Systems & Control Letters 119 (2018) 92–100

Contents lists available at ScienceDirect

Systems & Control Letters

journal homepage: www.elsevier.com/locate/sysconle

Suboptimal linear estimation for continuous–discrete bilinear systems

Xue Luo [a], Xiuqiong Chen [b,1], Stephen S.-T. Yau [b,*]

[a] School of Mathematics and Systems Science, Beihang University, Beijing 100191, China
[b] Department of Mathematical Sciences, Tsinghua University, Beijing, 100084, China

ARTICLE INFO

Article history:
Received 26 September 2017
Received in revised form 19 June 2018
Accepted 16 July 2018

Dedicated to Professor T. Duncan on the occasion of his 75th birthday.

Keywords:
Nonlinear filtering
Bilinear systems
Carleman approach
The extended Kalman filter

ABSTRACT

In this paper we derive a suboptimal estimation for continuous–discrete bilinear systems. One of the motivations of this work is that the bilinear system has the simplest structure in the nonlinear class in some sense. Similar to the Kalman filter, our algorithm includes prediction and updating step. We show rigorously that our algorithm gives an unbiased estimate, the a-priori estimate approaches to the conditional expectation exponentially fast, and the posterior estimate minimizes the conditional variance error in the linear space spanned by the a-priori estimate and the innovation. Our algorithm is also applicable to solve the nonlinear filtering problems. The efficiency of our method is illustrated by the cubic sensor problem and Lorenz system with discrete observation. The results have been compared with the extended Kalman filter and the unscented Kalman filter.

1. Introduction

How to get the instantaneous and accurate estimation of the states of a stochastic system from the polluted measurements by the noise is of central importance in engineering and this is also the central problem in the field of filtering. A continuous–discrete filtering problem is modeled by the following Itô stochastic differential equation:

$$\begin{cases} dx(t) = f(x(t))dt + g(x(t))dv(t) \\ y(t_k) = h(x(t_k)) + w(t_k), \end{cases} \tag{1.1}$$

where $v(t)$ is Brownian motion with proper dimension, $x(t) \in \mathbb{R}^n$ is the state, $0 = t_0 < t_1 < \cdots < t_K = T$, $y(t_k) \in \mathbb{R}^m$ is the measurement, t_k, $k = 1, 2, \ldots, K$ are instants when the measurements arrive and $w(t_k) \in \mathbb{R}^m$ is white noise. When the function $f(x)$ and $h(x)$ are linear functions of x and $g(x)$ is constant, we call (1.1) a linear filtering problem and its study can be traced back to early 1960s when Kalman [1], Kalman and Bucy [2] published two most influential papers and proposed the classical Kalman filter and Kalman–Bucy filter. We refer the readers to the book [3] for excellent introduction to filtering theory. Though the linear filtering problem is completely solved in [1,2,4], the nonlinear filtering (NLF) problems are much more complicated and important in applications since most practical models are nonlinear.

One class of methods to solve NLF problems is the so-called global approaches which try to find out the conditional density function of the states by solving the Duncan–Mortensen–Zakai (DMZ) equation [5–7]. Based on the DMZ equation, more research articles follow this direction such as [8–13]. Numerical methods to solve this problem can also be found such as in [14].

Another class of methods to solve the NLF problems is referred as local approaches, which construct suboptimal filters in some sense. There are many approximate methods including unscented Kalman filter (UKF) [15,16], ensemble Kalman filter [17], particle filter [18] and the most widely used extended Kalman filter (EKF) [3,19], which is basically the Kalman–Bucy filter applied to a linearized system. However, EKF can only perform well if the initial estimation error and the disturbing noises are small enough due to its local nature.

Continuous–discrete filter, which is for stochastic differential systems with sampled measurements, is also of great significance and has many applications such as in tracking and finance since the measurements always come in discretely. There has been increasing interest in this system and many continuous–discrete filters can be found in the literatures, such as continuous–discrete EKF [3], continuous–discrete UKF [15], continuous–discrete Gaussian filter [20] and continuous–discrete cubature Kalman filter [21]. And the comparison of these different methods can refer [22].

Our motivations to study the bilinear system are two folds: on the one hand, many important processes, not only in engineering but also in socio-economics, biology and ecology, may be modeled by bilinear systems [23]. On the other hand, the bilinear structure

* Corresponding author.
E-mail addresses: xluo@buaa.edu.cn (X. Luo), cxq14@mails.tsinghua.edu.cn (X. Chen), yau@uic.edu (S.S.-T. Yau).
[1] X. Chen is the co-first author.

https://doi.org/10.1016/j.sysconle.2018.07.005
0167-6911/

seems to be the simplest and closest one to the linear one among all the nonlinearities. Thus, some well-established techniques can be extended to bilinear systems [24]. The estimation theory for bilinear system can also be used to solve NLF problems. For example, the nonlinear analytical system can be approximated by a bilinear system using Carleman approach [25].

Notice that [24] only deals with the continuous–continuous systems. Recently, the first and the last author of this paper [26,27] proposed a novel algorithm for solving the continuous NLF problems based on the idea in [24]. In this paper we derive a suboptimal filter for the continuous–discrete bilinear systems. Compared with the work of Cacace and his collaborators, we consider the filter rather than state predictor [28] and the bilinear system (3.1) in our algorithm is more general than that in [29]. We call the estimate obtained in this paper suboptimal linear estimate (SLE). Similar to EKF, our algorithm consists of two steps including predicting and updating. We call the estimate after prediction the a-priori estimate, while that after updating the posterior estimate. The suboptimality of our algorithm in the following sense: essentially, we show that under some mild conditions SLE has the following properties:

1. Both the a-priori and the posterior estimates are unbiased;
2. The a-priori estimate approaches to the conditional expectation exponentially fast;
3. The posterior estimate minimizes the conditional variance error in a linear space.

This paper is organized as follows. Our algorithm is described in Section 2.1. The suboptimality of SLE has been shown rigorously in Section 2.2. Section 3 presents the application of our algorithm to representative NLF problems, where we compare the performance of the proposed filter with EKF and UKF. We arrive at the conclusion in Section 4.

2. Suboptimal algorithm

The bilinear continuous–discrete system considered in probability space $(\Omega, \mathscr{F}, P)$ is as follows:

$$\begin{cases} dX(t) = \mathbf{A}X(t)dt + \mathbf{N}dt + \sum_{j=1}^{b}(\mathbf{B}_jX(t) + \mathbf{F}_j)dW_j(t), \; t \in [0, T], \\ Y(t_k) = \mathbf{C}X(t_k) + \mathbf{D} + \sum_{j=1}^{b}\mathbf{G}_jV_j(t_k), \; k = 0, 1, \ldots, K, \end{cases} \tag{2.1}$$

where $0 = t_0 < t_1 < \cdots < t_K = T, \mathbf{A} \in \mathbb{R}^{n\times n}, \mathbf{N} \in \mathbb{R}^{n\times 1}, \mathbf{B}_j \in \mathbb{R}^{n\times n}, \mathbf{F}_j \in \mathbb{R}^{n\times 1}, \mathbf{C} \in \mathbb{R}^{m\times n}, \mathbf{D} \in \mathbb{R}^{m\times 1}, \mathbf{G} \in \mathbb{R}^{m\times 1}$ are constant matrices. $X(t) \in \mathbb{R}^n$ is the state with the initial value X_0 whose mean is $\bar{X}_0$ and covariance matrix is $\bar{P}_0$, $Y(t_k) \in \mathbb{R}^m$ is discrete measurement, $V_j(t_k) \sim \mathcal{N}(0, R_j(t_k)), R_j(t_k) \in \mathbb{R}$, $k = 0, 1, \ldots, K$, are independent one-dimensional white noises and $W_j(t), j = 1, \ldots, b$, are independent standard Brownian motions. Let $\mathscr{F}_{t_k}$ be the σ-field generated by the observations, i.e. $\mathscr{F}_{t_k} \triangleq \sigma\{Y(t_0), Y(t_1), \ldots, Y(t_k)\}$. Kronecker algebra is used for concise notation and derivation. Its properties can be found in [30].

Recall that the probability space $(\Omega, \mathscr{F}, \mathbb{P})$ with finite second moment, with scalar product $\langle x, y\rangle = E[x^Ty]$ and norm $\|x\| := E^{1/2}[x^Tx]$ is a Hilbert space, denoted as $L^2(\Omega, \mathscr{F}, \mathbb{P})$. Suppose that the state $X(t) \in L^2(\Omega, \mathscr{F}, \mathbb{P})$. In our algorithm, we shall obtain a linear recursive estimate in the similar fashion of EKF. As explained in [28], the prediction of the state on the observation history is indeed a random variable. After the approximation of the conditional expectation of the nonlinear drift term $f(x_t)$ coarsely in EKF, i.e. $E[f(x_t)|\mathscr{F}_{t_{k-1}}] \approx f(E[x_t|\mathscr{F}_{t_{k-1}}])$, for $t > t_{k-1}$, it makes the state estimate satisfy an deterministic ordinary differential equation. This is the essential reason why all the estimates in our algorithm will be treated in a deterministic way. Let us clearly define the linear recursive estimate of $X(t)$ based on the observation history $\{Y(t_0), Y(t_1), \ldots, Y(t_{k-1})\}$ first:

Definition 2.1. We call $\hat{X}(t_k|t_k)$ the linear recursive estimate of $X(t_k)$ based on the observation $\{Y(t_0), Y(t_1), \ldots, Y(t_k)\}$, if

1. The a-priori estimate, denoted as $\hat{X}(t|t_{k-1})$, $t \in [t_{k-1}, t_k]$, is linearly dependent of the previous posterior estimate $\hat{X}(t_{k-1}|t_{k-1})$, i.e.

$$\hat{X}(t|t_{k-1}) = H_1(t)\hat{X}(t_{k-1}|t_{k-1}) + H_2(t), \tag{2.2}$$

where H_1 and H_2 are matrices of proper dimensions;

2. The posterior estimate $\hat{X}(t_k|t_k)$ lives in the linear space spanned by 1, the a-priori linear estimate $\hat{X}(t_k|t_{k-1})$ and the innovation $Y(t_k) - \hat{Y}(t_k|t_{k-1})$, where $\hat{Y}(t_k|t_{k-1}) = C\hat{X}(t_k|t_{k-1}) + D$. That is,

$$\hat{X}(t_k|t_k) = H_3\hat{X}(t_k|t_{k-1}) + H_4(Y(t_k) - \hat{Y}(t_k|t_{k-1})) + H_5, \tag{2.3}$$

where H_3, H_4 and H_5 are constant matrices of proper dimensions.

2.1. Algorithm

Our algorithm consists of two steps: prediction and updating. Throughout the process, we assume that

(As) $\mathbf{A}$ and $\mathbf{A}_{ex}$ are Hurwitz, where $\mathbf{A}_{ex} := \sum_{l=1}^{b}(\mathbf{B}_l \otimes \mathbf{B}_l) + I_n \otimes \mathbf{A} + \mathbf{A} \otimes I_n$.

We state our algorithm first:

(Al-1) Prediction In the interval $[t_{k-1}, t_k)$, the a-priori estimate $\hat{X}(t|t_{k-1})$ of $X(t)$ based on data $\{Y(t_0), Y(t_1), \ldots, Y(t_{k-1})\}$ satisfies

$$\dot{\hat{X}}(t|t_{k-1}) = \mathbf{A}\hat{X}(t|t_{k-1}) + \mathbf{N}, \tag{2.4}$$

$$\dot{Q}(t|t_{k-1}) = \mathbf{A}Q(t|t_{k-1}) + Q(t|t_{k-1})\mathbf{A}^T + \sum_{j=1}^{b}\Big[\mathbf{B}_jQ(t|t_{k-1})\mathbf{B}_j^T + \Big(\mathbf{B}_j\hat{X}(t|t_{k-1}) + \mathbf{F}_j\Big) \times \Big(\mathbf{B}_j\hat{X}(t|t_{k-1}) + \mathbf{F}_j\Big)^T\Big], \tag{2.5}$$

with the initial value $\hat{X}(t_{k-1}|t_{k-1})$ and $Q(t_{k-1}|t_{k-1})$ from previous updating, $\hat{X}(t_0|t_0) := \bar{X}_0$, and $Q(t_0|t_0) := \bar{P}_{t_0}$.

(Al-2) Updating The posterior estimate $\hat{X}(t_k|t_k)$ of $X(t_k)$ based on the observation history $\mathscr{F}_{t_k}$ satisfies

$$\hat{X}(t_k|t_k) = \hat{X}(t_k|t_{k-1}) + K_k\Big[Y(t_k) - \hat{Y}(t_k|t_{k-1})\Big], \tag{2.6}$$

with $\hat{Y}(t_k|t_{k-1}) = \mathbf{C}\hat{X}(t_k|t_{k-1}) + \mathbf{D}$, and the gain function K_k is given by

$$K_k = Q(t_k|t_{k-1})\mathbf{C}^T \times \left[\mathbf{C}Q(t_k|t_{k-1})\mathbf{C}^T + \sum_{j=1}^{b}\mathbf{G}_jR_j(t_k)(\mathbf{G}_j)^T\right]^{-1}. \tag{2.7}$$

Meanwhile, the matrix $Q(t_k|t_k)$ is updated by

$$Q(t_k|t_k) = (I_n - K_k\mathbf{C})\,Q(t_k|t_{k-1}). \tag{2.8}$$

where I_n is the identity matrix of dimension $n \times n$.