Optimal fixed-point smoother using covariance information for white Gaussian plus colored observation noise

Seiichi NAKAMORI

(Received October 1, 1988)

Department of Technology, Faculty of Education, Kagoshima University

The optimal fixed-point smoother is derived for the white Gaussian plus colored observation noise when the covariance information for the signal, the colored and white Gaussian observation noises, and the observed value are available.

1. Introduction

In the estimation problems of the state of a dynamic model, the Kalman filter is well-known (Kalman, 1960). The Wiener-Hopf integral equation is often used to derive a different type of estimator, which uses the covariance information of the signal and white Gaussian observation noise with the observed value (Nakamori and Sugisaka, 1977). The other approach for an estimator using covariance information is studied (Kailath, 1976), where the estimator is derived in the relation between the Kalman filter and the covariance information. There seems to be many cases when the white Gaussian and colored observation noises are corrupted with the signal during the transmission of the signal (Trees, 1968). Also, this kind of estimator is appropriate for estimation problems of stochastic signal obtained from environmental circumstances.

This paper derives the optimal sequential fixed-point smoother from the Wiener-Hopf integral equation for the white Gaussian plus colored observation noise. It is assumed that the autocovariance functions of the signal and the colored observation noise are expressed in the semi-degenerate kernel form. The sequential algorithm for an optimal impulse response function, which yields a linear least-squares smoothing estimate, is also presented, and uniqueness of the presented fixed-point smoothing algorithm is proved.

2. Problem formulation

Consider the following observation equation

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

where $z(t)$ is an n-dimensional zero-mean signal vector, $v_c(t)$ is a zero-mean colored observation noise, and $v(t)$ is a zero-mean white Gaussian observation noise with variance R.

$$E[v(t)v^T(s)] = R\delta(t-s),\ E[v(t)] = O \tag{2}$$

It is assumed that the signal $z(t)$, the white Gaussian observation noise $v(t)$ and the colored observation noise $v_c(t)$ are uncorrelated mutually.

$$E[z(t)v^T(s)] = 0,\ E[z(t)v_c^T(s)] = 0,\ E[v_c(t)v^T(s)] = 0,\ 0 \leqq s,\ t \leqq T \tag{3}$$

Let us assume that the fixed-point smoothing estimate $\hat{z}(t, T)$ of $z(t)$, at the fixed-point t, based on the measurement data $\{y(s),\ 0 \leqq s \leqq T\}$ is given by

$$\hat{z}(t, T) = \int_0^T h(t, s, T)\, y(s)\, ds, \tag{4}$$

where $h(t, s, T)$ denotes an n×n impulse response function. Minimizing the mean-square value of smoothing error $z(t) - \hat{z}(t, T)$

$$J = E[(z(t) - \hat{z}(t, T))^T (z(t) - \hat{z}(t, T))], \tag{5}$$

one obtains the Wiener-Hopf integral equation (Sage and Melsa, 1971)

$$E[z(t)y^T(s)] = \int_0^T h(t, s', T)\, E[y(s')y^T(s)]\, ds'. \tag{6}$$

Let $K(t, s)$ and $K_c(t, s)$ denote the autocovariance functions of $z(t)$ and $v_c(t)$ as

$$K(t, s) = E[z(t)z^T(s)],\ K_c(t, s) = E[v_c(t)v_c^T(s)]. \tag{7}$$

Substituting (1) into (6), and using (2) and (3), one obtains

$$h(t, s, T) R = K(t, s) - \int_0^T h(t, s', T)(K(s', s) + K_c(s', s)) ds' \tag{8}$$

(Trees, 1968).

If one applies an invariant imbedding method (Nakamori and Sugisaka, 1977), which transforms Volterra type integral equation of the second kind into a Cauchy system, to (8), one can derive an algorithm for the optimal impulse response function $h(t, s, T)$. Here, it is assumed that the autocovariance functions of the signal $z(t)$ and the colored observation noise $v_c(t)$ are expressed in semi-degenerate kernel forms (Nakamori and Sugisaka, 1977) as $K(t, s) = A(t) B^T(s)$ and $K_c(t, s) = C(t) D^T(s)$ for $0 \leqq s \leqq t$ respectively. Here, $A(t)$ and $B(s)$ are n×m bounded matrices, and $C(t)$ and $D(s)$ are n×m' bounded matrices.

3. Derivation of optimal fixed-point smoother

If one differentiates (8) with respect to T, one obtains

$$\partial h(t, s, T)/\partial T R = -h(t, T, T)(K(T, s) + K_c(T, s)) - \int_0^T \partial h(t, s', T)/\partial T (K(s', s) + K_c(s', s)) ds'. \tag{9}$$

Introducing a function $\Phi(s, T)$, which satisfies

$$\Phi(s, T) R = -(K(T, s) + K_c(T, s)) - \int_0^T \Phi(s', T)(K(s', s) + K_c(s', s)) ds', \tag{10}$$

one obtains a partial differential equation

$$\partial h(t, s, T)/\partial T = h(t, T, T)\Phi(s, T),\ h(t, s, O) = K(t, s) R^{-1}. \tag{11}$$

Let us introduce functions $J(s, T)$ and $L(s, T)$ which satisfy

$$J(s, T) R = -B^T(s) - \int_0^T J(s', T)(K(s', s) + K_c(s', s)) ds' \tag{12}$$

and

$$L(s,\ T)R=-D^{T}(s)-\int_{0}^{T}L(s',T)\ (K(s',s)+K_{c}(s',s))\ ds'. \tag{13}$$

From (10), (12) and (13), one notices that the function $\Phi(s,\ T)$ is expressed as

$$\Phi(s,\ T)=A(T)J(s,\ T)+C(T)L(s,\ T) \tag{14}$$

by referring to the semi-degenerate kernel forms of $K(t,\ s)$ and $K_{c}(t,\ s)$ for $0\leqq s\leqq t$. Differentiating (12) with respect to T, one obtains

$$\partial J(s,\ T)/\partial TR=-J(T,\ T)\ (K(T,\ s)+K_{c}(T,\ s))-\int_{0}^{T}\partial J(s',T)/\partial T\ (K(s',s)+K_{c}(s',s))\ ds'. \tag{15}$$

A partial differential equation for the function $J(s,\ T)$ becomes

$$\partial J(s,\ T)/\partial T=J(T,\ T)\Phi(s,\ T),\ J(s,\ O)=-B^{T}(s)R^{-1}, \tag{16}$$

from (10) and (15). If one differentiates (13) with respect to T, one obtains

$$\partial L(s,\ T)/\partial TR=-L(T,\ T)\ (K(T,\ s)+K_{c}(T,\ s))-\int_{0}^{T}\partial L(s',T)/\partial T\ (K(s',s)+K_{c}(s',s))\ ds'. \tag{17}$$

From (10) and (17), a partial differential equation for $L(s,\ T)$ is written as

$$\partial L(s,\ T)/\partial T=L(T,\ T)\Phi(s,\ T),\ L(s,\ O)=-D^{T}(s)R^{-1}. \tag{18}$$

The function $J(T,\ T)$ which appeared in (16) is developed as

$$J(T,\ T)R=-B^{T}(T)-\int_{0}^{T}J(s',T)\ (K(s',T)+K_{c}(s',T))\ ds'$$

$$=-B^{T}(T)-\int_{0}^{T}J(s',T)B(s')A^{T}(T)ds'-\int_{0}^{T}J(s',T)D(s')C^{T}(T)ds'.$$

$$= -B^T(T) - r(T)A^T(T) - f(T)C^T(T) \tag{19}$$

by introducing functions $r(T)$ and $f(T)$, which are expressed by (20) and (21).

$$r(T) = \int_0^T J(s', T)B(s')ds' \tag{20}$$

$$f(T) = \int_0^T J(s', T)D(s')ds' \tag{21}$$

If one differentiates (20) with respect to T, and uses (14) and (16), one can develop an ordinary differential equation for $r(T)$.

$$dr(T)/dT = J(T, T)B(T) + \int_0^T \partial J(s', T)/\partial T B(s')ds'$$

$$= J(T, T)B(T) + \int_0^T J(T, T)\Phi(s', T)B(s')ds'$$

$$= J(T, T)B(T) + J(T, T)\int_0^T (A(T)J(s', T) + C(T)L(s', T))B(s')ds' \tag{22}$$

If one introduces a function $g(T)$ expressed by

$$g(T) = \int_0^T L(s', T)B(s')ds', \tag{23}$$

and if one uses (20) and (23), one can rewrite (22) as

$$dr(T)/dT = J(T, T)(B(T) + A(T)r(T) + C(T)g(T)),\ r(O) = O. \tag{24}$$

If one differentiates (21) with respect to T, one obtains

$$df(T)/dT = J(T,T)D(T) + \int_0^T \partial J(s',T)/\partial T D(s')ds'$$

$$= J(T,T)D(T) + J(T,T)\int_0^T \Phi(s',T)D(s')ds'$$

$$= J(T,T)(D(T) + A(T)\int_0^T J(s',T)D(s')ds' + C(T)\int_0^T L(s',T) D(s')ds')$$

$$= J(T,T)(D(T) + A(T)f(T) + C(T)I(T)),\ f(O) = O, \quad (25)$$

by using (14), (16), (21) and a newly introduced function $I(T)$ given by

$$I(T) = \int_0^T L(s',T)D(s')ds'. \quad (26)$$

If one differentiates (23) with respect to T, one obtains

$$dg(T)/dT = L(T,T)B(T) + L(T,T)(A(T)r(T) + C(T)g(T)),\ g(O) = O, \quad (27)$$

by using (14), (18), (20) and (23). A differential equation for the function $I(T)$ becomes

$$dI(T)/dT = L(T,T)(D(T) + A(T)f(T) + C(T)I(T)),\ I(O) = O, \quad (28)$$

by differentiating (26) with respect to T and using (14), (18), (21) and (26).
The function $L(T,T)$, which appeared in (18), is developed as

$$L(T,T)R = -D^T(T) - \int_0^T L(s',T)(K(s',T) + K_c(s',T))ds'$$

$$= -D^T(T) - \int_0^T L(s',T)B(s')A^T(T)ds' - \int_0^T L(s',T)D(s')C^T(T)ds'.$$

$$= -D^T(T) - g(T)A^T(T) - I(T)C^T(T) \quad (29)$$

by using (23) and (26) with the property of the semi-degenerate kernels in $K(s', T)$ and $K_c(s', T)$ for $0 \leqq s' \leqq T$.

The fixed-point smoothing estimate $\hat{z}(t, T)$ was formulated in (4). If one differentiates (4) with respect to T, one obtains a partial differential equation for $\hat{z}(t, T)$ from (11) and (14).

$$\partial \hat{z}(t, T)/\partial T = h(t, T, T)y(T) + \int_0^T \partial h(t, s, T)/\partial T y(s)\,ds$$

$$= h(t, T, T)y(T) + h(t, T, T)\int_0^T \Phi(s, T)y(s)\,ds$$

$$= h(t, T, T)\,(y(T) + A(T)\int_0^T J(s, T)y(s)\,ds +$$

$$C(T)\int_0^T L(s, T)y(s)\,ds) \qquad (30)$$

Introducing functions $e(T)$ and $Q(T)$ expressed by

$$e(T) = \int_0^T J(s, T)y(s)\,ds \qquad (31)$$

and

$$Q(T) = \int_0^T L(s, T)y(s)\,ds, \qquad (32)$$

one can rewrite (30) as

$$\partial \hat{z}(t, T)/\partial T = h(t, T, T)\,(y(T) + A(T)e(T) + C(T)Q(T)), \ \hat{z}(t, 0) = 0. \qquad (33)$$

A differential equation for the function $e(T)$ is derived by differentiating (31) with respect to T and using (14), (16), (31) and (32) as

$$de(T)/dT = J(T,T)\,(y(T) + A(T)e(T) + C(T)Q(T)),\ e(O) = O. \quad (34)$$

If one differentiates (32) with respect to T and uses (14), (18), (31) and (32), one obtains a differential equation for the function $Q(T)$ as

$$dQ(T)/dT = L(T,T)\,(y(T) + A(T)e(T) + C(T)Q(T)),\ Q(O) = O. \quad (35)$$

The function $h(t, T, T)$ in (11) is still unkown. If one put $s = T$ in (8), one obtains

$$h(t,T,T)R = K(t,T) - \int_O^T h(t,s',T)\,(K(s',T) + K_c(s',T))\,ds'. \quad (36)$$

Since the autocovariance functions of the signal and the colored observation noise are expressed by $K(s',T) = B(s')A^T(T)$ and $K_c(s',T) = D(s')C^T(T)$ for $O \leqq s' \leqq T$ from the property of the semi-degenerate kernels, (36) is transformed into

$$h(t,T,T)R = K(t,T) - \int_O^T h(t,s',T)B(s')\,ds'A^T(T) - \int_O^T h(t,s',T)D(s')\,ds'\,C^T(T). \quad (37)$$

If one introduces functions $S(t, T)$ and $W(t, T)$ expressed by

$$S(t,T) = \int_O^T h(t,s',T)B(s')\,ds' \quad (38)$$

and

$$W(t,T) = \int_O^T h(t,s',T)D(s')\,ds', \quad (39)$$

one obtains an expression for $h(t, T, T)$ as

$$h(t,T,T) = (K(t,T) - S(t,T)A^T(T) - W(t,T)C^T(T))R^{-1}. \quad (40)$$

If one differentiates (38) with respect to T and uses (11), (14), (20) and (23), a partial differential equation for the function $S(t, T)$ is developed as

$$\partial S(t, T)/\partial T = h(t, T, T) B(T) + \int_0^T \partial h(t, s', T)/\partial T B(s')\, ds'$$

$$= h(t, T, T) (B(T) + \int_0^T \Phi(s', T) B(s')\, ds')$$

$$= h(t, T, T) (B(T) + A(T) \int_0^T J(s', T) B(s')\, ds' +$$

$$C(T) \int_0^T L(s', T) B(s')\, ds')$$

$$= h(t, T, T) (B(T) + A(T) r(T) + C(T) g(T)),\ S(t, O) = O. \quad (41)$$

If one differentiates (39) with respect to T, a partial differential equation for the function $W(t, T)$ becomes

$$\partial W(t, T)/\partial T = h(t, T, T) (D(T) + A(T) f(T) + C(T) I(T)),\ W(t, O) = O, \quad (42)$$

from (11), (14), (21) and (26).

If one wants to start with the filtering estimate $\hat{z}(t, t)$ at the fixed-point t in the calculation of the fixed-point smoothing estimate $\hat{z}(t, T)$ by the partial differential equation (33), one needs the value of $\hat{z}(t, t)$. Fundamentally, the filtering estimate $\hat{z}(t, t)$ is an estimate which is formulated by putting $T = t$ in (4) and is expressed as a linear integral transformation of the observed value set $\{y(s), O \leqq s \leqq t\}$. An optimal impulse response function $h(t, s, t)$, which calculates the linear least-squares filtering estimate, satisfies the following integral equation

$$h(t, s, t) R = K(t, s) - \int_0^T h(t, s', t) (K(s', t) + K_c(s', t))\, ds'. \quad (43)$$

It is noted that the inequality $O \leqq s < t$ holds in (43). Substituting the expression $K(t, s) = A(t) B^T(s)$ for $O \leqq s < t$, which is represented by the semi-degenerate property for the autocovariance function of the signal $z(t)$, into (43), one obtains

$$h(t, s, t) R = A(t) B^T(s) - \int_0^T h(t, s', t) (K(s', t) + K_c(s', t))\, ds'. \quad (44)$$

Comparing (44) with (12) leads to an expression for the optimal impulse response function $h(t, s, t)$

$$h(t, s, t) = -A(t) J(s, t). \tag{45}$$

If one substitutes (45) into an equation, which is obtained by putting $T=t$ in (4), and uses (31), one obtains an expression for the filtering estimate $\hat{z}(t, t)$.

$$\hat{z}(t, t) = -A(t) e(t) \tag{46}$$

Also, the initial conditions at $T=t$ in the partial differential equations (41) and (42) are $S(t, t) = -A(t) r(t)$ and $W(t, t) = -A(t) f(t)$ respectively by substituting (45) into (38) and (39), and by using (20) and (21).

Now, let us summarize above results. The fixed-point smoothing estimate $\hat{z}(t, T)$ and the filtering estimate $\hat{z}(t, t)$ are calculated sequentially by (47)-(59).

Fixed-point smoothing estimate : $\hat{z}(t, T)$

$$\partial \hat{z}(t, T) / \partial T = h(t, T, T) (y(T) + A(T) e(T) + C(T) Q(T)), \ \hat{z}(t, O) = O \tag{47}$$

(Initial condition of $\hat{z}(t, T)$ at $T=t$ is $\hat{z}(t, t)$.)

$$de(T) / dT = J(T, T) (y(T) + A(T) e(T) + C(T) Q(T)), \ e(O) = O \tag{48}$$

$$dQ(T) / dT = L(T, T) (y(T) + A(T) e(t) + C(T) Q(T)), \ Q(O) = O \tag{49}$$

$$h(t, T, T) = (B(t) A^T(T) - S(t, T) A^T(T) - W(t, T) C^T(T)) R^{-1}. \tag{50}$$

$$\partial S(t, T) / \partial T = h(t, T, T) (B(T) + A(T) r(T) + C(T) g(T)), \ S(t, O) = O \tag{51}$$

(Initial condition of $S(t, T)$ at $T=t$ is $S(t, t) = -A(t) r(t)$.)

$$\partial W(t, T) / \partial T = h(t, T, T) (D(T) + A(T) f(T) + C(T) I(T)), \ W(t, O) = O \tag{52}$$

(Initial condition of $W(t, T)$ at $T=t$ is $W(t, t) = -A(t) f(t)$.)

$$dr(T) / dT = J(T, T) (B(T) + A(T) r(T) + C(T) g(T)), \ r(O) = O \tag{53}$$

$$df(T) / dT = J(T, T) (D(T) + A(T) f(T) + C(T) I(T)), \ f(O) = O \tag{54}$$

$$J(T, T) = (-B^T(T) - r(T) A^T(T) - f(T) C^T(T)) R^{-1} \tag{55}$$

$$dg(T)/dT = L(T, T) (B(T) + A(T) r(T) + C(T) g(T)), \ g(0) = 0 \tag{56}$$

$$dI(T)/dT = L(T, T) (D(T) + A(T) f(T) + C(T) I(T)), \ I(0) = 0 \tag{57}$$

$$L(T, T) = (-D^T(T) - g(T) A^T(T) - I(T) C^T(T)) R^{-1} \tag{58}$$

$$\hat{z}(t, t) = -A(t) e(t) \ \text{(Filtering estimate)} \tag{59}$$

Also, the optimal impulse response function $h(t, s, T)$ is calculated by the following sequential algorithm.

$$\partial h(t, s, T)/\partial T = h(t, T, T) \Phi(s, T), \ h(t, s, 0) = K(t, s) R^{-1} \tag{60}$$

$$\Phi(s, T) = A(T) J(s, T) + C(T) L(s, T) \tag{61}$$

$$\partial J(s, T)/\partial T = J(T, T) \Phi(s, T), \ J(s, 0) = -B^T(s) R^{-1} \tag{62}$$

$$\partial L(s, T)/\partial T = L(T, T) \Phi(s, T), \ L(s, 0) = -D^T(s) R^{-1} \tag{63}$$

Here, the functions $J(T, T)$ and $L(T, T)$ are calculated by (55) and (58).

4. Smoothing error covariance function

Let us derive an equation for a smoothing error covariance function. The smoothing error covariance function is defined by

$$P(t, s, T) = E[(z(t) - \hat{z}(t, T))(z(s) - \hat{z}(s, T))^T], \ 0 \leqq s, t \leqq T. \tag{64}$$

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

$$P(t, s, T) = K(t, s) - E[\hat{z}(t, T) z^T(s)]. \tag{65}$$

Substituting (4) into (65), and using (1) and (3) with an expression of $K(s', s) = E[z(s') z^T(s)]$, one obtains

$$P(t, s, T) = K(t, s) - \int_0^T h(t, s', T) K(s', s) ds'. \tag{66}$$

If one differentiates (66) with respect to T and uses (60), one obtains

$$\begin{aligned} \partial P(t, s, T)/\partial T &= -h(t, T, T) K(T, s) - \int_0^T \partial h(t, s', T)/\partial T K(s', s) ds' \\ &= -h(t, T, T) \left(A(T) B^T(s) + \int_0^T \Phi(s', T) K(s', s) ds'\right) \end{aligned} \tag{67}$$

in terms of $K(T, s) = A(T) B^T(s)$. If one introduces a function $F(s, T)$ defined by

$$F(s, T) = \int_0^T \Phi(s', T) K(s', s) ds', \tag{68}$$

one obtains a partial differential equation for $P(t, s, T)$ as

$$\partial P(t, s, T)/\partial T = -h(t, T, T) (A(T) B^T(s) + F(s, T)), \ P(t, s, O) = K(t, s). \tag{69}$$

Substituting (61) into (68), one obtains

$$F(s, T) = A(T) \int_0^T J(s', T) K(s', s) ds' + C(T) \int_0^T L(s', T) K(s', s) ds'. \tag{70}$$

If one introduces functions $U(s, T)$ and $V(s, T)$ expressed by

$$U(s, T) = \int_0^T J(s', T) K(s', s) ds' \tag{71}$$

and

$$V(s, T) = \int_0^T L(s', T) K(s', s) ds', \tag{72}$$

one can rewite (70) as

$$F(s, T) = A(T) U(s, T) + C(T) V(s, T). \tag{73}$$

If one differentiates (71) with respect to T and uses (61), (62), (71) and (72), one obtains a partial differential equation for the function $U(s, T)$ as

$$\partial U(s, T) / \partial T = J(T, T) (A(T) B^T(s) + A(T) U(s, T) + C(T) V(s, T)), \; U(s, 0) = 0. \tag{74}$$

Also, a partial differential equation for the function $V(s, T)$ becomes

$$\partial V(s, T) / \partial T = L(T, T) (A(T) B^T(s) + A(T) U(s, T) + C(T) V(s, T)), \; V(s, 0) = 0, \tag{75}$$

from (61), (63), (71) and (72).

Therefore, the sequential algorithm for the smoothing error covariance function $P(t, s, T)$ consists of (50)-(58), (69), (73), (74) and (75).

Now, the smoothing error covariance function $P(t, s, T)$ is written as

$$\begin{aligned} P(t, s, T) &= K(t, s) - E[\hat{z}(t, T) \hat{z}^T(s, T)] \\ &= K(t, s) - P_z(t, s, T), \end{aligned} \tag{76}$$

where $P_z(t, s, T)$ denotes an autocovariance function of the fixed-point smoothing estimate $\hat{z}(t, T)$. $P_z(t, s, T)$ is a positive semi-definite matrix, and the smoothing error covariance matrix $P(t, s, T)$ is also positive semi-definite. Therefore, one notices that there exists a relationship

$$0 \leqq P_z(t, s, T) \leqq K(t, s). \tag{77}$$

This relation for the white Gaussian noise is shown (Kailath, 1976). According to a discussion about stability problems (Kailath, 1976), (77) ensures that the presented smoothing algorithm has a unique solution, since $P_z(t, s, T)$ is lower and upper bounded.

From (69) with (76), one finds that $P_z(t, s, T)$ satisfies a similar partial differential equation with that for $P(t, s, T)$ as

$$\partial P_z(t, s, T) / \partial T = h(t, T, T) (A(T) B^T(s) + F(s, T)), \; P_z(t, s, 0) = 0. \tag{78}$$

5. Conclusions

In this paper, the continuous linear optimal filter was designed based on the Wiener-Hopf integral equation in linear least-squares estimation theory. The fixed-point smoothing algorithm presented here only uses information of $A(T)$, $B(T)$, $C(T)$, $D(T)$ and R with the observed value $y(T)$. Therefore, the current fixed-point smoother does not need information of system matrix etc. in a state-space model. This point is the basic difference in approaches between the presented smoother and the Kalman filter based on the state-space model.

REFERENCES

Kailath, T. , 1976, Lectures on Linear Least-Squares Estimation (Springer-Verlag).
Kalman, R. E. , 1960, Trans. ASME, J. Basic Eng. , 82D, 34.
Nakamori, S. , and Sugisaka, M. , 1977, Automatica, 13, 623.
Sage, A. P. , and Melsa, J. L. , 1971, Estimation Theory with Applications to Communications and Control (McGraw-Hill).
Trees, H. L. , 1968, Detection, Estimation and Modulation Theory, Part 1 (Wiley).