Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 154 (2019) 311–318

Procedia Computer Science

www.elsevier.com/locate/procedia

8th International Congress of Information and Communication Technology, ICICT 2019

Learning Algorithm for Fractional Dynamical Systems with Autocorrelated Errors-in-Variables

Dmitriy V. Ivanov[a]* , Ilya L. Sandler[a], Natalya V. Chertykovtseva[a], Elena U. Bobkova[b]

[a]*Department of mechatronics, Samara State University of Transport, Samara, Russia*
[b]*Department of economics, Moscow State University of Education, Moscow, Russia*

Abstract

In this paper, the stochastic gradient algorithm for learning fractional-order dynamical systems with noisy input and output is proposed. The proposed algorithm allows estimating the parameters of fractional dynamic systems if the input and output noises are color. The proposed algorithm does not require knowledge of the noise distribution laws. The simulation results demonstrate the high accuracy of the proposed learning algorithm in comparison with the least squares learning algorithm.

Keywords: fractional difference, errors-in-variables, least square, autocorrelated noise, recursive estimation;

1. Introduction

Fractional calculations are used in mechanics, economics, hydrology, and many other sciences. Therefore, the development of methods for the identification of fractional order systems is an important problem, the development of methods for the identification of fractional order systems is an important problem.

In Cois et al., 2000 and Cois et al., 2001 presents methods for the identification of systems based on the method of least squares. An overview of methods for identifying fractional order systems is proposed in Malti et al., 2006.

* Corresponding author. Tel.: +7-927-70-77-091;
E-mail address: dvi85@list.ru

10.1016/j.procs.2019.06.045

Instrumental variables for fractional systems with correlated noise are presented in Victor et al., 2011 Recursive identification algorithms for white noise are proposed in Djouambi, 2012. The use of Kalman filter for white noise in Sierociuk, Dzienlinsk, 2006 and color noise in Sierociuk, Zubinski, 2014, Safarinejadian et al., 2016, Yang et al., 2018 is considered.

An excellent review of methods for identifying integer-order systems with errors in variables is given in Söderström. 2018.

Today there are a small number of articles on the identification of fractional systems with errors in variables Chetoui et al., 2012, Chetoui et al., 2013, Ivanov, 2013. In Ivanov, 2017 proposed a generalization of the results for the case of fractional errors in variables.

2. Fractional Calculus

The fractional derivative is a generalization of the derivative of an integer order. At the same time, unlike derivative of integer order, there are many non-identical definitions fractional derivatives in Samko et al., 1993, Kilbas et al., 2006 . These definitions are similar in form, but significantly differing in properties of fractional differential equations order.

To identify discrete fractional systems, usually use the discrete analogue of the Grünwald-Letnikov derivative

$$\Delta^{\alpha_m} x_i = \sum_{k=0}^{i} \binom{\alpha_m}{k} x_{i-k},$$

where $\binom{\alpha_m}{k}$ is generalized binomial coefficient determined by

$$\binom{\alpha_m}{k} = \frac{(-1)^k \Gamma(\alpha_m+1)}{\Gamma(k+1)\Gamma(\alpha_m+1-k)},$$

$\Gamma(\alpha_m)$ is Gamma function.

3. Problem statement

The linear fractional-order system is described by equations with fractional order differences:

$$z_i = \sum_{m=1}^{r} b^{(m)} \Delta^{\alpha_m} z_{i-1} + \sum_{m=1}^{r_1} a^{(m)} \Delta^{\beta_m} x_i, \quad y_i = z_i + \xi_i, \quad w_i = x_i + \zeta_i \quad (1)$$

A1. System (1) is asymptotically stable.

A2. The values of the orders of the system r, r_1, α_m, β_m are known.

A3. The sequences $\{\xi_i\}$, $\{\zeta_i\}$ with $\mathrm{E}(\xi_i)=0$, $\mathrm{E}(\zeta_i)=0$ and system parameters $\alpha^{(m)}$, $\beta^{(m)}$ satisfy the assumptions:

$$\frac{1}{N}\sum_{i=0}^{N+g} \xi_i \xi_{i+g} \xrightarrow[N\to\infty]{a.s.} h_\xi(g) < \infty, \quad \frac{1}{N}\sum_{i=0}^{N+g} \zeta_i \zeta_{i+g} \xrightarrow[N\to\infty]{a.s.} h_\zeta(g) < \infty,$$

$$\sum_{g=0}^{N}\left|h_{\xi}(g)\right|<\infty,\ \sum_{g=0}^{N}\left|h_{\zeta}(g)\right|<\infty,$$

$h_{\xi}(g)$, $h_{\zeta}(g)$ are autocorrelated functions;

A4. For the sequences $\{\xi_i\}$, $\{\zeta_i\}$, $\{z_i\}$, $\{x_i\}$ the following conditions are satisfied

$$\mathrm{E}\left(\xi_i z_i\right)=0,\ \mathrm{E}\left(\xi_i x_i\right)=0,$$

$$\mathrm{E}\left(\zeta_i z_i\right)=0,\ \mathrm{E}\left(\zeta_i x_i\right)=0,$$

where $\mathrm{E}(\bullet)$ is expectation operator.

A5. For the input sequence $\{x_i\}$ with $E(x_i)=0$ the following conditions is satisfied

$$\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\varphi_{zx}^{(i)}\left(\varphi_{zx}^{(i)}\right)^T=H_{zx}\ \text{a.s. (almost surely)},$$

$$\text{where }\varphi_{zx}^{(i)}=\begin{pmatrix}\varphi_z^{(i)}\\ \hdashline \varphi_x^{(i)}\end{pmatrix}\in\mathbb{R}^{r+r_1},\ \varphi_z^{(i)}=\left(\Delta^{\alpha_1}z_{i-1},\dots,\Delta^{\alpha_r}z_{i-1}\right)^T\in\mathbb{R}^r,\ \varphi_x^{(i)}=\left(\Delta^{\beta_1}x_i,\dots,\Delta^{\beta_{r_1}}x_i\right)^T\in\mathbb{R}^{r_1},$$

matrix H_{zx} is positive definite matrix.

For noisy sequences $\{z_i\}$, $\{x_i\}$ to recursively estimate the parameter vector

$$\hat{\theta}=\left(\hat{b}^{(1)},\dots,\hat{b}^{(r)}\ \vdots\ \hat{a}^{(1)},\dots,\hat{a}^{(r_1)}\right)^T\in\mathbb{R}^{r+r_1}.$$

4. Stochastic Gradient Algorithm for Learning a Fractional-Order Dynamical System with Autocorrelated Error- in-Variables

The error signal is defined

$$\varepsilon_i=y_i-\theta^T\varphi_i=\xi_i-\theta^T\varphi_{\xi\zeta}^{(i)},$$

where

$$\varphi^{(i)}=\begin{pmatrix}\varphi_y^{(i)}\\ \hdashline \varphi_w^{(i)}\end{pmatrix}\in\mathbb{R}^{r+r_1},\qquad \varphi_{\xi\zeta}^{(i)}=\begin{pmatrix}\varphi_\xi^{(i)}\\ \hdashline \varphi_\zeta^{(i)}\end{pmatrix}\in\mathbb{R}^{r+r_1},$$

$$\varphi_y^{(i)}=\left(\Delta^{\alpha_1}y_{i-1},\dots,\Delta^{\alpha_r}y_{i-1}\right)^T\in\mathbb{R}^r,\ \varphi_w^{(i)}=\left(\Delta^{\beta_1}w_i,\dots,\Delta^{\beta_{r_1}}w_i\right)^T\in\mathbb{R}^{r_1},$$

$$\varphi_{\xi}^{(i)}=\left(\Delta^{\alpha_1}\xi_{i-1},\ldots,\Delta^{\alpha_r}\xi_{i-1}\right)^T\in\mathbb{R}^r,\ \varphi_{\zeta}^{(i)}=\left(\Delta^{\beta_1}\zeta_i,\ldots,\Delta^{\beta_{r_1}}\zeta_i\right)^T\in\mathbb{R}^{r_1}.$$

Find the estimates $\hat{\theta}$ of the minimum criterion condition

$$\min_{\theta}\frac{\sum_{i=1}^{N}\left(y_i-\theta^T\varphi_i\right)^2}{\sigma_{\xi}^2+\theta^T H_{\xi\zeta}\theta-2\theta^T\tilde{h}_{\xi\zeta}}. \qquad (2)$$

where

$$H_{\xi\zeta}=\left(\begin{array}{c|c} H_{\xi} & 0_{r\times r_1} \\ \hline 0_{r_1\times r} & H_{\zeta}\end{array}\right),$$

$$H_{\xi}^{(mm')}=\lim_{N\to\infty}\frac{1}{N}\sum_{g=0}^{N-1}\sum_{i=0}^{N-1}\left(\sum_{k=0}^{i}\binom{\alpha^{(m)}}{k}\xi_{i-k}\cdot\sum_{k=0}^{i}\binom{\alpha^{(m')}}{k}\xi_{i-k-g}\right)=$$

$$=\lim_{N\to\infty}\sum_{k=0}^{N-1}\binom{\alpha^{(m)}}{k}\binom{\alpha^{(m')}}{k}h_{\xi}(0)\frac{N-k}{N}+\sum_{g=1}^{N-1}\sum_{k=0}^{N-1}\left(\binom{\alpha^{(m)}}{k}\binom{\alpha^{(m')}}{k-g}+\binom{\alpha^{(m)}}{k-g}\binom{\alpha^{(m')}}{k}\right)h_{\xi}(g)\frac{N-k-g}{N},$$

$$m=1,r,$$

$$H_{\zeta}^{(mm')}=\lim_{N\to\infty}\frac{1}{N}\sum_{g=0}^{N-1}\sum_{i=0}^{N-1}\left(\sum_{k=0}^{i}\binom{\beta^{(m)}}{k}\zeta_{i-k}\cdot\sum_{j=0}^{i}\binom{\beta^{(m')}}{k}\zeta_{i-k-g}\right)=$$

$$=\lim_{N\to\infty}\sum_{k=0}^{N-1}\binom{\beta^{(m)}}{k}\binom{\beta^{(m')}}{k}h_{\zeta}(0)\frac{N-k}{N}+\sum_{g=1}^{N-1}\sum_{k=0}^{N-1}\left(\binom{\beta^{(m)}}{k}\binom{\beta^{(m')}}{k-g}+\binom{\beta^{(m)}}{k-g}\binom{\beta^{(m')}}{k}\right)h_{\zeta}(g)\frac{N-k-g}{N},$$

$$m=1,r_1,$$

$$\tilde{h}_{\xi\zeta}=\left(\tilde{h}_{\xi}\ \vdots\ 0_{1\times r_1}\right)^T,$$

$$\tilde{h}_{\xi\zeta}^{(m)}=H_{\xi}^{(mm')}=\lim_{N\to\infty}\frac{1}{N}\sum_{g=0}^{N-1}\sum_{i=0}^{N-1}\left(\xi_i\cdot\sum_{j=0}^{i}\binom{\alpha^{(m')}}{j}\xi_{i-j-g}\right)=\sum_{g=1}^{N-1}\sum_{j=0}^{N-1}\binom{\alpha^{(m)}}{j}h_{\xi}(g+j)\frac{N-j-g}{N},\ m=1,r,$$

The minimum of the criterion (2) can be found with the help of stochastic approximation:

$$\hat{\theta}(i+1)=\hat{\theta}(i)-\delta_i\nabla_{\theta}J\left(\hat{\theta}(i)\right), \qquad (3)$$

where

$$J\left(\hat{\theta}(i)\right)=\frac{\left(y_i-\hat{\theta}(i)^T\varphi_i\right)^2}{\sigma_\xi^2+\hat{\theta}(i)^T H_{\xi\zeta}\hat{\theta}(i)^T-2\hat{\theta}(i)^T\tilde{h}_{\xi\zeta}},$$

For the sequence $\{\delta_i\}$, the assumptions are defined:

$$\textbf{A6}.\ \delta_i \geq \delta_{i+1},\ \sum_{i=0}^{\infty}\delta_i = 0 \text{ and } \sum_{i=0}^{\infty}\delta_i^{\gamma} < \infty,\quad \gamma > 1.$$

$$\textbf{A7}.\ \sum_{i=0}^{\infty}\delta_i\,\xi_i < \infty,\ \sum_{i=0}^{\infty}\delta_i\,\xi_i < \infty \text{ a.s.}$$

Theorem. The sequence of estimates of the coefficients $\left\{\hat{\theta}(i)\right\}$ *(3) converges to the true value* $\lim_{i\to\infty}\hat{\theta}(i)=\theta$ *or diverges if the assumptions* ***A1-A7*** *hold.*

Proof. The proof is based on the continuous model method Ivanov et al., 2017.

5. Stochastic Approximation with Averaging

The rate of convergence of the algorithm (3) depends on the choice of the value of the sequence $\{\delta_i\}$. The choice of the optimal sequence $\{\delta_i\}$ depends on the variance of the generalized error ε_i and the minimum eigenvalue of the matrix H_{zx}. In practice, these quantities are usually unknown or only estimates of these quantities are known. The new type of stochastic approximation is proposed a stochastic approximation with averaging in Polyak, 1990:

$$\hat{\theta}(i+1)=\hat{\theta}(i)-\frac{\delta}{i^{\gamma}}\nabla_{\theta}J\left(\hat{\theta}(i)\right), \tag{4}$$

$$\hat{\bar{\theta}}(i+1)=\frac{i-1}{i}\hat{\bar{\theta}}(i)+\frac{1}{i}\hat{\theta}(i+1). \tag{5}$$

Algorithm (4), (5) has an asymptotically optimal rate of convergence for any values of the sequence $\{\delta_i\}$. In practical problems of estimation can be accepted $l=0.5$ Polyak, Judiskiy, 1992. When using the averaging algorithm, the step size is chosen more than in the algorithm (3). This leads to a jump around the true vector θ of the sequence of estimates $\left\{\hat{\theta}(i)\right\}$. Averaging compensates for these jumps, and under general conditions, the rate of convergence $\left\{\hat{\bar{\theta}}(i)\right\}$ coincides with the maximum possible. The use of stochastic approximation with averaging gives good results for solving applied problems in Nazin, Piterskaya, 2002.

6. Simulation Results

The proposed algorithm was compared with the recursive least-squares (RLS) method in a test example.The fractional noiseless system is defined by the equation

$$z_i = 0.5\Delta^{0.9} z_{i-1} - 0.2\Delta^{1.7} z_{i-1} + \Delta^{0.6} x_i - 0.2\Delta^{1.2} x_i \quad (5)$$

The input and output noises are defined as

$$\xi_i - 0.2\xi_{i-1} - 0.3\xi_{i-2} = \upsilon_i + 3\upsilon_{i-1} + 7\upsilon_{i-2} + 10\upsilon_{i-3}, \quad (6)$$

$$\zeta_i - 0.5\zeta_{i-1} - 0.2\zeta_{i-2} = \tau_i + 4\tau_{i-1} - 4\tau_{i-2} + 2\tau_{i-3}, \quad (7)$$

where τ_i, υ_i are white noises.

The vector of the true coefficients of the system is $\theta = (0.5 \quad -0.2 \quad 1 \quad -0.2)^T$.

The input signal x_i is defined as

$$.x_i = -0.5x_{i-1} + e_i - 0.2e_{i-1} - 0.75 \cdot e_{i-2} + e_{i-4},$$

where e_i - is white noise with $\mathrm{E}(e_i) = 0$.

Initial values of the parameter vector is $\hat{\theta}(0) = (0 \quad 0 \quad 0 \quad 0)^T$.

Number of observations $N = 2000$.

Figures 1-3 gives plots of normalized root-mean-square error (NRMSE) estimation of parameter is defined by

$$\delta\theta_i = \sqrt{\frac{\left\|\hat{\theta}(i) - \theta\right\|^2}{\|\theta\|^2}} \cdot 100\%.$$

Signal-to-Noise Ratio is defined by

$$SNR_{out} = \frac{\sigma_\xi}{\sigma_z}, \; SNR_{in} = \frac{\sigma_\zeta}{\sigma_x}.$$

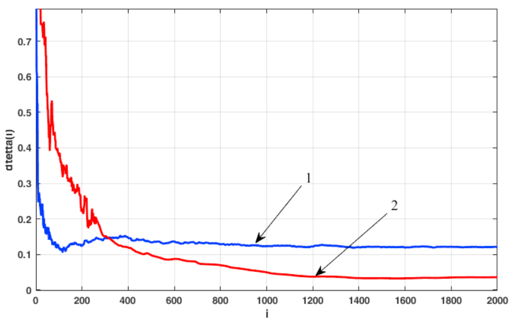

Fig. 1. NRMSE for SNRin= SNRout =0.2: 1-RLS; 2-proposed algorithm.

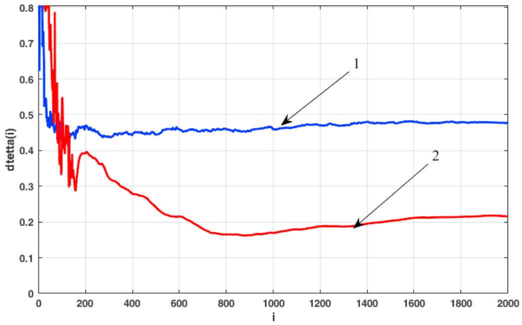

Fig. 2. NRMSE for SNRin= SNRout =0.7: 1-RLS; 2-proposed algorithm.

7. Conclusion

In this paper, the stochastic gradient algorithm for learning a fractional-order dynamical system in the presence of auto-correlated errors in variables is proposed. The simulation results showed that the proposed algorithm works better compared to the recursive least squares method. Further research should be directed to the development of algorithms that do not require knowledge of the autocorrelation functions of noises or their estimation in the learning process.

References

1. Chetoui, M., Thomassin, M., Malti, R., Aoun, M., Najar, S., Abdelkrim, M.N., Oustaloup, A., 2012, EIV methods for system identification with fractional models. IFAC Proceedings 45(16), pp. 1641–1646.
2. Chetoui, M., Thomassin, M., Malti, R., Aoun, M., Najar, S., Abdelkrim, M.N., Oustaloup, A., 2013, New consistent methods for order and coefficient estimation of continuous-time errors-in-variables fractional models. Computers & Mathematics with Applications, 66(5), pp. 860-872.
3. Cois, O, Oustaloup, A, Battaglia, E, Battaglia, J.L., 2000, Non-integer model from modal decomposition for time domain system identification. Proc Symp System Identification, SYSID 2000, p. 989–994.
4. Cois, O, Oustaloup, A, Battaglia, E, Poinot, T, Battaglia, J.L., 2001, Fractional state variable filter for system identification by fractional model. Proc. European Contr. Conf., ECC2001, p. 2481–2486.
5. Djouambi, A., Voda, A., Charef, A., 2012, Recursive prediction error identification of fractional order models, Communications in Nonlinear Science and Numerical Simulation 17(6), p. 2517-2524.
6. Ivanov, D,V., 2013, Identification discrete fractional order linear dynamic systems with errors-in-variables. In Proceedings of IEEE East-West Design & Test Symposium (EWDTS'13). Rostov-on-Don. Russia, pp. 374–377.
7. Ivanov, D.V., Ivanov, A.V., 2017, Identification Fractional Linear Dynamic Systems with fractional errors-in-variables. Journal of Physics: Conference Series, 803 (1).
8. Ivanov, D.V., Bobkova, E.U. and Zharkova, A.A., 2017, Recursive identification of bilinear dynamical systems with noise in output signal. In Proceeding 2017 IEEE East-West Design & Test Symposium (EWDTS), Novi Sad, pp. 1-5.
9. Kilbas, A.A., Srivastava, H.M., Trujillo, J.J., 2006, Theory and applications of fractional differential equations. Elsevier, Amsterdam.
10. Malti, R., Aoun, M., Sabatier, J., Oustaloup, A., 2006, Tutorial on system identification using fractional differentiation models. Proc. 14th IFAC Symposium on System Identification (SYSID’2006), p. 606-611.
11. Nazin, A.V., Piterskaya, E.V., 2002, Method of Stochastic Approximation with Averaging for Evaluating Sizes of Microparticles in Digital Image Autom. Remote Control, 63(11) pp. 1853–1859.
12. Polyak, B.T., 1990, A new method of stochastic approximation type. Autom. Remote Control, 51(7), pp. 937–946.
13. Polyak, B.T., Juditsky, A.B., 1992, Acceleration of stochastic approximation by averaging // SIAM J. Contr. Optimiz. 30(4), pp. 838–855.
14. Safarinejadian, B., Asad, M., Sadeghi, M.S., 2016, Simultaneous state estimation and parameter identification in linear fractional order systems using coloured measurement noise, International Journal of Control, 89(11), pp. 2277-2296.
15. Samko, S.G., Kilbas ,A.A., Marichev, O.I., 1993, Fractional Integrals and derivatives: Theory and Applications. Gordon and Breach Science Publishers, Yverdon.
16. Sierociuk, D., Dzieliński, A., 2006, Fractional Kalman filter algorithm for states, parameters and order of fractional system estimation. Int. J. Appl. Math. Comput. Sci. 16(1), p. 101–112.

17. Sierociuk, D., Ziubinski, P., 2014, Fractional Order Estimation Schemes for Fractional and Integer Order Systems with Constant and Variable Fractional Order Colored Noise. Circuits Syst Signal Process 33, pp. 3861–3882.
18. Victor, S., Malti, R., Melchior, P., Oustaloup, A., 2011, Instrumental Variable Identification of Hybrid Fractional Box-Jenkins Models. IFAC Proceedings Volumes 44(1), p. 4314-4319.
19. Yang, C., Gao, Z., Liu, F., 2018, Kalman filters for linear continuous-time fractional-order systems involving coloured noises using fractional-order average derivative. Applied Mathematics and Computation, 12(4), pp. 456 – 465.