

FRACTIONAL KALMAN FILTER ALGORITHM FOR THE STATES, PARAMETERS AND ORDER OF FRACTIONAL SYSTEM ESTIMATION

DOMINIK SIEROCIUK, ANDRZEJ DZIELIŃSKI

Institute of Control and Industrial Electronics, Faculty of Electrical Engineering
Warsaw University of Technology,
ul. Koszykowa 75, 00-662 Warsaw, Poland
e-mail: {dsieroci, adziel}@isep.pw.edu.pl

This paper presents a generalization of the Kalman filter for linear and nonlinear fractional order discrete state-space systems. Linear and nonlinear discrete fractional order state-space systems are also introduced. The simplified kalman filter for the linear case is called the fractional Kalman filter and its nonlinear extension is named the extended fractional Kalman filter. The background and motivations for using such techniques are given, and some algorithms are discussed. The paper also shows a simple numerical example of linear state estimation. Finally, as an example of nonlinear estimation, the paper discusses the possibility of using these algorithms for parameters and fractional order estimation for fractional order systems. Numerical examples of the use of these algorithms in a general nonlinear case are presented.

Keywords: discrete fractional state-space systems, fractional Kalman filter, parameters estimation, order estimation, extended fractional Kalman filter

1. Introduction

The idea of fractional calculus (a generalization of the traditional integer order integral and differential calculus) was mentioned in 1695 by Leibniz and L'Hospital. At the end of the 19-th century, Liouville and Riemann introduced the first definition of the fractional derivative. However, this idea started to be interesting for engineers only in the late 1960s, especially when it was observed that the description of some systems is more accurate when the fractional derivative is used. (For example, modeling the behavior of some materials like polymers and rubber, and especially macroscopic properties of materials with a very complicated microscopic structure (Bologna and Grigolini, 2003)). In (Sjöberg and Kari, 2002), the frequency dependence of the dynamics of the rubber isolator is modeled with success by a fractional calculus element. In (Reyes-Melo *et al.*, 2004a; 2004b), the relaxation phenomena of organic dielectric materials such as semi-crystalline polymers are successfully modeled by mechanical and dielectric fractional models. Relaxation processes in organic dielectric materials are associated with molecular motions into new structural equilibrium of less energy. The Lagrangian and Hamiltonian mechanics can be reformulated to include fractional order derivatives. This leads directly to equations of motion with nonconservative forces such as friction (Riewe, 1997).

In (Vinagre and Feliu, 2002), the electrochemical processes and flexible robot arm are modeled by fractional order models. Even for modeling traffic in information networks, fractional calculus is found to be a useful tool (Zaborovsky and Meylanov, 2001). More examples and areas of using fractional calculus (e.g. fractal modeling, Brownian motion, rheology, viscoelasticity, thermodynamics and others) are to be found in (Bologna and Grigolini, 2003; Hilfer, 2000). In (Moshrefi-Torbati and Hammond, 1998; Podlubny, 2002), some geometrical and physical interpretations of fractional calculus are presented.

Another area of interest for engineers which is developing very fast is the use of fractional order controllers, like $PI^\lambda D^\mu$ controllers (Podlubny *et al.*, 1997) or CRONE (Oustaloup, 1993). The $PI^\lambda D^\mu$ controller has both the differentiation and integration of fractional order, which gives an extra ability to tune control systems. In (Suarez *et al.*, 2003), the fractional PID controller is used to path-tracking problem of an industrial vehicle. In (Ferreira and Machado, 2003), fractional order algorithms are applied to position/force hybrid control of robotic manipulators.

It is also worth mentioning that fractional order polynomials, used in the analysis of discrete-time control systems, may be treated as nD linear systems (Gałkowski, 2005).

We have developed a fractional order dynamic model as a very useful tool for modeling some electrodynamic and electrothermal process. The model allows us to introduce nonlinear effects like friction and slipping in an easier way than any other dynamic model of integer order does. This model forms a basis for model-based state feedback control. In order to use the state-feedback control, when state variables are not directly measured from the plant, new estimation tools appropriate for fractional order models (FKF) are needed. When model parameters are unknown, the parameter/state estimation problem occurs. To solve this problem in these case one needs estimation tools suitable for nonlinear fractional order models (EFKF).

The identification of parameters in fractional order systems, and especially the fractional orders of these systems, is not as easy as in the case of integer order systems (because of a high nonlinearity). There are several algorithms trying to solve this problem, most of them using frequency domain methods (Vinagre and Feliu, 2002). In (Cois *et al.*, 2000), a time domain parametric identification of a non-integer order system is presented. In (Cois *et al.*, 2001), also the time domain approach is presented by using a fractional state variable filter.

The article is organized as follows: In Section 2, the fractional order model is introduced. The generalization of the Kalman filter for fractional order systems is presented in Section 3. Section 4 shows a basic example of state estimation. In Section 4.1, examples of realizations of fractional order state space systems and a fractional Kalman filter are presented and studied. The nonlinear fractional order model and the extended fractional Kalman filter are introduced in Section 5. Examples of nonlinear estimation regarding parameters and the fractional order are shown in Sections 6 and 7, respectively.

2. Fractional Calculus

In this paper, as a definition of the fractional discrete derivative, the Grünwald-Letnikov definition (Oldham and Spanier, 1974; Podlubny, 1999) will be used.

Definition 1. The *fractional order Grünwald-Letnikov difference* is given by

$$\Delta^n x_k = \frac{1}{h^n} \sum_{j=0}^k (-1)^j \binom{n}{j} x_{k-j}, \quad (1)$$

where $n \in \mathbb{R}$ is the order of the fractional difference, \mathbb{R} is the set of real numbers, h is the sampling interval, later assumed to be 1, and k is the number of samples for which the derivative is calculated. The factor $\binom{n}{j}$ can be

obtained from

$$\binom{n}{j} = \begin{cases} 1 & \text{for } j = 0, \\ \frac{n(n-1)\dots(n-j+1)}{j!} & \text{for } j > 0. \end{cases} \quad (2)$$

According to this definition, it is possible to obtain a discrete equivalent of the derivative (when n is positive), a discrete equivalent of integration (when n is negative) or, when n equals 0, the original function. More properties of the definition can be found in (Jun, 2001; Ostalczyk, 2000; 2004a; 2004b).

Now, we wish to present a generalization of the discrete state space model for fractional order derivatives, which will be used later. Let us assume a traditional (integer order) discrete linear stochastic state-space system

$$x_{k+1} = Ax_k + Bu_k + \omega_k, \quad (3)$$

$$y_k = Cx_k + \nu_k, \quad (4)$$

where x_k is the state vector, u_k is the system input, y_k is the system output, ω_k is the system noise and ν_k is the output noise at the time instant k .

Equation (3) can be rewritten as follows:

$$\Delta^1 x_{k+1} = A_d x_k + Bu_k + \omega_k,$$

where $\Delta^1 x_k$ is the first-order difference for the sample x_k , $A_d = A - I$ (where I is the identity matrix), and

$$\Delta^1 x_{k+1} = x_{k+1} - x_k.$$

The value of the space vector for the time instance $k + 1$ can be obtained from

$$x_{k+1} = \Delta^1 x_{k+1} + x_k.$$

Using this formula, the traditional discrete linear stochastic state-space system can be rewritten as follows:

$$\Delta^1 x_{k+1} = A_d x_k + Bu_k + \omega_k, \quad (5)$$

$$x_{k+1} = \Delta^1 x_{k+1} + x_k, \quad (6)$$

$$y_k = Cx_k + \nu_k. \quad (7)$$

In (5), the value of the state difference is calculated, and from this value the next state vector is obtained according to (6). The output equation (7) has the same form as (4).

The first-order difference can be generalized to the difference of any even noninteger order, according to Definition 1. In this way, the following discrete stochastic state-space system is introduced:

Definition 2. The discrete linear fractional order stochastic system in a state-space representation is given by

$$\Delta^n x_{k+1} = A_d x_k + B u_k + \omega_k, \quad (8)$$

$$x_{k+1} = \Delta^n x_{k+1} - \sum_{j=1}^{k+1} (-1)^j \binom{n}{j} x_{k+1-j}, \quad (9)$$

$$y_k = C x_k + \nu_k. \quad (10)$$

For the case when equation orders are not identical, the following generalized definition is introduced by analogy:

Definition 3. The generalized discrete linear fractional-order stochastic system in a state-space representation is given by

$$\Delta^\Upsilon x_{k+1} = A_d x_k + B u_k + \omega_k, \quad (11)$$

$$x_{k+1} = \Delta^\Upsilon x_{k+1} - \sum_{j=1}^{k+1} (-1)^j \Upsilon_j x_{k+1-j}, \quad (12)$$

$$y_k = C x_k + \nu_k, \quad (13)$$

where

$$\Upsilon_k = \text{diag} \left[\binom{n_1}{k} \quad \dots \quad \binom{n_N}{k} \right],$$

$$\Delta^\Upsilon x_{k+1} = \begin{bmatrix} \Delta^{n_1} x_{1,k+1} \\ \vdots \\ \Delta^{n_N} x_{N,k+1} \end{bmatrix},$$

n_1, \dots, n_N are the orders of system equations and N is the number of these equations.

3. Fractional Kalman Filter (FKF)

The Kalman filter is an optimal state vector estimator using the knowledge about the system model, input and output signals (Kalman, 1960). Estimation results are obtained by minimizing in each step the following cost function (Schutter *et al.*, 1999):

$$\hat{x}_k = \arg \min_x \left[(\tilde{x}_k - x) \tilde{P}_k^{-1} (\tilde{x}_k - x)^T + (y_k - Cx) R_k^{-1} (y_k - Cx)^T \right], \quad (14)$$

where

$$\tilde{x}_k = E[x_k | z_{k-1}^*] \quad (15)$$

is the state vector prediction at the time instant k , defined as the random variable x_k conditioned on the measurement stream z_{k-1}^* (Brown and Hwang, 1997).

In addition,

$$\hat{x}_k = E[x_k | z_k^*] \quad (16)$$

is the state vector estimate at the time instant k , defined as the random variable x_k conditioned on the measurement stream z_k^* . The measurement stream z_k^* contains the values of the measurement output y_0, y_1, \dots, y_k and the input signal u_0, u_1, \dots, u_k .

Furthermore,

$$\tilde{P}_k = E[(\tilde{x}_k - x_k)(\tilde{x}_k - x_k)^T] \quad (17)$$

is the prediction of the estimation error covariance matrix. The covariance matrix of the output noise ν_k in (13) is defined as

$$R_k = E[\nu_k \nu_k^T], \quad (18)$$

whereas the covariance matrix of the system noise ω_k in (11) (see Theorem 1 below) is defined as

$$Q_k = E[\omega_k \omega_k^T]. \quad (19)$$

Additionally,

$$P_k = E[(\hat{x}_k - x_k)(\hat{x}_k - x_k)^T] \quad (20)$$

is the estimation error covariance matrix.

All of those matrices are assumed to be symmetric.

Lemma 1. The state vector prediction \tilde{x}_{k+1} is given by

$$\begin{aligned} \Delta^\Upsilon \tilde{x}_{k+1} &= A_d \hat{x}_k + B u_k, \\ \tilde{x}_{k+1} &\cong \Delta^\Upsilon \tilde{x}_{k+1} - \sum_{j=1}^{k+1} (-1)^j \Upsilon_j \hat{x}_{k+1-j}. \end{aligned}$$

Proof. The state vector prediction presented here is obtained in much the same way as the state prediction in an integer order Kalman filter (Brown and Hwang, 1997; Haykin, 2001), where the state prediction is determined from the previous state estimate. We have

$$\begin{aligned} \tilde{x}_{k+1} &= E[x_{k+1} | z_k^*] \\ &= E \left[\left(A_d x_k + B u_k + \omega_k - \sum_{j=1}^{k+1} (-1)^j \Upsilon_j x_{k+1-j} \right) | z_k^* \right] \\ &= A_d E[x_k | z_k^*] + B u_k \\ &\quad - \sum_{j=1}^{k+1} (-1)^j \Upsilon_j E[x_{k+1-j} | z_k^*]. \end{aligned}$$

As regards the last term of the above equation, we may use the following simplifying assumption:

$$E[x_{k+1-j}, z_k^*] \cong E[x_{k+1-j}, z_{k+1-j}^*]$$

for $i = 1, \dots, k+1$. This assumption implies that the past state vector will not be updated using the newer data z_k . Using this assumption, the following relation is obtained:

$$\tilde{x}_{k+1} \cong A_d \hat{x}_k + B u_k - \sum_{j=1}^{k+1} (-1)^j \Upsilon_j \hat{x}_{k+1-j}.$$

But this is exactly the relation to be proved. ■

Theorem 1. *For the discrete fractional order stochastic system in the state-space representation introduced by Definition 3, the simplified Kalman filter (called the fractional Kalman filter) is given by the following set of equations:*

$$\Delta^\Upsilon \tilde{x}_{k+1} = A_d \hat{x}_k + B u_k, \quad (21)$$

$$\tilde{x}_{k+1} = \Delta^\Upsilon \tilde{x}_{k+1} - \sum_{j=1}^{k+1} (-1)^j \Upsilon_j \hat{x}_{k+1-j}, \quad (22)$$

$$\begin{aligned} \tilde{P}_k &= (A_d + \Upsilon_1) P_{k-1} (A_d + \Upsilon_1)^T \\ &+ Q_{k-1} + \sum_{j=2}^k \Upsilon_j P_{k-j} \Upsilon_j^T, \end{aligned} \quad (23)$$

$$\hat{x}_k = \tilde{x}_k + K_k (y_k - C \tilde{x}_k), \quad (24)$$

$$P_k = (I - K_k C) \tilde{P}_k, \quad (25)$$

where

$$K_k = \tilde{P}_k C^T (C \tilde{P}_k C^T + R_k)^{-1}$$

with the initial conditions

$$x_0 \in \mathbb{R}^N, \quad P_0 = E[(\tilde{x}_0 - x_0)(\tilde{x}_0 - x_0)^T].$$

Here ν_k and ω_k are assumed to be independent and with zero mean.

Proof. (a) Equations (21) and (22) follow directly from Lemma 1. The simplification used in the proof of Lemma 1 implies that the Kalman filter presented in Theorem 1 is only a suboptimal solution.

(b) To prove (24), the minimum of the cost function (14) has to be found. It is obtained by solving the following equation, in which the left-hand side is the first derivative of this function:

$$-2\tilde{P}_k^{-1}(\tilde{x}_k - \hat{x}_k) - 2C^T R_k^{-1}(y_k - C \hat{x}_k) = 0.$$

This yields

$$\hat{x}_k = (\tilde{P}_k^{-1} + C^T R_k^{-1} C)^{-1} (\tilde{P}_k^{-1} \tilde{x}_k + C^T R_k^{-1} y_k).$$

Using the Matrix Inversion Lemma, we get

$$\begin{aligned} \hat{x}_k &= (\tilde{P}_k - \tilde{P}_k C^T (C \tilde{P}_k C^T + R)^{-1} C \tilde{P}_k) \\ &\times (\tilde{P}_k^{-1} \tilde{x}_k + C^T R^{-1} y_k). \end{aligned}$$

Writing

$$K_k = \tilde{P}_k C^T (C \tilde{P}_k C^T + R_k)^{-1}, \quad (26)$$

which is called the Kalman filter gain vector, the following relation is obtained:

$$\hat{x}_k = \tilde{x}_k + \tilde{P}_k C^T R^{-1} y_k - K_k C - K_k C \tilde{P}_k C^T R^{-1} y_k.$$

This can be reduced using again (26) to finally produce the state estimation equation (24). We have

$$\hat{x}_k = \tilde{x}_k + K_k (y_k - C \tilde{x}_k).$$

As can be noticed, this equation is exactly the same as in the Kalman filter for integer order systems.

(c) The proof of (23) proceeds from (17). The term $\tilde{x}_k - x_k$ is calculated as

$$\begin{aligned} \tilde{x}_k - x_k &= A_d \hat{x}_{k-1} + B u_{k-1} - \sum_{j=1}^k [(-1)^j \Upsilon_j \hat{x}_{k-j}] \\ &- A_d x_{k-1} - B u_{k-1} - \omega_{k-1} \\ &+ \sum_{j=1}^k [(-1)^j \Upsilon_j x_{k-j}] \\ &= (A_d - \Upsilon_1)(\hat{x}_{k-1} - x_{k-1}) - \omega_{k-1} \\ &- \sum_{j=2}^k [(-1)^j \Upsilon_j (\hat{x}_{k-j} - x_{k-j})]. \end{aligned}$$

We assume the independence of each of noise sequences ω_k, ν_k in Theorem 1. Correlations of the terms $E[x_k x_j]$ for $k \neq j$ are very hard to determine and we assume that they do not have significant influence on the final results. That is why this correlation will be omitted in later expressions. This simplifying assumption, which will not be necessary when $E[\omega_k \omega_k^T] = 0$, implies that the expected values of the terms $(\hat{x}_l - x_l)(\hat{x}_m - x_m)^T$ are zero when $l \neq m$, which finally gives the following

equation:

$$\begin{aligned}
 \tilde{P}_k &= \mathbb{E}\left[(\tilde{x}_k - x_k)(\tilde{x}_k - x_k)^T\right] \\
 &= (A_d - \Upsilon_1)\mathbb{E}\left[(\hat{x}_{k-1} - x_{k-1})(\hat{x}_{k-1} - x_{k-1})^T\right] \\
 &\quad \times (A_d - \Upsilon_1)^T + \mathbb{E}[\omega_{k-1}\omega_{k-1}^T] \\
 &\quad + \sum_{j=2}^k \Upsilon_j \mathbb{E}\left[(\hat{x}_{k-j} - x_{k-j})(\hat{x}_{k-j} - x_{k-j})^T\right] \Upsilon_j^T \\
 &= (A_d + \Upsilon_1)P_{k-1}(A_d + \Upsilon_1)^T + Q_{k-1} \\
 &\quad + \sum_{j=2}^k \Upsilon_j P_{k-j} \Upsilon_j^T.
 \end{aligned}$$

As was shown, the prediction of the covariance error matrix depends on the values of the covariance matrices in previous time samples. This is the main difference in comparison with an integer order KF.

(d) To prove (25), the definition of the covariance error matrix in (20) is used. We get

$$\begin{aligned}
 P_k &= \mathbb{E}\left[(\hat{x}_k - x_k)(\hat{x}_k - x_k)^T\right] \\
 &= \mathbb{E}\left[(\tilde{x}_k + K_k(Cx_k + \nu_k - C\tilde{x}) - x_k) \right. \\
 &\quad \left. \times (\tilde{x}_k + K_k(Cx_k + \nu_k - C\tilde{x}) - x_k)^T\right] \\
 &= (I - K_k C)\mathbb{E}\left[(\tilde{x}_k - x_k)(\tilde{x}_k - x_k)^T\right] \\
 &\quad \times (I - K_k C)^T + K_k \mathbb{E}[\nu_k \nu_k^T] K_k^T \\
 &= (I - K_k C)\tilde{P}_k(I - K_k C)^T + K_k R_k K_k^T \\
 &= (I - K_k H_k)\tilde{P}_k \\
 &\quad + (-\tilde{P}_k H_k^T + K_k H_k \tilde{P}_k H_k^T + K_k R_k)K_k^T,
 \end{aligned}$$

which can be reduced using (26) and finally gives the relation (25):

$$P_k = (I - K_k C)\tilde{P}_k.$$

Again, there is no difference in comparison with the conventional KF. ■

The equations defined in Theorem 1 organize the recursive algorithm of the FKF. The algorithm starts from the initial values x_0 and P_0 , which represent our *a-priori* knowledge about the initial conditions of the estimated system. The matrix P_0 is usually a diagonal matrix with large entries, e.g., $100I$.

4. Example of State Estimation

In order to test the concept of the algorithm outlined in Section 3, let us try to estimate state variables of a system defined by the following matrices:

$$\begin{aligned}
 A_d &= \begin{bmatrix} 0 & 1 \\ -a_0 & -a_1 \end{bmatrix}, & B &= \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \\
 C &= \begin{bmatrix} b_0 & b_1 \end{bmatrix}, & \mathcal{N} &= \begin{bmatrix} n_1 & n_2 \end{bmatrix}^T,
 \end{aligned}$$

where

$$a_0 = 0.1, \quad a_1 = 0.2,$$

$$b_0 = 0.1, \quad b_1 = 0.3,$$

$$n_1 = 0.7, \quad n_2 = 1.2,$$

$$E[\nu_k \nu_k^T] = 0.3, \quad E[\omega_k \omega_k^T] = \begin{bmatrix} 0.3 & 0 \\ 0 & 0.3 \end{bmatrix}.$$

The parameters of the fractional Kalman filter used in the example are

$$\begin{aligned}
 P_0 &= \begin{bmatrix} 100 & 0 \\ 0 & 100 \end{bmatrix}, & Q &= \begin{bmatrix} 0.3 & 0 \\ 0 & 0.3 \end{bmatrix}, \\
 R &= \begin{bmatrix} 0.3 \end{bmatrix}.
 \end{aligned}$$

The results of the state estimation are shown in Fig. 2. As can be seen, the state variables were estimated with high accuracy. For comparison, in Fig. 1 the measured output is presented, based on which the estimates of the original states were obtained.

4.1. Practical Implementation

In practical realizations of discrete linear state-space systems, the number of elements in the sum in (12) has to be limited to a predefined value L . Equation (12) in this case has the following form:

$$x_{k+1} = \Delta^\Upsilon x_{k+1} - \sum_{j=1}^L (-1)^j \Upsilon_j x_{k-j+1}. \quad (27)$$

This simplification speeds up calculations and, in a real application, makes the calculus possible. However, it has an effect on the accuracy of the model realization (Sierociuk, 2005a). The example of using different L values is presented in Fig. 3. The system defined in Section 4 (without noise) is simulated for $L = 3, 6, 50, 200$. The square error of those realizations, as compared to a realization for $L = 200$ (which is the realization with the best accuracy in this case), is presented in Table 1. As can be

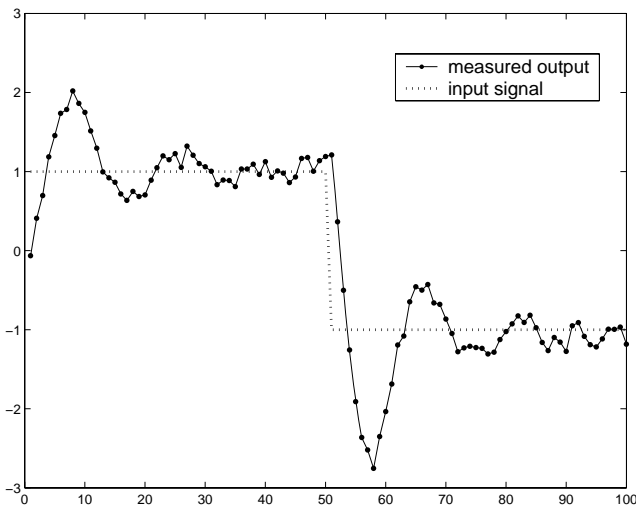


Fig. 1. Input and output signals from the plant.

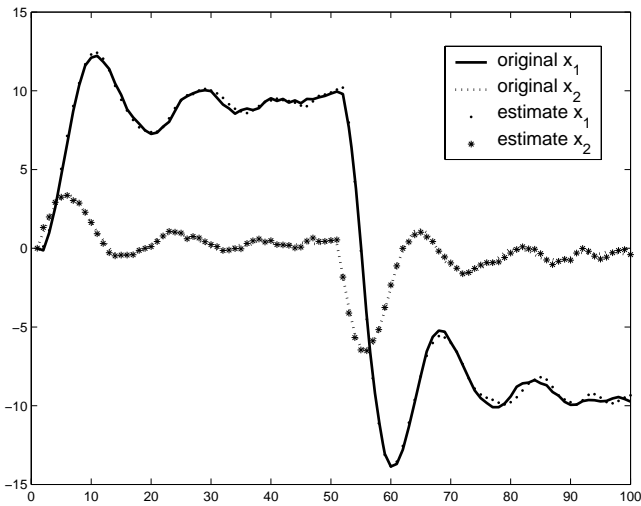


Fig. 2. Estimated and original state variables.

Table 1. Square error of realizations for different L .

L	error
3	10.173
6	0.81892
50	0.0025562

seen, the realization for $L = 50$ shows enough accuracy in that particular case. For different systems, the value L , which gives enough accuracy, may be different and depends on sampling times and system time constants.

In practical realizations of the fractional Kalman filter, the number of terms in the sums in (22) and (23) also has to be limited in the same way as in (27). The estimation results for different numbers of L are presented in

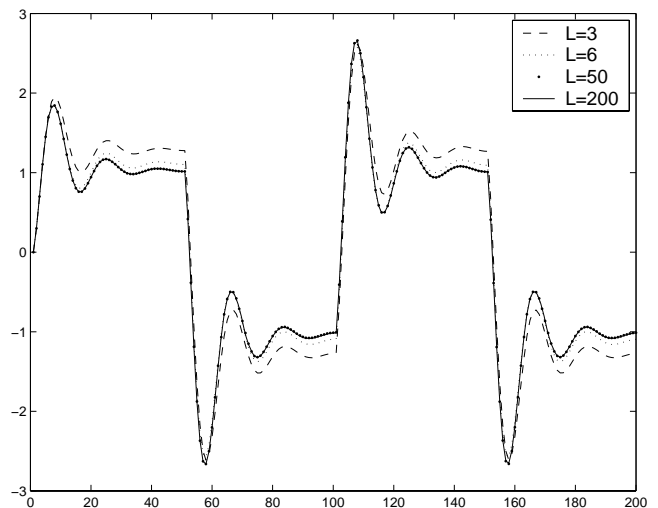


Fig. 3. Realizations of the fractional state space system for $L = 3, 6, 50, 200$.

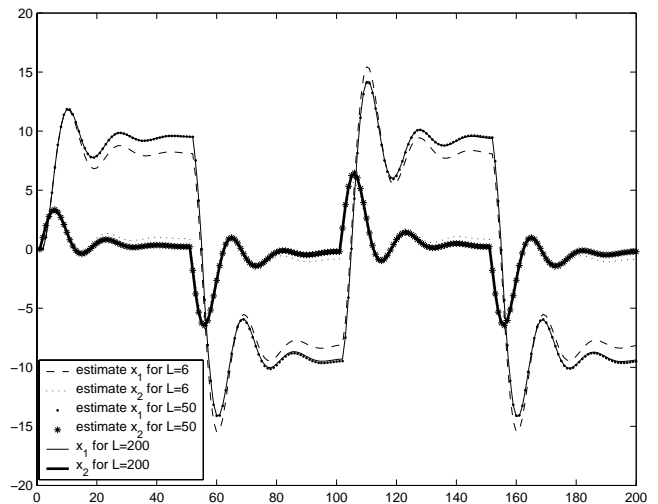


Fig. 4. Estimated and original state variables for different values of L .

Fig. 4. The square errors for $L = 6$ and $L = 50$ are equal to 247.0994 and 1.3819, respectively. Similarly to the system realization presented above, the realization of the FKF shows in this case enough accuracy for $L = 50$.

5. Nonlinear Estimation – Extended Fractional Kalman Filter

In previous sections, state estimation for a linear fractional order model was examined. In this section the same problem will be solved for a nonlinear fractional order model. The fractional-order nonlinear state-space system model is obtained analogously to the integer order one and defined as follows:

Definition 4. The nonlinear discrete stochastic fractional order system in a state-space representation is given by

$$\begin{aligned}\Delta^\Upsilon x_{k+1} &= f(x_k, u_k) + \omega_k, \\ x_{k+1} &= \Delta^\Upsilon x_{k+1} - \sum_{j=1}^{k+1} (-1)^j \Upsilon_j x_{k+1-j}, \\ y_k &= h(x_k) + \nu_k.\end{aligned}$$

The nonlinear functions $f(\cdot)$ and $h(\cdot)$, which are assumed to be of the class C^∞ , can be linearized according to the Taylor series expansion

$$f(x) = f(\tilde{x}) + \frac{\partial f(\tilde{x})}{\partial \tilde{x}}(\tilde{x} - x) + W, \quad (28)$$

where W stands for the higher order terms omitted in the linearization process.

In the previous section, the fractional Kalman filter for the linear model was presented. For the nonlinear model defined above, the fractional Kalman filter must be redefined in the same way as the extended Kalman filter for integer order models.

Lemma 2. The state vector prediction \tilde{x}_{k+1} for the system introduced by Definition 4 is given as

$$\begin{aligned}\Delta^\Upsilon \tilde{x}_{k+1} &= f(\hat{x}_k, u_k), \\ \tilde{x}_{k+1} &\cong \Delta^\Upsilon \tilde{x}_{k+1} - \sum_{j=1}^{k+1} (-1)^j \Upsilon_j \hat{x}_{k+1-j}.\end{aligned}$$

Proof. The state vector prediction presented in Lemma 2 is obtained in much the same way as the state prediction in the linear fractional Kalman filter in Lemma 1. We get

$$\begin{aligned}\tilde{x}_{k+1} &= \mathbb{E}[x_{k+1} | z_k^*] \\ &= \mathbb{E}\left[f(x_k, u_k) + \omega_k - \sum_{j=1}^{k+1} (-1)^j \Upsilon_j x_{k+1-j} \mid z_k^*\right].\end{aligned}$$

Linearizing $f(x_k, u_k)$ around the point \hat{x}_k according to (28), we obtain

$$\begin{aligned}\tilde{x}_{k+1} &= f(\hat{x}_k, u_k) - \frac{\partial f(\hat{x}_k, u_k)}{\partial \hat{x}_k}(\hat{x}_k - \mathbb{E}[x_k | z_k^*]) \\ &\quad - \sum_{j=1}^{k+1} (-1)^j \Upsilon_j \mathbb{E}[x_{k+1-j} | z_k^*].\end{aligned}$$

In the last term of the above equation, we may use the following simplifying assumption:

$$\mathbb{E}[x_{k+1-j}, z_k^*] \cong \mathbb{E}[x_{k+1-j}, z_{k+1-j}^*]$$

for $i = 1, \dots, k+1$.

This assumption implies that the past state vector will not be updated using newer data z_k and will not be necessary when $\mathbb{E}[\omega_k \omega_k^T] = 0$. Using this assumption, the following relation is obtained:

$$\tilde{x}_{k+1} \cong f(\hat{x}_k, u_k) - \sum_{j=1}^{k+1} (-1)^j \Upsilon_j \hat{x}_{k+1-j}.$$

This is exactly the relation to be proved. ■

Theorem 2. For the nonlinear discrete fractional order stochastic system in the state-space representation introduced by Definition 4, the extended fractional Kalman filter is given as follows:

$$\Delta^\Upsilon \tilde{x}_{k+1} = f(\hat{x}_k, u_k), \quad (29)$$

$$\tilde{x}_{k+1} = \Delta^\Upsilon \tilde{x}_{k+1} - \sum_{j=1}^{k+1} (-1)^j \Upsilon_j \hat{x}_{k+1-j}, \quad (30)$$

$$\begin{aligned}\tilde{P}_k &= (F_{k-1} + \Upsilon_1) P_{k-1} (F_{k-1} + \Upsilon_1)^T \\ &\quad + Q_{k-1} + \sum_{j=2}^k \Upsilon_j P_{k-j} \Upsilon_j^T, \quad (31)\end{aligned}$$

$$\hat{x}_k = \tilde{x}_k + K_k [y_k - h(\tilde{x}_k)], \quad (32)$$

$$P_k = (I - K_k H_k) \tilde{P}_k, \quad (33)$$

with the initial conditions

$$x_0 \in \mathbb{R}^N, \quad P_0 = \mathbb{E}[(\hat{x}_0 - x_0)(\hat{x}_0 - x_0)^T],$$

where

$$\begin{aligned}K_k &= \tilde{P}_k H_k^T (H_k \tilde{P}_k H_k^T + R_k)^{-1}, \\ F_{k-1} &= \left[\frac{\partial f(x, u_{k-1})}{\partial x} \right]_{x=\hat{x}_{k-1}}, \\ H_k &= \left[\frac{\partial h(x)}{\partial x} \right]_{x=\hat{x}_k},\end{aligned}$$

and the noise sequences ν_k and ω_k are assumed to be independent and zero mean.

Proof. (a) Equations (29) and (30) are defined in Lemma 2. The simplification used in the proof of Lemma 2 implies that the Kalman filter defined in Theorem 2 is only a suboptimal solution.

(b) To prove (32), the cost function (14) rewritten for the system given by Definition 4 has to be minimized. The cost function in that case has the form

$$\begin{aligned}\hat{x}_k &= \arg \min_x \left[(\tilde{x}_k - x) \tilde{P}_k^{-1} (\tilde{x}_k - x)^T \right. \\ &\quad \left. + (y_k - h(x)) R_k^{-1} (y_k - h(x))^T \right]. \quad (34)\end{aligned}$$

By expanding the nonlinear function $h(\cdot)$ to the Taylor series and omitting the higher order terms, the following expression is obtained:

$$\hat{x}_k = \arg \min_x \left[(\tilde{x}_k - x) \tilde{P}_k^{-1} (\tilde{x}_k - x)^T + \left(y_k - h(\tilde{x}_k) + \frac{\partial h(\tilde{x}_k)}{\partial \tilde{x}_k} (x_k - \tilde{x}_k) \right) R_k^{-1} \times \left(y_k - h(\tilde{x}_k) + \frac{\partial h(\tilde{x}_k)}{\partial \tilde{x}_k} (x_k - \tilde{x}_k) \right)^T \right].$$

Writing

$$H_k = \left[\frac{\partial h(x)}{\partial x} \right]_{x=\tilde{x}_k} \quad (35)$$

and equating the derivative of the cost function to zero, we get

$$-2\tilde{P}_k^{-1}(\tilde{x}_k - \hat{x}_k) - 2H_k^T R_k^{-1} [y_k - h(\tilde{x}_k) - H_k(x_k - \tilde{x}_k)] = 0.$$

According to the method presented in Section 3, using the Matrix Inversion Lemma and

$$K_k = \tilde{P}_k H^T (H \tilde{P}_k H^T + R_k)^{-1},$$

Eqn. (32) is concluded, i.e.,

$$\hat{x}_k = \tilde{x}_k + K(y_k - h(\tilde{x}_k)).$$

(c) The proof of (31) is similar to that of Theorem 1 (the linear case). It is obtained from (17).

The expression $\tilde{x}_k - x_k$ in (17) is calculated as follows:

$$\begin{aligned} \tilde{x}_k - x_k &= f(x_{k-1}, u_{k-1}) + \omega_{k-1} \\ &\quad - \sum_{j=1}^k (-1)^j \Upsilon_j x_{k-j} - f(\hat{x}_{k-1}, u_{k-1}) \\ &\quad + \sum_{j=1}^k (-1)^j \Upsilon_j \hat{x}_{k-j} \\ &= f(\hat{x}_{k-1}, u_{k-1}) + \omega_{k-1} \\ &\quad + \frac{\partial f(\hat{x}_{k-1}, u_{k-1})}{\partial \hat{x}_{k-1}} (x_{k-1} - \hat{x}_{k-1}) \\ &\quad - \sum_{j=1}^k (-1)^j \Upsilon_j x_{k-j} - f(\hat{x}_{k-1}, u_{k-1}) \\ &\quad + \sum_{j=1}^k (-1)^j \Upsilon_j \hat{x}_{k-j}. \end{aligned}$$

Denoting

$$F_{k-1} = \left[\frac{\partial f(x, u_{k-1})}{\partial x} \right]_{x=\hat{x}_{k-1}}, \quad (36)$$

the following expression is obtained:

$$\begin{aligned} \tilde{x}_k - x_k &= \omega_{k-1} - F_{k-1}(\hat{x}_{k-1} - x_{k-1}) \\ &\quad - \sum_{j=1}^k (-1)^j \Upsilon_j (\hat{x}_{k-j} - x_{k-j}). \end{aligned}$$

The independence of each noise ω_k, ν_k is assumed. The correlations of the terms $E[x_k x_j]$ for $k \neq j$ are very hard to determine and do not have significant influence on the final results. That is why this correlation will be omitted afterwards. This simplifying assumption, which will not be necessary when $E[\omega_k \omega_k^T] = 0$, implies that the expected values of the terms $(\hat{x}_l - x_l)(\hat{x}_m - x_m)^T$ are equal to zero when $l \neq m$, which finally gives the following equation:

$$\begin{aligned} \tilde{P}_k &= E[(\tilde{x}_k - x_k)(\tilde{x}_k - x_k)^T] \\ &= F_{k-1} E[(\hat{x}_{k-1} - x_{k-1})(\hat{x}_{k-1} - x_{k-1})^T] F_{k-1}^T \\ &\quad + E[\omega_{k-1} \omega_{k-1}^T] + \sum_{j=1}^k \Upsilon_j E[(\hat{x}_{k-j} - x_{k-j}) \\ &\quad \times (\hat{x}_{k-j} - x_{k-j})^T] \Upsilon_j^T. \end{aligned}$$

This directly leads to (31),

$$\begin{aligned} \tilde{P}_k &= (F_{k-1} + \Upsilon_1) P_{k-1} (F_{k-1} + \Upsilon_1)^T \\ &\quad + Q_{k-1} + \sum_{j=2}^k \Upsilon_j P_{k-j} \Upsilon_j^T. \end{aligned}$$

(d) To prove (33), the definition of the covariance error matrix in (20) is used. The expression $\hat{x}_k - x_k$ in this definition is evaluated as follows:

$$\begin{aligned} \hat{x}_k - x_k &= \tilde{x}_k + K_k (y_k - h(\tilde{x}_k)) - x_k \\ &= \tilde{x}_k + K_k \left[h(\tilde{x}_k) + \frac{\partial h(\tilde{x}_k)}{\partial \tilde{x}_k} (x_k - \tilde{x}_k) \right. \\ &\quad \left. + \omega_k - h(\tilde{x}_k) \right] - x_k \\ &= (I - K_k H_k)(\tilde{x}_k - x_k) + K_k \omega_k. \end{aligned}$$

Using the notation given by (35) and substituting the result in (20), the following relation is obtained:

$$\begin{aligned}
 P_k &= E\left[(\hat{x}_k - x_k)(\hat{x}_k - x_k)^T\right] \\
 &= (I - K_k H_k)E\left[(\tilde{x}_k - x_k)(\tilde{x}_k - x_k)^T\right] \\
 &\quad \times (I - K_k H_k)^T + K_k E[\omega_k \omega_k^T] K_k^T \\
 &= (I - K_k H_k) \tilde{P}_k (I - K_k H_k)^T + K_k R_k K_k^T \\
 &= (I - K_k H_k) \tilde{P}_k \\
 &\quad + \left(-\tilde{P}_k H_k^T + K_k H_k \tilde{P}_k H_k^T + K_k R_k\right) K_k^T,
 \end{aligned}$$

which finally gives (33). ■

6. Example of a Nonlinear Estimation – Parameter Estimation

When some parameters of the model are unknown or vary, it is possible to estimate them together with state variables. This is obtained by joining together state variables and estimated parameters in one state vector $x^w = [x^T w^T]^T$, where x^w is the new state vector and w is the vector containing the estimated parameters. This method is called *joint estimation* and leads to a nonlinear system.

For the system defined in Section 4 and the estimated parameter a_1 , nonlinear system equations are given as follows:

$$\begin{aligned}
 x_k^w &= [x_k^T a_1]^T, \\
 \Delta^\Upsilon x_{k+1}^w &= f(x_k^w, u_k) + \omega_k, \\
 x_{k+1}^w &= \Delta^\Upsilon x_{k+1}^w - \sum_{j=1}^{k+1} (-1)^j \Upsilon_j x_{k+1-j}^w, \\
 y_k &= h(x_k^w) + \nu_k,
 \end{aligned}$$

where

$$\begin{aligned}
 f(x_k^w, u_k) &= \begin{bmatrix} x_{2,k}^w \\ -a_0 x_{1,k}^w - a_1 x_{2,k}^w + u_k \\ 0 \end{bmatrix}, \\
 h(x_k) &= [b_0 x_{1,k}^w + b_1 x_{2,k}^w], \\
 \mathcal{N} &= [n_1 \quad n_2 \quad 1].
 \end{aligned}$$

Linearized matrices for the EKF are defined as

$$\begin{aligned}
 F_k &= \left[\frac{\partial F(x, u_k)}{\partial x} \right]_{x=\tilde{x}_k^w} = \begin{bmatrix} 0 & 1 & 0 \\ -a_0 & -a_1 & -\tilde{x}_{2,k}^w \\ 0 & 0 & 0 \end{bmatrix}, \\
 H_k &= \left[\frac{\partial H(x)}{\partial x} \right]_{x=\tilde{x}_k^w} = [b_0 \quad b_1 \quad 0].
 \end{aligned}$$

The parameters of the extended fractional Kalman filter used in the example are

$$\begin{aligned}
 P_0 &= \begin{bmatrix} 100 & 0 & 0 \\ 0 & 100 & 0 \\ 0 & 0 & 100 \end{bmatrix}, \\
 Q &= \begin{bmatrix} 0.3 & 0 & 0 \\ 0 & 0.3 & 0 \\ 0 & 0 & 0.0001 \end{bmatrix}, \\
 R &= [0.3].
 \end{aligned}$$

The results of the joint estimation are shown in Figs. 5 and 6. The final estimate of the parameter a_1 is equal to $a_1 = 0.2003$. The accuracy of the obtained results is very high. In addition to the parameters, estimated state variables are obtained which can be used, for example, to construct adaptive control algorithms.

7. System Order Estimation

The fractional order estimation problem is more complicated than parameter estimation. This is why the concept will be presented using a simpler model. Let us assume the discrete fractional linear system of the form

$$\Delta^n y_{k+1} = b u_k + \omega_k,$$

where b is a system parameter and ω is system noise.

In order to estimate the fractional order of the system defined above, the state vector and system equations are chosen as

$$x = [y, n]^T, \quad (37)$$

$$A_d = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0.3 \\ 0 \end{bmatrix}, \quad (38)$$

$$\Upsilon_j = \text{diag} \left[\begin{pmatrix} \hat{n}_{k-1} \\ j \end{pmatrix}, \begin{pmatrix} 1 \\ j \end{pmatrix} \right]. \quad (39)$$

We assume that the parameter b is known and in this example it is equal to $b = 0.3$.

There is a problem with such system matrices, because the FKF algorithm does not incorporate the knowledge of the fact that the order of the first state equation is

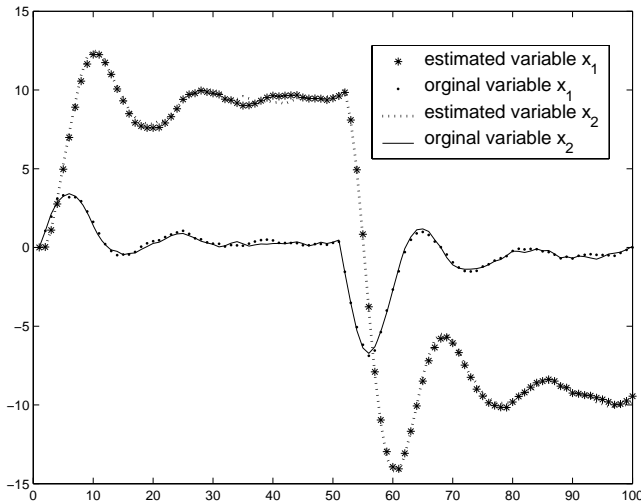


Fig. 5. Estimated and original state variables of the plant.

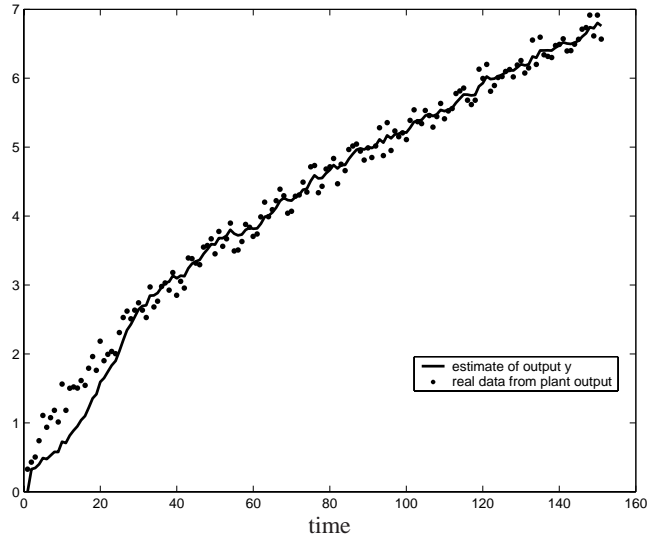


Fig. 7. Signal y – the real plant output and its estimate for order estimation.

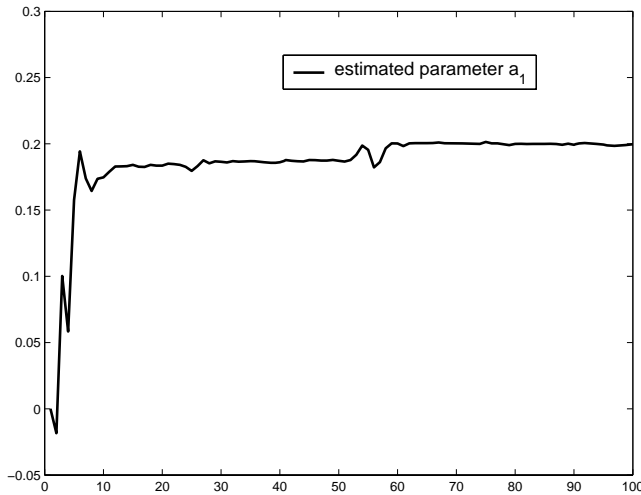


Fig. 6. Estimation of the parameter a_1 .

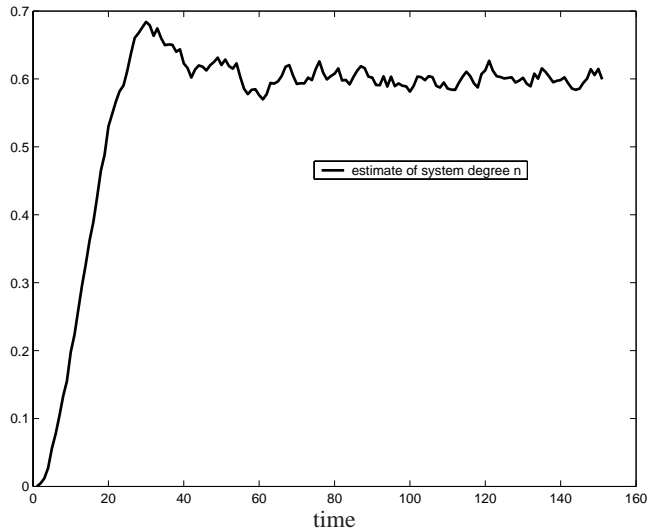


Fig. 8. Estimation of the order n .

an element n of the state vector. The knowledge is understood as the innovation of the prediction of the covariance matrix. Unfortunately, this dependence is very hard to resolve analytically. One of the solutions to this problem is to treat the dependence between the order of the first equation and the state variable n as noise and introduce in the matrix Q some value in the position representing this dependence.

In the following example, the matrix Q was defined as

$$Q = \begin{bmatrix} 0.55 & 0.09 \\ 0.09 & 0.1 \end{bmatrix}, \quad (40)$$

where a value of 0.09 corresponds to the additional noise described above.

The results of estimating the system order are shown in Figs. 7 and 8. Despite the simplification of the covariance matrix calculation, the final result $\hat{n} = 0.5994$, where the real value is $n = 0.6$, shows that this algorithm is useful.

In order to improve the results, the matrix Q can be changed according to the rule used in training neural networks by the KF algorithm. For example, the Robbins-Monro scheme (Haykin, 2001; Sum *et al.*, 1996),

$$Q_k = (1 - \alpha)Q_{k-1} + \alpha K_k (y_k - Hx_k)(y_k - Hx_k)^T K_k^T, \quad (41)$$

can be applied, where α is a small positive value. In this example, α is equal to 0.03.

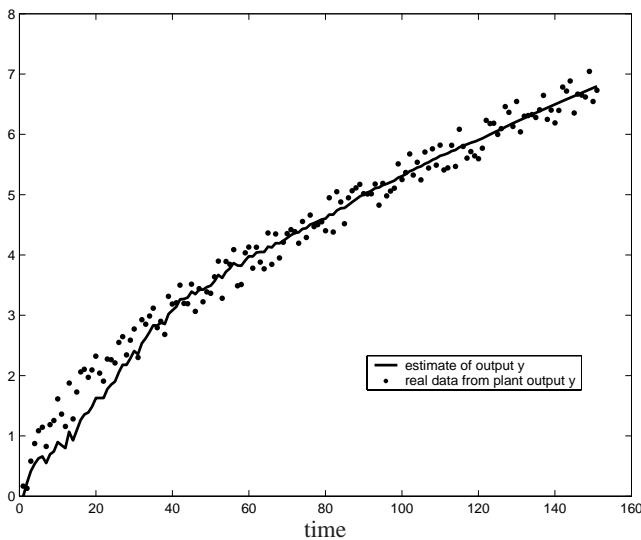


Fig. 9. Signal y – the real plant output and its estimate for order estimation using the Robbins-Monro scheme.

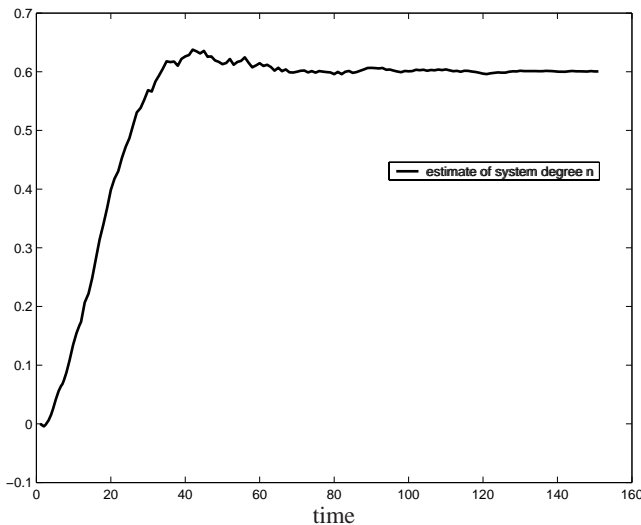


Fig. 10. Estimation of the order n with the Robbins-Monro scheme.

The noise of the output signal was increased in order to show better noise resistance for this algorithm. The other parameters are the same.

The results are shown in Figs. 9 and 10. It is easy to see that the learning rule which is used (41) improves the convergence, accuracy and robustness of the algorithm. The estimated order was equal to 0.6010.

8. Conclusions

The article presents the use of the Kalman filter algorithm for the estimation of parameters or the order of a fractional system. An example of parameter estimation

shows high accuracy of this estimation and its robustness to noise. This algorithm can also be used for the estimation of time-varying parameters, especially for adaptive control processes. The system order estimation problem was found to be more complicated. Despite the necessary simplifications of the algorithm, the obtained results are noise resistant. However, more studies and tests are needed. In particular, the sigma-point approach Kalman filters can be a more appropriate solution to this problem (Sierociuk, 2005b).

References

- Bologna M. and Grigolini P. (2003): *Physics of Fractal Operators*. — New York: Springer-Verlag.
- Brown R.G. and Hwang P.Y.C. (1997): *Introduction to Random Signals and Applied Kalman Filtering with Matlab Exercises and Solutions*. — New York: Wiley.
- Cois O., Oustaloup A., Battaglia E. and Battaglia J.-L. (2000): *Non-integer model from modal decomposition for time domain system identification*. — Proc. Symp. System Identification, SYSID 2000, Santa Barbara, CA, Vol. 3, p. 989–994.
- Cois O., Oustaloup A., Poinot T. and Battaglia J.-L. (2001): *Fractional state variable filter for system identification by fractional model*. — Proc. European Contr. Conf., ECC'2001, Porto, Portugal, pp. 2481–2486.
- Engheta N. (1997): *On the role of fractional calculus in electromagnetic theory*. — IEEE Trans. Antenn. Prop., Vol. 39, No. 4, pp. 35–46.
- Ferreira N.M.F. and Machado J.A.T. (2003): *Fractional-order hybrid control of robotic manipulators*. — Proc. 11-th Int. Conf. Advanced Robotics, ICAR'2003, Coimbra, Portugal, pp. 393–398.
- Gałkowski K. (2005): *Fractional polynomials and nD systems*. — Proc. IEEE Int. Symp. Circuits and Systems, IS-CAS'2005, Kobe, Japan, CD-ROM.
- Haykin S. (2001): *Kalman Filtering and Neural Networks*. — New York: Wiley.
- Hilfer R., Ed. (2000): *Application of Fractional Calculus in Physics*. — Singapore: World Scientific.
- Jun S.C. (2001): *A note on fractional differences based on a linear combination between forward and backward differences*. — Comput. Math. Applic., Vol. 41, No. 3, pp. 373–378.
- Kalman R.E. (1960): *A new approach to linear filtering and prediction problems*. — Trans. ASME J. Basic Eng., Vol. 82, Series D, pp. 35–45.
- Miller K.S. and Ross B. (1993): *An Introduction to the Fractional Calculus and Fractional Differential Equations*. — New York: Wiley.
- Moshrefi-Torbati M. and Hammond K. (1998): *Physical and geometrical interpretation of fractional operators*. — J. Franklin Inst., Vol. 335B, No. 6, pp. 1077–1086.

- Nishimoto K. (1984): *Fractional Calculus*. — Koriyama: De-cartes Press.
- Oldham K.B. and Spanier J. (1974): *The Fractional Calculus*. — New York: Academic Press.
- Ostalczyk P. (2000): *The non-integer difference of the discrete-time function and its application to the control system synthesis*. — Int. J. Syst. Sci., Vol. 31, No. 12, pp. 1551–1561.
- Ostalczyk P. (2004a): *Fractional-Order Backward Difference Equivalent Forms Part I – Horner’s Form*. — Proc. 1-st IFAC Workshop Fractional Differentiation and its Applications, FDA’04, Enseirb, Bordeaux, France, pp. 342–347.
- Ostalczyk P. (2004b): *Fractional-Order Backward Difference Equivalent Forms Part II – Polynomial Form*. — Proc. 1-st IFAC Workshop Fractional Differentiation and its Applications, FDA’04, Enseirb, Bordeaux, France, pp. 348–353.
- Oustaloup A. (1993): *Commande CRONE*. — Paris: Hermés.
- Oustaloup A. (1995): *La dérivation non entière*. — Paris: Hermés.
- Podlubny I. (1999): *Fractional Differential Equations*. — San Diego: Academic Press.
- Podlubny I. (2002): *Geometric and physical interpretation of fractional integration and fractional differentiation*. — Fract. Calc. Appl. Anal., Vol. 5, No. 4, pp. 367–386.
- Podlubny I., Dorcak L. and Kostial I. (1997): *On fractional derivatives, fractional-order systems and $PI^\lambda D^\mu$ -controllers*. — Proc. 36-th IEEE Conf. Decision and Control, San Diego, CA, pp. 4985–4990.
- Poinot T. and Trigeassou J.C. (2003): *Modelling and simulation of fractional systems using a non integer integrator*. — Proc. Design Engineering Technical Conferences, DETC’03, and Computers and Information in Engineering Conference, ASME 2003, Chicago, IL, pp. VIB-48390.
- Reyes-Melo M.E., Martinez-Vega J.J., Guerrero-Salazar C.A. and Ortiz-Mendez U. (2004a): *Application of fractional calculus to modelling of relaxation phenomena of organic dielectric materials*. — Proc. Int. Conf. Solid Dielectrics, Toulouse, France, Vol. 2, pp. 530–533.
- Reyes-Melo M.E., Martinez-Vega J.J., Guerrero-Salazar C.A. and Ortiz-Mendez U. (2004b): *Modelling of relaxation phenomena in organic dielectric materials. application of differential and integral operators of fractional order*. — J. Optoelect. Adv. Mat., Vol. 6, No. 3, pp. 1037–1043.
- Riewe F. (1997): *Mechanics with fractional derivatives*. — Phys. Rev. E, Vol. 55, No. 3, pp. 3581–3592.
- Riu D., Retière N. and Ivanés M. (2001): *Turbine generator modeling by non-integer order systems*. — Proc. IEEE Int. Electric Machines and Drives Conference, IEMDC 2001, Cambridge, MA, pp. 185–187.
- Samko S.G., Kilbas A.A. and Marichev O.I. (1993): *Fractional Integrals and Derivative. Theory and Applications*. — London: Gordon & Breach.
- Schutter J., Geeter J., Lefebvre T. and Bruynickx H. (1999): *Kalman filters: A tutorial*. — Available at <http://citeseer.ist.psu.edu/443226.html>
- Sierociuk D. (2005a): *Fractional order discrete state-space system Simulink toolkit user guide*. — Available at <http://www.ee.pw.edu.pl/dsieroci/fsst/fsst.htm>
- Sierociuk D. (2005b): *Application of a fractional Kalman filter to parameter estimation of a fractional-order system*. — Proc. 15-th Nat. Conf. Automatic Control, Warsaw, Poland, Vol. 1, pp. 89–94, (in Polish).
- Sjöberg M. and Kari L. (2002): *Non-linear behavior of a rubber isolator system using fractional derivatives*. — Vehicle Syst. Dynam., Vol. 37, No. 3, pp. 217–236.
- Suarez J.I., Vinagre B.M. and Chen Y.Q. (2003): *Spatial path tracking of an autonomous industrial vehicle using fractional order controllers*. — Proc. 11-th Int. Conf. Advanced Robotics, ICAR 2003, Coimbra, Portugal, pp. 405–410.
- Sum J.P.F., Leung C.S. and Chan L.W. (1996): *Extended Kalman filter in recurrent neural network training and pruning*. — Tech. Rep., Department of Computer Science and Engineering, Chinese University of Hong Kong, CS-TR-96-05.
- Vinagre B.M., Monje C.A. and Calderón A.J. (2002): *Fractional order systems and fractional order control actions*. — Lecture 3 IEEE CDC02 TW#2: Fractional calculus Applications in Automatic Control and Robotics’.
- Vinagre B.M. and Feliu V. (2002): *Modeling and control of dynamic system using fractional calculus: Application to electrochemical processes and flexible structures*. — Proc. 41-st IEEE Conf. Decision and Control, Las Vegas, NV, pp. 214–239.
- Zaborovsky V. and Meylanov R. (2001): *Informational network traffic model based on fractional calculus*. — Proc. Int. Conf. Info-tech and Info-net, ICII 2001, Beijing, China, Vol. 1, pp. 58–63.

Received: 10 April 2005
Revised: 8 December 2005