Aplikace matematiky

James Hardy Wilkinson

A survey of error analysis of matrix algorithms

Aplikace matematiky, Vol. 13 (1968), No. 1, 93–102

Persistent URL: http://dml.cz/dmlcz/103140

SVAZEK 13 (1968) APLIKACE MATEMATIKY ČÍSLO 1

A SURVEY OF ERROR ANALYSIS OF MATRIX ALGORITHMS

JAMES HARDY WILKINSON

1. INTODUCTION

Since the advent of high-speed digital computers enormous progress has been made in the solution of matrix problems. This is most marked in connexion with the algebraic eigenvalue problem. Up until 1950 surprisingly few problems involving matrices of orders greater than (say) six had been solved. Now on a modern high speed computer all the eigenvalues (real or complex) of a real matrix of order 50 can be found in less than a minute and rigorous bounds can be determined for the errors in the computed eigensystem if these are required.

Part of this advance has, of course, sprung directly from the very high speed of modern computers but at the same time a remarkable improvement has taken place in the algorithms themselves. It is interesting to note that since the latest computer was installed at the National Physical Laboratory almost all eigenvalue problems have been solved using algorithms which were unknown in 1950.

As long as the limitations of computing machines restricted the size of matrix that could be dealt with to order five or six, most of the algorithms that had been developed over the last hundred years could safely be used; these algorithms had mainly been developed in a purely theoretical context. Since the algebraic eigenvalue problem is theoretically equivalent to the calculation of the zeros of the characteristic polynomial it is at first sight attractive to compute the latter explicitly and then to use an algorithm designed specifically for the polynomial form. Most of the algorithms which were discussed in the early years of the electronic computer were of this type.

Experience with matrices, even of quite moderate orders brought into sharp relief the fact that problems which are mathematically equivalent may be very different from the point of view of numerical computation. This in turn led to a detailed consideration of the sensitivity of the eigensystem of a matrix with respect to perturbations of its elements and to a search for algorithms, based on transformations of the original matrix, which could be guaranteed to lead to matrices which were not more sensitive than the original. It is these considerations which have led to the development of the more successful algorithm now in use.

2. SENSITIVITY OF THE INVERSE OF A MATRIX AND OF ITS EIGENSYSTEM

The sensitivity of the inverse of a matrix with respect to changes in its elements is quite easy to determine. In fact if neither $A + E$ or A is singular we have

$$(2.1)\quad (A + E)^{-1} - A^{-1} = -A^{-1}(E - EA^{-1}E + EA^{-1}EA^{-1}E - \ldots)\, A^{-1}$$

giving

$$(2.2)\qquad \|(A + E)^{-1} - A^{-1}\|/\|A^{-1}\| \leqq \frac{\|A^{-1}\|\,\|E\|}{1 - \|A^{-1}\|\,\|E\|}$$

provided $\|A^{-1}\|\,\|E\| < 1$. If we assume $\|E\| = \alpha\|A\|$ we have

$$(2.3)\qquad \|(A + E)^{-1} - A^{-1}\|/\|A^{-1}\| \leqq \frac{\alpha\,\varkappa(A)}{1 - \alpha\,\varkappa(A)}$$

where

$$(2.4)\qquad \varkappa(A) = \|A\|\,\|A^{-1}\|\,.$$

Relations (2.3) and (2.4) give us a bound for the relative error in the norm. If we use the l_2 norm we have $\|QA\|_2 = \|A\|_2$ and $\|(QA)^{-1}\|_2 = \|A^{-1}\|_2$ for any unitary matrix. The sensitivity of the inverse *measured in the l_2 norm* is therefore invariant with respect to unitary transformations.

Turning to the eigenproblem we consider only diagonalizable matrices. We define X to be the non-singular matrix such that

$$(2.5)\qquad X^{-1}AX = \operatorname{diag}(\lambda_i)\,.$$

Then if $B = A + E$ and the eigenvalues of A are denoted by λ_i the Bauer-Fike theorem [1960] states that the eigenvalues of B lies in the discs with centres λ_i and radii equal to $\varkappa(X)\,\|E\|$. The overall sensitivity of the eigenvalues therefore depends on $\varkappa(X)$. Since X is not unique we can clearly use the minimum value of $\varkappa(X)$ for all permissible X.

This result suffers from the disadvantage that if any eigenvalue is sensitive then $\varkappa(X)$ must be large and the bound for all perturbations are accordingly affected. If we consider perturbations $A + \varepsilon E$ for small values of ε, then as ε tends to zero $A + \varepsilon E$ has an eigenvalue λ_i^1 such that

$$(2.6)\qquad \lambda_i^1 - \lambda_i \sim \varepsilon y_i^H E x_i / y_i^H x_i\,,$$

where y_i and x_i are normalized left-hand and right-hand eigenvectors corresponding to λ_i. Since $|y_i^H E x_i| \leqq \|E\|_2$ the sensitivities of the individual eigenvalues depend primarily upon the quantities $1/y_i^H x$.

When A is normal $y_i = x_i$ and hence the eigenvalues are always well determined. For hermitian matrices satisfying

$$B = A + C \tag{2.7}$$

if the eigenvalues of A, B, and C are α_i, β_i, and γ_i, each arranged in decreasing order, then Weyl's theorem [1912] gives

$$\gamma_n \leqq \beta_i - \alpha_i \leqq \gamma_1 \tag{2.8}$$

or

$$|\beta_i - \alpha_i| \leqq \|C\|_2 . \tag{2.9}$$

If B and A are normal matrices and satisfy (2.8) then the Hoffman-Wielandt theorem [1953] gives

$$(\sum|\beta_i - \alpha_i|^2)^{1/2} \leqq \|C\|_F \tag{2.10}$$

where $\|C\|_F$ is the Frobenius norm. If A and B are hermitian the result (2.10) is true when the α_i and β_i are arranged in decreasing order.

3. ALGORITHMS BASED ON ELEMENTARY TRANSFORMATIONS

A number of algorithms for inverting matrices, solving linear equations and computing eigensystems depend on the use of a sequence of transformations which reduce the original matrix to some condensed form, the solution of the problem for the original matrix being deducible from that for the transformed matrix. For matrix inversion and the solution of linear equations equivalence transformations are used, while for the calculation of eigensystems it is similarity transformations which are relevant.

For inversion, reduction to triangular form is adequate, though some algorithms give a complete reduction to diagonal form. For the eigenproblem, reduction to triangular or diagonal form (the latter may not exist) gives the solution immediately but the fundamental theorem of algebra shows that, in general, an infinite number of steps are required. In practice one continues until the matrix has negligible subdiagonal (or off-diagonal) elements. Reduction to Hessenberg form H (i.e. $h_{ij} = 0$, $i > j + 1$) can be achieved in a finite number of steps and gives a considerable simplification.

Two types of transformation matrix are in common use. The first type are the unitary transformations, and the matrices in question are either plane rotations or of the type I-$2ww^H$ where $\|w\|_2 = 1$. Transformations with unitary matrices have the advantages (i) the sensitivities of the matrix with respect to inversion (measured in the l_2 norm) and to the eigenvalue problem are invariant. We cannot therefore have a severe increase in the sensitivity as the reduction proceeds. (ii) In the case of

the eigenproblem the hermitian property is retained by unitary similarity transformations and hence at all stages one can work with only half the matrix, the super-diagonal or sub-diagonal half as proves convenient.

The second type of transformation is based on the use of elementary matrices of the type $I-ue_r^T$ and $I-e_ru^T$ where e_r is the rth column of the identity matrix. Transformations of this type are simpler than those based on the use of unitary matrices; they include all the algorithms usually referred to as *elimination methods*. If used in a straightforward manner the elements of u can be unbounded and the transformed matrices may be arbitrarily more sensitive than the original. However, when there is a related algorithm based on unitary transformations, elementary matrices of the type $I-ue_r^T$ may be used in combination with elementary permutation matrices (i.e. simple substitutions) to give algorithms in which the elements of u are bounded by unity. Such algorithms are usually said to include "pivoting" and the transformations have been called *stabilized elementary transformations*. (WILKINSON [1965]). It is worth commenting that although elementary transformations may lead to matrices with greater sensitivity than the original (even if stabilized) there is the possibility of reducing the sensitivity, whereas this cannot be done with unitary transformations. If for example we start with a non-normal but diagonalizable matrix having a very sensitive eigenvalue problem it can be reduced to diagonal form by elementary transformations (i.e. to an insensitive form) but not by unitary transformations.

4. ERROR ANALYSIS OF TRANSFORMATION ALGORITHMS

If a transformation type algorithm is used in practice, then because of rounding errors the computed sequence of transformed matrices $\bar{A}_2, \bar{A}_3, \ldots$ differs from the sequence $A_2, A_3, \ldots$ which would have resulted from exact computation. It might be felt that for a stable algorithm satisfactory bounds could be obtained for $\|A_s - \bar{A}_s\|$ but this is not so. In fact many of the more stable algorithms can produce an $\bar{A}_s$ such that $\|A_s - \bar{A}_s\|$ is of the same order of magnitude as $\|A_1\|$.

For the sake of definiteness we shall concentrate on similarity transformation and consider an algorithm which defines sequences of matrices A_r and P_r satisfying the relations

$$A_r = P_{r-1}^{-1} A_{r-1} P_{r-1} \tag{4.1}$$

where the final A_s is of some simple form.

A profitable line of approach is the following. At the $(r-1)$th stage the algorithm defines an exact transformation matrix Q_{r-1} corresponding to the computed $\bar{A}_{r-1}$. In the error analysis we attempt to obtain a bound for the difference between the computed $\bar{A}_r$ and the exact transform $Q_{r-1}^{-1}\bar{A}_{r-1}Q_{r-1}$ defined by the algorithm for the computed $\bar{A}_{r-1}$. The computed $\bar{A}_r$ differs from $Q_{r-1}^{-1}\bar{A}_{r-1}Q_{r-1}$ for two independent

reasons. First errors are made in computing Q_{r-1} and as a result $\bar{Q}_{r-1}$ is obtained instead. In general it is compartively easy to find bounds for $\|\bar{Q}_{r-1} - Q_{r-1}\|$. Then errors are made in computing $\bar{Q}_{r-1}^{-1}\bar{A}_{r-1}\bar{Q}_{r-1}$ using the computed $\bar{Q}_{r-1}$. With stable algorithms it is possible to find a satisfactory bound for $\|\bar{A}_r - \bar{Q}_{r-1}\bar{A}_{r-1}\bar{Q}_{r-1}\|$. Combining these two results a bound is obtained for $\|\bar{A}_r - Q_{r-1}^{-1}\bar{A}_{r-1}Q_{r-1}\|$. If we write

$$\bar{A}_r - Q_{r-1}\bar{A}_{r-1}Q_{r-1} = F_r \tag{4.2}$$

then combining relations of the type (4.2) for $r = 2, 3, \ldots$ we have

$$\bar{A}_r = F_r + S_{r-1}^{-1}F_{r-1}S_{r-1} + S_{r-2}^{-1}F_{r-2}S_{r-2} + \ldots + S_2^{-1}F_2S_2 + S_1^{-1}A_1S_1\,, \tag{4.3}$$

where

$$Q_kQ_{k+1}\cdots Q_{r-1} = S_k\,. \tag{4.4}$$

This may be written in the form

$$\bar{A}_r = G_r + S_1^{-1}A_1S_1 \tag{4.5}$$

where

$$G_r = F_r + S_{r-1}^{-1}F_{r-1}S_{r-1} + \ldots + S_2^{-1}F_2S_2 \tag{4.6}$$

showing that $\bar{A}_r$ differs from an exact similarity transformation of A_1 by the matrix G_r. In practice a bound for G_r is required.

Equation (4.5) may be written in the alternative form

$$\bar{A}_r = S_1^{-1}(A_1 + S_1G_rS_1^{-1})\,S_1 = S_1^{-1}(A_1 + H_1)\,S_1 \tag{4.7}$$

showing that the eigenvalues of $\bar{A}_r$ are exactly those of $A_1 + H_1$. If we can obtain a bound for H_1 we have a bound for the perturbation of A_1 which gives the same affect as the errors made in the course of the computation. The derivation of a bound for H_1 is usually referred to as a *backward* error analysis (WILKINSON [1965]). It has the advantage of putting the errors made in the course of the solution on the same footing as errors in the data.

5. ERROR BOUNDS FOR UNITARY TRANSFORMATIONS

General *a priori* error bounds for G_r and H_1 have been found for quite large classes of transformations based on the use of unitary transformations. For these algorithms the Q_i are exactly unitary since they are the exact transformations defined by the algorithms for the computed $\bar{A}_i$. From (4.4) we see that the S_k are exactly unitary and hence $\|H_1\| = \|G_r\|$ for either the l_2 or the Frobenius norm, since these are unitarily invariant. Further we have from (4.6)

$$\|G_r\| \leqq \|F_r\| + \|F_{r-1}\| + \ldots + \|F_2\|\,, \tag{5.1}$$

so that the effect of successive rounding errors is no more than additive.

In the case when the transformations are plane rotations it is comparatively trivial to show for a large class of algorithms that, using floating point computation with a t-digit mantissa,

$$\|Q_r - \bar{Q}_r\| \leqq \alpha \,.\, 2^{-t}, \tag{5.2}$$

where α is a constant of order unity which depends on the rounding procedure. The computation of $\bar{A}_{r+1}$ from $\bar{A}_r$ takes place in two steps

$$B_{r+1} = \mathrm{fl}\,(\bar{Q}_r^H \bar{A}_r) \tag{5.3}$$

and

$$\bar{A}_{r+1} = \mathrm{fl}\,(B_{r+1}\bar{Q}_r) \tag{5.4}$$

where the notation fl( ) means that the right-hand side is computed using floating-point arithmetic. Again it is easy to show that

$$B_{r+1} = \bar{Q}_r^H \bar{A}_r + K_r, \quad \|K_r\| \leqq \beta \,.\, 2^{-t}\|\bar{A}_r\| \tag{5.5}$$

$$\bar{A}_{r+1} = B_{r+1}\bar{Q}_r + L_r, \quad \|L_r\| \leqq \beta \,.\, 2^{-t}\|B_{r+1}\| \tag{5.6}$$

where again β is of order unity. This gives

$$\begin{aligned} \|B_{r+1}\| &\leqq \|\bar{Q}_r^H\|\,\|\bar{A}_r\| + \|K_r\| \leqq (1 + (\alpha + \beta)\,2^{-t})\,\|\bar{A}_r\| \\ &= (1 + \gamma \,.\, 2^{-t})\,\|\bar{A}_r\| \end{aligned} \tag{5.7}$$

and similarly

$$\|\bar{A}_{r+1}\| \leqq (1 + \gamma \,.\, 2^{-t})\,\|B_{r+1}\| \leqq (1 + \gamma \,.\, 2^{-t})^2\,\|\bar{A}_r\| \,. \tag{5.8}$$

Similarly

$$\|\bar{A}_{r+1}\| \geqq (1 - \gamma \,.\, 2^{-t})^2\,\|\bar{A}_r\| \,. \tag{5.9}$$

Equations (5.8) and (5.9) reveal an essential feature of algorithms based on plane rotations. Clearly we have

$$(1 - \gamma \,.\, 2^{-t})^{2r}\,\|A_1\| \leqq \|\bar{A}_{r+1}\| \leqq (1 + \gamma \,.\, 2^{-t})^{2r}\,\|A_1\|, \tag{5.10}$$

showing that the norm of successive $\bar{A}_{r+1}$ can vary only slowly. Combining the above bounds for successive values of r we have

$$\|H_1\| = \|G_r\| \leqq \gamma \,.\, 2^{-t}\,\|A_1\|\,(1 + (1 + \gamma \,.\, 2^{-t}) + \ldots + (1 + \gamma \,.\, 2^{-t})^{2r-1}) \tag{5.11}$$

giving certainly

$$\|H_1\| = \|G_r\| \leqq 2r\gamma \,.\, 2^{-t}\|A_1\|\,(1 + \gamma \,.\, 2^{-t})^{2r-2} \,. \tag{5.12}$$

Typical of algorithms covered by this analysis are Given's algorithms for the reduction of a general matrix to Hessenberg form or of a real symmetric matrix to tri-diagonal form. For these $r = \frac{1}{2}(n-1)(n-2)$ and hence we have certainly

$$\|H_1\| = \|G_r\| \leqq \gamma n^2 \,.\, 2^{-t} \|A_1\| (1 + \gamma \,.\, 2^{-t})^{n^2} \,. \tag{5.13}$$

By more sophisticated arguments the factor n^2 in the above bound can be reduced to $n^{3/2}$. More recently VOYEVODIN has shown that by simple modifications of algorithms of this type the n^2 factor may be further reduced. However these refinements should not be allowed to obscure the essential simplicity of the arguments.

In a similar manner it can be shown that for many algorithms based on the use of matrices of the type $I - 2ww^H$ one can obtain bounds of the form

$$\|H_1\| = \|G_r\| = 2r\delta \,.\, 2^{-t} \|A_1\| (1 + \delta \,.\, 2^{-t})^{2r-2} \tag{5.14}$$

provided inner-products are accumulated in double-precision and rounded on completion. Typical of algorithms covered by this analysis are Householder's algorithms for the reduction of a general matrix to Hessenberg form and of a real symmetric matrix to tri-diagonal form. For these $r = n - 2$ and hence we have certainly

$$\|H_1\| = \|G_r\| \leqq 2\delta n \,.\, 2^{-t} \|A_1\| (1 + \delta \,.\, 2^{-t})^{2n} \,. \tag{5.15}$$

When A_1 is normal the Wielandt-Hoffman theorem immediately gives *a priori* bounds for the errors in the eigenvalues themselves. When A_1 is hermitian either the Wielandt-Hoffman theorem or the Weyl theorem may be used to give *a priori* bounds for the errors in the eigenvalues. For non-normal matrices bounds for the errors in the eigenvalues depend on their sensitivities, as they must.

6. ALGORITHMS BASED ON ELEMENTARY TRANSFORMATIONS

It is a remarkable fact that for no algorithms based on elementary transformations have useful *a priori* bounds been found when A is a general matrix. For specialised matrices however there are stable algorithms using elementary transformations. We may mention in particular Gaussian elimination without pivoting for the inversion of positive definite matrices and Hyman's algorithm for the evaluation of determinants of Hessenberg matrices. It is interesting that pivoting is unnecessary with either of these algorithms and both of them are *scaling invariant.*

7. A POSTERIORI ERROR BOUNDS

The main purpose of *a priori* error analyses is to reveal the basic weaknesses of an algorithm and to gain some idea of its fundamental limitations. In practice the determination of *a posteriori* bounds is of much more value.

For matrix inversion this is trivial since from a computed inverse X we can determine E defined by

$$E = I - AX \tag{7.1}$$

and we have if $\|E\| < 1$ then

$$\|A^{-1} - X\|/\|A^{-1}\| \leqq \|E\| \tag{7.2}$$

and

$$\|A^{-1} - X\| \leqq \|X\| \, \|E\|/(1 - \|E\|) . \tag{7.3}$$

To obtain reliable information it is essential that E should be determined using accumulation of inner-products or the error made in determining E may be longer than its true value.

For hermitian matrices an error bound for an eigenvalue may be determined from any alleged eigenvector. For if

$$Ax - \lambda x = r , \quad \|x\|_2 = 1 , \quad \|r\|_2 = \varepsilon \tag{7.4}$$

then there is at least one eigenvalue of A in the disc with centre λ and radius ε. If further it is known that there is only one eigenvalue (possibly multiple) in this disc and all distinct eigenvalues are outside a disc with centre λ and radius a, then provided λ is the Rayleigh value corresponding to x there is at least one eigenvalue in the disc with centre λ and radius ε^2/a. Again to obtain the full benefit of this result r must be computed accurately. When a complete eigensystem has been determined then these results enable us to place each of the n eigenvalues in a separate disc provided no two discs overlap. When discs overlap the information given by the above inclusion theorem is incomplete. Wilkinson [1965] has shown that if $x_1, x_2, \ldots, x_s$ are orthonormal vectors and

$$Ax_i - \lambda_i x_i = r_i , \quad \|r_i\|_2 = \varepsilon_i , \quad i = 1, \ldots, s \tag{7.5}$$

then there are s eigenvalues $\mu_1, \mu_2, \ldots, \mu_s$ of A satisfying the relation

$$(\sum(\lambda_i - \mu_i)^2)^{1/2} \leqq 2^{1/2}(\sum \varepsilon_i^2)^{1/2} . \tag{7.6}$$

More recently KAHAN has shown that the factor $2^{1/2}$ is unnecessary.

When A is non-hermitian it is not easy to obtain precise information from a partial eigensystem. However if X is a complete set of computed eigenvectors and diag (λ_i) a complete set of eigenvalues we can proceed as follows. First compute R defined by

$$AX - X\Lambda = R . \tag{7.7}$$

This residual matrix must be computed accurately and an error bound determined for the computed $\bar{R}$. Then $X^{-1}\bar{R}$ is determined and error bounds for this matrix.

We then have

$$X^{-1}AX - \Lambda = S + E \tag{7.8}$$

where E is the total error bound for computed $X^{-1}R$. Here S is obtained explicitly but E is merely a set of bounds, in practice comprising a matrix which is far smaller than E which is itself smaller than A. The eigenvalues of A are then those of $A + S + E$ and this is almost diagonal. Gerschgorin's theorem may be used to obtain improved eigenvalues and rigorous bounds for their errors.

8. LIMITATIONS OF CURRENT ERROR ANALYSIS

The *a priori* error analysis have been discussed in connexion with obtaining a bound for a norm of the equivalent perturbation of the original matrix and the l_2 and Frobenius norms have been widely used. Though this is often satisfactory it is not invariably so and sometimes a very different norm may be appropriate.

As far as inversion is concerned we can deduce the inverse of A from the inverse of D_1AD_2 where D_1 and D_2 are any non singular diagonal matrices. The condition number $\varkappa(D_1AD_2)$ will in general be different from $\varkappa(A)$ and there is much to be said for choosing D_1 and D_2 so as to make $\varkappa(D_1AD_2)$ a minimum. This problem has been discussed by FORSYTHE and STRAUS [1955] and by BAUER [1963] for a number of different norms. In practice a useful stratagem is to scale all rows of A so that the scaled A has all its row sums equal. This is a justifiable strategy if A is inverted (or linear equations solved) using Householder's triangularization [1963] or Gaussian triangularization with partial pivoting (Wilkinson [1963]). In practice one need scale only by powers of two (or ten) chosen so as to make the row sums as nearly equal as possible.

For the eigenproblem one could work with $D^{-1}AD$ where D is any non-singular diagonal matrix. It would appear to be advantageous to choose D so that $D^{-1}AD$ has a matrix of eigenvectors X with minimum condition number $\varkappa(X)$. Since X is not known *ab initio* there would appear to be little chance of doing this. However, we can choose D so as to minimise some norm of A. The relationship of such a preliminary scaling to that which gives a minimum $\varkappa(X)$ is a little obscure. The problem has been discussed by OSBORNE [1960]. At NPL we have scaled so that $\sum\sum|a_{ij}|$ is a minimum for the scaled matrix, or at least as near the minimum as can be achieved using powers of two. For many eigenproblems this preliminary scaling is vital.

There are also eigenproblems in which it is important to obtain small eigenvalues with a low relative error and not with a low error relative to $\|A\|_2$. Small eigenvalues can sometimes be extremely insensitive to small relative variations in the elements of A. For the matrix of order 10 with

$$a_{ii} = (11 - i)^{12}\,, \quad a_{i,i+1} = a_{i+1,i} = 1\,, \quad a_{ij} = 0 \text{ otherwise}$$

the smallest eigenvalue is of order unity. It is fantastically insensitive to variations in a_{11} for example. This element can take any value such that $|a_{11}| \geqq 3$ without affecting the smallest eigenvalue by more than $2/(9!)^{24}$. Further study of the factors affecting the sensitivity of such small eigenvalues is very important.

References

[1912] *H. Weyl,* Das asymptotische Verteilungsgesetz der Eigenwerte. Math. Annalen *LXXI,* 441—479.

[1953] *A. J. Hoffman* and *H. W. Wielandt,* The variation of the spectrum of a normal matrix. Duke Math. J. *20,* 37—39.

[1955] *G. E. Forsythe* and *E. G. Straus,* On best conditioned matrices. Proc. Amer. Math. Soc. *6,* 340—345.

[1960] *F. L. Bauer* and *C. T. Fike,* Norms and exclusion theorems. Num. Math. *2,* 137—141.

[1960] *E. E. Osborne,* On pre-conditioning of matrices. J. Ass. Comp. Mach. *7,* 335—345.

[1963] *F. L. Bauer,* Optionally scaled matrices. Num. Math. *5,* 73—87.

[1963] *J. H. Wilkinson,* Rounding Errors in Algebraic Processes. Notes on Applied Science No. 32. Her Majesty's Stationery Office, London; Prentice Hall, New Jersey.

[1965] *J. H. Wilkinson,* The Algebraic Eigenvalue Problem, Oxford University Press, London.

Prof. *James Hardy Wilkinson,* Mathematics Division, National Physical Laboratory, Teddington, Middx., England.