Advances in Data Science and Adaptive Analysis
Vol. 12, No. 1 (2020) 2050002 (34 pages)

DOI: 10.1142/S2424922X20500023

Weak and Strong Compatibility in Data Fitting Problems Under Interval Uncertainty

Sergey P. Shary
Federal Research Center for Information and Computational Technologies
Academician M. A. Lavrentiev Avenue, 6,
630090 Novosibirsk, Russia
Novosibirsk State University,
1, Pirogova Str., 630090 Novosibirk, Russia
shary@ict.nsc.ru

Received 16 February 2019
Accepted 6 March 2019
Published 6 July 2020

In memory of my beloved wife Irene

For the data fitting problem under interval uncertainty, we introduce the concept of *strong compatibility* between data and parameters. It is shown that the new strengthened formulation of the problem reduces to computing and estimating the so-called tolerable solution set for interval systems of equations constructed from the data being processed. We propose a computational technology for constructing a "best-fit" linear function from interval data, taking into account the strong compatibility requirement. The properties of the new data fitting approach are much better than those of its predecessors: strong compatibility estimates have polynomial computational complexity, the variance of the strong compatibility estimates is almost always finite, and these estimates are rubust. An example considered in the concluding part of the paper illustrates some of these features.

Keywords: Data fitting problem; interval uncertainty; compatibility of data and parameters; strong compatibility; interval system of equations; tolerable solution set; recognizing functional; nondifferentiable optimization.

1. Introduction

1.1. *Problem statement*

The work is devoted to the development of methods for analyzing data that are inaccurate and have interval uncertainty. We consider a linear regression model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_m x_m, \tag{1}$$

in which $x_1, x_2, \ldots, x_m$ are independent variables (also called *exogenous*, *explanatory*, *predictor*, or *input* variables), y is a dependent variable (also called *endogenous*, *response*, *criterion*, or *output* variable), and $\beta_0, \beta_1, \ldots, \beta_m$ are some coefficients.

These unknown coefficients should be determined from a number of measurements (observations) of the values $x_1, x_2, \ldots, x_m$ and y.

The measurement results are inaccurate, and we assume that they have *bounded uncertainty* (see Milanese *et al.* [1996]) when we know only some intervals that provide us with two-sided bounds for the exact values of the measured quantities. Therefore, the ith measurement results in such intervals $\boldsymbol{x}_1^{(i)}, \boldsymbol{x}_2^{(i)}, \ldots, \boldsymbol{x}_m^{(i)}, \boldsymbol{y}^{(i)}$ that the actual value of x_1 is within $\boldsymbol{x}_1^{(i)}$, the actual value of x_2 is within $\boldsymbol{x}_2^{(i)}$, and so on, up to y, with its actual value being within $\boldsymbol{y}^{(i)}$.

In total, there are n measurements, so that the index i can take values from the set $\{1, 2, \ldots, n\}$. We need to find or somehow estimate the coefficients β_j, $j = 0, 1, \ldots, m$, for which the linear function (1) would "best approximate" the data. The ideal is, of course, the case when the graph of the constructed function (1) "passes through all measurement points", i.e. when the approximation of the data is indeed complete, in exactly the same way as, for example, in the interpolation.

1.2. *Main ideas and results of the work*

In the case when the data are inaccurate, when each measurement or observation represents an entire set of possible values rather than a single point, the very concept of "passing through measurement points" must be rethought. The fact is that now the sets of measurement uncertainty acquire a structure that makes it necessary to distinguish between different cases of passing a function graph through these sets. This is caused, in particular, by the inputs and outputs of the system (corresponding to independent arguments of the function and the dependent variables) differing from each other in their purpose. In addition, measurements of inputs and outputs can be performed in various ways, by various instruments and techniques, or even at different points in time.

In order to take into account these new realities, we introduce the concepts of *weak compatibility* and *strong compatibility* of data and parameters of the functional dependence. All parameters having weak compatibility with the data forms a set, which is known in interval analysis as the *united solution set* for an interval system of equations constructed from interval measurement data. On the other hand, the set of model parameters that satisfy the strong compatibility conditions is the so-called *tolerable solution set* for an interval system of equations constructed from interval measurement data.

The tolerable solution sets for interval systems of linear algebraic equations is relatively well studied. It is always a convex polyhedral set and there exist practical methods for recognizing whether it is empty or nonempty, as well as for its inner and outer estimations. Moreover, recognizing whether the tolerable solution set for an interval linear equation system is empty or nonempty is a polynomially complex problem, whereas recognition of the united solution set is NP-hard. As a consequence, the properties of the new data fitting approach are much better than those of its predecessors: strong compatibility estimates have polynomial computational complexity. Additionally, the variance of the strong compatibility estimates is almost always finite, and these estimates are rubust.

In our work, we discuss practical methods for the solution of the data fitting problem under the strong compatibility requirement. Our main tool is a technique that uses the so-called recognizing functional of the tolerable solution set to the interval system of linear equations constructed from the measurement data.

Although we study in detail the situation, when all the measurements are subject to the same compatibility conditions, the most general case in processing interval data is that some measurements with strong compatibility are combined with those where the usual weak compatibility takes place. Then the data fitting problem becomes even more complicated, and its analysis makes it necessary to consider the so-called AE solutions and AE solution sets for interval systems of equations. The corresponding mathematical theory, in fact, has already been developed, and there are computational methods for solving problems of recognition and estimation of the AE solution sets (see e.g. Sharaya and Shary [2016] and Shary [2002]). We postpone the detailed exposition of these results until future publications.

This work continues and supplements an earlier article [Shary (2017)], and our notation system corresponds to the informal international standard [Kearfott *et al.* (2010)]. In particular, intervals and interval objects are throughout indicated in bold type, while noninterval (point) values, quantities, and variables are not designated in any special way.

2. Data Fitting Under Interval Uncertainty

2.1. *Short review*

The data fitting problem is a popular and practically important problem, in which we are required to construct, according to empirical data, a functional dependence of a given type between "input" and "output" quantities. In our work, we consider in detail the simplest linear function of the form

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_m x_m, \tag{2}$$

although many constructions and conclusions are also valid in the general nonlinear case. It is necessary to determine the unknown coefficients β_i so that the resulting linear function "best-fits" a given set of values of the independent arguments and dependent variable,

$$\begin{matrix} x_1^{(1)}, & x_2^{(1)}, & \ldots, & x_m^{(1)}, & y^{(1)}, \\ x_1^{(2)}, & x_2^{(2)}, & \ldots, & x_m^{(2)}, & y^{(2)}, \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_1^{(n)}, & x_2^{(n)}, & \ldots, & x_m^{(n)}, & y^{(n)}. \end{matrix} \tag{3}$$

The above problem is often referred to as "linear regression problem" in statistics or as "parameter identification problem" in engineering language.

Substituting data (3) in equality (2), we obtain, after renaming $x_{ij} := x_j^{(i)}$ and $y_i := y^{(i)}$, the system of equations

$$\begin{cases} \beta_0 + x_{11}\beta_1 + \cdots + x_{1n}\beta_m = y_1, \\ \beta_0 + x_{21}\beta_1 + \cdots + x_{2n}\beta_m = y_2, \\ \quad\vdots \qquad\quad \vdots \qquad\quad \ddots \qquad\quad \vdots \qquad\quad \vdots \\ \beta_0 + x_{n1}\beta_1 + \cdots + x_{nm}\beta_m = y_n, \end{cases} \tag{4}$$

with the unknowns $\beta_0, \beta_1, \dots, \beta_m$, or briefly

$$X\beta = y, \tag{5}$$

with $n \times (m+1)$-matrix $X = (x_{ij})$, $(m+1)$-vector $\beta = (\beta_i)$, and n-vector $y = (y_i)$ such that

$$X = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1m} \\ 1 & x_{21} & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{nm} \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_m \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

(where the columns of the matrix X are, apparently, more convenient to be numbered from zero). A solution to systems (4) and (5), either ordinary or in a generalized sense, is taken as an estimate of the parameters $\beta_0, \beta_1, \dots, \beta_m$. A graphical illustration of the data fitting problem is shown in traditional Fig. 1: we have to find a straight line that "best approximates" the set of points with the coordinates (3).

In the practical data fitting problems, the data is almost always inaccurate, since the results of measurements and observations are influenced by external uncontrolled factors, the measuring devices themselves are not absolutely accurate, etc. Thus, in reality, we must deal with this or that *uncertainty* — the state of partial knowledge of the measured quantity, when we know some value, but it is approximate, and there is also some information (qualitative or quantitative) about the error of this value.

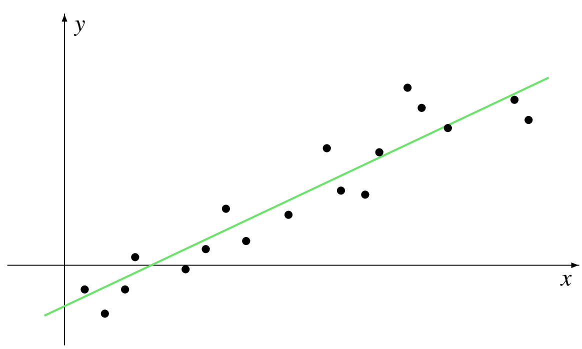

Fig. 1. Illustration of the data fitting problem.

How to describe this uncertainty? In other words, what "uncertainty model" do we accept for the data? The traditional choice is a probabilistic model of errors, the foundations of which were laid at the turn of the eighteenth and nineteenth centuries by C. F. Gauss and P.-S. Laplace. According to this approach, the errors in measurements and observations are random quantities that can be adequately described by the mathematical probability theory, and we (more or less) know the characteristics of these random variables. Over the past two centuries, the probabilistic model of measurement errors has been intensively developed by many outstanding mathematicians and statisticians. It has become very popular, turning into the main tool for data processing. Nevertheless, the application of this model puts a lot of nontrivial questions for both engineers and mathematicians, the answers to which are sometimes not entirely satisfactory.

In general, if the probabilistic description of the measurement errors is inadequate, it is often more convenient to work with uncertainties and inaccuracies in the data using interval analysis methods. In this approach, we suppose that interval estimates of the measurement results are given instead of probabilistic distributions, i.e. we know the smallest and largest bounds of possible values of the quantities of interest. In our data fitting problem, it is assumed that interval estimates are given for x_{ij} and y_i:

$$x_{ij} \in \boldsymbol{x}_{ij} = [\underline{\boldsymbol{x}}_{ij}, \overline{\boldsymbol{x}}_{ij}] \quad \text{and} \quad y_i \in \boldsymbol{y}_i = [\underline{\boldsymbol{y}}_i, \overline{\boldsymbol{y}}_i].$$

The pioneer of the new approach to data processing was Kantorovich, who, in 1962, first articulated the above principles and briefly outlined the formulation of the new problem and some methods for solving it in an article [Kantorovich (1962)]. The first Western article on this topic was authored by Schweppe [1968]. Later, a significant contribution to the development of the theory was made by many researchers, and the interested reader can find the necessary information on the current state of this area e.g. in the literature [Combettes (1993); Jaulin *et al.* (2001); Milanese *et al.* (1996); Polyak and Nazin (2006); Zhilin (2005, 2007)] (see also the references in the above articles). Our publications [Shary (2012, 2016); Shary and Sharaya (2013)], which develop the so-called maximum compatibility method, are devoted to this same problem.

2.2. *Definition of compatibility between parameters and data*

In the formulation of Kantorovich and his followers, the data fitting problem under interval uncertainty did not cover the most general case: the inaccuracies in the input data were absent, i.e. it was supposed that $\boldsymbol{x}_{ij} = x_{ij}$. Then, for the linear function (2), there should be

$$\underline{\boldsymbol{y}}_i \le \beta_0 + \beta_1 x_{i1} + \cdots + \beta_m x_{im} \le \overline{\boldsymbol{y}}_i, \tag{6}$$

$i = 1, 2, \ldots, n$. The compatibility of parameters and data was understood as the passage of the graph of the constructed functional dependence, i.e. of a straight line, through all the corridors of data uncertainty for the output variable (see Fig. 2).

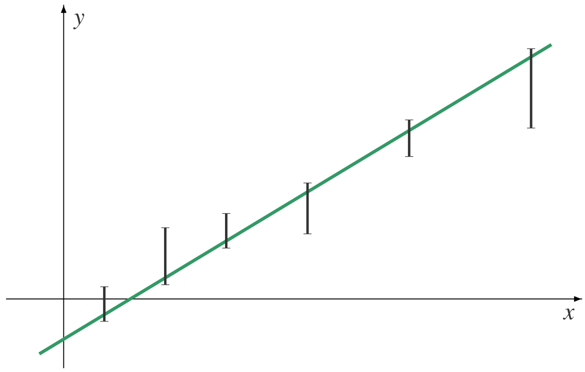

Fig. 2. Illustration of the compatibility between parameters of a linear model and interval measurement data for exact values of independent variable.

This particular case, nevertheless, is practically very important, and its careful solution facilitated the wide dissemination of the new approach. Mathematically, relations (6) form a system of linear inequalities, which can be solved, for example, by linear programming methods (this was proposed in Kantorovich [1962]). In the general case, when both input and output data have interval uncertainty, the following definition seems to be natural.

Definition 1. The parameters $\beta_0, \beta_1, \ldots, \beta_m$ of the linear function (2) are called *compatible* (or *weakly compatible*) with the interval experimental data $(\boldsymbol{x}_{i1}, \boldsymbol{x}_{i2}, \ldots, \boldsymbol{x}_{im}, \boldsymbol{y}_i)$, $i = 1, 2, \ldots, n$, if, for each measurement i, there exist such representatives $x_{i1} \in \boldsymbol{x}_{i1}, x_{i2} \in \boldsymbol{x}_{i2}, \ldots, x_{im} \in \boldsymbol{x}_{im}$ and $y_i \in \boldsymbol{y}_i$ within the measured intervals, that the equality

$$\beta_0 + \beta_1 x_{i1} + \cdots + \beta_m x_{im} = y_i$$

is valid.

According to this definition, the data of each measurement is a large point "inflated" to an axis aligned box in the space $\mathbb{R}^{m+1}$. The fact that the graph of the constructed dependence "passes" through such a point is understood as its intersection with this box (see Fig. 3).

Using the formal language of predicate logic (see e.g. Barker-Plummer *et al.* [2011]), the definition of the set of parameters $\beta = (\beta_0, \beta_1, \ldots, \beta_m)^\top$ compatible with the data (3) looks as follows:

$$\begin{aligned} \{\beta \in \mathbb{R}^{m+1} \mid & (\exists\, x_{11} \in \boldsymbol{x}_{11}) \cdots (\exists\, y_1 \in \boldsymbol{y}_1)(\beta_0 + x_{11}\beta_1 + \cdots = y_1) \,\& \\ & (\exists\, x_{21} \in \boldsymbol{x}_{21}) \cdots (\exists\, y_2 \in \boldsymbol{y}_2)(\beta_0 + x_{21}\beta_1 + \cdots = y_2) \,\& \ldots \& \\ & (\exists\, x_{n1} \in \boldsymbol{x}_{n1}) \cdots (\exists\, y_n \in \boldsymbol{y}_n)(\beta_0 + x_{n1}\beta_1 + \cdots = y_n)\}. \end{aligned} \tag{7}$$

Next, we transform the separating predicate, i.e. the logical formula that stands after the vertical line in the above definition of the set.

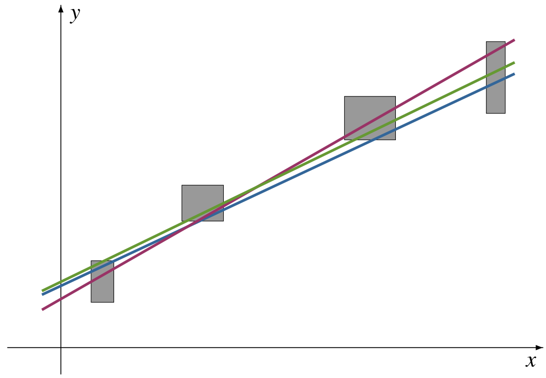

Fig. 3. Illustration of compatibility between parameters of a linear model and interval measurement data in the general case.

If P and Q are propositional formulas depending on the same variable v, then, as is well known, $(\exists\, vP(v))$ & $(\exists\, vQ(v))$ is not equivalent to $\exists\, v(P(v)\ \&\ Q(v))$ [Barker-Plummer *et al.* (2011)]. But the sets of variables that are members of individual conjunctions in formula (7) do not intersect each other. Because of this, we can use the weaker equivalence:

$$(\exists\, v' P(v'))\ \&\ (\exists\, v'' Q(v''))\ \ \Leftrightarrow\ \ \exists\, v'\, \exists\, v''(P(v')\ \&\ Q(v'')).$$

As a consequence, we obtain the formula equivalent to the separating predicate in the set (7):

$$\begin{aligned}&(\exists\, x_{11} \in \boldsymbol{x}_{11}) \cdots (\exists\, y_1 \in \boldsymbol{y}_1)(\exists\, x_{21} \in \boldsymbol{x}_{21}) \cdots (\exists\, y_2 \in \boldsymbol{y}_2) \cdots \\ &\quad (\exists x_{n1} \in \boldsymbol{x}_{n1}) \cdots (\exists\, y_n \in \boldsymbol{y}_n)((\beta_0 + x_{11}\beta_1 + \cdots = y_1)\& \\ &\quad (\beta_0 + x_{21}\beta_1 + \cdots = y_2)\& \ldots \&(\beta_0 + x_{n1}\beta_1 + \cdots = y_n)).\end{aligned} \tag{8}$$

If we organize, from the input data of the problem, an $n \times (m+1)$-matrix $\boldsymbol{X} = (\boldsymbol{x}_{ij})$ and an n-vector $\boldsymbol{y} = (\boldsymbol{y}_i)$, then the large quantifier prefix of formula (8) can be written briefly in the form $(\exists\, X \in \boldsymbol{X})(\exists\, y \in \boldsymbol{y})$, where X is an $n \times (m+1)$-matrix with the elements x_{ij}, and $y = (y_i)$ is an n-vector. Instead of the large formula (8), we thus get

$$\begin{aligned}&(\exists\, X \in \boldsymbol{X})(\exists\, y \in \boldsymbol{y})((\beta_0 + x_{11}\beta_1 + \cdots\ = y_1)\& \\ &\quad (\beta_0 + x_{21}\beta_1 + \cdots\ = y_2)\& \ldots \&(\beta_0 + x_{m1}\beta_1 + \cdots\ = y_n)).\end{aligned}$$

But the resulting conjunction of equalities standing after the quantifier prefix is nothing more than the vector equality $X\beta = y$. Therefore, we can finally conclude that the set of parameters compatible with the data in the sense of the first definition is a set determined as

$$\{\beta \in \mathbb{R}^{m+1} \mid (\exists\, X \in \boldsymbol{X})(\exists\, y \in \boldsymbol{y})(X\beta = y)\}.$$

In the interval analysis, it is called *united solution set* to the interval system of linear equations $\boldsymbol{X}\beta = \boldsymbol{y}$, denoted by $\Xi_{\text{uni}}(\boldsymbol{X}, \boldsymbol{y})$, and informally we can describe it as

$$\Xi_{\text{uni}}(\boldsymbol{X},\boldsymbol{y}) = \{\beta \in \mathbb{R}^{m+1} \mid X\beta = y \text{ for some } X \in \boldsymbol{X} \text{ and } y \in \boldsymbol{y}\}.$$

2.3. *Strong compatibility between parameters and data*

An important new circumstance is that the "swollen" data points acquire an additional structure that the initial infinitesimal points did not have. They become direct Cartesian products of intervals having different meanings, which correspond to input (independent) variables and output (dependent) variable. As a consequence, the different faces of the measurement uncertainty box have different meanings (in Fig. 3, these are the vertical and horizontal sides of the rectangles), and the data fitting problem under interval inaccuracy can take on various contexts. It becomes important how exactly the graph of the constructed function passes through the uncertainty box, which was first noticed, apparently, in Gutowski [2006].

If the process of measuring the values of the input and output is broken in time and, hence, divided into stages, when the outputs are measured *after* fixing the values of the inputs, then another understanding of "compatibility" is more adequate, in which the output constraint must be met *uniformly* at any value of the inputs. Formally, this situation is described by another definition.

Definition 2. The parameters $\beta_0, \beta_1, \ldots, \beta_m$ of the linear function (2) are *strongly compatible* with the interval experimental data $(\boldsymbol{x}_{i1}, \boldsymbol{x}_{i2}, \ldots, \boldsymbol{x}_{im}, \boldsymbol{y}_i)$, $i = 1, 2, \ldots, n$, if, for each measurement i and for any representatives $x_{i1} \in \boldsymbol{x}_{i1}, x_{i2} \in \boldsymbol{x}_{i2}, \ldots, x_{im} \in \boldsymbol{x}_{im}$, there exist $y_i \in \boldsymbol{y}_i$ within the measured intervals, that the equality

$$\beta_0 + x_{i1}\beta_1 + x_{i2}\beta_2 + \cdots + x_{im}\beta_m = y_i$$

is valid.

The set of parameters which are strongly compatible with the data according to the second definition is described, in the formal language, as follows:

$$\begin{aligned} \{\beta \in \mathbb{R}^{m+1} \mid (\forall\, x_{11} \in \boldsymbol{x}_{11}) \cdots (\exists\, y_1 \in \boldsymbol{y}_1)(\beta_0 + x_{11}\beta_1 + \cdots = y_1)\&\\ (\forall\, x_{21} \in \boldsymbol{x}_{21}) \cdots (\exists y_2 \in \boldsymbol{y}_2)(\beta_0 + x_{21}\beta_1 + \cdots = y_2)\& \ldots \&\\ (\forall\, x_{n1} \in \boldsymbol{x}_{n1}) \cdots (\exists\, y_n \in \boldsymbol{y}_n)(\beta_0 + x_{n1}\beta_1 + \cdots = y_n)\}. \end{aligned} \tag{9}$$

We perform equivalent transformations with the selecting predicate of this set, analogous to those carried out previously for Definition 1, using additionally the equivalence

$$(\forall\, uP(u))\ \&\ (\forall\, vQ(v)) \;\;\Leftrightarrow\;\; \forall\, u\, \forall\, v(P(u)\ \&\ Q(v)).$$

It turns out that the set (9) coincides with the set specified as

$$\{\beta \in \mathbb{R}^{m+1} \mid (\forall\, X \in \boldsymbol{X})(\exists\, y \in \boldsymbol{y})(X\beta = y)\},$$

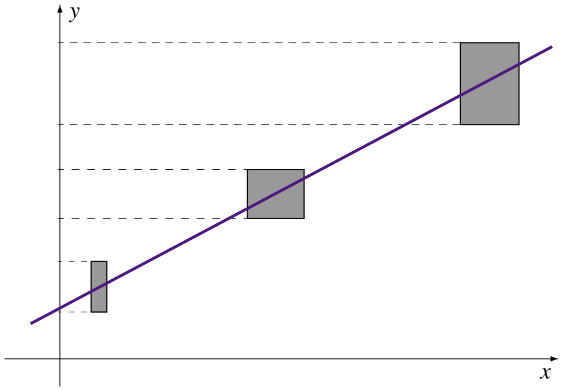

Fig. 4. Illustration of the strong compatibility between parameters of a linear model and interval measurement data.

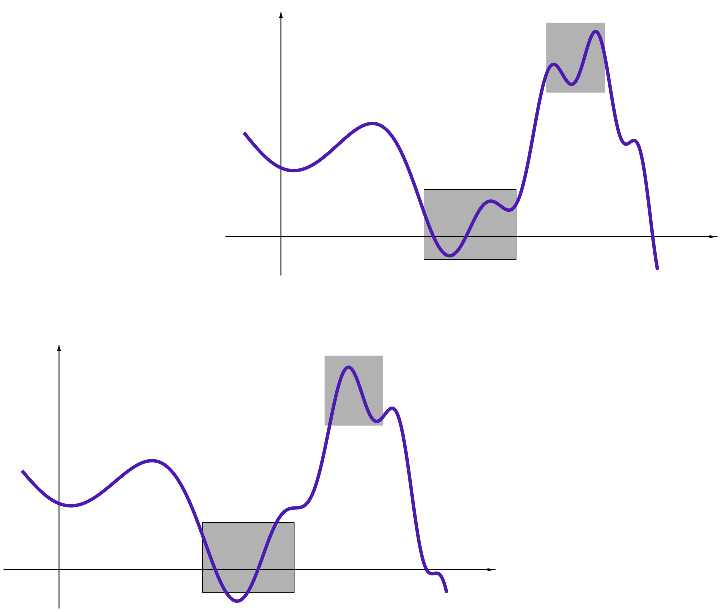

Fig. 5. Illustration of the weak compatibility (below) and strong compatibility (above) between parameters of a nonlinear model and interval measurement data.

where X is an $n \times (m+1)$-matrix with the elements x_{ij}, and $y = (y_i)$ is an n-vector. In interval analysis, this set is called the *tolerable solution set* $\Xi_{\text{tol}}(\boldsymbol{X}, \boldsymbol{y})$ of the interval linear system of equations $\boldsymbol{X}\beta = \boldsymbol{y}$, since historically it originated from practical problems in which the "tolerances" appear on the parameters of an object [Neumaier (1986); Shary (1995b, 2004, 2019)]. Informally,

$$\Xi_{\text{tol}}(\boldsymbol{X}, \boldsymbol{y}) = \{\beta \in \mathbb{R}^{m+1} \mid \text{for any } X \in \boldsymbol{X}, \text{ there holds } X\beta \in \boldsymbol{y}\}.$$

As one can see, the definition of the tolerable solution set differs from the definition of the united solution set by only one logical quantifier, which is applied to the matrix. But this leads to the fact that the properties of the tolerable solution set are much unlike the properties of the united solution set.

2.4. *Plan of the solution*

The specificity of the traditional data fitting problem, where we operate with the point (noninterval) values of measurements and observations, is the fact that the compatibility (consistency) between the parameters of the model and the data is an exceptional event that almost never takes place. In addition, even if there is compatibility, it collapses after an arbitrarily small perturbation of the data. But with an essential interval uncertainty, the set of parameters that are compatible (consistent) with data in typical situations has a nonzero measure, being stable to small perturbations in the data.

The solution of the data fitting problem from inaccurate data will be carried out according to the following general scheme:

(1) We introduce a quantitative "measure of strong compatibility" between parameters and data.
(2) As an estimate of the parameters, we take the point in which the maximum of this measure is achieved.

It is clear that, for a reasonable choice of the "compatibility measure", the evaluation of the parameters will always be performed by this method. But it is completely unessential that the actual compatibility of the obtained parameters and data will in fact take place. Similar to the traditional noninterval case, sometimes there may not exist a set of parameters that are compatible with the data in accordance with Definition 1 or Definition 2. In other words, then there is no line passing through all the uncertainty measurement boxes in the sense we need, either ordinary or strong.

The main question arising in connection with the intended plan is how to take the "measure of strong compatibility/incompatibility" of the data and parameters of a regression line? There are natural requirements that this measure should satisfy. With a nonempty solution set, it must be positive (or at least nonnegative) for points from this set on which "strong compatibility" is actually achieved. For points outside the solution set on which there is no "strong compatibility", it can be negative.

3. Interval Linear Systems of Equations

In this section, we consider in more detail the interval linear systems of equations, i.e. the main object that arises in the solution of the data fitting problem under interval uncertainty for the case of linear functional dependence.

3.1. *United and tolerable solution sets*

Applying the notation traditional for numerical analysis and linear algebra, we write an interval $n \times m$-system of linear algebraic equations in the form

$$\begin{cases} \boldsymbol{a}_{11}x_1 + \boldsymbol{a}_{12}x_2 + \cdots + \boldsymbol{a}_{1m}x_m = \boldsymbol{b}_1, \\ \boldsymbol{a}_{21}x_1 + \boldsymbol{a}_{22}x_2 + \cdots + \boldsymbol{a}_{2n}x_m = \boldsymbol{b}_2, \\ \quad \vdots \qquad\quad \vdots \qquad\quad \ddots \qquad \vdots \qquad\quad \vdots \\ \boldsymbol{a}_{n1}x_1 + \boldsymbol{a}_{n2}x_2 + \cdots + \boldsymbol{a}_{nm}x_m = \boldsymbol{b}_n, \end{cases} \tag{10}$$

or, briefly,

$$\boldsymbol{A}x = \boldsymbol{b}, \tag{11}$$

with interval $n \times m$-matrix $\boldsymbol{A} = (\boldsymbol{a}_{ij})$ and n-vector $\boldsymbol{b} = (\boldsymbol{b}_i)$. Both (10) and (11) are formal entries denoting a family of point linear systems $Ax = b$ of the same structure, with $A \in \boldsymbol{A}$ and $b \in \boldsymbol{b}$. Each system of linear algebraic equations $Ax = b$, whose matrix is taken from the interval matrix $\boldsymbol{A}$ and whose right-hand side b belongs to $\boldsymbol{b}$, can have solutions, and in many practical situations it makes sense to consider them together, as a single set, i.e. taking their union. In this way, we obtain the so-called *united solution set*

$$\Xi_{\text{uni}}(\boldsymbol{A}, \boldsymbol{b}) = \{x \in \mathbb{R}^m \mid \text{there exist such } A \in \boldsymbol{A} \text{ and } b \in \boldsymbol{b} \text{ that } Ax = b\}.$$

It corresponds, apparently, to the simplest and the most natural understanding of what is a “solution” to an interval system of equations. In a large number of works, this set is simply called the “solution set”, without the epithet “united” (see e.g. Mayer [2017], Moore *et al.* [2009] and Neumaier [1990]). Various techniques for estimating it and identifying whether it is empty or nonempty are well developed in modern interval analysis. In the formal language,

$$\Xi_{\text{uni}}(\boldsymbol{A}, \boldsymbol{b}) = \{x \in \mathbb{R}^m \mid (\exists\, A \in \boldsymbol{A})(\exists\, b \in \boldsymbol{b})(Ax = b)\},$$

or

$$\Xi_{\text{uni}}(\boldsymbol{A}, \boldsymbol{b}) = \{x \in \mathbb{R}^m \mid (\exists\, A \in \boldsymbol{A})(Ax \in \boldsymbol{b})\}.$$

But strong compatibility between parameters and data dictates a different understanding of the solution to the interval system of equations. It corresponds to the so-called *tolerable solution set* of the interval linear system of equations, the set defined as

$$\Xi_{\text{tol}}(\boldsymbol{A}, \boldsymbol{b}) = \{x \in \mathbb{R}^m \mid \text{for any } A \in \boldsymbol{A}, \text{ there holds the membership } Ax \in \boldsymbol{b}\}.$$

This is the set of solutions to all point systems $Ax = b$, for which the product Ax falls into the right-hand side intervals $\boldsymbol{b}$ for any $A \in \boldsymbol{A}$. In the formal language,

$$\Xi_{\text{tol}}(\boldsymbol{A}, \boldsymbol{b}) = \{x \in \mathbb{R}^m \mid (\forall\, A \in \boldsymbol{A})(\exists\, b \in \boldsymbol{b})(Ax = b)\},$$

or

$$\Xi_{\text{tol}}(\boldsymbol{A},\boldsymbol{b}) = \{x \in \mathbb{R}^m \mid (\forall A \in \boldsymbol{A})(Ax \in \boldsymbol{b})\}.$$

It is not hard to realize that if the membership $Ax \in \boldsymbol{b}$ is valid *for every* $A \in \boldsymbol{A}$, then it certainly holds *for some* $A \in \boldsymbol{A}$, i.e.

$$\{x \in \mathbb{R}^m \mid (\forall A \in \boldsymbol{A})(Ax \in \boldsymbol{b})\} \subseteq \{x \in \mathbb{R}^m \mid (\exists A \in \boldsymbol{A})(Ax \in \boldsymbol{b})\}.$$

The latter means that the following inclusion holds:

$$\Xi_{\text{tol}}(\boldsymbol{A},\boldsymbol{b}) \subseteq \Xi_{\text{uni}}(\boldsymbol{A},\boldsymbol{b}), \tag{12}$$

i.e. the tolerable solution set is always a subset of the united solution set. In terms of the data fitting problem under interval uncertainty, the above implies that if there is a strong compatibility between parameters and data, then the usual compatibility (which can be called "weak") obviously takes place.

The tolerable solution set and the united solution set coincide with each other if the matrix of the system is a point matrix, i.e. its width is zero:

$$\Xi_{\text{tol}}(A,\boldsymbol{b}) = \Xi_{\text{uni}}(A,\boldsymbol{b}) \quad \text{for any point matrix } A.$$

When the matrix of the system expands, i.e. its width grows, then the tolerable solution set decreases in size, while the united solution set increases, which is their principal distinction. For essentially interval matrices $\boldsymbol{A}$, with nonzero widths of the elements, the difference between the solution sets $\Xi_{\text{tol}}(\boldsymbol{A},\boldsymbol{b})$ and $\Xi_{\text{uni}}(\boldsymbol{A},\boldsymbol{b})$ can be considerable (see examples below).

The tolerable solution set can be empty if the intervals of the right-hand side are too narrow in comparison with the interval elements of the matrix. Then the product $\boldsymbol{A}x$ gets "large range", which may not fit into the corridors of the right-hand sides of the system. For example, for the interval equation $[1,2]x = [2,3]$, the tolerable solution set is empty, In fact, for any nonnegative real t, the ratio of the upper endpoint to the lower one is 2 in the product $[1,2]t = [t,2t]$, whereas this ratio is only $3/2$ for the right-hand side. At the same time, there can be no negative reals in the tolerable solution set, since the right-hand side $[2,3]$ is positive.

There are various ways to investigate whether the tolerable solution set is empty or non-empty for a given interval system of linear equations (see Rohn [1986] and Shary [1995b, 2004, 2012]), and we consider in detail one of them in Sec. 4.3 below.

3.2. *Analytical descriptions of the tolerable solution set*

The definitions of the solution sets given in the preceding section by means of logical formulas are convenient and well understood by practitioners. Nevertheless, they are not very suitable for solving some mathematical questions. For example, the needs to compute with the solution sets as well as to find their estimates require defining these sets through traditional arithmetic and analytical operations.

For the united solution set, there exist quite a lot of such equivalent reformulations of its definition (see Mayer [2017], Neumaier [1990], Rohn [2005] and Shary

[2019]). Also, its structure has been studied in detail. Below, we are presenting the results that give analytic descriptions of the tolerable solution set to the interval linear systems of equations.

The Rohn theorem [Rohn [1986, 2005]; Shary [2019]]. *A point $x \in \mathbb{R}^m$ belongs to the tolerable solution set of the interval $n \times m$-system of linear algebraic equations $\boldsymbol{A}x = \boldsymbol{b}$ if and only if $x = x' - x''$ for some vectors x', $x'' \in \mathbb{R}^m$ that satisfy the system of linear inequalities*

$$\begin{cases} \overline{\boldsymbol{A}}x' - \underline{\boldsymbol{A}}x'' \le \overline{\boldsymbol{b}}, \\ -\underline{\boldsymbol{A}}x' + \overline{\boldsymbol{A}}x'' \le -\underline{\boldsymbol{b}}, \\ x', x'' \ge 0, \end{cases}$$

where $\underline{\boldsymbol{A}}$, $\underline{\boldsymbol{b}}$, $\overline{\boldsymbol{A}}$, and $\overline{\boldsymbol{b}}$ are matrices and vectors made up of the lower and upper endpoints of the interval elements of $\boldsymbol{A}$ and $\boldsymbol{b}$, respectively.

To formulate the next result, we need the following notation. Let vert $\boldsymbol{a}$ denote the set of vertices of the interval vector $\boldsymbol{a} \in \mathbb{IR}^m$, i.e.

$$\text{vert}\, \boldsymbol{a} = \{a \in \mathbb{R}^m \mid \text{either } a_i = \underline{\boldsymbol{a}}_i \text{ or } a_i = \overline{\boldsymbol{a}}_i,\ i = 1, 2, \ldots, m\}.$$

Also, card S will denote the cardinality of a finite set S, i.e. the number of elements of S.

Theorem on the structure of the tolerable solutions set [Sharaya (2005a)]. *Let $\boldsymbol{A}_{i:}$ be the ith row of the interval matrix $\boldsymbol{A}$. For the interval $n \times m$-system of linear algebraic equations $\boldsymbol{A}x = \boldsymbol{b}$, the tolerable solution set $\Xi_{\text{tol}}(\boldsymbol{A}, \boldsymbol{b})$ can be represented in the form*

$$\Xi_{\text{tol}}(\boldsymbol{A}, \boldsymbol{b}) = \bigcap_{i=1}^{n} \bigcap_{a \in \text{vert}\, \boldsymbol{A}_{i:}} \{x \in \mathbb{R}^m \mid ax \in \boldsymbol{b}_i\},$$

i.e. as the intersection of hyperstrips, the number of which does not exceed

$$\sum_{i=1}^{n} \text{card vert}\, \boldsymbol{A}_{i:}$$

and, moreover, does not exceed $n \cdot 2^m$.

The term "hyperstrip" in the formulation of this theorem is quite adequate and justified by the fact that each of the inclusions $ax \in \boldsymbol{b}_i$ for $a \in \boldsymbol{A}_{i:}$ is equivalent to the two-sided inequality

$$\underline{\boldsymbol{b}}_i \le a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{im}x_m \le \overline{\boldsymbol{b}}_i,$$

which actually determines a "strip" between two hyperplanes in $\mathbb{R}^m$. The theorem on the structure of the tolerable solutions set gives, in essence, a representation of the tolerable solution set in the form of a solution set to a system of two-sided linear inequalities whose number is substantially smaller than the total number of extreme ("vertex") inequalities of the interval system, equal to $2^{n(m+1)}$. Overall, it follows from the above results that the tolerable solution set for an interval system of linear algebraic equations is a convex polyhedral set, no matter whether $m \le n$ or $m \ge n$. In particular, the tolerable solution set is always connected and cannot have disjoint parts.

Example 1. As an illustrative example, we consider the interval linear system of equations

$$\begin{pmatrix} [2,4] & [-2,1] \\ [-1,2] & [2,4] \\ [0,1] & [1,2] \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} [-1,2] \\ [-1,2] \\ [0,1] \end{pmatrix}. \tag{13}$$

Its united solution set and tolerable solution set are depicted in Fig. 6.

Example 2. An expressive three-dimensional example is provided by the interval system of linear algebraic equations

$$\begin{pmatrix} [2,3] & [-0.75,0.65] & [-0.75,0.65] \\ [-0.75,0.65] & [2,3] & [-0.75,0.65] \\ [-0.75,0.65] & [-0.75,0.65] & [2,3] \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} [-2,2] \\ [-2,2] \\ [-2,2] \end{pmatrix}. \tag{14}$$

It is a particular case of the test parametric system proposed in Shary [1995a]. The united and tolerable solution sets for (14) are shown in Figs. 7 and 8, and they are visualized with the use of the software packages `IntLinInc3D` [Sharaya (2014)].

Although the interval linear system of equations in the last example is square ($m = n$), while general rectangular systems are most common in data fitting

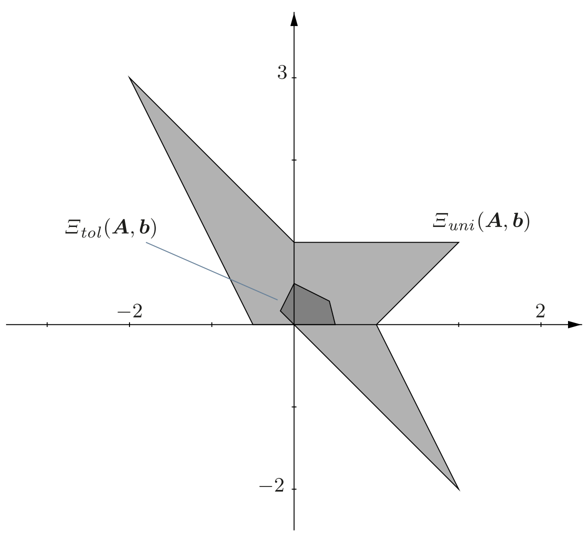

Fig. 6. United solution set and tolerable solution set for the interval linear system (13).

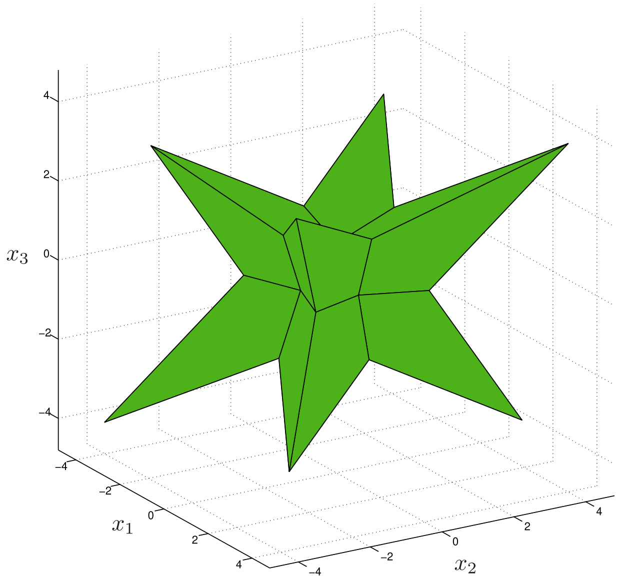

Fig. 7. United solution set for the interval system (14).

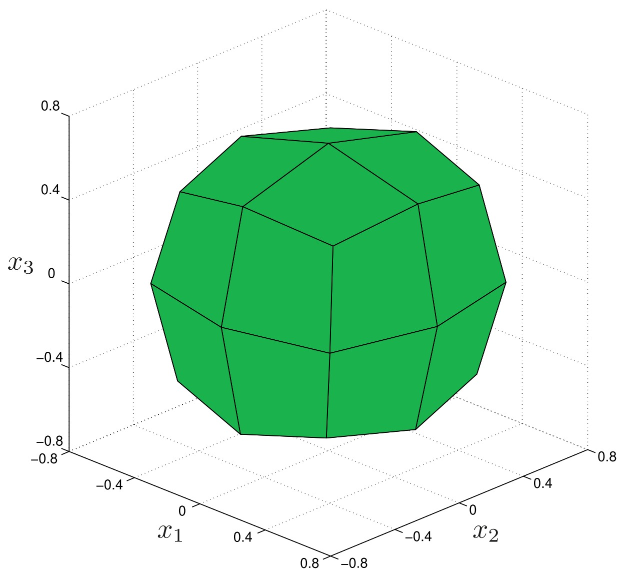

Fig. 8. Tolerable solution set for the interval system (14).

problems (with $m \neq n$), the form of the solution sets in Figs. 7 and 8 (and in Fig. 6 as well) is quite typical. They all are polyhedral sets that are bounded by pieces of hyperplanes. But the tolerable solution set is also convex, whereas the united solution set has only a convex intersection with each orthant of the space

$\mathbb{R}^m$, and it can be nonconvex as a whole (see the details in Mayer [2017], Neumaier [1990] and Shary [2019]). Moreover, the united solution set of interval linear systems with matrices of incomplete rank can be disconnected or unbounded, which is very unnatural for identification problems and data fitting. Readers can see specific examples in the manual for the software package `IntLinInc3D` [Sharaya (2014)].

The problem of solving systems of linear inequalities is known to have polynomial complexity (see, for example, Schrijver [1998]). As a consequence, it follows from the Rohn theorem that in general the recognition of the emptiness/nonemptiness of the tolerable solution set for interval linear systems (as well as finding a point from it) is also a polynomially solvable problem. Answering the same question for the united solution set is generally an NP-hard problem [Kreinovich *et al.* (1998)]. It is equally intractable to obtain outer estimates of the united solution set.

3.3. *Boundedness of the tolerable solution set*

To conclude this section, we give a simple and useful result on the tolerable solution set that allows us to investigate whether it is bounded or unbounded, i.e. whether the tolerable solution set is finite in size or extends infinitely.

Recall that a set of vectors of a linear space is said to be *linearly dependent* if one of the vectors in the set can be expressed as a linear combination, in which not all coefficients are equal to zero, of the other vectors. If no vector in the set can be expressed in this way, then the vectors are called *linearly independent*. An equivalent definition: a finite set of vectors is said to be linearly dependent if there exist scalars, not all of which are zeros, such that the linear combination of the vectors with these scalars is equal to zero vector.

Irene Sharaya's boundedness criterion [Sharaya (2005b)]. *Let the tolerable solution set to an interval linear system $\boldsymbol{A}x = \boldsymbol{b}$ be nonempty. It is unbounded if and only if the matrix $\boldsymbol{A}$ has linearly-dependent noninterval columns.*

The criterion of boundedness shows that the tolerable solution set is unbounded, in fact, under exceptional circumstances, which are almost never fulfilled in practice, when working with real-life interval data. That is, the tolerable solution set is mostly bounded.

4. The Method of Recognizing Functional

The results from the previous section — the Rohn theorem and the structural theorem of Sharaya, in principle, provide tools for investigating the tolerable solution set and working with it. In some situations, the first result is more convenient and preferable, while in other cases the second result is more appropriate. Nevertheless, the representation of the tolerable solution set through a system of linear inequalities has certain disadvantages. In particular, it is desirable to investigate the tolerable solution set and work with it in terms of entire data intervals from the problem statement, and not with their individual endpoints that have multiple occurrences in the system of inequalities.

In this section of our work, we briefly present the known results on the tolerable solution set published earlier in the literature [Shary (1995b, 2004, 2019)].

In the sequel, the classical interval arithmetic $\mathbb{IR}$ plays an important role. $\mathbb{IR}$ is an algebraic system formed by the intervals $\boldsymbol{x} = [\underline{\boldsymbol{x}}, \overline{\boldsymbol{x}}] \subset \mathbb{R}$ so that the result of any arithmetic operation "$\star$" between the intervals is defined by their representatives as

$$\boldsymbol{x} \star \boldsymbol{y} = \{x \star y \,|\, x \in \boldsymbol{x}, y \in \boldsymbol{y}\}, \quad \star \in \{+, -, \cdot, /\}.$$

Expanded constructive formulas for interval arithmetic operations are as follows (see e.g. Mayer [2017], Moore *et al.* [2009], Neumaier [1990] and Shary [2019]):

$$\begin{aligned}
\boldsymbol{x} + \boldsymbol{y} &= [\underline{\boldsymbol{x}} + \underline{\boldsymbol{y}}, \overline{\boldsymbol{x}} + \overline{\boldsymbol{y}}],\\
\boldsymbol{x} - \boldsymbol{y} &= [\underline{\boldsymbol{x}} - \overline{\boldsymbol{y}}, \overline{\boldsymbol{x}} - \underline{\boldsymbol{y}}],\\
\boldsymbol{x} \cdot \boldsymbol{y} &= [\min\{\underline{\boldsymbol{x}}\,\underline{\boldsymbol{y}}, \underline{\boldsymbol{x}}\,\overline{\boldsymbol{y}}, \overline{\boldsymbol{x}}\,\underline{\boldsymbol{y}}, \overline{\boldsymbol{x}}\,\overline{\boldsymbol{y}}\}, \max\{\underline{\boldsymbol{x}}\,\underline{\boldsymbol{y}}, \underline{\boldsymbol{x}}\,\overline{\boldsymbol{y}}, \overline{\boldsymbol{x}}\,\underline{\boldsymbol{y}}, \overline{\boldsymbol{x}}\,\overline{\boldsymbol{y}}\}],\\
\boldsymbol{x}/\boldsymbol{y} &= \boldsymbol{x} \cdot [1/\overline{\boldsymbol{y}}, 1/\underline{\boldsymbol{y}}] \quad \text{for } \boldsymbol{y} \not\ni 0.
\end{aligned}$$

4.1. *Derivation of the recognizing functional*

The starting point for the further constructions is the following characterization of points from the tolerable solution set (see e.g. Neumaier [1986], Sharaya [2005b], and Shary [1995b]): for the interval system of linear algebraic equations $\boldsymbol{A}x = \boldsymbol{b}$, a point $\tilde{x} \in \mathbb{R}^m$ belongs to the tolerable solution set $\Xi_{\text{tol}}(\boldsymbol{A}, \boldsymbol{b})$ if and only if

$$\boldsymbol{A} \cdot \tilde{x} \subseteq \boldsymbol{b}, \tag{15}$$

where "$\cdot$" is the interval matrix multiplication. The validity of this characterization follows from the properties of interval matrix–vector multiplication and the definition of the tolerable solution set. We transform the relation (15) into an analytical form.

First of all, we rewrite (15) as an equivalent system of component-wise inclusions. By definition of the interval matrix–vector product

$$(\boldsymbol{A} \cdot x)_i = \sum_{j=1}^{m} \boldsymbol{a}_{ij} x_j, \quad i = 1, 2, \ldots, n,$$

and then, instead of (15), we can write

$$\sum_{j=1}^{m} \boldsymbol{a}_{ij} x_j \subseteq \boldsymbol{b}_i, \quad i = 1, 2, \ldots, n.$$

Each right-hand side of these inclusions may be represented as the sum of the midpoint $\operatorname{mid} \boldsymbol{b}_i$ and the balanced (symmetric with respect to zero) interval

$[-\text{rad}\,\boldsymbol{b}_i, \text{rad}\,\boldsymbol{b}_i]$:

$$\sum_{j=1}^{m} \boldsymbol{a}_{ij} x_j \subseteq \text{mid}\,\boldsymbol{b}_i + [-\text{rad}\,\boldsymbol{b}_i, \text{rad}\,\boldsymbol{b}_i], \quad i = 1, 2, \ldots, n.$$

Adding $(-\text{mid}\,\boldsymbol{b}_i)$ to both sides of the above relations, we get

$$\sum_{j=1}^{m} \boldsymbol{a}_{ij} x_j - \text{mid}\,\boldsymbol{b}_i \subseteq [-\text{rad}\,\boldsymbol{b}_i, \text{rad}\,\boldsymbol{b}_i], \quad i = 1, 2, \ldots, n.$$

But inclusion of an interval into the balanced interval $[-\text{rad}\,\boldsymbol{b}_i, \text{rad}\,\boldsymbol{b}_i]$ is equivalent to the inequality on the absolute value. So,

$$\left| \sum_{j=1}^{m} \boldsymbol{a}_{ij} x_j - \text{mid}\,\boldsymbol{b}_i \right| \leq \text{rad}\,\boldsymbol{b}_i, \quad i = 1, 2, \ldots, n,$$

which implies

$$\text{rad}\,\boldsymbol{b}_i - \left| \sum_{j=1}^{m} \boldsymbol{a}_{ij} x_j - \text{mid}\,\boldsymbol{b}_i \right| \geq 0, \quad i = 1, 2, \ldots, n.$$

Therefore,

$$\boldsymbol{A}x \subseteq \boldsymbol{b} \;\;\Leftrightarrow\;\; \text{rad}\,\boldsymbol{b}_i - \left| \text{mid}\,\boldsymbol{b}_i - \sum_{j=1}^{m} \boldsymbol{a}_{ij} x_j \right| \geq 0, \quad i = 1, 2, \ldots, n.$$

Finally, we can convolve, over i, the conjunction of the inequalities in the right-hand side of the logical equivalence obtained:

$$\boldsymbol{A}x \subseteq \boldsymbol{b} \;\;\Leftrightarrow\;\; \min_{1 \leq i \leq n} \left\{ \text{rad}\,\boldsymbol{b}_i - \left| \text{mid}\,\boldsymbol{b}_i - \sum_{j=1}^{m} \boldsymbol{a}_{ij} x_j \right| \right\} \geq 0.$$

We thus arrive at the following result.

Theorem. *Let $\boldsymbol{A}$ be an interval $n \times m$-matrix and $\boldsymbol{b}$ be an interval n-vector. Then the expression*

$$\text{Tol}(x, \boldsymbol{A}, \boldsymbol{b}) = \min_{1 \leq i \leq n} \left\{ \text{rad}\,\boldsymbol{b}_i - \left| \text{mid}\,\boldsymbol{b}_i - \sum_{j=1}^{m} \boldsymbol{a}_{ij} x_j \right| \right\}$$

determines the mapping $\text{Tol} : \mathbb{R}^m \times \mathbb{IR}^{n \times m} \times \mathbb{IR}^n \to \mathbb{R}$, *such that the membership of a point* $x \in \mathbb{R}^m$ *in the tolerable solution set* $\Xi_{\text{tol}}(\boldsymbol{A}, \boldsymbol{b})$ *to the interval linear system* *of equations* $\boldsymbol{A}x = \boldsymbol{b}$ *is equivalent to nonnegativity of the mapping* Tol *in the point* x, *i.e.*

$$x \in \Xi_{\rm tol}(\boldsymbol{A}, \boldsymbol{b}) \quad \Leftrightarrow \quad \mathrm{Tol}(x, \boldsymbol{A}, \boldsymbol{b}) \geq 0.$$

The tolerable solution set $\Xi_{\rm tol}(\boldsymbol{A}, \boldsymbol{b})$ to the interval linear systems is therefore the "level set" (also called "Lebesgue set")

$$\{x \in \mathbb{R}^m \mid \mathrm{Tol}(x, \boldsymbol{A}, \boldsymbol{b}) \geq 0\}$$

of the mapping Tol. We call this mapping the *recognizing functional* of the tolerable solution set, since the range of values of the mapping is the numerical set $\mathbb{R}$, i.e. the real number line,[a] and Tol "recognizes", by means of the sign of its values, whether a point belongs to the solution set $\Xi_{\rm tol}(\boldsymbol{A}, \boldsymbol{b})$.

4.2. *Properties of the recognizing functional*

Below, we outline the main properties of the recognizing functional. Their detailed proofs can be found in Shary [1995b, 2004, 2019].

Proposition 1. *The functional* Tol *is continuous over all variables. The functional* Tol *is also Lipschitz continuous*, *i.e. continuous in a stronger sense.*

Proposition 2. *The functional* Tol *is concave with respect to* x *everywhere in* $\mathbb{R}^m$.

Proposition 3. *The functional* $\mathrm{Tol}(x, \boldsymbol{A}, \boldsymbol{b})$ *is a concave polyhedral function*, *i.e. its hypograph is a polyhedral set and its graph is made up of pieces of hyperplanes.*

As an illustration, Fig. 9 depicts the graph of the recognizing functional of the tolerable solution set for the interval system (13). It is clearly seen from the figure that the graph of the functional Tol really has a polyhedral shape.

The form of the expression for the functional Tol obviously implies that the functional is bounded from above:

$$\mathrm{Tol}(x, \boldsymbol{A}, \boldsymbol{b}) \leq \min_{1 \leq i \leq n} \mathrm{rad}\, \boldsymbol{b}_i,$$

since the subtracted absolute values are always nonnegative. In reality, even a stronger assertion is true.

Proposition 4. *The functional* $\mathrm{Tol}(x, \boldsymbol{A}, \boldsymbol{b})$ *attains a finite maximum over the entire space* $\mathbb{R}^m$.

Proposition 5. *If* $\mathrm{Tol}(x, \boldsymbol{A}, \boldsymbol{b}) > 0$, *then the point* x *belongs to the topological interior of the tolerable solution set*, *i.e.* $x \in \mathrm{int}\, \Xi_{\rm tol}(\boldsymbol{A}, \boldsymbol{b})$.

[a]In mathematics, a functional is a mapping defined on an arbitrary set and having a numeric range of values, usually the set of real numbers $\mathbb{R}$ or complex numbers $\mathbb{C}$.

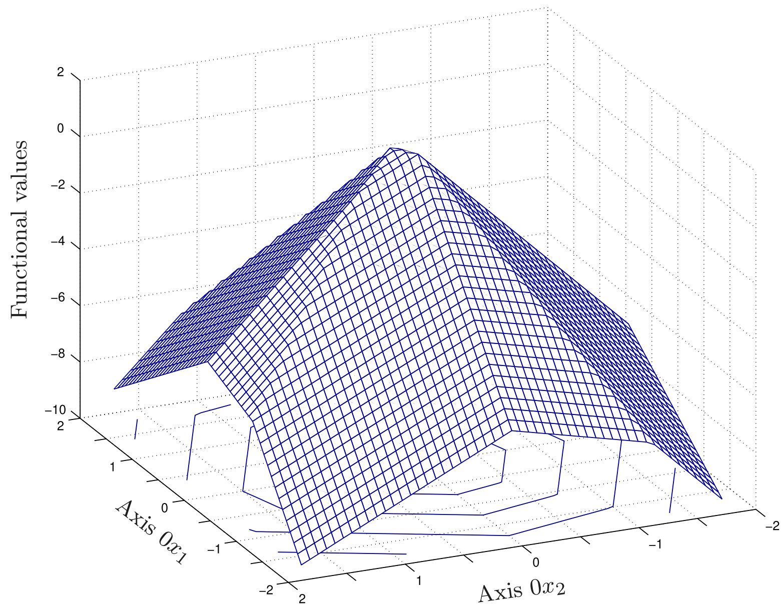

Fig. 9. The graph of the recognizing functional of the tolerable solution set to the system (13).

It should be clarified that any point of topological interior is a point of the set that belongs to it together with a ball (with respect to some norm) having the center at this point. Consequently, points from the interior of the set are “robust” points of the set, i.e. they remain within this set even after their small “perturbations”. This fact often turns out to be important for practice.

Proposition 6. *Let the interval linear system of equations* $\boldsymbol{A}x = \boldsymbol{b}$ *be such that, for each index* $i = 1, 2, \ldots, n$, *either there exists at least one nonzero element in the* i*th row of the matrix* $\boldsymbol{A}$ *or none of the endpoints of the corresponding component of the right-hand side* $\boldsymbol{b}_i$ *is zero. Then the membership* $x \in \operatorname{int}\Xi_{\text{tol}}(\boldsymbol{A}, \boldsymbol{b})$ *implies the strict inequality* $\operatorname{Tol}(x, \boldsymbol{A}, \boldsymbol{b}) > 0$.

4.3. *Solvability investigation*

As a consequence of the above results, we can use the recognizing functional to investigate whether the tolerable solution set is empty or nonempty. This can be done according to the following scheme.

For the interval linear system of equations $\boldsymbol{A}x = \boldsymbol{b}$, we solve the unconstrained maximization problem for the recognizing functional $\operatorname{Tol}(x, \boldsymbol{A}, \boldsymbol{b})$, with respect to x. Let $U = \max_{x\in\mathbb{R}^m} \operatorname{Tol}(x, \boldsymbol{A}, \boldsymbol{b})$, and it is attained at a point $\tau \in \mathbb{R}^m$. Then

- if $U \geq 0$, then $\tau \in \Xi_{\text{tol}}(\boldsymbol{A}, \boldsymbol{b}) \neq \varnothing$, i.e. the tolerable solution set to the system $\boldsymbol{A}x = \boldsymbol{b}$ is not empty and τ lies in it;

- if $U > 0$, then $\tau \in \operatorname{int} \Xi_{\mathrm{tol}}(\boldsymbol{A}, \boldsymbol{b}) \neq \varnothing$, and the membership of the point τ in the tolerable solution set is stable under small perturbations of $\boldsymbol{A}$ and $\boldsymbol{b}$;
- if $U < 0$, then $\Xi_{\mathrm{tol}}(\boldsymbol{A}, \boldsymbol{b}) = \varnothing$, i.e. the tolerable solution set to the interval linear system $\boldsymbol{A}x = \boldsymbol{b}$ is empty.

Next, we answer the question of what is the meaning of specific numerical values of the recognizing functional Tol.

As we have already seen, the criterion for the membership of a point $\tilde{x}$ in the tolerable solution set is the inclusion (15):

$$\boldsymbol{A} \cdot \tilde{x} \subseteq \boldsymbol{b}.$$

It is not difficult to show that the reserve of this inclusion, i.e. how strongly and with what margin this inclusion is fulfilled, is determined precisely by the value of the functional Tol at the point $\tilde{x}$ [Sharaya and Shary (2016)]. One can say that the values of the recognizing functional give a quantitative measure of the compatibility of the point $\tilde{x}$ and the data of the interval linear system, $\boldsymbol{A}$ and $\boldsymbol{b}$, with respect to its tolerable solution set.

5. Maximum Compatibility Method

5.1. *Formulation*

The results of the previous section can be used as a basis for the approach to computing such solutions to the data fitting problem under inaccuracy and uncertainty that satisfy the requirement of strong compatibility between data and parameters.

In accordance with the plan outlined in Sec. 2.4, we need to introduce a "measure of strong compatibility/incompatibility" between parameters and data. It is clear that, for a nonempty tolerable solution set, it must be positive for points from this set, on which the "strong compatibility" is actually achieved. For points outside the tolerable solution set, on which there is no "strong compatibility", it can be negative. Recalling the properties and meaning of the recognizing functional Tol presented in Sec. (4), we can see that it is very suitable for the role of the compatibility measure. In particular, Propositions 5 and 6 show that Tol distinguishes the boundary and interior of the tolerable solution set.

Suppose that an interval data fitting problem is specified for the linear function (1) with the parameters $\beta_0, \beta_1, \ldots, \beta_m$. In other words, we are given an interval $n \times (m+1)$-matrix $\boldsymbol{X} = (\boldsymbol{x}_{ij})$ and vector $\boldsymbol{y} = (\boldsymbol{y}_i)$, such that the first column in the matrix $\boldsymbol{X}$ consists of all ones and has the number 0 (see Sec. 2.1). Then we have to construct the recognizing functional of the tolerable solution set for the interval system of linear equations $\boldsymbol{X}\beta = \boldsymbol{y}$, i.e.

$$\operatorname{Tol}(\beta, \boldsymbol{X}, \boldsymbol{y}) = \min_{1 \leq i \leq n} \left\{ \operatorname{rad} \boldsymbol{y}_i - \left| \operatorname{mid} \boldsymbol{y}_i - \sum_{j=0}^{m} \boldsymbol{x}_{ij} \beta_j \right| \right\},$$

which should serve as the "strong compatibility measure" between the data $\boldsymbol{X}$, $\boldsymbol{y}$ and the parameter vector $\beta = (\beta_0, \beta_1, \ldots, \beta_m)^\top$.

The above motivates the following method for estimating the parameters of a linear functional dependence from inaccurate data, which we will call the "strong version" of the maximum compatibility method or simply *the maximum compatibility method* for brevity:

> As an estimate $\beta^\star$ of the parameters of the linear function (2), we take the point where the maximum of the recognizing functional Tol is reached.

In mathematical terms,

$$\beta^\star = \arg \max_{\beta \in \mathbb{R}^{m+1}} \mathrm{Tol}(\beta, \boldsymbol{X}, \boldsymbol{y}).$$

As a consequence of the theory of Sec. 3,

- if $\max \mathrm{Tol} \geq 0$, then the argument of the maximum lies in the set of parameters strongly compatible with the data;
- if $\max \mathrm{Tol} < 0$, then the set of parameters having strong compatibility with the data is empty, but the argument of the maximum minimizes the incompatibility (inconsistency) between the parameters and data.

The usual ("weak") version of the maximum compatibility method developed earlier in the literature [Kreinovich and Shary (2016); Shary (2012, 2016); Shary and Sharaya (2013, 2016)] is based on similar ideas. We need to maximize a measure of compatibility between the data and parameters of the function, which is also expressed by means of some recognizing functionals, called Uni and Uss. But the properties of weak compatibility estimate differ significantly from the properties of strong compatibility estimates. We have already mentioned that computing strong compatibility estimates is polynomially complex, while finding an estimate in the sense of weak compatibility is usually NP-hard. Additionally, we will see below in Sects. 5.4 and 5.5 that the strong compatibility estimates turn out to be more adequate to reality and actual properties of point estimates that we expect in practice.

5.2. *Interpretation of the maximum compatibility method*

Yet another interpretation of the maximum compatibility method in the case of the empty solution set $\Xi_{\mathrm{tol}}(\boldsymbol{X}, \boldsymbol{y})$ can be, for example, as follows: estimate of the parameters, i.e. the argument on which $\max \mathrm{Tol}$ is reached, is the first point that appears in the nonempty tolerable solution set after the uniform widening of the right-hand side vector with respect to its midpoint.

In fact, let us consider the expression for the recognizing functional Tol:

$$\mathrm{Tol}(\beta, \boldsymbol{X}, \boldsymbol{y}) = \min_{1 \leq i \leq n} \left\{ \mathrm{rad}\, \boldsymbol{y}_i - \left| \mathrm{mid}\, \boldsymbol{y}_i - \sum_{j=0}^{m} \boldsymbol{x}_{ij} \beta_j \right| \right\}.$$

The quantities $\operatorname{rad} \boldsymbol{y}_i$ enter, as summands, in all expressions over which we take $\min_{1 \le i \le n}$ when calculating the final value of the functional. Therefore, if we denote

$$\boldsymbol{e} = ([-1, 1], \ldots, [-1, 1])^\top,$$

then, for the interval system $\boldsymbol{X}\beta = \boldsymbol{y} + C\boldsymbol{e}$ with a widened right-hand side, we have

$$\operatorname{Tol}(\beta, \boldsymbol{X}, \boldsymbol{y} + C\boldsymbol{e}) = \operatorname{Tol}(\beta, \boldsymbol{X}, \boldsymbol{y}) + C,$$

since all the radii of the right-hand side components become equal to $\operatorname{rad} \boldsymbol{y}_i + C$, $i = 1, 2, \ldots, n$. Consequently,

$$\max_\beta \operatorname{Tol}(\beta, \boldsymbol{X}, \boldsymbol{y} + C\boldsymbol{e}) = \max_\beta \operatorname{Tol}(\beta, \boldsymbol{X}, \boldsymbol{y}) + C.$$

Expansion of the interval relative to the center is, actually, an increase in its uncertainty with the invariable value of the most representative point of the interval, its midpoint. As we can see, argument of the maximum of the recognizing functional is really the most promising point if we consider it with respect to variation in the accuracy of the output interval data.

5.3. *The maximum compatibility method generalizes Chebyshev data approximation*

In the limit case where there is no interval uncertainty in our measurements and we have usual point data, any good interval method should turn into a reasonable data fitting method for such data. The strong version of the maximum compatibility method, like the weak one, coincides with the so-called Chebyshev data approximation, which has long been successfully applied to data processing (see, for example, Iske [2018] and Remez [1962]).

In fact, if the data matrix $\boldsymbol{X}$ and the data vector $\boldsymbol{y}$ are point (noninterval), i.e. $\boldsymbol{X} = X = (x_{ij})$ and $\boldsymbol{y} = y = (y_i)$, then for all i, j

$$\operatorname{rad} \boldsymbol{y}_i = 0, \quad \operatorname{mid} \boldsymbol{y}_i = y_i, \quad \boldsymbol{x}_{ij} = x_{ij}.$$

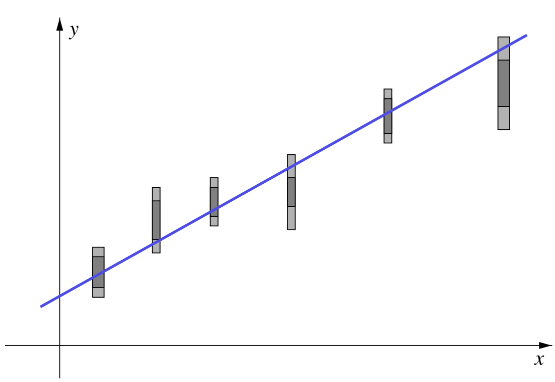

Fig. 10. Expanding the data uncertainty boxes along the output variables results in the strong compatibility.

Then the recognizing functional of the solution set (which is both united and tolerable simultaneously) takes the form

$$\begin{aligned}\mathrm{Tol}(\beta, X, y) &= \min_{1\le i\le n}\left\{-\left|y_i - \sum_{j=0}^{m} x_{ij}\beta_j\right|\right\}\\ &= -\max_{1\le i\le n}\left|y_i - \sum_{j=0}^{m} x_{ij}\beta_j\right|\\ &= -\max_{1\le i\le n}|(X\beta)_i - y_i| = -\|X\beta - y\|_\infty.\end{aligned}$$

In the above formula, $\|\cdot\|_\infty$ denotes the Chebyshev norm of a vector in the finite-dimensional space $\mathbb{R}^n$, which is defined as $\|z\|_\infty = \max_{1\le i\le n}|z_i|$ (it is also called ∞-*norm*, *uniform norm*, or *maximum norm*). Therefore,

$$\max \mathrm{Tol}(\beta) = \max_{\beta\in\mathbb{R}^{m+1}}(-\|X\beta - y\|_\infty) = -\min_{\beta\in\mathbb{R}^{m+1}}\|X\beta - y\,\|_\infty,$$

as long as $\max(-f(\beta)) = -\min f(\beta)$. Thus, the maximization of the recognizing functional is equivalent in this case to minimization of the Chebyshev norm of the residual, i.e. of the difference between the left-hand and right-hand sides of the equation system.

5.4. *Bounded variance of the strong compatibility estimates*

From a practical point of view, a strong version of the maximum compatibility method is more favorable for the solution of the data fitting problem with "overlapping" uncertainty boxes. The strong version allows to obtain a reasonable and bounded set of alternatives in such complex cases when the uncertainty boxes intersect each other.

Let us consider the situation when two uncertainty boxes intersect so that their intersection is solid, i.e. it is a box whose width is nonzero in each dimension, as shown in Fig. 11. Then, within this solid intersection, we can always take two points from the uncertainty boxes that have arbitrary mutual position, so that the straight line $y = \beta_0 + \beta_1 x$ passing through them will have the angular coefficient β_1 equal to any real number (or infinity as well). As a consequence, the set of parameters (β_0, β_1) compatible, in the sense of Definition 1, with the data from Fig. 11 is unbounded.

At the same time, the tolerable solution set for interval linear systems with essentially interval matrix should be bounded, which follows from the Sharaya boundedness criterion (see Sec. 3.3). Therefore, the set of parameters strongly compatible with the data (i.e. in the sense of Definition 2) is bounded for the case depicted in Fig. 11. This helps to reduce indeterminacy and ambiguity in estimating the parameters of the functional dependence, i.e. to choose the solution more definitely from a narrow collection of alternatives rather than from an unbounded set.

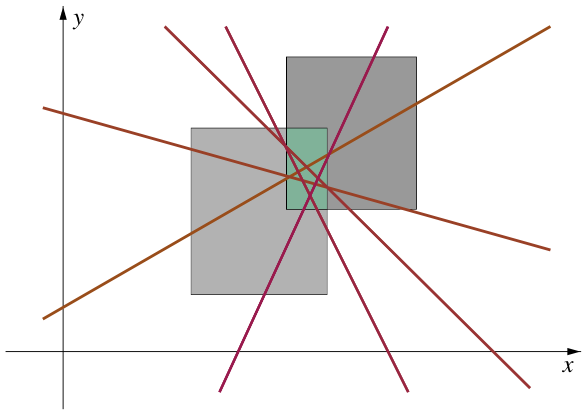

Fig. 11. The intersection of boxes may result in total indeterminacy of the angular coefficient of the line passing through the boxes in the sense of “weak compatibility”.

These ideas can be given a different form. The important concepts of *variance* and *standard deviation* are known to be one of the main characteristics of statistical estimates obtained using the methods of probability theory (see e.g. Cramér [1946]). They characterize the dispersion or variability of the estimate, or, put differently, its possible range of values. The analog of the variance and standard deviation in the statistics of interval data can be the size of the set of parameters compatible with the data, i.e. the size of the corresponding solution set to an interval equation system constructed from measurement data. Computation of enclosures of the solution sets to interval systems of equations can be performed using interval methods described in the literature [Mayer (2017); Moore *et al.* (2009); Neumaier (1990); Rohn (2005); Shary (2019)].

The relation (12), i.e. the property that the tolerable solution set is always included in the united solution set, can be interpreted as the fact that estimates in the sense of ordinary weak compatibility always have a greater “variance” than estimates in the sense of strong compatibility. In addition, the “variance” of the strong compatibility estimates is almost always finite, which follows from Sharaya’s criterion of boundedness of the tolerable solution set (see Sec. 3.3).

The above phenomenon is, in effect, a manifestation of the so-called “regularizing properties” of the tolerable solution set for interval systems of equations. It turns out that the tolerable solution set is the “most stable” among all the solution sets to the interval system of linear equations, which is discussed in detail in Shary [2018].

5.5. *Strong compatibility and the Demidenko paradox*

The “Demidenko paradox” is a paradoxical statement about the properties of the solution to the data fitting problem under interval uncertainty, first noted by Demidenko [1990] (see also Kreinovich and Shary [2016], Shary [2012], and Shary and Sharaya [2013]). Its essence can be expressed by the phrase “the worse, the better". More precisely, the wider the intervals of data uncertainty, i.e. the more uncertainty they represent, the easier it is to draw through them the graph of the constructed function.

Data uncertainty is undesirable because it distorts the true picture of reality. Therefore, reducing uncertainty, i.e. reducing the size of data uncertainty boxes, is a boon that should be welcomed in practice. On the other hand, for wider intervals of data, the united solution set of the interval equation system built from this data is also wider and, therefore, there are more opportunities to select model parameters from it, than for the case of narrow interval data. Thus, the higher the accuracy of the data, the lower the interval uncertainty and the worse it is to estimate the parameters. Conversely, the wider the interval uncertainty and the worse we know the exact values of the measured variables, the better the parameter estimation process and the richer the set of results that can be obtained. This situation is depicted in Fig. 12 where the uncertainty intervals at the lower picture are obtained by contracting the intervals of the upper picture. At the same time, the opportunity to draw a straight line passing through all uncertainty boxes is lost.

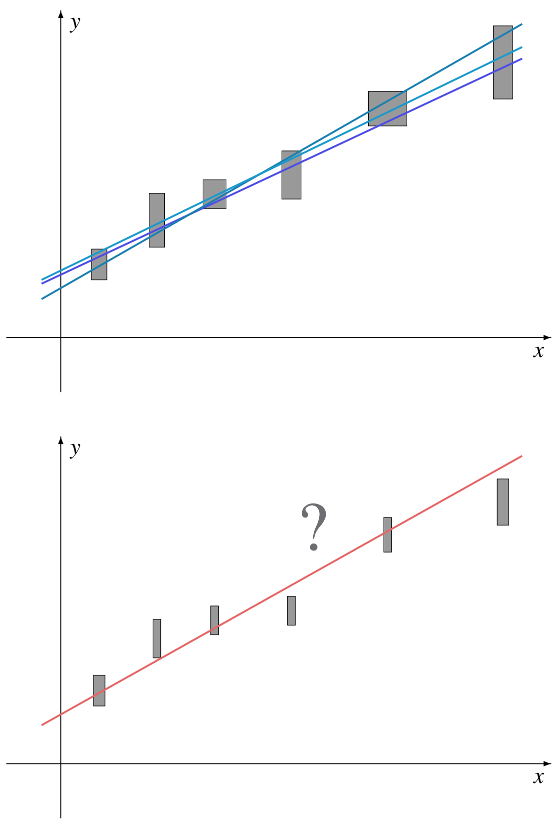

Fig. 12. Wide uncertainty boxes enable us to construct many models compatible with the data. For narrow uncertainty boxes, a model compatible with the data may not exist.

There are two basic ways to overcome the Demidenko paradox. The first one is based on the assumption that the intervals of the data adequately represent the boundaries of the measurement errors, so that the reduction of their width uncertainty is positive. Hence, the impossibility to choose the model parameters compatible with these interval data (where the solution set of the interval equation system is empty) indicates the inadequacy of the model used to describe the object. As a result, the model must be changed, and the process of parameter estimation must be repeated using another model.

The second way assumes that the uncertainty intervals of the data do not represent exactly the set of possible values of the corresponding variables. Therefore, one does not have to obtain full compatibility with the experimental data for the selected model of the object. As in the traditional case of noisy point (noninterval) data, a certain incompatibility (inconsistency) is acceptable, and then the problem of minimizing this incompatibility needs to be solved. Yet another situation where one has to go this way stems from the need to retain the selected model, form of a functional dependence between the considered variables about which it is *a priori* known that "this must be the case". Following this way, one has to select a numerical "incompatibility measure" between the data and model parameters. Then, for example, a point of the parameter space where the incompatibility (inconsistency) is minimal can be taken as the desired estimate.

Anyway, the Demidenko paradox is not fully applicable to the situation of strong compatibility between parameters and data, since the tolerable solution set, when changing data intervals, behaves quite differently from the united solution set.

As we already noted in Sec. 3.1, the tolerable solution set shrinks as the width of the intervals in the matrix of the equation system increases. Then, it becomes more difficult to construct a straight line that passes through the uncertainty boxes in the strong sense of Definition 2. This fact is well understood intuitively, from the consideration of Figs. 4 and 5 in which the widths of the boxes grow along the axis Ox. Thus, here we are in a situation where the increase of interval uncertainty at the input leads to the similar deterioration in the solution of the problem (it becomes more difficult to choose the desired function). The Demidenko paradox does not work.

6. Implementation

The theory developed in the preceding sections will be practical and really useful only if we have at our disposal effective methods for finding the maximum of the recognizing functional of the tolerable solution set, i.e. $\max \mathrm{Tol}$. The properties of the recognizing functional are considered in Sec. 4, and they are favorable for applying efficient numerical optimization methods.

In the general case, the problem of computing $\max \mathrm{Tol}$ is the problem of unconstrained maximization of a concave nonsmooth objective function. Its solution can be found by nonsmooth optimization methods, which many researchers have been intensively developing for several decades. We successfully used the algorithms designed by Shor and his co-workers in Kiev (see Shor and Zhurbenko [1971] and Stetsyuk [2014]).

In recent years, we freely circulate the program `tolsolvty`, accessible from the website "Interval Analysis and its Applications" — http://www.nsc.ru/interval (section "Software", then "Some interval programs on Scilab" or "Some interval programs on MATLAB"). The program is designed to numerically determine the unconditional maximum of the recognizing functional Tol and uses, as a basis, the code `ralgb5` developed by Stetsyuk (Institute of Cybernetics of the National Academy of Sciences of Ukraine; see an earlier article [Stetsyuk (2017)] specially devoted to this algorithm). In fact, `tolsolvty` is a very good and time-tested implementation of the maximum compatibility method in the "strong sense" that can be recommended for solving practical problems. Under the name `TOLSOLVTY2`, the international version of this program is also uploaded to our page of ResearchGate (see https://www.researchgate.net).

Recently, it has become possible to use the separating planes methods to find the maximum of the recognizing functional Tol. These methods were proposed by Nurminski [1997] and further developed and adapted by Vorontsova [2016, 2017]. The free program `tolspaclip` for maximizing the recognizing functional Tol that implements the separating planes method with additional clipping is posted on the website "Interval Analysis and its Applications". It is intended for the same purposes as `tolsolvty` and has roughly the same functionality.

Example 3. As a specific numerical example, we construct a homogeneous linear dependence of the form

$$y = y(x_1, x_2, x_3) = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 \tag{16}$$

from the interval observation data presented in the following table:

Observation	x_1	x_2	x_3	y
#1	$[11, 12]$	$[13, 14]$	$[15, 16]$	$[18, 22]$
#2	$[21, 22]$	$[23, 24]$	$[25, 26]$	$[28, 32]$
#3	$[31, 32]$	$[33, 34]$	$[35, 36]$	$[38, 42]$
#4	$[41, 42]$	$[43, 44]$	$[45, 46]$	$[48, 52]$

(17)

To determine the coefficients β_1, β_2, and β_3, we have to consider the interval linear 4×3-system of equations

$$\begin{pmatrix} [11, 12] & [13, 14] & [15, 16] \\ [21, 22] & [23, 24] & [25, 26] \\ [31, 32] & [33, 34] & [35, 36] \\ [41, 42] & [43, 44] & [45, 46] \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = \begin{pmatrix} [18, 22] \\ [28, 32] \\ [38, 42] \\ [48, 52] \end{pmatrix}. \tag{18}$$

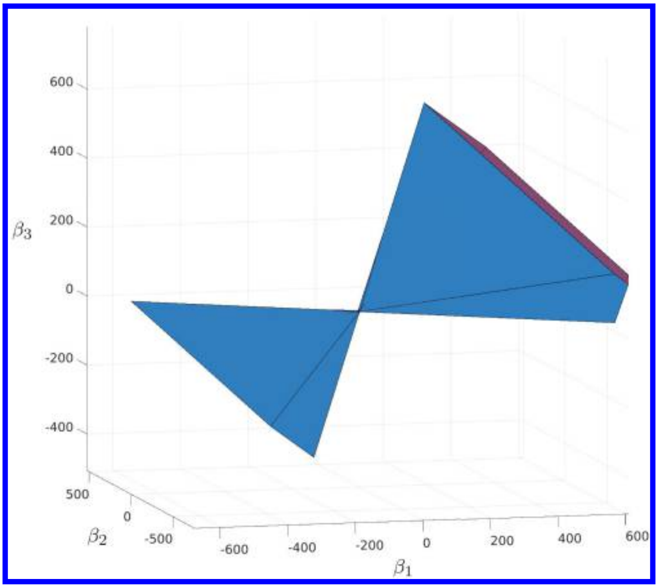

Fig. 13. The unbounded united solution set to the interval linear system (18).

The united solution set to the system (18) is unbounded (see Fig. 13), and the usual compatibility between data and parameters (in the sense of Definition 1) leads to a large indeterminacy in the choice of parameters we can take for the linear function (16). Obviously, most of the triples $(\beta_1, \beta_2, \beta_3)^\top$ that are present in the unbounded solution set will not have a physical meaning due to their large values. In essence, we have here a situation with "infinite variance" of the estimate described earlier in Sec. 5.4.

At the same time, the tolerable solution set to the system (18), depicted in Fig. 14, is bounded.[b] It provides us with quite a limited collection of values for the coefficients of the linear function (16).

The numerical results produced by the program `tolsolvty` (with all the stopping criteria of the order 10^{-10}) are the following:

$$\max_{\beta \in \mathbb{R}^3} \mathrm{Tol}(\beta) = 0.375, \quad \text{and it is attained at the point} \begin{pmatrix} -1.125 \\ 4.4 \cdot 10^{-12} \\ 2.125 \end{pmatrix}. \qquad (19)$$

Then the best fit linear function (16) for the interval data (17) should be

$$y = -1.125x_1 + 4.4 \cdot 10^{-12} x_2 + 2.125x_3.$$

[b]Again, the pictures of the solution sets are produced by the package `IntLinInc3D` [Sharaya (2014)].

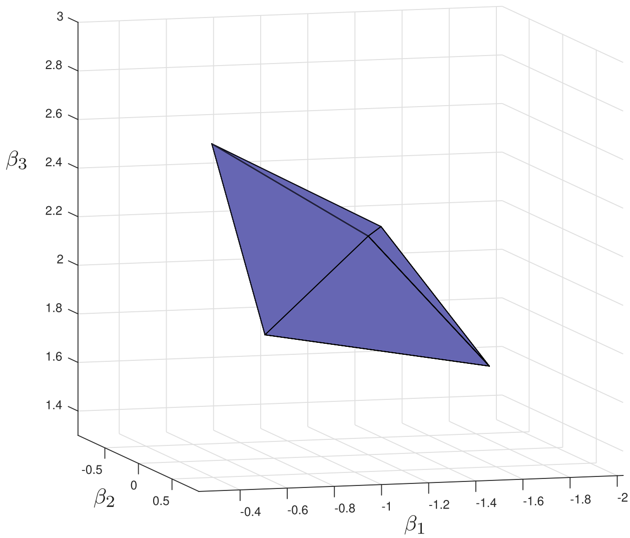

Fig. 14. The tolerable solution set to the interval linear system (18).

We may see that the second coefficient is almost zero, and Fig. 14 shows that the second component of points from the tolerable solution set is relatively small and varies around zero.

One can construct an inner interval box within the tolerable solution set, taking the point (19) as its center and using the methods described in the literature [Neumaier (1986); Shary (1995b, 2004, 2019)]:

$$\begin{pmatrix} [-1.1278409, -1.1221591] \\ [-0.0028409, 0.0028409] \\ [2.1221591, 2.1278409] \end{pmatrix}.$$

The above indicates a low significance of this coefficient in the linear function (16). If we were to consider a real-life problem with data (17), then the corresponding factor, perhaps, should be recognized as having no influence on the phenomenon we are studying.

7. Generalizations

Let us imagine a situation where, in some measurements, strong compatibility of parameters with the data is required, while the usual weak compatibility is sufficient in the other measurements. In formal mathematical language, this means that the logical quantifiers "$\forall$" are applied to a part of the input variables x_{ij}s, and the logical quantifiers "$\exists$" are applied to the other part of x_{ij}s in formula (8).

Then, instead of the united or tolerable solution sets, we naturally arrive at the solution sets in which quantifiers of different meanings acting on different input variables are intermixed. These are the so-called "quantifier solution sets" for the interval system of equations constructed from the data of the problem (see e.g. Shary [2002] and Sharaya [2018]). It can be shown that, in fact, the most general quantifier solutions do not arise in this situation, and we will have to do with their particular case, the so-called AE solutions of the interval systems of equations [Shary (2002, 2019)]. AE solutions are defined as quantifier solutions of interval systems of equations defined by logical formulas in which all occurrences of the universal quantifier "∀" precede the occurrences of the existence quantifier "∃".

For AE solution sets, it is also possible to construct "recognizing functionals" having properties that are analogous to the properties of the functional Tol for the tolerable solution set. This work has been done in Sharaya and Shary [2016], where the general recognizing functionals are constructed based on the idea of considering the "reserve" of the so-called characteristic inclusion for the corresponding AE solution sets. These functionals can serve to measure the degree of compatibility (consistency) between parameters and data in the case of more general requirements on the solution. Having found the unconditional maximum of such a recognizing functional, we obtain the point at which the maximum of compatibility is achieved, and this point can be taken as the desired estimate of the parameters. That is the general scheme for solving the problem, which, of course, needs to be specified and supplied with efficient computational algorithms.

8. Conclusions

In data fitting problems under interval uncertainty, it is necessary to distinguish between different types of compatibility (consistency) between interval data and parameters of the constructed functional dependence. In particular, it makes sense to introduce the concepts of "strong" and "weak" compatibilities of data and parameters that correspond to the different roles of input (predictor) variables and output (criterion) variables in the measurement process.

The maximum compatibility method is a promising method for parameter identification and data fitting under interval uncertainty, which is based on maximizing the recognizing functional of the solution set for the problem. It is a generalization of the Chebyshev data approximation and can serve as a good alternative to traditional methods of regression analysis using probabilistic models of data errors. In this paper, a modification is suggested for the case of "strong" compatibility (consistency) between parameters and data.

The strong version of the maximum compatibility method has several advantages over the usual ("weak") version. First, strong compatibility estimates have a polynomial computational complexity. Second, these estimates are robust and their variance is finite. Third, the strong compatibility estimation is only partially subject to the "Demidenko paradox", being in better agreement with the intuitive understanding of the meaning of estimates in interval data fitting.

An interesting open question: what is the probabilistic interpretation of the maximum compatibility method for the "strong case"?

For the case of weak compatibility between parameters and data, a probabilistic interpretation of the maximum compatibility method was given in a previous work [Kreinovich and Shary (2016)]. It was shown that the estimates produced by the maximum compatibility method coincide with those obtained from the maximum likelihood method for uniform distributions over data intervals. It would be extremely useful to derive a similar result for the strong compatibility.

References

Barker-Plummer, D., Barwise, J. and Etchemendy, J. (2011). *Language, Proof and Logic*, 2nd edn. CSLI Publications, Stanford, CA.

Combettes, P. (1993). The foundations of set theoretic estimation. *Proc. IEEE*, **81**(2): 182–208.

Cramér, H. (1946). *Mathematical Methods of Statistics*. Editions of 1946–1999. Princeton University Press, Princeton.

Demidenko, E. Z. (1990). Comment II to the Paper of A. P. Voshchinin, A. F. Bochkov, and G. R. Sotirov "A method of data analysis under interval nonstatistical error". *Ind. Lab.*, **56**(7): 83–84 (in Russian). Available at: http://www.nsc.ru/interval/Library/Thematic/DataProcs/VosBochSoti.pdf.

Gutowski, M. W. (2006). Interval experimental data fitting. *Focus on Numerical Analysis*, ed. J. P. Liu. Nova Science Publishers Inc., New York, pp. 27–70.

Iske, A. (2018). *Approximation Theory and Algorithms for Data Analysis*. Springer, Cham, Switzerland.

Jaulin, L., Kieffer, M., Didrit, O. and Walter, E. (2001). *Applied Interval Analysis — With Examples in Parameter and State Estimation, Robust Control and Robotics*. Springer, London.

Kantorovich, L. V. (1962). On some new approaches to numerical methods and processing observation data. *Sib. Math. J.*, **3**(5): 701–709 (in Russian). Available at: http://www.nsc.ru/interval/Introduction/Kantorovich62.pdf.

Kearfott, R. B., Nakao, M., Neumaier, A., Rump, S., Shary, S. P. and van Hentenryck, P. (2010). Standardized notation in interval analysis. *Comput. Technol.*, **15**(1): 7–13.

Kreinovich, V., Lakeyev, A., Rohn, J. and Kahl, P. (1998). *Computational Complexity and Feasibility of Data Processing and Interval Computations*. Kluwer Academic, Dordrecht and Springer Science + Business Media, Dordrecht.

Kreinovich, V. and Shary, S. P. (2016). Interval methods for data fitting under uncertainty: A probabilistic treatment. *Reliab. Comput.*, **23**: 105–140. Available at: http://interval.louisiana.edu/reliable-computing-journal/volume-23/reliable-computing-23-pp-105-140.pdf.

Mayer, G. (2017). *Interval Analysis and Automatic Result Verification*. De Gruyter, Berlin.

Milanese, M., Norton, J., Piet-Lahanier, H. and Walter, E. (eds.) (1996). *Bounding Approaches to System Identification*. Plenum Press, New York. Doi:10.1007/978-1-4757-9545-5.

Moore, R. E., Kearfott, R. B. and Cloud, M. J. (2009). *Introduction to Interval Analysis*. SIAM, Philadelphia.

Neumaier, A. (1986). Tolerance analysis with interval arithmetic. *Freiburger Intervall-Berichte*, **86**(9): 5–19.

Neumaier, A. (1990). *Interval Methods for Systems of Equations*. Cambridge University Press, Cambridge.

▶Nurminski, E. A. (1997). Separating plane algorithms for convex optimization. *Math. Program.*, **76**: 373–391. Doi:10.1007/BF02614389.

▶Polyak, B. T. and Nazin, S. A. (2006). Estimation of parameters in linear multidimensional systems under interval uncertainty. *J. Autom. Inf. Sci.*, **38**(2): 19–33.

Remez, E. Ya. (1962). *General Computational Methods of Chebyshev Approximation: The Problems with Linear Real Parameters*, Translation Series, Vol. 4491. U. S. Atomic Energy Commission, Division of Technical Information, Oak Ridge, TN.

Rohn, J. (1986). Inner solutions of linear interval systems. *Interval Mathematics 1985*, ed. K. Nickel, Lecture Notes in Computer Science, Vol. 212. Springer-Verlag, Berlin. Doi:10.1007/3-540-16437-5_15.

Rohn, J. (2005). *A Handbook of Results on Interval Linear Problems*. Czech Academy of Sciences, Prague. Available at: http://www.nsc.ru/interval/Library/Surveys/ILinProblems.pdf.

Schrijver, A. (1998). *Theory of Linear and Integer Programming*. John Wiley & Sons, Chichester.

▶Schweppe, F. S. (1968). Recursive state estimation: Unknown but bounded errors and system inputs. *IEEE Trans. Autom. Control*, **13**(1): 22–28.

Sharaya, I. A. (2005a). Structure of the tolerable solution set of an interval linear system. *Comput. Technol.* **10**(5): 103–119 (in Russian).

▶Sharaya, I. A. (2005b). On unbounded tolerable solution sets. *Reliab. Comput.*, **11**(5): 425–432. Doi:10.1007/s11155-005-0049-9.

Sharaya, I. A. (2014). `IntLinInc3D`, a software package for visualization of solution sets to interval linear 3D systems of relations. Available at: http://www.nsc.ru/interval/sharaya/.

Sharaya, I. A. (2018). Quantifier-free descriptions for quantifier solutions to interval linear systems of relations. Preprint, arXiv:1802.09199 [math.OC]. Available at: https://arxiv.org/abs/1802.09199.

▶Sharaya, I. A. and Shary, S. P. (2016). Reserve of characteristic inclusion as recognizing functional for interval linear systems. *Proc. 16th Int. Symp. Scientific Computing, Computer Arithmetic, and Validated Numerics: Revised Selected Papers*. eds. M. Nehmeier, J. Wolff von Gudenberg and W. Tucker, Lecture Notes in Computer Science, Vol. 9553. Springer, Cham, pp. 148–167.

▶Shary, S. P. (1995a). On optimal solution of interval linear equations. *SIAM J. Numer. Anal.*, **32**(2): 610–630.

▶Shary, S. P. (1995b). Solving the linear interval tolerance problem. *Math. Comput. Simul.*, **39**: 53–85. Doi:10.1016/0378-4754(95)00135-K.

▶Shary, S. P. (2002). A new technique in systems analysis under interval uncertainty and ambiguity. *Reliab. Comput.*, **8**(5): 321–418. Doi:10.1023/A:1020505620702.

▶Shary, S. P. (2004). An interval linear tolerance problem. *Autom. Remote Control*, **65**(10): 1653–1666. Doi:10.1023/B:AURC.0000044274.25098.da.

▶Shary, S. P. (2012). Solvability of interval linear equations and data analysis under uncertainty. *Autom. Remote Control*, **73**(2): 310–322. Doi:10.1134/S0005117912020099.

▶Shary, S. P. (2016). Maximum consistency method for data fitting under interval uncertainty. *J. Glob. Optim.*, **66**(1): 111–126. Doi:10.1007/s10898-015-0340-1.

Shary, S. P. (2017). Strong compatibility in data fitting problems with interval data. *Bull. South Ural State Univ. Ser. "Math. Mech. Phys."*, **9**(1): 39–48. (in Russian). Doi:10.14529/mmph170105.

Shary, S. P. (2018). Interval regularization for imprecise linear algebraic equations. Preprint, arXiv:1810.01481 [math.NA]. Available at: https://arxiv.org/abs/1810.01481.

Shary, S. P. (2019). *Finite-Dimensional Interval Analysis*. Institute of Computational Technologies SB RAS, Novosibirsk (in Russian). Available at: http://www.nsc.ru/interval/Library/InteBooks/SharyBook.pdf.

Shary, S. P. and Sharaya, I. A. (2013). Recognizing solvability of interval equations and its application to data analysis. *Comput. Technol.*, **18**(3): 80–109 (in Russian).

Shary, S. P. and Sharaya, I. A. (2016). On solvability recognition for interval linear systems of equations. *Optim. Lett.*, **10**(2): 247–260.

Shor, N. Z. and Zhurbenko, N. G. (1971). Minimization method using operation of space dilatation in the direction of difference of two sequential gradients. *Kibernetika*, 3: 51–59 (in Russain).

Stetsyuk, P. I. (2014). *Ellipsoids Methods and r-Algorithms*. Eureca, Chişinâu, Moldova (in Russian).

Stetsyuk, P. I. (2017). Subgradient methods `ralgb5` and `ralgb4` for minimization of ravine convex functions. *Comput. Technol.*, **22**(2): 127–149 (in Russian).

Vorontsova, E. (2016). Extended separating plane algorithm and NSO-solutions of PageRank problem. *Proc. 9th Int. Conf., Discrete Optimization and Operations Research*, eds. Y. Kochetov, M. Khachay, V. Beresnev, E. Nurminski and P. Pardalos, Lecture Notes in Computer Science, Vol. 9869. Springer International, Cham, Switzerland, pp. 547–560. Doi:10.1007/978-3-319-44914-2.43.

Vorontsova, E. A. (2017). Linear tolerance problem for input–output models with interval data. *Comput. Technol.*, **22**(2): 67–84 (in Russian).

Zhilin, S. I. (2005). On fitting empirical data under interval error. *Reliab. Comput.*, **11**(5): 433–442. Doi:10.1007/s11155-005-0050-3.

Zhilin, S. I. (2007). Simple method for outlier detection in fitting experimental data under interval error. *Chemom. Intell. Lab. Syst.*, **88**(1): 60–68.