COMPUTATIVE PARADOXES IN MODERN DATA ANALYSIS

Y. V. CHEBRAKOV AND V. V. SHMAGIN
Department of Mathematics, Technical University,
Nevsky 3-11, 191186, St-Petersburg, Russia
E-mail: chebra@phdeg.hop.stu.neva.ru

By developing F. Smarandache thema on paradoxes in mathematics it is stated, firstly, if in measurement (natural science) experiments the best solutions are found by using methods of modern data analysis theory, then some difficulties with the interpretation of the computation results are liable to occur; secondly, one is not capable to overcome these difficulties without a data analysis theory modification, consisted in the translation of this theory from Aristotelian "binary logic" into more progressive "fuzzy logic".

Key words: data analysis, revealing outliers, confidence interval, fuzzy logic.

1 Introduction

As generally known from history of science, a scientific theory may have crisis in process of its development, when it disjoints in a set of fragment theories, that weak-coordinate each other and, as a whole, form a collection of various non-integrated conceptions. For instance, as we assume, F. Smarandache mathematical notions and questions [1-2] help us to understand quite well that a stable equilibrium, observed in mathematics at the present time, is no more than fantasy. Thus, it falls in exactly with F. Smarandache views that the finding and investigating paradoxes in mathematics is a very effective way of approximating to the truth and so at present each of scientific researches, continuing F. Smarandache thema[2], should be considered as very actual one.

Let us assume that *computative paradoxes* in mathematics are mainly such computation results, obtained by using mathematical methods, which are contradicted some mathematical statements. The main goal of this paper is to demonstrate that the mentioned crisis, demanding practical action instead of debate, occurs in modern data analysis, which formally has its own developed mathematical theory, but does not capable "to cope worthily" with a large number of practical problems of quantitative processing results of measurement experiments.

Another goal of this paper is to equip the mathematicians and software designers, working in the data analysis field, with a set of examples, demonstrating dramatically that, if, for solving some problems on analysing data arrays, one uses the standard computer programmes and/or time-tested methods of modern data analysis theory, then a set of the paradoxical computative results may be obtained.

2 Approximative problems of data analysis

2.1 The main problems of regression analysis theory and standard solution methods

As generally known[3-7], for found experimental dependence $\{y_n, x_n\}$ ($n = 1, 2, \ldots, N$) and given approximative function $F(\mathbf{A}, x)$, in the measurement (natural science) experiments the main problems of regression analysis theory are finding estimates of $\mathbf{A}'$ and y' and variances of $\delta\mathbf{A}'$ and $\delta(y-y')$, where $\mathbf{A}'$ is an estimate of vector parameter $\mathbf{A}$ of the function $F(\mathbf{A}, x)$ and $\{y_n'\} = \{F(\mathbf{A}', x_n)\}$. In particular, if $F(\mathbf{A}, x) = \sum_{l=1}^{L} a_l h_l(x)$ $\{F(\mathbf{A}, x)$ is a *linear* model$\}$, where $h_l(x)$ are some functions on x, then in received regression analysis theory *standard* solution of discussed problems has form

$$\mathbf{A}' = (\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\,\mathbf{H}^{\mathrm{T}}\mathbf{Y}, \quad (\delta\mathbf{A}')^2 = s/(N-L)\ \mathrm{diag}(\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1} \tag{1}$$

$$\delta_p(y-y') = y' \pm t_p\, s\sqrt{1+\mathbf{H}_i^{\mathrm{T}}(\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{H}_i}\,,$$

where $\mathbf{H}$ is a matrix $L \times N$ in size with n-th row $(h_1(x_n), h_2(x_n), \ldots, h_L(x_n))$; $\mathbf{H}^{\mathrm{T}}$ is the transposed matrix $\mathbf{H}$; $\mathbf{Y} = \{y_n\}$; $s = \sum_{n=1}^{N}(y_n - y_n')^2 / (N-L)$; $\mathbf{H}_i = (h_1(x_i), h_2(x_i), \ldots, h_L(x_i))$; the value of t_p is determined by t-Student distribution table and generally depends on the assigned value of the significance level of p and the value of $N-L$ (a number of freedom degree); at the assigned value of the significance level of p the notation of $\delta_p(y-y')$ means confidence interval for possible deviations of experimental values of y from computed values $y' = F(\mathbf{A}', x_i)$. According to Gauss - Markov theorem[4,5], for classical data analysis model

$$y_n = F(\mathbf{A}, x_n) + e_n \tag{2}$$

the solution (1) is the best (gives minimum value of s), if the following conditions are fulfilled:

all values of $\{x_n\}$ are not random, mathematical expectation of random value $\{e_n\}$ is equal to zero and random values of $\{e_n\}$ are non-correlated and have the same dispersions σ^2.

Example 1. In table 1 we adduce an experimental data array, obtained by Russian chemist D.I.Mendeleev in 1881, when he investigated the solvability (y, relative units) of sodium nitrate ($NaNO_3$) on the water temperature (x, °C).

Table 1.

D.I.Mendeleev data array

n	x_n	y_n	$y_n - y_n'$	n	x_n	y_n	$y_n - y_n'$	n	x_n	y_n	$y_n - y_n'$
1	0	66.7	–0.80	4	15	80.6	0.05	7	36	99.4	0.58
2	4	71.0	0.02	5	21	85.7	–0.07	8	51	113.6	1.73
3	10	76.3	0.10	6	29	92.9	0.17	9	68	125.1	–1.56

By analysing the data array $\{y_n, x_n\}$, presented in table 1, Y.V.Linnik[3] states that these data, as it was noted by D.I.Mendeleev, are well-fitted by linear model $y' = 67.5 + 0.871x$ ($\delta A' = (0.5; 0.2)$), although the correspondence between experimental and computed on linear model values of y is slightly getting worse at the beginning and end of investigated temperature region (see the values $\{y_n - y'_n\}$ adduced in table 1). We add that for discussed data array Y.V.Linnik[3] computes the confidence interval of $\delta_p (y - y')$ from (1) at the significance level of $p = 0.9$:

$$\delta_{0.9} (y - y') = \pm 0.593 \sqrt{1 + (x - 26)^2 / 451.1} \ . \qquad (3)$$

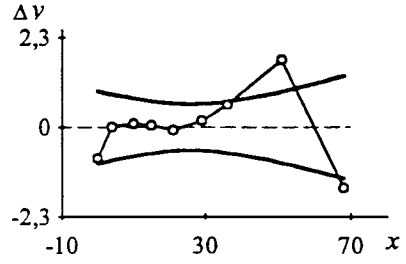

Figure 1. The plots of confidence interval of the deviation of y from y' (heavy lines) and residuals $y - y'$ (circles) for D.I.Mendeleev data array.

We show the plots of $\delta_{0.9} (y - y')$ on x by heavy lines in figure 1 and $\{y_n - y_n', x_n\}$ by the circles. Since the plot of $\{y_n - y_n', x_n\}$ steps over the heavy lines in figure 1, some computative difficulty is revealed:

the standard way (1), used by Y.V.Linnik[3] for determining the confidence interval of the deviations of y from y', is out of character with the discussed experimental data array.

It follows from results presented in table 1 and/or figure 1, if one assumes that $\delta(y - y') \geq \max |y_n - y_n'| = 1.73$ then the broken connections of the confidence interval $\delta(y - y')$ with D.I.Mendeleev data array will be pieced up. But values of $\delta A'$, calculated by Y.V.Linnik from (1), disagree with the values $\delta(y - y') \geq 1.73$, and, consequently,

standard values of $\delta A'$ is out of character with D.I.Mendeleev data array also.

2.2 Alternative methods of regression analysis theory

P.Huber[8] noted that, as the rule, 5 – 10% of all observations in the majority of analysing experimental arrays are anomalous or, in other words, the conditions of Gauss - Markov theorem, adduced above, are not fulfilled. Consequently, in practice instead of the standard solution (1), found by "least squares (LS) method", *alternative* methods, developed in the frames of received regression analysis theory, should be used. In particular, if the data array $\{y_n, x_n\}$ contains a set of

outliers, then for finding the best solution of discussed problem it is necessary[6, 7] or to remove all outliers from the analysing data array (*strategy 1*), or to compute the values of $\mathbf{A}'$ on the initial data array by means of M-robust estimators (*strategy 2*). For revealing outliers in the data array P.J.Rousseeuw and A.M.Leroy[9] suggest to use one of two combined statistical procedures, in which parameter estimates, minimising the median of the array $\{(y_n - y_n')^2\}$ (*the first procedure*) or the sum of K first elements of the same array (*the second procedure*), are considered as the best ones. If $F(\mathbf{A}, x)$ is a linear function (see above), then the robust M-estimates of $\mathbf{A}'$ are obtained as result of the solving of one from two minimisation problems[6-9]

$$S_\varphi(\mathbf{A}) = \sum_{n=1}^{N} \varphi(y_n - y_n') \Rightarrow \min \quad \text{or} \quad \partial S_\varphi / \partial a_l = \sum_{n=1}^{N} \psi(y_n - y_n')\, h_l(x_n) = 0, \qquad (4)$$

where function $\varphi(r)$ is symmetric concerning Y-axis, continuously differentiable with a minimum at zero and $\varphi(0) = 0$; $\psi(r)$ is a derivative of $\varphi(r)$ with respect to r.

Continued example 1. Since D.I.Mendeleev data array from table 1 contains outliers, we adduce results of quantitative processing this data by alternative methods, defined above.

1. Let in (4) Andrews function[10] be applied: $\varphi(r) = d(1-\cos(r/d))$ if $|r| \le d\pi$ and $\varphi(r) = 0$ if $|r| > d\pi$. It is articulate in figure 2 that in this case the values of the linear model parameters a_0 and a_1 depend on

a) the values of parameter d of Andrews function $\varphi(r)$;

b) the type of the minimisation robust regression problem (solutions of the first and second minimisation problem of (4) are marked respectively by triangles and circles in figure 2).

Thus, in this case a computative paradox declares itself in the fact, that

in actual practice the robust estimates are not robust

and so, as K.R.Draper and H.Smith[11] wrote already,

"unreasoning application of robust estimators looks like reckless application of ridge-estimators: they can be useful, but can be improper also. The main problem is such one, that we do not know, which robust estimators and at which types of supposes about errors are effectual to applicate; but some investigations in this direction have been done..."

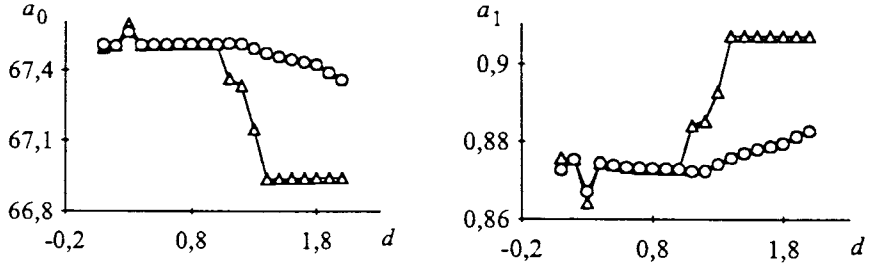

Figure 2. Dependences of parameters values of linear model a_0+a_1x on values internal parameter of robust Andrews estimator and the type of the minimisation problems (4).

2. Let us reveal outliers in D.I.Mendeleev data array by both combined statistical procedures[9], mentioned above.

Our computation results show

a) both procedures could not find the all four outliers (1, 7, 8 and 9) but the only three ones with numbers 1, 8 and 9;

b) if a set of readings with numbers 1, 8 and 9 is deleted from D.I.Mendeleev data array, then for the truncated data array the first procedure will not find a new outlier, but the second procedure will find two outliers yet that have numbers 2 and 6 in the initial data array.

Thus, in this case the main computative paradox is exhibited in the fact, that

revealing outliers problems solutions depend on a type of the used statistical procedures.

It remains for us to add, if one

a) computes y' by formula[6, 7]

$$y'(\xi,\ x) = g_\alpha(\ 68.12 + 0.02\xi + (0.85652 - 0.00046\xi)x\ ), \qquad (5)$$

then for each n the difference $|y_n - y_n'|$ will keep within the limit of the chosen above value for the confidence interval $\delta(y - y')$, where α=0.94; $0 \le \xi \le 35$; $g_\alpha(y) = 2\alpha[y/(2\alpha)] + 2\alpha$ at $|y - 2\alpha\ [y/(2\alpha)]| \ge \alpha$, otherwise $g_\alpha(y) = 2\alpha\ [y/(2\alpha)]$, [b] means integer part of b, Thus, another computative paradox occurs:

although for each contaminated data array a family of analytical solutions exists, the only single solution of the estimation problems is found in modern regression analysis theory.

b) puts the mentioned above extremal values of ξ in (5), one will be able to determine the exact limit of the variation for the linear model parameters a_0 and a_1: $a_0 = 67.77 \pm 0.35$ and $a_1 = 0.865 \pm 0.008$;

c) deletes a set of readings with numbers 1, 7, 8 and 9 from D.I.Mendeleev data array, one will obtain that in the truncated data array $\{y_n, x_n\}^*$ the difference of $|y_n - y_n'|$ for each n keeps within the limit of the error ε, where ε is the measuring error for readings $\{y_n\}^*$: $\varepsilon = 0.1$. Since in this case $\delta(y - y') \le \varepsilon$, the complete family of analytical solutions has form[6, 7]

$$y'(\xi, x) = g_\alpha(\ 67.566 + 0{,}002\xi + (0.870047 - 0.000097\xi)x\ ), \qquad (6)$$

where α=0.07; $0 \le \xi \le 45$ and, consequently, $a_0 = 67.521 \pm 0.045$ and $a_1 = 0.872 \pm 0.002$;

d) compares solutions (5) and (6) with the standard LS-solution, one can conclude that LS-estimations of parameter a_0 and a_1 $\{\mathbf{A}' = (67.5{\pm}0.5;\ 0.87{\pm}0.2)\}$ are pretty near equal of the mean values of these parameters in the general analytical solutions (6) and (7). However,

values of variances $\delta a_0'$ and $\delta a_1'$, computed by standard method, disagree with exact values determined by (5).

2.3 The main paradox of regression analysis theory

As it emerges from analysis of information presented in Sect. 2.2, *the main* paradox of modern regression analysis theory is exhibited in a contradiction between this theory statements, which guarantee uniqueness of data analysis problems solution, and multivarious solutions in actual practice. In this section we adduce yet several computative manifestations of this paradox.

Example 2. In table 2 a two-factors simulative data array is presented.

Table 2.

Simulative data array

n	x_n	y_n	n	x_n	y_n	n	x_n	y_n
1	–1.0	0.50	8	–0.3	0.75	15	0.4	0.87
2	–0.9	0.55	9	–0.2	0.77	16	0.5	0.89
3	–0.8	0.59	10	–0.1	0.79	17	0.6	0.90
4	–0.7	0.63	11	0.0	0.81	18	0.7	0.91
5	–0.6	0.66	12	0.1	0.83	19	0.8	0.92
6	–0.5	0.69	13	0.2	0.84	20	0.9	0.93
7	–0.4	0.72	14	0.3	0.86	21	1.0	0.94

Let the approximative model have form

$$y = (a_0 + a_1 x + a_2 x^2)/(1 + a_3 x + a_4 x^2)\,. \qquad (7)$$

To find vector parameter **A** estimates of the model (7) on the data array from table 2 we use two different estimation methods. As the first method we choose the estimation one, involved in the software CURVE-2.0, designed AISN. In this case we obtain, that

$$\mathbf{A}' = \{0.81;\ 0.008;\ -0.31;\ -0.22;\ -0.24\}.$$

As the second estimation method we select Marquardt method[12]. Using the value $\mathbf{A}'$, found above by the first estimation method, as initial value of **A** we obtain that in the second case

$$\mathbf{A}' = \{0.81;\ 0.55;\ 0.035;\ 0.45;\ 0.34\}.$$

Thus,

values of A', obtained by two different estimation methods, differ from each other.

Example 3. In table 3 yet one two-factors data array is presented. Let us select the model $y = a_1 x + e$ as approximative one and assume, that y is the random variable with the known density function p:

$$p = \exp\{-(y - a_1 x)/(2f(a_2))\}/\sqrt{2\pi f(a_2)}\,, \qquad (8)$$

where $f(a_2) =$ a) a_2; b) $a_2 x^2$; c) $a_2 x$.

We will find estimates of parameters a_1 and a_2 by method of *maximum likelihood*:

$$L = \prod_{n=1}^{N} \exp\{-(y_n - a_1 x_n\} / (2f(a_2)) / \sqrt{2\pi f(a_2)} \quad \Rightarrow \quad \max \tag{9}$$

or $\partial \ln L / \partial a_i = 0$, where symbol "ln" means the natural logarithm.

Table 3.

Two-factors data array

n	x_n	y_n	n	x_n	y_n	n	x_n	y_n
1	5	21	6	11	53	11	16	81
2	6	31	7	11	56	12	19	97
3	8	38	8	12	60	13	20	98
4	8	37	9	14	68	14	22	107
5	10	53	10	15	72	-	-	-

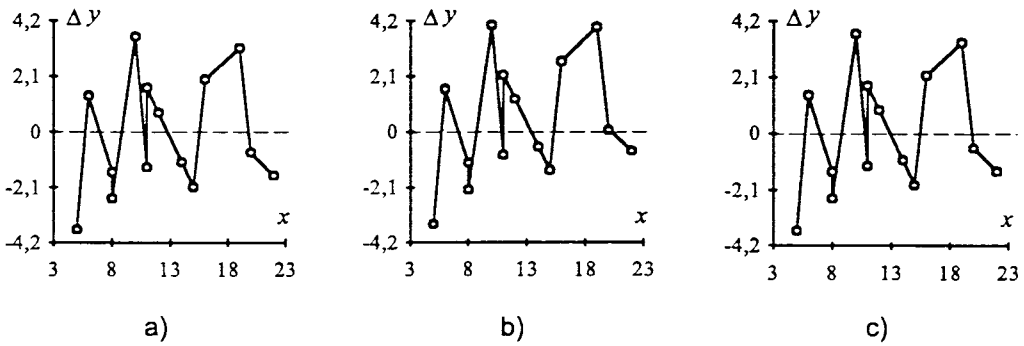

Figure 3. Dependences $\{y_n - a_1' x_n\}$ for different hypothesis about Law for the random variable y variance.

Computation results of V.I.Mudrov and V.P.Kushko[13] show, that in the discussed case the estimates values of parameters a_1 depend on hypothesis about Law for the random variable y variance: for case (a) in (8) $a_1' = 4.938$ ($L' = 4.995 \cdot 10^{-14}$); for case (b) $a_1' = 4.896$ ($L' = 4.421 \cdot 10^{-16}$) and for case (c) $a_1' = 4.927$ ($L' = 9.217 \cdot 10^{-15}$). By analysing obtained results the authors[13] conclude, that, since likelihood function (9) has maximum values for case (a), the more likelihood hypothesis about Law for the random variable y variance is the hypothesis (a): variance of y is the constant value.

We demonstrate in figure 3 that for cases (a), (b) and (c) dependences $\Delta y = \{e_n\} = \{y_n - a_1' x_n\}$ have practically the same form and, consequently,

the strong distinction of values L' for all mentioned cases does not tread on infirm ground.

It should be noted that

– the very apparent expression of the discussed main computative paradox of regression analysis theory one may find also in books[6, 7, 11], where, for the problem on finding the best linear multiple model, fitting Hald data array, a set of solutions,

found by various procedures and statistical tests of modern regression analysis theory, is adduced;

– the most impressive formulation of the main paradox of regression analysis theory is contained in Y.P.Adler introduction [14]:

"When the computation had arisen, the development of regression analysis algorithms went directly «up the stairs, being a descending road». Computer was improving and simultaneously new more advanced algorithms were yielded: whole regression method, step-by-step procedure, stepped method, etc., – it is impossible to name all methods. But again and again it appeared that all these tricks did not allow to obtain a correct solution. At least it became clear that in majority cases the regression problems belonged to a type of incorrect stated problems. Therefore either they can be regularised by exogenous information, or one must put up with ambiguous, multivarious solutions. So the regression analysis degraded ingloriously to the level of a heuristic method, in which the residual analysis and common sense of interpreter play the leading role. Automation of regression analysis problems came to a dead-lock".

3 Data analysis problems at unknown theoretical models

Let us assume, that a researcher is to carry out a quantitative analysis of a data array $\{\mathbf{X}_n\}$ in the absence of theoretical models. Further consideration will be based on the fact [6, 7] that the described situation demands a solution of following problems

– verification of the presence (or absence) of interconnections between analysed properties or phenomena;

– determining (in the case when the interconnection is obvious, a priori and logically plausible) in what force this interconnection is exhibited in comparison with other factors affecting the discussed phenomena;

– drawing a conclusion about the presence of a reliable difference between the selected groups of analysed objects;

– revealing object's characteristics irrelevant to analysed property or phenomenon;

– constructing a regression model describing interconnections between analysed properties or phenomena.

In following sections we consider some methods allowing to solve foregoing problems.

3.1 Correlation analysis

When one is to carry out a quantitative analysis of the data array $\{\mathbf{X}_n\}$ in the absence of theoretical models, it is usual to apply correlation analysis at the earlier

investigation stage, allowing to determine the structure and force of the connections between analysed variables[15-17].

Let, for instance, in an experiment each n-th state of the object be characterised by a pair of its parameters y and x. If relationship between y and x is unknown, it is sometimes possible to establish the existence and nature of their connection by means of such simple way as graphical. Indeed, for realising this way, it is sufficient to construct a plot of the dependence $\{y_n, x_n\}$ in rectangular coordinates $y - x$. In this case the plotted points determine a certain *correlation field*, demonstrating dependences $x = x(y)$ and/or $y = y(x)$ in a visual form.

To characterise the connection between y and x quantitatively one may use *the correlation coefficient* R, determined by the equation[15-17]

$$R_{yx} = \frac{\sum_{n=1}^{N}(y_n - \bar{y})(x_n - \bar{x})}{\sqrt{\sum_{n=1}^{N}(y_n - \bar{y})^2 \sum_{n=1}^{N}(x_n - \bar{x})^2}}. \qquad (10)$$

where $\bar{y}$ and $\bar{x}$ are the mean values of parameters y and x computed on all N readings of the array $\{y_n, x_n\}$. It can be demonstrated that absolute value of R_{yx} does not exceed a unit: $-1 \le R_{yx} \le 1$.

If variables y and x are connected by a strict linear dependence $y = a_0 + a_1 x$, then $R_{yx} = \pm 1$, where sign of R_{yx} is the same as that of the a_1 parameter. This can follow, for instance, from the fact that, using R_{yx}, one can rewrite the equation for the regression line in the following form[15-17]

$$y = \bar{y} + R_{yx}(S_y/S_x)(x - \bar{x}), \qquad (11)$$

where S_y and S_x are mean-square deviations of variables y and x respectively.

In a general case, when $-1 < R_{yx} < 1$, points $\{y_n, x_n\}$ will tend to approach the line (11) more closely with increasing of $|R_{yx}|$ value. Thus, correlation coefficient (10) characterises a linear dependence of y and x rather than an arbitrary one. To illustrate this statement we present in table 4 the values $R_{yx} = R_{yx}(\alpha)$ for the functional dependence $y = x^{\alpha}$, determined on x-interval [0.5; 5.5] in 11 points uniformly.

Table 4.

The values $R_{yx} = R_{yx}(\alpha)$ for the functional dependence $y=x^{\alpha}$, determined on interval [0.5; 5.5]

a	$R_{yx}(\alpha = -a)$	$R_{yx}(\alpha = a)$	a	$R_{yx}(\alpha = -a)$	$R_{yx}(\alpha = a)$
3.0	–0.570	0.927	1.0	–0.795	1.000
2.5	–0.603	0.951	0.5	–0.880	0.989
2.0	–0.650	0.974	0.0	0.0	0.0
1.5	–0.715	0.992	-	-	-

Let us clear up a question what influence has the presence of outliers in the data array $\{y_n, x_n\}$ on the value of correlation coefficient (10). To perform it let us analyse a data array

$$\{x_n\} = (-4; -3; -2; -1; 0; 10), \qquad (12)$$

$$\{y_n\} = (2.48; 0.73; -0.04; -1.44; -1.32; 0),$$

where, on simulation conditions[8], the reading with number 6 is *an extremal* outlier (such reading that contrasts sharply from others); approximative function $F(\mathbf{A}, \mathbf{X}) = a_0 + a_1 x$ and $\mathbf{A}_{\text{true}} = (-2; -1)$.

By computing the values of R_{yx} of (10) and s of (1), we determine the number i of a reading, which elimination from this data array leads to the maximum absolute value of R_{yx} and, consequently, to the minimum value of s (the most simple combinatoric-parametric *procedure Ps*, allowing to find one outlier[6,7] in a data array). Our calculations show that the desirable value $R_{yx} = -0.979$ and $i = 1$. We note, if extremal outlier y_6 is removed from the array (12), $R_{yx} = -0.960$, but s of (1) takes the minimum value. Presented results enable us to state that

procedure Ps loses its effectiveness when revealing the outlier is made not by test s, but by test R_{yx}.

Let us consider another case. For the array (12) the noise array $\{e_n\} = \{-2 - x_n - y_n\} = (-0.48;\ 0.27;\ 0.04;\ 0.44;\ -0.68;\ -12.0)$. We reduce by half the first 5 magnitudes of the noise array $\{e_n\}$: $\{e_n\}_{\text{new}} = (-0.24;\ 0.14;\ 0.02;\ 0.22;\ -0.34;\ -12.0)$; form a new array $\{y_n\}_{\text{new}} = \{-2 - x_n - (e_n)_{\text{new}}\}$ and determine again the number i of a reading, which elimination from the data array $\{y_n, x_n\}_{\text{new}}$ leads to the maximum absolute value of R_{yx}. In the described case R_{yx} reaches its maximum absolute value when the reading 6 (extremal outlier) is deleted from the array $\{y_n, x_n\}_{\text{new}}$ ($R_{yx} = -0.989$). If from the array $\{y_n, x_n\}_{\text{new}}$ we eliminate the reading 6, identified correctly by the test "the maximum absolute value of R_{yx}", then by this test we are able to identify correctly the sequent outlier (the reading 5) in the discussed array. Thus, we obtain finally

when a dependence between the analysed variables is to a certain extent close to a linear, one may use the correlation coefficient (10) for revealing outliers, presented in data arrays.

It is known[15-17], when the number of analysed variables $K > 2$, the structure and force of the connections between variables $x_1, x_2, \ldots, x_K$ are determined by computing all possible pairs of correlation coefficients $R_{x_i x_j}$ from (10). In this case all coefficients $R_{x_i x_j}$ are usually presented in the form of a square symmetric K by K matrix:

$$\mathbf{R} = \begin{bmatrix} R_{11} & R_{12} & \ldots & R_{1K} \\ \ldots & \ldots & \ldots & \ldots \\ R_{K1} & R_{K2} & \ldots & R_{KK} \end{bmatrix}, \qquad (13)$$

which is called *a correlation matrix* (we note that in this matrix diagonal elements $R_{ii} = 1$). Finding strong-interconnected pairs of variables $x_1, x_2, \ldots, x_K$ on the magnitudes of coefficients R_{ij} from the matrix **R** is a traditional use of matrix (13) in data analysis. But, obviously,

using the mentioned way, one should bear in mind all ideas presented above in outline concerning the correlation coefficient (10).

3.2 Discriminant analysis

Let a certain object W be characterised by a value of its vector parameter $\mathbf{X}_w = (x_1, x_2, \ldots, x_K)$; $W_1, W_2, \ldots, W_p$ be p classes and the object W must be ranged in a class W_j on the value of its vector parameter $\mathbf{X}_w$. In discriminant analysis the formulated problem is the *main* one [18-21].

The accepted technique for solving the mentioned problem entails construction of *a discriminant function* $D(\mathbf{A}, \mathbf{X})$. A form and coefficients $\{a_i\}_{i=1,2,\ldots,p}$ values of this function are determined from the requirement, that values of $D(\mathbf{A}, \mathbf{X})$ must have maximum dissimilarity, if parameters of objects, belonging to different populations $W_1, W_2, \ldots, W_p$, are used as arguments of this function.

It seems obvious that in a general case, firstly, $D(\mathbf{A}, \mathbf{X})$ may be either linear or non-linear function on $\{a_i\}$ and, secondly, must be some connection between the problem-solving techniques of discriminant and regression analyses. In particular, as stated [18-21], for solving problems of discriminant analysis one may use standard algorithms and programs of regression analysis. Thus, the similarity of techniques, used for solving problems of the regression and discriminant analyses, makes it possible in discriminant analysis to apply alternative algorithms and procedures of regression analysis and, consequently,

if data analysis problems are solved by discriminant analysis techniques then in practice the researcher may meet the same difficulties which are discussed in Sect. 2.

3.3 Regression analysis

In the absence of theoretical models it is usual to employ regression analysis in order to express in a mathematical form the connections existing between variables under analysis.

It happens with extreme frequency that researchers impose limitations on a type and form of approximative models or, in other words, approximative models are often chosen from a given set of ones. Evidently, in this case it is required to solve problem on finding the best approximative model from a given set of models. Let, for instance, it is required to find the best approximative multinomial with a minimal degree. With this in mind, in two examples below we consider some accepted

techniques, used for solving the mentioned problem in approximation and/or regression analysis theories.

Example 4. Let

$$\{x_n\} = \{-1+0.2(n-1)\};\ f(x) = \sin x;\ \{y_n\} = \{\,[\,kf(x_n)\,]/k\,\}, \qquad (14)$$

where square brackets mean the integer part; $n = 1, 2, \ldots, 11$; $f(x)$ is a given function, used for generating the array $\{y_n\}$; a factor $k = 10^3$ and its presence in (14) is necessary for "measuring" all values of y_n within error $\varepsilon = 10^{-3}$. It is required for the presented dependence $\{y_n, x_n\}$ to find the best approximative multinomial with a minimal degree.

A. As well known in approximation theory[22], if the type of the function $f(x)$ is given and either $(m+1)$-derivative of this function weakly varies on the realisation $\{x_n\}$, or on the x-interval $[-1, 1]$ the function $f(x)$ is presented in form of even-converging power series, then the problem of finding the best *even-approximating* multinomial for the discrete dependence $\{y_n, x_n\}$ offers no difficulty. Indeed, in the first case the solution of problem is an interpolative multinomial $P_M(\mathbf{B}, x)$ with a set of Chebyshev points (this multinomial is close to the best even-approximating one). In the second, case one may obtain the solution by the following *economical procedure* of an even-converging power series:

1. Choose the initial part of truncated Taylor series, approximating the function $f(x)$ within error $\varepsilon_M < \varepsilon$, as the multinomial $P_M(\mathbf{B}, x)$ (the multinomial with the degree M and vector parameter $\mathbf{B}$);

2. Replace ε_M with $\varepsilon_M - |b_M| / 2^{M-1}$, where b_M is a coefficient of the multinomial $P_M(\mathbf{B}, x)$ at x^M;

3. If $\varepsilon_M > 0$, then replace the multinomial $P_M(\mathbf{B}, x)$ with the multinomial

$$P_{M-1}(x) = P_M(x) - (b_M / 2^{M-1}) T_M(x), \qquad (15)$$

where $T_M(x)$ is Chebyshev multinomial: $T_0 = 1$, $T_1 = x$ and when $M \geq 2$ $T_M = 2xT_{M-1} - T_{M-2}$. Then decrement M by one and go to point 2. If $\varepsilon_M \leq 0$, then go to point 4;

4. End of computations: the multinomial $P_M(\mathbf{B}, x)$ is the desirable one.

By means of the foregoing economical procedure one may easy obtain that the multinomial with a minimal degree, even-approximating the function $\sin x$, given within error $\varepsilon = 10^{-3}$ on the x-interval $[-1, 1]$, has the following form

$$P_3(x) = (383/384)\,x - (5/32)\,x^3. \qquad (16)$$

B. Let $1 \leq M \leq 9$ and in the multinomial $P_\lambda(\mathbf{B}, x) = \sum_{m=0}^{M} \lambda_m b_m x^m$ all $\lambda_m = 1$. By determining LS-estimates of vector parameter $\mathbf{B}$ for each value of M on the formed above array $\{y_n, x_n\}$, we find that in all obtained approximative multinomials $P_\lambda(\mathbf{B}', x)$, as well as in Taylor series of function $\sin x$, the values of coefficients

$b'_{2l} = 0$ at $l = 0, 1, \ldots, 4$. Thus, regression analysis of the discussed array $\{y_n, x_n\}$ allows to determine a form of the best approximative multinomial:

$$P_{2l+1\,\mathrm{opt}}(\mathbf{B}, x) = b_1 x + b_3 x^3 + \ldots + b_{2l+1} x^{2l+1} + \ldots . \qquad (17)$$

We note, if $l = 1$, then the values of parameters b_1 and b_3, computed by regression analysis method (LS-method) for model (17), coincide with ones, shown in (16).

We remind, that in classical variant of regression analysis theory the best approximative multinomial is chosen by the minimal value of the test $s_\lambda = S_\lambda/(N-I_\lambda)$, where S_λ is residual sum-of-squares, $I_\lambda = \sum_{m=1}^{M} \lambda_m$; N is total number of readings; λ_m is such characteristic number that $\lambda_m = 0$, if the approximative multinomial $P_\lambda(\mathbf{B}, x)$ does not contain term $b_m x^m$, and $\lambda_m = 1$, otherwise. For approximative models (17) and $l = 0, 1, 2, 3$ the computation values of s are following

l	0	1	2	3
s	0.033	0.00063	0.00043	0.00054

Since s has the minimal value at $l = 2$, for the discussed array in the frame of classical variant of regression analysis theory, the multinomial

$$P_5(x) = b_1 x + b_3 x^3 + b_5 x^5 \qquad (18)$$

is the best approximative one.

C. Since, for each n in the data array (14), the difference of $|y_n - y_n'|$ must be kept within the limit of the error $\varepsilon = 10^{-3}$, the general solution of the discussed problem has the following form[6, 7]

$$y'(\xi, x) = g_\alpha\{(1.0012 - 0.0001\xi)\,x - (0.161200 - 0.000127\xi)\,x^3\}, \qquad (19)$$

where $\alpha = 0.001$; $0 \le \xi \le 49$ and, consequently, $b_1 = 0.9987 \pm 0.0025$ and $b_2 = -0.1582 \pm 0.0030$.

By analysing solutions (16), (18) and (19) we conclude that in the considered case

the solution of the problem on finding the best fitting multinomial depends on the type of the used mathematical theory.

Example 5. In some software products {for instance, in the different versions of software CURVE, designed by AISN} the solutions of problems on finding the best approximative models are found by the magnitude of *a determination coefficient R*, which value may be computed by a set of formulae

$$R_1 = \sqrt{1 - Q_r / Q},\ Q_r = \sum_{n=1}^{N} (y_n - y_n')^2,\ Q = \sum_{n=1}^{N} (y_n - \bar{y})^2 \qquad (20)$$

where $\bar{y} = \sum_{n=1}^{N} y_n / N$, y_n is n-th reading of dependent variable, y'_n is n-th value of dependent variable, computed on the fitting model; or by formula (firstly offered by K.Pearson)

$$R_2 = \frac{\sum_{n=1}^{N}(y_n - \bar{y})(y'_n - \bar{y}')}{\sqrt{\sum_{n=1}^{N}(y_n - \bar{y})^2 \sum_{n=1}^{N}(y_n - \bar{y}')^2}} \tag{21}$$

{evidently, one may easy obtain formula (21) from formula (10)}.

Table 5.

The simulative data array to example 5

n	y_n	x_n	n	y_n	x_n
1	85	11	7	205	3.8136396
2	105	5.6002132	8	225	3.5774037
3	125	5.0984022	9	245	3.4193292
4	145	4.7047836	10	265	3.2903451
5	165	4.3936608	11	285	3.1026802
6	185	4.0998636	-	-	-

There is a mathematical proof[23] of the equivalence of formulae (20) and (21). But, if the value of coefficient R_2^2 is computed within error $\geq 10^{-8}$,

in actual practice, for some data arrays, firstly, $R_2^2 \neq R_1^2$ and, secondly, $R_2^2 > 1$.

For instance, if one fits the simulative data array, presented in table 5, by the multinomial $P_M(\mathbf{B}, x)$ with $M = 8$, then software CURVE-2.0 will give the value $R_2^2 = 1.00040$.

4 Problems of quantitative processing experimental dependences found for heterogeneous objects

As it follows from the general consideration[24, 25], in practice at analysis of experimental dependences found for heterogeneous objects, three various situations can be realised: the heterogeneity of investigated objects causes a) no effect; b) a removable (local) inadequacy of postulated fitting model; c) an irremovable (global) inadequacy of the postulated model. In this section we discuss some computative difficulties which may occur at analysis of the mentioned experimental dependences.

Example 6. As we know from Sect. 2.1 if $F(\mathbf{A}, x)$ is a linear model $\{F(\mathbf{A}, x) = \sum_{l=1}^{L} a_l h_l(x)\}$ then the value of $\mathbf{A}'$, minimising residual sum-of-squares S, is

computed by (1). Let rank $\mathbf{H} < L$, or, in other words, there is a linear dependence between columns of matrix $\mathbf{H}$:

$$c_1 h_1 + c_2 h_2 + \ldots + c_L h_L = 0, \qquad (22)$$

where at least one coefficient $c_l \neq 0$. In this case matrix $(\mathbf{H}^T\mathbf{H})^{-1}$ does not exist, that means one cannot find $\mathbf{A}'$ from (1). Such situation is known as *strict multicollinearity.*

In the natural science investigations values of the independent variable $\mathbf{X}$ are always determined with a certain round-off error, although this error may be very small. Therefore, if even strict multicollinearity is present, in practice the equation (22) is satisfied only approximately and therefore rank $\mathbf{H} = L$. In such situation application of equation (1) to find the estimate of vector parameter $\mathbf{A}$ gives $\mathbf{A}'$ values drastically deviating from true coefficients values[6, 7, 23].

To correct this situation in regression on *characteristics roots*[26] it is suggested to obtain the information about the grade of matrix $\mathbf{H}^T\mathbf{H}$ conditioning from values of its eigennumbers λ_j and first elements V_{0j} of its eigenvector $\mathbf{V}_j$ and to exclude from regression such j-components, whose eigennumbers λ_j and elements V_{0j} are small. Following values are recommended to use as critical ones: $\lambda_{cr} = 0.05$ and $V_{cr} = 0.1$.

Let us demonstrate, that

in some practical computations the difference between A'_{CHR} and A'_{LS} of (1) can be explained not by the effects of multicollinearity, but by regression model inadequacy, which disappears simultaneously with the effects of multicollinearity after removing outliers.

Indeed, let data array be following

$$\{y_n, \mathbf{X}_n\} = \{1 + 0.5\, n + 0.05\, n^2 + 0.005\, n^3;\ n, n^2, n^3\}, \qquad (23)$$

$n = 1, 2, \ldots, 11$ and we introduce two outliers in (23), by means of increasing values y_3 and y_8 on 0.5. For this data array we obtain the following computation results:

$N_{step} = 1, N = 11$: $\mathbf{A}'_{LS} = (1.017; 0.542; 0.0462; 0.0050)$,

$n_a = \{3, 8\}$ $\mathbf{A}'_{CHR} = (1.046; 0.516; 0.0516; 0.0047)$;

$N_{step} = 2, N = 10$: $\mathbf{A}'_{LS} = (0.764; 0.801; -0.0169; 0.0090)$,

$n_a = \{3\}$ $\mathbf{A}'_{CHR} = (0.764; 0.801; -0.0169; 0.0090)$,

where N_{step} is a number of the step in used computative procedure; n_a is a vector to indicate the numbers of anomalous readings, contained in analysing array on the first and second steps of used computative procedure; N is the general quantity of analysing readings. In particular, after the first step of computative procedure from (23) the reading with number 8 is removed; after the second step — readings with numbers 3 and 8. And after the second step the values of $\mathbf{A}'$ are restored without any distortion by both examined algorithms.

By analysing obtained computation results one can conclude that

the difference between A'_{LS} and A'_{CHR} may be caused not only by multicollinearity but also, for instance, by a set of outliers presented in the data array.

Example 7. In table 6 we adduce an experimental data array, obtained by N.P.Bobrysheva[27], when she investigated magnetic susceptibility (χ, relative units) of polycrystalline system $V_xAl_{1-x}O_{1,5}$ (x = 0.078) on the temperature (T, K). Let us consider some computation results of quantitative processing this temperature dependence.

Table 6.

The experimental dependence of magnetic susceptibility of system $V_xAl_{1-x}O_{1,5}$ (x = 0.078) on temperature

n	T_n	χ_n	n	T_n	χ_n	n	T_n	χ_n
1	80	10.97	9	292	3.79	17	501	2.48
2	121	8.06	10	351	3.31	18	512	2.46
3	144	6.94	11	360	3.20	19	523	2.42
4	182	5.56	12	385	3.06	20	559	2.28
5	202	5.11	13	401	3.00	21	601	2.12
6	214	4.75	14	438	2.78	22	651	2.02
7	220	4.62	15	464	2.67	23	668	1.97
8	267	4.00	16	486	2.56	-	-	-

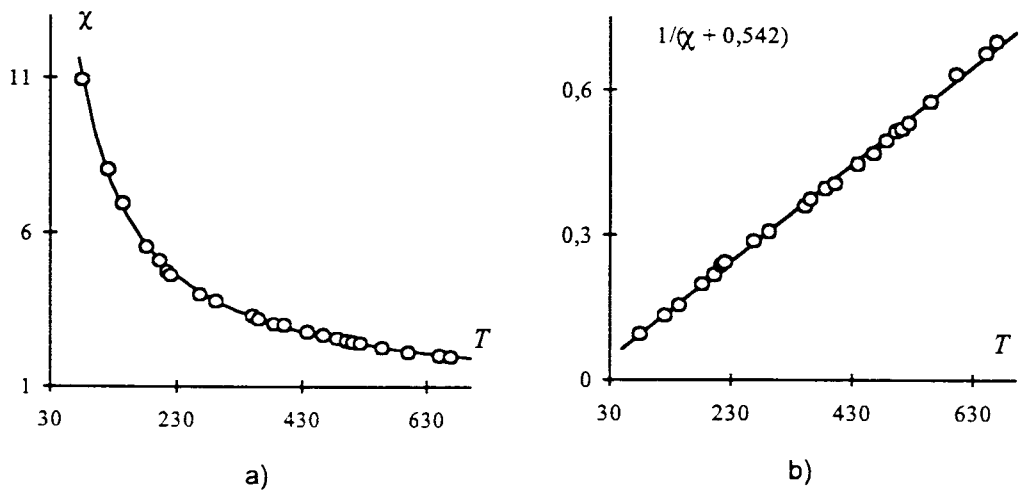

Figure 4. Experimental (circles) and analytical (continuous curves) plots of dependences $\chi(T)$ (a) and $1/(\chi + \chi_2) - T$ (b) for system $V_xAl_{1-x}O_{1,5}$ (x = 0.078).

A. In figure 4(*a*, *b*) for the discussed system experimental (circles) and analytical (continuous curves) plots of dependences $\chi(T)$ and $1/(\chi + \chi_2) - T$ are shown. For construction of analytical (continuous) curves we use modified Curie – Weiss law[6, 7, 24, 25]

$$\chi = \chi_0 + C/(T + \theta), \qquad (24)$$

where χ is the experimental magnitude of specific magnetic susceptibility; T is absolute temperature, K; C, θ and χ_0 are parameters: $C = 988$; $\theta = 14$, K; $\chi_0 = 0.54$. From analysing graphical information, presented in figure 4(*a*, *b*), one can conclude, that

the magnetic behaviour of system $V_xAl_{1-x}O_{1.5}$ *(x = 0.078) is well explained by the modified Curie – Weiss law (24).*

B. In figure 5(*a*) the dependence $\Delta\chi = \chi - C/(T + \theta) - \chi_0$ on T for system $V_xAl_{1-x}O_{1,5}$ (x = 0.078) is shown. Since $\Delta\chi_{max} \equiv 0.2 >> \varepsilon = 0.01$, where ε is the measurement error in the diccussed experiment, we obtain

in contradiction with the statement of point (A) in this case modified Curie – Weiss law (24) is an inadequate approximative model or, in other words, there is a set of outliers in the analysing experimental dependence.

C. After deleting first 5 readings from the initial data array the parameters values of modified Curie – Weiss law (24) have magnitudes $C = 1386$; $\theta = 89$, K; $\chi_0 = 0.14$. The plot of dependence $\Delta\chi = \chi - C/(T + \theta) - \chi_0$ with foregoing parameters values is shown in figure 5(*b*).

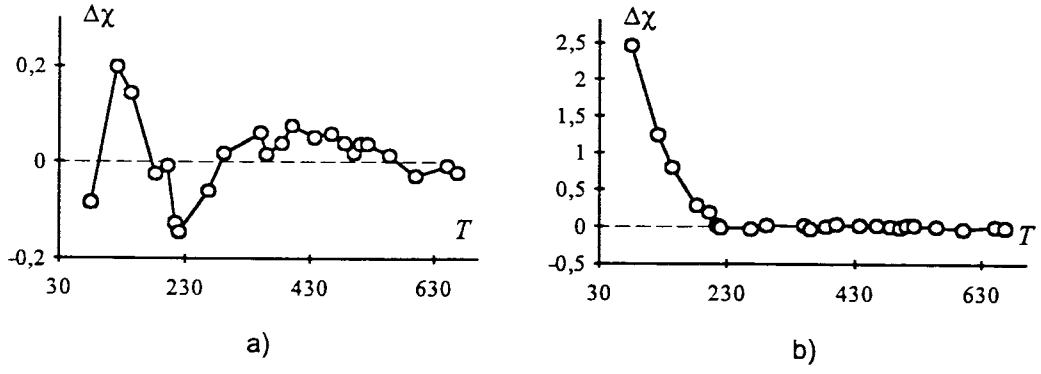

Figure 5. Plots of $\Delta\chi = \chi - C/(T + \theta) - \chi_0$ on T for system $V_xAl_{1-x}O_{1.5}$ (x = 0.078).

Analysing plots $\Delta\chi(T)$, presented in figure 5(*a*, *b*), and comparing with each other the parameters values of equation (24), mentioned in points (A) and (C), we conclude, that

neglect of the local inadequacy of the approximative model in the discussed experiment leads to distortion of both form of function $\Delta\chi(T)$ *and parameters values of the modified Curie – Weiss law (24).*

Thus, if, for proving well-fitted properties of equation (24), researchers[27–30] suggest to look at the graphic representation of dependences χ (T) or $1/\chi$ (T), for the proof completeness one should ask these researchers to present information about the measurement error of values χ and plots $\Delta\chi(T) = \chi - C/(T+\theta) - \chi_0$.

For the sake of convenience, the main causes, given rise to computative difficulties at analysis of experimental dependences found for heterogeneous objects, and methods of their overcoming are adduced in table 7 together. In this table all methods, overcoming computative difficulties, are marked by the symbol Θ, if at present they are *in the rough* or absent in modern data analysis theory.

Table 7.

Main causes and overcoming methods of computative difficulties in modern data analysis theory

	Main causes	Methods of overcoming
1.	Impossibility to take heed of preset measurement accuracy of dependent variable values in the frame of accepted data analysis model	Θ Modification of data analysis model
2.	Limited accuracy of computations	Increasing computation accuracy
3.	Point estimation of parameters	Replacing the point estimation of parameters by interval one
4.	The deficient measurement accuracy of dependent variable values	Increasing measurement accuracy of dependent variable values
5.	Ill-conditioning of estimation problem	Θ Using alternative estimations methods; Increasing measurement accuracy of dependent variable values; Θ Revealing and removing outliers; Designing experiments
6.	Presence of outliers in analysing data arrays	Θ Revealing and removing outliers; Θ Robust estimation of parameters
7.	Inadequacy of approximative model	Θ Eliminating inadequacy of approximative model; Θ Using advanced estimations methods
8.	Finding only single solution of the estimation problems for contaminated data array in the frame of modern data analysis theories	Θ Finding a family of solutions

Using information presented in table 7, let us clear up a question, whether one is able in the frame of modern data analysis theory to obtain reliable solutions for the problems of quantitative processing of experimental dependences, found for heterogeneous objects.

Let, when an investigated object is homogeneous, a connection between characteristics y and $\mathbf{X}$ exist and it be close to functional one: $y = F(\mathbf{A}, \mathbf{X})$. As we said already in beginning of this section, in the discussed experiments three various situations can be realised: the structural heterogeneity

1) has no effect on the experimental dependence $\{y_n, \mathbf{X}_n\}$ or, in other words, in this case it is impossible to distinguish the homogeneous objects from heterogeneous ones on the dependence $\{y_n, \mathbf{X}_n\}$;

2) leads to a distortion of the dependence $\{y_n, \mathbf{X}_n\}$ in some small region $\{\mathbf{X}_{n_1}\} \subset \{\mathbf{X}_n\}$ {the approximative model $F(\mathbf{A}, \mathbf{X})$ has *removable* (local) inadequacy}. In this case for extracting effects, connected with the presence of a homogeneity in the investigated objects, one may use the following way[6, 7]

i) solve the problem on revealing outliers $\{y_{n_1}, \mathbf{X}_{n_1}\}$;

ii) determine the value $\mathbf{A}'$ on readings $\{y_n, \mathbf{X}_n\} \setminus \{y_{n_1}, \mathbf{X}_{n_1}\}$ {we remind, that a set of $\{y_n, \mathbf{X}_n\} \setminus \{y_{n_1}, \mathbf{X}_{n_1}\}$ is to be well-fitted by the model $F(\mathbf{A}, \mathbf{X})$};

iii) detect a type and degree of the effects, connected with the presence of a homogeneity in the investigated objects, on the data array $\{y_{n_1} - F(\mathbf{A}', \mathbf{X}_{n_1}),\ \mathbf{X}_{n_1}\}$.

It follows from point 6 of table 7, that at solving problem (*i*) in actual practice some difficulties, which are unsurmountable in the frame of modern regression analysis theory, can be arisen;

3) leads to a distortion of the dependence $\{y_n, \mathbf{X}_n\}$ in a big region $\{\mathbf{X}_{n_1}\} \subseteq \{\mathbf{X}_n\}$: {the approximative model $F(\mathbf{A}, \mathbf{X})$ has *irremovable* (global) inadequacy}.

It follows from point 7 of table 7, that in this case it is impossible to find a reliable solution of the discussed problem in the frame of modern data analysis theory.

Summarising mentioned in points (1) – (3), we conclude

since at present the methods, marked by the symbol Θ in table 7, are not effective for overcoming computative difficulties or absent in modern data analysis theory, one is not able to obtain reliable solutions for the problems of quantitative processing of experimental dependences found for heterogeneous objects.

From our point of view, one of possible ways, overcoming computative difficulties in modern data analysis theory, is further development of this theory by means of translation of this theory from Aristotelian "binary logic" into more progressive "fuzzy logic"[6, 7, 24, 25, 31, 32].

References

1. Dumitrescu C and Seleacu V *Some notions and questions in number theory* (Erhus University Press, Vail, 1995)
2. Smarandache F *Paradoxist Mathematics* (Pennsylvania, 1985)
3. Linnik Y V *Least squares method and foundations of mathematical-statistical theory of observations processing* (Moscow, Publ. H. of phys.-math. lit., 1958) (In Russian)
4. Rao C P *Linear statistical inference and its applications.* 2nd ed. (New York, Wiley, 1973)
5. Ermakov S M and Zhiglavskiy A A *Mathematical theory of optimal experiment* (Moscow, Nauka, 1987) (In Russian)
6. Chebrakov Y V *The parameters estimation theory in the measuring experiments* (S.–Petersburg, S.–Petersburg State Univ. Press, 1997) (In Russian)

7. Chebrakov Y V and Shmagin V V *Regression data analysis for physicists and chemists* (S.–Petersburg, S.–Petersburg State Univ. Press, 1998)
8. Huber P J *Robust Statistics* (New York, Wiley, 1981)
9. Rousseeuw P J and Leroy A M *Robust regression and outlier detection* (New York, Wiley, 1987)
10. Andrews D F *Technometrics* (1974) **16** 523
11. Draper K R and Smith H *Applied regression analysis* (New York, Wiley, 1981)
12. Gill P E, Murray W and Wright M H *Practical Optimization* (New York – London, Academic Press, 1981)
13. Mudrov V I and Kushko V P *Methods of processing measurements: Quasi- believable estimates* (Moscow, Radio i svyaz, 1983) (In Russian)
14. Adler Y P *The introduction* to Russian edition of the book: Mosteller F and Tukey J W *Data analysis and regression* (Moscow, Finansy i statistika, 1982) (In Russian)
15. Aivazyan S A, Yenyukov I S and Meshalkin L D *Applied statistics. Analysis of dependences* (Moscow, Finansy i statistika, 1985) (In Russian)
16. Bendat J S and Piersol A G *Random data: Analysis and measurement procedures* (New York, Wiley, 1986)
17. Rozanov Y A *Probability theory, random processes and mathematical statistics: A textbook for higher education.* (Moscow, Nauka, 1989) (In Russian)
18. Afifi A A and Azen S P *Stastistical analysis: A computer-oriented approach.* 2nd ed. (New York – London, Academic Press, 1979)
19. Lloyd E and Lederman W *Handbook of applicable mathematics.* Volume IV: *Statistics.* Parts A and B (New York, Wiley, 1984)
20. Mandel I D *Cluster analysis* (Moscow, Nauka, 1988) (In Russian)
21. Seber G A F *Multivariate observations* (New York, Wiley, 1984)
22. Laurent P J *Approximation et optimisation* (Paris, Hermann, 1972)
23. Vuchkov I, Boyadzieva L and Solakov E *Applied linear regression analysis* (Moscow, Finansy i statistika, 1981) (In Russian)
24. Chebrakov Y V and Shmagin V V *J. Izv. Vysh. Uch. Zav. Fizika* (1995) **5** 40; (1998) **6** 110 (Engl. Transl. *Russian Phys. J.* (1995) **5** 470; (1998) **6**)
25. Chebrakov Y V *J. Izv. Vysh. Uch. Zav. Fizika* (1997) **7** 103 (Engl. Transl. *Russian Phys. J.* (1997) 7 687)
26. Webster J, Gunst R and Mason R *Technometrics* (1974) **16** 513
27. Bobrysheva N P *Magnetic dilution of a series of oxides containing elements of the first transitional period*: Condensed version of the PhD thesis (Leningrad, 1974) (In Russian)
28. Amine K, Tressand A, Imoto H et al. *J. of Solid State Chemistry* (1992) **96** 287
29. Miyata N, Kimishima Y, Akutsu N and Oguro I *J. Magnetism and Magnetic Materials* (1990) **90&91** 337
30. Zayachuk D M, Ivanchuk R D, Kempnik V I and Mikitjuk V I *Fizika tverdogo tela* (1996) **36** № 8 2502 (In Russian)
31. Chebrakov Y V and Shmagin V V *Correct solutions of fit problems in different experimental situations* In: Advanced Mathematical Tools in Metrology III (Singapore, World Scientific, 1997) 246
32. Chebrakov Y V and Shmagin V V *Some nontraditional approaches to solving a few data analysis problems* In: Integral methods in science and engineering (Edinburg, Longman, 1997) **2** 69.