

ある Bilateral Secretary Problem について.

阪大基礎工数理 穴太克則

(Katsunori Ano)

§1. はじめに.

最適停止問題のひとつに有名な Secretary Problem がある。その game 版とでもいうべき Model が最近 10 年ぐらいの間に研究されてきている。[cf. Roth, Kadane & DeGroot ('77), DeGroot & Kadane ('83), Sakaguchi ('84).]

Secretary Problem: ある意思決定者が Secretary を 1 人雇用したいとする。 n 人の girls が Secretary の position に対して、每期 1 人ずつ応募する。 girls の到着のし方は同様に等しい、すなわち、到着の順列は $n!$ 通り。意思決定者の知り得る情報はすでに応募してきた girls の相対的ランクのみとする。彼は何らかの主観的基準により応募してきた girls にランク付けができるとする。意思決定者は n 人の girls のなかで真に最も良いランク (Best と呼ぶとする) の人を採用する確率を最大にする最適停止政策を発見したい。採用回数は 1 回のみで recall はなし。

最適政策は、『 $i^* = \min\{1 \leq i \leq n \mid \sum_{j=i}^{n-1} \frac{1}{j} \leq 1\}$ なる i^* が存在し、 i^*-1 までの girls はすべて pass し、 i^* 以後最初に到着したキャンディット (相対ランク 1 の girl, G と書く) を採用し』というものである。

Bilateral Secretary Problem (B.S.P.): 意思決定者が 2 人, player I と II がいる。Secretary problem と同様に 2 人の前に n 人の girls (以下, objects と呼ぶ) が逐次に到着するとする。player I と II の知り得る情報は自分の目前にすでに到着し観測した objects の相対ランクのみとし, player I と II に共通とする。先ず, player I が object を accept するか reject するかの option をもつ。player I が reject したときのみ player II がその object に対する option をもつ。recall は許されない。一方の player が accept したならば, 他の player は accept することができない。"win" という事象を Best を accept するとする。"lose" という事象を win でないとする。利得は,

I	II	I	II
win	lose	+1	-1
lose	win	-1	+1
reject-reject で終了		0	0

で与えられる 0 和ゲームとする。

このとき, 両 player は期待利得を最大にしたい。最適政策は?

以上が基本となる Model であり, ここでは次の 3 タイプの B.S.P を考える。最適政策は一体どのようなものになり, どう違ってくるのかを研究する。

§2. Non-Zero Sum 版.

§3. Best か 2nd Best を accept できたとき "win".

§4. object が player の accept の申し出を拒否する確率をもつ.

§2. Non-Zero Sum 版 B.S.P.

両 player の利得は各々無関係に,

$$\left\{ \begin{array}{l} \text{player I が win するとき } +1, \text{ lose するとき } -1. \\ \text{player II が win するとき } +1, \text{ lose するとき } -1. \\ \text{両 player がすべての object を reject したとき } 0, 0. \end{array} \right.$$

I : II		I	II
win:		+1	0
lose:		-1	0
	win	0	+1
	lose	0	-1
rej.	rej.	0	0

$V(i), W(i)$ ≡ それぞれ player I, II が今までに accept して いず、
 i 番目の object に直面して いて、かつそれが C_i であり
 以後最適にふるまったときの最大期待利得.

最適性の原理から次の最適方程式が成立する.

$$(2.1) \quad V(i) = \max \left\{ 2\frac{i}{n} - 1, p_i \sum_{j=i+1}^n \frac{i}{j(j-1)} V(j) \right\}, \quad 1 \leq i \leq n-1, V(n) = 1.$$

$$(2.2) \quad W(i) = \max \left\{ 2\frac{i}{n} - 1, g_i \sum_{j=i+1}^n \frac{i}{j(j-1)} W(j) \right\}, \quad 1 \leq i \leq n-1, W(n) = 1.$$

$$(2.3) \quad p_i = \begin{cases} 1 \\ 0 \end{cases} \quad \text{if } W(i) \begin{cases} > \\ \leq \end{cases} 2\frac{i}{n} - 1, \quad 1 \leq i \leq n-1, p_n = 0.$$

$$(2.4) \quad g_i = \begin{cases} 1 \\ 0 \end{cases} \quad \text{if } V(i) \begin{cases} > \\ \leq \end{cases} 2\frac{i}{n} - 1, \quad 1 \leq i \leq n-1, g_n = 0.$$

p_i, g_i はそれぞれ player II, I が i 番目で C_i に直面して いるとき
 に reject する確率である.

定理 1: $i^* = \min \{ 1 \leq i \leq n \mid i \geq \frac{n}{2} \} = [\frac{n}{2}]$ とする. 最適 play のも
 とで,

player I の最適政策は, i^*-1 まですべて pass し, i^* 以後の最初の G を accept する. player II の最適政策は, i^*-1 まですべて pass し, i^* 以後の最初の G を accept する. 両 player が最適政策にしたがったときの期待利得は,

$$\text{player I: } V(i) = \frac{i^*-1}{n} \left(\sum_{j=i^*}^n \frac{2}{j-1} + 1 \right) - 1, \quad \text{player II: } W(i) = 0.$$

$$\lim_{n \rightarrow \infty} \frac{i^*}{n} = \alpha \text{ とすると, } \alpha = 0.5000, \quad \lim_{n \rightarrow \infty} V(i) = -2\alpha \log \alpha + \alpha - 1 \doteq 0.1931.$$

$$\lim_{n \rightarrow \infty} W(i) = 0.$$

(略証) $V(i), W(i)$ の解折は, g_i, p_i の解折に等しいことに注目して backward に g_i, p_i の値を調べる.

$$V(i) = \begin{cases} \frac{i^*-1}{n} \left(\sum_{j=i^*}^n \frac{2}{j-1} + 1 \right) - 1, & 1 \leq i \leq i^*-1 \quad (g_i=1) \\ 2\frac{i}{n}-1, & i^* \leq i \leq n \quad (g_i=0) \end{cases} \quad W(i) = \begin{cases} 0, & 1 \leq i \leq i^*-1 \quad (p_i=1) \\ 2\frac{i}{n}-1, & i^* \leq i \leq n \quad (p_i=0). \end{cases}$$

§3. Best か 2nd Best を accept したとき "win" である B.S.P.

win: Best か 2nd Best を accept した事象. lose: win でない事象 ($\text{rej} - \text{rej}$ は ~~除く~~)

利得は,

I \ II	I	II
win	lose	+1 -1
lose	win	-1 +1
rej.	rej.	0 0

ゼロ和ゲーム.

G_1, G_2 : 相対ランク 1, 2 の object.

$V_k(i) \equiv i$ 番目で $G_k (k=1,2)$ に直面して i までこれまでに accept しないで、以後最適にふるまった時の player I の期待利得.

i 番目で G_1, G_2 に直面して i まで、それを accept したときの player I の期待利得はそれぞれ $2g_{i-1}, 2h_{i-1}$. ここで g_i, h_i はそれぞれ i 番目で G_1, G_2 を accept したとき、それが Best か 2nd Best である

確率, すなわち, $g_i \equiv \frac{i}{n} + \frac{i(n-i)}{n(n-1)} = \frac{i(2n-i-1)}{n(n-1)}$, $h_i \equiv \frac{i(i-1)}{n(n-1)}$.

最適性の原理より次の最適方程式が成立する.

$$(3.1) \quad V_1(i) = \max\{2g_{i-1}, \min\{1-2g_i, u_i\}\}, \quad 2 \leq i \leq n-1, \quad V_1(n)=1,$$

$$V_1(1) = \max\{2g_1-1, \min\{1-2g_1, \frac{1}{2}[V_1(2)+V_2(2)]\}\},$$

$$(3.2) \quad V_2(i) = \max\{2h_{i-1}, \min\{1-2h_i, u_i\}\}, \quad 2 \leq i \leq n-1, \quad V_2(n)=1,$$

$$\text{ここで } u_i \equiv \sum_{j=i+1}^n \pi_{ij}[V_1(j)+V_2(j)] \quad \& \quad \pi_{ij} \equiv \frac{i(i-1)}{n(j-i)(j-2)}.$$

右辺 $\max\{\dots\}$ の中のオ1項は accept, オ2項は reject したときの player I の期待利得に対応している. 定理2を導くために使う補題を準備する.

補題1: $V_1(i) \geq 0$ for $1 \leq i \leq n$, $V_2(i) \geq 0$ for $2 \leq i \leq n$, & $u_i \geq 0$ for $2 \leq i \leq n-1$.

(証明) 略. Inductionによる.

定理2: 最適 play のもとで両 player の最適政策を特徴付ける four numbers $1 \leq d_2 \leq d_1 \leq \beta_2 \leq \beta_1 \leq n$ が存在し, 最適政策は,

player I: d_1-1 まですべて pass し, d_1 以後最初の C_1 を accept, もしくは, β_1 以後の最初の C_1 か C_2 を accept する.

player II: d_2-1 まですべて pass し, d_2 以後最初の C_1 を accept, もしくは, β_2 以後の最初の C_1 か C_2 を accept する.

$$\hat{\alpha}_1 = \lim_{n \rightarrow \infty} \frac{d_1}{n} = 1 - \frac{\sqrt{2}}{2} \doteq 0.2929, \quad \hat{\beta}_1 = \lim_{n \rightarrow \infty} \frac{\beta_1}{n} = \frac{\sqrt{2}}{2} \doteq 0.7071, \quad \hat{\beta}_2 = \lim_{n \rightarrow \infty} \frac{\beta_2}{n} \doteq 0.5564,$$

$$\hat{\alpha}_2 = \lim_{n \rightarrow \infty} \frac{d_2}{n} \doteq 0.2030, \quad \hat{\beta}_2, \hat{\alpha}_2 \text{ はそれぞれ } 4 \log x + \frac{9}{x} - \frac{1}{x^2} = 3 + 4 \log \hat{\beta}_1,$$

$$4 + 4 \hat{\alpha}_1 - 8 \log \hat{\alpha}_1 - 2/\hat{\alpha}_1 + 3/\hat{\beta}_2 - 4/\hat{\beta}_2 + 4 \log \hat{\beta}_2 = 4(x - \log x) \text{ の } (0,1) \text{ 間の単一}$$

$$\text{根. player I の期待利得は, } \lim_{n \rightarrow \infty} V_1(1) = 2\hat{\alpha}_2 - 4\hat{\alpha}_2 + 1 \doteq 0.2704.$$

(略証) backward に $V_1(i), V_2(i)$ の差分方程式を解く.

$$d_2 = \min\{i \mid 1-2g_i \leq u_i\}, \quad d_1 = \min\{i \mid 2g_i - 1 \geq 0\}, \quad \beta_1 = \min\{i \mid 2h_i - 1 \geq 0\},$$

$$\beta_2 = \min\{i \mid 1-2h_i \leq \sum_{j=i+1}^{\beta_1-1} \frac{2(n-j)}{(j-1)(j-2)} + \sum_{j=\beta_1}^n \frac{(n-1)(2j-n)}{j(j-1)(j-2)}\},$$

$$V_1(i) = \begin{cases} u_i & , 1 \leq i \leq d_2-1 \\ 1-2g_i & , d_2 \leq i \leq d_1-1 \\ 2g_i-1 & , d_1 \leq i \leq n \end{cases}$$

$$V_2(i) = \begin{cases} u_i & , 2 \leq i \leq \beta_2-1 \\ 1-2h_i & , \beta_2 \leq i \leq \beta_1-1 \\ 2h_i-1 & , \beta_1 \leq i \leq n \end{cases}$$

$$f(x) = \lim_{n \rightarrow \infty} V_1\left(\frac{i}{n}\right), \quad g(x) = \lim_{n \rightarrow \infty} V_2\left(\frac{i}{n}\right) \text{ とする,}$$

$$f(x) = \begin{cases} 2\hat{\alpha}_2^2 - 4\hat{\alpha}_2 + 1, & 0 < x \leq \hat{\alpha}_2 \\ 2x^2 - 4x + 1, & \hat{\alpha}_2 \leq x \leq \hat{\alpha}_1 \\ -2x^2 + 4x - 1, & \hat{\alpha}_1 \leq x < 1 \end{cases}$$

$$g(x) = \begin{cases} 2\hat{\alpha}_2^2 - 4\hat{\alpha}_2 + 1, & 0 < x \leq \hat{\alpha}_2 \\ -2x^2 + 4x \log x + K_2 x + 1, & \hat{\alpha}_2 \leq x \leq \hat{\alpha}_1 \\ 2x^2 - 4x \log x + K_1 x - 1, & \hat{\alpha}_1 \leq x \leq \hat{\beta}_2 \\ 1-2x^2, & \hat{\beta}_2 \leq x \leq \hat{\beta}_1 \\ 2x^2 - 1, & \hat{\beta}_1 \leq x < 1 \end{cases}$$

ここで K_1, K_2 はある定数.

□

数値例 (ばうすい)

n	α_2	α_1	β_2	β_1	$V_1(1)$
10	2	3	6	8	0.2803
20	4	6	12	15	0.2786
30	6	9	17	22	0.2762
100	21	30	56	71	0.2719

n	α_2	α_1	β_2	β_1	$V_1(1)$
200	41	59	112	142	0.2712
300	61	88	167	213	0.2710
500	102	147	279	354	0.2708
1000	203	293	557	708	0.2704

S4. object が player の accept の申し出を拒否する確率をもつ B.S.P.

win: Best を accept した事象.

利得構造は S1 の B.S.P と同じであるが 0 和 game. 各 object は two player の accept の申し出を, 確率 g ($0 \leq g \leq 1, p+g=1$) で拒否できる.

$V(i) \equiv$ i 番目に C_i に直屬しているとき以後最適にふるまったときの player I の期待利得.

次の最適方程式が成立.

$$(4.1) \quad V(i) = \max \begin{cases} p(2\frac{i}{n}-1) + g \min \left\{ p(1-2\frac{i}{n}) + g \sum_{j=i+1}^n \pi_{ij} V(j), \sum_{j=i+1}^n \pi_{ij} V(j) \right\}, \\ \min \left\{ p(1-2\frac{i}{n}) + g \sum_{j=i+1}^n \pi_{ij} V(j), \sum_{j=i+1}^n \pi_{ij} V(j) \right\}, \end{cases} \quad 1 \leq i \leq n-1$$

ここで, $\pi_{ij} \equiv \frac{i}{j(j-1)}$. $V(n) = p^2$.

asymptotic な性質に興味を移して, $x = \frac{i}{n}$, $n \rightarrow \infty$ とすると (4.1) は (4.2) となる. $f(x) = \lim_{n \rightarrow \infty} V(\frac{i}{n})$,

$$(4.2) \quad f(x) = \max \begin{cases} p(2x-1) + g \min \left\{ p(1-2x) + g \int_x^1 xy^{-2} f(y) dy, \int_x^1 xy^{-2} f(y) dy \right\}, \\ \min \left\{ p(1-2x) + g \int_x^1 xy^{-2} f(y) dy, \int_x^1 xy^{-2} f(y) dy \right\}, \end{cases} \quad 0 < x < 1$$

$f(1) = p^2$.

定理3: 最適 play のもとで, 両 player の最適政策を特徴付ける $0 < \alpha < \beta < 1$ が存在し, 最適政策は,

player I: β まですべて pass し, β 以後最初の C_i を accept する.

player II: α まですべて pass し, α 以後最初の C_i を accept する.

このときの player I の期待利得は, $K_1 \alpha^p + 1$, ここで α, β, K_1 は (4.5) (4.3) (4.4) で与えられる. 又, $p=1$ とすると, $\alpha \doteq 0.412$, $\beta = 0.5000$ I の期待利得 $\doteq 0.1756$.

(略証) (4.2) の微分方程式を解けば,

$$f(x) = \begin{cases} K_1 \alpha^p + 1 & , \quad 0 < x \leq \alpha \\ K_1 x^p + 1 & , \quad \alpha \leq x \leq \beta \\ \frac{p}{1+g} ((2-g^2)x^{1-g^2} - 1) & , \quad \beta \leq x < 1. \end{cases}$$

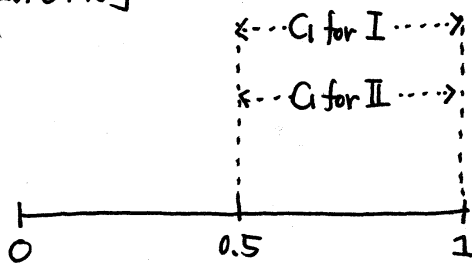
(4.3) β は $(4-2p)x - p(2-\delta^2)x^{1+\delta^2} = 2\delta$ の $(0,1)$ 間の unique root.

$$(4.4) \quad K_1 \equiv \frac{1}{1+\delta} [p(2-\delta^2)\beta^{p\delta} - 2\beta^p], \quad K_2 \equiv \frac{1}{1+\delta} p(2-\delta^2).$$

(4.5) α は, $\frac{K_1}{p-1}(x\beta^{p-1} - x^p) + \frac{K_2}{\delta^2}(x\beta^{-\delta^2} - x) + \frac{p}{1+\delta}(x - \frac{x}{\beta}) + 2x = 0$ の $(0,1)$ 間の unique root. $\square$

§5. 最適政策の比較

[§2のモデル]

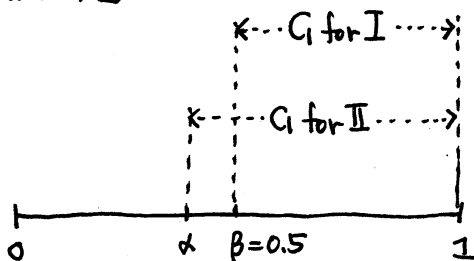


$n \rightarrow \infty$

Expected Payoff of I $\doteq 0.1931$
(E.P.I)

(記号) G for I ... player I はこの region で
最初に出た G を accept.

[§4のモデル]



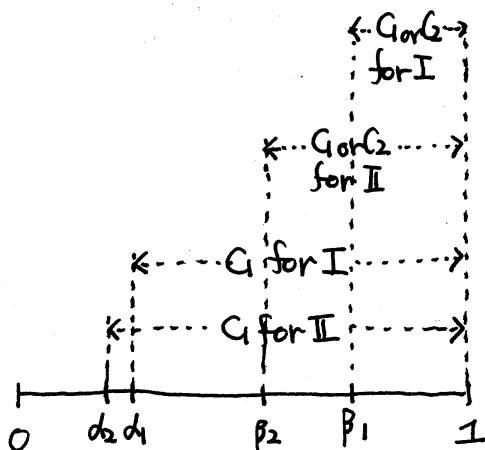
$n \rightarrow \infty$, 特に $p=1.0$ のとき

$$\alpha \doteq 0.4120$$

$$\beta \doteq 0.5000$$

$$\text{E.P.I} \doteq 0.1756.$$

[§3のモデル]



$n \rightarrow \infty$

$$\alpha_2 \doteq 0.2704$$

$$\alpha_1 \doteq 1 - \frac{\sqrt{2}}{2} \doteq 0.2929$$

$$\beta_2 \doteq 0.5564$$

$$\beta_1 = 1 - \alpha_1 \doteq 0.7071$$

$$\text{E.P.I.} \doteq 0.2704.$$

References

- [1] Roth, R., Kadane, J. B. and DeGroot, M. H., "Optimal peremptory challenges in trials by juries: A bilateral sequential process," *O.R.*, 25, 901-919 (1977).
- [2] DeGroot, M. H. and Kadane, J. B. "Optimal sequential decisions in problems involving more than one decision maker," in M. H. Rizvi, J. S. Rustagi. and D. Siegmund (ed.) Recent Advance in Statistics, Academic Press, London, 197-210 (1983).
- [3] Sakaguchi, M. "Bilateral sequential games related to the no-information secretary problem," *Mathematica Japonica*, 29, 961-973 (1984).
- [4] Enns, E. G. and Ferenstein, E. "The horse game," *J.O.R.S.J.*, 28, 51-62 (1985).
- [5] Smith, M. H. "A secretary problem with uncertain employment," *J.A.P.*, 12, 620-624 (1975).