NORTH-HOLLAND

Every Unit Matrix is a *LULU*

Gilbert Strang*
Department of Mathematics
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Submitted by Richard A. Brualdi

ABSTRACT

The four matrices $L_0 U_0 L_1 U_1$ at the end of the title are triangular with ones on their main diagonals. Their product has determinant one. Following a question and theorem of Toffoli, we show that any matrix with determinant one can be factored in this way. A transformation of the plane becomes a sequence of one-dimensional shears, with $n^2 - 1$ free parameters. © 1997 Elsevier Science Inc.

1. INTRODUCTION

Matrix factorizations now dominate the subject of linear algebra. They are part of the theory and part of the language. Often their history is obscure—a theorem is brought forward into its proper place by its applications. One common thread is that the total number of parameters is n^2 when factoring a (not quite arbitrary) matrix of order n:

*E-mail: gs@math.mit.edu.

LINEAR ALGEBRA AND ITS APPLICATIONS 265:165–172 (1997)

655 Avenue of the Americas, New York, NY 10010

0024-3795/97/$17.00
PII S0024-3795(96)00598-8

1. $A = LDU$ has $\dfrac{n^2-n}{2} + n + \dfrac{n^2-n}{2}$ parameters.

2. $A = QR$ has $\dfrac{n^2-n}{2} + \dfrac{n^2+n}{2}$ parameters.

3. $A = S\Lambda S^{-1}$ has $n^2 - n + n + 0$ parameters.

4. $A = UTU^H$ has $\dfrac{n^2-n}{2} + \dfrac{n^2+n}{2} + 0$ parameters.

5. $A = U\Sigma V^H$ has $\dfrac{n^2-n}{2} + n + \dfrac{n^2-n}{2}$ parameters.

6. $A = QH$ has $\dfrac{n^2-n}{2} + \dfrac{n^2+n}{2}$ parameters.

Each factorization is "generically" possible when complex numbers are allowed. The last three, Schur, SVD, and polar, are always possible. The first pair and last pair are real when A is real. There are special factorizations LDL^H and $U\Lambda U^H$ for Hermitian matrices, again with the correct parameter count [now $(n^2 + n)/2$]. And there are combinations like lower triangular L times symmetric H for which good applications have not been found.

We do not know a general theory of matrix factorizations. Such a study seems reasonable, but that is not at all our goal. The purpose of this note is to add an occasionally useful variation to the LDU factorization, by forcing $D = I$ (unit pivots) but then extending to more triangular factors—generically to ULU and exceptionally to $LULU$. The determinant of A is necessarily one, since all diagonal entries are ones. The factors are *shears*. These L's and U's are not repeated—they are different—so a better notation is $A = L_0U_0L_1U_1$.

Before describing these shears, we comment further on factorizations 1–6. The map from A to its factors is nonlinear. There is a choice of signs in the columns of Q and the diagonal entries of R. With the unit eigenvectors in S there is also freedom to reorder. More important is the possibility of *nonexistence*: elimination can fail and diagonalization can fail. We rescue diagonalization, as far as possible, by the 1's in the Jordan form. We rescue elimination by allowing a permutation matrix P. For numerical analysts it comes first: for algebraists it comes between L and U. If we permute to avoid small pivots, P may as well come first. Algebra prefers a canonical form, avoiding only zero pivots. The row and column operations on A are downward and rightward, accepting the first nonzeros as pivots. Then $L^{-1}AU^{-1}$ has at most one nonzero entry in each row and column, and P appears naturally.

A corresponding rescue will be needed for our factorization into shears. This is responsible for extending *ULU* to *LULU*. The generic case $A = U_0 L_1 U_1$ will have $n^2 - 1$ parameters (the determinant of A is 1), and an extra factor L_0 handles exceptional cases when certain submatrices are singular. For $n = 2$ and $n = 3$, we expect three and eight parameters in the factors of A:

$$U_0 L_1 U_1 = \begin{bmatrix} 1 & x \\ & 1 \end{bmatrix} \begin{bmatrix} 1 & \\ x & 1 \end{bmatrix} \begin{bmatrix} 1 & x \\ & 1 \end{bmatrix}, \tag{1}$$

$$U_0 L_1 U_1 = \begin{bmatrix} 1 & & x \\ & 1 & x \\ & & 1 \end{bmatrix} \begin{bmatrix} 1 & & \\ x & 1 & \\ x & x & 1 \end{bmatrix} \begin{bmatrix} 1 & x & x \\ & 1 & x \\ & & 1 \end{bmatrix}. \tag{2}$$

The factor U_0 has $n - 1$ nonzeros above the diagonal, all in the last column. Then the count for three factors is

$$(n - 1) + 2\left(\frac{n^2 - n}{2}\right) = n^2 - 1.$$

The *ULU* factorization has special importance for orthogonal matrices, as in

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} 1 & -\tan(\theta/2) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \sin\theta & 1 \end{bmatrix} \begin{bmatrix} 1 & -\tan(\theta/2) \\ 0 & 1 \end{bmatrix}. \tag{3}$$

This decomposition into shears is valuable in computer graphics, when a plane figure is to be turned. The rotation is effectively reduced to a series of translations in coordinate directions. Instead of interpolating between rotated pixels and original pixels, the processing of each shear is *one-dimensional*. No rescaling is needed with unit determinants. This three-pass implementation seems to have been discovered independently in [1–3]; the full history is unclear. The recent paper [4] develops a careful analysis of one-dimensional interpolation, leading to a good algorithm and impressive figures.

Note that a rotation by $\theta = \pi$ is not permitted in (3), because the tangent becomes infinite. The 2-by-2 matrix $A = -I$ is not a product of three shears. This is one of the exceptional cases requiring four shears.

For three-dimensional rotations, Toffoli [5] presented a generalization. Certainly A is a product of three plane rotations (through Euler angles). Each plane rotation is a product of three plane shears (making nine). By allowing more general triangular shears, Toffoli found that a three-pass factorization is again possible. His goal was the same: "The advantage remains that the address arithmetic for a shear (at the memory controller level) is much simpler than for a rotation (at the processor level)." In certain architectures the shear is a native operation.

It was natural to ask about matrices that are not rotations, and about orders $n > 3$. Here we extend the $A = ULU$ theorem to the generic case, and continue to $A = LULU$ for the exceptional cases. It was already known to algebraists that every unit matrix is a product of shears. It may not have been known that four shears are sufficient.

2. GENERIC CASE

Which matrices can be factored into $A_1 = L_1 U_1$ with *ones on the main diagonals of both factors*? Certainly $\det A_1 = 1$. More than that, every upper left submatrix must have $\det A_1^{(k)} = 1$. The reason is that these k-by-k submatrices also factor into $A_1^{(k)} = L_1^{(k)} U_1^{(k)}$. The ones are still on the diagonal, so all determinants equal 1. In the language of elimination, all pivots of A_1 are one with no row exchanges.

Suppose we attempt to change a given A into such a matrix A_1, by adding multiples of the last row of A to earlier rows. When the last row is $\mathbf{v}$, we add $c_i\mathbf{v}$ to row i (for each $i < n$). This operation will be $U_0^{-1}A$, producing A_1. The upper left k-by-k submatrix becomes

$$A^{(k)} + \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{bmatrix} \mathbf{v}_k = \text{first } k \text{ columns of} \begin{bmatrix} \mathbf{r}_1 + c_1\mathbf{v} \\ \mathbf{r}_2 + c_2\mathbf{v} \\ \vdots \\ \mathbf{r}_k + c_k\mathbf{v} \end{bmatrix}.$$

The determinant of this matrix is intended to equal one. We write $D_i^{(k)}$ for the determinant of the submatrix $A^{(k)}$ after $\mathbf{v}$ has replaced the ith row $\mathbf{r}_i$. Then the matrix above has (by multilinearity of determinants)

$$\text{determinant} = \det A^{(k)} + c_1 D_1^{(k)} + \cdots + c_k D_k^{(k)}. \tag{4}$$

We have a triangular system of $n-1$ equations for the $n-1$ coefficients c_i that yield submatrices with determinant one:

$$\begin{aligned} c_1 D_1^{(1)} &= 1 - \det A^{(1)}, \\ c_1 D_1^{(2)} + c_2 D_2^{(2)} &= 1 - \det A^{(2)}, \\ &\vdots \\ c_1 D_1^{(n-1)} + \cdots + c_{n-1} D_{n-1}^{(n-1)} &= 1 - \det A^{(n-1)}. \end{aligned} \tag{5}$$

If all coefficients $D_k^{(k)}$ on the diagonal of those systems are nonzero, the *ULU* factorization is not only possible but unique.

THEOREM 1. *If all $D_k^{(k)} \neq 0$, then the numbers c_i and the factors of A are uniquely determined:*

$$A = U_0 L_1 U_1 \qquad \textit{with} \quad U_0 = \begin{bmatrix} 1 & & & -c_1 \\ & 1 & & -c_2 \\ & & \ddots & \vdots \\ & & & 1 \end{bmatrix}. \tag{6}$$

The numbers c_i come from (5). *They appear in U_0^{-1}, taking A to A_1. These numbers become $-c_i$ when this "upward" operation is on the right side in $A = U_0 A_1$. Now A_1 has all upper left submatrices with determinant* 1. *Ordinary elimination then gives $A_1 = L_1 U_1$, and the three-shear factorization is established.*

Note that uniqueness fails for $A = I$. The coefficients $D_k^{(k)}$ are all zero because the last row **v** starts with zeros. There are many factorizations $I = U_0 I U_0^{-1}$. But the generic case has $n^2 - 1$ uniquely determined parameters in the three shears (only in the last column of U_0).

3. EXCEPTIONAL CASES

Two difficulties can arise in the above construction. Either the particular vector **v** in the last row of A fails to give nonzero coefficients $D_k^{(k)}$, or *no* vector **v** in that row can do so. When the fault lies in the particular vector **v**, we use a fourth (downward) shear L_0 to replace it by a better vector. When the fault is not in the last row, but in other rows, we include more nonzeros in U_0. Here are examples of both:

$$\text{fault in } \mathbf{v}: \quad A = \begin{bmatrix} 2 & 0 \\ 0 & 0.5 \end{bmatrix}; \qquad \text{fault in } A^{(2)}: \quad A = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}. \quad (7)$$

For the first matrix, L_0 can add row 1 to row 2. Then the last row $\mathbf{v}$ becomes satisfactory. Subtracting a multiple of $\mathbf{v}$ from row 1 yields the matrix A_1, with upper left entry equal to one. Then $A_1 = L_1 U_1$ as required.

For the second matrix, no last row would be satisfactory. Adding multiples to row 1 and 2 cannot produce a nonsingular 2-by-2 matrix in the corner. A remedy is available by including other upward operations in U_0: add row 3 to row 1. It is this possibility that we have to generalize.

Lemma 1 will remove the first difficulty, and Lemma 2 the second. We do not attempt a "minimal" adjustment in these exceptional cases when Theorem 1 does not succeed.

LEMMA 1. *Suppose* $\mathbf{v}$ *in* R^n *is not in the span of the first rows* $\mathbf{r}_1, \dots, \mathbf{r}_{n-1}$. *By downward row operations, the last row of A can be changed to a nonzero multiple of* $\mathbf{v}$.

Proof. Write $\mathbf{v}$ as a combination of $\mathbf{r}_1, \dots, \mathbf{r}_n$, which form a basis because $\det A = 1$. Dividing by the coefficient of $\mathbf{r}_n$ yields a vector $c\mathbf{v} = a_1\mathbf{r}_1 + \cdots + a_{n-1}\mathbf{r}_{n-1} + \mathbf{r}_n$, in which $\mathbf{r}_n$ has coefficient one. Downward row operations can produce this vector $c\mathbf{v}$ in the last row of A. If there is any acceptable $\mathbf{v}$ for the last row, this lemma puts it there by downward operations. Those are executed by (the inverse of) the fourth factor L_0.

LEMMA 2. *If a matrix has rank at least* $k - 1$, *then upward row operations can make its first* $k - 1$ *rows linearly independent.*

These upward steps can be done in a definite order. Remember that row n is not involved at this stage. The steps begin as follows:

1. The $n - 1$-by-2 matrix in the first two columns has rank at least 1. Make its first row nonzero.
2. The $n - 1$-by-3 matrix in the first three columns has rank at least 2. Make its second row independent of its first row.

When we reach the $n-1$-by-k matrix in the first k columns, the rank is at least $k-1$ (because including the nth row would give k complete columns of A, and those columns are independent). The first $k-2$ rows are already made independent. At that point, fix row $k-1$ using lower rows. Stop when this is done for $k=n-2$.

Now all the k-by-k determinants $D_k^{(k)}$, with **v** inserted in row k, are to be nonzero. Almost any **v** will make this true. For $k=1,\dots,n-1$, the vector **v** has to yield a kth row that is independent of the first $k-1$ rows. Lemma 1 assures that such a **v** is available.

Note that we may test the preliminary upward operations early, to determine an acceptable **v** for the nth row. L_0^{-1} puts a multiple of **v** in that row. Then U_0^{-1} does the upward operations first (not altering row n, as in Lemma 2) by using **v** to reach A_1. With the coefficients c_i from Equation (5), all upper left submatrices have det $A_1^{(k)}=1$. Ordinary elimination gives $A_1=L_1U_1$, and the four-shear factorization $A=L_0U_0A_1=L_0U_0L_1U_1$ is complete.

We end with an example that creates **v** in row 3 and then produces unit matrices (determinant 1) in rows 1 and 2:

$$A=\begin{bmatrix}0&1&0\\1&0&0\\0&0&-1\end{bmatrix}\to\begin{bmatrix}0&1&0\\1&0&0\\1&-1&-1\end{bmatrix}$$

$$\to\begin{bmatrix}1&0&-1\\0&1&1\\1&-1&-1\end{bmatrix}=A_1\to L_1U_1.$$

We thank Tom Toffoli for sending his "ULU theorem" for orthogonal matrices, and also Chris Leary for his invitation to teach a class at SUNY Geneseo—in which the 2-by-2 "LULU theorem" was proved (jointly with the class).

REFERENCES

1 H. Kiesewetter and A. Graf, Rotation in Digital Grids and Corresponding Models, Zentral Inst. für Kybernetik, Akad. der Wissenschaft, DDR, 1985.

2 A. W. Paeth, A fast algorithm for general raster rotation, in *Graphic Interface 86*, Canadian Information Processing Soc., Vancouver, pp. 77–81; reprinted with corrections in *Graphics Gems* (Glassner, Ed.), Academic, 1990, pp. 179–195.

3 P. E. Danielsson and M. Hammerin, High-accuracy rotation of images, *CVGIP Graphical Models and Image Process.* 54:340–344 (1992).
4 M. Unser, P. Thévenaz, and L. Yaroslavsky, Convolution-based interpolation for fast high-quality rotation of images, *IEEE Trans. Image Process.* submitted for publication.
5 T. Toffoli and J. Quick, Three-dimensional rotations by three shears, *Graphical Models and Image Process.* 59:89–96 (1997).
6 T. Toffoli, Almost every unit matrix is a ULU, Lin. Algebra and Its Appls. 259:31–38 (1997).
7 T. Ypma, A SAXPY formulation for plane rotations, Preprint, Western Washington Univ., 1994.

Received 18 April 1994; final manuscript accepted 13 November 1994