

## A Stable Algorithm for Updating Triangular Factors Under a Rank One Change

By R. Fletcher and S. P. J. Matthews

**Abstract.** An algorithm is presented for updating the  $LU$  factors of an  $n \times n$  matrix  $A$ , when  $A$  is changed by a matrix of rank one. The algorithm is based on the repeated use of triangular operations, and stability is obtained by allowing both row and column pivoting. The cost of the algorithm is approximately proportional to the maximum permitted depth for the pivot search. For well-conditioned matrices a maximum depth of 3 is sufficient to ensure stability. For substantially rank deficient matrices it is theoretically possible that pivots of any depth may be required, but in practice we find that a value of 5 is adequate. We suggest a pivot strategy, based on minimizing a growth bound, which penalizes deep pivots and imposes a maximum depth of pivot through a default value.

On well-behaved problems the asymptotic cost of the update is observed to be approximately  $2.6n^2$  compared with  $8n^2$  (or worse) for updating orthogonal factors. Given the accuracy obtained by the new algorithm, we feel that there are many applications in which the lower cost of triangular factors can be exploited. Comparison with *ab initio* factorization indicates that for  $n \geq 10$  updating triangular factors is advantageous.

**1. Introduction.** When solving a large system of nonlinear equations using quasi-Newton or secant techniques the solution of a system of equations

$$(1.1) \quad Ax = b, \quad A \in \mathbf{R}^{n \times n}, \mathbf{x}, \mathbf{b} \in \mathbf{R}^n,$$

is required at each stage. This problem arises not only in its own right but also as a subproblem in methods for the solution of ordinary differential equations. The matrix  $A$  is in general unsymmetric and is updated on each iteration using a rank one change. On the grounds of efficiency it is desirable to calculate factors for just the initial matrix  $A$ , and subsequently to update the relevant factors rather than recalculating them *ab initio*. The methods discussed in this paper for performing this update all generalize to the more general problem  $A \in \mathbf{R}^{n \times m}$ ,  $m \neq n$ .

When problem (1.1) arises in the above context it has become usual to use an orthogonal ( $QR$ ) factorization of the matrix  $A$  which can be updated using techniques described by Daniel, et al. [2] and Gill, et al. [4]. The disadvantage of this approach is that the  $QR$  factors are more expensive to calculate than triangular  $LU$  factors, and also the updating procedure is quite expensive. Alternatively, given the  $LU$  factors of a matrix  $A$ , it is possible to update these factors using an algorithm described by Bennett [1]. In this algorithm there is no pivoting and, hence, if the updated matrix does not possess triangular factors, then the method will fail. Even if

---

Received January 30, 1984; revised April 9, 1985.

1980 *Mathematics Subject Classification*. Primary 65F05.

*Key words and phrases*. Rank one update,  $LU$  factors, triangular factors, stability.

©1985 American Mathematical Society  
0025-5718/85 \$1.00 + \$.25 per page

the algorithm does not fail it is possible that the updated factors may grow rapidly, which implies that the round-off errors are magnified so that the algorithm is potentially unstable in practice.

In this paper we present a method for updating the  $LU$  factors which uses both column and row pivoting to maintain stability whilst to a large extent retaining the efficiency of Bennett's method. In Section 2 we show how Bennett's algorithm is essentially equivalent to the reduction of a product of two Hessenberg matrices into a product of two triangular matrices, using elementary triangular operations. We then generalize this technique so that it will reduce general matrix products to the desired form. This new algorithm is then combined with permutation operators to obtain an algorithm which updates triangular factors using elementary triangular and permutation operators. In Section 3 we relate the problem of pivot selection to that of error growth. We derive a quantity that gives an indication of the growth associated with the choice of any particular pivot. This quantity is used in Section 4 as the basis of a heuristic criterion for pivot selection which aims to limit error growth whilst discouraging the choice of pivots which disrupt the matrix structure. A selection of numerical experiments are described to choose parameters for this test and to illustrate the properties of the algorithm. These experiments indicate that the algorithm performs the required updates at a considerable saving compared with either updating the equivalent  $QR$  factors or recalculating the factors using Gaussian elimination, whilst at the same time maintaining stability and restricting error growth.

**2. The Rank One Update.** The rank one change to a matrix  $A$  may be written

$$(2.1) \quad A^* = A + \mathbf{xy}^T, \quad A^*, A \in \mathbf{R}^{n \times n}, \mathbf{x}, \mathbf{y} \in \mathbf{R}^n,$$

and we assume that we have the representation of  $A$

$$(2.2) \quad PAQ^T = L^0U^0 = \sum_{i=1}^n \mathbf{l}_i^0 \mathbf{u}_i^{0T},$$

where  $L^0 = [\mathbf{l}_1^0, \mathbf{l}_2^0, \dots, \mathbf{l}_n^0]$  is lower triangular,  $U^0 = [\mathbf{u}_1^0, \mathbf{u}_2^0, \dots, \mathbf{u}_n^0]^T$  is upper triangular and  $P$  and  $Q$  are permutation matrices. Note that neither  $L^0$  nor  $U^0$  is required to be unit triangular, so there is a small amount of redundant information in this representation. In the case when these factors have been calculated using Gaussian elimination with partial pivoting,  $Q$  is the identity matrix and  $L^0$  is a unit lower triangular matrix. Observe that (2.2) may be used to write  $A^*$  as the product of two Hessenberg matrices

$$(2.3) \quad P^*A^*Q^{*T} = L^0U^0 + P\mathbf{xy}^TQ^T = L^1U^1 = \sum_{i=1}^{n+1} \mathbf{l}_i^1 \mathbf{u}_i^{1T},$$

where  $L^1 = [P\mathbf{x} : L^0]$  and  $U^1 = [Q\mathbf{y} : U^{0T}]^T$ .

The aim of the algorithm is to restore this representation of  $A^*$  to the triangular form

$$(2.4) \quad P^*A^*Q^{*T} = L^*U^*.$$

Initially, we consider just the first stage of the transformation in which we obtain the first row of  $L^*$ , and column of  $U^*$ , by eliminating the elements  $l_{12}^1$  and  $u_{21}^1$ . (Subsequently, we omit the superscript 1 for clarity.) As in Fletcher and Matthews [3], we can eliminate these two elements by considering the first two terms of the summation in (2.3) and using two elementary triangular operations. These two operations can be combined to give

$$[l_1^+, l_2^+][u_1^+, u_2^+]^T = [l_1, l_2]BB^{-1}[u_1, u_2]^T,$$

where  $B$  is the matrix

$$(2.5) \quad B = \begin{bmatrix} 1 & -\frac{l_{12}}{l_{11}} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \frac{l_{11}u_{21}}{\alpha} & 1 \end{bmatrix} = \begin{bmatrix} \frac{l_{11}u_{11}}{\alpha} & -\frac{l_{12}}{l_{11}} \\ \frac{l_{11}u_{21}}{\alpha} & 1 \end{bmatrix}$$

and  $\alpha = l_{11}u_{11} + l_{12}u_{21} = (PA^*Q^T)_{11}$ . Thus the two columns of  $L$  and  $U^T$  are replaced by new quantities (superscripted +) defined by

$$(2.6) \quad [l_1^+, l_2^+] = [l_1, l_2]B$$

and

$$(2.7) \quad \begin{bmatrix} u_1^{+T} \\ u_2^{+T} \end{bmatrix} = B^{-1} \begin{bmatrix} u_1^T \\ u_2^T \end{bmatrix}.$$

If  $(PA^*Q^T)_{11} = 0$ , then the operation will fail, and if  $l_{11} = 0$ , we will have to interchange the two vectors  $l_1 \leftrightarrow l_2$  and  $u_1 \leftrightarrow u_2$  before carrying out the operation. Assuming that  $\alpha \neq 0$  the new vectors  $l_1^+$  and  $u_1^+$  are the first row and column of  $L^*$  and  $U^*$  respectively, and we define reduced matrices  $L^2, U^{2T} \in \mathbf{R}^{(n-1) \times n}$  by stripping off the first row and column of both  $L^+$  and  $U^+$ , where  $L^+ = [l_1^+, l_2^+, l_3, \dots, l_n]$ , etc. These new matrices have the same form as  $L^1$  and  $U^1$ , and so by repeating the above procedure we can calculate the new factors  $L^*$  and  $U^*$ . This algorithm is essentially equivalent to that proposed by Bennett.

The reason why this approach can fail is that after the rank one change the pivot element  $(PA^*Q^T)_{11}$  may be zero, so that the  $L^*U^*$  factors do not exist. It is therefore necessary to introduce some form of pivoting into the algorithm so as to bring a substantial element into this leading position. This may cause serious disruption to the Hessenberg structure of  $L^1$  and  $U^1$ . If we wish the algorithm to work for rank deficient matrices it must be able to handle cases such as

$$(2.8) \quad A^* = I - e_1 e_1^T$$

in which the entire first row and column of  $A^*$  are zero. In this case it is clear that neither row nor column pivoting will be sufficient on its own. If we assume that  $A$  is full rank then the leading  $2 \times 2$  submatrix of  $PAQ^T$  has rank 2, and so the leading  $2 \times 2$  submatrix of  $PA^*Q^T$  has at least rank 1 and must contain at least one nonzero element which is a feasible pivot element. Therefore, in this case only interchanges in row/column  $1 \leftrightarrow 2$  are required and the amount of disruption is small. However, in general we may wish to deal with singular or near singular matrices for which it is possible that all four of these potential pivots lead to large error growth. Therefore, we must allow the algorithm to select deeper pivots when necessary to retain stability.

It is clear that we cannot in general expect to retain the Hessenberg structure of  $L^1$  and  $U^1$ . Our next observation is, therefore, to show how a general matrix product  $A = XY$ , ( $X, Y^T \in \mathbf{R}^{n \times (n+1)}$ ) can be reduced to the triangular form  $PAQ^T = LU$ . Clearly, we need only describe the first stage in which off-diagonal elements in the first row of  $X$ , and first column of  $Y$ , are zeroed. We shall require two types of pivoting: *external pivoting*  $A \rightarrow PAQ^T$  in which we change the external permutation matrices which are associated with the representation of  $A$ , and *internal pivoting*  $A = XY \rightarrow XR^{-1}RY$  in which the external permutation is unchanged but the columns of  $X$ , and the corresponding rows of  $Y$ , are rearranged. By external pivoting we can ensure that the pivot element  $(PAQ^T)_{11}$  is substantial, and by internal pivoting that

$$(2.9) \quad |X_{11}Y_{11}| \geq |X_{1i}Y_{i1}|, \quad i = 2, \dots, n+1.$$

To simplify the presentation we will assume that all interchanges have been done previously and hence no pivots are necessary.

The first step is to use  $X_{11}$  to eliminate all other nonzero elements in the first row of  $X$ . This is equivalent to the matrix operation

$$\hat{X} = XT_1,$$

where

$$(2.10) \quad T_1 = \begin{bmatrix} 1 & -\frac{X_{12}}{X_{11}} & -\frac{X_{13}}{X_{11}} & \dots & -\frac{X_{1,n+1}}{X_{11}} \\ & 1 & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix}.$$

To retain the same matrix product we must carry out the inverse operations on the  $Y$  matrix giving rise to a matrix  $\hat{Y} = T_1^{-1}Y$ . The first column of  $\hat{Y}$  is  $(A_{11}/X_{11}, X_{21}, Y_{31}, \dots, Y_{n+1,1})^T$  and the second step is to use  $\hat{Y}_{11}$  to eliminate the other nonzero elements in this column. This can be written as the matrix operation

$$Y^+ = T_2\hat{Y},$$

where

$$(2.11) \quad T_2 = \begin{bmatrix} 1 & & & & \\ -\frac{X_{11}Y_{21}}{A_{11}} & 1 & & & \\ -\frac{X_{11}Y_{31}}{A_{11}} & & 1 & & \\ \vdots & & & \ddots & \\ -\frac{X_{11}Y_{n+1,1}}{A_{11}} & & & & 1 \end{bmatrix}.$$

Again the inverse operations must be carried out on  $\hat{X}$  to obtain  $X^+ = \hat{X}T_2^{-1}$ . We can now combine the two steps into one matrix operation

$$X^+ = XB, \quad Y^+ = B^{-1}X,$$

where

$$(2.12) \quad B = T_1 T_2^{-1} = \begin{bmatrix} \frac{X_{11}Y_{11}}{A_{11}} & -\frac{X_{12}}{X_{11}} & \dots & -\frac{X_{1,n+1}}{X_{11}} \\ \frac{X_{11}Y_{21}}{A_{11}} & 1 & & \\ \vdots & & \ddots & \\ \frac{X_{11}Y_{n+1,1}}{A_{11}} & & & 1 \end{bmatrix}.$$

It can be seen that if any of the positions,  $X_{1i}$  or  $Y_{i1}$  already contains a zero, then the corresponding columns or rows in  $B$  will be columns or rows of the identity matrix. This result also holds for  $B^{-1}$ , so those rows and columns of  $X$  or  $Y$  are left unchanged. We now strip off the first row and column of  $X^+$  and  $Y^+$  and repeat the above operations inductively to yield the triangular form  $PAQ^T = LU$ .

We now return to the main problem to consider the effect of allowing external pivoting in the algorithm. At the start of the  $k$ th stage we have the reduced matrix

$$(2.13) \quad A^k = L^k U^k$$

in which the matrices  $L^k$  and  $U^k$  typically have the form

$$(2.14) \quad p^k \left\{ \begin{bmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times & \times & \times \end{bmatrix} \right. \left. \begin{matrix} q^k \\ \left[ \begin{matrix} \times & \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times & \times \\ & \times & \times & \times & \times & \times & \times \\ & & \times & \times & \times & \times & \times \\ & & & \times & \times & \times & \times \\ & & & & \times & \times & \times \\ & & & & & \times & \times \\ & & & & & & \times \end{matrix} \right] \end{matrix} \right\},$$

where  $p^k$  and  $q^k$  are determined by the fill in during the previous stages of the algorithm. The first step of the general algorithm described above is to use external pivoting to bring a large element of  $A^k$  (indexed  $(r^k, s^k)$  say) to the (1,1) pivot position. If  $r^k \leq p^k$  and  $s^k \leq q^k$ , then no disruption of the structure occurs. However, if  $r^k > p^k$  or  $s^k > q^k$ , then more fill in will occur. To illustrate this, consider the case that  $(r^k, s^k) = (5,4)$  in which case, after the external pivoting, the structure becomes

$$(2.15) \quad \left[ \begin{matrix} \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & & \\ \times & \times & \times & \times & & \\ \times & \times & \times & \times & & \\ \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times & \times & \times \end{matrix} \right] \left[ \begin{matrix} \times & \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times & \times \\ \times & & \times & \times & \times & \times & \times \\ \times & & & \times & \times & \times & \times \\ & & & & \times & \times & \times \\ & & & & & \times & \times \\ & & & & & & \times \end{matrix} \right].$$

The next step is to use internal pivoting to ensure that (2.9) holds. Let the largest product  $l_{1i}u_{i1}$  occur in element  $j$ . If  $j \leq \min(p^k, q^k) + 1$  then no essential change in structure occurs. In this case, the final step of the current stage is to operate with the

$B$  matrix to introduce the required zeros. In our example the  $B$  matrix has the form

$$(2.16) \quad B = \begin{bmatrix} \frac{l_{51}^k u_{14}^k}{A_{54}^k} & -\frac{l_{52}^k}{l_{51}^k} & \dots & -\frac{l_{56}^k}{l_{51}^k} & 0 & \dots & 0 \\ \frac{l_{51}^k u_{24}}{A_{54}^k} & 1 & & & & & \\ \vdots & & \ddots & & & & \\ \frac{l_{51}^k u_{44}}{A_{54}^k} & & & 1 & & & \\ 0 & & & & 1 & & \\ \vdots & & & & & \ddots & \\ 0 & & & & & & 1 \end{bmatrix}$$

(using the indexing for  $L^k$  and  $U^k$  prior to external pivoting and assuming no internal pivoting). The updated factors have the form

$$(2.17) \quad \begin{bmatrix} \times & 0 & 0 & 0 & 0 & 0 \\ \times & \times & \times & \times & + & + \\ \times & \times & \times & \times & + & + \\ \times & \times & \times & \times & + & + \\ \times & \times & \times & \times & \times & + \\ \times & \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times & \times & \times \end{bmatrix} \begin{bmatrix} \times & \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times & \times \\ 0 & + & + & \times & \times & \times & \times \\ 0 & + & + & + & \times & \times & \times \\ & & & & \times & \times & \times \\ & & & & & \times & \times \\ & & & & & & \times \end{bmatrix}$$

where ‘+’ denotes fill in. After removing the first row and column of these matrices, we obtain reduced matrices  $L^{k+1}U^{k+1}$  for the next stage. In general, the values of  $p^{k+1}$  and  $q^{k+1}$  for the next iteration are given by

$$(2.18) \quad p^{k+1} = \max(p^k, r^k) - 1, \quad q^{k+1} = \max(q^k, s^k) - 1.$$

If  $j > \min(p^k, q^k) + 1$  in the above, then a similar process takes place, but it is possible to arrange the computation to reduce the size of  $p^{k+1}$  or  $q^{k+1}$ . We have illustrated cyclic pivoting, but it is clear that the same results may be obtained using interchanges which require less rearrangement. In practice, pivot depth is small and the choice is unlikely to be critical.

Finally, we consider the operational cost of this updating algorithm. If there is no pivoting, then the entire update takes  $2n^2 + O(n)$  multiplications and additions. If the pivot depth is at most  $r^k \leq c, s^k \leq c$ , then the cost is no greater than  $2cn^2 + O(n)$  multiplications and additions. Thus, it can be seen that the cost of the update increases with the depth of pivot chosen but not very rapidly.

**3. Pivot Choice and Stability.** The algorithm described is only of practical use if it performs the desired rank one update in some modest multiple of  $n^2$  operations. In this section we derive criteria for pivot selection which will prove satisfactory not only when  $A^*$  is well-conditioned, but also when  $A^*$  loses rank to any extent. Consider the application of our method in exact arithmetic to a problem for which  $\text{rank}(A^*) = k < n$ . Because we always select nonzero pivot elements, it follows that the reduced matrix  $A^{k+1}$  is zero and only columns 1 through  $k$  of  $L$  and  $U^T$  contain significant information. If the matrix is nearly rank-deficient, then after  $k$  stages the

reduced matrix becomes small relative to  $A^*$ , which is a desirable property for stability purposes. Thus, we are motivated to choose pivots which are large in a certain sense. Unfortunately, it is not possible to judge whether a pivot is large without examining the whole reduced matrix (complete pivoting), which is unduly expensive. Therefore, we aim to examine relatively few pivots to find one which gives an acceptable bound on the error growth.

Our first step, therefore, is to derive a bound on error growth. Because the  $B$  matrix in (2.16) can be written as the product of simple triangular operators, the error analysis results obtained for such operators in [3] are relevant to this problem. The main result is that the error growth is proportional to growth in terms of the form  $l_{is}^k u_{sj}^k$  at any stage of the calculation. For any potential choice of  $B$  we have

$$(3.1) \quad L^+ = L^k B, \quad U^+ = B^{-1} U^k.$$

Therefore, a bound on any term  $|l_{iq}^+ u_{qj}^+|$  is given by

$$(3.2) \quad |l_{iq}^+ u_{qj}^+| = \left| \sum_{t=1}^m l_{it} B_{tq} \right| \left| \sum_{p=1}^m B_{qp}^{-1} u_{pj} \right| \\ \leq \max_t (|B_{tq}|) \max_p (|B_{qp}^{-1}|) \sum_{t=1}^m \sum_{p=1}^m |l_{it} u_{pj}| \quad \forall q,$$

where we suppress superscript  $k$  for simplicity, and write  $m = n - k + 2$ . We therefore aim to choose a pivot element for which the quantity  $\max_q(\theta_q)$  is small, where

$$(3.3) \quad \theta_q = \max_t (|B_{tq}|) \max_p (|B_{qp}^{-1}|).$$

This is the product of the maximum modulus elements in column  $q$  of  $B$ , and row  $q$  of  $B^{-1}$ . The quantity  $\theta_q$  is cheap to calculate, and though it does not give a particularly strict bound on growth, it does work well in practice.

We now consider how  $\max_q(\theta_q)$  can be calculated. To simplify the presentation we write any possible pivot element as

$$(3.4) \quad \alpha_{rs} = A_{rs}^k,$$

and we can assume that internal pivoting is unnecessary so that

$$(3.5) \quad |l_{r1} u_{1s}| \geq |l_{ri} u_{is}|, \quad i = 2, \dots, m.$$

From (2.16) it follows that

$$(3.6) \quad B^{-1} = \begin{bmatrix} 1 & \frac{l_{r2}}{l_{r1}} & \frac{l_{r3}}{l_{r1}} & \dots & \frac{l_{rm}}{l_{r1}} \\ -\frac{l_{r1} u_{2s}}{\alpha_{rs}} & 1 - \frac{l_{r2} u_{2s}}{\alpha_{rs}} & -\frac{l_{r3} u_{2s}}{\alpha_{rs}} & \dots & -\frac{l_{rm} u_{2s}}{\alpha_{rs}} \\ -\frac{l_{r1} u_{3s}}{\alpha_{rs}} & -\frac{l_{r2} u_{3s}}{\alpha_{rs}} & 1 - \frac{l_{r3} u_{3s}}{\alpha_{rs}} & \dots & -\frac{l_{rm} u_{3s}}{\alpha_{rs}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\frac{l_{r1} u_{ms}}{\alpha_{rs}} & -\frac{l_{r2} u_{ms}}{\alpha_{rs}} & -\frac{l_{r3} u_{ms}}{\alpha_{rs}} & \dots & 1 - \frac{l_{rm} u_{ms}}{\alpha_{rs}} \end{bmatrix},$$

and hence for  $q = 1$ ,

$$(3.7) \quad \theta_1 = \frac{1}{|\alpha_{rs}|} \max_p (|l_{rp}|) \max_t (|u_{ts}|),$$

and for  $q > 1$ ,

$$(3.8) \quad \theta_q = \max \left( 1, \left| \frac{l_{rq}}{l_{r1}} \right| \right) \max \left[ \max_{\substack{p \\ p \neq q}} \left( \left| \frac{l_{rp} u_{qs}}{\alpha_{rs}} \right| \right), \left| 1 - \frac{l_{rq} u_{qs}}{\alpha_{rs}} \right| \right].$$

For  $q > 1$  we can express the right-hand side of (3.8) as a maximum of four terms which we will consider separately:

$$(3.9) \quad \begin{aligned} \text{(i)} \quad & 1 \cdot \max_{\substack{p \\ p \neq q}} \left( \left| \frac{l_{rp} u_{qs}}{\alpha_{rs}} \right| \right) \leq \theta_1, \\ \text{(ii)} \quad & \left| \frac{l_{rq}}{l_{r1}} \right| \max_{\substack{p \\ p \neq q}} \left( \left| \frac{l_{rp} u_{qs}}{\alpha_{rs}} \right| \right) = \left| \frac{l_{rq} u_{qs}}{l_{r1} u_{1s}} \right| \max_{\substack{p \\ p \neq q}} \left( \left| \frac{l_{rp} u_{1s}}{\alpha_{rs}} \right| \right) \leq \theta_1, \\ \text{(iii)} \quad & 1 \cdot \left| 1 - \frac{l_{rq} u_{qs}}{\alpha_{rs}} \right| \leq 1 + \theta_1, \quad \text{and} \\ \text{(iv)} \quad & \left| \frac{l_{rq}}{l_{r1}} \right| \left| 1 - \frac{l_{rq} u_{qs}}{\alpha_{rs}} \right| \leq \left| \frac{l_{rq} u_{1s}}{\alpha_{rs}} \right| \left| \frac{\alpha_{rs}}{l_{r1} u_{1s}} \right| + \left| \frac{l_{rq} u_{qs}}{l_{r1} u_{1s}} \right| \left| \frac{l_{rq} u_{1s}}{\alpha_{rs}} \right| \\ & \leq (\min(p^{k+1}, q^{k+1}) + 2) \theta_1, \end{aligned}$$

where we have used the fact that

$$(3.10) \quad \begin{aligned} |\alpha_{rs}| &= \left| \sum_{i=1}^m l_{ri} u_{is} \right| = \left| \sum_{i=1}^{\min(p^{k+1}, q^{k+1})+1} l_{ri} u_{is} \right| \\ &\leq (\min(p^{k+1}, q^{k+1}) + 1) |l_{r1} u_{1s}|. \end{aligned}$$

Therefore, it follows that

$$(3.11) \quad \max_q (\theta_q) \leq \max(1 + \theta_1, (\min(p^{k+1}, q^{k+1}) + 2) \theta_1),$$

and we can deduce that in all important cases (i.e., large growth)

$$(3.12) \quad \max_q (\theta_q) \leq c \theta_1,$$

where  $c$  is some small integer. We therefore use  $\theta_1$  to give an indication of the growth incurred by the choice of a particular pivot. From (3.7) it is clear that  $\theta_1$  is a function of  $r$  and  $s$ , so subsequently we suppress the subscript 1 and write

$$(3.13) \quad \theta(r, s) = \frac{1}{|\alpha_{rs}|} \max_p (|l_{rp}|) \max_t (|u_{ts}|).$$



We now have to decide how to use  $\theta(r, s)$  to choose a pivot element  $\alpha_{rs}$ . If we merely search for the element with smallest associated  $\theta(r, s)$ , this may give rise to a permutation which severely disrupts the structure of the factors and we lose efficiency. We therefore require a strategy which considers ‘shallow’ pivots first and only moves on to deeper pivots if no ‘satisfactory’ pivot is found. The criterion for accepting pivots should become less severe as depth increases. One particular strategy is suggested in the next section, where it is shown that the algorithm is relatively insensitive to the search pattern used.

In choosing  $\theta(r, s)$  we have so far restricted ourselves to the particular  $B$  matrix given in (2.12), derived from elementary operators. It is also possible to eliminate the desired elements by using a more general family of  $B$  matrices. The conditions that a suitable matrix must satisfy are

$$(3.14) \quad [X_{11}, X_{12}, \dots, X_{1,n+1}]B = [X_{11}^+, 0, \dots, 0]$$

and

$$(3.15) \quad B^{-1}[Y_{11}, Y_{21}, \dots, Y_{n+1,1}]^T = [Y_{11}^+, 0, \dots, 0]^T.$$

Because the right-hand sides of both (3.14) and (3.15) are multiples of the first row and column of the identity matrix, and  $B^{-1}B = I$ , we can deduce that the first row of  $B^{-1}$  must be  $1/X_{11}^+[X_{11}, X_{12}, \dots, X_{1,n+1}]$  and the first column of  $B$  must be  $1/Y_{11}^+[Y_{11}, Y_{21}, \dots, Y_{n+1,1}]^T$ . We can therefore calculate  $\theta(r, s)$  for the general matrix using the fact that  $X_{11}^+ Y_{11}^+ = \alpha_{rs}$ . This yields

$$(3.16) \quad \theta(r, s) = \frac{1}{|X_{11}^+ Y_{11}^+|} \max_t (|X_{1t}|) \max_p (|Y_{p1}|) = \frac{1}{|\alpha_{rs}|} \max_t (|X_{1t}|) \max_p (|Y_{p1}|),$$

and if we take  $X = L^k$  and  $Y = U^k$ , we obtain the same  $\theta(r, s)$  as for the elimination operator. Thus  $\theta(r, s)$  is independent of the choice of  $B$  matrix. This suggests that it is unlikely that any advantage in stability can be obtained by going to a more general family of  $B$  matrices, as against the more efficient triangular operator formulation of Section 2.

**4. Numerical Results and Discussion.** All the results in this section have been calculated using a DEC 10 computer with relative precision  $\epsilon = 2^{-27} \approx 7.5_{10^{-9}}$ . To test the algorithm we have produced a set of test problems of the following form. We take a ‘‘target matrix’’ and alter it directly using a series of  $m$  random rank one changes generated using random numbers in  $[-1, 1]$ . This altered matrix is regarded as the ‘‘initial matrix’’ and its triangular factors are calculated using Gaussian elimination with partial pivoting. Then we repeat the original series of rank one changes with the sign reversed, and update the factors subject to these changes using the algorithm in Section 2. After this calculation is completed, we obtain factors of the target matrix in the form (2.4). We consider four classes of target matrix, which are:

(i) Random rank  $k$  matrix—This is a matrix constructed as the sum of  $k \leq n$  rank one terms generated using random numbers in  $[-1, 1]$ . Results for this class of matrix are given as the average of 10 different problems.

(ii) Hilbert matrix—

$$(4.1) \quad H_{ij} = 1/(i + j - 1).$$

In this problem, the initial matrix is well-conditioned and the target matrix is very ill-conditioned, being numerically singular to this precision for any dimension greater than about 12.

(iii) Pascal matrix—

$$(4.2) \quad \begin{aligned} S_{i1} &= 1, & 1 \leq i \leq n, \\ S_{ij} &= 0, & i < j, \\ S_{ij} &= S_{i-1,j} + S_{i-1,j-1}, & j \leq i, i > 1, j > 1. \end{aligned}$$

This target matrix is a lower triangular matrix whose rows consist of the rows of a Pascal triangle. The main features of this matrix are that it is extremely ill-conditioned and has very large elements, e.g., for  $n = 20$  the largest element is  $1.8_{10}^5$ .

(iv) Wilkinson matrix—

$$(4.3) \quad \begin{aligned} W_{ij} &= -a, & i > j, 1 < j < n, \\ W_{ij} &= a, & j = i, j = n, 1 \leq i \leq n, \\ W_{ij} &= 0, & \text{otherwise.} \end{aligned}$$

This matrix is essentially that given by Wilkinson [5] as an example of a matrix for which maximum growth occurs when Gaussian elimination with partial pivoting is used. For all results we have used the value  $a = 1.1$ .

The results are presented in terms of certain quantities:

EMAX—maximum error in the factors of  $A^*$  defined by

$$(4.4) \quad \text{EMAX} = \max_{i,j} |E_{ij}|, \quad E = P^*A^*Q^{*T} - L^*U^*.$$

AVDEPTH—Average depth of pivots chosen, taken over all pivot choices. Since the disruption of the Hessenberg structure is roughly proportional to  $r^k + s^k$  we define the “depth” of a pivot to be

$$(4.5) \quad d^k = r^k + s^k - 1.$$

MAXDEPTH—Deepest pivot selected during the updating process.

OPCOUNT—Total number of operations (that is one ‘+’ and one ‘\*’) taken by the pivot search and the elimination procedures to complete the updates.

OPTIME—We wish to compare the cost of updating the factors with that of calculating them *ab initio*. Since the organization of the latter calculation is considerably simpler, comparing operation counts does not tell the entire story. Hence, for some results, we also present the CPU time taken by the algorithm in milliseconds. All the above quantities (except EMAX) are taken over all  $m$  updates, where  $m$  is defined at the start of this section.

Our first test aims to obtain some experience of the effect of pivot depth on stability and efficiency for this class of test problems. It uses the above test problems with  $n = 20$ ,  $m = 50$  and random matrices of full rank. These problems are solved using the algorithm with the simple pivot test in which the pivot  $\alpha_{r,k_s^k}$  is selected

where

$$(4.6) \quad \theta(r^k, s^k) \leq \theta(r, s) \quad \forall r, s \text{ such that } r + s - 1 \leq \text{MAXDEPTH},$$

where we have preset a value for MAXDEPTH. The results from these tests are tabulated in Table 4.1. From these results it can be seen that although a certain freedom in pivoting greatly improves the accuracy of the results, a break even point is quickly reached and deeper pivot searches do not return better results. Indeed, there is some evidence in the MAXDEPTH = 40 entries that the extra computation associated with deep pivots causes errors to increase. Also the prohibitive cost of comparing all possible pivots is demonstrated by the fact that the ratio

$$(\text{AVDEPTH for MAXDEPTH} = 40 / \text{AVDEPTH for MAXDEPTH} = 4) \approx 2.5,$$

whereas

$$(\text{OPCOUNT for MAXDEPTH} = 40 / \text{OPCOUNT for MAXDEPTH} = 4) \approx 15.$$

We are therefore motivated to derive a pivot selection criterion which discriminates against deep pivots, whilst allowing them in very occasional extreme circumstances. After evaluating several possibilities, we have selected a heuristic based on two constants,  $\beta$  and  $\gamma$ . A pivot  $\alpha_{rs}$  is judged to be 'satisfactory' if either

$$(4.7) \quad \begin{aligned} & \text{(a) } r < p^k, s < q^k \text{ and } \theta(r, s) \leq \beta, \text{ or} \\ & \text{(b) } \theta(r, s) \leq \beta \cdot \gamma^{\text{search depth}} \text{ and } r + s - 1 \leq \text{search depth.} \end{aligned}$$

Test (a) allows any pivot for which no additional fill in is generated, and the associated growth bound is no worse than  $\beta$ . In test (b) search depth is initially one, and is then increased until either a satisfactory pivot is discovered, or all possible pivots have been tested. If no element satisfies test (4.7) we terminate if the reduced matrix is zero, and if not, we select the element with minimum  $\theta(r, s)$ .

TABLE 4.1  
Effect of 50 updates on 20 × 20 matrices

Target Matrix		Random rank $n$	Hilbert	Pascal	Wilkinson
EMAX	<i>MAXDEPTH</i>				
	1	$9.11_{10^{-4}}$	$1.41_{10^{-4}}$	$2.49_{10^{-2}}$	$1.56_{10^1}$
	2	$4.22_{10^{-6}}$	$3.53_{10^{-6}}$	$3.91_{10^{-3}}$	$4.28_{10^{-1}}$
	3	$3.27_{10^{-6}}$	$1.53_{10^{-6}}$	$1.46_{10^{-2}}$	$1.76_{10^{-6}}$
	4	$3.45_{10^{-6}}$	$1.60_{10^{-6}}$	$5.43_{10^{-3}}$	$2.28_{10^{-6}}$
	40	$5.87_{10^{-5}}$	$3.98_{10^{-5}}$	$2.37_{10^{-2}}$	$9.15_{10^{-6}}$
OPCOUNT $\times 10^{-2}$	1	438	438	438	438
	2	552	575	579	565
	3	790	825	796	783
	4	1012	1035	1061	1017
	40	15540	15843	15857	15636
	AVDEPTH	1	1.000	1.000	1.000
2		1.430	1.517	1.539	1.469
3		2.185	2.324	2.206	2.177
4		2.631	2.712	2.720	2.636
40		7.007	7.313	7.502	7.214

TABLE 4.2  
Effect of different parameter selections in test (4.7)

Target Matrix			Random rank $n$	Hilbert	Pascal	Wilkinson
	$\beta$	$\gamma$				
	EMAX	1.0	1.0	$8.68_{10^{-6}}$	$3.96_{10^{-6}}$	$1.29_{10^{-2}}$
1.0		1.5	$3.20_{10^{-6}}$	$1.97_{10^{-6}}$	$5.37_{10^{-3}}$	$3.16_{10^{-6}}$
1.5		1.3	$3.65_{10^{-6}}$	$3.63_{10^{-6}}$	$9.82_{10^{-3}}$	$1.28_{10^{-6}}$
2.0		1.5	$4.26_{10^{-6}}$	$3.17_{10^{-6}}$	$7.81_{10^{-3}}$	$4.29_{10^{-6}}$
2.0		3.0	$7.07_{10^{-6}}$	$8.19_{10^{-6}}$	$8.33_{10^{-3}}$	$5.76_{10^{-6}}$
OPCOUNT $\times 10^{-2}$	1.0	1.0	3877	4955	4210	4623
	1.0	1.5	558	1033	559	582
	1.5	1.3	545	1138	592	557
	2.0	1.5	496	801	515	533
	2.0	3.0	480	580	475	490
OPTIME $\times 10^{-1}$	1.0	1.0	1244	1610	1436	1512
	1.0	1.5	134	320	136	147
	1.5	1.3	121	339	150	119
	2.0	1.5	105	229	109	111
	2.0	3.0	102	143	102	101
AVDEPTH	1.0	1.0	4.895	5.676	4.724	5.247
	1.0	1.5	1.431	2.105	1.433	1.484
	1.5	1.3	1.444	2.192	1.531	1.468
	2.0	1.5	1.265	1.817	1.322	1.381
	2.0	3.0	1.175	1.404	1.165	1.222
MAXDEPTH	1.0	1.0	25.9	28	24	23
	1.0	1.5	5.3	14	8	5
	1.5	1.3	6.0	17	8	6
	2.0	1.5	4.6	13	6	6
	2.0	3.0	3.3	7	4	3

The results in Table 4.2 are obtained by solving the same problem set as for Table 4.1 using a selection of  $(\beta, \gamma)$  pairs. From these results we can observe that EMAX is relatively insensitive to the parameter choice. The choice  $(\beta, \gamma) = (1, 1)$  usually causes extensive pivot searches to occur, because  $\gamma = 1$  gives no weight to deeper pivots. The error EMAX is seen to be satisfactory in all cases. (It should be noted that the errors in the Pascal matrix are substantially larger because they are relative to matrix elements with maximum size  $1.8_{10^5}$ . The relative error for this problem is in fact  $\approx 10^{-5}$ , which is quite reasonable for an error over 50 complete rank one updates.) With regard to OPTIME and OPCOUNT, it can be seen that except for the Hilbert matrix problem, the cost of pivoting over a sequence of updates is very small. In particular, a comparison with the results for Bennett's algorithm (that is, MAXDEPTH = 1 in Table 4.1) shows that the introduction of pivoting into the algorithm increases the computational cost by only a modest factor. For the Hilbert matrix problem, and also for other highly rank deficient problems, we observe that there is an extra cost, and this is due to the selection of deeper pivots (see AVDEPTH). However, this deeper pivoting is not necessary to retain small errors, which suggests that the bound (3.2), on which  $\theta(r, s)$  is based, is over-pessimistic. Comparing Tables 4.1 and 4.2 indicates that for these test problems a maximum

depth of three would be sufficient to ensure stability. Thus, we prefer to impose a limit on depth, and inform the user when this limit is attained.

Although the choice is not critical, we suggest on the basis of Tables 4.1 and 4.2 that the algorithm is implemented with default values  $(\beta, \gamma) = (2.0, 1.5)$  and  $\text{MAXDEPTH} \leq 5$ . Using these values we now investigate the dependence of the errors and costs of the algorithm with respect to  $m$  and  $n$ . We use the same set of problems as above with  $m \leq 50$ , but consider the range of values  $n = 10, 20, 50$ . We have also added a class of Random rank 5 problems to the experiment. The results of these experiments are given in Table 4.3. It can be seen that no untoward deterioration occurs as  $n$  is increased, and that the errors appear to increase at a rate just slightly faster than the  $\sqrt{m}$  rate which would be expected from statistical considerations. The only exception to this is the  $50 \times 50$  Wilkinson matrix for which the error growth is slightly faster over the last few updates. Inspection of AVDEPTH over these updates reveals a sharp rise as MAXDEPTH is repeatedly attained. This suggests that the bound on MAXDEPTH should be relaxed, and if we increase MAXDEPTH to a value of 7 the final EMAX is reduced to  $2.6_{10^{-5}}$ . For the extremely rank deficient matrices it can be seen that the bound on MAXDEPTH is frequently attained, but the AVDEPTH shows no sharp increase and no untoward

TABLE 4.3  
Effect on algorithm of varying  $n$  and  $m$

$n$		10			20			50		
Iteration ( $m$ )		1	5	50	1	5	50	1	5	50
Random rank $n$	EMAX	$1.5_{10^{-7}}$	$2.3_{10^{-7}}$	$1.4_{10^{-6}}$	$3.3_{10^{-7}}$	$6.3_{10^{-7}}$	$4.2_{10^{-6}}$	$1.0_{10^{-6}}$	$3.1_{10^{-6}}$	$2.3_{10^{-5}}$
	OPCOUNT	2.4	12	133	9.1	46	496	55	278	3056
	OPTIME	0.3	3.1	35	1.9	9.3	105	9.3	47	546
	AVDEPTH	1.055	1.053	1.208	1.116	1.119	1.267	1.135	1.135	1.399
	MAXDEPTH	1.3	1.9	3.3	2.0	2.6	4.4	2.5	3.3	5.0
Random rank 5	EMAX	$1.6_{10^{-7}}$	$2.9_{10^{-7}}$	$1.1_{10^{-6}}$	$3.6_{10^{-7}}$	$6.5_{10^{-7}}$	$3.0_{10^{-6}}$	$6.2_{10^{-7}}$	$2.4_{10^{-6}}$	$9.7_{10^{-6}}$
	OPCOUNT	2.4	12	138	9.0	45	577	56	291	4322
	OPTIME	0.3	2.6	40	1.8	9.1	142	11	52	913
	AVDEPTH	1.055	1.069	1.272	1.100	1.091	1.516	1.225	1.307	2.187
	MAXDEPTH	1.3	2.0	4.9	1.8	2.3	5.0	2.7	3.6	5.3
Hilbert	EMAX	$1.2_{10^{-7}}$	$1.8_{10^{-7}}$	$1.7_{10^{-6}}$	$3.8_{10^{-7}}$	$8.3_{10^{-7}}$	$3.2_{10^{-6}}$	$6.9_{10^{-7}}$	$1.9_{10^{-6}}$	$5.4_{10^{-6}}$
	OPCOUNT	2.4	12	151	11	46	612	56	287	4630
	OPTIME	0	1.8	45	1.9	9.1	139	9.0	47	952
	AVDEPTH	1.000	1.067	1.422	1.263	1.095	1.614	1.265	1.331	2.333
	MAXDEPTH	1	2	5	3	3	5	3	4	5
Pascal	EMAX	$9.5_{10^{-7}}$	$1.9_{10^{-6}}$	$3.8_{10^{-6}}$	$1.1_{10^{-3}}$	$2.9_{10^{-3}}$	$7.8_{10^{-3}}$	-	-	-
	OPCOUNT	2.4	12	132	9.1	46	513	-	-	-
	OPTIME	0	1.8	37	1.8	8.9	109	-	-	-
	AVDEPTH	1.000	1.044	1.204	1.053	1.126	1.306	-	-	-
	MAXDEPTH	1	2	4	2	3	5	-	-	-
Wilkinson	EMAX	$2.4_{10^{-7}}$	$2.4_{10^{-7}}$	$9.1_{10^{-7}}$	$2.2_{10^{-7}}$	$4.8_{10^{-7}}$	$4.3_{10^{-6}}$	$5.1_{10^{-7}}$	$2.1_{10^{-6}}$	$3.2_{10^{-4}}$
	OPCOUNT	2.4	12	131	9.4	45	528	52	277	3477
	OPTIME	1.8	5.3	33	1.8	7.2	114	9.0	44	647
	AVDEPTH	1.000	1.044	1.173	1.263	1.095	1.378	1.061	1.192	1.663
	MAXDEPTH	1	2	4	2	2	5	2	3	5

(All entries for OPCOUNT are  $\times 10^{-2}$  and for OPTIME  $\times 10^{-1}$ )

error growth occurs. The operation count increases at a rate of  $\approx 2.6n^2$  for all well-behaved problems, and rises to a maximum observed value of  $3.7n^2$  for the  $50 \times 50$  Hilbert matrix. In all cases the CPU time taken by the algorithm is roughly proportional to the operation count.

We can also use these results to obtain a comparison between the errors and costs of our updating algorithm as against those for Gaussian elimination with partial pivoting. For this purpose we have factorized the Hilbert matrices and the Random full rank matrices using Gaussian elimination. These results are presented in Table 4.4. The quantity EMAX is defined as in Table 4.1, but OPCOUNT and OPTIME are presented as a multiple of the average cost of performing the corresponding rank one update in Table 4.3. The results show that for well-conditioned matrices a single update of a  $10 \times 10$  matrix is approximately the same cost as the *ab initio* calculation. (This agrees with a direct comparison between the operation counts, i.e.,  $2.6n^2 = \frac{1}{3}n^3 \Rightarrow n \approx 8$ .) As  $n$  increases the updating process becomes steadily more efficient. For the ill-conditioned Hilbert matrix the advantage is marginally smaller due to the increased pivot depth.

We can also compare our algorithm with the algorithms proposed in [4] for updating orthogonal factors. The more efficient of these algorithms takes approximately  $8n^2$  operations to complete an equivalent update, which indicates that considerable savings can be made by using triangular factors together with the above algorithm.

When coding our algorithm the storage requirements are firstly  $n^2 + 3n$  locations for the  $LU$  factors together with  $x$  and  $y$ , and secondly extra locations to store any subsequent fill in that occurs. With no limit on the pivot depth this would require an additional  $n^2$  locations but if we preset a maximum pivot depth then only a very small amount of extra storage is required. (E.g., if  $\text{MAXDEPTH} = 5$  then an extra 9 storage locations must be reserved.) Another problem that must be considered when coding is that of poor scaling, which can lead to underflow or overflow. To avoid this, it is desirable to rescale the  $x$  and  $y$  vectors before performing the update so that they are of comparable size.

As can be seen from Table 4.3 the errors in the updated factors are cumulative and therefore it will eventually be necessary to reinvert to get back to the original  $A$  matrix. However, in many applications—especially nonlinear problems—the calculated factors will be used to calculate some form of search direction and hence

TABLE 4.4  
*Ab initio calculation vs. updated factors*

$n$	Problem	Random rank $n$	Hilbert
10	EMAX	$3.28_{10^{-8}}$	$3.72_{10^{-9}}$
	OPCOUNT	1.07	0.94
	OPTIME	1.17	1.11
20	EMAX	$1.06_{10^{-7}}$	$3.72_{10^{-9}}$
	OPCOUNT	2.49	2.02
	OPTIME	2.45	1.89
50	EMAX	$4.41_{10^{-7}}$	$3.72_{10^{-9}}$
	OPCOUNT	6.61	4.36
	OPTIME	6.19	3.55

determine the next rank one change. This introduces a self-correcting mechanism into the algorithm and makes reinversion unnecessary.

If the original matrix  $X$  is altered by a rank  $k$  change, where  $k > 1$ , and we wish to obtain the new factors in a single update, as opposed to  $k$  rank one updates, then this can be done using a similar approach to that used in Section 2. The first step is to use the initial factors and the rank  $k$  update to define matrices  $L^1$  and  $U^{1T} \in \mathbf{R}^{n \times (n+k)}$ , where  $L^1$  and  $U^1$  are Hessenberg matrices with  $k$  elements above and below the diagonal. These matrices can be reduced to triangular form in exactly the manner described in Section 2 for the general matrices  $X$  and  $Y$ .

In this paper we have described a practical algorithm for updating triangular factors of matrices which requires considerably less computational effort than the updating of corresponding orthogonal factors. Given the accuracy of the results obtained using this algorithm, we feel that there are many applications in which this additional efficiency can be exploited.

**Acknowledgments.** We wish to thank the Royal Commission for the Exhibition of 1851 and the New Zealand University Grants Committee for their financial support.

Department of Mathematical Sciences  
University of Dundee  
Dundee DD1 4HN, Scotland

1. J. M. BENNETT, "Triangular factors of modified matrices," *Numer. Math.*, v. 7, 1965, pp. 217–221.
2. J. W. DANIEL, W. B. GRAGG, L. KAUFMAN & G. W. STEWART, "Reorthogonalization and stable algorithms for updating the Gram-Schmidt QR factorization," *Math. Comp.*, v. 30, 1976, pp. 772–795.
3. R. FLETCHER & S. P. J. MATTHEWS, *Stable Modification of Explicit LU Factors for Simplex Updates*, Department of Mathematics Report NA/64, Dundee University, 1983.
4. P. E. GILL, G. H. GOLUB, W. MURRAY & M. A. SAUNDERS, "Methods for modifying matrix factorizations," *Math. Comp.*, v. 28, 1974, pp. 505–535.
5. J. H. WILKINSON, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, 1965.