DTU Library

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

Proceedings of 25th Conference
on Decision and Control
Athens, Greece • December 1986

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

Søren T. Lyngsø A/S
Lyngsø Allé
2970 Hørsholm, Denmark

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 Söderström [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,

$$\begin{cases} x(t+1)=f(\theta,u;t,x(t))+v(t) & \text{(1-a)} \\ y(t)=h(\theta;t,x(t))+e(t) & \text{(1-b)} \end{cases}$$

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

$$\begin{aligned} &E\,v(t)=E\,e(t)=0 \\ &E\,v(t)\,v^T(t)=R_1(\theta)\,\delta_{tt} \\ &E\,e(t)\,e^T(t)=R_2(\theta)\,\delta_{tt} \\ &E\,v(t)\,e^T(t)=R_{12}(\theta)\,\delta_{tt} \end{aligned} \qquad \text{(1-c)}$$

The initial value of the state x(o) has the properties

$$\begin{cases} E\,x(0)=x_0(\theta) \\ E\,[x(0)-x_0(\theta)][x(0)-x_0(\theta)]^T=\Pi_0(\theta) \end{cases} \qquad \text{(1-d)}$$

From the nonlinear filtering theory [Jazwinski, 1970][Mayback, 1982] it is known that an attractive and appliable 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 θ is given, the predictor corresponding to (1-a, b) will be

$$\begin{cases} \hat{x}(t+1,\theta)=f(\theta,u;t,\hat{x}(t,\theta))+B_x(t)+K(t)[y(t)-h(\theta;t,\hat{x}(t,\theta))-B_y(t)] & \text{(2-a)} \\ \hat{y}(t|\theta)=h(\theta;t,\hat{x}(t,\theta)) & \text{(2-b)} \end{cases}$$

where the second order term $B_x(t)$ is the n_x-vector with K^{th} component

$$B_{xk}(t)=\frac{1}{2}\,tr\left\{\frac{\partial^2 f_k(\theta,u;t,\hat{x}(t,\theta))}{\partial x^2}\,P(t)\right\} \qquad \text{(2-c)}$$

and By(t) is the ny-vector with Kth component

$$B_{yk}(t)=\frac{1}{2}\,tr\left\{\frac{\partial^2 h_k(\theta;t,\hat{x}(t,\theta))}{\partial x^2}\,P(t)\right\} \qquad \text{(2-d)}$$

One finds that use of the recursive prediction error method by Ljung and Söderström, [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

$$\begin{cases} \hat{x}(t+1,\theta)=f(\theta,u;t,\hat{x}(t,\theta))+B_x(t)+K(t)[y(t)-H(\theta)\hat{x}(t,\theta)] & \text{(3-a)} \\ \hat{y}(t|\theta)=H(\theta)\hat{x}(t,\theta) & \text{(3-b)} \end{cases}$$

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:

 CH2344-0/86/0000-1637 $1.00

$$\varepsilon(t)=y(t)-\hat{y}(t) \tag{4-a}$$

$$R(t)=R(t-1)+\alpha(t)[\psi(t)S^{-1}(t)\psi^{T}(t)-R(t-1)] \tag{4-b}$$

$$\hat{\theta}(t)=\hat{\theta}(t-1)+\alpha(t)R^{-1}(t)\psi(t)S^{-1}(t)\varepsilon(t) \tag{4-c}$$

$$K(t)=[F_t P(t) H_t^T+R_{12}(t)]S^{-1}(t) \tag{4-d}$$

$$P(t+1)=F_t P(t) F_t^T+R_1(t)-K(t)S(t)K^T(t) \tag{4-e}$$

$$S(t)=H_t P(t) H_t^T+R_2(t) \tag{4-f}$$

$$\hat{x}(t+1)=f(\theta,u;t,\hat{x}(t,\theta))+B_x(t)+K(t)\varepsilon(t) \tag{4-g}$$

$$\hat{y}(t+1)=H_t\hat{x}(t+1) \tag{4-h}$$

$$W(t+1)=\overline{W}_x(t,\theta)+\overline{M}_\theta(t)+\overline{K}_t\varepsilon(t)-K(t)D_t \tag{4-i}$$

$$\psi^T(t+1)=H_t W(t+1)+D_t \tag{4-j}$$

where

$$F_t=\left.\frac{\partial f(\theta,u;t,\hat{x}(t,\theta))}{\partial x}\right|_{\theta=\hat{\theta}(t)} \tag{5-a}$$

$$H_t=H(\hat{\theta}(t)) \tag{5-b}$$

$$\overline{W}_x(t,\theta)=\left.\frac{\partial}{\partial\theta}\left\{f(\theta,u;t,\hat{x}(t,\theta))+B_x(t,\hat{x}(t,\theta))-K(t)H(\theta)\hat{x}(t,\theta)\right\}\right|_{\theta=\hat{\theta}(t)} \tag{5-c}$$

is the derivative of $x(t,\theta)$ in the right-hand side of (3-a) with respect to θ. Further

$$\overline{M}_\theta(t)=\left.\frac{\partial}{\partial\theta}\left\{f(\theta,u;t,\hat{x})+B_x(\theta;t)\right\}\right|_{\theta=\hat{\theta}(t)} \tag{5-d}$$

is the derivative of the parameter matrices in the bracket with respect to θ, and

$$\overline{K}_t=\left.\frac{\partial}{\partial\theta}K(t)\right|_{\theta=\hat{\theta}(t)} \tag{5-e}$$

$$D_t=\left.\frac{\partial}{\partial\theta}(H(\theta)\hat{x})\right|_{\theta=\hat{\theta}(t)} \tag{5-f}$$

$$B_x(t)\ \text{is defined in (2-c)} \tag{5-g}$$

This version of the filter (4-a~j) includes a calculation of the Kalman gains in (4-d,e,f) and $\overline{K}_t$ 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:

$$\begin{cases}\hat{x}(t+1,\theta)=f(\theta,u;t,\hat{x}(t,\theta))+B_x(t)+K(\theta)\varepsilon(t) & \text{(6-a)}\\ y(t)=H(\theta)\hat{x}(t,\theta)+\varepsilon(t) & \text{(6-b)}\end{cases}$$

where $\varepsilon(t)$ 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 compatations, and the algorithm corresponding to (6-a,b) will then be as follows:

$$\varepsilon(t)=y(t)-\hat{y}(t) \tag{7-a}$$

$$\hat{\Lambda}(t)=\hat{\Lambda}(t-1)+\alpha(t)[\varepsilon(t)\varepsilon^T(t)-\hat{\Lambda}(t-1)] \tag{7-b}$$

$$R(t)=R(t-1)+\alpha(t)[\psi(t)\hat{\Lambda}^{-1}(t)\psi^T(t)-R(t-1)] \tag{7-c}$$

$$\hat{\theta}(t)=\hat{\theta}(t-1)+\alpha(t)R^{-1}(t)\psi(t)\hat{\Lambda}^{-1}(t)\varepsilon(t) \tag{7-d}$$

$$\hat{x}(t+1)=f(\theta,u;t,\hat{x}(t,\theta))+B_x(t)+K_t\varepsilon(t) \tag{7-e}$$

$$\hat{y}(t+1)=H_t\hat{x}(t+1) \tag{7-f}$$

$$W(t+1)=\overline{W}_x(t,\theta)+\overline{M}_\theta(t)-K_t D_t \tag{7-g}$$

$$\psi^T(t+1)=H_t W(t+1)+D_t \tag{7-h}$$

where

$$K_t=K(\hat{\theta}(t))$$

$$H_t=H(\hat{\theta}(t))$$

$\overline{W}_x(t,\theta)$, $\overline{M}_\theta(t)$, D_t are defined in (5-c,d,f) respectively, and $B_x(t)$ is defined in (2-c). It is noted that in version (7-a~h) one has to use equations (4-e) and (4-f) in order to obtain the covariance matrix P(t) in $B_x(t)$. If the measurement vector y(t) has the same dimension as the state x, and the matrix H is an identity matrix then the covariance matrix is

$$P(t)=E\{(x(t)-\hat{x}(t))(x(t)-\hat{x}(t))^T\}$$
$$=E\{\varepsilon(t)\varepsilon^T(t)\}$$

Since $y(t)=H_t\hat{x}(t)=\hat{x}(t)$. Consequently, the matrix P(t) can be replaced by $\hat{\Lambda}(t)$ in this case, and P(t) need no longer be calculated.

3. Model and algorithm in continuous-discrete version

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:

$$\begin{cases}\dfrac{d}{dt}x(t|t_i)=f(\theta,u;t,x(t|t_i))+v(t|t_i) & \text{(8-a)}\\ y(t_{i+1})=H(\theta)x(t_{i+1})+e(t_{i+1}) & \text{(8-b)}\end{cases}$$

where f( ) is the nonlinear function of state. $v(t|t_i)$ is white process noise, $e(t_i)$ is uncorrelated measurement noise with statistics,

$$E\,v(t)=E\,e(t_i)=0$$

$$E\,v(t)v^T(\tau)=R_1(\theta)\delta(t-\tau)$$

$$E\,e(t_i)e^T(t_j)=R_2(\theta)\delta_{ij}$$

The second order predictor using an innovations model will be

$$\frac{d}{dt}\hat{x}(t|t_i,\theta)=f(\theta,u;t,\hat{x}(t|t_i,\theta))+B_x(t|t_i) \tag{9-a}$$

$$\varepsilon(t_{i+1})=y(t_{i+1})-H(\theta)\hat{x}(t_{i+1}^-,\theta) \tag{9-b}$$

$$\hat{x}(t_{i+1}^+,\theta)=\hat{x}(t_{i+1}^-,\theta)+K(t_{i+1},\theta)\varepsilon(t_{i+1}) \tag{9-c}$$

where $\varepsilon(t_{i+1})$ is the innovation due to measurement t_{i+1}, and $K(t_{i+1},\theta)$ comprise parameterized steady state Kalman gains. The algorithm corresponding to (9-a,b,c) will be as follows:

$$\frac{d}{dt}\hat{x}(t|t_i) = f(0,u;t,\hat{x}(t|t_i,\theta)) + B_x(t|t_i) \tag{10-a}$$

$$\frac{d}{dt}P(t|t_i) = F_i P(t|t_i) + P(t|t_i)F_i^T + R_1(t_i) \tag{10-b}$$

$$\frac{d}{dt}W(t|t_i) = \overline{W}_x^*(t|t_i) + \overline{M}_\theta(t_i) \tag{10-c}$$

After integration of (10-a,b,c), $\hat{x}(t_{i+1^-})$, $P(t_{i+1^-})$, $W(t_{i+1^-})$ are available, and

$$W(t_{i+1}^+) = [I - K(t_i,\theta)H_i]W(t_{i+1}^-) - N(t_{i+1},\theta) - K(t_i,\theta)D(\theta,\hat{x}(t_{i+1}^-)) \tag{10-d}$$

$$\psi^T(t_{i+1}) = H_i W(t_{i+1}^+) + D(\theta,\hat{x}(t_{i+1}^-)) \tag{10-e}$$

$$\hat{y}(t_{i+1}^-) = H_i\hat{x}(t_{i+1}^-) \tag{10-f}$$

$$\varepsilon(t_{i+1}) = y(t_{i+1}) - \hat{y}(t_{i+1}^-) \tag{10-g}$$

$$\hat{\Lambda}(t_{i+1}) = \hat{\Lambda}(t_i) + \alpha(t_{i+1})[\varepsilon(t_{i+1})\varepsilon^T(t_{i+1}) - \hat{\Lambda}(t_i)] \tag{10-h}$$

$$R(t_{i+1}) = R(t_i) + \alpha(t_{i+1})[\psi(t_{i+1})\hat{\Lambda}^{-1}(t_{i+1})\psi^T(t_{i+1}) - R(t_i)] \tag{10-i}$$

$$\hat{\theta}t_{i+1} = \hat{\theta}(t_i) + \alpha(t_{i+1})R^{-1}(t_{i+1})\psi(t_{i+1})\hat{\Lambda}^{-1}(t_{i+1})\varepsilon(t_{i+1}) \tag{10-j}$$

$$\hat{x}(t_{i+1}^+) = \hat{x}(t_{i+1}^-) + K(t_{i+1},\theta)\varepsilon(t_{i+1}) \tag{10-k}$$

$$P(t_{i+1}^+) = P(t_{i+1}^-) - K(t_{i+1},\theta)H_i P(t_{i+1}^-) \tag{10-l}$$

where $B_x(t|t_i)$ is the n_x-vector with k^{th} component

$$B_{xk}(t|t_i) = \frac{1}{2}\,\mathrm{tr}\left\{\frac{\partial^2 f_k(\theta,u;t,\hat{x}(t|t_i,\theta))}{\partial x^2}P(t)\right\}$$

and

$$\overline{W}_x^*(t|t_i,\theta) = \frac{\partial}{\partial\theta}\left\{f(\theta;t,\hat{x}(t|t_i,\theta)) + B_x(t|t_i)\right\}\Big|_{\theta=\hat{\theta}(t_i)}$$

is the derivative of $x(t|t_i,\theta)$ in the right-hand side of (9-a) with respect to θ.

$$\overline{M}_\theta(t_i) = \frac{\partial}{\partial\theta}\left\{f(\theta,u;t,\hat{x}) + B_x(\theta;t|t_i)\right\}\Big|_{\theta=\hat{\theta}(t_i)}$$

is the derivative of the parameter matrices in the right-hand side of (9-a) with respect to θ. Further the following notation is used

$$D(\theta,\hat{x}(t_{i+1}^-)) = \frac{\partial}{\partial\theta}\left\{H(\theta)\hat{x}\right\}\Big|_{\theta=\hat{\theta}(t_i)}$$

$$N(t_{i+1},\theta) = \left[\frac{\partial}{\partial\theta}K(t_{i+1},\theta)\right]\Big|_{\theta=\hat{\theta}(t_i)}\varepsilon(t_{i+1})$$

$$H_i = H(\hat{\theta}(t_i))$$

$$F_i = \frac{\partial f(\theta,u;t,\hat{x}(t|t_i,\theta)}{\partial x}\Big|_{\theta=\hat{\theta}(t_i)}$$

The same treatment will be used when H_i is an identity matrix and has the same dimension as the state vector x. In this case the $P(t_i)$ matrix will not be calculated any longer and is replaced by $\hat{\Lambda}(t_i)$.

4. Example

The ability of the nonlinear RPE method to estimate parameters and states of a nonlinear system of practical importance is demonstrated in this example.

The continuous discrete version of the nonlinear filter derived above is compared with the corresponding linear algorithm by Gavel and Azevedo [1982]. The results demonstrate the advantages in terms of bias correction of the nonlinear filter.

The nonlinear system considered is an equivalent to the ship speed equation. The parameters identified will, for the real ship, mean hull resistance and efficiency in utilizing the prime mover of the vessel for forward thrust. Both values are of major technical importance and as they change over time, they have vast impact on the ships's fuel economy and efficiency. The criteria for maintenence of the ship's hull, propeller, and prime mover system can be directly derived from these parameters, and it is hence of prime importance that they are estimated without bias.

$$\frac{d}{dt}x(t) = a\,x^2(t) + b\,u(t) + v(t) \tag{11}$$

The second order nonlinearity type of system is furthermore technically important when identifying propulsion losses of ships at sea aiming at autopilot and steering gear performance evaluation, Blanke [1981], Blanke and Sørensen [1984], Blanke [1986].

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

The matrices B_x, $\overline{W}_x^*$, $\overline{M}_\theta$, and N in the algorithm (10-a~l) corresponding to the example will be

$$\begin{aligned} B_x(t|t_i) &= a\,p(t_i) = a\,\hat{\Lambda}(t_i) \\ \overline{W}_x^*(t|t_i) &= 2a\hat{x}(t_i)w(t|t_i) \\ \overline{M}_\theta(t_i) &= [(\hat{x}^2(t_i) + P(t_i)),u(t_i),o] = [(\hat{x}^2(t_i) + \hat{\Lambda}(t_i)),u(t_i),o] \\ N(t_i) &= [o,o,\varepsilon(t_i)] \end{aligned} \tag{12}$$

Figure 1 shows results of identifying the parameters a and b in the nonlinear equation using the nonlinear filter. The curves plotted in figure 2 illustrate the performance of a linear RPE filter applied to the same nonlinear equation. Although the driving signal's pertubation is only 10 percent of its average, the bias of the linear estimator is apparent, and the superior performance of the nonlinear filter is obvious.

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 persued in a further study.

6. References

Åström, K.J.: Introduction to Stochastic Control Theory. Academic Press, 1970

Blanke, M.: Propulsion losses Related to Automatic Steering and Prime Mover Control. PhD thesis. Technical Univ. of Denmark, Dec. 1981.

Blanke, M.: Cross-Bispectrum Technique Identification of Nonlinear Ship Speed Dynamics. International Journal of Computers and Control, Vol. 2, 1986 pp. 54-62.

Blanke, M. & L.S. Sørensen: The Ljung Innovations Filter Used for Identification of Nonlinear Ship Speed Dynamics. 7th Ship Control Systems Symposium, Bath, U.K., 1984

Eykhoff, P.: System Identification - Parameter and State Estimation. John Wiley & Sons, 1977

Gavel, D.T. & S.G. Azevedo: Identification of Continuous Time Systems - An Application of Ljung's Corrected Extended Kalman Filter Sixth. IFAC Symposium on Identification and System Parameter Estimation, USA, June, 1982

Gelb, A.:Applied Optimal Estimation, MIT Press, 1974

Goodwin, G.C. & R.L. Payne: Dynamic System Identification: Experiment Design and Data Analysis. Academic Press, 1977

Jazwinski, A.H.: Stochastic Processes and Filtering Theory. Academic Press, 1970

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

Ljung, L. & T. Söderström: Theory and Practice of Recursive Identification. The MIT Press, 1983

Ljung, L.: Analysis of a General Recursive Prediction Error Identification Algorithm. Automatica, Vol. 17, No. 1, pp 89-99, 1981

Ljung, L.: Identification Methods ,Model Validation Recursive Identification Methods for Off-line Identification Problems. Sixth IFAC Symposium on Identification and System Parameter Estimation, USA, June, 1982

Ljung, L.: Analysis of Recursive Stochastic Algorithms. IEEE Trans. Automatic Control, Vol. 22, No. 4, 1977

Ljung, L.: Asymptotic Behavior of the Extended Kalman Filter as a Parameter Estimator for Linear Systems.

Maybeck, P.S.: Stochastic Models, Estimation, and Control. Volume 1, Academic Press, 1979 Volume 2, Academic Press, 1982

McGarty, T.P.: Stochastic Systems & State Estimation. John Wiley & Sons, 1974

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

Zhou, W.-W.: Filtering and Recursive Identification. Servolaboraty, Technical University of Denmark, October 1985.

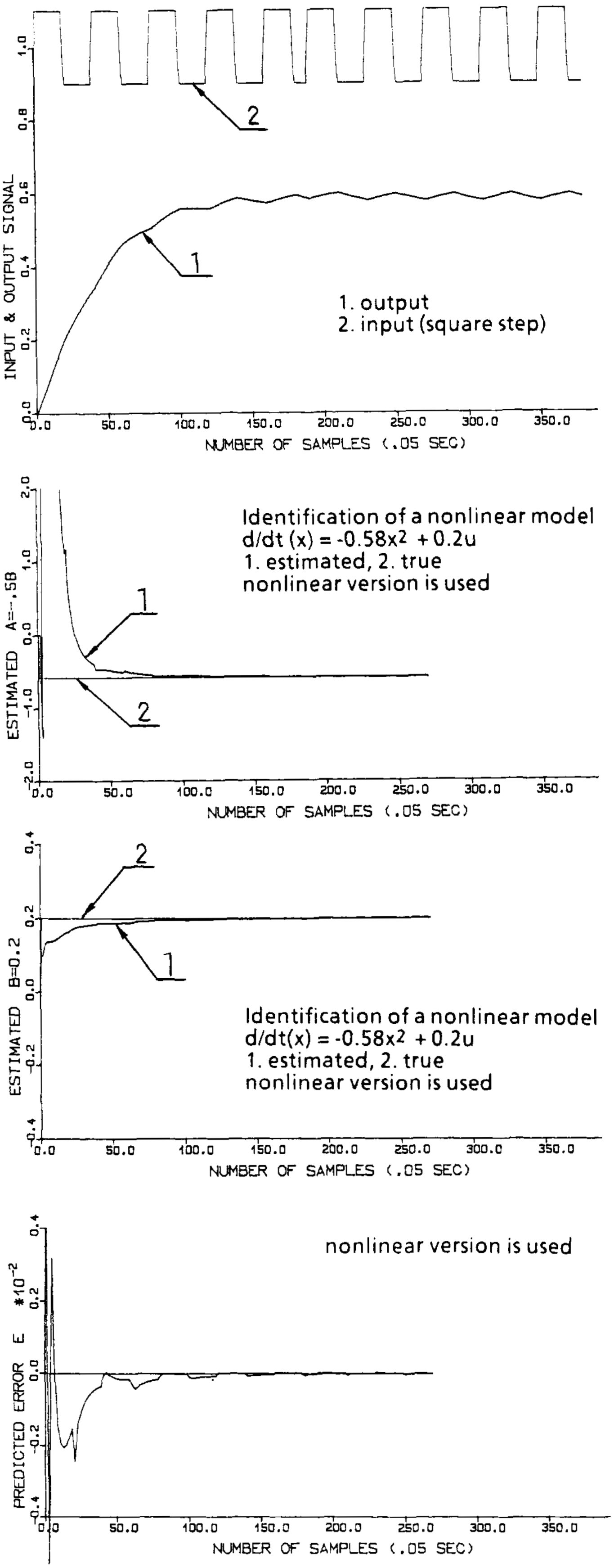

Figure 1 Identification of parameters a and b of equation 11 with square wave pertubation on the input signal.
The nonlinear estimator is used.

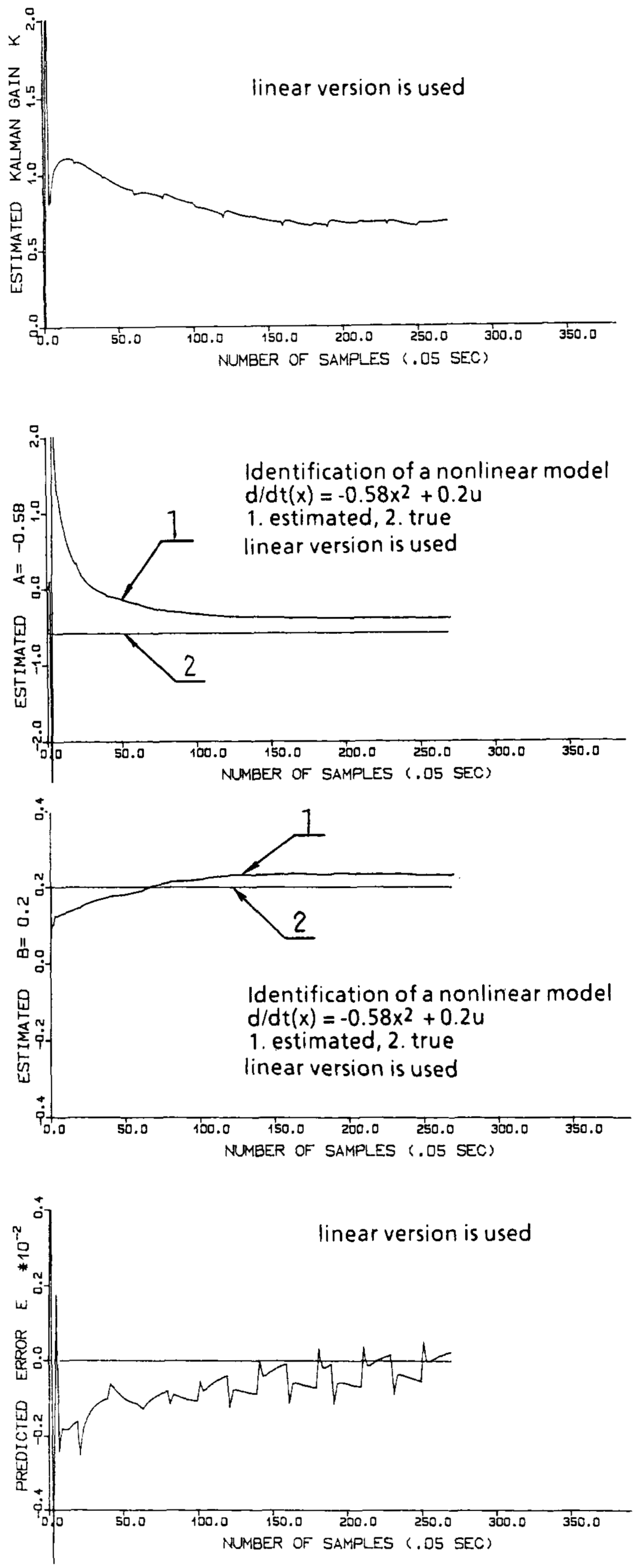

Figure 2 Identication of parameters a and b of equation 11 with similar excitation as in figure 1.
The nonlinear estimator is used.

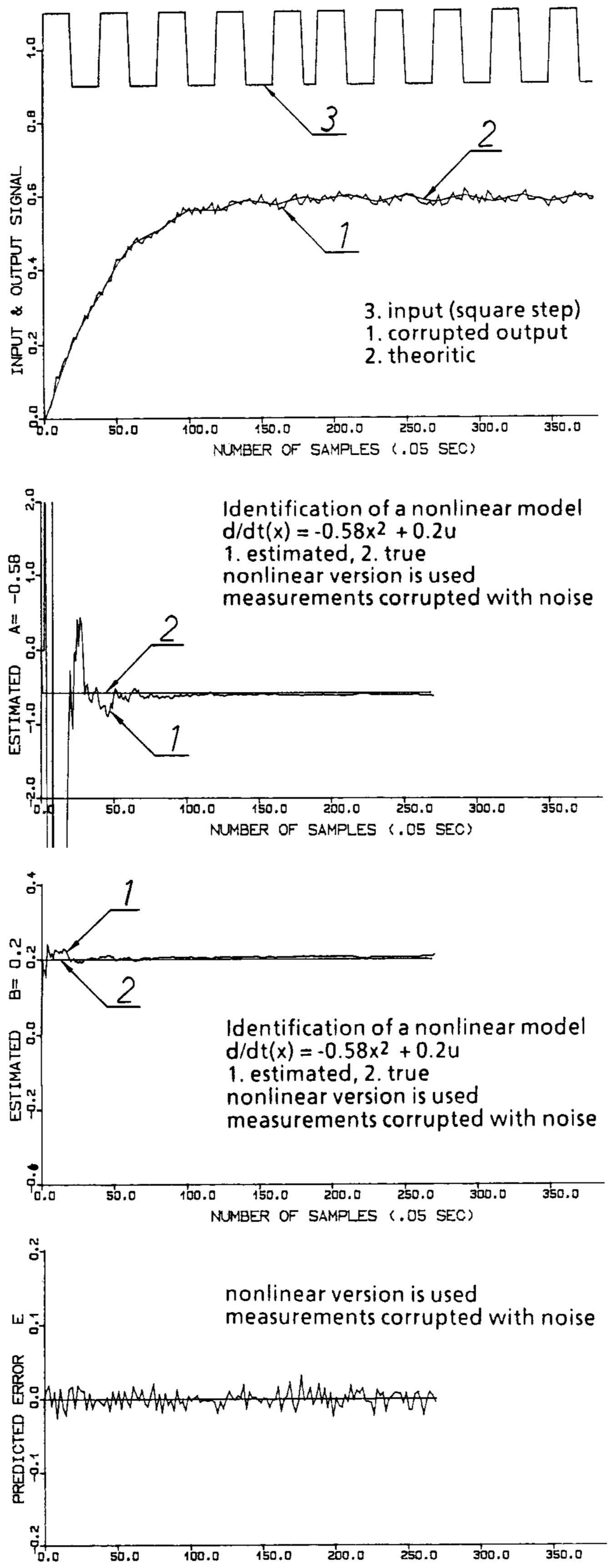

Figure 3 Same example as figure 1 with the nonlinear filter, but measurement corrupted with noise.

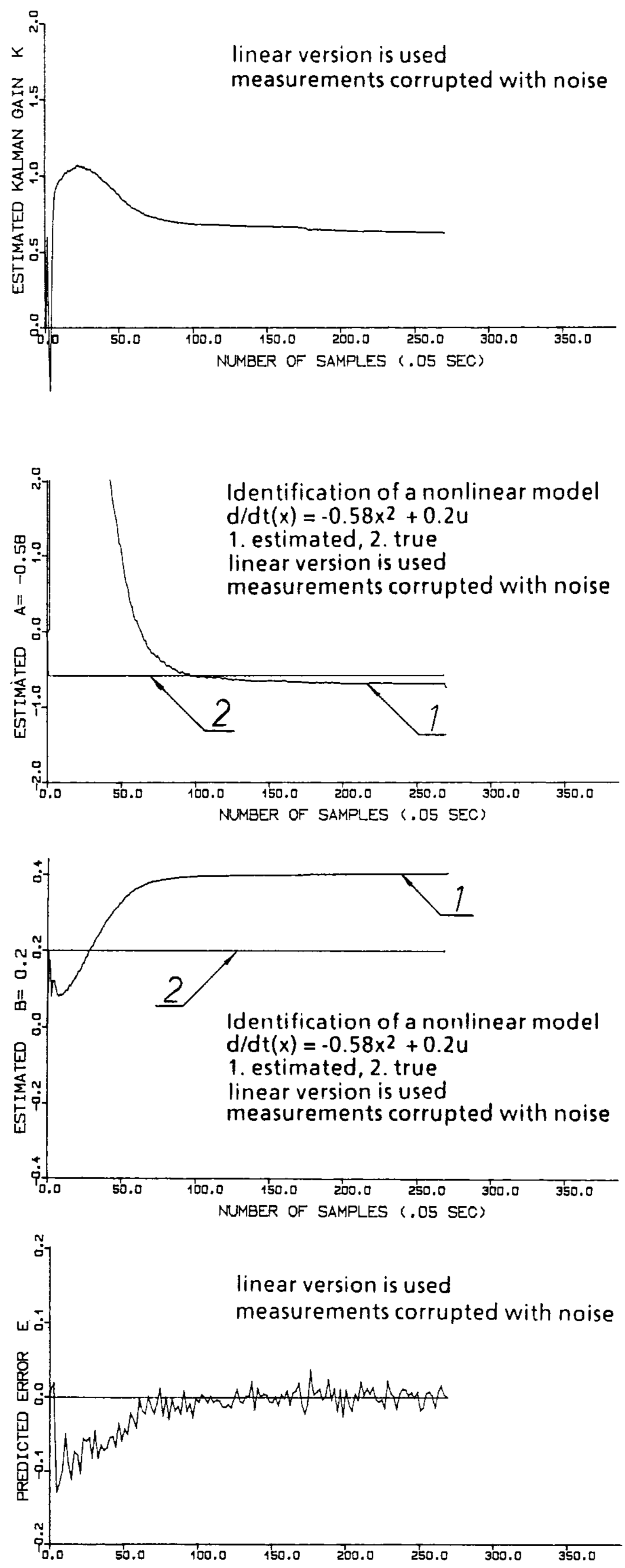

Figure 4 Same example as figure 2 with the linear filter, but measurement corrupted with noise