

МЕТОД ИСКЛЮЧЕНИЯ ГАУССА

Исключения Гаусса без выбора главного элемента. Вычислительная сложность метода исключения. Теорема об LU-разложении. Связь метода Гаусса и LU-разложения. Метод Гаусса с выбором главного элемента. Об устойчивости метода Гаусса. Вычисление определителей и обращение матриц.

Пусть задана система линейных алгебраических уравнений (СЛАУ)

$$Ax=b, \quad (1)$$

где A является матрицей порядка n , векторы x и b являются n -мерными векторами.

Метод Гаусса или метод последовательного исключения неизвестных относится к прямым методам решения СЛАУ. Общая идеология прямых методов решения СЛАУ состоит в преобразовании системы к некоторому более простому виду и в последующем решении получившейся более простой системы.

Схема вычислений без выбора главного элемента

Метод исключения Гаусса без выбора главного элемента, называемый просто методом Гаусса, состоит в том, что сначала система уравнений приводится к треугольному виду элементарными строчными преобразованиями без использования перестановки строк и/или столбцов (прямой ход), а затем неизвестные выражаются из полученной системы (обратный ход). Такая вычислительная схема накладывает некоторые ограничения на матрицу системы (1). Условия, при которых допустимы вычисления по формулам схемы, будут сформулированы несколько позже.

Запишем исходную систему (1) в развернутом виде

[illegible]

и опишем действия шагов прямого хода метода исключения Гаусса.

Шаг 1 (применяется к исходной системе). Исключим неизвестное x_1 из всех, кроме первого, уравнений системы (2). Для этого, предполагая $a_{11} \neq 0$, умножим первое уравнение на величину $\frac{a_{i1}}{a_{11}}$ и вычтем из i -го уравнения (2), $i=2,3,\dots,n$. В результате получим систему (первое уравнение оставляем без изменений)

$$\begin{aligned} a_{11}^{(0)}x_1 + a_{12}^{(0)}x_2 + \dots + a_{1n}^{(0)}x_n &= b_1^{(0)}, \\ a_{22}^{(1)}x_2 + \dots + a_{2n}^{(1)}x_n &= b_2^{(1)}, \end{aligned} \quad (3)$$

$$\dots\dots\dots$$

$$a_{n2}^{(1)}x_2 + \dots + a_{nn}^{(1)}x_n = b_n^{(1)}$$

(верхний индекс определяется номером шага, нулевой индекс приписан исходной системе). В последних $n-1$ уравнениях системы (3) не содержится неизвестное x_1 и их можно решать отдельно. К этой подсистеме с матрицей порядка $n-1$ применяется следующий шаг 2.

Шаг k , $2 \leq k \leq n-2$ (применяется к системе, полученной на $(k-1)$ -м шаге; в частности, при $k=2$ применяется к системе (3)). В последних, начиная с k -го, уравнениях системы, полученной на предыдущем, $(k-1)$ -м шаге, не содержатся неизвестные x_1, \dots, x_{k-1} . Исключим в уравнениях с номерами $k+1, \dots, n$ неизвестное x_k . Для этого, предполагая $a_{kk}^{(k-1)} \neq 0$, умножим k -е уравнение на величину $\frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}}$ и вычтем из i -го уравнения (2), $i=k+1, \dots, n$. В результате получим систему

$$\begin{aligned} a_{11}^{(0)}x_1 + a_{12}^{(0)}x_2 + \dots + a_{1n}^{(0)}x_n &= b_1^{(0)}, \\ a_{22}^{(1)}x_2 + \dots + a_{2n}^{(1)}x_n &= b_2^{(1)}, \\ \dots\dots\dots \\ a_{kk}^{(k-1)}x_k + \dots + a_{kn}^{(k-1)}x_n &= b_k^{(k-1)}, \\ a_{k+1,k+1}^{(k)}x_{k+1} + \dots + a_{k+1,n}^{(k)}x_n &= b_{k+1}^{(k)}, \\ \dots\dots\dots \\ a_{n,k+1}^{(k)}x_{k+1} + \dots + a_{nn}^{(k)}x_n &= b_n^{(k)} \end{aligned} \quad (4)$$

(при $k=2$ строка $a_{kk}^{(k-1)}x_k + \dots + a_{kn}^{(k-1)}x_n = b_k^{(k-1)}$ есть вторая строка системы (4)).

Шаг $n-1$ (применяется к системе (4) при $k=n-2$). В последних двух уравнениях системы, полученной на предыдущем, $(n-2)$ -м шаге, не содержатся неизвестные x_1, \dots, x_{n-2} . Исключим неизвестное x_{n-1} из последнего уравнения системы. Для этого, предполагая $a_{n-1,n-1}^{(n-2)} \neq 0$, умножим предпоследнее уравнение на $\frac{a_{i,n-1}^{(n-2)}}{a_{n-1,n-1}^{(n-2)}}$ и вычтем из последнего уравнения. В итоге получим систему

$$\begin{aligned} a_{11}^{(0)}x_1 + a_{12}^{(0)}x_2 + \dots + a_{1n}^{(0)}x_n &= b_1^{(0)}, \\ a_{22}^{(1)}x_2 + \dots + a_{2n}^{(1)}x_n &= b_2^{(1)}, \\ \dots\dots\dots \\ a_{nn}^{(n-1)}x_n &= b_n^{(n-1)}. \end{aligned} \quad (5)$$

Таким образом, прямой ход метода исключения заключается в преобразовании системы (2) к треугольному виду (5) и осуществляется с помощью следующих формул:

$$a_{ij}^{(0)} = a_{ij}, \quad i=1,2,\dots,n, \quad j=1,2,\dots,n;$$

$$b_i^{(0)} = b_i, \quad i=1,2,\dots,n;$$

$$k = 1, 2, \dots, n-1:$$

$$i = k+1, k+2, \dots, n:$$

$$l_{ik} = \frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}}, \quad (6)$$

$$b_i^{(k)} = b_i^{(k-1)} - l_{ik} b_k^{(k-1)}, \quad (7)$$

$$j = k+1, k+2, \dots, n:$$

$$a_{ij}^{(k)} = a_{ij}^{(k-1)} - l_{ik} a_{kj}^{(k-1)}. \quad (8)$$

Отметим, что при программировании существенными являются нижние индексы, а верхний индекс оставлен для наглядности, при программировании он не принимается во внимание.

Диагональный элемент $a_{kk}^{(k-1)}$ преобразуемой матрицы, полученный на $(k-1)$ -м шаге и играющий в вычислениях k -го шага особую роль, называется на k -м шаге прямого хода главным элементом. Главный элемент называют еще ведущим элементом. Позже будет рассмотрен вариант метода Гаусса, в котором главным (ведущим) элементом может быть выбран не обязательно $a_{kk}^{(k-1)}$.

Обратный ход заключается в получении значений неизвестных из треугольной системы (5): из последнего уравнения получаем значение неизвестного x_n ; подставляя его в предпоследнее уравнение, находим x_{n-1} ; подставляя оба этих неизвестных в третье с конца уравнение, находим x_{n-2} , и так далее. Вычисления осуществляются по формулам

$$x_n = \frac{b_n^{(n-1)}}{a_{nn}^{(n-1)}}, \quad x_i = \frac{1}{a_{ii}^{(i-1)}} \left(b_i^{(i-1)} - \sum_{j=i+1}^n a_{ij}^{(i-1)} x_j \right), \quad i=n-1, n-2, \dots, 1. \quad (9)$$

Замечание 1. В рассмотренной схеме исключения k -е уравнение системы на k -м шаге прямого хода не преобразуется. Часто рассматривают другой вариант исключения: на k -м шаге k -е уравнение делят на $a_{kk}^{(k-1)}$. В этом случае вместо системы (5) получится система с единичными коэффициентами при x_k в k -м уравнении.

Вычислительная сложность метода исключения

Вычислительная сложность (иногда называемая трудоемкостью) – это свойство алгоритма, которое характеризует требуемое для выполнения алгоритма число операций сложения, вычитания, умножения, деления, извлечения корня. Вычислительная сложность является одной из важных

характеристик любого численного метода. Часто подсчитывают только операции умножения, деления и извлечения корня, так как время их выполнения компьютером многократно превосходит время выполнения операций сложения и вычитания.

Замечание 2. Время решения задачи на компьютере (как на однопроцессорном, так и на многопроцессорном) зависит не только от числа выполняемых вычислительных операций, но и от других факторов: возможностей компьютера, степени использования памяти с быстрым доступом, согласованности структуры алгоритма с архитектурой компьютера.

Подсчитаем число умножений и делений, требуемых для нахождения решения системы из n уравнений по формулам (6)–(9). С учетом равенств

$$\sum_{k=1}^{n-1} (n-k) = 1 + 2 + \dots + (n-1) = \frac{n(n-1)}{2},$$

$$\sum_{k=1}^{n-1} (n-k)^2 = 1^2 + 2^2 + \dots + (n-1)^2 = \frac{n(n-1)(2n-1)}{6}$$

найдем число умножений и делений в формулах прямого хода (6)–(8):

$$\begin{aligned} \sum_{k=1}^{n-1} \sum_{i=k+1}^n \left(2 + \sum_{j=k+1}^n 1 \right) &= \sum_{k=1}^{n-1} \sum_{i=k+1}^n (2 + (n-k)) = \sum_{k=1}^{n-1} (n-k+2)(n-k) = \\ \sum_{k=1}^{n-1} (n-k)^2 + 2 \sum_{k=1}^{n-1} (n-k) &= \frac{(n-1)n(2n-1)}{6} + 2 \frac{n(n-1)}{2} = n(n-1) \left(\frac{n}{3} + \frac{5}{6} \right) = \\ &= \frac{n^3}{3} + \frac{n^2}{2} - \frac{5}{6}n. \end{aligned}$$

Для обратного хода из формул (9) получим:

$$n + \sum_{i=1}^{n-1} \sum_{j=i+1}^n 1 = n + \sum_{i=1}^{n-1} (n-i) = n + \frac{n(n-1)}{2} = \frac{n(n+1)}{2}.$$

Если требуется решить нескольких систем с различными правыми частями и одной и той же матрицей, то в алгоритм исключений для каждой новой правой части добавляется строка вида (7) и добавляются вычисления обратного хода вида (9). Если имеется P правых частей, то можно вместо вычислений вида (7) ввести в рассмотрение расширенную P столбцами матрицу системы и проводить преобразования типа (8) расширенной матрицы.

Выводы:

- число умножений и делений, выполняемых при нахождении решения системы из n уравнений методом исключений, составляет величину порядка $\frac{1}{3}n^3$ для прямого хода и величину порядка $\frac{1}{2}n^2$ для обратного хода;

- при больших n вычислительная сложность метода Гаусса решения одной СЛАУ составляет величину порядка $\frac{1}{3}n^3$ (основной объем вычислений приходится на прямой ход);
- при больших n вычислительная сложность решения P систем с одной и той же матрицей практически не отличается, если P небольшое, от вычислительной сложности решения одной системы; если P сравнимо с n , то вычислительную сложность решения P систем можно оценить величиной (доказательство не приводится) $\frac{4}{3}n^3$.

Замечание 3. Если при выполнении k -го шага процесса исключения неизвестных исключать неизвестное x_k не только из уравнений с номерами $k+1, \dots, n$, но и из уравнений с номерами $1, \dots, k-1$, то получим модификацию метода Гаусса, которая и носит название метода Гаусса–Жордана. При таком подходе в результате выполнения прямого хода матрица преобразуется не к треугольному, а к диагональному виду. Приведём алгоритм для случая P правых частей b_{ij} , $i=1,2,\dots,n$, $j=1,2,\dots,P$:

$$\begin{aligned}
 a_{ij}^{(0)} &= a_{ij}, \quad i=1,2,\dots,n, \quad j=1,2,\dots,n; \\
 a_{ij}^{(0)} &= b_{i,j-n}, \quad i=1,2,\dots,n, \quad j=n+1, n+2, \dots, n+P; \\
 k &= 1, 2, \dots, n: \\
 &\quad i = 1, 2, \dots, n: \\
 &\quad \quad \text{if } i \neq k: \\
 &\quad \quad \quad l = \frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}}, \\
 &\quad \quad \quad j = k+1, k+2, \dots, n+P: \\
 &\quad \quad \quad \quad a_{ij}^{(k)} = a_{ij}^{(k-1)} - l a_{kj}^{(k-1)}; \\
 x_{ij} &= \frac{a_{i,j+n}^{(n)}}{a_{ii}^{(i-1)}}, \quad i=1,2,\dots,n, \quad j=1,2,\dots, P.
 \end{aligned}$$

Удобен метод Гаусса–Жордана тем, что после приведения матрицы к диагональному виду практически отпадает необходимость организации обратного хода.

Вычислительная сложность алгоритма Гаусса–Жордана решения P систем с различными правыми частями и одной и той же матрицей, оценивается (выкладки опустим) величиной $\frac{1}{2}n^3$ при больших n и небольших P . Если P тоже большое и сравнимо с n , то вычислительную сложность можно оценить величиной $\frac{3}{2}n^3$.

LU-разложение матриц

Пусть задана квадратная матрица A . Обозначим через Δ_j угловой (другое название – главный) минор порядка j матрицы A :

$$\Delta_1 = a_{11}, \quad \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \quad \dots, \quad \Delta_n = \det A.$$

LU-разложением матрицы A называется её представление в виде $A=LU$, где L – нижняя треугольная матрица с единичной диагональю, U – верхняя треугольная матрица с ненулевыми диагональными элементами.

Теорема 1 (теорема об LU-разложении). *Для того чтобы существовало LU-разложение матрицы A необходимо и достаточно отличие от нуля всех её угловых миноров: $\Delta_1 \neq 0, \Delta_2 \neq 0, \dots, \Delta_n \neq 0$. LU-разложение единственно.*

Теорема 1 (другая формулировка). *LU-разложение матрицы A существует тогда и только тогда, когда отличны от нуля все её угловые миноры (т.е. когда $\Delta_1 \neq 0, \Delta_2 \neq 0, \dots, \Delta_n \neq 0$). LU-разложение единственно.*

Доказательство. Пусть все угловые миноры матрицы A отличны от нуля, т.е. $\Delta_1 \Delta_2 \dots \Delta_n \neq 0$. Достаточность этих условий для существования треугольного разложения $A=LU$ докажем методом математической индукции.

Покажем сначала, что LU-разложение существует при $n=2$ (при $n=1$ разложение тоже существует, но треугольность матриц специфическая: $A=a_{11}=1 \cdot a_{11}=L \cdot U$). Существование треугольного разложения

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ l_{21} & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} \\ 0 & u_{22} \end{bmatrix}$$

сводится к существованию решения относительно неизвестных $l_{21}, u_{11}, u_{12}, u_{22}$ системы уравнений

$$\begin{aligned} u_{11} &= a_{11}, \\ u_{12} &= a_{12}, \\ l_{21}u_{11} &= a_{21}, \\ l_{21}u_{12} + u_{22} &= a_{22}. \end{aligned}$$

Эта система имеет единственное решение $u_{11}=a_{11}=\Delta_1, u_{12}=a_{12}, l_{21}=\frac{a_{21}}{a_{11}}=\frac{a_{21}}{\Delta_1}$,

$u_{22} = a_{22} - l_{21}u_{12} = \frac{a_{22}a_{11} - l_{21}u_{12}a_{11}}{a_{11}} = \frac{a_{22}a_{11} - a_{21}a_{12}}{a_{11}} = \frac{\Delta_2}{\Delta_1}$. Так как по предположению $\Delta_1 \neq 0, \Delta_2 \neq 0$, то верхняя треугольная матрица U имеет ненулевые диагональные элементы.

Пусть теперь A_k – матрица порядка k и разложение $A_{k-1}=L_{k-1}U_{k-1}$ с требуемыми свойствами существует для матриц порядка $k-1$. Докажем справедливость LU-разложения $A_k=L_kU_k$ и для матриц порядка k .

Будем искать представление $A_k = L_k U_k$ в следующем блочном виде:

$$\begin{bmatrix} A_{k-1} & a_{k-1}^{col} \\ a_{k-1}^{line} & a_{kk} \end{bmatrix} = \begin{bmatrix} L_{k-1} & 0 \\ l_{k-1} & 1 \end{bmatrix} \begin{bmatrix} U_{k-1} & u_{k-1} \\ 0 & u_{kk} \end{bmatrix},$$

где a_{k-1}^{line} , l_{k-1} – строки, a_{k-1}^{col} , u_{k-1} – столбцы. После выполнения произведения блочных матриц придем к системе

$$L_{k-1} U_{k-1} = A_{k-1},$$

$$L_{k-1} u_{k-1} = a_{k-1}^{col}$$

$$l_{k-1} U_{k-1} = a_{k-1}^{line},$$

$$l_{k-1} u_{k-1} + u_{kk} = a_{kk}.$$

Первое из равенств системы выполняется в силу индуктивного предположения. Из второго и третьего равенств, учитывая невырожденность матриц L_{k-1} и U_{k-1} , найдем u_{k-1} и l_{k-1} , после чего из четвертого равенства получим u_{kk} . При этом $u_{kk} \neq 0$, так как по правилам вычисления определителей $\Delta_k = \det A_k = \det L_{k-1} \det U_{k-1} u_{kk}$ и по предположению $\Delta_k \neq 0$. С учетом вида матриц L_{k-1} и U_{k-1} заключаем, что матрица L_k – нижняя треугольная матрица с единичной диагональю, U_k – верхняя треугольная матрица с ненулевыми диагональными элементами.

Таким образом, предположения теоремы являются достаточными для существования разложения $A = LU$.

Докажем теперь, что для существования LU-разложения матрицы отличие от нуля всех угловых миноров необходимо.

Пусть разложение $A = LU$ существует. Зафиксируем произвольное, но меньшее n , целое число m и представим матрицы в блочном виде

$$A = \begin{bmatrix} A_m & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad L = \begin{bmatrix} L_m & 0 \\ L_{21} & L_{22} \end{bmatrix}, \quad U = \begin{bmatrix} U_m & U_{12} \\ 0 & U_{22} \end{bmatrix},$$

где A_m , L_m , U_m – матрицы порядка m , матрицы L_m и L_{22} – нижние треугольные с единичной диагональю, матрицы U_m и U_{22} – верхние треугольные с ненулевыми диагональными элементами. Смысл выделения матрицы A_m состоит в том, что $\det A_m$ есть угловой минор Δ_m (отличие от нуля которого требуется показать).

Разложение $A = LU$ в блочном представлении имеет вид

$$\begin{bmatrix} A_m & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} L_m & 0 \\ L_{21} & L_{22} \end{bmatrix} \begin{bmatrix} U_m & U_{12} \\ 0 & U_{22} \end{bmatrix} = \begin{bmatrix} L_m U_m & L_m U_{12} \\ L_{21} U_m & L_{21} U_{12} + L_{22} U_{22} \end{bmatrix}.$$

Отсюда следует, что $A_m = L_m U_m$. По правилам вычисления определителей получим $\Delta_m = \det A_m = \det L_m \det U_m = u_{11} \dots u_{mm} \neq 0$, $m = 1, 2, \dots, n$.

Осталось установить единственность разложения. Пусть существуют два разложения: $A = L_1 U_1 = L_2 U_2$. Тогда, учитывая невырожденность сомножителей, получим сначала $L_1 = L_2 U_2 U_1^{-1}$, затем $L_2^{-1} L_1 = U_2 U_1^{-1}$. Матрица

в левой части последнего равенства является нижней треугольной (как произведения двух нижних треугольных), а в правой – верхней треугольной (как произведения двух верхних треугольных). Поэтому равенство возможно лишь в том случае, когда оба произведения являются диагональными матрицами. Так как на диагонали произведения $L_2^{-1}L_1$ стоят единицы, то оба произведения $L_2^{-1}L_1$ и $U_2U_1^{-1}$ равны единичной матрице. Следовательно, $L_1=L_2$, $U_1=U_2$. Теорема доказана.

Замечание 4. Так как $\Delta_m = u_{11} \dots u_{mm} \neq 0$, $m=1,2,\dots,n$, то

$$u_{11} = \Delta_1, \quad u_{mm} = \frac{\Delta_m}{\Delta_{m-1}}, \quad m = 2, \dots, n.$$

Получим соотношения, связывающие элементы матриц L , U и A .

Из равенства

$$\begin{bmatrix} 1 & 0 & \dots & 0 \\ l_{21} & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{n,1} & l_{n,2} & \dots & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & \dots & u_{1n} \\ 0 & u_{22} & \dots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & u_{nn} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}.$$

имеем

$$a_{ij} = \sum_{k=1}^n l_{ik} u_{kj} = \sum_{k=1}^{\min(i,j)} l_{ik} u_{kj}.$$

Если $i \leq j$, то

$$a_{ij} = \sum_{k=1}^i l_{ik} u_{kj} = \sum_{k=1}^{i-1} l_{ik} u_{kj} + u_{ij};$$

если $i > j$, то

$$a_{ij} = \sum_{k=1}^j l_{ik} u_{kj} = \sum_{k=1}^{j-1} l_{ik} u_{kj} + l_{ij} u_{jj}.$$

Таким образом,

$$u_{ij} = a_{ij} - \sum_{k=1}^{i-1} l_{ik} u_{kj}, \quad i = 1, \dots, n, \quad j = i, \dots, n, \quad (10)$$

$$l_{ij} = \frac{1}{u_{jj}} \left(a_{ij} - \sum_{k=1}^{j-1} l_{ik} u_{kj} \right), \quad j = 1, \dots, n-1, \quad i = j+1, \dots, n. \quad (11)$$

Соотношения (10), (11) пригодны для вычисления элементов матриц L и U , если чередовать вычисление строк матрицы U по формулам (10) и столбцов матрицы L по формулам (11). В этом случае неизвестные элементы матриц U и L выражаются через a_{ij} и уже найденные l_{ik} и u_{kj} .

Замечание 5. В рассмотренном варианте LU-разложения L – нижняя треугольная матрица с единичной диагональю, U – верхняя треугольная матрица с ненулевыми диагональными элементами. Часто рассматривают другой вариант: L – нижняя треугольная матрица с ненулевыми

диагональными элементами, U – верхняя треугольная матрица с единичной диагональю.

Связь метода Гаусса и LU-разложения

Обозначим

$$L^G = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ l_{21} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{n,1} & l_{n2} & \cdots & 1 \end{bmatrix},$$

где элементы l_{ik} вычисляются по формулам (6) $l_{ik} = \frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}}$. Через U^G обозначим матрицу системы (5):

$$U^G = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ 0 & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u_{nn} \end{bmatrix} = \begin{bmatrix} a_{11}^{(0)} & a_{12}^{(0)} & \cdots & a_{1n}^{(0)} \\ 0 & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn}^{(n-1)} \end{bmatrix}.$$

Матрицы L^G и U^G получены в результате выполнения прямого хода метода исключений.

Теорема 2. В результате применения метода Гаусса к системе (1) $Ax=b$ получено LU-разложение $A=L^G U^G$ матрицы A .

Доказательство. При вычислении величин $b_i^{(k)}$, $1 \leq k \leq n-1$, $k+1 \leq i \leq n$, по формулам (7)

$k = 1, 2, \dots, n-1$:

$i = k+1, k+2, \dots, n$:

$$b_i^{(k)} = b_i^{(k-1)} - l_{ik} b_k^{(k-1)}$$

сначала фиксируется k , а затем вычисляются $b_{k+1}^{(k)}, \dots, b_n^{(k)}$ (обновление величин $b_i^{(k)}$ в правой части системы уравнений на k -м шаге алгоритма).

Если сначала фиксировать i , $2 \leq i \leq n$, то вычисляются $b_i^{(1)}, \dots, b_i^{(i-1)}$ (процесс изменения величин $b_i^{(k)}$ в правой части i -го уравнения):

$i = 2, \dots, n$:

$k = 1, \dots, i-1$:

$$b_i^{(k)} = b_i^{(k-1)} - l_{ik} b_k^{(k-1)}. \quad (7^*)$$

Детальное понимание алгоритма прямого хода позволяет сделать вывод о корректности такого изменения порядка вычислений: изменение правой части фиксированного уравнения никак не зависят от изменений в уравнениях с большим номером.

Для фиксированного i вычисления $b_i^{(k)}$ по формулам (7*) при $k=i-1$ зависят от вычислений при $k=i-2$, вычисления при $k=i-2$ зависят от вычислений при $k=i-3, \dots$, вычисления при $k=2$ зависят от вычислений при $k=1$). Поэтому

$$\begin{aligned} b_i^{(i-1)} &= b_i^{(i-2)} - l_{i,i-1} b_{i-1}^{(i-2)} = b_i^{(i-3)} - l_{i,i-2} b_{i-2}^{(i-3)} - l_{i,i-1} b_{i-1}^{(i-2)} = \dots \\ &= b_i^{(0)} - l_{i,1} b_1^{(0)} - \dots - l_{i,i-2} b_{i-2}^{(i-3)} - l_{i,i-1} b_{i-1}^{(i-2)}. \end{aligned}$$

Отсюда следует

$$b_i^{(0)} = b_i^{(i-1)} + l_{i1} b_1^{(0)} + \dots + l_{i,i-2} b_{i-2}^{(i-3)} + l_{i,i-1} b_{i-1}^{(i-2)} = \sum_{k=1}^{i-1} l_{ik} b_k^{(i-1)} + b_i^{(i-1)}.$$

Обозначим правую часть системы (5)

$$\begin{aligned} &a_{11}^{(0)} x_1 + a_{12}^{(0)} x_2 + \dots + a_{1n}^{(0)} x_n = b_1^{(0)}, \\ &\phantom{a_{11}^{(0)}} a_{22}^{(1)} x_2 + \dots + a_{2n}^{(1)} x_n = b_2^{(1)}, \\ &\phantom{a_{11}^{(0)} a_{22}^{(1)}} \dots\dots\dots \\ &\phantom{a_{11}^{(0)} a_{22}^{(1)}} a_{nn}^{(n-1)} x_n = b_n^{(n-1)}. \end{aligned}$$

полученной после прямого хода гауссовых исключений, через вектор u :

$$y = \left(b_1^{(0)}, b_2^{(1)}, \dots, b_n^{(n-1)} \right).$$

Тогда

$$b_i^{(0)} = \sum_{k=1}^{i-1} l_{ik} y_k + y_i, \quad i=2, \dots, n.$$

Запишем эти равенства и систему (5) в матрично-векторном виде: $b=L^G y$, $U^G x=y$. Действуя на левую часть последнего равенства матрицей L^G , получим $L^G U^G x=b$.

Так как x есть решение уравнения $Ax=b$, то $L^G U^G A^{-1}b=b$ для любого b . Поэтому $L^G U^G A^{-1}$ – единичная матрица, $A=L^G U^G$. Теорема доказана.

Следствие. Метод Гаусса (без выбора главного элемента) можно применять только тогда, когда все угловые миноры матрицы системы отличны от нуля.

Действительно, LU-разложение $A=L^G U^G$ получено применением метода Гаусса. Если предположить, что какой-то угловой минор равен нулю, но применение метода Гаусса приводит к LU-разложению $A=L^G U^G$, то получим противоречие: для существования LU-разложения матрицы отличие от нуля всех угловых миноров есть необходимое и достаточное условие (теорема 1).

Замечание 6. Для применимости метода Гаусса отличие от нуля всех угловых миноров является не только необходимым, но и достаточным условием (доказательство не приводится).

Таким образом, мы показали, что применение метода Гаусса к системе $Ax=b$ допускает следующую реализацию:

- получение LU-разложения $A=LU$ по формулам прямого хода (6) и (8); система примет вид $LUx=b$, или $Ly=b$, где $y=Ux$;

- решение системы $Ly=b$ с нижней треугольной матрицей (заменяет преобразование по формулам (7) вектора правой части системы уравнений);
- решение системы $Ux=y$ с верхней треугольной матрицей по формулам обратного хода (9) с учетом обозначений $u_{ij} = a_{ij}^{(i-1)}$, $y_i = b_i^{(i-1)}$.

Получение LU-разложения по формулам (6) и (8) требует порядка $\frac{1}{3}n^3$ операций (умножения и деления), решение систем $Ly=b$ и $Ux=y$ требует порядка $\frac{1}{2}n^2$ операций. Если требуется решить P систем с различными правыми частями и одной и той же матрицей, то при больших n и небольших P вычислительная сложность оценивается величиной $\frac{1}{3}n^3$. Если P сравнимо с n , то вычислительную сложность можно оценить величиной $\frac{1}{3}n^3 + Pn^2$. При любом числе правых частей LU-разложение осуществляется только один раз.

Распишем в явном виде формулы вида (6) и (8)

$$a_{ij}^{(0)} = a_{ij}, \quad i=1,2,\dots,n, \quad j=1,2,\dots,n;$$

$$k = 1, 2, \dots, n-1:$$

$$i = k+1, k+2, \dots, n:$$

$$l_{ik} = \frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}},$$

$$j = k+1, k+2, \dots, n:$$

$$a_{ij}^{(k)} = a_{ij}^{(k-1)} - l_{ik} a_{kj}^{(k-1)}$$

для получения LU-разложения:

$$a_{ij}^{(0)} = a_{ij}, \quad i=1,2,\dots,n, \quad j=1,2,\dots,n;$$

$$k = 1, 2, \dots, n-1:$$

$$j = k, \dots, n:$$

$$u_{kj} = a_{kj}^{(k-1)},$$

$$i = k+1, k+2, \dots, n:$$

$$l_{ik} = \frac{a_{ik}^{(k-1)}}{u_{kk}},$$

$$j = k+1, \dots, n:$$

$$a_{ij}^{(k)} = a_{ij}^{(k-1)} - l_{ik} u_{kj}$$

$$u_{nn} = a_{nn}^{(n-1)}.$$

Приведем ещё для получения LU-разложения так называемую компактную схему Гаусса. Компактность схемы означает использование при программной реализации только памяти, первоначально предназначенной для хранения исходной матрицы a_{ij} . Будем вместо обозначения l_{ik}

использовать обозначение $a_{ik}^{(k)}$ (ранее при $i > k$ это обозначение не использовалось); при программной реализации будем использовать ячейку памяти a_{ik} для хранения l_{ik} . Учитывая

$$l_{ik} = a_{ik}^{(k)}, \quad u_{kj} = a_{kj}^{(k-1)},$$

алгоритм (12) примет вид (существенными являются нижние индексы, а верхний индекс оставлен для наглядности, при программировании он не принимается во внимание)

$$a_{ij}^{(0)} = a_{ij}, \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, n;$$

$$k = 1, 2, \dots, n-1:$$

$$i = k+1, k+2, \dots, n:$$

$$a_{ik}^{(k)} = \frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}}, \quad (13)$$

$$j = k+1, k+2, \dots, n:$$

$$a_{ij}^{(k)} = a_{ij}^{(k-1)} - a_{ik}^{(k)} a_{kj}^{(k-1)}.$$

Алгоритмом (13) осуществляются преобразования прямого хода метода Гаусса, но на место нулевых элементов (ниже ведущего элемента) на каждом k -м шаге записываются величины $\frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}}$ (на которые умножается k -е уравнение перед вычитанием его из i -го уравнения), которые и есть элементы l_{ik} матрицы L . После завершения вычислений на месте элементов матрицы A расположены соответствующие элементы матриц L и U .

Метод Гаусса с выбором главного элемента. Об устойчивости метода Гаусса

Рассмотренный метод исключений Гаусса имеет недостатки, которые ограничивают использование метода на практике:

- неустойчивость алгоритма: в случае относительной малости главного элемента процесс вычислений приводит к сильному накоплению погрешностей;
- во многих случаях заранее неизвестно, все ли угловые миноры матрицы системы отличны от нуля и можно ли применять метод; при этом система может иметь единственное решение, даже если какой-либо из Δ_k равен нулю.

Преодолеть эти недостатки в значительной мере можно путем включения в алгоритм метода исключений некоторой схемы выбора главного элемента. Выбор осуществляется среди некоторого множества элементов преобразуемой матрицы. В качестве главного выбирается наибольший по модулю элемент. Укажем три основных варианта выбора главного элемента.

1. Метод Гаусса с выбором главного элемента по строке. Он эквивалентен применению обычного метода к системе, в которой на каждом шаге исключения проводится перенумерация переменных.

2. Метод Гаусса с выбором главного элемента по столбцу. Он эквивалентен применению обычного метода к системе, в которой на каждом шаге исключения проводится перенумерация уравнений.

3. Метод Гаусса с выбором главного элемента по всей подматрице. Он эквивалентен применению обычного метода к системе, в которой на каждом шаге исключения проводится соответствующая перенумерация и переменных, и уравнений.

Варианты 1 и 2 требуют порядка n^2 сравнений при поиске максимального по модулю элемента, вариант 3 требует порядка n^3 сравнений.

Устойчивость метода Гаусса исследуем с помощью эквивалентных возмущений. Известно, что решение системы (1) $Ax=b$ порядка n , вычисленное на компьютере методом Гаусса, в точности удовлетворяет «возмущенной» системе

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

причем для матрицы возмущения δA справедлива оценка

$$\|\delta A\|_{\infty} \leq f(n)2^{-t}\|A\|_{\infty}g(A).$$

Здесь $f(n)$ – некоторая медленно растущая функция (типа степени с небольшим показателем), t – число разрядов, отведенных для представления мантиссы, $g(A)$ – величина, оценивающая рост элементов промежуточных матриц по сравнению с элементами исходной матрицы A :

$$g(A) = \max_{0 \leq k \leq n-1} a_k / a_0, \\ a_0 = \max_{1 \leq i, j \leq n} |a_{ij}|, \quad a_k = \max_{i, j > k} |a_{ij}^{(k)}|.$$

Таким образом, ошибки δA , вносимые методом Гаусса, тем больше, чем больший рост элементов матриц допускается при прямом ходе. Согласно формуле (8)

$$|a_{ij}^{(k)}| = \left| a_{ij}^{(k-1)} - \frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}} a_{kj}^{(k-1)} \right| \leq |a_{ij}^{(k-1)}| + \left| \frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}} \right| |a_{kj}^{(k-1)}|.$$

Поэтому в схемах без выбора главного элемента рост элементов может быть сколь угодно велик, а в схемах с выбором он ограничен (за шаг максимальный модуль элемента матрицы может вырасти самое большое в 2 раза), метод Гаусса становится более устойчивым (т.е. менее чувствительным к погрешностям вычислений).

Из приведенных рассуждений и оценки относительной погрешности решения (формула (10) файла «Нормы_Обусловленность»)

$$\frac{\|\delta x\|}{\|x\|} \leq \frac{\|A\| \|A^{-1}\| \|\delta A\|}{1 - \|A\| \|A^{-1}\| \|\delta A\|} \frac{\|\delta A\|}{\|A\|}$$

следует, что схемы выбора главного элемента не гарантируют получения методом Гаусса близкого к точному решения в случае плохой

обусловленности матрицы, но могут в значительной степени уменьшить влияние вычислительной погрешности на отклонение полученного результата от настоящего решения задачи.

Наиболее используемым на практике является вариант с выбором главного элемента по столбцу, так как он не требует для своей реализации дополнительного вектора, в котором хранится информация о выполненных переименованиях неизвестных. Вариант с выбором главного элемента по всей подматрице приводит к меньшему, чем другие два варианта, накоплению погрешностей, но требует гораздо больших затрат на поиск максимального по модулю элемента.

Запишем модификацию алгоритма прямого хода (6)–(8) для случая выбора главного элемента по столбцу (в комментариях даны некоторые пояснения):

$$a_{ij}^{(0)} = a_{ij}, \quad i=1,2,\dots,n, \quad j=1,2,\dots,n;$$

$$b_i^{(0)} = b_i, \quad i=1,2,\dots,n;$$

$$k = 1, 2, \dots, n-1:$$

$$i = k, k+1, \dots, n:$$

$$c_i = a_{ik}^{(k-1)}, \quad // \text{ выделяется } k\text{-й столбец подматрицы}$$

Выбор наибольшего по модулю элемента в массиве c , перестановка строки с номером k и строки, содержащей наибольший по модулю элемент // ведущим элементом становится элемент c_k

$$i = k+1, k+2, \dots, n:$$

$$l_i = \frac{c_i}{c_k}, \quad // \text{ создается временный коэффициент } l_i = \frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}} \text{ путем}$$

деления i -го элемента k -го столбца матрицы на ведущий элемент

$$b_i^{(k)} = b_i^{(k-1)} - l_i b_k^{(k-1)},$$

$$j = k+1, k+2, \dots, n:$$

$$a_{ij}^{(k)} = a_{ij}^{(k-1)} - l_i a_{kj}^{(k-1)}.$$

Отметим, что этот алгоритм не предназначен для получения LU-разложения: осуществлялась перестановка строк, поэтому для получения LU-разложения требуется некоторая модификация этого алгоритма.

Вычисление определителей и обращение матриц

Применение к системе с матрицей A метода Гаусса без выбора главного элемента приводит к получению LU-разложения $A=L^G U^G$ (теорема 2). Поэтому

$$\det A = \det L^G \det U^G = u_{11} \dots u_{nn} = a_{11}^{(0)} \dots a_{nn}^{(n-1)},$$

т.е. определитель матрицы A равен произведению диагональных элементов матрицы U^G и равен произведению всех главных элементов.

Замечание 7. Для вычисления определителя матрицы A не требуется получение LU-разложения $A=L^G U^G$, требуется получить только матрицу U^G .

Если к системе с матрицей A применить метод Гаусса с выбором главного элемента (любой из вариантов выбора), то

$$\det A = (-1)^m a_{11}^{(0)} \dots a_{nn}^{(n-1)},$$

где m – количество перестановок, осуществляемых в процессе исключения (другие преобразования, применяемые в процессе исключения, не изменяют определитель матрицы).

Задача нахождения матрицы, обратной матрице A , эквивалентна задаче решения матричного уравнения $AX=E$, где X – искомая матрица (если $AX=E$, то $X=A^{-1}E=A^{-1}$). Матричное уравнение $AX=E$ с матрицей порядка n можно рассматривать как P систем с различными правыми частями и одной и той же матрицей, причем $P=n$, а правые части есть столбцы единичной матрицы.

Приведём один из вариантов алгоритма Гаусса–Жордана получения обратной матрицы (см. замечание 3):

$$a_{ij}^{(0)} = a_{ij}, \quad i=1,2,\dots,n, \quad j=1,2,\dots,n;$$

$$i = 1, 2, \dots, n:$$

$$j = 1, 2, \dots, n:$$

$$\text{if } j = i \text{ then } a_{i,j+n}^{(0)} = 1 \text{ else } a_{i,j+n}^{(0)} = 0;$$

$$k = 1, 2, \dots, n:$$

$$i = 1, 2, \dots, n:$$

$$\text{if } i \neq k:$$

$$l = \frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}},$$

$$j = k+1, k+2, \dots, 2n:$$

$$a_{ij}^{(k)} = a_{ij}^{(k-1)} - l a_{kj}^{(k-1)};$$

$$i = 1, 2, \dots, n:$$

$$j = 1, 2, \dots, n:$$

$$x_{ij} = \frac{a_{i,j+n}^{(n)}}{a_{ii}^{(i-1)}}.$$

Ранее отмечалось, что вычислительная сложность решения P систем с одной и той же матрицей при больших n и P , сравнимых с n , можно оценить величиной $\frac{4}{3}n^3$ при решении методом Гаусса и величиной $\frac{3}{2}n^3$ при решении методом Гаусса–Жордана. Можно воспользоваться спецификой правой части уравнения $AX=E$ (наличием нулевых компонент) и получить модификацию алгоритма обращения матриц, основанного на LU-разложении, с вычислительной сложностью n^3 : Упрощение вычислений получается потому, что при решении систем вида $Ly=b$ (один из этапов подхода, основанного на LU-разложении) некоторые компоненты вектора y вычислять не надо, так как они заведомо нулевые.

Список использованных источников

1. Самарский А.А., Гулин А.В. Численные методы. М: Наука. 1989. 432 с.
2. Репников В.И. Вычислительные методы алгебры. Курс лекций. Минск. Белгосуниверситет. Кафедра вычислительной математики. 2011.
3. Андреев В.Б. Численные методы. М: МГУ, 2013. 336 с.
4. Фалейчик Б.В. Методы вычислений: пособие. Минск: БГУ, 2014. 224 с.

Вопрос коллоквиума: Теорема об LU-разложении. Этапы решения СЛАУ на основе LU-разложения.

Ответ:

LU-разложением матрицы A называется её представление в виде $A=LU$, где L – нижняя треугольная матрица с единичной диагональю, U – верхняя треугольная матрица с ненулевыми диагональными элементами.

Теорема об LU-разложении. *Для того чтобы существовало LU-разложение матрицы необходимо и достаточно отличие от нуля всех её угловых миноров. LU-разложение единственно.*

Этапы решения системы $Ax=b$ на основе LU-разложения (фактически это вариант метода Гаусса):

- получение LU-разложения $A=LU$ по формулам прямого хода метода Гаусса; система примет вид $LUx=b$, или $Ly=b$, где $y=Ux$;
- решение системы $Ly=b$ с нижней треугольной матрицей;
- решение системы $Ux=y$ с верхней треугольной матрицей.

Вопрос к лабораторной работе «Метод Гаусса»: Этапы метода Гаусса решения СЛАУ. Вычислительная сложность метода. Об устойчивости метода.

Ответ:

Этапы решения системы $Ax=b$ методом Гаусса:

- выполнение прямого хода (приведение системы к треугольному виду);
- выполнение обратного хода (решение системы с верхней треугольной матрицей).

Вычислительная сложность метода Гаусса – величина порядка $\frac{1}{3}n^3$.

Алгоритм без выбора ведущего элемента неустойчив: в случае относительной малости главного элемента процесс вычислений приводит к сильному накоплению погрешностей. Схемы выбора главного элемента не гарантируют устойчивость (в случае плохой обусловленности матрицы), но могут в значительной степени уменьшить влияние вычислительной погрешности.