推計 二線型 시스템의 狀態推定

State Estimation of Stochastic Bilinear System

論 文
30～11～2

黃 春 植*

(Chun-Sik Hwang)

Abstract

Most of real world systems are highly non-linear. But due to difficulties in analyzing and dealing with it, only the linear system theory is well estabilished. Bilinear system where state and control are linear but not linear jointly is introduced. Here shows that optimal state estimation of stachastic bilinear system requirs infinite dimensional filter, thus onesub-optimal estimator for this system is suggested.

1. Structure of Bilinear System (BLS)

Almost of our real world systems have highly non linear structures. But because of difficulties in analyzing and dealing with it, it is general and standard procedures to linearize it about the operating steady state or nominal values using Taylor Series Expansion and to dealing with the resulting linearized systems. Consider a general non-linear system described by the state space model.

$$\dot{X}(t)=f(X,\ U,\ t),\ t\in(t_0,\ \infty) \tag{1}$$

where $X\in R^n$, $U\in R^m$, $f(.)$ is n-function.

Expansion of (1) in a Taylor series about nominal state, say, X_n and control U_n gives [1].

$$\Delta\dot{X}=\frac{\partial f()}{\partial X}\Big|_{U_n}^{X_n}\Delta X+\frac{\partial f()}{\partial U}\Big|_{U_n}^{X_n}\Delta U+\frac{\partial^2 f()}{\partial X\partial U}\Big|_{U_n}^{X_n}\Delta X\Delta U + \text{High Order Term} \tag{2}$$

If we neglect the high order term, we have a system where the state and the control are linear but not linear in the state and control jointly, i.e. mult iplicative mode term of state and control appears. Such a system is commonly called as a bilinear system (BLS) and can be described by the following differential equation form.

$$\dot{X}=AX+CU+B\ (X,\ U) \tag{3}$$

Where A is nxn matrix
C is nxm matrix
B ( ) is the bilinear operator

The bilinear term B ( ) can be rewritten as

$$B(X,\ U)=\sum_{i=1}^{m} B_iU_iX \tag{4}$$

Where B_i is nxn and U_i is the i-th component of input U.

Then BLS (3) becomes

$$\dot{X}=AX+\sum_{i=1}^{m} B_iU_iX+CU \tag{5}$$

Standard BLS model (5) is an alternative approximate linearization of a non-linear system (1). Thus it enjoys much amount of linear system properties since BLS is the nearest form to the linear system,

$$\dot{X}=AX+CU \tag{6}$$

We can expect that the various advantages of well estabilished linear system theory can be used for this system [2].

On the other hand, the non-linear structure offers some important advantages on the BLS, for examples controllability, optimalization and modeling etc..

System description by means of BLS can be used

* 正會員：陸軍士官學校 敎授部 電子工學科 敎授
接受日字：1981年 9月 2日

to model many naturally occuring physical processes, for example economics, socio-economics, ecologies, biology, life science, engineerings and military applications, etc,. One example of BLS model in the engineering part is the neutron reactivity control of nuclear reactor. The neutron population is described by a second order model.

$$\dot{n}=\frac{u-\beta}{l}n+\lambda c$$
$$\dot{c}=\beta/l\cdot n-\lambda c \quad (7)$$

where n=neutron population
c=precursor population
u=reactivity, which is control input and
l, λ, β contant.

Another application example of BLS modeling is in military area. A missile intercept problem or more generally pursuit-evasion problem can be modeled by bilinear system form. By denoting the angular rate of the missile and the target with respect to a non-rotating reference coordinates as input U, the dynanics are described by the following homo-geneous-in-the-state BLS.

$$\dot{X}=AX+BXU \quad (8)$$

When add linear observation $Y\in R^q$ to (5), we have

$$Y=HX \quad (9)$$

If the control and observation are corrupted by some noise, then we obtain an Ito-sense stochastic BLS (SBLS) equationas as

$$dX(t)=\left[(A+\sum_{i=1}^{m}B_iU_i)X+CU\right]dt+\sum_{i=1}^{m}B_iXdw_i+Cdw$$
$$dY(t)=HXdt+dv \quad (10)$$

where $A(t)$ contains correction term, and Wiener processes $w(t)$ and $v(t)$ have covariance

$$E[dw,\ dw^T(t)]=Q(t)dt$$
$$E[dv(t),\ dv^T(t)]=R(t)dt \quad (11)$$

where $E(\ )$ represent "expectation" operator and $Q\in R^{n\times n}$, $R\in R^{q\times q}$

Here our problem is to estimate the state $X(t)$ of the system by given the observation $Y(t)$.

2. State Estimation of Stochastic System

Consider again non-linear dynamic system and non-linear observation

$$dX(t)=f(X,\ t)dt+G(X,\ t)dw(t),\quad t\geq t_0,$$
$$X(t_0)=X_0 \quad (12)$$
$$dZ(t)=h(X,\ t)dt+dv(t) \quad (12')$$

If the observation is made over the time period $t_0\leq t\leq\tau$, then $Y\tau$ can be defined as

$$Y\tau=\{z(s),\ t_0\leq s\leq\tau\} \quad (13)$$

The state estimation of (12) based on $Y\tau$ is equivalent to determine the conditional probability density function $P(X,\ t\,|\,Y\tau)$.

This $P(X,\ t\,|\,Y\tau)$ can be obtained by the solution of the Fokker -Planck Type equation.

$$dP=L(P)dt+(h(t)-\hat{h})^TR^{-1}(t)\ \{dZ(t)-\hat{h}dt\}P \quad (14)$$

Where $L(\ )$ denotes Fokker-Planck equation operator.

$$LP=\frac{\partial P}{\partial t}=-\sum_{i=1}^{n}\frac{\partial(Pf_i)}{\partial X_i}+\frac{1}{2}\sum_{i,j=1}^{n}\frac{\partial^2(PGQG^T)_{ij}}{\partial X_i\ \partial X_j} \quad (15)$$

and where

$$\hat{h}=E[h(x,\ t\,|\,Y\tau)]=\int h(X,\ t)p(X,\ t\,|\,Y\tau)dX \quad (15')$$

Since the solution to the (14) is rarely available in general [4], we are satisfied with only the mean and covariance of $X(t)$. But it is known that the mean and covariance are dependent on the higher order term. It means that optimal estimator has infinite dimensional which is not practically implementable. Thus many approximation algorithms are developed to estimate $X(t)$ by many authors by the name of suboptimal estimator.

Here introduce one algorithm to estimate $X(t)$ of SBLS (10) with linear observation.

1) State Estimation of Non-linear System

With the non-linear system eq. (12) without observation (12'), the process $X(t)$ is difined by the density function $P(X,\ t)$ which satisfy the Fokker-Planck eq. (14) as we noted before. Mean and variance eq. of (12) are

$$\frac{d\hat{X}(t)}{dt}=E[\hat{f}(X,\ t)]=\hat{f}(X,\ t) \quad (16)$$

$$\frac{dP(t)}{dt}=\widehat{Xf^T}-\hat{X}\hat{f}^T+\widehat{fX^T}-\hat{f}\hat{X}^T+\widehat{GQG^T} \quad (17)$$

Where $\wedge$ denotes the expectation and T means transpose.

On the other hand, with the observation (12′), the conditional mean and variance of $X(t)$ satisfy the following stochastic differential equation.

$$d\hat{X}=\hat{f}dt+\{\widehat{Xh^T}-\hat{X}\hat{h}^T\}R^{-1}\ \{dZ-\hat{h}\}dt \quad (18)$$

$$\begin{aligned}\{dP\}_{ij}=&[\widehat{X_i f_j}-\hat{X}_i\hat{f}_j+\widehat{f_i X_j}-\hat{f}_i\hat{X}_j+(\widehat{GQG^T})_{ij}\\&-(\widehat{X_i h}-\hat{X}_i\hat{h}^T)^T\ R^{-1}(\widehat{hX_j}-\hat{h}\hat{X}_j)]dt+\\&(\widehat{X_iX_jh}-\widehat{X_iX_j}\hat{h}-\hat{X}_i\widehat{X_jh}-\hat{X}_j\widehat{X_ih}+2\hat{X}_i\\&\hat{X}_j\hat{h})^T\ [R^{-1}(dZ-\hat{h}dt)]\end{aligned} \quad (19)$$

Where X_i is the i-th element of $X(t)$ f_i is the i-th row of $f(\ )$.

$(\ )\ ij$ is the ij-th element of an n x n matrix.

As we mentioned before, mean and variance eq. (18) and (19) both involve conditional expectation $\hat{h}(t)$. This means the optimum of coupled moment equation to be solved. Which also mean such a filter dimension need infinite dimension to be implemented. Fortunately for certain class of non-linear stochastic system with Gaussian observation noise, the estimator is recursive and finite dimensional[5].

Very special case of non-linear system is linear system where the system and observation are both linear and the noises are Gaussian. In this linear case, Kalman filter algorithm provides a recursive optimal finite-dimensional estimator as will see in the following.

Since so good estimators are developed by many authors, we just mention about the main results of Kalman filter algorithm.

2) Stochastic Linear System Estimation [6]

Consider the linear stochastic system

$$dX(t)=F(t)X(t)dt+G(t)dw(t) \quad (20)$$

with noise currupted linear observation

$$dZ(t)=H(t)X(t)\ dt+dv(t) \quad (21)$$

where $w(t)$ and $v(t)$ are zero mean white noise

$$\text{cov}[w(t),\ w(\tau)]=Q(t)\ \delta(t-\tau)$$

$$\text{cov}[v(t),\ v(\tau)]=R(t)\ \delta(t-\tau)$$

$X(t)$, $w(t)$ and $v(t)$ are independent each other

$$E[X(t_0)]=\hat{X}(t_0),\ \text{Var}[X(t_0)]=P_0$$

Then the sequential linear estimator (Kalman filter) becomes

$$\dot{\hat{X}}(t)=F(t)\hat{X}(t)+K(t)[Z(t)-H(t)\hat{X}(t)] \quad (22)$$

Where $K(t)$ is Kalman filter gain and quantity

$$Z(t)-H(t)\hat{X}(t)$$

is a white noise stochastic process called "the innovation process" containing all of the new information in the observation $Z(t)$.

When $K(t)$ is selected so as to mininize the error variance i.e. minimize the cost function $J(t)$,

$$J(t)=tr[\text{var}\{\tilde{X}(t)\}]=tr[P(t)] \quad (23)$$

where $\tilde{X}(t)=X(t)-\hat{X}(t)$

then $K(t)$ becomes, so called, Kalman filter gain expressed as

$$\hat{K}(t)=PH^TR^{-1} \quad (24)$$

where error variance $P(t)$ satisfy the following Riccati type differential equation

$$P(t)=FP+PF^T+GQG^T-PH^TR^{-1}HP$$

$$P(t_0)=P_0 \quad (25)$$

3) State Estimation of a Stochastic Bilinear System

① Existance of the finite dimensional estimator

As far as state estimation of the SBLS problems are concerned, we can derive a more basic framework on the existance of moments and structure of the estimator using the Lie-algebra theory [7, 8].

Consider the next stochastic system and observation

$$d\zeta(t)=F\zeta(t)dt+Q^{1/2}dw(t)$$

$$dZ(t)=H\zeta(t)dt+R^{1/2}dv(t) \quad (26)$$

$$t\in[t_0,\ t_f]$$

Where $w(t)$, $v(t)$ are Wiener process with n and q dimensional. $\zeta(t_0)$ is Gaussian and $\zeta(t_0)$, w, v are mutual independent. $(F,\ Q^{\frac{1}{2}},\ H)$ is completely controllable and completely observable pair.

$\zeta(t)\in R^n$, $Z(t)\in R^q$ and Q and R are positive definite.

Here we are interested in estimation of signal process $\{X(t);\ X(t)\in R^m\}$ as following,

$$\dot{X}(t)=\Big[A+\sum_{i=1}^{n}B_i\zeta_i(t)\Big]X(t)$$

$$X(0)=I \quad (27)$$

where A, $B_1\cdots B_n$ are mxm constant matrix and $X(0)$ is independent of $\zeta(t_0)$, $w(t)$, $v(t)$.

Then the existance of the finite dimensional estimator problem can be described as following

Theorem 1.

Let a denotes the Lie-algebra generated by

$\{A,\ B_1 \cdots\cdots B_n\}$.

and $\bar{a}_0$ is the ideal in a generated by $\{B_1 \cdots\cdots B_n\}$.

If $\bar{a}_0$ is a nilpotent Lie-algebra with dimension n^*, then the conditional expectation

$$\hat{X}(t)=E[X(t)\,|\,Y\tau] \qquad (28)$$

can be estimated with a finite dimensional system of bilinear equation driven by the innovation having the form

$$d\hat{X}(t)=A\hat{X}(t)dt+K(t)dv(t) \qquad (29)$$

In particular if $A,\ B_1 \cdots\cdots B_n$ are all strictly upper triangular, then $\bar{a}_0$ and $\bar{a}$ are nilpotent Lie-algebra. If $\bar{a}_0$ is not nilpotent then the optimal estimation $X(t)$ is infinite dimensional.

Proof:

Proof of this theorem can be found in [8].

Instead of proving the theorem, several properties can be derived as following.

② **Structure and the properties of the estimator**

The least square filtered estimate (28) under the above existance assumption can be obtained from the finite dimensional bilinear stochastic differential equation of the following from

$$d\hat{X}^*(t)=\Big[A^*dt+\sum_{i=1}^{n}B_i^*\ \hat{\tau}_i(t)+\sum_{i=1}^{n}C_i^*\ dU_i(t)\Big]$$
$$X^*(t) \qquad (30)$$

where $\hat{X}^*(t)\in R^m,\ m^*\leq M\Big(\frac{(n^2n^*)^m-1}{n^2n^*-1}\Big)$

$$\hat{X}_i^*(0)=\begin{cases}E[X_i(0)], & \text{for } i\leq m\\ 0, & \text{for } i>m\end{cases}$$

$\hat{X}(t)=L(t)\hat{X}^*(t)$, L is $m\times m^*$, A^*, B_i^*, C_i^* are $m^*\times m^*$, and where * means the nilpotency dimension, U is modified innovation process.

We see that the optimal filter structure (30) is comparable to that of the model (27) and i) It is bilinear in both drift and diffusion terms and ii) It also possesses a nilpotency properties.

This results are analogeous to that of linear filtering problem in which a linear system model gives rise to the optimal filter which is also linear in both the drift and diffusion terms.

③ **A sub-optimal estimation of SBLS**

In this paper we will introduce a sub-optimal estimator which can be directly derived from the results of non-linear estimation which we discussed in section 2.

Let us rewrite the general form of SBLS (10) as

$$dX(t)=\Big[(A+\sum_{i=1}^{m}B_iU_i)X+CU\Big]dt+\sum_{i=1}^{m}B_iXdw_i$$
$$+Cdw \qquad (31)$$

and non-linear observation

$$dZ(t)=h(X,\ t)dt+dv(t) \qquad (32)$$

The state estimation is completed if conditional probability density function $P(X,\ t\,|\,Y\tau)$ can be found satisfying Fokker-planck eq (14). For this purpose the first two moments, mean and covariance functions of $X(t)$ are enough since solution of the partial differential equation of (14) is rarely possible to obtain as indicated. Let us define

$$F(t)=A+\sum_{i=1}^{m}B_iU_i \qquad (33)$$

$$G(t)=\sum_{i=1}^{m}B_iX+C \qquad (34)$$

then (31) reduces to

$$dX(t)=[FX+CU[dt+\sum_{i=1}^{m}B_iXdw_i+Cdw$$
$$=[FX+CU]dt+G(X,\ t)dw \qquad (35)$$

For this system we can apply the same procedures as the general nonlinear system where mean (18) and covariance (19) are derived from (12). The corresponding mean and covariance equation for the SBLS (35) and observation (32) become

$$d\hat{X}(t)=[F\hat{X}+CU]dt+(\widehat{Xh^T}-\hat{X}\hat{h}^T)R^{-1}(dZ$$
$$-\hat{h}dt) \qquad (36)$$

$$dP_{ij}=[(FP+PF^T+\widehat{GQG^T})_{ij}-(\widehat{X_ih}-\hat{X}_i\hat{h})R^{-1}$$
$$(\widehat{hX_j}-\hat{h}\hat{X}_j)]+(\widehat{X_iX_jh}-\hat{X}_i\widehat{X_jh}-\hat{X}_j\widehat{X_ih}$$
$$-\widehat{X_iX_j}\hat{h}+2\hat{X}_i\hat{X}_j\hat{h})^TR^{-1}(dZ-\hat{h}dt) \qquad (37)$$

Since $G(X,\ t)$ is linear in $X(t)$ from (34) GQG^T will only have up to 2nd order in $X(t)$.

When observation is assumed linear function of $X(t)$ i.e.

$$dZ(t)=HX(t)dt+dv(t) \qquad (38)$$

then mean eq (36) reduces to

$$d\hat{X}=[F\hat{X}+CU]dt+PH^TR^{-1}[dZ-H\hat{X}]dt \qquad (39)$$

$$dP_{ij}=[FP+PF^T+\widehat{GQG^T}-PH^TR^{-1}HP]dt$$
$$+P_{3ijl}H^TR^{-1}[dZ-H\hat{X}]dt \qquad (40)$$

where $P_{3ijl}=E[(X_i-\hat{X}_i)\ (X_j-\hat{X}_j)\ (X_l-\hat{X}_l)]$

with $P_{3ij}=\begin{bmatrix}P_{3ij_1}\\ \vdots\\ P_{3ij_n}\end{bmatrix}$, the 3rd central moment.

Note that optimal filter is dependent on the third order moment. In general r-th order moment for the SBLS requirs $(r+1)$-th order and the lower order moments for the optimal estimation. It means that the optimal estimation for SBLS is not available. This is why we must rely on the sub-optimal filter to estimate the state of this system.

One logical approximation to obtain finite dimensional estimation for this SBLS is to assume that for some r

$$P_{r+1}(t)=0 \tag{42}$$

For proper choice of r we can compromise between the computational complexity and the estimator requirment. If we choose $r=2$, then

$$P_3(t)=0 \tag{43}$$

Thus the following second order sub-optimal filter is derived from (39) and (40).

$$d\hat{X}(t)=[F\hat{X}+CU]dt+PH^TR^{-1}[dZ-H\hat{X}dt] \tag{44}$$

$$\dot{P}(t)=FP+PF^T+GQG^T-PH^TR^{-1}HP \tag{45}$$

This is final resultant estimation of the SBLS given (31) and (38). Several note can be derived from this results.

Note:

i) Moment eq (45) is no longer stochastic eq.

It is same type of matrix Riccati equation as linear Kalman filter (25), but term GQG^T and $F(t)$ are different here.

ii) Filter eq (44) is the same as Kalman filter (21), but $F(t)$ is given by (33) here.

iii) Truncated second order or Gaussion second order filter are indentical in this approximation.

State estimation problem of BLS or SBLS is dealt with by several authors for examples Lo [9], Marcas, Hsu and Willsky [10], Funahashi [11], Jurdjevic & Sussman [12] etc. Many are dealing with estimation problem based on the Lie-algebra or Lie-group theory.

Parameters identification is another topic which must be dealt with when we consider state estimation problem. But we will omit here. [13], [1], [14], [15] are all good references for this topic. Specially [16] used special function (Walsh function) method to identify BLS input, output functions.

3. Conclusion

Bilinear system is a simple class of non-linear system as well as it is the closet to the linear system. Many natural phenonenon can be modeled by this BLS. Stochastic BLS isformulated when noise is considered.

The existance of finite dimensional optimal estimator using Lie algebra is guaranteed when system is nilpotent. But in most case this condition is not satisfied. This means infinite dimensions are required to estimate optimally. Thus sub-optimal estimator is next feasible choice.

One sub-optimal estimator is introduced in this paper. The final resultant eq (44), (45) show that covariance equation is usual Riccati type equation and filter function is the same form as Kalman filter.

Reference

[1] A.S. Valavi; "On the state and parameter estimation of stochastic bilinear system", Ph. D. Thesis, Oregon State Univ. June 1978.

[2] C. Bruni, G. Dipillo, G. Koch; "Bilinear system; An appealing class of nearly linear systems in theory and application" IEEE. AC-19, No. 4, Aug. 1974.

[3] K.C. Wei; "Optimal control of bilinear systems with some aerospace applications", Ph. D. Thesis, Brown University, RI, June 1976.

[4] M.M.R. Williams; "Noise characteristics of nuclear reactors", Pergamon press, 10523 N.Y. 1974.

[5] S.I. Marcus; "Optimal non-linear estimation for a class of discrete linear stochastic systems", IEEE, AC-24, No. 2, Apr. 1979.

[6] A.P. Sage, C.C. White; "Optimum Systems Control", University of Virginia, 2nd ed. 1976.

[7] R.R. Mohler, W.J. Kolodziej, "An overview of stochastic bilinear system control processes", Dept. ofEE. Oregon State Univ. Or. 97331, 1977.

[8] Shirish D, Chikte, J. Ting-Ho Lo: "Optimal

filter for bilinear system with Nilpotent Lie-algebra", IEEE. AC-24, No. 6, Dec. 1979.

[9] James Ting-Ho Lo; "Signal detedtion for bilinear systems", Infor. Sciences. 9. p.249~278, '1975.

[10] S.I. Marcus, A.S. Willsky, K,Hsu; "The use of harmonic analysis in sub-optimal estimator design", IEEE. AC-23, No. 5, Oct. 1978.

[11] Y. Funahashi; "Stable state estimation for bilinear systems", Int. J. Control V 29, No. 2, pp.181~188, 1979.

[12] V. Jurdjevic, H.J. Sussman; "Control systems on Lie-groups", J. Differential Eq. 12(2), 1972.

[13] R.S. Baheti, R.R. Mohler, H.A. Spang; "A new cross correlation algorithms", IEEE. AC-24, No. 4, Aug. 1979.

[14] A.V. Barakrishnan; "Modeling and identification theory a flight control applications", Theory and App. of Variable Structure Sys. R.R. Mohler Ed. Academic Press, 1972.

[15] C. Bruni, G. Dipillo, G. Koch; "Math. models and identification of bilinear systems",

[16] V.R. Karanam, P.A. Frick, R.R. Mohler, "Bilinear system identification by Walsh function", IEEE, AC-23, No. 4, Aug. 1978.