

#### Министерство науки и высшего образования Российской Федерации Федеральное государственное бюджетное образовательное учреждение

### высшего образования «Московский государственный технический университет имени Н.Э. Баумана

(национальный исследовательский университет)» (МГТУ им. Н.Э. Баумана)

#### ФАКУЛЬТЕТ ИНФОРМАТИКА И СИСТЕМЫ УПРАВЛЕНИЯ

КАФЕДРА «ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ЭВМ И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ» (ИУ7)

НАПРАВЛЕНИЕ ПОДГОТОВКИ 09.03.01 Информатика и вычислительная техника

## ОТЧЕТ

	по лаоора	аторнои рас	ооте № _2	
Название:	Трудоёмкостн	ь алгоритм	ов умножения	я матриц
Дисциплина:	Анализ алгор	<u>ИТМОВ</u>		
Студент	ИУ7-52Б		(П	Е.В. Брянская
	(Группа)		(Подпись, дата)	(И.О. Фамилия)
Преподаватель				Л.Л. Волкова
			(Подпись, дата)	(И.О. Фамилия)

# Оглавление

$\mathbf{B}_{1}$	Введение		3	
1	Ана	алитическая часть	4	
2	Кон	нструкторская часть	5	
	2.1	Стандартный алгоритм умножения матриц	5	
	2.2	Алгоритм Винограда	6	
	2.3	Алгоритм Винограда (оптимизированный)	6	
	2.4	Требования к ПО	7	
	2.5	Заготовки тестов	7	
3	Tex	нологическая часть	10	
	3.1	Выбранный язык программирования	10	
	3.2	Листинг кода	10	
	3.3	Результаты тестов	12	
	3.4	Оценка трудоёмкости	16	
	3.5	Оценка времени	17	
4	Исс	следовательская часть	19	
За	аклю	рчение	21	
$\mathbf{C}$	писо	к литературы	22	

# Введение

В этой лабораторной работе будет оцениваться трудоёмкость алгоритмов умножения матриц.

**Трудоёмкость алгоритма** - это зависимость стоимости операций от линейного(ых) размера(ов) входа(ов) [1].

Модель вычислений трудоёмкости должна учитывать:

- 1) стоимость базовых операций. К ним относятся: =, +, -, \*, /, ==, !=, <, <=, >=, %, +=, -=, \*=, /=, [ ], < <, > >. Каждая из операций имеет стоимость равную 1.
- 2) оценку цикла. Она складывается из стоимости тела, инкремента и сравнения.
- 3) оценку условного оператора if. Положим, что стоимость перехода к одной из веток равной 0. В таком случае, общая стоимость складывается из подсчета условия и рассмотрения худшего и лучшего случаев.

Оценка характера трудоёмкости даётся по наиболее быстрорастущему слагаемому.

## 1. Аналитическая часть

**Цель** данной работы — оценить трудоёмкость алгоритмов умножения матриц и получить практический навык оптимизации алгоритмов.

Для достижения поставленной цели необходимо решить ряд следующих задач:

- 1) дать математическое описание;
- 2) описать алгоритмы умножения матриц;
- 3) дать теоретическую оценку трудоёмкости алгоритмов;
- 4) реализовать эти алгоритмы;
- 5) провести замеры процессорного времени работы алгоритмов на материале серии экспериментов;
- 6) провести сравнительный анализ алгоритмов.

Умножение осуществляется над матрицами  $A[M \times N]$  и  $B[N \times Q]$ . Число столбцов первой матрицы должно совпадать с числом строк второй, а таком случае можно осуществлять умножение. Результатом является матрица  $C[M \times Q]$ , в которой число строк столько же, сколько в первой, а столбцов, столько же, сколько во второй.

В основе стандартного алгоритма умножения матриц лежит формула (1.1).

$$c_{i,j} = \sum_{k=1}^{N} (a_{i,k} \times b_{k,j})$$
(1.1)

Существует и другой алгоритм умножения - алгоритм Винограда. Обозначим строку  $A_{i,*}$  как  $\overrightarrow{u}$ ,  $B_{*,j}$  как  $\overrightarrow{v}$ . Пусть  $u=(u_1,u_2,u_3,u_4)$  и  $v=(v_1,v_2,v_3,v_4)$ , тогда их произведение равно согласно формуле 1.2.

$$u \cdot v = u_1 \cdot v_1 + u_2 \cdot v_2 + u_3 \cdot v_3 + u_4 \cdot v_4 \tag{1.2}$$

Выражение (1.2) можно преобразовать в формулу 1.3.

$$u \cdot v = (u_1 + v_2) \cdot (u_2 + v_1) + (u_3 + v_1) \cdot (u_4 + v_3) - u_1 \cdot u_2 - u_3 \cdot u_4 - v_1 \cdot v_2 - v_3 \cdot v_4 \quad (1.3)$$

Причём, при нечётном значении N нужно учесть ещё одно слагаемое  $u_5 \cdot v_5$ .

Алгоритм Винограда основывается на раздельной работе со слагаемыми из выражения (1.3).

# 2. Конструкторская часть

Рассмотрим и оценим работу алгоритмов на матрицах  $A[M \times N]$  и  $B[N \times Q]$ .

### 2.1. Стандартный алгоритм умножения матриц

В основе этого алгоритма лежит формула (1.1). То есть для вычисления произведения двух матриц, каждая строка первой матрицы почленно умножается на каждый столбец второй, и затем подсчитывается сумма таких произведений, и полученный результат записывается в соответствующую ячейку результурующей матрицы.

Схема алгоритма представлена на Рис.2.1.

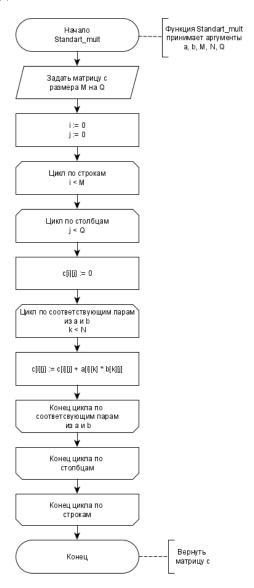


Рис. 2.1 — Стандартный алгоритм умножения матриц

## 2.2. Алгоритм Винограда

Цель данного алгоритма - сократить долю умножений в самом тяжёлом, затратном участке кода. Для этого используется формула (1.3).

Некоторые из слагаемых можно вычислить заранее и использовать повторно для каждой строки первой матрицы и для каждого столбца второй. Таким образом, трудо-ёмкость алгоритма уменьшается за счёт сокращения количества производимых операций.

В этом алгоритме важно учитывать, что при нечётном значении N, необходимо вычислять дополнительное слагаемое  $u_N \cdot v_N$ .

Схема алгоритма представлена на Рис.2.2.

# 2.3. Алгоритм Винограда (оптимизированный)

Алгоритм призван уменьшить трудоёмкость алгоритма, чтобы это сделать были использованы оптимизации:

- 1) видоизменён цикл по k, изменён шаг и условие. Таким образом, ушла необходимость в целочисленном делении, и в теле цикла не требуется больше умножать k на 2 каждый раз.
- 2) введена вспомогательная переменная buf, в которую записывается промежуточнее значение соответсвующей ячейки матрицы, и затем, конечный результат переносится в саму матрицу. Тем самым, уменьшается количество обращений к элементам матрицы, находящимся по конкретному адресу.
- 3) заранее высчитываются некоторые значения, например, n 1, которые далее используюся во вложенных циклах.
- 4) используется дополнительная переменная t=k 1, чтобы сократить число подсчетов этого значения на каждом шаге цикла.
- 5) объединён цикл 3 и 4, что позволило избежать ещё одного вложенного цикла.

Схема алгоритма представлена на Рис.2.3.

# 2.4. Требования к ПО

Для корректной работы алгоритмов и проведения тестов необходимо выполнить.

- 1) Обеспечить возможность ввода двух матриц через консоль и выбора алгоритма для умножения.
- 2) Вывести, в случае ввода размеров матриц, не удовлетворяющих главному условию, соответствующее сообщение. Программа не должна аварийно завершаться.
- 3) Рассчитать искомую матрицу и вывести её на экран.
- 4) Реализовать функцию замера процессорного времени, которое выбранный метод затрачивает на вычисление результата. Вывести результаты замеров на экран.

#### 2.5. Заготовки тестов

При проверке на корректность работы реализованных функций необходимо провести следующие тесты:

- умножение матриц размером  $1 \times 1$ ;
- квадратные матрицы;
- прямоугольные матрицы;
- ullet чётное и нечётное значение N.

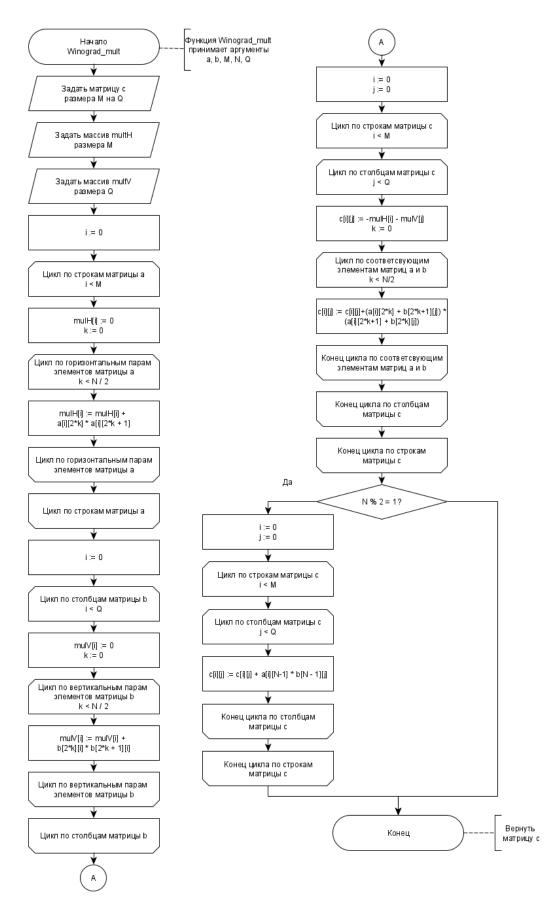


Рис. 2.2 — Алгоритм Винограда

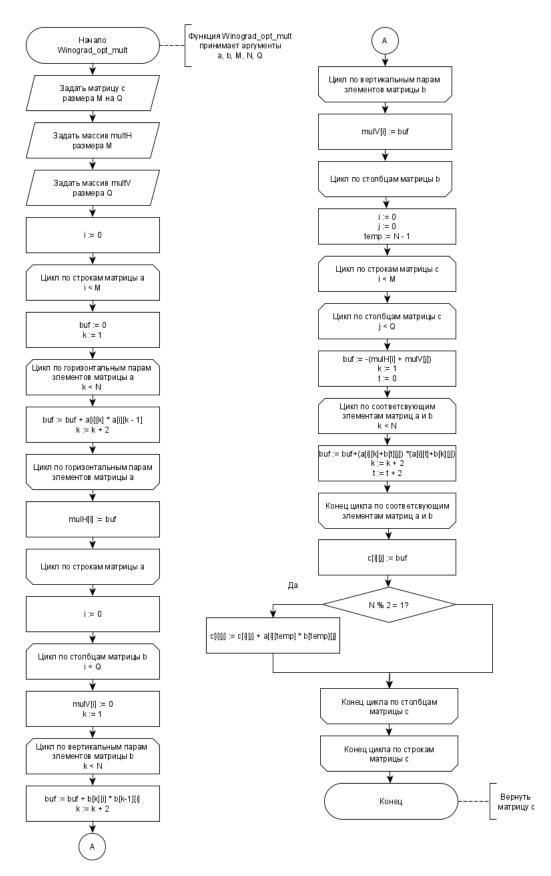


Рис. 2.3 — Оптимизированный алгоритм Винограда

## 3. Технологическая часть

## 3.1. Выбранный язык программирования

Для выполнения этой лабораторной работы был выбран язык программирования C++, так как есть большой навык работы с ним и с подключаемыми библиотеками, которые также использовались для проведения тестирования и замеров.

Использованная среда разработки - Visual Studio [3].

## 3.2. Листинг кода

Ниже представлены Листиги 3.1 - 3.3 функций, реализующих алгоритмы поиска расстояний.

Листинг 3.1 — Стандартный алгоритм умножения матриц

Листинг 3.2 — Алгоритм Винограда

```
matrix_t winograd_mult(matrix_t a, matrix_t b, int m, int n, int q)

arr_t mulH = create_array(m);
arr_t mulV = create_array(q);
matrix_t c = create_matrix(m, q);

for (int i = 0; i < m; i++)

{</pre>
```

```
mulH[i] = 0;
9
      for (int k = 0; k < n / 2; k++)
10
        mulH[i] = mulH[i] + a[i][2 * k] * a[i][2 * k + 1];
11
    }
12
13
    for (int i = 0; i < q; i++)
14
15
      mulV[i] = 0;
16
      for (int k = 0; k < n / 2; k++)
17
18
        mulV[i] = mulV[i] + b[2 * k][i] * b[2 * k + 1][i];
19
    }
20
^{21}
    for (int i = 0; i < m; i++)
22
      for (int j = 0; j < q; j++)
23
^{24}
        c[i][j] = -mulH[i] - mulV[j];
         for (int k = 0; k < n / 2; k++)
26
          c[i][j] = c[i][j] + (a[i][2 * k] + b[2 * k + 1][j]) *
27
                      (a[i][2 * k + 1] + b[2 * k][j]);
28
      }
29
30
    if (n % 2)
31
      for (int i = 0; i < m; i++)
32
         for (int j = 0; j < q; j++)
33
          c[i][j] = c[i][j] + a[i][n-1] * b[n-1][j];
34
35
36
    return c;
37 }
```

Листинг 3.3 — Оптимизированный алгоритм Винограда

```
matrix_t winograd_mult(matrix_t a, matrix_t b, int m, int n, int q)

arr_t mulH = create_array(m);
arr_t mulV = create_array(q);
double buf;

matrix_t c = create_matrix(m, q);

for (int i = 0; i < m; i++)</pre>
```

```
{
10
       buf = 0;
11
       for (int k = 1; k < n; k += 2)
12
         buf += a[i][k] * a[i][k - 1];
13
      mulH[i] = buf;
14
    }
15
16
    for (int i = 0; i < q; i++)
17
18
       buf = 0;
19
       for (int k = 1; k < n; k += 2)
^{20}
         buf += b[k][i] * b[k - 1][i];
21
      mulV[i] = buf;
22
    }
23
24
    int temp = n - 1;
^{25}
    int is odd = n \% 2;
27
    for (int i = 0; i < m; i++)
28
       for (int j = 0; j < q; j++)
29
30
         buf = -(mulH[i] + mulV[j]);
31
         for (int k = 1, t = 0; k < n; k += 2, t += 2)
32
           buf += (a[i][k] + b[t][j]) * (a[i][t] + b[k][j]);
33
         c[i][j] = buf;
34
35
         if (is odd)
36
           c[i][j] += a[i][temp] * b[temp][j];
37
      }
38
39
    return c;
40
41 }
```

## 3.3. Результаты тестов

Для тестирования были написаны функции, проверяющие, согласно заготовкам выше, случаи. Выводы о корректности работы делаются на основе сравнения результатов.

Все тесты пройдены успешно. Сами тесты представлены ниже (Листинг 3.4).

```
|boolmult_cmp(matrix_t a, matrix_t b, int m, int n, int q)
2 {
    matrix t c1 = standart mult(a, b, m, n, q);
3
    matrix t c2 = winograd mult(a, b, m, n, q);
    bool res = cmp matrix (c1, c2, m, q);
6
    free matrix(&c1, m, q);
8
    free matrix(&c2, m, q);
9
    return res;
11
12 }
_{14} // Матрицы размером 1 \times 1
void test size 1 1()
16 {
    int n = 1;
17
    matrix t a = create matrix(n, n);
19
    matrix t b = create matrix(n, n);
20
21
    a[0][0] = 15;
22
    b[0][0] = -7;
^{23}
24
    if (!mult cmp(a, b, n, n, n))
25
26
      cout << endl << __FUNCTION__ << " FAILED" << endl;
27
      free matrix(&a, n, n);
28
      free matrix(&b, n, n);
      return;
30
    }
31
32
    free matrix(&a, n, n);
33
    free matrix(&b, n, n);
34
35
    cout << end << __FUNCTION__ << " OK" << end |;
^{36}
37 }
38
39 // Квадратные матрицы
```

```
40 void test square matr()
41 {
    int n[] = { 2, 6, 10 };
42
43
    for (int i = 0; i < sizeof(n) / sizeof(n[0]); i++)
44
    {
45
      matrix t = random fill matrix(n[i], n[i]);
46
      matrix t b = random fill matrix (n[i], n[i]);
47
48
      if (!mult cmp(a, b, n[i], n[i], n[i]))
49
50
        cout << endl << FUNCTION__ << "FAILED" << endl;
51
        free matrix(&a, n[i], n[i]);
52
        free matrix(&b, n[i], n[i]);
53
         return;
54
      }
55
56
      free_matrix(&a, n[i], n[i]);
57
      free matrix(&b, n[i], n[i]);
58
59
      cout << end << __FUNCTION__ << " OK" << end I;
60
61
62 }
63
  // Прямоугольные матрицы
  void test rectangulat matr()
66 {
    int m[] = \{ 2, 6, 10 \};
67
    int n[] = \{ 1, 4, 7 \};
68
    int q[] = \{ 3, 4, 8 \};
70
    for (int i = 0; i < sizeof(n) / sizeof(n[0]); i++)
71
72
      matrix t a = random_fill_matrix(m[i], n[i]);
73
      matrix t b = random fill matrix (n[i], q[i]);
74
75
      if (!mult_cmp(a, b, m[i], n[i], q[i]))
76
77
        cout << endl << FUNCTION << " FAILED" << endl;</pre>
78
        free matrix(&a, m[i], n[i]);
79
```

```
free_matrix(&b, n[i], q[i]);
80
          return;
81
82
       free matrix(&a, m[i], n[i]);
83
       free matrix(\&b, n[i], q[i]);
84
85
       cout << endl << FUNCTION << "OK" << endl;
86
87
88 }
89
90 // Матрицы с чётным размером
   void test even size()
92 {
     int m[] = \{ 2, 4 \};
93
     int n[] = \{ 6, 2 \};
94
     int q[] = \{ 2, 8 \};
95
96
     for (int i = 0; i < sizeof(n) / sizeof(n[0]); i++)
97
     {
98
       matrix t a = random fill matrix(m[i], n[i]);
99
       matrix t b = random fill matrix (n[i], q[i]);
100
101
       if (!mult cmp(a, b, m[i], n[i], q[i]))
102
       {
103
         cout << endl << FUNCTION << " FAILED" << endl;</pre>
104
         free _ matrix(&a, m[i], n[i]);
105
         free matrix(\&b, n[i], q[i]);
106
          return;
107
       }
108
109
       free matrix(&a, m[i], n[i]);
110
       free _ matrix(&b, n[i], q[i]);
111
112
       cout << endl << __FUNCTION__ << "OK" << endl;
113
1\,1\,4
115 }
116
117 // Матрицы с нечётным размером
118 void test odd size()
119 {
```

```
int m[] = { 3, 3 };
120
     int n[] = {3, 1};
121
     int q[] = \{ 5, 7 \};
122
123
     for (int i = 0; i < sizeof(n) / sizeof(n[0]); i++)
124
     {
125
       matrix t a = random \ fill \ matrix(m[i], n[i]);
126
       matrix t b = random fill matrix (n[i], q[i]);
127
128
       if (!mult_cmp(a, b, m[i], n[i], q[i]))
129
130
         cout << endl << FUNCTION << " FAILED" << endl;</pre>
131
         free matrix(\&a, m[i], n[i]);
132
         free matrix(&b, n[i], q[i]);
133
         return;
134
       }
135
136
       free matrix(&a, m[i], n[i]);
137
       free matrix(&b, n[i], q[i]);
138
139
       cout << endl << FUNCTION << "OK" << endl;
140
141
142 }
143
  void run tests()
145 {
     test size 1 1();
146
     test square matr();
147
     test rectangulat matr();
148
     test even size();
149
     test odd size();
150
151 }
```

# 3.4. Оценка трудоёмкости

Произведём оценку трудоёмкости приведённых алгоритмов. Рассмотрим умножение матриц  $A[M \times N]$  и  $B[N \times Q]$ .

#### Стандартный алгоритм

$$f_{st} = 2 + M(2 + 2 + Q(3 + 2 + 2 + N(2 + 1 + 2 + 2 + 1 + 2)))$$
  
$$f_{st} = 2 + 4M + 7MQ + 10MNQ$$

#### Алгоритм Винограда (неоптимизированный)

$$f_{w} = 2 + M(2 + 3 + 2 + \frac{N}{2}(12 + 3)) + 2 + Q(2 + 3 + 2 + \frac{N}{2}(12 + 3)) + 2 + M(2 + 2 + Q(7 + 2 + 3 + \frac{N}{2}(23 + 3))) + 1 + \begin{bmatrix} 0, & \text{s.c.} \\ 2 + M(2 + 2 + Q(13 + 2)), & \text{x.c.} \end{bmatrix}$$

$$f_{w} = 7 + 11M + 7Q + \frac{15}{2}MN + \frac{15}{2}NQ + 12MQ + 13MNQ + \begin{bmatrix} 0, & \text{s.c.} \\ 2 + 4M + 15MQ, & \text{x.c.} \end{bmatrix}$$

#### Алгоритм Винограда (оптимизированный)

$$f_{wop} = 2 + M(2 + 1 + 2 + \frac{N}{2}(7 + 2) + 2) + 2 + Q(2 + 1 + 2 + \frac{N}{2}(7 + 2) + 2) + 2 + 2 + M(2 + 2) + Q(2 + 4 + 3 + \frac{N}{2}(3 + 12) + 3 + 1 + \begin{bmatrix} 0, & \text{s.c.} \\ 8, & \text{s.c.} \end{bmatrix})$$

$$f_{wop} = 8 + 11M + 7Q + 4.5MN + 4.5NQ + 13MQ + 7.5MNQ + \begin{bmatrix} 0, & \text{s.c.} \\ 8MQ, & \text{s.c.} \end{bmatrix}$$

## 3.5. Оценка времени

Процессорное время измеряется с помощью функции QueryPerformanceCounter библиотеки windows.h [2]. Осуществление замеров показано ниже (Листинг 3.5).

Листинг 3.5 — Замеры процессорного времени

```
void test time(matrix t(*f)(matrix t, matrix t, int, int, int), int n)
_{2}| {
    matrix t a = random fill matrix(n, n);
    matrix t b = random fill matrix(n, n);
    matrix t c;
    int num = 0;
    start measuring();
    while (get measured() < 3 * 1000)
10
1\,1
      c = f(a, b, n, n, n);
12
      free matrix(&c, n, n);
13
      num++;
14
    }
15
16
```

```
double t = get measured() / 1000;
17
    cout << "Выполнено" << num << " операций за " << t << " секунд" << endl;
18
    cout << "Время: " << t / num << endl;
20
    free matrix(&a, n, n);
21
    free matrix(&b, n, n);
23 }
^{24}
void test range (vector < int > &n)
26 \mid \left\{ \right.
    for (int key : n)
27
28
      cout << endl << "Размер тестируемых матриц: " << key << "x" <<
29
     key \ll endl;
30
      cout << endl << "-----Standart-----" << endl;</pre>
31
      test time(standart mult, key);
      cout << endl << "------Winograd------" << endl;
33
       test time(winograd mult, key);
34
       cout << endl << "-----Winograd (improved)-----" << endl;</pre>
35
       test time(winograd opt mult, key);
36
    }
38 }
```

# 4. Исследовательская часть

#### Характеристики ПО

При проведении замеров времени использовался компьютер, имеющий следующие характеристики:

- OC Windows 10 Pro
- Процессор Inter Core i7 10510U (1800 МГц)
- Объём ОЗУ 16 Гб

#### Измерения

Для проведения замеров процессорного времени использовались квадратные матрицы. Их содержимое генерируется случайным образом. Было проведено две серии экспериментов, ориентированных на выявление чувствительности алгоритмов к чётным и нечётным значениям N.

- 1) 50, 100, 200, 300, 400, 500, 600, 700
- 2) 51, 101, 201, 301, 401, 501, 601, 701

Каждый замер проводится 5 раз для получения более точного среднего результата. В таблице 4.1 и таблице 4.2 представлены результаты замеров процессорного времени работы реализаций алгоритмов (в сек).

Таблица 4.1 — Результаты измерений (чётная размерность)

Размер n								
	50	100	200	300	400	500	600	700
Алгоритм								
Стандартный	$5.1 * 10^{-4}$	$4.1 * 10^{-3}$	0.037	0.133	0.322	0.777	1.08	1.445
Виноград	$3.5 * 10^{-4}$	$2.9 * 10^{-3}$	0.027	0.096	0.237	0.559	0.727	1.226
Виноград	$3.4 * 10^{-4}$	$2.5 * 10^{-3}$	0.024	0.084	0.207	0.474	0.601	1.101
(оптимизированный)								

Таблица 4.2 — Результаты измерений (нечётная размерность)

Размер п								
/	51	101	201	301	401	501	601	701
Алгоритм								
Стандартный	$5.4 * 10^{-4}$	$4.2 * 10^{-3}$	0.037	0.133	0.329	0.838	0.982	1.628
Виноград	$4.2 * 10^{-4}$	0.003	0.028	0.098	0.240	0.6	0.852	1.465
Виноград	$3.5 * 10^{-4}$	$2.5 * 10^{-3}$	0.025	0.082	0.206	0.512	0.705	1.052
(оптимизированный)								

Согласно полученным данным можно сделать следующие выводы.

- 1) Оптимизированный алгоритм Винограда осуществляет умножение матриц быстрее, чем два других сравниваемых алгоритма (на 36% быстрее стандартного).
- 2) Оптимизированный алгоритм Винограда показывает результаты, примерно на 14%, чем неоптимизированный, что подтверждает разницу в рассчётах трудоёмкости обоих алгоритмов.
- 3) Результаты на матрицах с чётным и нечётным размером N отличаются на предложенном множестве значений, но не существенно (около 9%).
- 4) Наихудшие результаты измерений показал стандартный алгоритм умножения.

### Заключение

В ходе лабораторной работы была достигнута поставленная цель, а именно, оценена трудоёмкость алгоритмов умножения матриц и рассмотрены возможные оптимизации алгоритма Винограда.

В процессе выполнения были решены все задачи. Описаны все рассматриваемые алгоритмы, дана теоретическая оценка трудоёмкости каждого. Все проработанные алгоритмы реализованы, кроме того, были проведены замеры процессорного времени работы на материале серии экспериментов и проведён сравнительный анализ, сделаны выводы:

- 1) оптимизированный алгоритм Винограда, как и ожидалось, выполняет умножение матриц быстрее, чем два других сравниваемых алгоритма (на 36% он менее затратен по процессорному времени, чем стандартный и на 14%, чем неоптимизированный);
- оптимизированный алгоритм Винограда осуществляет умножение мтариц быстрее, чем неоптимизированный, что подтверждает разницу в рассчётах трудоёмкости обоих алгоритмов;
- 3) результаты отличаются несущественно на матрицах с чётным и нечётным размером N (разница примерно в 9%);
- 4) наихудшие результаты измерений времени показал стандартный алгоритм умножения.

# Список литературы

- 1. Трудоёмкость алгоритмов и временные оценки [Электронный ресурс]. Режим доступа: http://techn.sstu.ru/kafedri/подразделения/1/MetMat/shaturn/theoralg/5.htm, свободный (дата обращения: 29.09.20).
- 2. QueryPerformanceCounter function [Электронный ресурс]. Режим доступа: https://docs.microsoft.com/en-us/windows/win32/api/profileapi/nf-profileapi-queryperformancecounter, свободный (дата обращения: 01.10.2020).
- 3. Документация по Visual Studio 2019 [Электронный ресурс]. Режим доступа: https://docs.microsoft.com/ru-ru/visualstudio/windows/?view=vs-2019, свободный (дата обращения: 01.10.2020)