Ограничения метода Гаусса-Зейделя в комплексном случае.

В.С. Дронов
Алтайский государственный университет
e-mail: `<planeswalker@rambler.ru>`

Аннотация

Работа выполнена в рамках интервального анализа. Рассматривается обобщение интервального метода Гаусса-Зейделя на случай комплексных круговых интервалов. В работе рассматривается класс матриц, на котором метод обеспечивает эффективный результат и делается вывод об ограничениях данного метода.

Необходимые определения. Одним из классических способов учета неопределенностей является использование интервального анализа, обладающего рядом преимуществ с вычислительной точки зрения. Удобство интервального подхода 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. Два наиболее частых базовых подхода - выделение в качестве основных объектов или прямоугольников на комплексной плоскости (интервальные значения комплексной и мнимой частей), либо кругов (интервальное значение радиуса) – но встречаются и более более сложные объекты, например, круговые сектора (см, например [3]), круговые кольца [4] и тому подобные объекты, так или иначе допускающие задание через интервальные параметры. Фактически, в зависимости от выбранной метрики, на роль комплексного интервала может попасть чуть ли не любой объект,

Основным объектом данной работы будет выступать круговой комплексный интервал $\langle c, r\rangle$ (где $c \in C$ $r \geq 0,\ r \in R$) – круговая область комплексной плоскости $\{x \in C : |x - c| \leq r\}$. (Здесь и 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 $|c| + r$ (максимальную удалённость элемента интервала от нуля). С помощью данных операторов можно выписать формулы для арифметических операций над круговыми интервалами в явном виде. В частности, для умножения:

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

Отметим также прямое следствие данного определения:

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

Для обращения комплексного нольнесодержащего кругового интервала используется формула:

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

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

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

Перед тем, как перейти к применению этих характеристик, отметим также весьма удобное свойство круговых интервалов, напрямую следующее из определения модуля и формулы результата умножения выше:

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

Под интервальными векторами и матрицами ниже подразумеваются веторам и матрицы, состоящие из круговых комплексных интервалов (возможно, вырожденных). Под результатом операций над матрицами, как и в случае с интервалами, подразумевается матрица, включающая результат действий над представителями участвующих в операции матриц. Наконец, в тексте ниже будет встречаться понятие «ширины» интервального вектора – в рамках данного подхода за этот показатель может быть принята любая норма на интервальных векторах. Удобным для представления вариантом может быть максимум радиусов компонент рассматриваемого вектора, но все рассуждения далее остаются в силе для произвольной нормы.

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:

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

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 для действительного интервального случая существуют удачные обобщения, в числе которых метод Гаусса-Зейделя для внешнего оценивания множеств решений. Для действительных интервалов его алгоритм можно описать следующим псевдокодом:

На входе: система $\mathbf{A}\mathbf{x} = \mathbf{b}$, начальная оценка множества решений y с шириной d, требуемая точность ϵ. В качестве расстояния между интервальными векторами $dist$ может быть принята произвольная метрика на комплексных интервальных векторах – в рамках данного алгоритма это не существенно.

$DO\ WHILE\ d > \epsilon$
 $FOR\ i = 1\ TO\ n$
  $\mathbf{y_i} = \mathbf{x_i} \cap (\mathbf{b_i} - \sum_{\mathbf{j=1}}^{\mathbf{i-1}} \mathbf{a_{ij}y_j} - \sum_{\mathbf{j=i+1}}^{\mathbf{n}} \mathbf{a_{ij}y_j})/\mathbf{a_{ii}}$
 $IF\ x_i = \emptyset\ THEN\ STOP$ *(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 [5], утверждающая, что если матрица $\mathbf{A}$ в исходной системе относится к классу так называемых M-матриц, то метод Гаусса-Зейделя сходится к оптимальной внешней оценке для любого начального приближения, включающего данное множество решений, и теорема Ноймайера [6], утверждающая, что если исходная матрица не относится к классу H-матриц (обобщающих понятие M-матриц), то существует сколь угодно широкая начальная оценка, не улучшаемая методом Гаусса-Зейделя.

В работе [7] автором было показано, что метод Гаусса-Зейделя в псевдокоде выше может быть обобщён на случай комплексных круговых интервалов, причём изменения (несмотря на иной характер комплексных операций) могут быть минимальны - фактически, для сохранения работоспособности метода достаточно ввести один дополнительный шаг, страхующий от появления при пересечении интервалов объектов, не являющихся комплексными круговыми интервалами. Так как для пересечения круговых интервалов форма минимального включающего их кругового интервала может быть выписана в явном виде.

Ограничения в комплексном случае. Ограничения интервального метода Гаусса-Зейделя в действительном случае задаются теоремой Ноймайера, использующей понятие H-матрицы, обобщающее понятие матрицы монотонного вида (M-матрицы). Понятие H-матрицы может быть введено различными способами, но в конечном итоге сводится к преобладанию диагонали матрицы над прочей частью в спектральном смысле. Из-за иной природы комплексных интервалов прямой перенос теоремы, естественно, невозможен, однако можно построить некий аналог с использованием понятия модуля:

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

Теорема 1. Если в системе интервальных уравнений $\mathbf{A}\mathbf{x} = \mathbf{0}$ 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 К-матриц. Следовательно, существует ненулевой вектор $\mathbf{U}$ с нулевыми центрами интервалов, для которого

$$|\sum_{i\neq j}\mathbf{a}_{ij}\mathbf{U}_j| \geq |\mathbf{a}_{ii}\mathbf{U}_i|.$$

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 $c > 1$.

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

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

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

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

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

Покажем это. Проверим условие включения.

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

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

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

откуда

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

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.

Данная теорема играет для метода Гаусса-Зейделя в комплексном случае ту же роль, что и теоремаНоймайера в действительном, ограничивая класс матриц, на котором данный метод будет эффективен. Утверждение теоремы 1 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 $|\sum_{i\neq j}\mathbf{a}_{ij}\mathbf{U}_j| > \tau|\mathbf{a}_{ii}\mathbf{U}_i|$.

Теорема 2. В системе интервальных уравнений $\mathbf{Ax} = \mathbf{b}$ с произвольной правой частью $\mathbf{b}$ при матрице $\mathbf{A}$, существенно отличной от К-матрицы с коэффициентом большим

$$\max_{i=1..n}\{\frac{|mid(\mathbf{a}_{ii})| + rad^2(\mathbf{a}_{ii})}{|mid^2(\mathbf{a}_{ii})| - rad^2\mathbf{a}_{ii}}\sum_{i\neq j}|\mathbf{a}_{ij}|\},$$

существуют сколь угодно широкие начальные приближения, не улучшаемые методом Гаусса-Зейделя.

Доказательство. Аналогично теореме 1 возьмем в качестве начального приближения вектор $\mathbf{U}'$, полученный «раздуванием» радиусов компонент $\mathbf{U}$ в c раз. Улучшения по первой компоненте не произойдет, если

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

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

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

что с учётом условий включения превращается в

$$|mid(b_i)| \leq cR_1 + rad(\mathbf{b}_1) + \frac{c(|mid(\mathbf{a}_{11})| + r_{11}^2)}{|mid^2(\mathbf{a}_{11})| - r_{11}^2} \sum_{j=2}^{n} |\mathbf{a}_{1j}| R_j.$$

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

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

то есть

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

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. Теорема доказана.

Таким образом, для уверенной работы метода Гаусса-Зейделя желательна система с К-матрицей или матрицей близкой к ней.

Утверждение. Класс К-матриц пуст.

Доказательство. Пусть - квадратная интервальная матрица, и $\mathbf{U}$ – центрированный на нуле вектор, для которого $\sum_{i \neq j} |\mathbf{a}_{ij}\mathbf{U}_j| < |\mathbf{a}_{ii}\mathbf{U}_i|, \; i = 1..n$. В силу свойства модулей произведения круговых интервалов, отмеченного в первой части работы и определения модуля, данное равенство может быть записано как $\sum_{i \neq j} |\mathbf{a}_{ij}| rad(\mathbf{U}_j) < |\mathbf{a}_{ii}| rad(\mathbf{U}_i), \; i = 1..n$. Как легко заметить, увеличивая ширину любой не i-той компоненты вектора $\mathbf{U}$, а прочие сохраняя неизменными, мы рано или поздно получим вектор, нарушающий характеристическое условие К-матрицы для . Утверждение доказано.

Таким образом для комплексного случая класс матриц, соответствующий классу Н-матриц оказывается пустым, а метод Гаусса-Зейделя, несмотря на свою работоспособность, не гарантирует сходимости к оптимальной (т.е. наиболее узкой) оценке.

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. Barth W., Nuding E. Optimale Leosung von Intervallgleichungssystemen // Computing. 1974. Vol. 12.
6. Neumaier A. Interval methods for systems of equations. Cambridge. 1990.
7. Дронов В.С. О теореме Ноймайера в комплексной круговой арифметике. // Сборник научных статей межригиональной школы-семинара «Ломносовские чтения на Алтае», ч.1, Барнаул, 2010, стр 184-190.