STATE RECONSTRUCTION IN NONLINEAR STOCHASTIC SYSTEMS BY EXTENDED KALMAN FILTER

R. Unbehauen
Friedrich Alexander University Erlangen-Nuremberg, Germany

Keywords: Kalman filtering, state estimator, state observer, state space description of nonlinear stochastic systems, stochastic difference equations, stochastic differential equations, stochastic process, Riccati differential equation

Contents

Summary

In this chapter the extended Kalman filter is introduced and examined as an estimation method widely used in many areas of control, signal processing and optimization. The extended Kalman filter can be considered as a general state estimator for nonlinear stochastically excited systems in the continuous-time case as well as in the discrete-time case. It originated from the Kalman-Bucy filter developed primarily for linear system applications and is an extension of the concept to the nonlinear estimation problem.

It will be pointed out that under certain assumptions: namely that the initial estimation error and the disturbing noise terms are small enough, the estimation error as a function of time is bounded in a specific sense. The presented results are illustrated by numerical simulations.

1. Introduction

Many analytical procedures for control design are based on state feedback and it is assumed that the full state vector is available for measurement. However, in a lot of applications of control the state variables of a system that have to be controlled are not accessible for direct measurement or the number of measuring devices is limited, in order to apply state feedback to stabilize, to optimize or to decouple the system. In other cases the measured state values are often corrupted by noise. In such situations a reasonable substitute for the state vector has to be found. A device that constructs an approximation of the state vector based on measured output signals is called a state estimator or a state observer. The design of an observer is possible for linear systems, in fact, for continuous-time systems as well as for discrete-time systems.

In the nonlinear case the estimation of the state vector corrupted by noise is commonly carried out by a device which is a compromise between accuracy and practical computational complexity. Many possibilities have been proposed to extend observers originally developed for linear systems to nonlinear systems. The proposal most used in practical applications is the extended Kalman filter, since it can be easily implemented and provides very good estimates. Although the original Kalman-Bucy filter was developed for linear systems only, it can be extended to nonlinear systems in a relatively simple manner. It is common to linearize the nonlinear system at the actual estimate in order to design a Kalman-Bucy filter for this linearization. This approach leads to the so-called extended Kalman filter which will be treated in the following. State estimation for nonlinear deterministic systems without disturbing noise is closely related the zero noise case.

2. The Continuous-time Extended Kalman Filter

2.1. State Estimation of Stochastically Excited Nonlinear Systems

2.1.1. Preparations

To treat randomly excited nonlinear continuous-time systems, the mathematical theory of stochastic differential equations is necessary. In the mathematical literature, the so-called Ito calculus is used. The special features of the stochastic Ito differential equations concern, among other things, the fact that the integrals are not Riemannian integrals and the chain rule is substituted by the so-called Ito formula. In the following, however, we restrict ourselves to those cases where the aforementioned particularities are of no importance.

The system to be considered is described by the equations

$$\frac{dz(t)}{dt} = f\big(z(t), x(t), t\big) + G(t)u(t), \tag{1a}$$

$$y(t) = g\big(z(t), x(t), t\big) + v(t) \tag{1b}$$

where z denotes the vector of the q state variables, x is the input m-vector and y the output r- vector; t denotes continuous time. The terms u and v denote noise disturbances. The nonlinear vector functions f and g are assumed to be given and to be continuously differentiable such that the differential equation (1a) has a unique solution in the stochastic sense. The quantity $G(t)$ means a specified time-variant $q \times k$ matrix. The initial state $z(0) = z_0$ is an unknown deterministic vector. Furthermore, the k-vector $u(t)$ and r-vector $v(t)$ are uncorrelated (statistically independent), zero mean white noise disturbance processes, i.e. we have

$$E\left[u(t)\right] = 0, \tag{2a}$$

$$E\left[v(t)\right] = 0, \tag{2b}$$

$$E\left[u(t)u^{\mathrm{T}}(\tau)\right] = M(t)\delta(t-\tau), \tag{3a}$$

$$E\left[v(t)v^{\mathrm{T}}(\tau)\right] = N(t)\delta(t-\tau), \tag{3b}$$

$$E\left[u(t)v^{\mathrm{T}}(\tau)\right] = 0 \tag{3c}$$

where $M(t)$ and $N(t)$ are specified time-variant positive definite matrices, $\delta(t)$ denotes the unit impulse function at t = 0. We choose the state observer (estimator) in the form

$$\frac{d\hat{z}(t)}{dt} = f\left(\hat{z}(t), x(t), t\right) + K(t)\left[y(t) - g\left(\hat{z}(t), x(t), t\right)\right] \tag{4}$$

where $\hat{z}(t)$ means the vector of the state estimate and $K(t)$ is the amplification (gain) matrix. Now, we develop the nonlinear functions in the following form:

$$f\left(z(t), x(t), t\right) - f\left(\hat{z}(t), x(t), t\right) = A(t)\left(z(t) - \hat{z}(t)\right) + r_f\left(z(t), \hat{z}(t), x(t), t\right) \tag{5a}$$

$$g\left(z(t), x(t), t\right) - g\left(\hat{z}(t), x(t), t\right) = C(t)\left(z(t) - \hat{z}(t)\right) + r_g\left(z(t), \hat{z}(t), x(t), t\right) \tag{5b}$$

with

$$A(t) = \left(\frac{\partial f}{\partial z}\right)_{\hat{z}(t), x(t), t}, \tag{6a}$$

$$C(t) = \left(\frac{\partial g}{\partial z} \right)_{\hat{z}(t), x(t), t} \tag{6b}$$

where $r_f\left(z(t), \hat{z}(t), x(t), t\right)$ and $r_g\left(z(t), \hat{z}(t), x(t), t\right)$ represent the summaries of the second and higher order terms. The estimation error is defined as

$$w(t) = z(t) - \hat{z}(t). \tag{7}$$

Subtracting Eq. (4) from Eq. (1a) and considering Eqs. (1b), (5a,b) and (6a,b), (7) one obtains the following differential equation for the estimation error

$$\frac{dw(t)}{dt} = \left[A(t) - K(t)C(t) \right] w(t) + r_N(t) + r_R(t), \tag{8}$$

where

$$r_N(t) = r_f\left(z(t), \hat{z}(t), x(t), t\right) - K(t) r_g\left(z(t), \hat{z}(t), x(t), t\right) \tag{9a}$$

includes the nonlinear terms and

$$r_R(t) = G(t)u(t) - K(t)v(t) \tag{9b}$$

combines the noise terms.

2.1.2. Design equations

The design equations of the extended Kalman filter follow simply from those of the Kalman-Bucy filter by substitution of the system matrices through the corresponding Jacobian matrices. This is in accordance with the linearization procedure of the nonlinear system about the present estimate $\hat{z}(t)$. The application of the extended Kalman filter requires the solution of the differential equation (4) for the estimate

$$\frac{d\hat{z}(t)}{dt} = f\left(\hat{z}(t), x(t), t\right) + K(t)\left[y(t) - g\left(\hat{z}(t), x(t), t\right) \right] \tag{10}$$

(assuming that the functions *x* and *y* are always available) simultaneously with the Riccati differential equation

$$\frac{dP(t)}{dt} = A(t)P(t) + P(t)A^{\mathrm{T}}(t) + Q(t) - P(t)C^{\mathrm{T}}(t)R^{-1}(t)C(t)P(t) \tag{11}$$

where we have the definitions

$$A(t)=\left(\frac{\partial f}{\partial z}\right)_{\hat{z}(t),x(t),t}, \tag{12a}$$

$$C(t)=\left(\frac{\partial g}{\partial z}\right)_{\hat{z}(t),x(t),t} \tag{12b}$$

and the amplification matrix is given as

$$K(t)=P(t)C^{\mathrm{T}}(t)R^{-1}(t). \tag{13}$$

Furthermore, $Q(t)$, $R(t)$ and the initial value $P(t_0)$ for the solution of the Riccati differential equation (11) have to be chosen as positive definite matrices.

The following remarks can be made:

1. The differential equations (10) and (11) are generally stochastic differential equations which can be solved numerically, for instance, with the aid of an Euler discretization and a probability generator for the noise processes.

2. Frequently $Q(t)$ and $R(t)$ are chosen as the covariance matrices of the noise processes $u(t)$ and $v(t)$, respectively, i.e.

$$Q(t)=G(t)M(t)G^{\mathrm{T}}(t), \tag{14a}$$

$$R(t)=N(t), \tag{14b}$$

where $M(t)$ and $N(t)$ are given by Eqs. (3a) and (3b). In certain cases, for example in estimating nonlinear systems without noise terms (i.e. $M(t)=\mathbf{0}$ and $N(t)=\mathbf{0}$), however, also in state estimation of stochastically excited nonlinear systems, a different choice of the matrices $Q(t)$ and $R(t)$ is not only possible but also significant. The matrices $Q(t)$ and $R(t)$ are, therefore, design parameters.

2.1.3. Dynamics of the estimation error

The Kalman-Bucy filter for linear systems is designed so that the error covariance matrix is minimized in a certain sense. For nonlinear systems the error covariance $E\left[w^T(t)w(t)\right]$ is generally not minimized; but it remains bounded provided that certain requirements are satisfied. In the following theorem this statement is made precise.

Theorem 1: *Consider a nonlinear randomly excited continuous-time system described* *by Eqs. (1a) - (3c) and an extended Kalman filter as defined by Eqs.(10)- (13). Assume that the following conditions are satisfied:*

1. *There exist real positive numbers* $c_{\max}, p_{\min}, p_{\max}, q_{\min}, r_{\min}$ *such that the following bounds are satisfied for every* $t \geq t_0$*:*

$$\left\| C(t) \right\| \leq c_{\max}, \tag{15}$$

$$p_{\min} I \leq P(t) \leq p_{\max} I, \tag{16}$$

$$q_{\min} I \leq Q(t), \tag{17a}$$

$$r_{\min} I \leq R(t), \tag{17b}$$

where I *denotes the unit matrix.*

2. *There exist real positive numbers* $\varepsilon_{fg}, k_f, k_g$ *such that the nonlinear functions* r_f *and* r_g *introduced in Eqs.(5a,b) are bounded by*

$$\left\| r_f\left(z, \hat{z}, x, t\right) \right\| \leq k_f \left\| z - \hat{z} \right\|^2, \tag{18a}$$

$$\left\| r_g\left(z, \hat{z}, x, t\right) \right\| \leq k_g \left\| z - \hat{z} \right\|^2 \tag{18b}$$

for $\left\| z - \hat{z} \right\| \leq \varepsilon_{fg}$ *and every* $x \in R^m$ *and* $t, z, \hat{z} \in R^q$.

Then there exist for every $\varepsilon_w > 0$ *two constants* $\delta_w, \delta_R > 0$ *such that the estimation error defined by Eq. (7) is bounded for all* $t \geq t_0$ *according to*

$$E\left[\left\| w(t) \right\|^2 \right] \leq \varepsilon_w^2 \tag{19}$$

provided that the initial error and the covariance matrices of the noise terms given by Eqs. (3a,b) are sufficiently small such that the inequalities

$$\left\| w\left(t_0\right) \right\| \leq \delta_w, \tag{20a}$$

$$G(t) M(t) G^{\mathrm{T}}(t) \leq \delta_R^2 I, \tag{20b}$$

$$N(t) \leq \delta_R^2 I \tag{20c}$$

are satisfied.

Remarks

1. The inequalities (17a,b) can be satisfied by a suitable choice of the matrices $Q(t)$ and $R(t)$, for example, time-invariant positive definite matrices can be chosen. On the other hand the conditions (15) and (16) must be numerically checked.

2. With the aid of standard evaluations one can show that the inequalities (18a,b) are satisfied if the functions f and g are twice continuously differentiable and the norms of the Hessian matrices belonging to f and g are bounded. Therefore, in many practical applications one can assume that the inequalities (18a,b) are satisfied.

3. It is to be noticed in contrast to the linear case that the solution of the Riccati differential equation is generally not identical to the covariance matrix of the estimation error, i.e. we have in general

   $$E\left[w(t) w^{\mathrm{T}}(t) \right] \not\equiv P(t) \tag{21}$$

4. If the system in question is linear and one takes

   $$Q(t) = G(t) M(t) G^{\mathrm{T}}(t), \tag{22a}$$

   $$R(t) = N(t), \tag{22b}$$

   then

   $$P(t) \equiv E\left[w(t) w^{\mathrm{T}}(t) \right], \tag{23}$$

   and the boundedness of the covariance of the estimation error follows directly from Eq.(16).

5. In other words, Theorem 1 states the following: Under suitable assumptions the estimation error remains bounded and the bound can be chosen arbitrarily small in so far as the initial estimation error and the noise disturbance terms are sufficiently small. However, if the initial estimation error or the noise becomes too strong, the estimation error may eventually diverge. That behavior will now be demonstrated by two examples.

-
-
-

Bibliography

Anderson B.D.O. and Moore J.B. (1979). *Optimal Filtering*, 357 pp.. Englewood Cliffs: Prentice-Hall. [This book is an advanced level text on filtering. It includes a discussion of the extended Kalman filter.]

Arnold L. (1973). Stochastische Differentialgleichungen, 239 pp.. München: Oldenbourg. [This book presents the essential aspects and approaches to the theory and application of stochastic differential equations.]

Brown R.G. and Hwang P.Y. (1992). Introduction to Random Signals and Applied Kalman Filtering, 502 pp.. New York: John Wiley & Sons. [This is an excellent textbook on Kalman filtering. It also describes the extended Kalman filter.]

Gard T.C. (1988). Introduction to Stochastic Differential Equations, 234 pp.. New York: Marcel Dekker. [This is a widely used book presenting the fundamental theory of stochastic differential equations in the simplest setting. It contains the stochastic version of Heun's method.]

Gardiner C.W. (1990). Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences, 442 pp.. Berlin: Springer-Verlag. [This book presents all essential aspects for describing nonlinear stochastic systems. Strong emphasis is placed on systematic approximation methods for solving problems.]

Gelb A., ed. (1974). Applied Optimal Estimation, 374 pp.. Cambridge, MA: MIT Press. [This book is an introduction to optimal estimation with its major emphasis on applications, treating the subject more from an engineering than a mathematical orientation.]

Grewal M.S. and Andrews A.P. (1993). Kalman Filtering: Theory and Practice, 381 pp.. Englewood Cliffs: Prentice-Hall. [This book treats applications of Kalman-Bucy filtering.]

Jazwinski A.H. (1970). Stochastic Processes and Filtering Theory, 376 pp., New York: Academic Press. [This book presents a unified treatment of linear and nonlinear filtering theory for engineers with an emphasis on applications.]

Kalman R.E. (1963). New methods in Wiener filtering theory. (Proceedings of 1st Symp. on English Applications of Random Function Theory and Probability, J. Bogdano and F. Kozin, Eds.) pp. 270-388. New York: John Wiley & Sons.

Kushner H. (1967). Stochastic Stability and Control, 161 pp., New York: Academic Press. [In this monograph the stochastic Lyapunov approach is developed which plays the same role in the study of stochastic processes as Lyapunov functions do for deterministic processes.]

Lewis, F.L. (1986). Optimal Estimation: with an Introduction to Stochastic Control Theory, 376 pp., New York: John Wiley & Sons. [This book treats the optimal control of systems in a clear and direct manner. The basic idea of the extended Kalman filter is contained.]

Ljung L. (1979). Asymptotic behavior of the extended Kalman filter as a parameter estimator for linear systems, IEEE Trans. Autom. Contr., AC-24, 36-50. [The paper describes a one-step formulation of the discrete-time extended Kalman filter.]

Reif K., Günther S., Yaz E. and Unbehauen, R. (1998). Stabilität des zeitkontinuierlichen erweiterten

Kalman-Filters. Automatisierungstechnik 46 (12) 592-601. [This paper presents a proof of Theorem 1.]

Reif, K., Günther, S., Yaz, E. and Unbehauen, R. (1999). Stochastic stability of the discrete time extended Kalman filter. IEEE Trans. Autom. Contr. 44, 714-728. [The authors analyze the error behavior of the discrete-time extended Kalman filter for general nonlinear systems in a stochastic framework.]

Rümelin, W. (1982). Numerical treatment of stochastic differential equations, SIAM J. Numer. Anal., 19, pp. 604-613. [The paper describes the stochastic version of Heun's method.]

Unbehauen, R. (1998). System theory 2, Munich: 743pp. Oldenbourg Verlag. [This book presents detailed proofs of Theorems 2 and 3.]

Biographical Sketch

Rolf Unbehauen received the diploma degree in mathematics, Ph.D. in electrical engineering, and the Habilitate Doctorate in electrical engineering from Stuttgart University, Germany, in 1954, 1957, and 1964, respectively. From 1955 to 1966, he was a Member of the Institute of Mathematics, the Computer Center, and the Institute of Electrical Engineering at Stuttgart University, where he was appointed Associate Professor in 1965. Since 1966, he has been a Full Professor of Electrical Engineering at the University of Erlangen-Nürnberg, Erlangen, Germany. He has published a great number of papers on his research results. He has authored four books in German and is co-author of *MOS Switched-Capacitor and Continuous-Time Integrated Circuits and Systems,* Berlin: Springer-Verlag, 1989; *Neural Networks for Optimization and Signal Processing*, New York: Teubner-Verlag and Wiley, 1993; and *Applied Neural Networks for Signal Processing*, New York: Cambridge Univ. Press, 1997. His teaching and research interests include network theory and its application, system theory, and electromagnetics. From 1990 to 1991 Dr. Unbehauen was an Associate Editor of IEEE Transactions on Circuits and Systems. From 1996 to 1998 he served as an Associate Editor of IEEE Transactions on Neural Networks. Currently, he is an Associate Editor of Multidimensional Systems and Signal Processing. In 1959, he received the NTG Best Paper Award, in 1994 a honorary doctor from Technical University Cluj-Napoca, Romania, and in 2001 the Coales Premium Award. He is a Life Fellow of IEEE, a Member of the Informationstechnische Gesellschaft of Germany and of URSI, Commission C: Signals and Systems.