

Triangular Factorization and Inversion by Fast Matrix Multiplication*

By James R. Bunch and John E. Hopcroft

Abstract. The fast matrix multiplication algorithm by Strassen is used to obtain the triangular factorization of a permutation of any nonsingular matrix of order n in $< C_1 n^{1.0827}$ operations, and, hence, the inverse of any nonsingular matrix in $< C_2 n^{1.0827}$ operations.

1. Introduction. Strassen [3] has given an algorithm using noncommutative multiplication which computes the product of two matrices of order 2 by 7 multiplications and 18 additions. Then 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.

Let an operation be a multiplication, division, addition, or subtraction. Strassen showed that the product of two matrices could be computed by $< (4.7)n^{1.0827}$ operations. The usual method for matrix multiplication requires $2n^3$ operations. Thus, $(4.7)n^{1.0827} < 2n^3$ if and only if $n^{3-1.0827} > 2.35$, i.e. if $n \geq (2.35)^5 \approx 100$.

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, $\leq (6/5)m^3 7^k - m2^k$ multiplications, and $\leq (6/5)(5 + m)m^2 7^k - 7(m2^k)^2$ additions. The inverse of a matrix of order n could then be computed by $\leq (5.64)n^{1.0827}$ arithmetic operations.

Let

$$A \equiv \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} I & 0 \\ A_{21} A_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} A_{11} & 0 \\ 0 & \Delta \end{bmatrix} \begin{bmatrix} I & A_{11}^{-1} A_{12} \\ 0 & I \end{bmatrix}$$

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

$$\begin{aligned} A^{-1} &= \begin{bmatrix} I & -A_{11}^{-1} A_{12} \\ 0 & I \end{bmatrix} \begin{bmatrix} A_{11}^{-1} & 0 \\ 0 & \Delta^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ -A_{21} A_{11}^{-1} & I \end{bmatrix} \\ &= \begin{bmatrix} A_{11}^{-1} + A_{11}^{-1} A_{12} \Delta^{-1} A_{21} A_{11}^{-1} & -A_{11}^{-1} A_{12} \Delta^{-1} \\ -\Delta^{-1} A_{21} A_{11}^{-1} & \Delta^{-1} \end{bmatrix} \end{aligned}$$

if A_{11} and Δ are nonsingular.

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

Received February 26, 1973.

AMS (MOS) subject classifications (1970). Primary 15A06, 15A09; Secondary 65F05, 68A20.

Key words and phrases. LU decomposition, Strassen's method, fast matrix inversion, linear equations, Gaussian elimination, computational complexity.

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

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

Every principal submatrix in every reduced matrix is nonsingular if A is symmetric positive definite, strictly diagonally dominant, or irreducibly diagonally dominant (Varga [4, p. 23]). However, if A is only nonsingular, 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 nonsingular, 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 or LU factorizations (cf. Forsythe and Moler [1, p. 36]).

We shall ignore lower order terms. Then the usual method for inversion requires $2n^3$ operations [1, pp. 77–79]. Thus, $(5.64)n^{1.0827} < 2n^3$ if $n \geq (2.82)^5 \approx 180$.

However, one rarely needs the inverse of a matrix; rather, one usually wants to solve systems of linear equations, and here Gaussian elimination (i.e. obtaining the LU decomposition of a permutation of A) is more efficient [1, p. 79], since the LU decomposition requires $\frac{2}{3}n^3$ operations (and solving the two triangular systems requires $2n^2$ operations).

Thus, a fairer comparison is

$$(5.64)n^{1.0827} < \frac{2}{3}n^3 \quad \text{if } n \geq 3^5(2.82)^5 \approx 43,340.$$

If A is symmetric positive definite, then one should compare with symmetric Gaussian elimination or Cholesky's method.

In Sections 2 and 3, we show that, by employing pivoting, we can use Strassen's fast matrix multiplication algorithm to obtain the triangular factorization (LU decomposition) of a permutation of any nonsingular matrix of order $n = 2^k$ in $< (3.64)n^{1.0827}$ operations, and hence its inverse in $< (10.18)n^{1.0827}$ operations, 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 $< (2.04)n^{1.0827}$ operations and inverses in $< (5.70)n^{1.0827}$ operations when $n = 2^k$. Then, for arbitrary n , we can find triangular factorizations in $< (2.45)n^{1.0827}$ operations and inverses in $< (6.84)n^{1.0827}$ operations.

In Section 5 we show that matrix multiplication, triangular factorization, and inversion are equivalent in computational complexity.

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, \dots, P^{n-1} of permutation matrices so that $M = LUP$, i.e., $MP^{-1} = LU$, where $P \equiv P^1P^2 \dots P^{n-1}$ is a permutation matrix, $L \equiv L^1L^2 \dots L^{n-1}$ is unit lower triangular, U is upper triangular, and $\det M = (\det P) \det U = \pm \prod_{i=1}^n u_i$. Since $(P^i)^{-1} = P^i$ here,

$$P^{-1} = P^{n-1} \dots P^2P^1,$$

and

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

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

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

Let

$$B_i = \{j : i_j = 1, i = i_{k-1}2^{k-1} + i_{k-2}2^{k-2} + \dots + i_12^1 + i_02^0\};$$

let

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

Then

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

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 $N^i \equiv M^{i-1}P^i$, $n_{i,i} \neq 0$, and N^i can be partitioned as

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

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

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

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

where I_j is the identity matrix of order j .

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

$$M^i = \left[\begin{array}{cc|cc} U^i & & V^i & \\ \hline 0 & E^i & & F^i \\ \hline & 0 & & J^i \\ \hline X^i & & & Y^i \end{array} \right] \quad \text{where } J^i = H^i - Z^i F^i.$$

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

3. Operation Count. Finding the permutation P^i requires at most $n - i$ comparisons, and, if $P_{ii}^i = 0$, 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 $n \times n$ matrices, to multiply an $n \times n$ matrix by an upper triangular $n \times n$ matrix, and to invert an $n \times n$ nonsingular 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 $I_T(2^k) = 2I_T(2^{k-1}) + 2M_T(2^{k-1})$,

$$M_T(2^k) = 2 \sum_{i=0}^{k-1} 4^i M(2^{k-i-1}) < \left(\frac{14}{3}\right)7^k$$

and

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

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

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

operations. Forming all the multipliers Z^i for $1 \leq i \leq n - 1$ requires

$$2^{k-1} \sum_{i=0}^{k-1} \frac{1}{2^i} M_T(2^i) \leq \left(\frac{14}{15}\right)7^k$$

operations. Forming all the reduced matrices J^i for $1 \leq i \leq n - 1$ requires

$$\begin{aligned} \sum_{i=0}^{k-1} \left\{ \sum_{l=0}^{(n-2^i)/2^{i+1}} [2^k - (2l + 1)2^i] \right\} \frac{M(2^i)}{2^i} &\leq 2^{2(k-1)} \sum_{i=0}^{k-1} \frac{M(2^i)}{2^{2i}} \\ &= \left(\frac{7}{4}\right)2^{2k} \sum_{i=0}^{k-1} \left(\frac{7}{4}\right)^i < \frac{7}{4} 2^{2k} \left(\frac{7}{4}\right)^k \left(\frac{4}{3}\right) = \left(\frac{7}{3}\right)7^k. \end{aligned}$$

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

Inverting U requires $< (28/15)7^k$ operations and $U^{-1}L^{-1}$ requires

$$2^{2k} \sum_{i=0}^{k-1} \frac{M(2^i)}{2^{2i+1}} = \left(\frac{7}{2}\right) 2^{2k} \sum_{i=0}^{k-1} \left(\frac{7}{4}\right)^i < \left(\frac{7}{2}\right) 2^{2k} \left(\frac{7}{4}\right)^k \left(\frac{4}{3}\right) = \left(\frac{14}{3}\right) 7^k$$

operations.

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

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

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 $n \leq m2^k$.

First, let $n = 2^r = m2^k$. 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 $\min_{0 \leq s \leq r} f(s) = f(3) = 192/49$, we take $m = 8$, $k = r - 3$, and use regular multiplication and inversion for submatrices of order ≤ 8 . Then, $M(2^r) \leq (192/49)7^r$ for $r \geq 0$ (rather than $M(2^r) \leq (7)7^r$ in Section 3). Hence, each coefficient in Section 3 is multiplied by $(1/7)(192/49)$.

Triangular factorization requires

$$< \left(\frac{91}{25}\right)\left(\frac{192}{343}\right)7^r < (2.04)n^{\log_2 7}$$

operations, and inversion requires

$$< \left(\frac{763}{75}\right)\left(\frac{192}{343}\right)7^r < (5.70)n^{\log_2 7}.$$

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

Now,

$$M_T(m2^k) < 2(5 + 2m)m^2 7^{k-1} \sum_{i=0}^{k-1} \left(\frac{4}{7}\right)^i < \frac{2}{3}(5 + 2m)m^2 7^k$$

and

$$I_T(m2^k) \leq \frac{4}{21}(5 + 2m)m^2 7^{k-1} \sum_{i=0}^{k-1} \left(\frac{2}{7}\right)^i < \frac{4}{15}(5 + 2m)m^2 7^k.$$

Triangular factorization thus requires $< (39/75)(5 + 2m)m^2 7^k < (2.45)n^{\log_2 7}$ operations, and inversion requires $< (109/75)(5 + 2m)m^2 7^k < (6.84)n^{\log_2 7}$ 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 Sections 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 where $\log_u v > 2$, then algorithms similar to those in Sections 2 and 4 could find the triangular factorization of a permutation of any nonsingular matrix, and hence the inverse of any nonsingular matrix, in $< cn^{10 \log_u v}$ operations. The algorithms would be expressed in terms of the expansions of integers modulo u and hold over any ring.

Similarly, given an algorithm for multiplying two $n \times n$ matrices in $O(n^\alpha)$ operations, then algorithms similar to those in Sections 2 and 4 can find the triangular factorization of a permutation of any $n \times n$ nonsingular matrix in $O(n^\alpha)$ operations, and hence the inverse of any $n \times n$ nonsingular matrix in $O(n^\alpha)$ operations, where $2 < \alpha \leq 3$. Conversely, given an algorithm for calculating the inverse of an $n \times n$ nonsingular matrix in $\leq cn^\alpha$ operations, the product of two $n \times n$ matrices A, B can be computed in $\leq (3^\alpha c)n^\alpha$ operations, since

$$\begin{bmatrix} I & A & 0 \\ 0 & I & B \\ 0 & 0 & I \end{bmatrix}^{-1} = \begin{bmatrix} I & -A & AB \\ 0 & I & -B \\ 0 & 0 & I \end{bmatrix}$$

(Winograd [5]).

Thus, matrix multiplication, triangular factorization, and inversion are equivalent in computational complexity.

Addendum. A. Schönhage has proposed another method for forming the inverse of any nonsingular matrix (over the reals or complexes), based on the generalized inverse identity $M^{-1} = (M^*M)^{-1}M^*$. Since M^*M is symmetric positive definite, Strassen's inversion algorithm will not fail (but the condition number [1, pp. 20–26] of the problem will be squared). Strassen's matrix multiplication algorithm is used to form M^*M and to multiply $(M^*M)^{-1}$ by M^* . This gives the inverse of any nonsingular matrix in $\leq cn^{10 \log_2 7}$ operations (but the constant c is larger than the ones given above).

Department of Computer Science
Cornell University
Ithaca, New York 14850

1. G. E. FORSYTHE & C. B. MOLER, *Computer Solution of Linear Algebraic Equations*, Prentice-Hall, Englewood Cliffs, N.J., 1967. MR 36 #2306.
2. A. S. HOUSEHOLDER, *The Theory of Matrices in Numerical Analysis*, Blaisdell, New York, 1964. MR 30 #5475.
3. V. STRASSEN, "Gaussian elimination is not optimal," *Numer. Math.*, v. 13, 1969, pp. 354–356. MR 40 #2223.
4. R. S. VARGA, *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, N.J., 1962. MR 28 #1725.
5. S. WINOGRAD, Private communication.