Design of recursive fixed-point smoother using covariance information in linear discrete-time systems

SEIICHI NAKAMORI*

(Received September 19, 1990)

This paper designs the recursive fixed-point smoother and filter, in linear discrete-time systems, which use the observed value and the covariance information of signal and observation noise for the white Gaussian and white Gaussian plus coloured observation noises.

1. Introduction

In the previous research, the recursive fixed-point smoother, which is different from the Kalman filter in approach, is reported in linear continuous systems (Nakamori 1990). This smoother does not use the state-space model of a signal but the covariance information of signal and coloured noise in addition to the observed value. Therefore, this novel estimation technique is applicable not only to the signal processes, which can be estimated by the Kalman filter, but also to those, for which the realization by the state-space model is impossible. It is considered that the discrete-time estimator is appropriate for the digital processing by a digital computer.

This paper, at first, presents the recursive least-squares fixed-point smoother and filter using the covariance information for the white Gaussian observation noise in linear discrete-time systems. Secondly, the relation between the filter using the covariance information and the Kalman filter is clarified by introducing the innovations theory (Kailath 1968, Sage and Melsa 1971). The innovations theory leads to the expression of the discrete-time Kalman gain using the covariance information. The recursive algorithm for the impulse response function is also obtained for the white Gaussian observation noise. Finally, the recursive fixed-point smoother and filter are proposed in the linear discrete-time systems for the white Gaussian plus coloured observation noise when the signal and the coloured noise are uncorrelated.

In this paper, it is assumed that the autocovariance functions of the signal and the coloured noise are expressed by the semi-degenerate kernel forms. Since the semi-degenerate kernel can express the autocovariance functions of nonstationary or stationary stochastic processes, the presented estimators are suitable for estimating those general signal processes.

Received September 19, 1990

*Department of Technology, Faculty of Education, Kagoshima University, 1-20-6, Kohrimoto, Kagoshima 890, Japan

2. Discrete-time fixed-point smoothing problems

The observation of the signal is assumed to occur at discrete points in time in accordance with the linear relationship

$$y(k)=z(k)+v(k), \tag{1}$$

where $y(k)$ represents an $n\times 1$ observation vector, $z(k)$ is a zero-mean signal vector and $v(k)$ is a zero-mean white Gaussian observation noise vector. It is assumed that $z(k)$ is uncorrelated with $v(s)$ as

$$\mathrm{E}[z(k)v^{\mathrm{T}}(s)]=0. \tag{2}$$

The autocovariance function of $v(k)$ is given by

$$\mathrm{E}[v(k)v^{\mathrm{T}}(s)]=\mathrm{R}\delta(k-s), \tag{3}$$

where it is assumed that R is a positive definite matrix. The autocovariance function of $z(k)$ is expressed by using the semi-degenerate kernel form as

$$\mathrm{K}(k,s)=\begin{cases}\mathrm{A}(k)\mathrm{B}^{\mathrm{T}}(s), & 0\leqq s\leqq k\\ \mathrm{B}(k)\mathrm{A}^{\mathrm{T}}(s), & 0\leqq k\leqq s.\end{cases} \tag{4}$$

Here, $\mathrm{A}(k)$ and $\mathrm{B}(s)$ are bounded $n\times \mathrm{M}'$ matrix functions.

Let the discrete-time fixed-point smoothing estimate $\hat{z}(k,\mathrm{L})$ of the signal $z(k)$ be given by

$$\hat{z}(k,\mathrm{L})=\sum_{i=1}^{\mathrm{L}}h(k,i,\mathrm{L})y(i) \tag{5}$$

as a linear transformation of the observation data set $\{y(i), 1\leqq i\leqq \mathrm{L}\}$, where $h(k,i,\mathrm{L})$ is called the impulse response function. It can be shown that minimizing the mean square value

$$\mathrm{J}=\mathrm{E}[(z(k)-\hat{z}(k,\mathrm{L}))^{\mathrm{T}}(z(k)-\hat{z}(k,\mathrm{L}))] \tag{6}$$

of the smoothing error $z(k)-\hat{z}(k,\mathrm{L})$ leads to the Wiener-Hopf equation

$$\mathrm{E}[z(k)y^{\mathrm{T}}(s)]=\sum_{i=1}^{\mathrm{L}}h(k,i,\mathrm{L})\mathrm{E}[y(i)y^{\mathrm{T}}(s)] \tag{7}$$

from the orthogonal projection lemma

$$z(k)-\sum_{i=1}^{\mathrm{L}}h(k,i,\mathrm{L})y(i)\perp y(s),\ 0\leqq s,\ k\leqq \mathrm{L}, \tag{8}$$

(Sage and Melsa 1971). By using (1)−(3), one obtains

$$h(k,s,\mathrm{L})\mathrm{R}=\mathrm{K}(k,s)-\sum_{i=1}^{\mathrm{L}}h(k,i,\mathrm{L})\mathrm{K}(i,s) \tag{9}$$

from (8). (9) is the basic equation for the optimal impulse response function $h(k,s,\mathrm{L})$ in the linear discrete-time fixed-point smoothing problems.

In section 3, the Cauchy system for the fixed-point smoothing estimate is presented by starting with (9).

3. Design of fixed-point smoother for white Gaussian observation noise

[Theorem 1] presents the recursive fixed-point smoothing algorithm for the white Gaussian observation noise.

[Theorem 1]

If the autocovariance function of the signal is expressed by (4) in the semi-degenerate kernel form, if the signal $z(\cdot)$ and the white Gaussian observation noise $v(\cdot)$ are uncorrelated, and if the variance of the white Gaussian observation noise $v(k)$ is R, then the discrete-time fixed-point smoothing estimate $\hat{z}(k, L)$ is calculated by (10)—(16) recursively.

$$\hat{z}(k,L)=\hat{z}(k,L-1)+h(k,L,L)(y(L)-A(L)O(L-1)) \tag{10}$$

$$h(k,L,L)=(B(L)A^T(L)-P(k,L-1)A^T(L))(R+(B(L)-A(L)r(L-1))A^T(L))^{-1} \tag{11}$$

$$P(k,L)=P(k,L-1)+h(k,L,L)(B(L)-A(L)r(L-1)) \tag{12}$$

$$P(L,L)=A(L)r(L) \tag{13}$$

$$J(L,L)=(B^T(L)-r(L-1)A^T(L))(R+(B(L)-A(L)r(L-1))A^T(L))^{-1} \tag{14}$$

$$r(L)=r(L-1)+J(L,L)(B(L)-A(L)r(L-1)) \tag{15}$$

$$O(L)=O(L-1)+J(L,L)(y(L)-A(L)O(L-1)) \tag{16}$$

Here, $h(k,L,L)$, $P(k,L)$, $r(L)$ and $J(L,L)$ are $n\times n$, $n\times M'$, $M'\times M'$ and $M'\times n$ matrices respectively.

The filtering estimate $\hat{z}(L,L)$ is given by

$$\hat{z}(L,L)=A(L)O(L), \tag{17}$$

where $O(L)$ is an $M'\times 1$ vector.

The initial conditions on the difference equations (15) and (16) for $r(L)$ and $O(L)$ at $L=0$ are as follows.

$$\left.\begin{array}{l} r(0)=0 \\ O(0)=0 \end{array}\right\} \tag{18}$$

The proof of [Theorem 1] is deferred to the appendix.

4. Innovations theory and filtering error covariance

Let us introduce the innovations theory in relation to the estimation problems using the covariance information for the white Gaussian observation noise. The innovations process is defined by

$$\nu(k)=y(k)-\hat{z}(k,k-1) \tag{19}$$

(Kailath 1968), where $\hat{z}(k,k-1)$ is the least-squares estimate of the signal $z(k)$, at the discrete-time k, based on the observation data set $\{y(i),\ 1\leqq i\leqq k-1\}$. The filtering estimate $\hat{z}(k,k)$ is expressed by

$$\hat{z}(k,k)=\sum_{i=1}^{k}g(k,i)\nu(i) \tag{20}$$

as a linear transformation of the innovations process $\nu(i)$, $i=1, 2, \cdots, k$, by an impulse response function $g(k,i)$. If one rewrites (20), one obtains

$$\hat{z}(k,k)=\sum_{i=1}^{k-1}g(k,i)\nu(i)+g(k,k)(y(k)-\hat{z}(k,k-1)) \tag{21}$$

from (19). Let $W(k)$ denote the Kalman gain, then $g(k,k)$ is seen to be equal to $W(k)$ from (21). The Kalman gain $W(k)$ is formulated as

$$W(k)=A(k)J(k,k)$$
$$=(A(k)B^T(k)-A(k)r(k)A^T(k))R^{-1} \quad (22)$$

from (16), (17) and (A12). From Sage and Melsa (1971), it is found that the variance function $S(k,k)$ of the filtering error $\tilde{z}(k,k)(=z(k)-\hat{z}(k,k))$ is obtained as

$$S(k,k)=var(\tilde{z}(k,k))$$
$$=W(k)R$$
$$=A(k)B^T(k)-A(k)r(k)A^T(k) \quad (23)$$

by the covariance information. The filtering error variance of $z(k)$ is related to the covariance of $\tilde{z}(k,k-1)$, e. g.

$$var(\tilde{z}(k,k))=var(\tilde{z}(k,k-1))(I+R^{-1}var(\tilde{z}(k,k-1)))^{-1} \quad (24)$$

(Sage and Melsa 1971). If one uses an expression of $W(k)=var(\tilde{z}(k,k))R^{-1}$ and introduces

$$var(\nu(k))=var(\tilde{z}(k,k-1))+R, \quad (25)$$

which denotes the variance of the innovations process $\nu(k)$, one obtains the autocovariance function $\Omega(k)$ of the innovations process as

$$\Omega(k)=var(\nu(k))$$
$$=((I-W(k))^{-1}W(k)+I)R. \quad (26)$$

By the way, the filtering estimate $\hat{z}(L,L)$ is written as

$$\hat{z}(L,L)=A(L)O(L-1)+A(L)J(L,L)(y(L)-A(L)O(L-1)) \quad (27)$$

by (16) and (17). Comparing the discrete-time Kalman filter (Sage and Melsa 1971) with (21) and (27), one obtains

$$\hat{z}(L,L-1)=A(L)O(L-1)$$
$$=\sum_{i=1}^{L-1}g(L,i)\nu(i). \quad (28)$$

Therefore, (27) is written as

$$\hat{z}(L,L)=\hat{z}(L,L-1)+W(L)(y(L)-\hat{z}(L,L-1))$$
$$=\hat{z}(L,L-1)+W(L)\nu(L) \quad (29)$$

in terms of the innovations process $\nu(L)$. Thus, the innovations state-space model becomes

$$\hat{z}(L+1,L)=\Phi(L+1,L)\hat{z}(L,L-1)+\Phi(L+1,L)W(L)\nu(L), \quad (30)$$

where $\Phi(L+1,L)$ is the state-transition matrix of the state-space model for $z(\cdot)$. One finds that $\Phi(L+1,L)$ is given by

$$\Phi(L+1,L)=A(L+1)A^{-1}(L) \quad (31)$$

on the condition that $A(L)$ is a positive definite matrix of order n for $n=M'$. It is clear that the observation equation is given by

$$y(L)=\hat{z}(L,L-1)+\nu(L), \quad (32)$$

for the innovations state-space model of (30).

It is an advantage that this innovations model can be determined only by the covariance information. It is also interesting that the whitening filter is constructed from the innovations model, where this whitening filter yields the white Gaussian output innovations process from the observation process.

5. Optimal impulse response function $h(k,s,k)$ and $h(k,s,\mathrm{L})$

Let us present algorithms for the optimal impulse response functions $h(k,s,k)$ and $h(k,s,\mathrm{L})$ in [Theorem 2].

[Theorem 2]

Let the covariance information of the signal $z(\cdot)$ and the white Gaussian observation noise $v(\cdot)$ be given, then the impulse response function $h(k,s,k)$, which is useful in computing the filtering estimate $\hat{z}(k,k)$ by (5) for $\mathrm{L}=k$, is calculated by (14), (15), (33) and (34).

$$h(k,s,k)=\mathrm{A}(k)\mathrm{J}(s,k),\ 0\leqq s\leqq k \tag{33}$$

$$\mathrm{J}(s,\mathrm{L})=\mathrm{J}(s,\mathrm{L}-1)-\mathrm{J}(\mathrm{L},\mathrm{L})\mathrm{A}(\mathrm{L})\mathrm{J}(s,\mathrm{L}-1). \tag{34}$$

The initial condition on the partial differential equation (34) at $\mathrm{L}-1=s$ is $\mathrm{J}(s,s)$. $\mathrm{J}(s,s)$ is calculated by (14) and (15).

Also the optimal impulse response function $h(k,s,\mathrm{L})$, which yields the fixed-point smoothing estimate of (5) is calculated by (11), (14), (15), (34) and (35).

$$h(k,s,\mathrm{L})=h(k,s,\mathrm{L}-1)-h(k,\mathrm{L},\mathrm{L})\mathrm{A}(\mathrm{L})\mathrm{J}(s,\mathrm{L}-1) \tag{35}$$

(Proof)

(33) is clear from (A22). (34) is derived by substituting (A7) into (A9).

(35) is derived by (A.4) and (A.7) (Q. E. D.).

6. Fixed-point smoother for white Gaussian plus colored observation noise

Let an n dimensional observation equation be expressed by

$$y(k)=z(k)+v(k)+v_c(k), \tag{36}$$

for the white Gaussian plus coloured observation noise, where $v_c(k)$ is a zero-mean coloured observation noise, and $y(k)$, $z(k)$ and $v(k)$ are the same quantities introduced in (1) for the white Gaussian observation noise. It is assumed that $z(\cdot)$, $v(\cdot)$ and $v_c(\cdot)$ are uncorrelated mutually.

The autocovariance function of the coloured observation noise $v_c(k)$ is also expressed by the semi-degenerate kernel form

$$\mathrm{K}_c(k,s)=\begin{cases}\alpha(k)\beta^{\mathrm{T}}(s),\ 0\leqq s\leqq k,\\ \beta(k)\alpha^{\mathrm{T}}(s),\ 0\leqq k\leqq s,\end{cases} \tag{37}$$

where $\alpha(\cdot)$ and $\beta(\cdot)$ are $n\times \mathrm{N}'$ bounded matrices.

The fixed-point smoothing estimate $\hat{z}(k,\mathrm{L})$ is expressed by (5) also for the white Gaussian plus coloured observation noise. An equation, which corresponds to (9) for the white Gaussian observation noise, becomes

$$h(k,s,\mathrm{L})\mathrm{R}=\mathrm{K}(k,s)-\sum_{i=1}^{\mathrm{L}}h(k,i,\mathrm{L})(\mathrm{K}(i,s)+\mathrm{K}_c(i,s)) \tag{38}$$

for the white Gaussian plus coloured observation noise.

The fixed-point smoothing algorithm for the white Gaussian plus colored observation noise is summarized in [Theorem 3].

[Theorem 3]

Let the autocovariance functions of the signal $z(k)$ and the coloured observation noise $v_c(k)$ be expressed by (4) and (37) respectively in the semi-degenerate kernel forms, and let R denote the variance of the white Gaussian observation noise, then the fixed-point smoothing estimate $\hat{z}(k,L)$ is calculated by (39)−(48) recursively.

$$\hat{z}(k,L)=\hat{z}(k,L-1)+h(k,L,L)(y(L)-A(L)O(L-1)-\alpha(L)Q(L-1)) \tag{39}$$

$$h(k,L,L)=(B(L)A^T(L)-E(k,L-1)\alpha^T(L)-F(k,L-1)A^T(L))(R+(B(L)-A(L)r(L-1)-\alpha(L)C(L-1))A^T(L)+(\beta(L)-A(L)G(L-1)-A(L)D(L-1)))\alpha^T(L))^{-1} \tag{40}$$

$$O(L-1)=O(L-2)+J(L-1,L-1)(y(L-1)-A(L-1)O(L-2)-\alpha(L-1)Q(L-2)) \tag{41}$$

$$Q(L-1)=Q(L-2)+I(L-1,L-1)(y(L-1)-A(L-1)O(L-2)-\alpha(L-1)Q(L-2)) \tag{42}$$

$$r(L-1)=r(L-2)+J(L-1,L-1)(B(L-1)-A(L-1)r(L-2)-\alpha(L-1)C(L-2)) \tag{43}$$

$$G(L-1)=G(L-2)+J(L-1,L-1)(B(L-1)-A(L-1)G(L-2)-\alpha(L-1)D(L-2)) \tag{44}$$

$$C(L-1)=C(L-2)+I(L-1,L-1)(B(L-1)-A(L-1)r(L-2)-\alpha(L-1)C(L-2)) \tag{45}$$

$$D(L-1)=D(L-2)+I(L-1,L-1)(B(L-1)-A(L-1)G(L-2)-\alpha(L-1)D(L-2)) \tag{46}$$

$$J(L-1,L-1)=(B^T(L-1)-r(L-2)A^T(L-1)-G(L-2)\alpha^T(L-1))(R+(B(L-1)-A(L-1)r(L-2)-\alpha(L-1)C(L-2))A^T(L-1)+(\beta(L-1)-A(L-1)G(L-2)-\alpha(L-1)D(L-2))\alpha^T(L-1))^{-1} \tag{47}$$

$$I(L-1,L-1)=(B^T(L-1)-C(L-2)A^T(L-1)-D(L-2)\alpha^T(L-1))(R+(B(L-1)-A(L-1)r(L-2)-\alpha(L-1)C(L-2))A^T(L-1)+(\beta(L-1)-A(L-1)G(L-2)-\alpha(L-1)D(L-2))\alpha^T(L-1))^{-1} \tag{48}$$

The filtering estimate $\hat{z}(L,L)$ is given by

$$\hat{z}(L,L)=A(L)O(L). \tag{49}$$

$$E(k,L)=E(k,L-1)+h(k,L,L)(B(L)-A(L)r(L-1)-\alpha(L)C(L-1)) \tag{50}$$

$$E(L,L)=A(L)r(L) \tag{51}$$

$$F(k,L)=F(k,L-1)+h(k,L,L)(B(L)-A(L)G(L-1)-\alpha(L)D(L-1)) \tag{52}$$

$$F(L,L)=A(L)G(L) \tag{53}$$

Here, $O(L-1)$ and $Q(L-1)$ are $M'\times 1$ and $N'\times 1$ vectors. The functions $r(L-1)$, $G(L-1)$, $C(L-1)$, $D(L-1)$, $J(L-1,L-1)$, $I(L-1,L-1)$, $E(k,L)$, $F(k,L)$ and $h(k,L,L)$ are $M'\times M'$, $M'\times N'$, $N'\times M'$, $N'\times N'$, $M'\times n$, $N'\times n$, $n\times M'$, $n\times N'$ and $n\times n$ matrices respectively.

The initial conditions on the difference equations (41), (42), (43), (44), (45) and (46) are as follows.

$$\left.\begin{aligned}&O(0)=0\\&Q(0)=0\\&r(0)=0\\&G(0)=0\\&C(0)=0\\&D(0)=0\end{aligned}\right\} \quad (54)$$

[Theorem 3] is derived by using an invariant imbedding method as in the proof of [Theorem 1], then its proof is omitted.

7. A stabilty consideration

Let us consider a stability problem. The smoothing error covariance function is defined by

$$U(k,s,L)=E[(z(k)-\hat{z}(k,L))(z(s)-\hat{z}(s,L))^T],\ 0\leqq s,\ k\leqq L. \quad (55)$$

From an orthogonal projection lemma that the smoothing error $z(k)-\hat{z}(k,L)$ is orthogonal to the smoothing estimate $\hat{z}(s,L)$ for $0\leqq s,\ k\leqq L$, one obtains

$$\begin{aligned}U(k,s,L)&=K(k,s)-E[\hat{z}(k,L)z^T(s)]\\&=K(k,s)-E[\hat{z}(k,L)\hat{z}^T(s,L)]\\&=K(k,s)-U_z(k,s,L),\end{aligned} \quad (56)$$

where $U_z(k,s,L)$ denotes an autocovariance function of the fixed-point smoothing estimate $\hat{z}(k,L)$. $U_z(k,s,L)$ and $U(k,s,L)$ are positive semi-definite matrices. Therefore, one has a relationship

$$0\leqq U_z(k,s,L)\leqq K(k,s). \quad (57)$$

According to a discussion on stability problems (Kailath 1976), one notices that (57) ensures that the presented smoothing algorithm has a unique solution, since $U_z(k,s,L)$ is lower and upper bounded.

It can be shown that the smoothing error covariance function $U_z(k,k,L)$ is given by

$$U_z(k,k,L)=h(k,L,L)R \quad (58)$$

after some manipulations. $h(k,L,L)$ is calculated by (40), (43)−(48) and (50)−(53).

8. A numerical example

Let us show a numerical example of [Theorem 3] for the white Gaussian plus coloured observation noise. The observation equation is given by (36). The scalar observation process is considered here.

The autocovariance functions of the signal $z(k)$ and coloured noise $v_c(k)$ are given by

$$K(k,s)=1.026\cdot 0.95^{(k-s)} \quad (59)$$

and

$$K_c(k,s)=0.1\cdot 0.5^{(k-s)} \quad (60)$$

for $0\leqq s\leqq k$ respectively. The functions, which constitute the autocovariance functions of the signal and the coloured observation noise, are as follows.

$$A(k)=1.026\cdot 0.95^k,\ B(s)=1/0.95^s,\ \alpha(k)=0.1\cdot 0.5^k,\ \beta(s)=1/0.5^s. \quad (61)$$

Substituting (61) into the smoothing algorithm of [Theorem 3], one can calculate the fixed-point

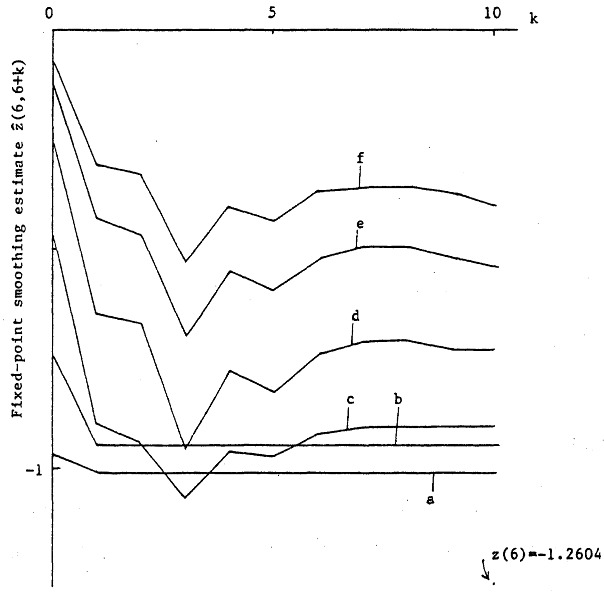

Fig. 1 Fixed-point smoothing estimate $\hat{z}(6,6+k)$ vs. k.

a ...... $\hat{z}(6,6+k)$ for the white Gaussian observation noise $N(0,0.1^2)$.

b ...... $\hat{z}(6,6+k)$ for the white Gaussian observation noise $N(0,0.3^2)$.

c ...... $\hat{z}(6,6+k)$ for the white Gaussian observation noise $N(0,1)$.

d ...... $\hat{z}(6,6+k)$ for the white Gaussian observation noise $N(0,3^2)$.

e ...... $\hat{z}(6,6+k)$ for the white Gaussian observation noise $N(0,5^2)$.

f ...... $\hat{z}(6,6+k)$ for the white Gaussian observation noise $N(0,7^2)$.

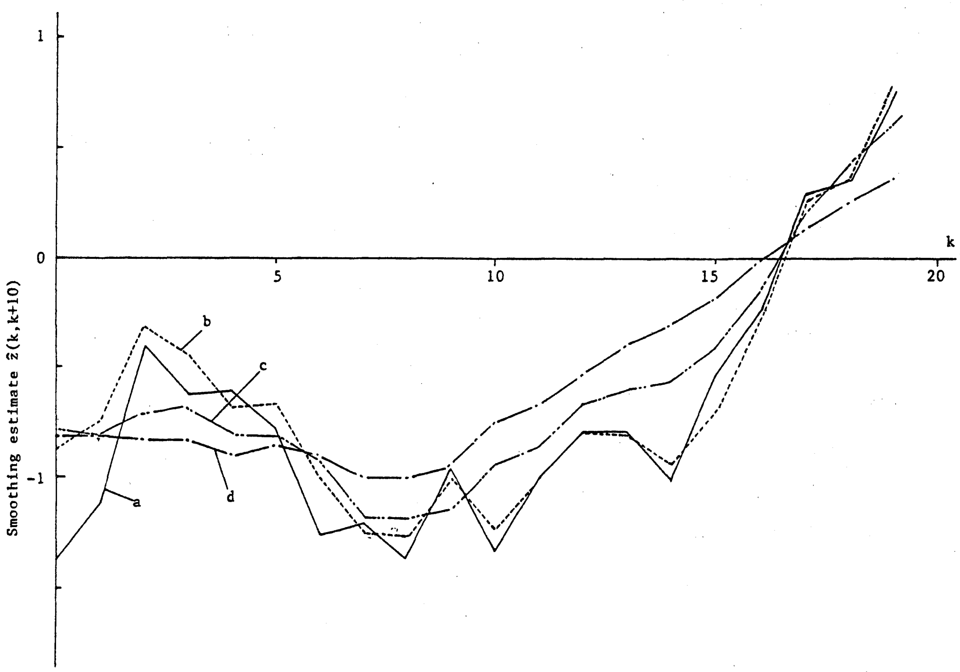

Fig. 2 Fixed-point smoothing estimate $\hat{z}(k,k+10)$ vs. k.

a ...... Signal process.

b ...... $\hat{z}(k,k+10)$ for the white Gaussian observation noise $N(0,0.1^2)$.

c ...... $\hat{z}(k,k+10)$ for the white Gaussian observation noise $N(0,0.5^2)$.

d ...... $\hat{z}(k,k+10)$ for the white Gaussian observation noise $N(0,1)$.

smoothing estimate. Fig. 1 illustrates $\hat{z}(6,6+k)$ vs. k for the white Gaussian observation noises $N(0,0.1^2)$, $N(0,0.3^2)$, $N(0,1)$, $N(0,3^2)$, $N(0,5^2)$ and $N(0,7^2)$. In Fig. 1, the fixed-point is 6. In Fig. 2, the fixed-lag smoothing estimate $\hat{z}(k,k+10)$ is plotted by repetitive uses of the proposed fixed-point smoothing algorithm for the white Gaussian observation noises $N(0,0.1^2)$, $N(0,0.5^2)$ and $N(0,1)$ in the interval of $0 \leqq k \leqq 19$. In Fig. 2, the fixed-point corresponds to k. Here, it is needless to say that the fixed-lag smoothing estimate is calculated by a fixed-lag smoother usually.

For reference, the state-space realization for the signal $z(k)$ from the covariance information (59) is expressed by

$$z(k+1)=0.95z(k)+u(k),\ E[u(k)u(s)]=0.1\delta(k-s). \quad (62)$$

Also, the state-space model for the coloured noise $v_c(k)$ is

$$v_c(k+1)=0.5v_c(k)+v_1(k),\ E[v_1(k)v_1(s)]=0.075\delta(k-s). \quad (63)$$

9. Conclusions

The numerical example has shown that the linear discrete-time fixed-point smoother presented in this paper is feasible.

The presented filter in this paper can be classified as the recursive Wiener filter (Kailath 1974) in linear discrete-time systems.

References

CASTI, J.L., KALABA, R. and MURTHY, V.K., 12, IEEE Trans. Information Theory, IT-18, 515.

KAILATH, T. 1968, IEEE Trans. Autom. Control, AC-13, 645.

KAILATH, T. 1974, IEEE Trans. Information Theory, IT-20, 146.

KAILATH, T. 1976, Lectures on linear least-squares estimation. (Berlin: Springer-Verlag).

NAKAMORI, S. and SUGISAKA, M. 1977, Automatica, 13, 623.

NAKAMORI, S. 1990, Int. J. Systems Sci., 21, 527.

SAGE, A.P. and MELSA, J.L. 1971, Estimation theory with applications to communications and control (New York: McGraw-Hill).

Appendix (proof of [Theorem 1])

Putting $L \rightarrow L-1$ in (9), one obtains

$$h(k,s,L-1)R=K(k,s)-\sum_{i=1}^{L-1}h(k,i,L-1)K(i,s). \quad (A1)$$

Subtracting (A1) from (9) yields

$$(h(k,s,L)-h(k,s,L-1))R=-h(k,L,L)K(k,s)-\sum_{i=1}^{L-1}(h(k,i,L)-h(k,i,L-1))K(i,s). \quad (A2)$$

Introducing the equation

$$\Lambda(s,L-1)R=K(L,s)-\sum_{i=1}^{L-1}\Lambda(i,L-1)K(i,s), \quad (A3)$$

which the function $\Lambda(s,L-1)$ satisfies, and using (A2) and (A3), one obtains

$$h(k,s,L)=h(k,s,L-1)-h(k,L,L)\Lambda(s,L-1) \quad (A4)$$

after some manipulations. If one takes into account of the property of the semi-degenerate kernel of (4), (A3) is rewritten as

$$\Lambda(s,L-1)R=A(L)B^T(s)-\sum_{i=1}^{L-1}\Lambda(i,L-1)K(i,s). \tag{A5}$$

Similarly, from (A5) and

$$J(s,L-1)R=B^T(s)-\sum_{i=1}^{L-1}J(i,L-1)K(i,s), \tag{A6}$$

$\Lambda(s,L-1)$ is given by

$$\Lambda(s,L-1)=A(L)J(s,L-1). \tag{A7}$$

Putting $L\to L-1$ in (A6) and subtracting the obtained equation from (A6), one has

$$(J(s,L-1)-J(s,L-2))R=-J(L-1,L-1)K(L-1,s)-\sum_{i=1}^{L-2}(J(i,L-1)-J(i,L-2))K(i,s). \tag{A8}$$

Thus the difference equation for $J(s,L-1)$ becomes

$$J(s,L)=J(s,L-1)-J(L,L)\Lambda(s,L-1) \tag{A9}$$

by (A3) and (A8). Putting $s=L-1$ in (A6), one obtains

$$J(L-1,L-1)R=B^T(L-1)-\sum_{i=1}^{L-1}J(i,L-1)B(i)A^T(L-1) \tag{A10}$$

with the property of the semi-degenerate kernel of (4). Introducing

$$r(L-1)=\sum_{i=1}^{L-1}J(i,L-1)B(i), \tag{A11}$$

one can rewrite (A10) as

$$J(L-1,L-1)R=B^T(L-1)-r(L-1)A^T(L-1). \tag{A12}$$

Putting $L\to L-1$ in (A11) and subtracting the obtained equation from (A11), one has

$$r(L-1)-r(L-2)=J(L-1,L-1)B(L-1)+\sum_{i=1}^{L-2}(J(i,L-1)-J(i,L-2))B(i). \tag{A13}$$

Substituting (A9) into (A13) and using (A7) and (A11), one obtains (15). $J(L,L)$ of (14) is derived by using (A12) and (15). The initial condition on the difference equation (15) for $r(L)$ at $L=0$ is $r(0)=0$ from (A11). If one puts $s=L$ in (9), the result is

$$h(k,L,L)R=K(k,L)-\sum_{i=1}^{L}h(k,i,L)K(i,L). \tag{A14}$$

Using the property of the semi-degenerate kernel of (4), one can rewrite (A14) as

$$h(k,L,L)R=(B(k)-\sum_{i=1}^{L}h(k,i,L)B(i))A^T(L). \tag{A15}$$

Let us introduce the function

$$P(k,L)=\sum_{i=1}^{L}h(k,i,L)B(i). \tag{A16}$$

Then (A15) is transformed into

$$h(k,L,L)R=(B(k)-P(k,L))A^T(L). \tag{A17}$$

Putting $L\to L-1$ in (A16) and subtracting the obtained equation from (A16), one has

$$P(k,L)-P(k,L-1)=h(k,L,L)B(L)+\sum_{i=1}^{L-1}(h(k,i,L)-h(k,i,L-1))B(i). \tag{A18}$$

(A18) is rewitten as

$$P(k,L)=P(k,L-1)+h(k,L,L)(B(L)-A(L)r(L-1)) \tag{A19}$$

by (A4), (A7) and (A11). The initial condition on $P(k,L)$ at $k=L$ is given by

$$P(L,L)=\sum_{i=1}^{L}h(L,i,L)B(i) \tag{A20}$$

from (A16).

It is shown that $h(L,i,L)$ satisfies

$$h(L,i,L)R = A(L)B^T(i) - \sum_{n=1}^{L} h(L,n,L)K(n,i) \quad \text{(A21)}$$

from (4) and (9). One can express $h(L,i,L)$ as

$$h(L,i,L) = A(L)J(i,L) \quad \text{(A22)}$$

by comparing (A21) with (A6). Substituting (A22) into (A20) and using (A11), one obtains

$$P(L,L) = A(L)r(L). \quad \text{(A23)}$$

The filtering estimate $\hat{z}(L,L)$ is given by

$$\hat{z}(L,L) = \sum_{i=1}^{L} h(L,i,L)y(i) \quad \text{(A24)}$$

from (5). Substituting (A22) into (A24) and introducing

$$O(L) = \sum_{i=1}^{L} J(i,L)y(i), \quad \text{(A25)}$$

one obtains (17) for the filtering estimate. Putting $L \to L-1$ in (A25) and subtracting the obtained equation from (A25), one has

$$O(L) - O(L-1) = J(L,L)y(L) + \sum_{i=1}^{L-1} (J(i,L) - J(i,L-1))y(i). \quad \text{(A26)}$$

Then (16) is derived by substituting (A9) into (A26) and by using (A7) and (A25). The initial condition on the difference equation for $O(L)$ at $L=0$ is $O(0)=0$ from (A25).

Finally, the difference equation for the fixed-point smoothing estimate is obtained as follows. Putting $L \to L-1$ in (5) and subtracting the obtained equation from (5), one has

$$\hat{z}(k,L) - \hat{z}(k,L-1) = h(k,L,L)y(L) + \sum_{i=1}^{L-1} (h(k,i,L) - h(k,i,L-1))y(i). \quad \text{(A27)}$$

Then (10) is derived from (A4), (A7), (A25) and (A27) (Q. E. D.).