'88 KACC 1988.10.21~22

ADAPTIVE CHANDRASEKHAR FILTER FOR LINEAR DISCRETE-TIME STATIONALY STOCHASTIC SYSTEMS

Masanori Sugisaka

Department of Electrical Engineering
Faculty of Engineering
Oita University
Oita, 870-11, Japan

Abstract: This paper considers the design problem of adaptive filters based on the state-space models for linear discrete-time stationary stochastic signal processes. The adaptive state estimator consists of both the predictor and the sequential prediction error estimator. The discrete Chandrasekhar filter developed by author is employed as the predictor and the nonlinear least-squares estimator is used as the sequential prediction error estimator. Two models are presented for calculating the parameter sensitivity functions in the adaptive filter. One is the exact model called the linear innovations model and the other is the simplified model obtained by neglecting the sensitivities of the Chandrasekhar $\mathcal{X}$ and $\mathcal{Y}$ functions with respect to the unknown parameters in the exact model.

1. Introduction

In many digital control systems or signal processes, the state-space models(SSM's) are not known completely and, in general, involve the unknown parameters. In the state estimation problems for such control systems or signal processes both filter and estimation algorithm to estimate unknown parameters must be used simultaneously in order to avoid the instability of the filter, viz., the divergence of the filter due to the insufficient models with unknown parameters.

Although for this problem the digital adaptive filter based on the SSM has been proposed[1] using the Kalman filter, this adaptive filter has disadvantages that the computational storage and time are large and the calculation of the filter becomes unstable. The digital Chandrasekhar-type filters, which have less computational time and storage than the Kalman filter, have been presented based on the SSM [2].

This paper considers the design problem of digital adaptive filters using vector Chandrasekhar-type filters where the system parameters are partially unknown. Namely, the vector adaptive Chandrasekhar-type filtering algorithms based on the SSM are derived using the sequential nonlinear least-squares algorithms(SNLSA's) as the parameter estimators. This paper is an extension of the scalar digital adaptive Chandrasekhar-type filters[3],[4]. The computational time and storage in calculating the proposed discrete adaptive Chandrasekhar filter is less than that in calculating the adaptive Kalman filter. The reason is that the number of the difference equations to be solved in the Chandrasekhar filter is {n+2nm} and that in the Kalman filter is {n+n(n+1)/2} where n is the dimension of the system states and m is the dimension of the measured outputs. If $n \gg m$, then the computational time and storage in calculating the discrete adaptive Chandrasekhar filter is greatly reduced compared with the adaptive Kalman filter.

In this paper especially two special cases are considered. The first case is that the unknown parameters are involved in only the state matrix and the second case is that the parameters are involved in only the observation matrix. It is shown that for the two special cases the adaptive Chandrasekhar filter becomes simpler than that for the general case.

2. Chandrasekhar Filter

Consider the linear stationary discrete-time dynamic system and measurements with unknown parameters given by

$$x(t+1)=F(\theta)x(t)+G(\theta)u(t),x(0)=x_0(\text{I.C.}), \quad (2.1)$$

$$z(t)=H(\theta)x(t), \quad (2.2)$$

$$y(t)=z(t)+v(t), t\in[0,c], \quad (2.3)$$

where t and c are integers and y, z, $v\in R^m$, $x\in R^n$, $u\in R^p$, $x_0\in R^n$, $F(\theta)\in R^{n\times n}$, $G(\theta)\in R^{n\times p}$, and $H(\theta)\in R^{m\times n}$ are measurement, signal, measurement noise, state, input noise, initial condition, system matrix, input matrix, and output matrix, respectively. It is assumed that the noises have the following properties.

$$E[x_0]=E[u(t)]=E[v(t)]=0, \quad (2.4a)$$

$$E[u(t)x_0^T]=0,\ E[v(t)x_0^T]=0,\ E[u(t)v^T(s)]=0, \quad (2.4b)$$

$$E[x_0x_0^T]=\Pi_0\in R^{n\times n}, \quad (2.4c)$$

$$E[u(t)u^T(s)]=Q\delta_{t,s}\in R^{p\times p}\ (Q\geq 0), \quad (2.4d)$$

$$E[v(t)v^T(t)]=R\delta_{t,s}\in R^{m\times m}\ (R>0), \quad (2.4e)$$

Notations T and $\delta_{t,s}$ denote transpose and Kroneker's delta function. Suppose that the system, input, and output matrices contain the unknown parameters θ. For the case where the estimates $\hat{\theta}$ of the unknown parameters have been obtained, the Chandrasekhar filtering equations are represented by[2]

$$\hat{x}(t+1)=F(\hat{\theta}(t))x(t)+\mathcal{K}(t+1)[y(t+1)-H(\hat{\theta}(t))\hat{x}(t+1/t)], \quad (2.5a)$$

$$\hat{z}(t)=H(\hat{\theta}(t))\hat{x}(t)=\hat{y}(t), \quad (2.5b)$$

$$\hat{x}(t+1/t)=F(\hat{\theta}(t))\hat{x}(t), \quad (2.5c)$$

$$\hat{x}(0)=\mathcal{Y}(0)y(0)(I.C.), \quad (2.5d)$$

where the Chandrasekhar $\mathcal{K}$ function appeared in (2.5a) is calculated from a pair of difference equations for the Chandrasekhar $\mathcal{K}$ and $\mathcal{Y}$ functions,

$$\mathcal{K}(t+1)=\mathcal{K}(t)-\mathcal{Y}(t+1)\times[H(\hat{\theta}(t))F(\hat{\theta}(t))\mathcal{Y}(t)R]^T R^{-1}, \quad (2.6a)$$

$$\mathcal{K}(0)=\Pi_0 H^T(\hat{\theta}(0))[R+H(\hat{\theta}(0))\Pi_0 H^T(\hat{\theta}(0))]^{-1} \quad (I.C.), \quad (2.6b)$$

$$\mathcal{Y}(t+1)=[F(\hat{\theta}(t))-\mathcal{K}(t+1)H(\hat{\theta}(t))F(\hat{\theta}(t))]\mathcal{Y}(t) \quad (2.7a)$$

$$=[I-\mathcal{K}(t)H(\hat{\theta}(t))]F(\hat{\theta}(t))\mathcal{Y}(t)\times[I-R(H(\hat{\theta}(t))F(\hat{\theta}(t))\mathcal{Y}(t))^T R^{-1}\times H(\hat{\theta}(t))F(\hat{\theta}(t))\mathcal{Y}(t)]^{-1}, \quad (2.7b)$$

$$\mathcal{Y}(0)=\mathcal{K}(0)(I.C.), \quad (2.7c)$$

or alternatively a pair of difference equations given by

$$\mathcal{K}(t+1)=[\mathcal{K}(t)-F(\hat{\theta}(t))\mathcal{Y}(t)(H(\hat{\theta}(t))F(\hat{\theta}(t))\times\mathcal{Y}(t)R)^T R^{-1}]\times[I-(H(\hat{\theta}(t))F(\hat{\theta}(t))\mathcal{Y}(t))R\times(H(\hat{\theta}(t))F(\hat{\theta}(t))\mathcal{Y}(t))^T R^{-1}]^{-1}, \quad (2.8)$$

$$\mathcal{Y}(t+1)=[I-\mathcal{K}(t+1)H(\hat{\theta}(t))]F(\hat{\theta}(t))\mathcal{Y}(t). \quad (2.9)$$

We note that either (2.6)-(2.7) or (2.8)-(2.9) is used to evaluate the Chandrasekhar $\mathcal{K}$ and $\mathcal{Y}$ functions in the Chandrasekhar filtering. Figure 1 shows the blockdiagram representation for the discrete-time dynamics of the Chandrasekhar $\mathcal{K}$ and $\mathcal{Y}$ functions. The Chandrasekhar filter shown above is used as a predictor in the adaptive filtering scheme.

3. Sequential Prediction Error Parameter Estimator[1]

In the preceeding section, the vector measurement case was considered for the Chandrasekhar filtering technique. In the suceedings, the scalar measurement case is considered in the derivations of the adaptive filter for notational simplicity. Let the measurement be specified by

$$y(t)=H(\hat{\theta}(t))\hat{x}(t)+v(t), \quad (3.1a)$$
$$=\hat{y}(t)+v(t), \quad (3.1b)$$

$$\hat{y}(t)=H(\hat{\theta}(t))\hat{x}(t), \quad (3.1c)$$

where y is scalar measurement and $H\in R^{1\times n}$ is the column vector. As the sequential prediction error estimation, the SNLSA is employed to update the estimates of the unknown paramater. The algorithm for the scalar measurement case is specified by

$$\hat{\theta}(t)=\hat{\theta}(t-1)+P(t-1)\Phi(t-1)[y(t)-y(t,\hat{\theta}(t-1))], \quad (3.2)$$

$$P(t-1)=P(t-2)-[P(t-2)\Phi(t-1)\Phi(t-1)^T P(t-2)]/[1+\Phi(t-1)^T P(t-2)\Phi(t-1)], \quad (3.3)$$

where

$$\hat{y}(t,\theta)=H(\theta)\hat{x}(t), \quad (3.4)$$

$$\Phi(t-1)=[d\hat{y}(t,\theta)/d\theta]\Big|_{\theta=\hat{\theta}(t-1)}, \quad (3.5a)$$

$$=-[d\varepsilon(t,\theta)/d\theta]\Big|_{\theta=\hat{\theta}(t-1)}, \quad (3.5b)$$

$$\varepsilon(t,\theta)=y(t)-\hat{y}(t,\theta), \quad (3.6)$$

and the initial condition are given by

$$\hat{\theta}(0)=0, \quad (3.7)$$

$$p(0)=\alpha I(\alpha\gg1), \quad (3.8)$$

respectively. In the next section, the calculation procedure is presented for the Φ function in the SNLSA.

4. Calculation Procedure for Φ-Function

Let the number of the unknown paramaters to be estimated in the system be q and then the Φ-function has q elements such that

$$\Phi(t)=[\Phi_1(t),\Phi_2(t),\cdots,\Phi_q(t)]^T, \quad (4.1)$$

where

$$\Phi_i(t-1)=[d\hat{y}(t,\theta)/d\theta_i]\Big|_{\theta_i=\hat{\theta}_i(t-1)},\quad i=1,2,\cdots,q. \quad (4.2)$$

Hence, the calculation of the Φ_i-function yields

$$\Phi_i(t-1)=H_{\theta i}(\theta)\hat{x}(t)+H(\theta)\hat{x}_{\theta i}(t), \quad (4.3)$$

where the notations

$$H_{\theta i}(\theta)=\partial H(\theta)/\partial\theta_i, \quad (4.4a)$$
$$x_{\theta i}(t)=\partial x(t)/\partial\theta_i, \quad (4.4b)$$

were used for the partial differentiation. We present the algorithms for calculating the sensitivity functions of the output vector H and the state x with respect to the unknown parameters θ_i.

Differentiate the filtering up-date equation (2.5a) with respect to θ_i to obtain

$$\hat{x}_{\theta i}(t+1)=[-\mathcal{X}_{\theta i}(t+1)H(\theta)-\mathcal{X}(t+1)H_{\theta i}(\theta)]F(\theta)x(t)$$
$$+[I-\mathcal{X}(t+1)H(\theta)]F_{\theta i}(\theta)x(t)$$
$$+[I-\mathcal{X}(t+1)H(\theta)]F(\theta)x_{\theta i}(t)$$
$$+\mathcal{X}_{\theta i}(t+1)y(t+1), \qquad (4.5a)$$

$$\hat{x}_{\theta i}(0)=0, i=1,2,\cdots,q (I.C.). \qquad (4.5b)$$

In the above expression, the unknown sensitivity function $x_{\theta i}$ is involved and therefore in order to obtain this value, we differentiate (2.6a) with respect to θ_i to obtain

$$\mathcal{X}_{\theta i}(t+1)=\mathcal{X}_{\theta i}(t)-\mathcal{Y}_{\theta i}(t+1)[HF\mathcal{Y}(t)]^{\prime}$$
$$-\mathcal{Y}(t+1)[H_{\theta i}F\mathcal{Y}(t)+HF_{\theta i}\mathcal{Y}(t)+HF\mathcal{Y}_{\theta i}(t)]^{\prime}. \qquad (4.6)$$

In the meanwhile, the sensitivity function $\mathcal{Y}_{\theta i}$ can be calculated from (2.7b). The result is

$$\mathcal{Y}_{\theta i}(t+1)=[1-(HF\mathcal{Y}(t))^{\prime}(HF\mathcal{Y}(t))]^{-2}$$
$$\times\{(-\mathcal{X}_{\theta i}(t)H-\mathcal{X}(t)H_{\theta i})F\mathcal{Y}(t)$$
$$+(I-\mathcal{X}(t)H)F_{\theta i}\mathcal{Y}(t)$$
$$+(I-\mathcal{X}(t)H)F\mathcal{Y}_{\theta i}(t)\}\times\{1-(HF\mathcal{Y}(t))^{\prime}(HF\mathcal{Y}(t))\}$$
$$+(I-\mathcal{X}(t)H)F\mathcal{Y}(t)$$
$$\times\{(-H_{\theta i}F\mathcal{Y}(t)-HF_{\theta i}\mathcal{Y}(t)+HF\mathcal{Y}_{\theta i}(t))^{\prime}(HF\mathcal{Y}(t))$$
$$+(HF\mathcal{Y}(t))^{\prime}(H_{\theta i}F\mathcal{Y}(t)+HF_{\theta i}\mathcal{Y}(t)+HF\mathcal{Y}_{\theta i}(t)\}]. \qquad (4.7)$$

We note that the algorithm given by (4.5)-(4.7) in addition to the adaptive filtering equations corresponds to the linear innovations model[1] for the adaptive Kalman filtering technique. This algorithm is a little complicated and therefore an approximated model is proposed for neglecting the terms which do not affect extensively on the calculations of numerical values in the sensitivity equations. The approximated model is called the simplified model which neglects the sensitivity function of the Chandrasekhar $\mathcal{X}$-function with respect to the unknown parameter.

Equating $\mathcal{X}_{\theta}=0$ in (4.5a), we have

$$\hat{x}_{\theta i}(t+1)=-\mathcal{X}(t+1)H_{\theta i}F(\theta)\hat{x}(t)$$
$$+[I-\mathcal{X}(t+1)H(\theta)]F_{\theta i}\hat{x}(t)$$
$$+[I-\mathcal{X}(t+1)H(\theta)]F(\theta)\hat{x}_{\theta i}(t). \qquad (4.8)$$

It is noticed that this simplified model provides us with an algorithm that yields less computational time and storage than the innovation models.

5. Special Cases

5-1. Case 1($F=F(\theta)$)

Consider the case when the unknown parameters are involved in only the system matrix such that

$$F=F(\theta). \qquad (5.1)$$

For this case, the algorithm can be easily obtained by equating $H(\theta)=H$ and $H_{\theta}(\theta)=0$ in (2.5 a-b), (2.6), (2.7), (4.3), (4.5a), (4.5)-(4.7). The results for both the linear innovations model and the simplified model are obviously represented as follows. Namely, the linear innovations model is specifed by the following equations.

$$\hat{z}(t)=H\hat{x}(t)=\hat{y}(t), \qquad (5.2)$$

$$\hat{x}(t+1)=F(\hat{\theta}(t))\hat{x}(t)+\mathcal{X}(t+1)[y(t+1)-HF(\hat{\theta}(t))\hat{x}(t), \qquad (5.3a)$$
$$=(I-\mathcal{X}(t+1)H)F(\hat{\theta}(t))\hat{x}(t)+\mathcal{X}(t+1)y(t+1), \qquad (5.3b)$$
$$\hat{x}(0)=\mathcal{Y}(0)y(0), \qquad (5.3c)$$

$$\mathcal{X}(t+1)=\mathcal{X}(t)-\mathcal{Y}(t+1)[HF(\hat{\theta}(t))\mathcal{Y}(t)R]^{\prime}R^{-1}, \qquad (5.4)$$

$$\mathcal{Y}(t+1)=[I-\mathcal{X}(t)H]F(\hat{\theta}(t))\mathcal{Y}(t)$$
$$\times[I-R(HF(\hat{\theta}(t))\mathcal{Y}(t))^{\prime}R^{-1}(HF(\hat{\theta}(t)))\mathcal{Y}(t)]^{-1}, \qquad (5.5)$$

$$\Phi_i(t-1)=Hx_{\theta i}(t)\Big|_{\theta_i=\hat{\theta}_i(t-1)}, \qquad (5.6)$$

$$\hat{x}_{\theta i}(t+1)=-\mathcal{X}_{\theta i}(t+1)HF(\hat{\theta}(t))\hat{x}(t)$$
$$+[I-\mathcal{X}(t+1)H]F_{\theta i}(\hat{\theta}(t))\hat{x}(t)$$
$$+[I-\mathcal{X}(t+1)H]F(\hat{\theta}(t))\hat{x}_{\theta i}(t)$$
$$+\mathcal{X}_{\theta i}(t+1)y(t+1), \qquad (5.7)$$

$$\mathcal{X}_{\theta i}(t+1)=\mathcal{X}_{\theta i}(t)-\mathcal{Y}_{\theta i}(t+1)[HF\mathcal{Y}(t)]^{\prime}$$
$$-\mathcal{Y}(t+1)\{HF_{\theta i}\mathcal{Y}(t)+HF\mathcal{Y}_{\theta i}(t)]^{\prime}, \qquad (5.8)$$

$$\mathcal{Y}_{\theta i}(t+1)=[I-(HF\mathcal{Y}(t))^{\prime}(HF\mathcal{Y}(t))]^{-1}$$
$$\times\{-\mathcal{X}_{\theta i}(t)HF\mathcal{Y}(t)$$
$$+(I-\mathcal{X}(t)H)F_{\theta i}\mathcal{Y}(t)$$
$$+(I-\mathcal{X}(t)H)F\mathcal{Y}_{\theta i}(t)\}$$
$$\times\{1-(HF\mathcal{Y}(t))^{\prime}(HF\mathcal{Y}(t))\}$$
$$+(I-\mathcal{X}(t)H)F\mathcal{Y}(t)$$
$$+\{(-HF_{\theta i}\mathcal{Y}(t)+HF\mathcal{Y}_{\theta i}(t))^{\prime}(HF\mathcal{Y}(t))$$
$$+(HF\mathcal{Y}(t)^{\prime}(HF_{\theta i}\mathcal{Y}(t)+HF\mathcal{Y}_{\theta i}(t))\}]. \qquad (5.9)$$

However, the simplified model is given by the simple equations such that

$$\hat{x}_{\theta i}(t+1)=[I-\mathcal{X}(t+1)H]F_{\theta i}(\hat{\theta}(t))\hat{x}(t)$$
$$+[I-\mathcal{X}(t+1)H]F(\hat{\theta}(t))\hat{x}_{\theta i}(t), \qquad (5.10a)$$

$$\mathcal{X}_{\theta i}(t)=\mathcal{Y}_{\theta i}(t)=0. \qquad (5.10b)$$

5-2. Case 2($H=H(\theta)$)

The second case is that the unknown parameters are assumed to be involved in only the observation matrix. Namely,

$$H=H(\theta).$$

Then we also obtain the adaptive Chandraskhar filtering equation in equating $F(\theta)=F$ and $F_{\theta i}(\theta)=0$ in the corresponding equations. The results for both models mentioned in the preceeding sections can be easily obtained.The results are omitted for the reason of space

6. Conclusions

In this paper the adaptive Chandrasekhar filter based on the SSM was developed for the adaptive estimation of the states and parameters in linear discrete time-invariant dynamical systems. Two types of models were presented to calculate the sensitivity functions with respect to the unknown parameters. One was the conventional linear innovations model and the other was the simplified model which has less computational time and storage than the innovations model. It was explained that the adaptive Chandrasekhar filtering technique is superior to the adaptive Kalman filtering technique from a computational point of view. Also, the adaptive Chandrasekhar filter is stable in the numerical calculations.

References

[1] G. C. Goodwin and K. S. Sin: Adaptive Filtering Prediction and Control, Prentice Hall, New Jersey, Chapt.9, pp.366(1984)

[2] M. Sugisaka and M. Nabeshima: A Filter Using Chandrasekhar's X and Y Functions for Linear Stationary Discrete-Time Systems, Preprints of 17th SICE Symp. on Control Theory, Tokyo, June 9-11, pp.205-208(1988)

[3] M. Sugisaka, A. Tominaga, and M. Nabeshima: Digital Chandrasekhar Filter and Smoother Based on State Space Model, Preprints of 18th JAACE Symp., Stochastic Systems Theory and Its Application, Tokyo, Oct., 29-31, pp.35-38(1986)

[4] M. Sugisaka and S. Abe: Adaptive Digital Chandrasekhar Filter Based on the State-Space Model, Report of the Faculty of Eng., oita Univ., Vol.16, pp.29-33(1987)

[5] M. Sugisaka: Digital Adaptive Chandrasekhar-Type Filters Based on Covariance and State Space Models, 8th IFAC/IFORS Symp. on Identification and System Parameter Estimation, August 27-31, Beijing, China(1988)

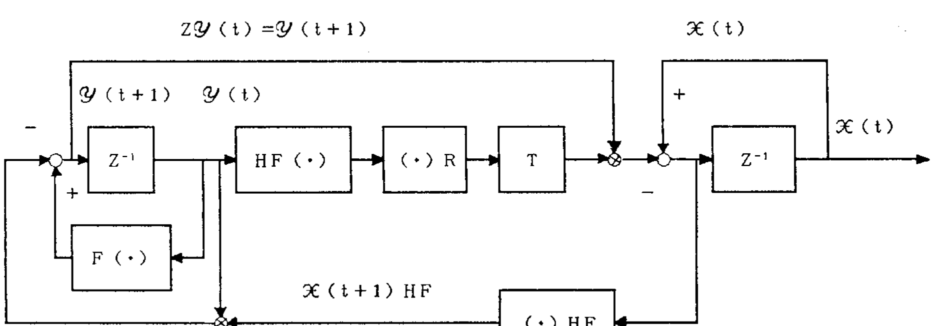

Fig. 1 The blockdiagram representation for a pair of difference equations of Chandrasekhar 𝒳 and 𝒴 functions ((2.6a) and (2.7a))