Санкт-Петербургский политехнический университет имени Петра Великого

Институт прикладной математики и механики

Кафедра прикладной математики

Интервальный анализ
Отчёт по лабораторной работе №5

Выполнил:

Студент: Дамаскинский Константин
Группа: 3630102/70201

Принял:

к. ф.-м. н., доцент
Баженов Александр Николаевич

2021 г.

Содержание

1. Постановка задачи

1.1. Задача 1

Требуется найти формальное решение ИСЛАУ с матрицей

$$\mathbf{A} = \begin{pmatrix} [0.9, 1.1] & [0.9, 1.1] & [0.9, 1.1] \\ [0.9, 1.1] & [0.9, 1.1] & [0.0, 0.0] \\ [0.9, 1.1] & [0.0, 0.0] & [0.0, 0.0] \end{pmatrix} \tag{1.1.1}$$

и правой частью, содержащей неправильные интервалы:

$$\mathbf{b} = \begin{pmatrix} [2.8, 3.2] \\ [2.2, 1.8] \\ [0.9, 1.1] \end{pmatrix} \tag{1.1.2}$$

1.2. Задача 2

Дана точечная матрица $A \in \mathbb{R}^{256\times 36}$ и свободный столбец $b \in \mathbb{R}^{256}$.
Требуется объинтервалить вектор b, получив $\mathbf{b} \in \mathbb{KR}^{256}$, такой что

$$\mathbf{b}_k = [b_k - \varepsilon, b_k + \varepsilon] \tag{1.2.1}$$

и найти такой $\mathbf{x} \in \mathbb{KR}^{36}$, что:

$$A\mathbf{x} \subset \mathbf{b} \tag{1.2.2}$$

2. Теория

2.1. Концепция решения интервальных уравнений

Речь идёт о решении интервальных уравнений вида $\mathbf{C}x = \mathbf{d}$ в полной интервальной арифметике.

Данная задача могла бы быть решена классическими численными методами, если бы не одно обстоятельство: даже полная интервальная арифметика Каухера не является линейным пространством. Ввиду данного обстоятельства предлагается следующий подход: необходимо сформировать взаимно-однозначное отображение из арифметики Каухера в некоторое линейное пространство. Поскольку речь идёт о конечномерном случае, в качестве линейного пространства можно рассматривать пространство вещественных чисел, поскольку все конечномерные линейные пространства изометрически изоморфны. Далее задача решается в линейном пространстве ранее изученными методами, а затем полученный результат обратно отображается в арифметику Каухера.

Ясно, что n-мерное пространство Каухера $\mathbb{KR}^n$ должно отображаться в $2n$-мерное евклидово пространство $\mathbb{R}^{2n}$ ввиду природы элементов арифметики Каухера (интервалов).

Далее речь пойдёт о конкретных отображениях и сопутствующей терминологии.

2.2. Погружение

Пусть φ – некоторое отображение на интервальном пространстве, задействованное при решении задачи:

$$\varphi : \mathbb{KR}^n \to \mathbb{KR}^n \tag{2.2.1}$$

Пусть имеется отображение:

$$i : \mathbb{KR}^n \to \mathbb{R}^{2n} \tag{2.2.2}$$

Тогда образ данной задачи в линейном пространстве будет решать с использованием отображения:

$$(i \circ \varphi \circ i^{-1})(y), \;\; y \in \mathbb{R}^{2n} \tag{2.2.3}$$

Поскольку отображение i должно быть обратимым, оно обязано быть биективным.

Определение 1. Биективное отображение, действующее из $\mathbb{KR}^n$ в $\mathbb{R}^{2n}$, называется **погружением** в линейное пространство.

Определение 2. **Стандартным погружением** называется отображение

$$\mathtt{sti}(\boldsymbol{x}) : \begin{pmatrix} \boldsymbol{x}_1 & \ldots & \boldsymbol{x}_n \end{pmatrix} \to \begin{pmatrix} -\underline{\boldsymbol{x}_1} & \ldots & -\underline{\boldsymbol{x}_n} & \overline{\boldsymbol{x}_1} & \ldots & \overline{\boldsymbol{x}_n} \end{pmatrix} \tag{2.2.4}$$

Оно действительно является инъективным и сюръективным, что легко проверяется по определению этих понятий [1, Лекция 11].

Кроме того, данное отображение индуцирует частичный порядок на $\mathbf{R}^{2n}$, что необходимо для корректного определения понятия выпуклости, используемого при построении численных методов оптимизации и решения ИСЛАУ [2, стр. 580].

Рассмотрим случай, когда оператор φ задаётся квадратной точечной матрицей Q:

$$\varphi(x) = Qx \tag{2.2.5}$$

В таком случае оператор $\mathtt{sti} \circ \varphi \circ \mathtt{sti}^{-1}$ будет являться линейным, определённым *знаково-блочной матрицей*:

$$\tilde{Q} = \left(\begin{array}{c|c} Q^+ & Q^- \\ \hline Q^- & Q^+ \end{array}\right) \tag{2.2.6}$$

Покажем это.

Лемма 1. *[2, стр. 583] Имеет место следующее свойство умножения интервала на число:*

$$\begin{cases} \underline{q \cdot \boldsymbol{x}} = q^+ \underline{\boldsymbol{x}} - q^- \underline{\boldsymbol{x}} \\ \overline{q \cdot \boldsymbol{x}} = -q^- \overline{\boldsymbol{x}} + q^+ \overline{\boldsymbol{x}} \end{cases} \tag{2.2.7}$$

где

$$q^+ = \begin{cases} q & q \geq 0 \\ 0 & q < 0 \end{cases} \tag{2.2.8}$$

$$q^- = \begin{cases} -q & q \leq 0 \\ 0 & q > 0 \end{cases} \tag{2.2.9}$$

Доказательство – по определению умножения числа на интервал в арифметике Каухера.

Пусть $y \in \mathbb{R}^{2n}$. Тогда:

$$\mathtt{sti}^{-1}(y) = \big([-y_1, y_{n+1}] \ldots [-y_n, y_{2n}]\big) \tag{2.2.10}$$

$$(\varphi \circ \mathtt{sti}^{-1})(y) = \begin{pmatrix} q_{11}[-y_1, y_{n+1}] + \cdots + q_{1n}[-y_n, y_{2n}] \\ \vdots \\ q_{n1}[-y_1, y_{n+1}] + \cdots + q_{nn}[-y_n, y_{2n}] \end{pmatrix} \tag{2.2.11}$$

Далее по лемме 1:

$$(\mathtt{sti} \circ \varphi \circ \mathtt{sti}^{-1})(y) = \begin{pmatrix} q^+_{11} y_1 + \cdots + q^+_{1n} y_n - q^-_{11} y_{n+1} - \cdots - q^-_{1n} y_{2n}] \\ \vdots \\ q^+_{n1} y_1 + \cdots + q^+_{nn} y_n - q^-_{n1} y_{n+1} - \cdots - q^-_{nn} y_{2n}] \end{pmatrix} = \tilde{Q} y \tag{2.2.12}$$

Кроме того, имеет место теорема [1, Лекция 11]:

Теорема 1. *Для точечной матрицы $Q \in \mathbb{R}^{n \times n}$ следующие условия равносильны:*

- *$Q\mathbf{x} = 0$ в интервальном пространстве $\mathbb{KR}^n$ тогда и только тогда, когда $\mathbf{x} = 0$*
- *Матрица $\tilde{Q} \in \mathbb{R}^{2n \times 2n}$, знаково-блочная для Q, является неособенной*
- *Неособенной является $|Q|$*

Следствие 1. *Оператор φ, введённый ранее, необратим тогда и только тогда, когда его матрица является абсолютно неособенной.*

Таким образом, теперь мы можем перейти от решения ИСЛАУ $A\mathbf{x} = \mathbf{b}$ с абсолютно неособенной матрицей A к решению стандартной СЛАУ в линейном пространстве.

2.3. Решение ИСЛАУ с точечной матрицей с помощью погружений

Получим явный вид оператора, обратного к φ.

Пусть $y \in \mathbb{R}^{2n}$. Тогда:

$$(\mathtt{sti} \circ \varphi \circ \mathtt{sti}^{-1})(y) = Qy \Leftrightarrow \tag{2.3.1}$$

$$Q^{-1}y = (\mathtt{sti}^{-1} \circ \varphi^{-1} \circ \mathtt{sti})(y) \Leftrightarrow \tag{2.3.2}$$

$$(\mathtt{sti}^{-1} \circ Q^{-1})y = (\varphi^{-1} \circ \mathtt{sti})(y) \Leftrightarrow \tag{2.3.3}$$

$$[\mathbf{x} = \mathtt{sti}^{-1}(y),\ x \in \mathbb{KR}^n] \Leftrightarrow \tag{2.3.4}$$

$$\varphi^{-1}(\mathbf{x}) = (\mathtt{sti}^{-1}(Q^{-1}(\mathtt{sti}(\mathbf{x})))) \tag{2.3.5}$$

Итак, перезапишем ИСЛАУ $Q\mathbf{x} = b$ её в виде: $\varphi(\mathbf{x}) = b$. Тогда $\mathbf{x} = \varphi^{-1}(b)$. Окончательный ответ:

$$\mathbf{x} = (\mathtt{sti}^{-1}(Q^{-1}(\mathtt{sti}(\mathbf{b})))) \tag{2.3.6}$$

Таким образом, мы свели решение интервальной СЛАУ к поиску обратной точечной матрицы. Однако зачастую легче решать не задачу поиска обратной матрицы, а задачу поиска решения СЛАУ: для этой задачи существует множество точных (LDR-, LU-разложение, методы вращений и отражений) и численных методов (методы простых итераций, Якоби, Зейделя). В таком случае 2.3.6 лучше представить как совокупность двух задач:

$$Qz = \mathtt{sti}(\mathbf{b}),\ z \in \mathbb{R}^{2n} \tag{2.3.7}$$

$$\mathbf{x} = (\mathtt{sti}^{-1}(z)) \tag{2.3.8}$$

2.4. Частичный порядок в $\mathbb{KR}^n$. Порядковая выпуклость. Условия разрешимости ИСЛАУ

Рассмотрим следующий подход к решению ИСЛАУ. Пусть сформулирована задача: $\mathbf{Cx} = \mathbf{d}$. Рассмотрим оператор:

$$\mathcal{F}(y) = \mathtt{sti}(\mathbf{C}\mathtt{sti}^{-1}(y) \ominus \mathtt{d}) : \mathbb{R}^{2n} \to \mathbb{R}^{2n} \tag{2.4.1}$$

Приравнивая $\mathcal{F}$ к нулю мы получаем задачу, индуцированную погружением исходной задачи в линейной пространство.

Хорошо развиты методы решения задача поиска корня выпуклого функционала. Однако предложенная постановка задачи составлена именно относительно отображения, действующего на конечномерном пространстве высокой размерности.

Рассмотренные ранее подходы, применимые к функционалам, как выясняется, можно применить и в этой ситуации, однако для этого необходимо перенести сопутствующую терминологию на многомерный случай.

Определение 3. Пусть $F : \mathbb{R}^p \to \mathbb{R}^q$.

Пусть также в $\mathbb{R}^q$ задан частичный порядок $\preccurlyeq$. Тогда отображение F называется **порядково выпуклым относительно** $\preccurlyeq$, если

$$F(\lambda u + (1-\lambda)v) \preccurlyeq \lambda F(u) + (1-\lambda)F(v) \; \forall u, v \in \mathbb{R}^p \; \forall \lambda \in [0;1]$$

.

Свойство (порядковой) выпуклости необходимо для построения ньютоновских методов. Однако получить его можно далеко не во всех случаях. Кроме того, необходимо описать механизм, задающий частичный порядок.

Введём для этого следующее понятие:

Определение 4. Будем называть квадратную интервальную матрицу размера n **построчно однородной**, если в каждой её строке все элементы являются либо только правильными интервалами, либо только неправильными.

Частичный порядок определяется построчно однородной матрицей $\Subset$ следующим образом: пусть $\mathcal{I}'$ – номера строк, в которых все интервалы являются правильными, а $\mathcal{I}''$ – где неправильными. Тогда $\mathbf{u} \Subset \mathbf{v}$, $\mathbf{u}, \mathbf{v} \in \mathbb{KR}^n$, если и только если

$$\begin{cases} \mathbf{u}_i \subseteq \mathbf{v}_i, & i \in \mathcal{I}' \\ \mathbf{v}_i \subseteq \mathbf{u}_i, & i \in \mathcal{I}'' \end{cases} \tag{2.4.2}$$

Стандартное погружение также индуцирует отношение порядка $\ll$ на евклидовом пространстве на основании отношения $\Subset$.

Теперь мы обладаем необходимой терминологией и способом совместного упорядочивания интервального и евклидова пространства.

Сформулируем условия, при которых отображение $\mathcal{F}$ будет порядково выпуклым.

Теорема 2. *Если матрица* $\mathbf{C}$ *обладает условием построчной однородности, то индуцированное отображение* $\mathcal{F}$ *является порядково выпуклым относительно порядка* $\ll$.

Определение 5. Пусть на $\mathbb{R}^q$ задан частичный порядок $\preccurlyeq$. Пусть $F : \mathbb{R}^p \to \mathbb{R}^q$. Тогда его **субдифференциалом** называется такое множество линейных операторов $D : \mathbb{R}^p \to \mathbb{R}^q$, что:

$$D(v) \preccurlyeq F(x+v) - F(x) \; \forall v \in \mathbb{R}^p \tag{2.4.3}$$

Элементы множества D называются **субградиентом** и обозначаются $\partial_{\ll} F(x)$.

2.5. Субдифференциальный метод Ньютона

Данный метод является модификацией метода градиентного спуска, который однако не накладывает на отображение требования дифференцируемости, а требует лишь порядковой выпуклости.

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

где $D^{(k-1)}\mathcal{F}(x^{(k-1)})$ – какой-нибудь субградиент в $x^{(k-1)})$, $\tau \in [0;1]$ – постоянный коэффициент. Алгоритм заканчивает работу, когда $\|\mathcal{F}(x^{(k)})\| < \varepsilon$.

В [2] утверждается, что при единичном τ метод даёт точный результат для полиэдральных функций, коей является $\mathcal{F}$, однако вариация этого коэффициента может помочь ускорить сходимость для других функций.

Для сходимости метода критично, чтобы $\mathcal{F}(x) \geq 0$. Если обеспечить выполнение этого условия на начальной итерации, то метод сойдётся. В качестве такой начальной точки подойдёт решение точечной СЛАУ:

$$(\mathrm{mid}\mathbf{C})\tilde{}\, x = \mathtt{sti}(\mathbf{d}) \tag{2.5.1}$$

([2, стр. 607])

2.6. Вычисление субградиента

Для вычисления субградиента используются формулы:

$$\partial \mathcal{F}_i = -\sum_{j=1}^{n} \partial(\underline{\mathbf{c}_{ij}[-x_j, x_{j+n}]}),\; i = \overline{1,n} \tag{2.6.1}$$

$$\partial \mathcal{F}_i = \sum_{j=1}^{n} \partial(\overline{\mathbf{c}_{ij}[-x_j, x_{j+n}]}),\; i = \overline{n+1, 2n} \tag{2.6.2}$$

Более детальное описание можно найти в [2] на стр. 610-613.

2.7. Адаптация субдифференциального метода Ньютона на случай прямоугольных матриц

В случае переопределённой матрицы самым простым способом решения задачи будет произвольный выбор строк матрицы ИСЛАУ и соответствующей компоненты правого столбца с проверкой соблюдения условий сходимости (построчной однородности выбранной матрицы).

Можно произвести этот процесс несколько раз, так, чтобы всех его итерациях по итогу в решении оказались задействованы все условия. Тогда итоговое решение можно представить как пересечение всех найденных.

3. Реализация

Лабораторная работа выполнена с помощью библиотек языка Python `scipy`, `numpy` и реализации `tolsolvty` Максима Смольского. Операционная система Ubuntu 20.04.

Ссылка на исходный код лабораторной работы и отчёта находится в разделе “Приложения”.

4. Результаты

Вначале убедимся, что у исходной матрицы пустое допусковое множество. Максимальное значение распознающего функционала равно

$$\texttt{tolmax = -0.201} \tag{4.0.1}$$

с точностью до тысячных, а значит, решений действительно нет.

В результате работы программы получен вектор:

$$w = \begin{pmatrix} 0 & 4.767 & 0 \end{pmatrix}^T \tag{4.0.2}$$

С помощью полученного результата исходная система была скорректирована. Максимум распознающего функционала стал неотрицательным и достиг значения

$$\texttt{tolmax = 0.009} \tag{4.0.3}$$

в точке

$$\texttt{argmax = (0.403, 0.107)}^T \tag{4.0.4}$$

(с точностью до тысячных).

Итоговая ИСЛАУ:

$$\begin{pmatrix} 5 & 6 \\ 3 & 2.5 \\ 7.2 & 4.6 \end{pmatrix} \cdot x = \begin{pmatrix} 2.5 \\ [0.524, 1.476] \\ 3.4 \end{pmatrix} \tag{4.0.5}$$

5. Обсуждение

В результате проделанной работы была скорректирована правая часть ИСЛАУ с точечной матрицей. Найденное решение в самом деле обеспечило непустоту доускового множества решения, а значит, сработало корректно.

В то же время стоит отметить, что радиус интервала одной из компонент был увеличен очень значительно: с `0.100` до `0.476`, что может быть неприемлемо в реальных задачах. Однако, остальные компоненты получили нулевой радиус.

Произведём простейшую оценку корректности полученного результата. Радиусы правой части составляют:

$$\mathrm{rad}\mathbf{b} = \begin{pmatrix} 0.25 & 0.1 & 0.1 \end{pmatrix}^T \tag{5.0.1}$$

.

Значение максимума распознающего функционала в исходной системе равно приблизительно -0.2. Значит, используя жадный подход, можно расширить все радиусы на 0.2 и гарантированно достичь неотрицательности распознающего функционала и, следовательно, непустоты допускового множества решений. Таким образом, жадный вектор масштабных коэффициентов:

$$\mathbf{w_{greedy}} = \begin{pmatrix} \frac{0.25+0.2}{0.25} & \frac{0.1+0.2}{0.1} & \frac{0.1+0.2}{0.1} \end{pmatrix}^T = \begin{pmatrix} 1.8 & 3 & 3 \end{pmatrix}^T \tag{5.0.2}$$

и его норма: $\|\mathbf{w_{greedy}}\|_1 = 7.8$

Найденное симплекс-методом решение оптимальнее жадного практически в два раза (в смысле l_1-нормы).

Список литературы

[1] А. Н. Баженов. *Лекции по интервальному анализу*. 2020.

[2] С. П. Шарый. *Конечномерный интервальный анализ*. Издательство “XYZ”, 2020. URL: http://www.nsc.ru/interval/Library/InteBooks/SharyBook.pdf.

6. Приложения

1. Репозиторий с кодом программы и кодом отчёта:

   https://github.com/kystyn/interval

2. `tolsolvty` Максима Смольского для Python:

   https://github.com/MaximSmolskiy/tolsolvty