数理解析研究所講究録
680巻 1989年 199-207

Switching cost をもつ 2-parameter optimal stopping problem について

九州大学 理学部 田中輝雄 (Teruo TANAKA)

§1 序

2-parameter optimal stopping problemは、２つのシステム（AとB）を対象にし、その２つのシステムを切り替えながら進むとしたとき、各ステップでの各々のシステムの状態から決まる期待利得を最大にする停止規則を見つけるというものである。つまり、2-armed bandit problem を一般化したもので、時間を2次元 $T=\mathbb{N}^2, \mathbb{R}^2_+$にした確率過程 $\{X_z, z\in T\}$ を考える。例えば、次のような過程を考えればよい：

$$z=(s,t),\ X_{(s,t)}=\sum_{r=0}^{s} A(r)+\sum_{r=0}^{t} B(r) \qquad (1)$$

$$X_{(s,t)}=f(A(s),B(t)) \qquad (2)$$

但し、$A=\{A(r), r=0,1,2,\cdots\}$,$B=\{B(r), r=0,1,2,\cdots\}$ は、実数値確率過程

f は、$\mathbb{R}^2$上の関数

(1)については、次のように解釈する。A,Bという２つのシステムがあり、次の様な進み方をするとする：

(i) 初期状態 $X_{(0,0)}=A(0)+B(0)$

(ii) Aの方に１ステップ進む $X_{(1,0)}=\sum_{r=0}^{1} A(r)+B(0)$

(iii) Bの方に１ステップ進む $X_{(1,1)}=\sum_{r=0}^{1} A(r)+\sum_{r=0}^{1} B(r)$

(iv) Bの方に１ステップ進む $X_{(1,2)}=\sum_{r=0}^{1} A(r)+\sum_{r=0}^{2} B(r)$

(v) 以下、上の様な手順で進む

ここで注意することは、システムAを動かすときは、システムBは停止している（逆についても同様）ということである。つまり、各ステップで、一方のシステムを動かし、そして、必ず１ステップ動かすということである。

2-parameter optimal stopping problem を考えるときは、上の例で述べたようなシステムの選択の仕方（戦略）を考える必要があり、《switch》ということが重要になる。

ここでは、各ステップで得られるrunning rewardとは別に、システムをAからBに（BからAに）切り替えたときにのみ加わるcost(switching cost)を新に考え、期待利得を最大にする戦略を見つける問

題について論じる。簡単の為、離散時間についてのみ述べることにする。§２では問題の定式化、§３では、一般の過程についてdynamic programming equationを示し、最適解を明らかにする。§４、§５では、Markov chain（2-parameter stochastic processの場合は、bi-Markov chain と言う）の場合について述べる。

§２　定義と定式化

$(\Omega, \mathcal{B}, P)$: complete probablity space

$\{\mathcal{B}_z, z \in T\}$: complete filtration　　$T = \mathbb{N}^2 \cup \{\infty\}$

random variable $\tau : \Omega \to T$: Markov point

$\overset{\text{def}}{\Leftrightarrow}$ $\{\tau \leqq z\} \in \mathcal{B}_z$　　$\forall z \in T$

$\{\sigma_t, t=0,1,2,\cdots\}$: strategy starting at z

$\overset{\text{def}}{\Leftrightarrow}$ (1) σ_t : Markov point　　$\forall t$

(2) $\sigma_{t+1} = \sigma_t + e_i$　for some i　但し　$e_1=(1,0)$ $e_2=(0,1)$

(3) $\sigma_0 = z$

$(\{\sigma_t\}, \tau)$: tactic starting at z

$\overset{\text{def}}{\Leftrightarrow}$ (1) $\{\sigma_t\}$: strategy starting at z

(2) $\tau : \Omega \to \mathbb{N} \cup \{\infty\}$: stopping time w.r.t. $\{\mathcal{G}_t, t=0,1,\cdots\}$

但し、$\mathcal{G}_t = \{A \in \mathcal{B} : A \cap \{\sigma_t \leqq z\} \in \mathcal{B}_z,\ \forall z\}$

注意　$\sigma_{t+1} : \mathcal{G}_t$-measurable

$\Sigma = \{$tactic starting at 0　全体$\}$

$\Sigma z = \{$tactic starting at z　全体$\}$

$\{X_z, z \in T\}$: $(\mathcal{B}_z)$-adapted uniformly bounded process

(running reward)

$\{Y_z, z \in T\}$: $(\mathcal{B}_z)$-adapted uniformly bounded process

(terminal reward)

switching cost

$C: \{e_i, i=1,2\}^2 \to \mathbb{R}$

$C(e_k, e_j) = 0 \quad k=j$

$\neq 0 \quad k \neq j$

$\Theta_t = \sigma_{t+1} - \sigma_t$ for strategy $\{\sigma_t\}$

PROBLEM

$\pi = (\{\sigma_t\}, \tau) \in \Sigma$

$$J(\pi) = E\{\sum_{t=0}^{\tau-1} \{a^t X(\sigma_t) + a^t C(\Theta_{t-1}, \Theta_t)\} + a^\tau Y(\sigma_\tau)\}$$

Find $\xi \in \Sigma$ such that $J(\xi) = \sup_\pi J(\pi)$

但し、$0<a<1$:discount factor

§3　一般の過程について

$z>0$,

$$U_z^i = \operatorname*{esssup}_{\pi \in \Sigma_z} E\left[\sum_{t=0}^{\tau-1} a^t X(\sigma_t) + a^\tau Y(\sigma_\tau) + C(e_i, \Theta_0) + \sum_{t=1}^{\tau-1} a^t C(\Theta_{t-1}, \Theta_t) \mid \mathcal{B}_z\right]$$

とおくと、次の定理が言える。

定理　3-1

$i=1,2$　$z>0$　に対して、

$$U_z^i = \max\left\{Y_z, \max_{k=1,2}\left\{E\left[X_z + C(e_i, e_k) + aU_{z+e_k}^k \mid \mathcal{B}_z\right]\right\}\right\}$$

$z=0, ne_i, n\in\mathbb{N}$ に対しては、switching cost を除いたもので定義する。つまり、

$$U_0 = \operatorname*{esssup}_{\pi \in \Sigma} E\left[\sum_{t=0}^{\tau-1} a^t X(\sigma_t) + a^\tau Y(\sigma_\tau) + \sum_{t=1}^{\tau-1} a^t C(\Theta_{t-1}, \Theta_t) \mid \mathcal{B}_0\right]$$

$$U_{ne_i} = U_z^i$$

とおくと、同様に次の事が言える。

定理　3-2

$$U_0 = \max\left\{Y_0, \max_{i=1,2}\left\{E\left[X_0 + aU_{e_i} \mid \mathcal{B}_0\right]\right\}\right\}$$

$$U_{ne_i} = \max\left\{Y_{ne_i}, \max_{k=1,2}\left\{E\left[X_{ne_i} + C(e_i, e_k) + aU_{ne_i+e_k}^k \mid \mathcal{B}_0\right]\right\}\right\}$$

定義　3-3

$\pi=(\{\sigma_t\},\tau)\in\Sigma$: admissible tactic

$\overset{\text{def}}{\Leftrightarrow}$(1)　$U_0=E\,[X_0+aU_{\sigma_1}^{\delta(\theta_0)}\mid\mathcal{B}_0]$

(2)　$U_{\sigma_t}^{\delta(\theta_{t-1})}=E\,[X_{\sigma_t}+C\,(\Theta_{t-1},\Theta_t)+aU_{\sigma_{t+1}}^{\delta(\theta_t)}\mid\mathcal{B}_{\sigma_t}]$　　on $\{t<\tau\}$

(3)　$U_{\sigma_\tau}^{\delta(\theta_{\tau-1})}=Y_{\sigma_\tau}$　　但し、$\delta(e_i)=i$

定理3-4

admissible tactic は最適である。

定理3-5

$$\Phi_z^i=\begin{cases}U_0 & z=0\\ U_z^i & z\neq 0\end{cases}$$

とおくと、過程 $\{\Phi_z^i,z\}$ i=1,2 は、次の条件を満たす有界な過程のうちで、最小の過程である。

(1)　$\Psi_z^i\geqq Y_z$　, $\forall z$

(2)　$\Psi_z^i\geqq\max_k E\,[X_z+C\,(e_i,e_k)+a\Psi_{z+e_k}^k\mid\mathcal{B}_z]$　　$z\neq 0$

(3)　$\Psi_0^i\geqq\max_k E\,[X_0+a\Psi_{e_k}^k\mid\mathcal{B}_0]$

§4 bi-Markov chain の構成について

$i=1,2$

$(\Omega^i, \mathcal{A}^i, \mathcal{A}^i_n, X^i_n, P^i_x)$: homogeneous Markov chain with state space $(E^i, \mathcal{E}^i)$

T^i_n : transition operator $\quad T^i_n f(x) = E^i_x f(X^i_n)$

この二つの、Markov chain に対して、次のように直積（tensor product）によって構成する。

$\Omega = \Omega^1 \otimes \Omega^2 \quad \mathcal{B} = \mathcal{A}^1 \otimes \mathcal{A}^2 \quad E = E^1 \otimes E^2 \quad \mathcal{E} = \mathcal{E}^1 \otimes \mathcal{E}^2$

$\mathcal{B}_{(s,t)} = \mathcal{A}^1_s \otimes \mathcal{A}^2_t$

$P_x = P^1_{x_1} \otimes P^2_{x_2} \quad x = (x_1, x_2) \in E$

$X_{(s,t)}(\omega) = (X^1_s(\omega_1), X^2_t(\omega_2)) \quad \omega = (\omega_1, \omega_2) \in \Omega$

$T_{(s,t)} = T^1_s \otimes T^2_t$

このとき、次の性質が言える：

(1) $\forall (z_1, z_2), s \geqq 0$ に対して、$E_x[f(X_{z+(s,0)}) \mid \mathcal{B}_z] = E^1_{x_1}[f(X_{z+(s,0)}) \mid \mathcal{A}^1_{z_1}]$

第二成分についても同様。

(2) $\forall (z_1, z_2), s \geqq 0$ に対して、$E_x[f(X_{z+(s,0)}) \mid \mathcal{B}_z] = E_{X_z}[f(X_{(s,0)})]$

第二成分についても同様。

(3) Markov property

$\forall z, w \in \mathbb{N}^2$,

$E_x[f(X_{z+w}) \mid \mathcal{B}_z] = E_{X_z}[f(X_w)]$

(4) strong Markov property

$\forall T$: Markov point, $z \in \mathbb{N}^2$ に対して、

$$E_x[f(X_{T+w}) \mid \mathcal{B}_z] = E_{X_T}[f(X_w)]$$

定義 4-1

上で構成した組 $(\Omega, \mathcal{B}, \mathcal{B}_z, X_z, P_x)$ を、$(E, \mathcal{E})$ をstate space にもつ、bi-Markov chain と言う。

§5 bi-Markov chain の場合について

f, g : bounded $\mathcal{E}$-mesurable function on E

Problem

$x \in E$, $\pi = (\{\sigma_t\}, \tau) \in \Sigma$

$$E_x\{\sum_{t=0}^{\tau-1} \{a^t f(X(\sigma_t)) + a^t C(\Theta_{t-1}, \Theta_t)\} + a^\tau g(X(\sigma_\tau))\} \rightarrow \mathrm{Max}$$

を考える。

(classcal)optimal stopping problem の場合のように、状態空間Eを分割する。

まず、E上の関数列 $\{V_n\}_n$, $\{W_n\}_n$ を次のように定義する：

$$V_0 = W_0 = g$$

$$V_{n+1} = \max\{V_n, aT_1^1 V_n + f, aT_1^2 W_n + f + C(e_1, e_2)\}$$

$$W_{n+1} = \max\{W_n, aT_1^2 W_n + f, aT_1^1 V_n + f + C(e_2, e_1)\}$$

このとき、明らかに、

$$V_{n+1} = \max\{g, aT_1^1 V_n + f, aT_1^2 W_n + f + C(e_1, e_2)\}$$

$$W_{n+1} = \max\{g, aT_1^2 W_n + f, aT_1^1 V_n + f + C(e_2, e_1)\}$$

また、$V_{n+1} \geqq V_n$ ，$W_{n+1} \geqq W_n$

故に、$V = \lim_{n\to\infty} V_n$　　$W = \lim_{n\to\infty} W_n$　が存在し、次を満たす：

$$V = \max\{g, aT_1^1 V + f, aT_1^2 W + f + C(e_1, e_2)\} \qquad (1)$$

$$W = \max\{g, aT_1^2 W + f, aT_1^1 V + f + C(e_2, e_1)\} \qquad (2)$$

この(1),(2)を基にして、次のように状態空間を分割する：

$\Gamma_V = \{V = g\}$　　$\Gamma_W = \{W = g\}$　…………Stopping domain

$\Gamma_1^1 = \{V = aT_1^1 V + f\} \setminus \Gamma_V$　　$\Gamma_2^2 = \{W = aT_1^2 W + f\} \setminus \Gamma_W$　……Continuing domain

$\Gamma_2^1 = \{V = aT_1^2 W + f + C(e_1, e_2)\} \setminus (\Gamma_V \cup \Gamma_1^1)$

………Switching domain

$\Gamma_1^2 = \{W = aT_1^2 V + f + C(e_2, e_1)\} \setminus (\Gamma_W \cup \Gamma_2^2)$

この分割を基にして、次の様にtacticを定める：

$\sigma_0 = 0$ ，　$\sigma_1 = e_j$　if $X_0 \in \Lambda_j$

$$\sigma_{t+1} = \begin{cases} \sigma_t + e_j & \text{if } \sigma_t \in \Gamma_j^1, \ \Theta_{t-1} = e_1 \\ \sigma_t + e_j & \text{if } \sigma_t \in \Gamma_j^2, \ \Theta_{t-1} = e_2 \end{cases}$$

$$\tau = \begin{cases} 0 & \text{if } X_0 \in \Lambda \\ \inf\{t \geqq 1 : X_{\sigma_t} \in \Gamma_V, \ \Theta_{t-1} = e_1\} \wedge \inf\{t \geqq 1 : X_{\sigma_t} \in \Gamma_W, \ \Theta_{t-1} = e_2\} & \text{if } X_0 \in \Lambda \end{cases}$$

但し、$U=\max\{g, aT_1^1V+f, aT_2^2W+f\}$

$\Lambda=\{U=g\}$

$\Lambda_1=\{U=aT_1^1V+f\}\setminus\Lambda$, $\Lambda_2=\{U=aT_2^2W+f\}\setminus(\Lambda\cup\Lambda_1)$

命題 5-1

$\{U(X_0), V(X_z)\quad z\neq0\}$ と、$\{U(X_0), W(X_z)\quad z\neq0\}$ は、定理3-5の条件を満たす最小の過程である。

命題 5-2

上で構成したtacticは、admissible である。

定理 5-3

上で構成したtacticは、最適である。

注意　上で構成したV,Wは、次の条件を満たす最小の関数である：

$$x\geqq g \qquad\qquad y\geqq g$$

$$x\geqq aT_1^1x+f \qquad\qquad y\geqq aT_1^2y+f \qquad \cdots\cdots(a, f)\text{-i-excessive}$$

$$x\geqq aT_1^2\,y+f+C(e_1, e_2) \qquad\qquad y\geqq aT_1^1\,x+f+C(e_2, e_1)$$