Comput. Math. Applic. Vol. 15, No. 12, pp. 977–989, 1988
Printed in Great Britain.

0097-4943/88 \$3.00 + 0.00

COMPLEXITIES OF SPECIAL MATRIX MULTIPLICATION PROBLEMS

J. TAKCHE
Department of Mathematics, Faculty of Sciences, P.O. Box 72, Mansourieh-el Metn, Lebanon

(*Received* 21 *January* 1986; *in revised form* 21 *September* 1987)

Communicated by E. Y. Rodin

Abstract—This paper develops optimal algorithms to multiply an $n \times n$ symmetric tridiagonal matrix by: (i) an arbitrary $n \times m$ matrix using $2nm - m$ multiplications; (ii) a symmetric tridiagonal matrix using $6n - 7$ multiplications; and (iii) a tridiagonal matrix using $7n - 8$ multiplications. Efficient algorithms are also developed to multiply a tridiagonal matrix by an arbitrary matrix, and to multiply two tridiagonal matrices.

1. INTRODUCTION

Let K be a ring and let $\mathbf{x} = (x_1, \ldots, x_p)^T$ and $\mathbf{y} = (y_1, \ldots, y_q)^T$ be two column vectors of indeterminates. A system $\mathbf{B} = \{B_1, \ldots, B_m\}$ of bilinear forms is given by

$$B_i = \sum_{j=1}^{p} \sum_{k=1}^{q} a_{ijk} x_j y_k = \mathbf{x}^T \mathbf{G}_i \mathbf{y},$$

where $\mathbf{G}_i$ is a $p \times q$ matrix with elements in K. The complexity $\delta(\mathbf{B})$ of the set of bilinear forms is the smallest number of nonscalar multiplications required to compute the B_is. One can show that δ is equal to the smallest number of rank one matrices necessary to include the $\mathbf{G}_i$s in their span [1].

The problem can be reformulated by introducing the characteristic matrix

$$G(\mathbf{s}) = \sum_{i=1}^{m} s_i \mathbf{G}_i,$$

where $\{s_i\}$ is a set of new indeterminates. We denote by $\delta(G(\mathbf{s}))$ the complexity of the associated set of bilinear forms. If $\mathbf{P}$, $\mathbf{Q}$ and $\mathbf{R}$ are respectively $p \times p$, $q \times q$ and $m \times m$ nonsingular matrices, then $\delta(\mathbf{PG(Rs)Q}) = \delta(G(\mathbf{s}))$.

One of the main open problems in algebraic complexity theory is the matrix multiplication problem: given an $m \times n$ matrix $\mathbf{X} = \{x_{ij}\}$ and an $n \times p$ matrix $\mathbf{Y} = \{y_{jk}\}$, we want to compute their product, an $m \times p$ matrix whose entries are the bilinear forms

$$z_{ik} = \sum_{j=1}^{n} x_{ij} y_{jk}.$$

Except for the trivial cases, the only known optimal algorithms correspond to the multiplication of 2×2 by $2 \times n$ matrices using $\lceil 7n/2 \rceil$ multiplications, over Z or Z_2 [2]. In this paper we attack several special cases. In the next section, we develop optimal algorithms to multiply a symmetric tridiagonal matrix respectively by an arbitrary matrix, a symmetric tridiagonal matrix and a tridiagonal matrix. In Section 3, we find elegant and efficient algorithms to multiply a tridiagonal matrix by an arbitrary matrix and to multiply two tridiagonal matrices.

Before closing this section we state one theorem from [1] which will be used frequently.

Theorem 1.1 [1]

Let $K = F$ be a field and let $G(\mathbf{s})$ be a proper characteristic matrix. Then, for

$$G(\mathbf{s}) = [G_1(\mathbf{s}) \vdots G_2(\mathbf{s})], \quad G(\mathbf{s}) = \begin{bmatrix} G_1(\mathbf{s}) \\ G_2(\mathbf{s}) \end{bmatrix} \quad \text{or} \quad G(\mathbf{s}) = G_1(\mathbf{u}) + G_2(\mathbf{v}),$$

we have respectively

$$\delta[G_1(\mathbf{s}) \vdots G_2(\mathbf{s})] \geqslant \min_M \delta[G_1(\mathbf{s}) + G_2(\mathbf{s})M] + \text{column rank } (G_2), \tag{1}$$

$$\delta\begin{bmatrix} G_1(\mathbf{s}) \\ G_2(\mathbf{s}) \end{bmatrix} \geqslant \min_N \delta[G_1(\mathbf{s}) + NG_2(\mathbf{s})] + \text{row rank } (G_2), \tag{2}$$

$$\delta(G_1(\mathbf{u}) + G_2(\mathbf{v})) \geqslant \min_T \delta(G_1(\mathbf{u}) + G_2(T\mathbf{u})) + \dim \mathbf{v}. \tag{3}$$

2. SYMMETRIC TRIDIAGONAL MATRIX MULTIPLICATION PROBLEM

Let $\mathbf{X}$ be an $n \times n$ symmetric tridiagonal matrix, i.e.

$$\mathbf{X} = \begin{bmatrix} x_1 & x_2 & & & & \\ x_2 & x_3 & x_4 & & & \\ & x_4 & x_5 & \cdot & & \\ & & \cdot & \cdot & \cdot & \\ & & & \cdot & \cdot & \cdot \\ & & & x_{2n-4} & x_{2n-3} & x_{2n-2} \\ & & & & x_{2n-2} & x_{2n-1} \end{bmatrix}.$$

First, consider the problem of multiplying $\mathbf{X}$ by an arbitrary n-dimensional vector

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{bmatrix}.$$

One equivalent form of the characteristic matrix is $\mathbf{X}$ itself. Let E_{ij} be the $n \times n$ matrix that consists of 1 in the (i, j) entry and 0 elsewhere. Then

$$\begin{aligned} G(\mathbf{x}) = \mathbf{X} &= \sum_{i=1}^{n} x_{2i-1}E_{ii} + \sum_{i=1}^{n-1} x_{2i}(E_{i,i+1} + E_{i+1,i}) \\ &= \sum_{i=1}^{n} x_{2i-1}E_{ii} + \sum_{i=1}^{n-1} x_{2i}(E_{i,i+1} + E_{i+1,i} + E_{ii} + E_{i+1,i+1}) \\ &\qquad - \sum_{i=1}^{n-1} x_{2i}(E_{ii} + E_{i+1,i+1}) = \mathbf{D} + \sum_{i=1}^{n-1} x_{2i}M_{i,i+1}, \end{aligned}$$

where $M_{i,i+1} = E_{i,i+1} + E_{i+1,i} + E_{ii} + E_{i+1,i+1}$ is a rank one matrix and $\mathbf{D}$ is a diagonal matrix. Hence, $\delta(G(\mathbf{x})) \leqslant n - 1 + \delta(\mathbf{D}) \leqslant n - 1 + n = 2n - 1$. On the other hand, $\dim \mathbf{x} = 2n - 1$. Hence we have the following lemma.

Lemma 2.1

The complexity of multiplying an $n \times n$ symmetric tridiagonal matrix by an arbitrary vector is $2n - 1$, over any ring.

Corollary 2.1

The complexity δ_{nm} of multiplying an $n \times n$ symmetric tridiagonal matrix by an arbitrary $n \times m$ matrix satisfies

$$\delta_{nm} \leqslant (2n - 1)m.$$

Theorem 2.2

The complexity δ_{nm} of multiplying an $n \times n$ symmetric tridiagonal matrix by an arbitrary $n \times m$ matrix is equal to $(2n-1)m$.

Proof. We have to show that $(2n-1)m$ is a lower bound. The characteristic matrix is:

$$G_{nm}(\mathbf{s}) = \begin{bmatrix} s_1 \dots s_m & & & & & \\ s_{m+1} \dots s_{2m} & s_1 \dots s_m & & & & \\ & s_{m+1} \dots s_{2m} & & & & \\ & s_{2m+1} \dots s_{3m} & s_{m+1} \dots s_{2m} & & & \\ & & s_{2m+1} \dots s_{3m} & & & \\ & & s_{3m+1} \dots s_{4m} & & & \\ & & & \ddots & & \\ & & & & s_{nm-3m+1} \dots s_{nm-2m} & \\ & & & & s_{nm-2m+1} \dots s_{nm-m} & \\ & & & & s_{nm-m+1} \dots s_{nm} & s_{nm-2m+1} \dots s_{nm-m} \\ & & & & & s_{nm-m+1} \dots s_{nm} \end{bmatrix}.$$

Note that $G_{n-1,m}(\mathbf{s})$ is obtained from $G_{nm}(\mathbf{s})$ by deleting the last two rows and the last m columns. Using Theorem 1.1, we get $\delta(G_{nm}(\mathbf{s})) \geqslant m + \delta(\bar{G}(\mathbf{s}))$ where $\bar{G}(\mathbf{s})$ is obtained from $G_{nm}(\mathbf{s})$ by deleting the last m columns and adding linear combinations of them to the remaining ones. In $\bar{G}(\mathbf{s})$ replace $s_{nm-m+1}, s_{nm-m+2}, \dots, s_{nm}$ by linear combinations of the remaining ones and delete the last 2 rows, we get $\delta(\bar{G}(\mathbf{s})) \geqslant m + \delta(G_{n-1,m}(\mathbf{s}))$. Therefore, $\delta(G_{nm}(\mathbf{s})) \geqslant 2m + \delta(G_{n-1,m}(\mathbf{s}))$; and this is true $\forall n \geqslant 1$. Hence $\delta(G_{nm}(\mathbf{s})) \geqslant 2m(n-1) + \delta(G_{1m}(\mathbf{s}))$. But $G_{1m}(\mathbf{s}) = [s_1, s_2, \dots, s_m]$, therefore $\delta(G_{1m}(\mathbf{s})) = m$, hence $\delta(G_{nm}(\mathbf{s})) \geqslant (2n-1)m$.

Corollary 2.2

The complexity of multiplying an $n \times n$ symmetric tridiagonal matrix by an arbitrary $n \times n$ matrix is

$$\delta(G_n(\mathbf{s})) = 2n^2 - n.$$

Secondly, consider now the problem of multiplying two $n \times n$ symmetric tridiagonal matrices. Let us start with the case when $n = 2$:

$$\mathbf{X} = \begin{bmatrix} x_1 & x_2 \\ x_2 & x_3 \end{bmatrix} \quad \text{and} \quad \mathbf{X}' = \begin{bmatrix} y_1 & y_2 \\ y_2 & y_3 \end{bmatrix}.$$

The characteristic matrix is

$$G_2(\mathbf{s}) = \begin{bmatrix} s_1 & s_2 & \\ s_3 & s_1 + s_4 & s_2 \\ & s_3 & s_4 \end{bmatrix}.$$

Lemma 2.3

$\delta(G_2(\mathbf{s})) = 5$, over any field of characteristic $\neq 2$.

Proof.

$$\delta(G_2(\mathbf{s})) \geqslant 1 + \delta \begin{bmatrix} s_1 & s_2 + \alpha s_3 & \alpha s_4 \\ s_3 & s_4 + s_1 + \beta s_3 & s_2 + \beta s_4 \end{bmatrix}$$

$$\geqslant 1 + \delta \begin{bmatrix} s_1 & s_2 + \alpha s_3 \\ s_3 & s_4 + s_1 + \beta s_3 \end{bmatrix} = 1 + \delta \begin{bmatrix} s_1 & s_2' \\ s_3 & s_4' \end{bmatrix} = 5.$$

For the upper bound, consider the following five rank-one matrices:

$$\mathbf{n}_1 = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{n}_2 = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$\mathbf{n}_3 = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{n}_4 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix},$$

and

$$\mathbf{n}_5 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

Then

$$G_2(\mathbf{s}) = \tfrac{1}{2}(\mathbf{n}_2 + \mathbf{n}_3)s_1 + (\tfrac{1}{2}\mathbf{n}_3 - \tfrac{1}{2}\mathbf{n}_2 - \mathbf{n}_1)s_2 + (\tfrac{1}{2}\mathbf{n}_5 - \tfrac{1}{2}\mathbf{n}_4 + \mathbf{n}_1)s_3 + \tfrac{1}{2}(\mathbf{n}_4 + \mathbf{n}_5)s_4.$$

Therefore $\delta = 5$.

Before we attack the general case, we need the following lemma.

Lemma 2.4

If

$$H(\mathbf{s}) = \begin{bmatrix} s_1 & s_2 & s_3 \\ s_2 & s_1 & s_4 \\ s_5 & s_4 & s_1 \end{bmatrix}$$

then $\delta(H(\mathbf{s})) = 5$.

Proof. $\delta(H(\mathbf{s})) \geqslant 5$ because $\dim \mathbf{s} = 5$. For the upper bound consider the following five rank-one matrices:

$$\mathbf{n}_0 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad \mathbf{n}_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$\mathbf{n}_3 = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix}, \quad \mathbf{n}_4 = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 1 \\ -1 & 1 & 1 \end{bmatrix},$$

and

$$\mathbf{n}_5 = \begin{bmatrix} -1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & -1 \end{bmatrix}.$$

Then

$$H(\mathbf{s}) = (\mathbf{n}_1 + \mathbf{n}_2 + \tfrac{1}{2}\mathbf{n}_3 + \tfrac{1}{2}\mathbf{n}_4)s_1 + (\mathbf{n}_1 + \mathbf{n}_2 + \tfrac{1}{2}\mathbf{n}_3 + \tfrac{1}{2}\mathbf{n}_5)s_2 + \mathbf{n}_2 s_3 + (\mathbf{n}_1 + \mathbf{n}_2 + \tfrac{1}{2}\mathbf{n}_4 + \tfrac{1}{2}\mathbf{n}_5)s_4 + \mathbf{n}_1 s_5.$$

Therefore $\delta = 5$.

Theorem 2.5

The complexity of multiplying two $n \times n$ symmetric tridiagonal matrices is $6n - 7$, over any field of characteristic $\neq 2$.

Proof. The characteristic matrix is:

$$G_n(\mathbf{s}) = \begin{bmatrix}
s_1 & s_2 & & & & & & & & & \\
s_4 & s_1+s_5 & s_2 & s_3 & & & & & & & \\
 & s_4 & s_5 & s_6 & & & & & & & \\
 & s_8 & s_9 & s_5+s_{10} & s_6 & s_7 & & & & & \\
 & & & s_9 & \cdot & \cdot & & & & & \\
 & & & s_{13} & \cdot & \cdot & & & & & \\
 & & & & & & \ddots & & & & \\
 & & & & & & & s_{5n-15}+s_{5n-10} & s_{5n-14} & s_{5n-13} & \\
 & & & & & & & s_{5n-11} & s_{5n-10} & s_{5n-9} & \\
 & & & & & & & s_{5n-8} & s_{5n-7} & s_{5n-10}+s_{5n-6} & s_{5n-9} \\
 & & & & & & & & & s_{5n-7} & s_{5n-6}
\end{bmatrix}.$$

$$n \text{ blocks} = \begin{cases} 2 & 2 \times 2 \text{ blocks and} \\ n-2 & 3 \times 3 \text{ blocks.} \end{cases}$$

Let

$$H_n(\mathbf{s}) = \begin{bmatrix}
s_1 & s_2 & & & & & & & & & \\
s_2 & s_1+s_5 & s_4 & s_3 & & & & & & & \\
 & s_4 & s_5 & s_6 & & & & & & & \\
 & s_8 & s_6 & s_5+s_{10} & s_9 & s_7 & & & & & \\
 & & & s_9 & \cdot & \cdot & & & & & \\
 & & & s_{13} & \cdot & \cdot & & & & & \\
 & & & & & & \ddots & & & & \\
 & & & & & & & s_{5n-15}+s_{5n-10} & s_{5n-11} & s_{5n-13} & \\
 & & & & & & & s_{5n-11} & s_{5n-10} & s_{5n-9} & \\
 & & & & & & & s_{5n-8} & s_{5n-9} & s_{5n-10}+s_{5n-6} & s_{5n-7} \\
 & & & & & & & & & s_{5n-7} & s_{5n-6}
\end{bmatrix},$$

and let

$$K_n(\mathbf{s}) = \begin{bmatrix}
0 & 0 & & & & & & & & & \\
s_4-s_2 & 0 & s_2-s_4 & 0 & & & & & & & \\
 & 0 & 0 & 0 & & & & & & & \\
 & 0 & s_9-s_6 & 0 & s_6-s_9 & 0 & & & & & \\
 & & & 0 & 0 & 0 & & & & & \\
 & & & 0 & s_{14}-s_{11} & 0 & & & & & \\
 & & & & & & \ddots & & & & \\
 & & & & & & & 0 & s_{5n-14}-s_{5n-11} & 0 & \\
 & & & & & & & 0 & 0 & 0 & \\
 & & & & & & & 0 & s_{5n-7}-s_{5n-9} & 0 & s_{5n-9}-s_{5n-7} \\
 & & & & & & & & & 0 & 0
\end{bmatrix}.$$

Then $G_n(\mathbf{s}) = H_n(\mathbf{s}) + K_n(\mathbf{s})$. $K_n(\mathbf{s})$ has $n-1$ nonzero rows, each of them can be generated by one rank-one matrix. Therefore $\delta(K_n(\mathbf{s})) = n - 1$.

On the other hand,

$$\delta(H_n(\mathbf{s})) \leqslant \delta\begin{bmatrix} s_1 & s_2 \\ s_2 & s_1 \end{bmatrix} + \delta\begin{bmatrix} s_5 & s_4 & s_3 \\ s_4 & s_5 & s_6 \\ s_8 & s_6 & s_5 \end{bmatrix} + \cdots + \delta\begin{bmatrix} s_{5n-10} & s_{5n-11} & s_{5n-13} \\ s_{5n-11} & s_{5n-10} & s_{5n-9} \\ s_{5n-8} & s_{5n-9} & s_{5n-10} \end{bmatrix} + \begin{bmatrix} s_{5n-6} & s_{5n-7} \\ s_{5n-7} & s_{5n-6} \end{bmatrix}.$$

$$\delta\begin{bmatrix} s_1 & s_2 \\ s_2 & s_1 \end{bmatrix} = 2, \quad \text{in fact,} \quad \begin{bmatrix} s_1 & s_2 \\ s_2 & s_1 \end{bmatrix} = \tfrac{1}{2}(n_1 + n_2)s_1 + \tfrac{1}{2}(n_1 - n_2)s_2,$$

where

$$n_1 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad n_2 = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

Similarly,

$$\delta\begin{bmatrix} s_{5n-6} & s_{5n-7} \\ s_{5n-7} & s_{5n-6} \end{bmatrix} = 2.$$

By Lemma 2.4, the complexity of each of the remaining 3×3 blocks is equal to 5. Therefore, $\delta(H_n(\mathbf{s})) \leqslant 4 + 5(n-2) = 5n - 6$. Hence $\delta(G_n(\mathbf{s})) \leqslant n - 1 + 5n - 6 = 6n - 7$.

For the lower bound, note that

$$G_n(\mathbf{s}) = \left[\begin{array}{ccc|cc} & & & 0 & 0 \\ & & & \vdots & \vdots \\ & & & 0 & 0 \\ & G_{n-1}(\bar{\mathbf{s}}) & & s_{5n-13} & 0 \\ & & & s_{5n-9} & 0 \\ \hline 0 \ldots 0 & s_{5n-8} & s_{5n-7} & s_{5n-10} + s_{5n-6} & s_{5n-9} \\ 0 \ldots 0 & 0 & 0 & s_{5n-7} & s_{5n-6} \end{array}\right],$$

where $G_{n-1}(\bar{\mathbf{s}})$ is the characteristic matrix of multiplying two $(n-1) \times (n-1)$ symmetric tri-diagonal matrices, and

$$\bar{\mathbf{s}} = \mathbf{s} - \{s_{5n-13}, s_{5n-9}, s_{5n-8}, s_{5n-7}, s_{5n-6}\}.$$

Add linear combinations of the last column to the remaining ones, and use Theorem 1.1, then delete the last row, we get

$$\delta(G_n(\mathbf{s})) \geqslant 1 + \delta\left[\begin{array}{ccc|c} & & & 0 \\ & & & \vdots \\ & & & 0 \\ & G_{n-1}(\bar{\mathbf{s}}) & & s_{5n-13} \\ & & & s_{5n-9} \\ \hline * \ldots * & s_{5n-8} + \alpha_3 s_{5n-9} & s_{5n-7} + \alpha_2 s_{5n-9} & s_{5n-10} + s_{5n-6} + \alpha_1 s_{5n-9} \end{array}\right],$$

Replace $s_{5n-6}, s_{5n-7}, s_{5n-8}, s_{5n-9}$ and s_{5n-13} by linear combinations of the remaining ones, then delete the last row and last column, we get:

$$\delta(G_n(\mathbf{s})) \geqslant 6 + \delta(G_{n-1}(\bar{\mathbf{s}})),$$

and this is true $\forall n \geqslant 2$. Therefore,

$$\delta(G_n(\mathbf{s})) \geqslant 6(n-2) + \delta(G_2(\mathbf{s})).$$

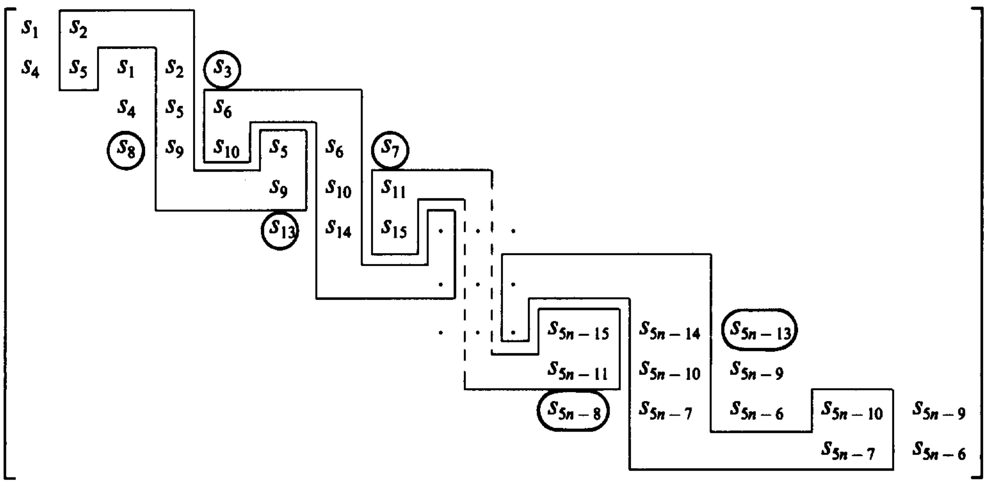
By Lemma 2.3, $\delta(G_2(\mathbf{s})) = 5$, hence, $\delta(G_n(\mathbf{s})) \geqslant 6n - 7$.

Thirdly, let $\mathbf{Y}$ be a tridiagonal matrix, i.e.

$$\mathbf{Y}=\begin{bmatrix} y_1 & y_2 & & & & \\ y_3 & y_4 & y_5 & & & \\ & y_6 & y_7 & y_8 & & \\ & & \ddots & \ddots & \ddots & \\ & & & y_{3n-6} & y_{3n-5} & y_{3n-4} \\ & & & & y_{3n-3} & y_{3n-2} \end{bmatrix}.$$

The characteristic matrix of multiplying the symmetric tridiagonal matrix $\mathbf{X}$ by $\mathbf{Y}$ is

$G_n(\mathbf{s}) =$



Lemma 2.6

If

$$\mathbf{H}=\begin{bmatrix} s_1 & \\ s_2 & s_1 \\ & s_2 \end{bmatrix}$$

then $\delta(\mathbf{H}) = 3$.

Proof. Notice that $\mathbf{H}=(\mathbf{n}_1+\mathbf{n}_2)s_1+(\mathbf{n}_2+\mathbf{n}_3)s_2$, where

$$\mathbf{n}_1=\begin{bmatrix} 1 & 0 \\ -1 & 0 \\ 0 & 0 \end{bmatrix},\quad \mathbf{n}_2=\begin{bmatrix} 0 & 0 \\ 1 & 1 \\ 0 & 0 \end{bmatrix}\quad \text{and}\quad \mathbf{n}_3=\begin{bmatrix} 0 & 0 \\ 0 & -1 \\ 0 & 1 \end{bmatrix}.$$

Therefore, $\delta(\mathbf{H}) \leqslant 3$. On the other hand, $\delta(\mathbf{H}) \geqslant 3$ because row rank $(\mathbf{H}) = 3$.

Lemma 2.7

If

$$\mathbf{H}=\begin{bmatrix} s_1 & & \\ s_2 & s_1 & \\ & s_2 & \\ & s_3 & s_2 \\ & & s_3 \end{bmatrix}$$

then $\delta(\mathbf{H}) = 5$.

Proof. Consider the following five rank-one matrices:

$$\mathbf{n}_1 = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{n}_2 = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$\mathbf{n}_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{n}_4 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

and

$$\mathbf{n}_5 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Then $H = (\mathbf{n}_1 + \mathbf{n}_2)s_1 + (\mathbf{n}_2 + \mathbf{n}_3 + \mathbf{n}_4)s_2 + (\mathbf{n}_4 + \mathbf{n}_5)s_3$. Therefore $\delta(\mathbf{H}) \leqslant 5$. But row rank $(\mathbf{H}) = 5$, therefore $\delta(\mathbf{H}) = 5$.

Theorem 2.8

The complexity of multiplying a symmetric tridiagonal matrix by a tridiagonal matrix is $7n - 8$ over any ring.

Proof. For the upper bound, notice that $G_n(\mathbf{s})$ can be decomposed into:

(a) $n - 2$ blocks of the form

$$\begin{bmatrix} s_2 & & \\ s_5 & s_2 & \\ & s_5 & \\ & s_9 & s_5 \\ & & s_9 \end{bmatrix},$$

(b) the 2 blocks

$$\begin{bmatrix} s_1 & \\ s_4 & s_1 \\ & s_4 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} s_{5n-9} & \\ s_{5n-6} & s_{5n-9} \\ & s_{5n-6} \end{bmatrix},$$

(c) the $2(n-2)$ 1×1 blocks

$$s_3, s_7, s_8, \ldots, s_{5n-13}, s_{5n-12}, s_{5n-8}.$$

Using the previous two lemmas, we get

$$\delta(G_n(\mathbf{s})) \leqslant 5(n-2) + 6 + 2(n-2) = 7n - 8.$$

For the lower bound, notice that $G_{n-1}(\mathbf{s})$ is obtained from $G_n(\mathbf{s})$ by deleting the last two rows and the last three columns:

$$G_n(\mathbf{s}) = \left[\begin{array}{cc|ccc} \multicolumn{2}{c|}{G_{n-1}(\mathbf{s})} & s_{5n-13} & & \\ & & s_{5n-9} & & \\ \hline s_{5n-8} & s_{5n-7} & s_{5n-6} & s_{5n-10} & s_{5n-9} \\ & & & s_{5n-7} & s_{5n-6} \end{array}\right],$$

Delete the last two columns and add linear combinations to the remaining ones, we get

$$\delta(G_n(\mathbf{s})) \geqslant 2 + \delta(\bar{G}_n(\mathbf{s})).$$

Replace s_{5n-13}, s_{5n-9}, s_{5n-8}, s_{5n-7} and s_{5n-6} by linear combinations of the remaining indeterminates, then delete the last two rows and the last column of $\bar{G}_n(\mathbf{s})$, we get

$$\delta(G_n(\mathbf{s})) \geqslant 7 + \delta(G_{n-1}(\mathbf{s}))$$

and this is true $\forall n \geqslant 3$. Hence $\delta(G_n(\mathbf{s})) \geqslant 7(n-2) + \delta(G_2(\mathbf{s}))$, where

$$G_2(\mathbf{s}) = \begin{bmatrix} s_1 & s_2 & & \\ s_3 & s_4 & s_1 & s_2 \\ & & s_3 & s_4 \end{bmatrix}.$$

But

$$\delta(G_2(\mathbf{s})) \geqslant 2 + \delta \begin{bmatrix} s_1 & s_2 \\ s_3 + \alpha_1 s_1 + \beta_1 s_2 & s_4 + \alpha_2 s_1 + \beta_2 s_2 \\ \alpha_1 s_3 + \beta_1 s_4 & \alpha_2 s_3 + \beta_2 s_4 \end{bmatrix} \geqslant 2 + \delta \begin{bmatrix} s_1 & s_2 \\ s_3' & s_4' \end{bmatrix} = 6.$$

Therefore

$$\delta(G_n(\mathbf{s})) \geqslant 7(n-2) + 6 = 7n - 8.$$

3. TRIDIAGONAL MATRIX MULTIPLICATION PROBLEM

Let $\mathbf{X}$ be a tridiagonal matrix, i.e.

$$\mathbf{X} = \begin{bmatrix} x_1 & x_3 & & & & \\ x_2 & x_4 & x_6 & & & \\ & x_5 & x_7 & x_9 & & \\ & & \cdot & \cdot & \cdot & \\ & & & \cdot & \cdot & \cdot \\ & & & & \cdot & \cdot & \cdot \\ & & & & & x_{3n-5} & x_{3n-3} \\ & & & & & x_{3n-4} & x_{3n-2} \end{bmatrix}.$$

First consider the problem of multiplying $\mathbf{X}$ by an arbitrary $n \times n$ matrix. Let us start with the case when $n = 4$. The characteristic matrix is

$$G_4(\mathbf{s}) = \left[\begin{array}{cccccccccccccccc}
s_1 & s_2 & s_3 & s_4 & & & & & & & & & & & & \\
s_5 & s_6 & s_7 & s_8 & & & & & & & & & & & & \\
& & & & s_1 & s_2 & s_3 & s_4 & & & & & & & & \\
& & & & s_5 & s_6 & s_7 & s_8 & & & & & & & & \\
& & & & s_9 & s_{10} & s_{11} & s_{12} & & & & & & & & \\
& & & & & & & & s_5 & s_6 & s_7 & s_8 & & & & \\
& & & & & & & & s_9 & s_{10} & s_{11} & s_{12} & & & & \\
& & & & & & & & s_{13} & s_{14} & s_{15} & s_{16} & & & & \\
& & & & & & & & & & & & s_9 & s_{10} & s_{11} & s_{12} \\
& & & & & & & & & & & & s_{13} & s_{14} & s_{15} & s_{16}
\end{array}\right],$$

Lemma 3.1

If

$$H(\mathbf{s}) = \begin{bmatrix} 0 & s_2 & & \\ s_3 & s_4 & & \\ & & s_1 & s_2 \\ & & s_3 & s_4 \end{bmatrix}$$

then $\delta(H(\mathbf{s})) = 6$, over any ring.

Proof. Consider the following rank-one matrices:

$$\mathbf{n}_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{n}_2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad \mathbf{n}_3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\mathbf{n}_4 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & -1 \\ 0 & 1 & 0 & 1 \end{bmatrix}, \quad \mathbf{n}_5 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix}, \quad \mathbf{n}_6 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 1 & -1 & 1 \\ 0 & -1 & 1 & -1 \end{bmatrix}.$$

Then

$$H(\mathbf{s}) = \mathbf{n}_1 s_1 + (\mathbf{n}_1 + \mathbf{n}_2 + \mathbf{n}_5 + \mathbf{n}_6)s_2 + (\mathbf{n}_1 + \mathbf{n}_3 + \mathbf{n}_4 + \mathbf{n}_6)s_3 + (\mathbf{n}_1 + \mathbf{n}_4 + \mathbf{n}_5 + \mathbf{n}_6)\, s_4.$$

Therefore $\delta \leqslant 6$. On the other hand,

$$\delta \geqslant 2 + \delta \begin{bmatrix} s_1 & s_2 \\ s_3 & s_4 \end{bmatrix} = 6.$$

Lemma 3.2

If

$$\bar{H}(\mathbf{s}) = \begin{bmatrix} s_3 & s_4 + s_2 & & \\ & s_2 & s_1 & s_2 \\ & -s_2 & s_3 & s_4 \end{bmatrix}$$

then $\delta(\bar{H}(\mathbf{s})) = 5$.

Proof. From the proof of Lemma 3.1, one can see that

$$\delta(\bar{H}(\mathbf{s})) = \delta(H(\mathbf{s}) - \mathbf{n}_2 s_2).$$

But

$$H(\mathbf{s}) - \mathbf{n}_2 s_2 = \mathbf{n}_1 s_1 + (\mathbf{n}_1 + \mathbf{n}_5 + \mathbf{n}_6) s_2 + (\mathbf{n}_1 + \mathbf{n}_3 + \mathbf{n}_4 + \mathbf{n}_6) s_3 + (\mathbf{n}_1 + \mathbf{n}_4 + \mathbf{n}_5 + \mathbf{n}_6) s_4.$$

Therefore, $\delta(\bar{H}(\mathbf{s})) = 5$.

Lemma 3.3

If

$$K(\mathbf{s}) = \begin{bmatrix} s_1 & s_2 & & & & & \\ s_3 & s_4 & & & & & \\ & & s_1 & s_2 & & & \\ & & & s_4 & & & \\ & & & s_6 & s_7 & & \\ & & & & & s_4 & s_5 \\ & & & & & s_6 & s_7 \end{bmatrix}$$

then $\delta(K(\mathbf{s})) = 11$.

Proof.

$$\delta \leqslant 1 + \delta \left[\begin{array}{cccc|ccc} s_1 & s_2 & & -s_4 & & & \\ s_3 & s_4 & & s_4 & & & \\ & & s_1 & s_2 + s_4 & & & \\ \hline & & & s_6 + s_4 & s_7 & & \\ & & & s_4 & & s_4 & s_5 \\ & & & -s_4 & & s_6 & s_7 \end{array}\right].$$

By the previous lemma, each block can be generated by five rank-one matrices. Therefore, $\delta \leqslant 11$.

Using Theorem 1.1, one can easily show that

$$\delta \geqslant 3 + \delta \begin{bmatrix} s_1 & s_2 & & \\ s_3 & s_4 & & \\ & & s_4 & s_5 \\ & & s_6 & s_7 \end{bmatrix}.$$

Replacing s_1, s_2, s_3, s_5, s_6, s_7 by multiples of s_4 and using Theorem 1.1, we get

$$\delta \geqslant 9 + \delta \begin{bmatrix} * & * & & \\ * & s_4 & & \\ & & s_4 & * \\ & & * & * \end{bmatrix} \geqslant 11.$$

Theorem 3.4

The complexity of multiplying a 4×4 tridiagonal matrix by an arbitrary 4×4 matrix satisfies

$$\delta(G_4(\mathbf{s})) \leqslant 34.$$

Proof. It is enough to notice that $G_4(\mathbf{s})$ can be decomposed into two blocks of the same form as in Lemma 3.3, and two blocks of the same form as in Lemma 3.1.

Corollary 3.4.1

If $n = 4m$ then the complexity of multiplying a 4×4 tridiagonal matrix by an arbitrary $4 \times n$ matrix satisfies

$$\delta(G_{4n}(\mathbf{s})) \leqslant \tfrac{17}{2} n.$$

Corollary 3.4.2

Assume that $n = 4m$ and let

$$\bar{\mathbf{X}} = \left[\begin{array}{cccc|cccc|c}
x_1 & x_3 & & & & & & & \\
x_2 & x_4 & x_6 & & & & & & \\
 & x_5 & x_7 & x_9 & & & & & \\
 & & x_8 & x_{10} & 0 & & & & \\
\hline
 & & & 0 & x_{13} & x_{15} & & & \\
 & & & & x_{14} & x_{16} & x_{18} & & \\
 & & & & & x_{17} & x_{19} & x_{21} & \\
 & & & & & & x_{20} & x_{22} & 0 \\
\hline
 & & & & & & & 0 & \ddots
\end{array}\right].$$

Then the complexity of multiplying $\bar{\mathbf{X}}$ by an arbitrary $n \times n$ matrix satisfies

$$\bar{\delta} \leqslant \tfrac{17}{8} n^2.$$

Corollary 3.4.3

If $n = 4m$ then the complexity of multiplying an $n \times n$ tridiagonal matrix by an arbitrary $n \times n$ matrix satisfies

$$\delta(G_n(\mathbf{s})) \leqslant \tfrac{21}{8} n^2 - 2n.$$

Corollary 3.4.4

$$\delta(G_n(\mathbf{s})) \leqslant \tfrac{21}{8} n^2 + 0(n).$$

Remark 1

By Theorem 2.2, we have

$$\delta(G_n(\mathbf{s})) \geqslant 2n^2 - n,$$

over any ring.

Remark 2

One can easily show that

$$\delta(G_n(\mathbf{s})) \geqslant 2n^2 - \left\lfloor \frac{n}{2} \right\rfloor,$$

over $\mathscr{Z}$ or $\mathscr{Z}_2$.

Secondly, consider now the problem of multiplying two $n \times n$ tridiagonal matrices. The characteristic matrix is:

$$H_n(\mathbf{s}) = \begin{bmatrix} s_1 & s_2 & & & & & & & & & & & \\ s_4 & s_5 & & & & & & & & & & & \\ & & s_1 & s_2 & s_3 & & & & & & & & \\ & & s_4 & s_5 & s_6 & & & & & & & & \\ & & s_8 & s_9 & s_{10} & & & & & & & & \\ & & & & s_5 & s_6 & s_7 & & & & & & \\ & & & & s_9 & s_{10} & s_{11} & & & & & & \\ & & & & s_{13} & s_{14} & s_{15} & & & & & & \\ & & & & & & & \ddots & & & & & \\ & & & & & & & s_{5n-15} & s_{5n-14} & s_{5n-13} & & \\ & & & & & & & s_{5n-11} & s_{5n-10} & s_{5n-9} & & \\ & & & & & & & s_{5n-8} & s_{5n-7} & s_{5n-6} & & \\ & & & & & & & & & & s_{5n-10} & s_{5n-9} \\ & & & & & & & & & & s_{5n-7} & s_{5n-6} \end{bmatrix}.$$

By Lemma 3.1, each of the $n-1$ blocks can be generated by six rank-one matrices. The remaining terms can be generated by $2n-3$ rank-one matrices. Therefore,

$$\delta(H_n(\mathbf{s})) \leqslant 6(n-1) + 2n - 3 = 8n - 9.$$

Using the same arguments as in the proof of Theorem 2.8, one can show that

$$\delta(H_n(\mathbf{s})) \geqslant 7(n-2) + \delta \begin{bmatrix} s_1 & s_2 & & \\ s_4 & s_5 & & \\ & & s_1 & s_2 \\ & & s_4 & s_5 \end{bmatrix} = 7n - 7.$$

Let us summarize the above results in the following theorem.

Theorem 3.5

The complexity of multiplying two $n \times n$ tridiagonal matrices satisfies

$$7n - 7 \leqslant \delta(H_n(\mathbf{s})) \leqslant 8n - 9,$$

over any ring.

REFERENCES

1. R. Brockett and D. Dobkin, On the optimal evaluation of a set of bilinear forms. *Linear Algebra Applic.* **19**, 207–235 (1978).
2. J. Hopcroft and L. Kerr, On minimizing the number of multiplications necessary for matrix multiplication. *SIAM Jl appl. Math.* **20**, 30–36 (1971).