Math-Net.Ru

Общероссийский математический портал

В. Б. Алексеев, А. В. Смирнов, О точной и приближенной билинейных сложностях умножения матриц размеров 4×2 и 2×2 , *Совр. пробл. матем.*, 2013, выпуск 17, 135–152

DOI: <https://doi.org/10.4213/spm47>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением
<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 52.90.164.192

26 апреля 2023 г., 11:27:22

О точной и приближенной билинейных сложностях умножения матриц размеров 4×2 и 2×2

В. Б. Алексеев ^a, А. В. Смирнов ^b

^a Московский государственный университет им. М. В. Ломоносова

^b Российский федеральный центр судебной экспертизы при МЮ РФ

Памяти Анатолия Алексеевича Карацубы

1. Введение

В 1960 г. А. А. Карацуба [1] нашел алгоритм умножения двух n -разрядных чисел с числом операций $O(n^{\log_2 3})$. Первый автор узнал об этом результате летом 1963 г., когда, окончив 9-й класс, приехал на летнюю математическую школу, организованную под Москвой академиком А. Н. Колмогоровым. Одним из преподавателей в эту школу был приглашен молодой тогда А. А. Карацуба. Его результат произвел на всех огромное впечатление и в большой степени определил научные интересы первого автора. А. А. Карацуба показал, что для простых широко распространенных арифметических и алгебраических операций существуют нестандартные асимптотически более быстрые алгоритмы. Поиск таких алгоритмов стал одной из важных задач в математике. И если для умножения чисел задачу удалось продвинуть почти до конца, то с умножением матриц ситуация оказалась намного сложнее.

Данная работа в основном посвящена задаче умножения матрицы размера 4×2 на матрицу размера 2×2 . В работе показано, что билинейная сложность этой задачи равна 14 над любым полем. С другой стороны, показано, что для этой задачи существуют приближенные билинейные алгоритмы с билинейной сложностью 13. Для задачи умножения матрицы размера 5×2 на матрицу размера 2×2 приведен приближенный билинейный алгоритм с билинейной сложностью 16 над любым полем характеристики, отличной от 2.

Пусть $\|a_{ij}\|_{m \times n}$ обозначает матрицу размера $m \times n$ над некоторым полем. Задача умножения матрицы $\|a_{ij}\|_{m \times n}$ на матрицу $\|b_{kl}\|_{n \times p}$ состоит (по определению) в вычислении mp билинейных форм вида $\sum_{j=1}^n a_{ij}b_{jl}$. Если вычислять эти формы непосредственно по указанным формулам, то необходимо сделать mnp умножений и почти столько же сложений. При $m = n = p$ всего операций больше, чем n^3 . В 1969 г. Штрассен [2] показал, что существуют более быстрые по порядку алгоритмы умножения матриц, построив алгоритм с числом операций над элементами поля $O(n^{\log_2 7})$. К 1986 г. эта оценка трудами многих авторов была понижена до $O(n^{2.38})$ [3] (см. обзор [4]), однако с тех пор эта оценка существенно не улучшена, несмотря на большой интерес к задаче.

Алгоритм Штрассена [2] для умножения матриц порядка n со сложностью $O(n^{\log_2 7})$ основан на найденном им алгоритме умножения двух квадратных матриц порядка 2 с 7 умножениями вместо 8 в обычном алгоритме. Затем этот алгоритм рекурсивно использовался для умножения матриц порядка 2^k с числом умножений 7^k . Этот прием идейно аналогичен методу А. А. Карацубы для умножения n -разрядных чисел с числом операций $O(n^{\log_2 3})$, который сейчас принято называть “разделяй и властвуй”.

Для того чтобы можно было применять рекурсию, Штрассен построил специального вида алгоритм умножения матриц порядка 2, а именно так называемый билинейный алгоритм.

Работа первого автора выполнена при поддержке Российского фонда фундаментальных исследований (грант № 12-01-91331-ННИО-а).

ОПРЕДЕЛЕНИЕ. Пусть F – некоторое кольцо, и пусть имеется 2 множества переменных $A = \{a_1, a_2, \dots, a_r\}$ и $B = \{b_1, b_2, \dots, b_s\}$. *Билинейными алгоритмами* над A и B и кольцом F называются алгоритмы, в которых сначала вычисляются произведения $(\sum_{i=1}^r \alpha_i^t a_i) \times (\sum_{j=1}^s \beta_j^t b_j)$ некоторых линейных форм (с коэффициентами из F) от первого множества переменных на некоторые линейные формы (с коэффициентами из F) от второго множества переменных, где $t = 1, 2, \dots, d$, а на втором этапе вычисляются некоторые линейные комбинации этих d произведений. При этом число умножений d называется *билинейной сложностью алгоритма*.

ОПРЕДЕЛЕНИЕ. Будем говорить, что билинейный алгоритм над кольцом F *вычисляет систему билинейных форм* $C_k = \sum_{i=1}^r \sum_{j=1}^s c_{ij}^k a_i b_j$, $k = 1, \dots, h$, где c_{ij}^k – произвольные константы из F , если каждая из этих билинейных форм оказывается вычисленной на втором этапе алгоритма. *Билинейной сложностью задачи* вычисления системы билинейных форм над кольцом F называется минимальная билинейная сложность алгоритмов над F , вычисляющих данную систему билинейных форм.

Требование билинейности при изучении сложности умножения матриц связано с тем, что при рекурсии вместо переменных подставляются матрицы, которые могут не коммутировать.

Кроме обычных (точных) билинейных алгоритмов, определенных выше, рассматривают также приближенные билинейные алгоритмы. Интерес к ним связан как с возможностью практического использования, так и с тем фактом, что, имея быстрый приближенный билинейный алгоритм для умножения матриц, можно получать асимптотически быстрый точный алгоритм для умножения матриц [5] (см. также обзор [4]).

Напомним, что лорановскими многочленами от переменной x называют выражения, отличающиеся от обычных многочленов тем, что в них разрешаются также целые отрицательные степени x .

ОПРЕДЕЛЕНИЕ. *Приближенным билинейным алгоритмом* над двумя множествами переменных A и B и кольцом F называется любой билинейный алгоритм над теми же множествами переменных, в котором в качестве коэффициентов вместо кольца F используется кольцо лорановских многочленов от x с коэффициентами из F . При этом число умножений линейных форм d также называется билинейной сложностью приближенного алгоритма.

ОПРЕДЕЛЕНИЕ. Пусть зафиксированы множества переменных A и B и кольцо F . Будем говорить, что приближенный билинейный алгоритм приближенно вычисляет систему билинейных форм C_k над F , $k = 1, \dots, h$, если на втором этапе этот билинейный алгоритм строит выражения вида $C_k + O(x)$ для всех $k = 1, \dots, h$, где $O(x)$ – многочлены, содержащие только положительные степени x (коэффициентами многочленов автоматически являются билинейные формы).

ОПРЕДЕЛЕНИЕ. *Приближенной билинейной сложностью задачи* вычисления системы билинейных форм над кольцом F называется минимальная билинейная сложность приближенных билинейных алгоритмов над F , приближенно вычисляющих данную систему билинейных форм.

Из определения видно, что если мы работаем, например, в поле действительных или комплексных чисел, то, выбирая конкретное сколь угодно близкое к 0 число x , мы можем из приближенного билинейного алгоритма получить обычный билинейный алгоритм той же сложности, сколь угодно точно вычисляющий заданную систему билинейных форм. При этом выбор $x = 0$ может оказаться запрещенным из-за наличия в коэффициентах отрицательных степеней x .

Выше отмечено, что уже более 20 лет не удается существенно улучшить результат $O(n^{2.38})$ Копперсмита и Винограда для сложности умножения матриц порядка n . В связи с этим возникает необходимость рассмотрения различных подходов к разработке быстрых алгоритмов

умножения матриц. В частности, интересно исследовать (для последующих обобщений), как устроены оптимальные алгоритмы умножения матриц малых размеров, как для них строить алгоритмы умножения с малой сложностью и какие существуют преграды в виде нижних оценок. В данной работе исследуются билинейные алгоритмы и, соответственно, билинейная сложность. Отметим, что полученные результаты позволяют лучше понять ситуацию в данной области, но не дают непосредственно улучшений в сложности умножения больших матриц.

Как сказано выше, Штрассен построил билинейный алгоритм для задачи умножения двух матриц порядка 2 с билинейной сложностью 7 (над произвольным полем). Виноград [6] показал, что эту оценку нельзя понизить, т.е. билинейная сложность задачи умножения двух матриц порядка 2 равна 7.

Несмотря на простоту постановки, задача определения билинейной сложности умножения двух матриц оказывается тяжелой даже при малых размерах матриц. Так, про билинейную сложность задачи умножения двух матриц порядка 3 над любым полем известно только, что она не больше 23 [7] и не меньше 19 [8]. Имеются также некоторые общие нижние и верхние оценки для билинейной сложности умножения двух матриц порядка k и некоторые верхние оценки для билинейной сложности умножения матриц конкретных порядков. В работе [9] построен приближенный билинейный алгоритм для умножения двух матриц порядка 3 (над произвольным полем) с билинейной сложностью 21 (см. также [4]). Одним из авторов данной работы (А.В. Смирновым) построен приближенный билинейный алгоритм для умножения двух матриц порядка 3 с билинейной сложностью 20, который будет опубликован в другой статье.

Поскольку билинейный алгоритм является частным случаем приближенного билинейного алгоритма, то для любой задачи приближенная билинейная сложность не больше, чем билинейная сложность. Известно, что приближенная билинейная сложность может быть меньше, чем билинейная сложность. Одним из простых примеров является следующий. Пусть требуется перемножить 2 матрицы порядка 2, причем один элемент в одной из них равен 0. Обычный алгоритм (“строка на столбец”) требует 6 умножений элементов. Нетрудно показать, что любой билинейный алгоритм для этой задачи имеет билинейную сложность не менее 6. В то же время, справедливо следующее утверждение.

УТВЕРЖДЕНИЕ 1. *Для умножения матрицы порядка 2 с одним нулевым элементом на матрицу порядка 2 над любым полем существует приближенный билинейный алгоритм с билинейной сложностью 5.*

Такой алгоритм построен в работе [10] (см. также [4]).

Если необходимо умножить матрицу размера $m \times 2$ на матрицу размера 2×2 , то можно разбить первую матрицу на блоки размера 2×2 ; при m нечетном последний блок будет иметь размер 1×2 . Тогда умножение матрицы размера $m \times 2$ на матрицу размера 2×2 сведется к $\lfloor m/2 \rfloor$ умножениям матриц порядка 2 и (при m нечетном) умножению матрицы размера 1×2 на матрицу размера 2×2 (и некоторым сложениям). Из результата Штрассена легко вытекает, что для этого достаточно $\lceil 7m/2 \rceil$ умножений, т.е. мы получаем билинейный алгоритм с билинейной сложностью $\lceil 7m/2 \rceil$. Хопкрофт и Керр [11] показали, что эту оценку нельзя понизить для задачи умножения таких матриц над полем из 2 элементов. Для произвольных полей и произвольных m не известно, является ли эта оценка неулучшаемой.

В работе одного из авторов [12] было показано, что при $m = 3$ любой билинейный алгоритм имеет билинейную сложность не меньше 11 над произвольным полем, т.е. билинейная сложность задачи умножения матрицы размера 3×2 на матрицу размера 2×2 равна 11 над любым полем. В то же время существует приближенный билинейный алгоритм для этой задачи с билинейной сложностью 10. Его можно построить следующим образом.

Пусть нужно умножить матрицу $A = \|a_{ij}\|_{3 \times 2}$ на матрицу $B = \|b_{kl}\|_{2 \times 2}$. Пусть матрица A_1 получается из A заменой a_{22} , a_{31} , a_{32} на 0, а матрица A_2 получается из A заменой a_{11} ,

a_{12}, a_{21} на 0. Тогда $A = A_1 + A_2$, и задача сводится к вычислению произведений A_1B и A_2B . Полностью нулевые строки в матрицах A_1 и A_2 порождают соответствующие нулевые строки в матрицах A_1B и A_2B . Поэтому задача фактически сводится к двум умножениям матриц порядка 2, причем в каждом произведении один элемент в одной из матриц равен 0 (и некоторым дополнительным сложениям). По утверждению 1 для каждого из произведений можно построить приближенный билинейный алгоритм с билинейной сложностью 5. Объединяя эти алгоритмы, получим приближенный билинейный алгоритм для задачи умножения матрицы размера 3×2 на матрицу размера 2×2 с билинейной сложностью 10 (этот алгоритм построен в [10]).

Использованная конструкция показывает также, что для задачи умножения матрицы порядка 2 с одним нулевым элементом на матрицу порядка 2 действительно не существует билинейного алгоритма (обычного) с билинейной сложностью 5. Если бы такой алгоритм существовал, то так же, как выше, из него можно было бы получить билинейный алгоритм для задачи умножения матрицы размера 3×2 на матрицу размера 2×2 с билинейной сложностью 10, что невозможно в соответствии с результатом из [12].

Отметим, что если считать переменные a_{ij} и b_{kl} коммутирующими и отказаться от билинейности, то для задачи умножения матрицы размера 3×2 на матрицу размера 2×2 можно построить алгоритм с 10 умножениями [13].

Данная работа в основном посвящена задаче умножения матрицы размера 4×2 на матрицу размера 2×2 . В этом случае верхняя оценка билинейной сложности равна 14 (дважды используется алгоритм Штрассена). В работе [13] показано, что при использовании коммутативности можно построить алгоритм с 13 умножениями. Мы покажем, что для билинейных алгоритмов оценка 14 неуплучшаема над любым полем. С другой стороны, мы покажем, что существуют приближенные билинейные алгоритмы для этой задачи с билинейной сложностью 13.

Отметим еще, что все результаты допускают обобщение с использованием двойственности. Задачу умножения матрицы размера $m \times n$ на матрицу размера $n \times p$ можно задать тройкой чисел m, n, p . Теорема о двойственности [14] (см. также [4]) утверждает, что билинейная сложность и приближенная билинейная сложность не изменяются при любых перестановках чисел m, n, p .

2. Приближенные алгоритмы

ТЕОРЕМА 1. *Для задачи умножения матрицы $A = \|a_{ij}\|$ размера 4×2 на матрицу $B = \|b_{kl}\|$ размера 2×2 над любым полем существуют приближенные билинейные алгоритмы с билинейной сложностью 13. Примером такого алгоритма является следующий приближенный билинейный алгоритм:*

$$\begin{aligned} D_1 &= (-xa_{12} + a_{22} + a_{42})(-b_{11} + b_{22}), \\ D_2 &= (xa_{12} - a_{22})(-xb_{21} + b_{22}), \\ D_3 &= (x^2a_{11} + xa_{12} - xa_{21} - a_{22} - a_{42})b_{11}, \\ D_4 &= (-xa_{41} - xa_{12} + a_{22} + a_{42})(-b_{11} + xb_{21}), \\ D_5 &= a_{42}b_{22}, \\ D_6 &= a_{21}(-xb_{11} + xb_{21} + b_{12}), \\ D_7 &= (x^2a_{31} + xa_{32} + a_{42})(b_{12} - xb_{21}), \\ D_8 &= ((x + x^2)a_{31} + a_{41})(xb_{11} + b_{12} - x^2b_{21}), \\ D_9 &= (xa_{32} + a_{42})(-b_{12} + xb_{21} + xb_{22}), \end{aligned}$$

$$\begin{aligned}
D_{10} &= (x^2 a_{11} + a_{22})(b_{12} - x^2 b_{21} + x b_{22}), \\
D_{11} &= (a_{21} + a_{22})(b_{12} + x b_{21}), \\
D_{12} &= (x a_{32} - x a_{31} + a_{42} - a_{41}) b_{12}, \\
D_{13} &= a_{22} b_{12}, \\
C_{11} &= x^{-2}(-D_1 - D_2 + D_3 + D_5 - D_6 + D_{11} - D_{13}), \\
C_{21} &= x^{-1}(-D_6 + D_{11} - D_{13}), \\
C_{31} &= x^{-2}(D_1 + D_2 - D_4 - D_5 - D_7 + D_8 + D_{12}), \\
C_{41} &= x^{-1}(-D_1 - D_2 + D_4 + D_5), \\
C_{12} &= x^{-2}(x D_2 + D_{10} - D_{13}), \\
C_{22} &= D_1 - D_3 - D_5 + D_6, \\
C_{32} &= x^{-2}(-x D_5 + D_7 + D_9), \\
C_{42} &= D_5 + D_8.
\end{aligned}$$

Для доказательства этой теоремы нетрудно провести указанные вычисления и проверить, что $C_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + O(x)$ для всех $i = 1, \dots, 4$ и $j = 1, 2$, что соответствует определению приближенного алгоритма для умножения матриц.

Ниже мы приведем еще одну конструкцию приближенного билинейного алгоритма с билинейной сложностью 13 для задачи умножения матрицы размера 4×2 на матрицу размера 2×2 , основанную на следующей теореме.

ТЕОРЕМА 2. *Для задачи умножения матрицы $A = \|a_{ij}\|$ размера 3×2 с одним нулевым элементом (пусть $a_{32} = 0$) на матрицу $B = \|b_{kl}\|$ размера 2×2 над любым полем характеристики, отличной от 2, существуют приближенные билинейные алгоритмы с билинейной сложностью 8. В то же время билинейная сложность точного умножения в этой задаче над любым полем равна 9.*

ДОКАЗАТЕЛЬСТВО. Рассмотрим следующий алгоритм:

$$\begin{aligned}
D_1 &= (-\tfrac{1}{2} x a_{11} - \tfrac{1}{2} a_{21} + x^{-1} a_{22})(-b_{11} + b_{12} + x b_{21}), \\
D_2 &= (x^{-1} a_{22} + \tfrac{1}{2} x^{-1} a_{31})(b_{11} - b_{12}), \\
D_3 &= (x a_{11} + a_{12} - \tfrac{1}{2} a_{21} - x^{-1} a_{22})(b_{11} + b_{12} + x b_{22}), \\
D_4 &= (\tfrac{1}{2} x a_{11} - a_{12} - \tfrac{1}{2} a_{21} + x^{-1} a_{22})(b_{11} + b_{12} - x b_{21}), \\
D_5 &= (-x a_{11} + a_{21} - x^{-1} a_{31}) b_{11}, \\
D_6 &= (\tfrac{1}{2} a_{21} + x^{-1} a_{22})(-b_{11} + b_{12} + x b_{22}), \\
D_7 &= (-a_{12} + x^{-1} a_{22} + \tfrac{1}{2} x^{-1} a_{31})(b_{11} + b_{12}), \\
D_8 &= (a_{21} + x^{-1} a_{31}) b_{12}, \\
C_{11} &= x^{-1}(D_1 + D_2 + D_4 - D_7 + D_8), \\
C_{12} &= x^{-1}(D_2 + D_3 + D_5 + D_6 + D_7), \\
C_{21} &= D_1 + D_2 + \tfrac{1}{2}(D_5 + D_8), \\
C_{22} &= D_2 + D_6 + \tfrac{1}{2}(D_5 + D_8), \\
C_{31} &= -x D_5, \\
C_{32} &= x D_8.
\end{aligned}$$

Проводя указанные вычисления, нетрудно проверить, что выполняются равенства $C_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + O(x)$ для $i = 1, 2$, $j = 1, 2$ и $C_{31} = a_{31}b_{11} + O(x)$, $C_{32} = a_{31}b_{12} + O(x)$, что соответствует определению приближенного алгоритма для умножения указанных матриц. Отсюда следует первая часть утверждения теоремы.

Если разбить матрицу A на матрицу порядка 2 и матрицу с одним ненулевым элементом и использовать алгоритм Штрассена, то потребуется 9 умножений. С другой стороны, если бы существовал билинейный алгоритм для рассматриваемой задачи с билинейной сложностью 8, то, добавляя к нему 2 произведения $a_{32}b_{21}$ и $a_{32}b_{22}$, мы получили бы билинейный алгоритм для задачи умножения матрицы размера 3×2 на матрицу размера 2×2 с билинейной сложностью 10, что невозможно в соответствии с результатом из [12]. Теорема 2 полностью доказана.

Пусть теперь надо умножить матрицу $A = \|a_{ij}\|$ размера 4×2 на матрицу $B = \|b_{kl}\|$ размера 2×2 . Пусть матрица A_1 получается из A заменой a_{32} , a_{41} , a_{42} на 0, а матрица A_2 получается из A , наоборот, заменой на 0 всех остальных элементов. Тогда $A = A_1 + A_2$, и задача сводится к вычислению произведений A_1B и A_2B . Полностью нулевые строки в матрицах A_1 и A_2 порождают соответствующие нулевые строки в матрицах A_1B и A_2B . Поэтому задача фактически сводится к умножению матрицы размера 3×2 с одним нулевым элементом на матрицу порядка 2 и умножению матрицы порядка 2 с одним нулевым элементом на матрицу порядка 2 (и некоторым дополнительным сложениям). По теореме 2 и утверждению 1 для этих задач можно построить приближенные билинейные алгоритмы с билинейными сложностями соответственно 8 и 5. Объединяя эти алгоритмы, получим еще один приближенный билинейный алгоритм для задачи умножения матрицы размера 4×2 на матрицу размера 2×2 с билинейной сложностью 13. Заметим, что этот алгоритм, в отличие от алгоритма из теоремы 1, можно реализовать в любых полях характеристики, отличной от 2.

Используя теорему 2, легко получить следующий результат.

ТЕОРЕМА 3. *Для задачи умножения матрицы A размера 5×2 на матрицу B размера 2×2 над любым полем характеристики, отличной от 2, существует приближенный билинейный алгоритм с билинейной сложностью 16.*

ДОКАЗАТЕЛЬСТВО. Разрежем (как выше) матрицу A размера 5×2 на 2 матрицы A_1 и A_2 . К первой отнесем элементы первых двух строк и элемент a_{31} (остальные обнулим), ко второй отнесем элемент a_{32} и элементы 4-й и 5-й строк. Тогда $A = A_1 + A_2$, и задача сводится к вычислению произведений A_1B и A_2B . Получаем, что фактически задача сводится к двум умножениям матрицы размера 3×2 с одним нулевым элементом на матрицу порядка 2. Применяя дважды алгоритм из теоремы 2, получим приближенный билинейный алгоритм для задачи умножения матрицы размера 5×2 на матрицу размера 2×2 с билинейной сложностью 16. Теорема доказана.

Отметим, что имеющаяся верхняя оценка для билинейной сложности точного умножения в этой задаче равна 18.

СЛЕДСТВИЕ. *Для задачи умножения матрицы A размера $t \times 2$ на матрицу B размера 2×2 над любым полем характеристики, отличной от 2, существует приближенный билинейный алгоритм с билинейной сложностью $\lceil 16t/5 \rceil$.*

ДОКАЗАТЕЛЬСТВО. Для $t = 5k, 5k + 3, 5k + 4$ разрежем матрицу A на k матриц размера 5×2 и (для $t = 5k + 3, 5k + 4$) 1 матрицу размера 3×2 или 4×2 соответственно. Для $t = 5k + 1, 5k + 2$ разрежем матрицу A на $k - 1$ матриц размера 5×2 и дополнительно либо на 2 матрицы размера 3×2 , либо 1 матрицу размера 3×2 и 1 матрицу размера 4×2 соответственно. Для умножения матриц размера 5×2 на матрицу B размера 2×2 используем приближенный билинейный алгоритм с билинейной сложностью 16 из теоремы 3. Для умножения матриц размеров 4×2 и 3×2 на матрицу B размера 2×2 используем приближенные

билинейные алгоритмы с билинейной сложностью 13 (из теоремы 1) и 10 (см. раздел 1) соответственно. Легко проверить, что суммарная билинейная сложность построенных алгоритмов во всех случаях выражается формулой $\lceil 16m/5 \rceil$. Следствие доказано.

Отметим, что для полей характеристики 2 в этом следствии можно аналогичным образом, используя теорему 1, получить приближенный билинейный алгоритм с билинейной сложностью $\lceil 13m/4 \rceil$.

В заключение раздела укажем, что приведенные в доказательствах теорем 3 и 2 приближенные алгоритмы получены на компьютере с помощью эвристического вычислительного метода. Детали этого метода планируется представить в отдельной статье.

3. Нижняя оценка

В этом разделе мы покажем, что любой билинейный алгоритм (точный) для задачи умножения матрицы размера 4×2 на матрицу размера 2×2 над любым полем имеет билинейную сложность не менее 14 (еще раз отметим, что в [11] такой результат получен только для поля из 2 элементов). Мы покажем, что для этой задачи не существует билинейного алгоритма с билинейной сложностью 13. Идея доказательства состоит в том, чтобы из произвольного решения сложности 13 получить решение более простого вида, а уже для этого более простого вида доказать невозможность. Опишем сначала способы представления и преобразования решений.

Пусть заданы 2 матрицы $A = \|a_{ij}\|_{m \times n}$ и $B = \|b_{kl}\|_{n \times p}$; фактически заданы 2 набора переменных: a_{ij} (их mn) и b_{kl} (их np). Любое выражение вида $\sum_{i=1}^m \sum_{j=1}^n \sum_{k=1}^n \sum_{l=1}^p \alpha_{(ij)(kl)} a_{ij} b_{kl}$ можно рассматривать как билинейную форму от двух множеств переменных $\{a_{ij}\}$ и $\{b_{kl}\}$. Матрица этой билинейной формы задается коэффициентами $\alpha_{(ij)(kl)}$. Для того чтобы точно задать эту матрицу, необходимо упорядочить все переменные a_{ij} и все переменные b_{kl} . Мы упорядочим переменные a_{ij} по строкам матрицы A : $a_{11}, a_{12}, \dots, a_{1n}, a_{21}, a_{22}, \dots, a_{2n}, \dots, a_{m1}, a_{m2}, \dots, a_{mn}$, и сопоставим их в этом порядке строкам матрицы коэффициентов билинейной формы. Переменные b_{kl} упорядочим по столбцам матрицы B : $b_{11}, b_{21}, \dots, b_{n1}, b_{12}, b_{22}, \dots, b_{n2}, \dots, b_{1p}, b_{2p}, \dots, b_{np}$, и сопоставим их в этом порядке столбцам матрицы коэффициентов билинейной формы. Таким образом, переменной a_{ij} соответствует строка с номером $(i-1)n+j$, а переменной b_{kl} соответствует столбец с номером $k+(l-1)n$. Тогда в матрице билинейной формы на пересечении строки, соответствующей a_{ij} , и столбца, соответствующего b_{kl} , будет (как обычно) ставиться коэффициент $\alpha_{(ij)(kl)}$ в билинейной форме при $a_{ij}b_{kl}$.

Умножение матрицы $A = \|a_{ij}\|_{m \times n}$ на матрицу $B = \|b_{kl}\|_{n \times p}$ состоит в вычислении mp билинейных форм вида $\sum_{j=1}^n a_{ij}b_{jl}$. Посмотрим, как выглядят матрицы этих билинейных форм при указанной упорядоченности a_{ij} и b_{kl} . Разобьем матрицу билинейной формы размера $mn \times np$ на блоки размера $n \times n$. Тогда ее можно рассматривать как блочную матрицу размера $m \times p$ с блоками размера $n \times n$. При этом в матрице билинейной формы $\sum_{j=1}^n a_{ij}b_{jl}$ при фиксированных i, l единицы стоят только в клетках с координатами $((i-1)n+j, j+(l-1)n)$, $j=1, 2, \dots, n$, т.е. в il -м блоке на диагонали стоят 1, а все остальные элементы матрицы равны 0. Другими словами, il -й блок является единичной матрицей E_n порядка n , а все остальные блоки – нулевые матрицы. Обозначим такую матрицу размера $mn \times np$ при фиксированных m, n, p через E_{il} ($1 \leq i \leq m, 1 \leq l \leq p$).

Нам нужно, чтобы все mp билинейных форм $\sum_{j=1}^n a_{ij}b_{jl}$ выражались как линейные комбинации некоторых разложимых билинейных форм вида $(\sum_{i=1}^m \sum_{j=1}^n \alpha_{ij}a_{ij})(\sum_{k=1}^n \sum_{l=1}^p \beta_{kl}b_{kl})$ с коэффициентами из заданного поля. Это равносильно тому, что матрицы этих mp билинейных форм выражаются как линейные комбинации матриц разложимых билинейных форм. В связи с этим возникает вопрос: как выглядят матрицы разложимых билинейных форм?

ЛЕММА 1. Ненулевую билинейную форму $C = \sum_{i=1}^r \sum_{j=1}^s c_{ij} x_i y_j$ над произвольным полем F можно представить в виде произведения двух линейных форм соответственно от первого и второго множества переменных $C = (\sum_{i=1}^r \alpha_i x_i)(\sum_{j=1}^s \beta_j y_j)$, тогда и только тогда, когда матрица c_{ij} этой билинейной формы имеет ранг 1.

ДОКАЗАТЕЛЬСТВО. Если билинейная форма C разложима, т.е.

$$C = \sum_{i=1}^r \sum_{j=1}^s c_{ij} x_i y_j = \left(\sum_{i=1}^r \alpha_i x_i \right) \left(\sum_{j=1}^s \beta_j y_j \right),$$

то $(c_{i1}, c_{i2}, \dots, c_{is}) = \alpha_i(\beta_1, \beta_2, \dots, \beta_s)$, поэтому все строки матрицы $\|c_{ij}\|$ билинейной формы C пропорциональны и матрица $\|c_{ij}\|$ имеет ранг 1. Обратно, если матрица $\|c_{ij}\|$ имеет ранг 1, то существует такой вектор $(\beta_1, \beta_2, \dots, \beta_s)$ и такие константы α_i , что $(c_{i1}, c_{i2}, \dots, c_{is}) = \alpha_i(\beta_1, \beta_2, \dots, \beta_s)$. Тогда

$$C = \sum_{i=1}^r \sum_{j=1}^s c_{ij} x_i y_j = \left(\sum_{i=1}^r \alpha_i x_i \right) \left(\sum_{j=1}^s \beta_j y_j \right).$$

Лемма доказана.

Таким образом, задача определения билинейной сложности умножения матрицы $A = \|a_{ij}\|_{m \times n}$ на матрицу $B = \|b_{kl}\|_{n \times p}$ (над фиксированным полем F) может быть сведена к следующей задаче.

ЗАДАЧА E . Найти d матриц $D_1, D_2, \dots, D_d$ ранга 1 (над фиксированным полем F), линейными комбинациями которых являются все указанные выше матрицы E_{il} (при заданных параметрах m, n, p). Это число d будем называть сложностью найденного решения.

В результате получаем, что билинейная сложность умножения матрицы $A = \|a_{ij}\|_{m \times n}$ на матрицу $B = \|b_{kl}\|_{n \times p}$ равна наименьшей сложности решения задачи E (при заданных параметрах m, n, p).

Для удобства введем следующее обозначение. В дальнейшем выражение $|P_1|P_2| \dots |P_m|$ будет обозначать матрицу, разрезанную на m вертикальных блоков одинаковой ширины n . Для произвольной матрицы P порядка mn и для любых $1 \leq i \leq m, 1 \leq l \leq p$ определим матрицы P_{il} размера $mn \times np$ следующим образом. Пусть $P = |P_1|P_2| \dots |P_m|$ – разбиение матрицы P на вертикальные блоки ширины n . Тогда положим $P_{il} = |0| \dots |0|P_i|0| \dots |0|$, где все блоки имеют ширину n , l -й блок матрицы P_{il} равен P_i , т.е. i -му блоку матрицы P , а остальные блоки матрицы P_{il} являются нулевыми матрицами. Важно заметить, что указанные выше матрицы E_{il} являются частным случаем матриц P_{il} при $P = E$, где E – единичная матрица порядка mn . Рассмотрим следующую задачу.

ЗАДАЧА P . Найти наименьшее число d матриц ранга 1, линейными комбинациями которых являются все указанные выше матрицы P_{il} .

ЗАМЕЧАНИЕ 1. Если $D_1, D_2, \dots, D_d$ – решение задачи P и все матрицы умножить на произвольные ненулевые константы, то, очевидно, снова получится решение задачи P .

ЗАМЕЧАНИЕ 2. Указанная выше задача E является частным случаем задачи P при $P = E$.

ЛЕММА 2. Над любым полем F наименьшая сложность решения задачи P при фиксированных m, n, p одинакова для всех невырожденных матриц P и равна билинейной сложности умножения матрицы $A = \|a_{ij}\|_{m \times n}$ на матрицу $B = \|b_{kl}\|_{n \times p}$. Если $D_1, D_2, \dots, D_d$ – решение задачи P и C – невырожденная матрица, то $CD_1, CD_2, \dots, CD_d$ – решение задачи CP .

ДОКАЗАТЕЛЬСТВО. Пусть C – невырожденная матрица порядка mn . Имеем

$$CP = C \cdot (|P_1| |P_2| \cdots |P_m|) = |CP_1| |CP_2| \cdots |CP_m|.$$

Отсюда $(CP)_i = C \cdot P_i$ для всех i и $(CP)_{il} = C \cdot P_{il}$ при всех i и l . Поэтому если матрица P_{il} является линейной комбинацией матриц $D_1, D_2, \dots, D_d$ ранга 1, то $(CP)_{il} = C \cdot P_{il}$ является той же линейной комбинацией матриц $CD_1, CD_2, \dots, CD_d$ ранга 1. Это значит, что если матрицы $D_1, D_2, \dots, D_d$ ранга 1 дают решение задачи P , то матрицы $CD_1, CD_2, \dots, CD_d$ ранга 1 дают решение задачи CP . Последнее утверждение леммы доказано. Если P' и P'' – две невырожденные матрицы порядка mn , то, рассматривая матрицы $C_1 = P''(P')^{-1}$ и $C_2 = P'(P'')^{-1}$, получаем, что из решения задачи P' можно получить решение задачи P'' той же сложности, и наоборот. Отсюда следует первая половина первого утверждения леммы. Поскольку матрица E невырожденная и наименьшая сложность решения задачи E равна билинейной сложности умножения матрицы $A = \|a_{ij}\|_{m \times n}$ на матрицу $B = \|b_{kl}\|_{n \times p}$, то лемма полностью доказана.

Лемма 2 позволяет из произвольного решения задачи P получать решения более простого вида (возможно для другой матрицы). Еще 3 преобразования решений отражены в леммах 3–5. В дальнейшем мы будем считать фиксированным некоторое поле F , над которым рассматриваются все матрицы, и не будем отмечать его в формулировках утверждений.

ЛЕММА 3. Пусть $D_1, D_2, \dots, D_d$ – решение задачи P для некоторой невырожденной матрицы P порядка mn . Пусть $D_t = |D_t^1| |D_t^2| \cdots |D_t^p|$ – разбиение каждой матрицы D_t на блоки ширины p . Пусть H – невырожденная матрица порядка n . Тогда набор матриц первого ранга $D_t' = |D_t^1 H| |D_t^2 H| \cdots |D_t^p H|$ является решением задачи вида P для невырожденной матрицы $|P_1 H| |P_2 H| \cdots |P_m H|$. (Заметим, что рассматриваемое преобразование означает просто представление всех векторов длины n в новом базисе.)

ДОКАЗАТЕЛЬСТВО. Докажем сначала корректность формулировки. Заметим, что всегда столбцы матрицы DH являются линейными комбинациями столбцов матрицы D . Так как по определению решения задачи P все матрицы D_t имеют ранг 1, то при каждом t все столбцы всех матриц $D_t^1, D_t^2, \dots, D_t^p$ пропорциональны, а значит, пропорциональны и все столбцы всех матриц $D_t^1 H, D_t^2 H, \dots, D_t^p H$. Следовательно, матрица D_t' имеет ранг 1 при каждом t . Далее имеем $|P_1 H| |P_2 H| \cdots |P_m H| = |P_1| |P_2| \cdots |P_m| \cdot K$, где K – матрица размера $mn \times mn$, разбитая на блоки размера $n \times n$, в которой диагональные блоки равны H , а остальные блоки равны 0. Так как по условию матрица H невырожденная, то и матрица K невырожденная, а так как матрица P невырожденная, то и матрица $|P_1 H| |P_2 H| \cdots |P_m H| = PK$ невырожденная. Корректность формулировки доказана.

Если некоторая линейная комбинация матриц $D_1, D_2, \dots, D_d$ дает матрицу $P_{il} = |0| \cdots |0| P_i |0| \cdots |0|$, то такая же линейная комбинация матриц $D_1', D_2', \dots, D_d'$ даст матрицу $|0| \cdots |0| P_i H |0| \cdots |0|$, где блок $P_i H$ также стоит на l -м месте. Отсюда следует утверждение леммы.

ЛЕММА 4. Пусть $D_1, D_2, \dots, D_d$ – решение задачи P для некоторой невырожденной матрицы $P = |P_1| |P_2| \cdots |P_m|$ порядка mn (блоки ширины n). Пусть каждый блок P_i заменяется на блок P'_i , являющийся линейной комбинацией блоков P_i . Тогда тот же набор матриц $D_1, D_2, \dots, D_d$ является решением задачи P' для матрицы $P' = |P'_1| |P'_2| \cdots |P'_m|$.

ДОКАЗАТЕЛЬСТВО. Так как матрица P'_i является линейной комбинацией матриц P_i , то при любом l матрица P'_{il} является линейной комбинацией матриц P_{il} . Так как $D_1, D_2, \dots, D_d$ – решение задачи P , то каждая матрица P_{il} является линейной комбинацией матриц D_t , а значит, и каждая матрица P'_{il} является линейной комбинацией матриц D_t . Лемма доказана.

ЗАМЕЧАНИЕ 3. При применении леммы 4 нужно будет следить за тем, чтобы матрица P' оставалась невырожденной.

ЛЕММА 5. Пусть $D_1, D_2, \dots, D_d$ – решение задачи P для некоторой невырожденной матрицы P . Пусть $D_t = |D_t^1| |D_t^2| \cdots |D_t^p|$ – разбиение каждой матрицы D_t на блоки шириной p .

Пусть одновременно во всех матрицах D_t делается одно и то же невырожденное преобразование с блоками: $\bar{D}_t^j = \sum_{i=1}^p l_{ij} D_t^i$, $j = 1, \dots, p$. Пусть матрица этого преобразования $L = \|l_{ij}\|$ невырожденна. Тогда полученные матрицы $\bar{D}_t$ также являются решением задачи P для той же матрицы P . (Матрицы $\bar{D}_t$ очевидно, остаются матрицами первого ранга.)

ДОКАЗАТЕЛЬСТВО. Описанное преобразование означает, что все матрицы D_t умножаются справа на одну и ту же блочную матрицу L' размера $pr \times pr$ с блоками размера $n \times n$, в которой ij -й блок равен $l_{ij} E_n$, где E_n – единичная матрица порядка n , l_{ij} – некоторые числа. По определению решения задачи P каждая матрица P_{ij} является линейной комбинацией матриц D_t : $P_{ij} = \sum_{t=1}^d c_{ij}^t D_t$. Умножая это равенство справа на матрицу L' , получим $P'_{ij} = \sum_{t=1}^d c_{ij}^t \bar{D}_t$, где $P'_{ij} = |l_{j1} P_i | l_{j2} P_i | \dots | l_{jp} P_i |$. Поскольку матрица $L = \|l_{ij}\|$ невырожденна, то для нее существует обратная матрица $R = \|r_{ij}\|$ такая, что $\sum_{k=1}^p r_{jk} l_{ks}$ равна 0 при $j \neq s$ и равна 1 при $j = s$. Тогда

$$\sum_{k=1}^p r_{jk} P'_{ik} = \left| \left(\sum_{k=1}^p r_{jk} l_{k1} \right) P_i \right| \dots \left| \left(\sum_{k=1}^p r_{jk} l_{kp} \right) P_i \right| = P_{ij}.$$

Таким образом, каждая матрица P_{ij} является линейной комбинацией матриц P'_{ij} , а следовательно, и матриц $\bar{D}_t$. Лемма доказана.

ТЕОРЕМА 4. Любой билинейный алгоритм (точный) для умножения матрицы размера 4×2 на матрицу размера 2×2 над произвольным полем имеет билинейную сложность не менее 14.

ДОКАЗАТЕЛЬСТВО. Мы докажем от противного, что для указанной задачи не существует билинейного алгоритма с 13 умножениями. По лемме 2 это утверждение равносильно отсутствию решения у следующей задачи.

Пусть P – произвольная невырожденная матрица размера 8×8 , разбитая на подматрицы размера 8×2 : $P = |P_1|P_2|P_3|P_4|$. Пусть P_{i1} – матрица размера 8×4 вида $P_{i1} = |P_i|0|$, а P_{i2} – матрица размера 8×4 вида $P_{i2} = |0|P_i|$ (все блоки ширины 2). Спрашивается, существуют ли 13 матриц первого ранга $D_1, \dots, D_{13}$ размера 8×4 таких, что все 8 матриц P_{i1} и P_{i2} , $j = i, \dots, 4$, являются линейными комбинациями матриц $D_1, \dots, D_{13}$.

Надо доказать, что эта задача не имеет решения ни для какой невырожденной матрицы P . Идея доказательства состоит в том, чтобы из произвольного решения этой задачи получить решение более простого вида, а уже для этого более простого вида доказать невозможность.

Допустим, что для некоторой невырожденной матрицы P размера 8×8 такое решение $D_1, \dots, D_{13}$ из матриц ранга 1 существует. Это означает, что существуют такие коэффициенты $\alpha_{it}^1, \alpha_{it}^2$ (из рассматриваемого поля), что для $i = 1, \dots, 4$ выполняются равенства

$$P_{i1} = \sum_{t=1}^{13} \alpha_{it}^1 D_t, \quad P_{i2} = \sum_{t=1}^{13} \alpha_{it}^2 D_t.$$

Разобьем все матрицы D_t на 2 блока размера 8×2 : $D_t = |D_t^1|D_t^2|$. Тогда последние равенства равносильны равенствам

$$P_i = \sum_{t=1}^{13} \alpha_{it}^1 D_t^1, \quad \sum_{t=1}^{13} \alpha_{it}^1 D_t^2 = 0, \quad \sum_{t=1}^{13} \alpha_{it}^2 D_t^1 = 0, \quad P_i = \sum_{t=1}^{13} \alpha_{it}^2 D_t^2. \quad (22)$$

Так как матрица P невырожденная, то ее столбцы линейно независимы, поэтому, в частности, линейно независимы подматрицы P_1, P_2, P_3, P_4 . Тогда из последнего равенства в (22) следует, что 4 вектора коэффициентов $(\alpha_{i1}^2, \alpha_{i2}^2, \dots, \alpha_{i,13}^2)$, $i = 1, \dots, 4$, линейно независимы. Но тогда

из третьего равенства в (22) следует, что среди матриц $D_1^1, \dots, D_{13}^1$ не более 9 линейно независимых. С другой стороны, в матрицах P_1, P_2, P_3, P_4 имеется 8 линейно независимых столбцов, которые согласно первому равенству в (22) должны быть линейными комбинациями столбцов из $D_1^1, \dots, D_{13}^1$. Поскольку матрицы $D_1^1, \dots, D_{13}^1$ имеют ранг 1, то это означает, что среди $D_1^1, \dots, D_{13}^1$ есть 8 ненулевых матриц первого ранга, построенных на 8 линейно независимых столбцах. Тогда эти 8 матриц линейно независимы. Получаем, что возможны только 2 случая: максимальное число линейно независимых матриц среди матриц $D_1^1, \dots, D_{13}^1$ равно 8 или 9. Рассмотрим эти случаи отдельно.

Случай 1. Среди матриц $D_1^1, \dots, D_{13}^1$ имеется 8 линейно независимых матриц (пусть это матрицы $D_1^1, \dots, D_8^1$), а матрицы $D_9^1, \dots, D_{13}^1$ являются линейными комбинациями матриц $D_1^1, \dots, D_8^1$. Это значит, что существуют такие коэффициенты b_t^p, k_i^p (из рассматриваемого поля), что выполняются равенства (см. (22))

$$D_t^1 = \sum_{p=1}^8 b_t^p D_p^1, \quad t = 9, \dots, 13, \quad P_i = \sum_{t=1}^8 k_i^t D_t^1, \quad i = 1, \dots, 4. \quad (23)$$

Согласно (23) из столбцов матриц $D_1^1, \dots, D_8^1$ линейными комбинациями должны получаться все 8 линейно независимых столбцов матрицы P . Так как все матрицы $D_1^1, \dots, D_8^1$ имеют ранг 1, то все матрицы $D_1^1, \dots, D_8^1$ должны быть ненулевыми и построенными на 8 линейно независимых столбцах. Каждую матрицу D_t первого ранга можно представить в виде произведения $D_t = f_t \cdot g_t$, где f_t – некоторый вектор-столбец, а g_t – некоторая вектор-строка. Тогда $CD_t = (Cf_t) \cdot g_t$. Поскольку вектор-столбцы $f_1, f_2, \dots, f_8$ линейно независимы, то можно найти такую невырожденную матрицу C , что $Cf_t = e_t$ при всех $t = 1, \dots, 8$, где e_t – вектор-столбец с одной 1 на t -м месте и нулями на остальных местах. Тогда в матрице CD_t единственной ненулевой строкой будет строка с номером t . По лемме 2 матрицы CD_t , $t = 1, \dots, 8$, будут опять давать решение задачи P , но с заменой матрицы P на матрицу CP . При этом равенства (23) останутся справедливыми для матрицы CP и матриц CD_t . Чтобы избежать громоздких обозначений, можем считать, что решение $D_1, \dots, D_{13}$ уже для исходной матрицы P имеет такой вид, т.е. в матрице D_t , $t = 1, \dots, 8$, единственной ненулевой строкой является строка с номером t .

Пусть

$$N_t = D_t - \sum_{p=1}^8 b_t^p D_p, \quad t = 9, \dots, 13. \quad (24)$$

Тогда из (23) получаем, что N_t имеет вид $N_t = |0|N_t^2|$, где N_t^2 – подматрица размера 8×2 . Так как среди матриц $D_1^1, \dots, D_{13}^1$ ровно 8 линейно независимых матриц, то пространство линейных комбинаций $\sum_{t=1}^{13} \mu_t D_t$ таких, что $\sum_{t=1}^{13} \mu_t D_t^1 = 0$, имеет размерность $13 - 8 = 5$. Но 5 линейных комбинаций $D_t - \sum_{p=1}^8 b_t^p D_p$, $t = 9, \dots, 13$, выражающих N_t , входят в это пространство и линейно независимы, поскольку каждое слагаемое $D_9, \dots, D_{13}$ входит ровно в одну из этих линейных комбинаций. Следовательно, каждая матрица, которая является линейной комбинацией матриц $D_1, \dots, D_{13}$ и имеет вид $|0|H|$, является линейной комбинацией матриц $N_9, \dots, N_{13}$. В частности, существуют такие константы e_i^t (из рассматриваемого поля), что выполняются равенства

$$P_{i2} = |0|P_i| = \sum_{t=9}^{13} e_i^t N_t, \quad P_i = \sum_{t=9}^{13} e_i^t N_t^2, \quad i = 1, \dots, 4. \quad (25)$$

ОПРЕДЕЛЕНИЕ. Будем говорить, что матрица $N_t = |0|F_t|$ имеет *тип* F , если для нее в (24) все коэффициенты $b_t^p = 0$. Будем говорить, что матрица $N_t = |0|G_t|$ имеет *тип* G , если для нее в (24) существует коэффициент $b_t^p \neq 0$.

Напомним, что каждая из матриц $D_1, \dots, D_8$ отлична от 0 только в одной строке (каждая в своей).

ЛЕММА 6. Если матрица $N_t = |0|F_t|$ имеет тип F , то ранг $\text{rk}(N_t) = 1$. Если матрица $N_t = |0|D_t|$ имеет тип G , то она может быть отлична от 0 только в таких строках, в которых соответствующие строки из $D_1^1, \dots, D_8^1$ пропорциональны.

ДОКАЗАТЕЛЬСТВО. В первом случае $N_t = D_t$ и ранг $\text{rk}(N_t) = \text{rk}(G_t) = 1$. Во втором случае из (23) имеем $D_t^1 = \sum_{p=1}^8 b_t^p D_p^1 \neq 0$ и D_t^1 имеет ранг 1. Это значит, что все строки в ней пропорциональны. Значит, в сумме $\sum_{p=1}^8 b_t^p D_p^1$ с ненулевыми коэффициентами b_t^p могут участвовать только те матрицы D_p^1 , у которых единственные ненулевые строки пропорциональны. Но тогда матрица D_t^1 равна 0 во всех остальных строках, а поскольку $D_t^1 \neq 0$ и матрица D_t имеет ранг 1, то и вся матрица D_t равна 0 во всех остальных строках. При этом и матрица $N_t = D_t - \sum_{p=1}^8 b_t^p D_p$ равна 0 во всех остальных строках. Лемма доказана.

Так как в P_1, P_2, P_3, P_4 имеется 8 линейно независимых столбцов, то из (25) и леммы 6 вытекает следующее утверждение.

ЛЕММА 7. В матрицах $N_9^2, \dots, N_{13}^2$ имеется 8 линейно независимых столбцов. В частности, среди $N_9, \dots, N_{13}$ должно быть не менее 3 матриц типа G .

Пусть для $j = 1, 2, \dots, 8$ матрица D_j равна $|p_j|q_j|$ в j -й строке, где p_j, q_j – векторы размерности 2 (в остальных строках она равна 0). Для краткости будем изображать набор матриц $D_1, \dots, D_8$ одной матрицей в виде

$$\begin{array}{l|l|l} D_1: & p_1 & q_1 \\ D_2: & p_2 & q_2 \\ \dots & \dots & \dots \\ D_8: & p_8 & q_8 \end{array}$$

Рассмотрим несколько случаев.

1) Пусть среди $p_1, \dots, p_8$ есть 5 пропорциональных. Пусть это векторы $p_1, \dots, p_5$. Тогда, переходя к новому базису на векторах длины 2, в соответствии с леммой 3 можно получить решение сложности 13 для задачи P для некоторой невырожденной матрицы P , в котором все эти строки будут пропорциональны вектору $(1, 0)$. Будем считать, что наше решение имеет именно такой вид. Тогда, учитывая вторую часть равенств (23), получаем, что во всех матрицах P_1, P_2, P_3, P_4 первые 5 элементов вторых столбцов равны 0. Но тогда столбцы с номерами 2, 4, 6, 8 в матрице P имеют ранг не более 3 и матрица P вырожденная – противоречие.

2) Пусть среди $p_1, \dots, p_8$ есть 4 пропорциональных вектора, которые не пропорциональны остальным векторам из $p_1, \dots, p_8$. Можно считать, что это векторы $p_1, \dots, p_4$. Тогда с учетом леммы 3 и замечания 1 перед леммой 2 можно считать, что все они равны $(1, 0)$. Учитывая вторую часть равенств (23), получаем

$$\begin{array}{l|l|l} D_1: & 10 & ** \\ D_2: & 10 & ** \\ D_3: & 10 & ** \\ D_4: & 10 & ** \\ D_5: & *1 & ** \\ D_6: & *1 & ** \\ D_7: & *1 & ** \\ D_8: & *1 & ** \end{array}, \quad P: \begin{array}{l|l|l|l} *0 & *0 & *0 & *0 \\ *0 & *0 & *0 & *0 \\ *0 & *0 & *0 & *0 \\ *0 & *0 & *0 & *0 \\ ** & ** & ** & ** \\ ** & ** & ** & ** \\ ** & ** & ** & ** \\ ** & ** & ** & ** \end{array}.$$

Обозначим подматрицу матрицы P_i размера 4×2 , расположенную в строках 1–4, через $\overline{P}_i$, а подматрицу, расположенную в строках 5–8, через $\overline{\overline{P}}_i$. Аналогично определим подматрицы

$\overline{N}_t^2, \overline{\overline{N}}_t^2, \overline{P}$ и $\overline{\overline{P}}$. Тогда из равенства (25) получаем

$$\overline{P}_i = \sum_{t=9}^{13} e_i^t \overline{N}_t^2, \quad \overline{\overline{P}}_i = \sum_{t=9}^{13} e_i^t \overline{\overline{N}}_t^2.$$

Матрица $\overline{P}$ имеет ранг 4, поскольку матрица P невырожденная, поэтому первые столбцы в матрицах $\overline{P}_i$ линейно независимы и матрицы $\overline{P}_1, \overline{P}_2, \overline{P}_3, \overline{P}_4$ линейно независимы. Следовательно, среди матриц $\overline{N}_t^2$ не менее 4 ненулевых. Столбцы с номерами 2, 4, 6, 8 в матрице P линейно независимы. Поэтому вторые столбцы в матрицах $\overline{P}_i$ линейно независимы и матрицы $\overline{\overline{P}}_1, \overline{\overline{P}}_2, \overline{\overline{P}}_3, \overline{\overline{P}}_4$ линейно независимы. Следовательно, среди матриц $\overline{\overline{N}}_t^2$ не менее 4 ненулевых. Таким образом, среди 10 матриц $\overline{N}_t^2$ и $\overline{\overline{N}}_t^2$ не менее 8 ненулевых. По лемме 6, если N_t имеет тип G , то либо $\overline{N}_t^2 = 0$, либо $\overline{\overline{N}}_t^2 = 0$. Поэтому получаем, что среди $N_9, \dots, N_{13}$ не более 2 имеют тип G . Получаем противоречие с леммой 7.

3) Пусть среди $p_1, \dots, p_8$ есть 3 пропорциональных вектора, которые не пропорциональны остальным векторам из $p_1, \dots, p_8$. Тогда так же, как в случае 2), можно считать, что

$$\begin{array}{l} D_1: \\ D_2: \\ D_3: \\ D_4: \\ D_5: \\ D_6: \\ D_7: \\ D_8: \end{array} \begin{array}{|c|c|c|} \hline 10 & ** & \\ \hline 10 & ** & \\ \hline 10 & ** & \\ \hline *1 & ** & \\ \hline *1 & ** & \\ \hline *1 & ** & \\ \hline *1 & ** & \\ \hline *1 & ** & \\ \hline \end{array}, \quad P: \begin{array}{|c|c|c|c|} \hline *0 & *0 & *0 & *0 \\ \hline *0 & *0 & *0 & *0 \\ \hline *0 & *0 & *0 & *0 \\ \hline ** & ** & ** & ** \\ \hline ** & ** & ** & ** \\ \hline ** & ** & ** & ** \\ \hline ** & ** & ** & ** \\ \hline ** & ** & ** & ** \\ \hline \end{array}.$$

Так же, как в случае 2), разобьем все матрицы на 2 подматрицы – одну в строках 1–3, другую в строках 4–8. Первые будем помечать одной чертой, вторые – двумя. Из равенства (25) опять получаем

$$\overline{P}_i = \sum_{t=9}^{13} e_i^t \overline{N}_t^2, \quad \overline{\overline{P}}_i = \sum_{t=9}^{13} e_i^t \overline{\overline{N}}_t^2.$$

Матрица $\overline{P}$ имеет ранг 3, поэтому первые столбцы в матрицах $\overline{P}_i$ имеют ранг 3. Следовательно, среди матриц $\overline{N}_t^2$ не менее 3 ненулевых. Столбцы с номерами 2, 4, 6, 8 в матрице P линейно независимы. Поэтому вторые столбцы в матрицах $\overline{P}_i$ линейно независимы и матрицы $\overline{\overline{P}}_1, \overline{\overline{P}}_2, \overline{\overline{P}}_3, \overline{\overline{P}}_4$ линейно независимы. Следовательно, среди матриц $\overline{\overline{N}}_t^2$ не менее 4 ненулевых. По лемме 6, если N_t имеет тип G , то либо $\overline{N}_t^2 = 0$ либо $\overline{\overline{N}}_t^2 = 0$, а по лемме 7 среди N_t не менее 3 матриц типа G . Получаем, что возможен единственный вариант (с точностью до перенумерации)

$$N_9^2 = \left| \begin{array}{c} \overline{F}_9 \\ \overline{F}_9 \end{array} \right|, \quad N_{10}^2 = \left| \begin{array}{c} \overline{F}_{10} \\ \overline{F}_{10} \end{array} \right|, \quad N_{11}^2 = \left| \begin{array}{c} \overline{G}_{11} \\ 0 \end{array} \right|, \quad N_{12}^2 = \left| \begin{array}{c} 0 \\ \overline{G}_{12} \end{array} \right|, \quad N_{13}^2 = \left| \begin{array}{c} 0 \\ \overline{G}_{13} \end{array} \right|,$$

причем N_9 и N_{10} – матрицы типа F , которые по лемме 6 имеют ранг 1. Тогда получаем, что матрицы $\overline{P}_1, \overline{P}_2, \overline{P}_3, \overline{P}_4$ являются линейными комбинациями матриц $\overline{F}_9, \overline{F}_{10}, \overline{G}_{11}$. Поскольку среди $\overline{P}_1, \overline{P}_2, \overline{P}_3, \overline{P}_4$ есть 3 линейно независимых, то в свою очередь матрицы $\overline{F}_9, \overline{F}_{10}, \overline{G}_{11}$ являются линейными комбинациями матриц $\overline{P}_1, \overline{P}_2, \overline{P}_3, \overline{P}_4$. Тогда в матрице $\overline{G}_{11}$ второй столбец нулевой и матрица N_{11}^2 имеет ранг 1. Но тогда в матрицах $N_9^2, \dots, N_{13}^2$ не более 7 линейно независимых столбцов, что противоречит лемме 7.

4) Пусть среди $p_1, \dots, p_8$ есть 2 пропорциональных вектора, которые не пропорциональны остальным векторам из $p_1, \dots, p_8$. Тогда так же, как в предыдущих случаях, можно считать, что

$$P: \begin{array}{l} D_1: \\ D_2: \\ D_3: \\ D_4: \\ D_5: \\ D_6: \\ D_7: \\ D_8: \end{array} \begin{array}{|c|c|c|} \hline 10 & ** & \\ \hline 10 & ** & \\ \hline *1 & ** & \\ \hline *1 & ** & \\ \hline *1 & ** & \\ \hline *1 & ** & \\ \hline *1 & ** & \\ \hline *1 & ** & \\ \hline \end{array}, \quad \begin{array}{l} *0 \\ *0 \\ ** \\ ** \\ ** \\ ** \\ ** \\ ** \end{array} \begin{array}{|c|c|c|c|} \hline *0 & *0 & *0 & *0 \\ \hline *0 & *0 & *0 & *0 \\ \hline ** & ** & ** & ** \\ \hline ** & ** & ** & ** \\ \hline ** & ** & ** & ** \\ \hline ** & ** & ** & ** \\ \hline ** & ** & ** & ** \\ \hline ** & ** & ** & ** \\ \hline \end{array}.$$

Так же, как в случаях 2) и 3), разобьем все матрицы на 2 подматрицы – одну в строках 1–2, другую в строках 3–8. Первые будем обозначать одной чертой, вторые – двумя. Из равенства (25) опять получаем

$$\bar{P}_i = \sum_{t=9}^{13} e_i^t \bar{N}_t^2, \quad \overline{\bar{P}}_i = \sum_{t=9}^{13} e_i^t \overline{\bar{N}}_t^2.$$

Матрица $\bar{P}$ имеет ранг 2, поэтому первые столбцы в матрицах $\bar{P}_i$ имеют ранг 2. Следовательно, среди матриц $\bar{N}_t^2$ не менее 2 ненулевых. Столбцы с номерами 2, 4, 6, 8 в матрице P линейно независимы. Поэтому вторые столбцы в матрицах $\bar{P}_i$ линейно независимы и матрицы $\bar{P}_1, \bar{P}_2, \bar{P}_3, \bar{P}_4$ линейно независимы. Следовательно, среди матриц $\bar{N}_t^2$ не менее 4 ненулевых. По лемме 6, если N_t имеет тип G , то либо $\bar{N}_t^2 = 0$, либо $\overline{\bar{N}}_t^2 = 0$, а по лемме 7 среди N_t не менее 3 матриц типа G . Возможны следующие 2 варианта.

4а) Среди матриц $N_9, \dots, N_{13}$ 1 матрица типа F и 4 матрицы типа G . Тогда получаем, что возможен единственный вариант (с точностью до перенумерации)

$$N_9^2 = \begin{vmatrix} \bar{F}_9 \\ \bar{F}_9 \end{vmatrix}, \quad N_{10}^2 = \begin{vmatrix} \bar{G}_{10} \\ 0 \end{vmatrix}, \quad N_{11}^2 = \begin{vmatrix} 0 \\ \bar{G}_{11} \end{vmatrix}, \quad N_{12}^2 = \begin{vmatrix} 0 \\ \bar{G}_{12} \end{vmatrix}, \quad N_{13}^2 = \begin{vmatrix} 0 \\ \bar{G}_{13} \end{vmatrix},$$

причем N_9 – матрица типа F , которая по лемме 6 имеет ранг 1. Тогда получаем, что матрицы $\bar{P}_1, \bar{P}_2, \bar{P}_3, \bar{P}_4$ являются линейными комбинациями матриц $\bar{F}_9, \bar{G}_{10}$. Поскольку среди $\bar{P}_1, \bar{P}_2, \bar{P}_3, \bar{P}_4$ есть 2 линейно независимых, то в свою очередь матрицы $\bar{F}_9, \bar{G}_{10}$ являются линейными комбинациями матриц $\bar{P}_1, \bar{P}_2, \bar{P}_3, \bar{P}_4$. Тогда в матрицах $\bar{G}_{10}$ и $\bar{F}_9$ второй столбец нулевой. А так как $\bar{F}_9 \neq 0$ и матрица N_9^2 имеет ранг 1, то и в матрице $\bar{F}_9$ второй столбец нулевой. Но тогда в матрицах $N_9^2, \dots, N_{13}^2$ вторые столбцы имеют ранг не более 3. Поскольку вторые столбцы матриц P_1, P_2, P_3, P_4 являются их линейными комбинациями (согласно (25)), то столбцы с номерами 2, 4, 6, 8 в матрице P имеют ранг не более 3, что противоречит невырожденности матрицы P .

4б) Среди матриц $N_9, \dots, N_{13}$ 2 матрицы типа F и 3 матрицы типа G . Тогда получаем, что возможны следующие 2 подварианта (с точностью до перенумерации).

4б1) Имеем

$$N_9^2 = \begin{vmatrix} \bar{F}_9 \\ \bar{F}_9 \end{vmatrix}, \quad N_{10}^2 = \begin{vmatrix} \bar{F}_{10} \\ \bar{F}_{10} \end{vmatrix}, \quad N_{11}^2 = \begin{vmatrix} 0 \\ \bar{G}_{11} \end{vmatrix}, \quad N_{12}^2 = \begin{vmatrix} 0 \\ \bar{G}_{12} \end{vmatrix}, \quad N_{13}^2 = \begin{vmatrix} 0 \\ \bar{G}_{13} \end{vmatrix},$$

причем N_9 и N_{10} – матрицы типа F , которые по лемме 6 имеют ранг 1. Тогда получаем, что матрицы $\bar{P}_1, \bar{P}_2, \bar{P}_3, \bar{P}_4$ являются линейными комбинациями матриц $\bar{F}_9, \bar{F}_{10}$. Поскольку среди $\bar{P}_1, \bar{P}_2, \bar{P}_3, \bar{P}_4$ есть 2 линейно независимых, то, в свою очередь, матрицы $\bar{F}_9, \bar{F}_{10}$ являются линейными комбинациями матриц $\bar{P}_1, \bar{P}_2, \bar{P}_3, \bar{P}_4$. Тогда в матрицах $\bar{F}_9$ и $\bar{F}_{10}$

второй столбец нулевой. А так как $\bar{F}_9 \neq 0$ и $\bar{F}_{10} \neq 0$ и обе матрицы N_9^2, N_{10}^2 имеют ранг 1, то в матрицах N_9^2, N_{10}^2 второй столбец нулевой. Но тогда в матрицах $N_9^2, \dots, N_{13}^2$ вторые столбцы имеют ранг не более 3. Поскольку вторые столбцы матриц P_1, P_2, P_3, P_4 являются их линейными комбинациями (согласно (25)), то столбцы с номерами 2, 4, 6, 8 в матрице P имеют ранг не более 3, что противоречит невырожденности матрицы P .

462) Имеем

$$N_9^2 = \begin{vmatrix} \bar{F}_9 \\ \bar{F}_9 \end{vmatrix}, \quad N_{10}^2 = \begin{vmatrix} \bar{F}_{10} \\ \bar{F}_{10} \end{vmatrix}, \quad N_{11}^2 = \begin{vmatrix} 0 \\ \bar{G}_{11} \end{vmatrix}, \quad N_{12}^2 = \begin{vmatrix} 0 \\ \bar{G}_{12} \end{vmatrix}, \quad N_{13}^2 = \begin{vmatrix} \bar{G}_{13} \\ 0 \end{vmatrix},$$

причем N_9 и N_{10} – матрицы типа F , которые по лемме 6 имеют ранг 1. Тогда получаем, что матрицы $\bar{P}_1, \bar{P}_2, \bar{P}_3, \bar{P}_4$ являются линейными комбинациями матриц $\bar{F}_9, \bar{F}_{10}, \bar{G}_{11}, \bar{G}_{12}$. Поскольку среди $\bar{P}_1, \bar{P}_2, \bar{P}_3, \bar{P}_4$ есть 4 линейно независимых, то в свою очередь матрицы $\bar{F}_9, \bar{F}_{10}, \bar{G}_{11}, \bar{G}_{12}$ являются линейными комбинациями матриц $\bar{P}_1, \bar{P}_2, \bar{P}_3, \bar{P}_4$. Согласно (23) матрицы $\bar{P}_1, \bar{P}_2, \bar{P}_3, \bar{P}_4$ являются линейными комбинациями матриц $D_3^1, \dots, D_8^1$. Поэтому и матрицы $\bar{F}_9, \bar{F}_{10}, \bar{G}_{11}, \bar{G}_{12}$ являются линейными комбинациями матриц $D_3^1, \dots, D_8^1$. По лемме 6 те строки, которые отличны от 0 в матрице $\bar{G}_{11}$, пропорциональны в $D_3^1, \dots, D_8^1$. А так как $\bar{G}_{11}$ является линейной комбинацией матриц $D_3^1, \dots, D_8^1$, то в $\bar{G}_{11}$ все строки пропорциональны. Значит, матрица N_{11}^2 имеет ранг 1. Тогда в матрицах $N_9^2, \dots, N_{13}^2$ не более 7 линейно независимых столбцов, что противоречит лемме 7.

5) Пусть среди $p_1, \dots, p_8$ нет пропорциональных векторов. Тогда по лемме 6 любая матрица N_t типа G будет ненулевой только в одной строке и, следовательно, будет матрицей первого ранга. Тогда все матрицы $N_9^2, \dots, N_{13}^2$ будут ранга 1 и будут иметь не более 5 линейно независимых столбцов, что противоречит лемме 7.

Случай 2. Среди матриц $D_1^1, \dots, D_{13}^1$ имеется 9 линейно независимых матриц, пусть это матрицы $D_1^1, \dots, D_9^1$, а матрицы $D_{10}^1, \dots, D_{13}^1$ являются линейными комбинациями матриц $D_1^1, \dots, D_9^1$. Тогда из (1) вытекает, что существуют такие коэффициенты b_t^p, k_i^p (из рассматриваемого поля), что выполняются равенства

$$D_t^1 = \sum_{p=1}^9 b_t^p D_p^1, \quad t = 10, \dots, 13, \quad P_i = \sum_{t=1}^9 k_i^t D_t^1, \quad i = 1, \dots, 4. \quad (26)$$

Так как 8 столбцов в матрицах P_i линейно независимы, то из (26) следует, что среди матриц первого ранга $D_1^1, \dots, D_9^1$ есть 8 ненулевых матриц, построенных на 8 линейно независимых столбцах (пусть это матрицы $D_1^1, \dots, D_8^1$). Умножая слева все матрицы (P и D_t) на матрицу C , переводящую эти 8 столбцов в столбцы с одной 1, получим по лемме 2 решение $CD_1, \dots, CD_{13}$ для задачи CP . При этом равенства (26) сохранятся. Будем считать, что наше исходное решение уже имеет такой вид. Тогда так же, как в случае 1, получаем, что каждая матрица $D_t, t = 1, \dots, 8$, отлична от 0 только в строке с номером t . Пусть в равенстве (26) среди $k_1^9, k_2^9, k_3^9, k_4^9$ есть отличные от 0. Будем считать, что это k_1^9 . Тогда, вычитая P_1 из P_2, P_3, P_4 с подходящими коэффициентами, можно для новых блоков получить

$$k_2^9 = k_3^9 = k_4^9 = 0. \quad (27)$$

По лемме 4 те же матрицы $D_1, \dots, D_{13}$ будут давать решение для полученной после вычитания блоков матрицы P' , причем легко видеть, что матрица P' останется невырожденной. Можно считать, что мы изначально рассматриваем такую матрицу, т.е. что равенства (27) выполняются (если изначально все $k_1^9, k_2^9, k_3^9, k_4^9$ равны 0, то это тоже выполняется).

Пусть

$$N_t = D_t - \sum_{p=1}^9 b_t^p D_p, \quad t = 10, \dots, 13.$$

Тогда из (26) получаем, что N_t имеет вид $N_t = |0|N_t^2|$, где N_t^2 – подматрица размера 8×2 . Так как среди матриц $D_1^1, \dots, D_{13}^1$ ровно 9 линейно независимых матриц, то пространство линейных комбинаций $\sum_{t=1}^{13} \mu_t D_t$ таких, что $\sum_{t=1}^{13} \mu_t D_t^1 = 0$, имеет размерность $13 - 9 = 4$. Но 4 линейные комбинации $D_t - \sum_{p=1}^9 b_t^p D_p$, $t = 10, \dots, 13$, выражающие N_t , входят в это пространство и линейно независимы, поскольку каждое слагаемое $D_{10}, \dots, D_{13}$ входит ровно в одну из этих линейных комбинаций. Следовательно, каждая матрица, которая является линейной комбинацией матриц $D_1, \dots, D_{13}$ и имеет вид $|0|H|$, является линейной комбинацией матриц $N_{10}, \dots, N_{13}$. В частности, существуют такие константы e_i^t (из рассматриваемого поля), что выполняются равенства

$$P_{i2} = |0|P_i| = \sum_{t=10}^{13} e_i^t N_t, \quad P_i = \sum_{t=10}^{13} e_i^t N_t^2, \quad i = 1, \dots, 4. \quad (28)$$

ЛЕММА 8. Пусть $D_1, \dots, D_{13}$ – решение задачи P для некоторой невырожденной матрицы P порядка 8. Пусть $D_t = |D_t^1|D_t^2|$ – разбиение каждой матрицы D_t на блоки ширины 2. Тогда ни для какого t невозможно $D_t^1 \equiv 0$ (и, аналогично, $D_t^2 \equiv 0$).

ДОКАЗАТЕЛЬСТВО. Допустим, что для некоторого k выполняется $D_k^1 \equiv 0$. Так как мы доказали, что случай 1 невозможен, то остается только вариант, что среди матриц $D_1^1, \dots, D_{13}^1$ имеется 9 линейно независимых. Поскольку $D_k^1 \equiv 0$, то D_k^1 не входит в их число. В равенстве (26) для D_k^1 все коэффициенты b_k^p будут равны 0 и, следовательно, для соответствующей матрицы N_k получим $N_k = D_k$ и N_k будет матрицей ранга 1. Тогда 4 матрицы N_t^2 в равенстве (28) будут иметь не более 7 линейно независимых столбцов. В равенстве (28) получим противоречие, поскольку в матрицах P_1, P_2, P_3, P_4 8 линейно независимых столбцов. Лемма доказана.

Из (26) следует, что матрица

$$P_{i1} - \sum_{t=1}^9 k_i^t D_t = |P_i|0| - \sum_{t=1}^9 k_i^t |D_t^1|D_t^2|$$

имеет вид $|0|H|$. В то же время она является линейной комбинацией матриц $D_1, \dots, D_{13}$, поскольку матрица P_{i1} является линейной комбинацией матриц $D_1, \dots, D_{13}$. Следовательно, она является линейной комбинацией матриц $N_{10}, \dots, N_{13}$. Значит, для некоторых констант k_i^t , $t = 10, \dots, 13$, выполняется

$$P_{i1} = \sum_{t=1}^9 k_i^t D_t + \sum_{t=10}^{13} k_i^t N_t, \quad i = 1, \dots, 4. \quad (29)$$

Из (28) следует, что P_1, P_2, P_3, P_4 выражаются линейными комбинациями 4 матриц $N_{10}^2, N_{11}^2, N_{12}^2, N_{13}^2$, и матрицы P_1, P_2, P_3, P_4 линейно независимы. Тогда матрицы $N_{10}^2, N_{11}^2, N_{12}^2, N_{13}^2$ в свою очередь выражаются как линейные комбинации матриц P_1, P_2, P_3, P_4 , т.е. существуют такие константы l_t^j , что

$$N_t^2 = \sum_{j=1}^4 l_t^j P_j, \quad N_t = \sum_{j=1}^4 l_t^j P_{j2}, \quad t = 10, \dots, 13. \quad (30)$$

Из (29) и (30) получаем, что существуют такие константы r_i^j , что

$$P_{i1} = \sum_{t=1}^9 k_i^t D_t + \sum_{j=1}^4 r_i^j P_{j2}, \quad i = 1, \dots, 4. \quad (31)$$

Тогда из (26), (27) и (31) получаем

$$P_i = \sum_{t=1}^8 k_i^t D_t^1, \quad P_{i1} = \sum_{t=1}^8 k_i^t D_t + \sum_{j=1}^4 r_i^j P_{j2}, \quad i = 2, 3, 4. \quad (32)$$

Если среди r_2^1, r_3^1, r_4^1 есть не равные 0, то пусть, например, $r_2^1 \neq 0$. Пусть тогда $v_3 = r_3^1/r_2^1$. Если среди r_2^1, r_3^1, r_4^1 все равны 0, то положим $v_3 = 0$. В любом случае $r_3^1 - v_3 r_2^1 = 0$. Заменим в матрице P блок P_3 на блок $P'_3 = P_3 - v_3 P_2$ (заметим, что равенства (27) сохраняются для новых блоков). По лемме 4 тот же набор матриц D_t является решением задачи P для новой невырожденной матрицы P' . Учитывая (32), получаем

$$\begin{aligned} P'_{31} &= |P'_3|_0 = P_{31} - v_3 P_{21} = \sum_{t=1}^8 (k_3^t - v_3 k_2^t) D_t + \sum_{j=1}^4 (r_3^j - v_3 r_2^j) P_{j2} \\ &= \sum_{t=1}^8 (k_3^t - v_3 k_2^t) D_t + (r_3^2 - v_3 r_2^2) P_{22} + (r_3^3 - v_3 r_2^3) P_{32} + (r_3^4 - v_3 r_2^4) P_{42}. \end{aligned}$$

Поскольку $P_3 = P'_3 + v_3 P_2$, то $P_{32} = P'_{32} + v_3 P_{22}$. Подставляя это выражение в предыдущее равенство вместо P_{32} , получим, что для некоторых констант $k_1, \dots, k_8, w_2, w_3, w_4$ выполняется равенство

$$P'_{31} = \sum_{t=1}^8 k_t D_t + w_2 P_{22} + w_3 P'_{32} + w_4 P_{42}. \quad (33)$$

Матрицы P_{22}, P'_{32}, P_{42} в первых двух столбцах равны 0. Поэтому в этих столбцах получаем $P'_3 = \sum_{t=1}^8 k_t D_t^1$. Поскольку $P'_3 \neq 0$, то среди $k_1, \dots, k_8$ есть ненулевые. В 3-м и 4-м столбцах равенство (33) дает

$$0 = \sum_{t=1}^8 k_t D_t^2 + w_2 P_2 + w_3 P'_3 + w_4 P_4. \quad (34)$$

Из (32) следует, что все матрицы P_2, P'_3, P_4 являются линейными комбинациями матриц $D_1^1, \dots, D_8^1$. Поэтому из (34) получаем, что для некоторых констант $h_1, \dots, h_8$ выполняется равенство

$$\sum_{t=1}^8 k_t D_t^2 + \sum_{t=1}^8 h_t D_t^1 \equiv 0, \quad (35)$$

причем среди $k_1, \dots, k_8$ есть ненулевые. Не ограничивая общности, пусть $k_1 \neq 0$. Так как для $t = 1, \dots, 8$ каждая матрица D_t отлична от 0 только в строке с номером t , то из (35), рассматривая первую строчку, получаем

$$k_1 D_1^2 + h_1 D_1^1 \equiv 0. \quad (36)$$

Проведем одновременно во всех матрицах $D_t, t = 1, \dots, 13$, следующее преобразование над блоками:

$$\begin{cases} \overline{D}_t^1 = h_1 D_t^1 + k_1 D_t^2, \\ \overline{D}_t^2 = D_t^1. \end{cases}$$

Определитель этого преобразования –

$$\begin{vmatrix} h_1 & k_1 \\ 1 & 0 \end{vmatrix} = -k_1 \neq 0.$$

Поэтому это преобразование обратимое и по лемме 5 новые матрицы $\overline{D}_t = |\overline{D}_t^1| \overline{D}_t^2|, t = 1, \dots, 13$, дают решение для той же невырожденной матрицы P' . При этом из (36) следует, что $\overline{D}_1^1 \equiv 0$, что невозможно по лемме 8. Теорема 4 полностью доказана.

ТЕОРЕМА 5. *Над любым полем билинейная сложность каждой из задач умножения: матрицы размера 4×2 на матрицу размера 2×2 , матрицы размера 2×4 на матрицу размера 4×2 , матрицы размера 2×2 на матрицу размера 2×4 , равна 14.*

ДОКАЗАТЕЛЬСТВО. Так как для задачи умножения матрицы размера 4×2 на матрицу размера 2×2 можно построить билинейный алгоритм билинейной сложности 14 (достаточно дважды применить алгоритм Штрассена билинейной сложности 7 для умножения двух матриц порядка 2), то из теоремы 4 следует, что билинейная сложность задачи умножения матрицы размера 4×2 на матрицу размера 2×2 равна 14. Для других двух задач утверждение теоремы вытекает из равенства билинейной сложности для двойственных задач умножения матриц [14] (см. также [4]).

Список литературы

- [1] А. А. Карацуба, Ю. П. Офман, “Умножение многозначных чисел на автоматах”, *ДАН СССР*, **145:2** (1962), 293–294.
- [2] V. Strassen, “Gaussian elimination is not optimal”, *Numer. Math.*, **13:4** (1969), 354–356; В. Штрассен, “Алгоритм Гаусса не оптимален”, Кибернетический сборник, **7**, Мир, М., 1970, 67–70.
- [3] D. Coppersmith, S. Winograd, “Matrix Multiplication via Arithmetic Progressions”, *J. Symbolic Comput.*, **9:3** (1990), 251–280.
- [4] В. Б. Алексеев, “Сложность умножения матриц. Обзор”, Кибернетический сборник, **25**, Мир, М., 1988, 189–236.
- [5] D. Bini, “Relations between exact and approximate bilinear algorithms. Applications”, *Calcolo*, **17** (1980), 87–97.
- [6] S. Winograd, “On multiplication of 2×2 matrices”, *Linear Algebra and Appl.*, **4:4** (1971), 381–388.
- [7] J. D. Laderman, “A noncommutative algorithm for multiplying 3×3 matrices using 23 multiplications”, *Bull. Amer. Math. Soc.*, **82:1** (1976), 126–128.
- [8] M. Bläser, “On the complexity of the multiplication of matrices of small formats”, *J. Complexity*, **19:1** (2003), 43–60.
- [9] A. Schönhage, “Partial and total matrix multiplication”, *SIAM J. Comput.*, **10:3** (1981), 434–455.
- [10] D. Bini, M. Capovani, F. Romani, G. Lotti, “ $O(n^{2.7799})$ complexity for $n \times n$ approximate matrix multiplication”, *Inform. Process. Lett.*, **8:5** (1979), 234–235.
- [11] J. E. Hopcroft, L. R. Kerr, “On minimizing the number of multiplications necessary for matrix multiplication”, *SIAM J. Appl. Math.*, **20:1** (1971), 30–36.
- [12] V. B. Alekseyev, “On the complexity of some algorithms of matrix multiplication”, *J. Algorithms*, **6:1** (1985), 71–85.
- [13] A. Waksman, “On Winograd’s algorithm for inner products”, *IEEE Trans. Comput.*, **C-19:4** (1970), 360–361.
- [14] J. E. Hopcroft, J. Musinski, “Duality applied to the complexity of matrix multiplication and other bilinear forms”, *SIAM J. Comput.*, **2:3** (1973), 159–173.