数理解析研究所講究録
978巻 1997年 64-70

On the memory - length - one rule for a secretary problem with uncertain employment

玉置 光司　　愛知大学 経営学部

1. はじめに

よく知られているように、秘書問題の代表的な最適化基準は通常次の二つである。

(1) 成功確率最大化
(2) 期待順位最小化

一般に応募者数nが大きくなると、評価値の漸近解が美しい形になる。応募者に拒否権が無い場合と有る場合（応募者数は順位に関係なく確率1-pでオファーを拒否するものとする）の(1), (2)に対応する結果をまとめると以下のようになる。

成功確率最大化		期待順位最小化	
拒否権が無い場合	$e^{-1} \cong 0.368$	拒否権が無い場合	$\prod_{j=1}^{\infty}\left(1+\frac{2}{j}\right)^{\frac{1}{j+1}} \cong 3.8695$
拒否権が有る場合	$p^{1/(1-p)}$	拒否権が有る場合	?

ここでは、未解決？の箇所、すなはち、期待順位最小化問題で応募者が拒否権を有する場合を考察し、漸近的評価値の１つの上限として

$$\frac{1}{p(1+p)}\left[(1+p)^{2(1+1/p)}-1\right]$$

を求めた(モデル 2)。別の上限も求めたが、それは閉じた形で表わすことができない(モデル 1)。4節で両モデルの数値例を与えた。ここで求めたものはあくまで上限であり、真の値は依然未知である。

他の３つの場合の既知の結果については、Gilbert and Mosteller(1966), Smith(1975), Chow, Robbins, Moriguti and Samuels(1964)を参照されたい。また、秘書問題全般に関するサーベイとしてはFerguson(1989), Samuels(1991)を勧める。

2. Rubin and Samuels' problem

拒否権が無い場合の期待順位最小化問題を少し振り返っておこう。Rubin and Samuels(1977)によれば、期待順位最小化問題のそもそもの関心は、$n\to\infty$の時、最小期待順位が有限の値に留まるかどうかというところにあったという（彼らは、ある定められた値より大きな相対順位の応募者を採用しないという政策の範囲では期待順位を有限にすることはできないと述べている）。期待順位最小化問題の結果を最初に導いたのはChow et al.(1964)であるが、その最適政策を実行するには、現在の応募者の相対順位をいつも正確に認識する必要がある（nが大きいとき、これは大変な作業で多くのmemoryを必要とする）。Rubin and Samuels(1977)はmemory-length-oneの範囲で期待順位を有限にする例を与えた。memory-length-oneとは過去の応募者を（比較の対象として）一人しか記憶しておくことができない場合を意味する。したがって、この場合、現在の応募者は記憶されているものと比較して、良いか悪いかという形で評価される。毎回、３通りの決定Accept（採用する）Reject（パスする）Remember（旧い記憶をすて、現在のものを記憶する）のどれかを選択することとする。政策は選択の列$\{W_r/B_r;\ r=2,3,\ldots,n-1\}$によって記述される。ここで、$W_r$、$B_r$は r-番目の応募者が、既に記憶されている者と比較して、悪い場合、良い場合に対応して選択される決定をあらわす（ただし、最初はRememberで、W_n/B_n=Accept/Accept）。W_r/B_rは本来９通りの決定が可能であるがRubin and Samuels(1977)は３通りの決定W_r/B_r = Reject/Remember, Reject/Accept, Remember/Remember に限定して、最適政策を求め、次のような形になることを示した。

$$W_r/B_r=\left\{\begin{array}{ll}\text{Reject/Remember} & (r<a_n)\\ \text{Reject/Accept} & (a_n\le r<r_n)\\ \text{Remember/Remember} & (r=r_n)\\ W'_{r-r_n}/B'_{r-r_n} & (r>r_n)\end{array}\right\}$$

ここで、$\{W'_i/B'_i;\ i=1,2,\ldots,n-r_n\}$は応募者の総数が$n-r_n+1$の場合に対応する最適政策。このことは、$n\to\infty$の時、

$$0=R_0<A_1<R_1<\ldots<A_k<R_k<\ldots<1$$

となる無限列$\{R_k\}_{k=0}^{\infty}$, $\{A_k\}_{k=1}^{\infty}$ が存在して、区間(R_{k-1}, A_k) では最良のものを記憶し、区間(A_k, R_k)では記憶されているものより良いものが出たら採用するという政策に対応し、R_k, A_kが

$$R_{k+1}=R_k+R_1(1-R_1)^k,\quad A_{k+1}=R_k+\alpha R_1(1-R_1)^k$$

となるように選ばれることを意味する。これらは２数$R_1=\beta$, αで決定されるので、この政策を(α,β)政策と呼ぶ。RubinandSamuelsは上記の問題において、最適値は$\beta=0.456$, $\alpha=0.296$で、このとき、期待順位は7.41375となることを示した(この問題をinfinite problemとして論じた）。

3. Uncertain employment

Rubin and Samuels(1977)のアプローチを拒否権が有る場合に拡張する。拒否権が有る場合の(α, β)政策を次のように定義する。

時間区間$(0, A_1)$では、最初の応募者を記憶し、以後ベターな者が出現したら、順次、記憶を更新する　(したがって、区間$(0, A_1)$の最後には、それまでのベストが記憶されている)。区間(A_1, R_1)では、記　憶されている者よりベターな者が出現したらオファーを与える。オファーが断わられたら、次のベターな者の出現を待って改めてオファーを与える。オファーが受け入れられたら試行終了。区間$(0, R_1)$で試行が終了しなかった場合、以降、同様のパターンを繰り返す。

Rubin and Samuels(1977)と同様にInfinite problem として考える (Infinite problemに関してはGianini and Samuels (1976) 参照)。政策(α, β)のもとでの、停止時刻を T 、選択したものの(絶対) 順位を X とすると、

$$E[X] = E[XI_{\{T\le\beta\}}] + P\{T>\beta\}\frac{E[X]}{1-\beta} \qquad \text{すなはち、} \qquad E[X] = \frac{\bar{\beta}E[XI_{\{T\le\beta\}}]}{\bar{\beta}-P\{T>\beta\}}$$

となる。ただし、$\bar{\beta}=1-\beta$ (α, pにかんしても、以後同様の記法を用いる)。
(A_1, R_1)でのオファーの与え方により、2つのモデルが考えられる。

モデル　1。　区間(A_1, R_1)でオファーが拒否されても、記憶を変更しない場合。

次の確率変数を導入する。

Q: 区間$(0, A_1)$で最終的に記憶された者の、区間$(0, R_1)$における順位。

明らかに、

$$P\{Q=m+1\} = \left(\frac{A_1}{R_1}\right)\left(1-\frac{A_1}{R_1}\right)^m, \qquad m=0,1,\ldots$$

他方、$A_1<t<R_1$に対して、

$$P\{T\le t\mid Q=m+1\} = \sum_{j=1}^{\infty}\left(1-\bar{p}^j\right)\binom{m}{j}\left(\frac{t-A_1}{R_1-A_1}\right)^j\left(1-\frac{t-A_1}{R_1-A_1}\right)^{m-j}$$

$$= 1-\left\{\bar{p}\left(\frac{t-A_1}{R_1-A_1}\right)+\left(1-\frac{t-A_1}{R_1-A_1}\right)\right\}^m.$$

したがって、Qが与えられたときの、Tのp.d.f. $f_T(t\mid Q)$は以下のようになる。

$$f_T(t \mid Q=m+1) = \left(\frac{mp}{R_1 - A_1}\right)\left(\bar{p}\left(\frac{t - A_1}{R_1 - A_1}\right) + \left(1 - \frac{t - A_1}{R_1 - A_1}\right)\right)^{m-1}$$

この密度は時刻T=tでの出現者の順位が1,…,mの誰でも構わない場合であるから、特定の順位k(k=1,...,m)の者がT=tで出現してオファーを受け入れる密度は$f_T(t \mid Q=m+1)/m$。また、このとき、この者の全体での真の順位の期待値は、k/R_1となり、次式を得る。

$$E[XI_{\{T\le R_1\}}] = E[E[XI_{\{T\le R_1\}} \mid Q]]$$

$$= \sum_{m=1}^{\infty} P\{Q=m+1\} \sum_{k=1}^{m} \int_{A_1}^{R_1} \left(\frac{k}{R_1}\right)\frac{1}{m} f_T(t \mid Q=m+1)dt$$

$$= \frac{\alpha}{2\beta}\left[\alpha^{-2} - (p+\bar{p}\alpha)^{-2}\right]$$

また、

$$P\{T>R_1\} = \sum_{m=0}^{\infty} P\{T>R_1 \mid Q=m+1\}P\{Q=m+1\} = \frac{\alpha}{1-\bar{p}\alpha}$$

以上より、次の結果を得る。

補題　1

$$E[X] = \frac{\bar{\alpha}p(2\alpha+\bar{\alpha}p)}{2\alpha(1-\bar{\alpha}p)^2}\frac{\bar{\beta}}{\beta\left(\bar{\beta}-\frac{\alpha}{1-\bar{\alpha}\bar{p}}\right)}$$

最適なα, β は次の関係を満足する。

$$\alpha = \frac{p\bar{\beta}^2}{1-\bar{p}\bar{\beta}^2}$$

$\bar{\beta}$ は次式 $g_1(x)=0$ の根 x.

$$g_1(x) = \bar{p}x^4(x^2-x-1) + px^3 + x^2 + x - 1$$

モデル　2。　区間(A_1, R_1)でオファーが拒否されると、拒否した者が今までの記憶に取って変わる（したがって、いつも、それまでのベストが記憶として保存されている）。

この場合、

$$f_T(t) = \sum_{m=1}^{\infty} \bar{p}^{m-1} p \int_{A_1}^{t} \frac{A_1}{y_1} dy_1 \int_{y_1}^{t} \frac{dy_2}{y_2} \cdots \int_{y_{m-2}}^{t} \frac{dy_{m-1}}{y_{m-1}} \left(\frac{1}{t}\right)^2 = \sum_{m=1}^{\infty} \frac{pA_1}{t^2} \frac{(\bar{p}\log(t/A_1))^{m-1}}{(m-1)!} = \frac{pA_1{}^p}{t^{1+p}}$$

したがって、

$$E[XI_{\{T \le R_1\}}] = \int_{A_1}^{R_1} \frac{1}{t} f_T(t)dt = \left(\frac{p}{1+p}\right)\left(\frac{\alpha^{-1}-\alpha^p}{\beta}\right)$$

$$P\{T>R_1\} = 1 - \int_{A_1}^{R_1} f_T(t)dt = \alpha^p$$

以上より、次の結果を得る。

補題　2

$$E[X] = \left(\frac{p}{1+p}\right)\frac{\left(\alpha^{-1}-\alpha^p\right)\bar{\beta}}{\beta\left(\bar{\beta}-\alpha^p\right)}$$

最適なα, β は次の関係を満足する。

$$\alpha = \left(\bar{\beta}\right)^{2/p}$$

$\bar{\beta}$ は次式 $g_2(x)=0$ の根 x.

$$g_2(x) = px^{2(1+1/p)} - (1+p)x + 1.$$

また、E[T]の 1 つの上限が次のように与えられる。

$$E[X] \le \frac{1}{p(1+p)}\left[(1+p)^{2(1+1/p)} - 1\right].$$

4.　2つのモデルの比較と数値例

下の数値例が示すように、モデル 1 のほうがモデル 2 より、良い政策となっている。

モデル 1

p	E[X]	β	α	$(1-\alpha)/\alpha$

0.1	47.95	0.266	0.105	8.57
0.5	12.39	0.395	0.225	3.45
0.9	7.99	0.447	0.284	2.52

モデル 2

p	E[X]	β	α	$-\log\alpha$
0.1	63.02	0.074	0.214	1.54
0.5	13.60	0.291	0.253	1.38
0.9	8.10	0.430	0.288	1.25

モデル１、２において、区間(A_1, R_1)，すなはち、区間$(\alpha\beta, \beta)$における、オファー対象者の数をそれぞれS, Tで表わす。Sの分布、平均は次のようになる。

$$P\{S = s\} = \alpha(1-\alpha)^s, \quad s=0,1,\ldots$$

$$E[S] = \sum_{s=1}^{\infty} s\alpha(1-\alpha)^s = (1-\alpha)/\alpha .$$

Tの分布は複雑なので省略。平均は次のようになる。

$$E[T] = E[E[T \mid S]] = \sum_{s=1}^{\infty}\left(\sum_{j=1}^{s}\frac{1}{j}\right)\alpha(1-\alpha)^s = -\log\alpha$$

参考文献

Chow,Y.S.,Robbins,H.,Moriguti,H.&Samuels,S.M.(1964) Optimal selection based on relative rank(the "secretary problem"), Israel J.Math. 2, 81-90.

Ferguson,T.S.(1989) Who solved the secretary problem ?, Statistical Science 4, 282-289.

Gianini,J.&Samuels,S.M.(1976) The infinite secretary problem, Ann. Probab. 4, 418-432.

Gilbert,J.& Mosteller,F.(1966) Recognizing the maximum of a sequence, J. Amer. Statiat.

Assoc. 61, 35-73.

Rubin,H&Samuels,S.M.(1977) The infinite-memory secretary problem, Ann.Probab. 5, 627-635.

Samuels,S.M.(1991) Secretary problems. Handbook of Sequential Analysis. 381-405(Chapter 16), B.K.Ghosh and P.K.Sen, eds., Marcel Dekker, Boston.

Smith M.H.(1975) A secretary problem with uncertain employment. J. Appl. Probab. 12, 620-624.