8. CLASS OF G_8-ESTIMATORS OF THE SOLUTIONS OF SYSTEMS OF LINEAR ALGEBRAIC EQUATIONS (SLAE)

Let some system S with input vector $\vec{x}^T = (x_1, \ldots, x_m)$ and output variables y be given. As a mathematical model M_1 of the system S, it is natural to take the equation $y = A(\vec{x}) + \varepsilon$ where $A(\vec{x})$ is some operator, and ε is an error of such representation. Choosing different input vectors $\vec{x}_1, \ldots, \vec{x}_n$ we have a system of equations

$$\vec{y} = A + \vec{\varepsilon},$$

where

$$A = \{A(\vec{x}_1), \ldots, A(\vec{x}_n)\}$$

is an operator acting in a space of vectors $\vec{x}$ with values in a space of vectors $\vec{y}^T = (y_1, \ldots, y_n)$, and $\vec{\varepsilon}^T = \{\varepsilon_1, \ldots, \varepsilon_n\}$ is a vector of errors of the model M_1. If $y = f(\vec{x})$, where f is an unknown analytic function, then for simplification of the calculations we can take the operator A to be

$$A\vec{x} = \sum_{i=1}^{m} c_i x_i; \quad A\vec{x} = \sum_{i=1}^{m} c_i \varphi_i(x_i)$$

or

$$A\vec{x} = \sum_{i=1}^{m} c_i x_i + \sum_{i,j=1}^{m} c_{ij} x_i x_j + \ldots + \sum_{i_1,\ldots,i_k=1}^{m} c_{i_1,\ldots,i_k} x_{i_1} \ldots x_{i_k},$$

where c_i, c_{ij}, $c_{i_1,\ldots,i_k}$ are unknown coefficients; φ_i are known functions. We note that in all these cases $A\vec{x} = \vec{c}^T \vec{z}$, where $\vec{c}$ is an unknown vector and $\vec{z}$ is a known vector. Thus, we arrive at model M_1 which is linear with respect to the unknown parameters:

$$\vec{y} = X\vec{c} + \tilde{\vec{\varepsilon}},$$

where $X^T = [\vec{z}_1, \ldots, \vec{z}_n]$ and $\tilde{\vec{\varepsilon}}$ is a vector of errors. In this section we formulate the methods of finding coefficients c_i if we have the observations of y and the input vectors $\vec{x}$.

8.1. The classical least squares method

Assume that a mathematical model of a system S has the form

$$y = \vec{x}^T \vec{c} + \varepsilon,$$

where $\vec{x}$ is an m-dimensional vector of input parameters, $\vec{c}$ is an unknown m-dimensional vector; y is the observable variable of a system S, and ε is a model error. Let n observations $y_1, \ldots, y_n$ of a system S under the values $\vec{x}_1, \ldots, \vec{x}_n$ of a vector $\vec{x}$ be made. Then for the unknown vector $\vec{c}$ we get the system of equations

$$\vec{y} = X\vec{c} + \vec{\varepsilon}, \tag{8.1}$$

where $\vec{\varepsilon}^T = (\varepsilon_1, \ldots, \varepsilon_n)$ is the observation error. The vectors $\vec{c}$ and $\vec{\varepsilon}$ in the system of equations (8.1) are unknown. This system of equations is undetermined with respect to the unknown vectors $\vec{c}$ and ε and in the general case has an infinite set of solutions. Calculating the vector $\vec{c}$ it is desirable to know the value of the vector $\vec{\varepsilon}$. However, because of the indeterminancy of the system (8.1), it is difficult to find the true value of the vector $\vec{\varepsilon}$ without any auxiliary conditions. We can reduce the system (8.1) to the form

$$\vec{y} = X\hat{\vec{c}}, \tag{8.2}$$

where the vector $\vec{c}$ is replaced by a new vector $\hat{\vec{c}}$ which is different from $\vec{c}$ in general. The preliminary investigations of the system (8.1) were made in the following way. In general the solution $\hat{\vec{c}}$ of a system (8.2) may not exist. However it is not necessary to find a solution of this system. We need to find the value of $\hat{\vec{c}}$ which minimizes some quality criterion of an estimator $F\left\{\vec{y} - X\hat{\vec{c}}\right\}$. For the simplification of calculations as the quality criterion the function

$$I(\vec{u}) = \vec{u}^T\vec{u} = \|\vec{u}\|^2$$

is usually chosen. If the inverse matrix $\left(X^TX\right)^{-1}$ exists, then we can obtain the minimizer of $\left\|\vec{y} - X\hat{\vec{c}}\right\|$ as

$$\hat{\vec{c}} = (X^TX)^{-1}X^T\vec{y}. \tag{8.3}$$

This formula explains the name, the "Least Squares Method". The estimation is

$$\hat{\vec{c}} - \vec{c} = (X^TX)^{-1}X^T\vec{\varepsilon}. \tag{8.4}$$

If the inverse matrix $\left(X^TX\right)^{-1}$ does not exist, then the function $\varphi(\vec{c}) := \|\vec{y} - X\vec{c}\|^2$ can have uncountable points of minimum. Again, to simplify calculations among all points of the minimum, the vector $\tilde{\vec{c}}$ with the smallest Euclidean norm is chosen. We can find this vector in the following way: consider the function

$$\varphi(\vec{c}, \alpha) := \|\vec{y} - X\vec{c}\|^2 + \alpha\|\vec{c}\|^2, \ \ \alpha > 0$$

instead of the function $\varphi(\vec{c}) := \|\vec{y} - X\vec{c}\|^2$. Because $\alpha > 0$, the minimum of function $\varphi(\vec{c}, \alpha)$ is unique and the vector $\vec{c}_\alpha$, under which the function $\varphi(\vec{c}, \alpha)$ will take the minimal value, is defined by the formula

$$\vec{c}_\alpha = \left(\alpha I + X^TX\right)^{-1}X^T\vec{y}. \tag{8.5}$$

It is easy to prove that $\lim_{\alpha\downarrow 0}\vec{c}_\alpha = \tilde{\vec{c}}$.

As the G-estimators of the regularized pseudo-solutions

$$\vec{x}_\alpha = \left[I\alpha + X^TX\right]^{-1}X^T\vec{b},$$

we choose a regularized solution

$$\vec{y}_\theta = \mathrm{Re}\left[I\left(\theta + \mathrm{i}\varepsilon\right) + \Xi^T\Xi\right]^{-1}\Xi^T\vec{b},$$

where $\varepsilon \neq 0$ and θ are real parameters, $\Xi = \left(\xi_{ij}^{(n)}\right)$ is the observation of the random matrix $X + H$, where H is a certain random matrix. The G-estimators of the values $\vec{x}_\alpha$ belong to the class of $\tilde{G}_8$ -estimators and are denoted by G_8. In this section, the following G_8-estimator of $\tilde{G}_8$-class is proposed

$$\vec{G}_8 = \mathrm{Re}\left[I\left(\hat{\theta}_1 + \mathrm{i}\varepsilon\right) + \Xi^T\Xi\right]^{-1}\Xi^T\vec{b}. \tag{8.6}$$

Here $\hat{\theta}_1$ is the maximal real solution of the equation

$$f_n(\theta) = \alpha, \tag{8.7}$$

where $\alpha \geq 0$,

$$f_n(\theta) = \theta \mathrm{Re}\left[1 + \delta_1 a\left(\theta\right)\right]^2 - \varepsilon \mathrm{Im}\left[1 + \delta_1 a\left(\theta\right)\right]^2 + \left(\delta_1 - \delta_2\right)\left[1 + \delta_1 \mathrm{Re} a(\theta)\right],$$

$$a\left(\theta\right) = \frac{1}{n}\mathrm{Tr}\left[I\left(\theta + \mathrm{i}\varepsilon\right) + \Xi^T\Xi\right]^{-1}, \; \delta_1 = \sigma_n^2 n\ , \; \delta_2 = \sigma_n^2 m,$$

σ_n^2 is the variance of entries $\xi_{ij}^{(n)}$ of the matrix $\Xi = \left(\xi_{ij}^{(n)}\right)$. We call equation (8.7) the main equation for the G_8-estimator.

It is proved [Gir44, Gir54, Gir69, Gir84] that under certain conditions, for every $\gamma > 0$

$$\lim_{\varepsilon\to 0}\lim_{n\to\infty}\mathbf{P}\left\{|\vec{d}[\vec{G}_8 - (I\alpha + X^TX)^{-1}X^T\vec{b}]| > \gamma\right\} = 0,$$

where $\vec{d}$ is an arbitrary vector such that $\vec{d}^T\vec{d} \leq c < \infty$.

8.2. Modified G_8-estimator of the solution of SLAE

In this section, the following modified G_8-estimator from the $\tilde{G}_8$-class for

$$\vec{x}_\alpha = [I\alpha + A^TA]^{-1}A^T\vec{b}$$

is proposed,

$$\vec{G}_8(\alpha, \varepsilon, B, C) = \mathrm{Im}\int_0^B\left\{\frac{\mathrm{e}^{|sp|}}{\pi}\int_{-C}^{C}\mathrm{Im}\left([I\hat{\theta} + X^TX]^{-1}X^T\vec{b}\right)\mathrm{e}^{-\mathrm{i}tp}\mathrm{d}t\right\}\mathrm{e}^{-p(\alpha-\mathrm{i}\varepsilon)}\mathrm{d}p.$$

Here $\hat{\theta}_1$ is the measurable complex solution of the equation

$$\hat{\theta}\left\{1 + \frac{\sigma^2}{n}\mathrm{Tr}\left[I\hat{\theta} + X^TX\right]^{-1}\right\}^2 + \left(1 - \frac{m_n}{n}\right)\left\{1 + \frac{\sigma^2}{n}\mathrm{Tr}\left[I\hat{\theta} + X^TX\right]^{-1}\right\} = -z,$$

σ_n^2 is the variance of entries $x_{ij}^{(n)}$ of observation $X = \left(x_{ij}^{(n)}\right)$ of matrix $A + \Xi$, $z = t + \mathrm{i}s$, $s \geq c$, c is a certain constant.

Under certain conditions we have ([Gir44], [Gir54], [Gir69], [Gir84])

$$\lim_{B\to\infty}\lim_{C\to\infty} p\lim_{n\to\infty}\vec{d}^T\left[\vec{G}_8(\alpha, \varepsilon, B, C) - \mathrm{Re}\left[I(\alpha + \mathrm{i}\varepsilon) + A^TA\right]^{-1}A^T\vec{b}\right] = 0.$$

8.3. G_8-estimator of the solutions of SLAE with block structure

For linear forms $\vec{d}^T\vec{x}_\alpha$ of regularized pseudo-solutions $\vec{x}_\alpha = [I\alpha + A^TA]^{-1}A^T\vec{b}$ of the systems of linear algebraic equations $A\vec{x} = \vec{b}$ with block structure, the following G_8-estimator

$$\vec{d}^T\vec{G}_8 = -\mathrm{Re}\vec{d}^T\left[C_1 + \mathrm{i}\varepsilon I_m + Z_s^T\left(C_2 - \mathrm{i}\varepsilon I_n\right)^{-1}Z_s\right]^{-1}Z_s^T\left(C_2 - \mathrm{i}\varepsilon I_n\right)^{-1}\vec{b},$$

is suggested. Here A is a matrix of the size $np\times mq$, $n \geq m$, $\vec{x}$ and $\vec{b}$ are vectors, $\alpha > 0$ is a parameter of regularization, $\varepsilon > 0$; $\vec{b}\in R^{np}$; $\vec{d}^T \in R^{mq}$; $Z_s = s^{-1}\sum_{i=1}^{s}X_i$; X_i are independent observations of the matrix $A+\Xi$, $\Xi = \left(\Xi_{ij}^{(n)}\right)_{i=1,\ldots,n}^{j=1,\ldots,m}$ is a random matrix with independent blocks $\Xi_{ij}^{(n)}$, $\mathbf{E}\,\Xi_{ij}^{(n)} = 0$, $\mathbf{E}\,\left\|\Xi_{ij}^{(n)}\right\|^2 < \infty$; and $C_1 = \left(C_{1i}\delta_{ij}\right)_{i,j=1}^{m}$, $C_2 = \left(C_{2i}\delta_{ij}\right)_{i,j=1}^{n}$ are block diagonal real matrices that are arbitrary measurable solutions of the system of nonlinear equations

$$C_{1l} + \mathrm{Re}\sum_{j=1}^{n}\left[\frac{1}{s}\mathbf{E}\,\Xi_{jl}^{(n)T}\left\{Q_{jj}\right\}\Xi_{jl}\right]_{Q=\left[C_2 - \mathrm{i}\varepsilon I_n + \tilde{X}(C_1+\mathrm{i}\varepsilon I_m)^{-1}\tilde{X}^T\right]^{-1}} = \alpha I;$$

$$C_{2k} + \mathrm{Re}\sum_{j=1}^{m}\frac{1}{s}\left[\mathbf{E}\,\Xi_{kj}^{(n)}\left\{\Theta_{jj}\right\}\Xi_{kj}^T\right]_{\Theta=\left[C_1+\mathrm{i}\varepsilon I_m + \tilde{X}^T(C_2-\mathrm{i}\varepsilon I_n)^{-1}\tilde{X}\right]^{-1}} = I,$$

$$k = 1,\ldots,n;\ p = 1,\ldots,m,\ \tilde{X} = Z_s.$$

It is proved [Gir84, p.236] that under certain conditions, for every $\gamma > 0$,

$$\lim_{\varepsilon\to 0}\lim_{n\to\infty}\mathbf{P}\left\{\left|\vec{d}^T\left(\vec{G}_8 - \vec{x}_\alpha\right)\right| > \gamma\right\} = 0.$$

8.4. G_8-estimator of the solutions of SLAE with symmetric block structure

Let $A\vec{x} = \vec{b}$ be a SLAE, where $A_{pq\times pq} = \left(A_{ks}^{(n)}\right)_{k,s=1}^{p}$, $A_{ks}^{(n)} = A_{ks}^{(n)T}$ and $A_{ks}^{(n)}$; $k \geq s$, $k,s = 1,...,p$ are blocks of the dimension q, and let $\vec{x}$, $\vec{b}$ be vectors. We consider the linear form of the regularized solution of such a system

$$\vec{d}^T\vec{x}_\varepsilon = \vec{d}^T\mathrm{Re}\left[A_{pq\times pq} + \mathrm{i}\varepsilon I_n\right]^{-1}\vec{b};\ \ \vec{d}\in R^n;\ n = pq;\ \varepsilon > 0.$$

For linear forms $\vec{d}^T\vec{x}_\varepsilon$ of regularized pseudo-solutions,

$$\vec{x}_\varepsilon = \mathrm{Re}\left[A_{pq\times pq} + \mathrm{i}\varepsilon I_n\right]^{-1}\vec{b},$$

of the systems of linear algebraic equations $A\vec{x} = \vec{b}$ with block structure, the following G_8–estimator

$$\vec{d}^T\vec{G}_8 = -\mathrm{Re}\left[X_{pq\times pq} + C\left(\varepsilon\right) + \mathrm{i}\varepsilon I_n\right]^{-1}\vec{b}$$

is considered. Here, $X_{pq\times pq}$ is an observation of matrix $\Xi_{pq\times pq} + A_{pq\times pq}$, $\Xi_{pq\times pq} = \left(\Xi_{ks}^{(n)}\right)_{k,s=1}^{p}$, $\Xi_{ks}^{(n)} = \Xi_{ks}^{(n)T}$ and $\Xi_{ks}^{(n)}$; $k \geq s$, $k.s = 1,...,p$ are independent random blocks of the dimension q, $C_{pq\times pq}(\varepsilon)=\left(\delta_{ij}C_{jj}^{(n)}(\varepsilon)\right)_{i,j=1}^{p}$ and the matrix-blocks $C_{ss}(\varepsilon)$ satisfy for $z=\mathrm{i}\varepsilon$ the canonical equation

$$C_{jj}(\varepsilon)=\operatorname{Re}\mathbf{E}\sum_{s=1}^{p}\Xi_{js}^{(n)}Q_{ss}\,\Xi_{js}^{(n)T}\Big|_{Q=[X_{pq\times pq}+C_{pq\times pq}(\varepsilon)+\mathrm{i}\varepsilon I_n]^{-1}}.$$

It is proven in [Gir84, p.250] that under certain conditions, for every $\gamma>0$

$$\lim_{\varepsilon\downarrow 0}\lim_{n\to\infty}\mathbf{P}\left\{\left|\vec{d}^{T}\left(\vec{G}_8-\vec{x}_{\varepsilon}\right)\right|>\gamma\right\}=0.$$

Random Matrices. Kiev University Publishing, Ukraine, 1975, 448pp. (in Russian).

Theory of Random Determinants. Kiev University Publishing, Ukraine, 1975, 368pp. (in Russian).

Limit Theorems for Functions of Random Variables. “Higher School” Publishing, Kiev, Ukraine, 1983, 207pp. (in Russian).

Multidimensional Statistical Analysis. “Higher School” Publishing, Kiev, Ukraine, 1983, 320pp. (in Russian).

Spectral Theory of Random Matrices. “Science” Publishing, Moscow, Russia, 1988, 376pp. (in Russian).

Theory of Random Determinants. Kluwer Academic Publishers, The Netherlands, 1990, 677pp.

Theory of Systems of Empirical Equations. “Lybid” Publishing, Kiev, Ukraine, 1990, 264pp. (in Russian).

Statistical Analysis of Observations of Increasing Dimensions. Kluwer Academic Publishers, The Netherlands, 1995, 286pp.

Theory of Linear Algebraic Equations with Random Coefficients. Allerton Press, Inc, New York, U.S.A. 1996, 320pp.

An Introduction to Statistical Analysis of Random Arrays. VSP, Utrecht, The Netherlands. 1998, 673pp.

Theory of Stochastic Canonical equations. Volume I and II. Kluwer Academic Publishers, (xxiv + 497, xviii + 463).(2001).