Лекция 8. Решение систем линейных алгебраических уравнений

1. Основные понятия и определения

Системы линейных алгебраических уравнений (СЛАУ) являются важной математической моделью линейной алгебры. На их базе ставятся такие практические математические задачи, как:

1. непосредственное решение линейных систем;
2. вычисление определителей матриц;
3. вычисление элементов обратных матриц;
4. определение собственных значений и собственных векторов матриц.

Решение линейных систем является одной из самых распространенных задач вычислительной математики. К их решению сводятся многочисленные практические задачи нелинейного характера, решения дифференциальных уравнений и др.

Обычно СЛАУ n-го порядка записывается в виде

$$\sum_{j=1}^{n} a_{ij}x_j = b_i; \quad i=\overline{1,n}$$

или в развернутой форме

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \qquad\qquad \ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n \end{cases} \tag{1}$$

или в векторной форме

$$A\bar{x} = \bar{b}\,, \tag{2}$$

где

$$A = \begin{bmatrix} a_{11}\ a_{12} \cdots a_{1n} \\ a_{21}\ a_{22} \cdots a_{2n} \\ \cdots \\ a_{n1}\ a_{n2} \cdots a_{nn} \end{bmatrix}; \quad \bar{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}; \quad \bar{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

В соотношениях (2):

A называется основной матрицей системы с n^2 элементами;

$\bar{x} = (x_1, x_2, \ldots, x_n)^{\mathrm{T}}$ – вектор-столбец неизвестных;

$\bar{b} = (b_1, b_2, \ldots, b_n)^{\mathrm{T}}$ – вектор-столбец свободных членов.

Определителем (детерминантом – det) матрицы A n-го порядка называется число D ($\det A$), равное

$$|A| = D = \det A = \begin{vmatrix} a_{11}\ a_{12} \cdots a_{1n} \\ a_{21}\ a_{22} \cdots a_{2n} \\ \cdots \\ a_{n1}\ a_{n2} \cdots a_{nn} \end{vmatrix} = \sum (-1)^k a_{1\alpha} a_{2\beta} \cdots a_{n\omega} \ .$$

Здесь индексы α, β, ..., ω пробегают все возможные n! перестановок номеров 1, 2, ..., n;

k – число инверсий в данной перестановке.

Первоначальным при решении СЛАУ (1) является анализ вида исходной матрицы A и вектора-столбца свободных членов $\bar{b}$ в (2).

Если все свободные члены равны нулю, т.е. $\bar{b} = 0$, то система $A\bar{x} = 0$ называется ***однородной***.

Если же $\bar{b} \neq 0$, или хотя бы одно $b_i \neq 0$ $(i = \overline{1,n})$, то система (2) называется ***неоднородной***.

Квадратная матрица A называется ***невырожденной***, или *неособенной*, если ее определитель $|A| \neq 0$.

При этом система (1) имеет единственное решение.

При $|A|=0$ матрица A называется ***вырожденной***, или *особенной*, а система (1) не имеет решения, либо имеет бесконечное множество решений.

Если $|A|\approx 0$ система (1) называется ***плохо обусловленной***, т.е. решение очень чувствительно к изменению коэффициентов системы.

Малые погрешности в вычислениях могут привести к большим погрешностям в решении, и задача решения СЛАУ является неустойчивой.

ЧИСЛО ОБУСЛОВЛЕННОСТИ cond(A)

При численном решении различных прикладных задач исследователи часто сталкиваются с таким понятием как число обусловленности. Это понятие описывается в учебниках по матричной алгебре. Число обусловленности является важнейшим «индикатором» для определения устойчивости решения той или иной задачи.

Известные американские математики Форсайт и Молер по поводу числа обусловленности пишут следующее: «Широко распространено заблуждение, что малость ***det(A)*** влечет за собой плохую обусловленность матрицы А». Далее, «…значение ***cond(A)*** является гораздо более важным критерием трудности решения линейной системы ***Ax=b***, чем малость ***det(A)***, либо громадность порядка *n*».

Обусловленность оценивает близость матрицы коэффициентов ***A*** к вырожденной.

Число обусловленности ***cond(A)*** является количественной оценкой обусловленности.

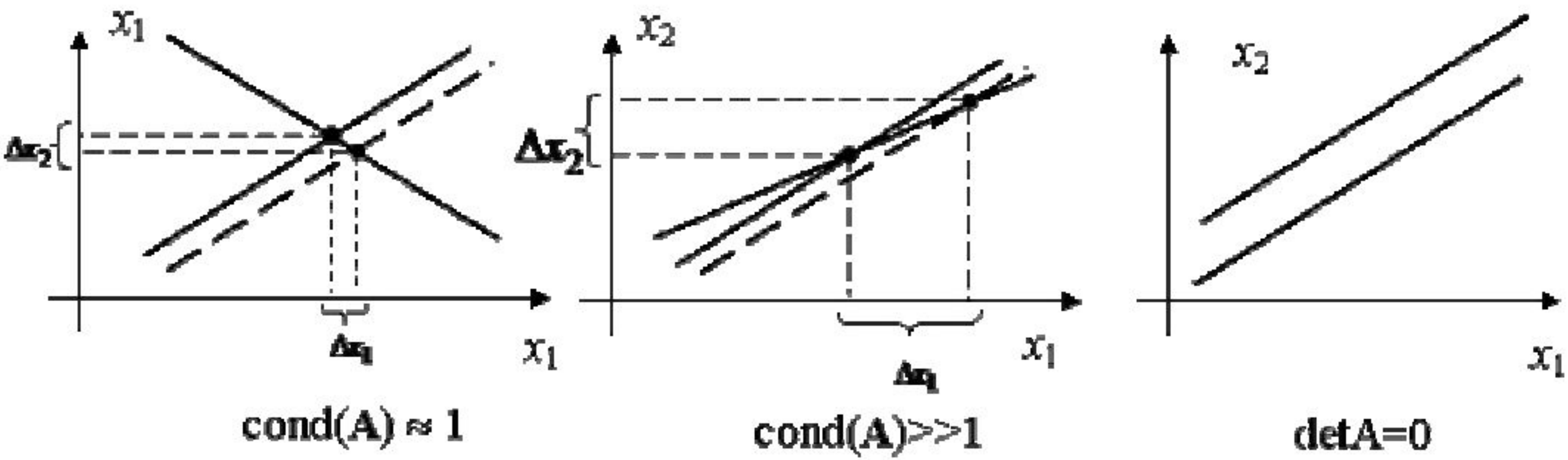
Всегда ***cond(A)≥1***.

Если $\boldsymbol{cond(A) \geq 10^3}$, то говорят, что матрица $\boldsymbol{A}$ плохо обусловлена.

Если $\boldsymbol{1 \leq cond(A) \leq 100}$, то матрица считается хорошо обусловленной.

Причина появления больших погрешностей при решении плохо обусловленных систем хорошо иллюстрируется на примере следующей системы линейных алгебраических уравнений (СЛАУ) с двумя неизвестными:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1, \\ a_{21}x_1 + a_{22}x_2 = b_2 . \end{cases}$$



а) *б*) *в*)

Рис. 1. Иллюстрация СЛАУ с двумя неизвестными:
а – хорошо обусловленная; *б* – плохо обусловленная; *в* – вырожденная система уравнений.

На рис. 1 случай (*а*) соответствует случаю хорошо обусловленной системы уравнений. Случай (*в*) -- это случай системы с вырожденной матрицей A ($\det(A) = 0$); здесь прямые, отвечающие каждому из уравнений, параллельны друг другу (уравнения линейно зависимы). Примером плохо обусловленной системы уравнений является случай (*б*) – прямые, соответствующие двум уравнениям, почти параллельны.

На рис. 1 случай (***а***) соответствует случаю хорошо обусловленной системы уравнений.

Случай (***в***) -- это случай системы с вырожденной матрицей ***A (det(A)=0)***; здесь прямые, отвечающие каждому из уравнений, параллельны друг другу (уравнения линейно зависимы).

Примером плохо обусловленной системы уравнений является случай (***б***) – прямые, соответствующие двум уравнениям, почти параллельны.

Штриховые прямые (***а***) и (***б***) на рис. 1 отвечают одному из уравнений, в котором немного изменены коэффициенты $\boldsymbol{a_{ij}}$ или правая часть $\boldsymbol{b_j}$.

Как видно, в случае хорошо обусловленной СЛАУ малые возмущения в величинах $\boldsymbol{a_{ij}}$ и $\boldsymbol{b_i}$ приводят к небольшим изменениям решения (точка пересечения прямых смещается незначительно).

В случае плохо обусловленной системы уравнений малые изменения в коэффициентах ведут к большим изменениям в решении (точка пересечения прямых смещается сильно).

Число обусловленности матрицы и устойчивость решения системы уравнений

- Система уравнений считается ***хорошо обусловленной***, если малые изменения в коэффициентах матрицы или в правой части вызывают малые изменения в решении.
- Система уравнений считается ***плохо обусловленной***, если малые изменения в коэффициентах матрицы или в правой части вызывают большие изменения в решении.

- Рассмотрим несколько примеров [2].

Пример 1. Дана система уравнений:

$$\begin{bmatrix} 1 & 2 \\ 2 & 3{,}999 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 7{,}999 \end{bmatrix}.$$

Решением уравнения (1) является следующее:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

Теперь внесем небольшие изменения в правую часть системы (1):

$$\begin{bmatrix} 1 & 2 \\ 2 & 3{,}999 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4.001 \\ 7{,}998 \end{bmatrix}. \qquad (2)$$

Решением системы (2) является

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -3{,}999 \\ 4{,}000 \end{bmatrix}.$$

Если теперь внесем малые изменения в коэффициентах системы (1)

$$\begin{bmatrix} 1{,}001 & 2{,}001 \\ 2{,}001 & 3{,}998 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 7{,}999 \end{bmatrix}, \tag{3}$$

То получим следующее решение

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3{,}994 \\ 0{,}001388 \end{bmatrix}.$$

Очевидно, система (1) является «плохо обусловленной», так как малые изменения, внесенные в коэффициентах матрицы или в правую часть, повлекли за собой большие изменения в решении системы.

Пример 2. Дана система уравнений:

$$\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 7 \end{bmatrix}. \tag{4}$$

Решением уравнения (1) является следующее:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

После внесения малых изменений в правую часть системы (1)

$$\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4{,}001 \\ 7{,}001 \end{bmatrix}, \tag{5}$$

получим решение

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1{,}999 \\ 1{,}001 \end{bmatrix}.$$

Если внесем малые изменения в коэффициентах матрицы

$$\begin{bmatrix} 1{,}001 & 2{,}001 \\ 2{,}001 & 3{,}001 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 7 \end{bmatrix}, \tag{6}$$

то получим решение

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2{,}003 \\ 0{,}997 \end{bmatrix}.$$

Как видно, система (4) является «хорошо обусловленной», так как малые изменения, внесенные в коэффициентах матрицы или в правую часть, повлекли за собой малые изменения в решении системы.

Как видно, система (4) является «хорошо обусловленной», так как малые изменения, внесенные в коэффициентах матрицы или в правую часть, повлекли за собой малые изменения в решении системы.

Эти простые примеры показывают, насколько небольшие изменения коэффициентов матрицы или правой части (вызванные различными источниками погрешности: ошибки округления, погрешности различных численных методов и т.д.) могут существенно изменить решение. И, действительно, с этой проблемой исследователи часто сталкиваются при численном решении различных прикладных задач.

Теперь рассмотрим, как можно вычислить число обусловленности матрицы.

Число обусловленности матрицы напрямую связано с понятием **норма матрицы**.

Как и определитель квадратной матрицы, норма матрицы – это число (скаляр).

Норма матрицы (квадратной, прямоугольной, обратимой или необратимой) всегда является положительным числом.

Понятие нормы универсально для любой **матрицы**, квадратной или неквадратной, **матрицы**-столбца или строки, размерность также может быть любой.

Эту характеристику используют в качестве оценочной величины для анализа изменяемости **матрицы** в каком-либо расчетном процессе или совокупности нескольких матриц.

Можно сказать, что норма является показателем «мощности» **матрицы**.
Она обозначается ‖**A**‖ и равна действительному числу, которое должно соответствовать определенному набору условий:

1) $\|A\| \geq 0$, где $[A]$ – матрица,

2) $\|kA\| = |k|\|A\|$, где $[A]$ – матрица, k – скаляр,

3) $\|A + B\| \leq \|A\| + \|B\|$, где $[A]$ и $[B]$ – матрицы одинакового порядка,

4) $\|AB\| \leq \|A\|\|B\|$, где $[A]$ и $[B]$ – матрицы, которые можно умножить друг на друга.

Существует три вида норм: бесконечная, первая норма и вторая (евклидова).

Все они являются каноническими, т.е. их значения не меньше по модулю любого матричного элемента. На практике обычно вычисляют только один из видов, этого достаточно для объективной оценки.

Различают несколько норм матрицы, например:

1. ∞–норма матрицы – это максимальная сумма модулей элементов каждой из строк матрицы:

$$\|A\|_\infty = \max_{1 \le i \le m} \sum_{j=1}^{n} |a_{ij}| .$$

2. 1-норма – это максимальная сумма модулей элементов каждого из столбцов матрицы:

$$\|A\|_1 = \max_{1 \le j \le n} \sum_{i=1}^{m} |a_{ij}| .$$

3. 2-норма (евклидова норма) – длина вектора в n-мерном пространстве (корень суммы квадратов всех элементов матрицы):

$$\|A\|_2 = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} (a_{ij})^2}$$

Чтобы найти **норму матрицы**, нужно воспользоваться одним из ниже приведенных способов для каждого вида. Все они основаны на расчете суммы элементов **матрицы**, но каждый подразумевает собственный алгоритм

Для расчета **бесконечной** нормы просуммируйте по модулю значения элементов отдельно по каждой строке и выберите из них максимальное

Найдите **первую норму**, поступив аналогично с элементами по каждому столбцу

Расчет **евклидовой** нормы подразумевает три действия: возведение каждого элемента в квадрат, суммирование и извлечение квадратной корня из общего результата

Пример: вычислите все виды норм для данной **матрицы**.

$$A = \begin{pmatrix} 5 & 6 \\ -3 & 9 \\ 4 & -1 \end{pmatrix}$$

Решение

1) $a_{11}+a_{12}=11;$

$a_{21}+a_{22}=12;$

$a_{31}+a_{32}=5 \rightarrow \|A\|_{\infty} = 12;$

Решение

2) $a_{11}+a_{21}+a_{31}=12;$
$a_{12}+a_{22}+a_{32}=16$
$\|A\|_1 = 16;$

Решение

$a_{11}+a_{12}=11$; $a_{21}+a_{22}=12$; $a_{31}+a_{32}=5 \rightarrow \|A\|\_1 = 12$;

$a_{11}+a_{21}+a_{31}=12$; $a_{12}+a_{22}+a_{32}=16 \rightarrow \|A\|\_2 = 16$;

$\|A\|_2 = \sqrt{25+36+9+81+16+1} = \sqrt{168} \approx 13.$

Пример 3. Вычислим ∞–норму матрицы A:

$$A = \begin{bmatrix} 10 & -7 & 0 \\ -3 & 2{,}009 & 6 \\ 5 & -1 & 5 \end{bmatrix}.$$

∞–норма матрицы A равна:

$$\|A\|_\infty = \max_{1 \le i \le 3} \sum_{j=1}^{3} |a_{ij}| = \max\left[\left(|10| + |-7| + |0|\right), \left(|-3| + |2{,}099| + |6|\right), \left(|5| + |-1| + |5|\right)\right] =$$

$$= \max\left[(17), (11{,}099), (11)\right] = 17.$$

Определим зависимость числа обусловленности от нормы матрицы

Вернемся к рассмотрению плохо обусловленной системы (1)

$$\begin{bmatrix} 1 & 2 \\ 2 & 3{,}999 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 7{,}999 \end{bmatrix}. \qquad (1)$$

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

Запишем систему (1) в новых обозначениях

$$[A][X] = [C]$$

Найдем ∞–нормы векторов X и C:

$$\|X\|_{\infty} = 2\,,\ \|C\|_{\infty} = 7{,}999\,.$$

После внесения малых изменений в правую часть системы (1) мы получили систему (2)

$$\begin{bmatrix} 1 & 2 \\ 2 & 3{,}999 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4.001 \\ 7{,}998 \end{bmatrix}. \qquad (2)$$

Обозначим систему (2) так:

$$[A][X'] = [C'].$$

Найдем изменение правой части и решения

Изменение правой части равно:

$$[\Delta C] = [C'] - [C] =$$

а изменение решения равно:

$$[\Delta X] = [X'] - [X]$$

Изменение правой части равно:

$$[\Delta C] = [C'] - [C] = \begin{bmatrix} 4{,}001 \\ 7{,}998 \end{bmatrix} - \begin{bmatrix} 4 \\ 7{,}999 \end{bmatrix} = \begin{bmatrix} 0{,}001 \\ -0{,}001 \end{bmatrix},$$

Изменение решения равно:

$$[\Delta X] = [X'] - [X] = \begin{bmatrix} -3{,}999 \\ 4 \end{bmatrix} - \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -5{,}999 \\ 3 \end{bmatrix}.$$

Найдем ∞–нормы векторов ΔX и ΔC

$\|\Delta C\|_\infty$=0,001 $\|\Delta X\|_\infty$=5,999

Отношение ∞–норм векторов ΔX , ΔC соответственно к X, C равно:

$$\frac{\|\Delta X\|_\infty}{\|X\|_\infty} = \frac{5{,}999}{2} = 2{,}9995\,,$$

$$\frac{\|\Delta C\|_\infty}{\|C\|_\infty} = \frac{0{,}001}{7{,}999} = 1{,}250 \times 10^{-4}\,.$$

Отношение ∞–норм векторов ΔX, ΔC соответственно к X, C равно:

$$\frac{\|\Delta X\|_\infty}{\|X\|_\infty} = \frac{5{,}999}{2} = 2{,}9995,$$

$$\frac{\|\Delta C\|_\infty}{\|C\|_\infty} = \frac{0{,}001}{7{,}999} = 1{,}250 \times 10^{-4}.$$

Как видно, малое изменение в правой части $1{,}250 \times 10^{-4}$ вызвало большое изменение в решении: 2,9995. Отношение изменения решения к изменению правой части равно:

$$\frac{\|\Delta X\|_\infty / \|X\|_\infty}{\|\Delta C\|_\infty / \|C\|_\infty} = \frac{2{,}9995}{1{,}250 \times 10^{-4}} = 23993.$$

Вернемся теперь к хорошо обусловленной системе (4). Найдем ∞–нормы векторов X и C:

$$\|X\|_{\infty} = 2, \ \|C\|_{\infty} = 7.$$

После внесения малых изменений в правую часть системы (4) мы получили систему (5). Найдем изменение правой части:

$$[\Delta C] = [C'] - [C] = \begin{bmatrix} 4{,}001 \\ 7{,}001 \end{bmatrix} - \begin{bmatrix} 4 \\ 7 \end{bmatrix} = \begin{bmatrix} 0{,}001 \\ 0{,}001 \end{bmatrix},$$

и изменение решения:

$$[\Delta X] = [X'] - [X] = \begin{bmatrix} 1{,}999 \\ 1{,}001 \end{bmatrix} - \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -0{,}001 \\ 0{,}001 \end{bmatrix}$$

Найдем ∞–нормы векторов ΔX и ΔC: $\|\Delta C\|_{\infty} = 0{,}001$, $\|X\|_{\infty} = 0{,}001$.

$$\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 7 \end{bmatrix}. \qquad (4)$$

$$\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4{,}001 \\ 7{,}001 \end{bmatrix}, \qquad (5)$$

Отношение ∞–норм векторов ΔX, ΔC соответственно к X, C равно:

$$\frac{\|\Delta X\|_\infty}{\|X\|_\infty} = \frac{0{,}001}{2} = 5\times 10^{-4},$$

$$\frac{\|\Delta C\|_\infty}{\|C\|_\infty} = \frac{0{,}001}{7} = 1{,}429\times 10^{-4}.$$

Как видно, малое изменение в правой части $1{,}429\times 10^{-4}$ вызвало небольшое изменение в решении: 5×10^{-4}. Отношение изменения решения к изменению правой части равно:

$$\frac{\|\Delta X\|_\infty / \|X\|_\infty}{\|\Delta C\|_\infty / \|C\|_\infty} = \frac{5\times 10^{-4}}{1{,}429\times 10^{-4}} = 3{,}5.$$

Очевидно, между $\dfrac{\|\Delta X\|}{\|X\|}$ и $\dfrac{\|\Delta C\|}{\|C\|}$, а также между $\dfrac{\|\Delta X\|}{\|X\|}$ и $\dfrac{\|\Delta A\|}{\|A\|}$ имеется некоторая связь. Э

Эта связь выражается в следующих неравенствах:

$$\frac{\|\Delta X\|}{\|X + \Delta X\|} \le \|A\| \|A^{-1}\| \frac{\|\Delta C\|}{\|C\|}, \qquad (7)$$

$$\frac{\|\Delta X\|}{\|X\|} \le \|A\| \|A^{-1}\| \frac{\|\Delta A\|}{\|A\|} \qquad (8).$$

Из неравенств (7), (8) следует, что отношение изменений нормы вектора правой части или нормы матрицы коэффициентов может увеличиваться почти в $\|A\|\|A^{-1}\|$-раз. Число $\|A\|\|A^{-1}\|$ и называется **числом обусловленности матрицы.** Используя число обусловленности вместе с машинным эпсилон, мы можем определить точность решения уравнения (1).

Из приведенных неравенств (7), (8), (11), (12) видно, что чем больше число обусловленности, тем сильнее сказывается на решении линейной системы ошибка в исходных данных. Грубо говоря, если $cond(A) \approx 10^p$ и исходные данные имеют погрешность в l-ом знаке после запятой, то независимо от способа решения системы в результате можно гарантировать не более $(l-p)$ знаков после запятой.

Как отмечалось выше, если число $cond(A)$ велико, то система считается плохо обусловленной. Оценка снизу для числа обусловленности дает 1, т.е. $cond(A)$ не может быть меньше 1. Для конкретной ЭВМ можно указать также верхнюю границу, превышение которой может привести к заведомо ложным решениям: решение считается ненадежным, если $cond(A) \geq (u)^{-1}$ или даже $cond(A) \geq (u)^{-0,5}$, где u – единичная ошибка округления (машинная точность). При этом важно отметить, что масштабирование матрицы А путем умножения на скаляр a не меняет ее число обусловленности.

пропустить

Пример 4. Найдем число верных цифр в решении следующей СЛАУ:

$$\begin{bmatrix} 1 & 2 \\ 2 & 3{,}999 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}.$$

Найдем число обусловленности для матрицы A: $[A] = \begin{bmatrix} 1 & 2 \\ 2 & 3{,}999 \end{bmatrix}$.

Сначала найдем обратную матрицу и норму матрицы A и ее обратной матрицы:

$$[A]^{-1} = \begin{bmatrix} -3999 & 2000 \\ 2000 & -1000 \end{bmatrix},$$

$$\|A\|_\infty = 5{,}999,$$

$$\|A^{-1}\|_\infty = 5999,$$

Число обусловленности равно

$$Cond(A) = \|A\|_\infty \|A^{-1}\|_\infty =$$

$$= 5{,}999 \times 5999{,}4 = 35990.$$

пропустить

Если учесть, что при представлении вещественных чисел на ЭВМ используются 24 значащие цифры, то машинная погрешность равна:

$$u = 2^{-23} = 0{,}119209 \times 10^{-6},$$

$$Cond(A) \times u = 35990 \times 0{,}119209 \times 10^{-6} = 0{,}4290 \times 10^{-2}.$$

Сравним полученный результат со значением $0{,}5 \times 10^{-m}$

$$0{,}5 \times 10^{-m} < 0{,}4290 \times 10^{-2}$$

$$m \leq 2.$$

Таким образом, в векторе решения будем иметь не более двух верных цифр после запятой.

Итак, понятие числа обусловленности очень важно при решении различных прикладных задач.

Число обусловленности является более важным критерием плохой обусловленности СЛАУ, чем малость определителя матрицы или ее большой порядок.

Понятие числа обусловленности очень важно при решении различных прикладных задач.

Число обусловленности является более важным критерием плохой обусловленности СЛАУ, чем малость определителя матрицы или ее большой порядок.

Итак, чем больше число обусловленности, тем сильнее сказывается на решении линейной системы ошибка в исходных данных.

Решение СЛАУ

Итак, обычно СЛАУ n-го порядка записывается в виде

$$\sum_{j=1}^{n} a_{ij}x_j = b_i; \quad i=\overline{1,n}$$

или в развернутой форме

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \quad\quad\quad \ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n \end{cases} \quad (1)$$

В ряде случаев получаются системы уравнений с матрицами специальных видов:

- диагональные,
- трехдиагональные (частный случай ленточных),
- симметричные ($a_{ij} = a_{ji}$),
- единичные (частный случай диагональной),
- треугольные и др.

Решение системы (2) заключается в отыскании вектора-столбца $\bar{x} = (x_1, x_2, \ldots, x_n)^{\mathrm{T}}$, который обращает каждое уравнение системы в тождество.

Существует два критерия близости приближенного решения к точному.

1. Если все приближенные значения $(x^*_1, x^*_2, \dots, x^*_n)$ близки к точным значениям $(x_1, x_2, \dots, x_n)$, то для любого $k = 1,\dots, n$ все погрешности $\Delta x_k = |x^*_k - x_k|$ достаточно малы. Для применения данного критерия получают еще одно приближенное решение, но более точное, чем имеющееся, например полученное с большим количеством значащих цифр. Тогда в качестве погрешности приближенного решения берут модуль разности между этими двумя значениями.
2. При подстановке с систему уравнений приближенные решения обращают уравнения в приближенные тождества, то есть **модули** величин $r_k = a_{k1}x^*_1 + a_{k2}x^*_2 + \dots + a_{kn}x^*_n - b_k$, именуемых **невязками**, достаточно близки к нулю. Данный критерий сводится к вычислению невязок, которые служат мерой, определяющей, на сколько приближенное решение обращает уравнения системы в тождества. Невязки рассматриваются по модулю и округляются **по правилам округления погрешностей.**

В случае приближенных значений исходных данных определитель матрицы из коэффициентов a_{ij} может быть не равен, но близок к нулю. В этом случае малые погрешности в вычислениях могут (но не обязательно) привести к большим погрешностям в решении, и задача решения системы уравнений является неустойчивой. Системы, чувствительные к погрешностям вычислений, называются **плохо обусловленными.**

Существуют две величины, характеризующие степень отклонения полученного решения от точного, которые появляются в связи с округлением и ограниченностью разрядной сетки ЭВМ, – погрешность ε и «невязка» r:

$$\begin{cases} \varepsilon = \bar{x} - \overline{x^*}; \\ r = \bar{B} - A\overline{x^*}; \end{cases} \qquad (3)$$

где $\overline{x^*}$ – вектор решения.

Как правило, значения вектора $\bar{x}$ – неизвестны.

Доказано, что если $\varepsilon \approx 0$, то и $r = 0$. Обратное утверждение не всегда верно. Однако если система не плохо обусловлена, для оценки точности решения используют невязку r.

2. Методы решения СЛАУ

Методы решения СЛАУ делятся на две группы:

– прямые (точные) методы;

– итерационные (приближенные) методы.

К ***прямым (точным)*** методам относятся такие методы, которые, в предположении, что вычисления ведутся без округлений, позволяют получить точные значения неизвестных за конечное число арифметических операций.

Они просты, универсальны и используются для широкого класса систем. Однако они не применимы к системам больших порядков ($n < 200$) и к плохо обусловленным системам из-за возникновения больших погрешностей.

К ним можно отнести:

- *правило* ***Крамера***,
- метод ***обратных матриц***,
- метод последовательного исключения неизвестных (метод ***Гаусса)***,
- метод ***Холецкого***,
- метод ***LU-разложения*** (использование мультипликативных разложений матриц),
- метод ***прогонки***,
- метод ***квадратного корня***, метод *наименьших квадратов* и др.

К ***приближенным*** относятся методы, которые даже в предположении, что вычисления ведутся без округлений, позволяют получить решение системы лишь с заданной точностью. Это итерационные методы (методы последовательных приближений) позволяют получать решение систем с помощью бесконечных сходящихся процессов.

К таким методам относятся:

- метод ***итераций*** ***(метод последовательных приближений)***
- метод ***Зейделя***
- метод ***Ричардсона с чебышевским набором параметров***
- метод ***минимальных поправок***
- метод ***скорейшего спуска***
- метод ***релаксации***

Прямые методы решения СЛАУ

1. Правило Крамера

Рассмотрим систему (1). Как отмечалось выше, если определитель этой системы не равен нулю, то будет иметь место единственное решение. Это необходимое и достаточное условие. Тогда по правилу Крамера

$$x_k = \frac{D_k}{D}, \ k = \overline{1, n}, \qquad (4)$$

где D_k – определитель, получающийся из D при замене элементов a_{1k}, a_{2k}, ..., a_{nk} k-го столбца (соответствующими) свободными членами b_1, b_2, ..., b_n из (1), или

$$D_k = \sum_{i=1}^{n} A_{ik} b_i \ , \ k = \overline{1,n} \quad ,$$

где A_{ik} алгебраическое дополнение элемента a_{ik} в определителе D.

Стоит существенная проблема вычисления определителей высоких порядков.

2. Метод обратных матриц

Дана система $A\bar{x} = \bar{b}$. Умножим левую и правую части этого выражения на A^{-1}:

$$A^{-1} A \bar{x} = A^{-1} \bar{b}; \qquad \bar{x} = A^{-1} \bar{b}.$$

При его реализации стоит проблема нахождения обратной матрицы A^{-1}, с выбором экономичной схемы ее получения и с достижением приемлемой точности.

3. Метод Гаусса

Этот метод является наиболее распространенным методом решения СЛАУ. В его основе лежит идея последовательного исключения неизвестных, в основном, приводящая исходную систему к треугольному виду, в котором все коэффициенты ниже главной диагонали равны нулю. Существуют различные вычислительные схемы, реализующие этот метод. Наибольшее распространение имеют схемы с выбором главного элемента либо по строке, либо по столбцу, либо по всей матрице. С точки зрения простоты реализации, хотя и с потерей точности, перед этими схемами целесообразней применять так называемую **схему единственного деления**. Рассмотрим ее суть.

Посредством первого уравнения системы (1) исключается x_1 из последующих уравнений. Далее посредством второго уравнения исключается x_2 из последующих уравнений и т.д. Этот процесс называется ***прямым ходом Гаусса***. Исключение неизвестных повторяется до тех пор, пока в левой части последнего n-го уравнения не останется одно неизвестное x_n

$$a'_{nn} x_n = b', \tag{5}$$

где a'_{nn} и b' – коэффициенты, полученные в результате линейных (эквивалентных) преобразований.

Прямой ход реализуется по формулам

$$\left.\begin{aligned} a^*_{mi} &= a_{mi} - a_{ki}\,\frac{a_{mk}}{a_{kk}}, \quad k=\overline{1,n-1}; \quad i=\overline{k,n} \\ b^*_m &= b_m - b_k\,\frac{a_{mk}}{a_{kk}}, \quad m=\overline{k+1,n} \end{aligned}\right\} \qquad (6)$$

где m – номер уравнения, из которого исключается x_k;

k – номер неизвестного, которое исключается из оставшихся $(n - k)$ уравнений, а также обозначает номер уравнения, с помощью которого исключается x_k;

i – номер столбца исходной матрицы;

a_{kk} – главный (ведущий) элемент матрицы.

Во время счета необходимо следить, чтобы $a_{kk} \neq 0$.

В противном случае прибегают к перестановке строк матрицы.

Обратный ход метода Гаусса состоит в последовательном вычислении x_n, x_{n-1}, …, x_1, начиная с (5) по алгоритму

$$x_n = b' / a'_{nn}; \quad x_k = \frac{1}{a'_{kk}}\left[b'_k - \sum_{i=k+1}^{n} a'_{ki}\, x_i\right], \quad k = \overline{n-1,1}. \qquad (7)$$

Точность полученного решения оценивается посредством «невязки» (3). В векторе невязки $(r_1, r_2, \ldots, r_n)^{\mathrm{T}}$ отыскивается максимальный элемент и сравнивается с заданной точностью ε.

Приемлемое решение будет, если $r_{\max} < \varepsilon$. В противном случае следует применить схему уточнения решения.

Уточнение корней

Полученные методом Гаусса приближенные значения корней можно уточнить.

Пусть для системы $A\bar{x} = \bar{b}$ найдено приближенное решение $\overline{x_0}$, не удовлетворяющее по «невязке».

Положим тогда $\bar{x} = \overline{x_0} + \bar{\delta}$. Для получения поправки $\delta = (\delta_1,\ \delta_2,\ \ldots,\ \delta_n)^{\mathrm{T}}$ корня $\overline{x_0}$ следует рассмотреть новую систему

$$A\,(\overline{x_0 + \delta}) = \bar{b} \qquad \text{или} \quad A\bar{\delta} = \bar{\varepsilon},$$

где $\bar{\varepsilon} = \bar{b} - A\overline{x_0}$ – невязка для исходной системы.

Таким образом, решая линейную систему с прежней матрицей A и новым свободным членом $\bar{\varepsilon} = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)^T$, получим поправки $(\delta_1, \delta_2, \ldots, \delta_n)$.

завершить

4. Модифицированный метод Гаусса

В данном случае помимо соблюдения требования $a_{kk} \neq 0$ при реализации формул (6) накладываются дополнительные требования, чтобы ведущий (главный) элемент в текущем столбце в процессе преобразований исходной матрицы имел максимальное по модулю значение. Это также достигается перестановкой строк матрицы.

Пример. В качестве иллюстрации преимущества модифицированного метода Гаусса, рассмотрим систему третьего порядка:

$$\begin{cases} 10x_1 - 7x_2 = 7; \\ -3x_1 + 2x_2 + 6x_3 = 4; \\ 5x_1 - x_2 + 5x_3 = 6. \end{cases} \qquad (a)$$

5. Метод прогонки

Данный метод также является модификацией метода Гаусса для частного случая *разреженных* систем – систем с матрицей трехдиагонального типа (краевая задача ДУ).

Каноническая форма их записи

$$a_i x_{i-1} + b_i x_i + c_i x_{i+1} = d_i; \quad i=\overline{1,n}; \; a_1 = c_n = 0, \tag{9}$$

или в развернутом виде

$$\begin{cases} b_1 x_1 + c_1 x_2 & = d_1; \\ a_2 x_1 + b_2 x_2 + c_2 x_3 & = d_2; \\ \qquad a_3 x_2 + b_3 x_3 + c_3 x_4 & = d_3; \\ \qquad\qquad \ldots & \\ \qquad\qquad a_{n-1} x_{n-2} + b_{n-1} x_{n-1} + c_{n-1} x_n & = d_{n-1}; \\ \qquad\qquad\qquad a_n x_{n-1} + b_n x_n & = d_n . \end{cases} \tag{10}$$

При этом, как правило, все коэффициенты $b_i \neq 0$.

Метод реализуется в два этапа – прямой и обратный ходы.

Прямой ход.

Каждое неизвестное x_i выражается через x_{i+1}

$$x_i = A_i \cdot x_{i+1} + B_i \text{ для } i = 1,2, \ldots, n-1, \tag{11}$$

посредством прогоночных коэффициентов A_i и B_i. Определим алгоритм их вычисления.

Из первого уравнения системы (10) находим x_1

$$x_1 = -\frac{c_1}{b_1} x_2 + \frac{d_1}{b_1}.$$

Из уравнения (11) при i=1: $x_1=A_1 \cdot x_2+B_1$.

Следовательно

$$A_1 = -\frac{c_1}{b_1}; \qquad B_1 = \frac{d_1}{b_1}. \tag{12}$$

Из второго уравнения системы (10) определяем x_2 через x_3, подставляя найденное значение x_1

$$a_2\,(A_1 x_2 + B_1) + b_2 x_2 + c_2 x_3 = d_2,$$

откуда

$$x_2 = \frac{-c_2 x_3 + d_2 - a_2 B_1}{a_2 A_1 + b_2}; \tag{12*}$$

и согласно (11) при $i = 2$: $x_2 = A_2 \cdot x_3 + B_2$, следовательно

$$A_2 = -\frac{c_2}{e_2}; \qquad B_2 = \frac{d_2 - a_2 B_1}{e_2}, \text{ где } e_2 = a_2 \cdot A_1 + b_2 .$$

Ориентируясь на соотношения индексов при коэффициентах (12) и (12*) можно получить эти соотношения для общего случая

$$A_i = -\frac{c_i}{e_i}; \qquad B_i = \frac{d_i - a_i B_{i-1}}{e_i},$$

где $e_i = a_i \cdot A_{i-1} + b_i$ (i=2,3, ..., n–1). (13)

Обратный ход.

Из последнего уравнения системы (10) с использованием (11) при $i = n–1$

$$x_n = \frac{d_n - a_n B_{n-1}}{b_n + a_n A_{n-1}} . \tag{14}$$

Далее посредством (11) и прогоночных коэффициентов (12), (13) последовательно вычисляем $x_{n-1}, x_{n-2}, \ldots, x_1$.

При реализации метода прогонки нужно учитывать, что при условии

$$| b_i | \geq | a_i | + | c_i | , \tag{15}$$

или хотя бы для одного b_i имеет место строгое неравенство (15), деление на «0» исключается, и система имеет единственное решение.

Заметим, что условие (15) является достаточным, но не необходимым.

В ряде случаев для хорошо обусловленных систем (10) метод прогонки может быть устойчивым и при несоблюдении условия (15).

Схема алгоритма метода прогонки может иметь вид, представленный на рис. 2.2.

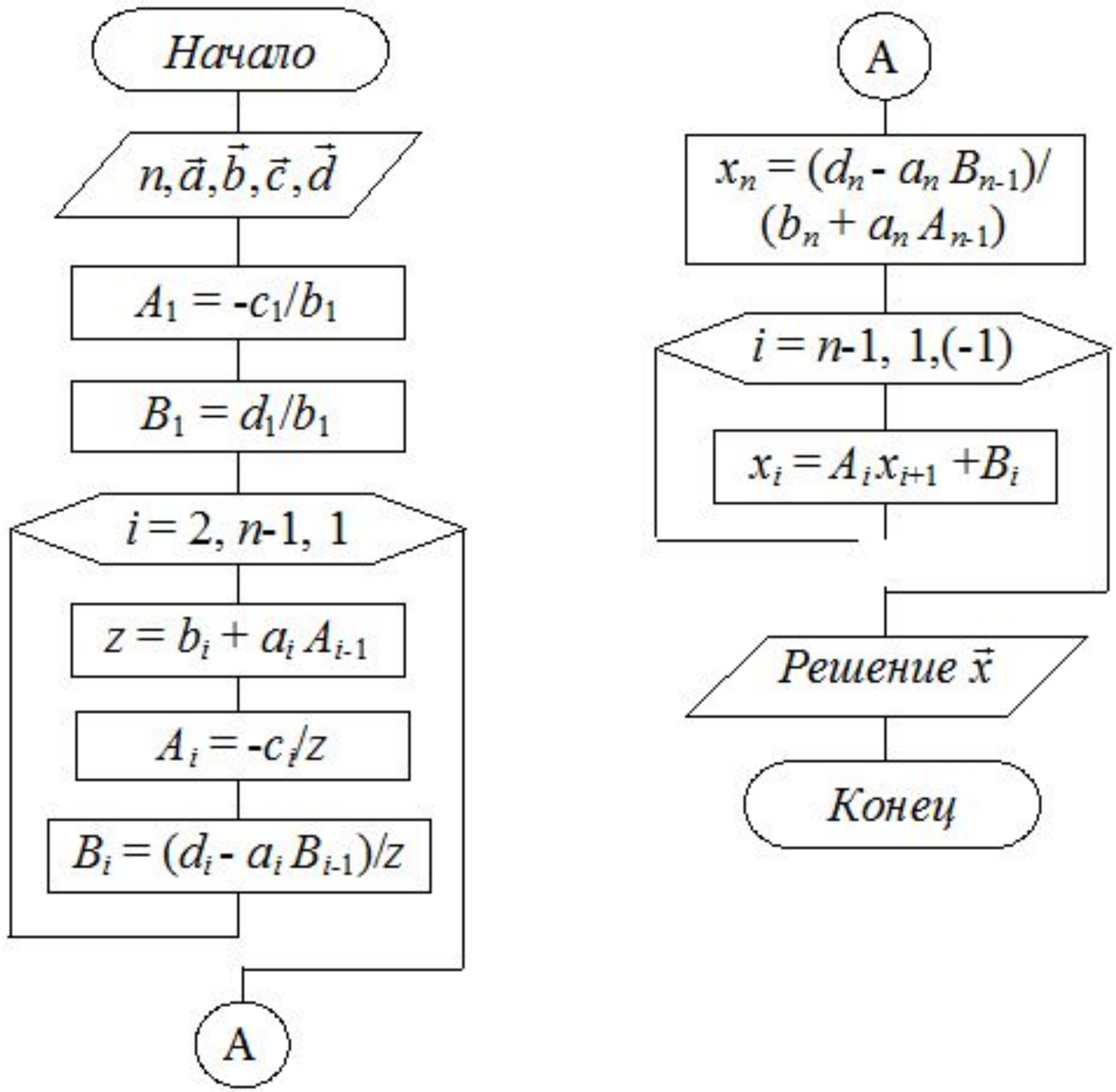


Рис. 2.2. Блок-схема метода прогонки