GAUSSIAN ELIMINATION WITH PIVOTING IS OPTIMAL *

Narayana Sadananda Kamath

TR 73-176

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University for the Degree of

Master of Science

July 1973

* This research has been supported in part by the National Science Foundation grants GJ-28176 and GJ-27528.

BIOGRAPHICAL SKETCH

The author was born on January 13, 1948 in Cochin, India. After completing preliminary education at T.D.High School, Cochin and Maharaja's College, Cochin he studied Electrical Engineering at Karnataka Regional Engineering College, Suratkal and graduated from the University of Mysore in 1970. He joined the graduate school at Cornell University in September 1971.

DEDICATION

To my parents

ACKNOWLEDGMENT

I thank Professor John E. Dennis for chairing my special committee, for suggesting the topic of this thesis and for rendering unusual encouragement. I also thank Professor Mark Eisner of Operations Research Department and Professor Ellis Horowitz for being on my special commi--ttee and Professor John E. Hopcroft for the discussions I had with him during the course of this work.

This work was supported by National Science Foudation under grants GJ-28176 and GJ-27528.

GAUSSIAN ELIMINATION WITH

PIVOTING IS OPTIMAL

by

Narayana Sadananda Kamath, M.S.

Cornell University, 1973

Abstract:

V. Strassen discovered that two matrices of order 2 could be multiplied using 7 multiplications and 18 additions of numbers and has shown that two matrices of order n could be multiplied in less than $4 \cdot 7\ n^{\log_2 7}$ operations, an operation being defined as a multiplication, division, subtraction or addition. Also he has shown that the classical Gaussian elimination is not optimal by giving an algorithm to compute the inverse of a nonsingular matrix with certain principal submatrices nonsingular in less than $5 \cdot 64\ n^{\log_2 7}$ operations. In other words, Strassen's algorithm provides no room for pivoting. J. Bunch and J. Hopcroft have got rid of the above anomaly and have shown how to obtain the triangular factorization of a permutation of a nonsingular matrix in less than $2 \cdot 44\ n^{\log_2 7}$ operations and the inverse in less than $6 \cdot 83\ n^{\log_2 7}$ operations. In this thesis it is shown, using the results of Strassen and Bunch and Hopcroft, that Gaussian elimination with pivoting is optimal in the sense that the bound for the number of operations required to do Gaussian elimination is the least for "sufficiently large" systems of equations. Also expressions are derived for the various coefficients in the bounds for the various procedures that arise in solving linear systems of equations with the general assumption that two matrices of order u could be multiplied in p multiplications and q additions of numbers.

CONTENTS

Gaussian Elimination with

Pivoting Is Optimal

CHAPTER I

1.1 INTRODUCTION

This work focuses on the solution of the linear system

$$AX = B \qquad (1.1.1)$$

where A is an n x n nonsingular matrix and X and B are matrices of compatible dimensions.

Gaussian elimination with partial pivoting solves (1.1.1) by decomposing PA = LU where P is a permutation matrix of order n, L is a lower triangular matrix of order n and U is a unit upper triangular matrix of order n and computing x as

$$H = L^{-1}PB$$

and (1.1.2)

$$X = U^{-1}H$$

A popular computational scheme to do Gaussian elimination is known as the Crout reduction and an algorithmic statement of the scheme is given on page 94 of Linear Algebra, Wilkinson and Reinsch, Springer-Verlag, 1971. Conditions for the existence and uniqueness of the triangular factorization can be found in many standard textbooks on matrix theory or linear algebra such as Matrix Theory, Franklin, J.N., Prentice-Hall, 1968.

It is well-known that solving (1.1.1) with B nxr and X nxr by Crout reduction takes $n(n-1)(n+1)/3 + n^2r$ long operations, a long operation being a multiplication or division. Also familiar is the fact that computing A^{-1} from its triangular factors and solving the resulting 2n triangular system of equations and computing $X = A^{-1}B$ takes $n(n-1)(2n+1)/2 + n^2r$ long operations. Thus it is obvious that Gaussian elimination saves $n(n-1)(4n+1)/6$ long operations as compared to inversion.

In [3], Klyuyev and Kokovkin-Shcherbak assert that Gaussian elimination is optimal if one restricts oneself to operations on rows and columns as a whole. Winograd [6] modifies the usual algorithms for matrix multiplication and inversion and for solving systems of linear equations trading roughly half of the multiplications for additions and subtractions. But still it takes $O(n^3)$ long operations to solve (1.1.1).

Strassen [5] discovered that two matrices of order 2 could be multiplied in 7 multiplications and 18 additions of numbers and showed how to compute the product of two matrices of order n in less than $4\cdot 7\, n^{\log_2 7}$ operations, an operation being a multiplication, division, addition or subtraction. He also showed how to compute the inverse of a nonsingular matrix of order n in less than $5\cdot 64\, n^{\log_2 7}$ operations. But his inversion algorithm fails if certain principal submatrices are nonsingular and the computational scheme admits no pivoting. J. Bunch and J. Hopcroft [1] have circumvented the anomaly and have shown that for any nonsingular matrix A, AP = LU can be computed in less than $2\cdot 44\, n^{\log_2 7}$ operations and $(AP)^{-1} = U^{-1}L^{-1}$ can be computed in fewer than $6\cdot 83\, n^{\log_2 7}$ operations.

In this thesis it is shown, using the algorithms due to Strassen and Bunch and Hopcroft that Gaussian elimination with pivot-

-ing is optimal in the sense that the bound for the number of operations required to do Gaussian elimination is the least for "sufficiently large" systems of equations.

But first we explore Strassen's results.

1.2 FAST MATRIX MULTIPLICATION AND INVERSION

We define algorithm $\alpha_{m,k}$ which multiplies two matrices of order $m2^k$ by induction on k: $\alpha_{m,o}$ is the usual algorithm for matrix multiplication requiring m^3 multiplications and $m^2(m-1)$ additions. The algorithm $\alpha_{m,k}$ being known, define $\alpha_{m,k+1}$ as follows:

If A and B are matrices of order $m2^{k+1}$, write

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} , \qquad B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} ,$$

$$AB = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} ,$$

where A_{ik}, B_{ik} and C_{ik} are matrices of order $m2^k$. Then compute

$$I = (A_{11} + A_{22})(B_{11} + B_{22})$$

$$II = (A_{21} + A_{22}) B_{11}$$

$$III = A_{11} (B_{12} - B_{22})$$

$$IV = A_{22} (-B_{11} + B_{21})$$

$$V = (A_{11} + A_{12}) B_{22}$$

$$VI = (-A_{11} + A_{21})(B_{11} + B_{12})$$

$$VII = (A_{12} - A_{22})\ (B_{21} + B_{22})$$

$$C_{11} = I + IV - V + VII$$

$$C_{21} = II + IV$$

$$C_{12} = III + V$$

$$C_{22} = I + III - II + VI$$

using $\alpha_{m,k}$ for multiplication and the usual algorithms for addition and subtraction of two matrices of order $m2^k$.

Note that in order to compute C it takes 7 multiplications and 18 additions and subtractions of matrices of order $m2^k$.

Lemma 1.2.1: The algorithm $\alpha_{m,k}$ computes the product of two matrices of order $m2^k$ with $m^3 7^k$ multiplications and $(5+m)m^2 7^k - 6(m2^k)^2$ additions and subtractions of numbers.

Proof: (by induction on k)

Assume true for k.

For k + 1, in forming I thru VII it requires 7 multiplications of matrices of order $m2^k$ and each multiplication takes $m^3 7^k$ multiplications of numbers. Hence it takes $7 \cdot m^3 7^k = m^3 7^{k+1}$ multiplications of numbers for k + 1. Note that for k = 1 and m = 1 it takes 7 multiplications and for k = o it takes m^3 multiplications in compliance with the definition of $\alpha_{m,o}$.

In forming I thru VII it takes $7\{(5+m)m^2 7^k - 6(m2^k)^2\}$ additions and subtractions of numbers in forming the seven products of matrices of order $m2^k$. In addition it takes $18(m2^k)^2$ additions and subtractions of numbers to form C_{ik} since each A_{ik}, B_{ik} has $(m2^k)^2$ elements. Hence it requires

$$7\{(5+m)m^2 7^k - 6(m2^k)^2\} + 18(m2^k)^2$$
$$= (5+m)m^2 7^{k+1} - 6(m2^{k+1})^2$$

additions and subtractions of numbers for multiplying two matrices of order $m2^{k+1}$. To complete the proof note that for $m = 1$, $k = 1$, it takes 18 additions and subtractions of numbers, and for $k = 0$ it takes $(5 + m)m^2 - 6m^2 = m^2(m - 1)$ additions and subtractions in compliance with the definition of $\alpha_{m,o}$. ∎

Lemma 1.2.2: The product of two matrices of order n may be computed in less than $4 \cdot 7\ n^{\log_2 7}$ operations.

Proof: Put

$$k = [\log n - 4]$$
$$m = [n2^{-k}] + 1 \qquad (1.2.1)$$

where $[\alpha]$ denotes the largest integer smaller than or equal to α.

Then $n \leq m2^k$.

Imbedding matrices A of order n, $m2^{k-1} < n \leq m2^k$, as

$$\begin{bmatrix} A & O \\ O & I_{m2^k-n} \end{bmatrix}$$

where I_j is the identity matrix of order j, our task reduces to that of estimating the number of operations for $\alpha_{m,k}$.

By lemma 1.2.1 this number is

$$M(n) = (5+m)m^2 7^k - 6(m2^k)^2 + m^3 7^k$$
$$= (5+2m)m^2 7^k - 6(m2^k)^2$$

$$< (5+2m)m^2 7^k$$

$$< (5+2(n2^{-k}+1))\ (n2^{-k}+1)\ 7^k$$

$$< (2n^3 2^{-3k} + 11n^2 2^{-2k} + 16n2^{-k} + 7)\ 7^k$$

$$< 2n^3(\frac{7}{8})^k + 11n^2(\frac{7}{4})^k + \frac{16}{16} n^2(\frac{7}{4})^k + \frac{7}{16^2} n^2(\frac{7}{4})^k ,$$

using (1.2.1)

$$< (2(\frac{8}{7})^{-k} \frac{n^3}{7^k} + (11+1+ \frac{7}{256})(\frac{4}{7})^{-k} \frac{n^2}{7^k}\ )\ 7^k$$

$$< (2(\frac{8}{7})^{\log_2 n-k} + 12{\cdot}03\ (\frac{4}{7})^{\log_2 n-k}\ )\ n^{\log_2 7}$$

(here we have used $n^3 = 8^{\log_2 n}$, $n^2 = 4^{\log_2 n}$, and $O(7^k) = O(n^{\log_2 7}) = O(7^{\log_2 n})$).

Letting $t = \log_2 n-k$ we have,

$$M(n) < (2(\frac{8}{7})^t + 12{\cdot}03(\frac{4}{7})^t)\ n^{\log_2 7}$$

Using $k = [\log_2 n - 4] \Rightarrow 4 \le t \le 5$ we have,

$$M(n) < \max_{4\le t\le 5} (2(\frac{8}{7})^t + 12{\cdot}03\ (\frac{4}{7})^t)\ n^{\log_2 7} \qquad (1.2.2)$$

$$< 4{\cdot}7\ n^{\log_2 7}$$

by a convexity argument. ∎

Next we define by induction on k, algorithms $\beta_{m,k}$ which inverts matrices of order $m2^k$; $\beta_{m,o}$ is the usual Gaussian elimination algorithm. Define $\beta_{m,k+1}$ from $\beta_{m,k}$ as follows:

If A is a matrix of order $m2^{k+1}$ to be inverted, write

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} ,$$

and

$$A^{-1} = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix},$$

where A_{ik}, C_{ik} are matrices of order $m2^k$. Then compute

$$I = A_{11}^{-1}$$

$$II = A_{21}A_{11}^{-1}$$

$$III = IA_{12}$$

$$IV = A_{21}III$$

$$V = IV - A_{22}$$

$$VI = V^{-1}$$

$$C_{12} = III \quad VI$$

$$C_{21} = VI \quad II$$

$$VII = III \quad C_{21}$$

$$C_{11} = I - VII$$

$$C_{22} = -VI$$

using $\alpha_{m,k}$ for multiplication and $\beta_{m,k}$ for inversion and the usual algorithms for addition and subtraction of two matrices of order $m2^k$.

Lemma 1.2.3: The algorithm $\beta_{m,k}$ computes the inverse of a matrix of order $m2^k$ with $m2^k$ divisions and less than or equal to $(6/5)m^3 7^k - m2^k$ multiplications and less than or eaual to $(6/5)\ (5 + m)m^2 7^k - 7(m2^k)^2$ additions and subtractions of numbers.

Proof: (by induction on k)

For $m = k = 1$ it takes 2 divisions, less than or equal to $(6/5)\ \ 7 - 2$, i.e. less than or equal to 6 multiplications and less than or equal to $(6/5)\cdot 6\cdot 7 - 7\cdot 4$, i.e. less than or equal to 22 additions and subtractions to do inversion of a matrix of order 2 which is true by observation.

Now assume true for k. For k + 1, we need only two inverses of matrices of order $m2^k$, i.e. $m2^{k+1}$ divisions. Also the number of multiplications required to find the inverse of a matrix of order $m2^{k+1}$ is

$$\leq 2\{\frac{6}{5} m^3 7^k - m2^k\} + 6m^3 7^k$$
$$\leq \frac{6}{5} m^3 7^{k+1} - m2^{k+1} \quad .$$

The number of additions and subtractions for inverting a matrix of order $m2^{k+1}$ is

$$\leq 2\{\frac{6}{5}(5+m) m^2 7^k - 7(m2^k)^2\}$$
$$+ 6\{(5+m) m^2 7^k - 6(m2^k)^2\}$$
$$+ 2(m2^k)^2$$
$$\leq \frac{6}{5}(5+m) m^2 7^{k+1} - 7(m2^{k+1})^2 \quad .$$

■

<u>Lemma 1.2.4</u>: The inverse of a matrix of order n may be computed with fewer than $5\cdot 64\ n^{\log_2 7}$ operations.

<u>Proof</u>: From lemma 1.2.2 and lemma 1.2.3 it follows that the inverse of a matrix or order n may be computed with less than $(6/5)\ (4\cdot 7)\ n^{\log_2 7}$, i.e. less than $5\cdot 64\ n^{\log_2 7}$ operations. ■

CHAPTER II

It is evident from the recursive partitioning nature of the above algorithms that the inversion algorithm fails if certain principal submatrices are singular.

For example, the algorithm fails for even as simple a matrix as

$$\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} .$$

In this chapter we explore the algorithms for finding the triangular factorization LU of a nonsingular matrix that overcomes the above mentioned difficulty, and then compute the inverse as $U^{-1}L^{-1}$ using Strassen's algorithm for fast matrix multiplication. Naturally, the algorithm involves pivoting (here the word pivoting is used to mean the interchanging of columns to make certain (sub)matrices nonsingular).

But before plunging into the details of the algorithm let us observe the following well-known facts.

Every principal submatrix in every reduced matrix is nonsingular if it is symmetric positive definite, strictly diagonally dominant, or irreducibly diagonally dominant. If A is nonsingular, then there exist permutation matrices P_1, P_2, Q_1, Q_2 such that AP_1, Q_1A, Q_2AP_2 have the above property.

2.1 THE BASIC ALGORITHM

The basic algorithm for triangular factorization introduces zeros after each iteration as follows:

$$
\begin{bmatrix}
x & x & x & x & x\ldots x \\
0 & x & x & x & x\ldots x \\
x & x & x & x & x\ldots x \\
x & x & x & x & x\ldots x \\
x & x & x & x & x\ldots x \\
\vdots & & & & \vdots \\
x & x & x & x & x\ldots x
\end{bmatrix},
\quad
\begin{bmatrix}
x & x & x & x & x & x\ldots x \\
0 & x & x & x & & x\ldots\ldots x \\
0 & 0 & x & x & & x\ldots\ldots x \\
0 & 0 & x & x & & x\ldots\ldots x \\
x & x & x & x & & x\ldots\ldots x \\
\vdots & & & & & \vdots \\
x & x & x & x & & x\ldots\ldots x
\end{bmatrix},
$$

$$
\begin{bmatrix}
x & x & x & x & x\ldots x \\
0 & x & x & x & x\ldots x \\
0 & 0 & x & x & x\ldots x \\
0 & 0 & 0 & x & x\ldots x \\
x & x & x & x & x\ldots x \\
\vdots & & & & \\
x & x & x & x & x\ldots x
\end{bmatrix}, \ldots,
\begin{bmatrix}
x & & & & & & \\
& x & & & \huge\times & & \\
& & x & & & & \\
& & & x & & & \\
& \huge 0 & & & x & & \\
& & & & & \ddots & \\
& & & & & & x
\end{bmatrix}.
$$

Formally, we describe the algorithm as follows:

For simplicity, let M be of order $n = 2^k$ with $\det M \neq 0$. Let $M^0 \equiv M$. We shall construct a sequence $P^1, P^2, \ldots, P^{n-1}$ of permutation matrices so that $M = LUP$, i.e. $MP^{-1} = LU$ where $P = P^1P^2\ldots P^{n-1}$ is a permutation matrix, $L \equiv L^1L^2\ldots L^{n-1}$ is a unit lower triangular matrix, U is an upper triangular matrix and $\det M = \det P \det U = \pm \prod_{i=1}^{n} u_{ii}$ Since $(P^i)^{-1} = P^i$ here, $P^{-1} = P^{n-1}\ldots P^2P^1$ and

$$
\begin{aligned}
M^{-1} &= P^{-1}U^{-1}L^{-1} \\
&= P^{n-1}\ldots P^2P^1U^{-1}(L^{n-1})^{-1}\ldots(L^2)^{-1}(L^1)^{-1}
\end{aligned}
$$

where $(L^i)^{-1} = 2I - L^i$.

We define the algorithm sequentially for $1 \leq i \leq n-1$ as follows:

Consider the binary representation of i as a string of 1 and 0. Let the positions in this string beginning from the right end be marked $0,1,\ldots,j,\ldots,k-1$. Let B_i be the set of integers j, $0 \leq j \leq k-1$ such that the j-th position in the string is occupied by a 1, i.e. let $B_i = \{j: i_j = 1,\ i = i_{k-1}2^{k-1} + \ldots + i_1 2^1 + i_0 2^0\}$.

Let $t = \max\{j : j \in B_i\}$, $s = \min\{j : j \in B_i\}$ and

$$r = \begin{cases} t & \text{if } s \neq t \\ t-1 & \text{if } s = t \end{cases} .$$

Then

$$M^{i-1} = \left[\begin{array}{c|c} M_{11}^{i-1} & M_{12}^{i-1} \\ \hline 0 & \\ \cline{1-1} & M_{22}^{i-1} \\ M_{21}^{i-1} & \end{array}\right] ,$$

where M_{11}^{i-1} is a nonsingular upper triangular matrix of order $i-1$, M_{12}^{i-1} is $(i - 1) \times (n - i + 1)$, 0 is the $(2^{r+1} - i + 1) \times (i - 1)$ zero matrix, M_{21}^{i-1} is $(n - 2^{r+1}) \times (i - 1)$, M_{22}^{i-1} is $(n - i + 1) \times (n - i + 1)$ and M^{i-1} is nonsingular.

Since $2^{r+1} - i + 1 > 0$ and M^{i-1} is nonsingular there exists a nonzero element in the first row of M_{22}^{i-1}. Hence there exists a permutation matrix P^i such that $N^i = M^{i-1}P^i$, $n_{ii} \neq 0$ and N^i can be written as

$$N^i = \left[\begin{array}{c|c|c} U^i & \multicolumn{2}{c}{V^i} \\ \hline & E^i & F^i \\ \cline{2-3} 0 & G^i & H^i \\ \hline X^i & \multicolumn{2}{c}{Y^i} \end{array}\right]$$

where U^i is $(i - 2^s) \times (i - 2^s)$, V^i is $(i - 2^s) \times (n - i + 2^s)$, E^i and G^i are $2^s \times 2^s$, F^i and H^i are $2^s \times (n - i)$, 0 is the $(2^r - 1 - i + 2^s) \times (i - 2^s)$ zero matrix, X^i is $(n - 2^{r+1}) \times (i - 2^s)$ and Y^i is $(n - i - 2^s) \times (N - i + 2^s)$. Further U^i and E^i are nonsingular upper triangular.

Let $Z^i = G^i(E^i)^{-1}$ and

$$L^i = \left[\begin{array}{c|c|c} I_{i-2s} & \multicolumn{2}{c}{0} \\ \hline & I_{2s} & \\ \cline{2-2} & Z^i & I_{n-1} \end{array}\right],$$

where I_j is the identity matrix of order j.

Define $M^i \equiv (L^i)^{-1}N^i$.

Then

$$M^i = \left[\begin{array}{c|c|c} U^i & \multicolumn{2}{c}{V^i} \\ \hline 0 & E^i & F^i \\ \cline{2-3} & 0 & J^i \\ \hline X^i & \multicolumn{2}{c}{Y^i} \end{array}\right],$$

where $J^i = H^i - Z^iF^i$.

At the last step $U \equiv M^{n-1}$ is nonsingular and upper triangular.

Example

$$M = \begin{bmatrix} 1 & 2 & 1 & 3 \\ 2 & 3 & 1 & 2 \\ 4 & 1 & 1 & 3 \\ 3 & 3 & 2 & 1 \end{bmatrix}$$

<u>i = 1</u>

$$M^{i-1} = M^0 \equiv M = \begin{bmatrix} 1 & 2 & 1 & 3 \\ 2 & 3 & 1 & 2 \\ 4 & 1 & 1 & 3 \\ 3 & 3 & 2 & 1 \end{bmatrix}$$

$$M_{11}^0 = M_{12}^0 = M_{21}^0 = \phi \, , \; M_{22}^0 = M^0 ,$$

$$P^i = P^1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

and

$$N^i = N^1 = M^0 P^1 = \begin{bmatrix} 1 & 2 & 1 & 3 \\ 2 & 3 & 1 & 2 \\ 4 & 1 & 1 & 3 \\ 3 & 3 & 2 & 1 \end{bmatrix}$$

with $U^1 = V^1 = X^1 = \phi$, $E^1 = [1]$, $F^1 = [\,2 \quad 1 \quad 3]$, $G^1 = [2]$, $H^1 = [\,3 \quad 1 \quad 2]$, and $Y^1 = \begin{bmatrix} 4 & 1 & 1 & 3 \\ 3 & 3 & 2 & 1 \end{bmatrix}$.

$$L^i = L^1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} , \quad (L^1)^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

and

$$M^i = M^1 = (L^1)^{-1} N^1$$

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

<u>i = 2</u>

$$M^{i-1} = M^1 = \begin{bmatrix} 1 & 2 & 1 & 3 \\ 0 & -1 & -1 & -4 \\ 4 & 1 & 1 & 3 \\ 3 & 3 & 2 & 1 \end{bmatrix},$$

$$M^1_{11} = [1], \quad M^1_{12} = [2 \quad 1 \quad 3], \quad M^1_{21} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}.$$

$$M^1_{22} = \begin{bmatrix} -1 & -1 & -4 \\ 1 & 1 & 3 \\ 3 & 2 & 1 \end{bmatrix}, \qquad P^i = P^2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

and

$$N^i = N^2 = M^1P^2 = \begin{bmatrix} 1 & 2 & 1 & 3 \\ 0 & -1 & -1 & -4 \\ 4 & 1 & 1 & 3 \\ 3 & 3 & 2 & 1 \end{bmatrix}$$

with $U^2 = V^2 = X^2 = Y^2 = \phi$, $E^2 = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}$, $F^2 = \begin{bmatrix} 1 & 3 \\ -1 & -4 \end{bmatrix}$

$G^2 = \begin{bmatrix} 4 & 1 \\ 3 & 3 \end{bmatrix}$ and $H^2 = \begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix}$.

$$L^i = L^2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 4 & 7 & 1 & 0 \\ 3 & 3 & 0 & 1 \end{bmatrix}, \qquad (L^2)^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -4 & -7 & 1 & 2 \\ -3 & -3 & 0 & 1 \end{bmatrix}$$

and

$$M^i = M^2 = (L^2)^{-1}N^2$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -4 & -7 & 1 & 0 \\ -3 & -3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 & 3 \\ 0 & -1 & -1 & -4 \\ 4 & 1 & 1 & 3 \\ 3 & 3 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & -1 & -1 & -4 \\ 0 & 0 & 4 & 19 \\ 0 & 0 & 2 & 4 \end{bmatrix}$$

<u>i = 3</u>

$$M^{i-1} = M^2 = \begin{bmatrix} 1 & 2 & 1 & 3 \\ 0 & -1 & -1 & -4 \\ 0 & 0 & 4 & 19 \\ 0 & 0 & 2 & 4 \end{bmatrix},$$

$$M_{11}^2 = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}, \qquad M_{12}^2 = \begin{bmatrix} 1 & 3 \\ -1 & -4 \end{bmatrix}$$

$$M_{22}^2 = \begin{bmatrix} 4 & 19 \\ 2 & 4 \end{bmatrix}, \qquad M_{21}^1 = \phi,$$

$$P^i = P^3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

and

$$N^i = N^3 = M^3P^2 = \begin{bmatrix} 1 & 2 & 1 & 3 \\ 0 & -1 & -1 & -4 \\ 0 & 0 & 4 & 19 \\ 0 & 0 & 2 & 4 \end{bmatrix}.$$

with $U^3 = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}$, $V^3 = \begin{bmatrix} 1 & 3 \\ -1 & -4 \end{bmatrix}$, $E^3 = [4]$, $F^3 = [19]$,

$G^3 = [2]$, $H^3 = [4]$ and $X^3 = Y^3 = \phi$.

$$L^i = L^3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{2}{4} & 1 \end{bmatrix} , \quad (L^3)^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{-2}{4} & 1 \end{bmatrix}$$

and

$$M^i = M^3 = (L^3)^{-1} N^3$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{-2}{4} & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 & 3 \\ 0 & -1 & -1 & -4 \\ 0 & 0 & 4 & 19 \\ 0 & 0 & 2 & 4 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 2 & 1 & 3 \\ 0 & -1 & -1 & -4 \\ 0 & 0 & 4 & 19 \\ 0 & 0 & 0 & \frac{-22}{4} \end{bmatrix}$$

Therefore we have

$$\begin{bmatrix} 1 & 2 & 1 & 3 \\ 2 & 3 & 1 & 2 \\ 4 & 1 & 1 & 3 \\ 3 & 3 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 4 & 7 & 1 & 0 \\ 3 & 3 & \frac{2}{4} & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 & 3 \\ 0 & -1 & -1 & -4 \\ 0 & 0 & 4 & 19 \\ 0 & 0 & 0 & \frac{-22}{4} \end{bmatrix}$$

2.2 OPERATION COUNT

Finding the permutation matrices P^i requires at most (n - i) comparisons, and if $p^i_{ii} = 0$ then permutation involves n element interchanges. Hence at most n(n-1)/2 comparisons and n(n-1) element interchanges are required to obtain M = LUP. The computation of M^{-1} would require at most n(n-1) element interchanges.

Let an operation be, as before, a multiplication, division, addition or subtraction. Let M(n), $M_T(n)$, and $I_T(n)$ be the number of operations required to multiply two n x n matrices, to multiply an n x n matrix by an upper triangular matrix, and to invert an n x n nonsingular upper triangular matrix.

From lemma 1.2.1,

$$\begin{aligned} M(2^k) &= (5+1)1^2 \cdot 7^k - 6(1\cdot 2^k)^2 + 1^3 \cdot 7^k \\ &< 7^{k+1} \text{ for } k \geq 1 \end{aligned} \tag{2.2.1}$$

Let $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ be a full matrix of order 2^k and let

$B = \begin{bmatrix} B_{11} & B_{12} \\ 0 & B_{22} \end{bmatrix}$ be an upper triangular matrix of order 2^k with

A_{ik}, B_{ik} of order 2^{k-1}.

Then we have

$$\begin{aligned} M_T(2^k) &= 4M_T(2^{k-1}) + 2M(2^{k-1}) + 2(2^{k-1})^2 \\ &< 2 \sum_{j=0}^{k-1} 4^j \, M(2^{k-j-1}) \\ &< (2)\ 7^k \sum_{j=0}^{k-1} (\tfrac{4}{7})^j, \text{ using (2.2.1)} \end{aligned} \tag{2.2.2}$$

$$< \left(\frac{14}{3}\right) 7^k \quad .$$

We have

$$A^{-1} = \begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1} A_{12} A_{22}^{-1} \\ 0 & A_{22}^{-1} \end{bmatrix}$$

and we have

$$\begin{aligned} I_T(2^k) &= 2\, I_T(2^{k-1}) + 2\, M_T(2^{k-1}) \\ &= 2 \sum_{j=0}^{k-1} 2^j\, M_T(2^{k-j-1}) \qquad \text{.} \qquad (2.2.3) \\ &< 2\left(\frac{14}{3}\right) \sum_{j=0}^{k-1} 2^j\, 7^{k-j-1} \\ &< 2\left(\frac{14}{3}\right)\left(\frac{2}{7}\right) 7^k \sum_{j=0}^{k-1} \left(\frac{2}{7}\right)^j \\ &< \left(\frac{28}{15}\right) 7^k \quad . \end{aligned}$$

To calculate the number of operations required to invert all the E^i, $1 \leq i \leq n-1$, note that the order of E^i follows the sequence 1; 2,1; 4,1,2,1; 8,1,2,1,4,1,2,1; 16,1,2,1,4,1,2,1,8,1,2,1,4,1,2,1; ...;2^{k-1},1,2,1,4,1,2,1,...,2^{k-2},...,1,2,1,4,1,2,1.

Therefore, by induction, it is easy to see that the number of operations to invert all the E^i, $1 \leq i \leq n-1$, is

$$= 2^{k-1} I_T(2^0) + 2^{k-2} I_T(2^1) + 2^{k-3} I_T(2^2) + \dots + 2^1 I_T(2^{k-2}) + 2^0 I_T(2^{k-1})$$

$$= 2^{k-1} \sum_{j=0}^{k-1} \frac{1}{2^j} I_T(2^j)$$

$$< 2^{k-1} \left(\frac{28}{15}\right) \sum_{j=0}^{k-1} \left(\frac{7}{2}\right)^j$$

$$< \left(\frac{28}{75}\right) \quad 7^k$$

Similarly forming all the multipliers Z^i, $1 \leq i \leq n-1$, requires

$$2^{k-1} \sum_{j=0}^{k-1} \frac{1}{2^j} \; M(2^j)$$

$$< \left(\frac{14}{15}\right) 7^k \qquad \text{operations.}$$

The number of operations required to form all the reduced matrices J^i, $1 \leq i \leq n-1$, neglecting terms of lower that account for subtractions in $H^i - Z^i F^i$, is

$$= \sum_{j=0}^{k-1} \left\{ \sum_{\ell=0}^{\frac{n-2^j}{2^{j+1}}} [2^k - 2(\ell+1)2^j] \right\} \frac{M(2^j)}{2^j}$$

$$\leq 2^{2(k-1)} \sum_{j=0}^{k-1} \frac{M(2^j)}{2^j}$$

$$< (\frac{7}{4})\ 2^{2k} \sum_{j=0}^{k-1} (\frac{7}{4})^j$$

$$< (\frac{7}{3})\ 7^k$$

Note that the sequence $L^1, L^2, \ldots, L^{n-1}$ has the structure

$$\begin{bmatrix} 1 & & & \bigcirc \\ x & \cdot & & \\ & \cdot & \cdot & \\ \bigcirc & & \cdot & 1 \end{bmatrix} , \quad \begin{bmatrix} 1 & & & \\ 0 & 1 & \bigcirc & \\ x & x & \cdot & \\ x & x & \bigcirc & 1 \end{bmatrix} , \ldots$$

and the product $L = L^1 L^2 \ldots L^{n-1}$ is just the superposition of L^i and at the maximum takes $O(n^2)$ operations. Therefore, for $n = 2^k$, triangular factorization takes

$$< (\frac{28}{75} + \frac{14}{15} + \frac{7}{3})\ 7^k$$

$$< (\frac{91}{25})\ 7^k \text{ operations}$$

$$< (3.64)\ n^{\log_2 7} \text{ operations} \quad .$$

Inverting U requires $< (\frac{28}{15})\ 7^k$ operations and $U^{-1}L^{-1}$ by computing $U^{-1}(L^{n-1})^{-1} \ldots (L^2)^{-1}(L^1)^{-1}$ requires

$$2^{2k} \sum_{j=0}^{k-1} \frac{M(2^j)}{2^{2j+1}}$$

$$< (\frac{7}{2})\ 2^{2k} \sum_{j=0}^{k-1} (\frac{7}{4})^j$$

$$< (\frac{14}{3})\ 7^k \text{ operations} \qquad (2.2.4)$$

On the other hand computing $L = (L^1+\ldots+L^{n-1}) - (n-2)I$ and inverting L and then computing $U^{-1}L^{-1}$ requires

$$< (\frac{28}{15})7^k + 2\sum_{j=0}^{k-1} 2^j M_T(2^{k-j-1}) + \sum_{j=0}^{k-1} 2^j M(2^{k-j-1})$$

$$< (\frac{77}{15})7^k \text{ operations} \qquad (2.2.5)$$

So we prefer the first scheme to compute $U^{-1}L^{-1}$. Hence for $n = 2^k$ inversion takes

$$< (\frac{91}{25} + \frac{28}{15} + \frac{14}{3})\ 7^k$$

$$< (\frac{763}{75})\ 7^k \text{ operations}$$

$$< 10\cdot 18\ n^{\log_2 7} \text{ operations} .$$

For arbitrary n, $2^k<n<2^{k+1}$, let

$$\mathcal{M} = \left[\begin{array}{c|c} M & O \\ \hline O & I_{2^{k+1}-n} \end{array}\right]$$

We can find the triangular factorization of a permutation of $\mathcal{M}$, and hence of a permutation of M by $<(3\cdot 64)7^{k+1}$, i.e. $<25\cdot 48\ n^{\log_2 7}$ operations, and the inverse of a permutation of $\mathcal{M}$, and hence of permutation of M by $<(10\cdot 18)7^{k+1}$, i.e. $<(71\cdot 22)\ n^{\log_2 7}$ operations.

2.3 OPTIMIZATION OF THE COEFFICIENTS IN OPERATIONS COUNT -- A MODIFIED ALGORITHM

We can modify the algorithm presented in 2.1 so that the coef-

-ficients of $n^{\log_2 7}$ is smaller in operation count. The idea is to find, for $n \leq m2^k$, suitable m and k such that the coefficients are the least possible. We consider two cases with $n = 2^r$ and n arbitrary.

Let $n = 2^r = m2^k$. Then $m = 2^{r-k} = 2^s$ and $M(2^r) < (5+2m)m^2 7^k = f(s)7^r$ where $f(s) = (5+2^{s+1})2^{2s}7^{-s}$. Now $\min_{0 \leq s \leq r} f(s) = f(3) = \frac{192}{49}$.

So we take $m = 2^3 = 8$ and $k = r - 3$ and use regular multiplication and inversion for 8 x 8 matrices.

Then $M(2^r) \leq (\frac{192}{49})7^r$ instead of 7^{r+1}. Hence each coefficient in 2·2 is multiplied by $\frac{1}{7}(\frac{192}{49})$. Thus triangular factorization requires $< \frac{1}{7}(\frac{192}{49})(3{\cdot}64)\, n^{\log_2 7}$, i.e. $(2{\cdot}04)\, n^{\log_2 7}$ operations, and inversion requires $< \frac{1}{7}(\frac{192}{49})(10{\cdot}18)\, n^{\log_2 7}$, i.e. $(5{\cdot}70)n^{\log_2 7}$ operations.

Now let n be arbitrary. Taking $k = [\log_2 n - 4]$ and $m = [n2^{-k}] + 1$, we have $n \leq m2^k$ and $(5+2m)m^2 7^k < 4{\cdot}7n^{\log_2 7}$.

From lemma 1.2.1 and (2.2.2) it follows that

$$M_T(m2^k) < 2(5+2m)m^2 7^k \sum_{j=1}^{k-1} \left(\frac{4}{7}\right)^j$$

$$< \frac{2}{3}(5+2m)m^2 7^k \quad . \qquad (2.3.1)$$

From (2.2.3) and (2.3.1) it follows that

$$I_T(m2^k) < 2 \sum_{j=1}^{k-1} 2^j \frac{2}{3} (5+2m)m^2 7^{k-j-1}$$

$$< \frac{4}{15}(5+2m)m^2 7^k \qquad (2.3.2)$$

We will now derive the number of operations for triangular factorization and inversion along the similar lines as in 2.2. Thus to compute all the E^i, $1 \leq i \leq n-1$, we need to do

$$2^{k-1}\sum_{j=0}^{k-1}\frac{1}{2^j} I_T(m2^j)$$

$$< 2^{k-1}\sum_{j=0}^{k-1}\frac{1}{2^j}\frac{4}{15}(5+2m)m^2 7^j, \text{ using (2.3.2)}$$

$$< \frac{4}{75}(5+2m)m^2 7^k \text{ operations} .$$

Similarly forming all the multipliers Z^i, $1\leq i\leq n-1$, requires

$$2^{k-1}\sum_{j=0}^{k-1}\frac{1}{2^j}M_T(m2^j)$$

$$< 2^{k-1}\sum_{j=0}^{k-1}\frac{2}{3}(5+2m)m^2\left(\frac{1}{2}\right)^j, \text{ using (2.3.1)}$$

$$< \frac{2}{15}(5+2m)m^2 7^k \text{ operations.}$$

Forming all the reduced matrices J^i, $1\leq i\leq n-1$ requires

$$\sum_{j=0}^{k-1}\left\{\sum_{\ell=0}^{\frac{n-2^r}{2^{r+1}}}[2^k - (2\ell + 1)2^j]\right\}\frac{M(m2^j)}{2^j}$$

$$\leq 2^{2(k-1)}\sum_{j=0}^{k-1}\frac{M(m2^j)}{2^{2j}}$$

$$< \frac{1}{3}(5+2m)m^2 7^k \text{ operations} .$$

By a similar argument as in 2.2, computing L takes $O(n^2)$ operations and is neglected.

Forming U^{-1} takes, from (2.3.2),

$$\frac{4}{15}(5+2m)m^2 7^k \text{ operations} ,$$

and forming $U^{-1}L^{-1}$ takes

$$2^{2k} \sum_{j=0}^{k-1} \frac{M(m2^j)}{2^{2j+1}}$$

$$< \frac{2}{3}(5+2m)\, m^2 7^k \quad \text{operations}$$

Therefore triangular factorization could be done in

$$< (\frac{4}{75} + \frac{2}{15} + \frac{1}{3}) \quad (5+2m)\, m^2 7^k$$

$$< (\frac{39}{75}) \quad (5+2m)\, m^2 7^k \quad \text{operations}$$

$$< (\frac{39}{75}) \quad (4 \cdot 7)\, n^{\log_2 7}$$

$$< 2 \cdot 44 \; n^{\log_2 7} \quad \text{operations}$$

and inversion could be done in

$$< (\frac{39}{75} + \frac{4}{15} + \frac{2}{3}) \quad (5+2m)\, m^2 7^k$$

$$< (\frac{109}{75}) \quad (5+2m)\, m^2 7^k$$

$$< (6 \cdot 83) \; n^{\log_2 7} \quad \text{operations} \quad .$$

2.4 GAUSSIAN ELIMINATION IS OPTIMAL

Crout reduction mentioned in 1.1 produces the factors L and U of PA and X is computed solving the two triangular systems

$$\begin{aligned} LH &= PB \\ UX &= H \end{aligned} \qquad (2.4.1)$$

Since L and U are nonsingular, Gaussian elimination is thus equivalent to computing X as

$$X = U^{-1}(L^{-1}PB) \quad . \qquad (2.4.2)$$

Inverting PA and computing X is equivalent to computing X as

$$X = (U^{-1}L^{-1})PB \tag{2.4.3}$$

Computing X as per (2.4.2) requires

$$< \left(\frac{39}{75} + 2\left(\frac{4}{15}\right) + 2\left(\frac{2}{3}\right)\right) (5+2m)m^2 7^k$$

$$< \left(\frac{175}{75}\right) (5+2m)m^2 7^k \quad \text{operations}$$

$$< (11 \cdot 21)\ n^{\log_2 7} \quad \text{operations, using lemma 1.2.2.}$$

Computing X as per (2.4.3) requires

$$< \left(\frac{109}{75} + 1\right)(5+2m)m^2 7^k \quad \text{operations}$$

$$< \frac{184}{75} (5+2m)m^2 7^k \quad \text{operations}$$

$$< 11 \cdot 53\ n^{\log_2 7} \quad \text{operations, using lemma 1.2.2.}$$

Thus Gaussian elimination with pivoting saves us

$$\frac{5}{75} \cdot (5+2m)m^2 7^k$$

$$= (0.313)\ n^{\log_2 7} \quad \text{operations, using lemma 1.2.2.}$$

and is optimal as compared to inversion, unless Strassen's bound on matrix multiplication is improved.

CHAPTER III

In the previous chapters we used 2 x 2 partitioning in developing algorithms for matrix multiplications and inversion. But in general it is possible to use u x u partitioning and then develop algorithms to do inversion and multiplication. It is not yet known whether for $u > 2$ there exist algorithms which are more economical then Strassen's algorithm. So far, to the author's knowledge, nobody has come up with a matrix multiplication algorithm or inversion algorithm which is better than Strassen's using partitioning. The purpose of this chapter is to derive expressions for the bounds on the number of operations required to do the precedures of Chapters I and II when u x u partitioning is used.

Suppose two matrices of order n could be multiplied in p multiplication of numbers and q additions of numbers.

3.1 MATRIX MULTIPLICATION

Consider two matrices of order $n = u^k$. As in the previous chapter it is easy to see by induction that the number of multiplications to compute the matrix product is p^k.

Let A(n) denote the number of additions and subtractions required to multiply two matrices of order $n = u^k$.

$$\text{Then } A(n) = q(\tfrac{n}{u})^2 + pA(\tfrac{n}{u})$$

$$= q(\tfrac{n}{u})^2 + p\,q(\tfrac{n}{u})^2(\tfrac{1}{u})^2 + p^2A(\tfrac{n}{u^2})$$

$$= \cdots$$

$$= \cdots$$

$$= \sum_{i=0}^{k-1} n^2 \; q(\tfrac{1}{u})^2 p^i (\tfrac{1}{u})^{2i}$$

$$= q \, \frac{u^{2k}}{u^2} \quad \frac{\frac{p^k}{u^{2k}} - 1}{\frac{p}{u^2} - 1} \; .$$

(here we have used $n^2 = u^{2k}$)

$$= \frac{q}{p-u^2}\, p^k - \frac{q}{p-u^2}\, u^{2k} \; .$$

This expression tallies with the one on page 427, The Art of Computer Programming, Vol.II, Knuth, D.E., Addison-Wesley, 1971, for the case $u = 2$.

Defining $\alpha_{m,k}$ and $\beta_{m,k}$ along the similar lines as in Chapter I, we have,

Lemma 3.1.1: The algorithm $\alpha_{m,k}$ computes the product of two matrices of order mu^k with m^3p^k multiplications and

$$(\frac{q}{p-u^2} - 1 + m)\, m^2 p^k - \frac{\alpha}{p-u^2} (mu^k)^2$$

additions and subtractions of numbers.

The proof by induction follows as in lemma 2.1.1.

Now we derive an expression of the form $C\, n^{\log_u p}$ for the number of operations $M(n)$ for matrix multiplication.

From lemma 3.1.1

$$M(n) = (\frac{q}{p-u^2} - 1 + m)\, m^2 p^k - \frac{q}{p-u^2} (mu^k)^2 + m^3 p^k$$

$$< (\frac{q}{p-u^2} - 1 + 2m)\, m^2 p^k \; .$$

Letting $m = [nu^{-k}] + 1$, where $[\alpha]$ denotes largest integer less than or equal to α,

$$M(n) < \left(\frac{q}{p-u^2} - 1 + 2(nu^{-k}+1)\right)(nu^{-k}+1)^2 p^k$$

$$< \left(\frac{-q}{p-u^2} + 1 + 2nu^{-k}\right)(n^2u^{-2k} + 2nu^{-k} + 1)p^k$$

$$< \{2n^3u^{-3k} + 2\left(\frac{q}{p-u^2} + 1\right)nu^{-k} + 2nu^{-k} + 4n^2u^{-2k}$$

$$+ \left(\frac{q}{p-u^2} + 1\right)n^2u^{-2k} + \frac{q}{p-u^2} + 1\}\, p^k$$

$$< \{2n^3u^{-3k} + \left(5 + \frac{q}{p-u^2}\right)n^2u^{-2k} + 2\left(\frac{q}{p-u^2} + 2\right)nu^{-k}$$

$$+ \frac{q}{p-u^2} + 1\}\, p^k \quad .$$

Letting $k = \log_u n - \log_u 2\left(\frac{q}{p-u^2} + 2\right)$, we have $2\left(\frac{q}{p-u^2} + 2\right)u^k \leq n$, and

$$M(n) < \{2n^3u^{-3k} + \left(5+\frac{q}{p-u^2}\right)n^2u^{-2k} + n^2u^{-2k}$$

$$+ \frac{\left(\frac{q}{p-u^2} + 1\right)}{4\left(\frac{q}{p-u^2} + 2\right)^2}\, n^2u^{-2k}\}p^k$$

$$< \{2n^3u^{-3k} + \frac{4\left(\frac{q}{p-u^2} + 2\right)^2\left(5 + \frac{q}{p-u^2} + 1\right)n^2u^{-2k}}{4\left(\frac{q}{p-u^2} + 2\right)^2}$$

$$+ \frac{\left(\frac{q}{p-u^2} + 1\right)}{4\left(\frac{q}{p-u^2} + 2\right)^2}\, n^2u^{-2k}\}\, p^k$$

$$< \{2n^3u^{-3k} + \{\frac{4(\frac{q^2}{(p-u^2)^2} + 4\frac{q}{p-u^2} + 4)(6 + \frac{q}{p-u^2}) + (\frac{q}{p-u^2} + 1)}{4(\frac{q^2}{p-u^2}) + 2)^2}\} n^2u^{-2k}\} p^k \quad .$$

i.e.,

$$M(n) < \{2n^3u^{-2k} + \{\frac{4(\frac{10q^2}{(p-u^2)^2} + \frac{q^3}{(p-u^2)^3}) + 96 + 112\frac{q}{p-u^2} + (\frac{q}{p-u^2} + 1)}{4(\frac{q}{p-u^2})^2}\} n^2u^{-2k}\} p^k$$

$$< \{2n^3u^{-3k} + \frac{[4(\frac{10q^2}{(p-u^2)^2} + \frac{q^3}{(p-u^2)^3}) + 97 + 113\frac{q}{p-u^2}]n^2u^{-2k}}{4(\frac{q}{p-u^2} + 2)^2}\}p^k$$

Letting $C_1 = \frac{[4(\frac{10q^2}{(p-u^2)^2} + \frac{q^3}{(p-u^2)^3} + 113\frac{q}{p-u^2} + 97]}{4(\frac{q}{p-u^2} + 2)^2}$

and $C_2 = 2$

and using the identities

$$n^3 = u^{3\log_u n}$$

$$n^2 = u^{2\log_u n}$$

and

$$O(p^k) = O(p^{\log_u n}) = O(n^{\log_u p}).$$

we have

$$M(n) < \left(C_1\left(\frac{u^2}{p}\right)^{\log_u n-k} + C_2\left(\frac{u^3}{p}\right)^{\log_u n-k}\right) n^{\log_u p} \quad .$$

Let $t = \log_u n - k$.

Then

$$M(n) < \left(\frac{u^2}{p}\right)^t (C_1 + C_2 u^t) n^{\log_u p} \qquad (3.1.1)$$

$$M(n) < \max_{\log_u 2\left(\frac{q}{p-u^2} + 2\right) \leq t \leq 1 + \log_u 2\left(\frac{q}{p-u^2} + 2\right)} \left(\frac{u^2}{p}\right)^t (C_1 + C_2 u^t) n^{\log_u p} \qquad (3.1.2)$$

Now we shall not invoke Strassen's convexity argument, for, without knowing the values of u, β and q it is not possible to find the maximum. However, the coefficient of $n^{\log_u p}$ in (3.1.2) is a smooth function of t in the neighborhood of the above interval.

A bound on the number of operations to invert a matrix of order n can be derived as in Chapter I. In fact it turns out to be less than

$$\left(\frac{p'}{\frac{q}{p-u^2} - 1}\right) M(n)$$

where p' is the number of multiplications required to invert a matrix of order u.

The basic algorithms for triangular factorization and inversion similar to the ones in Chapter II could easily be conceived. The algorithm would be expressed in terms of the expansions of integers modulo u. We do not go into the details of the computational scheme involved. The following section gives expressions for bounds on the number of operations required.

3.2 OPERATION COUNT

Following the notation in 2.2, we shall derive the expressions for $M_T(mu^k)$ and $I_T(mu^k)$.

First consider $M_T(mu^k)$.

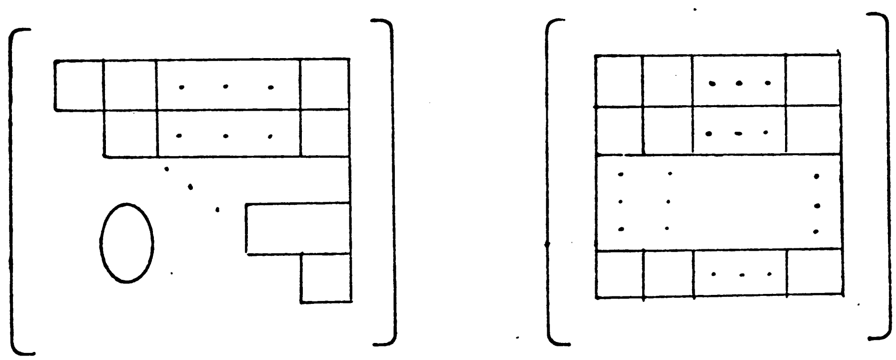


Figure 3.2.1

Let the two matrices, figure 3.2.1, be of order $\leq mu^k$ with submatrices, partitioned as shown, of order mu^{k-1}.

(i) Since each of the u upper triangular matrices of order $\leq mu^{k-1}$ along the diagonal multiplies u matrices of order mu^{k-1}, $M_T(mu^k)$ involves $u^2 M_T(mu^{k-1})$.

(ii) There are u(u-1)/2 off-diagonal full matrices of order mu^{k-1} in the first matrix and each of them multiplies u matrices of order $\leq mu^{k-1}$ in the second matrix. Hence it takes

$$\frac{u^2(u-1)}{2} M (mu^{k-1})$$

to do $M_T(mu^k)$.

(iii) Each submatrix has $(mu^{k-1})^2$ elements and hence the number of additions required is $\frac{u^2(u-1)}{2}(mu^{k-1})^2$.

Therefore

$$M_T(mu^k) = u^2 M_T(mu^{k-1}) + \frac{u^2(u-1)}{2} M(mu^{k-1}) + m^2u^{2k}(\frac{u-1}{2})$$

$$< \frac{u^2(u-1)}{2} \sum_{j=0}^{k-1} (u^2)^j M(mu^{k-j-1})$$

(here we have neglected the last term)

$$< \frac{u^2(u-1)}{2} \sum_{j=0}^{k-1} (u^2)^j (\frac{q}{p-u^2} - 1 + 2m)m^2p^{k-j-1}$$

(here we have used lemma 3.1.1)

$$< \frac{u^2(u-1)}{2} p^k \frac{1}{p} (\frac{q}{p-u^2} - 1 + 2m)m^2 \sum_{j=0}^{k-1} (\frac{u^2}{p})^j$$

$$< \frac{u^2(u-1)}{2(p-u^2)} (\frac{q}{p-u^2} - 1 + 2m)m^2p^k \quad .$$

Now we consider inversion of an upper triangular matrix of order $\leq mu^k$. The computational scheme is illustrated by the following figures:

$$\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} \\ 0 & A_{22}^{-1} \end{bmatrix}$$

$$\begin{bmatrix} A_{11} & A_{12} & A_{13} \\ 0 & A_{22} & A_{23} \\ 0 & 0 & A_{33} \end{bmatrix}^{-1} = \begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} & A_{11}(A_{12}A_{22}^{-1}A_{23}A_{33}^{-1}-A_{13}A_{33}^{-1}) \\ 0 & A_{22}^{-1} & -A_{22}^{-1}A_{23}A_{33}^{-1} \\ 0 & 0 & A_{33}^{-1} \end{bmatrix}$$

To compute the inverse of an upper triangular matrix it thus takes

(i) $uI_T(mu^{k-1})$ for the u matrices along the diagonal,

(ii) $2(u-1)\ M_T\ (mu^{k-1})$ for elements above the diagonal for the (u-1) matrices of order $\leq mu^{k-1} + 2(u-2)\ M_T\ (mu^{k-1})$ ... 2 above the diagonal ... (u-2) matrices ... + ... i.e. totally

$$\frac{2u(u-1)}{2} M_T(mu^{k-1})$$

for forming the off-diagonal elements, and

(iii) $(u-2)\ M\ (mu^{k-1})$ for forming elements 2 above the diagonal

$+ (u-3)\ M\ (mu^{k-1})$ ... 3 above ... + ...

i.e. totally

$$\frac{(u-1)(u-2)}{2}\ M\ (mu^{k-1})$$

for forming the off-diagonal elements.

Therefore we have

$$I_T(mu^k) = uI_T(mu^{k-1}) + u(u-1)M_T(mu^{k-1}) + \frac{(u-1)(u-2)}{2}\ M\ (mu^{k-1})$$

$$< u(u-1)\sum_{j=0}^{k-1} u^j M_t(mu^{k-j-1}) + \frac{(u-1)(u-2)}{2}\sum_{j=0}^{k-1} u^j M(mu^{k-j-1})$$

$$< u(u-1)\sum_{j=0}^{k-1} u^j \frac{u^2(u-1)}{2(p-u^2)}\left(\frac{q}{p-u^2} - 1 + 2m\right)m^2p^{k-j-1}$$

$$+\frac{(u-1)(u-2)}{2}\sum_{j=0}^{k-1} u^j \left(\frac{q}{p-u^2} - 1 + 2m\right)m^2p^{k-j-1}$$

$$< \left\{\frac{u^3(u-1)^2}{2(p-u^2)} + \frac{(u-1)(u-2)}{2}\right\}\left(\frac{q}{p-u^2} - 1 + 2m\right)\frac{m^2}{p-u^2}\quad p^k$$

$$< \left\{\frac{u^5 - 3u^4 + 4u^3 + (p-2)u^2 - 3pu + 2p}{2(p-u^2)}\right\}\left(\frac{q}{p-u^2} - 1 + 2m\right)\left(\frac{m^2}{p-u}\right)\ p^k\ .$$

CHECK: $m = 1,\ u = 2,\ p = 7,\ q = 18$

(i) $M_T(2^k) \quad \frac{4\ 1\ 7}{2\ 3}\ 7^k$

$$\frac{14}{3}7^k$$

(ii) $I_T(2^k) \quad \frac{7}{5}\ \frac{32-(3)(16)+(4)(8)+(5)(4)-(3)(7)(2)+(2)(7)}{(2)(3)}\ 7^k$

$$(\frac{28}{15})\ 7^k \qquad [\text{cf. } 2.2]$$

Now we derive expressions for bounds on the number of operations required for triangular factorization.

Inverting all the E^i, $1 \leq i \leq n-1$, following Chapter II, requires

$$u^{k-1} \sum_{j=0}^{k-1} \frac{1}{u^j} I_T(mu^j)$$

$$< u^{k-1} \ \frac{u^3(u-1)^2}{2(p-u^2)} + \frac{(u-1)(u-2)}{2} \ (\frac{q}{p-u^2} - 1 + 2m)\frac{m^2}{p-u} \sum_{j=0}^{k-1} (\frac{p}{u})^j$$

$$< \{\frac{u^3(u-1)^2}{2(p-u^2)} + \frac{(u-1)(u-2)}{2}\} \ (\frac{q}{p-u^2} - 1 + 2m)\frac{m^2}{(p-u)^2} \ p^k \text{ operations .}$$

(3.2.1)

Forming all the J^i, $1 \leq i \leq n-1$, requires

$$\sum_{j=0}^{k-1} \{ \sum_{\ell=0}^{\frac{n-u^j}{u^{j+1}}} [u^k - (2\ell+1)u^j]\} \frac{M(mu^j)}{u^j}$$

$$\leq u^{2(k-1)} \sum_{j=0}^{k-1} \frac{M(mu^j)}{u^{2j}}$$

$$\leq u^{2(k-1)} \sum_{j=0}^{k-1} (\frac{q}{p-u^2} - 1 + 2m)m^2 \ (\frac{p}{u^2})^j$$

$$\leq (\frac{q}{p-u^2} - 1 + 2m)\frac{m^2}{p-u^2} \ p^k \text{ operations} \qquad (3.2.2)$$

Forming all the z^i, $1 \le i \le n-1$, requires

$$< u^{k-1} \sum_{j=0}^{k-1} \frac{1}{n^j} M_T(mu^j)$$

$$< u^{k-1} \Sigma \frac{1}{n^j} \frac{u^2(u-1)}{2(p-u^2)} \left(\frac{q}{p-u^2} - 1 + 2m\right) m^2 p^j$$

$$< \frac{u^2(u-1)}{2(p-u^2)(p-u)} \left(\frac{q}{p-u^2} -1 + 2m\right) m^2 p^k \quad \text{operations} \qquad (3.2.3)$$

Therefore triangular factorization requires

$$< \left[\left\{\frac{u^3(u-1)^2}{2(p-u^2)} + \frac{(u-1)(u-2)}{2}\right\} \frac{1}{(p-u)^2} + \frac{1}{p-u^2} + \frac{u^2(u-1)}{2(p-u^2)(p-u)}\right]$$

$$\left(\frac{q}{p-u^2} - 1 + 2m\right) m^2 p^k \quad \text{operations}$$

$$< \frac{\left(\frac{q}{p-u^2} - 1 + 2m\right) m^2}{2(p-u^2)(p-u)^2} \left\{u^5 - 4u^4 + (p+5)u^3 - 7pu + 2p(p+1)\right\} p^k \quad \text{operations}$$

(3.2.4)

Finding $U^{-1}L^{-1}$, from U^{-1} and $L^1, L^2, \ldots, L^{n-1}$, requires

$$u^{2k} \sum_{j=0}^{k-1} \frac{M(mu^j)}{u^{2j+1}}$$

$$< u^{2k} \sum_{j=0}^{k-1} \frac{1}{u} \left(\frac{q}{p-u^2} - 1 + 2m\right) m^2 \left(\frac{p}{u^2}\right)^j$$

$$< \left(\frac{q}{p-u^2} - 1 + 2m\right) m^2 \frac{u}{p-u^2} \; p^k \quad \text{operations} \qquad (3.2.5)$$

Thus inversion requires

$$< \frac{\left(\frac{q}{p-u^2} -1+2m\right) m^2}{2(p-u^2)(p-u)^2} \left\{u^5-4u^4+(p+5)u^3-7pu+2p(p+1)\right\} p^k$$

$$+ \frac{\left(\frac{q}{p-u^2} - 1 + 2m\right)m^2}{2(p-u^2)(p-u)} \{u^5 - 3u^4 + 4u^3 + (p-2)u^2 - 3pu + 2p\} p^k$$

$$+ \left(\frac{q}{p-u^2} - 1 + 2m\right)m^2 \left(\frac{u}{p-u^2}\right) p^k \quad \text{operations}$$

$$< \frac{\left(\frac{q}{p-u^2} - 1 + 2m\right)m^2}{2(p-u^2)(p-u)^2} \{-u^6 + (p+4)u^5 - (3p+8)u^4 + (4p+9)u^3$$

$$+(p^2-3p)u^2 - (p^2+9p)u + 4p^2 + 2p\} p^k \quad \text{operations} \qquad (3.2.5)$$

CHECK: $u = 2$, $p = 7$, $q = 18$.

(i) Triangular Factorization

(3.2.4) gives $\frac{(5+2m)m^2}{2 \cdot 3 \cdot 5^2} \{ 32 - 64 - + 96 - 98 + 112\} 7^k$

$$= \frac{78}{(2)(3)(5^2)} (5 + 2m)m^2 \; 7^k$$

$$= \left(\frac{39}{75}\right) (5 + 2m) \; m^2 \; 7^k$$

(ii) Inversion

(3.2.5) gives $\frac{(5+2m)m^2}{(2)(3)(5^2)} \{-64 + (11)(32) - (29)(16)$

$$+ (37)(8) + (28)(4) - (112)(2) + 210 \}7^k$$

$$= \left(\frac{970 - 752}{(2) \; 3 \cdot 5^2}\right)(5 + 2m)m^2 7^k$$

$$= \left(\frac{109}{75}\right) (5+2m)m^2 \; 7^k$$

[cf. 2.3]

3.3 GAUSSIAN ELIMINATION VERSUS INVERSION TO SOLVE AX = B

To compute $X = U^{-1}(L^{-1}PB)$ it requires

$$< \frac{\left(\frac{q}{p-u^2} - 1 + 2m\right)m^2}{2(p-u^2)\ (p-u)^2} \{u^5 - 4u^4 + (p+5)u^3 - 7pu + 2p(p+1)\}$$

operations to do triangular factorization,

$$< \frac{2\left(\frac{q}{p-u^2} - 1 + 2m\right)m^2}{2(p-u^2)(p-u)^2} \{u^5 - 3u^4 + 4u^3 + (p-u)u^2 - 3pu + 2p\}\ p^k$$

operations to compute U^{-1} and L^{-1} and

$$< \frac{2u^2(u-1)}{2(p-u^2)} \left(\frac{q}{p-u^2} - 1 + 2m\right)m^2p^k$$

operations to do two multiplications of a full matrix by a triangular matrix.

Thus Gaussian elimination requires

$$< \frac{\left(\frac{q}{p-u^2} - 1 + 2m\right)m^2}{2(p-u^2)(p-u)^2} \{ u^5 - 4u^4 + (p+5)u^3 - 7pu + 2p(p+1)$$

$$+ 2(p-u)(u^5 - 3u^4 + 4u^3 + (p-2)u^2 - 3pu + 2p)$$

$$+ (2)(p-u)^2u^2(u-1) \}\ p^k$$

$$< \frac{\left(\frac{q}{p-u^2} - 1 + 2m\right)m^2}{2(p-u)^2(p-u)^2} \{-2u^6 + (9+2p)u^5 - (14+10p)u^4 + (2p^2+11p+9)u^3$$

$$+ 2pu^2 - (6p^2+11p)u + 6p^2 + 2p\ \}p^k \quad \text{operations} . \tag{3.3.1}$$

To compute $X = (U^{-1}L^{-1})PB$ it requires

$$< \frac{\left(\frac{q}{p-u^2} - 1 + 2m\right)m^2}{2(p-u^2)(p-u)^2} \{-u^6 + (p+4)u^5 - (3p+8)u^4 + (4p+9)u^3$$

$$+ (p^2 - 3p)u^2 - (p^2+9p)u + 4p^2 + 2p\}\ p^k$$

$$+ \left(\frac{q}{p-u^2} - 1 + 2m\right) m^2 \; p^k \quad \text{operations}$$

$$< \frac{\left(\frac{q}{p-u^2} - 1 + 2m\right) m^2}{2(p-u^2)(p-u)^2} \{-u^6 + (p+4)u^5 - (3p+10)u^4 + (8p+9)u^3$$

$$- (p+p^2)u^2 - (5p^2+9p)u + 2p(p^2+2p+1)\} \; p^k \quad \text{operations.} \tag{3.3.2}$$

Thus Gaussian eliminations saves us

$$\frac{\left(\frac{q}{p-u^2} - 1 + 2m\right) m^2}{2(p-u^2)(p-u)^2} \{u^6 - (p+5)u^5 + (7p+4)u^4 - (2p^2+3p)u^3$$

$$- (3p+p^2)u^2 + (p^2+2p)u + 2p^3 - 2p^2 \} p^k \quad \text{operations .} \tag{3.3.3}$$

<u>CHECK</u>: $u = 2$, $p = 7$, $q = 18$

(i) (3.3.1) gives us

$$\frac{(5+2m)m^2}{2 \; 3 \; 5^2} \{-128 + 736 - 1344 + 1472 + 56 - 742 + 294 + 14\} \; 7^k$$

$$= \left(\frac{2572 - 2214}{2 \; 3 \; 5^2}\right) (5 + 2m) m^2 7^k$$

$$= \left(\frac{179}{75}\right) (5 + 2m) m^2 7^k$$

(ii) (3.3.2) gives us

$$\frac{(5+2m)m^2}{2 \; 3 \; 5^2} \{-64 + 352 - 496 + 520 - 224 - 616 + 896 \} \; 7^k$$

$$= \left(\frac{1768-1400}{2 \; 3 \; 5^2}\right) (5 + 2m) \; m^2 \; 7^k$$

$$= \left(\frac{184}{75}\right) (5 + 2m) \; m^2 \; 7^k$$

(iii) (3.3.3) gives us

$$\frac{(5+2m)m^2}{2\ 3\ 5^2} \{64 - 384 + 848 - 952 - 280 + 126 + 686 - 98\} 7^k$$

$$= \left(\frac{1724 - 1714}{2\ 3\ 5^2}\right) (5 + 2m)\ m^2\ 7^k$$

$$= \frac{5}{75} (5 + 2m)\ m^2\ 7^k \quad . \qquad [\text{cf. } 2.4]$$

3.4 OPTIMIZATION OF COEFFICIENTS

Following the arguments of 2.3, in order to optimize the coefficients of p^k we should find the minimum of

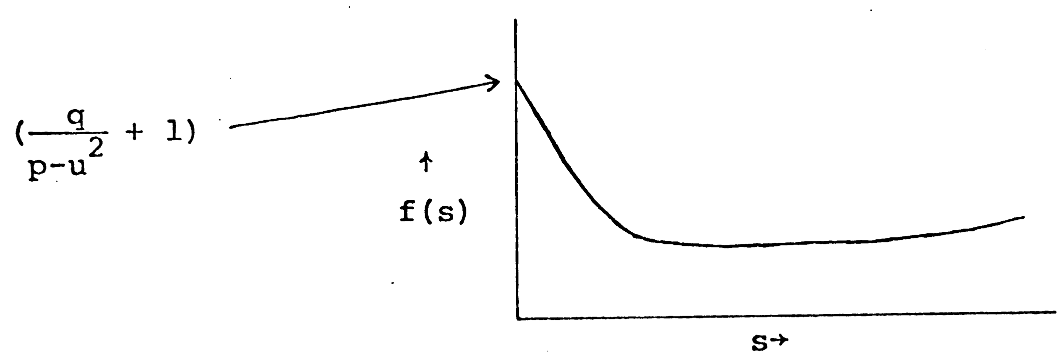
$$f(s) = \left(\frac{q}{p-u^2} - 1 + 2u^s\right) u^{2s} p^{-s} \ , \quad 0 \le s \le r$$

where $n = u^r = mu^k = u^s u^k$, the order of the matrix under consideration.

Differentiating f(s) and equating to zero, we have

$$s = \log_u \frac{-\left(\frac{q}{p-u^2} - 1\right)\left(1 + \ln \frac{u^2}{p}\right)}{2 \ln \frac{u^3}{p}} \qquad (3.4.1)$$

Since s has to be an integer, choose either the ceiling or the floor of the RHS of (3.4.1) as the value of s. That this choice does optimize the coefficients follows from a convexity argument since f(s) has the shape



3.5 OTHER APPLICATIONS

Suppose we know $A = LU$, $n \times n$, and add on m equations and m unknowns:

$$M = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right], \quad B \text{ is } n \times m,\ C \text{ is } m \times n,\ D \text{ is } m \times m.$$

Then $M = \mathscr{L}\mathscr{U}$

$$= \left[\begin{array}{c|c} L & O \\ \hline X & \tilde{L} \end{array}\right] \left[\begin{array}{c|c} U & Y \\ \hline O & \tilde{U} \end{array}\right] = \left[\begin{array}{c|c} LU & LY \\ \hline XU & XY + \tilde{L}\tilde{U} \end{array}\right]$$

$$= \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right].$$

We know L, U.

Solve: (1) $LY = B$

(2) $XU = C$

(3) Multiply XY

(4) Decompose $\tilde{L}\tilde{U} = D - XY$

(1) - (4) could be done in $O(n^{\log_u p})$ operations as compared to $O(n^2 m)$ operations using the classical method.

Also we have

$$M^{-1} = \left[\begin{array}{c|c} A^{-1}+A^{-1}BW^{-1}CA^{-1} & -A^{-1}BW^{-1} \\ \hline -W^{-1}CA^{-1} & W^{-1} \end{array}\right]$$

where $W = D - CA^{-1}B$.

Knowing A^{-1}, M^{-1} could be found in $O(n^{\log_u p})$ operations. To choose or not to choose the approach depicted in this thesis depends on the value of m.

3.6 REMARKS

Nonnegativity and integrality constraints on the coefficients or a part thereof imply many inequality relations between u, p and q. To fish out a few which imply all is a tedious task indeed. These relations, however, should provide guidance in a search for new algorithms based on the choice of u. It should be interesting to know the error bounds that would incur in using the algorithms described in this thesis.

The main result should be viewed in contrast to [3].

REFERENCES

[1] Bunch, J.R., and J. E. Hopcroft, "Triangular factorization and inversion using fast matrix multiplication," Computer Science Dept. Technical Report TR 72 - 152, Cornell University, December 1972.

[2] Franklin, J.N., Matrix Theory, Prentice-Hall, 1968.

[3] Klyuyev, V.V., and N.I. Kokovkin-Shcherbak, "On the minimization of the number of arithmetic operations needed for the solution of linear algebraic equations," translation by G.I. Tei, Technical Report CS 24, June 1965, Computer Science Dept., Stanford University.

[4] Knuth, D.E., The Art of Computer Programming, Vol. II, Addison-Wesley, 1971.

[5] Strassen, V., "Gaussian elimination is not optimal," Numerische Mathematik, 13, 1969, page 354.

[6] Winograd, S., "A new algorithm for inner product," IBM Research Report RC-1943, Nov. 1967.