



Операции над матрицами

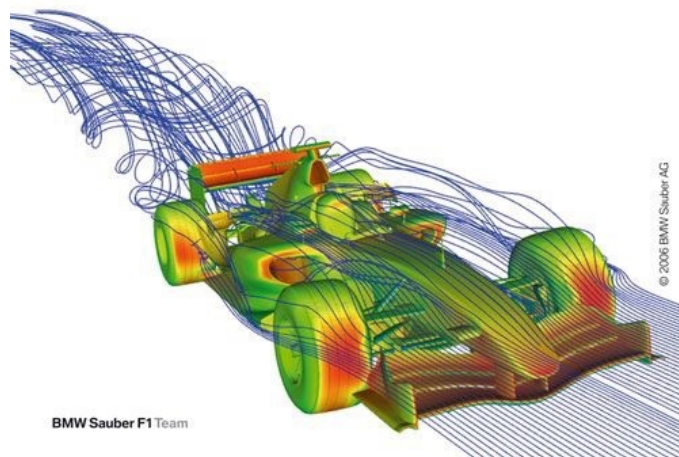
1 Введение

Ранее мы начали рассматривать матрицы как инструмент для решения линейных систем, но с матрицами можно работать как с самостоятельными объектами. Например, над ними можно определить простые арифметические операции. У результатов таких операций есть геометрический смысл, но о нём мы поговорим немного позднее.

Операции над матрицами, рассматриваемые в данном разделе, помимо всего прочего, служат основой для получения матричных разложений. Такие разложения на практике позволяют делать процесс последовательного поиска решений большого количества систем более быстрым. Это, например, очень актуально в вычислительной гидродинамике (Computational Fluid Dynamics, CFD). Данная дисциплина изучает поведение потоков воздуха/газа/жидкости. Она помогает смоделировать явления, возникающие при обтекании потоками различных твёрдых тел, а также оценить соответствующие температуру, скорость и давление.

Гидродинамическое моделирование помогает учёным исследовать цунами и потоки лавы. В последнее время всё больше и больше средств вкладываются в развитие CFD-лабораторий. Без таких технологий трудно представить современные аэро- и космическую промышленность.

Возможности, которые предоставляет вычислительная гидродинамика, используются и в Формуле 1. CFD позволяет просчитать аэродинамическую эффективность всего болида и каждого элемента по отдельности без необходимости изготавливать реальные детали. Существующие методы моделирования позволяют обрабатывать большой объём данных, описывающих характеристики поверхности машины. По сути этот процесс эквивалентен решению миллионов линейных систем вида $Ax = b$ с постоянно меняющейся правой частью b . Для решения таких задач на практике даже современным мощным суперкомпьютерам могут потребоваться дни и недели непрерывных расчётов. Метод LU -разложения матрицы, о котором пойдёт речь в этой главе, позволяет сделать этот процесс более эффективным.



2 Арифметические операции с матрицами

Определение 1

Две матрицы $A = (a_{ij})$ и $B = (b_{ij})$ называются **равными**, если они имеют один и тот же размер $m \times n$ и состоят из одних элементов, то есть $a_{ij} = b_{ij}, \forall i = \overline{1, m}, \forall j = \overline{1, n}$. Обозначение: $A = B$.

Определение 2

Матрица $-A = (-a_{ij})$ называется **противоположной** к матрице $A = (a_{ij})$.

Матрицы одного размера можно складывать и вычитать. Множество матриц размера $m \times n$ обозначим M_{mn} .

Определение 3

Суммой матриц $A = (a_{ij}) \in M_{mn}$ и $B = (b_{ij}) \in M_{mn}$ называется матрица $C = (c_{ij}) \in M_{mn}$, элементы которой определены равенством $c_{ij} = a_{ij} + b_{ij}, \forall i = \overline{1, m}, \forall j = \overline{1, n}$. Обозначение: $C = A + B$.

$$C = A + B = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn} \end{pmatrix}$$

Пример 1

$$A = \begin{pmatrix} 0 & 2 & 3 \\ 4 & -1 & 1 \end{pmatrix} \in M_{23}, B = \begin{pmatrix} 1 & 4 & -2 \\ 0 & 2 & 5 \end{pmatrix} \in M_{23},$$

$$A + B = \begin{pmatrix} 0+1 & 2+4 & 3-2 \\ 4+0 & -1+2 & 1+5 \end{pmatrix} = \begin{pmatrix} 1 & 6 & 1 \\ 4 & 1 & 6 \end{pmatrix} \in M_{23}.$$

Определение 4

Матрица $O \in M_{mn}$, у которой все элементы равны нулю, называется **нулевой**.

Важно отметить, что нулевая матрица играет такую же роль при выполнении действий над матрицами, как число 0 при соответствующих арифметических операциях.

Теорема 1

Операция сложения матриц обладает следующими свойствами:

1. $A + B = B + A$ (сложение матриц коммутативно);
2. $(A + B) + C = A + (B + C)$ (сложение матриц ассоциативно);
3. $A + O = O + A = A$;
4. $A + (-A) = -A + A = O$.

Определение 5

Произведением матрицы $A = (a_{ij}) \in M_{mn}$ на скаляр (число) λ называется матрица $C = (c_{ij}) \in M_{mn}$, элементы которой определены равенством $c_{ij} = \lambda a_{ij}, \forall i = \overline{1, m}, \forall j = \overline{1, n}$. Обозначение: $C = \lambda A$.

$$C = \lambda A = \begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \cdots & \lambda a_{1n} \\ \lambda a_{21} & \lambda a_{22} & \cdots & \lambda a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda a_{m1} & \lambda a_{m2} & \cdots & \lambda a_{mn} \end{pmatrix}$$

Пример 2

$$A = \begin{pmatrix} 3 & -5 \\ 0 & 1 \\ 2 & 7 \end{pmatrix} \in M_{32}, \quad \lambda = 4,$$
$$\lambda A = \begin{pmatrix} 4 \cdot 3 & 4 \cdot (-5) \\ 4 \cdot 0 & 4 \cdot 1 \\ 4 \cdot 2 & 4 \cdot 7 \end{pmatrix} = \begin{pmatrix} 12 & -20 \\ 0 & 4 \\ 8 & 28 \end{pmatrix} \in M_{32}.$$

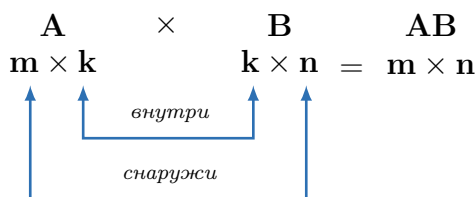
Теорема 2

Операция умножения матрицы на число обладает следующими свойствами:

1. $1 \cdot A = A$;
2. $(\alpha\beta)A = \alpha(\beta A)$;
3. $\alpha(A + B) = \alpha A + \alpha B$ (умножение матрицы на число дистрибутивно относительно сложения матриц);
4. $(\alpha + \beta)A = \alpha A + \beta A$ (умножение матрицы на число дистрибутивно относительно сложения чисел);
5. $-A = (-1)A$.

Определение 6

Произведением матриц $A = (a_{ij}) \in M_{mk}$ и $B = (b_{ij}) \in M_{kn}$ называется матрица $C = (c_{ij}) \in M_{mn}$, элементы которой определены равенством $c_{ij} = \sum_{s=1}^k a_{is}b_{sj}$, $i = \overline{1, m}, j = \overline{1, n}$. Обозначение: $C = AB$.



Произведение AB определено лишь в том случае, когда размеры матриц A и B согласованы специальным образом: количество столбцов матрицы A должно совпадать с количеством строк матрицы B .

Пример 3

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

$$\underset{2 \times 2}{A} \cdot \underset{2 \times 3}{B} = \underset{2 \times 3}{C} = \left(\begin{array}{c|c|c} 0 \cdot 1 + 2 \cdot (-1) & 0 \cdot 3 + 2 \cdot 2 & 0 \cdot (-2) + 2 \cdot 0 \\ \hline 4 \cdot 1 + 6 \cdot (-1) & 4 \cdot 3 + 6 \cdot 2 & 4 \cdot (-2) + 6 \cdot 0 \end{array} \right) = \begin{pmatrix} -2 & 4 & 0 \\ -2 & 24 & -8 \end{pmatrix}.$$

Обрати внимание, для приведенных выше матриц не получится найти произведение BA , так как количество столбцов матрицы B не совпадает с числом строк матрицы A .

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

$\mathbf{2 \times 3} \qquad \mathbf{2 \times 2}$

Замечание. Произведение матриц зависит от порядка сомножителей, то есть в общем случае $AB \neq BA$ (даже, если определены и AB , и BA). Например,

$$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \neq \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

Определение 7

Матрицы A и B , для которых $AB = BA$, называются **перестановочными** или **коммутирующими**.

Таким образом, когда мы говорим об умножении, всегда нужно уточнять, с какой стороны мы умножаем матрицу A на B — справа (AB) или слева (BA).

Теорема 3

Операция умножения матриц обладает следующими свойствами:

1. $\forall A \in M_{mk}, \forall B \in M_{kl}, \forall C \in M_{ln}$ верно, что $(AB)C = A(BC)$ (свойство ассоциативности);
2. $\lambda(AB) = (\lambda A)B = A(\lambda B), \forall \lambda \in \mathbb{R}, \forall A \in M_{mk}, \forall B \in M_{kn}$;
3. $\forall A \in M_{mk}, \forall B, C \in M_{kn}$ выполняется $A(B + C) = AB + AC$ (свойство дистрибутивности).

3 Транспонирование матрицы

Помимо арифметических действий из предыдущего раздела для матриц можно определить операции, у которых нет аналогов на множестве действительных чисел. Одна из таких операций — транспонирование.

Определение 8

Пусть $A = (a_{ij}) \in M_{mn}$. Матрица $A^T = (a_{ij}^t) \in M_{nm}$ называется **транспонированной** к матрице A , если $a_{ij}^t = a_{ji}, i = \overline{1, n}, j = \overline{1, m}$.

Обозначение: A^T .

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \quad A^T = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}.$$

По сути матрица A^T получается из A путём замены каждой её строки на столбец с тем же номером. Другими словами, транспонирование — это симметричное отображение относительно главной диагонали.

Пример 4

$$B = \begin{pmatrix} 1 & 3 & -2 \\ -1 & 2 & 0 \end{pmatrix}, \quad B^T = \begin{pmatrix} 1 & -1 \\ 3 & 2 \\ -2 & 0 \end{pmatrix}.$$

Теорема 4

Операция транспонирования обладает следующими свойствами:

1. $(A + B)^T = A^T + B^T$;
2. $(\lambda A)^T = \lambda A^T, \quad \forall \lambda \in \mathbb{R}$;
3. $(A^T)^T = A$;
4. $(AB)^T = B^T A^T$.

Первые 3 свойства очевидны. Последнее, на первый взгляд, может показаться необычным. Приведём его доказательство и покажем, что оно является совершенно естественным.

Доказательство

Пусть $A = (a_{ij}) \in M_{mk}$, $B = (b_{ij}) \in M_{kn}$, произведение AB имеет смысл. Произведение $B^T A^T$ также имеет смысл, при этом размеры матриц $(AB)^T$ и $B^T A^T$ совпадают. Элементы этих матриц, стоящие в одинаковых позициях, равны, поскольку

$$\{(AB)^T\}_{ij} = \{AB\}_{ji} = \sum_{s=1}^k a_{js} b_{si} = \sum_{s=1}^k b_{si} a_{js} = \sum_{s=1}^k b_{is}^t a_{sj}^t = \{B^T A^T\}_{ij}.$$

Определение 9

Матрица называется **симметричной**, если $A = A^T$.

Пример 5

$$A = A^T = \begin{pmatrix} 3 & 2 \\ 2 & 3 \end{pmatrix}$$

4 Обратные матрицы

На множестве вещественных чисел для любого ненулевого числа можно найти ему обратное, то есть для числа a всегда найдется a^{-1} такое, что $a \cdot a^{-1} = a^{-1} \cdot a = 1$. Есть ли аналог такого обратного элемента на множестве матриц? Прежде, чем искать его, давай обсудим, какой математический объект будет аналогом 1 на множестве матриц.

Определение 10

Единичной матрицей называется квадратная матрица, у которой все элементы на главной диагонали равны 1, а все остальные равны 0.

Существуют разные обозначения для единичной матрицы. В курсе будет использоваться общепринятое — I (от слова identity). Если важно отметить размер матрицы, например, $n \times n$, будем обозначать её как I_n .

$$I = I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

Единичные матрицы на множестве матриц играют ту же роль, что 1 играет на множестве вещественных чисел. А именно, легко увидеть, что если умножить произвольную матрицу A на единичную матрицу совместимого размера, то в результате получим ту же матрицу A . Или более формально, если $A \in M_{mn}$, то $A \cdot I_n = I_m \cdot A = A$.

Теперь мы готовы определить обратную матрицу.

Определение 11

Матрица A^{-1} называется **обратной** к матрице A , если $AA^{-1} = A^{-1}A = I$. Матрица A , для которой существует обратная матрица, называется **обратимой**.

Не для любого числа существует обратное число. Так же и с матрицами: не для любой можно найти обратную матрицу. Если A^{-1} не существует, то A называется **необратимой**.

Замечание. Обратная матрица определена только для квадратных матриц.

Это объясняется тем, что равенство $AA^{-1} = A^{-1}A$ возможно лишь для квадратных матриц A и A^{-1} одинакового размера.

Но не любая квадратная матрица $A \neq O$ обратима. Так, матрица $A = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$ при умножении справа на любую матрицу даёт матрицу с нулевой первой строкой, то есть ни для какой матрицы B произведение AB не может совпадать с единичной матрицей I .

Замечание. Обозначение A^{-1} нельзя интерпретировать как $1/A$. Деление на матрицы в том же виде, в котором мы производим такие операции для вещественных чисел, не определено.

Лемма 1

Для обратной матрицы верны следующие свойства:

1. $(AB)^{-1} = B^{-1}A^{-1}$;
2. $(A^T)^{-1} = (A^{-1})^T$;
3. $(A^{-1})^{-1} = A$.

Доказательство

1. Проверим, что $B^{-1}A^{-1}$ является обратной матрицей для AB по определению.

$$\begin{aligned}(B^{-1}A^{-1})AB &= B^{-1}(A^{-1}A)B = B^{-1}IB = B^{-1}B = I; \\ AB(B^{-1}A^{-1}) &= A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I.\end{aligned}$$

- 2.

$$\begin{aligned}(A^{-1})^T A^T &= (AA^{-1})^T = I; \\ A^T (A^{-1})^T &= (A^{-1}A)^T = I.\end{aligned}$$

3. Свойство выполняется, так как $AA^{-1} = A^{-1}A = I$.



Теорема 5

Если матрица $A \in M_{nn}$ обратима, то система уравнений $Ax = b$ имеет единственное решение $x = A^{-1}b$. Здесь $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ — столбец переменных, $b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$ — столбец свободных членов.

Получается, для того, чтобы решить систему линейных уравнений, достаточно знать чему равна обратная матрица. Как её найти?

Допустим, что матрица A обратима. Наша задача — найти матрицу $B = A^{-1}$ такую, что $AB = C = I$.

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}, \quad A^{-1} = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{pmatrix}.$$

По правилам произведения матриц, например, для элементов первого столбца матрицы C должно быть выполнено следующее:

$$\begin{aligned}c_{11} &= a_{11}b_{11} + a_{12}b_{21} + \cdots + a_{1n}b_{n1} = 1 \\ c_{21} &= a_{21}b_{11} + a_{22}b_{21} + \cdots + a_{2n}b_{n1} = 0 \\ &\vdots \\ c_{n1} &= a_{n1}b_{11} + a_{n2}b_{21} + \cdots + a_{nn}b_{n1} = 0\end{aligned}$$

Чтобы найти первый столбец A^{-1} , а именно b_{11}, \dots, b_{n1} , необходимо решить систему уравнений с расширенной матрицей.

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & 1 \\ a_{21} & a_{22} & \dots & a_{2n} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} & 0 \end{array} \right)$$

Аналогично, для второго столбца нужно решить систему:

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & 0 \\ a_{21} & a_{22} & \dots & a_{2n} & 1 \\ a_{31} & a_{32} & \dots & a_{3n} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} & 0 \end{array} \right)$$

Всего будет n систем. При этом, вместо того, чтобы решать эти системы по отдельности и делать двойную работу, используя одни и те же преобразования для приведения матриц к ступенчатому виду, мы можем решить их одновременно. Другими словами, мы можем переформулировать задачу и записать её в виде:

$$(A|I) = \left(\begin{array}{cccc|cccc} a_{11} & a_{12} & \dots & a_{1n} & 1 & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2n} & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} & 0 & 0 & \dots & 1 \end{array} \right)$$

Для всей матрицы целиком можно применить метод Гаусса-Жордана, который включает в себя обратный ход метода Гаусса, таким образом, что матрица A слева преобразуется в единичную матрицу I . Тем временем, справа окажется матрица A^{-1} .

Замечание. Таким образом, получилось преобразование $(A|I) \Rightarrow (I|A^{-1})$

По теореме 5 решение для $(A|I)$ существует и единственно, в случае если A обратима.

Пример 6

$$A = \begin{pmatrix} 1 & -1 & 0 \\ 5 & -1 & 3 \\ 0 & 1 & 1 \end{pmatrix}, \quad A^{-1} = ?$$

$$\begin{aligned} \left(\begin{array}{ccc|ccc} 1 & -1 & 0 & 1 & 0 & 0 \\ 5 & -1 & 3 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{array} \right) &\sim \left(\begin{array}{ccc|ccc} 1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 4 & 3 & -5 & 1 & 0 \end{array} \right) \sim \left(\begin{array}{ccc|ccc} 1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & -1 & -5 & 1 & -4 \end{array} \right) \sim \\ \left(\begin{array}{ccc|ccc} 1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 5 & -1 & 4 \end{array} \right) &\sim \left(\begin{array}{ccc|ccc} 1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -5 & 1 & -3 \\ 0 & 0 & 1 & 5 & -1 & 4 \end{array} \right) \sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -4 & 1 & -3 \\ 0 & 1 & 0 & -5 & 1 & -3 \\ 0 & 0 & 1 & 5 & -1 & 4 \end{array} \right) \\ A^{-1} &= \begin{pmatrix} -4 & 1 & -3 \\ -5 & 1 & -3 \\ 5 & -1 & 4 \end{pmatrix}. \end{aligned}$$

Пример 7

Необходимо решить систему уравнений.

$$\begin{cases} x - y = 1 \\ 5x - y + 3z = 2 \\ y + z = 3 \end{cases}$$

Казалось бы, можно долго не рассуждать и сразу использовать метод Гаусса. Но мы посмотрим на задачу под другим углом. Перепишем систему в матричном виде:

$$\begin{pmatrix} 1 & -1 & 0 \\ 5 & -1 & 3 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

С помощью теоремы 5 находим решение:

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 \\ 5 & -1 & 3 \\ 0 & 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} -4 & 1 & -3 \\ -5 & 1 & -3 \\ 5 & -1 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} -11 \\ -12 \\ 15 \end{pmatrix}.$$

5 Матричные уравнения

Когда мы искали обратную матрицу, мы по сути решали уравнение вида $AX = I$, где A , I и неизвестная X являлись матрицами размера $n \times n$. Такое уравнение — частный случай матричного уравнения вида

$$AX = B.$$

Это аналог обычного числового уравнения $ax = b$. Только в случае чисел мы можем найти x как $\frac{b}{a}$, а для матриц операция деления не определена.

Чтобы решить уравнение вида $AX = B$, необходимо домножить обе части уравнения на A^{-1} **слева**. Тогда $A^{-1}AX = A^{-1}B$, значит $X = A^{-1}B$.

Аналогично для решения уравнения $XA = B$ нужно домножить — на A^{-1} **справа**. Отсюда $X = BA^{-1}$.

Пример 8

$$\underbrace{\begin{pmatrix} 1 & -1 & 0 \\ 5 & -1 & 3 \\ 0 & 1 & 1 \end{pmatrix}}_{A \in 3 \times 3} X = \underbrace{\begin{pmatrix} 1 & 0 \\ 2 & 1 \\ 3 & -1 \end{pmatrix}}_{B \in 3 \times 2}$$

Для совместимости по умножению матрица X должна иметь размер 3×2 .

$$X = A^{-1}B = \begin{pmatrix} -4 & 1 & -3 \\ -5 & 1 & -3 \\ 5 & -1 & 4 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 2 & 1 \\ 3 & -1 \end{pmatrix} = \begin{pmatrix} -11 & 4 \\ -12 & 4 \\ 15 & -5 \end{pmatrix}$$

Пример 12

$$T_{12} = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \sim (1) \leftrightarrow (1) + 2 \cdot (3),$$
$$T_{12}A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = \begin{pmatrix} 15 & 18 & 21 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}.$$

Замечание. Аналогично можно показать, что умножение матрицы A на матрицы элементарных преобразований P_{ij} , D_i , T_{ij}^T **справа** эквивалентно элементарным преобразованиям соответствующих столбцов матрицы A .

Следующие две теоремы являются одними из основополагающих для многих результатов, которые будут обсуждаться в дальнейшем.

Теорема 6

Любая матрица элементарных преобразований обратима, причём обратная тоже является матрицей элементарных преобразований.

Теорема 7

Матрица $A \in M_{nn}$ обратима тогда и только тогда, когда она представима в виде произведения матриц элементарных преобразований.

7 LU-разложение

При решении разного рода задач на компьютере важными составляющими являются скорость и точность вычислений, а также объём используемой памяти. С этой точки зрения, для некоторых алгоритмов, которые мы обсуждаем в этом курсе, есть более эффективные и устойчивые альтернативы. Так, альтернативой методу Гаусса для решения линейных систем и поиска обратной матрицы при выполнении некоторых предпосылок является LU-разложение.

Определение 12

LU-разложением квадратной матрицы $A \in M_{nn}$ называется представление матрицы в виде произведения $A = LU$, где $L = (l_{ij})$ — нижнетреугольная матрица, а $U = (u_{ij})$ — верхнетреугольная матрица с единицами на главной диагонали.

$$A = LU = \begin{pmatrix} * & 0 & 0 \\ * & * & 0 \\ * & * & * \end{pmatrix} \begin{pmatrix} 1 & * & * \\ 0 & 1 & * \\ 0 & 0 & 1 \end{pmatrix}.$$

Сначала покажем, как LU-разложение помогает решить линейную систему, а затем разберём, как это LU-разложение построить.

В случае $A = LU$ уравнение $Ax = b$ переписывается в виде $L(Ux) = b$. Обозначив результат перемножения $Ux = y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$, можно найти x с помощью последовательного решения 2 систем. Первым действием находим y из

$$Ly = b,$$

а затем x из

$$Ux = y.$$

Такой метод, в отличие от прямого решения «в лоб» не требует громоздких вычислений, так как системы с соответствующими треугольными матрицами решаются легко.

Пример 13

С помощью LU-разложения решим следующую линейную систему:

$$\begin{pmatrix} -2 & -2 & 2 \\ 2 & -1 & 2 \\ 8 & -1 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 4 \end{pmatrix}$$

$$A = \begin{pmatrix} -2 & -2 & 2 \\ 2 & -1 & 2 \\ 8 & -1 & 5 \end{pmatrix} = \underbrace{\begin{pmatrix} -2 & 0 & 0 \\ 2 & -3 & 0 \\ 8 & -9 & 1 \end{pmatrix}}_L \underbrace{\begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & -4/3 \\ 0 & 0 & 1 \end{pmatrix}}_U$$

Таким образом, систему можно переписать в виде $LUx = b$.

$$\underbrace{\begin{pmatrix} -2 & 0 & 0 \\ 2 & -3 & 0 \\ 8 & -9 & 1 \end{pmatrix}}_L \underbrace{\begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & -4/3 \\ 0 & 0 & 1 \end{pmatrix}}_U \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 4 \end{pmatrix}$$

Произведём замену $y = Ux$.

$$\begin{pmatrix} -2 & 0 & 0 \\ 2 & -3 & 0 \\ 8 & -9 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 4 \end{pmatrix}$$

Полученное матричное уравнение эквивалентно системе:

$$\begin{cases} -2y_1 & = 2 \\ 2y_1 - 3y_2 & = 1 \\ 8y_1 - 9y_2 + y_3 & = 4 \end{cases}$$

Прямой подстановкой находим решение:

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ 3 \end{pmatrix}$$

Вернёмся к тождеству $Ux = y$.

$$\begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & -4/3 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ 3 \end{pmatrix},$$

что, в свою очередь, эквивалентно системе:

$$\begin{cases} x_1 + x_2 - x_3 = -1 \\ x_2 - \frac{4}{3}x_3 = -1 \\ x_3 = 3 \end{cases} \Rightarrow x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -1 \\ 3 \\ 3 \end{pmatrix}.$$

Замечание. Не у всякой квадратной матрицы существует LU-разложение. При этом, если такое разложение существует, оно может быть не единственным.

Чтобы сформулировать и доказать теорему об условиях существования LU-разложения, понадобятся следующие результаты.

Лемма 2

Произведение нижнетреугольных матриц является нижнетреугольной матрицей, а произведение верхнетреугольных — верхнетреугольной.

Лемма 3

Матрица, обратная к нижнетреугольной, является нижнетреугольной. Матрица, обратная к верхнетреугольной, является верхнетреугольной.

Теорема 8

Если квадратную матрицу A можно привести методом Гаусса к U без использования элементарных преобразований 1-го типа (перестановка строк местами), то её можно представить в виде $A = LU$, где L — нижнетреугольная, а U — верхнетреугольная матрицы.

Доказательство

В качестве обозначений матриц элементарных преобразований, вне зависимости от типа (2-го или 3-го), используем E . Рассмотрим произвольную матрицу A , которую можно привести к U без помощи перемены строк местами (элементарных преобразований 1-го типа). Это означает, что существуют такие матрицы элементарных преобразований E_1, E_2, \dots, E_k , что

$$E_k \cdot E_{k-1} \cdot \dots \cdot E_1 \cdot A = U$$

В силу теоремы 6 матрицы элементарных преобразований обратимы. Таким образом,

$$A = E_1^{-1} \cdot E_2^{-1} \cdot \dots \cdot E_k^{-1} U.$$

Обозначим $L = E_1^{-1} \cdot E_2^{-1} \cdot \dots \cdot E_k^{-1}$. По условию матрицы E_j соответствуют только преобразованиям 2-го или 3-го типов, то есть либо получаются из единичной матрицы умножением какой-то из строк на скаляр ($\neq 0$), либо умножением какой-то из строк на скаляр и добавлением её к другой строке. В каждом из таких случаев полученная матрица — нижнетреугольная. Используя леммы 2 и 3, получаем, что L — тоже нижнетреугольная. ◀

Пример 14

Найдём матрицы L и U для LU-разложения матрицы $A = \begin{pmatrix} -2 & -2 & 2 \\ 2 & -1 & 2 \\ 8 & -1 & 5 \end{pmatrix}$.

Для этого будем поэтапно проводить элементарные преобразования со строками матрицы 2-го и 3-го типов и записывать соответствующие матрицы элементарных преобразований вместе с матрицами, обратными к ним.

1.

$$(1) \leftrightarrow \left(-\frac{1}{2}\right) \cdot (1) \Rightarrow \begin{pmatrix} 1 & 1 & -1 \\ 2 & -1 & 2 \\ 8 & -1 & 5 \end{pmatrix},$$
$$E_1 = \begin{pmatrix} -1/2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad E_1^{-1} = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

2.

$$(2) \leftrightarrow (2) - 2 \cdot (1) \Rightarrow \begin{pmatrix} 1 & 1 & -1 \\ 0 & -3 & 4 \\ 8 & -1 & 5 \end{pmatrix},$$

$$E_2 = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad E_2^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ \mathbf{2} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

3.

$$(3) \leftrightarrow (3) - 8 \cdot (1) \Rightarrow \begin{pmatrix} 1 & 1 & -1 \\ 0 & -3 & 4 \\ 0 & -9 & 13 \end{pmatrix},$$

$$E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -8 & 0 & 1 \end{pmatrix}, \quad E_3^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \mathbf{8} & 0 & 1 \end{pmatrix}.$$

4.

$$(2) \leftrightarrow -\frac{1}{3} \cdot (2) \Rightarrow \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & -4/3 \\ 0 & -9 & 13 \end{pmatrix},$$

$$E_4 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1/3 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad E_4^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \mathbf{-3} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

5.

$$(3) \leftrightarrow (3) + 9 \cdot (2) \Rightarrow \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & -4/3 \\ 0 & 0 & 1 \end{pmatrix},$$

$$E_5 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 9 & 1 \end{pmatrix}, \quad E_5^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \mathbf{-9} & 1 \end{pmatrix}.$$

Таким образом,

$$U = \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & -4/3 \\ 0 & 0 & 1 \end{pmatrix}$$

$$L = E_1^{-1} \cdot E_2^{-1} \cdot \dots \cdot E_5^{-1} =$$

$$= \begin{pmatrix} \mathbf{-2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ \mathbf{2} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \mathbf{8} & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \mathbf{-3} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \mathbf{-9} & 1 \end{pmatrix} = \begin{pmatrix} \mathbf{-2} & 0 & 0 \\ \mathbf{2} & \mathbf{-3} & 0 \\ \mathbf{8} & \mathbf{-9} & 1 \end{pmatrix}.$$

Представление

$$\begin{pmatrix} -2 & -2 & 2 \\ 2 & -1 & 2 \\ 8 & -1 & 5 \end{pmatrix} = \begin{pmatrix} -2 & 0 & 0 \\ 2 & -3 & 0 \\ 8 & -9 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & -4/3 \\ 0 & 0 & 1 \end{pmatrix}$$

является LU-разложением матрицы A .

Замечание. Чтобы найти матрицу L , необязательно перемножать все матрицы E_j . Если аккуратно проводить элементарные преобразования строк, то записывая множители, соответствующие преобразованиям 2-го типа, на главной диагонали, а множители, соответствующие преобразованиям 3-го типа, под главной диагональю, можно легко составить матрицу L (см. элементы, выделенные цветом).

Многие компьютерные алгоритмы для решения линейных систем используют элементарные преобразования строк 1-го типа (перемены строк местами). В этом случае существование LU-разложения не гарантировано. Эта проблема решается тем, что перестановки строк делаются до поиска LU-разложения.

Иными словами, сначала составляем матрицу необходимых перестановок P и проводим эти перестановки с помощью эквивалентного вычисления PA . LU-разложение мы находим уже для новой матрицы:

$$PA = LU. \quad (1)$$

Матрица P является произведением матриц элементарных преобразований, поэтому является обратной. Системы уравнений $Ax = b$ и $PAx = Pb$ имеют одни и те же решения. Последняя система может быть решена с помощью LU-разложения.

Тождество (1) ещё иногда записывают в виде $A = P^{-1}LU$, а соответствующее разложение матрицы A называют PLU-разложением.

8 Вычислительная сложность алгоритмов решения линейных систем

При реализации алгоритмов на компьютере, особенно тех, что используют большие объёмы данных, важно учитывать время, которое требуется машине для вычислений.

Это время зависит как от мощности процессора, так и от вычислительной сложности алгоритма, то есть от количества арифметических операций, которые нужны для получения результата.

«Время — деньги». Выбор правильного алгоритма, который сможет сэкономить время на вычислениях, может быть очень критичным и в научных исследованиях, и при решении прикладных задач в индустрии.

Рассчитаем вычислительную сложность¹ метода Гаусса и метода Гаусса-Жордана для решения системы $Ax = b$, $A \in M_{nn}$. Для простоты потребуем, чтобы матрица A была обратима, и чтобы для её приведения к ступенчатому виду не использовались операции 1-го типа (перестановка строк местами). Ниже при обсуждении операции умножения и деления мы будем называть «операциями умножения», а операции сложения и вычитания — «операциями сложения».

Начнём с прямого хода метода Гаусса.

¹Под арифметическими операциями будем подразумевать операции с плавающей запятой (*F*loating *p*oint *O*Peration *p*er *S*econd, *flops*). От них произошло название единицы измерения вычислительной мощности компьютеров «флопс». Она показывает, сколько операций с плавающей запятой в секунду выполняет компьютер. Производительность самого мощного суперкомпьютера с сегодня достигает 4 эксафлопс = $4 \cdot 10^{18}$ (4 квинтиллиона) операций в секунду. Для сравнения процессоры персональных компьютеров разгоняются до нескольких сотен Гфлоп в секунду (1 Гигафлоп = 10^9 флопс).

Шаг №1. Чтобы добиться 1 на месте 1-го главного элемента матрицы, требуется n умножений.

$$\left(\begin{array}{cccc|cc} 1 & \times & \times & \dots & \times & \times & \times \\ \bullet & \bullet & \bullet & \dots & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \dots & \bullet & \bullet & \bullet \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \bullet & \bullet & \bullet & \dots & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \dots & \bullet & \bullet & \bullet \end{array} \right)$$

Шаг №2. Чтобы получить 0 в каждой строке под главным элементом в 1-м столбце, необходимо n умножений и n сложений. Всего у нас $n - 1$ строк под первым главным элементом. Таким образом, число операций на этом шаге для 1-го столбца $2n \cdot (n - 1)$.

В итоге для 1-го столбца общее число операций составляет $n + 2n(n - 1) = 2n^2 - n$.

Аналогично повторяем шаги для 2-го столбца. Только теперь у нас становится на 1 строку и столбец меньше. Общее число операций равно $2(n - 1)^2 - (n - 1)$.

Те же самые шаги повторяем для каждого последующего столбца.

$$\left(\begin{array}{cccc|cc} 1 & \bullet & \bullet & \dots & \bullet & \bullet & \bullet \\ 0 & \times & \times & \dots & \times & \times & \times \\ 0 & \times & \times & \dots & \times & \times & \times \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \times & \times & \dots & \times & \times & \times \\ 0 & \times & \times & \dots & \times & \times & \times \end{array} \right)$$

В итоге получаем выражение

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

Замечание. Напомним некоторые формулы сумм:

$$1 + 2 + 3 + \dots + n = \frac{n(n+1)}{2}$$

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

Значит, общее количество арифметических операций, которые требуются для реализации прямого хода метода Гаусса, равно

$$\frac{2}{3}n^3 + \frac{1}{2}n^2 - \frac{1}{6}n.$$

Для обратного хода метода Гаусса вычислений на каждом этапе становится меньше.

$$\left(\begin{array}{cccccc|c} 1 & \bullet & \bullet & \dots & \bullet & 0 & \times \\ 0 & 1 & \bullet & \dots & \bullet & 0 & \times \\ 0 & 0 & 1 & \dots & \bullet & 0 & \times \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 & \times \\ 0 & 0 & 0 & \dots & 0 & 1 & \bullet \end{array} \right)$$

Так, для n -го столбца, чтобы добиться нулей над последним диагональным элементом, необходимо $n - 1$ умножений и $n - 1$ сложений. В результате — $2(n - 1)$ операций. Для $n - 1$ -го столбца количество операций равно $2(n - 2)$ и т.д. Таким образом, общее количество операций для обратного хода метода Гаусса составляет

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

Метод Гаусса-Жордана целиком требует $(\frac{2}{3}n^3 - \frac{1}{6}n) + n^2 - n = \frac{2}{3}n^3 + \frac{5}{6}n^2 - \frac{7}{6}n$ арифметических операций. Асимптотически (то есть в пределе при $n \rightarrow \infty$) вычислительная сложность метода Гаусса-Жордана эквивалентна $\frac{2}{3}n^3$.

Следствие

Асимптотическая вычислительная сложность решения линейной системы с помощью LU-разложения эквивалентна $\frac{2}{3}n^3$.

Зачем использовать LU-разложение, если его вычислительная сложность совпадает со сложностью стандартного метода Гаусса? Дело в том, что у LU-разложения есть целый ряд преимуществ:

1. В методах Гаусса и Гаусса-Жордана для решения системы $Ax = b$ используется расширенная матрица $(A|b)$. Поэтому изначально для работы алгоритма должен быть задан столбец свободных членов b . В LU-разложении используется только матрица A . Как только разложение получено, его можно использовать для решения системы с любым b .

2. Для матриц с большим количеством нулей и ненулевыми элементами, сконцентрированными вокруг главной диагонали (**разряженных матриц**), существуют методы, которые сокращают арифметическую сложность LU-разложения, делая метод более эффективным, чем метод Гаусса
3. LU-разложение позволяет оптимизировать хранение используемой информации благодаря форме матриц. Все элементы нижнетреугольной матрицы L и верхнетреугольной матрицы U можно записать в одну матрицу того же размера. Это свойство полезно, если критичен объём занимаемой памяти.