ЕЩЕ РАЗ О ВНУТРЕННЕМ ОЦЕНИВАНИИ МНОЖЕСТВ РЕШЕНИЙ ИНТЕРВАЛЬНЫХ ЛИНЕЙНЫХ СИСТЕМ*

С. П. Шарый
Институт вычислительных технологий СО РАН, Новосибирск, Россия
e-mail: shary@ict.nsc.ru

For interval linear equations systems of the form $\mathbf{A}x = \mathbf{b}$, we consider the problem of inner estimation of its *solution set* formed by all the solutions to the point systems $Ax = b$ with $A \in \mathbf{A}$ and $b \in \mathbf{b}$. The so-called "center" approach to the problem is developed that constructs an interval solution around an a priori known "center point" from the solution set. Finally, we demonstrate that the initial problem can be reduced to minimization of a special quasiconcave function

Введение

Предметом рассмотрения в настоящей статье являются интервальные системы линейных алгебраических уравнений (ИСЛАУ) вида

$$\mathbf{A}x = \mathbf{b}, \tag{1}$$

где $\mathbf{A} = (\,\mathbf{a}_{ij})$ — интервальная $m \times n$-матрица и $\mathbf{b} = (\,\mathbf{b}_i)$ — интервальный m-вектор. Известно, что для интервальных систем уравнений решения и множества решений могут быть определены весьма разнообразными способами (см., к примеру, [1–6]), но ниже мы ограничимся так называемым *объединенным множеством решений* для (1), которое образовано всевозможными решениями x точечных систем $Ax = b$, когда матрица A и вектор b независимо пробегают $\mathbf{A}$ и $\mathbf{b}$ соответственно. Объединенное множество решений определяется строго как

$$\Xi(\mathbf{A}, \mathbf{b}) = \{\, x \in \mathbb{R}^n \mid (\exists\, A \in \mathbf{A})(\exists\, b \in \mathbf{b})(Ax = b)\}, \tag{2}$$

и ниже мы будем называть его просто *множеством решений* интервальной линейной системы (1), так как другие множества решений нами не исследуются. Точное описание множества решений может расти экспоненциально с размерностью вектора неизвестных n, а потому является практически невозможным уже при n, превосходящем несколько десятков. С другой стороны, в большинстве реальных постановок задач точное описание на самом деле и не нужно. На практике бывает вполне достаточно нахождения *оценки* для

*Поддержано грантом Президента России № НШ-2314.2003.1

множества решений, т. е. приближенного описания, удовлетворяющего содержательному смыслу рассматриваемой задачи.

Здесь нас интересует нахождение возможно более широких *внутренних* интервальных оценок (в виде подмножеств) для множества решений, т. е. мы решаем следующую задачу:

$$\boxed{\begin{array}{c}\text{Найти (по возможности бóльший) брус } \mathbf{U}, \\ \text{содержащийся во множестве решений } \Xi(\mathbf{A}, \mathbf{b}) \\ \text{интервальной линейной системы } \mathbf{A}x = \mathbf{b}.\end{array}} \tag{3}$$

Брусами здесь и далее в работе называются прямоугольные параллелепипеды в пространстве $\mathbb{R}^n$ со сторонами, параллельными координатным осям, т. е. геометрические образы интервальных векторов.

Известны несколько подходов к решению задачи внутреннего оценивания множеств решений интервальных линейных систем уравнений, среди которых для квадратных систем (при $m = n$) вычислительной эффективностью и общностью выделяется так называемый *формальный* (алгебраический) подход, развитый в [3, 5–7]. Тем не менее для произвольных интервальных линейных систем с прямоугольной неквадратной матрицей, когда $m \neq n$, внутреннее оценивание множеств решений по-прежнему представляет собой важную и актуальную задачу. Основываясь на геометрически наглядных соображениях, мы предлагаем ниже простой и весьма общий способ построения бруса, вписанного в $\Xi(\mathbf{A}, \mathbf{b})$, вокруг a priori известной точки-центра из этого множества решений (рис. 1). Показано, что рассматриваемая задача сводится к нахождению максимума некоторой специальной квазивогнутой функции, приближенное значение которого может быть получено весьма элементарными средствами.

Мы не предполагаем при этом у интервальной матрицы $\mathbf{A}$ никаких свойств неособенности, полноранговости и т.п., допуская случай неограниченного множества решений $\Xi(\mathbf{A}, \mathbf{b})$. Единственное необременительное требование на $\mathbf{A}$ состоит в том, что она не должна иметь целиком нулевых строк.

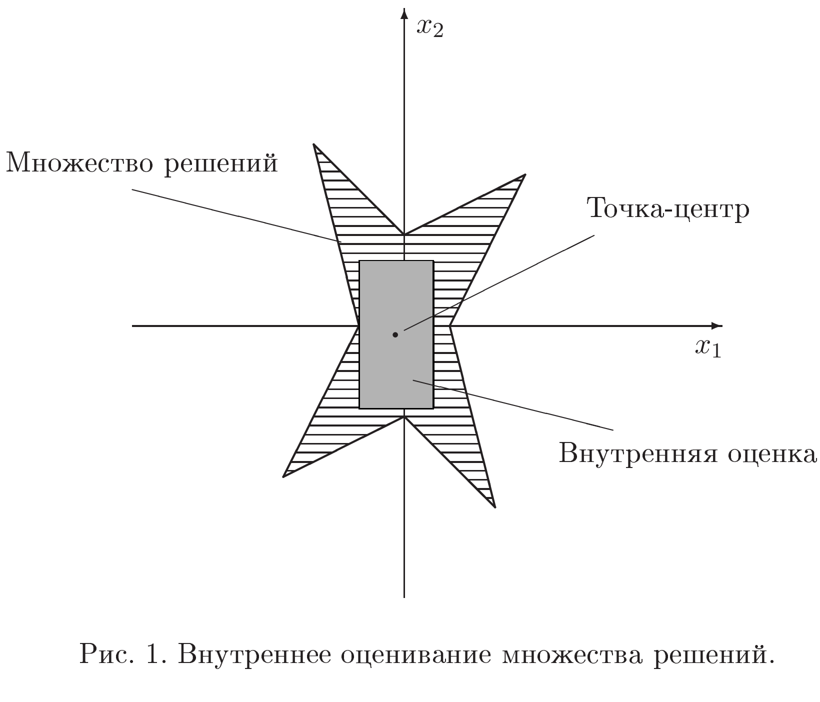

Рис. 1. Внутреннее оценивание множества решений.

Используемые нами обозначения соответствуют недавно принятому проекту международного стандарта [8]. В частности, интервалы и интервальные величины мы обозначаем буквами жирного шрифта — $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, ..., $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$, тогда как неинтервальные (точечные) величины специально не выделяются. Арифметические операции с интервальными величинами — это операции классической интервальной арифметики $\mathbb{IR}$ (см., например, [1, 9–11]). Подчеркивание и надчеркивание — $\underline{\mathbf{a}}$, $\overline{\mathbf{a}}$ — обозначают нижний и верхний концы интервала $\mathbf{a}$, кроме того,

$$\operatorname{mid}\mathbf{a} = \frac{1}{2}(\overline{\mathbf{a}} + \underline{\mathbf{a}}) \ \text{— середина (центр) интервала,}$$

$$\operatorname{rad}\mathbf{a} = \frac{1}{2}(\overline{\mathbf{a}} - \underline{\mathbf{a}}) \ \text{— радиус интервала,}$$

$$|\mathbf{a}| = \max\{\,|\overline{\mathbf{a}}|, |\underline{\mathbf{a}}|\,\} \ \text{— абсолютное значение (модуль) интервала,}$$

$$\langle\mathbf{a}\rangle = \begin{cases} \min\{\,|\overline{\mathbf{a}}|, |\underline{\mathbf{a}}|\,\}, & \text{если } 0 \notin \mathbf{a}, \\ 0 & \text{иначе,} \end{cases} \quad \text{— } \begin{array}{l} \text{наименьшее отклонение точек} \\ \text{интервала от нуля, антипод} \\ \text{абсолютной величины интервала.} \end{array}$$

К интервальным векторам операции взятия середины, радиуса и абсолютного значения будут применяться покомпонентно.

Предполагаемое у читателя знание основ интервального анализа может быть почерпнуто, к примеру, из книг [1, 9–11].

1. Уточнение постановки задачи

В приложениях постановка задачи (3) часто содержит дополнительную информацию о желаемой форме бруса $\mathbf{U} = (\,\mathbf{U}_1, \mathbf{U}_2, \dots, \mathbf{U}_n)^\top$, который должен оценивать $\Xi(\mathbf{A}, \mathbf{b})$ изнутри: ширины компонент $\mathbf{U}$ предполагаются пропорциональными соответствующим компонентам некоторого вещественного положительного вектора

$$w = (w_1, w_2, \dots, w_n), \qquad w_i > 0\,.$$

Иными словами, в постановке (3) дополнительно вводятся весовые коэффициенты w_i для ширин (или радиусов) компонент бруса внутренней оценки $\mathbf{U}$, такие что

$$\operatorname{rad}\mathbf{U}_i / \operatorname{rad}\mathbf{U}_j = w_i / w_j, \qquad i, j = 1, 2, \dots, n.$$

Оказывается, что посредством масштабирования неособенной диагональной матрицей

$$W = \operatorname{diag}\{w_1, w_2, \dots, w_n\}$$

с элементами w_1, w_2, ..., w_n по главной диагонали рассмотрение этого случая может быть сведено к простейшей ситуации, когда $w = (1, 1, \dots, 1)$ и брус $\mathbf{U}$ превращается в кубик, который мы должны вписывать во множество решений некоторой модифицированной интервальной системы уравнений. Более точно, справедливо

Предложение. *Пусть*

$$\tilde{\mathbf{A}} = \mathbf{A}W.$$

Интервальный вектор $\tilde{\mathbf{U}}$ *с одинаковыми ширинами компонент, т. е. такой, что*

$$\operatorname{rad}\tilde{\mathbf{U}}_i = \operatorname{rad}\tilde{\mathbf{U}}_j, \qquad i,j = 1,2,\ldots,n,$$

является решением задачи внутреннего оценивания (3) *для модифицированной интервальной системы* $\tilde{\mathbf{A}}x = \mathbf{b}$ *тогда и только тогда, когда интервальный вектор* $\mathbf{U} = W\tilde{\mathbf{U}}$ *с желаемым отношением ширин компонент есть решение задачи внутреннего оценивания* (3) *для исходной системы* $\mathbf{A}x = \mathbf{b}$.

Доказательство. Чтобы обосновать сделанное утверждение, воспользуемся характеризацией Х. Бека [11] множества решений интервальной линейной системы (1):
для $x \in \mathbb{R}^n$

$$x \in \Xi(\mathbf{A},\mathbf{b}) \qquad \Longleftrightarrow \qquad \mathbf{A}x \cap \mathbf{b} \neq \emptyset. \tag{4}$$

В частности, для модифицированной системы уравнений

$$\tilde{x} \in \Xi(\tilde{\mathbf{A}},\mathbf{b}) \qquad \Longleftrightarrow \qquad \tilde{\mathbf{A}}\tilde{x} \cap \mathbf{b} \neq \emptyset. \tag{5}$$

Умножение на матрицу W задает взаимно-однозначное соответствие между точками брусов $\mathbf{U}$ и $\tilde{\mathbf{U}}$ по правилу

$$x \;\rightleftarrows\; \tilde{x} = Wx$$

для $x \in \mathbf{U}$ и $\tilde{x} \in \tilde{\mathbf{U}}$. Далее, для каждой пары взаимно соответствующих друг другу x и $\tilde{x}$ справедливо

$$\mathbf{A}x = \mathbf{A}WW^{-1}x = \tilde{\mathbf{A}}\tilde{x},$$

так что соотношения в правых частях эквивалентностей (4) и (5) выполняются или не выполняются одновременно. Кроме того, для любых $i,j = 1,2,\ldots,n$ в самом деле

$$\operatorname{rad}\mathbf{U}_i / \operatorname{rad}\mathbf{U}_j = w_i/w_j,$$

как и требовалось. $\square$

Итак, всюду ниже мы вправе считать, что в задаче (3) внутреннего оценивания множества решений ИСЛАУ требуется отыскание интервального вектора $\mathbf{U}$ с компонентами равной ширины, такого что $\mathbf{U} \subseteq \Xi(\mathbf{A},\mathbf{b})$.

2. Идея подхода

Если мы найдем какую-нибудь точку из множества решений $\Xi(\mathbf{A},\mathbf{b})$, то далее сможем использовать ее как "центр", вокруг которого будет построено интервальное решение задачи (3) (см. рис. 1). Это основная идея развиваемого нами подхода, который, таким образом, может быть назван "центровым", совершенно аналогично подходу, примененному в [1, 4] для внутреннего оценивания так называемого *допустимого множества решений* ИСЛАУ. Итак,

сначала ищем некоторую точку $t \in \Xi(\mathbf{A},\mathbf{b})$,

затем используем известные координаты t для вычисления размера
кубика внутренней оценки, который имеет центр в этой точке t.

Формула для размеров интервального решения задачи (3) будет выведена нами ниже (см. § 4), и основную роль в ней будет играть взятие максимума рационального выражения с модулями по некоторому брусу, так что при известной точке $t \in \Xi(\mathbf{A},\mathbf{b})$ решение задачи о внутреннем оценивании множества решений сводится, по существу, к оптимизации на брусе. Мы подробно рассмотрим его в § 5.

3. Выбор центра внутренней оценки

Как известно, задачи распознавания того, пусто или непусто множество решений интервальной линейной системы $\Xi(\mathbf{A},\mathbf{b})$, а также нахождения хотя бы одной точки из $\Xi(\mathbf{A},\mathbf{b})$, в общем случае являются NP-трудными [2, 12]. Универсальный метод решения этих задач может быть основан на том факте, что пересечения множеств решений ИСЛАУ с каждым из ортантов пространства $\mathbb{R}^n$ являются выпуклыми полиэдральными множествами, для которых уравнения граничных гиперплоскостей легко выписываются по матрице и правой части ИСЛАУ (см., к примеру, [13, 14]). Следовательно, пустота или непустота пересечения $\Xi(\mathbf{A},\mathbf{b})$ с каждым из ортантов $\mathbb{R}^n$ может быть выявлена путем решения некоторой системы линейных неравенств, например, хорошо разработанными методами линейного программирования. В целом распознавание множества решений ИСЛАУ и нахождение какой-либо его точки потребуют не более чем 2^n решений систем линейных неравенств, причем этот результат не может быть принципиально улучшен.

Таким образом, в самой общей ситуации отыскание и корректировка точки из множества решений ИСЛАУ являются весьма непростым делом, и поэтому ниже имеет смысл дать набор частных рецептов для решения проблемы в тех или иных конкретных ситуациях.

Рассмотрим прежде всего случай квадратной $n\times n$-матрицы $\mathbf{A}$. Если она неособенная (т. е. неособенны все $A\in\mathbf{A}$), то точку t из $\Xi(\mathbf{A},\mathbf{b})$ можно получить, решив какую-нибудь точечную систему уравнений $At=b$ с A из $\mathbf{A}$ и b из $\mathbf{b}$, скажем, “среднюю” систему

$$(\operatorname{mid}\mathbf{A})\,t=\operatorname{mid}\mathbf{b}.$$

Проверка неособенности интервальной матрицы $\mathbf{A}$ может быть выполнена, например, методами, описанными в [15].

Предположим теперь, что интервальная матрица $\mathbf{A}$ — особенная, т. е. содержит особенные точечные матрицы. Хорошо известно, что во множестве всех вещественных $n\times n$-матриц особенные матрицы образует гладкое многообразие коразмерности 1, являясь весьма “тощим” множеством с лебеговой мерой нуль в $\mathbb{R}^{n\times n}$. Следовательно, если все элементы данной интервальной матрицы $\mathbf{A}$ имеют ненулевые ширины, то путем подходящего варьирования элементов точечной $n\times n$-матрицы в пределах $\mathbf{A}$ мы всегда можем надеяться попасть на какую-нибудь неособенную матрицу A. И вновь для отыскания точки t достаточно решить систему $At=b$ с каким-то $b\in\mathbf{b}$.

Что делать в случае прямоугольной системы уравнений? Иногда здесь оказывается полезной техника так называемых *распознающих функционалов*, разработанная автором в [4, 16]. Напомним некоторые факты и понятия.

Теорема. *Пусть* $\mathbf{A}$ *— интервальная* $m\times n$*-матрица,* $\mathbf{b}$ *— интервальный* m*-вектор, и выражением*

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

задается функционал $\operatorname{Uni}:\mathbb{R}^n\to\mathbb{R}$*. Принадлежность точки* x *множеству решений интервальной линейной системы* $\mathbf{A}x=\mathbf{b}$ *равносильна неотрицательности в* x *функционала* Uni,

$$x\in\Xi(\mathbf{A},\mathbf{b})\qquad\Longleftrightarrow\qquad\operatorname{Uni}(x,\mathbf{A},\mathbf{b})\ge 0,$$

т. е. множество решений $\Xi(\mathbf{A},\mathbf{b})$ *есть лебегово множество* $\{\,x\in\mathbb{R}^n\mid\operatorname{Uni}(x,\mathbf{A},\mathbf{b})\ge 0\,\}$ *функционала* Uni.

Если из контекста понятно, о какой интервальной системе идет речь, то мы будем писать просто $\mathrm{Uni}(x)$ вместо $\mathrm{Uni}(x, \mathbf{A}, \mathbf{b})$.

Доказательство. Точка x принадлежит множеству решений $\Xi(\mathbf{A}, \mathbf{b})$ тогда и только тогда, когда существует матрица $\tilde{A} = (\tilde{a}_{ij}) \in \mathbf{A}$, такая, что

$$\tilde{A}x \in \mathbf{b}.$$

После детального расписывания матрично-векторного произведения и представления интервалов правой части в центральной форме эта принадлежность примет вид

$$\sum_{j=1}^{n} \tilde{a}_{ij} x_j \in \mathrm{mid}\,\mathbf{b}_i + [\,-\mathrm{rad}\,\mathbf{b}_i, \mathrm{rad}\,\mathbf{b}_i\,], \qquad i = 1, 2, \dots, m.$$

Добавив теперь по $(-\mathrm{mid}\,\mathbf{b}_i)$ к обеим частям полученных включений, придем к эквивалентным соотношениям

$$\sum_{j=1}^{n} \tilde{a}_{ij} x_j \; - \mathrm{mid}\,\mathbf{b}_i \in [\,-\mathrm{rad}\,\mathbf{b}_i, \mathrm{rad}\,\mathbf{b}_i\,], \qquad i = 1, 2, \dots, m,$$

которые, в свою очередь, равносильны

$$\left| \sum_{j=1}^{n} \tilde{a}_{ij} x_j \; - \mathrm{mid}\,\mathbf{b}_i \right| \leq \mathrm{rad}\,\mathbf{b}_i \;,$$

и поэтому

$$\mathrm{rad}\,\mathbf{b}_i - \left| \mathrm{mid}\,\mathbf{b}_i - \sum_{j=1}^{n} \tilde{a}_{ij} x_j \right| \geq 0 \tag{6}$$

для всех $i = 1, 2, \dots, m$.

Итак, $x \in \Xi(\mathbf{A}, \mathbf{b})$, если и только если для каждого фиксированного i существуют такие $\tilde{a}_{ij} \in \mathbf{a}_{ij}$, $j = 1, 2, \dots, n$, что оказываются справедливыми неравенства (16). Но это эквивалентно выполнению для $i = 1, 2, \dots, m$ условий

$$\max_{\substack{\tilde{a}_{ij} \in \mathbf{a}_{ij},\\ j=1,2,\dots,n}} \left\{ \mathrm{rad}\,\mathbf{b}_i - \left| \mathrm{mid}\,\mathbf{b}_i - \sum_{j=1}^{n} \tilde{a}_{ij} x_j \right| \right\} \;\geq\; 0. \tag{7}$$

Внося максимум внутрь скобки и учитывая, что естественное интервальное расширение выражения под знаком модуля совпадает с его областью значений, вместо (7) получим для $i = 1, 2, \dots, m$

$$\left\{ \mathrm{rad}\,\mathbf{b}_i - \left\langle \mathrm{mid}\,\mathbf{b}_i - \sum_{j=1}^{n} \mathbf{a}_{ij} x_j \right\rangle \right\} \;\geq\; 0. \tag{8}$$

Далее взятием минимума мы можем свернуть m условий (8) в одно, заключая, что точка x принадлежит множеству $\Xi(\mathbf{A}, \mathbf{b})$ действительно в том и лишь в том случае, если

$$\min_{1 \leq i \leq m} \left\{ \mathrm{rad}\,\mathbf{b}_i - \left| \mathrm{mid}\,\mathbf{b}_i - \sum_{j=1}^{n} \mathbf{a}_{ij} x_j \right| \right\} \;\geq\; 0,$$

как и требовалось. $\square$

Получается, что с помощью знака своих значений функционал Uni "распознает" принадлежность точки множеству $\Xi(\mathbf{A}, \mathbf{b})$. Именно поэтому мы используем по отношению к нему эпитет "распознающий". Справедливы, кроме того, следующие свойства [16]:

1. Функционал Uni — вогнутый в каждом ортанте $\mathbb{R}^n$, а если в интервальной матрице $\mathbf{A}$ некоторые столбцы — целиком вещественные, то $\mathrm{Uni}(x, \mathbf{A}, \mathbf{b})$ вогнут и на объединениях нескольких ортантов.

2. Функционал $\mathrm{Uni}(x, \mathbf{A}, \mathbf{b})$ достигает конечного максимума на всем пространстве $\mathbb{R}^n$.

3. Если $\mathrm{Uni}(x, \mathbf{A}, \mathbf{b}) > 0$, то x — точка топологической внутренности $\mathrm{int}\;\Xi(\mathbf{A}, \mathbf{b})$ множества решений.

4. При некоторых дополнительных ограничениях на $\mathbf{A}$, $\mathbf{b}$ и x верно и обратное: из принадлежности $x \in \mathrm{int}\;\Xi(\mathbf{A}, \mathbf{b})$ следует $\mathrm{Uni}(x, \mathbf{A}, \mathbf{b}) > 0$.

Последние два свойства распознающего функционала позволяют использовать его для исследования принадлежности точек внутренности множества решений. Это имеет особую важность потому, что телесная внутренняя оценка множества решений по нашей методике может быть построена только вокруг точки-центра t, который лежит во внутренности $\mathrm{int}\;\Xi(\mathbf{A}, \mathbf{b})$ множества решений!

Как следствие вышеизложенных результатов получается такой практический рецепт коррекции точки-центра t в "центровом" подходе к решению задачи (3): находим какое-нибудь начальное приближение, а затем, пользуясь градиентным подъемом, пытаемся достичь бóльшего значения распознающего функционала Uni. Если полученное новое значение строго больше нуля, то мы оказались во внутренности множества решений.

Мы не обсуждаем вопрос об оптимизации (наилучшем выборе) центра интервального решения, так как он отчасти выходит за рамки нашего исследования и, с другой стороны, тесно связан с потребностями пользователей, решающих те или иные практические задачи.

4. Формула для размеров внутренней оценки

Теорема. *Если некоторая точка $t \in \mathbb{R}^n$ принадлежит множеству решений интервальной линейной системы $\mathbf{A}x = \mathbf{b}$, т. е. $t \in \Xi(\mathbf{A}, \mathbf{b})$, то*

$$\varrho = \min_{1 \le i \le m} \max_{A \in \mathbf{A}} \left\{ \frac{\mathrm{rad}\,\mathbf{b}_i - \left| \, \mathrm{mid}\,\mathbf{b}_i - \sum_{j=1}^{n} a_{ij} t_j \, \right|}{\sum_{j=1}^{n} |\, a_{ij}|} \right\} \;\ge\; 0 \qquad (9)$$

и интервальный вектор $\mathbf{U} = (t + \varrho\mathbf{e})$, $\mathbf{e} = ([-1, 1], \ldots, [-1, 1])^\top$, с центром в t целиком содержится во множестве решений $\Xi(\mathbf{A}, \mathbf{b})$.

Выражение под знаками экстремумов в (9) выглядит очень внушительно, но оно несет ясный содержательный смысл, о котором стоит упомянуть. Именно, вектор $|\,\mathrm{mid}\,\mathbf{b} - At\,|$ — это абсолютные величины отклонений произведения At от середины вектора правой части интервальной системы. Знаки разности между радиусами правой части и этими отклонениями, даваемые компонентами вектора $(\operatorname{rad}\mathbf{b} - |\operatorname{mid}\mathbf{b} - At\,|)$, указывают на принадлежность вектору $\mathbf{b}$ образа At точки t под действием линейного преобразования A. Фактически все это уже знакомо нам с предыдущего параграфа, где использовалось при выводе распознающего функционала Uni. Но, будучи отнесенными к сумме модулей элементов каждой строки матрицы A, компоненты этого вектора $(\operatorname{rad}\mathbf{b} - |\operatorname{mid}\mathbf{b} - At\,|)$ характеризуют уже нечто новое — степень чувствительности значений распознающего функционала к вариациям своего первого аргумента. Более точно, минимум этих отношений по i дает величину "грубости к возмущениям", показывающую, насколько можно сдвинуть точку t равномерно по всем координатам, чтобы она еще оставалась во множестве решений интервальной системы $Ax = \mathbf{b}$.

Доказательство. Коль скоро матрица ИСЛАУ не содержит нулевых строк,

$$\sum_{j=1}^{n} |\,a_{ij}| \;>\; 0,$$

и неотрицательность ϱ равносильна неотрицательности выражения

$$\min_{1\le i\le m}\ \max_{A\in\mathbf{A}} \left\{ \operatorname{rad}\mathbf{b}_i - \left|\ \operatorname{mid}\mathbf{b}_i - \sum_{j=1}^{n} a_{ij}t_j\ \right| \right\},$$

которое в силу теоремы, приведенной в § 3, определяет значение распознающего функционала Uni в точке t. Оно действительно неотрицательно при $t \in \Xi(\mathbf{A},\mathbf{b})$.

Приступая к обоснованию второго утверждения теоремы, предположим сначала, что в рассматриваемой нами задаче (3) матрица $\mathbf{A}$ имеет нулевую ширину, т. е. является просто вещественной, $\mathbf{A} = A = (\,a_{ij})$. Обозначая

$$\varrho_A = \min_{1\le i\le m} \left\{ \frac{\operatorname{rad}\mathbf{b}_i - \left|\ \operatorname{mid}\mathbf{b}_i - \sum_{j=1}^{n} a_{ij}t_j\ \right|}{\sum_{j=1}^{n} |\,a_{ij}|} \right\}, \tag{10}$$

представим каждый $x \in \mathbf{U}$ в виде $x = t + y$, где $\max_{1\le i\le n} |\,y_i| \le \varrho_A$.

Поскольку

$$|\,y_i| \;\le\; \varrho_A \;\le\; \frac{\operatorname{rad}\mathbf{b}_i - \left|\ \operatorname{mid}\mathbf{b}_i - \sum_{j=1}^{n} a_{ij}t_j\ \right|}{\sum_{j=1}^{n} |\,a_{ij}|},$$

для каждого $i = 1, 2, \ldots, m$ выполняется цепочка неравенств

$$\begin{aligned} |\,(Ay)_i| \;&=\; \left|\ \sum_{j=1}^{n} a_{ij}y_j \right| \;\le\; \sum_{j=1}^{n} |\,a_{ij}|\,|\,y_j| \;\le\; \varrho_A \cdot \sum_{j=1}^{n} |\,a_{ij}| \le \\ &\le\; \operatorname{rad}\mathbf{b}_i - \left|\ \operatorname{mid}\mathbf{b}_i - \sum_{j=1}^{n} a_{ij}t_j\ \right| = \\ &=\; \operatorname{rad}\mathbf{b}_i - |\operatorname{mid}\mathbf{b}_i - (At)_i\,|. \end{aligned}$$

Так как $Ay = Ax - At$, мы получаем

$$(At)_i - \operatorname{rad}\mathbf{b}_i + |\operatorname{mid}\mathbf{b}_i - (At)_i\,| \;\le\; (Ax)_i \;\le\; (At)_i + \operatorname{rad}\mathbf{b}_i - |\operatorname{mid}\mathbf{b}_i - (At)_i\,|$$

или, что равносильно,

$$\begin{aligned} \underline{\mathbf{b}}_i - (\,\mathrm{mid}\,\mathbf{b}_i - (At)_i) + |\,\mathrm{mid}\,\mathbf{b}_i - (At)_i\,| \le \\ \le (Ax)_i \le \\ \le \overline{\mathbf{b}}_i - (\,\mathrm{mid}\,\mathbf{b}_i - (At)_i) - |\,\mathrm{mid}\,\mathbf{b}_i - (At)_i\,|. \end{aligned} \tag{11}$$

Принимая во внимание тот факт, что

$$-z + |z| \ge 0 \quad \text{и} \quad -z - |z| \le 0$$

для любого вещественного z, неравенство (11) влечет для всех $i = 1, 2, \ldots, m$

$$\underline{\mathbf{b}}_i \le (Ax)_i \le \overline{\mathbf{b}}_i,$$

т. е. $Ax \in \mathbf{b}$. Это означает принадлежность точки x множеству решений интервальной линейной системы $Ax = \mathbf{b}$. Итак, формула (9) обоснована для систем (1), у которых интервальность присутствует только в правой части.

Предположим теперь, что в рассматриваемой ИСЛАУ матрица $\mathbf{A}$ является существенно интервальной матрицей, имеющей ненулевую ширину, множество решений $\Xi(\mathbf{A}, \mathbf{b})$ непусто и $t \in \Xi(\mathbf{A}, \mathbf{b})$. Рассмотрим совокупность всевозможных систем $Ax = \mathbf{b}$ с вещественными матрицами $A \in \mathbf{A}$ и внутренними оценками $\mathbf{U}_A$ их множеств решений $\Xi(A, \mathbf{b})$. В силу того что

$$\Xi(\mathbf{A}, \mathbf{b}) = \bigcup_{A \in \mathbf{A}} \Xi(A, \mathbf{b}),$$

объединение всех или некоторых внутренних оценок для множеств $\Xi(A, \mathbf{b})$ при $A \in \mathbf{A}$ также является внутренней оценкой для $\Xi(\mathbf{A}, \mathbf{b})$.

Пусть $\mathbf{U}_A$ — это кубики, имеющие фиксированный центр t. Ясно, что такие внутренние оценки существуют не для всех множеств решений $\Xi(A, \mathbf{b})$ с $A \in \mathbf{A}$, а лишь для тех, которые содержат точку t. Но объединение тех кубиков внутренних оценок $\mathbf{U}_A \subseteq \Xi(A, \mathbf{b})$, которые все-таки существуют для данного t, находится особенно просто: это тоже кубик с тем же центром t, размер которого есть максимум размеров объединяемых кубиков (рис. 2). В частности, если размеры кубиков определяются формулой (10), то брус

$$\mathbf{U} = t + \varrho \mathbf{e}$$

также целиком лежит во множестве решений $\Xi(\mathbf{A}, \mathbf{b})$ при

$$\varrho = \max_{A \in \mathbf{A}} \varrho_A = \max_{A \in \mathbf{A}} \min_{1 \le i \le m} \left\{ \frac{\mathrm{rad}\,\mathbf{b}_i - \left|\,\mathrm{mid}\,\mathbf{b}_i - \sum_{j=1}^n a_{ij} t_j\,\right|}{\sum_{j=1}^n |\,a_{ij}|} \right\}. \tag{12}$$

В этом выражении мы имеем право брать максимум относительно A по всей интервальной матрице $\mathbf{A}$ потому, что $\varrho_A < 0$ при $t \notin \Xi(\mathbf{A}, \mathbf{b})$ и это отрицательное значение никак не влияет на величину неотрицательного максимума (12).

Наконец, мы можем переставить в (12) операции взятия минимума и максимума, так как для разных индексов i максимумы от выражений в фигурных скобках берутся по

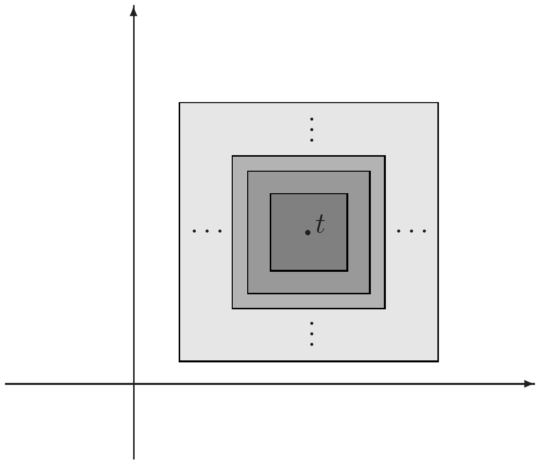

Рис. 2. Объединение кубиков с общим центром.

непересекающимся множествам аргументов, а именно по различным строкам матрицы **A**. Окончательно

$$\varrho \;=\; \min_{1\leq i\leq m}\;\max_{A\in\mathbf{A}}\;\left\{\;\frac{\operatorname{rad}\mathbf{b}_i - \left|\;\operatorname{mid}\mathbf{b}_i - \sum_{j=1}^{n} a_{ij}t_j\;\right|}{\sum_{j=1}^{n}|\,a_{ij}|}\;\right\}.$$

Этим и завершается доказательство теоремы. □

Нельзя не отметить красивую двойственность этого результата с формулой, выведенной в [1, 4] для размеров внутренней оценки уже упоминавшегося нами *допустимого множества решений* интервальной линейной системы (1), которое определяется как

$$\begin{aligned}\Xi_{tol}(\mathbf{A},\mathbf{b}) \;&=\; \{\,x\in\mathbb{R}^n \mid (\forall\,A\in\mathbf{A})(\exists\,b\in\mathbf{b})(Ax=b)\} =\\ &=\; \{\,x\in\mathbb{R}^n \mid (\forall\,A\in\mathbf{A})(Ax\in\mathbf{b})\} =\\ &=\; \{\,x\in\mathbb{R}^n \mid \mathbf{A}x\subseteq\mathbf{b}\,\}\end{aligned}$$

и имеет важные приложения в теории автоматического управления [17]. Оказывается, что если $t\in\Xi_{tol}(\mathbf{A},\mathbf{b})$, то

$$\varrho = \min_{1\leq i\leq m}\;\min_{A\in\mathbf{A}}\;\left\{\;\frac{\operatorname{rad}\mathbf{b}_i - \left|\;\operatorname{mid}\mathbf{b}_i - \displaystyle\sum_{j=1}^{n} a_{ij}t_j\;\right|}{\displaystyle\sum_{j=1}^{n}|\,a_{ij}|}\;\right\}\;\geq\;0 \tag{13}$$

и интервальный вектор $(t+\varrho\mathbf{e})$ включается в допустимое множество решений $\Xi_{tol}(\mathbf{A},\mathbf{b})$. Как видим, смена логического квантора при матрице в определении множества решений приводит к смене смысла внутреннего экстремума — вместо максимума по $A\in\mathbf{A}$ появляется минимум.

В выражении (9) взятие минимума по $i\in\{1,2,\dots,m\}$ не представляет трудностей, так что при отыскании ϱ центральной задачей является вычисление для каждого i внутренних максимумов

$$\max_{(a_{i1},\dots,a_{in})\in(\mathbf{a}_{i1},\dots,\mathbf{a}_{in})} \left\{ \frac{\operatorname{rad}\mathbf{b}_i - \left|\, \operatorname{mid}\mathbf{b}_i - \sum_{j=1}^n a_{ij}t_j \,\right|}{\sum_{j=1}^n |\, a_{ij}|} \right\}$$

или их оценивание снизу.

Для удобства дальнейших рассмотрений обозначим брус $(\mathbf{a}_{i1}, \mathbf{a}_{i2}, \dots, \mathbf{a}_{in})$ через

$$(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n) = \mathbf{X}$$

вне зависимости от индекса i, тогда как целевую функцию $\mathbb{R}^n \to \mathbb{R}$ — выражение внутри фигурных скобок в (9) и (13) — будем обозначать

$$\Phi(x) \;=\; \frac{R - \left|\, M - \sum\limits_{j=1}^n x_j t_j \,\right|}{\sum\limits_{j=1}^n |\, x_j|}, \tag{14}$$

где $R = \operatorname{rad}\mathbf{b}_i$, $M = \operatorname{mid}\mathbf{b}_i$ — вещественные константы. Таким образом, построение внутренней интервальной оценки множества решений ИСЛАУ вокруг известной точки-центра сводится к решению оптимизационной задачи

$$\boxed{\text{Найти } \max_{x\in\mathbf{X}} \Phi(x) \text{ или хотя бы оценить его снизу.}} \tag{15}$$

5. Вычисление размеров внутренней оценки

Ясно, что в (15) оценкой искомого $\max_{x\in\mathbf{X}} \Phi(x)$ снизу может служить значение целевой функции $\Phi(x)$ в любой точке бруса $\mathbf{X}$. Поэтому в случае, когда мы не хотим ввязываться в трудоемкие вычисления, простейшим способом решения задачи (15) является взятие максимального из значений целевой функции в нескольких выделенных точках области определения $\mathbf{X}$.

Обозначим

$$G(x) = R - \left|\, M - \sum_{j=1}^n x_j t_j \,\right|, \qquad H(x) = \sum_{j=1}^n |\, x_j|,$$

так что

$$\Phi(x) \;=\; \frac{G(x)}{H(x)}.$$

$G(x)$ и $H(x)$ представляют собой довольно простые выражения, имеющие лишь по одному вхождению каждой переменной x_j, а потому их экстремумы на $\mathbf{X}$ могут быть несложно вычислены — как левый или правый концы естественных интервальных расширений $G(\mathbf{X})$ и $H(\mathbf{X})$ соответствующих выражений. В частности,

$$\max_{x\in\mathbf{X}} G(x) \;=\; \overline{G(\mathbf{X})} \;=\; \overline{\left(R - \left\langle\, M - \sum_{j=1}^n \mathbf{X}_j t_j \,\right\rangle \right)}$$

и

$$\min_{x\in\mathbf{X}} H(x) \;=\; \underline{H(\mathbf{X})} \;=\; \sum_{j=1}^{n} \langle\, \mathbf{X}_j \rangle.$$

Далее, помимо самих значений этих экстремумов мы можем найти и значения аргументов, их доставляющие, отследив, какие из концов интервалов $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ в результате операций над ними — сложения, вычитания, умножения, взятия модуля и его антипода — дают концы естественных интервальных расширений $G(\mathbf{X})$ и $H(\mathbf{X})$. В итоге оценка решения задачи (15) может быть взята, к примеру, как максимум значений целевой функции $\Phi(x)$:

— в центре бруса $\mathbf{X}$;
— в точке минимума знаменателя дроби $H(x)$;
— в точке максимума числителя дроби $G(x)$.

Обратимся теперь к более развитым методам решения оптимизационной задачи (15). Напомним

Определение [18]. Пусть D — выпуклое множество в пространстве $\mathbb{R}^n$. *Функция* $f : D \to \mathbb{R}$ *называется квазивогнутой, если для любых* $x, y \in D$ *и* $0 \leq \lambda \leq 1$ *имеет место*

$$f(\,\lambda x + (1-\lambda)y\,) \geq \min\{\, f(x), f(y)\,\}.$$

Известно [18], что функция $f : D \to \mathbb{R}$ квазивогнута тогда и только тогда, когда для любого $\alpha \in \mathbb{R}$ ее лебеговы множества

$$\{\, x \in D \mid f(x) \geq \alpha \,\}$$

выпуклые (рис. 3). В частности, квазивогнутые функции не могут иметь несколько отличающихся по величине локальных максимумов, и, найдя один локальный максимум такой функции, мы можем быть уверенными в его глобальности.

Теорема. *Пусть брус* $\mathbf{X} \subseteq \mathbb{R}^n$ *не содержит начала координат. Множество* $\mathfrak{D}$ *всех точек из* $\mathbf{X}$*, на которых функция* $\Phi(x)$*, определяемая посредством (14), принимает неотрицательные значения, является выпуклым и* $\Phi(x)$ *квазивогнута на* $\mathfrak{D}$.

Доказательство. Для заданного фиксированного уровня $\alpha \geq 0$ обозначим через

$$S_\alpha = \{\, x \in \mathbf{X} \subset \mathbb{R}^n \mid \Phi(x) \geq \alpha \,\}$$

лебегово множество исследуемой функции $\Phi(x)$. В частности, $S_0 = \mathfrak{D}$.

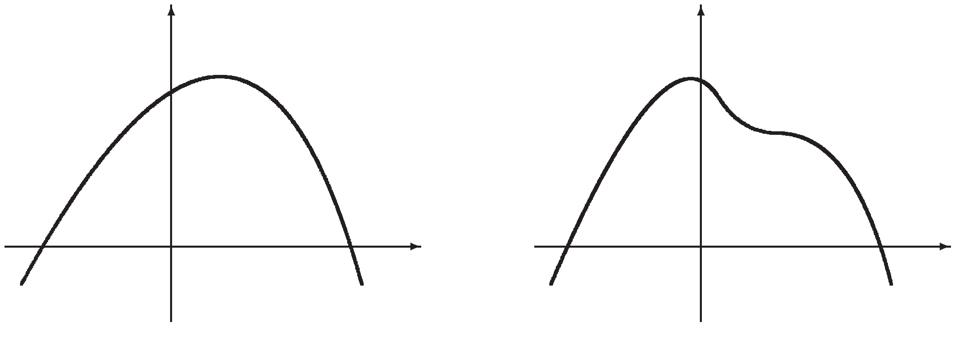

Рис. 3. Графики вогнутой и квазивогнутой функций.

Если S_α пусто, то рассуждать не о чем. Если же $S_\alpha \neq \emptyset$, то пусть точки x, y (не обязательно различные) принадлежат множеству S_α, так что $\Phi(x) \geq \alpha$, $\Phi(y) \geq \alpha$. При этом

$$R - \left|\; M - \textstyle\sum_{j=1}^n x_j\, t_j \;\right| \;\;\geq\;\; \alpha \sum_{j=1}^n |\, x_j|,$$

$$R - \left|\; M - \textstyle\sum_{j=1}^n y_j\, t_j \;\right| \;\;\geq\;\; \alpha \sum_{j=1}^n |\, y_j|.$$

Взяв какое-нибудь значение $\lambda \in [0, 1]$ и сложив эти неравенства с неотрицательными весами λ и $(1-\lambda)$, получим неравенство-следствие того же смысла:

$$R - \lambda \left|\; M - \sum_{j=1}^n x_j\, t_j \;\right| - (1-\lambda) \left|\; M - \sum_{j=1}^n y_j\, t_j \;\right| \;\geq\; \alpha \left(\lambda \sum_{j=1}^n |\, x_j| + (1-\lambda) \sum_{j=1}^n |\, y_j| \right). \quad (16)$$

Далее, применяя неравенство треугольника для абсолютного значения интервала, мы можем заменить левую часть неравенства (16) на бóльшую величину

$$R - \left|\; \lambda \left(M - \sum_{j=1}^n x_j\, t_j \right) + (1-\lambda) \left(M - \sum_{j=1}^n y_j\, t_j \right) \right|,$$

а правую часть (16) можем заменить (в силу $\alpha \geq 0$) на меньшую величину

$$\alpha \left(\sum_{j=1}^n |\, \lambda x_j + (1-\lambda)\, y_j| \right).$$

Окончательно имеем

$$R - \left|\; M - \sum_{j=1}^n (\lambda x_j + (1-\lambda)\, y_j)\, t_j \;\right| \;\geq\; \alpha \left(\sum_{j=1}^n |\, \lambda x_j + (1-\lambda)\, y_j| \right),$$

что равносильно

$$\Phi(\, \lambda x + (1-\lambda) y\,) \geq \alpha.$$

Следовательно, точка $\lambda x + (1-\lambda) y$ также лежит во множестве S_α, т. е. оно выпукло. Это и завершает доказательство теоремы. □

Отметим, что условие неотрицательности значений $\Phi(x)$ не является столь уж обременительным для применения доказанного результата, так как отрицательность $\Phi(x)$ при всех $x \in \mathbf{X}$ возможна лишь в том малоинтересном для нас случае, когда точка-центр t не лежит во множестве решений. Это следует из того, что отрицательность $\Phi(x)$ равносильна отрицательности числителя дроби (14) и, значит, “распознающего” функционала Uni в точке t (см. § 3). При этом следует озаботиться лучшим выбором центра t.

Присутствие в выражении (14) модулей делает целевую функцию $\Phi(x)$ негладкой, но она непрерывна и почти всюду дифференцируема на всей своей области определения. Таким образом, на факте квазивогнутости $\Phi(x)$ можно основать какие-нибудь градиентные методы для оценивания решения задачи (15), к примеру простейший метод

$$x^{(k+1)} := x^{(k)} + \gamma^{(k)}\, \nabla\Phi(\, x^{(k)}), \qquad k = 0, 1, 2, \dots, \quad (17)$$

с подходящим выбором величины шага $\gamma^{(k)} \in \mathbb{R}$ (см. [18]). При этом компоненты градиента $\nabla\Phi(x)$ имеют, как нетрудно проверить, следующий вид:

$$\Bigl(\nabla\Phi(x)\Bigr)_i \;=\; \frac{t_i\cdot\operatorname{sgn}\left(M-\sum_{j=1}^{n}x_j\,t_j\right)\left(\sum_{j=1}^{n}|\,x_j|\right)-\left(R-\left|\,M-\sum_{j=1}^{n}x_j\,t_j\,\right|\right)\operatorname{sgn}x_i}{\left(\sum_{j=1}^{n}|\,x_j|\right)^2},$$

где sgn — функция знака числа; $i=1,2,\dots,n$.

Естественно, что в качестве начального приближения $x^{(0)}$ для процесса (17) следует взять такое значение аргумента, на котором целевая функция неотрицательна. Как его найти?

Как следует из результатов § 3, принадлежность точки t множеству решений $\Xi(\mathbf{A},\mathbf{b})$ эквивалентна

$$\operatorname{Uni}(t,\mathbf{A},\mathbf{b}) = \min_{1\le i\le m}\left\{\operatorname{rad}\mathbf{b}_i-\left\langle\operatorname{mid}\mathbf{b}_i-\sum_{j=1}^{n}\mathbf{a}_{ij}\,t_j\right\rangle\right\}\;\ge\;0,$$

что, в свою очередь, равносильно справедливости такого же неравенства для отдельно рассматриваемой нами i-й строки матрицы $\mathbf{A}$:

$$R-\left\langle M-\sum_{j=1}^{n}\mathbf{X}_j\,t_j\right\rangle\;\ge\;0, \tag{18}$$

$R=\operatorname{rad}\mathbf{b}_i$, $M=\operatorname{mid}\mathbf{b}_i$. Поэтому для нахождения точки неотрицательности целевой функции Φ нам нужно, подобно тому как это было рекомендовано в начале параграфа, отследить, на каких именно концах интервалов $\mathbf{X}_1$, $\mathbf{X}_2$, …, $\mathbf{X}_n$ достигается значение выражения в левой части (18). Найденные числа и образуют искомый вектор начального приближения $x^{(0)}$ для градиентного подъема (17).

Список литературы

[1] Добронец Б. С., Шайдуров В. В. Двусторонние численные методы. Новосибирск: Наука, 1990.

[2] Лакеев А. В., Носков С. И. Описание множества решений линейного уравнения с интервально заданными оператором и правой частью // Докл. РАН. 1993. Т. 330, № 4. С. 430–433.

[3] Шарый С. П. Алгебраический подход к анализу линейных статических систем с интервальной неопределенностью // Изв. РАН. Теория и системы управления. 1997. № 3. С. 51–61.

[4] Shary S. P. Solving the linear interval tolerance problem // Mathematics and Computers in Simulation. 1995. Vol. 39. P. 53–85.

[5] Shary S. P. Algebraic approach to the interval linear static identification, tolerance and control problems, or One more application of Kaucher arithmetic // Reliable Comp. 1996. Vol. 2, No. 1. P. 3–33.

[6] Shary S. P. A new technique in systems analysis under interval uncertainty and ambiguity // Reliable Comp. 2002. Vol. 8, No. 5. P. 321–418.

[7] Kupriyanova L. Inner estimation of the united solution set of interval linear algebraic system // Reliable Comp. 1995. Vol. 1, No. 1. P. 15–31.

[8] Kearfott R. B., Nakao M. T., Neumaier A. et al. Standardized notation in interval analysis // Reliable Comp. (В печати) (`http://www.mat.univie.ac.at/~neum/software/int`)

[9] Алефельд Г., Херцбергер Ю. Введение в интервальные вычисления. М.: Мир, 1987.

[10] Калмыков С. А., Шокин Ю. И., Юлдашев З. Х. Методы интервального анализа. Новосибирск: Наука, 1986.

[11] Neumaier A. Interval Methods for Systems of Equations. Cambridge: Cambridge Univ. Press, 1990.

[12] Kreinovich V., Lakeyev A., Rohn J., Kahl P. Computational Complexity and Feasibility of Data Processing and Interval Computations. Dordrecht: Kluwer, 1997.

[13] Шарый С. П. Оптимальное внешнее оценивание множеств решений интервальных систем уравнений. Ч. 1 // Вычисл. технологии. 2002. Т. 7, № 6. С. 90–113.

[14] Oettli W. On the solution set of a linear system with inaccurate coefficients // SIAM J. Numer. Analysis. 1965. Vol. 2, No. 1. P. 115–118.

[15] Rex G., Rohn J. Sufficient conditions for regularity and singularity of interval matrices // SIAM J. Matrix Analysis and Appls. 1999. Vol. 20. P. 437–445.

[16] Шарый С. П. О характеризации объединенного множества решений интервальной линейной алгебраической системы. Красноярск, 1990. Деп. в ВИНИТИ. № 726-В91. 20 с.

[17] Дугарова И. В., Смагина Е. М. Обеспечение устойчивости системы с неопределенными параметрами // Автоматика и телемеханика. 1990. № 11. С. 176–181.

[18] Базара М., Шетти К. Нелинейное программирование. Теория и алгоритмы. М.: Мир, 1982.

Поступила в редакцию 9 апреля 2003 г.