Identification of MIMO Hammerstein Models in the Presence of Piecewise Polynomial Disturbances using Kaczmarz Algorithm

Vojislav Filipovic[1*], Vladimir Djordjevic[1]
[1]Department for Automatic Control, Robotics and Fluid Technique, Faculty of Mechanical and Civil Engineering, University of Kragujevac, Kraljevo (Serbia)

The Kaczmarz algorithm, originally, is an iterative projection algorithm for solving a linear system of equations. After the publications of the result, it was observed that it is possible to design simple recursive algorithm for parameters estimation. Because of its simplicity algorithm found different applications including tomography, distributed computation, computer vision and medicine (scanners) to name a few. In this paper, we try to expand the area of applications. Namely, we consider the problem of Hammerstein model identification. It is supposed that system has MIMO (multi-input multi-output) structure. The nonlinear part of Hammerstein model has general structure and disturbance has piecewise polynomial form. The algorithm is novel and the key ingredient of the algorithm is the Shannon – Kulebakin operator.

Keywords: MIMO Hammerstein models, Piecewise polynomial disturbance, Shannon-Kulebakin operator, Kaczmarz algorithm

1. INTRODUCTION

Multivariable systems represent an important class of systems in practice. Special attention is devoted to their identification [1]. In this paper, we consider identification of nonlinear multiple-input multiple-output (MIMO) systems. A wide class of nonlinear systems is modeled as a linear model cascading with a static nonlinear function. In this paper, we consider the situation when the static function is prior to the linear subsystem. Such kind of models is known as a Hammerstein models and are relevant for engineering systems, social systems, biological systems, machine learning, pattern recognition and others. The Hammerstein models belong to the block-oriented models [2]. Here the static nonlinear block is given in the general polynomial form while the linear model is an output error (OE) model.

Special attention is given to modeling disturbances. The common assumption is that disturbance has Gaussian distribution [3]. In references [4]-[5] it is assumed that the stochastic disturbance is non-Gaussian. This assumption is confirmed in practice [6] where is shown that in a population of observations there are rare large observations (outliers). Owing to that fact the great effort has been invested in creating algorithms which have a low sensitivity to changes in the stochastic disturbance distribution. That part of identification theory is based on robust statistics [7].

In this paper, we consider a different mode to modeling the disturbance. In classical control theory, disturbances are described as steps, ramps, and sinusoids. Here we model disturbances with signals which share properties both with deterministic signals generated by difference equations and with stochastic processes [8]. It is proved that a wide class of continuous function can be represented in the form of solutions of homogeneous differential equations [9]. That is the consequence of investigations of differential analyzers. Similar results are given by Kulebakin for electrical driver [10]. From those results, it is possible to construct Shannon-Kulebakin operator which can be very useful for the design of invariant control systems [11]. That operator is compensation matrix and has, as will be seen, a great influence on system identification [12].

Estimation of the unknown parameters is performed by Kaczmarz algorithm [13]. That is very simple and important algorithm in different areas: tomography, computer vision, synchronization in sensor networks, learning and adaptive control to name a few. In this paper, we extend the application of the algorithm to the identification of process models. The algorithm is new. As far as authors know that is the first application to the identification of MIMO models. The main feature of the algorithm is simplicity and version of the algorithm for SISO systems is applied to scanners.

2. MATHEMATICAL MODELS OF MIMO HAMMERSTEIN SYSTEM

Suppose that MIMO Hammerstein model has next structure

$$\mathbf{y}_k = \mathbf{F}^{-1}\left(q^{-1}\right)\mathbf{B}\left(q^{-1}\right)\mathbf{f}\left(\mathbf{u}_k\right)+\mathbf{d}_k . \tag{1}$$

where $\mathbf{B}\left(q^{-1}\right)$ and $\mathbf{F}\left(q^{-1}\right)$ are matrix polynomial and q^{-1} denotes the shift-back operator ($q^{-1}\mathbf{x}_k = \mathbf{x}_{k-1}$). Orders of polynomials $\mathbf{B}\left(q^{-1}\right)$ and $\mathbf{F}\left(q^{-1}\right)$ are m and n respectively

$$\mathbf{B}\left(q^{-1}\right) = \mathbf{B}_1 q^{-1} + \ldots + \mathbf{B}_m q^{-m} \tag{2}$$

$$\mathbf{F}\left(q^{-1}\right) = \mathbf{I} + \mathbf{F}_1 q^{-1} + \ldots + \mathbf{F}_n q^{-n} \tag{3}$$

where $\mathbf{B}_i$ ($i = 1,2,\ldots,m$) are $r \times r$ matrices and $\mathbf{F}_i$ ($i = 1,2,\ldots,n$) are $r \times r$ matrices. The disturbance $\mathbf{d}_k$ is piecewise polynomial disturbances. The Hammerstein model is given in the Figure 1.

*Corresponding author: Dositejeva 19, Kraljevo (Serbia),v.filipovic@mts.rs

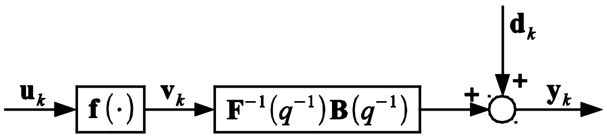

Figure 1: MIMO Hammerstein model

The function $\mathbf{f}(\mathbf{u}_k)$ is a nonlinear vector function

$$\mathbf{f}(\mathbf{u}_k)=\left[f_1\left(u_k^1\right),f_2\left(u_k^2\right),\ldots,f_r\left(u_k^r\right)\right]^T,\quad \mathbf{f}(\mathbf{u}_k)\in R^r \quad (4)$$

where $f_i\left(u_k^i\right)$ ($i=1,2,\ldots,r$) is a nonlinear function of a known basis ($\gamma_1,\gamma_2,\ldots,\gamma_r$)

$$f_i\left(u_k^i\right)=d_1^i\gamma_1\left(u_k^i\right)+d_2^i\gamma_2\left(u_k^i\right)+\ldots+d_{n_j}^i\gamma_{n_j}\left(u_k^i\right) \quad (5)$$

where d_j^i ($i=1,2,\ldots,r$; $j=1,2,\ldots,n_j$) are unknown parameters.

Let us introduced

$$s=\max_j\left\{n_j\right\} \quad (6)$$

For nonlinear vector function $\mathbf{f}(\mathbf{u}_k)$ from relations (5) and (6), it follows that

$$\mathbf{f}(\mathbf{u}_k)=\begin{bmatrix} f_1\left(u_k^1\right)\\ f_2\left(u_k^2\right)\\ \vdots\\ f_r\left(u_k^r\right)\end{bmatrix}= \quad (7)$$

$$=\begin{bmatrix} d_1^1\gamma_1\left(u_k^1\right)+d_2^1\gamma_2\left(u_k^1\right)+\ldots+d_s^1\gamma_s\left(u_k^1\right)\\ d_1^2\gamma_1\left(u_k^2\right)+d_2^2\gamma_2\left(u_k^2\right)+\ldots+d_s^2\gamma_s\left(u_k^2\right)\\ \vdots\\ d_1^r\gamma_1\left(u_k^r\right)+d_2^r\gamma_2\left(u_k^r\right)+\ldots+d_s^r\gamma_s\left(u_k^r\right)\end{bmatrix}$$

where some matrix elements, according to relation (6), are equal to zero.

Let us define

$$\mathbf{D}_i=\begin{bmatrix} d_i^1 & & \mathbf{0}\\ & d_i^2 & \\ & & \ddots & \\ \mathbf{0} & & & d_i^r\end{bmatrix},\quad i=1,2,\ldots,s \quad (8)$$

$$\mathbf{\Gamma}_i(u_k)=\begin{bmatrix}\gamma_1\left(u_k^1\right)\\ \gamma_2\left(u_k^1\right)\\ \vdots\\ \gamma_s\left(u_k^1\right)\end{bmatrix},\quad i=1,2,\ldots,s \quad (9)$$

Using relation (8) and (9) one can get

$$\mathbf{D}_i\mathbf{\Gamma}_i(\mathbf{u}_k)=\begin{bmatrix} d_i^1\gamma_1\left(u_k^1\right)\\ d_i^2\gamma_2\left(u_k^2\right)\\ \vdots\\ d_i^r\gamma_s\left(u_k^r\right)\end{bmatrix},\quad i=1,2,\ldots,s \quad (10)$$

From (7) and (10) if follows that

$$\mathbf{f}(\mathbf{u}_k)=\sum_{i=1}^{s}\mathbf{D}_i\mathbf{\Gamma}_i(\mathbf{u}_k) \quad (11)$$

Remark 1: The nonlinear function in relation (11) has general form and the first time in the literature is proposed in [14].

In what follows we will shortly describe disturbance $\mathbf{d}_k^T=\left[d_k^1,d_k^2,\ldots,d_k^r\right]$. The piecewise polynomial disturbances have next analytical form [8]

$$d_k^i=\begin{cases} d_k^{0,i}=d_0^i+d_0^ik+\ldots+\left(d_0^i\right)^{m_0}k^{m_0}, & 0\le k<n_0\\ d_k^{1,i}=d_1^i+d_1^ik+\ldots+\left(d_1^i\right)^{m_1}k^{m_1}, & n_0\le k<n_1\\ \vdots & \vdots\\ d_k^{r,i}=d_r^i+d_r^ik+\ldots+\left(d_r^i\right)^{m_r}k^{m_r}, & n_{r-1}\le k<n_r\end{cases} \quad (12)$$

where d_k^i is the i-th component of disturbance $\mathbf{d}_k$.

It is possible to distinguish two class of piecewise polynomial disturbances. The first one is the continuous disturbances (Figure 2).

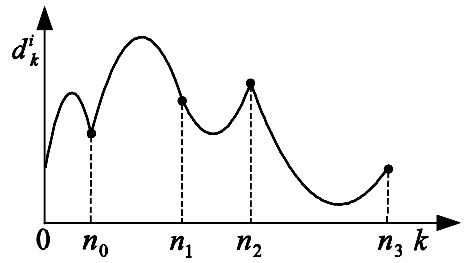

Figure 2: Piecewise polynomial disturbance (continuous case)

The second class of disturbance is discontinuous disturbances (Figure 3).

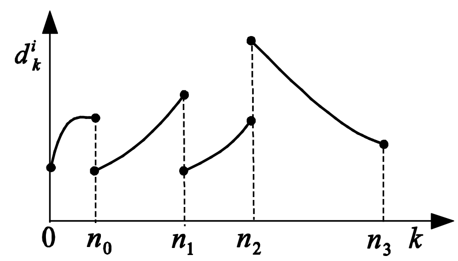

Figure 3: Piecewise polynomial disturbance (discontinuous case)

3. FUNDAMENTAL ASPECTS OF KACZMARZ ALGORITHM

The Kaczmarz method is an iterative projection algorithm for solving a linear system of equations [15]. Comprehensive treatment of algorithm and some generalization is presented in reference [16]. Due to its simplicity, the Kaczmarz algorithm has found numerous applications. The Kaczmarz algorithm is independently rediscovered in the field of image reconstruction under the name ART (algebraic reconstruction technique) [17].

Let $\mathbf{A} \in R^{m \times n}$ and $\mathbf{b} \in R^m$. The Kaczmarz method operates as follows: Initially, it starts with an arbitrary vector $\mathbf{x}^{(0)} \in R^n$. In each iteration, the Kaczmarz method runs through the rows of $\mathbf{A}$ in a cyclic manner and for each selected row, say the i-th row $\mathbf{A}^{(i)}$, it orthogonally projects the current estimate vector onto the affine hyperplane defined by the constraint of $\mathbf{Ax} = \mathbf{b}$, i.e. $\left(\mathbf{x} \middle| \left\langle \mathbf{A}^{(i)}, \mathbf{x} \right\rangle = \mathbf{b}_i \right)$, where $\langle \cdot, \cdot \rangle$ Euclidian inner product. To be more specific assuming that the i_k-th ran has been selected at the k-th iteration, than the $(k+1)$-th estimate vector $\mathbf{x}^{(k+1)}$ is obtained as

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \lambda_k \frac{\mathbf{b}_{i_k} - \left\langle \mathbf{A}^{(i_k)}, \mathbf{x}^{(k)} \right\rangle}{\left\| \mathbf{A}^{(i_k)} \right\|^2} \tag{13}$$

where $\lambda_k \in R^1$ is the relaxation parameter and $\|\cdot\|$ denotes Euclidean norm. The original Kaczmarz method corresponds to $\lambda_k = 1$ for $k \geq 0$ and all other setting of the λ_k are referred to as the relaxed Kaczmarz method [18].

The geometrical interpretation is given in the next figure.

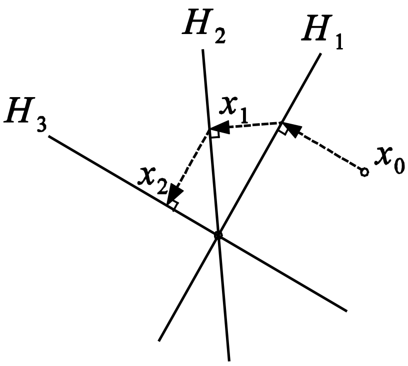

Figure 4: Geometrical interpretation of Kaczmarz algorithm

The hyperplane H_i is defined as

$$H_i = \left\{ \mathbf{x} \middle| \mathbf{a}_i^T \mathbf{x} = b_i \right\} \tag{14}$$

where the i-th row of matrix $\mathbf{A}$ is denoted with $\mathbf{a}_i^T$ and i-th element of $\mathbf{b}$ as b_i. The main task is to find solution of equation

$$\mathbf{Ax} = \mathbf{b} \tag{15}$$

where $\mathbf{A} \in R^{m \times n}$, $m \geq n$, is of full column rank and $\mathbf{b} \in R^m$. Geometrically, the solution of (15) can be thought as the intersection of all hyperplanes $\{H_i\}_{i=1}^m$.

4. RECURSIVE FORM OF KACZMARZ ALGORITHM FOR IDENTIFICATION OF MIMO HAMMERSTEIN MODELS

Generally, identification procedure depends on the disturbance. The disturbance can, however, be rejected to eliminate their effect on the system identification. Such case will be considered in this paper.

Let us suppose that disturbance satisfies a difference equation

$$\mathbf{d}_k = \mathbf{D}\left(q^{-1}\right) \mathbf{d}_{k-l-1} \tag{16}$$

where $D\left(q^{-1}\right)$ is a ($r \times r$) matrix with elements in form of polynomials. The equation (16) is the equation of extrapolation or prediction and the matrix $\mathbf{D}\left(q^{-1}\right)$ is a prediction matrix.

We can introduce Shannon-Kulebakin operator

$$\mathbf{T}\left(q^{-1}\right) = 1 - q^{-(l+1)} \mathbf{D}\left(q^{-1}\right) \tag{17}$$

From (16) and (17) it follows that

$$\mathbf{T}\left(q^{-1}\right) \mathbf{d}_k = 0 \tag{18}$$

The Shannon-Kulebakin operator we will call compensation operator or compensation matrix. In the presence of sufficiently complete information on disturbance a matrix $\mathbf{T}\left(q^{-1}\right)$ is quite easily determined.

The form of matrix $\mathbf{T}\left(q^{-1}\right)$ is

$$\mathbf{T}\left(q^{-1}\right) = \begin{bmatrix} t_1\left(q^{-1}\right) & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & t_p\left(q^{-1}\right) \end{bmatrix} \tag{19}$$

where $t_i\left(q^{-1}\right)$ are polynomials. For sake of simplicity we suppose that

$$t_1\left(q^{-1}\right) = \ldots = t_p\left(q^{-1}\right) = t\left(q^{-1}\right) \tag{20}$$

From (1) and (18) it follows that

$$\begin{aligned} \mathbf{T}\left(q^{-1}\right)\mathbf{y}_k &= \mathbf{T}\left(q^{-1}\right)\mathbf{F}^{-1}\left(q^{-1}\right)\mathbf{B}\left(q^{-1}\right)\mathbf{f}\left(\mathbf{u}_k\right) + \\ &+ \mathbf{T}\left(q^{-1}\right)\mathbf{d}_k = \mathbf{T}\left(q^{-1}\right)\mathbf{F}^{-1}\left(q^{-1}\right)\mathbf{B}\left(q^{-1}\right)\mathbf{f}\left(\mathbf{u}_k\right) \end{aligned} \tag{21}$$

From relation (20) and (21) we have

$$\mathbf{T}\left(q^{-1}\right)\mathbf{y}_k = \mathbf{F}^{-1}\left(q^{-1}\right)\mathbf{B}\left(q^{-1}\right)T\left(q^{-1}\right)\mathbf{f}\left(\mathbf{u}_k\right) \tag{22}$$

and finally

$$\mathbf{F}\left(q^{-1}\right)\mathbf{T}\left(q^{-1}\right)\mathbf{y}_k = \mathbf{B}\left(q^{-1}\right)\mathbf{T}\left(q^{-1}\right)\mathbf{f}\left(\mathbf{u}_k\right) \tag{23}$$

Using last relation we can define filtered output and input

$$\mathbf{y}_k^f = \mathbf{T}\left(q^{-1}\right)\mathbf{y}_k \tag{24}$$

$$\mathbf{f}^f\left(\mathbf{u}_k\right) = \mathbf{T}\left(q^{-1}\right)\mathbf{f}\left(\mathbf{u}_k\right) \tag{25}$$

Now equation (23) becomes

$$\mathbf{F}\left(q^{-1}\right)\mathbf{y}_k^f=\mathbf{B}\left(q^{-1}\right)\mathbf{f}\left(\mathbf{u}_k\right) \tag{26}$$

The identification problem is to estimation matrix polynomials $\mathbf{F}\left(q^{-1}\right)$ and $\mathbf{B}\left(q^{-1}\right)$. One can see that, through filtration, the influence of disturbance is removed.

We now will find the vector form of relation (26). Using relations (11), (26) and next fact that

$$q^{-l}\mathbf{\Gamma}_i\left(\mathbf{u}_k\right)=\mathbf{\Gamma}_i\left(\mathbf{u}_{k-l}\right),\qquad l=1,2,\ldots,m \tag{27}$$

we have

$$\begin{aligned}
&\mathbf{B}\left(q^{-1}\right)\mathbf{f}^f\left(\mathbf{u}_k\right)=\left(\mathbf{B}_1q^{-1}+\mathbf{B}_2q^{-2}+\ldots+\mathbf{B}_mq^{-m}\right)\mathbf{f}^f\left(\mathbf{u}_k\right)=\mathbf{B}_1\mathbf{f}^f\left(\mathbf{u}_{k-1}\right)+\mathbf{B}_2\mathbf{f}^f\left(\mathbf{u}_{k-2}\right)+\ldots+\mathbf{B}_m\mathbf{f}^f\left(\mathbf{u}_{k-m}\right)=\\
&=\mathbf{B}_1\left[\mathbf{T}\left(q^{-1}\right)\mathbf{D}_1\mathbf{\Gamma}_1\left(\mathbf{u}_{k-1}\right)+\mathbf{T}\left(q^{-1}\right)\mathbf{D}_2\mathbf{\Gamma}_2\left(\mathbf{u}_{k-1}\right)+\ldots+\mathbf{T}\left(q^{-1}\right)\mathbf{D}_s\mathbf{\Gamma}_s\left(\mathbf{u}_{k-1}\right)\right]+\\
&+\mathbf{B}_2\left[\mathbf{T}\left(q^{-1}\right)\mathbf{D}_1\mathbf{\Gamma}_1\left(\mathbf{u}_{k-2}\right)+\mathbf{T}\left(q^{-1}\right)\mathbf{D}_2\mathbf{\Gamma}_2\left(\mathbf{u}_{k-2}\right)+\ldots+\mathbf{T}\left(q^{-1}\right)\mathbf{D}_s\mathbf{\Gamma}_s\left(\mathbf{u}_{k-2}\right)\right]+\\
&\vdots\\
&\mathbf{B}_m\left[\mathbf{T}\left(q^{-1}\right)\mathbf{D}_1\mathbf{\Gamma}_1\left(\mathbf{u}_{k-m}\right)+\mathbf{T}\left(q^{-1}\right)\mathbf{D}_2\mathbf{\Gamma}_2\left(\mathbf{u}_{k-m}\right)+\ldots+\mathbf{T}\left(q^{-1}\right)\mathbf{D}_s\mathbf{\Gamma}_s\left(\mathbf{u}_{k-m}\right)\right]=\\
&=\left[\mathbf{B}_1\mathbf{T}\left(q^{-1}\right)\mathbf{D}_1\mathbf{\Gamma}_1\left(\mathbf{u}_{k-1}\right)+\mathbf{B}_2\mathbf{T}\left(q^{-1}\right)\mathbf{D}_1\mathbf{\Gamma}_1\left(\mathbf{u}_{k-2}\right)+\ldots+\mathbf{B}_m\mathbf{T}\left(q^{-1}\right)\mathbf{D}_1\mathbf{\Gamma}_1\left(\mathbf{u}_{k-m}\right)\right]+\\
&+\left[\mathbf{B}_1\mathbf{T}\left(q^{-1}\right)\mathbf{D}_2\mathbf{\Gamma}_2\left(\mathbf{u}_{k-1}\right)+\mathbf{B}_2\mathbf{T}\left(q^{-1}\right)\mathbf{D}_2\mathbf{\Gamma}_2\left(\mathbf{u}_{k-2}\right)+\ldots+\mathbf{B}_m\mathbf{T}\left(q^{-1}\right)\mathbf{D}_2\mathbf{\Gamma}_2\left(\mathbf{u}_{k-m}\right)\right]+\\
&\vdots\\
&+\left[\mathbf{B}_1\mathbf{T}\left(q^{-1}\right)\mathbf{D}_s\mathbf{\Gamma}_s\left(\mathbf{u}_{k-1}\right)+\mathbf{B}_2\mathbf{T}\left(q^{-1}\right)\mathbf{D}_s\mathbf{\Gamma}_s\left(\mathbf{u}_{k-2}\right)+\ldots+\mathbf{B}_m\mathbf{T}\left(q^{-1}\right)\mathbf{D}_s\mathbf{\Gamma}_s\left(\mathbf{u}_{k-m}\right)\right]
\end{aligned} \tag{28}$$

We also have

$$\mathbf{F}\left(q^{-1}\right)\mathbf{y}_k^f=\mathbf{y}_k^f+\mathbf{F}_1\mathbf{y}_{k-1}^f+\ldots+\mathbf{F}_n\mathbf{y}_{k-n}^f \tag{29}$$

Let us notice that

$$\begin{aligned}\mathbf{T}\left(q^{-1}\right)\mathbf{D}_i\mathbf{\Gamma}_i\left(\mathbf{u}_{k-i}\right)&=\mathbf{D}_i\left(\mathbf{T}\left(q^{-1}\right)\mathbf{\Gamma}_i\left(\mathbf{u}_{k-i}\right)\right)=\\&=\mathbf{D}_i\mathbf{\Gamma}_i^f\left(\mathbf{u}_{k-i}\right)\end{aligned} \tag{30}$$

Let us introduce the vector

$$\begin{aligned}\mathbf{X}_k^T=\Big[&-\left(\mathbf{y}_{k-1}^f\right)^T,-\left(\mathbf{y}_{k-2}^f\right)^T,\ldots,-\left(\mathbf{y}_{k-n}^f\right)^T,\\
&\left(\mathbf{\Gamma}_1^f\left(\mathbf{u}_{k-1}\right)\right)^T,\left(\mathbf{\Gamma}_1^f\left(\mathbf{u}_{k-2}\right)\right)^T,\ldots,\left(\mathbf{\Gamma}_1^f\left(\mathbf{u}_{k-m}\right)\right)^T,\\
&\left(\mathbf{\Gamma}_2^f\left(\mathbf{u}_{k-1}\right)\right)^T,\left(\mathbf{\Gamma}_2^f\left(\mathbf{u}_{k-2}\right)\right)^T,\ldots,\left(\mathbf{\Gamma}_2^f\left(\mathbf{u}_{k-m}\right)\right)^T,\\
&\left(\mathbf{\Gamma}_s^f\left(\mathbf{u}_{k-1}\right)\right)^T,\left(\mathbf{\Gamma}_s^f\left(\mathbf{u}_{k-2}\right)\right)^T,\ldots,\left(\mathbf{\Gamma}_s^f\left(\mathbf{u}_{k-m}\right)\right)^T\Big]\end{aligned} \tag{31}$$

and the matrix of parameters

$$\begin{aligned}\left(\mathbf{\theta}^M\right)^T=\big[&\mathbf{F}_1,\mathbf{F}_2,\ldots,\mathbf{F}_n,\mathbf{B}_1\mathbf{D}_1,\mathbf{B}_2\mathbf{D}_1,\ldots,\\
&\mathbf{B}_m\mathbf{D}_1,\mathbf{B}_1\mathbf{D}_2,\mathbf{B}_2\mathbf{D}_2,\ldots,\mathbf{B}_m\mathbf{D}_2,\ldots,\mathbf{B}_1\mathbf{D}_s,\\
&\mathbf{B}_2\mathbf{D}_s,\ldots,\mathbf{B}_m\mathbf{D}_s\big]\end{aligned} \tag{32}$$

From (26)-(32) it follows that

$$\mathbf{y}_k=\left(\mathbf{\theta}^M\right)^T\mathbf{x}_k^0 \tag{33}$$

Let us introduce the matrix

$$\mathbf{\varphi}_k=\begin{bmatrix}\mathbf{x}_k^T & & \mathbf{0}\\ & \ddots & \\ \mathbf{0} & & \mathbf{x}_k^T\end{bmatrix}=\mathbf{I}\otimes\mathbf{x}_k^T \tag{34}$$

where the symbol $\otimes$ denotes the Kronecker product.

Let us introduce

$$\mathbf{\theta}=\operatorname{vect}\mathbf{\theta}^M \tag{35}$$

Finally we have

$$\mathbf{y}_k=\mathbf{\varphi}_k\mathbf{\theta} \tag{36}$$

Where in last relation, according to (35), $\mathbf{\theta}$ is a vector.

The projection algorithm follows from the following optimization problem. Given estimate $\mathbf{\theta}_{k-1}$ and $\mathbf{y}_k$. Determine $\mathbf{\theta}_k$ so that criterion

$$J=\frac{1}{2}\left\|\mathbf{\theta}_k-\mathbf{\theta}_{k-1}\right\|^2 \tag{37}$$

is minimized subject to

$$\mathbf{y}_k=\mathbf{\varphi}_k\mathbf{\theta}_k \tag{38}$$

Using relations (37), (38) and introducing a Lagrange multiplier one can get next functional

$$J_c=\frac{1}{2}\left\|\mathbf{\theta}_k-\mathbf{\theta}_{k-1}\right\|^2+\mathbf{\lambda}\left[\mathbf{y}_k-\mathbf{\varphi}_k\mathbf{\theta}_k\right] \tag{39}$$

The necessary conditions for a minimum are [19]-[20]

$$\frac{\partial J_c}{\partial\mathbf{\theta}_k}=0 \tag{40}$$

$$\frac{\partial J_c}{\partial\lambda}=0 \tag{41}$$

Having in mind that

$$\mathbf{\varphi}_k=\mathbf{\varphi}_k^T \tag{42}$$

from (40) and (41) it follows that

$$\mathbf{\theta}_k-\mathbf{\theta}_{k-1}-\mathbf{\lambda}\mathbf{\varphi}_k=0 \tag{43}$$

$$\mathbf{y}_k-\mathbf{\varphi}_k^T\mathbf{\theta}_k=0 \tag{44}$$

From (43) and (44) it follows that

$$\mathbf{\lambda}=\frac{\mathbf{y}_k-\mathbf{\varphi}_k\mathbf{\theta}_{k-1}}{\mathbf{\varphi}_k^T\mathbf{\varphi}_k} \tag{45}$$

If (45) substitute to relation (43) one can get fundamental form of Kaczmarz algorithm

$$\boldsymbol{\theta}_k = \boldsymbol{\theta}_{k-1} + \frac{\boldsymbol{\varphi}_k}{\|\boldsymbol{\varphi}_k\|^2}\left(\mathbf{y}_k - \boldsymbol{\varphi}_k \boldsymbol{\theta}_{k-1}\right) \quad (46)$$

Modification of (46) is given in [21]

$$\boldsymbol{\theta}_k = \boldsymbol{\theta}_{k-1} + \frac{a\boldsymbol{\varphi}_k}{c + \|\boldsymbol{\varphi}_k\|^2}\left(\mathbf{y}_k - \boldsymbol{\varphi}_k \boldsymbol{\theta}_{k-1}\right),\ \exists c > 0,\ a \in (0,2) \quad (47)$$

Geometrical interpretation is given in next figure

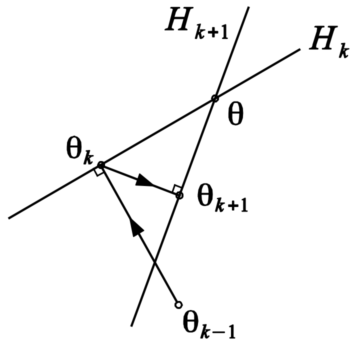

Figure 5: Geometric interpretation of recursive Kaczmarz algorithm

Remark 2: Further investigations of Kaczmarz algorithm is given in references [22]-[24].

5. CONCLUSION

The paper considers identification of MIMO Hammerstein models by using Kaczmarz algorithm. The algorithm is very simple and especially suitable for parameter estimation in the case of large data sets. It is considered very general MIMO models. The further investigations will be performed for next problems

(i) Relaxation of condition (20) in the paper
(ii) Design of adaptive invariant systems
(iii) Design of robust stochastic recursive Kaczmarz algorithms.

Also, in the field of block-oriented methods, the potential application of iterative Kaczmarz algorithm is very important (especially randomized Kaczmarz algorithm).

ACKNOWLEDGEMENTS

This paper is part of projects TR33026 and TR33027 at the University of Kragujevac, Faculty of Mechanical and Civil Engineering in Kraljevo, and it was supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia.

REFERENCES

[1] Y. Zhu, "Multivariable System Identification for Process Control", Elsevier Science, London (UK), (2000)

[2] E. Giri and E. Bai (Eds.), "Block-oriented Nonlinear System Identification", Springer, Berlin (Germany), (2010)

[3] L. Ljung, "System Identification: Theory for the User", Prentice Hall, New Jersey (USA), (1999)

[4] V. Filipovic, "Recursive Identification of Multivariable ARX Models in the Presence of a Priori Information: Robustness and Regularization", Signal Processing, Vol. 116, pp. 351-357, (2015)

[5] V. Filipovic, "A Global Convergent Outlier Robust Adaptive Predictor for MIMO Hammerstein Models", International Journal of Robust and Nonlinear Control, DOI: 10.1002/rnc.3705, (2016)

[6] R. Pearson, "Exploring Data in Engineering, the Sciences and Medicine", Oxford University Press, Oxford (UK), (2011)

[7] P. Huber, "Robust Estimation of a Location Parameters", Annals of Mathematical Statistics, Vol. 35, pp. 73-101, (1964)

[8] V. Filipovic, "Industrial Controllers", University of Kragujevac, Kraljevo, (2016), (in Serbian)

[9] C. Shannon, "Mathematical Theory of the Differential Analyzer", Journal of Mathematics and Physics, Vol. 20, pp. 337-352, (1941)

[10] V. Kulebakin, "On the Behaviour of the Continuously Distributed Automatic Linear Systems", Proceeding of the USSR Academy of Science, Vol. 68, pp. 73-79, (1949)

[11] Ya. Tsypkin, "Adaptive-Invariant Control Systems", Automatika i Telemehanika, Vol. 55, pp. 96-124, (1991)

[12] Ya. Tsypkin, J. Manson and K. Warwick, "Identification of Linear Systems in the Presence of Piecewise Polynomial Disturbances", IEE Proceedings – Control Theory Application, Vol. 143, Issue 4, pp. 305-308, (1996), (in Russian)

[13] S. Kaczmarz, "Angenäherte Auflösung von Systemen Linearer Gleichungen", Bulletin International de l'Academie Polonaise des Sciences et des Lettres, Vol. 35, pp. 355-357, (1937)

[14] V. Filipovic, "Robust Stochastic Approximation Algorithm for Identification of MIMO Hammerstein OE Models", Nonlinear Dynamics, Accepted for Publication, (2017)

[15] P. Parks, "S. Kaczmarz (1895-1939)", International Journal of Control, Vol. 57, pp. 1263-1267, (1993)

[16] A. Galantal, "Projectors and Projection Methods", Koluwer, Boston (USA), (2004)

[17] R. Gordon, R. Bender and G. Herman, "Algebraic Reconstruction Technique (ART) for Three-dimensional Electron Microscopy and X-ray Photography", Journal of Theoretical Biology, Vol. 29, Issue 3, pp. 471-481, (1970)

[18] Y. Censor, "Row-Action Methods for Huge and Sparse Systems and Their Application", SIAM Review, Vol. 23, Issue 4 pp. 444-466, (1981)

[19] G. Goodwin and K. Sin, “Adaptive Filtering, Prediction and Control”, Prentice Hall, New Jersey (USA), (1984)

[20] E. Aved’Yan, “Learning Systems”, Springer, Berlin (Germany), (1996)

[21] T. Strohmer and R. Vershynin, “A Randomized Kaczmarz Algorithm with Exponential Convergence”, J. Fourier Anal. Appl. Vol. 15, pp. 266-278, (2009)

[22] V. Filipovc and V. Djordjevic, “Kaczmarz Algorithm Based on Synergy of Fuzzy Identification and Machine Learning”, (2017), to be Published

[23] V. Filipovic, “Convergence Properties of Kaczmarz Algorithm for Identification of MIMO Hammerstein Models”, (2017), to be published

[24] V. Filipovic and V. Djordjevic, “Design of Multivariable PID controller for Boiler Plant Using Convex Analysis and Theory of Invariant Systems”, (2017), to be published