

означает, что каждое решение первой системы является решением второй системы, и каждое решение второй системы является решением первой).

Две несовместные системы считаются эквивалентными.

Преобразование, применение которого превращает систему в новую систему, эквивалентную исходной, называется **эквивалентным** или **равносильным преобразованием**. Примерами эквивалентных преобразований могут служить следующие преобразования: перестановка местами двух уравнений системы, перестановка местами двух неизвестных вместе с коэффициентами у всех уравнений, умножение обеих частей какого-либо уравнения системы на отличное от нуля число.

Метод Гаусса решения систем линейных уравнений

Рассмотрим квадратную систему

$$(0.2) \quad \begin{cases} x_1 - x_2 + 3x_3 + 2x_4 = 11 \\ 4x_1 + 6x_2 - x_3 = -1 \\ 3x_1 + 2x_2 + 2x_3 - x_4 = 3 \\ 5x_1 - x_2 + 2x_3 + x_4 = 2 \end{cases}$$

У этой системы коэффициент a_{11} отличен от нуля. Если бы это условие не выполнялось, то чтобы его получить, нужно было бы переставить местами уравнения, поставив первым то уравнение, у которого коэффициент при x_1 не равен нулю.

Проведем следующие преобразования системы:

- 1) поскольку $a_{11} \neq 0$, первое уравнение оставим без изменений;
- 2) вместо второго уравнения запишем уравнение, получающееся, если из второго уравнения вычесть первое, умноженное на 4;
- 3) вместо третьего уравнения запишем разность третьего и первого, умноженного на 3;
- 4) вместо четвертого уравнения запишем разность четвертого и первого, умноженного на 5.

Полученная новая система эквивалентна исходной и имеет во всех уравнениях, кроме первого, нулевые коэффициенты при x_1 (это и являлось целью преобразований 1 – 4):

$$(0.3) \quad \begin{cases} x_1 - x_2 + 3x_3 + 2x_4 = 11 \\ 10x_2 - 13x_3 - 8x_4 = -45 \\ 5x_2 - 7x_3 - 7x_4 = -30 \\ 4x_2 - 13x_3 - 9x_4 = -53 \end{cases}$$

Можно доказать, что замена любого уравнения системы новым, получающимся прибавлением к данному уравнению любого другого уравнения системы, умноженного на любое число, является эквивалентным преобразованием системы.

Для приведенного преобразования и для всех дальнейших преобразований не следует целиком переписывать всю систему, как это только что было сделано. Можно избавиться от необходимости выписывать всякий раз обозначения неизвестных. Для этого коэффициенты системы надо выписывать столбцами, договорившись, что коэффициенты при x_1 всегда располагаются в первом столбце, при x_2 – во втором и т.д. Свободные члены, стоящие в правых частях уравнений, запишем в виде последнего дополнительного столбца. Тогда исходную систему (0.2) можно представить в виде *расширенной матрицы*:

$$A_p = \begin{pmatrix} 1 & -1 & 3 & 2 & 11 \\ 4 & 6 & -1 & 0 & -1 \\ 3 & 2 & 2 & -1 & 3 \\ 5 & -1 & 2 & 1 & 2 \end{pmatrix}.$$

Введем обозначения:

$$A = \begin{pmatrix} 1 & -1 & 3 & 2 \\ 4 & 6 & -1 & 0 \\ 3 & 2 & 2 & -1 \\ 5 & -1 & 2 & 1 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} 11 \\ -1 \\ 3 \\ 2 \end{pmatrix}, \quad \vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

A представляет собой матрицу коэффициентов при неизвестных и называется *матрицей коэффициентов системы* или просто *матрицей системы*. Векторы-столбцы $\vec{b}$, $\vec{x}$ называются *столбцами свободных членов и неизвестных* соответственно. Введенные обозначения позволяют записать исходную систему (0.2) в очень компактной матричной форме:

$$A\vec{x} = \vec{b}.$$

Эта форма записи будет широко использоваться в дальнейших разделах курса. Расширенная матрица системы получается приписыванием справа столбца свободных членов к матрице системы:

$$A_p = (A | \vec{b})$$

Каждую систему m линейных уравнений с n неизвестными можно представить в виде расширенной матрицы, содержащей m строк и $n+1$ столбцов. Каждую матрицу можно считать расширенной матрицей или матрицей коэффициентов некоторой системы линейных уравнений. Системе (0.3) соответствует расширенная матрица

$$\begin{pmatrix} 1 & -1 & 3 & 2 & 11 \\ 0 & 10 & -13 & -8 & -45 \\ 0 & 5 & -7 & -7 & -30 \\ 0 & 4 & -13 & -9 & -53 \end{pmatrix}.$$

Преобразуем эту матрицу следующим образом:

- 1) первые две строки оставим без изменения, поскольку элемент a_{22} не равен нулю;
- 2) вместо третьей строки запишем разность между второй строкой и удвоенной третьей;
- 3) четвертую строку заменим разностью между удвоенной второй строкой и умноженной на 5 четвертой.

В результате получится матрица, соответствующая системе, у которой неизвестная x_1 исключена из всех уравнений, кроме первого, а неизвестная x_2 — из всех уравнений кроме первого и второго:

$$\begin{pmatrix} 1 & -1 & 3 & 2 & 11 \\ 0 & 10 & -13 & -8 & -45 \\ 0 & 0 & 1 & 6 & 15 \\ 0 & 0 & 39 & 29 & 175 \end{pmatrix}.$$

Теперь исключим неизвестную x_3 из четвертого уравнения. Для этого последнюю матрицу преобразуем так:

- 1) первые три строки оставим без изменения, так как $a_{33} \neq 0$;
- 2) из четвертой строки вычтем третью, умноженную на 39:

$$\begin{pmatrix} 1 & -1 & 3 & 2 & 11 \\ 0 & 10 & -13 & -8 & -45 \\ 0 & 0 & 1 & 6 & 15 \\ 0 & 0 & 0 & -205 & -410 \end{pmatrix}.$$

Полученная матрица соответствует системе

$$(0.4) \quad \begin{cases} x_1 - x_2 + 3x_3 + 2x_4 = 11 \\ 10x_2 - 13x_3 - 8x_4 = -45 \\ x_3 + 6x_4 = 15 \\ -205x_4 = -410 \end{cases}$$

Из последнего уравнения этой системы получаем $x_4 = 2$. Подставив это значение в третье уравнение, получим $x_3 = 3$. Теперь из второго уравнения следует, что $x_2 = 1$, а из первого — $x_1 = -1$. Очевидно, что полученное решение единственно (так как единственным образом определяется значение x_4 , затем x_3 и т. д.).

Назовем элементарными преобразованиями матрицы следующие преобразования:

- 1) перемена местами двух строк;
- 2) умножение строки на число, отличное от нуля;
- 3) замена строки матрицы суммой этой строки с любой другой строкой, умноженной на некоторое число.

Если матрица **A** является расширенной матрицей некоторой системы, и путем ряда элементарных преобразований матрица **A** переводится в матрицу **B**, являющуюся расширенной матрицей некоторой другой системы, то эти системы эквивалентны.

Назовем квадратную матрицу, у которой на главной диагонали стоят числа, отличные от нуля, а под главной диагональю – нули, **треугольной матрицей**. Матрица коэффициентов системы (0.4) – треугольная матрица.

Если с помощью элементарных преобразований матрицу коэффициентов квадратной системы можно привести к треугольной матрице, то система совместна и определена.

Рассмотрим другой пример:

$$\begin{cases} x_1 - 2x_2 + 3x_3 - x_4 + 4x_5 = -2 \\ 2x_1 + x_2 + x_3 - 2x_4 + 3x_5 = 1 \\ 3x_1 - x_2 + 2x_3 + x_4 + x_5 = 7 \\ x_1 - 2x_2 + 5x_3 - 5x_4 + 10x_5 = -10 \\ 2x_1 + x_2 - x_3 + 2x_4 - 3x_5 = 9 \end{cases}$$

Проведем следующие преобразования расширенной матрицы системы:

- 1) первую строку оставим без изменения;
- 2) вместо второй строки запишем разность между второй строкой и удвоенной первой;
- 3) вместо третьей строки запишем разность между третьей строкой и утроенной первой;
- 4) четвертую строку заменим разностью между четвертой и первой;
- 5) пятую строку заменим разностью пятой строки и удвоенной первой.

В результате преобразований получим матрицу

$$\begin{pmatrix} 1 & -2 & 3 & -1 & 4 & -2 \\ 0 & 5 & -5 & 0 & -5 & 5 \\ 0 & 5 & -7 & 4 & -11 & 13 \\ 0 & 0 & 2 & -4 & 6 & -8 \\ 0 & 5 & -7 & 4 & -11 & 13 \end{pmatrix}.$$

Оставив без изменения первые две строки этой матрицы, приведем ее элементарными преобразованиями к следующему виду:

$$\begin{pmatrix} 1 & -2 & 3 & -1 & 4 & -2 \\ 0 & 5 & -5 & 0 & -5 & 5 \\ 0 & 0 & 2 & -4 & 6 & -8 \\ 0 & 0 & 2 & -4 & 6 & -8 \\ 0 & 0 & 2 & -4 & 6 & -8 \end{pmatrix}.$$

Если теперь, следуя методу Гаусса, который также называют и методом последовательного исключения неизвестных, с помощью третьей строки привести к нулю коэффициенты при x_3 в четвертой и пятой строках, то после деления всех элементов второй строки на 5 и деления всех элементов третьей строки на 2 получим матрицу

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

Каждая из двух последних строк этой матрицы соответствует уравнению $0x_1+0x_2+0x_3+0x_4+0x_5 = 0$. Это уравнение удовлетворяется любым набором чисел $x_1, x_2, \dots, x_5$, и его следует удалить из системы. Таким образом, система с только что полученной расширенной матрицей эквивалентна системе с расширенной матрицей вида

$$(0.5) \quad \begin{pmatrix} 1 & -2 & 3 & -1 & 4 & -2 \\ 0 & 1 & -1 & 0 & -1 & 1 \\ 0 & 0 & 1 & -2 & 3 & -4 \end{pmatrix}.$$

Последняя строка этой матрицы соответствует уравнению $x_3 - 2x_4 + 3x_5 = -4$. Если неизвестным x_4 и x_5 придать произвольные значения: $x_4 = r$; $x_5 = s$, то из последнего уравнения системы, соответствующей матрице (0.5), получим $x_3 = -4 + 2r - 3s$. Подставив выражения x_3 , x_4 , и x_5 во второе уравнение той же системы, получим $x_2 = -3 + 2r - 2s$. Теперь из первого уравнения можно получить $x_1 = 4 - r + s$. Окончательно решение системы представляется в виде

$$\begin{cases} x_1 = 4 - r + s \\ x_2 = -3 + 2r - 2s \\ x_3 = -4 + 2r - 3s \\ x_4 = r \\ x_5 = s \end{cases}.$$

Рассмотрим прямоугольную матрицу $\mathbf{A}$, у которой число столбцов m больше, чем число строк n . Если матрицу $\mathbf{A}$ можно разделить вертикальной чертой на две матрицы: стоящую слева треугольную матрицу размера m и стоящую справа прямоугольную матрицу, то матрицу $\mathbf{A}$ назовем **трапециевидной** или **трапецеидальной**. Очевидно, что матрица (0.5) — трапециевидная матрица.

Если при применении эквивалентных преобразований к системе уравнений хотя бы одно уравнение приводится к виду

$$0x_1 + 0x_2 + \dots + 0x_n = b_j \quad (b_j \neq 0),$$

то система несовместна или противоречива, так как ни один набор чисел $x_1, x_2, \dots, x_n$ не удовлетворяет этому уравнению.

Если при преобразовании расширенной матрицы системы матрица коэффициентов приводится к трапецеидальному виду и при этом система не

получается противоречивой, то система совместна и является неопределенной, то есть имеет **бесконечно много решений**.

В последней системе можно получить все решения, придавая конкретные числовые значения параметрам r и s .

Те переменные, коэффициенты при которых стоят на главной диагонали трапецеидальной матрицы (это значит, что эти коэффициенты отличны от нуля), называются **базисными**. В рассмотренном выше примере это неизвестные x_1, x_2, x_3 . Остальные неизвестные называются **свободными**. В рассмотренном выше примере это неизвестные x_4 и x_5 . Свободным неизвестным можно придавать любые значения или выражать их через параметры, как это сделано в последнем примере.

Базисные неизвестные единственным образом выражаются через свободные неизвестные.

Если свободным неизвестным приданы конкретные числовые значения и через них выражены базисные неизвестные, то полученное решение называется **частным решением**.

Если свободные неизвестные выражены через параметры, то получается решение, которое называется **общим решением**.

Все бесконечное множество решений системы можно получить, придавая свободным неизвестным любые числовые значения и находя соответствующие значения базисных неизвестных.

Если всем свободным неизвестным приданы нулевые значения, то полученное решение называется **базисным**.

Одну и ту же систему иногда можно привести к разным наборам базисных неизвестных. Так, например, можно поменять местами 3-й и 4-й столбцы в матрице (0.5). Тогда базисными будут неизвестные x_1, x_2, x_4 , а свободными – x_3 и x_5 .

Если получены два различных набора базисных неизвестных при различных способах нахождения решения одной и той же системы, то эти

наборы обязательно содержат одно и то же число неизвестных, называемое **рангом системы**.

Рассмотрим еще одну систему, имеющую бесконечно много решений:

$$\begin{cases} x_1 - x_2 + 2x_3 + 4x_4 + x_5 = 2 \\ 2x_1 + 3x_3 + 7x_4 - x_5 = 7 \\ 2x_1 - 4x_2 + 5x_3 + 9x_4 + 5x_5 = 1 \\ 3x_1 - x_2 + 5x_3 + 14x_4 - 2x_5 = 8 \end{cases}$$

Проведем преобразование расширенной матрицы системы по методу Гаусса:

$$\begin{pmatrix} 1 & -1 & 2 & 4 & 1 & 2 \\ 0 & 2 & -1 & -1 & -3 & 3 \\ 0 & 2 & -1 & -1 & -3 & 3 \\ 0 & 2 & -1 & 2 & -5 & 2 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & -1 & 2 & 4 & 1 & 2 \\ 0 & 2 & -1 & -1 & -3 & 3 \\ 0 & 0 & 0 & 3 & -2 & -1 \end{pmatrix}.$$

Как видно, мы не получили трапецеидальной матрицы, однако последнюю матрицу можно преобразовать, поменяв местами третий и четвертый столбцы:

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

Эта матрица уже является трапецеидальной. У соответствующей ей системы две свободных неизвестных – x_3 , x_5 и три базисных – x_1 , x_2 , x_4 . (Ранг системы равен 3). Решение исходной системы представляется в следующем виде:

$$\begin{cases} x_1 = \frac{14}{3} - \frac{3}{2}r - \frac{11}{6}s \\ x_2 = \frac{4}{3} + \frac{1}{2}r + \frac{11}{6}s \\ x_3 = r \\ x_4 = -\frac{1}{3} + \frac{2}{3}s \\ x_5 = s \end{cases}.$$

Приведем пример не имеющей решения системы:

$$\begin{cases} 2x_1 - 3x_2 + x_3 = 7 \\ 3x_1 + 2x_2 - x_3 = 5 \\ 4x_1 + 7x_2 - 3x_3 = 4 \end{cases}.$$

Преобразуем матрицу системы по методу Гаусса:

$$\begin{pmatrix} 2 & -3 & 1 & 7 \\ 3 & 2 & -1 & 5 \\ 4 & 7 & -3 & 4 \end{pmatrix} \Rightarrow \begin{pmatrix} 2 & -3 & 1 & 7 \\ 0 & 13 & -5 & -11 \\ 0 & 13 & -5 & -10 \end{pmatrix} \Rightarrow \begin{pmatrix} 2 & -3 & 1 & 7 \\ 0 & 13 & -5 & -11 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Последняя строка последней матрицы соответствует не имеющему решения уравнению $0x_1 + 0x_2 + 0x_3 = 1$. Следовательно, исходная система несовместна.

Сформулируем теперь кратко суть метода Гаусса. Полагая, что в системе коэффициент a_{11} отличен от нуля (если это не так, то следует на первое место поставить уравнение с отличным от нуля коэффициентом при x_1 и переобозначить коэффициенты), преобразуем систему следующим образом: первое уравнение оставляем без изменения, а из всех остальных уравнений исключаем неизвестную x_1 с помощью эквивалентных преобразований описанным выше способом.

В полученной системе

[illegible]

считая, что $a_{22}^* \neq 0$ (что всегда можно получить, переставив уравнения или слагаемые внутри уравнений и переобозначив коэффициенты системы), оставляем без изменений первые два уравнения системы, а из остальных уравнений, используя второе уравнения, с помощью элементарных преобразований исключаем неизвестную x_2 . Во вновь полученной системе

$$\left\{ \begin{array}{l} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots a_{1n}x_n = b_1 \\ a_{22}^*x_2 + a_{23}^*x_3 + \dots a_{2n}^*x_n = b_2^* \\ a_{33}^{**}x_3 + \dots a_{3n}^{**}x_n = b_3^{**} \\ \dots\dots\dots \\ a_{m3}^{**}x_3 + \dots a_{mn}^{**}x_n = b_m^{**} \end{array} \right.$$

при условии $a_{33}^{**} \neq 0$ оставляем без изменений первые три уравнения, а из всех остальных с помощью третьего уравнения элементарными преобразованиями исключаем неизвестную x_3 .

Этот процесс продолжается до тех пор, пока не реализуется один из трех возможных случаев:

1) если в результате приходим к системе, одно из уравнений которой имеет нулевые коэффициенты при всех неизвестных и отличный от нуля свободный член, то исходная система несовместна;

2) если в результате преобразований получаем систему с матрицей коэффициентов треугольного вида, то система совместна и является определенной;

3) если получается система с трапецеидальной матрицей коэффициентов (и при этом не выполняется условие пункта 1), то система совместна и неопределенна.

ЗАДАЧИ И УПРАЖНЕНИЯ 3.

Решить методом Гаусса:

$$1. \begin{cases} x_1 + 2x_2 + x_3 = 8 \\ -x_1 + x_2 - x_3 = -2 \\ 3x_1 + 3x_2 = 9 \end{cases}$$

$$2. \begin{cases} x_1 + x_2 + x_3 = 2 \\ -x_1 + x_2 + 2x_3 = 2 \\ x_1 + 3x_2 + 2x_3 = 6 \end{cases}$$

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

$$4. \begin{cases} x_1 + 2x_2 + x_3 = 5 \\ -x_1 + 2x_2 + 2x_3 = -3 \\ 2x_1 + 3x_2 + 2x_3 = 8 \end{cases}$$

$$5. \begin{cases} x_1 + x_2 + x_3 = 6 \\ 2x_1 + 2x_2 + 3x_3 = 15 \\ 3x_1 - 3x_2 - x_3 = -6 \\ 5x_1 - 2x_3 = -1 \end{cases}$$

$$6. \begin{cases} x_1 + 2x_2 + 3x_3 = 8 \\ 2x_1 + 2x_2 + 3x_3 = 9 \\ 3x_1 - 3x_2 - x_3 = -4 \\ 2x_1 - 3x_2 - x_3 = -10 \end{cases}$$

$$7. \begin{cases} x_1 + 2x_2 + 3x_3 = 4 \\ 2x_1 + x_2 - x_3 = 3 \\ 3x_1 + 3x_2 + 2x_3 = 7 \end{cases}$$

$$8. \begin{cases} x_1 - 2x_2 + x_3 = 4 \\ 2x_1 + 3x_2 - x_3 = 3 \\ 4x_1 - x_2 + x_3 = 11 \end{cases}$$

$$9. \begin{cases} x_1 + x_3 - 2x_4 = 0 \\ 2x_1 + x_2 + 2x_3 + x_4 = 6 \\ x_1 - x_2 - 2x_3 + x_4 = -4 \end{cases}$$

Сколько в системе базисных неизвестных?

Сколько свободных?

Укажите частное решение, удовлетворяющее условию $x_1=0$.

Укажите общее решение системы.

Ответы

1) (1,2,3); 2) (0,2,0); 3) (-1,2,1); 4) (3,2,-2); 5) (1,2,3); 6) система несовместна;
7) $(2+5z)/3, (5-7z)/3, z$; 8) $(t, 7-3t, 18-7t)$; 9) 3, 1, $(0,1,2,1)$, $((2-2t)/3, 6-5t, (8t-2)/3, t)$.

Матричные уравнения

Как уже отмечалось ранее, систему из m уравнений с n неизвестными можно представить в матричном виде:

$$(0.6) \quad A\vec{x} = \vec{b}$$

где

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}; \quad \vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}; \quad \vec{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

Ниже будет показано, что, записывая систему в сжатом виде, кроме краткости написания мы получаем и другие очень важные преимущества.

Пусть имеются две квадратные матрицы одинаковой размерности:

$$A = \begin{pmatrix} 1 & 2 & 1 \\ -1 & 3 & 2 \\ 2 & 3 & 1 \end{pmatrix}; \quad B = \begin{pmatrix} 2 & 1 & 7 \\ 3 & -5 & 9 \\ -1 & 4 & 6 \end{pmatrix}.$$

Требуется найти матрицу $\mathbf{X}$, удовлетворяющую матричному уравнению

$$\mathbf{AX} = \mathbf{B}.$$

Из правила умножения матриц следует, что матрица $\mathbf{X}$ должна быть квадратной матрицей той же размерности, что и матрицы $\mathbf{A}$ и $\mathbf{B}$:

$$\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{pmatrix}.$$

Из правила умножения матриц и из определения равенства матриц следует, что последнее матричное уравнение распадается на три системы линейных уравнений:

$$(0.7) \quad \begin{cases} x_{11} + 2x_{21} + x_{31} = 2 \\ -x_{11} + 3x_{21} + 2x_{31} = 3 \\ 2x_{11} + 3x_{21} + x_{31} = -1 \end{cases}; \quad \begin{cases} x_{12} + 2x_{22} + x_{32} = 1 \\ -x_{12} + 3x_{22} + 2x_{32} = -5 \\ 2x_{12} + 3x_{22} + x_{32} = 4 \end{cases}$$

$$\begin{cases} x_{13} + 2x_{23} + x_{33} = 7 \\ -x_{13} + 3x_{23} + 2x_{33} = 9 \\ 2x_{13} + 3x_{23} + x_{33} = 6 \end{cases}.$$

Все три системы (0.7) имеют одинаковые матрицы коэффициентов, что дает возможность решать их одновременно, введя матрицу

$$\begin{pmatrix} 1 & 2 & 1 & 2 & 1 & 7 \\ -1 & 3 & 2 & 3 & -5 & 9 \\ 2 & 3 & 1 & -1 & 4 & 6 \end{pmatrix}.$$

Здесь первые четыре столбца образуют расширенную матрицу первой системы, первые три столбца вместе с пятым столбцом образуют расширенную матрицу второй системы, а первые три столбца вместе с шестым – расширенную матрицу третьей системы.

Применим для решения **метод Жордана-Гаусса** который является модификацией метода Гаусса.

Первый шаг преобразования матрицы по методу Жордана-Гаусса совпадает с первым шагом преобразований по методу Гаусса. Оставляем без изменений первую строку матрицы, а во второй и третьей “организуем” нули в первом столбце:

$$\begin{pmatrix} 1 & 2 & 1 & 2 & 1 & 7 \\ 0 & 5 & 3 & 5 & -4 & 16 \\ 0 & 1 & 1 & 5 & -2 & 8 \end{pmatrix}.$$

Теперь, следуя методу Жордана-Гаусса, **оставляем без изменения лишь вторую строку** (так как $a_{22} \neq 0$) и получаем с помощью второй строки в **первой и третьей строках во втором столбце нули**. Для этого вместо первой строки пишем сумму первой строки, умноженной на 5, и второй строки, умноженной на -2 . Вместо третьей строки пишем сумму третьей строки, умноженной на 5, и второй строки, умноженной на -1 . После деления полученной третьей строки на 2 получаем матрицу

$$\begin{pmatrix} 5 & 0 & -1 & 0 & 13 & 3 \\ 0 & 5 & 3 & 5 & -4 & 16 \\ 0 & 0 & 1 & 10 & -3 & 12 \end{pmatrix}.$$

Чтобы в первой и второй строках в третьем столбце получить нули, проведем следующие преобразования последней матрицы. Оставив третью строку без изменений, заменим вторую строку разностью второй строки и утроенной

третьей, а первую – суммой первой и третьей строк. После деления первой и второй строк преобразованной матрицы на 5 получится матрица

$$(0.8) \quad \begin{pmatrix} 1 & 0 & 0 & 2 & 2 & 3 \\ 0 & 1 & 0 & -5 & 1 & -4 \\ 0 & 0 & 1 & 10 & -3 & 2 \end{pmatrix}.$$

При преобразовании системы по методу Жордана-Гаусса матрица коэффициентов приводится (если это возможно) к единичной матрице.

Если взять первые четыре столбца матрицы (3), то получится матрица, в которую преобразовалась расширенная матрица первой из систем уравнений (2). Из нее следует: $x_{11}=2$; $x_{21}=-5$; $x_{31}=10$. Матрица, образованная первыми тремя столбцами вместе с пятым столбцом матрицы (3), дает решение второй системы уравнений (2): $x_{12}=2$; $x_{22}=1$; $x_{32}=-3$. И, наконец, матрица, образованная первыми тремя столбцами вместе с шестым столбцом матрицы (3), дает решение третьей системы уравнений (2): $x_{13}=3$; $x_{23}=-4$; $x_{33}=12$.

Из сказанного можно сделать очень интересный и важный вывод: последние три столбца матрицы (3) образуют искомую матрицу X .

$$X = \begin{pmatrix} 2 & 2 & 3 \\ -5 & 1 & -4 \\ 10 & -3 & 12 \end{pmatrix}.$$

Нулевой матрицей называется матрица, у которой все элементы – нули. Очевидно равенство $A + (-1)A = 0$. Здесь в правой части через 0 обозначена нулевая матрица той же размерности, что и матрица A .

Напомним, что квадратная матрица E размера n называется **единичной**, если все её элементы, стоящие на главной диагонали, равны единице, а все остальные – нули.

$$\mathbf{E} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}.$$

Легко проверить справедливость равенств: $\mathbf{E}\mathbf{A} = \mathbf{A}\mathbf{E} = \mathbf{A}$. Здесь $\mathbf{A}$ – квадратная матрица, и размеры $\mathbf{A}$ и $\mathbf{E}$ одинаковы.

Пусть $\mathbf{A}$ – квадратная матрица. **Обратной матрицей** к матрице $\mathbf{A}$ называется такая матрица $\mathbf{A}^{-1}$, для которой справедливы равенства:

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{E}.$$

Очевидно, что $\mathbf{A}^{-1}$ – квадратная матрица того же размера, что и матрица $\mathbf{A}$. Сразу заметим, что **не всякая квадратная матрица имеет обратную матрицу**.

Поставим задачу: найти обратную матрицу к матрице

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

Условие

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 4 \\ 2 & -1 & 3 \end{pmatrix} \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

где

$$\begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{pmatrix} = \mathbf{A}^{-1},$$

сводится к трём системам уравнений, которые будем решать одновременно, используя метод Жордана-Гаусса. Матрица, представляющая расширенные матрицы всех трёх систем, примет вид

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

Подвергая её преобразованиям по методу Жордана-Гаусса, последовательно будем получать:

$$\begin{aligned} \begin{pmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -1 & 1 & -1 & 0 \\ 0 & 5 & 3 & 2 & 0 & -1 \end{pmatrix} &\Rightarrow \begin{pmatrix} 1 & 0 & 5 & -1 & 2 & 0 \\ 0 & 1 & -1 & 1 & -1 & 0 \\ 0 & 0 & 8 & -3 & 5 & 1 \end{pmatrix} \Rightarrow \\ \Rightarrow \begin{pmatrix} 8 & 0 & 0 & 7 & -9 & 5 \\ 0 & 8 & 0 & 5 & -3 & 1 \\ 0 & 0 & 8 & -3 & 5 & -1 \end{pmatrix} &\Rightarrow \begin{pmatrix} 1 & 0 & 0 & \frac{7}{8} & -\frac{9}{8} & \frac{5}{8} \\ 0 & 1 & 0 & \frac{5}{8} & -\frac{3}{8} & -\frac{1}{8} \\ 0 & 0 & 1 & -\frac{3}{8} & \frac{5}{8} & -\frac{1}{8} \end{pmatrix} \end{aligned} \quad (4)$$

Как и в предыдущем примере, можно сказать, что три последних столбца образуют искомую матрицу, то есть

$$\mathbf{A}^{-1} = \begin{pmatrix} \frac{7}{8} & -\frac{9}{8} & \frac{5}{8} \\ \frac{5}{8} & -\frac{3}{8} & -\frac{1}{8} \\ -\frac{3}{8} & \frac{5}{8} & -\frac{1}{8} \end{pmatrix}.$$

Теперь сформулируем правило, по которому находится матрица, обратная к квадратной матрице $\mathbf{A}$ размера n .

Нужно выписать матрицу размерности $n \times 2n$, первые n столбцов которой образованы матрицей $\mathbf{A}$, а последние n столбцов образуют единичную матрицу $\mathbf{E}$. Построенная таким образом матрица преобразуется по методу Жордана-Гаусса так, чтобы на месте матрицы $\mathbf{A}$ получилась единичная матрица, если это возможно. Тогда на месте матрицы $\mathbf{E}$ получается матрица $\mathbf{A}^{-1}$.

Если матрицу A нельзя методом Жордана-Гаусса преобразовать к единичной матрице, то A^{-1} не существует. Так матрица

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 1 \\ 4 & 5 & 7 \end{pmatrix}$$

не имеет обратной. Читатель может в этом убедиться самостоятельно.

Решение систем линейных уравнений с использованием обратной матрицы.

Пусть имеется матричное уравнение

$$(0.9) \quad AX=B$$

с квадратной матрицей A , для которой существует обратная матрица A^{-1} . (X и B могут быть векторами или матрицами подходящих размеров.)

Умножим уравнение (0.9) *слева* на матрицу A^{-1} и проведем очевидные преобразования:

$$(0.10) \quad A^{-1}AX = A^{-1}B \Rightarrow EX = A^{-1}B \Rightarrow X = A^{-1}B$$

Таким образом, для нахождения решения системы (0.9) достаточно найти матрицу, обратную к A , и умножить ее на матрицу (столбец) B . При ручной реализации данный метод решения не дает каких-либо преимуществ по сравнению с методом Жордана-Гаусса, однако, он с успехом реализуется на компьютере, так как операции получения обратной матрицы и умножения матриц легко формализуются. В частности данный способ решения матричных уравнений легко реализуется при помощи табличного процессора Excel.

Пример. Требуется решить матричное уравнение

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

Используя процедуру, описанную выше, находим обратную матрицу:

$$\begin{pmatrix} 1 & 0 \\ -3 & 2 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 \\ 1,5 & 0,5 \end{pmatrix}$$

Далее

$$X = \begin{pmatrix} 1 & 0 \\ 1,5 & 0,5 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & -4 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix}$$

ЗАДАЧИ И УПРАЖНЕНИЯ 4.

Решить матричные уравнения методом Жордана-Гаусса:

1. $\begin{pmatrix} 1 & 3 \\ 0 & 2 \end{pmatrix} X = \begin{pmatrix} 4 & 4 \\ 2 & 2 \end{pmatrix}$

6. $\begin{pmatrix} 1 & 0 & 4 \\ 0 & 1 & 6 \\ 5 & 2 & 0 \end{pmatrix} X = \begin{pmatrix} 1 & 3 & 5 \\ 0 & 4 & 7 \\ 5 & 23 & 7 \end{pmatrix}$

2. $\begin{pmatrix} 1 & -2 \\ 0 & -1 \end{pmatrix} X = \begin{pmatrix} -3 & -5 \\ -2 & -4 \end{pmatrix}$

7. $\begin{pmatrix} 1 & 7 & -3 \\ 3 & -1 & 6 \\ -2 & 2 & 8 \end{pmatrix} X = \begin{pmatrix} 18 & -4 & -3 \\ -5 & -5 & 6 \\ -6 & -10 & 8 \end{pmatrix}$

3. $\begin{pmatrix} 1 & -1 \\ 3 & 2 \end{pmatrix} X = \begin{pmatrix} 2 & 4 \\ 1 & 2 \end{pmatrix}$

4. $\begin{pmatrix} -1 & 3 \\ -2 & 2 \end{pmatrix} X = \begin{pmatrix} 12 & -2 \\ 8 & -8 \end{pmatrix}$

8. $\begin{pmatrix} 1 & 0 & 4 \\ 1 & -3 & 6 \\ 5 & 2 & 0 \end{pmatrix} X = \begin{pmatrix} 1 & 0 & 12 \\ 1 & -6 & 18 \\ 5 & 4 & 0 \end{pmatrix}$

5. $\begin{pmatrix} 1 & 7 & 4 \\ 3 & 9 & 6 \\ 5 & 2 & 8 \end{pmatrix} X = \begin{pmatrix} 15 & 31 & 4 \\ 21 & 45 & 6 \\ 9 & 23 & 8 \end{pmatrix}$

При помощи метода Жордана-Гаусса найти обратную матрицу:

9. $\begin{pmatrix} 1 & 3 \\ 0 & 2 \end{pmatrix}$

13. $\begin{pmatrix} 1 & 3 & 4 \\ 3 & 0 & 6 \\ 0 & 2 & 3 \end{pmatrix}$

10. $\begin{pmatrix} 1 & -2 \\ 0 & -1 \end{pmatrix}$

14. $\begin{pmatrix} 2 & 3 & 4 \\ 0 & 2 & 6 \\ 0 & 0 & 3 \end{pmatrix}$

11. $\begin{pmatrix} 1 & -1 \\ 3 & 2 \end{pmatrix}$

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

12. $\begin{pmatrix} -1 & 3 \\ -2 & 2 \end{pmatrix}$

$$16. \begin{pmatrix} 2 & 3 & 0 \\ 1 & 2 & 6 \\ -1 & -2 & 3 \end{pmatrix}$$

Методом обратной матрицы решить системы уравнений:

$$17. \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} X = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

$$18. \begin{pmatrix} 2 & 3 & 0 \\ 1 & 2 & 6 \\ -1 & -2 & 3 \end{pmatrix} X = \begin{pmatrix} 1 \\ 9 \\ 9 \end{pmatrix}$$

Ответы:

$$1) \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}; \quad 2) \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}; \quad 3) \begin{pmatrix} 1 & 2 \\ -1 & -2 \end{pmatrix}; \quad 4) \begin{pmatrix} 0 & 5 \\ 4 & 1 \end{pmatrix}; \quad 5) \begin{pmatrix} 1 & 3 & 0 \\ 2 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix}; \quad 6)$$

$$\begin{pmatrix} 1 & 3 & 1 \\ 0 & 4 & 1 \\ 0 & 0 & 1 \end{pmatrix}; \quad 7) \begin{pmatrix} 1 & 0 & 0 \\ 2 & -1 & 0 \\ -1 & -1 & 1 \end{pmatrix}; \quad 8) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}; \quad 9) \begin{pmatrix} 1 & -1,5 \\ 0 & 0,5 \end{pmatrix}; \quad 10) \begin{pmatrix} 1 & -2 \\ 0 & -1 \end{pmatrix};$$

$$11) \begin{pmatrix} 0,4 & 0,2 \\ -0,6 & 0,2 \end{pmatrix}; \quad 12) \begin{pmatrix} 0,5 & -0,75 \\ 0,5 & -0,25 \end{pmatrix}; \quad 13) \begin{pmatrix} 4/5 & 1/15 & -6/5 \\ 3/5 & -1/5 & -2/5 \\ -2/5 & 2/15 & 3/5 \end{pmatrix}; \quad 14)$$

$$\begin{pmatrix} 1/2 & -3/4 & 5/6 \\ 0 & 1/2 & -1 \\ 0 & 0 & 1/3 \end{pmatrix}; \quad 15) \begin{pmatrix} 6/31 & -9/31 & 10/31 \\ 9/31 & 2/31 & -16/31 \\ -2/31 & 3/31 & 7/31 \end{pmatrix}; \quad 16) \begin{pmatrix} 2 & -1 & 2 \\ -1 & 2/3 & -4/3 \\ 0 & 1/9 & 1/9 \end{pmatrix};$$

$$17) (0,8; 0,6); \quad 18) (11, -7, 2).$$

Определители

Рассмотрим систему двух линейных уравнений с двумя неизвестными в общем виде:

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

Найдем x_1 следующим образом: чтобы исключить x_2 , умножим первое уравнение на a_{22} и из полученного уравнения вычтем второе, умноженное на a_{12} :

$$(0.11) \quad (a_{11}a_{22} - a_{12}a_{21})x_1 = b_1a_{22} - b_2a_{12}.$$

Обозначим $\Delta = a_{11}a_{22} - a_{12}a_{21}$, $\Delta_1 = b_1a_{22} - b_2a_{12}$.

Для определения x_2 поступим так: умножим второе уравнение на a_{11} и из полученного уравнения вычтем первое, умноженное на a_{21} :

$$(0.12) \quad (a_{11}a_{22} - a_{12}a_{21})x_2 = a_{11}b_2 - a_{21}b_1.$$

Обозначим $\Delta_2 = a_{11}b_2 - a_{21}b_1$.

(0.11) и (0.12) можно переписать следующим образом:

$$(0.13) \quad \begin{cases} \Delta x_1 = \Delta_1 \\ \Delta x_2 = \Delta_2 \end{cases}$$

Из (0.13) видно, что если $\Delta \neq 0$, то каждое уравнение можно поделить на Δ , и система имеет в этом случае единственное решение, определяемое формулой

$$x_i = \frac{\Delta_i}{\Delta}, i = 1, 2.$$

Если же $\Delta=0$, то делить на Δ нельзя, и возможны два случая:

1. $\Delta_1=\Delta_2=0$ — система совместна, но не определена (имеет бесконечное множество решений);
2. $\Delta_1 \neq 0$ или $\Delta_2 \neq 0$ — система несовместна.

Величина Δ называется **определителем квадратной матрицы второго порядка**

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

Вообще **определителем произвольной квадратной матрицы второго порядка** $\begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix}$ называется число, которое обозначается $\begin{vmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{vmatrix}$ и равно произведению двух чисел, стоящих на главной диагонали минус произведение двух чисел, стоящих на другой диагонали: $\alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21}$.

Например,

$$\begin{vmatrix} 3 & -4 \\ -2 & -5 \end{vmatrix} = (3 \times (-5)) - ((-2) \times (-4)) = -15 - 8 = -23.$$

Из сказанного следует, что величины Δ_1 и Δ_2 в (3) тоже являются определителями:

$$\Delta_1 = \begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}; \quad \Delta_2 = \begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}.$$

Эти определители получаются из определителя Δ путем замены первого или второго столбца соответственно на столбец свободных членов.

Оказывается, что при решении квадратных систем порядка большего, чем 2, получается совершенно аналогичная картина. Решение квадратной системы порядка n может быть сведено тем же способом к системе

$$\Delta x_i = \Delta_i, \quad i = 1, \dots, n$$

где $\Delta = |A|$, а Δ_i получается из Δ заменой i -го столбца на столбец свободных членов. Если определитель системы Δ не равен нулю, то единственное решение системы задается *формулами Крамера*:

$$(0.14) \quad x_i = \frac{\Delta_i}{\Delta}, \quad i = 1, \dots, n$$

До сих пор было показано, как вычислять определитель второго порядка. Чтобы вычислить определитель более высоких порядков, пользуются формулой Лапласа разложения определителя по строке или столбцу.

Пусть дан определитель n -го порядка (международное название определителя – determinant):

$$\det A = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

тогда его разложение по *первой* строке выглядит следующим образом:

$$\det A = a_{11}A_{11} + a_{12}A_{12} + \dots + a_{1n}A_{1n} = \sum_{j=1}^n a_{1j}A_{1j}$$

где A_{ij} называется **алгебраическим дополнением** элемента a_{ij} .
Алгебраическое дополнение вычисляется по следующим правилам:

$$A_{ij} = (-1)^{i+j} M_{ij}$$

где M_{ij} называется **минором** и равняется определителю порядка $(n-1)$, который получается из определителя $\det A$, если вычеркнуть из него i -ю строку и j -й столбец. В общем случае раскрытия по i -ой строке справедлива формула:

$$(0.15) \quad \det A = a_{i1}(-1)^{i+1}M_{i1} + a_{i2}(-1)^{i+2}M_{i2} + \dots + a_{in}(-1)^{i+n}M_{in}$$

А в случае раскрытия по j -му столбцу:

$$(0.16) \quad \det A = a_{1j}(-1)^{1+j}M_{1j} + a_{2j}(-1)^{2+j}M_{2j} + \dots + a_{nj}(-1)^{n+j}M_{nj}$$

Пример. Пусть

$$\det A = \begin{vmatrix} 1 & -1 & 2 \\ 2 & 1 & 0 \\ 3 & 0 & 4 \end{vmatrix}$$

Разложим его по первому столбцу.

$$\begin{aligned} \det A &= 1 \cdot (-1)^{1+1} \begin{vmatrix} 1 & 0 \\ 0 & 4 \end{vmatrix} + 2 \cdot (-1)^{2+1} \begin{vmatrix} -1 & 2 \\ 0 & 4 \end{vmatrix} + 3 \cdot (-1)^{3+1} \begin{vmatrix} -1 & 2 \\ 1 & 0 \end{vmatrix} = \\ &= (4 - 0) - 2(-4 - 0) + 3(0 - 2) = 4 + 8 - 6 = 6 \end{aligned}$$

Приведенный способ раскрытия был не самым эффективным. Выгоднее было бы раскрыть определитель по второму столбцу, так как он содержит ноль. Вообще, чем больше нулей содержит строка (столбец) определителя, тем выгоднее ее использовать для раскрытия.

Определитель обладает следующими свойствами:

1. Если поменять местами две строки определителя (два столбца), то получим новый определитель, равный исходному, умноженному на (-1) .

2. Определитель, имеющий две равных строки (два равных столбца), равен нулю.

3. Если одну из строк определителя умножить на какое-либо число, то получится определитель, равный исходному, умноженному на это число.

4. Определитель транспонированной матрицы равен определителю исходной матрицы.

5. Если в определителе вместо любой строки записать сумму этой строки и любой другой строки, умноженной на некоторое число, то полученный новый определитель будет равен исходному.

Вернемся к последнему примеру. Прежде чем раскрывать определитель, используя правило 5, прибавим к первой строке вторую.

$$\det A = \begin{vmatrix} 1 & -1 & 2 \\ 2 & 1 & 0 \\ 3 & 0 & 4 \end{vmatrix} \stackrel{I+II}{=} \begin{vmatrix} 3 & 0 & 2 \\ 2 & 1 & 0 \\ 3 & 0 & 4 \end{vmatrix}$$

Теперь раскроем определитель по второму столбцу, так как он содержит наибольшее число нулей.

$$\det A = 1 \cdot (-1)^{2+2} \begin{vmatrix} 3 & 2 \\ 3 & 4 \end{vmatrix} = 12 - 6 = 6$$

Пример. Пусть Δ – определитель четвертого порядка:

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

Этот определитель разложим по третьей строке, так как там есть нуль и, что особенно важно, -1 . Если сделать это сразу, то придется считать три определителя третьего порядка. Поэтому предварительно увеличим количество нулей. Задача заключается в таком преобразовании определителя Δ , чтобы получить нули на месте a_{31} и a_{33} . К первому столбцу прибавим

второй столбец, умноженный на -2 , а к третьему столбцу прибавим второй столбец, умноженный на 3 . Второй столбец, с помощью которого проводились преобразования, остается без изменений.

Таким образом, вычисление определителя 4-го порядка сведено к вычислению только одного определителя 3-го порядка:

$$\Delta = \begin{vmatrix} -4 & 3 & 5 & 5 \\ -1 & 2 & 7 & 6 \\ 0 & -1 & 0 & 0 \\ -2 & 3 & 11 & 7 \end{vmatrix} = (-1)(-1)^{3+2} \begin{vmatrix} -4 & 5 & 5 \\ -1 & 7 & 6 \\ -2 & 11 & 7 \end{vmatrix}.$$

Пользуясь свойствами определителя и методом Лапласа, можно вычисление определителя n -го порядка свести к вычислению лишь одного определителя порядка $(n - 1)$.

Вернемся к решению СЛАУ по формулам Крамера (0.14). Пусть, например,

$$\begin{cases} 3x_1 + 2x_2 - x_3 = 2 \\ 2x_1 - x_2 + 2x_3 = 3; \\ 4x_1 + x_2 + 3x_3 = 5 \end{cases}$$

$$\Delta = \begin{vmatrix} 3 & 2 & -1 \\ 2 & -1 & 2 \\ 4 & 1 & 3 \end{vmatrix} = -17; \quad \Delta_1 = \begin{vmatrix} 2 & 2 & -1 \\ 3 & -1 & 2 \\ 5 & 1 & 3 \end{vmatrix} = -16;$$

$$\Delta_2 = \begin{vmatrix} 3 & 2 & -1 \\ 2 & 3 & 2 \\ 4 & 5 & 3 \end{vmatrix} = 3; \quad \Delta_3 = \begin{vmatrix} 3 & 2 & 2 \\ 2 & -1 & 3 \\ 4 & 1 & 5 \end{vmatrix} = -8;$$

$$x_1 = \frac{16}{17}; \quad x_2 = -\frac{3}{17}; \quad x_3 = \frac{8}{17}.$$

Вычисление обратной матрицы

Пусть $\mathbf{A} = (a_{ij})$ – квадратная матрица с определителем, не равным нулю. Тогда существует обратная матрица $\mathbf{A}^{-1}$, которая вычисляется по формуле

$$\mathbf{A}^{-1} = (c_{ij}) = \left(\frac{A_{ji}}{\det \mathbf{A}} \right).$$

Последняя формула означает, что **в i -й строке и j -м столбце** обратной матрицы располагается алгебраическое дополнение элемента, стоящего **в j -й строке и в i -м столбце** исходной матрицы, деленное на определитель исходной матрицы.

Напомним здесь, что $A_{pq} = (-1)^{p+q} M_{pq}$, где M_{pq} называется минором и представляет собой определитель, получающийся из определителя $\det \mathbf{A}$ вычеркиванием p -й строки и q -го столбца.

Рассмотрим пример:

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

$$\begin{aligned} A_{11} &= 6, & A_{12} &= 0, & A_{13} &= 0, \\ A_{21} &= -3, & A_{22} &= 3, & A_{23} &= 0, \\ A_{31} &= 0, & A_{32} &= -2, & A_{33} &= 2; \end{aligned} \quad \mathbf{A}^{-1} = \begin{pmatrix} 1 & -1/2 & 0 \\ 0 & 1/2 & -1/3 \\ 0 & 0 & 1/3 \end{pmatrix}.$$

Еще раз подчеркнем, что обратная матрица существует только для квадратной матрицы с определителем, отличным от нуля!

ЗАДАЧИ И УПРАЖНЕНИЯ 5.

Вычислить определители:

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

$$3. \begin{vmatrix} 2 & 4 & 5 \\ 1 & -3 & 6 \\ 2 & 4 & -1 \end{vmatrix}$$

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

$$4. \begin{vmatrix} 3 & 4 & 6 \\ 3 & 2 & 6 \\ -1 & 3 & -2 \end{vmatrix}$$

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

$$9. \begin{vmatrix} 3 & 1 & 3 & -1 \\ 3 & 1 & 3 & -3 \\ -1 & 0 & 1 & 2 \\ -1 & 0 & 2 & 1 \end{vmatrix}$$

$$6. \begin{vmatrix} 2 & 4 & 0 \\ 1 & 2 & 6 \\ -1 & 3 & 0 \end{vmatrix}$$

$$10. \begin{vmatrix} 2 & 0 & 4 & 0 \\ 3 & 1 & 3 & -3 \\ -1 & 0 & 1 & 2 \\ -1 & 2 & 2 & 1 \end{vmatrix}$$

$$7. \begin{vmatrix} 1 & 0 & 4 & 0 \\ 3 & 1 & 3 & -3 \\ -1 & 0 & 1 & 2 \\ 2 & 2 & 2 & 1 \end{vmatrix}$$

$$8. \begin{vmatrix} 2 & 0 & 1 & 0 \\ 3 & 1 & 3 & 1 \\ -1 & -1 & 1 & 2 \\ -1 & 2 & 2 & 1 \end{vmatrix}$$

11. Пусть A – квадратная матрица размера 4, и $\det(A)=3$. Найдите $\det(2A)$.

12. С матрицей A из предыдущей задачи произведены следующие преобразования: вторая строка матрицы заменена на разность между удвоенной второй и утроенной третьей строками. Каков определитель вновь полученной матрицы?

Решите системы при помощи формул Крамера:

$$13. \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$$

$$16. \begin{pmatrix} 3 & -1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 7 \\ 12 \end{pmatrix}$$

$$14. \begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -1 \\ 4 \end{pmatrix}$$

$$17. \begin{pmatrix} 4 & -1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 10 \\ -2 \end{pmatrix}$$

$$15. \begin{pmatrix} 1 & -2 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ -7 \end{pmatrix}$$

$$18. \begin{pmatrix} 2 & 2 \\ 2 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -2 \\ 4 \end{pmatrix}$$

По правилу Крамера найти x_2 :

$$19. \begin{pmatrix} 1 & -2 & 4 \\ 0 & 3 & 2 \\ 1 & -4 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 9 \\ 4 \\ 1 \end{pmatrix}$$

$$20. \begin{pmatrix} 1 & -2 & 0 \\ 0 & 3 & 2 \\ 1 & -4 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 5 \end{pmatrix}$$

По правилу Крамера найти x_3 :

$$21. \begin{pmatrix} 1 & -2 & 3 \\ 2 & 3 & 2 \\ 1 & -4 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \\ 1 \end{pmatrix}$$

$$22. \begin{pmatrix} 5 & -2 & 1 \\ 3 & -5 & 2 \\ 6 & -4 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 \\ 5 \\ 6 \end{pmatrix}$$

23. Пусть имеется уравнение $A\vec{x} = \vec{b}$, причем второй столбец матрицы A совпадает со столбцом свободных членов. Чему равно x_2 ?

24. В условиях предыдущей задачи найти x_1 .

25. Пусть имеется уравнение $A\vec{x} = \vec{b}$, причем второй столбец матрицы A совпадает с удвоенным столбцом свободных членов. Чему равно x_2 ?

Пусть $B = A^{-1}$. Вычислить указанный элемент обратной матрицы.

$$26. A = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}, \quad b_{12}$$

$$27. A = \begin{pmatrix} 8 & 4 \\ 5 & 3 \end{pmatrix}, \quad b_{21}$$

$$28. A = \begin{pmatrix} 1 & -3 & 2 \\ -5 & 1 & -2 \\ 0 & 2 & 1 \end{pmatrix}, \quad b_{31}$$

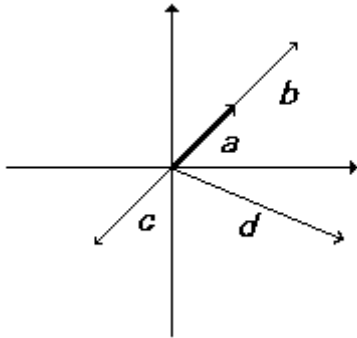
$$29. A = \begin{pmatrix} 1 & -1 & 2 \\ -2 & 1 & -2 \\ 0 & 2 & 1 \end{pmatrix}, \quad b_{32}$$

ОТВЕТЫ

1) -130, 2) -35, 3) 60, 4) 0, 5) -2, 6) -60, 7) 11, 8) -3, 9) -2, 10) 2, 11) 48, 12) 6, 13) (1, -1), 14) (1, 2), 15) (-2, -1), 16) (3, 2), 17) (2, -2), 18) (-3, 2), 19) 0, 20) -1, 21) 1, 22) 5, 23) 1, 24) 0, 25) 2, 26) 4/5, 27) -5/4, 28) 1/3, 29) 2/5.

Задача о собственных векторах и собственных значениях

Два ненулевых вектора x и y называются коллинеарными (лежащими на одной прямой), если существует такой коэффициент пропорциональности a , что выполнено соотношение $y = ax$, $a \in R$.



На рисунке изображены векторы в двумерном векторном пространстве. Векторы a , b и c являются коллинеарными. В то же время, вектор d не коллинеарен ни одному из них.

Как уже отмечалось, каждая квадратная матрица размера n задает линейное преобразование в n -мерном векторном пространстве. Преобразование, таким образом, может быть отождествлено с матрицей A . Под воздействием преобразования A каждый вектор x преобразуется в вектор $y = Ax$. Этот факт часто записывают следующим образом:

$$A: x \rightarrow y; \quad y = Ax$$

Определение. Собственным вектором данного линейного преобразования A называется всякий ненулевой вектор x , удовлетворяющий условию

$$(0.17) \quad Ax = \lambda x$$

где λ — какое-нибудь число.

Число λ называется *собственным значением* преобразования A , соответствующим данному собственному вектору x .

Для краткости говорят: « λ - собственное значение данного собственного вектора». Преобразование A каждый вектор векторного пространства переводит в какой-либо иной вектор того же пространства. *Собственный вектор переходит в коллинеарный ему вектор.* В действительном пространстве собственное значение показывает, во сколько раз собственный вектор «вытягивается» (при $|\lambda| < 1$ фактически «сжимается»). Отрицательность

собственного значения означает, что образ собственного вектора будет ориентирован в пространстве в противоположном направлении.

Нетрудно сообразить, что если x —собственный вектор, то ax —также собственный вектор при любом $a \neq 0$, то есть любой вектор коллинеарный собственному тоже является собственным. Множество векторов $y=ax$, где вектор x – собственный, а $a \in \mathbb{R}$, представляет собой инвариантное одномерное подпространство (инвариантную прямую).

Во многих вопросах алгебры и ее приложений встречается необходимость найти все собственные векторы данного линейного преобразования. Займемся этой задачей.

В силу того, что $Ex=x$, можно преобразовать уравнение (0.17) следующим образом:

$$\begin{aligned} Ax = \lambda x &\Rightarrow Ax = \lambda Ex \Rightarrow Ax - \lambda Ex = 0 \\ (0.18) \quad &(A - \lambda E)x = 0 \end{aligned}$$

Уравнение (0.18) представляет собой матричную запись однородной системы линейных уравнений.

Определение. Матрица $(A - \lambda E)$ системы (0.18) называется *характеристической матрицей* данного преобразования A , а ее определитель $\det(A - \lambda E)$ называется *характеристическим определителем* преобразования A .

Очевидно, что $\det(A - \lambda E)$ есть многочлен степени n относительно λ . Его называют характеристическим многочленом матрицы A (или преобразования A).

Однородная система всегда имеет нулевое решение $x=0$. Если определитель матрицы $(A - \lambda E)$ не равен нулю, то нулевое решение системы (0.18) будет ее единственным решением. Это нулевое решение называется *тривиальным*. Нетривиальные решения у системы будут только в том случае, когда определитель матрицы $(A - \lambda E)$ **равен нулю**. В этом случае система будет иметь бесконечное множество решений.

Общий план решения задачи о собственных векторах сводится теперь к следующему. Сначала составляется так называемое характеристическое уравнение

$$(0.19) \quad \det(A - \lambda E) = 0$$

Равенство (0.19) необходимо и достаточно для того, чтобы система (0.18) имела нетривиальные решения. В действительном пространстве собственными значениями являются все действительные корни (0.19) и только они. Допустим, что все корни $\lambda_1, \dots, \lambda_k$ найдены. Каждый из них подставим в систему (0.18). Каждый раз эта система получит определенные числовые коэффициенты.

Ранг полученной системы будет некоторым числом r , причем $r < n$, так что система будет иметь $n-r$ независимых решений. Найдя их, мы найдем тем самым $n-r$ независимых собственных векторов с одним и тем же собственным значением, равным взятому корню. Их линейная оболочка, с исключением из нее нулевого вектора, дает все собственные векторы с тем же собственным значением. Это следует из теоремы о множестве решений линейной однородной системы уравнений.

Перебрав таким способом все действительные корни характеристического уравнения, мы найдем вообще все собственные векторы данного преобразования.

Примеры. 1) Подобие представляет собой преобразование, для которого все ненулевые векторы являются собственными с одним и тем же собственным значением, равным коэффициенту подобия. Пусть в двумерном векторном пространстве преобразование подобия задано матрицей:

$$A_k = \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}, \text{ где } k - \text{коэффициент подобия.}$$

Характеристическое уравнение данного преобразования:

$$\begin{vmatrix} k - \lambda & 0 \\ 0 & k - \lambda \end{vmatrix} = (k - \lambda)^2 = 0$$

имеет корень $\lambda=k$, кратности 2.

Система (0.17) приобретает вид:

$$\begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = k \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \Rightarrow \begin{cases} kx_1 = kx_1 \\ kx_2 = kx_2 \end{cases}$$

и удовлетворяется при любых x_1 и x_2 .

2) Преобразование $B = \begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix}$ имеет характеристическое уравнение:

$$\begin{vmatrix} 2-\lambda & 3 \\ 1 & 4-\lambda \end{vmatrix} = (2-\lambda)(4-\lambda) - 3 = 0 \Rightarrow \lambda^2 - 6\lambda + 5 = 0,$$

дающее два собственных значения: $\lambda_1 = 1$, $\lambda_2 = 5$. Найдем собственные вектора.

$$\lambda_1 = 1 \Rightarrow \begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \Rightarrow \begin{cases} 2x_1 + 3x_2 = x_1 \\ x_1 + 4x_2 = x_2 \end{cases} \Rightarrow x_1 = -3x_2$$

$$\lambda_2 = 5 \Rightarrow \begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 5 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \Rightarrow \begin{cases} 2x_1 + 3x_2 = 5x_1 \\ x_1 + 4x_2 = 5x_2 \end{cases} \Rightarrow x_1 = x_2$$

Таким образом, собственными векторами преобразования являются:

$$v_1 = \alpha(3, -1), \quad v_2 = \beta(1, 1), \quad \alpha \in R, \quad \beta \in R, \quad \alpha \neq 0, \quad \beta \neq 0$$

3) Преобразование $C = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}$ не имеет собственных векторов, так как

характеристическое уравнение $\lambda^2 - 2\lambda + 3 = 0$ не имеет действительных корней.

ЗАДАЧИ И УПРАЖНЕНИЯ 6.

Ниже предложены матрицы линейных преобразований. Найдите собственные числа и собственные вектора этих преобразований. В двумерном случае изобразите собственные вектора на координатной плоскости.

1) $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

2) $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

3) $\begin{pmatrix} 1 & 3 \\ 2 & 0 \end{pmatrix}$

4) $\begin{pmatrix} 1 & 2 \\ 6 & 2 \end{pmatrix}$

$$5) \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \quad 6) \begin{pmatrix} 3 & 3 & 3 \\ 0 & 2 & 2 \\ 0 & 0 & 1 \end{pmatrix} \quad 7) \begin{pmatrix} 1 & -1 & 1 \\ 0 & 2 & -1 \\ 0 & 0 & 3 \end{pmatrix} \quad 8) \begin{pmatrix} 4 & 0 & 3 \\ 2 & -1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$$

Ответы:

- 1) $\lambda_{1,2} = \pm 1$, $(1,1)$, $(-1,1)$; 2) $\lambda_{1,2} = \pm 1$, $(1,0)$, $(0,1)$;
 3) $\lambda_1 = -2$, $\lambda_2 = 3$, $(1,-1)$, $(3,2)$; 4) $\lambda_1 = -2$, $\lambda_2 = 5$, $(2,-3)$, $(1,2)$;
 5) $\lambda = 1$, $(1,0,0)$; 6)
 $\lambda_1 = 1$, $\lambda_2 = 2$, $\lambda_3 = 3$, $(3,-4,2)$, $(-3,1,0)$, $(1,0,0)$;
 7) $\lambda_1 = 1$, $\lambda_2 = 2$, $\lambda_3 = 3$, $(1,0,0)$, $(1,-1,0)$, $(1,-1,1)$
 8) $\lambda_1 = 1$, $\lambda_2 = -1$, $\lambda_3 = 4$, $(1,0,-1)$, $(0,1,0)$, $(5,2,0)$

Equation Section (Next) Тема 3 Метрические пространства и некоторые способы статистической обработки данных

Векторное пространство U называется *нормированным*, если для каждого принадлежащего ему вектора a существует такое действительное число $|a|$ (**норма, абсолютная величина, модуль вектора a**), что из $a=b$ следует $|a|=|b|$ и для любых $a, b \in U$ выполнены следующие условия (аксиомы):

$$(1.1) \quad \begin{cases} |a| \geq 0 \\ |a| = 0 \Rightarrow a = 0 \\ |\alpha a| = |\alpha| |a| \quad \alpha \in R \\ |a+b| \leq |a| + |b| \quad (\text{Неравенство Минковского}) \end{cases}$$

В евклидовом пространстве в качестве нормы вектора принимается его длина, однако длина вектора является лишь одним из возможных, но не единственным вариантом нормы.

Пространство называется метрическим, если задан способ вычисления расстояний (метрика) $d(a, b)$ между любыми двумя векторами a и b . Каждое нормированное пространство является метрическим с метрикой $d(a, b) = |a - b|$.

Наличие метрики дает возможность говорить об удаленности векторов друг от друга. Близким векторам соответствуют малые расстояния, далеким – большие. Понятие расстояния можно использовать для классификации объектов, описываемых многомерными векторами.

Методы иерархической классификации

В качестве примера рассмотрим рисунок 1. На нем изображено множество точек на плоскости. Каждая точка описывается двумя координатами и, следовательно, является двумерным вектором.

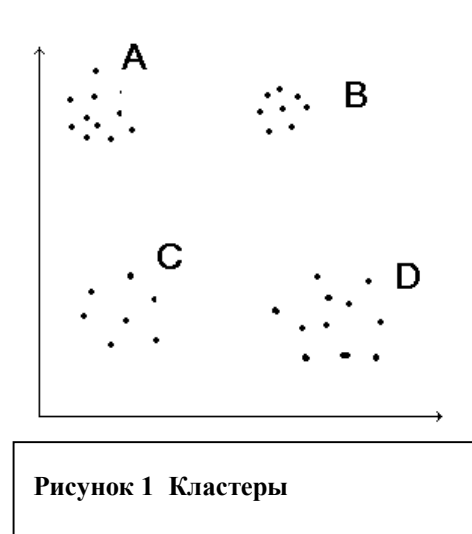


Рисунок 1 Кластеры

Из рисунка видно, что точки располагаются не беспорядочно, а имеют сгущения. Эти сгущения обозначены на рисунке буквами A, B, C, D. Расстояния между отдельными точками внутри сгущения относительно малы, в то время как расстояния между точками из разных сгущений существенно больше. Если координаты вектора суть численное выражение некоторых

свойств, описываемого им объекта, то можно утверждать, что объекты, входящие в состав одного сгущения обладают близкими свойствами. Сгущения точек, типа изображенных на рисунке 1, принято называть **кластерами**, а процесс их выявления из всего множества точек **кластерным анализом**. Кластерный анализ позволяет разбить все множество объектов на отдельные классы (кластеры), обладающие общностью свойств, то есть он решает задачу классификации объектов. Речь здесь идет не о распределении объектов по известным заранее классам, а о выявлении и формировании

самых классов. То есть кластерный анализ устанавливает классификацию, либо не существовавшую ранее, либо, если это желательно, игнорирующую предшествующие работы и пересматривающую данные заново.

Естественно, что кластерный анализ активно использует метрику пространства, то есть способ вычисления расстояний между векторами.

Сфера применения кластерного анализа очень широка. В качестве объектов кластеризации могут рассматриваться растения, животные, минералы, почвы, покупатели, страны, политические партии, политические деятели и многое другое.

Для лучшего понимания дальнейшего материала рассмотрим простой пример. Пусть ставится задача классификации депутатов Парламента по результатам их персональных голосований. Здесь полностью игнорируется партийная или фракционная принадлежность депутата, анализируется лишь его голосовательное поведение. Нелишне заметить, что декларируемая политическая позиция и позиция, занимаемая при голосовании могут сильно отличаться друг от друга.

Рассмотрим таблицу, в первом столбце которой все члены Парламента, а в первой строке выписаны законопроекты, поставленные на персональное голосование. Абстрагируемся от мелких проблем, связанных с возможной неявкой депутата или подаче голоса типа «воздержался». Будем считать, что голосуют всякий раз все депутаты, а голоса подаются лишь в двух возможных формах: «против» и «за». Закодируем голос «за» цифрой 1, а голос «против» - цифрой 0. Тогда получим таблицу типа той, что показана на иллюстрации.

	Законопроект 1	Законопроект 2	Законопроект 3	...	Законопроект n
Депутат 1	1	1	0	...	0
Депутат 2	0	1	0	...	1
...	...	...	...	...	...
Депутат m	0	0	1	...	1

Матрица, получаемая из таблицы исключением первой строки и первого столбца, называется матрицей исходных данных. Матрица состоит из векторов строк размерности n . Вектор-строка с номером i фиксирует голосовательное поведение депутата с номером i . Величина n обычно имеет порядок нескольких десятков, поэтому использовать визуальные методы выявления кластеров, как это показано на рисунке 1 для двумерных векторов, не представляется возможным. Кластерный анализ может производиться только с привлечением специальных математических методов, активно использующих понятие расстояния, и, безусловно, с использованием компьютера.

Введение и определения. Базовая модель.

Данные, предназначенные для кластерного анализа, как правило, состоят из множества элементов, обычно известных как индивиды или операционные таксономические единицы (ОТЕ), каждый из которых определяется набором признаков (в приведенном выше примере – вектором поданных голосов). Термин "признак" используется в широком смысле как обобщение понятия переменной. Множество индивидов предполагается неоднородным в том смысле, что его полезно рассматривать как состоящее из неизвестного числа подмножеств, которые надо найти. Члены одного подмножества должны быть более сходны друг с другом, чем с членами любого другого подмножества. Термин "сходны" включает ряд различных математических выражений, которые будут даны ниже. Этот подход обычно называют моделью минимальной дисперсии, хотя дисперсия в ее статистическом смысле может и не использоваться как мера сходства. Решение задачи обычно не единственно, и его лучше рассматривать не как структуру, являющуюся объективным свойством данных, а как структуру для исследователя в том смысле, что она предназначена для выяснения исследователем интересующих его свойств данных, слишком громоздких для непосредственного анализа.

Полезно различать два крайних типа интересов исследователя.

1. Исследователь хочет выяснить, в какой степени можно считать, что существуют различные подмножества при использовании данной численной модели. Система должна быть возможно меньше искажена, так чтобы если разделения на различные подмножества не существует, то оно и не должно быть найдено. Этот тип анализа относительно устойчив при изменении численной стратегии.

2. Исследователь подозревает, что истинных различных подмножеств не существует, однако для облегчения анализа очень большого числа элементов хочет иметь подмножества, полученные путем искусственного расчленения. Этот процесс, обычно называемый *рассечением*, очень чувствителен к изменению численной стратегии и поэтому не нравится математикам. Однако исследователи-прикладники применяют его довольно часто.

Иногда желательно транспонировать матрицу данных и классифицировать признаки на основе их значений для разных индивидов. Такая классификация называется *обратной* в противоположность *нормальной* классификации индивидов.

Типы признаков. Известны три главных типа. *Номинальные* признаки (иногда называемые "признаками с неупорядоченными состояниями") определяются рядом состояний, например: (песчанник, гранит, базальт, мел) или (врач, учитель, инженер, адвокат), так что, хотя состояния могут быть перенумерованы, номер состояния не несет никакой смысловой нагрузки. Обычно любой отдельный элемент может находиться только в одном из состояний данного признака. Частным случаем являются *бинарные* (или "качественные") признаки, представляющие собой номинальные признаки только с двумя состояниями (присутствие или отсутствие, да или нет). *Порядковые* признаки (иногда называемые "признаками с упорядоченными состояниями") определяются упорядоченным рядом состояний, например редкий, случайный, обычный, изобилующий. Порядок состояний имеет смысл, однако расстояния между состояниями не определены. *Численные*

признаки (иногда называемые "метрическими" или "количественными") представляют собой измеряемые или исчисляемые количества (вес в граммах, число листьев, уровень годового дохода и т.п.).

Наибольшие проблемы возникают при попытке кластеризации с

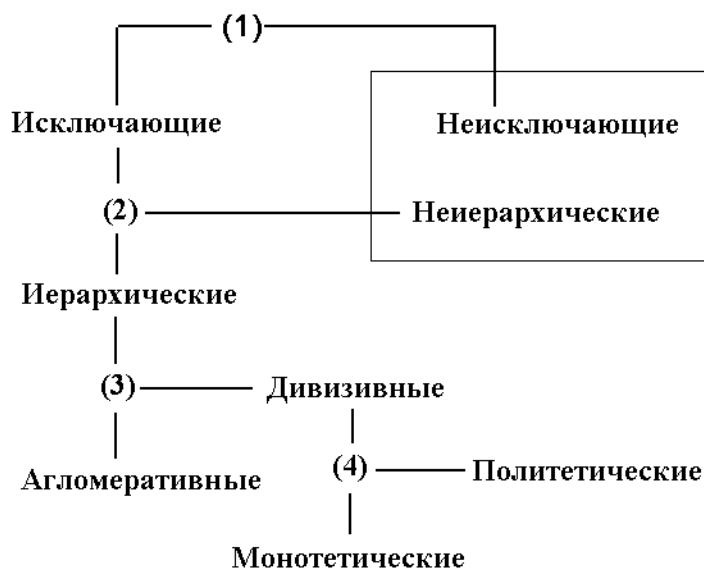


Рисунок 2 Типы классификации

использованием номинальных признаков. Их трудно разумным образом оцифровать (за исключением случая бинарных признаков), а без оцифровки невозможно использовать понятие расстояния. Иногда в подобных случаях используются иные меры сходства и различия, отличные от расстояния.

Однако в рамках нашего курса мы ограничимся тем случаем, когда применение понятия расстояния возможно.

Типы классификации. Численные классификации включают различные численные процессы, из которых выбирается наиболее подходящий. Последовательность выборов представлена на рис. 2, на котором они пронумерованы. Стратегии, заключенные в прямоугольник, и любые выборы внутри них не рассматриваются в данном пособии.

Решения принимаются в следующем порядке.

(1) *Исключающие/неисключающие.* В исключающей классификации один элемент может появиться в одном и только одном подмножестве. В неисключающем варианте один и тот же элемент может, появиться более чем в одном подмножестве. (В классификации пациентов больницы, основанной на болезнях, один и тот же пациент может иметь более одного заболевания.) Неисключающие классификации используются в некоторых видах

информационного поиска, особенно в библиотечной работе и медицинской диагностике. Они не носят иерархического характера и не будут далее рассматриваться в этой главе.

(2) *Иерархические/неиерархические*. В неиерархической классификации группы выбираются таким образом, чтобы каждая была возможно более однородной. Отношения между группами не выясняются. В иерархическом случае группы рассматриваются попарно как возможные кандидаты для объединения. Критерием объединения служит возможно меньшее увеличение неоднородности при объединении. Формально это обычно формулируется как утверждение, что неиерархическая классификация оптимизирует внутренние свойства подмножеств, а иерархическая — отношения между индивидами и всей популяцией. Неиерархическая классификация всегда итеративная и нами далее рассматриваться не будет. Почти все иерархические стратегии детерминированы в вычислительном отношении и, как следствие, менее требовательны к машинному времени по сравнению с неиерархическими стратегиями.

(3) *Агломеративные/дивизивные*. Различие между этими классификациями состоит в направлении. В агломеративной классификации индивиды объединяются в подмножества возрастающего объема до тех пор, пока вся популяция не объединится в одно множество. В дивизивной классификации исходная популяция элементов постепенно разделяется (обычно дихотомически) до тех пор, пока не будет получена желаемая степень разделения.

(4) *Монотетические/политетические*. В монотетической классификации деление производится на основании одного признака, имеющего максимальную информативность, тогда как в политетической классификации все признаки учитываются в равной степени. Агломеративные стратегии всегда политетические.

Каждый из перечисленных типов может быть разделен в свою очередь на различные стратегии, отличающиеся используемыми численными моделями, выбор которых зависит от обрабатываемых данных.

Дополнительные требования. В начале классификации почти во всех приложениях все признаки считаются равноправными. Однако после того, как процесс закончен, обнаруживается, что вклад в полученный результат различных признаков неодинаков. Отсюда вытекает потребность в диагностической системе, которая следила бы за ходом политетической классификации и выдавала бы упорядоченный список вкладов признаков для каждого очередного объединения или деления. Может возникнуть потребность в том, чтобы взаимные отношения между конечными подмножествами были отображены в евклидово пространство низкой размерности. Этот процесс называется ординацией.

Меры сходства и различия.

Мы рассмотрим меры сходства между двумя индивидами или, что менее обычно, между двумя признаками. Строго говоря, увеличение положительной меры соответствует увеличению сходства в случае меры сходства и уменьшению сходства в случае меры различия, однако это различие не всегда соблюдается, и часто в обоих случаях говорят о *мерах сходства*. На практике все современные системы имеют дело с мерами различия. Различают два главных класса. Первый включает (i,j) -меры, т.е. меры, которые могут быть вычислены для двух элементов или групп элементов до их объединения, если известны характеристики этих элементов или групп. Второй класс включает (ij,k) -меры, связывающие характеристики групп до и после объединения. В литературе можно найти огромное число предлагаемых мер, некоторые из которых довольно необычны, однако преобладают четыре главных класса. Это следующие меры: 1) коэффициент корреляции; 2) меры, основанные на евклидовой метрике; 3) меры, основанные на манхеттенской мере (называемой также мерой в городских

кварталах или мерой на решетке) и 4) информационные статистики. Мы рассмотрим каждый из этих классов. Элементы матрицы исходных данных будем обозначать x_{ij} , где i – номер индивида, а j – номер признака. В случае бинарных признаков будем использовать (a, b, c, d) -обозначения, где a — число признаков, которыми обладают оба индивида; d — число признаков, не встречающихся ни у одного из двух индивидов; b и c - числа признаков, которыми обладает только один из индивидов. Будем называть $(1-x)$ "дополнением x до единицы". Другие символы и термины будут определяться по мере надобности.

Коэффициент корреляции. Коэффициент корреляции используется в классической статистике как мера связи между двумя переменными. Пусть имеется всего n индивидов и пусть $\bar{x}_j$ и $\bar{x}_k$ — средние по всем индивидам значения для j -го и k -го признаков соответственно. Тогда коэффициент корреляции между признаками j и k вычисляют по формуле

$$r_{jk} = \frac{\sum_i (x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k)}{\sqrt{\sum_i (x_{ij} - \bar{x}_j)^2 \sum_i (x_{ik} - \bar{x}_k)^2}}$$

В бинарном случае, когда $x_{ij} \in \{0,1\}$, мы можем использовать обозначения a, b, c, d . Формула для коэффициента корреляции, известного в этом случае обычно как ϕ -коэффициент Пирсона, принимает тогда вид

$$(ad - bc) / \sqrt{(a+b)(a+c)(b+d)(c+d)}$$

Эта формула нашла некоторое применение в исследованиях по экологии растений. Коэффициент корреляции применяется как мера сходства между признаками. Один и тот же признак измеряется у всех индивидов в одних и тех же единицах. В этом случае понятие среднего значения признака имеет ясный содержательный смысл. Если же предпринимается попытка использовать коэффициент корреляции как меру сходства между индивидами, то здесь могут возникать сложности, связанные с

неоднородность признаков, их измерением в разных физических единицах. В этом случае понятие среднего значения может оказаться бессмысленным.

Евклидова метрика. Это метрика Минковского второго порядка, определяемая формулой (расстояние между индивидами 1 и 2):

$$\sqrt{\sum_j |x_{1j} - x_{2j}|^2}$$

Поскольку желательно, чтобы стратегии классификации были аддитивными по отношению к признакам, часто используется квадрат евклидова расстояния:

$$\sum_j |x_{1j} - x_{2j}|^2$$

который обычно делится на число признаков, чтобы избежать получения слишком больших чисел.

Унификация и стандартизация числовых признаков. Если признаки измерены в различных физических единицах, то предварительно они должны быть сделаны безразмерными, иначе операция сложения будет бессмысленной. Если они измерены в одинаковых единицах, но имеют очень разные диапазоны значений, то должны быть стандартизованы, чтобы признаки с большим разбросом не оказались полностью доминирующими. Обычно оба эти требования выполняются после деления исходных значений на некоторые величины той же размерности, в результате чего вклады отдельных признаков в некотором смысле выравниваются. Желательно, чтобы эти величины были достаточно устойчивы по отношению к небольшим изменениям в данных и по отношению к выборочным ошибкам.

Порядковые признаки. Если популяция достаточно велика, то в принципе можно было бы применить преобразование нормализации. Однако это довольно громоздкий прием, поэтому на практике обычно считают состояния равноудаленными и обрабатывают порядковые переменные как числовые.

Манхеттенская метрика. Эта метрика, являющаяся частным случаем метрики Минковского порядка 1, определяется формулой

$$\sum_j |x_{1j} - x_{2j}|$$

Приведенное выражение часто делится на число признаков.

Стандартизация числовых признаков осуществляется разными способами. К числу наиболее известных относятся *канберровская метрика* и *метрика Брея-Кертиса*.

Метрика Брея—Кертиса. Эта метрика задается формулой

$$\frac{\sum_j |x_{1j} - x_{2j}|}{\sum_j x_{1j} + \sum_j x_{2j}}$$

Хотя в этом случае значения и заключены между 0 и 1, все же признаки с большим размахом будут оказывать доминирующее влияние. Обычно перед использованием этой метрики данные стандартизуют (первоначально метрика была предназначена для случая, когда все x_{1j} были процентными долями). Данные после стандартизации должны быть неотрицательными. Процедура используется не часто, однако имеет своих приверженцев.

Канберровская метрика. Эта метрика задается выражением

$$(1.2) \quad \sum_j \frac{|x_{1j} - x_{2j}|}{x_{1j} + x_{2j}}$$

Она автоматически стандартизована и в случае неотрицательных x , заключена между 0 и 1. Ее нечувствительность к сильно отклоняющимся значениям делает ее подходящей в случае сильно асимметричных данных. Она имеет особенность в нуле в том смысле, что при $x_{1j} = 0$ ее величина равна 1 независимо от значения x_{2j} . Обычно эту трудность обходят, заменяя нулевые x_{1j} значениями, меньшими, чем наименьшее из всех x_{ij} в матрице данных. В случае использования данных разных знаков знаменатель в (1.2) можно заменить на $|x_{1j}| + |x_{2j}|$, однако тогда мера

принимает максимальное значение, равное 1 во всех случаях, когда x_{1j} и x_{2j} имеют противоположные знаки.

Порядковые переменные. Обращаются как числовые.

Номинальные переменные. Представляет интерес лишь ситуация, когда результат сравнения по номинальному признаку равен 1, если состояния пересекаются, и 0 – в остальных случаях. В этом случае наиболее подходящей будет мера Брея — Кертиса. В полностью бинарном случае мера Брея-Кертиса принимает вид $(b+c)/(2a+b+c)$. Канберровская мера аналогичным образом принимает вид $(b+c)/(a+b+c)$.

Информационная статистика. Этот случай требует специального рассмотрения, поскольку переменные, значения которых получены путем счета или путем измерения на непрерывной шкале, не могут обрабатываться таким же образом. Удобно начать с полностью бинарного случая. Может быть использовано определение информации Шеннона, основанное на формуле $n \ln n$. или Бриллюэна, основанное на формуле $\ln n!$. В большинстве случаев возникающие различия несущественны. Чаще используется определение Шеннона.

Бинарные признаки. Пусть имеется группа из n элементов, описываемых присутствием или отсутствием s признаков, и пусть a_j элементов обладают признаком j . Определим информационное содержание такой группы выражением

$$I = s \cdot n \cdot \ln n - \sum_j \left[a_j \ln a_j + (n - a_j) \ln (n - a_j) \right]$$

Пусть информационное содержание двух групп A и B равно соответственно I_A и I_B , и пусть обе группы объединены в одну группу C с информационным содержанием I_C . Тогда можно определить информационный выигрыш от слияния двух групп формулой

$$\Delta I = I_C - (I_A + I_B)$$

Информационное содержание одного элемента или группы одинаковых элементов в этой модели всегда равно нулю.

Пусть отдельный элемент, подлежащий классификации, сам представляет популяцию из n индивидов, разделяющихся на категории (например, на виды растений или животных), так что в j -ю категорию попадает a_j индивидов и $\sum_j a_j = n$. Определим информационное содержание полного элемента (называемое также разнообразием) как

$$\left(\sum_j a_j \right) \ln \left(\sum_j a_j \right) - \sum_j (a_j \ln a_j)$$

Информационный выигрыш определяется, как и прежде. Запишем его в явном виде. Пусть два объединяющихся элемента представлены векторами-строками (a_{1j}) и (a_{2j}) , r_i - суммы по строкам ($i = 1, 2$), c_j - сумма по столбцу и N - общая сумма. Тогда

$$\Delta I = N \ln N - \sum_j (c_j \ln c_j) - \sum_i (r_i \ln r_i) + \sum_i \sum_j (a_{ij} \ln a_{ij})$$

Эта величина называется также *переданной информацией*.

Стратегии объединения (агломеративные системы).

Начальные действия во всех агломеративных системах одинаковы. Для n индивидов вычисляются все $n(n - 1)/2$ мер различия и пара индивидов с наименьшей мерой объединяется в одну группу. Необходимо затем определить подходящую меру различия между этой группой и остальными $n - 2$ индивидами, а на более поздних стадиях, очевидно, будет необходимо определить меру различия между индивидом и группой любого объема, а также между любыми двумя группами. На каждом шаге классификации осуществляется то объединение (между двумя индивидами, между индивидом и группой или между двумя группами), для которого мера различия минимальна среди всех оставшихся к данному шагу. Мера должна быть такой, чтобы индивид мог рассматриваться как группа из одного элемента. Стратегия объединения определяется именно мерой различия между группами.

Искажение пространства. Мы рассматриваем индивиды как точки в многомерном пространстве, оси которого соответствуют признакам. Некоторые стратегии объединения устанавливают границы между группами точек, но не изменяют относительного расположения точек в пространстве. Такие стратегии называют стратегиями, *сохраняющими метрику пространства*. В других стратегиях пространство вокруг группы растягивается по мере того, как группа растет, так что кажется, что группа отступает от остальных точек. Такие стратегии называются *растягивающими пространство*. Они интенсифицируют классификацию, и получающиеся группы кажутся более различающимися, чем это имеет место в действительности. Есть также стратегии, в которых пространство как бы сжимается вокруг группы по мере ее роста (*сжимающие пространство стратегии*). Естественная кластеризация в этом случае "размывается", и может появиться множество "цепочек", состоящих из последовательно соединившихся друг с другом индивидов.

Комбинаторные решения. Если при заданной начальной матрице различий все остальные меры между индивидом и группой или между двумя группами вычисляются, исходя из нее, рекурсивно, то процедура называется рекурсивной. (i,j) -меры обычно могут рассматриваться с позиций одной линейной модели. Пусть имеются две группы i и j с n_i и n_j элементами соответственно; мера различия между этими группами обозначается d_{ij} . Допустим, что d_{ij} — минимальная мера из всех оставшихся, так что i и j объединяются и образуют новую группу k с $n_k = n_i + n_j$ элементами. При объединении должны быть вычислены меры различия вновь образованной группы со всеми остальными группами. Рассмотрим некоторую другую группу h с n_h элементами. Перед объединением известны значения $d_{hi}, d_{hj}, d_{ij}, n_h, n_i, n_j$. Положим

$$d_{hk} = \alpha_i d_{hi} + \alpha_j d_{hj} + \beta d_{ij} + \gamma |d_{hi} - d_{hj}|$$

где параметры $\alpha_i, \alpha_j, \beta, \gamma$ определяют сущность стратегии. Для некоторых стратегий параметры представляют собой просто числа, однако в большинстве случаев они являются простыми алгебраическими функциями от некоторых величин n_h, n_i, n_j, n_k . Конкретные значения или выражения приведены ниже при описании отдельных стратегий.

Для информационной статистики, которая представляет истинные (ij, k) -меры, не существует комбинаторного решения, и в этом случае данные должны быть сохранены для вычислений, связанных с объединением в группы, в течение всего процесса классификации.

Монотонность. Чтобы графически представить весь процесс объединения, индивиды (или группы, если представление усеченное) должны быть размещены в соответствующем порядке вдоль оси абсцисс. Но последовательность объединений (*иерархия* или *дендрограмма*) требует, чтобы каждое объединение было связано с некоторым значением ординаты. Обычно для этого используют меру различия, соответствующую данному объединению, хотя в случае стратегий, растягивающих пространство, это может привести к ошибкам в интерпретации. Дендрограмма будет очень ненаглядной, если она не будет расти при каждом объединении, что соответствует требованию, чтобы ряд мер различия, связанных с последовательными объединениями, был монотонным. Если $\gamma \neq 0$, то все известные комбинаторные стратегии монотонны; остальные стратегии монотонны, если $\alpha_i + \alpha_j + \beta \geq 1$. Хотя стратегия, основанная на информационной статистике, не обязательно монотонна, но отсутствие монотонности в этом случае настолько редко, что им можно пренебречь.

Недостатки. В большинстве приложений только верхние уровни, т.е. последние несколько объединений, представляют интерес. Однако, чтобы их получить, необходимо произвести все объединения, начиная с отдельных индивидов. Как следствие этого агломеративные программы относительно медленны, и значительная часть машинного времени расходуется на

получение результатов, не представляющих существенного интереса. Кроме того, объединения начинаются на низшем уровне информации, где возможны наибольшие ошибки. Так как иерархические объединения не пересматриваются, некоторые ранние объединения могут в последующем оказаться бесполезными. Возникновение по этой причине отдельных "ошибок классификации" не является редкостью, особенно в случае пропущенных значений.

Конкретные стратегии объединения. Мере, основанной на информационной статистике, соответствуют свои собственные стратегии объединения, не требующие отдельного описания. Они сильно растягивают пространство и, хотя и очень популярны, должны интерпретироваться с осторожностью.

Стратегия ближайшего соседа. Расстояние между двумя группами определяется в этом случае как расстояние между двумя самыми близкими индивидами из этих групп. Параметры равны: $\alpha_i = \alpha_j = 1/2$, $\beta = 0$, $\gamma = -1/2$. Стратегия монотонна, сильно сжимает пространство и привлекательна в теоретическом отношении. Ее целесообразно использовать, если хотят получить что-то вроде минимального дерева вместо разбиения на группы. Она также оказывается полезной в случаях, когда понятие смежности является важным в рассматриваемой задаче. Например, когда два стада овец становятся одним, если ближайшие соседи из этих стад оказываются очень близко друг к другу. Однако сжатие пространства стратегией делает ее не очень подходящей в стандартных задачах классификации.

Стратегия дальнего соседа. В этой стратегии расстояние между двумя группами определяется как расстояние между двумя самыми удаленными представителями этих групп. Это монотонная, сильно растягивающая пространство стратегия, которая в значительной степени может быть

заменена гибкой стратегией (см. ниже). Ее параметры имеют следующие значения: $\alpha_i = \alpha_j = 1/2$, $\beta = 0$, $\gamma = 1/2$

Стратегия группового среднего. Если одна группа состоит из m_1 , а другая из m_2 индивидов, то расстояние между этими группами в данной стратегии определяется как среднее арифметическое из $m_1 \cdot m_2$ расстояний между индивидами из этих групп. Параметры стратегии равны:

$\alpha_i = \frac{n_i}{n_k}$, $\alpha_j = \frac{n_j}{n_k}$, $\beta = \gamma = 0$. Стратегия монотонна и сохраняет метрику пространства. Она рекомендуется, когда искусственное разделение групп не нужно.

Центроидная стратегия. В евклидовой модели расстояние между двумя группами определяется как расстояние между их центрами. Параметры стратегии равны: $\alpha_i = \frac{n_i}{n_k}$, $\alpha_j = \frac{n_j}{n_k}$, $\beta = -\alpha_i \alpha_j$, $\gamma = 0$. Стратегия сохраняет метрику пространства, но не монотонна. Нарушения монотонности возникают часто и доставляют неудобства, поэтому стратегия почти не используется.

Стратегия, основанная на приращении суммы квадратов. В евклидовой модели межгрупповое расстояние определяется как увеличение внутригрупповой суммы квадратов (суммы квадратов расстояний от соответствующих центров) при объединении. Стратегия монотонна, и растягивает пространство.

Гибкая стратегия. Применима для любой меры различия и определяется четырьмя ограничениями $\alpha_i + \alpha_j + \beta = 1$, $\alpha_i = \alpha_j$, $\beta < 1$, $\gamma = 0$.

Стратегия монотонна, и ее свойства полностью зависят от β . Если $\beta=0$, то стратегия сохраняет метрику пространства. Если β принимает положительные значения, то стратегия сжимает пространство, а если отрицательные, то растягивает. На практике обычно используют значение $\beta = -0,25$.

Разделяющие (дивизивные) стратегии.

Теоретические преимущества дивизивных стратегий перед агломеративными состоят в следующем: 1) процесс начинается с использования максимального информационного содержания; 2) деление не обязательно должно продолжаться до тех пор, пока вся популяция будет разделена до отдельных индивидов; 3) если число признаков меньше, чем число индивидов, то количество вычислений уменьшается, так как время, требуемое для реализации дивизивного процесса, приблизительно пропорционально квадрату числа признаков s^2 , тогда как для агломеративного - квадрату числа индивидов n^2 . Однако до недавнего времени все дивизивные программы были монотетическими, т.е. дихотомизация популяции проводилась на каждом шаге на основании только одного признака. Такая процедура чувствительна к ошибкам определения или кодирования признака, по которому проводится деление. Все индивиды с отклоняющимися значениями этого признака порождают "ложные" ветви иерархии.

Приведем пример кластеризации на основе объединяющей стратегии.

Пусть имеется 5 законодателей, обозначенных буквами от V до Z и результаты их персональных голосований по 8 законопроектам, обозначенным буквами от a до h. Эти результаты в оцифрованном виде, приведены в следующей таблице:

	a	b	c	d	e	f	g	h
V	1	0	1	1	1	0	0	1
W	1	0	1	1	0	1	0	1
X	0	1	1	0	1	0	1	0
Y	0	1	1	1	1	0	1	1
Z	0	1	1	1	1	0	0	0

Будем использовать Манхеттенскую метрику, то есть будем считать, что мера различия между законодателями есть число несовпадений в их голосовательных векторах. Подсчитаем все меры различия. Для того, чтобы подсчитать меру различия между V и W, сравним соответствующие вектора-строки: число несовпадений – 2. Продолжая процесс, получим следующую матрицу расстояний:

	V	W	X	Y	Z
V	0				
W	2	0			
X	5	7	0		
Y	3	5	2	0	
Z	3	5	2	2	0

Будем использовать для объединения законодателей стратегию группового среднего. Минимальное из расстояний равно 2. Начнем объединение с образования группы VW, состоящей из V и W. Необходимо рассчитать расстояния между новой группой и остальными законодателями. Рассчитаем расстояние $d(VW, X)$ между VW и X. Согласно стратегии группового среднего оно равно среднему арифметическому из расстояний $d(V, X)$ и $d(W, X)$:

$$d(VW, X) = \frac{5+7}{2} = 6$$

Рассчитав аналогичным образом остальные расстояния, получаем новую матрицу расстояний:

	VW	XY	Y	Z
VW	0			
X	6	0		
Y	4	2	0	
Z	4	2	2	0

Минимальное из расстояний по-прежнему равно 2. Объединим пару XY, соответствующую этому расстоянию. После перерасчета расстояний получаем:

	VW	XY	Z
VW	0		
XY	5	0	
Z	4	2	0

Объединяем XY и Z.

	VW	XYZ
VW	0	
XYZ	4,667	0

Полученные группы можно слить в одну. Однако это объединение происходит на качественно ином уровне. Если до сих пор объединение происходило при значении расстояния 2, то есть практически при одинаковой степени близости, то сейчас расстояние между кластерами составляет уже 4,667, что составляет резкий скачок. Разумно, поэтому, остановиться на данном этапе кластеризации. Окончательной структурой кластеров считаем структуру VW-XYZ. Процесс объединения можно проиллюстрировать дендрограммой:

