数理解析研究所講究録
第 452 巻 1982 年 20-40

M/G/1 待ち行列における停止規則の選択を含めた最適制御問題

高知大 理学部　　大坪 義夫

§1. 序

M/G/1 待ち行列システムにおける最適制御問題は、多数の人々によって研究されてきている (cf. Crabill－Gross－Magazine [2])。その中で特に、Mitchell [8] と Doshi [3] 等は、サービス率を制御する問題を、Prabhu [11] は、最適停止問題を取り扱った。しかしこれまでは、これらの問題は各々別個に議論がなされてきた。この報告では、サービス率を制御する政策と選択可能な停止規則を対として扱う制御問題について述べる。このような政策と停止規則を対として考える問題は、次のような観点から起った。サービス率を制御する最適化問題において、ある時刻にサービス機関の営業を停止せざるをえなかったり、または、停止することが有益である場合がある。このような問題は、仮の吸収状態ともとの状態空間の和集合を新しい状態空間とすることによって、

新たな制御問題として考えることもできる。しかし、このように定義された制御問題では、もとの制御問題に関して、確率1で停止するとは限らない。したがって、確率1で停止する制御問題に制限するならば、上で考えた新しい制御問題は不適当である。

一般のマルコフ過程において、政策と停止規則を対として考える最適化問題は、Hordijk [5], Furukawa-Iwamoto [4], Krylov [7], Nisio [9] 等によって研究されている。

この報告では、評価関数として、期待総利得関数と平均型期待利得関数を考えるが、ともに、利得率、ジャンプ利得、終端利得によって定められる。ジャンプに伴う利得を加味する問題は、Bieber [1] で取り扱われている。

ここでの主な目的は、次の3点である；

(1) 最適期待総利得関数を逐次近似の極限として与え、それの性質を調べる。(§3)

(2) 政策と停止規則のある対が、最適であるための必要条件と十分条件を、生成作用素を含む形で与える(§4)

(3) 平均型評価関数に対して、政策と停止規則のある対が最適であるための十分条件を与える。(§5)

§2. 問題の定式化

M/G/1 待ち行列システムにおける最適制御問題は、次の7つの組 $(S, A, r, \rho, g, \lambda, H)$ で表される；

(i) state space S は、仕事量（残余分）x の集合で、$S \equiv [0,\infty)$。$Z \equiv \mathbb{R}^+ \times S$，但し、$\mathbb{R}^+ \equiv [0,\infty)$ は time space。

(ii) action space A は、$A \equiv [a_1, a_2]$，$0 \leqq a_1 < a_2 < \infty$。

(iii) reward rate r は、$Z \times A$ 上の有界連続関数。

(iv) jump reward fct ρ は、$Z \times S$ 上の有界連続関数。

(v) terminal reward fct g は、Z 上の有界連続関数。

(vi) jump rate λ；客の到着間分布は、parameter $\lambda : Z \to \mathbb{R}^+$ をもつ指数分布であり、λ は次をみたす；

(a) λ は Z 上の連続関数

(b) $\exists M < \infty$ s.t. $0 < \lambda(z) < M$ for $\forall z \in Z$

(vii) H；順次到着する客のサービス量（仕事量）は、互いに独立、かつ、到着間分布に独立な確率変数で、同一分布 $H(\cdot)$ をもつ。

定義 2.1. $f : Z \to \mathbb{R}$ が次をみたすとき、right-left continuous on Z という；各 $(s,x) \in Z$ に対して、

$$\lim_{\substack{(t,y)\to(s,x)\\ t \geqq s,\ y \leqq x}} f(t,y) = f(s,x) \quad 。$$

定義2.2.　policy $\pi : Z \to A$ は、次をみたすものとする；

(a)　π は、Z 上で right-left continuous、

(b)　π は、Z の任意の compact subset 上で 2変数関数の意味で有限個の不連続点をもつ、

(c)　各 $s \geq 0$ に対して、$\pi(s, \cdot)$ は、S 上で単調非減少かまたは、連続区間において Lipschitz 連続である。

(注)　条件 (c) は path の一意性のために加えたものである。

policy π の全体を Π と表す。

各々の $\pi \in \Pi$ と $(s, x) \in Z$ に対して、微分方程式

$$\begin{cases} \dfrac{d}{dt} x(t) = -\pi(t, x(t)), & t \geq s, \\ (s) = x \end{cases}$$

の解が存在して、一意である。これより、policy $\pi \in \Pi$ を用いて、初期状態が (s, x) のとき、時間 $[s, s+t]$ の間に、客が到着しなければ、時刻 $t : s \leq t \leq s+t$ の状態は、

$$y(t) = \max(x(t), 0)$$

である。必要ならば、この $y(t)$ を $y^{\pi}_{(s,x)}(t)$ とかく。

policy $\pi \in \Pi$ に対応する制御過程 $\{Z^{\pi}_t\}_{t \geq 0} \equiv \{(t, X^{\pi}_t)\}_{t \geq 0}$ を、次の確率法則で定める；

$$P^{\pi}_{(s,x)}[\,\nu(s,x) \leqq t\,] = 1 - \exp\left(-\int_s^{s+t} \lambda(w, y(w))\,dw\right),$$

$$P^{\pi}_{(s,x)}[\,y(s+t) \leqq X^{\pi}_{s+t} \leqq y(s+t)+v \mid \nu(s,x) = t\,] = H(v),$$

$$\text{for } (s,x) \in Z,\ t \geqq 0;\ v \geqq 0,$$

ここで、$y(w) = y^{\pi}_{(s,x)}(w)$, $w \geqq s$, であり、$\nu(s,x)$ は (s,x) を start したときに、次の客の到着するまでの経過時間を表す。

<u>定理 2.1.</u>　各 $\pi \in \Pi$ に対して、

(i) $\{Z^{\pi}_t\}_{t \geqq 0}$ の a.a. の sample paths は、左極限をもち、右連続で、かつ有限区間で有限個の不連続点をもつ

(ii) $\{Z^{\pi}_t\}_{t \geqq 0}$ は、strong Markov process である。

（証明）

(i) は、$\{Z^{\pi}_t\}_{t \geqq 0}$ の定義から明らかである。

(ii) Ito [6] の §2.4 における strong Markov 性に対する十分条件を与えている定理の証明方法により、

「任意の $f \in F$ に対して、もし各 $(s,x) \in Z$ に対して

$$\lim_{t \downarrow 0} f(Z^{\pi}_{s+t}) = f(s,x) \quad P^{\pi}_{(s,x)}\text{-a.s.},$$

であるならば、各 $(s,x) \in Z$ と任意の $t > 0$ に対して

$$\lim_{t \downarrow 0} T^{\pi}_t f(Z^{\pi}_{s+t}) = f(s,x) \quad P^{\pi}_{(s,x)}\text{-a.s.}」$$

が成立することを示せばよい。ここで、F は、Z 上の有界可測な関数の全体であり、

$$T^{\pi}_t f(s,x) \equiv E^{\pi}_{(s,x)}[\,f(Z^{\pi}_{s+t})\,]$$

である。 以下略。

sample space Ω を、$\omega:[0,\infty)\to Z$（可測）の全体とし、$\mathcal{F}_t^s$ を、集合 $\{\omega\in\Omega \mid \omega(s')\in B\}$ $(s\le s'\le t,\ B\in\mathcal{B}(Z))$ で生成された σ-algebra とする。但し、$\mathcal{B}(Z)$ は、Z の Borel-algebra である。

$C_s(\pi)$ を、$P^\pi_{(s,x)}[s\le\tau<\infty]=1$ $(\forall x\in S)$ をみたす $\{\mathcal{F}_t^s\}_{t\ge s}$ に関する stopping time τ の全体とし、

$$\Lambda(s)\equiv\{(\pi,\tau)\mid \pi\in\Pi,\ \tau\in C_s(\pi)\}$$

とおく。

各 $(s,x)\in Z$, $(\pi,\tau)\in\Lambda(s)$ に対して、

$$D^\pi_\tau f(s,x)\equiv E^\pi_{(s,x)}\Big[\int_s^\tau r(Z^\pi_t,\pi(Z^\pi_t))\,dt + \sum_{j=1}^\infty \rho(\tau_j,\bar{X}^\pi_{\tau_j},X^\pi_{\tau_j})\,\chi(\tau_j\le\tau)+f(Z^\pi_\tau)\Big]$$

とおく。ここで、τ_j は、(s,x) を start してからの jth の客の到着時刻を表し、χ は定義関数であり、$\bar{X}^\pi_t\equiv\lim_{w\uparrow t}X^\pi_w$。

このとき、

期待総利得関数を、$\varphi^\pi_\tau(s,x)\equiv D^\pi_\tau g(s,x)$, $\big((s,x)\in Z,\ (\pi,\tau)\in\Lambda(s)\big)$

最適総利得関数を、$u^*(s,x)\equiv\sup_{(\pi,\tau)\in\Lambda(s)}\varphi^\pi_\tau(s,x)$, $\big((s,x)\in Z\big)$

とおく。

定義2.3.

(π^*, τ^*) が optimal であるとは、$(\pi^*, \tau^*) \in \Lambda(s)$ $(\forall s \geqq 0)$ かつ、$u^*(z) = \varphi^{\pi^*}_{\tau^*}(z)$ $(\forall z \in Z)$ のときにいう。

これより以後、$r^\pi(z) \equiv r(z, \pi(z))$, $\rho(X^\pi_{\tau_j}) \equiv \rho(\tau_j, X^\pi_{\tau_j^-}, X^\pi_{\tau_j})$ とかく。

次の仮定を課す；

仮定1. 各 $(s,x) \in Z$ に対して、

$$\sup_{\pi \in \Pi} E^\pi_{(s,x)}\left[\sup_{t \geqq s}\left(\int_s^t r^\pi(Z^\pi_w)\,dw + \sum_{j=1}^\infty \rho(X^\pi_{\tau_j})\chi(\tau_j \leqq t)\right)^+\right] < +\infty,$$

ここで、$b^+ \equiv \max(0, b)$。

仮定2. $(s,x) \in Z$, $(\pi, \tau) \in \Lambda(s)$ に対して、

$$E^\pi_{(s,x)}\left[\int_s^\tau r^\pi(Z^\pi_t)\,dt + \sum_{j=1}^\infty \rho(X^\pi_{\tau_j})\chi(\tau_j \leqq \tau)\right] \neq -\infty$$

$$\Rightarrow$$

$$\liminf_{N \to \infty} \int_{\{\tau > N\}} \left[\int_s^N r^\pi(Z^\pi_t)\,dt + \sum_{j=1}^\infty \rho(X^\pi_{\tau_j})\chi(\tau_j \leqq N)\right]^- dP^\pi_{(s,x)} = 0,$$

ここで、N は自然数、$b^- \equiv \max(0, -b)$。

仮定1，2をともにみたす簡単な例としては、

(i) $r \leqq 0$, $\rho \leqq 0$.

(ii) $r(s,x,a) = e^{-\alpha_1 s}\hat{r}(x,a)$, $\rho(s,x,y) = e^{-\alpha_2 s}\hat{\rho}(x,y)$

($\hat{r}$, $\hat{\rho}$ は有界で、$\alpha_1 > 0$, $\alpha_2 > M$)

§3. 最適総利得と逐次近似

この節では、最適総利得関数を得るための逐次近似法を与える。結果だけ述べるが、証明は Ohtsubo[10] を参照。

整数 $n \geq 0$，自然数 N に対して、$v_N^n(\pi)$ を次で定める；

$$\begin{cases} v_N^n(\pi)(s,x) \equiv \max\{g(s,x),\ D^{\pi}_{k_n(s)} v_N^n(\pi)(s,x)\} \\ \qquad\qquad \text{for } (s,x) \in [0,N) \times S \equiv Z_N, \\ v_N^n(\pi)(s,x) \equiv g(s,x) \qquad \text{for } (s,x) \in Z - Z_N, \end{cases}$$

ここで、$k_n(s) \equiv \min_k \{k \cdot 2^{-n} \mid k \cdot 2^{-n} > s,\ k = 1, 2, \cdots\}$。

$v_N^n(\pi)$ は、"backward-inductively" に求められ、各 n に対して単調である。

$$v_N(\pi)(z) \equiv \lim_{n \to \infty} v_N^n(\pi)(z) \quad , \ z \in Z,$$

とおくと、

定理3.1. 各 $N < \infty$ と任意の $\pi \in \Pi$ に対して、

(i)
$$v_N(\pi)(s,x) = \begin{cases} \sup_{\tau \in C_s^N(\pi)} \varphi_\tau^\pi(s,x), & (s,x) \in Z_N \\ g(s,x), & (s,x) \in Z - Z_N \end{cases}$$

ここで、$C_s^N(\pi) \equiv \{\tau \in C_s(\pi) \mid P^\pi_{(s,x)}(\tau \leq N) = 1 \quad {}^\forall x \in S\}$。

i.e. $v_N(\pi)$ は、有限区間における停止問題の最適利得。

(ii) $\varepsilon > 0$ を任意にとる。このとき、任意の $(s,x) \in Z$ に対し、

$$v_N(\pi)(s,x) \leq \varphi^\pi_{\sigma_N^\varepsilon}(s,x) + \varepsilon,$$

ここで、$\sigma_N^\varepsilon \equiv \inf\{t \geq s \mid v_N(\pi)(Z_t^\pi) \leq g(Z_t^\pi) + \varepsilon\}$。

同様にして、u_N^n を次で定める；

$$\begin{cases} u_N^n(s,x) \equiv \max\{g(s,x),\ \sup_{\pi\in\Pi} D_{t_n(s)}^{\pi} u_N^n(s,x)\} & \text{for } (s,x)\in Z_N \\ u_N^n(s,x) \equiv g(s,x) & \text{for } (s,x)\in Z - Z_N \end{cases}$$

$\{u_N^n\}$ は、各 n, N に対して、単調であるから、

$$u(z) \equiv \lim_{N\to\infty}\lim_{n\to\infty} u_N^n(z), \quad z\in Z,$$

とおくと、

定理3.2.

(i) $u^* = u = \lim_{N\to\infty} \sup_{\pi\in\Pi} v_N(\pi)$ on Z。

(ii) u^* が Z 上で有界かつ、right-left continuous であるとき、

(a) $u^* \geq g$ on Z かつ、u^* は、additively excessive

i.e. 任意の $(s,x)\in Z$，$\pi\in\Pi$ に対して、

$$u^*(s,x) \geq D_t^{\pi} u^*(s,x), \quad t\geq s,$$

$$\liminf_{t\downarrow s} u^*(Z_t^{\pi}) \geq u^*(s,x) \quad P_{(s,x)}^{\pi}\text{-a.s.}。$$

(b) h が additively excessive かつ、$h\geq g$ on Z をみたすならば、$h \geq u^*$ on Z （u^* の最小性）。

(iii) 任意に固定した $s\geq 0$ に対して、$u^*(s,\cdot)$ は S 上で universally measurable な関数である。

§4.　最適性に対する必要条件と十分条件

F を、Z 上の有界 Borel 可測な関数の全体とし、

$$F^\pi \equiv \{ f \in F \mid \lim_{t \downarrow 0} T_t^\pi f(z) = f(z) \quad \forall z \in Z \}, \quad \pi \in \Pi,$$

とおく。ここで、$T_t^\pi f(s,x) = E_{(s,x)}^\pi [f(Z_{s+t}^\pi)]$。

$$\underline{\mathcal{A}^\pi f(z)} \equiv \lim_{t \downarrow 0} \frac{1}{t} [T_t^\pi f(z) - f(z)], \quad \pi \in \Pi, \quad z \in Z,$$

$$\underline{\mathcal{U}^\pi f(z)} \equiv \lim_{t \downarrow 0} \frac{1}{t} [D_{s+t}^\pi f(z) - f(z)], \quad \pi \in \Pi, \quad z = (s,x) \in Z,$$

上の作用素は、右辺が存在して、F^π に属するときに定義し、

$\mathcal{A}^\pi$ の定義域を $\mathcal{D}^\pi (\subset F^\pi)$ とする。

<u>補題 4.1.</u>　$f \in F$ に対して次を仮定する；

(i)　$\frac{\partial}{\partial s} f$, $\frac{\partial^-}{\partial x} f$ が存在して、有界、

(ii)　$\frac{\partial}{\partial s} f$, $\frac{\partial^-}{\partial x} f$ はともに Z 上で right-left continuous。

このとき、$f \in \mathcal{D} \equiv \bigcap_{\pi \in \Pi} \mathcal{D}^\pi$ かつ

$$\mathcal{A}^\pi f(s,x) = \frac{\partial}{\partial s} f(s,x) - \pi(s,x) \frac{\partial^-}{\partial x} f(s,x)\, \delta(x)$$
$$+ \lambda(s,x) \int_0^\infty \big(f(s, x+v) - f(s,x) \big) H(dv),$$

ここで、$\delta(x) = 1$ if $x > 0$, $= 0$ if $x = 0$。

さらに、

$$\mathcal{U}^\pi f(s,x) = \mathcal{A}^\pi f(s,x) + r^\pi(s,x) + \lambda(s,x) \int_0^\infty \rho(s, x, x+v) H(dv).$$

証明は、$\mathcal{A}^\pi$, $\mathcal{U}^\pi$ の定義にしたがって、実際に計算すればよい。

補題4.2. f を、Z 上の additively excessive fct とする。このとき、各 $(s,x)\in Z$, $\pi\in\Pi$ と、$\tau\leqq\sigma$ $(P^{\pi}_{(s,x)}\text{-a.s.})$ となる任意の $\tau,\sigma\in C_s(\pi)$ に対して、

$$D^{\pi}_{\tau} f(s,x) \geqq D^{\pi}_{\sigma} f(s,x) \quad 。$$

証明は、Širjaev [13] の Lemma III.5 を参照。

任意に固定した t_0 $(0<t_0\leqq\infty)$ に対して

$$\xi \equiv s+\min\left(t_0,\ \nu(Z^{\pi}_s)\right), \qquad s\geqq 0,\ \pi\in\Pi,$$

とし、また、

$$\Gamma^* \equiv \{z\in Z \mid u^*(z)=g(z)\},$$

$$\sigma^* \equiv \inf\{t\geqq s \mid Z^{\pi}_t\in\Gamma^*\}, \qquad s\geqq 0,\ \pi\in\Pi,$$

とおく。

定理4.1. （最適性に対する必要条件）

(π^*,τ^*) が optimal であって

(i) $\varphi^{\pi^*}_{\tau^*}$ が上から有界で、Z 上で right-left continuous

(ii) $\frac{\partial}{\partial s}\varphi^{\pi^*}_{\tau^*}$ と $\frac{\partial}{\partial x}\varphi^{\pi^*}_{\tau^*}$ がともに存在して、有界かつ Z 上で、right-left continuous

であるならば、以下が成り立つ；

(a) $\tau^*\geqq\sigma^*$ $\quad P^{\pi^*}_z\text{-a.s.}$ $\quad(\forall z\in Z)$,

(b) (π^*,σ^*) も optimal,

(c) $\varphi^{\pi^*}_{\tau^*}\left(=\varphi^{\pi^*}_{\sigma^*}\right)$ は、次の2組の関係式をみたす；

$$\begin{cases} \max_{\pi\in\Pi} \mathcal{U}^{\pi} \varphi_{\tau^*}^{\pi^*}(z) = 0, & z \notin \Gamma^* \\ \varphi_{\tau^*}^{\pi^*}(z) = g(z), & z \in \Gamma^* \end{cases}$$

$$\begin{cases} \max_{\pi\in\Pi} \mathcal{U}^{\pi} \varphi_{\tau^*}^{\pi^*}(z) \leqq 0, & \\ \varphi_{\tau^*}^{\pi^*}(z) \geqq g(z), & z \in Z, \\ [\varphi_{\tau^*}^{\pi^*}(z) - g(z)][\max_{\pi\in\Pi} \mathcal{U}^{\pi} \varphi_{\tau^*}^{\pi^*}(z)] = 0. & \end{cases}$$

（証明）

(a) 定理3.2の(ii)から、$u^* = \varphi_{\tau^*}^{\pi^*}$ は additively excessive であり、$u^* \geqq g$ on Z であるから、補題4.2より、

$$u^*(z) \geqq D_{\tau^*}^{\pi^*} u^*(z) \geqq D_{\tau^*}^{\pi^*} g(z)。$$

また、(π^*, τ^*) の最適性より、

$$u^*(z) = \varphi_{\tau^*}^{\pi^*}(z) = D_{\tau^*}^{\pi^*} g(z), \quad z \in Z。$$

したがって

$$u^*(z) = D_{\tau^*}^{\pi^*} u^*(z) = D_{\tau^*}^{\pi^*} g(z), \quad z \in Z。$$

明らかに、$u^*(Z_{\tau^*}^{\pi^*}) \geqq g(Z_{\tau^*}^{\pi^*}) \quad P_z^{\pi^*}\text{-a.s.} \quad (z \in Z)$。

しかるに、ある $z_0 \in Z$ に対して、

$$P_{z_0}^{\pi^*}[u^*(Z_{\tau^*}^{\pi^*}) > g(Z_{\tau^*}^{\pi^*})] > 0$$

と仮定すると

$$u^*(z_0) = D_{\tau^*}^{\pi^*} u^*(z_0) > D_{\tau^*}^{\pi^*} g(z_0)。$$

これは矛盾である。したがって、各 $z \in Z$ に対して、

$$u^*(Z_{\tau^*}^{\pi^*}) = g(Z_{\tau^*}^{\pi^*}) \qquad P_z^{\pi^*}\text{-a.s.}。$$

よって、σ^* の定義より、$\tau^* \geqq \sigma^* \quad P_z^{\pi^*}\text{-a.s.} \quad (\forall z \in Z)$。

(b) (a) の結果と、補題 4.2 から、 各 $z \in Z$ に対して、

$$u^*(z) \geq \varphi^{\pi^*}_{\sigma^*}(z) = D^{\pi^*}_{\sigma^*} g(z) = D^{\pi^*}_{\sigma^*} u^*(z) \geq D^{\pi^*}_{\tau^*} u^*(z) = u^*(z).$$

したがって、 $u^*(z) = D^{\pi^*}_{\sigma^*} g(z) = \varphi^{\pi^*}_{\sigma^*}(z)$, $z \in Z$, つまり、(π^*, σ^*) は、optimal である。

(c) (b) より、 $u^* = \varphi^{\pi^*}_{\tau^*} = \varphi^{\pi^*}_{\sigma^*}$ であるから、$\varphi^{\pi^*}_{\sigma^*}$ に対して証明する。 $\varphi^{\pi^*}_{\sigma^*}$ が、additively excessive であることと、補題 4.2 より、 任意の $\pi \in \Pi$ に対して、

$$\begin{aligned}
\varphi^{\pi^*}_{\sigma^*}(s,x) &\geq D^{\pi}_{\xi} \varphi^{\pi^*}_{\sigma^*}(s,x) \\
&= [1 - G^{\pi}(s,x:h)]\left[\int_s^{s+h} r^{\pi}(w, y(w))\,dw + \varphi^{\pi^*}_{\sigma^*}(s+h, y(s+h))\right] \\
&\quad + \int_0^{\infty}\int_0^{h} G^{\pi}(s,x:dw)\left[\int_s^{s+w} r^{\pi}(w', y(w'))\,dw' \right. \\
&\quad \left. + \rho(s+w, y(s+w), y(s+w)+v) + \varphi^{\pi^*}_{\sigma^*}(s+w, y(s+w)+v)\right] dH(v)
\end{aligned}$$

ここで、 $G^{\pi}(s,x;h) = P^{\pi}_{(s,x)}[\nu(s,x) \leq h]$ 。

この式において、 $\pi = \pi^*$ のときは、strong Markov 性により、等式が成立する。したがって、上の両辺に $\frac{1}{h}$ を掛けて、極限操作をして、補題 4.1 を用いれば、得たい結果が証明される。

補題 4.3. (Dynkin の公式)

各 $s \geq 0$ に対して、$(\pi, \tau) \in \Lambda(s)$ とする。このとき、$f \in \mathcal{D}^{\pi}$ かつ、各 $z \in Z$ に対して、 $E^{\pi}_{z}[\tau] < \infty$ であるならば、

各 $(s,x)\in Z$ に対して

$$E^{\pi}_{(s,x)}\left[\int_s^{\tau}\mathcal{U}^{\pi}f(Z^{\pi}_t)\,dt\right]=D^{\pi}_{\tau}f(s,x)-f(s,x)\text{。}$$

定理4.2.（最適性に対する十分条件）

(π^*,τ^*) が存在して、

(i) $(\pi^*,\tau^*)\in\Lambda(s)$ $(\forall s\geqq 0)$,

(ii) $\varphi^{\pi^*}_{\tau^*}\in\mathcal{D}$, かつ,

(iii) $\begin{cases}\varphi^{\pi^*}_{\tau^*}(z)\geqq g(z), & \forall z\in Z,\\ \sup\limits_{\pi\in\Pi}\mathcal{U}^{\pi}\varphi^{\pi^*}_{\tau^*}(z)\leqq 0, & \forall z\notin\Gamma^*\end{cases}$

であるならば、(π^*,τ^*) は optimal である。

（したがって、定理4.1より、(π^*,σ^*) も optimal である。）

（証明）

$z\in\Gamma^*$ に対しては、明らかに

$$u^*(z)\geqq\varphi^{\pi^*}_{\tau^*}(z)\geqq g(z)=u^*(z)$$

であるから、$u^*(z)=\varphi^{\pi^*}_{\tau^*}(z)$ 。

$v_N(\pi)$, σ^{ε}_N を§3で定義されたものとし、

$$\Gamma^{\varepsilon}_N(\pi)\equiv\{z\in Z\mid v_N(\pi)(z)\leqq g(z)+\varepsilon\},\quad\varepsilon>0$$

とおく。明らかに、$u^*(z)\geqq v_N(\pi)(z)\geqq g(z)$ $(\forall\pi\in\Pi,\forall N<\infty,\forall z\in Z)$

であるから，任意の $\varepsilon>0$ に対して、$\Gamma^*\subset\Gamma^{\varepsilon}_N(\pi)$, $\sigma^{\varepsilon}_N\leqq\sigma^*$

かつ $\sigma^{\varepsilon}_N\leqq N$ 。したがって、$(s,x)\notin\Gamma^*$ に対して、

$$Z^{\pi}_t\notin\Gamma^*\qquad\text{for }\forall t\in[s,\sigma^{\varepsilon}_N]\qquad(a.s.)\text{。}$$

$(s,x) \notin \Gamma^*$ とする。上の議論と条件(iii)から、任意の $\pi \in \Pi$ と $N<\infty$, $\varepsilon>0$ に対して、

$$E^{\pi}_{(s,x)}\left[\int_s^{\sigma_N^{\varepsilon}} \mathcal{U}^{\pi} \varphi^{\pi^*}_{\tau^*}(Z^{\pi}_t)\,dt\right] \leqq 0 .$$

補題4.3より、

$$\varphi^{\pi^*}_{\tau^*}(s,x) \geqq D^{\pi}_{\sigma_N^{\varepsilon}} \varphi^{\pi^*}_{\tau^*}(s,x) .$$

$\varphi^{\pi^*}_{\tau^*} \geqq g$ であることと定理3.1から、

$$\varphi^{\pi^*}_{\tau^*}(s,x) \geqq D^{\pi}_{\sigma_N^{\varepsilon}} g(s,x) = \varphi^{\pi}_{\sigma_N^{\varepsilon}}(s,x) \geqq v_N(\pi)(s,x) - \varepsilon .$$

$\pi \in \Pi$ に対して sup をとり、$\varepsilon \downarrow 0$, $N \to \infty$ とすると、定理3.2から、 $\varphi^{\pi^*}_{\tau^*}(s,x) \geqq u^*(s,x)$ 。

以上より、(π^*, τ^*) は optimal である。

<u>定理4.3.</u> $u^* \in \mathcal{D}$ とする。そのとき、$\pi^* \in \Pi$ が存在して、

(i) $\sigma^* \in C_s(\pi^*)$ $(\forall s \geqq 0)$,

(ii) $\varphi^{\pi^*}_{\sigma^*} \geqq g$ on Z,

(iii) $\mathcal{U}^{\pi^*} u^*(z) = 0$, $\forall z \notin \Gamma^*$,

であるならば、(π^*, σ^*) は optimal である。

(証明)

$z \in \Gamma^*$ に対しては、明らかに $u^*(z) = \varphi^{\pi^*}_{\sigma^*}(z)$ 。

$(s,x) \notin \Gamma^*$ とする。各 $N<\infty$ に対して $\sigma^N \equiv \min(\sigma^*, N)$ とおく。条件(iii)から、各 $N<\infty$ に対して、

$$E^{\pi^*}_{(s,x)}\left[\int_s^{\sigma^N} \mathcal{U}^{\pi^*} u^*(Z^{\pi^*}_t)\,dt\right] = 0 .$$

補題4.3から

$$u^*(s,x) = D^{\pi^*}_{\sigma^*_N} u^*(s,x) .$$

$\sigma^* \in C_s(\pi^*)$ であるから、$N \to \infty$ とすると

$$u^*(s,x) = D^{\pi^*}_{\sigma^*} u^*(s,x) = D^{\pi^*}_{\sigma^*} g(s,x) = \varphi^{\pi^*}_{\sigma^*}(s,x) 。$$

以上より、(π^*, σ^*) は optimal である。

§5、平均型の最適化問題

この節で考える平均型は、Taylor[14]で導入されたものを、政策と停止規則を対とした M/G/1 待ち行列における最適化問題で議論したものである。目的は、平均型最適であるための十分条件を与えることである。

$\pi \in \Pi$, $(s,x) \in Z$ に対して、$\underline{\tilde{C}_{(s,x)}(\pi)}$ を $s < E^{\pi}_{(s,x)}[\tau] < \infty$ をみたす $\tau \in C_s(\pi)$ の全体とし、

$$\underline{\tilde{\Lambda}(s,x)} \equiv \{(\pi,\tau) \in \Lambda(s) \mid \tau \in \tilde{C}_{(s,x)}(\pi)\},$$

とおく。

このとき、平均型評価関数を次で定める；$(s,x) \in Z$, $(\pi,\tau) \in \tilde{\Lambda}(s,x)$ に対して、

$$\Phi^{\pi}_{\tau}(s,x) \equiv \varphi^{\pi}_{\tau}(s,x) \Big/ \left(E^{\pi}_{(s,x)}[\tau] - s\right) 。$$

また、平均型最適利得関数を

$$v^*(s,x) \equiv \sup_{(\pi,\tau) \in \tilde{\Lambda}(s,x)} \Phi^{\pi}_{\tau}(s,x) , \quad (s,x) \in Z,$$

とし、$(\pi^*, \tau^*) \in \tilde{\Lambda}(z_0)$ が、$z_0 \in Z$ で平均型最適であるとは、$v^*(z_0) = \Phi^{\pi^*}_{\tau^*}(z_0)$ のときにいう。

constant d, $z = (s, x) \in Z$, $(\pi, \tau) \in \Lambda(s)$ に対して、

$$\varphi^{\pi}_{\tau}[d](z) \equiv E^{\pi}_{(s,x)}\Big[\int_s^{\tau} (r^{\pi}(Z^{\pi}_t) - d)\, dt + \sum_{j=1}^{\infty} \rho(X^{\pi}_{\tau_j})\, \chi(\tau_j \leqq \tau) + g(Z^{\pi}_{\tau})\Big],$$

$$u^*[d](z) \equiv \sup_{(\pi,\tau) \in \Lambda(s)} \varphi^{\pi}_{\tau}[d](z),$$

$$\tilde{u}[d](z) \equiv \sup_{(\pi,\tau) \in \tilde{\Lambda}(z)} \varphi^{\pi}_{\tau}[d](z),$$

$$\Gamma_d \equiv \{z \in Z \mid u^*[d](z) = g(z)\}$$

$$\sigma_d \equiv \inf\{t \geqq s \mid Z^{\pi}_t \in \Gamma_d\}, \quad s \geqq 0, \quad \pi \in \Pi 。$$

<u>仮定3.</u>　constant d に対して、r を $r - d$ と置きかえることにより、仮定1、2をみたす。(このとき、d は仮定3をみたすということにする。)

<u>補題5.1.</u>　仮定3をみたす任意の constant d に対して、$\pi^* \in \Pi$ が存在して、

(i)　$\sigma_d \in C_s(\pi^*)$　$(\forall s \geqq 0)$,

(ii)　$\varphi^{\pi^*}_{\sigma_d}[d] \in \mathcal{D}$,

(iii)　$\varphi^{\pi^*}_{\sigma_d}[d] \geqq g$　on Z,

(iv) $\sup_{\pi\in\Pi} \mathcal{U}^{\pi} \varphi^{\pi^*}_{\sigma_\alpha}[\alpha](z) \leqq 0$, $\forall z \notin \Gamma_\alpha$,

であるならば、 $u^*[\alpha] = \varphi^{\pi^*}_{\sigma_\alpha}[\alpha]$ on Z 。

証明は、定理4.2からすぐ示される。

定理5.1. (Taylor の拡張)

$z_0 \in Z$ を固定し、constant α は仮定3をみたすものとする。このとき、$(\pi^*, \tau^*) \in \hat{\Lambda}(z_0)$ が存在して、

$$\hat{u}[\alpha](z_0) = \varphi^{\pi^*}_{\tau^*}[\alpha](z_0) = 0$$

であるならば、(π^*, τ^*) は(z_0で)平均型最適である、i.e.

$$v^*(z_0) = \Phi^{\pi^*}_{\tau^*}(z_0) = \alpha \text{。}$$

また、逆も真である。

(証明)

Φ^π_τ と $\varphi^\pi_\tau[\alpha]$ の定義から、明らかに $z=(s,x)\in Z$ に対し、

(*) $$\varphi^\pi_\tau[\alpha](z) = \left(\Phi^\pi_\tau(z) - \alpha\right)\left(E^\pi_z[\tau] - s\right), \quad (\pi,\tau)\in\hat{\Lambda}(z) \text{。}$$

これより、$\varphi^{\pi^*}_{\tau^*}[\alpha](z_0) = 0$ ならば、$\Phi^{\pi^*}_{\tau^*}(z_0) = \alpha$ 。また、任意の $(\pi,\tau)\in\hat{\Lambda}(z_0)$ に対して、$\hat{u}[\alpha](z_0) = 0 \geqq \varphi^\pi_\tau[\alpha](z_0)$ なので、$\Phi^\pi_\tau(z_0) \leqq \alpha$ 。したがって、$v^*(z_0) = \Phi^{\pi^*}_{\tau^*}(z_0) = \alpha$。

逆に、$\alpha = v^*(z_0) = \Phi^{\pi^*}_{\tau^*}(z_0)$ とおくと、(*)から、任意の $(\pi,\tau)\in\hat{\Lambda}(z_0)$ に対して

$$0 = \varphi^{\pi^*}_{\tau^*}[\alpha](z_0) \geqq \varphi^\pi_\tau[\alpha](z_0) \text{。}$$

したがって、$\hat{u}[\alpha](z_0) = \varphi^{\pi^*}_{\tau^*}[\alpha](z_0) = 0$ 。

次の定理では、Ross[12]とは異なる形で十分条件を与えている。

<u>定理5.2.</u> $z_0=(s_0,x_0)\in Z$ を固定する。このとき、constant α と $\pi^*\in\Pi$ が存在して、

(i) α は仮定3をみたす,

(ii) $u^*[\alpha](z_0)=0$

(iii) $\sigma_\alpha\in\widetilde{C}_{z_0}(\pi^*)$

(iv) $\varphi^{\pi^*}_{\sigma_\alpha}[\alpha]\in\mathcal{D}$

(v) $\begin{cases}\varphi^{\pi^*}_{\sigma_\alpha}[\alpha]\geq g & \text{on } [s_0,\infty)\times S,\\ \sup_{\pi\in\Pi}\mathcal{U}^\pi\varphi^{\pi^*}_{\sigma_\alpha}[\alpha](z)\leqq 0, & \forall z\notin\Gamma_\alpha\cup[0,s_0)\times S,\end{cases}$

であるならば、(π^*,σ_α) は、z_0で平均型最適であり、かつ $v^*(z_0)=\alpha$ 。

（証明）

$g(z_0)=0$ とすると、$u^*[\alpha](z_0)=0$ なので $\sigma_\alpha=s_0$。したがって $\sigma_\alpha\notin\widetilde{C}_{z_0}(\pi^*)$ 。これは条件(iii)に反する。よって $g(z_0)\neq 0$ 。ゆえに、$z_0\notin\Gamma_\alpha$ 。

補題5.1と条件(i)～(v)から

$$\varphi^{\pi^*}_{\sigma_\alpha}[\alpha](z_0)=u^*[\alpha](z_0)=0 。$$

また、$u^*[\alpha](z)\geq\widetilde{u}[\alpha](z)\geq\varphi^{\pi^*}_{\sigma_\alpha}[\alpha](z)$ であるから、

$$\widetilde{u}[\alpha](z_0)=\varphi^{\pi^*}_{\sigma_\alpha}[\alpha](z_0)=0 。$$

定理5.1より、$v^*(z_0)=\Phi^{\pi^*}_{\sigma_\alpha}(z_0)=\alpha$ 。

References

[1] Bieber,G., On optimal control of Markov processes with application to a class of service systems, Math. Operat. Statist., Ser. Optimization, 10, 143-162 (1979)

[2] Crabill,T.B.,D.Gross and M.J.Magazine, A classified bibliography of research on optimal design and control of queues, Operations Res., 25, 219-232 (1977)

[3] Doshi,B.T., Optimal control of the service rate in an M/G/1 queueing system, Adv. Appl. Prob., 10, 682-701 (1978)

[4] Furukawa,N. and S.Iwamoto, Stopped decision processes on complete separable metric spaces, J. Math. Anal. Appl., 31, 615-658 (1970)

[5] Hordijk,A., Dynamic Programming and Markov Potential Theory, Mathematisch Centrum, Amsterdom, 1974

[6] Ito,K., Lectures on Stochastic Processes, Tata Inst. Fundamental Res., Bombay, 1961

[7] Krylov,N.V., Control of a solution of a stochastic integral equation, Theory Prob. Appl., 17, 114-131 (1972)

[8] Mitchell,B., Optimal service-rate selection in an M/G/1 queue, SIAM J. Appl. Math., 24, 19-35 (1973)

[9] Nisio,M., On nonlinear semigroups for Markov processes associated with optimal stopping, Appl. Math. Opt., 4, 143-169 (1978)

[10] Ohtsubo,Y., Optimal control of the service rate and the stopping rule in an M/G/1 queue, Mem. Fac. Sci., Kochi Univ., Ser.A, 3, 75-103 (1982)

[11] Prabhu,N.U., Stochastic control of queueing systems, Naval Res. Logist. Quart., 21, 411-418 (1974)

[12] Ross,S.M., Infinitesimal look-ahead stopping rules, Ann. Math. Statist., 42, 297-303 (1971)

[13] Širjaev,A.N., Statistical Sequential Analysis, Transl. Math. Monogra.
Amer. Math. Soc., 1973

[14] Taylor,H.M., Optimal stopping in a Markov process, Ann. Math. Statist.
39, 1333-1344 (1968)