World Academy of Science, Engineering and Technology
International Journal of Mathematical and Computational Sciences
Vol:7, No:12, 2013

Kalman Filter for Bilinear Systems with Application

Abdullah E. Al-Mazrooei

Abstract—In this paper, we present a new kind of the bilinear systems in the form of state space model. The evolution of this system depends on the product of state vector by its self. The well known Lotak Volterra and Lorenz models are special cases of this new model. We also present here a generalization of Kalman filter which is suitable to work with the new bilinear model. An application to real measurements is introduced to illustrate the efficiency of the proposed algorithm.

Keywords—Bilinear systems, state space model, Kalman filter.

I. INTRODUCTION

BILINEAR systems are a very important class of nonlinear systems. It appears in many applications in science, engineering, economic and control system. It have used recently in weather forecasting, atmospheric studies and ocean studies [5],[6]. In recent years, the bilinear systems captured the attention in the research and application. Because, they are applicable in various aspects of real life. They also provide more flexible approximations to nonlinear systems than do linear systems. Moreover, bilinear systems have rich geometric and algebraic structures that create a fruitful field of research.
The bilinear systems were introduced in control theory in 1960's. Yet, their nonlinearity is subject to the product between the state vector and the input of the systems[6]. The kind of these systems are easier to deal, because they are reduced to linear systems according to use of a certain of Kronecker product. In this paper, we present a new bilinear model, the nonlinearity is subject to the product of the state of the system by its self. This technique allows for this model to generalize the Lotka-Volterra models and Lorenz models which they have applications in science and weather prediction [3],[4]. This means that, this bilinear model applicable more widely. Kalman filter is the optimal estimator that is used to estimate the state in linear systems. It was introduced by Kalman in 1970 [2]. Since, the system is nonlinear, then the classical Kalman filter is not suitable to be used, because it works with linear systems only. Here, we have a nonlinear model of a bilinear class, thus, Kalman filter does not work with our model. Therefore, we need to develop Kalman filter to work with the new bilinear model. The direct development of the recursions for the nonlinear filters is very complicated if not impossible altogether[1],[7],[8]. Instead, we develop our recursions based on a linearization of the quadratic term that uses the most current state estimate available.

The paper is organized as follows: In Section II, the new bilinear model is presented. In Section III, we describe a new generalization of Kalman filter algorithm. In section IV, a new application to real data is introduced to illustrate the efficiency of the proposed algorithm.

Abdullah E. Al-Mazrooei is with the Department of Mathematics, King Abdulaziz University, P.O. Box 80203, Jeddah, 21589, Saudi Arabia. e-mail: aealmazrooei@kau.edu.sa.

II. THE NEW BILINEAR MODEL

In this section, we present a new sort of the bilinear system. The nonlinearity of this model consists of bilinear interaction between the states of the model themselves. While, the nonlinearity of the earlier bilinear systems is interaction between the state of the systems and the system input. The well- known Lotka-Volterra models and Lorenz models are considered to be special cases of the new bilinear model. Here, considering a bilinear model in discrete state space form Which is given as follows

$$x_{k+1} = Ax_k + B(x_k \otimes x_k) + w_k; \tag{1}$$

$$y_k = Cx_k + v_k; \tag{2}$$

where;
$x_k \in \mathbb{R}^n$ is the state vector,
$y_k \in \mathbb{R}^p$ is the measurements vector,
while the matrices $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times \frac{n(n+1)}{2}}$ and $C \in \mathbb{R}^{p\times n}$ are the parameters of the model.
The noise corruption signals $w_k \in \mathbb{R}^n$ and $v_k \in \mathbb{R}^p$ are white, uncorrelated and Gaussian with zero mean and covariances Q and R respectively. That is

$$w_k \sim N(0, Q)$$

$$v_k \sim N(0, R)$$

Also,

$$E(w_k w_l^T) = \begin{cases} Q, \text{for} & \text{k=l} \\ 0, \text{for} & \text{k}\neq\text{l}, \end{cases}$$

$$E(v_k v_l^T) = \begin{cases} R, \text{for} & \text{k=l} \\ 0, \text{for} & \text{k}\neq\text{l}, \end{cases}$$

and

$$E(w_k v_k^T) = 0.$$

Here, $x_k \otimes x_k$ represents the Kronecker product of the state x_k with itself without repetition of the entries.

Note, for example, Lorenz model-69 with $n = 3$ can be written in the form of our bilinear model with

$$A = -I_3$$

and

$$B = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Now by using Taylor polynomial expansion for the bilinear term $z_k = x_k \otimes x_k$ at x_0 we get,

$$z_k = z_0 + z'(x_0)(x_k - x_0) + \frac{1}{2}\mathcal{H}(x_k, x_0)(x_k - x_0) \quad (3)$$

where $z'(x)$ is the $\frac{n(n+1)}{2} \times n$ gradient of $z(x)$ given by

$$z'(x) = [\frac{\partial x_i x_j}{\partial x_l}]_{i,j,l=1,2,\ldots,m},$$

and $\mathcal{H}(x_k, x_0)$ is given by

$$\mathcal{H}(x_k, x_0) = \begin{bmatrix} (x_k - x_0)^T D_1 \\ (x_k - x_0)^T D_2 \\ \vdots \\ (x_k - x_0)^T D_m \end{bmatrix}$$

and, $D_1, D_2, ..., D_m$ is the matrices of second derivatives of the entries of z_k, and $m = \frac{n}{n+1}$

To illustrate,suppose $n = 3$. Then

$$z(x) = (x_1^2, x_1x_2, x_1x_3, x_2^2, x_2x_3, x_3^2)^T,$$

$$z'(x) = \begin{bmatrix} 2x_1 & 0 & 0 \\ x_2 & x_1 & 0 \\ x_3 & 0 & x_1 \\ 0 & 2x_2 & 0 \\ 0 & x_3 & x_2 \\ 0 & 0 & 2x_3 \end{bmatrix}$$

and

$$D_1 = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, D_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$
$$D_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, D_4 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$
$$D_5 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, D_6 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

III. A GENERALIZATION OF KALMAN FILTER

In this work, we have a nonlinear system, and thus using the classical Kalman filter is not possible in this case since it is appropriate only for linear systems. Thus, we will derive a development of Kalman filter for the bilinear system (1) and (2).

We will adopt the following notations:

$$\begin{aligned} z_k &= x_k \otimes x_k \\ x_k^t &= E\{x_k|\{y\}_1^t\} = E_t(x_t) \\ P_k^t &= E_t\{(x_k - x_k^t)(x_k - x_k^t)^T\} \\ z_k^t &= E\{z_k|\{y\}_1^t\} = E_t(z_t) \\ \dot{P}_k^t &= E_t\{(z_k - z_k^t)(z_k - z_k^t)^T\} \\ \ddot{P}_k^t &= E_t\{(x_k - x_k^t)(z_k - z_k^t)^T\}, \end{aligned}$$

where,

$$1 \leq k \leq t$$

$$1 \leq t \leq n$$

and, $\{y\}_1^t$ is the measurements sequence

$$\{y\}_1^t = \{y_1, ..., y_t\}.$$

Now to compute equation(3), we approximate the second degree term $\mathcal{H}(x_k, x_0)$ by using the most current available state estimation of x_k. So we have two different cases.

- In the case of prediction, we take,
$$(x_k - x_0) \approx (x_k^{k-2} - x_0),$$
with $x_0 = x_k^{k-1}$ and $x_1^{-1} = 0$. This means that,
$$z_k \approx z_k^{k-1} + z'(x_k^{k-1})(x_k - x_k^{k-1})$$
$$+\frac{1}{2}\mathcal{H}(x_k^{k-2}, x_k^{k-1})(x_k - x_k^{k-1}),$$
- In the case of filtering, we take,
$$(x_k - x_0) \approx (x_k^{k-1} - x_0).$$
with $x_0 = x_k^k$, this means that,
$$z_k \approx z_k^k + z'(x_k^k)(x_k - x_k^k) + \frac{1}{2}\mathcal{H}(x_k^{k-1}, x_k^k)(x_k - x_k^k),$$

Thus, these cases can be summarized in the following linearization:

$$\begin{aligned} z_k &\approx z_k^j + z'(x_k^j)(x_k - x_k^j) + \frac{1}{2}\mathcal{H}(x_k^{j-1}, x_k^j)(x_k - x_k^j) \\ &= z_k^j + V_k^j(x_k - x_k^j), \end{aligned} \quad (4)$$

where,

$$V_k^j = z'(x_k^j) + \frac{1}{2}\mathcal{H}(x_k^{j-1}, x_k^j). \quad (5)$$

In the next theorem, we introduce a bilinear Kalman filter algorithm.

Theorem. For the bilinear state-space model defined by (1) and (2), we have

$$x_{k+1}^k = Ax_k^k + Bz_k^k \quad (6)$$
$$P_{k+1}^k = AP_k^kA^T + A\ddot{P}_k^kB^T + B(\ddot{P}_k^k)^TA^T + B\dot{P}_k^kB^T + Q. \quad (7)$$

with,

$$x_{k+1}^{k+1} = x_{k+1}^{k} + K_{k+1}[y_k - Cx_{k+1}^{k}] \quad (8)$$
$$P_{k+1}^{k+1} = [I - K_{k+1}C]P_{k+1}^{k} \quad (9)$$
$$\ddot{P}_{k+1}^{k+1} = P_{k+1}^{k+1}[V_{k+1}^{k+1}]^T \quad (10)$$
$$\dot{P}_{k+1}^{k+1} = (V_{k+1}^{k+1})\ddot{P}_{k+1}^{k+1} \quad (11)$$

where,

$$K_{k+1} = P_{k+1}^{k}C^T[CP_{k+1}^{k}C^T + R]^{-1} \quad (12)$$

and,

$$V_{k+1}^{k+1} = z'(x_{k+1}^{k+1}) + \frac{1}{2}\mathcal{H}(x_{k+1}^{k}, x_{k+1}^{k+1}), \quad (13)$$

for $k = 0, ..., t$.

Proof. First, to derive the forecast steps which are given by (5) and (6), we consider the case $t < k$ in the previous notations. By applying the conditional expectation to (1):

$$\begin{aligned} x_{k+1}^{k} &= E_k(x_{k+1}) \\ &= E_k(Ax_k + Bz_k + w_k) \\ &= AE_k(x_k) + BE_k(z_k) + E_k(w_k) \\ &= Ax_k^k + Bz_k^k. \end{aligned}$$

To obtain the error recursion (6), we proceed as follows

$$\begin{aligned} P_{k+1}^{k} &= E_k\{(x_{k+1} - x_{k+1}^{k})(x_{k+1} - x_{k+1}^{k})^T\} \\ &= E_k\{(Ax_k + Bz_k + w_k - Ax_k^k - Bz_k^k) \\ &(Ax_k + Bz_k + w_k - Ax_k^k - Bz_k^k)^T\} \\ &= E_k\{[A(x_k - x_k^k) + B(z_k - z_k^k) + w_k] \\ &[A(x_k - x_k^k) + B(z_k - z_k^k) + w_k]^T\} \\ &= E_k\{[A(x_k - x_k^k) + B(z_k - z_k^k) + w_k] \\ &[(x_k - x_k^k)^T A^T + (z_k - z_k^k)^T B^T + w_k^T]\} \\ &= E_k\{A(x_k - x_k^k)(x_k - x_k^k)^T A^T + A(x_k - x_k^k) \\ &(z_k - z_k^k)^T B^T + B(z_k - z_k^k)(x_k - x_k^k)^T A^T \\ &+ B(z_k - z_k^k)(z_k - z_k^k)^T B^T + w_k w_k^T\} \\ &= AP_k^k A^T + A\ddot{P}_k^k B^T + B(\ddot{P}_k^k)^T A^T + B\dot{P}_k^k B^T \\ &+ Q, \end{aligned}$$

Second, when $t = k$, we derive the filtering steps. Let

$$\begin{aligned} \rho_k &= y_k - E_{k-1}(y_k) \\ &= y_k - E_{k-1}(Cx_k - v_k) \\ &= y_k - Cx_k^{k-1} \\ &= Cx_k - Cx_k^{k-1} + v_k \\ &= C(x_k - x_k^{k-1}) + v_k, \end{aligned}$$

for $k = 1, ..., t$. Thus, we note that,

$$E_{k-1}(\rho_k) = \rho_k^{k-1} = 0 \quad (14)$$

and

$$\begin{aligned} \Sigma_{k+1} &= \text{Var}(\rho_{k+1}) \\ &= E\{[C(x_{k+1} - x_{k+1}^{k}) + v_k][C(x_{k+1} - x_{k+1}^{k}) + v_k]^T\} \\ &= CE\{[x_{k+1} - x_{k+1}^{k}][x_{k+1} - x_{k+1}^{k}]^T\}C^T + E(v_k v_k^T) \\ &= CP_{k+1}^{k}C^T + R. \end{aligned}$$

We also note that,

$$E_k(\rho_{k+1}y_k^T) = E_k((y_{k+1} - y_{k+1}^{k})y_k^T) = 0,$$

which means that the innovations are independent of the past measurements. On the other hand, the conditional covariance between x_{k+1} and ρ_{k+1} is computed as follows

$$\begin{aligned} \text{Cov}(x_{k+1}, \rho_{k+1}) &= \text{Cov}(x_{k+1} - x_{k+1}^{k}, C(x_{k+1} - x_{k+1}^{k}) + v_k) \\ &= E\{[(x_{k+1} - x_{k+1}^{k}) - E_k(x_{k+1} - x_{k+1}^{k})] \\ &= [C(x_{k+1} - x_{k+1}^{k}) + v_k - CE_k(x_{k+1} - x_{k+1}^{k})]^T\} \\ &= P_{k+1}^{k}C^T. \end{aligned}$$

From theses results, we conclude that x_{k+1} and ρ_{k+1} have a Gaussian joint distribution conditional on $\{y\}_1^k$. That is ,

$$\{\begin{pmatrix} x_{k+1} \\ \rho_{k+1} \end{pmatrix} |\{y\}_1^k\} \sim N\{\begin{pmatrix} x_{k+1}^{k} \\ 0 \end{pmatrix}, \begin{pmatrix} P_{k+1}^{k} & P_{k+1}^{k}C^T \\ CP_{k+1}^{k} & \Sigma_{k+1} \end{pmatrix}\}. \quad (15)$$

Now, since ρ_{k+1} and y_k are independent,

$$\begin{aligned} x_{k+1}^{k+1} &= E_{k+1}(x_{k+1}) \\ &= E_k\{x_{k+1}|\rho_{k+1}\} \\ &= E_k(x_{k+1}) + \text{Cov}(x_{k+1}, \rho_{k+1})\Sigma_{k+1}^{-1}\rho_{k+1} \\ &= x_{k+1}^{k} + P_{k+1}^{k}C^T[CP_{k+1}^{k}C^T + R]^{-1}\rho_{k+1} \\ &= x_{k+1}^{k} + K_{k+1}[y_{k+1} - Cx_{k+1}^{k}]; \end{aligned}$$

where,

$$K_{k+1} = P_{k+1}^{k}C^T[CP_{k+1}^{k}C^T + R]^{-1}$$

is the Kalman gain.
To derive (8), we will use the information of (15), We get

$$\begin{aligned} P_{k+1}^{k+1} &= \text{Cov}(x_{k+1}, \rho_{k+1}) \\ &= \text{Cov}(x_{k+1}) - \text{Cov}(x_{k+1}, \rho_{k+1})\Sigma_{k+1}^{-1}\text{Cov}(\rho_{k+1}, x_{k+1}) \\ &= P_{k+1}^{k} - P_{k+1}^{k}C^T[CP_{k+1}^{k}C^T + R]^{-1}CP_{k+1}^{k} \\ &= P_{k+1}^{k} - K_{k+1}CP_{k+1}^{k} \\ &= [I - K_{k+1}C]P_{k+1}^{k} \end{aligned}$$

To derive (9), we have,

$$\begin{aligned} \ddot{P}_{k+1}^{k+1} &= E((x_{k+1} - x_{k+1}^{k+1})(z_{k+1} - z_{k+1}^{k+1})^T) \\ &= E(((x_{k+1} - x_{k+1}^{k+1})(x_{k+1} - x_{k+1}^{k+1})^T)[V_{k+1}^{k+1}]^T \\ &= P_{k+1}^{k+1}[V_{k+1}^{k+1}]^T \end{aligned}$$

By using the same argument for deriving (10), we obtain,

$$\begin{aligned} \dot{P}_{k+1}^{k+1} &= E((z_{k+1} - z_{k+1}^{k+1})(z_{k+1} - z_{k+1}^{k+1})^T) \\ &= (V_{k+1}^{k+1})E((x_{k+1} - x_{k+1}^{k+1})(x_{k+1} - x_{k+1}^{k+1})^T)[V_{k+1}^{k+1}]^T \\ &= (V_{k+1}^{k+1})P_{k+1}^{k+1}[V_{k+1}^{k+1}]^T \\ &= (V_{k+1}^{k+1})\ddot{P}_{k+1}^{k+1} \end{aligned}$$

IV. APPLICATION TO TEMPERATURES

In this section, we will show the efficiency of the proposed algorithm by applying it to real measurements. These measurements are the daily average temperature for Jeddah city, which is located in western region of Saudi Arabia. Here, we have true measurements for a period of five years. By comparing the true measurements with the estimated values from the proposed algorithm , we find that the algorithm gives good results. The results are shown below in two figures. The first figure is for a period of hundred days. The second figure is for the whole period.

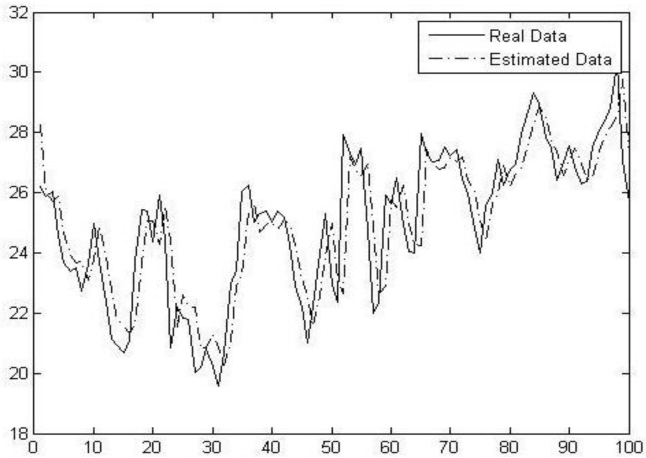

Fig. 1 Real data vs. estimated data for 100 days

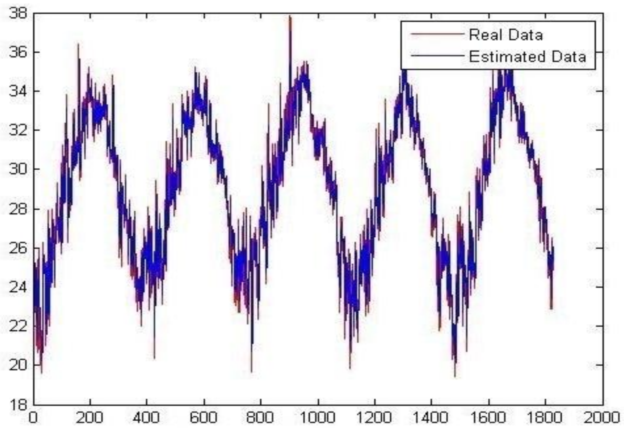

Fig. 2 Real data vs. estimated data for 5 years

V. CONCLUSION

The paper has introduced a new bilinear state space model. The bilinearity of this model depends on the product of the state vector by itself. This model generalizes Lotka volterra models and Lorenz models which have many applications in real life. Since the linear Kalman filter does not work with nonlinear systems, the paper has derived a new generalization of Kalman filter which work with the new bilinear model. The new algorithm depends on a linearization of the second order term by making use of the best available information about the state of the system. A new application to real data of temperatures are presented, which demonstrated that the proposed algorithm gives good results.

ACKNOWLEDGMENT

The author acknowledge with thanks King Abdulaziz University KAU for supporting.

REFERENCES

[1] B.D. Anderson and J.B. Moore, Optimal Filtering, Prentice-Hall, Englewood Cliffs, NJ, 1979.
[2] M. S. Gerwal and A. P. Andrews, Kalman Filtering: Theory and Practice Using MATLAB, John Wily and Sons Inc., 2001.
[3] N. S. Goel, et al., On The Volterra and Other Nonlinear Models of Interacting Populations, Academic Press Inc., 1971.
[4] D. Roy , Z.E. Musielak, Generalized Lorenz models and their routes to chaos. III, Energy-Conserving horizontal and vertical mode Truncations, Chaos, *Solitons and Fractals*, Vol. 33, pp. 1064–1070, 2007
[5] P.M. Pardalos and V.Yatsenko, Optimization and control of bilinear systems: theory, algorithms and applications, Springer, 2008
[6] S. H. Strgate, Nonlinear Dynamics and Chaos, with Applications to Physics, Biology, Chemistry, and Engineering, Presure Books l.l.c.,1994.
[7] H. Tanizaki, Nonlinear Filtering: Estimation and Applications, Springer-Verlag, 1996.
[8] N. Wiener, Nonlinear Problems in Random Theory, Boston, MA: The M.I.T. Press, 1958.

A. Al Mazrooei was porn in Jeddah, Saudi Arabia. He received the Bachelor degree and Master degree in mathematics from KAU, Saudi Arabia, in 1998 and 2006 respectively. He received the Ph.D. degrees in mathematics from KFUPM, Saudia Arabia, in 2011.