On the Problem of Variability of Interval Data

Václav Finěk
***Ctirad Matonoha**

Technical University of Liberec
Faculty of Science, Humanities and Education
Department of Mathematics and Didactics of Mathematics
Studentská 2, 461 17, Liberec, Czech Republic
vaclav.finek@tul.cz

*Institute of Computer Science
Academy of Sciences of the Czech Republic
Pod Vodárenskou věží 2, 182 07 Prague, Czech Republic
matonoha@cs.cas.cz

Abstract

In this contribution we will deal with the problem of variability of interval data. The issue is to find lower and upper bounds for the interval of possible values for the variance of given interval data. This leads to the problem of minimizing/maximizing the sum of squares of the distance between the components of the $n-$dimensional vector and their average value. As the maximization is more difficult than the minimization, we present here some theoretical results concerning the solution. Furthermore, we introduce preliminary algorithms for solving both problems which take into consideration their special structure.

Keywords: Interval data; computation of variance; theoretical analysis.

Introduction

In some cases, we have only intervals $[a_i, b_i]$ of possible values of x_i instead of the actual value x_i. For example, measured values usually include some measurement error with known upper bounds. Then, the actual value of x_i is unknown and we only know that its value is located within the interval determined by the upper bound of the measurement error. Therefore, we should work rather with these intervals than with these single values. Consequently, possible values of their average and their variance are also intervals. For more details on interval data see in [1]. While the computation of lower and upper bounds for the average of interval data is straightforward, the computation of lower and upper bounds for their variance is significantly complicated. In [4], it was proved that computing the variance for interval data is NP-hard. They also proposed algorithms with quadratic complexity for computing the lower bound of variance and for computing its upper bound in some special cases. Further interesting results were recently published in [2, 3, 7]. In this contribution, we present some theoretical results concerning the solution of a maximization problem and introduce preliminary algorithms for solving both problems which use their special structure.

We consider n intervals $I_i = [a_i, b_i]$ and define

$$K = I_1 \otimes I_2 \otimes I_2 \otimes \cdots \otimes I_n. \tag{1}$$

We would like to find:

$$\mathbf{x}^{min} = \arg\min_{\mathbf{x}\in K} \frac{1}{n} F(x), \tag{2}$$

$$\mathbf{x}^{max} = \arg\max_{\mathbf{x}\in K} \frac{1}{n} F(x), \tag{3}$$

where $F(x) = \sum_{i=1}^{n} (x_i - \overline{\mathbf{x}})^2$ with $\overline{\mathbf{x}} = \frac{1}{n}\sum_{i=1}^{n} x_i$ and the tensor product K of the given intervals forms the feasible set of optimization problems (2) and (3).
Note that the function F can be written in the following forms:

$$F(x) = \sum_{i=1}^{n} (x_i - \overline{\mathbf{x}})^2 = \sum_{i=1}^{n} x_i^2 - \frac{(x_1 + \cdots + x_n)^2}{n} = \sum_{i=1}^{n} x_i^2 - n\overline{\mathbf{x}}^2 = \frac{1}{n}\left(\sum_{j<i} (x_i - x_j)^2\right)$$

Problem (2) possesses a convex objective function F, thus any local solution is also a global one. In problem (3), the function F is concave and the solution has to be a vertex of K. This results in a rather different behavior. The analysis of this second case is discussed in Section 3. There are several apparent properties:

- If $\cap_i I_i \neq \emptyset$ then there exist either one or infinitely many solutions of (2). All elements of the solution are the same and belong to $\cap_i I_i$.
- The solution of (3) lies on the vertex of K.
- It holds $\sum_{i=1}^{n} (x_i - \overline{\mathbf{x}})^2 < \sum_{i=1}^{n} (x_i - d)^2$ for any $d \neq \overline{\mathbf{x}}$.

In this contribution we give some theoretical results of problems (2) and (3), respectively, and outline the algorithms for seeking the solutions. We will consider the following structure of those problems:

$$\begin{aligned} & \mathbf{x}^{min} = \arg\min_{x} F(x), \\ \text{subject to} \quad & x_i \in [a_i, b_i], \quad i = 1 \ldots n \end{aligned} \tag{4}$$

$$\begin{aligned} & \mathbf{x}^{max} = \arg\max_{x} F(x), \\ \text{subject to} \quad & x_i \in [a_i, b_i], \quad i = 1 \ldots n \end{aligned} \tag{5}$$

The following quantities are used throughout the paper:

$$c_i = \frac{a_i + b_i}{2}, \quad d_i = b_i - a_i > 0, \quad i = 1 \ldots n,$$

$$p_a = \frac{a_1 + \cdots + a_n}{n}, \quad p_b = \frac{b_1 + \cdots + b_n}{n}.$$

1 The Problem of Minimization

Both problems (4) and (5), respectively, are classical optimization problems with simple bounds that can be solved by a suitable optimization method, e.g. a variable metric method or trust-region method, see e.g. [6]. The basic optimization method is an iteration process starting from an initial point $x^{(0)}$ and generating a sequence of points $x^{(1)}, x^{(2)}, \ldots$ such that

$$x^{(k+1)} = x^{(k)} + \alpha^{(k)} d^{(k)},$$

where $d^{(k)}$ is a direction vector and $\alpha^{(k)} > 0$ is a step-length. The direction vector is determined on the basis of values $x^{(j)}, F(x^{(j)}), F'(x^{(j)}), F''(x^{(j)}),\ 0 \leq j \leq k,$ and the step-length is determined on the basis of behavior of the function F in the neighborhood of $x^{(k)}$.

Unfortunately, the Hessian matrix of F is dense which makes difficult to solve problems (4) and (5), respectively, for large n by standard optimization methods. We will use of a special structure of the problem and introduce algorithms that take into consideration this structure.

The problem of minimization (4) is simple and no significant difficulties arise. We need not to know anything special about the solution and the following algorithm works well on the test problems.

Algorithm 1 *Solving problem (4).*

1. **Initiation:** Set
$$p_a = \o a_i, \quad p_b = \o b_i, \quad p = \frac{p_a + p_b}{2}, \quad A^{\max} = \max\{a_i\}, \quad B_{\min} = \min\{b_i\},$$
$s_i = 0\ \forall i$ (indicator of the solution x_i^*: 0 - not found, 1 - found).

2. **Test on intersection of all intervals:** If $A^{\max} \leq B_{\min},$ then the solution $x_i^*\ \forall i$ is an arbitrary number in $[A^{\max}, B_{\min}]$ and $F = 0$.

3. Until we have not found all x_i^* (i.e. $\exists i$ such that $s_i = 0$):

   (a) **Iteration process:** For $i = 1 \dots n$ such that $s_i = 0$:
   If $p \in [a_i, b_i],$ set $x_i = p;$ else if $p < a_i,$ set $x_i = a_i;$ else set $x_i = b_i$.

   (b) **Reducing of all intervals:** For $i = 1 \dots n$ such that $s_i = 0$:
   If $a_i < p_a \leq b_i,$ set $a_i = p_a;$ if $b_i > p_b \geq a_i,$ set $b_i = p_b$.

   (c) **Test on components of the solution:** For $i = 1 \dots n$ such that $s_i = 0$:
   If $[a_i, b_i] \cap [p_a, p_b] = \emptyset$ or a_i or $b_i,$ set $a_i = b_i = x_i^* = x_i$ and $s_i = 1$.

4. **Computation of new values:** Set $p_a^{old} = p_a,\ p_b^{old} = p_b$ and update p_a, p_b.

5. **Test on termination:** If $\max\{|p_a - p_a^{old}|, |p_b - p_b^{old}|\} > \varepsilon$ (e.g. $= 10^{-6}$), update p and goto Step 3. Otherwise set $x_i^* = x_i$ for i such that $s_i = 0$.

2 The Problem of Maximization

The maximization problem is much harder than the minimization problem, so we will focus on the theoretical analysis and find useful information concerning the solution. The solution x^* to (5) lies on the vertex of K, see the second apparent property above or Lemma 1 below. Thus in the subsequent analysis we assume that x^* has components x_i^* equal to either a_i or b_i. First, we will study what the solution must satisfy.

Lemma 1 *The solution x^* of (5) lies on the vertex of K, i.e. $x^* = (x_1^* \dots x_n^*)$, where $x_i^* = a_i$ or $x_i^* = b_i$.*

Lemma 2 *The solution x^* of (5) has the property that there exists at least one i and at least one j, $i \neq j$ such that $x_i^* = a_i$ and $x_j^* = b_j$.*

The following lemma says that no component x_i^* of the solution x^* is equal to the average value $\overline{\mathbf{x}^*}$.

Lemma 3 *Let* $x \in \mathbb{R}^n$ *and* $\bar{\mathbf{x}}$ *be the average value of componenets* $x_1 \dots x_n$. *Suppose that there exists an index j such that* $x_j = \bar{\mathbf{x}}$. *Then if we take a component* $x_j + \tau$ *instead of* x_j *for some* $\tau \neq 0$, *we obtain a greater function value, i.e.* $F(x) < F(x_\tau)$, *where*

$$x_\tau = (x_1 \dots x_{j-1},\ x_j + \tau,\ x_{j+1} \dots x_n).$$

The following analysis concerns the differences among function values. The first lemma stands for the most general case.

Lemma 4 *Let* x, y *be arbitrary points. Denote* p_{xy} *the average value of all points* x_i, y_i, $i = 1 \dots n$, *and define the following sets:*

$$N_{aa} = \{i:\ x_i = a_i,\ y_i = a_i\},$$

$$N_{ab} = \{i:\ x_i = a_i,\ y_i = b_i\},$$

$$N_{ba} = \{i:\ x_i = b_i,\ y_i = a_i\},$$

$$N_{bb} = \{i:\ x_i = b_i,\ y_i = b_i\}.$$

It is evident that $N_{aa} \cup N_{ab} \cup N_{ba} \cup N_{bb} = \{1 \dots n\}$. *Then it holds*

$$F(x) - F(y) = 2\left[\sum_{i \in N_{ba}} d_i(c_i - p_{xy}) - \sum_{i \in N_{ab}} d_i(c_i - p_{xy})\right]$$

Suppose that we have some combination of points $x_1 \dots x_n$. Now we fix some j and ask if the function value is greater for $x_j = a_j$ or $x_j = b_j$.

Lemma 5 *Let* $x \in \mathbb{R}^n$ *and take an arbitrary index j. Denote* p_{x_i} *the average value of points* $\{x_i\}_{i \neq j}$. *Then*

$$F(\{x_i\}_{i \neq j}, b_j) - F(\{x_i\}_{i \neq j}, a_j) = 2\frac{n-1}{n} d_j(c_j - p_{x_i})$$

The consequence is that:

$$F(\{x_i\}_{i \neq j}, b_j) > F(\{x_i\}_{i \neq j}, a_j) \quad \Leftrightarrow \quad c_j > p_{x_i},$$

$$F(\{x_i\}_{i \neq j}, b_j) < F(\{x_i\}_{i \neq j}, a_j) \quad \Leftrightarrow \quad c_j < p_{x_i},$$

$$F(\{x_i\}_{i \neq j}, b_j) = F(\{x_i\}_{i \neq j}, a_j) \quad \Leftrightarrow \quad c_j = p_{x_i}.$$

The special case of this lemma is the following result.

Lemma 6 *Consider sets* $\{a_i\}, \{b_i\}$ *and let j be an arbitrary index. Then*

$$F(\{a_i\}_{i \neq j}, b_j) - F(\{a_i\}) = 2d_j(c_j - p_a - \frac{1}{2n} d_j),$$

$$F(\{b_i\}_{i \neq j}, a_j) - F(\{b_i\}) = 2d_j(p_b - c_j - \frac{1}{2n} d_j).$$

Both numbers on the right-hand side are equal if and only if $c_j = p := \frac{p_a + p_b}{2}$.
In general, if we compare both numbers on the right-hand side, we will derive relations

$$F(\{a_i\}_{i \neq j}, b_j) - F(\{a_i\}) > F(\{b_i\}_{i \neq j}, a_j) - F(\{b_i\}) \quad \Leftrightarrow \quad c_j > p,$$

$$F(\{a_i\}_{i\neq j},b_j)-F(\{a_i\})<F(\{b_i\}_{i\neq j},a_j)-F(\{b_i\}) \quad\Leftrightarrow\quad c_j<p,$$
$$F(\{a_i\}_{i\neq j},b_j)-F(\{a_i\})=F(\{b_i\}_{i\neq j},a_j)-F(\{b_i\}) \quad\Leftrightarrow\quad c_j=p.$$

As Lemma 2 says that the solution must contain at least one a_i and at least one b_j, the consequence is that

a_j can be replaced with b_j if and only if $p_a<c_j-\frac{1}{2n}d_j,$

b_j can be replaced with a_j if and only if $p_b>c_j+\frac{1}{2n}d_j.$

The following lemma gives the comparison of function values on each side of the set K.

Lemma 7 *It holds*

$$F(a_1,\dots,a_n)<F(b_1,\dots,b_n) \quad\Leftrightarrow\quad p<\frac{\sum(c_id_i)}{\sum d_i},$$
$$F(a_1,\dots,a_n)>F(b_1,\dots,b_n) \quad\Leftrightarrow\quad p>\frac{\sum(c_id_i)}{\sum d_i},$$
$$F(a_1,\dots,a_n)=F(b_1,\dots,b_n) \quad\Leftrightarrow\quad p=\frac{\sum(c_id_i)}{\sum d_i}.$$

Lemma 8 *Let j be an arbitrary index. Then*

$$F(\{a_i\}_{i\neq j},b_j)<F(\{b_i\}_{i\neq j},a_j) \quad\Leftrightarrow\quad (c_j-p)d_j<\textstyle\sum_{i\neq j}[(c_i-p)d_i],$$
$$F(\{a_i\}_{i\neq j},b_j)>F(\{b_i\}_{i\neq j},a_j) \quad\Leftrightarrow\quad (c_j-p)d_j>\textstyle\sum_{i\neq j}[(c_i-p)d_i],$$
$$F(\{a_i\}_{i\neq j},b_j)=F(\{b_i\}_{i\neq j},a_j) \quad\Leftrightarrow\quad (c_j-p)d_j=\textstyle\sum_{i\neq j}[(c_i-p)d_i].$$

Another approach to determine Lemma 5 consists in that we fix a set of points $\{x_i\}_{i\neq j}$ and study the function values on $[a_j,b_j]$. Denote

$$x_j=a_j+\tau_j(b_j-a_j)=a_j+\tau_jd_j, \quad \tau_j\in[0,1].$$

Then we have

$$f(\tau_j)\equiv F(\{x_i\}_{i\neq j},x_j)=\sum_{i\neq j}x_i^2+(a_j+\tau_jd_j)^2-\frac{1}{n}\left(\sum_{i\neq j}x_i+a_j+\tau_jd_j\right)^2$$

$$f'(\tau_j)=2d_j(a_j+\tau_jd_j)-\frac{2}{n}d_j\left(\sum_{i\neq j}x_i+a_j+\tau_jd_j\right)$$

$$f'(\tau_j)=0 \quad\Leftrightarrow\quad \tau_j^*=\frac{p_{x_i}-a_j}{b_j-a_j}$$

$$f''(\tau_j)=2\frac{n-1}{n}d_j^2 \quad\Rightarrow\quad f(\tau_j^*)=\min f(\tau_j)$$

From here we obtain properties mentioned in Lemma 5, that is

$$\tau_j^*<0.5 \quad\Leftrightarrow\quad c_j>p_{x_i} \quad\Leftrightarrow\quad F(\{x_i\}_{i\neq j},b_j)>F(\{x_i\}_{i\neq j},a_j)$$
$$\tau_j^*>0.5 \quad\Leftrightarrow\quad c_j<p_{x_i} \quad\Leftrightarrow\quad F(\{x_i\}_{i\neq j},b_j)<F(\{x_i\}_{i\neq j},a_j)$$
$$\tau_j^*=0.5 \quad\Leftrightarrow\quad c_j=p_{x_i} \quad\Leftrightarrow\quad F(\{x_i\}_{i\neq j},b_j)=F(\{x_i\}_{i\neq j},a_j)$$

Remark 1 *Function $f(\tau_j)$ satisfies*

$$f(\tau_j) = \frac{n-1}{n} d_j^2 \tau_j^2 - 2\frac{n-1}{n} d_j (p_{x_i} - a_j)\tau_j + \sum_{i\neq j} x_i^2 + a_j^2 - \frac{1}{n}\left(\sum_{i\neq j} x_i + a_j\right)^2 .$$

From Lemma 2 and Lemma 5 we can immediately determine some components of the solution. Denote

$p_j^a =$ the average value of $\{a_i\}_{i\neq j},\ j = 1\dots n$

$p_j^b =$ the average value of $\{b_i\}_{i\neq j},\ j = 1\dots n$

$p_\alpha =$ the average value of $\{\{a_i\}_{i\neq k}, b_k\}$, where k is such that $d_k = \min_i\{d_i\}$.

$p_\beta =$ the average value of $\{\{b_i\}_{i\neq k}, a_k\}$, where k is such that $d_k = \min_i\{d_i\}$.

Then it holds

If $c_j < \max\{p_j^a, p_\alpha\}$, then $x_j^* = a_j$.

If $c_j > \min\{p_j^b, p_\beta\}$, then $x_j^* = b_j$.

If $c_j \in P_j = [\max\{p_j^a, p_\alpha\}, \min\{p_j^b, p_\beta\}]$, an iteration process must be performed.

Now we give a preliminary algorithm for solving the maximization problem (5).

Algorithm 2 *Solving problem (5).*

1. **Initiation:** Set

   $$p_a = øa_i, \quad p_b = øb_i, \quad p = \frac{p_a + p_b}{2}, \quad c_i = \frac{a_i + b_i}{2}, \quad d_i = \frac{b_i - a_i}{2n}, \quad d = 2\min d_i,$$

   $s_i = 0\ \forall i$ (indicator of the solution x_i^*: 0 - not found, 1 - found).

2. For $i = 1\dots n$ such that $s_i = 0$:

   (a) **Test on components of the solution:**
   If $c_i > p_b - d$, then $x_i^* = a_i = b_i$ and $s_i = 1$.
   If $c_i < p_a + d$, then $x_i^* = b_i = a_i$ and $s_i = 1$.

   (b) Otherwise **– Iteration process:**
   If $c_i < p$, then $x_i = a_i$ and $b_i = \min\{b_i, p_a\}$.
   If $c_i > p$, then $x_i = b_i$ and $a_i = \max\{a_i, p_b\}$.

   (c) Otherwise **– Possible reduction of intervals:**
   If $a_i < p_a$ and $p_b < b_i$, then $a_i = p_b$ or $b_i = p_a$.

3. **Computation of new values:** Set $p^{old} = p$ and update p_a, p_b, p, c_i, d_i, d.

4. **Test on termination:** If $|p - p^{old}| > \varepsilon$ (e.g. $= 10^{-6}$), goto Step 2. Otherwise set $x_i^* = x_i$ for i such that $s_i = 0$.

Example 1 Let $n = 3$ and

$$[a_i,\, b_i] = [-3,\, 1],\, [-9/2,\, 3],\, [-1,\, 1/2]$$

It holds that

$$c_1 = -1, \quad d_1 = 4, \quad p_1^a = -11/4, \quad p_1^b = 7/4$$
$$c_2 = -3/4, \quad d_2 = 15/2, \quad p_2^a = -2, \quad p_2^b = 3/4$$
$$c_3 = -1/4, \quad d_3 = 3/2, \quad p_3^a = -15/4, \quad p_3^b = 2$$
$$p_\alpha = -7/3, \quad p_\beta = 1, \quad p = -2/3$$

We do not have immediately any component of the solution because $c_j \in P_j\ \forall j$. The solution satisfies

$$x^* = (-3,\, 3,\, -1), \quad \text{ø}x_i^* = -1/3, \quad F(x^*) = 6.\overline{2}$$

and the point just on the other sides of all interval satisfies

$$\tilde{x} = (1,\, -9/2,\, 1/2), \quad \text{ø}\tilde{x}_i = -1, \quad F(\tilde{x}) = 6.1\overline{6}$$

The function values are located close together which makes problems for the algorithm to identify the right solution.

Example 2 This example shows efficiency of our method and also that using a standard optimization method (in this case the line-search approach from the UFO system [5]) the right solution has not to be obtained. Let $n = 40$ and $[a_i,\, b_i] =$

$$[-47.5,\, 28.75],\, [47.25,\, 91.75],\, [38.5,\, 81.5],\, [-53.5,\, 45],\, [-46.5,\, 93.5],$$
$$[-98.25,\, -40.75],\, [-34.5,\, -28],\, [51,\, 64],\, [-88.5,\, -71.5],\, [-93.75,\, -33],$$
$$[40.75,\, 95],\, [-47,\, 30.5],\, [-95.75,\, -25.25],\, [-34,\, -28.25],\, [-1.25,\, 34.5],$$
$$[16.5,\, 81.25],\, [-21.75,\, 39.75],\, [30.25,\, 80.25],\, [-14.5,\, -6],\, [-22.5,\, 1.25],$$
$$[-10,\, 7.75],\, [-81.5,\, -72.25],\, [-94,\, -18.75],\, [-65.5,\, 22],\, [-3.25,\, 83.25],$$
$$[-90.25,\, -77],\, [-24.75,\, -1.25],\, [42.5,\, 79.75],\, [-11,\, -5],\, [-19.5,\, 6.75],$$
$$[27.75,\, 57],\, [49.75,\, 85.5],\, [12.75,\, 90.75],\, [18.5,\, 85.5],\, [-93,\, -83],$$
$$[-95.75,\, -52],\, [-66,\, -61.25],\, [-77.5,\, -42.75],\, [-32.5,\, -13.5],\, [-42.25,\, 20].$$

In the following table, we present the obtained results with differences in bold.

x^* using Algorithm 2	x^* using standard optimization method
-47.5, 91.75, 81.5, 45, 93.5, -98.25, -34.5, 64, -88.5, -93.75, 95, **-47**, -95.75, -34, 34.5, 81.25, 39.75, 80.25, -14.5, -22.5, 7.75, -81.5, -94, **-65.5**, 83.25, -90.25, -24.75, 79.75, -11, **-19.5**, 57, 85.5, 90.75, 85.5, -93, -95.75, -66, -77.5, -32.5, **-42.25**	**28.75**, 91.75, 81.5, 45, 93.5, -98.25, -34.5, 64, -88.5, -93.75, 95, **30.5**, -95.75, -34, 34.5, 81.25, 39.75, 80.25, -14.5, -22.5, 7.75, -81.5, -94, **22**, 83.25, -90.25, -24.75, 79.75, -11, **6.75**, 57, 85.5, 90.75, 85.5, -93, -95.75, -66, -77.5, -32.5, **20**
$F = 4911.99902$, #iterations = 2	$F = 4709.79625$, #iterations = 41

Note that we have found the same solution as in [3] in the second iteration while their genetic algorithm used several hundred iterations.

Conclusion

In this contribution we have presented some theoretical results and preliminary algorithms for computing variance of interval data described in the introduction. Although the problem can be solved using various optimization methods combining direction vectors and the step-length, developing a special algorithm is advantageous. Concerning the maximization problem, the solution satisfies useful properties which allow us to identify directly some components of the solution. The main task is to properly deal with the case when $c_j \in P_j$ and develop more robust algorithm to identify the right solution.

Acknowledgments

This work has been supported by the project ESF No. CZ.1.07/2.3.00/09.0155 “Constitution and improvement of a team for demanding technical computations on parallel computers at TU Liberec” and by the long-term strategic development financing of the Institute of Computer Science (RVO:67985807).

Literature

[1] ALEFELD, G. - HERZBERGER, J.: *Introduction to interval computations.* Academic Press, New York, 1983.

[2] ANTOCH, J.; BRZEZINA, M.; MIELE, R.: A note on variability of interval data. *Comput. Stat.* 25, No. 1, pp. 143-153, 2010.

[3] ANTOCH, J.; MIELE, R.: Use of Genetic Algorithms When Computing Variance of Interval Data. *Classification and Multivariate Analysis for Complex Data Structures*, Springer, pp. 371-377, 2011, ISBN 978-3-642-13311-4.

[4] FERSON, S.; GINZBURG, L.; KREINOVICH, V.; LONGPRE, L.; and AVILES, M.: Computing Variace for Interval Data is NP-Hard. *ACM SIGACT News* 33, pp. 108-118, 2002.

[5] LUKŠAN, L.; TŮMA, M.; VLČEK, J.; RAMEŠOVÁ, N.; ŠIŠKA, M.; HARTMAN, J.; MATONOHA, C.: *UFO 2011 - Interactive system for universal functional optimization*. Technical Report V-1151, ICS AS CR, Prague 2011.

[6] NOCEDAL, J.; WRIGHT, S.J.: *Numerical Optimization. Second Edition.* Springer, New York, 2006. ISBN 978-0387-30303-1; 0-387-30303-0.

[7] XIANG, G.: Fast Algorithm for Computing the Upper Endpoint of Sample Variance for Interval Data: Case of Sufficiently Accurate Measurements. *Reliable Computing* 12, pp. 59-64, 2006.

RNDr. Václav Finěk, Ph.D., RNDr. Ctirad Matonoha, Ph.D.

Problém variability intervalových dat

V příspěvku se zabýváme problémem variability intervalových dat. Pro daná intervalová data se jedná o nalezení dolních a horních mezí intervalu možných hodnot jejich rozptylu. Toto vede na problém minimalizace/maximalizace součtu čtverců rozdílu složek *n*-dimensionálního vektoru a jejich průměrné hodnoty. Jelikož je problém maximalizace mnohem obtížnější než problém minimalizace, uvedeme některé teoretické výsledky týkající se řešení. Také uvedeme předběžné algoritmy pro nalezení řešení obou problémů, které využívají jejich speciální vlastnosti.

Das Problem der Variabilität von Intervalldaten

In diesem Beitrag befassen wir uns mit dem Problem der Variabilität von Intervalldaten. Bei den vorliegenden Intervalldaten handelt es sich um die Auffindung der Unter- und Obergrenzen der möglichen Werte ihrer Zerstreuung. Dies führt zum Problem der Minimalisierung/ Maximalisierung der Summe der Quadrate des Unterschieds der Komponenten eines *n*-dimensionalen Vektors und deren Durchschnittswert. Da das Problem der Maximalisierung viel schwieriger ist das Problem der Minimalisierung, führen wir einige theoretische Ergebnisse an, welche mit der Lösung zu tun haben, ebenso vorläufige Algorithmen zur Findung einer Lösung für beide Probleme, welche deren speziellen Eigenschaften nutzen.

Problem zmienności danych interwałowych

W artykule przedstawiono problem zmienności danych interwałowych. W przypadku określonych danych interwałowych chodzi o znalezienie dolnych i górnych granic interwału możliwych wartości ich rozproszenia. W związku z tym pojawia się problem minimalizacji/maksymalizacji sumy kwadrarów różnicy elementów *n*-wymiarowego wektora oraz ich przeciętnej wartości. Problem maksymalizacji jest o wiele trudniejszy w porównaniu z problemem minimalizacji, dlatego wskazano niektóre teoretyczne wnioski dotyczące rozwiązania. Ponadto przedstawiono wstępne algorytmy służące do znalezienia rozwiązania obu problemów, które wykorzystują ich szczególne cechy.