数理解析研究所講究録
第1912巻 2014年 181-192

多変量ベイズ管理図の適応手法
(Adaptive methods for multivariate Bayesian Control Chart)

日本ピュアテック株式会社・品質保証部　佐々木　稔 (Minoru SASAKI)
Nippon Puretec Co., Ltd.
神奈川大学・理学部　堀口　正之 (Masayuki HORIGUCHI)
Faculty of Science, Kanagawa University
千葉大学・名誉教授　蔵野　正美 (Masami KURANO)
Professor Emeritus, Chiba University

1　はじめに

現在の品質管理に用いられる管理図の原型は, およそ82年前に考案されたシューハート管理図 (cf. [13]) である. シューハート管理図は, 通常, 中心線から両側へ 3σ の距離に管理限界線をもち, 規則的な時間間隔で工程からサンプリングされたデータをもとにプロットしたグラフからなる.

ベイズ推定を用いた適応型の品質管理については,多くの研究があり (cf.[2, 12, 17, 22]), 品質管理の現場でその有効性が報告されている. ベイズ推定を基本とした品質管理では, 蓄積された情報を基にして管理限界, サンプルサイズおよびサンプリング間隔を変更して事象や状況の変化に適応していく (cf. [17]).

品質管理図を設計する場合, 統計的手法に重きをおくか, あるいは経費的な側面に重きをおくかは重要であるが, それぞれ一長一短があることが知られている (cf. [21]). そこで, 統計的および経済的な両側面を考慮した管理図の作成が考えられるが, これに答えるためには, 問題を逐次決定過程としてとらえ, その解析結果を管理図に反映させる必要がある. この種の研究も多く行われている (cf. [1, 6, 10, 16, 18, 19]). 最近, V. Makis[9] は, システムの状態が既知のパラメータ $\theta(\theta > 0)$ をもつ指数分布に従って正常状態 (state in control) から不正常な状態 (state out of control) に移行する多変量管理モデルをマルコフ決定過程 (Markov decision process, MDP)(cf. [3, 11]) として定式化し, 長時間の平均期待コスト基準のもとでの最適な管理政策を求め, これにより, 多変量管理図の作成方法を提案した. ここでは, パラメータ θ が未知の場合の "Makis Model" を取り扱い, 平均コスト基準および割引きされた総期待コスト最小化問題に対する適応最適政策を議論して, 適応管理図の作成方法を提案する. 第2節では, Makis[9] が扱ったベイズモデルを述べ, 問題の定式化と重要な補題を述べる. 第3節は, ベイズモデルの最適化を考察し, これらの結果を利用して第4節で適応管理図を求める.

2　ベイズ管理モデル

本論で取り扱う品質管理モデルを述べ, 同値なベイズモデルによって定式化する.

システムの正常な状態を "0", 不正常な状態を "1" で表す. 状態0から状態1に移行 (故障原因の発生) する時間分布は, パラメータ θ の指数分布とする. θ の真値は未知で

$\theta \in \Theta = \{\theta_1, \theta_2, \ldots, \theta_r\}$ とする. ただし,$\theta_1, \theta_2, \ldots, \theta_r$ は互いに異なる正数とする. θ の事前分布に従う確率変数を $\widetilde{\theta}$ で表す. 時刻 $t(t \geqq 0)$ のシステムの状態を X_t で表す. 与えられた定数 $h > 0$ に対して, h の時間間隔で状態に対する情報 (大きさ n の q 次元データ) を取得して, システムの運用を継続 "to continue" するか (この行動を "0" で表す), システムの運用を停止して故障の有無を精査 "to stop and search" するか (この行動を "1" で表す) を選択する. 精査したとき, システムが正常であるか不正常であるかが正確に分かり, もし不正常ならば正常な状態に瞬間的に取り替え, 正常な状態からプロセスは再スタートする. 意思決定者 (decision maker) が, 状態に関する情報を得て, 0 か 1 の行動を選択する決定時点 (decision epoch) は, $ih(i = 1, 2, \ldots)$ である.

時点 ih (i=1,2,3,...) で取得する多次元正規情報:

大きさ n の q 次元の標本

$$(1) \qquad Y_i = \begin{bmatrix} \boldsymbol{y}_1^i \\ \boldsymbol{y}_2^i \\ \vdots \\ \boldsymbol{y}_1^n \end{bmatrix}, \quad \boldsymbol{y}_j^i = \left(y_{j1}^i, y_{j2}^i, \ldots, y_{jq}^i\right) \quad j = 1, 2, \ldots, n.$$

仮定:

$X_{ih} = 0$ (または 1) のとき, $\boldsymbol{y}_1^i, \boldsymbol{y}_2^i, \ldots, \boldsymbol{y}_n^i$ は互いに独立で各 $\boldsymbol{y}_j^i$ は同一の分布 $N_q(\mu_0, \Sigma)$ ($N_q(\mu_1, \Sigma)$) に従う.

ただし, $N_q(\mu_0, \Sigma), N_q(\mu_1, \Sigma)$ は分散共分散行列 (正値)Σ であり, それぞれ平均ベクトル $\mu_0 = (\mu_{01}, \mu_{02}, \ldots, \mu_{0q}), \mu_1 = (\mu_{11}, \mu_{12}, \ldots, \mu_{1q})$ をもつ q 次元正規分布で表す. ここで, μ_1 の μ_0 からの M-距離 d_1 について, 次を仮定する:

$$(2) \qquad d_1 := \left[(\mu_1 - \mu_0)\,\Sigma^{-1}\,(\mu_1 - \mu_0)\right]^{\frac{1}{2}} > 0.$$

コスト構造:

- システムの運用を停止して故障の有無を精査する費用 $A > 0$
- 状態 1(不正常) を状態 0(正常) に取り替える費用 $R \geqq 0$
- 不正常の状態のまま運用したときの単位時間当たりのコスト $M > 0$
- 大きさ n のサンプルをとる費用 $b + nc$ $(b, c \geqq 0)$

この決定過程は, 部分観測可能なマルコフ決定過程としてみることができる. 従って, $X_t = 1$(時刻 t で状態が不正常) である確率を新しい状態として状態空間 $S = [0, 1]$ に拡張して, ベイズの定理により事前状態分布を事後状態分布に変換することにより状態の推移を記述する完全観測のベイズモデルに同値に変換される (cf. [7, 20]).

同値なベイズモデル:

行動の決定時点は ih $(i = 1, 2, \ldots)$ で次の要素からなる MDP モデルを考える.

$$\begin{cases} S = [0, 1]\text{:状態空間} \\ A = \{0, 1\}\text{:行動空間} \\ \Theta = \{\theta_1, \theta_2, \ldots, \theta_r\}\text{:パラメータ空間} \\ c(p, a)\text{:}p \in S, a \in A\text{ のときのコスト} \end{cases}$$

標本空間を $\overline{\Omega} = \Theta \times \Omega, \Omega = S \times (A \times S)^\infty$ と表し, プロセスを表す確率変数を $\tilde{\theta}, \tilde{p}_0, \tilde{a}_0, \tilde{p}_1, \tilde{a}_1, \ldots$ とする. すなわち, $\overline{\Omega} \ni \omega = (\theta, p_0, a_0, p_1, a_1, p_2, \ldots)$ のとき, $\tilde{\theta}(\omega) = \theta, \tilde{p}_0(\omega) = p_0, \tilde{a}_0(\omega) = a_0, \tilde{p}_1(\omega) = p_1, \ldots$ である. ただし, $p_0 = 0$ として一般性を失わない.

mh 時点で状態 $\tilde{p}_m = p$ のとき, 行動 $\tilde{a}_m = 0(1)$ を選択し $(m+1)h$ 時点で $Y_{m+1} = y^{m+1}$ を観測した場合には, $(m+1)h$ 時点の状態は

$$\tilde{p}_{m+1} = T(p, y^{m+1}, 0)\ \big(T(p, y^{m+1}, 1)\big)$$

に推移する, ただし, ベイズの定理により事前-事後ベイズ作用素 T は次のように定まる (Lemma 1[9]).

$$(3)\qquad \begin{cases} T(p,z,0) = (1-(1-p)e^{-\theta h}h_1(z))/h(z|p), \\ T(p,y,1) = T(0,z,0), \\ \text{ただし}, \\ z = 2\sum_{j=1}^{n}(\boldsymbol{y}_j - \mu_0)\Sigma^{-1}(\mu_0-\mu_1)^T, \\ y = \begin{bmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \\ \vdots \\ \boldsymbol{y}_n \end{bmatrix}, \boldsymbol{y}_j = (y_{j1}, y_{j2}, \ldots, y_{jq}), \\ h(z|p) = (1-(1-p)e^{-\theta h})h_1(z) + (1-p)e^{-\theta h}h_0(z), \\ h_0(z) = N_1(0, 4nd_1^2), h_1(z) = N_1(-2nd_1^2, 4nd_1^2). \end{cases}$$

注: z は十分統計量で状態の推移は z の値のみに依存する. $X_{ih} = 0(1)$ のとき, z は平均 $0(-2nd_1^2)$, 分散 $4nd_1^2$ の 1 次元正規分布に従う.

θ が真のときのコストは次で与えられる:

$$(4)\qquad \begin{cases} c(p,0) = M\int_0^h I_{\{X_s=1\}}\,ds + b + nc = M\left[h - \frac{1-p}{\theta}(1-e^{-\theta h})\right] + b + nc, \\ c(p,1) = c_1(p) + c(0,0). \end{cases}$$

ただし, $c_1(p) = A + Rp$.

政策 (policy) は $\pi = (\pi_0, \pi_1, \ldots)$ で表し, $\pi_m(H_m) \in A = \{0,1\} (m \geqq 0), H_m = (\tilde{p}_0, \tilde{a}_0, \tilde{p}_1, \ldots, \tilde{p}_m)$ とする. 政策の全体を Π で表す.

$\mathcal{P}(\Theta)$ を Θ 上の確率分布の全体とする. 任意の政策 $\pi = (\pi_0, \pi_1, \ldots) \in \Pi$ に対して, 停止時刻 (stopping time) の系列 $\tau = (\tau_0, \tau_1, \tau_2, \ldots)$ が次によって定まる.

$$\tau_0 = 0, \tau_k = \min\{k - \tau_{k-1} | \pi_k(H_k) = 1, k > \tau_{k-1}\}.$$

明らかに, 政策 π と停止時刻の系列 (停止政策と呼ぶ) τ は, 1 対 1 に対応する. 従って, 以降では必要に応じて, 政策と対応する停止政策を同一視して取り扱う.

$\tilde{\theta}$ の分布 $q = (q(\theta_1), q(\theta_2), \ldots, q(\theta_r)) \in \mathcal{P}(\Theta)$ と初期状態分布 $p_0 = p \in S$ が与えられたときの政策 $\pi \in \Pi$ の平均期待コスト $\varphi(\pi|q, p_0)$ を次で定める:

$$(5)\qquad \varphi(\pi|q,p_0) = \limsup_{k\to\infty} \frac{1}{E(s_k)} E\left[\sum_{m=0}^{s_k} c(\tilde{p}_m, \tilde{a}_m)|q, p_0\right],$$

ただし, $s_k = \sum_{l=0}^{k} \tau_l (k \geqq 1)$.
さらに, 割引きされた総期待コスト $v(\pi|q, p_0)$ は次で定める:

$$(6) \qquad v(\pi|q,p_0) = \sum_{m=0}^{\infty} \beta^m E_\pi \left[c(\tilde{p}_m, \tilde{a}_m) | q, p_0 \right].$$

ただし, $\beta (0 < \beta < 1)$ は割引き率を表し, $E_\pi [\cdot|q,p]$ は, q, p および π が与えられたときの $\overline{\Omega}$ 上に定まる確率測度 $P_\pi(\cdot|q,p)$ に関する期待値である.
$\varphi(\pi|q,p), v(\pi|q,p)$ を最小にする政策 $\pi \in \Pi$ を求める議論は §3, §4 で行う.
この節の最後に, §4 で用いられる未知パラメータ $\tilde{\theta} \in \Theta$ の推定法に関する一致性 (consistency) の補題, およびコスト比の停止問題に関する補題を与える.
確率変数の系列 $X_1, X_2, \ldots$ は互いに独立で, X_k は確率密度関数 $f_k(\cdot|\theta)$ $(\theta \in \Theta)$ をもつとする. μ をルベーグ測度とする.

<u>**仮定 A**</u>:

$$D_{ij}^{(k)} := \{x | f_k(x|\theta_i) \neq f_k(x|\theta_j)\}.$$

このとき, $\mu(D_{ij}^{(k)}) > 0$ $(k \geqq 1, i \neq j, i, j = 1, 2, \ldots, r)$.
仮定 A のもとで, Hölder の不等式より

$$\int f_k(x|\theta_i)^{\frac{1}{2}} f_k(x|\theta_j)^{\frac{1}{2}} \, d\mu(x) < \left(\int f_k(x|\theta_i) \, d\mu(x) \right)^{\frac{1}{2}} \left(\int f_k(x|\theta_j) \, d\mu(x) \right)^{\frac{1}{2}} = 1.$$

この事実を踏まえて, 次の仮定 B を設定する.

<u>**仮定 B**</u>:

$$\gamma := \sup_{k \geqq 1} \max_{\substack{\theta_i, \theta_j \in \Theta \\ i \neq j}} \int f_k(x|\theta_i)^{\frac{1}{2}} f_k(x|\theta_j)^{\frac{1}{2}} \, d\mu(x) < 1.$$

$X_1 = x_1, X_2 = x_2, \ldots$ に対して, 次を定義する (最尤推定 cf. [4]).

$$\tilde{\theta}_k = \arg\max_{\theta \in \Theta} \prod_{l=1}^{k} f_l(x_l|\theta).$$

Θ 上の初期分布 $q_0 = (q_0(\theta_1), q_0(\theta_2), \ldots, q_0(\theta_r)) \in \mathcal{P}(\Theta)$ に対して, 事後分布の系列 $\{q_k\}_{k=1}^{\infty}$ を逐次的に次で定める (ベイズ推定 cf. [4]):

$$q_k(\theta_i) = \frac{f_k(x_k|\theta_i) q_{k-1}(\theta_i)}{\sum_{l=1}^{r} f_k(x_k|\theta_l) q_{k-1}(\theta_l)} \quad (i = 1, 2, \ldots, r).$$

このとき, 次が成り立つ.

Lemma 2.1 (consistency) 仮定 A,B のもとで, 次が成立する.

(i) $P\left(\tilde{\theta}_k \neq \theta_{i_0} | \tilde{\theta} = \theta_{i_0}\right) \leqq K_1 \gamma^k$ *for some* $K_1 > 0$ $(k \geqq 1)$.

(ii) $q_0(\theta_l) > 0$ $(l = 1, 2, \ldots, r)$ のとき, 任意の $\varepsilon > 0$ と $i \neq i_0$ なる i に対して, $P\left(q_k(\theta_i) > \varepsilon | \tilde{\theta} = \theta_{i_0}\right) \leqq K_2 \gamma^k$ $(k \geqq 1)$. ただし, $K_2 = \max_{i \neq i_0} \frac{1}{\varepsilon^{\frac{1}{2}}} \left(\frac{q_0(\theta_i)}{q_0(\theta_{i_0})} \right)^{\frac{1}{2}}$.

(証明) (ii) について証明を与える.

$$q_k(\theta_i) = \frac{q_0(\theta_i)\prod_{l=1}^k f_l(x_l|\theta_i)}{\sum_{j=1}^r q_0(\theta_j)\prod_{l=1}^k f_l(x_l|\theta_j)} \leqq \frac{q_0(\theta_i)\prod_{l=1}^k f_l(x_l|\theta_i)}{q_0(\theta_{i_0})\prod_{l=1}^k f_l(x_l|\theta_{i_0})}.$$

この関係と仮定 B より次を得る.

$$\begin{aligned}
P(q_k(\theta_i) > \varepsilon|\tilde{\theta} = \theta_{i_0}) &= P\left(\frac{1}{\varepsilon}q_k(\theta_i) > 1|\tilde{\theta} = \theta_{i_0}\right) \\
&\leqq \frac{1}{\varepsilon^{\frac{1}{2}}}E\left[q_k(\theta_i)^{\frac{1}{2}}|\tilde{\theta} = \theta_{i_0}\right] \\
&\leqq \frac{1}{\varepsilon^{\frac{1}{2}}}\left(\frac{q_0(\theta_i)}{q_0(\theta_{i_0})}\right)^{\frac{1}{2}}\int\prod_{l=1}^k f_l(x_l|\theta_i)^{\frac{1}{2}}f_l(x_l|\theta_{i_0})^{\frac{1}{2}}\,d\mu(x_1)\cdots d\mu(x_k) \\
&\leqq \frac{1}{\varepsilon^{\frac{1}{2}}}\left(\frac{q_0(\theta_i)}{q_0(\theta_{i_0})}\right)^{\frac{1}{2}}\gamma^k \quad (k \geqq 1).
\end{aligned}$$

(i) に関しても同様に証明される.∎

確率変数の系列 $X_1, X_2, \ldots, 0 < T_1 < T_2 < \cdots$ a.s. に対して, stopping times の集合 $\mathbb{C} := \{\tau|\tau \geqq 1, E[T_\tau] < \infty\}$ とする. このとき, 比 $E[X_\tau]/E[T_\tau]$ の最適停止問題について, 次が成り立つ (cf. [5]).

Lemma 2.2 (λ-maximization technique, Theorem 1 in Chap. 6, Ferguson[5]) ある λ について, $\sup_{\tau\in\mathbb{C}} E[\lambda T_\tau - X_\tau] = 0$ ならば, $\inf_{\tau\in\mathbb{C}}\dfrac{E[X_\tau]}{E[T_\tau]} = \lambda$ である. また, $\sup_{\tau\in\mathbb{C}} E[\lambda T_\tau - X_\tau] = E[\lambda T_{\tau^*} - X_{\tau^*}]$ ならば, $\dfrac{E[X_{\tau^*}]}{E[T_{\tau^*}]} = \lambda$ となる.

3 コスト比の最適停止

ここでは, 既知パラメータ $\tilde{\theta} = \theta$ の場合, すなわち $q_0 = I_{\{\theta\}}$ の場合の最適化について考察する. ただし, I_A は集合 A の指示関数を表す. $I_{\{\theta\}}$ を簡単に θ で表す. 政策 $\pi^* \in \Pi$ が θ-平均最適であるとは, すべての $p_0 \in S$ とすべての $\pi \in \Pi$ に対して $\varphi(\pi^*|\theta, p_0) \leqq \varphi(\pi|\theta, p_0)$ が成り立つ場合をいう. また, すべての $p_0 \in S$ とすべての $\pi \in \Pi$ に対して, $v(\pi^*|\theta, p_0) \leqq v(\pi|\theta, p_0)$ が成り立つとき, π^* を θ-割引き最適という.

$k \geqq 1$ に対して,

$$X_k := c(p_0) + c(p_1) + \cdots + c(p_{k-1}) + c_1(p_k) \ (p_0 = 0)$$

とする. ただし,

$$\begin{cases} c(p) = M\left[h - \dfrac{1-p}{\theta}(1 - e^{-\theta h})\right] + b + nc, \\ c_1(p) = A + Rp, \\ p_0 = 0, p_l = T(p_{l-1}, z_{l-1}, 0) \ l = 1, 2, \ldots, \\ z_l \text{ は, 密度関数 } h(z|p_{l-1}) \text{ に従う確率変数} \end{cases} \tag{7}$$

である. このとき, 次が成り立つ.

Theorem 3.1 (平均コスト基準) コスト比の最適停止問題:$\min_{\tau\in\mathbb{C}} E[X_\tau]/E[\tau]$ の最適停止時刻 τ^* に対して, $(\tau^*)^\infty = (\tau^*, \tau^*, \ldots)$ に対応する政策 τ^* は θ-平均最適である.

(証明) 政策 $\pi \in \Pi$ に対応する停止政策 $\tau = (\tau_1, \tau_2, \ldots)$ に対して, $\varphi(\pi|\theta, p_0)$ を次のように書き直せる.

$$\varphi(\pi|\theta, p_0) = \limsup_{m\to\infty} \frac{1}{E[S_m]} \sum_{k=1}^{m} E[X_{\tau_k}],$$

ただし, $S_m = \tau_1 + \tau_2 + \cdots + \tau_m$. 故に, $\dfrac{E[X_{\tau_i}]}{E[\tau_i]} \geqq \dfrac{E[X_{\tau^*}]}{E[\tau^*]}$ $(i = 1, 2, \ldots)$ より, $\varphi(\pi|\theta, p_0) \geqq \dfrac{E[X_{\tau^*}]}{E[\tau^*]}$ となる. 明らかに, $\varphi(\pi^*|\theta, p_0) = \dfrac{E[X_{\tau^*}]}{E[\tau^*]}$ より, 証明が終わる.∎

V. Makis[9] は, コスト比の停止問題:$\min_\tau \dfrac{E_{X_\tau}}{E[\tau]}$ について λ-maximization technique(Lemma 2.2) を適用して次の結果を得た.

Theorem 3.2 (V. Makis[9]) $A + R < \dfrac{M}{\theta}$ ならば, *control-limit* 型の θ-平均最適な政策が存在する. すなわち, $p_\theta^* \in (0, 1)$ が存在して, 決定関数 $f_\theta : S \to A$,

$$f_\theta(p) = \begin{cases} 0 & \textit{if } p < p_\theta^*, \\ 1 & \textit{if } p \geqq p_\theta^*. \end{cases} \tag{8}$$

による管理政策が θ-平均最適となる.

<u>**注意**</u>: $\theta \in \Theta$ が既知のとき, 決定関数 f_θ によって最適な管理図を作成することができる. すなわち, ih 時点の状態 $p_i \in S$ に対して得られた多変量データ $Y_i = y^i$ から事後分布 p_{i+1} を計算し, (8) によって $p_{i+1} < p_\theta^*$ ならば "continue," $p_{i+1} \geqq p_\theta^*$ ならば "stop and search" を選択することになる.

$k \geqq 0$ に対して

$$D_k := c(p_0) + \beta c(p_1) + \cdots + \beta^{k-1} c(p_{k-1}) + \beta^k c_1(p_k) \ (p_0 = 0)$$

とする. ただし, $c(p), p_l$ $(l \geqq 1)$ は (7) で与えられている.

Theorem 3.3 (割引きコスト基準) コスト比の最適停止問題:$\min_{\tau\in\mathbb{C}} \dfrac{E[D_\tau]}{1 - E[\beta^\tau]}$ の最適停止時刻 τ^* に対して $(\tau^*)^\infty = (\tau^*, \tau^*, \ldots)$ に対応する政策 τ^* は θ-割引き最適である.

(証明) §2 で定義したベイズモデルは, コンパクトな状態空間 S, 行動空間 A をもち, かつコスト関数 c は有界, 状態推移は連続であるので, θ-割引き最適な定常政策が存在する (cf. [3]). 今, 任意に与えられた定常政策に対応する停止政策 $(\tau)^\infty = (\tau, \tau, \ldots)$ に対して次の関係式が成り立つ.

$$v((\tau)^\infty|\theta, p_0) = E[D_\tau] + E[\beta^\tau] v((\tau)^\infty|\theta, p_0) \ \ (p_0 = 0).$$

これより, $v((\tau)^{\infty}|\theta, p_0) = E[D_\tau]/(1 - E[\beta^\tau])$. 故に, Theorem 3.3 の命題は明らかに成立する. ∎

次に, λ-maximization technique (§2 の Lemma 2.2) を適用して, θ-割引き最適な政策の構造を調べてみよう.

初期値 $p_0 = p \in S$ と $\lambda(-\infty < \lambda < \infty)$ との関数を次で定義する:

$$
\begin{aligned}
V(p, \lambda) &:= \max_{\tau \in \mathbb{C}} [\lambda(1 - E[\beta^\tau]) - E[D_\tau]], \\
V_m(p, \lambda) &:= \max_{\substack{\tau \in \mathbb{C} \\ 0 \leqq \tau \leqq m}} [\lambda(1 - E[\beta^\tau]) - E[D_\tau]] \quad (m \geqq 1).
\end{aligned}
$$

このとき, 動的計画法 (Dynamic Programming) の考え方を適用して, 次の最適方程式が成り立つ.

$$
\begin{aligned}
V(p, \lambda) &= \max\left\{ -A - Rp, \lambda(1-\beta) - M\left[h - \frac{1-p}{\theta}(1 - e^{-\theta h})\right] \right. \\
&\qquad \left. + \int V(T(p, z, 0), \lambda) h(z|p)\, dz \right\}, \\
V_m(p, \lambda) &= \max\left\{ -A - Rp, \lambda(1-\beta) - M\left[h - \frac{1-p}{\theta}(1 - e^{-\theta h})\right] \right. \\
&\qquad \left. + \int V_{m-1}(T(p, z, 0), \lambda) h(z|p)\, dz \right\} \quad (m \geqq 1).
\end{aligned} \tag{9}
$$

ただし, $V_0(p, \lambda) = -A - Rp$, $T, h(z|p)$ は §2 の式 (3) で与えられている.

このとき, 次が成り立つ.

Lemma 3.1

(i) $V_m(p, \lambda)$ は, 各 λ に関して $p \in S$ の凸かつ非増加関数である. また, 各 $p \in S$ に対して λ の凸かつ非減少関数である.

(ii) $V(p, \lambda)$ についても *(i)* と同様なことが成り立つ.

(証明) $V_m(p, \lambda)$ の p に関する凸性および非増加性は Lemma 2 in [9] と同じようにして証明される. また, $V_m(p, \lambda)$ の λ に関する凸性は, 一般論として Lemma 1 of Chapter 6 in [5] で証明されている. また, $V_m(p, \lambda)$ の λ に関する非減少性は, 定義式から明らかに成り立つ. これで, (i) が示された. $V_m(p, \lambda) \to V(p, \lambda)(m \to \infty)$ であるから, (ii) は明らかに成り立つ. ∎

Lemma 3.1 と最適方程式 (9) より, 次の方程式が容易に証明される.

Theorem 3.4 *control-limit* 型の θ-割引き最適な政策が存在する. すなわち, $\overline{p}_\theta \in (0, 1)$ が存在して, 最適な決定関数 $g_\theta : S \to A$ は次で与えられる:

$$
g_\theta(p) = \begin{cases} 0 & \text{if } p < \overline{p}_\theta, \\ 1 & \text{if } p \geqq \overline{p}_\theta. \end{cases} \tag{10}
$$

(証明) Lemma 3.1 と式 (9) より, 各 λ に対して $\overline{p}_\theta(\lambda) \in S$ が存在して, $\max_{\tau\in\mathbb{C}}(\lambda(1-E[\beta^\tau])-E[D_\tau])$ の最適停止時刻 $\tau^*(\lambda)$ は $\tau^*(\lambda)=\min\{k|p_k \geqq \overline{p}_\theta(\lambda)\}$ となる. $\lim_{\lambda\to-\infty} V(p_0,\lambda) = -\infty, \lim_{\lambda\to\infty} V(p_0,\lambda)=\infty$ かつ Lemma 3.1 より, $V(p_0,\lambda)$ は λ に関して非減少である. 故に, $V(p_0,\overline{\lambda})=0$ なる $\overline{\lambda}$ が存在する. §2 の Lemma 2.2 および Theorem 3.3 より $(\tau^*(\overline{\lambda}))^\infty$ に対応する政策 g_θ は $\overline{p}_\theta = \overline{p}_0(\overline{\lambda})$ によって式 (10) で与えられ, g_θ は θ-割引き最適である. ∎

4 適応手法による最適管理図

ここでは, 政策の定義を拡張して, 未知パラメータ θ が存在する場合の有効な管理図の作成方法を提案する.

§3 の政策 $\pi=(\pi_0,\pi_1,\dots)$ と対応する停止政策 $\tau=(\tau_1,\tau_2,\dots)$ を考える.

$t \geqq 0$ に対して, $G_t := (X_{hs_1}, X_{hs_2}, \dots, X_{hs_{\sigma(t)}})$, ただし, $\sigma(t) := \max\{k|s_k \leqq t\}$, $s_k = \sum_{i=1}^k \tau_i$. G_t は時刻 t までに "stop and search" をとることによって得られたシステムの状態に関する情報を表している.

(H_m, G_m) に依存して, $\overline{\pi}_m(H_m,G_m) \in A$ の系列 $\overline{\pi}=(\overline{\pi}_0,\overline{\pi}_1,\dots)$ の政策全体を $\overline{\Pi}$ とする.

任意の $\overline{\pi} \in \overline{\Pi}$ に対して,

$$\varphi(\overline{\pi}|\theta,p_0) := \limsup_{k\to\infty} \varphi_k(\overline{\pi}|\theta,p_0), \quad v(\overline{\pi}|\theta,p_0) := \limsup_{k\to\infty} v_k(\overline{\pi}|\theta,p_0), \tag{11}$$

ただし,

$$\begin{aligned}\varphi_k(\overline{\pi}|\theta,p_0) &= \frac{1}{E[s_k]}E\left[\sum_{m=0}^{s_k} c(\tilde{p}_m,\tilde{a}_m)\right], v_k(\overline{\pi}|\theta,p_0) = \sum_{m=1}^{\infty}\beta^{m-1}E[c(\tilde{p}_m,\tilde{a}_m)],\\ E[\cdot] &= E_{\overline{\pi}}[\cdot|\theta,p_0].\end{aligned} \tag{12}$$

$\overline{\pi} \in \overline{\Pi}$ が適応平均最適であるとは, すべての $\theta \in \Theta$ に対して, $\overline{\pi}^*$ が θ-平均最適である場合をいう, すなわち, すべての $\theta \in \Theta$ に対して次式が成り立つ.

$$\varphi(\overline{\pi}|\theta,p_0) = \inf_{\pi\in\Pi} \varphi(\pi|\theta,p_0). \tag{13}$$

$\overline{\pi}=(\overline{\pi}_0,\overline{\pi}_1,\dots)$ が学習割引き最適であるとは, 任意の $\varepsilon>0$ に対して, ある番号 N が存在して, すべての $\theta \in \Theta$ に関して,

$$E_{\overline{\pi}}(v(\overline{\pi}\{H_t,G_t\})|\theta,p_0)) \leqq \inf_{\pi\in\Pi} v(\pi|\theta,p_0)+\varepsilon \quad (t \geqq N) \tag{14}$$

が成り立つ場合をいう (cf. [8]). ただし, $(H_t,G_t)=(h_t,g_t)$ のとき,

$$\begin{cases}\overline{\pi}\{h_t,g_t\} = (\overline{\pi}\{h_t,g_t\}_0, \overline{\pi}\{h_t,g_t\}_1,\dots),\\ \overline{\pi}\{h_t,g_t\}_k(h'_k,g'_k) = \overline{\pi}_{t+k}((h_t-h'_k),(g_t,g'_k)),\\ (h_t-h'_k) = (p_0,a_0,p_1,\dots,a_{t-1},p'_0,a'_0,\dots,p'_k),\\ (g_t,g'_k)) = (x_1,x_2,\dots,x_{\sigma(t)},x'_1,x'_2,\dots,x'_{\sigma(k)}),\\ h_t = (p_0,a_0,p_1,\dots,p_t), h'_k=(p'_0,a'_0,p'_1,\dots,p'_k),\\ g_t=(x_1,x_2,\dots,x_{\sigma(t)}), g'_k = (x'_1,x'_2,\dots x'_1,x'_2,\dots,x'_{\sigma(k)}).\end{cases} \tag{15}$$

ここでは, 推定と制御の原理 (*principle of estimation and control*) (cf. [7, 14]) の考えにより, 上記2つの最適政策を求める.

任意の停止政策 $\tau = (\tau_1, \tau_2, \ldots)$ と $\theta \in \Theta$ に対して, 次を定義する:

$$f_k^{\tau_k}(0|\theta) := P(X_{s_k h} = 0|\theta, p_0) = e^{-\theta h \tau_k},$$
$$f_k^{\tau_k}(1|\theta) := P(X_{s_k h} = 1|\theta, p_0) = 1 - e^{-\theta h \tau_k}.$$

時点 $s_k h$ での θ の最尤推定 $\tilde{\theta}_k$ は, $X_{s_1 h} = x_1, X_{s_2 h} = x_2, \ldots, X_{s_k h} = x_k$ のとき,

$$\tilde{\theta}_k(x_1, x_2, \ldots, x_k) = \arg\max_{\theta \in \Theta} \prod_{l=1}^{k} f_l^{\tau_l}(x_l|\theta)$$

で与えられる. ただし, τ_l は $(x_1, x_2, \ldots, x_{l-1})$ に依存してもよい.

§3 の Theorem 3.2 より, 各 $\theta \in \Theta$ に対して, θ-平均最適な停止政策 $(\tau^*(\theta))^\infty$ が与えられる. 最尤推定 $\tilde{\theta}_k$ と $(\tau^*(\theta))^\infty$ を用いて, 新しい停止政策 $\overline{\tau}^* = (\tau^*(\tilde{\theta}_1), \tau^*(\tilde{\theta}_2), \ldots)$ を構成する ($\tilde{\theta}_0 \in \Theta$ は任意). $\overline{\tau}^*$ に対応する政策を $\overline{\pi}^* \in \overline{\Pi}$ とする. $\overline{\pi}^*$ は推定値 $\tilde{\theta}_k$ があたかも未知のパラメータの真値であると考えて最適な行動を選ぶ政策である.

$\theta_i, \theta_j \in \Theta$ $(i \neq j)$ と任意の停止時刻 τ に対して,

$$\begin{aligned} G_{ij}(\tau) &:= E\left[f_k^{\tau}(0|\theta_i)^{\frac{1}{2}} f_k^{\tau}(0|\theta_j)^{\frac{1}{2}} + f_k^{\tau}(1|\theta_i)^{\frac{1}{2}} f_k^{\tau}(1|\theta_j)^{\frac{1}{2}}\right] \\ &= E\left[\left(e^{-\theta_i h\tau} e^{-\theta_j h\tau}\right)^{\frac{1}{2}} + \left(\left(1 - e^{-\theta_i h\tau}\right)\left(1 - e^{-\theta_j h\tau}\right)\right)^{\frac{1}{2}}\right] \\ &< 1. \end{aligned}$$

これを用いて, $0 < \gamma < 1$ を次で定義する.

$$\gamma := \max_{\theta \in \Theta} \left[\max_{\substack{i,j \\ (i \neq j)}} G_{ij}\left(\tau^*(\theta)\right)\right].$$

このとき, §2 の Lemma 2.1 より, 次が成り立つ.

Lemma 4.1 次の不等式系を満たす K が存在する:

$$P_{\overline{\pi}^*}\left(\tilde{\theta}_k \neq \theta_{i_0} | \theta_{i_0}, p_0\right) \leqq K\gamma^k \quad (k \geqq 1).$$

Theorem 4.1 $\overline{\pi}^*$ は適応平均最適政策である.

(証明) 任意の $k \geqq 1$ に対して, $\overline{\tau}^{(k)} = (\tau^*(\tilde{\theta}_0), \tau^*(\tilde{\theta}_1), \ldots, \tau^*(\tilde{\theta}_k), \tau^*(\theta), \tau^*(\theta), \ldots)$, $\overline{\tau}^{(k)}$ に対応する政策を $\overline{\pi}^{(k)} \in \overline{\Pi}$ とする. 記号の簡略化のために, $P(\cdot|\overline{\pi}) = P_{\overline{\pi}}(\cdot|\theta, p_0), E[\cdot|\overline{\pi}] = E_{\overline{\pi}}[\cdot|\theta, p_0], \overline{\pi} \in \overline{\Pi}$ とする. Lemma 4.1 より,

$$P\left(\bigcup_{l=k+1}^{\infty} \left\{\tilde{\theta}_l \neq \theta\right\} \middle| \overline{\pi}^*\right) \leqq \sum_{l=k+1}^{\infty} P\left(\tilde{\theta}_l \neq \theta \mid \overline{\pi}^*\right) \leqq M\gamma^k/(1-\gamma) \quad (k \geqq 1).$$

これより, 次を得る.

(16)
$$P\left(\overline{\pi}^{(k)} \neq \overline{\pi}^{*}\right) \leqq M\gamma^{k}/(1-\gamma).$$

さらに, 次が成り立つ.

(17)
$$\begin{aligned} B_{k+m} &:= \left|\varphi_{k+m}(\overline{\pi}^{(k)}|\theta,p_0) - \varphi_{k+m}(\overline{\pi}^{*}|\theta,p_0)\right| \\ &\leqq \frac{1}{E[s_{k+m}|\overline{\pi}^{*}]}\left|E\left[\sum_{t=0}^{s_{k+m}} c(\tilde{p}_t,\tilde{a}_t) \mid \overline{\pi}^{(k)}\right] - E\left[\sum_{t=0}^{s_{k+m}} c(\tilde{p}_t,\tilde{a}_t) \mid \overline{\pi}^{*}\right]\right| \\ &\quad + \frac{1}{E[s_{k+m}|\overline{\pi}^{(k)}]E[s_{k+m}|\overline{\pi}^{*}]}\left|E[s_{k+m}|\overline{\pi}^{*}] - E[s_{k+m}|\overline{\pi}^{(k)}]\right| \times E\left[\sum_{t=0}^{s_{k+m}} c(p_t,a_t) \mid \overline{\pi}^{*}\right]. \end{aligned}$$

ここで, $M_1 := \max_{\overline{\theta}\in\Theta} E[\tau^*(\overline{\theta})|\overline{\pi}^*], M_2 := \min_{\overline{\theta}\in\Theta} E[\tau^*(\overline{\theta})|\overline{\pi}^*]$ とおくとき, $\overline{\pi}^*$ の定義と式 (16) より, 定数 c_1, c_2, c_3 が存在して次が成り立つ:

$$\begin{aligned} &\left|E\left[\sum_{t=0}^{s_{k+m}} c(\tilde{p}_t,\tilde{a}_t) \mid \overline{\pi}^{(k)}\right] - E\left[\sum_{t=0}^{s_{k+m}} c(\tilde{p}_t,\tilde{a}_t) \mid \overline{\pi}^{*}\right]\right| \leqq c_1 m\gamma^k/(1-\gamma), \\ &\left|E[s_{k+m}|\overline{\pi}^{(k)}] - E[s_{k+m}|\overline{\pi}^{*}]\right| \leqq c_2 m\gamma^k/(1-\gamma). \end{aligned}$$

故に, 式 (17) より,

$$B_{k+m} \leqq c_1\frac{m}{(k+m)M_2}\gamma^k/(1-\gamma) + c_3\frac{m(k+m)}{((k+m)M_2)^2}\gamma^k/(1-\gamma) \quad (k,m \geqq 1).$$

これより, 任意の $m \geqq 1$ に対して, $\lim_{k\to\infty} B_{k+m} = 0$. 故に, $\varphi(\overline{\pi}^*|\theta,p_0) = \limsup_{l\to\infty}\varphi_l(\overline{\pi}^*|\theta,p_0) = \limsup_{m\to\infty}\varphi_{k+m}(\overline{\pi}^*|\theta,p_0) = \varphi(\overline{\pi}^{(k)}|\theta,p_0)$. φ の定義より, $\varphi(\overline{\pi}^{(k)}|\theta,p_0) = \varphi((\pi^*(\theta))^\infty|\theta,p_0) = \inf_{\pi\in\Pi}\varphi(\pi|\theta,p_0)$. これより, $\varphi(\overline{\pi}^*|\theta,p_0) = \inf_{\pi\in\Pi}\varphi(\pi|\theta,p_0)$.∎

§3 の Theorem 3.4 で各 $\theta\in\Theta$ に対して, θ-割引き最適な停止政策 $(\overline{\tau}(\theta))^\infty$ が存在する. ここで, 新しい停止政策 $\widetilde{\tau} = (\overline{\tau}(\tilde{\theta}_0), \overline{\tau}(\tilde{\theta}_1), \ldots)$ を構成する. $\widetilde{\tau}$ に対応する政策を $\widetilde{\pi}$ とする.

Theorem 4.2 政策 $\widetilde{\pi}$ は, 学習割引き最適である.

(証明) $\widetilde{\tau}^{(k)} = (\overline{\tau}(\tilde{\theta}_0), \overline{\tau}(\tilde{\theta}_1), \ldots, \overline{\tau}(\tilde{\theta}_k), \overline{\tau}(\theta), \overline{\tau}(\theta), \ldots)$ として, $\widetilde{\tau}^{(k)}$ に対応する政策を $\widetilde{\pi}^{(k)}$ とする. Theorem 4.1 の証明と同じようにして, 次を得る.

$$P_{\widetilde{\pi}^*}\left(\widetilde{\pi}\{H_{s_k}, G_{s_k}\} \neq \widetilde{\pi}^*\right) \leqq M\gamma^k/(1-\gamma) \quad (k \geqq 1).$$

上の事実から, ある定数 $c_4 > 0$ が存在して, 次が成り立つ.

$$\begin{aligned} &\left|E\left[v(\widetilde{\pi}\{H_{s_k}, G_{s_k}\}|\theta,p_0) - v(\widetilde{\pi}^*|\theta,p_0)\right]\right| \\ &\leqq \frac{c_4}{1-\beta}\gamma^k/(1-\gamma) \to 0 \quad (\text{as } k\to\infty). \end{aligned}$$

これにより, 証明が終わる.∎

Theorem 4.1, 4.2で構成された政策$\pi^*, \tilde{\pi}$を用いて, 適応最適な管理図を作ることができる. たとえば, 平均期待コスト基準のもとでは次のようになる.

θ_i-最適管理図 $(i = 1, 2, \ldots, r)$:
$p_k < p^*_{\theta_i}$ ならば "continue," $p_k \geqq p^*_{\theta_i}$ ならば "stop and research"

適応管理図:
$\tilde{\theta}_k = \theta_i$ ならば θ_i-最適管理図を用いて
"continue," あるいは "stop and research" を選択せよ.

References

[1] J. A. Bather. Control charts and minimization of costs. *J. Roy. Statist. Soc. Ser. B*, 25:49–80, 1963.

[2] Robert V. Baxley, Jr. An application of variable sampling interval control charts. *Journal of Quality Technology*, 27(4):275–282, 1995.

[3] Dimitri P. Bertsekas and Steven E. Shreve. *Stochastic optimal control*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1978. The discrete time case.

[4] Morris H. DeGroot. *Optimal statistical decisions*. McGraw-Hill Book Co., New York, 1970.

[5] T. S. Ferguson. Optimal stopping and applications (electronic texts). `http://www.math.ucla.edu/~tom/Stopping/Contents.html`.

[6] M. A. Girshick and Herman Rubin. A Bayes approach to a quality control model. *Ann. Math. Statistics*, 23:114–125, 1952.

[7] O. Hernández-Lerma. *Adaptive Markov control processes*, volume 79 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1989.

[8] Masami Kurano. Adaptive policies in Markov decision processes with uncertain transition matrices. *J. Inform. Optim. Sci.*, 4(1):21–40, 1983.

[9] Viliam Makis. Multivariate Bayesian control chart. *Oper. Res.*, 56(2):487–496, 2008.

[10] E. L. Porteus and A. Angelus. Opportunities for improved statistical process control. *Management Sci.*, 43:1214–1228, 1997.

[11] Martin L. Puterman. *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons Inc., New York, 1994. A Wiley-Interscience Publication.

[12] Marion R. Reynolds, Jr., Jesse C. Arnold, Raid W. Amin, and Joel A. Nachlas. $\overline{X}$ charts with variable sampling intervals. *Technometrics*, 30(2):181–192, 1988.

[13] W.A. Schewhart. *Economic Control of Quality of Manufactured Product.* Van Nostrand, 1931. 白崎 文雄（訳）,「工業製品の経済的品質管理」, 日本規格協会, 1951.

[14] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction.* Adaptive Computation and Machine Learning. MIT Press, Cambridge, MA, 1998.

[15] George Tagaras. A dynamic programming approach to the economic design of X-charts. *IIE Trans.*, 26(3):48–56, 1994.

[16] George Tagaras. Dynamic control charts for finite production runs. *Europian J. Oper. Res.*, 91:38–55, 1998.

[17] George Tagaras and Yiannis Nikolaidis. Comparing the effectiveness of various Bayesian $\overline{X}$ control charts. *Oper. Res.*, 50(5):878–888, 2002.

[18] Howard M. Taylor. Markovian sequential replacement processes. *Ann. Math. Statist.*, 36:1677–1694, 1965.

[19] Howard M. Taylor. Statistical control of a Gaussian process. *Technometrics*, 9:29–41, 1967.

[20] K. M. van Hee. *Bayesian control of Markov chains*, volume 95 of *Mathematical Centre Tracts.* Mathematisch Centrum, Amsterdam, 1978.

[21] William H. Woodall, Thomas J. Lorenzen, and Lonnie C. Vance. Weaknesses of the economic design of control charts. *Technometrics*, 28(4):408–410, 1986.

[22] 佐々木 稔. **適応型の管理図とその品質マネジメントシステムへの応用に関する研究**. 放送大学大学院文化科学研究科修士論文, 2004.