J. Quant. Spectrosc. Radiat. Transfer Vol. 39, No. 5, pp. 415–419, 1988
Printed in Great Britain.

0022-4073/88 $3.00 + 0.00

COMPUTATION OF EIGENVALUES AND EIGENVECTORS FOR THE DISCRETE ORDINATE AND MATRIX OPERATOR METHODS IN RADIATIVE TRANSFER

KNUT STAMNES[1], SI-CHEE TSAY[2]† and TERUYUKI NAKAJIMA[3]

[1]Auroral Observatory, Institute of Mathematical and Physical Sciences, University of Tromsø, Norway, [2]Geophysical Institute and Department of Physics, University of Alaska-Fairbanks, Fairbanks, AK 99775-0800, U.S.A., and [3]Upper Atmosphere Research Laboratory, Tohoku University, Sendai 980, Japan

(*Received* 26 *May* 1987)

Abstract—Nakajima and Tanaka showed that the algebraic eigenvalue problem occurring in the discrete ordinate and matrix operator methods can be reduced to finding eigenvalues and eigenvectors of the product of two symmetric matrices, one of which is positive definite. Here, we show that the Cholesky decomposition of this positive definite matrix can be used to convert the eigenvalue problem into one involving a symmetric matrix. The Cholesky decomposition is extremely stable and is expected to improve the speed of the eigenvalue/eigenvector computation. After a careful comparison of the Nakajima and Tanaka procedure, our new Cholesky decomposition method and the original procedure suggested by Stamnes and Swanson, we find (contrary to our expectations) that the Stamnes and Swanson prescription is still the most accurate because it avoids round-off errors due to matrix multiplications needed to symmetrize the matrix in the two other procedures. We also find that, when the QR algorithm (used to solve the asymmetric eigenvalue problem in the Stamnes and Swanson procedure) is changed to avoid complex arithmetic, the speed becomes comparable to that of the two other procedures based on reduction to symmetric matrices.

INTRODUCTION

Several years ago, Stamnes and Swanson[1] showed that the system of coupled differential equations occurring in the discrete ordinate method pertinent to radiative transfer in plane-parallel media can be reduced to a standard algebraic eigenvalue problem. The same eigenvalue problem occurs in the matrix operator and the spherical harmonic method. The relationship between the discrete ordinate and the matrix operator methods has been discussed by Waterman,[2] Nakajima and Tanaka[3] and Stamnes,[4] and between the discrete ordinate and the spherical harmonic method by Karp et al[5] and Stamnes and Swanson.[1]

Because of the special structure of the matrix involved, which is a consequence of the reciprocity principle, the eigenvalues occur in positive/negative pairs, and the order of the algebraic eigenvalue problem can be reduced by a factor of 2. The reduced matrix, which is real but asymmetric, consists of a product of two matrices, which are also real, but asymmetric.[1] Since the reduced matrix is real but asymmetric, Stamnes and Swanson[1] adopted a solver which utilizes the double-QR algorithm. This algorithm applies to a general real matrix (which may have complex eigenvalues/eigenvectors) and uses complex arithmetic.

The fact that the eigenvalues are real, as proven by Kuscer and Vidav,[6] suggests that it is possible to transform the algebraic eigenvalue problem into one involving symmetric matrices. Recently, Nakajima and Tanaka[3] showed that it is, in fact, possible to introduce a scaling transformation which transforms the matrix for the reduced problem in Stamnes and Swanson[1] into a product of two real, *symmetric* matrices, one of which is positive definite.

Nakajima and Tanaka[3] showed that the eigenvalues and eigenvectors of the transformed reduced matrix can be obtained by solving two symmetric eigenvalue problems (and thereby avoiding complex arithmetic). However, in addition to two symmetric eigenvalue problems, the procedure of Nakajima and Tanaka[3] also involves five matrix multiplications and one square root of a matrix,

†Present address: Department of Atmospheric Science, Colorado State University, Fort Collins, CO 80523, U.S.A.

which makes their computation susceptible to computer round-off errors. Nakajima and Tanaka[3] erroneously stated that their procedure yields more accurate results than the Stamnes and Swanson procedure based on the QR algorithm. This error was due to the use of a faulty QR algorithm by Nakajima and Tanaka.

The purpose of this article is to show that it is possible to obtain eigenvalues and eigenvectors of the transformed reduced matrix of Nakajima and Tanaka by a simpler and faster procedure which involves the solution of only one symmetric eigenvalue problem, one inversion of a triangular matrix and the equivalent of one matrix multiplication. As we shall see, this simplification is obtained by utilizing the Cholesky decomposition of the positive definite matrix. In spite of this improvement, we find to our surprise that, in single-precision computations, the Stamnes and Swanson procedure is more accurate than the Cholesky decomposition procedure and we attribute this circumstance to the matrix multiplications needed to symmetrize the matrix in the Nakajima and Tanaka procedure and our present procedure based on the Cholesky decomposition.

MATRIX FORMULATION

We start with the discrete ordinate approximation to the homogeneous version of the radiative transfer equation, which may be written in matrix form as[1]

$$\begin{bmatrix} \dfrac{d\mathbf{u}^+}{d\tau} \\ \dfrac{d\mathbf{u}^+}{d\tau} \end{bmatrix} = \begin{bmatrix} -\boldsymbol{\alpha} & -\boldsymbol{\beta} \\ \boldsymbol{\beta} & \boldsymbol{\alpha} \end{bmatrix} \begin{bmatrix} \mathbf{u}^+ \\ \mathbf{u}^- \end{bmatrix}, \tag{1}$$

where the notation is as follows:[1,3]

$$\mathbf{u} = \{u^m(\tau, \pm\mu_i)\}, \quad i = 1, \ldots, N, \tag{2a}$$

$$\boldsymbol{\alpha} = \mathbf{M}^{-1}\{\mathbf{C}^+\mathbf{W} - \mathbf{I}\}, \tag{2b}$$

$$\boldsymbol{\beta} = \mathbf{M}^{-1}\mathbf{C}^-\mathbf{W}, \tag{2c}$$

$$\mathbf{M} = \{\mu_i\delta_{ij}\}, \quad i,j = 1, \ldots, N, \tag{2d}$$

$$\mathbf{W} = \{w_i\delta_{ij}\}, \quad i,j = 1, \ldots, N, \tag{2e}$$

$$\mathbf{C}^+ = \{C^m(\mu_i, \mu_j)\} = \{C^m(-\mu_i, \mu_j)\}, \quad i,j = 1, \ldots, N, \tag{2f}$$

$$\mathbf{C}^- = \{C^m(-\mu_i, \mu_j)\} = \{C^m(\mu_i, -\mu_j)\}, \quad i,j = 1, \ldots, N. \tag{2g}$$

The μ_i and w_i are the quadrature points and weights, and $u^m(\tau, \mu_i)$ is the **m**th Fourier component of the intensity at optical depth τ and direction μ_i. $C^m(\tau, \mu_i, \mu_j)$ contains information about the single scattering albedo and the phase function and is one-half for conservative, isotropic scattering [cf. Eq. (5) of Ref. 1]. We refer the reader to Stamnes and Swanson, Nakajima and Tanaka and Stamnes for further details.

The reduction of the order proceeds as follows (cf. Ref. 1). We rewrite Eq. (1) as

$$\frac{d(\mathbf{u}^+ + \mathbf{u}^-)}{d\tau} = -(\boldsymbol{\alpha} - \boldsymbol{\beta})(\mathbf{u}^+ - \mathbf{u}^-), \tag{3a}$$

$$\frac{d(\mathbf{u}^+ - \mathbf{u}^-)}{d\tau} = -(\boldsymbol{\alpha} + \boldsymbol{\beta})(\mathbf{u}^+ + \mathbf{u}^-), \tag{3b}$$

and, by combining Eqs. (3a) and (3b), we obtain

$$\frac{d^2(\mathbf{u}^+ + \mathbf{u}^-)}{d\tau^2} = (\boldsymbol{\alpha} - \boldsymbol{\beta})(\boldsymbol{\alpha} + \boldsymbol{\beta})(\mathbf{u}^+ + \mathbf{u}^-) \tag{4a}$$

and

$$\frac{d^2(\mathbf{u}^+ - \mathbf{u}^-)}{d\tau^2} = (\boldsymbol{\alpha} + \boldsymbol{\beta})(\boldsymbol{\alpha} - \boldsymbol{\beta})(\mathbf{u}^+ - \mathbf{u}^-), \tag{4b}$$

which completes the reduction of the order. The matrices $(\boldsymbol{\alpha} - \boldsymbol{\beta})$ and $(\boldsymbol{\alpha} + \boldsymbol{\beta})$ are both asymmetric and their products (Eqs. (4a) and (4b)] are both asymmetric. Stamnes and Swanson[1] solved Eq. (4a) and thus obtained the eigenvectors $(\mathbf{u}^+ + \mathbf{u}^-)$. They then used Eq. (3b) to determine $(\mathbf{u}^+ - \mathbf{u}^-)$.

TRANSFORMATION TO SYMMETRIC MATRICES

We now follow the notation and presentation of Nakajima and Tanaka[3] closely and start by rewriting Eqs. (3a) and (3b) as

$$\frac{d\boldsymbol{\phi}^{\pm}}{d\tau} = (-\boldsymbol{\alpha} \pm \boldsymbol{\beta})\boldsymbol{\phi}^{\mp} = \mathbf{M}^{-1}\mathbf{G}^{\mp}\mathbf{W}\boldsymbol{\phi}^{\mp}, \tag{5a}$$

where

$$\boldsymbol{\phi}^{\pm} = \mathbf{u}^+ \pm \mathbf{u}^-, \tag{5b}$$

$$\mathbf{G}^{\pm} = \mathbf{W}^{-1} - (\mathbf{C}^+ \pm \mathbf{C}^-). \tag{5c}$$

We note that, while $(-\boldsymbol{\alpha} \pm \boldsymbol{\beta})$ are asymmetric matrices, the $\mathbf{G}^{\pm}$ are symmetric. By introducing the scaling

$$\boldsymbol{\phi}^{\pm} = \sqrt{\mathbf{WM}}\boldsymbol{\phi}^{\pm} \tag{6a}$$

into Eq. (5a), we find

$$\frac{d\hat{\boldsymbol{\phi}}^{\pm}}{d\tau} = \mathbf{X}^{\mp}\hat{\boldsymbol{\phi}}^{\mp}, \tag{6b}$$

where

$$\mathbf{X}^{\pm} = \sqrt{\mathbf{WM}^{-1}}\mathbf{G}^{\pm}\sqrt{\mathbf{WM}^{-1}} \tag{6c}$$

are symmetric matrices since $\mathbf{G}^{\pm}$ are symmetric.

Elimination of $\boldsymbol{\phi}^-$ and $\boldsymbol{\phi}^+$ from Eq. (6b) yields

$$\frac{d^2\hat{\boldsymbol{\phi}}^+}{d\tau^2} = \mathbf{X}^-\mathbf{X}^+\hat{\boldsymbol{\phi}}^+ \tag{7a}$$

and

$$\frac{d^2\hat{\boldsymbol{\phi}}^-}{d\tau^2} = \mathbf{X}^+\mathbf{X}^-\hat{\boldsymbol{\phi}}^-, \tag{7b}$$

respectively.

We note that, except for the scaling [Eq. (6a)], Eqs. (7a) and (7b) are identical to Eqs. (4a) and (4b), respectively, and that the virtue of the scaling is to make $\mathbf{X}^+$ and $\mathbf{X}^-$ symmetric. We need to solve only one of the eigenvalue problems (7a) or (7b) since they are coupled through Eq. (6b). Nakajima and Tanaka obtained the eigenvalues and eigenvectors of $\mathbf{X}^-\mathbf{X}^+$ [Eq. (7a)] by a procedure involving two symmetric eigenvalue problems, five matrix multiplications and one square root of a matrix [Eqs. (17)–(23) in their paper]. Next, we present a simpler and computationally faster procedure to compute the eigenvalues and eigenvectors of $\mathbf{X}^+\mathbf{X}^-$ [Eq. (7b)].

REDUCTION TO ONE SYMMETRIC EIGENVALUE PROBLEM

As was noted by Nakajima and Tanaka[3] the matrix $\mathbf{X}^-$ is positive definite and we shall take advantage of this fact in the following analysis.

To obtain eigenvalues and eigenvectors of Eq. (7b), i.e.,

$$(\mathbf{X}^+\mathbf{X}^-)\mathbf{Q} = \mathbf{Q}\hat{\mathbf{Z}}, \quad \hat{\mathbf{Z}} = \{\lambda_i^2\delta_{ij}\} \tag{8a}$$

we use the Cholesky decomposition to write

$$\mathbf{X}^- = \mathbf{R}^T\mathbf{R}, \tag{8b}$$

where the superscript-T denotes transpose and $\mathbf{R}$ is an upper triangular matrix with positive diagonal elements.[7] Substitution of Eq. (8b) into Eq. (8a) yields

$$(\mathbf{X}^{+}\mathbf{R}^{\mathrm{T}}\mathbf{R})\mathbf{Q} = \mathbf{Q}\hat{\mathbf{Z}} \tag{8c}$$

and

$$(\mathbf{R}\mathbf{X}^{+}\mathbf{R}^{\mathrm{T}})\mathbf{R}\mathbf{Q} = \mathbf{R}\mathbf{Q}\hat{\mathbf{Z}}. \tag{8d}$$

Thus, we see that by solving the symmetric eigenvalue problem

$$(\mathbf{R}\mathbf{X}^{+}\mathbf{R}^{\mathrm{T}})\mathbf{V} = \mathbf{V}\mathbf{Z}, \tag{9a}$$

the eigenvalues and eigenvectors of Eq. (8a) are λ_i^2 and

$$\mathbf{Q} = \mathbf{R}^{-1}\mathbf{V}. \tag{9b}$$

This approach involves only the solution of one symmetric eigenvalue problem [Eq. (9a)], one decomposition and one inversion of a triangular matrix [Eq. (9b)].

DISCUSSION AND CONCLUSION

The advantage of solving the algebraic eigenvalue of one asymmetric matrix (4a,b) (which involves complex arithmetic) is that only one matrix multiplication is necessary. The transformation to symmetric matrices $\mathbf{X}$ and subsequent solution procedures introduce matrix operations in which the effect of rounding error in the numerical computation may become significant. However, the Cholesky decomposition of the positive definite matrix $\mathbf{X}^{-}$ is extremely stable and the symmetric matrix $\mathbf{S} = \mathbf{R}\mathbf{X}^{+}\mathbf{R}^{\mathrm{T}}$ can be obtained in 2/3 N^3 multiplications (N is the order of the matrix) by taking proper advantage of the symmetry. This includes the Cholesky decomposition (see Ref. 7, p. 338). To keep the effect of rounding errors at a low level, $\mathbf{S}$ is obtained by properly accumulating the inner products of $\mathbf{R}\mathbf{X}^{+}\mathbf{R}^{\mathrm{T}}$. Because $\mathbf{X}^{+}$ is symmetric and $\mathbf{R}$ is upper triangular, the upper triangular part of $\mathbf{S} = \{S_{ij};\ i = 1, \ldots, N,\ j = i, \ldots, N\}$ can be obtained as follows:

$$S_{ij} = \sum_{k=i}^{N} \sum_{l=l}^{N} r_{ik} x_{kl} r_{jl},$$

where

$$\mathbf{R} = \begin{vmatrix} r_{11} \ldots r_{1n} \\ 0 \ldots r_{nn} \end{vmatrix}. \quad \mathbf{X}^{+} = \begin{vmatrix} x_{11} \ldots x_{1n} \\ x_{1n} \ldots x_{nn} \end{vmatrix},$$

and the lower triangular part of $\mathbf{S}$ follows by symmetry.

We have compared our new procedure based on the Cholesky decomposition [Eqs. (8) and (9)] with Nakajima and Tanaka [Eqs. (17)–(23) in their paper] and Stamnes and Swanson. These comparative runs were made on a VAX computer in Alaska and a SX-1 super computer of NEC in Japan. Single- and double-precision runs were made on both machines which are 32-bits/single 64-bits/double computers and the results were compared with those of the doubling method,[8] which was executed in double precision for 32 streams to get high accuracy. The outcome of these comparisons may be summarized as follows: (1) all three methods work well when executed in double precision. (2) In single precision, the Stamnes and Swanson procedure is more accurate than the Cholesky procedure which, in turn, is slightly more accurate than the Nakajima and Tanaka procedure. (3) The Cholesky procedure is slightly faster than Nakajima and Tanaka, which is slightly faster than Stamnes and Swanson. Based on these findings, we decided to look into the possibility of speeding up the Stamnes and Swanson procedure, since it is the most accurate. A close examination then revealed that by eliminating complex arithmetic in the QR-algorithm, the speed of this procedure became comparable to that of the Cholesky procedure. We have also found that, on 32-bit machines, very accurate results are obtained (3–4 digits for fluxes and 2–3 digits for intensities) in single precision if the eigenvalues are computed in double precision. Thus, on 64-bit machines, single precision will yield all of the accuracy needed for most practical applications. Our conclusion is therefore that the Stamnes-Swanson prescription to compute eigenvalues and eigenvectors is still the best known procedure available. Finally, we point out that the apparent shortcomings of this procedure found by Nakajima and Tanaka[3] (cf. their Fig. 3) are due to their use of a faulty double QR-algorithm.

Acknowledgement—This research was supported by the National Science Foundation through grant DPP 84-06093 to the University of Alaska.

REFERENCES

1. K. Stamnes and R. A. Swanson, *J. Atmos. Sci.* **38,** 387 (1981).
2. P. C. Waterman, *JOSA* **71,** 410 (1981).
3. T. Nakajima and M. Tanaka, *JQSRT* **35,** 13 (1986).
4. K. Stamnes, *Rev. Geophys.* **24,** 299 (1986).
5. A. H. Karp, J. J. Greenstadt, and J. A. Filmore, *JQSRT* **24,** 391 (1980).
6. I. Kuscer and I. Vidav, *J. Math. Anal. Appl.* **25,** 80 (1969).
7. J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, p. 662, Oxford University Press, Oxford (1965).
8. W. J. Wiscombe, *JQSRT* **15,** 477 (1976).