Тема 3. Интервальные векторы и матрицы.

А.Н. Баженов

ФТИ им. А.Ф.Иоффе

a_bazhenov@inbox.ru

26.01.2023

ПЛАН

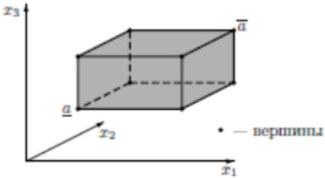
ПЛАН

- Основные определения и факты
- Нормы интервальных векторов и матриц
- Метрика и топология в интервальных пространствах
- Неособенные интервальные матрицы
- Сильно неособенные интервальные матрицы
- Обратные интервальные матрицы
- **Задание 1**
- Специальные матрицы

Интервальный вектор. Брус.

Интервальный вектор — упорядоченный кортеж из интервалов, расположенный вертикально (вектор-столбец) или горизонтально (вектор-строка).
Интервальные векторы из $\mathbb{IR}^n$ являются прямыми произведениями интервалов вещественной оси, а их геометрическим образом служат прямоугольные параллелепипеды в пространстве $\mathbb{R}^n$ с рёбрами, параллельными координатным осям.

Краткое название — **брусы**.



Интервальная матрица

Интервальная матрица — прямоугольная таблица, составленная из интервалов $\boldsymbol{a}_{ij}$:

$$\boldsymbol{A} = (\boldsymbol{a})_{ij}$$

.

Интервальные векторы отождествляются с интервальными матрицами размера $1 \times n$ или $n \times 1$.

Обозначения

Если $\boldsymbol{a} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_n)$, обозначают $\underline{\boldsymbol{a}} = (\underline{\boldsymbol{a}}_1, \underline{\boldsymbol{a}}_2, \dots, \underline{\boldsymbol{a}}_n)$ и $\overline{\boldsymbol{a}} = (\overline{\boldsymbol{a}}_1, \overline{\boldsymbol{a}}_2, \dots, \overline{\boldsymbol{a}}_n)$.

Например, вектор на плоскости.

$$\boldsymbol{a} = ([1, 2], [2, 4])$$

Аналогично, для интервальной матрицы $\boldsymbol{A} = \boldsymbol{a}_{ij}$ определяют точечные матрицы $\underline{\boldsymbol{A}} = (\underline{\boldsymbol{a}}_{ij})$ и $\overline{\boldsymbol{A}} = (\overline{\boldsymbol{a}}_{ij})$, образованные соответствующими точечными элементами.

Например, матрица растяжения по координате x на плоскости в $[1, 2]$ раза.

$$\boldsymbol{A} = \begin{pmatrix} [1, 2] & 0 \\ 0 & 1 \end{pmatrix}$$

Вершины интервального вектора и матрицы

Определение. **Вершинами интервального вектора** $\boldsymbol{a}$ из $\mathbb{IR}$ будем называть точечные n-векторы, i-ая компонента которых равна $\underline{\boldsymbol{a}}_{ij}$ или $\overline{\boldsymbol{a}}_{ij}$. Множество вершин интервального вектора обозначаем как

$$\texttt{vert}\ \boldsymbol{a} := \{a \in \mathbb{IR} \mid a_i \in \{\underline{\boldsymbol{a}}_i, \overline{\boldsymbol{a}_i}\}, i = 1, 2, \dots, n.\}$$

Определение. **Вершинами интервальной матрицы** $A = (a_{ij}$ из $\mathbb{IR}^{m\times n}$ назовём точечные $m \times n$-матрицы, ij-ым элементом которых является $\underline{\boldsymbol{a}}_{ij}$ или $\overline{\boldsymbol{a}}_{ij}$. Множество вершин интервальной матрицы обозначаем как

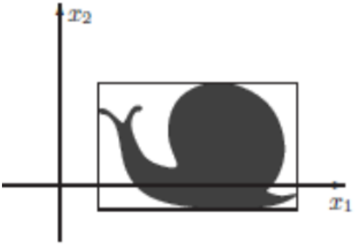
$$\texttt{vert}\ \boldsymbol{A} := \{A \in \mathbb{IR}^{m\times n}\ |A = (a_{ij})\ a_{ij} \in \{\underline{\boldsymbol{a}}_{ij}, \overline{\boldsymbol{a}}_{ij}\}\}$$

Обозначения

Определение. Интервальная оболочка множества S — это пересечение всех интервальных векторов (матриц), содержащих S:

$$\Box S = \bigcap \{\boldsymbol{a} \in \mathbb{IR} \mid \boldsymbol{a} \supseteq S\}$$

Пример. Интервальная облочка векторов на плоскости.



Интервальная оболочка

В некоторых ситуациях нужно не всё множество интервалов или интервальных векторов, а только лишь те из них, что лежат в заданной области рассмотрения.

Определение. Пусть D — некоторое подмножество пространства $\mathbb{IR}^n$. Через $\mathbb{ID}$ обозначают множество всех брусов $\boldsymbol{a} \in \mathbb{IR}$, содержащихся в D, т.е. таких, что $\boldsymbol{a} \subseteq D$.

Действия над матрицами

Сложение и умножение интервальных матриц определяются как естественные поэлементные операции.

«Эффект обёртывания» (ниже) может приводить к неожиданным следствиям в силу того, что умножение матриц включает умножения и сложения их элементов.
В целом это приводит к условию:

$$\mathbf{A} \star \mathbf{B} \;=\; \square\{\, A \star B \mid A \in \mathbf{A},\, B \in \mathbf{B} \,\}.$$

Следует иметь в виду, что интервальные векторы не образуют линейного пространства в привычном смысле. Этому мешает отсутствие дистрибутивности в интервальных арифметиках.

Предложение. Для любых интервальных матриц $\mathbf{A} = (\mathbf{a}_{ij}), \mathbf{B} = (\mathbf{b}_{ij}) \in \mathbb{IR}^{m\times n}$ множество $\{A \pm B \mid A \in \mathbf{A}, B \in \mathbf{B}\}$ совпадает с интервальной матрицей $\mathbf{C} = (\mathbf{c}_{ij}) \in \mathbb{IR}^{m\times n}$ такой что

$$\mathbf{c}_{ij} = \mathbf{a}_{ij} \pm \mathbf{b}_{ij}.$$

Для любых интервальных матриц $\mathbf{A} = (\mathbf{a}_{ij}) \in \mathbb{IR}^{m\times l}, \mathbf{B} = (\mathbf{b}_{ij}) \in \mathbb{IR}^{l\times n}$ множество $\square\{AB \mid A \in \mathbf{A}, B \in \mathbf{B}\}$ совпадает с интервальной матрицей $\mathbf{C} = (\mathbf{c}_{ij}) \in \mathbb{IR}^{m\times n}$, такой что

$$\mathbf{c}_{ij} = \sum_{k=1}^{l} \mathbf{a}_{ik}\mathbf{b}_{kj}.$$

Умножение интервальной матрицы на точечный вектор

Существует частный случай интервального матричного умножения, при котором его результат совпадает с множеством всевозможных точечных произведений «по представителям» —
умножение интервальной матрицы на точечный вектор:

для любых $\boldsymbol{A} \in \mathbb{IR}^n$, $b \in \mathbb{R}^n$

$$\boldsymbol{A}b = \{Ab \mid A \in \boldsymbol{A}\}.$$

Интервальная матричная операция умножения

Результатом интервальной матричной операции умножения

$$\sum_{k=1}^{l} \boldsymbol{a}_{ik}\boldsymbol{b}_{kj}.$$

является интервальная оболочка множества произведений представителей матриц-сомножителей, то есть внешняя оценка этого множества.

В эту матрицу входят и «лишние матрицы», которые однажды появившись в интервальной оценке, далее воспринимаются уже как её неотъемлемая часть, огрубляя окончательный ответ.

Эффект обёртывания

Рассмотрим процесс вращения со сжатием. Такая задача может возникать при решении дифференциального уравнения

$$\frac{dx}{dt} = Ax.$$

Исходно имеем брус со стороной ε, расположенный в точке $(1, 0)$.

$$\begin{aligned} \boldsymbol{x}^{(0)} &\leftarrow ([1-\varepsilon, 1+\varepsilon], [-\varepsilon, \varepsilon])^\top \\ \boldsymbol{x}^{(k+1)} &\leftarrow \tfrac{1}{1.15} R\boldsymbol{x}^{(k)}, \quad k = 0, 1, \dots \end{aligned}$$

С матрицей

$$R = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}$$

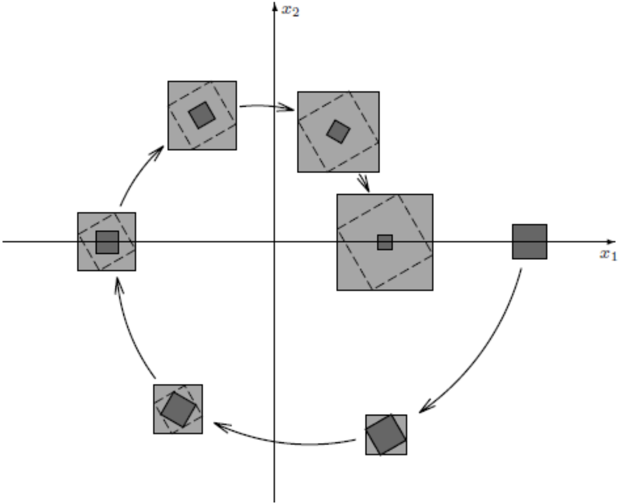
Эффект обёртывания

Итерации расходятся, их результатом являются неограниченно увеличивающиеся в размерах брусы, вращающиеся вокруг начала координат и постепенно его захватывающие. На рисунке они выделены светло-серой закраской.

Взятие интервальной оболочки точечных результатов при интервальном матрично-векторном умножении приводит к тому, что радиус бруса $x^{(k)}$ увеличивается примерно в 1.366 раза на каждом шаге.

Эффект обёртывания

процесс вращения со сжатием



Отсутствие эффекта обёртывания

В частном случае процесс вращения со сжатием с матрицей поворота на 90 градусов

$$R_{90} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

не будет приводить к возникновению эффекта обёртывания.

В этом случае квадратик проецируется строго на оси координат.

Эффект обёртывания и рекуррентные вычисления

Эффект обёртывания — паразитное увеличение оценивающего множества в сравнении с множеством идеальных математических результатов операции, выполненных «по представителям», возникающее вследствие несовпадения его формы с формой оценивающих множеств (непараллельность осям координат в 2- и 3-мерных случаях).

Эффект обёртывания особенно сильно проявляется в итерационных процессах либо рекуррентных вычислениях, где последовательные (пошаговые) замены множества решений на более простые оценивающие множества происходят многократно.

Эффект обёртывания — шаговый двигатель

Шаговый двигатель — это двигатель постоянного тока, ротор которого совершает дискретные перемещения при последовательном приложении напряжения к обмоткам статора.

Рассмотрим часто встречающееся на практике задание — поворот на заданный угол с последующим возвращением в исходное положение.

Пусть, например, эти действия выполняет шаговый двигатель `Nema` 17 (FL42STH), который имеет угловой шаг $(1.8 \pm 0.09)^\circ$ и диаметр вала 22 мм. Такая величина углового шага означает, что оборот на 360° градусов шаговый двигатель должен совершить за 200 шагов.

Эффект обёртывания — шаговый двигатель

Представим рассматриваемую задачу в математических терминах. Вращение объекта $x \in \mathbb{R}^2$ на плоскости на угол φ можно описать, используя матрицу поворота $R(\varphi)$

$$R(\varphi) = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}. \tag{1}$$

В таком случае выполнение задания описывается последовательностью

$$T = R(\varphi) \cdot R(-\varphi) = I, \tag{2}$$

где $R(\varphi)$ определяется выражением (1), I — единичная матрица.

Эффект обёртывания — шаговый двигатель

Однако на практике идеальный случай точного возврата в исходное положение (2) не реализуется. Докажем это утверждение, используя классическую интервальную арифметику $\mathbb{IR}$.

Будем считать, что центр координат находится в центре вала шагового двигателя. При повороте вала мы будем отслеживать изменение положения самой верхней точки вала, начальное положение которой описывается вектором $x_0 = (0, 11)^\top$.

Угол поворота будем считать равным $\varphi = \pi/3$, а для учёта погрешности углового шага представим этот угол в виде интервала $\boldsymbol{\varphi}$ с радиусом $\mathrm{rad}\,\boldsymbol{\varphi} = 1.6 \cdot 10^{-3}$ рад.

Эффект обёртывания — шаговый двигатель

В случае интервальных матриц $\boldsymbol{R}(\pm\varphi)$ (1) матрица $\boldsymbol{T}$ в формуле (2) равна

$$\boldsymbol{T} = \begin{pmatrix} [0.9972, 1.0028] & [-0.0032, 0.0032] \\ [-0.0032, 0.0032] & [0.9972, 1.0028] \end{pmatrix}. \tag{3}$$

Определим положение верхней точки вала после одного вращения на угол φ и одного вращения на угол $-\varphi$:

$$\boldsymbol{x}_1 = \boldsymbol{T} \cdot x_0 = \begin{pmatrix} [-0.0352, 0.0352] \\ [10.969, 11.031] \end{pmatrix}. \tag{4}$$

Кажется, что разница между начальным x_0 и конечным положениями $\boldsymbol{x}_1$ верхней точки вала после вращений крайне мала.

Эффект обёртывания — шаговый двигатель

Но после выполнения ста вращений на угол φ и ста обратных вращений на угол $-\varphi$, конечное положение точки будет описывать уже интервальный вектор

$$\boldsymbol{x}_{100} = \begin{pmatrix} [-4.7679, 4.7679] \\ [7.6773, 15.302] \end{pmatrix}, \tag{5}$$

и внешняя оценка позиционирования (5) составит более 5 мм!
Если мы продолжим вращать вал шагового двигателя в противоположные стороны, то ширина интервального вектора $\boldsymbol{x}$ будет неуклонно расти, что иллюстрирует «эффект обертывания» и подтверждает невозможность точного возвращения вала в исходное положение.

Монотонность по включению

Как и в одномерном случае, имеет место **монотонность по включению**:
для любых интервальных матриц $\boldsymbol{A}, \boldsymbol{A}', \boldsymbol{B}, \boldsymbol{B}'$ одинакового размера и любых операций $\star \subseteq \{+\,,-\,,\,\cdot\,,/\,\}$

$$\boldsymbol{A} \subseteq \boldsymbol{A}', \boldsymbol{B} \subseteq \boldsymbol{B}' \Longrightarrow \boldsymbol{A} \star \boldsymbol{B} \subseteq \boldsymbol{A}' \star \boldsymbol{B}'$$

Свойство монотонности по включению отражает тот факт, что расширение областей определения объектов неизбежно расширяет и область на которую отображаются результаты арифметических операций над объектами.

Действия над матрицами

Для любых интервальных матриц $\mathbf{A}$ и $\mathbf{B}$, одинакового размера справедливы:

$$\begin{aligned}(\mathbf{A}+\mathbf{B})+\mathbf{C}=\mathbf{A}+(\mathbf{B}+\mathbf{C}) &\quad \text{— ассоциативность сложения,}\\ \mathbf{A}+\mathbf{B}=\mathbf{B}+\mathbf{A} &\quad \text{— коммутативность сложения.}\end{aligned}$$

Для интервального матричного умножения нет еще и (кроме обычной коммутативности и интервальной дистрибутивности по сложению) ассоциативности.

Отсутствие ассоциативности интервального матричного умножения

Пример. Дано

$$A = \begin{pmatrix} 1, 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \boldsymbol{C} = \begin{pmatrix} [-1, 1] \end{pmatrix}$$

Варианты вычисления:

$$\begin{aligned} (AB)\boldsymbol{C} = 0 \cdot [-1, 1] \quad &= \quad 0 \\ A(B\boldsymbol{C}) = \begin{pmatrix} 1, 1 \end{pmatrix} \cdot \begin{pmatrix} [-1, 1] \\ [-1, 1] \end{pmatrix} \quad &= \quad [-2, 2] \end{aligned}$$

$$\boxed{(AB)\boldsymbol{C} \neq A(B\boldsymbol{C})}$$

Отсутствие ассоциативности интервального матричного умножения

Невозможность формального решения уравнения

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

как

$$b = \mathbf{A}^{-1}x$$

так как:

$$\mathbf{A}^{-1}(\mathbf{A}x) \neq \left(\mathbf{A}^{-1}\mathbf{A}\right)x$$

Нормы интервальных матриц и векторов

Норма линейного пространства – обобщение на многомерный случай понятия абсолютной величины числа и формализует такие интуитивно понятные свойства как «длина» вектора.

Нормой интервального вектора *a* называют вещественную величину $\|a\|$, удовлетворяющую аксиомам:

$$\|\mathbf{a}\| \geq 0, \text{причем} \|\mathbf{a}\| = 0 \Longleftrightarrow \mathbf{a} = 0 \qquad \text{— неотрицательность,}$$

$$\|\alpha \cdot \mathbf{a}\| = |\alpha| \cdot \|\mathbf{a}\| \qquad \text{— абсолютная однородность,}$$

$$\|\mathbf{a} + \mathbf{b}\| \leq \|\mathbf{a}\| + \|\mathbf{b}\| \qquad \text{— «неравенство треугольника».}$$

Нормы интервальных векторов

Наиболее популярные *векторные нормы*:

$$\begin{aligned}
\|\mathbf{a}\|_1 &= |\mathbf{a}_1| + |\mathbf{a}_2| + \ldots + |\mathbf{a}_n|, \\
\|\mathbf{a}\|_2 &= \left(|\mathbf{a}_1|^2 + |\mathbf{a}_2|^2 + \ldots + |\mathbf{a}_n|^2\right)^{\frac{1}{2}}, \\
\|\mathbf{a}\|_\infty &= \max_{1 \le i \le n} |\mathbf{a}_i|.
\end{aligned}$$

Нормы интервальных матриц

Нормой интервальной матрицы $\boldsymbol{A}$ называют вещественную величину $\|\boldsymbol{A}\|$, удовлетворяющую аксиомам:

$$\|\boldsymbol{A}\| \geq 0, \text{причем} \|\boldsymbol{A}\| = 0 \Longleftrightarrow \boldsymbol{A} = 0 \qquad \text{— неотрицательность,}$$

$$\|\alpha \cdot \boldsymbol{A}\| = |\alpha| \cdot \|\boldsymbol{A}\| \qquad \text{— абсолютная однородность,}$$

$$\|\boldsymbol{A} + \boldsymbol{B}\| \leq \|\boldsymbol{A}\| + \|\boldsymbol{B}\| \qquad \text{— «неравенство треугольника»,}$$

$$\|\boldsymbol{A} \cdot \boldsymbol{B}\| \leq \|\boldsymbol{A}\| \cdot \|\boldsymbol{B}\| \qquad \text{— субмультипликативность,}$$

$$\text{pro}A \subseteq \text{pro}B \Longleftrightarrow \|\boldsymbol{A}\| \leq \|\boldsymbol{B}\| \qquad \text{— монотонность по включению.}$$

Нормы интервальных матриц и векторов

Четвертая аксиома не имеет места для векторов, поскольку их умножение не определено. Если понимать ее в широком смысле, то ее можно использовать для произведения матрицы на вектор и согласования векторных и матричных норм.

$$\|\boldsymbol{A} \cdot \boldsymbol{B}\| \leq \|\boldsymbol{A}\| \cdot \|\boldsymbol{B}\| \quad \text{— субмультипликативность,}$$

$$\mathrm{pro}A \subseteq \mathrm{pro}B \Longleftrightarrow \|\boldsymbol{A}\| \leq \|\boldsymbol{B}\| \quad \text{— монотонность по включению.}$$

Пятая аксиома относится к т.н. «полной» интервальной арифметике.

Дуализация в арифметике Каухера

Правильные и неправильные интервалы, две половинки $\mathbb{KR}$, переходят друг в друга в результате **отображения дуализации** $\mathtt{dual} : \mathbb{KR} \to \mathbb{KR}$, меняющего местами (переворачивающего) концы интервала, т. е. такого что

$$\mathtt{dual}\ \boldsymbol{a} := [\overline{\boldsymbol{a}}, \underline{\boldsymbol{a}}].$$

Правильной проекцией интервала $\boldsymbol{a}$ называется величина

$$\mathrm{pro}\ \boldsymbol{a} := \begin{cases} \boldsymbol{a}, & \text{если } \boldsymbol{a} \text{ — правильный} \\ \text{dual } \boldsymbol{a}, & \text{если иначе.} \end{cases}$$

Согласованные нормы интервальных матриц и векторов

Нормы интервальных векторов и матриц согласованы друг с другом, если для любых матриц $\boldsymbol{A}$ и векторов $\boldsymbol{b}$, для которых определено произведение $\boldsymbol{A} \cdot \boldsymbol{b}$, имеет место неравенство:

$$\|\boldsymbol{A} \cdot \boldsymbol{b}\| \leq \|\boldsymbol{A}\| \cdot \|\boldsymbol{b}\|.$$

Метрика и топология в интервальных пространствах.

Аналогично определению расстояния между векторами в линейных пространствах расстояние между интервальными векторами можно ввести как норму вектора их разности:

$$\operatorname{dist}(\boldsymbol{a}, \boldsymbol{b}) = \|\boldsymbol{a} - \boldsymbol{b}\|.$$

На пространстве интервальных матриц аналогично можно ввести:

$$\operatorname{dist}(\boldsymbol{A}, \boldsymbol{B}) = \|\boldsymbol{A} - \boldsymbol{B}\|.$$

Свойства матриц

Свойства середины

$$\begin{aligned} \mathrm{mid}\,(\boldsymbol{A} \pm \boldsymbol{B}) &= \mathrm{mid}\,(\boldsymbol{A}) + \mathrm{mid}\,(\boldsymbol{B}) \\ \mathrm{mid}\,(A \cdot \boldsymbol{B}) &= A \cdot \mathrm{mid}\,(\boldsymbol{B}) \\ \mathrm{mid}\,(\boldsymbol{A} \cdot B) &= \mathrm{mid}\,(\boldsymbol{A}) \cdot B \end{aligned}$$

Свойства радиуса

$$\begin{aligned} \mathrm{rad}\,(\boldsymbol{A} \pm \boldsymbol{B}) &= \mathrm{rad}\,(\boldsymbol{A}) + \mathrm{rad}\,(\boldsymbol{B}) \\ \mathrm{rad}\,(A \cdot \boldsymbol{B}) &= |A| \cdot \mathrm{rad}\,(\boldsymbol{B}) \\ \mathrm{rad}\,(\boldsymbol{A} \cdot B) &= \mathrm{rad}\,(\boldsymbol{A}) \cdot |B| \end{aligned}$$

Оценка радиуса произведения матриц

$$L \le \mathrm{rad}\,(\boldsymbol{AB}) \le R,$$

где

$$\begin{aligned} L &= \max\{\mathrm{rad}\,(\boldsymbol{A}) \cdot |\mathrm{B}|, |A| \cdot \mathrm{rad}\ \boldsymbol{B}\} \\ R &= \mathrm{rad}\,(\boldsymbol{A}) \cdot |\mathrm{B}| + |A| \cdot \mathrm{rad}\ \boldsymbol{B}. \end{aligned}$$

Оценка радиуса произведения матриц

$$\mathrm{rad}\,(\boldsymbol{A}) \cdot |\mathrm{A}| \le \mathrm{rad}\,(\boldsymbol{AA}) \le \mathrm{rad}\,(\boldsymbol{A}) \cdot |\mathrm{A}| + |\mathrm{mid}\,(\boldsymbol{A})| \cdot \mathrm{rad}\,(\boldsymbol{A})$$

$$\boldsymbol{A} = \boldsymbol{B}, \quad \boldsymbol{A} = \begin{pmatrix} [1,2] & 0 \\ 0 & [1,2] \end{pmatrix}, \quad |\boldsymbol{A}| = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

$$\mathrm{mid}\,\boldsymbol{A} = \begin{pmatrix} 3/2 & 0 \\ 0 & 3/2 \end{pmatrix}, \quad \mathrm{rad}\,\boldsymbol{A} = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}$$

$$\boldsymbol{AA} = \begin{pmatrix} [1,4] & 0 \\ 0 & [1,4] \end{pmatrix}, \quad \boxed{\mathrm{rad}\,(\boldsymbol{AA}) = \begin{pmatrix} 3/2 & 0 \\ 0 & 3/2 \end{pmatrix}}$$

Оценка радиуса произведения матриц

$$\mathrm{rad}\,(\boldsymbol{A}) \cdot |\mathrm{A}| \le \mathrm{rad}\,(\boldsymbol{AA}) \le \mathrm{rad}\,(\boldsymbol{A}) \cdot |\mathrm{A}| + |\mathrm{mid}\,(\boldsymbol{A})| \cdot \mathrm{rad}\,(\boldsymbol{A})$$

$$\boxed{L \le \mathrm{rad}\,(\boldsymbol{AA}) \le R}$$

$$L = \mathrm{rad}\,(\boldsymbol{A}) \cdot |\boldsymbol{A}| = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$R = L + 9/4 \cdot \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix} = \begin{pmatrix} 17/8 & 0 \\ 0 & 17/8 \end{pmatrix}$$

$$\boxed{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \le \begin{pmatrix} 3/2 & 0 \\ 0 & 3/2 \end{pmatrix} \le \begin{pmatrix} 17/8 & 0 \\ 0 & 17/8 \end{pmatrix}}$$

Произведение интервальной матрицы на точечный вектор

Рассмотрим произведение интервальной матрицы на точечный вектор

$$\boldsymbol{A} \cdot x.$$

Имеем

$$\begin{aligned} \operatorname{mid}\left(\boldsymbol{A} \cdot x\right) &= \operatorname{mid} \boldsymbol{A} \cdot x \\ \operatorname{rad}\left(\boldsymbol{A} \cdot x\right) &= \operatorname{rad} \boldsymbol{A} \cdot x \end{aligned}$$

Окончательно

$$\boldsymbol{A} \cdot x = [\ \operatorname{mid} \boldsymbol{A} \cdot x - \operatorname{rad} \boldsymbol{A} \cdot |x|,\ \operatorname{mid} \boldsymbol{A} \cdot x + \operatorname{rad} \boldsymbol{A} \cdot |x|\ ].$$

Подчиненные матричные нормы

На интервальную матрицу $\mathbf{A}$ можно смотреть как на оператор $\phi_{\mathbf{A}}$, действующий из пространства $\mathbb{IR}^n$ в пространство $\mathbb{IR}^m$ по правилу

$$\phi_{\mathbf{A}} = \mathbf{A}\mathbf{x}.$$

Он является однородным, и потому для описания его действия напрашивается хорошо известная из линейной алгебры конструкция подчинённой операторной нормы

$$\|\phi_{\mathbf{A}}\| := \sup_{\|\mathbf{x}\|\leq 1} \|\mathbf{A}\mathbf{x}\|$$

Нормы интервальных векторов и матриц согласованы друг с другом, если для любых матриц и векторов, для которых определено произведение имеет место неравенство:

$$\|\mathbf{A}\mathbf{b}\| \leq \|\mathbf{A}\| \cdot \|\mathbf{b}\|$$

Примеры подчиненных матричных норм

$$\|A\|_1 := \max_{1\le j\le n}\left(\sum_{i=1}^{m}|a_{ij}|\right)$$

$$\|A\|_\infty := \max_{1\le i\le m}\left(\sum_{j=1}^{n}|a_{ij}|\right)$$

Сингулярные числа матрицы

Сингулярными числами матрицы называются неотрицательные квадратные корни из собственных чисел матрицы AA^T

Равносильное определение сингулярных чисел — элементы диагональной матрицы Σ в разложении $A = U\Sigma V^T$.

Можно показать, что и для норм интервальных векторов и матриц величина является нормой и подчинена 2-норме интервальных векторов

$$\|\boldsymbol{A}\| := \sigma_{\max}(\boldsymbol{A})$$

Спектральный радиус матрицы

Спектральным радиусом точечной матрицы называется наибольшее из абсолютных значений собственных чисел.

Спектральный радиус не превосходит ни одну из норм.

В частности,
$$\rho(A) \leq \|\boldsymbol{A}\|$$
для всякой $\boldsymbol{A} \in \mathbb{IR}$.

Определение. Матрица $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ называется **разложимой** , если существует разбиение множества $1, 2, ..., n$ первых n натуральных чисел на два непересекающихся подмножества I и J, таких что $a_{ij} = 0$ при $i \in I$ и $j \in J$.
Эквивалентное определение: матрица $A \in \mathbb{R}^{n \times n}$ разложима, если путём перестановок строк и столбцов она может быть приведена к блочно-треугольному виду

$$\begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}$$

с квадратными блоками A_{11} и A_{22}.

Разложимые и неразложимые матрицы

Матрицы, не являющиеся разложимыми, называются **неразложимыми**. Таких же определений разложимости и неразложимости мы будем придерживаться и для интервальных матриц.

Важнейший пример неразложимых матриц — это матрицы, все элементы которых не равны нулю, в частности, положительны.

Основной результат теории **неотрицательных** матриц

Теорема Перрона-Фробениуса.
Если $A \in \mathbb{R}$ — *неразложимая* неотрицательная матрица
(i) A имеет **положительное** собственное значение, которое равно спектральному радиусу $\rho(A)$;
(ii) существует **положительный** собственный вектор, отвечающий собственному значению $\rho(A)$;
(iii) собственное значение $\rho(A)$ имеет алгебраическую кратность 1, т. е. является простым корнем характеристического многочлена для A.

Свойства неотрицательных матриц

Предложение ($\star$). Пусть $A \in \mathbb{R}^{n\times n}$ — неотрицательная матрица, α- – положительное вещественное число. Тогда

$$\rho(\alpha) \leq \alpha \Longleftrightarrow (\exists v \in \mathbb{R}^n)(v > 0 \;\&\; Av \leq \alpha v)$$
$$\rho(\alpha) \geq \alpha \Longleftrightarrow (\exists v \in \mathbb{R}^n)(v > 0 \;\&\; Av \geq \alpha v)$$

где векторные неравенства понимаются покомпонентно.
Из Предложения ($\star$) легко выводится следующее важное свойство спектрального радиуса: для любых матриц $A, B \in \mathbb{R}^{n\times n}$

$$|A| \leq B \Longrightarrow \rho(A) \leq \rho(B).$$

На пространстве интервальных матриц можно ввести метрику

$$
\begin{aligned}
\operatorname{dist}(\mathrm{A},\mathrm{B}) &= \|\mathrm{A}\ominus\mathrm{B}\| \\
\operatorname{dist}_1(\mathrm{A},\mathrm{B}): &= \sum_{i,j}\operatorname{dist}(\mathrm{a}_{ij},\mathrm{b}_{ij}) \\
\operatorname{dist}_2(\mathrm{A},\mathrm{B}): &= \left(\sum_{i=1}^{n}\left(\operatorname{dist}(\mathrm{a}_{ij},\mathrm{b}_{ij})\right)^2\right)^{1/2}
\end{aligned}
$$

Мультиметрика

$$
\operatorname{Dist}(\boldsymbol{A},\boldsymbol{B}) := \begin{pmatrix}
\operatorname{dist}(\boldsymbol{a}_{11},\boldsymbol{b}_{11}) & \cdots & \operatorname{dist}(\boldsymbol{a}_{1n},\boldsymbol{b}_{1n}) \\
\operatorname{dist}(\boldsymbol{a}_{21},\boldsymbol{b}_{21}) & \cdots & \operatorname{dist}(\boldsymbol{a}_{2n},\boldsymbol{b}_{2n}) \\
\vdots & \ddots & \vdots \\
\operatorname{dist}(\boldsymbol{a}_{m1},\boldsymbol{b}_{m1}) & \cdots & \operatorname{dist}(\boldsymbol{a}_{mn},\boldsymbol{b}_{mn})
\end{pmatrix}
$$

Сжимающие отображения

Определение.
Отображение $f : X \longrightarrow X$ метрического пространства X с расстоянием $d : X \times X \to \mathbb{R}_+$ называется **сжимающим** (сжатием), если существует неотрицательная постоянная $\alpha < 1$, такая что для всех $x, y \in X$ имеет место неравенство

$$d(f(x), f(y)) \leq \alpha \cdot d(x, y)$$

Определение.
Отображение $f : X \longrightarrow X$ мультиметрического пространства X с расстоянием $D : X \times X \to \mathbb{R}_+$ называется P-**сжимающим** (сжатием), если существует неотрицательная матрица P со спектральным радиусом $\rho(P) < 1$ такая что для всех $x, y \in X$ имеет место неравенство

$$D(f(x), f(y)) \leq P \cdot D(x, y)$$

Теорема (Банаха о неподвижной точке).
Сжимающее отображение $f : X \to X$ полного метрического пространства X в себя имеет **единственную неподвижную точку**.
Она может быть найдена как предел последовательных приближений

$$x^{(k+1)} \leftarrow f(x^{(k)}), k = 0, 1, 2, ...,$$

при любом начальном приближении $x^{(0)} \in X$.

Теорема (конечномерная теорема Шредера о неподвижной точке).
Пусть отображение $\Phi : X \to X$ является P-сжимающим отображением полного мультиметрического пространства X с мультиметрикой $D : X \times X \to \mathbb{R}_+$. Тогда для любого x^0 последовательность итераций

$$x^{(k+1)} = \Phi\left(x^{(k)}\right), \quad k = 0, 1, 2, \ldots$$

сходится к **единственной неподвижной точке** x^* отображения Φ в X и имеет место оценка

$$D\left(x^{(k)}, x^*\right) \leq (I - P)^{-1} P \cdot D\left(x^{(k)}, x^{(k-1)}\right)$$

Неособенные интервальные матрицы

Определение.
Интервальная матрица $\mathbf{A} \in \mathbb{IR}$ называется **неособенной**, если неособенны все точечные матрицы $A \in \mathbf{A}$.
Интервальная матрица называется **особенной**, если она содержит особенную точечную матрицу.

Неособенные интервальные матрицы — пример

Пример. Рассмотрим 2 матрицы и определим, особенны ли они

$$\boldsymbol{A} = \begin{pmatrix} [0,1] & [2,3] \\ [4,5] & [6,7] \end{pmatrix}$$

$$\boldsymbol{B} = \begin{pmatrix} [1,2] & [3,4] \\ [5,6] & [7,8] \end{pmatrix}$$

$$\begin{aligned} \det \boldsymbol{A} &= [0,1]\cdot[6,7]-[4,5]\cdot[2,3] = [-15,-1] \quad 0 \notin [-15,-1] \\ \det \boldsymbol{B} &= [1,2]\cdot[7,8]-[3,4]\cdot[5,6] = [-17,1] \quad 0 \in [-17,1] \end{aligned}$$

Неособенные интервальные матрицы — наблюдение

$$\begin{aligned} \boldsymbol{B} &= \begin{pmatrix} [1,2] & [3,4] \\ [5,6] & [7,8] \end{pmatrix} \\ &= \boldsymbol{A} + k \cdot \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad k = 1. \end{aligned}$$

добавляя к элементам матрицы (большие) положительные числа, приближаем её к особенной

Интервальный признак Адамара

Определение. Интервальная матрица имеет **диагональное преобладание**, если все содержащиеся в ней точечные матрицы являются диагональное преобладающими.

Теорема (интервальный признак Адамара).
Интервальная матрица с диагональным преобладанием является неособенной.

Критерий Баумана

Теорема (критерий Баумана). Интервальная матрица $\boldsymbol{A}$ неособенна тогда и только тогда, когда
определители всех её крайних матриц имеют одинаковый знак, т. е.

$$(detA') \cdot (detA'') > 0$$

для любых $A', A'' \in \mathtt{vert}\,\boldsymbol{A}$.

Критерий Баумана - доказательство

Доказательство.
Необходимость покажем «от противного», предположив, что для каких-то крайних матриц A' и A'' интервальной матрицы $\boldsymbol{A}$ определители $detA'$ и $detA''$ имеют разный знак.
Соединим в пространстве $\mathbb{R}^{n\times n}$ матрицы A' и A'' отрезком прямой с параметрическим уравнением $(1-t)A' + tA''$, $t \in [0,1]$, и рассмотрим функцию $\psi(t) = det((1-t)A' + tA'')$.
Ясно, что $\psi(t)$ — непрерывная вещественнозначная функция, к тому же на концах интервала $[0,1]$ она принимает значения разных знаков.
Следовательно, в силу теоремы Больцано-Коши внутри $[0,1]$ обязательно найдётся точка ξ, в которой $\psi(\xi) = 0$.
Соответствующая матрица $(1-\xi)A' + \xi A''$ — особенная и лежит в интервальной матрице $\boldsymbol{A}$.

Критерий Баумана - доказательство

Доказательство.
Для доказательства достаточности используется факт линейной алгебры:
если в матрице некоторый столбец является суммой вектор-столбцов, то её определитель равен сумме определителей матриц, у которых на месте этого столбца стоят слагаемые вектор-столбцы.

Особенные интервальные матрицы

Теорема.
Интервальная матрица $\boldsymbol{A} \in \mathbb{IR}^{n\times n}$ особенна тогда и только тогда, когда система неравенств

$$| \, (\mathtt{mid}\boldsymbol{A})x \, | \leq (\mathrm{rad}\ \boldsymbol{A})|x|, \quad x \in \mathbb{R}^n$$

имеет ненулевое решение.
Неформально это звучит таким образом: «радиус матрицы больше её среднего».
Результат является необходимым и достаточным условием особенности интервальных матриц, тесно связанным с *характеризацией Оеттли-Прагера* для объединённого множества решений ИСЛАУ.

Доказательство теоремы 1/3

Доказательство. Необходимость.
Если $\boldsymbol{A}$ содержит особенную матрицу $\widetilde{A}$, то $\widetilde{A}\widetilde{x} = 0$ для некоторого ненулевого вектора $\widetilde{x}$.
Следовательно,

$$\begin{aligned} |(\mathrm{mid}\ \boldsymbol{A})\widetilde{x}| &= \left|(\mathrm{mid}\ \boldsymbol{A} - \widetilde{A})\widetilde{x}\right| \le \left|\widetilde{A} - \mathrm{mid}\ \boldsymbol{A}\right| \widetilde{x} \\ &\le (\mathrm{rad}\ \boldsymbol{A})\,|\widetilde{x}| \end{aligned}$$

поскольку $\left|\widetilde{A} - \mathrm{mid}\ \boldsymbol{A}\right| \le \mathrm{rad}\ \boldsymbol{A}$.
Достаточность. Если неравенство в теореме действительно имеет решение $\widetilde{x} = 0$, то для векторов $y = (y_i)$, $z = (z_j) \in \mathbb{R}^n$, таких что

$$y_i = \begin{cases} \frac{(\mathrm{mid}\ \boldsymbol{A}\cdot\widetilde{x})_i}{(\mathrm{rad}\ \boldsymbol{A}\cdot|\widetilde{x}|)_i}, & \text{если } (\mathrm{rad}\ \boldsymbol{A}\cdot|\widetilde{x}|) \ne 0 \\ 1, & \text{если } (\mathrm{rad}\ \boldsymbol{A}\cdot|\widetilde{x}|) = 0 \end{cases}$$

Доказательство теоремы 2/3

$$z_j = \begin{cases} 1, & \text{если } \widetilde{x}_j \geq 0 \\ -1, & \text{если } \widetilde{x}_j < 0 \end{cases}$$

$i, j = 1, 2, \ldots, n$, рассмотрим матрицу $\widetilde{A}$ с элементами

$$(\mathrm{mid}\ \boldsymbol{A})_{ij} - y_i z_j (\mathrm{rad}\ \boldsymbol{A})_{ij}$$

В матричном виде она может быть представлена как

$$\widetilde{A} = \mathrm{mid}\ \boldsymbol{A} - \mathrm{diag}\{y\} \cdot \mathrm{rad}\ \boldsymbol{A} \cdot \mathrm{diag}\{z\}$$

Доказательство теоремы 3/3

Так как все $|y_i z_j| \leq 1$, то, очевидно, $\widetilde{A}$ принадлежит $\boldsymbol{A}$. В то же время, она особенная, так как её произведение на ненулевой вектор $\widetilde{x}$ зануляется:

$$\begin{aligned}\widetilde{A}\widetilde{x} &= \mathrm{mid}\ \boldsymbol{A}\widetilde{x} - \mathrm{diag}\{y\} \cdot \mathrm{rad}\ \boldsymbol{A} \cdot \mathrm{diag}\{z\}\widetilde{x} \\ &= \mathrm{mid}\ \boldsymbol{A}\widetilde{x} - \mathrm{diag}\{y\} \cdot \mathrm{rad}\ \boldsymbol{A}|\widetilde{x}|\end{aligned}$$

причём i-ая компонента этого вектора должна быть равна разности

$$((\mathrm{mid}\ \boldsymbol{A})\widetilde{x})_i - ((\mathrm{mid}\ \boldsymbol{A})\widetilde{x})_i$$

если $((\mathrm{rad}\ \boldsymbol{A})|\widetilde{x}|)_i \neq 0$, или разности двух нулей по условию теоремы, если $((\mathrm{rad}\ \boldsymbol{A})|\widetilde{x}|)_i = 0$.

Определение. Ортантом пространства $\mathbb{R}^n$ называется множество точек, координаты которых имеют определённые фиксированные знаки, взятое вместе со своей границей.

Теорема.

Интервальная матрица A неособенна тогда и только тогда, когда для каждого ортанта $\mathcal{O}$ в $\mathbb{R}^n$ существует решение системы неравенств

$$|\,(\mathtt{mid}\boldsymbol{A})x\,| > (\mathrm{rad}\ \boldsymbol{A})|x, \quad x \in \mathbb{R}^n$$

удовлетворяющее $(\mathtt{mid}\boldsymbol{A})x \in \mathcal{O}$.

Практическое значение теоремы невелико, так как их применение требует, фактически, решения $2n$ линейных неравенств (в каждом из ортантов пространства). Но совместно с предыдущей, она составляет основу для конструирования более практичных достаточных признаков особенности/неособенности интервальных матриц.

признак Бекка

Теорема (признак Бекка). Пусть интервальная матрица $\boldsymbol{A} \in \mathbb{IR}^{n\times n}$ такова, что ее середина $\operatorname{mid} \boldsymbol{A}$ неособенна и

$$\rho \left(\left| (\operatorname{mid} \boldsymbol{A})^{-1} \right| \cdot \operatorname{rad} \boldsymbol{A} \right) < 1$$

Тогда $\boldsymbol{A}$ неособенна.
Если вычисления обратной средней матрицы $(\operatorname{mid} \boldsymbol{A})^{-1}$ производятся приближенно, неравенства мы, строго говоря, проверить не сможем.

признак Бекка

Теорема. Пусть интервальная матрица$\mathbf{A} \in \mathbb{IR}^{n\times n}$ такова, что её середина $\mathrm{mid}\ \mathbf{A}$ неособенна и

$$\max_{1\le j\le n} \left(\mathrm{rad}\ \mathbf{A} \cdot \left|(\mathrm{mid}\ \mathbf{A})^{-1}\right|\right)_{jj} \ge 1$$

Тогда $\mathbf{A}$ — особенная.

Теоремы Рона и Рекса

Теорема. Если существует матрица $R \in \mathbb{R}^{n\times n}$ такая что

$$\rho\left(|I - R \cdot \mathrm{mid}\, \boldsymbol{A}| + |R| \cdot \mathrm{rad}\, \boldsymbol{A}\right) < 1,$$

то интервальная матрица $\boldsymbol{A}$ — неособенная.
Теорема. Если существует матрица $R \in \mathbb{R}^{n\times n}$, такая что покомпонентное неравенство

$$\left(I + |I - \mathrm{mid}\, \boldsymbol{A} \cdot R|\right)_{:j} < \left(\mathrm{rad}\, \boldsymbol{A} \cdot |R|\right)_{:j},$$

выполнено для некоторого $j \in 1, 2, ..., n$, то интервальная матрица $\boldsymbol{A}$ — особенная.

Вычисление спектрального радиуса

Для вычисления спектрального радиуса можно с успехом применять степенные итерации, поскольку матрицы

$$(\mathrm{mid}\,\boldsymbol{A})^{-1} \cdot \mathrm{rad}\,\boldsymbol{A}$$
$$|I - R \cdot \mathrm{mid}\,\boldsymbol{A}| + |R| \cdot \mathrm{rad}\,\boldsymbol{A}$$

— неотрицательные, и потому в силу теории Перрона-Фробениуса они обладают неотрицательными доминирующими собственными значениями, которые превосходят по модулю все их остальные собственные значения.

Признак Румпа

Если средняя матрица $\mathrm{mid}\,\boldsymbol{A}$ близка к особенной, то тесты, использующие обратную к ней матрицу, могут оказаться неустойчивыми. Румп предложил условия неособенности, не опирающиеся на нахождение обратной средней, хотя это и достигается ценой вычисления информации о сингулярных числах матриц середин и радиусов для исследуемой матрицы.

Теорема. (Признак Румпа) Если для интервальной матрицы $\boldsymbol{A} \in \mathbb{IR}^{n\times n}$ имеет место

$$\sigma_{\max}(\mathrm{rad}\,\boldsymbol{A}) < \sigma_{\min}(\mathrm{mid}\,\boldsymbol{A})$$

Тогда $\boldsymbol{A}$ неособенна.

Неособенные интервальные матрицы

Интервальная матрица неособенная.

$$\mathbf{A} = \begin{pmatrix} [0,1] & [2,3] \\ [4,5] & [6,7] \end{pmatrix}, \quad \mathrm{mid}\ \mathbf{A} = \begin{pmatrix} 0.5 & 2.5 \\ 4.5 & 6.5 \end{pmatrix}$$

1) Признак Бека выполнен.

$$\mathrm{rad}\ \mathbf{A} = \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix}, \quad (\mathrm{mid}\,\mathrm{A})^{-1} = \begin{pmatrix} -0.8125 & 0.3125 \\ 0.5625 & -0.0625 \end{pmatrix}$$

$$\rho\left(\left|(\mathrm{mid}\ \mathbf{A})^{-1}\right| \cdot \mathrm{rad}\ \mathbf{A}\right) = 0.91 < 1$$

2) Признак Румпа невыполнен и не позволяет судить об неособенности.

$$\sigma_{max}(\mathrm{rad}\ \mathbf{A}) = 1, \quad \sigma_{min}(\mathrm{mid}\ \mathbf{A}) = 0.9697.$$

Неособенные интервальные матрицы

Интервальная матрица неособенная.

$$\boldsymbol{B} = \begin{pmatrix} 3.2 & [0,2] & [0,2] \\ [0,2] & 3.2 & [0,2] \\ [0,2] & [0,2] & 3.2 \end{pmatrix},$$

$$\mathrm{rad}\ \boldsymbol{B} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \quad \mathrm{mid}\ \boldsymbol{B} = \begin{pmatrix} 3.2 & 1 & 1 \\ 1 & 3.2 & 1 \\ 1 & 1 & 3.2 \end{pmatrix}$$

1) Признак Бека не выполнен и не позволяет судить об неособенности.

$$\rho \left(\left| (\mathrm{mid}\ \boldsymbol{B})^{-1} \right| \cdot \mathrm{rad}\ \boldsymbol{B} \right) = 1.0839 > 1$$

2) Признак Румпа выполнен.

$$\sigma_{\max}(\mathrm{rad}\ \boldsymbol{B}) = 2, \quad \sigma_{\min}(\mathrm{mid}\ \boldsymbol{B}) = 2.2$$

Сильно неособенные интервальные матрицы

Определение. Интервальная матрица **сильно неособенная** (сильно невырожденная, сильно регулярная), если интервальная матрица

$$(\mathrm{mid}\,\boldsymbol{A})^{-1}\boldsymbol{A}$$

существует и неособенная.

Почему в определении сильно неособенной матрицы фигурирует умножение на обратную к средней? Неформально это можно объяснить тем, что при таком выборе матрицы C в произведении CA значение диагонали будет сделано наибольшим возможным относительно внедиагональной части полученной матрицы.

Сильно неособенные интервальные матрицы

Определённый недостаток понятия неособенности интервальной матрицы, состоит в том, что он не позволяет оценить при годность интервальной матрицы для алгебраических манипуляций.

Матрица может быть неособенной, но её произведение на другую неособенную матрицу становится уже особенным.

Матрицы Ноймайера-Райхмана

Матрицы Ноймайера-Райхмана — представление. θ - неотрицательный вещественный параметр.

$$\begin{pmatrix} \theta & [0,2] & \dots & [0,2] \\ [0,2] & \theta & \dots & [0,2] \\ \dots & \dots & \dots & \dots \\ [0,2] & [0,2] & \dots & \theta \end{pmatrix} = \theta \cdot E + [0,2] \cdot \begin{pmatrix} 0 & 1 & \dots & 1 \\ 1 & 0 & \dots & 1 \\ \dots & \dots & \dots & \dots \\ 1 & 1 & \dots & 0 \end{pmatrix}$$

Можно показать, что $n \times n$-матрицы Ноймайера-Райхмана чётного порядка n неособенны при $\theta > n$, а матрицы нечётного порядка n неособенны при $\sqrt{n^2 - 1}$.

Сильно неособенные интервальные матрицы

С другой стороны, если умножить матрицу $\boldsymbol{B}$ слева на обратную к её середине, т. е. на

$$(\mathrm{mid}\ \boldsymbol{B})^{-1} = \begin{pmatrix} 0.4 & -0.1 & -0.1 \\ -0.1 & 0.4 & -0.1 \\ -0.1 & -0.1 & 0.4 \end{pmatrix}$$

то получим интервальную матрицу

$$(\mathrm{mid}\ \boldsymbol{B})^{-1}\boldsymbol{B} = \begin{pmatrix} [0.8, 1.2] & [-0.5, -0.5] & [-0.5, -0.5] \\ [-0.5, -0.5] & [0.8, 1.2] & [-0.5, -0.5] \\ [-0.5, -0.5] & [-0.5, -0.5] & [0.8, 1.2] \end{pmatrix}$$

Сильно неособенные интервальные матрицы

Матрица $(\mathrm{mid}\,\boldsymbol{B})^{-1}\boldsymbol{B}$ уже особенна:
она содержит особенные точечные матрицы.

$$\begin{pmatrix} 0.8 & -0.4 & -0.4 \\ -0.4 & 0.8 & -0.4 \\ -0.4 & -0.4 & 0.8 \end{pmatrix}, \qquad \begin{pmatrix} 1 & -0.5 & -0.5 \\ -0.5 & 1 & -0.5 \\ -0.5 & -0.5 & 1 \end{pmatrix},$$

и ещё много других.

Определение. Для неособенной интервальной матрицы A **обратной интервальной матрицей** называют

$$\mathbf{A}^{-1} := \square\left\{A^{-1} \mid A \in \mathbf{A}\right\}$$

то есть, интервальную оболочку множества всех обратных для точечных матриц, содержащихся в $\mathbf{A}$.

Обратные интервальные матрицы

$$A^{-1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} \frac{d}{ad-bc} & \frac{b}{bc-ad} \\ \frac{c}{bc-ad} & \frac{a}{ad-bc} \end{pmatrix}$$

Если a не обращается в нуль, то

$$\frac{a}{ad - bc} = \frac{1}{d - bc/a}$$

Это выражение содержит по одному вхождению каждой переменной в первой степени и поэтому его область значений совпадает с естественным интервальным расширением

$$\frac{1}{\mathbf{d} - \mathbf{bc}/\mathbf{a}}$$

Обратные интервальные матрицы

$$\mathrm{A} = \begin{pmatrix} [2,4] & [-2,1] \\ [-1,2] & [2,4] \end{pmatrix}$$

Проверим неособенность

$$\det \mathrm{A} = [2,4]\cdot[2,4] - [-2,1]\cdot[-1,2] = [4,16] - [-4,2] = [2,20] \not\ni 0$$

Обратная матрица

$$\boldsymbol{A}^{-1} = \begin{pmatrix} [2,4] & [-2,1] \\ [-1,2] & [2,4] \end{pmatrix}^{-1} = \begin{pmatrix} [\frac{1}{6},1] & [-\frac{1}{2},1] \\ [-1,\frac{1}{2}] & [\frac{1}{6},1] \end{pmatrix} \supseteq \begin{pmatrix} [\frac{1}{2}] & [\frac{1}{2}] \\ [\frac{1}{2}] & [\frac{1}{2}]. \end{pmatrix}$$

Получается, что обратная интервальная матрица, даже оптимально оценённая, может быть особенной в силу того, что взятие внешней интервальной оценки множества всех обратных матриц вовлекает в неё лишние элементы.

Задача 1

Дана матрица $\boldsymbol{A}$:

$$\mathrm{mid}\ \boldsymbol{A} = \begin{pmatrix} 1 & 1 \\ 1.05 & 0.95 \end{pmatrix}$$

Пусть $\mathrm{rad}\ \boldsymbol{a}_{ij} = \Delta, \quad \forall i, j$.
При каких значениях Δ матрица $\boldsymbol{A}$ будет содержать особенные матрицы?
Иначе, при каких значениях Δ определитель матрицы

$$\boldsymbol{A} = \begin{pmatrix} [1-\Delta, 1+\Delta] & [1-\Delta, 1+\Delta] \\ [1.05-\Delta, 1.05+\Delta] & [0.95-\Delta, 0.95+\Delta] \end{pmatrix}$$

содержит 0?

$$det(\boldsymbol{A}) \ni 0?$$

Можно использовать критерий Баумана.

Итерационный метод Шульца

Итерационный метод Шульца для обращения матриц

$$X^{k+1} \longleftarrow X^k(2I - AX^k) = X^k + X^k(I - AX^k)$$

Это применение метода Ньютона для решения системы уравнений

$$X^{-1} - A = 0$$

Простая интервализация метод Шульца сходится плохо.
Коррекция. Перепишем

$$\boldsymbol{X}^{k+1} \longleftarrow \mathrm{mid}\, \boldsymbol{X}^k(2I - A \cdot \mathrm{mid}\, \boldsymbol{X}^k)$$

Выбор начального приближения

Проиллюстрируем выбор начального приближения на примере точечной матрицы.

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \\ 0 & 2 & 2 \end{pmatrix}$$

«Угадаем» приближённое решение: на диагонали — обратные элементы, вне диагонали — обратные с противоположным знаком.

$$X^{(0)} = \begin{pmatrix} 1 & -0.5 & -1 \\ 0 & 1 & 0 \\ 0 & -0.5 & 0.5 \end{pmatrix}, \quad \lambda(X^{(0)}) = \begin{pmatrix} 1 \\ 0.5 \\ 1 \end{pmatrix}, \quad \rho(X^{(0)}) = 1.$$

Итерации дают

$$X^{(1)} = \begin{pmatrix} 1 & -0.5 & -0.5 \\ 0 & 1 & 0 \\ 0 & -1 & 0.5 \end{pmatrix}, \quad X^{(2)} = \begin{pmatrix} 1 & -0.5 & -1 \\ 0 & 1 & 0 \\ 0 & -1 & 0.5 \end{pmatrix}.$$

$$\boldsymbol{A} = \begin{pmatrix} 2 & -1 \\ [-5.000, -4.950] & 3 \end{pmatrix}, \tag{6}$$

обратной матрицей для которой является

$$\boldsymbol{A}^{-1} = \begin{pmatrix} [2.857, 3.000] & [0.952, 1.000] \\ [4.714, 5.000] & [1.905, 2.000] \end{pmatrix}. \tag{7}$$

Найдем обратную матрицу для $\boldsymbol{A}$, используя интервальный метод Шульца. Выберем интервальную матрицу, которую в дальнейшем будем использовать в качестве начального приближения для построения итерационного процесса:

$$\boldsymbol{X}^{(0)} = \begin{pmatrix} [2.500, 3.500] & [0.500, 1.500] \\ [3.500, 5.500] & [1.000, 2.500] \end{pmatrix}. \tag{8}$$

В этом случае

$$\mathbf{X}^{(1)} = \begin{pmatrix} [1.9250, 4.0000] & [0.4750, 1.5000] \\ [3.1000, 6.5000] & [1.0750, 2.7500] \end{pmatrix},$$
$$\mathbf{X}^{(2)} = \begin{pmatrix} [2.5880, 3.3407] & [0.8090, 1.1600] \\ [4.2700, 5.5400] & [1.6700, 2.2600] \end{pmatrix},$$
$$\vdots$$
$$\mathbf{X}^{(6)} = \begin{pmatrix} [2.8536, 3.0000] & [0.9510, 1.0000] \\ [4.7073, 5.0000] & [1.90240, 2.0000] \end{pmatrix}.$$

Результаты седьмой и последующих итераций уже не отличаются от интервальной матрицы $\mathbf{X}^{(6)}$ (в третьем знаке после запятой).

Итерационный метод Шульца

Сравним найденную интервальную матрицу $\boldsymbol{X}^{(7)}$ с обратной матрицей $\boldsymbol{A}^{-1}$ (7):

$$\boldsymbol{A}^{-1} = \begin{pmatrix} [2.857, 3.000] & [0.952, 1.000] \\ [4.714, 5.000] & [1.905, 2.000] \end{pmatrix}.$$

Расхождение между этими интервальными матрицами невелико, однако все же присутствует. Если такая точность вычислений достаточна, то интервальная матрица $\boldsymbol{X}^{(6)}$, найденная с помощью интервального метода Шульца, и есть искомая обратная матрица для $\boldsymbol{A}$ (6).

Итерационный метод Шульца

Стоит заметить, что чем «ближе» интервальная матрица $\boldsymbol{A}$ к точечной, тем точнее будет получаемый нами результат ее обращения с помощью интервального метода Шульца. Например, для интервальной матрицы

$$\boldsymbol{A} = \begin{pmatrix} 2 & -1 \\ [-5.000, -4.995] & 3 \end{pmatrix}$$

обратная матрица, найденная на пятой итерации при выборе того же начального приближения $\boldsymbol{X}^{(0)}$ (8), уже не отличается от обратной матрицы $\boldsymbol{A}^{-1}$ (с точностью до третьего знака после запятой):

$$\boldsymbol{X}^{(5)} = \begin{pmatrix} [2.98500, 3.00000] & [0.99500, 1.00000] \\ [4.97000, 5.00000] & [1.99000, 2.00000] \end{pmatrix},$$

$$\boldsymbol{A}^{-1} = \begin{pmatrix} [2.98507, 3.00000] & [0.99503, 1.00000] \\ [4.97015, 5.00000] & [1.99005, 2.00000] \end{pmatrix}.$$

$\boxed{\text{Матрицы специального вида}}$

М-матрицы Wiki

Основной вклад в теорию М-матриц вносят в основном математики и экономисты. М-матрицы используются в математике для установления границ собственных значений и установления критериев сходимости для итерационных методов решения больших разреженных систем линейных уравнений. М-матрицы возникают естественным образом при некоторых дискретизациях дифференциальных операторов, таких как лапласиан, и, как таковые, хорошо изучены в научных вычислениях. М-матрицы встречаются и при изучении решений линейной задачи дополнительности. Проблемы линейной комплементарности возникают в линейном и квадратичном программировании, вычислительной механике и в задаче нахождения точки равновесия в игре с двумя матрицами.

Наконец, М-матрицы встречаются при изучении конечных цепей Маркова в области теории вероятностей и исследования операций, таких как теория массового обслуживания. Между тем, экономисты изучили М-матрицы в связи с валовой взаимозаменяемостью, стабильностью общего равновесия и анализом ввода-вывода Леонтьева в экономических системах. Условие позитивности всех основных несовершеннолетних также известно в экономической литературе как условие Хокинса – Саймона. В технике М-матрицы встречаются также в задачах устойчивости по Ляпунову и управления с обратной связью в теории управления и связаны с матрицей Гурвица. В вычислительной биологии М-матрицы встречаются при изучении динамики популяций.

Определение 1. Матрица $A \in \mathbb{R}^{m \times n}$ является матрицей монотонного типа (**монотонной**), если для любых векторов $x', x'' \in \mathbb{R}^n$ из $Ax' < Ax''$ следует $x' < x''$

Определение 2. Матрица $A \in \mathbb{R}^{m \times n}$ называется **M-матрицей**, если она представима в виде $A = sI - P$, где P-неотрицательная матрица и $s > \rho(P)$.

M-матрицы неособенны. Обратную можно получить матричным рядом Неймана

$$A^{-1} = (sI - P)^{-1} = s^{-1}(I - P/s)^{-1} = s^{-1} \sum_{k=1}^{\infty} (P/s)^k$$

все члены которого неотрицательны. Поэтому неотрицательна и его сумма, и матрица A, таким образом, положительно обратима.

M-матрицы

M-матрицы имеют на главной диагонали строго положительные элементы, а вне главной диагонали — неположительные элементы. При этом главная диагональ доминирует над внедиагональными элементами, в смысле спектрального радиуса.

Предложение. (критерий Фань Цзы) Матрица $A \in \mathbb{R}^{n \times n}$ является M-матрицей тогда и только тогда, когда её внедиагональные элементы неположительны и существует положительный вектор $v \in \mathbb{R}^n$, $v > 0$, такой что $Av > 0$.

Из **Определения 2**.

Следствие. М-матрица не может иметь на главной диагонали нулевые элементы, так как в противном случае соответствующая этой строке компонента произведения Av не получится положительной.

Уравнение Леонтьева

Линейное уравнение межотраслевого баланса (уравнение Леонтьева)

$$x = Px + y$$

где x — вектор объёмов производства по отраслям, y — вектор конечного потребления, а матрица $P \geq 0$ называется матрицей коэффициентов прямых производственных затрат.
Матрица $(I - P)$ системы линейных уравнений

$$(I - P)x = y$$

для определения плана x оказывается M-матрицей.
Говорят, что неотрицательная матрица P — продуктивная, если существует положительный вектор $v > 0$, такой что $(I - P)v \geq 0$.

M-матрицы

Несколько равносильных определений M-матриц.
Именно, матрица $A \in \mathbb{R}^{n\times n}$ является M-матрицей, если она удовлетворяет любому из следующих эквивалентных условий
(i) $A = sI - P$, где P — неотрицательная матрица и $s > \rho(P)$;
(ii) внедиагональные элементы A неположительны и $A^{-1} \geq 0$;
(iii) внедиагональные элементы A неположительны и существует положительный вектор $u > 0$, такой что $Au > 0$;
(iv) внедиагональные элементы A неположительны и её собственные значения имеют положительные вещественные части;
(v) внедиагональные элементы A неположительны и все её главные миноры положительны;
(vi) . . . и т.д.

Подчеркивающие фильтры в обработке изображений — M-матрицы.
Матрица подчеркивающего фильтра имеет структуру

$$A = \begin{pmatrix} \dots & \dots & \dots \\ -\alpha & 1-2\alpha & -\alpha \\ \dots & \dots & \dots \end{pmatrix}, \quad \alpha > 0$$

Обратная к ней матрица — сглаживающий фильтр

$$A^{-1} = \begin{pmatrix} \dots & \dots & \dots \\ \beta & 1-2\beta & \beta \\ \dots & \dots & \dots \end{pmatrix}, \quad \beta \approx \alpha > 0$$

Дифференциальные операторы как M-матрицы

При численном решении дифференциальных уравнений в частных производных эллиптического типа, матрица оператора второй производной имеет структуру

$$D = \begin{pmatrix} \dots & \dots & \dots \\ 1 & -2 & 1 \\ \dots & \dots & \dots \end{pmatrix},$$

Соответственно,

$$-D = \begin{pmatrix} \dots & \dots & \dots \\ -1 & 2 & -1 \\ \dots & \dots & \dots \end{pmatrix},$$

— M-матрица.

Интервальные M-матрицы

Определение. Матрица $A \in \mathbb{IR}^{n \times n}$ называется **интервальной** M-**матрицей**, если каждая вещественная матрица $A \in \boldsymbol{A}$ является M-матрицей.

Предложение. Матрица $\boldsymbol{A} \in \mathbb{IR}^{n \times n}$ является M-матрицей тогда и только тогда, когда её внедиагональные элементы неположительны и существует положительный вектор $v > 0$, такой что $Av > 0$.

Теорема. Интервальная матрица $\boldsymbol{A} \in \mathbb{IR}^{n\times n}$ является M-матрицей тогда и только тогда, когда $\underline{\boldsymbol{A}}$ и $\overline{\boldsymbol{A}}$ — M-матрицы. Всякая интервальная M-матрица неособенна и

$$\boldsymbol{A}^{-1} = [\overline{\boldsymbol{A}}^{-1}, \underline{\boldsymbol{A}}^{-1}].$$

Получается, что интервальная M-матрица ведёт себя при обращении подобно одномерному интервалу вещественной оси, коль скоро

$$\boldsymbol{a}^{-1} = [\frac{1}{\overline{\boldsymbol{a}}}, \frac{1}{\underline{\boldsymbol{a}}}].$$

для любого $\boldsymbol{a} \in \mathbb{IR}$, $\boldsymbol{a} \not\ni 0$. Доказанная теорема допускает обобщение и на все положительно обратимые матрицы.

Теорема Куттлера.

Теорема. Пусть интервальная матрица $\mathbf{A} \in \mathbb{IR}^{n\times n}$ такова, что $\underline{\mathbf{A}}$ и $\overline{\mathbf{A}}$ неособенны и положительно обратимы. Тогда $\mathbf{A}$ также неособенная и

$$\mathbf{A}^{-1} = [\overline{\mathbf{A}}^{-1}, \underline{\mathbf{A}}^{-1}] \geq 0.$$

Теорема использовалась для решения задач математической физики: A Fourth-Order Finite-Difference Approximation for the Fixed Membrane Eigenproblem, 1971.

H-матрицы

Обобщением M-матриц являются так называемые H-матрицы. Они получаются из M-матриц ослаблением обременительного условия на знаки диагональных и внедиагональных элементов

Если из определения M-матриц удалить условие на знаки элементов, то в нём останется лишь идея преобладания, в спектральном смысле, диагонали над остальной частью матрицы. На эксплуатации этого условия и основаны все хорошие свойства H-матриц.

Обратимость H-матриц является гарантией сходимости итерационного метода Гаусса-Зайделя.

H-матрицы применения

Ljiljana Cvetkovic et al. A new subclass of H-matrices. Applied Mathematics and Computation 208 (2009) 206–210

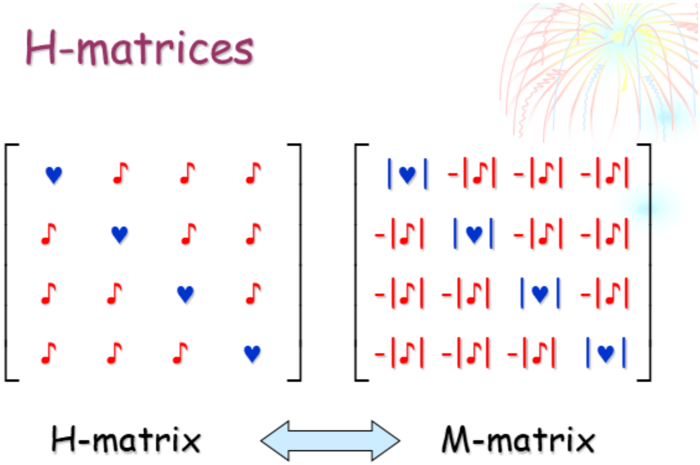
«Класс H-матриц играет важную роль в различных научных дисциплинах, в экономике, например. Однако этот класс можно использовать для получения различных преимуществ в других полях линейной алгебры, таких как оценка детерминанта, оценка корня Перрона, локализация собственных значений, улучшение области сходимости методов релаксации и т.д. »

Может оказаться, что предобуславливание с помощью умножения на Λ приведет к H-матрице . . .

$$\Lambda \cdot \boldsymbol{A}$$

H-матрицы иллюстрация

Ljiljana Cvetkovic. Презентация «H-matrix theory and its applications». 2006



Определение.
Компарантом матрицы $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ будем называть матрицу того же размера, обозначаемую $\langle A \rangle$, такую что

$$ij - \text{й элемент} \langle A \rangle := \begin{cases} |a_{ij}|, \text{ если } i = j \\ -|a_{ij}|, \text{ если } i \neq j \end{cases}$$

Операция взятия компаранта матрицы — это принудительное назначение «нужных» знаков для элементов матрицы, положительных для диагональных элементов и отрицательных для внедиагональных.
В частности, если $A \in \mathbb{R}^{n \times n}$ — вещественная M-матрица, то $\langle A \rangle = A$.

Интервальные H-матрицы

Определение.
Матрицу $A \in \mathbb{R}^{n \times n}$ назовём H-**матрицей**, если её компарант является M-матрицей.

Определение.
Матрица $\boldsymbol{A} \in \mathbb{IR}^{n \times n}$ называется **интервальной** H-**матрицей**, если каждая вещественная матрица $A \in \boldsymbol{A}$ является H-матрицей.

Компарант интервальной матрицы

Определение. Будем говорить, что точечная $n \times n$-матрица $\langle A \rangle$ есть **компарант** интервальной матрицы $\mathbf{A} = \mathbf{a}_{ij} \in \mathbb{IR}^{n \times n}$, если

$$ij - \text{й элемент} \langle \mathbf{A} \rangle := \begin{cases} \langle \mathbf{a}_{ij} \rangle, \text{ если } i = j \\ -\left| \mathbf{a}_{ij} \right|, \text{ если } i \neq j \end{cases}$$

Определение. **Мигнитудой** интервала $\mathbf{a}$ назовём наименьшее из абсолютных значений точек интервала $\mathbf{a}$ — величину

$$\langle \mathbf{a} \rangle := \min\{|a| \mid a \in \mathbf{a}\} = \begin{cases} \min\{|\overline{\mathbf{a}}|, |\underline{\mathbf{a}}|\}, \text{ если } 0 \notin \mathbf{a} \\ 0, \text{ если } 0 \in \mathbf{a} \end{cases}$$

Пример. Если

$$\boldsymbol{A} = \begin{pmatrix} [2,4] & [-2,1] \\ [-1,2] & [2,4] \end{pmatrix}, \text{ то } \langle \boldsymbol{A} \rangle = \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix}$$

Получается, что

$$\langle \boldsymbol{A} \rangle = \min_{A \in \boldsymbol{A}} \langle A \rangle$$

где минимум понимается в поэлементном смысле.
Поэтому всегда $\langle \boldsymbol{A} \rangle = \langle A \rangle$ для некоторой $A \in \boldsymbol{A}$ В частности,

$$\langle \boldsymbol{A} \rangle = \underline{\boldsymbol{A}}$$

для интервальной M-матрицы $\boldsymbol{A}$. В качестве следствия получаем
Предложение. Интервальная матрица $\boldsymbol{A} = \boldsymbol{a}_{ij} \in \mathbb{IR}^{n \times n}$ является H-матрицей тогда и только тогда, когда её компарант $\langle \boldsymbol{A} \rangle$ является M-матрицей.

Интервальные H-матрицы

Важный пример H-матриц — это *неособенные треугольные матрицы* из $\mathbb{IR}^{n\times n}$, верхние или нижние.
Действительно, для такой матрицы $\boldsymbol{A} = (\boldsymbol{aa})_{ij}$ положим

$$\alpha = \max\{\langle \boldsymbol{a}_{11}\rangle, \langle \boldsymbol{a}_{22}\rangle, \dots, \langle \boldsymbol{a}_{nn}\rangle\}$$

Тогда $\langle \boldsymbol{A}\rangle = \alpha I - P$, где P — неотрицательная треугольная матрица, имеющая на главной диагонали элементы с абсолютными значениями, строго меньшими α.

$$|P_{ii}| < \alpha, \quad i = 1, 2, \dots n.$$

Диагональные элементы треугольной матрицы — это её собственные значения, и поэтому спектральный радиус P тоже меньше α.

Интервальные H-матрицы

Теорема. Если $\mathbf{A} \in \mathbb{IR}^{n\times n}$ интервальная H-матрица, то она неособенна и

$$|\mathbf{A}^{-1}| < \langle \mathbf{A} \rangle^{-1},$$

причём это неравенство является точным.
Результат Теоремы позволяет грубо оценивать размеры множеств решений интервальных линейных систем уравнений
Теорема. Матрица $\mathbf{A} \in \mathbb{IR}^{n\times n}$ является H-матрицей тогда и только тогда, когда $\mathrm{mid}\ \mathbf{A}$ — это H-матрица и

$$\rho(\langle \mathrm{mid}\ \mathbf{A} \rangle^{-1} \mathrm{rad}\ \mathbf{A}) < 1,$$

Интервальные Н-матрицы

Спектральный радиус произведения обратной средней матрицы на матрицу радиусов уже встречался нам в признаке Бека неособенности интервальных матриц. Там, правда, присутствовало абсолютное значение обратной средней матрицы, но для интервальных H-матриц мы можем оценить его сверху через обратную для компаранта. Таким образом

Следствие. Всякая интервальная Н-матрица сильно неособенна.

Обратное неверно. Матрица C не является H-матрицей

$$\boldsymbol{C} = \begin{pmatrix} 0 & [1,2] \\ [1,2] & 0 \end{pmatrix},$$

$$(\mathrm{mid}\ \boldsymbol{C})^{-1} = \begin{pmatrix} 0 & \frac{2}{3} \\ \frac{2}{3} & 0 \end{pmatrix}, \quad (\mathrm{mid}\ \boldsymbol{C})^{-1}\boldsymbol{C} = \frac{2}{3}\begin{pmatrix} 0 & [1,2] \\ [1,2] & 0 \end{pmatrix}$$

— неособенная матрица.