2001年度日本オペレーションズ・リサーチ学会
春季研究発表会

予約期間も決め得る予約可能な最適停止問題

02202090　　愛知大学　齋藤 毅　SAITO Tsuyoshi

1. 過去の研究と本研究のモデル

筆者はこれまで，最適停止問題に「予約」という概念を導入した場合の最適行動の変化に関心を持ち，①一度予約したオファーは常にリコール可能 [1]，②予約は一定期間のみ有効 [2]，③探索終了後の残存期間に価値を認める [3]，の3モデルを研究してきた．しかしいずれのモデルにおいても，意思決定者はオファーを予約するかどうかを決めることはできても，その期間を決めることはできないと仮定していた．そこで今回の研究においてはこの仮定を外し，予約期間も決定項目とするモデルを考える．

意思決定者は今から t 期間内にオファーを1つ採択せねばならない．オファーの価値分布は既知の F，価値は最低 $a \geq 0$，最高 $b < \infty$，平均 μ とする．オファーは時点 (期間と期間の境界点) ごとに1つ現れ得るが，オファーを現すためには探索費用 $s > 0$ をその都度投資しておく必要がある．現れたオファーをパスしたり採択したりできるのはもちろんのこと，本研究では k 期間以内の任意の期間の予約を認める．すなわちオファー w(価値 w のオファー) に対し，その価値 w と予約期間 $l(\leq k)$ に応じた予約費用 $r(w,l)$ を投ずれば，それ以降の l 期間内ならばいつでもそのオファー w をリコールすることができる．なお，予約しなかったオファーはその後リコールできず，予約はしてもリコールはしなかったオファーに投じた予約費用は返還されず，予約費用はオファー価値と予約期間の増加関数であるとする．

彼の目的は今後の総割引期待利益の最大化であり，各時点において，その時点で現れたオファーを**カレントオファー** (current offer)，その時点で予約が有効なオファーの中で最大価値を持つものを**リーディングオファー** (leading offer) と定義する．すると最終時点を除く各時点において，彼が下すべき決定は次の4つのいずれかである : 1. カレントオファーの採択 (AS)，2. カレントオファーの予約 (RC)，3. リーディングオファーの採択 (PS)，4. 単に次のオファーを求める (PC)．なお，各決定の記号は Accept, Reserve, Pass-up, Stop, Continue からつけたものである．決定 RC を採る場合，**予約期間** (reserving duration) も決定せねばならない．また，最終時点において採り得るのは決定 AS 又は PS のみである．

このモデルにおける最適決定ルールの特徴の1つは「リーディングオファーは明日もまだリコールできるという状況ならば，今日はリコールすべきではない」というものである．これは予約期間を選べないモデル [2] の結論と同じである．

2. 最適方程式

現時点から i 期前に現れたオファーを x_i と表すと，予約可能期間は最大で k 期間のため，予約が有効たり得るオファーの系列は $\boldsymbol{x} = (x_1, x_2, \cdots, x_k)$ と表せ，これを**履歴ベクトル**と呼ぶ (この他のオファーは確実にリコール不可能である)．オファー x_k は現時点で予約が有効としても必ず満期を迎えているため，$\boldsymbol{y} = (x_1, x_2, \cdots, x_{k-1})$ が次時点でも予約が有効たり得るオファーの系列である．したがってカレントオファーを w とすると，次時点における履歴ベクトルは $(w, \boldsymbol{y}) = (w, x_1, x_2, \cdots, x_{k-1})$ となる．

オファー x_i の予約期間の残存分を d_i と表し，$\boldsymbol{d} = (d_1, d_2, \cdots, d_k)$ を**期間ベクトル**と呼び，$\boldsymbol{c} = (d_1, d_2, \cdots, d_{k-1})$ とする．現時点でオファー x_i の予約があと d_i 期間有効ならば，次時点における予約の残存期数は $d_i - 1$ となるため，カレントオファーを l 期間予約したとすると，次時点における期間ベクトルは $(l, \boldsymbol{c}) - \mathbb{1} = (l-1, d_1 - 1, d_2 - 1, \cdots, d_{k-1} - 1)$ となる．決定 PC を採った場合には $l = 0$ と考えればよい．また $d_i < 0$ の場合，$|d_i|$ には意味がなく，単にオファー x_i はリコール不可能であるという事実を示すに過ぎない．

ここで δ_i を，$d_i \geq 0$ ならば $\delta_i = 1$，$d_i < 0$ ならば $\delta_i = 0$ と定義すると，リーディングオファー $\hat{x}$ は $\hat{x} = \max\{\delta_1 x_1,\ \delta_2 x_2,\ \cdots,\ \delta_k x_k\}$ と表すことができる．

さて，現在は時点 $t \geq 0$，履歴ベクトルは $\boldsymbol{x}$，期間ベクトルは $\boldsymbol{d}$，カレントオファーは w とし，以降を最適決定ルールに従った場合の最大総割引期待利益を $u_t((w,\boldsymbol{x});\boldsymbol{d})$ とすると，$u_t((w,\boldsymbol{x});\boldsymbol{d})$ は

$$u_t((w,\boldsymbol{x});\boldsymbol{d}) = \max \left\{ \begin{array}{lll} \mathrm{AS} & : & w, \\ \mathrm{RC} & : & -r(w,\eta^*) - s + \beta v_{t-1}((w,\boldsymbol{y});(\eta^*,\boldsymbol{c})-\boldsymbol{1}), \\ \mathrm{PS} & : & \hat{x}, \\ \mathrm{PC} & : & -s + \beta v_{t-1}((w,\boldsymbol{y});(0,\boldsymbol{c})-\boldsymbol{1}) \end{array} \right\}, \qquad t \geq 0, \tag{2.1}$$

と表すことができる．ここで，

$$v_t(\boldsymbol{x};\boldsymbol{d}) = \int_a^b u_t((w,\boldsymbol{x});\boldsymbol{d})\mathrm{d}F(w), \quad t \geq 0; \qquad v_{-1}(\boldsymbol{x};\boldsymbol{d}) = -\infty. \tag{2.2}$$

$$\eta^* = \arg\max_{l \in K}\{-r(w,l) - s + \beta v_{t-1}((w,\boldsymbol{y});(l,\boldsymbol{c})-\boldsymbol{1})\}. \tag{2.3}$$

3. 最適決定ルールと主な特徴

採り得る 4 決定 (AS，RC，PS，PC) のいずれが最適決定であるかを決める際，(1) カレントオファーをパスするか否かを決め，(2-1) パスしない場合にはそれを採択して探索を終了するか予約して続行するかを決め，(2-2) パスする場合にはリーディングオファーをリコールして探索を終了するかリコールせずに続行するかを決める，という 2 段階を踏む方法が考えられる．与えられた $\boldsymbol{x}$ と $\boldsymbol{d}$ の下で，パスすべきではないカレントオファーの集合を $W_t(\boldsymbol{x},\boldsymbol{d})$，オファー x_i をリコールすることとしないことが無差別となる x_i の値を $\theta_t^i(\boldsymbol{x}_i,\boldsymbol{d})$，オファー w を採択することと予約することが無差別となる w の値を $\lambda_t(\boldsymbol{y},\boldsymbol{d})$ とすると，時点 t における最適決定ルールは次のように表すことができる：

a) $w \in W_t(\boldsymbol{x},\boldsymbol{d})$ のとき，

1. $\lambda_t(\boldsymbol{y},\boldsymbol{d}) < w$ ならば決定 AS (カレントオファー w の採択)．
2. $w \leq \lambda_t(\boldsymbol{y},\boldsymbol{d})$ ならば決定 RC (カレントオファー w を η^* 期間予約)．

b) $w \notin W_t(\boldsymbol{x},\boldsymbol{d})$ のとき，

3. $\theta_t^i(\boldsymbol{x}_i,\boldsymbol{d}) < x_i = \hat{x}$ となる x_i があれば決定 PS (リーディングオファー $\hat{x}$ のリコール)．
4. 全ての x_i で $x_i \leq \theta_t^i(\boldsymbol{x}_i,\boldsymbol{d})$ ならば決定 PC (単なる探索続行)．

このルールの主な特徴を整理すると，

① 「リーディングオファーは，明日もまだリコールできる」という状況ならば，今日はリコールしてはならない．すなわち，リーディングオファーをリコールすることが最適決定となり得るのはリーディングオファーの予約が満期を迎えている時点のみである．

② $-s + \beta\mu \leq a$ ならば，探索開始直後に現れたオファーを採択して探索を終了せよ．

これらは [2] のモデルにおける最適決定ルールの特徴と等しく，[1] と [3] のものともほぼ同一である．すなわち，リーディングオファーの予約が有効な限りはオファー探索を続行すべきであるという特徴は，予約期間の長さやその決定権の有無に関わらず，予約が可能な最適停止問題に備わる基幹的な特徴と言うことができる．

参考文献

[1] Saito, T. Optimal stopping problem with controlled recall. *Probability in the Engineering and Informational Sciences*, 12(1):91–108, 1998.

[2] Saito, T. Optimal stopping problem with finite-period reservation. *European Journal of Operational Research*, 118:605–619, 1999.

[3] Saito, T. Optimal stopping problem with reservation and remaining time value (in Japanese). *The Aichi Journal of Business*, 141:53–74, 2000.