離散時間形線形制御系における受動的適応制御方策

明 石 一*・今 井 弘 之*

Passive Adaptive Control Strategies in Discrete Time Linear Control Systems

Hajime AKASHI* and Hiroyuki IMAI*

In this paper, a passive adaptive control problem is considered from the Bayesian approach. It is shown that the problem can be understood as one of the special cases of a multi-stage decision problem with a non-classical information structure. A suboptimal passive adaptive control problem is formulated, and an algorithm which produces the passive adaptive control strategy is obtained, for the discrete time linear system with some unknown but constant parameters. A numerical example is presented.

1. 緒 言

多くの場合、制御の対象となるプラントの数学的モデルは未知パラメータを含んでいる。このような場合に取りうる一つの方式はプラントの運転中に得られる観測情報を積極的に利用して未知パラメータに関する学習を on-line で行なうことである。この学習の効果をも含めた一般的な最適化問題は適応制御過程として定式化される^{1), 2)}。そして問題は最適性の原理を適用することによって得られる関数方程式を解くことに帰着される。ところが、この方程式を直接解くことはごく限られた場合を除いて非常に困難であるので、多くの場合、準最適方策が用いられる^{3)~6)}。しかしながら、いずれにしてもこのような能動的な適応制御方式をとる限り、on-line での計算量が非常に大きなものとなることは避けられない。

これに対して、事前に十分な実験あるいは調査を行なって未知パラメータの値をできるだけ正確に定めることにし、なおかつ残ったパラメータのもつ不確実さに対しては、on-line での学習を行なわないが、それに対して鈍感な制御方策を用いることによって対処しようという考えがある。このほうが on-line での計算量を小さくおさえることができ、また soft memory

が少なくすむので、広い立場から評価すればより優れていると考えられる場合もある。このような考えは受動的適応制御⁷⁾と呼ばれ、従来では主として感度解析的な立場から接近された⁸⁾。

これに対して、本報告においては Bayesian の立場(注 1) から受動的適応制御方策を求める問題について考察する。そのためにはまず、学習を行なわないということを明確に表現する必要があるであろう。ここでは、これを、制御装置が学習に必要な on-line で得られる観測情報の記憶装置をもたないと表現する。すなわち制御装置は、その時点で得られた観測情報と先験情報(これは設計段階で得られている情報である)のみを用いて決定を行なう。

このような、記憶が不完全である制御過程は非古典的情報構造¹¹⁾をもつ制御過程の一つになるのであるが、このような制御過程に対して問題を一般的に定式化すると(これを以下では問題 PAO と呼ぶ)、それは容易には解けない。そこで、2.において、許される制御方策のクラスを状態変数に関して線形のクラスに限定し、さらに person by person optimal (以下では pbpo と略記する) feed-back ゲイン行列を求める問題として、問題 PA1 を定式化する。なお、考察の対象は離散時間形線形レギュレータに限定する。続いて、3.において、pbpo であるための必要十分条件を求め、それを用いて pbpo であるようなゲイン行列を求めるための繰返し計算アルゴリズムを与える。最後に、数値例を示す。

なお、学習の効果を考慮に入れない制御方策としては、ほかにも certainty equivalent 制御方策²⁾(以下では CE 制御方策と略記する)、あるいは open loop feedback 制御方策^{2), 9), 10)}などがある。前者においては不確実性の存在がまったく考慮に入れられていない。後者においては不確実性の存在は考慮に入れられており、したがってこれは受動的適応制御方策の一

(注 1) 未知パラメータに先験的確率を仮定し、それに関する評価関数の期待値によって議論を進める立場。

* 京都大学工学部 京都市左京区吉田本町

* Faculty of Engineering, Kyoto University, Kyoto
(Received October 18, 1973)

つであると考えられる。しかしながら、これは、未知パラメータの学習はもちろん行なわれないが、状態変数の観測もまだ行なわれない（すなわち、将来まったく情報が入らない）という条件のもとでの適応方策である。これに対して、本報告で扱う受動的適応制御方策は、状態変数の観測は行なわれるが、制御装置はそれらを記憶して未知パラメータの学習に役立てることができないという条件のもとでの最適方策である。

2. 問題の定式化

プラントの動特性方程式ならびに評価関数がつぎのように与えられる場合を考える。

$$\mathbf{x}(i+1) = A(\alpha, i)\mathbf{x}(i) + B(\alpha, i)\mathbf{u}(i) + \xi(i) \quad (1)$$

$$W = \sum_{i=1}^N W(i)$$

$$W(i) = \mathbf{x}^T(i)Q(i)\mathbf{x}(i) + \mathbf{u}^T(i-1)R(i)\mathbf{u}(i-1) \quad (2)$$

ここに、 $\mathbf{x}(i)$ は n 次元の状態ベクトル、 $\mathbf{u}(i)$ は m 次元の制御ベクトル、 $\xi(i)$ は n 次元の外乱ベクトルである。 i は時刻あるいは制御段を表わす。 A および B は未知定数パラメータベクトル α に依存する係数行列である。 $Q(i)$ は対称準正定値行列、 $R(i)$ は対称正定値行列である。 $\mathbf{x}(0)$, α , $\xi(i)$; $i=0, 1, \dots, N-1$ は互いに独立な確率変数でそれらの確率的性質は先験的に与えられているものとする。特に $\xi(i)$ については

$$E[\xi(i)] = 0, \text{cov}[\xi(i)] = V_\xi; i=0, 1, \dots, N-1$$

と与えられているものとする。ただし V_ξ は対称正定値行列とする。以上の先験的情報をまとめて I_0 と書くことにする。状態ベクトル $\mathbf{x}(i)$ は時刻 i において正確に観測されるものとする。

このような制御過程においては、制御の進行に伴って得られる on-line 情報 $\mathbf{x}(0), \mathbf{x}(1), \dots$ を用いて α に関する学習を行なうことが可能である。しかし、ここではすでに述べたように、そのような学習に必要な on-line 情報の記憶装置をもたない場合を考えている。すなわち、制御装置は各時点 i で $\mathbf{x}(i)$ と I_0 のみを用いて $\mathbf{u}(i)$ の決定を行なう。

したがって受動的適応制御問題を一般的に定式すれば

「問題 PAO:

$$J \triangleq E[W | I_0] \quad (3)$$

を最小にする制御方策

$$\mathbf{u}(i) = g(i, \mathbf{x}(i), I_0) \quad i=0, 1, \dots, N-1 \quad (4)$$

を求めよ。」

となるであろう。しかしながら、上に仮定したような情報構造のもとでは $g(i)$ は $\mathbf{x}(i)$ に関して必ずしも線形になるとは限らないことが知られており¹¹⁾、した

がって問題 PAO を解くことは非常に困難である。そこで許される制御方策のクラス $\Omega(g(i))$ を

$$\Omega(g(i)) = \{g(i); g(i) = -K(i, I_0)\mathbf{x}(i)\} \quad (5)$$

で表わされるような、 $\mathbf{x}(i)$ に関して線形のクラスに限定して考える。

さらに、この種の情報構造をもつ最適化問題において通常なされるように、ここでも person by person optimal (pbpo)¹²⁾ 方策を求めることにとどめる。pbpo 方策とは

$$\begin{aligned} & (\forall i)(\forall g(i)) [J(g^*(0), g^*(1), \dots, g^*(i), \dots, g^*(N-1)) \\ & \leq J(g^*(0), \dots, g^*(i-1), g(i), g^*(i+1), \dots, \\ & g^*(N-1))] \end{aligned} \quad (6)$$

であるような方策 $g^*(0), g^*(1), \dots, g^*(N-1)$ のことをいう。明らかに最適方策ならば pbpo 方策であるが、その逆は成立しない。

結局、本報告で扱う受動的適応制御問題はつぎのように定式化される。

「問題 PA1:

$$\begin{aligned} & (\forall i)(\forall K(i)) [J(K^*(0), K^*(1), \dots, K^*(i), \dots, \\ & K^*(N-1)) \leq J(K^*(0), \dots, K^*(i-1), \\ & K(i), K^*(i+1), \dots, K^*(N-1))] \end{aligned} \quad (7)$$

を満す $K^*(i)$; $i=0, 1, \dots, N-1$ を求めよ。ただし J は (3) 式によって与えられる。」

問題 PA1 を解くことによって得られる $K^*(i)$; $i=0, 1, \dots, N-1$ を PA1 ゲイン行列あるいは簡単に PA1 と呼ぶ。

3. 受動的適応制御方策

(1) 式を (3) 式に代入し、そこで

$$\mathbf{u}(i) = -K(i)\mathbf{x}(i)$$

とすると J はつぎのように書かれる。

$$\begin{aligned} J &= E \left[\sum_{i=1}^N W(i) | I_0 \right] \\ &= E \left[\sum_{i=0}^{N-1} \{ (A(i) - B(i)K(i))\mathbf{x}(i) \right. \\ & \quad \left. + \xi(i) \}^T Q(i+1) \{ (A(i) - B(i)K(i))\mathbf{x}(i) \right. \\ & \quad \left. + \xi(i) \} + \mathbf{x}^T(i) K^T(i) R(i+1) K(i) \mathbf{x}(i) | I_0 \right] \end{aligned} \quad (8)$$

J のなかで $K(i)$ に直接関係する項だけを取出してそれを $J(i)$ と書くことにすれば

$$J(i) = E \left[\sum_{j=i+1}^N W(j) | I_0 \right] \quad (9)$$

である。ここで (1) 式を繰返し用い、さらに $\xi(i), \xi(i+1), \dots, \xi(N-1)$ に関する期待値の演算を行なえば、 $J(i)$ はつぎのように書かれる。

$$\begin{aligned}
 J(i) = & E[\mathbf{x}^T(i) \{ (A(i) - B(i)K(i))^T P(i+1)(A(i) \\
 & - B(i)K(i)) + K^T(i)R(i+1)K(i) \} \mathbf{x}(i) | I_0] \\
 & + \sum_{j=i}^{N-1} \text{tr} [P(j+1)V_{\xi}] \quad (10)
 \end{aligned}$$

ただし $P(i)$ はつぎの漸化式によって算出される。

$$\begin{aligned}
 P(j) = & Q(j) + (A(j) - B(j)K(j))^T P(j+1)(A(j) \\
 & - B(j)K(j)) + K^T(j)R(j+1)K(j) \quad (11)
 \end{aligned}$$

ただし

$$P(N) = Q(N)$$

である。ついで、 $\xi(0), \xi(1), \dots, \xi(i-1)$ に関する期待値の演算を行ない、 $\mathbf{x}^T M \mathbf{x} = \text{tr} [\mathbf{x} \mathbf{x}^T M]$ の関係を用いれば

$$\begin{aligned}
 J(i) = & E[\text{tr} [\Psi(i)K^T(i)C(i)K(i) \\
 & - 2\Psi(i)D(i)K(i)] | I_0] + f(i) \quad (12)
 \end{aligned}$$

と書かれる。ここに

$$\begin{aligned}
 \Psi(i) \triangleq & E \begin{bmatrix} E & E & \dots & E \\ \xi(0) & \xi(1) & & \xi(i-1) \end{bmatrix} [\mathbf{x}(i)\mathbf{x}^T(i) | I_0] \\
 C(i) \triangleq & B^T(i)P(i+1)B(i) + R(i+1) \quad (13) \\
 D(i) \triangleq & A^T(i)P(i+1)B(i) \quad (14)
 \end{aligned}$$

である。また

$$\begin{aligned}
 f(i) \triangleq & E[\text{tr} [\Psi(i)A^T(i)P(i+1)A(i)] | I_0] \\
 & + \sum_{j=i+1}^N \text{tr} [P(j)V_{\xi}] \quad (15)
 \end{aligned}$$

は $K(i)$ に無関係な項をまとめて表わしたものである。 $\Psi(i)$ は次式によって漸化的に算出される。

$$\begin{aligned}
 \Psi(j) = & (A(j-1) \\
 & - B(j-1)K(j-1))^T \Psi(j-1)(A(j-1) \\
 & - B(j-1)K(j-1)) + V_{\xi} \quad (16)
 \end{aligned}$$

ただし

$$\Psi(0) = E[\mathbf{x}(0)\mathbf{x}^T(0) | I_0]$$

である。

さてここで行列 A の行展開を $\text{rs } A$ 、列展開を $\text{cs } A$ で表わすことにすれば、

$$\text{tr}[AB] = \text{rs } A \cdot \text{cs } B$$

と書くことができる。この性質を(12)式に用いれば、 $J(i)$ はつぎのように書かれる。

$$\begin{aligned}
 J(i) = & E[\text{rs} (\Psi(i)K^T(i)C(i)) \cdot \text{cs } K(i) \\
 & - 2 \text{rs} (\Psi(i)D(i)) \cdot \text{cs } K(i)] + f(i) \quad (17)
 \end{aligned}$$

続いて $A \otimes B$ を A と B とのクロネッカ積とすると

$$\text{rs}(ABC) = \text{rs } B(A^T \otimes C)$$

の性質¹²⁾がある。また

$$\text{rs } A^T = [\text{cs } A]^T$$

であるから、結局 $J(i)$ はつぎのように書かれる。

$$\begin{aligned}
 J(i) = & [\text{cs } K(i)]^T E[\Psi(i) \otimes C(i)] \cdot \text{cs } K(i) \\
 & - 2 \text{rs } E[\Psi(i)D(i)] \cdot \text{cs } K(i) + f(i) \quad (18)
 \end{aligned}$$

$\Psi(i)$ および $C(i)$ が正定値ならば、 $\Psi(i) \otimes C(i)$ もまた正定値となることは、Radner の補助定理¹³⁾を用い

て容易に示すことができる。したがって $E[\Psi(i) \otimes C(i) | I_0]$ もまた正定値であることが確かめられる。

ゆえに任意に与えられた $K(0), K(1), \dots, K(i-1), K(i+1), \dots, K(N-1)$ に対して J を最小にする $K(i)$ は

$$\text{cs } K(i) = \{E[\Psi(i) \otimes C(i)]\}^{-1} \text{cs } E[D^T(i)\Psi(i)] \quad (19)$$

によって与えられる。以上をまとめるとつぎの定理が得られる。

《定理》 $K(0), K(1), \dots, K(N-1)$ が PA 1 であるための必要十分条件は、それらが(11), (16), (19)式を満すことである。

(証明) 上述したことから自明。□

また、この定理の条件は、 $K(0), K(1), \dots, K(N-1)$ が最適な (J を最小にする) ゲイン行列であるための必要条件でもある。

結局、PA 1 を求める問題は連立方程式(11), (16), (19)を解くことに帰着されたのであるが、これを直接解くことは困難であるので、つぎの繰返しアルゴリズムによって解くのが便利であろう。

(i) まず、 $K(i): i=0, 1, \dots, N-1$ を適当に仮定する。

(ii) 上に仮定した $K(i): i=0, 1, \dots, N-2$ を(16)式に用いて $\Psi(N-1)$ を計算し、そして(13), (14)式から得られる $C(N-1), D(N-1)$ を用いて(19)式から $K(N-1)$ を決定する。

(iii) 新たに得た $K(N-1)$ を(11), (13), (14)式に用いて $C(N-2)$ と $D(N-2)$ を計算する。これらと(16)式から得られる $\Psi(N-2)$ とを(19)式に用いて $K(N-2)$ を決定する。

(iv) (iii)の操作を $i=N-3, N-2, \dots, 1, 0$ の順に繰返して、 $K(i): i=N-3, \dots, 1, 0$ を順次決定する。

(v) (ii)~(iv) で得られた $K(i): i=0, 1, \dots, N-1$ を新たな仮定として用いて(ii)~(iv)の操作を繰返す。

(vi) 適当に設定された収束の条件を満すまで、(ii)~(v)の操作を繰返す。

上記の方程式の解の唯一性についての保証はないので、最初の仮定が不適当であれば良好な結果が得られないかも知れない。したがって、いくつかの異なった仮定から出発して上の繰返し計算を行ない、そこで得られた結果の中から最も良いものを採用するという手続きが必要であろう。しかしながら、幸いなことに繰返し計算を重ねるごとに評価関数の期待値は確実に減少する。このことは(19)式から得られる $K(i)$ が pbpo

であることから明白である。そこで、既存の手法で得られるものの中から最も良いものを（たとえば CE 制御方策などを）最初の仮定として用いることにすれば、それより悪くなることはなく、また改善されることはほとんど確実である。したがって、上述した繰返し計算手続きは事実上は十分に有効であるといえる。

(19) 式に現われる期待値を解析的に評価することは非常に困難であるが、ガウスの積分公式¹⁴⁾などを用いれば効果的に評価することができる。あるいはモンテカルロ法を用いてもよい。

α が i に関して独立であるような時変不規則パラメータの場合には、上に得た結果は Gunkel¹⁵⁾, Tou¹⁶⁾らによって得られた結果に帰着される。すなわち (19) 式において $\Psi(i)$ と $C(i)$, $\Psi(i)$ と $D(i)$ は互いに独立となるから、期待値と積の演算が交換可能となり、したがって

$$\begin{aligned} \text{cs } K(i) &= [E[\Psi(i)] \otimes E[C(i)]]^{-1} \\ &\quad \text{cs } [E[D^T(i)] E[\Psi(i)]] \quad (20) \end{aligned}$$

と書かれる。ここでクロネッカ積に関する性質¹²⁾

$$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$$

$$\text{cs } (ABC) = (C^T \otimes A) \text{cs } B$$

を用いれば

$$\begin{aligned} \text{cs } K(i) &= \text{cs } [E[C(i)]^{-1} E[D^T(i)] \\ &\quad \times E[\Psi(i)] E[\Psi(i)]^{-1}] \\ &= \text{cs } [E[C(i)]^{-1} E[D^T(i)]] \quad (21) \end{aligned}$$

したがって

$$\begin{aligned} K(i) &= E[C(i)]^{-1} E[D^T(i)] \\ &= E[B^T(i) E[P(i+1)] B(i) \\ &\quad + R(i+1)]^{-1} E[B^T(i) E[P(i+1)] A(i)] \\ &\triangleq E[B^T(i) \Pi(i+1) B(i) \\ &\quad + R(i+1)]^{-1} E[B^T(i) \Pi(i+1) A(i)] \quad (22) \end{aligned}$$

となって、 $K(i)$ を求めるために二点境界値問題を解く必要はなくなった。ただし

$$\Pi(i) \triangleq E[P(i)]$$

とおいた。 $\Pi(i)$ に関する漸化式は、(11) および (22) 式からつぎようになる。

$$\begin{aligned} \Pi(j) &= Q(j) + E[(A(j) \\ &\quad - B(j)K(j))^T \Pi(j+1) (A(j) - B(j)K(j))] \\ &\quad + K^T(j) R(j+1) K(j) \\ &= Q(j) - E[A^T(j) \Pi(j+1) B(j)] K(j) \quad (23) \end{aligned}$$

ただし

$$\Pi(N) = Q(N)$$

である。(22), (23) 式は Gunkel, Tou のアルゴリズムにはかならない。

つぎに簡単な数値例を示す。

(数値例)

Fig. 1
Passive adaptive
gain sequences

Fig. 2
Expected values of
criterion function
obtained by CE,
PA 1 and PA 2

プラントの動特性方程式ならびに評価関数が

$$x(i+1) = x(i) + \alpha u(i) + \xi(i) \quad (24)$$

$$W(i) = x^2(i) + u^2(i-1) \quad (25)$$

と与えられる場合を考える。ただし、変数はすべてスカラーである。 α に関する先験的確率密度関数は $N_{0r}[1, V\alpha]$ と与えられているものとする。また $x(0)=1$, $V\xi=1$ とする。

$V\alpha=0.16$ および 0.25 に対する PA 1 を Fig. 1 に示した。同図で CE とあるのは CE 制御方策において得られる feed back ゲインである。PA 1 および CE を用いることによって得られる評価関数の期待値 J を Fig. 2 に示した。同図によれば、PA 1 では CE に比べて性能が大幅に改善されているのがわかる。

Fig. 1 において、PA 1 によって得られる $K(i)$ の値は CE によって得られる $K(i)$ より小さいことが観察される。これは $K(i)$ の値が小さいほど、不確かさの影響を小さくおさえることができることを意味しているのであるが、このことは (24) 式を観察すること

によって直感的に予想されることでもある。たとえば極端な場合として、 $K(i)=0$ となれば不確かさの影響はまったくなくなる。このような制御動作のもつ受動的適応効果は Jacobs¹⁷⁾らによって“caution”と呼ばれている。これに対して Florentin¹⁸⁾によれば、本例題と同様のシステムにおいて能動的適応制御を行なうとき、 $u(i)$ を大きくするほどより速く学習が進むことが指摘されている。この効果は“probing action”¹⁹⁾と呼ばれている。すなわち、本例題で扱った制御過程においては“probing action”と“cautions action”が正反対の動作を示すことになる。

このような動作は、もっと一般の制御過程においてもみられるであろうと推論される。なぜならば、学習が効果的に行なえる状況とは、パラメータの小さな相違が大きく結果（通常は出力、本稿では状態）に反映されるような状況であり、したがって受動的適応という立場からは好ましくない状況となるからである。すなわち、学習が容易な状態へ制御することと、受動的適応の立場から好ましい状態へと制御することとは二者択一の関係にあるのである。このような定性的性質から、本稿の主題である受動的適応制御とは反対に能動的適応制御を行なう場合には、手のこんだ準最適方よりも CE 方策の方が、結局は最適方策に近くな策ということもありうるであろう。

4. 結 言

本報告においては Bayesian の立場に立って受動的適応制御問題を定式化し、未知パラメータをもつ線形レギュレータについて受動的適応制御方策を求めるアルゴリズムを導いた。さらに 1 次元のシステムを例にとり、具体的に数値解を求め、そして CE 制御方策に比較して大幅に性能が改善されることを示した。

参 考 文 献

1) R. Bellman: Adaptive control processes, Princeton

- univ. press, New Jersey (1961)
- 2) M. Aoki: Optimization of stochastic systems, Academic press, New York (1967)
- 3) 榎木, 吉川: 不確実なパラメータを含む系の準最適制御, 制御工学, **11-12**, 657/668 (1967)
- 4) W.J. Murphy: Optimal stochastic control of discrete linear systems with unknown gain, IEEE Trans. Automatic Control, **AC-13-4**, 338/344 (1968)
- 5) E. Tse and Y. Bar-Shalom: An actively adaptive control for linear systems with random parameters via the dual control approach, IEEE Trans. Automatic Control, **AC-18-2**, 109/117 (1973)
- 6) D.L. Alspach: Dual control based on approximate a posteriori density functions, IEEE Trans. Automatic Control, **AC-17-5**, 689/693 (1972)
- 7) 杉浦: 感度解析とそれを考慮した最適制御系の研究, 計測と制御, **10-1**, 1/16 (1971)
- 8) 文献⁷⁾の参考文献を参照されたい。
- 9) Y. Bar-Shalom and R. Sivan: On the optimal control of discrete-time linear systems with random parameters, IEEE Trans. Automatic Control, **AC-14-1**, 3/8 (1969)
- 10) E. Tse and M. Athans: Adaptive stochastic control for a class of linear systems, IEEE Trans. Automatic Control, **AC-17-1**, 38/52 (1972)
- 11) H.S. Witsenhausen: A counterexample in stochastic optimum control, SIAM J. Control, **6-1**, 131/147 (1968)
- 12) 児玉, 佐藤, 須田: 制御工学者のためのマトリクス理論 (13), システムと制御, **16-9**, 711/718 (1972)
- 13) R. Radner: Team decision problems, Ann. Math. Statist., **33-3**, 857/881 (1962)
- 14) A.H. Stroud and D. Secrest: Gaussian Quadrature Formulas, Prentice-Hall, Inc. (1966)
- 15) J.T. Tou: Optimum design of digital control systems, Academic press, New York (1963)
- 16) T.L. Gunckel and G.F. Franklin: A general solution for linear, sampled-data control, Trans. ASME, J. Basic Engrg., Ser. D, **85**, 197/203 (1963)
- 17) O.L.R. Jacobs and J.W. Patchell: Caution and probing in stochastic control, Int. J. Control, **16-1**, 189/199 (1972)
- 18) J.J. Florentin: Optimal, probing, adaptive control of a simple Bayesian system, J. of Electronics and Control, **13-2**, 165/177 (1962)
- 19) A.A. Feldbaum: Optimal control systems, Academic press, New York (1965)