非有界コスト関数および非コンパクトな状態空間と決定空間におけるマルコフ決定過程について

Markov decision processes with unbounded cost functions and non-compact state space and action spaces

* 源 隆哉 (Ryuya Minamoto), * 鈴木陸斗 (Rikuto Suzuki),
* 王瀚東 (Wang Handong), *,† 堀口正之 (Masayuki Horiguchi)

1 はじめに

本研究では, 状態空間, 決定空間が非コンパクト集合のマルコフ決定過程についてと, マルコフ決定過程における推移確率の存在について扱う. 非コンパクト集合上で定義された連続関数の最大値や最小値の存在が一般には保証されない. そこで下半連続性や inf-compact 性とよばれる概念を定義し, それらを用いてマルコフ決定過程を定義する. マルコフ決定過程を定義した際に問題になってくるのはその確率過程は存在するのかである. 存在を保証する拡張定理を述べる. §2 では, 本研究で扱うマルコフ決定過程を定義を行い, §3 では定義した確率過程が存在することを証明するのによく用いられる Kolmogorov の拡張定理とその応用例を述べて, §4 では, マルコフ決定過程の存在を証明するのに用いられる Ionescu-Tulcea の拡張定理について述べ, §5 で今後の課題とまとめについて述べる.

2 マルコフ決定過程について

非コンパクト集合上のマルコフ決定過程について考える. 資産, 在庫, 人口など, どこまでも増加しうるものに対して最適政策などを考える場合, 状態空間に対してコンパクト性を仮定したマルコフ決定過程では考えることができない. 改めて状態空間が非コンパクト集合上で

* 神奈川大学大学院・理学研究科数学領域 (Field of Mathematics, Graduate School of Science, Kanagawa University),† 神奈川大学理学部理学科数学分野 (Department of Mathematics, Faculty of Science, Kanagawa University)

‡ 221-8686 横浜市神奈川区六角橋 3-27-1, (3-27-1, Rokkakubashi Kanagawa-ku, Yokohama City, Kanagawa Prefecture, 221-8686 Japan)

のマルコフ決定過程を考える必要がある. しかし, 非コンパクト集合上だと連続関数は必ずしも最大値を取るとは言えない. そこで, 論文 [1] を元にして, 下半連続性や inf-comapact 性と呼ばれる概念を定義し, 非コンパクト上でも常に最小値が存在するような状況設定を行う.

定義 2.1. マルコフ決定過程の設定

マルコフ決定過程 (X, A, Q, c) を設定する. ただし X, A, Q, c はそれぞれ, 状態空間, 決定空間, 推移確率, コスト関数を意味する.

(i) $(X, \mathscr{B}(X)), (A, \mathscr{B}(A))$ はそれぞれ Borel 空間とする.

(ii) 任意の状態 x に対して取りうることができる決定の集合 $A(x) \in \mathscr{B}(A)$ が存在する. つまり $A(x)$ の要素は状態 x で取ることが出来る決定全体.

(iii) 状態と決定のペア集合

$$\mathbb{K} := \{(x, a) | x \in X, a \in A(x)\}$$

は $(X \times A, \mathscr{B}(X \times A))$ 上の Borel 集合, つまり $\mathbb{K} \in \mathscr{B}(X \times A)$ である.

(iv) 推移確率 $Q(B|x, a)$ は $B \in \mathscr{B}(X)$ と $(x, a) \in \mathbb{K}$ に基づく X 上の確率核である.

(v) コスト関数 $c(x, a)$ は $\mathbb{K}$ 上の可測関数である. さらに下に有界とする. もし負の値を取る可能性があるならば十分大きな定数 $M > 0$ をとって $c'(x, a) := c(x, a) + M$ とすれば良い.

定義 2.2. 下半連続関数

関数 $f : X \to \mathbb{R}$ が X 上で**下半連続関数**であるとは任意の $r \in \mathbb{R}$ に対して以下の集合が X 上の閉集合になること.

$$\{x \in X | f(x) \leq r\}$$

X 上の下半連続関数全体を $L(X)$ と表し, その中でも非負な関数全体を $L_+(X)$ と表す.

例 2.1. 下半連続関数の例

関数

$$f(x) = \begin{cases} x^2 + 1, & x \neq 0, \\ 0, & x = 0 \end{cases}$$

は $\mathbb{R}$ 上で下半連続であるが連続ではない.

証明

任意の $r \in \mathbb{R}$ に対して集合

$$F_r := \{\, x \in \mathbb{R} \mid f(x) \leq r \,\}$$

が閉集合であることを示せば良い.

(i) $r < 0$ のとき, $\forall x \in \mathbb{R}$ に対して $f(x) \geq 0$ 従って $f(x) \leq r$ となる点は存在しない.

$$F_r = \varnothing$$

これは閉集合である．

(ii) $0 \leq r < 1$ のとき，$f(x) \leq r$ となるのは $x=0$ の場合のみ．したがって

$$F_r = \{0\}$$

これは閉集合である．

(iii) $r \geq 1$ のとき

$$F_r = [-\sqrt{r}, \sqrt{r}]$$

これは閉集合である．

以上より，任意の $r \in \mathbb{R}$ に対して F_r が閉集合になるので，f は下半連続である．ただし $x=0$ において左極限と右極限が一致しないため，f は連続ではない． □

定義 2.3. inf-compact 関数

関数 $v : \mathbb{K} \to \mathbb{R}$ が $\mathbb{K}$ 上 **inf-compact** とは任意の $r \in \mathbb{R}$ と任意の $x \in X$ に対して以下の集合がコンパクト集合になること.

$$\{a \in A(x) | v(x,a) \leq r\}$$

例 2.2. inf-compact の例

$X = A = \mathbb{R}$, 各 $x \in X$ に対して $A(x) = \mathbb{R}$ とする．このとき

$$v(x,a) = \begin{cases} x^2 + a^2 + 1 & x \neq 0, \\ x^2 + a^2 & x = 0 \end{cases}$$

は $\mathbb{K}$ 上で inf-compact である.

証明

任意の $x \in X$ と $r \in \mathbb{R}$ に対して

$$C_{x,r} := \{\, a \in \mathbb{R} \mid v(x,a) \leq r \,\}$$

がコンパクトであることを示せば良い.

(i) $x = 0$ のとき

$$C_{x,r} = \{\, a \in \mathbb{R} \mid a^2 \leq r\} = [-\sqrt{r}, \sqrt{r}]$$

となるが,$r < 0$ のときは, $C_{x,r} = \varnothing$ となり, $r \geq 0$ のときはハイネ・ボレルの定理より, $C_{x,r}$ は $\mathbb{R}$ 上のコンパクト集合である.

(ii) $x \neq 0$ のとき

$$C_{x,r} = \{a \in \mathbb{R} \mid v(x,a) \leq r\} = [-\sqrt{r - x^2 - 1}, \sqrt{r - x^2 - 1}]$$

上記と同様に $r - x^2 - 1 < 0$ のときは, $C_{x,r} = \varnothing$ となり, $r - x^2 - 1 \geq 0$ のときは, ハイネ・ボレルの定理より, $C_{x,r}$ は $\mathbb{R}$ 上のコンパクト集合である.

よって任意の $x \in X$ と $r \in \mathbb{R}$ に対して, $C_{x,r}$ がコンパクト集合なので v は $\mathbb{K}$ 上で inf-compact である. □

次の仮定をすると良い性質がいくつか導出される.

仮定 2.1.

(i)コスト関数 $c(x,a)$ は $\mathbb{K}$ 上, 下半連続, inf-compact である.

(ii)推移確率 Q は $\mathbb{K}$ 上弱連続である.

(iii)集合値関数 $x \to A(x)$ は下半連続である.

つまり, $x_n \to x, a \in A(x) \Rightarrow a_n \in A(x_n)$ s.t. $a_n \to a$

仮定 2.1 より, 次の定理が証明される. これらの定理はマルコフ決定過程が値反復法で解けることを証明する際に用いられる.

定理 2.1. ([1], Lemma 2.7, P.194)

(i) $v(x,a)$ が inf-compact かつ下半連続かつ下に有界とする. このとき関数

$$v^*(x) := \inf_{a \in A(x)} v(x,a)$$

は $L(X)$ に属し, さらにある関数 $f \in \mathbb{F}$ が存在して

$$v^*(x) = v(x, f(x)) \ \ \forall x \in X$$

が成り立つ.

(ii) $u \in L(X)$ が非負であるとする. このとき

$$v^*(x) := \inf_{a \in A(x)} \left[c(x,a) + \int_X u(y) Q(dy|x,a) \right]$$

は $L(X)$ に属し, ある関数 $f \in \mathbb{F}$ が存在して

$$u^*(x) = c(x, f(x)) + \int u(y) Q(dy|x, f(x)) \ \ \forall x \in X$$

が成り立つ.

(iii) 各 $n = 0, 1, ...$ に対して v_n が $\mathbb{K}$ 上で下半連続かつ下に有界かつ inf-compact であるとする. もし v_n が単調増加しながら v_∞ に収束するならば

$$\lim_{n \to \infty} \inf_{a \in A(x)} v_n(x,a) = \inf_{a \in A(x)} v_\infty(x,a) \ \forall x \in X$$

が成り立つ.

マルコフ決定過程を構成する際に問題になってくるのは考えてるマルコフ決定過程に適切な推移確率測度が存在するのかである. その問題について述べていく.

3 Kolmogorov の拡張定理とその応用例 (マルコフ連鎖の構成)

前節の内容とともに, マルコフ決定過程として確率過程を定義する際に問題になってくるのはその"無限次元における確率過程は存在するのか"である. その存在性を保証する定理を 2 つ取り上げる.

定理 3.1. Kolmogorov の拡張定理 ([5], 定理 7.1, P.193)
$\forall n \in \mathbb{N}$ に対して $(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))$ 上の確率測度 μ_n が以下を満たすとする:

$$\text{任意の } n = 1, 2, ... \text{ と任意の } A \in \mathscr{B}(\mathbb{R}^n) \text{ に対して，} \mu_{n+1}(A \times \mathbb{R}) = \mu_n(A)$$

確率測度が上記の条件を満たすとき, μ_n は**整合的**という.
このとき, $(\mathbb{R}^\infty, \mathscr{B}(\mathbb{R}^\infty))$ 上の確率測度 μ で

$$\mu(A \times \mathbb{R}^\infty) = \mu_n(A), \ \ A \in \mathscr{B}(\mathbb{R}^n)$$

を満たすもの一意に存在する. ただし,

$$A \times \mathbb{R}^\infty = \{(\omega_1, \omega_2, ..., \omega_n, ...) \in \mathbb{R}^\infty \mid (\omega_1, ..., \omega_n) \in A\} \subset \mathbb{R}^\infty$$

Kolomogorov の拡張定理の応用例としてマルコフ連鎖を構成する.

定義 3.1. マルコフ連鎖 ([5], 定義 7.6, P.196)
I : 高々可算集合 , $\mathbb{P}$: I 上の推移確率行列, ただし, $\mathbb{P} = (p_{ij})_{i,j \in I}$ と表記する.
確率空間 $(\Omega, \mathscr{F}, P)$ 上で定義された $I-$ 値確率過程 $\{X_n\}$ が**初期分布** μ と**推移確率** $\mathbb{P}$ を持つ**マルコフ連鎖**であるとは:

$$\forall n \in \mathbb{N} \text{ と } \forall x_0, x_1, ..., x_n \in I \text{ に対して } P(X_0 = x_0, X_1 = x_1, ..., X_n = x_n) = \mu_{x_0} p_{x_0 x_1} \cdots p_{x_{n-1} x_n}$$

が成り立つこと.

マルコフ連鎖に関して次の定理が成り立つ。

定理 3.2. マルコフ連鎖の存在定理 ([5], 命題 7.7, P.196)
適当な確率空間上に初期分布 μ と推移確率行列 $\mathbb{P}$ をもつ高々可算集合 I 上のマルコフ連鎖 $\{X_n\}$ が存在する.

証明
I は高々可算集合なので適当に番号をつけることで $I \subset \mathbb{N}$, 従って, $I \subset \mathbb{R}$ とみることができる. $n = 0, 1, 2, ...$ に対し $(\mathbb{R}^{n+1}, \mathscr{B}(\mathbb{R}^{n+1}))$ 上の確率測度 P_{n+1} を

$$P_{n+1}(A) := \sum_{(x_0, x_1, ..., x_n) \in I^{n+1} \cap A} \mu_{x_0} p_{x_0 x_1} \cdots p_{x_{n-1} x_n}, \ A \in \mathscr{B}(\mathbb{R}^{n+1})$$

と定義する. このとき,

$$P_{n+2}(A \times \mathbb{R}) = \left(\sum_{(x_0,x_1,...,x_n) \in I^{n+1} \cap A} \mu_{x_0} p_{x_0 x_1} \cdots p_{x_{n-1} x_n} \right) \sum_{x_{n+1} \in I} p_{x_n x_{n+1}} = P_{n+1}(A)$$

が成り立つ. よって $\{P_{n+1}\}$ は整合的, 従って Kolmogorov の拡張定理より $(\mathbb{R}^\infty, \mathscr{B}(\mathbb{R}^\infty))$ 上に確率測度 P が存在し, $P(A \times \mathbb{R}^\infty) = P_n(A),\ A \in \mathscr{B}(\mathbb{R}^n)$ が成り立つ. $(\mathbb{R}^\infty, \mathscr{B}(\mathbb{R}^\infty), P)$ 上の実数値確率変数列 $\{X_n\}$ を $\omega = (\omega_0, \omega_1, ...) \in \mathbb{R}^\infty$ に対して, $X_n(\omega) := \omega_n$ と定義すれば

$$\begin{aligned} P(X_0 = x_0, X_1 = x_1, ..., X_n = x_n) &= P(\omega_0 = x_0, \omega_1, ..., \omega_n = x_n) \\ &= P_{n+1}(\{x_0, ..., x_n\}) \\ &= \mu_{x_0} p_{x_0 x_1} \cdots p_{x_{n-1} x_n} \end{aligned}$$

なので, $\{X_n\}$ は初期分布 μ と推移確率 $\mathbb{P}$ を持つマルコフ連鎖が得られる. □

4 Ionescu-Tulcea の拡張定理とその証明

Kolmogorov の拡張定理は仮定として, 各時刻の有限次元分布が与えられている. 次に述べる拡張定理は初期分布と各時刻における推移確率が与えられている. マルコフ決定過程に適切な推移確率が存在することを保証するのに用いられる定理である.

定理 4.1. Ionescu-Tulcea の拡張定理 ([3], Proposition V.1.1,P.74)
$\{\Omega_t, \mathscr{F}_t\}$: 可測空間列, $P_{t+1}^{0,...,t} : \left(\prod_{s=0}^{t} \Omega_s, \bigotimes_{s=0}^{t} \mathscr{F}_s \right) \to (\Omega_{t+1}, \mathscr{F}_{t+1})$ を推移確率測度とする.
このとき,$\forall x_0 \in \Omega_0$ に対して, $\left(\prod_{s=0}^{\infty} \Omega_s, \bigotimes_{s=0}^{\infty} \mathscr{F}_s \right)$ 上に一意に確率測度 P_{x_0} が存在し,

$$A := \prod_{s=0}^{T} A_s \times \prod_{s=T+1}^{\infty} \Omega_s \in \bigotimes_{s=0}^{\infty} \mathscr{F}_s,\ \forall T \in \mathbb{N} \cup \{0\}$$

に対して

$$P_{x_0}(A) = 1_{A_0}(x_0) \int_{A_1} P_1^0(x_0 : dx_1) \int_{A_2} P_2^{0,1}(x_0, x_1 : dx_2) \cdots \int_{A_T} P_T^{0,...,T-1}(x_0, x_1, ..., x_{T-1} : dx_T)$$

が成り立つ.

証明の中心となる定理を述べる.

定理 4.2. 確率測度の拡張定理 ([4], 定理 2.4, P.54)
$\mathscr{A}$ を Ω 上の有限加法族とし, p を $(\Omega, \mathscr{A})$ 上の有限加法的確率測度とする. このとき p が $\sigma[\mathscr{A}]$ 上の確率測度に拡張できる必要十分条件は

$$B_1 \supset B_2 \supset \cdots,\ B_n \in \mathscr{A},\ \bigcap_{n=1}^{\infty} B_n = \phi \Rightarrow \lim_{n \to \infty} p(B_n) = 0$$

上記の条件を**共通点性**と呼ぶ.

定理 4.1 の証明に用いるその他の定理を述べておく.

定理 4.3.
$(X,\mathscr{F}),(Y,\mathscr{G})$: 可測空間, $f:X\times Y\to[0,\infty]$ は可測関数, Q を $(X,\mathscr{F})$ から $(Y,\mathscr{G})$ への確率核とする. このとき以下で定義する X 上の関数 g は可測関数になる.

$$g(x):=\int_Y f(x,y)Q(x:dy)$$

定理 4.4.
$\{f_n\}$ を測度空間 $(X,\mathscr{F},\mu)$ 上の可測関数列とする. $\forall n\in\mathbb{N}$ について X の各点で $f_n(x)\geq 0$ かつ $\displaystyle\sum_{n=1}^{\infty}f_n(x)$ が存在するならば

$$\int_X\sum_{n=0}^{\infty}f_n(x)\mu(dx)=\sum_{n=1}^{\infty}\int_X f_n(x)\mu(dx)$$

が成り立つ.

証明の流れをまとめる:

1. $\left(\displaystyle\prod_{s=0}^{T}\Omega_s,\ \bigotimes_{s=0}^{T}\mathscr{F}_s\right)$ 上に推移確率測度 $P_{T+1}^{0,\ldots,T}$ を用いて, 確率測度 $P_{x_0,\ldots,x_t}$ を構成する.
2. $\mathscr{B}:=\displaystyle\bigcup_{T=0}^{\infty}\bigotimes_{s=0}^{T}\mathscr{F}_s$ とする. 確率測度 $P_{x_0,\ldots,x_t}$ は $\left(\displaystyle\prod_{s=0}^{\infty}\Omega_s,\ \mathscr{B}\right)$ 上の有限加法的確率測度になる.
3. $\displaystyle\bigotimes_{s=0}^{\infty}\mathscr{F}_s=\sigma[\mathscr{B}]$ を示す.
4. $P_{x_0,\ldots,x_t}$ が確率測度の拡張定理の条件を $\left(\displaystyle\prod_{s=0}^{\infty}\Omega_s,\ \mathscr{B}\right)$ 上で満たすことを示す.
5. 確率測度の拡張定理を用いて $P_{x_0,\ldots,x_t}$ を $\left(\displaystyle\prod_{s=0}^{\infty}\Omega_s,\ \bigotimes_{s=0}^{\infty}\mathscr{F}_s\right)$ 上の確率測度へと拡張する.

証明
$\left(\displaystyle\prod_{s=0}^{T}\Omega_s,\bigotimes_{s=0}^{T}\mathscr{F}_s\right)$ 上に集合関数 $P_{x_0,\ldots,x_T}$ を次のように定義する:
$A=\displaystyle\prod_{s=0}^{T}E_s,\ E_s\in\mathscr{F}_s$ に対しては, $P_{x_0,\ldots,x_T}(A):=1_A(x_0,\ldots,x_T)$
$A=\displaystyle\prod_{s=0}^{t}E_s\times\prod_{s=t+1}^{T}\Omega_s,\ E_s\in\mathscr{F}_s$ に対しては
$P_{x_0,\ldots,x_t}(A):=\displaystyle\int_{\Omega_{t+1}}P_{x_0,\ldots,x_{t+1}}(A)P_{t+1}^{0,\ldots,t}(x_0,\ldots,x_t:dx_{t+1})$ と定義する. 簡単に状況を整理する. まだ証明はしていないが便宜上, 上記の集合関数の値を確率と呼ぶことにする.$T=2$ とし, 考えてる可測空間は $\Omega=\{1,2,3,4,5,6\}$ とすると $(\Omega^3,\ 2^{\Omega^3})$ である. サイコロ投げを考える. 観測結果として $x_0=1,\ x_1=3,\ x_2=5$ とする. このとき上記で定義した確率 $P_{1,3,5}$ は

$$P_{1,3,5}(A)=1_A((1,3,5))$$

である. これはすでに結果が出ているので, 指示関数の形にするのは自然である. 次に x_2 はまだ観測されてい場合を考える. つまり, $P_{1,3}((1,3,5))$ を考える.

$$
\begin{aligned}
P_{1,3}((1,3,5)) &= \sum_{x_2=5} P_2^{0,1}(x_0=1, x_1=3 : X_2=x_2) P_{1,3,x_2}((1,3,x_2)) \\
&= P_2^{0,1}(x_0=1, x_1=3 : x_2=5) \\
&= \text{サイコロを 1 回振って出る目が 5 の確率} \\
&= \frac{1}{6}
\end{aligned}
$$

つまり $P_{x_0,\ldots,x_t}$ とは逐次的に定義された集合関数である. この集合関数が $\left(\prod_{s=0}^{T}\Omega_s, \bigotimes_{s=0}^{T}\mathscr{F}_s\right)$ 上の確率測度になっていることを示す. $t=T$ のときは明らか.$P_{x_0,\ldots,x_{t+1}}$ が $\left(\prod_{s=0}^{T}\Omega_s, \bigotimes_{s=0}^{T}\mathscr{F}_s\right)$ 上の確率測度と仮定する. $P_{x_0,\ldots,x_t}$ が $\left(\prod_{s=0}^{T}\Omega_s, \bigotimes_{s=0}^{T}\mathscr{F}_s\right)$ 上の確率測度になることを示す.

非負性については
$P_{x_0,\ldots,x_{t+1}}$ が確率測度なので, 任意の $A \in \bigotimes_{s=0}^{T}\mathscr{F}_s$ に対して, $P_{x_0,\ldots,x_t}(A) \geq 0$ を得る.
次に標本全体の測度が 1 になることを示す. $\Omega = \prod_{s=0}^{T}\Omega_s$ とする. つまり, $\Omega = \prod_{s=0}^{t}\Omega_s \times \prod_{s=t+1}^{T}\Omega_s$ とみる.

$$
\begin{aligned}
P_{x_0,\ldots,x_t}(\Omega) &= \int_{\Omega_{t+1}} P_{x_0,\ldots,x_{t+1}}(\Omega) P_{t+1}^{0,\ldots,t}(x_0,\ldots,x_t : dx_{t+1}) \\
&= \int_{\Omega_{t+1}} P_{t+1}^{0,\ldots,t}(x_0,\ldots,x_t : dx_{t+1}) \\
&= 1
\end{aligned}
$$

が示される. σ 加法性については, 定理 4.3 により $P_{x_0,\ldots,x_{t+1}}(A)$ を $\{x_0,\ldots,x_t\}$ と $A \in \bigotimes_{s=0}^{T}\mathscr{F}_s$ を固定したとき x_{t+1} に関して可測関数だと仮定すると, $P_{x_0,\ldots,x_t}(A)$ は x_t に関して可測関数であることがすぐに証明でき, $P_{x_0,\ldots,x_T}(A)$ が x_T に関する可測関数になることは明らか. $P_{x_0,\ldots,x_{t+1}}(A)$ が x_{t+1} について可測関数だと仮定すると定理 4.3 より,$P_{x_0,\ldots,x_t}$ は x_t に関して可測関数になる. 従って, 任意の t に対して $P_{x_0,\ldots,x_t}(A)$ は x_t に関する可測関数である. $P_{x_0,\ldots,x_{t+1}}$

を $\left(\prod_{s=0}^{T}\Omega_s, \bigotimes_{s=0}^{T}\mathscr{F}_s\right)$ 上の確率測度と仮定する.

$$
\begin{aligned}
P_{x_0,\dots,x_t}\left(\sum_{s=0}^{\infty}A_s\right) &= \int_{\Omega_{t+1}} P_{x_0,\dots,x_{t+1}}\left(\sum_{s=0}^{\infty}A_s\right)P_{t+1}^{0,\dots,t}(x_0,\dots,x_t : dx_{t+1}) \\
&= \int_{\Omega_{t+1}}\sum_{s=0}^{\infty}P_{x_0,\dots,x_{t+1}}(A_s)P_{t+1}^{0,\dots,t}(x_0,\dots,x_t : dx_{t+1}) \\
&= \sum_{s=0}^{\infty}\int_{\Omega_{t+1}}P_{x_0,\dots,x_{t+1}}(A_s)P_{t+1}^{0,\dots,t}(x_0,\dots,x_t : dx_{t+1}) \quad (\because \text{定理 } 4.4) \\
&= \sum_{s=0}^{\infty}P_{x_0,\dots,x_t}(A_s)
\end{aligned}
$$

上記の $P_{x_0,\dots,x_t}$ を $P^{(T)}_{x_0,\dots,x_t}$ と表すことにすると, 任意の T に対して $P^{(T)}_{x_0,\dots,x_t}$ は $\left(\prod_{s=0}^{T}\Omega_s,\ \bigotimes_{s=0}^{T}\mathscr{F}_s\right)$ 上の確率測度であることが証明できた.

$T_1 \leq T_2$ とすると, $P^{(T_2)}_{x_0,\dots,x_t}$ は $P^{(T_1)}_{x_0,\dots,x_t}$ の拡張になっている.
実際, $A \in \bigotimes_{s=0}^{T_1}\mathscr{F}_s$ を $A = \prod_{s=0}^{T_1}A_s \times \prod_{s=T_1+1}^{T_2}\Omega_s$ と見なすことにすれば, $P^{(T_1)}_{x_0,\dots,x_t}(A) = P^{(T_2)}_{x_0,\dots,x_t}(A)$ となる. つまり $P^{(T_2)}_{x_0,\dots,x_t}$ は $\left(\prod_{s=0}^{T_1}\Omega_s,\ \bigotimes_{s=0}^{T_1}\mathscr{F}_s\right)$ 上の確率測度になっているので $P^{(T_2)}_{x_0,\dots,x_t}$ は $P^{(T_1)}_{x_0,\dots,x_t}$ の拡張になっている.
$\mathscr{B} := \bigcup_{T=0}^{\infty}\bigotimes_{s=0}^{T}\mathscr{F}_s$ とする. 証明は省略するが, $\mathscr{B}$ は $\prod_{s=0}^{\infty}\Omega_s$ 上の有限加法族になっている.
$A \in \mathscr{B}$ とすると定義から, A に対してある自然数 T が存在し, $A \in \bigotimes_{s=0}^{T}\mathscr{F}_s$ となる. このとき, $\left(\prod_{s=0}^{\infty}\Omega_s,\ \mathscr{B}\right)$ 上の集合関数 $P^{*}_{x_0,\dots,x_t}$ を以下のように定義する:

$$
P^{*}_{x_0,\dots,x_t}(A) := P^{(T)}_{x_0,\dots,x_t}(A)
$$

このとき, $P^{*}_{x_0,\dots,x_t}$ は $\left(\prod_{s=0}^{\infty}\Omega_s,\ \mathscr{B}\right)$ 上の有限加法的確率測度になる. 従って, 定理 4.2 より,

$$
B_1 \supset B_2 \supset \cdots,\ B_n \in \mathscr{B},\ \bigcap_{n=1}^{\infty}B_n = \phi \Rightarrow \lim_{n\to\infty}P^{*}_{x_0,\dots,x_t}(B_n) = 0
$$

を示せれば良い.
単調減少列な事象列 $\{B_n\} \subset \mathscr{B}$ について, $\bigcap_{s=0}^{\infty}B_s = \phi$ と仮定する.
このとき, $\lim_{n\to\infty}P^{*}_{x^*_0,\dots,x^*_t}(B_n) > 0$ となる t と $x^*_0,\dots,x^*_t$ が存在すると仮定する. 任意の $A \in \mathscr{B}$ に対して $P^{*}_{x_0,\dots,x_t}(A)$ は $\left(\prod_{s=0}^{T}\Omega_s,\ \bigotimes_{s=0}^{T}\mathscr{F}_s\right)$ 上の可測関数である. $\{B_n^c\}$ は単調増加な事象列なので

$\{P^*_{x_0,...,x_t}(B^c_n)\}$ は非負単調増加可測関数列である. 従って Lebesgue の単調収束定理を用いて次のような式変形をする.

$$\begin{aligned}\lim_{n\to\infty} P^*_{x_0^*,...,x_t^*}(B_n) &= \lim_{n\to\infty}\int_{\Omega_{t+1}} P^*_{x_0^*,...,x_t^*,x_{t+1}}(B_n)P^{0,...,t}_{t+1}(x_0^*,...,x_t^*:dx_{t+1})\\ &= \lim_{n\to\infty}\int_{\Omega_{t+1}} (1-P^*_{x_0^*,...,x_t^*,x_{t+1}^*}(B^c_n))P^{0,...,t}_{t+1}(x_0^*,...,x_t^*:dx_{t+1})\\ &= 1-\int_{\Omega_{t+1}} \lim_{n\to\infty} P^*_{x_0^*,...,x_t^*,x_{t+1}^*}(B^c_n)P^{0,...,t}_{t+1}(x_0^*,...,x_t^*:dx_{t+1})\\ &= \int_{\Omega_{t+1}} \lim_{n\to\infty} P^*_{x_0^*,...,x_t^*,x_{t+1}}(B_n)P^{0,...,t}_{t+1}(x_0^*,...,x_t^*:dx_{t+1})\end{aligned}$$

左辺が正なので, ある x^*_{t+1} が存在して, $\lim\limits_{n\to\infty} P^*_{x_0^*,...,x_t^*,x_{t+1}^*}(B_n) > 0$ となる. これを繰り返すことで $\omega^* := (x_0^*,...)$ という列が作れる. つまり, $\forall u > t$ に対して $\lim\limits_{n\to\infty} P^*_{x_0^*,...,x_u^*}(B_n) > 0$ となる ω^* が存在する. 任意に固定された n に対して, u が十分大きく, $B_n \in \bigotimes_{s=0}^{u} \mathscr{F}_s$ とすると, $P_{x_0,...,x_T}$ の定義より, $P^*_{x_0,...,x_u}(B_n) = 1_{B_n}(x_0,...,x_u)$ となる. $\lim\limits_{n\to\infty} P^*_{x_0^*,...,x_u^*}(B_n) > 0$ なので,$\forall n \in \mathbb{N}$ に対して, $\lim\limits_{n\to\infty} P^*_{x_0^*,...,x_u^*}(B_n) > 0$ つまり $1_{B_n}(x_0^*,...,x_u^*) = 1$ である. よって, $\forall n \in \mathbb{N}$ に対して, $\omega^* \in B_n$ となり, $\bigcap_{s=0}^{\infty} B_s \neq \phi$ となるので矛盾. よって $\lim\limits_{n\to\infty} P^*_{x_0,...,x_t}(B_n) = 0$ となる. 定理 4.2 より, $P^*_{x_0,...,x_t}$ は $\left(\prod_{s=0}^{\infty}\Omega_s, \bigotimes_{s=0}^{\infty}\mathscr{F}_s\right)$ 上の確率測度である. □

5　今後の課題

今回のような抽象空間上のマルコフ決定過程に関する理論を中心に修士論文に取組みたいと考えている. 特に状態空間, 決定空間が非コンパクト集合の場合や, 推移する時間間隔が確率変数になってる場合のマルコフ決定過程の最適政策の存在について取り組みたい.

参考文献

[1] Hernández-Lerma and Myram Muñoz De Ozak. Discrete-time Markov control processes with discounted unbounded costs: Optimality criteria. Kybernetika. 1992. pp, 191–212.

[2] Christilam Léonard. A set lecture notes on convex optimization with some applications to probability theory incomplete draft. 2006 May 06.

[3] Jacques Neveu. Mathematical foundations of the calculus of probability. Holden-day, inc. 1965.

[4] 伊藤清. 確率論. 岩波書店. 1991.

[5] 舟木直久. 確率論. 朝倉書店. 2004.

[6] 中出康一. マルコフ決定過程–理論とアルゴリズム–. コロナ社. 2019.