A note on the least squares solution of a band system of linear equations by Householder reductions

By J. K. Reid*

The straightforward use of Householder reductions for the solution of a band system of over-determined linear equations is wasteful of machine time. In this note a modification that greatly reduces the number of arithmetic operations is proposed.

We consider here the least squares solution of the system

$$A\,x = b \tag{1}$$

where A is a rectangular $m \times n(m > n)$ band matrix of bandwidth r. By this we mean that if $a_{ij} \neq 0$ and $a_{ik} \neq 0$ then $|j - k| \leqslant r$ and that if a_{ij} and a_{kl} are the first non-zero elements in the ith and kth rows with $i < k$ then $j \leqslant l$. By the use of suitable row permutations the latter condition can always be made to hold. Systems of equations of this form arise, for example, in the least squares fitting of a spline of given degree and given knot positions to a function defined on a set of discrete points (M. J. D. Powell, private communication).

The formation of A^TA and A^Tb followed by the solution of

$$A^TA\,x = A^Tb \tag{2}$$

involves

$$[\tfrac{1}{2}mr(r+1) + mr + \tfrac{1}{3}n(r-1)(r+2) + n(2r-1) + o(m) + o(n)]$$

multiplications if full advantage is taken of the band structure of A and the fact that A^TA is band and symmetric positive definite. However, since the matrix A^TA is often ill-conditioned a more satisfactory procedure is to use Householder reduction of A to upper triangular form,

$$QA = U = \begin{bmatrix} U_1 \\ \cdots \\ 0 \end{bmatrix}\begin{array}{l} n \times n \\ \\ (m-n) \times n \end{array} \tag{3}$$

so that (2) takes the form

$$U^TU\,x = U^TQb \tag{4}$$

which reduces to the form

$$U_1\,x = Q_1b, \tag{5}$$

if Q_1 denotes the first n rows of Q. This procedure is described, for example, by Businger and Golub (1965). This is inefficient for band matrices, however, since approximately rnm multiplications are required in the triangularization.

This difficulty may be overcome by suitably subdividing the Householder reductions. Group the rows according to the column in which the first non-zero element occurs; thus the jth block consists of those rows whose first non-zero element is in the jth column. Householder transformations are performed on the pivotal row and the rows of a single block. When the pth row is pivotal the blocks involved are $q, q+1, \ldots, p$ where $q = \max(1, p-r+1)$ and these blocks must be taken in this order. The form of the matrix after such a succession of reductions is as follows: the first p rows are band upper triangular, and a row from the remainder is zero if it belongs to one of blocks 1 to $p-r+1$ (provided $p > r-1$), has at most $r-s$ non-zero elements if it belongs to block $p-s+1$ ($s = 1, 2, \ldots, p-q+1$) and is unchanged, having at most r non-zeros, if it belongs to one of blocks $p+1$ onwards. We show in **Fig. 1** an example with $m=30$, $n=11$ and $r=4$ both before reduction and after the stage $p=6$. Each $\times$ is in general non-zero, a zero is represented by ○ or a space, and the blocks are shown separated by dotted lines. The reason for the inefficiency of the straightforward use of Householder reductions is that at the stage after the pth row is pivotal the rows of blocks 1 to p, other that rows 1 to p, will in general have $r-1$ non-zeros. In the example of Fig. 1 all elements that are zero in the modified procedure but are in general non-zero in the straightforward procedure are marked ○.

We note that the two procedures are essentially identical, for if A has rank n and the matrices obtained in the new procedure are Q' and U' with first n rows Q_1' and U_1' then $Q_1' = DQ_1$ and $U_1' = DU_1$ where D is a diagonal matrix with diagonal elements ± 1. To prove this we observe from (3) that

$$A = Q^TU = Q_1^TU_1 = Q_1'^TU_1'. \tag{6}$$

U_1 is non-singular since A has rank n so that from (6) we deduce that

$$Q_1'AU_1^{-1} = Q_1'Q_1^T = U_1'U_1^{-1}. \tag{7}$$

The matrix of (7) is orthogonal and upper-triangular so must be diagonal with diagonal elements ± 1 and is the matrix D mentioned above.

The total number of multiplications in the modified procedure is

$$[r(r+1)(m+\tfrac{2}{3}n) + rm + rn(4+s) + o(m) + o(n)]$$

if each evaluation of a square root takes s multiplications. This is about twice as much work as in the straightforward use of (2), an increase similar to that for full matrices.

* *Mathematics Division, University of Sussex, Falmer, Brighton, Sussex.*

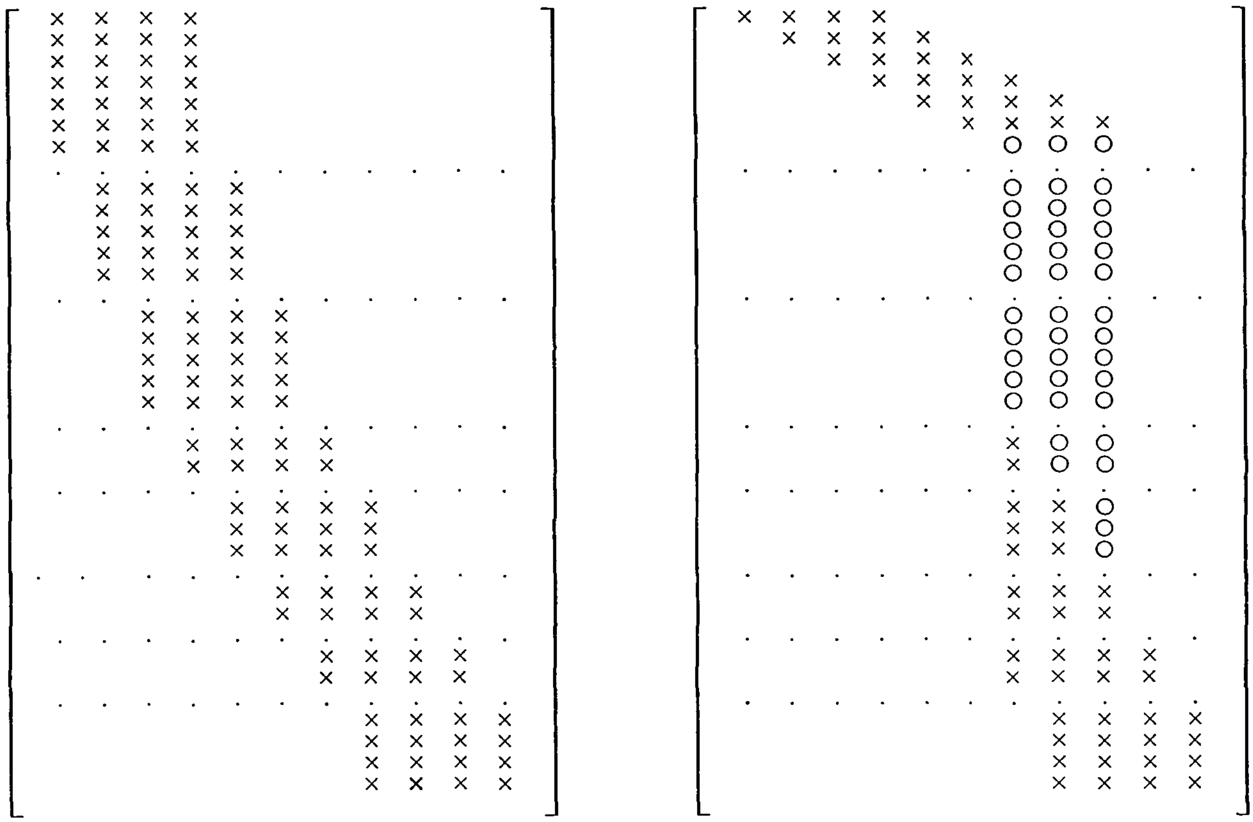

Fig. 1. The form of a band matrix before reduction is commenced and at a typical intermediate stage

Reference

BUSINGER, P., and GOLUB, G. H. (1965). "Linear Least Squares Solutions by Householder Transformations", *Num. Math.*, Vol. 7, p. 269.

Book Review (*Continued from p. 187*)

In more complex systems, where the selective information is greater than for the binary choice "fight or flee" this dichotomy may not be so easy to define. When the other essential computer component, the clock, becomes really important there is no longer a sharp distinction between excitation and inhibition. As far as Young reports Octopus is not a good judge of elapsed time, while vertebrates and particularly man are notable for their accuracy even without adventitious aids, and even in states of diminished awareness or distraction. Some of the electrochemical processes in higher brains look very like the unwinding and striking of an alarm clock, and an alarm clock is both excitatory—it wakes you up—and inhibitory—it avoids your awakening too soon.

The final conjecture in this admirable collection may apply not merely to us who regret that the arrow of time points to the grave, but also to those who are more prosaically concerned with the breeding and grooming of more sophisticated computers—the greater our capacity for learning the longer we must live in order to learn. In a sense men never grow up; if our great brains are not to destroy our species "We must become even more like little children".

W. GREY WALTER