回帰型非線形システムのパラメータ推定に関する一手法[†]

笠松 博史*・樋口 彰**・森本 滋郎**・古本 奈奈代***・田 渕 敏 明**

A Method on Parameter Estimation of Regression-Type Nonlinear Systems[†]

Hiroshi KASAMATSU*・Akira HIGUCHI**・Jiro MORIMOTO**・Nanayo FURUMOTO*** and Toshiaki TABUCHI**

This paper deals with the parameter estimation of regression-type nonlinear system model based on maximum *a posteriori* (MAP) estimation. This is equivalent to the minimization problem of the object function derived from *a posteriori* probability density function. It is shown that the minimizer based on Newton-method becomes to iterated extended Kalman filter (IEKF). A method on the gain adjustment of the IEKF is given so as to improve the performance of it. Then the parameter setting for the gain adjustment is given as the solution of a continuous-time algebraic Riccati equation (CARE). As a result, the proposed estimator has some effect of simulated annealing.

It is shown in numerical example that the proposed method indicates acceptable performance compared with conventional methods.

Key Words : nonlinear systems, parameter estimation, iterated extended Kalman filter, MAP estimate, gain adjustment

1. はじめに

回帰型非線形システムの未知パラメータベクトルの推定は、システム同定、信号処理などの分野で重要であり、種々の手法が考案されている。その代表的なものとして、最尤推定法(以後、ML推定法と略記)、最大事後推定法(以後、MAP推定法と略記)がある¹⁾²⁾。これらの手法は、最終的には、ある目的関数の最小化問題に帰着する。このとき、扱われるシステムモデルが未知パラメータに関して非線形の場合は、目的関数も未知パラメータに関して多峰性になる。この最小化にニュートン法などの勾配法を用いる場合には、得られる結果が最小点ではなく、極小点に過ぎない可能性がある。そのような場合には、極小点回避が必要になってくる。

これを行なう方法の一つとして、パラメータ推定量に確率的なゆらぎを与える方法がある³⁾⁴⁾。このときに問題になるのが、ゆらぎの強さをどの程度にしたらいいか、ということである。これが大きければ、極小点からの脱出は容易であるが、一方では推定値が振動的になり、最小点への収束が困難になる。他方、これが小さければ、極小点からの脱出ができない。

このようなことから、ゆらぎの強さを調整することは、モデルの非線形性に起因するモデルパラメータの推定精度劣化を軽減する観点から、重要なことである。

本論文では、MAP推定法による回帰型非線形システムの未知パラメータ推定問題が扱われ、極小点回避のための確率的ゆらぎの強さを調整する一つの方法が提案される。具体的には、パラメータ推定誤差ノルムの平均が0に収束するというところから逆算したときに、推定量が備えておくべき条件が導出され、この条件を満足するようなゆらぎの強さは、連続時間型の代数Riccati方程式(以後、CAREと略記)の解として得られることが示される。数値実験例では、本推定量が良好な結果を与えることが示される。

2. MAP推定とアルゴリズム

2.1 問題の記述

次式の回帰型非線形システムを考える^{注1)}。

$$y_t = g(\boldsymbol{\varphi}_t, \boldsymbol{\theta}_0) + e_t \quad (2.1)$$

上式において、 y_t はスカラーの観測データ、回帰ベクトル $\boldsymbol{\varphi}_t$ はシステムの入出力 u_t, y_t をもとに、次式で与えられる。

$$\boldsymbol{\varphi}_t = [y_{t-1}, \dots, y_{t-n_y}, u_t, \dots, u_{t-n_u}]^T \quad (2.2)$$

ここで、肩付の T はベクトルもしくは行列の転置を表わし、次数 n_y, n_u は既知とする。 u_t は外生入力であり、フィードバックはかかっていないとする。 e_t は平均が 0、分散が σ_e^2 の白色雑音で、 $\boldsymbol{\varphi}_t, \boldsymbol{\theta}_0$ とは統計的に独立である。また、 g は n 次元パラメータベクトル $\boldsymbol{\theta}_0$ に関して微分可能な無記憶の非線形関数で

^{注1)} 本論文では、確率変数、確率ベクトル、確率行列は立体で表記され、これらの実現値は斜体で表記される。また、確定的定数、変数は数値、零ベクトル $\mathbf{0}$ 、零行列 $\mathbf{O}$ を除いて斜体で表記される。

[†] 平成 14 年度電気関係学会四国支部連合大会で発表(2002・10)

* 徳島文理大学大学院工学研究科在学 さぬき市志度

** 徳島文理大学工学部 さぬき市志度

*** 徳島文理大学人間生活学部 徳島市山城町西浜傍示

* Graduate Student of Engineering, Tokushima Bunri University, Sanuki-shi

** Faculty of Engineering, Tokushima Bunri University, Sanuki-shi

*** Faculty of Human Life Science, Tokushima Bunri University, Yamashiro-cho, Tokushima

(Received October 15, 2002)

(Revised March 27, 2003)

あり, その形状は既知とする. いま, 表記を簡単にするために $\hat{y}_t(\theta_0) \triangleq g(\varphi_t, \theta_0)$ として, (2.1)式をつぎのように書く.

$$y_t = \hat{y}_t(\theta_0) + e_t \quad (2.3)$$

このとき, 問題は $z^t \triangleq \{y_i, u_i; i \leq t\}$ を用いて n 次元未知パラメータベクトル θ_0 を推定することである.

2.2 MAP 推定法によるパラメータ推定アルゴリズム

一般に確率変数 x, y に対して, 確率密度関数を簡単のために $p(x) \triangleq p_x(x), p(x|y) \triangleq p_{x|y}(x|y)$ のように確率変数を表わす下付きの添字を略して表記するとき, $p(\theta|z^t)$ は Bayes の定理を用いてつぎのように表わされる.

$$p(\theta|z^t) = \text{const} \cdot p(y_t|u_t, z^{t-1}, \theta) p(\theta|z^{t-1}) \quad (2.4)$$

いま, $p(y_t|u_t, z^{t-1}, \theta), p(\theta|z^{t-1})$ はそれぞれつぎの正規分布であることを仮定する.

$$\left. \begin{aligned} p(y_t|u_t, z^{t-1}, \theta) &= N[\hat{y}_t(\theta), \sigma_e^2] \\ p(\theta|z^{t-1}) &= N[\hat{\theta}_{t-1}, P_{t|t-1}] \end{aligned} \right\} \quad (2.5)$$

ただし

$$P_{t|t-1} = P_{t-1} + Q_{t-1}, (Q_{t-1} \geq 0) \quad (2.6)$$

であり, Q_{t-1} が確率的ゆらぎの強さを調整するパラメータである. (2.5)式の表記に関し, 一般に n 次元ベクトル x に対して

$$p(x) = N[m, P] = \frac{1}{\sqrt{(2\pi)^n |P|}} e^{-\frac{1}{2}(x-m)^T P^{-1}(x-m)}$$

を表わす. (2.5)式を考慮するとき, (2.4)式の対数は次式となる.

$$\begin{aligned} \ell_t(\theta) &\triangleq \log p(\theta|z^t) \\ &= \text{const} - \frac{1}{2} f_t(\theta, \hat{\theta}_{t-1}, \sigma_e^2) \end{aligned} \quad (2.7)$$

ただし

$$f_t(\theta, \hat{\theta}_{t-1}, \sigma_e^2) \triangleq \frac{\hat{e}_t^2(\theta)}{\sigma_e^2} + \|\theta - \hat{\theta}_{t-1}\|_{P_{t|t-1}^{-1}}^2 \quad (2.8)$$

$$\hat{e}_t(\theta) \triangleq y_t - \hat{y}_t(\theta) \quad (2.9)$$

であり, 一般に $\|x\| = \sqrt{x^T x}, \|x\|_A^2 = x^T A x$ である. MAP 推定は (2.7)式の θ に関する最大点を, したがって (2.8)式の最小点を求めるものである. この最小化にニュートン法を用いるとき, $f_t(\theta, \hat{\theta}_{t-1}, \sigma_e^2)$ およびこの θ に関する導関数の値の計算が必要になるが, σ_e^2 が未知のため, これらが計算できない. そこで, (2.8)式における σ_e^2 を, この ML 推定値 $\hat{\sigma}_{t-1}^2$ で代用した $f_t(\theta, \hat{\theta}_{t-1}, \hat{\sigma}_{t-1}^2)$ が最小化される.

$f_t(\theta, \hat{\theta}_{t-1}, \hat{\sigma}_{t-1}^2)$ の最小化アルゴリズムとして, 各 t でニュートン法を用いるとき, 次式を得る.

$$\hat{\theta}_t^{(l)} = \hat{\theta}_t^{(l-1)} - G_t^{(l-1)^{-1}} d_t^{(l-1)} \quad (2.10)$$

ただし, 肩付の (l) は各 t における繰返し数を表わし, また

$$\left. \begin{aligned} d_t^{(l-1)} &\triangleq d_t(\theta, \hat{\theta}_{t-1}, \hat{\sigma}_{t-1}^2) \Big|_{\theta=\hat{\theta}_t^{(l-1)}} \\ G_t^{(l-1)} &\triangleq G_t(\theta, \hat{\theta}_{t-1}, \hat{\sigma}_{t-1}^2) \Big|_{\theta=\hat{\theta}_t^{(l-1)}} \end{aligned} \right\} \quad (2.11)$$

である. 上式右辺を具体的に評価するとき, 次式を得る.

$$d_t(\theta, \hat{\theta}_{t-1}, \hat{\sigma}_{t-1}^2) \triangleq \frac{\partial f_t(\theta, \hat{\theta}_{t-1}, \hat{\sigma}_{t-1}^2)}{\partial \theta}$$

$$= 2 \left[-\frac{\psi_t(\theta) \hat{e}_t(\theta)}{\hat{\sigma}_{t-1}^2} + P_{t|t-1}^{-1} (\theta - \hat{\theta}_{t-1}) \right] \quad (2.12)$$

$$\begin{aligned} G_t(\theta, \hat{\theta}_{t-1}, \hat{\sigma}_{t-1}^2) &\triangleq \frac{\partial^2 f_t(\theta, \hat{\theta}_{t-1}, \hat{\sigma}_{t-1}^2)}{\partial \theta \partial \theta^T} \\ &= 2 \left[P_{t|t-1}^{-1} + \frac{\psi_t(\theta) \psi_t^T(\theta) - \Gamma_t(\theta) \hat{e}_t(\theta)}{\hat{\sigma}_{t-1}^2} \right] \end{aligned} \quad (2.13)$$

$$\psi_t(\theta) \triangleq \frac{\partial \hat{y}_t(\theta)}{\partial \theta}, \quad \Gamma_t(\theta) \triangleq \frac{\partial^2 \hat{y}_t(\theta)}{\partial \theta \partial \theta^T} \quad (2.14)$$

(2.12)式, (2.13)式をもとに (2.11)式を評価して, それらを (2.10)式に用いるとき, MAP 推定値が得られるが, このときに, (2.13)式における $\Gamma_t(\theta) \hat{e}_t(\theta)$ は $G_t(\theta, \hat{\theta}_{t-1}, \hat{\sigma}_{t-1}^2)$ の正值性を崩して数値計算を不安定にさせる恐れがあるために, この項は無視する. その結果, (2.10)式はつぎのように表わされる.

$$\hat{\theta}_t^{(l)} = \hat{\theta}_{t-1} + k_t^{(l-1)} \varepsilon_t^{(l-1)}, \quad \hat{\theta}_t^{(0)} = \hat{\theta}_{t-1} \quad (2.15a)$$

$$\varepsilon_t^{(l-1)} = \hat{e}_t^{(l-1)} + \psi_t^{(l-1)^T} (\hat{\theta}_t^{(l-1)} - \hat{\theta}_{t-1}) \quad (2.15b)$$

$$\hat{e}_t^{(l-1)} = y_t - \hat{y}_t^{(l-1)} \quad (2.15c)$$

$$k_t^{(l-1)} = \frac{P_{t|t-1} \psi_t^{(l-1)}}{r_t^{(l-1)}} \quad (2.15d)$$

$$P_{t|t-1} = P_{t-1} + Q_{t-1} \quad (2.15e)$$

$$P_t^{(l-1)} = (I - k_t^{(l-1)} \psi_t^{(l-1)^T}) P_{t|t-1}, \quad P_t^{(0)} = P_{t|t-1} \quad (2.15f)$$

$$r_t^{(l-1)} = \psi_t^{(l-1)^T} P_{t|t-1} \psi_t^{(l-1)} + \hat{\sigma}_{t-1}^2 \quad (2.15g)$$

$$\hat{\sigma}_t^2 = \left(1 - \frac{1}{l}\right) \hat{\sigma}_{t-1}^2 + \frac{1}{l} \hat{e}_t^2 \quad (2.15h)$$

ここで, 一般に $A_t^{(l-1)} = A_t(\hat{\theta}_t^{(l-1)})$ であり, 上式が各 t で収束判定条件 $\|\hat{\theta}_t^{(l)} - \hat{\theta}_t^{(l-1)}\| < \delta, (\delta = 10^{-10} \sim 10^{-12})$ を満足するまで繰返される. この条件を満足したときの $\hat{\theta}_t^{(l)}$ を $\hat{\theta}_t$ と表記するとき, これが θ_0 の MAP 推定値である. (2.15)式は繰返型の拡張カルマンフィルタ (以後, IEKF と略記) であり, $l=1$ とした場合は拡張カルマンフィルタ (以後, EKF と略記) である. (2.15)式の繰返し l に関する収束形は, 肩付量における肩付部を取って表記するとき, 推定量の形で次式となる.

$$\hat{\theta}_t = \hat{\theta}_{t-1} + k_t \varepsilon_t \quad (2.16a)$$

$$\varepsilon_t = \hat{e}_t + \psi_t^T (\hat{\theta}_t - \hat{\theta}_{t-1}) \quad (2.16b)$$

$$\hat{e}_t = y_t - \hat{y}_t \quad (2.16c)$$

$$k_t = \frac{P_{t|t-1} \psi_t}{r_t} \quad (2.16d)$$

$$P_{t|t-1} = P_{t-1} + Q_{t-1} \quad (2.16e)$$

$$P_t = (I - k_t \psi_t^T) P_{t|t-1} \quad (2.16f)$$

$$r_t = \psi_t^T P_{t|t-1} \psi_t + \hat{\sigma}_{t-1}^2 \quad (2.16g)$$

$$\hat{\sigma}_t^2 = \left(1 - \frac{1}{l}\right) \hat{\sigma}_{t-1}^2 + \frac{1}{l} \hat{e}_t^2 \quad (2.16h)$$

2.3 ゆらぎ調整アルゴリズム

(2.15)式の実行に際しては, (2.15e)式におけるゆらぎの強さ調整用パラメータ Q_{t-1} が既知でなければならない. 本論文ではこの値の設定が (2.16b)式, (2.16g)式をもとに

$$v_i \triangleq \frac{\varepsilon_i}{\sqrt{r_{i,-}}} \quad (2.17)$$

$$r_{i,-} \triangleq r_i \Big|_{\mathbf{Q}_{i-1}=\mathbf{0}} = \boldsymbol{\Psi}_i^T \mathbf{P}_{i-1} \boldsymbol{\Psi}_i + \hat{\sigma}_{v,i}^2 \quad (2.18)$$

で定義される $\{v_i\}$ の統計的挙動監視にもとづいて、つぎの①～⑦で設定される。これらの導出は、3.2節～3.5節で与えられる。

①(2.19)式による AIC の計算と MAICE の特定

$$\left. \begin{aligned} \text{AIC}(0, 1) &= \text{const} + L(\hat{\mu}_i^2 + \hat{\sigma}_{v,i}^2) \\ \text{AIC}(\mu_i, 1) &= \text{const} + L\hat{\sigma}_{v,i}^2 + 2 \\ \text{AIC}(0, \sigma_{v,i}^2) &= \text{const} + L \log(\hat{\mu}_i^2 + \hat{\sigma}_{v,i}^2) + L + 2 \\ \text{AIC}(\mu_i, \sigma_{v,i}^2) &= \text{const} + L \log \hat{\sigma}_{v,i}^2 + L + 4 \end{aligned} \right\} \quad (2.19)$$

ただし

$$\left. \begin{aligned} \text{const} &= L \log 2\pi, \quad \hat{\mu}_i = \frac{1}{L} \sum_{i=i-L+1}^i v_i \\ \hat{\sigma}_{v,i}^2 &= \frac{1}{L} \sum_{i=i-L+1}^i v_i^2 - \hat{\mu}_i^2 \end{aligned} \right\} \quad (2.20)$$

であり、MAICE(Minimum AIC Estimate) は(2.19)式の最小値のことである。 L は v_i の使用個数である。

②(2.21)式による v_i の分散の暫定推定値 $\hat{\tau}_{i,-}$ の算出

$$\hat{\tau}_{i,-} = \begin{cases} 1 & , \text{ if MAICE} = \text{AIC}(0, 1) \\ 1 + \hat{\mu}_i^2 & , \text{ if MAICE} = \text{AIC}(\mu_i, 1) \\ \hat{\sigma}_{v,i}^2 & , \text{ if MAICE} = \text{AIC}(0, \sigma_{v,i}^2) \\ \hat{\mu}_i^2 + \hat{\sigma}_{v,i}^2 & , \text{ if MAICE} = \text{AIC}(\mu_i, \sigma_{v,i}^2) \end{cases} \quad (2.21)$$

③(2.22)式の A_i , (2.23)式の B_i の算出

$$A_i \triangleq \frac{1}{L} \sum_{i=i-L+1}^i \frac{\boldsymbol{\Psi}_i \boldsymbol{\Psi}_i^T}{r_{i,-}} + \delta_Q I \quad (2.22)$$

$$B_i \triangleq \frac{1}{L} \sum_{i=i-L+1}^i \alpha_i \frac{\boldsymbol{\Psi}_i \boldsymbol{\Psi}_i^T}{\|\boldsymbol{\Psi}_i\|^4 + \delta_\Psi} \varepsilon_i^2 + \delta_Q I \quad (2.23)$$

ただし

$$\alpha_i \triangleq \left(\sqrt{\hat{\tau}_{i,-}} - \frac{1}{\sqrt{\hat{\tau}_{i,-}}} \right)^2 \quad (2.24)$$

であり、 $\delta_\Psi = 10^{-5} \sim 10^{-7}$, $\delta_Q = 10^{-10} \sim 10^{-12}$ で設定すればよい。

④(2.25)式による $\bar{\tau}_i$ の算出

$$\bar{\tau}_i \triangleq 1 + \frac{1}{n} \sqrt{\frac{\lambda_{\min}(B_i)}{\lambda_{\max}(A_i)}} \frac{\|\boldsymbol{\Psi}_i\|^2}{r_{i,-}} \quad (2.25)$$

ただし、 $\lambda_{\max}(\bullet)$, $\lambda_{\min}(\bullet)$ はそれぞれ $(\bullet)$ の最大固有値、最小固有値である

⑤(2.26)式による v_i の分散の推定値 $\hat{\tau}_i$ の算出

$$\hat{\tau}_i = \begin{cases} \hat{\tau}_{i,-} & , \text{ if } \hat{\tau}_{i,-} \leq \bar{\tau}_i \\ \bar{\tau}_i & , \text{ if } \hat{\tau}_{i,-} > \bar{\tau}_i \end{cases} \quad \text{for } \hat{\tau}_{i,-} \geq 1$$

$$\frac{1}{\hat{\tau}_i} = \begin{cases} \frac{1}{\hat{\tau}_{i,-}} & , \text{ if } \frac{1}{\hat{\tau}_{i,-}} \leq \bar{\tau}_i \\ \bar{\tau}_i & , \text{ if } \frac{1}{\hat{\tau}_{i,-}} > \bar{\tau}_i \end{cases} \quad \text{for } \hat{\tau}_{i,-} < 1$$

$$\quad (2.26)$$

⑥(2.26)式の $\hat{\tau}_i$ を用いて

$$\alpha_i \triangleq \left(\sqrt{\hat{\tau}_i} - \frac{1}{\sqrt{\hat{\tau}_i}} \right)^2 \quad (2.27)$$

の算出と、これを用いての(2.23)式の B_i の算出

⑦(2.28)式第1式の CARE の求解

$$\mathbf{Q}_i \mathbf{A}_i \mathbf{Q}_i = \mathbf{B}_i, \quad (i \geq L), \quad \mathbf{Q}_i = \mathbf{0}, \quad (i < L) \quad (2.28)$$

3. アルゴリズムの性質と導出

3.1 パラメータ推定アルゴリズムの性質

(2.16)式の特徴の一つは、 $\hat{\boldsymbol{\theta}}_i$ が目的関数 $f_i(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}_{i-1}, \hat{\sigma}_{v,i}^2)$ の停留点になっていることである。これはつぎのようにしてわかる。まず、(2.16a)式を(2.16b)式に代入して $\varepsilon_i = \hat{\varepsilon}_i / (1 - \boldsymbol{\Psi}_i^T \mathbf{k}_i)$ を得る。ここで、 $\boldsymbol{\Psi}_i^T \mathbf{k}_i$ の部分に(2.16d)式、(2.16g)式を用いるとき、次式を得る。

$$\varepsilon_i = \frac{r_i}{\hat{\sigma}_{v,i}^2} \hat{\varepsilon}_i \quad (3.1)$$

これを(2.12)式に代入するとき

$$d_i(\hat{\boldsymbol{\theta}}_i, \hat{\boldsymbol{\theta}}_{i-1}, \hat{\sigma}_{v,i}^2) = \mathbf{0} \quad (3.2)$$

となり、 $\hat{\boldsymbol{\theta}}_i$ は $f_i(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}_{i-1}, \hat{\sigma}_{v,i}^2)$ の停留点となる。このときの停留値は(2.8)式より $f_i(\hat{\boldsymbol{\theta}}_i, \hat{\boldsymbol{\theta}}_{i-1}, \hat{\sigma}_{v,i}^2) = \varepsilon_i^2 / r_i$ となり、この両辺に z' が与えられたときの $\boldsymbol{\theta}_0$ に関する条件付期待値をとるとき、 $f_i(\hat{\boldsymbol{\theta}}_i, \hat{\boldsymbol{\theta}}_{i-1}, \hat{\sigma}_{v,i}^2)$ は z' の関数であるので、次式を得る。

$$f_i(\hat{\boldsymbol{\theta}}_i, \hat{\boldsymbol{\theta}}_{i-1}, \hat{\sigma}_{v,i}^2) = E_{\boldsymbol{\theta}_0} \left\{ \frac{\varepsilon_i^2}{r_i} \mid z' \right\} \quad (3.3)$$

また、 $\hat{\boldsymbol{\theta}}_i = \boldsymbol{\theta}_0$, $\hat{\sigma}_{v,i}^2 = \sigma_v^2$ の場合は(2.8)式をもとに

$$E_{\boldsymbol{\theta}_0, z'} \{ f_i(\boldsymbol{\theta}_0, \boldsymbol{\theta}_0, \sigma_v^2) \} = 1 \quad (3.4)$$

となる。なお、(3.2)式、(3.3)式が成立するのは IEKF の場合であり、EKF ではこれらは成立しない。IEKF は EKF に比べて $\hat{\mathbf{y}}_i(\boldsymbol{\theta}_0)$ の $\boldsymbol{\theta}_0$ に関する非線形性に起因する推定精度劣化を軽減し、 $\hat{\boldsymbol{\theta}}_i \triangleq \hat{\boldsymbol{\theta}}_0 - \hat{\boldsymbol{\theta}}_i$ とするときつぎの近似が成立する^{3),6)}。

$$E_{\boldsymbol{\theta}_0} \{ \hat{\boldsymbol{\theta}}_i \mid z' \} \approx \mathbf{0}, \quad E_{\boldsymbol{\theta}_0} \{ \hat{\boldsymbol{\theta}}_i^T \mid z' \} = \mathbf{P}_i \quad (3.5)$$

ところで、(2.16b)式の ε_i の中身はつぎのようになる。まず、(2.3)式を $\hat{\boldsymbol{\theta}}_i$ の回りで Taylor 展開して、各量の実現値が計算可能な部分、そうでない部分を両辺にわけるとき次式を得る。

$$\hat{\varepsilon}_i + \boldsymbol{\Psi}_i^T \hat{\boldsymbol{\theta}}_i = \boldsymbol{\Psi}_i^T \boldsymbol{\theta}_0 + \zeta_i + e_i \quad (3.6)$$

ただし

$$\zeta_i \triangleq \frac{1}{2} \hat{\boldsymbol{\theta}}_i^T \boldsymbol{\Gamma}_i \hat{\boldsymbol{\theta}}_i + \text{higher order terms} \quad (3.7)$$

であり、 ζ_i は $\hat{\mathbf{y}}_i(\boldsymbol{\theta})$ が $\boldsymbol{\theta}$ に関して非線形であるために現れる項である。(3.6)式に対する1段予測誤差が(2.16b)式であり、その中身は次式となる。

$$\varepsilon_i = \boldsymbol{\Psi}_i^T \hat{\boldsymbol{\theta}}_{i-1} + \zeta_i + e_i \quad (3.8)$$

したがって、 $\hat{\boldsymbol{\theta}}_i$ が大きい場合は(3.7)式より ζ_i の分散も大きくなり、結果として ε_i の分散も大きくなる。

3.2 ゆらぎ調整の指標

(2.16a)式は(2.16d)式、(2.16e)式を用いるとき、つぎのように表わされる。

$$\hat{\boldsymbol{\theta}}_i = \hat{\boldsymbol{\theta}}_{i-1} + \frac{\mathbf{P}_{i-1} \boldsymbol{\Psi}_i}{r_i} \varepsilon_i + \hat{\mathbf{w}}_{i-1|i} \quad (3.9)$$

ただし

$$\hat{\mathbf{w}}_{t-1|t} \triangleq \mathbf{Q}_{t-1} \frac{\boldsymbol{\Psi}_t}{\mathbf{r}_t} \boldsymbol{\varepsilon}_t \quad (3.10)$$

である。このとき、本論文で言う確率的ゆらぎとは、(3.9)式右辺第3項のことである。今、(2.16)式において持続的励振条件

$$\sum_{k=1}^m \frac{\boldsymbol{\Psi}_{t-k+1}^T \boldsymbol{\Psi}_{t-k+1}}{\hat{\sigma}_{t-k}^2} \geq \beta \mathbf{I} (>0), \quad (t \geq m) \quad (3.11)$$

が満足されるとする。このとき、(3.9)式において $\hat{\mathbf{w}}_{t-1|t} = \mathbf{0}$ すなわち $\mathbf{Q}_{t-1} = \mathbf{0}$ の場合は $\mathbf{P}_t \rightarrow \mathbf{0}, (t \rightarrow \infty)$ となるので、 $\hat{\boldsymbol{\theta}}_t \rightarrow \hat{\boldsymbol{\theta}}$ となる。このとき、 $\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}_0$ であればよいが、 $\hat{\boldsymbol{\theta}}$ が目的関数の極小点に過ぎない場合は、 $\hat{\mathbf{w}}_{t-1|t}$ の投入による極小点からの脱出が必要になる。その投入指針としては、 $\hat{\boldsymbol{\theta}}_t \rightarrow \mathbf{0}, (t \rightarrow \infty)$ とすることが考えられる。そこで、まず $\hat{\boldsymbol{\theta}}_t$ に関して次式を得る。

$$\begin{aligned} E_{\theta_0, z_t} \{ \|\hat{\boldsymbol{\theta}}_t\| \} &= E_{\theta_0, z_t} \left\{ \left\| \mathbf{P}_t^{-\frac{1}{2}} \mathbf{P}_t^{-\frac{1}{2}} \hat{\boldsymbol{\theta}}_t \right\| \right\} \\ &\leq E_{\theta_0, z_t} \left\{ \left\| \mathbf{P}_t^{-\frac{1}{2}} \right\| \left\| \mathbf{P}_t^{-\frac{1}{2}} \hat{\boldsymbol{\theta}}_t \right\| \right\} \end{aligned} \quad (3.12)$$

ここで、一般に $\|\mathbf{A}\| = \sqrt{\lambda_{\max}(\mathbf{A}^T \mathbf{A})}$ である。(3.12)式にシュワルツの不等式を適用するとき、次式を得る。

$$\begin{aligned} E_{\theta_0, z_t} \{ \|\hat{\boldsymbol{\theta}}_t\| \} &\leq \sqrt{E_t \left\{ \left\| \mathbf{P}_t^{-\frac{1}{2}} \right\|^2 \right\} E_{\theta_0, z_t} \left\{ \left\| \mathbf{P}_t^{-\frac{1}{2}} \hat{\boldsymbol{\theta}}_t \right\|^2 \right\}} \\ &= \sqrt{E_t \{ \|\mathbf{P}_t\| \} V_t} \end{aligned} \quad (3.13)$$

ただし

$$V_t \triangleq E_{\theta_0, z_t} \{ \hat{\boldsymbol{\theta}}_t^T \mathbf{P}_t^{-1} \hat{\boldsymbol{\theta}}_t \} \quad (3.14)$$

である。(3.13)式よりつぎの関係を得る。

$$\lim_{t \rightarrow \infty} E_t \{ \|\mathbf{P}_t\| \} = 0, \lim_{t \rightarrow \infty} V_t = V \Rightarrow \lim_{t \rightarrow \infty} E_{\theta_0, z_t} \{ \|\hat{\boldsymbol{\theta}}_t\| \} = 0 \quad (3.15)$$

これより、上式前提部が満足されれば、(2.16)式の推定量は真値に収束する。すなわち、(2.15)式で得られる推定値は真値に収束する。したがって、 $\mathbf{Q}_{t-1}$ の設定指針として、(3.15)式前提部の成立ということが得られ、2.3節の①～⑦はこの指針にしたがつて、ゆらぎ $\hat{\mathbf{w}}_{t-1|t}$ の強さを調整するものである。

次節では、(3.15)式前提部第2式の具体化と、それに基づく $\mathbf{Q}_{t-1}$ の具体的設定手法が述べられる。

3.3 設定指針の具体化

(3.15)式前提部第1式が成立するためには $\mathbf{Q}_t \rightarrow \mathbf{0}, (t \rightarrow \infty)$ のほかに、(3.11)式の持続的励振条件が満足されればよい。前提部第2式の $\lim_{t \rightarrow \infty} V_t = V$ を成立させる条件を見いだすために、(3.5)式が正確に成立するとして V_t の時間進化を評価するとき、導出が付録に与えられるように次式となる。

$$V_t = V_{t-1} - \left[E_t \left\{ \frac{\boldsymbol{\Psi}_t^T \mathbf{P}_{t-1} \boldsymbol{\Psi}_t}{\mathbf{r}_t} \left(E_{\theta_0} \left\{ \frac{\boldsymbol{\varepsilon}_t^2}{\mathbf{r}_t} \mid z^t \right\} - 1 \right) \right\} \right] + V_{t-1} - V_{t|t-1} \quad (3.16)$$

ここで、 $V_{t|t-1} \triangleq E_{\theta_0, z_{t-1}} \{ \hat{\boldsymbol{\theta}}_{t-1}^T \mathbf{P}_{t-1}^{-1} \hat{\boldsymbol{\theta}}_{t-1} \}$ であり $V_{t-1} - V_{t|t-1} \geq 0$ である。これを考慮するとき、(3.16)式よりつぎの関係を得る。

$$E_{\theta_0} \left\{ \frac{\boldsymbol{\varepsilon}_t^2}{\mathbf{r}_t} \mid z^t \right\} = 1 \Rightarrow V_{t-1} \geq V_t (\geq 0) \Rightarrow \lim_{t \rightarrow \infty} V_t = V \quad (3.17)$$

したがって、(3.15)式前提部第2式を成立させるための具体化と

して、次式を得る。

$$E_{\theta_0} \left\{ \frac{\boldsymbol{\varepsilon}_t^2}{\mathbf{r}_t} \mid z^t \right\} = 1 \quad (3.18)$$

(3.5)式が正確に成立しない場合でも、 $\mathbf{Q}_t$ の設定指針として(3.18)式を用いるということは、(3.3)式、(3.4)式を考慮するとき、次式を満足するように $\mathbf{Q}_t$ を設定するということである。

$$f_t(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\theta}}_{t-1}, \hat{\sigma}_{t-1}^2) = E_{\theta_0, z_t} \{ f_t(\boldsymbol{\theta}_0, \boldsymbol{\theta}_0, \sigma_0^2) \} \quad (3.19)$$

上式は、 $\hat{\boldsymbol{\theta}}_t$ の生成において、真値 $\boldsymbol{\theta}_0, \sigma_0^2$ での目的関数値の平均値とつねに同じ値を与えるような点を $\hat{\boldsymbol{\theta}}, \hat{\sigma}_{t-1}^2$ として生成させることを意味している。したがって、(3.18)式を $\mathbf{Q}_t$ の設定指針として用いることは合理的なことである。

3.4 CARE の導出

(3.18)式は次のように書ける。

$$E_{\theta_0} \left\{ \frac{\boldsymbol{\varepsilon}_t}{\mathbf{r}_t} \mathbf{r}_t \frac{\boldsymbol{\varepsilon}_t}{\mathbf{r}_t} \mid z^t \right\} = 1 \quad (3.20)$$

これを $\mathbf{Q}_{t-1}$ に関する陽表示にするために、つぎのようにする。まず、(3.10)式と(2.16)式を用いて(2.18)式を考慮するとき、次式を得る。

$$\boldsymbol{\Psi}_t^T \hat{\mathbf{w}}_{t-1|t} = \boldsymbol{\Psi}_t^T \mathbf{Q}_{t-1} \frac{\boldsymbol{\Psi}_t}{\mathbf{r}_t} \boldsymbol{\varepsilon}_t \quad (3.21)$$

$$= \left(1 - \frac{\mathbf{r}_{t-}}{\mathbf{r}_t} \right) \boldsymbol{\varepsilon}_t \quad (3.22)$$

(3.22)式をもとに

$$\hat{\mathbf{w}}_{t-1|t} \triangleq \frac{\boldsymbol{\Psi}_t}{\|\boldsymbol{\Psi}_t\|^2} \left(1 - \frac{\mathbf{r}_{t-}}{\mathbf{r}_t} \right) \boldsymbol{\varepsilon}_t \quad (3.23)$$

を定義するとき

$$\boldsymbol{\Psi}_t^T \hat{\mathbf{w}}_{t-1|t} = \boldsymbol{\Psi}_t^T \hat{\mathbf{w}}_{t-1|t} \quad (3.24)$$

となる。これを考慮するとき、(3.21)式より次式を得る。

$$\frac{\boldsymbol{\varepsilon}_t}{\mathbf{r}_t} = \frac{\boldsymbol{\Psi}_t^T \hat{\mathbf{w}}_{t-1|t}}{\boldsymbol{\Psi}_t^T \mathbf{Q}_{t-1} \boldsymbol{\Psi}_t} = \frac{\boldsymbol{\Psi}_t^T \hat{\mathbf{w}}_{t-1|t}}{\boldsymbol{\Psi}_t^T \mathbf{Q}_{t-1} \boldsymbol{\Psi}_t} \quad (3.25)$$

これを(3.20)式に代入するとき、(3.18)式の $\mathbf{Q}_{t-1}$ に関する陽表示として次式を得る。

$$\begin{aligned} \boldsymbol{\Psi}_t^T E_{\theta_0} \left\{ \frac{\hat{\mathbf{w}}_{t-1|t} \hat{\mathbf{w}}_{t-1|t}^T}{\mathbf{r}_t} \mid z^t \right\} \boldsymbol{\Psi}_t \\ \boldsymbol{\Psi}_t^T \mathbf{Q}_{t-1} \frac{\boldsymbol{\Psi}_t \boldsymbol{\Psi}_t^T}{\mathbf{r}_t} \mathbf{Q}_{t-1} \boldsymbol{\Psi}_t = 1 \end{aligned} \quad (3.26)$$

したがって

$$\mathbf{Q}_{t-1} \frac{\boldsymbol{\Psi}_t \boldsymbol{\Psi}_t^T}{\mathbf{r}_t} \mathbf{Q}_{t-1} = E_{\theta_0} \left\{ \frac{\hat{\mathbf{w}}_{t-1|t} \hat{\mathbf{w}}_{t-1|t}^T}{\mathbf{r}_t} \mid z^t \right\} \quad (3.27)$$

であれば、(3.18)式が成立する。上式は、 $\boldsymbol{\Psi}_t \boldsymbol{\Psi}_t^T / \mathbf{r}_t$ の実現値および右辺が既知であれば $\mathbf{Q}_{t-1}$ に関する CARE である。しかし、 $\mathbf{r}_t$ には $\mathbf{P}_{t|t-1}$ を通じて $\mathbf{Q}_{t-1}$ が含まれており、計算不可能である。そして、この計算不可能な $\mathbf{r}_t$ が $\hat{\mathbf{w}}_{t-1|t}$ の中にも含まれているので、(3.27)式右辺も計算不可能であり、(3.27)式はこのままでは解けない。そこで、 $\mathbf{r}_t$ のかわりに(2.18)式の $\mathbf{r}_{t-}$ を用いるために、つぎのようにする。まず、次式を定義する。

$$\tau_t \triangleq E_{\theta_0} \left\{ \frac{\boldsymbol{\varepsilon}_t^2}{\mathbf{r}_{t-}} \mid z^t \right\} \left(= \frac{E_{\theta_0} \{ \boldsymbol{\varepsilon}_t^2 \mid z^t \}}{\mathbf{r}_{t-}} \right) \quad (3.28)$$

したがって、 τ_t は(3.8)式における ε_t の実際の分散と ε_t がいないとした場合の ε_t の計算上の分散との比であり、 ε_t の中に残留す

る非線形性 ζ_t の多少に関する情報を有している。 $\tau_t > 1$ の場合は ε_t の中に ζ_t が多く残留しており、推定がうまく行なわれていない状況下であり、 $\tau_t = 1$ の場合は ε_t の中に ζ_t の残留がなく、推定がうまく行なわれている状況にあることを示している。反対に、 $\tau_t < 1$ の場合は ε_t の中に ζ_t が残留しているが、 $\mathbf{P}_{t-1}$ が大きめに出ているために、(3.9)式右辺第2項が $\hat{\theta}_t$ の更新に生かされており、この意味で推定がうまく行なわれていることを示している。したがって、 τ_t はモデルの非線形性に起因して生じる ζ_t が ε_t の中にどの程度残留しているか、という情報を有しており、推定量の作動状況の良否に関する指標の一つになる。本論文ではこれを仮に非線形性残留度と呼び、そのような情報をゆらぎ調整に生かすことは有用なことである。そこで、 τ_t と $\mathbf{Q}_{t-1}$ を次式で関係づける。

$$\tau_t \triangleq \begin{cases} \frac{r_t}{r_{t,-}}, & \text{if } \tau_t \geq 1 \\ \frac{r_{t,-}}{r_t}, & \text{if } \tau_t < 1 \end{cases} \quad (3.29)$$

このとき、 τ_t と $\mathbf{Q}_{t-1}$ の関係が次式のように陽になる。

$$\left. \begin{aligned} \tau_t &= 1 + \frac{\Psi_t^T \mathbf{Q}_{t-1} \Psi_t}{r_{t,-}} \leq \bar{\tau}_t, & \text{if } \tau_t \geq 1 \\ \frac{1}{\tau_t} &= 1 + \frac{\Psi_t^T \mathbf{Q}_{t-1} \Psi_t}{r_t} \leq \bar{\tau}_t, & \text{if } \tau_t < 1 \end{aligned} \right\} \quad (3.30)$$

ただし

$$\bar{\tau}_t \triangleq 1 + \frac{\lambda_{\max}(\mathbf{Q}_{t-1}) \|\Psi_t\|^2}{r_{t,-}} \quad (3.31)$$

である。(3.30)式より、 $\tau_t = 1$ の場合、すなわち ε_t の中に非線形性の影響が残留していない場合は $\mathbf{Q}_{t-1} = \mathbf{O}$ に、そうでない場合は $\mathbf{Q}_{t-1} > 0$ に設定することが必要になることがわかる。

さて、(3.29)式、(3.23)式を考慮するとき、(3.27)式は $\tau_t \geq 1$ あるいは $\tau_t < 1$ のいずれの場合にもつぎのように表わされる。

$$\mathbf{Q}_{t-1} \frac{\Psi_t \Psi_t^T}{r_{t,-}} \mathbf{Q}_{t-1} = \alpha_t \frac{\Psi_t \Psi_t^T}{\|\Psi_t\|^4} E\{\varepsilon_t^2 | z^t\} \quad (3.32)$$

ただし

$$\alpha_t \triangleq \left(\sqrt{\tau_t} - \frac{1}{\sqrt{\tau_t}} \right)^2 \quad (3.33)$$

であり、 τ_t を推定値で代用したものが(2.24)式あるいは(2.27)式である。いま、 $\mathbf{Q}_{t-1}$ の調整は緩やかに変化させるとして、時間区間 L 内では $\mathbf{Q}_{t-1}$ は一定とする。すなわち

$$\mathbf{Q}_{t-1} = \mathbf{Q}_t, \quad i \in [t-L+1, t] \quad (3.34)$$

を仮定する。このもとで、 $[t-L+1, t]$ で(3.32)式の標本移動平均をとり、数値計算上の不安定を避けるために修正して、 $\mathbf{Q}_t$ が求解可能になるようにしたのが(2.28)式第1式である。

3.5 非線形性残留度 τ_t の推定

(2.17)式のもとで、(3.28)式は $\tau_t = E\{v_t^2 | z^t\}$ と表わされる。これを推定するための方法の一つとして、 z^t が与えられたときの v_t の確率分布モデルをいくつか用意し、実現値 $\{v_t\}$ がいくつかのモデルに最もよく適合するかをAICを用いて特定し、特定されたモデルをもとに推定することが考えられる。

いま、 v_t の L 個の実現値

$$v_{[t, t-L+1]} \triangleq \{v_t, v_{t-1}, \dots, v_{t-L+1}\} \quad (3.35)$$

をもとに、これらを生み出している確率分布モデルとして、つぎの4種類の正規分布モデルを用意する。

$$N[0, 1], N[\mu_t, 1], N[0, \sigma_{v,t}^2], N[\mu_t, \sigma_{v,t}^2] \quad (3.36)$$

このとき、(2.19)式は、 $v_{[t, t-L+1]}$ を上式の4種類のモデルに当てはめたときのAICであり、 $v_{[t, t-L+1]}$ が最もよく適合する確率分布モデルは、つぎのように特定できる⁷⁾。

$$v_{[t, t-L+1]} \sim \begin{cases} N[0, 1] & , \text{if MAICE} = \text{AIC}(0, 1) \\ N[\mu_t, 1] & , \text{if MAICE} = \text{AIC}(\mu_t, 1) \\ N[0, \sigma_{v,t}^2] & , \text{if MAICE} = \text{AIC}(0, \sigma_{v,t}^2) \\ N[\mu_t, \sigma_{v,t}^2] & , \text{if MAICE} = \text{AIC}(\mu_t, \sigma_{v,t}^2) \end{cases} \quad (3.37)$$

ここで、一般に $x \sim N[m, \sigma^2]$ は確率変数 x が $N[m, \sigma^2]$ の正規分布に従うことを示す。(3.37)式より $v_{[t, t-L+1]}$ が最もよく適合する確率分布モデルが特定できるが、特定されたモデルにもとづく τ_t の暫定推定値 $\hat{\tau}_{t,-}$ が(2.21)式である。この $\hat{\tau}_{t,-}$ はあくまでも推定値であるので、 $\hat{\tau}_{t,-}$ あるいは $1/\hat{\tau}_{t,-}$ は(3.30)式の上限 $\bar{\tau}_t$ を越える可能性がある。そのような場合は、 $\bar{\tau}_t$ を越えないように $\hat{\tau}_{t,-}$ に修正を加える必要があり、それが2.3節ではつぎのように行なわれている。

まず、(3.34)式を考慮するとき、(2.28)式第1式に対して、固有値に関する次の不等式が成立する⁸⁾。

$$\lambda_i(\mathbf{Q}_{t-1} \mathbf{A}_t \mathbf{Q}_{t-1}) = \lambda_i(\mathbf{B}_t) \leq \lambda_{\max}(\mathbf{A}_t) \lambda_i^2(\mathbf{Q}_{t-1}) \quad (3.38)$$

これと、 $\lambda_{\max}(\mathbf{Q}_{t-1}) \geq \text{Tr}(\mathbf{Q}_{t-1})/n$ の関係をを用いるとき、次式を得る。

$$\begin{aligned} \lambda_{\max}(\mathbf{Q}_{t-1}) &\geq \frac{\text{Tr}(\mathbf{Q}_{t-1})}{n} \geq \frac{\sum_{i=1}^n \sqrt{\lambda_i(\mathbf{B}_t)}}{n \sqrt{\lambda_{\max}(\mathbf{A}_t)}} \\ &\geq \frac{1}{n} \sqrt{\frac{\lambda_{\min}(\mathbf{B}_t)}{\lambda_{\max}(\mathbf{A}_t)}} \end{aligned} \quad (3.39)$$

これを(3.31)式に用いるとき、(2.25)式を定義するならば、 $\bar{\tau}_t$ は計算可能で、しかも $\bar{\tau}_t \leq \bar{\tau}_t$ となり、(3.30)式の上限の内側にある。したがって、 τ_t の推定値 $\hat{\tau}_t$ は、 $\hat{\tau}_t$ あるいは $1/\hat{\tau}_t$ が $\bar{\tau}_t$ を越えないようにしておけば、(3.30)式は自動的に満足される。これを具体化したのが(2.26)式である。

Fig.1は、(2.27)式の α_t の形状である。これにおいて、 $\hat{\tau}_t \equiv 1$ のときは $\alpha_t \equiv 0$ となり、(2.23)式より $\mathbf{B}_t \equiv \delta_Q \mathbf{I}$ となる。すなわち、 ε_t の中に ζ_t の残留がないと推定された場合は、 $\mathbf{Q}_t$ は $\mathbf{Q}_t \mathbf{A}_t \mathbf{Q}_t = \delta_Q \mathbf{I}$ の解として得られ、 $\mathbf{Q}_t$ は $\mathbf{O}$ ではないが非常に小さく設定される。逆に、 $\hat{\tau}_t \gg 1$ あるいは $\hat{\tau}_t \ll 1$ の場合は α_t が大きくなり、 $\mathbf{B}_t$ も大きくなる。したがって、 ε_t の中に ζ_t が多く残留していると推定された場合は、 $\mathbf{Q}_t$ は大きく設定される。

4. 数値実験例

Fig.2のシステム⁹⁾に対して実験を行なった。すなわち

$$\hat{y}_t(\theta) = \frac{1}{1 + e^{-(\theta_1 u_t + \theta_2)}} = \frac{1}{1 + e^{-\Phi^T \theta}} \quad (4.1)$$

である。ただし、

$$\Phi^T = [u_t \quad 1] \quad (4.2)$$

である。推定すべき未知パラメータの真値は

Fig.1 Shape of the α_i of eq.(2.27)

Fig.2 System to be identified

Fig.3 Shape of the object function of eq.(4.5)

$$\theta_0 = [\theta_{0,1} \ \theta_{0,2}]^T = [4 \ 0.2]^T \quad (4.3)$$

であり, u_t, e_t はそれぞれつぎの正規乱数を用いた.

$$u_t \sim N[0, 1], \quad e_t \sim N[0, 0.1] \quad (4.4)$$

Fig.3 は最小 2 乗規範

$$S(\theta) = \frac{1}{10000} \sum_{t=1}^{10000} \{y_t - \hat{y}_t(\theta)\}^2 \quad (4.5)$$

の形状を示す. 本数値実験は(4.3)式の θ_0 を推定するものであるが, これは Fig.3 において, ある出発点 S から $S(\theta)$ 上を降下して行き, $S(\theta)$ の最下点に到達することと同じである. 実験は, 推定性能比較のために, つぎの 4 種類の推定量を用いた.

IEKF1: ゆらぎ調整を行なう IEKF(本手法)

IEKF0: ゆらぎ調整を行なわない IEKF((2.15)式で $Q_{t-1} \equiv \mathbf{0}$ としたもの)

EKF1: ゆらぎ調整を行なう EKF((2.15)式で $l \equiv 1$ としたもの)

EKF0: ゆらぎ調整を行なわない EKF((2.15)式で $Q_{t-1} \equiv \mathbf{0}$,

Fig.4 Time evolutions of parameter estimates by IEKF

Fig.5 Time evolutions of parameter estimates by EKF

$l \equiv 1$ としたもの)

また, 実験は 100 本の見本系列に対して行ない, 推定値は次式で評価した.

$$\hat{\theta}_t = \frac{1}{100} \sum_{j=1}^{100} \hat{\theta}_t^{(j)} \quad (4.6)$$

ただし, $\hat{\theta}_t^{(j)}$ は j 番目の見本系列に対する $\hat{\theta}_t$ である. 用いた推定量の初期値などは, つぎの通りである.

$$\left. \begin{aligned} \hat{\theta}_0 &= [-20 \ -20]^T, \quad P_0 = I, \quad \hat{\sigma}_0^2 = 1 \\ \delta &= 10^{-12}, \quad L = 50, \quad \delta_\psi = 10^{-5}, \quad \delta_Q = 10^{-12} \end{aligned} \right\} \quad (4.7)$$

Fig.4, Fig.5 はそれぞれ IEKF, EKF による θ_0 の(4.6)式に基づく評価結果である. これらの結果より, 本手法は良好な推定結果を与えていることが確認できる. EKF1, IEKF0, EKF0 は推定結果がよくないが, EKF1 については IEKF0, EKF0 に比べると推定精度は幾分向上している.

Fig.4, Fig.5 において, 4 種類の推定量性能に大きい違いが出てくるのは, 推定環境がつぎのような状況下にあることによる. まず, (4.1)式より次式を得る.

$$\psi_t(\theta) = \varphi_t \hat{y}_t(\theta) \frac{e^{-\varphi_t^T \theta}}{1 + e^{-\varphi_t^T \theta}} \quad (4.8)$$

これより, $\varphi_t^T \theta \ll 0$ の場合は $e^{-\varphi_t^T \theta} \gg 1$ となり, (4.1)式を考慮するとき $\psi_t(\theta) \approx 0$ となるので, (3.11)式が満足されなくなる.

Fig. 6 Time evolutions of gain adjustment factor Q_i

その結果、推定値更新のための情報がほとんど入ってこなく、(2.15d)式のゲインが $k_i^{(i-1)} = 0$ となり、推定値が更新されない。本実験においては、 $\hat{\theta}_0$ として(4.7)式のように与えているので $e^{-\theta^T \theta} = e^{20u_1 + 20}$ となり、 $u_1 \geq 0$ の場合は $\psi_1 = 0$ となって $t=1$ から推定環境が非常に厳しい状況にある。このような中で、推定量の個別的な特性はつぎようになる。

IEKF1については(2.28)式に基づき $t \geq L (=50)$ でゆらぎ調整が実施されるが、 $t < L$ ではゆらぎ調整が実施される前である。したがって、Fig. 4においてIEKF1は $t < L$ ではIEKF0と同じ性能で $\|\hat{\theta}_i\|$ は大きい。しかし、 $t = L$ でゆらぎ調整が開始されると、Fig. 6におけるように Q_i が一気に大きく設定され、 k_i も大きくなる。その結果、 $\hat{\theta}_i$ の更新のために大量の $\hat{w}_{i-1|i}$ が投入され、しかもその投入のされ方は(3.17)式あるいは(3.19)式を満足させるように行なわれるので、 $\hat{\theta}_i$ の精度が向上する。このことが、Fig. 4における $t \geq L$ の $\hat{\theta}_i$ の挙動である。 $\|\hat{\theta}_i\|$ が小さくなると、(2.26)式の $\hat{\tau}_i$ は $\hat{\tau}_i \approx 1$ となり、(2.27)式の α_i は $\alpha_i = 0$ となる。したがって、(2.28)式は $Q_i A_i Q_i = \delta_Q I = 10^{-12} I$ となり、 Q_i は小さくなる。これがFig. 6におけるIEKF1の $t \geq 600$ の状況である。すなわち、IEKF1は $\|\hat{\theta}_i\|$ が大きい場合は Q_i を大きくして $\|\hat{\theta}_i\|$ を小さくするようにゆらぎ調整をし、 $\|\hat{\theta}_i\|$ が小さくなれば Q_i も小さくする。これがFig. 4のIEKF1の状況である。Fig. 7は、 e_i における $\hat{\tau}_i$ の残留状況を

$$\hat{\tau}_i = \frac{1}{100} \sum_{j=1}^{100} \hat{\tau}_i^{(j)} \quad (4.9)$$

をで評価したものであり、 $\hat{\tau}_i^{(j)}$ は j 番目の見本系列に対する $\hat{\tau}_i$ のことである。Fig. 7より $\hat{\tau}_i$ は1に漸近している状況が確認でき、このことは $\hat{\tau}_i$ が消滅しつつあることを示している。

EKF1についてもFig. 6からうかがえるように、 $t \geq L$ からは Q_i が大きく設定されてゆらぎ調整が実施され、 $\hat{\theta}_i$ の更新のために多量の $\hat{w}_{i-1|i}$ が投入されている。実際、Fig. 5から $t \geq L$ で $\hat{\theta}_i$ が更新されている状況が確認でき、また、Fig. 7からは $\hat{\tau}_i$ が1に漸近している状況が確認できる。しかし、EKF1がIEKF1と決定的に異なる点は、EKF1はIEKF1に比べて $\hat{w}_{i-1|i}$ の投入が必ずしも(3.17)式あるいは(3.19)式を満足させるように行なわれるとは限らない、ということである。したがって、 Q_i が0で

Fig. 7 Time evolutions of $\hat{\tau}_i$

ないかぎり $\hat{w}_{i-1|i}$ は投入され続けるが、それが $\|\hat{\theta}_i\|$ の減少に繋がる保証はない。その結果、 $\psi_i = 0$ の状況下での(2.28)式は、(2.22)式、(2.23)式を考慮するとき $\delta_Q Q_i = \delta_Q I$ となり、 $Q_i = I$ となる。Fig. 6は100本の見本系列に対する平均的な挙動であり、いくつかの見本系列に対してはEKF1は良好に作動したが、多くの見本系列に対して不良作動をし、全体としてはIEKF1に比べて $Q_i \approx I$ となった。これがFig. 6のEKF1における $t \geq 200$ の状況である。

IEKF0, EKF0については $Q_i = 0$ であるので、 $\psi_i = 0$ の状況が続いても $\hat{w}_{i-1|i}$ は投入されず、 $k_i \approx 0$ となって $\hat{\theta}_i \approx \hat{\theta}_{i-1}$ の状況が続き、推定値はほとんど更新されない。これがそれぞれFig. 4, Fig. 5におけるIEKF0, EKF0の状況である。元来、EKF0は $\|\hat{\theta}_i\|$ が小さい場合には良好に作動し、IEKF0は $\|\hat{\theta}_i\|$ を可能な限り小さくするように作動する。したがって、IEKF0の方がEKF0に比べて性能がよくなると言われている¹⁰⁾。しかし、これも程度もので、Fig. 4からも確認できるように、 $\|\hat{\theta}_i\|$ が大きい場合はIEKF0も性能はよくなる。これを改善するためには何らかの策を講じることが必要になるが、その一つとして本手法は有効であると考えられる。

5. おわりに

本論文では、回帰型非線形システムにおける未知パラメータベクトルの逐次型MAP推定量において、確率的ゆらぎを加える場合に、ゆらぎの強さを調整するための一つの手法として、2.3節①～⑦を提案した。その考え方の出発点は、(3.13)式を0に収束させるということであった。これを具体化したとき、(3.15)式を経由して(3.18)式を満足するように調整すればよい、という指針が得られた。これを調整用パラメータ Q_i で陽に表示するとき、(2.28)式第1式のCAREとなり、本推定量はこのCAREの求解と連動させる形となった。

本手法の有効性検証については、理論的性能評価が困難であるので、Fig. 2のシステムに対して数値実験による検証が行なわれた。その結果、つぎのことが確認された。

- ① IEKFは、ゆらぎを与えると良好に作動する。
- ② EKFは、ゆらぎを与えても良好に作動するとは限らない。

③IEKF, EKFは, ゆらぎを与えない場合は, 設定初期値, データ数にもよるが, 良好に作動しない場合が多い.

これらのことより, 本手法はゆらぎ調整に関して有効な一つの手法であると考えられる.

参考文献

- 1) T.L.Vincent and P.P.Khargonekar: A Class of Nonlinear Filtering Problems Arising from Drifting Sensor Gains, IEEE Trans. Automat. Contr., AC-44, 510/520 (1999)
- 2) S. Amari, H.Park and K.Fukumizu: Adaptive Method of Realizing Natural Gradient Learning for Multilayer Perceptrons, Neural Computation, 12, 1399/1409 (2000)
- 3) 豊田秀樹: 非線形多変量解析, 62/63, 朝倉書店 (1998)
- 4) 船橋誠壽: ニューロコンピューティング入門, 第2章, オーム社, (1992)
- 5) B.M.Bell and F.W.Cathey: The Iterated Kalman Filter Update as a Gauss-Newton Method, IEEE Trans. Automat. Contr., AC-38, 294/297 (1993)
- 6) B.M.Bell: The Iterated Kalman Smoother as a Gauss-Newton Method, SIAM J. Optimization, 4, 626/636 (1994)
- 7) 坂本, 石黒, 北川: 情報量統計学, 110/113, 共立出版 (1983)
- 8) U.Helmke and J.B.Moore: Optimization and Dynamical Systems, 310/311, Springer-Verlag (1994)
- 9) C-S.Leung, A-C.Tsoi and L.W.Chan: Efficient Training of Neural Nets for Nonlinear Adaptive Filtering Using a Recursive Levenberg-Marquardt Algorithm, IEEE Trans. Signal Processing, SP-48, 1915/1927 (2000)
- 10) J.M.Mendel: Lessons in Digital Estimation Theory, 248/254, Prentice-Hall (1987)

付録 (3.16)式の導出

まず, (2.16a)式をもとに

$$\tilde{\theta}_t = \tilde{\theta}_{t-1} - \mathbf{k}_t \varepsilon_t \quad (\text{A.1})$$

である. これを用いるとき, (3.14)式はつぎのようになる.

$$\begin{aligned} V_t &= E_{\theta_0, z_t} \left\{ (\tilde{\theta}_{t-1} - \mathbf{k}_t \varepsilon_t)^T \mathbf{P}_{t-1}^{-1} (\tilde{\theta}_{t-1} - \mathbf{k}_t \varepsilon_t) \right\} \\ &= E_{\theta_0, z_t} \left\{ \tilde{\theta}_{t-1}^T \mathbf{P}_{t-1}^{-1} \tilde{\theta}_{t-1} - 2 \mathbf{k}_t^T \mathbf{P}_{t-1}^{-1} \tilde{\theta}_{t-1} \varepsilon_t \right. \\ &\quad \left. + \mathbf{k}_t^T \mathbf{P}_{t-1}^{-1} \mathbf{k}_t \varepsilon_t^2 \right\} \quad (\text{A.2}) \end{aligned}$$

ここで, (2.16f)式, (2.16d)式は

$$\mathbf{P}_t^{-1} = \mathbf{P}_{t-1}^{-1} + \frac{\Psi_t \Psi_t^T}{\hat{\sigma}_{t-1}^2}, \quad \mathbf{k}_t = \frac{\mathbf{P}_t \Psi_t}{\hat{\sigma}_{t-1}^2}, \quad \frac{\Psi_t^T \mathbf{P}_t \Psi_t}{\hat{\sigma}_{t-1}^2} = \frac{\Psi_t^T \mathbf{P}_{t-1} \Psi_t}{r_t}$$

と表わされることを用いるとき, (A.2)式はつぎのようになる.

$$\begin{aligned} V_t &= V_{t|t-1} + E_{\theta_0, z_t} \left\{ \frac{(\Psi_t^T \tilde{\theta}_{t-1})^2}{\hat{\sigma}_{t-1}^2} - 2 \frac{\Psi_t^T \tilde{\theta}_{t-1}}{\hat{\sigma}_{t-1}^2} \varepsilon_t + \frac{\Psi_t^T \mathbf{k}_t}{\hat{\sigma}_{t-1}^2} \varepsilon_t^2 \right\} \\ &= V_{t|t-1} + E_{\theta_0, z_t} \left\{ \frac{[\varepsilon_t - (\zeta_t + e_t)]^2}{\hat{\sigma}_{t-1}^2} - 2 \frac{\varepsilon_t - (\zeta_t + e_t)}{\hat{\sigma}_{t-1}^2} \varepsilon_t \right. \\ &\quad \left. + \frac{1}{\hat{\sigma}_{t-1}^2} \frac{\Psi_t^T \mathbf{P}_{t-1} \Psi_t}{r_t} \varepsilon_t^2 \right\} \\ &= V_{t|t-1} - E_{\theta_0, z_t} \left\{ \frac{\varepsilon_t^2}{r_t} - \frac{(\zeta_t + e_t)^2}{\hat{\sigma}_{t-1}^2} \right\} \quad (\text{A.3}) \end{aligned}$$

ここで(3.6)式からの $\zeta_t + e_t = \hat{e}_t - \Psi_t^T \tilde{\theta}_t$ と(3.1)式および(3.5)式の等号での成立の仮定を用いて, (A.3)式はつぎのようになる.

$$\begin{aligned} V_t &= V_{t|t-1} - E_{\theta_0, z_t} \left\{ \frac{\varepsilon_t^2}{r_t} - \frac{(\hat{e}_t - \Psi_t^T \tilde{\theta}_t)^2}{\hat{\sigma}_{t-1}^2} \right\} \\ &= V_{t|t-1} - E_{\theta_0, z_t} \left\{ \frac{\varepsilon_t^2}{r_t} - \frac{\hat{\sigma}_{t-1}^2 \varepsilon_t^2 - \Psi_t^T \mathbf{P}_t \Psi_t}{\hat{\sigma}_{t-1}^2} \right\} \\ &= V_{t|t-1} - E_{\theta_0, z_t} \left\{ \frac{\varepsilon_t^2}{r_t} - \frac{\hat{\sigma}_{t-1}^2 \varepsilon_t^2 - \Psi_t^T \mathbf{P}_{t-1} \Psi_t}{r_t} \right\} \\ &= V_{t-1} - E_{z_t} \left\{ \frac{\Psi_t^T \mathbf{P}_{t-1} \Psi_t}{r_t} \left(E_{\theta_0} \left\{ \frac{\varepsilon_t^2}{r_t} \mid z_t \right\} - 1 \right) \right\} \\ &\quad + V_{t-1} - V_{t|t-1} \quad (\text{A.4}) \end{aligned}$$

これが(3.16)式である.

「著者紹介」

笠 松 博 史 (学生会員)

2001年, 徳島文理大学家政学部卒業. 同年徳島文理大学大学院工学研究科入学. 現在, 同研究科在学中. 適応システムの解析・設計とその人間生活への応用に関する内容を研究中.

樋 口 彰 (正会員)

1963年, 東京電機大学工学部卒業. 防衛大学校勤務を経て89年徳島文理大学勤務. 現在工学部助教授. 二輪車の自動制御装置の研究開発に従事.

森 本 滋 郎 (正会員)

1998年, 徳島文理大学大学院工学研究科修了. 同年徳島文理大学工学部助手, 2003年同講師, 現在に至る. 適応信号処理のアルゴリズム開発とその画像処理問題への応用を研究中(博士(工学)).

古 本 奈 奈 代

1980年, 神戸大学理学部数学科卒業. 徳島大学勤務を経て87年徳島文理大学勤務. 現在人間生活学部教授. 非線形非定常データの統計的解析の研究に従事.(博士(工学)).

田 淵 敏 明 (正会員)

1976年, 徳島大学大学院工学研究科修了. 徳島大学などの勤務を経て86年徳島文理大学勤務. 現在工学部教授. 非線形システムのモデリングとそのデータ解析への応用に関する内容を研究中.(工学博士)