УДК 519.6

Численное нахождение алгебраического решения интервальных линейных систем

С. П. Шарый

Вычислительный Центр СО РАН,
Академгородок, 660036 Красноярск, Россия

Аннотация

Работа посвящена численному нахождению *алгебраического решения* интервальных линейных систем в расширенной интервальной арифметике Каухера. Необходимость рассмотрения подобных постановок естественно возникает в минимаксных задачах исследования операций и теории принятия решений. Предлагаются два подхода к построению стационарных итерационных методов для нахождения алгебраического решения, основанных на расщеплении матрицы системы, обсуждаются вопросы реализации и сходимости этих методов.

1 Введение

Основным объектом изучения в настоящей работе являются интервальные системы линейных алгебраических уравнений (ИСЛАУ) в расширенной интервальной арифметике Каухера $\mathbb{IR}$, т.е. системы уравнений вида

$$\left\{ \begin{array}{ccccccccc} \mathbf{c}_{11}\mathbf{x}_1 & + & \mathbf{c}_{12}\mathbf{x}_2 & + & \ldots & + & \mathbf{c}_{1n}\mathbf{x}_n & = & \mathbf{d}_1, \\ \mathbf{c}_{21}\mathbf{x}_1 & + & \mathbf{c}_{22}\mathbf{x}_2 & + & \ldots & + & \mathbf{c}_{2n}\mathbf{x}_n & = & \mathbf{d}_2, \\ \vdots & & & & \ddots & & & & \vdots \\ \mathbf{c}_{m1}\mathbf{x}_1 & + & \mathbf{c}_{m2}\mathbf{x}_2 & + & \ldots & + & \mathbf{c}_{mn}\mathbf{x}_n & = & \mathbf{d}_m, \end{array} \right. \tag{1}$$

с $\mathbf{c}_{ij}, \mathbf{x}_i, \mathbf{d}_i \in \mathbb{IR}$. Удобно для краткости обозначать такие системы уравнений через

$$\mathbf{C}\mathbf{x} = \mathbf{d}, \tag{2}$$

где $\mathbf{C}$ — $m \times n$-матрица и $\mathbf{d}$ — m-вектор с элементами из $\mathbb{IR}$.

Работа выполнена при финансовой поддержке Красноярского краевого фонда науки.

Определение 1. Интервальный вектор называется *алгебраическим решением* интервальной системы уравнений, если его подстановка в эту систему и выполнение всех интервальных арифметических операций приводят к верному равенству.

Развитие алгоритмов для отыскания алгебраических решений интервальных систем уравнений вида (1)–(2) — главная задача нашей работы. При этом в целях упрощения изложения мы ограничимся только случаем квадратных систем с $m = n$.

Необходимость рассмотрения подобных постановок возникает, в частности, при моделировании линейных статических систем в условиях интервальной неопределённости их параметров, как следствие *алгебраического подхода* к их анализу [9, 15], предусматривающего замену исходной задачи оценки внутренних состояний системы на задачу нахождения алгебраических решений вспомогательного уравнения в арифметике Каухера $\mathbb{IR}$. В этой работе мы не останавливаемся на определении интервальной арифметики Каухера и обсуждении её многочисленных замечательных свойств. Заинтересованному читателю следует обратиться либо к оригинальным работам [11, 10], либо к конспективному изложению в [9, 15, 4, 12].

Всюду ниже мы будем выделять интервалы и интервальные величины (векторы, матрицы) жирным шрифтом (например, $\mathbf{A}, \mathbf{B}, \mathbf{C}, \ldots, \mathbf{x}, \mathbf{y}, \mathbf{z}$), под $\underline{\mathbf{x}}$, $\overline{\mathbf{x}}$ будут пониматься операции взятия левого и правого концов интервала, прочие обозначения также следуют работам [9, 10, 14, 15]. В оставшейся части введения мы напомним основные теоретические факты, касающиеся интервальных линейных систем в арифметике Каухера $\mathbb{IR}$ [15, 9].

Несмотря на простую структуру системы уравнений (2), мы лишь в очень частных ситуациях можем воспользоваться для её решения какими-либо методами исключения, пакетами символьных преобразований и т.п. Причина — недостаточные алгебраические свойства $\mathbb{IR}$. Отсутствие в арифметике Каухера полноценной дистрибутивности делает в общем случае невозможной даже такую простейшую операцию как приведение подобных членов. Именно по этой причине рассматриваемые ниже методы являются существенно *численными*. Отметим, что недавно А.В. Лакееву удалось показать, что задача нахождения алгебраического решения интервальной системы линейных уравнений в арифметике Каухера является NP-трудной [13].

Вычислительная математика накопила богатый арсенал как теоретических подходов, так и практических алгоритмов, но подавляющее большинство из них относится к операторным уравнениям в линейных пространствах.

Формально эти методы неприменимы к задаче вычисления алгебраического решения системы (2), поскольку $\mathbb{IR}^n$ не является линейным пространством. В действительности мы можем легко обойти это затруднение, осуществив вложение пространства $\mathbb{IR}^n$ в обычное хорошо изученное евклидово пространство $\mathbb{R}^{2n}$.

Нетрудно видеть, что если U — какое-нибудь линейное пространство, то любая биекция $\iota : \mathbb{IR}^n \to U$ порождает биекцию

$$\iota^{\natural} : (\mathbb{IR}^n)^{\mathbb{IR}^n} \to U^U$$

из множества всех отображений на $\mathbb{IR}^n$ в множество всех отображений на U, при которой каждому оператору $\psi : \mathbb{IR}^n \to \mathbb{IR}^n$ сопоставляется единственный оператор $\psi^{\iota} : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$, называемый *индуцированным оператором* для ψ, такой что

$$\psi^{\iota} = \iota \circ \psi \circ \iota^{-1}$$

($\circ$ — знак композиции). Посредством погружения исследование отображений $\mathbb{IR}^n \to \mathbb{IR}^n$ сводится к исследованию отображений $U \to U$ линейного пространства, т.е. мы приходим к ситуации, более привычной для современных вычислителей. В частности, если обозначить $\ominus$ — внутреннее вычитание в $\mathbb{IR}$ (прибавление противоположного элемента), то стоящая перед нами исходная задача отыскания нулей отображения $\phi : \mathbb{IR}^n \to \mathbb{IR}^n$, такого что

$$\phi(\mathbf{x}) = \mathbf{C}\mathbf{x} \ominus \mathbf{d},$$

заменяется на задачу решения уравнения

$$\Phi(x) = 0 \tag{3}$$

в $\mathbb{R}^{2n}$, такого что $\Phi = \iota \circ \phi \circ \iota^{-1} : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$, т.е.

$$\Phi(x) = \iota\big(\mathbf{C} \cdot \iota^{-1}(x) \ominus \mathbf{d}\big) = \iota\big(\mathbf{C} \cdot \iota^{-1}(x)\big) - \iota(\mathbf{d}). \tag{4}$$

Основной вопрос, касающийся конструирования вложения ι, — как обеспечить разумный компромисс между его простотой и удобной формой индуцированных отображений ψ^{ι}. Мы рекомендуем для практической реализации погружение $\sigma : \mathbb{IR}^n \to \mathbb{R}^{2n}$, которое действует следующим образом

$$(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n) \mapsto (-\underline{\mathbf{x}}_1, -\underline{\mathbf{x}}_2, \ldots, -\underline{\mathbf{x}}_n, \overline{\mathbf{x}}_1, \overline{\mathbf{x}}_2, \ldots, \overline{\mathbf{x}}_n), \tag{5}$$

т.е. когда 1-ая, 2-ая, ..., n-ая компоненты вектора $\sigma(\mathbf{x})$ полагаются равными левым концам $\mathbf{x}_1$, $\mathbf{x}_2$, ..., $\mathbf{x}_n$, взятым с противоположными знаками, а $(n+1)$-ая, ..., $2n$-ая компоненты $\sigma(\mathbf{x})$ полагаются равными правым концам $\mathbf{x}_1$, $\mathbf{x}_2$, ..., $\mathbf{x}_n$ соответственно.

Определение 2. Назовём отображение $\sigma : \mathbb{IR}^n \to \mathbb{R}^{2n}$, действующее по правилу (5), *стандартным погружением* интервального пространства $\mathbb{IR}^n$ в евклидово пространство $\mathbb{R}^{2n}$.

В результате погружения естественный порядок по включению на $\mathbb{IR}^n$ переносится и в $\mathbb{R}^{2n}$: для элементов $x, y \in \mathbb{R}^{2n}$ мы будем считать, что x "не превосходит" y, если $\iota^{-1}(x) \subseteq \iota^{-1}(y)$ в $\mathbb{IR}^n$. Непосредственной проверкой можно убедиться, что в случае стандартного погружения σ такой порядок совпадает с обычным покомпонентным упорядочением $\mathbb{R}^{2n}$, задаваемым конусом векторов с неотрицательными компонентами, когда

$$x \text{ не превосходит } y \quad \Longleftrightarrow \quad x_i \leq y_i,\ i = 1, 2, \ldots, 2n,$$

т.е. выглядит особенно просто.

Если $\psi : \mathbb{IR}^n \to \mathbb{IR}^n$ — оператор умножения на точечную матрицу $W = (\, w_{ij}) \in \mathbb{R}^{n\times n}$, т.е. $\psi(\mathbf{x}) = W\mathbf{x}$, то, как нетрудно показать, индуцированный оператор $\psi^\iota = \iota \circ \psi \circ \iota^{-1}$ является линейным преобразованием пространства $\mathbb{R}^{2n}$. Для стандартного погружения σ соответствующая $2n \times 2n$-матрица линейного преобразования $\psi^\sigma : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ имеет следующий блочный вид

$$W^\sigma \stackrel{\text{def}}{=} \left(\begin{array}{c|c} W^+ & W^- \\ \hline W^- & W^+ \end{array} \right), \tag{6}$$

где матрицы $W^+ = (\, w^+_{ij})$ и $W^- = (\, w^-_{ij})$ — это отрицательная и положительная части W, т.е.

$$w^+_{ij} = \max\{\, w_{ij}, 0\,\} \qquad \text{и} \qquad w^-_{ij} = \max\{-w_{ij}, 0\,\}.$$

Изотонные по включению линейные операторы на $\mathbb{IR}^n$ и только они соответствуют при стандартном погружении (5) неотрицательным точечным $2n\times 2n$-матрицам.

При каких условиях оператор умножения на точечную матрицу в $\mathbb{IR}^n$ обратим? Поскольку он является линейным, то для этого, как известно, необходимо и достаточно, чтобы

$$\psi(z) = 0 \quad \Longleftrightarrow \quad z = 0, \tag{7}$$

т.е. чтобы его ядро было тривиальным. Важно отметить, что невырожденность точечной матрицы $W \in \mathbb{R}^{n\times n}$ в смысле классической линейной алгебры не обязательно влечёт то, что соответствующий оператор умножения на W в $\mathbb{IR}^n$ также невырожден в смысле (7). Например, матрица

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \tag{8}$$

имеет ненулевой определитель, но

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} [-1,1] \\ [1,-1] \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Чтобы различать подобные случаи, вводится

Определение 3. Назовём матрицу $W \in \mathbb{R}^{n\times n}$ *ι-невырожденной*, если умножение на эту матрицу удовлетворяет условиям невырожденности (7) в $\mathbb{IR}^n$, т.е. если

$$W\mathbf{x} = 0 \quad \Longleftrightarrow \quad \mathbf{x} = 0 \in \mathbb{IR}^n.$$

Иначе будем говорить, что матрица W является *ι-вырожденной*.

Оказывается, что точечная матрица $W \in \mathbb{R}^{n\times n}$ ι-невырождена тогда и только тогда, когда матрица $W^\sigma \in \mathbb{R}^{2n\times 2n}$ вида (6) невырождена в обычном смысле, т.е. её определитель не равен нулю.

Например, единичная матрица является ι-невырожденной, а матрица (8) — ι-вырожденная. Все неотрицательные невырожденные матрицы также ι-невырождены. Нетрудно показать, что ι-невырожденные точечные матрицы образуют открытое всюду плотное подмножество во множестве всех квадратных матриц данного размера, т.е. их множество достаточно представительно. В целом важность понятия ι-невырожденной матрицы определяется тем, что им выделяется широкий класс простейших обратимых операторов на $\mathbb{IR}^n$, отталкиваясь от которых, можно изучать обратимость и более сложных (составных) операторов.

2 Субдифференциальный метод Ньютона

В этом разделе работы мы кратко излагаем результаты о *субдифференциальном методе Ньютона* для нахождения алгебраических решений интервальных линейных систем.

Напомним некоторые определения и факты [1]. Пусть U, V — вещественные линейные пространства, причём V — упорядоченное линейное пространство с порядком $\preceq$. Отображение $F : U \to V$ называется *(порядково) выпуклым*, если

$$F(\lambda x + (1-\lambda)y) \preceq \lambda F(x) + (1-\lambda)F(y)$$

для любых x, $y \in U$ и $\lambda \in (0,1)$. *Субдифференциалом* выпуклого отображения $F : U \to V$ в точке x называется множество $\partial_{\preceq} F(x)$ всех линейных операторов $D : U \to V$ таких, что

$$D(y-x) \preceq F(y) - F(x)$$

для любых $y \in U$. Известно, что для каждой внутренней точки x эффективной области определения выпуклой функции F субдифференциал является непустым множеством, элементы которого называют *субградиентами* функции F в точке x.

Предположим, что интервальная матрица $\mathbf{C}$ такова, что в каждой её строке все элементы либо только правильные, либо только неправильные, а множества натуральных чисел $I' = \{\, i'_1, i'_2, \dots, i'_\alpha \,\}$ и $I'' = \{\, i''_1, i''_2, \dots, i''_\beta \,\}$, $I' \cap I'' = \emptyset$, $\alpha + \beta = n$, представляют номера строк $\mathbf{C}$ с правильными и неправильными интервалами соответственно:

$$\mathbf{c}_{ij} \text{ является } \begin{cases} \text{правильным,} & \text{если } i \in I' \\ \text{неправильным,} & \text{если } i \in I''. \end{cases}$$

Если задать на $\mathbb{R}^{2n}$ частичный порядок “$\sqsubseteq$” следующим образом

$$x \sqsubseteq y \iff \begin{cases} x_i \le y_i & \text{для } i \in \{\, i'_1, \dots, i'_\alpha, 2i'_1, \dots, 2i'_\alpha \,\}, \\ x_i \ge y_i & \text{для } i \in \{\, i''_1, \dots, i''_\beta, 2i''_1, \dots, 2i''_\beta \,\}, \end{cases}$$

то, как следствие свойств суб- и супердистрибутивности в арифметике Каухера, исследуемое отображение $\Phi(x)$, задаваемое посредством (4), оказывается выпуклым относительно “$\sqsubseteq$”. Соответственно, тогда в любой точке $x \in \mathbb{R}^{2n}$ будет существовать непустой субдифференциал $\partial_{\sqsubseteq}\Phi(x)$, легко вычислимый, поскольку $\Phi(x)$ — полиэдральное (кусочно-аффинное) отображение. Эффективным средством для нахождения алгебраических решений таких ИСЛАУ зарекомендовал себя *субдифференциальный метод Ньютона.* Он является дальнейшим развитием известных результатов о монотонно сходящихся методах ньютоновского типа в упорядоченных линейных пространствах.

Субдифференциальный метод Ньютона
со специальным начальным приближением

В качестве начального приближения $x^{(0)}$ возьмём решение "средней" системы

$$(\mathrm{mid}\,\mathbf{C})^{\sigma}\,\mathrm{x} = \sigma(\mathbf{d}).$$

Если k-е приближение $x^{(k)}$, $k = 0, 1, \ldots$, уже известно, то находим какой-нибудь субградиент $D^{(k)} \in \partial_{\sqsubseteq}\Phi(\,x^{(k)})$ и затем полагаем

$$x^{(k+1)} := x^{(k)} - \tau\left(\,D^{(k)}\right)^{-1}\left(\Phi(\,x^{(k)})\right).$$

Здесь τ — некоторый релаксационный параметр из $(0, 1]$. Справедлива

Теорема 1. Если в каждой строке интервальной матрицы $\mathbf{C}$ все элементы только правильные или только неправильные, а pro $\mathbf{C}$ содержит только $\imath$-невырожденные точечные матрицы и достаточно узка (т.е. $\|\mathrm{rad}\ \mathbf{C}\|$ достаточно мала), то субдифференциальный метод Ньютона сходится к $\sigma(\mathbf{x}_a)$, где $\mathbf{x}_a$ — алгебраическое интервальное решение системы $\mathbf{Cx} = \mathbf{d}$.

В применении к более частным типам интервальных систем доказательство сходимости субдифференциального метода Ньютона и его обстоятельное обсуждение содержатся в работе [15]. На практике субдифференциальный метод Ньютона сходится исключительно быстро, за небольшое конечное число итераций (особенно для τ близких или равных единице), что связано с полиэдральностью функции $\Phi(x)$. Другое достоинство этого метода — отсутствие проблем с выбором начального приближения. В целом, субдифференциальный метод Ньютона имеет огромное практическое значение, но в этой работе мы будем рассматривать его, главным образом, как вспомогательный алгоритм, начальное звено в составе численных процессов с более широкой областью сходимости.

3 Методы на основе теоремы об обобщённых сжатиях

Как находить алгебраические интервальные решения ИСЛАУ с матрицами, имеющими произвольное распределение правильных и неправильных элементов? В общем случае для этой цели применима, например, универсальная схема метода простой итерации со всеми её многочисленными модификациями — Зейделя, Якоби и т.п. [8]. При реализации подобных методов иногда можно итерировать непосредственно в $\mathbb{IR}^n$, даже не погружая его в линейное пространство. С другой стороны, получаемая в подобных алгоритмах сходимость является качественно более медленной, чем в субдифференциальном методе Ньютона.

В соответствии с общей схемой одношаговых стационарных итерационных методов приведём исходное уравнение (2) к виду

$$\mathbf{x} = T(\mathbf{x}), \tag{9}$$

$T : \mathbb{IR}^n \to \mathbb{IR}^n$, а затем, выбрав начальное приближение $\mathbf{x}^{(0)}$, станем итерировать

$$\mathbf{x}^{(k+1)} = T(\,\mathbf{x}^{(k)}). \tag{10}$$

При некоторых специальных условиях на оператор перехода T (когда он является сжатием и т.п.) и на начальное приближение $\mathbf{x}^{(0)}$ последовательность $\mathbf{x}^{(k)}$ сходится к его неподвижной точке, т.е. к искомому решению уравнения (2). Но, в отличие от вещественного случая, из-за слабых алгебраических свойств арифметики Каухера не вполне тривиальной задачей является приведение ИСЛАУ (2) к виду (9). Проблема состоит в том, что по меньшей мере *два* члена с переменной $\mathbf{x}$ в формуле (9) (которая эквивалентна $\mathbf{x} \ominus T(\mathbf{x}) = 0$) должны в итоге свернуться в выражение $\mathbf{Cx} \ominus \mathbf{d}$, содержащее лишь *одно* вхождение переменной $\mathbf{x}$. А это, при отсутствии полноценной возможности приводить подобные члены, требует специальных средств для преобразования исходного уравнения к виду (9).

Один из возможных подходов к конструированию итерационных схем для решения уравнения (2) заключается в том, чтобы пойти обратным путём — от возможных расщеплений интервальной матрицы $\mathbf{C}$, т.е. от представлений $\mathbf{C} = \mathbf{G} + \mathbf{H}$, где $\mathbf{G}$, $\mathbf{H} \in \mathbb{IR}^{n\times n}$, удовлетворяют условию

$$\mathbf{Cx} = \mathbf{Gx} + \mathbf{Hx}$$

при любых $\mathbf{x} \in \mathbb{IR}^n$. Мы будем называть такие представления *дистрибутивными* расщеплениями матрицы $\mathbf{C}$. Если для оператора умножения на интервальную матрицу $\mathbf{G}$ легко может быть построен обратный оператор $\Xi : \mathbb{IR}^n \to \mathbb{IR}^n$, такой что

$$\Xi\,(\mathbf{Gx}) = \mathbf{x} \quad \text{при всех } \mathbf{x} \in \mathbb{IR}^n,$$

то из уравнения

$$\mathbf{Gx} + \mathbf{Hx} \ominus \mathbf{d} = 0,$$

эквивалентного (2), получим

$$\mathbf{x} = \Xi\,(\mathbf{d} \ominus \mathbf{H}\mathbf{x}).$$

Соответственно, итерационный процесс можно организовывать по формуле

$$\mathbf{x}^{(k+1)} = \Xi\,(\,\mathbf{d} \ominus \mathbf{H}\mathbf{x}^{(k)}). \tag{11}$$

Предположим, что интервальная матрица $\mathbf{C}$ невырождена в том смысле, что невырождены все вещественные матрицы из её правильной проекции $\mathrm{pro}\,\mathbf{C}$, т.е. матрицы составленной из элементов

$$\mathrm{pro}\,\mathbf{C} = \begin{cases} \mathbf{c}_{ij}, & \text{если } \mathbf{c}_{ij} \text{ правильный}, \\ \mathrm{dual}\,\mathbf{c}_{ij}, & \text{если } \mathbf{c}_{ij} \text{ неправильный}. \end{cases}$$

Тогда существуют по крайней мере две возможности для дистрибутивного расщепления $\mathbf{C}$ с легко обратимым членом $\mathbf{G}\mathbf{x}$:

1. $\mathbf{G}$ — вещественная матрица, которую поэтому можно обозначать просто через G, $\mathbf{H} = \mathbf{C} - G$, а ненулевые элементы G и $\mathbf{H}$ имеют одинаковые знаки (за счёт чего и обеспечивается дистрибутивность).

   Последнему условию можно удовлетворить, например, если $G = \lfloor \mathbf{C} \rfloor$, т.е. если G образована поэлементным применением к $\mathbf{C}$ унарной операции $\lfloor\cdot\rfloor$, такой что

   $$\lfloor \mathbf{x} \rfloor = \begin{cases} \max\{\,\underline{\mathbf{x}},\, \overline{\mathbf{x}}\,\}, & \text{если } \mathbf{x} < 0, \\ 0, & \text{если } 0 \in \mathrm{pro}\,\mathbf{x}, \\ \min\{\,\underline{\mathbf{x}},\, \overline{\mathbf{x}}\,\}, & \text{если } \mathbf{x} > 0. \end{cases}$$

   (взятие ближайшего к нулю элемента правильной проекции интервала). Ясно, что при сделанном нами предположении матрица G невырождена. Если она к тому же $\imath$-невырождена, то обратный оператор Ξ соответствует умножению на матрицу $(G^{\sigma})^{-1}$ в $\mathbb{R}^{2n}$. В любом случае у нас всегда есть возможность сделать матрицу G $\imath$-невырожденной путём небольшого уменьшения абсолютной величины её ненулевых элементов, не нарушающего субдистрибутивности расщепления матрицы $\mathbf{C}$.

2. $\mathbf{G}$ и $\mathbf{H}$ — это верхняя и нижняя треугольные интервальные матрицы соответственно, причём $\mathbf{G}$ имеет обратимые элементы на главной диагонали, а у $\mathbf{H}$ главная диагональ нулевая (возможно, для этого сначала потребуется поменять местами уравнения системы).

Обратный оператор Ξ при этом таков, что результат $\mathbf{y}$ его действия на элемент $\mathbf{x} \in \mathbb{IR}^n$ определяется по формулам обратного хода метода Гаусса для системы $\mathbf{G}\mathbf{y} = \mathbf{x}$. Естественно называть такое дистрибутивное расщепление $\mathbf{C}$ *треугольным.*

В обоих рассмотренных случаях обратный оператор $\Xi : \mathbb{IR}^n \to \mathbb{IR}^n$, вообще говоря, не может быть задан умножением на интервальную матрицу.

Выпишем формулы соответствующих итерационных процессов (11) в явном виде. В первом из рекомендованных рецептов дистрибутивного расщепления $\mathbf{C}$ нам будет удобно перейти в пространство $\mathbb{R}^{2n}$. Имеем

$$x^{(k+1)} = (G^{\sigma})^{-1}\sigma\big(\,\mathbf{d} \ominus \mathbf{H}\sigma^{-1}(\,x^{(k)})\big), \tag{12}$$

где $G = \lfloor \mathbf{C} \rfloor$, $\mathbf{H} = \mathbf{C} - G$.

Рассчётные формулы итерационного метода (11) с треугольным дистрибутивным расщеплением матрицы $\mathbf{C}$ имеют в $\mathbb{IR}^n$ вид:

$$\mathbf{p}_1 = \mathbf{d}_1, \qquad \mathbf{p}_i = \mathbf{d}_i \ominus \sum_{j=1}^{i-1} \mathbf{h}_{ij}\mathbf{x}_j^{(k)}, \qquad i = 2, \ldots, n;$$

$$\mathbf{x}_n^{(k+1)} = \mathbf{p}_n \oslash \mathbf{g}_{nn},$$

$$\mathbf{x}_i^{(k+1)} = \left(\,\mathbf{p}_i \ominus \sum_{j=i+1}^{n} \mathbf{g}_{ij}\mathbf{x}_j^{(k+1)}\right) \oslash \mathbf{g}_{ii}, \qquad i = n-1, \ldots, 1,$$

где $\oslash$ — внутреннее деление в $\mathbb{IR}$, т.е. умножение на обратный интервал (если он существует).

Имеет место

Теорема 2. Пусть $(G^{\sigma})^{-1} = \mathfrak{G} = (\mathfrak{g}_{ij})$, а Γ – это вещественная $n\times n$-матрица, образованная элементами

$$\gamma_{ij} = \max\{\,|\mathfrak{g}_{ij}|, |\mathfrak{g}_{i,j+n}|, |\mathfrak{g}_{i+n,j}|, |\mathfrak{g}_{i+n,j+n}|\,\}, \qquad i,j = 1,2,\ldots,n.$$

Тогда, если спектральный радиус $n\times n$-матрицы $\Gamma\,|\mathbf{H}|$ меньше единицы, то итерации (12) сходятся к алгебраическому интервальному решению системы $\mathbf{C}x = \mathbf{d}$ из любого начального приближения $\mathbf{x}^{(0)}$.

Доказательство. Введем на $\mathbb{R}^{2n}$ псевдометрику $\varrho : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ следующим образом

$$\varrho(x,y) = \begin{pmatrix} \max\{\,|x_1-y_1|, |x_{n+1}-y_{n+1}|\,\} \\ \vdots \\ \max\{\,|x_n-y_n|, |x_{2n}-y_{2n}|\,\} \\ \max\{\,|x_1-y_1|, |x_{n+1}-y_{n+1}|\,\} \\ \vdots \\ \max\{\,|x_n-y_n|, |x_{2n}-y_{2n}|\,\} \end{pmatrix} = \begin{pmatrix} |\,\sigma^{-1}(x)\ominus\sigma^{-1}(y)\,| \\ |\,\sigma^{-1}(x)\ominus\sigma^{-1}(y)\,| \end{pmatrix}.$$

Покажем, что относительно такой псевдометрики оператор перехода итерационной схемы (12)

$$T(x) = \mathfrak{G}\,\sigma\bigl(\,\mathbf{d}\ominus\mathbf{H}\sigma^{-1}(x)\bigr)$$

удовлетворяет условиям теоремы Шрёдера об обобщённых сжатиях [5, 8].

Имеем

$$\begin{aligned} |\,T_i(x)-T_i(y)| &= |\,T(x)-T(y)|_i \\ &= \bigl|\,\mathfrak{G}\,\sigma(\,\mathbf{d}\ominus\mathbf{H}\sigma^{-1}(x)) - \mathfrak{G}\,\sigma(\,\mathbf{d}\ominus\mathbf{H}\sigma^{-1}(y))\bigr|_i \\ &= \bigl|\,\mathfrak{G}\,\bigl(\sigma(\,\mathbf{d}\ominus\mathbf{H}\sigma^{-1}(x)) - \sigma(\,\mathbf{d}\ominus\mathbf{H}\sigma^{-1}(x))\bigr)\,\bigr|_i \\ &= \bigl|\,\mathfrak{G}\,\sigma(\,\mathbf{H}\sigma^{-1}(x)\ominus\mathbf{H}\sigma^{-1}(y))\,\bigr|_i \\ &\le \bigl(\,|\mathfrak{G}|\;|\sigma(\,\mathbf{H}\sigma^{-1}(x)\ominus\mathbf{H}\sigma^{-1}(y))|\,\bigr)_i \\ &= \left(\,|\mathfrak{G}|\cdot\begin{pmatrix} |\,\underline{\mathbf{H}\sigma^{-1}(x)\ominus\mathbf{H}\sigma^{-1}(y)}\,| \\ |\,\overline{\mathbf{H}\sigma^{-1}(x)\ominus\mathbf{H}\sigma^{-1}(y)}\,| \end{pmatrix}\right)_i. \end{aligned}$$

Отметим, что для любых $\mathbf{u},\mathbf{v}\in\mathbb{IR}^n$

$$\max\bigl\{\,\underline{\mathbf{Hu}\ominus\mathbf{Hv}},\ \overline{\mathbf{Hu}\ominus\mathbf{Hv}}\,\bigr\} = |\,\mathbf{Hu}\ominus\mathbf{Hv}| \le |\mathbf{H}|\cdot|\mathbf{u}\ominus\mathbf{v}|.$$

Поэтому, продолжая выкладки, получим:

$$\begin{aligned}
|\,T_i(x)-T_i(y)| \;&\le\; \left(|\mathfrak{G}|\cdot\begin{pmatrix} |\mathbf{H}|\,|\sigma^{-1}(x)\ominus\sigma^{-1}(y)\,| \\ |\mathbf{H}|\,|\sigma^{-1}(x)\ominus\sigma^{-1}(y)\,| \end{pmatrix}\right)_i \\
&=\; \left(|\mathfrak{G}|\,|\mathbf{H}|^{\sigma}\begin{pmatrix} |\sigma^{-1}(x)\ominus\sigma^{-1}(y)\,| \\ |\sigma^{-1}(x)\ominus\sigma^{-1}(y)\,| \end{pmatrix}\right)_i \\
&=\; \big(\, |\mathfrak{G}|\,|\mathbf{H}|^{\sigma}\varrho(x,y)\,\big)_i \\
&=\; \big(\, i\text{-ая строка матрицы } |\mathfrak{G}|\,|\mathbf{H}|^{\sigma}\big)\cdot\varrho(x,y).
\end{aligned}$$

Следовательно,

$$\begin{aligned}
&\max\{\,|\,T_i(x)-T_i(y)|,\,|\,T_{i+n}(x)-T_{i+n}(y)|\,\} \\
&\quad=\;\max\big\{\,\big(|\mathfrak{G}|\,|\mathbf{H}|^{\sigma}\,\varrho(x,y)\big)_i,\;\big(|\mathfrak{G}|\,|\mathbf{H}|^{\sigma}\,\varrho(x,y)\big)_{i+n}\,\big\} \\
&\quad=\;\max\left\{ \begin{pmatrix} i\text{-ая строка} \\ \text{матрицы} \\ |\mathfrak{G}|\,|\mathbf{H}|^{\sigma} \end{pmatrix}\varrho(x,y),\; \begin{pmatrix} (i+n)\text{-ая строка} \\ \text{матрицы} \\ |\mathfrak{G}|\,|\mathbf{H}|^{\sigma} \end{pmatrix}\varrho(x,y) \right\}.
\end{aligned}$$

А поскольку внедиагональные блоки матрицы $\mathfrak{G}$ совпадают, то в целом

$$\varrho(T(x),T(y)) \le \mathfrak{G}\,|\mathbf{H}|^{\sigma}\varrho(x,y).$$

Для завершения доказательства осталось отметить, что спектральные радиусы матриц $\mathfrak{G}\,|\mathbf{H}|^{\sigma}$ и $\Gamma\,|\mathbf{H}|$ равны. $\square$

Предмет основной заботы разработчиков итерационных методов вида (10) — как можно сильнее уменьшить спектральный радиус оператора Липшица для оператора перехода T, чтобы, во-первых, обеспечить сходимость итераций, и, во-вторых, ускорить эту сходимость там, где она уже есть. Как следует из доказательства Теоремы 2, для схемы (12) матрица этого оператора Липшица равна $\mathfrak{G}\,|\mathbf{H}|^{\sigma}$ и её норма будет уменьшена при уменьшении $|\mathbf{H}|$ и увеличении $|G|$. Следовательно, для улучшения сходимости итерационной схемы (12) желательно локализовать элементы с наибольшей абсолютной величиной в матрице $\mathbf{G}$, а в матрице $\mathbf{H}$ оставить элементы с малой абсолютной величиной. Этого можно достичь за счёт предварительной перестановки строк и столбцов в $\mathbf{C}$.

4 Методы, использующие гетеротонность оператора

Ещё один плодотворный подход к построению вычислительных алгоритмов для нахождения алгебраического решения ИСЛАУ может быть основан на том, что пространство $\mathbb{IR}^n$ несёт дополнительную структуру частичного порядка по включению. Нахождению неподвижных точек монотонных и родственных им операторов в упорядоченных пространствах посвящены развитые разделы современного численного анализа [5, 6, 8]. Как правило, стандартные подходы к таким задачам основываются на фактах, являющихся вариациями широко известной *леммы Канторовича* (см., например, [8]).

Напомним следующее определение [7]:

Определение 4. Отображение $F : U \to V$, действующее из частично упорядоченного множества U в частично упорядоченное множество V называется *гетеротонным*, если существует отображение $\tilde{F} : U \times U \to V$, называемое *сопутствующим* для F, такое что

(i) $\tilde{F}(x, x) = F(x)$,

(ii) $\tilde{F}(y, z)$ является изотонным по y и антитонным по z.

Мы опираемся на общий результат [5, 6] об итерационных процессах для локализации неподвижных точек гетеротонных операторов (задействующий и теорему Брауэра). Применительно к нашей ситуации его удобно переформулировать в следующем виде:

Пусть $T(\mathbf{x})$ — непрерывный гетеротонный относительно порядка по включению оператор на $\mathbb{IR}^n$, $\tilde{T}$ — его сопутствующий оператор и для некоторого порядкового отрезка $[\,\mathbf{y}^{(0)}, \mathbf{z}^{(0)}]$ справедливо

$$\mathbf{y}^{(0)} \subseteq \tilde{T}(\,\mathbf{y}^{(0)}, \mathbf{z}^{(0)}) \subseteq \tilde{T}(\,\mathbf{z}^{(0)}, \mathbf{y}^{(0)}) \subseteq \mathbf{z}^{(0)}.$$

Тогда для последовательностей

$$\mathbf{y}^{(k)} := \tilde{T}(\,\mathbf{y}^{(k)},\, \mathbf{z}^{(k)}) \quad \textit{и} \quad \mathbf{z}^{(k)} := \tilde{T}(\,\mathbf{z}^{(k)},\, \mathbf{y}^{(k)})$$

имеем

$$\mathbf{y}^{(k)} \subseteq \mathbf{z}^{(k)}, \qquad \mathbf{y}^{(k)} \nearrow \mathbf{y}^*, \qquad \mathbf{z}^{(k)} \searrow \mathbf{z}^*, \qquad \mathbf{y}^* \subseteq \mathbf{z}^*,$$

а на отрезке $[\,\mathbf{y}^*, \mathbf{z}^*]$ находится неподвижная точка $\mathbf{x}^*$ оператора T.

Обычно на практике $\mathbf{y}^* = \mathbf{z}^*$ и неопределённость с точным значением неподвижной точки T не возникает.

Теорема 3. Операторы перехода $T : \mathbb{IR}^n \to \mathbb{IR}^n$ итерационной схемы (11) с рассмотренными в §3 двумя способами дистрибутивного расщепления $\mathbf{C}$ — гетеротонные.

Доказательство. В качестве оператора, сопутствующего оператору перехода итерационной схемы (11), мы можем взять

$$\tilde{T}(y,z) = \mathfrak{G}^{-}\sigma\big(\,\mathbf{d} \ominus \mathbf{H}\sigma^{-1}(y)\big) + \mathfrak{G}^{+}\sigma\big(\,\mathbf{d} \ominus \mathbf{H}\sigma^{-1}(z)\big), \tag{13}$$

где $\mathfrak{G}^{-}$ и $\mathfrak{G}^{+}$ — отрицательная и положительная части матрицы $\mathfrak{G} = (\,G^{\sigma})^{-1}$ соответственно. Ясно, что первое слагаемое представления (13) — изотонный относительно естественного порядка на $\mathbb{R}^{2n}$ оператор, а второе слагаемое — антитонный оператор.

Обратимся теперь к оператору перехода схемы (11) для случая треугольного дистрибутивного расщепления матрицы $\mathbf{C}$. Естественное упорядочение $\mathbb{IR}^n$ по включению является прямым произведением порядков по включению на $\mathbb{IR}$ [3]. Поэтому требуемая гетеротонность оператора T будет обоснована, если мы докажем гетеротонность каждой отдельной его компоненты $T_i(\mathbf{x})$ как функции $\mathbf{x}$. Это доказательство легко провести индукцией по номеру компоненты i. $\square$

Отметим также следующую использованную Зюзиным [4] и Куприяновой [12] итерационную схему (в $\mathbb{IR}^n$) для решения уравнения (2):

$$\mathbf{x}_i^{(k+1)} = \left(\mathbf{d}_i \ominus \sum_{j=1,\, j\neq i}^{n} \mathbf{c}_{ij}\mathbf{x}_j^{(k)}\right) \oslash \mathbf{c}_{ii}, \qquad i = 1, 2, \ldots, n. \tag{14}$$

Она является частным случаем схемы (11) с треугольным дистрибутивным расщеплением $\mathbf{C}$, когда слагаемое $\mathbf{G}$ берётся в виде *диагональной* интервальной матрицы. Достоинство схемы (14) — антитонность оператора перехода, благодаря чему её реализация особенно проста.

Зюзин [4] был первый, кто использовал лемму Канторовича и близкие к ней результаты для отыскания алгебраического интервального решения ИСЛАУ (см. также [12]). Он же столкнулся и с трудной проблемой выбора начального порядкового отрезка $[\,\mathbf{y}^{(0)}, \mathbf{z}^{(0)}]$. В заключение нашей работы мы рассмотрим один достаточно общий приём выбора начальных приближений для итерационной схемы (14). Он опирается на использование уравнений с мажорирующими и минорирующими операторами, к которым уже может быть успешно применён субдифференциальный метод Ньютона.

Для реализации схемы (14) предварительно требуется знать инвариантный порядковый отрезок $[\,\mathbf{y}^{(0)},\ \mathbf{z}^{(0)}\,]$ для оператора перехода $T : \mathbb{IR}^n \to \mathbb{IR}^n$, такого что

$$T_i(\mathbf{x}) = \left(\mathbf{d}_i \ominus \sum_{j=1,\,j\neq i}^{n} \mathbf{c}_{ij}\mathbf{x}_j \right) \oslash \mathbf{c}_{ii}, \qquad i = 1, 2, \ldots, n. \tag{15}$$

Определение 5. Пусть R, S — отображения из множества U в частично упорядоченное множество V с порядком $\preceq$ и $R(x) \preceq S(x)$ для всех $x \in U$. Тогда станем называть R *минорантой* (минорирующим отображением) для S, а S — *мажорантой* (мажорирующим отображением) для R.

Предположим, что для оператора T, задаваемого посредством (15), мы сумели построить миноранту R и мажоранту S относительно порядка по включению, для которых легко могут быть найдены неподвижные точки $\mathbf{y}'$ и $\mathbf{z}'$, причём $\mathbf{y}' \subseteq \mathbf{z}'$. Тогда имеем

$$\mathbf{y}' = R(\mathbf{y}') \subseteq T(\mathbf{y}') \qquad \text{и} \qquad T(\mathbf{z}') \subseteq S(\mathbf{z}') = \mathbf{z}'.$$

Если дополнительно оказываются выполненными неравенства

$$\mathbf{y} \subseteq T(\mathbf{z}') \qquad \text{и} \qquad T(\mathbf{y}') \subseteq \mathbf{z},$$

то порядковый отрезок $[\,\mathbf{y}',\ \mathbf{z}'\,]$ пригоден в качестве начального приближения для итерационной схемы (14).

Далее, построение минорант и мажорант по включению для оператора T также не представляет трудностей. Действительно, поскольку операции $\ominus$ и $\oslash$ антитонны по второму аргументу, то для любых матриц $\mathbf{R}, \mathbf{S} \in \mathbb{IR}^{n\times n}$, удовлетворяющих условию

$$\mathbf{S} \subseteq \mathbf{C} \subseteq \mathbf{R},$$

операторы $R, S : \mathbb{IR}^n \to \mathbb{IR}^n$, определяемы следующим образом

$$R_i(\mathbf{x}) = \left(\mathbf{d}_i \ominus \sum_{j=1,\,j\neq i}^{n} \mathbf{r}_{ij}\mathbf{x}_j \right) \oslash \mathbf{r}_{ii}, \qquad i = 1, 2, \ldots, n,$$

и

$$S_i(\mathbf{x}) = \left(\mathbf{d}_i \ominus \sum_{j=1,\,j\neq i}^{n} \mathbf{s}_{ij}\mathbf{x}_j \right) \oslash \mathbf{s}_{ii}, \qquad i = 1, 2, \ldots, n,$$

являются, соответственно, минорантой и мажорантой для T. Неподвижные точки операторов R, S совпадают, как легко видеть, с алгебраическими решениями интервальных уравнений

$$\mathbf{Rx} = \mathbf{d} \qquad \text{и} \qquad \mathbf{Sx} = \mathbf{d},$$

и теперь нужно позаботиться о применимости к этим уравнениям субдифференциального метода Ньютона.

Произведём с матрицей $\mathbf{C}$ исходной системы следующую процедуру. Если некоторая строка $\mathbf{C}$ состоит целиком из правильных или целиком из неправильных элементов, то оставим эту строку $\mathbf{C}$ без изменений. Если же некоторая строка $\mathbf{C}$ состоит как из правильных, так и из неправильных элементов, то все неправильные элементы оставим неизменными, а правильные заменим на их среднюю точку $\operatorname{mid} \mathbf{c}_{ij} = (\underline{\mathbf{c}}_{ij} + \overline{\mathbf{c}}_{ij})$. В результате получится матрица, содержащаяся в $\mathbf{C}$, в каждой строке которой все элементы уже только правильные или только неправильные. Её можно взять в качестве матрицы $\mathbf{S}$.

Напротив, для конструирования матрицы $\mathbf{R}$ произведем с $\mathbf{C}$ процедуру, двойственную вышеописанной. Если некоторая строка $\mathbf{C}$ состоит целиком из правильных или целиком из неправильных элементов, то оставим эту строку $\mathbf{C}$ без изменений. Если же некоторая строка $\mathbf{C}$ состоит как из правильных, так и из неправильных элементов, то все правильные элементы оставим неизменными, а неправильные заменим на их среднюю точку $\operatorname{mid} \mathbf{c}_{\mathrm{ij}} = (\underline{\mathbf{c}}_{\mathrm{ij}} + \overline{\mathbf{c}}_{\mathrm{ij}})$. В результате получится матрица, содержащая $\mathbf{C}$, в каждой строке которой все элементы уже только правильные или только неправильные.

Список литературы

[1] *Акилов Г.П., Кутателадзе С.С.* Упорядоченные векторные пространства. – Новосибирск: Наука, 1978.

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

[3] *Биркгоф Г.* Теория решеток. – М.: Наука, 1984.

[4] *Зюзин В.С.* Итерационный метод решения системы алгебраических сегментных уравнений первого порядка. // Дифференциальные уравнения и теория функций. – Саратов: Издательство СГУ, 1989, с. 72–82.

[5] *Коллатц Л.* Функциональный анализ и вычислительная математика. – М.: Мир, 1969.

[6] *Курпель Н.С., Шувар Б.А.* Двусторонние операторные неравенства и их приложения. – Киев: Наукова думка, 1980.

[7] *Опойцев В.И.*, Равновесие и устойчивость в моделях коллективного поведения. – М.: Наука, 1977.

[8] *Ортега Дж., Рейнболдт В.* Итерационные методы решения нелинейных систем уравнений со многими неизвестными. – М.: Мир, 1975.

[9] *Шарый С.П.* Алгебраический подход к анализу линейных статических систем с интервальной неопределённостью. // Актуальные проблемы информатики, прикладной математики и механики. – Красноярск: ВЦ СО РАН, 1995, с. 331–356.

[10] *Gardeñes E., Trepat A.* Fundamentals of SIGLA, an interval computing system over the completed set of intervals. // Computing, Vol. 24, 1980, pp. 161–179.

[11] *Kaucher E.* Interval analysis in the extended interval space $\mathbb{IR}$. // Computing Suppl. 2, 1980, pp. 33–49.

[12] *Kupriyanova L.V.* Inner estimation of the united solution set to interval linear algebraic systems // Interval Computations, 1995, в печати.

[13] *Lakeyev A.V.* Linear algebraic equation in Kaucher arithmetic. // Reliable Computing, 1995, Supplement (Extended Abstracts of APIC'95: International Workshop on Applications of Interval Computations, El Paso, TX, Febr. 23–25, 1995), pp. 130–133.

[14] *Neumaier A.* Interval methods for systems of equations. – Cambridge: Cambridge University Press, 1990.

[15] *Shary S.P.* Algebraic approach to the interval linear static identification, tolerance and control problems, or One more application of Kaucher arithmetic. // Interval Computations, 1995, в печати.