Design of Predictor Using Covariance Information in Continuous-Time Stochastic Systems with Nonlinear Observation Mechanism

Seiichi NAKAMORI

(Received September 29, 1997)

Abstract

This paper proposes a new design method of a nonlinear prediction algorithm in continuos-time stochastic systems. The observd value consists of nonlinearly modulated signal and additive white Gaussian observation noise. The prediction algorithm is designed based on the same idea as the extended Kalman filter is obtained from the recursive least-squares Kalman filter in linear continuous-time stochastic systems. The proposed predictor necessitates the information of the autocovariance function of the signal, the variance of the observation noise, the nonlinear observation function and its differentiated one with respect to the signal. The proposed predictor is compared in estimation accuracy with the MAP predictor from both theoretical and numerical aspects.

1. Introduction

In linear least-squares estimation problem, recursive least-squares estimation algorithms [3] have been devised. They use the covariance information of the signal and observation noise. The linear estimation algorithms are used limitedly within linear systems and can not be applied to estimation problem in nonlinear systems directly. From this respect, this paper presents a new design method of the nonlinear predictor by use of the covariance information for the signal with nonlinear observation mechanism in continuous-time stochastic systems. Here, the observed value consists of the nonlinear function of the signal and additive white Gaussian observation noise. In this paper, the prediction algorithm is devised based on same idea as the extended Kalman filter [1],[2],[4] for nonlinear estimation problem. The extended Kalman filter is designed based on the recursive least-squares Kalman filter in linear stochas

Department of Technology (Electric Technology), Faculty of Education, Kagoshima University, Kagoshima 890-0065, Japan

tic systems. In this paper, the extended predictor using the covariance information is designed in the relation with the recursive linear least-squares predictor in [Theorem 1] by use of the covariance information in linear continuous-time stochastic systems. The present predictor uses the information of the autocovariance function of the signal, the variance of the observation noise, the nonlinear function of the signal and its differentiated one with respect to the signal. The proposed nonlinear prediction algorithm estimates the signal recursively via an updated observed value.

Incidentally, as an approach, which uses the same information as the present predictor, to the nonlinear estimation and modulation problem, the MAP (maximum a posterior) estimation technique [5] is developed. The current nonlinear predictor is compared in estimation accuracy with the nonlinear MAP predictor in continuous-time stochastic systems both theoretically and numerically.

2. Recursive least-squares prediction problem in linear continuous-time stochastic systems

As a step toward the extended prediction problem, we consider recursive least-squares prediction problem in linear continuous-time stochastic systems.

Let an observation equation be given by

$$y(t) = H(t)x(t) + v(t), \tag{1}$$

where $y(t)$ is an observed value of dimension m, $x(t)$ is a zero-mean signal of dimension n and $v(t)$ is white Gaussian observation noise with the variance R.

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

Here, the symbol "T" represents transpose and the symbol δ $(t\text{-}s)$ the Dirac delta function. We assume that the signal $x(t)$ is uncorrelated with the observation noise $v(s)$ as

$$E[x(t)v^T(s)] = 0, \quad 0 \le s,t < \infty. \tag{3}$$

Let $\hat{x}(t,t+\alpha)$ represent the α time ahead prediction estimate of the signal $x(t)$. Let $\hat{x}(t,t+\alpha)$ be expressed by

$$\hat{x}(t,t+\alpha) = \int_0^t h(t,s)y(s)ds, \tag{4}$$

where $h(t,s)$ is an impulse response function. Minimizing the mean-square value of the prediction error $x(t+\alpha) - \hat{x}(t,t+\alpha)$

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

we obtain the Wiener-Hopf integral equation [4]

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

If we substitute (1) into (6), use the stochastic property of the signal $x(t)$ and observation noise $v(t)$ and represent the autocovariance function of the signal by $K_x(t,s)(=E[x(t)x^T(s)])$, we obtain

$$h(t,s)R = K_x(t+\alpha,s)H^T(s) - \int_0^t h(t,s')H(s')K_x(s',s)ds' H^T(s). \tag{7}$$

We assume that the autocovariance function of the signal is expressed in the semi-degenerate kernel from [3] of

$$K_x(t,s) = E[x(t)x^T(s)] = \begin{cases} A(t)B^T(s), & 0 \le s \le t, \\ B(t)A^T(s), & 0 \le t \le s, \end{cases} \tag{8}$$

where $A(t)$ and $B(s)$ are n-by-l bounded matrices.

In section 3, we derive the recursive least-squares predictor under the above assumptions in linear continuous-time stochastic systems as a step toward the nonlinear extended predictor.

3. Recursive least-squares prediction algorithm in linear continuous-time stochastic systems

[Theorem 1] presents the recursive least-squares prediction algorithm in linear continuous-time stochastic systems.

[Theorem 1]

Let the observation equation be given by (1), let the autocovariance function $K_x(t,s)$ of the signal $x(t)$ be expressed in the semi-degenerate kernel form of (8) and let the signal be uncorrelated with the observation noise $v(s)$ as (3). Then, the recursive least-squares algorithm for the α time ahead prediction estimate $\hat{x}(t,t+\alpha)$ of the signal $x(t)$ consists of (9)~(13) in linear continuous-time stochastic systems.

α time ahead prediction estimate of $x(t)$: $\hat{x}(t,t+\alpha)$

$$\hat{x}(t,t+\alpha) = A(t+\alpha)e(t) \tag{9}$$

Filtering estimate of $x(t)$: $\hat{x}(t,t)$

$$\hat{x}(t,t) = A(t)e(t) \tag{10}$$

$$\frac{de(t)}{dt} = J(t,t)(y(t) - H(t)\hat{x}(t,t)), \quad e(0) = 0 \tag{11}$$

$$J(t,t) = (B^T(t) - r(t)A^T(t))H^T(t)R^{-1} \tag{12}$$

$$\frac{dr(t)}{dt}=J(t,t)H(t)(B(t)-A(t)r(t)),\quad r(0)=0 \tag{13}$$

Proof

Since $K_x(t,s)=A(t)B^T(s)$ for $0\le s\le t$ from(8), (7) is written as

$$h(t,s)R=A(t+\alpha)B^T(s)H^T(s)-\int_0^t h(t,s')H(s')K_x(s',s)ds'H^T(s). \tag{14}$$

If we introduce the function $J(t,s)$ which satisfies

$$J(t,s)R=B^T(s)H^T(s)-\int_0^t J(t,s')H(s')K_x(s',s)ds'H^T(s), \tag{15}$$

we obtain

$$h(t,s)=A(t+\alpha)J(t,s) \tag{16}$$

for the impulse response function $h(t,s)$.

If we differentiate (15) with respect to t, we have

$$\frac{\partial J(t,s)}{\partial t}R=-J(t,t)H(t)K_x(t,s)H^T(s)-\int_0^t\frac{\partial J(t,s')}{\partial t}H(s')K_x(s',s)ds'H^T(s). \tag{17}$$

From (15) and (17), we have a partial differential equation

$$\frac{\partial J(t,s)}{\partial t}=-J(t,t)H(t)A(t)J(t,s) \tag{18}$$

for $J(t,s)$.

If we put $s=t$ in (15), we have

$$J(t,t)R=B^T(t)H^T(t)-\int_0^t J(t,s')H(s')K_x(s',t)ds'H^T(t). \tag{19}$$

Since $K_x(t,s)=B(t)A^T(s)$ for $0\le t\le s$ from (8), (19) is written as

$$J(t,t)R=B^T(t)H^T(t)-\int_0^t J(t,s')H(s')B(s')ds'A^T(t)H^T(t). \tag{20}$$

If we introduce the function $r(t)$ given by

$$r(t)=\int_0^t J(t,s')H(s')B(s')ds', \tag{21}$$

we obtain (12) for the function $J(t,t)$.

If we differentiate (21) with respect to t and use (18), we have

$$\frac{dr(t)}{dt}=J(t,t)H(t)B(t)-J(t,t)H(t)A(t)\int_0^t J(t,s')H(s')B(s')ds'. \tag{22}$$

From (21) and (22), we obtain the differential equation (13) for $r(t)$. The initial condition on the differential equation for $r(t)$ at $t=0$ is $r(0)=0$ from (21).

If we substitute (16) into (4) and introduce the function $e(t)$ given by

$$e(t)=\int_0^t J(t,s')y(s')ds', \tag{23}$$

we obtain (9) for the α time ahead prediction estimate $\hat{x}(t,t+\alpha)$ of $x(t)$.

If we differentiate (23) with respect to t and use (18), we obtain the differential equation (11) for $e(t)$. The initial condition on the differential equation for $e(t)$ at $t=0$ is $e(0)=0$ from (23). □

In section 4, referring to the linear prediction algorithm of [Theorem 1], we design the extended predictor using the covariance information.

4. Design of extended predictor for estimating signal with nonlinear observation mechanism

Let the signal be observed with nonlinear mechanism by

$$y(t)=f(x(t),t)+v(t), \tag{24}$$

where the signal $x(t)$ and the observation noise $v(t)$ have the same stochastic properties as those in section 2.

(24) can be expressed as a Taylor series by expanding about nominal trajectory $x_n(t)$ [1],[2]:

$$\begin{aligned} y(t)-y_n(t) &= f(x(t),t)-f(x_n(t),t)+v(t), \quad y_n(t)=f(x_n(t),t) \\ &= \left.\frac{\partial f(x(t),t)}{\partial x(t)}\right|_{x(t)=x_n(t)}(x(t)-x_n(t))+h.o.t+v(t) \\ &= H(t)(x(t)-x_n(t))+h.o.t+v(t), \quad H(t)=\left.\frac{\partial f(x(t),t)}{\partial x(t)}\right|_{x(t)=x_n(t)}. \end{aligned} \tag{25}$$

Here, "h.o.t." are terms in powers of $x(t)-x_n(t)$ greater than one and $H(t)$ is m-by-n matrix. A linearization of this relation to the first order yields the perturbation measurement model of

$$\delta y(t)=H(t)\delta x(t)+v(t), \quad \delta y(t)=y(t)-y_n(t), \quad \delta x(t)=x(t)-x_n(t). \tag{26}$$

The output of predictor, which is called the linearlized predictor or perturbation predictor,based on the observation equation (26) would be the optimal prediction estimate of $\delta x(t)$. This could be added to the nominal value $x_n(t)$ to obtain the estimate of the signal $x(t)$. From (26) we find that we can estimate $x(t)$ in terms of $y(t)-y_n(t)+H(t)x_n(t)$, $y_n(t)=f(x_n(t),t)$, through the relationship $y(t)-y_n(t)+H(t)x_n(t)=H(t)x(t)+v(t)$. In the continuous-time systems, $x_n(t)$ is set to the filtering estimate $\hat{x}(t,t)$ of $x(t)$ [1],[2].Since $y(t)-f(\hat{x}(t,t),t)+H(t)$

$\hat{x}(t,t)(=H(t)x(t)+v(t))$ has an equivalent information to $y(t)$, the filtering estimate $\hat{x}(t,t)$ of $x(t)$ can be estimated by the observed value $y(t)$.

As in the extended Kalman filter we also use the linearlized observation matrix $H(t)$ in the current filtering equations. Although (12) and (13) in [Theorem 1] do not include the term of the observed value $y(t)$, $H(t)$, given by $H(t)=\left.\frac{\partial f(x(t),t)}{\partial x(t)}\right|_{x(t)=\hat{x}(t,t)}$, in these equations might be dependent on the observed value. According to the guideline in the design of the extended Kalman filter[1],[2], $H(t)$ in (11),(12) and (13) are replaced with $H(t)=\left.\frac{\partial f(x(t),t)}{\partial x(t)}\right|_{x(t)=\hat{x}(t,t)}$ in the current nonlinear extended prediction problem using the covariance information.

Also, based on the procedure in the design of the extended Kalman filter from the Kalman filter, the term "$H(t)\hat{x}(t,t)$" in (11) is replaced with "$f(\hat{x}(t,t),t)$" in the current nonlinear extended predictor. (9) is included as it is also in the extended predictor using the covariance information.

[Theorem 2] summarizes the extended prediction algorithm using the covariance information. The validity of the above design technique is examined by a numerical simulation example in section 6.

[Theorem 2]

Let the nonlinear observation equation be given by (24), let the autocovariance function of the signal $x(t)$ be expressed by (8) in the semi-degenerate kernel form, let $H(t)$ be given by $H(t)=\left.\frac{\partial f(x(t),t)}{\partial x(t)}\right|_{x(t)=\hat{x}(t,t)}$ and let the signal $x(t)$ and the observation noise $v(s)$ be uncorrelated as (3). Then, the recursive least-squares algorithm for the α time ahead prediction estimate $\hat{x}(t,t+\alpha)$ of the signal $x(t)$ consists of (27)~(31) in continuous-time stochastic systems with nonlinear observation mechanism of the signal.

α time ahead prediction estimate of $x(t)$: $\hat{x}(t,t+\alpha)$

$$\hat{x}(t,t+\alpha)=A(t+\alpha)e(t) \tag{27}$$

Filtering estimate of $x(t)$: $\hat{x}(t,t)$

$$\hat{x}(t,t)=A(t)e(t) \tag{28}$$

$$\frac{de(t)}{dt}=J(t,t)(y(t)-f(\hat{x}(t,t),t)),\quad e(0)=0 \tag{29}$$

(Filtering estimate of $f(x(t),t)$: $f(\hat{x}(t,t),t)$)

$$J(t,t) = (B^T(t) - r(t)A^T(t))H^T(t)R^{-1} \tag{30}$$

$$\frac{dr(t)}{dt} = J(t,t)H(t)(B(t) - A(t)r(t)), \quad r(0) = 0 \tag{31}$$

5. Comparison of estimation accuracy of current predictor with MAP prediction algorithm

In this section, we introduce the MAP predictor [5] in estimating stochastic signal with nonlinear observation mechanism.

In estimating the α time ahead signal, the MAP prediction estimate of the signal $x(t)$ is given by

$$\hat{x}(t,t+\alpha) = \int_0^t K_x(t+\alpha,s')\left.\frac{\partial f^T(x(s'),s')}{\partial x(s')}\right|_{x(s')=\hat{x}(s',s')} R^{-1}(y(s') - f(\hat{x}(s',s'),s'))ds'. \tag{32}$$

If we substitute the relations $H(s') = \left.\frac{\partial f(x(s'),s')}{\partial x(s')}\right|_{x(s')=\hat{x}(s',s')}$ and $K_x(t+\alpha,s') = A(t+\alpha)B^T(s')$, $0 \le s \le t$, from (8) into (32), we have

$$\hat{x}(t,t+\alpha) = A(t+\alpha)\int_0^t B^T(s')H^T(s')R^{-1}(y(s') - f(\hat{x}(s',s'),s'))ds'. \tag{33}$$

If we introduce the function $q(t)$ expressed by

$$q(t) = \int_0^t B^T(s')H^T(s')R^{-1}(y(s') - f(\hat{x}(s',s'),s'))ds' \tag{34}$$

and differentiate (34) with respect to t, we obtain the recursive MAP prediction equations as follows:

$$\hat{x}(t,t+\alpha) = A(t+\alpha)q(t), \tag{35}$$

$$\frac{dq(t)}{dt} = B^T(t)H^T(t)R^{-1}(y(t) - f(\hat{x}(t,t),t)), \quad q(0) = 0, \tag{36}$$

where the initial condition on the differential equation (36) at $t=0$ is clear from (34). The MAP prediction estimate is calculated by (35) and (36) in terms of the same information as in [Theorem 2] for the nonlinear extended predictor using the covariance information.

Let us compare the predictor in [Theorem 2] with the MAP predictor from the theoretical point of estimation accuracy.

If we differentiate (27) with respect to t and use (29), we have

$$\frac{d\hat{x}(t,t+\alpha)}{dt} = \frac{dA(t+\alpha)}{dt}e(t) + A(t+\alpha)J(t,t)(y(t) - f(\hat{x}(t,t),t)), \tag{37}$$

where $A(t+\alpha)J(t,t)$ represents the predictor gain. From (16) and (30), the predictor gain $h(t,t)(=A(t+\alpha)J(t,t))$ might be expressed by

$$h(t,t)=(A(t+\alpha)B^T(t)-A(t+\alpha)r(t)A^T(t))H^T(t)R^{-1}. \tag{38}$$

Let $h_{MAP}(t,t)$ represent the prediction gain for the MAP predictor. Then we have $h_{MAP}(t,t)=A(t+\alpha)B^T(t)H^T(t)R^{-1}$ from (35) and (36).

If we compare the MAP prediction equations (35) and (36) with the proposed prediction algorithms in [Theorem 2], we find that the MAP prediction equations are obtained by setting $r(t)=0$ in the prediction algorithms of [Theorem 2].

Let $P_{\tilde{x}}(t+\alpha,t+\alpha)$ represent the prediction error variance function of the proposed predictor. From (4), (16) and (21), $P_{\tilde{x}}(t+\alpha,t+\alpha)$ might be developed as

$$\begin{aligned} P_{\tilde{x}}(t+\alpha,t+\alpha) &= E[(x(t+\alpha)-\hat{x}(t,t+\alpha))(x(t+\alpha)-\hat{x}(t,t+\alpha))^T] \\ &= E[(x(t+\alpha)-\hat{x}(t,t+\alpha))x(t+\alpha)^T] \\ &= K_x(t+\alpha,t+\alpha)-\int_0^t h(t,s)E[y(s)x(t+\alpha)^T]ds \\ &= K_x(t+\alpha,t+\alpha)-A(t+\alpha)r(t)A^T(t+\alpha). \end{aligned} \tag{39}$$

Let $P_{\hat{x}}(t+\alpha,t+\alpha)$ represent the autovariance function of the prediction estimate $\hat{x}(t,t+\alpha)$, i.e., $P_{\hat{x}}(t+\alpha,t+\alpha)=A(t+\alpha)r(t)A^T(t+\alpha)$. Let $P_{\tilde{x}_{MAP}}(t+\alpha,t+\alpha)$ represent the prediction error variance function of the MAP predictor. It is clear that the autovariance function of the prediction estimate for the MAP predictor is the zero square matrix of order n and $P_{\tilde{x}_{MAP}}(t+\alpha, t+\alpha)=K_x(t+\alpha,t+\alpha)$. Since $P_{\tilde{x}}(t+\alpha,t+\alpha)$ and $P_{\hat{x}}(t+\alpha,t+\alpha)$ are nonnegative definite matrices, the relationship

$$0 \le P_{\tilde{x}}(t+\alpha,t+\alpha) \le P_{\tilde{x}_{MAP}}(t+\alpha,t+\alpha)(=K_x(t+\alpha,t+\alpha)) \tag{40}$$

is valid. Hence, it is seen that the proposed predictor is superior or equal to the MAP predictor in estimation accuracy.

6. A numerical simulation example

In this section a simulation example is demonstrated. The example shows the prediction for a stochastic signal with a nonlinear observation mechanism.

Let the observation equation be given by

$$y(t)=f(x(t),t)+v(t),\ \ f(x(t),t)=(1+x(t))^2, \tag{41}$$

where we assume that the process of the signal $x(t)$ is stationary stochastic with the autocovariance function given by

$$K_x(t,s) = K_x(t-s)$$
$$= Pe^{-k(t-s)}, \; P = 10, \; k = 0.5 \; . \tag{42}$$

From (8) and (42), we have the functions $A(t)$ and $B(s)$ expressed by

$$A(t) = Pe^{-kt}, \; B(s) = e^{ks}. \tag{43}$$

$H(t)$ is given by

$$H(t) = \left.\frac{\partial f(x(t),t)}{\partial x(t)}\right|_{x(t)=\hat{x}(t,t)}$$
$$= 2(1+\hat{x}(t,t)). \tag{44}$$

If we substitute the expressions for $f(\hat{x}(t,t),t)$, $H(t)$, $A(t)$ and $B(t)$ above into the prediction algorithm of [Theorem 2], we can calculate the prediction estimate $\hat{x}(t,t+\alpha)$ and the filtering estimate $f(\hat{x}(t,t),t)$ of $f(x(t),t)$ recursively.

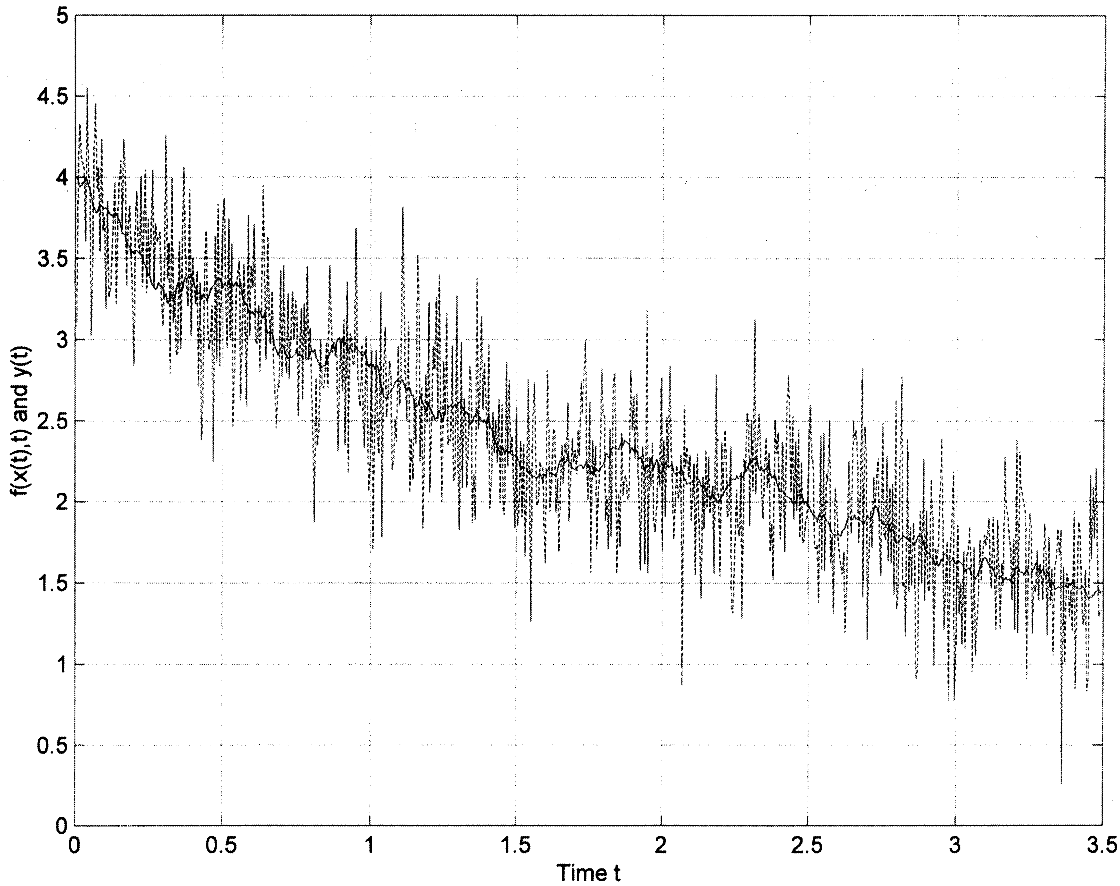

Fig.1 Sequences of $f(x(t),t)$ (solid line) and the observed value $y(t)$ (dashed line) for the observation noise $N(0,0.4^2)$.