Pole-Assignment Fixed-Interval Kalman Smoother and Wiener Smoother[1)]

SUN Shu-Li[1]　DENG Zi-Li[2]
[1](Department of Automation, Heilongjiang University, Harbin 150080)
[2](Institute of Applied Mathematics, Heilongjiang University, Harbin 150080)
(E-mail: sunsl@hlju.edu.cn)

Abstract Based on steady-state Kalman filter and white noise estimators, and according to the pole-assignment principle of the control theory, the pole-assignment fixed-interval steady-state Kalman smoother and Wiener smoother are presented. They avoid computation of the initial optimal smoothing estimates and can rapidly eliminate the effects of the initial smoothing estimates by assigning the poles of the smoothers, so that they have a practical stability in the finite fixed interval. A simulation example shows their effectiveness.

Key words Fixed-interval Kalman smoother, fixed-interval Wiener smoother, pole-assignment, Kalman filtering method

1 Introduction

Fixed-interval smoothers can be applied to reconstruction for orbits of man-made earth satellite and missile. Several fixed-interval Kalman smoothers have been presented in [1～4], their disadvantages are to require the computation of optimal initial values, which increases computational complexity. In this paper, based on steady-state Kalman filter, the pole-assignment principle of the control theory is applied to the state estimation for stochastic systems, and the pole-assignment fixed-interval steady-state Kalman smoother and Wiener smoother are presented for the completely observable and completely controllable systems. They avoid computation of optimal smoothing initial values and can rapidly eliminate the effects of the initial smoothing estimates by assigning the poles of the smoothers, so that they have a practical stability in the finite interval.

Consider the discrete time-invariant linear stochastic system

$$x(t+1) = \Phi x(t) + \Gamma w(t) \tag{1}$$

$$y(t) = Hx(t) + v(t) \tag{2}$$

where the state $x(t) \in R^n$, the measurement $y(t) \in R^m$, Φ, Γ, H are the constant matrices with compatible dimensions.

Assumption 1. $w(t) \in R^r$ and $v(t) \in R^m$ are independent white noises with zero mean and variance matrices $Q>0$ and $R>0$, respectively.

Assumption 2. The initial state $x(0)$ with mean $\mathrm{E}x(0)=\mu$ is independent of $w(t)$ and $v(t)$.

Assumption 3. The system is completely observable and completely controllable.

The problem is to find the steady-state fixed-interval smoothers $\hat{x}(t\mid N)$ based on measurements $(y(N), y(N-1), \cdots, y(0))$, where N is fixed, $t=0,1,\cdots,N$.

2 Lemmas

Lemma 1[1]. For system by (1) and (2) with Assumptions 1～3, the steady-state Kal-

1) Supported by National Natural Science Foundation of P. R. China (69774019), and by Natural Science Foundation of Heilongjiang Province (F01-15)
Received June 18, 2002; in revised form April 22, 2003
收稿日期　2002-06-18；收修改稿件日期　2003-04-22

man predictor is given by

$$\hat{x}(t+1 \mid t) = \Phi \hat{x}(t \mid t-1) + K_P \varepsilon(t), \quad \varepsilon(t) = y(t) - H\hat{x}(t \mid t-1) \tag{3}$$

$$K = \Sigma H^{\mathrm{T}}[H\Sigma H^{\mathrm{T}} + R]^{-1}, \quad K_P = \Phi K \tag{4}$$

where the steady-state prediction error variance matrix Σ is the unique positive definite solution of the following steady-state Riccati equation

$$\Sigma = \Phi[\Sigma - \Sigma H^{\mathrm{T}}(H\Sigma H^{\mathrm{T}} + R)^{-1} H\Sigma]\Phi^{\mathrm{T}} + \Gamma Q \Gamma^{\mathrm{T}} \tag{5}$$

Lemma 2[1]. For system by (1) and (2) with Assumptions 1~3, the unified steady-state white noise estimators are given by

$$\hat{\theta}(t \mid t+N) = \hat{\theta}(t \mid t+N-1) + M_\theta(N)\varepsilon(t+N), \quad \theta = w, v \tag{6}$$

where $\hat{\theta}(t|t+N)=\mathbf{0}$, $N<0$, and $M_w(0)=0$, $M_v(0)=R[H\Sigma H^{\mathrm{T}}+R]^{-1}$,

$$M_\theta(N) = D_\theta[(I_n - KH)^{\mathrm{T}}\Phi^{\mathrm{T}}]^{N-1} H^{\mathrm{T}}[H\Sigma H^{\mathrm{T}} + R]^{-1}, \quad N \geqslant 1 \tag{7}$$

$$D_w = Q\Gamma^{\mathrm{T}}, \quad D_v = -RK^{\mathrm{T}}\Phi^{\mathrm{T}} \tag{8}$$

and the innovation $\varepsilon(t)$ can be recursively computed by (3) with arbitrary initial value $\hat{x}(0|-1)$.

Unified white noise Wiener smoothers are given by

$$\hat{\theta}(t \mid t+N) = \psi_p^{-1}(q^{-1})M_N^\theta(q^{-1})A(q^{-1})y(t+N), \quad \theta = w, v \tag{9}$$

where $\Psi_P=(\Phi-K_P H)$ is a state matrix[1], *i. e.*, its eigenvalues all lie inside the unit circle. q^{-1} is the backward shift operator, and we define $\psi_p(q^{-1})=\det(I_n-\Psi_P q^{-1})$, $A(q^{-1})=\psi_p(q^{-1})I_m - H\mathrm{adj}(I_n-q^{-1}\Psi_P)K_P q^{-1}$, and

$$M_N^\theta(q^{-1}) = \sum_{i=0}^{N} M_\theta(i)q^{i-N} \tag{10}$$

Lemma 3[1]. For system by (1) and (2) with Assumptions 1~3, the steady-state fixed-point Kalman smoother is given by

$$\hat{x}(t \mid t+j) = \hat{x}(t \mid t+j-1) + K_j\varepsilon(t+j) \tag{11}$$

with the initial value $\hat{x}(t|t-1)$, $j=0,1,2,\cdots$, and $K_j=\Sigma[(I_n-KH)^{\mathrm{T}}\Phi^{\mathrm{T}}]^j H^{\mathrm{T}}[H\Sigma H^{\mathrm{T}}+R]^{-1}$.

3 Pole-assignment fixed-interval steady-state Kalman smoother

Theorem 1. For system by (1) and (2) with Assumptions 1~3, the pole-assignment fixed-interval steady-state Kalman smoother is given by

$$\hat{x}(t \mid N) = \Psi\hat{x}(t-1 \mid N) + \Gamma\hat{w}(t-1 \mid N) - T_0 y(t-1) + T_0\hat{v}(t-1 \mid N) \tag{12}$$

where $\Psi=\Phi+T_0 H$, and we can choose matrix T_0 to assign the eigenvalues of Ψ arbitrarily, and make Ψ stable. In the finite fixed-interval $[0,N]$, it has a practical stability in the following sense: case 1) All eigenvalues of Ψ_p are near origin. For the arbitrary smoothing initial estimate $\hat{x}(0|N)$, a matrix T_0 is chosen such that all eigenvalues of Ψ are sufficiently small, then the effects of the initial values $\hat{x}(0|N)$ and $\hat{x}(0|-1)$ of smoother (12) will rapidly decay to zero. Case 2) There exists an eigenvalue of Ψ_p near the unite circle. If $\hat{x}(0|-1)$ is taken as the optimal initial value $\hat{x}(0|-1)=\boldsymbol{\mu}$, and all eigenvalues of Ψ are assigned sufficiently small, then the effect of the initial value $\hat{x}(0|N)$ will rapidly decay to zero.

Proof. Taking the projection of each term on both sides of (1) yields the steady-state fixed-interval Kalman smoother

$$\hat{x}(t \mid N) = \Phi\hat{x}(t-1 \mid N) + \Gamma\hat{w}(t-1 \mid N) \tag{13}$$

and premultiplying (2) by an $n\times m$ matrix T_0 and taking the projection operation yields $T_0 y(t-1)=T_0 H\hat{x}(t-1|N)+T_0\hat{v}(t-1|N)$, which plus (13) yields (12). From Assumption 3, (Φ, H) is a completely observable pair. Hence, by choosing T_0 such that the eigenvalues of Ψ are assigned arbitrarily specified λ_i, $|\lambda_i|<1$[1], Ψ is a stable matrix. From (6) we know that the computation of white noise estimator $\hat{w}(t|N)$ is dependent on the innovation process $\varepsilon(t)$, while the computation of the innovation $\varepsilon(j)$ is dependent on the initial values $\hat{\boldsymbol{x}}(0|-1)$, so that the computation of smoother (12) is dependent on both initial values $\hat{\boldsymbol{x}}(0|-1)$ and $\hat{\boldsymbol{x}}(0|N)$. Arbitrarily taking two sets of initial values of (12) as $(\hat{\boldsymbol{x}}^{(i)}(0|N), \hat{\boldsymbol{x}}^{(i)}(0|-1))$, $i=1,2$, from (3), (6) and the initial values $\hat{\boldsymbol{x}}^{(i)}(0|-1)$, we have the corresponding Kalman predictors $\hat{\boldsymbol{x}}^{(i)}(t|t-1)$, innovations $\boldsymbol{\varepsilon}^{(i)}(t)$, white noise estimators $\hat{\boldsymbol{w}}^{(i)}(t-1|N)$ and $\hat{\boldsymbol{v}}^{(i)}(t-1|N)$, $i=1,2$, and from (12) we have the corresponding two fixed-interval Kalman smoothers $\hat{\boldsymbol{x}}^{(i)}(t|N)$. From (11) we have the non-recursively steady-state fixed-interval smoother

$$\hat{\boldsymbol{x}}(t \mid N) = \hat{\boldsymbol{x}}(t \mid t-1) + \sum_{j=0}^{N-t} K_j \boldsymbol{\varepsilon}(t+j) \tag{14}$$

For the above initial value $\hat{\boldsymbol{x}}^{(i)}(0|-1)$, $i=1,2$, using Lemma 1 and (14) yields two corresponding non-recursive smoothers $\hat{\boldsymbol{x}}_o^{(i)}(t \mid N)$, and from the asymptotic stability of the steady-state Kalman predictor $\hat{\boldsymbol{x}}(t|t-1)$[1], with time t increasing, we have

$$\boldsymbol{\delta}_o(t) = [\hat{\boldsymbol{x}}_o^{(1)}(t \mid N) - \hat{\boldsymbol{x}}_o^{(2)}(t \mid N)] \rightarrow 0 \tag{15}$$

and for case 1), we have that $\boldsymbol{\delta}_o(t)$ rapidly converges to zero. If the smoothing initial value of (12) is taken as $\hat{\boldsymbol{x}}_o^{(i)}(0 \ N)$, $i=1,2$, then applying the uniqueness of projection yields the results computed by (12) to be numerically identical to $\hat{\boldsymbol{x}}_o^{(i)}(t|N)$, so that $\hat{\boldsymbol{x}}_o^{(i)}(t|N)$ also satisfies (12).

Setting $\boldsymbol{\delta}^{(i)}(t)=\hat{\boldsymbol{x}}^{(i)}(t|N)-\hat{\boldsymbol{x}}_o^{(i)}(t|N)$, from (12) we have the difference equations

$$\boldsymbol{\delta}^{(i)}(t) = \boldsymbol{\Psi}\boldsymbol{\delta}^{(i)}(t-1), \quad i=1,2 \tag{16}$$

where $\boldsymbol{\delta}^{(i)}(t)=[\delta_1^{(i)}(t),\cdots,\delta_n^{(i)}(t)]^{\mathrm{T}}$. Assuming that $\boldsymbol{\Psi}$ is assigned to have n different real eigen values $\lambda_i, i=1,2,\cdots,n$, $|\lambda_i|<1$, and the corresponding n linearly independent eigenvectors are $\boldsymbol{\alpha}_i=[\alpha_{1i},\alpha_{2i},\cdots,\alpha_{ni}]^{\mathrm{T}}, i=1,2,\cdots,n$, the difference equation (16) has the general solutions [5]

$$\boldsymbol{\delta}^{(i)}(t) = c_1^{(i)}\boldsymbol{\alpha}_1\lambda_1^{\mathrm{T}} + c_2^{(i)}\boldsymbol{\alpha}_2\lambda_2^{\mathrm{T}} + \cdots + c_n^{(i)}\boldsymbol{\alpha}_n\lambda_n^{\mathrm{T}}, \quad i=1,2 \tag{17}$$

where the coefficients $c_j^{(i)}$ are determined by the initial values $\boldsymbol{\delta}^{(i)}(0)$. We rewrite (17) in the component form

$$\delta_j^{(i)}(t) = c_1^{(i)}\alpha_{j1}\lambda_1^{\mathrm{T}} + c_2^{(i)}\alpha_{j2}\lambda_2^{\mathrm{T}} + \cdots + c_n^{(i)}\alpha_{jn}\lambda_n^{\mathrm{T}}, \quad j=1,2,\cdots,n,\ i=1,2 \tag{18}$$

Letting $\lambda_m=\max(|\lambda_1|,|\lambda_2|,\cdots,|\lambda_n|), c_\lambda=\max(|c_k^{(i)}\alpha_{jk}\ |j,k=1,\cdots,n;i=1,2)$, we have

$$|\delta_j^{(i)}(t)| \leqslant nc_\lambda\lambda_m^{\mathrm{T}}, \quad j=1,2,\cdots,n,\ i=1,2 \tag{19}$$

Taking sufficiently small λ_i with $|\lambda_i|<1$, such that λ_m is sufficiently small, in finite interval $[0,N]$, with time t increasing, $\boldsymbol{\delta}^{(i)}(t)\rightarrow\boldsymbol{0}$ rapidly. Noting that

$$\boldsymbol{\delta}(t) = \hat{\boldsymbol{x}}^{(1)}(t \mid N) - \hat{\boldsymbol{x}}^{(2)}(t \mid N) = \boldsymbol{\delta}^{(1)}(t) - \boldsymbol{\delta}^{(2)}(t) + \boldsymbol{\delta}_o(t) \tag{20}$$

we also have $\boldsymbol{\delta}(t)\rightarrow 0$ rapidly, i.e., the effects of the initial values decay rapidly.

For case 2), $\boldsymbol{\delta}_o(t)$ will converge to zero slowly, which will delay the convergent process of $\boldsymbol{\delta}(t)$. To eliminate the bad effect of $\boldsymbol{\Psi}_p$ to the convergent rate, we should set the initial value $\hat{\boldsymbol{x}}(0|-1)$ of (12) as the optimal initial value $\hat{\boldsymbol{x}}_o(0|-1)=\boldsymbol{\mu}$. In fact, on the one hand, based on the optimal initial values $\hat{\boldsymbol{x}}_o(0|-1)$ and (3) and (14), we can obtain the steady-state optimal smoother $\hat{\boldsymbol{x}}_o(t|N)$. According to the uniqueness of projection, if (12) with the initial value $\hat{\boldsymbol{x}}(0|N)=\hat{\boldsymbol{x}}_o(0|N)$ is computed, $\hat{\boldsymbol{x}}_o(t|N)$ also satisfies (12). On the other hand, based on the optimal initial values $\hat{\boldsymbol{x}}_o(0|-1)$ and arbitrary initial value $\hat{\boldsymbol{x}}(0|N)$, using (12) yields the corresponding smoother $\hat{\boldsymbol{x}}(t|N)$. Since $\hat{\boldsymbol{x}}(t|N)$ and $\hat{\boldsymbol{x}}_o(t|N)$ both satisfy (12) with the same initial value $\hat{\boldsymbol{x}}_o(0|-1)$, setting $\boldsymbol{\delta}(t)=\hat{\boldsymbol{x}}(t|N)-\hat{\boldsymbol{x}}_o(t|N)$, we have that $\boldsymbol{\delta}(t)=\boldsymbol{\Psi}\boldsymbol{\delta}(t-1)$. Hence we can assign the eigenvalues of $\boldsymbol{\Psi}$ sufficiently near origin, so $\boldsymbol{\delta}(t)$ converges to zero rapidly with time t increasing, which means that $\hat{\boldsymbol{x}}(t|N)$ can approximate to the optimal smoother $\hat{\boldsymbol{x}}_o(t|N)$ rapidly. The proof is completed. □

Remark 1. Given the initial value $\hat{x}(0|-1)$, since the second and fourth terms on the right side of (12) can be computed with the initial value $\hat{x}(0|-1)$, they are known. Hence the poles of the smoother (12) are determined by $\det(I_n - q^{-1}\Psi)=0$ or eigenvalues of Ψ. Therefore the smoother (12) with eigenvalues of Ψ is called pole-assignment smoother.

4 Pole-assignment fixed-interval Wiener smoother

Theorem 2. For system by (1) and (2) with Assumptions 1～3, the pole-assignment fixed-interval Wiener smoother is given by

$$\psi_p(q^{-1})\psi(q^{-1})\,\hat{x}(t \mid N) = K_{N,t}(q^{-1})y(N) \qquad (21)$$

where q^{-1} only operates on the time t of $\hat{x}(t|N)$, $\psi_p(q^{-1})$ is defined by lemma 2, Ψ is defined by Theorem 1, and

$$\psi(q^{-1}) = \det(I_n - q^{-1}\Psi),\quad K_{N,t}(q^{-1}) = \mathrm{adj}(I_n - q^{-1}\Psi)T_{N,t}(q^{-1}),$$

$$T_{N,t}(q^{-1}) = \Gamma M^{w}_{N+1-t}(q^{-1})A(q^{-1}) - \psi_p(q^{-1})T_0 q^{t-1-N} + T_0 M^{v}_{N+1-t}(q^{-1})A(q^{-1}) \qquad (22)$$

where $A(q^{-1})$ is defined by Lemma 2, T_0 is defined by Theorem 1. If all eigenvalues of Ψ_p are near origin and we choose T_0 to assign sufficiently small eigenvalues of Ψ, then the Wiener smoother (21) can rapidly eliminate the effect of arbitrary smoothing initial values $(\hat{x}(0|N),\cdots,\hat{x}(2n-1|N))$, and has a practical stability.

Proof. (12) can be expressed in the transfer function form as

$$\hat{x}(t \mid N) = (I_n - q^{-1}\Psi)^{-1}[\Gamma\hat{w}(t-1 \mid N) - T_0 y(t-1) + T_0\hat{v}(t-1 \mid N)] \qquad (23)$$

Substituting (9) and $(I_n - q^{-1}\Psi)^{-1} = \mathrm{adj}(I_n - q^{-1}\Psi)/\psi(q^{-1})$ into (23) yields (21) and (22). Similar to the proof of the case 1) in Theorem 1, the practical stability of (21) can be proved. The proof is completed. □

Remark 2. $K_{N,t}(q^{-1})$ in the fixed-interval Wiener smoother (21) is a polynomial matrix with time-varying degree.

5 Simulation results

Consider system by (1) and (2) with $\Phi=\begin{bmatrix}1 & 0\\ 2 & 0\end{bmatrix}$, $\Gamma=\begin{bmatrix}1\\ 1\end{bmatrix}$, $H=[1]$, $Q=0.1$, $R=0.1$, $x(t)=[x_1(t),x_2(t)]^{\mathrm{T}}$, and the initial value $x(0)=[0,0]^{\mathrm{T}}$, where $w(t)$ and $v(t)$ are independent Gaussian white noises. Setting $N=200$, the problem is to find the pole-assignment fixed-interval steady-state Kalman smoother $\hat{x}(t|200)$.

We can easily verify that this system is completely observable and completely controllable, and the transition matrix Φ is an unstable matrix. Hence from Theorem 1 we have the pole-assignment fixed-interval Kalman smoother (12). We can obtain the eigenvalues $\lambda_1=-0.1459$, $\lambda_2=0$ of $\Psi_p=\Phi-K_pH$, which are near origin, so the initial value $\hat{x}(0|-1)$ can be chosen arbitrarily. Case 1): choose $T_0=[-0.24 \quad -0.46]^{\mathrm{T}}$ to assign the eigenvalues of $\Psi=\Phi+T_0H$ as $\lambda_1=0.1$, $\lambda_2=0.2$. Setting the smoothing initial value $\hat{x}(0|200)=[10,50]^{\mathrm{T}}$, the simulation results are shown in Fig. 1 and Fig. 2, where the solid lines denote the true values, the dashed lines denote the estimation values. We see that the smoother $\hat{x}(t|200)$ can rapidly approximate to the true value $x(t)$ by assigning the eigenvalues of Ψ near origin. Case 2): choose $T_0=[-0.0067,\ 0.7067]^{\mathrm{T}}$ to assign the eigenvalues of Ψ as $\lambda_1=0.8$, $\lambda_2=0.9$. Again Still setting the smoothing initial value $\hat{x}(0|200)=[10,50]^{\mathrm{T}}$, the simulation results are shown by the dash dot lines in Fig. 1 and Fig. 2. We see that the smoother $\hat{x}(t|200)$ will have a transition process gradually to approximate to the true value $x(t)$, if the assigned eigenvalues of Ψ are near the unite circle. This shows that the smoother defined by (12) can rapidly eliminate the effect of the initial smoothing estimate $\hat{x}(0|200)$ as long as the assigned eigenvalues of Ψ are sufficiently small.

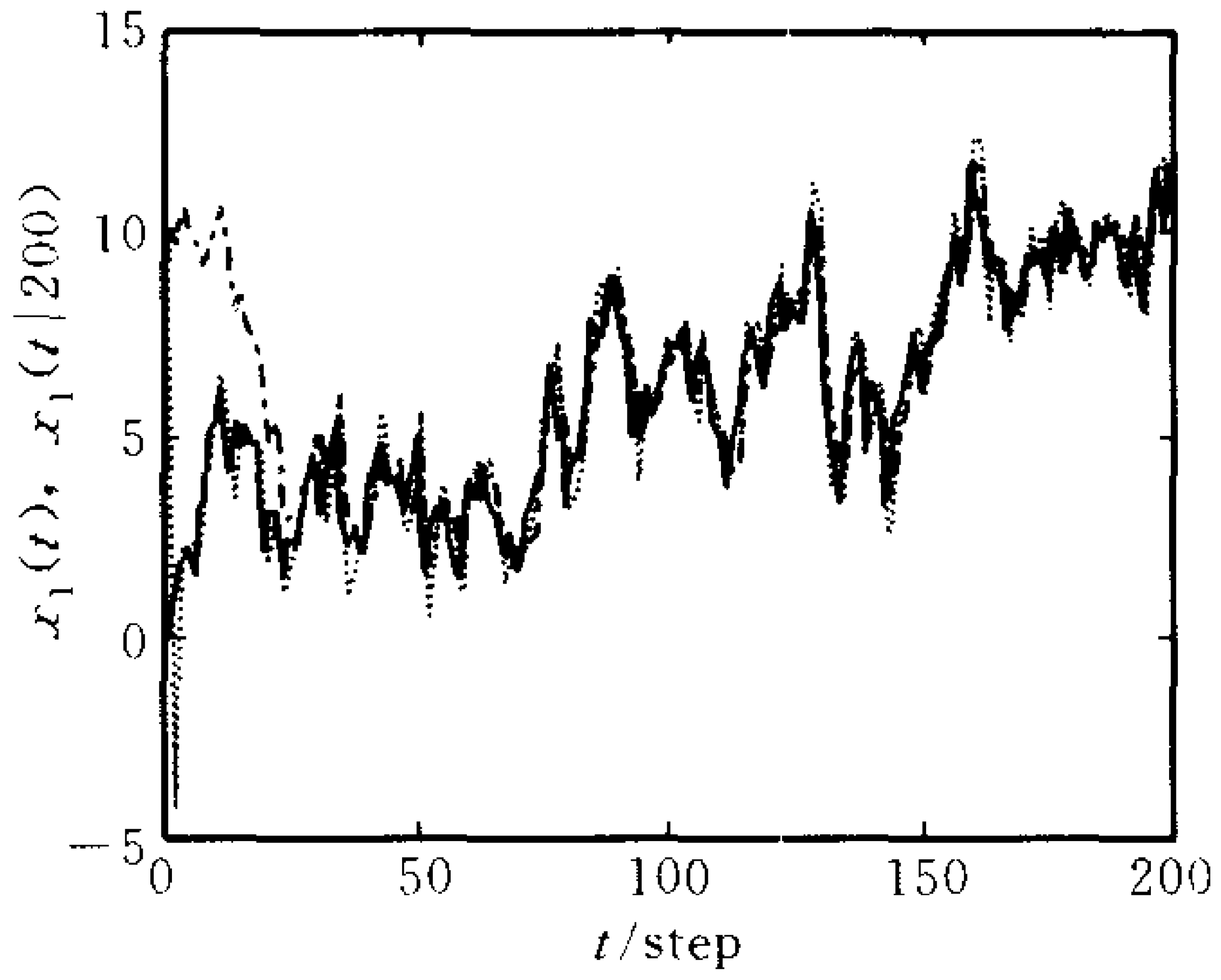

Fig. 1 $x_1(t)$ and pole-assignment fixed-interval Kalman smoother $\hat{x}_1(t|200)$

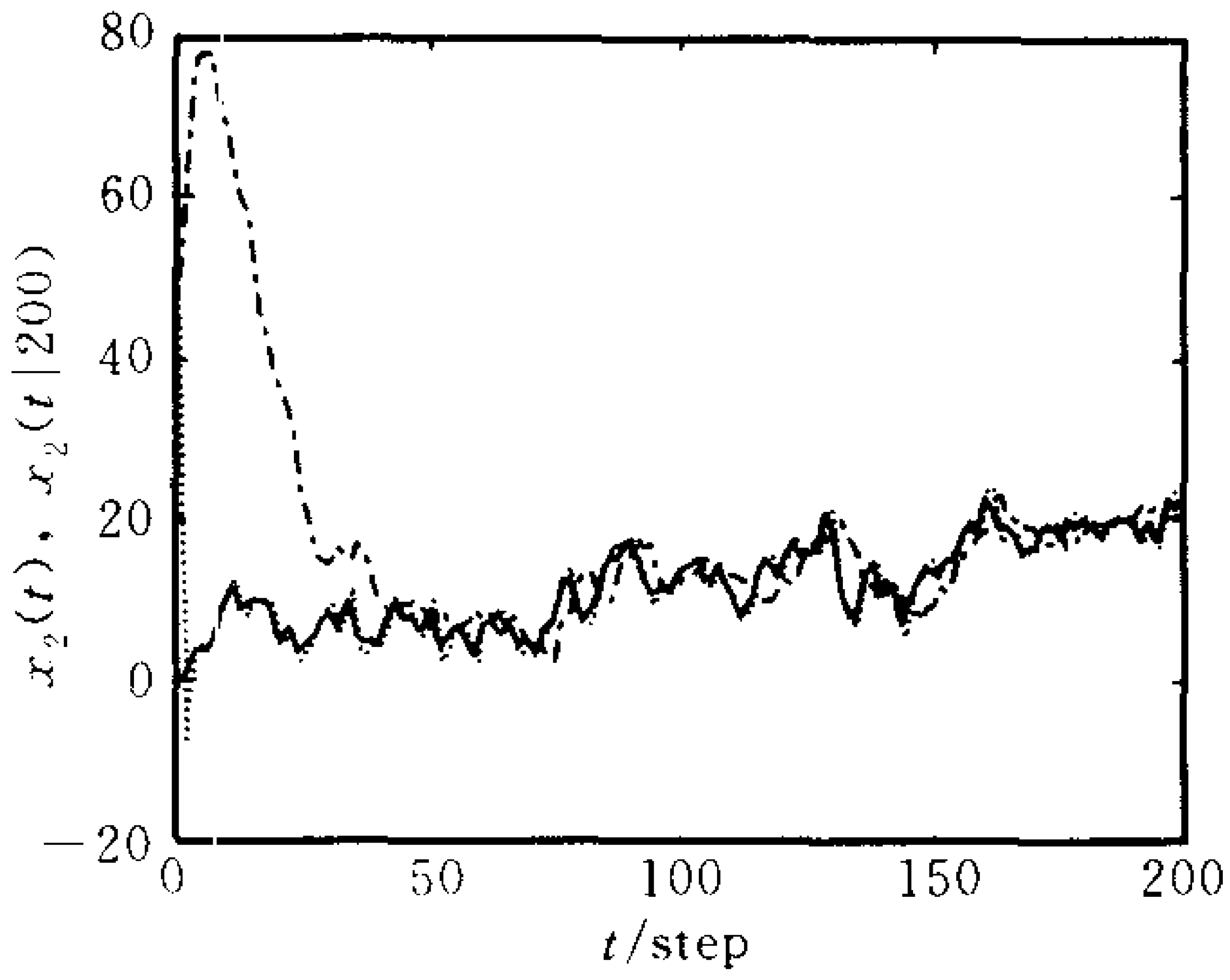

Fig. 2 $x_2(t)$ and pole-assignment fixed-interval Kalman smoother $\hat{x}_2(t|200)$

References

1 Deng Z L. Kalman Filtering and Wiener Filtering— Modern Time Series Analysis Method. Harbin: Harbin Institute of Technology Press, 2001. 1～396 (in Chinese)

2 Mendel J M. Lessons in Estimation Theory for Signal Processing, Communications, and Control. Englewood Cliffs, New Jersey: Prentice-Hall, 1995

3 Deng Z L, Liu Y M. New fixed-point and fixed-interval Kalman smoothers. *Acta Automatica Sinica*, 1999, **25**(1): 32～37(in Chinese)

4 Deng Z L, Wang L H. New fixed-interval Kalman smoothering algorithms. *Control Theory and Applications*, 2000, **19**(5):777～780 (in Chinese)

5 Wang Y. Discrete Control Systems. Beijing: Science Press, 1987 (in Chinese)

SUN Shu-Li Received his master degree from Heilongjiang University in 1999. He is a Ph. D. candidate in Harbin Institute of Technology. His research interests include Kalman filtering and Wiener filtering.

DENG Zi-Li Professor. His research interests include state estimation and optimal filtering.

极点配置固定区间 Kalman 平滑器和 Wiener 平滑器

孙书利[1] 邓自立[2]

[1](黑龙江大学自动化系 哈尔滨 150080)

[2](黑龙江大学应用数学研究所 哈尔滨 150080)

(E-mail: sunsl@hlju.edu.cn)

摘 要 基于稳态 Kalman 滤波器和白噪声估值器，根据控制理论中的极点配置原理，提出了极点配置固定区间稳态 Kalman 平滑器和 Wiener 平滑器. 它们避免了计算最优平滑初值，且通过配置平滑器的极点，可快速消除初始平滑估值的影响，因而它们具有在有限固定区间上的实用稳定性，仿真例子说明了它们的有效性.

关键词 固定区间 Kalman 平滑器，固定区间 Wiener 平滑器，极点配置，Kalman 滤波方法

中图分类号 O211.64