

Внутреннее оценивание множеств решений неотрицательных интервальных линейных систем

С. П. Шарый

УДК 519.6

Шарый С.П. Внутреннее оценивание множеств решений неотрицательных интервальных линейных систем // Сиб. журн. вычисл. математики / РАН. Сиб. отд.-ние. — Новосибирск, 2006. — Т. 9, № 2. — С. 189–206.

В работе предлагается способ нахождения максимальных по включению внутренних оценок для множеств решений интервальных линейных систем с неотрицательными матрицами, основанный на свойстве *монотонности конфигурации* этих множеств решений.

Ключевые слова: *интервальная линейная система, неотрицательная матрица, множество решений, внутреннее оценивание, максимальная оценка.*

Shary S.P. Inner estimation of solution sets to non-negative interval linear systems // Siberian J. of Numer. Mathematics / Sib. Branch of Russ. Acad. of Sci. — Novosibirsk, 2006. — Vol. 9, № 2. — P. 189–206.

This paper presents a new technique for constructing the maximum (with respect to an inclusion) inner estimates of the solution sets to the interval linear equations systems having non-negative matrices, based on the *shape monotonicity* of these solution sets.

Key words: *interval linear system, non-negative matrix, solution set, inner estimation, maximum estimate.*

1. Введение

Предмет рассмотрения нашей работы — интервальные системы линейных алгебраических уравнений (ИСЛАУ) вида

$$Ax = b \quad (1.1)$$

с интервальной $m \times n$ -матрицей $A = (a_{ij})$ и интервальным m -вектором $b = (b_i)$. Мы будем рассматривать их как совокупность обычных (точечных) систем уравнений вида

$$Ax = b \quad (1.2)$$

той же структуры, что и (1.1), у которых элементы матрицы a_{ij} и правой части b_i могут независимо друг от друга изменяться в пределах интервалов a_{ij} и b_i соответственно.

Отметим, что всюду в нашей работе интервалы и интервальные объекты обозначаются в соответствии с проектом неформального международного стандарта [8] буквами жирного шрифта (A, B, C, \dots, x, y, z), тогда как обычные неинтервальные (точечные) величины никак специально не выделяются. Арифметические операции с интервальными величинами — это операции соответствующих интервальных арифметик, либо классической интервальной арифметики \mathbb{IR} (см., напр., [1, 3, 12]), либо расширенной интервальной арифметики Кахана [7, 11]. Наконец, подчеркивание и надчеркивание — \underline{a} , \bar{a} — обозначают нижний и верхний концы интервала a .

По мере того как параметры системы (1.1) — a_{ij} и b_i — варьируются в пределах соответствующих интервалов, решение вещественной системы (1.2) также дрейфует в пространстве \mathbb{R}^n , “заметая” некоторое множество, которое мы будем называть *множеством решений* интервальной системы (1.1) и обозначать $\Xi(\mathbf{A}, \mathbf{b})$. Таким образом, множество решений интервальной системы линейных уравнений (1.1) образовано всевозможными решениями систем $Ax = b$ с $A \in \mathbf{A}$ и $b \in \mathbf{b}$, и строгое его определение имеет вид

$$\Xi(\mathbf{A}, \mathbf{b}) := \{ x \in \mathbb{R}^n \mid (\exists A \in \mathbf{A})(\exists b \in \mathbf{b})(Ax = b) \}. \quad (1.3)$$

Часто его называют также *объединенным множеством решений*, поскольку существуют и другие множества решений интервальных систем уравнений (см., напр., [5, 15]). Мы не будем рассматривать их в нашей работе, и потому для краткости опускаем эпитет “объединенное”.

В качестве примера практического возникновения множества решений (1.3), важного для нашей работы, рассмотрим так называемую задачу идентификации в условиях интервальной неопределенности.



Рис. 1. Структурная схема объекта идентификации

Пусть имеется статический объект (рис. 1), входы и выходы которого описываются конечномерными векторами $(a^{(1)}, a^{(2)}, \dots, a^{(n)}) \in \mathbb{R}^n$ и $(b^{(1)}, b^{(2)}, \dots, b^{(r)}) \in \mathbb{R}^r$ соответственно. Будем предполагать также, что зависимость “вход-выход” является линейной:

$$b^{(k)} = \sum_{l=1}^n x_{kl} a^{(l)}, \quad k = 1, 2, \dots, r,$$

с некоторыми постоянными коэффициентами x_{kl} . При функционировании объекта (или его экспериментальном исследовании) мы можем измерять его входные сигналы и соответствующие выходные отклики, и одна из практически наиболее важных задач, возникающих в подобных ситуациях, — это *задача идентификации*: требуется найти (или как-то оценить) значения x_{kl} , т. е. идентифицировать параметры объекта. Иными словами, мы должны сформировать математическую модель объекта на основе результатов измерений (наблюдений), например, с целью предсказания его будущего поведения и т. п.

Если серия наблюдений входа-выхода объекта является “достаточно представительной”, то мы получаем в свое распоряжение набор отношений с числовыми параметрами, связывающий x_{kl} , т. е., фактически, некоторую систему уравнений относительно x_{kl} , из которой можно найти их значения.

Рассмотрим далее для простоты модель объекта, в которой на x_{kl} не накладывается никаких дополнительных ограничений, в частности, их можно считать независимыми

Если $\mathbf{A} = (\mathbf{a}_{ij})$ — интервальная $m \times n$ -матрица, составленная из m результатов измерений входов, $\mathbf{b} = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m)^\top$ — интервальный вектор m измерений выходов, то семейство всех векторов параметров, согласующихся с интервально заданными экспериментальными данными, может быть представлено в виде

$$\{x \in \mathbb{R}^n \mid (\exists \mathbf{A} \in \mathbf{A})(\exists \mathbf{b} \in \mathbf{b})(\mathbf{A}x = \mathbf{b})\},$$

то есть как множество решений всевозможных точечных систем $\mathbf{A}x = \mathbf{b}$ с $\mathbf{A} \in \mathbf{A}$ и $\mathbf{b} \in \mathbf{b}$. Специалистами по теории идентификации это множество часто называется *информационным множеством*, (*апостериорным*) *множеством возможных значений параметров* [10] и т. п. Но для нас наиболее существенно то, что это множество является ни чем иным, как определенным в (1.3) множеством решений $\Xi(\mathbf{A}, \mathbf{b})$ интервальной линейной системы уравнений $\mathbf{A}x = \mathbf{b}$.

2. Постановка задачи

Известно, что пересечение множества решений $\Xi(\mathbf{A}, \mathbf{b})$ с каждым из ортантов пространства \mathbb{R}^n является выпуклым и полиэдральным [13], так что в целом множество решений может быть представлено как объединение не более чем 2^n полиэдров. Следовательно, длина точного прямого описания множества решений ИСЛАУ, при котором мы скрупулезно выписываем в каждом ортанте все уравнения ограничивающих $\Xi(\mathbf{A}, \mathbf{b})$ гиперплоскостей, может расти экспоненциально с размерностью интервальной системы n , а потому оно представляется практически невозможным уже при n , превосходящем несколько десятков. В действительности, подобное точное описание и не нужно в большинстве реальных постановок задач. Пользователи, как правило, ограничиваются задачами нахождения оценок, в том или ином смысле, для множеств решений, т. е. заменяют задачу точного описания $\Xi(\mathbf{A}, \mathbf{b})$ на задачу его приближенного описания в соответствии со смыслом решаемой практической постановки [1–3, 7, 12, 15].

В зависимости от стоящей перед нами задачи возможны и различные способы оценивания множеств решений. К примеру, если в рассмотренной во введении задаче идентификации линейного статического объекта в условиях интервальной неопределенности мы собираемся использовать результаты для выбора параметров объекта на стадии его проектирования, то естественно оценивать $\Xi(\mathbf{A}, \mathbf{b})$ с помощью подмножеств, т. е. таким образом, чтобы наверняка не захватить “лишние” векторы параметров, не имеющие отношения к идентифицируемому объекту. С другой стороны, если мы планируем использовать результаты идентификации для гарантированной оценки по полученной модели наибольших возможных отклонений выходов в процессе функционирования объекта, то множество решений $\Xi(\mathbf{A}, \mathbf{b})$ следует оценивать “извне”, с помощью объемлющих множеств, которые содержат все векторы результатов идентификации. Подробное обсуждение этих вопросов заинтересованный читатель может найти, в частности, в [2, 10].

В этой работе мы будем рассматривать задачу внутреннего (посредством подмножеств) оценивания множеств решений ИСЛАУ (см. рис. 2), причем в качестве оценивающих подмножеств возьмем интервальные векторы, т. е. брусы в \mathbb{R}^n со сторонами, параллельными координатным осям:

<p>Для интервальной линейной системы $\mathbf{A}x = \mathbf{b}$ найти интервальный вектор, содержащийся во множестве решений $\Xi(\mathbf{A}, \mathbf{b})$.</p>	(2.6)
---	-------

Отметим, что интервальное представление ответа — внутренней оценки — является, как правило, весьма удобным для “лиц, принимающих решения” (конструктора, оператора технологического процесса и т. п.). Вдобавок, формулировка задачи в чисто интервальных терминах позволяет привлечь для ее решения развитый аппарат интервального анализа. Ясно, что внутренняя интервальная оценка неединственна, и на практике наибольшую ценность имеют *максимальные по включению* (неулучшаемые) внутренние оценки.

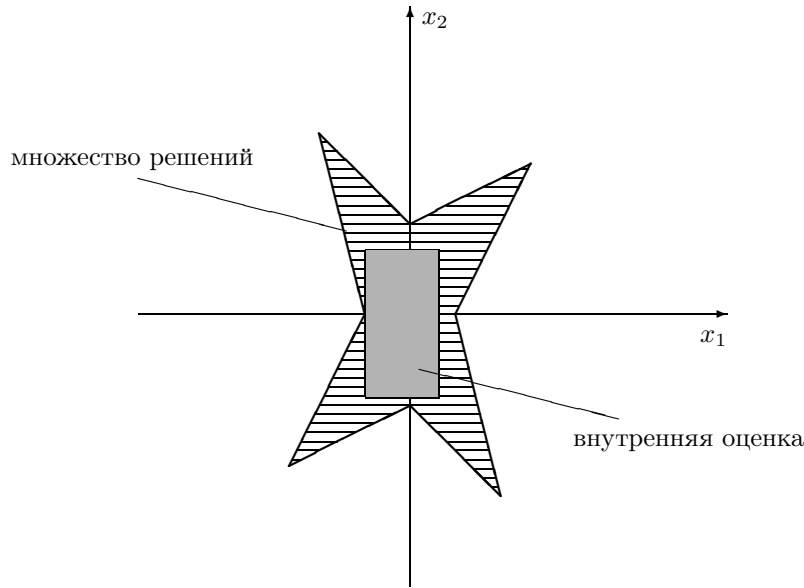


Рис. 2. Внутреннее оценивание множества решений

Задачи внутреннего интервального оценивания множеств решений ИСЛАУ рассматривались ранее, например, в [5, 15], где автором развивался подход, основанный на сведении исходной задачи оценивания к задаче нахождения формального (алгебраического) решения некоторой вспомогательной интервальной системы. Ниже мы развиваем еще одну методику внутреннего интервального оценивания множеств решений ИСЛАУ с неотрицательными не обязательно квадратными матрицами, которая в ряде случаев может оказаться разумной альтернативой формально-алгебраическому подходу из [5, 15]. Она основана на свойстве *монотонности конфигурации* множеств решений ИСЛАУ с неотрицательными матрицами (впервые отмеченном в [14]) и позволяет находить максимальные по включению внутренние оценки множеств решений.

Кроме неотрицательности, мы не накладываем никаких ограничений (квадратность, неособенность и т. п.) на интервальную матрицу \mathbf{A} системы уравнений (1.1), но при этом постановка задачи должна быть скорректирована на случай неограниченного множества решений. В этой ситуации мы будем искать внутреннюю оценку для пересечения множества решений с некоторым заранее заданным интервальным вектором-брусом $\mathbf{U} = (U_1, U_2, \dots, U_n)^T \in \mathbb{IR}^n$, т. е. решать не (2.6), а задачу:

Для интервальной линейной системы уравнений $\mathbf{Ax} = \mathbf{b}$ найти интервальный вектор, содержащийся в пересечении множества решений $\Xi(\mathbf{A}, \mathbf{b})$ с некоторым а priori заданным брусом \mathbf{U} .

(2.7)

3. Теоретическая основа

Зафиксируем натуральный индекс $\nu \in \{1, 2, \dots, n\}$ и рассмотрим в пространстве \mathbb{R}^n прямую линию l с параметрическим уравнением

$$\begin{cases} x_1 &= r_1, \\ \dots & \\ x_{\nu-1} &= r_{\nu-1}, \\ x_\nu &= t, \\ x_{\nu+1} &= r_{\nu+1}, \\ \dots & \\ x_n &= r_n, \end{cases} \quad (3.1)$$

где $r_1, \dots, r_{\nu-1}, r_{\nu+1}, \dots, r_n$ — вещественные константы, а t — параметр, пробегающий вещественную ось \mathbb{R} . Эта прямая, которую мы будем называть “пробной”, параллельна ν -й координатной оси и полностью задается указанием $(n-1)$ -мерного вещественного вектора $r = (r_1, \dots, r_{\nu-1}, r_{\nu+1}, \dots, r_n)^\top$. Для явного указания параметров этой прямой мы будем использовать обозначение $l(r)$.

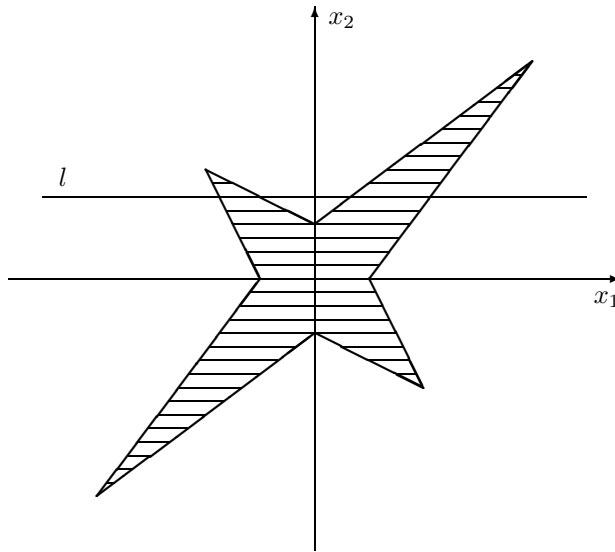


Рис. 3. Пересечение множества решений пробной прямой

Пусть также

$$\underline{\Omega}_\nu(r) = \min\{x_\nu \mid x \in \Xi(\mathbf{A}, \mathbf{b}) \cap l(r)\},$$

$$\overline{\Omega}_\nu(r) = \max\{x_\nu \mid x \in \Xi(\mathbf{A}, \mathbf{b}) \cap l(r)\}$$

суть соответственно наименьшее и наибольшее значения ν -й координаты точек из пересечения $l(r)$ с множеством решений интервальной линейной системы (1.1), см. рис. 3 (в случае $\Xi(\mathbf{A}, \mathbf{b}) \cap l(r) = \emptyset$ мы полагаем $\underline{\Omega}_\nu(r) = +\infty$ и $\overline{\Omega}_\nu(r) = -\infty$). Наша ближайшая цель — вывод явных выражений для функций $\underline{\Omega}_\nu(r)$ и $\overline{\Omega}_\nu(r)$, которые, в некотором роде, “отслеживают” конфигурацию границы множества решений ИСЛАУ.

“Подставим” параметрическое уравнение пробной прямой (3.1) в исходную интервальную систему $\mathbf{Ax} = \mathbf{b}$, которая превратится при этом в “распавшуюся” систему m одномерных линейных уравнений с интервальными коэффициентами и одной переменной t :

$$\begin{cases} \mathbf{a}_{1\nu}t + \sum_{j=1, j \neq \nu}^n \mathbf{a}_{1j}r_j = \mathbf{b}_1, \\ \vdots \\ \mathbf{a}_{m\nu}t + \sum_{j=1, j \neq \nu}^n \mathbf{a}_{mj}r_j = \mathbf{b}_m. \end{cases} \quad (3.2)$$

Содержательный смысл этой процедуры состоит в следующем. При подстановке параметрического уравнения (3.1) в точечную систему $Ax = b$ мы получаем некоторую систему из m одномерных уравнений, которая совпадает по структуре с (3.2), но имеет вещественные коэффициенты. Далее варьируем элементы a_{ij} матрицы и элементы b_i вектора правой части в пределах заданных для них границ \mathbf{a}_{ij} и \mathbf{b}_i соответственно. Ясно, что множество всех полученных таким образом точечных систем уравнений как раз таки образует (3.2).

При сделанных нами предположениях о неотрицательности элементов матрицы интервальной линейной системы все интервалы $\mathbf{a}_{i\nu}$, $i = 1, 2, \dots, m$, также неотрицательны, и множество решений i -го уравнения этой системы может быть вычислено по правилам интервальной арифметики как

$$\left(\mathbf{b}_i - \sum_{j=1, j \neq \nu}^n \mathbf{a}_{ij}r_j \right) / \mathbf{a}_{i\nu} \quad (3.3)$$

с единственной оговоркой, что в качестве “/” может потребоваться операция деления в расширенной интервальной арифметике Кахана [7, 11], допускающая нульсодержащие делители.

Напомним, что помимо обычных интервалов из \mathbb{IR} элементами интервальной арифметики Кахана являются множества вида $] - \infty, p] \cup [q, +\infty[$, $p \leq q$, и $] - \infty, p]$, $[q, +\infty[$. Результаты сложения, вычитания, умножения и деления \mathbf{a}/\mathbf{b} при $0 \notin \mathbf{b}$ в классической интервальной арифметике и арифметике Кахана полностью совпадают. Но в арифметике Кахана дополнительно определено деление обычных интервалов \mathbf{a} и \mathbf{b} , \mathbf{a}/\mathbf{b} с $0 \in \mathbf{b}$. Чтобы сделать нашу работу самодостаточной, мы выпишем соответствующие результаты в развернутой форме:

$$\mathbf{a}/\mathbf{b} = \frac{[\underline{\mathbf{a}}, \overline{\mathbf{a}}]}{[\underline{\mathbf{b}}, \overline{\mathbf{b}}]} = \begin{cases} \mathbf{a} \cdot [1/\overline{\mathbf{b}}, 1/\underline{\mathbf{b}}], & \text{если } 0 \notin \mathbf{b}, \\] - \infty, +\infty[, & \text{если } 0 \in \mathbf{a} \text{ и } 0 \in \mathbf{b}, \\ [\overline{\mathbf{a}}/\underline{\mathbf{b}}, +\infty[, & \text{если } \overline{\mathbf{a}} < 0 \text{ и } \underline{\mathbf{b}} < \overline{\mathbf{b}} = 0, \\] - \infty, \overline{\mathbf{a}}/\underline{\mathbf{b}}] \cup [\overline{\mathbf{a}}/\underline{\mathbf{b}}, +\infty[, & \text{если } \overline{\mathbf{a}} < 0 \text{ и } \underline{\mathbf{b}} < 0 < \overline{\mathbf{b}}, \\] - \infty, \overline{\mathbf{a}}/\overline{\mathbf{b}}], & \text{если } \overline{\mathbf{a}} < 0 \text{ и } 0 = \underline{\mathbf{b}} < \overline{\mathbf{b}}, \\] - \infty, \underline{\mathbf{a}}/\underline{\mathbf{b}}], & \text{если } 0 < \underline{\mathbf{a}} \text{ и } \underline{\mathbf{b}} < \overline{\mathbf{b}} = 0, \\] - \infty, \underline{\mathbf{a}}/\underline{\mathbf{b}}] \cup [\underline{\mathbf{a}}/\overline{\mathbf{b}}, +\infty[, & \text{если } 0 < \underline{\mathbf{a}} \text{ и } \underline{\mathbf{b}} < 0 < \overline{\mathbf{b}}, \\ [\underline{\mathbf{a}}/\overline{\mathbf{b}}, +\infty[, & \text{если } 0 < \underline{\mathbf{a}} \text{ и } 0 = \underline{\mathbf{b}} < \overline{\mathbf{b}}, \\ \emptyset, & \text{если } 0 \notin \mathbf{a} \text{ и } 0 = \mathbf{b}. \end{cases} \quad (3.4)$$

Подробное описание интервальной арифметики Кахана, формализующей операции над интервалами и их дополнениями, аналогичные (3.4), можно найти в [11], но для наших целей вполне достаточно использовать только формулы (3.4), а затем пересекать результат с обычным интервалом.

Каждое из одномерных уравнений, образующих систему (3.2), мы можем решить отдельно от других, а затем пересечь все получившиеся при этом множества решений (3.3) друг с другом (а также с \mathbf{U}_ν , если решаем постановку задачи (2.7)). В пределах всех интервалов, входящих в систему (3.2), соответствующие коэффициенты изменяются независимо друг от друга, как и в исходной ИСЛАУ, так что замкнутое множество \mathcal{S} , полученное в результате описанного выше раздельного решения уравнений и пересечения их одномерных множеств решений, является в точности множеством значений ν -й координаты точек из $\Xi(\mathbf{A}, \mathbf{b}) \cap l(r)$. Оно может оказаться пустым, если система (3.2) несовместна, но в любом случае

$$\underline{\Omega}_\nu(r) = \min \mathcal{S} \quad \text{и} \quad \overline{\Omega}_\nu(r) = \max \mathcal{S}.$$

Если интервалы $\mathbf{a}_{i\nu}$, $i = 1, 2, \dots, m$, не содержат нуля в своей внутренности, то все множества (3.3) — *связные* интервалы вида $[p, q]$ или $] - \infty, p]$, или $[q, +\infty[$, или же вся вещественная ось \mathbb{R} . Это следует из формул для арифметических операций в классической интервальной арифметике и формул (3.4) для деления в расширенной интервальной арифметике Кахана. Далее, пересечение получающихся интервалов-результатов тривиально организуется по формуле

$$\mathbf{x} \cap \mathbf{y} = [\max\{\underline{\mathbf{x}}, \underline{\mathbf{y}}\}, \min\{\overline{\mathbf{x}}, \overline{\mathbf{y}}\}].$$

Таким образом, в точках эффективной области определения функции $\underline{\Omega}_\nu(r)$ (т. е. когда пересечение множеств (3.3) непусто) мы имеем

$$\underline{\Omega}_\nu(r) = \max_{1 \leq i \leq m} \left\{ \left(\mathbf{b}_i - \sum_{j=1, j \neq \nu}^n \mathbf{a}_{ij} r_j \right) / \mathbf{a}_{i\nu} \right\}, \quad (3.5)$$

если множество решений ограничено, и

$$\underline{\Omega}_\nu(r) = \max \left\{ \max_{1 \leq i \leq m} \left\{ \left(\mathbf{b}_i - \sum_{j=1, j \neq \nu}^n \mathbf{a}_{ij} r_j \right) / \mathbf{a}_{i\nu} \right\}, \underline{\mathbf{U}}_\nu \right\}, \quad (3.6)$$

если множество решений неограничено, и мы решаем задачу (2.7) (тогда подчеркивание означает взятие инфимума, а не нижнего конца интервала).

Аналогично, в точках эффективной области определения функции $\overline{\Omega}_\nu(r)$ имеет место

$$\overline{\Omega}_\nu(r) = \min_{1 \leq i \leq m} \left\{ \overline{\left(\mathbf{b}_i - \sum_{j=1, j \neq \nu}^n \mathbf{a}_{ij} r_j \right) / \mathbf{a}_{i\nu}} \right\}, \quad (3.7)$$

если множество решений ограничено, и

$$\overline{\Omega}_\nu(r) = \min \left\{ \min_{1 \leq i \leq m} \left\{ \overline{\left(\mathbf{b}_i - \sum_{j=1, j \neq \nu}^n \mathbf{a}_{ij} r_j \right) / \mathbf{a}_{i\nu}} \right\}, \overline{\mathbf{U}}_\nu \right\}, \quad (3.8)$$

если множество решений неограничено, и мы решаем задачу (2.7) (тогда надчеркивание означает взятие супремума).

Предложение. Если в интервальной линейной системе $Ax = b$ матрица A неотрицательна, то все функции $\underline{\Omega}_\nu(r)$ и $\bar{\Omega}_\nu(r)$, $\nu = 1, 2, \dots, n$, являются монотонно невозрастающими по любой переменной на эффективных областях своего определения.

Доказательство. Оно основывается на следующем простом факте: как нижняя, так и верхняя огибающие (т. е. поточечные минимум и максимум) любого семейства монотонно невозрастающих функций также являются невозрастающими функциями.

Заметим, что если $a_{ij} \geq 0$ и $a_{i\nu} \geq 0$, то для всех i, j и ν выражения

$$\frac{\left(\text{конец интервала } b_i \right) - \sum_{j=1, j \neq \nu}^n \left(\text{конец интервала } a_{ij} \right) r_j}{\text{конец интервала } a_{i\nu}} \tag{3.9}$$

являются монотонно невозрастающими по любому из аргументов r_j , $j = 1, \dots, \nu - 1, \nu + 1, \dots, n$ (при условии, что остальные аргументы фиксированы). Интервальные арифметические операции в \mathbb{IR} определяются, как известно, таким образом, что концы результирующих интервалов являются либо минимумами, либо максимумами результатов арифметических операций между концами интервалов-операндов. Отсюда следует, что для любого $i = 1, 2, \dots, m$ функции

$$\underline{\omega}_{i\nu}(r) = \frac{\left(b_i - \sum_{j=1, j \neq \nu}^n a_{ij} r_j \right)}{a_{i\nu}}$$

суть нижние огибающие для (3.9), а функции

$$\bar{\omega}_{i\nu}(r) = \frac{\left(b_i - \sum_{j=1, j \neq \nu}^n a_{ij} r_j \right)}{a_{i\nu}}$$

суть верхние огибающие для (3.9). При этом все они также невозрастающие по любому своему аргументу r_j , $j = 1, \dots, \nu - 1, \nu + 1, \dots, n$.

Далее, для всех ν такими же невозрастающими являются функции $\underline{\Omega}_\nu(r)$, которые в силу (3.5)–(3.6) — верхние огибающие всех $\underline{\omega}_{i\nu}(r)$, $i = 1, 2, \dots, m$ (и, возможно, константы \underline{U}_ν), а также функции $\bar{\Omega}_\nu(r)$, которые, в силу (3.7)–(3.8) — нижние огибающие для всех $\underline{\omega}_{i\nu}(r)$, $i = 1, 2, \dots, m$ (и, возможно, константы \bar{U}_ν). \square

Например, для известной интервальной линейной системы Хансена [3, 6]

$$\begin{pmatrix} [2, 3] & [0, 1] \\ [1, 2] & [2, 3] \end{pmatrix} x = \begin{pmatrix} [0, 120] \\ [60, 240] \end{pmatrix}$$

множество решений, изображенное на рис. 4, в целом является невыпуклым, но имеет веретенообразную “монотонную” конфигурацию. В то же время, в силу Предложения форма множества решений, подобная той, что представлена на рис. 3, с выпирающими в разных направлениях шипами, невозможна для множеств решений двумерных интервальных линейных систем с неотрицательными матрицами.

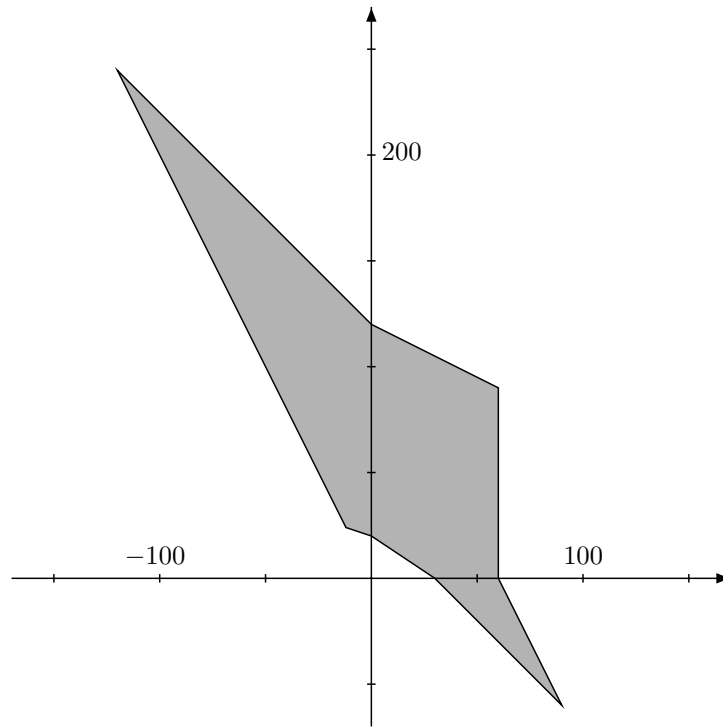


Рис. 4. Множество решений системы Хансена

Еще более наглядной иллюстрацией Предложения является интервальная линейная система

$$\begin{pmatrix} 3.5 & [0, 2] & [0, 2] \\ [0, 2] & 3.5 & [0, 2] \\ [0, 2] & [0, 2] & 3.5 \end{pmatrix} x = \begin{pmatrix} [-1, 1] \\ [-1, 1] \\ [-1, 1] \end{pmatrix},$$

множество решений которой изображено на рис. 7 (оно воспроизведено также на суперобложке книги А. Ноймайера [12], но в несколько другой проекции). Внимательный наблюдатель заметит, что, несмотря на кажущуюся беспорядочность и неструктурированность конфигурации этого множества, ограничивающие его поверхности все равно монотонны! Рассмотренные примеры иллюстрируют также еще одну особенность функций $\underline{\Omega}_\nu(r)$ и $\overline{\Omega}_\nu(r)$ — их возможную разрывность, которая возникает из-за того, что концами интервальных элементов матрицы ИСЛАУ являются нули.

Теорема. Если в интервальной линейной системе $\mathbf{A}x = \mathbf{b}$ матрица \mathbf{A} неотрицательна, то для любых двух точек $y, z \in \Xi(\mathbf{A}, \mathbf{b})$, таких что $y \leq z$ в покомпонентном смысле, интервальный вектор $[y, z]$ тоже является подмножеством множества решений $\Xi(\mathbf{A}, \mathbf{b})$.

Доказательство. Из самого определения функций $\underline{\Omega}_\nu(r)$ и $\overline{\Omega}_\nu(r)$ следует, что для любого $r \in \mathbb{R}^{n-1}$ и любого индекса $\nu \in \{1, 2, \dots, n\}$

$$\underline{\Omega}_\nu(r) \leq \{x_\nu \mid x \in \Xi(\mathbf{A}, \mathbf{b}) \cap l(r)\} \leq \overline{\Omega}_\nu(r),$$

причем оба неравенства точны. Но при сделанном нами допущении о неотрицательности матрицы \mathbf{A} справедливо даже большее:

$$\{x_\nu \mid x \in \Xi(\mathbf{A}, \mathbf{b}) \cap l(r)\} = [\underline{\Omega}_\nu(r), \overline{\Omega}_\nu(r)],$$

так как множество $\{x_\nu \mid x \in \Xi(\mathbf{A}, \mathbf{b}) \cap l(r)\}$ связно. Следовательно, множество решений $\Xi(\mathbf{A}, \mathbf{b})$ есть в точности пересечение надграфика функции $\underline{\Omega}_\nu(r)$ и подграфика функции $\overline{\Omega}_\nu(r)$ (см. рис. 5). Утверждение доказываемой теоремы вытекает поэтому из того факта, что функции $\underline{\Omega}_\nu(r)$ и $\overline{\Omega}_\nu(r)$ являются монотонно невозрастающими. \square

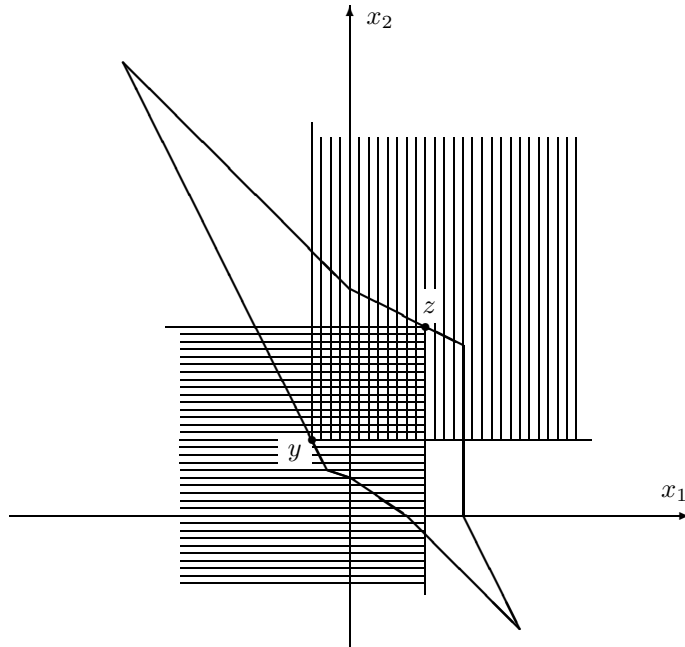


Рис. 5. Иллюстрация утверждения теоремы

Таким образом, структура множества решений ИСЛАУ с неотрицательной матрицей является весьма специальной, и для его внутреннего оценивания могут быть построены эффективные алгоритмы, даже полиномиальные по сложности в случае, когда уже известна какая-то точка из $\Xi(\mathbf{A}, \mathbf{b}) \neq \emptyset$.

4. Алгоритм

Псевдокод алгоритма для построения максимальной внутренней интервальной оценки множеств решений ИСЛАУ с неотрицательными матрицами, вытекающий из теории п. 3, приведен в таблице. Знаком “ \leftarrow ” мы всюду далее обозначаем оператор присваивания.

Дадим необходимые пояснения к этому псевдокоду. Алгоритм осуществляет построение нижней y и верхней z границ интервального вектора $[y, z]$ внутренней оценки для $\Xi(\mathbf{A}, \mathbf{b})$, отправляясь от некоторой начальной точки $\tilde{x} \in \Xi(\mathbf{A}, \mathbf{b})$. Первоначально полагаем $y \leftarrow \tilde{x}$, $z \leftarrow \tilde{x}$, а далее i -й, $i = 1, 2, \dots, n$, шаг алгоритма “раздвигает” точки y и z по i -й координате (см. рис. 6), так что в результате n -го шага получается, вообще говоря, строгое покомпонентное неравенство $y < z$.

Алгоритм NonNeg для внутреннего интервального оценивания
множеств решений ИСЛАУ с неотрицательными матрицами

Вход

Интервальная линейная система $\mathbf{Ax} = \mathbf{b}$ с неотрицательной матрицей \mathbf{A} .

Точка \tilde{x} из оцениваемого множества решений $\Xi(\mathbf{A}, \mathbf{b})$.

Параметры $\lambda, \mu \in]0, 1]$.

Выход

Нижняя y и верхняя z границы интервального вектора $[y, z]$ внутренней
оценки множества решений $\Xi(\mathbf{A}, \mathbf{b})$.

Алгоритм

$y \leftarrow \tilde{x}$;

$z \leftarrow \tilde{x}$;

DO FOR $k = 1$ TO n

$\mathbf{Y} \leftarrow \mathbb{R}$;

$\mathbf{Z} \leftarrow \mathbb{R}$;

 DO FOR $i = 1$ TO n

$$\mathbf{Y} \leftarrow \mathbf{Y} \cap \left(\left(\mathbf{b}_i - \sum_{j=1, j \neq k}^n \mathbf{a}_{ij} y_j \right) / \mathbf{a}_{ik} \right) ;$$

$$\mathbf{Z} \leftarrow \mathbf{Z} \cap \left(\left(\mathbf{b}_i - \sum_{j=1, j \neq k}^n \mathbf{a}_{ij} z_j \right) / \mathbf{a}_{ik} \right) ;$$

 END DO

 IF ($k < n$) THEN

$$y_k \leftarrow \lambda \mathbf{Y} + (1 - \lambda) \tilde{x}_k ;$$

$$z_k \leftarrow (1 - \mu) \tilde{x}_k + \mu \mathbf{Z} ;$$

 ELSE

$$y_k \leftarrow \mathbf{Y} ;$$

$$z_k \leftarrow \mathbf{Z} ;$$

 END IF

END DO

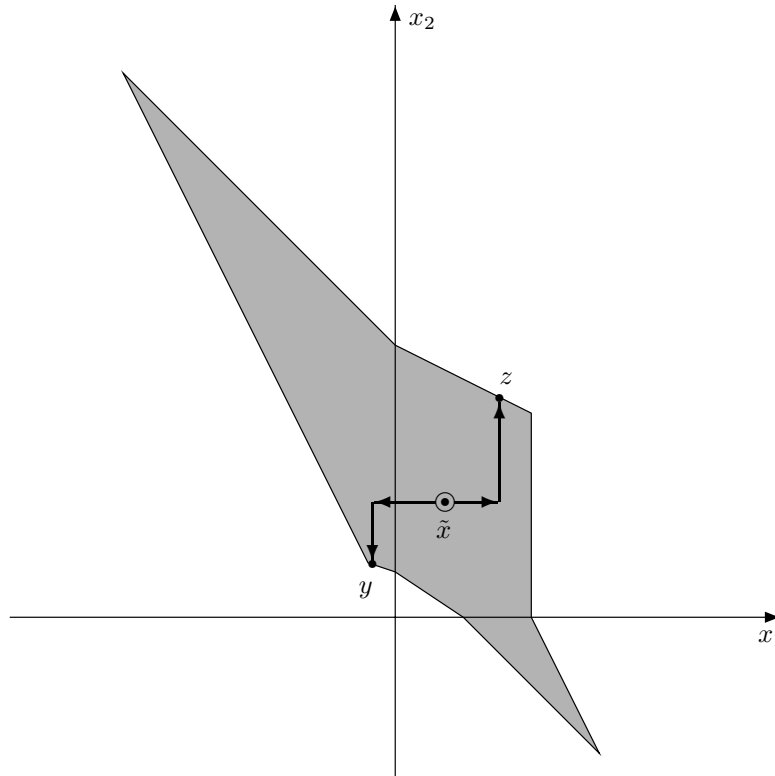


Рис. 6. Как работает алгоритм NonNeg

С помощью вспомогательных безразмерных параметров λ и μ пользователь имеет возможность изменять форму интервальной оценки $[y, z]$ и ее расположение внутри множества решений. Эти параметры регулируют то, насколько отличными от \tilde{x}_i будут сделаны на i -м шаге алгоритма значения y_i и z_i соответственно. Величина $\lambda = 1$ или $\mu = 1$ задает максимально возможное в пределах множества решений $\Xi(\mathbf{A}, \mathbf{b})$ отклонение y_i от \tilde{x}_i в сторону уменьшения и z_i от \tilde{x}_i в сторону увеличения, а нулевые λ или μ соответствовали бы $z_i = y_i = \tilde{x}_i$. Конкретная величина сдвига y_i и z_i относительно \tilde{x}_i определяется из информации о пересечении с множеством решений ИСЛАУ прямых, параллельных i -й координатной оси и проходящих через насчитанные к текущему шагу точки y и z . Методика вычисления таких пересечений подробно изложена нами в п. 3.

Отдельного пояснения требует факт различного подхода к обработке i -й, $i = 1, 2, \dots, n - 1$, и последней n -й компонент векторов y и z . Для получения максимального по включению внутреннего бруса имеет смысл взять точки y и z на границе множества решений, а потому по n -й координате эти точки раздвигаются максимально далеко друг от друга, на противоположные границы $\Xi(\mathbf{A}, \mathbf{b})$, так что параметры λ и μ уже никак не влияют на выполнение этого завершающего шага алгоритма. Из нахождения y и z на границе множества решений следует и максимальность по включению получающейся внутренней оценки.

Для получения “телесной” формы бруса внутренней оценки множества решений наш вычислительный опыт рекомендует выбирать “средние” между 0 и 1 значения для λ и μ , т. е. в районе 0.3–0.7. Слишком близкие к 0 или к 1 значения λ и μ могут привести к “сплющиванию” оценивающего бруса по некоторым координатам. Впрочем, нередко построение удовлетворяющей пользователю оценки может стать лишь результатом интерактивной процедуры, включающей многократный выбор \tilde{x} и варьирование λ и μ .

Приведенная в таблице версия алгоритма рассчитана на интервальные линейные системы с неотрицательными квадратными матрицами и случай ограниченных множеств решений. Для ИСЛАУ с неотрицательными прямоугольными $m \times n$ -матрицами внутренний цикл “DO FOR” следует выполнять до m . Если же предполагается неограниченность множества решений, и мы решаем постановку (2.7), то в начале внешнего цикла “DO FOR” по k нужно инициализировать интервалы \mathbf{Y} и \mathbf{Z} не всей числовой осью \mathbb{R} , а интервалами \mathbf{U}_k , т. е. k -ми компонентами ограничивающего бруса \mathbf{U}_k , данного нам из самой постановки (2.7).

Наконец, отметим, что трудоемкость исполнения алгоритма NonNeg при известной начальной точке \tilde{x} составляет всего $O(mn^2)$ арифметических и логических операций.

5. Выбор начальной точки

Для получения телесной внутренней оценки множества решений желательно, чтобы начальная точка \tilde{x} алгоритма NonNeg лежала во внутренности $\text{int } \Xi(\mathbf{A}, \mathbf{b})$ множества решений. В этом параграфе мы обсудим, как проверять, действительно ли $\tilde{x} \in \text{int } \Xi(\mathbf{A}, \mathbf{b})$, а в некоторых случаях и корректировать положение \tilde{x} .

В самой общей ситуации корректировка точки из множества решений ИСЛАУ представляет собой непростую задачу, так как NP-трудной задачей является распознавание самих этих множеств решений [9]. В качестве начального приближения для точки \tilde{x} из $\Xi(\mathbf{A}, \mathbf{b})$ можно взять, например, решение какой-нибудь точечной системы уравнений с коэффициентами из \mathbf{A} и \mathbf{b} , скажем, “средней” системы

$$(\text{mid } \mathbf{A})x = \text{mid } \mathbf{b}, \quad (5.1)$$

матрица и правая часть которой образованы серединами $\text{mid } \mathbf{a}_{ij}$ и $\text{mid } \mathbf{b}_i$ интервальных элементов \mathbf{a}_{ij} и \mathbf{b}_i соответственно. Если этот выбор по тем или иным причинам неудовлетворителен, можно воспользоваться техникой так называемых “распознающих функционалов”, разработанной автором, в частности, в [4]. Напомним некоторые понятия и факты из этой работы.

Пусть \mathbf{A} — интервальная $m \times n$ -матрица, \mathbf{b} — интервальный m -вектор, и $\langle \mathbf{a} \rangle$ обозначает наименьшее расстояние точек интервала \mathbf{a} до нуля, а $\text{rad } \mathbf{a}$ — радиус интервала \mathbf{a} , т. е.

$$\langle \mathbf{a} \rangle := \begin{cases} \min\{|\underline{\mathbf{a}}|, |\overline{\mathbf{a}}|\}, & \text{если } \mathbf{a} \not\supset 0, \\ 0, & \text{если } \mathbf{a} \supset 0. \end{cases} \quad \text{rad } \mathbf{a} := \frac{1}{2}(\overline{\mathbf{a}} - \underline{\mathbf{a}}).$$

Тогда выражением

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

задается функционал $\text{Uni} : \mathbb{R}^n \rightarrow \mathbb{R}$, такой что принадлежность точки x множеству решений интервальной линейной системы $\mathbf{A}x = \mathbf{b}$ равносильна неотрицательности в x функционала Uni:

$$x \in \Xi(\mathbf{A}, \mathbf{b}) \iff \text{Uni}(x, \mathbf{A}, \mathbf{b}) \geq 0.$$

Таким образом, множество решений соответствующей ИСЛАУ есть лебегово множество $\{x \in \mathbb{R}^n \mid \text{Uni}(x, \mathbf{A}, \mathbf{b}) \geq 0\}$ функционала Uni.

Функционал Uni является вогнутым в каждом ортанте пространства \mathbb{R}^n , а если в интервальной матрице \mathbf{A} некоторые столбцы целиком точечные, то $\text{Uni}(x, \mathbf{A}, \mathbf{b})$ вогнут

и на объединениях нескольких ортантов. Кроме того, функционал $\text{Uni}(x, \mathbf{A}, \mathbf{b})$ достигает конечного максимума на всем пространстве \mathbb{R}^n . Если $\text{Uni}(x, \mathbf{A}, \mathbf{b}) > 0$, то x — точка топологической внутренней $\text{int } \Xi(\mathbf{A}, \mathbf{b})$ множества решений ИСЛАУ. При некоторых дополнительных ограничениях на \mathbf{A} и \mathbf{b} верно и обратное: из принадлежности $x \in \text{int } \Xi(\mathbf{A}, \mathbf{b})$ следует $\text{Uni}(x, \mathbf{A}, \mathbf{b}) > 0$.

Как следствие этих результатов, мы естественно приходим к следующему практическому рецепту коррекции начальной точки \tilde{x} для алгоритма **NonNeg**. Если по каким-либо причинам \tilde{x} не удовлетворяет нас (например, не находится во внутренней множества решений, о чем может свидетельствовать значение $\text{Uni}(\tilde{x}) \leq 0$), то, пользуясь градиентным подъемом, пытаемся достичь лучшего значения распознающего функционала Uni . В случае, когда полученное новое значение функционала строго больше нуля, мы можем быть уверенными, что оказались во внутренней множества решений.

6. Численные эксперименты

Пример 1. Для интервальной линейной системы Хансена

$$\begin{pmatrix} [2, 3] & [0, 1] \\ [1, 2] & [2, 3] \end{pmatrix} x = \begin{pmatrix} [0, 120] \\ [60, 240] \end{pmatrix}$$

применение алгоритма **NonNeg** с параметрами $\lambda = \mu = 1$ и начальной точкой в виде решения “средней системы” (5.1) приводит к ответу

$$\begin{pmatrix} [-25.909, 60] \\ [51.818, 90] \end{pmatrix},$$

а с параметрами $\lambda = \mu = 0.7$ получается внутренняя оценка

$$\begin{pmatrix} [-13.022, 47.114] \\ [26.045, 96.443] \end{pmatrix}.$$

Как видим, получающиеся интервальные векторы внутренней оценки в обоих случаях являются максимальными по включению и покрывают значительные куски множества решений.

Пример 2. В качестве второго примера рассмотрим интервальную линейную систему с $n \times n$ -матрицей

$$\begin{pmatrix} t & [0, 2] & \cdots & [0, 2] \\ [0, 2] & t & \cdots & [0, 2] \\ \vdots & \vdots & \ddots & \vdots \\ [0, 2] & [0, 2] & \cdots & t \end{pmatrix},$$

которая была предложена А. Ноймайером в книге [12]. Здесь t — вещественный параметр, с помощью которого можно управлять свойствами матрицы. В частности, при $t > n$ для четных n и при $t > \sqrt{n^2 - 1}$ для нечетных n матрица Ноймайера неособенна (невырождена), т. е. содержит только неособенные вещественные матрицы. Кроме того, матрица Ноймайера неотрицательна, и потому интервальные линейные системы с ней могут служить тестовыми для развитой нами в этой работе методики.

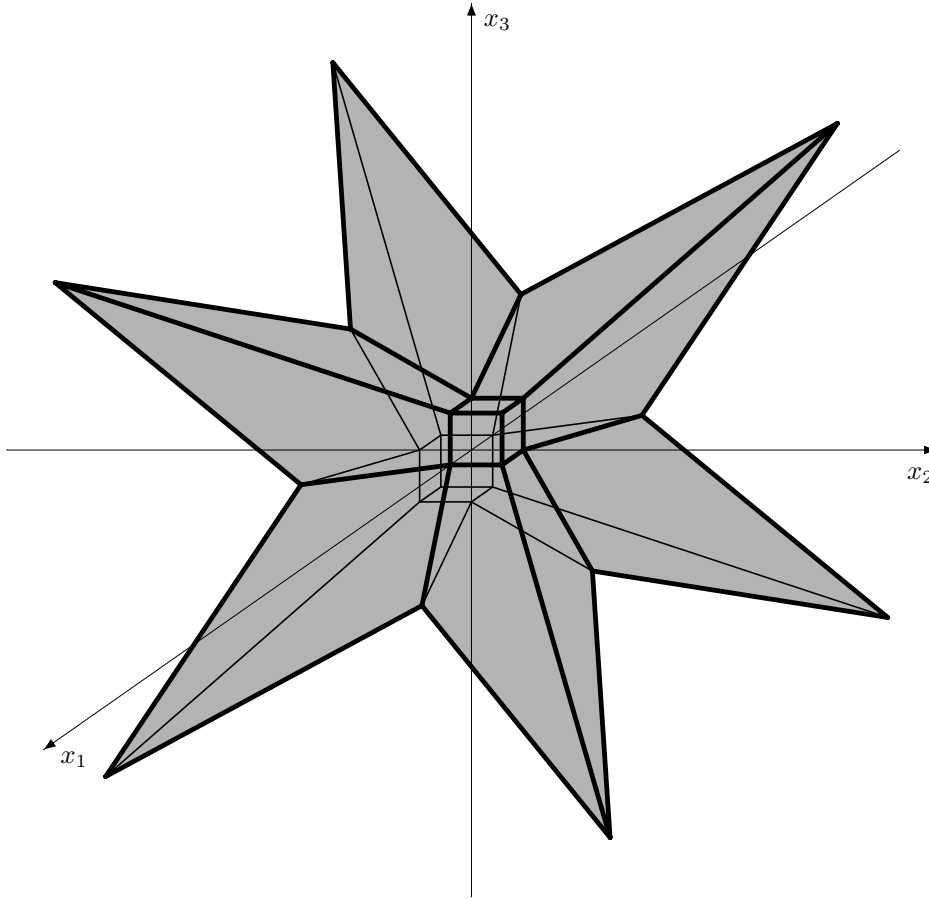


Рис. 7. Множество решений трехмерной системы Ноймайера

При $t = 3.5$, размерности 3 и правой части $([-1, 1], [-1, 1], [1, 1])^T$ множество решений этой интервальной системы имеет вид, изображенный на рис. 7. Взяв в качестве начальной точки \tilde{x} решение “средней системы” — нулевой вектор, — после применения алгоритма внутреннего оценивания NonNeg с параметрами $\lambda = \mu = 1$ получим результат

$$\begin{pmatrix} [-0.285714, 0.285714] \\ [-0.285714, 0.285714] \\ [-0.285714, 0.285714] \end{pmatrix},$$

совпадающий с точной внутренней оценкой той части множества решений (“кубика”), которая прилегает к началу координат в положительном и отрицательном ортантах. Аналогично предыдущему примеру, столь хорошие результаты оценивания при граничных значениях параметров λ и μ имеют причиной специальную конфигурацию множества решений. Именно, так как в матрице ИСЛАУ “много” элементов имеют нулевые концы, то некоторые из граней множества решений оказываются параллельными координатным плоскостям.

Наши выводы подтверждает

Пример 3. Рассмотрим интервальную линейную систему

$$\begin{pmatrix} 3.5 & [1, 2] & [1, 2] \\ [1, 2] & 3.5 & [1, 2] \\ [1, 2] & [1, 2] & 3.5 \end{pmatrix} x = \begin{pmatrix} [-1, 1] \\ [-1, 1] \\ [-1, 1] \end{pmatrix},$$

которая отличается от системы Ноймайера тем, что вместо нулей левыми концами внедиагональных элементов матрицы взяты единицы. Для нее внутреннее оценивание множества решений алгоритмом NonNeg с нулевой начальной точкой \tilde{x} и параметрами $\lambda = \mu = 1$ приводит к “сплюсненному” интервальному ответу

$$\begin{pmatrix} [-0.285714, 0.285714] \\ 0 \\ 0 \end{pmatrix}.$$

Ясно, что в этом случае правильнее взять параметры λ и μ внутри отрезка $]0, 1[$. Например, мы получаем телесную внутреннюю оценку

$$\begin{pmatrix} [-0.2, 0.2] \\ [-0.16, 0.16] \\ [-0.14, 0.14] \end{pmatrix}$$

при использовании алгоритма NonNeg с $\lambda = \mu = 0.7$.

Список литературы

- [1] **Алефельд Г., Херцбергер Ю.** Введение в интервальные вычисления. — М.: Мир, 1987.
- [2] **Воцинин А.П., Сотиров Г.Р.** Оптимизация в условиях неопределенности. — Москва, София: Изд-во МЭИ “Техника”, 1989.
- [3] **Калмыков С.А., Шокин Ю.И., Юлдашев З.Х.** Методы интервального анализа. — Новосибирск: Наука, 1986.
- [4] **Шарый С.П.** О характеристике объединенного множества решений интервальной линейной алгебраической системы. — Красноярск, 1990. — Деп. в ВИНТИ № 726–В91.
- [5] **Шарый С.П.** Алгебраический подход к анализу линейных статических систем с интервальной неопределенностью // Известия РАН. Теория и системы управления. — 1997. — № 3. — С. 51–61.
- [6] **Hansen E.R.** On linear algebraic equations with interval coefficients // Topics in Interval Analysis / E. Hansen, ed. — Oxford: Clarendon Press, 1969. — P. 35–46.
- [7] **Kearfott R.B.** Rigorous global search: continuous problems. — Dordrecht: Kluwer, 1996.
- [8] **Kearfott R.B., Nakao M., Neumaier A., Rump S., Shary S.P., van Hentenryck P.** Standardized notation in interval analysis. — <http://www.ict.nsc.ru/interval/InteNotation.ps>.
- [9] **Kreinovich V., Lakeyev A., Rohn J., Kahl P.** Computational complexity and feasibility of data processing and interval computations. — Dordrecht: Kluwer, 1997.
- [10] **Куржанский А. Б.** Задача идентификации — теория гарантированных оценок // Автоматика и телемеханика. — 1991. — № 4. — С. 3–26.

- [11] **Laveuve S.E.** Definition einer Kahan-Arithmetic und ihre Implementierung // Interval Mathematics / K. Nickel, ed. — Berlin: Springer, 1975. — (Lect. Notes in Computer Science. — Vol. 29. — P. 236–245).
- [12] **Neumaier A.** Interval methods for systems of equations. — Cambridge: Cambridge University Press, 1990.
- [13] **Oettli W.** On the solution set of a linear system with inaccurate coefficients // SIAM Journal on Numerical Analysis. — 1965. — Vol. 2, №1. — P. 115–118.
- [14] **Shary S.P.** Solving interval linear systems with nonnegative matrices // Scientific Computations and Mathematical Modelling // Proc. of the International Conference MMSC-93 / S. M. Markov, ed. — Sofia: DATECS Publishing, 1993. — P. 179–182.
- [15] **Shary S.P.** A new technique in systems analysis under interval uncertainty and ambiguity // Reliable Computing. — 2002. — Vol. 8, № 5. — P. 321–418.
— <http://www.ict.nsc.ru/shary/Papers/ANewTech.ps>.

Институт вычислительных технологий СО РАН,
просп. Акад. Лаврентьева, 6,
Новосибирск, 630090,
E-mail: shary@ict.nsc.ru

*Статья поступила
20 июля 2005 г.
Переработанный вариант
11 августа 2005 г.*