TRIANGULAR FACTORIZATION AND INVERSION BY FAST MATRIX MULTIPLICATION

James R. Bunch

and

John E. Hopcroft

TR 72 - 152

December 1972

Computer Science Department
Cornell University
Ithaca, New York 14850

TRIANGULAR FACTORIZATION AND INVERSION BY FAST MATRIX MULTIPLICATION*

James R. Bunch
and
John E. Hopcroft

Department of Computer Science
Cornell University

Abstract

The fast matrix multiplication algorithm by Strassen is used to obtain the triangular factorization of a permutation of any non-singular matrix of order n in $<C_1 n^{\log_2 7}$ operations, and hence the inverse of any non-singular matrix in $<C_2 n^{\log_2 7}$ operations.

*This research was supported by ONR grant N00014-67-A-0077-0021.

TRIANGULAR FACTORIZATION AND INVERSION BY FAST MATRIX MULTIPLICATION

James R. Bunch

and

John E. Hopcroft

1. INTRODUCTION

Strassen [3] has given an algorithm using non-commutative multiplication which computes the product of two matrices of order 2 by 7 multiplications and 18 additions. Thus the product of two matrices of order $m2^k$ could be computed by $m^3 7^k$ multiplications and $(5+m)m^2 7^k - 6(m2^k)^2$ additions.

Strassen uses block LDU factorization (Householder [2], p. 126) recursively to compute the inverse of a matrix of order $m2^k$ by $m2^k$ divisions, $\leq \frac{6}{5}m^3 7^k - m2^k$ multiplications, and $\leq \frac{6}{5}(5+m)m^2 7^k - 7(m2^k)^2$ additions. The inverse of a matrix of order n could then be computed by $\leq 5.64n^{\log_2 7}$ arithmetic operations.

Let

$$A \equiv \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} I & O \\ A_{21}A_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} A_{11} & O \\ O & \Delta \end{bmatrix} \begin{bmatrix} I & A_{11}^{-1}A_{12} \\ O & I \end{bmatrix}$$

where $\Delta = A_{22} - A_{21}\, A_{11}^{-1}\, A_{12}$, and

$$A^{-1} = \begin{bmatrix} I & -A_{11}^{-1}\, A_{12} \\ O & I \end{bmatrix} \begin{bmatrix} A_{11}^{-1} & O \\ O & \Delta^{-1} \end{bmatrix} \begin{bmatrix} I & O \\ -A_{21}\, A_{11}^{-1} & I \end{bmatrix} ;$$

$$\begin{bmatrix} A_{11}^{-1} + A_{11}^{-1} A_{12} \Delta^{-1} A_{21} A_{11}^{-1} & -A_{11}^{-1} A_{12} \Delta^{-1} \\ - \Delta^{-1} A_{21} A_{11}^{-1} & \Delta^{-1} \end{bmatrix},$$

if A_{11} and Δ are non-singular.

Since the algorithm is applied recursively, it will fail whenever the inversion of a singular principal submatrix in any of the reduced matrices is required.

For example, the block LDU factorization fails to exist for a matrix as simple as

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

Every principal submatrix in every reduced matrix is non-singular if A is symmetric positive definite, strictly diagonally dominant, or irreducibly diagonally dominant (Varga [4], p. 23). However, if A is only non-singular, then we must, in general, pivot (i.e., interchange rows or columns) in order to obtain a (point or block) LDU factorization. If A is non-singular, then there exist permutation matrices P_1, P_2, Q_1, Q_2 such that AP_1, Q_1A, Q_2AP_2 have (point or block) LDU factorizations (cf. Forsythe and Moler [1], p. 36).

In Sections 2 and 3 we show that by employing pivoting we can use Strassen's fast matrix multiplication algorithm to obtain the LU, or LDU, decomposition of any non-singular matrix of order $n = 2^k$ in $< (3.64)n^{\log_2 7}$ operations, and hence its inverse in $< (10.18)n^{\log_2 7}$ operation, where an operation is defined to be a multiplication, division, addition, or subtraction.

In Section 4 we modify the algorithm so that we can find triangular factorizations in $< (2.04)n^{\log_2 7}$ operations and inverses in $< (5.70)n^{\log_2 7}$ operations when $n = 2^k$. Then, for arbitrary n, we can find triangular factorizations in $< (3.07)n^{\log_2 7}$ operations and inverses in $< (7.46)n^{\log_2 7}$ operations.

2. THE BASIC ALGORITHM

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 $P \equiv P^1 P^2 \cdots P^{n-1}$ is a permutation matrix, $L \equiv L^1 L^2 \cdots L^{n-1}$ is unit lower triangular, U is upper triangular, and $\det M = (\det P) \det U = \pm \prod_{i=1}^{n} u_{ii}$. Since $(P^i)^{-1} = P^i$ here, $P^{-1} = P^{n-1} \ldots P^2 P^1$, and

$$M^{-1} = P^{-1}U^{-1}L^{-1} = P^{n-1}\ldots P^2P^1U^{-1}(L^{n-1})^{-1} \ldots (L^2)^{-1}(L^1)^{-1},$$

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

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

Let $B_i = \{j: i_j=1,\ i=i_{k-1}2^{k-1}+i_{k-2}2^{k-2}+\ldots+i_1 2^1+i_0 2^0\}$;

let $t = \max_j\{j:j \in B_i\}$, $s = \min_j\{j:j \in B_i\}$, and $r = \begin{cases} t & \text{if } s \neq t \\ t-1 & \text{if } s=t \end{cases}$.

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

where M_{11}^{i-1} is a non-singular 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 non-singular.

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

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

, where U^i is $(i-2^s)x(i-2^s)$, V^i is $(i-2^s)x(n-i+2^s)$, E^i and G^i are $2^s x 2^s$, F^i and H^i are $2^s x(n-i)$, O is the $(2^{r+1}-i+2^s)x(i-2^s)$ zero matrix, X^i is $(n-2^{r+1})x(i-2^s)$, and Y^i is $(n-i-2^s)x(n-i+2^s)$. Further, U^i and E^i are non-singular upper triangular.

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

$$L^i = \left[\begin{array}{c|c|c} I_{i-2^s} & \multicolumn{2}{c}{O} \\ \hline & \multicolumn{1}{c}{I_{2^s}} & O \\ \cline{2-3} O & Z^i & \\ \cline{2-2} & O & I_{n-i} \end{array}\right],$$

where I_j is the identity matrix of order j.

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

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

, where $J^i = H^i - Z^i F^i$.

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

3. OPERATION COUNT.

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

Let an operation be 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 nxn matrices, to multiply an nxn matrix by an upper triangular nxn matrix, and to invert an nxn non-singular upper triangular matrix (we shall ignore lower order terms). Then $M(1) = 1$ and $M(2^k) = 7^{k+1}$ for $k \geq 1$. Since $M_T(2^k) = 4M_T(2^{k-1}) + 2M(2^{k-1}) + 2^{2k-1}$ and $I_T(2^k) = 2\ I_T(2^{k-1}) + 2\ M_T(2^{k-1})$, $M_T(2^k) = 2 \sum_{j=0}^{k-1} 4^j M(2^{k-j-1}) < (\frac{14}{3})7^k$ and $I_T(2^k) = 2 \sum_{j=0}^{k-1} 2^j\ M_T(2^{k-j-1}) < (\frac{28}{15})7^k$.

Inverting all the U^i for $1 \leq i \leq n-1$ requires

$$2^{k-1} \sum_{j=0}^{k-1} \frac{1}{2^j}\ I_T(2^j) \leq 2^{k-1}\ (\tfrac{28}{15}) \sum_{j=0}^{k-1} (\tfrac{7}{2})^j < (\tfrac{14}{15})2^k(\tfrac{2}{5})\frac{7^k}{2^k} = (\tfrac{28}{75})7^k$$

operations. Forming all the multipliers Z^i for $1 \leq i \leq n-1$ requires $2^{k-1} \sum_{j=0}^{k-1} \frac{1}{2^j}\ M_T(2^j) \leq (\frac{14}{15})7^k$ operations. Forming all the reduced matrices J^i for $1 \leq i \leq n-1$ requires

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

$$= \frac{7}{4}\, 2^{2k} \sum_{j=0}^{k-1} (\tfrac{7}{4})^j < \frac{7}{4}\, 2^{2k}\, (\tfrac{7}{4})^k\, (\tfrac{4}{3}) = (\tfrac{7}{3})\, 7^k.$$

Hence, for $n = 2^k$, triangular factorization requires $< (\frac{91}{25}) 7^k = 3.64 n^{\log_2 7}$ operations.

Inverting U requires $< (\frac{28}{15}) 7^k$ operations and $U^{-1} L^{-1}$ requires $2^{2k} \sum_{j=0}^{k-1} \frac{M(2^j)}{2^{2j+1}} = (\frac{7}{2})\, 2^{2k} \sum_{j=0}^{k-1} (\frac{7}{4})^j <$ $(\frac{7}{2}) 2^{2k}\, (\frac{7}{4})^k\, (\frac{4}{3}) = (\frac{14}{3}) 7^k$ operations.

Hence, for $n = 2^k$, inversion requires $(\frac{763}{75}) 7^k < (10.18) n^{\log_2 7}$ operations.

If M is a non-singular matrix of order n, where $2^k < n < 2^{k+1}$, then let $\mathcal{M} = M \oplus I_{2^{k+1}-n}$. We can find the triangular factorization of a permutation of $\mathcal{M}$, and hence of a permutation of M, by $< (\frac{91}{25}) 7^{k+1} = (\frac{637}{25}) 7^k < (25.48) n^{\log_2 7}$, and the inverse of $\mathcal{M}$, and hence of M, by $< (\frac{763}{75}) 7^{k+1} =$ $(\frac{5341}{75}) 7^k < (71.22) n^{\log_2 7}$.

4. A MODIFIED ALGORITHM

We can modify the algorithm in Section 2 so that the coefficient of $n^{\log_2 7}$ is smaller in the operation counts of Section 3. In particular, we find m and k such that the number of operations is minimized subject to the constraint $n \le m\, 2^k$.

First, let $n = 2^r = m\, 2^k$. Then $m = 2^{r-k} \equiv 2^s$, and $M(2^r) < (5+2m)m^2 7^k = f(s)7^r$, where $f(s) = (5+2m)m^2 7^{-s} = (5+2^{s+1})2^{2s}7^{-s}$. Since $\min\limits_{0\le s\le r} f(s) = f(3) = \frac{192}{49}$, we take

$m = 8$, $k = r-3$, and use regular multiplication and inversion for 8x8 matrices. Then $M(2^r) \le (\frac{192}{49})7^r$ for $r \ge 0$ (rather than $M(2^r) \le (7)7^r$ in Section 3). Hence each coefficient in Section 3 is multiplied by $\frac{1}{7}(\frac{192}{49})$.

Triangular factorization requires $< \frac{91}{25}(\frac{192}{343})7^r < (2.04)n^{\log_2 7}$ operations, and inversion requires $< (\frac{763}{75})(\frac{192}{343})7^r < (5.70)n^{\log_2 7}$.

Now let n be arbitrary. Taking $k = [\log_2 n-4]$ and $m = [n2^{-k}]+1$ (cf. [3]), we have $n \le m2^k$ and $(5+2m)m^2 7^k < (4.7)n^{\log_2 7}$.

Now $$M_T(m2^k) < 2(5+2m)m^2 7^{k-1} \sum_{j=0}^{k-1} (\tfrac{4}{7})^j < \tfrac{2}{3}(5+2m)m^2 7^k$$

and $$I_T(m2^k) \le \tfrac{4}{21}(5+2m)m^2 7^{k-1} \sum_{j=0}^{k-1} (\tfrac{2}{7})^j < \tfrac{4}{15}(5+2m)m^2 7^k.$$

Triangular factorization thus requires $< \frac{49}{75}(5+2m)m^2 7^k < (3.07)n^{\log_2 7}$ operations, and inversion requires $< \frac{119}{75}(5+2m)m^2 7^k < (7.46)n^{\log_2 7}$ operations.

5. REMARKS.

As seen above, the coefficient of $n^{\log_2 7}$ is very sensitive to the implementation of the algorithm. Another modification of the algorithm might reduce the coefficient. Further, the bounds we have given on the coefficient are pessimistic.

The algorithm as stated in Section 2 and 4 may not be numerically stable since we cannot guarantee that the elements in the reduced matrices are bounded. However, there may be a modification of our algorithm which guarantees stability; this question deserves further investigation.

If a fast matrix multiplication algorithm were given for multiplying two matrices of order u in v multiplications, then algorithms similar to those in Sections 2 and 4 could find the triangular factorization of a permutation of any non-singular matrix, and hence the inverse of any non-singular matrix, in $< c\, n^{\log_u v}$ operations. The algorithms would be expressed in terms of the expansions of integers modulo u.

REFERENCES

[1] Forsythe, G.E. and C.B. Moler, Computer Solution of Linear Algebraic Equations, Prentice-Hall, 1967.

[2] Householder, A.S., The Theory of Matrices in Numerical Analysis, Blaisdell, 1964.

[3] Strassen, V., "Gaussian elimination is not optimal", Numerische Mathematik, 13, 1969, pp. 354-356.

[4] Varga, R.S., Matrix Iterative Analysis, Prentice-Hall, 1962.