Identification of a class of nonlinear state-space models using RPE techniques

Zhou, W. W.; Blanke, Mogens

Published in:
25th IEEE Conference on Decision and Control

Link to article, DOI:
10.1109/CDC.1986.267185

Publication date:
1986

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):
Identification of a Class of Nonlinear State-Space Models using RPE Techniques

W.W. Zhou and M. Blanke

Servolaboratory, build. 326
Technical University of Denmark
2800 Lyngby
on leave from
Jimei Navigation Institute
Xiamen, China

Summary

The recursive prediction error methods in state-space form have been efficiently used as parameter identifiers for linear systems, and especially Ljung's innovations filter using a Newton search direction has proved to be quite ideal. In this paper, the RPE method in state-space form is developed to the nonlinear case and extended to include the exact form of a nonlinearity, thus enabling structure preservation for certain classes of nonlinear systems. Both the discrete and the continuous-discrete versions of the algorithm in an innovations model are investigated, and a nonlinear simulation example shows a quite convincing performance of the filter as combined parameter and state estimator.

1. Introduction

In this paper we present two parameter identifiers for nonlinear discrete and continuous-discrete state-space models. These algorithms are investigated by using the linear recursive prediction error (RPE) method, Ljung and Soderstrom [1983], in combination with nonlinear second-order filtering theory Jazwinski [1970], Mayback [1982], Zhou [1985].

2. Model and algorithm in discrete version

We assume a nonlinear discrete state-space model of the following form,

\[ x(t+1) = f(x(t);u(t),x(t),t) + v(t) \]
\[ y(t) = h(0;x(t);t) + e(t) \]

where \( f() \) and \( h() \) are nonlinear functions of the state, \( v(t) \) is white process noise, and \( e(t) \) is uncorrelated measurement noise with statistics

\[ E(v(t) = E(e(t) = 0) \]
\[ E(v(t)v(t)^\top) = R_v(t) \delta_{11} \]
\[ E(e(t)e(t)^\top) = R_e(t) \delta_{11} \]
\[ E(v(t)e(t)^\top) = R_{ve}(t) \delta_{11} \]

The initial value of the state \( x(0) \) has the properties

\[ E(x(0) = x_0(0) \]
\[ E(x(0)x(0)^\top) - x_0(0) = R_0(0) \delta_{11} \]

From the nonlinear filtering theory [Jazwinski, 1970; Mayback, 1982] it is known that an attractive and applicable nonlinear filter is the first-order filter with bias correction term (FOFBC), which is based on using first-order covariance and gain computations, but with the second-order terms in state expectation and prediction error equations. In this study we use the FOFBC method for identification of the nonlinear model (1-a, b). When a fixed value \( \theta \) is given, the predictor corresponding to (1-a, b) will be

\[ \hat{x}(t+1,\theta) = f(\theta;u(t),\hat{x}(t,\theta)) + B_x(t) + K(t)[y(t) - h(\theta,t,\hat{x}(t,\theta))] - B_y(t) \]
\[ \hat{y}(t;\theta) = h(\theta,t,\hat{x}(t,\theta)) \]

where the second order term \( B_x(t) \) is the \( n_x \)-vector with \( K \)th component

\[ B_x(t) = \frac{1}{2} tr \left[ \frac{\partial^2 f(\theta;u(t),\hat{x}(t,\theta))}{\partial x^2} P(t) \right] \]

and \( B_y(t) \) is the \( n_y \)-vector with \( K \)th component

\[ B_y(t) = \frac{1}{2} tr \left[ \frac{\partial^2 h(\theta;\hat{x}(t,\theta))}{\partial x^2} P(t) \right] \]

One finds that use of the recursive prediction error method by Ljung and Soderstrom, [1983], directly on the nonlinear predictor model (2-a, b) is hardly feasible, due to computational complexity. If a linear measurement equation is chosen instead, however, complexity of the algorithm is reduced significantly. Then the predictor has the following form

\[ \hat{x}(t+1,\theta) = f(\theta;u(t),\hat{x}(t,\theta)) + B_x(t) + K(t)[y(t) - H(\theta,\hat{x}(t,\theta))] - B_y(t) \]
\[ \hat{y}(t;\theta) = H(\theta,\hat{x}(t,\theta)) \]

The assumption of a linear measurement is valid in a wide class of practical applications. Then the recursive prediction error method using a Newton search direction for parameter updating can be applied to the model (3-a, b). The algorithm will consist of the following set of recursive equations:
\( \epsilon(t) = y(t) - \hat{y}(t) \)

\( R(t) = R(t - 1) + \alpha(t) \psi(t) S^{-1}(t) \psi^T(t) - R(t - 1) \)

\( \hat{\theta}(t) = \hat{\theta}(t - 1) + \alpha(t) R^{-1}(t) \psi(t) S^{-1}(t) \epsilon(t) \)

\( K(t) = F_t \psi(t) H_t^T + R_{12}(t) S^{-1}(t) \)

\( P(t + 1) = F_t P(t) F_t^T + R(t) - K(t) S(t) K^T(t) \)

\( S(t) = H_t P(t) H_t^T + R(t) \)

\( \hat{x}(t + 1) = f(\theta, u, t, \hat{x}(t), \theta) + B_{x}(t) + K(t) \epsilon(t) \)

\( \gamma(t + 1) = H_t \hat{x}(t + 1) \)

\( W(t + 1) = \hat{W}_x(t, \theta) + \hat{M}_x(t, \theta) - K(t) D_t \)

\( \psi^T(t + 1) = H_t W(t + 1) + D_t \)

where

\[ F_t = \left. \frac{\partial \epsilon(t, \theta)}{\partial x} \right|_{\theta = \hat{\theta}(t)} \]

\[ H_t = \left. H(\hat{\theta}(t)) \right|_{\theta = \hat{\theta}(t)} \]

\[ \hat{W}_x(t, \theta) = \left. \frac{\partial}{\partial \theta} \left[ f(\theta, u, t, \hat{x}(t), \theta) + B_{x}(t) + K(t) \epsilon(t) \right] \right|_{\theta = \hat{\theta}(t)} \]

\[ \hat{K}_t = \left. \frac{\partial}{\partial \theta} K(t) \right|_{\theta = \hat{\theta}(t)} \]

\[ D_t = \left. \left( H(\hat{\theta}(t)) \right) \right|_{\theta = \hat{\theta}(t)} \]

\( B_x(t) \) is defined in \((2-c)\).

This version of the filter \((4-a-\overline{1})\) includes a calculation of the Kalman gains in \((4-d, e, f)\) and \( K_t \) is calculated from \((4-d, e, f)\). As per the suggestion given by Ljung (1979), the parameter identifier can assume an innovations model of the form:

\[ \hat{x}(t + 1) = f(\theta, u, t, \hat{x}(t), \theta) + B_{x}(t) + K(\theta) \epsilon(t) \]

\[ y(t) = H(\theta) \hat{x}(t, \theta) + \epsilon(t) \]

where \( \epsilon(t) \) is the innovation due to measurement \( t \), and \( K(\theta) \) is a set of (as yet undetermined) steady state Kalman gains, which is parameterized and will be identified directly along with the system parameters. This gives less complex computations, and the algorithm corresponding to \((6-a, b)\) will then be as follows:

\[ \epsilon(t) = y(t) - \hat{y}(t) \]

\[ \hat{x}(t) = \hat{x}(t - 1) + \alpha(t) [\psi(t) S^{-1}(t) \psi^T(t) - \hat{x}(t - 1)] \]

\[ R(t) = R(t - 1) + \alpha(t) \psi(t) S^{-1}(t) \psi^T(t) - R(t - 1) \]

\[ \hat{\theta}(t) = \hat{\theta}(t - 1) + \alpha(t) R^{-1}(t) \psi(t) S^{-1}(t) \epsilon(t) \]

\[ K(t) = F_t \psi(t) H_t^T + R_{12}(t) S^{-1}(t) \]

\[ P(t + 1) = F_t P(t) F_t^T + R(t) - K(t) S(t) K^T(t) \]

\[ S(t) = H_t P(t) H_t^T + R(t) \]

\[ \hat{x}(t + 1) = f(\theta, u, t, \hat{x}(t), \theta) + B_{x}(t) + K(t) \epsilon(t) \]

\[ \gamma(t + 1) = H_t \hat{x}(t + 1) \]

\[ W(t + 1) = \hat{W}_x(t, \theta) + \hat{M}_x(t, \theta) - K(t) D_t \]

\[ \psi^T(t + 1) = H_t W(t + 1) + D_t \]

where

\[ \epsilon(t) = y(t) - \hat{y}(t) \]

\[ R(t) = R(t - 1) + \alpha(t) [\psi(t) S^{-1}(t) \psi^T(t) - R(t - 1)] \]

\[ \hat{\theta}(t) = \hat{\theta}(t - 1) + \alpha(t) R^{-1}(t) \psi(t) S^{-1}(t) \epsilon(t) \]

\[ K(t) = F_t \psi(t) H_t^T + R_{12}(t) S^{-1}(t) \]

\[ P(t + 1) = F_t P(t) F_t^T + R(t) - K(t) S(t) K^T(t) \]

In most applications involving the identification of parameters of a physical continuous time system, it is generally preferable to use a continuous-discrete algorithm. The reasons are primarily structure preservation of known parts of the system and the possibility to include bounds on parameter estimates of physical parameters whose constraints are known. The latter is a practical way to overcome part of the difficulties with possible local minima when identifying parameters of nonlinear systems. As in the presentation in section 2, the discrete measurement equation will be chosen in its linear version, and an innovations model is employed. We hence assume the nonlinear continuous-discrete state-space model of the form:

\[ \begin{align*}
    d \frac{\partial}{\partial t} x(t) &= f(\theta, u, t, x(t), \theta) + v(t) \\
    y(t + 1) &= H(\theta) x(t + 1) + v(t + 1)
\end{align*} \]

where \( f(\cdot) \) is the nonlinear function of state, \( v(t) \) is white process noise, \( v(t) \) is uncorrelated measurement noise with statistics, \( E v(t) = 0 \)

\[ E \epsilon(t) = 0 \]

The second order predictor using an innovations model will be

\[ \hat{x}(t, \theta) = f(\theta, u, t, \hat{x}(t), \theta) + B_{x}(t) \]

\[ y(t + 1) = H(\theta) \hat{x}(t, \theta) + \epsilon(t, \theta) \]

where \( \epsilon(t + 1) \) is the innovation due to measurement \( t + 1 \), and \( K(t + 1, \theta) \) comprise parameterized steady state Kalman gains. The algorithm corresponding to \((9-a, b, c)\) will be as follows:
After integration of (10-a,b,c), \( \dot{x}(t_i) \), \( P(t_i) \), \( W(t_i) \) are available, and

\[
\begin{align*}
W(t_i) &= I - K(t_i) \Theta(t_i) \Theta(t_i) - N(t_i) - N(t_i) \Theta(t_i) \Theta(t_i) - K(t_i) \Theta(t_i) \Theta(t_i) - N(t_i) - N(t_i) \\
\Sigma(t_i) &= H_i W(t_i) + D(t_i) \Theta(t_i) W(t_i) \\
\lambda(t_i) &= \hat{\lambda}(t_i) + a(t_i) + P(t_i) \\
B_i &= a(t_i) + a(t_i) + P(t_i) \\
K_i &= K(t_i) \\
N_i &= N(t_i) \\
M_i &= M(t_i) \\
V_i &= V(t_i) \\
N_i &= N(t_i) - K(t_i) \\
W_i &= W(t_i) - K(t_i) \\
\text{Figure 1 shows results of identifying the parameters } a \text{ and } b \text{ in the nonlinear equation using the nonlinear filter.} \\
\text{The curves plotted in figure 2 illustrate the performance of a linear RPE filter applied to the same nonlinear equation.} \\
\text{Although the driving signal's perturbation is only 10 percent of its steady state value, the practical equivalent to this experiment would be a stepwise increase/decrease in propeller thrust.}
\end{align*}
\]

The responses and parameter estimates below were obtained using a square wave perturbation to the input \( u(t) \). The amplitude of the perturbation is 10 percent of its steady state value. The practical equivalent to this experiment would be a stepwise increase/decrease in propeller thrust.

\[
\begin{align*}
B_x(t_i) &= \frac{1}{2} \left( \frac{\partial^2 \dot{x}(t_i \Theta(t_i))}{\partial x^2} \right) P(t_i) \\
\text{and} \\
\overline{W}_x^*(t_i) &= \frac{\partial}{\partial \Theta} \left( \frac{\partial^2 \dot{x}(t_i \Theta(t_i))}{\partial x^2} + B_x(t_i \Theta(t_i)) \right) \bigg|_{\Theta = \hat{\Theta}(t_i)} \\
\text{is the derivative of } x(t_i) \text{ with respect to } \Theta. \\
\overline{M}_g(t_i) &= \frac{\partial}{\partial \Theta} \left( f(\Theta, u(t), x(t_i)) + B_x(t_i \Theta(t_i)) \right) \bigg|_{\Theta = \hat{\Theta}(t_i)} \\
\text{is the derivative of the parameter matrices in the right-hand side of (9-a) with respect to } \Theta. \text{ Further the following notation is used} \\
D(\Theta, \dot{x}(t_i \Theta(t_i))) &= \frac{\partial}{\partial \Theta} \left( H(\Theta \dot{x}) \right) \bigg|_{\Theta = \hat{\Theta}(t_i)} \\
N(\dot{x}(t_i \Theta(t_i))) &= \frac{\partial}{\partial \Theta} \left( K(t_i) \dot{x}(t_i \Theta(t_i)) \right) \bigg|_{\Theta = \hat{\Theta}(t_i)} \\
H_1 &= H(\hat{\Theta}(t_i)) \\
F_i &= \frac{\partial^2 \Theta(t_i \Theta(t_i))}{\partial x} \bigg|_{\Theta = \hat{\Theta}(t_i)} \\
\text{The same treatment will be used when } H_1 \text{ is an identity matrix and has the same dimension as the state vector } x. \text{ In this case the } P(t_i) \text{ matrix will not be calculated any longer and is replaced by } \dot{\lambda}(t_i). \\
\end{align*}
\]
5. Conclusions

This paper has presented two algorithms for identifying parameters of a nonlinear discrete state-space system model and a nonlinear continuous-discrete state-space system model. Both versions are treated using a linear discrete measurement equation. These algorithms were investigated with reference to the theory of linear RPE methods and the theory of nonlinear filtering. The innovations model formulation was found to be attractive, and the algorithms were implemented and tested against computer simulations showing excellent convergence, and bias properties that by far exceed those of a linear continuous/discrete filter. The analysis of the convergence properties of the nonlinear estimator and further tests of applications of these algorithms should be pursued in a further study.

6. References


Ljung, L.: System Identification - theory for the user. Lecture Notes, University of Linköping, Sweden 1984


Young, P.: Recursive Estimation and Time-Series Analysis. Springer-Verlag, 1984

Identification of a nonlinear model
\[ \frac{d}{dt}(x) = -0.58x^2 + 0.2u \]
1. estimated, 2. true
nonlinear version is used

Identification of a nonlinear model
\[ \frac{d}{dt}(x) = -0.58x^2 + 0.2u \]
1. estimated, 2. true
linear version is used

Figure 1 Identification of parameters a and b of equation 11 with square wave perturbation on the input signal. The nonlinear estimator is used.

Figure 2 Identification of parameters a and b of equation 11 with similar excitation as in figure 1. The nonlinear estimator is used.
Identification of a nonlinear model
\[ \frac{d}{dt}(x) = -0.58x^2 + 0.2u \]
1. estimated, 2. true nonlinear version is used measurements corrupted with noise

**Figure 3** Same example as figure 1 with the nonlinear filter, but measurement corrupted with noise.