Глава II. Системы линейных уравнений

§ 5. Метод Гаусса

Б.М.Верников

Уральский федеральный университет,
Институт естественных наук и математики,
кафедра алгебры и фундаментальной информатики

Вступительные замечания

В этом параграфе излагается метод решения произвольной системы линейных уравнений, известный под названием *метода Гаусса*. Он назван в честь великого немецкого математика Карла Фридриха Гаусса, жившего с 1777 по 1855 г. Хотя это название и является общепринятым, Гаусс не является его автором: метод был известен задолго до него. Первое его описание имеется в китайском трактате «Математика в девяти книгах», который составлен между II в. до н. э. и I в. н. э. и представляет собой компиляцию более ранних трудов, написанных в X–II вв. до н. э.

В дальнейшем мы постоянно будем использовать метод Гаусса при решении самых разных задач.

5.1. Основные понятия

Определения

Линейным уравнением (или *уравнением 1-го порядка*) с n неизвестными $x_1, x_2, \dots, x_n$ называется уравнение вида

$$a_1x_1 + a_2x_2 + \dots + a_nx_n = b. \qquad (1)$$

Величины $a_1, a_2, \dots, a_n$ называются *коэффициентами* при неизвестных, а b — *свободным членом* уравнения (1). Коэффициенты при неизвестных и свободный член предполагаются известными.

Произвольная система линейных уравнений записывается следующим образом:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m. \end{cases} \qquad (2)$$

Всюду далее мы будем предполагать, что коэффициенты при неизвестных и свободные члены в системах линейных уравнений лежат в некотором поле F и называть систему (4) *системой линейных уравнений над полем F*. Элементы поля F мы будем называть *скалярами*.

Частное и общее решение системы. Совместные и несовместные системы

Определения

Частным решением (или просто *решением*) системы (4) называется упорядоченный набор скаляров $(x_1^0, x_2^0, \ldots, x_n^0)$ из поля F такой, что при подстановке в любое уравнение системы (4) x_1^0 вместо x_1, x_2^0 вместо x_2, $\ldots$, x_n^0 вместо x_n получается верное равенство. Система линейных уравнений (4) называется *совместной*, если она имеет хотя бы одно частное решение, и *несовместной* в противном случае. *Общим решением* системы (4) называется множество всех ее частных решений.

- Общее решение есть у любой системы. В частности, у несовместной системы общим решением является пустое множество.
- Решить систему линейных уравнений — значит найти ее общее решение.

Как мы увидим в дальнейшем, во многих задачах, а также при анализе строения общего решения произвольной системы линейных уравнений важную роль играют системы, у которых правые части всех уравнений равны 0. Такие системы имеют специальное название.

Определение

Система линейных уравнений, в которой правые части всех уравнений равны 0, называется *однородной*.

Очевидно, что если в любое уравнение однородной системы вместо всех неизвестных подставить 0, то получится верное равенство. Иначе говоря, набор скаляров $(\underbrace{0, 0, \ldots, 0}_{n \text{ раз}})$, где n — число неизвестных в системе, является частным решением любой однородной системы. Это решение называется *нулевым* решением. Из сказанного вытекает

Замечание 5.1

Любая однородная система линейных уравнений совместна. □

Определение

Если в системе (4) все свободные члены заменить нулями, то мы получим однородную систему

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0, \end{cases} \tag{3}$$

которую мы будем называть *однородной системой, соответствующей системе* (4).

Основная матрица, столбец неизвестных и столбец свободных членов системы линейных уравнений

Определения

Матрица

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

называется *основной матрицей* (или просто *матрицей*) *системы* (4).
Первый из столбцов

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \text{и} \quad \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

называется *столбцом неизвестных* системы (4), а второй — *столбцом свободных членов* этой системы.

Обозначим основную матрицу, столбец неизвестных и столбец свободных членов системы (4) через A, X и B соответственно. Матрицы A и X имеют размеры $m \times n$ и $n \times 1$, соответственно. По определению произведения матриц, матрица AX существует и имеет размер $m \times 1$ — такой же, как у матрицы B. Запишем матрицу AX:

$$AX = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \dotfill \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix}$$

и приравняем ее к матрице B:

$$\begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \dotfill \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Если расписать полученное матричное равенство поэлементно, то получится не что иное, как система (4). Таким образом, это матричное равенство равносильно системе (4). Равенство $AX = B$ (при условии, что X и B — матрицы, состоящие из одного столбца) называется *матричной записью системы линейных уравнений* (4).

Определение

Пусть дана система линейных уравнений

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m. \end{cases} \tag{4}$$

Матрица

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{pmatrix} \tag{5}$$

называется *расширенной матрицей* этой системы.

Расширенная матрица системы линейных уравнений (стандартная запись)

Обычно, записывая расширенную матрицу, ее последний столбец отделяют от остальной части матрицы вертикальной чертой:

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{array}\right).$$

Это делается, чтобы подчеркнуть, что слева и справа от вертикальной черты стоят скаляры, играющие принципиально разную роль в системе: слева — коэффициенты при неизвестных, а справа — свободные члены уравнений.

Расширенная матрица системы линейных уравнений (комментарий)

В расширенной матрице системы каждая строка соответствует какому-то уравнению, каждый столбец, кроме последнего, — это набор коэффициентов при некотором неизвестном в различных уравнениях системы, а последний столбец — это совокупность свободных членов системы. Таким образом,

- расширенная матрица системы содержит в себе полную информацию о системе.

Иными словами, не только по системе линейных уравнений однозначно выписывается ее расширенная матрица, но и наоборот, по произвольной матрице A, содержащей более одного столбца[1], однозначно восстанавливается система линейных уравнений, расширенной матрицей которой является матрица A.

Определение

Мы будем говорить, что система (4) *соответствует* матрице (5).

[1] Эта оговорка необходима, так как требуется один столбец для свободных членов и как минимум один столбец для коэффициентов при неизвестных.

5.2. Общая схема и корректность метода Гаусса

Приступим к изложению метода Гаусса. Его можно реализовывать двумя способами — на языке линейных уравнений и на языке матриц. Мы изложим второй из этих способов. В самом общем виде метод Гаусса можно описать как последовательность из следующих четырех шагов:

1) записываем расширенную матрицу данной системы линейных уравнений;
2) с помощью некоторых преобразований (называемых *элементарными преобразованиями матрицы*) приводим эту матрицу к некоторому специальному виду (так называемой *ступенчатой* матрице);
3) восстанавливаем систему, соответствующую полученной ступенчатой матрице;
4) решаем систему, полученную на предыдущем шаге.

При этом оказывается, что общее решение системы, соответствующей полученной ступенчатой матрице, совпадает с общим решением исходной системы, а система, соответствующая произвольной ступенчатой матрице, решается легко.

Шаги 1) и 3) тривиальны, и мы их далее комментировать не будем (отметим, что при решении конкретных задач шаг 3), как правило, в явном виде не осуществляют).

Определение

Элементарными преобразованиями матрицы называются следующие действия:

1) умножение строки на ненулевой скаляр;
2) прибавление одной строки к другой;
3) перестановка двух строк;
4) перестановка двух столбцов;
5) вычеркивание или добавление нулевой строки.

Определение

Матрицы A и B называются *эквивалентными*, если одна из них может быть получена из другой с помощью конечного числа элементарных преобразований. Тот факт, что матрицы A и B эквивалентны, обозначается так: $A \sim B$.

Обоснование корректности метода Гаусса (1)

Определение

Системы линейных уравнений называются *равносильными*, если они имеют одно и то же общее решение.

Следующее несложно проверяемое утверждение принципиально важно, так как оно обосновывает корректность метода Гаусса.

Предложение 5.1

Если матрица B получена из матрицы A с помощью элементарных преобразований типов 1)–3) и 5), то системы линейных уравнений, соответствующие матрицам A и B равносильны.

Доказательство предложения приведено на двух следующих слайдах.

Доказательство. Договоримся называть систему линейных уравнений, соответствующую матрице A, старой, а систему, соответствующую матрице B, — новой. Достаточно рассмотреть случай, когда матрица B получена из A с помощью одного элементарного преобразования. В зависимости от типа этого преобразования возможны 4 случая.

Случай 1: *B получена из A умножением i-й строки на ненулевой скаляр t.* В этом случае новая система получена из старой умножением i-го уравнения на t. Ясно, что всякое решение старой системы является решением новой. Поскольку старая система получается из новой умножением i-го уравнения на ненулевой скаляр $\frac{1}{t}$, верно и обратное утверждение.

Случай 2: *B получена из A прибавлением j-й строки к i-й.* Поскольку сумма двух верных равенств является верным равенством, всякое решение старой системы является решением новой. Далее, матрицу A можно получить из матрицы B выполнением трех элементарных преобразований — сначала умножаем j-ю строку матрицы B (совпадающую с j-й строкой матрицы A!) на -1, затем прибавляем полученную строку к i-й строке матрицы B, и, наконец, еще раз умножаем j-ю строку матрицы B на -1. В силу сказанного выше, всякое решение новой системы является и решением старой.

Обоснование корректности метода Гаусса (3)

Случай 3: *B получена из A перестановкой строк.* В этом случае системы, соответствующие матрицам A и B, различаются лишь порядком записи уравнений, что, очевидно, не влияет на общее решение системы.

Случай 4: *B получена из A вычеркиванием или добавлением нулевой строки.* Это означает, что новая система получена из старой вычеркиванием или добавлением «тривиального» уравнения, т. е. уравнения вида $0 \cdot x_1 + 0 \cdot x_2 + \cdots + 0 \cdot x_n = 0$. Очевидно, что эта операция никак не может повлиять на общее решение системы. $\square$

В предложении 5.1 не упоминается об элементарном преобразовании типа 4) (перестановке столбцов). Дело в том, что если при элементарных преобразованиях расширенной матрицы системы переставить местами последний столбец матрицы с некоторым другим ее столбцом, то система, соответствующая полученной матрице, может оказаться не равносильной исходной системе. Как мы увидим ниже, при приведении матрицы к ступенчатому виду всегда можно обойтись без этого элементарного преобразования. Но исключать его из числа элементарных преобразований невыгодно, так как им бывает удобно пользоваться при решении задач, не связанных с решением систем линейных уравнений.

5.3. Приведение матрицы к ступенчатому виду

Введем понятие, которое будет играть важную роль в дальнейшем.

Определение

Матрица называется *ступенчатой*, если выполнено следующее условие: если некоторая строка матрицы, отличная от первой, не является нулевой, то в начале этой строки стоит больше нулей, чем в начале предыдущей строки.

Из определения непосредственно вытекает, что если некоторая строка ступенчатой матрицы является нулевой, то и все ее последующие строки — нулевые. В самом деле, если в матрице после нулевой строки идет ненулевая, то для этой ненулевой строки нарушено условие из определения ступенчатой матрицы (в начале этой строки стоит меньше нулей, чем в начале предыдущей).

Общий вид ступенчатой матрицы

Произвольная ступенчатая матрица имеет следующий вид:

$$\begin{pmatrix}
0 \ldots 0 & * & & & & \\
0 \ldots\ldots\ldots 0 & & * & & & \\
\ldots\ldots\ldots\ldots\ldots\ldots & & & \ddots & & \\
0 \ldots\ldots\ldots\ldots\ldots\ldots 0 & & & & * & \\
0 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 0 & & & & & * \\
0 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 0 & & & & & \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & & & & & \\
0 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 0 & & & & &
\end{pmatrix}.$$

Элементы, обозначенные звездочками, не должны быть равны 0. Напротив, все элементы, стоящие в «серой зоне» (левее и ниже синей ломаной линии), обязаны быть равны 0. Элементы, стоящие выше ломаной (кроме тех, что обозначены звездочками), могут быть любыми. Нулевых столбцов в левой части матрицы, как и нулевых строк в ее нижней части, может не быть.

- Ломаная линия, ограничивающая сверху «нулевую часть» ступенчатой матрицы, имеет вид ступенек. Именно этим объясняется термин «ступенчатая матрица».

Если переходить от ненулевой строки ступенчатой матрицы к следующей за ней ненулевой строке (до тех пор, пока это возможно), то мы каждый раз будем сдвигаться ровно на одну строку вниз и на, вообще говоря, произвольное число столбцов вправо. Следовательно, справедливо

Замечание 5.2

В любой ступенчатой матрице число ненулевых строк не превышает числа столбцов. □

Приведение матрицы к ступенчатому виду (1)

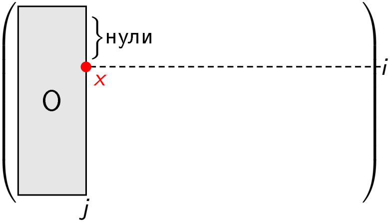
Ключевую роль в методе Гаусса играет следующее утверждение.

Предложение 5.2

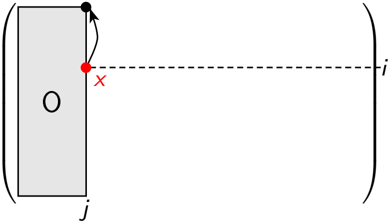
Любую матрицу с помощью конечного числа элементарных преобразований можно привести к ступенчатому виду. При этом достаточно использовать только элементарные преобразования типов 1)–3).

Доказательство этого утверждения дает алгоритм приведения произвольной матрицы к ступенчатому виду, который будет постоянно использоваться в дальнейшем при решении самых разных задач (в том числе таких, которые не связаны напрямую с решением систем линейных уравнений).

Доказательство. Пусть A — произвольная матрица. Если A — нулевая матрица, то она уже является ступенчатой. Поэтому далее будем считать, что A содержит хотя бы один ненулевой элемент. Выберем в A самый левый ненулевой столбец (обозначим его номер через j), а в этом столбце — самый верхний ненулевой элемент. Обозначим этот элемент через x, а номер строки, в которой он стоит, — через i.

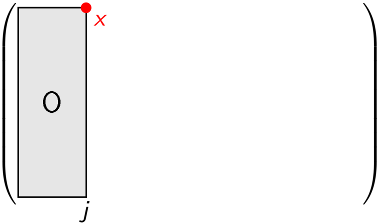


Если $i > 1$, поменяем местами первую и i-ю строки.



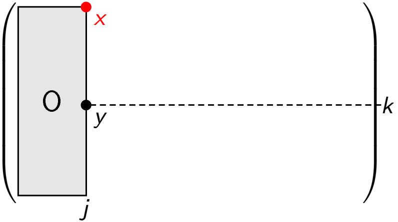
Приведение матрицы к ступенчатому виду (4)

Теперь матрица выглядит так:



Наша следующая цель — обнулить все элементы j-го столбца, стоящие ниже первой строки.

Предположим, что в j-м столбце есть ненулевой элемент y, стоящий в k-й строке, где $k > 1$.

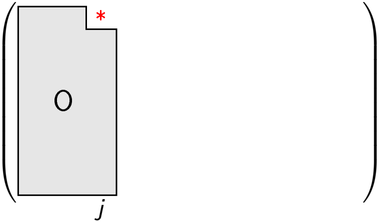


Прибавим к k-й строке, умноженной на x, первую строку, умноженную на $-y$. На самом деле мы выполнили здесь последовательность из четырех элементарных преобразований: сначала умножили первую строку на $-y$, затем умножили k-ю строку на x, затем прибавили первую строку к k-й, и, наконец, умножили первую строку на $-\frac{1}{y}$ (возвращая ее в исходное состояние). После этого в k-й строке и j-м столбце будет стоять элемент $xy - yx = 0$.

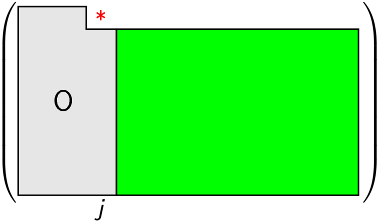
Действуя таким образом, обнулим в j-м столбце все элементы, расположенные ниже первой строки.

Приведение матрицы к ступенчатому виду (6)

«Нулевая зона» продвинется на одну строку вниз и как минимум на один столбец вправо.

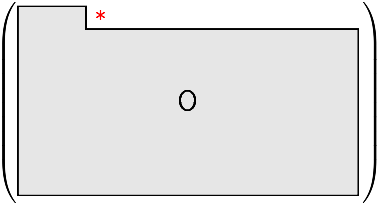


Рассмотрим часть полученной матрицы, расположенную правее j-го столбца и ниже первой строки.

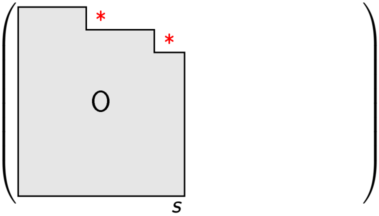


Приведение матрицы к ступенчатому виду (8)

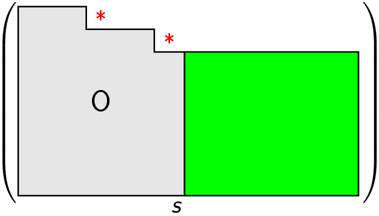
Если все элементы в этой части матрицы равны 0, то полученная матрица является ступенчатой.



В противном случае повторим для этой части матрицы все описанные выше действия. А именно: найдем в этой части матрицы самый левый ненулевой столбец. Обозначим его номер (во всей матрице) через s. В этом столбце найдем самый верхний ненулевой элемент (расположенный ниже первой строки всей матрицы). Обозначим этот элемент через x'. Номер строки (во всей матрице), в которой стоит x', обозначим через r. Если $r > 2$, поменяем местами вторую и r-ю строки. Обнулим все ненулевые элементы s-го столбца, расположенные ниже второй строки, прибавляя к строкам, в которых стоят эти элементы, вторую строку, умноженную на подходящее число. В результате заполненная нулями зона в левой нижней части матрицы продвинется на одну строку вниз и как минимум на один столбец вправо.

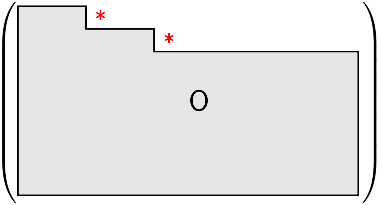


Рассмотрим часть полученной матрицы, расположенную правее s-го столбца и ниже второй строки.



Приведение матрицы к ступенчатому виду (11)

Если все элементы в этой части матрицы равны 0, то полученная матрица является ступенчатой.



В противном случае применим к этой части матрицы описанные выше действия. Будем продолжать этот процесс. Рано или поздно он оборвется, поскольку либо мы получим, что часть матрицы, расположенная ниже очередной строки и правее очередного столбца, состоит из нулей, либо в матрице не останется больше строк, либо в ней не останется больше столбцов. В любом случае полученная матрица будет ступенчатой. □

Комментарий к алгоритму приведения матрицы к ступенчатому виду

В доказательстве предложения 5.2 используются только первые три элементарных преобразования. Таким образом, при приведении матрицы к ступенчатому виду можно обойтись не только без перестановки столбцов (что принципиально важно с точки зрения предложения 5.1), но и без вычеркивания или добавления нулевых строк. Но совсем отказываться от возможности применить последнее элементарное преобразование невыгодно: вместо того, чтобы, строго придерживаясь алгоритма приведения матрицы к ступенчатому виду, «сбрасывать» нулевые строки в нижнюю часть матрицы, их можно вычеркивать, тем самым экономя время и место (а при компьютерной реализации метода Гаусса — объем используемой памяти).

5.4. Нахождение общего решения

Для того, чтобы завершить изложение метода Гаусса, нам осталось объяснить, как искать общее решение системы линейных уравнений, соответствующей ступенчатой матрице. Здесь возможны три случая.

Случай 1: ступенчатая матрица содержит строку, в которой все элементы, кроме последнего, равны 0, а последний элемент нулю не равен. Эта строка соответствует уравнению вида $0 \cdot x_1 + 0 \cdot x_2 + \cdots + 0 \cdot x_n = b$, где $b \neq 0$. Ясно, что это уравнение, а значит и произвольная система, его содержащая, решений не имеет. Учитывая предложение 5.1, получаем, что

- *в рассматриваемом случае система несовместна.*

Для удобства будем называть строки матрицы, в которых все элементы, кроме последнего, равны 0, а последний элемент нулю не равен, *плохими*. Заметим, что

- если при приведении расширенной матрицы системы к ступенчатому виду плохая строка возникла в тот момент, когда матрица еще не является ступенчатой, то продолжать преобразования не имеет смысла, так как уже в этот момент стало ясно, что система несовместна.

Нахождение общего решения системы по ступенчатой матрице: переход к двум оставшимся случаям

- *Всюду в дальнейшем мы будем считать, что при приведении расширенной матрицы системы к ступенчатому виду плохих строк не возникло и мы довели матрицу до ступенчатого вида.*

Определение

Совокупность всех столбцов расширенной матрицы системы, кроме ее последнего столбца, мы будем называть *основной частью* расширенной матрицы.

В силу замечания 5.2 возможны два случая: в основной части полученной ступенчатой матрицы число ненулевых строк либо равно числу столбцов, либо меньше этого числа.

Случай 2: число ненулевых строк ступенчатой матрицы равно числу столбцов в ее основной части. Вычеркнем из матрицы нулевые строки (если они в ней есть). Система линейных уравнений, соответствующая полученной матрице, имеет вид

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ \qquad\quad a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \qquad\qquad \dots\dots\dots\dots\dots\dots \\ \qquad\qquad\qquad\quad a_{nn}x_n = b_n, \end{cases}$$

где $a_{11}, a_{22}, \dots, a_{nn} \neq 0$. Из последнего уравнения этой системы находим x_n: $x_n = \frac{b_n}{a_{nn}}$. После этого из предпоследнего уравнения находим x_{n-1}:

$$x_{n-1} = \frac{b_{n-1} - a_{n-1\,n}x_n}{a_{n-1\,n-1}}.$$

Продолжая двигаться по системе снизу вверх, мы из третьего с конца уравнения найдем x_{n-2}, из четвертого с конца — x_{n-3}, ..., наконец, из первого — x_1. На каждом шаге очередное неизвестное определяется однозначно. Следовательно,

- *в рассматриваемом случае система имеет единственное решение.*

Случай 3: число ненулевых строк ступенчатой матрицы меньше числа столбцов в ее основной части. Вычеркнем из матрицы нулевые строки (если они в ней есть). Система линейных уравнений, соответствующая полученной матрице, может быть схематично записана в виде

$$\begin{cases} a_{1i_1}x_{i_1} + \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots = b_1, \\ \qquad\qquad a_{2i_2}x_{i_2} + \dots\dots\dots\dots\dots\dots = b_2, \\ \qquad\qquad\qquad \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \qquad\qquad\qquad\qquad a_{mi_m}x_{i_m} + \cdots = b_m, \end{cases}$$

где $i_1 < i_2 < \cdots < i_m$ и $a_{1i_1}, a_{2i_2}, \ldots, a_{mi_m} \neq 0$. При этом в систему входит как минимум одна неизвестная, отличная от $x_{i_1}, x_{i_2}, \ldots, x_{i_m}$, так как в противном случае число ненулевых строк было бы равно числу столбцов в основной части матрицы. Перенесем все неизвестные, кроме $x_{i_1}, x_{i_2}, \ldots, x_{i_m}$, в правые части равенств с обратным знаком. Получим систему, которую можно схематично записать в виде

$$\begin{cases} a_{1i_1}x_{i_1} + \dots\dots\dots\dots\dots\dots = b_1 - \cdots, \\ \qquad a_{2i_2}x_{i_2} + \dots\dots\dots = b_2 - \cdots, \\ \qquad\qquad \dots\dots\dots\dots\dots\dots\dots \\ \qquad\qquad a_{mi_m}x_{i_m} = b_m - \cdots . \end{cases} \tag{6}$$

Нахождение общего решения системы по ступенчатой матрице: случай неопределенной системы (2)

Переменные, входящие в правые части уравнений системы (6), называются *свободными*, а переменные x_{i_1}, x_{i_2}, ..., x_{i_m} — *основными* или *связанными*. Придадим свободным переменным произвольные значения, подставим их в систему (6) и обозначим правые части полученных равенств через b'_1, b'_2, ..., b'_m. Получим систему m линейных уравнений с m неизвестными:

$$\begin{cases} a_{1i_1}x_{i_1} + \dots\dots\dots\dots\dots\dots\dots = b'_1, \\ \qquad\qquad a_{2i_2}x_{i_2} + \dots\dots\dots\dots = b'_2, \\ \qquad\qquad\qquad\qquad \dots\dots\dots\dots\dots\dots\dots \\ \qquad\qquad\qquad\qquad\qquad a_{mi_m}x_{i_m} = b'_m. \end{cases}$$

В ступенчатой матрице, соответствующей этой системе, число ненулевых строк равно числу столбцов в основной части матрицы. Как мы видели выше при рассмотрении случая 2, эта система имеет единственное решение. Найдя его и объединив полученные значения переменных x_{i_1}, x_{i_2}, ..., x_{i_m} с теми значениями, которые мы подставили вместо свободных переменных в правые части системы (6), мы найдем одно частное решение исходной системы. Значения свободным переменным можно придавать многими разными способами. Следовательно, система имеет более одного решения. Таким образом,

- *в рассматриваемом случае система является неопределенной.*

Нахождение общего решения системы по ступенчатой матрице: случай неопределенной системы (3)

Остается понять, как в рассматриваемом случае можно записать общее решение системы. Для простоты обозначений будем считать, что свободными являются переменные $x_{m+1}, \ldots, x_n$, а связанными — переменные $x_1, x_2, \ldots, x_m$ (если это не так, мы можем переименовать неизвестные). Иными словами, будем считать, что система (6) имеет вид

$$\left\{\begin{array}{rl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m = & b_1 - a_{1\,m+1}x_{m+1} - \cdots - a_{1n}x_n, \\ a_{22}x_2 + \cdots + a_{2m}x_m = & b_2 - a_{2\,m+1}x_{m+1} - \cdots - a_{2n}x_n, \\ \multicolumn{2}{r}{\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots} \\ a_{mm}x_m = & b_m - a_{m\,m+1}x_{m+1} - \cdots - a_{mn}x_n. \end{array}\right. \qquad (7)$$

Пусть $k = n - m$. Положим $x_{m+1} = c_1$, $x_{m+2} = c_2$, $\ldots$, $x_n = c_k$, где $c_1, c_2, \ldots, c_k$ — произвольные константы. Из последнего уравнения системы (7) имеем $x_m = \frac{b_m - a_{m\,m+1}c_1 - \cdots - a_{mn}c_k}{a_{mm}}$. Подставив правую часть этого равенства вместо x_m в предпоследнее уравнение системы (7), мы сможем выразить x_{m-1} через $c_1, c_2 \ldots, c_k$. Продолжая двигаться по системе (7) снизу вверх, мы последовательно выразим через эти константы переменные $x_{m-2}, \ldots, x_1$. Объединив полученные равенства, выражающие переменные $x_1, x_2, \ldots, x_m$ через константы $c_1, c_2, \ldots, c_k$, с равенствами $x_{m+1} = c_1$, $\ldots$, $x_n = c_k$, мы получим запись общего решения системы.

Таким образом, общее решение неопределенной системы можно записать в виде

$$\begin{cases} x_1 &= f_1(c_1, c_2, \ldots, c_k), \\ x_2 &= f_2(c_1, c_2, \ldots, c_k), \\ \multicolumn{2}{l}{\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots} \\ x_m &= f_m(c_1, c_2, \ldots, c_k), \\ x_{m+1} &= c_1, \\ x_{m+2} &= c_2, \\ \multicolumn{2}{l}{\ldots\ldots\ldots\ldots\ldots} \\ x_n &= c_k, \end{cases} \tag{8}$$

где $f_1(c_1, c_2, \ldots, c_k), f_2(c_1, c_2, \ldots, c_k), \ldots, f_m(c_1, c_2, \ldots, c_k)$ — выражения для переменных $x_1, x_2, \ldots, x_m$ через константы $c_1, c_2, \ldots, c_k$, способ получения которых описан на предыдущем слайде. Равенства (8) называются *координатной записью* общего решения системы линейных уравнений. Другой способ записи общего решения системы будет указан в конце § 25.

Мы видим, что каждая из свободных переменных независимо от других может принимать любое значение из поля F. Из этого вытекает следующее утверждение.

Следствие 5.1

а) *Неопределенная система линейных уравнений над бесконечным полем имеет бесконечно много решений.*

б) *Неопределенная система линейных уравнений над конечным полем F имеет $|F|^k$ решений, где k — число свободных переменных в этой системе.* □

Из сказанного при рассмотрении случая 3 вытекает важный для дальнейшего вывод:

Замечание 5.3

Если система линейных уравнений является неопределенной, то число ее свободных переменных равно $n - m$, где n — число столбцов в основной матрице системы (или, что то же самое, число неизвестных в системе), а m — число ненулевых строк в матрице, полученной из расширенной матрицы системы приведением к ступенчатому виду. □

- Расширенную матрицу системы можно привести к ступенчатому виду многими различными способами, причем полученные ступенчатые матрицы могут различаться. Возникает вопрос: однозначно ли определено число свободных переменных в системе. Иначе говоря, верно ли, что приводя матрицу к ступенчатому виду различными способами, мы всегда будем получать ступенчатые матрицы с одним и тем же числом ненулевых строк. Ответ на этот вопрос положителен, но доказать мы это сможем только в § 24 (это вытекает из приводимого там доказательства теоремы о ранге матрицы).

Замечания об однородных системах

Если решать методом Гаусса однородную систему линейных уравнений, то последний столбец расширенной матрицы системы на всех этапах будет нулевым. Переписывать его все время нет никакого смысла. Поэтому

- при решении однородных систем, как правило, выписывают и приводят к ступенчатому виду основную матрицу системы, а при нахождении общего решения «вспоминают», что в матрице неявно присутствует еще последний нулевой столбец.

Замечание 5.4

Если в однородной системе линейных уравнений число уравнений меньше числа неизвестных, то она имеет по крайней мере одно ненулевое решение.

Доказательство. Запишем основную матрицу системы и приведем ее к ступенчатому виду. В исходной матрице число строк равно числу уравнений, а число столбцов — числу неизвестных. По условию первое число меньше второго. При приведении матрицы к ступенчатому виду число ее ненулевых строк может разве что уменьшиться. Следовательно, и в полученной ступенчатой матрице число ненулевых строк будет меньше числа столбцов. Иными словами, мы попадаем в условия рассмотренного выше случая 3, в котором исходная система имеет более одного решения. Все эти решения, кроме одного, являются ненулевыми. $\square$

Некоторые типы матриц

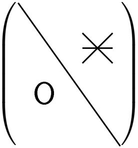
5.5. Метод Гаусса–Жордана

Введем в рассмотрение несколько новых типов матриц, которые часто будут возникать в дальнейшем по самым разным поводам.

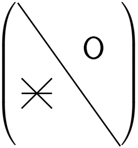
Определения

Квадратная матрица называется: *верхнетреугольной* [*нижнетреугольной*], если все ее элементы, расположенные ниже [выше] главной диагонали, равны 0; *треугольной*, если она либо верхнетреугольна, либо нижнетреугольна; *диагональной*, если все ее элементы, не лежащие на главной диагонали, равны 0.

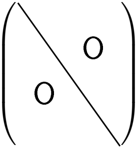
Ниже слева направо символически изображены верхнетреугольная, нижнетреугольная и диагональная матрицы:



Верхнетреугольная матрица



Нижнетреугольная матрица



Диагональная матрица

Мы переходим к изложению модификации метода Гаусса, которая называется *методом Гаусса–Жордана*.

Обозначим через A матрицу, полученную из расширенной матрицы системы приведением к ступенчатому виду, а через B — матрицу, полученную из A вычеркиванием нулевых строк, последнего столбца (содержащего свободные члены) и столбцов, соответствующие свободным переменным (если они существуют). Из замечания 5.3 вытекает, что матрица B квадратна. Кроме того, она верхнетреугольна, поскольку, очевидно, всякая квадратная ступенчатая матрица верхнетреугольна. Для краткости будем называть матрицу B *базовой частью* матрицы A. Идея метода Гаусса–Жордана состоит в том, что

- после приведения расширенной матрицы системы к ступенчатому виду можно продолжить элементарные преобразования и довести базовую часть матрицы сначала до диагонального, а затем и до единичного вида. После этого общее решение системы находится очень легко.

Метод Гаусса–Жордана: реализация

В самом деле, предположим, что расширенная матрица системы приведена к ступенчатому виду. Обозначим число ее ненулевых строк через n, а число столбцов в ее основной части — через m. В силу замечания 5.2 $n \leqslant m$. Вычеркнем из матрицы все нулевые строки. Будем считать, что базовая часть матрицы занимает первые m столбцов в основной части матрицы (этого всегда можно добиться, переставив при необходимости столбцы в основной части матрицы). При этом на главной диагонали базовой части все элементы не равны 0. К каждой строке матрицы, кроме m-й, можно прибавить m-ю строку, умноженную на подходящее число (свое для каждой строки), таким образом, чтобы все элементы m-го столбца выше m-й строки оказались равны 0. После этого аналогичным образом, за счет $(m-1)$-й строки, можно обнулить все элементы $(m-1)$-го столбца, стоящие выше $(m-1)$-й строки. Продолжая этот процесс и двигаясь снизу вверх и справа налево (в отличие от «прямого хода» в методе Гаусса, когда мы, приводя матрицу к ступенчатому виду, двигались сверху вниз и слева направо), мы в конце концов добьёмся того, что базовая часть станет диагональной матрицей с ненулевыми элементами на главной диагонали. Разделив после этого i-ю строку на элемент a_{ii} (для всех $i = 1, 2, \ldots, m$), мы приведем базовую часть к единичному виду.

Метод Гаусса–Жордана: случай неопределенной системы

Для простоты обозначений будем далее считать, что столбцы в базовой части матрицы соответствуют неизвестным $x_1, \ldots, x_m$ (если это не так, мы можем переименовать неизвестные). Рассмотрим сначала случай, когда $n > m$. Ясно, что мы находимся в условиях рассмотренного выше случая 3, и потому система является неопределенной. Она имеет вид

$$\begin{cases} x_1 \qquad\qquad\qquad + a_{1\,m+1}x_{m+1} + \cdots + a_{1n}x_n = b_1, \\ \qquad x_2 \qquad\qquad + a_{2\,m+1}x_{m+1} + \cdots + a_{2n}x_n = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \qquad\qquad x_m + a_{m\,m+1}x_{m+1} + \cdots + a_{mn}x_n = b_m. \end{cases}$$

Переменные $x_1, \ldots, x_m$ являются основными, а переменные $x_{m+1}, \ldots, x_n$ — свободными. Положим $k = n - m$. Перенося слагаемые, содержащие свободные переменные, в правые части уравнений и полагая $x_{m+1} = c_1$, $\ldots$, $x_n = c_k$, получаем координатную запись общего решения системы:

$$\begin{cases} x_1 = b_1 - a_{1\,m+1}c_1 - \cdots - a_{1n}c_k, \\ x_2 = b_2 - a_{2\,m+1}c_1 - \cdots - a_{2n}c_k, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_m = b_m - a_{m\,m+1}c_1 - \cdots - a_{mn}c_k, \\ x_{m+1} = c_1, \\ \ldots\ldots\ldots\ldots \\ x_n = c_k. \end{cases}$$

Предположим теперь, что $n = m$, т. е. что после описанных выше действий базовая часть матрицы совпадает с ее основной частью. В этом случае мы находимся в условиях рассмотренного выше случая 2, и потому система имеет единственное решение. Она имеет вид

$$\begin{cases} x_1 \qquad\qquad\quad = b_1, \\ \quad x_2 \qquad\quad\;\; = b_2, \\ \qquad \dots\dots\dots\dots \\ \qquad\qquad x_n = b_n. \end{cases}$$

Ясно, что с содержательной точки зрения это не система линейных уравнений, а ее (единственное) решение. Исходя из этого, получаем следующий алгоритм, на котором в дальнейшем будут основаны алгоритмы решения некоторых важных задач.

Алгоритм 5.1

Пусть дана система линейных уравнений, имеющая единственное решение. Запишем ее расширенную матрицу и с помощью элементарных преобразований всей матрицы приведем ее основную часть к единичному виду (в рассматриваемом случае это всегда можно сделать). В этот момент в последнем столбце расширенной матрицы будет стоять (единственное) решение системы.

5.6. Матричные уравнения

Матричным уравнением называется уравнение, в котором неизвестным является матрица. Мы рассмотрим матричные уравнения трех видов. Начнем с уравнения вида

$$AX = B, \tag{9}$$

где A и B — известные матрицы, а X — неизвестная. Из определения произведения матриц видно, что число строк в матрице AX равно числу строк в матрице A. Следовательно, если число строк в матрицах A и B различно, то уравнение (9) решений не имеет. Поэтому всюду в дальнейшем, говоря об этом уравнении, мы будем считать, что *матрицы A и B имеют одинаковое число строк*.

Из определения произведения матриц видно также, что число столбцов в матрице AX равно числу столбцов в матрице X. Следовательно, если X — решение уравнения (9), то матрицы X и B содержат одинаковое число столбцов. Как мы видели выше в данном параграфе, если матрицы X и B содержат один столбец, то уравнение (9) есть просто другой способ записи системы линейных уравнений. Обозначим через k число столбцов в матрицах X и B. Для всякого $i = 1, 2, \ldots, k$ обозначим i-й столбец матрицы X через X_i, а i-й столбец матрицы B — через B_i. Из определения произведения матриц вытекает, что i-й столбец матрицы AX равен AX_i.

Поэтому

!! *в общем случае уравнение* (9) *равносильно совокупности* k *систем линейных уравнений вида*

$$AX_1 = B_1, AX_2 = B_2, \ldots, AX_k = B_k. \tag{10}$$

Для того чтобы решить каждую из этих систем методом Гаусса, надо записать расширенную матрицу системы и с помощью элементарных преобразований привести ее к ступенчатому виду. Если при этом окажется, что хотя бы одна из систем несовместна, то и исходное матричное уравнение не имеет решений. Если же все системы совместны, то, решив каждую из них, мы найдем все столбцы матрицы X, а значит и саму эту матрицу. Но у всех решаемых систем основная матрица одна и та же — матрица A. Это позволяет решать все системы одновременно, действуя по алгоритму, который приведен на следующем слайде.

Алгоритм 5.2

Пусть дано уравнение (9), в котором A — матрица размера $n \times m$, а B — матрица размера $n \times k$. Запишем матрицу $(A \mid B)$, т. е. матрицу размера $n \times (m + k)$, в которой в первых m столбцах стоит матрица A, а в последних k столбцах — матрица B. С помощью элементарных преобразований всей этой матрицы приведем ее левую часть (т. е. первые m столбцов) к ступенчатому виду. После этого для всякого $i = 1, 2, \ldots, k$ можно найти i-й столбец матрицы X, решив систему линейных уравнений вида $A'X_i = B'_i$, где A' — левая часть полученной матрицы, а B'_i — i-й столбец правой части полученной матрицы. Если при этом окажется, что хотя бы одна из этих систем несовместна, то уравнение $AX = B$ решений не имеет.

Транспонирование матрицы

Введем одно важное для дальнейшего понятие.

Определение

Пусть $A = (a_{ij})$ — матрица размера $m \times n$. Матрицей, *транспонированной* к A, называется матрица $B = (b_{ij})$ размера $n \times m$, определяемая равенством $b_{ij} = a_{ji}$ для всех $i = 1, 2, \ldots, n$ и $j = 1, 2, \ldots, m$. Иными словами, матрица B получается из A заменой строк на столбцы: первая строка матрицы A становится первым столбцом матрицы B, вторая строка матрицы A — вторым столбцом матрицы B и т. д. Матрица, транспонированная к A, обозначается через $A^\top$.

Очевидно, что

- матрица, транспонированная к квадратной матрице, является квадратной матрицей того же порядка, что и исходная матрица.

Отметим, что транспонирование матрицы является унарной операцией на множестве всех матриц и на множестве всех квадратных матриц данного порядка (но не на множестве всех матриц размера $m \times n$, где $m \neq n$).

Свойства операции транспонирования

Свойства операции транспонирования

Пусть A и B — матрицы над (одним и тем же) кольцом R, а $t \in R$. Тогда:

1) $(A^\top)^\top = A$;
2) *если матрицы A и B имеют один и тот же размер, то* $(A+B)^\top = A^\top + B^\top$;
3) $(tA)^\top = tA^\top$;
4) *если произведение матриц AB определено, то* $(AB)^\top = B^\top A^\top$. □

Мы не будем доказывать эти свойства, поскольку они непосредственно вытекают из определений операций.

Рассмотрим теперь матричное уравнение вида $XA = B$, где A и B — известные матрицы, а X — неизвестная. Транспонируя обе части равенства $XA = B$ и используя свойство 4) операции транспонирования, получаем уравнение $A^\top X^\top = B^\top$, т. е. уравнение вида (9). Решив его одним из двух описанных выше способов, мы найдем матрицу $X^\top$ (естественно, второй способ можно применять лишь в случае, когда A — квадратная матрица, после приведения которой к ступенчатому виду не возникает нулевых строк). Используя свойство 4) операции транспонирования, транспонируем матрицу $X^\top$ и найдем искомую матрицу X.

Рассмотрим, наконец, уравнение $AXB = C$, где матрицы A, B и C известны, а матрица X неизвестна. Положим $Y = XB$. Тогда наше уравнение принимает вид $AY = C$. Его можно решить, используя алгоритм 5.1. Обозначим решение этого уравнения через D. Остается решить уравнение $XB = D$, что можно сделать способом, указанным в предыдущем абзаце.

В дальнейшем мы дважды вернемся к рассмотрению матричных уравнений — в § 7 и 23.