Interval Computations
No 4(6), 1992

SOME QUESTIONS OF APPLICATION OF INTERVAL MATHEMATICS IN PARAMETER ESTIMATION AND DECISION MAKING

A.P.Voshchinin

In the paper the application of interval mathematics to the practice problem solving in which a solution vector $X \in \mathbb{R}^n$ depends on interval input data. These problems are investigated from three points of view. The application of regression analysis is studied, as well as that of interval regression analysis. The problem of comparing and decision-making under interval indefiniteness is considered.

НЕКОТОРЫЕ ВОПРОСЫ ПРИМЕНЕНИЯ ИНТЕРВАЛЬНОЙ МАТЕМАТИКИ В ПАРАМЕТРИЧЕСКОМ ОЦЕНИВАНИИ И ПРИНЯТИИ РЕШЕНИЙ

А.П.Вощинин

В работе рассматривается применение интервальной математики к решению практических задач, в которых вектор решения $X \in \mathbb{R}^n$ зависит от интервальных входных данных. Такие задачи исследуются с трех точек зрения. Анализируется применение регрессионного анализа и интервального регрессионного анализа, а также рассматривается проблема сравнения и принятия решений в условиях интервальной неопределенности.

0. Introduction

There are many practical applications of interval mathematics for solving problems with interval input data when real values and operations with them are replaced resp. by real compact intervals:

$$[a] \triangleq [a] \in \mathbb{IR} \triangleq [\underline{a}; \overline{a}] \triangleq \{a \in R : \underline{a} \leq a \leq \overline{a}\} \tag{1}$$

and by interval operations $[a] * [b]$ where $* \in \{+, -, \times, /\}$:

$$[a] + [b] = [\underline{a} + \underline{b}, \overline{a} + \overline{b}]; \tag{2}$$

$$[a] - [b] = [\underline{a} - \overline{b}, \overline{a} - \underline{b}]; \tag{3}$$

$$[a] \times [b] = [\mathrm{Min}\{\underline{a} \cdot \underline{b}, \underline{a} \cdot \overline{b}, \overline{a} \cdot \underline{b}, \overline{a} \cdot \overline{b}\}, \mathrm{Max}\{\underline{a} \cdot \underline{b}, \underline{a} \cdot \overline{b}, \overline{a} \cdot \underline{b}, \overline{a} \cdot \overline{b}\}]; \tag{4}$$

$$[a]/[b] = [\underline{a}, \overline{a}] \times [1/\overline{b}, 1/\underline{b}]. \tag{5}$$

In all these cases the property of isotonicity is satisfied

$$([a], [b] \in \mathbb{IR}, a \in [a], b \in [b]) \rightarrow (a * b \in [a] * [b]). \tag{6}$$

Using interval operations (2)–(4) one can easily define the operations with interval vectors $[\mathbf{c}] \in \mathbb{IR}^n$ and interval matrices $[\mathbf{A}] \in \mathbb{IR}^{mn}$:

$$[b][\mathbf{A}] \triangleq ([b] \times [a_{ij}]), \ [b] \in \mathbb{IR}, \ [\mathbf{A}] \in \mathbb{IR}^{mn}; \tag{7}$$

$$[\mathbf{A}] + [\mathbf{B}] \triangleq ([a_{ij}] + [b_{ij}]), \ [\mathbf{A}], [\mathbf{B}] \in \mathbb{IR}^{mn}; \tag{8}$$

$$[A][B] \triangleq (\sum_{l=1}^{p} [a_{il}] \times [b_{lj}]), \ [\mathbf{A}] \in \mathbb{IR}^{mp} \ [B] \in \mathbb{IR}^{pn}. \tag{9}$$

The tools for interval mathematics [1] are widely used for solving of many practical problems when the solution vector $x \in \mathbb{R}^n$ depends on interval input data, i.e. $\mathbf{x} \triangleq \mathbf{x}(\mathbf{A}, \mathbf{b}), \ \mathbf{A} \in [\mathbf{A}], \ \mathbf{b} \in [b]$.

In particular case of interval linear system

$$[\mathbf{A}]\mathbf{x} = [\mathbf{b}], \ [\mathbf{A}] \in \mathbb{IR}^{nn}, [\mathbf{b}] \in \mathbb{IR}^n \tag{10}$$

the solution set is defined as

$$\mathbf{X}([\mathbf{A}], [\mathbf{b}]) = \{\mathbf{x} \in \mathbb{IR}^n : \ \mathbf{Ax} = \mathbf{b}, \ \mathbf{A} \in [\mathbf{A}], \ \mathbf{b} \in [\mathbf{b}]\}. \tag{11}$$

In general case the solution set

$$\Xi \triangleq \{\mathbf{x} \in \mathbb{R}^n : \ \mathbf{x} = \mathbf{x}(\mathbf{A}, \mathbf{b}), \ \mathbf{A} \in [\mathbf{A}], \ b \in [\mathbf{b}]\} \qquad (12)$$

may have a complete structure and because of this the corresponding methods have to be used for computing of its outer and inner bounds, i.e. such interval vectors resp. Ξ^+ and Ξ^-, that $\Xi^- \subseteq \Xi \subseteq \Xi^+$.

When solving concrete practical problem the answers on the following questions should be given:

1. What is the model of errors and how to define the bounds of errors?

2. What is the special structure of input data and what is the appropriate interval method for computing?

3. How do we interpret the solution set Ξ?

Two problems are discussed below from this point of view.

1. Regression analysis

The basic concepts

Problem definition. Using experimental data $\mathbf{x}^i, y_i, i = 1, \ldots, n$, to estimate of unknown parameter vector c of the given function

$$f(\mathbf{x}, c) = c_1\varphi_1(\mathbf{x}) + c_2\varphi_2(\mathbf{x}) \ldots + c_m\varphi_m(\mathbf{x}) = c^T\varphi(\mathbf{x}), c \in \mathbb{R}^m, \ \mathbf{x} \in \mathbb{R}^k \qquad (13)$$

under following assumptions:

$H-1$: output values y_i are received with additive errors e_i

$$y_i = f(\mathbf{x}^i, c) + e_i; \qquad (14)$$

$H-2$: e_i are random normal distributed errors such as

$$M(e_i) = 0, \ M(e_ie_j) = 0, \ M(e_i^2) = \sigma^2, \ \forall i, j = 1, \ldots, n; \qquad (15)$$

$H-3$: vector of basic functions $\varphi(\mathbf{x}) = \{\varphi_1(\mathbf{x}), \varphi_2(\mathbf{x}), \ldots, \varphi_m(\mathbf{x})\}$ is known.

Under these assumptions the OLS-method (Ordinary Least Squares) provides the optimal estimation c_{ols} which can be received as a solution of linear system

$$\mathbf{A}\cdot\mathbf{c}_{\text{ols}} = \mathbf{b}, \tag{16}$$

where $\mathbf{A} \triangleq \mathbf{F}^t\mathbf{F} \in \mathbb{R}^{mm}$, $\mathbf{b} = \mathbf{F}^t\mathbf{y} \in \mathbb{R}^m$, $\mathbf{F} \triangleq \mathbf{F}(\mathbf{x}) = \{f_{ij} = \varphi_j(\mathbf{x}^i), i = \overline{1,n}; j = \overline{1,m}\}$.

As a result of regression analysis we have the optimal estimation $\mathbf{c}_{\text{ols}}$ and its covariance matrix D, i.e.

$$\mathbf{c}_{\text{ols}} = \mathbf{A}^{-1}\mathbf{b}, \qquad D(\mathbf{c}_{\text{ols}}) = \sigma^2\mathbf{A}^{-1}. \tag{17}$$

The confidence set Ω_{ols} and its outer bounds $[\mathbf{c}]_{\text{ols}}$ for given probability α and known variance σ^2 can be also obtained for truth vector $\mathbf{c}$

$$\Omega_{\text{ols}} = \left\{\mathbf{c} \in \mathbb{R}^m :\ (\mathbf{c} - \mathbf{c}_{\text{ols}})^T \cdot \mathbf{A} \cdot (\mathbf{c} - \mathbf{c}_{\text{ols}}) \le \sigma^2\chi^2(\alpha, m)\right\}, \qquad P(\mathbf{c} \in \Omega_{\text{ols}}) \le \alpha. \tag{18}$$

$$[\mathbf{c}]_{\text{ols}} = \left\{[c_i] = [c^i_{\text{ols}} - \chi(\alpha,m)(d_{ii})^{1/2}, \quad c^i_{\text{ols}} + \chi(\alpha,m)(d_{ii})^{1/2}],\ i = \overline{1,m}\right\}, \tag{19}$$

where $\chi^2(\alpha, m)$ is corresponding quantity of Pearson's distribution, d_{ii} is diagonal element of covariance matrix D.

Using the interval approach to regression model we can formally replace matrix $\mathbf{A}$ and vector $\mathbf{b}$ in (16) by resp. interval matrix $[\mathbf{A}]$ and interval vector $[\mathbf{b}]$ and to receive an interval linear system

$$[\mathbf{A}]\mathbf{c}_{\text{ols}} = [\mathbf{b}], \quad [\mathbf{A}] \in \mathbb{IR}^{mm},\ [\mathbf{b}] \in \mathbb{IR}^m, \tag{20}$$

but as it was mentioned above in this case it is necessary to define the sources, type and bounds of errors, structure of data, etc.

Model of errors. Analyzing the possible ways of collecting, handling and transmission of experimental data from object to computer two cases should be distinguished.

For the first one there is direct connection between object and computer through A/D card, i.e.

object — A/D card — computer.

In the second case data collects and transmits according with the scheme which is included the user as one of components:

object — instruments — user — computer.

Obviously each component within both these schemes is a source of different errors which will be noted as: e^0 (object), e^t (A/D card), e^i (instruments), e^u (user), e^c (computer). It is easily to state that:

e^0 are an integrative result of many errors and noises and consequently they are random normally distributed errors;

e^t are bounded A/D transform errors;

e^u, e^c are bounded rounding errors of respective component.

Let $\Delta^t, \Delta^i, \Delta^u, \Delta^c$ denote such values that for bounded errors take place

$$| e^t | \leq \Delta^t, \; | e^i | \leq \Delta^i, \; | e^u | \leq \Delta^u, \; | e^c | \leq \Delta^c. \qquad (21)$$

By using of corresponding specifications and information from user it is possible to define these values and respectively interval matrix $[\mathbf{F}]$ and interval vector $[\mathbf{y}]$ related with interval linear system (16).

It is necessary to underline that when using interval mathematics for computing OLS -estimations the random errors e_i^0 should not be taken into account. This is first of all because of their normal distribution and besides that because of OLS-method by itself involves these random errors by the optimal way, i.e. OLS-estimations are the best one in the class of all linear estimations under the random errors.

Special structure of input data. To receive interval linear system (16) for given interval matrix $[\mathbf{F}]$ and interval vector $[\mathbf{y}]$ we first have to calculate outer bounds for $\mathbf{A}$ and $\mathbf{b}$

$$[\mathbf{A}] = [\mathbf{F}]^T[\mathbf{F}], \quad [\mathbf{b}] = [\mathbf{F}]^T[\mathbf{y}] \qquad (22)$$

and then to obtain the outer bounds Ξ^+ of the solution set of interval linear system (16) taking into account dependencies of errors in first equation and the fact that matrix $\mathbf{A}$ is a square, symmetric, positive defined and as a rule bad-conditioned matrix. The most appropriated method for defining of outer bounds Ξ^+ under these conditions is described in [2].

Interpretation of the solution. Obviously the outer set solution Ξ^+ reflects the influence of the bounded errors e^t, e^i, e^u, e^c and conversion errors on OLS-estimations. Comparing the sets $[\mathbf{c}]_{\text{ols}}$ and Ξ^+ it is easily for user to evaluate the effect of errors mentioned above and to make corresponding decision. In particular, if $[c_i]_{\text{ols}} \cong \Xi_i^+$ it means that the influence of rounding errors on OLS-estimations is very essential.

2. Regression analysis with interval data

This problem is similar to one mentioned above. The difference is that vector $\mathbf{y}$ is given in interval form $[\mathbf{y}]$ and it is assumed that unknown truth output value in every trial belongs to the given interval $[\mathbf{y}_i]$, i.e.

$$\underline{y}_i \le f(\mathbf{x}^i, c) \le \overline{y}_i, \ \forall i = 1, \dots, n. \tag{23}$$

Using vector notations we can rewrite (23) in form

$$\mathbf{Fc} = [\mathbf{y}], \ \mathbf{F} \in \mathbb{R}^{nm}, \ [\mathbf{y}] \in \mathbb{IR}^n, \tag{24}$$

which presents overdefined interval linear system with the solution set

$$\Omega = \left\{ \mathbf{c} \in \mathbb{R}^m : \ \underline{y}_i \le f(\mathbf{x}^i, \mathbf{c}) \le \overline{y}_i, \ \forall i = 1, \dots, n \right\} \tag{25}$$

As every solution $\mathbf{c} \in \Omega$ provides a model $f(\mathbf{x}, \mathbf{c})$ which passes through all interval output measurements it can be consider as set of possible values of the unknown parameters $\mathbf{c}$.

It can be proved that:

1. If rank $\mathbf{F} < m$ (in particular, when $n < m$) Ω is unlimited set;
2. If rank $\mathbf{F} = m$ and $\Omega \neq \emptyset$ the solution set Ω is compact;
3. If $m = n$, rank $\mathbf{F} = m$ the solution set can be presented in the form

$$\Omega = \{ \mathbf{c} \in \mathbb{R}^m : \ \mathbf{c} = \mathbf{F}^{-1}\mathbf{y}, \ \mathbf{y} \in [\mathbf{y}] \} \tag{26}$$

and Ω is symmetric convex politope with 2^k top points and with the central point $\overline{c} = \mathbf{F}^{-1} \cdot \overline{\mathbf{y}}$, where $\mathbf{y}$ denotes the midpoint of $[\mathbf{y}]$.

Using ***interval approach*** under the assumptions that the errors in F are much smaller then in y we can state the following.

Model of errors. The sources of errors are similar to the case of regression analysis. The errors in output vector y are bounded errors which include all errors related respectively with object (e^0), computer (e^c), A/D card (e^t) or instruments (e^i) and user (e^u). The bounds of these errors should be given by the user.

Special structure of input data. As a rule $\mathbf{F}$ is bad-conditioned matrix with $n \geq m$ and rank $\mathbf{F} = \mathrm{m}$. As we have assumed that there are no errors in matrix $\mathbf{F}$ its structure should not be taken into account. Obviously there are no dependencies between components of vector $[\mathbf{y}]$.

If $\mathbf{F}$ is $m \times m$ matrix and rank $\mathbf{F} = m$ any general method for interval linear system can be used for computing of outer set $[\mathbf{c}] \supseteq \Omega$.

If $\mathbf{F}$ is $n \times m$ matrix rigorous outer bounds for Ω can be obtained as a solution of $2m$ linear programming problems [2]

$$\begin{aligned} &\min(\max) c_i \\ &\text{subject to } \underline{y}_i \leq f(\mathbf{x}^i, \mathbf{c}) \leq \overline{y}_i, \ \forall i = 1, \dots, n. \end{aligned} \tag{27}$$

Interpretation of solution. The outer set solution $[\mathbf{c}]$ is a set of possible values of the unknown parameters $\mathbf{c}$ with the influence of all bounded errors e^t, e^i, e^u, e^c. Analyzing this set user can evaluate a sensitivity of parameter vector to these errors. If, for instance, $0 \in [c_i]$ it is possible to accept the hypothesis $H : c_i = 0$ and consequently to "turn to zero" corresponding coefficient in $f(\mathbf{x}, c)$.

When model will be used for prediction of y outer set solution $[\mathbf{c}]$ allows to receive the guaranteed error of prediction. If model will be included as a goal function in some optimization problem the inner set solution can be also useful.

2. Comparison

In many practical applications it is necessary to compare intervals and to choose one of them according to some criteria. For example let be defined an optimization problem

$$[f(\mathbf{x})] \xrightarrow{\mathbf{x} \in \mathbf{X}} \min, \tag{28}$$

where $[f(\mathbf{x})] \triangleq [\underline{f}(\mathbf{x}), \overline{\overline{f}}(\mathbf{x})]$ is known interval function, $\mathbf{c} \in \mathbb{IR}^m$, $\mathbf{x} \in \mathbb{R}^k$, $\mathbf{X} = \{\mathbf{x}^1, \dots, \mathbf{x}^i, \dots, \mathbf{x}^n\}$ is discrete set of feasible solutions. Obviously in this case to each solution $\mathbf{x}^i$ corresponds interval estimation of the goal function $[a^i] = [f(\mathbf{x}^i)]$. The discrete set of such interval estimations we'll note as $A = \{[a^1], \dots, [a^i], \dots, [a^n]\}$. For computing the set of optimal solutions

$$\mathbf{X}^0 \triangleq \{\mathbf{x}^i \in \mathbf{X} : \ \mathbf{x}^i = \arg\min f(\mathbf{x}), \ f(\mathbf{x}) \in [f(\mathbf{x})]\}. \tag{29}$$

Interval version of corresponding methods can be used including that ones which are based on the comparison of the feasible solutions $\mathbf{x}^i \in \mathbf{X}$. In this case such partial ordering relations can be defined on $\mathbf{X}$ and on $\mathbf{A}$.

$$[a^i] < [a^j] \quad [a^i] \bigcap [a^j] = \varnothing, \quad \overline{a}^i < \underline{a}^j, \ \forall [a^i], \ [a^j] \in \mathbf{A}; \tag{30}$$

$$[a^i] \cong [a^i] \quad [a^i] \bigcap [a^j] \neq \varnothing; \tag{31}$$

$$\mathbf{x}^i < \mathbf{x}^j \quad [a^i] < [a^j], \ \forall \mathbf{x}^i, \mathbf{x}^j \in \mathbf{X}; \tag{32}$$

$$\mathbf{x}^i \cong \mathbf{x}^j \quad [a^i] \cong [a^j], \ \forall \mathbf{x}^i, \mathbf{x}^j \in \mathbf{X}. \tag{33}$$

It can be proved that relations (30) and (32) are antisymmetric and transitive relations, resp. (31) and (33) are symmetric and reflexive relations. It means that on $\mathbf{A}$ resp. $\mathbf{X}$ a partial ordering relation is defined. In this case one can define the set of optimal solutions as

$$\mathbf{X}^0 \triangleq \{\mathbf{x}^i \in \mathbf{X} : \ \exists \mathbf{x}^j < \mathbf{x}^i, \mathbf{x}^j \in \mathbf{X}\}. \tag{34}$$

It can be proved that (29) and (34) are presented the same set, i.e. $\mathbf{X}^0_1 = \mathbf{X}^0$ besides that if we define the set

$$\Upsilon(z) \triangleq \{\mathbf{z}, \mathbf{x} \in \mathbf{X} : \ \mathbf{x} \cong \mathbf{z}\}, \tag{35}$$

then $\mathbf{X}^0 \subseteq \Upsilon(\mathbf{x}^0)$ for any $\mathbf{x}^0 = \arg\min f(\mathbf{x})$, $\mathbf{x}^0 \in \mathbf{X}$, $f(\mathbf{x}) \in [f(\mathbf{x})]$.

Refering to (33) we can obtain that

$$\Upsilon(\mathbf{x}^0) = \{\mathbf{z} \in \mathbf{X} : \ [f(\mathbf{z})] \bigcap [f(\mathbf{x}^0)] \neq \varnothing\}. \tag{36}$$

Depending on how the interval function $[f(\mathbf{x})]$ is defined and what is its structure we can distinquish three cases.

1. $f(\mathbf{x})$ is *linear combination of known bounds* of interval function $[f(\mathbf{x})]$, i.e a note $f(\mathbf{x}) \in [f(\mathbf{x})]$ means that

$$f(\mathbf{x}) = \alpha \underline{f}(\mathbf{x}) + (1-\alpha)\cdot \overline{f}(\mathbf{x}), \ \alpha \in [0,1]. \tag{37}$$

It is proved that is *Pareto set for bikriteria problem*

$$\underline{\mathbf{f}}(\mathbf{x}) \xrightarrow{\mathbf{x}\in\mathbf{X}} \min, \qquad \overline{f}(\mathbf{x}) \xrightarrow{\mathbf{x}\in\mathbf{X}} \min \tag{38}$$

and $\mathbf{x}^i \cong \mathbf{x}^j$ either $[f(\mathbf{x}^i)] \subset [f(\mathbf{x}^j)]$ or $[f(\mathbf{x}^j)] \subset [f(\mathbf{x}^i)]$.

2. $f(\mathbf{x})$ *is any function which passes within known bounds,* i.e. $f(\mathbf{x}) \in [f(\mathbf{x})]$ means that $\underline{f}(\mathbf{x}) \le f(\mathbf{x}) \le \overline{f}(\mathbf{x})$, $\forall \mathbf{x} \in \mathbf{X}$. In this case

$$\mathbf{x}^i \cong \mathbf{x}^j \qquad [f(\mathbf{x}^j)] \bigcap [f(\mathbf{x}^j)] \neq \varnothing, \tag{39}$$

and the solution set is defined as

$$\mathbf{X}^0 = \{\mathbf{x} \in \mathbf{X} : \ f(\mathbf{x}) \le \arg\min \overline{f}(\mathbf{x})\}. \tag{40}$$

3. $f(\mathbf{x})$ *is linear parametrized function with interval parameters,* i.e.

$$[f(\mathbf{x})] = [c_1]\varphi_1(\mathbf{x}) + [c_2]\varphi_2(\mathbf{x}) + \ldots + [c_m]\varphi_m(\mathbf{x}) = \varphi^T(\mathbf{x})[\mathbf{c}]. \tag{41}$$

In this case the relation (33) can be present in the form

$$\mathbf{x}^i \cong \mathbf{x}^j \qquad \exists \mathbf{c} \in [\mathbf{c}] : \{\varphi(\mathbf{x}^i) - \varphi(\mathbf{x}^j)\}^T[\mathbf{c}]. \tag{42}$$

Together with (36) it allows in some cases to evaluate the outer set solution for (28).

References

1. Moore, R.E. *Methods and applications of interval analysis.* SIAM, Philadelphia, 1979.
2. Jansson, C. *Rigorous sensitivity analysis for real symmetric matrices with uncertain data.* Computer Arithmetic, IMACS, 1991.
3. Voshchinin, A.P., Botchkov, A.F. and Sotirov, G.R. *Method of data analysis with a nonstatistical interval errors.* Indusrtial Laboratory **56** (7) (1990).