RECURSIVE IDENTIFICATION FRACTIONAL ORDER HAMMERSTEIN SYSTEMS WITH NOISE IN OUTPUT SIGNAL

C D.V. Ivanov, O.A. Katsyuba
Samara State University of Transport, Samara, Russia

Recursive algorithm is proposed for identification single-input-single-output (SISO) fractional order Hammerstein systems with noise in output signal. The estimates are proved to be convergent to the true values with probability one. The results of a simulated example indicate that the proposed algorithm provides good estimates/

Keywords: consistency, fractional order system, Hammerstein system, nonlinear least square, output-error, parametric identification, error in equation.

Introduction

Throughout the last few decades, the identification methods of the linear systems have been explored intensively. But in the field of nonlinear systems, the modelling issue is more complicated and harder.

For many systems, when there is a wide operating area rather than a unique operating point, a linear model cannot be used. In this case non-linear models such as Volterra series, neural networks, Hammerstein and Wiener-type models can be used. Hammerstein models are composed of a memoryless static nonlinearity followed by a linear dynamical system.

They have been used e.g., for modeling biological processes [1], [2], chemical processes [3], and in signal processing applications [4]. Hammerstein models have also been shown to be useful for control problems [5].

As far as the applications of fractional calculus are concerned there is a large number of research on viscoelasticity/damping, [6, 7] and chaos/fractals [8], dielectric materials [9], electrochemical processes and flexible robot [10], traffic in information networks [11].

Due to their long memory behavior, the identification of fractional order models is more difficult compared with those of the integer order models.

In this paper, a Hammerstein model, with a static non-linear part connected to a fractional linear part, is considered. Non-recursive identification methods considered in [12,13,14]. The aim of this paper is to present new method recursive identification the output-error Hammerstein system with error in the equation.

The rest of the paper is organized as follows. Section 1 presents problem statement and introduces some notational conventions. Recursive algorithms are given and their strong consistency is proved in Section 2. To justify theoretical assertions some numerical simulation results are demonstrated in Section 3. Some short concluding remarks are finally given in Section 4.

Problem Statement

The structure of Hammerstein systems is shown in figure (1). It consists of a static non-linear part connected to a dynamic linear part:

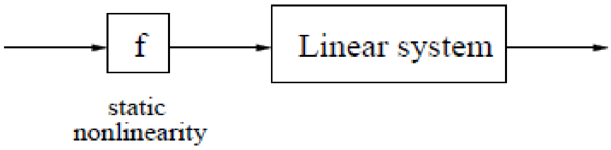

Fig. 1. Hammerstein system consists of a memoryless static non-linearity followed by a linear dynamical system

Let's consider the fractional Hammerstein system described by the stochastic equations

$$z_i = \sum_{m=1}^{r} b_0^{(m)} \Delta^{\alpha_m} z_{i-1} + \sum_{m=1}^{r_1} a_0^{(m)} \Delta^{\beta_m} f_m(x_i) + \zeta_i,$$

$$y_i = z_i + \xi_i, \qquad (1)$$

where $\Delta^{\alpha_m} z_i = \sum_{j=0}^{i} (-1)^j \binom{\alpha_m}{j} z_{i-j}$, $0 < \alpha_1 \ldots < \alpha_r$, $\Delta^{\beta_m} f_m(x_i) = \sum_{j=0}^{i} (-1)^j \binom{\beta_m}{j} f_m(x_{i-j})$, $0 < \beta_1 \ldots < \beta_{r_1}$,

the Euler's function Γ is defined as $\Gamma(\alpha) = \int_0^{\infty} e^{-t} t^{\alpha-1} dt$,

Newton's binomial is generalized to non-integer orders using the Euler's function

$$\binom{\alpha_m}{j} = \frac{\Gamma(\alpha_m + 1)}{\Gamma(j+1)\Gamma(\alpha_m - j + 1)}, \quad \binom{\beta_m}{j} = \frac{\Gamma(\beta_m + 1)}{\Gamma(j+1)\Gamma(\beta_m - j + 1)},$$

$f_m(\bullet)$ nonlinear function.

When $f_m(\bullet)$ is unknown, it can be approximated with a polynomial expansion

$$f_m(x_i) = \sum_{k}^{M} c_k^{(m)} x_i^k .$$

This class of models includes the class of Hammerstein integer order models.

The following assumptions are introduced:

1. The dynamic system (1) is asymptotically stable.

2. Noises $\{\xi_i\}$ and $\{\zeta_i\}$ are statistically independent sequences with $E\{\xi_i\} = 0$, $E\{\zeta_i\} = 0$, $E\{\xi_i^2\} = \sigma_\xi^2 < \infty$, $E\{\zeta_i^2\} = \sigma_\zeta^2 < \infty$ a.s., where E is the expectation operator.

3. The sequences $\{\xi_i\}$ and $\{\zeta_i\}$ are mutually do not depend and do not depend with z_i , x_i , respectively.

4. Sequence $\{x_i\}$ are random signals with $E\left\{f_m^2(x_i)\right\}<\infty$, $\left|f_m(x_i)\right|<\infty$ and

$$\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\left(\begin{array}{c}\varphi_z^{(i)}\\ \hdashline \varphi_{f(x)}^{(i)}\end{array}\right)\left(\left(\varphi_z^{(i)}\right)^T \;\vdots\; \left(\varphi_{f(x)}^{(i)}\right)^T\right)=H \text{ a.s.,}$$

$$\varphi_z^{(i)}=\left(\sum_{j=0}^{i}(-1)^j\binom{\alpha_1}{j}z_{i-j-1},\ldots,\sum_{j=0}^{i}(-1)^j\binom{\alpha_r}{j}z_{i-j-1}\right)^T,$$

$$\varphi_{f(x)}^{(i)}=\left(\sum_{j=0}^{i}(-1)^j\binom{\beta_1}{j}f_1\left(x_{i-j}\right),\ldots,\sum_{j=0}^{i}(-1)^j\binom{\beta_{r_1}}{j}f_{r_1}\left(x_{i-j}\right)\right)^T,$$

matrix H is restricted positive definite.

5. Ratio $\gamma=\sigma_\zeta^2/\sigma_\xi^2$ is known apriori.

It is required to recursive estimate unknown vector of coefficients b,a fractional order Hammerstein system, described by the equation (1) in observable sequences $\{y_i\}$, $\{x_i\}$.

Recursive identification algorithm

Equation (1) is represented in the vector form as a linear regression:

$$y_i=\varphi_i^T\theta_0+\varepsilon_i, \tag{2}$$

where $\varphi_i=\left(\left(\varphi_y^{(i)}\right)^T \;\vdots\; \left(\varphi_{f(x)}^{(i)}\right)^T\right)^T$,

$$\varphi_y^{(i)}=\left(\sum_{j=0}^{i}(-1)^j\binom{\alpha_1}{j}y_{i-j-1},\ldots,\sum_{j=0}^{i}(-1)^j\binom{\alpha_r}{j}y_{i-j-1}\right)^T,$$

$$\theta_0=\left(b_0^T \;\vdots\; a_0^T\right)^T=\left(b_0^{(1)},\ldots,b_0^{(r)} \;\vdots\; a_0^{(1)},\ldots,a_0^{(r_1)}\right)^T,$$

$$\varepsilon_i=\xi_i+\zeta_i-b_0^T\varphi_\xi^{(i)},\quad \varphi_\xi^{(i)}=\left(\sum_{j=0}^{i}(-1)^j\binom{\alpha_1}{j}\xi_{i-j-1},\ldots,\sum_{j=0}^{i}(-1)^j\binom{\alpha_r}{j}\xi_{i-j-1}\right)^T.$$

From requirement 2 it follows that generalized error ε_i has zero mean. We obtain that variance of generalized error equal to:

$$\begin{aligned}\sigma_\varepsilon^2&=\sigma_\zeta^2+\sigma_\xi^2+\sigma_\xi^2 b_0^T H_\xi b_0=\\ &=\sigma_\xi^2\left(1+\gamma+b_0^T H_\xi b_0\right)=\sigma_\xi^2\omega(b_0),\end{aligned}$$

$$H_\xi=E\left[\sum_{i=1}^{N}\varphi_\xi^{(i)}\left(\varphi_\xi^{(i)}\right)^T\right]\Big/\sigma_\xi^2=\begin{pmatrix}h_\xi^{(11)} & \cdots & h_\xi^{(r1)}\\ \vdots & \ddots & \vdots\\ h_\xi^{(1r)} & \cdots & h_\xi^{(rr)}\end{pmatrix},$$

$$h_{\xi}^{(mk)} = E\left(\sum_{j=0}^{i}(-1)^j\binom{\alpha_m}{j}\xi_{i-j-1}\cdot\sum_{j=0}^{i}(-1)^j\binom{\alpha_k}{j}\xi_{i-j-1}\right)=$$

$$= \lim_{N\to\infty}\frac{1}{N}\left(\sum_{j=0}^{N-1}\binom{\alpha_m}{j}\binom{\alpha_k}{j}\cdot\frac{N-j}{N}\right)\quad m=\overline{1,r},\ k=\overline{1,r}.$$

Let's define the estimates $\hat{\theta}$ of the unknown true value from requirement of minimum considered square variance σ_{ε}^2 weighted $\omega(b,a)$ [14]:

$$\min_{\theta\in B}\sum_{i=1}^{N}\frac{\left(y_i-\varphi_i^T\theta\right)^2}{1+\gamma+b^T H_{\xi}(N)b}.$$

Theorem1. The estimates of unknown vector θ can be obtained by means of the stochastically gradient algorithm of the functional minimization:

$$\hat{\theta}(i+1)=\hat{\theta}(i)-\delta_i\nabla_{\theta}\left[\frac{\left(y_{i+1}-\varphi_{i+1}^T\hat{\theta}(i)\right)^2}{1+\gamma+\hat{b}^T(i)H_{\xi}(i+1)\hat{b}(i)}\right], \qquad (3)$$

where α_i sequence for which the requirements:

6. $\sum_{i=0}^{\infty}\delta_i=\infty,$ $\delta_i\ge\delta_{i+1}$ и $\sum_{i=0}^{\infty}\delta_i^l<\infty$ at $l>1.$

7. $\sum_{i=1}^{\infty}\delta_i\xi_i<\infty$, $\sum_{i=1}^{\infty}\delta_i\zeta_i<\infty$ a.s.

are met, then the estimates defined by algorithm (3) or $\hat{\theta}(i)\xrightarrow[i\to\infty]{}\theta_0$ a.s., or $\hat{\theta}(i)\xrightarrow[i\to\infty]{}\infty.$

Proof. Function (2) can be represented as:

$$J(\theta)=\sigma_{\xi}^2+\frac{(\theta-\theta_0)^T H(\theta-\theta_0)}{1+\gamma+b^T H_{\xi}b},$$

as follows from 1, 5 [15].

In this case the asymptotic continuous deterministic model can be represented as [16]:

$$\dot{\theta}=-\nabla_{\theta}J(\theta).$$

Suppose Lyapunov function is equal to $V(\theta)=J(\theta)$, as $\dot{V}(\theta)=\nabla_{\theta}^T V(\theta)J(\theta)=-\left\|\nabla_{\theta}^T J(\theta)\right\|^2$.

To go from (3) to a continuous model to show that for $\{\xi_i\}$, $\{\zeta_i\}$ and $\{\alpha_i\}$ condition hold [16]:

$$\lim_{T\to 0}\lim_{n\to\infty}\sup\frac{1}{T}\left\|\sum_{i=n}^{k(n,T)}\delta_i\varepsilon_i\left(\hat{\theta}(i),\xi_i,\zeta_i\right)\right\|=0, \qquad (4)$$

where for $T>0,\ \ k(n,T)=\max\left\{k:\sum_{i=n}^{k}\delta_i\le T\right\},$

$\|\bullet\|-$ euclidean norm.

For equality (4) a priori requires the boundedness of $\{\hat{\theta}(i)\}$, which is an implicit requirement for the growth rate of $\nabla_\theta J(\theta)$ as $\|\theta\|\to\infty$, and in our case

$$\lim_{\|\theta\|\to\infty}\nabla_\theta J(\theta)=0.$$

From the boundedness sum provided in the 7 and the sequence $\{\hat{\theta}(i)\}$implies that the sum $\sum_{i=1}^{\infty}\delta_i\varepsilon_i\left(\hat{\theta}(i),\xi(i)\right)<\infty,$ of which implies the validity of (4).

Of the theorems given in [16], it follows that at 1-7, the sequence $\{\hat{\theta}(i)\}$ is bounded and tends to $i\to\infty,$ $\{\hat{\theta}(i)\}$ the points of the set $B=\{\theta\in R^{r+r_1}:\dot{V}(\theta)=0\}$ consists of stationary points of functional $J(\theta)$ [17].

However, theorem 3.15 [17] implies that probable accumulation points of algorithm (3) are the points of set

$$B_*=\left\{\theta\in R^{r+r_1}:\dot{V}(\theta)=0,\ \ -\nabla^2 J(\theta)\le 0\right\}$$

Let's show that $B_*=\{\theta\in R^{r+r_1}:\theta=\theta_0\}$, i.e. set B_* consists of a single point θ_0.

For this purpose let's take functional

$$J'\ (u)=\frac{u^T\overline{H}_\varphi u}{u^T H'_\xi u}\ ,$$

where $u\ =(u_1,...,u_{r+r_1+1})^T,$

$$\overline{H}_\varphi=\lim_{i\to\infty}E\left[\begin{pmatrix}-y_i\\ \hdashline \varphi_i\end{pmatrix}\begin{pmatrix}-y_i & \vdots & \varphi_i^T\end{pmatrix}\right],$$

$$H'_\xi=\begin{pmatrix}1+\gamma & 0_{1\times r+1} & 0_{1\times r_1}\\ 0_{r\times 1} & H_\xi & 0_{r\times r_1}\\ 0_{r_1\times 1} & 0_{r_1\times r} & 0_{r_1\times r_1}\end{pmatrix},$$

It is obvious that

$$\min_{\theta} J(\theta) = \min_{u} J'(u) = J(\theta_0) = \Lambda_{\min}, \qquad (4)$$

where $\Lambda_{\min}$ is a minimal eigenvalue of the regular pencil of matrices [18], i.e. $\Lambda_{\min}$ is the minimal root of equation $\det(\overline{H}_{\varphi} - \Lambda H'_{\xi}) = 0$. Assume that $\Lambda_{\min} = \Lambda^{(1)} \leq \ldots \leq \Lambda^{(r+r_1+1)} = \Lambda_{\max}$ and $u_1,\ldots,u_{r+r_1+1}$ are their eigenvectors. Then, Λ_k, where $k = \overline{1, r+r_1+1}$ are stationary values of function $J'(u)$, which can be obtained at u, equal to $u_1,\ldots,u_{r+r_1+1}$ respectively. Hence, stationary values of function $J(\theta)$; $\nabla_{\theta} J(\theta) = 0$ are obtained at points

$$\theta_1 = \left(u_1^{(2)} \big/ u_1^{(1)}, \ldots, u_1^{(r+r_1+1)} \big/ u_1^{(1)}\right)^T, \ldots,$$

$$\theta_{r+r_1+1} = \left(u_{r+r_1+1}^{(2)} \big/ u_{r+r_1+1}^{(1)}, \ldots, u_{r+r_1+1}^{(r+r_1+1)} \big/ u_{r+r_1+1}^{(1)}\right)^T.$$

It follows from (4) that $\theta_1 = \theta$.

It remains to show that

$$\left\{\nabla^2 J(\theta) \geq 0, \forall \theta \in \left\{\theta : \theta_1, \ldots, \theta_{r+r_1+1}\right\}\right\} \qquad (5)$$

only at one stationary point $\theta = \theta_1 = \theta_0$.

The problem of determining minimum of function $J(\theta)$ is equivalent to the constrained extremum problem

$$\begin{aligned} &\min u^T \overline{H}_{\varphi} u, \\ &u^T H'_{\xi} u = 1. \end{aligned} \qquad (6)$$

Problem (6) can be solved by means of Lagrange multiplier method. Then, the required conditions will be written as follows

$$\begin{cases} (\overline{H}_{\varphi} - \lambda H'_{\xi}) u = 0, \\ u^T H'_{\xi} u = 1, \end{cases} \qquad (7)$$

where λ is Lagrange undetermined multiplier. Solution set of system (7), are $\lambda \in \left\{\Lambda_1, \ldots, \Lambda_{r+r_1+1}\right\}$ and their eigenvectors $u_1,\ldots,u_{r+r_1+1}$.

Let's test matrix $\overline{H}_{\varphi} - \lambda H'_{\xi}$ for its positive definiteness. From (4) it follows that

$$\Lambda^{(1)} \left|\overline{H}_{\varphi}\right| < \Lambda^{(1)} \left|H_{\varphi}\right|,$$

where $\Lambda^{(1)}\left|\overline{H}_{\varphi}\right|$ and $\Lambda^{(1)}\left|H_{\varphi}\right|$ are minimal eigenvalue of matrices $\overline{H}_{\varphi}$ and H_{φ} respectively.

Again, Sturm's theorem [19] implies that

$$\Lambda^{(1)}\left|H_{\varphi}\right| \leq \Lambda^{(2)}\left|\overline{H}_{\varphi}\right| \text{ or } \Lambda^{(1)}\left|H_{\varphi}\right| < \Lambda^{(2)}\left|\overline{H}_{\varphi}\right|. \qquad (8)$$

From (8) it follows that matrix $\overline{H}_{\varphi} - \lambda H'_{\xi}$ is non-negatively defined only at $\lambda = \Lambda_{\min}$ and (5) is performed $\theta_1 = \theta$, i.e. for all $\lambda > \Lambda_{\min}$ matrix $\overline{H}_{\varphi} - \lambda H'_{\xi}$ has negative eigenvalues, hence (3). The main feature of the algorithm which enables us to prove global convergence of the simple stochastically gradient algorithm is that function $J(\theta)$ is limited both from above and below. Among all stationary points of function $J(\theta)$ only point θ_0 is the minimum point, whereas all other points are saddle points and one point is a maximum point.

Simulation results

The offered algorithm has been realized in Matlab and compared to recursive least square (RLS) method.

The dynamic system is defined by the equations

$$z_i = 0.5\Delta^{0.9} z_{i-1} - 0.2\Delta^{1.7} z_{i-1} + \Delta^{0.7} f(x_i) - 0.2\Delta^{1.4} f(x_i) + \zeta_i, \quad f(x_i) = x_i + \tanh(x_i), \quad y_i = z_i + \xi_i.$$

The noise-free input is defined as

$$x_i + 0.5 \cdot x_{i-1} = \zeta_i - 0.2 \cdot \zeta_{i-1} - 0.75 \cdot \zeta_{i-2} + \zeta_{i-4},$$

where ζ_i - is white noise.

Initial parameters are equal to 0.

Fig. 2-4 gives plots of the root mean squared error (RMSE) which is defined by

$$\delta\theta_i = \sqrt{\left\|\hat{\theta}_i - \theta_0\right\|^2 \Big/ \left\|\theta_0\right\|^2} \cdot 100\,\%.$$

Conclusion

In this paper, the recursive algorythm of consistent estimation of Hammerstein system has been studied. Simulation results indicate that the RLS method gives biased results. Hammerstein systems have received much attention because of its important applications in signal processing, communications and control systems.

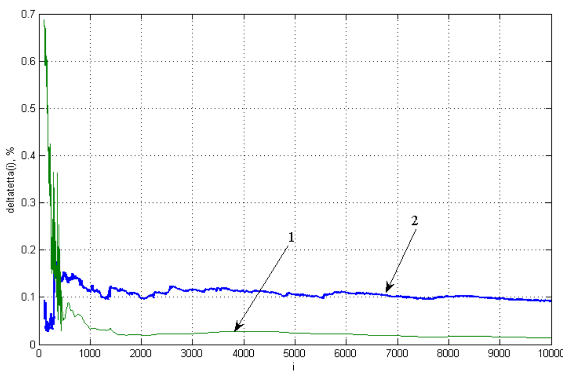

Fig. 2. RMSE of parameter estimates for $\sigma_{\xi}/\sigma_z = 0.2$, $\sigma_{\zeta}/\sigma_z = 0.2$: 1-proposed algorithm; 2-RLS;

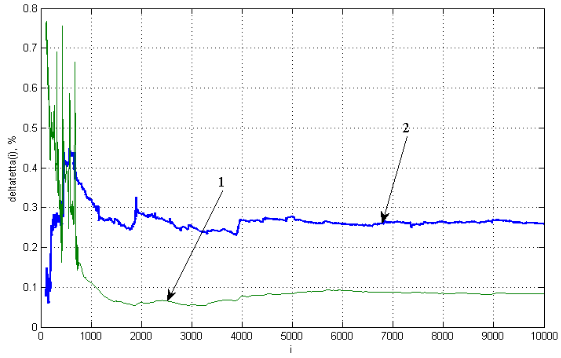

Fig. 3. RMSE of parameter estimates for $\sigma_\xi / \sigma_z = 0.5$, $\sigma_\zeta / \sigma_z = 0.2$: 1-proposed algorithm; 2-RLS;

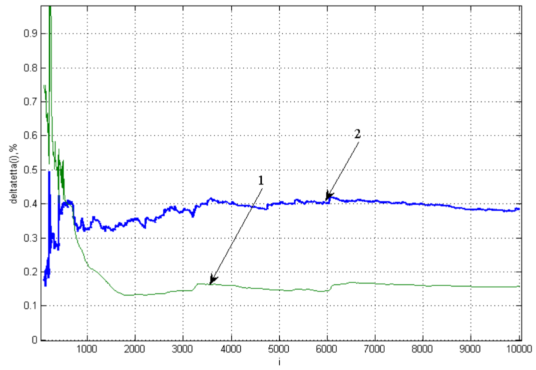

Fig. 1. RMSE of parameter estimates for $\sigma_\xi / \sigma_z = 0.7$, $\sigma_\zeta / \sigma_z = 0.2$: 1-proposed algorithm; 2-RLS.

References

1. Korenberg, M.J. Recent advances in the identification of nonlinear systems: Minimum-variance approximation by Hammerstein models. Proc. IEEE EMBS. – 1991. – Vol.13. – P. 2258-2259.
2. Westwick, D., Kearney, R., Identification of a Hammerstein model of the stretch reflex EMG using separable least squares. Proc. World Congr. Medical Physics and Biomedical Engineering, Chicago, IL. – 2000. – Vol. 3. – P. 1901-1904.
3. Eskinat, E. , Johnson, S.H., Luyben, W.L. Use of Hammerstein models in identification of nonlinear systems, AIChE Journal. – 1991. – Vol. 37(2). – P. 255–268.
4. Stapleton, J.C., Bass, S.C. Adaptive noise cancellation for a class of nonlinear dynamic reference channels. IEEE Transactions on Circuits and Systems. – 1985. – Vol. 32(2). – P. 143-150.
5. Lang, Z.H. Controller design oriented model identification method for Hammerstein systems. Automatica. – 1993. – Vol. 29. – P. 767-771.
6. Stiassnie, M. On the application of fractional calculus for the formulation of viscoelastic models. Applied Mathematical Modelling. – 1979. – Vol. 3. – P. 300-302.
7. Bagley, R.L. Fractional calculus – a different approach to the analysis of viscoelastically damped structures. AIAA Journal. – 1983. Vol. 21. – P. 741-748.
8. Le Mehaute, A. Fractal Geometries, Theory and Applications, Penton Press, New York, 1991.
9. Reyes-Melo, M.E., Martinez-Vega, J.J., Guerrero-Salazar, C.A., Ortiz-Mendez, U. Application of fractional calculus to modelling of relaxation phenomena of organic dielectric materials. Proc Int Conf Solid Dielectrics. – 2004. – Vol. 2. – P. 530-533.
10. Vinagre BM, FeliuV. Modeling and control of dynamic system using fractional calculus: Application to electrochemical processes and flexible structures. Proc 41-st IEEE Conf Decision Control, Las Vegas, NV; 2002: 214–239.
11. Zaborovsky V., Meylanov R. Informational network traffic model based on fractional calculus. Proc Int Conf Info-tech and Info-net, ICII 2001. – 2001. – Vol. 1. – P. 58–63.
12. Aoun, M., Malti, R., Cois, O., Oustaloup, A. System identification using fractional Hammerstein models. Proc. 15th IFAC World Congress, Barcelona, Spain. – 2002. P. 265-270.

13. Li Y, Zhai L, Chen Y, Ahn H. Fractional-order iterative learning control and identification for fractional-order Hammerstein system. Proc. of the 11th World Congresson Intelligent Control and Automation. – 2014. – P. 840–845.
14. Ivanov DV. Identification discrete fractional order Hammerstein systems. Proc. International Siberian Conference on Control and Communications (SIBCON'2015). – 2015. – P. 1-4.
15. Wilks, S. Mathematical statistics. Princeton University Press, NJ, 1943.
16. Chen, H.F. Stochastic Approximation and Its Applications. Kluwer, Dordrecht. 2005.
17. Derevitskiy, D.P., Fradkov, A.L. Applied theory of discrete adaptive control systems. Moscow: Nauka, 1991 (in Russian).
18. Gantmacher, F.R. The theory of matrices. AMS Chelsea Publishing: Reprinted by American Mathematical Society. 2000.
19. Bellman, R. Introduction to matrix analisys. McGraw-Hill, NJ, 1960.