Modelling in the Class of Regression Equations Systems in Conditions of Structural Uncertainty

Alexander Sarychev

Institute of Technical Mechanics of the National Academy of Sciences of Ukraine,
15 Leshko-Popel St., Dnipropetrovsk, 49600, Ukraine.

Sarychev@prognoz.dp.ua

Abstract. *The problem of search of optimum complexity system of regression equations by principles of the Group Method of Data Handling surveyed. The criterion of quality of a system of regression equations which is system analogue of criterion of the regularity is offered . The criterion is researched in the scheme of repeated observations.*

Keywords

Group Method of Data Handling, criterion of the regularity for regression equations system, repeated observations scheme.

1 Introduction

The problem of construction of a regression equations system in conditions of structural uncertainty by quantity and structure of regressors in regression equations is one of objects of researches in a Group Method of Data Handling (GMDH) [1, 2, 3].

For the solution of a problem of statistical modeling in the given class of models in conditions of structural uncertainty it is necessary: 1) to specify a method of estimation of coefficients in models; 2) to construct algorithm of generation of models structures; 3) to accept criterion of quality for models comparison.

As is known, in a GMDH the comparison of quality of models is based on partition of initial sample of the data on learning and verifying parts: on learning sample the coefficients of model estimate, and on verifying the quality of model estimates.

Pursuant to principles of modeling in GMDH for the substantiation of adequacy of criterion of quality of models it is necessary: 1) to calculate mathematical expectation of researched criterion for given structure of model; 2) to research behavior of mathematical expectation of this criterion depending on structure of models; 3) to prove existence of model of optimum complexity; 4) to receive a condition of a reduction (simplification) of model of optimum complexity.

In the given article the criterion of quality of a regression equations system being system analogue of criterion of regularity GMDH is offered.

In [4] the problem of estimation of coefficients in a regression equations system in the supposition surveyed, that the errors of observations of output variable simulated object statistically are dependent, the output variables can be determined, generally speaking, by different sets of regressors, and the dispersion matrix of errors of observations of output variables is unknown. In these conditions for determination of coefficients the iterative scheme of estimation is constructed, which efficiency is confirmed by a method of statistical tests. In [5] this problem surveyed at the additional supposition that the dispersion matrix of errors of observations of output variables is known to within a scalar multiplier. This supposition has allowed to receive the analytical formulas for estimations of coefficients in a regression of equations system and analytically to research their properties.

Based on results [4, 5], it is possible to state a problem of construction of system criterion of regularity for search of optimum set of regressors in a regression models system.

2 Statement of the Problem

Let researched static object be described by set m of input variables $\overset{o}{X} = \{\overset{o}{x}_1, \overset{o}{x}_2, ..., \overset{o}{x}_m\}$ and set h output variables $Y = \{y(1), y(2), ..., y(h)\}$.

Let the model of investigated object have the form

$$y(k) = \overset{o}{y}(k) + \xi(k) = \sum_{j=1}^{m(k)} \overset{o}{\theta}_j(k) \overset{o}{x}_j(k) + \xi(k), \quad k = 1, 2, ..., h, \tag{1}$$

where k is a number of an output variable; h is a volume of output variables; $y(k)$ is a measured with an error k-th output variable; $\overset{o}{y}(k)$ is a not noisy (unobservable) k-th output variable; $\xi(k)$ is random unobservable measurement error k-th output variable; $\overset{o}{x}_j(k)$ is an j-th input variable from set of input variables $\overset{o}{X}(k) \neq \varnothing$ ($\varnothing$ – empty set), participating in formation a k-th output variable; $m(k)$ is a volume input variable, inhering to set $\overset{o}{X}(k)$; $\overset{o}{\boldsymbol{\theta}}(k) = (\overset{o}{\theta}_1(k), \overset{o}{\theta}_2(k), ..., \overset{o}{\theta}_{m(k)}(k))^{\mathrm{T}}$ is a vector of unknowns not equal to zero of coefficients.

Let as a result of observation of object for everyone output variable $y(k)$, $k = 1, 2, ..., h$, for two sample of observations – learning (A) and verifying (B) are obtained: 1) $\overset{o}{\mathbf{X}}(D,k)$ is a $(n(D) \times m(k))$-matrix $n(D)$ of observations $m(k)$ of inputs of set $\overset{o}{X}(k)$, having a full rank equal $m(k)$; 2) $\mathbf{y}(D,k)$ is $(n(D) \times 1)$-vector of the conforming observations of an output variable $y(D,k)$, where $D = A, B$. Pursuant to model (1)

$$\mathbf{y}(D,k) = \overset{o}{\mathbf{y}}(D,k) + \boldsymbol{\xi}(D,k) = \overset{o}{\mathbf{X}}(D,k)\overset{o}{\boldsymbol{\theta}}(k) + \boldsymbol{\xi}(D,k), \; k = 1, 2, ..., h, \tag{2}$$

where $\overset{o}{\mathbf{y}}(D,k)$ is $(n(D) \times 1)$-vector of values no noisy (unobservable) by a k-th output variable; $\boldsymbol{\xi}(D,k)$ is a $(n(D) \times 1)$-vector of random unobservable measurement errors by a k-th output variable.

Let the vector random variable $\xi(D) = (\xi(D,1), \xi(D,2), ..., \xi(D,h))^{\mathrm{T}}$ is distributed under the h-dimensional normal law: $\xi(D) \sim N(\mathbf{0}_h, \boldsymbol{\Sigma})$, and concerning $(n(D) \times 1)$- vectors $\boldsymbol{\xi}(D,k)$, $k = 1, 2, ..., h$, assumptions are executed

$$E\{\boldsymbol{\xi}(D,k)\} = \mathbf{0}_{n(D)}, \quad E\{\boldsymbol{\xi}(D,k)\boldsymbol{\xi}^{\mathrm{T}}(D,k)\} = \sigma_{kk}\, \mathbf{I}_{n(D)}, \quad k = 1, 2, ..., h; \tag{3}$$

$$E\{\boldsymbol{\xi}(D,k)\boldsymbol{\xi}^{\mathrm{T}}(D,q)\} = \sigma_{kq}\, \mathbf{I}_{n(D)}, \quad k, q = 1, 2, ..., h, \quad k \neq q; \tag{4}$$

$$E\{\xi_{i_1}(D,k)\xi_{i_2}(D,k)\} = 0, \; i_1, i_2 = 1, 2, ..., n(D), \; E\{\boldsymbol{\xi}(A,k)\boldsymbol{\xi}^{\mathrm{T}}(B,q)\} = \mathbf{O}_{(n(A) \times n(B))}, \tag{5}$$

where $E\{\cdot\}$ is a symbol of a mathematical expectation on all probable realizations of random vectors $\boldsymbol{\xi}(D,k)$ and $\boldsymbol{\xi}(D,q)$; $\mathbf{0}_h$ is $(h \times 1)$-vector consisting of zero; $\boldsymbol{\Sigma}$ is a unknown covariance $(h \times h)$-matrix; σ_{kk} – unknown final size, a dispersion of a random variable $\xi(D,k)$; σ_{kq} – unknown finit covariance of random variables $\xi(D,k)$ and $\xi(D,q)$; $\mathbf{I}_{n(D)}$ is the $(n(D) \times n(D))$ unit matrix; $\mathbf{O}_{(n(A) \times n(B))}$ is a zero $(n(A) \times n(B))$ matrix.

Let's write down (2)–(5) in the generalized kind. For this purpose we shall enter designations

$$\mathbf{Y}(D) = [\mathbf{y}(D,1), \mathbf{y}(D,2), ..., \mathbf{y}(D,h)], \tag{6}$$

$$\overset{o}{\mathbf{Y}}(D) = [\overset{o}{\mathbf{y}}(D,1), \overset{o}{\mathbf{y}}(D,2), ..., \overset{o}{\mathbf{y}}(D,h)], \tag{7}$$

$$\boldsymbol{\Xi}(D) = [\boldsymbol{\xi}(D,1), \boldsymbol{\xi}(D,2), ..., \boldsymbol{\xi}(D,h)]. \tag{8}$$

Then (2) it is possible to write down as

$$\mathbf{Y}(D)=\overset{o}{\mathbf{Y}}(D)+\mathbf{\Xi}(D), \tag{9}$$

and assumptions (3)–(5) –

$$E\{\mathbf{\Xi}(D)\}=\mathbf{O}_{(n(D)\times h)}, \tag{10}$$

$$E\{\mathbf{\Xi}^{\mathrm{T}}(D)\mathbf{\Xi}(D)\}=n(D)\mathbf{\Sigma}=n(D)\begin{bmatrix}\sigma_{11} & \sigma_{12} & \cdots & \sigma_{1h}\\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2h}\\ \vdots & \vdots & \ddots & \vdots\\ \sigma_{h1} & \sigma_{h2} & \cdots & \sigma_{hh}\end{bmatrix}, \tag{11}$$

where $\mathbf{O}_{(n(D)\times h)}$ is a zero $(n(D)\times h)$-matrix.

Let's consider, that the covariance matrix $\mathbf{\Sigma}$ looks like

$$\mathbf{\Sigma}=\rho\overset{o}{\mathbf{\Sigma}}, \tag{12}$$

where the matrix $\overset{o}{\mathbf{\Sigma}}$ is supposed set, and ρ is an unknown scalar multiplier.

Let's formulate **statement of a problem:** it is required to construct criterion of regularity MGDH which would allow to find optimum set of regressors for the regression equations system in conditions (1)–(12).

Let's assume, that the algorithm of full generation of analyzed and selected sets of regressors is applied.

2 Estimation of Coefficients in *SA*-optimal Regression Equations Systems

The decision of a task in view assumes estimation coefficients of regression equations system on training sample A and estimation of the received the regression equations system on verifying sample B.

For estimation coefficients on training sample A we shall take advantage of results of work [5] where SA-optimum estimations of coefficients of regression equations system in assumptions (1)–(12) are received as a result of minimization entered *SA*-functional.

Let $\hat{\mathbf{d}}(A,k)=(\hat{d}_1(A),\hat{d}_2(A),...,\hat{d}_{m(k)}(A)),\ k=1,2,...,h,$ – some estimations of unknown coefficients $\overset{o}{\boldsymbol{\theta}}(k),\ k=1,2,...,h$, received on sample A. For outputs corresponding regression models it is carried out

$$\mathbf{y}(A,k)=\hat{\mathbf{y}}(A,k)+\mathbf{u}(A,k)=\overset{o}{\mathbf{X}}(A,k)\hat{\mathbf{d}}(A,k)+\mathbf{u}(A,k),\ k=1,2,...,h,$$

where $\mathbf{u}(A,k)$ is a so-called $(n\times 1)$ - vector of the rests of k-th model.

Let's enter a designation $\mathbf{U}(A)=[\mathbf{u}(A,1),\mathbf{u}(A,2),...,\mathbf{u}(A,h)]$.

Definition 1. *SA*-optimum regression equations system is defined as system for coefficients which $\mathbf{d}_{SA}(A,1),\mathbf{d}_{SA}(A,2),...,\mathbf{d}_{SA}(A,h)$ is carried out

$$(\mathbf{d}_{SA}(A,1),\mathbf{d}_{SA}(A,2),...,\mathbf{d}_{SA}(A,h))=\arg\min_{\hat{\mathbf{d}}(A,1),\hat{\mathbf{d}}(A,2),...,\hat{\mathbf{d}}(A,h)}\Phi_{SA},$$

where $\Phi_{SA}=\dfrac{1}{h}\ln(\det[\mathbf{U}^{\mathrm{T}}(A)\mathbf{U}(A)])$.

Estimations of coefficients of *SA*-optimum regression equations system, received on sample A, have the form (see results (27)–(32) in [5])

$$\mathbf{d}_{SA}(A)=(\mathbf{A}(A)^{\mathrm{T}}\mathbf{A}(A))^{-1}\mathbf{A}(A)^{\mathrm{T}}\mathbf{B}(A)\ \mathbf{y}(A), \tag{13}$$

where

$$\mathbf{y}(A)=\begin{pmatrix}\mathbf{y}(A,1)\\ \hdashline \mathbf{y}(A,2)\\ \hdashline \vdots \\ \hdashline \mathbf{y}(A,h)\end{pmatrix}; \tag{14}$$

matrixes $\mathbf{A}_{k\bullet}(A)$ and $\mathbf{B}(A)$ are the block matrixes such, that for k -th a "block line" $\mathbf{A}_{k\bullet}(A)$ matrix $\mathbf{A}(A)$ is carried out

$$\mathbf{A}_{k\bullet}(A)=\left[\alpha_{1k}\mathbf{P}(A,k)\overset{o}{\mathbf{X}}(A,1) \;\vdots\; \alpha_{2k}\mathbf{P}(A,k)\overset{o}{\mathbf{X}}(A,2) \;\vdots\; \cdots \;\vdots\; \alpha_{hk}\mathbf{P}(A,k)\overset{o}{\mathbf{X}}(A,h)\right], \tag{15}$$

for k -th a "block line" $\mathbf{B}_{k\bullet}(A)$ matrixes $\mathbf{B}(A)$ is carried out

$$\mathbf{B}_{k\bullet}(A)=\left[\alpha_{1k}\,\mathbf{P}(A,k) \;\vdots\; \alpha_{2k}\,\mathbf{P}(A,k) \;\vdots\; \cdots \;\vdots\; \alpha_{hk}\,\mathbf{P}(A,k)\right]. \tag{16}$$

Let's note, that the size of a block matrix $\mathbf{A}(A)$ is equal $h\times h$ blocks or $M\times M$ elements where $M=m(1)+m(2)+...+m(h)$, and the size of a block matrix $\mathbf{B}(A)$ is equal $h\times h$ blocks or $M\times N(A)$ elements where $N(A)=n(A)\times h$.

In formulas (15) and (16) sizes α_{qk}, $k,q=1,2,...,h,$ are elements of matrixes, the opposites of the covariance matrixes set in (12)

$$\alpha_{qk}=[\overset{o}{\mathbf{\Sigma}}{}^{-1}]_{qk}, \tag{17}$$

and $\mathbf{P}(A,k)$ – a matrix of projection at independent estimation coefficients k -th model on a method of the least squares:

$$\mathbf{P}(A,k)=(\overset{o}{\mathbf{X}}{}^{\mathrm{T}}(A,k)\overset{o}{\mathbf{X}}(A,k))^{-1}\overset{o}{\mathbf{X}}{}^{\mathrm{T}}(A,k). \tag{18}$$

If to enter a designation

$$\mathbf{C}(A)=(\mathbf{A}^{\mathrm{T}}(A)\mathbf{A}(A))^{-1}\mathbf{A}^{\mathrm{T}}(A)\mathbf{B}(A) \tag{19}$$

(the size of a matrix $\mathbf{C}(A)$ is equal $h\times h$ blocks or $M\times N(A)$ elements), then the *SA*-optimum estimations of coefficients received on sample A, it is possible to write down as

$$\mathbf{d}_{SA}(A)=\begin{pmatrix}\mathbf{d}_{SA}(A,1)\\ \hdashline \mathbf{d}_{SA}(A,2)\\ \hdashline \vdots \\ \hdashline \mathbf{d}_{SA}(A,h)\end{pmatrix}=\begin{bmatrix}\mathbf{C}_{1\bullet}(A)\\ \hdashline \mathbf{C}_{2\bullet}(A)\\ \hdashline \vdots \\ \hdashline \mathbf{C}_{h\bullet}(A)\end{bmatrix}\begin{pmatrix}\mathbf{y}(A,1)\\ \hdashline \mathbf{y}(A,2)\\ \hdashline \vdots \\ \hdashline \mathbf{y}(A,h)\end{pmatrix},\ \text{ i.e. }\ \mathbf{d}_{SA}(A,k)=\mathbf{C}_{k\bullet}(A)\begin{pmatrix}\mathbf{y}(A,1)\\ \hdashline \mathbf{y}(A,2)\\ \hdashline \vdots \\ \hdashline \mathbf{y}(A,h)\end{pmatrix}, \tag{20}$$

where

$$\mathbf{C}_{k\bullet}(A)=\left[\mathbf{C}_{k1}(A) \;\vdots\; \mathbf{C}_{k2}(A) \;\vdots\; \cdots \;\vdots\; \mathbf{C}_{kh}(A)\right] - \tag{21}$$

k -th "block row" matrixes $\mathbf{C}(A)$.

For outputs conforming of regression models is executed

$$\mathbf{y}(A,k)=\hat{\mathbf{y}}(A,k)+\mathbf{u}(A,k)=\overset{o}{\mathbf{X}}(A,k)\hat{\mathbf{d}}(A,k)+\mathbf{u}(A,k),\ k=1,2,...,h, \tag{22}$$

where vectors of the rests for which it is carried out ($\mathbf{0}_{n(A)}$ – a zero $(n(D)\times 1)$ vector)

$$E\{\mathbf{u}(A,k)\}=\mathbf{0}_{n(A)}. \tag{23}$$

Let's enter a matrix of outputs of models

$$\hat{\mathbf{Y}}(A)=[\hat{\mathbf{y}}(A,1),\ \hat{\mathbf{y}}(A,2),...,\ \hat{\mathbf{y}}(A,h)], \tag{24}$$

the incorporated matrix регрессоров

$$\overset{o}{\mathbf{X}}(A)=[\overset{o}{\mathbf{X}}(A,1),\overset{o}{\mathbf{X}}(A,2),...,\overset{o}{\mathbf{X}}(A,h)] \tag{25}$$

and a matrix of the rests of models

$$\mathbf{U}(A)=[\mathbf{u}(A,1),\ \mathbf{u}(A,2),...,\ \mathbf{u}(A,h)]. \tag{26}$$

Then (22) it is possible to write down in the generalized kind:

$$\mathbf{Y}(A)=\hat{\mathbf{Y}}(A)+\mathbf{U}(A)=\overset{o}{\mathbf{X}}(A)\mathbf{D}_{SA}(A)+\mathbf{U}(A)\,, \tag{27}$$

where

$$\mathbf{D}_{SA}(A)=\begin{bmatrix}\mathbf{d}_{SA}(A,1) & \mathbf{0}_{m(1)} & \cdots & \mathbf{0}_{m(1)}\\ \mathbf{0}_{m(2)} & \mathbf{d}_{SA}(A,2) & \cdots & \mathbf{0}_{m(2)}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{0}_{m(h)} & \mathbf{0}_{m(h)} & \cdots & \mathbf{d}_{SA}(A,h)\end{bmatrix}. \tag{28}$$

Let's consider a matrix $\mathbf{W}(A)$ in which the (k,q)-element is calculated on columns matrixes of the rests (26) with numbers k and q:

$$[\mathbf{W}(A)]_{kq}=[\mathbf{U}^{\mathrm{T}}(A)\mathbf{U}(A)]_{kq}=\left[[\mathbf{Y}(A)-\hat{\mathbf{Y}}(A)]^{\mathrm{T}}[\mathbf{Y}(A)-\hat{\mathbf{Y}}(A)]\right]_{kq}=\mathbf{u}^{\mathrm{T}}(A,k)\,\mathbf{u}(A,q) \tag{29}$$

and has a mathematical expectation

$$\begin{aligned}[\boldsymbol{\Omega}(A)]_{kq}=E\{[\mathbf{W}(A)]_{kq}\}=n(A)\;\sigma_{kq}-\sum_{s=1}^{h}\sigma_{sk}\ \mathrm{tr}\ [\overset{o}{\mathbf{X}}(A,q)\ \mathbf{C}_{qs}(A)]-\\ -\sum_{r=1}^{h}\sigma_{qr}\ \mathrm{tr}\ [\overset{o}{\mathbf{X}}(A,k)\ \mathbf{C}_{kr}(A)]+\\ +\sum_{r=1}^{h}\sum_{s=1}^{h}\sigma_{rs}\ \mathrm{tr}\left[\overset{o}{\mathbf{X}}(A,q)\mathbf{C}_{qs}(A)[\mathbf{C}_{kr}(A)]^{\mathrm{T}}\ \overset{o}{\mathbf{X}}{}^{\mathrm{T}}(A,k)\right].\end{aligned} \tag{30}$$

The covariance matrix of the rests (26) as follows from (30) is determined on the set matrixes of plans $\overset{o}{\mathbf{X}}(A,1),\ \overset{o}{\mathbf{X}}(A,2),...,\ \overset{o}{\mathbf{X}}(A,h)$ and covariance matrix of mistakes of supervision $\boldsymbol{\Sigma}$.

We use results (13)–(30) for the construction of system criterion of regularity GMDH intended for search of optimum set of regressors in system of regressive models.

3 System Criterion of Regularity GMDH

Let set $V=(V(1),V(2),...,V(h))$ is some analyzed set regressors in system regression the equations ($V\subseteq X$, X is the set of regressors such, that $\overset{o}{X}\subseteq X$), and a matrix

$$\mathbf{V}(D)=[\mathbf{V}(D,1),\mathbf{V}(D,2),...,\mathbf{V}(D,h)] \tag{31}$$

is matrix of regressors observations on the sample $D\ (D=A,B)$, corresponding to set regressors V.

Let's consider a $(n(B)\times h)$-matrix

$$\mathbf{U}(B/A,V)=\mathbf{Y}(B)-\hat{\mathbf{Y}}(B/A,V)=\mathbf{Y}(B)-\mathbf{V}(B)\mathbf{D}_{SA}(A,V)\,. \tag{32}$$

where $\mathbf{Y}(B)$ is a matrix of observations of output variables (9) on sample B; $\hat{\mathbf{Y}}(B/A,V)$ is a matrix of outputs of regression models system on the verifying sample B, designed on model (regression models system) which $\mathbf{D}_{SA}(A,V)$ estimations of coefficients are received on training sample A for analyzed set of regressors V according to (13)–(21).

Definition 2. A random variable

$$ARS(V)=\frac{1}{h}\ln\Big(\det\left[\mathbf{U}^{\mathrm{T}}(B/A,V)\mathbf{U}(B/A,V)\right]\Big) \tag{33}$$

is defined as system criterion of regularity for GMDH method.

For a mathematical expectation of a matrix

$$\mathbf{W}(B/A,V)=\mathbf{U}^{\mathrm{T}}(B/A,V)\mathbf{U}(B/A,V) \tag{34}$$

is carried out

$$E\{\mathbf{W}(B/A,V)\}=E\left\{\left[\mathbf{Y}(B)-\mathbf{V}(B)\hat{\mathbf{D}}(A,V)\right]^{\mathrm{T}}\left[\mathbf{Y}(B)-\mathbf{V}(B)\hat{\mathbf{D}}(A,V)\right]\right\}=$$

$$=E\left\{\left[\overset{\circ}{\mathbf{Y}}(B)+\mathbf{\Xi}(B)-\mathbf{V}(B)\left[\begin{array}{c|c|c|c}\hat{\mathbf{d}}(A,V;1) & \mathbf{0}_{m(1)} & \cdots & \mathbf{0}_{m(1)}\\ \hline \mathbf{0}_{m(2)} & \hat{\mathbf{d}}(A,V;2) & \cdots & \mathbf{0}_{m(2)}\\ \hline \vdots & \vdots & \ddots & \vdots\\ \hline \mathbf{0}_{m(h)} & \mathbf{0}_{m(h)} & \cdots & \hat{\mathbf{d}}(A,V;h)\end{array}\right]\right]^{\mathrm{T}}[\cdots]\right\}=$$

$$=E\left\{\left[\overset{\circ}{\mathbf{Y}}(B)+\mathbf{\Xi}(B)-\mathbf{V}(B)\left[\begin{array}{c|c|c|c}\mathbf{C}_{1\bullet}(A,V)\mathbf{y}(A) & \mathbf{0}_{m(1)} & \cdots & \mathbf{0}_{m(1)}\\ \hline \mathbf{0}_{m(2)} & \mathbf{C}_{2\bullet}(A,V)\mathbf{y}(A) & \cdots & \mathbf{0}_{m(2)}\\ \hline \vdots & \vdots & \ddots & \vdots\\ \hline \mathbf{0}_{m(h)} & \mathbf{0}_{m(h)} & \cdots & \mathbf{C}_{h\bullet}(A,V)\mathbf{y}(A)\end{array}\right]\right]^{\mathrm{T}}[\cdots]\right\}=$$

$$=(\Delta\overset{\circ}{\mathbf{Y}}(B,V))^{\mathrm{T}}\Delta\overset{\circ}{\mathbf{Y}}(B,V)+n_B\,\mathbf{\Sigma}+\mathbf{M}(B/A,V)=\mathbf{\Omega}(B/A,V)\,, \tag{35}$$

where it is accepted $E\{[\mathbf{H}]^{\mathrm{T}}[...]\}=E\{[\mathbf{H}]^{\mathrm{T}}[\mathbf{H}]\}$ for reduction of record of a bulky matrix $\mathbf{H}$;

$$\Delta\overset{\circ}{\mathbf{Y}}(B)=\overset{\circ}{\mathbf{Y}}(B)-\mathbf{V}(B)\left[\begin{array}{c|c|c|c}\mathbf{C}_{1\bullet}(A,V)\overset{\circ}{\mathbf{y}}(A) & \mathbf{0}_{m(1)} & \cdots & \mathbf{0}_{m(1)}\\ \hline \mathbf{0}_{m(2)} & \mathbf{C}_{2\bullet}(A,V)\overset{\circ}{\mathbf{y}}(A) & \cdots & \mathbf{0}_{m(2)}\\ \hline \vdots & \vdots & \ddots & \vdots\\ \hline \mathbf{0}_{m(h)} & \mathbf{0}_{m(h)} & \cdots & \mathbf{C}_{h\bullet}(A,V)\overset{\circ}{\mathbf{y}}(A)\end{array}\right]; \tag{36}$$

$$[\mathbf{M}(B/A,V)]_{kq}=\sum_{r=1}^{h}\sum_{s=1}^{h}\sigma_{rs}\,\mathrm{tr}\left[\mathbf{V}(B,q)\mathbf{C}_{qr}(A,V)[\mathbf{C}_{ks}(A,V)]^{\mathrm{T}}\mathbf{V}^{\mathrm{T}}(B,k)\right] \tag{37}$$

is (k,q)-element of a matrix $\mathbf{M}(B/A,V)$.

Let's consider the scheme of repeated supervision [6] in which for each vector of observations of input variables the pair observations of output variables is executed, and the "first" observation of each pair form sample A, and the "second" observation form sample B, i.e.

$$n_A=n_B=n,\quad \mathbf{V}(A)=\mathbf{V}(B)=\mathbf{V},\quad \mathbf{C}(A)=\mathbf{C}(B)=\mathbf{C}\,. \tag{38}$$

In the scheme of repeated supervision for $E\{\mathbf{W}^*(B/A,V)\}$ it is carried out

$$E\{\mathbf{W}^*(B/A,V)\}=E\left\{\left[\mathbf{Y}(B)-\mathbf{V}\hat{\mathbf{D}}(A,V)\right]^{\mathrm{T}}\left[\mathbf{Y}(B)-\mathbf{V}\hat{\mathbf{D}}(A,V)\right]\right\}=$$

$$=E\left\{\left[\overset{\circ}{\mathbf{Y}}+\mathbf{\Xi}(B)-\mathbf{V}\left[\begin{array}{c|c|c|c}\mathbf{C}_{1\bullet}(V)\mathbf{y}(A) & \mathbf{0}_{m(1)} & \cdots & \mathbf{0}_{m(1)}\\ \hline \mathbf{0}_{m(2)} & \mathbf{C}_{2\bullet}(V)\mathbf{y}(A) & \cdots & \mathbf{0}_{m(2)}\\ \hline \vdots & \vdots & \ddots & \vdots\\ \hline \mathbf{0}_{m(h)} & \mathbf{0}_{m(h)} & \cdots & \mathbf{C}_{h\bullet}(V)\mathbf{y}(A)\end{array}\right]\right]^{\mathrm{T}}[\cdots]\right\}=$$

$$=(\Delta\overset{\circ}{\mathbf{Y}}(V))^{\mathrm{T}}\Delta\overset{\circ}{\mathbf{Y}}(V)+n\,\mathbf{\Sigma}(B)+\mathbf{M}^*(B/A,V)=\mathbf{\Omega}^*(B/A,V)\,, \tag{39}$$

where

$$\Delta\overset{\circ}{\mathbf{Y}}(V)=\overset{\circ}{\mathbf{Y}}(V)-\mathbf{V}\left[\begin{array}{c|c|c|c}\mathbf{C}_{1\bullet}(V)\overset{\circ}{\mathbf{y}} & \mathbf{0}_{m(1)} & \cdots & \mathbf{0}_{m(1)}\\ \hline \mathbf{0}_{m(2)} & \mathbf{C}_{2\bullet}(V)\overset{\circ}{\mathbf{y}} & \cdots & \mathbf{0}_{m(2)}\\ \hline \vdots & \vdots & \ddots & \vdots\\ \hline \mathbf{0}_{m(h)} & \mathbf{0}_{m(h)} & \cdots & \mathbf{C}_{h\bullet}(V)\overset{\circ}{\mathbf{y}}\end{array}\right], \tag{40}$$

and the (k,q) -element of a matrix $\mathbf{M}^*(B/A,V)$ looks like

$$[\mathbf{M}^*(B/A,V)]_{kq} = \sum_{r=1}^{h}\sum_{s=1}^{h}\sigma_{rs}\ \mathrm{tr}\ \left[\mathbf{V}(q)\mathbf{C}_{qr}(V)[\mathbf{C}_{ks}(V)]^{\mathrm{T}}\mathbf{V}^{\mathrm{T}}(k)\right]. \tag{41}$$

Definition 3. The optimal set of regressors is defined as the set regressors $V_0 \subseteq X$ for which

$$V_0 = \arg\min_{V\subseteq X} E\{ARS(V)\}. \tag{42}$$

Definition 4. The optimal regression equations system with respect to number and composition of the regressors is defined as regression equations system constructed on the set of regressors V_0.

Let's receive a condition of a reduction (simplification) of the regression equations system of optimum complexity for system criterion of regularity.

4 Condition of Reduction (Simplification) of Regression Equations System Optimum Complexity

Let's consider two sets of regressors V_1 and V_2

$$V_1 = \overset{\mathrm{o}}{X} = \{\overset{\mathrm{o}}{X}(1),\overset{\mathrm{o}}{X}(2),...,\overset{\mathrm{o}}{X}(h-1),\overset{\mathrm{o}}{X}(h)\},\ \ V_2 = V = \{\overset{\mathrm{o}}{X}(1),\overset{\mathrm{o}}{X}(2),...,\overset{\mathrm{o}}{X}(h-1),\overline{X}(h)\}, \tag{43}$$

where $\overline{X}(h)$ is a subset of regressors such, that $\overset{\mathrm{o}}{X}(h) = \overline{X}(h)\bigcup\overline{M}(h)$, $\overline{X}(h)\bigcap\overline{M}(h) = \varnothing$ ($\varnothing$ is the empty set), i.e. $\overline{M}(h)$ is set missed regressors in the regression equation with number h.

Let's calculate a difference

$$\Delta(V_1,V_2) = \Delta(\overset{\mathrm{o}}{X},V) = E\{ARS^*(\overset{\mathrm{o}}{X})\} - E\{ARS^*(V)\} =$$
$$= \frac{1}{h}E\left\{\ln\left(\det\left[\mathbf{W}_1^*(B/A,\overset{\mathrm{o}}{X})\right]\right)\right\} - \frac{1}{h}E\left\{\ln\left(\det\left[\mathbf{W}_2^*(B/A,V)\right]\right)\right\} =$$
$$= \frac{1}{h}\ln\left(\det\left[E\left\{\mathbf{W}_1^*(B/A,\overset{\mathrm{o}}{X})\right\}\right]\cdot\prod_{k=1}^{h}(n-k)\right) - \frac{1}{h}\ln\left(\det\left[E\left\{\mathbf{W}_2^*(B/A,V)\right\}\right]\cdot\prod_{k=1}^{h}(n-k)\right) =$$
$$= \frac{1}{h}\ln\left(\frac{\det\left[E\left\{\mathbf{W}_1^*(B/A,\overset{\mathrm{o}}{X})\right\}\right]}{\det\left[E\left\{\mathbf{W}_2^*(B/A,V)\right\}\right]}\right) = \frac{1}{h}\ln\left(\frac{\det\left[\mathbf{\Omega}_1^*(B/A,\overset{\mathrm{o}}{X})\right]}{\det\left[\mathbf{\Omega}_2^*(B/A,V)\right]}\right). \tag{44}$$

For calculation in (44) mathematical expectation of a determinant of the matrix having Wishart distribution, results [8, the theorem 7.5.3, by page 236-237] are applied.

Taking into account (31)–(43), and also that fact, that at $V_1 = \overset{\mathrm{o}}{X}$ from (40) follows $\Delta\overset{\mathrm{o}}{\mathbf{Y}}(\overset{\mathrm{o}}{X}) = \mathbf{O}_{(n\times h)}$ is a zero $(n\times h)$ -matrix, for matrixes $\mathbf{\Omega}_1^*(B/A,\overset{\circ}{X})$ and $\mathbf{\Omega}_2^*(B/A,V)$ in (44) is received

$$\mathbf{\Omega}_1^*(B/A,\overset{\mathrm{o}}{X}) = n\mathbf{\Sigma} + \mathbf{M}^*(B/A,\overset{\mathrm{o}}{X}), \tag{45}$$

$$\mathbf{\Omega}_2^*(B/A,V_2) = (\Delta\overset{\mathrm{o}}{\mathbf{Y}}(V))^{\mathrm{T}}\Delta\overset{\mathrm{o}}{\mathbf{Y}}(V) + n\mathbf{\Sigma} + \mathbf{M}^*(B/A,V) =$$
$$= \left[\begin{array}{c|c|c|c} 0 & 0 & \cdots & 0 \\ \hline 0 & 0 & \cdots & 0 \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline 0 & 0 & \cdots & (\delta\overset{\mathrm{o}}{\mathbf{y}}(h))^{\mathrm{T}}\delta\overset{\mathrm{o}}{\mathbf{y}}(h) \end{array}\right] + n\mathbf{\Sigma} + \mathbf{M}^*(B/A,V), \tag{46}$$

where

$$\delta\overset{o}{\mathbf{y}}(h)=\overset{o}{\mathbf{y}}(h)-\overline{\mathbf{y}}(h)\,, \tag{47}$$

$$\overline{\mathbf{y}}(h)=\overline{\mathbf{X}}(h)\sum_{q=1}^{h}\mathbf{C}_{hq}(V)\overset{o}{\mathbf{y}}(q)=\overline{\mathbf{X}}(h)\sum_{q=1}^{h}\mathbf{C}_{hq}(V)\overset{o}{\mathbf{X}}(q)\overset{o}{\boldsymbol{\theta}}(q)\,, \tag{48}$$

and for elements of matrixes $\mathbf{M}^{*}(B/A,\overset{o}{X})$ and $\mathbf{M}^{*}(B/A,V)$ it is carried out $(k,q=1,2,...,h)$:

$$[\mathbf{M}^{*}(B/A,\overset{o}{X})]_{qk}=\sum_{r=1}^{h}\sum_{s=1}^{h}\sigma_{rs}\,\mathrm{tr}\left[\overset{o}{\mathbf{X}}(q)\mathbf{C}_{qr}(\overset{o}{X})[\mathbf{C}_{ks}(\overset{o}{X})]^{\mathrm{T}}\overset{o}{\mathbf{X}}{}^{\mathrm{T}}(k)\right], \tag{49}$$

$$[\mathbf{M}^{*}(B/A,V)]_{qk}=\sum_{r=1}^{h}\sum_{s=1}^{h}\sigma_{rs}\,\mathrm{tr}\left[\mathbf{V}(q)\mathbf{C}_{qr}(V)[\mathbf{C}_{ks}(V_2)]^{\mathrm{T}}\mathbf{V}^{\mathrm{T}}(k)\right]. \tag{50}$$

The condition $\Delta(\overset{o}{X},V)>0$ represents a condition of a reduction (simplification) of the regression equations system of optimum complexity.

Unfortunately, to simplify a condition of a reduction in view of representation (43) it is extremely difficult because of block character of a matrix $\mathbf{C}(V)$. But such condition can be received for a special case orthogonal regressors.

4.1 Case of orthogonal regressors

As is known, the reduction regression models of optimum complexity (its simplification on number included regressors) in conditions confined samples of observations can occur for two reasons: first, because of dependence between regressors, and, second, because of smallness the contribution given regressor in model and commensurabilities of this contribution with a dispersion of mistakes of observations of a output variable. The assumption of orthogonality regressors leaves possible only the second reason of a reduction of model.

Let's assume, that the set missed regressors $\overline{M}(h)$ in (43) will consist only of one regressor, i.e. for corresponding matrixes regressors and vectors of coefficients of model splitting are carried out

$$\overset{o}{\mathbf{X}}(h)=\left[\overline{\mathbf{X}}(h)\;\vdots\;\mathbf{m}\right],\quad \overset{o}{\boldsymbol{\theta}}(h)=\begin{pmatrix}\overset{o}{\boldsymbol{\theta}}_{\overline{X}}(h)\\ \hdashline \overset{o}{\theta}_{m}(h)\end{pmatrix}, \tag{51}$$

where $\overline{\mathbf{X}}(h)$ is $(n\times(m(h)-1))$-matrix of observations of set regressors $\overline{M}(h)$; $\mathbf{m}$ is $(n\times 1)$-vector of observations missed regressor; $\overset{o}{\boldsymbol{\theta}}_{\overline{X}}(h)$ is $((m(h)-1)\times 1)$-vector of factors for set regressors $\overline{X}(h)$; $\overset{o}{\theta}_{m}(h)$ is a coefficient at missed regressor.

Let's calculate consistently matrixes $\mathbf{P}(\overset{o}{X},k)$, $\mathbf{A}(\overset{o}{X})$, $\mathbf{B}(\overset{o}{X})$, $\mathbf{A}^{\mathrm{T}}(\overset{o}{X})\mathbf{A}(\overset{o}{X})$, $\mathbf{A}^{\mathrm{T}}(\overset{o}{X})\mathbf{B}(\overset{o}{X})$, $\mathbf{C}(\overset{o}{X})$ in view of the assumption of orthogonality of regressors

$$\mathbf{P}(\overset{o}{X},k)=\left[\overset{o}{\mathbf{X}}{}^{\mathrm{T}}(k)\overset{o}{\mathbf{X}}(k)\right]^{-1}\overset{o}{\mathbf{X}}{}^{\mathrm{T}}(k)=\left[\overset{o}{\mathbf{X}}{}^{\mathrm{T}}(h)\,[\overset{o}{\mathbf{x}}_1,\overset{o}{\mathbf{x}}_2,...,\overset{o}{\mathbf{x}}_{m(k)}]\right]^{-1}\overset{o}{\mathbf{X}}{}^{\mathrm{T}}(k)=$$

$$=\begin{bmatrix}\overset{o}{\mathbf{x}}{}_1^{\mathrm{T}}(k)\overset{o}{\mathbf{x}}_1(k) & 0 & \cdots & 0\\ 0 & \overset{o}{\mathbf{x}}{}_2^{\mathrm{T}}(k)\overset{o}{\mathbf{x}}_2(k) & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \overset{o}{\mathbf{x}}{}_{m(k)}^{\mathrm{T}}(k)\overset{o}{\mathbf{x}}_{m(k)}(k)\end{bmatrix}^{-1}\overset{o}{\mathbf{X}}{}^{\mathrm{T}}(k)=$$

$$=\mathbf{F}^{-1}(\overset{o}{X},k)\overset{o}{\mathbf{X}}{}^{\mathrm{T}}(k)\,,\quad k=1,2,...,h\,; \tag{52}$$

$$\mathbf{A}(\overset{o}{X})=\begin{bmatrix}\alpha_{11}\mathbf{I}_{m(1)} & \mathbf{O}_{m(1)\times m(2)} & \cdots & \mathbf{O}_{m(1)\times m(h)}\\ \mathbf{O}_{m(2)\times m(1)} & \alpha_{22}\mathbf{I}_{m(2)} & \cdots & \mathbf{O}_{m(2)\times m(h)}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{O}_{m(h)\times m(1)} & \mathbf{O}_{m(h)\times m(2)} & \cdots & \alpha_{hh}\mathbf{I}_{m(h)}\end{bmatrix}; \tag{53}$$

where $\mathbf{I}_{m(k)}$ is the $(m(k)\times m(k))$ unit matrix; $\mathbf{O}_{m(k)\times m(q)}$ is the $(m(k)\times m(q))$ -zero matrix;

$$\mathbf{B}(\mathring{X})=\left[\begin{array}{c|c|c|c}\alpha_{11}\mathbf{F}^{-1}(\mathring{X},1)\mathring{\mathbf{X}}^{\mathrm{T}}(1) & \alpha_{21}\mathbf{F}^{-1}(\mathring{X},1)\mathring{\mathbf{X}}^{\mathrm{T}}(1) & \cdots & \alpha_{h1}\mathbf{F}^{-1}(\mathring{X},1)\mathring{\mathbf{X}}^{\mathrm{T}}(1)\\ \hline \alpha_{12}\mathbf{F}^{-1}(\mathring{X},2)\mathring{\mathbf{X}}^{\mathrm{T}}(2) & \alpha_{22}\mathbf{F}^{-1}(\mathring{X},2)\mathring{\mathbf{X}}^{\mathrm{T}}(2) & \cdots & \alpha_{h2}\mathbf{F}^{-1}(\mathring{X},2)\mathring{\mathbf{X}}^{\mathrm{T}}(2)\\ \hline \vdots & \vdots & \ddots & \vdots\\ \hline \alpha_{1h}\mathbf{F}^{-1}(\mathring{X},h)\mathring{\mathbf{X}}^{\mathrm{T}}(h) & \alpha_{2h}\mathbf{F}^{-1}(\mathring{X},h)\mathring{\mathbf{X}}^{\mathrm{T}}(h) & \cdots & \alpha_{hh}\mathbf{F}^{-1}(\mathring{X},h)\mathring{\mathbf{X}}^{\mathrm{T}}(h)\end{array}\right]; \tag{54}$$

$$\mathbf{A}^{\mathrm{T}}(\mathring{X})\mathbf{A}(\mathring{X})=\left[\begin{array}{c|c|c|c}\alpha_{11}^{2}\,\mathbf{I}_{m(1)} & \mathbf{O}_{m(1)\times m(2)} & \cdots & \mathbf{O}_{m(1)\times m(h)}\\ \hline \mathbf{O}_{m(2)\times m(1)} & \alpha_{22}^{2}\,\mathbf{I}_{m(2)} & \cdots & \mathbf{O}_{m(2)\times m(h)}\\ \hline \vdots & \vdots & \ddots & \vdots\\ \hline \mathbf{O}_{m(h)\times m(1)} & \mathbf{O}_{m(h)\times m(2)} & \cdots & \alpha_{hh}^{2}\,\mathbf{I}_{m(h)}\end{array}\right]; \tag{55}$$

$$\left(\mathbf{A}^{\mathrm{T}}(\mathring{X})\mathbf{A}(\mathring{X})\right)^{-1}=\left[\begin{array}{c|c|c|c}\alpha_{11}^{-2}\,\mathbf{I}_{m(1)} & \mathbf{O}_{m(1)\times m(2)} & \cdots & \mathbf{O}_{m(1)\times m(h)}\\ \hline \mathbf{O}_{m(2)\times m(1)} & \alpha_{22}^{-2}\,\mathbf{I}_{m(2)} & \cdots & \mathbf{O}_{m(2)\times m(h)}\\ \hline \vdots & \vdots & \ddots & \vdots\\ \hline \mathbf{O}_{m(h)\times m(1)} & \mathbf{O}_{m(h)\times m(2)} & \cdots & \alpha_{hh}^{-2}\,\mathbf{I}_{m(h)}\end{array}\right]; \tag{56}$$

$$\mathbf{A}^{\mathrm{T}}(\mathring{X})\mathbf{B}(\mathring{X})=\left[\begin{array}{c|c|c|c}\alpha_{11}^{2}\mathbf{F}^{-1}(\mathring{X},1)\mathring{\mathbf{X}}^{\mathrm{T}}(1) & \mathbf{O}_{m(1)\times m(2)} & \cdots & \mathbf{O}_{m(1)\times m(h)}\\ \hline \mathbf{O}_{m(2)\times m(1)} & \alpha_{22}^{2}\mathbf{F}^{-1}(\mathring{X},2)\mathring{\mathbf{X}}^{\mathrm{T}}(2) & \cdots & \mathbf{O}_{m(2)\times m(h)}\\ \hline \vdots & \vdots & \ddots & \vdots\\ \hline \mathbf{O}_{m(h)\times m(1)} & \mathbf{O}_{m(h)\times m(2)} & \cdots & \alpha_{hh}^{2}\mathbf{F}^{-1}(\mathring{X},h)\mathring{\mathbf{X}}^{\mathrm{T}}(h)\end{array}\right]; \tag{57}$$

$$\mathbf{C}(\mathring{X})=\left[\begin{array}{c|c|c|c}\mathbf{F}^{-1}(\mathring{X},1)\mathring{\mathbf{X}}^{\mathrm{T}}(1) & \mathbf{O}_{m(1)\times m(2)} & \cdots & \mathbf{O}_{m(1)\times m(h)}\\ \hline \mathbf{O}_{m(2)\times m(1)} & \mathbf{F}^{-1}(\mathring{X},2)\mathring{\mathbf{X}}^{\mathrm{T}}(2) & \cdots & \mathbf{O}_{m(2)\times m(h)}\\ \hline \vdots & \vdots & \ddots & \vdots\\ \hline \mathbf{O}_{m(h)\times m(1)} & \mathbf{O}_{m(h)\times m(2)} & \cdots & \mathbf{F}^{-1}(\mathring{X},h)\mathring{\mathbf{X}}^{\mathrm{T}}(h)\end{array}\right]. \tag{58}$$

Using (52)–(58), for a matrix $\mathbf{M}^{*}(B/A,\mathring{X})$ in (45) and (49) it is received

$$\left[\mathbf{M}^{*}(B/A,\mathring{X})\right]_{kk}=\sigma_{kk}\,m(k),\ \left[\mathbf{M}^{*}(B/A,\mathring{X})\right]_{qk}=0,\ \ k,q=1,2,...,h,\ \ q\neq k\ . \tag{59}$$

Similarly for a matrix $\mathbf{M}^{*}(B/A,V)$ in (46) and (50) with the account (51) it is received

$$\left[\mathbf{M}^{*}(B/A,V)\right]_{qk}=0,\ \ k,q=1,2,...,h,\ \ q\neq k\ ; \tag{60}$$

$$\left[\mathbf{M}^{*}(B/A,V)\right]_{kk}=\sigma_{kk}\,m(k),\ \ k=1,2,...,h-1;\ \left[\mathbf{M}^{*}(B/A,V)\right]_{hh}=\sigma_{hh}\,(m(h)-1)\,. \tag{61}$$

Let's calculate now $\delta\mathring{\mathbf{y}}(h)$ from (46)–(47). With the account (51) for $\mathring{\mathbf{y}}(h)$ it is carried out

$$\mathring{\mathbf{y}}(h)=\mathring{\mathbf{X}}(h)\mathring{\boldsymbol{\theta}}(h)=\left[\overline{\mathbf{X}}(h)\ \vdots\ \mathbf{m}\right]\begin{pmatrix}\mathring{\boldsymbol{\theta}}_{\overline{X}}(h)\\ \hline \mathring{\theta}_{m}(h)\end{pmatrix}=\overline{\mathbf{X}}(h)\mathring{\boldsymbol{\theta}}_{\overline{X}}(h)+\mathbf{m}\mathring{\theta}_{m}(h)\,. \tag{62}$$

According to (48) and with the account (52)–(58) for $\overline{\mathbf{y}}(h)$ it is carried out

$$\overline{\mathbf{y}}(h)=\overline{\mathbf{X}}(h)\sum_{q=1}^{h}\mathbf{C}_{hq}(V)\mathring{\mathbf{X}}(q)\mathring{\boldsymbol{\theta}}(q)=\overline{\mathbf{X}}(h)\mathbf{F}^{-1}(\overline{X},h)\mathring{\mathbf{X}}(h)\mathring{\boldsymbol{\theta}}(h)=$$

$$=\overline{\mathbf{X}}(h)\mathring{\boldsymbol{\theta}}_{\overline{X}}(h)+\overline{\mathbf{X}}(h)\mathbf{F}^{-1}(\overline{X},h)\overline{\mathbf{X}}^{\mathrm{T}}(h)\mathbf{m}\mathring{\theta}_{m}(h)\,. \tag{63}$$

Taking into account (62) and (63), for $\delta \overset{o}{\mathbf{y}}(h)$ it is received

$$\delta \overset{o}{\mathbf{y}}(h) = \left[\mathbf{I}_n - \overline{\mathbf{X}}(h)\mathbf{F}^{-1}(\overline{X},h)\overline{\mathbf{X}}^{\mathrm{T}}(h)\right]\mathbf{m}\overset{o}{\theta}_m(h)\,, \tag{64}$$

and for $(\delta \overset{o}{\mathbf{y}}(h))^{\mathrm{T}}\delta \overset{o}{\mathbf{y}}(h)$ в (46) it is received

$$(\delta \overset{o}{\mathbf{y}}(h))^{\mathrm{T}}\delta \overset{o}{\mathbf{y}}(h) = \overset{o}{\theta}{}_m^2(h)\mathbf{m}^{\mathrm{T}}\left[\mathbf{I}_n - \overline{\mathbf{X}}(h)\mathbf{F}^{-1}(\overline{X},h)\overline{\mathbf{X}}^{\mathrm{T}}(h)\right]\mathbf{m}\,. \tag{65}$$

Results (62)–(65) are received without taking into account that, as $\overline{\mathbf{X}}(h)$ and $\mathbf{m}$ orthogonal. If to take into account this assumption for $\delta \overset{o}{\mathbf{y}}(h)$ and $(\delta \overset{o}{\mathbf{y}}(h))^{\mathrm{T}}\delta \overset{o}{\mathbf{y}}(h)$ it is received

$$\delta \overset{o}{\mathbf{y}}(h) = \mathbf{m}\overset{o}{\theta}_m(h)\,, \tag{66}$$

$$(\delta \overset{o}{\mathbf{y}}(h))^{\mathrm{T}}\delta \overset{o}{\mathbf{y}}(h) = \overset{o}{\theta}{}_m^2(h)\mathbf{m}^{\mathrm{T}}\mathbf{m}\,. \tag{67}$$

The matching (45) and (46) with the count (59), (61) and (67) proves, that the matrixes $\mathbf{\Omega}_1^*(B/A,\overset{o}{X})$ and $\mathbf{\Omega}_2^*(B/A,V)$ differ only $(h\times h)$ by units, for which the ratio is fair

$$\left[\mathbf{\Omega}_2^*(B/A,V)\right]_{hh} = \left[\mathbf{\Omega}_1^*(B/A,\overset{o}{X})\right]_{hh} + \overset{o}{\theta}{}_m^2(h)\mathbf{m}^{\mathrm{T}}\mathbf{m} - \sigma_{hh}\,. \tag{68}$$

Let's enter $(h\times 1)$ vector

$$\mathbf{b} = \begin{pmatrix} \mathbf{0}_{h-1} \\ \hline (\mathrm{abs}\,\{\gamma\})^{1/2} \end{pmatrix}, \tag{69}$$

where $\gamma = \overset{o}{\theta}{}_m^2(h)\ \mathbf{m}^{\mathrm{T}}\mathbf{m} - \sigma_{hh}$; $\mathbf{0}_{h-1}$ is zero $((h-1)\times 1)$ -vector.

With count (69) ratio (68) is possible to write as

$$\mathbf{\Omega}_2^*(B/A,V) = \mathbf{\Omega}_1^*(B/A,\overset{o}{X}) + \mathrm{sign}\{\gamma\}\,\mathbf{b}\mathbf{b}^{\mathrm{T}}\,. \tag{70}$$

Let's take into account a ratio (70) and we receive a condition of a reduction, prolonging calculation (44). For this purpose we shall take advantage of a rule of calculation of a continuant

$$\det\,[\mathbf{A} + \mathbf{b}\mathbf{b}^{\mathrm{T}}] = \det\,[\mathbf{A}]\,\left(1 + \mathbf{b}^{\mathrm{T}}\mathbf{A}^{-1}\,\mathbf{b}\right). \tag{71}$$

Applying (71) to (44) and taking into account (70), we obtain

$$\Delta(\overset{o}{X},V) = E\{ARS^*(\overset{o}{X})\} - E\{ARS^*(V)\} =$$

$$= \frac{1}{h}\ln\left(\frac{\det\left[\mathbf{\Omega}_1^*(B/A,\overset{o}{X})\right]}{\det\left[\mathbf{\Omega}_1^*(B/A,\overset{o}{X}) + \mathrm{sign}\{\gamma\}\mathbf{b}\mathbf{b}^{\mathrm{T}}\right]}\right) =$$

$$= -\frac{1}{h}\ln\left(1 + \mathrm{sign}\{\gamma\}\left[\mathbf{\Omega}_1^*(B/A,\overset{o}{X})\right]_{hh}\mathrm{abs}\,\{\gamma\}\right). \tag{72}$$

From (72) follows, that the condition of a reduction in a regression equations system is determined is familiar values $\gamma = \overset{o}{\theta}{}_m^2(h)\ \mathbf{m}^{\mathrm{T}}\mathbf{m} - \sigma_{hh}$

$$\text{если } \gamma > 0\,,\ \text{то } \Delta(\overset{o}{X},V) < 0\,, \tag{73}$$

i.e. the set of regressors $\overset{o}{X}$ is ""better" than set of regressors V ;

$$\text{если } \gamma < 0\,,\ \text{то } \Delta(\overset{o}{X},V) > 0\,, \tag{74}$$

i.e. the set of regressors V is ""better" than set of regressors $\overset{o}{X}$, and in this case regression equations system of optimum complexity becomes simpler on number live in it regressors.

5 Conclusions

The problem of search of a regression equations system of optimum complexity by principles of a Group Method of Data Handling surveyed. The criterion of quality of a regression equations system is offered which is system analogue of criterion of regularity. The criterion is researched in the scheme of repeated observations.

The structural uncertainty in a problem of construction of regression equations system can reveal in two kinds: 1) uncertainty on a degree of statistical relation between random by additive amounting in output variables – this kind of uncertainty surveyed in [4, 5], where for an arising problem of estimation of factors the conforming method designed; 2) uncertainty by quantity and structure of regressors in a regression equations system – this kind of uncertainty is object of research of [7] and this article.

To record all cases of a combination of the missed and redundant regressors in current set of regressors V is similar, how it was made in article [6] for one regression equation, it is extremely difficult by virtue of block character of a matrix $\mathbf{C}(V)$. Concerning the obtained condition of a reduction (74) it is possible to mark, that it will coincide condition of a reduction of *J*-optimum model (12) in [6], if for number of the missed regressors in (12) in [6] to put $p_2 = 1$ and in addition to assume orthogonality of regressors.

References

[1] A. G. Ivakhnenko and V. S. Stepashko, Pomekhoustoychivost' modelirovaniya (Noice Immunity of Modelling), Naukova dumka, Kiev (1985).
[2] A. G. Ivakhnenko and Yu. P. Yurachkovskiy, Modelirovaniye slozhnykh system po eksperimental'nym dannym (Modelling of Complex Systems for Experimental Data), Radio i Svyaz', Moscow (1987).
[3] Muller J.-A., Lemke F. Self-organizing Data Mining. Extracting Knowledge from Data. Libri, Hamburg (2000).
[4] A. P. Sarychev, Estimation of Coefficients in Systems of Regression Models, Problems of Control and Informatics, **4**, 74–82 (2003) (Journal of Automation and Information Sciences, Vol. 35, **7** (2003)).
[5] A. P. Sarychev, Estimation of Coefficients in Systems of Regression Models. A Case of Covarience Matrix, Known to Within Scalar Parameter, Artificial Intelligence (Ukraine), **4**, 182–190 (2004).
[6] A. P. Sarychev, Determination of *J*-optimal set of Regressors by the Repeated Observation Samples, Automatika, **3**, 58–66 (1993) (Soviet J. of Automation and Information Sciences, Vol. 25, **3** (1993)).
[7] A. P. Sarychev, System Regularity Criterion of Group Method of Data Handling, Problems of Control and Informatics, **6**, 25–37 (2006) (Journal of Automation and Information Sciences, Vol. 38, **11** (2006)).
[8] T.W. Anderson, An Introduction to Multivariate Statistical Analysis. – New York, John Wiley & Sons. Inc. (1958).