INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN ENGINEERING, VOL. 19, 943-948 (1983)

# SHORT COMMUNICATION

## A METHOD FOR EIGENVALUES OF SPARSE $\lambda$-MATRICES

WEI H. YANG

*Department of Mechanical Engineering and Applied Mechanics, The University of Michigan, Ann Arbor, Michigan, U.S.A.*

### SUMMARY

The matrix $N(\lambda)$ whose elements are functions of a parameter $\lambda$ is called the $\lambda$-matrix. Those values of $\lambda$ that make the matrix singular are of great interest in many applied fields. An efficient method for those eigenvalues of a $\lambda$-matrix is presented. A simple explicit convergence criterion is given, as well as the algorithm and two numerical examples.

### INTRODUCTION

The most general algebraic eigenvalue problem takes the form

$$N(\lambda)v = 0 \tag{1}$$

where $N(\lambda)$ is a square matrix whose elements are functions of a parameter $\lambda$. The problem is to find those values of $\lambda$ such that equation (1) admits non-trivial vectors $v$. The special cases, $N(\lambda) = A - \lambda I$ and $N(\lambda) = A - \lambda B$, involving constant matrices $A$, $B$ and the identity matrix $I$, are the most often encountered ones in physical science applications. Equation (1) may arise directly from an applied problem or is converted from a usual eigenvalue problem for the purpose of reducing the dimension.[1]

Osborne and Michaelson[2] gave an iterative method which simultaneously iterates an eigenpair of equation (1). In many applications, the eigenvectors are of less interest than the eigenvalues. It seems wasteful to compute those eigenvectors if they are not needed. In the case when $N(\lambda)$ is Hermitian, the eigenvectors are in general complex, although the eigenvalues are real. Additional computation for those complex eigenvectors further decreases the efficiency of a method if only eigenvalues are desired.

The work of Kublanovskaya[3] presents elegance and efficiency. It uses unitary matrices to triangularize the $\lambda$-matrix for a given estimate $\lambda$. The Newton's iteration is then applied to the smallest diagonal element of the lower triangular matrix to improve the estimated eigenvalue.

The method presented in this paper uses the $LU$ decomposition[6] of the $\lambda$-matrix and an idea of small perturbation on the determinant. It directly calculates an eigenvalue of the $\lambda$-matrix without involving eigenvectors. A simple convergence criterion is developed.

### METHOD

Consider a $n \times n$ matrix $N(\lambda)$ whose elements are functions of $\lambda$ and are differentiable at least once. The method described here will obtain iteratively an eigenvalue $\lambda^*$ of $N(\lambda)$ if an initial estimate $\lambda_0$ is sufficiently close to $\lambda^*$.

0029-5981/83/060943-06$01.00

*Received 13 July 1982*
*Revised 10 August 1982*

Differentiating the $\lambda$-matrix and denoting it by $N'(\lambda)$, an approximate expression can be written

$$N(\lambda+\Delta\lambda)\approx N(\lambda)+\Delta\lambda N'(\lambda) \tag{2}$$

For a given estimate of $\lambda$, we seek a correction $\Delta\lambda$ to make $\det(N(\lambda+\Delta\lambda))=0$. Let $\mu=\Delta\lambda$. $\mu$ should satisfy

$$\det(N(\lambda)+\mu N'(\lambda))=0 \tag{3}$$

This is equivalent to solving an eigenvalue problem of the form

$$Ax=\mu Bx \tag{4}$$

where $A=N(\lambda)$ and $B=-N'(\lambda)$.

There are many good methods[4] to solve equation (4). For our purpose, only the smallest $|\mu|$ is needed. A method most efficient for this purpose is developed.

For a given $\lambda_k$, it is always possible to factorize $N(\lambda_k)$ into a product of three matrices with suitable choice of a permutation matrix $P$ such that

$$N(\lambda_k)=P^{\mathrm{T}}LU \tag{5}$$

where $L$ is a lower triangular matrix with unit diagonal and $U$ is an upper triangular matrix. If one or more diagonal elements of $U$ equals zero, then $\lambda_k$ is an eigenvalue, since

$$\det(N(\lambda_k))=\prod_{i=1}^{n}u_{ii} \tag{6}$$

We shall assume that $u_{ii}\neq 0$, $i=1,2,\ldots,n$. Equation (3) can be written in the form

$$\det(I+\mu E)=0 \tag{7}$$

since $N(\lambda_k)$ is assumed invertible, where $E=L^{-1}PN'U^{-1}$.

The matrix $\mu E$ may be regarded as a small perturbation of the identity matrix $I$ for a small value of $\mu$. The effect of the perturbation on the determinant may be expressed by

$$\det(I+\mu E)\approx\det(I)+\sum_{j=1}^{n}\sum_{i=1}^{n}\frac{\partial}{\partial d_{ij}}\det(I)\mu e_{ij} \tag{8}$$

where $d_{ij}$ and $e_{ij}$ are the elements of the matrices $I$ and $E$, respectively.

We may denote the partial derivatives by

$$\frac{\partial}{\partial d_{ij}}\det(I)=\operatorname*{Lim}_{\Delta d_{ij}\to 0}\left\{\frac{1}{\Delta d_{ij}}[\det(I+\Delta d_{ij}e_ie_j^{\mathrm{T}})-\det(I)]\right\} \tag{9}$$

where $e_i$ and $e_j$ are the unit base vectors in $E^n$.

Note that $\det(I+\Delta d_{ij}e_ie_j^{\mathrm{T}})=\det(I)$ for $i\neq j$. It is simple to deduce that

$$\frac{\partial}{\partial d_{ij}}\det(I)=\begin{cases}0 & i\neq j\\ 1 & i=j\end{cases} \tag{10}$$

Equation (8) can be reduced to give

$$1+\mu\sum_{i=1}^{n}e_{ii}=0 \tag{11}$$

Hence, we obtain

$$\mu(\lambda_k) = -\left(\sum_{i=1}^{n} e_{ii}\right)^{-1} \tag{12}$$

We can now improve the approximate eigenvalue by

$$\lambda_{k+1} = \lambda_k + \mu(\lambda_k) \qquad k = 0, 1, 2, \ldots \tag{13}$$

The convergence of the above iteration is discussed in the next section.

In the case when $N(\lambda)$ is Hermitian, the method applies with a small modification. All algebraic operations remain real except the procedure of forming the factorization in (5). The matrices $L$ and $U$ are complex in this case, but the diagonal elements of $E$ remain real since eigenvalues of a Hermitian matrix are real. The complex arithmetic required does not add a significant amount of computation.

There remains the probability that all $e_{ii} = 0$, $i = 1, 2, \ldots, n$. This situation shall be discussed later. First, we shall show that iteration (13) is Newton's iteration in disguise, that is the function $\mu(\lambda)$ from equation (12) can be expressed by

$$\mu(\lambda) = -\det(N(\lambda)) \Big/ \frac{\mathrm{d}}{\mathrm{d}\lambda}[\det(N(N))] \tag{14}$$

This can be easily shown since

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}\lambda}\det(N) &= \operatorname*{Lim}_{\Delta\lambda \to 0} \frac{1}{\Delta\lambda}[\det(N + \Delta\lambda N') - \det(N)] \\ &= \det(N) \operatorname*{Lim}_{\Delta\lambda \to 0} \frac{1}{\Delta\lambda}[\det(I + \Delta\lambda E) - 1] \\ &= \det(N) \sum_{i=1}^{n} e_{ii} \end{aligned} \tag{15}$$

Here, we have used the fact that $N = P^{\mathrm{T}}LU$ and $\det(U) \neq 0$.

Hence, the expression (13) is the Newton's iteration on the scalar equation $\det(N(\lambda)) = 0$. Direct construction of the Newton's iteration from such an equation faces the difficulty of differentiating the determinant. The present method circumvents this difficulty.

## CONVERGENCE

Convergence of the Newton's iteration (13) is obviously dependent on the initial estimate $\lambda_0$. We may use the Kantorovic theorem[5] to establish a convergence criterion for the algorithm of computation.

Denote the determinant by

$$f(\lambda) = \det(N(\lambda)) \tag{16}$$

For a given $\lambda_0$, we may compute $\mu_0 = \mu(\lambda_0)$. In the process we have computed

$$f_0 = f(\lambda_0) = \prod_{i=1}^{n} u_{ii} \tag{17}$$

Then obtain

$$f_0' = f'(\lambda_0) = -f_0/\mu_0 \tag{18}$$

as well as $\lambda_1 = \lambda_0 + \mu_0$. We still need an estimate for $f''(\lambda_0)$ which can be expressed in the form

$$f_0'' = f''(\lambda_0) = (1 + \mu'(\lambda_0)) f_0 / u_0^2 \tag{19}$$

Unfortunately, $\mu'(\lambda_0)$ is not obtainable exactly. An approximation

$$\mu'(\lambda_0) \approx \frac{1}{\varepsilon} [\mu(\lambda_0 + \varepsilon) - \mu(\lambda_0)] \tag{20}$$

can be computed, where $\varepsilon$ is a small number. It requires an additional $LU$ factorization on the matrix $N(\lambda_0 + \varepsilon)$.

Now we shall state the Kantorovic theorem that if

$$h_0 = \frac{|\lambda_1 - \lambda_0| \cdot |f''(\lambda_0)|}{|f'(\lambda_0)|} \leq \frac{1}{2} \tag{21}$$

then the Newton sequence (13) starting from $\lambda_0$ will converge to a solution $\lambda^*$ which exists in the neighbourhood defined by

$$|\lambda - \lambda_0| \leq \frac{1 - \sqrt{(1 - 2h_0)}}{h_0} |\lambda_1 - \lambda_0| \tag{22}$$

The above statement is a sufficient condition for convergence. If the condition (21) is not satisfied, the convergence cannot be guaranteed. In some instances, when $\lambda_0$ does not satisfy (21), Newton's iteration actually converges with a subsequent iterate satisfying condition (21). It is therefore practical to do a certain amount of computation on faith in the eventual convergence of the process. A prudent programmer, however, will set limits on the total number of iterations and amount of computer time for such trial ventures. This should be considered in the computer program.

## ALGORITHM

The method applies to a general $\lambda$-matrix with complex elements. A real symmetric and a Hermitian matrix arise more frequently in applications. It is practical to construct an algorithm for such special cases where the iterates remain real. In the case of a general $\lambda$-matrix, a complex eigenvalue can sometimes be approached only from a complex initial estimate in the neighbourhood of the eigenvalue.

In the process of forming the $LU$ factorization, it is probable that a diagonal element of $U$ becomes zero even with all possible rows pivoting. In such a case the process should be terminated and the current value of $\lambda$ be chosen as an eigenvalue. In practice, a very small diagonal element can also be regarded as zero provided that the smallness is measured relatively to the maximum diagonal element.

The permutation matrix $P$ which produces row pivoting in the $LU$ factorization is essential so that the process is numerically stable. The algorithm is stated as follows:

1. Assume an itial estimate $\lambda_0$ and compute $\lambda_1$ by iteration (13).
2. Compute $\mu'(\lambda_0)$ by (20) and test the adequacy of $\lambda_0$ by criterion (21).
3. If (21) is satisfied, continue. Otherwise, assume another $\lambda_0$ and go to step 1.
4. At a subsequent $k$th iteration ($k = 2, 3, 4, \ldots$), if the elements of $U$ satisfy $|u|_{\max} \cdot |u|_{\min} \leq \varepsilon_d$ (where $\varepsilon_d$ is a properly chosen small number), go to step 6. Otherwise continue.
5. If $\mu(\lambda_k)/\lambda_k \leq \varepsilon_\lambda$ (assuming $\lambda_k \neq 0$ and $\varepsilon_\lambda$ is properly chosen), go to step 6. Otherwise increase $k$ by 1 and go to step 4.
6. Set $\lambda^* = \lambda_k$.

A practical approach may be implemented. A convergence factor $\eta$ may be used such that

$$\lambda_{k+1} = \lambda_k + \eta\mu(\lambda_k) \tag{23}$$

$\eta$ can be properly chosen to accelerate the convergence.

## EXAMPLES

Two examples are presented here to demonstrate the method and the rate of convergence. First, consider the matrix

$$N(\lambda) = \begin{bmatrix} e^{\lambda} & 1 \\ 1 & \lambda \end{bmatrix} \tag{24}$$

which has a real eigenvalue $\lambda^* = 0{\cdot}56714329$ obtainable by solving $\det(N) = \lambda e^{\lambda} - 1 = 0$. We shall use the method described in this paper to obtain the eigenvalue. Three initial estimates are chosen to be $-2{\cdot}0$, 0, $3{\cdot}0$. The results are shown in Table I. On examining the $\det(N)$, the function has a minimum at $\lambda = -1$. The iterative process diverges for $\lambda_0 \leqslant -1$.

Table I

| $k$ | $\lambda_k$ | $\mu(\lambda_k)$ | $\lambda_k$ | $\mu(\lambda_k)$ | $\lambda_k$ | $\mu(\lambda_k)$ |
|---|---|---|---|---|---|---|
| 0 | −0·20000000 | 1·77675345 | 0·00000000 | 1·00000000 | 3·00000000 | −0·73755323 |
| 1 | 1·57675345 | −0·53171892 | 1·00000000 | −0·31606028 | 2·26244677 | −0·66157441 |
| 2 | 1·04503453 | −0·33904310 | 0·68393972 | −0·10648524 | 1·60087236 | −0·53795485 |
| 3 | 0·70599143 | −0·12448625 | 0·57745448 | −0·01022474 | 1·06291751 | −0·34779433 |
| 4 | 0·58150517 | −0·01419465 | 0·56722974 | −0·86441E−4 | 0·71512318 | −0·13176390 |
| 5 | 0·56731052 | −0·16721E−3 | 0·56714330 | −0·61124E−8 | 0·58335928 | −0·01600307 |
| 6 | 0·56714331 | | 0·56714329 | | 0·56735621 | |

The second example is to use the method to solve a linear eigenvalue problem $Ax = \lambda x$, where the matrix

$$A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \tag{25}$$

has eigenvalues $\lambda^* = 1, 3$. This problem can be written in the form of a nonlinear eigenvalue problem with

$$N(\lambda) = \begin{bmatrix} 2-\lambda & -1 \\ -1 & 2-\lambda \end{bmatrix} \tag{26}$$

With two initial estimates $\lambda_0 = 0$ and 4, we have the iterations shown in Table II. $\lambda_0 = 2$ does not lead to convergence for this problem for a reason to be given in the discussion. For $\lambda_0 = 2 \pm \varepsilon$, where $\varepsilon$ is a small positive number, the iterated sequence converges to either the eigenvalue 3 or 1 depending on the $\pm$ sign chosen. The value 2 happens to lie on the dividing point of the domains of attraction of the two eigenvalues.

Although the examples given are trivial in nature, they are merely to demonstrate the convergence properties of the iteration. The algorithm presented in this paper has been successfully applied to problems of waves in the periodic medium.

Table II

| $k$ | $\lambda_k$ | $\mu(\lambda_k)$ | $\lambda_k$ | $\mu(\lambda_k)$ |
|---|---|---|---|---|
| 0 | 0 | $\frac{3}{4}$ | 4 | $-\frac{3}{4}$ |
| 1 | $\frac{3}{4}$ | $\frac{9}{40}$ | $3\left(\frac{13}{12}\right)$ | $-\frac{9}{40}$ |
| 2 | $\frac{39}{40}$ | $\frac{81}{3280}$ | $3\left(\frac{121}{120}\right)$ | $-\frac{81}{3280}$ |
| 3 | $\frac{3279}{3280}$ | | $3\left(\frac{9841}{9840}\right)$ | |

## CONCLUSION

In both examples there is a point for which $\mathrm{d}/\mathrm{d}\lambda \det(N) = 0$. $\lambda_0 = -1$ is such a point for the first example; $\lambda_0 = 2$ for the second example. The iteration, of course, diverges immediately if such a $\lambda_0$ is chosen. The similar situation for other problems will be detected by the algorithm, since all diagonal elements of the matrix $E$ vanish at those points.

If an initial estimate is chosen close to those points, the iteration may converge, as is indicated in the second example. In a general problem with many eigenvalues, such convergence is not desirable, since the eigenvalue being converged to is not likely in the neighbourhood of the initial estimate. This possibility is precluded, since the criterion (21) is not satisfied for such an initial estimate.

The amount of work for each iteration involved in this method should be less than that of Kublanovskaya's method since the $LU$ factorization requires fewer operations than the unitary transformation. Some savings can be further realized when the matrix $N(\lambda)$ has sparse structures. The convergence criterion (21) is so easy to apply that it provides convenient tests on the adequacy of an initial estimate. This is an additional feature of the new method. A by-product of this paper is a method for the smallest eigenvalue of equation (4).

## REFERENCES

1. G. H. Golub, L. Jennings and W. H. Yang, 'Waves in periodically structed media', *J. Comp. Phys.* **17**(4) (1975).
2. M. R. Osborne and S. Michaelson, 'The numerical solution of eigenvalue problem in which the eigenvalue parameter appears nonlinearly, with an application to differential equations', *Comp. J.* **7**, 58–65 (1964).
3. V. N. Kublanovskaya, 'On an approach to the solution of the generalized latent value problem for $\lambda$-matrices', *SIAM J. Number. Anal.* **7**(4) (1970).
4. C. B. Moler and G. W. Stewart, 'An algorithm for generalized matrix eigenvalue problems', *SIAM J. Number. Anal.* **10**(2), (1973).
5. L. B. Rall, *Computational Solution of Nonlinear Operator Equations*, Wiley, London and New York, 1969.
6. G. Forsythe and C. B. Moler, *Computer Solution of Linear Algebraic Systems*, Prentice-Hall, Englewood Cliffs, N.J., 1967.