

Метод дробления параметров для интервальных линейных систем со связями

С.П. Шарый*

Институт вычислительных технологий СО РАН
Россия, 630090 г. Новосибирск, просп. Лаврентьева, 6
shary@ict.nsc.ru

Аннотация Работа посвящена развитию методов дробления параметров (PPS-методов) для оптимального (точного) внешнего покоординатного оценивания множеств решений интервальных линейных систем уравнений, на элементы которых наложены дополнительные связи специального вида.

1 Введение

Станем говорить, что задана *интервальная величина* (интервальный параметр), если имеется переменная, изменяющаяся в пределах некоторого интервала. Интервальные величины $x_1 \in \mathbf{x}_1, x_2 \in \mathbf{x}_2, \dots, x_n \in \mathbf{x}_n$ назовём *независимыми* (несвязанными), если кортеж из соответствующих переменных (x_1, x_2, \dots, x_n) принимает любые значения из декартова произведения интервалов их изменения $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$, т.е. из бруса $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$. В противном случае интервальные величины называются *зависимыми* или *связанными*. Будем говорить также, что на них *наложены связи*, если имеются в виду какие-то соотношения для x_1, x_2, \dots, x_n в виде равенств, неравенств и т.п.

Конкретный вид связанности (зависимости) интервальных величин удобно представлять наглядно графически на чертеже, изображающем множество всевозможных значений кортежа (x_1, x_2, \dots, x_n) на фоне декартова произведения $\mathbf{x}_1 \times \mathbf{x}_2 \times \dots \times \mathbf{x}_n$. Мы будем называть такие чертежи *диаграммами связанности* (диаграммами зависимости).

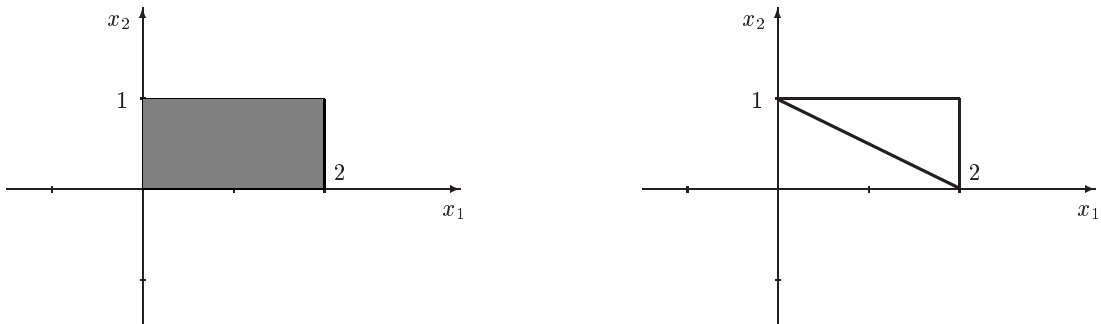


Рис. 1. Диаграммы связанности для независимых (слева) и связанных (справа) интервальных величин.

К примеру, пусть переменные x_1 и x_2 изменяются так, что имеет место равенство

$$x_1 + 2x_2 = 2. \quad (1)$$

При варьировании x_1 в пределах $[0, 2]$ переменная x_2 также изменяется в пределах $[0, 1]$, а множество их совместных значений — пар (x_1, x_2) — “заметает” диагональ в прямоугольнике $([0, 2], [0, 1])^\top$. Диаграмма

* Работа поддержана грантом Президента России №НШ-2314.2003.1.

связанности интервальных величин $x_1 \in [0, 2]$ и $x_2 \in [0, 1]$ выглядит в этом случае так, как изображено на правом чертеже Рис. 1, а для тех же самых, но независимых $x_1 \in [0, 2]$ и $x_2 \in [0, 1]$ диаграмма связанности представлялась бы полным прямоугольником на левом чертеже Рис. 1.

Связанность переменных является весьма распространённым явлением в окружающем нас мире, но в традиционном интервальном анализе задачи со связанными переменными практически не рассматривались. Это было вызвано, прежде всего, тем обстоятельством, что подобные задачи гораздо более сложны, чем обычные постановки с независимыми входными данными, и, кроме того, классическая интервальная арифметика и другие интервальные арифметики, которые являются основным инструментом интервального анализа, нацелены на обработку именно независимых величин.

Предмет представляемой работы — задачи оценивания множеств решений интервальных систем линейных алгебраических уравнений (ИСЛАУ), имеющих вид

$$\begin{cases} \mathbf{a}_{11}x_1 + \mathbf{a}_{12}x_2 + \dots + \mathbf{a}_{1n}x_n = \mathbf{b}_1, \\ \mathbf{a}_{21}x_1 + \mathbf{a}_{22}x_2 + \dots + \mathbf{a}_{2n}x_n = \mathbf{b}_2, \\ \quad \quad \quad \vdots \quad \quad \quad \ddots \quad \quad \quad \vdots \\ \mathbf{a}_{n1}x_1 + \mathbf{a}_{n2}x_2 + \dots + \mathbf{a}_{nn}x_n = \mathbf{b}_n, \end{cases} \quad (2)$$

с интервалами \mathbf{a}_{ij} и \mathbf{b}_i , или, в краткой форме,

$$\mathbf{A}x = \mathbf{b} \quad (3)$$

с интервальной $n \times n$ -матрицей $\mathbf{A} = (\mathbf{a}_{ij})$ и интервальным n -вектором правой части $\mathbf{b} = (\mathbf{b}_i)$, причём на элементы матрицы будут наложены связи некоторого частного вида. Мы подробно рассмотрим случай интервальных симметричных линейных систем, а на некоторые другие типы ИСЛАУ со связями — имеющими кососимметричные, персимметричные, ганкелевы, тёплицевы или циркулянтные матрицы — наши рассуждения переносятся с минимальными модификациями.

Интервальной симметричной линейной системой будем называть интервальную систему линейных алгебраических уравнений (2)–(3) с матрицей \mathbf{A} , симметричной относительно главной диагонали, и такой, что в её пределах также рассматриваются только вещественные матрицы $A \in \mathbf{A}$, обладающие свойством симметричности, $A = A^T$. Для интервальных систем уравнений существуют различные определения решений и множеств решений, для которых, в свою очередь, возможны те или иные способы оценивания, но ниже мы ограничимся так называемым *объединённым множеством решений* для (2)–(3), которое образовано всевозможными решениями x точечных систем $Ax = b$, когда матрица A и вектор b пробегает \mathbf{A} и \mathbf{b} соответственно, подчиняясь наложенным на них связям. Объединённое множество решений интервальной симметричной линейной системы определяется строго как

$$\Xi_{sym}(\mathbf{A}, \mathbf{b}) = \{ x \in \mathbb{R}^n \mid (\exists A \in \mathbf{A})(\exists b \in \mathbf{b})(A = A^T \text{ и } Ax = b) \}, \quad (4)$$

и мы будем называть его просто *множеством решений* интервальной линейной системы (2)–(3), так как другие множества решений нами не исследуются.

В этой заметке нас интересует нахождение для множества решений возможно более узких *внешних* оценок в виде объёмлющих множеств, в качестве которых берутся *брусы* — прямоугольные параллелепипеды в пространстве \mathbb{R}^n со сторонами, параллельными координатным осям, т.е. геометрические образы интервальных векторов. Таким образом, мы решаем задачу:

Найти (по возможности меньший) брус \mathbf{U} ,
содержащий множество решений $\Xi_{sym}(\mathbf{A}, \mathbf{b})$
интервальной линейной системы $\mathbf{A}x = \mathbf{b}$.

(5)

В свою очередь, задача (5) может быть редуцирована к следующей покомпонентной форме:

Для интервальной линейной системы уравнений $\mathbf{A}x = \mathbf{b}$ найти
как можно более точные оценки для $\min\{x_\nu \mid x \in \Xi_{sym}(\mathbf{A}, \mathbf{b})\}$
снизу и для $\max\{x_\nu \mid x \in \Xi_{sym}(\mathbf{A}, \mathbf{b})\}$ сверху, $\nu = 1, 2, \dots, n$.

(6)

При этом в постановке (6) можно даже ограничиться требованием вычисления только минимума $\min\{x_\nu \mid x \in \Xi_{sym}(\mathbf{A}, \mathbf{b})\}$, поскольку

$$\max\{x_\nu \mid x \in \Xi_{sym}(\mathbf{A}, \mathbf{b})\} = -\min\{x_\nu \mid x \in \Xi_{sym}(\mathbf{A}, -\mathbf{b})\}.$$

Ниже мы предполагаем невырожденность интервальной матрицы \mathbf{A} . Наконец, используемые нами обозначения соответствуют недавно принятому проекту международного стандарта [19].

2 Методы дробления параметров для интервальных линейных систем

Изучение интервальных линейных систем со связями и первые попытки решения для них задачи внешнего оценивания (5) восходят к началу 90-х годов прошлого века [14,16]. Предложенный Алефельдом и Майером в [14] интервальный метод Холесского является естественным интервальным расширением вещественного метода Холесского (известного также как “метод квадратного корня”), и численные эксперименты демонстрируют его невысокое качество. Далее в большом цикле работ [7,8,9,10,11,12,13] и других Алефельд, Крейнвич и Майер обратились к теоретическому исследованию ИСЛАУ с симметричными, кососимметричными, тёплицевыми и даже более общими видами связей, дав характеристику соответствующих множеств решений. Они установили, в частности, что границы множеств решений интервальных симметричных и кососимметричных линейных систем в общем случае составлены из кусков гиперплоскостей и поверхностей второго порядка (в отличие от кусочно-плоских границ множеств решений ИСЛАУ с независимыми данными). В более общем случае аффинных связей на элементы матрицы и правой части множество решений ИСЛАУ является, как было показано, в [10,11], полуалгебраическим множеством.

Для точной характеристики множеств решений ИСЛАУ со связями специального вида Алефельд, Крейнвич и Майер разработали метод исключения, родственной процессу исключения Фурье-Мощкина для систем линейных неравенств. В силу огромной трудоёмкости этого метода он имеет, главным образом, теоретическое значение: для 3×3 -системы, например, результирующая система неравенств, описывающая пересечение Ξ_{sym} с одним ортантом пространства \mathbb{R}^n имеет аж 44 неравенства [12]. Таким образом, задача практического вычисления оценок множеств решений ИСЛАУ с симметричными, кососимметричными и т.п. матрицами со связанными элементами стоит весьма остро.

Для уточнения внешних оценок множеств решений интервальных систем уравнений автором в работах [4,6,22] были предложены и развиты так называемые *методы дробления параметров*. Они основаны на часто используемой в интервальном анализе идее адаптивного (т.е. подстраивающегося под задачу и текущий результат) дробления интервальных входных данных. Для интервальных линейных систем с независимыми величинами на входе вычислительная схема методов дробления параметров выглядит особенно просто, так как экстремальные оценки множеств решений достигаются на крайних матрицах и правых частях (теорема Никеля). Дробление интервальных параметров системы сводится в этом случае, фактически, к распадению интервала на его концы.

Но для систем со связанными параметрами этот факт уже неверен. Например, для интервальной симметричной линейной системы

$$\begin{pmatrix} 1 & [0, 1] \\ [0, 1] & [-4, -1] \end{pmatrix} x = \begin{pmatrix} [0, 2] \\ [0, 2] \end{pmatrix}, \quad (7)$$

Алефельд, Крейнвич и Майер показали в [12], что её множества решений имеют вид, изображенный на Рис. 2, и, кроме того,

$$\max\{x_1 \mid x \in \Xi_{sym}(\mathbf{A}, \mathbf{b})\} = 1 + \sqrt{2},$$

причём этот максимум достигается на вещественной системе

$$\begin{pmatrix} 1 & \sqrt{2} - 1 \\ \sqrt{2} - 1 & -1 \end{pmatrix} x = \begin{pmatrix} 2 \\ 2 \end{pmatrix},$$

матрица которой не является крайней (при этом $x_2 = -1$). Таким образом, алгоритмы метода дробления параметров, описанные в [4,6,22], неприменимы напрямую для уточнения оценок множеств решений интервальных линейных систем со связями. Тем не менее, путем несложной модификации исходной схемы мы можем исправить это положение, получив в результате методы для вычисления точных (оптимальных) оценок множеств решений ИСЛАУ со связями.

Основная идея этой модификации проста и естественна:

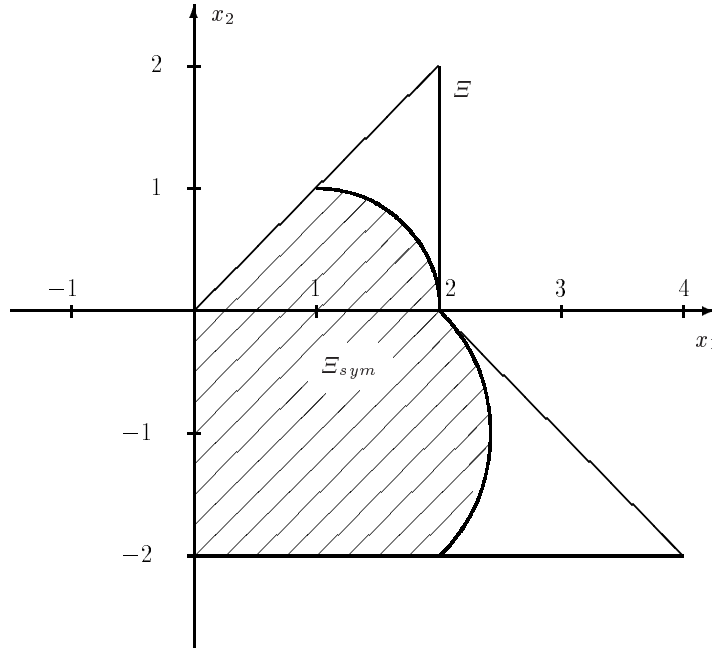


Рис. 2. Множество решений интервальной симметричной линейной системы уравнений (7) в сравнении с множеством решений такой же системы без связей.

- мы отказываемся от поконцевого дробления элементов интервальной системы и дробим, как это обычно делается, интервальные параметры системы на подинтервалы ненулевой длины, в объединении дающие исходный дробимый интервал,
- мы дробим интервальную систему так, чтобы получающиеся системы-потомки соответствовали связям, накладываемым на систему.

Например, если рассматривается интервальная симметричная система уравнений, то в единичном акте дробления должны одновременно дробиться симметричные относительно главной диагонали элементы матрицы, чтобы получающиеся ИСЛАУ снова имели интервальные симметричные системы. Если матрица исходной системы была интервальной кососимметричной (персимметричной, ганкелевой, тёплицевой, циркулянтной и т.п.), то и системы-потомки на каждом шаге алгоритма должны порождаться так, чтобы иметь кососимметричную (персимметричную, ганкелеву, тёплицеву, циркулянтную и т.п.) матрицу.

Приступая к реализации наших задумок, обозначим, аналогично тому, как это сделано в [4,6,22],

$Encl$ — какой-нибудь фиксированный метод внешнего оценивания множеств решений ИСЛАУ (мы будем называть его *базовым*);

$Encl(\mathbf{Q}, \mathbf{r})$ — получаемый с его помощью интервальный вектор внешней оценки для множества решений $\Xi_{sym}(\mathbf{Q}, \mathbf{r})$ системы $\mathbf{Q}x = \mathbf{r}$, т.е. $Encl(\mathbf{Q}, \mathbf{r}) \in \mathbb{I}\mathbb{R}^n$ и

$$Encl(\mathbf{Q}, \mathbf{r}) \supseteq \Xi_{sym}(\mathbf{Q}, \mathbf{r});$$

$\Upsilon(\mathbf{Q}, \mathbf{r})$ — нижний конец ν -ой компоненты (для заданного фиксированного номера $\nu \in \{1, 2, \dots, n\}$) внешней интервальной оценки множества решений $\Xi_{sym}(\mathbf{Q}, \mathbf{r})$, получаемой методом $Encl$, т.е.

$$\Upsilon(\mathbf{Q}, \mathbf{r}) := \underline{Encl(\mathbf{Q}, \mathbf{r})}_{\nu}. \quad (8)$$

Подчеркнем, что, коль скоро $\Xi_{sym}(\mathbf{A}, \mathbf{b}) \subseteq \Xi(\mathbf{A}, \mathbf{b})$, то в качестве базовых методов можно брать традиционные и хорошо разработанные методы внешнего оценивания множеств решений ИСЛАУ без связей, описанные, к примеру, в [1,2,5,15,17,18,20,21]. Получается, что алгоритм оценивания множеств решений ИСЛАУ со связями конструируется из более простых и грубых методов решения ИСЛАУ с независимыми данными.

Мы потребуем лишь от базового метода *Encl* удовлетворения следующим условиям:

$$\begin{aligned} &\text{оценка } \Upsilon(\mathbf{Q}, \mathbf{r}) \text{ монотонна по включению} \\ &\text{относительно матрицы } \mathbf{Q} \text{ и вектора } \mathbf{r}, \\ &\text{т. е. для всех } \mathbf{Q}', \mathbf{Q}'' \in \mathbb{IR}^{n \times n} \text{ и } \mathbf{r}', \mathbf{r}'' \in \mathbb{IR}^n \\ &\text{при } \mathbf{Q}' \subseteq \mathbf{Q}'' \text{ и } \mathbf{r}' \subseteq \mathbf{r}'' \text{ верно неравенство} \end{aligned} \quad (C1)$$

$$\Upsilon(\mathbf{Q}', \mathbf{r}') \geq \Upsilon(\mathbf{Q}'', \mathbf{r}'')$$

и

$$\begin{aligned} &\text{оценка } \Upsilon(\mathbf{Q}, \mathbf{r}) \text{ является точной на вещественных} \\ &\text{линейных алгебраических системах, т. е.} \\ &\Upsilon(Q, r) = (Q^{-1}r)_\nu \text{ для всех } Q \in \mathbb{R}^{n \times n} \text{ и } r \in \mathbb{R}^n. \end{aligned} \quad (C2)$$

Если базовый метод *Encl* является естественным интервальным расширением какого-нибудь вещественного метода (как, например, интервальный метод Гаусса), или, более общо, дерево Канторовича базового метода *Encl* своими узлами имеет только интервальные арифметические операции, то свойство (C1) очевидным образом выполняется в силу свойства монотонности интервальных арифметик по включению. Иначе, если в алгоритме базового метода встречаются неинтервальные операции, то свойство (C1) может и нарушаться. Проверку того, что данный конкретный базовый метод обладает свойством (C1), мы возлагаем на разработчиков программ.

Имеем

$$\min\{x_\nu \mid x \in \Xi_{sym}(\mathbf{A}, \mathbf{b})\} = (\tilde{A}^{-1}\tilde{\mathbf{b}})_\nu,$$

для некоторых точечной симметричной матрицы $\tilde{A} = (\tilde{a}_{ij}) \in \mathbb{R}^{n \times n}$ и точечного вектора $\tilde{\mathbf{b}} = (\tilde{b}_i) \in \mathbb{R}^n$, составленных из представителей элементов матрицы \mathbf{A} и вектора \mathbf{b} , причём по самому определению оценки Υ

$$\Upsilon(\tilde{A}, \tilde{\mathbf{b}}) \leq (\tilde{A}^{-1}\tilde{\mathbf{b}})_\nu.$$

Предположив, что в матрице \mathbf{A} симметричные относительно главной диагонали элементы \mathbf{a}_{ij} и \mathbf{a}_{ji} имеют ненулевую ширину, обозначим

$$\begin{aligned} \mathbf{A}' &\text{ — матрицу, полученную из } \mathbf{A} \text{ заменой элементов } \mathbf{a}_{ij} \text{ и } \mathbf{a}_{ji} \text{ на } [\underline{\mathbf{a}}_{ij}, \text{mid } \mathbf{a}_{ij}], \\ \mathbf{A}'' &\text{ — матрицу, полученную из } \mathbf{A} \text{ заменой элементов } \mathbf{a}_{ij} \text{ и } \mathbf{a}_{ji} \text{ на } [\text{mid } \mathbf{a}_{ij}, \bar{\mathbf{a}}_{ij}], \\ \tilde{\mathbf{A}}' &\text{ — матрицу, полученную из } \tilde{A} \text{ заменой элементов } \tilde{a}_{ij} \text{ и } \tilde{a}_{ji} \text{ на } [\underline{\mathbf{a}}_{ij}, \text{mid } \mathbf{a}_{ij}], \\ \tilde{\mathbf{A}}'' &\text{ — матрицу, полученную из } \tilde{A} \text{ заменой элементов } \tilde{a}_{ij} \text{ и } \tilde{a}_{ji} \text{ на } [\text{mid } \mathbf{a}_{ij}, \bar{\mathbf{a}}_{ij}]. \end{aligned}$$

Интервальные системы линейных алгебраических уравнений $\mathbf{A}'x = \mathbf{b}$ и $\mathbf{A}''x = \mathbf{b}$, полученные из исходной системы путём рассечения пары симметричных относительно главной диагонали интервальных элементов пополам, мы будем называть *системами-потомками* для $\mathbf{A}x = \mathbf{b}$.

Далее, так как

$$\tilde{\mathbf{A}}' \subseteq \mathbf{A}' \subseteq \mathbf{A}, \quad \tilde{\mathbf{A}}'' \subseteq \mathbf{A}'' \subseteq \mathbf{A},$$

и $\tilde{\mathbf{b}} \subseteq \mathbf{b}$, то условие (C1) имеет своим следствием неравенства

$$\Upsilon(\mathbf{A}, \mathbf{b}) \leq \Upsilon(\mathbf{A}', \mathbf{b}) \leq \Upsilon(\tilde{\mathbf{A}}', \tilde{\mathbf{b}})$$

и

$$\Upsilon(\mathbf{A}, \mathbf{b}) \leq \Upsilon(\mathbf{A}'', \mathbf{b}) \leq \Upsilon(\tilde{\mathbf{A}}'', \tilde{\mathbf{b}}).$$

Следовательно, беря почленный минимум от соответствующих частей неравенств, мы получим

$$\Upsilon(\mathbf{A}, \mathbf{b}) \leq \min\{\Upsilon(\mathbf{A}', \mathbf{b}), \Upsilon(\mathbf{A}'', \mathbf{b})\} \leq \min\{\Upsilon(\tilde{\mathbf{A}}', \tilde{\mathbf{b}}), \Upsilon(\tilde{\mathbf{A}}'', \tilde{\mathbf{b}})\}. \quad (9)$$

Кроме того, поскольку матрица \tilde{A} обязательно содержится либо в \tilde{A}' , либо в \tilde{A}'' , то справедливо, по крайней мере, одно из неравенств

$$\gamma(\tilde{A}', \tilde{b}) \leq \gamma(\tilde{A}, \tilde{b}) \quad \text{или} \quad \gamma(\tilde{A}'', \tilde{b}) \leq \gamma(\tilde{A}, \tilde{b}),$$

так что

$$\min\{\gamma(\tilde{A}', \tilde{b}), \gamma(\tilde{A}'', \tilde{b})\} \leq \gamma(\tilde{A}, \tilde{b}) \leq (\tilde{A}^{-1}\tilde{b})_{\nu} = \min\{x_{\nu} \mid x \in \Xi_{sym}(\mathbf{A}, \mathbf{b})\}. \quad (10)$$

Сопоставление неравенств (9) и (10) приводит к соотношению

$$\gamma(\mathbf{A}, \mathbf{b}) \leq \min\{\gamma(\mathbf{A}', \mathbf{b}), \gamma(\mathbf{A}'', \mathbf{b})\} \leq \min\{x_{\nu} \mid x \in \Xi_{sym}(\mathbf{A}, \mathbf{b})\},$$

и, как следствие, к следующему практическому рецепту:

решив две интервальных “системы-потомка” $\mathbf{A}'x = \mathbf{b}$ и $\mathbf{A}''x = \mathbf{b}$, мы можем прийти, вообще говоря, к более точной оценке снизу для искомого значения $\min\{x_{\nu} \mid x \in \Xi_{sym}(\mathbf{A}, \mathbf{b})\}$ в виде

$$\min\{\gamma(\mathbf{A}', \mathbf{b}), \gamma(\mathbf{A}'', \mathbf{b})\}.$$

Такой же эффект имеет и дробление в векторе правых частей \mathbf{b} какого-нибудь интервального элемента \mathbf{b}_i на подинтервалы $[\underline{\mathbf{b}}_i, \text{mid } \mathbf{b}_i]$ и $[\text{mid } \mathbf{b}_i, \overline{\mathbf{b}}_i]$, что может быть обосновано выкладками, совершенно аналогичными (9)–(10). Поэтому впредь для единообразия договоримся обозначать ИСЛАУ–“потомки”, получающиеся из $\mathbf{A}x = \mathbf{b}$ дроблением пополам одного интервального элемента в матрице \mathbf{A} либо в векторе \mathbf{b} , через $\mathbf{A}'x = \mathbf{b}'$ и $\mathbf{A}''x = \mathbf{b}''$.

Процедуру улучшения оценки для $\min\{x_{\nu} \mid x \in \Xi_{sym}(\mathbf{A}, \mathbf{b})\}$ посредством дробления элементов интервальной системы (3) можно повторить по отношению к системам-потомкам $\mathbf{A}'x = \mathbf{b}'$ и $\mathbf{A}''x = \mathbf{b}''$, затем снова разбить потомков от $\mathbf{A}'x = \mathbf{b}'$ и $\mathbf{A}''x = \mathbf{b}''$ и снова улучшить оценку и т. д. Мы оформим подобный процесс последовательного улучшения оценки снизу для $\min\{x_{\nu} \mid x \in \Xi_{sym}(\mathbf{A}, \mathbf{b})\}$ так, как это делается в широко известном в комбинаторной оптимизации “методе ветвей и границ” [3] и как это было адаптировано для интервальных методов глобальной оптимизации (см., к примеру, книги [17,18]):

- во-первых, организуем все интервальные системы $\mathbf{Q}x = \mathbf{r}$, которые возникают в процессе дробления исходной ИСЛАУ $\mathbf{A}x = \mathbf{b}$, вместе с их оценками $\gamma(\mathbf{Q}, \mathbf{r})$ в некоторый список \mathcal{L} ;
- во-вторых, дробить будем лишь ту интервальную систему $\mathbf{Q}x = \mathbf{r}$ из списка \mathcal{L} , которая обеспечивает рекордную (наилучшую) на данный момент оценку $\gamma(\mathbf{Q}, \mathbf{r})$ для искомой величины $\min\{x_{\nu} \mid x \in \Xi_{sym}(\mathbf{A}, \mathbf{b})\}$;
- в-третьих, в подвергаемой дроблению ИСЛАУ мы будем дробить лишь самый широкий из интервальных элементов.

Итак, в процессе выполнения алгоритма мы будем поддерживать список \mathcal{L} , состоящий из записей-троек вида

$$(\mathbf{Q}, \mathbf{r}, \gamma(\mathbf{Q}, \mathbf{r})), \quad (11)$$

где

- \mathbf{Q} — интервальная $n \times n$ -матрица, $\mathbf{Q} \subseteq \mathbf{A}$,
- \mathbf{r} — интервальный n -вектор, $\mathbf{r} \subseteq \mathbf{b}$.

Кроме того, образующие \mathcal{L} записи будут упорядочены по возрастанию значений оценки $\gamma(\mathbf{Q}, \mathbf{r})$, а первую запись списка, так же как и соответствующие ИСЛАУ $\mathbf{Q}x = \mathbf{r}$ и оценку $\gamma(\mathbf{Q}, \mathbf{r})$ (рекордную в списке), мы будем называть *ведущими* на данном шаге. Полный псевдокод получающегося нового алгоритма, который мы аналогично [6] также назовём *методом дробления параметров*, представлен в Таблице (где через “ \leftarrow ” обозначен оператор присваивания). Он отличается от метода дробления параметров, представленного в [4,22,6] способом порождения интервальных систем-потомков из ведущей интервальной системы и условием остановки.

То, насколько близкими окажутся результат работы алгоритма и искомый $\min\{x_{\nu} \mid x \in \Xi_{sym}(\mathbf{A}, \mathbf{b})\}$, зависит, с одной стороны, от способа получения оценки $\gamma(\mathbf{Q}, \mathbf{r})$, т. е. от выбранного нами базового метода

Простейший метод дробления параметров
для интервальных симметричных линейных систем

Вход

Интервальная симметричная линейная система $\mathbf{Ax} = \mathbf{b}$.
Номер оцениваемой компоненты $\nu \in \{1, 2, \dots, n\}$.
Метод *Encl*, формирующий оценку \mathcal{T} по правилу (8).
Константа $\delta > 0$.

Выход

Оценка Z снизу для $\min\{x_\nu \mid x \in \Xi_{sym}(\mathbf{A}, \mathbf{b})\}$.

Алгоритм

присваиваем $\mathbf{Q} \leftarrow \mathbf{A}$ и $\mathbf{r} \leftarrow \mathbf{b}$;
вычисляем оценку $v \leftarrow \mathcal{T}(\mathbf{Q}, \mathbf{r})$;
инициализируем рабочий список $\mathcal{L} \leftarrow \{(\mathbf{Q}, \mathbf{r}, v)\}$;
DO WHILE (максимум ширины элементов в \mathbf{Q} и \mathbf{r} превосходит δ)
 в матрице $\mathbf{Q} = (\mathbf{q}_{ij})$ и векторе $\mathbf{r} = (\mathbf{r}_i)$ выбираем интервальный элемент \mathbf{s} , имеющий наибольшую ширину;
 порождаем интервальные системы-потомки $\mathbf{Q}'x = \mathbf{r}'$ и $\mathbf{Q}''x = \mathbf{r}''$:
 если $\mathbf{s} = \mathbf{q}_{kl}$ для некоторых $k, l \in \{1, 2, \dots, n\}$, то полагаем
 $\mathbf{q}'_{ij} \leftarrow \mathbf{q}''_{ij} \leftarrow \mathbf{q}_{ij}$ для $(i, j) \neq (k, l)$ или $(i, j) \neq (l, k)$,
 $\mathbf{q}'_{lk} \leftarrow \mathbf{q}'_{kl} \leftarrow [\underline{\mathbf{q}}_{kl}, \text{mid } \mathbf{q}_{kl}]$, $\mathbf{q}''_{lk} \leftarrow \mathbf{q}''_{kl} \leftarrow [\text{mid } \mathbf{q}_{kl}, \overline{\mathbf{q}}_{kl}]$,
 $\mathbf{r}' \leftarrow \mathbf{r}'' \leftarrow \mathbf{r}$;
 если $\mathbf{s} = \mathbf{r}_k$ для некоторого $k \in \{1, 2, \dots, n\}$, то полагаем
 $\mathbf{Q}' \leftarrow \mathbf{Q}'' \leftarrow \mathbf{Q}$, $\mathbf{r}'_k \leftarrow [\underline{\mathbf{r}}_k, \text{mid } \mathbf{r}_k]$, $\mathbf{r}''_k \leftarrow [\text{mid } \mathbf{r}_k, \overline{\mathbf{r}}_k]$,
 $\mathbf{r}'_i \leftarrow \mathbf{r}''_i \leftarrow \mathbf{r}_i$ для $i \neq k$;
 вычисляем оценки $v' \leftarrow \mathcal{T}(\mathbf{Q}', \mathbf{r}')$ и $v'' \leftarrow \mathcal{T}(\mathbf{Q}'', \mathbf{r}'')$;
 удаляем из \mathcal{L} бывшую ведущей запись $(\mathbf{Q}, \mathbf{r}, v)$;
 вносим записи $(\mathbf{Q}', \mathbf{r}', v')$ и $(\mathbf{Q}'', \mathbf{r}'', v'')$ в список \mathcal{L} , сохраняя его упорядоченность по возрастанию третьего поля;
 обозначаем первую запись списка через $(\mathbf{Q}, \mathbf{r}, v)$;
END DO
 $Z \leftarrow v$;

для решения ИСЛАУ-потомков, а с другой — от обусловленности точечных систем, образующих последнюю ведущую систему (её можно оценивать в процессе выполнения алгоритма). В частности, для того, чтобы при $\delta \rightarrow 0$ вычисленное алгоритмом значение стремилось бы к $\min\{x_\nu \mid x \in \Xi_{sym}(\mathbf{A}, \mathbf{b})\}$, необходимо и достаточно выполнения условия (С2). Если же в исходной ИСЛАУ суммарная ширина интервальных элементов оказывается “большой” в сравнении с δ , то, как правило, простейший метод дробления параметров не будет прорабатывать до конца, и потому целесообразней рассматривать его как итеративную уточняющую процедуру.

3 Тест на монотонность

Пусть дана интервальная система линейных алгебраических уравнений $\mathbf{Q}x = \mathbf{r}$, и нам известны

$$\frac{\partial x_\nu(\mathbf{Q}, \mathbf{r})}{\partial q_{ij}} \quad \text{и} \quad \frac{\partial x_\nu(\mathbf{Q}, \mathbf{r})}{\partial r_i}$$

— интервальные расширения соответствующих производных

$$\frac{\partial x_\nu(Q, r)}{\partial q_{ij}} \quad \text{и} \quad \frac{\partial x_\nu(Q, r)}{\partial r_i}$$

от ν -ой компоненты вектора решения системы уравнений $Qx = r$ по ij -ому элементу матрицы $Q = (q_{ij})$ и i -ому элементу вектора $r = (r_i)$, взятых с учётом наложенных на систему связей. Если интервальные $n \times n$ -матрица $\tilde{\mathbf{Q}} = (\tilde{q}_{ij})$ и n -вектор $\tilde{\mathbf{r}} = (\tilde{r}_i)$ образованы из элементов

$$\tilde{q}_{ij} = \begin{cases} [\underline{q}_{ij}, \overline{q}_{ij}], & \text{при} \quad \frac{\partial x_\nu(\mathbf{Q}, \mathbf{r})}{\partial q_{ij}} \geq 0, \\ [\overline{q}_{ij}, \underline{q}_{ij}], & \text{при} \quad \frac{\partial x_\nu(\mathbf{Q}, \mathbf{r})}{\partial q_{ij}} \leq 0, \\ q_{ij}, & \text{при} \quad \text{int} \frac{\partial x_\nu(\mathbf{Q}, \mathbf{r})}{\partial q_{ij}} \ni 0, \end{cases} \quad (12)$$

$$\tilde{r}_i = \begin{cases} [\underline{r}_i, \overline{r}_i], & \text{при} \quad \frac{\partial x_\nu(\mathbf{Q}, \mathbf{r})}{\partial r_i} \geq 0, \\ [\overline{r}_i, \underline{r}_i], & \text{при} \quad \frac{\partial x_\nu(\mathbf{Q}, \mathbf{r})}{\partial r_i} \leq 0, \\ r_i, & \text{при} \quad \text{int} \frac{\partial x_\nu(\mathbf{Q}, \mathbf{r})}{\partial r_i} \ni 0, \end{cases} \quad (13)$$

где “int” обозначает внутренность интервала, то, очевидно,

$$\min\{x_\nu \mid x \in \Xi_{sym}(\tilde{\mathbf{Q}}, \tilde{\mathbf{r}})\} = \min\{x_\nu \mid x \in \Xi_{sym}(\mathbf{Q}, \mathbf{r})\}.$$

А поскольку количество существенно интервальных (с ненулевой шириной) элементов в $\tilde{\mathbf{Q}}$ и $\tilde{\mathbf{r}}$ может быть, вообще говоря, меньшим, чем в \mathbf{Q} и \mathbf{r} , то, переходя от исходной ИСЛАУ $\mathbf{Q}x = \mathbf{r}$ к системе $\tilde{\mathbf{Q}}x = \tilde{\mathbf{r}}$, мы упрощаем задачу вычисления $\min\{x_\nu \mid x \in \Xi_{sym}(\mathbf{Q}, \mathbf{r})\}$.

Ранее в интервальном анализе в ряде численных методик уже использовались производные решений систем линейных уравнений по элементам матрицы и правой части (см., к примеру, Главу 17 книги [1]). Выведем формулы для этих производных с учётом симметричности матрицы системы уравнений.

Пусть k и l — некоторые фиксированные индексы, такие что $1 \leq k \leq l \leq n$. Распишем систему уравнений $Qx = r$ в виде

$$\sum_{j=1}^n q_{ij}x_j = r_i, \quad i = 1, 2, \dots, n, \quad (14)$$

и продифференцируем по q_{kl} . Учитывая, что

$$\frac{\partial}{\partial q_{kl}}(q_{ij}x_j) = \frac{\partial q_{ij}}{\partial q_{kl}}x_j + q_{ij}\frac{\partial x_j}{\partial q_{kl}},$$

где

$$\frac{\partial q_{ij}}{\partial q_{kl}} = \begin{cases} 0, & \text{если } (i, j) \neq (k, l), \\ 1, & \text{если } (i, j) = (k, l), \end{cases}$$

мы получим из (14)

$$\left\{ \begin{array}{l} \sum_{j=1}^n q_{ij} \frac{\partial x_j}{\partial q_{kl}} = 0, \text{ если } i \neq k \text{ и } i \neq l, \\ \sum_{j=1}^n q_{ij} \frac{\partial x_j}{\partial q_{kl}} + x_l = 0, \text{ если } i = k, \\ \sum_{j=1}^n q_{ij} \frac{\partial x_j}{\partial q_{kl}} + x_k = 0, \text{ если } i = l. \end{array} \right.$$

Таким образом, если

$$\frac{\partial x}{\partial q_{kl}} = \left(\frac{\partial x_1}{\partial q_{kl}}, \dots, \frac{\partial x_n}{\partial q_{kl}} \right)^\top,$$

то

$$Q \cdot \frac{\partial x}{\partial q_{kl}} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ -x_l \\ 0 \\ \vdots \\ 0 \\ -x_k \\ 0 \\ \vdots \\ 0 \end{pmatrix} \begin{array}{l} \leftarrow k\text{-е место} \\ \\ \leftarrow l\text{-е место.} \end{array}$$

По этой причине

$$\frac{\partial x}{\partial q_{kl}} = Q^{-1} \cdot (0, \dots, 0, -x_l, 0, \dots, 0, -x_k, 0, \dots, 0)^\top,$$

и если $Y = (y_{ij})$ — обратная матрица для Q , то производные решения вещественной симметричной линейной системы $Qx = r$ по элементам матрицы даются формулами

$$\frac{\partial x_\nu}{\partial q_{kl}} = -y_{\nu k}x_l - y_{\nu l}x_k.$$

Дифференцирование уравнений (14) по r_k приводит к более простым соотношениям

$$\left\{ \begin{array}{l} \sum_{j=1}^n q_{ij} \frac{\partial x_j}{\partial r_k} = 0, \text{ если } i \neq k, \\ \sum_{j=1}^n q_{ij} \frac{\partial x_j}{\partial r_k} = 1, \text{ если } i = k. \end{array} \right.$$

Таким образом, если

$$\frac{\partial x}{\partial r_k} = \left(\frac{\partial x_1}{\partial r_k}, \dots, \frac{\partial x_n}{\partial r_k} \right)^\top,$$

то

$$Q \cdot \frac{\partial x}{\partial r_k} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow k\text{-е место.}$$

По этой причине

$$\frac{\partial x}{\partial r_k} = Q^{-1} \cdot (0, \dots, 0, 1, 0, \dots, 0)^\top,$$

и если $Y = (y_{ij})$ — обратная матрица для Q , то производные решения вещественной симметричной линейной системы $Qx = r$ по элементам вектора правой части даются формулами

$$\frac{\partial x_\nu}{\partial r_k} = y_{\nu k}.$$

Следовательно, если $\mathbf{Y} = (y_{ij})$ — “обратная интервальная матрица” для \mathbf{Q} , т. е. внешняя интервальная оценка для множества обратных матриц из Q ,

$$\mathbf{Y} \supseteq \{Q^{-1} \mid Q \in \mathbf{Q}\},$$

а \mathbf{x}_k и \mathbf{x}_l — k -ая и l -ая компоненты некоторого интервального вектора $\mathbf{x} \supseteq \Xi_{sym}(\mathbf{Q}, \mathbf{r})$, то мы можем принять следующие интервальные оценки производных

$$\frac{\partial x_\nu(\mathbf{Q}, \mathbf{r})}{\partial q_{kl}} = -y_{\nu k} \mathbf{x}_l - y_{\nu l} \mathbf{x}_k, \quad \frac{\partial x_\nu(\mathbf{Q}, \mathbf{r})}{\partial r_k} = y_{\nu k}.$$

Для интервальных линейных систем с кососимметричными матрицами аналогичные интервальные оценки производных, как нетрудно убедиться, имеют вид

$$\frac{\partial x_\nu(\mathbf{Q}, \mathbf{r})}{\partial q_{kl}} = -y_{\nu k} \mathbf{x}_l + y_{\nu l} \mathbf{x}_k, \quad \frac{\partial x_\nu(\mathbf{Q}, \mathbf{r})}{\partial r_k} = y_{\nu k}.$$

Вычисление “обратной интервальной матрицы” для \mathbf{Q} можно выполнить как решение задачи внешнего оценивания объединённого множества решений интервального матричного уравнения

$$\mathbf{Q}\mathbf{Y} = I, \quad I — \text{единичная матрица,}$$

применив, к примеру, n раз (для каждого столбца) тот же самый метод внешнего оценивания *Encl*, который выбран базовым для всего алгоритма. Соответствующие инструкции — пересчёт обратной интервальной матрицы и проверку на монотонность с возможным последующей заменой некоторых интервальных элементов на их концы по формулам (12)–(13) — следует выполнять сразу после порождения интервальных систем потомков в основном цикле алгоритма.

4 Стратегия дробления

В простейшем методе дробления параметров мы рассекали на каждом шаге самый широкий из интервальных элементов ведущей ИСЛАУ. Можно ли путём специального выбора элемента для дробления обеспечить наиболее значительное улучшение целевой функции на каждом шаге алгоритма? Строгая и точная оптимизация методов дробления параметров в таком виде трудна и вряд ли целесообразна в полном объёме. Мы будем решать этот вопрос, руководствуясь следующими эвристическими соображениями: величина произведения ширины (или радиуса) интервального элемента на модуль интервального расширения соответствующей производной решения может служить, в некотором смысле, мерой того, как бисекция элементов из \mathbf{Q} либо \mathbf{r} влияет на $\min\{x_\nu \mid x \in \Xi_{sym}(\mathbf{Q}, \mathbf{r})\}$ и размеры объединённого множества решений ИСЛАУ.

Далее, оценки объединённого множества решений ИСЛАУ, получаемые по большинству существующих методов, являются тем более точными, чем меньше размеры этого множества решений. С подобными

базовыми методами уменьшение размеров множества решений $\Xi_{sym}(\mathbf{Q}, \mathbf{r})$ должно приводить к аналогичному и сравнимому по величине изменению в оценке $\mathcal{Y}(\mathbf{Q}, \mathbf{r})$. При этом, следовательно, требование наиболее быстрого улучшения целевой функции за один шаг метода дробления параметров становится, по существу, эквивалентным условию наиболее быстрого уменьшения размеров множества решений при дроблении ведущей ИСЛАУ.

Учитывая сделанные выше заключения, мы рекомендуем дробить ведущие ИСЛАУ по элементам, на которых достигается максимальная из величин

$$\left| \frac{\partial x_\nu(\mathbf{Q}, \mathbf{r})}{\partial q_{ij}} \right| \cdot \text{wid } \mathbf{q}_{ij}, \quad \left| \frac{\partial x_\nu(\mathbf{Q}, \mathbf{r})}{\partial r_i} \right| \cdot \text{wid } \mathbf{r}_i, \quad \text{wid} — \text{ширина интервалов}, \quad (15)$$

$i, j \in \{1, 2, \dots, n\}$, т. е. по элементам, на которых достигается максимум произведения оценки производной решения на ширину интервала, а не просто по самым широким элементам ИСЛАУ. Стратегия дробления, требующая максимизации величин (15), была впервые рассмотрена автором в статье [22].

5 Численный пример

В качестве примера применения разработанного нами подхода рассмотрим интервальную симметричную линейную систему

$$\begin{pmatrix} [1, 10] & [0, 1] \\ [0, 1] & [-4, -1] \end{pmatrix} x = \begin{pmatrix} [0, 2] \\ [0, 2] \end{pmatrix}, \quad (16)$$

которая отличается от системы (7) из работы [12] единственным элементом \mathbf{a}_{11} в матрице. Кажется, что Алефельд, Крейнович и Майер намеренно взяли его точечным, так как иначе выполненные ими вручную выкладки (и без того весьма длинные), приводящие к точному описанию формы множества решений системы (7) ещё бы значительно усложнились. Мы не стеснены подобными ограничениями и легко можем вычислить оценки множеств решений этой системы с существенно интервальным элементом \mathbf{a}_{11} .

Метод дробления параметров, в котором в качестве базового метода использовался интервальный метод Гаусса-Зейделя (описание которого можно найти, к примеру, в [18,20]), снабженный проверкой монотонности (из §3) и модифицированной стратегией выбора дробимого элемента (из §4), надёжно находит внешнюю оценку множества решений этой ИСЛАУ в виде

$$\begin{pmatrix} [-7.61868 e-8, 2.41421] \\ [-2.00000, 1.00000] \end{pmatrix}$$

(мы сохранили по шесть значащих цифр результата). При этом брус начальной оценки, вычисленный с помощью процедуры Хансена-Рона [21], оказался равным

$$\begin{pmatrix} [-14.9333, 23.7333] \\ [-44.6667, 38.0000] \end{pmatrix},$$

т.е. весьма грубым. Для нахождения нижних оценок множества решений системы (16) мы затратили по каждой координате по 1000 итераций, и уточнение этих оценок шло исключительно за счет бисекции элементов ИСЛАУ, хотя (любопытный факт!) диагностика свидетельствует, что тест на монотонность (12)–(13) почти на каждом шаге сужал те или иные интервальные элементы в системах-потомках, появляющихся по ходу работы алгоритма. Но, как ни странно, это не привело к повышению эффективности всего алгоритма в целом. Что касается верхних оценок, то здесь тесты монотонности сыграли свою положительную роль, и количества шагов метода дроблений параметров оказались равными всего 40 по первой компоненте и 33 по второй.

Как видим, оценки множества решений системы (16) не изменились в сравнении с системой (7), т.е. интервализация элемента \mathbf{a}_{11} никак не повлияла на них.

Отметим в заключение, что при проведении вычислительных экспериментов с подобного сорта задачами одной из основных трудностей является в настоящее время отсутствие тестовых примеров.

Список литературы

1. АЛЕФЕЛЬД Г., ХЕРЦБЕРГЕР Ю. *Введение в интервальные вычисления*. – М.: Мир, 1987.
2. КАЛМЫКОВ С. А., ШОКИН Ю. И., ЮЛДАШЕВ З. Х. *Методы интервального анализа*. – Новосибирск: Наука, 1986.
3. ПАПАДИМИТРИУ Х., СТАЙГЛИЦ К. *Комбинаторная оптимизация. Алгоритмы и сложность*. – М.: Мир, 1985.
4. ШАРЫЙ С. П. Новый класс алгоритмов для оптимального решения интервальных линейных систем // Конференция “Актуальные проблемы прикладной математики”, Саратов, 20–22 мая 1991 г. – Саратов, 1991. – С. 113–119.
5. ШАРЫЙ С. П. Алгебраический подход во “внешней задаче” для интервальных линейных систем // *Вычислительные Технологии*. – 1998. – Т. 3, №2. – С. 67–114.
6. ШАРЫЙ С.П. Оптимальное внешнее оценивание множеств решений интервальных систем уравнений. Часть 1 // *Вычислительные Технологии*. – 2002. – Т. 7, №6. – С. 90–113.
7. ШАРЫЙ С.П. Оптимальное внешнее оценивание множеств решений интервальных систем уравнений. Часть 2 // *Вычислительные Технологии*. – 2003. – Т. 8, №1. – С. 84–109.
7. ALEFELD G., KREINOVICH V., MAYER G. Symmetric linear systems with perturbed input data // *Numerical Methods and Error Bounds*, Alefeld G. and Herzberger J., eds. – Berlin: Akademie Verlag, 1996. – P. 16–22.
8. ALEFELD G., KREINOVICH V., MAYER G. The shape of the symmetric solution set // *Applications of Interval Computations*, Kearfott R.B. and Kreinovich V., eds. – Dordrecht: Kluwer, 1996. – P. 61–79.
9. ALEFELD G., KREINOVICH V., MAYER G. On the shape of the symmetric, persymmetric, and skew-symmetric solution set // *SIAM J. Matrix Anal. Appl.* – 1997. – Vol. 18. – P. 693–705.
10. ALEFELD G., KREINOVICH V., MAYER G. The shape of the solution set for systems of interval linear equations with dependent coefficients // *Mathematische Nachrichten*. – 1998. – Bd. 192. – P. 23–36.
11. ALEFELD G., KREINOVICH V., MAYER G. A comment on the shape of the solution set for systems of interval linear equations with dependent coefficients // *Reliable Computing*. – 2001. – Vol. 7, No. 3. – P. 275–277.
12. ALEFELD G., KREINOVICH V., MAYER G. On symmetric solution sets // *Computing Supplement 16*, Herzberger J., ed. – Wien, New York: Springer, 2003. – P. 1–23.
13. ALEFELD G., KREINOVICH V., MAYER G. On the solution sets of particular classes of linear interval systems // *Journal of Computational and Applied Mathematics*. – 2003. – Vol. 152, No. 1–2. – P. 1–15. (электронная версия доступна в Интернете на <http://www.cs.utep.edu/vladik/1996/abstr96.html>).
14. ALEFELD G., MAYER G. The Cholesky method for interval data // *Linear Algebra and its Applications*. – 1993. – Vol. 194. – P. 161–182.
15. GAY D. M. Solving interval linear equations // *SIAM J. Numer. Analysis*. – 1982. – Vol. 19, №4. – P. 858–870.
16. JANSSON CH. Interval linear systems with symmetric matrices, skew-symmetric matrices, and dependencies in the right hand side // *Computing*. – 1991. – Vol. 46. – P. 265–274.
17. HANSEN E. *Global optimization using interval analysis*. – New York: Marcel Dekker, 1992.
18. KEARFOTT R. B. *Rigorous global search: Continuous problems*. – Dordrecht: Kluwer, 1996.
19. KEARFOTT R.B., NAKAO M.T., NEUMAIER A., RUMP S.M., SHARY S.P., VAN HENTENRYCK P. Standardized notation in interval analysis // *Reliable Computing*. – в печати (электронная версия доступна в Интернете на <http://www.mat.univie.ac.at/~neum/software/int>).
20. NEUMAIER A. *Interval methods for systems of equations*. – Cambridge: Cambridge University Press, 1990.
21. NEUMAIER A. A simple derivation of Hansen-Blik-Rohn-Ning-Kearfott enclosure for linear interval equations // *Reliable Computing*. – 1999. – Vol. 5, №2. – P. 131–136.
22. SHARY S. P. A new class of algorithms for optimal solution of interval linear systems // *Interval Computations*. – 1992. – №2(4). – P. 18–29.