

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ ОБРАЗОВАТЕЛЬНОЕ  
УЧРЕЖДЕНИЕ ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ  
«ЮЖНЫЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»

С. Н. Овчинникова, И. В. Моршнева

## **Численные методы**

Вычислительные методы линейной алгебры

УЧЕБНОЕ ПОСОБИЕ

Часть II

Ростов-на-Дону

2008 год

Рецензенты:

доцент кафедры фундаментальной и прикладной математики

РГЭУ "РИНХ",

кандидат физ.-мат. наук

Рогожин С.В.

профессор кафедры теоретической гидроаэромеханики ЮФУ,

доктор физ.-мат. наук

Батищев В.А.

**Овчинникова С. Н., Моршнева И. В.**

Численные методы. Часть II. Вычислительные методы линейной алгебры: Учебное пособие. – Ростов-на-Дону, 2008.–75с.

В пособии излагаются численные методы решения систем линейных алгебраических уравнений. Теоретический материал иллюстрируется рисунками, для некоторых методов приводятся алгоритмы, их реализующие. В каждом разделе предлагаются контрольные вопросы и примеры для самостоятельной работы. В приложении приводятся некоторые теоретические сведения, которые используются в основном тексте. Пособие предназначено для студентов математических и инженерных специальностей. Оно может быть полезно для широкого круга читателей, интересующихся вычислительной математикой.

## Содержание

<b>1</b>	<b>Вычислительные методы алгебры</b>	<b>5</b>
<b>2</b>	<b>Методы численного решения систем линейных алгебраических уравнений</b>	<b>6</b>
2.1	Нормы матриц и векторов . . . . .	9
2.2	Погрешность приближенного решения системы и обусловленность матриц . . . . .	11
2.3	Матрицы специального вида. . . . .	18
2.4	Методы решения СЛАУ . . . . .	22
2.5	Вопросы и упражнения для самопроверки . . . . .	22
<b>3</b>	<b>Точные методы решения СЛАУ</b>	<b>24</b>
3.1	LU - разложение матрицы . . . . .	25
3.2	Метод Гаусса . . . . .	31
3.3	Ортогональные (QR) разложения матриц . . . . .	34
3.3.1	Преобразование вращения(преобразование Гивенса) . . . . .	34
3.3.2	Преобразование отражения (преобразования Хаусхолдера) . . . . .	37
3.4	Ортогональные методы решения СЛАУ . . . . .	39
3.5	Метод Холесского (метод квадратного корня) . . . . .	42
3.6	Вопросы и упражнения для самопроверки . . . . .	44
<b>4</b>	<b>Итерационные методы решения СЛАУ</b>	<b>46</b>
4.1	Метод простой итерации . . . . .	49
4.2	Метод Якоби . . . . .	51
4.3	Метод Зейделя . . . . .	53
4.4	Методы спуска для решения СЛАУ . . . . .	56
4.4.1	Метод покоординатного спуска. . . . .	57
4.4.2	Метод градиентного спуска . . . . .	59

4.4.3	Метод наискорейшего градиентного спуска . . . . .	60
4.5	Методы релаксации . . . . .	61
4.6	Практическая оценка погрешности итерационного метода. . . . .	63
4.7	Вопросы и упражнения для самопроверки . . . . .	67
<b>5</b>	<b>Приложение</b>	<b>69</b>
5.0.1	Необходимые сведения из линейной алгебры и теории матриц . . . . .	69
5.0.2	Градиент функционала . . . . .	71
5.0.3	Ответы, указания, решения . . . . .	72

**Цель вычислений в проникновении в суть, а не в цифрах.**

**Цель расчетов — понимание, а не числа.**

**Р. Хемминг.**

## **1 Вычислительные методы алгебры**

Вычислительные методы алгебры давно вошли в число основных инструментов, используемых при исследованиях в различных областях математики и её приложений. Этим, по-видимому, объясняется существование многочисленных пакетов стандартных подпрограмм (СП), реализующих вычислительные алгоритмы алгебры. Обычно такие программы тщательно протестированы и для большого класса задач эффективны, т.е. обеспечивают заданную точность расчета (см., например, пакеты IMSL и LINPACK из США, NAG из Англии, написанные на языке fortran). Поэтому, как правило, нет необходимости писать собственные программы. Однако, чтобы понимать принятую в стандартных подпрограммах информацию о вырожденных ситуациях или об ошибках пользователя, требуются теоретические знания как самих методов вычисления, так и реализующих их алгоритмов. Такие же знания полезны при выборе подходящих СП, гарантирующих требуемую точность вычисленного решения, ведь эффективность метода зависит от того, насколько полно в нем используется специфика решаемой задачи. Кроме того, иногда появляются задачи, для которых требуется модифицировать старые или создавать новые методы и алгоритмы.

В этом разделе курса лекций "Численные методы" рассматривается несколько наиболее распространенных методов решения задач линейной алгебры.

## 2 Методы численного решения систем линейных алгебраических уравнений

Пусть требуется найти решение системы линейных алгебраических уравнений (СЛАУ)

$$A \bar{x} = \bar{b}, \quad (2.1)$$

где  $A = \{a_{ij}\}_{i,j=1}^m$  — квадратная матрица размера  $(m \times m)$ , а  $\bar{x} = \{x_i\}_{i=1}^m$  и  $\bar{b} = \{b_i\}_{i=1}^m$  — вектор-столбцы ( $\bar{x} \in R^m$  и  $\bar{b} \in R^m$ ).

Теоретически при решении СЛАУ (2.1) возможны три ситуации:

1. ранг матрицы  $A$  равен  $m$  и равен рангу расширенной матрицы  $\{Ab\}$  — система (2.1) имеет единственное решение;
2. ранг расширенной матрицы  $\{Ab\}$  равен рангу  $A$  и меньше  $m$  — система (2.1) имеет бесконечное множество решений;
3. ранг матрицы  $\{Ab\}$  не равен рангу  $A$  — система (2.1) не имеет решений (несовместна).

Решить систему линейных уравнений второго порядка ( $m = 2$ )

$$a_{11}x_1 + a_{12}x_2 = b_1$$

$$a_{21}x_1 + a_{22}x_2 = b_2$$

геометрически означает найти точку пересечения двух прямых, заданных уравнениями системы. Для такой системы три возможные при решении СЛАУ ситуации можно изобразить следующим образом

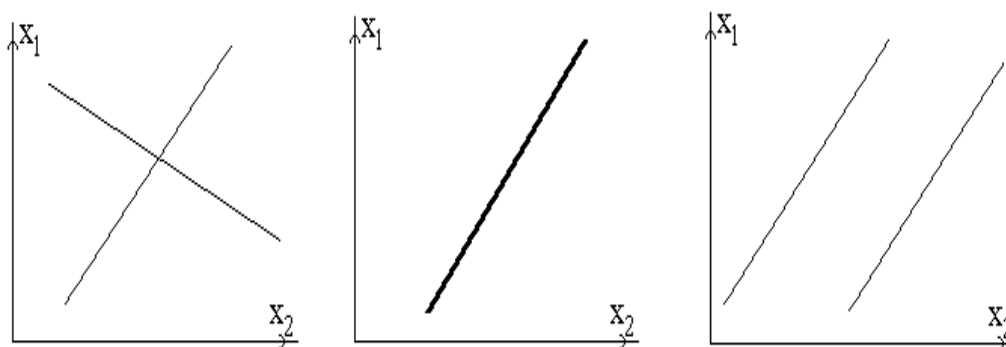


Рис. 1: Различные ситуации, возникающие при решении систем линейных алгебраических систем второго порядка

В первом случае система имеет единственное решение (СЛАУ не вырождена), во втором — бесконечно много решений и в третьем случае — ни одного (в обоих этих случаях система вырождена).

В практических вычислениях возможен ещё один случай. Рассмотрим пример системы

$$5x_1 + 7x_2 = 12,$$

$$7x_1 + 10x_2 = 17,$$

которая имеет единственное решение  $x_1 = 1$ ,  $x_2 = 1$ .

Пусть расчет ведется в множестве  $F$  чисел с плавающей запятой, у которого  $t = 2$ ,  $\beta = 10$  ( $t$  — количество разрядов для записи мантиссы числа в  $F$ ,  $\beta$  — основание системы счисления принятое в  $F$ . Подробнее определение множества чисел с плавающей запятой см. [14] или [18]). Если подставить в систему  $x_1 = 2.4$  и  $x_2 = 0$ , то правые части будут равны 12 и 16.8, т.е. ошибкам в правой части 0 и 0.2 соответствуют ошибки решения 1.4 и  $-1$ . Таким образом, в рассматриваемом игрушечном множестве  $F$   $x_1 = 2.4$  и  $x_2 = 0$  являются приближенным решением системы в  $F$ . Дело в том, что прямые, соответствующие этим уравнениям на некотором отрезке почти сливаются. Системы такого вида называются плохо обусловленными. В этом случае найти численно единственное решение с заданной точностью трудно.

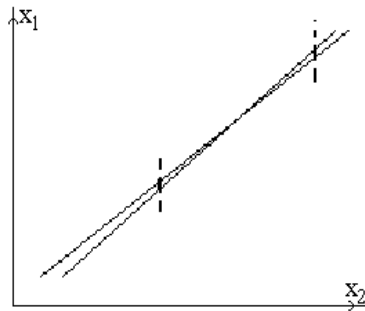


Рис. 2: Изображение плохо обусловленной системы линейных алгебраических систем второго порядка

Следует отметить, что для многих задач (например, экономических) при вычислении решения важно, чтобы при подстановке его в систему правые части мало отличались бы от их точных значений (в этом случае говорят, что мала невязка). То есть годится любое приближенное решение из отмеченной на рисунке полосы.

Чем больше порядок системы, тем труднее выявлять и, тем более, интерпретировать причины большой потери точности при решении СЛАУ, вызванной малыми ошибками входных данных (причины плохой обусловленности системы).

Например, представьте себе в трехмерном пространстве три параллельные прямые, проходящие через вершины треугольника и перпендикулярные его плоскости. Три плоскости, содержащие пары этих прямых, не параллельны, но и не имеют ни одной общей точки. Если же одну из плоскостей слегка наклонить, то три плоскости пересекутся в одной точке. Соответствующая система уравнений будет плохо обусловленной. Плохая обусловленность проявляется в том, что в зависимости от порядка действий можно вычислить решение и с различной точностью. Если, решая задачу, сначала найти линию пересечения плоскостей, перпендикулярных плоскости треугольника, а затем точку пересечения найденной прямой с третьей плоскостью, то решение будет найдено с заданной точностью. Если же изменить порядок действий, т. е. найти сначала линию пересечения слегка на-



клоненной и перпендикулярной плоскостей, то получается плохо обусловленная задача, так как найденная линия пересечения почти параллельна третьей плоскости.

Очевидно, что при любых расчетах требуется получить решение с заданной точностью. Для оценок погрешности методов требуется ввести определение норм матриц и векторов.

## 2.1 Нормы матриц и векторов

Пусть вектор  $\bar{x} \in R^m$ , а квадратная матрица  $A$  имеет размер  $m \times m$ .

*Определение.* Нормой вектора  $\bar{x}$  называется поставленное ему в соответствие неотрицательное число  $\|\bar{x}\|$  такое, что

1.  $\|\bar{x}\| > 0$ , при  $\bar{x} \neq 0$ ,  $\|\bar{x}\| = 0 \Leftrightarrow \bar{x} = 0$ ,
2.  $\|c\bar{x}\| = |c| \|\bar{x}\|$ , если  $c - \text{const}$ ,
3.  $\|\bar{x} + \bar{y}\| \leq \|\bar{x}\| + \|\bar{y}\|$ ,  $\bar{x}, \bar{y} \in R^m$ .

При численных расчетах наиболее часто употребляются следующие нормы векторов

1.  $\|\bar{x}\|_1 = \sum_{i=1}^m |x_i|$ ,
2.  $\|\bar{x}\|_2 = \sqrt{\sum_{i=1}^m |x_i|^2}$  – евклидова норма,
3.  $\|\bar{x}\|_\infty = \max_{1 \leq i \leq m} |x_i|$  – чебышевская норма.

Эти нормы связаны друг с другом следующими неравенствами

$$\begin{aligned} \|\bar{x}\|_\infty \leq \|\bar{x}\|_1 \leq m \|\bar{x}\|_\infty, \quad \|\bar{x}\|_\infty \leq \|\bar{x}\|_2 \leq \sqrt{m} \|\bar{x}\|_\infty, \\ \|\bar{x}\|_1 / \sqrt{m} \leq \|\bar{x}\|_2 \leq \|\bar{x}\|_1. \end{aligned} \tag{2.2}$$

*Определение.* Нормой квадратной матрицы  $A$  называется поставленное ей в соответствие неотрицательное число  $\|A\|$  такое, что

1.  $\|A\| > 0$ , при  $A \neq 0$ ,  $\|A\| = 0 \Leftrightarrow A = 0$  (у нуль-матрицы все элементы нулевые),

2.  $\|cA\| = |c| \|A\|$ , если  $c - \text{const}$ ,
3.  $\|A + B\| \leq \|A\| + \|B\|$ ,  $B$  – квадратная матрица размера  $m \times m$ ,
4.  $\|A B\| \leq \|A\| \|B\|$ .

Наиболее часто употребляются следующие нормы матриц

1.  $\|A\|_1 = \max_{1 \leq j \leq m} \sum_{i=1}^m |a_{ij}|$  – максимальная сумма модулей элементов столбцов,

2.  $\|A\|_2 = \sqrt{\max_{1 \leq i \leq m} (\lambda_{AA^*})_i}$ , здесь  $(\lambda_{AA^*})_i$  –  $i$ -ое собственное число матрицы  $AA^*$ ,

3.  $\|A\|_\infty = \max_{1 \leq i \leq m} \sum_{j=1}^m |a_{ij}|$  – максимальная сумма модулей элементов строк.

*Определение.* Нормы матриц и векторов, удовлетворяющие неравенству  $\|A\bar{x}\| \leq \|A\| \|\bar{x}\|$ , называются согласованными.

*Определение.* Если  $\|A\| = \sup_{\|\bar{x}\| \neq 0} \|A\bar{x}\| / \|\bar{x}\|$ , то норма  $\|A\|$  называется подчиненной норме  $\|\bar{x}\|$  вектора  $\bar{x}$ .

Все введенные нормы матриц подчинены соответствующим нормам векторов, т.е.  $\|A\|_\alpha = \sup_{\|\bar{x}\|_\alpha \neq 0} \|A\bar{x}\|_\alpha / \|\bar{x}\|_\alpha$ , где  $\alpha = 1, 2, \infty$ .

Отметим, что подчиненная норма матрицы не больше согласованной с той же нормой вектора. Действительно, пусть  $\|A\|_\alpha$  подчинена норме вектора  $\|\bar{x}\|_\alpha$ , а норма  $\|A\|_\beta$  согласована с той же нормой вектора  $\|\bar{x}\|_\alpha = \|\bar{x}\|_\beta$ , то есть

$$\|A\|_\alpha = \sup_{\|\bar{x}\| \neq 0} \|A\bar{x}\|_\alpha / \|\bar{x}\|_\alpha, \quad \|A\bar{x}\|_\beta \leq \|A\|_\beta \|\bar{x}\|_\beta.$$

Пусть  $\sup_{\|\bar{x}\| \neq 0} \|A\bar{x}\|_\alpha / \|\bar{x}\|_\alpha$  достигается на векторе  $\bar{x}_0$ , и  $\|\bar{x}_0\|_\alpha = \|\bar{x}_0\|_\beta = 1$ , тогда для вектора  $\bar{x}_0$  выполняется  $\|A\|_\alpha = \|A\bar{x}_0\|_\alpha = \|A\bar{x}_0\|_\beta \leq \|A\|_\beta \|\bar{x}_0\|_\beta = \|A\|_\beta$ .

В приложениях часто используются еще две нормы

1.  $M(A) = m \max_{1 \leq i, j \leq m} |a_{ij}|$ .

Эта норма согласована со всеми введенными нормами векторов, но не подчинена ни одной из них. Отсутствие подчиненности следует из соотно-

шений для единичной матрицы  $E$

$$\|E\| = \sup_{\|\bar{x}\| \neq 0} \|E\bar{x}\|/\|\bar{x}\| = 1, \quad M(E) = m.$$

2.  $S(A) = \sqrt{\sum_{i,j=1}^m |a_{ij}|^2}$  — евклидова норма (норма Фробениуса, норма Шура).

Эта норма также согласована со всеми введенными нормами векторов, но не подчинена ни одной из них, так как для единичной матрицы  $E$  выполняются равенства

$$\|E\| = \sup_{\|\bar{x}\| \neq 0} \|E\bar{x}\|/\|\bar{x}\| = 1, \quad S(E) = \sqrt{m}.$$

## 2.2 Погрешность приближенного решения системы и обусловленность матриц

Пусть требуется решить систему

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

При любом расчете одним из основных является вопрос о погрешности. Неизбежные вычислительные ошибки, связанные с округлением, могут сильно исказить вычисленное решение даже в случае, когда оно единственно и используется точный метод, который при отсутствии ошибок округления за конечное число операций позволяет найти точное решение.

Это происходит, если матрица  $A$  плохо обусловлена, то есть малые погрешности (возмущения) входных данных (значений элементов матрицы или компонент вектора правых частей) могут вызвать большие ошибки в разыскиваемом решении. Априори выяснять, какая ситуация реализуется в решаемом примере, часто трудно, а иногда и невозможно.

Ошибки округления появляются, как правило, уже при вводе данных в память вычислительной машины (см. [14] или [18]), то есть вместо системы

(2.1) разыскивается решение системы

$$(A + \Delta A)\bar{x} = \bar{b} + \overline{\Delta b}, \quad (2.3)$$

где  $\Delta A$  — матрица, элементы которой являются ошибками соответствующих элементов матрицы  $A$ ,  $\overline{\Delta b}$  — вектор, компоненты которого являются ошибками соответствующих компонент вектора  $\bar{b}$ . Пусть известны оценки  $\|\Delta A\|$  и  $\|\overline{\Delta b}\|$ , и матрицы  $A$  и  $A + \Delta A$  не вырождены, то есть существуют  $A^{-1}$  и  $(A + \Delta A)^{-1}$ .

Пусть  $\bar{x}$  — точное решение системы (2.1),  $\bar{x}^*$  — точное решение системы (2.3). Оценим погрешность  $\overline{\Delta x} = \bar{x} - \bar{x}^*$ , вызываемую лишь погрешностями входных данных  $\|\Delta A\|$  и  $\|\overline{\Delta b}\|$  (неизбежные ошибки округления при вычислении решения не учитываются). В силу линейности задачи можно рассмотреть отдельно влияние ошибок правой части и ошибок элементов матрицы.

1. Пусть  $\|\Delta A\| = 0$ , а  $\|\overline{\Delta b}\| \neq 0$ . Из (2.1) и (2.3) следует, что

$$\begin{aligned} \bar{x} &= A^{-1} \bar{b}, \\ \bar{x}^* &= A^{-1}(\bar{b} + \overline{\Delta b}). \end{aligned}$$

Отсюда получается цепочка равенств

$$\bar{x}^* - \bar{x} = A^{-1}(\bar{b} + \overline{\Delta b}) - A^{-1}\bar{b} = A^{-1}\overline{\Delta b},$$

$$\|\overline{\Delta x}\| = \|\bar{x}^* - \bar{x}\| = \|A^{-1}\overline{\Delta b}\| \leq \|A^{-1}\| \|\overline{\Delta b}\| = \|A^{-1}\| \|\bar{b}\| \frac{\|\overline{\Delta b}\|}{\|\bar{b}\|},$$

$$\|\bar{b}\| = \|A\bar{x}\| \leq \|A\| \|\bar{x}\|,$$

$$\|\overline{\Delta x}\| \leq \|A^{-1}\| \|A\| \|\bar{x}\| \frac{\|\overline{\Delta b}\|}{\|\bar{b}\|}.$$

Таким образом получается оценка относительной погрешности решения СЛАУ, вызванная относительной погрешностью правой части

$$\frac{\|\overline{\Delta x}\|}{\|\bar{x}\|} \leq \|A^{-1}\| \|A\| \frac{\|\overline{\Delta b}\|}{\|\bar{b}\|}. \quad (2.4)$$

2. Пусть  $\|\overline{\Delta b}\| = 0$ , а  $\|\Delta A\| \neq 0$ . Из (2.1) и (2.3) следует

$$\begin{aligned}\bar{x} &= A^{-1} \bar{b}, \\ \bar{x}^* &= (A + \Delta A)^{-1} \bar{b}.\end{aligned}$$

Отсюда получается

$$\bar{x}^* - \bar{x} = ((A + \Delta A)^{-1} - A^{-1}) \bar{b}.$$

Преобразуем выражение

$$\begin{aligned}(A + \Delta A)^{-1} - A^{-1} &= A^{-1} A (A + \Delta A)^{-1} - A^{-1} (A + \Delta A) (A + \Delta A)^{-1} = \\ &= A^{-1} (A - (A + \Delta A)) (A + \Delta A)^{-1} = -A^{-1} \Delta A (A + \Delta A)^{-1}.\end{aligned}$$

Теперь

$$\begin{aligned}\overline{\Delta x} &= -A^{-1} \Delta A (A + \Delta A)^{-1} \bar{b} = -A^{-1} \Delta A \bar{x}^*, \\ \|\overline{\Delta x}\| &= \|A^{-1} \Delta A \bar{x}^*\| \leq \|A^{-1}\| \|A\| \frac{\|\Delta A\|}{\|A\|} \|\bar{x}^*\|.\end{aligned}$$

Для оценки относительной погрешности решения СЛАУ, вызванной относительной погрешностью матрицы системы, получается выражение

$$\frac{\|\overline{\Delta x}\|}{\|\bar{x}^*\|} \leq \|A^{-1}\| \|A\| \frac{\|\Delta A\|}{\|A\|}. \quad (2.5)$$

Вспомним, что относительные погрешности  $\delta(x) = \|\overline{\Delta x}\|/\|\bar{x}\|$  и  $\delta(b) = \|\overline{\Delta b}\|/\|\bar{b}\|$  отличаются друг от друга на бесконечно малые порядка  $O(\|\Delta x\|^2)$  и считаются неразличимыми. Из (2.4) и (2.5) следует, что

$$\delta(x) \leq \text{cond}(A) \|\Delta A\|/\|A\| \quad \text{или} \quad \delta(x) \leq \text{cond}(A) \|\overline{\Delta b}\|/\|\bar{b}\|,$$

где  $\text{cond}(A) = \|A^{-1}\| \|A\|$  — число обусловленности матрицы  $A$ . Если погрешности  $\|\overline{\Delta b}\| \rightarrow 0$  и  $\|\Delta A\| \rightarrow 0$ , то и погрешность решения  $\|\overline{\Delta x}\| \rightarrow 0$ . Заметим, что, если норма матрицы подчинена норме вектора, то полученные оценки нельзя улучшить, так как существуют такие  $\|\overline{\Delta b}\| \rightarrow 0$  и  $\|\Delta A\| \rightarrow 0$ , что оценки погрешности (2.4) и(или) (2.5) превращаются в равенства.

Если число обусловленности  $\text{cond}(A) = \|A^{-1}\| \|A\|$  мало, то система (2.1) хорошо обусловлена, если же оно велико, то плохо обусловлена. Какое значение числа обусловленности считается угрожающе большим зависит от точности вычислительной машины (количества цифр  $t$  для записи мантисы числа). Так, если вычисления проводятся на компьютере с точностью  $t$  (погрешность входных данных  $O(2^{-t})$ ), то при значении числа обусловленности  $O(2^{t/2})$  нельзя гарантировать лучшую точность, чем  $O(2^{-t/2})$ .

Числа обусловленности  $\text{cond}(A) = \|A^{-1}\| \|A\|$  обладают следующими свойствами:

1.  $\text{cond}(A) \geq 1$ . Действительно,  $A A^{-1} = E \Rightarrow 1 = \|E\| = \|A A^{-1}\| \geq \|A\| \|A^{-1}\| = \text{cond}(A)$ .

2.  $\text{cond}(c A) = \text{cond}(A)$ ,  $c - \text{const}$ .

3.  $\text{cond}(A B) \leq \text{cond}(A) \text{cond}(B)$ .

4.  $\text{cond}(A^{-1}) = \text{cond}(A)$ .

5.  $\text{cond}(D) = \frac{\max_{1 \leq i \leq m} d_{ii}}{\min_{1 \leq i \leq m} d_{ii}}$ , если  $D$  — диагональная матрица ( $D = \text{diag}(d_{ii})$ ).

Очевидно, число обусловленности зависит от того, какая используется норма. Если у числа обусловленности не написаны индексы выбранной нормы, то по умолчанию  $\text{cond}(A) = \|A\|_2 \|A^{-1}\|_2$ .

Вычислять  $\text{cond}(A)$  для произвольной матрицы трудная, а иногда невыполнимая задача, поэтому полученные оценки погрешности решения (2.1) не имеют практического значения, но помогают понять, как происходит потеря точности.

Поясним смысл числа обусловленности на примере симметричной положительно определенной матрицы  $A = A^* > 0$ . У такой матрицы жорданова форма диагональна. Собственные числа положительны  $\lambda_i > 0$ ,  $i = 1, \dots, m$ , а соответствующие им собственные векторы  $\bar{\varphi}_i$ ,  $i = 1, \dots, m$

образуют ортогональный базис в пространстве  $R^m$ . Выпишем норму симметричной положительно определенной матрицы

$$\|A\|_2 = \sqrt{\max_{1 \leq i \leq m} (\lambda_{AA^*})_i} = \sqrt{\max_{1 \leq i \leq m} \lambda_i^2} = \max_{1 \leq i \leq m} \lambda_i,$$

и ее число обусловленности

$$\text{cond}(A) = \frac{\max_{1 \leq i \leq m} \lambda_i}{\min_{1 \leq i \leq m} \lambda_i}.$$

Разложим искомое решение системы (2.1) и вектор  $\bar{b}$  по базису

$$\bar{x} = \sum_{i=1}^m C_i \bar{\varphi}_i, \quad \bar{b} = \sum_{i=1}^m \beta_i \bar{\varphi}_i.$$

Подставим эти разложения в (2.1)

$$\begin{aligned} A \left( \sum_{i=1}^m C_i \bar{\varphi}_i \right) &= \sum_{i=1}^m \beta_i \bar{\varphi}_i \Rightarrow \sum_{i=1}^m C_i A(\bar{\varphi}_i) = \sum_{i=1}^m \beta_i \bar{\varphi}_i \Rightarrow \\ \sum_{i=1}^m C_i \lambda_i \bar{\varphi}_i &= \sum_{i=1}^m \beta_i \bar{\varphi}_i \Rightarrow C_i \lambda_i = \beta_i \Rightarrow C_i = \frac{\beta_i}{\lambda_i}. \end{aligned}$$

Если значение  $\beta_i$  задано с погрешностью  $\varepsilon_i$ , то в разложении решения  $\bar{x}$  у коэффициента  $C_i = \frac{\beta_i}{\lambda_i} + \frac{\varepsilon_i}{\lambda_i}$  появляется погрешность  $\Delta C_i = \frac{\varepsilon_i}{\lambda_i}$ .

Пусть погрешности  $|\varepsilon_i| \leq \varepsilon$ ,  $i = 1, \dots, m$ ,  $\lambda_k$  — максимальное, а  $\lambda_l$  — минимальное собственные значения матрицы  $A$ . Тогда минимальная ошибка  $\Delta C_k$  решения  $\bar{x}$  соответствует направлению вектора  $\bar{\varphi}_k$ , а максимальная ошибка  $\Delta C_l$  — направлению вектора  $\bar{\varphi}_l$ .

Большая величина числа обусловленности означает, что ошибки округления вдоль направлений  $\bar{\varphi}_k$  и  $\bar{\varphi}_l$  сильно искажают решение. Наглядно это можно представить так,

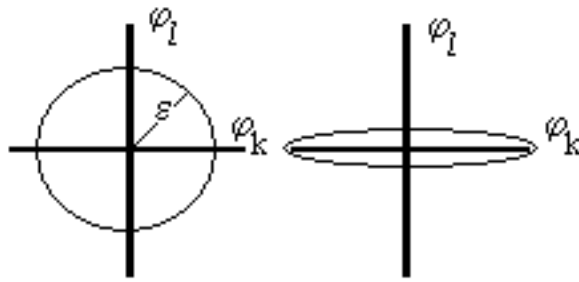


Рис. 3: Слева изображена область в плоскости  $(\bar{\varphi}_k, \bar{\varphi}_l)$ , в которой лежат ошибки входных данных, справа изображено как преобразуется эта область под действием обратной матрицы  $A^{-1}$  для плохо обусловленной СЛАУ второго порядка.

если в плоскости  $(\bar{\varphi}_k, \bar{\varphi}_l)$  ошибки входных данных лежат внутри круга радиуса  $\varepsilon$ , то действие обратной матрицы  $A^{-1}$  (вычисление решения системы  $A\bar{x} = \bar{b}$ ) превращает этот круг в вытянутый эллипс, который тем длиннее в направлении вектора  $\bar{\varphi}_l$ , чем больше число обусловленности. Это означает, что ошибка решения системы может быть значительно больше  $\varepsilon$ . При использовании точных методов ситуация улучшается, если вектор правых частей ортогонален „плохим“ собственным векторам.

Иногда разумно применять метод регуляризации. Зададимся некоторым малым  $\delta$  (трудный вопрос: каким?) и вместо системы

$$A \bar{x} = \bar{b},$$

с решением  $\bar{x}$ , в разложении которого коэффициенты  $C_i = \frac{\beta_i}{\lambda_i}$ , будем решать систему

$$(A + \delta E) \bar{x} = \bar{b},$$

решению  $\bar{x}_\delta$  которой соответствуют коэффициенты  $C_{\delta i} = \frac{\beta_i}{\lambda_i + \delta}$ .

Сравним решения обеих систем, считая ортогональную систему векторов  $\bar{\varphi}_i$  нормированной ( $\|\bar{\varphi}_i\| = 1$ ). Сначала вычислим

$$\begin{aligned} \|\bar{x}\|^2 &= (\bar{x}, \bar{x}) = \sum_{i=1}^m C_i^2 = \sum_{i=1}^m \left( \frac{\beta_i}{\lambda_i} \right)^2 \\ \|\bar{x}_\delta\|^2 &= (\bar{x}_\delta, \bar{x}_\delta) = \sum_{i=1}^m C_{\delta i}^2 = \sum_{i=1}^m \left( \frac{\beta_i}{\lambda_i + \delta} \right)^2 \end{aligned}$$



Если  $\lambda_i + \delta \neq 0$ , то  $\lim_{\delta \rightarrow 0} \|\bar{x}_\delta\| = \|\bar{x}\|$ . Так как  $C_i - C_{\delta i} = \frac{\delta \beta_i}{\lambda_i(\lambda_i + \delta)}$ , то для больших значений  $\lambda_i$  введение  $\delta$  не оказывает существенного влияния, а для значений  $\lambda_i \ll \delta$

$$|C_{\varepsilon i}| = \frac{|\beta_i|}{|\lambda_i + \delta|} \ll |C_i| = \frac{|\beta_i|}{|\lambda_i|}$$

то есть вклад слагаемых, соответствующих малым  $\lambda_i$ , в погрешность решения станет меньше.

## 2.3 Матрицы специального вида.

### Треугольные матрицы.

*Определение.*

Матрица называется правой или верхней (левой или нижней) треугольной, если для ее элементов  $a_{ij}$  выполняются равенства  $a_{ij} = 0$  при  $i > j$  ( $i < j$ ). Эти матрицы имеют вид

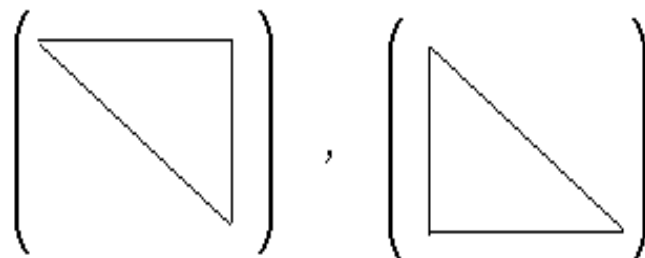


Рис. 4: Слева схематически изображена правая (верхняя) треугольная матрица, справа — левая (нижняя) треугольная матрица.

Квадратные треугольные матрицы обладают следующими свойствами

1. сумма и произведение одноименных треугольных матриц есть треугольная матрица того же наименования;
2. определитель треугольной матрицы равен произведению диагональных элементов;
3. собственные числа треугольной матрицы совпадают с её диагональными элементами;
4. невырожденная (см. приложения стр. 69) треугольная матрица легко обращается и обратная к ней тоже треугольная матрица того же наименования.

### Элементарные ортогональные матрицы

*Определение.* Вещественная матрица  $U$  называется ортогональной, если транспонированная матрица  $U^T$  совпадает с обратной  $U^{-1}$ , т.е.  $U^T U = U U^T = E$ .

Модуль определителя матрицы  $|\det(U)| = 1$ , т.к.  $1 = \det(E) = \det(UU^T) = \det(U) \det(U^T) = \det^2(U)$ .

Все собственные числа ортогональной матрица по модулю равны 1. Действительно, если  $\vec{\varphi}$  — собственный вектор, соответствующий собственному числу  $\lambda$ , то есть  $U\vec{\varphi} = \lambda\vec{\varphi}$ , тогда

$$\left. \begin{aligned} (U\vec{\varphi}, U\vec{\varphi}) &= (\lambda\vec{\varphi}, \lambda\vec{\varphi}) = |\lambda|^2(\vec{\varphi}, \vec{\varphi}) \\ (U\vec{\varphi}, U\vec{\varphi}) &= (\vec{\varphi}, U^T U\vec{\varphi}) = (\vec{\varphi}, \vec{\varphi}) \end{aligned} \right\} \Rightarrow |\lambda|^2 = 1$$

Заметим, что действие матрицы  $U$  на вектор не изменяет его длину

$$\|U\vec{\varphi}\| = \sqrt{(U\vec{\varphi}, U\vec{\varphi})} = \sqrt{(\vec{\varphi}, \vec{\varphi})} = \|\vec{\varphi}\|$$

Произведение ортогональных матриц также ортогональная матрица, поскольку  $(UV)(UV)^T = UVV^T U^T = UU^T = E$ .

### Примеры элементарных ортогональных матриц.

1. *Элементарная матрица перестановок.* Матрица перестановок  $P_{ij}$  получается, если у единичной матрицы переставить  $i$ -ую и  $j$ -ую строки (или  $i$ -ый и  $j$ -ый столбцы), например, матрица  $P_{24}$  размера  $5 \times 5$  имеет вид

$$P_{24} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

При умножении матрицы  $A$  размера  $5 \times 5$  на  $P_{ij}$  слева у матрицы  $A$  переставляются  $i$ -ая и  $j$ -ая строки, а при умножении справа у  $A$  переставляются  $i$ -ый и  $j$ -ый столбцы. Проверьте, что матрица  $P_{ij}$  — ортогональная.

2. *Матрицы вращения.* Вещественная матрица  $Q_{ij}(\varphi)$  элементарного вращения отличается от единичной матрицы четырьмя элементами, расположенными на пересечении строк и столбцов с номерами  $i$  и  $j$ , например,

матрица  $Q_{24}(\varphi)$  размера  $5 \times 5$  имеет вид

$$Q_{24}(\varphi) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \cos \varphi & 0 & -\sin \varphi & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & \sin \varphi & 0 & \cos \varphi & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Такие матрицы называются также матрицами простого поворота, матрицами Гивенса. Проверьте, что матрица  $Q_{ij}$  ортогональная.

3. *Матрицы отражения.* Матрица отражения — это такая ортогональная матрица  $P$  (размера  $m \times m$ ), что для произвольного ненулевого  $m$ -мерного вектор-столбца  $\bar{v} = (v_1, v_2, \dots, v_m)^T$  справедливо равенство

$$P\bar{v} = \sigma \|\bar{v}\| \bar{e}, \quad (2.6)$$

где  $\bar{e} = (1, 0, \dots, 0)^T$ , а константа  $\sigma = \begin{cases} -1, & \text{если } v_1 \geq 0, \\ 1, & \text{если } v_1 < 0. \end{cases}$

Покажем, что, если матрица  $P$  имеет вид

$$P = E - 2 \frac{\bar{u} \bar{u}^T}{\bar{u}^T \bar{u}}, \quad (2.7)$$

где  $\bar{u} = \bar{v} - \sigma \|\bar{v}\| \bar{e}$ , то выполняется равенство (2.6), т.е.  $P$  является матрицей отражения (или матрицей Хаусхолдера).

Сначала вычислим

$$\begin{aligned} \bar{u}^T \bar{u} &= (\bar{v}^T - \sigma \|\bar{v}\| \bar{e}^T) \cdot (\bar{v} - \sigma \|\bar{v}\| \bar{e}) = \\ &= \|\bar{v}\|^2 - 2\sigma v_1 \|\bar{v}\| + \|\bar{v}\|^2 = -2\|\bar{v}\|(\sigma v_1 - \sigma^2 \|\bar{v}\|) = -2\sigma \|\bar{v}\| u_1. \end{aligned}$$

Теперь запишем вектор  $\bar{v}$  в виде  $\bar{v} = \bar{u} + \sigma \|\bar{v}\| \bar{e}$ , и вычислим

$$\begin{aligned} P\bar{v} &= \bar{u} + \sigma \|\bar{v}\| \bar{e} - 2 \left( \frac{\bar{u} \bar{u}^T \bar{u}}{\bar{u}^T \bar{u}} \right) - 2\sigma \|\bar{v}\| \left( \frac{\bar{u} \bar{u}^T \bar{e}}{\bar{u}^T \bar{u}} \right) = \\ &= \bar{u} + \sigma \|\bar{v}\| \bar{e} - 2\bar{u} + 2\sigma \|\bar{v}\| \frac{\bar{u} u_1}{2\sigma \|\bar{v}\| u_1} = \sigma \|\bar{v}\| \bar{e}. \end{aligned}$$

Выбор константы  $\sigma$  не позволяет первой компоненте  $u_1$  вектора  $\bar{u}$  обратиться в нуль, более того  $|u_1| \geq \|\bar{v}\|$ , что уменьшает влияние ошибок округления.

Матрица  $P$  симметрична и ортогональна, действительно

$$P^T = E - 2 \frac{(\bar{u}^T)^T \bar{u}^T}{\bar{u}^T \bar{u}} = E - 2 \frac{\bar{u} \bar{u}^T}{\bar{u}^T \bar{u}},$$

$$P P^T = E - 4 \frac{\bar{u} \bar{u}^T}{\bar{u}^T \bar{u}} + 4 \frac{\bar{u} \bar{u}^T \bar{u} \bar{u}^T}{(\bar{u}^T \bar{u})^2} = E - 4 \frac{\bar{u} \bar{u}^T}{\bar{u}^T \bar{u}} + 4 \frac{\bar{u} (\bar{u}^T \bar{u}) \bar{u}^T}{(\bar{u}^T \bar{u})^2} = E.$$

С геометрической точки зрения преобразование Хаусхолдера (умножение матрицы отражения  $P$  на вектор) есть отражение произвольного вектора  $\bar{y} \in R^m$  относительно  $m-1$ -мерного подпространства  $S$  ортогонального вектору  $\bar{u}$ . Такой вывод следует из следующих соотношений:

1. если  $\bar{s} \in S$ , то  $P\bar{s} = \bar{s} - 2 \frac{\bar{u} \bar{u}^T}{\bar{u}^T \bar{u}} \bar{s} = \bar{s}$ , так как  $\bar{u}^T \bar{s} = 0$ ,

2.  $P\bar{u} = \bar{u} - 2 \frac{\bar{u} \bar{u}^T}{\bar{u}^T \bar{u}} \bar{u} = -\bar{u}$ ,

3. произвольный вектор  $\bar{y} \in R^m$  можно представить в виде  $\bar{y} = \alpha \bar{u} + \beta \bar{s}$  ( $\alpha$  и  $\beta$  – некоторые константы), тогда  $P\bar{y} = \alpha \bar{u} - \beta \bar{s}$ .

### Почти треугольные матрицы.

*Определение.* Матрица называется правой (левой) почти треугольной или правой (левой) матрицей Хессенберга, если для ее элементов  $A_{ij}$  выполняются соотношения  $a_{ij} = 0$  при  $i > j + 1$  ( $i < j - 1$ ).

Такие матрицы имеют вид

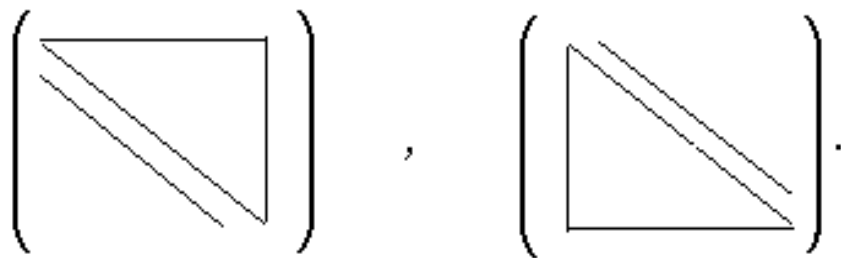


Рис. 5: Слева схематически изображена правая (верхняя) почти треугольная матрица, справа — левая (нижняя) почти треугольная матрица

Сумма и произведение одноименных почти треугольных квадратных матриц есть почти треугольная матрица того же наименования. Любая матрица ортогонально подобна (определение подобия см. 71) почти треугольной матрице.

## 2.4 Методы решения СЛАУ

Методы решения СЛАУ можно условно разделить на две группы методов:

1. Точные (прямые) методы, позволяющие при отсутствии ошибок округления получить точное решение за конечное число операций. Такие методы обычно применяются при решении „небольших“ систем. Системы какого порядка можно считать небольшими зависит от доступной вычислительной техники.

2. Итерационные методы, с помощью которых решение получается как предел последовательных приближений к решению. Эти методы обычно применяются при решении „больших“ систем.

## 2.5 Вопросы и упражнения для самопроверки

1. Сформулируйте определение нормы вектора и нормы матрицы. Докажите, что  $\max_{1 \leq i, j \leq m} |a_{ij}|$  не является нормой матрицы.

2. Проверьте, что  $M(A)$  удовлетворяет определению нормы и согласована с нормами вектора  $\|\bar{x}\|_\alpha$  для  $\alpha = 1, 2, 3$ .

3. Проверьте, что  $S(A)$  удовлетворяет определению нормы и согласована с нормами вектора  $\|\bar{x}\|_\alpha$  для  $\alpha = 1, 2, 3$ .

4. Докажите, что нормы векторов удовлетворяют неравенствам (2.2).

$$\|\bar{x}\|_\infty \leq \|\bar{x}\|_1 \leq m \|\bar{x}\|_\infty,$$

$$\|\bar{x}\|_\infty \leq \|\bar{x}\|_2 \leq \sqrt{m} \|\bar{x}\|_\infty,$$

$$\|\bar{x}\|_1 / \sqrt{m} \leq \|\bar{x}\|_2 \leq \|\bar{x}\|_1.$$

5. Изобразите на плоскости множество  $\bar{x} : \|\bar{x}\|_\alpha = 1$  для  $\alpha = 1, 2, 3$ .
6. Докажите подчиненность норм матриц  $\|A\|_\alpha$  норме вектора  $\|\bar{x}\|_\alpha$  для  $\alpha = 1, 2, 3$ .
7. Перечислите свойства чисел обусловленности и докажите их.
8. Можно ли утверждать, что если определитель матрицы мал, то матрица плохо обусловлена? Приведите пример.
9. Какой метод решения СЛАУ называется точным?
10. Почему при решении СЛАУ точными методами на вычислительной технике появляется погрешность решения?
11. Что характеризует число обусловленности при приближенном решении СЛАУ?
12. Как зависит обусловленность системы от точности вычислительной техники?
13. Докажите, что для норм матриц выполняются неравенства

$$\frac{M(A)}{m} \leq \|A\|_\infty \leq M(A)$$

$$\frac{S(A)}{\sqrt{m}} \leq \|A\|_2 \leq S(A)$$

14. Вычислите число обусловленности матрицы  $A = \begin{pmatrix} 5 & 7 \\ 7 & 10 \end{pmatrix}$

- 15 Рассмотрите систему

$$x_1 + 0.99 x_2 = 1.99,$$

$$0.99 x_1 + 0.98 x_2 = 1.97,$$

у которой точным решением является  $x_1 = 1$  и  $x_2 = 1$ . Вычислите каким ошибкам правой части отвечает решение  $x_1 = 3$  и  $x_2 = -1.0203$ .

### 3 Точные методы решения СЛАУ

Большинство точных (прямых) методов решения СЛАУ основано на преобразовании системы (2.1) к эквивалентной системе с треугольной матрицей.

Так, если матрица  $A = L U$ , где  $L$  – левая (left, нижняя),  $U$  – верхняя (upper, правая) треугольные матрицы, то исходную СЛАУ (2.1) можно записать в виде

$$L U \bar{x} = \bar{b}.$$

Решение такой системы сводится к решению двух СЛАУ

$$L \bar{y} = \bar{b} \quad (\text{прямой ход}),$$

$$U \bar{x} = \bar{y} \quad (\text{обратный ход}).$$

На таком  $LU$  - разложении основан хорошо известный метод Гаусса. Каждая из этих систем решается проще чем (2.1), так как они треугольные. Действительно, запишем систему  $L \bar{y} = \bar{b}$  в виде

$$\begin{aligned} l_{11}y_1 &= b_1, \\ l_{21}y_1 + l_{22}y_2 &= b_2, \\ \dots & \dots \dots \dots \dots \dots \dots, \\ l_{m-1,1}y_1 + l_{m-1,2}y_2 + \dots + \dots + l_{m-1,m-1}y_{m-1} &= b_{m-1}, \\ l_{m1}y_1 + l_{m2}y_2 + \dots + \dots + l_{m,m-1}y_{m-1} + l_{mm}y_m &= b_m. \end{aligned}$$

Из первого уравнения вычисляется  $y_1 = b_1/l_{11}$ , а затем все остальные значения  $y_i$  по формуле

$$y_i = \frac{b_i - \sum_{k=1}^{i-1} l_{ik}y_k}{l_{ii}} \quad i = 2, \dots, m.$$



Аналогично, при обратном ходе решается система  $U\bar{x} = \bar{y}$

$$\begin{aligned} u_{11}x_1 + u_{12}x_2 + u_{13}x_3 + \dots + \dots + u_{1m}x_m &= y_1, \\ u_{21}x_2 + u_{23}x_3 + \dots + \dots + u_{2m}x_m &= y_2, \\ \dots \quad \dots &\dots \quad \dots, \\ u_{m-1,m-1}x_{m-1} + u_{mm}x_m &= y_{m-1} \\ u_{mm}x_m &= y_m. \end{aligned}$$

Из последнего уравнения вычисляется  $x_m = y_m/u_{mm}$ , затем все остальные компоненты  $x_i$  вектора  $\bar{x}$  по формуле

$$x_i = \frac{y_i - \sum_{k=i+1}^m u_{ik}x_k}{u_{ii}} \quad i = m-1, m-2, \dots, 1.$$

В ортогональных методах используется разложение на произведение ортогональной матрицы  $Q$  и правой треугольной  $R$  (right) ( $QR$ -разложение матрицы  $A$ ). В этом случае СЛАУ (2.1) записывается в виде

$$QR\bar{x} = \bar{b},$$

и решение системы (2.1) получается описанным выше обратным ходом из системы с правой треугольной матрицей

$$R\bar{x} = Q^T\bar{b}.$$

### 3.1 LU - разложение матрицы

Пусть  $A_m$  — квадратная матрица размера  $m \times m$ .

*Теорема об LU - разложении матрицы.*

Если все главные миноры матрицы  $A_m$  отличны от нуля, т.е.

$$a_{11} \neq 0, \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \neq 0, \dots, \det(A_m) \neq 0,$$

то матрица  $A_m$  представима в виде  $A_m = LU$ .

Заметим, что если матрица  $A_m$  не вырождена, то перестановкой строк и столбцов можно добиться, чтобы все ведущие (главные) миноры были отличны от нуля.

*Доказательство.* Используем метод математической индукции.

Если  $m = 1$ , то  $a_{11} = l_{11} \cdot u_{11}$ .

Пусть для  $m - 1$  справедливо  $A_{m-1} = L_{m-1} \cdot U_{m-1}$ , покажем, что  $A_m = L_m \cdot U_m$ . Запишем матрицу  $A_m$  в виде

$$A_m = \begin{pmatrix} \begin{pmatrix} a_{11} & \dots & \dots & a_{1,m-1} \\ a_{21} & a_{22} & \dots & a_{2,m-1} \\ \dots & \dots & \dots & \dots \\ a_{m-1,1} & \dots & \dots & a_{m-1,m-1} \end{pmatrix} & \begin{pmatrix} a_{1m} \\ a_{2m} \\ \dots \\ a_{m-1,m} \end{pmatrix} \\ \begin{pmatrix} a_{m-1} & \dots & \dots & a_{m-1,m} \end{pmatrix} & a_{mm} \end{pmatrix} = \begin{pmatrix} A_{m-1} & \bar{u} \\ \bar{v}^T & a_{mm} \end{pmatrix},$$

где  $\bar{u} \in R^{m-1}$  – вектор столбец, а  $\bar{v}^T \in R^{m-1}$  – вектор строка. Будем разыскивать такие векторы  $\bar{x}^T$ ,  $\bar{y}$  и числа  $l_{mm}$ ,  $u_{mm}$ , чтобы выполнялось равенство

$$A_m = \underbrace{\begin{pmatrix} L_{m-1} & 0 \\ \bar{x}^T & l_{mm} \end{pmatrix}}_L \underbrace{\begin{pmatrix} U_{m-1} & \bar{y} \\ 0 & u_{mm} \end{pmatrix}}_U = \begin{pmatrix} A_{m-1} & \bar{u} \\ \bar{v}^T & a_{mm} \end{pmatrix}.$$

После перемножения матриц  $L$  и  $U$ , получаются равенства

$$\begin{aligned} A_{m-1} &= L_{m-1}U_{m-1}, \\ L_{m-1}\bar{y} &= \bar{u}, \quad \bar{x}^T U_{m-1} = \bar{v}^T, \\ \bar{x}^T \bar{y} + l_{mm}u_{mm} &= a_{mm}. \end{aligned}$$

Первое из этих равенств верно по предположению. Из следующих двух определяются векторы  $\bar{y} = L_{m-1}^{-1}\bar{u}$  и  $\bar{x}^T = \bar{v}^T U_{m-1}^{-1}$ . Существование  $L_{m-1}^{-1}$  и

$U_{m-1}^{-1}$  следует из условий теоремы. Значения  $l_{mm}$  и  $u_{mm}$  определяются из последнего равенства неединственно.

Таким образом, LU - разложение матрицы  $A_m$  существует, но оно неоднозначно.  $\square$

*Теорема об LDU – разложении матрицы.*

Если все главные миноры матрицы  $A$  отличны от нуля, то есть

$$a_{11} \neq 0, \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \neq 0, \dots, \det(A) \neq 0,$$

то матрица  $A$  представима единственным образом в виде  $A = LDU$ , где  $L$ ,  $U$  - левая и правая треугольные матрицы с единичными диагональными элементами  $l_{ii} = 1$  и  $u_{ii} = 1$ , а  $D$  - диагональная матрица с ненулевыми элементами.

*Доказательство.* Очевидно, что разложение  $A = LU$ , можно записать в виде  $A = LDU$ , где у матрицы  $L$  диагональные элементы  $l_{ii} = 1$  и у матрицы  $U$  —  $u_{ii} = 1$ . Требуется доказать, что такое разложение единственно. Пусть это не так, т.е. существуют разложения  $A = L_1D_1U_1$  и  $A = L_2D_2U_2$ . Из равенства  $L_1D_1U_1 = L_2D_2U_2$  следует, что  $U_1U_2^{-1} = D_1^{-1}L_1^{-1}L_2D_2$ . Правая треугольная матрица  $U_1U_2^{-1}$  равна левой  $D_1^{-1}L_1^{-1}L_2D_2$ , если они обе диагональные. Так как диагональные элементы матриц  $U_1$  и  $U_2$  равны единице, то получается  $U_1U_2^{-1} = D = E \Rightarrow U_1 = U_2$ . Аналогично из равенства  $D_1^{-1}L_1^{-1}L_2D_2 = E$  следует, что  $L_1^{-1}L_2 = D_1D_2^{-1}$ . Так как диагональные элементы матриц  $L_1$  и  $L_2$  равны единице, то  $L_1^{-1}L_2 = E \Rightarrow L_1 = L_2$  и  $D_1 = D_2$ .  $\square$

### **Алгоритм построения LU - разложения матрицы**

Этот алгоритм состоит из  $m$  шагов. На каждом шаге соответствующий диагональный элемент преобразуется в 1, а элементы, стоящие под ним, преобразуются в 0.

На первом шаге матрица

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & \dots & \dots & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & \dots & \dots & \dots & a_{2m} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{m-1,1} & a_{m-1,2} & \dots & \dots & \dots & \dots & a_{m-1,m} \\ a_{m1} & a_{m2} & \dots & \dots & \dots & \dots & a_{mm} \end{pmatrix}$$

преобразуется в матрицу

$$A^{(1)} = L_1 D_1 A = \begin{pmatrix} 1 & a_{12}^{(1)} & \dots & \dots & \dots & \dots & a_{1m}^{(1)} \\ 0 & a_{22}^{(1)} & \dots & \dots & \dots & \dots & a_{2m}^{(1)} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & a_{m-1,2}^{(1)} & \dots & \dots & \dots & \dots & a_{m-1,m}^{(1)} \\ 0 & a_{m2}^{(1)} & \dots & \dots & \dots & \dots & a_{mm}^{(1)} \end{pmatrix},$$

$$\text{где } D_1 = \begin{pmatrix} 1/a_{11} & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}, L_1 = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ -a_{21} & 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ -a_{m1} & 0 & 0 & \dots & 1 \end{pmatrix}.$$

Подробно рассмотрим  $k$ -ый шаг. После  $k - 1$  шага матрица  $A^{(k-1)}$  имеет

ВИД

$$A^{(k-1)} = L_{k-1}D_{k-1} \dots L_1D_1A = \begin{pmatrix} 1 & a_{12}^{(1)} & \dots & \dots & \dots & \dots & a_{1m}^{(1)} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 1 & 0 & a_{k-1,k}^{(k-1)} & \dots & \dots & a_{k-1,m}^{(k-1)} \\ 0 & 0 & 1 & a_{kk}^{(k-1)} & \dots & \dots & a_{k,m}^{(k-1)} \\ 0 & 0 & 0 & a_{k+1,k}^{(k-1)} & \dots & \dots & a_{k+1,m}^{(k-1)} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & a_{mk}^{(k-1)} & \dots & \dots & a_{mm}^{(k-1)} \end{pmatrix}$$

Умножим матрицу  $A^{(k-1)}$  сначала на диагональную матрицу  $D_k$

$$D_k = \text{diag}(1, \dots, 1, 1/a_{kk}^{(k-1)}, \dots, 1),$$

а затем на левую треугольную матрицу  $L_k$

$$L_k = \begin{pmatrix} 1 & \dots & \dots & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 1 & 0 & \dots & 0 \\ 0 & \dots & -a_{k+1,k}^{(k-1)} & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & -a_{mk}^{(k-1)} & 0 & \dots & 1 \end{pmatrix}.$$

В результате получим

$$A^{(k)} = L_kD_kL_{k-1}D_{k-1} \dots L_1D_1A =$$

$$= \begin{pmatrix} 1 & \cdots & a_{1,k-1}^1 & a_{1,k}^1 & \cdots & a_{1,m-1}^1 & a_{1,m}^1 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 1 & a_{k-1,k}^{(k-1)} & \cdots & a_{k-1,m-1}^{(k-1)} & a_{k-1,m}^{(k-1)} \\ 0 & \cdots & 0 & 1 & \cdots & a_{k,m-1}^{(k)} & a_{k,m}^{(k)} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 0 & 0 & \cdots & a_{m-1,m-1}^{(k)} & a_{m-1,m}^{(k)} \\ 0 & \cdots & 0 & 0 & \cdots & a_{m,m-1}^{(k)} & a_{m,m}^{(k)} \end{pmatrix}$$

Выполнив  $(m-1)$  таких шагов ( $k = 1, \dots, m-1$ ), и, умножив полученную таким образом матрицу на  $D_m$ , получим правую треугольную матрицу  $U$  с единичными диагональными элементами

$$U = D_m L_{m-1} D_{m-1} \cdots L_1 D_1 A =$$

$$= \begin{pmatrix} 1 & \cdots & a_{1,k-1}^1 & a_{1,k}^1 & \cdots & a_{1,m-1}^1 & a_{1,m}^1 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 1 & a_{k-1,k}^{(k-1)} & \cdots & a_{k-1,m-1}^{(k-1)} & a_{k-1,m}^{(k-1)} \\ 0 & \cdots & 0 & 1 & \cdots & a_{k,m-1}^{(k)} & a_{k,m}^{(k)} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 0 & 0 & \cdots & 1 & a_{m-1,m}^{(m-1)} \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

Обозначим  $L^{-1} = D_m L_{m-1} D_{m-1} \cdots L_1 D_1 A$ . Матрица  $L^{-1}$  левая треугольная как произведение левых треугольных и диагональных матриц. Так как обратная левой треугольной матрице также левая треугольная, то получим разложение

$$A = LU.$$

Вычислим число обусловленности матрицы  $U$

$$\begin{aligned} \text{cond}(U) &= \text{cond}(L^{-1}A) = \text{cond}(D_m L_{m-1} D_{m-1} \dots L_1 D_1 A) \leq \\ &\leq \prod_{i=1}^m \text{cond}(D_i) \prod_{i=1}^{m-1} \text{cond}(L_i) \text{cond}(A) \end{aligned} \quad (3.1)$$

Число обусловленности  $\text{cond}(L_i) \geq 1$ , следовательно значение  $\prod_{i=1}^{m-1} \text{cond}(L_i)$  не может уменьшить  $\text{cond}(U)$ . Для диагональной матрицы  $D_i$  число обусловленности

$$\text{cond}(D_i) = \begin{cases} 1/|a_{ii}^{(i)}|, & \text{если } |a_{ii}^{(i)}| < 1 \\ |a_{ii}^{(i)}|, & \text{если } |a_{ii}^{(i)}| > 1 \end{cases}$$

Это означает, что, если на некотором шаге алгоритма разложения матрицы ведущий элемент оказывается малым, то  $\text{cond}(D_i)$  соответствующей матрицы становится большим, и может существенно увеличиться  $\text{cond}(U)$  по сравнению с  $\text{cond}(A)$ .

Заметим, что рассмотренный алгоритм построения LU-разложения позволяет понять опасности, связанные с ошибками округления. Обычно реализуется более экономичный алгоритм (см. ниже) приведения СЛАУ к эквивалентной системе с правой треугольной матрицей, который не требует построения матриц  $L_i$ ,  $D_i$  и умножения на них.

## 3.2 Метод Гаусса

Название „метод Гаусса“ является собирательным для большой группы алгоритмов решения СЛАУ. Построение LU-разложения — это то общее, что в той или иной степени присутствует в каждом из таких алгоритмов.

В очень распространенном варианте метода Гаусса система (2.1)

$$\begin{aligned}
 a_{11}x_1 + a_{12}x_2 + \dots + a_{1m}x_m &= b_1 \\
 a_{21}x_1 + a_{22}x_2 + \dots + a_{2m}x_m &= b_2 \\
 &\dots\dots\dots \\
 a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mm}x_m &= b_m
 \end{aligned}$$

приводится к треугольному виду с правой треугольной матрицей  $U$ , с помощью последовательного исключения неизвестного  $x_k$  из всех уравнений, начиная с  $k+1$ -ого,  $k=1, \dots, n-1$ . В этом случае удобно записать расширенную матрицу системы, добавив к матрице  $A$  столбец свободных членов  $a_{i,n+1} = b_i$ , тогда один из вариантов алгоритма метода Гаусса записывается с помощью вложенных циклов:

1. цикл по  $k$  от 1 до  $m - 1$ ,
2.     цикл по  $i$  от  $k + 1$  до  $m + 1$ ,
3.          $r := a_{ik}/a_{kk}$ ;
4.         цикл по  $j$  от  $k + 1$  до  $m + 1$ ,
5.              $a_{ij} := a_{ij} - r a_{kj}$ ;
6.         конец цикла по  $j$ ,
7.     конец цикла по  $i$ ,
8.   конец цикла по  $k$ ,
9.  $x_n := a_{n,n+1}/a_{n,n}$ ;
10.   цикл по  $k$  от  $n - 1$  до 1,
11.      $x_k := \left( a_{k,n+1} - \sum_{j=k+1}^n a_{kj}x_j \right) / a_{kk}$ .
12.   конец цикла по  $k$ .

Процесс исключения неизвестных и приведение системы к верхнему треугольному виду называется прямым ходом метода Гаусса, а вычисление решения полученной системы обратным ходом. Такой алгоритм метода Гаусса позволяет одновременно решать несколько СЛАУ с одинаковой матрицей  $A$  и различными вектор столбцами правых частей.



При преобразовании исходной системы к системе  $U\bar{x} = \bar{y}$  неизбежна погрешность, вызванная ошибками округления. Влияние этой погрешности на точность вычисленного решения зависит от  $\text{cond}(A)$ .

Вычислительная погрешность решения эквивалентной системы с правой треугольной  $U\bar{x} = \bar{y}$  зависит от числа обусловленности матрицы  $U$  (см. (3.1)). Это число может быть существенно больше числа обусловленности исходной матрицы  $A$ , если на некотором шаге исключения модуль ведущего элемента окажется малым. Чтобы исправить ситуацию, можно на каждом  $k$ -ом шаге перед исключением  $x_k$  выбирать максимальный по модулю элемент  $|a_{ln}|^{(k)} = \max_{k \leq i, j \leq m} |a_{ij}|^{(k)}$ , переставлять  $k$  и  $l$  строки,  $k$  и  $n$  столбцы. Тогда ведущим элементом станет максимальный по модулю из возможных. При этом надо помнить, что переменные с индексами  $k$  и  $n$  меняются местами.

Реализация метода Гаусса требует порядка  $O\left(\frac{2}{3}m^3\right)$  операций умножения и деления.

*Практическая оценка погрешности решения.* Обозначим  $\bar{x}^*$  – точное решение, а  $\bar{x}^{(1)}$  – решение вычисленное с помощью метода Гаусса. При умножении матрицы  $A$  на  $\bar{x}^{(1)}$  получится вектор  $\bar{b} + \overline{\Delta b} = A \cdot \bar{x}^{(1)}$ .

Ясно, что погрешность  $\bar{r}^{(1)} = \bar{x}^{(1)} - \bar{x}^*$  является решением СЛАУ

$$A\bar{r}^{(1)} = \overline{\Delta b}.$$

Если окажется, что  $\|\bar{r}^{(1)}\| \leq \varepsilon$ , то  $\bar{x}^{(1)}$  обычно считается приближенным решением, вычисленным с точностью  $\varepsilon$ .

Можно организовать итерационный процесс уточнения решения, пока не будет достигнута нужная точность. Вычислим  $\bar{x}^{(2)} = \bar{x}^{(1)} + \bar{r}^{(1)}$  и, умножив  $\bar{x}^{(2)}$  на матрицу  $A$ , получим вектор  $\overline{\Delta b}^{(2)} = A \cdot \bar{x}^{(2)} - \bar{b}$ . Теперь поправка  $\bar{r}^{(2)} = \bar{x}^{(2)} - \bar{x}^*$  и все следующие являются решениями систем

$$A\bar{r}^{(i)} = \overline{\Delta b}^{(i)}, \quad i = 2, 3, \dots,$$

где  $\overline{\Delta b}^{(i)} = A \cdot \bar{x}^{(i)} - \bar{b}$ ,  $\bar{x}^{(i)} = \bar{x}^{(i-1)} + \bar{r}^{(i-1)}$

Уточнение продолжается пока  $\|\bar{r}^{(i)}\|$  убывает. Если  $\|\bar{r}^{(i)}\| \leq \varepsilon$ , то соответствующее  $\bar{x}^{(i)}$  принимается за приближенное с точностью  $\varepsilon$  решение системы (2.1). Если же  $\|\bar{r}^{(i)}\|$  начинает возрастать или точность не достигается за некоторое фиксированное число итераций, то расчет прекращается, и выдается соответствующая информация.

### 3.3 Ортогональные (QR) разложения матриц

Построим ортогональное разложение матрицы  $A$

$$A = QR,$$

где  $Q$  - ортогональная матрица, а  $R$ - правая треугольная (right) матрица.

#### 3.3.1 Преобразование вращения(преобразование Гивенса)

Получим разложение матрицы  $A$  с помощью преобразования вращения (преобразование Гивенса). Рассмотрим несколько шагов этого алгоритма на примере квадратной матрицы 5-ого порядка.

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} \end{pmatrix}$$

Все элементы первого столбца невырожденной матрицы  $A$  не могут быть нулевыми. Если  $a_{11} = a_{21} = 0$ , то используя ортогональные матрицы перестановок, добьемся, чтобы  $a_{11} \neq 0$ . Умножим  $A$  слева на матрицу элемен-

тарного вращения

$$Q_{12}(\varphi_{12}) = \begin{pmatrix} \cos \varphi_{12} & -\sin \varphi_{12} & 0 & 0 & 0 \\ \sin \varphi_{12} & \cos \varphi_{12} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

У матрицы  $A_{12} = Q_{12}A$  изменяются только элементы первой и второй строки

$$A_{12} = \begin{pmatrix} a_{11} \cos \varphi_{12} - a_{21} \sin \varphi_{12} & \cdot & \cdot & \cdot & a_{15} \cos \varphi_{12} - a_{25} \sin \varphi_{12} \\ a_{11} \sin \varphi_{12} + a_{21} \cos \varphi_{12} & \cdot & \cdot & \cdot & a_{15} \cos \varphi_{12} + a_{25} \sin \varphi_{12} \\ a_{31} & a_{32} & \cdot & \cdot & \cdot \\ a_{41} & a_{42} & \cdot & \cdot & \cdot \\ a_{51} & a_{52} & \cdot & \cdot & \cdot \end{pmatrix}$$

Преобразованные элементы выражаются через элементы того же столбца, в котором находятся, и сумма их квадратов не изменяется. Действительно,

$$\begin{aligned} \left(a_{1i}^{(12)}\right)^2 + \left(a_{2i}^{(12)}\right)^2 &= (a_{1i} \cos \varphi_{12} - a_{2i} \sin \varphi_{12})^2 + (a_{1i} \sin \varphi_{12} + a_{2i} \cos \varphi_{12})^2 = \\ &= a_{1i}^2 + a_{2i}^2 \end{aligned}$$

Выберем угол  $\varphi_{12}$  так, чтобы элемент  $a_{21}^{(12)}$  матрицы  $A_{12}$  обращался в нуль, т.е.

$$a_{11} \sin \varphi_{12} + a_{21} \cos \varphi_{12} = 0.$$

Из этого равенства получаем выражения

$$\sin \varphi_{12} = -\frac{a_{21}}{\sqrt{a_{11}^2 + a_{21}^2}}, \quad \cos \varphi_{12} = \frac{a_{11}}{\sqrt{a_{11}^2 + a_{21}^2}}.$$

У матрицы  $A_{12}$  элемент первого столбца  $a_{21}^{(12)} = 0$ , а  $a_{11}^{(12)} = \sqrt{a_{11}^2 + a_{21}^2} > 0$ .

Теперь  $A_{12}$  умножим слева на матрицу

$$Q_{13}(\varphi_{13}) = \begin{pmatrix} \cos \varphi_{13} & 0 & -\sin \varphi_{13} & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ \sin \varphi_{13} & 0 & \cos \varphi_{13} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

У матрицы  $A_{13} = Q_{13}A_{12}$  изменятся элементы первой и третьей строки. Аналогично предыдущему выберем угол  $\varphi_{13}$  так, чтобы  $a_{31}^{(13)} = 0$ , тогда элемент  $a_{11}^{(13)} = \sqrt{\left(a_{11}^{(12)}\right)^2 + a_{31}^2} = \sqrt{a_{11}^2 + a_{21}^2 + a_{31}^2} > 0$ .

Повторим этот процесс еще два раза, и получим матрицу

$$A_1 = \begin{pmatrix} a_{11}^{(15)} & a_{12}^{(15)} & a_{13}^{(15)} & a_{14}^{(15)} & a_{15}^{(15)} \\ 0 & a_{22}^{(12)} & a_{23}^{(12)} & a_{24}^{(12)} & a_{25}^{(12)} \\ 0 & a_{32}^{(13)} & a_{33}^{(13)} & a_{34}^{(13)} & a_{35}^{(13)} \\ 0 & a_{42}^{(14)} & a_{43}^{(14)} & a_{44}^{(14)} & a_{45}^{(14)} \\ 0 & a_{52}^{(15)} & a_{53}^{(15)} & a_{54}^{(15)} & a_{55}^{(15)} \end{pmatrix} = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & a_{14}^{(1)} & a_{15}^{(1)} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} & a_{24}^{(1)} & a_{25}^{(1)} \\ 0 & a_{32}^{(1)} & a_{33}^{(1)} & a_{34}^{(1)} & a_{35}^{(1)} \\ 0 & a_{42}^{(1)} & a_{43}^{(1)} & a_{44}^{(1)} & a_{45}^{(1)} \\ 0 & a_{52}^{(1)} & a_{53}^{(1)} & a_{54}^{(1)} & a_{55}^{(1)} \end{pmatrix}$$

у которой все элементы первого столбца, начиная со второго равны нулю, а  $a_{11}^{(1)} = \sqrt{a_{11}^2 + \dots + a_{15}^2}$ . Матрица  $A_1$  получается из  $A$  с помощью ортогонального преобразования

$$A_1 = Q_{15} \cdot Q_{14} \cdot Q_{13} \cdot Q_{12} \cdot A = Q_1 \cdot A,$$

где  $Q_1 = Q_{15} \cdot Q_{14} \cdot Q_{13} \cdot Q_{12}$  — ортогональная матрица, так как она равна произведению ортогональных.

Аналогично строится ортогональная матрица  $Q_2 = Q_{25} \cdot Q_{24} \cdot Q_{23}$ , такая, что у матрицы  $A_2 = Q_2A_1 = Q_2Q_1A$  нулевыми становятся поддиагональные элементы в первых двух столбцах. Заметим, что при умножении матрицы  $A_1$  на матрицу  $Q_2$  элементы первого столбца и первой строки не изменяются, а диагональный элемент  $a_{22}^{(2)} = \sqrt{\left(a_{22}^{(1)}\right)^2 + \dots + \left(a_{25}^{(1)}\right)^2} > 0$ .

Аналогично строятся матрицы  $Q_3$  и  $Q_4$ , и в результате получается треугольная матрица  $A_4 = Q_4 \cdot Q_3 \cdot Q_2 \cdot Q_1 A$ , у которой все поддиагональные элементы равны нулю, а диагональные, кроме, быть может, последнего неотрицательны.

Если квадратная матрица  $A$  размера  $m \times m$ , то она аналогично преобразуется к виду

$$A_{m-1} = Q_{m-1} \cdot \dots \cdot Q_1 \cdot A = \tilde{Q} \cdot A,$$

где  $\tilde{Q}$  — ортогональная матрица, а матрица  $A_{m-1}$  — правая треугольная матрица. Таким образом, с помощью преобразования вращения получено ортогональное разложение матрицы  $A = QR$ , где  $Q = \tilde{Q}^{-1}$ ,  $R = A_{m-1}$ .

### 3.3.2 Преобразование отражения (преобразования Хаусхолдера)

Используем преобразование отражения (Хаусхолдера) с матрицей  $P$  для построения  $QR$ -разложения матрицы  $A$ . Матрица отражения  $P$  (2.7) имеет вид

$$P = E - 2 \frac{\bar{u} \bar{u}^T}{\bar{u}^T \bar{u}}$$

где  $\bar{u} = \bar{v} - \sigma \|\bar{v}\| \bar{e}$ ,  $\sigma = \begin{cases} -1, & \text{если } v_1 \geq 0, \\ 1, & \text{если } v_1 < 0. \end{cases}$ ,  $\bar{e} = (1, 0, \dots, 0)^T$ .

Произвольный отличный от нулевого  $m$ -мерный вектор-столбец  $\bar{v} = (v_1, v_2, \dots, v_m)^T$  с помощью матрицы  $P$  преобразуется в  $\sigma \|\bar{v}\| \bar{e}$ . Введение константы  $\sigma$  позволяет уменьшить влияние ошибок округления.

Выберем в качестве  $m$ - мерного вектора  $\bar{v}_1$  первый столбец матрицы  $A$

$$\bar{v}_1 = (a_{11}, a_{21}, \dots, a_{m1})^T$$

и построим матрицу  $P_1$  размера  $m \times m$ , такую, что  $P_1 \bar{v}_1 = \sigma_1 \|\bar{v}_1\| \bar{e}_1$ , тогда

$$A_1 = P_1 \cdot A = \begin{pmatrix} \sigma_1 \|\bar{v}_1\| & a_{12}^{(1)} & \cdot & \cdot & a_{1m}^{(1)} \\ 0 & a_{22}^{(1)} & \cdot & \cdot & a_{2m}^{(1)} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & a_{m2}^{(1)} & \cdot & \cdot & a_{mm}^{(1)} \end{pmatrix}$$

Теперь возьмем  $(m-1)$ -мерный вектор  $\bar{v}_2$

$$\bar{v}_2 = (a_{22}^{(1)}, a_{32}^{(1)}, \dots, a_{m2}^{(1)})^T$$

и построим матрицу  $P$  размера  $(m-1) \times (m-1)$ , такую, что  $P \bar{v}_2 = \sigma_2 \|\bar{v}_2\| \bar{e}_2$ , где  $\bar{e}_2 = (1, 0, \dots, 0)^T$  —  $(m-1)$ -мерный вектор-столбец. Далее возьмем матрицу  $P_2$  размера  $m \times m$

$$P_2 = \begin{pmatrix} 1 & 0 \\ 0 & P \end{pmatrix},$$

тогда

$$A_2 = P_2 \cdot A_1 = P_2 P_1 \cdot A = \begin{pmatrix} \sigma_1 \|\bar{v}_1\| & a_{12}^{(1)} & a_{13}^{(1)} & \cdot & a_{1m}^{(1)} \\ 0 & \sigma_2 \|\bar{v}_2\| & a_{23}^{(2)} & \cdot & a_{2m}^{(2)} \\ 0 & 0 & a_{32}^{(2)} & \cdot & a_{3m}^{(2)} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & a_{m3}^{(2)} & \cdot & a_{mm}^{(2)} \end{pmatrix}$$

Продолжая этот процесс, получим

$$\begin{aligned}
 A_{m-1} &= P_{m-1} \cdot A_{m-1} = P_{m-1} \dots P_1 \cdot A = \\
 &= \begin{pmatrix} \sigma_1 \|\bar{v}_{(1)}\| & a_{12}^{(1)} & a_{13}^{(1)} & \cdot & a_{1m}^{(1)} \\ 0 & \sigma_2 \|\bar{v}_2\| & a_{23}^{(2)} & \cdot & a_{2m}^{(2)} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \sigma_{m-1} \|\bar{v}_{m-1}\| & a_{m-1,m}^{(m-1)} \\ 0 & 0 & \cdot & 0 & a_{mm}^{(m-1)} \end{pmatrix}
 \end{aligned}$$

Обозначим  $A_{m-1} = R$  и  $Q = P_1^T \cdot P_2^T \dots \cdot P_{m-1}^T$ , тогда получим искомое разложение

$$A = Q \cdot R.$$

В процессе реализации алгоритмов метода отражений (также как и метода вращений) вычисляется матрица  $A_{m-1}$  — правая треугольная матрица. Чтобы вычислить ортогональную матрицу  $Q$ , надо сначала положить ее равной единичной матрице, а затем на каждом шаге умножать её слева на  $P_i^T$ ,  $i = 1, \dots, m - 1$ , тогда будет вычислена и матрица  $Q$ .

### 3.4 Ортогональные методы решения СЛАУ

Если известно разложение матрицы  $A = QR$ , где  $Q$  — ортогональная матрица, а  $R$  — правая (right) треугольная матрица, то решение СЛАУ

$$Ax = b$$

сводится к решению эквивалентной системы

$$Rx = Q^T b \tag{3.2}$$

с правой треугольной матрицей. Заметим, что  $\text{cond}(A) = \text{cond}(R)$ . В зависимости от способа построения ортогонального разложения получаются различные методы решения СЛАУ.

**Метод вращений** получается, если для построения QR - разложения матрицы  $A$  использовались преобразования вращения.

В методе вращений можно уменьшить число операций, заменив умножение матрицы  $A$  на матрицу вращений  $Q_{ij}$  преобразованием только  $i$ -ой и  $j$ -ой строк по формулам

$$\begin{aligned} b_{ik}^{(1)} &= b_{ik} \cos \varphi_{ij} - a_{jk} \sin \varphi_{ij} \\ b_{jk}^{(1)} &= b_{ik} \sin \varphi_{ij} + a_{jk} \cos \varphi_{ij} \end{aligned} \quad k = 1, \dots, m.$$

Матрица  $Q$  вычисляется одновременно с матрицей  $A_{m-1} = R$ , если сначала положить ее равной единичной матрице, а затем на каждом шаге построения QR разложения преобразовывать ее элементы также как элементы матрицы  $A$ .

Приведем один из вариантов алгоритма построения QR разложения с помощью преобразований вращения:

1. присвоить матрице  $\tilde{Q}$  единичную матрицу,
2. цикл по  $i$  от 1 до  $m - 1$ ,
3. найти в цикле по  $k$  от  $i$  до  $m$  первый ненулевой элемент  $a_{ik}$  ( $|a_{ik}| > \varepsilon$ , где  $\varepsilon$  — точность расчета, конец цикла по  $k$ ),  
если окажется, что  $k > i$ , то переставить  $i$  и  $k$  столбцы и перейти к пункту 4.,  
если такого элемента нет, то переставить элементы  $i$ -ого столбца с элементами  $(m + 1 - i)$ -ого, увеличить  $i$  на единицу и продолжить цикл по  $i$ .
4. цикл по  $j$  от  $i + 1$  до  $m$ ,
5.  $r := \sqrt{a_{ii}^2 + a_{ji}^2}$ ;  $c := \frac{a_{ii}}{r}$ ;  $s := -\frac{a_{ji}}{r}$
6. цикл по  $k$  от  $i + 1$  до  $m$ ,
7.  $r := a_{ik} * c - a_{jk} * s$ ;  $a_{jk} := a_{ik} * s + a_{jk} * c$ ;  $a_{ik} := r$ ;
8.  $r := \tilde{q}_{ik} * c - \tilde{q}_{jk} * s$ ;  $\tilde{q}_{jk} := \tilde{q}_{ik} * s + \tilde{q}_{jk} * c$ ;  $\tilde{q}_{ik} := r$ ;
9. конец цикла по  $k$ ,
10. конец цикла по  $j$ ,



11. конец цикла по  $i$ .

После выполнения этих вложенных циклов получаются такие матрицы  $R = A_{m-1}$  (правая треугольная) и  $Q = \tilde{Q}^T$  (ортогональная), что  $A = QR$ .

Заметим, что, если матрица  $A$  вырожденная, то в процессе построения разложения на некотором шаге  $i$  возможно появление столбца, у которого диагональный элемент и все поддиагональные равны нулю. В этом случае элементы  $i$ -ого столбца переставляются с элементами последнего ненулевого столбца, что эквивалентно умножению справа на ортогональную матрицу перестановок. У матрицы  $A_{m-1}$  нулевыми окажутся элементы последних столбцов, начиная с диагональных. Количество ненулевых диагональных элементов у матрицы  $A_{m-1}$  равно рангу матрицы  $A$ .

Решение системы 3.2 выполняется также как обратный ход в методе Гаусса на странице 32.

Замечание. Если матрица  $A$  правая почти треугольная (правая матрица Хессенберга), то для построения ее  $QR$  разложения потребуется существенно меньше операций (умножений на матрицы  $Q_{ij}$ ).

Число операций умножения и деления, необходимых для реализации метода вращений, имеет порядок  $O(2m^3)$ .

**Метод отражений.** Метод отражений получается, если  $QR$  - разложение матрицы  $A$  построено с помощью преобразований отражения (см. стр. 37).

В методе отражений можно уменьшить число операций, создав процедуру перемножения миноров матриц. Тогда не потребуется строить матрицы  $P_i$  размера  $m \times m$ . Кроме того, так как миноры имеют меньший порядок, то при их перемножении требуется меньше операций.

В методе отражений требуется выполнить  $O\left(\frac{4}{3}m^3\right)$  операций умножения и деления.

### 3.5 Метод Холецкого (метод квадратного корня)

Некоторые методы используют особые свойства матриц. Рассмотрим случай симметричных матриц  $A = A^*$ .

#### Теорема о разложении Холецкого

Пусть  $A = A^*$  и все ее главные миноры отличны от нуля, тогда справедливо разложение  $A = L D L^*$ , где  $D$  – диагональная матрица (с элементами  $d_{ii} = 1$  или  $d_{ii} = -1$ ),  $L$  – левая треугольная матрица.

*Доказательство.* Любую матрицу с отличными от нуля главными минорами можно представить в виде

$$A = L U.$$

Из этого разложения следует цепочка равенств

$$A = L U = A^* = U^* L^* \Rightarrow L U = U^* L^* \Rightarrow U (L^*)^{-1} = L^{-1} U^*.$$

В полевом равенстве слева стоит правая, а справа — левая треугольная матрица, и это равенство возможно, если обе матрицы диагональны, то есть

$$U (L^*)^{-1} = L^{-1} U^* = D \Rightarrow U = D L^* \Rightarrow A = L D L^*.$$

□

Метод Холецкого (метод квадратного корня) основан на построении разложения  $A = L D L^*$ , где

$$L = \begin{pmatrix} l_{11} & 0 & \cdot & 0 \\ l_{21} & l_{22} & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot \\ l_{m1} & l_{m2} & \cdot & l_{mm} \end{pmatrix}, \quad L^* = \begin{pmatrix} l_{11} & l_{21}^* & \cdot & l_{m1}^* \\ 0 & l_{22} & \cdot & l_{m2}^* \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & l_{mm} \end{pmatrix},$$

$D = \text{diag}(d_{11}, d_{22}, \dots, d_{mm})$ ,  $d_{ii} = 1$  или  $d_{ii} = -1$  ( $i = 1, \dots, m$ ), символ \* над элементом означает комплексное сопряжение. В силу симметричности

матрицы  $A$  ее диагональные элементы вещественны, даже если она комплекснозначная.

Построим алгоритм вычисления элементов матрицы  $L$  для вещественной матрицы  $A$ . По правилам умножения матриц имеем

$$a_{ij} = a_{ji} = \sum_{k=1}^m l_{ik} d_{kk} l_{jk} = \sum_{k=1}^i l_{ik} d_{kk} l_{jk}, \quad i \leq j,$$

здесь используется симметричность матрицы  $A$  и равенства  $l_{ik} = 0$ , если  $k > i$ .

Сначала положим  $i = 1$ , тогда

$$a_{1j} = a_{j1} = l_{11} d_{11} l_{j1},$$

отсюда следует

$$a_{11} = l_{11}^2 d_{11}, \quad \text{и} \quad l_{j1} = \frac{a_{1j}}{l_{11} d_{11}}, \quad 2 \leq j \leq m.$$

Выберем  $d_{11} = \text{sgn}(a_{11})$ , тогда  $l_{11} = \sqrt{|a_{11}|}$  будет вещественным числом. Все элементы первого столбца матрицы  $L$  вычислены.

Пусть вычислены первые  $r - 1$  столбцы матрицы  $L$ , то есть вычислены  $l_{ij}$  для  $1 \leq i \leq r - 1$  и  $r - 1 \leq j \leq m$ .

Вычислим элементы  $r$ -ого столбца, используя равенство

$$a_{rj} = \sum_{k=1}^r l_{rk} d_{kk} l_{jk} = \sum_{k=1}^{r-1} l_{rk} d_{kk} l_{jk} + l_{rr} d_{rr} l_{jr}, \quad r \leq j \leq m.$$

Сначала положим  $j = r$ , тогда это равенство примет вид

$$a_{rr} = \sum_{k=1}^{r-1} l_{rk}^2 d_{kk} + l_{rr}^2 d_{rr}.$$

Заметим, что все слагаемые под знаком суммы вычислены на предыдущих шагах. Выберем  $d_{rr} = \text{sgn}(a_{rr} - \sum_{k=1}^{r-1} l_{rk}^2 d_{kk})$ , тогда

$$l_{rr} = \sqrt{\left| a_{rr} - \sum_{k=1}^{r-1} l_{rk}^2 d_{kk} \right|}, \quad \text{а} \quad l_{jr} = \frac{a_{rj} - \sum_{k=1}^{r-1} l_{rk} d_{kk} l_{jk}}{l_{rr} d_{rr}}, \quad r < j \leq m.$$

Для такого построения разложения Холецкого требуется порядка  $O\left(\frac{1}{3}m^3\right)$  арифметических операций, т. е. вдвое меньше, чем в методе Гаусса. При решении СЛАУ это разложение позволяет также, как  $LU$  разложение, свести решение системы (2.1) к решению двух систем с треугольными матрицами.

Особенно удобна эта схема для симметричных слабо заполненных (ленточных) матриц.

### 3.6 Вопросы и упражнения для самопроверки

1. Постройте  $LU$  разложение матрицы  $A = \begin{pmatrix} 2 & -2 & -6 \\ -1 & 2 & 5 \\ 3 & -2 & -9 \end{pmatrix}$

2. Вычислите определитель матрицы  $A$  из пункта 1.

3. Сформулируйте и докажите теорему о  $UL$  разложении матрицы (аналогичную теореме о  $LU$  разложении матрицы).

4. Найдите с помощью метода Гаусса решение системы  $A\bar{x} = \bar{b}$ , где

$$A = \begin{pmatrix} 5.01 & 8 & 5 \\ 2 & 3 & 5 \\ 3 & 4.99 & 5 \end{pmatrix}, \bar{b} = \begin{pmatrix} -19.01 \\ -1 \\ -7.97 \end{pmatrix}.$$

Вычисления проводите с тремя значащими цифрами. Результат сравните с точным решением  $\bar{x}^* = (-1, -3, 2)^T$ .

5. Постройте  $QR$  разложение матрицы  $A = \begin{pmatrix} 1.0 & 2 & 3 \\ 0 & 0 & 1 \\ 2 & 3 & 4 \end{pmatrix}$  с помощью

преобразования вращения (Гивенса). Вычисления проводите с тремя значащими цифрами.

6. Постройте  $QR$  разложение матрицы матрицы  $A$  из пункта 5 с помощью преобразования отражения (Хаусхолдера). Вычисления проводите с тремя значащими цифрами.

7. Найдите с помощью метода Холецкого решение системы  $A\bar{x} = \bar{b}$ , где

$$A = \begin{pmatrix} 1.0 & -1.0 & 2.0 \\ 1.0 & 5.0 & 0 \\ 2.0 & 0 & 6.0 \end{pmatrix}, \bar{b} = \begin{pmatrix} 6.0 \\ -14 \\ 10 \end{pmatrix}.$$

Вычисления проводите с тремя значащими цифрами.

8. Измените алгоритм метода Гаусса на странице 32 для решения  $l$  систем линейных уравнений с одинаковой матрицей  $A$  и  $l$  разными столбцами правых частей.

9. Постойте алгоритм  $QR$  разложения матрицы с помощью метода отражений.

## 4 Итерационные методы решения СЛАУ

Суть итерационных методов состоит в выборе начального приближения  $\bar{x}^{(0)}$  к точному решению  $\bar{x}^*$  системы

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

и построении алгоритма вычисления последовательных приближений (итераций, от латинского слова “итерацио” – повторение)  $\bar{x}^{(n)}$  таких, чтобы  $\lim_{n \rightarrow \infty} \bar{x}^{(n)} = \bar{x}^*$ . Вычисление итераций продолжается до тех пор, пока не будет достигнута заданная точность, т.е. пока не будет выполнено условие  $\|\bar{x}^{(n)} - \bar{x}^*\| < \varepsilon$ . Заметим, что при расчете можно использовать лишь вычисляемую (практическую) оценку погрешности, величина которой достаточно близка к значению  $\|\bar{x}^{(n)} - \bar{x}^*\|$ .

Итерационные методы особенно важны при решении больших систем, так как они обычно требуют меньшее число операций  $O(m^2)$ , т.е. на порядок меньше, чем точные методы ( $O(m^3)$ ).

Чтобы организовать вычисления  $n$ -ой итерации  $\bar{x}^{(n)}$  выберем некоторую невырожденную матрицу  $B$ , набор параметров  $\tau_n$ ,  $n = 1, 2, \dots$  и рассмотрим итерационный процесс

$$B \frac{\bar{x}^{(n)} - \bar{x}^{(n-1)}}{\tau_n} + A \bar{x}^{(n-1)} = \bar{b}, \quad n = 1, 2, \dots \quad (4.1)$$

Если полученная таким способом последовательность  $\bar{x}^{(n)}$  сходится к некоторому пределу  $\bar{x}^*$ , то этот предел является точным решением системы  $A \bar{x} = \bar{b}$ . Действительно, переходя к пределу в системе (4.1) при  $n \rightarrow \infty$ , получим  $A \bar{x}^* = \bar{b}$ .

Построенный процесс (4.1) вычисления  $\bar{x}^{(n)}$  называется двухслойным нестационарным итерационным методом решения СЛАУ. Если  $\tau_n = \tau$ , то итерационный метод называется стационарным. Очевидно, что параметры  $\tau_n$ ,  $n = 1, 2, \dots$  и матрицу  $B$  надо выбрать так, чтобы итерационный процесс сходиллся и как можно быстрее, причем вычисление матрицы  $B^{-1}$

должно быть существенно проще вычисления  $A^{-1}$ . Выразим  $\bar{x}^{(n)}$  из (4.1)

$$\bar{x}^{(n)} = (E - \tau_n B^{-1} A) \bar{x}^{(n-1)} + \tau_n B^{-1} \bar{b},$$

матрица  $S_n = E - \tau_n B^{-1} A$  называется матрицей перехода.

Обозначим через  $\bar{r}^{(n)} = \bar{x}^{(n)} - \bar{x}^*$  ошибку  $n$ -ой итерации. Задача вычисления приближенного решения системы (4.1) с точностью  $\varepsilon$  сводится к нахождению такого значения  $n = n(\varepsilon)$ , что

$$\|\bar{r}^{(n)}\| = \|\bar{x}^{(n)} - \bar{x}^*\| \leq \varepsilon \|\bar{x}^{(0)} - \bar{x}^*\| \leq \varepsilon \|\bar{r}^{(0)}\|.$$

Если для вычисления  $k$ -ой итерации требуется  $q_k$  арифметических действий, то для вычисления приближенного решения с точностью  $\varepsilon$  требуется  $q(\varepsilon) = \sum_{k=1}^{n(\varepsilon)} q_k$  операций. Итерационный метод будет экономичным, если минимизировать  $q(\varepsilon)$  за счет выбора параметров  $\tau_n$ ,  $n = 1, 2, \dots$

Чтобы получить систему для вычисления  $r^{(n)}$  – ошибки  $n$ -ой итерации, выпишем две системы

$$B \frac{\bar{x}^{(n)} - \bar{x}^{(n-1)}}{\tau_n} + A \bar{x}^{(n-1)} = \bar{b},$$

$$B \frac{\bar{x}^* - \bar{x}^*}{\tau_n} + A \bar{x}^* = \bar{b}.$$

Вычитая из первого векторного уравнения второе, получим

$$B \frac{\bar{r}^{(n)} - \bar{r}^{(n-1)}}{\tau_n} + A \bar{r}^{(n-1)} = 0 \Rightarrow \bar{r}^{(n)} = (E - \tau_n B^{-1} A) \bar{r}^{(n-1)} = S_n \bar{r}^{(n-1)}$$

Далее будем рассматривать стационарные итерационные методы, т.е.  $\tau_n = \tau$  для любого  $n$ , а матрица перехода  $S = E - \tau B^{-1} A$ . Тогда

$$\bar{x}^{(n)} = (E - \tau B^{-1} A) \bar{x}^{(n-1)} + \tau B^{-1} \bar{b} = S \bar{x}^{(n-1)} + \bar{c}, \quad (4.2)$$

а

$$\bar{r}^{(n)} = S \bar{r}^{(n-1)} = S S \bar{r}^{(n-2)} \dots = S^n \bar{r}^{(0)}.$$

*Теорема о достаточном условии сходимости стационарного итерационного метода.*

Если норма матрицы  $S$  меньше единицы, то итерационный процесс (4.2) с матрицей перехода  $S$  сходится.

*Доказательство.* Погрешность  $\bar{r}^{(n)} = S^n \bar{r}^{(0)}$ , отсюда для согласованных норм матрицы и вектора получается  $\|\bar{r}^{(n)}\| \leq \|S\|^n \|\bar{r}^{(0)}\|$ . Если  $\|S\| \leq 1$ , то  $\|\bar{r}^{(n)}\| \rightarrow 0$  при  $n \rightarrow \infty$ .  $\square$

*Теорема о необходимом и достаточном условии сходимости стационарного итерационного метода.*

Для того чтобы итерационный процесс (4.2) с матрицей перехода  $S$  сошелся для любого начального приближения необходимо и достаточно, чтобы все собственные числа матрицы  $S$  были по модулю меньше единицы.

*Доказательство.* Достаточность.

Пусть подобное преобразование с матрицей  $G$  приводит невырожденную матрицу  $S$  к жордановой форме  $J = G S G^{-1}$ , где

$$J = \begin{pmatrix} J_{\alpha_1} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & J_{\alpha_k} \end{pmatrix}, \quad \alpha_1 + \dots + \alpha_k = m.$$

Жорданова клетка  $J_{\alpha_k}$  размера  $\alpha_k \times \alpha_k$  имеет вид

$$J_{\alpha_k} = \begin{pmatrix} \mu_k & 1 & 0 & \dots & 0 \\ 0 & \mu_k & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \mu_k \end{pmatrix},$$

где  $\mu_k$  – собственное число матрицы  $S$ .

Покажем на примере жордановой клетки размера  $3 \times 3$ , что  $J_{\alpha_k}^n \rightarrow 0$  при  $n \rightarrow \infty$ , если  $|\mu_k| < 1$ . Действительно, вычислим несколько степеней  $J_{\alpha_k}$



$$J_{\alpha_k} = \begin{pmatrix} \mu_k & 1 & 0 \\ 0 & \mu_k & 1 \\ 0 & 0 & \mu_k \end{pmatrix}, \quad J_{\alpha_k}^2 = \begin{pmatrix} \mu_k^2 & 2\mu_k & 1 \\ 0 & \mu_k^2 & 2\mu_k \\ 0 & 0 & \mu_k^2 \end{pmatrix},$$

$$J_{\alpha_k}^3 = \begin{pmatrix} \mu_k^3 & 3\mu_k^2 & 3\mu_k \\ 0 & \mu_k^3 & 3\mu_k^2 \\ 0 & 0 & \mu_k^3 \end{pmatrix}, \quad J_{\alpha_k}^4 = \begin{pmatrix} \mu_k^4 & 4\mu_k^3 & 6\mu_k^2 \\ 0 & \mu_k^4 & 4\mu_k^3 \\ 0 & 0 & \mu_k^4 \end{pmatrix},$$

и так далее  $J_{\alpha_k}^n = \begin{pmatrix} \mu_k^n & n\mu_k^{n-1} & C_n^2 \mu_k^{n-2} \\ 0 & \mu_k^n & n\mu_k^{n-1} \\ 0 & 0 & \mu_k^n \end{pmatrix}.$

Отсюда следует, что  $J^n \rightarrow 0$  при  $n \rightarrow \infty$ , если у матрицы  $S$  все собственные числа  $|\mu_k| < 1$ .

Для погрешности  $\bar{r}^{(n)}$  получаем

$$\bar{r}^{(n)} = S^n \bar{r}^{(0)} = G^{-1} \underbrace{JGG^{-1}}_E JG \dots G^{-1} JG \bar{r}^{(0)} = G^{-1} J^n G \bar{r}^{(0)}.$$

Отсюда

$$\|\bar{r}^{(n)}\| = \|G^{-1} J^n G \bar{r}^{(0)}\| \leq \|G^{-1}\| \|J^n\| \|G\| \|\bar{r}^{(0)}\| \rightarrow 0 \text{ при } n \rightarrow \infty.$$

Необходимость.

Пусть у матрицы  $S$  есть собственное число  $|\mu_p| > 1$ , а  $\bar{\varphi}_p$  — соответствующий ему собственный вектор. Выберем в качестве начального приближения  $\bar{x}^{(0)} = \bar{x}^* + \varepsilon \bar{\varphi}_p$ , где  $\bar{x}^*$  — точное решение СЛАУ,  $\varepsilon$  — малая константа. Тогда  $\bar{r}^{(0)} = \varepsilon \bar{\varphi}_p$ , а  $\bar{r}^{(n)} = \varepsilon S^n \bar{\varphi}_p = \varepsilon \mu_p^n \bar{\varphi}_p$  не стремится к 0 при  $n \rightarrow \infty$ , т.е. нет сходимости.  $\square$

#### 4.1 Метод простой итерации

Если  $B = E$ , то стационарный итерационный процесс с матрицей перехода  $S = E - \tau A$  называется методом простой итерации. Скорость его

сходимости тем выше, чем меньше максимум модулей собственных чисел матрицы перехода  $S$  ( $\max_k |\mu_k|$ ). Значит надо выбрать  $\tau$  так, чтобы  $\max_k |\mu_k|$  был минимальным.

*Пример.* Покажем как наилучшим образом выбрать  $\tau$ , если  $A = A^* > 0$  и известны границы спектра матрицы  $A$ , то есть  $0 < \gamma_1 \leq \lambda_k \leq \gamma_2$ ,  $k = 1, \dots, m$ .

Если  $\vec{\varphi}^k$  – собственный вектор, отвечающий  $\lambda_k$ , то он также является собственным вектором матрицы перехода  $S$ , и ему соответствует собственное число  $\mu_k = 1 - \tau \lambda_k$ .

Если  $0 < \tau < 2/\gamma_2$ , то  $|\mu_k| = |1 - \tau \lambda_k| < 1$  и выполняется необходимое и достаточное условие сходимости итераций. Для ускорения итерационного процесса надо на интервале  $0 < \tau < 2/\gamma_2$  выбрать  $\tau = \tau_*$  так, чтобы  $|\mu_*| = \min_{0 < \tau < 2/\gamma_2} \max_{1 \leq k \leq m} |1 - \tau \lambda_k|$ .

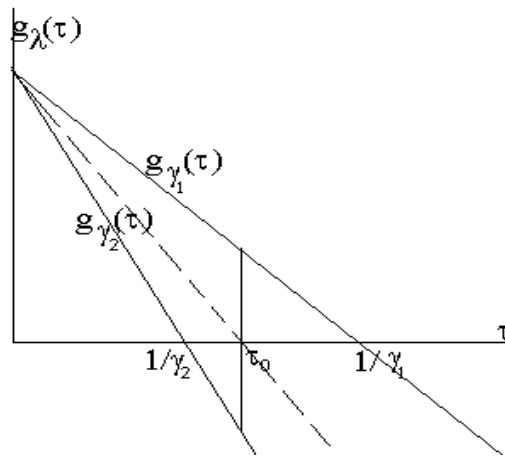


Рис. 6: Область изменения параметра  $\tau$ , в которой находится значение  $\tau = \tau_*$ , обеспечивающее самую быструю сходимость метода простых итераций для  $A = A^* > 0$ .

Рассмотрим на интервале  $\gamma_1 < \lambda < \gamma_2$  функцию  $g_\lambda(\tau) = 1 - \tau \lambda$ . Выберем такое значение  $\tau = \tau_*$ , чтобы  $|g_\lambda(\tau_*)| \leq |g_\lambda(\tau)|$  для  $\gamma_1 < \lambda < \gamma_2$  и  $0 < \tau < 2/\gamma_2$ . Заметим, что на интервале  $0 < \tau < 1/\gamma_2$  обе функции положительны, т.е.  $|g_{\gamma_2}(\tau)| \leq |g_{\gamma_1}(\tau)|$ , и отрицательны на  $\tau > 1/\gamma_1$ , т.е.  $|g_{\gamma_1}(\tau)| \leq |g_{\gamma_2}(\tau)|$ .

Теперь на отрезке  $1/\gamma_2 \leq \tau \leq 1/\gamma_1$  найдем такое значение  $\tau_0$ , при кото-

ром

$$|g_{\gamma_2}(\tau_0)| = |g_{\gamma_1}(\tau_0)|, \text{ то есть } 1 - \tau_0\gamma_1 = \tau_0\gamma_2 - 1 \Rightarrow \tau_0 = \frac{2}{\gamma_1 + \gamma_2}.$$

Покажем, что  $\tau_0$  и есть искомое  $\tau_*$ . Действительно, для любого  $\gamma_1 \leq \lambda \leq \gamma_2$  выполняются неравенства

$$|g_\lambda(\tau)| \leq |g_{\gamma_1}(\tau_0)| \quad 0 < \tau \leq \tau_0$$

$$|g_\lambda(\tau)| \leq |g_{\gamma_2}(\tau_0)| \quad \tau \geq \tau_0$$

Из этих неравенств следует, что  $|g_\lambda(\tau_0)| = \min_{0 \leq \tau \leq 2/\gamma_2} \max_{\gamma_1 \leq \lambda \leq \gamma_2} |g_\lambda(\tau)|$  (см. рис.6).

Таким образом, найдено такое значение параметра  $\tau_0$ , что

$$\max_{1 \leq k \leq m} |1 - \tau_0 \lambda_k| \leq \max_{1 \leq k \leq m} |1 - \tau \lambda_k| < 1$$

для любого значения  $0 < \tau < 2/\gamma_2$ .

Оценим для этого примера норму матрицы перехода  $S$

$$\|S\| = \max_{\gamma_1 \leq \lambda \leq \gamma_2} \left| 1 - \frac{2}{\gamma_1 + \gamma_2} \lambda_i \right| \leq \left| 1 - \frac{2\gamma_2}{\gamma_1 + \gamma_2} \right| = \frac{\gamma_2 - \gamma_1}{\gamma_1 + \gamma_2} = \frac{\zeta - 1}{\zeta + 1}$$

где  $\zeta = \frac{\gamma_2}{\gamma_1} \approx \text{cond}(A)$ . Это означает: чем больше  $\text{cond}(A)$ , тем хуже сходимость, так как  $\|S\| \rightarrow 1$  при  $\zeta \rightarrow \infty$ , т.е. для плохо обусловленных систем метод простой итерации сходится медленно (см. теорему на стр 48).

## 4.2 Метод Якоби

Запишем систему  $A\bar{x} = \bar{b}$  в виде

$$a_{ii}x_i + \sum_{j=1, j \neq i}^m a_{ij}x_j = b_i, \quad i = 1, \dots, m.$$

Если  $a_{ii} \neq 0$  ( $i = 1, \dots, m$ ) и задано  $\bar{x}^{(n)}$  –  $n$ -ое приближение к решению системы, то в итерационном методе Якоби  $i$ -ая компонента  $(n+1)$ -ого приближения вычисляется из  $i$ -ого уравнения системы

$$a_{ii}x_i^{(n+1)} + \sum_{j=1, j \neq i}^m a_{ij}x_j^{(n)} = b_i, \quad i = 1, \dots, m.$$

Чтобы получить достаточные условия сходимости метода Якоби, сведем его к двухслойному итерационному методу. Обозначим  $B = \text{diag}(a_{11}, \dots, a_{mm})$ , тогда метод Якоби можно записать в виде

$$B(\bar{x}^{(n+1)} - \bar{x}^{(n)}) + A\bar{x}^{(n)} = \bar{b} \Rightarrow \bar{x}^{(n+1)} = (E - B^{-1}A)\bar{x}^{(n)} + B^{-1}\bar{b},$$

где  $S = E - B^{-1}A$  — матрица перехода.

Из теоремы о достаточном условии сходимости получается, что метод Якоби сходится, если матрица  $A$  имеет диагональное преобладание по строкам или по столбцам. Действительно, достаточное условие сходимости для первой нормы матриц имеет вид

$$\|S\|_1 = \max_{1 \leq i \leq m} \sum_{j=1}^m |s_{ij}| < 1 \Rightarrow \sum_{j=1, j \neq i}^m \frac{|a_{ij}|}{|a_{ii}|} < 1 \Rightarrow |a_{ii}| > \sum_{j=1, j \neq i}^m |a_{ij}|,$$

а для другой нормы

$$\|S\|_\infty = \max_{1 \leq j \leq m} \sum_{i=1}^m |s_{ij}| < 1 \Rightarrow \sum_{i=1, i \neq j}^m \frac{|a_{ij}|}{|a_{ii}|} < 1 \Rightarrow |a_{ii}| > \sum_{i=1, i \neq j}^m |a_{ij}|.$$

Если задано начальное приближение  $\bar{x}^{(0)}$ , то алгоритм метода Якоби можно записать, например, так:

1.  $n := 0$ ,
2. цикл вычисления приближенного решения СЛАУ,
3. цикл по  $i$  от 1 до  $m$ ,
4. 
$$x_i^{(1)} = \frac{b_i - \sum_{j=1, j \neq i}^m a_{ij}x_j^{(0)}}{a_{ii}},$$
5. конец цикла по  $i$ ,
6.  $\bar{x}^{(0)} := \bar{x}^{(1)}$ ,  $n := n + 1$ ,
7. вычисление практической оценки погрешности по формуле (4.7) (см. ниже, стр 66),
8. если точность достигнута, то  $\bar{x}^{(1)}$  — приближенное решение системы  $A\bar{x} = \bar{b}$ , вычисленное с точностью  $\varepsilon$ , счет прекращается (выход

на пункт 11).

9. иначе (точность не достигнута), если число вычисленных итераций  $n$  не превосходит максимально допустимое число итераций  $N$ , заданное вычислителем, то вычисление приближенного решения СЛАУ продолжается,
10. иначе ( $n > N$ ) выдается информация о том, что метод расходится и вычисление приближенного решения СЛАУ прекращается (выход на пункт 11). Часто полезно выдать на печать последние две итерации, так как может оказаться, что максимально допустимое число итераций нужно незначительно увеличить.
11. конец цикла вычисления приближенного решения СЛАУ.

### 4.3 Метод Зейделя

Пусть у матрицы  $A$  системы  $A\bar{x} = \bar{b}$  все диагональные элементы  $a_{ii} \neq 0$ , и известно  $\bar{x}^{(n)}$  —  $n$ -ое приближение к решению системы. В итерационном методе Зейделя  $i$ -ая компонента  $(n + 1)$ -ого приближения вычисляется из  $i$ -ого уравнения системы, при этом используются уже вычисленные компоненты  $x_j^{(n+1)}$ ,  $j = 1, 2, \dots, i - 1$ ,

$$\begin{aligned}
 a_{11}x_1^{(n+1)} + a_{12}x_2^{(n)} + a_{13}x_3^{(n)} + \dots + a_{1m}x_m^{(n)} &= b_1 \\
 a_{21}x_1^{(n+1)} + a_{22}x_2^{(n+1)} + a_{23}x_3^{(n)} + \dots + a_{2m}x_m^{(n)} &= b_2 \\
 \dots & \\
 a_{m1}x_1^{(n+1)} + a_{m2}x_2^{(n+1)} + a_{m3}x_3^{(n+1)} + \dots + a_{mm}x_m^{(n+1)} &= b_m
 \end{aligned} \tag{4.3}$$

Чтобы выяснить условие сходимости метода Зейделя, введем обозначе-

ния

$$B = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mm} \end{pmatrix} \quad \text{и} \quad G = \begin{pmatrix} 0 & a_{12} & \dots & a_{1m} \\ 0 & 0 & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \end{pmatrix}.$$

и запишем (4.3) в виде

$$B\bar{x}^{(n+1)} + G\bar{x}^{(n)} = \bar{b}.$$

Если все диагональные элементы  $a_{ii} \neq 0$ , то матрица  $B$  обратима, и

$$\bar{x}^{(n+1)} = -B^{-1}G\bar{x}^{(n)} + B^{-1}\bar{b},$$

т.е.  $S = -B^{-1}G$  — матрица перехода. Необходимое и достаточное условие сходимости этого итерационного процесса  $|\mu_i| < 1$  для любого  $i$ , где  $\mu_i$  — собственные числа матрицы перехода  $S$ . Числа  $\mu_i$  являются корнями характеристического многочлена  $\det(S - \mu E) = 0$ . Так как

$$\begin{aligned} \det(S - \mu E) &= \det(-B^{-1}G - \mu E) = \det(-B^{-1}(G + \mu B)) = \\ &= \det(-B^{-1}) \det(G + \mu B), \end{aligned}$$

то характеристическое уравнение принимает вид

$$\det(G + \mu B) = \begin{vmatrix} \mu a_{11} & a_{12} & \dots & a_{1m} \\ \mu a_{21} & \mu a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ \mu a_{m1} & \mu a_{m2} & \dots & \mu a_{mm} \end{vmatrix} = 0.$$

Метод Зейделя сходится, если корни этого уравнения по модулю меньше единицы. Заметим, что, перестановка уравнений в исходной системе может изменить и характеристическое уравнение, и его корни.

Рассмотрим геометрическую интерпретацию метода Зейделя на примере системы второго порядка ( $m = 2$ ). Изобразим на плоскости  $(x_1, x_2)$  прямые,

первая из них:  $a_{11}x_1 + a_{12}x_2 = b_1$  ( $I$ ), а вторая:  $a_{21}x_1 + a_{22}x_2 = b_2$  ( $II$ ). Пусть выбрано начальное приближение  $\bar{x}^0 = (x_1^{(0)}, x_2^{(0)})$ . Компонента  $x_1^{(1)}$  первого приближения получается из уравнения

$$a_{11}x_1^{(1)} + a_{12}x_2^{(0)} = b_1,$$

т.е. разыскивается точка пересечения прямой  $I$  с прямой  $x_2 = x_2^{(0)}$ . Далее компонента  $x_2^{(1)}$  первого приближения получается из второго уравнения

$$a_{21}x_1^{(1)} + a_{22}x_2^{(1)} = b_2,$$

т.е. разыскивается точка пересечения прямой  $II$  с прямой  $x_1 = x_1^{(1)}$ . Аналогично разыскиваются компоненты следующих приближений.

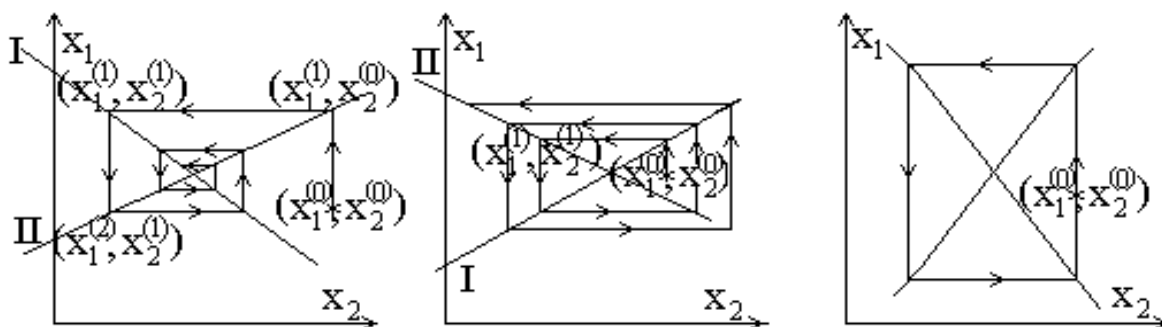


Рис. 7: Геометрическая интерпретация метода Зейделя для решения систем линейных алгебраических уравнений второго порядка

На первом рисунке изображен сходящийся итерационный процесс. На втором рисунке изображено решение той же системы, но у нее переставлены первое и второе уравнения (первая прямая стала второй, а вторая - первой), метод Зейделя стал расходиться. И на третьем рисунке изображен самый неприятный с точки зрения вычислений случай – закливание.

Если задано начальное приближение  $\bar{x}^0$ , то можно построить такой же алгоритм метода Зейделя, что и алгоритм метода Якоби (см. стр. 52):

1.  $n := 0$ ,
- 2 в цикле по  $i$  от 1 до  $m$  присвоить  $x_i^{(1)}$  значения  $x_i^{(0)}$ ,

3. начало цикла вычисления приближенного решения СЛАУ
4. цикл по  $i$  от 1 до  $m$ ,
5. 
$$x_i^{(1)} = \frac{b_i - \sum_{j=1, j \neq i}^m a_{ij} x_j^{(1)}}{a_{ii}}, \quad x_i^{(0)} := x_i^{(1)},$$
6. конец цикла по  $i$ ,
7.  $n := n + 1$ ,
8. далее все пункты совпадают с пунктами 7–11 на стр. 52.

#### 4.4 Методы спуска для решения СЛАУ

Суть методов спуска состоит в том, что решение системы уравнений  $A\bar{x} = \bar{b}$  заменяется поиском минимума некоторого функционала  $\Phi(\bar{x})$ , который достигается на точном решении  $\bar{x}^*$  системы ( $\min_{\bar{x}} \Phi(\bar{x}) = \Phi(\bar{x}^*)$ ).

Методы спуска можно использовать и для решения систем нелинейных уравнений, если удастся построить такой функционал.

В случае, когда матрица  $A$  симметричная и положительно определенная ( $A = A^* > 0$ ), функционал  $F(\bar{x}) = (A\bar{x}, \bar{x}) - 2(\bar{b}, \bar{x})$  достигает своего минимума, на точном решении  $\bar{x}^*$  СЛАУ  $A\bar{x} = \bar{b}$ .

Действительно, для любого  $\bar{x} \in R^m$ ,  $\bar{x} \neq \bar{x}^*$  выполняется цепочка равенств

$$\begin{aligned} F(\bar{x}) - F(\bar{x}^*) &= (A\bar{x}, \bar{x}) - 2(\bar{b}, \bar{x}) - (A\bar{x}^*, \bar{x}^*) + 2(\bar{b}, \bar{x}^*) = \\ &= (A\bar{x}, \bar{x}) - 2(A\bar{x}^*, \bar{x}) - (A\bar{x}^*, \bar{x}^*) + 2(A\bar{x}^*, \bar{x}^*) = \\ &= (A(\bar{x} - \bar{x}^*), \bar{x}) - (A\bar{x}^*, (\bar{x} - \bar{x}^*)) = (A(\bar{x} - \bar{x}^*), (\bar{x} - \bar{x}^*)) > 0. \end{aligned} \tag{4.4}$$

Отсюда следует, что  $\min_{\bar{x}} F(\bar{x}) = F(\bar{x}^*)$ . В случае, если матрица не является симметричной и положительно определенной, то можно выбрать, например, функционал  $\Phi(\bar{x}) = (A\bar{x} - \bar{b}, A\bar{x} - \bar{b})$  для любого  $\bar{x} \in R^m$ .

Трудные вопросы сходимости методов спуска рассматриваются для каждой задачи, в этом разделе остановимся лишь на алгоритмах вычисления итераций для методов спуска.



#### 4.4.1 Метод покоординатного спуска.

Пусть известно приближение  $\bar{x}^{(n)} = (x_1^{(n)}, x_2^{(n)}, \dots, x_m^{(n)})$  к решению  $\bar{x}^*$ , доставляющему минимум функционалу  $F(\bar{x})$ . Этот функционал можно рассматривать как функцию многих переменных

$$F(\bar{x}) = F(x_1, x_2, \dots, x_m).$$

Минимум  $F(\bar{x})$  будем разыскивать следующим образом. На первом шаге зафиксируем все переменные, кроме  $x_1$ , получим функцию одной переменной

$$\varphi_1(x_1) = F(x_1, x_2^{(n)}, \dots, x_m^{(n)}),$$

и найдем значение  $x_1^{(n+1)}$ , доставляющее минимум этой функции

$$\varphi_1(x_1^{(n+1)}) = \min_{x_1} \varphi_1(x_1).$$

На втором шаге зафиксируем все переменные, кроме  $x_2$ , получим функцию

$$\varphi_2(x_2) = F(x_1^{(n+1)}, x_2, x_3^{(n)}, \dots, x_m^{(n)})$$

и найдем ее минимум

$$\varphi_2(x_2^{(n+1)}) = \min_{x_2} \varphi_2(x_2).$$

За  $m$  шагов такого покоординатного спуска будут вычислены все компоненты  $(n+1)$ -ого приближения.

Рассмотрим геометрическую интерпретацию сходящегося метода покоординатного спуска для  $m = 2$  и матрицы  $A = A^* > 0$ . Построим линии уровня  $F(\bar{x}) = (A\bar{x}, \bar{x}) - 2(\bar{b}, \bar{x}) = C - \text{const}$ . Для такой матрицы они представляют собою вложенные друг в друга эллипсы. Точка  $(x_1^*, x_2^*)$ , отвечающая  $C = 0$ , является точным решением системы.

Выберем начальное приближение  $\bar{x}^{(0)} = (x_1^{(0)}, x_2^{(0)})$ , затем, двигаясь от точки  $A_0(x_1^{(0)}, x_2^{(0)})$  вдоль прямой  $x_1 = x_1^{(0)}$ , найдем точку  $a_1(x_1^{(1)}, x_2^{(0)})$ , которой соответствует минимальное значение константы  $C$  вдоль этой прямой. Далее, двигаясь от точки  $a_1$  вдоль прямой  $x_2 = x_2^{(0)}$ , найдем точку

$A_1(x_1^{(1)}, x_2^{(1)})$  (первое приближение), соответствующую минимуму константы  $C$ . Продолжая такое движение вдоль прямых, параллельных координатным осям, получим приближенное решение системы  $A\bar{x} = \bar{b}$

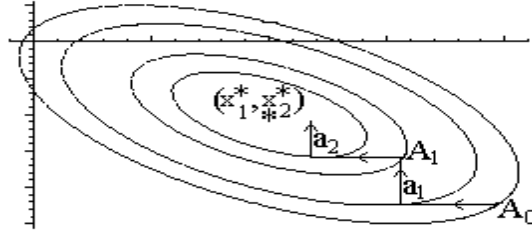


Рис. 8: Геометрическая интерпретация метода покоординатного спуска для решения системы линейных алгебраических уравнений второго порядка  $A\bar{x} = \bar{b}$ , где  $A = A^* > 0$ .

Распишем  $k$ -ый шаг алгоритма покоординатного спуска для симметричной матрицы. На  $k$ -ом шаге разыскивается  $k$ -ая координата  $n + 1$ -ого приближения  $x_k^{(n+1)}$ , доставляющая минимум функции

$$\varphi_k(x_k) = F(x_1^{(n+1)}, \dots, x_{k-1}^{(n+1)}, x_k, x_{k+1}^{(n)}, \dots, x_m^{(n)}).$$

Искомая координата  $n + 1$ -ого приближение находится из уравнения

$$\frac{\partial \varphi_k(x_k)}{\partial x_k} = 0 \implies \frac{\partial}{\partial x_k} F(x_1^{(n+1)}, \dots, x_{k-1}^{(n+1)}, x_k, x_{k+1}^{(n)}, \dots, x_m^{(n)}) = 0.$$

Чтобы вычислить  $\frac{\partial F}{\partial x_k}$ , запишем  $F = (Ax, x) - 2(b, x)$  в виде

$$F = \sum_{i=1}^m Ax_i x_i - 2 \sum_{i=1}^m b_i x_i = \sum_{i=1}^m \sum_{j=1}^m a_{ij} x_j x_i - 2 \sum_{i=1}^m b_i x_i,$$

и получим

$$\frac{\partial F}{\partial x_k} = \sum_{j=1}^m a_{kj} x_j + \sum_{i=1}^m a_{ik} x_i - 2b_k = 2 \left( \sum_{j=1}^m a_{jk} x_j - b_k \right) = 2(A\bar{x} - \bar{b})_k. \quad (4.5)$$

Значение  $x_k^{(n+1)}$  вычисляется из уравнения

$$\sum_{i=1}^{k-1} a_{ik} x_i^{(n+1)} + a_{kk} x_k^{(n+1)} + \sum_{i=k+1}^m a_{ik} x_i^n = b_k.$$

Заметим, что алгоритм построенного метода покоординатного спуска совпадает с алгоритмом метода Зейделя. Известно, что метод покоординатного спуска (а значит и метод Зейделя) сходится для симметричной положительно определенной матрицы ( $A = A^* > 0$ ).

#### 4.4.2 Метод градиентного спуска

Известно, что функционал  $F(\bar{x})$  в малой окрестности точки  $\bar{x}$  (локально) быстрее всего убывает в направлении противоположном градиенту  $\text{grad } F = \left( \frac{\partial F}{\partial x_1}, \frac{\partial F}{\partial x_2}, \dots, \frac{\partial F}{\partial x_m} \right)$  (см. стр. 71). Очередную итерацию  $\bar{x}^{(n+1)}$  будем разыскивать в виде

$$\bar{x}^{(n+1)} = \bar{x}^{(n)} - \alpha_n \text{grad } F(\bar{x}^{(n)}),$$

где величина  $\alpha_n$  характеризует величину шага по направлению антиградиента. Значение  $\alpha_n$  можно выбирать так:

1. Сначала  $\alpha_n := 1$ ,
2. Если  $\|F(\bar{x}^{(n+1)})\| \leq \|F(\bar{x}^{(n)})\|$ , то будем считать, что итерация  $\bar{x}^{(n+1)}$  вычислена. Если же  $\|F(\bar{x}^{(n+1)})\| \geq \|F(\bar{x}^{(n)})\|$ , т.е. сделан слишком большой шаг по направлению  $\text{grad } F(\bar{x}^n)$ , то  $\alpha_n := \alpha_n/2$ .
3. Значение  $\bar{x}^{(n+1)}$  пересчитывается с новым значением  $\alpha_n$ . Такое деление  $\alpha_n$  продолжается до тех пор пока  $\|F(\bar{x}^{(n+1)})\| \leq \|F(\bar{x}^{(n)})\|$  или пока не будет исчерпано допустимое число  $N$  изменений  $\alpha_n$ . Если число изменений  $\alpha_n$  превзойдет  $N$ , то метод расходится и нужно прекратить расчет, выдав соответствующую информацию.
4. После вычисления  $\bar{x}^{(n+1)}$  проверяется достигнута ли требуемая точность  $\varepsilon$ . Здесь возможны два варианта выхода по достижению точности  $\varepsilon$ :

выход по „невязке“, когда

$$\|F(\bar{x}^{(n+1)}) - F(\bar{x}^n)\| \leq \varepsilon, \quad \text{если } \|F(\bar{x}^0)\| \leq 1,$$

$$\frac{\|F(\bar{x}^{(n+1)}) - F(\bar{x}^n)\|}{\|F(\bar{x}^{(n+1)})\|} \leq \varepsilon, \quad \text{если } \|F(\bar{x}^0)\| \geq 1,$$

выход по „шагу итерации“, когда

$$\begin{aligned} \|\bar{x}^{(n+1)} - \bar{x}^n\| &\leq \varepsilon, & \text{если } \|\bar{x}^0\| &\leq 1, \\ \frac{\|\bar{x}^{(n+1)} - \bar{x}^n\|}{\|\bar{x}^{(n+1)}\|} &\leq \varepsilon, & \text{если } \|\bar{x}^0\| &\geq 1. \end{aligned}$$

#### 4.4.3 Метод наискорейшего градиентного спуска

Можно  $\alpha_n$  определять так, чтобы наименьшим было значение  $|\varphi(\alpha_n)|$ , где функция

$$\varphi(\alpha_n) = F(\bar{x}^{(n+1)}) = F(\bar{x}^{(n)} - \alpha_n \text{grad } F(\bar{x}^{(n)})).$$

Будем разыскивать такое  $\alpha_n$  в случае  $A = A^* > 0$ . Для симметричной матрицы ( см. (4.5))

$$\text{grad } F(\bar{x}^{(n)}) = \left( \frac{\partial F}{\partial x_1}(\bar{x}^{(n)}), \frac{\partial F}{\partial x_2}(\bar{x}^{(n)}), \dots, \frac{\partial F}{\partial x_m}(\bar{x}^{(n)}) \right) = 2(A\bar{x}^{(n)} - \bar{b})$$

тогда  $\bar{x}^{(n+1)} = \bar{x}^{(n)} - 2\alpha_n(A\bar{x}^{(n)} - \bar{b})$  и  $\varphi(\alpha_n) = F(\bar{x}^{(n)} - 2\alpha_n(A\bar{x}^{(n)} - \bar{b}))$ .

Значение  $\alpha_n$ , доставляющее минимум  $\varphi(\alpha_n)$ , удовлетворяет условию

$$\frac{d\varphi}{d\alpha_n} = \frac{d}{d\alpha_n} F(\bar{x}^{(n+1)}) = \frac{d}{d\alpha_n} ((A\bar{x}^{(n+1)}, \bar{x}^{(n+1)}) - 2(\bar{b}, \bar{x}^{(n+1)})) = 0.$$

Учитывая симметричность матрицы  $A$ , получаем

$$\begin{aligned} \left( \left( A \frac{d}{d\alpha_n} \bar{x}^{(n+1)}, \bar{x}^{(n+1)} \right) + A \left( \bar{x}^{(n+1)}, \frac{d}{d\alpha_n} \bar{x}^{(n+1)} \right) - 2 \left( \bar{b}, \frac{d}{d\alpha_n} \bar{x}^{(n+1)} \right) \right) = \\ = 2 \left( A\bar{x}^{(n+1)} - \bar{b}, \frac{d}{d\alpha_n} \bar{x}^{(n+1)} \right) = 0. \end{aligned}$$

Вычислим производную

$$\frac{d}{d\alpha_n} \bar{x}^{(n+1)} = -2(A\bar{x}^{(n)} - \bar{b}).$$

Подставляя эту производную и  $\bar{x}^{(n+1)}$  в выведенное условие, получим

$$\left( A \underbrace{(\bar{x}^{(n)} - \alpha_n(A\bar{x}^{(n)} - \bar{b}))}_{\bar{x}^{(n+1)}} - \bar{b}, A\bar{x}^{(n)} - \bar{b} \right) = 0.$$

Отсюда искомое значение  $\alpha_n$

$$\alpha_n = \frac{(A\bar{x}^{(n)} - \bar{b}, A\bar{x}^{(n)} - \bar{b})}{(A(A\bar{x}^{(n)} - \bar{b}), A\bar{x}^{(n)} - \bar{b})}$$

На рис.9 приведена геометрическая интерпретация метода наискорейшего градиентного спуска для решения СЛАУ с симметричной положительно определенной матрицей второго порядка.

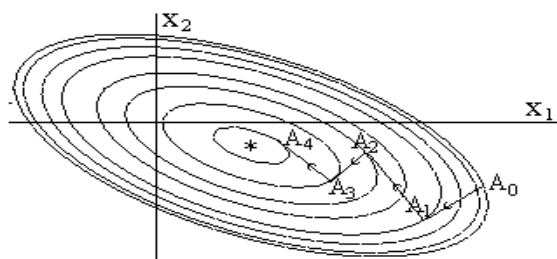


Рис. 9: Геометрическая интерпретация метода наискорейшего градиентного спуска для решения системы линейных алгебраических уравнений второго порядка  $A\bar{x} = \bar{b}$ , где  $A = A^* > 0$ .

Пусть задано начальное приближение, т.е. известна точка  $A_0 = (x_1^0, x_2^0)$ . Любой шаг по направлению  $\text{grad } F(\bar{x}^0)$  (вдоль прямой  $A_0A_1$ ), который не больше длины отрезка  $A_0A_1$  позволяет вычислить первое приближение  $(x_1^1, x_2^1)$  ( точку  $A_1$ ) к искомому решению (обозначенному на рисунке символом \*). Если выбран шаг, равный длине отрезка  $A_0A_1$ , то получается метод наискорейшего градиентного спуска, если выбран меньший шаг, то – метод градиентного спуска.

#### 4.5 Методы релаксации

Методы релаксации используются при решении СЛАУ с симметричной положительно определенной матрицей  $A = A^* > 0$ . Пусть  $\bar{x}^{(n)} = (x_1^{(n)}, x_2^{(n)}, \dots, x_m^{(n)})$  – известное  $n$ -ое приближение к искомому решению  $\bar{x}^*$ . Будем разыскивать  $\bar{x}^{(n+1)}$  в виде

$$\bar{x}^{(n+1)} = (x_1^{(n)}, \dots, x_i^{(n)} - \alpha, \dots, x_m^{(n)}) = \bar{x}^{(n)} - \alpha \bar{e}_i,$$

где  $\bar{e}_i = (0, \dots, \underbrace{1}_{i\text{-место}}, 0, \dots, 0)$ . Выберем параметр  $\alpha$  так, чтобы

$$|F(\bar{x}^{(n+1)})| = |(A\bar{x}^{(n+1)}, \bar{x}^{(n+1)}) - 2(\bar{b}, \bar{x}^{(n+1)})|$$

убывало с возрастанием  $n$ . Это эквивалентно требованию убывания (см. (4.4))

$$|F(\bar{x}^{(n+1)}) - F(\bar{x}^*)| = |(A(\bar{x}^{(n+1)} - \bar{x}^*), \bar{x}^{(n+1)} - \bar{x}^*)| = |(A\bar{r}^{(n+1)}, \bar{r}^{(n+1)})|,$$

где ошибка  $(n+1)$ -ой итерации

$$\bar{r}^{(n+1)} = \bar{x}^{(n+1)} - \bar{x}^* = \bar{r}^{(n)} - \alpha\bar{e}_i. \quad (4.6)$$

Сделаем преобразования

$$\begin{aligned} (A\bar{r}^{(n+1)}, \bar{r}^{(n+1)}) &= (A\bar{r}^{(n)} - \alpha A\bar{e}_i, \bar{r}^{(n)} - \alpha\bar{e}_i) = \\ &= (A\bar{r}^{(n)}, \bar{r}^{(n)}) - \alpha(A\bar{r}^{(n)}, \bar{e}_i) - \alpha(A\bar{e}_i, \bar{r}^{(n)}) + \alpha^2(A\bar{e}_i, \bar{e}_i). \end{aligned}$$

Обозначая  $z_i^n = (A\bar{r}^{(n)}, \bar{e}_i)$ , используя симметричность матрицы, получим

$$\begin{aligned} (A\bar{r}^{(n+1)}, \bar{r}^{(n+1)}) &= (A\bar{r}^{(n)}, \bar{r}^{(n)}) - 2\alpha z_i^n + \alpha^2 a_{ii} = \\ &= (A\bar{r}^{(n)}, \bar{r}^{(n)}) + \frac{1}{a_{ii}}(z_i^n - \alpha a_{ii})^2 - \frac{(z_i^n)^2}{a_{ii}}. \end{aligned}$$

Отсюда следует, что  $(A\bar{r}^{(n+1)}, \bar{r}^{(n+1)})$  убывает, если  $|z_i^n - \alpha a_{ii}| < |z_i^n|$ . У положительно определенной матрицы  $a_{ii} > 0$ , значит  $\alpha$  можно выбрать из интервала

$$0 < \alpha < 2\frac{z_i^n}{a_{ii}},$$

или, если положить  $\alpha = \omega z_i^n / a_{ii}$ , то  $0 < \omega < 2$ . Для построения алгоритма вычисления  $(n+1)$ -ого приближения преобразуем

$$z_i^{(n+1)} = (A\bar{r}^{(n+1)}, \bar{e}_i) = (A(\bar{x}^{(n+1)} - \bar{x}^*), \bar{e}_i) = (A\bar{x}^{(n+1)}, \bar{e}_i) - (\bar{b}, \bar{e}_i).$$

С другой стороны, учитывая (4.6), преобразуем  $z_i^{(n+1)}$  также к виду

$$\begin{aligned} z_i^{(n+1)} &= (A(\bar{x}^{(n+1)} - \bar{x}^*), \bar{e}_i) = (A(\bar{x}^{(n)} - \bar{x}^* - \alpha \bar{e}_i), \bar{e}_i) = (A(\bar{x}^{(n)} - \bar{x}^*), \bar{e}_i) - \\ &\quad - \alpha(A\bar{e}_i, \bar{e}_i) = (A\bar{x}^{(n)} - b, \bar{e}_i) - \alpha a_{ii} = (A\bar{x}^{(n)}, \bar{e}_i) - (b, \bar{e}_i) - \omega z_i^n = \\ &= (A\bar{x}^{(n)}, \bar{e}_i) - (b, \bar{e}_i) - \omega(A\bar{x}^{(n)}, \bar{e}_i) = (A\bar{x}^{(n)}, \bar{e}_i) - (b, \bar{e}_i) - \\ &\quad - \omega(A(\bar{x}^{(n)} - \bar{x}^*), \bar{e}_i) = (1 - \omega)(A\bar{x}^{(n)}, \bar{e}_i) - (1 - \omega)(b, \bar{e}_i). \end{aligned}$$

Отсюда получается

$$(A(\bar{x}^{(n+1)} - (1 - \omega)\bar{x}^{(n)}), \bar{e}_i) = \omega(\bar{b}, \bar{e}_i), \quad 0 \leq \omega \leq 2,$$

то есть  $\bar{x}^{(n+1)}$  является решением СЛАУ

$$\begin{aligned} (A(\bar{x}^{(n+1)} - (1 - \omega)\bar{x}^{(n)}))_i &= \omega \bar{b}_i, \Rightarrow (A((\bar{x}^{(n+1)} - \bar{x}^{(n)}) + \omega \bar{x}^{(n)}))_i = \omega \bar{b}_i, \\ i &= 1, 2, \dots, m. \end{aligned}$$

Здесь надо вспомнить, что  $\bar{x}^{(n+1)} = (x_1^{(n)}, \dots, x_{i-1}^{(n)}, x_i^{(n+1)}, x_{i+1}^{(n)}, \dots, x_m^{(n)})$ . Таким образом,  $i$ -ая компонента  $(n+1)$ -ого приближения вычисляется из уравнения

$$a_{ii}x_i^{(n+1)} = (1 - \omega)a_{ii}x_i^{(n)} - \omega \sum_{j=1, j \neq i}^m a_{ij}x_j^{(n)} + \omega b_i$$

Если  $0 < \omega < 1$ , то говорят о нижней релаксации, если  $1 < \omega < 2$ , то говорят о верхней релаксации. При  $\omega = 1$  получается метод покоординатного спуска (метод Зейделя).

#### 4.6 Практическая оценка погрешности итерационного метода.

Рассмотрим стационарный итерационный метод

$$\bar{x}^{(n)} = S\bar{x}^{(n-1)} + \bar{c},$$

где  $S$  – матрица перехода, а вектор  $\bar{c}$  – преобразованная правая часть системы  $A\bar{x} = \bar{b}$  (2.1).

Если выполнено условие  $\|\bar{r}^{(n)}\| = \|\bar{x}^{(n)} - \bar{x}^*\| \leq \varepsilon$ , то вектор  $\bar{x}^{(n)}$  можно считать искомым приближением к точному решению СЛАУ (2.1)  $\bar{x}^*$  с точностью  $\varepsilon$ .

Однако погрешность  $\|\bar{r}^{(n)}\| = \|\bar{x}^{(n)} - \bar{x}^*\|$  вычислить невозможно, поэтому требуется ввести практическую вычисляемую погрешность  $\bar{v}^{(n)}$  такую, что

$$\frac{\|\bar{v}^{(n)} - \bar{r}^{(n)}\|}{\|\bar{r}^{(n)}\|} \rightarrow 0 \text{ при } n \rightarrow \infty.$$

Построим  $\bar{v}^{(n)}$  для матрицы  $S$  простой структуры, у которой жорданова форма диагональна, т.е. существует полный набор собственных векторов, которые можно принять за базис в  $R^m$ .

Пусть собственные значения  $\mu_i$  матрицы перехода  $S$  упорядочены

$$1 > |\mu_1| > |\mu_2| \geq |\mu_3| \geq \dots \geq |\mu_m|,$$

каждому  $\mu_i$  соответствует линейно независимый собственный вектор  $\bar{\varphi}_i$ .

Разложим ошибку нулевого приближения по собственным векторам

$$\bar{r}^{(0)} = \sum_{i=1}^m C_i \bar{\varphi}_i,$$

тогда

$$\bar{r}^{(n)} = S^n \bar{r}^{(0)} = \sum_{i=1}^m C_i S^n \bar{\varphi}_i = \sum_{i=1}^m C_i \mu_i^n \bar{\varphi}_i = C_1 \mu_1^n \bar{\varphi}_1 + O(\mu_2^n).$$

Здесь под знаком  $O(\mu_2^n)$  понимается вектор, норма которого имеет порядок  $O(|\mu_2|^n)$ .

Обозначим  $\bar{w}^{(n)} = C_1 \mu_1^n \bar{\varphi}_1$  (заметим, что  $\|\bar{w}^{(n)}\| = O(|\mu_1|^n)$ ) и выпишем ряд равенств

$$\begin{aligned} \bar{x}^{(n-2)} - \bar{x}^* &= \bar{r}^{(n-2)} = C_1 \mu_1^{n-2} \bar{\varphi}_1 + O(\mu_2^n) = \mu_1^{-2} \bar{w}^{(n)} + O(\mu_2^n), \\ \bar{x}^{(n-1)} - \bar{x}^* &= \bar{r}^{(n-1)} = C_1 \mu_1^{n-1} \bar{\varphi}_1 + O(\mu_2^n) = \mu_1^{-1} \bar{w}^{(n)} + O(\mu_2^n), \\ \bar{x}^{(n)} - \bar{x}^* &= \bar{r}^{(n)} = C_1 \mu_1^n \bar{\varphi}_1 + O(\mu_2^n) = \bar{w}^{(n)} + O(\mu_2^n). \end{aligned}$$



Отсюда получим

$$\begin{aligned}\bar{x}^{(n)} - \bar{x}^{(n-1)} &= (1 - \mu_1^{-1})\bar{w}^{(n)} + O(\mu_2^n), \\ \bar{x}^{(n-1)} - \bar{x}^{(n-2)} &= \mu_1^{-1}(1 - \mu_1^{-1})\bar{w}^{(n)} + O(\mu_2^n).\end{aligned}$$

Введем величину

$$\mu_1^{(n)} = \frac{(\bar{x}^{(n)} - \bar{x}^{(n-1)}, \bar{x}^{(n)} - \bar{x}^{(n-1)})}{(\bar{x}^{(n-1)} - \bar{x}^{(n-2)}, \bar{x}^{(n)} - \bar{x}^{(n-1)})},$$

и вычислим скалярные произведения

$$\begin{aligned}(\bar{x}^{(n)} - \bar{x}^{(n-1)}, \bar{x}^{(n)} - \bar{x}^{(n-1)}) &= (1 - \mu_1^{-1}) (\bar{w}^{(n)}, \bar{w}^{(n)}) + O(|\mu_2|^n \|\bar{w}^{(n)}\|), \\ (\bar{x}^{(n-1)} - \bar{x}^{(n-2)}, \bar{x}^{(n)} - \bar{x}^{(n-1)}) &= \mu_1^{-1} (1 - \mu_1^{-1}) (\bar{w}^{(n)}, \bar{w}^{(n)}) + O(|\mu_2|^n \|\bar{w}^{(n)}\|).\end{aligned}$$

Определим предел  $\mu_1^{(n)}$  при  $n \rightarrow \infty$

$$\begin{aligned}\mu_1^{(n)} &= \frac{(1 - \mu_1^{-1}) (\bar{w}^{(n)}, \bar{w}^{(n)}) + O(|\mu_2|^n \|\bar{w}^{(n)}\|)}{\mu_1^{-1} (1 - \mu_1^{-1}) (\bar{w}^{(n)}, \bar{w}^{(n)}) + O(|\mu_2|^n \|\bar{w}^{(n)}\|)} = \\ &= \frac{(1 - \mu_1^{-1})^2 \|\bar{w}^{(n)}\|^2 + O(|\mu_1 \mu_2|^n)}{\mu_1^{-1} (1 - \mu_1^{-1})^2 \|\bar{w}^{(n)}\|^2 + O(|\mu_1 \mu_2|^n)} = \\ &= \frac{(1 - \mu_1^{-1})^2 \|\bar{w}^{(n)}\|^2 \left(1 + O\left(\frac{|\mu_2|^n}{|\mu_1|^n}\right)\right)}{\mu_1^{-1} (1 - \mu_1^{-1})^2 \|\bar{w}^{(n)}\|^2 \left(1 + O\left(\frac{|\mu_2|^n}{|\mu_1|^n}\right)\right)} = \\ &= \mu_1 \left(1 + O\left(\left(\frac{|\mu_2|}{|\mu_1|}\right)^n\right)\right) \rightarrow \mu_1 \quad \text{при } n \rightarrow \infty,\end{aligned}$$

так как  $\frac{|\mu_2|}{|\mu_1|} < 1$ . Отсюда следует, что  $\mu_1^{(n)} - \mu_1 = O\left(\left(\frac{|\mu_2|}{|\mu_1|}\right)^n\right)$ .

Теперь построим выражение для  $\bar{v}^{(n)}$

$$\bar{v}^{(n)} = \frac{\bar{x}^{(n)} - \bar{x}^{(n-1)}}{1 - \left(\mu_1^{(n)}\right)^{-1}},$$

и преобразуем его

$$\begin{aligned}
\bar{v}^{(n)} &= \frac{(1 - \mu_1^{-1})\bar{w}^{(n)} + O(\mu_2^n)}{1 - (\mu_1^{(n)})^{-1}} = \\
&= \frac{\left(1 - (\mu_1^{(n)})^{-1}\right)\bar{w}^{(n)} + \left((\mu_1^{(n)})^{-1} - \mu_1^{-1}\right)\bar{w}^{(n)} + O(\mu_2^n)}{1 - (\mu_1^{(n)})^{-1}} = \\
&= \bar{w}^{(n)} + \frac{(\mu_1 - \mu_1^{(n)})\bar{w}^{(n)}}{\mu_1^{(n)}\mu_1\left(1 - (\mu_1^{(n)})^{-1}\right)} + O(\mu_2^n) = \bar{w}^{(n)} + O(\mu_2^n).
\end{aligned}$$

Покажем, что  $\bar{v}^n$  искомая практическая (вычисляемая) оценка погрешности  $n$ -ой итерации

$$\begin{aligned}
\frac{\|\bar{v}^{(n)} - \bar{r}^{(n)}\|^2}{\|\bar{r}^{(n)}\|^2} &= \frac{\|(\bar{w}^{(n)} + O(\mu_2^n)) - (\bar{w}^{(n)} + O(\mu_2^n))\|^2}{(\bar{w}^{(n)} + O(\mu_2^n), (\bar{w}^{(n)} + O(\mu_2^n))} = \\
&= \frac{O(\mu_2^{2n})}{\|\bar{w}^{(n)}\|^2 + O(\|\bar{w}^{(n)}\|\mu_2^n)} = \frac{O(\mu_2^{2n})}{O(\mu_1^{2n})} = O\left(\left(\frac{\mu_2}{\mu_1}\right)^{2n}\right) \rightarrow 0
\end{aligned}$$

при  $n \rightarrow \infty$ , так как  $|\mu_2/\mu_1| < 1$ . Таким образом, построена близкая к  $\bar{r}_n$  вычисляемая оценка  $\bar{v}_n$ , которую при практическом счете записывают в виде

$$\bar{v}^{(n)} = \frac{\bar{x}^{(n)} - \bar{x}^{(n-1)}}{\mu_1^{(n)} - 1} \quad (4.7)$$

где

$$\mu_1^{(n)} = \frac{(\bar{x}^{(n)} - \bar{x}^{(n-1)}, \bar{x}^{(n)} - \bar{x}^{(n)})}{(\bar{x}^{(n-1)} - \bar{x}^{(n-2)}, \bar{x}^{(n-1)} - \bar{x}^{(n-2)})}.$$

Отметим, что реальная погрешность может оказаться значительно больше значения  $\|\bar{x}^{(n)} - \bar{x}^{(n-1)}\|$ , если  $\mu_1$  близко к 1.

*Замечание 1.* При расчете с помощью итерационных методов приближенного решения СЛАУ с точностью  $\varepsilon$  разумно требовать выполнения условия

$$\|\bar{v}^{(n)}\| \leq \varepsilon_1, \quad \text{где} \quad \varepsilon_1 = \begin{cases} \varepsilon & , \text{ если } \|x^{(0)}\| \leq 1, \\ \varepsilon\|x^{(0)}\| & , \text{ если } \|x^{(0)}\| > 1. \end{cases}$$

*Замечание 2.* При реализации метода последовательных приближений, как и любого другого итерационного процесса, надо указывать максимально допустимое число итераций. Это объясняется тем, что при решении практических задач, для некоторых из них итерационный метод может расходиться.

#### 4.7 Вопросы и упражнения для самопроверки

1. Каковы различия между точными и итерационными методами решения СЛАУ?

2. Сформулируйте достаточное условие сходимости двухслойного итерационного метода решения СЛАУ.

3. Сформулируйте необходимое и достаточное условие сходимости двухслойного итерационного метода решения СЛАУ.

4. При каких  $\alpha$  и  $\beta$  сходится метод простой итерации  $\bar{x}^{(k+1)} = S \bar{x}^{(k)} + \bar{c}$ , где

$$S = \begin{pmatrix} \alpha & \beta & 0 \\ \beta & \alpha & \beta \\ 0 & \beta & \alpha \end{pmatrix}?$$

5. Показать, что для СЛАУ  $A\bar{x} = \bar{b}$  с матрицей

$$A = \begin{pmatrix} 2 & 0.3 & 0.5 \\ 0.1 & 3 & 0.4 \\ 0.1 & 0.1 & 4.8 \end{pmatrix},$$

сходится метод простой итерации, если у матрицы перехода  $S = E - \tau A$  параметр  $0 < \tau < 0.4$ .

6. Как влияет на сходимость итерационных методов решения СЛАУ обусловленность матрицы?

7. При каких  $\alpha$  и  $\beta$  выполняется необходимое и достаточное условие сходимости метода Якоби для решения СЛАУ из пункта 4?

8. При каких  $\alpha$  и  $\beta$  выполняется необходимое и достаточное условие сходимости метода Зейделя для решения СЛАУ из пункта 4?

9. Переставьте в системе из пункта 4 первое и второе уравнения. Определите, при каких  $\alpha$  и  $\beta$  выполняется необходимое и достаточное условие сходимости метода Зейделя? Сравните полученный ответ с ответом в пункте 8.

10. Доказать, что для СЛАУ второго порядка ( $n = 2$ ) методы Якоби и Зейделя сходятся и расходятся одновременно.

11. В чем идея методов спуска при решении СЛАУ?

## 5 Приложение

В этом приложении приводятся некоторые сведения, которые используются в основном тексте, а также указания для решения примеров.

### 5.0.1 Необходимые сведения из линейной алгебры и теории матриц

Пусть  $A = \{a_{ij}\}_{i,j=1}^m$  – квадратная матрица размера  $(m \times m)$ , а  $\bar{x} = \{x_i\}_{i=1}^m$ . Обозначим сопряженную к  $A$  матрицу через  $A^*$ . Если  $A$  – вещественная матрица, то  $A^* = A^T$ , где  $A^T$  – транспонированная к  $A$  матрица. Если  $A$  комплексная матрица, то для элементов  $A^*$  справедливо равенство  $a_{ij}^* = \bar{a}_{ji}$  ( $\bar{a}_{ji}$  комплексно сопряжено с  $a_{ji}$ ). Определитель матрицы  $A$  будем обозначать  $\det(A)$ , а обратную к  $A$  матрицу через  $A^{-1}$ . Если существует  $A^{-1}$ , то матрица  $A$  называется невырожденной.

Эквивалентны следующие утверждения.

1. Матрица  $A$  является невырожденной.
2. Определитель  $\det(A) \neq 0$ .
3. СЛАУ  $A\bar{x} = 0$  имеет только тривиальное решение.
4. При любом векторе  $\bar{b} \neq 0$  СЛАУ  $A\bar{x} = \bar{b}$  имеет единственное решение.
5. Все столбцы (строки) матрицы  $A$  линейно независимы.
6. Ранг матрицы  $A$  равен  $m$  (ранг матрицы – количество линейно независимых столбцов (или строк)).

Вещественное или комплексное число  $\lambda$  и вектор  $\bar{\varphi} \neq 0$  называются соответственно собственным значением (числом) и собственным вектором матрицы  $A$ , если

$$A\bar{\varphi} = \lambda\bar{\varphi}.$$

Собственные числа матрицы  $A$  являются корнями характеристического многочлена степени  $m$

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

Множество всех собственных чисел матрицы  $A$  называется ее спектром, а

величина

$$\rho(A) = \max_i |\lambda_i|$$

называется спектральным радиусом матрицы  $A$ .

Сумма элементов квадратной матрицы, стоящих на главной диагонали, называется следом матрицы и обозначается  $tr A$  или  $Sp A$  ( $tr A = \sum_{i=1}^m a_{ii}$ ). След матрицы  $A$  совпадает с суммой корней характеристического многочлена матрицы  $A$ .

Вычисление собственных значений является, вообще говоря, сложной задачей, которая легко решается, если матрица  $A$  правая (или левая) треугольная матрица. Для таких матриц собственные числа являются диагональными элементами. Если матрица  $A$  диагональная, то ее собственные числа являются также диагональными элементами, и собственному значению, равному  $a_{jj}$ , отвечает собственный вектор  $\bar{e}_j = (0, \dots, 0, \underbrace{1}_{j\text{место}}, 0, \dots, 0)$ .

Перечислим некоторые свойства собственных чисел и собственных векторов.

Если собственные числа  $\lambda_k$  и  $\lambda_j$  матрицы  $A$  различны ( $\lambda_k \neq \lambda_j$ ), то соответствующие им собственные векторы линейно независимы. Если все собственные значения  $\lambda_k$  простые, то собственные векторы  $\varphi_k$  ( $k = 1, \dots, m$ ) линейно независимы, следовательно, образуют базис в  $R^m$ . При наличии кратных собственных чисел система всех собственных векторов может не быть базисом в  $R^m$ . Матрица, у которой полная система собственных векторов образует базис в  $R^m$ , называется матрицей простой природы (структуры), её жорданова форма диагональна.

Собственные значения  $\bar{\lambda}_k$  сопряженной матрицы  $A^*$  комплексно сопряжены числам  $\lambda_k$ , а собственные векторы  $\psi_k$  ( $k = 1, \dots, m$ ) ортогональны векторам  $\varphi_k$  ( $k = 1, \dots, m$ ), т.е.

$$(\varphi_k, \psi_j) = 0, \quad \text{если } k \neq j.$$

Собственные числа симметричной матрицы  $A$  вещественны, а собствен-

ные векторы, соответствующие различным собственным значениям, ортогональны.

Две квадратные матрицы  $A$  и  $B$  одинаковых размеров называются подобными, если существует такая невырожденная матрица  $P$ , что  $B = P^{-1}AP$ . Все подобные матрицы имеют одни и те же собственные значения. Если  $P$  – ортогональная матрица и  $B = P^*AP$ , то матрицы  $A$  и  $B$  называются ортогонально подобными.

*Неравенство Коши-Буняковского* имеет вид

$$\sum_{k=1}^m x_k y_k \leq \sqrt{\sum_{k=1}^m x_k^2} \sqrt{\sum_{k=1}^m y_k^2}, \quad \text{или} \quad (\bar{x}, \bar{y}) \leq \|\bar{x}\| \|\bar{y}\|. \quad (5.1)$$

### 5.0.2 Градиент функционала

Будем некоторый функционал  $F : R^m \rightarrow R$  рассматривать как функцию многих переменных  $F(x_1, \dots, x_m)$ . Возьмем произвольный единичный вектор  $\bar{y} = (y_1, \dots, y_m)$  ( $\|\bar{y}\| = 1$ ) и рассмотрим скорость изменения  $F(\bar{x})$  по направлению  $\bar{y}$ . Производной функционала  $F(\bar{x})$  в точке  $x = (x_1, \dots, x_m)$  по направлению  $\bar{y}$  называется

$$\begin{aligned} \frac{\partial F(\bar{x})}{\partial \bar{y}} &= \lim_{t \rightarrow 0} \frac{F(x_1 + ty_1, \dots, x_m + ty_m) - F(\bar{x})}{t} = \\ &= \left( \frac{\partial}{\partial t} F(x_1 + ty_1, \dots, x_m + ty_m) \right) \Big|_{t=0}. \end{aligned}$$

Вычислим производную

$$\frac{\partial}{\partial t} F(x_1 + ty_1, \dots, x_m + ty_m) = \frac{\partial F}{\partial x_1} y_1 + \frac{\partial F}{\partial x_2} y_2 + \dots + \frac{\partial F}{\partial x_m} y_m = (\text{grad } F, \bar{y}),$$

где вектор  $\text{grad } F = \left( \frac{\partial F}{\partial x_1}, \frac{\partial F}{\partial x_2}, \dots, \frac{\partial F}{\partial x_m} \right)$ . Теперь

$$\frac{\partial F(\bar{x})}{\partial \bar{y}} = (\text{grad } F, \bar{y}) = \|\text{grad } F\| \|\bar{y}\| \cos \alpha = \|\text{grad } F\| \cos \alpha,$$

где  $\alpha$  угол между направлением  $\text{grad } F$  и вектором  $\bar{y}$ .

Это означает, что  $-\|\text{grad } F\| \leq \frac{\partial F(\bar{x})}{\partial \bar{y}} \leq \|\text{grad } F\|$ , то есть  $F(\bar{x})$  убывает быстрее всего вдоль направления противоположного градиенту. Заметим только, что это свойство локально (справедливо только в достаточно малой окрестности точки  $x$ ). Направление градиента совпадает с направлением внешней нормали к поверхности  $F(\bar{x}) = c$ .

*Замечание.* Уравнение нормали поверхности  $F(\bar{x}) = c$  в точке  $x^0 = (x_1^0, x_2^0, \dots, x_m^0)$  имеет вид

$$\frac{x - x_1^0}{\frac{\partial F}{\partial x_1}(x^0)} = \frac{x - x_2^0}{\frac{\partial F}{\partial x_2}(x^0)} = \dots = \frac{x - x_m^0}{\frac{\partial F}{\partial x_m}(x^0)}.$$

### 5.0.3 Ответы, указания, решения

#### Раздел 2.

1. Функционал  $\max_{1 \leq i, j \leq m} |a_{ij}|$  не является нормой матрицы, так как он не удовлетворяет четвертому свойству из определения нормы матрицы. Например, для единичной матрицы  $E$  получается  $\|E E\| > \|E\| \|E\|$ . Можно привести пример матрицы второго порядка.

3. Первые три условия из определения нормы  $S(A) = \sqrt{\sum_{i,j=1}^m |a_{ij}|^2}$  очевидны, так как эту норму можно рассматривать как норму вектора  $m^2$ -мерного арифметического пространства. Четвертое свойство получается с помощью неравенства Коши-Буняковского (см. 5.1), действительно

$$\begin{aligned} S^2(AB) &= \sum_{i,j=1}^m \left| \sum_{k=1}^m a_{ik} b_{kj} \right|^2 \leq \sum_{i,j=1}^m \sum_{k=1}^m |a_{ik}|^2 \sum_{k=1}^m |b_{kj}|^2 = \\ &= \sum_{i=1}^m \sum_{k=1}^m |a_{ik}|^2 \sum_{j=1}^m \sum_{k=1}^m |b_{kj}|^2 = S^2(A) S^2(B) \end{aligned}$$

Согласованность  $S(A)$  и  $\|x\|_2$  следует из неравенства

$$\|Ax\|_2 = \sum_{i=1}^m \left| \sum_{j=1}^m a_{ij} x_j \right|^2 \leq \sum_{i,k=1}^m |a_{ik}|^2 \sum_{k=1}^m |x_k|^2 = S(A) \|x\|_2^2$$

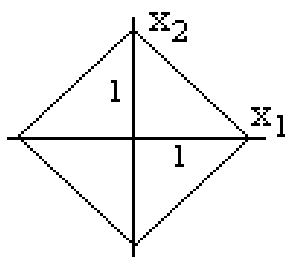


4. Доказательство неравенств (2.2) достаточно простое. Докажем лишь  $\|\bar{x}\|_1/\sqrt{m} \leq \|\bar{x}\|_2$ , используя неравенства Коши-Буняковского (см. 5.1).

Чтобы оценить  $\|x\|_1$ , преобразуем

$$\begin{aligned} (|x_1| + |x_2| + \dots + |x_m|)^2 &= (|x_1| \cdot 1 + |x_2| \cdot 1 + \dots + |x_m| \cdot 1)^2 \leq \\ &\leq (|x_1|^2 + |x_2|^2 + \dots + |x_m|^2)(1^2 + 1^2 + \dots + 1^2) = \\ &= m(|x_1|^2 + |x_2|^2 + \dots + |x_m|^2) \end{aligned}$$

5. Множество векторов  $\{\bar{x} : \|\bar{x}\|_1 = 1\}$  на плоскости имеет вид.



6. Докажем подчиненность нормы матрицы  $\|A\|_2 = \sqrt{\max_{1 \leq i \leq m} (\lambda_{AA^*})_i}$  норме вектора  $\|\bar{\varphi}\| = \sqrt{(\bar{\varphi}, \bar{\varphi})}$ . Известно, что у положительной симметричной матрицы  $AA^*$  все собственные числа положительные. Пусть  $\lambda_1$  наибольшее положительное число  $AA^*$ , которому отвечает собственный вектор  $\bar{\varphi}_1$  ( $\|\bar{\varphi}_1\| = 1$ ). Тогда  $\|A\|_2 = \sqrt{\lambda_1}$ , и  $\|A\bar{\varphi}\| = \sqrt{(A\bar{\varphi}, A\bar{\varphi})} = \sqrt{(\bar{\varphi}, A^*A\bar{\varphi})} = \sqrt{\lambda_1} \sqrt{(\bar{\varphi}, \bar{\varphi})} = \sqrt{\lambda_1} \|\bar{\varphi}\|$ .

8. Нет. Примером может служить хорошо обусловленная диагональная матрица  $diag(\varepsilon, \varepsilon, \dots, \varepsilon)$  порядка  $m \times m$  ( $m \gg 1, \varepsilon \ll 1$ ).

13. Правая часть неравенства  $\frac{S(A)}{m} \leq \|A\|_2 \leq S(A)$  очевидна. Заметим, что  $S(A) = \sqrt{tr AA^*}$ . Доказательство левой части следует из следующих преобразований

$$\|A\|_2^2 = \max_{1 \leq i \leq m} \lambda_i \geq \frac{1}{m}(\lambda_1 + \dots + \lambda_m) = \frac{1}{m} tr AA^* = S^2(A)$$

Здесь  $\lambda_i$  собственное значение матрицы  $AA^*$ .

14. Чтобы вычислить число обусловленности, постройте обратную матрицу или найдите ее собственные значения.

**Раздел 3.**

$$1. L = \begin{pmatrix} 2 & 0 & 0 \\ -1 & 1 & 0 \\ 3 & 1 & -2 \end{pmatrix}, U = \begin{pmatrix} 1 & -1 & -3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$$

2. -4.

5. Сначала надо умножить матрицу  $A$  слева на матрицу перестановок  $P_{23}$ , а затем выполнить одно преобразование вращения. В зависимости от организации вычислений в ответе последние знаки могут различаться

$$Q = \begin{pmatrix} 0.447 & -0.894 & 0 \\ 0 & 0 & 1 \\ 0.894 & 0.447 & 0 \end{pmatrix}, R = \begin{pmatrix} 2.24 & 3.58 & 4.92 \\ 0 & -0.447 & -0.894 \\ 0 & 0 & 1 \end{pmatrix}$$

7. Точное решение  $\bar{x}^* = (-1, -3, 2)^T$ .

**Раздел 4.**

4. Собственные числа матрицы перехода  $S$  должны удовлетворять теореме о необходимом и достаточном условии сходимости двухслойного итерационного метода, т.е.  $|\alpha| < 1$  и  $|\alpha \pm \sqrt{2} \beta| < 1$ .

8. Необходимое и достаточное условие сходимости метода Зейделя для заданной системы имеет вид  $|\lambda| < 1$ , где  $\lambda$  решение характеристическое уравнение

$$\begin{vmatrix} a \lambda & b \lambda & 0 \\ b \lambda & a \lambda & \beta \\ 0 & b \lambda & a \lambda \end{vmatrix} = a \lambda^2 (2\beta^2 \lambda^2 - a^2) = 0.$$

Отсюда следует, что  $\alpha$  и  $\beta$  должны удовлетворять неравенству  $|\beta/\alpha| < 1/\sqrt{2}$ .

## Список литературы

- [1] Бахвалов Н.С. Численные методы. М.: Наука, 1973.
- [2] Воеводин В.В. Вычислительные основы линейной алгебры. - М. : Наука, 1977.
- [16] Коновалов А.Н. Введение в вычислительные методы линейной алгебры. Учебное пособие. НГУ, 1983, Новосибирск.
- [4] Бахвалов Н.С., Жидков Н.П., Кобельков Г.М. Численные методы. М.: Наука, 1989г.
- [5] Калиткин Н.Н. Численные методы. М.: Наука, 1978
- [6] Самарский А.А., Гулин А.В. Численные методы. М.: Наука, 1989.
- [7] Вержбицкий В.М. Основы численных методов. - М.: Высш.шк., 2002.
- [8] Дж. Ортега, У. Пул Введение в численные методы решения дифференциальных уравнений. М.: Наука. Гл. ред. физ.-мат. лит., 1986.-288с.
- [9] Бахвалов Н.С., Лапин А.В., Чижонков Е.В. Численные методы в задачах и упражнениях. М., "Высшая школа ", 2000.
- [10] Крылов В.И., Бобков В.В. Вычислительные методы. Т. II, М.: Наука, 1977
- [11] Фаддеев Д.К., Фаддеев В.Н. Вычислительные методы линейной алгебры. М. Физматгиз, 1963.
- [12] Дж. Х Уилкинсон. Алгебраическая проблема собственных значений. М., Наука, 1970г.
- [13] Уилкинсон Дж., Райнш Л. Справочник алгоритмов на языке АЛГОЛ. Линейная алгебра.- М., Машиностроение, 1976г.
- [14] Форсайт Дж., Малькольм М., Моулер К. Машинные методы математических вычислений. - М., Мир, 1980г.
- [15] Х.Д.Икрамов Численное решение матричных уравнений. - М., "Наука", 1984г.
- [16] Дробышев В.И., Дымников В.П., Ривин Г.С. Задачи по вычислительной математики. - М., "Наука", Главная редакция физ-мат. литературы, 1980.
- [17] Б.П. Демидович, И.Л. Марон Основы вычислительной математики. - М., Наука", 1966.
- [18] С.Н. Овчинникова, Численные методы. Часть I. Численные методы математического анализа. Курс лекций. ЮФУ, 2008г.