Fractional Kalman Filter algorithm for states, parameters and order of fractional system estimation

Dominik Sierociuk

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

Abstract

This paper presents a generalization of the Kalman Filter for linear and nonlinear fractional order discrete state-space systems. The linear and nonlinear discrete fractional order state-space systems are also introduced. The simplified Kalman Filter for linear case is called Fractional Kalman Filter and its nonlinear extension is named Extended Fractional Kalman Filter respectively. The background and motivation for using such techniques are given and the algorithms are derived and proved. The paper also shows a simple numerical example of linear state estimation. Finally, as an example of nonlinear estimation paper discusses the possibility of using these algorithms to parameters and fractional order estimation for fractional order systems. Numerical examples of the use of these algorithms in 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 fractional calculus (generalization of a traditional integer order integral and differential calculus) idea has been mentioned in 1695 by Leibniz and L'Hospital. In the end of 19th century Liouville and Riemann introduced first definition of fractional derivative. However, only just in late 60-ties of the 20th century this idea started to be interesting for engineers. Especially, when it was observed that the description of some systems is more accurate when the fractional derivative is used. For example, modeling of behavior of some materials like polymers and rubber, and especially macroscopic properties of materials with very complicated microscopic structure (Bologna & Grigolini 2003) . In (Sjöberg & 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. 2004b) and (Reyes-Melo et al. 2004a) the relaxation phenomena of organic dielectric materials such as the semi-crystalline polymers are successfully modeled by mechanical and dielectric fractional models. The relaxation processes in organic dielectric materials are associated with molecular motions to a 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 & Feliu 2002) the electrochemical processes and flexible robot arm are modeled by a fractional order models. Even for modeling of traffic in information network fractional calculus is found to be a useful tool (Zaborovsky & Meylanov 2001).

More examples and areas of using fractional calculus (eg. fractal modeling, Brownian motion, rheology, viscoelasticy, thermodynamics and others) are to be found in (Bologna & Grigolini 2003) and (Hilfer 2000).

In (Podlubny 2002) and (Moshrefi-Torbati & Hammond 1998) some geometrical and physical interpretation of fractional calculus are presented.

Another area of engineers interest, very fast developing, 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 differentiation and integration of fractional order, which gives extra ability to tune control systems. In (Suarez et al. 2003) the fractional PID controller is used to pathtracking problem of an industrial vehicle. In (Ferreira & Machado 2003) the fractional-order algorithms in position/force hybrid control of a robotic manipulators are applied.

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).

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

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

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

2 Fractional calculus

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

Definition 1 The fractional order Grünwald-Letnikov difference is given by the following equation

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

Where *n* is a fractional order and *h* is a sampling time later equal to 1, *k* is a number of sample for which the derivative is calculated. The factor $\binom{n}{j}$ can be obtained from relation:

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

According to this definition it is possible to obtain the discrete equivalent of derivative (when n is positive), the discrete equivalent of integration (when n is negative) or when n equal to 0 the original function.

More properties of the definition are to be found in (Ostalczyk 2000),(Ostalczyk 2004*a*),(Ostalczyk 2004*b*) and (Jun 2001).

Now the generalization of the discrete state space model for fractional order derivatives, which will be used later, is presented.

Let us assume a traditional (integer order) discrete linear stochastic state-space system,

$$x_{k+1} = Ax_k + Bu_k + \omega_k \tag{3}$$

$$y_k = Cx_k + \nu_k \tag{4}$$

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

Equation (3) could be rewritten as follows:

$$\Delta^1 x_{k+1} = A_d x_k + B u_k + \omega_k$$

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

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

Value of the space vector for time instance k+1 could be obtained from the relation,

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

Using this formula the traditional discrete linear stochastic state space system could be rewritten as follows

$$\Delta^1 x_{k+1} = A_d x_k + B u_k + \omega_k \tag{5}$$

$$x_{k+1} = \Delta^1 x_{k+1} + x_k \tag{6}$$

$$y_k = Cx_k + \nu_k \tag{7}$$

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

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

Definition 2 *The linear fractional order stochastic discrete state-space system is given by the following set of equations*

$$\Delta^n x_{k+1} = A_d x_k + B u_k + \omega_k \tag{8}$$

$$x_{k+1} = \Delta^{n} x_{k+1} - \sum_{k=1}^{k+1} (-1)^{j} {n \choose k} x_{k+1-i}$$
(9)

$$-\sum_{j=1}^{j} (-1)^{j} {\binom{j}{j}} x_{k+1-j} \qquad (9)$$

$$y_k = Cx_k + \nu_k \tag{10}$$

For the case when orders of equations are not identical, the following generalized definition is introduced analogically:

Definition 3 The generalized linear fractional order stochastic discrete state-space system is given by the following

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

$$-\sum_{j=1}^{k+1} (-1)^j \Upsilon_j x_{k+1-j} \qquad (12)$$

$$y_k = Cx_k + \nu_k \tag{13}$$

where

3

$$\Upsilon_{k} = \operatorname{diag} \begin{bmatrix} \binom{n_{1}}{k} & \dots & \binom{n_{N}}{k} \end{bmatrix}$$
$$\Delta^{\Upsilon} x_{k+1} = \begin{bmatrix} \Delta^{n_{1}} x_{1,k+1} \\ \vdots \\ \Delta^{n_{N}} x_{N,k+1} \end{bmatrix}$$

and $n_1...n_N$ are orders of system equations.

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). Results of estimation are obtained by minimizing in each step the following cost function (Schutter et al. 1999):

$$\hat{x}_{k} = \arg\min_{x} [(\tilde{x}_{k} - x)\tilde{P}_{k}^{-1}(\tilde{x}_{k} - x)^{T} + (y_{k} - Cx)R_{k}^{-1}(y_{k} - Cx)^{T}]$$
(14)

where

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

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

$$\hat{x}_k = \mathbf{E}[x_k | z_k^*] \tag{16}$$

is a state vector estimation at time instant k, defined as the random variable x_k conditioned on the measurement stream z_k^* .

The measurement stream z_k^* contain values of the measurement output y_0, y_1, \ldots, y_k and input signal u_0, u_1, \ldots, u_k .

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

is a prediction of an estimation error covariance matrix.

$$R_k = \mathbf{E} \left[\nu_k \nu_k^T \right] \tag{18}$$

is a covariance matrix of an output noise ν_k in (13).

$$Q_k = \mathbf{E} \left[\omega_k \omega_k^T \right] \tag{19}$$

is a covariance matrix of a system noise ω_k in (11) (see Theorem 1 below).

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

is an 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 the following relation

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

Proof:

The state vector prediction presented in Lemma 1 is obtained analogically to the state prediction in integer order Kalman Filter (Haykin 2001, Brown & Hwang 1997), where the state prediction is obtained from the previous state estimate.

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

In the last term of the equation above 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 causes that the past state vector will not be updated using newer data z_k . Using this assumption the following relation is obtained

$$\tilde{x}_{k+1} \cong A_d \hat{x}_k + B u_k$$

$$- \sum_{j=1}^{k+1} (-1)^j \Upsilon_j \hat{x}_{k+1-j}$$

This is exactly the relation to be proved. \Box

Theorem 1 For the fractional order stochastic discrete state-space system defined by Definition 3 the simplified Kalman Filter (called Fractional Kalman Filter) is given by the set of following equations

$$\Delta^{\Upsilon} \tilde{x}_{k+1} = A_d \hat{x}_k + B u_k$$

$$\tilde{x}_{k+1} = \Delta^{\Upsilon} \tilde{x}_{k+1}$$
(21)

$$-\sum_{j=1}^{k+1} (-1)^{j} \Upsilon_{j} \hat{x}_{k+1-j}$$
(22)

$$\tilde{P}_{k} = (A_{d} + \Upsilon_{1}) P_{k-1} (A_{d} + \Upsilon_{1})^{T}$$

+
$$Q_{k-1} + \sum_{j=2} \Upsilon_j P_{k-j} \Upsilon_j^T$$
 (23)

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

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

where

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

with initial conditions

$$x_0$$

 $P_0 = \mathbf{E}[(\tilde{x}_0 - x_0)(\tilde{x}_0 - x_0)^T]$

and ν_k and ω_k are assumed to be independent and with zero expected value.

Proof:

- a) The equations (21) and (22) follow directly from the Lemma 1. The simplification used in proof of Lemma 1 implies that the Kalman Filter defined in Theorem 1 is only the suboptimal solution.
- b) To proof the equation (24) the minimum of the cost function (14) has to be found. It is obtained by solving the following equation in which 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 Matrix Inversion Lemma one can find

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

Denoting

$$K_{k} = \tilde{P}_{k}C^{T}(C\tilde{P}_{k}C^{T} + R_{k})^{-1}$$
 (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 relation (26) and finally gives the state estimation equation (24).

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

As it could be seen this equation is exactly the same as in the Kalman Filter for integer order systems.

c) The proof of the equation (23) is developed from the equation (17).

The term $(\tilde{x}_k - x_k)$ is calculated as

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

The independence of each of noises ω_k, ν_k is assumed in Theorem 1. The 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 terms $(\hat{x}_l - x_l)(\hat{x}_m - x_m)^T$ are equal to zero when $l \neq m$, what finally gives the following equation:

$$\begin{split} \tilde{P}_{k} &= \mathrm{E}\left[(\tilde{x}_{k} - x_{k})(\tilde{x}_{k} - x_{k})^{T}\right] \\ &= (A_{d} - \Upsilon_{1})\mathrm{E}[(\hat{x}_{k-1} - x_{k-1}) \\ & (\hat{x}_{k-1} - x_{k-1})^{T}](A_{d} - \Upsilon_{1})^{T} \\ &+ \mathrm{E}[\omega_{k-1}\omega_{k-1}^{T}] + \sum_{j=2}^{k}\Upsilon_{j}\mathrm{E}[(\hat{x}_{k-j} - x_{k-j}) \\ & (\hat{x}_{k-j} - x_{k-j})^{T}]\Upsilon_{j}^{T} \\ &= (A_{d} + \Upsilon_{1})P_{k-1}(A_{d} + \Upsilon_{1})^{T} + Q_{k-1} \\ &+ \sum_{j=2}^{k}\Upsilon_{j}P_{k-j}\Upsilon_{j}^{T} \end{split}$$

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

d) To proof the equation (25) the definition of the covariance error matrix in equation (20) is used.

$$P_{k} = E[(\hat{x}_{k} - x_{k})(\hat{x}_{k} - x_{k})^{T}]$$

$$= E[(\tilde{x}_{k} + K_{k}(Cx_{k} + \nu_{k} - C\tilde{x}) - x_{k}) (\tilde{x}_{k} + K_{k}(Cx_{k} + \nu_{k} - C\tilde{x}) - x_{k})^{T}]$$

$$= (I - K_{k}C)E[(\tilde{x}_{k} - x_{k})(\tilde{x}_{k} - x_{k})^{T}] (I - K_{k}C)^{T} + K_{k}E[\nu_{k}\nu_{k}^{T}]K_{k}^{T}$$

$$= (I - K_{k}C)\tilde{P}_{k}(I - K_{k}C)^{T} + K_{k}R_{k}K_{k}$$

$$= (I - K_{k}H_{k})\tilde{P}_{k} + (-\tilde{P}_{k}H_{k}^{T} + K_{k}H_{k}\tilde{P}_{k}H_{k}^{T} + K_{k}R_{k})K_{k}^{T}$$

what 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 to conventional KF.

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 initial conditions of the estimated system. The matrix P_0 is usually a diagonal matrix with large entries eg. 100*I*.

4 Example of state estimation

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

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

where

$$a_0 = 0.1$$
 $a_1 = 0.2$
 $b_0 = 0.1$ $b_1 = 0.3$
 $n_1 = 0.7$ $n_2 = 1.2$

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

Fractional Kalman Filter parameters used in the example are:

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

Results of the state estimation are shown in Fig. 2. As it could 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 original states were obtained.

4.1 Practical realization

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



Figure 1: Input and output signals from the plant



Figure 2: Estimated and original state variables

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

This simplification speeds up calculus and in real application make the calculus possible. However it has an effect on the accuracy of the model realization (Sierociuk 2005*a*). The example of using different L values is presented in Fig. 3. The system defined in Section 4 (without noises) is simulated for L = 63, 6, 50, 200. The square error of those realizations in compare to realization for L = 200 (what is the ideal case) is presented in table 1. As it is could be seen the realization for L = 50 has enough accuracy in that particular case. For different systems the value L which

L	error
3	10.173
6	0.81892
50	0.0025562

Table 1: Square error of realizations for different L

gives enough accuracy could be different and depends on sample time and system time constants.



Figure 3: Realizations for L=3,6,50,200 of fractional state space system



Figure 4: Estimated and original state variables for different values of L

In practical realization of the Fractional Kalman Filter the number of terms in sums in equations (22) and (23) has to be also limited in the same way as in (27). The results of estimation for different numbers of L are presented in Fig. 4. The square error for L = 6 is equal to 247.0994 and for L = 50 to 1.3819. The same like for system realization presented above the realization of the FKF has in this case enough accuracy for L = 50. And also could be different for the different systems.

5 Nonlinear estimation - Extended Fractional Kalman Filter

In previous sections the state estimation for 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 analogically to the integer order one and defined as follows:

Definition 4 The nonlinear stochastic discrete fractional order state-space system is given by the following set of equations:

$$\Delta^{\Upsilon} x_{k+1} = f(x_k, u_k) + \omega_k$$

$$x_{k+1} = \Delta^{\Upsilon} x_{k+1}$$

$$- \sum_{j=1}^{k+1} (-1)^j \Upsilon_j x_{k+1-j}$$

$$y_k = h(x_k) + \nu_k$$

The nonlinear functions f(.) and h(.), which are assumed to be of C^{∞} class, could be linearized according to Taylor series expansion,

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

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

In 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 the integer order models.

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

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

$$\tilde{x}_{k+1} \cong \Delta^{\Upsilon} \tilde{x}_{k+1}$$

$$- \sum_{j=1}^{k+1} (-1)^j \Upsilon_j \hat{x}_{k+1-j}$$

Proof:

The state vector prediction presented in Lemma 2 is obtained analogically to the state prediction in linear Fractional Kalman Filter in Lemma 1.

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

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

$$\hat{x}_{k+1} = f(\hat{x}_k, u_k) - \frac{\partial f(\hat{x}_k, u_k)}{\partial \hat{x}_k} (\hat{x}_k - \mathbf{E}[x_k | z_k^*]) - \sum_{j=1}^{k+1} (-1)^j \Upsilon_j \mathbf{E}[x_{k+1-j} | z_k^*]$$

In the last term of the equation above 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 cause that the past state vector will not be updated using newer data z_k and will not be necessary when $E[\omega_k \omega_k^T] = 0$. 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 fractional order stochastic discrete state-space system given by the Definition 4 the Extended Fractional Kalman Filter is given by the following equations

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

$$\tilde{x}_{k+1} = \Delta^{\Upsilon} \tilde{x}_{k+1}$$
(29)

$$-\sum_{j=1}^{k+1} (-1)^j \Upsilon_j \hat{x}_{k+1-j}$$
(30)

$$\tilde{P}_{k} = (F_{k-1} + \Upsilon_{1}) P_{k-1} (F_{k-1} + \Upsilon_{1})^{T}$$

+
$$Q_{k-1} + \sum_{j=2} \Upsilon_j P_{k-j} \Upsilon_j^T$$
 (31)

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

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

with initial conditions

$$x_0$$

 $P_0 = \mathbf{E}[(\hat{x}_0 - x_0)(\hat{x}_0 - x_0)^T]$

where

$$K_{k} = \tilde{P}_{k}H_{k}^{T}(H_{k}\tilde{P}_{k}H_{k}^{T} + R_{k})^{-1}$$
$$F_{k-1} = \left[\frac{\partial f(x, u_{k-1})}{\partial x}\right]_{x=\hat{x}_{k-1}}$$
$$H_{k} = \left[\frac{\partial h(x)}{\partial x}\right]_{x=\tilde{x}_{k}}$$

and noises ν_k and ω_k are assumed to be independent and with zero expected value.

Proof:

- a) The equations (29) and (30) are defined in Lemma2. The simplification used in proof of Lemma 2 implies that the Kalman Filter defined in Theorem 2 is only the suboptimal solution.
- b) To proof the equation (32) the cost function (14) rewritten for the system given by the Definition 4 has to be minimized. The cost function in that case has the form

$$\hat{x}_{k} = \arg\min_{x} [(\tilde{x}_{k} - x)\tilde{P}_{k}^{-1}(\tilde{x}_{k} - x)^{T} + (y_{k} - h(x))R_{k}^{-1}(y_{k} - h(x))^{T}]$$
(34)

By expanding the nonlinear function h(.) to the Taylor series and omitting the higher order terms the following expression is obtained

$$\hat{x}_{k} = \arg \min_{x} \left[(\tilde{x}_{k} - x) \tilde{P}_{k}^{-1} (\tilde{x}_{k} - x)^{T} + \left(y_{k} - h(\tilde{x}_{k}) + \frac{\partial h(\tilde{x}_{k})}{\partial \tilde{x}_{k}} (x_{k} - \tilde{x}_{k}) \right) R_{k}^{-1} \right]$$
$$\left(y_{k} - h(\tilde{x}_{k}) + \frac{\partial h(\tilde{x}_{k})}{\partial \tilde{x}_{k}} (x_{k} - \tilde{x}_{k}) \right)^{T} \right]$$

Denoting

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

and equaling the derivative of the cost function to zero, the following expression is achieved.

$$- 2\tilde{P}_{k}^{-1}(\tilde{x}_{k} - \hat{x}_{k}) - 2H_{k}^{T}R_{k}^{-1}[y_{k} - h(\tilde{x}_{k}) - H_{k}(x_{k} - \tilde{x}_{k})] = 0$$

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

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

the equation (32) is concluded

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

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

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

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

Denoting

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

the following expression is obtained

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

The independence of each of noises ω_k, ν_k is assumed in Theorem 2. 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 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 terms $(\hat{x}_l - x_l)(\hat{x}_m - x_m)^T$ are equal to zero when $l \neq m$, what finally gives the following equation:

$$\begin{split} \tilde{P}_{k} &= \mathrm{E}[(\tilde{x}_{k} - x_{k})(\tilde{x}_{k} - x_{k})^{T}] \\ &= F_{k-1}\mathrm{E}[(\hat{x}_{k-1} - x_{k-1}) \\ & (\hat{x}_{k-1} - x_{k-1})^{T}]F_{k-1}^{T} + \mathrm{E}[\omega_{k-1}\omega_{k-1}^{T}] \\ &+ \sum_{j=1}^{k} \Upsilon_{j}\mathrm{E}[(\hat{x}_{k-j} - x_{k-j}) \\ & (\hat{x}_{k-j} - x_{k-j})^{T}]\Upsilon_{j}^{T} \end{split}$$

This leads directly to the equation (31)

$$\tilde{P}_{k} = (F_{k-1} + \Upsilon_{1}) P_{k-1} (F_{k-1} + \Upsilon_{1})^{T}$$

$$+ Q_{k-1} + \sum_{j=2}^{k} \Upsilon_{j} P_{k-j} \Upsilon_{j}^{T}$$

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

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

Using the notation given by equation (35) and substituting to the equation (20) the following relation are obtained

$$P_{k} = E[(\hat{x}_{k} - x_{k})(\hat{x}_{k} - x_{k})^{T}]$$

$$= (I - K_{k}H_{k})E[(\tilde{x}_{k} - x_{k})(\tilde{x}_{k} - x_{k})^{T}]$$

$$(I - K_{k}H_{k})^{T} + K_{k}E[\omega_{k}\omega_{k}^{T}]K_{k}$$

$$= (I - K_{k}H_{k})\tilde{P}_{k}(I - K_{k}H_{k})^{T}$$

$$+ K_{k}R_{k}K_{k}^{T}$$

$$= (I - K_{k}H_{k})\tilde{P}_{k} + (-\tilde{P}_{k}H_{k}^{T} + K_{k}H_{k}\tilde{P}_{k}H_{k}^{T} + K_{k}R_{k})K_{k}^{T}$$

what finally gives the equation (33)

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

6 Example of nonlinear estimation - parameters estimation

When the parameter or parameters of the model are unknown or are changing it is possible to estimate them together with state variables. It 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 a new state vector and w is a vector containing estimated parameters. This method is called Joint Estimation and leads to the nonlinear system.

For the system defined in Section 4 and estimated parameter a_1 , the 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

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

Linearized matrices for EFKF are defined as

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

Parameters of the Extended Fractional Kalman Filter used in the example are:

$$P_0 = \begin{bmatrix} 100 & 0 & 0 \\ 0 & 100 & 0 \\ 0 & 0 & 100 \end{bmatrix} Q = \begin{bmatrix} 0.3 & 0 & 0 \\ 0 & 0.3 & 0 \\ 0 & 0 & 0.0001 \end{bmatrix}$$

$$R = \left[\begin{array}{c} 0.3 \end{array} \right]$$

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



Figure 5: Estimated and original state variables of the plant



Figure 6: Estimation of a parameter a_1

7 System order estimation

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

$$\Delta^n y_{k+1} = bu_k + \omega_k$$

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

In order to estimate the fractional order of the system defined above the state vector and system equations could be chosen as:

$$x = [y, n]^T \tag{37}$$

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

$$\Upsilon_{j} = \operatorname{diag}\left[\left(\begin{array}{c}\hat{n}_{k-1}\\j\end{array}\right), \left(\begin{array}{c}1\\j\end{array}\right)\right]$$
(39)

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

For such system matrices there exists a problem, because the algorithm of FKF does not incorporate the knowledge on the fact that order of first state equation is an element n of state vector. Knowledge is understood as an innovation of prediction of covariance matrix. Unfortunately this dependency is very hard to resolve analytically. One of the solutions of this problem is to treat the dependency between the order of the first equation and the state variable n as a noise and introduce in matrix Q some value in the position representing this dependency.

In following example matrix Q was defined as

$$Q = \left[\begin{array}{cc} 0.55 & 0.09\\ 0.09 & 0.1 \end{array} \right] \tag{40}$$

Where value 0.09 corresponds to additional noise described above.

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

In order to improve the results matrix Q could be changed according to the rule used in training of neural



Figure 7: Signal y – real plant output and its estimate for order estimation



Figure 8: Estimation of order n

networks by KF algorithm. For example Robbins-Monro scheme (Haykin 2001) (Sum et al. 1996) given by the relation:

$$Q_{k} = (1 - \alpha)Q_{k-1} + \alpha K_{k}(y_{k} - Hx_{k})(y_{k} - Hx_{k})^{T}K_{k}^{T}$$
 (41)

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

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

Results are shown in Fig. 9 and 10. As it is clear to see, learning rule which is used (41) improves resulting



Figure 9: Signal y – real plant output and its estimate for order estimation with Robbins-Monro scheme



Figure 10: Estimation of order n with Robbins-Monro scheme

convergence and accuracy and it also improve robustness of the algorithm. The estimated order was equal to 0.6010.

8 Conclusions

The article presents use of the Kalman filter algorithm to the estimation of parameters or order of fractional system. Parameters estimation example shows high accuracy of this estimation and its robustness with respect to noise. This algorithm could be also used for estimation of time varying parameters, especially for the adaptive control processes. The system order estimation problem was found to be more complicated. Despite of the necessary simplifications of the algorithm, obtained results found it very useful and also noise resistant. However, more study and tests are needed. In particular, the sigma-point approach Kalman Filters could be a more appropriate solution for this problem (Sierociuk 2005*b*).

References

- Bologna, M. & P. Grigolini (2003), *Physics of Fractal Operators*, Springer-Verlag.
- Brown, Robert Grover & Patrick Y.C. Hwang (1997), Introduction to Random Signals and Applied Kalman Filtering. With Matlab Exercises and Solutions, John Wiley & Sons, Inc.
- Cois, O., A. Oustaloup, E. Battaglia & J.-L. Battaglia (2000), Non integer model from modal decomposition for time domain system identification, *in* 'Proceedings of Symposium on System Identification', Vol. 3, SYSID 2000, p. 989.
- Cois, O., A. Oustaloup, T. Poinot & J.-L. Battaglia (2001), Fractional state variable filter for system identification by fractional model, *in* 'Proceedings of European Control Conference', ECC'2001, Porto, Portugal.
- Engheta, N. (1997), 'On the role of fractional calculus in electromagnetic theory', *IEEE Antennas and Propagation* **39**(4).
- Ferreira, N. M. Fonseca & J. A. Tenreiro Machado (2003), Fractional-order hybrid control of robotic manipulators, *in* 'Proceedings of The 11th International Conference on Advanced Robotics', Coimbra, Portugal.
- Gałkowski, K. (2005), Fractional polynomials and nd systems, *in* 'Proceedings of IEEE Symposium on Circuits and Systems', ISCAS-2005, Kobe (Japan).
- Haykin, S. (2001), Kalman Filtering and Neural Networks, John Wiley & Sons, Inc.
- Hilfer, R., ed. (2000), *Application of Fractional Calculus in Physics*, World Scientific.

- Jun, S.C. (2001), 'A note on fractional differences based on a linear combination between forward and backward differences', *Computers and Mathematics with Applications* **41**, 373–378.
- Kalman, Rudolph, Emil (1960), 'A new approach to linear filtering and prediction problems', *Transactions of the ASME–Journal of Basic Engineering* 82(Series D), 35–45.
- Miller, K.S. & B. Ross (1993), An Introduction to the Fractional Calculus and Fractional Differenctial Equations, John Wiley & Sons Inc., New York.
- Moshrefi-Torbati, M. & K. Hammond (1998), 'Physical and geometrical interpretation of fractional operators', *Journal of the Franklin Institute* **335B**(6), 1077–1086.
- Nishimoto, K. (1984), *Fractional Calculus*, Decartes Press, Koriyama.
- Oldham, K. B. & J. Spanier (1974), *The Fractional Calculus*, Academic Press.
- Ostalczyk, P. (2000), 'The non-integer difference of the discrete-time function and its application to the control system synthesis', *International Journal of Systems Science* **31**(12), 1551–1561.
- Ostalczyk, Piotr (2004*a*), Fractional-Order Backward Difference Equivalent Forms Part I - Horner's Form, *in* 'Proceedings of 1st IFAC Workshop on Fractional Differentation and its Applications', FDA'04, Enseirb, Bordeaux, France.
- Ostalczyk, Piotr (2004*b*), Fractional-Order Backward Difference Equivalent Forms Part II - Polynomial Form, *in* 'Proceedings of 1st IFAC Workshop on Fractional Differentation and its Applications', FDA'04, Enseirb, Bordeaux, France.
- Oustaloup, A. (1993), *Commande CRONE*, Hermés, 'Paris.
- Oustaloup, A. (1995), *La dérivation non entiére*, Hermés, 'Paris.
- Podlubny, I. (1999), *Fractional Differential Equations*, Academic Press.
- Podlubny, I. (2002), 'Geometric and physical interpretation of fractional integration and fractional differentiation', *Fractional Calculus and Applied Analysis* **5**(4), 367–386.

- Podlubny, I., L. Dorcak & I. Kostial (1997), On fractional derivatives, fractional-order systems and $PI^{\lambda}D^{\mu}$ -controllers, *in* 'Proceedings of the $36^{t}h$ IEEE Conference on Decision and Control', San Diego, pp. 4985–4990.
- Poinot, T. & J.C. Trigeassou (2003), Modelling and simulation of fractional systems using a non integer integrator, *in* 'Proceedings of DETC.03 Design Engineering Technical Conferences and Computers and Information in Engineering Conference', ASME 2003, Chicago, Illinois, USA.
- Reyes-Melo, M.E., J.J. Martinez-Vega, C.A. Guerrero-Salazar & U. Ortiz-Mendez (2004*a*), Application of fractional calculus to modelling of relaxation phenomena of organic dielectric materials, *in* 'Proceedings of International Conference on Solid Dielectrics', Toulouse, France.
- Reyes-Melo, M.E., J.J. Martinez-Vega, C.A. Guerrero-Salazar & U. Ortiz-Mendez (2004b), 'Modelling of relaxation phenomena in organic dielectric materials. application of differential and integral operators of fractional order', *Journal of Optoelectronics and Advanced Materials* 6(3), 1037–1043.
- Riewe, F. (1997), 'Mechanics with fractional derivatives', *Physical Review E* **55**(3), 3581–3592.
- Riu, D., N. Retiére & M. Ivanés (2001), Turbine generator modeling by non-integer order systems, *in* 'Proceedings of IEEE International Electric Machines and Drives Conference', IEEE IEMDC 2001, Cambridge, Massachussets (USA).
- Samko, S.G., A.A. Kilbas & O.I. Maritchev (1987), Fractional Integrals and Derivative. Theory and Applications, Gordon & Breach Sci. Publishers.
- Schutter, J., J. Geeter, Tine Lefebvre & Herman Bruynickx (1999), 'Kalman filters: A tutorial'. *http://citeseer.ist.psu.edu/443226.html
- Sierociuk, Dominik (2005*a*), 'Fractional order discrete state-space system simulink toolkit user guide'. *http://www.ee.pw.edu.pl/ dsieroci/fsst/fsst.htm
- Sierociuk, Dominik (2005b), Użycie ułamkowego filtru kalmana do estymacji parametrów układu ułamkowego rzędu, *in* 'XV Krajowa Konferencja Automatyki', Warszawa, Polska.

- Sjöberg, M. & L. Kari (2002), 'Non-linear behavior of a rubber isolator system using fractional derivatives', *Vehicle System Dynamics* 37(3), 217 – 236.
- Suarez, J.I., B. M. Vinagre & Y. Q. Chen (2003), Spatial path tracking of an autonomous industrial vehicle using fractional order controllers, *in* 'Proceedings of the 11th International Conference on Advanced Robotics', ICAR 2003.
- Sum, J.P.F., C.S. Leung & L.W. Chan (1996), Extended kalman filter in recurrent neural network training and pruning, Technical report, Department of Computer Science and Engineering, The Chinese University of Hong Kong. CS-TR-96-05.
- Vinagre, B.M., C.A. Monje & A.J. Calderón (2002), Fractional order systems and fractional order control actions, *in* 'Lecture 3 of the IEEE CDC02 TW#2: Fractional Calculus Applications in Automatic Control and Robotics'.
- Vinagre, B.M. & V. Feliu (2002), Modeling and control of dynamic system using fractional calculus: Application to electrochemical processes and flexible structures, *in* 'Proceedings of 41st IEEE Conference on Decision and Control', Las Vegas.
- Zaborovsky, V. & R. Meylanov (2001), Informational network traffic model based on fractional calculus, *in* 'Proceedings of the International Conference on Info-tech and Info-net', ICII 2001, Beijing.