Control and Cybernetics
vol. **49** (2020) No. 3

Bayesian multidimensional-matrix polynomial empirical regression*

by
Vladimir S. Mukha

Belarusian State University of Informatics and Radioelectronics
P. Brovki str. 6, Minsk, 220013, Republic of Belarus
mukha@bsuir.by

Abstract: The problem of parameter estimation for the polynomial in the input variables regression function is formulated and solved. The input and output variables of the regression function are multidimensional matrices. The parameters of the regression function are assumed to be random independent multidimensional matrices with Gaussian distribution and known mean value and variance matrices. The solution to this problem is a multidimensional-matrix system of the linear algebraic equations in multidimensional-matrix unknown regression function parameters. We consider the particular cases of constant, affine and quadratic regression function, for which we have obtained formulas for parameter calculation. Computer simulation of the quadratic regression function is performed for the two-dimensional matrix input and output variables.

Keywords: regression function, parameter estimation, maximum likelihood estimation, Bayesian estimation, multidimensional matrices

1. Introduction

To date, the most popular methods of estimating the parameters of the regression function are the maximum likelihood method and the least squares method, see, e.g. Ermakov and Zhiglyavskii (1987) and Klepikov and Sokolov (1961). The estimations, obtained with the use of these methods have good asymptotic properties and this is the justification of their application. But the use of these methods becomes problematic in the case of small sizes of the samples. In this case the Bayesian approach to the estimation of the parameters of the regression function is more attractive. The case of a small size of the sample is very important for the problem of dual control of regression objects when the regression object is being studied and controlled simultaneously starting from the initial moment in time (Feldbaum, 1966; Mukha and Sergeev, 1974, 1976). The

*Submitted: March 2019. Accepted: September 2020.

interest in Bayesian inference appears also in econometrics, see Zellner (1971), as well as in other areas, see Fahrmeir and Kneib (2011), or Wakefield (2013).

The existing investigations of the Bayesian approach to date relate mainly to the regression functions that are linear in the parameters and in the input variables. Such regression functions can be represented in terms of usual matrices. The regression functions nonlinear in the input variables can be represented as a scalar product of the vector of the parameters and the vector of the basis functions, see Mukha and Sergeev (1974). However, this approach is poorly formalized and does not feature the algorithmic generality; i.e. the mathematical expression for the vector of the basis functions is not determined and the software implementation is inapplicable for any number of variables and any degree of the polynomial. It is a manual approach that is suitable for objects with a small number of variables. This disadvantage is overcome in the present paper within the framework of multidimensional-matrix mathematical approach. The multidimensional-matrix approach uses the concept of a multidimensional matrix and it is the approach that places the respective technique in the context of "big data".

The introduction to the theory of multidimensional matrices was developed in the work of Sokolov (1972). This theory has its roots in the works of A. Cayley (1842), R.F. Scott (1879/80), and other scientists. An extensive list of literature in English is available in the work of Sokolov (1960). The multidimensional-matrix approach has found effective application in many areas of research (see Mukha, 2005, 2006, 2007a,b, 2008, 2011, 2012, 2017). We present in the Appendix the main definitions of this theory for a better understanding of the content of the article.

2. Problem statement

Let us consider some object with q-dimensional-matrix input variable $\bar{x} = (x_j)$, $j = (j_1, j_2, ..., j_q)$, p-dimensional-matrix output variable $\bar{\eta} = (\eta_i)$, $i = (i_1, i_2, ..., i_p)$, see Sokolov (1972), Mukha (2004), and (A.1), and suppose that the output variable $\bar{\eta}$ is stochastically dependent on the input variable $\bar{x}$, so that there is a conditional probability density $f(\bar{\eta}/\bar{x})$. We denote by $\bar{y} = \phi(\bar{x})$ the regression function $\bar{\eta}$ with respect to $\bar{x}$ and assume that the dependence of $\bar{\eta}$ on $\bar{x}$ can be represented in the form $\bar{\eta} = \phi(\bar{x}) + \bar{\varepsilon}$, where $\bar{\varepsilon}$ is a p-dimensional random matrix. Assume that for some values $\bar{x}_1, \bar{x}_2, ..., \bar{x}_n$ of the input variable $\bar{x}$ we obtained the values $\bar{y}_1, \bar{y}_2, ..., \bar{y}_n$ of the output variable $\bar{\eta}$ (observations, measurements) as follows:

$$\bar{y}_\mu = \phi(\bar{x}_\mu) + \bar{z}_\mu, \quad \mu = 1, ..., n, \tag{1}$$

where $\bar{z}_\mu$ is a realization of the random matrix $\bar{\varepsilon}$, which we will refer to as errors of the measurements, and μ is the sequential number of the observation. We will consider the Gaussian distribution of the matrix $\bar{\varepsilon}$ with zero mean value and variance matrix R_ε, see Mukha (2004), and (A.7).

Hereinafter we will use the following notations for indices of multidimensional matrices: $i_1, i_2, ...,$ are separate indices, $\bar{i}_{(p)} = (i_1, i_2, ..., i_p)$ is the set of p indices (p-multi-index), $\bar{\bar{i}}_{(p,k)} = (\bar{i}_{(p),1}, \bar{i}_{(p),2}, ..., \bar{i}_{(p),k})$ is the set of k p-multi-indices.

Let the hypothetic regression function be the polynomial of m-th degree, Mukha (2004):

$$\phi(\bar{x}) = \sum_{k=0}^{m} {}^{0,kq}(C_{(p,kq)}\bar{x}^k) = \sum_{k=0}^{m} {}^{0,kq}(\bar{x}^k C_{(kq,p)}), \quad m = 0, 1, 2, ..., \tag{2}$$

where $C_{(p,kq)}$ and $C_{(kq,p)}$ are multidimensional-matrix parameters of the regression function, $C_{(p,kq)}$ is $(p + kq)$-multidimensional matrix:

$$C_{(p,kq)} = (c_{\bar{i}_{(p)},\bar{\bar{j}}_{(q,k)}}), \quad \bar{i}_{(p)} = (i_1, i_2, ..., i_p), \quad \bar{\bar{j}}_{(q,k)} = (\bar{j}_{(q)1}, \bar{j}_{(q)2}, ..., \bar{j}_{(q)k}).$$

It is symmetric relative to p-multi-indices $\bar{j}_{(q)1}, \bar{j}_{(q)2}, ..., \bar{j}_{(q)k}$. The matrix $C_{(kq,p)}$ is the transpose of the matrix $C_{(p,kq)}$, i.e.

$$C_{(p,kq)} = (C_{(kq,p)})^{H_{p+kq,kq}}, \quad C_{(kq,p)} = (C_{(p,kq)})^{B_{p+kq,kq}},$$

where $H_{p+kq,kq}$ and $B_{p+kq,kq}$ are transpose substitutions of the types 'back' and 'onward', respectively (Mukha, 2004, (A.4), (A.5), (A.6)). We also denote by ${}^{0,kq}(C_{(p,kq)}x^k)$ the $(0, kq)$-folded product of matrices $C_{(p,kq)}$, and $\bar{x}^k$, $\bar{x}^k = {}^{0,0}\bar{x}^k$ is the $(0, 0)$-folded degree of the matrix $\bar{x}$ (Sokolov, 1972; Mukha, 2004; (A.2)).

In these conditions the measurement $\bar{y}_\mu$, (1), has the probability density

$$f(\bar{y}_\mu/\bar{x}_\mu, C_{(p,0)}, C_{(p,q)}, ..., C_{(p,mq)}) =$$

$$= C_y \exp\left(-\frac{1}{2}\,{}^{0,2p}(R_\varepsilon^{-1}(\bar{y}_\mu - \sum_{k=0}^{m} {}^{0,kq}(C_{(p,kq)}\bar{x}_\mu^k))^2)\right), \quad \mu = 1, ..., n, \tag{3}$$

where C_y is a normalizing constant, R_ε^{-1} is $(0, p)$-inverse to R_ε matrix (Sokolov, 1972; Mukha, 2004; (A.3)).

The problem consists in finding the estimations of parameters $C_{(p,kq)}$ $(C_{(kq,p)})$ of the regression function (2) by using the given measurements $(\bar{x}_1, \bar{y}_1), (\bar{x}_2, \bar{y}_2)$ $, ..., (\bar{x}_n, \bar{y}_n)$.

3. Maximum likelihood multidimensional-matrix polynomial empirical regression

In this section we will find the maximum likelihood estimations (ML-estimations) of the parameters of the polynomial multidimensional-matrix regression function. We formulate the main result in the form of the following theorem.

Theorem 1 *Let $\phi(\bar{x})$ be a multidimensional-matrix regression under conditions (1)–(3). Then, the ML-estimations $\hat{C}_{(p,0)}, \hat{C}_{(p,q)}, ..., \hat{C}_{(p,mq)}$ of its parameters $C_{(p,0)}, C_{(p,q)}, ..., C_{(p,mq)}$ are defined as the solution of the following multidinensional-matrix linear system of equations:*

$$\sum_{k=0}^{m}{}^{0,kq}(C_{(p,kq)} S_{x^k x^\lambda}) = S_{yx^\lambda}, \quad \lambda = 0, 1, ..., m, \tag{4}$$

where

$$S_{yx^\lambda} = \sum_{\mu=1}^{n}{}^{0,0}(y_\mu x_\mu^\lambda), \quad S_{x^k x^\lambda} = \sum_{\mu=1}^{n}{}^{0,0}(x_\mu^k x_\mu^\lambda). \tag{5}$$

Proof. The logarithmic likelihood function is defined as follows:

$$\ln f_n(C_{(p,0)}, ..., C_{(p,mq)}) = \sum_{\mu=1}^{n} \ln f(\bar{y}_\mu / \bar{x}_\mu, C_{(p,0)}, ..., C_{(p,mq)}).$$

The necessary conditions of the maximum of the logarithmic likelihood function are defined by the following system of equations:

$$\frac{d}{dC_{(p,\lambda q)}} \sum_{\mu=1}^{n} \ln f(\bar{y}_\mu / \bar{x}_\mu, C_{(p,0)}, ..., C_{(p,mq)}) = 0, \quad \lambda = 0, 1, ..., m.$$

We will obtain this system of equation for our case. The likelihood function, as it follows from the distribution (3), is defined by the expression

$$\ln f(\bar{y}_\mu / \bar{x}_\mu, C_{(p,0)}, ..., C_{(p,mq)}) \sim -\frac{1}{2}{}^{0,2p}\left(R_\varepsilon^{-1} \left(\bar{y}_\mu - \sum_{k=0}^{m}{}^{0,kq}(C_{(p,kq)} \bar{x}_\mu^k) \right)^2 \right),$$

where symbol $\sim$ means equivalence for the maximization. Now we find the derivatives of this function on the parameters $C_{(p,\lambda q)}$. For this, we denote

$$W_\mu^{(\lambda)} = {}^{0,\lambda q}(C_{(p,\lambda q)} \bar{x}_\mu^\lambda) = (w_{\bar{i}_{(p)}}).$$

By the rules of the multidimensional-matrix differentiation, see Mukha (2004), we have

$$\frac{W_\mu^{(\lambda)}}{dC_{(p,\lambda q)}} = U_\mu^{(\lambda)} = \left(\frac{dw_{\bar{i}_{(p)}}}{dc_{\bar{j}_{(p)}, \bar{\bar{t}}_{(q,\lambda)}}}\right) = (u_{\bar{i}_{(p)}, \bar{j}_{(p)}, \bar{\bar{t}}_{(q,\lambda)}}) =$$

$$={}^{0,\lambda q}\,(E(0,p+\lambda q)x_\mu^\lambda)^T=(Z_\mu^{(\lambda)})^T,$$

where $Z_\mu^{(\lambda)}=\ {}^{0,\lambda q}(E(0,p+\lambda q)x_\mu^\lambda)$, $E(0,p+\lambda q)$ is the $(0,p+\lambda q)$-identity matrix (Sokolov, 1972; Mukha, 2004), T denotes transpose substitution of the 'back' type on the $2p+\lambda q$ indices $H_{p+\lambda q+p,p+\lambda q}$ (Mukha, 2004), i. e.

$$T=H_{2p+\lambda q,p+\lambda q}=\begin{pmatrix}\bar{i}_{(p)},\bar{j}_{(p)},\bar{\bar{t}}_{(q,\lambda)}\\ \bar{j}_{(p)},\bar{\bar{t}}_{(q,\lambda)},\bar{i}_{(p)}\end{pmatrix}.$$

We convert the matrix $Z_\mu^{(\lambda)}$ considering the properties of the identity matrix $E(0,p+\lambda q)$:

$$Z_\mu^{(\lambda)}={}^{0,\lambda q}\,(E(0,p+\lambda q)\bar{x}_\mu^\lambda)=$$

$$=(\sum_{\bar{\bar{r}}_{(q,\lambda)}}e_{\bar{i}_{(p)},\bar{\bar{t}}_{(q,\lambda)},\bar{j}_{(p)},\bar{\bar{r}}_{(q,\lambda)}}x_{\bar{\bar{r}}_\lambda})=(z_{\bar{i}_{(p)},\bar{\bar{t}}_\lambda,\bar{j}})=(\sum_{\bar{\bar{r}}_\lambda}e_{\bar{i}_{(p)},\bar{j}_{(p)}}e_{\bar{\bar{t}}_{(q,\lambda)},\bar{\bar{r}}_{(q,\lambda)}}x_{\bar{\bar{r}}_{(q,\lambda)}})=$$

$$=(e_{\bar{i}_{(p)},\bar{j}_{(p)}}x_{\bar{\bar{t}}_{(q,\lambda)}})={}^{0,0}(E(0,p)\bar{x}_\mu^\lambda).$$

From the transpose relationship

$$(Z_\mu^{(\lambda)})^T=(Z_\mu^{(\lambda)})^{\begin{pmatrix}\bar{i}_{(p)},\bar{j}_{(p)},\bar{\bar{t}}_{(q,\lambda)}\\ \bar{j}_{(p)},\bar{\bar{t}}_{(q,\lambda)},\bar{i}_{(p)}\end{pmatrix}}=U_\mu^{(\lambda)}$$

it follows that $u_{\bar{i}_{(p)},\bar{j}_{(p)},\bar{\bar{t}}_{(q,\lambda)}}=z_{\bar{i}_{(p)},\bar{\bar{t}}_{(q,\lambda)},\bar{j}_{(p)}}=e_{\bar{i}_{(p)},\bar{j}_{(p)}}x_{\bar{\bar{t}}_{(q,\lambda)}}$, so that

$$U_\mu^{(\lambda)}={}^{0,0}(E(0,p)\bar{x}_\mu^\lambda). \tag{6}$$

Further, considering the previous results, we have

$$\frac{d}{dC_{(p,\lambda q)}}{}^{0,2p}\left(R_\varepsilon^{-1}\left(\bar{y}_\mu-\sum_{k=0}^{m}{}^{0,kq}(C_{(p,kq)}\bar{x}_\mu^k)\right)^2\right)=$$

$$={}^{0,p}\left({}^{0,p}\left(R_\varepsilon^{-1}\left(\bar{y}_\mu-\sum_{k=0}^{m}{}^{0,kq}(C_{(p,kq)}\bar{x}_\mu^k)\right)\right)U_\mu^\lambda\right)=$$

$$={}^{0,p}\left({}^{0,p}\left(R_\varepsilon^{-1}\bar{y}_\mu\right)U_\mu^\lambda\right)-{}^{0,p}\left({}^{0,p}\left(R_\varepsilon^{-1}\sum_{k=0}^{m}{}^{0,kq}(C_{(p,kq)}\bar{x}_\mu^k)\right)U_\mu^\lambda\right).$$

If we take into account the equation (6), we obtain

$$\frac{d}{dC_{(p,\lambda q)}}\ln f(\bar{y}_\mu/\bar{x}_\mu,C_{(p,0)},...,C_{(p,mq)})=$$

$$={}^{0,p}\left({}^{0,p}\left(R_\varepsilon^{-1}\bar{y}_\mu\right){}^{0,0}(E(0,p)\bar{x}_\mu^\lambda)\right)-$$

$$-{}^{0,p}\left({}^{0,p}\left(R_\varepsilon^{-1}\sum_{k=0}^{m}{}^{0,kq}(C_{(p,kq)}\bar{x}_\mu^k)\right){}^{0,0}(E(0,p)\bar{x}_\mu^\lambda)\right)=$$

$$={}^{0,0}\left({}^{0,p}\left(R_\varepsilon^{-1}\bar{y}_\mu\right)\bar{x}_\mu^\lambda)\right)-{}^{0,0}\left({}^{0,p}\left(R_\varepsilon^{-1}\sum_{k=0}^{m}{}^{0,kq}(C_{(p,kq)}\bar{x}_\mu^k)\right)\bar{x}_\mu^\lambda)\right)=$$

$$={}^{0,p}\left(R_\varepsilon^{-1}\,{}^{0,0}(\bar{y}_\mu\bar{x}_\mu^\lambda)\right)-{}^{0,p}\left(R_\varepsilon^{-1}\sum_{k=0}^{m}{}^{0,0}({}^{0,kq}(C_{(p,kq)}\bar{x}_\mu^k)\bar{x}_\mu^\lambda)\right)=$$

$$={}^{0,p}\left(R_\varepsilon^{-1}\,{}^{0,0}(\bar{y}_\mu\bar{x}_\mu^\lambda)\right)-{}^{0,p}\left(R_\varepsilon^{-1}\sum_{k=0}^{m}{}^{0,kq}(C_{(p,kq)}\,{}^{0,0}(\bar{x}_\mu^k\bar{x}_\mu^\lambda))\right).$$

Now, by summing over $\mu=1,...,n$ and denoting

$$\sum_{\mu=1}^{n}{}^{0,0}(\bar{y}_\mu\bar{x}_\mu^\lambda)=S_{yx^\lambda},$$

$$\sum_{\mu=1}^{n}{}^{0,0}(\bar{x}_\mu^k\bar{x}_\mu^\lambda)=S_{x^kx^\lambda},$$

$$\sum_{\mu=1}^{n}{}^{0,0}(\bar{x}_\mu^0\bar{x}_\mu^0)=S_{x^0x^0}=n$$

we obtain the derivative

$$\frac{d}{dC_{(p,\lambda q)}}\sum_{\mu=1}^{n}\ln f(\bar{y}_\mu/\bar{x}_\mu,C_{(p,0)},...,C_{(p,mq)})=$$

$$={}^{0,p}(R_\varepsilon^{-1}\,S_{yx^\lambda})-{}^{0,p}(R_\varepsilon^{-1}\sum_{k=0}^{m}{}^{0,kq}(C_{(p,kq)}S_{x^kx^\lambda}))\tag{7}$$

and the system of equations for the parameters of the maximum likelihood regression

$${}^{0,p}(R_\varepsilon^{-1}\sum_{k=0}^{m}{}^{0,kq}(C_{(p,kq)}S_{x^kx^\lambda}))={}^{0,p}(R_\varepsilon^{-1}\,S_{yx^\lambda}),\quad\lambda=0,1,...,m.$$

If we multiply both sides of these equations from the left by R_ε in the sense of $(0,p)$-folded multiplication, we obtain the system of equations (4). Thus, we have proven Theorem 1. □

We notice that this system of equations coincides with the one obtained by the least squares method in Mukha (2007b) and also with the one obtained on the basis of the best theoretical multidimensional-matrix polynomial regression (Mukha, 2007a, 2011).

4. Bayesian multidimensional-matrix polynomial empirical regression

Let us consider multidimensional-matrix polynomial regression (2) in the form

$$\phi(\bar{x}) = \sum_{k=0}^{m} {}^{0,kq}(\tilde{C}_{(p,kq)}\bar{x}^k) = \sum_{k=0}^{m} {}^{0,kq}(\bar{x}^k \tilde{C}_{(kq,p)}), \quad m = 0,1,2,... \quad . \tag{8}$$

In addition to the assumptions (1)–(3) we will consider the parameter $\tilde{C}_{(p,kq)}$ of the multidimensional-matrix polynomial regression (8) as a random matrix with Gaussian a priori probability density

$$f_a(C_{(p,kq)}) = K_{(p,kq)} \exp(-\frac{1}{2}({}^{0,2(p+kq)}(R^{-1}_{a,(p,kq)}(C_{(p,kq)} - C_{a,(p,kq)})^2), \tag{9}$$

$$k = 0,1,2,...,m, \quad m = 0,1,2,... \quad ,$$

where $K_{(p,kq)}$ is a normalizing constant,

$C_{a,(p,kq)} = (C_{a,(kq,p)})^{H_{p+kq,kq}}$, $\quad C_{a,(kq,p)} = (C_{a,(p,kq)})^{B_{p+kq,kq}}$ is a priori mean value ($(p+kq)$-dimensional matrix),

$R_{a,(p,kq)} = (R_{a,(kq,p)})^{(H_{p+kq,kq},H_{p+kq,kq})}$, $\quad R_{a,(kq,p)} = (R_{a,(p,kq)})^{(B_{p+kq,kq},B_{p+kq,kq})}$ is a priori variance matrix ($2(p+kq)$-dimensional matrix),

$R^{-1}_{a,(p,kq)} = (R^{-1}_{a,(kq,p)})^{(H_{p+kq,kq},H_{p+kq,kq})}$, $\quad R^{-1}_{a,(kq,p)} = (R^{-1}_{a,(p,kq)})^{(B_{p+kq,kq},B_{p+kq,kq})}$ are $(0,p+kq)$-inverses to the $R_{a,(p,kq)}$, $R_{a,(kq,p)}$ matrices, respectively.

We will assume that the parameters $\tilde{C}_{(p,0)}$, $\tilde{C}_{(p,q)}, \ldots$, $\tilde{C}_{(p,mq)}$ are independent, i.e.

$$f_a(C_{(p,0)},...,C_{(p,mq)}) = \prod_{k=0}^{m} f_a(C_{(p,kq)}). \tag{10}$$

With these assumptions, on the basis of measurements $(\bar{x}_1,\bar{y}_1)$, $(\bar{x}_2,\bar{y}_2)$, ..., $(\bar{x}_n,\bar{y}_n)$ we will find the Bayesian estimations $\hat{C}_{(p,0)}$, $\hat{C}_{(p,q)}$, $\ldots$, $\hat{C}_{(p,mq)}$ of the unknown values $C_{(p,0)}$, $C_{(p,q)}, \ldots$, $C_{(p,mq)}$ of the parameters $\tilde{C}_{(p,0)}$, $\tilde{C}_{(p,q)}, \ldots$, $\tilde{C}_{(p,mq)}$, i.e. the estimations minimizing the average risk:

$$r = E(W(\tilde{C}_{(p,0q)},...,\tilde{C}_{(p,mq)},\hat{C}_{(p,0q)},...,\widehat{C}_{(p,mq)})),$$

where $W(\tilde{C}_{(p,0q)},...,\tilde{C}_{(p,mq)},\hat{C}_{(p,0q)},...,\hat{C}_{(p,mq)})$ is the loss function, and E is the symbol of mathematical expectation.

Theorem 2 *Under conditions (1)–(3), (9), (10) relative to the multidimensional-matrix polynomial regression (8) and quadratic loss function, the Bayesian estimations* $\hat{C}_{(p,0)}, \hat{C}_{(p,q)}, ..., \hat{C}_{(p,mq)}$ *of the parameters* $\tilde{C}_{(p,0)}, \tilde{C}_{(p,q)}, ..., \tilde{C}_{(p,mq)}$ *satisfy the following system of linear multidimensional-matrix equations:*

$$ {}^{0,(p+\lambda q)}(R^{-1}_{a,(p,\lambda q)} C_{(p,\lambda q)}) + \sum_{k=0}^{m} {}^{0,(p+kq)}(V_{k,\lambda}^{T_{k,\lambda}} C_{(p,kq)}) = $$

$$ = {}^{0,p}(R_{\varepsilon}^{-1} S_{yx^{\lambda}}) + {}^{0,(p+\lambda q)}(R^{-1}_{a,(p,\lambda q)} C_{a,(p,\lambda q)}), \quad \lambda = 0, 1, ..., m, \tag{11} $$

where $S_{yx^{\lambda}}$ *and* $S_{x^k x^{\lambda}}$ *are defined by expressions (5),* $V_{k,\lambda}$ *is* $(2p + kq + \lambda q)$*-dimensional matrix,*

$$ V_{k,\lambda} = {}^{0,0}(R_{\varepsilon}^{-1} S_{x^k x^{\lambda}}), $$

R_{ε}^{-1} *is* $(0, p)$*-inverse to the* R_{ε} *matrix,* $V_{k,\lambda}^{T_{k,\lambda}}$ *is transposed in accordance with substitution* $T_{k,\lambda}$ *matrix* $V_{k,\lambda}$*, and*

$$ T_{k,\lambda} = \begin{pmatrix} \bar{i}_{(p)}, \bar{\bar{v}}_{(q,\lambda)}, \bar{j}_{(p)}, \bar{\bar{t}}_{(q,k)} \\ \bar{i}_{(p)}, \bar{j}_{(p)}, \bar{\bar{t}}_{(q,k)}, \bar{\bar{v}}_{(q,\lambda)} \end{pmatrix}. $$

Proof. The posterior probability density of the parameters, considering the independence of the parameters, (10), is defined as follows

$$ f_n(C_{(p,0)}, ..., C_{(p,mq)}) = \frac{\prod\limits_{k=0}^{m} f_a(C_{(p,kq)}) \prod\limits_{\mu=1}^{n} f(\bar{y}_{\mu}/\bar{x}_{\mu}, C_{(p,0)}, ..., C_{(p,mq)})}{\int\limits_{\bar{C}(\omega)} \prod\limits_{k=0}^{m} f_a(C_{(p,kq)}) \prod\limits_{\mu=1}^{n} f(\bar{y}_{\mu}/\bar{x}_{\mu}, C_{(p,0)}, ..., C_{(p,mq)}) d\omega}. \tag{12} $$

As it is known, Bayesian estimations are defined as a posteriori mean value when the loss function is quadratic. The posterior probability density will be Gaussian due to the linearity of the regression function in the parameters, and the estimations may be obtained as maximizing the a posteriori probability density, i.e. by maximizing the numerator in the formula (12). The numerator in the formula (12) will be referred to as the joint likelihood function and we will use the logarithmic joint likelihood function:

$$ \ln f_n(C_{(p,0)}, ..., C_{(p,mq)}) = \sum_{k=0}^{m} \ln f_a(C_{(p,kq)}) + \sum_{\mu=1}^{n} \ln f(\bar{y}_{\mu}/\bar{x}_{\mu}, C_{(p,0)}, ..., C_{(p,mq)}). \tag{13} $$

The necessary conditions of the maximum of the logarithmic joint likelihood function (13) are presented as the following system of equations:

$$\frac{d}{dC_{(p,\lambda q)}}\sum_{k=0}^{m}\ln f_a(C_{(p,kq)})+\frac{d}{dC_{(p,\lambda q)}}\sum_{\mu=1}^{n}\ln f(\bar{y}_\mu/\bar{x}_\mu, C_{(p,0)},...,C_{(p,mq)})=0,$$
$$\lambda=0,1,...,m. \tag{14}$$

This system of equations defines the parameters of the Bayesian multidimensional-matrix polynomial empirical regression. It differs from the system (4) in the first summand. In the light of the a priori probability density (9) we will have for the first summand:

$$\ln f_a(C_{(p,\lambda q)})\sim -\frac{1}{2}({}^{0,2(p+\lambda q)}(R^{-1}_{a,(p,\lambda q)}(C_{(p,\lambda q)}-C_{a,(p,\lambda q)})^2),$$

$$\frac{d}{dC_{(p,\lambda q)}}\ln f_a(C_{(p,\lambda q)}) = -{}^{0,(p+\lambda q)}(R^{-1}_{a,(p,\lambda q)}(C_{(p,\lambda q)}-C_{a,(p,\lambda q)}))=$$
$$= -{}^{0,(p+\lambda q)}(R^{-1}_{a,(p,\lambda q)}C_{(p,\lambda q)})+{}^{0,(p+\lambda q)}(R^{-1}_{a,(p,\lambda q)}C_{a,(p,\lambda q)}). \tag{15}$$

Taking into account the expressions (15), (7), we obtain the following system of equations instead of (4):

$${}^{0,(p+\lambda q)}(R^{-1}_{a,(p,\lambda q)}C_{(p,\lambda q)})+{}^{0,p}(R_\varepsilon^{-1}\sum_{k=0}^{m}{}^{0,kq}(C_{(p,kq)}S_{x^kx^\lambda}))=$$
$$={}^{0,p}(R_\varepsilon^{-1}\,S_{yx^\lambda})+{}^{0,(p+\lambda q)}(R^{-1}_{a,(p,\lambda q)}C_{a,(p,\lambda q)}),\quad \lambda=0,1,...,m. \tag{16}$$

Now we transform the second summand on the left-hand side of the system (16) to look like the first one. For this, we perform the following transformations for the second summand on the left-hand side of the system (16):

$$Z_k={}^{0,p}(R_\varepsilon^{-1}\,{}^{0,kq}(C_{(p,kq)}S_{x^kx^\lambda}))=$$
$$=(\sum_{\bar{j}_{(p)}}r_{\bar{i}_{(p)},\bar{j}_{(p)}}\sum_{\bar{\bar{t}}_{(q,k)}}c_{\bar{j}_{(p)},\bar{\bar{t}}_{(q,k)}}s_{\bar{\bar{t}}_{(q,k)},\bar{\bar{v}}_{(q,\lambda)}})=(z_{\bar{i}_{(p)},\bar{\bar{v}}_{(q,\lambda)}})$$
$$=(\sum_{\bar{j}_{(p)}}\sum_{\bar{\bar{t}}_{(q,k)}}r_{\bar{i}_{(p)},\bar{j}}s_{\bar{\bar{t}}_{(q,k)},\bar{\bar{v}}_{(q,\lambda)}}c_{\bar{j}_{(p)},\bar{\bar{t}}_{(q,k)}})={}^{0,(p+kq)}(V_{k,\lambda}^{T_{k,\lambda}}C_{(p,kq)}), \tag{17}$$

where

$$V_{k,\lambda}={}^{0,0}(R_\varepsilon^{-1}S_{x^kx^\lambda}),\quad k,\lambda=0,1,...,m, \tag{18}$$

R_ε^{-1} is $(0,p)$-inverse to the R_ε matrix, $V_{k,\lambda}^{T_{k,\lambda}}$ is transposed, in accordance with substitution $T_{k,\lambda}$, matrix $V_{k,\lambda}$. We find the substitution $T_{k,\lambda}$. Since, in accordance with the formula (18), $V_{k,\lambda}=(v_{\bar{i}_{(p)},\bar{j}_{(p)},\bar{\bar{t}}_{(q,k)},\bar{\bar{v}}_{(q,\lambda)}})$ and, in accordance with the formula (18), $V_{k,\lambda}^{T_{k,\lambda}} = (v^{T_{k,\lambda}}_{\bar{i}_{(p)},\bar{\bar{v}}_{(q,\lambda)},\bar{j}_{(p)},\bar{\bar{t}}_{(q,k)},})$, then we have the equality (Sokolov, 1960) $v^{T_{k,\lambda}}_{\bar{i}_{(p)},\bar{\bar{v}}_{(q,\lambda)},\bar{j}_{(p)},\bar{\bar{t}}_{(q,k)}} = v_{\bar{i}_{(p)},\bar{j}_{(p)},\bar{\bar{t}}_{(q,k)},\bar{\bar{v}}_{(q,\lambda)}}$. It means that

$$T_{k,\lambda} = \left(\begin{array}{c} \bar{i}_{(p)}, \bar{\bar{v}}_{(q,\lambda)}, \bar{j}_{(p)}, \bar{\bar{t}}_{(q,k)} \\ \bar{i}_{(p)}, \bar{j}_{(p)}, \bar{\bar{t}}_{(q,k)}, \bar{\bar{v}}_{(q,\lambda)} \end{array} \right).$$

In the light of the performed transformations the system of equations (16) takes the form (11). This completes the proof of Theorem 2. □

5. Bayesian multidimensional-matrix quadratic empirical regression

Assumption of $m = 2$ in the expression (8) gives us the quadratic regression function:

$$\bar{y} = \tilde{C}_{(p,0q)} + {}^{0,p}(\tilde{C}_{(p,1q)}\bar{x}) + {}^{0,2p}(\tilde{C}_{(p,2q)}\bar{x}^2). \tag{19}$$

We obtain the Bayesian estimations of the parameters for the regression function (19).

The system of equations (11) for these parameters contains three equations:

$$\begin{aligned}
&{}^{0,(p+0q)}(R^{-1}_{a,(p,0q)} C_{(p,0q)}) + {}^{0,(p+0q)}(V^{T_{0,0}}_{0,0} C_{(p,0q)}) + {}^{0,(p+1q)}(V^{T_{1,0}}_{1,0} C_{(p,1q)}) + \\
&+ {}^{0,(p+2q)}(V^{T_{2,0}}_{2,0} C_{(p,2q)}) = \\
&= {}^{0,p}(R^{-1}_{\varepsilon} S_{yx^0}) + {}^{0,(p+0q)}(R^{-1}_{a,(p,0q)} C_{a,(p,0q)}), \\
&{}^{0,(p+1q)}(R^{-1}_{a,(p,1q)} C_{(p,1q)}) + {}^{0,(p+0q)}(V^{T_{0,1}}_{0,1} C_{(p,0q)}) + {}^{0,(p+1q)}(V^{T_{1,1}}_{1,1} C_{(p,1q)}) + \\
&+ {}^{0,(p+2q)}(V^{T_{2,1}}_{2,1} C_{(p,2q)}) = \\
&= {}^{0,p}(R^{-1}_{\varepsilon} S_{yx^1}) + {}^{0,(p+1q)}(R^{-1}_{a,(p,1q)} C_{a,(p,1q)}), \\
&{}^{0,(p+2q)}(R^{-1}_{a,(p,2q)} C_{(p,2q)}) + {}^{0,(p+0q)}(V^{T_{0,2}}_{0,2} C_{(p,0q)}) + {}^{0,(p+1q)}(V^{T_{1,2}}_{1,2} C_{(p,1q)}) + \\
&+ {}^{0,(p+2q)}(V^{T_{2,2}}_{2,2} C_{(p,2q)}) = \\
&= {}^{0,p}(R^{-1}_{\varepsilon} S_{yx^2}) + {}^{0,(p+2q)}(R^{-1}_{a,(p,2q)} C_{a,(p,2q)}).
\end{aligned}$$

Upon collecting similar terms we obtain the following system of equations:

$$\begin{aligned}
&{}^{0,p}((R^{-1}_{a,(p,0q)} + V^{T_{0,0}}_{0,0}) C_{(p,0q)}) + {}^{0,(p+q)}(V^{T_{1,0}}_{1,0} C_{(p,1q)}) + {}^{0,(p+2q)}(V^{T_{2,0}}_{2,0} C_{(p,2q)}) = \\
&= {}^{0,p}(R^{-1}_{\varepsilon} S_{yx^0}) + {}^{0,p}(R^{-1}_{a,(p,0q)} C_{a,(p,0q)}),
\end{aligned}$$

$$
{}^{0,p}(V_{0,1}^{T_{0,1}} C_{(p,0q)}) + {}^{0,(p+q)}((R_{a,(p,1q)}^{-1} + V_{1,1}^{T_{1,1}})C_{(p,1q)}) + {}^{0,(p+2q)}(V_{2,1}^{T_{2,1}} C_{(p,2q)}) =
$$

$$
= {}^{0,p}(R_\varepsilon^{-1} S_{yx^1}) + {}^{0,(p+q)}(R_{a,(p,1q)}^{-1} C_{a,(p,1q)}),
$$

$$
{}^{0,p}(V_{0,2}^{T_{0,2}} C_{(p,0q)}) + {}^{0,(p+q)}(V_{1,2}^{T_{1,2}} C_{(p,1q)}) + {}^{0,(p+2q)}((R_{a,(p,2q)}^{-1} + V_{2,2}^{T_{2,2}})C_{(p,2q)}) =
$$

$$
= {}^{0,p}(R_\varepsilon^{-1} S_{yx^2}) + {}^{0,(p+2q)}(R_{a,(p,2q)}^{-1} C_{a,(p,2q)}).
$$

With notations

$$
R_{(p,0q)} = (R_{a,(p,0q)}^{-1} + V_{0,0}^{T_{0,0}}),
$$

$$
R_{(p,1q)} = (R_{a,(p,1q)}^{-1} + V_{1,1}^{T_{1,1}}),
$$

$$
R_{(p,2q)} = (R_{a,(p,2q)}^{-1} + V_{2,2}^{T_{2,2}}),
$$

$$
B_{(p)} = {}^{0,p}(R_\varepsilon^{-1} S_{yx^0}) + {}^{0,p}(R_{a,(p,0q)}^{-1} C_{a,(p,0q)}),
$$

$$
B_{(p+q)} = {}^{0,p}(R_\varepsilon^{-1} S_{yx^1}) + {}^{0,(p+q)}(R_{a,(p,1q)}^{-1} C_{a,(p,1q)}),
$$

$$
B_{(p+2q)} = {}^{0,p}(R_\varepsilon^{-1} S_{yx^2}) + {}^{0,(p+2q)}(R_{a,(p,2q)}^{-1} C_{a,(p,2q)}),
$$

we rewrite this system in the form:

$$
\begin{cases}
{}^{0,p}(R_{(p,0q)} C_{(p,0q)}) + {}^{0,(p+q)}(V_{1,0}^{T_{1,0}} C_{(p,1q)}) + {}^{0,(p+2q)}(V_{2,0}^{T_{2,0}} C_{(p,2q)}) = B_{(p)}, \\
{}^{0,p}(V_{0,1}^{T_{0,1}} C_{(p,0q)}) + {}^{0,(p+q)}(R_{(p,1q)} C_{(p,1q)}) + {}^{0,(p+2q)}(V_{2,1}^{T_{2,1}} C_{(p,2q)}) = B_{(p+q)}, \\
{}^{0,p}(V_{0,2}^{T_{0,2}} C_{(p,0q)}) + {}^{0,(p+q)}(V_{1,2}^{T_{1,2}} C_{(p,1q)}) + {}^{0,(p+2q)}(R_{(p,2q)} C_{(p,2q)}) = B_{(p+2q)}.
\end{cases}
\tag{20}
$$

This system of equations can be solved by the elimination method.

Let us eliminate the variable $C_{(p,0q)}$ from the second equation of the system (20). Multiplying the first equation on the left by ${}^{0,p}R_{(p,0q)}^{-1}$ in the sense of $(0,p)$-folded product, and then by $V_{0,1}^{T_{0,1}}$ in the sense of $(0,p)$-folded product gives us the equation

$$
{}^{0,p}(V_{0,1}^{T_{0,1}} C_{(p,0q)}) + {}^{0,p}(V_{0,1}^{T_{0,1}\,0,p}({}^{0,p}R_{(p,0q)}^{-1}\,{}^{0,(p+q)}(V_{1,0}^{T_{1,0}} C_{(p,1q)})))+
$$

$$
+{}^{0,p}(V_{0,1}^{T_{0,1}\,0,p}({}^{0,p}R_{(p,0q)}^{-1}\,{}^{0,(p+2q)}(V_{2,0}^{T_{2,0}} C_{(p,2q)}))) = {}^{0,p}(V_{0,1}^{T_{0,1}\,0,p}({}^{0,p}R_{(p,0q)}^{-1} B_{(p)})),
$$

which is transformed to the equation

$$
{}^{0,p}(V_{0,1}^{T_{0,1}} C_{(p,0q)}) + {}^{0,(p+q)}({}^{0,p}({}^{0,p}(V_{0,1}^{T_{0,1}\,0,p}R_{(p,0q)}^{-1})V_{1,0}^{T_{1,0}})C_{(p,1q)})+
$$

$$
+{}^{0,(p+2q)}({}^{0,p}({}^{0,p}(V_{0,1}^{T_{0,1}\,0,p}R_{(p,0q)}^{-1})V_{2,0}^{T_{2,0}})C_{(p,2q)}) = {}^{0,p}(V_{0,1}^{T_{0,1}\,0,p}({}^{0,p}R_{(p,0q)}^{-1} B_{(p)})).
$$

By subtracting this equation from the second equation of the system (20), instead of the second equation of the system (20) we will have the following equation:

$$ {}^{0,(p+q)}(A_{2,2}C_{(p,1q)}) + {}^{0,(p+2q)}(A_{2,3}C_{(p,2q)}) = D_2, $$

where

$$ A_{2,2} = R_{(p,1q)} - {}^{0,p}({}^{0,p}(V_{0,1}^{T_{0,1}}\,{}^{0,p}R^{-1}_{(p,0q)})V_{1,0}^{T_{1,0}}), $$

$$ A_{2,3} = V_{2,1}^{T_{2,1}} - {}^{0,p}({}^{0,p}(V_{0,1}^{T_{0,1}}\,{}^{0,p}R^{-1}_{(p,0q)})V_{2,0}^{T_{2,0}}), $$

$$ D_2 = B_{(p+q)} - {}^{0,p}({}^{0,p}(V_{0,1}^{T_{0,1}}\,{}^{0,p}R^{-1}_{(p,0q)})B_{(p)}). $$

Now we shall eliminate the variable $C_{(p,0q)}$ from the third equation of the system (20). Multiplying the first equation on the left by ${}^{0,p}R^{-1}_{(p,0q)}$ in the sense of $(0,p)$-folded product, and then by $V_{0,2}^{T_{0,2}}$ in the sense of $(0,p)$-folded product gives us the equation

$$ {}^{0,p}(V_{0,2}^{T_{0,2}}C_{(p,0q)}) + {}^{0,p}(V_{0,2}^{T_{0,2}}\,{}^{0,p}({}^{0,p}R^{-1}_{(p,0q)}\,{}^{0,(p+q)}(V_{1,0}^{T_{1,0}}C_{(p,1q)})))+ $$

$$ +{}^{0,p}(V_{0,2}^{T_{0,2}}\,{}^{0,p}({}^{0,p}R^{-1}_{(p,0q)}\,{}^{0,(p+2q)}(V_{2,0}^{T_{2,0}}C_{(p,2q)}))) = {}^{0,p}(V_{0,2}^{T_{0,2}}\,{}^{0,p}({}^{0,p}R^{-1}_{(p,0q)}B_{(p)})), $$

which is transformed to the equation

$$ {}^{0,p}(V_{0,2}^{T_{0,2}}C_{(p,0q)}) + {}^{0,(p+q)}({}^{0,p}({}^{0,p}(V_{0,2}^{T_{0,2}}\,{}^{0,p}R^{-1}_{(p,0q)})V_{1,0}^{T_{1,0}})C_{(p,1q)})+ $$

$$ +{}^{0,(p+2q)}({}^{0,p}({}^{0,p}(V_{0,2}^{T_{0,2}}\,{}^{0,p}R^{-1}_{(p,0q)})V_{2,0}^{T_{2,0}})C_{(p,2q)}) = {}^{0,p}(V_{0,2}^{T_{0,2}}\,{}^{0,p}({}^{0,p}R^{-1}_{(p,0q)}B_{(p)})). $$

By subtracting this equation from the third equation of the system (20), instead of the third equation of the system (20) we will have the following equation:

$$ {}^{0,(p+q)}(A_{3,2}C_{(p,1q)}) + {}^{0,(p+2q)}(A_{3,3}C_{(p,2q)}) = D_3, $$

where

$$ A_{3,2} = V_{1,2}^{T_{1,2}} - {}^{0,p}({}^{0,p}(V_{0,2}^{T_{0,2}}\,{}^{0,p}R^{-1}_{(p,0q)})V_{1,0}^{T_{1,0}}), $$

$$ A_{3,3} = R_{(p,2q)} - {}^{0,p}({}^{0,p}(V_{0,2}^{T_{0,2}}\,{}^{0,p}R^{-1}_{(p,0q)})V_{2,0}^{T_{2,0}}), $$

$$ D_3 = B_{(p+2q)} - {}^{0,p}({}^{0,p}(V_{0,2}^{T_{0,2}}\,{}^{0,p}R^{-1}_{(p,0q)})B_{(p)}). $$

As a result, instead of (20), we obtain the following system:

$$\begin{cases} {}^{0,p}(R_{(p,0q)}C_{(p,0q)}) + {}^{0,(p+q)}(V_{1,0}^{T_{1,0}}C_{(p,1q)}) + {}^{0,(p+2q)}(V_{2,0}^{T_{2,0}}C_{(p,2q)}) = B_{(p)}, \\ {}^{0,(p+q)}(A_{2,2}C_{(p,1q)}) + {}^{0,(p+2q)}(A_{2,3}C_{(p,2q)}) = D_2, \\ {}^{0,(p+q)}(A_{3,2}C_{(p,1q)}) + {}^{0,(p+2q)}(A_{3,3}C_{(p,2q)}) = D_3. \end{cases} \tag{21}$$

Further, we will eliminate the variable $C_{(p,1q)}$ from the third equation of the system (21). We multiply the second equation on the left by ${}^{0,(p+q)}A_{2,2}^{-1}$ in the sense of $(0,p+q)$-folded product, and then by $A_{3,2}$ in the sense of $(0,p+q)$-folded product and obtain the equation

$$\begin{aligned} &{}^{0,(p+q)}(A_{3,2}C_{(p,1q)}) + {}^{0,(p+q)}(A_{3,2}{}^{0,(p+q)}({}^{0,(p+q)}A_{2,2}^{-1}{}^{0,(p+2q)}(A_{2,3}C_{(p,2q)}))) = \\ &= {}^{0,(p+q)}(A_{3,2}{}^{0,(p+q)}({}^{0,(p+q)}A_{2,2}^{-1}D_2)), \end{aligned}$$

which is transformed to the form

$$\begin{aligned} &{}^{0,(p+q)}(A_{3,2}C_{(p,1q)}) + {}^{0,(p+2q)}({}^{0,(p+q)}(A_{3,2}{}^{0,(p+q)}A_{2,2}^{-1})A_{2,3})C_{(p,2q)}) = \\ &= {}^{0,(p+q)}(A_{3,2}{}^{0,(p+q)}({}^{0,(p+q)}A_{2,2}^{-1}D_2)). \end{aligned}$$

By subtracting this equation from the third equation of the system (21), instead of third equation of the system (21) we will have the following equation:

$${}^{0,(p+2q)}(F_{3,3}C_{(p,2q)}) = G_3,$$

where

$$F_{3,3} = A_{3,3} - {}^{0,(p+q)}({}^{0,(p+q)}(A_{3,2}{}^{0,(p+q)}A_{2,2}^{-1})A_{2,3}),$$

$$G_3 = D_3 - {}^{0,(p+q)}({}^{0,(p+q)}(A_{3,2}{}^{0,(p+q)}A_{2,2}^{-1})D_2).$$

This completes the process of elimination. We have transformed our system of equations (20) to the following form

$$\begin{cases} {}^{0,p}(R_{(p,0q)}C_{(p,0q)}) + {}^{0,(p+q)}(V_{1,0}^{T_{1,0}}C_{(p,1q)}) + {}^{0,(p+2q)}(V_{2,0}^{T_{2,0}}C_{(p,2q)}) = B_{(p)}, \\ {}^{0,(p+q)}(A_{2,2}C_{(p,1q)}) + {}^{0,(p+2q)}(A_{2,3}C_{(p,2q)}) = D_2, \\ {}^{0,(p+2q)}(F_{3,3}C_{(p,2q)}) = G_3. \end{cases} \tag{22}$$

Now we find the solution of the system of equations. From the third equation of the system (22) we obtain the estimation $\hat{C}_{(p,2q)}$ of the parameter $C_{(p,2q)}$ of the quadratic regression function (19):

$$\hat{C}_{(p,2q)} = {}^{0,(p+2q)}({}^{0,(p+2q)}F^{-1}_{3,3}G_3). \tag{23}$$

Then we obtain the estimation $\hat{C}_{(p,1q)}$ of the parameter $C_{(p,1q)}$ from the second equation of the system (22):

$$\hat{C}_{(p,1q)} = {}^{0,(p+q)}({}^{0,(p+q)}A^{-1}_{2,2}(D_2 - {}^{0,(p+2q)}(A_{2,3}\hat{C}_{(p,2q)})). \tag{24}$$

Finally, we obtain the estimation $\hat{C}_{(p,0q)}$ of the parameter $C_{(p,0q)}$ from the first equation of the system (22):

$$\hat{C}_{(p,0q)} = {}^{0,p}({}^{0,p}R^{-1}_{(p,0q)}(B_{(p)} - {}^{0,(p+q)}(V^{T_{1,0}}_{1,0}\hat{C}_{(p,1q)}) - {}^{0,(p+2q)}(V^{T_{2,0}}_{2,0}\hat{C}_{(p,2q)}))). \tag{25}$$

It should be noted that the solution to the system of equations (20) always exists due to the positive definiteness of the matrices $R_{(p,0q)}$, $R_{(p,1q)}$, $R_{(p,2q)}$. This statement also applies to the system of equations (11). This is the advantage of the Bayesian approach, which provides the ability to use the samples of small sizes.

6. Numerical simulation of the Bayesian quadratic empirical regression

We perform the numerical simulation of the Bayesian quadratic empirical regression function (19) with two-dimensional input and output variables $\bar{x}$ and $\bar{y}$, i.e. we assume that in (19) $p = q = 2$. In this case the regression function (19) has the following form

$$\bar{y} = \phi(\bar{x}) = \tilde{C}_{(p,0q)} + {}^{0,2}(\tilde{C}_{(p,1q)}\bar{x}) + {}^{0,4}(\tilde{C}_{(p,2q)}\bar{x}^2), \tag{26}$$

where $\tilde{C}_{(p,0q)}$ is a two-dimensional matrix, $\tilde{C}_{(p,1q)}$ is a four-dimensional matrix and $C_{(p,2q)}$ is a six-dimensional matrix. We present the results of the simulation for the matrices $\bar{x}$, $\bar{y}$, $\tilde{C}_{(p,0q)}$, $\tilde{C}_{(p,1q)}$, $\tilde{C}_{(p,2q)}$ of order two. In this case the number of the parameters for estimation is $2^2 + 2^4 + 2^6 = 84$. The matrices $C_{(p,0q)}$, $\tilde{C}_{(p,1q)}$, $\tilde{C}_{(p,2q)}$ are supposed to be random independent Gaussian.

The following mean values $C_{a,(p,0q)}$, $C_{a,(p,1q)}$, $C_{a,(p,2q)}$ of the matrices $\tilde{C}_{(p,0q)}$, $\tilde{C}_{(p,1q)}$, $\tilde{C}_{(p,2q)}$ are used. The matrix $C_{a,(p,0q)}$ is $C_{a,(p,0q)} = \begin{pmatrix} 3 & 2 \\ 2 & 4 \end{pmatrix}$. The matrix $C_{a,(p,1q)} = (C_{a,(p,1q),i_1,i_2,i_3,i_4})$ has the element $C_{a,(p,1q),1,1,2,1} = 3$ and all other elements are equal to 1. The matrix $C_{a,(p,2q)} = (C_{a,(p,2q),i_1,i_2,i_3,i_4,i_5,i_6})$ has the elements

$C_{a,(p,2q),1,1,2,1,2,2} = 0.5$,
$C_{a,(p,2q),1,2,2,1,2,1} = 0$,
$C_{a,(p,2q),1,2,1,2,2,1} = 0$

and all other elements are equal to 0.1.

The real values of the parameters $\tilde{C}_{(p,0q)}$, $C_{(p,1q)}$, $C_{(p,2q)}$ were chosen as follows: $C_{(p,0q)} = \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}$. The matrix $C_{(p,1q)} = (C_{(p,1q),i_1,i_2,i_3,i_4})$ has the element $C_{(p,1q),1,1,2,1} = 5$ and all other elements are equal to 1. The matrix $C_{(p,2q)} = (C_{(p,2q),i_1,i_2,i_3,i_4,i_5,i_6})$ has the elements
$C_{a,(p,2q),1,1,2,1,2,2} = 0$,
$C_{a,(p,2q),1,1,2,2,2,1} = 0$,
$C_{a,(p,2q),1,2,2,1,2,1} = 2$,
$C_{a,(p,2q),1,2,1,2,2,1} = 3$,
$C_{a,(p,2q),1,2,2,1,1,2} = 3$

and all other elements are equal to 0.1.

The elements of the matrices $\bar{C}_{(p,0q)}$, $\tilde{C}_{(p,1q)}$, $\tilde{C}_{(p,2q)}$ are supposed to be independent. The variances of the elements of the matrix $\tilde{C}_{(p,0q)}$ are equal to 4, the variances of the elements of the matrix $\tilde{C}_{(p,1q)}$ are equal to 5 and the variances of the elements of the matrix $\tilde{C}_{(p,2q)}$ are equal to 3.

The measurements were simulated according to the formulas (26) and (1). The elements of the two-dimensional matrix of the errors of the measurements $\bar{\varepsilon}$ are supposed to be independent with variances equal to 4.

Bayesian estimations $\hat{C}_{(p,0q)}$, $\hat{C}_{(p,1q)}$, $\hat{C}_{(p,2q)}$ were calculated by the formulas (23), (24), (25), and the empirical Bayesian regression – by the formula

$$\hat{\bar{y}} = \hat{C}_{(p,0q)} + {}^{0,2}(\hat{C}_{(p,1q)}\bar{x}) + {}^{0,4}(\hat{C}_{(p,2q)}\bar{x}^2). \tag{27}$$

The Bayesian empirical regression function $\hat{y}_{1,1} = \hat{Y}_{1,1}(x_{1,1}, x_{1,2})$, (27), when $x_{2,1} = x_{2,2} = -5$, constructed on the four measurements at the points

$$x_1 = \begin{pmatrix} -10 & -10 \\ -5 & -5 \end{pmatrix}, \quad x_2 = \begin{pmatrix} 10 & -10 \\ -5 & -5 \end{pmatrix},$$
$$x_3 = \begin{pmatrix} -10 & 10 \\ -5 & -5 \end{pmatrix}, \quad x_4 = \begin{pmatrix} 10 & 10 \\ -5 & -5 \end{pmatrix}$$

is shown in Fig. 1 (top surface). The measurements are marked as asterisks in the figure. The bottom surface in the figure is a general (true) regression function $y_{1,1} = y_{1,1}(x_{1,1}, x_{1,2})$, (26), when $x_{2,1} = x_{2,2} = -5$. This figure confirms the correctness of the theoretical results and illustrates the capacity of the Bayesian approach to make use of the samples of small sizes.

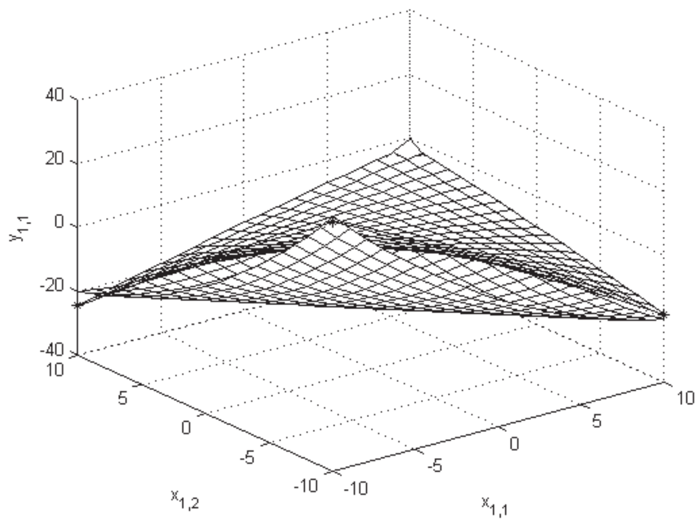

Figure 1. Bayesian empirical quadratic regression function (27) built on four measurements (the top surface) compared with the true regression function (the bottom surface)

7. Conclusion

In conclusion, we outline the main results of this work and note their particularities.

1. The problem of constructing the maximum likelihood multidimensional-matrix polynomial regression was formulated and solved. This regression has the following particularities, when compared the existing regressions: 1) a more general multidimensional-matrix polynomial regression function, when input and output variables are the multidimensional matrices, is considered; 2) a new non-traditional multidimensional-matrix form of the representation of the regression function in the manner of multidimensional-matrix polynomial is used. The general solution of the problem is a system of linear multidimensional-matrix equations relative to the multidimensional-matrix parameters of the regression function (Theorem 1).

2. The problem of constructing the Bayesian multidimensional-matrix polynomial regression with the same particularities as in the point 1 above was formulated and solved. Besides, the prior distributions of the multidimensional-matrix parameters of the regression function are supposed to be Gaussian. The general solution of this problem is the system of linear multidimensional-matrix equations relative the multidimensional-matrix parameters of the regression function (Theorem 2).

3. On the basis of the general solution, the algorithm of calculation of the parameters of the Bayesian multidimensional-matrix quadratic empirical regression functions was obtained.

4. Simulation of the quadratic Bayesian empirical regressions function with two-dimensional input and output variables was performed. The simulation confirmed the correctness of the theoretical results and illustrated the important benefits of the Bayesian approach in terms of having the algorithmic generality and obtaining the estimations for the cases with the small number of measurements.

Some aspects of the work, reported in this article were presented in a short form at the international conference, Mukha (2020).

8. Appendix

For more details on the subjects here considered, see Sokolov (1972), Mukha (2004).

The definition of a multidimensional matrix. A multidimensional (p-dimensional) matrix is a system of numbers or variables $a_{i_1,i_2,...,i_p}$, $i_\alpha = 1,2,...,n_\alpha$, $\alpha = 1,2,...,p$, located at the points of the p-dimensional space defined by the coordinates $i_1, i_2, ..., i_p$.

The p-dimensional matrix is denoted as

$$A = (a_{i_1,i_2,...,i_p}), \quad i_\alpha = 1,2,...,n_\alpha, \quad \alpha = 1,2,...,p, \tag{A.1}$$
or
$$A = (a_i),$$

where $i = (i_1, i_2, ..., i_p)$ is a multi-index, $i_\alpha = 1,2,...,n_\alpha$, $\alpha = 1,2,...,p$.

If $n_1 = n_2 = ...n_p = n$, then the matrix (A.1) is called p-dimensional matrix of the order n (a hyper-square matrix). In this connection, the matrix (A.1) with different $n_1, n_2, ...n_p$ could be called a hyper-rectangular matrix.

Thus, a zero-dimensional matrix is a scalar, a one-dimensional matrix is a vector and a two-dimensional matrix is an ordinary matrix in traditional notation.

Matrices associated with multidimensional matrices. Let the p-dimensional matrix $A=(a_{i_1,i_2,...,i_p})$ of the order n be represented in the form of $A=(a_{l,s,c})$, where $l=(l_1,l_2,...,\ l_\kappa)$, $s=(s_1,s_2,...,s_\lambda)$, $c=(c_1,...,c_\mu)$ are multi-indexes, $\kappa+\lambda+\mu=p$. In this case we say that the matrix A has the (κ,λ,μ)-structure and is denoted $A_{(\kappa,\lambda,\mu)}$. The multi-indexes l,s,c of this matrix have n^κ, n^λ and n^μ values, respectively. Let us arrange the values of l,s,c in some way:

$$\tilde{l}=l^{(1)},l^{(2)},...,l^{(n^\kappa)},$$

$$\tilde{s}=s^{(1)},s^{(2)},...,s^{(n^\lambda)},$$

$$\tilde{c}=c^{(1)},c^{(2)},...,c^{(n^\mu)}.$$

The cell-diagonal matrix

$$\tilde{A}_{(\kappa,\lambda,\mu)}=diag\left\{\ A^{(1)}_{(\kappa,0,\mu)},A^{(2)}_{(\kappa,0,\mu)},...,A^{(n^\lambda)}_{(\kappa,0,\mu)}\right\},$$

consisting of the elements of the matrix A, where the diagonal cells $\mathrm{A}^{(h)}_{(\kappa,0,\mu)}$, $h=1,2,...,n^\lambda$, are two-dimensional $(n^\kappa\times n^\mu)$-matrices

$$A^{(h)}_{(\kappa,0,\mu)}=(a_{\tilde{l},s^{(h)},\tilde{c}}),\quad \tilde{l}=l^{(1)},l^{(2)},...l^{(n^\kappa)},\ \ \tilde{c}=c^{(1)},c^{(2)},...,c^{(n^\mu)},$$

is called (κ,λ,μ)-associated matrix with the matrix $A_{(\kappa,\lambda,\mu)}$.

The associated matrix $\tilde{A}_{(\kappa,\lambda,\mu)}$ represents fully the initial multidimensional matrix $A_{(\kappa,\lambda,\mu)}$, because it contains all of its elements.

Addition of multidimensional matrices. If $A=(a_{i_1,i_2,...,i_p})$, $B=(b_{i_1,i_2,...,i_p})$, then $C=A+B=(c_{i_1,i_2,...,i_p})$, where $c_{i_1,i_2,...,i_p}=a_{i_1,i_2,...,i_p}+b_{i_1,i_2,...,i_p}$, $i_\alpha=1,2,...,n_\alpha$, $\alpha=1,2,...,p$.

Multiplication of multidimensional matrix by a scalar. If t is some scalar number or a variable and A is a p-dimensional matrix, then $C=tA=(c_{i_1,i_2,...,i_p})$, where $c_{i_1,i_2,...,i_p}=ta_{i_1,i_2,...,i_p}$, $i_\alpha=1,2,...,n_\alpha$, $\alpha=1,2,...,p$.

Multiplication of two multidimensional matrices. If a p-dimensional matrix A is represented in the form of $A=(a_{i_1,i_2,...,i_p})=(a_{l,s,c})$, where $l=(l_1,l_2,...,\ l_\kappa)$, $s=(s_1,s_2,...,s_\lambda)$, $c=(c_1,...,c_\mu)$ are multi-indices, $\kappa+\lambda+\mu=p$, and a q-dimensional matrix B is represented in the form of $B=(b_{j_1,j_2,...,j_q})=(b_{c,s,m})$, where $m=(m_1,...,m_\nu)$ is a multi-index, $\lambda+\mu+\nu=q$, then the matrix

$D = (d_{l,s,m})$ is called a (λ,μ)-folded product of the matrices A and B, if its elements are defined by the expression

$$d_{l,s,m} = \sum_{c} a_{l,s,c} b_{c,s,m} = \sum_{c_1}\sum_{c_2}\cdots\sum_{c_\mu} a_{l,s,c} b_{c,s,m}.$$

The (λ,μ)-folded product of the matrices A and B is denoted ${}^{\lambda,\mu}(AB)$. Thus,

$$D = {}^{\lambda,\mu}(AB) = (\sum_{c} a_{l,s,c} b_{c,s,m}) = (d_{l,s,m}). \tag{A.2}$$

In the case of the $(0,0)$-folded product we often omit the left upper indices and write AB instead of ${}^{0,0}(AB)$.

In the general case ${}^{\lambda,\mu}(AB) \neq {}^{\lambda,\mu}(BA)$.

The associative law of multiplication of the multidimensional matrices is fulfilled. If $\lambda' + \mu' \leq \nu$, then

$${}^{\lambda',\mu'}({}^{\lambda,\mu}(AB)\,C) = {}^{\lambda,\mu}(A^{\lambda',\mu'}(BC)).$$

The distributive law of multiplication of the multidimensional matrices takes the form as follows:

$${}^{\lambda,\mu}(A(B+C)) = {}^{\lambda,\mu}(AB) + {}^{\lambda,\mu}(A\,C).$$

The degree of multidimensional matrix. The matrix

$$D = {}^{\lambda,\mu}(AA) = {}^{\lambda,\mu}A^2$$

is called a (λ,μ)-folded square of the matrix A, and the matrix

$$D = {}^{\lambda,\mu}(A^{\lambda,\mu}(A\cdots{}^{\lambda,\mu}(AA))) = {}^{\lambda,\mu}A^k$$

is called a (λ,μ)-folded k-th degree of the matrix A. If it is $(0,0)$-folded k-th degree of the matrix A, then we omit the left upper indices and write A^k instead of ${}^{0,0}A^k$.

A multidimensional identity matrix. The matrix $E(\lambda,\mu)$ is called a (λ,μ)-identity matrix if for any multidimensional matrix A the equality

$${}^{\lambda,\mu}(AE(\lambda,\mu)) = {}^{\lambda,\mu}(E(\lambda,\mu)A) = A$$

is fulfilled. $E(\lambda,\mu)$ is a $(\lambda+2\mu)$-dimensional matrix, whose elements are defined by the formula

$$E(\lambda,\mu) = (e_{c,s,m}) = \left(\begin{cases} 1, & if\ \ c = m, \\ 0, & if\ \ c \neq m \end{cases} \right),$$

$$c = (c_1, ..., c_\mu), \quad s = (s_1, s_2, ..., s_\lambda), \quad m = (m_1, ..., m_\mu).$$

A multidimensional inverse matrix. The matrix $A^{-1}(\lambda, \mu)$ (or ${}^{\lambda,\mu}A^{-1}$) is called the (λ, μ)-inverse to the matrix A, if the equalities

$$^{\lambda,\mu}(AA^{-1}(\lambda, \mu)) = {}^{\lambda,\mu}(A^{-1}(\lambda, \mu)A) = E(\lambda, \mu) \tag{A.3}$$

are satisfied.

The transpose of a multidimensional matrix. The matrix $A^T = (a^T_{i_1,i_2,...,i_p})$ is composed of the elements $a^T_{i_1,i_2,...,i_p}$, which are connected with the elements $a_{i_1,i_2,...,i_p}$ of the matrix $A = (a_{i_1,i_2,...,i_p})$ by the equalities

$$a^T_{i_1,i_2,...,i_p} = a_{i_{\alpha_1},i_{\alpha_2},...,i_{\alpha_p}}, \tag{A.4}$$

where $i_{\alpha_1}, i_{\alpha_2}, ..., i_{\alpha_p}$ is some permutation of the indices $i_1, i_2, ..., i_p$ and is called transposed according to the substitution

$$T = \begin{pmatrix} i_1, ..., & i_p \\ i_{\alpha_1}, ..., & i_{\alpha_p} \end{pmatrix}$$

of matrix A.

In Mukha (2004) some standard substitutions are introduced that allow us to form various substitutions: of the types ‘onward’, ‘back’, and ‘onward-back’.

The substitution on the p indices, the lower string of which is formed from the upper string by the transfer of the r left indices to the right (onward) is called substitution of the type ‘onward’:

$$B_{p,r} = \begin{pmatrix} i_1, & i_2, & ..., & i_{p-r}, & i_{p-r+1}, & ..., & i_p \\ i_{r+1}, & i_{r+2}, & ..., & i_p, & i_1, & ..., & i_r \end{pmatrix}, \quad p \geq r. \tag{A.5}$$

The substitution on the p indices, the lower string of which is formed from the upper string by the transfer of the r right indices to the left (back) is called substitution of the type ‘back’:

$$H_{p,r} = \begin{pmatrix} i_1, & i_2, & ..., & i_r, & i_{r+1}, & ..., & i_p \\ i_{p-r+1}, & i_{p-r+2}, & ..., & i_p, & i_1, & ..., & i_{p-r} \end{pmatrix}, \quad p \geq r. \tag{A.6}$$

The substitution on the p indices, the lower string of which is formed from the upper string by the transfer of the r left indices to the right (onward) and the

s right indices to the left (back) is called substitution of the type 'onward-back':

$$B_r H_s = \begin{pmatrix} i_1, & ..., & i_r, & ..., & i_{p-s+1}, & ..., & i_p \\ i_{p-s+1}, & ..., & i_p, & ..., & i_1, & ..., & i_r \end{pmatrix},$$

$$p \geq r + s.$$

The Matlab's function ipermute.m performs a transpose of a multidimensional array in accordance with the definition (A.4).

A multidimensional-matrix derivative (Sokolov, 1960). Let $Y = (y_m)$, $m = (m_1, m_2, ..., m_p)$, be a p-dimensional matrix depending on a q-dimensional matrix $X = (x_k)$, $k = (k_1, k_2, ..., k_q)$. The derivative of the matrix Y with respect to the matrix X is a $(p+q)$-dimensional matrix Z, defined by the expression:

$$Z(X) = (z_{m,k}(X)) = \frac{dY(X)}{dX} = Y'(X) = \left(\frac{\partial y_m}{\partial x_k}\right).$$

The derivatives of the higher orders are defined by the sequential differentiation:

$$\frac{d^n Y(X)}{dX^n} = \frac{d}{dX}\left(\frac{d^{n-1}Y(X)}{dX^{n-1}}\right),$$

or $Y^{(n)}(X) = (Y^{(n-1)}(X))'$.

The derivative of the (λ, μ)*-folded product of the matrices* (Sokolov, 1960). If $F(X)$ and $\Phi(X)$ are p- and r-dimensional matrices, respectively, depending on a q-dimensional matrix X, then the derivative of the (λ, μ)-folded product of these matrices with respect to the matrix X is defined by the following expression:

$$\frac{d}{dX}{}^{\lambda,\mu}(F\Phi) = {}^{\lambda,\mu}\left(\left(\frac{dF}{dX}\right)^{B_{p+q,q}}\Phi\right)^{H_{p+q+r-\lambda-2\mu,q}} + {}^{\lambda,\mu}\left(F\frac{d\Phi}{dX}\right),$$

where $B_{p+q,q}$ and $H_{p+q+r-\lambda-2\mu,q}$ are substitutions of the types 'onward' and 'back', respectively.

The derivative of the composition of two multidimensional-matrix functions (Mukha, 2004). Let

$$F(Y) = (f_s(Y)), \;\; s = (s_1, s_2, ..., s_r),$$

be an r-dimensional matrix depending on the p-dimensional matrix

$$Y(X) = (y_m(X)), \;\; m = (m_1, m_2, ..., m_p),$$

while $Y(X)$ depends on the q-dimensional matrix

$$X = (x_k), \;\; k = (k_1, k_2, ..., k_q),$$

with different elements, then the derivative dF/dX is defined as the following $(0,p)$-folded product:

$$\frac{dF(Y)}{dX} = {}^{0,p}\left(\frac{dF}{dY}\frac{dY}{dX}\right).$$

If the function F is explicitly dependent on the matrix X, i.e. if $F = F(X, Y)$, then

$$\frac{dF(X,Y)}{dX} = {}^{0,p}\left(\frac{\partial F}{\partial Y}\frac{dY}{dX}\right) + \frac{\partial F}{\partial X}.$$

The derivative of the implicitly defined multidimensional-matrix function (Mukha, 2004). Let

$$Y(X) = (y_l), \;\; l = (l_1, l_2, ..., l_p),$$

be a p-dimensional matrix function,

$$X = (x_s), \;\; s = (s_1, s_2, ..., s_q),$$

be a q-dimensional matrix and function $Y(X)$ be defined implicitly by the equality

$$F(X, Y) = 0,$$

where F is an r-dimensional matrix function,

$$F = (f_k), \;\; k = (k_1, k_2, ..., k_r)\,,$$

then the derivative dY/dX is the $(p+q)$-dimensional matrix that is defined as the solution of the following multidimensional-matrix equation:

$${}^{0,p}\left(\frac{\partial F}{\partial Y}\frac{dY}{dX}\right) + \frac{\partial F}{\partial X} = 0.$$

The derivatives of some multidimensional-matrix functions (Mukha, 2004). If X is a q-dimensional matrix and A is an r-dimensional constant matrix, then

$$\frac{dA}{dX} = 0, \quad \frac{dX}{dX} = E(0, q), \quad \frac{d}{dX}{}^{0,q}(AX) = A,$$

where 0 is an $(r+q)$-dimensional zero matrix and $E(0, q)$ is a $(0, q)$-identity matrix ($2q$-dimensional matrix).

If X is a q-dimensional matrix, $B = (b_{l,c_1,c_2,...,c_m})$ is a $(p+mq)$-dimensional constant matrix, such that each multi-index $c_1, c_2, ..., c_m$ of it contains q indices and B is symmetric relative to these multi-indices, then

$$\frac{d}{dX}^{0,mq}(AX^m) = m\,^{0,(m-1)q}(AX^{m-1}).$$

The multidimensional-matrix Gaussian distribution (Mukha, 2004). The random p-dimensional matrix ξ of the order n with probability density of the form

$$f_\xi(\bar{x}) = \frac{1}{\sqrt{(2\pi)^{n^p}|d_\xi|}} \exp(-\tfrac{1}{2}\,^{0,2p}(d_\xi^{-1}(\bar{x} - a_\xi)^2)) \tag{A.7}$$

is called Gaussian or normal random matrix. In this definition, a_ξ is the mean value of the ξ, d_ξ is the variance matrix of the ξ, d_ξ^{-1} is the matrix $(0,p)$-inverse to the d_ξ, and $|d_\xi|$ is the determinant of the matrix d_ξ (the determinant of the matrix $(0,p)$-associated with the matrix d_ξ).

References

Cayley A. (1842/9) On the notations and properties of certain functions resolvable into a series of determinants. *Trans. Cambridge Philos. Soc.* 8, 85–88.

Ermakov, S. M. and Zhiglyavskii, A. A. (1987) *Mathematical theory of optimal experiment.* Moscow, Publishing house “Science”. In Russian.

Fahrmeir, L. and Kneib, Th. (2011) *Bayesian Smoothing and Regression for Longitudinal, Spatial and Event History Data.* Oxford Statistical Science Series, Oxford University Press.

Feldbaum, A. A. (1966) *The foundations of the theory of optimal automatic systems.* Moscow, Publishing house “Science”. In Russian.

Klepikov, N. P. and Sokolov, S. N. (1961) *Analysis and Planning of Experiments by the Method of Maximum Likelihood.* Berlin, Akademie-Verlag.

Mukha, V. S. and Sergeev, E. V. (1974) Bayesian estimations of the parameters of the regression objects. *Proceedings of LETI*, 149, 26–29. In Russian.

Mukha, V. S. and Sergeev E. V. (1976) Dual control of regression objects. *Proceedings of LETI*, 202, 58–65. In Russian.

Mukha, V. S. (2004) *Analysis of multidimensional data.* Minsk, Technoprint. In Russian.

Mukha, V. S. (2005) Vector Multiconnected Markov Chains. *Journal of Automation and Information Sciences*, **37**, 63–69.

Mukha, V. S. (2006) Multidimensional Matrix Approach in Parallel Factor Analysis. *Journal of Automation and Information Sciences*, **38**, 21–29.

Mukha, V. S. (2007a) The Best Polynomial Multidimensional-matrix Regression. *Cybernetics and Systems Analysis*, **43**, 3, 427–432.

Mukha, V. S. (2007b) Multidimensional-matrix polynomial regression analysis. Parameters estimations. *Proceedings of the National Academy of Sciences of Belarus. Physics and Mathematics series*, 1, 45–51. In Russian.

Mukha, V. S. (2008) PARAFAC-regression Analysis for Determining the Concentration of Substances in Solutions. *Journal of Automation and Information Sciences*, **40**, 25–40.

Mukha, V. S. (2011) Minimum average risk and efficiency of optimal polynomial multidimensional-matrix predictors. *Cybernetics and Systems Analysis*, **47**, 2, 277–285.

Mukha, V. S. (2012) Educational Timetable Problem: Formulation and Solving. *Journal of Automation and Information Sciences*, **44**, 12, 55–67.

Mukha, V.S. (2017) On Statistical Forecasting of the Atmospheric Temperature for Month. *American Journal of Environmental Engineering and Science*, **4**, 6, 71–77.

Mukha, V. S. (2020) On the Bayesian multidimensional-matrix polynomial empirical regression / Vladimir S. Mukha, Irina G. Malikova // Scientific research of the SCO countries: synergy and integration: Materials of the International Conference, Beijing, May 14, 2020. China, 2020, 159–165. DOI: https://doi.org/10.34660/INF.2020.28.63870.

Scott, R.F. (1879/80) On cubic determinants and other determinants of higher class and on determinants of alternate numbers. *Proc. London Math. Soc. (1)*, 11, 17–29.

Sokolov, N. P. (1960) *Spatial matrices and their applications.* Moscow, Fizmatgiz. In Russian.

Sokolov, N. P. (1972) *Introduction to the theory of multidimensional matrices.* Kiev, Naukova dumka. In Russian.

Wakefield, J. (2013) *Bayesian and Frequentist Regression Methods.* Springer.

Zellner, A. (1971) *An Introduction to Bayesian Inference in Econometrics.* John Wiley and Sons, Inc.