

САНКТ-ПЕТЕРБУРГСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ ИМЕНИ ПЕТРА
ВЕЛИКОГО

ФИЗИКО-МЕХАНИЧЕСКИЙ ИНСТИТУТ
ВЫСШАЯ ШКОЛА ПРИКЛАДНОЙ МАТЕМАТИКИ И ФИЗИКИ

**Отчет по курсовой работе
по дисциплине "Интервальный анализ"
Тема: Субдифференциальный метод Ньютона**

Выполнил:

Студент: Байрамов Самир

Группа: 5030102/00201

Принял:

к. ф.-м. н., доцент

Баженов Александр Николаевич

Санкт-Петербург
2023 г.

Содержание

1	Постановка задачи	2
2	Теория	2
2.1	Введение	2
2.2	Алгоритм субдифференциального метода Ньютона	3
3	Результаты	4
3.1	4
3.2	6
3.2.1	Коррекция вектора правой части.	6
3.2.2	Коррекция матрицы.	6
3.2.3	Уменьшение колебаний	7
4	Вывод	8
5	Исходный код	9

1 Постановка задачи

1. Задана ИСЛАУ $x = \mathbf{C}x + \mathbf{d}$, где

$$\mathbf{C} = \begin{pmatrix} [12, 16] & [2, 3] \\ [2, 3] & [26, 30] \end{pmatrix}, \quad \mathbf{d} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$$

Необходимо найти решение данной системы интервальных линейных уравнений с помощью субдифференциального метода Ньютона.

2. Задана ИСЛАУ $Ax = b$, где

$$\mathbf{A} = \begin{pmatrix} [3, 4] & [5, 6] \\ [-1, 1] & [-3, 1] \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} [-3, 4] \\ [-1, 2] \end{pmatrix}$$

Субдифференциальный метод Ньютона порождает осциллирующую последовательность:

$$\mathbf{x}^k = \begin{pmatrix} [-0.3333, 1] \\ [-0.3333, 0] \end{pmatrix}, \quad \mathbf{x}^{k+1} = \begin{pmatrix} [0, 0.5] \\ [-0.5, 0.3333] \end{pmatrix}$$

Начальное приближение:

$$\mathbf{x}^0 = \begin{pmatrix} [0, 1] \\ [0, 1] \end{pmatrix}$$

Необходимо исследовать возможность коррекции задачи для достижения разрешимости: коррекция правой части и матрицы при исходной правой части.

2 Теория

2.1 Введение

Мы имеем уравнение, заданное в рекуррентном виде,

$$x = \mathbf{C}x + \mathbf{d}, \quad \mathbf{C} = (I \ominus \mathbf{A}^c), \quad \mathbf{d} = \mathbf{b}^c,$$

и мы можем записать равносильное ему в $\mathbb{R}\mathbb{R}$:

$$\mathbf{C}x \ominus x + \mathbf{d} = 0.$$

Для расширения наших возможностей при нахождении формального решения мы перейдем из нелинейного пространства $\mathbb{K}\mathbb{R}^n$ в некое линейное пространство U , построив биективное отображение, называемое **вложением**:

$$\iota : \mathbb{K}\mathbb{R}^n \rightarrow U,$$

причем для каждого отображения $\phi : \mathbb{K}^n \rightarrow \mathbb{K}\mathbb{R}^n$ существует лишь единственное индуцированное отображение

$$\iota \circ \phi \circ \iota^{-1} : U \rightarrow U$$

где « \circ » служит для обозначения композиции отображений.

Тесная связь свойств отображений ϕ и $\iota \circ \phi \circ \iota^{-1}$ позволит нам заменить исходную задачу решения уравнения в $\mathbb{K}\mathbb{R}$

$$\mathbf{f}(x) = 0, \quad \mathbf{f} : \mathbf{x} \mapsto \mathbf{C}\mathbf{x} \ominus \mathbf{x} + \mathbf{d}$$

на задачу решения индуцированного уравнения в U :

$$\begin{aligned} \mathcal{F}(y) &= \iota(0), \\ \mathcal{F} &= \iota \circ \mathbf{f} \circ \iota^{-1} : U \rightarrow U, \\ \mathcal{F}(y) &= \iota(\mathbf{C}\iota^{-1}(y) \ominus \iota^{-1}(y) + \mathbf{d}), \end{aligned}$$

и однозначно восстановить формальное интервальное решение \mathbf{x}^* по y^* из соотношения $\mathbf{x}^* = \iota^{-1}(y^*)$.

Определение. Биективные вложения $\iota : \mathbb{K}^n \rightarrow U$, которые

- сохраняют аддитивную алгебраическую структуру \mathbb{K}^n , т. е. такие, что $\iota(\mathbf{u} + \mathbf{v}) = \iota(\mathbf{u}) + \iota(\mathbf{v})$ для любых $\mathbf{u}, \mathbf{v} \in \mathbb{K}^n$,
- сохраняют топологическую структуру \mathbb{K}^n , т. е. такие, что само отображение $\iota : \mathbb{K}^n \rightarrow U$ и его обратное $\iota^{-1} : U \rightarrow \mathbb{K}^n$ непрерывны,

будем называть **погружениями интервального пространства** \mathbb{K}^n в линейное пространство U .

Определение свойств погружения позволяет сделать два важных вывода. Во-первых, оно однозначно задает линейное пространство: U должно быть евклидовым пространством \mathbb{R}^{2n} . Во-вторых, из определения погружения следует, что любые два погружения \mathbb{K}^n в \mathbb{R}^{2n} одинаковы с точностью до неособенного линейного преобразования \mathbb{R}^{2n} .

Определение. Погружение $sti : \mathbb{K}^n \rightarrow \mathbb{R}^{2n}$, которое действует по правилу

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

т.е. такое, при котором взятые с противоположным знаком левые концы интервалов $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ становятся первой, второй, ..., n -ой компонентами точечного $2n$ -вектора, а правые концы $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ становятся $(n+1)$ -ой, ..., $2n$ -ой компонентами точечного $2n$ -вектора соответственно, будем называть **стандартным погружением интервального пространства** \mathbb{K}^n в \mathbb{R}^{2n} .

2.2 Алгоритм субдифференциального метода Ньютона

Алгоритм субдифференциального метода Ньютона имеет следующий вид:

1. Выберем некоторое начальное приближение $x^{(0)} \in \mathbb{R}^{2n}$.
2. Если $(k-1)$ -е приближение $x^{(k-1)} \in \mathbb{R}^{2n}$, $k = 1, 2, \dots$, уже найдено, то вычисляем какой-нибудь субградиент $D^{(k-1)}$ отображения \mathcal{F} в точке $x^{(k-1)}$ и полагаем

$$x^{(k)} \leftarrow x^{(k-1)} - \tau (D^{(k-1)})^{-1} \mathcal{F}(x^{(k-1)}),$$

где $\tau \in [0, 1]$ - релаксационный параметр.

3 Результаты

3.1

Поскольку все элементы матрицы \mathbf{C} являются правильными интервалами, то в качестве начального приближения $x^{(0)}$ мы можем выбрать решение специальной $2n \times 2n$ -системы точечных уравнений:

$$(I - (\text{mid } \mathbf{C})^\sim) x = \text{sti}(\mathbf{d}).$$

Результаты, получаемые на каждой итерации субдифференциального метода Ньютона (для выбранного значения релаксационного параметра $\tau = 1$):

k	$x^{(k)}$
0	$\begin{pmatrix} [-0.1232, -0.1232] \\ [-0.1342, -0.1342] \end{pmatrix}$
1	$\begin{pmatrix} [-0.1079, -0.1549] \\ [-0.1267, -0.1476] \end{pmatrix}$

Полученные результаты позволяют сделать два вывода.

1. Получение искомого корня — интервального вектора $x^{(1)}$ — уже на первой итерации свидетельствует о *высокой сходимости* субдифференциального метода Ньютона.
2. Можно заметить, что обе компоненты вектора-решения $x^{(1)}$ являются неправильными интервалами.

Попробуем изменять элементы интервальных матрицы \mathbf{C} и вектора \mathbf{d} и посмотрим, как эти изменения отразятся на решении системы

Сначала будем изменять вектор \mathbf{d} . В первую очередь увеличим радиус компоненты \mathbf{d}_1 на 6. Тогда ИСЛАУ примет вид

$$x = \begin{pmatrix} [12, 16] & [2, 3] \\ [2, 3] & [26, 30] \end{pmatrix} x + \begin{pmatrix} [-4, 8] \\ 4 \end{pmatrix}$$

Решение достигается уже на третьей итерации:

k	$x^{(k)}$
0	$\begin{pmatrix} [-0.5753, 0.3287] \\ [-0.0931, -0.1753] \end{pmatrix}$
1	$\begin{pmatrix} [0.3004, -0.5203] \\ [-0.1690, -0.0975] \end{pmatrix}$
2	$\begin{pmatrix} [0.4089, -0.7084] \\ [-0.1661, -0.1033] \end{pmatrix}$

Теперь вернемся к исходной компоненте \mathbf{d}_1 , равной 2, и «расширим» компоненту вектора \mathbf{d}_2 так же на 6:

$$x = \begin{pmatrix} [12, 16] & [2, 3] \\ [2, 3] & [26, 30] \end{pmatrix} x + \begin{pmatrix} 2 \\ [-2, 10] \end{pmatrix}$$

k	$x^{(k)}$
0	$\begin{pmatrix} [-0.0821, -0.1643] \\ [-0.3561, 0.0876] \end{pmatrix}$
1	$\begin{pmatrix} [-0.1502, -0.0894] \\ [0.0845, -0.3386] \end{pmatrix}$
2	$\begin{pmatrix} [-0.1463, -0.1107] \\ [0.0975, -0.3911] \end{pmatrix}$

Тогда получаемый на третьей итерации субдифференциального метода Ньютона интервальный вектор-решение примет вид

Таким образом, увеличивая радиус компоненты \mathbf{d}_1 , мы «превратили» компоненту вектора-решения \mathbf{x}_2 в правильный интервал. И наоборот: «расширяя» компоненту \mathbf{d}_2 , мы добились того, что компонента \mathbf{x}_1 стала правильным интервалом. Аналогичного результата можно добиться путем поочередного увеличения радиусов элементов интервальной матрицы \mathbf{C} .

Означает ли это, что если мы «расширим» обе компоненты вектора \mathbf{d} , вектор \mathbf{x} будет содержать только правильные интервальные компоненты?

Решением ИСЛАУ

$$\mathbf{x} = \begin{pmatrix} [12, 16] & [2, 3] \\ [2, 3] & [26, 30] \end{pmatrix} \mathbf{x} + \begin{pmatrix} [-4, 8] \\ [-2, 10] \end{pmatrix}$$

является интервальный вектор \mathbf{x}

k	$x^{(k)}$
0	$\begin{pmatrix} [-0.53424, 0.2876] \\ [-0.3150, 0.0465] \end{pmatrix}$
1	$\begin{pmatrix} [0.2582, -0.4741] \\ [0.0422, -0.2957] \end{pmatrix}$
2	$\begin{pmatrix} [0.3542, -0.6642] \\ [0.0516, -0.3468] \end{pmatrix}$

и обе его компоненты являются неправильными векторами.

Сравним последнее полученное нами решение \mathbf{x} ИСЛАУ с расширенным вектором \mathbf{d} с решением $\mathbf{x}^{(1)}$ исходной ИСЛАУ:

$$\mathbf{x}^{(1)} = \begin{pmatrix} [-0.1079, -0.1549] \\ [-0.1267, -0.1476] \end{pmatrix}$$

Поскольку к неправильному интервалу в полной интервальной арифметике Каухера неприменимо понятие «ширина», определим правильные проекции сравниваемых векторов:

$$pro \mathbf{x} = \begin{pmatrix} [-0.6642, 0.3542] \\ [-0.3468, 0.0516] \end{pmatrix}, \quad pro \mathbf{x}^{(1)} = \begin{pmatrix} [-0.1549, -0.1079] \\ [-0.1476, -0.1267] \end{pmatrix}$$

Легко заметить, что радиусы элементов вектора $pro \mathbf{x}$, относящегося к ИСЛАУ с расширенным вектором \mathbf{d} , значительно больше радиусов элементов вектора $pro \mathbf{x}^{(1)}$, который является решением исходной ИСЛАУ с точечным вектором \mathbf{d} .

3.2

3.2.1 Коррекция вектора правой части.

Для коррекции вектора правой части использовался следующий метод:

1. Вычисляется середина-вектор правой части $\mathbf{b} - \text{mid } \mathbf{b}$.
2. Генерируется 4-мерная сетка для левых и правых радиусов для соответствующих компонент: $[0, \text{rad}(\mathbf{b}_1)]^2 \times [0, \text{rad}(\mathbf{b}_2)]^2$ с шагом дискретизации 0.1 .
3. В качестве метрики для оценки расстояния между новыми и исходными данными выбрана:

$$\rho(\mathbf{b}_{\text{new}}) = \|\text{dist}(\mathbf{b}_{\text{new}}, \mathbf{b})\|,$$

где $\text{dist}(\mathbf{c}, \mathbf{d})$ - покомпонентный максимум из разностей верхних и нижних границ интервалов \mathbf{c}, \mathbf{d} .

Ищется вектор \mathbf{b}_{new} : $\min_{\tau=1} \rho(\mathbf{b}_{\text{new}}) \& \exists \mathbf{x}_b : \mathbf{A}\mathbf{x}_b = \mathbf{b}_{\text{new}}$. Эксперименты проводились для

Полученные результаты: $\mathbf{b}_{\text{opt}} = \begin{pmatrix} [-2.5, 3.5] \\ [-1, 1.5] \end{pmatrix}$; решение: $\mathbf{x} = \begin{pmatrix} [-0.255, 0.75] \\ [-0.25, 0.085] \end{pmatrix}$.

$$\mathbf{A}\mathbf{x} = \mathbf{b}_{\text{opt}}$$

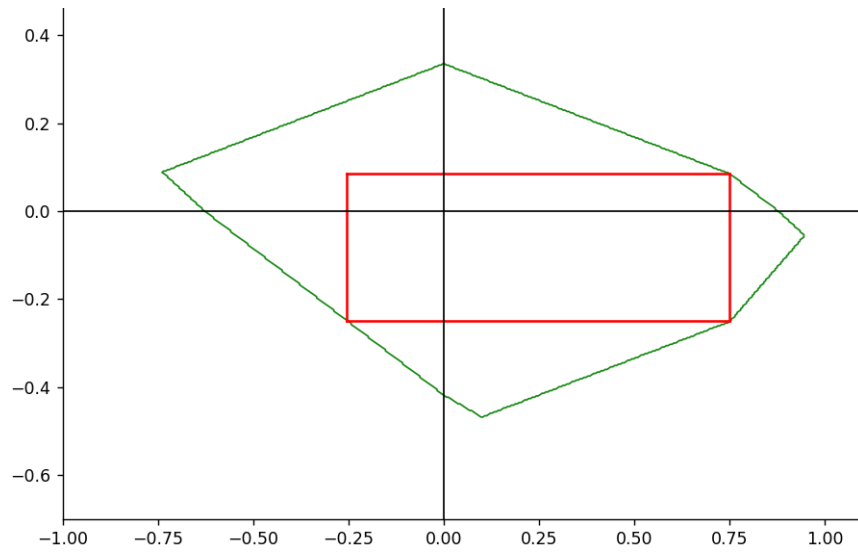


Рис. 1: Решение и допусковое множество

3.2.2 Коррекция матрицы.

Производилась вариация радиусов матричных элементов с исходной правой частью, и подбором были получены матрицы, для которых итерационный процесс сходится. В

частности, $\mathbf{A}'_{\text{opt}} = \begin{pmatrix} [3, 4] & [5, 6] \\ [-0.6, 0.6] & [-3, 1] \end{pmatrix}$, для которой достигается решение:

$$\mathbf{x} = \begin{pmatrix} [0.05, 0.71] \\ [-0.52, 0.19] \end{pmatrix}.$$

Расширением \mathbf{a}_{21} на величины 0.1 влево и 0.4 вправо при дискретности шага 0.01 была найдена новая матрица, более близкая к исходной: $\mathbf{A}_{\text{opt}} = \begin{pmatrix} [3, 4] & [5, 6] \\ [-0.75, 1] & [-3, 1] \end{pmatrix}$,

$$\text{решение } x = \begin{pmatrix} [-0.15, 0.8] \\ [-0.4, 0.133] \end{pmatrix}$$

$$\mathbf{A}_{\text{opt}} \mathbf{x} = \mathbf{b}$$

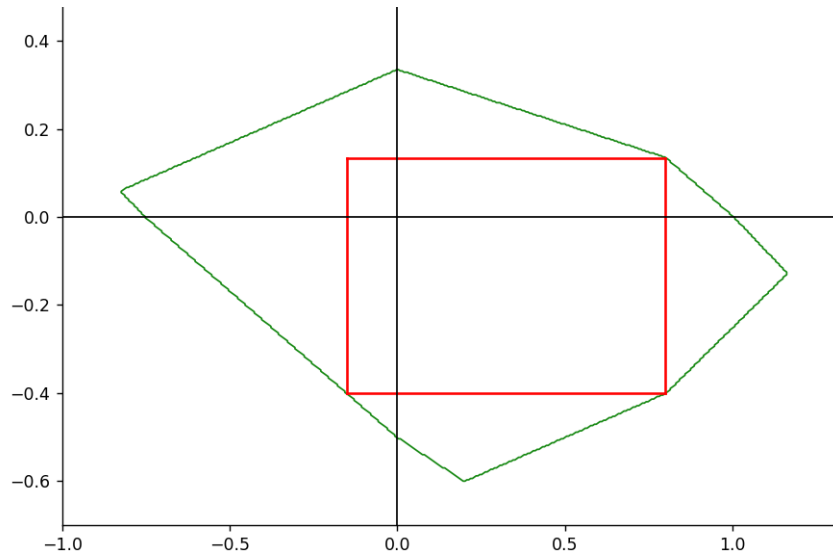


Рис. 2: Решение и допустовое множество

3.2.3 Уменьшение колебаний

Приведём промежуточные варианты, для которых колебания решений есть, но меньше, чем для исходной системы:

$$1. \mathbf{b} = \begin{pmatrix} [-2.7, 3.9] \\ [-1, 2] \end{pmatrix} \text{ Решение } \mathbf{x}_k = \begin{pmatrix} [-0.18, 0.35] \\ [0.18, -0.02] \end{pmatrix}, \mathbf{x}_{k+1} = \begin{pmatrix} [0.14, 0.5] \\ [-0.5, 0.33] \end{pmatrix}$$

$$\mathbf{A} \mathbf{x} = \mathbf{b}$$

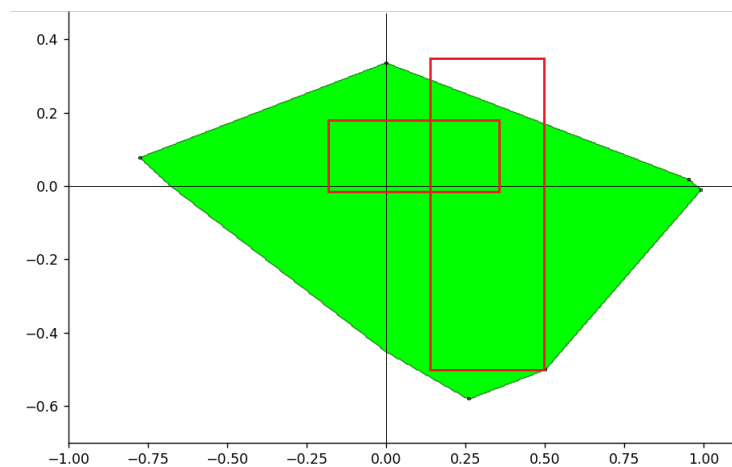


Рис. 3: Решение и допустовое множество

$$2. \mathbf{b} = \begin{pmatrix} [-2.7, 3.9] \\ [-1, 1.7] \end{pmatrix},$$

$$\text{Решение } \mathbf{x}_k = \begin{pmatrix} [0.15, 0.35] \\ [0.15, 0.02] \end{pmatrix}, \mathbf{x}_{k+1} = \begin{pmatrix} [0.3, 0.25] \\ [0.25, 0.02] \end{pmatrix}$$

$$\mathbf{Ax} = \mathbf{b}$$

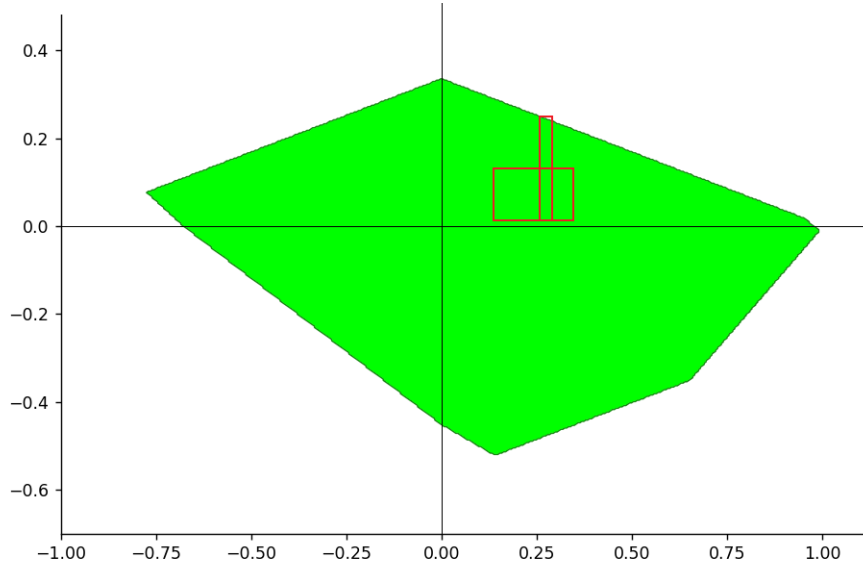


Рис. 4: Решение и допусковое множество

4 Вывод

1. Реализован субдифференциальный метод Ньютона на языке программирования Си.
2. Полученные результаты свидетельствуют о высокой скорости сходимости субдифференциального метода Ньютона.
3. Варьируя ширину или, что равносильно, радиусы элементов интервальных матрицы \mathbf{C} и вектора \mathbf{d} , можно в достаточно широких пределах управлять решением ИСЛАУ вида $x = \mathbf{C}x + \mathbf{d}$
4. Рассмотрен пример задачи, в которой есть сложности с нахождением устойчивого формального решения ИСЛАУ субдифференциальным методом Ньютона. Приведены способы получения решения: коррекция матрицы или правой части.

5 Исходный код

Ссылка на [github](#)

Пояснение к построению графика:

1. Создаются двумерные массивы x и y , представляющие собой координаты сетки
2. Вычисляется значение функционала f для текущей пары координат $(x[i][j], y[i][j])$ с использованием функции `Tol` из модуля `intvalpy.linear`
3. Если значение f больше или равно нулю, оно присваивается соответствующему элементу массива z
4. Наконец, используется метод `contour` объекта осей (ax), чтобы построить график