環境変化を伴う線形系の最適制御問題

明 石 一*・能 勢 和 夫*

Optimal Control Problem for a Linear System in Switching Environment

Hajime AKASHI* and Kazuo NOSE*

This article presents a new result in the optimal control of a linear stochastic system with respect to a quadratic performance criterion.

It is assumed that the system is in "switching environment" as defined by Ackerson and Fu. The switching environment is expressed in terms of a stationary two state Markov chain with unknown transition probability. According to the Markov dependent switching environment, the system operates in the presence of nongaussian system noise and measurement noise which are timewise and mutually correlated, and are dependent on the nongaussian initial state.

Separability, neutrality and certainty equivalence previously proved for the so-called Linear-Quadratic-Gaussian problem are established in this stochastic optimal control problem. In this connection, a counterexample is found for the conjecture by Patchell and Jacobs that neutrality may be a sufficient condition for certainty equivalence.

1. 緒 言

周知のように、確率的最適制御問題は動的計画法を使って、Bellmanの汎関数方程式を解く問題に帰着させることができる¹⁾。しかしながら、この方程式を解析的手法を用いて解くことはしばしば困難であり、また、数値的に解くことは原理的には可能であっても、多くの場合、計算機に要求される記憶量は膨大なものとなり、したがって、何らかの近似最適制御が必要となってくる。このような意味から、Patchell & Jacobs²⁾は確率的最適制御問題が容易に解ける場合、すなわち

- (i) separability
- (ii) neutrality
- (iii) certainty equivalence

が成立する場合の相互関係について比較検討を行ない、certainty equivalence が成立するときには必ず separability は成立することを明らかにすると共に、neutrality は certainty equivalence が成り立つための十分条件ではないかという推測を行なっている。

ところで、最近、Ackerson & Fu³⁾は Kalman フィルタの一つの拡張として、系への外乱ならびに観測雑音の統計的性質が不規則に変化する場合の最適状態推定問題を考察し、準最適推定アルゴリズムを提案している。また、長田⁴⁾は線形確率制御系において、大きな非定常観測雑音によって観測値が異常な値を示す場合が起こり得るとき、evidence function の概念を使って、観測系の信頼性を考慮に入れた最適制御アルゴリズムについて論じている。

本論文では、Ackerson & Fu が考察した状態推定問題の結果を適用して、2次形式の評価基準のもとに、系のおかれている環境の不規則な変化に伴って、系の初期状態、系への外乱ならびに観測雑音の統計的性質が不規則に変化する場合の線形力学系の確率的最適制御問題を考察する。

まず、2.で、制御系のおかれている環境が不規則に変化することを特徴づける環境変数を導入して、上述の最適制御問題の定式化を行ない、3.において、問題が neutral であることを用いて最適制御方を誘導し、separability が成立することを示す。4.では、certainty equivalence もまた成り立つことを示し、すでに報告されている certainty equivalence, 等に関する研究との比較検討を行なう。また、最適制御方策を実施する場合の問題点についてもふれる。最後に、5.において、Patchell & Jacobs の推測は成立しないこと、すなわち、neutrality は certainty equivalence の十分条件ではないことを反例によって示す。

2. 問題の記述

はじめに、separability, 等の定義を文献 2) に従って与える。

* 京都大学工学部 京都市左京区吉田本町

* Faculty of Engineering, Kyoto University, Kyoto
(Received March 8, 1974)

【定義】 確率的最適制御問題がつぎの条件 (i), (ii) あるいは (iii) を満すとき, それぞれ, (i) separable, (ii) neutral あるいは (iii) certainty equivalent であるという.

- (i) 最適制御方策が, 問題中に存在する確率変数の2次以上の統計的モーメントを含まず, 事後平均値のみの関数となるとき.
- (ii) 問題に現われる確率変数の事後情報に基づくエントロピーが過去の制御方策に依存しないとき.
- (iii) 問題中に存在する確率変数を確定値とみなした場合の確定的最適制御問題の最適制御方策において, 確定値に対応する確率変数の事後平均値で置き換えると, 元の確率的最適制御問題の最適制御方策と一致するとき.

考察の対象とする系は次の数学モデルで記述されるものとする.

$$x(t+1) = F(t)x(t) + G(t)u(t) + \xi(t) \quad (1)$$

$$y(t) = H(t)x(t) + \eta(t) \quad t=0, 1, \dots, T-1 \quad (2)$$

ここに, $x(t)$ は時刻 t における d_1 次元状態ベクトル, $u(t)$ は d_2 次元制御ベクトル, $y(t)$ は d_3 次元観測ベクトル, $\xi(t)$ は系への外乱を表わす d_1 次元不規則ベクトル, $\eta(t)$ は観測雑音を表わす d_3 次元不規則ベクトル, $F(t)$ は $d_1 \times d_1$ 状態遷移行列, $G(t)$ は $d_1 \times d_2$ 駆動行列, $H(t)$ は $d_3 \times d_1$ 観測行列である.

ここで, (1) 式および (2) 式で記述される制御系のおかれている環境の変化を特徴づけるものとして, スカラの不規則変数 $\zeta(t)$ を導入する. $\zeta(t)$ は2状態1または2をとり(注1), 確率過程 $\{\zeta(t)\}$ は推移確率行列

$$\begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} \quad (3)$$

に従う定常な一重マルコフ連鎖とする. ここに

$$p_{ij} \triangleq P(\zeta(t)=i | \zeta(t-1)=j) \quad i, j=1, 2 \quad (4)$$

を意味している. ただし, ここでは, 推移確率 p_{ij} は未知で, 事前情報として, 確率密度関数

$$p(p_{ij}) \quad 0 \leq p_{ij} \leq 1 \quad (5)$$

が与えられているものとする. また, 初期分布は

$$P(\zeta(0)=i) = p_i \quad i=1, 2 \quad (6)$$

で与えられているとする. ただし, (4), (6) 式において

$$\sum_{i=1}^2 p_{ij} = 1 \quad \sum_{j=1}^2 p_i = 1 \quad (7)$$

という関係が成立している. 本論文では, $\zeta(t)$ を環境変数と呼ぶことにする.

(注1) 以下の議論は $\zeta(t)$ が有限個の状態をとる場合にも容易に拡張される.

系の初期状態 $x(0)$, 系への外乱 $\xi(t)$ ならびに観測雑音 $\eta(t)$ の統計的性質は, 環境変数 $\zeta(t)$ を介して, つぎのように与えられているものとする.

$$\left. \begin{aligned} p(x(0) | \zeta(0)=i) &= p(x_i(0)) \\ &\sim N\{\bar{x}_i(0), \bar{\Sigma}_i(0)\} \\ p(\xi(t) | \zeta(t+1)=i) &= p(\xi_i(t)) \\ &\sim N\{\mu_i(t), Q_i(t)\} \\ p(\eta(t) | \zeta(t)=i) &= p(\eta_i(t)) \\ &\sim N\{\lambda_i(t), R_i(t)\} \end{aligned} \right\} \quad \begin{matrix} \text{(注2)} \\ \text{(8)} \end{matrix}$$

ここに, $N\{m, C\}$ は適当な次元の平均値ベクトル m , 共分散行列 C の正規確率密度関数を表わしている. 上式は, 環境変数 $\zeta(t)$ が状態 i であるとき, 系への外乱 $\xi(t-1)$ および観測雑音 $\eta(t)$ はそれぞれ $N\{\mu_i(t-1), Q_i(t-1)\}$ および $N\{\lambda_i(t), R_i(t)\}$ なる統計的性質をもつ雑音源から発生されていることを意味している. 初期状態 $x(0)$ についても同様である. なお, $x_i(0), \xi_i(t)$ および $\eta_i(t)$ は時間的および相互に独立で, $\{\zeta(t)\}$ とも相関はないものと仮定する.

評価関数は状態ベクトルと制御ベクトルの2次形式の和の確率平均

$$J = E \left\{ \sum_{t=1}^T [x'(t)V(t)x(t) + u'(t-1)W(t-1)u(t-1)] \right\} \quad (9)$$

とする. ここに, $V(t)$ および $W(t)$ はそれぞれ対称な非負定値および正定値の重み行列であり, 記号 $E\{\cdot\}$ は数学的期待値を, また $'$ はベクトルおよび行列の転置を意味している.

上述の制御系の制御性能は評価関数(9)式の値が小さいほど良いものとし, 意志決定者(Decision Maker)は, 物理的あるいは経済的理由から, 環境変化を直接に知ることはできず, (2) 式によって与えられる観測データ $Y_t \triangleq \{y(0), y(1), \dots, y(t)\}$, 過去の制御方策 $U_{t-1} \triangleq \{u(0), u(1), \dots, u(t-1)\}$ およびすべての事前情報に基づいて, 制御方策 $u(t)$ を決定するものとする. なお, これらの情報を以後 D_t と記すことにする. すなわち

$$D_t \triangleq \{Y_t, U_{t-1}, \text{すべての事前情報}\} \quad (10)$$

とする.

意志決定者の利用しうる情報 D_t の中にマルコフ的に変化する環境変数系列 $\{\zeta(t)\}$ の実現値に関する情報が含まれていないことから, 意志決定者には, 系へ

(注2) 系への外乱 $\xi(t)$ の確率密度関数を $p(\xi(t) | \zeta(t+1))$ で与えたのは, 以下の議論において $\xi(t-1)$ と $\eta(t)$ とが対となって表われてくるため, 表記上の利点から, 環境変数の時刻を系への外乱に対して1時刻進めただけのことで本質的な問題ではない.

の外乱 $\xi(t)$ および観測雑音 $\eta(t)$ は互いに相関のある非正規性有色雑音で、初期状態 $x(0)$ とも独立ではないとみなされる点に注意しておく。

3. 最適制御方策

本節では、まず、2. で定式化した確率的最適制御問題が neutralであることを示し、つぎに、最適制御方策を導出する。

《定理 1》(1)～(8)式および(10)式によって記述される制御系は neutral である。

(証明) いま、不規則変数系列 $\{\zeta(0), \zeta(1), \dots, \zeta(t)\}$ の一つの見本系列を Z_t, Ω_t を Z_t を要素とする集合とする。このとき、状態 $x(t)$ の最小分散推定値

$$\hat{x}(t) \triangleq E\{x(t) | D_t\} \quad (11)$$

はつぎのように書き換えることができる。

$$\hat{x}(t) = \sum_{Z_t \in \Omega_t} \hat{x}(t | Z_t) P(Z_t | D_t) \quad (12)$$

ただし

$$\hat{x}(t | Z_t) \triangleq E\{x(t) | D_t, Z_t\} \quad (13)$$

である。

ここで、Ackerson & Fu³⁾ の結果を適用すると、条件つき確率密度関数 $p(x(t) | D_t, Z_t)$ は正規分布

$$N\{\hat{x}(t | Z_t), \hat{\Sigma}(t | Z_t)\} \quad (14)$$

に従い、Kalman フィルタと同様な関係式(15)～(19)が成立する。

$$\begin{aligned} \hat{x}(t | Z_t) &= \bar{x}(t | Z_t) + K(t | Z_t) \\ &\quad \times [y(t) - H(t)\bar{x}(t | Z_t) - \lambda_i(t)] \end{aligned} \quad (15)$$

$$\begin{aligned} \bar{x}(t | Z_t) &= F(t-1)\bar{x}(t-1 | Z_{t-1}) \\ &\quad + G(t-1)u(t-1) + \mu_i(t-1) \end{aligned} \quad (16)$$

$$\hat{\Sigma}(t | Z_t) = \bar{\Sigma}(t | Z_t) - K(t | Z_t)H'(t)\bar{\Sigma}(t | Z_t) \quad (17)$$

$$\begin{aligned} \bar{\Sigma}(t | Z_t) &= F(t-1)\bar{\Sigma}(t-1 | Z_{t-1})F'(t-1) \\ &\quad + Q_i(t-1) \end{aligned} \quad (18)$$

$$\begin{aligned} K(t | Z_t) &= \bar{\Sigma}(t | Z_t)H'(t)[H(t)\bar{\Sigma}(t | Z_t)H'(t) \\ &\quad + R_i(t)]^{-1} \end{aligned} \quad (19)$$

ここに

$$Z_t = \{Z_{t-1}, \zeta(t) = i\}$$

である。

つぎに、条件つき確率 $P(Z_t | D_t)$ について考える。これは Bayes 則を使って

$$\begin{aligned} P(Z_t | D_t) &= \int_0^1 \int_0^1 P(\bar{Z}_t | D_t, p_{11}, p_{22}) \\ &\quad \times p(p_{11}, p_{22} | D_t) dp_{11} dp_{22} \end{aligned} \quad (20)$$

と書き表わすことができる。そこで、まず、条件つき確率 $P(Z_t | D_t, p_{11}, p_{22})$ に着目すると、これは

$$\begin{aligned} P(Z_t | D_t, p_{11}, p_{22}) &= \frac{L(t | Z_t) p_{ij} P(Z_{t-1} | D_{t-1}, p_{11}, p_{22})}{\sum_{Z_t \in \Omega_t} [\text{分子}]} \end{aligned} \quad (21)$$

によって逐次的に求めることができる。ただし

$$Z_{t-1} = \{Z_{t-2}, \zeta(t-1) = j\}$$

であり、また、 $L(t | Z_t)$ は次式で与えられる。

$$\begin{aligned} L(t | Z_t) &\triangleq p(y(t) | D_{t-1}, u(t-1), Z_t) \\ &= \frac{1}{(2\pi)^{d/2} |H(t)\bar{\Sigma}(t | Z_t)H'(t) + R_i(t)|^{1/2}} \\ &\quad \times \exp \left\{ -\frac{1}{2} [y(t) - H(t)\bar{x}(t | Z_t) - \lambda_i(t)]' \right. \\ &\quad \times [H(t)\bar{\Sigma}(t | Z_t)H'(t) + R_i(t)]^{-1} \\ &\quad \left. \times [y(t) - H(t)\bar{x}(t | Z_t) - \lambda_i(t)] \right\} \end{aligned} \quad (22)$$

一方、条件つき確率密度関数 $p(p_{11}, p_{22} | D_t)$ に対しては

$$\begin{aligned} p(p_{11}, p_{22} | D_t) &= \frac{\sum_{Z_t \in \Omega_t} L(t | Z_t) p_{ij} P(Z_{t-1} | D_{t-1}, p_{11}, p_{22})}{\int_0^1 \int_0^1 [\text{分子}] dp_{11} dp_{22}} \\ &\quad \times \frac{p(p_{11}, p_{22} | D_{t-1})}{p(p_{11}, p_{22} | D_{t-1})} \end{aligned} \quad (23)$$

なる関係式が成立する。未知推移確率 p_{ij} は上式を用いて逐次修正される。

さて

$$\sum_{Z_t \in \Omega_t} P(Z_t | D_t) = 1$$

であるから

$$\hat{x}(t | Z_t) \triangleq x(t) - \hat{x}(t | Z_t) \quad (24)$$

と定義すれば、状態 $x(t)$ の最小分散推定誤差

$$\hat{x}(t) \triangleq x(t) - \hat{x}(t) \quad (25)$$

は $\hat{x}(t | Z_t)$ の加重平均

$$\hat{x}(t) = \sum_{Z_t \in \Omega_t} \hat{x}(t | Z_t) P(Z_t | D_t) \quad (26)$$

として与えられる。また、推定誤差の確率密度関数は

$$p(\hat{x}(t) | D_t) = \sum_{Z_t \in \Omega_t} p(\hat{x}(t) | D_t, Z_t) P(Z_t | D_t) \quad (27)$$

となる。ここに、 $p(\hat{x}(t) | D_t, Z_t)$ は(14)、(17)～(19)式から正規分布 $N[0, \hat{\Sigma}(t | Z_t)]$ に従い、過去の制御方策 U_{t-1} には依存しない。したがって、条件つき確率 $P(Z_t | D_t)$ が U_{t-1} に依存しなければ、状態 $x(t)$ の事後情報 D_t に基づくエントロピー

$$\begin{aligned}
 H(x(t)|D_i) \\
 &\triangleq - \int_{-\infty}^{\infty} p(x(t)|D_i) \log p(x(t)|D_i) dx(t) \\
 &= - \int_{-\infty}^{\infty} p(\bar{x}(t)|D_i) \log p(\bar{x}(t)|D_i) d\bar{x}(t)
 \end{aligned}
 \quad (28)$$

は過去の制御方策に依存しない。

条件つき確率 $P(Z_i|D_i)$ が U_{i-1} に依存しないことを示すためには、一連の関係式(20)~(23)式から、条件つき確率密度 $L(t|Z_i)$ が U_{i-1} に無関係であることを示せばよい。そこで、(22)式において

$$y(t) - H(t)\bar{x}(t|Z_i) - \lambda_i(t)$$

に着目する⁵⁾。これは(1)、(2)式および(24)式を使うと

$$\begin{aligned}
 &H(t)[F(t-1)\bar{x}(t-1|Z_{i-1}) \\
 &+ \xi(t-1) - \mu_i(t-1)] + \eta(t) - \lambda_i(t)
 \end{aligned}$$

と書き表わすことができる。ここに、(1)、(15)式および(16)式を(24)式に用いると

$$\begin{aligned}
 \bar{x}(t|Z_i) \\
 &= [I - K(t|Z_i)H(t)][F(t-1)\bar{x}(t-1|Z_{i-1}) \\
 &+ \xi(t-1) - \mu_i(t-1)] - K(t|Z_i)[\eta(t) - \lambda_i(t)]
 \end{aligned}$$

なる関係が成立し、さらに、 $\bar{x}(0|Z_0)$ は制御方策には無関係であることから、 $\bar{x}(t|Z_i)$ は U_{i-1} に依存しない。ただし、上式において、 I は単位行列、 $\lambda_i(t)$ および $\mu_i(t-1)$ はそれぞれ仮説 Z_i のもとでの観測雑音および外乱の平均値である。ゆえに、 $L(t|Z_i)$ はいかなる仮説 $Z_i \in \Omega_i$ のもとでも過去の制御方策 U_{i-1} には依存しないことがわかる。

最後に、外乱 $\xi(t-1)$ および観測雑音 $\eta(t)$ の事後情報 D_i に基づくエントロピー

$$\begin{aligned}
 &H(\xi(t-1)|D_i) \\
 &\triangleq - \int_{-\infty}^{\infty} p(\xi(t-1)|D_i) \\
 &\quad \times \log p(\xi(t-1)|D_i) d\xi(t-1) \\
 &H(\eta(t)|D_i) \\
 &\triangleq - \int_{-\infty}^{\infty} p(\eta(t)|D_i) \log p(\eta(t)|D_i) d\eta(t)
 \end{aligned}
 \quad (29)$$

が、共に、過去の制御方策 U_{i-1} に無関係であることも容易に確かめられる。 Q. E. D.

[系] 情報 D_i に基づく状態 $x(t)$ の最小分散推定値の推定誤差およびその共分散行列ならびに系への外乱 $\xi(\tau)$ ($\tau \geq t$) の最小分散予測値の予測誤差およびその共分散行列はいずれも過去の制御方策 U_{i-1} に依存しない。

この事実は、つぎの最適制御方策の導出に際して重要な役割りを果たす。

《定理2》 (1)~(10)式で記述される制御系の最適制御方策 $u^*(t)$ は次式で与えられる。

$$\begin{aligned}
 u^*(t) = &-[G'(t)M(t+1)G(t) + W(t)]^{-1}G'(t) \\
 &\times [M(t+1)F(t)\bar{x}(t) + \hat{\theta}(t)]
 \end{aligned}
 \quad (30)$$

$$\text{ここに} \quad \hat{\theta}(t) \triangleq E\{\theta(t)|D_i\} \quad (31)$$

であり、 $\theta(t)$ は再帰関係

$$\left. \begin{aligned}
 &\theta(t) = M(t+1)\xi(t) + S(t+1)\theta(t+1) \\
 &\dots\dots\dots \\
 &\theta(T-1) = M(T)\xi(T-1)
 \end{aligned} \right\} \quad (32)$$

によって与えられる確率変数である。ただし、 $M(t)$ は

$$\left. \begin{aligned}
 &M(t) = F'(t)M(t+1)F(t) + V(t) - N(t) \\
 &N(t) = F'(t)M(t+1)G(t)[G'(t)M(t+1)G(t) \\
 &\quad + W(t)]^{-1}G'(t)M(t+1)F(t) \\
 &M(T) = V(T)
 \end{aligned} \right\} \quad (33)$$

なる漸化式から算出される。また、 $S(t)$ は

$$\begin{aligned}
 S(t) = &F'(t)\{I - M(t+1)G(t)[G'(t) \\
 &\times M(t+1)G(t) + W(t)]^{-1}G'(t)\}
 \end{aligned}
 \quad (34)$$

で与えられる。

(証明) いま

$$\begin{aligned}
 J^0_{T-(i-1)} \triangleq &\min_{u(t-1)\dots u(T-1)} E \left\{ \sum_{\tau=t}^T [x'(\tau)V(\tau)x(\tau) \right. \\
 &\left. + u'(\tau-1)W(\tau-1)u(\tau-1)] | D_{i-1} \right\}
 \end{aligned}$$

と定義すれば、動的計画法によって、前節で定式化した最適制御問題は拘束条件(1)式のもとに つぎの汎関数方程式を解く問題に帰着される¹⁾。

$$\left. \begin{aligned}
 &J^0_{T-(i-1)} = \min_u E \{ x'(t)V(t)x(t) \\
 &\quad + u'W(t-1)u + J^0_{T-i} | D_{i-1} \} \\
 &J^0_0 = 0
 \end{aligned} \right\} \quad (35)$$

上式から、最終段の最適制御方策 $u^*(T-1)$ は次式で最小にするものである。

$$\begin{aligned}
 J_1 = &E \{ x'(T)V(T)x(T) \\
 &+ u'(T-1)W(T-1)u(T-1) | D_{T-1} \}
 \end{aligned}$$

拘束条件(1)式のもとで、上式の最小化を実行することにより

$$\begin{aligned}
 u^*(T-1) \\
 = &-[G'(T-1)M(T)G(T-1) + W(T-1)]^{-1} \\
 &\times G'(T-1)[M(T)F(T-1)\bar{x}(T-1) \\
 &+ \hat{\theta}(T-1)]
 \end{aligned}$$

を得る。このとき

$$\begin{aligned}
 J_1^0 = &\min_{u(T-1)} J_1 \\
 = &E \{ x'[F'M(T)F - N]x + 2x'S\theta \\
 &+ \text{trace} \{ M(T)\xi\xi' - G[G'M(T)G \\
 &+ W]^{-1}G'\{\theta\theta' + [M(T)F\bar{x} + \hat{\theta}] \\
 &\times [M(T)F\bar{x} + \hat{\theta}]\} \} | D_{T-1} \}
 \end{aligned}
 \quad (36)$$

となっていることは容易に確かめられる。ただし、簡単のため $(T-1)$ は省略してある。また

$$\bar{\theta}(t) \triangleq \theta(t) - \hat{\theta}(t)$$

である。(36)式において、 $\bar{x}(T-1)$ および $\bar{\theta}(T-1)$ は前述の系によって制御方策 U_{T-2} には依存しない。したがって、制御方策 U_{T-2} を決定する際には(36)式中の trace の項は無視することができる。

以上から、一般の場合の最適制御方策 $u^*(t)$ が(30)~(34)式で与えられることを導くことは容易である。

Q. E. D.

(30)~(34)式によって与えられる最適制御方策 $u^*(t)$ は状態 $x(t)$ と系への外乱 $\{\xi(\tau); t \leq \tau \leq T-1\}$ の事後情報 D_t に基づく条件つき平均値のみの関数であり、したがって、(1)~(10)式で記述される確率的最適制御問題は separable である。次節では、さらに、certainty equivalent でもあることを、まず、示すことにする。

4. 考 察

いま、(1)式において、 $x(0)$ および $\{\xi(t)\}$ は確定値で、状態 $x(t)$ は正確に観測できるものとする。このとき、評価関数

$$J = \sum_{t=1}^T [x'(t)V(t)x(t) + u'(t-1)W(t-1)u(t-1)]$$

を最小にする確定的最適制御問題を考えると、最適制御方策 $u^*(t)$ は

$$u^*(t) = -[G'(t)M(t+1)G(t) + W(t)]^{-1} \times G'(t)[M(t+1)F(t)x(t) + \theta(t)] \quad (37)$$

となる。ここに、 $\theta(t)$ および $M(t)$ はそれぞれ(32)式および(33)式によって与えられる。(30)式と(37)式とを比較することによりつぎの結果を得る。

《定理 3》 (1)~(10)式で記述される確率的最適制御問題は certainty equivalent である。

ところで、いわゆる「線形・2次形式・正規性」という条件のもとでの確率的最適制御問題に対して neutrality ならびに certainty equivalence したがって separability が成立することはよく知られている⁶⁾。さらに、Alspach & Sorenson⁵⁾、Root⁷⁾ は初期状態、付加雑音が非正規性の場合にも certainty equivalence が成立することを示している。ただし、文献 5) で変数 $x(0)$ 、 $\xi(t)$ および $\eta(t)$ は時間的ならびに相互に独立、確率立で、 $\int f(m)N\{m; C\} dm$ なる分布に従う場合に限られているが、文献 7) では、 $\xi(t-1)$ と $\eta(t)$ との相互の独立性の仮定および分布形の制限は除かれ

ている。

これらに対して、本論文では、初期状態 $x(0)$ 、付加雑音 $\xi(t)$ および $\eta(t)$ は時間的ならびに相互に関があり、かつ、非正規分布に従う場合に対しても neutrality ならびに certainty equivalence が成り立つことを示した。ただし、非正規分布は有限個の正規分布の加重平均として与えられる場合に限っていたが、 $x(0)$ 、 $\xi(t)$ および $\eta(t)$ の分布形や独立性、などの統計的性質に何の制限も課することのない(注 3)より一般の場合にも定理 1~3 は適用できる(付録参照)。これは、 $x(0)$ 、 $\xi(t)$ および $\eta(t)$ の統計的性質は、「線形・2次形式・付加雑音」制御問題の neutrality, certainty equivalence, 等の基本的性質には何の影響も与えないことを意味するものである。

なお、(30)式で与えられる最適制御方策を実施するためには、最小分散推定値 $\hat{x}(t)$ ならびに $\hat{\theta}(t)$ をオンラインで算出しなければならないが、これらの計算に必要な記憶情報は 2^{t+1} 個の $x(t|Z_t)$ 、 $\hat{\Sigma}(t|Z_t)$ および $P(Z_t|D_t)$ である。これは t が大きくなると、計算機の記憶容量の点から、最適制御方策の実施が困難となることを意味している。この問題点に関して、記憶必要量を減らすための一つの近似計算法が文献 3) で提案されている。また、ランダム・サンプリングの適用⁸⁾なども考えられる。これについては現在検討中である。

5. Patchell & Jacobs の推測に対する反例

最近、Patchell & Jacobs²⁾ は「確率的最適制御問題が neutral であるときには、必ず、certainty equivalence が成立するのではないか」という推測を行なっている。本節では、彼らの推測が成立しないことを反例によって示すことにする。

簡単のため、プラント方程式、観測方程式および評価関数が次式で与えられる一次元確率系を考える。

$$\left. \begin{aligned} x(t+1) &= x(t) + \gamma(t)u(t) + v(t) \\ y(t) &= x(t) + w(t) \\ J &= E \left\{ \sum_{t=1}^T [x^2(t) + \gamma(t-1)u^2(t-1)] \right\} \end{aligned} \right\} \quad (38)$$

ここに、 $v(t)$ および $w(t)$ はともに平均値 0、分散 1 の正規性白色雑音とし、初期状態 $x(0)$ は平均値 $\bar{x}(0)$ 、分散 $\sigma(0)$ の正規分布をしているものとする。また、 $\gamma(t)$ は 0 または 1 をとる不規則変数で、制御装置の間欠故障、などによって制御動作が中断する場合のあることを表現するものである⁹⁾。この中断過程 $\{\gamma(t)\}$ は

$$P(\gamma(t) = i | \gamma(t-1) = j) \neq 0 \quad i, j = 0, 1 \quad (39)$$

(注 3) ただし、 $x(0)$ 、 $\xi(t)$ および $\eta(t)$ はいずれも有界な共分散行列をもつものとする。

なる推移確率に従うマルコフ連鎖とする。さらに、制御方策 $u(t)$ を決定する際には、 Y_t, U_{t-1} の外に中断過程の情報 $\Gamma_{t-1} \triangleq \{\gamma(0), \gamma(1), \dots, \gamma(t-1)\}$ も利用できるものとする。すなわち

$$D_t \triangleq \{Y_t, U_{t-1}, \Gamma_{t-1}, \text{すべての事前情報}\} \quad (40)$$

とする。なお、 $x(0)$, $\{v(t)\}$, $\{w(t)\}$ および $\{\gamma(t)\}$ は互いに独立であると仮定する。

以上のように設定された確率的最適制御問題に対して neutrality が成立することは

$$\left. \begin{aligned} p(x(t)|D_t) &\sim N\{\bar{x}(t), \bar{\sigma}(t)\} \\ \bar{x}(t) &= \bar{x}(t-1) + k(t)[\gamma(t) - \bar{x}(t-1)] \\ \bar{x}(t) &= \bar{x}(t-1) + \gamma(t-1)u(t-1) \\ \bar{\sigma}(t) &= \bar{\sigma}(t-1) - k(t)\bar{\sigma}(t-1) \\ \bar{\sigma}(t) &= \bar{\sigma}(t-1) + 1 \\ k(t) &= \frac{\bar{\sigma}(t)}{\bar{\sigma}(t) + 1} \end{aligned} \right\} \quad (41)$$

という関係が成立することから明らかである(注4)。

また、最適制御方策 $u^o(t)$ が

$$u^o(t) = -\frac{\phi(t+1; t)}{\phi(t+1; t) + 1} \bar{x}(t) \quad (42)$$

で与えられることも、neutrality が成立することから容易に導かれる。ここに、 $\phi(t+1; t)$ はつぎの漸化式からオフラインで算出される⁹⁾。

$$\left. \begin{aligned} \phi(s; t) &= \phi(s+1; t) + 1 \\ -P(\gamma(s)=1|\gamma(t)=1) \frac{\phi^2(s+1; s)}{\phi(s+1; s) + 1} &\quad (t < s < T) \\ \phi(T; t) &= 1 \end{aligned} \right\} \quad (43)$$

さて、確率的最適制御問題(38)~(40)式に対応する確定的最適制御問題

$$\left. \begin{aligned} x(t+1) &= x(t) + \gamma(t)u(t) + v(t) \\ J &= \sum_{t=1}^T [x^2(t) + \gamma(t-1)u^2(t-1)] \\ \left(\begin{aligned} \text{ただし、}\gamma(t) \text{ と } v(t) \text{ は既知で、状態} \\ x(t) \text{ は正確に観測できるものとする。} \end{aligned} \right) \end{aligned} \right\} \quad (44)$$

の最適制御方策 $u(t)$ において、 $x(\cdot)$, $\gamma(\cdot)$ および $v(\cdot)$ を情報 D_t に基づく条件つき平均値でおきかえると

$$u(t) = -\frac{\phi(t+1|t)}{P(\gamma(t)=1|D_t)\phi(t+1|t) + 1} \bar{x}(t) \quad (45)$$

となる。ただし $\phi(t+1|t)$ は

$$\left. \begin{aligned} \phi(s|t) &= \phi(s+1|t) + 1 \\ -\frac{P(\gamma(s)=1|D_t)\phi^2(s+1|t)}{P(\gamma(s)=1|D_t)\phi(s+1|t) + 1} &\quad (t < s < T) \\ \phi(T|t) &= 1 \end{aligned} \right\} \quad (46)$$

(注4) たとえば $H(x(t)|D_t) = \log \sqrt{2\pi e \bar{\sigma}(t)}$ 。

なる漸化式からオンラインで計算しなければならない。

(42), (43)式と(45), (46)式とを比較すれば、(38)~(40)式で記述される確率的最適制御問題は neutral ではあるが、certainty equivalent ではないことがわかる。

6. 結 言

系のおかれている環境が不規則に変化することによって、初期状態、系への外乱および観測雑音の統計的性質が不規則に変化する場合の線形系の最適制御問題を考察した。環境の変化がマルコフ的であるため、制御系に存在する不確定要素は、意志決定者には、時間的および相互に相関があり、かつ、非正規分布をしているとみなされるにもかかわらず、制御系は neutral であり、さらに、2次形式の評価基準のもとでは、certainty equivalence したがって、separability が成立することを示した。また、Patchell & Jacobs の推測が成立しないこと、すなわち、neutrality は certainty equivalence の十分条件ではないことを反例を挙げて示した。

最後に、本論文に対して討論を頂いた本学数理工学教室藤重 悟氏に感謝致します。

参 考 文 献

- 1) R. Bellman: Adaptive Control Processes; A Guided Tour, Princeton University Press (1961)
- 2) J. W. Patchell & O. L. R. Jacobs: Separability, Neutrality and Certainty Equivalence, Int. J. Control, **13**-2, 337/342 (1971)
- 3) G. A. Ackerson & K. S. Fu: On State Estimation in Switching Environments, IEEE Trans. on Automatic Control, **AC-15**-1, 10/17 (1970)
- 4) 長田 正: 観測系の信頼度と最適制御, 電気学会雑誌, **89**-1-964, 170/178 (1969)
- 5) D. L. Alspach & H. W. Sorenson: Stochastic Optimal Control for Linear but Nongaussian Systems, Int. J. Control, **13**-6, 1169/1181 (1971)
- 6) M. Aoki: Optimization of Stochastic Systems, Academic Press (1967)
- 7) J. G. Root: Optimum Control of Non-Gaussian Linear Stochastic Systems with Inaccessible State Variables, SIAM J. Control, **7**-2, 317/323 (1969)
- 8) 明石 一, 熊本博光, 能勢和夫: 状態依存性観測雑音を受ける線形系の状態推定—ランダム・サンプリングによるアルゴリズム—, 第6回統計学的制御理論シンポジウム講演論文集, 37/40 (1974)
- 9) 明石 一, 能勢和夫, 藤本隆一: 不確実性を有する線形確率系の準最適制御, 第16回自動制御連合講演会前刷, 143/144 (1973)
- 10) W. M. Wonham: On the Separation Theorem of Stochastic Control, SIAM J. Control, **6**-2, 312/326 (1968)

《付 録》

$x(t)$ をつぎのように $x^*(t)$ と $x^\#(t)$ とに分割する.

$$x(t) = x^*(t) + x^\#(t) \quad (\text{A} \cdot 1)$$

ただし, $x^*(t)$ および $x^\#(t)$ は, それぞれ

$$x^*(t+1) = F(t)x^*(t) + G(t)u(t), \quad x^*(0) = 0 \quad (\text{A} \cdot 2)$$

および

$$x^\#(t+1) = F(t)x^\#(t) + \xi(t), \quad x^\#(0) = x(0) \quad (\text{A} \cdot 3)$$

で定義されるものである.

いま

$$\begin{aligned} y^\#(t) &\triangleq y(t) - H(t)x^*(t) \\ &= H(t)x^\#(t) + \eta(t) \end{aligned} \quad (\text{A} \cdot 4)$$

とおき, $D_i^\#$ を

$$D_i^\# \triangleq \{y^\#(0), y^\#(1), \dots, y^\#(t), \text{すべての事前情報}\} \quad (\text{A} \cdot 5)$$

と定義すると, 文献 10) と同様の議論を経て

$$D_i = D_i^\# \quad (\text{A} \cdot 6)$$

(厳密には, (2) 式によって与えられる確率過程 $\{y(t)\}$ の生成する最小 σ -代数と, (A・4) 式で定義される確率過程 $\{y^\#(t)\}$ が生成する最小 σ -代数とが等しいこと) が証明される. さらに

$$\begin{aligned} \tilde{x}(t) &= x(t) - E\{x(t) | D_i\} \\ &= x^\#(t) - E\{x^\#(t) | D_i^\#\} \triangleq \tilde{x}^\#(t) \end{aligned} \quad (\text{A} \cdot 7)$$

を得る. 上式は, 推定誤差 $\tilde{x}(t)$ は力学系を (A・3) 式とし, これに対する観測方程式を (A・4) 式とする推定問題の推定誤差 $\tilde{x}^\#(t)$ に等しいことを示している. すなわち, $\tilde{x}(t)$ は, $x(0), \xi(t)$ ならびに $\eta(t)$ の統計的性質のいかんにかかわらず, 過去の制御方策 U_{t-1} に依存しない. 以上の議論で, 定理 1~3 は, $x(0), \xi(t)$ ならびに $\eta(t)$ の統計的性質に関係なく成立することがわかる.