

特集◎最適タイミングの数理

ディンキンゲームと その多数回停止モデルへの拡張

Krzysztof Szajowski 安田 正實

1 はじめに

ディンキンゲームとよばれる最適停止時刻問題(最適停止問題)とは、確率過程に関する停止時刻から定めた期待利得を、2人プレーヤによる零和ゲーム問題へと変形した最適化問題をさす。本特集号にみられるよう、最適停止問題にはさまざまな興味ある問題やその拡張などが多く議論され発展しているが、そのうちの一つの基本的な変形である。このディンキンゲームの由来は E. B. Dynkin が 1969 年(英訳)に停止問題のゲーム変形として発表したものがオリジナルであると思われるが、今日ディンキンゲームとよばれる最適停止問題は、ディンキン型ではなく、むしろヌブー型とよぶべきであろう。1975 年 J. Neveu が今日の数理ファイナンスの理論では基幹的な概念となってマルティンゲールのテキストで取り上げたタイプである。

本論ではこのディンキンゲームの古典的オリジナル問題を述べ、その後の主要な結果を解説し、ひとつの拡張として停止という決定を多数回可能とした場合のモデルについて議論しよう。

[筆者紹介]



クリスチョフ サジョスキ。1973 年プロツコフ工科大学、応用数学修士(M. Sc.)取得。1979 年同大学(数学)Ph. D. 取得。1973 年同大学リサーチアシスタント、その後同大学講師を経て、同大教授、現在は同大学数学研究所所長。この間オランダ・デルフェスト工科大学客員教授を兼務。研究テーマは確率過程の最適停止問題、確率モデリング、数理ファイナンス、医学統計等。1981 年ポーランド科学教育大臣賞受賞、ポーランド数学会、アメリカ数学会各フェロー会員、日本数理科学協会編集委員。



やすだ まさみ。1969 年千葉大学文理学部数学科卒業。1971 年九州大学大学院理学研究科修了、同年鹿児島大学助手、その後千葉大学講師、助教授を経て、1989 年同大学教授、現在に至る。研究テーマは確率モデルにおける最適化問題、ファジィ環境でのモデル化とその解析。理学博士(九州大学)。日本数学会、日本 OR 学会、日本統計学会、日本応用数理学会各会員。Journal of Information and Optimization Sciences 編集委員。

2 マルコフ型決定過程の最適性方程式

マルコフ過程あるいは連鎖と限定して、その推移(運動)法則に決定者の意志を反映させるよう、決定変数を導入し、その選択した決定と各時点での状態変化により得られる利得を、計画期間全体にわたって総計した総利得を最大にする問題がマルコフ型決定過程とよばれる。いわゆる統計的逐次決定過程の流れをくむ多段決定問題のひとつである。

いま決定者が2つの選択肢のうちから一つの決定を選ぶ。一つはそのまま過程の観測をし続けること、もう一つは現在状態へと変化したことを考慮し、停止の決定をとって、このときに得られる利得で過程の観測を終了すること。このように観測をしつづけるか、停止するか2つの選択肢をもつ2決定問題はマルコフ型決定過程のなかで決定個数が最も少ない場合である。したがって一般に動的計画法での最適性の原理を適用して得られる「最適性方程式」が簡潔な場合、すなわち決定が2選択の場合に問題が帰着される。これが最適停止問題の特徴でもある。

与えられた状態空間 $(E, \mathcal{B})$ 上の確率空間 $(\Omega, \mathcal{F}, \mathbf{P})$ での時間一様なマルコフ連鎖を $(X_n, \mathcal{F}_n)$, $n=0, 1, \dots, N$ と表す。この確率過程の実現値に対して、ある利得関数 $\varphi(x)$, $x \in E$ を用いて最適停止問題が定式化できる。一般に利得関数は、 $\varphi: E \rightarrow R$ の B_+ 可測な関数である。また確率空間に附随する停止時刻 $\tau = \tau(\omega)$ とは、連続(離散)時間の値をとる確率変数であって、確率過程に対する合成関数として新たな確率過程 $X_\tau(\omega) = X_{\tau(\omega)}(\omega)$ が定義される。われわれの目的はこの停止時刻から定めた確率過程の期待値を最大化すること。すなわち $v(x) = \sup_\tau \mathbf{E}_x \varphi(X_\tau)$, ここでは停止した時刻のみでの利得の値しか考慮せず、この時刻までの継続期間の決定に対しては利得がゼロである。この形でも一般性を失わない。

よく知られた動的計画法の最適性原理から、つぎの最適性方程式(あるいはベルマン方程式)とよばれる関数方程式が導かれる。これは繰り返しの状況を反映した再帰関係式であって、マルコフ過程を対象とする場合には、簡潔に表現され、決定に応じた要素が表れる。すなわち

$$v(x) = \max \{ \varphi(x), \mathbf{P}v(x) \}, \quad x \in E$$

であり、この $\max$ 演算は、どちらがより多くの利得を得ることができるかという決定に応じている。カッコ内の2項は、前者が停止による利得で現状態値に対応する値が利得として得られるもの、後者の部分が現時刻では継続することで、その後同じ最適政策をとることで得られる総利得に対応する値を意味する。

この最適性方程式はつぎの形にも書き直すことができる。 $\forall x \in E$ に対し、

$$v(x) \geq \varphi(x), \quad v(x) \geq \mathbf{P}v(x), \quad (v(x) - \varphi(x))(v(x) - \mathbf{P}v(x)) = 0$$

これはいわゆる変分不等式とよばれ、さまざまな最適化問題の特徴付けなどに

については有効な解析手段である。

ディンキンゲームの場合(ヌブー型)での変分不等式は、モデルで与えた2つの利得関数(2オブジェクトを考える)に適切な条件を課した2オブスタクル問題となる。同じように自由境界問題として既に Grigelionis & Shiryaev[7], R. Myneni[13] が定式化している。Benssousan & Lions[1] には詳しい議論が述べられている。また凸計画問題形のマルコフ型決定過程として論文[23]がある。さらに関連したいくつかの話題については、参考文献のリストを御覧いただきたい。

3 ディンキンゲームのはじまり

確率過程の最適停止問題はマルティンゲールとの関係が知られていたが、1969年のディンキンが定式化したゲーム版としてのモデルの期待利得は、

$$E[I_{\{n < \tau_0, X(n) \geq 0\}} a(\tau_1) + I_{\{n < \tau_0, X(n) \leq 0\}} a(\tau_2)]$$

のタイプとして与えた。目的は、それぞれのプレーヤが停止時刻を選び、一方が最大、他方が最小となるようにしたゼロ和ゲームである。ある確率過程 $\{X(n)\}$ に対して、一つの利得関数 $\{a(n)\}$ と正と負の領域に分割した停止できる領域を考えている。ここで I_A は集合 A の示性関数とする。

1975年のヌブーでは、停止できる領域を制限することではなく、2種類の利得関数 $\{a(n), b(n)\}$ を取り入れる。ただしこの利得には意味をもたせるために大小関係を仮定する。すなわち $a(n) \leq b(n)$ とする。この仮定の根拠には触れていないが、後述している停止時刻(戦略)のランダム化と関係をもつことが分かる。

$$E[I_{\{n \leq \tau_0\}} a(\tau_1) + I_{\{n > \tau_0\}} b(\tau_2)]$$

このようなタイプの利得関数に関する停止問題が通常ディンキンゲームとよばれているものである。

さらに1985年の論文[22]では

$$E[I_{\{n < \tau_0\}} a(\tau_1) + I_{\{n > \tau_0\}} b(\tau_2) + I_{\{n = \tau_0\}} c(\tau_1)]$$

として議論した。とくに3つの利得関数 $\{a(n), b(n), c(n)\}$ には大小関係を仮定しないことが特徴である。しかし、この仮定をしない結果、戦略の拡張を考える必要が起ってくる。利得関数 $a(n)$ と $b(n)$ は一方が他方に先駆けて停止した場合に得られる利得を表す。どちらが先に明示したかにより、それぞれ異なる利得が対応する。もうひとつの利得 $c(n)$ は同時に停止をした場合である。ヌブー型問題のように $a(n) \leq b(n)$ を仮定すると、同時に停止することが均衡解にならず、したがって、停止時刻を純戦略とみなし、この枠組みの中で考えておけば十分であることが分かる。またこの大小関係の意味することは変分不等式問題でも理解できるが、自分の決定が他方の決定に対してコンフリクトの起らないように調整されている。同様のことは、マルコフ型決定問題の例題と

して、あるいはオペレーションリサーチでの古典的な在庫問題でよく知られているよう、発注と在庫の兼ね合いから、最適な注文量とその在庫水準を自由境界問題としてとらえたいいわゆる (s, S) 政策がある。ここでいうゲーム版の最適停止問題はこれを拡張した両サイド (s, S) 政策に対応している。

4 ランダム化した停止時刻

停止時刻をランダム化することは、通常最適停止時刻問題にはその必要を生じない (Chow/Robbins/Siegmund[3]) ことが知られている。しかし、ゲーム変形にした場合には必要性が起ってくる。この考え方は、いわゆる行列ゲームにおける均衡解の存在を混合戦略のなかで考えることの延長とみなされる。よく知られているよう、2人ゼロ和ゲーム、すなわち行列ゲームにおけるミニマックス均衡を考えると、一般には純戦略ではなく混合戦略のなかに最適戦略が存在する。これと同じことが、逐次決定のゲーム版である最適停止時刻問題でも起っている。通常確率変数である、停止時刻が純戦略に対応し、そのランダム化した停止時刻 (停止時刻のランダム化 Irle[9]) が混合戦略に対応する。

なぜならば、ゲーム変形した停止時刻問題の最適性方程式は、 $\max, \min$ を同時に考慮した行列ゲームの値 val をつかえば表現されるから。具体的には支払い行列が $v_n(x)$ (present stage) から $\mathbf{P}v_{n+1}(x)$ (next stage) への再帰的關係式として

$$v_n(x) = \text{val} \begin{bmatrix} c & a \\ b & \mathbf{P}v_{n+1}(x) \end{bmatrix}$$

の形で与えられるから、いわゆる行列ゲームのミニマックス均衡解がすぐに連想される。このようなゼロ和ゲームではなく、自然な拡張として非ゼロ和ゲーム[16]も考えられる。

5 多数回停止選択が可能なゲーム版への拡張

5.1 多数回の停止時刻問題

最適停止問題において、多数回の停止を考えた参考論文として、Haggstrom[8], Nikolaev[15] などが挙げられる。われわれの目的は、このような多数回をヌプー型に拡張してみることである。

ある時間一様なマルコフ連鎖 $(X_n, \mathcal{F}_n, \mathbf{P}_x)_{n=0}^N$ が状態空間 $(E, \mathcal{B})$ 上の確率空間 $(\Omega, \mathcal{F}, \mathbf{P})$ で定義され、さらに $f: E \times E \rightarrow \mathcal{R}$ は、 $\mathcal{B} \times \mathcal{B}$ 実数値可測関数が与えられたとする。また計画期間 N は有限と仮定する。このマルコフ連鎖の観測を続けていき、関数 f による評価値が最も高くなる “the best realizations” を選択することが目的である。集合族 $\{\mathcal{F}_n\}_{n=0}^N$ に関するマルコフ時間の集まりを $\mathcal{S}^N$ と表す。ここでは $\tau \in \mathcal{S}^N$ に対し、 $\mathbf{P}_x(\tau \leq N) < 1$ を認めるとする。つまりマルコフ連鎖が正の確率で停止しないことも許す。この集合 $\mathcal{S}^N$ の要素

がここでの停止時刻問題のとり得る戦略である.

2 回の停止回数をもつ問題では集合 $\mathcal{S}^N$ から停止時間の組を選んでつぎの期待利得を最大にする. つまり

$$v_N^{(2)}(x) := \sup_{\substack{\tau_1, \tau_2 \in \mathcal{S}^N \\ \tau_1 < \tau_2}} \mathbf{E}_x \tilde{f}(X_{\tau_1}, X_{\tau_2}) \quad x \in E \quad (5.1)$$

とし, 最適戦略 $\tau_1^*, \tau_2^* \in \mathcal{S}^N$ はすべての $x \in E$ に対し, $\mathbf{E}_x \tilde{f}(X_{\tau_1^*}, X_{\tau_2^*}) = v_N^{(2)}(x)$ となるものを求める. ただし,

$$\tilde{f}(X_{\tau_1}, X_{\tau_2}) := \begin{cases} f(X_{\tau_1}, X_{\tau_2}) & \text{if } \max(\tau_1, \tau_2) \leq N, \\ f_1(X_{\tau_1}) & \text{if } \tau_1 \leq N, \tau_2 = \infty, \\ f_0 & \text{if } \tau_1 \leq \infty, \tau_2 = \infty, \end{cases}$$

かつ, $f_1(x) = \inf_{y \in E} f(x, y)$, $f_0 = \inf_{x, y \in E} f(x, y)$, $\{\tau = \infty\} = \{\tau \leq N\}$.

2 回停止問題の解を構成するためにつぎの補助問題を解く. まず $\mathcal{S}_k^N := \{\tau \in \mathcal{S}^N : \tau \geq k\}$, $\tilde{\mathcal{S}}^N := \{(\tau, \{\sigma_n\}) : \tau \in \mathcal{S}^N, \sigma_n \in \mathcal{S}_{n+1}^N \text{ for every } n\}$ とおき, $\mathfrak{s} \in \tilde{\mathcal{S}}^N$ に対して, $\varphi_N(x, \mathfrak{s}) := \mathbf{E}_x \tilde{f}(X_\tau, X_{\sigma_\tau})$ とおく. 式(5.1)から $v_N^{(2)}(x) = \sup_{\mathfrak{s} \in \tilde{\mathcal{S}}^N} w(x, \mathfrak{s})$ と $\tilde{v}_N^{(2)}(n, x) = \mathbf{E}_x v_{N-n-1}^{(2)}(X_1)$ を得る. また

$$v_N(x, y) := \sup_{\tau \in \mathcal{S}^N} \mathbf{E}_y \tilde{f}(x, X_\tau) \quad (5.2)$$

とすれば, 問題(5.2)の解はつぎの動的計画法で構成される.

$$Q_y f(x, y) := \max \{f(x, y), T_y f(x, y)\}$$

ここで $T_y f(x, y) := \mathbf{E}_y f(x, X_1)$ ([19]をみよ)とおき, これから, N 回の繰り返しで $Q_y^N = Q_y(Q_y^{N-1})$ を定義すると, $v_N(x, y) = Q_y^N f(x, y)$ が成り立ち, 最適なマルコフ時間は $\tau^* = \min \{n \leq N : v_{N-n}(x, X_n) = f(x, X_n)\}$ となる.

この手順を利用して, 2 回停止の問題における最適停止時刻の第 2 成分が構成できる. 一方第 1 成分は $g(n, x) = \mathbf{E}_x v_{N-n-1}(x, X_1)$ ([8], [15]をみよ)におけるマルコフ連鎖の最適停止問題の解として得られる. いま $Qg(n, x) := \max \{g(n, x), Tg(n, x)\}$, $Tg(n, x) := \mathbf{E}_x g(n+1, X_1)$ とおく. 制限した停止時刻 $\mathcal{S}_n^N$ に対しての利得関数 $g(n, x)$ をもつ最適停止問題の最適値は $\tilde{v}_N(n, x)$ である. この解は $\tilde{v}_N(n, x) = Q^{N-n} g(n, x)$ で与えられ, ベルマン方程式 $\tilde{v}_N(n, x) = \max \{g(n, x), T\tilde{v}_N(n+1, x)\}$ を満たすことがわかる.

5.2 優先権をもつ場合の最適停止ゲーム問題

プレーヤ 1 は 1 回の停止だけであるが, プレーヤ 2 は 2 回の停止を許す場合を考えよう. そのときの利得関数は $f: E \times E \times E \rightarrow \mathfrak{R}$ とし, このようなゲームを $\mathcal{G}_{pn}^{(12)}$ と表す. この関数はそれぞれの選んだ状態の値に依存している. 確率変数 X_n の実現値を x_n とする. もし両方のプレーヤがまだこの実現値で採択をしていないときには, 時刻 n におけるマルコフ連鎖の状態を評価して, 2 つのオプションを選べる. 観測した値を採択するか, 棄却するかである. もし

同時に両者が採択をしようとしたならば、ルールとして、プレーヤ 1 がこれを獲得できるものとする。しかしプレーヤ 1 がこの候補を棄却するときには、プレーヤ 2 の手番になり、採択か棄却かの決定をくだす。両者の一方が確率変数 X_n の実現値 x_n を採択したならば、他方はこの通知を受けるとする。プレーヤ 1 が最初に採択をしてしまったならば、これから後はプレーヤ 2 だけの問題となる。逆にプレーヤ 2 が採択をすれば、それぞれ 1 回停止の選択をもち、プレーヤ 1 が優先される場合の 2 人ゼロ和ゲームに帰着される。このときの決定で、プレーヤ 1 は採択の値が x 、またプレーヤ 2 は y, z で採択をしたとするならば、対応するプレーヤ 2 がプレーヤ 1 へ $f(x, y, z)$ を支払う。同様にプレーヤ 1 が状態の値 x 、プレーヤ 2 が $(y, z \text{ or only } y)$ で採択したとすれば、仮定によって、プレーヤ 1 は $f_1(x) = \sup_{y, z \in E} f(x, y, z)$ ($f_2(y, z) = \inf_{x \in E} f(x, y, z)$ あるいは $f_{21}(y) = \inf_{z \in E} f_2(y, z)$) を獲得する。また 1 回のみしか選択をしていないならば、利得は $f_3(x, y) = \sup_{z \in E} f(x, y, z)$ となる。全く選択をせずに決定過程が終了したときには、両者の利得は 0 であるとする。

プレーヤのとり戦略としては状態を観測して得るから、 $(\mathcal{I}_n)_{n=0}^N$ に関する停止時刻を考えることが自然であろう。いまの場合では一つだけの確率変数列 $\{X_n\}_{n=0}^N$ を対象にしているから、各時刻 n ではプレーヤは X_n の実現値 x_n を得ている。優先順位のあるばあい、すなわち両プレーヤが同時に対象の割り当て、獲得を得たいときの問題は、[5], [20], [21] など議論されている。優先権があれば、実現値 x_n を得ることができ、そうでなければ、時刻 $n < N$ における x_n の選択は拒まれ、その後表れる実現値だけが選択できる。

リコールなし、すなわち現時点から過去にさかのぼって選択することはできない、と仮定する。マルコフ過程によって記述されるという特徴をもった対象の探索という意味での決定過程と考えることができよう。ここでは両方のプレーヤを合わせて 3 回までの対象を選択できるから、この決定過程としては、プレーヤの選択としての戦略は、ゲームの異なった状態の値に準備しておかねばならない。

いま Λ_k^N と M_k^N を $\mathcal{S}_k^N$ のコピー ($\mathcal{S}^N := \mathcal{S}_0^N$) とおく。戦略の組 $\tilde{\Lambda}^N = \{(\lambda, \{v_n^1\}, \{v_{nm}^1\}) : \lambda \in \Lambda^N, v_n^1 \in \Lambda_{n+1}^N, v_{nm}^1 \in \Lambda_{m+1}^N; n < m\}$ および $\tilde{M}^N = \{(\mu, \{v_n^2\}, \{v_{nm}^2\}) : \mu \in M^N, v_n^2 \in M_{n+1}^N, v_{nm}^2 \in M_{m+1}^N; n < m\}$ をそれぞれプレーヤ 1 と 2 の戦略と定義する。各々の $s \in \tilde{\Lambda}^N$ と $t \in \tilde{M}^N$ に対して、プレーヤ 1 の停止時刻 τ およびプレーヤ 2 の停止時刻の組 (σ^1, σ^2) に対応した時刻は

$$\begin{aligned} \{\tau = n\} &= \{\lambda = n, \mu \geq n\} \cup \bigcup_{i=0}^{n-1} \{\mu = i, v_i^1 = n, v_i^2 \geq n\} \\ &\quad \cup \bigcup_{k=0}^{n-1} \bigcup_{i=k+1}^{n-1} \{\mu = k, v_k^2 = i, v_{ki}^1 = n\} \\ \{\sigma_1 = n\} &= \{\mu = n, \lambda \geq n\} \cup \bigcup_{i=0}^{n-1} \{\lambda = i, v_i^2 = n\} \end{aligned}$$

$$\begin{aligned} \{\sigma_2=m\} = & \bigcup_{i=0}^{n-1} \{\mu=i, v_k^2=m, \lambda \geq m\} \cup \bigcup_{k=0}^{n-1} \bigcup_{i=k+1}^{n-1} \{\lambda=k, v_k^2=i, \mu_{ki}^2=m\} \\ & \cup \bigcup_{k=0}^{n-1} \bigcup_{i=k+1}^{n-1} \{\mu=k, v_k^1=i, \mu_{ki}^2=m\} \end{aligned}$$

と定義できる.

補題 5.1 確率変数 τ, σ_1 および σ_2 は集合族 $(\tilde{\mathcal{F}}_n)_{n=0}^N$ に関するマルコフ時間であって $\sigma_i \neq \tau, i=1, 2, \sigma_1 \neq \sigma_2$ が成り立つ.

(証明) なぜならば, 各 n について $\{\tau=n\} = \{\lambda=n, \mu \leq n\} \cup \bigcup_{i=0}^{n-1} \{\mu=i, v_i^1=n, v_i^2 \geq n\} \cup \bigcup_{k=0}^{n-1} \bigcup_{i=k+1}^{n-1} \{\mu=k, v_k^2=i, v_{ki}^1=n\} \in \tilde{\mathcal{F}}_n$ であるから. 同様に各 n について, $\{\sigma_1=n\} \in \tilde{\mathcal{F}}_n, \{\sigma_2=m\} \in \tilde{\mathcal{F}}_m$ また $\sigma_1 < \sigma_2$ である. したがって τ, σ_1 および σ_2 はマルコフ時間で $\tau \neq \sigma_i, i=1, 2$ を得る. (終)

各々の $n, m=0, 1, \dots, N$ と $x \in E$ について,

$$\mathbf{E}_x f_1^+(X_n) < \infty, \quad \mathbf{E}_x f_2^-(X_n, X_m) < \infty,$$

$$\mathbf{E}_x f_{21}^-(X_n) < \infty, \quad \text{および} \quad \mathbf{E}_x f_3^+(X_n, X_m) < \infty$$

と仮定し, また $\mathfrak{s} \in \tilde{A}^N, t \in \tilde{M}^N$ とする. プレーヤ 1 が得る期待利得として $\tilde{R}(x, \mathfrak{s}, t) := \mathbf{E}_x f(X_{\tau_1}, X_{\tau_2}, X_{\tau_3})$ を定義する.

これによって正規形のゲーム $(\tilde{A}^N, \tilde{M}^N, \tilde{R}(x, \mathfrak{s}, t))$ が定義できる. このゲームを $\mathcal{G}_{pn}^{(12)}$ と表す. すなわちマルコフ過程に対する双形停止の問題は $\mathcal{G}_{pn}^{(12)}$ がこのモデルを表している.

定義 5.1 ゲーム $\mathcal{G}_{pn}^{(12)}$ における組 $(\mathfrak{s}^*, t^*), \mathfrak{s}^* \in \tilde{A}^N, t^* \in \tilde{M}^N$ が均衡解であるとは, 各々の $x \in E$ に対して, $\mathfrak{s} \in \tilde{A}^N$ と $t \in \tilde{M}^N$ が

$$\tilde{R}(x, \mathfrak{s}, t^*) \leq \tilde{R}(x, \mathfrak{s}^*, t^*) \leq \tilde{R}(x, \mathfrak{s}^*, t).$$

を満たすとき.

均衡解の組 $(\mathfrak{s}^*, t^*)$ を構成することが目的である. つぎのような接近法を提案する. 時刻 n での実現値 x_n をプレーヤ 1 が採択したならば, プレーヤ 2 は他のプレーヤの妨げなどを受けることなく, 自分の利得が最大となるよう 2 回の選択を実行できる. もしプレーヤ 2 が時刻 n での実現値 x_n を採択したならば, この 2 人のプレーヤたちは, プレーヤ 1 が優先権をもつゼロ和ゲーム $\mathcal{G}_{pn}^{(11)}$ をすることとなる. このことは, もし時刻 n まで状態の採択をしていないならば, X_n の実現値 x_n を採択あるいは棄却する前に, 相手方のプレーヤのとり戦略からの影響を考慮に入れなければならないということである. この節の終わりにつぎの補助問題を導入しよう. この問題には $\mathcal{G}_{pn}^{(11)}$ と 2 回停止問題の解をもちいる.

5.2.1 部分ゲーム問題 $\mathcal{G}_{pn}^{(1)}$ の解

ゲーム $\mathcal{G}_{pn}^{(1)}$ においては, その利得が $g_{y_1}(x, y) = f(x, y_1, y)$ (cf. [21]) で与えられていたことに注意する. このとき $s_0(x, y_1, y) = S_0(x, y_1, y) = f(x, y_1, y)$ とおき, さらにすべての $x, y_1, y_2 \in E$, $n = 1, 2, \dots, N$ で

$$s_n^{(1)}(x, y_1, y_2) := \inf_{\tau \in \mathcal{S}^n} \mathbf{E}_{y_2} f(x, y_1, X_\tau),$$

$$S_n^{(1)}(x, y_1, y_2) := \sup_{\tau \in \mathcal{S}^n} \mathbf{E}_x f(X_\tau, y_1, y_2)$$

とおく.

マルコフ過程の最適停止問題[19]によって, 関数 $s_n^{(1)}(x, y_1, y_2)$, $S_n^{(1)}(x, y_1, y_2)$ は再帰的な手順により構成できる. つまり $s_n^{(1)}(x, y_1, y_2) = Q_{\min, y_1}^n f(x, y_1, y_2)$, $S_n^{(1)}(x, y_1, y_2) = Q_{\max, y_1}^n f(x, y_1, y_2)$. ただし, ここで $Q_{\min, y_1} f(x, y_1, y_2) = f(x, y_1, y_2) \wedge T_2 f(x, y_1, y_2)$, $Q_{\max, y_1} f(x, y_1, y_2) = f(x, y_1, y_2) \vee T_1 f(x, y_1, y_2)$ および $T_2 f(x, y_1, y_2) = \mathbf{E}_{y_2} f(x, y_1, X_1)$, $T_1 f(x, y_1, y_2) = \mathbf{E}_x f(X_1, y_1, y_2)$, ($\wedge, \vee$ はそれぞれ minimum と maximum を表す). この作用素 $\wedge$ と T_2 ($\vee$ と T_1) は可測性を保つ. したがって $s_n(x, y_1, y_2)$, $S_n(x, y_1, y_2)$ は $\mathcal{B} \otimes \mathcal{B} \otimes \mathcal{B}$ 可測である.

プレーヤ 1 がまず時刻 n ($n = 0, 1, \dots, N-1$) で x を採択したとするならば, その期待利得は

$$h(n, x, y_1, y_2) = \mathbf{E}_{y_2} s_{N-n-1}^{(1)}(x, y_1, X_1) \quad (5.3)$$

であり, $h(N, x, y_1, y_2) = f(x, y_1, y_2)$ となる. プレーヤ 2 のほうが最初であるとするれば, プレーヤ 1 の期待利得は,

$$H(n, x, y_1, y_2) = \mathbf{E}_x s_{N-n-1}^{(1)}(X_1, y_1, y_2) \quad (5.4)$$

さらに $H(N, x, y_1, y_2) = f(x, y_1, y_2)$ となる. 関数 $h(n, x, y_1, y_2)$ と $H(n, x, y_1, y_2)$ の定義は well defined であることに注意する. これらの関数は, 2, 3 番目の変数について $\mathcal{B} \otimes \mathcal{B} \otimes \mathcal{B}$ 可測であり, $h(n, X_1, y_1, y_2)$ と $H(n, x, y_1, X_1)$ は $\mathbf{P}_x$ で積分可能である.

ゲーム $\mathcal{G}_{pn}^{(1)}$ におけるプレーヤ 1 と 2 の戦略の集合をそれぞれ Λ^N と M^N とおく. この $\lambda \in \Lambda^N$ と $\mu \in M^N$ に対して, 支払い関数を

$$r(y_1, \lambda, \mu) := \begin{cases} h(\lambda, X_\lambda, y_1, X_\lambda) \mathbf{I}_{\{\lambda \leq \mu\}} \\ + H(\mu, X_\mu, y_1, X_\mu) \mathbf{I}_{\{\lambda > \mu\}} & \text{if } \lambda \leq N \text{ or } \mu \leq N, \\ 0 & \text{otherwise,} \end{cases} \quad (5.5)$$

で定義する. ゲームの解としてはつぎの関係式; $\forall x \in E$ に対し,

$$R^{(1)}(x, y_1, \lambda, \mu^*) \leq R^{(1)}(x, y_1, \lambda^*, \mu^*) \leq R^{(1)}(x, y_1, \lambda^*, \mu) \quad (5.6)$$

を満たすような均衡解 (λ^*, μ^*) を求めることである. ただし $R^{(1)}(x, y_1, \lambda, \mu) := \mathbf{E}_x r(y_1, \lambda, \mu)$ とする. 戦略の集合を拡張し, 支払い関数を (5.5) とすれば, 戦略集合 Λ^N と M^N をもつゲーム $\mathcal{G}_a$ は Yasuda[22] が議論した Neveu's stopping problem[14] と同値となる. マルコフ過程を観測して, $E \times E$ 上で定めた列 $w_n^{(1)}(x, y_1)$, $n = 0, 1, \dots, N+1$ を以下のように $w_{N+1}^{(1)}(x, y_1) := 0$ から始めて,

順次 $n=N, \dots, 1, 0$ と続ければ,

$$w_n^{(1)}(x, y_1) := \text{val} \begin{bmatrix} h(n, x, y_1, x) & h(n, x, y_1, x) \\ H(n, x, y_1, x) & Tw_{n+1}^{(1)}(x, y_1) \end{bmatrix} \quad (5.7)$$

によって定義できる. ここで

$$Tw_n^{(1)}(x, y_1) := \mathbf{E}_{xw}^{(1)}(X_1, y_1), \quad \tilde{w}^{(1)}(n, y) := Tw_{N-n-1}^{(1)}(X_1, y)$$

とし, さらに記号 $\text{val } A$ の意味は, 支払い行列 A における 2 人ゼロ和ゲームの値である ([12], [22] をみよ).

このような構成である理由を説明する. なぜなら (5.7) における支払い行列が,

$$A = \begin{bmatrix} a & a \\ b & c \end{bmatrix} \quad (5.8)$$

なる形で与えられていて, a, b, c を実数値とし, 行と列をそれぞれ s (stop) および f (forward) とおいて, 直接行列ゲームの値を計算してチェックすることができるから. つまり

補題 5.2 2 人零和ゲームで支払い行列 A が (5.8) で与えられているならば, 純戦略のなかで均衡解 (ε, δ) が存在する.

$$(\varepsilon, \delta) = \begin{cases} (s, s) & \text{if } a \geq b, \\ (s, f) & \text{if } c \leq a < b, \\ (f, s) & \text{if } a < b \leq c, \\ (f, f) & \text{if } a < c < b. \end{cases}$$

この補題の結果と $w_{n+1}^{(1)}$ が可測であることに注意する. 式 (5.7) から関数 $w_i^{(1)}, i=N, N-1, \dots, n+1$ は可測であると仮定する. いま

$$A_n^{\text{ss}} := \{(x, y_1) : h(n, x, y_1, x) \geq H(n, x, y_1, x)\}$$

$$A_n^{\text{sf}} := \{(x, y_1) : Tw_{n+1}^{(1)}(x, y_1) \leq h(n, x, y_1, x) < H(n, x, y_1, x)\}$$

$$A_n^{\text{fs}} := \{(x, y_1) : h(n, x, y_1, x) < H(n, x, y_1, x) \leq Tw_{n+1}^{(1)}(x, y_1)\}$$

とおき, $A_n^{\text{ff}} := E \setminus (A_n^{\text{ss}} \cup A_n^{\text{sf}} \cup A_n^{\text{fs}})$ とする. これらの集合について $A_n^{\text{ss}}, A_n^{\text{sf}}, A_n^{\text{fs}} \in \mathcal{B} \otimes \mathcal{B}$ であり, 補題 5.2 から $w_n^{(1)}(x, y_1) = h(n, x, y_1, x)(I_{A_n^{\text{ss}}}(x, y_1) + I_{A_n^{\text{sf}}}(x, y_1)) + H(n, x, y_1, x)I_{A_n^{\text{fs}}}(x, y_1) + Tw_{n+1}^{(1)}(x, y_1)I_{A_n^{\text{ff}}}(x, y_1)$, であるから, $\mathcal{B} \otimes \mathcal{B}$ 可測となる.

集合 $\{X_n = y_1\}$ 上で $v_n^{1*} := \inf \{k > n | (X_k, y_1) \in A_n^{\text{ss}} \cup A_n^{\text{sf}}\}$ および $v_n^{2*} := \inf \{k > n | (y_1, X_k) \in A_n^{\text{ss}} \cup A_n^{\text{fs}}\}$ と定める.

定理 5.1 支払い関数 (5.5) のゲーム $\mathcal{G}_{pn}^{(11)}$ においてプレーヤ 1 と 2 のそれぞれの戦略 Λ^N と M^N は解をもつ. 集合 $\{X_n = y_1\}$ 上で組 (v_n^{1*}, v_n^{2*}) は均衡点であり, 関数 $w_n^{(1)}(x, y_1), n=0, 1, \dots, N$ がゲームの値である.

定理の証明は[21]の結果を用いる. 均衡点の形は補題 5.2 から得られる.

5.2.2 2回停止の補助問題に対する解

第 5.1 節で述べた結果とゲーム $\mathcal{G}_{pn}^{(12)}$ における利得関数 $f(x, y, z)$ によってプレーヤ 1 が最初に時刻 n で状態 x で採択を選んだとき, プレーヤ 2 がまだ 2 回の停止が可能であれば, 支払い関数 $g_x(y, z) = f(x, y, z)$ を保つことになる. マルコフ連鎖 $(X_n, \mathcal{F}_n, \mathbf{P}_x)$ が時間一様であることを考慮すれば, 時刻 n で状態が $X_n = x$ である場合から, 2 回停止の問題を解くことになる. このプレーヤ 1 の期待利得は $\tilde{\varphi}^{(2)}(n, x, x) = \mathbf{E}_x \varphi_{N-n-1}^{(2)}(x, X_1)$ である. ただし

$$\varphi_N^{(2)}(x, y) := \inf_{\substack{\sigma_1, \sigma_2 \in \mathcal{S}^N \\ \sigma_1 < \sigma_2}} \mathbf{E}_y f(x, X_{\sigma_1}, X_{\sigma_2}) \quad (5.9)$$

とする.

5.2.3 部分ゲーム $\mathcal{G}_{pn}^{(12)}$ の解

第 5.2.1 節での議論と同様に空間 $\mathbf{E}$ 上の関数列 $w_n^{(2)}(x)$, $n=0, 1, \dots, N+1$ を逐次的に $w_{N+1}^{(2)}(x) := 0$ および

$$w_n^{(2)}(x) := \text{val} \begin{bmatrix} \tilde{\varphi}(n, x) & \tilde{\varphi}(n, x) \\ \tilde{w}^{(1)}(n, x) & Tw_{n+1}^{(2)}(x) \end{bmatrix} \quad (5.10)$$

$n=0, 1, \dots, N$ で定める. ここで $Tw^{(2)}(x) = \mathbf{E}_x w^{(1)}(X_1)$, $\tilde{w}^{(2)}(n, y) = Tw_{n-1}^{(2)}(X_1)$ とする. ゲーム $\mathcal{G}_{pn}^{(12)}$ の均衡戦略 (s^*, t^*) は, 再帰的方程式 (5.10) および (5.7) の解から導くことができる.

行列ゲーム (5.10) の停止 (均衡) 集合により, まず最初の採択する停止時刻が定まる. 関数 $w_{N+1}^{(2)}$ は可測である. 関数 $w_i^{(2)}$, $i=N, N-1, \dots, n+1$ は可測であると仮定する. また

$$\begin{aligned} B_n^{\text{ss}} &:= \{x \in \mathbf{E} : \tilde{\varphi}(n, x) \geq \tilde{w}^{(1)}(n, x)\} \\ B_n^{\text{sf}} &:= \{x \in \mathbf{E} : Tw_{n+1}^{(2)}(x) \leq \tilde{\varphi}(n, x) < \tilde{w}^{(1)}(n, x)\} \\ B_n^{\text{fs}} &:= \{x \in \mathbf{E} : \tilde{\varphi}(n, x) < \tilde{w}^{(1)}(n, x) \leq Tw_{n+1}^{(2)}(x)\} \\ B_n^{\text{ff}} &:= \mathbf{E} \setminus (B_n^{\text{ss}} \cup B_n^{\text{sf}} \cup B_n^{\text{fs}}) \end{aligned}$$

とおく. 集合 $B_n^{\text{ss}}, B_n^{\text{sf}}, B_n^{\text{fs}} \in \mathcal{B}$ と補題 5.2 によって,

$w_n^{(2)}(x) = \tilde{\varphi}(n, x)(I_{B_n^{\text{ss}}}(x) + I_{B_n^{\text{sf}}}(x)) + \tilde{w}^{(1)}(n, x)I_{B_n^{\text{fs}}}(x) + Tw_{n+1}^{(2)}(x)I_{B_n^{\text{ff}}}(x)$ が得られるから, よって $w_n^{(2)}(x)$ は $\mathcal{B}_-$ 可測である. また $\lambda^* := \inf \{n | X_n \in B_n^{\text{ss}} \cup B_n^{\text{sf}}\}$, $\mu^* := \inf \{n | X_n \in B_n^{\text{ss}} \cup B_n^{\text{fs}}\}$ と定義する.

定理 5.2 支払い関数 $f(x, y, z)$ と各プレーヤの戦略を $\hat{A}^N, \tilde{M}^N$ とするゲーム $\mathcal{G}_{pn}^{(12)}$ は均衡解をもつ. 組 (s^*, t^*) は, 最初の座標成分を λ^*, μ^* で定め, 第 2 の座標成分は, 定理 5.1 より定義される $\{v_n^{1*}\}, \{v_n^{2*}\}$ で定める. また第 5.2 節で記述した第 3 番目の座標成分は均衡点であり, $w_0^{(2)}(x)$ がこのゲームの値に等しい.

6 終わりに

この拙論は、後半に述べている多数回のディンキンゲームというアイデアで意見交換をしつつ第1著者がまとめたものが主題であって、その翻訳とよぶべきものかも知れません。第2著者は多数人数による停止規則のもとでのナッシュ均衡解を計算したことがあります。また共同論文となった直接のきっかけは投票手順停止問題として“Voting Procedure on Stopping Games of Markov Chain”, K. Szajowski and M. Yasuda, LN in Economics and Math System 445, pp. 68-80, Springer 1997 を発表したことが縁となっています。また千葉大学の蔵野正美先生、中神潤一先生、北九州市立大学の吉田祐治先生とは、ファジィを入れた場合の停止問題も考えています。物事のあいまいさや不確実性を加味するモデルにはファジィの概念を取り入れることも多いので、このような最適停止問題においても、状態の把握をファジィ関数とみなし、さらにファジィ確率変数列に対する停止時刻を考えています([24])。ここでは省略させていただきます。ご参考までに <http://neyman.im.pwr.wroc.pl/~szajow/> や <http://www.math.s.chiba-u.ac.jp/~yasuda/> をご覧いただければ幸いです。

いくつかの誤りを指摘下さった査読者に謝辞を述べたいと存じます。最後にこの機会を与えていただいた、土肥正先生(広島大学)と玉置光司先生(愛知大学)に感謝申し上げます。

[参考文献]

- [1] Bensoussan, A., and Lions, J. L., Applications of Variational Inequalities in Stochastic Control, North-Holland, Amsterdam, 1982.
- [2] Bismut, J. M., Sur un probleme de Dynkin Z., Warsch. V. Geb., 39(1977), 31-53.
- [3] Chow, Y. S., Robbins, H., and Siegmund, D., Great Expectations: The Theory of Optimal Stopping, Houghton Mifflin Co., Boston, 1971.
- [4] Dynkin, E. B., Game variant of a problem on optimal stopping, Soviet Math. Dokl., 10 (1969), 270-274.
- [5] Von Enns, E. G., and Ferenstein, E. Z., On a multi-person time-sequential game with priorities, Sequential Analysis, 6(1987), 239-256.
- [6] Ferenstein, E. Z., A variation of the Dynkin's stopping game, Mathematica Japonica, 38(1993), 371-379.
- [7] Grigelionis, B. I., and Shiryaev, A. N., On Stefan's problem and optimal stopping rules for Markov processes, Theory Probab. Appl., 11(1966), 541-558.
- [8] Haggstrom, G. W., Optimal sequential procedures when more then one stop is required, Ann. Math. Stat., 38(1967), 1618-1626.
- [9] Irle, A., Games of stopping with infinite horizon, ZOR-Math. Methods Oper. Res., 42 (1995), 345-359.
- [10] Karatzas, I., A pathwise approach to Dynkin Games Statistics, Probability and Game Theory, IMS Lecture Notes, 30(1996), 115-125.
- [11] Lepeltier, J. P., and Maingueneau, M. A., Le jeu de Dynkin en théorie générale sans l'hypothese de Mokobodski, Stochastics, 13(1984), 25-44.
- [12] Luce, R. D., and Raiffa, H., Games and decisions, John Wiley & Sons, New York, 1957.
- [13] Myneni, R., The pricing of the American option, Ann. Appl. Probab., 2(1992), 1-23.
- [14] Neveu, J., Discrete-parameter martingales, North-Holland, Amsterdam, 1975.
- [15] Nikolaev, M. L., On optimal multiple stopping of Markov sequences, Theory of Probability and Its Appl., 43(1999), 298-306.
- [16] Ohtsubo, Y., A nonzero-sum extension of Dynkin's stopping problem, Math. Opera-

- tions Research, 12(1987), 277-296.
- [17] Rosenberg, D., Solan, E., and Vielille, N., Stopping games with randomized strategies, Probab. Theory Relat. Fields, 119(2001), 433-451.
- [18] Sakaguchi, M., Optimal stopping games—a review, Math Japonica, 42(1995), 343-351, Correction : 42(1995), 573.
- [19] Shiryaev, A. N., Optimal stopping rules, Springer-Verlag, New York, Heidelberg, Berlin, 1978.
- [20] Szajowski, K., Double stop by two decision makers, Adv. Appl. Probab., 25(1993), 438-452.
- [21] —, Optimal stopping of a discrete Markov process by two decision makers, SIAM J. on Control and Optimization, 33(1995), no. 5, 1392-1410.
- [22] Yasuda, M., On a randomized strategy in Neveu's stopping problem, Stoch. Proc. Appl., 21(1985), 159-166.
- [23] —, A Markov decision process with convex reward and its associated Stopping Game, J. Information & Optimization Sciences, 18(1997), 34-40.
- [24] Yoshida, Y., Yasuda, M., Kurano, M., and Nakagami, J., Optimal Stopping Problems in a Stochastic and Fuzzy System, J. Math. Anal. Appl., 246(2000), 135-149.

[Abstract]

A problem of optimal stopping of the discrete time Markov process by two decision makers in a competitive situation is considered. We briefly follow Dynkin Game and discuss its extension to a multiple stopping model. The double stopping problem is recalled and then a stopping game with priority for Player 1 and the two stops for Player 2 will be formulated. By using the solution of an auxiliary double stopping problem, our aim is to construct the equilibrium pair in the game. The examples of such decision problems is related to variation of the best choice problem and also is a generalization of the other stopping games.