Note

Evaluation of $m^\dagger(h - \epsilon)^{-1}m$ for a Large Hermitian Matrix h

The matrix expression indicated in the title occurs in linear expansion methods for bound state or scattering solutions of Schrödinger's equation. A method of evaluation is described that is efficient and accurate for matrices h much larger than available random access memory in a computer. Expansion of the lower triangle of h or transposition is avoided and all matrix processing is sequential. The proposed method uses triangular decomposition of the Hermitian matrix, but avoids complex arithmetic unless the original matrix is complex. In comparison with direct use of Gaussian elimination for $(h - \epsilon)^{-1}m$ the proposed method avoids an entire step of matrix processing.

This note describes an efficient procedure for evaluating the expression

$$m^\dagger(h - \epsilon)^{-1}\, m, \tag{1}$$

where h is an Hermitian matrix of very large dimension, m is a rectangular matrix, and ϵ denotes ϵI where I is the unit matrix.

An expression of this kind occurs in variational calculations in scattering theory [1], where h is the bound-bound Hamiltonian matrix (connecting normalizable functions), m is the bound-free Hamiltonian matrix (connecting normalizable and non-normalizable functions), and ϵ is the specified total energy of target system plus scattered particle.

Equation (1) arises in general from variational solution of Schrödinger's equation in a partitioned linear space [2]. If there are two partitions, subspaces A and B, the linear eigenvalue equation system can be symbolized by

$$(H_{AA} - E_A)\, x_A + H_{AB}x_B = 0,$$
$$H_{BA}x_A + (H_{BB} - E_B)\, x_B = 0. \tag{2}$$

The second equation can be solved formally for the vector x_B,

$$x_B = -(H_{BB} - E_B)^{-1}\, H_{BA}x_A, \tag{3}$$

which is substituted into the first equation to give

$$(H_{AA} + H'_{AA} - E_A)\, x_A = 0, \tag{4}$$

where

$$H'_{AA} = -H_{AB}(H_{BB} - E_B)^{-1}H_{BA}. \tag{5}$$

The eigenvalue condition

$$E_A = E_B = E \tag{6}$$

must be satisfied by iteration of the matrix equations in the case of discrete eigenvalues. Matrix H'_{AA} is of the form considered in Eq. (1).

The most straightforward approach to Eq. (1) is to solve the inhomogeneous equations

$$(h - \epsilon) y = m \tag{7}$$

for the matrix y considered as a collection of vectors, and then to evaluate

$$m^{\dagger} y = m^{\dagger}(h - \epsilon)^{-1} m. \tag{8}$$

If h is a matrix of very large dimension, this procedure leads to grave difficulties of data-handling. For example, if the Hermitian matrix h is given originally as the lower-triangular array of its independent elements, it must first be expanded to a full square array for application of standard methods to Eq. (7). This expansion requires nonsequential access to the original elements of h. When only a small fragment of h can be stored in the available random access memory of a computer, the entire matrix must be scanned time after time through a buffer system in order to produce the transpose or the expanded square matrix.

The alternative procedure proposed here makes use of the triangular factorization

$$h - \epsilon = t\sigma t^{\dagger}, \tag{9}$$

where t is a lower triangular matrix whose diagonal elements are positive real numbers and σ is a diagonal sign matrix (elements ± 1 only). An auxiliary rectangular matrix b is computed such that

$$tb = m \quad \text{or} \quad b = t^{-1}m. \tag{10}$$

Then Eq. (1) is evaluated as

$$\begin{aligned} m^{\dagger}(h - \epsilon)^{-1} m &= m^{\dagger}(t^{\dagger})^{-1} \sigma t^{-1}m \\ &= b^{*}\sigma b. \end{aligned} \tag{11}$$

In practical applications only a small fraction of the elements of σ are negative. It is convenient to represent this matrix as a list of pointers to these negative elements, to be used in the logic of computations indicated by Eqs. (9) and (11).

When $h - \epsilon$ is positive definite, σ reduces to a unit matrix, and Eq. (9) describes the well-known Cholesky factorization, advocated by Wilkinson [3] on grounds of numerical stability as a step in matrix inversion or solution of inhomogeneous

equations. In contrast, because the proof of stability fails for general Hermitian matrices, Wilkinson specifically warns against the use of triangular factorization except for positive definite matrices [3]. One purpose of the present note is to report that for matrices arising in scattering theory [1], the procedure using triangular factorization, Eqs. (9)–(11), has been found to be numerically more stable than the alternative procedure, Eqs. (7) and (8), with a careful choice of pivots in solution of Eq. (7) by Gaussian elimination.

Equations (9)–(11) can be implemented entirely by sequential data processing using $h - \epsilon$ as a lower triangular array. The elements t_{ij} of t can be obtained in their natural sequence (ordered by rows of a lower triangular matrix). From Eq. (9),

$$t_{ij} = \left(h_{ij} - \sum_{k=1}^{j-1} \sigma_k t_{ik} t_{jk}^* \right) \Big/ \sigma_j t_{ij}, \qquad j = 1,\dots, i-1 \tag{12}$$

$$\left. \begin{aligned} \sigma_i &= \mathrm{sgn} \left(h_{ii} - \epsilon - \sum_{k=1}^{i-1} \sigma_k t_{ik} t_{ik}^* \right) \\[2ex] t_{ii} &= + \left| h_{ii} - \epsilon - \sum_{k=1}^{i-1} \sigma_k t_{ik} t_{ik}^* \right|^{1/2}, \qquad i = 1,\dots, n, \end{aligned} \right\} \tag{13}$$

where n is the linear dimension of h. If the entire lower triangle h is accessible at random, it can be replaced element by element by t without any additional working space except for the negative sign pointers for σ. When this is not possible, the input matrix h must be scanned only once sequentially, while the output matrix t must be scanned sequentially for each new element t_{ij} from its origin up to the location of the new element. This can be organized efficiently by using a large buffer area for as many rows of t as can be accessed concurrently. The corresponding segment of h is read into this window area and all elements in the area converted concurrently from h to t during a single sequential scan of the segment of t preceding the window segment.

The data processing required by Eq. (10) is closely analogous to that described for Eq. (9). The elements $b_{i\alpha}$ of b are obtained in natural sequence (ordered by rows),

$$b_{i\alpha} = \left(m_{i\alpha} - \sum_{k=1}^{i-1} t_{ik} b_{k\alpha} \right) \Big/ t_{ii}, \qquad \alpha = 1,\dots, n'; \quad i = 1,\dots, n, \tag{14}$$

where n' is the second dimension of the rectangular matrix b. If random access to all of m is possible, it can be replaced element by element by b. Otherwise, m can be scanned once sequentially while b is built up through a window buffer, requiring a sequential scan of b from its origin for each window segment.

In practical applications, n' is very much smaller than n, and the matrix defined by Eq. (1) is to be subtracted from an existing matrix stored in random access memory. Then the matrix contraction given by Eq. (11) requires only a single sequential scan through the rectangular matrix b.

In comparison with Gaussian elimination, as applied to Eq. (7), the procedure described here saves an entire step of matrix manipulation in reaching the final result given by Eq. (1). Since the eliminated step, back-substitution in the usual procedure, requires the equivalent of transposition of a triangular matrix, the saving in data-handling is very important, in addition to eliminating arithmetic operations.

To see this, suppose that the first step in Gaussian elimination (reduction to triangular form) is carried out by multiplying Eq. (7) by t^{-1}, defined by Eq. (9), to give

$$\sigma t^\dagger y = t^{-1} m = b. \tag{15}$$

The work involved in reaching this form is precisely that of the two stages of matrix computation required to obtain t from Eq. (9) and b from Eq. (10). Gaussian elimination would proceed to solve Eq. (15) for the rectangular matrix y by back-substitution, and then y would be used in the unsymmetrical contraction, Eq. (8), to obtain the final result, Eq. (1). Direct use of b in Eq. (11) eliminates the back-substitution step.

The method described here is being used in a many-electron, many-channel formulation of electron-atom scattering theory [4]. It is also being incorporated into a procedure designed greatly to extend the practical limits of dimensionality of complex Hermitian eigenvalue equations in energy band calculations by the OPW method [5].

It should be noted that if h is a real symmetric matrix and if m is real, only real numbers occur in the matrices t and b.

The method can be applied to the generalized form of Eqs. (2) that includes a metric matrix S. Then Eq. (5) is replaced by

$$H'_{AA} = -(H_{AB} - E_B S_{AB})(H_{BB} - E_B S_{BB})^{-1}(H_{BA} - E_B S_{BA}). \tag{16}$$

This generalization requires a trivial modification of Eqs. (12) and (13).

If any leading submatrix of $h - \epsilon$ has a zero eigenvalue, Eq. (12) fails. However, in practical applications using floating point arithmetic, this requires a very unlikely numerical coincidence. In general, for the matrix $h - \epsilon$ or for its leading submatrices, the number of negative elements σ is equal to the number of negative eigenvalues.

EXAMPLE.   In order to demonstrate the numerical stability of the present

method, it has been applied to a simple example with a known exact solution. Given linear dimension n, the matrix h is taken to be

$$h_{ij} = 1 + 2(i - 1)\,\delta_{ij}\,; \qquad i, j = 1,..., n. \tag{17}$$

The matrix m is taken to be a single column vector,

$$m_i = n + 2(i - 1) - \epsilon; \qquad i = 1,..., n. \tag{18}$$

This vector has been chosen so that the vector y, Eq. (7), has all elements equal to unity. Then, from Eq. (8),

$$m^+(h - \epsilon)^{-1}\,m = m^+y = \Sigma_i m_i = n(2n - 1 - \epsilon). \tag{19}$$

Eigenvalues of h, for $n = 5$, 10, 20, and 40, were found by a Jacobi matrix diagonalization, and the eigenvectors were used to compute Eq. (1) directly, verifying the formula given above to eight significant decimals for all values of ϵ considered. A parallel calculation, using the method of the present paper, gave identical results to eight decimals. The values of ϵ considered scanned through the fourth eigenvalue of each h matrix with increments of 0.1×10^{-6}, affecting the eighth significant decimal digit in each case. 64 bit floating point arithmetic was used for these calculations. With $n = 40$, the fourth eigenvalue of h is computed by the Jacobi method to be 6.4030997, rounded to eight digits. The present method gives the correct value 2903.8760 for $m^+(h - \epsilon)^{-1}m$ when ϵ is either 6.4030996 or 6.4030997. The number of negative elements σ_i increases by one in this interval, indicating agreement between the present method and the Jacobi method with regard to the location of the eigenvalue.

These calculations show no evidence of numerical instability of the method proposed here when applied to matrices with negative eigenvalues, even when an eigenvalue of $h - \epsilon$ is made to be very nearly equal to zero.

These examples deal with relatively small matrices, for which alternative methods are available. In applications to electron-atom scattering theory, for which the present method is designed, matrices of linear dimension several thousand can occur. Such matrices, which contain several millions of nonzero elements, cannot be stored in the random access memory of existing computers. For such matrices the data handling aspects of the present method make it possible to carry out calculations that would be impossible or extremely inefficient by alternative methods.

REFERENCES

1. R. K. NESBET, *Phys. Rev.* **179** (1969), 60.
2. P.-O. LÖWDIN, *J. Chem. Phys.* **19** (1951), 1396; in "Perturbation Theory and its Applications in Quantum Mechanics" (C. H. Wilcox, Ed.), p. 255, John Wiley and Sons, New York, 1966.

3. J. H. WILKINSON, "The Algebraic Eigenvalue Problem," pp. 221–233, Oxford University Press, New York, 1965.
4. R. K. NESBET, *Bull. Amer. Phys. Soc.* **15** (1970), 1524.
5. W. E. RUDGE, to be published.

R. K. NESBET

IBM Research Laboratory,
Monterey and Cottle Roads,
San Jose, California 95114