УДК 519.235 + 519.237.5

В.С. Дронов

О методе Гаусса-Зейделя в случае комплексных круговых интервалов

V.S. Dronov

The Gauss-Seidel Method in Circular Complex Intervals

Работа выполнена в рамках интервального анализа. Рассматривается обобщение интервального метода Гаусса-Зейделя на случай комплексных круговых интервалов. Доказывается эффективность метода и устанавливаются границы его применимости.

Ключевые слова: интервальный анализ, метод Гаусса-Зейделя, комплексные интервалы.

The work is made in accordance with interval iterations method. Generalization of the Gauss-Seidel interval method is examined in a case of circular complex sets. The author proves the efficiency of this method and set up its applicability range.

Key words: interval analysis, Gauss-Seidel iterations, complex sets.

Необходимые определения. Одним из классических способов учета неопределенностей является использование интервального анализа, обладающего рядом преимуществ с вычислительной точки зрения. Удобство интервального подхода (в котором в качестве основного объекта берется интервал) привело к тому, что в настоящий момент методы работы с действительными интервальными системами уравнений хорошо разработаны. Тем не менее ряд задач приводит к задачам со схожим типом неопределенности над полем комплексных чисел. Примерами могут служить задачи мезомеханики [1], оценки диэлектрической проницаемости или теплопереноса [2] и др. Для комплексного случая теория решения систем даже линейных интервальных уравнений развита в значительно меньшей степени. Если в действительном случае понятие интервала естественно, то в комплексном подобного единства нет. В различных работах в качестве базового объекта выступают как круги на комплексной плоскости (круговые интервалы), так и прямоугольники либо более сложные объекты, например, круговые сектора (см., например [3]), круговые кольца [4] и тому подобные объекты, так или иначе допускающие задание через интервальные параметры. Основным объектом данной работы будет круговой комплексный интервал $\langle c, r\rangle$ (где $r \geq 0$, $r \in R$) – круговая область комплексной плоскости $\{x \in C : |x-c| \leq r\}$. (Далее интервалы и интервальные объекты всюду обозначаются жирным шрифтом). Результатом арифметической операции над интервалами будем называть наименьший по включению интервал данного типа, включающий в себя множество всех результатов этой арифметической операции над представителями данных интервалов. Под операторами rad и mid будем понимать выделение радиуса r и центра c интервала $\langle c, r\rangle$; под модулем интервала $|\langle c, r\rangle|$ – действительное число $|c|+r$. В дальнейших рассуждениях будет играть роль определение произведения круговых интервалов:

$$\langle a, r\rangle \cdot \langle b, R\rangle = \langle ab, |a|R + |b|r + Rr\rangle,$$

и его прямое следствие:

$$\langle a, r\rangle \cdot \langle b, cR\rangle \supseteq \langle ab, crR\rangle,$$

а также формула для обращения комплексного нольнесодержащего кругового интервала:

$$\frac{1}{\langle a, r\rangle} = \langle \frac{a^*}{|a|^2 - r^2}, \frac{r}{|a|^2 - r^2}\rangle.$$

(Звездочка в последней формуле означает сопряжение). Помимо этого, обратим внимание также на условие включения:

$$\langle a, r\rangle \subseteq \langle b, R\rangle \Leftrightarrow |b-a| \leq R - r.$$

Далее рассматривается система уравнений вида $\mathbf{Ax} = \mathbf{b}$, где $\mathbf{A}$ – интервальная матрица размерности $m \times n$, а $\mathbf{b}$ – интервальный вектор длины m. Для систем подобного рода нет однозначного понятия решения, так как задача может быть поставлена и в виде нахождения x, подходящего для любого значения, входящего в интервалы правой и левой частей, так и для хоть какого-то из значений, попадающих в эти интервалы, и т.д. В данной работе рассматривается в первую очередь так называемое объединенное множество решений:

$$\Xi_{uni}(\mathbf{A}, \mathbf{b}) = \{x \in C | \exists A \in \mathbf{A}, \exists \mathbf{b} \in \mathbf{B} : \mathbf{Ax} = \mathbf{b}\},$$

которое далее будет называться просто множеством решений. Поскольку сами по себе множества решений могут иметь весьма причудливую форму даже для задач малой размерности, обычно в интервальных методах подразумевается нахождение минимального интервала, включающего данное множество (внешнее оценивание) или максимального по включению, целиком содержащегося в нем (внутреннее оценивание).

Интервальный метод Гаусса-Зейделя. Прямой перенос методов решения систем уравнений с "точечного" случая на интервальный обычно непродуктивен даже в действительном случае – так прямой аналог метода Гаусса (полученный заменой обычных операций на интервальные) "разваливается" уже на некоторых квадратных матрицах размерности 3. Тем не менее существуют удачные обобщения, в числе которых хорошо изученный метод Гаусса-Зейделя для внешнего оценивания множеств решений. Для действительного случая его алгоритм можно описать следующим псевдокодом:

На входе: система $\mathbf{Ax} = \mathbf{b}$, начальная оценка множества решений y с шириной d, требуемая точность ϵ.

$DO\ WHILE\ d > \epsilon$
  $FOR\ i = 1\ TO\ n$
    $\mathbf{y_i} = \mathbf{x_i} \cap (\mathbf{b_i} - \sum_{\mathbf{j=1}}^{\mathbf{i-1}} \mathbf{a_{ij} y_j} - \sum_{\mathbf{j=i+1}}^{\mathbf{n}} \mathbf{a_{ij} y_j}) / \mathbf{a_{ii}}$
  $IF\ x_i = \emptyset\ THEN\ STOP$ *(Решений нет)*
  $END\ IF$
$END\ FOR$
$d = dist(\mathbf{x}, \mathbf{y})$
$\mathbf{x} = \mathbf{y}$
$END\ DO$

Ключевыми результатами для действительного случая являются теорема Барта-Нудинга [6], утверждающая, что если матрица $\mathbf{A}$ в исходной системе относится к классу так называемых М-матриц, то метод Гаусса-Зейделя сходится к оптимальной внешней оценке для любого начального приближения, включающего данное множество решений, и теорема Ноймайера [5], утверждающая, что если исходная матрица не относится к классу Н-матриц (обобщающих понятие М-матриц), то существует сколь угодно широкая начальная оценка, не улучшаемая методом Гаусса-Зейделя.

Рассмотрим комплексную систему интервальных линейных уравнений $\mathbf{Ax} = \mathbf{b}$, где $\mathbf{A}$ – квадратная матрица размерности n, компонентами которой служат комплексные круговые интервалы (возможно, вырожденные), $\mathbf{b}$ – аналогичный вектор правых частей. Предположим также, что элементы главной диагонали матрицы системы не являются нульсодержащими интервалами.

Комплексный случай. С учетом того, что для круговых интервалов определены все операции, задействованные в алгоритме Гаусса-Зейделя, можно попробовать осуществить его перенос на комплексный случай. Единственным проблемным моментом будет взятие пересечения в цикле FOR – пересечение двух круговых интервалов не является в общем случае круговым интервалом. Тем не менее эта проблема может быть нейтрализована путем замены пересечения на его оболочку (*hull*) – минимальный круговой интервал, включающий данное множество.

Утверждение. В случае пересечения двух круговых интервалов, оболочка от них может быть выписана явно:

$$hull(\langle c_1, r_1 \rangle, \langle c_2, r_2 \rangle) = \langle z, R \rangle,$$

где

$$R = \frac{\sqrt{(r_1 + r_2 + |c_1 - c_2|)(r_1 + r_2 - |c_1 - c_2|)}}{2} \frac{\sqrt{(r_1 - r_2 + |c_1 - c_2|)(r_2 + |c_1 - c_2| - r_1)}}{||c_1 - c_2||},$$

$$z = \frac{c_1 \sqrt{{r_2}^2 - R^2} + c_2 \sqrt{{r_1}^2 - R^2}}{\sqrt{{r_2}^2 - R^2} + \sqrt{{r_1}^2 - R^2}}.$$

В дальнейшем под пересечением круговых интервалов всюду будет пониматься оболочка пересечения.

Заметим также, что в силу выпуклости круговых интервалов радиус оболочки заведомо не превосходит меньшего из радиусов пересекаемых интервалов.

Теорема 1. Построенный описанным выше способом аналог метода Гаусса- Зейделя является методом улучшения внешней оценки множества решений (как минимум не ухудшая ее на каждом шаге).

Доказательство. Пусть $\mathbf{x}$ – внешняя оценка множества решений, $z \in \Xi_{uni} \cap \mathbf{x}$, тогда для некоторых $A \in \mathbf{A}$ и $b \in \mathbf{b}$ $Az = b$, т.е.

$$\sum_{i=1}^{n} a_{ij} z_i = b_i.$$

Перенося в правые части этих равенств все слагаемые, кроме соответствующих диагональным элементам матрицы, и производя деление (допустимое, так как диагональные интервалы не содержат нулей), получим:

$$z_i = (b_i - \sum_{i \neq j} a_{ij} z_j)/a_{ii}.$$

По введенному выше определению действий над комплексными интервалами, подобное выражение лежит в пределах интервала

$$\mathbf{x'_i} = (\mathbf{b_i} - \sum_{\mathbf{i \neq j}} \mathbf{a_{ij} z_j})/\mathbf{a_{ii}},$$

т.е. $z_i \in \mathbf{x'_i}$, откуда $\Xi_{uni} \subseteq \mathbf{x'}$, т.е. $\mathbf{x'}$ – новая внешняя оценка Ξ_{uni}. Так как пересечение ее со старой оболочкой не увеличивает радиус, а данное пересечение и есть результат применения построенного метода, то метод обеспечивает неухудшение оценки на каждом шаге. Теорема доказана.

К сожалению, традиционный для интервальных методов "эффект обертывания" проявляется и в этом методе, причем сильнее, чем для действительного случая, так как произведение невырожденных круговых интервалов само по себе является оболочкой точного результата.

Рассмотрим теперь ограничения предлагаемого метода. Теорема Ноймайера в действительном случае использует понятие H-матрицы, обобщающее понятие матрицы монотонного вида (M-матрицы), которое может быть введено различными способами, но в конечном итоге сводится к преобладанию диагонали матрицы над прочей частью в спектральном смысле.

Определение. Будем называть K-матрицей квадратную матрицу комплексных круговых интервалов размерности n, для которой $\forall \mathbf{U}$ – ненулевого вектора комплексных круговых интервалов размерности n, $mid(U_i) = 0$ – выполняется условие $|\sum_{i \neq j} \mathbf{a_{ij} U_j}| < |\mathbf{a_{ii} U_i}|$.

Теорема 2. Если в системе интервальных уравнений $\mathbf{Ax} = \mathbf{0}$ матрица $\mathbf{A}$ не является K-матрицей, то существует исходная внешняя оценка $\mathbf{x}$ для множества решений системы, со сколь угодно широкими компонентами, не улучшаемая при помощи комплексного аналога метода Гаусса-Зейделя.

Доказательство. Пусть матрица $\mathbf{A}$ не относится к классу K-матриц. Следовательно, существует ненулевой вектор $\mathbf{U}$ с нулевыми центрами интервалов, для которого

$$|\sum_{i \neq j} \mathbf{a_{ij} U_j}| \geq |\mathbf{a_{ii} U_i}|.$$

Рассмотрим ненулевой вектор $\mathbf{U'}$, связанный с $\mathbf{U}$ соотношением $mid(\mathbf{U'}_i) = mid(\mathbf{U}_i), rad(\mathbf{U'}_i) = c \cdot rad(\mathbf{U}_i)$, где $c > 1$.

$$\sum_{i \neq j} \mathbf{a_{ij} U'_j} \supseteq \mathbf{a_{ii} U'_i},$$

для всех $i = 1, 2 \ldots n$.

Возьмем начальной оценкой множества решений вектор $\mathbf{U'}$. Первой компонентой следующего приближения по методу Гаусса-Зейделя будет интервал

$$\mathbf{y'_1} = \mathbf{U'_1} \cap (\mathbf{0} - \sum_{\mathbf{j=2}}^{\mathbf{n}} \mathbf{a_{1j} U'_j})/\mathbf{a_{11}}.$$

Но из включения выше и свойств операций над комплексными интервалами следует, что:

$$(-\sum_{j=2}^{n} \mathbf{a_{1j} U'_j})/\mathbf{a_{11}} \supseteq \mathbf{U'_1}.$$

Покажем это. Проверим условие включения.

$$mid(\sum_{j \geq 2}^{n} \mathbf{a_{1j} U'_j} - \mathbf{a_{11} U'_1}) = \mathbf{0},$$

по определению $\mathbf{U}$ как центрированного на нуле.

$$rad(\sum_{j \geq 2} \mathbf{a_{1j} U'_j}) - rad(\mathbf{a_{11} U'_1}) =$$
$$= c \cdot rad(\sum_{j \geq 2} \mathbf{a}_{1j} \mathbf{U}_j) - rad(\mathbf{a_{11} U_1}) \geq \mathbf{0},$$

откуда

$$|mid(\sum_{j \geq 2}^{n} \mathbf{a_{1j} U'_j} - \mathbf{a_{11} U'_1}| \leq$$
$$\leq rad(\sum_{j \geq 2} \mathbf{a_{1j} U'_j}) - rad(\mathbf{a_{11} U'_1})).$$

Итого $\mathbf{y'_1} = \mathbf{U'_1}$, т.е. улучшения оценки по первой координате не происходит. Аналогичные рассуждения показывают отсутствие улучшения оценки по каждой из прочих координат. Таким образом, $\mathbf{y'} = \mathbf{U'}$, т.е. оценка $\mathbf{U'}$, могущая быть сколь угодно "расширенной" за счет увеличения константы c, не поддается улучшению с помощью комплексного аналога метода Гаусса-Зейделя. Теорема доказана.

Теорема 2 играет для метода Гаусса-Зейделя в комплексном случае ту же роль, что и теорема Ноймайера в действительном, ограничивая класс матриц, на котором данный метод будет эффективен. Утверждение теоремы 2 может быть обобщено на случай ненулевой правой части.

Определение. Будем называть квадратную матрицу комплексных круговых интервалов существенно отличной от К-матрицы с коэффициентом отличия τ, если для нее $\exists \mathbf{U}$, $mid(U_i) = 0$, такой, что $|\sum_{i \neq j} \mathbf{a_{ij}} \mathbf{U_j}| > \tau |\mathbf{a_{ii}} \mathbf{U_i}|$.

Утверждение. В системе интервальнх уравнений $\mathbf{Ax} = \mathbf{b}$ с произвольной правой частью $\mathbf{b}$ при матрице $\mathbf{A}$, существенно отличной от К-матрицы с коэффициентом большим некоторого значения, существуют сколь угодно широкие начальные приближения, не улучшаемые методом Гаусса-Зейделя.

Доказательство. Аналогично теореме 2 возьмем в качестве начального приближения вектор $\mathbf{U}'$, полученный "раздуванием" радиусов компонент $\mathbf{U}$ в c раз. Улучшения по первой компоненте не произойдет, если

$$\mathbf{b_1} - \sum_{\mathbf{j=2}}^{\mathbf{n}} \mathbf{a_{1j}} \mathbf{U'_j} / \mathbf{a_{11}} \supseteq \mathbf{U'_1}.$$

Обозначив $rad(\mathbf{U_i}) = R_i$, $rad(\mathbf{a_{ij}}) = r_{ij}$ и расписав соответствующие выражения по правилам действий над круговыми интервалами, получим:

$$\langle 0, cR_1 \rangle \subseteq \langle mid(\mathbf{b}_1), rad(\mathbf{b_1}) + \\ + \frac{c(|mid(\mathbf{a_{11}})| + \mathbf{r_{11}^2})}{|mid^2(\mathbf{a_{11}})| - \mathbf{r_{11}^2}} \sum_{j=2}^{n} |\mathbf{a_{1j}}| R_j \rangle,$$

что с учётом условий включения превращается в

$$|mid(b_i)| \leq cR_1 + rad(\mathbf{b_1}) + \\ + \frac{c(|mid(\mathbf{a_{11}})| + \mathbf{r_{11}^2})}{|mid^2(\mathbf{a_{11}})| - \mathbf{r_{11}^2}} \sum_{j=2}^{n} |\mathbf{a_{1j}}| R_j.$$

Как легко заметить, правая часть неравенства растет с ростом c, если

$$R_1 - \frac{|mid(\mathbf{a_{11}})| + \mathbf{r_{11}^2}}{|mid^2(\mathbf{a_{11}})| - \mathbf{r_{11}^2}} \sum_{j=2}^{n} |\mathbf{a_{1j}}| \mathbf{R_j} > 0,$$

то есть

$$\tau > \frac{|mid(\mathbf{a_{11}})| + \mathbf{r_{11}^2}}{|mid^2(\mathbf{a_{11}})| - \mathbf{r_{11}^2}}.$$

За счет увеличения c таким образом можно получить сколь угодно широкое по первой координате начальное приближение, не улучшаемое методом Гаусса-Зейделя. Проделав аналогичные действия с прочими координатами и выбрав τ, подходящее под все ограничения, получим, что для любой матрицы $\mathbf{A}$, существенно отличной от К-матрицы с коэффициентом отличия τ и более, можно подобрать сколь угодно широкую оценку, не улучшаемую рассматриваемым методом. Утверждение доказано.

Библиографический список

1. Dessombz O., Thouverez F., Laine J.-P., Jezequel L. Analysis of Mechanical Systems using Interval Computations applied to Finite Elements Methods // Journal of Sound and Vibration. – 2001. – №5.
2. Candau Y., Raissi T., Ramdani N. and Ibos L. Complex interval arithmetic using polar form // Reliable Computing. – 2006. – №1.
3. Klatte P., Ullrich Ch. Complex sector arithmetic // Computing. – 1980. – Vol. 24.
4. Petkovic M.S., Mitrovic Z.M., Petkovic L.B. Arithmetic of circular rings // Interval Mathematics. – New York, 1986.
5. Neumaier A. Interval methods for systems of equations // Cambridge, 1990.
6. Barth W., Nuding E. Optimale Leosung von Intervallgleichungssystemen // Computing. – 1974 – Vol. 12.