A DIRECT METHOD FOR SOLVING LINEAR ALGEBRAIC EQUATIONS

E. E. Osborne

This talk concerns a method for solving the set of simultaneous linear equations

$$(A - \alpha I)x = b,$$

in which the n x n matrix A, the scaler α, the vector b and the unknown vector x can either be all single precision or all double precision. The bulk of the arithmetic involved is single precision. The computing time in either case is comparable to that of other single precision direct methods.

The procedure combines Gaussian elimination with an automatic correcting feature, plus iterations to improve the accuracy. The correcting feature involves a technique very similar to one used for matrix pre-conditioning [K. Eisemann, Removal of Ill-conditioning for Matrices, QUART. APPL. MATH., 15 (1957), pp. 225-230].

The term "in full precision" shall mean double precision if the problem is double precision and single precision if the problem is single precision.

A key subroutine is one which performs precise accumulation of products. It accumulates double precision products of single precision numbers. The resulting sum is accompanied by an integer telling how many digits have cancelled in the accumulation.

A theorem (at the present time is being refined) could be stated concerning matrices L and U such that LA = U with U triangular. The theorem says, in effect, that single precision accuracy suffices for L and U provided that pains have been taken to maintain significant digits except possibly in certain rare circumstances. The method:

a. Triangularization.

Initialize the single precision temporary matrix

$$T^{(r)} = (U^{(r)}, L^{(r)})$$

by taking $U^{(1)}$ to be the more significant half of $A - \alpha I$ and $L^{(1)} = I$. Then for $r = 2, \ldots, n$ carry out the steps.

i. In Single Precision obtain $T^{(r)}$ from $T^{(r-1)}$ by performing in the usual way the operations upon the rows numbered r-1 to n so as to produce zeros below the diagonal element in the rth column of $T^{(r)}$.

ii. Re-compute the element $U^{(r)}_{rr}$. If it suffers a cancellation of more than a prescribed amount the entire rth row of $U^{(r)}$ is re-computed in Double Precision using the rth row of $L^{(r)}$ and the full precision matrix $A - \alpha I$. Otherwise return to step i for the (r + 1)st step.

iii. If the rth row has been re-computed, eliminate its elements in columns 1, 2, ..., r-1 by subtracting the appropriate multiples of the 1st, 2nd, ...(r-1)st rows respectively of $T^{(r)}$ from the rth row. This may be done in single precision arithmetic. Return to i for the (r+1st step.

This process yields finally the matrices $L = L^{(n)}$ and $U = U^{(n)}$ which are then employed in the following way.

b. Solving the modified set of equations.

Take $R_o = b$ and $x_o = 0$ then;

i. Form $\hat{R}_i = LR_i$. This should be done using precise accumulation of products. But R_i is rounded to single precisions.

ii. Solve for the single precision δx_i in

$$U\delta x_i = \hat{R}_i$$

iii. Form, in full precision, the new estimate

$$x_{i+1} = x_i + \delta x_i$$

iv. Form, using double precision and with full precision $\hat{A}$ the residual

$$R_{i+1} = b - \hat{A} x_{i+1}$$

v. Test for convergence.

Results obtained thus far indicate that the program fulfills its expected performance.