

Глава III. Матрицы

1. ДЕЙСТВИЯ НАД МАТРИЦАМИ

Понятие матрицы. Матрицей называется любая прямоугольная таблица вида

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

компоненты которой a_{ij} могут быть элементами любого множества. Помимо обозначения a_{ij} для элемента i -й строки и j -го столбца матрицы A мы будем использовать обозначение A_{ij} . Особенно это будет удобно при рассмотрении матриц, полученных из других матриц при помощи различных операций; например, $(A(B+C))_{ij}$ означает элемент i -й строки и j -го столбца матрицы $A(B+C)$.

Множество всех матриц с компонентами из множества \mathcal{A} , состоящих из m строк и n столбцов, обозначается через $\mathcal{A}^{m \times n}$. Часто бывает удобно рассматривать и матрицы с пустыми множествами строк или столбцов; это может понадобиться, например, когда мы выбрасываем из матрицы несколько строк или столбцов, и если не допускать матрицы с 0 строк или 0 столбцов, в рассуждениях придется вводить ограничения на рассматриваемые матрицы и разбирать различные случаи, что, как правило, ничем не оправдано. Конечно, у таких матриц нет никаких компонент, и ясно, множества $\mathcal{A}^{0 \times n}$, $\mathcal{A}^{m \times 0}$ должны состоять каждое из единственного элемента.

Действия над матрицами. В этом пункте мы рассматриваем матрицы с компонентами из фиксированного ассоциативного коммутативного кольца с единицей Λ . Определим три действия над матрицами: сложение и умножение матриц и умножение матрицы на элемент из Λ .

Пусть матрицы A, B принадлежат $\Lambda^{m \times n}$. Их суммой называется матрица $A+B \in \Lambda^{m \times n}$, компоненты которой равны суммам соответствующих компонент матриц A и B :

$$(A+B)_{ij} = A_{ij} + B_{ij} \quad (1 \leq i \leq m, \ 1 \leq j \leq n).$$

Аналогично, произведением элемента $a \in \Lambda$ на матрицу A называется матрица $aA \in \Lambda^{m \times n}$, элементы которой равны произведениям a на соответствующие элементы матрицы A :

$$(aA)_{ij} = aA_{ij} \quad (1 \leq i \leq m, \ 1 \leq j \leq n).$$

На первый взгляд, произведение матриц определяется менее естественно, но позднее мы поймем, почему оно определяется именно так. Пусть $A \in \Lambda^{m \times n}$, $B \in \Lambda^{n \times k}$; произведением матриц A, B называется матрица $AB \in \Lambda^{m \times k}$, компоненты которой вычисляются по формуле

$$(AB)_{ij} = \sum_{s=1}^n A_{is} B_{sj}, \quad (1 \leq i \leq m, \ 1 \leq j \leq k).$$

Таким образом, перемножать матрицы можно только тогда, когда число столбцов первого сомножителя равно числу строк второго сомножителя, и у получившейся матрицы будет столько строк, сколько у первого сомножителя, и столько столбцов, сколько у второго сомножителя. Компоненты произведения вычисляются по правилу "строка на столбец": для того, чтобы получить элемент i -й строки и j -го столбца произведения, надо перемножить первый элемент i -й строки первого сомножителя и первый элемент j -го столбца второго сомножителя, второй элемент i -й строки первого сомножителя и второй элемент j -го столбца второго

сумматора, ..., n -й элемент i -й строки первого сомножителя и n -ый элемент j -го столбца второго сомножителя, и сложить все полученные произведения.

Для каждого натуральных m, n через $\mathbf{0}^{m \times n}$ обозначим матрицу из m строк и n столбцов, все компоненты которой равны 0. Такая матрица называется нулевой. Из контекста обычно ясно, сколько строк и сколько столбцов в матрице, поэтому, как правило, мы будем писать $\mathbf{0}$ вместо $\mathbf{0}^{m \times n}$, хотя надо всегда помнить, что для каждого m, n своя нулевая матрица. Более того, в дальнейшем мы будем обозначать нулевую матрицу через 0, а не $\mathbf{0}$, но пока все же будем использовать различные символы для нуля кольца Λ и нулевой матрицы.

Единичной матрицей порядка n называется квадратная матрица $E_n \in \Lambda^{n \times n}$, все диагональные элементы которой равны 1, а недиагональные - 0:

$$(E_n)_{ii} = 1, \quad (E_n)_{ij} = 0 \quad \text{при } i \neq j \quad (1 \leq i, j \leq n).$$

Как и для нулевой матрицы, мы обычно опускаем указание на порядок единичной матрицы и пишем E вместо E_n .

Введем еще одно обозначение. Для матрицы $A \in \Lambda^{m \times n}$ через $-A$ будем обозначать матрицу из $\Lambda^{m \times n}$, которая получается заменой всех элементов матрицы на противоположные:

$$(-A)_{ij} = -A_{ij} \quad (1 \leq i \leq m, \quad 1 \leq j \leq n).$$

Свойства действий над матрицами. Всюду ниже в этом пункте A, B, C – матрицы с компонентами из коммутативного ассоциативного кольца с единицей Λ , $a, b \in \Lambda$. Напомним, что не любые две матрицы можно сложить или перемножить: складывать можно только матрицы с одинаковым числом строк и одинаковым числом столбцов, а перемножать можно лишь такие матрицы, что в первой из них столько же столбцов, сколько строк во второй; поэтому все перечисленные ниже свойства надо читать так: если определена одна из частей равенства, то определена и другая, и они равны (кроме свойств 3, 12, в которых надо требовать лишь, чтобы была определена левая часть, и свойств 4, 6, 7, 8, которые всегда осмысленны).

- (1) $(A + B) + C = A + (B + C)$;
- (2) $A + B = B + A$;
- (3) $A + \mathbf{0} = A$;
- (4) $A + (-A) = \mathbf{0}$;
- (5) $a(A + B) = aA + bB$;
- (6) $(a + b)A = aA + bA$;
- (7) $1 \cdot A = A$;
- (8) $(ab)A = a(bA)$;
- (9) $(AB)C = A(BC)$;
- (10) $A(B + C) = AB + AC$;
- (11) $(A + B)C = AC + BC$;
- (12) $EA = A, AE = A$;
- (13) $a(AB) = (aA)B = A(aB)$.

Доказательства большинства из этих свойств тривиальны; свойства (1)-(8) вообще сводятся к тому, что выполнены аналогичные свойства для элементов кольца Λ . Все же для примера докажем (1). Пусть $A \in \Lambda^{m \times n}$, и пусть имеют смысл все действия в левой части равенства $(A + B) + C = A + (B + C)$; тогда $B \in \Lambda^{m \times n}$ (иначе сумма $A + B$ не определена), и потому $A + B \in \Lambda^{m \times n}$. Для того, чтобы имела смысл сумма $(A + B) + C$ необходимо, чтобы и матрица C принадлежала множеству $\Lambda^{m \times n}$. Поэтому определена и сумма $A + (B + C)$, и обе матрицы $(A + B) + C, A + (B + C)$ состоят из одинакового числа строк и одинакового числа столбцов. Осталось показать, что соответствующие элементы обеих матриц совпадают; но это немедленно следует из ассоциативности сложения в кольце Λ :

$$\begin{aligned}
((A + B) + C)_{ij} &= (A + B)_{ij} + C_{ij} = (A_{ij} + B_{ij}) + C_{ij} = \\
&= A_{ij} + (B_{ij} + C_{ij}) = A_{ij} + (B + C)_{ij} = (A + (B + C))_{ij}.
\end{aligned}$$

Остальные из свойств (1)-(8) доказываются совершенно аналогично, и даже несколько проще. Самое громоздкое из свойств – это ассоциативность умножения матриц (9); проведем его доказательство полностью. Пусть определена, например, правая часть равенства $(AB)C = A(BC)$, и пусть $B \in \Lambda^{n \times k}$. Поскольку определено произведение BC , матрица C состоит из k строк; пусть r число ее столбцов, так что $BC \in \Lambda^{n \times r}$. Из того, что произведение $A(BC)$ имеет смысл, следует, что матрица A имеет n столбцов; пусть m – число ее строк. Теперь видно, что определено и произведение $(AB)C$, и что обе матрицы $(AB)C$, $A(BC)$ состоят из m строк и r столбцов. Проверим, что соответствующие элементы этих матриц равны. В приведенной ниже выкладке через I обозначено декартово произведение множеств $\{1, \dots, n\}$ и $\{1, \dots, k\}$. Первые два шага – выражение элемента произведения матриц AB , C через элементы этих матриц и выражение элементов матрицы AB через элементы матриц A , B ; затем мы пользуемся дистрибутивностью умножения относительно сложения в кольце Λ , заменяем двойную сумму на сумму по декартову произведению множеств индексов этих сумм и пользуемся ассоциативностью умножения в кольце Λ , после чего начинается "обратный ход" – мы производим те же операции в обратном порядке:

$$\begin{aligned}
((AB)C)_{ij} &= \sum_{t=1}^k (AB)_{it} C_{tj} = \sum_{t=1}^k \left(\sum_{s=1}^n A_{is} B_{st} \right) C_{tj} = \sum_{t=1}^k \left(\sum_{s=1}^n (A_{is} B_{st}) C_{tj} \right) = \\
&= \sum_{(s,t) \in I} (A_{is} B_{st}) C_{tj} = \sum_{(s,t) \in I} A_{is} (B_{st} C_{tj}) = \sum_{s=1}^n \sum_{t=1}^k A_{is} (B_{st} C_{tj}) = \\
&= \sum_{s=1}^n A_{is} \left(\sum_{t=1}^k A_{is} B_{st} C_{tj} \right) = \sum_{s=1}^n A_{is} (BC)_{sj} = (A(BC))_{ij}.
\end{aligned}$$

В свойствах (10), (13) мы опустим исследование структуры матриц, ограничившись лишь доказательством того, что соответствующие элементы левой и правой частей совпадают; в первом случае используется лишь дистрибутивность умножения относительно сложения, а в последнем – еще и то, что в коммутативном ассоциативном кольце произведения $a(A_{is}B_{sj})$, $(aA_{is})B_{sj}$, $A_{is}(aB_{sj})$ равны:

$$\begin{aligned}
(A(B + C))_{ij} &= \sum_{s=1}^n A_{is} (B + C)_{sj} = \sum_{s=1}^n (A_{is} B_{sj} + A_{is} C_{sj}) = \sum_{s=1}^n A_{is} B_{sj} + \sum_{s=1}^n A_{is} C_{sj} = \\
&= (AB)_{ij} + (AC)_{ij} = (AB + AC)_{ij}, \\
(a(AB))_{ij} &= a(AB)_{ij} = a \sum_{s=1}^n A_{is} B_{sj} = \sum_{s=1}^n a(A_{is} B_{sj}) = \sum_{s=1}^n (aA_{is}) B_{sj} = ((aA)B)_{ij}, \\
(a(AB))_{ij} &= a(AB)_{ij} = a \sum_{s=1}^n A_{is} B_{sj} = \sum_{s=1}^n a(A_{is} B_{sj}) = \sum_{s=1}^n A_{is} (aB_{sj}) = (A(aB))_{ij}
\end{aligned}$$

(в этих равенствах через n обозначается число столбцов матрицы A). Свойство (11) доказывается так же, как (10), а о свойстве (12) поговорим чуть подробнее.

Пусть $A \in \Lambda^{m \times n}$; напомним, что обозначение E применяется не для единственной матрицы, а для единичной матрицы E_n любого порядка n . Произведение EA определено, когда $E = E_m$, а произведение AE – когда $E = E_n$; таким образом, надо доказать, что $E_m A = A E_n = A$. Все матрицы в этом равенстве имеют m

строк и n столбцов, так что нужно проверить лишь совпадение их соответствующих элементов:

$$(E_m A)_{ij} = \sum_{s=1}^m (E_m)_{is} A_{sj} = (E_m)_{ii} A_{ij} = A_{ij} = A_{ij} (E_n)_{jj} = \sum_{t=1}^n A_{it} (E_n)_{tj} = (A E_n)_{ij},$$

потому что диагональные элементы $(E_m)_{ii}$, $(E_n)_{jj}$ матриц E_m , E_n равны 1, а все остальные элементы этих матриц равны 0.

Отметим, что умножение матриц, вообще говоря, не коммутативно. Во-первых, произведение AB может быть определено, а произведение BA – нет (например, если $A \in \Lambda^{2 \times 3}$, $B \in \Lambda^{3 \times 4}$); во-вторых, если даже оба произведения определены, они могут иметь разную структуру (например, при $A \in \Lambda^{2 \times 3}$, $B \in \Lambda^{3 \times 2}$ произведение AB является квадратной матрицей порядка 2, а произведение BA – квадратной матрицей порядка 3). Но даже когда со структурой все в порядке, произведения AB , BA могут быть различны, в чем убеждаешься на примере произведений почти любых первых попавшихся квадратных матриц порядка 2:

$$\begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 7 & 3 \\ 3 & 1 \end{pmatrix}, \quad \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 4 & 6 \\ 3 & 4 \end{pmatrix}.$$

Отметим еще, что матрицы могут быть "делителями 0"; например,

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

Кольцо квадратных матриц порядка n . Если количество строк и количество столбцов матрицы равно одному и тому же числу n , то такая матрица называется квадратной матрицей порядка n . Множество $\Lambda^{n \times n}$ всех квадратных матриц порядка n обозначается через Λ_n . Из определений ясно, что сумма и произведение матриц из Λ_n снова принадлежат Λ_n , а справедливость свойств (1)-(4), (9)-(12) действий над матрицами показывает, что Λ_n – ассоциативное кольцо с единицей $E = E_n$. Мы только что убедились, что уже при $n = 2$ кольцо Λ_n некоммутативно.

До сих пор все встречавшиеся у нас кольца были коммутативны. Теперь у нас появилось много примеров некоммутативных колец. Бывают и неассоциативные кольца; примером является множество векторов трехмерного пространства относительно обычного сложения и векторного умножения.

Транспонирование матриц. Пусть A – матрица из $\mathcal{A}^{m \times n}$; транспонированной по отношению к A матрицей называется матрица $A^T \in \mathcal{A}^{n \times m}$, элементы которой симметричны элементам матрицы A относительно диагонали:

$$(A^T)_{ij} = A_{ji}, \quad (1 \leq j \leq m, \quad 1 \leq i \leq n).$$

Таким образом, при транспонировании строки матрицы становятся ее столбцами, а столбцы – строками.

Далее в этом пункте A и B будут матрицами с компонентами из коммутативного ассоциативного кольца с единицей Λ , а a – элементом из Λ . Следующие свойства показывают, как связано транспонирование с действиями над матрицами. Как и выше, свойства (15) и (16) надо читать так: если определена одна из сторон равенства, то определена и другая, и они равны.

$$(14) \quad (aA)^T = aA^T;$$

$$(15) \quad (A + B)^T = A^T + B^T;$$

$$(16) \quad (AB)^T = B^T A^T;$$

$$(17) \quad (A^T)^T = A.$$

Приведем лишь доказательство свойства (16) (остальные тривиальны). Пусть $A \in \Lambda^{m \times n}$; если имеет смысл левая часть равенства, то определено произведение AB , а это значит, что у матрицы B число строк равно n ; пусть k – число ее

столбцов. Тогда $AB \in \Lambda^{m \times k}$, а $(AB)^T \in \Lambda^{k \times m}$. Далее, в этой ситуации $B^T \in \Lambda^{k \times n}$, $A^T \in \Lambda^{n \times m}$, поэтому произведение $B^T A^T$ определено и принадлежит $\Lambda^{k \times m}$. Аналогичное рассмотрение показывает, что если определена правая часть равенства (16), то определена и его левая часть, и обе этих матрицы имеют одинаковое число строк и одинаковое число столбцов. Остается убедиться, что равны соответствующие компоненты левой и правой сторон равенства (16). Пользуясь коммутативностью умножения, получаем:

$$((AB)^T)_{ij} = (AB)_{ji} = \sum_{s=1}^n A_{js} B_{si} = \sum_{s=1}^n B_{si} A_{js} = \sum_{s=1}^n (B^T)_{is} (A^T)_{sj} = (B^T A^T)_{ij}.$$

Отметим, что это – единственное из всех свойств действий над матрицами (кроме части свойства (13)), в котором существенна коммутативность кольца Λ ; все остальные свойства без труда переносятся на случай матриц над некоммутативным, но ассоциативным кольцом. Это показывает, что транспонирование матриц естественно только для матриц над коммутативными кольцами; в случае некоммутативных колец иногда удается подправить определение транспонированной матрицы так, чтобы свойство (16) сохранилось, но для этого надо, чтобы на Λ была некоторая дополнительная структура.

Нумерация строк и столбцов элементами множеств. Обычно строки и столбцы матрицы нумеруются первыми натуральными числами; в этом случае они естественным образом упорядочены. Однако, иногда бывает удобно нумеровать их элементами из некоторых множеств S и T , не обязательно упорядоченных. В этом случае матрица с компонентами из множества \mathcal{A} понимается как отображение $S \times T \rightarrow \mathcal{A}$. Образ A_{st} элемента $(s, t) \in S \times T$ при этом отображении считается элементом s -й строки и t -го столбца матрицы A . Обычные матрицы появляются как частный случай, когда $S = \{1, 2, \dots, m\}$, $T = \{1, 2, \dots, n\}$.

Множество матриц с компонентами из \mathcal{A} , строки и столбцы которых занумерованы элементами из множеств S и T , обозначается $\mathcal{A}^{S \times T}$.

Все действия над матрицами могут быть определены и тогда, когда их строки и столбцы занумерованы элементами из произвольных конечных множеств. Так, транспонированной к матрице $A \in \mathcal{A}^{S \times T}$ называется матрица $A^T \in \mathcal{A}^{T \times S}$, такая что

$$(A^T)_{ts} = A_{st} \quad \text{для всех } s \in S, t \in T.$$

Пусть, далее, Λ – коммутативное ассоциативное кольцо с 1, S, T, U, V – конечные множества, и пусть $A \in \Lambda^{S \times T}$, $B \in \Lambda^{U \times V}$. Если $a \in \Lambda$, то произведением a на A называется матрица $aA \in \Lambda^{S \times T}$, элементы которой равны произведениям a на соответствующие элементы матрицы A :

$$(aA)_{st} = aA_{st} \quad (s \in S, t \in T).$$

Сумма $A + B$ матриц A, B определена тогда и только тогда, когда $S = U$, $T = V$; она тоже принадлежит $\Lambda^{S \times T}$, и ее элементы равны суммам соответствующих элементов матриц A, B :

$$(A + B)_{st} = A_{st} + B_{st} \quad (s \in S, t \in T).$$

Произведение матриц A, B определено тогда и только тогда, когда $T = U$; оно принадлежит $\Lambda^{S \times V}$ и его элементы вычисляются по формуле

$$(AB)_{sv} = \sum_{t \in T} A_{st} B_{tv} \quad (s \in S, v \in V).$$

Все свойства действий над матрицами без труда переносятся на матрицы, строки и столбцы которых занумерованы элементами из конечных множеств. В частности, множество $\Lambda_S = \Lambda^{S \times S}$ квадратных матриц, строки и столбцы которых

занумерованы элементами из S , является ассоциативным, но, вообще говоря, не коммутативным кольцом с единицей $E = E_S$.

2. ОБОБЩЕНИЕ ДЕЙСТВИЙ НАД МАТРИЦАМИ

Достаточные условия для выполнимости действий над матрицами. Наше первоначальное определение матрицы состояло в том, что компоненты матрицы могут быть элементами любого множества. В частности, пусть для каждого i, j задано свое непустое множество \mathcal{A}_{ij} ($1 \leq i \leq m, 1 \leq j \leq n$). Тогда мы можем рассматривать множество матриц вида

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

компоненты $a_{ij} = A_{ij}$ которых принадлежат соответствующему множеству \mathcal{A}_{ij} . Если на каждом из множеств \mathcal{A}_{ij} определено сложение (в частности, если они являются аддитивно записанными абелевыми группами), то подобные матрицы можно складывать, и получившаяся сумма тоже будет матрицей того же типа.

Для определения произведения матриц в подобной более общей ситуации нам надо уметь перемножать элементы, принадлежащие, вообще говоря, разным множествам. Мы говорим, что на паре множеств \mathcal{A}, \mathcal{B} , задано умножение со значениями в множестве \mathcal{C} , если зафиксировано любое отображение $\Phi : \mathcal{A} \times \mathcal{B} \rightarrow \mathcal{C}$. Мы будем записывать произведение $\Phi(a, b) \in \mathcal{C}$ элементов $a \in \mathcal{A}, b \in \mathcal{B}$ через ab .

Пусть $\mathcal{A}_{ij}, \mathcal{B}_{jk}, \mathcal{C}_{ik}$ ($1 \leq i \leq m, 1 \leq j \leq n, 1 \leq k \leq p$) – непустые множества, причем на каждом из множеств \mathcal{C}_{ik} задано коммутативное ассоциативное сложение, и для каждой тройки индексов i, j, k задано умножение $\mathcal{A}_{ij} \times \mathcal{B}_{jk} \rightarrow \mathcal{C}_{ik}$. Пусть теперь A – матрица из m строк и n столбцов, такая что $A_{ij} \in \mathcal{A}_{ij}$ для любых i, j , а B – матрица из n строк и p столбцов, такая что $B_{jk} \in \mathcal{B}_{jk}$ для любых j, k . Тогда определена матрица AB , состоящая m строк и p , компоненты которой $(AB)_{ik}$ принадлежат соответствующим множествам \mathcal{C}_{ik} и находятся по формулам

$$(AB)_{ik} = \sum_{s=1}^n A_{is} B_{sk}.$$

Действительно, все произведения $A_{is} B_{sk}$ принадлежат множеству \mathcal{C}_{ik} , и их можно сложить в \mathcal{C}_{ik} , причем результат сложения n слагаемых не зависит от того, в каком порядке мы эти сложения выполняем, так что мы можем эту сумму записывать, используя знак суммирования $\sum_{s=1}^n$.

Пример: матрицы, разбитые на блоки. Пусть Λ – кольцо, и пусть $S_1, \dots, S_m; T_1, \dots, T_n$ – два набора попарно не пересекающихся непустых конечных множеств. Пусть

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

– матрица, каждая компонента a_{ij} которой сама является матрицей и принадлежит соответствующему множеству $\Lambda^{S_i \times T_j}$. Обозначим через S объединение множеств S_1, \dots, S_m , а через T – объединение множеств T_1, \dots, T_n . Пусть $A \in \Lambda^{S \times T}$ – матрица, компоненты A_{st} которой определены следующим образом: если $s \in S_i, t \in T_j$, то $A_{st} = (a_{ij})_{st}$ (заметим, что любой элемент s из объединения S попарно не пересекающихся множеств $S_i, 1 \leq i \leq m$, принадлежит единственному из этих

множеств, и точно так же каждый элемент $t \in T$ принадлежит единственному из множеств T_j , $1 \leq j \leq n$). Мы будем говорить, что матрица A разбита на блоки a_{ij} , соответствующие разбиениям $S = S_1 \cup \dots \cup S_m$, $T = T_1 \cup \dots \cup T_n$ их множеств строк и столбцов, а a является матрицей, составленной из блоков матрицы A , отвечающих этим разбиениям.

Следующее утверждение очень полезно, и мы будем часто пользоваться им.

Предложение 1. Пусть Λ – кольцо, и пусть S, T, R – непустые конечные множества.

(1) Пусть $A, B \in \Lambda^{S \times T}$, и пусть a, b – матрицы, составленные из блоков матриц A, B , отвечающих разбиениям $S = S_1 \cup \dots \cup S_m$, $T = T_1 \cup \dots \cup T_n$ множеств строк и столбцов этих матриц в объединения попарно не пересекающихся непустых подмножеств. Тогда $a + b$ – матрица, составленная из блоком матрицы $A + B \in \Lambda^{S \times T}$, отвечающих тем же разбиениям множества S строк и множества T столбцов.

(2) Пусть $A \in \Lambda^{S \times T}$, $B \in \Lambda^{T \times R}$, и пусть a, b – матрицы, составленные из блоков матриц A, B, C , отвечающих разбиениям

$$S = S_1 \cup \dots \cup S_m, \quad T = T_1 \cup \dots \cup T_n, \quad R = R_1 \cup \dots \cup R_p$$

множеств строк и столбцов этих матриц в объединения попарно не пересекающихся непустых подмножеств. Тогда ab – матрица, составленная из блоком матрицы $AB \in \Lambda^{S \times R}$, отвечающих тем же разбиениям множества S ее строк и множества R ее столбцов.

Доказательство. Это было труднее сформулировать, чем доказать; мы ограничимся доказательством утверждения (2) (утверждение (1) намного проще). Пусть $s \in S_i \subseteq S$, $r \in R_j \subseteq R$; тогда

$$\begin{aligned} (AB)_{sr} &= \sum_{t \in T} A_{st} B_{tr} = \sum_{l=1}^n \left(\sum_{t \in T_l} A_{st} B_{tr} \right) = \sum_{l=1}^n \left(\sum_{t \in T_l} (a_{il})_{st} (b_{lj})_{tr} \right) = \\ &= \sum_{l=1}^n (a_{il} b_{lj})_{sr} = \left(\sum_{l=1}^n a_{il} b_{lj} \right)_{sr} = ((ab)_{ij})_{sr}. \end{aligned}$$

Свойства обобщенных действий над матрицами. Большинство свойств действий над матрицами с компонентами из кольца Λ остаются справедливыми и в более общей ситуации, описанной в начале параграфа. Грубо говоря, если все умножения дистрибутивны и какое-то из свойств действий выполняется для одноэлементных матриц, то оно выполняется всегда, когда левая или правая часть соответствующего соотношения определена. Мы не будем пытаться точно формулировать и доказывать эти свойства в самом общем случае, хотя сделать это несложно, а ограничимся лишь ситуацией, в которой мы в дальнейшем будем эти свойства неоднократно применять.

Пусть \mathcal{A} – абелева группа относительно сложения; тогда очевидно, что сложение матриц с компонентами из \mathcal{A} ассоциативно и коммутативно.

Перейдем теперь к свойствам, в которых участвует произведение матриц. Сначала дадим определение дистрибутивного умножения. Пусть $\mathcal{A}, \mathcal{B}, \mathcal{C}$ – аддитивно записанные абелевы группы; умножение $\mathcal{A} \times \mathcal{B} \rightarrow \mathcal{C}$ называется дистрибутивным, если

$$(a_1 + a_2)b = a_1b + a_2b, \quad a(b_1 + b_2) = ab_1 + ab_2$$

для любых $a, a_1, a_2 \in \mathcal{A}$, $b, b_1, b_2 \in \mathcal{B}$. Если это умножение дистрибутивно, то для любых матриц $A \in \mathcal{A}^{m \times n}$, $B, B' \in \mathcal{B}^{n \times k}$ выполняется соотношение дистрибутивности $A(B + B') = AB + AB'$; действительно,

$$(A(B + B'))_{ij} = \sum_{s=1}^n A_{is}(B + B')_{sj} = \sum_{s=1}^n (A_{is}B_{sj} + A_{is}B'_{sj}) =$$

$$= \sum_{s=1}^n A_{is} B_{sj} + \sum_{s=1}^n A_{is} B'_{sj} = (AB)_{ij} + (AB')_{ij} = (AB + AB')_{ij}.$$

Аналогично, для любых матриц $A, A' \in \mathcal{A}^{m \times n}$, $B \in \mathcal{B}^{n \times k}$ выполняется соотношение дистрибутивности $(A + A')B = AB + A'B$.

Ассоциативность умножения матриц может быть доказана в следующем контексте. Пусть $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}, \mathcal{E}, \mathcal{F}$ – абелевы группы относительно сложения, на которых заданы дистрибутивные умножения

$$\mathcal{A} \times \mathcal{B} \rightarrow \mathcal{D}, \quad \mathcal{D} \times \mathcal{C} \rightarrow \mathcal{F}, \quad \mathcal{B} \times \mathcal{C} \rightarrow \mathcal{E}, \quad \mathcal{A} \times \mathcal{E} \rightarrow \mathcal{F},$$

причем $(ab)c = a(bc)$ для любых $a \in \mathcal{A}$, $b \in \mathcal{B}$, $c \in \mathcal{C}$. Тогда для любых матриц $A \in \mathcal{A}^{m \times n}$, $B \in \mathcal{B}^{n \times k}$, $C \in \mathcal{C}^{k \times r}$ выполняется соотношение ассоциативности $(AB)C = A(BC)$. Действительно, в доказательстве ассоциативности умножения матриц над ассоциативным кольцом Λ , которое мы повторяем ниже, использовались только эти свойства (напомним, что через I обозначено декартово произведение множеств $\{1, \dots, n\}$ и $\{1, \dots, k\}$):

$$\begin{aligned} ((AB)C)_{ij} &= \sum_{t=1}^k (AB)_{it} C_{tj} = \sum_{t=1}^k \left(\sum_{s=1}^n A_{is} B_{st} \right) C_{tj} = \sum_{t=1}^k \left(\sum_{s=1}^n (A_{is} B_{st}) C_{tj} \right) = \\ &= \sum_{(s,t) \in I} (A_{is} B_{st}) C_{tj} = \sum_{(s,t) \in I} A_{is} (B_{st} C_{tj}) = \sum_{s=1}^n \sum_{t=1}^k A_{is} (B_{st} C_{tj}) = \\ &= \sum_{s=1}^n A_{is} \left(\sum_{t=1}^k B_{st} C_{tj} \right) = \sum_{s=1}^n A_{is} (BC)_{sj} = (A(BC))_{ij}. \end{aligned}$$

3. МАТРИЦЫ НАД ПОЛЕМ

Элементарные матрицы. Всюду в этом параграфе k является фиксированным полем, и все матрицы будут матрицами с элементами из k .

Элементарными матрицами над k называются квадратные матрицы одного из видов

$$\begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & \alpha & & \\ & & & \ddots & \\ & & & & 1 \\ & & & & & \ddots \\ & & & & & & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & \dots & \lambda \\ & & & \ddots & \vdots \\ & & & & 1 \\ & & & & & \ddots \\ & & & & & & 1 \end{pmatrix},$$

в которых α, λ – элементы из k , причем $\alpha \neq 0$, все не отмеченные недиагональные элементы равны 0, а не отмеченные диагональные элементы равны 1. Таким образом, если $A \in k_n$ – элементарная матрица первого типа, то существует индекс i , $1 \leq i \leq n$, и элемент $\alpha \neq 0$ из поля k , такие что $A_{ii} = \alpha$, $A_{jj} = 1$ при $j \neq i$, $A_{lj} = 0$ при $l \neq j$. Если $A \in k_n$ – элементарная матрица второго типа, то существуют индексы $i \neq j$, $1 \leq i, j \leq n$, и элемент $\lambda \in k$, такие что $A_{ij} = \lambda$, $A_{ss} = 1$ для всех s , $A_{st} = 0$ для всех пар $(s, t) \neq (i, j)$, $s \neq t$.

Рассматривают еще и элементарные матрицы третьего типа; матрица $A \in k_n$ называется элементарной матрицей третьего типа, если существуют такие числа $i \neq j$, $1 \leq i, j \leq n$, что $A_{ij} = A_{ji} = 1$, $A_{ss} = 1$ при $s \neq i, j$, а все остальные элементы матрицы A равны 0. Таким образом, элементарная матрица третьего

типа имеет вид

$$\begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 0 & \dots & 1 \\ & & \vdots & \ddots & \vdots \\ & & 1 & \dots & 0 \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix}.$$

Впрочем, элементарные матрицы третьего типа не так важны, потому что они могут быть представлены в виде произведения элементарных матриц первых двух типов. Для сокращения записи покажем это не в общем случае, а на примере матриц порядка 2:

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

Элементарные преобразования над строками и над столбцами матрицы. Элементарным преобразованием над строками матрицы называется ее умножение слева на элементарную матрицу, а элементарным преобразованием над столбцами матрицы называется умножение матрицы на элементарную матрицу справа.

Рассмотрим элементарные преобразования подробнее. Пусть A – матрица с компонентами из k , состоящая из m строк и n столбцов, и пусть матрица B получена из нее умножением на элементарную матрицу X . Поскольку элементарная матрица квадратная, а произведение $B = XA$ определено только если число строк второго сомножителя равно числу столбцов первого сомножителя, мы получаем, что $X \in k_n$. Имеется три типа элементарных матриц. Пусть сначала X – элементарная матрица первого типа; это значит, что существуют такие i , $1 \leq i \leq m$, и $\alpha \in k$, $\alpha \neq 0$, что $X_{ii} = \alpha$, $X_{jj} = 1$ при $j \neq i$, $X_{st} = 0$ при $s \neq t$ ($1 \leq j, s, t \leq m$). Тогда

$$B_{st} = (XA)_{st} = \sum_{j=1}^m X_{sj}A_{jt} = X_{ss}A_{st} = \begin{cases} A_{st}, & \text{если } s \neq i; \\ \alpha A_{it}, & \text{если } s = i, \end{cases}$$

потому что при фиксированном s из элементов X_{sj} только диагональный элемент X_{ss} отличен от 0, и он равен 1, если $s \neq i$, и равен α при $s = i$. Таким образом, при элементарном преобразовании первого типа все элементы i -й строки матрицы A умножаются на элемент $\alpha \neq 0$ из поля k , а остальные элементы матрицы A не меняются.

Пусть теперь X – элементарная матрица второго типа; это значит, что существуют такие индексы $i \neq j$ ($1 \leq i, j \leq m$) и элемент $\lambda \in k$, что $X_{ij} = \lambda$, $X_{ss} = 1$, $X_{st} = 0$ при $s \neq t$, $(s, t) \neq (i, j)$ ($1 \leq s, t \leq m$). Тогда

$$B_{st} = (XA)_{st} = \sum_{l=1}^m X_{sl}A_{lt} = \begin{cases} X_{ss}A_{st} = A_{st}, & \text{если } s \neq i; \\ X_{ii}A_{it} + X_{ij}A_{jt} = A_{it} + \lambda A_{jt}, & \text{если } s = i, \end{cases}$$

потому что при фиксированном $s \neq i$ из элементов X_{sl} только диагональный элемент X_{ss} отличен от 0, и он равен 1, а среди элементов X_{il} не равны 0 только $X_{ii} = 1$ и $X_{ij} = \lambda$. Таким образом, при элементарном преобразовании второго типа ко всем элементам i -й строки матрицы A прибавляются соответствующие элементы j -й строки, умноженные на элемент λ из поля k , а остальные элементы матрицы A не меняются.

Наконец, если X – элементарная матрица третьего типа, то существуют такие индексы $i \neq j$ ($1 \leq i, j \leq m$), что среди элементов X_{it} отличен от 0 только элемент

$X_{ij} = 1$, среди элементов X_{jt} отличен от 0 только элемент $X_{ji} = 1$, а при $s \neq i, j$ из элементов X_{st} только диагональный элемент $X_{ss} = 1$ не равен 0; поэтому

$$B_{st} = (XA)_{st} = \sum_{l=1}^m X_{sl}A_{lt} = \begin{cases} X_{ss}A_{st} = A_{st}, & \text{если } s \neq i, j; \\ X_{ij}A_{jt} = A_{jt}, & \text{если } s = i, \\ X_{ji}A_{it} = A_{it}, & \text{если } s = j. \end{cases}$$

Таким образом, при элементарном преобразовании третьего типа i -я и j -я строки меняются местами, а остальные элементы матрицы A не меняются.

Рассуждая аналогично, мы находим, что существуют три типа элементарных преобразований над столбцами матрицы A :

- (1) все элементы j -го столбца матрицы A умножаются на элемент $\alpha \neq 0$ из поля k , а остальные элементы матрицы A не меняются;
- (2) ко всем элементам i -го столбца матрицы A прибавляются соответствующие элементы j -го столбца, умноженные на элемент λ из поля k , а остальные элементы матрицы A не меняются;
- (3) i -й и j -й столбцы матрицы A меняются местами, а остальные элементы матрицы A не меняются.

Приведение матрицы элементарными преобразованиями. Прежде, чем сформулировать основную теорему этого пункта, дадим одно определение. Матрица $A \in k^{m \times n}$ называется *лестницеобразной*, если существуют такие числа r, j_1, \dots, j_r , что $0 \leq r \leq m, 1 \leq j_1 < \dots < j_r \leq n$ и выполняются условия:

- (1) $A_{1j_1} = \dots = A_{rj_r} = 1$;
- (2) если $s \leq r$ и $i \neq s$, то $A_{ij_s} = 0$;
- (3) если $i > r$, то $A_{ij} = 0$ для всех $j, 1 \leq j \leq n$;
- (4) если $i \leq r, j < j_i$, то $A_{ij} = 0$.

Если $r = 0$, то количество индексов j_1, \dots, j_r равно 0, т.е. их попросту нет; в этом случае лестницеобразная матрица оказывается нулевой матрицей. Если число строк или число столбцов матрицы равно 0, то все условия выполняются при $r = 0$, так что матрицы с пустым множеством строк или столбцов являются лестницеобразными.

Следующий пример, в котором $m = 6, n = 9, r = 4, j_1 = 2, j_2 = 3, j_3 = 6, j_4 = 8$, иллюстрирует это определение и, можно надеяться, дает полное представление о том, что такое лестницеобразная матрица (на позициях "под лестницей", отмеченных знаком \cdot , стоят нули):

$$\begin{pmatrix} \cdot & 1 & 0 & a_{14} & a_{15} & 0 & a_{17} & 0 & a_{19} \\ \cdot & \cdot & 1 & a_{24} & a_{25} & 0 & a_{27} & 0 & a_{29} \\ \cdot & \cdot & \cdot & \cdot & \cdot & 1 & a_{37} & 0 & a_{39} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & a_{49} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix}.$$

Отметим, что длины ступенек могут быть различными (в нашем примере они равнялись 1, 3, 2, 2).

Следующее простое утверждение немедленно следует из определения.

Лемма 1. Если к лестницеобразной матрице с $r \geq 0$ ненулевыми строками приписать справа столбец, все элементы которого, кроме $(r+1)$ -го, равны 0, а $(r+1)$ -й элемент равен 1, то получившаяся матрица тоже будет лестницеобразной.

Теорема 1. Всякую матрицу над полем можно последовательностью элементарных преобразований над строками привести к лестницеобразной матрице.

Доказательство. Для всякой матрицы B обозначим через $B^{[i]}$ матрицу, составленную из первых i столбцов матрицы B . Индукцией по i докажем следующее утверждение:

(*) Пусть k – поле, и пусть $A \in k^{m \times n}$. Для всякого $j \leq n$ матрицу A можно последовательностью элементарных преобразований над строками перевести в такую матрицу B , что $B^{[j]}$ – лестницеобразная матрица.

При $j = n$ мы получаем утверждение теоремы; для $j = 0$ утверждение бесспорно. Пусть $1 < j \leq n$, и матрица A последовательностью элементарных преобразований над строками уже приведена к такому виду B , что $B^{[j-1]}$ – лестницеобразная матрица. Обозначим через s число ступенек (т.е. число ненулевых строк) матрицы $B^{[j]}$. Если $B_{ij} = 0$ для всех $i > s$, то матрица $B^{[j]}$ уже является лестницеобразной, и никаких дополнительных преобразований делать не надо. Если же найдется такое $i > s$, что элемент $\alpha = B_{ij}$ отличен от 0, то делаем над строками матрицы B элементарные преобразования

$$B \xrightarrow{(1)} C \xrightarrow{(2)} D \xrightarrow{(3)} F,$$

описания которых мы сейчас дадим.

- (1) Умножаем все элементы i -й строки на α^{-1} ; тогда $C_{ij} = 1$, а все остальные элементы j -го столбца матрицы C – такие же, как у матрицы B .
- (2) (На самом деле это не одно, а $m - 1$ элементарное преобразование). Для каждого $l \neq i$, $1 \leq l \leq m$, прибавляем ко всем элементам l -й строки матрицы C соответствующие элементы i -й строки, умноженные на $-C_{lj} = -B_{lj}$; в получившейся матрице D все элементы D_{lj} j -го столбца, кроме элемента $D_{ij} = 1$, равны 0.
- (3) Переставляем i -ю и $(s + 1)$ -ю строки; в получившейся матрице F все элементы j -го столбца равны 0, кроме элемента $F_{s+1,j}$, который равен 1.

Очевидно, все эти преобразования не затрагивают первые $j - 1$ столбцов матрицы; поэтому $F^{[j-1]} = B^{[j-1]}$ – лестницеобразная матрица с s ненулевыми строками. По лемме 1, матрица $F^{[j]}$, получающаяся из $F^{[j-1]}$ приписыванием справа столбца, все элементы которого, кроме $(s + 1)$ -го, равны 0, а $(s + 1)$ -й элемент равен 1, тоже является лестницеобразной. Утверждение (*) доказано; вместе с ним доказана и теорема 1.

Отметим, что доказательство теоремы не только показывает возможность приведения матрицы к лестницеобразному виду, но и указывает алгоритм, при помощи которого это можно сделать. Этот алгоритм является одним из важнейших алгоритмов линейной алгебры и используется для решения многих задач. В частности, на этом алгоритме основаны почти все методы решения систем линейных уравнений.

Теорема 2. *Всякую матрицу над полем можно последовательностью элементарных преобразований над строками и столбцами привести к виду*

$$\begin{pmatrix} E_r & \mathbf{0}^{r \times t} \\ \mathbf{0}^{s \times r} & \mathbf{0}^{s \times t} \end{pmatrix},$$

где E_r – единичная матрица порядка $r \geq 0$, а $\mathbf{0}^{r \times t}$, $\mathbf{0}^{s \times r}$, $\mathbf{0}^{s \times t}$ – нулевые матрицы.

Доказательство. По теореме 1 любую матрицу элементарными преобразованиями над строками можно превратить в лестницеобразную матрицу B . Пусть r – число ступенек этой матрицы, и пусть $B_{1j_1} = \dots = B_{rj_r} = 1$ – начала этих ступенек, так что $1 \leq j_1 < \dots < j_r$, $B_{ij} = 0$ при $i > r$ и $B_{sj_s} = 1$ – единственный ненулевой элемент в j_s -м столбце $1 \leq s \leq r$. Теперь элементарными преобразованиями над столбцами матрицы "убьем" все остальные элементы первых r строк;

точнее, для каждого $s \leq r$ и для каждого $j > j_s$ прибавим ко всем элементам j -го столбца матрицы B соответствующие элементы j_s -го столбца, умноженные на $-B_{sj}$. В результате получится матрица C , в которой $C_{1j_1} = \dots = C_{rj_r} = 1$, а все остальные элементы равны 0. Остается переставить столбцы матрицы C , чтобы получить матрицу, указанную в формулировке теоремы.

Поскольку элементарное преобразование над строками матрицы равносильно умножению слева на элементарную матрицу, а элементарное преобразование над столбцами – умножению на элементарную матрицу справа, мы можем переформулировать теорему 2 следующим образом.

Теорема 3. Пусть k – поле, и пусть $A \in k^{m \times n}$. Тогда существуют такое натуральное число r и такие матрицы $X \in k_m$, $Y \in k_n$, каждая из которых раскладывается в произведение элементарных матриц, что

$$XAY = \begin{pmatrix} E_r & \mathbf{0}^{r \times t} \\ \mathbf{0}^{s \times r} & \mathbf{0}^{s \times t} \end{pmatrix},$$

где через s и t обозначены разности $m - r$, $n - r$.

4. ОБРАТИМЫЕ МАТРИЦЫ

Обратная матрица. В этом параграфе мы, как и в предыдущем, рассматриваем только матрицы с компонентами из некоторого поля k . Матрица B называется обратной к матрице $A \in k^{m \times n}$, если AB и BA – единичные матрицы. Поскольку число строк матрицы AB равно m , а число столбцов матрицы BA равно n , эти единичные матрицы должны иметь соответственно порядки m и n : $AB = E_m$, $BA = E_n$.

Предложение 1. Если для матрицы $A \in k^{m \times n}$ существует обратная матрица, то только одна.

Доказательство. Пусть B, B_1 – обратные к A матрицы, то $AB = AB_1 = E_m$, $BA = B_1A = E_n$, и потому $B_1 = B_1E_m = B_1(AB) = (B_1A)B = E_nB = B$.

Матрица A , у которой есть обратная, называется обратимой, а единственная обратная к ней матрица обозначается A^{-1} .

Предложение 2. (1) Единичная матрица E_n обратима, и $E_n^{-1} = E_n$;

(2) если A, B – обратимые матрицы, и произведение AB определено, то AB – обратимая матрица, причем $(AB)^{-1} = B^{-1}A^{-1}$;

(3) если A – обратимая матрица, то A^{-1} – тоже обратимая матрица, причем $(A^{-1})^{-1} = A$;

(4) если A – обратимая матрица, то A^T – тоже обратимая матрица, причем $(A^T)^{-1} = (A^{-1})^T$.

Доказательство. Утверждение (1) очевидно. Если A, B – обратимые матрицы, причем произведение AB определено, то

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AEA^{-1} = AA^{-1} = E,$$

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}EB = B^{-1}B = E,$$

т.е. $B^{-1}A^{-1}$ – обратная к AB матрица. Если матрица A обратима, то равенства $AA^{-1} = E$, $A^{-1}A = E$ показывают не только то, что A^{-1} – обратная к A матрица, но и что A – обратная к A^{-1} матрица. Наконец, транспонируя равенства $AA^{-1} = E$, $A^{-1}A = E$ и пользуясь тем, что, очевидно, $E^T = E$, получим соотношения

$$(A^{-1})^T A^T = E^T = E, \quad A^T (A^{-1})^T = E^T = E,$$

которые показывают, что у матрицы A^T есть обратная матрица $(A^{-1})^T$.

Предложение 3. *Элементарные матрицы обратимы, причем обратные к ним матрицы тоже являются элементарными матрицами.*

Доказательство. Элементарные матрицы

$$\begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & \alpha^{-1} & & \\ & & & \ddots & \\ & & & & 1 \\ & & & & & \ddots \\ & & & & & & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & \dots & -\lambda \\ & & & \ddots & \vdots \\ & & & & 1 \\ & & & & & \ddots \\ & & & & & & 1 \end{pmatrix}$$

являются, как легко проверить, обратными к элементарным матрицам первого и второго типов

$$\begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & \alpha & & \\ & & & \ddots & \\ & & & & 1 \\ & & & & & \ddots \\ & & & & & & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & \dots & \lambda \\ & & & \ddots & \vdots \\ & & & & 1 \\ & & & & & \ddots \\ & & & & & & 1 \end{pmatrix}$$

(здесь $\alpha, \lambda \in k$, $\alpha \neq 0$). Элементарная матрица третьего типа, очевидно, сама является обратной к себе матрицей.

Предложение 4. *Пусть*

$$A = \begin{pmatrix} E_r & \mathbf{0}^{r \times t} \\ \mathbf{0}^{s \times r} & \mathbf{0}^{s \times t} \end{pmatrix},$$

где E_r – единичная матрица порядка $r \geq 0$, а $\mathbf{0}^{r \times t}$, $\mathbf{0}^{s \times r}$, $\mathbf{0}^{s \times t}$ – нулевые матрицы. Матрица A обратима тогда и только тогда, когда $s = t = 0$.

Доказательство. Для любой матрицы B последние s строк произведения

$$AB = \begin{pmatrix} E_r & \mathbf{0}^{r \times t} \\ \mathbf{0}^{s \times r} & \mathbf{0}^{s \times t} \end{pmatrix} B$$

состоят лишь из нулей; поэтому это произведение может быть единичной матрицей только если $s = 0$. Следовательно, если $s \neq 0$, то матрица A не обратима. Точно так же, A не обратима, если $t \neq 0$. Если же $s = t = 0$, то матрица A превращается в единичную матрицу E_r , которая, конечно, обратима.

Критерий обратимости матрицы. . В последующих главах мы увидим, что обратимость матрицы равносильна многим разнообразным условиям, формулируемым в совершенно различных терминах. Сейчас мы укажем одно из таких условий.

Теорема 1. *Пусть k – поле. Матрица $A \in k^{m \times n}$ обратима тогда и только тогда, когда $m = n$ и матрица A является произведением элементарных матриц.*

Доказательство. Все элементарные матрицы обратимы, а потому их произведение – тоже обратимая матрица. Обратно, пусть матрица A обратима. По теореме

3.3 существуют такое натуральное число r и такие матрицы $X \in k_m$, $Y \in k_n$, каждая из которых раскладывается в произведение элементарных матриц, что

$$XAY = \begin{pmatrix} E_r & \mathbf{0}^{r \times t} \\ \mathbf{0}^{s \times r} & \mathbf{0}^{s \times t} \end{pmatrix},$$

где через s и t обозначены разности $m - r$, $n - r$. Матрицы X , Y являются произведениями элементарных матриц и потому обратимы; следовательно, и произведение трех обратимых матриц

$$\begin{pmatrix} E_r & \mathbf{0}^{r \times t} \\ \mathbf{0}^{s \times r} & \mathbf{0}^{s \times t} \end{pmatrix} = XAY$$

– тоже обратимая матрица. По предложению 4 это возможно только тогда, когда $s = t = 0$ и потому $m = r = n$. Теперь, вспоминая, что матрицы X и Y обратимы, получаем:

$$A = E_m A E_n = (X^{-1}X)A(YY^{-1}) = X^{-1}(XAY)Y^{-1} = X^{-1}E_n Y^{-1} = X^{-1}Y^{-1}.$$

Каждая из матриц X , Y является произведением элементарных матриц; поэтому обратные к ним матрицы X^{-1} , Y^{-1} являются произведениями (в обратном порядке) матриц, обратных к элементарным. Но, как мы знаем, матрица, обратная к элементарной, сама является элементарной, и потому матрицы X^{-1} , Y^{-1} , а вместе с ними и их произведение $A = X^{-1}Y^{-1}$, раскладываются в произведения элементарных матриц.

Вычисление обратной матрицы. Изложенные выше идеи позволяют дать алгоритм, позволяющий выяснить, обратима ли квадратная матрица, и если она обратима, найти обратную.

Пусть A – квадратная матрица порядка n над некоторым полем. Припишем к ней справа единичную матрицу порядка n и получившуюся матрицу

$$B = (A | E)$$

элементарными преобразованиями над строками приведем к лестницеобразной матрице D , которую тоже разобьем на два квадратных блока:

$$D = (U | V).$$

Каждое элементарное преобразование над строками матрицы B сводится к умножению слева на элементарную матрицу, а вся последовательность элементарных преобразований, переводящих матрицу B в матрицу D – к умножению на произведение X этих элементарных матриц. Таким образом,

$$(U | V) = D = XB = X(A | E) = (XA | XE),$$

откуда следует, что $U = XA$, $V = XE = X$. Но произведение X элементарных матриц является обратимой матрицей; из равенства $U = XA$ мы теперь получаем, что

$$A = (X^{-1}X)A = X^{-1}(XA) = X^{-1}U = V^{-1}U.$$

Возможны два случая.

(1) $U = E$. Тогда матрица A обратима, и $A^{-1} = V$. Действительно, в этом случае $A = V^{-1}U = V^{-1}E = V^{-1}$; по предложению 2 матрица A , обратная к обратимой матрице V , сама обратима, и $A^{-1} = (V^{-1})^{-1} = V$.

(2) $U \neq E$. Тогда матрица A необратима. Действительно, матрица U , получающаяся из лестницеобразной матрицы D выбрасыванием нескольких последних столбцов, сама лестницеобразная; но легко видеть, что единственной квадратной лестницеобразной матрицей, у которой все строки ненулевые, является единичная матрица. Поскольку $U \neq E$, отсюда следует, что у матрицы U есть нулевые строки, и поэтому будут нулевые строки и у любого произведения UY , т.е. это произведение никогда не сможет стать единичной матрицей. Поэтому матрица

$U = XA$ необратима, а значит, необратима и матрица A : в противном случае произведение XA двух обратимых матриц было бы обратной матрицей.

5. Тело кватернионов

Построение тела кватернионов. До сих пор мы встречались в нашем курсе с многими полями. Напомним, что поле – это коммутативное ассоциативное кольцо с 1, в котором каждый ненулевой элемент обратим. Однако, встречаются и другие алгебраические структуры, которые отличаются от полей только тем, что умножение в них не обязательно коммутативно; они называются телами. Таким образом, множество T , на котором заданы сложение и умножение, называется телом, если выполняются следующие свойства:

- (1) $a + (b + c) = (a + b) + c$ для любых $a, b, c \in T$ (ассоциативность сложения);
- (2) $a + b = b + a$ для любых $a, b \in T$ (коммутативность сложения);
- (3) существует такой элемент $0 \in T$, что $0 + a = a$ для любого $a \in T$;
- (4) для любого $a \in T$ существует такой элемент $-a \in T$, что $a + (-a) = 0$;
- (5) $a(bc) = (ab)c$ для любых $a, b, c \in T$ (ассоциативность умножения);
- (6) существует такой элемент $1 \in T$, что $1 \cdot a = a \cdot 1 = a$ для любого $a \in T$;
- (7) $a(b + c) = ab + ac$, $(a + b)c = ac + bc$, $a(b \cdot c) = ab \cdot c$ для любых $a, b, c \in T$ (дистрибутивность умножения относительно сложения);
- (8) для всякого $a \in T$, $a \neq 0$, существует элемент $a^{-1} \in T$, такой что $aa^{-1} = a^{-1}a = 1$.

Подчеркнем, что в аксиоме 7 участвуют два соотношения – левая дистрибутивность и правая дистрибутивность, а в аксиоме 8 мы требуем, чтобы единице 1 равнялись оба произведения aa^{-1} и $a^{-1}a$, а не одно из них; конечно, это связано с тем, что умножение не обязано быть коммутативным. Всякое коммутативное тело является полем.

Мы сейчас построим пример некоммутативного тела, используя для этой цели матрицы. В кольце матриц второго порядка над полем комплексных чисел \mathbb{C} рассмотрим подмножество \mathbb{H} , состоящее из всех матриц вида

$$\begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix}, \quad \text{где } \alpha, \beta \in \mathbb{C}.$$

Предложение 1. *Относительно матричных сложения и умножения множество \mathbb{H} представляет собой некоммутативное тело.*

Доказательство. Покажем, что сумма, разность и произведение любых двух матриц из T снова принадлежат T . Действительно, пусть

$$X = \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix}, \quad Y = \begin{pmatrix} \gamma & \delta \\ -\bar{\delta} & \bar{\gamma} \end{pmatrix} \quad (\alpha, \beta, \gamma, \delta \in \mathbb{C});$$

тогда

$$X \pm Y = \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} \pm \begin{pmatrix} \gamma & \delta \\ -\bar{\delta} & \bar{\gamma} \end{pmatrix} = \begin{pmatrix} \alpha \pm \gamma & \beta \pm \delta \\ -\bar{\beta} \mp \bar{\delta} & \bar{\alpha} \pm \bar{\gamma} \end{pmatrix},$$

$$XY = \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} \begin{pmatrix} \gamma & \delta \\ -\bar{\delta} & \bar{\gamma} \end{pmatrix} = \begin{pmatrix} \alpha\gamma - \beta\bar{\delta} & \alpha\delta + \beta\bar{\gamma} \\ -\bar{\beta}\gamma - \bar{\alpha}\bar{\delta} & -\bar{\beta}\delta + \bar{\alpha}\bar{\gamma} \end{pmatrix},$$

и мы видим, что получившиеся матрицы принадлежат T .

Единичная и нулевая матрицы, очевидно, принадлежат T , и аксиомы (1)-(7) выполняются в T , потому что они выполняются во всем кольце матриц \mathbb{C}_2 . Осталось доказать, что любая ненулевая матрица из T обратима, и обратная к ней матрица тоже принадлежит T . Пусть $X = \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix}$ – ненулевая матрица; тогда хотя бы одно из чисел α, β отлично от 0, и потому вещественное число $R = \alpha\bar{\alpha} + \beta\bar{\beta} = |\alpha|^2 + |\beta|^2$ тоже не равно 0. Мы явно укажем обратную матрицу:

$$\begin{pmatrix} \bar{\alpha}/R & -\beta/R \\ \bar{\beta}/R & \alpha/R \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} = \begin{pmatrix} (\bar{\alpha}\alpha + \beta\bar{\beta})/R & 0 \\ 0 & (\bar{\beta}\beta + \alpha\bar{\alpha})/R \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

так что матрица $\begin{pmatrix} \bar{\alpha}/R & -\beta/R \\ \bar{\beta}/R & \alpha/R \end{pmatrix}$ является обратной к матрице X , и ясно, что она принадлежит T .

Для того, чтобы лучше понять структуру этого тела T , представим его элементы в немного другом виде. Пусть $X = \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix}$ – произвольный элемент из T , и пусть $\alpha = a + bi$, $\beta = c + di$, где a, b, c, d – вещественные числа. Тогда

$$X = \begin{pmatrix} a + bi & c + di \\ -c + di & a - bi \end{pmatrix} = a \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + b \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} + c \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} + d \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.$$

Матрица, коэффициентом при которой служит a , является единичной матрицей E ; обозначим через I, J, K матрицы, которые входят в предыдущее выражение с коэффициентами b, c, d . Таким образом, любой элемент из T представляется в виде $aE + bI + cJ + dK$, где a, b, c, d – вещественные числа, причем такое представление, очевидно, единственно. Непосредственным вычислением находим, что

$$I^2 = J^2 = K^2 = -E, \quad IJ = -JI = K, \quad JK = -KJ = I, \quad KI = -IK = J.$$

В частности, $IJ = -JI \neq JI$, так что T – некоммутативное тело. Заметим еще, что все скалярные матрицы aE ($a \in \mathbb{R}$) перестановочны со всеми матрицами из T .

Телом кватернионов называется тело \mathbb{H} , содержащее поле вещественных чисел \mathbb{R} и такие элементы i, j, k , что

- (1) всякий элемент из \mathbb{H} однозначно представляется в виде $a + bi + cj + dk$ с вещественными a, b, c, d ;
- (2) если $a \in \mathbb{R}$, $x \in \mathbb{H}$, то $ax = xa$;
- (3) $i^2 = j^2 = k^2 = -1$, $ij = -ji = k$, $jk = -kj = i$, $ki = -ik = j$.

Из этого описания еще не следует, что наше определение тела кватернионов непротиворечиво, и что такой объект действительно является телом. Но тело T , построенное выше, удовлетворяет всем требованиям определения тела кватернионов: скалярные матрицы aE , где $a \in \mathbb{R}$, образуют подполе T , которое можно отождествить с \mathbb{R} , а матрицы I, J, K удовлетворяют тем тождествам, которым должны удовлетворять элементы i, j, k . Таким образом, тело кватернионов существует и действительно является телом.

Название "кватернион" имеет тот же корень, что и слова квадрат, квартет, кварта и многие другие; этот латинский корень означает "четыре", а кватернион задается четверкой вещественных чисел, откуда и происходит его название. Обозначение \mathbb{H} напоминает нам о том, что впервые кватернионы были открыты ирландским математиком Уильямом Гамильтоном (W.Hamilton).

Векторы и кватернионы. Пусть $x = a + bi + cj + dk$ – кватернион (здесь a, b, c, d – вещественные числа); число a называется вещественной, или скалярной, частью кватерниона x , а сумма остальных трех компонент $bi + cj + dk$ – его векторной частью. Причины такого названия ясны: всякий вектор в трехмерном пространстве представляется в виде линейной комбинации ортов $\vec{i}, \vec{j}, \vec{k}$, и кватернион $bi + cj + dk$ естественно отождествить с вектором $b\vec{i} + c\vec{j} + d\vec{k}$, что мы в дальнейшем и будем делать. Кватернион с нулевой скалярной частью мы будем называть векторным кватернионом или просто вектором. Таким образом, всякий кватернион представляется в виде суммы вещественного числа a и вектора в трехмерном пространстве $v = bi + cj + dk$.

Умножение векторных кватернионов очень красиво интерпретируется геометрически. Мы будем обозначать через (u, v) скалярное произведение векторов u и v трехмерного пространства, а через $u \times v$ – их векторное произведение. Пусть $u = b_1i + c_1j + d_1k$, $v = b_2i + c_2j + d_2k$ – два векторных кватерниона; тогда

$$\begin{aligned} uv &= (b_1i + c_1j + d_1k)(b_2i + c_2j + d_2k) = b_1b_2i^2 + b_1c_2ij + b_1d_2ik + c_1b_2ji + c_1c_2j^2 + \\ &+ c_1d_2jk + d_1b_2ki + d_1c_2kj + d_1d_2k^2 = -(b_1b_2 + c_1c_2 + d_1d_2) + (c_1d_2 - c_2d_1)i + \\ &+ (-b_1d_2 + b_2d_1)j + (b_1c_2 - b_2c_1)k = -(u, v) + u \times v. \end{aligned}$$

Таким образом, вещественная часть произведения векторных кватернионов равна их скалярному произведению, взятому со знаком $-$, а векторная часть – их векторному произведению.

Сопряжение кватернионов. Норма и тригонометрическая форма записи кватерниона. Пусть $x = a + bi + cj + dk \in \mathbb{H}$, где a, b, c, d – вещественные числа; сопряженным к кватерниону x называется кватернион $\bar{x} = a - bi - cj - dk$. Иначе говоря, если $x = a + u$, где a – вещественная, а u – векторная части кватерниона x , то $\bar{x} = a - u$. Пусть $y = a_1 + u_1$ – другой кватернион (здесь опять все коэффициенты a_1 и u_1 – вещественная и векторная части кватерниона y); тогда

$$\begin{aligned} \overline{x+y} &= \overline{(a+a_1) + (u+u_1)} = (a+a_1) - (u+u_1) = (a-u) + (a_1-u_1) = \bar{x} + \bar{y}, \\ \overline{xy} &= \overline{(a+u)(a_1+u_1)} = \overline{aa_1 + au_1 + a_1u + uu_1} = \overline{aa_1 - (u, u_1) + au_1 + a_1u + (u \times u_1)} = \\ &= aa_1 - (u, u_1) - au_1 - a_1u - (u \times u_1) = aa_1 - (u_1, u) - au_1 - a_1u + (u_1 \times u) = \\ &= aa_1 - au_1 - a_1u + u_1u = (a-u)(a_1-u_1) = \bar{y}\bar{x}. \end{aligned}$$

Обратим внимание на то, что при сопряжении произведения сомножители не только заменяются на сопряженные, но еще и переставляются в обратном порядке.

Нормой кватерниона $x = a + bi + cj + dk$, где a, b, c, d – вещественные числа, называется неотрицательное вещественное число $\|x\| = a^2 + b^2 + c^2 + d^2$. Обозначая, как и выше, через u вектор $bi + cj + dk$, найдем, что

$$\begin{aligned} x\bar{x} &= (a+u)(a-u) = a^2 - au + au - u^2 = a^2 - u^2 = a^2 - (-(u, u) + (u \times u)) = \\ &= a^2 + (u, u) = a^2 + b^2 + c^2 + d^2 = \|x\|, \end{aligned}$$

и точно так же показывается, что $\bar{x}x = \|x\|$.

Иногда бывает удобно представлять кватернионы в форме, аналогичной тригонометрической форме комплексного числа. $x = a + bi + cj + dk$, где a, b, c, d – вещественные числа, причем $x \neq 0$, т.е. $\|x\| > 0$. Пусть $r = \sqrt{\|x\|}$. Как и раньше, обозначим через u вектор $bi + cj + dk$. Тогда $r^2 = \|x\| = a^2 + b^2 + c^2 + d^2 = a^2 + \|u\|^2$, где через $\|u\|$ обозначена длина вектора u ; поэтому

$$\left(\frac{a}{r}\right)^2 + \left(\frac{\|u\|}{r}\right)^2 = 1,$$

и существует такое вещественное число φ , что

$$\cos \varphi = \frac{a}{r}, \quad \sin \varphi = \frac{\|u\|}{r}.$$

Заметим, что при нашем определении $\sin \varphi = \|u\|/r \geq 0$. Если $\sin \varphi > 0$, положим $v = u/r \sin \varphi$; тогда длина $\|v\|$ вектора v равна $\|u\|/r \sin \varphi = 1$. Если же $\sin \varphi = 0$, то выберем в качестве v любой вектор длины 1. В обоих случаях $x = r(\cos \varphi + v \sin \varphi)$, где v – векторный кватернион длины 1. Аналогия с тригонометрической формой комплексного числа станет еще большей, когда мы заметим, что $v^2 = -(v, v) + v \times v = -\|v\|^2 = -1$. Это, между прочим, показывает, что в теле кватернионов уравнение $x^2 = -1$ имеет не два, как было бы в поле, а бесконечно много решений.

Формула Эйлера для произведения сумм четырех квадратов. Используя кватернионы, мы легко получаем формулу, представляющую произведение двух сумм четырех квадратов в виде суммы четырех квадратов.

Лемма. Для любых кватернионов $x, y \in \mathbb{H}$ норма $\|xy\|$ их произведения равна произведению норм сомножителей.

Доказательство. Пользуясь тем, что вещественное число можно в произведении переставить с любым кватернионом, получим:

$$\|xy\| = (xy)(\overline{xy}) = xy\overline{y}\overline{x} = x\|y\|\overline{x} = x\overline{x}\|y\| = \|x\| \cdot \|y\|.$$

Пусть теперь $x = a + bi + cj + dk$, $y = a_1 + b_1i + c_1j + d_1k$ – два кватерниона (здесь $a, b, c, d, a_1, b_1, c_1, d_1$ – вещественные числа). Сосчитаем их произведение:

$$xy = (a + bi + cj + dk)(a_1 + b_1i + c_1j + d_1k) = (aa_1 + bb_1 + cc_1 + dd_1) + \\ + (ab_1 + ba_1 + cd_1 - dc_1)i + (ac_1 - bd_1 + ca_1 + db_1)j + (ad_1 + bc_1 - cb_1 + da_1)k.$$

По лемме, произведение норм кватернионов x, y равно норме кватерниона xy , т.е. выполняется тождество

$$(a^2 + b^2 + c^2 + d^2)(a_1^2 + b_1^2 + c_1^2 + d_1^2) = (aa_1 + bb_1 + cc_1 + dd_1)^2 + \\ + (ab_1 + ba_1 + cd_1 - dc_1)^2 + (ac_1 - bd_1 + ca_1 + db_1)^2 + (ad_1 + bc_1 - cb_1 + da_1)^2.$$

Это тождество называется тождеством Эйлера. Наше доказательство проходит лишь для вещественных чисел, но на самом деле его легко распространить на все коммутативные ассоциативные кольца. Заметим еще, что левая и правая части этого тождества представляют собой многочлены от $a, b, c, d, a_1, b_1, c_1, d_1$; мы показали, что все их значения равны, но по теореме о формальном и функциональном равенстве для многочленов от нескольких переменных отсюда следует, что и сами многочлены равны.

Тождества, аналогичные тождеству Эйлера, существуют еще только для сумм двух и сумм восьми квадратов. Для сумм двух квадратов доказательство аналогично приведенному выше рассуждению; надо только вместо нормы кватерниона использовать квадрат модуля комплексного числа.

Представление рационального числа в виде суммы четырех квадратов. Мы используем кватернионы для доказательства следующего теоретико-числового результата.

Теорема 1. Всякое положительное рациональное число может быть представлено в виде суммы квадратов четырех рациональных чисел.

Замечание. На самом деле справедливо и более сильное утверждение: всякое натуральное число может быть представлено в виде суммы квадратов четырех натуральных чисел. Но мы ограничимся здесь доказательством лишь ослабленной формы этого классического результата.

Доказательство. Достаточно доказать, что любое положительное целое число является суммой квадратов четырех рациональных чисел. Действительно, если $a/b \in \mathbb{Q}$, где a, b положительные целые числа, и целое число ab равно сумме квадратов рациональных чисел $\alpha, \beta, \gamma, \delta$, то

$$a/b = (\alpha/b)^2 + (\beta/b)^2 + (\gamma/b)^2 + (\delta/b)^2$$

– представление a/b в виде суммы квадратов четырех рациональных чисел.

Предположим, что наше утверждение неверно; пусть p – наименьшее положительное целое число, не представимое в виде суммы квадратов четырех рациональных чисел. Покажем, что p – нечетное простое число. В самом деле, $p \neq 1, 2$, так как $1 = 1^2 + 0^2 + 0^2 + 0^2$, $2 = 1^2 + 1^2 + 0^2 + 0^2$. Если p – не простое число, то p раскладывается в произведение $p = ab$ натуральных чисел a, b , каждое из

которых строго меньше p и поэтому представляется в виде суммы квадратов четырех рациональных чисел; но тогда по тождеству Эйлера и их произведение p тоже представляется в виде суммы квадратов четырех рациональных чисел, что противоречит нашему предположению.

Лемма 1. *Существуют такие целые числа a, b, c , не все равные 0, что*

$$a^2 + b^2 + c^2 \div p, \quad a^2 + b^2 + c^2 < p^2.$$

Доказательство. Если -1 является квадратичным вычетом по модулю p , то существует целое число a , такое что $0 < a \leq p-1$ и $a^2 \equiv -1 \pmod{p}$. Тогда

$$a^2 + 1^2 + 0^2 = a^2 + 1 \div p, \quad a^2 + 1^2 + 0^2 \leq (p-1)^2 + 1 = p^2 - 2p < p^2.$$

Пусть теперь -1 – квадратичный невычет по модулю p , и пусть d – наименьший квадратичный невычет по модулю p ; тогда $2 \leq d < p$, а $d-1$ и $-d$ квадратичные вычеты (последнее потому, что произведение квадратичных невычетов является квадратичным вычетом). Следовательно, существуют такие целые числа a_1, b_1 , что $a_1^2 \equiv d-1 \pmod{p}$, $b_1^2 \equiv p-d \pmod{p}$, $0 < a_1, b_1 < p$. Пусть a – то из чисел $a_1, p-a_1$, которое меньше $p/2$, и аналогично b – то из чисел $b_1, p-b_1$, которое меньше $p/2$. Тогда

$$a^2 + b^2 + 1^2 \equiv (d-1) + (p-d) + 1 \equiv 0 \pmod{p}, \quad a^2 + b^2 + 1^2 < 3p^2/4 < p^2.$$

Вернемся к доказательству теоремы. По лемме, существует целое число m , такое что $0 < m < p$ и $a^2 + b^2 + c^2 = mp$. Поскольку по нашему предположению p – наименьшее положительное целое число, не представимое в виде суммы квадратов четырех рациональных чисел, существуют такие рациональные числа $\alpha, \beta, \gamma, \delta$, что $m = \alpha^2 + \beta^2 + \gamma^2 + \delta^2$. Тогда числа $m, mp, 1/m^2$ являются нормами кватернионов $\alpha + \beta i + \gamma j + \delta k, a + bi + cj, 1/m$, поэтому их произведение p равно норме кватерниона

$$(\alpha + \beta i + \gamma j + \delta k)(a + bi + cj)/m,$$

все компоненты которого, очевидно, рациональны. Итак, в противоречие с предположением число p оказалось равным сумме квадратов четырех рациональных компонент некоторого кватерниона.