

**МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РФ
Федеральное государственное бюджетное
образовательное учреждение
высшего профессионального образования
ПЕНЗЕНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ**

Н. Ю. Кудряшова, Н. В. Печникова

ДОПОЛНИТЕЛЬНЫЕ ГЛАВЫ АЛГЕБРЫ

Учебное пособие

ПЕНЗА 2012

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РФ
Федеральное государственное бюджетное
образовательное учреждение
высшего профессионального образования
«Пензенский государственный университет» (ПГУ)

Н. Ю. Кудряшова, Н. В. Печникова

ДОПОЛНИТЕЛЬНЫЕ ГЛАВЫ АЛГЕБРЫ

Учебное пособие

Пенза
Издательство ПГУ
2012

Р е ц е н з е н т ы:

доктор технических наук, профессор,
заведующий кафедрой «Математика и математическое моделирование»
Пензенского государственного университета
архитектуры и строительства
А.М. Данилов

доктор физико-математических наук,
профессор, заведующий кафедрой «Физика»
Пензенского государственного университета
В.Д. Кревчик

Кудряшова, Н.Ю.

Дополнительные главы алгебры: учебное пособие /
Н.Ю.Кудряшова, Н.В.Печникова. – Пенза: Изд-во ПГУ, 2012. – 138 с.

Излагаются основные численные методы решения систем линейных алгебраических уравнений и проблемы собственных значений. Рассматриваются основные алгебраические структуры, такие как группы, кольца и поля. Приводятся методические указания для выполнения лабораторных работ.

Учебное пособие подготовлено на кафедре «Высшая и прикладная математика» и предназначено для студентов, обучающихся по направлению 231300 «Прикладная математика» (профиль подготовки «Математическое моделирование в экономике и технике»).

УДК

ПРЕДИСЛОВИЕ

Предлагаемое учебное пособие предназначено, главным образом, для студентов, обучающихся по направлению 231300 «Прикладная математика» (профиль подготовки «Математическое моделирование в экономике и технике») при изучении курса «Дополнительные главы алгебры». В учебное пособие включены основные теоретические сведения по изучению численных методов линейной алгебры и основных алгебраических структур, а также предлагается лабораторный практикум по основным численным методам линейной алгебры.

Первая глава учебного пособия посвящена рассмотрению основных численных методов решения систем линейных алгебраических уравнений и проблемы собственных значений.

К численным методам алгебры относятся численные методы решения систем линейных уравнений, обращения матриц, вычисления определителей, нахождения собственных векторов и собственных значений. При формальном подходе решение этих задач не встречает затруднений: решение системы можно найти, раскрыв определители в формуле Крамера; для нахождения собственных значений матрицы достаточно выписать характеристическое уравнение и найти его корни. Однако на практике, как правило, указанные рекомендации встречают ряд затруднений. Так, например, при непосредственном раскрытии определителей решение системы с n неизвестными требует порядка $n!/n$ арифметических операций. Другой причиной, по которой эти классические способы неприменимы даже при малых n , является сильное влияние округлений при вычислениях на окончательный результат. Точно также обстоит дело при нахождении собственных значений матрицы с использованием явного выражения характеристического многочлена.

Методы решения алгебраических задач разделяются на точные и приближенные (или итерационные). Под точными методами понимаются методы, которые дают решение задачи при помощи конечного числа элементарных арифметических операций. При этом, если исходные данные заданы точно и вычисления выполняются точно, то решение также получается точное. Приближенные методы являются средством нахождения приближенного решения. В этом случае решение получается как предел последовательных приближений, вычисляемых некоторым единообразным процессом. При применении итерационных методов существенным является не только сходимость построенных последовательных приближений, но и быстрота сходимости. В этом отношении каждый итерационный метод не является универсальным: давая быструю сходимость для одних матриц, он может сходиться медленно или даже совсем не сходиться для других. Поэтому при применении итерационных методов важную роль играет предварительная подготовка системы, т.е. замена данной системы ей эквивалентной, устроенной так, чтобы для нее выбранный процесс сходил по возможности быстро.

Во второй главе пособия изучению подвергаются немногие, наиболее важные типы алгебраических систем, т. е. множеств, составленных из элементов какой-либо природы, для которых определены некоторые алгебраические операции. Таковы, в частности, поля, кольца и группы. Теория полей оказалась естественной областью для дальнейшего развития теории уравнений, а ее основные ветви — теория полей алгебраических чисел и теория полей алгебраических функций — связали ее соответственно с теорией чисел и теорией функций комплексного переменного. Более широким, чем понятие поля, является понятие кольца. Простейшими примерами колец служат совокупность всех целых чисел (включая и отрицательные), система многочленов от одного неизвестного и система действительных функций действительного переменного. Теория колец включает в себя такие старые ветви алгебры, как теория гиперкомплексных систем и теория идеалов, она связана с рядом математических наук, в частности с функциональным анализом, и уже нашла некоторые выходы в физику. Еще большую область применений имеет теория групп. Группы играли большую роль уже в теории Галуа, в вопросе о разрешимости уравнений в радикалах, сейчас же они являются важным орудием в теории полей, во многих разделах геометрии, в топологии, а также и в кристаллографии, в теоретической физике. Вообще, по широте области приложений теория групп занимает среди всех ветвей алгебры следующее после линейной алгебры место.

В третьей главе предлагается лабораторный практикум. При выполнении лабораторных работ было бы интересно сравнить результаты, получаемые для решаемых задач при использовании стандартных математических пакетов, с результатами, полученными по изучаемым в данном курсе численными методами.

ОСНОВНЫЕ ПОНЯТИЯ

НОРМЫ ВЕКТОРОВ И МАТРИЦ

Нормой вектора X с координатами $(x_1, x_2, \dots, x_n)$ называется неотрицательное число $\|X\|$, удовлетворяющее следующим требованиям:

1. $\|X\| > 0$ при $X \neq 0$ и $\|X\| = 0$ при $X = 0$.
2. $\|CX\| = |C| \cdot \|X\|$ при $\forall C$.
3. $\|X + Y\| \leq \|X\| + \|Y\|$ – неравенство треугольника.

$$\|X\| = \sqrt{(X, X)}; (X, Y) = \sum_{i=1}^n x_i y_i$$

Будем использовать три основные нормы:

I. $\|x\|_I = \max |x_i|$ – кубическая

II. $\|x\|_{II} = \sum_{i=1}^n |x_i| = |x_1| + |x_2| + \dots + |x_n|$ – октаэдрическая

III. $\|x\|_{III} = \sqrt{\sum_{i=1}^n |x_i|^2}$ – сферическая

Нормой квадратной матрицы $A = (a_{ij})$, $i = \overline{1, m}$, $j = \overline{1, n}$ называется неотрицательное число $\|A\|$, удовлетворяющее условиям:

- 1) $\|A\| \geq 0$ и $\|A\| = 0 \Leftrightarrow A = 0$.
- 2) $\|cA\| = |c| \cdot \|A\|$.
- 3) $\|A + B\| \leq \|A\| + \|B\|$.
- 4) $\|A \cdot B\| \leq \|A\| \cdot \|B\|$.

Обозначение диагональной матрицы $D = \begin{pmatrix} d_{11} & 0 & \dots & 0 \\ 0 & d_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & d_{nn} \end{pmatrix} = [d_{11}, d_{22}, \dots, d_{nn}]$.

Матрица называется положительно определенной, если для $\forall x$ $(Ax, x) \geq 0$ и $(Ax, x) = 0 \Leftrightarrow x = 0$.

Будем рассматривать три основные нормы:

$$1. \|A\|_I = \sqrt{\sum_{i,j=1}^n |a_{ij}|^2}; \quad 2. \|A\|_{II} = \max_i \sum_{j=1}^n |a_{ij}|; \quad 3. \|A\|_{III} = \max_j \sum_{i=1}^n |a_{ij}|.$$

Глава I. ЧИСЛЕННЫЕ МЕТОДЫ ЛИНЕЙНОЙ АЛГЕБРЫ

1. Погрешности вычислений. Обусловленность матриц. Точные методы решения систем линейных алгебраических уравнений (СЛАУ)

Погрешность результата численного решения задачи. Отклонение истинного решения от приближенного назовем погрешностью.

Решение задач всегда имеют погрешность, связанную со следующими причинами:

1) созданием математической модели (любая модель имеет свою степень точности);

2) получением исходных данных (т.к. начальные данные являются «результатом измерений», следовательно, возникают измерительные погрешности);

3) использованием вычислительной техники (ошибки округления, возникающие из-за ограниченной разрядной сетки и ошибки, связанные с самими методами).

Неустраняемая погрешность: погрешности исходных данных; погрешности математической модели.

Полная погрешность: неустраняемая погрешность; погрешность метода; вычислительная погрешность.

Неустраняемую погрешность и погрешность метода необходимо контролировать, чтобы не осуществлять расчеты с избыточной точностью.

Характеристиками точности результата решения задачи являются абсолютная и относительная погрешности.

Если x – точное значение некоторого числа, x^* – приближенное, то **абсолютной погрешностью** приближения x^* назовем величину: $\Delta x^* \geq |x - x^*|$, т.е. точное значение числа x заключено в границах $x^* - \Delta x^* \leq x \leq x^* + \Delta x^*$.

Отношение абсолютной погрешности к абсолютному значению приближенной величины есть **относительная погрешность** (т.е. доля истинного значения): $\delta x^* = \frac{\Delta x^*}{|x^*|}$, при условии, что $|x^*| \neq 0$.

Обусловленность матриц. При численном решении систем линейных уравнений даже точными методами возникает несколько источников неточности решения. Одним из них является необходимость округления чисел в процессе вычисления. При этом может случиться, что по ходу вычисления придется столкнуться с явлением исчезновения значащих цифр в результате вычитания двух величин, близких друг к другу. Исчезновение значащих цифр может быть причиной настолько значительного снижения точности результата, что из-за этого иногда бывает необходимо изменить схему вычисления или переделать работу с большим числом значащих цифр в промежуточных выкладках.

Второй источник появляется в условиях, когда система линейных уравнений возникает в процессе решения практической задачи, так что элементы матрицы коэффициентов, так же как и свободные члены, известны лишь приближенно, с некоторой степенью точности. Неточность самих исходных данных порождает ошибки в решении, так как изменение самих коэффициентов системы в пределах заданной точности влечет за собой изменение решения. Займемся более детально исследованием этой причины неточности в решении системы.

Теоретическое решение системы $AX = F$ дается формулой $X = A^{-1}F$, где A^{-1} матрица, обратная к A . Как известно, A^{-1} существует в том и только том случае, если $|A| \neq 0$. Однако, если элементы матрицы A заданы приближенно, возможно, что даже сам вопрос о том, имеет ли матрица A отличный от нуля определитель или нет, лишен смысла. Именно, может случиться, что при точном вычислении определителя, исходя из приближенных значений элементов матрицы, принятых за точные, определитель оказывается

отличным от нуля, но изменение элементов в пределах точности их задания может привести к матрице с нулевым определителем. Ясно, что система с матрицей, обладающей указанным свойством, не может быть решена со сколько-нибудь удовлетворительной точностью. Система практически оказывается несовместной.

Мы будем называть обратную матрицу *устойчивой*, если малым изменениям в элементах матрицы будут соответствовать малые изменения в элементах обратной матрицы.

Очевидно, что для устойчивости обратной матрицы, во всяком случае, необходимо, чтобы определитель матрицы был не слишком мал. Степень малости определителя, препятствующую получению устойчивой обратной матрицы, трудно определить раз и навсегда, ибо само понятие «не слишком мал» вряд ли может быть определено.

Будем называть матрицу *плохо обусловленной*, если соответствующая ей обратная матрица будет неустойчивой.

Плохо обусловленная матрица может оказаться практически особенной, если ее элементы заданы приближенно.

Естественно, что система линейных уравнений с плохо обусловленной матрицей мало устойчива, т. е. ее решение сильно изменяется при малых изменениях как коэффициентов, так и свободных членов.

Для характеристики матрицы с точки зрения её обусловленности несколькими авторами предложены различные количественные характеристики, так называемые *числами обусловленности*. Во-первых, это два числа Тюринга:

$$1) \ N - \text{число} : \frac{1}{n} N(A) N(A^{-1}) = \nu(A),$$

$$2) \ M - \text{число} : \frac{1}{n} M(A) M(A^{-1}) = \mu(A),$$

где $N(A) = \sqrt{\text{Tr} A^T A}$ и $M(A) = n \cdot \max |a_{ij}|$, A^T - транспонированная к матрице

A , $\text{Tr} A = a_{11} + \dots + a_{nn}$ - след матрицы.

$$3) \text{ Число Тодда } P - \text{число} : \frac{\max |\lambda_i|}{\min |\lambda_i|} = \rho(A),$$

где λ_i – собственные значения матрицы A , и

$$4) H - \text{число} : \sqrt{\frac{\mu_1}{\mu_n}} = \eta(A),$$

где μ_1 и μ_n наибольшее и наименьшее собственные значения матрицы $A^T A$,

где A^T –транспонированная с A матрица.

Основные типы методов решения СЛАУ. В настоящее время имеется большое количество методов численного решения систем линейных уравнений, и работа над их усовершенствованием интенсивно ведется в наши дни.

Численные методы, решающие указанные задачи, делятся на две группы: точные и итерационные методы. Под **точными методами** мы подразумеваем методы, которые дают решение задачи при помощи конечного числа элементарных арифметических операций. При этом, если исходные данные, определяющие задачу, заданы точно (например, если они целые или рациональные числа, представленные в виде обыкновенных дробей) и вычисления выполняются точно (например, по правилам действий над обыкновенными дробями), то решение также получается точное. В точных методах число необходимых для решения задачи вычислительных операций зависит только от вида вычислительной схемы и от порядка матрицы, определяющей данную задачу.

Исторически первым методом этой группы является метод, основанный на идее исключения неизвестных. В применении к решению линейной неоднородной системы алгоритм этого метода, связанный с именем Гаусса, состоит из цепочки последовательных исключений, посредством которых данная система превращается в систему с треугольной матрицей, решение которой не составляет труда.

В настоящее время разработано много различных вычислительных схем метода Гаусса. Среди этих схем следует отметить так называемые компактные схемы, требующие минимальной записи или запоминания промежуточных результатов.

Близким к методу Гаусса является метод, основанный на идее окаймления. В основе этого метода лежит представление данной матрицы в виде последовательно вложенных друг в друга матриц, начиная с матрицы первого порядка.

Различные модификации метода исключения, так же как и метод окаймления, по существу связаны с разложением матрицы в произведение двух треугольных матриц.

К точным методам обращения матриц относится и метод, основанный на идее постепенного изменения обрабатываемой матрицы с внесением последовательно поправок к обратной (метод пополнения).

В последние годы получили большое распространение точные методы, основанные на идее построения вспомогательной системы векторов, ортогональных в той или иной метрике.

Итерационные методы представляют собой средство для приближенного решения системы линейных уравнений. Решение системы при помощи итерационных методов получается как предел последовательных приближений, вычисляемых некоторым единообразным процессом. При применении итерационных методов существенным является не только сходимость построенных последовательных приближений, но и быстрота сходимости. В этом отношении каждый итерационный метод не является универсальным: давая быструю сходимость для одних матриц, он может сходиться медленно или даже расходиться для других матриц. Поэтому при применении итерационных методов важную роль играет предварительная подготовка системы, т. е. замена данной системы ей эквивалентной, устроенной так, чтобы для нее выбранный процесс сходил по возможности быстро, а также различные приемы ускорения.

2. Метод Гаусса

Метод Гаусса основан на последовательном исключении неизвестных и имеет много различных вычислительных схем.

1. Схема единственного деления

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = f_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = f_2, \\ \dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = f_n \end{cases} \quad (2.1)$$

Пусть $a_{11} \neq 0$. Разделим первое уравнение на a_{11} (a_{11} – ведущий элемент первого шага).

Обозначим

$$b_{1j} = \frac{a_{1j}}{a_{11}} \quad (j > 1), \quad g_1 = \frac{f_1}{a_{11}} \quad (2.2)$$

Получим новое уравнение

$$x_1 + b_{12}x_2 + \dots + b_{1n}x_n = g_1 \quad (2.3)$$

Исключим x_1 из всех уравнений (2.1), начиная со второго, посредством вычитания из этих уравнений уравнения (2.3), умноженного соответственно на $a_{21}, a_{31}, \dots, a_{n1}$.

Преобразованная система примет вид:

$$\begin{cases} a_{22}^{(1)}x_2 + \dots + a_{2n}^{(1)}x_n = f_2^{(1)}, \\ \dots \\ a_{n2}^{(1)}x_2 + \dots + a_{nn}^{(1)}x_n = f_n^{(1)} \end{cases} \quad (2.4)$$

где

$$a_{ij}^{(1)} = a_{ij} - a_{i1}b_{1j}, \quad (i, j \geq 2); \quad f_i^{(1)} = f_i - a_{i1}g_1 \quad (2.5)$$

Далее разделим коэффициенты первого из преобразованных уравнений системы на ведущие элемент второго шага $a_{22}^{(1)}$, который будем считать $a_{22}^{(1)} \neq 0$. Получим уравнение:

$$x_2 + b_{23}x_3 + \dots + b_{2n}x_n = g_2, \text{ где } b_{2j} = \frac{a_{2j}^{(1)}}{a_{22}^{(1)}}, \quad g_2 = \frac{f_2^{(1)}}{a_{22}^{(1)}}.$$

Исключая x_2 из системы (2.4), начиная со второго уравнения, мы
придем к уравнениям:

$$\begin{cases} a_{33}^{(2)} x_3 + \dots + a_{3n}^{(2)} x_n = f_3^{(2)}, \\ \dots\dots\dots, \text{ где} \\ a_{n3}^{(2)} x_3 + \dots + a_{nn}^{(2)} x_n = f_n^{(2)} \end{cases} \quad \begin{cases} a_{ij}^{(2)} = a_{ij}^{(1)} - a_{i2}^{(2)} b_{2j} \\ f_i^{(2)} = f_i^{(1)} - a_{i2}^{(1)} g_2 \end{cases} \quad (i, j \geq 3)$$

Продолжаем процесс по той же схеме. На m -том шаге мы получим уравнения:

$$\left\{ \begin{array}{l} x_m + b_{m,m+1}x_{m+1} + ... + b_{mn}x_n = g_m, \\ a^{(m)}_{m+1,m+1}x_{m+1} + ... + a^{(m)}_{m+1,n}x_n = f^{(m)}_{m+1}, \\ \\ a^{(m)}_{n,m+1}x_{m+1} + ... + a^{(m)}_{nn}x_n = f^{(m)}_n \end{array} \right. \quad (2.6)$$

где:

$$\begin{aligned} b_{mj} &= \frac{a_{mj}^{(m-1)}}{a_{mm}^{(m-1)}}, \quad j \geq m+1; & g_m &= \frac{f_m^{(m-1)}}{a_{mm}^{(m-1)}}; \\ a_{ij}^{(m)} &= a_{ij}^{(m-1)} - a_{jm}^{(m-1)} b_{mj}; & f_i^{(m)} &= f_i^{(m-1)} - a_{im}^{(m-1)} g_m \end{aligned} \quad (2.7)$$

Объединив все первые уравнения каждого шага, получим систему

$$\left\{ \begin{array}{l} x_1 + b_{12}x_2 + b_{13}x_3 + \dots + b_{1n}x_n = g_1, \\ \quad x_2 + b_{23}x_3 + \dots + b_{2n}x_n = g_2, \\ \quad \quad \quad \dots \\ \quad \quad \quad \quad x_n = g_n \end{array} \right. \quad (2.8)$$

с треугольной матрицей, эквивалентную исходной системе.

Из системы (2.8) значения для неизвестных находим последовательно от x_n к x_1 по формулам:

$$x_m = g_m - b_{m,m+1}x_{m+1} - \dots - b_{mn}x_n, \quad m = n-1, \dots, 1 \quad (2.9)$$

Процесс нахождения коэффициентов треугольной системы мы будем называть прямым ходом, а процесс получения ее решения – обратным.

Схема единственного деления не является универсальной в том смысле, что для ее применения нужно, чтобы все ведущие элементы были не равны нулю. Близость к нулю ведущих элементов может быть причиной значительной потери точности.

Наилучшим вариантом является схема единственного деления по главным элементам. В этой схеме в качестве исключаемой на каждом шаге неизвестной выбирается та, коэффициент при которой на предыдущем шаге был наибольшим по модулю. При вычислении по схеме главных элементов исчезновение значащих цифр может происходить, только если система плохо обусловлена, так что происходящая при этом потеря точности неизбежна.

Последовательное исключение неизвестных, преобразующее данную систему в систему с треугольной матрицей, можно проводить и по другим вычислительным схемам.

3. Компактные схемы для решения неоднородных линейных систем

Решение системы линейных уравнений по схеме единственного деления свелось к определению коэффициентов $a_{ij}^{(k)}$ преобразованных уравнений и коэффициентов b_{ij} уравнений окончательной треугольной системы. При этом для получения решения данной системы нам нужны только коэффициенты b_{ij} , а числа $a_{ij}^{(k)}$ играют вспомогательную роль, и нужны лишь для определения b_{ij} .

Покажем, что числа b_{ij} можно получить процессом накопления, позволяющим избежать вычисления и записи всех коэффициентов $a_{ij}^{(k)}$.

Выделим элементы первого столбца каждой вспомогательной матрицы $a_{ij}^{(j-1)}$, $i \geq j$, обозначим их через c_{ij} , $i \geq j$, $c_{ij} = a_{ij}^{(j-1)}$, $i \geq j$.

Анализируя процесс вычисления коэффициентов вспомогательных матриц, получим, что

$$\begin{aligned} a_{ij}^{(k)} &= a_{ij}^{(k-1)} - a_{ik}^{(k-1)} b_{kj} = a_{ij}^{(k-1)} - c_{ik} b_{kj} = a_{ij}^{(k-2)} - c_{i,k-1} b_{k-1,j} - c_{ik} b_{kj} = \dots = \\ &= a_{ij} - c_{i1} b_{1j} - c_{i2} b_{2j} - \dots - c_{ik} b_{kj} = a_{ij} - \sum_{l=1}^k c_{il} b_{lj} \end{aligned} \quad (3.1)$$

Для самих элементов c_{ij} , $i \geq j$ и b_{ij} , $i < j$ имеют место рекуррентные формулы:

$$c_{ij} = a_{ij}^{(i-1)} = a_{ij} - \sum_{l=1}^{j-1} c_{il} b_{lj}, \quad (i \geq j) \quad (3.2)$$

$$b_{ij} = \frac{a_{ij}^{(i-1)}}{a_{ii}^{(i-1)}} = \frac{a_{ij} - \sum_{l=1}^{j-1} c_{il} b_{lj}}{c_{ii}}, \quad (i < j) \quad (3.3)$$

Схема проведения прямого хода по формулам (3.2) и (3.3) называется компактной.

4. Метод Гаусса-Жордана. Вычисление определителей. Обращение матриц

Метод Гаусса-Жордана. Данный метод похож на метод Гаусса с той лишь разницей, что матрица системы приводится не к треугольному виду, а к диагональному. Тогда решением системы будет преобразованный вектор правой части.

Если в результате преобразований в какой-то строке имеем на диагонали элемент, очень близкий к нулю, хотя в векторе правой части на соответствующей позиции стоит ненулевой элемент, то можно сделать вывод о несовместности системы.

Вычисление определителей. Метод Гаусса, развитый в предыдущем параграфе для решения линейной системы, может быть применен и для вычисления определителей. Мы остановимся отдельно на описании соответствующей схемы деления, так как вычисление определителей часто встречается в приложениях. Пусть

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{vmatrix}, a_{11} \neq 0.$$

Вынесем элемент a_{11} из первой строки

$$\Delta = a_{11} \begin{vmatrix} 1 & b_{12} & b_{13} & \dots & b_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{vmatrix}.$$

Затем из каждой строки вычтем первую, умноженную соответственно на первый элемент этой строки. Мы получим, очевидно,

$$\Delta = a_{11} \begin{vmatrix} 1 & b_{12} & b_{13} & \dots & b_{1n} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} & \dots & a_{2n}^{(1)} \\ 0 & a_{32}^{(1)} & a_{33}^{(1)} & \dots & a_{3n}^{(1)} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & a_{n2}^{(1)} & a_{n3}^{(1)} & \dots & a_{nn}^{(1)} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22}^{(1)} & \dots & a_{2n}^{(1)} \\ \dots & \dots & \dots \\ a_{n2}^{(1)} & \dots & a_{nn}^{(1)} \end{vmatrix}.$$

Далее с определителем $(n-1)$ -го порядка поступаем совершенно таким же образом, если только $a_{22}^{(1)} \neq 0$. Продолжая процесс, мы получим, что искомый определитель равен произведению ведущих элементов: $\Delta = a_{11} \cdot a_{22}^{(1)} \cdot \dots \cdot a_{nn}^{(1)} \neq 0$. Если на каком-то шаге окажется, что $a_{ii}^{(i-1)} = 0$ или $a_{ii}^{(i-1)} = 0$ близко к нулю (что влечет за собой уменьшение точности вычислений), можно предварительно изменить порядок строк и столбцов определителя так, чтобы в левом верхнем углу оказался неисчезающий элемент.

Связь метода Гаусса с разложением матрицы на множители. Метод Гаусса, проведенный с фиксированным порядком ведущих элементов, состоит в том, что данная система заменяется равносильной треугольной системой посредством линейного комбинирования уравнений, что сводится к линейному комбинированию строк матрицы A . При этом приходится (пользуясь схемой единственного деления) кроме делений на ведущие элементы добавлять к элементам строк числа, пропорциональные элементам предшествующих строк, т. е. совершать над матрицей A **элементарные преобразования**.

Результат нескольких преобразований этого вида, равносильен умножению матрицы A слева на левую треугольную матрицу

$$\Gamma = \begin{bmatrix} \gamma_{11} & 0 & \dots & 0 \\ \gamma_{21} & \gamma_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ \gamma_{n1} & \gamma_{n2} & \dots & \gamma_{nn} \end{bmatrix}$$

В результате этих преобразований мы переходим к системе с правой треугольной матрицей

$$B = \begin{bmatrix} 1 & b_{12} & \cdots & b_{1n} \\ 0 & 1 & \cdots & b_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

Таким образом, $\Gamma A = B$, т.е. $A = \Gamma^{-1}B$ и, следовательно, матрица A разлагается в произведение двух треугольных матриц.

Компактная схема метода Гаусса осуществляет это разложение. А именно справедливы соотношения:

$$a_{ij} = c_{ij} + \sum_{k=1}^{j-1} c_{ik} b_{kj} = \sum_{k=1}^j c_{ik} b_{kj}, \quad (i \geq j) \quad (4.1)$$

$$a_{ij} = c_{ii} b_{ij} + \sum_{k=1}^{j-1} c_{ik} b_{kj} = \sum_{k=1}^i c_{ik} b_{kj} \quad (i < j) \quad (4.2)$$

Последние формулы означают, что $A = CB$, где $C = \Gamma^{-1}$. Так как диагональные элементы матрицы B равны единице, такое разложение единственно.

Обращение матриц. Как уже говорилось в предисловии, задача решения линейной неоднородной системы и задача обращения матрицы тесно связаны друг с другом. Действительно, если для матрицы A известна её обратная матрица A^{-1} , то умножая равенство $AX = F$ на A^{-1} слева, мы получим $X = A^{-1}F$.

Обратно, определение элементов обратной матрицы можно свести к решению n систем вида $\sum_{k=1}^n a_{ik} a_{kj} = \delta_{ij}$, $i = \overline{1, n}$, $j = \overline{1, n}$, где δ_{ij} - символ Кронекера. Последнее вытекает из определения обратной матрицы $AA^{-1} = E$ и правила умножения матриц.

Теорема о разложении матрицы в произведение двух треугольных матриц дает возможность построить **компактную схему** для вычисления элементов обратной матрицы.

Пусть $A=CB$, причем элементы треугольных матриц C и B определяются по формулам:

$$c_{ij} = a_{ij} - \sum_{k=1}^{j-1} c_{ik} b_{kj}, \quad (i \geq j); \quad b_{ij} = \frac{a_{ij} - \sum_{k=1}^{j-1} c_{ik} b_{kj}}{c_{ii}}, \quad (i < j; b_{ij} = 1) \quad (4.3)$$

Обозначим элементы обратной матрицы $A^{-1}=D$ через d_{ij} . Имеем, очевидно, что $D=B^{-1}C^{-1}$ и $BD=C^{-1}$. Так как матрица C^{-1} треугольная, то $\frac{n(n-1)}{2}$ её элементов будут нулями. Легко видеть, что система, полученная объединением упомянутых выше $\frac{n(n-1)}{2}$ равенств системы $DC=B^{-1}$ и $\frac{n(n-1)}{2}$ равенств системы $BD=C^{-1}$, является рекуррентной системой, дающей возможность определить n^2 элементов обратной матрицы.

Выпишем её для $n=4$.

$i =$	1	2	3	4		$j =$	2	3	4
$c_{11}d_{i1} + c_{21}d_{i2} + c_{31}d_{i3} + c_{41}d_{i4} =$	1	0	0	0		$d_{1j} + b_{21}d_{2j} + b_{13}d_{3j} + b_{14}d_{4j} =$	0	0	0
$c_{22}d_{i2} + c_{32}d_{i3} + c_{42}d_{i4} =$		1	0	0		$d_{2j} + b_{23}d_{3j} + b_{24}d_{4j} =$			
$c_{33}d_{i3} + c_{43}d_{i4} =$						$d_{3j} + b_{34}d_{4j} =$			
$c_{44}d_{i4} =$				1					

Из уравнений первой группы определяются последовательно элементы $d_{44}, d_{43}, d_{42}, d_{41}$ при $i=4$. Затем из уравнений второй группы определяем,

d_{34}, d_{24}, d_{14} при $j=4$. Далее процесс идет аналогично, и мы пользуемся по очереди формулами первой и второй группы. Именно, из уравнений первой группы определяются d_{33}, d_{32}, d_{31} при $i=3$ и из уравнений второй группы - d_{23}, d_{13} при $j=3$; из уравнений первой группы определяются d_{22}, d_{21} при $i=2$ и из уравнений второй группы d_{12} при $j=2$. Наконец, из уравнений первой группы определяем d_{11} при $i=1$.

Уточнение обратной матрицы. Обращение матрицы по приведенной выше схеме не дает уверенности в точности полученных результатов из-за неизбежных округлений, влияние которых на конечный результат трудно оценить. Для контроля точности матрицы D_0 , полученной из матрицы A каким-либо процессом обращения, следует составить произведение AD_0 . Отклонение этого произведения от единичной матрицы указывает на степень неточности полученных результатов.

Это контрольное вычисление показывает, что приближение D_0 к A^{-1} таково, что $\|R_0\| \leq k < 1$, где $R_0 = E - AD_0$.

В качестве нормы матрицы удобно брать первую и вторую нормы, как наиболее легко вычисляемые. При этом условии элементы обратной матрицы A^{-1} могут быть вычислены со сколь угодно большой точностью при помощи следующего итерационного процесса.

Образуем последовательность матриц

$$D_1 = D_0(E + R_0) \qquad R_0 = E - AD_0$$

$$D_2 = D_1(E + R_1) \qquad R_1 = E - AD_2$$

.....

$$D_m = D_{m-1}(E + R_{m-1}) \qquad R_m = E - AD_m$$

Проверим, что $R_m = E - AD_m = R_0^{2^m}$. Действительно,

$$R_m = E - AD_m = E - AD_{m-1}(E + R_{m-1}) = E - (E - R_{m-1})(E + R_{m-1}) = R_{m-1}^2 = \dots = R_0^{2^m}$$

Отсюда следует что $D_m = A^{-1}(E - R_0^{2^m})$.

Последняя формула показывает, что D_m стремится к A^{-1} , причем сходимость процесса очень быстрая. Дадим оценку погрешности, принимая во внимание, что $A^{-1} = D_0 (AD_0)^{-1} = D_0 (E - R_0)^{-1}$:

$$\|D_m - A^{-1}\| = \|-A^{-1}R_0^{2m}\| = \|D_0(E - R_0)^{-1}R_0^{2m}\| \leq \|D_0\| \|(E - R_0)^{-1}\| \|R_0^{2m}\| \leq \|D_0\| \frac{k^{2m}}{1-k}$$

Из этой оценки видно, что как только начальное приближение выбрано так, что $\|E - AD_0\| \leq k < 1$, число верных десятичных знаков возрастает в геометрической прогрессии.

5. Метод квадратных корней. Метод ортогонализации

Метод квадратных корней. Пусть матрица A симметрическая. Тогда ее можно представить в виде произведения двух треугольных матриц. Итак, пусть $A = S^T S$, где

$$S = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1n} \\ 0 & s_{22} & \cdots & s_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & s_{nn} \end{bmatrix}.$$

Определим элементы матрицы S . Имеем, в силу правила умножения матриц,

$$a_{ij} = s_{1i}s_{1j} + s_{2i}s_{2j} + \cdots + s_{ii}s_{ij}, \quad i < j \quad (5.1)$$

$$a_{ii} = s_{1i}^2 + s_{2i}^2 + \cdots + s_{ii}^2, \quad i = j \quad (5.2)$$

Отсюда получаем формулы для определения s_{ij} :

$$\begin{aligned} s_{11} &= \sqrt{a_{11}}; \quad s_{1j} = \frac{a_{1j}}{s_{11}}; \quad s_{ii} = \sqrt{a_{ii} - \sum_{j=1}^{i-1} s_{ji}^2}, \quad (i > 1); \\ s_{ij} &= \frac{a_{ij} - \sum_{l=1}^{i-1} s_{li}s_{lj}}{s_{ii}}, \quad (j > i); \quad s_{ij} = 0, \quad (i > j) \end{aligned} \quad (5.3)$$

Решение системы сводится к решению двух треугольных систем. Действительно, равенство $AX = F$ равносильно двум равенствам $S^T K = F$ и $SX = K$.

Элементы вектора K определяются по рекуррентным формулам, аналогичным формулам для s_{ij} , т.е.

$$k_1 = \frac{f_1}{s_{11}}, \dots, k_i = \frac{f_i - \sum_{l=1}^{i-1} s_{li}k_l}{s_{ii}} \quad (i > 1). \quad (5.4)$$

Окончательное решение находится по формулам:

$$x_n = \frac{k_n}{s_{nn}}, \dots, x_i = \frac{k_i - \sum_{l=1}^{i-1} s_{li} x_l}{s_{ii}} \quad (i < n) \quad (5.5)$$

Метод ортогонализации. С решением $X = (x_1, \dots, x_n)$ системы уравнений

[illegible]

связывается вектор $Z = (x_1, \dots, x_n, 1)^T = (X, 1)^T$ арифметического $(n+1)$ -мерного пространства R_{n+1} в естественном представлении. Уравнения системы понимаются как условия ортогональности вектора Z с векторами $A_i = (a_{i1}, \dots, a_{in}, -f_i)^T$, $(i=1, 2, 3, \dots, n)$.

Решение находят следующим образом.

1. Шаг за шагом строятся базисы подпространств убывающих размерностей $R_{n+1} = R^{(0)} \supset R^{(1)} \supset \dots \supset R^{(n)}$, где $R^{(k)}$ -подпространство, состоящее из векторов, ортогональных векторам $A_1, A_2, \dots, A_k$.

2. Каждый базис $V_{k+1}^{(k)}, \dots, V_{n+1}^{(k)}$ строится из предыдущего $V_k^{(k-1)}, \dots, V_{n+1}^{(k-1)}$ в виде двучленной линейной комбинации

$$V_i^{(k)} = V_i^{(k-1)} - c_i^{(k)} V_k^{(k-1)} \quad (i = k+1, \dots, n+1). \quad (5.7)$$

3. Коэффициенты $c_i^{(k)}$ определяются из условия ортогональности к вектору A_k , что дает

$$c_i^{(k)} = \frac{(A_k, V_i^{(k-1)})}{(A_k, V_k^{(k-1)})}. \quad (5.8)$$

Для осуществимости процесса нужно, чтобы все скалярные произведения $\left(A_k, V_k^{(k-1)}\right)$ были отличными от нуля.

В качестве базиса $R^{(0)} = R_{n+1}$ берется естественный базис $V_1^{(0)} = (1, 0, \dots, 0)^T, \dots, V_n^{(0)} = (0, 0, \dots, 1)^T$.

Из хода процесса ясно, что на всех шагах вектор $V_{n+1}^{(k)}$ будет иметь $(n+1)$ -ю компоненту, равную единице. Единственный базисный вектор $V_{n+1}^{(n)}$ подпространства $R^{(n)}$ ортогонален ко всем векторам $A_1, A_2, \dots, A_n$ и имеет последнюю компоненту, равную единице. Таким образом, вектор $V_{n+1}^{(n)}$ дает численное решение системы.

6. Принципы построения итерационных процессов. Метод простой итерации

Принципы построения итерационных процессов. Основные итерационные процессы для решения линейных систем могут быть описаны посредством следующей общей схемы. Пусть дана система линейных уравнений $AX = F$ с неособенной матрицей A . Строится последовательность векторов $X^{(1)}, X^{(2)}, \dots, X^{(k)}, \dots$ по рекуррентным формулам

$$X^{(k)} = X^{(k-1)} + H^{(k)}(F - AX^{(k-1)}), \quad (6.1)$$

выбор последовательностей $H^{(k)}$ приводит к различным итерационным процессам.

Итерационные процессы, протекающие по формуле (6.1), обладают тем свойством, что для каждого из них точное решение X^* является неподвижной точкой. Это значит, что если за начальное приближение $X^{(0)}$ взято X^* , то все последующие приближения будут так же равны X^* .

Простейшими среди итерационных процессов являются стационарные итерационные процессы, в которых матрицы $H^{(k)}$ не зависят от номера шага k . В частности, при $H^{(k)} = E$ получается классический процесс последовательных приближений. Любой стационарный процесс с $H \neq E$ можно рассматривать как процесс последовательных приближений, примененный к равносильной системе $HAX = HF$, так сказать «подготовленной» к применению метода последовательных приближений. Конечно, осуществлять такую подготовку на самом деле обычно нет необходимости, и такого рода рассмотрение стационарных процессов лишь дает удобное средство для их теоретического исследования.

Близкими к стационарным итерационным процессам являются циклические, в которых матрицы $H^{(k)}$ периодически повторяются через

некоторое число p шагов. Ясно, что из каждого циклического процесса можно получить равносильный ему стационарный, принимая за один шаг стационарного процесса результат применения полного цикла из p шагов исходного циклического процесса.

Нестационарные итерационные процессы, в свою очередь, можно подразделить на два типа. Это, во-первых, нестационарные итерационные процессы в буквальном смысле, когда изменение матрицы $H^{(k)}$ осуществляется на каждом шагу. Во-вторых, сюда можно включить стационарные процессы с ускорением сходимости посредством замены, время от времени, стационарной матрицы H , определяющей процесс, на некоторые, специальным образом подобранные, матрицы $H^{(k)}$.

Выбор матрицы H для стационарного процесса и матриц $H^{(k)}$ для нестационарного может осуществляться многими различными способами на основании различных принципов.

Возможно построение матриц $H^{(k)}$ так, чтобы итерационный процесс сходил к решению для возможно более широкого класса систем уравнений.

Важным принципом построения итерационных процессов является принцип релаксации. Под этим понимается принцип выбора матриц $H^{(k)}$ из некоторого заранее очерченного класса матриц так, чтобы на каждом шагу процесса уменьшалась какая-либо величина, характеризующая точность решения системы.

Среди релаксационных методов наиболее разработаны координатные, в которых матрицы $H^{(k)}$ подобраны так, что на каждом шагу меняются одна или несколько компонент последовательных приближений, и градиентные, в которых матрицы $H^{(k)}$ являются скалярными.

О точности приближенного решения X системы $AX = F$ естественно судить по величине (в том или другом смысле) вектора ошибки $Y = X^* - X$, где X^* – точное решение системы. Однако вектор ошибки не может быть вычислен без знания точного решения системы и может лишь оцениваться. Вектором, характеризующим точность приближенного решения X системы $AX = F$, может служить также вектор невязки (невязка) $r = F - AX$. Ясно, что $r = AY$. Таким образом, релаксация может быть построена на уменьшении любой нормы каждого из этих векторов.

При положительно-определенных матрицах A удобной мерой точности является так называемая функция ошибки $f(X) = (AY, Y) = (Y, r) = (A^{-1}r, r)$. В силу положительной определенности матрицы A всегда $f(X) > 0$, причем $f(X) = 0$ только при $X = X^*$.

Метод простой итерации.

Рассмотрим систему уравнений

[illegible]

Разделим каждое уравнение системы (6.2) на диагональный элемент a_{11} .

Получим систему

[illegible]

или в матричной форме

$$X = BX + G, \quad (6.4)$$

где

$$B = \begin{pmatrix} 0 & -\frac{a_{12}}{a_{11}} & \dots & -\frac{a_{1n}}{a_{11}} \\ -\frac{a_{21}}{a_{22}} & 0 & \dots & -\frac{a_{2n}}{a_{22}} \\ \dots & \dots & \dots & \dots \\ -\frac{a_{n1}}{a_{nn}} & -\frac{a_{n2}}{a_{nn}} & \dots & 0 \end{pmatrix}, G = \begin{pmatrix} \frac{f_1}{a_{11}} \\ \frac{f_2}{a_{22}} \\ \vdots \\ \frac{f_n}{a_{nn}} \end{pmatrix} \quad (6.5)$$

Данное преобразование системы (6.2) в систему (6.4) равносильно умножению системы (6.2) слева на матрицу

$$H = \begin{pmatrix} \frac{1}{a_{11}} & 0 & \dots & 0 \\ 0 & \frac{1}{a_{22}} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \frac{1}{a_{nn}} \end{pmatrix}.$$

Таким образом, $H = D^{-1}$, где D - диагональная матрица $[a_{11}, \dots, a_{nn}]$.

Выберем начальное приближение $X^{(0)} = G$. Строим последовательные приближения по формулам

$$X^{(k)} = BX^{(k-1)} + G, \quad k = 1, 2, 3, \dots \quad (6.6)$$

или в развернутом виде

$$x_i^{(k)} = \sum_{j=1}^n b_{ij} x_j^{(k-1)} + g_i, \quad i = \overline{1, n}. \quad (6.7)$$

7. Необходимые и достаточные условия сходимости метода простой итерации. Подготовка системы к виду, удобному для итераций

Необходимые и достаточные условия сходимости метода простой итерации

Теорема 1. *Для того чтобы метод простой итерации сходил, необходимо и достаточно чтобы все собственные значения матрицы $B = E - D^{-1}A$ были по модулю меньше единицы.*

Так как условие теоремы 1 трудно проверяется, судить о сходимости процесса последовательных приближений лучше при помощи достаточных признаков, связанных непосредственно с элементами матрицы B . Некоторые достаточные признаки вытекают из теоремы 1.

Теорема 2. *Для того чтобы процесс последовательных приближений сходил, достаточно, чтобы какая-либо норма матрицы B была меньше единицы.*

Доказательство:

Действительно, если $\|B\| < 1$, то все собственные значения матрицы B меньше единицы и потому на основании теоремы 1 процесс последовательных приближений сходится. Если выполняется одно из условий:

$$\text{I. } \|B\|_I = \sum_{j=1}^n |b_{ij}| \leq \mu < 1 \text{ при } i = \overline{1, n} \quad (7.1)$$

$$\text{II. } \|B\|_{II} = \sum_{i=1}^n |b_{ij}| \leq \nu < 1 \text{ при } j = \overline{1, n} \quad (7.2)$$

$$\text{III. } \|B\|_{III} = \sum_{i,j=1}^n b_{ij}^2 \leq \rho < 1, \quad (7.3)$$

то процесс последовательных приближений сходится. Это равносильно преобладанию диагональных элементов в исходной матрице A .

Дадим теперь оценки быстроты сходимости процесса последовательных приближений в терминах нормы. При этом выбор нормы векторов совершенно безразличен, но норма матриц должна быть согласована с выбранной нормой векторов.

Теорема 3.

$$\text{Если } \|B\| < 1, \text{ то } \|X^* - X^{(k)}\| \leq \|B\|^k \|X^{(0)}\| + \frac{\|G\| \cdot \|B\|^k}{1 - \|B\|} \quad (7.4)$$

Доказательство:

Имеем

$$\begin{aligned} \|X^* - X^{(k)}\| &= \|(E - B)^{-1}G - (E + B + \dots + B^{k-1})G - B^k X^{(0)}\| \leq \\ &\leq \|(E - B)^{-1}G - (E + B + \dots + B^{k-1})G\| + \|B^k X^{(0)}\| \leq \\ &\leq \|(E - B)^{-1}G - (E + B + \dots + B^{k-1})G\| + \|B\|^k \|X^{(0)}\| \leq \\ &\leq \|B\|^k \|X^{(0)}\| + \frac{\|G\| \cdot \|B\|^k}{1 - \|B\|}. \end{aligned}$$

Часто бывает важно сравнить результаты двух последовательных приближений, то есть сравнить величины $\|X^* - X^{(k)}\|$ и $\|X^* - X^{(k-1)}\|$. Такое сравнение можно проводить на основании следующей теоремы.

$$\text{Теорема 4.} \quad \|X^* - X^{(k)}\| \leq \|B\| \|X^* - X^{(k-1)}\| \quad (7.5)$$

Доказательство:

Действительно, из равенств $X^* = BX^* + G$, $X^{(k)} = BX^{(k-1)} + G$ следует, что $X^* - X^{(k)} = B(X^* - X^{(k-1)})$.

$$\text{Отсюда } \|X^* - X^{(k)}\| = \|B(X^* - X^{(k-1)})\| \leq \|B\| \|X^* - X^{(k-1)}\|.$$

Подготовка системы к виду, удобному для итераций. Пусть матрица A положительно определена. Тогда система $AX = F$ всегда может быть подготовлена к виду, в котором метод последовательных приближений будет сходящимся. Подготовка состоит в переходе от данной системы $AX = F$ к

равносильной системе $HAX = HF$, где H - некоторая неособенная матрица, которая выбирается так, чтобы матрица HA была бы близка к единичной.

Положим

$$H = \frac{2}{\mu} E, \quad (7.6)$$

где μ - любая из норм матрицы A .

Тогда система уравнений преобразуется к виду

$$X = \left(E - \frac{2}{\mu} A \right) X + \frac{2}{\mu} F = BX + G \quad (7.7)$$

Собственные значения матрицы

$$B = E - \frac{2}{\mu} A \quad (7.8)$$

будут заключены в открытом интервале $(-1;1)$, в силу того, что собственные значения положительно определенной матрицы A находятся в интервале $(0;\mu)$.

Следовательно, метод последовательных приближений для системы (7.7) будет сходящимся.

8. Метод Зейделя

Пусть система линейных уравнений $AX = F$ представлена в виде:

$$X = BX + G, \quad (8.1)$$

где $B = E - A$, $G = F$.

Обозначим компоненты числового вектора решения через $x_1, x_2, \dots, x_n$. Одношаговый циклический процесс (метод Зейделя) напоминает процесс последовательных приближений, с той разницей, что при вычислении k -го приближения для i -ой компоненты учитываются вычисленные уже ранее k -ые приближения для компонент $x_1^{(k)}, \dots, x_{i-1}^{(k)}$.

Приближения строятся по формулам:

$$x_i^{(k)} = \sum_{j=1}^{i-1} b_{ij} x_j^{(k)} + \sum_{j=i}^n b_{ij} x_j^{(k-1)} + g_i \quad (8.2)$$

Одношаговый циклический процесс может быть истолкован двумя способами, как разновидность общего итерационного процесса. За один шаг процесса принять переход от вектора: $(x_1^{(k)}, \dots, x_{i-1}^{(k)}, x_i^{(k-1)}, \dots, x_n^{(k-1)})$ к вектору: $(x_1^{(k)}, \dots, x_i^{(k)}, x_{i+1}^{(k-1)}, \dots, x_n^{(k-1)})$; или же за один шаг можно считать результат применения полного цикла, т.е. переход от вектора $(x_1^{(k-1)}, \dots, x_n^{(k-1)})$ к вектору $(x_1^{(k)}, \dots, x_n^{(k)})$.

В первом истолковании метод не будет стационарным, но будет циклическим. Именно в этом случае в качестве матриц, определяющих процесс, берутся по очереди матрицы $\ell_{11}, \ell_{22}, \dots, \ell_{nn}$:

$$l_{ii} = \begin{bmatrix} 0 & \dots & 0 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \dots & 1 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \dots & 0 & \dots & 0 \end{bmatrix} \quad (8.3)$$

$$H_{n(k-1)+i} = \ell_{ii}.$$

Во втором истолковании процесс будет стационарным. Исследуя его подробнее, уравнение $X = BX + G$ представим в виде

$$X = (M + N)X + G, \quad (8.4)$$

где

$$M = \begin{bmatrix} 0 & 0 & \dots & 0 & 0 \\ b_{21} & 0 & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ b_{n1} & b_{n2} & \dots & b_{n,n-1} & 0 \end{bmatrix}; \quad N = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ 0 & b_{22} & \dots & b_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & b_{nn} \end{bmatrix} \quad (8.5)$$

В этих обозначениях формулы (8.2), представленные в матричной форме будут иметь вид:

$$X^{(k)} = MN^{(k)} + NX^{(k-1)} + G \quad (8.6)$$

Отсюда следует:

$$X^{(k)} = (E - M)^{-1} NX^{(k-1)} + (E - M)^{-1} G \quad (8.7)$$

Таким образом, один полный цикл одношагового циклического процесса для системы (8.4) оказывается равносильным одному шагу процесса последовательных приближений, примененного к системе:

$$X = (E - M)^{-1} NX + (E - M)^{-1} G, \quad (8.8)$$

которая равносильна исходной системе $X = (M + N)X + G$ и может быть получена из неё умножением слева на неособенную матрицу $(E - M)^{-1}$. Из такого представленного процесса следует, что для сходимости необходимо и достаточно чтобы все собственные значения матрицы $S = (E - M)^{-1} N$ были по модулю меньше единицы. Эти собственные значения матрицы являются корнями полинома $|S - tE| = 0$. Умножаем обе части этого уравнения на $|E - M|$ и воспользуемся теоремой о произведении определителей двух матриц, преобразуем уравнение к виду: $|N - (E - M)t| = 0$ или

$$\begin{vmatrix} b_{11}-t & b_{12} & \dots & b_{1n} \\ b_{21}t & b_{22}-t & \dots & b_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ b_{n1}t & b_{n2}t & \dots & b_{nn}-t \end{vmatrix} = 0 \quad (8.9)$$

Таким образом, для сходимости одношагового процесса необходимо и достаточно чтобы все корни уравнения (8.9) были по модулю меньше единицы.

Рассмотрим одно достаточное условие сходимости одношагового циклического процесса:

$$\|B\|_I = \max_i \sum_{j=1}^n |b_{ij}| \leq \mu < 1. \quad (8.10)$$

В этом случае для метода последовательного приближения справедливо:

$$\|X^* - X^{(k)}\| \leq \mu \|X^* - X^{(k-1)}\|, \quad (8.11)$$

где в качестве нормы вектора взята норма $\|X\| = \max_i |x_i|$.

Покажем, что при этом условии одношаговый циклический процесс сходится, и для него имеет место несколько лучшая оценка.

Если $X^* = (x_1, \dots, x_n)$ - точное решение, то

$$x_i = \sum_{j=1}^n b_{ij} x_j + g_i, \quad i = \overline{1, n} \quad (8.12)$$

Вычтем из (8.12) равенство (8.2) и оценим получившуюся разность:

$$\begin{aligned} x_i - x_i^{(k)} &= \sum_{j=1}^{i-1} b_{ij} (x_j - x_j^{(k)}) + \sum_{j=1}^n b_{ij} (x_j - x_j^{(k-1)}), \\ |x_i - x_i^{(k)}| &\leq \sum_{j=1}^{i-1} |b_{ij}| |x_j - x_j^{(k)}| + \sum_{j=1}^n |b_{ij}| |x_j - x_j^{(k-1)}| \end{aligned} \quad (8.13)$$

Обозначим $\sum_{j=1}^{i-1} |b_{ij}| = \beta_i$; $\sum_{j=1}^n |b_{ij}| = \gamma_i$. Тогда из (8.13) следует:

$$|x_i - x_i^{(k)}| \leq \beta_i \|X^* - X^{(k)}\| + \gamma_i \|X^* - X^{(k-1)}\|.$$

Взяв для i то значение i_o , при котором $|x_i - x_i^{(k)}|$ достигает максимума, получим

$$\|X^* - X^{(k)}\| \leq \frac{\gamma_{i_o}}{1 - \beta_{i_o}} \|X^* - X^{(k-1)}\|, \quad (8.14)$$

так как $\|X^* - X^{(k)}\| = \max_i |x_i - x_i^{(k)}| = |x_{i_o} - x_{i_o}^{(k)}|$.

Обозначим $\max \frac{\gamma_{i_o}}{1 - \beta_{i_o}} = \mu'$, тогда:

$$\|X^* - X^{(k)}\| \leq \mu' \|X^* - X^{(k-1)}\|, \quad (8.15)$$

Покажем, что $\mu' \leq \mu$.

Действительно,

$$\beta_i + \gamma_i = \sum_{j=1}^n |b_{ij}| \leq \mu \quad \text{и} \quad \beta_i + \gamma_i - \frac{\gamma_i}{1 - \beta_i} = \frac{\beta_i(1 - \beta_i - \gamma_i)}{1 - \beta_i} \geq 0, \quad \beta_i > 0$$

Тогда

$$\mu \geq \max(\beta_i + \gamma_i) \geq \max \frac{\gamma_i}{1 - \beta_i} = \mu' \quad (8.16)$$

Здесь знак « $\Rightarrow$ » возможен, только если $\max_i \sum_{j=1}^n |b_{ij}|$ достигается при $i=1$.

Однако, одношаговый циклический процесс не всегда оказывается более выгодным, чем процесс последовательного приближения. Иногда одношаговый циклический процесс сходится медленнее процесса последовательного приближения.

9. Методы полной и неполной релаксации.

Управление релаксацией

Методы полной релаксации. Пусть X – точное решение системы $AX = F$ с положительно-определенной матрицей A , X – некоторый вектор, $f(X)$ – функция ошибки. Ставится задача, как изменить i -компоненту вектора X , чтобы для измененного вектора X' значение функции ошибки было бы наименьшим.

Пусть

$$X' = X + \alpha e_i \quad (9.1)$$

тогда $Y' = Y - \alpha e_i$ и

$$\begin{aligned} f(X') &= (AY', Y') = (A(Y - \alpha e_i), Y - \alpha e_i) = \\ &= (AY, Y) - 2\alpha (AY, e_i) + \alpha^2 (Ae_i, e_i) = \\ &= (AY, Y) + \alpha^2 a_{ii} - 2\alpha r_i = f(X) + \frac{1}{a_{ii}} (a_{ii}\alpha - r_i)^2 - \frac{r_i^2}{a_{ii}} \end{aligned} \quad (9.2)$$

где r_i – i -компонента вектора невязки для приближения X .

Ясно, что $f(X')$ будет иметь минимальное значение при

$$\alpha = \frac{r_i}{a_{ii}} \quad (9.3)$$

И это минимальное значение равно

$$f(X') = f(X) - \frac{r_i^2}{a_{ii}} \quad (9.4)$$

Вычислим теперь i -компоненту невязки для приближения X' . Она равна:

$$\begin{aligned} (F - AX', e_i) &= (F - AX - \alpha(Ae_i, e_i)) = \\ &= (F - AX, e_i) - \alpha(Ae_i, e_i) = r_i - \alpha a_{ii} = 0 \end{aligned} \quad (9.5)$$

Т.е. приближение X' удовлетворяет i -му уравнению системы $AX = F$. Другими словами, i -я компонента приближения X' может быть вычислена из

i -го уравнения системы $AX = F$, в которое вместо остальных неизвестных подставлены компоненты вектора X .

Неполная релаксация. Вместо полной минимизации функции ошибки на каждом отдельном шаге процесса можно заботиться лишь об уменьшении функции ошибки. Процессы, построенные исходя из этого принципа, называются процессами неполной релаксации.

Выясним, как изменить одну компоненту с тем, чтобы функция ошибки уменьшилась.

Пусть $X' = X + \alpha e_i$. Тогда $f(X') - f(X) = \frac{1}{a_{ii}}(a_{ii}\alpha - r_i)^2 - \frac{r_i^2}{a_{ii}}$. Для того, чтобы значение $f(X') - f(X)$ было отрицательным, необходимо и достаточно, чтобы $(a_{ii}\alpha - r_i)^2 < r_i^2$, откуда $|a_{ii}\alpha - r_i| < |r_i|$. Это неравенство будет выполняться, если

$$\alpha = q \frac{r_i}{a_{ii}} \text{ при } 0 < q < 2 \quad (9.6)$$

При $q = 1$ будет иметь место полная минимизация функции ошибки, или, как говорят, полная релаксация. Полная релаксация называется нижней, если $0 < q < 1$ и верхней при $1 < q < 2$.

При неполной релаксации функция ошибки изменятся по формуле:

$$f(X') = f(X) - \frac{q(2-q)}{a_{ii}} r_i^2 \quad (9.7)$$

На каждом отдельном шаге процесса метод полной релаксации является наивыгоднейшим, так как он обеспечивает максимальное уменьшение функции ошибки за один шаг. Однако при проведении большего числа шагов может оказаться, что неполная релаксация дает лучший результат. Число q при неполной релаксации может изменяться от шага к шагу. Если процесс неполной релаксации берется циклическим при постоянном или циклически меняющемся q , то процесс можно рассматривать как частный случай общего одношагового циклического процесса, примененного к системе, подготовленной к виду:

$$X = (E - QD^{-1}A)X + QD^{-1}F \quad (9.8)$$

где $D = [a_{11} \dots a_{nn}]$, $Q = [q_1 \dots q_n]$ – диагональные матрицы, $q_1, \dots, q_n$ значения множителей релаксации.

Действительно, формулы для вычисления компонент результата k -го цикла будут:

$$x_i^{(k)} = x_i^{(k+1)} + q_i \frac{f_i - (a_{i1}x_1^{(k)} + \dots + a_{i,i-1}x_{i-1}^{(k)} + a_{ii}x_i^{(k-1)} + \dots + a_{in}x_n^{(k-1)})}{a_{ii}}, \quad i = \overline{1, n} \quad (9.9)$$

Легко найти необходимое и достаточное условие сходимости процесса.

Действительно:

$$M = -QD^{-1}L \quad (9.10)$$

$$N = E - QD^{-1}(D + R), \quad (9.11)$$

где Q – левая треугольная матрица с нулями на главной диагонали, R – правая треугольная матрица с нулями на главной диагонали.

$$\begin{aligned} S &= (E - M)^{-1}N = (E + QD^{-1}L)^{-1} \cdot (E - Q - QD^{-1}R) = \\ &= (D + QL)^{-1}(D - DQ - DR) \end{aligned} \quad (9.12)$$

Так что необходимым и достаточным условием сходимости процесса (9.9) будет требование, чтобы все собственные значения матрицы $(D + QL)^{-1}(D - DQ - DR)$ были по модулю меньше 1.

Собственные значения этой матрицы, очевидно, являются корнями уравнения:

$$\begin{vmatrix} (t + q_1 - 1) \cdot a_{11} & q_1 a_{12} & \dots & q_1 a_{1n} \\ tq_2 a_{21} & (t + q_2 - 1) \cdot a_{22} & \dots & q_2 a_{2n} \\ \dots & \dots & \dots & \dots \\ tq_n a_{n1} & tq_n a_{n2} & \dots & (t + q_n - 1) a_{nn} \end{vmatrix} = 0 \quad (9.13)$$

Быстрота сходимости метода будет зависеть от величины наибольшего модуля корней этого уравнения.

Полученный критерий сходимости процесса (9.9) представляет интерес главным образом в случае, если матрицы системы не положительно определены, так как в случае положительно определенной матрицы циклический процесс неполной релаксации всегда сходится.

Управление релаксацией. Вместо того, чтобы при проведении процесса координатной релаксации задаваться последовательностью ведущих индексов i_k *a priori* (как это делается, например, в циклическом процессе или в процессе, где управление задается десятичной записью числа e), можно выбирать индекс i_k на каждом шагу, исходя из результатов предыдущего шага. Так, например, самое быстрое убывание функции ошибки при переходе от $X^{(k-1)}$ к $X^{(k)}$,

получается, что если выбирать индекс i_k так, чтобы число $\frac{(r_{i_k}^{(k-1)})^2}{a_{i_k i_k}}$ было бы

наибольшим среди всех чисел $\frac{(r_{i_k}^{(k-1)})^2}{a_{ii}}$ где $i = 1, \dots, n$. Иными словами, индекс

i_k выбирается так, чтобы число $\frac{|r_{i_k}^{(k-1)}|}{\sqrt{a_{i_k i_k}}}$ было бы наибольшим. Процесс

релаксации с таким управлением носит название процесса Зейделя.

С именем Гаусса связывается процесс, при котором ведущий индекс

выбирается из условия $\max \frac{|r_i^{(k-1)}|}{a_{ii}}$. Наконец, правило управления Сауселла

определяется вычислением $\max |r_i^{(k-1)}|$.

10. Релаксация по длине вектора невязки. Групповая релаксация

Релаксация по длине вектора невязки. Рассмотрим теперь процессы, основанные на минимизации, или уменьшении длины, вектора невязки за счет изменения на каждом шаге одной компоненты предыдущего приближения. Будем по-прежнему полагать, что матрица A – положительно определенная. Пусть X – некоторое приближение и i – номер изменяемой компоненты. Положим:

$$X' = X + \delta e_i \quad (10.1)$$

тогда

$$r' = r - \delta A_i. \quad (10.2)$$

Значит

$$\begin{aligned} (r', r') &= (r - \delta A_i, r - \delta A_i) = (r, r) - 2\delta(r, A_i) + \delta^2(A_i, A_i) = \\ &= (r, r) + \frac{1}{(A_i, A_i)} \cdot [(r, A_i) - \delta(A_i, A_i)]^2 - \frac{(r, A_i)^2}{(A_i, A_i)}. \end{aligned} \quad (10.3)$$

Здесь A_i – i -й столбец матрицы A . Величина (r', r') будет минимальна при:

$$\delta = \frac{(r, A_i)}{(A_i, A_i)}. \quad (10.4)$$

Так что в случае полной релаксации можно взять

$$X' = X + \frac{(r, A_i)}{(A_i, A_i)} \cdot e_i \quad (10.5)$$

Если заданная система предварительно приведена к такому виду, что все $(A_i, A_i) = 1$ (что всегда можно сделать посредством замены неизвестных

$x_i = \frac{\bar{x}_i}{\sqrt{(A_i, A_i)}}$), мы будем иметь:

$$X' = X + (r, A_i) \cdot e_i \quad (10.6)$$

и

$$(r', r') = (r, r) - (r, A_i)^2 \quad (10.7)$$

Неполная релаксация в этом случае может проводиться по формулам:

$$X' = X + q(r, A_i) e_i, \quad (0 < q < 2) \quad (10.8)$$

Причем

$$(r', r') = (r, r) - q(2 - q)(r, A_i)^2 \quad (10.9)$$

Нетрудно видеть, что метод релаксации по длине вектора невязки равносильен релаксационному методу по функции ошибки для системы $A^T A X = A^T F$, полученной из системы $A X = F$ первой трансформацией Гаусса. Действительно: $(r, r) = (A Y, A Y) = (A^T A Y, Y)$, где Y - вектор ошибки, так что квадрат длины вектора невязки для системы $A X = F$ есть функция ошибки для системы $A^T A X = A^T F$.

Если последовательность номеров изменяемых компонент задана а priori (например, если применять циклический процесс), то проведение процесса не требует фактического выполнения трансформации Гаусса, т. е. вычисления матрицы $A^T A$. Если же осуществлять управление процессом, требуя на каждом шагу максимального убывания длины вектора невязки за счет выбора номера изменяемой компоненты, то целесообразно предварительно вычислить матрицу $A^T A$ и вести процесс (при $(A_i, A_i) = 1$) по формулам $r^{(k)} = r^{(k-1)} + r_{i_k}^{(k-1)} \cdot B_{i_k}$, где i_k - номер наибольшей по модулю компоненты вектора $r^{(k-1)}$, B_i есть i -й столбец матрицы $A^T A$. Это равносильно проведению релаксационного процесса с управлением для системы $A^T A X = A^T F$.

Групповая релаксация. Пусть $A X = F$ данная система. Отдельный шаг метода групповой релаксации заключается в следующем. Выделяется группа G индексов и при переходе от предшествующего приближения к следующему изменяются только компоненты с индексами из выбранной группы. Изменение осуществляется так, что уравнения, индексы которых

входят в группу G , удовлетворяются точно. Иначе говоря, изменяемые компоненты суть решения системы уравнений

$$\sum_{j \in G} a_{ij} x'_j = f_i - \sum_{m \in G} a_{im} x_m, \quad (10.10)$$

где индекс i пробегает всю группу G .

Легко видеть, что если матрица A положительно-определенная, то один шаг групповой релаксации минимизирует функцию ошибок в подпространстве, натянутом на единичные векторы с индексами, образующими группу G .

Метод групповой релаксации допускает много модификаций, в зависимости от принципа выбора групп G на каждом шагу. Простейшей модификацией является циклический групповой процесс, в котором все индексы раз навсегда разбиваются на несколько не перекрывающихся групп и по ходу процесса эти группы циклически чередуются. Этот процесс может рассматриваться как процесс последовательных приближений при надлежащей подготовке системы. Для того чтобы показать это, положим, для простоты, что группы состоят из последовательных индексов и чередование осуществляется в порядке их возрастания. Обозначим через D квазидиагональную матрицу, „вырезанную” из матрицы A в соответствии с разбиением индексов на группы, через L - матрицу, состоящую из элементов матрицы A , лежащих на лево от элементов матрицы D , через R - матрицу, состоящую из элементов A , лежащих направо от элементов матрицы D . Тогда

$$A = L + D + R \quad (10.11)$$

Пусть $X^{(k)}$ - приближение, полученное после завершения k циклов процесса. Тогда $X^{(k+1)}$ и $X^{(k)}$, очевидно, связаны соотношением $LX^{(k+1)} + DX^{(k+1)} + RX^{(k)} = F$. И потому

$$X^{(k+1)} = (L + D)^{-1} F - (L + D)^{-1} RX^{(k)} \quad (10.12)$$

Таким образом, изучаемый процесс равносильен процессу последовательных приближений, примененному к системе, подготовленной к виду

$$X = -(L + D)^{-1}RX + (L + D)^{-1}F. \quad (10.13)$$

Отсюда вытекает необходимое и достаточное условие сходимости процесса, которое заключается в том, чтобы собственные значения матрицы $(L + D)^{-1}R$ были по модулю меньше 1. Если матрица A положительно-определенная эрмитова матрица, то процесс сходится всегда.

Действительно, в этом случае собственные значения матрицы $(L + D)^{-1}R = (R^* + D)^{-1}R$ по модулю не превосходят $\frac{A_0}{\sqrt{\lambda_0 d_0 + A_0^2}}$. Здесь через A_0 обозначено $\|R\| = \|R^*\|$, λ_0 - наименьшее собственное значение матрицы A , d_0 - наименьшее собственное значение матрицы D .

При свободной групповой релаксации допускается не только не циклическое чередование заранее выбранных групп, но и изменение состава групп на каждом шаге. Как мы видели, отдельный шаг с выбранной группой $G^{(k)}$ осуществляется по формулам

$$\sum_{j \in G} a_{ij} x_j^{(k+1)} = f_i - \sum_{m \in G} a_{im} x_m^k, \quad i \in G^{(k)}, \quad x_j^{(k+1)} = x_j^{(k)}, \quad j \in G^{(k)}.$$

Эти формулы равносильны формулам

$$\sum_{j \in G^{(k)}} a_{ij} [x_j^{(k+1)} - x_j^{(k)}] = f_i - \sum_{m=1}^n a_{im} x_m^{(k)} = r_i^{(k)}, \quad i \in G^{(k)}. \quad (10.14)$$

Обозначив $x_j^{(k+1)} - x_j^{(k)} = \delta_j^{(k)}$, получим $x_j^{(k+1)} = x_j^{(k)} + \delta_j^{(k)}$, $x_j^{(k+1)} = x_j^{(k)}$, где $\delta_j^{(k)}$ есть решение системы

$$\sum_{j \in G^{(k)}} a_{ij} \delta_j^{(k)} = r_i^{(k)}, \quad i \in G^{(k)} \quad (10.15)$$

При неполной релаксации вычисление проводится по формулам

$$x_j^{(k+1)} = x_j^{(k)} + q_k \delta_j^{(k)}, \quad j \in G^{(k)}, \quad x_j^{(k+1)} = x_j^{(k)}, \quad \text{где } 0 < q_k < 2. \quad (10.16)$$

11. Градиент функционала. Градиентные методы. Метод наискорейшего спуска. Метод с наименьшими невязками

Градиент функционала. Пусть $F(X)$ функционал, вообще говоря, не линейный, определенный в вещественном евклидовом пространстве R и принимающий вещественные значения. Пусть вектор X задан координатами в некотором ортонормированном базисе. Тогда функционал $F(X)$ есть функция $F(x_1, x_2, \dots, x_n)$ от переменных координат вектора X . Мы будем предполагать, что функционал $F(X)$ дифференцируем, для чего достаточно потребовать, чтобы функция $F(x_1, x_2, \dots, x_n)$ имела непрерывные частные производные по всем аргументам.

Пусть Y - произвольный вектор единичной длины с координатами $y_1, y_2, \dots, y_n$. Производной от функционала F в точке X по направлению Y называется выражение

$$\frac{\partial F(X)}{\partial Y} = \lim_{t \rightarrow 0} \frac{F(X + tY) - F(X)}{t} = \left. \frac{d}{dt} F(X + tY) \right|_{t=0}. \quad (11.1)$$

Производная $\frac{\partial F(X)}{\partial Y}$ характеризует скорость изменения функционала F при изменении «аргумента» в направлении вектора Y . Имеем далее $F(X + tY) = F(x_1 + ty_1, \dots, x_n + ty_n)$.

И потому

$$\frac{\partial F(X)}{\partial Y} = \left. \frac{d}{dt} F(x_1 + ty_1, \dots, x_n + ty_n) \right|_{t=0} = \frac{\partial F}{\partial x_1} y_1 + \dots + \frac{\partial F}{\partial x_n} y_n = (Z, Y). \quad (11.2)$$

Где Z - вектор с координатами $\frac{\partial F}{\partial x_1}, \dots, \frac{\partial F}{\partial x_n}$. Этот вектор называется градиентом функционала $F(X)$. Из последнего равенства вытекает, что

$$\frac{\partial F(X)}{\partial Y} \leq |Z| \cos(Z, Y) \text{ т.к. } |Y| = 1. \text{ Отсюда следует, что}$$

$$-|Z| \leq \frac{\partial F(X)}{\partial Y} \leq |Z|. \quad (11.3)$$

Причем $\frac{\partial F(X)}{\partial Y} = |Z|$, если направление Y совпадает с направлением градиента, и $\frac{\partial F(X)}{\partial Y} = -|Z|$, если направление противоположно направлению градиента. Поэтому направление градиента есть направление наибольшей скорости роста функционала F в данной точке, а направление, противоположное градиенту, есть направление наибольшей скорости убывания.

Метод наискорейшего спуска. Пусть A положительно-определенная матрица и пусть

$$AX = F \quad (11.4)$$

Метод заключается в минимизации функционала ошибки в направлении, противоположном градиенту этого функционала.

Заданная линейная система связывается с задачей отыскания вектора, дающего минимум функционалу

$$H(X) = (AX, X) - 2(F, X), \quad (11.5)$$

отличающемся лишь постоянным, но заранее известным слагаемым (AX^*, X^*) от функции ошибки $f(X) = (AY, Y)$. Здесь X^* - точное решение системы, совпадающее с вектором, реализующем минимум $H(X)$, $Y = X^* - X$ - вектор ошибки.

Поставленная задача решается следующим образом. Выбирается произвольный вектор X_0 . Вычисляется направление, противоположное градиенту функционала $H(X)$ (или, что то же самое, градиенту функции ошибки) в этой точке, которое совпадает с направлением невязки $r_0 = F - AX_0$ выбранного начального приближения. Из точки X_0 мы движемся в выбранном

направлении до точки X_1 , в которой функционал $H(X)$ становится минимальным. Так как

$$\begin{aligned}
 H(X_0 + \alpha r_0) &= (AX_0 + \alpha Ar_0, X_0 + \alpha r_0) - 2(F, X_0 + \alpha r_0) = \\
 &= (AX_0, X_0) + 2\alpha (AX_0, r_0) + \alpha^2 (Ar_0, r_0) - \\
 &\quad - 2(F, X_0) - 2\alpha (F, r_0) = \\
 &= H(X_0) - 2\alpha (r_0, r_0) + \alpha^2 (Ar_0, r_0) = \\
 &= H(X_0) - \frac{(r_0, r_0)^2}{(Ar_0, r_0)} + (Ar_0, r_0) \left[\alpha - \frac{(r_0, r_0)}{(Ar_0, r_0)} \right]^2.
 \end{aligned} \tag{11.6}$$

Это выражение достигает минимума при

$$\alpha = \alpha_0 = \frac{(r_0, r_0)}{(Ar_0, r_0)}, \tag{11.7}$$

и этот минимум равен

$$H(X_0) - \frac{(r_0, r_0)^2}{(Ar_0, r_0)}. \tag{11.8}$$

Итак, $X_1 = X_0 + \alpha_0 r_0$, где $r_0 = F - AX_0$, $\alpha_0 = \frac{(r_0, r_0)}{(Ar_0, r_0)}$,

$$H(X_1) = H(X_0) - \frac{(r_0, r_0)^2}{(Ar_0, r_0)}, \tag{11.9}$$

$$f(X_1) = f(X_0) - \frac{(r_0, r_0)^2}{(Ar_0, r_0)}. \tag{11.10}$$

Далее определяется $X_2 = X_1 + \alpha_1 r_1$, $r_1 = F - AX_1$, $\alpha_1 = \frac{(r_1, r_1)}{(Ar_1, r_1)}$ и процесс

продолжается далее по формулам

$$X_{k+1} = X_k + \alpha_k r_k \tag{11.11}$$

$$r_k = F - AX_k = r_{k-1} - \alpha_{k-1} Ar_{k-1}, \text{ где } \alpha_k = \frac{(r_k, r_k)}{(Ar_k, r_k)}. \tag{11.12}$$

При этом

$$f(X_{k+1}) = f(X_k) - \frac{(r_k, r_k)^2}{(Ar_k, r_k)} \quad (11.13)$$

Заметим сразу, что при фактическом проведении процесса векторы r_k , особенно при большом порядке матрицы системы, удобнее вычислять по формуле

$$r_k = r_{k-1} - \alpha_{k-1} Ar_{k-1}. \quad (11.14)$$

Однако, вследствие ошибок округления, так вычисленные векторы r_k после нескольких шагов процесса могут начать отклоняться от истинных невязок $F - AX_k$. Поэтому следует время от времени вычислять r_k по формуле $r_k = F - AX_k$.

Градиентный метод с наименьшими невязками. Пусть A - положительно-определенная матрица, X_0 - начальное приближение к решению системы $AX = F$. Следующее приближение X_1 ищется также, как и в методе наискорейшего спуска, в виде $X_0 + \beta r_0$, но параметр β подбирается так, чтобы минимизировалась длина вектора невязки $|r|$ или, что то же самое, $(r, r)^2 = |r|^2$. Таким образом, здесь минимизируется функционал (r, r) в направлении, противоположном градиенту другого функционала, именно функции ошибок. После выполнения первого шага процесс повторяется.

Выведем формулы, связывающие соседние приближения. Пусть

$$X_{k+1} = X_k + \beta_k r_k. \quad (11.15)$$

Тогда

$$r_{k+1} = r_k - \beta_k Ar_k. \quad (11.16)$$

Следовательно,

$$\begin{aligned} (r_{k+1}, r_{k+1}) &= (r_k, r_k) - 2\beta_k (Ar_k, r_k) + \beta_k^2 (Ar_k, Ar_k) = \\ &= (r_k, r_k) - \frac{(Ar_k, r_k)^2}{(Ar_k, r_k)} + (Ar_k, Ar_k) \left[\beta_k - \frac{(Ar_k, r_k)}{(Ar_k, Ar_k)} \right]^2. \end{aligned} \quad (11.17)$$

Тогда (r_{k+1}, r_{k+1}) принимает минимальное значение равное $(r_k, r_k) - \frac{(Ar_k, r_k)^2}{(Ar_k, Ar_k)}$ при $\beta_k = \frac{(Ar_k, r_k)}{(Ar_k, Ar_k)}$. Итак, рабочими формулами метода являются следующие:

$$X_{k+1} = X_k + \beta_k r_k, \quad r_k = F - AX_k = r_{k-1} - \beta_{k-1} Ar_{k-1}, \quad \beta_k = \frac{(Ar_k, r_k)}{(Ar_k, Ar_k)} \quad (11.18)$$

Градиентные методы с неполной релаксацией. Пусть по-прежнему A -положительно-определенная матрица, и ищется решение системы $AX = F$.

Рассмотрим итерационный процесс, в котором каждое последующее приближение получается из предыдущего изменением в направлении, противоположном градиенту функции ошибок, причем так, что на каждом шаге функция ошибок уменьшается. Формулы для получения последовательных приближений, очевидно, должны иметь вид

$$X_{k+1} = X_k + \gamma_k r_k. \quad (11.19)$$

Для дальнейшего исследования нам удобно положить

$$\gamma_k = q_k \alpha_k, \quad (11.20)$$

где α_k коэффициент в методе наискорейшего спуска.

Имеем

$$\begin{aligned} f(X_{k+1}) &= f(X_k) - 2\gamma_k (r_k, r_k) + \gamma_k^2 (Ar_k, r_k) = \\ &= f(X_k) - 2q_k \alpha_k (r_k, r_k) + q_k^2 \alpha_k^2 (Ar_k, r_k) = \\ &= f(X_k) - 2q_k \frac{(r_k, r_k)^2}{(Ar_k, r_k)} + q_k^2 \frac{(r_k, r_k)^2}{(Ar_k, r_k)} = \\ &= f(X_k) - (2q_k - q_k^2) \frac{(r_k, r_k)^2}{(Ar_k, r_k)} \end{aligned} \quad (11.21)$$

Из этой формулы ясно, что для того, чтобы $f(X_{k+1})$ было меньше, чем $f(X_k)$, необходимо и достаточно выполнение для множителей релаксации q_k неравенств

$$0 < q_k < 2. \quad (11.22)$$

Будем называть группу методов, в которых не все q_k равны 1 методами неполной градиентной релаксации (одношаговыми). Если же все множители релаксации $q_k \leq 1$, но не все равны 1, метод называется методом нижней релаксации, если все $q_k \geq 1$, но не все $q_k = 1$ – методом верхней релаксации.

Так, например, метод с минимальными невязками является методом нижней релаксации, так как в нем

$$q_k = \frac{\beta_k}{\alpha_k} = \frac{(Ar_k, r_k)^2}{(Ar_k, Ar_k)(r_k, r_k)} \leq 1 \quad (11.23)$$

по неравенству Коши-Буняковского.

Здесь знак равенства возможен, только если r_k есть собственный вектор матрицы A .

В группу методов неполной релаксации входит и градиентный метод с постоянным множителем $\gamma_k = \gamma$, если этот множитель удовлетворяет неравенству

$$0 < \gamma < \frac{2}{M}, \quad (11.24)$$

где M - наибольшее собственное значение матрицы A .

12. Методы сопряженных направлений. S-шаговые градиентные методы

Метод сопряженных направлений. Методами сопряженных направлений для решения систем линейных уравнений

$$AX = F \quad (12.1)$$

называются методы, в которых решение X представляется в виде линейной комбинации векторов, ортогональных в некоторой метрике, так или иначе связанной с матрицей системы. Термин «сопряженные направления» происходит от того, что направления векторов, ортогональных в некоторой метрике R , сопряжены по отношению к поверхности второго порядка

$$(RX, X) = \text{const}. \quad (12.2)$$

Методы сопряженных направлений укладываются в следующую общую схему. Пусть $R = CAB$ – положительно определенная матрица. Пусть построена система R -ортогональных векторов $s_1, \dots, s_n$ (векторов R -сопряженных направлений). Ищем решение системы в виде

$$X = X_0 + B \cdot \sum_{i=1}^n a_i s_i. \quad (12.3)$$

Здесь X_0 – начальное приближение, выбираемое, вообще говоря, произвольно. Обычно берется $X_0 = 0$. Подстановка (12.3) в систему дает:

$$AB \sum_{i=1}^n a_i s_i = F - AX_0 = r_0. \quad (12.4)$$

Умножая последнее равенство на C , получим

$$R \sum_{i=1}^n a_i s_i = Cr_0. \quad (12.5)$$

Откуда коэффициенты a_i легко определяются

$$a_i = \frac{(Cr_0, s_i)}{(Rs_i, s_i)}. \quad (12.6)$$

Решение (12.3) можно представить также в виде

$$X = X_0 + \sum_{i=1}^n a_i B s_i \quad (12.7)$$

Метод ортогональных векторов. В этом методе матрица A положительно-определенная, $R = A$. Соответствующие формулы получаются при $C = B = E$, $R = A$. Векторами направлений будут A -ортогональные векторы

$$s_i = e_i - \gamma_{i,1}s_1 - \dots - \gamma_{i,i-1}s_{i-1}, \quad \gamma_{ij} = \frac{(Ae_i, s_j)}{(Ae_j, s_j)}, \quad i = 2, \dots, n, \quad j = 1, \dots, i-1 \quad (12.8)$$

Если $X_0 = 0$, то

$$X = \sum_{i=1}^n a_i s_i \quad (12.9)$$

$$a_i = \frac{(F, s_i)}{(s_i, A s_i)} = \frac{(F, s_i)}{(e_i, A e_i)}. \quad (12.10)$$

Отметим, что векторы Ae_i суть не что иное, как столбцы A_i матрицы A . Таким образом,

$$\dots\dots\dots \gamma_{ij} = \frac{(A_i, s_j)}{(A_j, s_j)}, \quad (12.11)$$

$$a_i = \frac{(F, s_i)}{(s_i, A_i)}. \quad (12.12)$$

Контроль вычислений осуществляется вычислением величин (A_i, s_j) , которые теоретически должны равняться нулю при $j > i$.

Метод сопряженных градиентов. В методе сопряженных градиентов $R = A$, матрица A является положительно-определенной, $B = C = E$. A -ортогональные векторы $s_1, \dots, s_n$ теоретически строятся в процессе A -ортогонализации невязок $r_0, r_1, \dots, r_{n-1}$ последовательных приближений $X_0, X_1, \dots, X_{n-1}$, определяемых по формулам метода сопряженных направлений, т.е.:

$$X_1 = X_0 + \sum_{j=1}^i a_j s_j, \quad (12.13)$$

где $a_j = \frac{(r_0, s_j)}{(s_j, As_j)}$ – система векторов, подвергающихся процессу ортогонализации. Система не задается заранее, а строится параллельно с построением векторов сопряженных направлений и соответствующих им последовательных приближений. Отметим, что название метода связано с тем обстоятельством, что невязка r_i является градиентом функции ошибок, вычисленным в точке X_i . Очевидно, что метод сопряженных градиентов естественно связывается с решением лишь одной системы $AX = F$, хотя знание базиса $s_1, \dots, s_n$ позволяет решать так же и системы с отличными от F свободными членами.

Процесс построения последовательных приближений оборвется, как только некоторая невязка r_k окажется равной нулю, т. е. как только приближение совпадает с точным решением системы. Предполагается и другая причина остановки процесса, не приводящая к точному решению. Именно, процесс может остановиться, если некоторая невязка r_k окажется линейной комбинацией предыдущих. Тогда получится $s_{k+1} = 0$ и $r_{k+1} = r_k$. Однако это невозможно, в силу следующей теоремы.

Теорема 1. *Ненулевые невязки $r_0, \dots, r_{k-1}$ взаимно ортогональны.*

Из теоремы следует, что если $r_k \neq 0$, то r_k не может быть линейной комбинацией невязок $r_0, \dots, r_{k-1}$, так что обращение невязки в ноль является единственной причиной остановки процесса.

Выведем теперь расчетные формулы процесса. Прежде всего заметим, что из формул

$$X_i = X_{i-1} + a_i s_i, \quad a_i = \frac{(r_0, s_i)}{(As_i, s_i)} \quad (12.14)$$

следует

$$r_i = r_{i-1} - a_i A s_i, \quad (12.15)$$

причем $a_i \neq 0$, $r_{i-1} \neq r_i$. Допустим, что уже построены векторы $s_1, \dots, s_i, r_0, \dots, r_i$.

Покажем, что следующий вектор направления s_{i+1} строится по формуле

$$s_{i+1} = r_i + b_i s_i, \quad (12.16)$$

где

$$b_i = -\frac{(r_i, A s_i)}{(s_i, A s_i)} \quad (12.17)$$

Для того чтобы убедиться в этом, достаточно доказать, что так полученный вектор s_{i+1} будет A -ортогонален к векторам $s_1, s_2, \dots, s_i$. Ясно, что

$$(s_{i+1}, A s_i) = (r_i, A s_i) - \frac{(r_i, A s_i)}{(s_i, A s_i)} (s_i, A s_i) = 0.$$

Рассмотрим теперь $(s_{i+1}, A s_j)$ при $j = 1, 2, \dots, i-1$. Прежде всего отметим, что $(s_{i+1}, A s_j) = (r_i, A s_j)$. Далее по формуле (12.15) имеем

$$A s_j = \frac{1}{a_j} (r_i, r_{j-1}) - \frac{1}{a_j} (r_i, r_j) = 0 \quad \text{в силу того, что } j < i, j-1 < i \text{ и невязки}$$

ортогональны. Итак, формула (12.16) справедлива.

Отметим, что вычисление коэффициентов a_i, b_i можно производить и по формулам

$$a_i = \frac{(s_i, r_{i-1})}{(s_i, A s_i)} = \frac{(r_{i-1}, r_{i-1})}{(s_i, A s_i)}, \quad (12.18)$$

$$b_i = \frac{(r_i, r_i)}{(r_{i-1}, r_{i-1})}, \quad (12.19)$$

справедливость которых легко устанавливается.

При практическом применении метода коэффициенты следует вычислять

по формуле $a_i = \frac{(r_{i-1}, r_{i-1})}{(s_i, A s_i)}$, пользуясь $a_i = \frac{(s_i, r_{i-1})}{(s_i, A s_i)}$ для контроля.

Формула $a_i = \frac{(r_0, s_i)}{(s_i, As_i)}$ чувствительна к ошибкам округления и ее

применение нецелесообразно.

Таким образом, последовательность действий по методу сопряженных градиентов такова.

1. Выбирается произвольное приближение X_0 к решению системы $AX = F$ и вычисляется невязка $r_0 = F - AX_0$.

2. По рекуррентным соотношениям

$$\begin{aligned} s_1 = r_0, \quad r_i = r_{i-1} - a_i As_i, \quad a_i = \frac{(r_{i-1}, r_{i-1})}{(s_i, As_i)} = \frac{(s_i, r_{i-1})}{(s_i, As_i)}, \\ s_{i+1} = r_i + b_i s_i, \quad b_i = \frac{(r_i, r_i)}{(r_{i-1}, r_{i-1})} = -\frac{(r_i, As_i)}{(s_i, As_i)} \end{aligned} \quad (12.20)$$

строятся векторы направлений $s_1, s_2, \dots; r_1, r_2, \dots$. Процесс заканчивается тем, что при некотором $m \leq n$ окажется $r_m = 0$.

3. Решение системы получается по формуле

$$X = X_0 + \sum_{i=1}^m a_i s_i. \quad (12.21)$$

Как правило, процесс протекает без вырождения, так что $m=n$. Контроль вычислений производится при помощи вычисления скалярных произведений (s_i, As_i) или скалярных произведений (r_i, r_j) , которые должны быть нулями.

S-шаговые градиентные методы. Поставим вопрос о том, как определить множители $\gamma_1, \dots, \gamma_s$ так, чтобы, исходя из начального приближения X_0 , получить наилучший результат после s шагов процесса и как объединить эти s шагов в один шаг нового вычислительного процесса. Иначе говоря, как построить вычислительный процесс, один шаг которого равносильен s шагам одношагового градиентного метода, выбранным согласно наилучшей стратегии.

Пусть

$$\begin{aligned}
X_0 &= X^{(0)} \\
X_1 &= X_0 + \gamma_0(F - AX_0) \\
X_2 &= X_1 + \gamma_1(F - AX_1) \\
&\dots\dots\dots \\
X^{(1)} &= X_S = X_{S-1} + \gamma_{S-1}(F - AX_{S-1})
\end{aligned} \tag{12.22}$$

Тогда, переходя к векторам ошибки, получим

$$\begin{aligned}
Y_1 &= Y_0 - \gamma_0 AY_0 = (E - \gamma_0 A)Y_0 \\
Y_2 &= Y_1 - \gamma_1 AY_1 = (E - \gamma_1 A)Y_1 \\
&\dots\dots\dots \\
Y^{(1)} &= Y_S = Y_{S-1} - \gamma_{S-1} AY_{S-1} = (E - \gamma_{S-1} A)Y_{S-1},
\end{aligned} \tag{12.23}$$

откуда

$$\begin{aligned}
Y^{(1)} &= (E - \gamma_0 A)(E - \gamma_1 A) \dots (E - \gamma_{S-1} A)Y_0 = \\
&= (E + c_1 A + c_2 A^2 + \dots + c_s A^s)Y_0 = \\
&= Y_0 + c_1 r_0 + \dots + c_s A^{s-1} r_0,
\end{aligned} \tag{12.24}$$

где $r_0 = F - AX^{(0)}$ и

$$X^{(1)} = X^{(0)} - c_1 r_0 - \dots - c_s A^{s-1} r_0. \tag{12.25}$$

Таким образом, задача сводится к определению коэффициентов $c_1, \dots, c_s$ так, чтобы значение $f(X^{(1)})$ функции ошибки было бы наименьшим. s -е приближение метода сопряженных градиентов минимизирует функцию ошибки среди векторов $X^{(0)} + V$, где вектор V принадлежит подпространству, натянутому на вектора $r_0, Ar_0, \dots, A^{s-1}r_0$.

Итак, результат одного шага s -шагового метода наискорейшего спуска совпадает с результатом s -го приближения метода сопряженных градиентов, а именно

$$X^{(1)} = X^{(0)} + \sum_{j=1}^s a_j s_j, \tag{12.26}$$

$$\text{где } s_1 = r_0, \quad r_i = r_{i-1} - a_i A s_i, \quad a_i = \frac{(s_i, r_{i-1})}{(s_i, A s_i)} = \frac{(r_{i-1}, r_{i-1})}{(s_i, A s_i)}.$$

На следующем шаге

$$s_{i+1} = r_i + b_i s_i, \quad b_i = \frac{(r_i, r_i)}{(r_{i-1}, r_{i-1})} = -\frac{(r_i, As_i)}{(s_i, As_i)} \quad i = (1, 2, \dots, s-1) \quad (12.27)$$

Поэтому при фактическом проведении s -шагового метода наискорейшего спуска не нужны ни коэффициенты $c_1, \dots, c_s$, ни тем более коэффициенты $\gamma_1, \dots, \gamma_s$. Многошаговый метод наискорейшего спуска впервые был описан в работах Л. В. Канторовича. В этих работах для составления последовательных приближений фактически находились коэффициенты $c_1, \dots, c_s$ посредством решения системы s линейных уравнений. Эта система для первого шага имеет вид

[illegible]

Для каждого последующего шага начальная невязка r_0 должна быть заменена невязкой, полученной в результате предыдущего шага.

13. Полная проблема собственных значений. Устойчивость проблемы собственных значений. Метод Крылова. Нахождение собственных векторов по методу Крылова

Полная проблема собственных значений. Под полной проблемой собственных значений понимается проблема нахождения всех собственных значений матрицы A , так же как и соответствующих этим собственным значениям собственных векторов (или векторов, образующих канонический базис). Напомним, что собственными значениями матрицы A называются корни ее характеристического полинома, т. е. корни уравнения

$$|A - tE| = \begin{vmatrix} a_{11} - t & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - t & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} - t \end{vmatrix} = (-1)^n [t^n - p_1 t^{n-1} - \dots - p_n] = 0$$

Определение компонент собственного вектора требует решения системы n однородных уравнений с n неизвестными; для вычисления всех собственных векторов матрицы, вообще говоря, требуется решить n систем вида $(A - \lambda_i E)X_i = 0$, где $X_i = (x_{1i}, \dots, x_{ni})$ - собственный вектор матрицы A , соответствующий собственному значению λ_i .

Коэффициенты p_i характеристического полинома являются, с точностью до знака, суммами всех миноров определителя матрицы A порядка i , опирающихся на главную диагональ. Непосредственное вычисление коэффициентов p_i является чрезвычайно громоздким и требует огромного числа операций.

Совершенно естественно, поэтому, появление специальных вычислительных приемов, упрощающих численное решение поставленных задач. Большинство методов, дающих решение полной проблемы собственных значений, включает предварительное вычисление коэффициентов характеристического полинома, которое осуществляется теми или иными средствами, минуя вычисление многочисленных миноров. Собственные

значения вычисляются затем по какому-либо методу для приближенного вычисления корней полинома. Одним из лучших способов приближенного вычисления корней полинома $\varphi(t) = a_0 t^n + a_1 t^{n-1} + \dots + a_{n-1} t + a_n$ является способ Ньютона. Именно, если t_0 есть некоторое исходное приближение к корню, то последовательные приближения $t_1, t_2, \dots$ вычисляются по формуле

$$t_i = t_{i-1} - \frac{\varphi(t_{i-1})}{\varphi'(t_{i-1})}, \quad i = 1, 2, \dots$$

Если t_0 взято достаточно близко к исходному корню λ , то последовательные приближения t_i сходятся к λ с квадратичной сходимостью, т.е. погрешность последующего приближения будет иметь порядок квадрата предыдущей:

$$|\lambda - t_i| \leq c |\lambda - t_{i-1}|^2, \quad \text{где } c - \text{некоторая константа.}$$

Вычисление значений полинома $\varphi(t)$ и его производной в данной точке t_0 следует проводить по схеме Горнера:

$$\begin{array}{ccccccc} a_0 & a_1 & a_2 & \dots & a_{n-1} & a_n & \| t_0 \\ a_0 & b_1 & b_2 & \dots & b_{n-1} & b_n & \\ a_0 & c_1 & c_2 & \dots & c_{n-1} & & \end{array}$$

которая заполняется по рекуррентным формулам:

$$b_i = b_{i-1} t_0 + a_i \quad (i = 1, 2, \dots, n), \quad b_0 = a_0,$$

$$c_i = c_{i-1} t_0 + b_i \quad (i = 1, 2, \dots, n-1), \quad c_0 = a_0.$$

Тогда $b_n = \varphi(t_0)$, $c_{n-1} = \varphi'(t_0)$.

Числа же $a_0, b_1, \dots, b_{n-1}$ будут коэффициентами частного $\varphi_1(t_0)$ от деления полинома $\varphi(t)$ на $t - t_0$. Соответственно числа $a_0, c_1, \dots, c_{n-1}$ будут коэффициентами частного $\varphi_2(t)$ от деления $\varphi_1(t)$ на $t - t_0$. По большей части собственные векторы матрицы удастся определить, используя промежуточные результаты вычислений, проведенных для определения коэффициентов характеристического полинома. Конечно, для определения собственного вектора, принадлежащего тому или другому собственному

значению, это собственное значение должно быть уже вычислено. Методы этой группы являются точными, т. е. если их осуществлять для матриц, элементы которых заданы точно (рациональными числами) и вычисления проводить точно (по правилам действий над обыкновенными дробями), то в результате будет получено точное значение коэффициентов характеристического полинома, и компоненты собственных векторов окажутся выраженными точными формулами через собственные значения.

Наряду с точными методами для решения проблемы собственных значений имеются методы итерационные, в которых собственные значения получаются как пределы некоторых числовых последовательностей, так же как и компоненты соответствующих им собственных векторов. В итерационных методах, как правило, собственные значения вычисляются непосредственно, без предварительного вычисления коэффициентов характеристического полинома. Это существенно упрощает задачу, так как вычисление корней полинома, коэффициенты которого известны, достаточно трудоемко.

Однако итерационные методы более приспособлены к решению частичной проблемы собственных значений. Под частичной проблемой мы подразумеваем задачу нахождения одного или нескольких собственных значений и соответствующих им собственных векторов.

Полная и частичная проблемы собственных значений совершенно различны как по методам их решения, так и по области приложений. Решение полной проблемы для матриц даже не очень высокого порядка неизбежно оказывается весьма громоздким, и возможность решения частичной проблемы, минуя тяжести решения полной, является очень ценной для практики.

Устойчивость проблемы собственных значений. При постановке проблемы собственных значений для матриц, элементы которых заданы приближенно, естественно возникает вопрос об устойчивости полученного решения, иными словами, вопрос о том, как изменяются собственные

значения и собственные векторы при изменении элементов данной матрицы в пределах допустимой погрешности.

То, что в отдельных случаях проблема собственных значений не может быть устойчивой, ясно из следующих соображений. Допустим, что данная матрица, если ее численное задание рассматривать как точное, имеет лишь простые собственные значения, однако, при некотором определенном изменении ее элементов в пределах точности задания можно прийти к матрице, имеющей кратное собственное значение, с нелинейным элементарным делителем. В этом случае, каноническая форма матрицы при изменении ее элементов в пределах точности задания претерпевает качественное изменение, переходя от чисто диагональной формы к общей канонической форме. В частности, даже число собственных векторов изменяется скачкообразно. В этих условиях, конечно, полная проблема собственных значений, вместе с определением собственных векторов, просто теряет смысл. В условиях же, близких к описанной ситуации, проблема определения собственных векторов наверное не имеет устойчивого решения.

Пусть A - данная матрица и $A + dA$ - близкая к ней матрица. Выясним как изменяются собственные значения и собственные векторы матрицы A , когда она получает приращение dA . Проведем подсчет в предположении, что все собственные значения матрицы A различны, отбрасывая величины второго порядка малости, т. е. будем рассматривать dA (и соответственно dX и $d\lambda$) как дифференциалы, а не как конечные приращения.

Пусть

$$AX_i = \lambda_i X_i \quad (i=1, 2, \dots, n). \quad (13.1)$$

Тогда

$$(dA)X_i + AdX_i = \lambda_i dX_i + d\lambda_i X_i. \quad (13.2)$$

Пусть $V_1, \dots, V_n$ - собственные векторы сопряженной матрицы A^* , соответствующие собственным значениям $\overline{\lambda_1}, \dots, \overline{\lambda_n}$.

Тогда

$$\left((dA)X_i, V_j\right) + \left(A(dX_i), V_j\right) = \lambda_i \left(dX_i, V_j\right) + d\lambda_i \left(X_i, V_j\right). \quad (13.3)$$

Положив в равенстве (13.3) $i=j$, получим

$$\left((dA)X_i, V_i\right) + \left(A(dX_i), V_i\right) = \left(dX_i, \overline{\lambda_i} V_i\right) + d\lambda_i \left(X_i, V_i\right), \text{ откуда}$$

$$d\lambda_i = \frac{\left((dA)X_i, V_i\right)}{\left(X_i, V_i\right)} \quad (13.4)$$

ибо $AV_i = \overline{\lambda_i} V_i$. Положим теперь $i \neq j$. Тогда, в силу равенств $\left(X_i, V_j\right) = 0$ и

$$\left(A(dX_i), V_j\right) = \lambda_j \left(dX_i, V_j\right), \text{ получим } \left(\lambda_i - \lambda_j\right) \left(dX_i, V_j\right) = \left((dA)X_i, V_j\right), \text{ откуда}$$

$$\left(dX_i, V_j\right) = \frac{\left((dA)X_i, V_j\right)}{\left(\lambda_i - \lambda_j\right)}.$$

Пусть

$$dX_i = \sum_{j=1}^n a_{ij} X_j, \quad (13.5)$$

тогда $\left(dX_i, V_j\right) = a_{ij} \left(X_j, V_j\right)$ и, следовательно,

$$a_{ij} = \frac{\left((dA)X_i, V_j\right)}{\left(X_j, V_j\right) \left(\lambda_i - \lambda_j\right)}. \quad (13.6)$$

Коэффициент a_{ii} остается, естественно, неопределенным, в силу неоднозначности собственного вектора, и без нарушения общности можно считать, что $a_{ii} = 0$. Перейдем теперь к оценкам. Из формулы (13.4) получим

$$|d\lambda_i| \leq \frac{\|dA\| \cdot |X_i| \cdot |V_i|}{\left|(X_i, V_i)\right|} = c_i \|dA\|, \text{ где}$$

$$c_i = \frac{|X_i| \cdot |V_i|}{\left|(X_i, V_i)\right|}. \quad (13.7)$$

Ясно, что $c_i \geq 1$. Если собственные векторы вещественны, то $c_i = \frac{1}{|\cos \varphi_i|}$, где φ_i

угол между векторами X_i и V_i .

Число c_i назовем коэффициентом перекоса матрицы A , соответствующим собственному значению λ_i . Таким образом, изменение λ_i при данной $\|dA\|$ может быть тем больше, чем больше соответствующий этому собственному значению коэффициент перекоса c_i . Для нормальных матриц, в частности, для эрмитовых и унитарных матриц $|d\lambda_i| \leq \|dA\|$, ибо для нормальных матриц $X_i = V_i$. Поэтому для нормальных матриц проблема определения собственных значений всегда устойчива. Для произвольных же матриц задача определения собственных значений будет не устойчивой только при большом коэффициенте перекоса.

Что же касается определения собственных векторов, то, как показывают формулы (13.5) и (13.6),

$$|dX_i| \leq \|dA\| \cdot |X_i| \cdot \sum_{j=1}^n \frac{c_j}{|\lambda_i - \lambda_j|}, \quad (13.8)$$

так что задача может быть неустойчивой, только если велик хоть один коэффициент перекоса или если имеются близкие собственные значения.

Метод Крылова определения собственных значений. Идея А.Н. Крылова заключалась в предварительном преобразовании уравнения

$$\varphi(t) = \begin{vmatrix} a_{11} - t & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - t & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & a_{nn} - t \end{vmatrix} = 0 \quad (13.9)$$

в эквивалентное ему, вообще говоря, уравнение вида

$$D(t) = \begin{vmatrix} b_{11} - t & b_{12} & \dots & b_{1n} \\ b_{21} - t^2 & b_{22} & \dots & b_{2n} \\ \dots & \dots & \dots & \dots \\ b_{n1} - t^n & b_{n2} & \dots & b_{nn} \end{vmatrix} = 0, \quad (13.10)$$

развертывание которого по степеням t осуществляется, очевидно, значительно проще при разложении определителя по минорам 1-го столбца.

Равенство нулю определителя

$$\varphi(t) = \begin{vmatrix} a_{11} - t & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - t & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & a_{nn} - t \end{vmatrix} \quad (13.11)$$

есть необходимое и достаточное условие для того, чтобы система однородных уравнений:

$$\begin{cases} tx_1 = a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n, \\ tx_2 = a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n, \\ \dots \\ tx_n = a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n \end{cases} \quad (13.12)$$

имела решение $x_1, x_2, \dots, x_n$, отличное от нулевого.

Преобразуем систему (13.12) следующим образом. Умножим первое выражение на t и заменим $tx_1, tx_2, \dots, tx_n$ их выражениями (13.12) через $x_1, x_2, \dots, x_n$.

Это дает

$$t^2 x_1 = b_{21}x_1 + b_{22}x_2 + \dots + b_{2n}x_n, \quad (13.13)$$

где

$$b_{2k} = \sum_{s=1}^n a_{1s} a_{sk}. \quad (13.14)$$

Умножим далее выражение (13.13) на t и заменим снова $tx_1, tx_2, \dots, tx_n$ их выражениями через $x_1, x_2, \dots, x_n$. Мы получим $t^3 x_1 = b_{31}x_1 + b_{32}x_2 + \dots + b_{3n}x_n$. Повторяя этот процесс $(n-1)$ раз, мы перейдем от системы (13.14) к системе

$$\begin{cases} tx_1 = b_{11}x_1 + b_{12}x_2 + \dots + b_{1n}x_n, \\ t^2 x_1 = b_{21}x_1 + b_{22}x_2 + \dots + b_{2n}x_n, \\ \dots \\ t^n x_1 = b_{n1}x_1 + b_{n2}x_2 + \dots + b_{nn}x_n, \end{cases} \quad (13.15)$$

коэффициенты которой b_{ik} будут определяться по рекуррентным формулам

$$b_{ik} = a_{ik}, \quad b_{ik} = \sum_{s=1}^n b_{i-1,s} a_{sk}, \quad (i = 2, \dots, n; k = 1, \dots, n). \quad (13.16)$$

Очевидно, что определитель системы (13.15) будет иметь вид (13.10). Система (13.15) имеет ненулевое решение для всех значений t , удовлетворяющих уравнению $\varphi(t) = 0$. Таким образом, $D(t)$ обращается в нуль при всех t , являющихся корнями уравнения $\varphi(t) = 0$. Покажем, что

$$\frac{D(t)}{\varphi(t)} = \begin{vmatrix} 1 & 0 & \dots & 0 \\ b_{11} & b_{12} & \dots & b_{1n} \\ \dots & \dots & \dots & \dots \\ b_{n-1,1} & b_{n-1,2} & \dots & b_{n-1,n} \end{vmatrix} = N \quad (13.17)$$

т.е. при $N \neq 0$ $D(t)$ отличается от искомого характеристического полинома только численным множителем.

Пусть все корни $\varphi(t)$ различны. Так как все корни $\varphi(t)$ являются корнями $D(t)$, то $D(t)$ делится на $\varphi(t)$. Так как, кроме того, степени $D(t)$ и $\varphi(t)$ одинаковы, частное должно быть постоянным (не зависеть от t). Сравнивая коэффициенты при t^n , получим $\frac{D(t)}{\varphi(t)} = N$.

В случае, если $\varphi(t)$ имеет кратные корни, равенство

$$D(t) = N\varphi(t) \quad (13.18)$$

сохраняется, что следует хотя бы из соображений непрерывности.

Можно проверить это равенство и непосредственно умножением входящих в него определителей, если при этом использовать соотношения (13.16).

Из равенства (13.18) видно, что если $N = 0$, то $D(t)$ тождественно равно нулю.

В этом случае указанное преобразование ничего не дает. Однако и при $N = 0$ А.Н. Крылов предлагает особый прием, алгебраическая сущность которого будет выяснена ниже.

Обратимся теперь к коэффициентам b_{ik} , определяющим $D(t)$. Введем в рассмотрение векторы B_i с компонентами $b_{i1}, b_{i2}, \dots, b_{in}$. Равенства

$$b_{ik} = \sum_{s=1}^n b_{i-1,s} a_{sk} \quad (i = 2, \dots, n) \text{ показывают, что}$$

$$B_i = A^T B_{i-1}, \quad (13.19)$$

где A^T – матрица, транспонированная к данной.

Из равенства (13.19) следует, что $B_i = (A^T)^{i-1} B_1$, ($i = 2, \dots, n$).

В свою очередь $B_1 = A^T B_0$, где $B_0 = (1, 0, \dots, 0)^T$. Таким образом, окончательно,

$$B_i = (A^T)^i B_0 \quad (i=1,2,\dots,n). \quad (13.20)$$

Очевидно, что преобразовывать систему (13.12) можно, исходя, например, из второго уравнения этой системы. В этом случае t войдет во второй столбец определителя $D(t)$, а коэффициенты b_{ik} будут определяться по формулам (13.20), где $B_0 = (1, 0, \dots, 0)^T$.

Метод А.Н. Крылова естественным образом обобщается, если ввести в рассмотрение вместо вектора B_0 специального вида произвольный вектор $B_0 = (b_{01}, b_{02}, \dots, b_{0n})^T$.

Пусть

$$u = b_{01}x_1 + b_{02}x_2 + \dots + b_{0n}x_n, \quad (13.21)$$

где $x_1, x_2, \dots, x_n$ решение системы (13.12).

Тогда, повторяя прежние рассуждения, получим

$$\begin{aligned} u &= b_{01}x_1 + b_{02}x_2 + \dots + b_{0n}x_n \\ tu &= b_{11}x_1 + b_{12}x_2 + \dots + b_{1n}x_n \\ t^2u &= b_{21}x_1 + b_{22}x_2 + \dots + b_{2n}x_n \\ &\vdots \\ t^nu &= b_{n1}x_1 + b_{n2}x_2 + \dots + b_{nn}x_n, \end{aligned} \tag{13.22}$$

где $B_1 = (b_{i1}, b_{i2}, \dots, b_{in})^T = (A^T)^i B_0$.

Рассматривая $(n+1)$ равенство (13.22) как систему линейных однородных уравнений с $n+1$ неизвестными $u, x_1, \dots, x_n$, получим, что ненулевое решение возможно в том и только в том случае, когда определитель

$$D(t) = \begin{vmatrix} 1 & b_{01} & \dots & b_{0n} \\ t & b_{11} & \dots & b_{1n} \\ \dots & \dots & \dots & \dots \\ t^n & b_{n1} & \dots & b_{nn} \end{vmatrix} = 0. \quad (13.23)$$

Повторяя прежние рассуждения, найдем, что $D(t) = \varphi(t)N$, где на этот раз

$$N = \begin{vmatrix} b_{01} & b_{02} & \dots & b_{0n} \\ b_{11} & b_{12} & \dots & b_{1n} \\ \dots & \dots & \dots & \dots \\ b_{n-1,1} & b_{n-1,2} & \dots & b_{n-1,n} \end{vmatrix} \quad (13.24)$$

Так же как и для рассмотренного выше частного случая, преобразование ничего не дает, если $N=0$.

Предположим поэтому сначала, что $N \neq 0$. На основании равенства $D(t) = \varphi(t)N$, коэффициенты p_i характеристического полинома определяются

как отношения $p_i = \frac{(-1)^{n-1} N_i}{N}$, где N_i суть алгебраические дополнения

элементов t^{n-i} в определителе $D(t)$. Определение коэффициентов характеристического полинома через указанные отношения и составляет сущность работы А.Н.Крылова. Однако проведенное исследование дает возможность определить искомые коэффициенты, минуя вычисления миноров, сократив при этом число операций.

Действительно, ввиду того, что элементы строк определителя (13.24) являются компонентами векторов $B_0, B_1, \dots, B_{n-1}$, условие $N \neq 0$ равносильно линейной независимости этих векторов. Поэтому при $N \neq 0$ векторы $B_0, B_1, \dots, B_{n-1}$

образуют базис пространства. Следовательно, вектор B_n является их линейной комбинацией:

$$B_n = q_1 B_{n-1} + \dots + q_n B_0. \quad (13.25)$$

Покажем, что коэффициенты этого соотношения являются коэффициентами p_i характеристического полинома, записанного в виде

$$\varphi(t) = (-1)^n [t^n - p_1 t^{n-1} - \dots - p_n].$$

Действительно, отняв от последней строки определителя $D(t)$ линейную комбинацию предыдущих строк с соответствующими коэффициентами $q_1, q_2, \dots, q_n$, получим, на основании равенства (13.25), что

$$D(t) = \begin{vmatrix} 1 & b_{01} & \dots & b_{0n} \\ \dots & \dots & \dots & \dots \\ t^{n-1} & b_{n-1,1} & \dots & b_{n-1,n} \\ t^n - q_1 t^{n-1} - \dots - q_n & 0 & \dots & 0 \end{vmatrix} = (-1)^n [t^n - q_1 t^{n-1} - \dots - q_n] N.$$

Отсюда $\varphi(t) = \frac{D(t)}{N} = (-1)^n [t^n - q_1 t^{n-1} - \dots - q_n]$, т.е. $q_i = p_i, i = \overline{1, n}$ что и требовалось доказать.

Равенство (13.25) позволяет находить коэффициенты $q_1 = p_1, q_2 = p_2, \dots, q_n = p_n$ как решения системы линейных уравнений, эквивалентной этому векторному равенству.

Очевидно, что вместо системы (13.25) для определения коэффициентов p_i можно употреблять систему $C_n = p_1 C_{n-1} + \dots + p_n C_0$, где векторы C_n определяются равенствами $C_k = A^k C_0$.

В случае, если $N=0$, система, эквивалентная равенству (13.25), не дает возможности определить коэффициенты характеристического полинома, так как определитель этой системы как раз равен N .

Алгебраическая сущность упомянутого приема А.Н. Крылова заключается в том, что возможно определить коэффициенты полинома наименьшей степени $\theta(\lambda)$ такого, что $\theta(A)C_0 = 0$, т.е. коэффициенты минимального аннулирующего

полинома. Вообще говоря, это будет минимальный полином матрицы, и его корни будут совпадать со всеми корнями характеристического полинома, но будут иметь меньшую кратность. Однако при неудачном выборе вектора C_0 , вместо минимального полинома может получиться какой-либо его делитель, и тогда часть корней уравнения $|A - tE| = 0$ может быть потеряна. Как показали Н.Н. Лузин и И.Н. Хлодовский, в качестве полинома $\theta(t)$ при специальном выборе вектора C_0 можно получить любой делитель минимального полинома.

Если минимальный полином матрицы не совпадает с характеристическим полиномом, то $N \neq 0$ при любом выборе вектора C_0 . Действительно, в этом случае $\psi(A)C_0 = 0$, где $\psi(A)$ - минимальный многочлен матрицы, и, так как степень полинома $\psi(t)$ меньше n , векторы $C_0, AC_0, \dots, A^{n-1}C_0$ линейно зависимы. Что же касается дополнительного вырождения, то его можно избежать за счет изменения начального вектора C_0 .

Итак, метод А.Н. Крылова дает возможность определить коэффициенты характеристического полинома, если $N \neq 0$, или некоторого его делителя, вообще говоря, минимального полинома, если $N = 0$.

Практически обстоятельство $N = 0$ обнаружится само собой в процессе прямого хода при решении системы (13.25) по методу Гаусса. Именно, в части уравнений исключаются все коэффициенты одновременно, так что эти уравнения обратятся в тождества $0=0$. Эти уравнения (пусть число их равно $n - m$) нужно отбросить. В оставшейся системе надо отбросить $n - m$ последних столбцов, начиная со столбца свободных членов (т.е. со столбца из компонент вектора C_n). Последний из оставшихся столбцов, составленный из компонент вектора C_m , надо принять за свободный член новой системы. Решение системы дает коэффициенты *линейной* зависимости C_m от $C_0, C_1, \dots, C_{m-1}$, т.е. коэффициенты минимального аннулирующего вектор C_0 полинома.

Нахождение собственных векторов по методу Крылова. Предположим, что методом А.Н. Крылова найдены коэффициенты характеристического полинома, и что все собственные числа вычислены и оказались различными. Покажем, как, используя проведенные вычисления, определить собственные векторы матрицы. Пусть C_0 есть исходный вектор в процессе А.Н.Крылова и пусть $X_1, X_2, \dots, X_n$ собственные векторы матрицы A , принадлежащие $\lambda_1, \lambda_2, \dots, \lambda_n$.

Разложим вектор C_0 по собственным векторам:

$$C_0 = \alpha_1 X_1 + \alpha_2 X_2 + \dots + \alpha_n X_n. \quad (13.26)$$

Тогда

$$\begin{aligned} C_1 &= AC_0 = \alpha_1\lambda_1X_1 + \alpha_2\lambda_2X_2 + ... + \alpha_n\lambda_nX_n \\ \\ C_{n-1} &= A^{n-1}C_0 = \alpha_1\lambda_1^{n-1}X_1 + \alpha_2\lambda_2^{n-1}X_2 + ... + \alpha_n\lambda_n^{n-1}X_n. \end{aligned} \tag{13.27}$$

Векторы $C_1, C_2, \dots, C_{n-1}$ вычислены в процессе нахождения собственных чисел. Покажем, что собственные векторы могут быть найдены в виде линейных комбинаций $x_i = \beta_{i0}C_{n-1} + \beta_{i1}C_{n-2} + \dots + \beta_{i,n-1}C_0$ ($i=1, 2, \dots, n$) при подходящем выборе коэффициентов β_{ij} .

Рассмотрим линейную комбинацию

$$\begin{aligned}
&\beta_{10}C_{n-1} + \beta_{11}C_{n-2} + \dots + \beta_{1,n-1}C_0 = \\
&\alpha_1\left(\beta_{10}\lambda_1^{n-1} + \beta_{11}\lambda_1^{n-2} + \dots + \beta_{1,n-1}\right)X_1 + \\
&+\alpha_2\left(\beta_{10}\lambda_2^{n-1} + \beta_{11}\lambda_2^{n-2} + \dots + \beta_{1,n-1}\right)X_2 + \\
&+ \dots + \\
&+\alpha_n\left(\beta_{10}\lambda_n^{n-1} + \beta_{11}\lambda_n^{n-2} + \dots + \beta_{1,n-1}\right)X_n = \\
&\alpha_1\varphi_1(\lambda_1)X_1 + \alpha_2\varphi_1(\lambda_2)X_2 + \dots + \alpha_n\varphi_1(\lambda_n)X_n,
\end{aligned} \tag{13.28}$$

где

$$\varphi_1(t) = \beta_{10}t^{n-1} + \beta_{11}t^{n-2} + \dots + \beta_{1,n-1}. \quad (13.29)$$

Подберем коэффициенты $\beta_{10}, \dots, \beta_{1,n-1}$ так, чтобы

$$\varphi_1(\lambda_1) \neq 0, \varphi_1(\lambda_2) = \dots = \varphi_1(\lambda_n) = 0. \quad (13.30)$$

Для этого достаточно взять в качестве $\varphi_1(t)$ полином

$$\begin{aligned}\varphi_1(t) &= (t - \lambda_2) \dots (t - \lambda_n) = \frac{(t - \lambda_1)(t - \lambda_2) \dots (t - \lambda_n)}{(t - \lambda_1)} = \\ &= \frac{(-1)^n \varphi(t)}{(t - \lambda_1)} = \frac{t^n - p_1 t^{n-1} - \dots - p_n}{(t - \lambda_1)}.\end{aligned}\quad (13.31)$$

Здесь $\varphi(t)$ - характеристический полином, коэффициенты и корни которого уже вычислены.

Коэффициенты частного (13.31) легко вычисляются по схеме Горнера, т.е. по рекуррентным формулам

$$\beta_{10} = 1, \quad \beta_{1j} = \lambda_1 \beta_{1,j-1} - p_j, \quad j = 1, \dots, n-1. \quad (13.32)$$

Таким образом, $\beta_{10}C_{n-1} + \beta_{11}C_{n-2} + \dots + \beta_{1,n-1}C_0 = \alpha_1 \varphi_1(\lambda_1) X_1$, т.е. составленная нами линейная комбинация есть собственный вектор X_1 с точностью до численного множителя. Конечно, коэффициент α_1 должен быть отличным от нуля; это обеспечивается успешным завершением процесса А.Н. Крылова. Так как собственный вектор определен с точностью до постоянного множителя, мы можем принять за собственный вектор построенную линейную комбинацию.

Аналогично,

$$X_i = \sum_{j=0}^{n-1} \beta_{ij} C_{n-1-j}, \quad (13.33)$$

где

$$\beta_{i0} = 1, \quad \beta_{ij} = \lambda_i \beta_{i,j-1} - p_j, \quad j = 1, \dots, n-1. \quad (13.34)$$

13. Метод Леверье и его модификация Фаддеевым

Пусть

$$\varphi(t) = (-1)^n [t^n - p_1 t^{n-1} - p_2 t^{n-2} - \dots - p_n] \quad (14.1)$$

характеристический полином матрицы, и $\lambda_1, \lambda_2, \dots, \lambda_n$ его корни, среди которых могут быть равные. Обозначим

$$\sum_{l=1}^n \lambda_l^k = s_k. \quad (14.2)$$

Тогда справедливы соотношения, известные под названием формул Ньютона:

$$k p_k = s_k - p_1 s_{k-1} - \dots - p_{k-1} s_1 \quad (k=1, \dots, n). \quad (14.3)$$

Если числа s_k известны, то, решая рекуррентную систему (14.3), мы сможем найти нужные нам коэффициенты p_k .

Покажем, как определяются числа s_k . Имеем $s_1 = \lambda_1 + \lambda_2 + \dots + \lambda_n = \text{Tr} A$.

Характеристические числа матрицы A^k будут $\lambda_1^k, \lambda_2^k, \dots, \lambda_n^k$.

Следовательно,

$$s_k = \lambda_1^k + \lambda_2^k + \dots + \lambda_n^k = \text{Tr} A^k. \quad (14.4)$$

Таким образом, процесс вычисления сводится к последовательному вычислению степеней матрицы A , затем к вычислению их следов и, наконец, к решению рекуррентной системы (14.3). Вычисление n степеней матрицы A (у последней матрицы A^n нужно вычислить только диагональные элементы) требует большого числа операций, и потому метод Леверье является очень трудоемким.

Изложим теперь видоизменение метода, предложенное Д. К. Фаддеевым, которое кроме упрощений при вычислении коэффициентов характеристического полинома, позволит нам определить обратную матрицу и собственные векторы матрицы.

Будем вместо последовательности $A, A^2, \dots, A^n$ вычислять последовательность $A_1, A_2, \dots, A_n$, построенную следующим образом:

$$\begin{aligned} A_1 &= A, & Tr A_1 &= q_1, & B_1 &= A_1 - q_1 E \\ A_2 &= AB_1, & \frac{Tr A_2}{2} &= q_2, & B_2 &= A_2 - q_2 E \\ &\dots\dots\dots & & & & \\ A_{n-1} &= AB_{n-2}, & \frac{Tr A_{n-1}}{n-1} &= q_{n-1}, & B_{n-1} &= A_{n-1} - q_{n-1} E \\ A_n &= AB_{n-1}, & \frac{Tr A_n}{n} &= q_n, & B_n &= A_n - q_n E. \end{aligned} \tag{14.5}$$

Докажем, что

- a) $q_1 = p_1, q_2 = p_2, \dots, q_n = p_n;$

- б) B_n - нулевая матрица;

- в) если A неособенная матрица, то $A^{-1} = \frac{B_{n-1}}{p_n}$.

- а) Докажем методом математической индукции. Очевидно, что $p_1 = TrA = q_1$. Предположим, что $q_1 = p_1, q_2 = p_2, \dots, q_{k-1} = p_{k-1}$, и докажем, что $q_k = p_k$. Согласно нашей конструкции:

$$A_k = A^k - q_1 A^{k-1} - \dots - q_{k-1} A = A^k - p_1 A^{k-1} - \dots - p_{k-1} A.$$

Следовательно,

$$TrA_k = kq_k = TrA^k - p_1 TrA^{k-1} - \dots - p_{k-1} TrA = s_k - p_1 s_{k-1} - \dots - p_{k-1} s_1.$$

Отсюда, в силу формул Ньютона, $kq_k = kp_k$ и, следовательно, $q_k = p_k$.

- б) В силу соотношения Кели-Гамильтона (матрица является корнем своего характеристического многочлена),

$$B_n = A^n - p_1 A^{n-1} - \dots - p_n E = 0. \quad (14.6)$$

- в) Из равенства (14.6) следует, что $AB_{n-1} = A_n = B_n + p_n E = p_n E$, так что

$$A^{-1} = \frac{1}{p_n} B_{n-1}. \quad (14.7)$$

Перейдем теперь к определению собственных векторов матрицы A . Пусть собственные числа уже вычислены и при этом оказались различными. Построим матрицу

$$Q_k = \lambda_k^{n-1}E + \lambda_k^{n-2}B_1 + \dots + B_{n-1}, \quad (14.8)$$

где B_i - матрицы, вычисленные в процессе нахождения коэффициентов характеристического полинома, а λ_k есть k -е собственное число матрицы A .

Можно доказать, в предположении, что все $\lambda_1, \dots, \lambda_n$ различны, что матрица Q_k ненулевая.

Покажем, что каждый столбец матрицы Q_k состоит из компонент собственного вектора, принадлежащего собственному числу λ_k .

Действительно,

$$\begin{aligned} (\lambda_k E - A)Q_k &= (\lambda_k E - A)(\lambda_k^{n-1}E + \lambda_k^{n-2}B_1 + \dots + B_{n-1}) = \\ &= \lambda_k^n E + \lambda_k^{n-1}(B_1 - A) + \lambda_k^{n-2}(B_2 - AB_1) + \dots - AB_{n-1} = \\ &= \lambda_k^n E - p_1 \lambda_k^{n-1} E - p_2 \lambda_k^{n-2} E - \dots - p_n E = 0. \end{aligned}$$

Отсюда следует, что $(\lambda_k E - A)u = 0$, где u любой столбец построенной матрицы Q_k , т.е. что

$$\lambda_k u = Au. \quad (14.9)$$

Это равенство показывает, что u есть собственный вектор.

14. Эскалаторный метод

Оригинальный метод определения собственных чисел и собственных векторов матрицы известен под названием эскалаторного метода. Этот метод дает индуктивную конструкцию, посредством которой, зная собственные числа и собственные векторы матрицы A_{k-1} и ее транспонированной, можно составить уравнение для определения собственных чисел матрицы A_k , полученной из A_{k-1} окаймлением, и затем вычислить по несложным формулам компоненты собственных векторов для матрицы A_k и ее транспонированной. Применение эскалаторного метода начинается с отыскания собственных векторов матрицы 2-го порядка. Эта задача решается совсем просто.

Метод основан на использовании свойств ортогональности собственных векторов матрицы и ее транспонированной. Ограничимся рассмотрением перехода от матрицы 3-го порядка к матрице 4-го порядка. Для удобства компоненты векторов будем обозначать, различными буквами. Предположим, что все собственные числа матрицы A_3 вещественны и различны.

Итак, пусть λ_r ($r=1;2;3$) - собственные числа матрицы A_3 и A_3' , где

$$A_3 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \quad (15.1)$$

а A_3' - матрица, транспонированная к A .

Пусть далее $X_r = (x_r, y_r, z_r)$ и $X_r' = (x_r', y_r', z_r')$ ($r=1;2;3$) - совокупность собственных векторов этих матриц.

Эти собственные векторы могут быть нормированы так, что

$$\begin{bmatrix} x_1' & x_2' & x_3' \\ y_1' & y_2' & y_3' \\ z_1' & z_2' & z_3' \end{bmatrix} \cdot \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{bmatrix} = \begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{bmatrix} \cdot \begin{bmatrix} x_1' & x_2' & x_3' \\ y_1' & y_2' & y_3' \\ z_1' & z_2' & z_3' \end{bmatrix} = E. \quad (15.2)$$

Пусть A_4 - матрица 4-ого порядка, полученная из A_3 окаймлением, $X_r = (x, y, z, u)$ её собственный вектор, принадлежащий собственному числу λ .

Имеем

$$\begin{cases} \lambda_x = a_{11}x + a_{12}y + a_{13}z + a_{14}u, \\ \lambda_y = a_{21}x + a_{22}y + a_{23}z + a_{24}u, \\ \lambda_z = a_{31}x + a_{32}y + a_{33}z + a_{34}u, \\ \lambda_u = a_{41}x + a_{42}y + a_{43}z + a_{44}u \end{cases} \quad (15.3)$$

Умножим первые три уравнения системы (15.3), соответственно, на x'_r, y'_r, z'_r и сложим. Получим

$$\begin{aligned} \lambda (xx'_r + yy'_r + zz'_r) = & (a_{11}x'_r + a_{21}y'_r + a_{31}z'_r)x + (a_{12}x'_r + a_{22}y'_r + a_{32}z'_r)y + \\ & + (a_{13}x'_r + a_{23}y'_r + a_{33}z'_r)z + (a_{14}x'_r + a_{24}y'_r + a_{34}z'_r)u, \end{aligned}$$

откуда в силу того, что (x'_r, y'_r, z'_r) есть собственный вектор для матрицы A'_3

$$\lambda (xx'_r + yy'_r + zz'_r) = \lambda_r (xx'_r + yy'_r + zz'_r)x + (a_{14}x'_r + a_{24}y'_r + a_{34}z'_r)u,$$

и, следовательно,

$$xx'_r + yy'_r + zz'_r = -\frac{P'_r u}{\lambda_r - \lambda}, \quad (15.4)$$

где

$$P'_r = a_{14}x'_r + a_{24}y'_r + a_{34}z'_r. \quad (15.5)$$

Пусть

$$P_r = a_{41}x_r + a_{42}y_r + a_{43}z_r. \quad (15.6)$$

Тогда, в силу свойств ортогональности (15.2), будет справедливо следующее соотношение:

$$\sum_{r=1}^3 P_r (x'_r x + y'_r y + z'_r z) = P, \quad (15.7)$$

где

$$P = a_{41}x + a_{42}y + a_{43}z = -(a_{44} - \lambda)x. \quad (15.8)$$

Действительно,

$$\begin{aligned} & \sum_{r=1}^3 (a_{41}x_r + a_{42}y_r + a_{43}z_r) (x_r'x + y_r'y + z_r'z) = \\ & = a_{41} (x_1x_1' + x_2x_2' + x_3x_3')x + a_{42} (x_1y_1' + x_2y_2' + x_3y_3')y + a_{43} (x_1z_1' + x_2z_2' + x_3z_3')z + \dots = \\ & = a_{41}x + a_{42}y + a_{43}z. \end{aligned}$$

Заменив в уравнении (15.7) выражение $x_r'x + y_r'y + z_r'z$ на $-\frac{P_r'u}{\lambda_r - \lambda}$, согласно (15.4), получим следующее уравнение для определения собственных чисел матрицы A_4 :

$$a_{44} - \lambda = \sum_{r=1}^3 \frac{P_r P_r'}{\lambda_r - \lambda} \quad (15.9)$$

Это уравнение называют эскалаторной формой характеристического уравнения или эскалаторным уравнением.

Далее, умножая (15.4) последовательно на x_r, y_r, z_r ($r=1,2,3$) и складывая, получим, опять принимая во внимание свойства ортогональности (15.2):

$$\frac{x}{u} = -\sum_{r=1}^3 \frac{P_r'x_r}{\lambda_r - \lambda}; \quad \frac{y}{u} = -\sum_{r=1}^3 \frac{P_r'y_r}{\lambda_r - \lambda}; \quad \frac{z}{u} = -\sum_{r=1}^3 \frac{P_r'z_r}{\lambda_r - \lambda}. \quad (15.10)$$

Аналогично

$$\frac{x'}{u'} = -\sum_{r=1}^3 \frac{P_r x_r'}{\lambda_r - \lambda}; \quad \frac{y'}{u'} = -\sum_{r=1}^3 \frac{P_r y_r'}{\lambda_r - \lambda}; \quad \frac{z'}{u'} = -\sum_{r=1}^3 \frac{P_r z_r'}{\lambda_r - \lambda}. \quad (15.11)$$

Таким образом, найдя собственное число λ из уравнения (15.9), мы находим по формулам (15.10) и (15.11) собственные вектора матриц A_4 и A_4' , принадлежащие этому числу, с точностью до численного множителя. Для возможности продолжения процесса мы должны еще нормировать их в смысле

формулы (15.2). Нетрудно проверить (снова используя свойства (15.2)), что

$$\frac{xx' + yy' + zz'}{uu'} = \sum_{r=1}^3 \frac{P_r P'_r}{(\lambda_r - \lambda)^2}.$$

Следовательно, $\frac{xx' + yy' + zz' + uu'}{uu'} = 1 + \sum_{r=1}^3 \frac{P_r P'_r}{(\lambda_r - \lambda)^2}$. Таким образом, мы

удовлетворим условию нормированности при $\frac{1}{uu'} = 1 + \sum_{r=1}^3 \frac{P_r P'_r}{(\lambda_r - \lambda)^2}$.

Заметим, что если левую часть эскалаторного уравнения обозначить через $f(\lambda)$,

$$f(\lambda) = -a_{44} + \lambda + \sum_{r=1}^3 \frac{P_r P'_r}{(\lambda_r - \lambda)}, \quad (15.12)$$

то $f'(\lambda) = 1 + \sum_{r=1}^3 \frac{P_r P'_r}{(\lambda_r - \lambda)^2}$. Таким образом $\frac{1}{uu'} = f'(\lambda)$.

Не теряя общности, мы можем считать, что $u = \pm u'$, выбирая знак так, чтобы $\frac{1}{u^2} = \pm f'(\lambda)$ было положительным. Это даёт $u = u' = \frac{1}{\sqrt{f'(\lambda)}}$, если

$f'(\lambda) > 0$; $u = -u' = \frac{1}{\sqrt{-f'(\lambda)}}$, если $f'(\lambda) < 0$.

В заключение отметим контрольные равенства

$$\begin{aligned} \sum_{r=1}^3 P_r x'_r &= a_{41}; \quad \sum_{r=1}^3 P_r y'_r = a_{42}; \quad \sum_{r=1}^3 P_r z'_r = a_{43}; \\ \sum_{r=1}^3 P'_r x_r &= a_{14}; \quad \sum_{r=1}^3 P'_r y_r = a_{24}; \quad \sum_{r=1}^3 P'_r z_r = a_{34}. \end{aligned}$$

Последние равенства показывают, что все «новые» элементы матрицы A_4 употребляются нами для контроля. Хорошее совпадение контрольных формул гарантирует правильность вычислений на каждом шагу. При окончании процесса полезно проверить для найденных векторов выполнение условия ортогональности.

16. Определение наибольшего по модулю собственного значения и принадлежащего ему собственного вектора

Изложим метод, позволяющий вычислять наибольшее по модулю собственное значение матрицы и принадлежащий ему собственный вектор при помощи вычисления последовательности итераций произвольного вектора. Излагаемый метод называется степенным и является простейшим итерационным процессом для решения частичной проблемы собственных значений. Он применим для произвольной матрицы, хотя ход итерационного процесса существенно зависит от того, как входит наибольшее по модулю собственное значение матрицы в ее каноническую форму Жордана. В связи с этим приходится различать несколько возможных случаев.

Для упрощения изложения мы будем предполагать, что все собственные значения матрицы, кроме, может быть, наибольшего по модулю, имеют линейные элементарные делители. Будем также считать, что элементы исследуемых матриц вещественны.

Наибольшее по модулю собственное значение вещественное и простое. В этом случае наибольшему по модулю собственному значению соответствует один линейный элементарный делитель, так что все элементарные делители матрицы линейны. Поэтому существует базис из собственных векторов $U_1, U_2, \dots, U_n$, принадлежащих собственным значениям $\lambda_1, \lambda_2, \dots, \lambda_n$, расположенным в порядке убывания модулей, причём $|\lambda_1| > |\lambda_2|$, но среди остальных могут быть равные. Возьмём произвольный вектор Y_0 и образуем последовательность его итераций матрицей A : $AY_0, A^2Y_0, \dots, A^kY_0, \dots$

Напишем разложение вектора Y_0 по собственным векторам

$$Y_0 = a_1U_1 + a_2U_2 + \dots + a_nU_n \quad (16.1)$$

Среди чисел a_i некоторые могут равняться нулю. Предположим, однако, что $a_1 \neq 0$. Очевидно, что

$$(16.2)$$

Обозначим $A^k Y_0 = Y_k = (y_{1k}, y_{2k}, \dots, y_{nk})^T$ и выясним структуру компонент

$$U_1 = (u_{11}, u_{21}, \dots, u_{n1})^T$$

$$U_2 = (u_{12}, u_{22}, \dots, u_{n2})^T$$

$$U_n = (u_{1n}, u_{2n}, \dots, u_{nn})^T$$

Тогда из (16.2) получим $y_{ik} = a_1 u_{i1} \lambda_1^k + a_2 u_{i2} \lambda_2^k + \dots + a_n u_{in} \lambda_n^k$.

Коэффициент при λ_1^k по крайней мере в одной из компонент не равен

$$y_k = c_1 \lambda_1^k + c_2 \lambda_2^k + \dots + c_n \lambda_n^k, \quad (16.3)$$

Причём коэффициент c_i не зависит от индекса k и $c_1 \neq 0$. Рассмотрим

$$(16.4)$$

где

$$(16.5)$$

Произведя деление и удерживая члены до порядка α_2^{2k} и α_3^k

ВКЛЮЧИТЕЛЬНО, ПОЛУЧИМ

$$\frac{y_{k+1}}{y_k} = \lambda_1 \left[1 - b'_2 \alpha_2^k - b'_3 \alpha_3^k + b_2 b'_2 \alpha_3^{2k} \right] + o(\alpha_3^k + \alpha_2^{2k}). \quad (16.6)$$

где

$$\begin{aligned} b'_2 &= b_2 (1 - \alpha_2), \\ b'_3 &= b_3 (1 - \alpha_3) \end{aligned} \quad (16.7)$$

Отсюда мы видим, что если k достаточно велико, то

$$\lambda_1 \approx \frac{y_{k+1}}{y_k}. \quad (16.8)$$

Так как обычно все компоненты вектора U_1 отличны от нуля, то в качестве y_k может быть взята, как правило, любая компонента вектора Y_k . Таким образом, первое собственное значение приближенно равно отношению любых соответствующих компонент двух соседних достаточно высоких итераций произвольного вектора матрицей A .

При практическом выполнении итераций следует вычислять отношения $\frac{y_{k+1}}{y_k}$ для нескольких компонент. Хорошее совпадение этих отношений будет показывать, что в выражении (16.6) различие значений коэффициентов b'_2, b'_3 уже перестало играть заметную роль.

Быстрота сходимости процесса в рассматриваемом случае определяется величиной отношения $\frac{\lambda_2}{\lambda_1}$ и может быть медленной, если это отношение близко к единице.

Для того чтобы избежать роста компонент, иногда целесообразно при вычислении итерации тем или другим способом нормировать на каждом шагу получаемые векторы. Удобными нормировками являются деление вектора на его первую компоненту, или на наибольшую компоненту, или, наконец, нормировка к единичной длине. При этом вместо последовательности Y_k мы получим последовательность $\tilde{Y}_k = p_k Y_k$, где p_k есть нормирующие множители, и для получения λ_1 надо брать отношения компонент векторов $A\tilde{Y}_k$ и $\tilde{Y}_k$.

Описанный процесс даёт возможность определить также и все компоненты собственного вектора, принадлежащего наибольшему собственному числу. Именно, отношения компонент вектора Y_k стремятся к отношениям компонент этого собственного вектора.

Действительно, при $a_1 \neq 0$.

$$\begin{aligned} Y_k = A^k Y_0 &= \lambda_1^k \left[a_1 U_1 + a_2 \left(\frac{\lambda_2}{\lambda_1} \right) U_2 + \dots + a_n \left(\frac{\lambda_n}{\lambda_1} \right)^k U_n \right] = \\ &= a_1 \lambda_1^k \left[U_1 + O \left(\frac{\lambda_2}{\lambda_1} \right)^k \right]. \end{aligned} \quad (16.9)$$

Наибольшее по модулю собственное значение вещественное, кратное.

В этом случае формула (16.3) остаётся верной, но несколько первых членов можно соединить вместе, так что

$$y_k = c_1 \lambda_1^k + c_{r+1} \lambda_{r+1}^k + \dots + c_n \lambda_n^k,$$

где r - кратность λ_1 . Все дальнейшие рассуждения остаются в силе и потому

$$\frac{y_{k+1}}{y_k} = \lambda_1 + O \left(\frac{\lambda_{r+1}}{\lambda_1} \right)^k. \quad (16.10)$$

Таким образом, и в этом случае, при условии $a_1 \neq 0$ отношение $\frac{y_{k+1}}{y_k}$ даёт

приближённое значение наибольшего собственного числа.

Вопрос о кратности корня не может быть решён без более детального исследования.

Векторы $Y_k = A^k Y_0$, так же как и в предыдущем случае, сходятся по направлению к одному из собственных векторов, принадлежащих λ_1 , именно к собственному вектору, лежащему в циклическом подпространстве, порожденном вектором Y_0 . Исходя из различных начальных векторов, мы придём, вообще говоря, к различным собственным векторам.

Два наибольшие по модулю собственные значения вещественны и противоположны по знаку. Из равенства (16.3) мы видим, что в этом случае

четные и нечетные итерации имеют различные коэффициенты при соответствующих степенях λ_1 , так как

$$y_{2k} = (c_1 + c_2)\lambda_1^{2k} + c_3\lambda_3^{2k} + \dots + c_n\lambda_n^{2k},$$

$$y_{2k+1} = (c_1 - c_2)\lambda_1^{2k+1} + c_3\lambda_3^{2k+1} + \dots + c_n\lambda_n^{2k+1},$$

и потому две соседние итерации не могут быть использованы для определения λ_1 .

Однако мы можем определить λ_1^2 по одной из следующих формул:

$$\lambda_1^2 = \frac{y_{2k+2}}{y_{2k}} \text{ или } \lambda_1^2 = \frac{y_{2k+1}}{y_{2k-1}} \quad (16.11)$$

Для нахождения собственных векторов, принадлежащих λ_1 и $\lambda_2 = -\lambda_1$, целесообразно построить векторы $Y_k + \lambda_1 Y_{k-1}$ и $Y_k - \lambda_1 Y_{k-1}$. Отношения компонент этих векторов будут стремиться, соответственно, к отношению компонент векторов U_1 и U_2 , принадлежащих собственным числам λ_1 и λ_2 .

Действительно, в силу равенства

$$Y_k = a_1\lambda_1^k U_1 + a_2(-\lambda_1)^k U_2 + a_3\lambda_3^k U_3 + \dots \quad (16.12)$$

имеем

$$\begin{aligned} Y_k + \lambda_1 Y_{k-1} &= 2a_1\lambda_1^k U_1 + a_3(\lambda_3 + \lambda_1)\lambda_3^{k-1} U_3 + \dots = \\ &= \lambda_1^k \left[2a_1 U_1 + O\left(\frac{\lambda_3}{\lambda_1}\right)^k \right], \\ Y_k - \lambda_1 Y_{k-1} &= 2a_2(-\lambda_1)^k U_2 + a_3(\lambda_3 - \lambda_1)\lambda_3^{k-1} U_3 + \dots = \\ &= (-\lambda_1)^k \left[2a_2 U_2 + O\left(\frac{\lambda_3}{\lambda_1}\right)^k \right]. \end{aligned} \quad (16.13)$$

Наибольшие по модулю собственные значения образуют комплексную пару. Пусть λ_1 и λ_2 комплексно-сопряженные, наибольшие по модулю собственные значения и $|\lambda_3| < |\lambda_1|$. Согласно формуле (16.3)

$$y_k = c_1\lambda_1^k + c_2\lambda_2^k + \dots + c_n\lambda_n^k,$$

причём в этом случае c_1 и c_2 комплексно сопряжены. Пусть

$$c_1 = \operatorname{Re}^{i\alpha}, c_2 = \operatorname{Re}^{-i\alpha}, \lambda_1 = re^{i\theta}, \lambda_2 = re^{-i\theta}, \quad (16.14)$$

тогда

$$y_k = 2Rr^k \cos(k\theta + \alpha) + c_3 \lambda_3^k. \quad (16.15)$$

Присутствие множителя $\cos(k\theta + \alpha)$ будет причиной того, что значения y_k будут сильно колебаться как по величине, так и по знаку.

Таким образом, наличие комплексных корней, наибольших по модулю, сразу обнаруживается при составлении итераций. Положим

$$p = -(\lambda_1 + \lambda_2) = -2r \cos \theta, \quad q = \lambda_1 \lambda_2 = r^2. \quad (16.16)$$

Тогда λ_1 и λ_2 будут корнями квадратного уравнения $t^2 + pt + q = 0$. Коэффициенты p и q могут быть определены из следующих соображений.

Пусть k настолько велико, что $y_k \approx c_1 \lambda_1^k + c_2 \lambda_2^k$. Тогда

$$y_{k+1} + p y_k + q y_{k-1} \approx c_1 \lambda_1^{k-1} [\lambda_1^2 + p \lambda_1 + q] + c_2 \lambda_2^{k-1} [\lambda_2^2 + p \lambda_2 + q] = 0. \quad (16.17)$$

Здесь приближённое равенство справедливо с точностью до $O(|\lambda_3|^k)$.

Аналогичное приближённое равенство

$$z_{k-1} + p z_k + q z_{k-1} \approx 0 \quad (16.18)$$

будет справедливо для любой другой компоненты z_k вектора $Y_k = A^k Y_0$. В качестве z_k можно также взять $z_k = y_{k+1}$ или, в более общем виде, любую компоненту вектора $Z_k = A^k Z_0$, где Z_0 - произвольный начальный вектор.

Из равенств (16.17) и (16.18) получим, вообще говоря,

$$p \approx -\frac{y_{k-1} z_{k+1} - z_{k-1} y_{k+1}}{y_{k-1} z_k - z_{k-1} y_k}, \quad q \approx \frac{y_k z_{k+1} - z_k y_{k+1}}{y_{k-1} z_k - z_{k-1} y_k}. \quad (16.19)$$

В частности, если взять $z_k = y_{k+1}$, получим

$$p \approx -\frac{y_{k-1} y_{k+2} - y_k y_{k+1}}{y_{k-1} y_{k+1} - y_k^2}, \quad q \approx \frac{y_k y_{k+2} - y_{k+1}^2}{y_{k-1} y_{k+1} - y_k^2}. \quad (16.20)$$

Дадим более строгое обоснование формул (16.19) и (16.20). Это позволит выяснить условия их применимости и оценить погрешность.

Пусть

$$y_k = c_1 \lambda_1^k + c_2 \lambda_2^k + O(|\lambda_3|^k),$$

$$z_k = d_1 \lambda_1^k + d_2 \lambda_2^k + O(|\lambda_3|^k).$$

Тогда

$$\begin{aligned} y_{k-1} z_{k+1} - z_{k-1} y_{k+1} &= \left(c_1 \lambda_1^{k-1} + c_2 \lambda_2^{k-1} + O(|\lambda_3|^k) \right) \left(d_1 \lambda_1^{k+1} + d_2 \lambda_2^{k+1} + O(|\lambda_3|^k) \right) - \\ &\quad - \left(c_1 \lambda_1^{k+1} + c_2 \lambda_2^{k+1} + O(|\lambda_3|^k) \right) \left(d_1 \lambda_1^{k-1} + d_2 \lambda_2^{k-1} + O(|\lambda_3|^k) \right) = \\ &= (c_1 d_2 - c_2 d_1) (\lambda_2^2 - \lambda_1^2) \lambda_1^{k-1} \lambda_2^{k-1} + O(|\lambda_1^k \lambda_3^k|). \end{aligned}$$

Аналогично

$$y_{k-1} z_k - z_{k-1} y_k = (c_1 d_2 - c_2 d_1) (\lambda_2 - \lambda_1) \lambda_1^{k-1} \lambda_2^{k-1} + O(|\lambda_1^k \lambda_3^k|),$$

$$y_k z_{k+1} - z_k y_{k+1} = (c_1 d_2 - c_2 d_1) (\lambda_2 - \lambda_1) \lambda_1^k \lambda_2^k + O(|\lambda_1^k \lambda_3^k|).$$

Поэтому
$$-\frac{y_{k-1} z_{k+1} - z_{k-1} y_{k+1}}{y_{k-1} z_k - z_{k-1} y_k} = -(\lambda_1 + \lambda_2) + O\left(\left|\frac{\lambda_3}{\lambda_2}\right|^k\right) = p + O\left(\left|\frac{\lambda_3}{\lambda_2}\right|^k\right),$$

$$\frac{y_k z_{k+1} - z_k y_{k+1}}{y_{k-1} z_k - z_{k-1} y_k} = \lambda_1 \lambda_2 + O\left(\left|\frac{\lambda_3}{\lambda_2}\right|^k\right) = q + O\left(\left|\frac{\lambda_3}{\lambda_2}\right|^k\right), \text{ если только } c_1 d_2 - c_2 d_1 \neq 0.$$

Аналогично

$$-\frac{y_{k-1} y_{k+2} - y_k y_{k+1}}{y_{k-1} y_{k+1} - y_k^2} = p + O\left(\left|\frac{\lambda_3}{\lambda_2}\right|^k\right), \quad \frac{y_k y_{k+2} - y_{k+1}^2}{y_{k-1} y_{k+1} - y_k^2} = q + O\left(\left|\frac{\lambda_3}{\lambda_2}\right|^k\right). \quad (16.21)$$

Отметим, что для формул (16.21) условие $c_1 d_2 - c_2 d_1 \neq 0$ всегда выполняется, ибо $d_1 = \lambda_1 c_1$, $d_2 = \lambda_2 c_2$, так что $c_1 d_2 - c_2 d_1 = c_1 c_2 (\lambda_1 - \lambda_2) \neq 0$.

Собственные значения λ_1 и λ_2 можно определить, минуя вычисление p и решение квадратного уравнения. Именно, определив $q = r^2$ по одной из формул (16.19) или (16.20), найдём r и вычислим выражение

$$\begin{aligned}\mu_k &= \frac{1}{2} \left[r y_{k-1} + r^{-1} y_{k+1} \right] \approx \frac{1}{2} \left[r (c_1 \lambda_1^{k-1} + c_2 \lambda_2^{k-1}) + r^{-1} (c_1 \lambda_1^{k+1} + c_2 \lambda_2^{k+1}) \right] = \\ &= \frac{1}{2} \left[c_1 r^k e^{ik\theta} + c_2 r^k e^{-ik\theta} \right] \left[e^{-i\theta} + e^{i\theta} \right] = y_k \cos \theta.\end{aligned}\quad (16.22)$$

Отсюда находим

$$\cos \theta \approx \frac{\mu_k}{y_k} \quad (16.23)$$

с точностью до величин порядка $\left(\frac{\lambda_3}{\lambda_1} \right)^k$.

После того как собственные значения λ_1 и λ_2 определены, соответствующие им собственные векторы легко определяются. Именно, из приближённых равенств

$$\begin{aligned}Y_k &\approx a_1 \lambda_1^k U_1 + a_2 \lambda_2^k U_2, \\ Y_{k+1} &\approx a_1 \lambda_1^{k+1} U_1 + a_2 \lambda_2^{k+1} U_2\end{aligned}$$

находим

$$\begin{aligned}Y_{k+1} - \lambda_2 Y_k &\approx a_1 \lambda_1^k (\lambda_1 - \lambda_2) U_1, \\ Y_{k+1} - \lambda_1 Y_k &\approx a_2 \lambda_2^k (\lambda_2 - \lambda_1) U_2,\end{aligned}\quad (16.24)$$

откуда следует, что $Y_{k+1} - \lambda_2 Y_k$ и $Y_{k+1} - \lambda_1 Y_k$ с точностью до малых слагаемых, являются собственными векторами, соответствующими собственными значениями λ_1 и λ_2 .

16. Метод скалярных произведений. Метод координатной релаксации

Метод скалярных произведений. Этот прием особенно удобен в применении к симметричной матрице, однако мы изложим его без этого предположения. Пусть наряду с последовательностью итераций вектора матрицей A

$$Y_0, Y_1 = AY_0, Y_2 = A^2Y_0, \dots, Y_k = A^kY_0, \dots \quad (17.1)$$

вычислена также последовательность итераций матрицей A^T

$$Z_0, Z_1 = A^T Z_0, Z_2 = (A^T)^2 Z_0, \dots, Z_k = (A^T)^k Z_0, \dots \quad (17.2)$$

Введём базисы $U_1, \dots, U_n$ и $V_1, \dots, V_n$, составленные из собственных векторов матриц A и A^T , причём будем предполагать, что эти базисы удовлетворяют условиям ортогональности и нормированности.

Пусть

$$\begin{aligned} Y_0 &= a_1 U_1 + a_2 U_2 + \dots + a_n U_n, \\ Z_0 &= b_1 V_1 + b_2 V_2 + \dots + b_n V_n. \end{aligned} \quad (17.3)$$

$$\begin{aligned} \text{Тогда } (Y_k, Z_k) &= \left(A^k Y_0, (A^T)^k Z_0 \right) = \left(A^{2k} Y_0, Z_0 \right) = \\ &= \left(a_1 \lambda_1^{2k} + a_2 \lambda_2^{2k} U_2 + \dots + a_n \lambda_n^{2k} U_n, b_1 V_1 + b_2 V_2 + \dots + b_n V_n \right). \end{aligned}$$

Далее, в силу свойств ортогональности и нормированности системы векторов $U_1, \dots, U_n$ и $V_1, \dots, V_n$, имеем

$$(Y_k, Z_k) = a_1 b_1 \lambda_1^{2k} + a_2 b_2 \lambda_2^{2k} + \dots + a_n b_n \lambda_n^{2k} \quad (17.4)$$

Аналогично

$$(Y_{k-1}, Z_k) = a_1 b_1 \lambda_1^{2k-1} + a_2 b_2 \lambda_2^{2k-1} + \dots + a_n b_n \lambda_n^{2k-1}. \quad (17.5)$$

Из равенств (17.4) и (17.5) получаем

$$\frac{(Y_k, Z_k)}{(Y_{k-1}, Z_k)} = \frac{a_1 b_1 \lambda_1^{2k} + a_2 b_2 \lambda_2^{2k} + \dots + a_n b_n \lambda_n^{2k}}{a_1 b_1 \lambda_1^{2k-1} + a_2 b_2 \lambda_2^{2k-1} + \dots + a_n b_n \lambda_n^{2k-1}} = \lambda_1 + O\left(\frac{\lambda_2}{\lambda_1}\right)^{2k}. \quad (17.6)$$

Из этой оценки видно, что образование скалярного произведения сокращает число шагов итерации, нужных для определения λ_1 с данной точностью, почти вдвое. Однако при этом требуется дополнительное вычисление последовательности (17.2).

В случае симметричной матрицы, при $Z_0 = Y_0$ последовательности (17.1) и (17.2) совпадают, и поэтому в этом случае применение метода скалярных произведений особенно целесообразно. Начиная с некоторого шага процесса, нужно вычислять соответствующие скалярные произведения и определять λ_1 через их отношения. Именно,

$$\lambda_1 \approx \frac{(A^k Y_0, A^k Y_0)}{(A^{k-1} Y_0, A^k Y_0)}. \quad (17.7)$$

Метод координатной релаксации. Метод применим только в случае, если матрица A симметрична. В некотором роде он близок к итерационным методам для решения линейных систем, основанным на релаксации того или другого функционала. В данном случае роль такого функционала играет отношение Релея.

Вычислим прежде всего, как изменяется отношение Релея $\mu(X) = \frac{(AX, X)}{(X, X)}$ при изменении X в определённом направлении. Пусть

$$X' = X + \alpha Y, \quad (17.8)$$

где Y – некоторый фиксированный вектор, определяющий направление изменения вектора X . Тогда

$$\begin{aligned} (AX', X') &= (AX + \alpha AY, X + \alpha Y) = (AX, X) + \alpha (AX, Y) + \alpha (AY, X) + \alpha^2 (AY, Y) = \\ &= (AX, X) + 2\alpha (AX, Y) + \alpha^2 (AY, Y) \end{aligned}$$

и, следовательно,

$$\mu(X') = \frac{(AX', X')}{(X', X')} = \frac{(AX, X) + 2\alpha (AX, Y) + \alpha^2 (AY, Y)}{(X, X) + 2\alpha (AX, Y) + \alpha^2 (Y, Y)}. \quad (17.9)$$

Подберём теперь множитель α так, чтобы отношение $\mu(X')$ достигало наибольшего значения. Вычисляя производную $\mu(X')$ по α , получим

$$\frac{d\mu(X')}{d\alpha} = \frac{[2(AX, Y) + 2\alpha(AY, Y)][(X, X) + 2\alpha(X, Y) + \alpha^2(Y, Y)]}{[(X, X) + 2\alpha(X, Y) + \alpha^2(Y, Y)]^2} - \frac{[2(X, Y) + 2\alpha(Y, Y)][(AX, X) + 2\alpha(AX, Y) + \alpha^2(AY, Y)]}{[(X, X) + 2\alpha(X, Y) + \alpha^2(Y, Y)]^2}$$

и поэтому для определения α получим уравнение

$$[(AX, Y)(Y, Y) - (AY, Y)(X, Y)]\alpha^2 + [(AX, X)(Y, Y) - (AY, Y)(X, X)]\alpha + (AX, X)(X, Y) - (AX, Y)(X, X) = a\alpha^2 + b\alpha + c = 0, \quad (17.10)$$

где $a = (AX, Y)(Y, Y) - (AY, Y)(X, Y)$, $b = (AX, X)(Y, Y) - (AY, Y)(X, X)$,
 $c = (AX, X)(X, Y) - (AX, Y)(X, X)$.

Исследование уравнения (17.10) показывает, что его корни всегда вещественны и различны.

Координатный релаксационный метод заключается в том, что за вектор Y на каждом шагу процесса берётся один из координатных векторов e_i .

Пусть $X = (x_1, \dots, x_n)^T$. Введём обозначения

$$\begin{aligned} F &= AX = (f_1, \dots, f_n)^T, \\ p &= (AX, X), \\ q &= (X, X). \end{aligned} \quad (17.11)$$

Тогда

$$\begin{aligned} (X, AY) &= (X, Ae_i) = (AX, e_i) = f_i, \\ (AY, Y) &= (Ae_i, e_i) = a_{ii}, \\ (Y, Y) &= (e_i, e_i) = 1, \\ (X, Y) &= (X, e_i) = x_i. \end{aligned} \quad (17.12)$$

Квадратное уравнение для определения α будет иметь вид

$$\alpha^2 [f_i - a_{ii}x_i] + \alpha [p - a_{ii}q] + px_i - f_iq = 0. \quad (17.13)$$

Выбрав соответствующий корень, определим следующее приближение

$$X' = X + \alpha e_i. \quad (17.14)$$

Для этого приближения будем иметь

$$F' = AX' = AX + \alpha Ae_i = F + \alpha A_i, \quad (17.15)$$

где A_i есть i -й столбец матрицы A . Далее

$$\begin{aligned} p' &= (AX', X') = (F', X') = p + 2\alpha f_i + \alpha^2 a_{ii}, \\ q' &= (X', X') = q + 2\alpha x_i + \alpha^2. \end{aligned} \quad (17.16)$$

Таким образом, величины p' , q' и компоненты вектора F' , нужные для проведения следующего шага, легко вычисляются.

Если на следующем шаге менять j -ю компоненту, $j \neq i$, то в квадратном уравнении нужно заменить a_{ii} на a_{jj} , x_i на x_j' , f_i на f_j' , q на q' и p на p' . Выбор номеров изменяемых компонент может осуществляться различно. Простейшей возможностью являются циклическое чередование индексов. Ввиду возможного накопления ошибок округления время от времени нужно вычислять величины q , p и f_i ($i = 1, 2, \dots, n$) непосредственно по определяющим их формулам.

Отметим также соотношение

$$\mu' = \frac{p - f_i x_i + \alpha a}{q - x_i^2} = \frac{p'}{q'}, \quad (17.17)$$

которое можно использовать для контроля вычислений.

17. Метод λ -разности. Уточнение отдельного собственного значения и принадлежащего ему собственного вектора

Метод λ -разности. Метод λ -разности даёт возможность, зная наибольшее по модулю собственное значение λ_1 , находить следующее собственное значение λ_2 и принадлежащий ему собственный вектор при условии, что

$$|\lambda_1| > |\lambda_2| > |\lambda_3|.$$

Метод состоит в следующем. Пусть вычислена последовательность

$$y_1, y_2, \dots, y_m, \dots, y_k, \dots \quad (18.1)$$

и из неё определено $\lambda_1 \approx \frac{y_{k+1}}{y_k}$. Здесь y_k - любая компонента вектора $Y_k = A^k Y_0$.

Образую разность

$$\Delta y_k = y_{k+1} - \lambda_1 y_k = c_2 (\lambda_2 - \lambda_1) \lambda_2^k + \dots + c_n (\lambda_n - \lambda_1) \lambda_n^k. \quad (18.2)$$

Так как λ_2 по модулю больше, чем все остальные собственные значения, и $c_2 \neq 0$, то первый член этой разности будет преобладать, и можно определить λ_2 аналогично тому, как мы определяли λ_1 :

$$\lambda_2 \approx \frac{y_{k+1} - \lambda_1 y_k}{y_k - \lambda_1 y_{k-1}}. \quad (18.3)$$

Однако при таком определении λ_2 происходит исчезновение значащих цифр, так как в числителе и знаменателе отношения (18.3) приходится вычитать величины, близкие друг к другу. Практически целесообразно, найдя λ_1 из отношения y_{k+1} и y_k , вернуться назад и определить λ_2 из отношения

$$\lambda_2 \approx \frac{y_{m+1} - \lambda_1 y_m}{y_m - \lambda_1 y_{m-1}}, \quad m < k, \quad (18.4)$$

беря в качестве m наименьшее из чисел, при котором преобладание λ_2 над следующими собственными числами уже начинает сказываться. Указанный приём даёт для λ_2 довольно грубые значения, однако часто достаточные для

нужд практики. Теоретически возможно при помощи аналогичного процесса определять и следующие собственные числа.

Очевидно, что для определения второго собственного вектора процесс составления λ -разности надо произвести в последовательности $AY_0, A^2Y_0, \dots, A_kY_0, \dots$

Действительно, разность

$$A^{k-1}Y_0 - \lambda_1 A^k Y_0 = a_2(\lambda_2 - \lambda_1)\lambda_2^k X_2 + \dots + a_n(\lambda_n - \lambda_1)\lambda_n^k X_n$$

показывает, что компоненты вектора X_2 могут быть найдены аналогично тому, как мы определяли компоненты вектора X_1 .

Метод λ -разности может быть обобщён следующим образом. Пусть известны λ_1 и λ_2 и требуется определить λ_3 , причём известно также, что $|\lambda_3| > |\lambda_4|$.

В этом случае мы составляем «вторую $\lambda_1\lambda_2$ -разность», т.е.

$$\Delta^2 y_k = (y_{k+2} - \lambda_1 y_{k+1}) - \lambda_2 (y_{k+1} - \lambda_1 y_k) = y_{k+2} - (\lambda_1 + \lambda_2) y_{k+1} + \lambda_1 \lambda_2 y_k.$$

Легко видеть, что

$$\Delta^2 y_k = c_3(\lambda_3 - \lambda_1)(\lambda_3 - \lambda_2)\lambda_3^k + \dots + c_n(\lambda_n - \lambda_1)(\lambda_n - \lambda_2)\lambda_n^k$$

и при достаточно большом k

$$\lambda_3 \approx \frac{\Delta^2 y_{k+1}}{\Delta^2 y_k}.$$

Пропадание знаков при составлении второй $\lambda_1\lambda_2$ -разности будет ещё значительнее, чем при использовании первой λ_1 -разности, так что в случае вещественных λ_1 и λ_2 метод почти не применим. Относительно хорошие результаты получаются лишь в случае, когда λ_1 и λ_2 образуют комплексную пару. В этом случае вторую $\lambda_1\lambda_2$ -разность целесообразно искать в виде $\Delta^2 y_k = y_{k+2} + py_{k+1} + qy_k$.

Уточнение отдельного собственного значения. Пусть λ и U суть приближённые значения для некоторого собственного значения матрицы A и

принадлежащего этому собственному значению собственного вектора. Обозначим через $\lambda^* = \lambda + \Delta\lambda$ и $U^* = U + \Delta U$ точные значения того и другого. Пусть

$$AU - \lambda U = r. \quad (18.5)$$

Для определения $\Delta\lambda$ и ΔU имеем нелинейное уравнение $AU^* = \lambda^* U^*$ или

$$A(U + \Delta U) = (\lambda + \Delta\lambda)(U + \Delta U). \quad (18.6)$$

Отбрасывая члены 2-ого порядка малости, получим

$$AU + A\Delta U = \lambda U + \Delta \lambda U + \lambda \Delta U.$$

или, принимая во внимание (18.5),

$$r + A\Delta U - \Delta\lambda U - \lambda\Delta U = 0. \quad (18.7)$$

Без нарушения общности можно считать, что первая компонента вектора ΔU равна 0, ибо собственный вектор U^* определён с точностью до постоянного множителя. Записав равенство (18.7) в компонентах, мы получим систему n линейных уравнений с n неизвестными $\Delta\lambda, \Delta u_2, \dots, \Delta u_n$. Эта система имеет вид

[illegible]

Найдя $\Delta\lambda, \Delta u_1, \dots, \Delta u_n$, можно повторить процесс.

Глава II. ОСНОВНЫЕ АЛГЕБРАИЧЕСКИЕ СТРУКТУРЫ

2. Группы

1.1 Определение и примеры. Пусть G — какое-то множество элементов (например, множество чисел, или функций, или каких-нибудь объектов геометрической природы и т. д.) Говорят, что в множестве G определена *алгебраическая операция*, если каждому двум элементам a, b из G , взятым в определенном порядке, однозначным образом поставлен в соответствие некоторый третий элемент c из G .

Примеры алгебраических операций: сложение целых чисел (каждым двум целым числам ставится в соответствие их сумма), их вычитание; сложение векторов на плоскости; сложение или умножение квадратных матриц n -го порядка; векторное умножение векторов трехмерного пространства.

Алгебраическая операция, определенная в множестве G , обычно называется или *умножением*, или *сложением*. В первом случае, если элементам $a, b \in G$ поставлен в соответствие элемент $c \in G$, пишут $c = ab$, во втором случае пишут $c = a + b$.

Множество G с определенной в нем алгебраической операцией является *группой*, если выполнены следующие условия:

- 1) для любых трех элементов $a, b, c \in G$

$$a(bc) = (ab)c$$

(ассоциативность умножения);

- 2) в множестве G существует такой элемент e , что для любого $a \in G$

$$ae = ea = a;$$

- 3) для любого $a \in G$ существует такое $a' \in G$, что

$$aa' = a'a = e.$$

Элемент e называется *единицей* группы, элемент a' называется элементом, *обратным* a , и обозначается обычно через a^{-1} . Если операция, определенная в группе, называется сложением, то вместо единицы группы говорят о *нуле* группы (это элемент, обозначаемый символом 0 и

обладающий свойством $a + 0 = 0 + a = a$ для любого $a \in G$), а вместо элемента a^{-1} , обратного элементу a , говорят об элементе $(-a)$, *противоположном* a ($a + (-a) = 0$).

Множество X с заданной на нем ассоциативной алгебраической операцией называется *полугруппой*. Полугруппу с единичным (нейтральным) элементом называют *моноидом*.

Пример 2.1.1 Группами являются:

1) множество всех целых чисел относительно операции сложения (так называемая *аддитивная группа* всех целых чисел);

2) множество всех четных чисел относительно операции сложения (аддитивная группа всех четных чисел);

3) множество всех отличных от нуля рациональных чисел относительно операции умножения (*Мультипликативная группа* отличных от нуля рациональных чисел);

4) множество всех векторов на плоскости относительно обычного сложения векторов;

5) множество всех невырожденных квадратных матриц n -го порядка с действительными элементами относительно операции умножения матриц.

Целые числа относительно операции умножения группы не образуют, так как для целого числа, отличного от ± 1 , не существует обратного ему целого числа. По той же причине не образуют группы относительно умножения все квадратные матрицы n -го порядка. Не является группой и множество всех векторов трехмерного пространства относительно векторного умножения векторов, так как эта операция не ассоциативна.

Если операция, определенная в группе, *коммутативна* (т.е. для любых элементов a, b группы $ab = ba$), то сама группа называется *коммутативной* или *абелевой*. Группы приведенных выше примеров 1, 2, 3, 4 абелевы, группа примера 5 не абелева. Еще пример не абелевой группы — множество всех вращений трехмерного пространства вокруг начала

координат, где произведением двух вращений называется их последовательное выполнение.

Если группа G состоит из конечного числа элементов, то она называется *конечной группой*, а число элементов в ней называется *порядком группы*.

Пример 2.1.2. Произведением двух подстановок n -й степени называется результат их последовательного выполнения (например, если $n=4$ и

$$a = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 3 & 2 \end{pmatrix}, \quad \text{то} \quad ab = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 3 & 1 \end{pmatrix}, \quad \text{так как,}$$

например, число 1 при подстановке a переходит в 2, а число 2 при подстановке b переходит в 4, т.е. в итоге 1 переходит в 4, и т. д.). Это — алгебраическая операция, относительно которой множество всех подстановок n -й степени является группой. Единицей группы служит

тождественная подстановка, $\begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix}$, элементом, обратным к

подстановке $\begin{pmatrix} 1 & 2 & \dots & n \\ i_1 & i_2 & \dots & i_n \end{pmatrix}$ является подстановка $\begin{pmatrix} i_1 & i_2 & \dots & i_n \\ 1 & 2 & \dots & n \end{pmatrix}$. Эта

группа называется *симметрической группой n -й степени* и обычно обозначается через S_n .

1.2. Изоморфизм групп. Говорят, что между элементами множеств M и N установлено *взаимно однозначное соответствие*, если каждому элементу множества M поставлен в соответствие некоторый вполне определенный элемент множества N , причем различным элементам из M соответствуют различные элементы в N и всякий элемент из N соответствует некоторому элементу из M .

Две группы G и H называются *изоморфными*, если между их элементами можно установить взаимно однозначное соответствие, при котором для любых элементов $a, b \in G$ и соответствующих им элементов $a', b' \in H$ элементу $ab = c$ будет соответствовать элемент $c' = a'b'$.

Пример 2.1.3 Аддитивная группа G всех целых чисел изоморфна аддитивной группе H всех четных чисел (для установления изоморфизма

между ними можно каждому числу $k \in G$ поставить в соответствие число $2k \in H$).

Пример 2.1.4. Мультипликативная группа всех положительных действительных чисел изоморфна аддитивной группе всех действительных чисел (изоморфизм: $a \rightarrow \lg a$).

Изоморфное отображение группы на себя называется *автоморфизмом* этой группы.

Пример 2.1.5. Одним из автоморфизмов аддитивной группы всех целых чисел является отображение, при котором каждому целому числу a ставится в соответствие число $(-a)$.

1.3. Гомоморфизм. Пусть каждому элементу группы G соответствует однозначно определенный элемент группы H , причем если элементам $a, b \in G$ соответствуют элементы $a', b' \in H$, то элементу $ab = c$ соответствует элемент $c' = a'b'$. Такое отображение $G \rightarrow H$ называется *гомоморфизмом*; говорят, что *группа G гомоморфно отображена в группу H* . Если при этом на каждый элемент группы H отображается хотя бы один элемент группы G , то говорят о *гомоморфном отображении* группы G на группу H ; гомоморфизм в этом случае называется *эпиморфизмом*.

Вообще говоря, при гомоморфизме в данный элемент группы H могут переходить различные элементы группы G , а также может не переходить ни одного.

Пример 2.1.6. Если каждому четному числу поставить в соответствие число 1, а каждому нечетному поставить в соответствие число -1 , то получится гомоморфное отображение аддитивной группы всех целых чисел в мультипликативную группу всех отличных от нуля рациональных чисел. Это отображение будет также гомоморфным отображением аддитивной группы всех целых чисел на мультипликативную группу, состоящую из чисел $-1, 1$.

Пример 2.1.7. Если каждой невырожденной квадратной матрице n -го порядка с действительными элементами поставить в соответствие

определитель этой матрицы, то получится гомоморфное отображение группы (по умножению) всех действительных невырожденных квадратных матриц n -го порядка на мультипликативную группу всех отличных от нуля действительных чисел.

При всяком гомоморфном отображении группы G в группу H единица группы G (или нуль, если групповая операция — сложение) переходит в единицу (в нуль) группы H . Совокупность всех элементов группы G , переходящих в единицу (в нуль) группы H , называется *ядром* данного гомоморфизма. В приведенном выше примере 2.1.6 ядро гомоморфизма составляют все четные числа, в примере 2.1.7 — все матрицы с определителем, равным 1.

Гомоморфное отображение группы в себя называется *эндоморфизмом* этой группы. Примером эндоморфизма может служить отображение, при котором каждому элементу группы ставится в соответствие единица этой группы.

Если группа G гомоморфно отображена на какое-то множество M , в котором определена алгебраическая операция, то относительно этой алгебраической операции множество M само необходимо является группой.

1.4. Подгруппы. Циклические группы. Подмножество A группы G называется *подгруппой* этой группы, если вместе с каждым элементом a оно содержит также обратный ему элемент a^{-1} и вместе с каждыми двумя элементами a, b оно содержит и их произведение ab . Эти два требования можно заменить одним: для любых элементов $a, b \in A$ элемент $(ab)^{-1}$ должен лежать в A .

Всякая подгруппа данной группы G сама является группой относительно той операции, которая определена в G .

Пример 2.1.8. Аддитивная группа всех четных чисел является подгруппой аддитивной группы всех целых чисел, которая сама является подгруппой аддитивной группы всех комплексных чисел.

Пример 2.1.9. Множество всех четных подстановок n элементов является подгруппой в группе всех подстановок n элементов. Эта подгруппа называется *знакопеременной группой, n -й степени* и обычно обозначается через A_n .

В любой группе подмножество, состоящее только из единичного элемента группы, является подгруппой. Эта подгруппа называется *единичной подгруппой* данной группы и обычно обозначается символом E . Сама группа также всегда является своей подгруппой. Всякая подгруппа, отличная от всей группы, называется *истинной подгруппой* этой группы.

Если в группе G взять какой-нибудь элемент g и все степени этого элемента (или все его кратные, если операция в группе — сложение), то также получится подгруппа группы G . Эта подгруппа называется *циклической подгруппой, порожденной элементом g* , и обозначается через $\{g\}$. Если подгруппа $\{g\}$ совпадает со всей группой G , то сама группа G называется *циклической группой*.

Пусть G — произвольная группа, a — ее некоторый элемент. Имеются две возможности.

- 1) Все степени элемента a различны, т.е. $m \neq n \Rightarrow a^m \neq a^n$. В этом случае говорят, что элемент $a \in G$ имеет *бесконечный порядок*.
- 2) Имеются совпадения $a^m = a^n$ при $m \neq n$. Если, например, $m > n$, то существуют положительные степени элемента $a \in G$, равные единичному элементу. Пусть q — наименьший положительный показатель, для которого $a^q = e$. Тогда говорят, что a — элемент *конечного порядка q* .

В конечной группе все элементы будут конечного порядка.

Пример 2.1.10. Аддитивная группа всех целых чисел есть бесконечная циклическая группа (она состоит из всех кратных числа 1).

Пример 2.1.11. Все значения корня n -й степени из 1 образуют (относительно умножения) циклическую группу порядка n , порожденную любым из первообразных корней n -й степени из 1.

1.5. Смежные классы. Разложение группы по подгруппе.

Если A — подгруппа, g — произвольный элемент группы G , то через gA обозначается множество всех элементов группы G , получающихся при умножении элемента g на всевозможные элементы из подгруппы A (т.е., множество всех элементов вида ga , где $a \in A$). Это множество называется *левым смежным классом группы G по подгруппе A , определяемым элементом g* . Аналогично *правым смежным классом Ag группы G по подгруппе A , определяемым элементом g* , называется множество всех элементов вида ag , где $a \in A$. Если групповая операция — сложение, то левые и правые смежные классы обозначаются соответственно через $g+A$ и $A+g$.

Пример 2.1.12. В группе всех подстановок третьей степени возьмем подгруппу A , состоящую из подстановок $e = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$ и $a = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$, и

возьмем элемент $g = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$. Тогда левый смежный класс gA будет

состоять из подстановок $ge = g = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$ и $ga = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$, а правый

смежный класс Ag — из подстановок $eg = g = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$ и $ag = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$.

Каждый левый смежный класс определяется любым из входящих в него элементов, т. е. если $g_1 \in gA$, то $g_1A = gA$.

Два любых левых смежных класса группы G по подгруппе A или совпадают, или не имеют ни одного общего элемента. Одним из левых смежных классов группы G по подгруппе A является сама подгруппа A ($A = eA$, где e — единица группы G); все остальные левые смежные классы по подгруппе A подгруппами группы G не являются.

Эти же утверждения верны для правых смежных классов.

Число всех различных левых смежных классов группы G по подгруппе A всегда равно числу всех различных правых смежных классов группы G по этой же подгруппе (в бесконечном случае это означает, что мощность

множества всех различных левых смежных классов группы G по подгруппе A совпадает с мощностью множества правых смежных классов). Это число (в бесконечном случае—мощность) называется *индексом* подгруппы A в группе G . Если группа G конечна, то ее порядок равен произведению порядка любой ее подгруппы A на индекс этой подгруппы в группе G (*теорема Лагранжа*). Отсюда следует, что порядок всякой подгруппы конечной группы является делителем порядка группы. Также порядок любого элемента конечной группы является делителем порядка этой группы. Обратно, если порядок конечной группы G делится на простое число p , то G обладает элементами порядка p (*теорема Коши*).

Если в произвольной группе G выбрана какая-то подгруппа A , то все элементы группы G можно разбить на непересекающиеся классы, объединяя вместе те элементы, которые лежат в одном и том же левом смежном классе группы G по подгруппе A . Такое разбиение называется *левосторонним разложением группы G по подгруппе A* . Если вместо левых смежных классов взять правые смежные классы по подгруппе A , то получится *правостороннее разложение группы G по подгруппе A* .

Пример 2.1.13. В симметрической группе 3-й степени S_3 , элементы которой

$$a_1 = e = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \quad a_2 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \quad a_3 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \quad a_4 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix},$$

$$a_5 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \quad a_6 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \quad \text{возьмем подгруппу } A, \text{ состоящую из}$$

элементов e и a_6 . Тогда при левостороннем разложении группы S_3 по подгруппе A множество ее элементов разобьется на классы: 1) e, a_6 (смежный класс $eA = A$), 2) a_2, a_5 (смежный класс a_2A), 3) a_3, a_4 (смежный класс a_3A), а при правостороннем разложении получатся классы: 1) e, a_6 (смежный класс $Ae = A$), 2) a_2, a_4 (смежный класс Aa_2), 3) a_3, a_5 (смежный класс Aa_3). Индекс подгруппы A в группе S_3 равен 3.

1.6. Нормальный делитель группы. Если при левостороннем и при правостороннем разложении группы G по некоторой ее подгруппе H

классы, на которые распадаются элементы группы G , получаются одинаковыми, то подгруппа H называется *нормальным делителем* группы G (или *инвариантной подгруппой*.)

Пример 2.1.14. В примере 2.1.13 подгруппа A группы S_3 не является нормальным делителем этой группы.

Подгруппа H тогда и только тогда является нормальным делителем группы G , когда для любого элемента g группы G

$$gH = Hg.$$

Это равенство означает, что для всякого элемента h из H можно найти в H такие элементы h' и h'' , что

$$gh = h'g, \quad hg = gh''.$$

Пример 2.1.15. Если группа G абелева, то всякая ее подгруппа H является нормальным делителем.

1.7. Фактор-группа. Если в множестве всех смежных классов группы G по нормальному делителю H ввести операцию по правилу $(g_1H)(g_2H) = (g_1g_2)H$ (при аддитивной записи $(g_1+H)+(g_2+H) = (g_1+g_2)+H$, то это будет алгебраическая операция, относительно которой множество всех смежных классов группы G по нормальному делителю H само окажется группой. Эта группа называется *фактор-группой группы G по нормальному делителю H* и обозначается через G/H . Единицей фактор-группы G/H является смежный класс H . Элементом, обратным элементу фактор-группы gH , является смежный класс $g^{-1}H$. Порядок фактор-группы G/H равен индексу H в G .

1.8. Группы подстановок. *Группы подстановок* — это подгруппы симметрических групп. При их изучении обычно для удобства пользуются записью подстановок в виде произведений циклов. *Циклом* или *циклической подстановкой* называется такая подстановка чисел $1, 2, \dots, n$, которая одно из этих чисел i_1 переводит в число i_2 , i_2 переводит в i_3 и т.д., i_{k-1} переводит в i_k ($k \leq n$) i_k переводит в i_1 а все остальные числа оставляет на месте. Эта подстановка обозначается через $(i_1 i_2, \dots, i_k)$. Циклы $(i_1 i_2, \dots, i_k)$ и, например, $(i_2, \dots, i_k, i_1)$ равны между собой. Число k называется *длиной* цикла. Цикл

длины 1 — это тождественная подстановка. Чтобы представить произвольную подстановку чисел $1, 2, \dots, n$ в виде произведения циклов, нужно взять любое из этих чисел (например, 1), затем число, в которое оно переводится данной подстановкой (пусть это число — j_2), затем число, в которое j_2 переходит при этой подстановке, и т.д. до числа j_l , переводящегося данной подстановкой в первое из взятых чисел (в нашем случае в 1), и выписать цикл $(1, j_2, \dots, j_l)$, после этого нужно взять какое-нибудь из чисел $1, 2, \dots, n$, которое пока еще не встречалось (если такое число существует), и, начиная с него, сделать то же самое и т.д., пока не будут использованы все числа $1, 2, \dots, n$. Тогда мы получим несколько циклов, не имеющих общих действительно перемещаемых ими чисел, произведению которых (в любом порядке) и будет равна данная подстановка. Циклы, не содержащие общих действительно перемещаемых ими чисел, называются *независимыми*. Те циклы, входящие в произведение, длина которых равна 1, обычно не пишут. При этом условии всякая подстановка разлагается в произведение независимых циклов единственным образом.

Пример 2.1.16. $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 8 & 2 & 1 & 5 & 7 & 4 & 6 & 3 \end{pmatrix} = (183)(4576).$

2. Кольца

2.1. Определения, примеры Множество R с двумя определенными в нем алгебраическими операциями, сложением и умножением, называется *кольцом*, если относительно операции сложения оно является абелевой группой, а операция умножения связана с операцией сложения законами *дистрибутивности*, т. е. для любых трех элементов $a, b, c \in R$

$$a(b + c) = ab + ac \text{ и } (b+c)a = ba + ca. \quad (2.1)$$

Умножение, определенное в кольце, не обязано быть ни ассоциативным, ни коммутативным. Если умножение, определенное в кольце R , ассоциативно (т. е. для любых трех элементов $a, b, c \in R$ $a(bc) = (ab)c$), то кольцо R называется *ассоциативным кольцом*. Если, кроме того, умножение,

определенное в R , коммутативно, то R называется *коммутативным кольцом*. В коммутативном кольце второе из равенств (2.1) является следствием первого.

Если в кольце R для любого элемента a выполнено условие $a^2 = 0$ и для любых трех элементов a, b, c $a(bc) + b(ca) + c(ab) = 0$ (*тождество Якоби*), то R называется *кольцом Ли*.

Пример 2.1.1. Все целые числа относительно обычных операций сложения и умножения чисел образуют коммутативное кольцо.

Пример 2.2.2. Все рациональные числа, все действительные числа, все комплексные числа относительно обычных операций сложения и умножения образуют коммутативные кольца.

Пример 2.2.3. Все многочлены от одного переменного с произвольными числовыми коэффициентами относительно обычных операций сложения и умножения многочленов образуют коммутативное кольцо.

Пример 2.2.4. Ассоциативное, но не коммутативное, кольцо образуют все квадратные матрицы n -го порядка с произвольными числовыми элементами.

Пример 2.2.5. Множество всех векторов трехмерного пространства, где векторы складываются обычным образом, а произведением двух векторов называется их векторное произведение, является неассоциативным кольцом. Это кольцо есть кольцо Ли.

Абелева группа, которая получится, если в кольце рассматривать только одну операцию сложения, называется *аддитивной группой* кольца. Нулевой элемент этой группы называется *нулем кольца*. Произведение любого элемента a кольца на нуль равно нулю кольца: $a \cdot 0 = 0$.

Если для элементов a, b некоторого кольца $ab = 0$, но $a \neq 0$ и $b \neq 0$, то a и b называются *делителями нуля* (a — левый делитель нуля, b — правый). Если в кольце R делителей нуля нет, то R называется *кольцом без делителей нуля*. Коммутативное кольцо без делителей нуля называется *областью целостности*.

Пример 2.2.6. Всякое кольцо, в котором элементы — числа, а операции — обычное сложение и умножение чисел, является областью целостности.

Пример 2.2.7. Все функции, определенные и непрерывные на отрезке $[—1, 1]$, относительно обычных операций сложения и умножения функций образуют кольцо с делителями нуля (например, произведение функций

$$f_1(x) = \begin{cases} 0 & \text{при } -1 \leq x \leq 0, \\ x & \text{при } 0 < x \leq 1 \end{cases}; \quad f_2(x) = \begin{cases} x & \text{при } -1 \leq x \leq 0, \\ 0 & \text{при } 0 < x \leq 1 \end{cases};$$

ни одна из которых не равна нулю кольца, является нулем).

Если для элементов a, b_1, b_2 кольца выполнено равенство $ab_1 = ab_2$ (или $b_1a = b_2a$), причем $a \neq 0$ не есть левый (соответственно — правый) делитель нуля, то $b_1 = b_2$, т. е. на отличные от нуля элементы, не являющиеся делителями нуля, равенства можно сокращать. Производить сокращение на элемент, являющийся делителем нуля, нельзя.

Пример 2.2.8. В кольце всех квадратных матриц 2-го порядка (пример

2.2.4) для матриц $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, B_1 = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}, B_2 = \begin{pmatrix} 1 & 2 \\ 4 & 0 \end{pmatrix}$ справедливо

равенство $AB_1 = AB_2$, хотя $B_1 \neq B_2$. Здесь A — левый делитель нуля: например,

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Элемент e кольца R называется *единицей* этого кольца, если для любого элемента $a \in R$ $ae = ea = a$. Единицы в кольце может и не быть. Если в кольце R единица есть, то R называется *кольцом с единицей*.

Пример 2.2.9. Кольцо всех целых чисел есть кольцо с единицей.

Пример 2.2.10. Все четные числа образуют кольцо без единицы.

В кольце с единицей e для элемента $a \neq 0$ может существовать *обратный* ему элемент a^{-1} со свойством $aa^{-1} = a^{-1}a = e$ (но может такого элемента и не быть). Элементы кольца с единицей, для которых в этом кольце обратный элемент существует, называются *делителями единицы*.

Пример 2.2.11. В кольце всех квадратных матриц n -го порядка единицей является единичная матрица. Обратный элемент существует для всякой

невырожденной матрицы; для вырожденных матриц обратных им элементов не существует.

2.2 Изоморфизм. Гомоморфизм. Кольца R и Q называются *изоморфными*, если между их элементами можно установить такое взаимно однозначное соответствие, что если элементам $a_1, a_2 \in R$, соответствуют элементы $b_1, b_2 \in Q$, то элементу $a_1 + a_2$ соответствует элемент $b_1 + b_2$ и элементу $a_1 a_2$ — элемент $b_1 b_2$.

Пример 2.2.12. Кольцо всех квадратных матриц n -го порядка с действительными элементами изоморфно кольцу всех линейных преобразований, действующих в действительном n -мерном линейном векторном пространстве

Говорят, что кольцо R *гомоморфно отображено* в кольцо Q , если каждому элементу кольца R поставлен в соответствие однозначно определенный элемент кольца Q , причем если элементам a_1, a_2 кольца R соответствуют элементы $b_1, b_2 \in Q$, то элементу $a_1 + a_2$ соответствует $b_1 + b_2$, элементу $a_1 a_2$ соответствует $b_1 b_2$. Если при этом на каждый элемент кольца Q отображается по крайней мере один элемент кольца R , то говорят о гомоморфном отображении кольца R на кольцо Q ; гомоморфизм в этом случае называется *эпиморфизмом*.

2.3. Подкольца. Идеалы. Подгруппа A аддитивной группы кольца R называется *подкольцом* этого кольца, если она вместе с каждым элементом $a \in A$ содержит также их произведение $a_1 a_2$. Подкольцо A кольца R называется *левым* (соответственно *правым*) *идеалом* этого кольца, если оно вместе с каждым элементом a содержит также все элементы вида ra (вида ar), где r — произвольный элемент кольца R . В коммутативных кольцах понятия левого и правого идеала совпадают. Если подкольцо A произвольного кольца одновременно является и левым, и правым идеалом, то оно называется *двусторонним идеалом* этого кольца.

Пример 2.2.13 В кольце всех целых чисел числа, кратные некоторому фиксированному числу n , составляют двусторонний идеал.

Пример 2.2.14 В кольце всех квадратных матриц n -го порядка множество всех матриц, у которых последний столбец состоит из нулей, является левым идеалом, а множество всех матриц, у которых последняя строка состоит из нулей, является правым идеалом

В любом кольце элемент 0 и само кольцо являются двусторонними идеалами. Если других двусторонних идеалов в кольце нет, оно называется *простым кольцом*.

Пример 2.2.15 Простым кольцом служит кольцо всех квадратных матриц n -го порядка с любыми комплексными (или с любыми действительными) элементами

Пример 2.2.16. Всякое поле является простым кольцом.

Если в кольце R дано некоторое множество элементов N , то наименьший (левый, правый или двусторонний) идеал кольца R , содержащий все элементы из N , называется (соответственно левым, правым или двусторонним) *идеалом, порожденным множеством N* ; этот идеал является пересечением всех (соответственно левых, правых или двусторонних) идеалов кольца R , содержащих множество N . Идеал, порожденный одним элементом a , называется *главным идеалом* (обозначение: $(a)_l$, $(a)_r$ или (a) , в зависимости от того, идеал левый, правый или двусторонний). Если R — коммутативное кольцо, то главный идеал (a) , порожденный элементом a этого кольца, состоит из всех элементов вида $ra+na$, где $r \in R$, n — целое число. Если R — коммутативное кольцо с единицей, то (a) состоит просто из всех элементов вида ra , где $r \in R$.

3. Поля. Тела

3.1. Поля. *Полем* называется коммутативное кольцо, состоящее не только из нуля, в котором для любого элемента $a \neq 0$ и любого элемента b существует ровно один такой элемент x , что $ax = b$. Элемент x называется *частным* от деления элемента b на элемент a (обозначение: $x = \frac{b}{a}$).

Примерами полей служат: кольцо всех рациональных чисел, кольцо всех действительных чисел, кольцо всех комплексных чисел. Все комплексные числа, являющиеся корнями многочленов с рациональными коэффициентами, также образуют поле, называемое *полем алгебраических чисел*. Кольцо всех целых чисел полем не является

Всякое поле обладает единицей. Для любого отличного от нуля элемента поля существует обратный ему элемент. Множество всех отличных от нуля элементов поля образует относительно умножения, определенного в поле, абелеву группу (*мультипликативную группу поля*). Никакое поле не содержит делителей нуля. Единственными идеалами поля являются нулевой идеал и само поле.

Множество с двумя алгебраическими операциями, изоморфное полю, само является полем. Всякое гомоморфное отображение одного поля в другое является или изоморфизмом, или отображением, переводящим все элементы поля в нуль.

Если некоторое целое положительное кратное единичного элемента e поля P $ne = e + e + \dots + e$ (n слагаемых) равно нулю, то наименьшее целое положительное число p со свойством $pe = 0$ называется *характеристикой поля P* ; p всегда является простым числом. Если никакое целое положительное кратное единичного элемента поля P нулю не равно, то P называется полем *характеристики нуль*. Пример поля характеристики p — поле классов вычетов по модулю p , пример поля характеристики нуль — любое числовое поле (например, поле всех действительных чисел).

Множество P' элементов поля P называется *подполем* этого поля, если оно само является полем по отношению к тем операциям, которые определены в P ; P' тогда и только тогда является подполем поля P , когда оно вместе с любыми двумя элементами a, b содержит также $a + b$, ab , $a - b$ и $\frac{a}{b}$ (если $b \neq 0$). Если P' — подполе поля P , то P называется *расширением* поля P' . Поле, не имеющее никаких подполей, кроме него самого, называется *простым*

полем. Примеры простых полей — поле классов вычетов по модулю p , поле всех рациональных чисел.

Всякое поле характеристики p содержит подполе, изоморфное полю классов вычетов по модулю p , а всякое поле характеристики нуль содержит подполе, изоморфное полю всех рациональных чисел.

Если в поле P даны подполе P' и множество элементов N , то наименьшее подполе P'' поля P , содержащее P' и N , называется полем, полученным *присоединением* к полю P' множества N (обозначение: $P''=P'(N)$). Поле $P'(N)$ состоит из всех элементов, получающихся из элементов поля P' и множества N путем сложения, вычитания, умножения и деления, и является пересечением всех подполей поля P , содержащих P' и N . Если множество N состоит из одного элемента α , то P'' называется *простым расширением* поля P' . Если при этом α является корнем некоторого многочлена $f(x)$ с коэффициентами из поля P' , то $P'' = P'(\alpha)$ называется *простым алгебраическим расширением* поля P' , а элемент α называется *алгебраическим относительно P'* ; если же не существует многочлена $f(x)$ с коэффициентами из P' , корнем которого был бы элемент α , то $P'' = P'(\alpha)$ называется *простым трансцендентным расширением* поля P' , а элемент α называется *трансцендентным относительно P'* .

Пример 2.3.1. Простое алгебраическое расширение поля рациональных чисел, полученное присоединением к нему числа $\alpha = \sqrt[3]{2}$, состоит из всех чисел вида $a + b\sqrt[3]{2} + c\sqrt[3]{4}$, где a, b, c — рациональные числа.

Пример 2.3.2 Поле всех комплексных чисел является простым алгебраическим расширением поля всех действительных чисел, полученным из него присоединением корня i многочлена x^2+1 .

3.2. Тела. *Телом* называется ассоциативное кольцо, состоящее не только из нуля, в котором для любого элемента $a \neq 0$ и любого элемента b существуют такой элемент x и такой элемент y , что $ax = b$, $ya = b$. Отличие тела от поля состоит в том, что умножение в теле не обязано быть коммутативным.

3.3. Алгебры. Кольцо A называется *алгеброй* над полем P , если его аддитивная группа есть векторное пространство над полем P и если умножение в A связано с умножением на элементы из P формулой $\alpha(ab) = (\alpha a)b = a(\alpha b)$ ($a, b \in A, \alpha \in P$).

Если векторное пространство, которое служит аддитивной группой алгебры A , является n -мерным, то число n называется *рангом* алгебры A . Алгебры конечного ранга называются также *гиперкомплексными системами*. Если алгебра A является кольцом Ли, то она называется *алгеброй Ли*. Если какая-то алгебра является не только кольцом, но даже телом, она называется *алгеброй с делением*.

Пример 2.3.3. Множество всех векторов трехмерного пространства, в котором векторы складываются и умножаются на числа обычным образом, а произведением двух векторов является их векторное произведение, есть алгебра Ли над полем действительных чисел.

Пример 2.3.4. Множество всех квадратных матриц n -го порядка с комплексными элементами, в котором обычным образом определены операции сложения и умножения матриц и операция умножения матрицы на комплексное число, является алгеброй ранга n^2 над полем комплексных чисел.

Пример 2.3.5. Все комплексные числа относительно обычных операций сложения и умножения комплексных чисел и обычной операции умножения комплексных чисел на действительные числа образуют алгебру с делением ранга 2 над полем действительных чисел.

Глава III. ЛАБОРАТОРНЫЙ ПРАКТИКУМ

Требования к оформлению отчета о выполнении лабораторной работы

В отчет о выполнении лабораторной работы необходимо включить следующие пункты:

1. Тема лабораторной работы.
2. Постановка задачи и вариант задания.
3. Математическое описание метода решения поставленной задачи.
4. Листинг программы.
5. Результаты выполнения программы.
6. Анализ результатов и выводы.

Лабораторная работа № 1

Точные методы решения систем линейных алгебраических уравнений

Цель работы – решение систем алгебраических уравнений методом Гаусса, Гаусса – Жордана или методом квадратных корней.

Теоретическая часть

1) Метод Гаусса основан на последовательном исключении неизвестных. Этот метод имеет много вычислительных схем. Рассмотрим схему единственного деления.

Пусть дана система уравнений $AX = F$, или в развернутом виде

[illegible]

Предположим, что $a_{11} \neq 0$. Исключим неизвестное x_1 из всех уравнений (1.1), начиная со 2-го. Для этого 1-е уравнение разделим на a_{11} , и из всех остальных уравнений вычтем получившееся, умноженное на соответствующий коэффициент a_{j1} , $j = \overline{2, n}$.

В результате эти уравнения преобразуются к виду $\sum_{j=2}^n a_{ij}^{(1)} x_j = f_i^{(1)}, i = \overline{2, n}.$

Затем аналогично исключаем неизвестное x_2 из всех уравнений, начиная с 3-го и т.д.

В результате система преобразуется к треугольному виду

[illegible]

Алгоритм нахождения решения системы (1.2) представляется очевидным.

Указанная схема проста и удобна. Однако не является универсальной, в том смысле, что для применимости нужно, чтобы все ведущие элементы были отличны от нуля. Близость к нулю ведущих элементов может быть причиной значительной потери точности. Поэтому на практике чаще всего применяется схема единственного деления по главным элементам. В этой схеме в качестве исключаемой на m -м шаге неизвестной выбирается та, коэффициент при которой на предыдущем шаге был наибольшим по модулю.

2) Метод Гаусса – Жордана заключается в преобразовании исходной системы к диагональному виду в результате исключения на каждом шаге переменной $x_j (j = \overline{1, n})$ из всех уравнений, за исключением уравнения с номером j .

3) Метод квадратных корней.

Предположим, что матрица A системы симметрична¹. В этом случае её можно разложить в произведение двух транспонированных друг другу треугольных матриц.

Пусть

$$A = S^T S, \quad (1.3)$$

где

$$S = \begin{pmatrix} s_{11} & s_{12} & \cdots & s_{1n} \\ 0 & s_{22} & \cdots & s_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & s_{nn} \end{pmatrix} \quad (1.4)$$

Определим элементы матрицы S .

Имеем

$$\begin{aligned} a_{ij} &= s_{1i}s_{1j} + s_{2i}s_{2j} + \cdots + s_{ii}s_{ij}, i < j; \\ a_{ii} &= s_{1i}^2 + s_{2i}^2 + \cdots + s_{ii}^2, i = j. \end{aligned}$$

Тогда

$$s_{11} = \sqrt{a_{11}}, s_{1j} = \frac{a_{1j}}{s_{11}}, s_{ii} = \sqrt{a_{ii} - \sum_{k=1}^{i-1} s_{ki}^2}, (i > 1);$$

¹ Этого всегда можно добиться путем умножения системы уравнений слева на матрицу A^T .

$$\overline{s_{ij}} = \frac{a_{ij} - \sum_{k=1}^{i-1} \overline{s_{ki}} \overline{s_{kj}}}{s_{ii}}, \quad (j > i)$$

$$\overline{s_{ij}} = 0, \quad (i > j).$$

Далее, решение системы сводится к решению двух треугольных систем.

Равенство $AX = F$ равносильно двум равенствам $S^T K = F$ и $SX = K$.

Элементы вектора K определим по рекуррентным формулам:

$$K_1 = \frac{f_1}{s_{11}}, \quad K_i = \frac{f_i - \sum_{l=1}^{i-1} \overline{s_{li}} K_l}{\overline{s_{li}}}, \quad i > 1$$

Тогда решение исходной системы находится по формулам

$$x_n = \frac{K_n}{s_{nn}}, \quad x_i = \frac{K_i - \sum_{l=i+1}^n s_{il} x_l}{s_{ii}}, \quad (i < n)$$

Рассмотрим вопрос об устойчивости решения относительно изменения элементов матрицы.

Теоретическое решение системы $AX = F$ дается формулой $X = A^{-1}F$,

где A^{-1} – матрица, обратная к A . Обратную матрицу называют устойчивой, если малым изменениям в элементах матрицы соответствуют малые изменения в элементах обратной матрицы. Очевидно, что необходимым условием устойчивости обратной матрицы является то, чтобы определитель матрицы не был бы близок к нулю. Но это условие не является достаточным. В качестве меры близости к вырожденности матрицы A рассматривают числа обусловленности, например $P = \frac{\max |\lambda_i|}{\min |\lambda_i|}$, где λ_i – собственное значение

матрицы A .

Матрицу называют плохо обусловленной, если соответствующая ей обратная матрица неустойчива. Чем больше числа обусловленности, тем хуже обусловленность матрицы.

На практике этим определением обусловленности воспользоваться достаточно трудно, т.к. это связано с нахождением обратной матрицы и собственных значений матрицы. Поэтому обычно ограничиваются проверкой условия $\det A \approx 0$. Для этого систему нормируют, т.е. i -е уравнение системы делят на величину $\sqrt{\sum_{j=1}^n a_{ij}^2}$, а затем определитель полученной матрицы сравнивают с единицей. Малость указанного определителя по сравнению с единицей является признаком плохой обусловленности системы.

Порядок выполнения работы

1. Изучить теоретическую часть.
2. Написать программу решения системы одним из указанных методов. Отладить её на модельной задаче.
3. Вычислить значение нормированного определителя матрицы.
4. Найти решение системы. Вычислить норму вектора невязки $r = F - A\tilde{X}$, где $\tilde{X}$ - найденное решение системы.
5. Внести случайную погрешность в элементы матрицы системы и определить решение для измененной матрицы.
6. Найти решение системы, воспользовавшись одним из стандартных математических пакетов: MathCad, MathLab и т. п. Сравнить результаты с результатами, полученными с помощью выше описанных методов.
7. Проанализировать результаты. Сделать выводы о точности метода и устойчивости системы.

Варианты заданий

$$1) \quad A = \begin{pmatrix} 67.09 & -40.45 & -16.77 & 9.87 \\ -40.45 & 24.66 & 9.86 & -5.92 \\ -16.77 & 9.86 & 4.94 & -2.97 \\ 9.87 & -5.92 & -2.97 & 1.98 \end{pmatrix}, \quad F = \begin{pmatrix} 1.05 \\ 21.33 \\ 7.11 \\ -0.23 \end{pmatrix}$$

$$2) \quad A = \begin{pmatrix} 5.01 & 7.01 & 5.99 & 5.00 \\ 7.01 & 10.02 & 8.01 & 7.02 \\ 5.99 & 8.01 & 10.01 & 8.98 \\ 5.00 & 7.02 & 8.98 & 10.00 \end{pmatrix}, F = \begin{pmatrix} 0.08 \\ 21.52 \\ -7.31 \\ 51.63 \end{pmatrix}$$

$$3) \quad A = \begin{pmatrix} 0.23 & 1.28 & -5.13 & 0.14 \\ 2.28 & 6.31 & -7.15 & 0.01 \\ 2.51 & 3.15 & 1.06 & 0.86 \\ 4.28 & 7.65 & 5.41 & 2.37 \end{pmatrix}, F = \begin{pmatrix} 1.21 \\ 2.78 \\ -6.35 \\ 0.63 \end{pmatrix}$$

$$4) \quad A = \begin{pmatrix} 2.46 & -1.27 & 15.32 & 6.12 \\ -0.73 & -10.15 & 2.29 & 3.30 \\ -7.41 & -0.81 & -1.27 & 6.31 \\ 0.06 & 0.01 & -10.21 & 7.45 \end{pmatrix}, F = \begin{pmatrix} 3.28 \\ 10.24 \\ -3.58 \\ 0.51 \end{pmatrix}$$

$$5) \quad A = \begin{pmatrix} 2.38 & 7.41 & -5.15 & 0.15 \\ -6.31 & 0.28 & 2.31 & 0.45 \\ -7.11 & 28.31 & 43.25 & 16.01 \\ -0.07 & 5.21 & 0.01 & 0.04 \end{pmatrix}, F = \begin{pmatrix} 15.32 \\ 0.15 \\ 23.16 \\ 0.15 \end{pmatrix}$$

$$6) \quad A = \begin{pmatrix} 13.21 & 6.15 & 7.11 & 0.82 \\ -10.15 & -3.16 & 2.22 & -0.18 \\ -7.10 & 2.32 & -0.91 & 0.98 \\ 3.45 & 16.21 & 96.17 & 103.22 \end{pmatrix}, F = \begin{pmatrix} 1.12 \\ 2.45 \\ 126.17 \\ 0.15 \end{pmatrix}$$

$$7) \quad A = \begin{pmatrix} 3.42 & -4.15 & 2.22 & 1.16 \\ 1.43 & 2.14 & -0.72 & -1.28 \\ 0.45 & 16.21 & 7.32 & 15.61 \\ 2.34 & -0.07 & 0.17 & 0.49 \end{pmatrix}, F = \begin{pmatrix} 1.06 \\ 2.35 \\ 4.21 \\ -10.11 \end{pmatrix}$$

$$8) \quad A = \begin{pmatrix} 2.34 & 15.11 & 2.43 & 6.15 \\ 1.43 & 7.21 & -1.13 & 1.28 \\ -2.35 & 7.15 & 1.17 & 2.38 \\ 1.16 & 5.14 & 2.32 & 7.16 \end{pmatrix}, \quad F = \begin{pmatrix} 0.01 \\ 18.22 \\ 7.45 \\ 5.21 \end{pmatrix}$$

$$9) \quad A = \begin{pmatrix} 3.22 & 7.46 & -1.28 & 1.36 \\ 2.58 & 6.22 & -7.10 & 2.31 \\ 3.28 & 1.43 & 1.55 & 0.01 \\ 2.45 & -7.11 & 9.95 & -3.34 \end{pmatrix}, \quad F = \begin{pmatrix} 1.15 \\ 0.07 \\ 11.21 \\ -4.43 \end{pmatrix}$$

$$10) \quad A = \begin{pmatrix} 7.43 & 1.24 & -1.17 & -0.03 \\ 1.12 & 0.27 & -6.31 & -10.05 \\ 11.22 & -1.17 & 5.34 & 1.26 \\ 0.01 & 0.07 & 1.24 & 1.17 \end{pmatrix}, \quad F = \begin{pmatrix} 1.23 \\ 1.15 \\ 0.07 \\ -1.27 \end{pmatrix}$$

$$11) \quad A = \begin{pmatrix} 1.26 & 1.15 & 2.46 & 28.15 \\ -0.43 & -15.13 & 1.24 & 17.43 \\ 23.16 & -1.12 & 1.15 & -10.16 \\ 0.13 & 21.22 & 1.46 & 7.15 \end{pmatrix}, \quad F = \begin{pmatrix} 1.16 \\ -2.51 \\ 1.17 \\ 8.46 \end{pmatrix}$$

$$12) \quad A = \begin{pmatrix} 15.26 & -1.21 & 17.13 & -7.78 \\ 2.31 & -2.41 & 11.28 & 1.15 \\ 5.46 & -1.24 & 11.74 & -0.01 \\ 0.04 & 1.13 & 2.41 & 6.37 \end{pmatrix}, \quad F = \begin{pmatrix} 78.13 \\ -0.05 \\ 11.12 \\ 0.14 \end{pmatrix}$$

$$13) \quad A = \begin{pmatrix} 1.46 & 1.21 & -0.08 & 0.12 \\ -21.62 & 1.47 & 27.16 & -3.21 \\ 1.28 & 7.45 & 3.51 & -28.63 \\ 11.16 & 17.51 & -0.07 & -1.48 \end{pmatrix}, \quad F = \begin{pmatrix} 1.49 \\ -0.07 \\ 12.56 \\ 11.32 \end{pmatrix}$$

$$14) \quad A = \begin{pmatrix} 4.48 & -2.34 & -17.61 & 11.31 \\ 5.48 & 6.38 & 21.76 & 31.87 \\ -0.01 & 1.21 & 17.45 & -27.13 \\ 15.21 & 0.05 & -0.78 & 1.28 \end{pmatrix}, F = \begin{pmatrix} 1.58 \\ 27.13 \\ 0.08 \\ -2.43 \end{pmatrix}$$

$$15) \quad A = \begin{pmatrix} 3.46 & 1.12 & -2.28 & 16.43 \\ -0.25 & 0.73 & -11.23 & 14.42 \\ -1.28 & 7.67 & 1.13 & 2.43 \\ 1.16 & 2.34 & -0.09 & 13.44 \end{pmatrix}, F = \begin{pmatrix} 6.43 \\ -1.23 \\ -0.73 \\ 1.15 \end{pmatrix}$$

$$16) \quad A = \begin{pmatrix} 1.67 & 7.43 & -2.31 & 0.99 \\ 1.01 & -2.34 & -7.45 & 0.35 \\ 2.34 & 11.21 & 7.15 & 6.34 \\ 4.39 & 17.56 & 3.21 & 11.28 \end{pmatrix}, F = \begin{pmatrix} 0.01 \\ 11.57 \\ 0.34 \\ 12.63 \end{pmatrix}$$

$$17) \quad A = \begin{pmatrix} 15.21 & 1.46 & 17.34 & -2.34 \\ 0.08 & 1.49 & 13.28 & -0.97 \\ 11.16 & 21.73 & 3.51 & 6.28 \\ 7.48 & 3.96 & 11.21 & 9.99 \end{pmatrix}, F = \begin{pmatrix} 1.28 \\ 7.48 \\ -1.13 \\ 21.64 \end{pmatrix}$$

$$18) \quad A = \begin{pmatrix} 1.48 & 2.31 & 1.16 & -1.14 \\ 2.34 & 0.07 & 2.31 & -0.86 \\ 11.21 & -6.38 & 11.28 & 1.17 \\ 5.47 & 16.21 & 7.51 & 0.02 \end{pmatrix}, F = \begin{pmatrix} 1.58 \\ 28.46 \\ -1.15 \\ 2.38 \end{pmatrix}$$

$$19) \quad A = \begin{pmatrix} 15.49 & 6.37 & 16.21 & 7.48 \\ 0.08 & -2.31 & -16.28 & 7.15 \\ 1.28 & 0.13 & 5.48 & 16.11 \\ 0.01 & 0.08 & 0.21 & -15.46 \end{pmatrix}, F = \begin{pmatrix} 1.99 \\ -2.31 \\ 11.16 \\ 5.35 \end{pmatrix}$$

$$20) \quad A = \begin{pmatrix} 2.97 & -1.13 & 0.08 & 0.97 \\ 16.41 & 1.38 & 0.18 & 1.29 \\ -13.51 & 7.64 & 9.28 & 6.31 \\ 0.02 & 9.97 & -7.13 & 28.14 \end{pmatrix}, \quad F = \begin{pmatrix} 43.15 \\ 28.14 \\ -0.09 \\ 19.23 \end{pmatrix}$$

$$21) \quad A = \begin{pmatrix} 17.31 & -5.34 & 9.87 & 16.31 \\ 0.08 & 10.21 & -7.13 & 16.21 \\ 14.07 & 15.11 & -10.21 & 0.07 \\ -13.21 & -1.08 & 2.43 & 5.14 \end{pmatrix}, \quad F = \begin{pmatrix} 15.11 \\ -0.17 \\ 21.13 \\ 1.14 \end{pmatrix}$$

$$22) \quad A = \begin{pmatrix} 13.51 & 2.87 & 16.15 & 0.011 \\ 15.31 & 0.09 & -9.35 & -0.18 \\ 11.49 & 16.51 & 7.49 & 28.13 \\ 5.96 & 7.43 & -0.03 & 9.07 \end{pmatrix}, \quad F = \begin{pmatrix} 10.16 \\ -0.13 \\ 21.61 \\ 7.38 \end{pmatrix}$$

$$23) \quad A = \begin{pmatrix} 10.16 & -2.34 & 0.99 & 7.46 \\ 23.99 & -7.16 & 7.28 & -0.18 \\ 27.15 & -0.08 & 27.16 & -1.13 \\ 0.97 & 15.16 & 0.01 & 23.16 \end{pmatrix}, \quad F = \begin{pmatrix} 10.29 \\ 17.35 \\ 11.87 \\ -0.93 \end{pmatrix}$$

$$24) \quad A = \begin{pmatrix} 1.16 & -0.18 & 0.93 & 9.27 \\ 0.18 & 0.29 & 3.15 & 7.16 \\ 23.19 & 7.15 & -9.13 & 0.01 \\ 3.41 & 9.21 & 7.15 & 0.28 \end{pmatrix}, \quad F = \begin{pmatrix} 11.17 \\ 23.19 \\ -3.48 \\ 0.17 \end{pmatrix}$$

$$25) \quad A = \begin{pmatrix} 77.16 & 2.31 & -0.17 & 3.51 \\ 0.09 & 1.15 & 1.21 & -7.15 \\ 2.49 & 0.07 & -0.29 & 1.31 \\ 16.25 & 1.27 & 5.34 & 6.48 \end{pmatrix}, \quad F = \begin{pmatrix} 5.31 \\ -0.19 \\ 1.16 \\ 7.41 \end{pmatrix}$$

$$B = \begin{pmatrix} 0 & -\frac{a_{12}}{a_{11}} & \dots & -\frac{a_{1n}}{a_{11}} \\ -\frac{a_{21}}{a_{22}} & 0 & \dots & -\frac{a_{2n}}{a_{22}} \\ \dots & \dots & \dots & \dots \\ -\frac{a_{n1}}{a_{nn}} & -\frac{a_{n2}}{a_{nn}} & \dots & 0 \end{pmatrix}, G = \begin{pmatrix} \frac{f_1}{a_{11}} \\ \frac{f_2}{a_{22}} \\ \vdots \\ \frac{f_n}{a_{nn}} \end{pmatrix} \quad (2.3)$$

Данное преобразование системы (2.1) в систему (2.2) равносильно умножению системы (2.1) слева на матрицу

$$H = \begin{pmatrix} \frac{1}{a_{11}} & 0 & \dots & 0 \\ 0 & \frac{1}{a_{22}} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \frac{1}{a_{nn}} \end{pmatrix}$$

Таким образом, $H = D^{-1}$, где D – диагональная матрица $[a_{11}, \dots, a_{nn}]$.

Выберем начальное приближение $X^{(0)} = G$. Строим последовательные приближения по формулам

$$X^{(k)} = BX^{(k-1)} + G, \quad (2.4)$$

$$k = 1, 2, 3, \dots \text{ или в развернутом виде } x_i^{(k)} = \sum_{j=1}^n b_{ij} x_j^{(k-1)} + g_i, i = \overline{1, n}$$

Необходимое и достаточное условие сходимости процесса простой итерации состоит в том, что все собственные значения матрицы были

$$B = E - D^{-1}A$$

по модулю меньше единицы.

Укажем достаточные признаки сходимости процесса последовательных приближений:

Если выполняется одно из условий:

$$\text{I. } \|B\|_{\text{I}} = \sum_{j=1}^n |b_{ij}| \leq \mu < 1 \quad \text{при} \quad i = \overline{1, n}$$

$$\text{II. } \|B\|_{\text{II}} = \sum_{i=1}^n |b_{ij}| \leq \nu < 1 \quad \text{при} \quad j = \overline{1, n}$$

$$\text{III. } \|B\|_{\text{III}} = \sum_{i,j=1}^n b_{ij}^2 \leq \rho < 1,$$

то процесс последовательных приближений сходится.

Это равносильно преобладанию диагональных элементов в исходной матрице A .

2) Метод Зейделя.

Пусть система уравнений $AX = F$ представлена в виде

$$X = BX + G, \quad (2.5)$$

где $B = E - A$, $G = F$.

Метод Зейделя похож на метод простой итерации с той лишь разницей, что при вычислении k -го приближения для i -й компоненты учитываются вычисленные уже ранее k -е приближения для компонент $x_1^{(k)}, \dots, x_{i-1}^{(k)}$.

Вычисление последовательных приближений ведется по формулам

$$x_i^{(k)} = \sum_{j=1}^{i-1} b_{ij} x_j^{(k)} + \sum_{j=i}^n b_{ij} x_j^{(k-1)} + g_i \quad (2.6)$$

3) Метод полной релаксации.

Пусть X^* - точное решение системы

$$AX = F$$

с положительно-определенной матрицей A . (Для того, чтобы получить систему с положительно-определенной матрицей, эквивалентную исходной, нужно умножить систему слева на транспонированную матрицу A^T).

Пусть X - некоторый вектор, $f(X) = (AY, Y)$ - функция ошибки,

где $Y = X^* - X$ - вектор ошибки.

Изменим i -ю компоненту вектора X так, чтобы для измененного вектора X' значение функции ошибки было бы наименьшим.

Пусть $X' = X + \alpha e_i$.

Тогда $Y' = Y - \alpha e_i$, и

$$\begin{aligned} f(X') &= (AY', Y') = (A(Y - \alpha e_i), Y - \alpha e_i) = \\ &= (AY, Y) - 2\alpha(A Y, e_i) + \alpha^2(A e_i, e_i) = (AY, Y) + \alpha^2 a_{ii} - 2\alpha r_i = \\ &= f(X) + \frac{(a_{ii}\alpha - r_i)^2 - r_i^2}{a_{ii}}, \end{aligned} \quad (2.7)$$

где $r_i - i$ - компонента вектора невязки

$$r = F - AX$$

для приближения X .

$f(X')$ будет иметь минимальное значение при $\alpha = \frac{r_i}{a_{ii}}$.

В этом случае i -я компонента вектора невязки для приближения X' обратится в ноль.

Выбор номеров изменяемых компонент можно осуществлять либо циклически от 1 до n , либо на каждом шаге выбирать компоненту с некоторым номером k , если на предыдущем шаге k -я компонента вектора невязки оказалась наибольшей по модулю.

Подготовка системы уравнений к виду, удобному для применения метода последовательных приближений.

Пусть матрица A положительно определена. Тогда система $AX = F$ всегда может быть подготовлена к виду, в котором метод последовательных приближений будет сходящимся. Подготовка состоит в переходе от данной системы $AX = F$ к равносильной системе $HAX = HF$, где H - некоторая неособенная матрица, которая выбирается так, чтобы матрица HA была бы близка к единичной.

Положим

$$H = \frac{2}{\mu} E, \quad (2.8)$$

где μ - норма матрицы A .

Тогда система уравнений преобразуется к виду

$$X = (E - \frac{2}{\mu} A)X + \frac{2}{\mu} F = BX + G \quad (2.9)$$

Собственные значения матрицы

$$B = E - \frac{2}{\mu} A$$

будут заключены в открытом интервале $(-1,1)$, в силу того собственные значения положительно определенной матрицы A находятся в интервале $(0, \mu)$. Следовательно, метод последовательных приближений для системы (2.9) будет сходящимся.

Порядок выполнения работы

1. Изучить теоретическую часть.
2. Написать программу решения системы одним из указанных методов.
Отладить программу на модельной задаче.
3. Определить три нормы матрицы.
4. Найти решение системы с заданной точностью. Для этого выход из цикла, осуществляющего итерационный процесс, производить, как только норма соответствующего вектора невязки меньше наперед заданной точности.

Выход из цикла можно также осуществлять, как только норма разности между векторами решений, полученных на данном шаге и на предыдущем станет меньше заданной погрешности.

5. Исследовать скорость сходимости метода в зависимости от числа итерации.
6. Сделать выводы о точности метода и быстроте его сходимости.
7. Найти решение системы, воспользовавшись одним из стандартных математических пакетов: MathCad, MathLab и т. п. Сравнить результаты с результатами, полученными с помощью итерационных методов.

Воспользоваться вариантами заданий для лабораторной работы № 1.

Лабораторная работа № 3

Градиентные методы решения систем уравнений

Цель работы – изучение градиентных итерационных методов, решение системы уравнений одним из градиентных методов.

Теоретическая часть

Градиентные методы основаны на минимизации выбранного функционала, характеризующего точность решения, (например, функционала ошибки, длины вектора невязки) в направлении, противоположном направлению градиента функционала в данной точке. Такой выбор направления минимизации связан с тем, что направление, противоположное направлению градиента функционала, в некоторой точке, обеспечивает в окрестности этой точки наиболее быстрое убывание функционала.

1) Метод наискорейшего спуска.

Пусть A положительно определенная матрица и пусть

$$AX = F \quad (3.1)$$

заданная линейная система.

Ставится задача нахождения вектора, дающего минимум функционалу

$$H(X) = (AX, X) - 2F(X) \quad (3.2)$$

Т.к. этот функционал отличается от функции ошибки $f(X) = (AY, Y)$ на постоянное слагаемое (AX^*, X^*) , где X^* - точное решение, то эта задача равносильна нахождению вектора, дающего минимум функции ошибки.

Выбирается начальное приближение X_0 . Вычисляется направление, противоположное градиенту функционала $H(X)$ в этой точке, которое совпадает с направлением вектора невязки $r_0 = F - AX_0$ выбранного начального приближения. Из точки X_0 движемся в выбранном направлении до точки X_1 , дающей минимум функционалу $H(X)$.

Положим

$$X_1 = X_0 + \alpha r_0 \quad (3.3)$$

Нетрудно убедиться, что функционал $H(X_0 + \alpha r_0)$ достигает минимального значения при

$$\alpha = \alpha_0 = \frac{(r_0, r_0)}{(Ar_0, r_0)} \quad (3.4)$$

Далее определяется $X_2 = X_1 + \alpha_1 r_1$, где $r_1 = F - AX_1$ и $\alpha_1 = \frac{(r_1, r_1)}{(Ar_1, r_1)}$

и процесс продолжается по формулам

$$\begin{aligned} X_{k+1} &= X_k + \alpha_k r_k, \\ r_k &= F - AX_k = r_{k-1} - \alpha_{k-1} Ar_{k-1}, \end{aligned} \quad (3.5)$$

$$\text{где } \alpha_k = \frac{(r_k, r_k)}{(Ar_k, r_k)} \quad (3.6)$$

2) Градиентный метод с минимальными невязками.

Пусть A положительно-определенная матрица, X_0 - начальное приближение к решению системы $AX = F$. Следующее приближение X_1 ищется так же как и в методе наискорейшего спуска, в виде $X_1 = X_0 + \beta r_0$, но параметр β подбирается так, чтобы минимизировалась длина вектора невязки (r). Таким образом, здесь минимизируется функционал (r, r) в направлении, противоположном градиенту функционала ошибки.

После выполнения первого шага процесс повторяется.

Пусть

$$X_{k+1} = X_k + \beta_k r_k, \quad (3.7)$$

тогда

$$r_{k+1} = r_k - \beta_k Ar_k, \quad (3.8)$$

Нетрудно убедиться, что функционал (r_{k+1}, r_{k+1}) принимает минимальное значение равно

$$(r_{k+1}, r_{k+1}) = (r_k, r_k) - \frac{(Ar_k, r_k)^2}{(Ar_k, Ar_k)}$$

При

$$\beta_k = \frac{(Ar_k, r_k)}{(Ar_k, Ar_k)} \quad (3.9)$$

Порядок выполнения работы

1. Изучить теоретическую часть.
2. Написать программу решения системы одним из градиентных методов.
Отладить программу на модельной задаче.
3. Найти решение системы с заданной точностью.
4. Исследовать скорость сходимости метода в зависимости от числа итераций.
5. Сделать выводы о точности метода, скорости его сходимости. Сравнить результаты с результатами решения с помощью точных методов и методов последовательных приближений .
6. Найти решение системы, воспользовавшись одним из стандартных математических пакетов: MathCad, MathLab и т. п. Сравнить результаты с результатами, полученными с помощью градиентных методов.

Воспользоваться вариантами заданий к лабораторной работе №1.

Лабораторная работа № 4

Методы решения полной проблемы собственных значений

Цель работы – изучение методов решения полной проблемы собственных значений, нахождение всех собственных значений и собственных векторов матрицы методом Крылова или Леверье.

Теоретическая часть

Под полной проблемой собственных значений понимается проблема нахождения всех собственных значений матрицы A и принадлежащих им собственных векторов.

Собственными значениями матрицы A называются корни ее характеристического уравнения:

$$|A - tE| = \begin{vmatrix} a_{11-t} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22-t} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn-t} \end{vmatrix} = (-1)^n (t^n - p_1 t^{n-1} - \dots - p_n) = 0$$

Определение компонент собственного вектора требует решения системы n однородных уравнений с n неизвестными.

Коэффициенты p_i характеристического полинома являются, с точностью до знака, суммами всех миноров определителя матрицы A порядка i , опирающихся на главную диагональ. Поэтому непосредственное вычисление коэффициентов p_i является очень громоздким.

В связи с этим применяются специальные приемы, упрощающие численное решение поставленных задач. Большинство методов, дающих решение полной проблемы собственных значений, заключается в вычислении коэффициентов характеристического полинома, которое осуществляется теми или иными способами, минуя вычисление многочисленных миноров. Собственные значения вычисляются затем по какому-либо методу для приближенного вычисления корней полинома. Одним из простейших методов

приближенного нахождения корней многочлена является метод половинного деления.

Пусть дано уравнение

$$\varphi(t) = 0 \quad (4.1)$$

где функция $\varphi(t)$ непрерывна на $[a, b]$ и $\varphi(a)\varphi(b) < 0$.

Для нахождения корня уравнения (4.1), принадлежащего отрезку $[a, b]$, делим этот отрезок пополам. Если $\varphi\left(\frac{a+b}{2}\right) = 0$, то $t_1 = \frac{a+b}{2}$ является корнем уравнения. Если $\varphi\left(\frac{a+b}{2}\right) \neq 0$, то выбираем ту из половин исходного отрезка, на концах которого функция $\varphi(t)$ имеет различные знаки.

Новый отрезок снова делим пополам и процесс повторяем.

В результате получаем на каком-то этапе или точный корень уравнения (1), или же бесконечную последовательность вложенных друг в друга отрезков $\{[a_n, b_n]\}$, таких что $\varphi(a_n)\varphi(b_n) < 0$, $n=1, 2, \dots$ и $b_n - a_n = \frac{b-a}{2^n}$.

Если требуется найти корень уравнения (4.1) с заданной точностью, то следует продолжать процесс до тех пор, пока длина интервала $b_n - a_n$ не станет меньше заданной точности.

Вычисление значений полинома $\varphi(t)$ в данной точке t_0 следует проводить по схеме Горнера:

$$\begin{array}{cccccc} a_0 & a_1 & a_2 & \dots & a_{n-1} & a_n \\ a_0 & b_1 & b_2 & \dots & b_{n-1} & b_n \end{array},$$

которая заполняется по рекуррентным формулам:

$$b_i = b_{i-1}t_0 + a_i, \quad i = \overline{1, n}$$

Тогда $b_n = \varphi(t_0)$.

Числа $a_0, b_1, \dots, b_{n-1}$ будут коэффициентами частного $\varphi_1(t)$ от деления полинома $\varphi(t)$ на $t-t_0$.

1) Метод Крылова.

Идея метода заключается в предварительном преобразовании уравнения

$$\varphi(t) = \begin{vmatrix} a_{11-t} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22-t} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{2n} & \dots & a_{nn-t} \end{vmatrix} = 0 \quad (4.2)$$

в эквивалентное ему уравнение вида:

$$D(t) = \begin{vmatrix} b_{11-t} & b_{12} & \dots & b_{1n} \\ b_{21-t^2} & b_{22} & \dots & b_{2n} \\ \dots & \dots & \dots & \dots \\ b_{n1-t^n} & b_{2n} & \dots & b_{nn} \end{vmatrix} = 0, \quad (4.3)$$

развертывание которого по степеням t осуществляется гораздо проще.

Нетрудно проверить, что при выборе коэффициентов b_{ik} по следующим формулам

$$\begin{aligned} b_{1k} &= a_{1k}, \\ b_{ik} &= \sum_{j=1}^n b_{i-1,j} a_{jk}, \quad i = \overline{1, n}, \quad k = \overline{1, n} \end{aligned} \quad (4.4)$$

справедливо равенство

$$\frac{D(t)}{\varphi(t)} = \begin{vmatrix} 1 & 0 & \dots & 0 \\ b_{11} & b_{12} & \dots & b_{1n} \\ \dots & \dots & \dots & \dots \\ b_{n-1,1} & b_{n-1,2} & \dots & b_{n-1,n} \end{vmatrix} = N, \quad (4.5)$$

т.е. при $N \neq 0$ $D(t)$ отличается от искомого характеристического полинома только численным множителем.

Зафиксируем вектор

$$B_0 = (1, 0, \dots, 0)^T,$$

тогда справедливо равенство

$$\begin{aligned} B_i &= (b_{i1}, b_{i2}, \dots, b_{in})^T = (A^T)^i B_0, \\ i &= \overline{1, n} \end{aligned}$$

При $N \neq 0$ векторы $B_0, B_1, \dots, B_{n-1}$ образуют базис пространства. Тогда вектор B_n является их линейной комбинацией:

$$B_n = q_1 B_{n-1} + \dots + q_n B_0 \quad (4.6)$$

Нетрудно показать, что коэффициенты этого соотношения и являются коэффициентами p_i характеристического полинома

$$\varphi(t) = (-1)^n [t^n - p_1 t^{n-1} - \dots - p_n]$$

Таким образом, решив систему уравнений (4.6), найдем коэффициенты характеристического полинома

$$q_1 = p_1, \dots, q_n = p_n.$$

Очевидно, что вместо системы (4.6) можно использовать систему

$$C_n = p_1 C_{n-1} + \dots + p_n C_0, \quad (4.7)$$

где векторы C_n определяются равенствами

$$C_k = A^k C_0.$$

Определим теперь собственные векторы $X_1, X_2, \dots, X_n$ матрицы A , принадлежащие собственным числам $\lambda_1, \lambda_2, \dots, \lambda_n$.

Т.к. $X_1, \dots, X_n$ линейно независимы, то справедливо разложение

$$C_0 = \alpha_1 X_1 + \alpha_2 X_2 + \dots + \alpha_n X_n \quad (4.8)$$

Тогда

[illegible]

Нетрудно показать, что собственные векторы могут быть получены в виде линейных комбинаций

$$X_i = \sum_{j=0}^{n-1} \beta_{ij} C_{n-1-j}, \quad (4.9)$$

$$\text{где} \quad \beta_{10}=1, \beta_{ij}=\lambda_i \beta_{i,j-1}-\rho_j, \quad (4.10)$$

$$j = \overline{1, n-1}$$

2) Метод Лаврье и его модификация Д.К. Фаддеевым.

Пусть $\lambda_1, \lambda_2, \dots, \lambda_n$ - корни характеристического полинома матрицы.

Обозначим

$$\sum_{j=1}^n \lambda_j^k = S_k \quad (4.11)$$

Тогда справедливы соотношения Ньютона

$$kp_k = S_k - p_1 S_{k-1} - \dots - p_{k-1} S_1 \quad (4.12)$$

$$k = \overline{1, n}$$

определив числа S_k , можно найти коэффициенты характеристического полинома, решая рекуррентную систему (4.12)

Очевидно, что

$$S_1 = \lambda_1 + \lambda_2 + \dots + \lambda_n = \text{Tr}A = \sum_{i=1}^n a_{ii}$$

Т.к. характеристическими числами матрицы A^k будут

$$\lambda_1^k, \lambda_2^k, \dots, \lambda_n^k, \text{ то}$$

$$S_k = \lambda_1^k + \lambda_2^k + \dots + \lambda_n^k = \text{Tr}A^k \quad (4.13)$$

Таким образом, процесс вычислений сводится к последовательному вычислению степеней матрицы A , затем к вычислению их следов и к решению рекуррентной системы (4.12).

Рассмотрим видоизменение метода, предложенное Фаддеевым, которое кроме упрощений при вычислении коэффициентов характеристического полинома позволяет определить собственные векторы матрицы.

Строится последовательность матриц $A_1, A_2, \dots, A_n$ следующим образом:

$$A_1 = A, \quad \text{Tr}A_1 = q_1, \quad B_1 = A_1 - q_1 E$$

$$A_2 = AB_1, \quad \frac{\text{Tr}A_2}{2} = q_2, \quad B_2 = A_2 - q_2 E$$

$$\dots\dots\dots (4.14)$$

$$A_n = AB_{n-1}, \quad \frac{\text{Tr}A_n}{n} = q_n, \quad B_n = A_n - q_n E$$

Нетрудно показать, что справедливы равенства

$$p_1 = q_1, p_2 = q_2, \dots, p_n = q_n$$

Для определения собственных векторов для каждого собственного значения λ_k строится матрица

$$Q_k = \lambda_k^{n-1}E + \lambda_k^{n-1}B_1 + \dots + B_{n-1} \quad (4.15)$$

Нетрудно убедиться, что каждый столбец матрицы Q_k состоит из компонент собственного вектора, принадлежащего собственному числу λ_k .

Порядок выполнения работы

1. Изучить теоретическую часть.
2. Написать программу нахождения собственных чисел и собственных векторов матрицы. Отладить её на модельной задаче.
3. Найти собственные числа и собственные векторы матрицы.
4. Внести случайную погрешность в коэффициенты матрицы. Определить собственные значения и собственные векторы для измененной матрицы.
5. Сделать выводы о точности метода и устойчивости матрицы.

Варианты заданий

$$1) \quad A = \begin{pmatrix} -5.51 & 1.87 & 0.42 & 0.01 \\ 0.29 & -11.81 & 5.71 & 0.06 \\ 0.05 & 4.31 & -12.97 & 0.23 \\ 0.01 & 0.27 & 1.40 & -17.60 \end{pmatrix}$$

$$2) \quad A = \begin{pmatrix} -7.46 & 11.24 & -0.54 & 14.31 \\ 0.16 & 17.15 & -1.11 & 10.16 \\ 23.17 & 10.26 & 1.19 & 7.43 \\ -28.16 & 0.15 & 0.97 & -3.41 \end{pmatrix}$$

$$3) \quad A = \begin{pmatrix} 10.95 & 1.21 & -7.41 & 0.94 \\ 3.51 & 9.78 & 1.16 & 2.58 \\ 1.95 & 7.36 & 11.21 & -7.43 \\ 10.54 & 11.21 & 7.17 & 2.38 \end{pmatrix}$$

$$4) \quad A = \begin{pmatrix} 1.46 & 23.14 & -0.78 & 1.13 \\ 2.31 & 1.58 & 6.73 & 1.61 \\ -0.13 & -9.21 & 7.41 & 1.23 \\ 0.96 & 1.23 & 3.79 & 5.46 \end{pmatrix}$$

$$5) \quad A = \begin{pmatrix} 7.58 & 16.31 & 2.91 & -0.48 \\ 3.51 & 1.92 & 7.15 & -0.38 \\ -9.61 & 2.34 & 1.15 & 0.97 \\ 7.43 & 2.39 & 5.16 & -1.17 \end{pmatrix}$$

$$6) \quad A = \begin{pmatrix} 3.46 & 9.21 & 7.63 & 3.21 \\ -10.14 & 5.31 & 7.43 & -1.28 \\ 16.54 & 17.23 & 1.28 & 0.38 \\ 1.58 & 7.36 & -2.92 & 0.97 \end{pmatrix}$$

$$7) \quad A = \begin{pmatrix} 5.51 & 2.34 & 7.49 & -6.24 \\ 3.73 & -1.28 & 0.13 & 0.79 \\ 2.49 & 6.51 & 3.22 & 0.34 \\ 4.28 & 0.76 & 2.31 & -10.11 \end{pmatrix}$$

$$8) \quad A = \begin{pmatrix} 93.11 & 0.58 & 1.28 & 2.47 \\ 1.15 & 2.97 & -1.81 & 2.73 \\ 0.38 & -0.43 & 0.15 & 0.97 \\ 2.53 & 9.63 & 11.22 & 7.15 \end{pmatrix}$$

$$9) \quad A = \begin{pmatrix} 21.46 & 3.98 & 11.24 & 0.73 \\ 0.58 & -9.41 & 2.77 & -0.21 \\ 2.38 & 9.97 & 5.41 & 10.11 \\ -1.22 & 0.79 & 6.31 & 22.17 \end{pmatrix}$$

$$10) \quad A = \begin{pmatrix} 7.48 & 16.21 & 9.34 & 0.58 \\ 6.34 & 1.47 & 6.94 & -0.77 \\ -0.15 & 1.24 & 9.98 & 7.63 \\ 10.28 & 1.38 & -7.51 & -2.99 \end{pmatrix}$$

$$11) \quad A = \begin{pmatrix} 51.15 & 2.13 & -1.16 & 0.15 \\ 2.77 & 1.15 & -0.73 & -1.97 \\ 3.48 & 0.97 & 16.21 & 9.32 \\ 3.76 & 0.21 & -0.58 & 0.76 \end{pmatrix}$$

$$12) \quad A = \begin{pmatrix} 0.78 & 0.93 & -9.51 & 6.32 \\ 7.56 & 23.12 & 77.14 & 5.31 \\ 13.21 & 0.79 & 17.16 & 23.19 \\ 3.48 & 0.56 & -27.19 & 0.37 \end{pmatrix}$$

$$13) \quad A = \begin{pmatrix} 7.56 & 23.17 & 6.43 & 0.93 \\ 9.87 & -0.73 & 3.19 & -0.74 \\ 0.21 & 0.79 & -2.34 & 3.78 \\ 5.64 & 1.22 & 0.73 & 1.12 \end{pmatrix}$$

$$14) \quad A = \begin{pmatrix} 10.51 & 23.17 & -1.21 & 0.97 \\ 2.34 & -2.77 & 0.93 & 9.87 \\ 0.43 & 4.31 & 2.78 & 1.73 \\ 10.11 & 23.05 & 7.08 & 0.02 \end{pmatrix}$$

$$15) \quad A = \begin{pmatrix} 2.58 & 10.66 & -2.79 & 0.58 \\ 6.33 & -1.12 & 5.78 & -0.98 \\ 10.52 & 0.38 & -3.43 & 15.22 \\ 1.28 & 0.43 & -1.22 & 1.15 \end{pmatrix}$$

$$16) \quad A = \begin{pmatrix} 5.97 & 2.34 & 1.15 & -1.21 \\ 0.77 & 6.34 & 1.43 & 5.21 \\ 0.34 & 2.48 & 0.77 & -0.12 \\ 2.34 & -5.58 & 7.78 & 2.35 \end{pmatrix}$$

$$17) \quad A = \begin{pmatrix} 9.49 & -2.21 & 0.77 & 5.34 \\ 2.78 & 5.41 & 9.28 & -0.73 \\ 6.28 & 0.73 & 10.52 & 3.48 \\ 0.15 & 2.74 & 0.21 & -5.38 \end{pmatrix}$$

$$18) \quad A = \begin{pmatrix} 5.92 & 0.16 & -1.13 & 10.28 \\ 5.07 & 6.34 & 0.28 & -7.42 \\ 9.54 & 3.49 & 0.16 & -9.21 \\ 0.06 & 0.08 & -3.41 & 4.28 \end{pmatrix}$$

$$19) \quad A = \begin{pmatrix} 5.92 & 0.16 & -1.13 & 10.28 \\ 5.07 & 6.34 & 0.28 & -7.42 \\ 9.54 & 3.49 & 0.16 & -9.21 \\ 0.06 & 0.08 & -3.41 & 4.28 \end{pmatrix}$$

$$20) \quad A = \begin{pmatrix} 10.29 & 0.38 & -0.77 & -5.14 \\ -9.27 & 16.34 & 0.99 & 0.13 \\ 15.54 & -2.17 & 0.97 & 0.18 \\ 4.48 & 2.97 & 6.38 & 1.28 \end{pmatrix}$$

$$21) \quad A = \begin{pmatrix} 0.77 & 15.21 & -3.41 & 0.14 \\ 0.07 & 0.29 & 16.11 & 2.34 \\ 9.77 & 0.19 & 3.51 & -6.24 \\ 3.21 & 1.15 & 1.28 & 0.76 \end{pmatrix}$$

$$22) \quad A = \begin{pmatrix} 9.32 & 1.28 & 7.46 & -1.15 \\ 0.77 & 2.34 & -5.41 & 0.14 \\ -10.21 & 3.42 & 6.54 & 7.28 \\ 3.48 & 6.92 & 0.34 & 0.07 \end{pmatrix}$$

$$23) \quad A = \begin{pmatrix} 2.73 & -1.15 & 2.34 & 0.28 \\ 7.56 & 2.34 & 1.94 & 0.07 \\ 5.68 & 0.39 & 7.49 & 2.34 \\ -1.58 & 0.37 & 0.01 & 2.97 \end{pmatrix}$$

$$24) \quad A = \begin{pmatrix} 1.16 & 2.34 & -0.92 & 1.17 \\ 0.56 & 1.13 & 1.28 & 3.43 \\ 10.07 & 6.34 & -1.27 & 7.49 \\ 0.08 & 3.58 & 4.92 & -1.28 \end{pmatrix}$$

$$25) \quad A = \begin{pmatrix} 9.91 & 0.34 & 7.43 & -2.51 \\ 0.98 & 7.35 & 4.21 & -0.12 \\ 2.34 & 7.39 & 5.77 & 6.34 \\ 2.91 & 10.15 & 3.28 & 0.79 \end{pmatrix}$$

СОДЕРЖАНИЕ

ПРЕДИСЛОВИЕ.....	5
Основные понятия.....	7
Глава I. ЧИСЛЕННЫЕ МЕТОДЫ ЛИНЕЙНОЙ АЛГЕБРЫ.....	8
1. Погрешности вычислений. Обусловленность матриц. Точные методы решения систем линейных алгебраических уравнений (СЛАУ).....	8
2. Метод Гаусса.....	13
3. Компактные схемы для решения неоднородных линейных систем.....	16
4. Метод Гаусса-Жордана. Вычисление определителей. Обращение матриц.....	17
5. Метод квадратных корней. Метод ортогонализации.....	22
6. Принципы построения итерационных процессов. Метод простой итерации.....	25
7. Необходимые и достаточные условия сходимости простой итерации. Подготовка системы к виду, удобному для итераций.....	29
8. Метод Зейделя.....	32
9. Методы полной и неполной релаксации. Управление релаксации.....	36
10. Релаксация по длине вектора невязки. Групповая релаксация.....	40
11. Градиент функционала. Градиентные методы. Метод наискорейшего спуска. Метод с наименьшими невязками.....	44
12. Методы сопряженных направлений. S-шаговые градиентные методы.....	50
13. Полная проблема собственных значений. Устойчивость проблемы собственных значений. Метод Крылова. Нахождение собственных векторов по методу Крылова	57
14. Метод Леверье и его модификация Фаддеевым.....	71
15. Эскалаторный метод.....	74
16. Определение наибольшего по модулю собственного значения и принадлежащего ему собственного вектора.....	78
17. Метод скалярных произведений. Метод координатной релаксации.....	86
18. Метод λ -разности. Уточнение отдельного собственного значения и принадлежащего ему собственного вектора.....	90
Глава II. ОСНОВНЫЕ АЛГЕБРАИЧЕСКИЕ СТРУКТУРЫ.....	93
1. Группы.....	93
1.1. Определение и примеры.....	93
1.2. Изоморфизм групп.....	95

1.3. Гомоморфизм.....	96
1.4. Подгруппы. Циклические группы	97
1.5. Смежные классы. Разложение группы по подгруппе.....	99
1.6. Нормальный делитель группы.....	100
1.7. Фактор-группа.....	101
1.8. Группы подстановок.....	101
2. Кольца.....	102
2.1. Определения, примеры.....	102
2.2. Изоморфизм. Гомоморфизм.....	105
2.3. Подкольца. Идеалы.....	105
3. Поля. Тела.....	106
3.1. Поля.....	106
3.2. Тела.....	109
3.3. Алгебры.....	109
Глава III. ЛАБОРАТОРНЫЙ ПРАКТИКУМ	110
Требования к оформлению отчета о выполнении лабораторной работы....	110
<i>Лабораторная работа №1</i>	
Точные методы решения систем линейных алгебраических уравнений.....	111
<i>Лабораторная работа №2.</i>	
Итерационные методы решения систем уравнений.....	119
<i>Лабораторная работа № 3.</i>	
Градиентные методы решения систем уравнений.....	124
<i>Лабораторная работа № 4.</i>	
Методы решения полной проблемы собственных значений.....	127
СПИСОК ЛИТЕРАТУРЫ.....	137