有限未知パラメータを含む線形確率系に 対する開ループ帰還最適制御†

石 原 正*・阿 部 健 一*・竹 田 宏*

The Open Loop Feedback Optimal Control for Unknown Linear
Systems with Finite Parameter Set

Tadashi ISHIHARA*, Ken-ichi ABE* and Hiroshi TAKEDA*

The optimal control problem of linear stochastic systems with unknown parameters is essentially nonlinear, so that an explicit solution of dynamic programming is not available. Thus, various suboptimal approaches have been suggested in treating this class of problems.

In this paper, a suboptimal algorithm based on the use of the open-loop-feedback-optimal (OLFO) method is derived on the assumption that unknown parameters take finite values. It is shown that the resulting control system consists of

1) an optimal state-parameter estimator formed by the parallel operation of Kalman filters;

2) an, OLFO adaptive feedback gain which is computed on-line based on the a posteriori probabilities of unknown parameters.

The OLFO algorithm is compared with existing algorithms by using the Monte Carlo method. The results are summarized as follows.

1) In the average performance, the OLFO control is a little superior to others for stable systems but inferior for unstable systems.

2) For the both stable and unstable systems, the standard deviation of the performance of the OLFO control is smaller than that of others.

3) The convergence rate of the a posteriori probabilities of unknown parameters is primarily dependent on the stability of the system rather than the choice of the algorithm. For unstable systems, the convergence rate is faster than that of stable systems.

1. はじめに

未知パラメータを含む線形確率系の最適制御問題は、理論的に興味深い問題であり、また応用上も重要である。この問題は、本質的に非線形問題であり最適制御を実現することが困難であるため、実現可能な近似最適制御を求める試みがなされてきた。特に、未知パラメータが連続分布に従う確率変数である系に対し、多くの近似最適制御が提案されている^{1)~5)}。これらの近似最適制御において、状態が完全に観測できる場合¹⁾、あるいは、駆動行列のみが未知パラメータを含む場合²⁾には、カルマンフィルタにより状態および未知パラメータの最適推定を行うことができる。しかし、システム行列が未知パラメータを含む場合には、最適推定を実現することは困難となり、拡張カルマンフィルタなどの近似フィルタにより推定を行わざるを得ない。この場合、制御成績は、推定精度に依存するため、近似フィルタの精度を十分検討しておく必要がある。

一方、未知パラメータが有限個の値を取る確率変数である場合には、駆動行列のみならず、システム行列、観測行列、外乱の統計量などが未知パラメータを含むような系に対しても最適推定を実現できる。このような系に対し、Saridis-Dao⁶⁾は、ベルマンの方程式に最適評価関数値の下界を用いることにより近似最適制御を求めている。また、Deshpande ら⁷⁾は、各パラメータ値に対応する最適制御を未知パラメータの事後分布について平均することにより、近似最適制御を得ることを提案している。これらの近似最適制御は、計算が容易であるが、近似の意味が明確でない。

本論文では、未知パラメータが有限個の値を取る確率変数である、線形確率系に対する近似最適制御の一つとして、各時刻において得られているデータに基づ

† 電子通信学会情報部門全国大会で発表 (昭 52・8)

* 東北大学工学部 仙台市荒巻字青葉

* Faculty of Engineering, Tohoku University, Sendai
(Received November 15, 1977)

いて開ループ最適制御を決定する。いわゆる開ループ帰還最適 (Open Loop Feedback Optimal, OLFO) 制御について考察する。OLFO 制御は、各時刻において、将来の観測を無視するという明確な近似の意味を有しているため、確率系に対する近似最適制御として重要視され、さらに高度の機能 (いわゆる能動的学習機能⁴⁾) を備えた制御を実現するための基礎となるものである。未知パラメータが連続分布に従う場合の OLFO 制御は、すでに Tse-Athans²⁾, Ku-Athans³⁾ により扱われているが、有限未知パラメータの場合の研究はまだ報告されていない。本論文では、まず、有限未知パラメータの場合における OLFO 制御を算出するアルゴリズムを求め、次いで OLFO 制御の性質について考察する。最後に本論文で求めた OLFO 制御と Saridis-Do⁶⁾ および Deshpande ら⁷⁾ の近似最適制御との比較をシミュレーションにより行い検討を加える。

2. 問題の記述

つぎのような未知パラメータを含む離散時間線形確率系を考える。

$$x(t+1) = A(t, \theta)x(t) + B(t, \theta)u(t) + C(t, \theta)w(t) \\ t=0, 1, \dots, N-1 \quad (1)$$

$$y(t) = H(t, \theta)x(t) + G(t, \theta)v(t) \\ t=1, 2, \dots, N \quad (2)$$

ここで、 $x(t)$ は n 次元状態ベクトル、 $u(t)$ は p 次元制御ベクトル、 $y(t)$ は m 次元観測ベクトルである。 $w(t)$, $v(t)$ はそれぞれ力学系、観測系に対する外乱ベクトルであり、それぞれ r_1 , r_2 次元である。 θ は未知パラメータベクトルであり、 $\{\theta_1, \theta_2, \dots, \theta_M\}$ のいずれかを取る確率ベクトルで、事前確率 $\{p(\theta_i), i=1, 2, \dots, M\}$ が与えられているものとする。 $x(0)$, $\{w(t), t=0, 1, \dots, N-1\}$, $\{v(t), t=1, 2, \dots, N\}$ は、互いに独立な正規性確率ベクトルであり、つぎの統計量に従う。

$$\left. \begin{aligned} E[x(0)] &= \hat{x}(0|0), \text{cov}[x(0)] = \Sigma(0|0) \\ E[w(t)] &= 0, \text{cov}[w(t)] = I_{r_1} \\ E[v(t)] &= 0, \text{cov}[v(t)] = I_{r_2} \end{aligned} \right\} \quad (3)$$

ここで、 $E[\cdot]$, $\text{cov}[\cdot]$ は、平均および共分散を意味する。 I_{r_i} ($i=1, 2$) は、 $r_i \times r_i$ 単位行列である。さらに、 $G(t, \theta_i)G'(t, \theta_i) > 0$ (正定行列), $i=1, 2, \dots, M$ であるものと仮定する。記号' は転置を表す。

評価規準として、つぎのような 2 次形式評価関数が与えられているものとする。

$$J \triangleq E \left\{ \|x(N)\|_F^2 + \sum_{s=0}^{N-1} (\|x(s)\|_Q^2 + \|u(s)\|_R^2) \right\}$$

(4)

ここで、 $n \times n$ 行列 F , $Q(s)$ は非負定対称行列であり、 $p \times p$ 行列 $R(s)$ は正定対称行列である。

以上の設定のもとでは、状態および未知パラメータの事後分布を正確に知ることが可能である。すべての事前情報を Σ^0 で表し、 t -時刻において制御に利用できる情報を

$$\Sigma^t \triangleq \{\Sigma^0; y(1), y(2), \dots, y(t); \\ u(0), u(1), \dots, u(t-1)\} \quad (5)$$

で表すことにする。 Σ^t が与えられた時の状態および未知パラメータの結合確率密度関数 $p(x(t), \theta_i | \Sigma^t)$ は、Bayes rule により、つぎのように書ける。

$$p(x(t), \theta_i | \Sigma^t) = p(x(t) | \theta_i, \Sigma^t) p(\theta_i | \Sigma^t) \quad (6)$$

ここで、 $p(x(t) | \theta_i, \Sigma^t)$ は正規分布であり、その平均値および分散をそれぞれ $\hat{x}(t | t, i)$, $\Sigma(t | t, i)$ で表すと、これらは $\theta = \theta_i$ とした場合のカルマンフィルタにより求められる。また、未知パラメータの事後確率 $p(\theta_i | \Sigma^t)$ は、Bayes rule により、つぎの関係を満たす。

$$p(\theta_i | \Sigma^{t+1}) = \frac{p(y(t+1) | \theta_i, \Sigma^t, u(t)) p(\theta_i | \Sigma^t)}{\sum_{j=1}^M p(y(t+1) | \theta_j, \Sigma^t, u(t)) p(\theta_j | \Sigma^t)} \quad (7)$$

ただし、 $p(\theta_i | \Sigma^0) = p(\theta_i)$, $i=1, 2, \dots, M$

ここで、 $p(y(t+1) | \theta_i, \Sigma^t, u(t))$ は平均値および分散が次式で与えられる正規分布密度関数である。

$$\hat{y}(t+1 | t, i) \triangleq H(t+1, \theta_i)[A(t, \theta_i)\hat{x}(t | t, i) \\ + B(t, \theta_i)u(t)] \quad (8)$$

$$\Sigma_y(t+1 | t, i) \triangleq H(t+1, \theta_i)[A(t, \theta_i)\Sigma(t | t, i)A'(t, \theta_i) \\ + C(t, \theta_i)C'(t, \theta_i)]H_i(t+1, \theta_i) \\ + G(t+1, \theta_i)G'(t+1, \theta_i) \quad (9)$$

したがって、未知パラメータの事後確率は $\theta = \theta_i$ ($i=1, 2, \dots, M$) とした M 個のカルマンフィルタより得られる $\hat{x}(t | t, i)$, $\Sigma(t | t, i)$ ($i=1, 2, \dots, M$) を (8), (9) 式に用いることにより、(7) 式からオンラインで容易に計算される。結局、事後分布は有限個の統計量 $\{\hat{x}(t | t, i), \Sigma(t | t, i), p(\theta_i | \Sigma^t) \mid i=1, 2, \dots, M\}$ (10)

により完全に記述される。

以上のように、事後分布は完全に知り得るが、ベルマンの方程式を解き (4) 式を最小とする制御を求めることは困難である。ここでは、近似最適制御の一つとして開ループ帰還最適 (OLFO) 制御を求める。すなわち、各時刻で得られている情報に基づいて開ループ最適制御を求める。

3. OLFO 制御問題の解

t -時刻に得られているデータ $\mathcal{Z}^t$ に基づく開ループ最適制御は、 t -時刻以降の制御 $\{u(s|t), s=t, t+1, \dots, N-1\}$ がすべて $\mathcal{Z}^t$ のみに依存するものとし、開ループ評価規準

$$J_{\text{OLF}}(\mathcal{Z}^t) \triangleq E \left\{ \|x(N)\|^2_F + \sum_{s=t}^{N-1} (\|x(s)\|^2_{Q(s)} + \|u(s|t)\|^2_{R(s)}) \mid \mathcal{Z}^t \right\} \quad (11)$$

を最小とする制御である。(11)式は、未知パラメータの事後確率 $\{p(\theta_i | \mathcal{Z}^t), i=1, 2, \dots, M\}$ を用いてつぎのように書ける。

$$\begin{aligned} J_{\text{OLF}}(\mathcal{Z}^t) &= \sum_{i=1}^M E \left\{ \|x(N)\|^2_F + \sum_{s=t}^{N-1} (\|x(s)\|^2_{Q(s)} + \|u(s|t)\|^2_{R(s)}) \mid \theta_i, \mathcal{Z}^t \right\} p(\theta_i | \mathcal{Z}^t) \\ &= \sum_{i=1}^M \left\{ \|\hat{x}(N|t, i)\|^2_F + \sum_{s=t}^{N-1} (\|\hat{x}(s|t, i)\|^2_{Q(s)} + \|u(s|t)\|^2_{R(s)}) \right\} p(\theta_i | \mathcal{Z}^t) \\ &\quad + \sum_{i=1}^M \left\{ \text{tr } F \Sigma(N|t, i) + \sum_{s=t}^{N-1} \text{tr } Q(s) \Sigma(s|t, i) \right\} p(\theta_i | \mathcal{Z}^t) \end{aligned} \quad (12)$$

ここで、

$$\hat{x}(s|t, i) \triangleq E[x(s) | \theta_i, \mathcal{Z}^t, u(\tau|t), \tau=t, t+1, t+2, \dots, s-1] \quad (13)$$

$$\Sigma(s|t, i) \triangleq \text{cov}[x(s) | \theta_i, \mathcal{Z}^t, u(\tau|t), \tau=t, t+1, \dots, s-1] \quad (14)$$

であり、それぞれつぎの関係式を満たす。

$$\hat{x}(s+1|t, i) = A(s, \theta_i) \hat{x}(s|t, i) + B(s, \theta_i) u(s|t) \quad s=t, t+1, \dots, N-1 \quad (15)$$

$$\begin{aligned} \Sigma(s+1|t, i) &= A(s, \theta_i) \Sigma(s|t, i) A'(s, \theta_i) \\ &\quad + C(s, \theta_i) C'(s, \theta_i) \quad s=t, t+1, \dots, N-1 \end{aligned} \quad (16)$$

$\{\Sigma(s|t, i), s=t, t+1, \dots, N, i=1, 2, \dots, M\}$ は、制御に依存しないオフラインで計算可能な量であるから、(12)式第2項は、 $J_{\text{OLF}}(\mathcal{Z}^t)$ の最小化には関係しない。したがって(12)式の最小化は、(15)式の拘束のもとで

$$\tilde{J}_{\text{OLF}}(\mathcal{Z}^t) \triangleq \sum_{i=1}^M \left\{ \|\hat{x}(N|t, i)\|^2_F \right.$$

$$\left. + \sum_{s=t}^{N-1} (\|\hat{x}(s|t, i)\|^2_{Q(s)} + \|u(s|t)\|^2_{R(s)}) \right\} p(\theta_i | \mathcal{Z}^t) \quad (17)$$

を最小化することと等価である。この最小化は、以下のように容易に行うことができる。まず、つぎの諸量を定義する。

$$\hat{\mathbf{X}}(s|t) \triangleq (\hat{x}'(s|t, 1), \hat{x}'(s|t, 2), \dots, \hat{x}'(s|t, M))' \quad (18)$$

$$\mathbf{A}(s) \triangleq \text{diag. } [A(s, \theta_1), A(s, \theta_2), \dots, A(s, \theta_M)] \quad (19)$$

$$\mathbf{B}(s) \triangleq [B'(s, \theta_1), B'(s, \theta_2), \dots, B'(s, \theta_M)]' \quad (20)$$

$$\mathbf{Q}(s|t) \triangleq \text{diag. } [p(\theta_1 | \mathcal{Z}^t) Q(s), p(\theta_2 | \mathcal{Z}^t) Q(s), \dots, p(\theta_M | \mathcal{Z}^t) Q(s)] \quad (21)$$

$$\mathbf{F}(N|t) \triangleq \text{diag. } [p(\theta_1 | \mathcal{Z}^t) F, p(\theta_2 | \mathcal{Z}^t) F, \dots, p(\theta_M | \mathcal{Z}^t) F] \quad (22)$$

(15)式は、(18)~(20)式を用いるとつぎのように書ける。

$$\begin{aligned} \hat{\mathbf{X}}(s+1|t) &= \mathbf{A}(s) \hat{\mathbf{X}}(s|t) + \mathbf{B}(s) u(s|t), \\ s &= t, t+1, \dots, N-1 \end{aligned} \quad (23)$$

さらに、 $\tilde{J}_{\text{OLF}}(\mathcal{Z}^t)$ は(18), (21), (22)式を用いて

$$\begin{aligned} \tilde{J}_{\text{OLF}}(\mathcal{Z}^t) &= \|\hat{\mathbf{X}}(N|t)\|^2_{\mathbf{F}(N|t)} \\ &\quad + \sum_{s=t}^{N-1} (\|\hat{\mathbf{X}}(s|t)\|^2_{\mathbf{Q}(s|t)} + \|u(s|t)\|^2_{R(s)}) \end{aligned} \quad (24)$$

と書ける。したがって、(24)式を(23)式の拘束のもとで最小化する確定的問題を解くことにより開ループ最適制御が求められる。これは、よく知られた線形-2次形式の問題である。 $0 \leq p(\theta_i | \mathcal{Z}^t) \leq 1, F \geq 0, Q(s) \geq 0$ (非負定行列) から $\mathbf{F}(N|t) \geq 0, \mathbf{Q}(s|t) \geq 0$ さらに $R(s) > 0$ に注意すれば、解の一意性も保証され、求める開ループ最適制御 $\{u^*(s|t), s=t, t+1, \dots, N-1\}$ は、つぎのように求められる。

$$\begin{aligned} u^*(s|t) &= -[R(s) + \mathbf{B}'(s) \mathbf{S}(s+1|t) \mathbf{B}(s)]^{-1} \\ &\quad \times \mathbf{B}'(s) \mathbf{S}(s+1|t) \mathbf{A}(t) \hat{\mathbf{X}}(s|t) \end{aligned} \quad (25)$$

ここで、 $\mathbf{S}(s+1|t)$ はつぎの関係式から求められる。

$$\begin{aligned} \mathbf{S}(\tau|t) &= \mathbf{Q}(\tau|t) + \mathbf{A}'(\tau) [\mathbf{S}(\tau+1|t) \\ &\quad - \mathbf{S}(\tau+1|t) \mathbf{B}(\tau) (R(\tau) + \mathbf{B}'(\tau) \mathbf{S}(\tau \\ &\quad + 1|t) \mathbf{B}(\tau))^{-1} \mathbf{B}'(\tau) \mathbf{S}(\tau+1|t) \mathbf{A}(\tau) \end{aligned} \quad (26)$$

$$\tau = t, t+1, \dots, N-1 \quad (27)$$

$$\mathbf{S}(N|t) \triangleq \mathbf{F}(N|t) \quad (27)$$

t -時刻における OLFO 制御は、

$$u_{\text{OLFO}}(t) \triangleq u^*(t|t) \quad (28)$$

として求められる。

4. OLFO 制御の性質

本節では、前節で導いた OLFO 制御のもつ性質について考察する。

[1] 未知パラメータが連続分布に従う確率変数である場合の OLFO 制御は、駆動ベクトルが未知である 1 入力系に対しては Tse-Athans²⁾により Matrix Minimum Principle を用いて求められている。Ku-Athans³⁾は、システム行列および駆動ベクトルが未知である 1 入力系にまで Tse-Athans の結果を拡張したが、この場合最適推定を実現できないうえにシステム行列はコンパニオン形に限られ、しかもアルゴリズムが複雑である。前節で求めた有限未知パラメータの場合の OLFO 制御は、1 入力系に限られず未知パラメータの含まれ方に制限がないうえに、アルゴリズムが簡単であるという特徴がある。

[2] (25)式における $\hat{X}(t|t)$ は、各パラメータ値に対する M 個のカルマンフィルタの並列計算により求められる。また、 $S(t+1|t)$ は、未知パラメータの事後確率 $\{p(\theta_i|\mathcal{Z}^t), i=1, 2, \dots, M\}$ に基づいて、(26)、(27)式によりオンラインで計算される必要がある。したがって、Saridis-Dao⁶⁾、Deshpande ら⁷⁾の近似最適制御と比べオンラインでの計算量は多い。 $S(\cdot|t)$ の計算量は、状態の次元 n 、未知パラメータの取りうる値の個数 M が多くなるにつれ増加するが、逆行列の計算量は、入力個数のみに依存し、増加しない。

[3] システム行列および駆動行列が未知パラメータを含まないときには、オンラインで $S(\cdot|t)$ を計算する必要はない。この場合、比較的簡単な計算により、OLFO 制御はつぎのように求められる。

$$u_{\text{OLFO}}(t) = L(t) \sum_{i=1}^M x(t|t, i) p(\theta_i|\mathcal{Z}^t) \quad (29)$$

ここで、 $L(t)$ はシステム行列、駆動行列および評価関数の重み $F, Q(s), R(s)$ から定まる線形-2 次形式問題の最適ゲインであり、オフラインで計算可能である。 $\sum_{i=1}^M x(t|t, i) p(\theta_i|\mathcal{Z}^t)$ は、状態の最適推定値であり、(29)式は分離定理の成立を仮定した場合の制御と一致する。また、Saridis-Dao⁶⁾、Deshpande ら⁷⁾の近似最適制御もこの場合(29)式に一致することが容易に確かめられる。

[4] $p(\theta_i|\mathcal{Z}^t) = \delta_{ij}$, $i=1, 2, \dots, M$ (δ_{ij} はクロネッカーのデルタ) すなわち、パラメータが既知で θ_j である場合には、OLFO 制御は θ_i に対する L - Q - G 問題の最適制御と一致する。したがって、無限時間問題において、 $p(\theta_i|\mathcal{Z}^t) \rightarrow \delta_{ij}$ ($t \rightarrow \infty$) であれば、OLFO 制御は漸近的に最適である。

[5] 最適評価関数値を J^* 、OLFO 制御を用いた場合の評価関数値を J_{OLFO} とすると、明らかに

$$J^* \leq J_{\text{OLFO}} \quad (30)$$

の関係が成立する。 J^*, J_{OLFO} 共に解析的に求めることは困難であるが、つぎのように容易に計算できる下界 J_L が存在する。

$$J_L \leq J^* \leq J_{\text{OLFO}} \quad (31)$$

$$J_L \triangleq \sum_{i=1}^M J^*(\theta_i) p(\theta_i) \quad (32)$$

ここで、 $J^*(\theta_i)$ は、 $\theta = \theta_i$ に対する L - Q - G 問題の最適評価関数値である。この下界は、シミュレーションにより近似最適制御の評価を行う場合、制御の良さを判断する規準として用いることができる。未知パラメータが連続分布に従う確率変数である場合には、

Table 1 Summary of the Monte Carlo simulation

θ_1	θ_2	Algorithm	Average Performance	Standard Deviation of Performance	Average Miss Distance Squared	Theoretical Lower Bound
1	$\begin{bmatrix} 1.6 \\ 0.5 \end{bmatrix}$	OLFO	287.63	97.00	1.1955	73.57
		Saridis	192.65	158.00	1.1984	
		Deshpande	195.23	166.30	1.1986	
2	$\begin{bmatrix} 1.5 \\ 0.5 \end{bmatrix}$	OLFO	227.30	87.48	1.2170	82.95
		Saridis	174.08	104.50	1.2181	
		Deshpande	174.07	107.19	1.2206	
3	$\begin{bmatrix} 0.9 \\ 0.5 \end{bmatrix}$	OLFO	26.51	38.96	0.5087	13.94
		Saridis	28.00	44.64	0.5442	
		Deshpande	26.69	40.71	0.5161	
4	$\begin{bmatrix} 0.9 \\ 0.7 \end{bmatrix}$	OLFO	25.60	32.90	0.4988	13.81
		Saridis	25.72	33.50	0.5022	
		Deshpande	25.99	34.11	0.5080	

(32)式に相当する下界を計算するのは困難である。

5. シミュレーションと検討

本節では、3節で導いた OLFO 制御と Saridis-Dao⁶⁾ および Deshpande⁷⁾ の近似最適制御との比較をシミュレーションにより行った結果を報告する。

対象とする系は

$$x(t+1) = ax(t) + bu(t) + cw(t), \quad t=0, 1, \dots, 19. \quad (33)$$

$$y(t) = x(t) + qv(t), \quad t=1, 2, \dots, 20 \quad (34)$$

なる一変数系とし、評価規準は

$$J = E \left\{ 50x^2(20) + \sum_{s=0}^{19} u^2(s) \right\} \quad (35)$$

で与えられるものとする。未知パラメータベクトルは、 $\theta = (a, b)'$ で、 $\theta_i = (a_i, b_i)'$ ($i=1, 2$) のいずれかの値をとるものとする。多くの例についてシミュレーションを行ったが、特に $p(\theta_1) = p(\theta_2) = 0.5$, $x(0|0) = 3.0$, $\Sigma(0|0) = 1.0$, $c = 0.25$, $q = 1.0$ の場合に4組の (θ_1, θ_2) について、それぞれ100個のサンプルをもと

Fig. 1 Evolution of the state for the unstable system (one sample run for case 1)

Fig. 2 Control history (one sample run for case 1)

に得られた結果を Table 1 に示す。表中の理論的下界とは、(32)式で与えられる下界である。

シミュレーションにより得られた結果を要約すると、つぎようになる。

[1] OLFO 制御は、不安定系に対しては、他の近似最適制御と比べ制御成績が悪い。これは、主に評価関数中の制御に関する項が大きい値を取ることが原因している。終端誤差は、他の近似最適制御とほとんど同じである。Fig. 1, Fig. 2, Fig. 3 に、case 1 の場合の同一乱数列に対する状態、制御、未知パラメータの事後確率の推移を示す。いずれの図とも、Saridis-Dao および Deshpande らの近似最適制御が、ほとんど同じ傾向を示しているのに対し、OLFO 制御は、かなり異なる傾向を示している。

[2] 安定系に対しては、三つの近似最適制御間には、不安定系の場合ほど顕著な差は認められないが、OLFO 制御が多少良い制御成績を示す。Saridis-Dao と Deshpande らの近似最適制御は、きわめて近い傾

Fig. 3 Evolution of the a posteriori probability $p(\theta_i | Z^t)$ for the unstable system (one sample run for case 1)

Fig. 4 Evolution of the state for the stable system (one sample run case 3)

Fig. 5 Control history (one sample run for case 3)

Fig. 6 Evolution of the a posteriori probability $p(\theta_2 | Z^t)$ for the stable system (one sample run for case 3)

向を示すので、Deshpandeらの近似最適制御とOLFO制御に対する状態、制御、事後確率の推移例を Fig. 4, Fig. 5, Fig. 6 に示す。

〔3〕 未知パラメータの事後確率の収束の速さは、各近似最適制御により多少異なるものの、系の安定性が大きく影響する。不安定系においては、収束はきわめて速く、安定系では遅い。これは、不安定系の場合、状態および制御が比較的大きい値を取るため、観測値の信号対雑音比が安定系の場合と比べて良く、未知パラメータに関する情報を抽出しやすいためと考えられる。

〔4〕 OLFO 制御は、他の近似最適制御を用いた場合に比べ評価関数値の標準偏差が小さい。この性質は、Fig. 2, Fig. 5 が示すように OLFO 制御が他の近似最適制御と比べ大きな値をとり、しかも変化が大きく、パラメータ推定に関しては望ましい制御となっ

ていることによるものと考えられる。特に不安定系においてこのことは顕著である。

以上の結果は、未知パラメータが連続分布に従う場合に、OLFO 制御と分離定理の成立を仮定した制御との比較を行った Ku-Athans³⁾ のシミュレーション結果と同様の傾向を示している。しかし、そこでは拡張カルマンフィルタにより推定を行っているため、推定誤差の発散を防ぐ必要上(33)式における a の初期分散がきわめて小さい場合しか扱われていない。

6. おわりに

有限未知パラメータを含む線形確率系に対するOLFO 制御のアルゴリズムを導き、その性質について考察してきた。パラメータが有限個の値をとる場合には、系へのパラメータの関与の仕方、入力の数などの点で、連続パラメータよりも広い線形確率系のクラスに対して OLFO 制御が簡単に求められる。また、Saridis-Dao, Deshpande らの近似最適制御との比較をシミュレーションによって行った結果、OLFO 制御は、評価関数値のばらつきが小さいことが確められた。

本論文における OLFO 制御および比較を行った近似最適制御は、いずれも能動的学習機能⁴⁾を備えていない。今後、この機能を備えた、より最適に近い制御を構成する必要があるであろう。

参考文献

- 1) J.B. Farison, R.E. Graham and R.C. Shelton: Identification and Control of Linear Discrete Systems, IEEE Trans., **AC-14-4**, 438/442 (1967)
- 2) E. Tse and M. Athans: Adaptive Stochastic Control for a Class of Linear Systems, IEEE Trans., **AC-17-1**, 38/51 (1972)
- 3) R. Ku and M. Athans: On the Adaptive Control of Linear Systems Using the Open-Loop-Feedback-Optimal Approach, IEEE Trans., **AC-18-5**, 489/493 (1973)
- 4) E. Tse and Y. Barshalom: An Actively, Adaptive Control for Linear Systems with Random Parameters via Dual Control Approach, IEEE Trans., **AC-18-2**, 109/117 (1973)
- 5) 砂原, 大住, 立森: 線形確率システムの適応制御の方法, 計測自動制御学会論文集, **12-6**, 666/673 (1976)
- 6) G. Saridis and T. Dao: A Learning Approach to the Parameter-Adaptive-Self-Organizing Control Problem, Automatica, **8**, 589/597 (1972)
- 7) J. Deshpande, T. Upadhyay and D. Lainiotis: Adaptive Control of Linear Stochastic Systems, Automatica, **9**, 107/115 (1973)