МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ

Национальный исследовательский
Нижегородский государственный университет им. Н.И. Лобачевского

Н.Р. Стронгина

КУРС «ЧИСЛЕННЫЕ МЕТОДЫ»

Итерационные методы решения СЛАУ для вычислительно-трудоемких задач

(Модули 10 – 11)

Учебно-методическое пособие

Рекомендовано методической комиссией
Института информационных технологий, математики и механики
для студентов ННГУ, обучающихся по направлению
01.03.02 «Прикладная математика и информатика»

Нижний Новгород
2021

УДК 519.6
ББК 22.19
С-86

С-86 Стронгина Н.Р. Курс «Численные методы»: Итерационные методы решения СЛАУ для вычислительно-трудоемких задач (Модули 10 – 11): Учебно-методическое пособие. – Нижний Новгород: Нижегородский госуниверситет, 2021. – 79 с.

Рецензент: к.ф.-м.н., доцент **А.А. Перов**

Пособие является компонентом учебно-методического комплекса по дисциплине «Численные методы». В нем представлен теоретический аппарат линейной алгебры, необходимый для анализа обусловленности систем линейных алгебраических уравнений (СЛАУ), построения итерационных методов их решения и анализа их скорости сходимости. Рассмотрены примеры методов, построенных на основе спектрального и вариационного подходов. С целью подготовки студентов к самостоятельной программной реализации численных методов решения модельных задач, имеющих большую размерность, рассмотрен наглядный пример СЛАУ малой размерности и приведен пошаговый разбор решения.

Пособие предназначено для студентов университета, обучающихся по направлению подготовки 01.03.02 «Прикладная математика и информатика», а также для преподавателей. Может быть рекомендовано студентам магистратуры, изучающим параллельные численные методы на основе технологий параллельного программирования.

Ответственный за выпуск:
председатель методической комиссии
Института информационных технологий, математики и механики ННГУ
к.ф.-м.н., доцент А.В. Грезина

УДК 519.6
ББК 22.19

СОДЕРЖАНИЕ

Введение……………………………………………………………….......... 4

Модуль 10. Теоретический аппарат для построения и применения итерационных методов решения СЛАУ ……………………………….. 7

10.1. Нормы векторов и матриц, согласованные и подчиненные нормы ………………………………………… 7

10.2. Собственные числа и спектральный радиус ………………….. 9

10.3. Оценка собственных чисел по теореме Гершгорина ………… 10

10.4. Симметричные матрицы, их нормы ………………………… 11

10.5. Норма обратной матрицы ……………………………………… 12

10.6. Симметричные положительно определенные матрицы ……... 13

10.7. Число обусловленности и его влияние на свойства СЛАУ …. 15

10.8. Примеры плохой обусловленности …………………………… 18

10.9. Механизм влияния обусловленности на погрешность решения СЛАУ ……………………………… 19

10.10. Нормализация СЛАУ ……………………………………………. 20

10.11. Скорость сходимости (определение) ………………………… 21

10.12. Презентация модельных задач ……………………………….. 22

Модуль 11. Примеры итерационных методов решения СЛАУ для задач большой размерности ………………………………………………... 23

11.1. Метод простой итерации ……………………………………. 23

11.2. Метод минимальных невязок ……………………………….. 37

11.3. Метод с чебышевским набором параметров (k параметров) … 44

11.4. Метод сопряженных градиентов ……………………………… 58

Модуль 11 – Практикум по темам 11.1-11.4. Решение типовых задач на примере метода простой итерации ……………………………….. 68

Литература …………………………………………………………….. 77

ВВЕДЕНИЕ

Начавшееся в середине XX столетия развитие вычислительной техники открыло качественно новые возможности изучения сложных реальных объектов методами вычислительного эксперимента. Современные высокопроизводительные вычислительные системы открывают возможность решения все более сложных математических задач и соответственно изучения методом вычислительного эксперимента новых, более сложных реальных объектов [18, 19].

Машинный вычислительный эксперимент как новый метод научного исследования предполагает дискретизацию исходной задачи и требует специальной проработки численного алгоритма (корректность, устойчивость, точность, сходимость). Поэтому на современном этапе подготовки специалистов основная цель дисциплины «Численные методы» по направлению «Прикладная математика и информатика» – изучение фундаментальных принципов построения численных алгоритмов, подходов к анализу их свойств, подготовка студентов к разработке и применению эффективных вычислительных комплексов, необходимых для математического моделирования сложных систем.

В Институте информационных технологий, математики и механики ННГУ в системе подготовки бакалавров по указанному выше направлению дисциплина «Численные методы» изучается на 3-м курсе в течение двух семестров. Обучение включает лекции, практические и лабораторные занятия, самостоятельную работу, зачеты и экзамен. Содержание дисциплины соответствует требованиям федеральных государственных образовательных стандартов и обновляется с учетом проблематики научных исследований и технологий программирования. Фундаментальные основы курса соответствует требованиям типовой программы по направлению «Прикладная математика и информатика», разработанной под руководством академика РАН А.А. Самарского [20].

Курс содержит изучение основ машинной арифметики, анализ структуры погрешности, подходы и методы приближенного вычисления функций, численное дифференцирование и интегрирование, численное решение систем линейных алгебраических уравнений, задач на собственные значения, решение нелинейных алгебраических уравнений и систем. Особое внимание уделяется инструментам математического моделирования: методам численного решения задачи Коши и краевых задач для обыкновенных дифференциальных уравнений (ОДУ), решению уравнений в частных производных, а также структуре соответствующих вычислительных комплексов.

В связи с успешным многолетним применением в ННГУ практико-ориентированного подхода и на основе принципа «образование как исследование», вытекающего из положения Гумбольдта «образование на основе исследований» [18], фундаментальный курс «Численные методы» имеет в ННГУ уровневую структуру. С одной стороны, в нем представлены все основные разделы численного анализа. С другой стороны, актуальные приложения требуют одновременного использования разных методов. Поэтому основой курса является системное изучение модельных задач, описывающих свойства реальных объектов различной природы. График освоения разделов построен таким образом, чтобы в течение каждого семестра студенты могли самостоятельно подготовить программную реализацию численного алгоритма для решения модельной задачи, провести вычислительный эксперимент и подготовить отчет.

Нижегородский государственный университет является участником Суперкомпьютерного консорциума университетов России [2]. Студентов 3-го курса, изучающих дисциплину «Численные методы», знакомят с подходами к организации параллельных вычислений. Глубокое изучение этих подходов опирается на тот же комплект модельных задач, но проводится на старших курсах после освоения дисциплин, посвященных технологиям и методам параллельного программирования.

При освоении курса «Численные методы» у студентов должны быть сформированы компетенции разработки и применения программных средств разного уровня сложности. Поэтому требования к программам, подготовленным студентами, реализуют практико-ориентированный подход: программа должна быть написана на алгоритмическом языке высокого уровня; код, реализующий алгоритм, должен быть подготовлен студентом самостоятельно; объектно-ориентированный подход приветствуется; программа и способ работы с ней должны быть пригодны не только для выполнения конкретного расчета, но также для проверки корректной реализации метода и результатов вычислительного эксперимента, и затем для изучения свойств метода и свойств моделируемого объекта.

Ряд заданий выполняются сначала с помощью программы-тренажера, затем – с помощью программы, подготовленной студентом.

Требования самостоятельной программной реализации алгоритма и последующего самостоятельного проведения вычислительного эксперимента предполагают, что при рассмотрении теоретического материала, проведении практических занятий, выполнении заданий в рамках самостоятельной работы необходимо уделить больше внимания анализу понятийного аппарата дисциплины, доказательной базе, рассмотрению «простых» примеров и разбору по шагам решений специально подобранных задач. Решение именно этой учебной задачи поддерживает предлагаемое пособие.

Изучение тематического модуля, представленного в пособии, опирается на дисциплины «Дифференциальные уравнения», «Математический анализ», «Геометрия и алгебра» и «Программирование на ЭВМ» и осуществляется одновременно с изучением дисциплин «Уравнения математической физики» и «Функциональный анализ».

Нумерация разделов пособия соответствует установленной в настоящее время нумерации тематических модулей электронного учебного курса «Численные методы», представленного в системе электронного обучения ННГУ (СЭО ННГУ) на базе платформы Moodle. В период весеннего семестра 2019-20 учебного года и осеннего семестра 2020-21 учебного года дистанционная организация учебного процесса по дисциплине «Численные методы» выстраивалась на базе этого электронного курса [21].

Пособие предназначено для студентов университета, обучающихся по направлению подготовки 01.03.02 «Прикладная математика и информатика», изучающих курс «Численные методы», и преподавателей.

Материал пособия может быть полезен студентам, изучающим в вузе численные методы на различных направлениях подготовки, а также студентам магистратуры ИИТММ, изучающим параллельные численные методы на основе технологий параллельного программирования.

Материал разработан таким образом, чтобы обеспечить освоение тематического модуля фундаментальной дисциплины в случаях очного, смешанного, а также дистанционного способа организации обучения.

Модуль 10. Теоретический аппарат для построения и применения итерационных методов решения СЛАУ

Нормы векторов и матриц, согласованные и подчиненные нормы. Собственные числа, их свойства и оценка. Спектральный радиус, оценка нормы матрицы. Норма обратной матрицы. Симметричные и симметричные положительно определенные матрицы, их свойства и нормы. Число обусловленности, его влияние на свойства СЛАУ. Примеры плохой обусловленности. Механизм влияния обусловленности на погрешность решения СЛАУ. Нормализация. Скорость сходимости. Презентация модельных задач

Рассмотрим систему линейных алгебраических уравнений (СЛАУ) вида

$$Ax = b \tag{10.1}$$

Здесь $x \in R^n$, $b \in R^n$, $A\ (n \times n)$, $\det A \neq 0$ (невырожденная матрица).

Через x^* обозначим решение задачи (10.1), $x^* \in R^n$.

Итерационные методы решения СЛАУ генерируют последовательность приближенных решений задачи (10.1):

$$x^{(0)}, x^{(1)}, \ldots x^{(s)} \ldots$$

Определение. **Итерационный метод называется сходящимся**, если при любом выборе начального приближения последовательность приближенных решений сходится к точному решению:

$$\forall x^{(0)} \in R^n \quad \lim_{s \to +\infty} \left\| x^{(s)} - x^* \right\| = 0.$$

Далее представлен теоретический аппарат, необходимый для построения итерационных методов и изучения их сходимости.

10.1. Нормы векторов и матриц, согласованные и подчиненные нормы

Определение. **Нормой вектора** $x \in R^n$ называется функционал $\| x \|$, удовлетворяющий трем аксиомам нормы:

1. $\| x \| \geq 0; \| x \| = 0 \Leftrightarrow x = 0$
2. $\| \alpha x \| = | \alpha | \cdot \| x \|$
3. $\| x + y \| \leq \| x \| + \| y \|$

Здесь $x, y \in R^n$, α – число.

Примеры. Обычно используют нормы

$$\| x \|_1 = \sum_{i=1}^{n} | x_i | \qquad \| x \|_2 = \sqrt{(x, x)} \qquad \| x \|_\infty = \max_{i=1,n} | x_i |$$

Определение. **Нормой матрицы** $A\ (n \times n)$ называется функционал $\| A \|$, удовлетворяющий четырем аксиомам нормы:

1. $\| A \| \geq 0; \| A \| = 0 \Leftrightarrow A = 0$

2. $\| \alpha A \| = | \alpha | \cdot \| A \|, \forall \alpha \in R$

3. $\| A + B \| \leq \| A \| + \| B \|$

4. $\| AB \| \leq \| A \| \cdot \| B \|$

Здесь A, B – матрицы $(n \times n)$, α – число.

Определение. Норма матрицы **согласована** с нормой вектора, если $\forall x \in R^n$

$$\| Ax \| \leq \| A \| \cdot \| x \| \tag{10.2}$$

Определение. Нормой матрицы, **подчиненной** норме вектора, называют функционал

$$\| A \| = \sup_{x \neq 0} \frac{\| Ax \|}{\| x \|} \tag{10.3}$$

Примеры

Если в (10.3) использовать $\| x \|_1$ и $\| Ax \|_1$, получим норму $\| A \|_1$.

Если в (10.3) использовать $\| x \|_2$ и $\| Ax \|_2$, получим норму $\| A \|_2$.

Если в (10.3) использовать $\| x \|_\infty$ и $\| Ax \|_\infty$, получим норму $\| A \|_\infty$.

Для подчиненных норм справедливо следующее:

Утверждение 1. Функционал (10.3), то есть **подчиненная** норма, соответствует всем 4-м аксиомам нормы.

Утверждение 2. **Подчиненная** норма является **согласованной:** если $\| A \|$ удовлетворяет (10.3), то $\forall x \in R^n$ справедливо (10.2).

При изучении сходимости методов обычно используют подчиненные нормы. Это связано с тем, что оценки вида (10.2), записанные в подчиненных нормах, **оптимальны**: улучшить такие оценки можно только за счет сужения области их применения:

Утверждение 3. Пусть A – матрица $(n \times n)$. Рассмотрим оценки вида

$$\| Ax \| \leq M \cdot \| x \| \tag{10.4}$$

Если выбрано $M \geq \| A \|$, оценка (10.4) верна для $\forall x$.

Если выбрано $M < \| A \|$, оценка (10.4) не может быть верна для $\forall x$.

При $M = \| A \|$ получим (10.2) и улучшить (10.2) нельзя.

(утверждения 1-3 доказать самостоятельно).

10.2. Собственные числа и спектральный радиус

Определение. **Собственными числами** матрицы A $(n \times n)$ называются корни характеристического уравнения

$$\det(A - \lambda E) = 0 \tag{10.5}$$

Обозначим их через $\lambda_i(A), i = 1,\ldots n$, и напомним основные свойства:

1) Для любой квадратной матрицы A $(n \times n)$ количество собственных чисел с учетом кратности корней *характеристического уравнения* (10.5) равно n. Если элементы матрицы A являются действительными числами, собственные числа могут быть как действительными, так и комплексными.

2) **Алгебраической кратностью** собственного числа λ называется кратность соответствующего корня характеристического уравнения (10.5).

3) Каждое собственное число λ имеет хотя бы один (с точностью до постоянного множителя) собственный вектор v.

4) Ненулевой вектор v называется собственным и ему соответствует собственное число λ, если $Av = \lambda v$.

5) Количество линейно независимых собственных векторов, соответствующих некоторому собственному числу λ, называют **геометрической кратностью** собственного числа.

6) **Геометрическая кратность** собственного числа не меньше, чем 1, и не превосходит его **алгебраической кратности.**

7) Вырожденная матрица имеет нулевое собственное число:

$(\det A = 0) \Rightarrow \det(A - \lambda E) = 0$ имеет корень $\lambda = 0$.

8) Матрица A невырождена тогда и только тогда, когда все ее собственные числа отличны от нуля :

$$(\det A \neq 0) \Leftrightarrow (\lambda_i(A) \neq 0, i = 1,\ldots n) \tag{10.6)}$$

Покажем, как связаны собственные числа и согласованные матричные нормы.

Определение. **Спектром** матрицы A называют **множество всех ее собственных чисел** (на комплексной плоскости).

Определение. **Спектральным радиусом** $\rho(A)$ матрицы A называют **модуль максимального по модулю собственного числа матрицы** A:

$$\rho(A) = \max_{i=1,n} |\lambda_i(A)| \tag{10.7}$$

(спектральный радиус есть расстояние от нуля (на комплексной плоскости) до наиболее удаленного от нуля собственного числа).

Утверждение 4. Если норма матрицы согласована с нормой вектора, норма матрицы не может быть меньше, чем спектральный радиус:

$$\| A \| \geq \rho(A). \tag{10.8}$$

Доказательство

Пусть v – собственный вектор матрицы A, λ – соответствующее ему собственное число: $Av = \lambda v$. Тогда по аксиомам нормы вектора

$$\| Av \| = \| \lambda v \| = | \lambda | \cdot \| v \|$$

По свойству согласованности норм

$$\| Av \| \le \| A \| \cdot \| v \|$$

откуда следует, что $|\lambda| \le \| A \|$. Это справедливо для любого собственного числа, в том числе для максимального по модулю, откуда следует (10..6).

Утверждение 5. Если норма матрицы подчинена евклидовой норме вектора, то

$$\| A \|_2 = \sqrt{\rho(A^T A)} \tag{10.9}$$

(доказательство в учебной литературе).

10.3. Оценка собственных чисел по теореме Гершгорина

Теорема Гершгорина. Собственные числа матрицы $A = \{a_{ij}\}, i, j = 1,\ldots n$ расположены на комплексной плоскости в кругах следующего вида:

$$|z - a_{ii}| \le \sum_{\substack{j=1, \\ j \ne i}}^{n} |a_{ij}|, i = 1,\ldots n \tag{10.10}$$

Если множество кругов (10.10) состоит из нескольких связных компонент, каждая связная компонента содержит столько собственных чисел, сколько кругов ее составляют.

Пример

$$A = \begin{bmatrix} 7 & -3 & 0.5 \\ 1 & 9 & 1 \\ 2 & -1 & 15 \end{bmatrix}$$

Круги Гершгорина на комплексной плоскости имеют вид:

$|z - 7| \le 3.5$ $\quad |z - 9| \le 2$ $\quad |z - 15| \le 3$

Круги образуют *два связных множества*: объединение кругов с центрами в точках 7 и 9, и круг с центром в точке 15. Два собственных числа расположены в объединении двух кругов, и одно собственное число – в изолированном круге с центром в точке 15. *Из расположения кругов следует:*

1) *спектральный радиус* (модуль максимального по модулю собственного числа) можно оценить сверху и снизу:

$12 \le \rho\,(A) \le 18$), см. (10.7);

2) модуль *минимального по модулю собственного числа* можно оценить сверху и снизу:

$$3.5 \le \min_{i=1,2,3} |\lambda_i(A)| \le 11;$$

3) матрица не является вырожденной: $\det A \ne 0$, потому что нулевого собственного числа нет, см. (10.6).

10.4. Симметричные матрицы, их нормы

В приложениях часто встречаются СЛАУ с **симметричной** матрицей A и, в частности, с **симметричной, положительно определенной** матрицей A. Для решения указанных классов задач разработаны специальные прямые и итерационные методы.

Покажем, как связаны нормы таких матриц с их собственными числами.

Определение. Матрица называется симметричной, если она совпадает с результатом своего транспонирования: $A = A^T$.

Утверждение 6. Все собственные числа матрицы $A = A^T$ действительны, их можно упорядочить: $\lambda_1 \le \lambda_2 \le \ldots \le \lambda_n$ и существует ортонормированный базис из собственных векторов.

Спектральный радиус матрицы $A = A^T$ определяется ее «крайними» собственными числами, а именно **модулем минимального** или **модулем максимального** собственного числа:

$$\rho(A) = \max\{|\lambda_1|, |\lambda_n|\} \qquad (10.11)$$

(это следует из (10.7) и упорядоченности собственных чисел).

Утверждение 7. Пусть $A = A^T$. Тогда норма матрицы, согласованной с нормой вектора, *оценивается* модулем минимального или модулем максимального собственного числа:

$$\|A\| \ge \rho(A) = \max\{|\lambda_1|, |\lambda_n|\} \qquad (10.12)$$

Норма матрицы, подчиненная евклидовой норме вектора, *определяется* модулем минимального или модулем максимального собственного числа:

$$\|A\|_2 = \rho(A) = \max\{|\lambda_1|, |\lambda_n|\} \qquad (10.13)$$

Доказательство

(10.12) следует из (10.11) и утверждения 4, а (10.13) – из (10.11) и утверждения 5.

Для (10.13) проведем детальное обоснование. Во-первых, для любой квадратной матрицы A $(n \times n)$

$\|A\|_2 = \sqrt{\rho(A^T A)}$, где $\rho(A^T A) = \max\limits_{i=1,n} \left|\lambda_i(A^T A)\right|$ есть модуль максимального по модулю собственного числа матрицы $A^T A$.

Пусть $A = A^T$ (симметричная). Тогда $A^T A = A^2$ и собственными числами матрицы A^2 являются квадраты собственных чисел матрицы A: если $Av = \lambda v$, то $A^2 v = A(Av) = A(\lambda v) = \lambda(Av) = \lambda(\lambda v) = \lambda^2 v$. Тогда

$$\rho(A^T A) = \rho(A^2) = \max_{i=1,n} \left|\lambda_i(A^2)\right| = \max_{i=1,\ldots n} \left|[\lambda_i(A)]^2\right| =$$

$$= \max_{i=1,\ldots n} |\lambda_i(A)|^2 = \left(\max_{i=1,\ldots n} |\lambda_i(A)|\right)^2 = [\rho(A)]^2$$

что означает: для $A = A^T$ выполняется $\rho(A^2) = [\rho(A)]^2$.

Поэтому для симметричной матрицы $A = A^T$

$$\| A \|_2 = \sqrt{\rho(A^T A)} = \sqrt{\rho(A^2)} = \rho(A), \qquad (10.14)$$

откуда с учетом (10.11) следует (10.13).

10.5. Норма обратной матрицы

Выясним, каковы собственные числа и собственные векторы обратной матрицы.

Утверждение 8. Пусть A – матрица размерности $n \times n$, $\det A \neq 0$. Тогда матрицы A и A^{-1} имеют одинаковый комплект линейно независимых собственных векторов, а их собственные числа связаны соотношением

$$\lambda_i\ (A^{-1}) = \frac{1}{\lambda_i\ (A)},\ i = 1,\ldots n. \qquad (10.15)$$

Доказательство

Пусть v – собственный вектор, λ – собственное число матрицы A: $Av = \lambda v$. Умножим левую и правую части равенства на обратную матрицу (слева):

$A^{-1}A\, v = A^{-1}(\lambda v) = \lambda \cdot \left(A^{-1}v\right)$, что означает $v = \lambda \cdot \left(A^{-1}v\right)$.

Разделим левую и правую части равенства на число λ (потому что $\lambda \neq 0$):

$$A^{-1}v = \frac{1}{\lambda}v. \qquad (10.16)$$

Из (10.16) следует, что v – собственный вектор матрицы A^{-1} и число $1/\lambda$ – ее собственное число.

Повторим выкладки для всех линейно независимых собственных векторов матрицы A и получим некоторый комплект линейно независимых собственных векторов матрицы A^{-1}. Затем, повторив выкладки для всех собственных векторов матрицы A^{-1}, убеждаемся в том, что A и A^{-1} располагают одинаковым комплектом линейно независимых собственных векторов. Утверждение доказано.

Покажем, как оценить норму обратной матрицы.

Утверждение 9. **Спектральный радиус** обратной матрицы A^{-1} является величиной, **обратной модулю минимального по модулю собственного числа** исходной матрицы A:

$$\rho\,(A^{-1}) = \frac{1}{\min\limits_{i=1,n} |\ \lambda_i\,(A)\ |} \qquad (10.17)$$

Доказательство

$$\rho(A^{-1}) = \max_{i=1,\ldots n} \left|\ \lambda_i\,(A^{-1})\ \right| = \max_{i=1,\ldots n} \frac{1}{|\ \lambda_i\,(A)\ |} = \frac{1}{\min\limits_{i=1,n} |\ \lambda_i\,(A)\ |}.$$

Из (10.17) и утверждения 4 следует

Утверждение 10. Пусть A – невырожденная матрица размерности $n \times n$, $\det A \neq 0$. Тогда норма обратной матрицы, согласованная с нормой вектора, *оценивается* величиной, обратной модулю минимального по модулю собственного числа:

$$\| A^{-1} \| \geq \rho(A^{-1}) = \frac{1}{\min\limits_{i=1,n} |\lambda_i(A)|} \qquad (10.18)$$

Из (10.14), (10.17), и утверждения 5 следует

Утверждение 11. Пусть $A = A^T$ – невырожденная симметричная матрица размерности $n \times n$, $\det A \neq 0$. Тогда обратная матрица симметрична

$$A^{-1} = [A^{-1}]^T$$

и норма обратной матрицы, подчиненная евклидовой норме вектора, *определяется* величиной, обратной модулю минимального по модулю собственного числа:

$$\| A^{-1} \|_2 = \rho(A^{-1}) = \frac{1}{\min\limits_{i=1,n} |\lambda_i(A)|} \qquad (10.19)$$

10.6. Симметричные положительно определенные матрицы

Определение. Матрица A $(n \times n)$ называется **положительно определенной** (обозначается $A > 0$), если $\forall h \neq 0, h \in R^n$ $(Ah, h) > 0$.

Утверждение 12. Собственные числа положительно определенной матрицы положительны: $\lambda_i(A) > 0, i = 1,...n$, и такая матрица невырождена: $\det A \neq 0$.

Доказательство

1) Пусть v – некоторый собственный вектор матрицы A, λ – соответствующее ему собственное число: $Av = \lambda v$.

Так как $v \neq 0, v \in R^n$ и $A > 0$, получим $(Av, v) > 0$.

Используя свойства собственных чисел и собственных векторов, запишем $(Av, v) = (\lambda v, v) = \lambda(v, v) > 0$.

Учитывая $(v, v) > 0$, получим $\lambda > 0$.

2) $\det A \neq 0$ следует из того, что $A > 0$ не имеет нулевого собственного числа.

Наличие двух свойств одновременно (симметричность и положительная определенность) обозначается как $A = A^T > 0$.

Утверждение 13. Все собственные числа матрицы $A = A^T > 0$ положительны, их можно упорядочить: $0 < \lambda_1 \leq \lambda_2 \leq ... \leq \lambda_n$ и существует ортонормированный базис из собственных векторов.

Сформулируем и докажем критерий положительной определенности для симметричных матриц:

Утверждение 14. Если матрица симметрична $A = A^T$, необходимым и достаточным условием положительной определенности, т.е. выполнения условия

$$\forall h \neq 0, h \in R^n \;\; (Ah, h) > 0 \qquad (10.20)$$

является положительность всех ее собственных чисел:

$$\lambda_i(A) > 0, i = 1,...n \qquad (10.21)$$

Доказательство

1) из условия $\forall h \neq 0, h \in R^n \quad (Ah, h) > 0$ следует $\lambda_i(A) > 0, i = 1,...n$ (см. утверждение 14).

2) пусть $\lambda_i(A) > 0, i = 1,...n$. Покажем, что $\forall h \neq 0, h \in R^n \;\; (Ah, h) > 0$.

Для этого используем ортонормированный базис из собственных векторов $v_1, \; v_2 ... v_n$. Базис имеет свойства

$$(v_i, v_i) = 1, i = 1,...n$$

$$(v_i, v_l) = 0, i, l = 1,...n, i \neq l$$

$$Av_i = \lambda_i v_i, i = 1,...n$$

Запишем вектор $h \in R^n$, $h \neq 0$ в виде $h = \sum\limits_{i=1}^{n} \alpha_i v_i$. Среди коэффициентов разложения по базису есть хотя бы один, отличный от нуля: $\exists\, \alpha_s \neq 0$. Тогда

$$(Ah, h) = (A\sum_{i=1}^{n} \alpha_i v_i, \sum_{l=1}^{n} \alpha_l v_l) =$$

$$= (\sum_{i=1}^{n} \alpha_i (Av_i), \sum_{l=1}^{n} \alpha_l v_l) ==(\sum_{i=1}^{n} \lambda_i \alpha_i v_i, \sum_{l=1}^{n} \alpha_l v_l) =$$

$$= \sum_{i=1}^{n} \lambda_i \; \alpha_i^2 (v_i, v_i) \geq \lambda_s \alpha_s^2 > 0$$

Критерий доказан.

Перейдем к оценке нормы $A = A^T > 0$ и обратной матрицы.

Спектральный радиус матрицы $A = A^T > 0$ определяется ее **максимальным** собственным числом (оно положительное):

$$\rho(A) = \lambda_n(A) \qquad (10.22)$$

Спектральный радиус матрицы A^{-1}, обратной $A = A^T > 0$ определяется величиной, **обратной минимальному собственному числу** матрицы A (оно положительное):

$$\rho(A^{-1}) = \frac{1}{\lambda_1(A)} \qquad (10.23)$$

Указанные неравенства (10.22) и (10.23) вытекают из (10.11) и (10.17).

Из (10.22) и утверждения 7 следует

Утверждение 15. Пусть $A = A^T > 0$. Тогда норма матрицы, согласованной с нормой вектора, *оценивается* максимальным собственным числом:

$$\| A \| \geq \rho(A) = \lambda_n(A) \tag{10.24}$$

Норма матрицы, подчиненная евклидовой норме вектора, *определяется* максимальным собственным числом:

$$\| A \|_2 = \rho(A) = \lambda_n(A) \tag{10.25}$$

Утверждение 16. Матрица, обратная к $A = A^T > 0$, симметрична и положительно определена $A^{-1} = [A^{-1}]^T > 0$ (доказать самостоятельно).

Из утверждения 16, утверждения 7 и (10.23) следует

Утверждение 17. Пусть A – матрица размерности $n \times n$, $A = A^T > 0$. Тогда норма обратной матрицы, согласованная с нормой вектора, *оценивается* величиной, обратной минимальному собственному числу матрицы A:

$$\| A^{-1} \| \geq \rho(A^{-1}) = \frac{1}{\lambda_1(A)} \tag{10.26}$$

Норма обратной матрицы, подчиненная евклидовой норме вектора, *определяется* величиной, обратной минимальному собственному числу исходной матрицы:

$$\| A^{-1} \|_2 = \rho(A^{-1}) = \frac{1}{\lambda_1(A)} \tag{10.27}$$

10.7. Число обусловленности и его влияние на свойства СЛАУ

Определение. **Число обусловленности** невырожденной матрицы A обозначают μ_A и определяют как

$$\mu_A = \| A \| \cdot \| A^{-1} \| \tag{10.28}$$

В формуле (10.28) используются матричные нормы, согласованные с нормами векторов.

Определение. **Число обусловленности Тодта** невырожденной матрицы A размерности $n \times n$ обозначают μ_A^T и определяют как

$$\mu_A^T = \frac{\max\limits_{i=1,n} |\lambda_i(A)|}{\min\limits_{i=1,n} |\lambda_i(A)|}. \tag{10.29}$$

Утверждение 18. При выборе матричной нормы, согласованной с нормой вектора

$$\mu_A \geq \mu_A^T \geq 1. \tag{10.30}$$

(число обусловленности Тодта оценивает снизу число обусловленности μ_A и само оценивается снизу единицей).

Доказательство

Из (10.7), (10.8) и (10.17) и определения числа обусловленности следует

$$\| A \| \geq \rho(A) = \max_{i=1,n} |\lambda_i(A)|$$

$$\| A^{-1} \| \geq \rho(A^{-1}) = \frac{1}{\min\limits_{i=1,n} |\lambda_i(A)|}$$

$$\mu_A = \| A \| \cdot \| A^{-1} \| \geq \frac{\max\limits_{i=1,n} |\lambda_i(A)|}{\min\limits_{i=1,n} |\lambda_i(A)|} = \mu_A^T \geq 1$$

Утверждение 19. Пусть A – матрица размерности $n \times n$, $A = A^T > 0$. Пусть число обусловленности μ_A определено на основе матричной нормы, подчиненной евклидовой норме вектора. Тогда

$$\mu_A = \| A \|_2 \cdot \| A^{-1} \|_2 = \frac{\lambda_n}{\lambda_1} = \mu_A^T \geq 1 \tag{10.31}$$

Доказательство

Из (10.25), (10.27) и определения числа обусловленности следует

$$\| A \|_2 = \rho(A) = \lambda_n$$

$$\| A^{-1} \|_2 = \rho(A^{-1}) = \frac{1}{\lambda_1}$$

$$\mu_A = \| A \|_2 \cdot \| A^{-1} \|_2 = \frac{\lambda_n}{\lambda_1} = \mu_A^T \geq 1$$

Обусловленность матриц и СЛАУ

С числом μ_A связаны следующие свойства СЛАУ и итерационных методов.

Свойство 1. Если $\mu_A >> 1$, то при решении СЛАУ вида $Ax = b$ небольшие погрешности исходных данных могут приводить к значительным погрешностям решения (независимо от выбора метода решения), см. утверждение 20.

Свойство 2. Если $\mu_A >> 1$, то при решении СЛАУ вида $Ax = b$ итерационными методами скорость сходимости, как правило, медленная, см. теоремы о сходимости итерационных методов (Модуль 11).

В связи с указанными свойствами различают **хорошо** и **плохо обусловленные матрицы**, **хорошо** и **плохо обусловленные СЛАУ**.

Если $\mu_A >> 1$, матрица A и система линейных алгебраических уравнений $Ax = b$ называется **плохо обусловленной**.

Если $\mu_A \approx 1$, матрица A и система линейных алгебраических уравнений $Ax = b$ называется **хорошо обусловленной**.

Сформулируем и докажем утверждение, характеризующее влияние числа обусловленности на погрешность решения СЛАУ.

Утверждение 20. Пусть A – матрица размерности $n \times n$, $\det A \neq 0.$
Рассматривается исходная система уравнений
$Ax = b$, где x^* – решение, и система уравнений с возмущенной правой частью
$Ax = b + \Delta b$, где $x^* + \Delta x$ – решение.
Тогда

$$\frac{\| \Delta x \|}{\| x^* \|} \leq \mu_A \frac{\| \Delta b \|}{\| b \|} \qquad (10.32)$$

Здесь μ_A – число обусловленности матрицы A.

Величину $\| \Delta b \|$ называют абсолютным возмущением правой части,

величину $\| \Delta x \|$ – абсолютным возмущением решения. Аналогично:

$\dfrac{\| \Delta b \|}{\| b \|}$ – относительное возмущение правой части;

$\dfrac{\| \Delta x \|}{\| x^* \|}$ – относительное возмущение решения.

Доказательство

Так как $Ax^* = b$, запишем

$$\| b \| = \| A x^* \| \leq \| A \| \cdot \| x^* \|$$

и получим оценку

$$\| x^* \| \geq \frac{\| b \|}{\| A \|} \qquad (10.33)$$

(норма матрицы должна быть согласована с нормой вектора).

Так как $A(x^* + \Delta x) = b + \Delta b$, запишем $A\Delta x = \Delta b$ и выразим $\Delta x = A^{-1}\Delta b$.

Проведем оценку

$$\| \Delta x \| = \| A^{-1}\Delta b \| \leq \| A^{-1} \| \cdot \| \Delta b \| \qquad (10.34)$$

(норма матрицы согласована с нормой вектора).

Применим (10.33) и (10.34) для оценки относительного возмущения решения:

$$\frac{\| \Delta x \|}{\| x^* \|} \leq \frac{\| A^{-1} \| \cdot \| A \| \cdot \| \Delta b \|}{\| b \|} \leq \mu_A \frac{\| \Delta b \|}{\| b \|}$$

Неравенство (10.32) доказано.

Комментарии

Если $\mu_A \approx 1$, то небольшой относительной погрешности правой части соответствует небольшая относительная погрешность решения.

Если $\mu_A >> 1$, то небольшому относительному возмущению правой части может соответствовать значительное относительное возмущение решения.

Если оценка (10.32) получена на основе **подчиненных** норм, оценку улучшить нельзя: для $\forall A$ при некоторых b и Δb (10.32) выполняется как равенство.

10.8. Примеры плохой обусловленности

Пример плохо обусловленной матрицы

$$A = \begin{bmatrix} 100 & 99 \\ 99 & 98 \end{bmatrix}$$

Матрица симметрична: $A = A^T$. Собственные числа (решения $\det(A - \lambda E) = 0$):

$\lambda_1 = -0.00505$ и $\lambda_2 = 198.00505$

Для числа обусловленности (в любой согласованной норме) верно

$$\mu_A \ge \frac{|\lambda_2|}{|\lambda_1|} = 39208.92 >> 1$$

Число обусловленности (в норме, подчиненной евклидовой норме вектора) составит

$$\mu_A = 39208.92 >> 1.$$

Пример исходной и возмущенной СЛАУ, для которых малая относительная погрешность правой части СЛАУ влечет за собой большую относительную погрешность решения

Исходная система (с плохо обусловленной матрицей) имеет вид

$$\begin{bmatrix} 100 & 99 \\ 99 & 98 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 199 \\ 197 \end{bmatrix}$$

Ее решение: $x^* = \begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

Рассмотрим систему с возмущенной правой частью:

$$\begin{bmatrix} 100 & 99 \\ 99 & 98 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 199+1 \\ 197-1 \end{bmatrix} = \begin{bmatrix} 200 \\ 196 \end{bmatrix}$$

Ее решение: $x^* + \Delta x = \begin{bmatrix} x_1^* + \Delta x_1 \\ x_2^* + \Delta x_2 \end{bmatrix} = \begin{bmatrix} 1-197 \\ 1+199 \end{bmatrix} = \begin{bmatrix} -196 \\ 200 \end{bmatrix}$

Очевидно, что относительно малое возмущение правой части, а именно,

возмущение $\begin{bmatrix} +1 \\ -1 \end{bmatrix}$ к правой части исходной задачи $\begin{bmatrix} 199 \\ 197 \end{bmatrix}$, приводит к существенному относительному возмущению решения, а именно,

возмущению $\begin{bmatrix} -197 \\ +199 \end{bmatrix}$ к решению исходной задачи $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

10.9. Механизм влияния обусловленности на погрешность решения СЛАУ

Пусть $A = A^T > 0$ – симметричная положительно определенная матрица ($n \times n$), ее собственные числа положительны и упорядочены: $0 < \lambda_1 \le \lambda_2 \le \ldots \le \lambda_n$.

Пусть $v_1, v_2 \ldots v_n$ – ортонормированный базис из собственных векторов, собственный вектор v_i соответствует собственному числу λ_i, $i = 1, \ldots n$:

$$(v_i, v_i) = 1,\ i = 1, \ldots n$$

$$(v_i, v_l) = 0,\ i, l = 1, \ldots n,\ i \ne l$$

$$A v_i = \lambda_i v_i,\ i = 1, \ldots n$$

Рассмотрим задачу $Ax = b$, где $b = v_n$

(правая часть СЛАУ есть нормированный собственный вектор, соответствующий максимальному собственному числу)

Решением СЛАУ является $x^* = \dfrac{v_n}{\lambda_n}$.

Рассмотрим задачу с возмущенной правой частью:

$Ax = b + \Delta b$, где $b + \Delta b = v_n + \alpha \cdot v_1$, то есть $\Delta b = \alpha \cdot v_1$

(возмущение правой части СЛАУ есть нормированный собственный вектор, соответствующий минимальному собственному числу, умноженный на некоторое малое число α)

Решением возмущенной СЛАУ является

$$x^* + \Delta x = \frac{v_n}{\lambda_n} + \alpha \cdot \frac{v_1}{\lambda_1}.$$

Возмущение решения составило $\Delta x = \alpha \cdot \dfrac{v_1}{\lambda_1}$

Подсчитаем величины относительных возмущений для правой части СЛАУ и для ее решения.

Относительное возмущение правой части составит

$$\frac{\| \Delta b \|}{\| b \|} = \frac{\| \alpha \cdot v_1 \|}{\| v_n \|} = | \alpha | \frac{\| v_1 \|}{\| v_n \|} = | \alpha |$$

Относительное возмущение решения составит

$$\frac{\|\Delta x\|}{\|x^*\|} = |\alpha| \cdot \frac{|\lambda_n| \cdot \|v_1\|}{|\lambda_1| \cdot \|v_n\|} = |\alpha| \cdot \frac{\lambda_n}{\lambda_1} = |\alpha| \cdot \mu_A$$

Таким образом, относительное возмущение решения пропорционально относительному возмущению правой части с коэффициентом, значение которого в точности совпадает с величиной числа обусловленности (определенного в евклидовой норме):

$$\frac{\|\Delta x\|}{\|x^*\|} = \mu_A \frac{\|\Delta b\|}{\|b\|}$$

Нестрогое неравенство (10.32) в данном случае выполняется как равенство.

Если $\mu_A \approx 1$ (хорошо обусловленная матрица), относительное возмущение решения окажется примерно таким, как относительное возмущение правой части.

Если $\mu_A >> 1$ (плохо обусловленная матрица), относительное возмущение решения окажется много больше, чем относительное возмущение правой части.

Комментарий

В рассмотренном примере правая часть СЛАУ есть собственный вектор, соответствующий максимальному собственному числу, а возмущение правой части – собственный вектор, соответствующий минимальному собственному числу. Величина возмущения (число α) на выполнение (10.32) как равенства не влияет.

Задание

Приведите пример исходной и возмущенной СЛАУ с плохо обусловленной матрицей, когда малое относительное возмущение правой части приводит к малому относительному возмущению решения. При построении примера используйте базис из собственных векторов.

10.10. Нормализация СЛАУ

Рассмотрим систему линейных алгебраических уравнений (СЛАУ) вида (10.1): $Ax = b$

Здесь $x \in R^n, b \in R^n, A\,(n \times n), \det A \neq 0$ (невырожденная матрица).

Умножаем левую и правую части (10.1) слева на A^T и получим СЛАУ

$$A^T A x = A^T b \tag{10.35}$$

Утверждение 21. Матрица $A^T A$, где $A\,(n \times n), \det A \neq 0$, симметрична и положительно определена. Задачи (10.1) и (10.35) эквивалентны: имеют одинаковое и единственное решение.

Доказательство

$\left(A^T A\right)^T = A^T A^{TT} = A^T A$.

$\forall h \neq 0 \; Ah \neq 0$, потому что $\det A \neq 0$

$\forall h \neq 0 \;\; (A^T Ah, h) = (Ah, A^{TT} h) = (Ah, Ah) > 0.$

$\det A^T A \neq 0$, так как $\lambda_i (A^T A) > 0, i = 1,...n$

$$Ax^* = b \;\; \Rightarrow \;\; A^T Ax^* = A^T (Ax^*) = A^T b$$

Определение. Замену СЛАУ (10.1) с произвольной невырожденной матрицей на эквивалентную СЛАУ (10.35) с симметричной положительно определенной матрицей называют **нормализацией**.

Комментарий

Недостаток замены в том, что СЛАУ (10.35) может иметь существенно большее число обусловленности, чем исходная СЛАУ: $\mu_{A^T A} >> \mu_A$.

10.11. Скорость сходимости (определение)

Рассмотрим систему линейных алгебраических уравнений (СЛАУ) вида (10.1): $Ax = b$. Здесь $x \in R^n, b \in R^n, A\,(n \times n), \det A \neq 0$ (невырожденная матрица). Решение задачи обозначено через $x*$, $x* \in R^n$. Итерационные методы генерируют последовательность приближенных решений: $x^{(0)}, x^{(1)}, \ldots x^{(s)} \ldots$

Определение. Погрешность итерационного метода на шаге s обозначим через $z^{(s)}$ и определим как разность приближенного и точного решения задачи на шаге s:

$$z^{(s)} = x^{(s)} - x*.$$

Пусть для погрешности метода на шаге s верна оценка

$$\left\| z^{(s)} \right\| = \left\| x^{(s)} - x* \right\| \leq Mq^s \left\| z^{(0)} \right\| \tag{10.36}$$

где $M > 0$, $0 < q < 1$ - константы, характеризующие СЛАУ (10.1) и исследуемый метод и не зависящие от выбора начального приближения $x^{(0)}$.

Тогда величину $\lg\left(\frac{1}{q}\right)$ называют **скоростью сходимости метода**.

Если $q \approx 0$ (оценка погрешности быстро стремится к нулю), скорость сходимости велика:

$$\lg\left(\frac{1}{q}\right) >> 1.$$

Если $q \approx 1$ (оценка погрешности медленно стремится к нулю), скорость сходимости мала:

$$\lg\left(\frac{1}{q}\right) \approx 0.$$

Комментарий

Смысл показателя, именуемого «скорость сходимости», покажем на примере: сколько нужно сделать шагов, чтобы начальная погрешность решения сократилась в 10 или более раз?

Решение: пусть N – искомое число шагов. Нужно, чтобы

$$\frac{\left\| z^{(0)} \right\|}{\left\| z^{(N)} \right\|} \geq 10.$$

Так как

$$\frac{\left\| z^{(0)} \right\|}{\left\| z^{(N)} \right\|} \geq \frac{1}{M} \cdot \left(\frac{1}{q} \right)^N,$$

требуем, чтобы $\frac{1}{M} \cdot \left(\frac{1}{q} \right)^N \geq 10$. (10.37)

Логарифмируем неравенство (10.37):

$$N \cdot \lg\left(\frac{1}{q} \right) \geq \lg 10 + \lg M.$$

Для искомого N получим

$$N \geq \frac{1 + \lg M}{\lg\left(\frac{1}{q} \right)}. \qquad (10.38)$$

Таким образом, чем выше скорость сходимости, тем меньше шагов потребуется для того, чтобы начальная погрешность сократилась в 10 или более раз. Если методы имеют одинаковую скорость сходимости, то для метода с большим значением $M > 0$ нужно большее число шагов N.

При изучении сходимости итерационных методов также используются понятия

– **средней скорости сходимости** (за конечное число шагов)

– **асимптотической скорости сходимости** (когда число проведенных шагов стремится к бесконечности).

10.12. Презентация модельных задач

Итерационные методы применяем с целью численного решения задачи Дирихле для уравнения Пуассона: 1) в прямоугольной области, где известны собственные числа матрицы СЛАУ; 2) в области, включенной в прямоугольник, где известны их оценки.

Модуль 11. Примеры итерационных методов решения СЛАУ для задач большой размерности

Метод простой итерации. Метод минимальных невязок.
Метод с чебышевским набором параметров. Метод сопряженных градиентов.
Расчет параметров, теоремы о сходимости и построение сходящихся методов на основе оценок собственных чисел

11.1. Метод простой итерации

Теоремы об условиях сходимости и оценках сходимости.
Выбор оптимального параметра, оптимальная оценка сходимости.
Влияние числа обусловленности на сходимость метода.
Расчет параметра метода на основе оценок собственных чисел, оптимальный выбор параметра и оценка погрешности метода

Общий случай

Рассмотрим СЛАУ (систему линейных алгебраических уравнений) вида

$$Ax = b, \tag{11.1}$$

где $x \in R^n, b \in R^n, A\,(n \times n), \det A \neq 0$ (невырожденная матрица).

Через x^* обозначим точное решение системы, $x^* \in R^n$.

Методом простой итерации называют явный стационарный итерационный метод

$$\frac{x^{(s+1)} - x^{(s)}}{\tau} + Ax^{(s)} = b \tag{11.2}$$

где τ – число (постоянный параметр метода), $x^{(0)} \in R^n$ – начальное приближение для запуска итераций (его можно выбирать любым), $s = 0,1,...$ – номер шага метода.

Запись метода в виде (11.2) называется *канонической*. Для расчетов вместо (11.2) используется формула

$$x^{(s+1)} = x^{(s)} + \tau \cdot (b - Ax^{(s)}) = x^{(s)} - \tau \cdot r^{(s)} \tag{11.3}$$

где $r^{(s)} = Ax^{(s)} - b$ – невязка СЛАУ на текущем приближении $x^{(s)}$.

Основные свойства метода описывают следующие теоремы.

Теорема 1. При решении СЛАУ (11.1) **методом простой итерации** (11.2) оценка погрешности метода на шаге s имеет вид

$$\left\| z^{(s)} \right\|_2 \leq \| G(\tau) \|_2^{\,s} \left\| z^{(0)} \right\|_2 \tag{11.4}$$

достаточным условием сходимости метода является условие

$$\| G(\tau) \|_2 < 1. \tag{11.5}$$

Здесь $z^{(s)} = x^{(s)} - x^*$ – погрешность метода на шаге s, $z^{(0)} = x^{(0)} - x^*$ – погрешность метода на начальном шаге, $G(\tau) = E - \tau A$ – переходная матрица метода, $\| G(\tau) \|_2$ – норма переходной матрицы, подчиненная евклидовой норме вектора,

$\| G(\tau) \|_2^s$ – норма переходной матрицы, подчиненная евклидовой норме вектора и возведенная в степень s, $\| z^{(s)} \|_2$ и $\| z^{(0)} \|_2$ – евклидова норма векторов.

Доказательство

Выясним, как связаны $z^{(s)}$ и $z^{(0)}$. Запишем метод (11.2) в виде (11.3)

$$x^{(s+1)} = x^{(s)} + \tau \cdot (b - Ax^{(s)}) \tag{1}$$

Вычтем слева и справа x^* (это точное решение (11.1)), используем $b = Ax^*$:

$$x^{(s+1)} - x^* = x^{(s)} - x^* + \tau \cdot (Ax^* - Ax^{(s)}) \tag{2}$$

Напомним, что погрешностью метода на шаге s называют разность приближенного и точного решений: $z^{(s)} = x^{(s)} - x^*$. Аналогично, погрешность метода на шаге $s+1$ есть $z^{(s+1)} = x^{(s+1)} - x^*$.

Перепишем (2) в обозначениях погрешностей:

$$z^{(s+1)} = z^{(s)} - \tau \cdot Az^{(s)} = (E - \tau A)\, z^{(s)} \tag{3}$$

Введем обозначение

$$G(\tau) = E - \tau A. \tag{4}$$

Матрицу $G(\tau)$ называют **переходной матрицей метода:** она показывает, как связаны погрешности на текущем и следующем шаге:

$$z^{(s+1)} = G(\tau)\, z^{(s)} \tag{5}$$

На основе (5) выясним, как связана $z^{(s+1)}$ с начальной погрешностью:

$$z^{(s+1)} = G(\tau) z^{(s)} = (G(\tau))^2 z^{(s)} = \ldots = (G(\tau))^{s+1} z^{(0)} \tag{6}$$

(погрешности $z^{(s+1)}$ и $z^{(0)}$ связывает переходная матрица $G(\tau)$, возведенная в степень $s+1$). Поэтому для погрешности на шаге s справедливо

$$z^{(s)} = (G(\tau))^s z^{(0)} \tag{7}$$

Оценим $\| z^{(s)} \|_2$ – евклидову норму погрешности метода на шаге s. В силу согласованности нормы матрицы и евклидовой нормы вектора запишем

$$\| z^{(s)} \|_2 = \| (G(\tau))^s z^{(0)} \|_2 \le \| (G(\tau))^s \|_2 \| z^{(0)} \|_2 \tag{8}$$

В силу аксиомы нормы

$$
\left\| (G(\tau))^s \right\|_2 = \underbrace{\| G(\tau) \cdot G(\tau) ... G(\tau) \|_2}_{s\ раз} \le
$$

$$
\le \underbrace{\| G(\tau) \|_2 \cdot \| G(\tau) \|_2 \cdots \| G(\tau) \|_2}_{s\ раз} = \| G(\tau) \|_2^s \qquad (9)
$$

В неравенстве (8) $\left\| (G(\tau))^s \right\|_2$ есть евклидова норма переходной матрицы $G(\tau)$, **возведенной** в степень s, то есть $(G(\tau))^s$. В неравенстве (9) $\| G(\tau) \|_2^s$ есть «евклидова норма» переходной матрицы $G(\tau)$, **возведенная** в степень s.

Из (8) и (9) получаем

$$
\left\| z^{(s)} \right\|_2 \le \left\| (G(\tau))^s \right\|_2 \left\| z^{(0)} \right\|_2 \le \left\| G(\tau) \right\|_2^s \left\| z^{(0)} \right\|_2 \qquad (10)
$$

Оценка (11.4) доказана, и достаточным условием сходимости метода (11.2) является условие (11.5).

Если $\| G(\tau) \|_2 < 1$, то при $s \to +\infty$ $\| G(\tau) \|_2^s \to 0$ и $\left\| z^{(s)} \right\|_2 \to 0$.

Рассмотрим сходимость метода простой итерации для систем (11.1) с симметричной положительно определенной матрицей.

Теорема 2. При решении СЛАУ (11.1) с *симметричной, положительно определенной* матрицей $A = A^T > 0$ **методом простой итерации** (11.2) оценка погрешности метода на шаге s имеет вид

$$
\left\| z^{(s)} \right\|_2 \le \left(\max \{ | 1 - \lambda_1 \tau |, | 1 - \lambda_n \tau | \}\right)^s \left\| z^{(0)} \right\|_2 \qquad (11.6)
$$

необходимым и достаточным условием сходимости метода является

$$
\tau \in (0, \frac{2}{\lambda_n}) \qquad (11.7)
$$

Здесь $z^{(s)} = x^{(s)} - x^*$ – погрешность метода на шаге s, $z^{(0)} = x^{(0)} - x^*$ – погрешность метода на начальном шаге, $0 < \lambda_1 \le ... \le \lambda_n$ – собственные числа матрицы $A = A^T > 0$ (они положительны и упорядочены), λ_1 – минимальное из них, λ_n – максимальное из них.

Комментарии

1) Выражение в круглых скобках (см. (11.6)) есть норма переходной матрицы $G(\tau)$, подчиненная евклидовой норме вектора:

$$
\| G(\tau) \|_2 = \max \{ | 1 - \lambda_1 \tau |, \ | 1 - \lambda_n \tau | \} \qquad (11.8)
$$

Запись (11.8) следует из того, что собственные числа матрицы $G(\tau)$ можно выразить через собственные числа матрицы A как

$\lambda_i(G) = 1 - \tau \cdot \lambda_i(A),\ i = 1,\dots n$

и норма симметричной матрицы $G(\tau)$, подчиненная евклидовой норме вектора, определяется модулями ее «крайних» собственных чисел, а именно

$|\ \lambda_1(G)\ | = |\ 1 - \lambda_1(A)\cdot\tau\ |,\ |\ \lambda_n(G)\ | = |\ 1 - \lambda_n(A)\cdot\tau\ |.$

2) условия сходимости (11.7) получены из условия $\|\ G(\tau)\ \|_2 < 1$ как решение системы следующих двух неравенств:

$|\ 1 - \lambda_1\tau\ | < 1,\ \ |\ 1 - \lambda_n\tau\ | < 1.$

Доказательство

Покажем, что (11.6) выполняется при любом выборе параметра метода и условие (11.7) является достаточным для сходимости.

Рассмотрим собственные числа матрицы $G(\tau) = E - \tau A$.

Так как E и A симметричны, матрица $G(\tau)$ также симметрична и ее собственные числа действительны. Собственные числа матрицы A обозначим через $\lambda_i, i = 1,\dots n$, собственные векторы A обозначим через $v_i, i = 1,\dots n$: $Av_i = \lambda_i v_i, i = 1,\dots n.$

В силу симметричности и положительной определенности A все ее собственные числа положительны, их можно упорядочить: $0 < \lambda_1 \le \dots \le \lambda_n$.

Очевидно, что $G(\tau)v_i = (E - \tau A)v_i = v_i - \tau\lambda_i v_i = (1 - \tau\lambda_i)v_i$.

Это означает, что $G(\tau)$ имеет такие же собственные векторы, как A, и собственными числами $G(\tau)$ являются $\lambda_i(G) = 1 - \tau\lambda_i, i = 1,\dots n$.

С учетом того, что собственные числа $G(\tau)$ действительны, получим

$$\|G(\tau)\|_2 = \max_{i=1,\dots n} |\ \lambda_i(G)| = \max_{i=1,\dots n} |1 - \tau\lambda_i| = \max\{|1 - \tau\lambda_1|, |1 - \tau\lambda_n|\} \quad (11)$$

(максимум модуля собственных чисел матрицы $G(\tau)$ определяется ее максимальным или ее минимальным собственным числом и в итоге зависит от минимального и максимального собственных чисел матрицы A).

Из (11.4) и (11) следует (11.6).

Покажем, что условие (11.7) является достаточным для сходимости.

По теореме 1 для сходимости метода **достаточно** выполнения неравенства

$\|\ G(\tau)\ \|_2 < 1,$

что приводит к условию

$$\|\ G(\tau)\ \|_2 = \max\{\ |1 - \tau\lambda_1\ |, |1 - \tau\lambda_n\ |\ \} < 1 \quad (12)$$

Чтобы выбрать параметр τ, решим систему, состоящую из двух неравенств:

$$\begin{cases} |1-\tau\cdot\lambda_1|<1, \\ |\ 1-\tau\cdot\lambda_n\ |<1 \end{cases}, \qquad (13)$$

Раскрывая модули, запишем

$$-1<1-\tau\cdot\lambda_1<1,$$

$$-1<1-\tau\cdot\lambda_n<1.$$

Учитывая положительность собственных чисел, получим $0<\tau<\dfrac{2}{\lambda_1}$, $0<\tau<\dfrac{2}{\lambda_n}$.

Так как $\lambda_1 \le .\lambda_n$, решением (13) является более сильное ограничение

$$0<\tau<\frac{2}{\lambda_n}.$$

Таким образом, при выполнении условия (11.7), то есть выборе параметра $\tau\in(0,\dfrac{2}{\lambda_n})$, выполнено ограничение $\|\ G(\tau)\ \|_2<1$ и метод (11.2) сходится с оценкой (11.6).

Покажем, как доказывать необходимость. Доказательство основано на том, что оценка (8), а именно

$$\left\|\ z^{(s)}\ \right\|_2=\left\|\ (G(\tau))^s z^{(0)}\ \right\|_2\le\left\|\ (G(\tau))^s\ \right\|_2\left\|\ z^{(0)}\ \right\|_2$$

использует евклидову норму вектора и подчиненную ей («евклидову») норму матрицы. Оценку (8) нельзя улучшить: найдется такое начальное приближение, для которого (8) выполняется как равенство:

$$\left\|\ z^{(s)}\ \right\|_2=\left\|\ (G(\tau))^s z^{(0)}\ \right\|_2=\left\|\ (G(\tau))^s\ \right\|_2\left\|\ z^{(0)}\ \right\|_2 \qquad (14)$$

Для «евклидовых» норм симметричных матриц $G(\tau)$ и $(G(\tau))^s$ верно

$$\left\|\ (G(\tau))^s\ \right\|_2=\|\ G(\tau)\ \|_2^s \qquad (15)$$

Поэтому при $\tau\notin(0,\dfrac{2}{\lambda_n})$ получим $\left\|G(\tau)\right\|_2\ge 1$ и соответственно $\left\|\ (G(\tau))^s\ \right\|_2\ge 1$. Тогда при решении СЛАУ (11.1) найдется такое начальное приближение $x^{(0)}$, у которого такая погрешность $z^{(0)}$, что на каждом шаге метода (11.2) выполняется

$$\left\|\ z^{(s)}\ \right\|_2\ge\left\|\ z^{(0)}\ \right\|_2 \qquad (16)$$

и сходимость метода отсутствует.

К о м м е н т а р и и

Неравенство (16) следует из того, что

$$\| z^{(s)} \|_2 = \| (G(\tau))^s z^{(0)} \|_2 = \| (G(\tau))^s \|_2 \| z^{(0)} \|_2 = \| G(\tau) \|_2^s \| z^{(0)} \|_2,$$

где $\| G(\tau) \|_2 \geq 1$.

Равенство (15) вытекает из того, что $\rho(G) = \max\limits_{i=1,n} | \lambda_i(G) |$

$$\rho(G^s) = \max_{i=1,n} \left| \lambda_i(G^s) \right|, \ \lambda_i(G^s) = [\lambda_i(G)]^s, i = 1,...n, \ \rho(G^s) = [\rho(G)]^s$$

$$\| G \|_2 = \rho(G), \ \| G^s \|_2 = \rho(G^s) = [\rho(G)]^s = \| G \|_2^s.$$

Рассмотрим выбор оптимального параметра

Т е о р е м а 3 . При решении СЛАУ (11.1) с *симметричной, положительно определенной* матрицей $A = A^T > 0$ **методом простой итерации** (11.2) оптимальным является

$$\tau_{opt}^* = \frac{2}{\lambda_1 + \lambda_n} \tag{11.9}$$

для которого оценка погрешности метода на шаге s имеет вид

$$\| z^{(s)} \|_2 \leq \left(\frac{\mu_A - 1}{\mu_A + 1} \right)^s \| z^{(0)} \|_2 \tag{11.10}$$

Здесь $z^{(s)} = x^{(s)} - x^*$ – погрешность метода на шаге s, $z^{(0)} = x^{(0)} - x^*$ – погрешность метода на начальном шаге, $0 < \lambda_1 \leq ... \leq \lambda_n$ – собственные числа матрицы $A = A^T > 0$ (они положительны и упорядочены), λ_1 – минимальное из них, λ_n – максимальное из них.

Через μ_A обозначено число обусловленности матрицы A, определяемое на основе нормы матрицы, подчиненной евклидовой норме вектора (далее кратко – *число обусловленности, определяемое на основе евклидовой нормы*):

$$\mu_A = \| A \|_2 \| A^{-1} \|_2 = \frac{\lambda_n}{\lambda_1} \tag{11.11}$$

К о м м е н т а р и и

Оптимальным считается такое значение τ, при котором метод (11.2) сходится и оценка погрешности метода на шаге s является наилучшей. В данном случае для отыскания оптимального τ ставится задача минимизации

$$\| G(\tau) \|_2 \underset{\tau \in \left(0, \frac{2}{\lambda_n}\right)}{\longrightarrow} \min \tag{11.12}$$

Минимум функционала $\| G(\tau) \|_2$ достигается при $\tau^*_{opt} = \dfrac{2}{\lambda_1 + \lambda_n}$.

Минимальным значением является $\left\| G(\tau^*_{opt}) \right\|_2$, и в ходе доказательства теоремы эта величина будет найдена как

$$\min_{\tau \in R} \| G(\tau) \|_2 = \min_{\tau \in R} \max \{ |1 - \lambda_1 \tau|, \; |1 - \lambda_n \tau| \} =$$

$$= \max \left\{ \left| 1 - \lambda_1 \tau^*_{opt} \right|, \; \left| 1 - \lambda_n \tau^*_{opt} \right| \right\} =$$

$$= \max \left\{ \left| 1 - \frac{\lambda_1 \cdot 2}{\lambda_1 + \lambda_n} \right|, \; \left| 1 - \frac{\lambda_n \cdot 2}{\lambda_1 + \lambda_n} \right| \right\} = \frac{\mu_A - 1}{\mu_A + 1} \qquad (11.13)$$

откуда следует (11.10).

Доказательство

Рассмотрим принцип, в соответствии с которым метод (11.2) с параметром (11.9) считается оптимальным.

При решении СЛАУ (11.1) с матрицей $A = A^T > 0$ методом простой итерации (11.2) при любом выборе параметра τ для погрешности метода верна оценка

$$\left\| z^{(s)} \right\|_2 \le \| G(\tau) \|_2^{\,s} \left\| z^{(0)} \right\|_2 \qquad (17)$$

где в силу симметричности матрицы $G(\tau)$ для ее нормы выполняется

$$\| G(\tau) \|_2 = \max\{|1 - \tau\lambda_1|, |1 - \tau\lambda_n|\} \qquad (18)$$

Для заданной матрицы $A = A^T > 0$ оптимальным назовем такое значение τ, при котором норма переходной матрицы $\| G(\tau) \|_2$ минимальна и тем самым согласно (17) гарантирована наиболее высокая скорость убывания погрешности метода.

Чтобы найти оптимальное значение τ, записывают задачу оптимизации

$$\| G(\tau) \|_2 = \max\{ |1 - \tau \cdot \lambda_1|, |1 - \tau \cdot \lambda_n| \} \to \min_{\tau \in R} \qquad (19)$$

(читается так: нужно найти минимальное значение функционала, определяемого по формуле (18) и зависящего от τ, при условии, что τ принимает любые действительные значения; а также найти τ, при котором достигается указанный выше минимум).

Если $\tau \notin (0, \dfrac{2}{\lambda_n})$, норма переходной матрицы не обеспечивает сходимости: $\|G(\tau)\|_2 \ge 1$. Поэтому минимальное значение функционала можно искать только для сходящихся методов, то есть при $\tau \in (0, \dfrac{2}{\lambda_n})$, и заменить задачу (19) задачей

$$\| G(\tau) \|_2 = \max\{ |1-\tau \cdot \lambda_1|, |1-\tau \cdot \lambda_n| \} \underset{\tau\in(0,\frac{2}{\lambda_n})}{\rightarrow} \min \qquad (20)$$

Тем не менее, в данном случае удобнее решить задачу без ограничений на параметр τ, то есть задачу (19), и потом проверить найденное решение на соответствие условиям сходимости. Решим (19) графически, записав ее в следующих обозначениях:

$$\Phi(\tau) \underset{\tau\in R}{\rightarrow} \min \qquad (21)$$

где $\Phi(\tau) = \max\{ \varphi_1(\tau), \varphi_2(\tau) \}$, $\varphi_1(\tau) = |1-\tau \cdot \lambda_1|$, $\varphi_2(\tau) = |1-\tau \cdot \lambda_n|$

Для $\varphi_1(\tau), \varphi_2(\tau)$ раскроем модули и получим

$$\varphi_1(\tau) = \begin{cases} 1-\tau \cdot \lambda_1, \; если \;\; \tau \le \dfrac{1}{\lambda_1} \\ \tau \cdot \lambda_1 - 1, \; если \; \tau \ge \dfrac{1}{\lambda_1} \end{cases} \qquad \varphi_2(\tau) = \begin{cases} 1-\tau \cdot \lambda_n, \; если \;\; \tau \le \dfrac{1}{\lambda_n} \\ \tau \cdot \lambda_n - 1, \; если \;\; \tau \ge \dfrac{1}{\lambda_n} \end{cases}$$

Функции $\varphi_1(\tau), \varphi_2(\tau)$ показаны на графике зеленым и синим цветом. Красным цветом показана $\Phi(\tau)$ как максимум из указанных двух.

Функции $\varphi_1(\tau), \varphi_2(\tau)$ принимают одинаковые значения при двух значениях аргумента: при $\tau = 0$, когда $\varphi_1(0) = \varphi_2(0) = 1$, и при некотором положительном $\tau = \tau^*$, которое расположено между точками, в которых $\varphi_1(\tau), \varphi_2(\tau)$ обращаются в ноль.

Поэтому $\tau = \tau^*$ не меньше, чем $\dfrac{1}{\lambda_n}$ и не больше, чем $\dfrac{1}{\lambda_1}$

Значение τ^* можно найти из условия $\varphi_1(\tau^*) = \varphi_2(\tau^*)$, которое принимает вид

$$1 - \tau^* \lambda_1 = \tau^* \lambda_n - 1. \qquad (22)$$

позволяет найти $\tau^* = \dfrac{2}{\lambda_1 + \lambda_n}$ и

$$\varphi_1(\tau^*) = \varphi_2(\tau^*) = 1 - \frac{2\lambda_1}{\lambda_1 + \lambda_n} = \frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1} = \frac{\dfrac{\lambda_n}{\lambda_1} - 1}{\dfrac{\lambda_n}{\lambda_1} + 1} = \frac{\mu_A - 1}{\mu_A + 1} \qquad (23)$$

Здесь через μ_A обозначено число обусловленности матрицы A, а именно

$$\mu_A = \| A \|_2 \, \| A^{-1} \|_2 = \frac{\lambda_n}{\lambda_1}$$

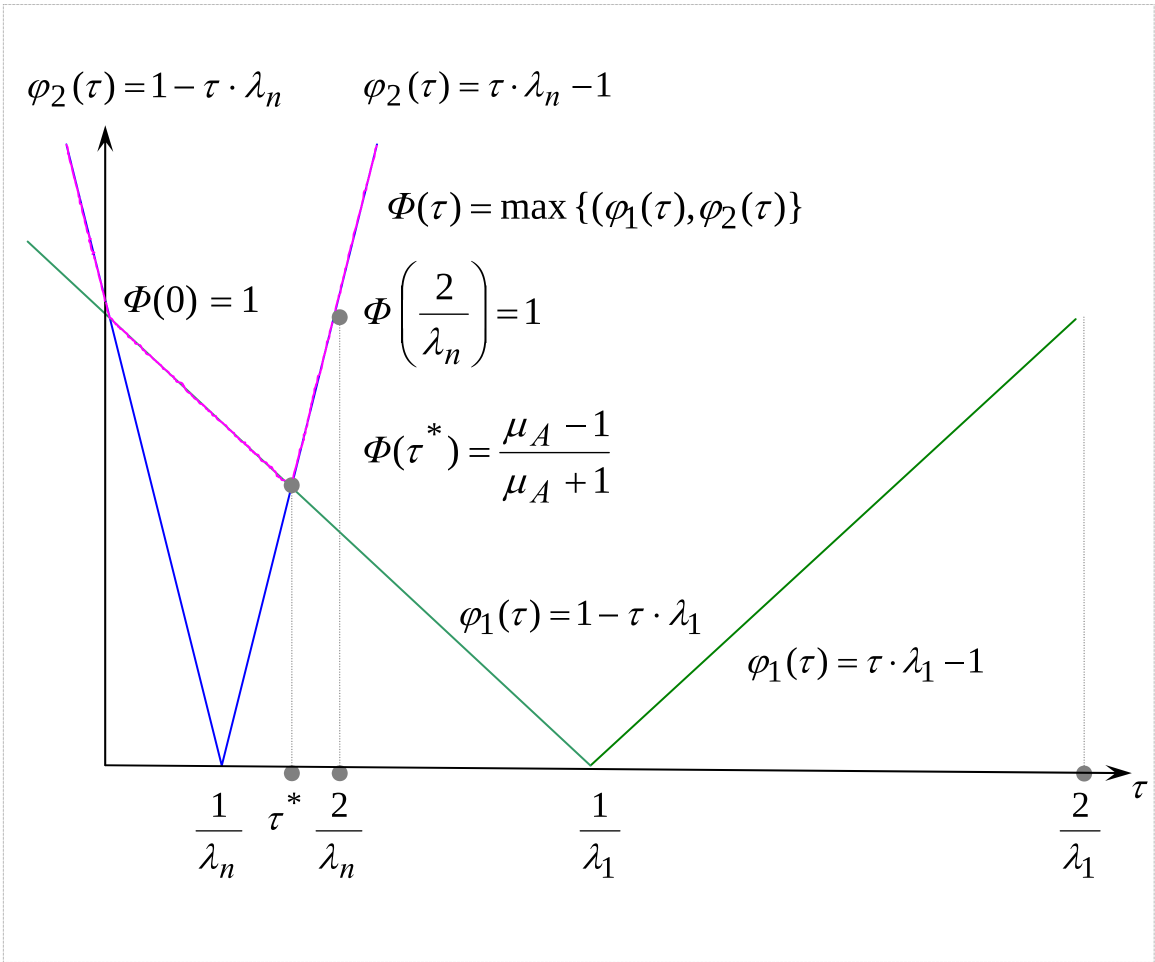

Рисунок 1

Функции $\varphi_1(\tau), \varphi_2(\tau)$ принимают одинаковые значения при двух значениях аргумента: при $\tau = 0$, когда $\varphi_1(0) = \varphi_2(0) = 1,$ и при некотором положительном $\tau = \tau^*$, которое расположено между точками, в которых $\varphi_1(\tau), \varphi_2(\tau)$ обращаются в ноль.

Поэтому $\tau = \tau^*$ не меньше, чем $\frac{1}{\lambda_n}$ и не больше, чем $\frac{1}{\lambda_1}$

Значение τ^* можно найти из условия $\varphi_1(\tau^*) = \varphi_2(\tau^*)$, которое принимает вид

$$1-\tau^*\lambda_1 = \tau^*\lambda_n - 1. \tag{22}$$

и позволяет найти $\tau^* = \frac{2}{\lambda_1 + \lambda_n}$ и

$$\varphi_1(\tau^*) = \varphi_2(\tau^*) = 1 - \frac{2\lambda_1}{\lambda_1+\lambda_n} = \frac{\lambda_n - \lambda_1}{\lambda_n+\lambda_1} = \frac{\frac{\lambda_n}{\lambda_1}-1}{\frac{\lambda_n}{\lambda_1}+1} = \frac{\mu_A-1}{\mu_A+1} \tag{23}$$

Здесь через μ_A обозначено число обусловленности матрицы A, а именно

$$\mu_A = \| A \|_2 \| A^{-1} \|_2 = \frac{\lambda_n}{\lambda_1}$$

Запишем, как устроена функция $\Phi(\tau) = \max\{ \varphi_1(\tau), \varphi_2(\tau) \}$, см. график:

$$\Phi(\tau) = \begin{cases} \varphi_2(\tau) = 1 - \tau\lambda_n, \ \text{если} \ \ \tau \le 0 \\ \varphi_1(\tau) = 1 - \tau\lambda_1, \ \ \text{если} \ \ 0 \le \tau \le \tau^* \\ \varphi_2(\tau) = \tau\lambda_n - 1, \ \text{если} \ \ \tau \ge \tau^* \end{cases} \tag{24}$$

Минимальное значение $\Phi(\tau)$ достигается при $\tau = \tau^*$, причем

$$\Phi(\tau^*) = \frac{\mu_A - 1}{\mu_A + 1} \tag{25}$$

Решением задачи (19) является $\tau = \tau^*$, причем минимально возможная норма переходной матрицы составит

$$\left\| G(\tau^*) \right\|_2 = \frac{\mu_A - 1}{\mu_A + 1} \tag{26}$$

Завершим доказательство формальными проверками.

Значение τ^* далее называем оптимальным и обозначаем τ^*_{opt}.

Метод (11.2) с параметром (11.9) является сходящимся, потому что τ^*_{opt} соответствует условиям сходимости (11.7):

$$\tau^*_{opt} = \frac{2}{\lambda_1 + \lambda_n} \in (0, \frac{2}{\lambda_n}). \tag{27}$$

Для погрешности метода (11.2) при любом выборе параметра верна оценка (11.4), и при выборе параметра τ^*_{opt} запишем:

$$\left\| z^{(s)} \right\|_2 \le \left\| G(\tau^*_{opt}) \right\|_2^s \left\| z^{(0)} \right\|_2 \tag{28}$$

При выборе параметра τ^*_{opt} норма переходной матрицы является минимальной и в силу (26)

$$\left\| G(\tau^*_{opt}) \right\|_2 = \frac{\mu_A - 1}{\mu_A + 1}. \tag{29}$$

Через μ_A обозначено число обусловленности матрицы A, а именно $\mu_A = \| A \|_2 \| A^{-1} \|_2 = \frac{\lambda_n}{\lambda_1}$

Из (28) и (29) следует оценка (11.10):

$$\left\| z^{(s)} \right\|_2 \le \left(\frac{\mu_A - 1}{\mu_A + 1} \right)^s \left\| z^{(0)} \right\|_2. \qquad (30)$$

Так как для любой невырожденной матрицы число обусловленности μ_A не меньше 1, получим

$$0 \le \frac{\mu_A - 1}{\mu_A + 1} \le 1 \qquad (31)$$

При $s \to \infty$ $\left(\frac{\mu_A - 1}{\mu_A + 1} \right)^s \to 0$ и погрешность метода $\left\| z^{(s)} \right\|_2 \to 0$.

Комментарии (после доказательства)

1) Чтобы найти оптимальный параметр τ^*_{opt}, нужно знать минимальное и максимальное собственные числа матрицы A.

2) Если $\mu_A \approx 1$, то $\left(\frac{\mu_A - 1}{\mu_A + 1} \right) \approx 0$ и метод сходится быстро. Если $\mu_A >> 1$, то $\left(\frac{\mu_A - 1}{\mu_A + 1} \right) \approx 1$ и сходимость медленная. Чем выше число обусловленности матрицы A, тем медленнее сходится метод.

3) Существуют такие начальные приближения $x^{(0)} \in R^n$, для которых $\left\| z^{(s)} \right\|_2 = \left(\frac{\mu_A - 1}{\mu_A + 1} \right)^s \left\| z^{(0)} \right\|_2$, и это означает, что оценку (1.10) нельзя улучшить.

Построение метода на основе оценок собственных чисел

Рассмотрим СЛАУ (11.1) с **симметричной, положительно определенной матрицей**. Как следует из Теоремы 2, чтобы построить сходящийся метод простой итерации, нужно знать максимальное собственное число (см. (11.7)), чтобы оценить погрешность метода – знать минимальное и максимальное собственные числа (см. (11.6)).

Покажем, как построить сходящийся метод **на основе оценок собственных чисел.**

Пусть собственные числа матрицы $A = A^T > 0$ неизвестны, но известны их оценки, а именно, положительные числа $M_{\min} > 0$ и $M_{\max} > 0$, такие, что

$\lambda_i(A) \in [M_{\min}, M_{\max}]$, $i = 1,...n$.

Тогда для λ_1 и λ_n – минимального и максимального собственных чисел – верно

$$0 < M_{\min} \le \lambda_1 \le \lambda_n \le M_{\max} \qquad (11.14)$$

(оценку вида (11.14) называют **оценкой границ спектра**).

Теорема 4. При решении СЛАУ (11.1) с *симметричной, положительно определенной* матрицей $A = A^T > 0$ **методом простой итерации** (11.2) в случае, когда известна оценка границ спектра $0 < M_{\min} \le \lambda_1 \le \lambda_n \le M_{\max}$, достаточным условием сходимости метода является

$$\tau \in (0, \frac{2}{M_{max}}) \tag{11.15}$$

Оценка погрешности метода на шаге s имеет вид

$$\left\| z^{(s)} \right\|_2 \le \left(\max\{\left|1-\tau M_{\min}\right|, \left|1-\tau M_{\max}\right|\}\right)^s \left\| z^{(0)} \right\|_2 \tag{11.16}$$

Здесь $z^{(s)} = x^{(s)} - x^*$ – погрешность метода на шаге s, $z^{(0)} = x^{(0)} - x^*$– погрешность метода на начальном шаге, $0 < \lambda_1 \le \ldots \le \lambda_n$ – собственные числа матрицы $A = A^T > 0$ (они положительны и упорядочены), λ_1 – минимальное из них, λ_n – максимальное из них.

Доказательство

Используем Теорему 2, условие сходимости (11.7) и оценку погрешности (11.6). Так как $0 < \frac{2}{M_{\max}} < \frac{2}{\lambda_n}$, интервал $(0, \frac{2}{M_{\max}})$ принадлежит области сходимости метода.

Далее для каждого фиксированного значения τ очевидна оценка

$$\max\{\left|1-\lambda_1\tau\right|, \left|1-\lambda_n\tau\right|\} \le \max_{\lambda\in[\lambda_1,\lambda_n]} \left|1-\lambda\tau\right| \le \max_{\lambda\in[M_{\min}, M_{\max}]} \left|1-\lambda\tau\right|$$

Здесь максимальное из двух неизвестных чисел $\left|1-\lambda_1\tau\right|, \left|1-\lambda_n\tau\right|$ оценивается максимальным значением функции $\left|1-\lambda\tau\right|$ с аргументом λ, пробегающим отрезок $\lambda \in [\lambda_1, \lambda_n]$, и затем оценивается максимальным значением той же функции $\left|1-\lambda\tau\right|$ с аргументом λ, пробегающим отрезок $\lambda \in [M_{\min}, M_{\max}]$, который включает в себя и значения λ_1 и λ_n, и отрезок $\lambda \in [\lambda_1, \lambda_n]$.

Известно, что максимальное значение непрерывной функции на отрезке достигается либо на концах отрезка, либо в точке локального максимума. Так как при фиксированном τ функция $\left|1-\lambda\tau\right|$ с аргументом $\lambda \in R$ непрерывна и не имеет локального максимума, максимальное значение функции $\left|1-\lambda\tau\right|$ достигается на концах отрезка. Поэтому

$$\max_{\lambda\in[M_{\min}, M_{\max}]} \left|1-\lambda\tau\right| = \max\{\left|1-\tau M_{\min}\right|, \left|1-\tau M_{\max}\right|\}$$

Таким образом, получена оценка

$$\max\{\left|1-\lambda_1\tau\right|, \left|1-\lambda_n\tau\right|\} \le \max\{\left|1-\tau\cdot M_{\min}\right|, \left|1-\tau\cdot M_{\max}\right|\} \tag{11.17}$$

Подставляя (11.17) в (11.6), получим (11.16).

Выбор оптимального параметра на основе оценок собственных чисел

Рассмотрим СЛАУ (11.1) с **симметричной, положительно определенной матрицей**. Как следует из Теоремы 3, чтобы построить метод простой итерации с оптимальной оценкой сходимости, нужно знать минимальное и максимальное собственные числа матрицы.

Рассмотрим возможности оптимизации на основе оценок собственных чисел.

Пусть собственные числа матрицы $A = A^T > 0$ неизвестны, но известны их оценки, а именно, положительные числа $M_{\min} > 0$ и $M_{\max} > 0$, такие, что

$\lambda_i(A) \in [M_{\min}, M_{\max}]$, $i = 1,...n$.

Тогда для λ_1 и λ_n – минимального и максимального собственных чисел – верно

$$0 < M_{\min} \le \lambda_1 \le \lambda_n \le M_{\max}$$

и для числа обусловленности μ_A, определяемого на основе евклидовой нормы, верно

$$1 \le \mu_A = \frac{\lambda_n}{\lambda_1} \le \frac{M_{\max}}{M_{\min}} = M \qquad (1.18)$$

Теорема 5. При решении СЛАУ (11.1) с *симметричной, положительно определенной* матрицей $A = A^T > 0$ **методом простой итерации** в случае, когда известна оценка границ спектра $0 < M_{\min} \le \lambda_1 \le \lambda_n \le M_{\max}$, метод (11.2) с параметром

$$\tilde{\tau}^* = \frac{2}{M_{\min} + M_{\max}} \qquad (11.19)$$

имеет оптимальные свойства и сходится с оценкой

$$\| z^{(s)} \|_2 \le \left(\frac{M-1}{M+1} \right)^s \| z^{(0)} \|_2 \qquad (11.20)$$

Здесь $z^{(s)} = x^{(s)} - x^*$ – погрешность метода на шаге s, $z^{(0)} = x^{(0)} - x^*$ – погрешность метода на начальном шаге, $0 < \lambda_1 \le ... \le \lambda_n$ – собственные числа матрицы $A = A^T > 0$ (они положительны и упорядочены), λ_1 – минимальное из них, λ_n – максимальное из них, число M является верхней оценкой числа обусловленности μ_A (определенного в евклидовой норме).

Доказательство

Используем Теорему 4, условие сходимости (11.15) и оценку погрешности (11.16).

Во-первых, значение $\tau = \tilde{\tau}^*$ принадлежит области сходимости метода:

$$\tilde{\tau}^* = \frac{2}{M_{\min} + M_{\max}} \in (0, \frac{2}{M_{\max}}) \subset (0, \frac{2}{\lambda_n}).$$

Во-вторых, подстановкой значения $\tau = \tilde{\tau}^*$ можно показать, что

$$\max\{|1 - \tilde{\tau}^* M_{\min}|, |1 - \tilde{\tau}^* M_{\max}|\} = \frac{M_{\max} - M_{\min}}{M_{\max} + M_{\min}} = \frac{M-1}{M+1}$$

поэтому из (11.16) следует (11.20).

Комментарий

Оптимальные свойства метода вытекают из следующих обстоятельств. Рассмотрим класс методов (11.2) с параметром (11.15):

$$\tau \in (0, \frac{2}{M_{\max}})$$

и оценкой погрешности (11.16):

$$\left\| z^{(s)} \right\|_2 \le \left(\max\{|1-\tau M_{\min}|, |1-\tau M_{\max}|\}\right)^s \left\| z^{(0)} \right\|_2$$

Оптимальным считается такое значение τ, при котором метод сходится и оценка погрешности метода на шаге s (см. (11.16)) является в некотором смысле оптимальной. В данном случае для отыскания оптимального значения τ ставится задача минимизации

$$\max\{|1-\tau M_{\min}|, |1-\tau M_{\max}|\} \underset{\tau \in \left(0, \frac{2}{M_{\max}}\right)}{\longrightarrow} \min \qquad (11.21)$$

Используя графики функций $|1-\tau M_{\min}|$ и $|1-\tau M_{\max}|$ с аргументом $\tau \in R$, несложно показать, что минимальное значение функционала достигается при

$$\tau = \tilde{\tau}^* = \frac{2}{M_{\min} + M_{\max}}$$

При этом минимальное значение функционала

$$\max\{|1-\tau M_{\min}|, |1-\tau M_{\max}|\}$$

может быть найдено подстановкой значения $\tau = \tilde{\tau}^*$:

$$\min_{\tau \in R} \max\{|1-\tau M_{\min}|, |1-\tau M_{\max}|\} =$$

$$= \max\{|1-\tilde{\tau}^* \cdot M_{\min}|, |1-\tilde{\tau}^* \cdot M_{\max}|\} = \frac{M-1}{M+1} \qquad (11.22)$$

Комментарий к параграфу 11.1

В формулировках теорем 1-5 погрешность метода на шаге s оценивается начальной погрешностью, см. (11.4), (11.6), (11.10), (11.16), 11.20), все оценки в евклидовой норме. В свою очередь, **погрешность метода на начальном шаге можно оценить по начальной невязке**, используя норму обратной матрицы или ее оценку:

$\left\| z^{(0)} \right\|_2 \le \left\| A^{-1} \right\|_2 \cdot \left\| r^{(0)} \right\|_2$, где $r^{(0)} = Ax^{(0)} - b$ и ее можно вычислить.

11.2. Метод минимальных невязок

Теоремы о выборе параметра, сходимости метода и оценке погрешности на основе собственных чисел или их оценок

Рассмотрим СЛАУ (систему линейных алгебраических уравнений) вида

$$Ax = b, \tag{11.1}$$

где $x \in R^n$, $b \in R^n$, $A\,(n \times n)$, $\det A \neq 0$ (невырожденная матрица).

Через x^* обозначим точное решение системы, $x^* \in R^n$.

Методом минимальных невязок называют явный нестационарный итерационный метод

$$\frac{x^{(s+1)} - x^{(s)}}{\tau_s} + Ax^{(s)} = b \tag{11.23}$$

где параметр τ_s (число) на каждом шаге метода выбирают так, чтобы для уже вычисленного приближения $x^{(s)} \in R^n$ следующая невязка $r^{(s+1)} \in R^n$, то есть $r^{(s+1)} = Ax^{(s+1)} - b$, была минимальной.

Если $A = A^T > 0$ и невязка измеряется в евклидовой норме, следует выбирать

$$\tau_s = \frac{(Ar^{(s)}, r^{(s)})}{(Ar^{(s)}, Ar^{(s)})} \tag{11.24}$$

Запись метода в виде (11.23), (1.24) называется *канонической*. Для расчетов вместо (11.23) используется формула

$$x^{(s+1)} = x^{(s)} + \tau_s \cdot (b - Ax^{(s)}) = x^{(s)} - \tau_s \cdot r^{(s)} \tag{11.25}$$

где $r^{(s)} = Ax^{(s)} - b$ – невязка СЛАУ на текущем приближении $x^{(s)}$, τ_s – параметр шага $s+1$.

Теорема 6. При решении СЛАУ (11.1) с *симметричной, положительно определенной* матрицей $A = A^T > 0$ **методом минимальных невязок** (11.23), (11.24) метод сходится и для погрешности метода на шаге s верна оценка

$$\left\| z^{(s)} \right\|_2 \le \mu_A \left(\frac{\mu_A - 1}{\mu_A + 1} \right)^s \left\| z^{(0)} \right\|_2 \tag{11.26}$$

Здесь $z^{(s)} = x^{(s)} - x^*$ – погрешность метода на шаге s, $z^{(0)} = x^{(0)} - x^*$ – погрешность метода на начальном приближении, $0 < \lambda_1 \le \ldots \le \lambda_n$ – собственные числа матрицы $A = A^T > 0$ (они положительны и упорядочены), λ_1 – минимальное из них, λ_n – максимальное из них. Через μ_A обозначено число обусловленности матрицы A, определяемое на основе евклидовой нормы:

$$\mu_A = \| A \|_2 \, \| A^{-1} \|_2 = \frac{\lambda_n}{\lambda_1}.$$

Д о к а з а т е л ь с т в о

Часть I

В доказательстве сходимости метода (11.23) и обосновании оценки погрешности (11.26) используется следующее обстоятельство: параметр метода τ_s, определяемый по формуле (11.24), минимизирует невязку шага $s+1$.

Поэтому сначала докажем формулу (11.24).

По формуле (11.25) запишем шаг метода:

$$x^{(s+1)} = x^{(s)} - \tau_s r^{(s)} \tag{1}$$

Умножим левую и правую части равенства (1) на матрицу A, $\det A \neq 0$:

$$Ax^{(s+1)} = Ax^{(s)} - \tau_s Ar^{(s)}. \tag{2}$$

Вычтем слева и справа $b \in R^n$ (правую часть СЛАУ):

$$Ax^{(s+1)} - b = (Ax^{(s)} - b) - \tau_s Ar^{(s)}. \tag{3}$$

Полученный результат перепишем в обозначениях невязки:

$$r^{(s+1)} = r^{(s)} - \tau_s Ar^{(s)} \tag{4}$$

Уравнение (4) показывает, как меняется невязка СЛАУ на очередном шаге метода. Чтобы оценить такое изменение, запишем (4) в виде

$$r^{(s+1)} = (E - \tau_s A)r^{(s)} \tag{5}$$

где $E - \tau_s A$ есть **переходная матрица метода на шаге** $s+1$.

Указанную матрицу было бы логично обозначить $G(\tau_s) = E - \tau_s A$ и записывать зависимость невязок в форме $r^{(s+1)} = G(\tau_s)\, r^{(s)}$.

Но для доказательства удобнее использовать (5).

Напомним **принцип**, по которому построен метод минимальных невязок: параметр τ_s выбирают так, чтобы для уже вычисленного приближения $x^{(s)} \in R^n$ невязка на следующем приближении $x^{(s+1)} \in R^n$, то есть $r^{(s+1)} \in R^n$, была минимальной. Чтобы найти такой τ_s, рассмотрим задачу минимизации *евклидовой нормы невязки*:

$$\left\| r^{(s+1)} \right\|_2 \to \min \tag{6}$$

где $\left\| r^{(s+1)} \right\|_2 = \sqrt{(r^{(s+1)}, r^{(s+1)})}$.

Задачу минимизации нормы, основанной на скалярном произведении, принято решать как задачу минимизации квадрата нормы:

$$\left\| r^{(s+1)} \right\|_2^2 \to \min \tag{7}$$

В задаче (7) функционал, значение которого нужно минимизировать, является гладким, а параметр, при котором достигается минимум (7), совпадает с параметром, при котором достигается минимум (6).

Запишем квадрат евклидовой нормы невязки как скалярный квадрат:

$$\left\| r^{(s+1)} \right\|_2^2 = (r^{(s+1)}, r^{(s+1)}) \tag{8}$$

где в силу представления (5)

$$\begin{aligned}(r^{(s+1)}, r^{(s+1)}) &= ((E-\tau_s A)r^{(s)}, (E-\tau_s A)r^{(s)}) = \\ &= (r^{(s)}, r^{(s)}) - 2\tau_s (Ar^{(s)}, r^{(s)}) + \tau_s^2 (Ar^{(s)}, Ar^{(s)})\end{aligned} \tag{9}$$

Рассмотрим шаг s, на котором вычислено приближение $x^{(s)} \in R^n$. Тогда

$(r^{(s)}, r^{(s)}),$

$(Ar^{(s)}, r^{(s)}),$,

$(Ar^{(s)}, Ar^{(s)})$

есть числа, их значения известны и выражение (9) есть полином степени не выше 2 с аргументом τ_s.

В связи с этим задача (8) – задача *минимизации квадрата евклидовой нормы* – есть задача отыскания *минимального значения полинома степени не выше* 2:

$$(r^{(s)}, r^{(s)}) - 2\tau_s (Ar^{(s)}, r^{(s)}) + \tau_s^2 (Ar^{(s)}, Ar^{(s)}) \underset{\tau_s \in R}{\to} \min \tag{10}$$

Рассмотрим два случая:

1) Если так вышло, что на шаге s было найдено точное решение СЛАУ, то есть $x^{(s)} = x^*$, тогда $r^{(s)} = 0$ (потому что невязка $Ax^* - b$ равна нулю).

В этом случае

$(r^{(s)}, r^{(s)}) = 0,$

$(Ar^{(s)}, r^{(s)}) = 0,$

$(Ar^{(s)}, Ar^{(s)}) = 0$

и невязка следующего шага в силу (5) (или в силу (9)) при любом выборе параметра τ_s равна нулю: $r^{(s+1)} = (E - \tau_s A)r^{(s)}$.

Заметим, что формула (11.24) в этом случае не работает: знаменатель обращается в ноль. Поэтому в случае, если точное решение найдено, в программной реализации метода дальнейшие шаги должны быть остановлены.

2) Если на шаге s точное решение задачи СЛАУ еще не найдено, то есть $x^{(s)} \neq x^*$, тогда $r^{(s)} \neq 0$ и коэффициенты полинома (9) таковы:

$(r^{(s)}, r^{(s)}) > 0$ в силу $r^{(s)} \neq 0$;

$(Ar^{(s)}, r^{(s)}) > 0$, в силу положительной определенности матрицы $A > 0$;

$(Ar^{(s)}, Ar^{(s)}) > 0$ в силу того, что $r^{(s)} \neq 0$ и $\det A \neq 0$.

В этом случае (9) представляет собой полином степени 2 с аргументом τ_s. Ветви параболы направлены вверх и решением задачи (10) является аргумент вершины параболы:

$$\tau_s = \frac{(Ar^{(s)}, r^{(s)})}{(Ar^{(s)}, Ar^{(s)})} \tag{10}$$

Минимальное значение функционала (8) найдем после подстановки (10) в (9), и для указанного значения верно

$$0 \leq (r^{(s+1)}, r^{(s+1)}) = (r^{(s)}, r^{(s)}) - \frac{(Ar^{(s)}, r^{(s)})^2}{(Ar^{(s)}, Ar^{(s)})} < (r^{(s)}, r^{(s)}) \tag{11}$$

На каждом шаге метода норма невязки убывает.

Часть II

Для изучения сходимости метода напомним уравнение (5)

$$r^{(s+1)} = (E - \tau_s A)\, r^{(s)}$$

и запишем евклидову норму левой и правой части данного уравнения:

$$\left\| r^{(s+1)} \right\|_2 = \left\| (E - \tau_s A)\, r^{(s)} \right\|_2.$$

Параметр τ_s, определенный формулой (10), соответствует минимальному значению функционала (7) и минимальному значению функционала (6).

Поэтому для $\forall \tau \neq \tau_s$ верно

$$\left\| (E - \tau_s \cdot A)\, r^{(s)} \right\| \leq \left\| (E - \tau \cdot A)\, r^{(s)} \right\|_2 \tag{12}$$

Перепишем (12) для $\tau = \tau^*_{opt}$ – оптимального параметра метода простой итерации:

$$\left\| (E - \tau_s \cdot A)\, r^{(s)} \right\| \leq \left\| (E - \tau^*_{opt} \cdot A)\, r^{(s)} \right\|_2 \tag{13}$$

Выражение в правой части (13) есть евклидова норма переходной матрицы $G(\tau^*_{opt})$ метода простой итерации с оптимальным параметром τ^*_{opt}, см. (11.1):

$$\left\| G(\tau^*_{opt}) \right\|_2 = \left\| (E - \tau^*_{opt} \cdot A)\, r^{(s)} \right\|_2 = \frac{\mu_A - 1}{\mu_A + 1} \tag{14}$$

Указанная норма выражается через μ_A – число обусловленности матрицы A, определяемое как

$$\mu_A = \| A \|_2 \, \| A^{-1} \|_2 = \frac{\lambda_n}{\lambda_1}. \tag{15}$$

Продолжим оценку (13), используя свойство согласованности норм и формулу (14):

$$\left\| (E - \tau_s \cdot A)\, r^{(s)} \right\|_2 \le \left\| (E - \tau_{opt}^{*} \cdot A) r^{(s)} \right\|_2 \le$$

$$\le \left\| (E - \tau_{opt}^{*} \cdot A) \right\|_2 \cdot \left\| r^{(s)} \right\|_2 = \left\| G(\tau_{opt}^{*}) \right\|_2 \cdot \left\| r^{(s)} \right\|_2$$

$$= \left(\frac{\mu_A - 1}{\mu_A + 1} \right) \cdot \left\| r^{(s)} \right\|_2$$

Таким образом, в методе минимальных невязок

$$\left\| r^{(s+1)} \right\|_2 \le \left(\frac{\mu_A - 1}{\mu_A + 1} \right) \cdot \left\| r^{(s)} \right\|_2 \le \ldots \le \left(\frac{\mu_A - 1}{\mu_A + 1} \right)^{s+1} \left\| r^{(0)} \right\|_2 \tag{16}$$

Неравенство (16) свидетельствует о **сходимости метода по невязке**:

при $s \to +\infty$ $\left(\frac{\mu_A - 1}{\mu_A + 1} \right)^{s+1} \to 0$ и $\left\| r^{(s+1)} \right\|_2 \to 0$.

Покажем **сходимость для погрешности** и докажем ее оценку.

На каждом шаге l любого метода для невязки и погрешности верно

$$Az^{(l)} = r^{(l)} \tag{17}$$

$$A^{-1} r^{(l)} = z^{(l)} \tag{18}$$

Для индекса $l = 0$ из (17) получим

$$\left\| r^{(0)} \right\|_2 = \left\| Az^{(0)} \right\|_2 \le \left\| A \right\|_2 \cdot \left\| z^{(0)} \right\|_2 \tag{19}$$

Для индекса $l = s + 1$ из (18) получим

$$\left\| z^{(s+1)} \right\|_2 = \left\| A^{-1} r^{(s+1)} \right\|_2 \le \left\| A^{-1} \right\|_2 \cdot \left\| r^{(s+1)} \right\|_2 \tag{20}$$

Умножим левую и правую части неравенства (16) на норму обратной матрицы:

$$\left\| A^{-1} \right\|_2 \cdot \left\| r^{(s+1)} \right\|_2 \le \left\| A^{-1} \right\|_2 \cdot \left(\frac{\mu_A - 1}{\mu_A + 1} \right)^{s+1} \left\| r^{(0)} \right\|_2 \tag{21}$$

Продолжим неравенство (21) вправо, используя (19), и влево, используя (20):

$$\left\| z^{(s+1)} \right\|_2 \le \left\| A^{-1} \right\|_2 \cdot \left\| r^{(s+1)} \right\|_2 \le$$

$$\le \left\| A^{-1} \right\|_2 \cdot \left(\frac{\mu_A - 1}{\mu_A + 1} \right)^{s+1} \left\| r^{(0)} \right\|_2 \le \tag{22}$$

$$\le \left\| A^{-1} \right\|_2 \cdot \left\| A \right\|_2 \cdot \left(\frac{\mu_A - 1}{\mu_A + 1} \right)^{s+1} \left\| z^{(0)} \right\|_2$$

С учетом (15) для погрешности метода минимальных невязок получена оценка

$$\left\| z^{(s+1)} \right\|_2 \le \mu_A \cdot \left(\frac{\mu_A - 1}{\mu_A + 1} \right)^{s+1} \left\| z^{(0)} \right\|_2 \tag{23}$$

где μ_A – число обусловленности.

Оценка (11.26) доказана, и **сходимость метода** (11.23), (11.24) следует из нее:

при $s \to +\infty$ $\left(\frac{\mu_A - 1}{\mu_A + 1} \right)^{s+1} \to 0$ и $\left\| z^{(s+1)} \right\|_2 \to 0$.

Оценка погрешности на основе оценок собственных чисел

Рассмотрим СЛАУ (11.1) с **симметричной, положительно определенной матрицей**. Как следует из Теоремы 6, для решения СЛАУ методом минимальных невязок знать собственные числа матрицы A не требуется, но для оценки погрешности метода – см. (11.26) – такая информация необходима.

Теорема 7. Пусть собственные числа *симметричной, положительно определенной* матрицы $A = A^T > 0$ неизвестны, но известны их оценки, то есть числа $M_{\min} > 0$ и $M_{\max} > 0$, такие, что $\lambda_i(A) \in [M_{\min}, M_{\max}]$, $i = 1,\ldots n$. Тогда для λ_1 и λ_n верно $0 < M_{\min} \le \lambda_1 \le \lambda_n \le M_{\max}$, а для числа обусловленности μ_A (основанного на евклидовой норме) верна оценка

$$1 \le \mu_A = \frac{\lambda_n}{\lambda_1} \le \frac{M_{\max}}{M_{\min}} = M \tag{11.27}$$

Для оценки погрешности метода минимальных невязок можно использовать неравенство

$$\left\| z^{(s)} \right\|_2 \le M \left(\frac{M - 1}{M + 1} \right)^s \left\| z^{(0)} \right\|_2 \tag{11.28}$$

Доказательство

Так как

$1 \le \mu_A \le M$ и

$$0 \le \frac{\mu_A - 1}{\mu_A + 1} \le \frac{M - 1}{M + 1} < 1$$

(оценивается возрастающая функция вида $\varphi(x)=\dfrac{x-1}{x+1}$ при $x \ge 1$)

из (11.26) получим

$$\left\| z^{(s)} \right\|_2 \le \mu_A \left(\frac{\mu_A - 1}{\mu_A + 1} \right)^s \left\| z^{(0)} \right\|_2 \le M \left(\frac{M-1}{M+1} \right)^s \left\| z^{(0)} \right\|_2$$

что и требовалось доказать.

К о м м е н т а р и й

1) Для СЛАУ с симметричной, положительно определенной матрицей метод (11.23), (11.24) сходится с оценкой (11.26) независимо от того, насколько точно известны собственные числа и известны ли они.

2) Если они известны неточно, вместо «реальной» оценки погрешности (11.26) приходится использовать более грубое неравенство (11.28), которое основано на оценке (11.27).

11.3. Метод с чебышевским набором параметров (k параметров)

Выбор параметров на основе собственных чисел или их оценок. Способ применения метода. Теоремы об оптимальных свойствах, оценках погрешности и сходимости метода. Пример задачи о построении полинома, наименее уклоняющегося от нуля. Пример построения метода.
Итерационный метод с чебышевским набором параметров (k – число параметров) на основе оценок границ спектра: выбор параметров, теорема о сходимости метода, оценках сходимости и оптимальных свойствах метода (формулировка)

Рассмотрим СЛАУ (11.1) с симметричной, положительно определенной матрицей A

$$Ax = b, \tag{11.1}$$

где $x \in R^n, b \in R^n, A\,(n \times n), A = A^T > 0$.

Через x^* обозначим точное решение системы, $x^* \in R^n$.

Пусть $0 < \lambda_1 \le \ldots \le \lambda_n$ – собственные числа матрицы $A = A^T > 0$ (они положительны и упорядочены), λ_1 – минимальное и λ_n – максимальное из них.

Явным нестационарным итерационным методом с чебышевским набором параметров (k параметров) называют метод

$$\frac{x^{(s+1)} - x^{(s)}}{\tau_s} + Ax^{(s)} = b \tag{11.29}$$

где k – натуральное число (параметр метода), $\tau_0, \tau_1, \ldots \tau_{k-1}$ – вещественные параметры метода (их количество равно k), $x^{(0)} \in R^n$ – начальное приближение, которое можно выбрать любым, $s = 0,1,\ldots$ – номер шага метода.

Значения параметров $\tau_0, \tau_1, \ldots \tau_{k-1}$ определяют как

$$\tau_s = \left(\frac{\lambda_1 + \lambda_n}{2} + \frac{\lambda_n - \lambda_1}{2} \cdot \cos\left(\frac{\pi}{2k}(1 + 2s) \right) \right)^{-1}, \; s = 0,\ldots k-1 \tag{11.30}$$

то есть (при $\lambda_1 < \lambda_n$) как величины, обратные корням полинома Чебышева – полинома степени k, наименее уклоняющегося от нуля на отрезке $[\lambda_1, \lambda_n], 0 < \lambda_1 < \lambda_n$ в классе полиномов степени k со свободным слагаемым, равным 1.

Для матрицы $A = A^T > 0$ случай $0 < \lambda_1 = \lambda_n$ не является типичным и заслуживает отдельного рассмотрения, так как в данном случае все собственные числа одинаковы: $0 < \lambda_1 = \lambda_2 = \ldots = \lambda_{n-1} = \lambda_n$ и каждый вектор $x \in R^n$ является собственным

Первые k шагов метода проводятся с использованием параметров $\tau_0, \tau_1, \ldots \tau_{k-1}$. При вычислении $x^{(1)}$ используют τ_0, при вычислении $x^{(2)}$ используют τ_1, ... при вычислении $x^{(k)}$ используют τ_{k-1}.

На последующих шагах метода указанные выше k параметров используют циклически:

при вычислении $x^{(k+1)}$ снова используется τ_0, при вычислении $x^{(k+2)}$ используем τ_1, … при вычислении $x^{(2k)}$ используется τ_{k-1} и т.д.

Таким образом, метод (11.29), (11.30) строит последовательность

$$x^{(0)}, x^{(1)}, x^{(2)}, \ldots x^{(k-1)}, x^{(k)} ..$$

$$x^{(k+1)}, x^{(k+2)}, x^{(k+3)}, \ldots x^{(2k-1)}, x^{(2k)} ..$$

$$x^{(2k+1)}, x^{(2k+2)}, x^{(2k+3)}, \ldots x^{(3k-1)}, x^{(3k)} .. \ldots \tag{11.31}$$

Приближенными решениями задачи (11.1) следует считать следующие элементы последовательности:

$$x^{(0)}, x^{(k)}, x^{(2k)}, .x^{(3k)}, x^{(4k)} \ldots x^{(Nk)} \ldots.. \tag{11.32}$$

Остальные элементы последовательности имеют вспомогательное значение.

Запись метода в виде (11.29), (11.30) называется канонической. Для расчетов вместо (11.29) используется формула

$$x^{(s+1)} = x^{(s)} + \tau_s \cdot (b - Ax^{(s)}) = x^{(s)} - \tau_s \cdot r^{(s)} \tag{11.33}$$

где $r^{(s)} = Ax^{(s)} - b$ – невязка СЛАУ на текущем приближении $x^{(s)}$, τ_s – параметр шага $s+1$.

Свойства метода описывает следующая теорема.

Теорема 8. При решении задачи (11.1) с симметричной, положительно определенной матрицей $A = A^T > 0$ методом (11.29) с параметрами (11.30) оценка погрешности метода на шаге k имеет вид

$$\| z^{(k)} \|_2 \le \frac{2\rho^k}{1+\rho^{2k}} \| z^{(0)} \|_2 \tag{11.34}$$

Оценка погрешности метода на шаге Nk (N есть количество циклов) имеет вид

$$\| z^{(Nk)} \|_2 \le \left(\frac{2\rho^k}{1+\rho^{2k}} \right)^N \| z^{(0)} \|_2 \tag{11.35}$$

и метод сходится в следующем смысле:

$$\forall x^{(0)} \in R^n \quad \lim_{N \to +\infty} \| z^{(Nk)} \|_2 = 0. \tag{11.36}$$

(к решению x^* сходятся элементы последовательности (11.32)).

В формулах (11.34)-(11.36) $z^{(Nk)} = x^{(Nk)} - x^*$ – погрешность метода на шаге Nk, $z^{(0)} = x^{(0)} - x^*$– погрешность метода на начальном шаге, $0 < \lambda_1 \le \ldots \le \lambda_n$ – собственные числа матрицы $A = A^T > 0$ (они положительны и упорядочены), λ_1 – минимальное и λ_n – максимальное из них.

Значение ρ в формулах (11.34), (11.35) определено как

$$\rho = \frac{\sqrt{\mu_A} - 1}{\sqrt{\mu_A} + 1} \tag{11.37}$$

Через μ_A обозначено число обусловленности матрицы A, определяемое на основе евклидовой нормы:

$$\mu_A = \| A \|_2 \, \| A^{-1} \|_2 = \frac{\lambda_n}{\lambda_1} \tag{11.38}$$

Оптимальное свойство метода состоит в следующем: среди всех методов вида (11.29) метод с параметрами (11.30) дает наилучшую гарантию убывания погрешности через k шагов, справедливую для всех симметричных положительно определенных матриц $A = A^T > 0$, собственные числа которых расположены в диапазоне $[\lambda_1, \lambda_n], 0 < \lambda_1 < \lambda_n$.

Доказательство

Шаг I – постановка задачи оптимизации

Рассмотрим методы вида (11.29) и для заданного k подберем такие $\tau_0, \tau_1, \dots \tau_{k-1}$, чтобы оценка погрешности метода на шаге k была **оптимальной**.

Такая задача возникает достаточно часто. Например, при ограниченных ресурсах: для решения СЛАУ выделены k итераций и метод должен гарантировать наилучший результат, возможный за k итераций.

Чтобы найти **оптимальные** параметры $\tau_0, \tau_1, \dots \tau_{k-1}$, выясним, как меняется погрешность метода на соседних шагах и как связаны погрешность метода на шаге k и начальная погрешность.

Пусть номер шага $s + 1 \le k$. Запишем метод в виде (11.33):

$$x^{(s+1)} = x^{(s)} - \tau_s A x^{(s)} + \tau_s b \tag{1}$$

Слева и справа вычтем x^* (решение СЛАУ), затем используем замену $b = Ax^*$:

$$x^{(s+1)} - x^* = (x^{(s)} - x^*) - \tau_s A x^{(s)} + \tau_s (A x^*)$$

Перепишем формулу в обозначениях погрешности:

$$z^{(s+1)} = z^{(s)} - \tau_s A z^{(s)} \tag{2}$$

Таким образом, на шаге $s + 1 \le k$ погрешности $z^{(s+1)}$ и $z^{(s)}$ связаны уравнением

$$z^{(s+1)} = (E - \tau_s A) z^{(s)} \tag{3}$$

где матрица $(E - \tau_s A)$ есть **матрица перехода на шаге** $s+1$.

Используя (3) для разных значений индекса, установим связь $z^{(s+1)}$ и начальной погрешности:

$$z^{(s+1)} = (E - \tau_s A) z^{(s)} = (E - \tau_s A)(E - \tau_{s-1} A) z^{(s-1)} = \ldots\ldots.$$
$$= (E - \tau_s A)(E - \tau_{s-1} A)\ldots(E - \tau_0 A) z^{(0)} \qquad (4)$$

Из формулы (4) следует, что на шаге k погрешности $z^{(k)}$ и $z^{(0)}$ связаны уравнением

$$z^{(k)} = (E - \tau_{k-1} A)(E - \tau_{k-21} A)\ldots(E - \tau_0 A)\, z^{(0)} \qquad (5)$$

Введем обозначение

$$G(\tau_0, \tau_1, \ldots \tau_{k-1}) = (E - \tau_{k-1} A)(E - \tau_{k-2} A)\ldots(E - \tau_0 A) \qquad (6)$$

Если метод использует параметры $\tau_0, \tau_1, \ldots \tau_{k-1}$, матрица $G(\tau_0, \tau_1, \ldots \tau_{k-1})$ описывает переход от начальной погрешности $z^{(0)}$ к погрешности $z^{(k)}$,

$$z^{(k)} = G(\tau_0, \tau_1, \ldots \tau_{k-1})\, z^{(0)} \qquad (7)$$

Матрицу $G(\tau_0, \tau_1, \ldots \tau_{k-1})$ называют **переходной матрицей метода**.

Рассмотрим евклидову норму погрешности $\left\| z^{(k)} \right\|_2$.

В силу **согласованности** норм матрицы и вектора

$$\left\| z^{(k)} \right\|_2 = \left\| G(\tau_0, \tau_1, \ldots \tau_{k-1})\, z^{(0)} \right\|_2 \le \left\| G(\tau_0, \tau_1, \ldots \tau_{k-1}) \right\|_2 \left\| z^{(0)} \right\|_2 \qquad (8)$$

Норма матрицы, указанная в (8), **подчинена** евклидовой норме вектора. Поэтому оценку (8) нельзя улучшить: существуют такие начальные приближения и соответственно такие начальные погрешности, что (8) выполняется как равенство.

Чтобы построить оптимальный метод вида (11.29), нужно найти такие $\tau_0, \tau_1, \ldots \tau_{k-1}$, чтобы норма переходной матрицы была минимальной:

$$\left\| G(\tau_0, \tau_1, \ldots \tau_{k-1}) \right\|_2 \to \min \qquad (9)$$

Шаг II – анализ нормы переходной матрицы

Исследуем норму $G(\tau_0, \tau_1, \ldots \tau_{k-1})$.

Несложно проверить, что $G(\tau_0, \tau_1, \ldots \tau_{k-1})$ симметрична.

В соответствии с формулой (6) она представляет собой полином степени k, в котором аргументом является матрица $A = A^T > 0$:

$$G(\tau_0, \tau_1, \ldots \tau_{k-1}) =$$
$$= E - \underbrace{(\tau_0 + \tau_1 + \ldots + \tau_{k-1})}_{\text{это число}} \cdot A + \ldots\ldots + \underbrace{(-1)^k (\tau_{k-1} \tau_{k-2} \ldots \tau_0)}_{\text{это число}} \cdot A^k$$

Каждый множитель вида

$$A^s = \underbrace{A \cdot A \cdot \ldots \cdot A}_{s\ \text{раз}},\ s = 1, \ldots k$$

является симметричным, каждое слагаемое полинома симметрично и поэтому симметрична сумма всех слагаемых:

$$G(\tau_0,\tau_1,...\tau_{k-1}) = [G(\tau_0,\tau_1,...\tau_{k-1})]^T \qquad (10)$$

Напомним, что «евклидова» норма симметричной матрицы совпадает с ее спектральным радиусом, то есть наибольшим по модулю собственным числом.

В силу симметрии (10) функционал задачи (9) имеет вид

$$\| G(\tau_0,\tau_1,...\tau_{k-1}) \|_2 = \rho(G) = \max_{i=1,...n} | \lambda_i(G) |, \qquad (11)$$

где $\lambda_i(G),\ i=1,...n$ – **собственные числа переходной матрицы**.

Шаг III – собственные числа переходной матрицы

Исследуем собственные числа $G(\tau_0,\tau_1,...\tau_{k-1})$.

Собственные векторы A обозначим через $v_i,\ i=1,...n$.

Собственные числа A обозначим через $\lambda_i,\ i=1,...n$.

Умножим $G(\tau_0,\tau_1,...\tau_{k-1})$ слева на собственный вектор матрицы A:

$$G(\tau_0,\tau_1,...\tau_{k-1})\, v_i = (E-\tau_{k-1}A)(E-\tau_{k-2}A)...(E-\tau_0 A)\, v_i \qquad (12)$$

В силу того, что

$$Av_i = \lambda_i v_i,\ \ i=1,...n,$$

не очень сложно показать, что

$$\begin{aligned}(E-\tau_{k-1}A)(E-\tau_{k-2}A)...(E-\tau_0 A)\, v_i &= \\ &= (1-\tau_0\lambda_i)(1-\tau_1\lambda_i)...(1-\tau_{k-1}\lambda_i)\, v_i\end{aligned} \qquad (13)$$

Из (12) и (13) получим

$$G(\tau_0,\tau_1,...\tau_{k-1})\, v_i = \underbrace{(1-\tau_{k-1}\lambda_i)(1-\tau_{k-2}\lambda_i)...(1-\tau_0\lambda_i)}_{\text{это число}}\, v_i$$

Следовательно, матрица $G(\tau_0,\tau_1,...\tau_{k-1})$ имеет такие же собственные векторы $v_i,\ i=1,...n$, что и матрица A, и собственными числами матрицы $G(\tau_0,\tau_1,...\tau_{k-1})$ являются вещественные числа

$$\lambda_i(G) = (1-\tau_0\lambda_i)(1-\tau_2\lambda_i)...(1-\tau_{k-1}\lambda_i),\ i=1,...n \qquad (14)$$

Других собственных чисел и других (с точностью до константы) собственных векторов у переходной матрицы $G(\tau_0,\tau_1,...\tau_{k-1})$ нет.

Примечание

Формула (14) означает следующее.

Если $\tau_0,\tau_1,...\tau_{k-1}$ выбраны, то на основе **упорядоченных, положительных** собственных чисел матрицы A

$$0 < \lambda_1 \le ... \le \lambda_n$$

могут быть вычислены собственные числа переходной матрицы.

Они принимают вещественные значения

$$\lambda_1(G) = (1-\tau_0\lambda_1)(1-\tau_2\lambda_1)...(1-\tau_{k-1}\lambda_1)$$

$$\lambda_2(G) = (1-\tau_0\lambda_2)(1-\tau_2\lambda_2)...(1-\tau_{k-1}\lambda_2)$$

… … …

$$\lambda_n(G) = (1-\tau_0\lambda_n)(1-\tau_2\lambda_n)...(1-\tau_{k-1}\lambda_n),$$

которые проиндексированы, но не упорядочены.

Если указанные значения еще не вычислены, определить, какое из них: будет наибольшим по модулю ($\lambda_1(G), \lambda_3(G)$ или $\lambda_7(G)$), не представляется возможным.

Шаг IV – задача о полиномах, наименее уклоняющихся от нуля (этот фрагмент доказательства не содержит доказательств)

Чтобы решить задачу (9) с функционалом (11), рассмотрим свойства полиномов, наименее уклоняющихся от нуля.

Уклонением непрерывной функции $\varphi(\lambda)$ от нуля на отрезке $\lambda \in [\lambda_1; \lambda_n]$ называется максимальное по модулю значение этой функции на данном отрезке:

$$\max_{\lambda \in [\lambda_1, \lambda_n]} |\varphi(\lambda)| \qquad (15)$$

Рассмотрим на отрезке $[\lambda_1; \lambda_n], 0 < \lambda_1 < \lambda_n$ класс полиномов степени не выше k со свободным слагаемым, равным 1. Обозначим указанный класс через $\hat{K}$:

$$\hat{K} = \{1 + a_1\lambda + ... + a_{k-1}\lambda^{k-1} + a_k\lambda^k\}. \qquad (16)$$

Элементами класса являются полиномы

$$P_k(\lambda) = 1 + a_1\lambda + ... + a_{k-1}\lambda^{k-1} + a_k\lambda^k$$

При $\lambda = 0$ они принимают значение 1: $P_k(0) = 1$

Полином $\hat{P}_k(\lambda) \in \hat{K}$ называют **наименее уклоняющимся от нуля**

на отрезке $[\lambda_1; \lambda_n], 0 < \lambda_1 < \lambda_n$ **в классе** $\hat{K}$,

если для любого (другого) полинома $P_k(\lambda) \in \hat{K}$ выполнено

$$\max_{\lambda \in [\lambda_1; \lambda_n]} |\hat{P}_k(\lambda)| \le \max_{\lambda \in [\lambda_1; \lambda_n]} |P_k(\lambda)| \qquad (17)$$

Задачу об отыскании полинома, наименее уклоняющегося от нуля на отрезке $[\lambda_1; \lambda_n], 0 < \lambda_1 < \lambda_n$ **в классе** $\hat{K}$ записывают в виде

$$\max_{\lambda \in [\lambda_1;\, \lambda_n]} |\, P_k(\lambda)\,| \underset{(a_1,\, a_2\, \ldots a_k\,) \in R^k}{\longrightarrow} \min, \qquad P_k(\lambda) \in \hat{K} \tag{18}$$

Запись (18) читается так: найти такие коэффициенты $(a_1, a_2, \ldots a_k) \in R^k$, чтобы полином $P_k(\lambda)$ из класса $\hat{K}$ с указанными выше коэффициентами имел на отрезке $\lambda \in [\lambda_1; \lambda_n]$ **минимально возможное максимальное по модулю значение**

$$\max_{\lambda \in [\lambda_1;\, \lambda_n]} |\, P_k(\lambda)\,|$$

Решение задачи (18) называют **полиномом, наименее уклоняющимся от нуля на отрезке** $[\lambda_1; \lambda_n], 0 < \lambda_1 < \lambda_n$ **в классе** $\hat{K}$

Установлено следующее:

Решением задачи (18), т.е. **полиномом, наименее уклоняющимся от нуля на отрезке** $[\lambda_1; \lambda_n], 0 < \lambda_1 < \lambda_n$ в классе $\hat{K}$, является полином Чебышева $\hat{T}_k(\lambda)$, корни которого вещественны, различны и принадлежат отрезку $[\lambda_1; \lambda_n]$.

Корнями являются следующие значения:

$$\hat{\lambda}_s = \frac{\lambda_1 + \lambda_n}{2} + \frac{\lambda_n - \lambda_1}{2} \cdot \cos\left(\frac{\pi}{2k}(1 + 2s) \right), \quad s = 0, \ldots k - 1 \tag{19}$$

Уклонение от нуля для полинома $\hat{T}_k(\lambda) \in \hat{K}$ на отрезке $[\lambda_1; \lambda_n], 0 < \lambda_1 < \lambda_n$ составит

$$\max_{\lambda \in [\lambda_1, \lambda_n]} |\, \hat{T}_k(\lambda)\,| = \frac{2\rho^k}{1 + \rho^{2k}}, \tag{20}$$

где $\rho = \dfrac{\sqrt{\dfrac{\lambda_n}{\lambda_1}} - 1}{\sqrt{\dfrac{\lambda_n}{\lambda_1}} + 1}$, причем

$$\rho \in (0; 1) \text{ и } \frac{2\rho^k}{1 + \rho^{2k}} \in (0; 1). \tag{21}$$

Уклонение от нуля для полинома $\hat{T}_k(\lambda) \in \hat{K}$ достигается в $k + 1$ точке отрезка, а именно, $k - 1$ точках локального экстремума, расположенных на интервале

$(\lambda_1;\ \lambda_n)$, а также на концах отрезка $[\lambda_1;\ \lambda_n]$, то есть в точках $\lambda = \lambda_1;\ \lambda = \lambda_n$.
Это означает, что

$$\max_{\lambda \in [\lambda_1, \lambda_n]} \left| \hat{T}_k(\lambda) \right| = \left| \hat{T}_k(\lambda_1) \right| = \left| \hat{T}_k(\lambda_n) \right| = \frac{2\rho^k}{1+\rho^{2k}} \tag{22}$$

Полином $\hat{T}_k(\lambda)$ имеет степень k и свободное слагаемое 1: $\hat{T}_k(0) = 1$.

Поэтому $\hat{T}_k(\lambda) \in \hat{K}$ можно записывать в виде

$$\hat{T}_k(\lambda) = \frac{(\lambda - \hat{\lambda}_0)(\lambda - \hat{\lambda}_1)\ldots(\lambda - \hat{\lambda}_{k-1})}{(-1)^k\, \hat{\lambda}_0 \cdot \hat{\lambda}_1 \ldots \cdot \hat{\lambda}_{k-1}} \tag{23}$$

где $\hat{\lambda}_s,\ s = 0,\ldots k-1$ - его корни, см. (19). Полином 4-й степени $\hat{T}_4(\lambda) \in \hat{K}$, наименее уклоняющийся от нуля на отрезке $[2, 15]$ и принимающий значение $\hat{T}_4(0) = 1$ (свободное слагаемое 1), в качестве примера приведен на рисунке.

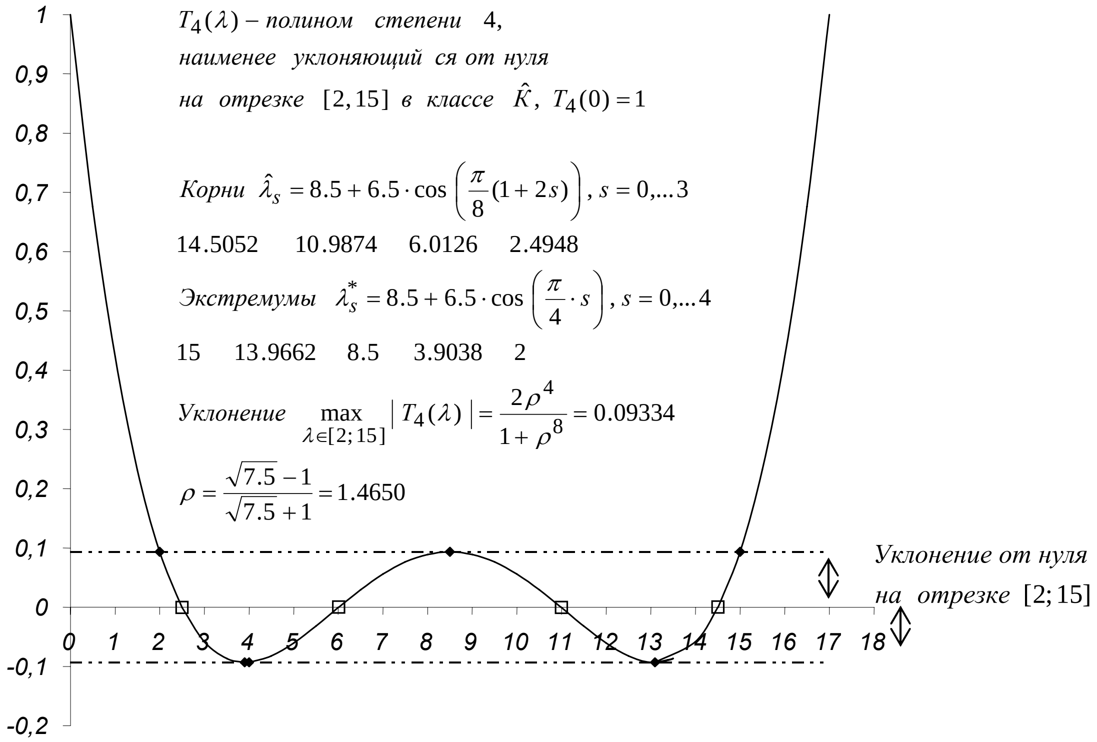

Рисунок 2

Шаг V – решение задачи оптимизации

Вернемся к задаче (9) с функционалом (11). Считаем, что параметры $\tau_0, \tau_1, ...\tau_{k-1}$ некоторым образом заданы, используем (14) и для нормы переходной матрицы $G(\tau_0, \tau_1, ...\tau_{k-1})$ запишем оценку

$$\| G(\tau_0, \tau_1, ...\tau_{k-1}) \|_2 = \max_{i=1,...n} | \lambda_i(G) | =$$
$$= \max_{i=1,...n} \underbrace{|(1-\tau_0\lambda_i)(1-\tau_2\lambda_i)...(1-\tau_{k-1}\lambda_i)|}_{\substack{\text{это собственные числа } \lambda_i(G) \\ \text{переходной матрицы } G(\tau_0,\,\tau_1,\,...\tau_{k-1})}} \le \qquad (24)$$
$$\le \max_{\lambda\in[\lambda_1,\lambda_n]} |(1-\tau_0\lambda)(1-\tau_1\lambda)...(1-\tau_{k-1}\lambda)|$$

Оценка верна, потому что в левой ее части под знаком модуля указаны значения некоторого полинома $P_k(\lambda)$ в точках $\lambda_i,\ i = 1,...n,$, **расположенных на отрезке** $[\lambda_1;\ \lambda_n]$, а в правой части под знаком модуля указаны значения того же полинома $P_k(\lambda)$ **в произвольной точке отрезка** $[\lambda_1;\ \lambda_n]$:

$$\| G(\tau_0, \tau_1, ...\tau_{k-1}) \|_2 = \max_{i=1,...n} | \lambda_i(G) | =$$
$$= \max_{i=1,...n} \underbrace{|(1-\tau_0\lambda_i)(1-\tau_2\lambda_i)...(1-\tau_{k-1}\lambda_i)|}_{\substack{\text{это } P_k(\lambda_i)\text{, где } \lambda_i - \text{положительное} \\ \text{собственное число матрицы } A}} \le$$
$$\le \max_{\lambda\in[\lambda_1,\lambda_n]} \underbrace{|(1-\tau_0\lambda)(1-\tau_1\lambda)...(1-\tau_{k-1}\lambda)|}_{\substack{\text{это } P_k(\lambda) \text{ при } \lambda\in[\lambda_1,\,\lambda_n]\text{, где} \\ \lambda_1 \text{ и } \lambda_n - \text{минимальное и максимальное} \\ \text{положительные собственные числа матрицы } A}}$$

Оценка верна, потому что максимум модуля функции $P_k(\lambda)$ на отрезке $[\lambda_1;\ \lambda_n]$ не может быть меньше модуля функции $P_k(\lambda)$ в конкретной точке того же отрезка $[\lambda_1;\ \lambda_n]$.

Далее поставим задачу найти такие параметры $\tau_0, \tau_1, ...\tau_{k-1}$, чтобы на заданном отрезке $[\lambda_1;\ \lambda_n]$, где $0 < \lambda_1 < \lambda_n$, **максимум** модуля функции $P_k(\lambda)$ принимал **минимально возможное значение**:

$$\max_{\lambda\in[\lambda_1,\lambda_n]} \underbrace{|(1-\tau_0\lambda)(1-\tau_1\lambda)...(1-\tau_{k-1}\lambda)|}_{\substack{\text{это } P_k(\lambda) - \text{полином степени} \\ \text{не выше } k\text{, такой, что } P_k(0)=1}} \xrightarrow[(\tau_0,\,\tau_1,\,...\tau_k)\,\in\,R^k]{} \min \qquad (25)$$

Задача (25) представляет собой разобранную выше (см. (18)) **задачу об отыскании полинома, наименее уклоняющегося от нуля на отрезке** $[\lambda_1;\ \lambda_n]$, $0 < \lambda_1 < \lambda_n$, **в классе полиномов степени не выше** k **со свободным слагаемым, равным 1.**

Как рассмотрено выше (Шаг IV), решением (25) является полином $\hat{T}_k(\lambda)$ степени

k, такой, что $\hat{T}_k(0)=1$. Его корни определены по формулам

$$\hat{\lambda}_s = \frac{\lambda_1+\lambda_n}{2} + \frac{\lambda_n-\lambda_1}{2}\cdot\cos\left(\frac{\pi}{2k}(1+2s)\right), \quad s=0,...k-1 \qquad \text{(см. (19))}$$

Чтобы искомый полином $P_k(\lambda)$ совпал с $\hat{T}_k(\lambda)$, нужно, чтобы совпали их корни. Так как $P_k(\lambda)$ обращается в ноль при значениях аргумента

$$\lambda = \frac{1}{\tau_s},\ s=0,...k-1,$$

полином $P_k(\lambda)$ совпадает с $\hat{T}_k(\lambda)$, если выбирать параметры $\tau_0, \tau_1, ... \tau_{k-1}$ из условия

$$\frac{1}{\tau_s} = \hat{\lambda}_s, \quad s=0,...k-1,$$

что означает

$$\tau_s = \left(\frac{\lambda_1+\lambda_n}{2} + \frac{\lambda_n-\lambda_1}{2}\cdot\cos\left(\frac{\pi}{2k}(1+2s)\right)\right)^{-1},\ s=0,...k-1$$

Таким образом, при выборе $\tau_0, \tau_1, ... \tau_{k-1}$ по формулам (11.30) функционал задачи (25) принимает минимально возможное значение, и оно равно

$$\max_{\lambda\in[\lambda_1,\lambda_n]} \left| (1-\tau_0\lambda)(1-\tau_1\lambda)...(1-\tau_{k-1}\lambda) \right| =$$

$$= \max_{\lambda\in[\lambda_1,\, \lambda_n]} \left| T_k(\lambda) \right| = \frac{2\rho^k}{1+\rho^{2k}} \qquad (26)$$

где $\rho = \dfrac{\sqrt{\dfrac{\lambda_n}{\lambda_1}}-1}{\sqrt{\dfrac{\lambda_n}{\lambda_1}}+1} = \dfrac{\sqrt{\mu_A}-1}{\sqrt{\mu_A}+1}$

Напомним, что для полинома $\hat{T}_k(\lambda)\in\hat{K}$ уклонение от нуля (максимальное по модулю значение) достигается в $k+1$ точке отрезка $[\lambda_1;\ \lambda_n]$, в том числе на границе отрезка, в точках $\lambda=\lambda_1;\ \lambda=\lambda_n$, см. (22):

$$\left| \hat{T}_k(\lambda_1) \right| = \left| \hat{T}_k(\lambda_n) \right| = \frac{2\rho^k}{1+\rho^{2k}}$$

Это означает, что при выборе $\tau_0, \tau_1, ... \tau_{k-1}$ по формулам (11.30)

$$\left| (1-\tau_0\lambda_1)(1-\tau_1\lambda_1)...(1-\tau_{k-1}\lambda_1) \right| = \frac{2\rho^k}{1+\rho^{2k}}$$

$$\left| (1-\tau_0\lambda_n)(1-\tau_1\lambda_n)...(1-\tau_{k-1}\lambda_n) \right| = \frac{2\rho^k}{1+\rho^{2k}} \qquad (27)$$

Шаг VI – проверка утверждений о сходимости

Рассмотрим неравенство (24) для метода (11.29) с параметрами (11.30). С учетом (26) для нормы переходной матрицы $G(\tau_0,\tau_1,...\tau_{k-1})$ доказана оценка

$$\| G(\tau_0,\tau_1,...\tau_{k-1}) \|_2 = \max_{i=1,...n} | \lambda_i(G) | =$$

$$= \max_{i=1,...n} \underbrace{\left| (1-\tau_0\lambda_i)(1-\tau_2\lambda_i)...(1-\tau_{k-1}\lambda_i) \right|}_{\text{это собственные числа } \lambda_i(G) \text{ переходной матрицы } G(\tau_0,\tau_1,...\tau_{k-1})} \le$$

$$\le \max_{\lambda\in[\lambda_1,\lambda_n]} \underbrace{\left| (1-\tau_0\lambda)(1-\tau_1\lambda)...(1-\tau_{k-1}\lambda) \right|}_{\text{здесь полином } \hat{T}_k(\lambda) \text{, наименее уклоняющийся от нуля на отрезке } [\lambda_1,\lambda_n]} = \qquad (28)$$

$$= \frac{2\rho^k}{1+\rho^{2k}} \quad \text{(это величина его уклонения)}$$

В силу (27)

$$\| G(\tau_0,\tau_1,...\tau_{k-1}) \|_2 = \max_{i=1,...n} | \lambda_i(G) | =$$

$$= \max_{i=1,...n} \underbrace{\left| (1-\tau_0\lambda_i)(1-\tau_2\lambda_i)...(1-\tau_{k-1}\lambda_i) \right|}_{\text{это собственные числа } \lambda_i(G) \text{ переходной матрицы } G(\tau_0,\tau_1,...\tau_{k-1})} =$$

$$= \underbrace{\left| (1-\tau_0\lambda_1)(1-\tau_1\lambda_1)(1-\tau_{k-1}\lambda_1) \right|}_{\text{это собственное число } \lambda_1(G) \text{ переходной матрицы } G \text{, оно максимальное по модулю}} =$$

$$= \underbrace{\left| (1-\tau_0\lambda_n)(1-\tau_1\lambda_n)(1-\tau_{k-1}\lambda_n) \right|}_{\text{это собственное число } \lambda_n(G) \text{ переходной матрицы } G \text{, оно также максимальное помодулю}} =$$

$$= \max_{\lambda\in[\lambda_1,\lambda_n]} \underbrace{\left| (1-\tau_0\lambda)(1-\tau_1\lambda)...(1-\tau_{k-1}\lambda) \right|}_{\text{здесь полином } \hat{T}_k(\lambda) \text{, наименее уклоняющийся от нуля на отрезке } [\lambda_1,\lambda_n]} =$$

$$= \frac{2\rho^k}{1+\rho^{2k}} \quad \text{(это величина его уклонения)} \qquad (29)$$

Таким образом, для метода (11.29) с параметрами (11.30)

$$\| G(\tau_0,\tau_1,...\tau_{k-1}) \|_2 = \frac{2\rho^k}{1+\rho^{2k}} \qquad (30)$$

Из (8) и (30) следует приведенная в Теореме 8 оценка погрешности метода:

$$\| z^{(k)} \|_2 \le \frac{2\rho^k}{1+\rho^{2k}} \| z^{(0)} \|_2$$ см. (11.34)

Оценку (11.35) для погрешности метода на шаге Nk (N есть количество циклов) получим по индукции.

Пусть $N = 2.$ В соответствии с (11.32) начальным приближением для второго цикла является $x^{(k)}$. Поэтому погрешность $z^{(k)} = x^{(k)} - x*$ является начальной погрешностью второго цикла. Через k шагов будет получен элемент $x^{(2k)}$, и погрешность метода следует обозначить $z^{(2k)} = x^{(2k)} - x*$. В соответствии с (11.34)

$$\| z^{(2k)} \|_2 \le \frac{2\rho^k}{1+\rho^{2k}} \| z^{(k)} \|_2 \le \left(\frac{2\rho^k}{1+\rho^{2k}} \right)^2 \| z^{(0)} \|_2 .$$

Аналогично доказывается оценка погрешности метода по завершении каждого последующего цикла. По завершении цикла с номером N получим

$$\| z^{(Nk)} \|_2 \le \frac{2\rho^k}{1+\rho^{2k}} \| z^{((N-1)k)} \|_2 \le \ldots \le \left(\frac{2\rho^k}{1+\rho^{2k}} \right)^N \| z^{(0)} \|_2$$

Сходимость метода следует из (21) и (11.35).

Так как $\frac{2\rho^k}{1+\rho^{2k}} \in (0,1)$, при $N \to \infty$ получим $\left(\frac{2\rho^k}{1+\rho^{2k}} \right)^N \to 0$.

Откуда следует, что для $\forall x^{(0)} \in R^n$ при $N \to \infty$ (при увеличении числа циклов)

$$\lim_{N \to +\infty} \| z^{(Nk)} \|_2 = 0 .$$

Последовательность $x^{(0)}, x^{(k)}, x^{(2k)}, .x^{(3k)}, x^{(4k)} ... x^{(Nk)}$..... сходится к $x*$ (то есть к решению СЛАУ $Ax = b$) с оценкой (11.35)

Построение метода на основе оценок собственных чисел

Рассмотрим СЛАУ (11.1) с симметричной, положительно определенной матрицей. Как следует из Теоремы 8, для построения метода и оценки погрешности нужно знать минимальное и максимальное собственные числа матрицы A.

Предположим, что собственные числа матрицы $A = A^T > 0$ неизвестны, но известны их оценки, то есть известны числа $M_{\min} > 0$ и $M_{\max} > 0$, такие, что

$\lambda_i(A) \in [M_{\min}, M_{\max}]$, $i = 1,...n$.

Рассмотрим метод (11.29) с натуральным параметром k и вещественными параметрами

$$\tilde{\tau}_s = \left(\frac{M_{\min} + M_{\max}}{2} + \frac{M_{\max} - M_{\min}}{2} \cdot \cos\left(\frac{\pi}{2k}(1+2s) \right) \right)^{-1}, \; s = 0,...k-1 \quad (11.39)$$

Т е о р е м а 9. При решении задачи (11.1) с симметричной, положительно определенной матрицей A методом (11.29) с параметрами (11.39) оценка погрешности метода на шаге k имеет вид

$$\left\| z^{(k)} \right\|_2 \le \frac{2\tilde{\rho}^k}{1+\tilde{\rho}^{2k}} \left\| z^{(0)} \right\|_2 \quad (11.40)$$

Оценка погрешности метода на шаге Nk (N есть количество циклов) имеет вид

$$\left\| z^{(Nk)} \right\|_2 \le \left(\frac{2\tilde{\rho}^k}{1+\tilde{\rho}^{2k}} \right)^N \left\| z^{(0)} \right\|_2 \quad (11.41)$$

и метод сходится в следующем смысле:

$$\forall x^{(0)} \in R^n \quad \lim_{N \to +\infty} \left\| z^{(Nk)} \right\|_2 = 0. \quad (11.42)$$

(к решению x^* сходятся элементы последовательности (11.32)).

В формулах (11.40)-(11.42) значение $\tilde{\rho}$ определено как

$\tilde{\rho} = \dfrac{\sqrt{M}-1}{\sqrt{M}+1}$, где $M = \dfrac{M_{\max}}{M_{\min}}$.

Оптимальное свойство метода (11.29), (11.39) состоит в следующем: среди всех методов вида (11.29) метод с параметрами (11.39) дает наилучшую гарантию убывания погрешности за k шагов, справедливую для всех симметричных положительно определенных матриц $A = A^T > 0$, собственные числа которых расположены в диапазоне $[M_{\min}, M_{\max}], 0 < M_{\min} < M_{\max}$.

К о м м е н т а р и й

> Если собственные числа $A = A^T > 0$ расположены в диапазоне $[\lambda_1, \lambda_n]$, $0 < \lambda_1 < \lambda_n$, метод (11.29) с параметрами (1.30) обеспечит лучшую гарантию убывания погрешности, чем метод (11.29) с параметрами (11.39).

Действительно, в данном случае $1 < \mu_A \le M$ и $0 < \dfrac{\sqrt{\mu_A} - 1}{\sqrt{\mu_A} + 1} \le \dfrac{\sqrt{M} - 1}{\sqrt{M} + 1} < 1$

(оценивается возрастающая функция вида $\varphi(x) = \dfrac{\sqrt{x} - 1}{\sqrt{x} + 1}$ при $x > 1$).

Поэтому для множителей из оценок (11.34) и (11.40), (11.35) и (11.41) верно

$$0 < \frac{2\rho^k}{1+\rho^{2k}} \le \frac{2\tilde{\rho}^k}{1+\tilde{\rho}^{2k}} < 1,\ 0 < \left(\frac{2\rho^k}{1+\rho^{2k}}\right)^N \le \left(\frac{2\tilde{\rho}^k}{1+\tilde{\rho}^{2k}}\right)^N < 1$$

Метод (11.29) с параметрами (11.30) (тот, что лучше сходится) требует знания собственных чисел, а метод с параметрами (11.39) использует их оценки.

Пример (k=4)

$Ax = b$, где $x \in R^n, b \in R^n,\ A\ (n \times n),\ A = A^T > 0$, $\lambda_i(A) \in [2;15]$, $i = 1,\dots n$.

Метод $x^{(s+1)} = x^{(s)} - \tilde{\tau}_s \cdot r^{(s)}$, где $k = 4$, $s = 0,\dots 3$. Используем оценки спектра.

$$\tilde{\tau}_s = \left(\frac{2+15}{2} + \frac{15-2}{2} \cdot \cos\left(\frac{\pi}{8}(1+2s)\right)\right)^{-1},\ s = 0,\dots 3$$

$$\tilde{\tau}_0 = \left(8.5 + 6.5 \cdot \cos\left(\frac{\pi}{8}\right)\right)^{-1} = 0.06894 \qquad \tilde{\tau}_1 = \left(8.5 + 6.5 \cdot \cos\left(\frac{3\pi}{8}\right)\right)^{-1} = 0.09101$$

$$\tilde{\tau}_2 = \left(8.5 + 6.5 \cdot \cos\left(\frac{5\pi}{8}\right)\right)^{-1} = 0.16632 \qquad \tilde{\tau}_3 = \left(8.5 + 6.5 \cdot \cos\left(\frac{7\pi}{8}\right)\right)^{-1} = 0.40084$$

Оценим погрешность метода на шаге $4N$ (N есть количество циклов):

$$\left\| z^{(4N)} \right\|_2 \le \left(\frac{2\rho^4}{1+\rho^8}\right)^N \left\| z^{(0)} \right\|_2 = (0.09334)^N \cdot \left\| z^{(0)} \right\|_2,$$

Здесь $\rho = \dfrac{\sqrt{\frac{15}{2}} - 1}{\sqrt{\frac{15}{2}} + 1} = 0.46504$. Метод сходится: $\forall x^{(0)} \in R^n \quad \lim\limits_{N \to +\infty} \left\| z^{(4N)} \right\|_2 = 0$.

Через $N = 5$ циклов, то есть за 20 шагов, для погрешности метода верно

$$\left\| z^{(20)} \right\|_2 \le (0.09334)^5 \cdot \left\| z^{(0)} \right\|_2 = 0.71 \cdot 10^{-5} \cdot \left\| z^{(0)} \right\|_2.$$

Погрешность начального приближения снизится более чем в 100 000 раз.

11.4. Метод сопряженных градиентов

Сведение решения СЛАУ к решению задачи оптимизации. Сопряженные направления и их свойства. Сведение k-мерной (многомерной) задачи оптимизации к решению k одномерных оптимизационных задач. Описание метода сопряженных градиентов. Основные свойства метода. Комментарии о применении метода

Рассмотрим систему линейных алгебраических уравнений (СЛАУ) с симметричной положительно определенной матрицей

$$Ax = b \tag{11.43}$$

где $x \in R^n, b \in R^n, A\,(n \times n), A = A^T > 0$.

Через x^* обозначим точное решение СЛАУ, $x^* \in R^n$.

Введем функционал $F(x) = (Ax, x) - 2(b, x)$ и рассмотрим задачу оптимизации

$$F(x) = (Ax, x) - 2(b, x) \to \min \tag{11.44}$$

Используя свойства $F(x)$, несложно показать, что единственным решением (11.44) является тот самый $x^* \in R^n$, который является единственным решением (11.43).

В силу этого свойства методы решения задач оптимизации могут быть использованы для отыскания решения СЛАУ.

Докажем эквивалентность задач

Утверждение 1. Каждая из задач (11.43) и (11.44) имеет единственное решение и эти решения совпадают.

Доказательство

Так как $\det A \neq 0$, решение СЛАУ (11.43) существует и единственно. Обозначим его x^*. Для $\forall x, \forall h \in R^n$ верно

$$\begin{aligned} F(x+h) &= (A(x+h), x+h) - 2(b, x+h) = \\ &= (Ax, x) + (Ah, x) + (Ax, h) + (Ah, h) - 2(b, x) - 2(b, h) = \\ &= F(x) + (Ah, h) + (Ah, x) + (Ax, h) - 2(b, h) \end{aligned} \tag{1}$$

В силу коммутативности скалярного произведения и симметричности матрицы A второе и третье слагаемые равны: $(Ah, x) = (Ax, h)$. Действительно,

$$(Ah, x) = (x, Ah) = (A^T x, h) = (Ax, h). \tag{2}$$

Поэтому для $\forall x, \forall h \in R^n$

$$F(x+h) = F(x) + (Ah, h) + 2(Ax - b, h). \tag{3}$$

Рассмотрим (3) с аргументом $x = x^*$. Тогда для $\forall h \in R^n$

$$F(x^* + h) = F(x^*) + (Ah, h) + 2(Ax^* - b, h) \tag{4}$$

Так как $Ax^* = b$ и для $\forall h \neq 0$ $(Ah, h) > 0$, из (4) для $\forall h \neq 0$ получим

$$F(x^* + h) = F(x^*) + (Ah, h) > F(x^*). \qquad (5)$$

Это означает, что для $\forall x \neq x^*$ $F(x) > F(x^*)$, то есть решение задачи оптимизации (11.44) существует, единственно и совпадает с решением (11.43).

Для решения задачи оптимизации определим сопряженные направления

Определение 1. Пусть $A = A^T > 0$. Направления (векторы) $h', h'' \in R^n, h', h'' \neq 0$ называются **сопряженными** относительно A, если $(Ah', h'') = 0$.

Комментарий

В данном определении направления $h', h'' \in R^n$ «равноправны». Они оба должны быть отличны от нуля и если $(Ah', h'') = 0$, то в силу коммутативности скалярного произведения и симметричности матрицы A получим $(Ah'', h') = 0$. Действительно,

$$(Ah'', h') = (h'', A^T h') = (h'', Ah') = (Ah', h'') = 0$$

Определение 2. Пусть $A = A^T > 0$. Ненулевые направления (векторы) $h^{(0)}, \ldots h^{(k-1)} \in R^n$ называются **взаимно сопряженными** относительно A, если $(Ah^{(i)}, h^{(j)}) = 0$ для $\forall i, j = \overline{0, k-1}, i \neq j$.

Решение задач оптимизации опирается на линейную независимость сопряженных направлений

Утверждение 2. Если ненулевые направления $h^{(0)}, \ldots h^{(k-1)} \in R^n$ взаимно сопряжены относительно $A = A^T > 0$, тогда они линейно независимы. Если ненулевые направления $h^{(0)}, \ldots h^{(n-1)} \in R^n$, взятые в количестве n, взаимно сопряжены относительно $A = A^T > 0$, тогда они образуют базис в R^n.

Доказательство (от противного)

Пусть ненулевые взаимно сопряженные направления $h^{(0)}, \ldots h^{(k-1)} \in R^n$ линейно зависимы.

Тогда существует такое сопряженное направление $h^{(l)} \neq 0$, которое можно представить в виде линейной комбинации остальных сопряженных направлений:

$$h^{(l)} = \sum_{i=0, i\neq l}^{k-1} \alpha_i h^{(i)}.$$

Так как $A > 0$ и $h^{(l)} \neq 0$, имеет место строгое неравенство $(Ah^{(l)}, h^{(l)}) > 0$.

В силу взаимной сопряженности направлений получим $(Ah^{(l)}, h^{(l)}) = 0$. Действительно,

$$(Ah^{(l)}, h^{(l)}) = (Ah^{(l)}, \sum_{i=0, i\neq l}^{k-1} \alpha_i h^{(i)}) = \sum_{i=0,\ i\neq l}^{k-1} \alpha_i (Ah^{(l)}, h^{(i)}) = 0\,.$$

Обнаружено противоречие. Значит, ненулевые взаимно сопряженные направления $h^{(0)},...h^{(k-1)} \in R^n$ линейно независимы.

Если $A = A^T > 0$ и найдутся n ненулевых направлений, взаимно сопряженных относительно A, то указанные направления в силу их линейной независимости и в силу их количества образуют базис в R^n.

Решение задач оптимизации строится на базе сопряженных направлений

Теорема 10. Пусть $h^{(0)},...h^{(k-1)} \in R^n$, $k \le n$, представляют собой ненулевые взаимно сопряженные направления относительно $A = A^T > 0$ (они линейно независимы) и пусть $x^{(0)} \in R^n$.

Определим в R^n линейное многообразие $L_k(x_0, h^{(0)},....h^{(k-1)})$ размерности k, элементами которого являются суммы направления (вектора) $x^{(0)} \in R^n$ и линейной комбинации сопряженных направлений $h^{(0)},...h^{(k-1)} \in R^n$, то есть

$x = x^{(0)} + \alpha_0 h^{(0)} + ... + \alpha_{k-1} h^{(k-1)}$, где $\alpha_i \in R, i = 0,...k-1$,

Тогда решением *k*-мерной (многомерной) задачи оптимизации

$$F(x) = \underset{x \in L_k(x^{(0)},\ h^{(0)},...\ h^{(k-1)})}{\rightarrow} \min \tag{11.45}$$

(такую задачу называют задачей минимизации на многообразии) является направление

$x^{(k)} = x^{(0)} + \alpha_0^* h^{(0)} + ... + \alpha_{k-1}^* h^{(k-1)}$,

$x^{(k)} \in L_k(x_0, h^{(0)},....h^{(k-1)})$,

где коэффициенты $\alpha_s^*, s = 0,...k-1$, есть решения k одномерных задач оптимизации

$$\alpha_s^2 (Ah^{(s)}, h^{(s)}) + 2\alpha_s (Ax^{(0)} - b, h^{(s)}) \underset{\alpha_s \in R}{\rightarrow} \min,\ s = 0, k-1. \tag{11.46}$$

(в каждой из задач минимизируется значение полинома степени 2 с аргументом α_s).

При этом коэффициенты $\alpha_s^*, s = 0,...k-1$ вычисляются как аргументы вершины параболы:

$$\alpha_s^* = -\frac{(Ax^{(0)} - b, h^{(s)})}{(Ah^{(s)}, h^{(s)})}, s = 0,...k-1 \tag{11.47}$$

Доказательство

Пусть $A = A^T > 0$ и ненулевые направления $h^{(0)},...h^{(k-1)} \in R^n$ взаимно сопряжены относительно A. По условию теоремы количество взаимно сопряженных направлений равно $k \le n$ и в соответствии с утверждением 2 они линейно независимы. Пусть выбран некоторый $x^{(0)} \in R^n$.

Определим в R^n линейное многообразие размерности k и обозначим его через $L_k(x_0, h^{(0)},....h^{(k-1)})$. Элементы данного многообразия имеют вид

$$x = x^{(0)} + \alpha_0 h^{(0)} + ... + \alpha_{n-1} h^{(k-1)},$$

где $\alpha_i \in R, i = 0,...k-1$ – числа, $h^{(0)},...h^{(k-1)} \in R^n$ – ненулевые линейно независимые направления, $x^{(0)} \in R^n$ – выбранный выше элемент.

Рассмотрим на линейном многообразии $L_k(x^{(0)}, h^{(0)},....h^{(k-1)})$ задачу оптимизации (11.45):

$$F(x) \underset{x \in L_k(x^{(0)},\, h^{(0)},...\, h^{(k-1)})}{\to} \min$$

В силу формулы (3) для $\forall x \in L_k(x^{(0)}, h^{(0)},....h^{(k-1)})$ верно представление

$$F(x) = F(x^{(0)} + \sum_{s=0}^{k-1} \alpha_s h^{(s)}) =$$

$$= F(x^{(0)}) + (A \sum_{s=0}^{k-1} \alpha_s h^{(s)}, \sum_{l=0}^{k-1} \alpha_l h^{(l)}) + 2(Ax^{(0)} - b, \sum_{s=0}^{k-1} \alpha_s h^{(s)})$$

В силу взаимной сопряженности направлений

$$(A \sum_{s=0}^{k-1} \alpha_s h^{(s)}, \sum_{l=0}^{k-1} \alpha_l h^{(l)}) = \sum_{s=0}^{k-1} \alpha_s^2 (Ah^{(s)}, h^{(s)}).$$

Тогда функционал $F(x)$ принимает вид

$$F(x) = F(x^{(0)}) + \sum_{s=0}^{k-1} \alpha_s^2 (Ah^{(s)}, h^{(s)}) + 2 \sum_{s=0}^{k-1} \alpha_s (Ax^{(0)} - b, h^{(s)}).$$

Слагаемые, зависящие от одного и того же коэффициента α_s и одного и того же сопряженного направления $h^{(s)} \in R^n$, можно сгруппировать:

$$F(x) = F(x^{(0)}) + \sum_{s=0}^{k-1} \{\alpha_s^2 (Ah^{(s)}, h^{(s)}) + 2\alpha_s (Ax^{(0)} - b, h^{(s)})\}.$$

Теперь каждое слагаемое, указанное в фигурных скобках, зависит только от одного коэффициента α_s и только одного направления $h^{(s)} \in R^n$.

Поэтому решение k-мерной задачи оптимизации (11.45) сводится к независимому решению k одномерных задач оптимизации вида (11.46):

$$\alpha_s^2(Ah^{(s)},h^{(s)})+2\alpha_s(Ax^{(0)}-b,h^{(s)}) \underset{\alpha_s\in R}{\to} \min,\ s=0,k-1$$

Каждая из (11.46) есть задача минимизации значения полинома степени 2 (минимизируется полином с аргументом $\alpha_s \in R$).

Так как в каждой из задач вида (11.46) $h^{(s)} \neq 0$ и $(Ah_s,h_s)>0$, ветви параболы направлены вверх и решением задачи является вершина параболы. Ее аргумент

$$\alpha_s^* = -\frac{2(Ax^{(0)}-b,h^{(s)})}{2(Ah^{(s)},h^{(s)})} = -\frac{(Ax^{(0)}-b,h^{(s)})}{(Ah^{(s)},h^{(s)})},$$

Минимальное значение функционала из задачи (11.46) составит

$$\left[\alpha_s^*\right]^2(Ah^{(s)},h^{(s)})+2\alpha_s^*(Ax^{(0)}-b,h^{(s)}) = -\frac{(Ax^{(0)}-b,h^{(s)})^2}{(Ah^{(s)},h^{(s)})}.$$

Таким образом, решением задачи (11.45) является $x^{(k)}=x^{(0)}+\alpha_0^*h^{(0)}+...+\alpha_{k-1}^*h^{(k-1)}$ с коэффициентами (11.47), а минимальное значение функционала из задачи (11.45) составит

$$F(x^{(k)})=F(x^{(0)})-\sum_{s=0}^{k-1}\frac{(Ax^{(0)}-b,h^{(s)})^2}{(Ah^{(s)},h^{(s)})}.$$

Теорема доказана.

Комментарий

Очевидно, что $F(x^{(k)}) \le F(x^{(0)})$. При этом $F(x^{(k)}) = F(x^{(0)})$ в случаях:

1) $x^{(0)} \in R^n$ является решением $Ax=b$, то есть $x^{(0)}=x^*$

2) $x^{(0)} \in R^n$ не является решением $Ax=b$, но невязка $Ax^{(0)}-b$ ортогональна каждому из сопряженных направлений $h^{(0)},...h^{(k-1)} \in R^n$. Тогда ненулевая погрешность $z^{(0)}=x^{(0)}-x^*$ сопряжена с каждым из $h^{(0)},...h^{(k-1)} \in R^n$ и система взаимно сопряженных направлений может быть дополнена направлением $h^{(k)}=z^{(0)}=x^{(0)}-x^*$. Тогда значение функционала $F(x)$ может быть уменьшено на линейном многообразии более высокой размерности.

Рассмотрим следствие из теоремы: оно показывает, что с помощью сопряженных направлений решение (11.44) и (11.43) сводится к явному решению задач одномерной оптимизации

Следствие. Пусть $h^{(0)},...h^{(n-1)} \in R^n$ представляют собой ненулевые взаимно сопряженные относительно $A = A^T > 0$ направления и пусть $x^{(0)} \in R^n$. Тогда указанные выше взаимно сопряженные направления $h^{(0)},...h^{(n-1)} \in R^n$ образуют базис в R^n, линейное многообразие $L_n(x_0, h^{(0)},....h^{(n-1)})$ совпадает с пространством R^n, решением *n*-мерной (многомерной) задачи оптимизации

$$F(x) \underset{x \in R^n}{\rightarrow} \min$$

является направление

$$x^* = x^{(0)} + \alpha_0^* h^{(0)} + ... + \alpha_{k-1}^* h^{(n-1)}$$

где коэффициенты $\alpha_s^*, s = 0,...n-1$, есть решения n одномерных задач оптимизации

$$\alpha_s^2 (Ah^{(s)}, h^{(s)}) + 2\alpha_s (Ax^{(0)} - b, h^{(s)}) \underset{\alpha_s \in R}{\rightarrow} \min, \ s = 0, n-1. \qquad (11.48)$$

и вычисляются как аргументы вершины параболы:

$$\alpha_s^* = -\frac{(Ax^{(0)} - b, h^{(s)})}{(Ah^{(s)}, h^{(s)})}, s = 0,...n-1 \qquad (11.49)$$

Комментарий

Для того, чтобы решение многомерной задачи оптимизации (11.44) было сведено к явному решению нескольких одномерных оптимизационных задач, нужно знать сопряженные направления

Описание метода сопряженных градиентов

Идея пошагового построения ненулевых взаимно сопряженных направлений реализована в **методе сопряженных градиентов**. В качестве $x^{(0)} \in R^n$ выбирают элемент R^n, «удобный» как начальное приближение к решению x^*.

Первый шаг метода

Приближение $x^{(1)} \in R^n$ найдем по формуле

$$x^{(1)} = x^{(0)} + \alpha_0 h^{(0)} \qquad (11.50)$$

где направление (вектор) $h^{(0)} \in R^n$ определяется начальной невязкой:

$$h^{(0)} = -r^{(0)} = Ax^{(0)} - b. \qquad (11.51)$$

Чтобы найти α_0, решаем одномерную задачу оптимизации

$$F(x^{(0)} + \alpha_0 h^{(0)}) \underset{\alpha_0 \in R}{\rightarrow} \min \qquad (11.52)$$

(задача минимизации по аргументу α_0, где $x^{(0)}, h^{(0)} \in R^n$ известны).

Так как

$$F(x^{(0)}+\alpha_0 h^{(0)})=F(x^{(0)})+\alpha_0^2(Ah^{(0)},h^{(0)})+2\alpha_0(Ax^{(0)}-b,h^{(0)})$$

(полином второй степени относительно α_0), решением (11.52) является

$$\alpha_0=-\frac{(Ax^{(0)}-b,h^{(0)})}{(Ah^{(0)},h^{(0)})} \qquad (11.53)$$

(аргумент вершины параболы).

Второй шаг метода

Приближение $x^{(2)}\in R^n$ найдем по формуле

$$x^{(2)}=x^{(1)}+\alpha_1 h^{(1)} \qquad (11.54)$$

где направление (вектор) $h^{(1)}\in R^n$ определяется невязкой предыдущего шага

$$r^{(1)}=Ax^{(1)}-b$$

и предыдущим направлением $h^{(0)}$:

$$h^{(1)}=-r^{(1)}+\beta_1 h^{(0)} \qquad (11.55)$$

Нужно, чтобы $h^{(1)},h^{(0)}$ были сопряженными относительно $A=A^T>0$:

$$(Ah^{(0)},h^{(1)})=0.$$

Получим условие

$$(Ah^{(0)},h^{(1)})=(Ah^{(0)},h^{(0)})\beta_1-(Ah^{(0)},r^{(1)})=0,$$

откуда следует

$$\beta_1=\frac{(Ah^{(0)},r^{(1)})}{(Ah^{(0)},h^{(0)})} \qquad (11.56)$$

Чтобы найти α_1, решаем одномерную задачу оптимизации

$$F(x^{(1)}+\alpha_1 h^{(1)})\underset{\alpha_1\in R}{\to}\min \qquad (11.57)$$

(задача минимизации по аргументу α_1, где $x^{(1)},h^{(1)}\in R^n$ известны). Так как

$$F(x^{(1)}+\alpha_1 h^{(1)})=F(x^{(1)})+\alpha_1^2(Ah^{(1)},h^{(1)})+2\alpha_1(Ax^{(1)}-b,h^{(1)})$$

(полином второй степени относительно α_1), решением (11.57) является

$$\alpha_1=-\frac{(Ax^{(1)}-b,h^{(1)})}{(Ah^{(1)},h^{(1)})} \qquad (11.58)$$

(аргумент вершины параболы).

Третий шаг и далее

Шаг $s+1$, где $s \geq 2$ (т.е. третий шаг и далее) аналогичен шагу 2. Приближение $x^{(s+1)} \in R^n$ находим в виде

$$x^{(s+1)} = x^{(s)} + \alpha_s h^{(s)} \tag{11.59}$$

где направление (вектор) $h^{(s)} \in R^n$ определяется невязкой предыдущего шага

$$r^{(s)} = Ax^{(s)} - b$$

и предыдущим направлением $h^{(s-1)}$:

$$h^{(s)} = -r^{(s)} + \beta_s h^{(s-1)} \tag{11.60}$$

Нужно, чтобы $h^{(s)}, h^{(s-1)}$ были сопряженными относительно $A = A^T > 0$:

$$(Ah^{(s-1)}, h^{(s)}) = 0.$$

Получим условие

$$(Ah^{(s-1)}, h^{(s)}) = (Ah^{(s-1)}, h^{(s-1)})\beta_s - (Ah^{(s-1)}, r^{(s)}) = 0$$

откуда следует

$$\beta_s = \frac{(Ah^{(s-1)}, r^{(s)})}{(Ah^{(s-1)}, h^{(s-1)})} \tag{11.61}$$

Чтобы найти α_s, решаем одномерную задачу оптимизации

$$F(x^{(s)} + \alpha_s h^{(s)}) \underset{\alpha_s \in R}{\rightarrow} \min \tag{11.62}$$

(задача минимизации по аргументу α_s, где $x^{(s)}, h^{(s)} \in R^n$ известны).

Так как

$$F(x^{(s)} + \alpha_s h^{(s)}) = F(x^{(s)}) + \alpha_s^2 (Ah^{(s)}, h^{(s)}) + 2\alpha_s (Ax^{(s)} - b, h^{(s)})$$

(полином второй степени относительно α_s), решением (11.62) является

$$\alpha_s = -\frac{(Ax^{(s)} - b, h^{(s)})}{(Ah^{(s)}, h^{(s)})} \tag{11.63}$$

(аргумент вершины параболы).

Результат работы метода

На шаге $s+1$ будет построен $x^{(s+1)} \in R^n$, его связь с начальным приближением $x^{(0)} \in R^n$ описывается формулой

$$
\begin{aligned}
x^{(s+1)} &= x^{(s)} + \alpha_s h^{(s)} = x^{(s-1)} + \alpha_{s-1} h^{(s-1)} + \alpha_s h^{(s)} = .. \\
&= x^{(0)} + \alpha_0 h^{(0)} + \ldots + \alpha_s h^{(s)}
\end{aligned}
\qquad (11.64)
$$

Здесь сопряженные направления $h^{(i)}, i = 0,\ldots s$ вычислены по формулам (11.51), (11.55) и (11.60), коэффициенты $\alpha_i, i = 0,\ldots s$ вычислены по формулам (11.53), (11.58) и (11.63), коэффициенты $\beta_i, i = 1,\ldots s$, необходимые для расчета сопряженных направлений – по формулам (11.56) и (11.61).

Основные свойства метода

Свойство 1. Если на шаге $s+1$ получено $r^{(s+1)} = 0$, то $x^{(s+1)}$ – точное решение СЛАУ (11.43) и задачи оптимизации (11.44).

Свойство 2. Если в процессе работы метода получены невязки $r^{(0)},\ldots\ r^{(s+1)} \neq 0$ (то есть за $s+1$ шагов точное решение задач (11.43) и (11.44) еще не найдено), тогда:

– невязки $r^{(0)},\ldots\ r^{(s+1)} \in R^n$ взаимно ортогональны;

– векторы $h^{(0)},\ldots h^{(s)} \in R^n$ взаимно сопряжены;

– значение коэффициента α_s, заданного формулой (11.53), (11.58) или (11.63), совпадает со значением, заданным формулой (11.49), то есть.

$$
\alpha_s = -\frac{(Ax^{(s)} - b, h^{(s)})}{(Ah^{(s)}, h^{(s)})} = -\frac{(Ax^{(0)} - b, h^{(s)})}{(Ah^{(s)}, h^{(s)})};
$$

– приближение $x^{(s+1)} \in R^n$ обеспечивает минимальное значение функционала $F(x) = (Ax, x) - 2(b, x)$ на многообразии $L_{s+1}(x_0, h^{(0)}, \ldots. h^{(s)})$.

Свойство 3. Не позднее чем на шаге n метод сопряженных градиентов строит точное решение СЛАУ (11.43) и задачи оптимизации (11.44).

Комментарии к параграфу 11.4

1) Метод сопряженных градиентов может использоваться как прямой метод или как итерационный.

Если на шаге n или ранее получено точное решение задач (11.43) и (11.44) – значит, метод использован как прямой.

Если каждое $x^{(s)}$ рассматривается как приближенное решение задач (11.43) и (11.44) – метод используется как итерационный.

2) Приближение $x^{(s+1)}$ лучше, чем предыдущее приближение $x^{(s)}$, так как обеспечивает минимальное значение функционала $F(x)$ на многообразии $L_{s+1}(x_0, h^{(0)}, \ldots. h^{(s)})$ размерности $s+1$, включающем предыдущее многообразие $L_s(x_0, h^{(0)}, \ldots. h^{(s-1)})$ размерности s, и соответственно значение функционала $F(x)$ с каждым шагом убывает (не возрастает).

3) В силу накопления вычислительной погрешности взаимно сопряженные

направления строятся приближенно и при большом числе шагов s векторы

$$h^{(0)},...h^{(s)} \in R^n$$

теряют свойство взаимной сопряженности.

Если есть проблемы отыскания $x^{(s+1)}$, текущее приближение $x^{(s)}$ можно принять за начальное приближение и запустить метод заново (то есть заново строить взаимно сопряженные направления).

4) В качестве численного решения СЛАУ (11.43) и задачи оптимизации (11.44) может быть выбрано такое приближение $x^{(s)}$, для которого:

– выполнен критерий отыскания точного решения: $r^{(s)} = 0$, то есть $x^{(s)} = x^*$,

– либо выполнен критерий остановки по точности: $\left\| x^{(s)} - x^{(s-1)} \right\| \le \varepsilon$;

– либо выполнен критерий остановки по числу шагов: $s+1 > N_{max}$;

– решение СЛАУ найдено с достаточно малой невязкой: $\left\| r^{(s)} \right\| \le \varepsilon^*$ и др.

5) погрешность решения СЛАУ на шаге s можно оценить по текущей невязке, используя норму обратной матрицы или ее оценку:

$\left\| z^{(s)} \right\| \le \left\| A^{-1} \right\| \cdot \left\| r^{(s)} \right\|$ (нормы матрицы и вектора должны быть согласованы):

6) оценки погрешности метода на шаге s по *начальной невязке* см. в литературе.

Модуль 11 – Практикум по темам 11.1-11.4.
Решение типовых задач на примере метода простой итерации

Пусть СЛАУ $Ax = b$ имеет вид

$$\begin{bmatrix} 3 & -0.8 & 0.2 \\ -0.8 & 9 & 1.8 \\ 0.2 & 1.8 & 13 \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3.2 \\ 1 \\ 13.2 \end{bmatrix} \qquad x^{(0)} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

Необходимо:

1. Оценить собственные числа и норму, указать свойства матрицы на основе кругов Гершгорина
2. Выписать метод простой итерации и подобрать параметр для решения СЛАУ,
   указать способ оценки погрешности метода
3. Выполнить два шага метода, используя указанное выше начальное приближение
   На каждом шаге вычислить невязку и точность, вычислить норму невязки (max)
4. Оценить собственные числа и норму, указать свойства обратной матрицы на основе кругов Гершгорина
5. Построить оценку погрешности метода на шаге 2 по текущей невязке
6. Построить оценку погрешности метода на шаге 2 по начальной невязке
7. Оценить число шагов, гарантирующих решение задачи с погрешностью не более 0.001

Решение

1. Оценить собственные числа и норму, указать свойства матрицы на основе кругов Гершгорина

1) Собственные числа (их 3, так как матрица 3×3) расположены на комплексной плоскости в кругах Гершгорина с центрами в точках 3, 9, 13 и радиусами 1, 2.4, 2:

$$|z-3| \le |0.2| + |-0.8| \Rightarrow |z-3| \le 1$$

$$|z-9| \le |1.8| + |-0.8| \Rightarrow |z-9| \le 2.4$$

$$|z-13| \le |0.2| + |1.8| \Rightarrow |z-13| \le 2$$

Объединение кругов состоит из двух связных компонент, одна компонента (один круг), содержит одно собственное число, вторая компонента (объединение двух кругов) содержит два собственных числа

2) Число 0 в круги не попадает. Значит, нет нулевого собственного числа и $\det A \ne 0$. Решение СЛАУ $\exists!$

3) Матрица симметрична, поэтому собственные числа действительные.

Одно из них на участке [2; 4], остальные два на участке [6.6; 15].

Минимальное собственное число положительно, оно не меньше 2 и не больше 4

Максимальное собственное число положительное , оно не меньше 6.6 и не больше 15

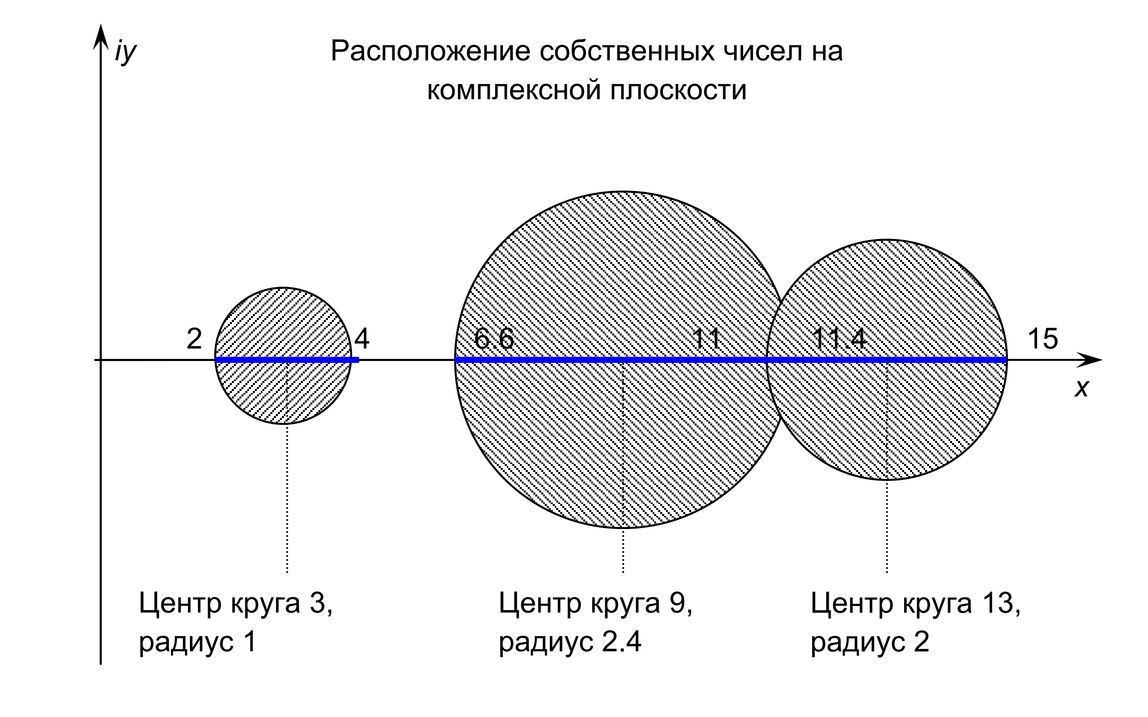

Рисунок 3

4) Матрица симметрична, все собственные числа положительные, следовательно, матрица положительно определена: для нее верно $\forall h \neq 0, h \in R^n \ \ (Ah, h) > 0$

5) Спектральный радиус матрицы есть расстояние от нуля (на комплексной плоскости) до наиболее удаленного от нуля собственного числа:

$\rho\,(A) = \max\limits_{i=1,n} \left| \lambda_i\,(A) \right|$, в нашем случае n = 3

Спектральный радиус симметричной положительно определенной матрицы равен ее максимальному собственному числу (потому что именно оно является наиболее удаленным от нуля)

Спектральный радиус матрицы СЛАУ не меньше 6.6 и не больше 15.

6) «Евклидова» норма симметричной матрицы равна спектральному радиусу

$$\| A \|_2 = \rho(A)$$

«Евклидова» норма матрицы СЛАУ оценивается как $6.6 \le \| A \|_2 \le 15$

Число обусловленности симметричной матрицы

$$\mu_A = \| A \|_2 \, \| A^{-1} \|_2 = \frac{\lambda_n}{\lambda_1}$$

Его можно оценить

$$1.65 = \frac{6.6}{4} \le \mu_A = \frac{\lambda_n}{\lambda_1} \le \frac{15}{2} = 7.5$$

матрица неплохо обусловлена

2. Выписать метод простой итерации и подобрать параметр для решения СЛАУ, указать способ оценки погрешности метода

СЛАУ $Ax = b$, $\det A \neq 0$

Метод простой итерации имеет вид

$$x^{(s+1)} = x^{(s)} - \tau \cdot r^{(s)}$$

$r^{(s)} = Ax^{(s)} - b$ – невязка СЛАУ на текущем шаге

τ – число (постоянный параметр метода)

Так как по условию задачи собственные числа вычислять не положено, применим результаты, основанные на оценках собственных чисел.

Для подбора метода используем Теоремы 4 и 5 (Модуль 11).

Матрица СЛАУ симметрична и положительно определена.

Все собственные числа в диапазоне от 2 до 15.

В формулировке Теоремы 4 и 5 берем $M_{\min} = 2, M_{\max} = 15$

По Теореме 4 достаточное условие сходимости метода $\tau \in (0; \frac{2}{M_{\max}})$

Оценка погрешности

$$\left\| z^{(s)} \right\|_2 \leq \left(\max\{\left|1 - \tau M_{\min}\right|, \left|1 - \tau M_{\max}\right|\}\right)^s \left\| z^{(0)} \right\|_2$$

По Теореме 5 оптимальный (для тех оценок собственных чисел, которыми располагает исследователь) выбор есть

$$\tilde{\tau}^* = \frac{2}{M_{\min} + M_{\max}} \in (0; \frac{2}{M_{\max}})$$

Оценка погрешности

$$\left\| z^{(s)} \right\|_2 \leq \left(\frac{M - 1}{M + 1}\right)^s \left\| z^{(0)} \right\|_2$$

число M является верхней оценкой числа обусловленности μ_A (определенного в евклидовой норме).

$$1 \leq \mu_A = \frac{\lambda_n}{\lambda_1} \leq \frac{M_{\max}}{M_{\min}} = M$$

Подберем метод по Теореме 4:

Достаточные условия сходимости $\tau \in (0; \frac{2}{15})$, например, $\tau = \frac{1}{8}$

Оценка погрешности метода

$$\left\| z^{(s)} \right\|_2 \le \left(\max\left\{ \left| 1 - \frac{1}{8} M_{\min} \right|, \left| 1 - \frac{1}{8} M_{\max} \right| \right\} \right)^s \left\| z^{(0)} \right\|_2$$

$$\left\| z^{(s)} \right\|_2 \le \left(\max\left\{ \left| 1 - \frac{2}{8} \right|, \left| 1 - \frac{15}{8} \right| \right\} \right)^s \left\| z^{(0)} \right\|_2$$

$$\left\| z^{(s)} \right\|_2 \le \left(\max\left\{ \frac{3}{4}, \frac{7}{8} \right\} \right)^s \left\| z^{(0)} \right\|_2$$

$$\left\| z^{(s)} \right\|_2 \le \left(\frac{7}{8} \right)^s \left\| z^{(0)} \right\|_2$$

При $s \to +\infty$ $\left(\frac{7}{8} \right)^s \to 0$ и поэтому для любого начального приближения

$\left\| z^{(s)} \right\|_2 \to 0$.

Метод сходится.

Подберем метод по Теореме 5:

$$\tilde{\tau}^* = \frac{2}{15+2} = \frac{2}{17} \in \left(0; \frac{2}{15}\right)$$

$$M = \frac{M_{\max}}{M_{\min}} = \frac{15}{2} = 7.5$$

$$\frac{M-1}{M+1} = \frac{6.5}{8.5} = \frac{65}{85} = 0.76471$$ (с округлением в большую сторону)

Оценка погрешности метода

$$\left\| z^{(s)} \right\|_2 \le \left(\frac{65}{85} \right)^s \left\| z^{(0)} \right\|_2$$

При $s \to +\infty$ $\left(\frac{65}{85} \right)^s \to 0$ и поэтому для любого начального приближения

$\left\| z^{(s)} \right\|_2 \to 0$.

Метод сходится.

Сравним 2 способа

По теореме 4 $\tau = \dfrac{1}{8}$ $\quad \| z^{(s)} \|_2 \le \left(\dfrac{7}{8}\right)^s \| z^{(0)} \|_2$

По теореме 5 $\tilde{\tau}^* = \dfrac{2}{17}$ $\quad \| z^{(s)} \|_2 \le \left(\dfrac{65}{85}\right)^s \| z^{(0)} \|_2$

7/8 = 0.875 65/85 = 0.76471

По теореме 5 оценка сходимости лучше. Далее используем $\tilde{\tau}^* = \dfrac{2}{17}$

3. Выполнить <u>два шага</u> метода, используя указанное выше начальное приближение

На каждом шаге вычислить невязку и точность, вычислить норму невязки (max)

$$x^{(0)} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$ начальное приближение

$$r^{(0)} = \begin{bmatrix} 3 & -0.8 & 0.2 \\ -0.8 & 9 & 1.8 \\ 0.2 & 1.8 & 13 \end{bmatrix} \times \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} - \begin{bmatrix} 3.2 \\ 1 \\ 13.2 \end{bmatrix} = \begin{bmatrix} -0.8 \\ 9 \\ 1.8 \end{bmatrix} - \begin{bmatrix} 3.2 \\ 1 \\ 13.2 \end{bmatrix} = \begin{bmatrix} -4.0 \\ 8 \\ -11.4 \end{bmatrix}$$

(начальная невязка)

Первый шаг метода

$x^{(1)} = x^{(0)} - \dfrac{2}{17} \cdot r^{(0)}$ – расчет на шаге 1

$$x^{(1)} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} - \frac{2}{17} \cdot \begin{bmatrix} -4.0 \\ 8 \\ -11.4 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} - \begin{bmatrix} -0.4706 \\ 0.9412 \\ -1.3412 \end{bmatrix} = \begin{bmatrix} 0.4706 \\ 0.0588 \\ 1.3412 \end{bmatrix}$$ – первый шаг завершен

$$r^{(1)} = \begin{bmatrix} 3 & -0.8 & 0.2 \\ -0.8 & 9 & 1.8 \\ 0.2 & 1.8 & 13 \end{bmatrix} \times \begin{bmatrix} 0.4706 \\ 0.0588 \\ 1.3412 \end{bmatrix} - \begin{bmatrix} 3.2 \\ 1 \\ 13.2 \end{bmatrix} = \begin{bmatrix} 1.6329 \\ 2.5671 \\ 17.6353 \end{bmatrix} - \begin{bmatrix} 3.2 \\ 1 \\ 13.2 \end{bmatrix} = \begin{bmatrix} -1.5671 \\ 1.5671 \\ 4.4353 \end{bmatrix}$$

(невязка на шаге 1)

Второй шаг метода

$x^{(2)} = x^{(1)} - \dfrac{2}{17} \cdot r^{(1)}$ – расчет на шаге 2

$$x^{(2)} = \begin{bmatrix} 0.4706 \\ 0.0588 \\ 1.3412 \end{bmatrix} - \frac{2}{17} \cdot \begin{bmatrix} -1.5671 \\ 1.5671 \\ 4.4353 \end{bmatrix} = \begin{bmatrix} 0.4706 \\ 0.0588 \\ 1.3412 \end{bmatrix} - \begin{bmatrix} -0.1844 \\ 0.1844 \\ 0.5218 \end{bmatrix} = \begin{bmatrix} 0.6549 \\ -0.1255 \\ 0.8194 \end{bmatrix}$$

Второй шаг завершен.

$$r^{(2)} = \begin{bmatrix} 3 & -0.8 & 0.2 \\ -0.8 & 9 & 1.8 \\ 0.2 & 1.8 & 13 \end{bmatrix} \times \begin{bmatrix} 0.6549 \\ -0.1255 \\ 0.8194 \end{bmatrix} - \begin{bmatrix} 3.2 \\ 1 \\ 13.2 \end{bmatrix} = \begin{bmatrix} 2.2291 \\ -0.1789 \\ 10.5569 \end{bmatrix} - \begin{bmatrix} 3.2 \\ 1 \\ 13.2 \end{bmatrix} = \begin{bmatrix} -0.9709 \\ -1.1789 \\ -2.6431 \end{bmatrix}$$

(невязка на шаге 2)

Подсчет нормы невязки на каждом шаге (норма max)

$$\left\| r^{(s)} \right\|_\infty = \max \{ \left| r_1^{(s)} \right|. \left| r_2^{(s)} \right|, \left| r_3^{(s)} \right| \}$$

Начальная невязка (невязка на шаге 0)

$$\left\| r^{(0)} \right\|_\infty = \max \{ |-4|. |8|, |-11.4| \} = 11.4$$

Невязка на шаге 1

$$\left\| r^{(1)} \right\|_\infty = \max \{ |-1.5671|. |1.5671|, |4.4353| \} = 4.4353$$

Невязка на шаге 2

$$\left\| r^{(2)} \right\|_\infty = \max \{ |-0.9709|. |-1.1789|, |-2.6431| \} = 2.6431$$

Невязка нужна для оценки погрешности. Как и следовало ожидать для сходящегося метода, в данном случае невязка убывает.

Точность на шаге (так называется критерий остановки метода)

$$\varepsilon_s = \left\| x^{(s)} - x^{(s-1)} \right\|_\infty = \max \{ \left| x_1^{(s)} - x_1^{(s-1)} \right|. \left| x_2^{(s)} - x_2^{(s-1)} \right|, \left| x_3^{(s)} - x_3^{(s-1)} \right| \}$$

Точность на шаге 1

$$\varepsilon_1 = \left\| x^{(1)} - x^{(0)} \right\|_\infty = \max \{ |0.4706 - 0|. |0.0588 - 1|, |1.3412 - 0| \} = 0.9412$$

Точность на шаге 2

$$\varepsilon_2 = \left\| x^{(2)} - x^{(1)} \right\|_\infty =$$

$$= \max \{ |0.6549 - 0.4706|, |-0.1255 - 0.0588|, |0.8194 - 1.3412| \} = 0.5218$$

Как и следовало ожидать для сходящегося метода, показатель ε_s убывает.

4. Оценить <u>собственные числа</u> и <u>норму</u>, указать <u>свойства</u> обратной матрицы на основе кругов Гершгорина

1) Собственные числа обратной матрицы являются обратными величинами по отношению к собственным числам матрицы СЛАУ (см. Модуль 10)

$$\lambda_i\ (A^{-1}) = \frac{1}{\lambda_i\ (A)},\ i = 1,...n.$$

В данном случае собственные числа обратной матрицы действительные и положительные.

Максимальное собственное число обратной матрицы не меньше 1/4 и не больше 1/2

Минимальное собственное число обратной матрицы не меньше 1/15 и не больше 10/66

2) Спектральный радиус обратной матрицы равен ее максимальному собственному числу (потому что именно оно является наиболее удаленным от нуля)

Спектральный радиус обратной матрицы СЛАУ не меньше 1/4 и не больше 1/2.

3) Обратная матрица симметрична, потому что матрица СЛАУ симметрична:

$$(A^{-1})^T = (A^T)^{-1} = A^{-1}$$

4) «Евклидова» норма симметричной матрицы равна спектральному радиусу, поэтому для обратной матрицы верно

$$\| A^{-1} \|_2 = \rho(A^{-1})$$

«Евклидова» норма обратной матрицы оценивается как

$$\frac{1}{4} \le \| A^{-1} \|_2 \le \frac{1}{2}$$

Эта оценка нужна для того, чтобы оценить погрешность по невязке

Число обусловленности обратной матрицы такое же, как исходной матрицы:

Если (по определению) в евклидовых нормах $\mu_A = \| A \|_2 \| A^{-1} \|_2$ то

$$\mu_{A^{-1}} = \| A^{-1} \|_2 \| \left[A^{-1}\right]^{-1} \|_2 = \left\|A^{-1}\right\|_2 \left\|A\right\|_2 = \mu_A$$

Число обусловленности обратной матрицы можно оценить тем же границами, что и число обусловленности матрицы СЛАУ, то есть

$$1.65 = \frac{6.6}{4} \le \mu_A = \mu_{A^{-1}} \le \frac{15}{2} = 7.5$$

Можно перепроверить оценку

$$1.65 = \frac{6.6}{4} \le \mu_{A^{-1}} = \frac{\lambda_n(A^{-1})}{\lambda_1(A^{-1})} \le \frac{\left(\frac{1}{2}\right)}{\left(\frac{1}{15}\right)} = \frac{15}{2} = 7.5$$

5. Построить оценку погрешности метода на шаге 2 по <u>текущей невязке</u>

Оценка погрешности по текущей невязке одинакова для любого метода

$$\left\| z^{(s)} \right\|_2 \le \left\| A^{-1} \right\|_2 \cdot \left\| r^{(s)} \right\|_2,$$

где $r^{(s)} = Ax^{(s)} - b$ и ее можно вычислить.

Оценим погрешность метода на шаге 2

$$\left\| z^{(2)} \right\|_2 \le \left\| A^{-1} \right\|_2 \cdot \left\| r^{(2)} \right\|_2,$$

$\left\| A^{-1} \right\|_2 \le \frac{1}{2}$ разобрано выше, используем верхнюю оценку

$$\left\| z^{(2)} \right\|_2 \le \frac{1}{2} \cdot \left\| r^{(2)} \right\|_2$$

$$\left\| r^{(2)} \right\|_2 = \sqrt{(-0.9709)^2 + (-1.1789)^2 + (-2.6431)^2} = 3.0526$$

$\left\| z^{(2)} \right\|_2 \le \frac{1}{2} \cdot 3.0526 = 1.5263$ оценили по текущей невязке

6. Построить оценку погрешности метода на шаге 2 по <u>начальной невязке</u>

Оценим погрешность метода по начальной погрешности (см. теорему о сходимости), а начальную погрешность – по начальной невязке

Оценим начальную погрешность

$$\left\| z^{(0)} \right\|_2 \le \left\| A^{-1} \right\|_2 \cdot \left\| r^{(0)} \right\|_2,$$

$\left\| A^{-1} \right\|_2 \le \frac{1}{2}$ разобрано выше, используем верхнюю оценку

$$\left\| z^{(0)} \right\|_2 \le \frac{1}{2} \cdot \left\| r^{(0)} \right\|_2$$

$$\left\| r^{(0)} \right\|_2 = \sqrt{(-4)^2 + 8^2 + (11.4)^2} = \sqrt{16 + 64 + 12\,996} = \sqrt{209.96} = 14.49000$$

$$\left\| z^{(0)} \right\|_2 \le \frac{1}{2} \cdot 14.49 = 7.245$$

Оценка погрешности метода на шаге 2

$$\left\| z^{(s)} \right\|_2 \le \left(\frac{65}{85}\right)^s \left\| z^{(0)} \right\|_2$$

$$\left\| z^{(s)} \right\|_2 \le \left(\frac{65}{85}\right)^s \left\| z^{(0)} \right\|_2 \le \left(\frac{65}{85}\right)^s \cdot 7.245$$

$$\left\| z^{(2)} \right\|_2 \le \left(\frac{65}{85}\right)^2 \left\| z^{(0)} \right\|_2 \le \left(\frac{65}{85}\right)^2 \cdot 7.245 = 0.5847751 \cdot 7.245 = 4.23670$$

Сравним оценки

$\left\| z^{(2)} \right\|_2 \le 1.5263$ оценка погрешности шага 2 по текущей невязке

$\left\| z^{(2)} \right\|_2 \le 4.23670$ оценка погрешности шага 2 по начальной невязке

Первая оценка точнее (лучше).

7. Оценить <u>число шагов</u>, <u>гарантирующих</u> решение задачи с погрешностью <u>не более 0.001</u>

$$\left\| z^{(N)} \right\|_2 \le \left(\frac{65}{85}\right)^N \left\| z^{(0)} \right\|_2 \le \left(\frac{65}{85}\right)^N \cdot 4.2367$$

Найдем такое N, чтобы $\left(\frac{65}{85}\right)^N \cdot 4.2367 \le 0.001$

$$N \cdot \lg\left(\frac{65}{85}\right) + \lg(4.2367) \le \lg(0.001)$$

$$N \cdot \ge \frac{\lg(0.001) - \lg(4.2367)}{\lg(65) - \lg(85)} \qquad N \cdot \ge \frac{-3.62703}{0.11651} = 31.1318 \qquad N \ge 32$$

Чтобы гарантировать погрешность не более 0.001, нужно 32 шага.

ЛИТЕРАТУРА

а) литература по тематическому блоку

1. Бахвалов Н.С., Жидков Н.П., Кобельков Г.М. Численные методы. – 7-е изд. М.: БИНОМ. Лаборатория знаний, 2011. – 636 с.
2. Вержбицкий В.М. Основы численных методов. – М.: Высшая школа, 2002. – 840 с.
3. Воеводин В.В. Вычислительные основы линейной алгебры. М.: Главная редакция физико-математической литературы изд-ва «Наука», 1973. – 304 с.
4. Демидович Б.П., Марон И.А. Основы вычислительной математики. М.: Физматгиз, 1963. – 660 с.
5. Марчук Г.И. Методы вычислительной математики. М.: Наука, 1980. – 536 с.
6. Методы вычислительной математики / Ред. Г.И. Марчук, Ж.-Л. Лионс. Новосибирск: Наука, 1975.
7. Самарский А.А., Вабищевич П.Н. Численные методы решения задач конвекции-диффузии. – М.: Книжный дом «Либроком», 2015. – 248 с.
8. Самарский А.А., Гулин А.В. Численные методы математической физики. М.: Научный мир, 2000. – 316 с.
9. Самарский А.А., Гулин А.В. Численные методы. М.: Наука, 1989.
10. Самарский А.А. Теория разностных схем. М.: Наука, 1989. – 616 с.
11. Фаддеев Д.К., Фаддеева В.Н. Вычислительные методы линейной алгебры. СПб: Лань, 2002. – 736 с.
12. Стронгина Н.Р., Баркалов К.А. Численные методы. Семестр 7. ЭУК, учебно-методический комплекс. Фонд электронных образовательных ресурсов ННГУ. Н. Новгород, 2014. Ид.н. 815Е.14.08.
13. Стронгина Н.Р., Баркалов К.А. Практикум по курсу «Численные методы». Применение итерационных методов решения разностных схем на примере решения задачи Дирихле для уравнения Пуассона.: Учебно-методическое пособие. – Нижний Новгород: Нижегородский госуниверситет, 2011. – 40 с.
14. Практические приложения численных методов линейной алгебры. Учебно-методический комплекс для поддержки общего курса «Численные методы» (направление «Прикладная математика и информатика»). Результаты образовательного проекта лаборатории «Информационные технологии» факультета ВМК ННГУ (отчет). Авторы: Стронгина Н.Р., Балабанов А.С., Баркалов К.А., Ирхина А.Л., Федоткин А.М., Юсов Е.А. / Под ред. Н.Р. Стронгиной. - Н. Новгород: Нижегородский госуниверситет, 2004. – 281 с.
15. Перов А.А., Протогенов А.П. Численные методы в физических исследованиях. – Н. Новгород: Нижегородский университет, 2019. – 69 с.

16. Морозов О.А., Семин Ю.А. Моделирование физических процессов и систем. Часть 2: моделирование систем с несколькими пространственными измерениями: Учебно-методическое пособие. – Нижний Новгород: Нижегородский госуниверситет, 2014. – 43 с.
17. Параллельные вычисления: технологии и численные методы: Учебное пособие в 4-х томах. Авторы: Гергель В.П., Баркалов К.А., Мееров И.Б. и др. Том 3: Элементы компьютерной арифметики. Прямые и итерационные методы решения СЛАУ. – Нижний Новгород: Изд-во Нижегородского университета, 2013. – 415 с.

б) литература об организации учебного процесса по дисциплине

18. Садовничий В.А. Международный форум «Университеты, общество и будущее человечества». Доклад ректора МГУ имени М.В. Ломоносова академика В.А. Садовничего на Международном форуме «Университеты, общество и будущее человечества» 25 марта 2019 года. М.: Издательство Московского университета, 2019. – 36 с.
19. Высокопроизводительные параллельные вычисления. 100 заданий для расширенного лабораторного практикума. М.: ФИЗМАТЛИТ, 2018. – 248 с.
20. Программы дисциплин по направлению «Прикладная математика и информатика». Учебно-методическое объединение Университетов. Учебно-методический совет по прикладной математике и информатике. М.: Изд-во факультета ВМиК МГУ, 2002. С. 59 – 62.
21. Стронгина Н.Р. Цифровизация и качество обучения на примере фундаментальной дисциплины «Численные методы» // Научные вести. 2021. №2 (31). – С. 85 – 103.

Наталья Романовна **Стронгина**

КУРС «ЧИСЛЕННЫЕ МЕТОДЫ»

Итерационные методы решения СЛАУ для вычислительно-трудоемких задач

(Модули 10 – 11)

Учебно-методическое пособие

Федеральное государственное автономное образовательное учреждение
высшего образования «Национальный исследовательский
Нижегородский государственный университет им. Н.И. Лобачевского».
603950, Нижний Новгород, пр. Гагарина, 23.