ВВЕДЕНИЕ

История развития тех разделов линейной алгебры, о которых идет речь в настоящем пособии, насчитывает не одно столетие. Метод Гаусса[1] и правило Крамера[2] разработаны в 18 веке, метод квадратных корней, называемый также методом Холецкого, – в начале 20 века, а многие методы решения систем линейных алгебраических уравнений (СЛАУ), основная матрица которых имеет те или иные особенности (симметричность, ленточность, разреженность и т.д.), разработаны в последние десятилетия. Исследования по созданию эффективных численных методов решения СЛАУ и других задач линейной алгебры активно продолжаются и в настоящее время, публикации по этой тематике исчисляются сотнями единиц.

В настоящем пособии не ставилась цель познакомить читателя с максимальным числом методов решения СЛАУ и проблемы собственных значений. Задача пособия двоякая: с одной стороны – изучение наиболее широко применяемых на практике методов решения некоторых классов СЛАУ, с другой - акцентирование внимания читателя на проблемах, которые возникают при реализации на ЭВМ алгоритмов решения задач не только линейной алгебры, но и других разделов математики. Некоторые из обсуждаемых проблем – количество вычислительных операций, расположение данных в памяти ЭВМ – наиболее актуальны для задач линейной алгебры. Число неизвестных в СЛАУ может иметь порядок 10^5–10^6, число элементов основной матрицы СЛАУ – соответственно 10^{10}–10^{12}. Матрицы с такими размерами, как правило, имеют большое число нулевых элементов (являются разреженными). Учет этих особенностей позволяет строить алгоритмы, экономные по затратам памяти и числу вычислительных операций.

Еще одна проблема, возникающая при реализации на ЭВМ алгоритмов решения тех или иных задач, – вычислительная устойчивость численных методов. Порождается она процедурой округления чисел в ЭВМ, которая и приводит к вычислительным ошибкам. Например известно, что метод Гаусса позволяет найти решение СЛАУ, если оно существует и единственно, за конечное число шагов. Естественно, при этом предполагается, что вычисления на каждом шаге выполняются абсолютно точно. Приближенное решение этой СЛАУ на ЭВМ, полученное методом Гаусса, может существенно отличаться от точного. Более того, процесс решения может даже прекратиться из-за неограниченного роста ошибок вычислений.

[1]Gauss Karl Friedrich (30.4.1777 –23.2.1855) – немецкий математик (исследования по алгебре, геометрии, теоретической физике; фундаментальные результаты по астрономии и геодезии).
[2]Cramer Gabriel (31.7.1704 – 4.1.1752) – швейцарский математик (исследования систем линейных уравнений, теория определителей).

Все это должен учитывать специалист из той или иной предметной области при выборе метода и программного обеспечения для решения прикладной задачи.

Значком ⊗ обозначены окончания доказательства теорем и утверждений.

1. ЧИСЛЕННЫЕ МЕТОДЫ РЕШЕНИЯ СИСТЕМ ЛИНЕЙНЫХ АЛГЕБРАИЧЕСКИХ УРАВНЕНИЙ

1.1. Метод квадратных корней

1.1.1. Идея метода

Пусть

$$\mathbf{AX} = \mathbf{B} \qquad (1.1)$$

- система линейных алгебраических уравнений (СЛАУ) с симметричной основной матрицей $\mathbf{A} = (\mathbf{a_{ij}})$ размером $\mathbf{n \times n}$; $\mathbf{B} = (\mathbf{b_1, b_2, ..., b_n})^\mathbf{T}$ – вектор свободных членов СЛАУ и $\mathbf{X} = (\mathbf{x_1, x_2, ..., x_n})^\mathbf{T}$ – вектор неизвестных.

Представим матрицу $\mathbf{A}$ в виде произведения двух треугольных матриц

$$\mathbf{A} = \mathbf{R^T R}, \qquad (1.2)$$

где $\mathbf{R}$ – верхняя треугольная матрица, а $\mathbf{R^T}$ – транспонированная к ней, т.е.

$$\mathbf{R} = \begin{pmatrix} r_{11} & r_{12} & ... & r_{1n} \\ 0 & r_{22} & ... & r_{2n} \\ . & . & . & . \\ 0 & 0 & ... & r_{nn} \end{pmatrix}, \quad \mathbf{R^T} = \begin{pmatrix} r_{11} & 0 & ... & 0 \\ r_{12} & r_{22} & ... & 0 \\ . & . & . & . \\ r_{1n} & r_{2n} & ... & r_{nn} \end{pmatrix}.$$

Разложение (1.2) матрицы $\mathbf{A}$ в произведение матриц $\mathbf{R^T R}$ называется ***факторизацией*** матрицы $\mathbf{A}$. [Возможно оно, как будет показано ниже, если $\mathbf{r_{ii} \neq 0}$].

С учетом (1.2) СЛАУ (1.1) принимает следующий вид

$$\left(\mathbf{R^T R}\right)\mathbf{X} = \mathbf{B}.$$

Используя свойство ассоциативности произведения матриц, можем записать

$$\mathbf{R^T}\left(\mathbf{RX}\right) = \mathbf{B}.$$

Введем обозначение

$$\mathbf{RX} = \mathbf{Y}.$$

В результате СЛАУ (1.1) будет эквивалентна двум СЛАУ с треугольными матрицами:

$$\mathbf{AX = B} \sim \begin{cases} \mathbf{R^T Y = B}, \\ \mathbf{RX = Y}. \end{cases}$$

Решая первую СЛАУ $\mathbf{R^T Y = B}$, находим вспомогательный вектор $\mathbf{Y}$ - вектор свободных членов для СЛАУ $\mathbf{RX = Y}$. Решая СЛАУ $\mathbf{RX = Y}$, находим искомый вектор **X.**

Таким образом, для решения СЛАУ (1.1) с симметричной матрицей **A** необходимо:

1) осуществить факторизацию матрицы $\mathbf{A}$: $\mathbf{A = R^T R}$;
2) решить последовательно две СЛАУ с треугольными матрицами коэффициентов:

$$\mathbf{R^T Y = B \Rightarrow Y, RX = Y \Rightarrow X.}$$

1.1.2. Факторизация матрицы A

Из условия $\mathbf{R^T R = A}$

$$\mathbf{R^T R} = \begin{pmatrix} r_{11} & 0 & \dots & 0 \\ r_{12} & r_{22} & \dots & 0 \\ . & . & . & . \\ r_{1n} & r_{2n} & \dots & r_{nn} \end{pmatrix} \cdot \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ 0 & r_{22} & \dots & r_{2n} \\ . & . & . & . \\ 0 & 0 & \dots & r_{nn} \end{pmatrix} =$$

$$= \begin{pmatrix} r_{11}^2 & r_{11}r_{12} & \dots & r_{11}r_{1n} \\ r_{11}r_{12} & r_{12}^2 + r_{22}^2 & \dots & r_{12}r_{1n} + r_{22}r_{2n} \\ . & . & . & . \\ r_{11}r_{1n} & r_{12}r_{1n} + r_{22}r_{2n} & \dots & r_{1n}^2 + r_{2n}^2 + r_{nn}^2 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{12} & a_{22} & \dots & a_{2n} \\ . & . & . & . \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{pmatrix}$$

получаем систему $\frac{n(n+1)}{2}$ уравнений для нахождения $\frac{n(n+1)}{2}$ элементов $r_{ij}, i = \overline{1,n}; j = i,n$ матрицы $\mathbf{R}$[1]:

$$\begin{array}{lll} r_{11}^2 = a_{11}, & r_{11}r_{1j} = a_{1j}, & j = 2,3,\dots,n, \\ r_{12}^2 + r_{22}^2 = a_{22}, & r_{12}r_{1j} + r_{22}r_{2j} = a_{2j}, & j = 3,4,\dots,n, \\ . & . & . \\ r_{1k}^2 + r_{2k}^2 + \dots + r_{kk}^2 = a_{kk}, & r_{1k}r_{1j} + r_{2k}r_{2j} + \dots + r_{kk}r_{kj} = a_{kj}, & j = k, k+1,\dots,n, \\ . & . & . \\ r_{1n}^2 + r_{2n}^2 + \dots + r_{nn}^2 = a_{nn}. & & \end{array} \quad (1.3)$$

[1] Запись $\mathbf{i = \overline{1,n}}$ означает, что $\mathbf{i}$ принимает значения $\mathbf{1;2;\dots;n}$; соответственно $\mathbf{j = \overline{1,n} \Leftrightarrow j \in \{i; i+1; \dots n\}}$.

Из (1.3) последовательно находим

$$\begin{cases} r_{11} = \sqrt{a_{11}}; \quad r_{1j} = \dfrac{a_{1j}}{r_{11}}, \quad j = \overline{2,n}, \\ r_{22} = \sqrt{a_{22} - r_{12}^2}; \quad r_{2j} = \dfrac{a_{1j} - r_{12}r_{1j}}{r_{11}}, \quad j = \overline{3,n}, \\ r_{kk} = \sqrt{a_{kk} - (r_{1k}^2 + r_{2k}^2 + \dots + r_{k-1,k}^2)}, \\ r_{kj} = \dfrac{a_{kj} - (r_{1k}r_{1j} + r_{2k}r_{2j} + \dots + r_{k-1,k}r_{k-1,j})}{r_{kk}}, \; k = \overline{3,n-1}, j > k, \\ r_{nn} = \sqrt{a_{nn} - (r_{1n}^2 + r_{2n}^2 + \dots + r_{n-1,n}^2)}. \end{cases} \quad (1.4)$$

Из (1.4) следует, что система уравнений (1.3) однозначно разрешима, если $\mathbf{r_{kk} \neq 0, \forall k \in \{1,2,\dots,n\}}$. Это, в свою очередь, достаточное условие существования и единственности решения СЛАУ (1.1).

Формулы (1.4) не очень удобны для запоминания. Целесообразно пользоваться векторной формой записи этих формул для произвольного $\mathbf{k \in \{2,3,\dots,n\}}$.

Введем обозначения

$$\begin{aligned} \overline{r_k} &= (r_{1k}, r_{2k}, \dots, r_{k-1,k}), \\ \overline{r_j} &= (r_{1j}, r_{2j}, \dots, r_{k-1,j}). \end{aligned} \quad (1.5)$$

Здесь $\overline{\mathbf{r_k}}$ - вектор элементов **k**-го столбца матрицы **R**, расположенных в первых $(\mathbf{k-1})$-х строках (выделены жирным шрифтом в матрице **R**). Аналогично определяется вектор $\overline{\mathbf{r_j}}$ - вектор элементов **j**-го столбца $(\mathbf{j \in \{k+1, k+2, \dots, n\}})$, расположенных в первых $(\mathbf{k-1})$-х строках (также выделены жирным шрифтом).

$$R = \begin{pmatrix} r_{11} & r_{12} & \dots & \mathbf{r_{1k}} & \dots & \mathbf{r_{1j}} & \dots & r_{1n} \\ 0 & r_{22} & \dots & \mathbf{r_{2k}} & \dots & \mathbf{r_{2j}} & \dots & r_{2n} \\ . & . & . & . & . & . & . & . \\ 0 & 0 & \dots & \mathbf{r_{k-1,k}} & \dots & \mathbf{r_{k-1,j}} & \dots & r_{k-1,n} \\ 0 & 0 & \dots & r_{kk} & \dots & r_{kj} & \dots & r_{kn} \\ . & . & . & . & . & . & . & . \\ 0 & 0 & \dots & 0 & \dots & r_{jj} & \dots & r_{jn} \\ . & . & . & . & . & . & . & . \\ 0 & 0 & \dots & 0 & . & 0 & \dots & r_{nn} \end{pmatrix}.$$

Тогда

$$\begin{aligned} r_{kk} &= \sqrt{a_{kk} - (\overline{r_k}, \overline{r_k})},\ k = \overline{2,n}, \\ r_{kj} &= \frac{a_{kj} - (\overline{r_k}, \overline{r_j})}{r_{kk}},\quad j = \overline{k+1,n}. \end{aligned} \tag{1.6}$$

Здесь $(\overline{r_k}, \overline{r_k})$, $(\overline{r_k}, \overline{r_j})$- скалярные произведения арифметических векторов. Действительно,

$$\det A = \det(R^T R) = \det R^T \cdot \det R = r_{11}^2 \cdot r_{22}^2 \cdot \ldots \cdot r_{nn}^2$$

и если $r_{kk} \neq 0\ \forall k \in \{1,2,\ldots,n\}$, то $\det A \neq 0$, $rg(A|B) = rg(A) = n$ и в соответствии с теоремой Кронекера-Капелли СЛАУ (1.1) совместна и имеет единственное решение.

Выясним теперь, при каких дополнительных условиях действительная матрица A $(a_{ij} \in R)$ разлагается в произведение $R^T \cdot R$ с действительной матрицей R. Из формул (1.4) следует, что комплексные элементы r_{ij} в матрице R могут появиться, если среди чисел $a_{11}; a_{kk} - (r_{1k}^2 + r_{2k}^2 + \ldots + r_{k-1,k}^2) = r_{kk}^2$, $k = \overline{2,n}$, будет хотя бы одно отрицательное число.

Справедлива следующая теорема.

<u>Теорема 1.1.</u> Действительная матрица A разлагается в произведение $R^T \cdot R$ с действительной матрицей R тогда и только тогда, когда матрица A определена положительно.

<u>Доказательство</u>. Введем обозначения

$$A_k = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1k} \\ a_{12} & a_{22} & \ldots & a_{2k} \\ . & . & . & . \\ a_{1k} & a_{2k} & \ldots & a_{kk} \end{pmatrix},\ R_k = \begin{pmatrix} r_{11} & r_{12} & \ldots & r_{1k} \\ 0 & r_{22} & \ldots & r_{2k} \\ . & . & . & . \\ 0 & 0 & \ldots & r_{kk} \end{pmatrix}.$$

Матрицы A_k и R_k являются подматрицами порядка k матриц A и R соответственно, расположенными в левых верхних углах этих матриц. Легко установить, что если имеет место разложение $A = R^T R$, то и $A_k = R_k^T R_k$. При этом

$$\begin{aligned} |A| &= |R^T R| = |R^T|\,|R| = \prod_{i=1}^{n} r_{ii}^2; \\ |A_k| &= |R_k^T R_k| = |R_k^T|\,|R_k| = \prod_{i=1}^{k} r_{ii}^2. \end{aligned} \tag{1.7}$$

Обозначим $\Delta_k = |A_k|$ - угловые миноры матрицы A. Из (1.7) с учетом (1.3) последовательно получаем

$\Delta_1 = |A_1| = a_{11} = r_{11}^2;$

$\Delta_2 = |A_2| = |R_2^T| \cdot |R_2| = r_{11}^2 \cdot r_{22}^2 = \Delta_1 \cdot r_{22}^2;$

$\Delta_3 = |A_3| = |R_3^T| \cdot |R_3| = r_{11}^2 \cdot r_{22}^2 \cdot r_{33} = |r_{11}^2 \cdot r_{22}^2 = \Delta_2| = \Delta_2 \cdot r_{33};$

. . . .

$\Delta_n = |A_n| = |R_n^T| \cdot |R_n| = r_{11}^2 \cdot r_{22}^2 \ldots r_{nn}^2 = |r_{11}^2 \cdot r_{22}^2 \ldots r_{n-1,n-1}^2 = \Delta_{n-1}| = \Delta_{n-1} \cdot r_{nn},$

откуда имеем $\quad r_{11}^2 = a_{11} = \Delta_1;\ \ r_{kk}^2 = \dfrac{\Delta_k}{\Delta_{k-1}},\ k = \overline{2,n}.$

Последние равенства позволяют сделать искомый вывод: если числа r_{kk}^2, $k = \overline{1,n}$, положительны, то и все угловые миноры Δ_k, $k = \overline{1,n}$, положительны и, наоборот – из положительности угловых миноров Δ_k, $k = \overline{1,n}$, следует положительность чисел r_{kk}^2. ⊗

Таким образом, положительная определенность симметричной матрицы A обеспечивает существование действительных значений корней $\sqrt{a_{11}}$, $\sqrt{a_{kk} - (r_{1k}^2 + r_{2k}^2 + \ldots + r_{k-1,k}^2)}$, $k = \overline{2,n}$. В противном случае матрица R будет комплексной $(r_{ij} \in C)$, даже если матрица A действительная.

1.1.3. Решение треугольных СЛАУ

СЛАУ $R^T Y = B$ и $RX = Y$ являются “треугольными“ и потому их решения находятся сразу (обратный ход метода Гаусса)

$$\begin{cases} r_{11}y_1 = b_1, \\ r_{12}y_1 + r_{22}y_2 = b_2, \\ \ldots\ldots\ldots\ldots \\ r_{1n}y_1 + r_{2n}y_2 + \ldots + r_{nn}y_n = b_n, \end{cases} \Rightarrow \begin{cases} y_1 = b_1/r_{11}, \\ y_2 = (b_2 - r_{12}y_1)/r_{22}, \\ \ldots\ldots\ldots\ldots \\ y_n = \left(b_n - \sum\limits_{i=1}^{n-1} r_{in}y_i\right)\Big/ r_{nn}, \end{cases} \tag{1.8}$$

$$\begin{cases} r_{11}x_1 + r_{12}x_{21} + \ldots + r_{1n}x_n = y_1, \\ r_{12}x_2 + \ldots + r_{2n}x_n = y_2, \\ \ldots\ldots\ldots\ldots \\ r_{n-1,n-1}x_{n-1} + r_{n-1,n}x_n = y_{n-1}, \\ r_{nn}x_n = y_n \end{cases} \Rightarrow \begin{cases} x_n = y_n/r_{nn}, \\ x_{n-1} = (y_{n-1} - r_{n-1,n}x_n)/r_{n-1,n-1}, \\ \ldots\ldots\ldots\ldots \\ x_1 = \left(y_1 - \sum\limits_{j=2}^{n} r_{1j}x_j\right)\Big/ r_{11}. \end{cases} \tag{1.9}$$

1.1.4. Оценивание затрат памяти ЭВМ и числа вычислительных операций

Для хранения коэффициентов a_{ij} (верхнего или нижнего “треугольника”, так как $a_{ij} = a_{ji}, \forall i \neq j$), необходимо $n + (n-1) + (n-2) + \ldots + 1 = \frac{n(n+1)}{2}$ ячеек памяти. В этих же ячейках можно поместить и элементы r_{ij} матрицы $\mathbf{R}$. Еще n ячеек необходимо для хранения коэффициентов $b_i, i = \overline{1,n}$ и в них же - векторов $\mathbf{Y}$ и $\mathbf{X}$. В результате необходимо $\frac{n(n+3)}{2}$ ячеек основной памяти ЭВМ.

Вычисление числа операций. Здесь и далее символами $N(\pm)$, $N(*)$, $N(/)$, $N(\sqrt{\ })$ обозначены соответственно число операций сложения, умножения, деления и извлечения квадратного корня.

А) факторизация матрицы $\mathbf{A}$ [(формулы 1.4)]

А.1) вычисление коэффициентов r_{kk}, $k = \overline{1,n}$:

$$
\begin{aligned}
&N(\sqrt{\ }) = n,\\
&N(/) = 0,\\
&N(*) = 1 + 2 + \ldots + (n-1) = \frac{(n-1)n}{2},\\
&N(\pm) = \frac{(n-1)n}{2};
\end{aligned}
\tag{1.10}
$$

А.2) вычисление коэффициентов $r_{kj}, k = \overline{1, n-1}, j = \overline{k+1, n}$

$$N(\sqrt{\ }) = 0,$$

$$N(/) = (n-1) + (n-2) + \ldots + 1 = \frac{n(n-1)}{2},$$

$$N(*) = 1 \cdot (n-2) + 2 \cdot (n-3) + \ldots + (n-2) \cdot 1 = \frac{n(n-1)(n-2)}{6}, \tag{1.11}$$

$$N(\pm) = N(*) = \frac{n(n-1)(n-2)}{6}.$$

В формуле (1.11): $1 \cdot (n-2)$ – число умножений, необходимое для вычисления коэффициентов r_{2j}, $j = \overline{3,n}$; $2 \cdot (n-3)$ –число умножений, необходимое для вычисления коэффициентов $r_{3j}, j = \overline{4,n}$, и т.д; $(n-2) \cdot 1$ – число умножений, необходимое для вычисления коэффициента $r_{n-1,n}$.

<u>Доказательство формулы(1.11):</u>

$$
\begin{aligned}
&1(n-2) + 2(n-3) + \ldots + (n-2) \cdot 1 = n + 2 \cdot n + \ldots + (n-2)n -\\
&- [1 \cdot 2 + 2 \cdot 3 + \ldots + (n-2)(n-1)] = \frac{(n-2)(n-1)n}{2} - [1 \cdot (1+1) + 2 \cdot (2+1) + \ldots +\\
&+ (n-2)((n-1)+1)] = \frac{(n-2)(n-1)n}{2} - \left(1^2 + 2^2 + \ldots + (n-2)^2\right) -\\
&- (1 + 2 + \ldots + (n-2)) = \frac{(n-2)(n-1)n}{2} - \frac{(n-2)(n-1)(2n-3)}{6} - \frac{(n-1)(n-2)}{2} =\\
&= \frac{(n-2)(n-1)}{6}(3n - 2n + 3 - 3) = \frac{n(n-1)(n-2)}{6}.
\end{aligned}
$$

При выводе формулы (1.11) были использованы следующие очевидные соотношения $(n-2) = (n-2)(n-(n-1))$; $(n-3) = (n-3)(n-(n-2))$ и формула $\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$.

Основные затраты времени ЭВМ при реализации вычислительных операций приходятся на “медленные” операции умножения, деления и нахождения квадратного корня. Общее число таких операций на этапе факторизации матрицы $\mathbf{A}$ равно:

$$N(\sqrt{\ };*;/) = n + \frac{n(n-1)}{2} + \frac{n(n-1)}{2} + \frac{n(n-1)(n-2)}{6} =$$

$$= \frac{n(n+1)(n+2)}{6} = \frac{n^3}{6} + 0(n^2).$$

В) Решение треугольных СЛАУ

Состав и число операций при решении каждой треугольной СЛАУ совпадают. Для одной СЛАУ имеем из (1.8)

$$N(/) = n,$$

$$N(*) = 1 + 2 + \ldots + (n-1) = \frac{n(n-1)}{2},$$

$$N(\pm) = N(*) = \frac{n(n-1)}{2}.$$

С) Общее число вычислительных операций

С.1) медленных

$$N(\sqrt{\ };*;/) = \frac{n(n+1)(n-2)}{6} + 2\left(n + \frac{n(n-1)}{2}\right) = \frac{n(n+1)(n+8)}{6} = \frac{n^3}{6} + 0(n^2);$$

С.2) быстрых

$$N\left(\sqrt{\ };*;/\right)=\underbrace{\frac{n(n-1)}{2}+\frac{n(n-1)(n-2)}{6}}_{\text{Факторизация A}}+\underbrace{2\frac{n(n-1)}{2}}_{\text{Решение СЛАУ}}=\frac{n(n-1)(n+7)}{6}=$$

$$=\frac{n^3}{6}+O\left(n^2\right).$$

1.1.5. Выводы

1. В методе Гаусса число медленных операций равно $\frac{n^3}{3}+0\left(n^2\right)$. В результате метод квадратного корня быстрее метода Гаусса приблизительно в 2 раза.
2. Метод Гаусса применим для исследования совместности и решения произвольных СЛАУ. Метод квадратного корня применим для решения “квадратных” СЛАУ с <u>*симметричной*</u> основной матрицей СЛАУ. Естественно, эти ограничения существенно сужают область применения метода квадратного корня. Вместе с тем, при решении некоторых, достаточно широко распространенных задач, появляется СЛАУ вида

$$\left(K^TK\right)X=F \tag{1.12}$$

с основной матрицей СЛАУ $A=K^TK$, где $\dim\ K=m\times n$ и $m<n$. Если матрица K является матрицей полного ранга, т.е. $rgK=m,(m<n)$, то матрица $A=K^TK$ будет <u>***положительно определенной***</u>. Кроме того, матрица A - симметрична. Поэтому для решения СЛАУ (1.12) <u>***можно***</u> и <u>***целесообразно***</u> использовать метод квадратного корня.

СЛАУ вида (1.12) появляется, например, в методе наименьших квадратов и при решении интегральных уравнений Фредгольма первого рода.

Пусть $y_i=a_1x_{1i}+...+a_nx_{ni}+\varepsilon_i$, $i=\overline{1,N}$, - уравнение множественной линейной регрессии или в векторно-матричной форме $Y=A^TX+E$, где Y – вектор значений случайной величины $y\left(Y=(y_1,y_2,...,y_n)^T\right)$,

A – вектор неизвестных параметров $\left(A=(a_1,a_2,...,a_n)^T\right)$,

X – матрица значений факторов $x_k\left(X=(x_{ij})_{N\times n}\right)$,

E – вектор случайных отклонений $\left(E=(\varepsilon_1,...,\varepsilon_N)^T\right)$. Неизвестные коэффициенты $a_1,...,a_n$ уравнения регрессии находятся из условия минимума суммы квадратов отклонений

$$F(A)=\sum_{i=1}^{N}\left(y_i-(a_1x_{1i}+...+a_nx_{ni})\right)^2=(Y-XA)^T(Y-XA)\to\min_A.$$

Из необходимого условия минимума функции $F(A)$ n переменных $\frac{dF(A)}{dA}=0$ получаем

$$\frac{dF}{dA}=-2X^TY+2X^TXA=0$$

или

$$X^TXA=X^TY.$$

А это СЛАУ вида (1.12) с положительно определенной матрицей $\left(X^TX\right)$. Решение интегрального уравнения Фредгольма первого рода $\int_a^b K(t,\tau)x(\tau)d\tau=f(t)$, $a\le t\le b$ или в операторной форме $K\cdot x=f$, $Kx\equiv\int_a^b K(t,\tau)x(\tau)d\tau$ находится в результате решения уравнения Эйлера

$$\left(K^*K+\alpha G\right)X=K^*f. \tag{1.13}$$

Здесь K^* – оператор, сопряженный к оператору K, α – параметр регуляризации, G – дифференциальный оператор. Дискретным аналогом уравнения Эйлера на заданной (t,τ)-сетке является СЛАУ

$$\left(K^TK+\alpha G\right)X=K^TF\ . \tag{1.14}$$

Матрица $K^TK+\alpha G$ этой СЛАУ при соответствующем задании оператора G будет симметричной и положительно определенной. Поэтому решать СЛАУ (1.14) также целесообразно методом квадратного корня.

1.1.6. Примеры

<u>***Пример 1.1***</u>. Решить СЛАУ

$$\begin{cases}4x_1+2x_2-2x_3\qquad=4,\\2x_1+2x_2+x_3-2x_4=3,\\-2x_1+x_2+6x_3-3x_4=2,\\\qquad-2x_2-3x_3+6x_4=1.\end{cases}$$

Матрица A СЛАУ является симметричной

$$A=\begin{pmatrix}4&2&-2&0\\2&2&1&-2\\-2&1&6&-3\\0&-2&-3&6\end{pmatrix}.$$

1. Проверим ее на знакоопределенность:

$$\Delta_1 = 4 > 0,\ \Delta_2 = \begin{vmatrix} 4 & 2 \\ 2 & 2 \end{vmatrix} = 4 > 0,\ \Delta_3 = \begin{vmatrix} 4 & 2 & -2 \\ 2 & 2 & 1 \\ -2 & 1 & 6 \end{vmatrix} = \begin{vmatrix} 0 & -2 & -4 \\ 2 & 2 & 1 \\ 0 & 3 & 7 \end{vmatrix} = 4 > 0,$$

$$\Delta_4 = \begin{vmatrix} 4 & 2 & -2 & 0 \\ 2 & 2 & 1 & -2 \\ -2 & 1 & 6 & -3 \\ 0 & -2 & -3 & 6 \end{vmatrix} = \begin{vmatrix} 0 & -2 & -4 & 4 \\ 2 & 2 & 1 & -2 \\ 0 & 3 & 7 & -5 \\ 0 & -2 & -3 & -6 \end{vmatrix} = 4 > 0.$$

В соответствии с критерием Сильвестра матрица **A** является положительно определенной и потому результатом ее факторизации будет действительная матрица **R**.

2. Факторизация

$$r_{11} = \sqrt{4} = 2,\ r_{12} = \frac{2}{2} = 1,\ r_{13} = \frac{-2}{2} = -1,\ r_{14} = \frac{0}{2} = 0;$$

$$r_{22} = \sqrt{2 - 1^2} = 1,\ r_{23} = (1 - 1\cdot(-1))/1 = 2,\ r_{24} = (-2 - 1\cdot 0)/1 = -2;$$

$$r_{33} = \sqrt{6 - (-1)^2 - (2)^2} = 1,\ r_{34} = (-3 - (-1)\cdot 0 - 2(-2))/1 = 1;$$

$$r_{44} = \sqrt{6 - 0^2 - (-2)^2 - (1)^2} = 1.$$

$$R = \begin{pmatrix} 2 & 1 & -1 & 0 \\ 0 & 1 & 2 & -2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}, R^T = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ -1 & 2 & 1 & 0 \\ 0 & -2 & 1 & 1 \end{pmatrix}.$$

3. Решение треугольных СЛАУ

$$R^T Y = B: \begin{cases} 2y_1 = 4, \\ y_1 + y_2 = 3, \\ -y_1 + 2y_2 + y_3 = 2, \\ -2y_2 + y_3 + y_4 = 1, \end{cases} \Rightarrow \begin{cases} y_1 = 2, \\ y_2 = 3 - 2 = 1, \\ y_3 = 2 + 2 - 2 = 2, \\ y_4 = 1 + 2\cdot 1 - 2 = 1. \end{cases}$$

$$RX = Y: \begin{cases} 2x_1 + x_2 - x_3 = 2, \\ x_2 + 2x_3 - 2x_4 = 1, \\ x_3 + x_4 = 2, \\ x_4 = 1, \end{cases} \Rightarrow \begin{cases} x_4 = 1, \\ x_3 = 2 - 1 = 1, \\ x_2 = 1 - 2\cdot 1 + 2\cdot 1 = 1, \\ x_1 = (2 - 1 + 1)/2 = 1. \end{cases}$$

Ответ: $X = (1,\ \ 1,\ \ 1,\ \ 1)^T$.

1.2. Метод итераций

1.2.1. Алгоритм метода итераций. Примеры

Пусть

$$AX = B \tag{1.15}$$

СЛАУ порядка **n**, т.е.

$$\begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n = b_1, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}x_1 + \ldots + a_{nn}x_n = b_n. \end{cases} \tag{1.16}$$

Допустим, что $a_{ii} \neq 0,\ \forall i = \overline{1,n}$. Разделим обе части каждого уравнения (1.16) на диагональные коэффициенты a_{ii} и преобразуем СЛАУ (1.16) к следующему виду:

$$\begin{cases} x_1 = \dfrac{b_1}{a_{11}} + 0\cdot x_1 + \left(-\dfrac{a_{12}}{a_{11}}\right)x_2 + \ldots + \left(-\dfrac{a_{1n}}{a_n}\right)x_n, \\ x_2 = \dfrac{b_2}{a_{22}} + \left(-\dfrac{a_{21}}{a_{22}}\right)x_1 + 0\cdot x_2 + \ldots + \left(-\dfrac{a_{2n}}{a_{22}}\right)x_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ x_n = \dfrac{b_n}{a_{nn}} + \left(-\dfrac{a_{n1}}{a_{nn}}\right)x_1 + \left(-\dfrac{a_{n2}}{a_{nn}}\right)x_n + \ldots + 0\cdot x_n. \end{cases} \tag{1.17}$$

Введем обозначения: $\beta_i = \dfrac{b_i}{a_{ii}}, i = \overline{1,n};\quad \gamma_i = \dfrac{a_{ij}}{a_{ii}},\ i, j = \overline{1,n}, i \neq j;\ \gamma_{ii} = 0,$ $i = \overline{1,n}.$

С учетом введенных обозначений СЛАУ (1.17) будет иметь следующий вид:

$$\begin{cases} x_1 = \beta_1 + 0x_1 + \gamma_{12}x_2 + \ldots + \gamma_{1n}x_n, \\ x_2 = \beta_2 + \gamma_{21}x_1 + 0x_2 + \ldots + \gamma_{2n}x_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ x_n = \beta_n + \gamma_{n1}x_1 + \gamma_{n2}x_2 + \ldots + 0x, \end{cases}$$

или в матричной форме

$$X = \beta + \Gamma X, \tag{1.18}$$

где $\beta = (\beta_{11}, \beta_{12}, \ldots, \beta_n)^T, \Gamma = (\gamma_{ij})_{n\times n}$. СЛАУ (1.18) равносильна СЛАУ (1.15). Организуем итерационный процесс по следующей схеме. Положим $X^{(0)} = \beta$. Далее

$$X^{(1)} = \beta + \Gamma X^{(0)},$$
$$X^{(2)} = \beta + \Gamma X^{(1)} = \beta + \Gamma(\beta + AX^{(0)}) = \beta + \Gamma\beta + \Gamma^2 X^{(0)},$$
$$\ldots\ldots\ldots\ldots\ldots\ldots \tag{1.19}$$

$$X^{(m)} = \beta + \Gamma X^{(m-1)} = \beta + \Gamma\beta + \Gamma^2\beta + \ldots + \Gamma^{m-1}\beta + \Gamma^m X^{(0)}.$$

В результате получим последовательность векторов $\left\{X^{(m)}\right\}_{m=0}^{\infty}$. Естественно поставить два вопроса:

1. Сходится ли последовательность $\left\{X^{(m)}\right\}$, т.е. существует ли вектор

$$X^* \in R^n : X^* = \lim_{m\to\infty} X^{(m)} \Leftrightarrow \lim_{m\to\infty} \rho_{R^n}\left(X^{(m)}, X^*\right) = 0\,?$$

2. Будет ли построенный вектор X^*(если $\exists \lim\limits_{m\to\infty} X^{(m)}$) решением СЛАУ (1.18), а значит СЛАУ (1.15)?

Прежде чем дать обоснованный ответ на эти вопросы, рассмотрим два примера - “хороший” и “плохой”.

<u>Пример 1.2</u>

$$\begin{cases} 5x_1 + x_2 - 2x_3 = 15, \\ -x_1 + 5x_2 - 2x_3 = 5, \\ x_1 + x_2 + 5x_3 = 10. \end{cases}$$

Решением данной СЛАУ является вектор $X^* = (3{,}2{,}1)^T$. Организуем вычисление векторов $X^{(m)}$по итерационному алгоритму (1.19).

1. Преобразование СЛАУ к виду (1.18)

$$\begin{cases} x_1 = \dfrac{15}{5} - \dfrac{1}{5}x_2 + \dfrac{2}{5}x_3, \\ x_2 = \dfrac{5}{5} + \dfrac{1}{5}x_1 + \dfrac{2}{5}x_3, \\ x_3 = \dfrac{10}{5} - \dfrac{1}{5}x_1 - \dfrac{1}{5}x_2 \end{cases}$$

или

$$X = \underbrace{\begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}}_{\beta} + \frac{1}{5}\underbrace{\begin{pmatrix} 0 & -1 & 2 \\ 1 & 0 & 2 \\ -1 & -1 & 0 \end{pmatrix}}_{A}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

2. Положим $X^{(0)} = \beta = (3;1;2)^T$.

Имеем далее в соответствии с (1.19)

$$X^{(1)} = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} + \frac{1}{5}\begin{pmatrix} 0 & -1 & 2 \\ 1 & 0 & 2 \\ -1 & -1 & 0 \end{pmatrix}\begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} + \frac{1}{5}\begin{pmatrix} 3 \\ 7 \\ -4 \end{pmatrix} = \begin{pmatrix} 18/5 \\ 12/5 \\ 6/5 \end{pmatrix} = \begin{pmatrix} 3{,}6 \\ 2{,}4 \\ 1{,}2 \end{pmatrix},$$

$$\rho\left(X^{(1)}, X^*\right) = \sqrt{(3{,}6-3)^2 + (2{,}4-2)^2 + (1{,}2-1)^2} \approx 0{,}748;$$

$$X^{(2)} = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} + \frac{6}{25}\begin{pmatrix} 0 & -1 & 2 \\ 1 & 0 & 2 \\ -1 & -1 & 0 \end{pmatrix}\begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} + \frac{6}{25}\begin{pmatrix} 0 \\ 5 \\ -5 \end{pmatrix} = \begin{pmatrix} 3 \\ 11/5 \\ 4/5 \end{pmatrix} = \begin{pmatrix} 3{,}0 \\ 2{,}2 \\ 0{,}8 \end{pmatrix},$$

$$\rho\left(X^{(2)}, X^*\right) = \sqrt{(3-3)^2 + (2{,}2-2)^2 + (0{,}8-1)^2} \approx 0{,}283;$$

$$X^{(3)} = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} + \frac{1}{25}\begin{pmatrix} 0 & -1 & 2 \\ 1 & 0 & 2 \\ -1 & -1 & 0 \end{pmatrix}\begin{pmatrix} 15 \\ 11 \\ 4 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} + \frac{1}{25}\begin{pmatrix} -3 \\ 23 \\ -26 \end{pmatrix} = \frac{1}{25}\begin{pmatrix} 72 \\ 48 \\ 24 \end{pmatrix} \approx \begin{pmatrix} 2{,}88 \\ 1{,}92 \\ 0{,}96 \end{pmatrix},$$

$$\rho\left(X^{(3)}, X^*\right) = \sqrt{(2{,}88-3)^2 + (1{,}92-2)^2 + (0{,}96-1)^2} \approx 0{,}150;$$

$$X^{(4)} = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} + \frac{24}{125}\begin{pmatrix} 0 & -1 & 2 \\ 1 & 0 & 2 \\ -1 & -1 & 0 \end{pmatrix}\begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} + \frac{24}{125}\begin{pmatrix} 0 \\ 5 \\ -5 \end{pmatrix} = \begin{pmatrix} 3 \\ 49/25 \\ -26/25 \end{pmatrix} = \begin{pmatrix} 3 \\ 1{,}96 \\ 1{,}04 \end{pmatrix},$$

$$\rho\left(X^{(4)}, X^*\right) = \sqrt{(3-3)^2 + (1{,}96-2)^2 + (1{,}04-1)^2} \approx 0{,}057;$$

$$X^{(5)} = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} + \frac{1}{125}\begin{pmatrix} 0 & -1 & 2 \\ 1 & 0 & 2 \\ -1 & -1 & 0 \end{pmatrix}\begin{pmatrix} 75 \\ 49 \\ 26 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} + \frac{1}{125}\begin{pmatrix} 3 \\ 127 \\ -124 \end{pmatrix} \approx \begin{pmatrix} 3{,}024 \\ 2{,}016 \\ 1{,}008 \end{pmatrix},$$

$$\rho\left(X^{(5)}, X^*\right) = \sqrt{(0{,}024)^2 + (0{,}016)^2 + (0{,}008)^2} \approx 0{,}030.$$

Имеем $\rho\left(X^{(1)}, X^*\right) > \rho\left(X^{(2)}, X^*\right) > \ldots > \rho\left(X^{(5)}, X^*\right)$.

Можно <u>предположить</u>, что последовательность $\left\{\rho\left(X^{(m)}, X^*\right)\right\}$ в данном примере является монотонно убывающей. Она ограничена снизу нулем и потому должен существовать $\lim\limits_{m\to\infty} X^{(m)} = X^*$. Естественно ожидать, что $X^* = (3;2;1)^T$.

<u>Пример 1.3</u>

$$\begin{cases} x_1 - x_2 + x_3 = 2, \\ 2x_1 + x_2 - x_3 = 1, \\ 2x_1 - x_2 + 4x_3 = 12. \end{cases}$$

Решением этой СЛАУ является вектор $X^* = (1;2;3)^T$ (проверьте!).

1) Преобразованная СЛАУ

$$\begin{cases} x_1 = 2 + x_2 - x_3, \\ x_2 = 1 - 2x_1 + x_3 x_3 = 3 - \dfrac{1}{2}x_1 + \dfrac{1}{4}x_2 \end{cases} \quad \text{или } X = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} + \begin{pmatrix} 0 & 1 & -1 \\ -2 & 0 & 1 \\ -\dfrac{1}{2} & \dfrac{1}{4} & 0 \end{pmatrix} X.$$

$$2)\ X^{(0)}=\begin{pmatrix}2\\1\\3\end{pmatrix},\quad X^{(1)}=\begin{pmatrix}2\\1\\3\end{pmatrix}+\begin{pmatrix}0&1&-1\\-2&0&1\\-\frac{1}{2}&\frac{1}{4}&0\end{pmatrix}\begin{pmatrix}2\\1\\3\end{pmatrix}=\begin{pmatrix}2\\1\\3\end{pmatrix}+\begin{pmatrix}-2\\-1\\-3/4\end{pmatrix}=\begin{pmatrix}0\\0\\9/4\end{pmatrix},$$

$$\rho\left(X^{(1)},X^*\right)=\sqrt{(1-0)^2+(2-0)^2+(3-9/4)^2}\approx 2{,}358;$$

$$X^{(2)}=\begin{pmatrix}2\\1\\3\end{pmatrix}+\begin{pmatrix}0&1&-1\\-2&0&1\\-\frac{1}{2}&\frac{1}{4}&0\end{pmatrix}\begin{pmatrix}0\\0\\9/4\end{pmatrix}=\begin{pmatrix}2\\1\\3\end{pmatrix}+\begin{pmatrix}-9/4\\9/4\\0\end{pmatrix}=\begin{pmatrix}-1/4\\13/4\\3\end{pmatrix}=\begin{pmatrix}-0{,}25\\3{,}25\\3\end{pmatrix},$$

$$\rho\left(X^{(2)},X^*\right)=\sqrt{\left(1+\frac{1}{4}\right)^2+\left(2-\frac{13}{4}\right)^2+(3-3)^2}\approx 1{,}767;$$

$$X^{(3)}=\begin{pmatrix}2\\1\\3\end{pmatrix}+\begin{pmatrix}0&1&-1\\-2&0&1\\-\frac{1}{2}&\frac{1}{4}&0\end{pmatrix}\begin{pmatrix}-1/4\\13/4\\3\end{pmatrix}=\begin{pmatrix}2\\1\\3\end{pmatrix}+\begin{pmatrix}1/4\\7/2\\15/16\end{pmatrix}=\begin{pmatrix}9/4\\9/2\\63/16\end{pmatrix}=\begin{pmatrix}2{,}25\\4{,}5\\3{,}94\end{pmatrix},$$

$$\rho\left(X^{(3)},X^*\right)=\sqrt{(1-2{,}25)^2+(4{,}5-2)^2+(3{,}94-3)^2}\approx 2{,}949;$$

$$X^{(4)}=\begin{pmatrix}2\\1\\3\end{pmatrix}+\begin{pmatrix}0&1&-1\\-2&0&1\\-\frac{1}{2}&\frac{1}{4}&0\end{pmatrix}\begin{pmatrix}9/4\\9/2\\63/16\end{pmatrix}=\begin{pmatrix}2\\1\\3\end{pmatrix}+\begin{pmatrix}9/16\\-9/16\\0\end{pmatrix}=\begin{pmatrix}41/16\\7/16\\3\end{pmatrix}=\begin{pmatrix}2{,}56\\0{,}44\\3{,}0\end{pmatrix},$$

$$\rho\left(X^{(4)},X^*\right)=\sqrt{(1-2{,}56)^2+(2-0{,}44)^2+(3-3)^2}\approx 2{,}206;$$

$$X^{(5)}=\begin{pmatrix}2\\1\\3\end{pmatrix}+\begin{pmatrix}0&1&-1\\-2&0&1\\-\frac{1}{2}&\frac{1}{4}&0\end{pmatrix}\begin{pmatrix}41/16\\7/16\\3\end{pmatrix}=\begin{pmatrix}2\\1\\3\end{pmatrix}+\begin{pmatrix}-41/16\\-34/16\\-27/64\end{pmatrix}=\begin{pmatrix}-9/16\\-9/8\\165/64\end{pmatrix}\approx\begin{pmatrix}-0{,}56\\-1{,}125\\2{,}58\end{pmatrix},$$

$$\rho\left(X^{(5)},X^*\right)\approx 1{,}837.$$

Говорить о сходимости данного итерационного процесса, по крайней мере по приведенным первым шагам, сложно. Вместе с тем отсутствие монотонного убывания последовательности $\rho\left(x^{(m)},x^*\right)$ является основанием для предположения о расходимости итерационного процесса.

Приведенные примеры показывают, что вопросы, поставленные в п.1.2.1, действительно требуют разрешения. Используя результаты общей теории СЛАУ, можно предположить, что ответ на них связан со свойствами основной матрицы **A** СЛАУ (1.16) или матрицы **Г** приведенной СЛАУ (1.17). Решение всех вопросов, связанных со сходимостью

итерационных методов решения СЛАУ, рассматривается в следующем п.1.2.2 с самых общих позиций.

1.2.2. Принцип сжимающих отображений

Пусть $\mathbf{X}$ – ***полное метрическое пространство*** с метрикой $\rho_{\mathbf{X}}(\cdot,\cdot)$, $\tilde{\mathbf{A}}$ – некоторый оператор, действующий в этом пространстве $\left(\tilde{\mathbf{A}}:\mathbf{X}\to\mathbf{X}\right)$. Например, $\mathbf{X}=\mathbf{R}^{\mathbf{n}}$, $\rho_{\mathbf{R}^n}(\bar{\mathbf{x}},\bar{\mathbf{y}})=\sqrt{(\mathbf{x}_1-\mathbf{y}_1)^2+\ldots+(\mathbf{x}_n-\mathbf{y}_n)^2}$, $\tilde{\mathbf{A}}\bar{\mathbf{x}}=\beta+\Gamma\mathbf{X}$, $\beta=(\beta_1,\ldots,\beta_n)^{\mathbf{T}}$, $\Gamma=(\gamma_{ij})_{n\times n}$, $\mathbf{X}=(\mathbf{x}_1,\ldots,\mathbf{x}_n)^{\mathbf{T}}$, $\bar{\mathbf{x}}=(\mathbf{x}_1,\ldots,\mathbf{x}_n)$, $\tilde{\mathbf{A}}:\mathbf{R}^{\mathbf{n}}\to\mathbf{R}^{\mathbf{n}}$.

Определим условия, при которых операторное уравнение

$$\tilde{\mathbf{A}}\bar{\mathbf{x}}=\bar{\mathbf{x}}\qquad(1.20)$$

имеет решение $\bar{\mathbf{x}}^*$ (в нашем случае это уравнение $\beta+\Gamma\mathbf{X}=\mathbf{X}$).

Одновременно построим алгоритм нахождения приближенного решения $\bar{\mathbf{x}}^{(m)}$ с заданной степенью точности:

$$\rho\left(\bar{\mathbf{x}}^{(m)},\bar{\mathbf{x}}^*\right)<\varepsilon.$$

<u>Определение.</u> Точка $\overline{\mathbf{x}^*}$ называется <u>неподвижной</u> точкой оператора $\tilde{\mathbf{A}}$, действующего в <u>полном</u> пространстве $\mathbf{X}$, если

$$\tilde{\mathbf{A}}\bar{\mathbf{x}}^*=\bar{\mathbf{x}}^*.\qquad(1.21)$$

<u>Определение.</u> Оператор $\tilde{\mathbf{A}}:\mathbf{X}\to\mathbf{X}$ называется <u>оператором сжатия</u> или сжимающим оператором, если существует $\alpha\in\mathbf{R}_+:\mathbf{0}<\alpha<\mathbf{1}$ и

$$\forall\bar{\mathbf{x}},\bar{\mathbf{y}}\in\mathbf{X}\Rightarrow\rho_{\mathbf{X}}\left(\tilde{\mathbf{A}}\bar{\mathbf{x}},\tilde{\mathbf{A}}\bar{\mathbf{y}}\right)\le\alpha\rho_{\mathbf{X}}\left(\bar{\mathbf{x}},\bar{\mathbf{y}}\right).\qquad(1.22)$$

Неравенство (1.22) означает, что при сжимающем отображении расстояние между образами любых двух элементов пространства $\mathbf{X}$ уменьшается по сравнению с расстоянием между их прообразами.

<u>Теорема 1.2 (С. Банаха о неподвижной точке).</u> Сжимающий оператор $\tilde{\mathbf{A}}$, действующий в полном метрическом пространстве $\mathbf{X}$, имеет единственную неподвижную точку $\bar{\mathbf{x}}^*\in\mathbf{X}$, т.е. уравнение $\tilde{\mathbf{A}}\bar{\mathbf{x}}=\bar{\mathbf{x}}$ имеет единственное решение.

<u>Доказательство.</u> Возьмем произвольную точку $\bar{\mathbf{x}}^{\mathbf{0}}\in\mathbf{X}$ и построим последовательность точек $\left\{\bar{\mathbf{x}}_{\mathbf{k}}\right\}$ по правилу

$$\bar{\mathbf{x}}_{\mathbf{k}}=\tilde{\mathbf{A}}\bar{\mathbf{x}}_{\mathbf{k-1}}\qquad(1.23)$$

1. Покажем, что такая последовательность точек $\left\{\bar{\mathbf{x}}_{\mathbf{k}}\right\}$ является сходящейся (при условии, что $\tilde{\mathbf{A}}$ - оператор сжатия).

Поскольку пространство $\mathbf{X}$ полное, то достаточно доказать, что последовательность $\{\bar{\mathbf{x}}_\mathbf{k}\}$ является фундаментальной, т.е.

$$\forall\varepsilon > 0 \exists \mathbf{K}(\varepsilon): \forall \mathbf{k} > \mathbf{K}(\varepsilon) \wedge \forall \mathbf{m} \in \mathbf{N} \Rightarrow \rho(\bar{\mathbf{x}}_\mathbf{k}, \bar{\mathbf{x}}_\mathbf{k+m}) < \varepsilon. \qquad (1.24)$$

Имеем

$$\rho(\bar{\mathbf{x}}_\mathbf{k}, \bar{\mathbf{x}}_\mathbf{k+m}) \le \Big| \text{ в силу неравенства треугольника } \Big| \le$$

$$\le \rho(\bar{\mathbf{x}}_\mathbf{k}, \bar{\mathbf{x}}_\mathbf{k+m}) + \rho(\bar{\mathbf{x}}_\mathbf{k+1}, \bar{\mathbf{x}}_\mathbf{k+m}) \le \Big| \text{ аналогично } \Big| \le$$

$$\begin{aligned} &\le \rho(\bar{\mathbf{x}}_\mathbf{k}, \bar{\mathbf{x}}_\mathbf{k+1}) + \rho(\bar{\mathbf{x}}_\mathbf{k+1}, \bar{\mathbf{x}}_\mathbf{k+2}) + \rho(\bar{\mathbf{x}}_\mathbf{k+2}, \bar{\mathbf{x}}_\mathbf{k+m}) \le \ldots \le \\ &\le \rho(\bar{\mathbf{x}}_\mathbf{k}, \bar{\mathbf{x}}_\mathbf{k+1}) + \rho(\bar{\mathbf{x}}_\mathbf{k+1}, \bar{\mathbf{x}}_\mathbf{k+2}) + \ldots + \rho(\bar{\mathbf{x}}_\mathbf{k+m-1}, \bar{\mathbf{x}}_\mathbf{k+m}). \end{aligned} \qquad (1.25)$$

Поскольку $\bar{\mathbf{x}}_\mathbf{k} = \tilde{\mathbf{A}}\bar{\mathbf{x}}_\mathbf{k-1}$ (1.23) и $\tilde{\mathbf{A}}$- оператор сжатия, т.е. $\forall \bar{\mathbf{x}}_\mathbf{k+s}$ и $\bar{\mathbf{x}}_\mathbf{k+s+1}, \mathbf{s} \in \mathbf{N},$ выполняется условие (1.22) ,то

$$\begin{aligned} \rho(\bar{\mathbf{x}}_\mathbf{k+s}, \bar{\mathbf{x}}_\mathbf{k+s+1}) &= \rho(\tilde{\mathbf{A}}\mathbf{x}_\mathbf{k+s-1}, \tilde{\mathbf{A}}\mathbf{x}_\mathbf{k+s}) \le \alpha\rho(\bar{\mathbf{x}}_\mathbf{k+s-1}, \bar{\mathbf{x}}_\mathbf{k+s}) = \\ &= \alpha\rho(\tilde{\mathbf{A}}\mathbf{x}_\mathbf{k+s-2}, \tilde{\mathbf{A}}\mathbf{x}_\mathbf{k+s-1}) \le \alpha^2\rho(\bar{\mathbf{x}}_\mathbf{k+s-2}, \bar{\mathbf{x}}_\mathbf{k+s-1}) \le \ldots \le \\ &\le \alpha^{\mathbf{k+s}}\rho\left(\bar{\mathbf{x}}^\mathbf{0}, \tilde{\mathbf{A}}\mathbf{x}_\mathbf{0}\right). \end{aligned} \qquad (1.26)$$

Используя (1.26), преобразуем оценку $\rho(\bar{\mathbf{x}}_\mathbf{k}, \bar{\mathbf{x}}_\mathbf{k+m})$ в (1.25)

$$\rho(\bar{\mathbf{x}}_\mathbf{k}, \bar{\mathbf{x}}_\mathbf{k+m}) \le \alpha^\mathbf{k}\rho(\bar{\mathbf{x}}^\mathbf{0}, \tilde{\mathbf{A}}\bar{\mathbf{x}}^\mathbf{0}) + \alpha^\mathbf{k+1}\rho(\bar{\mathbf{x}}^\mathbf{0}, \tilde{\mathbf{A}}\bar{\mathbf{x}}^\mathbf{0}) + \ldots + \alpha^\mathbf{k+m-1}\rho(\bar{\mathbf{x}}^\mathbf{0}, \tilde{\mathbf{A}}\bar{\mathbf{x}}^\mathbf{0}) =$$

$$= \alpha^\mathbf{k}(\mathbf{1} + \alpha + \ldots + \alpha^\mathbf{m-1})\rho(\bar{\mathbf{x}}^\mathbf{0}, \tilde{\mathbf{A}}\bar{\mathbf{x}}^\mathbf{0}) = \frac{\alpha^\mathbf{k}(\mathbf{1} - \alpha^\mathbf{m})}{\mathbf{1} - \alpha}\rho(\bar{\mathbf{x}}^\mathbf{0}, \tilde{\mathbf{A}}\bar{\mathbf{x}}^\mathbf{0}).$$

Поскольку $\mathbf{0} < \alpha < \mathbf{1}$ и $\mathbf{1} - \alpha^\mathbf{m} < \mathbf{1}$, то получаем оценку

$$\rho(\bar{\mathbf{x}}_\mathbf{k}, \bar{\mathbf{x}}_\mathbf{k+m}) < \frac{\alpha^\mathbf{k}}{\mathbf{1} - \alpha}\rho(\bar{\mathbf{x}}^\mathbf{0}, \tilde{\mathbf{A}}\bar{\mathbf{x}}^\mathbf{0}). \qquad (1.27)$$

Из (1.27) с учетом $\mathbf{0} < \alpha < \mathbf{1}$, следует, что за счет выбора $\mathbf{k}$ можно всегда добиться выполнения условия (1.24) фундаментальности последовательности $\{\bar{\mathbf{x}}_\mathbf{k}\}$.

Действительно, пусть $\rho(\bar{\mathbf{x}}_\mathbf{k}, \bar{\mathbf{x}}_\mathbf{k+m}) < \dfrac{\alpha^\mathbf{k}}{\mathbf{1} - \alpha}\rho(\bar{\mathbf{x}}^\mathbf{0}, \tilde{\mathbf{A}}\bar{\mathbf{x}}^\mathbf{0}) < \varepsilon$,

откуда $\alpha^\mathbf{k} < \dfrac{(\mathbf{1} - \alpha)\varepsilon}{\rho(\bar{\mathbf{x}}^\mathbf{0}, \tilde{\mathbf{A}}\bar{\mathbf{x}}^\mathbf{0})} \Rightarrow \mathbf{k} > \log_\alpha \dfrac{(\mathbf{1} - \alpha)\varepsilon}{\rho(\bar{\mathbf{x}}^\mathbf{0}, \tilde{\mathbf{A}}\bar{\mathbf{x}}^\mathbf{0})}$ и $\mathbf{K}(\varepsilon) = \left[\log_\alpha \dfrac{(\mathbf{1} - \alpha)\varepsilon}{\rho(\bar{\mathbf{x}}^\mathbf{0}, \tilde{\mathbf{A}}\bar{\mathbf{x}}^\mathbf{0})}\right]$.

Таким образом, последовательность $\bar{\mathbf{x}}_\mathbf{k}$ является <u>фундаментальной в полном</u> пространстве и потому сходится к некоторому элементу $\bar{\mathbf{x}}^* \in \mathbf{X}$.

2. Покажем, что $\bar{x}^*$ - неподвижная точка оператора $\tilde{\mathbf{A}}$, т.е. $\tilde{\mathbf{A}}\bar{\mathbf{x}}^* = \bar{\mathbf{x}}^*$. Используя неравенство треугольника, получаем оценку:

$$\begin{aligned} \rho(\bar{\mathbf{x}}^*, \tilde{\mathbf{A}}\bar{\mathbf{x}}^*) &\le \rho(\bar{\mathbf{x}}^*, \bar{\mathbf{x}}_\mathbf{k}) + \rho(\bar{\mathbf{x}}_\mathbf{k}, \tilde{\mathbf{A}}\bar{\mathbf{x}}^*) = \rho(\bar{\mathbf{x}}^*, \bar{\mathbf{x}}_\mathbf{k}) + \rho(\tilde{\mathbf{A}}\bar{\mathbf{x}}_\mathbf{k-1}, \mathbf{A}\bar{\mathbf{x}}^*) \le \\ &\le \rho(\bar{\mathbf{x}}^*, \bar{\mathbf{x}}_\mathbf{k}) + \alpha\rho(\bar{\mathbf{x}}^*, \bar{\mathbf{x}}_\mathbf{k-1}). \end{aligned}$$

Так как $\lim\limits_{\mathbf{k}\to\infty} \rho(\bar{\mathbf{x}}_\mathbf{k}, \bar{\mathbf{x}}^*) = \mathbf{0}$, то $\forall\varepsilon > \mathbf{0}, \exists\mathbf{K}(\varepsilon): \forall\mathbf{k} > \mathbf{K}(\varepsilon) \Rightarrow \rho(\bar{\mathbf{x}}_\mathbf{k}, \bar{\mathbf{x}}^*) < \varepsilon.$ Имеем окончательно $\rho(\bar{\mathbf{x}}^*, \tilde{\mathbf{A}}\bar{\mathbf{x}}^*) < \varepsilon + \alpha\varepsilon = \varepsilon(\mathbf{1} + \alpha) = \tilde{\varepsilon}$. Это неравенство должно выполняться для любого сколь угодно малого $\varepsilon > \mathbf{0}$, поэтому $\rho(\bar{\mathbf{x}}^*, \mathbf{A}\bar{\mathbf{x}}^*) = \mathbf{0}$ и $\tilde{\mathbf{A}}\bar{\mathbf{x}}^* = \bar{\mathbf{x}}^*$.

3. Единственность неподвижной точки $\bar{\mathbf{x}}^*$ легко доказывается методом от противного. Пусть $\mathbf{y}^* \in \mathbf{X} \wedge \tilde{\mathbf{A}}\bar{\mathbf{y}}^* = \bar{\mathbf{y}}^*$. Тогда

$\rho(\tilde{\mathbf{A}}\bar{\mathbf{x}}^*, \tilde{\mathbf{A}}\bar{\mathbf{y}}^*) = \rho(\bar{\mathbf{x}}^*, \bar{\mathbf{y}}^*)$. С другой стороны, $\rho(\tilde{\mathbf{A}}\bar{\mathbf{x}}^*, \tilde{\mathbf{A}}\bar{\mathbf{y}}^*) \le \alpha\rho(\bar{\mathbf{x}}^*, \bar{\mathbf{y}})$, где $\mathbf{0} < \alpha < \mathbf{1}$. Получаем соотношение
$\rho(\bar{\mathbf{x}}^*, \bar{\mathbf{y}}^*) = \rho(\tilde{\mathbf{A}}\bar{\mathbf{x}}^*, \tilde{\mathbf{A}}\bar{\mathbf{y}}^*) \le \alpha\rho(\bar{\mathbf{x}}^*, \mathbf{y}^*) \Rightarrow (\mathbf{1} - \alpha)\rho(\bar{\mathbf{x}}^*, \bar{\mathbf{y}}^*) \le \mathbf{0}$, а это неравенство неверное. ⊗

<u>Замечание 1</u>. Из неравенства (1.27) $\rho(\bar{\mathbf{x}}_\mathbf{k}, \bar{\mathbf{x}}_\mathbf{k+m}) < \dfrac{\alpha^\mathbf{k}}{\mathbf{1} - \alpha}\rho(\bar{\mathbf{x}}^\mathbf{0}, \tilde{\mathbf{A}}\bar{\mathbf{x}}^\mathbf{0})$ можно получить оценку $\mathbf{k}$-го приближения $\bar{\mathbf{x}}_\mathbf{k}$ к $\bar{\mathbf{x}}^*$. Переходя к пределу в последнем неравенстве при $\mathbf{m} \to +\infty$, получаем

$$\rho\left(\bar{\mathbf{x}}_\mathbf{k}, \bar{\mathbf{x}}^*\right) < \frac{\alpha^\mathbf{k}}{\mathbf{1} - \alpha}\rho(\bar{\mathbf{x}}^\mathbf{0}, \tilde{\mathbf{A}}\bar{\mathbf{x}}^\mathbf{0}). \qquad (1.28)$$

<u>Замечание 2.</u> Итерационная последовательность $\{\bar{\mathbf{x}}_\mathbf{k} = \tilde{\mathbf{A}}\bar{\mathbf{x}}_\mathbf{k-1}\}$ сходится к $\bar{\mathbf{x}}^*$ из любой начальной точки $\bar{\mathbf{x}}^\mathbf{0} \in \mathbf{X}$. Выбор точки $\bar{\mathbf{x}}^\mathbf{0}$ может повлиять лишь на скорость сходимости. Поэтому ошибка, допущенная в вычислениях на некоторой итерации, может интерпретироваться как выбор новой начальной точки $\bar{\mathbf{x}}^\mathbf{0} \in \mathbf{X}$.

1.2.3. Применение принципа сжимающих отображений к СЛАУ

Пусть

$$\mathbf{AX} = \mathbf{B} \qquad (1.29)$$

- СЛАУ порядка $\mathbf{n}$,

$$\mathbf{X} = \beta + \Gamma\mathbf{X} \qquad (1.30)$$

- СЛАУ, полученная из (1.29) в результате преобразования (1.17). Введем оператор $\tilde{\mathbf{A}} = \beta + \Gamma(\cdot)$, т.е.

$$\tilde{\mathbf{A}}\mathbf{X} = \beta + \Gamma\mathbf{X}. \qquad (1.31)$$

С учетом (1.31) СЛАУ (1.30) можно записать в операторной форме

$$\tilde{\mathbf{A}}\mathbf{X} = \mathbf{X}. \qquad (1.32)$$

В соответствии с доказанной теоремой операторное уравнение будет иметь единственное решение, если оператор $\tilde{\mathbf{A}}$ сжимающий, т.е.

$$\underset{\mathbf{0}<\alpha<\mathbf{1}}{\exists\alpha \in \mathbf{R}_+}: \rho_{\mathbf{R}^\mathbf{n}}(\tilde{\mathbf{A}}\mathbf{X}, \tilde{\mathbf{A}}\mathbf{Y}) \le \alpha \cdot \rho_{\mathbf{R}^\mathbf{n}}(\mathbf{X}, \mathbf{Y}), \forall\mathbf{X}, \mathbf{Y} \in \mathbf{R}^\mathbf{n}. \qquad (1.33)$$

Поскольку

$$\rho_{R^n}\left(\tilde{A}X,\tilde{A}Y\right)=\rho_{R^n}(\beta+\Gamma X,\beta+\Gamma Y)=\|(\beta+\Gamma X)-(\beta+\Gamma Y)\|_{R^n}= \\ =\|\Gamma(X-Y)\|_{R^n}\le\|\Gamma\|\cdot\|X-Y\|_{R^n}=\|\Gamma\|\cdot\rho_{R^n}(X,Y), \qquad (1.34)$$

из сопоставления (1.33) и (1.34) заключаем, что достаточным условием существования решения операторного уравнения (1.32), отвечающего СЛАУ (1.30), является условие

$$0<\|\Gamma\|<1, \qquad (1.35)$$

т.е. норма $\|\Gamma\|$ матрицы Γ преобразованной СЛАУ (1.30) должна быть меньше единицы. Заметим, что нарушение условия (1.35) может приводить к расходимости итерационного процесса в методе последовательных приближений. Однако это еще не означает, что СЛАУ (1.29) и равносильная ей (1.30) являются несовместными (не имеют решений) или совместными, но неопределенными (имеют бесконечно много решений). Условие (1.35) является <u>достаточным</u> для существования <u>единственного</u> решения СЛАУ (1.29) и, в частности, гарантирует его нахождение с помощью итерационного процесса $X_k=\tilde{A}X_{k-1}$, $k=1,2,...$, с любой степенью точности.

В качестве нормы матрицы A используются нормы:

$$\text{норма 1: } \|\Gamma\|_1=\max_i\sum_{j=1}^n|\gamma_{ij}|,$$

$$\text{норма 2: } \|\Gamma\|_2=\max_j\sum_{i=1}^n|\gamma_{ij}|, \qquad (1.36)$$

$$\text{норма 3: } \|\Gamma\|_3=\sqrt{\sum_{i=1}^n\sum_{j=1}^n\gamma_{ij}^2}.$$

С учетом (1.36) условие (1.35) можно конкретизировать:

$$0<\max_i\sum_{j=1}^n|\gamma_{ij}|<1, \text{ или } 0<\max_j\sum_{i=1}^n|\gamma_{ij}|<1, \text{ или } 0<\sqrt{\sum_{i=1}^n\sum_{j=1}^n\gamma_{ij}^2}<1. \qquad (1.37)$$

Выполним анализ рассмотренных в п. 1.2 примеров с учетом полученных результатов, а именно выполнимости достаточных условий (1.37) сходимости итерационного процесса.

<u>Пример 1.2</u>

$$\Gamma=\begin{pmatrix}0 & -1/5 & 2/5\\ 1/5 & 0 & 2/5\\ -1/5 & -1/5 & 0\end{pmatrix},$$

$$\|\Gamma\|_1=\max\left\{\left|-\frac{1}{5}\right|+\frac{2}{5};\frac{1}{5}+\frac{2}{5},\left|-\frac{1}{5}\right|+\left|-\frac{1}{5}\right|\right\}=\frac{3}{5}<1,$$

$$\|\Gamma\|_2=\max\left\{\frac{2}{5},\frac{2}{5},\frac{4}{5}\right\}=\frac{4}{5}<1,$$

$$\|\Gamma\|_3=\sqrt{\frac{1}{25}+\frac{4}{25}+\frac{1}{25}+\frac{4}{25}+\frac{1}{25}+\frac{1}{25}}=\frac{2}{5}\sqrt{3}<1.$$

Таким образом, в примере 1.2 достаточные условия сходимости итерационного процесса выполнены. Заметим, что достаточно были вычислить лишь одну из трех матричных норм.

<u>Пример 1.3</u>

$$\Gamma=\begin{pmatrix}0 & 1 & -1\\ -2 & 0 & 1\\ -\frac{1}{2} & \frac{1}{4} & 0\end{pmatrix},$$

$$\|\Gamma\|_1=\max\left\{2,3,\frac{3}{4}\right\}=3>1,\ \|\Gamma\|_2=\max\left\{\frac{5}{2},\frac{5}{4},2\right\}=2>1,$$

$$\|\Gamma\|_3=\sqrt{1+1+4+1+\frac{1}{4}+\frac{1}{16}}>1.$$

В примере 1.3 достаточные условия не выполняются.

В соответствии с теоремой о неподвижной точке итерационный процесс $X^{(m)}=\beta+\Gamma X^{(m-1)}$ в примере 1.2 сходится к решению исходной СЛАУ, а в примере 1.3 он может расходиться, так как оператор $\tilde{A}$ в этом примере не является сжимающим.

<u>Замечание.</u> Приведенная СЛАУ $X=\beta+\Gamma X$ получается из СЛАУ $AX=B$ умножением на коэффициент $1/a_{ii}$ и последующим выражением x_i через остальные переменные. Но тогда для нормы 1 имеем

$$\max_i\sum_{j=1}^n|\alpha_{ij}|=\max_{i=1,n}\left\{\left|\frac{a_{i1}}{a_{ii}}\right|+\left|\frac{a_{i2}}{a_{ii}}\right|+...+\left|\frac{a_{i,i-1}}{a_{ii}}\right|+\left|\frac{a_{i,i+1}}{a_{ii}}\right|+...+\left|\frac{a_{in}}{a_{ii}}\right|\right\}<1.$$

Приведенное условие эквивалентно следующему: $\forall i=1,2,...,n$ $|a_{i1}|+|a_{i2}|+...+|a_{i,i-1}|+|a_{i,i+1}|+...+|a_{in}|<|a_{ii}|$, т.е. диагональный элемент a_{ii} каждой строки матрицы A должен быть больше суммы модулей остальных элементов этой строки (уже не преобразованной матрицы Γ, а исходной A). Говорят, что матрица A должна иметь диагональное преобладание.

Пример 1.2

$$A=\begin{pmatrix}5 & 1 & -2\\ -1 & 5 & -2\\ 1 & 1 & 5\end{pmatrix},$$

$$|a_{11}|=5>|a_{12}|+|a_{13}|=3,\quad |a_{22}|=5>|a_{12}|+|a_{23}|=3,$$

$$|a_{33}|=5>|a_{13}|+|a_{23}|=2.$$

Здесь матрица **A** имеет диагональное преобладание.

Пример 1.3

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

$$|a_{11}|=1<|-1|+1=2,\quad |a_{22}|=2<4+|-2|=6,$$

$$|a_{33}|=4>2+|-1|=3.$$

Данная матрица не имеет диагонального преобладания.

1.2.4.Определение затрат памяти ЭВМ и числа вычислительных операций

Затраты памяти ЭВМ:

-n^2 ячеек для хранения матрицы **A** и в ней же матрицы $\Gamma=(\gamma_{ij})_{n\times n}$;

-n ячеек для хранения вектора **B** и в нем же вектора β;

-n ячеек для вектора **X**.

Общий объем оперативной памяти ЭВМ, необходимой для реализации метода итераций, равен n^2+2n ячеек.

Число вычислительных операций

Если исключить затраты на преобразование СЛАУ (1.15) к виду (1.18), то получим чистые затраты времени на реализацию метода итераций. Будем иметь:

- n^2 операций умножения (умножение матрицы Γ на вектор **X**) для получения одного приближения $X^{(m)}$;

-$n(n-1)+n=n^2$ операций сложения для получения оценки $X^{(m)}$ вектора **X** на m-й итерации.

Если общее число итераций будет равно k, то число операций умножения и сложения увеличивается в k раз, т.е. суммарное число операций в методе итераций будет равно $N(*)=N(+)=kn^2$.

1.2.5. Сравнение метода итераций с другими методами

Затраты памяти ЭВМ в методе итераций такие же, как и в методе Гаусса, и в два раза больше (для СЛАУ с симметричной основной матрицей **A**), чем в методе квадратных корней.

Напомним, что в методе Гаусса число медленных операций равно $\frac{n^3}{3}+0(n^2)$, в методе квадратного корня - $\frac{n^3}{6}+0(n^2)$. Число медленных операций в методе итераций (kn^2) зависит от скорости сходимости итерационного процесса (числа итераций k). Если $k<<n$, то метод итераций существенно превосходит и метод Гаусса, и метод квадратных корней. Если же k имеет тот же порядок, что и n, то затраты времени на решение СЛАУ $AX=B$ методом итераций и методом Гаусса будут сопоставимы по величине. Так, если $n=10^4$, $k=50$, то число медленных операций в методе Гаусса будет равно $10^{12}/3$, а в методе итераций $5\cdot10^9$ медленных операций или в 67 раз меньше. Такие СЛАУ - 10000 уравнений с 10000 неизвестных - возникают, например, при решении задач восстановления спектров. Результаты измерения спектров с помощью оже-спектрометра или масс-спектрометра описываются интегральным уравнением Фредгольма первого рода. Решение этой некорректной задачи методом регуляризации с построением дискретного аналога уравнения Эйлера (1.13) приводит к СЛАУ (1.14). При восстановлении двумерных изображений порядок СЛАУ может быть больше 10^4 $(n>10^4)$. При $n\sim100$ и меньше преимущества метода итераций в быстродействии могут теряться.

1.2.6.Модификации метода итераций

Одной из наиболее известных модификаций является метод Зейделя. Смысл модификации заключается в том, что при вычислении компонент $x_i^{(k)}$ вектора $X^{(k)}$ на k-й итерации сразу же учитываются предыдущие оценки компонент $x_1^{(k)},x_2^{(k)},\dots,x_{i-1}^{(k)}$ - вектора $X^{(k)}$,вычисленные на предыдущих шагах итерационного процесса. Если $X^{(k)}=\beta+\Gamma X^{(k-1)}$, то в методе Зейделя

$$\begin{cases} x_1^{(k)} = \beta_1 + \sum\limits_{j=1}^{n} \gamma_{1j} x_j^{(k-1)}, \\ x_2^{(k)} = \beta_2 + \gamma_{21} x_1^{(k)} + \sum\limits_{j=2}^{n} \gamma_{2j} x_j^{(k-1)}, \\ x_3^{(k)} = \beta_3 + \gamma_{31} x_1^{(k)} + \gamma_{32} x_2^{(k)} + \sum\limits_{j=3}^{n} \gamma_{3j} x_j^{(k-1)}, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ x_n^{(k)} = \beta_n + \sum\limits_{j=1}^{n-1} \gamma_{nj} x_j^{(k)} + \gamma_{nn} x_n^{(k-1)}. \end{cases} \tag{1.38}$$

Во многих случаях метод Зейделя сходится быстрее метода итераций и даже может сходиться в тех случаях, когда метод итераций расходится. Однако есть примеры и обратного: метод Зейделя расходится, а метод итераций сходится.

Пример 1.4. Решить СЛАУ

$$\begin{cases} 5x_1 + x_2 - 2x_3 = 15, \\ -x_1 + 5x_2 - 2x_3 = 5 \\ x_1 + x_2 + 5x_3 = 10. \end{cases}$$

методом Зейделя.

Приведенная СЛАУ. $X = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} + \frac{1}{5} \begin{pmatrix} 0 & -1 & 2 \\ 1 & 0 & 2 \\ -1 & -1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$.

Точное решение $X^* = (3,2,1)^T$. Пусть $X^0 = (3,1,2)^T$. В соответствии с алгоритмом метода Зейделя (1.38) имеем далее

$$x_1^{(1)} = 3 + \frac{1}{5}(0 \quad -1 \quad 2)\begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} = \frac{18}{5} = 3{,}60,$$

$$x_2^{(1)} = 1 + \frac{1}{5}(1 \quad 0 \quad 2)\begin{pmatrix} 18/5 \\ 1 \\ 2 \end{pmatrix} = \frac{63}{25} = 2{,}520,$$

$$x_3^{(1)} = 2 + \frac{1}{5}(-1 \quad -1 \quad 0)\begin{pmatrix} 18/5 \\ 63/25 \\ 2 \end{pmatrix} = \frac{97}{125} = 0{,}776,$$

$$\rho\left(X^{(1)}, X^*\right) = \sqrt{(3{,}6-3)^2 + (2{,}52-2)^2 + (0{,}776-1)^2} = 0{,}825,$$

$$x_1^{(2)} = 3 + \frac{1}{5}(0 \quad -1 \quad 2)\begin{pmatrix} 3{,}600 \\ 2{,}520 \\ 0{,}776 \end{pmatrix} = 2{,}8064,$$

$$x_2^{(2)} = 1 + \frac{1}{5}(1 \quad 0 \quad 2)\begin{pmatrix} 2{,}8064 \\ 2{,}5200 \\ 0{,}7760 \end{pmatrix} = 1{,}8717,$$

$$x_3^{(2)} = 2 + \frac{1}{5}(-1 \quad -1 \quad 0)\begin{pmatrix} 2{,}8064 \\ 1{,}8717 \\ 0{,}7760 \end{pmatrix} = 1{,}0644,$$

$$\rho\left(X^{(2)}, X^*\right) = \sqrt{(2{,}8064-3)^2 + (1{,}8717-2)^2 + (1{,}0644-1)^2} = 0{,}241;$$

$$x_1^{(3)} = 3 + \frac{1}{5}(0 \quad -1 \quad 2)\begin{pmatrix} 2{,}8064 \\ 1{,}8717 \\ 1{,}0644 \end{pmatrix} = 3{,}0514,$$

$$x_2^{(3)} = 1 + \frac{1}{5}(1 \quad 0 \quad 2)\begin{pmatrix} 3{,}0514 \\ 1{,}8717 \\ 1{,}0644 \end{pmatrix} = 2{,}0360,$$

$$x_3^{(3)} = 2 + \frac{1}{5}(-1 \quad -1 \quad 0)\begin{pmatrix} 3{,}0514 \\ 2{,}0360 \\ 1{,}0644 \end{pmatrix} = 0{,}9825,$$

$$\rho\left(X^{(3)}, X^*\right) = \sqrt{(3{,}0514-3{,}0)^2 + (2{,}036-2{,}0)^2 + (0{,}9825-1)^2} = 0{,}065;$$

$$x_1^{(4)} = 3 + \frac{1}{5}(0 \quad -1 \quad 2)\begin{pmatrix} 3{,}0514 \\ 2{,}0360 \\ 0{,}9825 \end{pmatrix} = 2{,}9858,$$

$$x_2^{(4)} = 1 + \frac{1}{5}(1 \quad 0 \quad 2)\begin{pmatrix} 2{,}9858 \\ 2{,}0360 \\ 0{,}9825 \end{pmatrix} = 1{,}9902,$$

$$x_3^{(4)} = 2 + \frac{1}{5}(-1 \quad -1 \quad 0)\begin{pmatrix} 2{,}9858 \\ 1{,}9902 \\ 0{,}9825 \end{pmatrix} = 1{,}0048,$$

$$\rho\left(X^{(4)}, X^*\right) = \sqrt{(2{,}9858-3{,}0)^2 + (1{,}9902-2{,}0)^2 + (1{,}0048-1)^2} = 0{,}018.$$

Решение этой же СЛАУ методом итераций дало следующие результаты (пример 1.2):

$$\rho\left(X^{(1)},X^*\right)\approx 0{,}748;\ \rho\left(X^{(2)},X^*\right)\approx 0{,}283;\ \rho\left(X^{(3)},X^*\right)\approx 0{,}150;$$
$$\rho\left(X^{(4)},X^*\right)\approx 0{,}057;\ \rho\left(X^{(5)},X^*\right)\approx 0{,}030.$$

Как видим, решение методом Зейделя сходится к X^* быстрее, чем методом простой итерации.

1.3. Нормальные СЛАУ

Определение. СЛАУ $AX = B$ и соответствующая матрица A называются нормальными, если

1) $A^T = A$, т.е. матрица A симметричная;

2) матрица A определена положительно, т.е. $\forall X \in R^n$, $X \neq 0 \Rightarrow X^T AX > 0$.

Имеет место следующая теорема.

Теорема. Если обе части СЛАУ $AX = B$ с невырожденной матрицей A умножить слева на транспонированную матрицу A^T, то полученная СЛАУ

$$\left(A^T A\right)X = A^T B \tag{1.39}$$

будет нормальной.

Доказательство. 1. Докажем, что матрица $A^T A$ симметричная, т.е. $\left(A^T A\right)^T = A^T A$. Имеем $\left(A^T A\right)^T = A^T (A^T)^T = A^T A$.

2. Докажем, что матрица $A^T A$ определена положительно, т.е. $\forall X \in R^n,\ X \neq 0 \Rightarrow X^T A^T AX > 0$.

Имеем

$$X^T A^T AX = (AX)^T (AX) = \left|\ \text{обозначим}\ AX = Y\right| = Y^T Y > 0. \tag{1.40}$$

$Y^T Y > 0$, поскольку $Y = AX \neq 0$ ($AX \neq 0$, так как $\det A \neq 0$ по условию теоремы и потому СЛАУ $AX = 0$ имеет только нулевые решения $X = 0$, а в (1.40) $X \neq 0$.)

В заключении к подразделу 1.1 уже отмечалось, что в практике решения прикладных математических задач часто встречаются СЛАУ с симметричной и положительно определенной основной матрицей СЛАУ - нормальные СЛАУ. Для их решения целесообразно использовать метод квадратных корней. Вместе с тем при определенных условиях, например большой порядок СЛАУ $\left(n > 10^3\right)$, выгоднее, с точки зрения затрат машинного времени, использовать метод итераций.

Приведем без доказательства следующую теорему.

Теорема. Если СЛАУ $AX = B$ является нормальной, то процесс Зейделя для эквивалентной ей приведенной СЛАУ

$$X = \beta + \Gamma X$$

сходится при любом выборе начального вектора X^0.

1.4. Метод прогонки

1.4.1. Постановка задачи

Один частный, но очень важный случай представляют СЛАУ $AX = B$, матрица A которых имеет следующий вид:

$$A = \begin{pmatrix} a_{11} & a_{12} & 0 & \dots & 0 & 0 \\ a_{21} & a_{22} & a_{23} & \dots & 0 & 0 \\ 0 & a_{32} & a_{33} & \dots & 0 & 0 \\ . & . & . & . & . & . \\ 0 & 0 & 0 & \dots & a_{n,n-1} & a_{nn} \end{pmatrix}. \tag{1.41}$$

Матрицы A, у которых все или часть элементов на главной диагонали и на части диагоналей, параллельных главной (кодиагонали), отличны от нуля, а остальные элементы равны нулю, называются ленточными.

Общее число диагоналей с отличными от нуля элементами называется шириной ленты.

Матрица (1.41) имеет ширину ленты $L = 3$ (трехдиагональная матрица).

Системы уравнений с ленточными матрицами возникают при решении многих прикладных задач. Например, рассматривавшееся выше интегральное уравнение Фредгольма 1-го рода, но с разностным ядром $\int_a^b k(t-\tau)y(\tau)d\tau = f(t)$ порождает при построении дискретного аналога отвечающего ему уравнения Эйлера СЛАУ (1.14) с ленточной матрицей.

Для решения СЛАУ с ленточными матрицами можно использовать те же методы, что и в общем случае: метод Гаусса, метод квадратного корня (если A - симметричная и положительно определенная матрица), метод итераций и т.д. Если при этом программная реализация алгоритма учитывает ленточность матрицы A, то при $L \ll n$ (n - порядок СЛАУ) достигается существенный выигрыш в числе вычислительных операций.

Системы уравнений $AX = B$ с трехдиагональной матрицей A возникают при решении разностных уравнений, при разностной аппроксимации дифференциальных уравнений второго порядка и задач на собственные значения.

Пусть, например, решается задача Штурма-Лиувилля на собственные значения: найти λ и отвечающие им нетривиальные решения $y(x)$:

$$\begin{cases} y'' + \lambda y = 0, a < x < b, \\ y(a) = y(b) = 0. \end{cases} \tag{1.42}$$

Введем на отрезке $[a,b]$ равномерную сетку $a = x_0 < x_1 < x_2 < ... < x_n = b$,

$$\Delta x_i = \frac{b-a}{n} = h,\ x_i = x_0 + (i-1)\ h, i = \overline{1,n}. \tag{1.43}$$

Обозначим $y(x_i) = y_i$, $i = 1,2,...,n-1$. При этом $y_0 = y_n = 0$ в силу краевых условий. Полагая $y''(x_i) = \dfrac{y_{i-1} - 2y_i + y_{i+1}}{h^2}$, $i = 1,2,...,n-1$, получаем разностный аналог задачи (1.42)

$$\begin{cases} \dfrac{y_{i-1} - 2y_i + y_{i+1}}{h^2} + \lambda y_i = 0,\ i = 1,2,...,n-1, \\ y_0 = y_n = 0. \end{cases} \tag{1.44}$$

Или в матричной форме

$$\left(A + \lambda h^2 E\right) Y = 0, \tag{1.45}$$

где

$$A = \begin{pmatrix} 1 & 0 & 0 & ... & 0 & 0 \\ 1 & -2 & 1 & ... & 0 & 0 \\ 0 & 1 & -2 & ... & 0 & 0 \\ 0 & 0 & 0 & ... & 0 & 1 \end{pmatrix}, \quad Y = \begin{pmatrix} y_0 \\ y_1 \\ y_2 \\ . \\ . \\ y_n \end{pmatrix}.$$

Система линейных алгебраических уравнений (1.45) - дискретный аналог задачи на собственные значения (1.42). Одновременно при известном λ - это однородная СЛАУ с трехдиагональной матрицей: $A' = A + \lambda h^2 E$.

<u>Другой пример</u>. Пусть

$$\begin{cases} y'' + p(x) y = f(x), \quad a < x < b, \\ y(a) = A, \quad y(b) = B \end{cases} \tag{1.46}$$

- краевая задача. Ее дискретным аналогом будет система уравнений

$$\begin{aligned} &\frac{y_{i-1} - 2y_i + y_{i+1}}{h^2} + p_i y_i = f_i, \quad i = \overline{1, n-1}, \\ &y_0 = A, \quad y_n = B. \end{aligned} \tag{1.47}$$

Здесь $p_i = p(x_i)$, $f_i = f(x_i)$. В матричной форме система уравнений (1.47) будет иметь следующий вид:

$$\begin{pmatrix} 1 & 0 & 0 & ... & 0 & 0 \\ 1 & -2 + p_1 h^2 & 1 & ... & 0 & 0 \\ 0 & 1 & -2 + p_2 h^2 & ... & 0 & 0 \\ . & . & . & . & . & . \\ 0 & 0 & 0 & ... & -2 + p_{n-1} h^2 & 1 \\ 0 & 0 & 0 & ... & 0 & 1 \end{pmatrix} \begin{pmatrix} y_0 \\ y_1 \\ y_2 \\ . \\ . \\ . \\ y_{n-1} \\ y_n \end{pmatrix} = \begin{pmatrix} A \\ h^2 f_1 \\ h^2 f_2 \\ . \\ . \\ . \\ h^2 f_{n-1} \\ B \end{pmatrix}.$$

И в первом, и во втором примерах получились СЛАУ с трехдиагональными матрицами. Для решения таких СЛАУ предложен эффективный алгоритм, называемый <u>методом прогонки.</u>

Системы алгебраических уравнений (1.44),(1.47) с трехдиагональными матрицами A вида (1.41) принято записывать в следующем виде:

$$a_i y_{i-1} - c_i y_i + b_i y_{i+1} = f_i, \quad i = 1,2,...,n-1, \tag{1.48}$$

$$y_0 = \chi_0 y_1 + \mu_0, \quad y_n = \chi_n y_{n-1} + \mu_n, \tag{1.49}$$

где $a_i, c_i, b_i, f_i, i = \overline{1, n-1}$, и $\chi_0, \chi_n;\ \mu_0, \mu_n$ - заданные числа, а $y_0, y_1, ..., y_n$ - неизвестные.

Матрица A и вектор F матричной записи $AY = F$ СЛАУ (1.48),(1.49) имеют следующий вид:

$$A = \begin{pmatrix} 1 & -\chi_0 & 0 & .. & 0 & 0 \\ a_1 & -c_1 & b_1 & .. & 0 & 0 \\ 0 & a_2 & -c_2 & .. & 0 & 0 \\ . & . & . & . & . & . \\ 0 & 0 & 0 & .. & -c_{n-1} & b_{n-1} \\ 0 & 0 & 0 & .. & -\chi_n & 1 \end{pmatrix}, F = \begin{pmatrix} \mu_0 \\ f_1 \\ f_2 \\ . \\ . \\ . \\ f_{n-1} \\ \mu_n \end{pmatrix}, \dim A = (n+1) \times (n+1).$$

Нетрудно видеть, что полученные выше в примерах разностные уравнения (1.44) и (1.47) имеют тот же вид, что и уравнение (1.48). В связи с этим уравнение (1.48) называется <u>разностным уравнением второго порядка или трехточечным разностным уравнением.</u>

Уравнения (1.49) называются <u>краевыми условиями</u>.

В приведенных примерах краевые условия имели следующий вид: $y_0 = 0,\ y_n = 0$ - в первом примере (задача Штурма-Лиувилля) и $y_0 = A, y_n = B$ - во втором примере (1.46). Эти условия получаются из уравнений (1.49) при $\chi_0 = \chi_n = \mu_0 = \mu_n = 0$ - для первого примера и $\chi_0 = \chi_n = 0,\ \mu_0 = A, \mu_n = B$ - для второго.

Задачи (1.48),(1.49) в целом называются разностными краевыми задачами.

Разностные краевые задачи возникают прежде всего при приближенном решении дифференциальных уравнений. Однако такие же СЛАУ (1.48), (1.49) возникают при построении сплайнов и в некоторых других задачах.

1.4.2. Алгоритм метода прогонки

Прямой ход метода прогонки. Запишем первые два уравнения СЛАУ (1.48), (1.49)

$$y_0 = \chi_0 y_1 + \mu_0,$$
$$a_1 y_0 - c_1 y_1 + b_1 y_2 = f_1.$$

Подставив правую часть первого уравнения вместо y_0 во второе уравнение, получим $a_1(\chi_0 y_1 + \mu_0) - c_1 y_1 + b_1 y_2 = f_1$ или $a_1\mu_0 - f_1 + b_1 y_2 = (c_1 - a_1\chi_0)y_1$, откуда

$$y_1 = \chi_1 y_2 + \mu_1, \tag{1.50}$$

где

$$\chi_1 = \frac{b_1}{c_1 - a_1\chi_0},\ \mu_1 = \frac{a_1\mu_0 - f_1}{c_1 - a_1\chi_0}.$$

Далее подставляем правую часть $(\chi_1 y_2 + \mu_1)$ уравнения (1.50) вместо y_1 в третье $a_2 y_1 - c_2 y_2 + b_2 y_3 = f_2$, т.е. $a_2(\chi_1 y_2 + \mu_1) - c_2 y_2 + b_2 y_3 = f_2$ и выражаем из него y_2

$$y_2 = \frac{b_2}{c_2 - a_2\chi_1} y_3 + \frac{a_2\mu_1 - f_2}{c_2 - a_2\chi_1}.$$

Аналогично вводим обозначение $\chi_2 = \dfrac{b_2}{c_2 - a_2\chi_1}$,

$$\mu_2 = \frac{a_2\mu_1 - f_2}{c_2 - a_2\chi_1} \Rightarrow y_2 = \chi_2 y_3 + \mu_2$$

и т.д. В частности, на некотором k-м шаге будем иметь

$$y_k = \chi_k y_{k+1} + \mu_k, \tag{1.51}$$

где

$$\chi_k = \frac{b_k}{c_k - a_k\chi_{k-1}},\ \mu_k = \frac{a_k\mu_{k-1} - f_k}{c_k - a_k\chi_{k-1}},\ k = 1,2,\dots,n-1. \tag{1.52}$$

Коэффициенты χ_k, μ_k, $k = 1,2,\dots,n-1$, легко вычисляются, поскольку они выражаются через заданные коэффициенты $a_k, c_k, b_k, f_k, \chi_0, \mu_0$ и рекуррентно вычисляемые χ_{k-1}, μ_{k-1}.

Процесс вычисления коэффициентов χ_k, μ_k, $k = 1,2,\dots,n-1$, называется прямым ходом метода прогонки.

Обратный ход метода прогонки. При $k = n-1$ имеем в соответствии с (1.51) $y_{n-1} = \chi_{n-1} y_n + \mu_{n-1}$. Подставим правую часть этого выражения во второе краевое условие (1.49)

$$y_n = \chi_n(\chi_{n-1} y_n + \mu_{n-1}) + \mu_n,$$

откуда

$$y_n = \frac{\mu_n + \chi_n\mu_{n-1}}{1 - \chi_n\chi_{n-1}}. \tag{1.53}$$

Далее по формулам (1.51), но в обратном порядке:

$$y_k = \chi_k y_{k+1} + \mu_k,\ k = n-1, n-2, \dots, 1, 0 \tag{1.54}$$

находим остальные неизвестные $y_{n-1}, y_{n-2}, \dots, y_1, y_0$.

1.4.3. Условия существования и свойства решений разностной краевой задачи

Теорема 1.3. Если коэффициенты разностной краевой задачи (1.48),(1.49) удовлетворяют условиям

$$a_i \ne 0, b_i \ne 0, |c_i| \ge |a_i| + |b_i| \ge |a_i| > 0, i = \overline{1, n-1}, \tag{1.55}$$

$$|\chi_0| < 1,\ |\chi_n| \le 1, \tag{1.56}$$

то эта задача (СЛАУ) имеет единственное решение.

Доказательство. Формулы (1.51)-(1.53) справедливы (корректны), если знаменатели выражений (1.52) и (1.53) не обращаются в нуль. Поскольку $|\chi_0| < 1$, то для $|\chi_1|$ имеем в соответствии с (1.52)

$|\chi_1| = \dfrac{|b_1|}{|c_1 - a_1\chi_0|} \le \dfrac{|c_1| - |a_1|}{|c_1| - |a_1||\chi_0|} < \dfrac{|c_1| - |a_1|}{|c_1| - |a_1|} = 1$, т.е. $|\chi_1| < 1$. При доказательстве этого неравенства использовано, что в соответствии с (1.55) $0 \le |b_1| \le |c_1| - |a_1|$ и $|c_1 - a_1\chi_0| \ge |c_1| - |a_1\chi_0|$.

Аналогично доказывается, что $|\chi_2| < 1$, $|\chi_3| < 1 \dots |\chi_{n-1}| < 1$ и, следовательно, $c_k - a_k\chi_{k-1} \ne 0, \forall k = 1,2,\dots,n-1$, т.е. формулы (1.52) корректны при выполнении условий (1.55),(1.56). Поскольку $|\chi_n| \le 1$, $|\chi_{n-1}| < 1$, то и в формуле (1.53) знаменатель не равен нулю, более того: $1 - \chi_n\chi_{n-1} > 0$.

Таким образом, решение СЛАУ (1.48),(1.49) существует и находится оно по формулам (1.51)-(1.53). Единственность решения вытекает из тех же условий (1.55),(1.56). В этих условиях решение (1.51)-(1.53) СЛАУ существует для любой правой части F. А это равносильно тому, что $rgA = n+1$ или, $\det A \ne 0$. В соответствии с теоремой Кронекера-Капелли СЛАУ $AY = F$ имеет в этом случае $(rgA = n+1)$ единственное решение. ⊗

Замечание. Из неравенств $|\chi_k|<1$, $\forall k=1,2,...,n-1$, $|\chi_n|\le 1$ следует, что решение СЛАУ методом прогонки устойчиво к вычислительным ошибкам и ошибкам округлений. Действительно, пусть на некотором шаге $k+1$ получено $\tilde{y}_{k+1}:=y_{k+1}+\Delta y_{k+1}$ приближенное значение и Δy_{k+1}- ошибка приближения. В соответствии с (1.51) на следующем шаге будем иметь $y_k=\chi_k(y_{k+1}+\Delta y_{k+1})+\mu_k$. Погрешность этого шага $\Delta y_k=\chi_k\Delta y_{k+1}$, откуда $|\Delta y_k|=|\chi_k||\Delta y_{k+1}|<|\Delta y_{k+1}|$, т.е. погрешность не возрастает.

1.4.4.Оценивание затрат памяти ЭВМ и числа вычислительных операций

Для хранения матрицы **A** порядка $(n+1)$ необходимо $(3n+1)$ ячеек; $(n+1)$ ячейка нужна для хранения вектора **F** правой части СЛАУ, а $2n$ ячеек необходимо для хранения коэффициентов χ_k и μ_k. Всего требуется $2(3n+1)$ ячеек оперативной памяти.

Все вычисления в методе прогонки проводятся по формулам (1.51)-(1.53). Сохраняя принятые ранее обозначения для арифметических операций, получаем

формулы (1.52) - прямой ход: $N(*)=2(n-1)$, $N(/)=2(n-1)$, $N(\pm)=2(n-1)$;

формулы (1.53),(1.51) - обратный ход: $N(*)=2+n$, $N(/)=1$, $N(\pm)=2+n$.

В результате общее число медленных операций будет равно $4(n-1)+2+n+1=5n-1$, а быстрых -$3n$.

1.4.5. Примеры

Пример 1.5. Решить краевую задачу

$$\begin{cases} y''=6x, \\ y(0)=0,\ y(1)=2. \end{cases} \tag{1.57}$$

Решение. Сначала найдем аналитическое решение задачи

$$\begin{cases} y(x)=x^3+c_1x+c_2, \\ 0=c_2, \\ 2=1+c_1 \end{cases} \Rightarrow y(x)=x^3+x. \tag{1.58}$$

Построим равномерную сетку $\begin{cases} x_i=ih,\ i=0,1,...,5, \\ h=1/5=0,2. \end{cases}$.

Аппроксимируем $y''(x_i)=\dfrac{y_{i-1}-2y_i+y_{i+1}}{0,04}$.

В результате получим разностную краевую задачу

$$\begin{cases} y_{i-1}-2y_i+y_{i+1}=0,04\cdot 6\cdot x_i,\ i=1,2,...,4 \\ y_0=0,\ y_5=2. \end{cases} \tag{1.59}$$

Этой задаче отвечает СЛАУ

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 & 0 \\ 0 & 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{pmatrix}=\begin{pmatrix} 0 \\ 0,048 \\ 0,096 \\ 0,144 \\ 0,192 \\ 2 \end{pmatrix} \tag{1.60}$$

Имеем $\chi_0=\mu_0=0,\chi_5=0,\mu_5=2$. Будем решать СЛАУ (1.60) методом прогонки.

Прямой ход. В соответствии с формулами (1.52) вычисляем χ_k,μ_k:

$$\chi_1=\frac{1}{2-0}=0,5;\ \mu_1=\frac{-0,048}{2-0}=-0,024;$$
$$\chi_2=\frac{1}{2-0,5}=0,667;\ \mu_2=\frac{-0,024-0,096}{1,5}=-0,08;$$
$$\chi_3=\frac{1}{2-0,667}\approx 0,750;\ \mu_3=\frac{-0,08-0,144}{1,333}=-0,168;$$
$$\chi_4=\frac{1}{2-0,75}=0,8;\ \mu_4=\frac{-0,168-0,192}{1,25}=-0,288.$$

Обратный ход. По формуле (1.53) находим $y_5=\dfrac{2+0(-0,288)}{1-0\cdot 0,8}=2$.

Далее по формуле (1.51) последовательно вычисляем y_4,y_3,y_2,y_1,y_0:

$$y_4=2\cdot 0,8-0,288=1,312;$$
$$y_3=1,312\cdot 0,75-0,168=0,816;$$
$$y_2=0,816\cdot 0,667-0,08=0,464;$$
$$y_1=0,464\cdot 0,5-0,024=0,208;$$
$$y_0=0,208\cdot 0-0=0.$$

В таблице 1.1 приведены результаты вычисления $y(x_i)$ по формуле (1.58) в узлах сетки и отвечающие им результаты решения разностной краевой задачи методом прогонки.

Таблица 1.1

x_i	0	0,2	0,4	0,6	0,8	1,0
$y(x_i)=x_i^3+x_i$	0	0,208	0,464	0,816	1,312	2,0
y_i	0	0,208	0,464	0,816	1,312	2,0

Рассмотрим еще один пример решения разностной краевой задачи методом прогонки.

Пример 1.6. Решить краевую задачу

$$\begin{cases} y''-y'=-x, \\ y(0)=3; \quad y(2)=6+e^2. \end{cases} \tag{1.61}$$

Решение. Решением краевой задачи является функция

$y(x)=e^x+\dfrac{x^2}{2}+x+2$ (проверьте!).

Построим приближенное решение задачи (1.61). Для этого по аналогии с примером 1.5 сводим задачу (1.62) к разностной краевой задаче на равномерной сетке с 11 узлами:

$$\begin{cases} x_i=ih, \quad i=0,1,2,...,10, \\ h=\dfrac{2}{10}=0,2. \end{cases}$$

Аппроксимируем $y'(x)$ и $y''(x)$:

$$\hat{y}''(x_i)=\frac{y_{i-1}-2y_i+y_{i+1}}{0,04}; \; \hat{y}'(x_i)=\frac{y_{i+1}-y_{i-1}}{0,4}, \tag{1.62}$$

где

$$y_i=y(x_i); \quad y_{i-1}=y(x_{i-1}), \quad y_{i+1}=y(x_{i+1}).$$

Заметим, что аппроксимация производной $y'(x)$ осуществлена по формуле $\hat{y}'(x_i)=\dfrac{y(x_{i+1})-y(x_{i-1})}{2h}$.

Эта формула дает погрешность в оценивании производной $y''(x_i)$ равную $\dfrac{h^2}{6}y'''(\xi_i), \xi_i\in(x_{i-1},x_{i+1})$, тогда как формулы $\hat{y}'(x_i)=\dfrac{y_i-y_{i-1}}{h}$ и $\hat{y}'(x_i)=\dfrac{y_{i+1}-y_i}{h}$ дают погрешность $\dfrac{h}{2}y''(\xi_i)$.

Заменяя производные y', y'' в уравнении (1.61) их оценками (1.62), получаем разностный аналог краевой задачи:

$$\begin{cases} \dfrac{y_{i-1}-2y_i+y_{i+1}}{0,04}-\dfrac{y_{i+1}-y_{i-1}}{0,4}=-0,2i, \; i=\overline{1,9}, \\ y_0=3, \quad y_{10}=6+e^2 \end{cases}$$

или

$$\begin{cases} 1,1y_{i-1}-2y_i+0,9y_{i+1}=-0,008i, \; i=\overline{1,9} \\ y_0=3, \quad y_{10}=6+e^2. \end{cases} \tag{1.63}$$

Сопоставляя СЛАУ (1.63) с системой уравнений (1.48), (1.49), заключаем, что $a_i=1.1,\ c_i=2.0,\ b_i=0.9;\ \chi_0=\chi_{10}=0,\ \mu_0=3,\ \mu_{10}=6+e^2.$ Поскольку $|c_i|=|a_i+b_i|>|a_i|>0$ и $|\chi_0|<1, |\chi_{10}|<1$, то в соответствии с теоремой 1.3 СЛАУ (1.63) имеет единственное решение, которое может быть найдено методом прогонки.

Прямой ход. В соответствии с формулами (1.52) вычисляем коэффициенты χ_k, μ_k

$$\chi_1=\frac{0,9}{2-0\cdot 1,1}=0,45; \qquad \mu_1=\frac{1,1\cdot 3+0,008}{2}=1,654;$$

$$\chi_2=\frac{0,9}{2-1,1\cdot 0,45}\approx 0,598; \qquad \mu_2=\frac{1,1\cdot 1,654+0,016}{1,505}\approx 1,2195.$$

Результаты вычисления остальных коэффициентов сведены в табл. 1.2.

Таблица 1.2

i	0	1	2	3	4	5	6
χ_i	0	0,45	0,598	0,6706	0,7130	0,7403	0,759
μ_i	3	1,654	1,2195	1,0175	0,9120	0,8581	0,8366

Окончание табл. 1.2

7	8	9	10
0,7725	0,7824	0,7899	0
0,8379	0,8569	0,8905	13,3891

Обратный ход. По формулам (1.51) последовательно вычисляем $\hat{y}_9, \hat{y}_8, ..., \hat{y}_1, \hat{y}_0$ $(y_{10}=13,389)$

$$\hat{y}_9=0,7899\cdot 13,3891+0,8905\approx 11,4666;$$
$$\hat{y}_8=0,7824\cdot 11,4666+0,8569\approx 9,8284$$

и т.д.

В таблице 1.3 приведены результаты вычисления остальных значений $\hat{y}_i$, $i = 7,6,...,1$, а также значения функции $y(x) = e^x + \frac{x^2}{2} + x + 2$ в узлах сетки и отклонения $\Delta y_i = \hat{y}_i - y(x_i)$.

Таблица 1.3

i	1	2	3	4	5	6
$\hat{y}_i$	3,4407	3,9705	4,6004	5,3428	6,2143	7,2352
$y(x_i)$	3,4414	3,9718	4,6021	5,3455	6,2183	7,2401
Δy_i	-0,0007	-0,0013	-0,0017	0,0027	-0,004	-0,0049

Окончание табл. 1.3

7	8	9
8,4302	9,8284	11,4666
8,4352	9,8330	11,4696
-0,005	-0,0046	-0,003

1.5. Вычислительная устойчивость в численных методах

На примере решения СЛАУ численными методами удобно рассмотреть вопрос о вычислительной устойчивости численных методов.

Методы решения СЛАУ разделяются на конечные (иногда их называют еще точными) и приближенные.

Конечные методы (метод Гаусса, метод квадратного корня, формулы Крамера) обеспечивают получение за конечное число шагов точного решения X^* СЛАУ, естественно, если это решение существует, единственно и вычисления выполняются абсолютно точно. Приближенные методы позволяют получать, вообще говоря, приближенное решение $X^{(k)}$ с заданной степенью точности: $\left\| X^{(k)} - X^* \right\| < \varepsilon$. К приближенным методам относятся метод итераций и его модификации.

При реализации на ЭВМ любого метода решения СЛАУ деление их на точные и приближенные в значительной мере теряет смысл. Поскольку числа в ЭВМ записываются приближенно в соответствии с размером разрядной сетки, то все методы, как приближенные по своей сути, так и точные, оказываются приближенными. Более того, все методы, в том числе и точные, ведут себя по-разному в отношении ошибок округления. В некоторых алгоритмах ошибки округления могут быстро расти и приводить к существенным ошибкам в получаемом "точном" или "приближенном" решении. В других алгоритмах ошибки не накапливаются. Алгоритмы первой группы (любые, не обязательно конечные) называются неустойчивыми к вычислительным ошибкам, а алгоритмы второй группы - устойчивыми.

Осуществим краткий анализ вычислительной устойчивости метода Гаусса. Как известно, алгоритм метода Гаусса состоит из прямого и обратного ходов. На прямом ходе матрица A СЛАУ приводится к треугольному виду. В процессе приведения матрицы A к треугольному виду осуществляется деление внедиагональных элементов a_{ij} i-й активной подматрицы A_i на диагональный элемент a_{ii}, $i = 1,2,...,n$. Если элемент a_{ii} мал (по сравнению с "машинным нулем"), то результаты деления на него элементов a_{ij}, $j > i$ этой же строки могут быть очень большими (т.е. величины a_{ij} / a_{ii}). В конечном итоге это может приводить к неограниченному росту ошибок округления вплоть до аварийного останова процесса вычислений. С учетом деления вычислительных алгоритмов на устойчивые и неустойчивые к ошибкам округления метод Гаусса нужно отнести к неустойчивым вычислительным методам.

Для повышения вычислительной устойчивости численных методов используют различные приемы. В методе Гаусса, например, это выбор ведущего элемента a_{ii}^* на каждом шаге прямого хода метода Гаусса - максимального по модулю элемента i-го столбца активной подматрицы $A_i\left(a_{ii}^* = \max\limits_{i \le k \le n} |a_{ki}| \right)$ или максимального элемента во всей подматрице

$$A_i\left(a_{ii}^* = \max_{i \le k,s \le n} |a_{ks}| \right).$$

Метод квадратного корня, в отличие от метода Гаусса, является устойчивым к ошибкам округлений. В методе квадратного корня также присутствует деление на элементы r_{ii} [см. формулы (1.7)]. Однако положительная определенность матрицы A равносильна положительности диагональных элементов r_{ii} треугольной матрицы R. Действительно, в пункте 1.1.2 показано, что если $r_{ii}^2 > 0$, то и $\Delta_i > 0$, где Δ_i - i-й главный минор матрицы A и, наоборот, если матрица A определена положительно, то $\Delta_i > 0$ $\forall i = \overline{1,n}$ и, как следствие, $r_{ii} > 0$.

Вычислительная устойчивость метода квадратного корня является дополнительным аргументом в пользу его применения для решения нормальных СЛАУ.

Рассмотрим еще один подход к анализу влияния ошибок округления на вычислительную устойчивость метода. Этот подход позволит взглянуть на проблему вычислительной устойчивости методов решения СЛАУ с других позиций. Суть этого подхода заключается в замене прямой задачи учета ошибок округления в данном методе задачей анализа ошибок, возникающих при решении возмущений СЛАУ:

$$(A + \Delta A)X = B + \Delta B. \quad (1.64)$$

Здесь ΔA и ΔB - возмущения матрицы A и вектора B соответственно. Пусть $\tilde{X}_д$ - приближенное решение СЛАУ методом Гаусса, X^* - точное решение СЛАУ, $\delta X_д = \left\|\tilde{X}_д - X^*\right\| / \left\|X^*\right\|$ - относительная ошибка приближенного решения по методу Гаусса, $\tilde{X}_B$ - приближенное решение возмущенной СЛАУ (1.64). При этом возмущения ΔA и ΔB предполагаются такими, что $\delta\tilde{X}_B = \delta\tilde{X}_д$ $\left(\delta X_B = \left\|\tilde{X}_B - X^*\right\| / \left\|X^*\right\|\right)$.

Возмущения ΔA и ΔB, удовлетворяющие условию $\delta\tilde{X}_B = \delta\tilde{X}_д$, называются соответственно матрицей (ΔA) и вектором (ΔB) <u>эквивалентного возмущения</u> метода (в данном случае метода Гаусса). Подчеркнем еще раз, что эквивалентность здесь понимается в том смысле, что решение $\tilde{X}_B$ возмущенной системы отличается от точного X^* на такую же величину, что и приближенное решение $\tilde{X}_д$ в методе Гаусса, т.е. $\left\|\tilde{X}_B - X^*\right\| = \left\|\tilde{X}_д - X^*\right\|$. При этом смещение приближенного решения $\tilde{X}_B$ от точного X^* обусловлено <u>только</u> <u>возмущениями</u> матрицы $A\,(\Delta A)$ и вектора $B\,(\Delta B)$, а вычисления $\tilde{X}_B$ проводятся <u>точно.</u> Смещение приближенного решения $\tilde{X}_д$ в методе Гаусса от точного X^*, наоборот, обусловлено <u>только ошибками округлений</u>, а матрица A и вектор B СЛАУ предполагаются заданными точно.

Для относительной ошибки δX_B решения $\tilde{X}_B$ возмущенной СЛАУ (1.64) относительно точного решения X^* справедлива следующая оценка:

$$\delta X_B \le \frac{\|A\| \cdot \left\|A^{-1}\right\|}{1 - \left\|A^{-1}\right\| \cdot \|\Delta A\|}\left(\delta_A + \delta_B\right), \text{ где } \delta_A = \frac{\|\Delta A\|}{\|A\|},\ \delta_B = \frac{\|\Delta B\|}{B}. \qquad (1.65)$$

Из (1.65) следует, что относительная ошибка δX_B, как и следовало ожидать, пропорциональна уровню входных возмущений: возмущения матрицы (δ_A) и возмущения вектора (δ_B). Одновременно эта ошибка определяется коэффициентом $\|A\| \cdot \left\|A^{-1}\right\|$, который называется числом обусловленности матрицы A и обозначается ν_A. Формулу (1.65) можно записать еще в следующем виде

$$\delta X_B \le \frac{\nu_A}{1 - \nu_A \delta_A}\left(\delta_A + \delta_B\right).$$

Из этой формулы следует, что если число обусловленности ν_A велико $(\nu_A >> 1)$, а δ_A мало так, что $\nu_A \delta_A << 1$, то именно число обусловленности будет определять уровень δX_B относительной ошибки решения $\tilde{X}_B$ возмущенного уравнения. Большие значения коэффициента ν_A отвечают матрицам A, определитель $\det A$ которых близок к нулю (почти вырожденные матрицы).

Решение СЛАУ с почти вырожденными матрицами A методом Гаусса может приводить к большим ошибкам в получаемом решении $\tilde{X}_д$. Если такая матрица A задана точно $(\delta A = 0)$, но $\nu_A >> 1$, то даже небольшие возмущения вектора B могут приводить к большим значениям $\delta X_B\,(\delta X_B \sim \nu_A \delta_B)$.

Для получения устойчивых решений СЛАУ с почти вырожденной основной матрицей A (устойчивых к малым возмущениям δB правой части СЛАУ) используются специальные методы регуляризации. Они в данном пособии не обсуждаются.

2. ЧИСЛЕННЫЕ МЕТОДЫ РЕШЕНИЯ ПРОБЛЕМЫ СОБСТВЕННЫХ ЗНАЧЕНИЙ

2.1. Задачи, приводящие к проблеме собственных значений

Проблема собственных значений (с.з.) и собственных векторов (с.в.) линейного оператора является одной из центральных задач линейной алгебры. Решение этой проблемы имеет не только теоретическое, но большое и весьма разнообразное прикладное значение.

Так, квадратичная форма $H(\bar{x}) = X^T AX,\ \bar{x} = (x_1, \ldots, x_n)$ приводится к каноническому виду $H(\tilde{x}) = \lambda_1 \tilde{x}_1^2 + \lambda_2 \tilde{x}_2^2 + \ldots + \lambda_n \tilde{x}_n^2$ в базисе из <u>собственных</u> <u>векторов матрицы</u> A квадратичной формы. Этот результат используется для приведения к главным осям (к каноническому виду) алгебраических кривых и поверхностей второго порядка.

Решение систем линейных дифференциальных уравнений с постоянными коэффициентами $\frac{dY}{dx} = A_n Y$ может быть найдено в виде линейной комбинации $(C_1 Y_1 + \ldots + C_n Y_n)$ векторов-функций $Y_k = H_k e^{\lambda_k x}$, где λ_k - <u>с.з. матрицы</u> A, а H_k - отвечающие <u>с.в. матрицы</u> A системы дифференциальных уравнений.

Решение многих краевых задач для уравнений в частных производных может быть получено в виде обобщенного ряда Фурье по системе <u>собственных функций</u> дифференциального оператора, соответствующего данному уравнению (<u>собственных функций</u> задачи Штурма-Лиувилля, порождаемой краевой задачей).

Пусть

$$y(x) - \lambda\int_a^b K(x,s)\,y(s)\,ds = f(x), \qquad a \le x \le b, \qquad (2.1)$$

- интегральное уравнение Фредгольма 2-го рода с симметричным ядром $\mathbf{K(x,s)}$. Пусть далее $\tilde{\mathbf{A}}\mathbf{y} \equiv \int_a^b \mathbf{K(x,s)y(s)ds}$ - оператор, отвечающий этому уравнению. Тогда, если $\lambda \neq \frac{1}{\lambda_n}$, λ_n - с.з. оператора $\mathbf{A}$, то интегральное уравнение (2.1) имеет единственное решение $\mathbf{y(x)}$ при любой правой части $\mathbf{f(x)} \in \mathbf{L_2[a,b]}$ и это решение является рядом Фурье по системе собственных функций интегрального оператора $\tilde{\mathbf{A}}$.

В рассмотренных примерах задач на применение с.з. и с.в. линейных операторов используются все с.з. и с.в. линейного оператора. Задачу отыскания всех с.з. и с.в. линейного оператора называют полной проблемой собственных значений. В отдельных случаях для решения основной задачи достаточно знать одно или несколько первых максимальных (минимальных) по модулю с.з. линейных операторов. В этом случае говорят о решении частичной проблемы собственных значений.

Определение. Число λ называется с.з. линейного оператора $\tilde{\mathbf{A}}$, действующего в линейном пространстве $\mathbf{X}$, если существует ненулевой вектор $\bar{\mathbf{x}} \in \mathbf{X}$ такой, что

$$\tilde{\mathbf{A}}\bar{\mathbf{x}} = \lambda\bar{\mathbf{x}}. \qquad (2.2)$$

Вектор $\bar{\mathbf{x}}$ при этом называется с.в. оператора $\tilde{\mathbf{A}}$, отвечающим с.з. λ.

Далее рассмотрение проблемы собственных значений ограничим случаем операторов $\tilde{\mathbf{A}}$, действующих в конечномерных пространствах $\mathbf{X^n}$.

Если в пространстве $\mathbf{X^n}$ фиксирован базис $\bar{\mathbf{e}}_1, \bar{\mathbf{e}}_2, ..., \bar{\mathbf{e}}_n$, то оператору $\tilde{\mathbf{A}}$, действующему в этом пространстве, отвечает квадратная матрица $\mathbf{A}$, а вектору $\bar{\mathbf{x}}$ - $\mathbf{n}$-мерный арифметический вектор $\mathbf{X} = (\mathbf{x_1, x_2, ..., x_n})^T$ координат разложения вектора $\bar{\mathbf{x}}$ по базису $\{\bar{\mathbf{e}}_k\}$: $\bar{\mathbf{x}} = \mathbf{x_1}\bar{\mathbf{e}}_1 + ... + \mathbf{x_n}\bar{\mathbf{e}}_1$. Соотношению (2.2) при этом отвечает матричное уравнение $\mathbf{AX} = \lambda\mathbf{X}$ или

$$(\mathbf{A} - \lambda\mathbf{E})\mathbf{X} = \mathbf{0}. \qquad (2.3)$$

Как известно, ненулевое решение $\mathbf{X}$ однородной СЛАУ (2.3) существует тогда и только тогда, когда

$$\det(\mathbf{A} - \lambda\mathbf{E}) = \mathbf{0}. \qquad (2.4)$$

Таким образом, число λ будет с.з. оператора $\tilde{\mathbf{A}}$ $\left(\tilde{\mathbf{A}} : \mathbf{X^n} \to \mathbf{X^n}\right)$ или матрицы $\mathbf{A}$ тогда и только тогда, когда оно будет корнем характеристического уравнения (2.4). Собственный вектор $\bar{\mathbf{X}}$ оператора $\tilde{\mathbf{A}}$ (собственный

вектор $\mathbf{X} = (\mathbf{x_1, ..., x_n})^T$ матрицы $\mathbf{A}$), отвечающий с.з. λ, находится в результате решения СЛАУ (2.3).

Основные трудности решения полной проблемы собственных значений связаны с решением характеристического уравнения (2.4). При разложении определителя $\det(\mathbf{A} - \lambda\mathbf{E})$ в левой части (2.4) по степеням λ получается многочлен $\mathbf{n}$-й степени относительно λ:

$$\mathbf{P_n}(\lambda) \equiv (-1)^n \left[\lambda^n - \mathbf{p_1}\lambda^{n-1} + \mathbf{p_2}\lambda^{n-2} + ... + (-1)^n \mathbf{p_n}\right], \qquad (2.5)$$

где

$$\mathbf{p_1} = \sum_{i=1}^{n} \mathbf{a_{ii}}, \ \mathbf{p_2} = \sum_{1 \le i \le j \le n} \begin{vmatrix} \mathbf{a_{ii}} & \mathbf{a_{ij}} \\ \mathbf{a_{ji}} & \mathbf{a_{jj}} \end{vmatrix}, ..., \mathbf{p_n} = |\mathbf{A}|,$$

т.е. $\mathbf{p_k}$ - есть сумма всех диагональных миноров порядка $\mathbf{k}$ матрицы $\mathbf{A}$.

В соответствии с (2.6) необходимо вычислить $\mathbf{2^n - 1}$ определителей различных порядков (от первого до $\mathbf{n}$-го), а именно по $\mathbf{C_n^k} = \mathbf{n!/(k!(n-k)!)}$ определителей $\mathbf{k}$-го порядка, $\mathbf{k = 1,2,...,n}$. Отметим, что $\sum_{k=1}^{n} \mathbf{C_n^k} = \mathbf{2^n - 1}$.

Задача вычисления такого количества определителей при больших значениях $\mathbf{n}$ требует огромных затрат машинного времени.

Разработаны различные методы получения разложения определителя $\det(\mathbf{A} - \lambda\mathbf{E})$ в многочлен $\mathbf{P_n}(\lambda)$ без процедуры вычисления определителей (2.6) (методы Данилевского, Крылова, Леверье и др.). Они сокращают число вычислительных операций, необходимых для решения полной проблемы с.з. Однако и с применением этих методов полная проблема с.з. остается сложной в вычислительном отношении.

Для решения частичной проблемы с.з. разработаны эффективные итерационные алгоритмы.

2.2. Методы решения частичной проблемы собственных значений

2.2.1. Нахождение наибольшего по модулю с.з. матрицы и отвечающего ему с.в.

Предварительно приведем некоторые свойства с.з. и с.в. действительных матриц $(\mathbf{A} = (\mathbf{a_{ij}}), \mathbf{a_{ij}} \in \mathbf{R})$. Известно, что с.з. даже действительной матрицы могут быть комплексными и при этом еще и кратными.

Однако для симметричных матриц с действительными элементами и матриц с действительными положительными элементами справедливы следующие свойства.

Теорема 2.1. Собственные значения симметричной матрицы с действительными элементами действительны.

Доказательство. Пусть $\mathbf{X}$- собственное значение матрицы $\mathbf{A}$, а $\mathbf{X} = (\mathbf{x_1}, \mathbf{x_2}, ..., \mathbf{x_n})^{\mathbf{T}}$ - отвечающий ему собственный вектор.

Тогда $\mathbf{AX} = \lambda\mathbf{X}$ или $(\lambda\mathbf{X}, \mathbf{X}) = (\mathbf{X}, \lambda\mathbf{X})$. Поскольку $(\lambda\mathbf{X}, \mathbf{X}) = \lambda(\mathbf{X}, \mathbf{X})$, а $(\mathbf{X}, \lambda\mathbf{X}) = \bar{\lambda}(\mathbf{X}, \mathbf{X})$, где $\bar{\lambda}$ - сопряженное λ комплексное число, то $\lambda(\mathbf{X}, \mathbf{X}) = \bar{\lambda}(\mathbf{X}, \mathbf{X})$ или $(\lambda - \bar{\lambda})(\mathbf{X}, \mathbf{X}) = \mathbf{0}$. Так как собственный вектор $\mathbf{X}$ - ненулевой, то $(\mathbf{X}, \mathbf{X}) > \mathbf{0}$ и потому $\lambda - \bar{\lambda} = \mathbf{0}$ или $\bar{\lambda} = \lambda$, т.е. λ - действительное число. ⊗

Теорема 2.2. Собственные векторы симметричной матрицы с действительными элементами, отвечающие различным собственным значениям, попарно ортогональны.

Доказательство. Пусть λ и μ - два различных $(\lambda \neq \mu)$ собственных значения матрицы $\mathbf{A}$; $\mathbf{X}$ и $\mathbf{Y}$ - отвечающие им собственные векторы. Докажем, что $(\mathbf{X}, \mathbf{Y}) = \mathbf{0}$.

Имеем

$$\mathbf{AX} = \lambda\mathbf{X}, \ \mathbf{AY} = \mu\mathbf{Y}. \tag{2.6}$$

Составим скалярное произведение $(\mathbf{AX}, \mathbf{Y})$. Так как $\mathbf{A}^{\mathbf{T}} = \mathbf{A}$, то $(\mathbf{AX}, \mathbf{Y}) = (\mathbf{X}, \mathbf{AY})$ или с учетом (2.6) и теоремы 2.1:

$(\lambda\mathbf{X}, \mathbf{Y}) = (\mathbf{X}, \mu\mathbf{Y}) \Rightarrow \lambda(\mathbf{X}, \mathbf{Y}) = \mu(\mathbf{X}, \mathbf{Y}) \Rightarrow (\lambda - \mu)(\mathbf{X}, \mathbf{Y}) = \mathbf{0}$. Но $\lambda - \mu \neq \mathbf{0}$, поэтому

$$(\mathbf{X}, \mathbf{Y}) = \mathbf{0}.$$

⊗

Теорема 2.3. Действительная симметричная матрица является положительно определенной тогда и только тогда, когда все собственные значения ее <u>положительны</u>.

Теорема 2.4 (Перрона). Если все элементы квадратной матрицы *положительны*, то наибольшее по модулю собственное значение ее также положительно и является простым корнем характеристического уравнения матрицы, причем ему соответствует собственный вектор с положительными координатами.

Рассмотрим итерационный алгоритм нахождения наибольшего по модулю собственного значения λ_1 квадратной действительной матрицы $\mathbf{A}$.

Пусть с.з. $\lambda_1, \lambda_2, ..., \lambda_n$ действительной матрицы $\mathbf{A}$ удовлетворяют условию

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq ... \geq |\lambda_n| \tag{2.7}$$

и пусть $\mathbf{X_1}, \mathbf{X_2}, ..., \mathbf{X_n}$ - отвечающие им с.в. Возьмем произвольный вектор $\mathbf{Y}^{(0)} \in \mathbf{X}^{\mathbf{n}}$ и разложим его по с.в. $\{\mathbf{X_k}\}$ матрицы $\mathbf{A}$:

$$\mathbf{Y}^{(0)} = \sum_{i=1}^{n} \alpha_i \mathbf{X_i}. \tag{2.8}$$

Векторы $\mathbf{Y}^{(1)}, \mathbf{Y}^{(2)}, ..., \mathbf{Y}^{(k)}, ...$ будем строить по единому рекуррентному правилу

$$\mathbf{Y}^{(k)} = \mathbf{AY}^{(k-1)}. \tag{2.9}$$

Поскольку $\mathbf{X_i}$ - с.в. матрицы $\mathbf{A}$, то $\mathbf{AX_i} = \lambda_i \mathbf{X_i}$ и потому

$$\mathbf{Y}^{(1)} = \mathbf{AY}^{(0)} = |(\mathbf{2.8})| = \sum_{i=1}^{n} \alpha_i (\mathbf{AX_i}) = \sum_{i=1}^{n} \alpha_i \lambda_i \mathbf{X_i}.$$

Аналогично

$$\mathbf{Y}^{(2)} = \mathbf{AY}^{(1)} = \sum_{i=1}^{n} \alpha_i \lambda_i (\mathbf{AX_i}) = \sum_{i=1}^{n} \alpha_i (\lambda_i)^2 \mathbf{X_i},$$

$$, \dots \dots \dots \dots \dots \dots \dots \dots \tag{2.10}$$

$$\mathbf{Y}^{(k)} = \mathbf{AY}^{(k-1)} = \sum_{i=1}^{n} \alpha_i (\lambda_i)^{k-1} (\mathbf{AX_i}) = \sum_{i=1}^{n} \alpha_i (\lambda_i)^k \mathbf{X_i},$$

откуда для $\mathbf{j}$ – го элемента вектора $\mathbf{Y}^{(k)}$ будем иметь

$$\mathbf{y}_j^{(k)} = \alpha_1 (\lambda_1)^k \mathbf{x_{1j}} + \alpha_2 (\lambda_2)^k \mathbf{x_{2j}} + ... + \alpha_n (\lambda_n)^k \mathbf{x_{nj}}.$$

Покажем, что $\forall \mathbf{j} \in \{1, 2, ..., n\}$ $\lim\limits_{m \to \infty} \dfrac{\mathbf{y}_j^{(k+1)}}{\mathbf{y}_j^{(k)}} = \lambda_1$.

Действительно

$$\lim \frac{\mathbf{y}_j^{(k+1)}}{\mathbf{y}_j^{(k)}} = \lim_{k \to \infty} \frac{\alpha_1 (\lambda_1)^{k+1} \mathbf{x_{1j}} + \alpha_2 (\lambda_2)^{k+1} \mathbf{x_{2j}} + ... + \alpha_n (\lambda_n)^{k+1} \mathbf{x_{nj}}}{\alpha_1 (\lambda_1)^k \mathbf{x_{1j}} + \alpha_2 (\lambda_2)^k \mathbf{x_{2j}} + ... + \alpha_n (\lambda_n)^k \mathbf{x_{nj}}} =$$

$$= \lim_{k \to \infty} \lambda_1 \frac{1 + \dfrac{\alpha_2 \mathbf{x_{2j}}}{\alpha_1 \mathbf{x_{1j}}} \left(\dfrac{\lambda_2}{\lambda_1}\right)^{k+1} + ... + \dfrac{\alpha_n \mathbf{x_{nj}}}{\alpha_1 \mathbf{x_{1j}}} \left(\dfrac{\lambda_n}{\lambda_1}\right)^{k+1}}{1 + \dfrac{\alpha_2 \mathbf{x_{2j}}}{\alpha_1 \mathbf{x_{1j}}} \left(\dfrac{\lambda_2}{\lambda_1}\right)^{k} + ... + \dfrac{\alpha_n \mathbf{x_{nj}}}{\alpha_1 \mathbf{x_{1j}}} \left(\dfrac{\lambda_n}{\lambda_1}\right)^{k}} = \lambda_1.$$

При вычислении последнего предела использованы неравенства (2.7), из которых следует, что $\lim\limits_{k \to \infty} \left(\dfrac{\lambda_m}{\lambda_1}\right)^k = \mathbf{0}, \forall \mathbf{m} = \mathbf{2}, \mathbf{3}, ..., \mathbf{n}$.

Таким образом, приближенно можем положить

$$\lambda_1 \approx \frac{\mathbf{y}_i^{(k+1)}}{\mathbf{y}_i^{(k)}}, \ \mathbf{i} \in \{1, 2, ..., n\} \tag{2.11}$$

или

$$\lambda_1 \approx \frac{1}{n} \sum_{i=1}^{n} \frac{\mathbf{y}_i^{(k+1)}}{\mathbf{y}_i^{(k)}}. \tag{2.12}$$

При этом приближенное значение λ_1 можно вычислить с любой степенью точности. Для этого на каждой итерации, начиная со второй, нужно вычислять $\lambda_1^{(k)}$ и следить за поведением цифр после запятой в десятичном представлении числа $\lambda_1^{(k)}$. Вычисления можно прекращать, когда стабилизируются цифры в заданном количестве разрядов после запятой.

Замечание 1. Выбор начального вектора $\mathbf{Y}^{(0)}$ может влиять на скорость сходимости итерационного процесса.

Замечание 2. Вектор $\tilde{\mathbf{Y}}^{(k)} = \dfrac{1}{c_1\lambda_1^{(k)}}\mathbf{Y}^{(k)}$ является приближением к с.в. $\mathbf{X}_1$, отвечающему с.з. λ_1, т.е., если $\lambda_1^{(k)}$ (2.11) или (2.12) - приближенное значение максимального по модулю с.з. λ_1, то $\tilde{\mathbf{Y}}^{(k)}$- отвечающее ему приближенное значение с.в. $\mathbf{X}_1$. Действительно, так как $\left(\dfrac{\lambda_m}{\lambda_1}\right)^k \to 0$ при $k \to \infty$, то в соответствии с (2.10) имеем

$$\frac{1}{\alpha_1\lambda_1^k}\mathbf{Y}^{(k)} = \frac{1}{\alpha_1\lambda_1^k}\left(\alpha_1\lambda_1^k\mathbf{X}_1 + \alpha_2\lambda_2^k\mathbf{X}_2 + \ldots + \alpha_n\lambda_n^k\mathbf{X}_n\right) =$$

$$= \mathbf{X}_1 + \lambda_2\left(\frac{\lambda_2}{\lambda_1}\right)^k\mathbf{X}_2 + \ldots + \lambda_n\left(\frac{\lambda_n}{\lambda_1}\right)^k\mathbf{X}_n \approx \mathbf{X}_1.$$

Замечание 3. Если наибольшее по модулю с.з. λ_1 является кратным $\lambda_1 = \lambda_2 = \ldots = \lambda_m$ и $|\lambda_1| > |\lambda_k|$, $\forall k > m$, то проводя аналогичные рассуждения, можно показать, что $\lambda_1 = \lim\limits_{m\to\infty}\dfrac{y_i^{(k+1)}}{y_i^{(k)}}$, а вектор $\mathbf{Y}^{(k)} / \left\|\mathbf{Y}^{(k)}\right\|$ будет одним из с.в., отвечающих с.з. λ_1.

Пример 2.1. Найти наибольшее с.з. и отвечающий ему с.в. матрицы

$$\mathbf{A} = \begin{pmatrix} 3 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

Решение. Пусть $\mathbf{Y}^{(0)} = (1 \quad 1 \quad 1)^T$,

$$\mathbf{Y}^{(1)} = \begin{pmatrix} 3 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \\ 2 \end{pmatrix},\ \mathbf{Y}^{(2)} = \begin{pmatrix} 3 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} 4 \\ 4 \\ 2 \end{pmatrix} = \begin{pmatrix} 16 \\ 14 \\ 6 \end{pmatrix},$$

$$\lambda_1^{(1)} = \frac{1}{3}\left(\frac{16}{4} + \frac{14}{4} + \frac{6}{2}\right) = 3{,}5;$$

$$\mathbf{Y}^{(3)} = \begin{pmatrix} 3 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} 16 \\ 14 \\ 6 \end{pmatrix} = \begin{pmatrix} 62 \\ 50 \\ 20 \end{pmatrix},\ \lambda_1^{(2)} = \frac{1}{3}\left(\frac{62}{16} + \frac{50}{14} + \frac{20}{6}\right) = 3{,}5933;$$

$$\mathbf{Y}^{(4)} = \begin{pmatrix} 3 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} 62 \\ 50 \\ 20 \end{pmatrix} = \begin{pmatrix} 236 \\ 182 \\ 70 \end{pmatrix},\ \lambda_1^{(3)} = \frac{1}{3}\left(\frac{236}{62} + \frac{182}{50} + \frac{70}{20}\right) = 3{,}6488.$$

Дальнейшие результаты итерационного процесса приведены в табл.2.1.

Таблица 2.1

$\mathbf{Y}^{(4)}$	$\mathbf{Y}^{(5)}$	$\mathbf{Y}^{(6)}$	$\mathbf{Y}^{(7)}$	$\mathbf{Y}^{(8)}$	$\mathbf{Y}^{(9)}$	$\mathbf{Y}^{(10)}$	$\mathbf{Y}^{(11)}$
236 182 70	890 670 252	3340 2482 922	12502 9226 3404	46732 34358 12630	174554 128078 46988	651740 477698 175066	2432918 1782202 652764
$\lambda_1^{(m)}$	3,6842	3,7053	3,7174	3,7241	3,7278	3,7297	3,7308

Окончание таблицы 2.1

$\mathbf{Y}^{(12)}$	$\mathbf{Y}^{(13)}$	$\mathbf{Y}^{(14)}$	$\mathbf{Y}^{(15)}$
9080956 6650086 2434966	33892954 24816094 9085052	126494956 92610194 33901146	472095062 345616490 126511340
3,7314	3,7317	3,7319	3,7319

Можно принять $\lambda_1 \approx 3{,}732.$ В качестве с.в. $\mathbf{X}_1$, отвечающего с.з. $\lambda_1 \approx 3{,}732$, можно взять вектор $\mathbf{Y}^{(15)} / \left\|\mathbf{Y}^{(15)}\right\|$, т.е. $\mathbf{X}_1 = (0{,}7887 \quad 0{,}5774 \quad 0{,}2113)^T$.

2.2.2. Нахождение наибольшего по модулю с.з. симметричной матрицы

Для нахождения наибольшего по модулю с.з. и отвечающего ему с.в. симметричной матрицы $\mathbf{A}$ можно предложить алгоритм, основанный на методе скалярных произведений.

Введем для арифметического вектора $\mathbf{X} = (x_1, \ldots, x_n)^T$ еще одно обозначение $\bar{\mathbf{x}} = (x_1, \ldots, x_n)$, более удобное для записи скалярного произведения арифметических векторов $(\bar{x}, \bar{y})$.

Справедлива следующая теорема.

Теорема 2.5. Если с.з. симметричной матрицы связаны соотношением

$$|\lambda_1| > |\lambda_2| > ... > |\lambda_n|, \tag{2.13}$$

то $\lambda_1 = \lim\limits_{k\to\infty} \dfrac{\left(\bar{y}^{(k)}, \bar{y}^{(k+1)}\right)}{\left(\bar{y}^{(k)}, \bar{y}^{(k)}\right)}$, где $\bar{y}^{(k)} = A \cdot \bar{y}^{(k-1)}$ $\left(Y^{(k)} = A \cdot Y^{(k-1)}\right)$ и $Y^{(0)}$- произвольный вектор пространства R^n.

Доказательство. Как известно, с.з. симметричной матрицы действительны, а отвечающие им с.в. ортогональны.

Пусть $\bar{x}_1, \bar{x}_2, ..., \bar{x}_n$ – нормированные с.в. матрицы A, отвечающие с.з. $\lambda_1, \lambda_2, ..., \lambda_n$ соответственно. Разложим вектор $\bar{y}^{(0)}$ (вектор $Y^{(0)}$) по базису из с.в. $\{\bar{x}_k\}$, $k = \overline{1,n}$.

$$\bar{y}^{(0)} = \sum_{i=1}^{n} \alpha_i \bar{x}_i .$$

Тогда $\bar{y}^{(1)} = A\bar{y}^{(0)} = \sum\limits_{i=1}^{n} \alpha_i A\bar{x}_i = \left|A\bar{x}_i = \lambda_i \bar{x}_i\right| = \sum\limits_{i=1}^{n} \alpha_i \lambda_i \bar{x}_i$,

$$\bar{y}^{(2)} = A\bar{y}^{(0)} = \sum_{i=1}^{n} \alpha_i \lambda_i A\bar{x}_i = \sum_{i=1}^{n} \alpha_i \lambda_i^2 \bar{x}_i, ..., \bar{y}^{(k)} = A\bar{y}^{(k-1)} = ... = \sum_{i=1}^{n} \alpha_i \lambda_k \bar{x}_i .$$

Вычислим скалярные произведения $\left(\bar{y}^{(k)}, \bar{y}^{(k)}\right)$ и $\left(\bar{y}^{(k)}, \bar{y}^{(k+1)}\right)$.

Поскольку

$$\left(\bar{x}_i, \bar{x}_j\right) = \delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases}$$

то

$$\left(\bar{y}^{(k)}, \bar{y}^{(k)}\right) = \left(\sum_{i=1}^{n} \alpha_i \lambda_i^k \bar{x}_i, \sum_{j=1}^{n} \alpha_j \lambda_j^k \bar{x}_i\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \alpha_j \lambda_i^k \lambda_j^k (\bar{x}_i, \bar{x}_j) = \sum_{i=1}^{n} \alpha_i^2 \lambda_i^{2k}. \tag{2.14}$$

Аналогично получаем, что

$$\left(\bar{y}^{(k)}, \bar{y}^{(k+1)}\right) = \sum_{i=1}^{n} \alpha_i^2 \lambda_i^{2k+1} . \tag{2.15}$$

Вычислим $\lim\limits_{k\to\infty} \dfrac{\left(\bar{y}^{(k)}, \bar{y}^{(k+1)}\right)}{\left(\bar{y}^{(k)}, \bar{y}^{(k)}\right)}$. С учетом (2.14) и (2.15) имеем

$$\lim_{k\to\infty} \frac{\alpha_1^2 \lambda_1^{2k+1} + \alpha_2^2 \lambda_2^{2k+1} + ... + \alpha_n^2 \lambda_n^{2k+1}}{\alpha_1^2 \lambda_1^{2k} + \alpha_2^2 \lambda_2^{2k} + ... + \alpha_n^2 \lambda_n^{2k}} =$$

$$= \lambda_1 \lim_{k\to\infty} \frac{\alpha_1^2 + \alpha_2^2 (\lambda_2/\lambda_1)^{2k+1} + ... + \alpha_n^2 (\lambda_n/\lambda_1)^{2k+1}}{\alpha_1^2 + \alpha_2^2 (\lambda_2/\lambda_1)^{2k} + ... + \alpha_n^2 (\lambda_n/\lambda_1)^{2k}} = \lambda_1.$$

При вычислении последнего предела использованы неравенства (2.13), из которых следует, что $\lim\limits_{k\to\infty} (\lambda_m/\lambda_1)^k = 0$ для всех $m = 2,3,...,n.$ $\otimes$

Замечание 1. Если матрица A симметричная и положительно определенная, то приближенные значения λ_1, можно находить не только по формуле

$$\lambda_1 \approx \lim_{k\to\infty} \frac{\left(\bar{y}^{(k)}, \bar{y}^{(k+1)}\right)}{\left(\bar{y}^{(k)}, \bar{y}^{(k)}\right)} = \frac{\left(A\bar{y}^{(k-1)}, A\bar{y}^{(k)}\right)}{\left\|A\bar{y}^{(k-1)}\right\|^2}, \tag{2.16}$$

но и по формуле

$$\lambda_1 \approx \frac{\left\|\bar{y}^{(k+1)}\right\|}{\left\|\bar{y}^{(k)}\right\|} = \frac{\left\|A\bar{y}^{(k)}\right\|}{\left\|A\bar{y}^{(k-1)}\right\|}. \tag{2.17}$$

Замечание 2. Векторы $\bar{y}^{(k)}$ являются приближениями к с.в. $\bar{x}_1$ матрицы A, отвечающему с.з. λ_1. Действительно,

$$A\bar{y}^{(k)} = A\sum_{i=1}^{n} \alpha_i \lambda_i^k \bar{x}_i = \sum_{i=1}^{n} \alpha_i \lambda_i^{k+1} \bar{x}_i = \alpha_1 \lambda_1^{k+1} \left(\bar{x}_1 + \sum_{i=2}^{n} \alpha_i (\lambda_i/\lambda_1)^{k+1} \bar{x}_i\right) \approx$$

$$\approx \alpha_1 \lambda_1^{k+1} \bar{x}_1.$$

Поскольку $(\lambda_i/\lambda_1)^{k+1} \to 0$ при $k \to \infty$ для всех $i = 2,3,...,n$ и по предположению $\bar{x}_1$ – с.в., отвечающий λ_1, то полученное приближенное равенство $A\bar{y}^{(k)} \approx \alpha_1 \lambda_1^{k+1} \bar{x}_1$ означает, что вектор $\bar{y}^{(k)}$ является с.в. (с точностью до множителя $1/\alpha_1 \lambda_1^k$) матрицы A, отвечающим с.з. λ_1.

2.3. Решение полной проблемы собственных значений

Если симметричная матрица A имеет различные действительные с.з. $\lambda_1, \lambda_2, ..., \lambda_n$, то рассмотренная процедура вычисления наибольшего по модулю с.з. λ_1 и отвечающего ему с.в. $\bar{x}_1$ легко распространяется на решение полной проблемы с.з.

Теорема 2.6. Если с.з. действительной матрицы $\mathbf{A}$ различны: $|\lambda_1| > |\lambda_2| > \ldots > |\lambda_n|$, то с.з. $\lambda_l, \mathbf{i} = \mathbf{m}+1, \mathbf{m}+2, \ldots, \mathbf{n}$, матрицы $\mathbf{A}$ являются одновременно с.з. матрицы $\mathbf{A}^{(m)}$:

$$\mathbf{A}^{(m)} = \mathbf{A} - \sum_{k=1}^{m} \lambda_k \mathbf{X}_k \mathbf{X}_k^T, \quad \mathbf{m} = 1,2,\ldots,\mathbf{n}-1, \tag{2.18}$$

где $\lambda_k, \mathbf{X}_k$ – соответственно $\mathbf{k}$–е с.з. и $\mathbf{k}$–й с.в. матрицы $\mathbf{A}$, причем в формуле (2.18) векторы $\mathbf{X}_k$ ортонормированы, т.е.

$$\mathbf{X}_k^T \mathbf{X}_j = \delta_{kj} = \begin{cases} 1, & j = k, \\ 0, & j \neq k. \end{cases} \tag{2.19}$$

При этом

$$\mathbf{A} = \sum_{k=1}^{n} \lambda_k \mathbf{X}_k \mathbf{X}_k^T. \tag{2.20}$$

Доказательство. Пусть λ_1 - наибольшее по модулю с.з. матрицы $\mathbf{A}$, а $\mathbf{X}_1$ – отвечающий ему с.в. матрицы $\mathbf{A}$ $(\mathbf{A}\mathbf{X}_1 = \lambda_1 \mathbf{X}_1)$. Образуем матрицу

$$\mathbf{A}^{(1)} = \mathbf{A} - \sum_{k=1}^{m} \lambda_1 \mathbf{X}_1 \mathbf{X}_1^T. \tag{2.21}$$

Покажем, что с учетом $\mathbf{A}^{(1)}\mathbf{X}_1 = 0$, а $\mathbf{A}^{(1)}\mathbf{X}_j = \lambda_j \mathbf{X}_j$, $j = 2,3,\ldots,n$. Имеем с учетом (2.19) и (2.21)

$$\begin{aligned} \mathbf{A}^{(1)}\mathbf{X}_1 &= \left(\mathbf{A} - \lambda_1 \mathbf{X}_1 \mathbf{X}_1^T\right)\mathbf{X}_1 = \mathbf{A}\mathbf{X}_1 - \lambda_1 \mathbf{X}_1\left(\mathbf{X}_1^T \mathbf{X}_1\right) = \\ &= \left| \begin{matrix} \mathbf{A}\mathbf{X}_1 = \lambda_1 \mathbf{X}_1, \\ \mathbf{X}_1^T \mathbf{X}_1 = 1 \end{matrix} \right| = \lambda_1 \mathbf{X}_1 - \lambda_1 \mathbf{X}_1 = 0; \end{aligned} \tag{2.22}$$

$$\begin{aligned} \mathbf{A}^{(1)}\mathbf{X}_j &= \left(\mathbf{A} - \lambda_1 \mathbf{X}_1 \mathbf{X}_1^T\right)\mathbf{X}_j = \mathbf{A}\mathbf{X}_j - \lambda_1 \mathbf{X}_1\left(\mathbf{X}_1^T \mathbf{X}_j\right) = \\ &= \left| \begin{matrix} \mathbf{A}\mathbf{X}_j = \lambda_j \mathbf{X}_j, \\ \mathbf{X}_1^T \mathbf{X}_j = 0, j > 1 \end{matrix} \right| = \lambda_j \mathbf{X}_j. \end{aligned} \tag{2.23}$$

Аналогично доказывается, что $\mathbf{A}^{(m)}\mathbf{X}_j = 0$ для всех $1 \le j \le m$ и $\mathbf{A}^{(m)}\mathbf{X}_j = \lambda_j \mathbf{X}_j$ для $j = m+1, m+2, \ldots, n.$. Поскольку $\dim \mathbf{A}^{(1)} = n \cdot n$, то матрица $\mathbf{A}^{(1)}$ имеет $\mathbf{n}$ с.з.: $\lambda_1^{(1)}, \lambda_2^{(1)}, \ldots, \lambda_n^{(1)}$. Из (2.22) и (2.23) следует, что $\lambda_1^{(1)} = 0$ и $\lambda_j^{(1)} = \lambda_j$, $j = 2,3,..;n$. Таким образом, с.з. $\lambda_2^{(1)}, \ldots, \lambda_n^{(1)}$ матрицы $\mathbf{A}^{(1)}$ совпадают с с.з. $\lambda_2, \lambda_3, \ldots, \lambda_n$ матрицы $\mathbf{A}$, а

вместо с.з. λ_1 матрица $\mathbf{A}^{(1)}$ имеет с.з. $\lambda_1^{(1)} = 0$. У матрицы $\mathbf{A}^{(m)}$ нулевыми будут первые $\mathbf{m}$ с.з.: $\lambda_1^{(m)} = \lambda_2^{(m)} = \ldots = \lambda_n^{(m)} = 0.$ Остальные с.з. матрицы $\mathbf{A}^{(m)}$ $(\lambda_{m+1}^{(m)}, \ldots, \lambda_n^{(m)})$ совпадают с соответствующими с.з. матрицы $\mathbf{A}$, т.е.

$$\lambda_{m+1}^{(m)} = \lambda_{m+1}, \ldots, \lambda_n^{(m)} = \lambda_n. \qquad \otimes$$

Из доказанной теоремы вытекает следующий алгоритм решения полной проблемы с.з. в случае симметричной матрицы.

Алгоритм

1. Найти наибольшее по модулю с.з. λ_1 и отвечающий ему с.в. $\mathbf{X}_1$ матрицы $\mathbf{A}$. Для решения этой задачи можно использовать либо алгоритм , приведенный в п.2.2.1, либо в п. 2.2.2, если матрица $\mathbf{A}$ симметричная.
2. Составить матрицу $\mathbf{A}^{(1)} = \mathbf{A} - \lambda_1 \mathbf{X}_1 \mathbf{X}_1^T$.
3. Найти наибольшее по модулю с.з. λ_2 и отвечающий ему с.в. $\mathbf{X}_2$ матрицы $\mathbf{A}^{(1)}$ (матрицы $\mathbf{A}$).
4. Составить матрицу $\mathbf{A}^{(2)} = \mathbf{A}^{(1)} - \lambda_2 \mathbf{X}_2 \mathbf{X}_2^T$.
5. Найти наибольшее по модулю с.з. λ_3 и отвечающий ему с.в. $\mathbf{X}_3$ матрицы $\mathbf{A}^{(2)}$ (матрицы $\mathbf{A}$) и т.д.

На $(\mathbf{n}-1)$ - м шаге будет получена матрица $\mathbf{A}^{(n-1)}$ и найдены последнее с.з. λ_n и отвечающий ему с.в. $\mathbf{X}_n$. Признаком правильности вычислений может служить равенство $\mathbf{A}^{(n)} = 0$ или $\sum_{i=1}^{n} \lambda_i \mathbf{X}_i \mathbf{X}_i^T = \mathbf{A}$.

Библиографический список

1. Демидович Б.П., Марон И.А. Основы вычислительной математики. М.: Наука, 1966.
2. Бахвалов Н.С., Жидков Н.П., Кобельков Г.М. Численные методы. М.: Наука, 1987.
3. Самарский А.А., Гулин А.В. Численные методы. М.: Наука, 1989.
4. Фаддеев Д.К., Фаддеева В.Н. Вычислительные методы линейной алгебры. М.: Физматгиз, 1960.

5. Воеводин В.В. Вычислительные основы линейной алгебры. М.: Наука, 1977.
6. Икрамов Х.Д. Численные методы для симметричных линейных систем. М.: Наука, 1988.
7. Икрамов Х.Д. Численное решение матричных уравнений. М.: Наука, 1984.

Оглавление

Н о в и к о в Анатолий Иванович

Численные методы линейной алгебры

Редактор: И.П. Перехрест
Корректор: Н.А. Орлова
Лицензия №020446 от 04.03.97
Подписано в печать Формат бумаги 60x84 1/16.
Бумага газетная. Печать трафаретная. Усл. печ. л. 3,25.
Уч.-изд. л. 3,25. Тираж 100 экз. Заказ
Рязанская государственная радиотехническая академия
391000, Рязань, ул. Гагарина, 59/1.
Редакционно-издательский центр РГРТА