態を直接知ることができる完全情報の場合を簡単にまとめる。3 節では、状態が部分観測可

能なマルコフ過程にしたがって状態が推移するときに、ベイズの定理にしたがった学習過程の性質や、状態の推移確率に関する性質などについても考える (Nakai [16])。この TP_2 については、Karlin and McGregor [4]、Karlin [3]、Karlin and Rinott [5] などでも、確率過程との関連について研究されている。最後に、TP_2 を用いたもう一つの多段決定問題への応用として、Sequential Investment Problem について考える。

2 動的な経済 (dynamic economy) におけるジョブ・サーチ

2.1 最適政策と期待利得

状態空間を $[0,S]$ とし、任意の状態 $s \in [0,S]$ に対して、状態空間 $[0,S]$ 上の確率分布の確率密度関数を $\boldsymbol{p}_s = (p_s(t))_{t\in[0,S]}$ とする。ここで、$p_s(t) \geq 0$ であり $\int_0^S p_s(t)dt = 1$ である $(t \in [0,S])$。このとき、これらの $\boldsymbol{p}_s = (p_s(t))_{t\in[0,S]}$ は、状態が $s \in [0,S]$ のときのマルコフ過程の推移法則を表し、$\boldsymbol{P} = (p_s(t))_{s,t\in[0,S]}$ を推移法則とする。マルコフ過程の状態 s が経済の状態を表すとすれば、この状態に依存する賃金を表す確率変数を X_s とする $(s \in [0,S])$。ジョブ・サーチとは、期待賃金を最大にする最適政策を求めることである。いま、ある人が仕事を探していて、費用 c を支払って一つの仕事が紹介され、最大で m 個の仕事が出現するまで続けることができる。ここでは、採択しなかった仕事は再び現れることはない、リコールがない場合を考える。現れる仕事の賃金は、マルコフ過程の状態に依存する。状態空間が $\{1,2,\cdots,n\}$ のマルコフ連鎖の場合には、Lippman and MacCall [7] では、(1)X_i は i に関して確率的に減少する。すなわち、任意の x に対して $F_1(x) \geq F_2(x) \geq \cdots \geq F_n(x)$ であり、(2)$\sum_{j=k}^{n} p_{ij}$ は、任意の $k(k=1,2,\cdots,K)$ に対して i に関して増加するときに議論している。

いま、n 個の仕事が残っていて、直面している仕事からの賃金が x のとき、この仕事を採択すれば利得 $u_n(x)$ が得られる。ここで、状態が s で、直面している仕事からの賃金が x のとき、このジョブ・サーチの状態を (s,x) という。また、次の仕事を探すためには費用 c が必要であり、割引率を $0<\beta<1$ とする。ここで、$v_n(s,x)$ を、n 個の仕事が残っていて、ジョブ・サーチの状態が (s,x) のとき、最適に振る舞って得られる β で割り引いた期待利得とすれば、最適性の原理 (Ross [18] など) から、$v_n(s,x)$ はつぎの最適方程式を満足する。

$$v_n(s,x) = \max\left\{u_n(x), -c+\beta\int_0^S p_s(t)dt\int_0^\infty v_{n-1}(t,y)dF_t(y)\right\} \tag{1}$$

ここで、$v_1(s,x) = u_1(x)$ とする。また、$u_n(x)$ は x と n に関する増加関数とする。たとえば、$u_n(x) = \dfrac{1-\delta^n}{1-\delta}x$ は条件を満足するが、これは資産 x を年利率 γ で n 年間預金したときの元利合計となっている $(\delta = 1+\gamma)$。状態を直接知ることができるときは、最適政策は可能な仕事の賃金の集合を 2 つの互いに素な集合に分割することによって定まり、n 個の仕事が残っていて、ジョブ・サーチの状態が (s,x) のとき、これらの集合は reservation wage $\alpha_n(s)$ により定まる。帰納法を用いれば、この $\alpha_n(i)$ が補題 1 を満足し、β で割り引いた期待利得 $v_n(i,x)$ は補題 2 を満たす。

補題 1 任意の状態 $s \in [0, S]$ と正整数 n に対して、$\alpha_{n+1}(s) \geq \alpha_n(s)$ であり、$s < t$ となる任意の状態 $s, t \in [0, S]$ と正整数 n に対して、$\alpha_n(s) \geq \alpha_n(t)$ である。

補題 2 任意の状態 $s \in [0, S]$ と正整数 n に対して、$v_{n+1}(s, x) \geq v_n(s, x)$ かつ $v_{n+1}(s, x) \geq v_{n+1}(t, x)$ である $(x > 0, s < t)$。また、$x > y$ ならば、$v_{n+1}(s, x) \geq v_{n+1}(s, y)$ である。

しかしここでは、状態を直接知ることができない場合を扱うので、推移法則と確率変数 X_s $(s \in [0, S])$ の分布に関して、仮定 1 と 2 のもとで考える。この点が、上記の場合と異なってくる。状態空間が $[0, S]$ で、推移法則が $\boldsymbol{P} = (p_s(t))_{s,t\in[0,S]}$ のとき、議論を簡単にするために確率変数 X_s は絶対連続で、密度関数 $f_s(x)$ を持つとする $(s \in [0, S])$ が、Nakai [13] で考えたようにこの仮定は一般化でき、いろいろな応用が考えられる (Nakai [10, 11, 12] など)。また、定義 1 において、全順序 $\geq$ が定義された完備で可分な距離空間上の確率変数のあいだに、尤度比を用いて確率的な順序関係を導入する。

定義 1 確率変数 X と Y が、それぞれ密度関数 $f(x)$ と $g(x)$ を持ち、$x \geq y$ となる任意の x と y に対して $f(y)g(x) \leq f(x)g(y)$ であれば、X は Y より尤度比の意味で大きいといい、$X \succeq Y$ と表す。

定義 2 関数 $\boldsymbol{P} = (p_s(t))_{s,t\in[0,S]}$ が、$s \leq t$ および $u \leq v$ となる任意の s, t, u と v に対して $(s, t, u, v \in [0, S])$、$p_s(u)\,p_t(v) \geq p_t(u)\,p_s(v)$ のとき、この $\boldsymbol{P}$ を TP_2 という。

定義 1 による順序が半順序となっていることは簡単に示され、この順序もまた TP_2 とよぶ。ここでは、確率変数 $\{X_s\}_{s\in[0,S]}$ は非負の実数値を取るものとし、2 つの仮定 (仮定 1 と 2) を設ける。これらの仮定は、学習過程としてベイズ学習を用いるために必要となるものである。

仮定 1 確率変数 $\{X_s \mid s \in [0, S]\}$ に対して、$s \leq t$ ならば $X_s \succeq X_t$ である $(s, t \in [0, S])$。すなわち、X_s は s に関して尤度比の意味で減少する。

仮定 2 推移法則 $\boldsymbol{P} = (p_s(t))_{s,t\in[0,S]}$ は TP_2 である。

仮定 1 において、$s \leq t$ となる s と t に対して $X_s \succeq X_t$ だから $(s, t \in [0, S])$、$x > y$ ならば $f_s(y)f_t(x) \leq f_s(x)f_t(y)$ であるから、s の値が大きくなるにつれ確率変数 X_s は小さい値を取りやすくなる。すなわち、状態 0 が最も良く、…、状態 S が最も悪いクラスである。仮定 2 より、このマルコフ過程が TP_2 であることを表し、現在の状態から良いクラスの状態へ推移する確率は、現在の状態が良くなるにつれて大きくなる。この仮定から、状態を表す s が大きくなるにつれ、悪い状態へ推移する確率が増加する。いっぽう、$\boldsymbol{p}_s = (p_s(u))$ および $\boldsymbol{p}_t = (p_t(u))$ とおけば、$\boldsymbol{P}$ が仮定 2 を満たすことから、$s \leq t$ となる任意の $s, t \in [0, S]$ に対して $\boldsymbol{p}_t \succeq \boldsymbol{p}_s$ であり、さらにつぎの性質が成り立つ。ここで $F_{\boldsymbol{\mu}}(x) = \int_0^S \mu(s)F_s(x)$ は weighted distribution function(De Vylder[2]) である。

補題 3 $h(x)$ が x の非減少の非負関数とする。$\boldsymbol{\mu} \succeq \boldsymbol{\nu}$ のとき、仮定 1 と 2 のもとで、

$$\int_0^\infty h(x)dF_{\boldsymbol{\mu}}(x) \le \int_0^\infty h(x)dF_{\boldsymbol{\nu}}(x)$$

となる $(\boldsymbol{\mu}, \boldsymbol{\nu} \in \mathcal{S})$。

2.2 状態への推移確率

つぎに、マルコフ過程の状態を知ることができるとき、n 期間後に状態が t となる確率を考える。はじめに、状態の推移のみに着目し、マルコフ過程の状態が s のとき、$\overline{p}_{s,n}(t)$ を n 期間後における状態を表す確率変数の確率密度とする $(s,t \in [0,S], n = 1,2,\cdots)$。このとき、この密度関数 $\overline{p}_{s,n}(t)$ が、初期条件を $\overline{p}_{s,1}(t) = p_s(t)$ とする再帰式 $\overline{p}_{s,n}(t) = \int_0^S p_s(u)\overline{p}_{u,n-1}(t)du$ を満たすことは簡単にわかる。ここで、関数 $\overline{\boldsymbol{P}}_n = \left(\overline{p}_{s,n}(t)\right)_{s,t\in[0,S]}$ に対して、$\overline{\boldsymbol{P}}_1 = \boldsymbol{P}$ であり $\overline{\boldsymbol{P}}_n = \langle \boldsymbol{P}, \overline{\boldsymbol{P}}_{n-1}\rangle$ となっている。いま、2 つの関数 $\boldsymbol{P} = (p_s(t))_{s,t\in[0,S]}$ と $\boldsymbol{Q} = (q_s(t))_{s,t\in[0,S]}$ に対して、

$$\langle \boldsymbol{P}, \boldsymbol{Q}\rangle = \left(\int_0^S p_s(u)q_u(t)du\right)_{s,t\in[0,S]}$$

と定義すれば、つぎの性質が成り立つ。

補題 4 $\boldsymbol{P} = (p_s(t))_{s,t\in[0,S]}$ と $\boldsymbol{Q} = (q_s(t))_{s,t\in[0,S]}$ が TP_2 であれば、$\langle \boldsymbol{P}, \boldsymbol{Q}\rangle$ もまた TP_2 である。

このとき、仮定 2 から $\boldsymbol{P} = (p_s(t))_{s,t\in[0,S]}$ が TP_2 だから、$\overline{\boldsymbol{P}}_{n-1} = (p_{s,n-1}(t))_{s,t\in[0,S]}$ が TP_2 であることが示されるので、n に関する帰納法と補題 4 より、$\overline{\boldsymbol{P}}_n = \langle \boldsymbol{P}, \overline{\boldsymbol{P}}_{n-1}\rangle = \left(\overline{p}_{s,n}(t)\right)_{s,t\in[0,S]}$ が TP_2 となる。

ここで、確率密度 $\overline{p}_{s,n}(t)$ は、マルコフ過程の状態が s のとき、部分観測可能なマルコフ過程にしたがって状態が推移して、n 期後の状態を表す確率変数の確率密度であった。つぎに、このマルコフ過程にしたがって状態が推移するジョブ・サーチにおいて、最適政策にしたがったときに、同様の確率を考える。現在の状態が s で、n 個の仕事が残っているとき、$\overline{p}_{s,n,m}(t)$ を最適政策にしたがったとき、m 期後の状態を表す確率変数の確率密度とする $(s,t \in [0,S]$ および $m \le n, n, m = 1,2,\cdots)$。前節で見たように状態を直接に知ることができるジョブ・サーチにおいては、最適政策は reservation wages $\alpha(s,n)$ によって定まり、$F_s(\alpha(s,n))$ は、状態が s で、n 個の仕事が残っているときに、直面している仕事を採択しない確率である。したがって、$\overline{\boldsymbol{p}}_{s,n,m} = (\overline{p}_{s,n,m}(t))_{t\in[0,S]}$ が、初期条件を $\overline{\boldsymbol{p}}_{s,n,1} = (\overline{p}_{s,n,1}(t))_{t\in[0,S]}$ とする再帰式

$$\overline{p}_{s,n,m}(t) = F_s(\alpha(s,n))\int_0^S p_s(x)\overline{p}_{x,n-1,m-1}(t)dx \tag{2}$$

を満たす。ただし、$\overline{p}_{s,n,1}(t) = F_s(\alpha(s,n))p_s(t)$ である。ここで、

$$\overline{\boldsymbol{P}}_{n,m} = \left(\overline{\boldsymbol{p}}_{s,n,m}\right)_{s\in[0,S]} = \left(\overline{p}_{s,n,m}(t)\right)_{s,t\in[0,S]}$$

とおけば、任意の $n(>0)$ に対して $\overline{\boldsymbol{P}}_{n,1} = (F_s(\alpha(s,n))\boldsymbol{p}_s)_{s\in[0,S]} = (F_s(\alpha(s,n))p_s(t))_{s,t\in[0,S]}$ であることと (2) 式より

$$\overline{\boldsymbol{P}}_{n,m} = (F_s(\alpha(s,n))\langle \boldsymbol{P}, \overline{\boldsymbol{P}}_{n-1,m-1}\rangle_s)_{s\in[0,S]} \tag{3}$$

となる。このとき、補題 4 から、これらの $\overline{\boldsymbol{P}}(n,m)$ は、つぎの性質を満足する。

命題 1 $\overline{\boldsymbol{P}}_{n,m} = \left(\overline{\boldsymbol{p}}_{s,m,n}\right)_{s\in[0,S]} = \left(\overline{p}_{s,m,n}(t)\right)_{s,t\in[0,S]}$ は TP_2 である。

3　不完備情報のジョブ・サーチ

3.1　最適政策と期待利得

前節で考えたマルコフ過程におけるジョブ・サーチで、状態を直接知ることができない場合、すなわち部分観測可能なマルコフ過程でのジョブ・サーチを考える。状態に関する情報は、状態空間 $[0,S]$ 上の確率分布 $\boldsymbol{\mu}$ で表され、$\mathcal{S}$ を情報全体の集合とすれば、

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

となる。これらの確率分布 $\boldsymbol{\mu}$ は密度関数を持つものとする。また、$\mathcal{S}$ に含まれる情報のあいだに、定義 1 によって尤度比を用いた順序関係を導入する。すなわち、状態空間 $[0,S]$ 上の 2 つの確率分布 $\boldsymbol{\mu},\boldsymbol{\nu}$ に対して、$\mu(t)\nu(s) \geq \mu(s)\nu(t)$ が任意の $s,t\,(s \leq t,\ s,t \in [0,S])$ に対して成り立ち、少なくとも 1 つの s と t の組み合わせに対して $\mu(t)\,\nu(s) \geq \mu(s)\,\nu(t)$ のとき、$\boldsymbol{\mu}$ は $\boldsymbol{\nu}$ より尤度比の意味で大きいといい、簡単に $\boldsymbol{\mu} \succ \boldsymbol{\nu}$ と表す。この順序関係は半順序であり、簡単に TP_2 という。定義 1 より、$\boldsymbol{\mu} \succeq \boldsymbol{\nu}$ ならば $(\boldsymbol{\mu},\boldsymbol{\nu} \in \mathcal{S})$、$t$ が大きくなるにしたがって、状態 t における密度の比 $\dfrac{\mu(t)}{\nu(t)}$ は $\nu(t) \neq 0$ の範囲で増加する。この関係は、一般的な部分観測可能なマルコフ過程においても定義でき、詳しくは多段決定問題への応用を含めて Nakai [13, 14] にある。

観測できない状態に関して、その状態に関する情報を得るための情報過程が存在すると考える。ここでは、確率変数 $\{X_s \mid s \in [0,S]\}$ が未知の状態に依存する仕事の賃金を表すから、これらの確率変数を観測することを情報過程と考える。したがって、観測できない状態に関して、この賃金を用いて情報を改良することになる。すなわち、事前情報が $\boldsymbol{\mu}$ で、直面する仕事の賃金が x のとき、学習過程としてベイズの定理を用いることによって、状態についての新しい情報を $\boldsymbol{\mu}(x) = (\mu(x,s))_{s\in[0,S]}$ と改良する。そのあと推移法則 $\boldsymbol{P}$ にしたがって状態が推移し、新しい状態へ移ると考える。このとき、つぎの決定時点における事前情報を $\overline{\boldsymbol{\mu}(x)} = (\overline{\mu(x,s)})_{s\in[0,S]}$ とする。この順序は、逆に考えることも可能であるが、基本的には同様である。つぎに、x と s の関数 $\boldsymbol{h}(x,s)$ に対して、定義 3 により単調性を定義する。

定義 3 $s \in [0,S]$ と $x \in \Re_+$ の非負関数 $\boldsymbol{h}(x) = (h(x,s))_{s\in[0,S]}$ に対して、$x < y$ ならば $s \leq t$ となる任意の t と s について $(s,t \in [0,S])$、$\boldsymbol{h}(x) \succeq \boldsymbol{h}(y)$ (または$\boldsymbol{h}(y) \succeq \boldsymbol{h}(x)$) すなわち

$h(x,t)\,h(y,s) \geq h(x,s)\,h(y,t)$ (または $h(x,t)\,h(y,s) \leq h(x,s)\,h(y,t)$) とする。このとき、関数 $\boldsymbol{h}(x,s)$ を x の減少関数 (または増加関数) という。

確率変数 $\{X_s \mid s \in [0,S]\}$ は密度関数 $\{f_s(x) \mid s \in [0,S]\}$ を持ち、仮定 1 を満たすから、$\boldsymbol{f}(x) = (f_s(x))_{s\in[0,S]}$ とおけば $\boldsymbol{f}(y) \succeq \boldsymbol{f}(x)$ となる。すなわち、$\boldsymbol{f}(x)$ は x に関する増加関数である。

事前情報 $\boldsymbol{\mu}$ と事後情報 $\overline{\boldsymbol{\mu}(x)}$ については、仮定 1 と 2 のもとで、次のような補題 5 が成り立ち、Nakai [13] などで求められている。

補題 5 $\boldsymbol{\mu} \succ \boldsymbol{\nu}$ ならば、任意の x に対して $\boldsymbol{\mu}(x) \succ \boldsymbol{\nu}(x)$ かつ $\overline{\boldsymbol{\mu}(x)} \succ \overline{\boldsymbol{\nu}(x)}$ である。任意の $\boldsymbol{\mu}$ に対して、$\boldsymbol{\mu}(x)$ と $\overline{\boldsymbol{\mu}(x)}$ は x の減少関数である。

補題 5 から、事前情報 $\boldsymbol{\mu}$ のあいだの順序関係 (定義 1) は改良された情報 $\boldsymbol{\mu}(x)$ と事後情報 $\overline{\boldsymbol{\mu}(x)}$ のあいだでも保存される。さらに、同じ事前情報 $\boldsymbol{\mu}$ であっても、仕事から得られる賃金 x が大きくなれば、事後情報 $\overline{\boldsymbol{\mu}(x)}$ は尤度比の意味で悪くなる (定義 1)。

つぎに、マルコフ過程の未知の状態に関する事前情報が $\boldsymbol{\mu}$ であるジョブ・サーチを考える。いま、n 個の仕事が残っていて、直面している仕事の賃金が x のとき、$v_n(\boldsymbol{\mu},x)$ を、最適政策を用いたときの β で割引された総期待利得とする $(0<\beta<1)$。最適性の原理より、この $v_n(\boldsymbol{\mu},x)$ は次の再帰方程式を満足する。

$$v_n(\boldsymbol{\mu},x) = \max\left\{u_n(x), c+\beta\int_0^\infty v_{n-1}(\overline{\boldsymbol{\mu}(x)},y)dF_{\overline{\boldsymbol{\mu}(x)}}(y)\right\} \tag{4}$$

ただし、$v_1(\boldsymbol{\mu},x) = \mathrm{E}_{\boldsymbol{\mu}}[u_1(X)] = \int_0^\infty u_1(x)dF_{\boldsymbol{\mu}}(x)$ とする。また、状態空間が $\mathcal{S}=[0,S]$ のとき、推移確率密度 $\boldsymbol{P}$ に対して $p_S(t) = I_S(t)$ であり、確率 1 で $X=0$ とすれば、状態 S は債務不履行の状態を表すと考えられる。ここで、$I_S(t)$ は t の indicator function である。

ここで

$$S(\boldsymbol{\mu},n) = \left\{x \,\middle|\, u_n(x) \geq c+\beta\int_0^\infty v_{n-1}(\overline{\boldsymbol{\mu}(x)},y)dF_{\overline{\boldsymbol{\mu}(x)}}(y)\right\}$$

および $C(\boldsymbol{\mu},n) = S(\boldsymbol{\mu},n)^c$ とおけば、$S(\boldsymbol{\mu},n)$ と $C(\boldsymbol{\mu},n)$ は、それぞれこのジョブ・サーチにおける停止領域と継続領域を表す。ここで、$u_n(x)$ は x の増加関数であり、$\overline{\boldsymbol{\mu}(x)}$ は x の減少関数であることに注意する。すなわち、$x>y$ ならば $\overline{\boldsymbol{\mu}(y)} \succeq \overline{\boldsymbol{\mu}(x)}$ である。また、被積分関数 $v_{n-1}(\overline{\boldsymbol{\mu}(x)},z)$ が z の増加関数であり、$\boldsymbol{\mu}$ の減少関数のとき、補題 3 より $x>y$ ならば

$$\int_0^\infty v_{n-1}(\overline{\boldsymbol{\mu}(x)},z)dF_{\overline{\boldsymbol{\mu}(x)}}(z) \geq \int_0^\infty v_{n-1}(\overline{\boldsymbol{\mu}(y)},z)dF_{\overline{\boldsymbol{\mu}(y)}}(z)$$

となる。したがって、これら 2 つの領域 $S(\boldsymbol{\mu},n)$ と $C(\boldsymbol{\mu},n)$ に関して、(4) 式から次の性質が得られる。

補題 6 $\boldsymbol{\mu} \succeq \boldsymbol{\nu}$ ならば $(\boldsymbol{\mu},\boldsymbol{\nu} \in \mathcal{S})$、$S(\boldsymbol{\nu},n) \subset S(\boldsymbol{\mu},n)$ および $S(\boldsymbol{\mu},n+1) \subset S(\boldsymbol{\mu},n)$ である。

任意の $\boldsymbol{\mu}$ と $n \geq 1$ に対して、$S(\boldsymbol{\mu},n) \cup C(\boldsymbol{\mu},n) = \Re_+$ であり、$S(\boldsymbol{\mu},n) \cap C(\boldsymbol{\mu},n) = \emptyset$ であるから、この補題から $C(\boldsymbol{\mu},n) \subset C(\boldsymbol{\nu},n)$ および $C(\boldsymbol{\mu},n) \subset C(\boldsymbol{\mu},n+1)$ となる。また、$v_n(\boldsymbol{\mu},x)$ は次の性質を持つ。

補題 7 $\boldsymbol{\mu} \succeq \boldsymbol{\nu}$ ならば $(\boldsymbol{\mu}, \boldsymbol{\nu} \in \mathcal{S})$、$v_n(\boldsymbol{\mu}, x) \leq v_n(\boldsymbol{\nu}, x)$ である。また、$x > y$ ならば、$v_{n+1}(\boldsymbol{\mu}, x) \geq v_n(\boldsymbol{\mu}, x)$ および $v_n(\boldsymbol{\mu}, x) \geq v_n(\boldsymbol{\mu}, y)$ である。

3.2 状態への推移確率–不完備情報の場合

2.2 節と同じように、状態が部分観測可能なマルコフ過程に従って推移する場合に、仮定 1 と 2 のもとで、n 期間後に状態が t となる確率を考える。はじめに、順序立てて考えるために、これらの確率を決定と未知の状態に関する学習過程を除いて考える。未知の状態に関する事前情報が $\boldsymbol{\mu}$ のとき、$\overline{\boldsymbol{P}}_m(\boldsymbol{\mu})$ を m 期間後の状態を表す確率変数の確率密度とする。初期条件として、$m = 1$ のときは $\overline{\boldsymbol{P}}_{\boldsymbol{\mu},1} = (\overline{\boldsymbol{P}}_{\boldsymbol{\mu},1}(t))_{t\in[0,S]}$ であり、$\overline{\boldsymbol{P}}_1(\boldsymbol{\mu})_t = \int_0^S \mu(s)p_s(t)ds = \langle\boldsymbol{\mu}, \boldsymbol{P}\rangle(t)$ となる。前節と同じように、$\boldsymbol{\mu} = (\mu(s))_{s\in[0,S]}$ と $\boldsymbol{P} = (p_s(t))_{s,t\in[0,S]}$ に対して、$\langle\boldsymbol{\mu}, \boldsymbol{P}\rangle$ を $\langle\boldsymbol{\mu}, \boldsymbol{P}\rangle = (\langle\boldsymbol{\mu}, \boldsymbol{P}\rangle(t))_{t\in[0,S]}$ とする。ただし、

$$\langle\boldsymbol{\mu}, \boldsymbol{P}\rangle(t) = \int_0^S \mu(s)p_s(t)ds$$

である。このとき、$\langle\langle\boldsymbol{\mu}, \boldsymbol{P}\rangle, \boldsymbol{Q}\rangle = \langle\boldsymbol{\mu}, \langle\boldsymbol{P}, \boldsymbol{Q}\rangle\rangle$ であることは明らかである。さらに、$\boldsymbol{P} = (p_s(t))_{s,t\in[0,S]}$ に対して、$\boldsymbol{P}^n$ を $\boldsymbol{P}^1 = \boldsymbol{P}$ および $\boldsymbol{P}^n = \langle\boldsymbol{P}, \boldsymbol{P}^{n-1}\rangle$ で定義する。このとき、$\overline{\boldsymbol{\mu}} = \langle\boldsymbol{\mu}, \boldsymbol{P}\rangle$ および $\overline{\mu(x)} = \langle\mu(x), \boldsymbol{P}\rangle$ となる。この関係式から $m = 2$ に対して $\overline{\boldsymbol{P}}_{\boldsymbol{\mu},2} = \overline{\boldsymbol{P}}_{\overline{\boldsymbol{\mu}},1} = \langle\overline{\boldsymbol{\mu}}, \boldsymbol{P}\rangle = \langle\boldsymbol{\mu}, \boldsymbol{P}^2\rangle$ であり、$\overline{\boldsymbol{P}}(\boldsymbol{\mu}, m)$ の再帰関係式はつぎのようになる。

$$\overline{\boldsymbol{P}}_{\boldsymbol{\mu},m} = \overline{\boldsymbol{P}}_{\overline{\boldsymbol{\mu}},m-1} = \overline{\boldsymbol{P}}_{\langle\boldsymbol{\mu},\boldsymbol{P}\rangle,m-1} = \langle\langle\boldsymbol{\mu}, \boldsymbol{P}\rangle, \boldsymbol{P}^{m-1}\rangle = \langle\boldsymbol{\mu}, \boldsymbol{P}^m\rangle \tag{5}$$

ここで、$\boldsymbol{P}$ が TP_2 だから、m に関する帰納法より $\overline{\boldsymbol{P}}_m(\boldsymbol{\mu}) = \langle\boldsymbol{\mu}, \boldsymbol{P}^m\rangle$ もまた TP_2 であることがわかる。よって、つぎの性質が成り立つ。

補題 8 $\boldsymbol{\mu} \succeq \boldsymbol{\nu}$ であり $(\boldsymbol{\mu}, \boldsymbol{\nu} \in \mathcal{S})$、$\boldsymbol{P}$ が TP_2 ならば $(\boldsymbol{\mu}, \boldsymbol{\nu} \in \mathcal{S})$、$\langle\boldsymbol{\mu}, \boldsymbol{P}\rangle \succeq \langle\boldsymbol{\nu}, \boldsymbol{P}\rangle$ である。

命題 2 $\boldsymbol{\mu} \succeq \boldsymbol{\nu}$ ならば $(\boldsymbol{\mu}, \boldsymbol{\nu} \in \mathcal{S})$、$\overline{\boldsymbol{P}}_{\boldsymbol{\mu},m} \succeq \overline{\boldsymbol{P}}_{\boldsymbol{\nu},m}$ である。

つぎに、同様の確率を、決定を除いて考える。すなわち、直面している仕事の賃金 x を用いて、未知の状態に関する学習過程を考慮する。事前情報が $\boldsymbol{\mu}$ のとき、はじめに状態に依存する x の値を知って、ベイズの定理に従って情報を改良する。そのあと、つぎの期へ進み、推移法則 $\boldsymbol{P} = (p_s(t))_{s,t\in[0,S]}$ にしたがって状態の推移が起こる。したがって、状態に関する事前情報が $\boldsymbol{\mu}$ であるというとき、推移は終わっているものと考える。状態に関する事前情報が $\boldsymbol{\mu}$ のとき、$\widehat{P}_{\boldsymbol{\mu},m}(t)$ を m 期間後の状態を表す確率変数の確率密度とし $(t \in [0, S])$、$\widehat{\boldsymbol{P}}_{\boldsymbol{\mu},m} = (\widehat{P}_{\boldsymbol{\mu},m}(t))_{t\in[0,S]}$ とする。

ここで、関数 $\boldsymbol{u}(x) = (u(x,s))_{s\in[0,S]}$ に対して、$\int_a^b u(x,s)dF(x)$ が任意の s に対して存在すれば $(s \in [0, S])$、$\int_a^b \boldsymbol{u}(x)dF(x)$ を簡単のために、

$$\int_a^b \boldsymbol{u}(x)dF(x) = \left(\int_a^b u(x,s)dF(x)\right)_{s\in[0,S]}$$

と表す。

未知の状態に関する事前情報が $\boldsymbol{\mu}$ のとき、$\widehat{\boldsymbol{P}}_{\boldsymbol{\mu},1} = (\widehat{P}_{\boldsymbol{\mu},1}(t))_{t\in[0,S]}$ がつぎの時点での状態を表す確率変数の確率密度だから、

$$\widehat{\boldsymbol{P}}_{\boldsymbol{\mu},1} = \int_0^\infty \langle \boldsymbol{\mu}(x), \boldsymbol{P}\rangle dF_{\boldsymbol{\mu}}(x) = \int_0^\infty \overline{\boldsymbol{\mu}(x)} dF_{\boldsymbol{\mu}}(x)$$

である。ある時点での事前情報が $\boldsymbol{\mu}$ で、直面する仕事の賃金が x のとき、つぎの期での事前情報が $\overline{\boldsymbol{\mu}(x)}$ であった。事前情報が $\boldsymbol{\mu}$ のとき、$\widehat{\boldsymbol{P}}_{\boldsymbol{\mu},m}$ が m 期間後の状態を表す確率変数の確率密度だから、$\widehat{\boldsymbol{P}}_{\boldsymbol{\mu},m}$ は (6) 式を満足する。

$$\widehat{\boldsymbol{P}}_{\boldsymbol{\mu},m} = \int_0^\infty \widehat{\boldsymbol{P}}_{\overline{\boldsymbol{\mu}(x)},m-1} dF_{\boldsymbol{\mu}}(x), \tag{6}$$

ここで、$\widehat{\boldsymbol{P}}_{\boldsymbol{\mu},1} = \int_0^\infty \overline{\boldsymbol{\mu}(x)} dF_{\boldsymbol{\mu}}(x)$ とする。このとき、$\boldsymbol{\mu} \succeq \boldsymbol{\nu}$ ならば $(\boldsymbol{\mu}, \boldsymbol{\nu} \in \mathcal{S})$、補題5より $\overline{\boldsymbol{\mu}} \succeq \overline{\boldsymbol{\nu}}$ および $\overline{\boldsymbol{\mu}(x)} \succeq \overline{\boldsymbol{\nu}(x)}$ となるから、$\widehat{\boldsymbol{P}}_{\boldsymbol{\mu},m}$ はつぎの性質を持つ。

命題 3 $\boldsymbol{\mu} \succeq \boldsymbol{\nu}$ ならば $(\boldsymbol{\mu}, \boldsymbol{\nu} \in \mathcal{S})$、$\widehat{\boldsymbol{P}}_{\boldsymbol{\mu},m}$ は $\boldsymbol{\mu}$ の増加関数である。すなわち、$\widehat{\boldsymbol{P}}_{\boldsymbol{\mu},m} \succeq \widehat{\boldsymbol{P}}_{\boldsymbol{\nu},m}$ である。

最後に、同様の確率を、決定と学習過程を含めて考える。すなわち、事前情報が $\boldsymbol{\mu}$ のとき、直面している仕事の賃金 x を知って、未知の状態について学習を行い、この情報のもとで、この仕事に就くかどうかを決定する。もし、見送れば1期進み、マルコフ過程は推移確率密度 $\boldsymbol{P}$ にしたがって新しい状態へと推移し、状態についての事後情報は $\overline{\boldsymbol{\mu}(x)}$ となる。

いま、n 個の仕事が残っていて、状態に関する事前情報が $\boldsymbol{\mu}$ のとき、直面している仕事の賃金を x とする。$(\widetilde{P}_{\boldsymbol{\mu},n,m}(t))_{t\in[0,S]}$ を最適政策にしたがったときの m 期間後の状態を表す確率変数の確率密度とする $(t \in [0,S], n, m = 1, 2, \cdots, m \le n)$。これらの $(\widetilde{P}_{\boldsymbol{\mu},n,m}(t))_{t\in[0,S]}$ を考えるために、はじめに $m=1$ の場合を考える。直面している仕事の賃金 x を観測し、この仕事を採択しなかったとしよう。このとき、状態について改良した情報を簡単のために $\boldsymbol{\mu}^* = (\mu^*(s))_{s\in[0,S]}$ とおく。$(\widetilde{P}'_{\boldsymbol{\mu}^*,n,1}(t))_{t\in[0,S]}$ を、推移をしたあとでの、つぎの期での状態を表す確率変数の確率密度とする $(s \in [0,S])$。このとき、$x \in C(\boldsymbol{\mu}, n)$ のときにのみ、つぎの期へ進むから、$\widetilde{\boldsymbol{P}}'_{\boldsymbol{\mu}^*,n,1} = (\widetilde{P}'_{\boldsymbol{\mu}^*,n,1}(t))_{t\in[0,S]}$ は $\widetilde{P}'_{\boldsymbol{\mu}^*,n,1}(t) = \int_0^S \mu^*(s) p_s(t)$ と $\widetilde{\boldsymbol{P}}'_{\boldsymbol{\mu}^*,n,1} = \langle \boldsymbol{\mu}^*, \boldsymbol{P}\rangle = \overline{\boldsymbol{\mu}^*}$ を満たす。直面する仕事の賃金 x を観測したとき、改良した情報は $\boldsymbol{\mu}(x)$ だから、

$$\widetilde{P}_{\boldsymbol{\mu},n,1}(t) = \int_{C(\boldsymbol{\mu},n)} \widetilde{P}'_{\boldsymbol{\mu}(x),n,1}(t) dF_{\boldsymbol{\mu}}(x) = \int_{C(\boldsymbol{\mu},n)} dF_{\boldsymbol{\mu}}(x) \int_0^S \mu(x)_s p_s(t) ds$$

となる。しがたって、$\widetilde{\boldsymbol{P}}_{\boldsymbol{\mu},n,1} = \int_{C(\boldsymbol{\mu},n)} \overline{\boldsymbol{\mu}(x)} dF_{\boldsymbol{\mu}}(x)$ となる。

n 個の仕事が残っているとき、状態に関する事前情報が $\boldsymbol{\mu}$ であれば、$(\widetilde{P}_{\boldsymbol{\mu},n,m}(t))_{t\in[0,S]}$ を、最適政策にしたがったときの m 期間後の状態を表す確率変数の確率密度とする $(t \in [0,S], n, m = 1, 2, \cdots, m \le n)$。このとき、新しい仕事が現れ、その賃金は状態に依存する。この値をもとに、この仕事に就くかどうかを決定する。ジョブ・サーチにおいては、$x \in C(\boldsymbol{\mu}, n)$ のとき、

つぎの仕事へと進むから、$\widetilde{\boldsymbol{P}}_{\boldsymbol{\mu},n,m} = \left(\widetilde{P}_{\boldsymbol{\mu},n,m}(t)\right)_{t\in[0,S]}$ が再帰方程式

$$\widetilde{P}_{\boldsymbol{\mu},n,m}(t) = \int_{C(\boldsymbol{\mu},n)} \widetilde{P}_{\overline{\boldsymbol{\mu}(x)},n-1,m-1}(t)dF_{\boldsymbol{\mu}}(x) \tag{7}$$

を満たす。ここで、$\int_{S(\boldsymbol{\mu},n)} dF_{\boldsymbol{\mu}}(x)$ は直面する仕事に就く確率だから、$\int_0^S \widetilde{P}_{\boldsymbol{\mu},n,m}(t)dt \le 1$ であることは明らかである。さらに、$\boldsymbol{\mu} \succeq \boldsymbol{\nu}$ ならば $C(\boldsymbol{\mu},n) \subset C(\boldsymbol{\nu},n)$ となる。すなわち、見送ってつぎの仕事を探す確率は、$\boldsymbol{\mu}$ が増加するにしたがって減少する。いっぽう、より悪い状態へ推移する確率は、$\boldsymbol{\mu}$ が増加すれば増加する。マルコフ過程の状態を直接知ることができる場合には、性質 1 より、$\overline{\boldsymbol{P}}_{n,m} = (\overline{\boldsymbol{P}}_{s,n,m})_{s\in[0,S]}$ は TP_2 であった。しかし、この確率 $(\widetilde{P}_{\overline{\boldsymbol{\mu}(x)},n-1,m-1}(t))_{t\in[0,S]}$ が観測した x によって変化するので、この場合には $\widetilde{\boldsymbol{P}}_{\boldsymbol{\mu},n,m}$ がその様な性質を持つことを示すことは難しい。

4 Sequential Investment Problem

消防や警察などの公的な部門に決められた予算の範囲で資源を投入することを考える。いっぽう、このような公的部門に対しては、これらのサービスに対する満足度と実際の設備や人員の数とのあいだの関係は、ある種の関数で表されている関係があると考える。いま、このサービスに対する満足度は $[0,1]$ 区間に含まれる値 s で表せるものとし、$s=1$ であれば、これらのサービスは要求に応えられていると考え、s の値が小さくなればなるほど、要求に応えられていないとする。一方、要求に応えるために新たな設備や人員の増加をしたとしても、おかれている状況が変化すれば要求がさらに大きくなり満足度が減少することもあり、反対に状況が変化することにより満足度が増すことも考えられる。このような問題をモデル化するために、満足度を表す値を状態と考え、この状態は新たな設備や人員の増加することによっても変化するが、環境が変化するなど制御できない要素によっても変わるものとする。ここでは、設備や人員などの資本の価値を減価償却などを含めて単純に累積したものではなく、その時点における満足度を達成するために必要な投資総額として考えることにする。これは、減価償却のみでなく要求度に応じてこれらの価値が変化することに対応するものと考えることになる。したがって、このようなモデルの状態空間を $[0,1]$ とし、それぞれの状態 $s \in [0,1]$ はこれらのサービスに対する満足度を表すと考える。

一方、これらのサービスを実現するために資金 x を使って設備や人員を一時に配置すれば、これらのサービスに対する満足度が $s(x)$ で表される関数で表現されているとする。すなわち、満足度が s であるときには、$s = s(x)$ であれば、資本を x 投下したことと等価であると考えることができる。

ここで、満足度を表す関数 $s(x)$ はつぎの条件を満たすものとする。

1. $s(x)$ は x に関する増加関数である
2. $s(x)$ は x に関して凹関数 (concave function) である

このとき、

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

とすれば、この $x(s)$ は満足度が s とするために必要な資本の量を表し、満足度の資金への換算値と言え、資本 $x(s)$ を投下したことに等しいと考える。また、$c(s,t)$ を満足度が s のときに、満足度を t にするための費用とする $(t \geq s)$。

いま、残りの計画期間が n のとき、予算 K の範囲内で資本を投下して設備や人員を増やすことによって満足度を増加させることを考える。このとき、最適に振る舞ったときに得られる効用を $v_n(s)$ とすれば、最適方程式は最適性の原理より

$$v_n(s) = \max_{0 \leq x \leq K} \{-c(s, s + d_s(x)) + v_{n-1}(s + d_s(x))\} \tag{8}$$

と表される。また、初期状態 $(n = 1)$ では、

$$v_1(s) = \max_{x \geq 0} \{-c(s, s + d_s(x)) + u(s + d_s(x))\}$$

である。ここで、$d_s(x)$ は、

$$d_s(x) = s(x + s(x)) - s(x)$$

とする。すなわち、満足度が s のとき、資本 x を追加して投下したときの満足度の増加分である。また、費用 $-c(s, s + d_s(x))$ は、満足度が s のときに、資本 x を追加して投下することにより満足度を上昇させ、その結果に費用が依存すると考えたときのものである。もし、$-c(s, s + d_s(x)) = x$ であれば、投入した資本の大きさを費用と考えることになる。このとき、これらの費用関数 $c(s,t)$ に対してつぎの性質が成り立つ $(s \leq t)$。

補題 9 $s < t$ のとき、$d_s(x) \geq d_t(x)$ である。

補題 10 $s < t$ のとき、任意の $x \geq 0$ に対して、$s + d_s(x) \leq t + d_t(x)$ である。

いっぽう、この最適方程式を追加して投下する資本ではなく、目標とする満足度で表せば $(n \geq 1)$、

$$v_n(s) = \max_{s \leq t \leq 1} \{-c(s,t) + v_{n-1}(t)\}$$

であり、初期状態が

$$v_1(s) = \max_{s \leq t \leq 1} \{-c(s,t) + v_0(t)\}$$

となる。ここで、満足度を表す関数として、つぎのような簡単な例を考える。

例 1 例えば、満足度を表す関数を

$$s(x) = 1 - e^{-x} \quad (0 \leq x)$$

とする。このとき、満足度が s のときの、資産の価値 $x(s)$ は

$$x = x(s) = -\log(1 - s) \quad (0 \leq s \leq 1) \tag{9}$$

となり、これらの関数に対して

$$\begin{aligned}\frac{ds(x)}{dx} &= e^{-x}\\ \frac{dx(s)}{ds} &= -\frac{1}{(1-s)}\end{aligned}$$

である。また、満足度を s から t へ変化させるとすれば、追加して投資しなければならない投資額は $-\log\dfrac{1-t}{1-s}$ と求められるから、

$$c(s,t) \equiv y - x = -\log\frac{1-t}{1-s}$$

と考えることにする。したがって、累積投資額が $x \to y$ のとき、すなわち x から y へと増加するときには、満足度はそれぞれ s から t へと変化し、その変化量は

$$s - t = e^{-y} - e^{-x}$$

である。

いっぽう、満足度が s のときには、この時点で投下した資本の総量を $x(s) = -\log(1-s)$ と考えて良いから、新たに資本 x を追加したときには、投下資本の総量が $-\log(1-s)+x$ となり、満足度は $s(-\log(1-s)+x) = 1-(1-s)e^{-x}$ となる。したがって、この場合には満足度は s から

$$s + d_s(x) = 1 - e^{-x}(1-s) = e^{-x}s + 1 - e^{-x} \geq 0$$

へと変化する。すなわち、

$$d_s(x) = e^{-x}s + 1 - e^{-x} - s = (1-e^{-x})(1-s) \geq 0$$

が満足度が s のときに、新たに資本 x を追加して投入したことによる満足度の増加分となる。

ここで、費用関数に関してつぎのことを仮定する。すなわち、$c(s,t)$ は t に関する増加関数であり、凸 (convex) 関数とし、s に関する減少関数であって、$t \to s$ のとき $c(s,t) \to 0$ とする。また、初期値 $v_0(t) = u(t)$ は、満足度が t であるときの全体の効用を表し、この関数は t に関する増加関数であり、凹 (concave) 関数とする。

このとき、上記の例にある (9) で表される関数で満足度が表されたときには、最適方程式は

$$v_n(s) = \max_{x\geq 0}\{-c(s, e^{-x}s + 1 - e^{-x}) + v_{n-1}(e^{-x}s + 1 - e^{-x})\} \tag{10}$$

であり、初期状態 $(n=1)$ では、

$$v_1(s) = \max_{x\geq 0}\{-c(s, e^{-x}s + 1 - e^{-x}) + u(e^{-x}s + 1 - e^{-x})\}$$

となる。

このとき、(8) 式によって表される最適方程式に関してつぎの性質が成り立つ。

補題 11 $v_n(s)$ 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 12 $v_n(s)$ は n に関する非減少関数である。すなわち、$n \geq 1$ のとき $v_n(s) \leq v_{n+1}(s)$ である。

例 2（続き）$v_n(s)$ を残りの計画期間が n のとき、（予算 K の範囲内で）資本を投下して、最適に振る舞ったときに得られる効用とすれば、次の最適方程式を満足する。

$$\begin{aligned} v_n(s) &= \max_{x\geq 0}\{-c(s, s+d_s(x)) + v_{n-1}(s+d_s(x))\} \quad (11) \\ &= \max_{x\geq 0}\{-c(s, e^{-x}s+1-e^{-x}) + v_{n-1}(e^{-x}s+1-e^{-x})\} \quad (12) \end{aligned}$$

ここで、$-c(s, s+d_s(x))$ は、満足度が s のときに、資本 x を追加して投下するときにかかる費用である。初期状態 $(n=1)$ では、

$$v_1(s) = \max_{x\geq 0}\{-c(s, e^{-x}s+1-e^{-x}) + u(e^{-x}s+1-e^{-x})\}$$

であるから、

$$\begin{aligned} &\frac{d}{dx}(-c(s, e^{-x}s+1-e^{-x}) + u(e^{-x}s+1-e^{-x})) \\ &= e^{-x}(1-s)(-\frac{d}{dt}c(s, e^{-x}s+1-e^{-x}) + u'(e^{-x}s+1-e^{-x})) \end{aligned}$$

を見ればよい。

いま、

$$\frac{d}{dt}(-c(s,t)+v_0(t)) = -\frac{d}{dt}c(s,t) + v_0'(t)$$

であり、

$$c(s,t) \equiv y - x = -\log\frac{1-t}{1-s}$$

としたから

$$\begin{aligned} \frac{d}{dt}(-c(s,t)) &= -\frac{1}{1-t} < 0 \\ \frac{d}{ds}(-c(s,t)) &= \frac{1}{1-s} > 0 \end{aligned}$$

となっているから、$-\frac{d}{dt}c(s, e^{-x}s+1-e^{-x}) < 0$ であり、$n=1$ のときの効用関数は単調増加だから、$\frac{d}{dx}u(e^{-x}s+1-e^{-x}) > 0$ であることから、最適政策を求めることができる。

4.1 費用関数 $c(x)$ の場合

つぎに、費用関数 (cost function) を x の関数である $c(x)$ の場合に限定して議論を進める。このときには、最適方程式は

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

と表すことができる。ただし、$c(x)$ は資本 x を投下する費用とする。さらに、費用関数 $c(x)$ は x の増加かつ凹関数であるとし、それに加え次の仮定を設ける。

仮定 3 $d_s(x)$ は s に関する凹 (*concave*) 関数である。

さきの例で見れば

$$\begin{aligned} s(x(s+u)+x)-s(x(s)+x) &= 1-(1-s-u)e^{-x}-(1-(1-s)e^{-x}) \\ &\geq 1-(1-t-u)e^{-x}-(1-(1-t)e^{-x}) \\ &= s(x(t+u)+x)-s(x(t)+x) \end{aligned}$$

すなわち

$$s(x(s+u)+x)-s(x(s)+x)=ue^{-x}=s(x(t+u)+x)-s(x(t)+x)$$

となっていることから、この仮定を満足することがわかる。このとき、つぎの性質が成り立つ。

補題 13 $v_n(s)$ は s に関する凹 (*concave*) 関数である。

補題 14 残りの計画期間が n で、満足度が s のときの、最適な投資額を $x_n^*(s)$ とすれば、$s \leq t$ のとき $x_n^*(s) \leq x_n^*(t)$ である。

注 1 ここで仮定したように費用関数が $c(x)$ ではなく、一般的な関数 $c(t,x)$ のときには、$s<t$ に対して $c(t,x)-c(s,x)$ が x に関する減少関数であれば、この補題 14 は同じように成立する。すなわち、$0 \leq x \leq x^*$ 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 *14* が成り立つ。

補題 15 残りの計画期間が n で、満足度が s のときの、最適な投資額を $x_n^*(s)$ とすれば、$x_{n-1}^*(s) \geq x_n^*(s)$ である。

補題 16 $s<t$ 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。

5 Sequential Investment Problem: Stochastic Case

つぎに、サービスを提供するために設備や人員を揃えるために資本を追加して投資をして満足度を上げたとしても、まわりの状況が変化することによって、必ずしも思ったようにサービスを受ける人にとっての満足度は変化することはなく、場合によっては満足度が減少するという状況も考えることにする。すなわち、満足度が状態を表すと考えたとき、これらの状態が確率過程 (マルコフ連鎖またはマルコフ過程) にしたがって推移する場合を考える。これまでと同じように、状態空間を $[0,1]$ とし、$(p_s(t))_{0\leq s\leq 1}$ を推移法則とするマルコフ過程を考える。このとき、この推移法則につぎの仮定を設ける。

仮定 4 $(p_s(t))_{0\le s\le 1}$ は TP_2 である。

つぎに、残りの計画期間が n のとき、予算 K の範囲内で資本を投下して、最適に振る舞ったときに得られる効用の期待値を $V_n(s)$ とすれば、最適方程式は

$$V_n(s)=\max_{x\ge 0}\{-c(s,s+x(s))+\int_0^1 p_{s+x(s)}(t)V_{n-1}(t)dt\} \tag{13}$$

となり、初期条件は

$$V_1(s)=\max_{x\ge 0}\{-c(s,s+x(s))+\int_0^1 p_{s+x(s)}(t)u(t)dt\}$$

である。ここで、推移法則が TP_2 であるときにつぎの性質が成り立つことが知られている。

補題 17 (*Kijima and Ohnishi[6]*) 仮定 4 のもとで、$u(t)$ 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 に関して増加かつ凹関数である。

補題 18 $u(t)$ が t に関する増加関数であれば、

$$\int_0^1 p_s(t)u(t)dt$$

は s に関して増加関数である。

したがって、補題 18 より、$V_{n-1}(t)$ が t に関する増加関数でることが示されれば、帰納法により $\int_0^1 p_s(t)V_{n-1}(t)dt$ もまた s に関して増加関数であることがわかる。このとき、つぎ述べる性質が帰納法を用いて成り立つことが示される。

補題 19 $V_n(s)$ は s に関する非減少関数である。

補題 20 $V_n(s)$ は n に関する非減少関数である。

補題 21 $V_n(s)$ は s に関する凹 (*concave*) 関数である。

補題 22 残りの計画期間が n で、満足度が s のときの、最適な投資額を $x_n^*(s)$ とすれば、$x\le y$ のとき $x_n^*(s)\le x_n^*(t)$ である。

補題 23 残りの計画期間が n で、満足度が s のときの、最適な投資額を $x_n^*(s)$ とすれば、$x_{n-1}^*(s)\ge x_n^*(s)$ である $(n\ge 1)$。

参考文献

[1] Albright, S. C. (1974). A Markov-Decision-Chain Approach to a Stochastic Assignment Problem, *Operations Research*, **22**, 61–64.

[2] F. De Vylder, (1983). Duality Theorem for Bounds in Integrals with Applications to Stop Loss Premiums, *Scandinavian Actuarial Journal*, 129–147.

[3] S. Karlin, (1968). *Total Positivity*, Stanford University Press, Stanford, California.

[4] S. Karlin and J. L. McGregor, (1960). Classical Diffusion Process and Total Positivity, *Journal of Mathematical Analysis and Applications*, **1**, 163–183.

[5] S. Karlin and Y. Rinott, (1981). 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.

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

[7] S. A. Lippman and J. J. McCall, (1976). Job Search in a Dynamic Economy, *Journal of Economic Theory*, **12**, 365–390.

[8] G. Monahan, (1980). Optimal Stopping in a Partially Observable Markov Processes with Costly Information, *Operations Research*, **28**, 1319–1334.

[9] T. Nakai, (1985). The Problem of Optimal Stopping in a Partially Observable Markov process, *Journal of Optimization Theory and Applications*, **45**, 425–442.

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

[11] T. Nakai, (1996). An Optimal Selection Problem on a Partially Observable Markov process, *Stochastic Modelling in Innovative Manufacturing*, Lecture Notes in Economics and Mathematical Systems **445**, Eds. A. H. Christer, S. Osaki and L. C. Thomas, 140–154, Springer-Verlag, Berlin.

[12] T. Nakai, (1999). An Optimal Assignment Problem for Multiple Objects per Period – Case of a Partially Observable Markov process, *Bulletin of Informatics and Cybernetics*, **31**, 23–34.

[13] T. Nakai, (2002). A Generalization of Multivariate Total Positivity of Order Two with an Application to Bayesian Learning Procedure, *Journal of Information & Optimization Sciences*, **23**, 163–176.

[14] T. Nakai, (2002). A Generalized MTP_2 and a Sequential Stochastic Model on a Partially Observable Markov Process, *Probabilistic Methods in Discrete Mathematics–Proceedings of the Fifth International Petrozavodsk Conference*, (Eds. V.F. Kolchin, V.Ya. Kozlov,

V.V. Mazalov, Yu.L. Pavlov and Yu.V. Prokhorov), VSP publishes, The Netherlands, 291–302.

[15] T. Nakai, (2002). Properties of a Partially Observable Markov process for a Job Search Problem in a Dynamic Economy, *Proceedings of the 2nd Euro Japanese Workshop on Stochastic Modelling for Finance, Insurance, Production and Reliability*, (Eds. T. Dohi, N. Limonios and S. Osaki), Systems Reliability Engineering Laboratory, Hiroshima University, 340–349.

[16] T. Nakai, Properties of Total Positivity and an Application to Job Search under Uncertainty, *Scientiae Mathematicae Japonicae*, 2004 (to appear).

[17] M. Ohnishi, H. Kawai and H. Mine, (1986). An Optimal Inspection and Replacement Policy under Incomplete State Information,*European Journal of Operations Research*, **27**, 117–128.

[18] S. M. Ross, (1970). *Applied Probability Models with Optimization Applications*, Holden-Day, San Francisco, California.