選択肢の価値分布が変化する場合の完全情報最良選択問題

來　島　愛　子

1. はじめに

最適停止問題は経済学、確率論、統計学、金融工学、オペレーションズ・リサーチなど幅広い分野において適用されている問題である。最適停止問題の代表的な問題である確率最大化を目的とした最良選択問題は古典的秘書問題、結婚問題などとよばれ、これまでさまざまな拡張が研究されている。

古典的秘書問題は1950年代半ばにアメリカ東海岸の研究者に知られていたようであるが、Gilbert and Mosteller（1966）によって論文として初めて紹介された。古典的秘書問題の設定において選択肢の総数は既知であるが、総数を未知とし、確率変数と考えその分布を既知とした問題がPresman and Sonin（1972）によって研究されている。また、総数が未知であり、選択肢がポアソン過程にしたがって到着する場合において有限期間内での最良を選択することを目的とした問題も考えられている。ポアソン過程の到着頻度が既知の問題がCowan and Zabczyk（1978）によって解かれている。到着頻度が未知で事前分布が指数分布の場合がBruss（1987）によって、事前分布を自然数パラメータのガンマ分布（アーラン分布）に拡張した問題がKurushima and Ano（2003a）により研究されている。

また、古典的秘書問題を含め、選択肢の順位のみが観測可能な問題は無情報問題とよばれている。一方、選択肢の価値そのものが観測可能な問題を完全情報問題という。2節において、古典的秘書問題の完全情報への拡張に対する問題の記述および最適停止規則を紹介する。完全情報問題に対する研究成果として、総数が未知でその分布が既知の問題に対してPorosinski（1987）によって最適停止規則が求められている。パラメータに事前分布を導入したポアソン到着問題に対してSakaguchi（1989）、Kurushima and Ano（2003b）で研究されているが、最適停止規則は求まっていない。そのほかにも、多数回停止可能な場合や選択肢が確実に得られない場合、なども考えられている。さらに、期待効用最大化を目的とした問題では、例えばベストだけでなくセカンドベストの選択肢を選んでもよいという問題、選択肢の順位の期待値を最小にする問題や、これまでの最良のものを保持している期間を最大にする所有期間最大化問題などがある。

最適停止規則が求められている多くの問題では、ある時点以降、あるいはある時点である値以上（以下）が満たされているならば、そのときに停止することを要求する停止規則が最適であることが示されている。ここで求められた時点、また値のことを閾値とよび、これらによって上の形に記述される停止規則を閾値規則という。閾値規則でない最適停止規則を求めることは困難な場合が多く、Sakaguchi（1989）、Kurushima and Ano（2003b）などの完全情報問題においては最適停止規則が閾値規則でないことが示されている。

今回、完全情報最良選択問題の拡張として選択肢の価値分布が選択過程の途中で変化する場合の最良選択問題を考える。この拡張はジョブサーチにおいてその過程で経済が縮小あるいは拡大する傾向に変化することが予想されている状況などに対応すると考えられる。本稿では、選択肢の価値が観測可能である問題において選択肢の価値分布が意思決定過程のなかで変化する場合を考える。まず、その変化点N_1と選択肢の総数N_2が既知とする。ここで、変化する前の選択肢数がN_1であるとする。さらに、価値分布の変

化に関して以下を仮定する。$i=1, 2, ..., N_1$ のとき、一様分布 $U(0, 1)$ にしたがい、$i=N_1+1, ..., N_2$ のとき、一様分布 $U(0, b)$、$0<b<1$ にしたがうものとする。

以上の設定をもつ価値を観測できる完全情報問題においてその価値分布が選択の途中で変化する場合の最良の選択肢を選ぶ確率を最大化する最適停止規則を求めることが目的である。この問題においては最適停止規則は閾値規則の形で求まることが示された。具体的な停止規則は3節において述べる。

2. 完全情報最良選択問題

まず、古典的秘書問題を完全情報の場合に拡張した基本となる問題の設定と最適停止規則について述べる。問題に関する記述は、無情報問題である古典的秘書問題についての Ferguson（1989）の記述を元にしたものである。

1. ある企業が1人の秘書を探している。この秘書のポストに対して、n 人の応募者が1人ずつ逐次に応募してくる。n 人の応募者の到着はランダムオーダーである。すなわち、到着の順列の総数は $n!$ 通りであり、各順列の起こる確率はそれぞれ $1/n!$ である。
2. 応募者が現れ次第、企業は逐次に面接を行う。企業はこの応募者に対して何らかの基準により応募者の価値を値で知ることができる。価値の分布は一様分布 $U(0, 1)$ とする。
3. 企業は面接が終了した後、直ちに採用か否かを決めなければならない。不採用とした応募者を採用することはできない。採用を決めたとき、この求人活動を打ち切る。
4. 応募者は企業の採用の申し出を拒否しない。
5. 企業は最後の n 人目まで面接したときには、この最後の応募者を必ず採用しなければならない。
6. 企業の目的は、n 人中の価値の最大値をもつ応募者を採用することである。すなわちベストの応募者を採用したとき、利得1を得て、それ以外の応募者を採用したときには利得0を得る。言いかえると、期待利得を最大にすることは、ベストの応募者を採用する確率を最大にすることと同等である。

この問題では今までの応募者の中で価値の最大値をもつ応募者のみが n 人中のベストになり得るので、今までの価値の最大値をもつ応募者を以下、候補者とよぶことにする。

この問題の最適停止規則を求めるため、詳しく問題を以下のように表す。逐次に現れる n 人の応募者の価値 $X_1, X_2, ..., X_n$ は既知の共通な一様分布 $U(0, 1)$ にしたがう独立同一確率変数とする。ここで、n 人中の最大の価値をもつ応募者を採用する確率を最大にする最適停止規則を求める。

完全情報最良選択問題を OLA（one-stage look-ahead）停止規則を用いて解く。ここで、OLA 停止規則とは、もう1期継続してから停止するより今停止するほうがよい場合には停止することを要求する停止規則である。

$n-j$ 番目の応募者の値が $X_{n-j}=x$ であり、今までに面接した $n-j$ 人中最も値が高いとき、この応募者を採用したときにその人が n 人中のベストである確率 y_{n-j} は

$$
\begin{aligned}
y_{n-j} &= P(X_{n-j}=\max(X_1, \ldots, X_n) \mid X_{n-j}=\max(X_1, ..., X_n)=x) \\
&= P(X_{n-j}>X_{n-j+1}, \ldots, X_{n-j}>X_n \mid X_{n-j}=x) \\
&= \prod_{k=1}^{j} P(x>X_{n-j+k}) \\
&= \prod_{k=1}^{j} \int_0^x dy = x^j
\end{aligned}
$$

$n-j$ 番目の応募者が今までの最大の値 $X_{n-j}=x$ をもち、この人を面接して最適にふるまってベストを得

る確率を $V_{n-j}(x)$ とする。このとき、今までの最大の値 $X_{n-j}=x$ をもつ応募者を採用せず、これ以降最適にふるまってベストを得る最大確率は

$$\sum_{k=1}^{j} P(n-j \text{ 番目に到着した候補者以降、初めての候補者が } n-j+k \text{ 番目に現れる} \mid n-j \text{ 番目が候補者}) EV_{n-j+k}(X_{n-j+k})$$
$$= \sum_{k=1}^{j} \int_0^x \cdots \int_0^x \int_x^1 V_{n-j+k}(y)dy$$
$$= \sum_{k=1}^{j} x^{k-1} \int_x^1 V_{n-j+k}(y)dy$$

と表される。よって最適方程式は

$$V_{n-j}(x) = \max\left\{x^j, \sum_{k=1}^{j} x^{k-1}\int_x^1 V_{n-j+k}(y)dy\right\},\ j=n-1, \ldots, 1,$$

$V_n(x)=1$ となる。$n-j$ 番目の応募者が値 x をもつ候補者である状態を $(n-j, x)$ で表すとき、OLA 停止領域 B は

$$B = \left\{(n-j, x) : x^j \geq \sum_{k=1}^{j} x^{k-1}\int_x^1 y^{n-j+k}dy\right\}$$
$$= \{(n-j, x) : G_j(x) \leq 1\}$$

で与えられる。ここで、$G_j(x)$ は

$$G_j(x) \equiv \sum_{k=1}^{j} \frac{1}{k}(x^{-k}-1)$$

とし、OLA 停止領域を特徴づける関数である（以後、これを OLA 関数とよぶ）。このとき、方程式 $G_1(x)=1$ について簡単に唯一解 $s_1=1/2$ が求まる。$G_j(x)=1$ は x について減少関数であり、j について増加関数、すなわち $n-j$ について減少関数であるので、x についての方程式 $G_j(x)=1$ は区間 $(s_{j-1}, 1)$ の間に唯一解 s_j をもつ。よって、

$$\frac{1}{2} = s_1 < s_2 < \cdots < s_n < \cdots < 1$$

であり、B は

$$B = \{(n-j, x) : 1 \geq G_j(x)\} = \{(n-j, x) : x \geq s_j\}$$

と書きなおすことができる。また、$k=1, \ldots, j$、$y \geq x$ に対して

$$P((n-j+k, y) \in B \mid (n-j, x) \in B) = P(G_{j-k}(y) \leq 1 \mid G_j(x) \leq 1) = 1$$

が成り立つので、B は 'closed'（閉じている）ということがわかり、OLA 停止規則が最適停止規則となるための十分条件を満たす。

したがって、最適停止規則は次のように求まる：残り j 人目の値が s_j 以上であり、その値がこれまでの最大値であったならば、すぐに停止しなさい。また、添え字の付け方を $i=n-j$ と書きかえると、最適停止規則は i 人目の値が s_{n-i} 以上であり、その値がこれまでの最大値であったならば、すぐに停止しなさい、と考えられる。

3. 変化点が既知の場合の最適停止規則

前節で述べた完全情報最良問題の拡張として、選択肢の価値が観測可能である問題において選択肢の価値分布が途中で変化する場合を考える。価値分布の変化に関して以下を仮定する。その変化点は既知とする。

- 選択肢の価値分布は時刻 N_1（既知）で変化する。
- 選択肢の総数は N_2（既知）とする。
- $i=1, 2, ..., N_1$ のとき、選択肢の価値分布は一様分布 $U(0, 1)$ にしたがう。
- $i=N_1+1, ..., N_2$ のとき、選択肢の価値分布は一様分布 $U(0, b)$、$0<b<1$ にしたがう。

まず、定式化のために以下の記号を定義する。(i, x) を i 番目の選択肢が到着していて、その価値が x で、これまでの最大値である状態、$U_i(x)$ を状態 (i, x) で選択肢を選んで成功する確率とする。状態は、次の 3 つの場合に分けて考えられる。

1. $i \leq N_1$、$x \geq b$ のとき
2. $i \leq N_1$、$x < b$ のとき
3. $i > N_1$、$x < b$ のとき

それぞれの場合について、OLA（one-step look-ahead）停止規則を求め、最適停止規則を考える。まず、準備として状態 (i, x) で選択肢を選んで成功する確率 $U_i(x)$ の計算をする。$U_i(x)$ はそれぞれの場合について以下のように表し、計算される。

1. $i \leq N_1$、$x \geq b$ のとき
$$U_i^1(x) = x^{N_1-i}$$
2. $i \leq N_1$、$x < b$ のとき
$$U_i^2(x) = x^{N_1-i}\left(\frac{x}{b}\right)^{N_2-N_1}$$
3. $i > N_1$、$x < b$ のとき
$$U_i^3(x) = \left(\frac{x}{b}\right)^{N_2-i}$$

続いて、第 1 のケース $i \leq N_1$、$x \geq b$ のときを考える。これは N_1 で終了する完全情報最良選択問題の場合と一致する。状態 (i, x) で停止したときの成功確率は
$$U_i^1(x) = x^{N_1-i}$$
次に初めて到着する最大値をもつ選択肢を採用して成功する確率は
$$\sum_{k=1}^{N_1-i} x^{k-1}\int_x^1 U_{i+k}^1(y)dy$$
であり、OLA 関数は
$$H_i^1(x) \equiv U_i^1(x) - \sum_{k=1}^{N_1-i} x^{k-1}\int_x^1 U_{i+k}^1(y)dy = x^{N_1-i}\left\{1-\sum_{j=1}^{N-i}\frac{1}{j}(x^{-j}-1)\right\} \tag{1}$$
で表される。また、OLA 停止領域 B_1 は

$$B_1=\left\{(i,x):x^{N_1-i}\left\{1-\sum_{k=1}^{N-i}\frac{1}{k}(x^{-k}-1)\right\}\geq 0\right\}$$
$$=\left\{(i,x):\sum_{k=1}^{N-i}\frac{1}{k}(x^{-k}-1)\leq 1\right\} \tag{2}$$

とかかれる。ここで、あらためて OLA 関数を

$$G_i^1(x)\equiv\sum_{k=1}^{N-i}\frac{1}{k}(x^{-k}-1)$$

とおくと、$G_i^1(x)$ は i、x についてそれぞれ減少関数である。よって B_1 は閉じている。

次に、第 2 のケース $i\leq N_1$、$x<b$ のときを考える。上で計算したように (i,x) で停止したときの成功確率は

$$U_i^2(x)=x^{N_1-i}\left(\frac{x}{b}\right)^{N_2-N_1}$$

であった。i 番目以降、初めて到着する最大値をもつ選択肢を採用して成功する確率は

$$\sum_{k=1}^{N_1-i}x^{k-1}\left(\int_b^1 U_{i+k}^1(y)dy+\int_x^b U_{i+k}^2(y)dy\right)$$
$$+\sum_{k=N_1-i+1}^{N_2-i}x^{N_1-i}\left(\frac{x}{b}\right)^{i+k-N_1-1}\int_x^b U_{i+k}^3(y)dy$$

であり、OLA 関数は

$$H_i^2(x)\equiv x^{N_1-i}\left(\frac{x}{b}\right)^{N_2-N_1}\left[1-\sum_{k=N_1-i+1}^{N_2-i}\frac{b}{k}\left\{\left(\frac{x}{b}\right)^{-k}-1\right\}\right]$$
$$-x^{N_1-i}\sum_{k=1}^{N-i}\frac{b}{k+N_2-N_1}x^{-k}\left(1-\left(\frac{x}{b}\right)^{k+N_2-N_1}\right)-x^{N_1-i}\sum_{k=1}^{N-i}\frac{1}{k}x^{-k}(1-b^k) \tag{3}$$

と計算される。OLA 停止領域 B_2 は

$$B_2=\left\{(i,x):\left(\frac{x}{b}\right)^{N_2-N_1}\left[1-\sum_{k=N_1-i+1}^{N_2-i}\frac{b}{k}\left\{\left(\frac{x}{b}\right)^{-k}-1\right\}\right]\right.$$
$$\left.-\sum_{k=1}^{N-i}\frac{b}{k+N_2-N_1}x^{-k}\left(1-\left(\frac{x}{b}\right)^{k+N_2-N_1}\right)-\sum_{k=1}^{N-i}\frac{1}{k}x^{-k}(1-b^k)\geq 0\right\} \tag{4}$$

と表される。あらたに書き直した OLA 関数を

$$G_i^2(x)\equiv\left(\frac{x}{b}\right)^{N_2-N_1}\left[1-\sum_{k=N_1-i+1}^{N_2-i}\frac{b}{k}\left\{\left(\frac{x}{b}\right)^{-k}-1\right\}\right]$$
$$-\sum_{k=1}^{N-i}\frac{b}{k+N_2-N_1}x^{-k}\left(1-\left(\frac{x}{b}\right)^{k+N_2-N_1}\right)-\sum_{k=1}^{N-i}\frac{1}{k}x^{-k}(1-b^k)$$

とおく。ここで、$G_i^2(x)$ は i、x についてそれぞれ増加関数であることが示され、よって B_2 は閉じている。

最後に $i>N_1$、$x<b$ のときを考える。残りの選択肢数で考えたときの N_2 で終了する完全情報最良選択問題と一致する。このとき、価値 X_i の分布は $U(0,b)$、$0<b<1$ であるから、すでに計算したように状態 (i,x) で停止した時の成功確率は

$$U_i^3(x)=\left(\frac{x}{b}\right)^{N_2-i}$$

状態（i, x）で停止せずに次に初めて到着する最大値をもつ選択肢を採用して成功する確率は

$$\sum_{k=1}^{N_2-i}\left(\frac{x}{b}\right)^{k-1}\int_x^b U_{i+k}^3(y)dy$$

で計算される。OLA 関数は

$$\begin{aligned}H_i^3(x)&\equiv U_i^3(x)-\sum_{k=1}^{N_2-i}\left(\frac{x}{b}\right)^{k-1}\int_x^b U_{i+k}^3(y)dy\\&=\left(\frac{x}{b}\right)^{N_2-i}\left[1-\sum_{k=N_1-i+1}^{N_2-i}\frac{b}{k}\left\{\left(\frac{x}{b}\right)^{-k}-1\right\}\right]\end{aligned}\tag{5}$$

となる。そこで、OLA 停止領域は

$$\begin{aligned}B_3&=\left\{(i,x):\left(\frac{x}{b}\right)^{N_2-i}\left[1-\sum_{k=N_1-i+1}^{N_2-i}\frac{b}{k}\left\{\left(\frac{x}{b}\right)^{-k}-1\right\}\right]\geq 0\right\}\\&=\left\{(i,x):\sum_{k=N_1-i+1}^{N_2-i}\frac{1}{k}\left\{\left(\frac{x}{b}\right)^{-k}-1\right\}\leq\frac{1}{b}\right\}\end{aligned}\tag{6}$$

と書かれ、

$$G_i^3(x)\equiv\sum_{k=N_1-i+1}^{N_2-i}\frac{1}{k}\left\{\left(\frac{x}{b}\right)^{-k}-1\right\}$$

とおくと、$G_i^3(x)$ は i、x について減少関数であることがわかる。よって、B_3 は閉じている。

以上で求めた各ケースにおける OLA 停止領域が問題の全体において閉じていることを確認する。まず、価値 $x=b$ での境界で停止領域が一致することを示す。はじめに求めた OLA 関数 $H_i^1(x)$、$H_i^2(x)$ において $H_i^1(b)=H_i^2(b)$ が成り立つことから示される。次に、時刻 $i=N_1$ での境界で停止領域が一致することを示す。価値の境界と同様にはじめの OLA 関数 $H_i^2(x)$ と $H_i^3(x)$ に対して $H_{N_1}^2(x)=H_{N_1}^3(x)$ が成立し、$i=N_1$ において OLA 停止領域 B_2 と B_3 は一致する。

さらに、$1\leq i\leq N_1$ における時刻の経過に伴うケースの変化に関しては第2のケースから第1のケースへ移る場合が起こり得る。そこで、$x<b$ のとき $H_i^2(x)\geq 0$ ならば、$b\leq y<1$ のとき $H_i^1(y)\geq 0$ を示す。上で示してきたとおり、$H_i^1(x)$、$H_i^2(x)$ はともに x についての増加関数であり、また $x=b$ での一致から $H_i^2(x)<H_i^2(b)=H_i^1(b)<H_i^1(y)$ が成り立つ。よって、示された。したがって、時点の経過に伴い、ケースが移った場合に対しても停止領域は閉じている。

これまで示してきたことより、この問題の最適停止規則は以下のようにまとめられる。

定理 1　選択肢の総数が N_2 であり、選択肢の価値分布が既知の時刻 N_1 で変化する完全情報最良選択問題に対して、価値分布が一様分布 $U(0, 1)$ から、時刻 N_1+1 以降において一様分布 $U(0, b)$、$0<b<1$ へ変化する場合、最適停止規則は以下の不等式を満たす最初の状態（i, x）で停止する、である。

1. $i \leq N_1$、$x \geq b$ のとき $\sum_{k=1}^{N-i} \frac{1}{k}(x^{-k}-1) \leq 1$
2. $i \leq N_1$、$x < b$ のとき

$$\left(\frac{x}{b}\right)^{N_2-N_1}\left[1-\sum_{k=N_1-i+1}^{N_2-i}\frac{b}{k}\left\{\left(\frac{x}{b}\right)^{-k}-1\right\}\right]$$
$$-\sum_{k=1}^{N-i}\frac{b}{k+N_2-N_1}x^{-k}\left(1-\left(\frac{x}{b}\right)^{k+N_2-N_1}\right)-\sum_{k=1}^{N-i}\frac{1}{k}x^{-k}(1-b^k) \geq 0$$

3. $i > N_1$、$x < b$ のとき $\sum_{k=N_1-i+1}^{N_2-i}\frac{1}{k}\left\{\left(\frac{x}{b}\right)^{-k}-1\right\} \leq \frac{1}{b}$

4. おわりに

3 節で本問題に対する最適停止規則を求めたが、パラメータ N_1、N_2、b の値のいくつかの組について停止規則における実際の閾値を数値計算によって求めることや総数 N_2 に対する変化時点 N_1 の割合を一定にしたまま総数 N_2 を無限に大きくした場合の閾値の極限を求めることなどが挙げられる。

また、今回の問題設定においては既知のある時点での価値分布の変化が一様分布 $U(0, 1)$ から一様分布 $U(0, b)$、$0 < b < 1$ であり、経済状態が縮小する状態にあたるものであった。今後の課題として、まずこの分布の変化が一様分布 $U(0, 1)$ から一様分布 $U(0, b)$、$b > 1$ となる場合、すなわち経済が拡大する問題を考えることができる。この場合は問題の場合分けが少し簡単になり、今回の結果と同様に最適停止規則閾値規則になることが予想される。本稿の問題とは異なり、時刻の経過に対して最良が現れる確率は明らかに減るわけではないので、総数 N_2 に対する変化時点 N_1 の割合や b の値によって、閾値、つまり停止の意思決定を行う時刻がどの程度影響を受けるかという点が興味深く、パラメータのいくつかの値に対して数値計算を行うことが必要と考えられる。

さらに、変化の時点が未知であり、変化点を探索しながら、停止の意思決定をしていく問題はより現実の状況に合致した問題になると考えられる。ジョブサーチを初めとする労働市場における意思決定問題としてさらなる問題設定の改良、あるいは資産売却問題などの現実的問題への適用として価値分布の変化を伴った確率モデルの作成および最適停止規則の導出なども関心の高い話題である。

参考文献

[1] Bruss, F. T. (1987), "On an optimal selection problem by Cowan and Zabczyk," *J. Appl. Prob.,* **24**, 918-928.

[2] Chow, Y. S. Robbins, H. and Siegmund, D. O., (1971), *Theory of Optimal Stopping,* Houghton Mifflin.

[3] Cowan, R. and Zabczyk, J. (1978), "An optimal selection problem associated with the Poisson process," *Theory Prob. Appl.,* **23**, 584-592.

[4] Ferguson, T. S. (1989), "Who solved the secretary problem?" *Statist. Sci.,* **4**, 282-289.

[5] Ferguson, T. S., Hardwick, J. P. and Tamaki, M. (1993), "Maximizing the duration of owning a relatively best object," *Contemporary Mathematics,* Bruss, F. T., Ferguson, T. S. and Samuels, S. M. eds., **125**, 37-58.

[6] Gilbert, J. and Mosteller, F. (1966), "Recognizing the maximum of a sequence," *J. Amer. Statist. Assoc.,* **61**, 35-73.

［7］ Kurushima, A. and Ano, K.（2003a）, “A Poisson arrival selection problem for gamma prior intensity with natural number parameter,” *Sci. Math. Japonicae,* **57**, 217-231.

［8］ Kurushima, A. and Ano, K.（2003b）, “A note on the full-information Poisson arrival selection problem,” *J. Appl. Prob.,* **40**, 1147-1154.

［9］ Porosinski, Z.（1987）, “The full-information best choice problem with a random numberof observations,” *Stoch. Processes Appl.*, **24**, 293-307.

［10］ Presman, E. L. and Sonin, I. M.（1972）, “The best choice problem for a random number of objects,” *Theory Prob. Appl.,* **17**, 657-668.

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

［12］ Sakaguchi, M.（1989）, “Some infinite problems in classical secretary problems,” *Math. Japo.,* **34**, 307-318.

［13］ Szajowski, K. and Tamaki, M.,（2006）, “Shelf life of candidates in the generalized secretary problem,” Working paper.

［14］ Tamaki, M., Pearce, E. M., and Szajowski, K.（1998）, “Multiple choice problem related to duration of the secretary problem,” *RIMS Kokyuroku,* **1068**, 75-86.

［15］ 穴太克則（2000）, タイミングの数理，朝倉書店 .