

Chapter 6

Algebraic eigenvalue problems

Das also war des Pudels Kern! GOETHE.

6.0. Introduction

Determination of eigenvalues and eigenvectors of matrices is one of the most important problems of numerical analysis. Theoretically, the problem has been reduced to finding the roots of an algebraic equation and to solving n linear homogeneous systems of equations. In practical computation, as a rule, this method is unsuitable, and better methods must be applied.

When there is a choice between different methods, the following questions should be answered:

- (a) Are both eigenvalues and eigenvectors asked for, or are eigenvalues alone sufficient?
- (b) Are only the absolutely largest eigenvalue(s) of interest?
- (c) Does the matrix have special properties (real symmetric, Hermitian, and so on)?

If the eigenvectors are not needed less memory space is necessary, and further, if only the largest eigenvalue is wanted, a particularly simple technique can be used. Except for a few special cases a direct method for computation of the eigenvalues from the equation $\det(\lambda I - A) = 0$ is never used. Further it turns out that practically all methods depend on transforming the initial matrix one way or other without affecting the eigenvalues. The table on p. 114 presents a survey of the most important methods giving initial matrix, type of transformation, and transformation matrix. As a rule, the transformation matrix is built up successively, but the resulting matrix need not have any simple properties, and if so, this is indicated by a horizontal line. It is obvious that such a compact table can give only a superficial picture; moreover, in some cases the computation is performed in two steps. Thus a complex matrix can be transformed to a normal matrix following Eberlein, while a normal matrix can be diagonalized following Goldstine-Horwitz. Incidentally, both these procedures can be performed simultaneously giving a unified method as a result. Further, in some cases we have recursive techniques which differ somewhat in principle from the other methods.

It is not possible to give here a complete description of all these methods because of the great number of special cases which often give rise to difficulties. However, methods which are important in principle will be treated carefully

Initial matrix	Technique	Transformation matrix	Method, originator
Real	Iteration and deflation		Power method
Real, symmetric	Diagonalization	Orthogonal	Jacobi
Hermitian	Diagonalization	Unitary	Generalized Jacobi
Normal	Diagonalization	Unitary	Goldstine-Horwitz
Real, symmetric	Tridiagonalization	Orthogonal (rotations)	Givens
Real, symmetric	Tridiagonalization	Orthogonal (reflections)	Householder
Real	Tridiagonalization		Lanczos
Real	Tridiagonalization	Unitary	La Budde
Real	Triangularization	Unitary	Ruthishauser (LR)
Hessenberg	Triangularization	Unitary (rotations)	Francis (QR)
Complex	Triangularization	Unitary (reflections)	Lotkin, Greenstadt
Complex	Reduction to Hessenberg form		Givens
Complex	Reduction to Hessenberg form		Householder
Complex	Reduction to normal matrix		Patricia Eberlein
Tridiagonal	Sturm sequence and interpolation		Givens, Wilkinson, Wielandt
Hessenberg	Recursive evaluation and interpolation		Hyman, Wielandt

and in other cases at least the main features will be discussed. On the whole we can distinguish four principal groups with respect to the kind of transformation used initially:

1. Diagonalization,
2. Almost diagonalization (tridiagonalization),
3. Triangularization,
4. Almost triangularization (reduction to Hessenberg form).

The determination of the eigenvectors is trivial in the first case and almost trivial in the third case. In the other two cases a recursive technique is easily established which will work without difficulties in nondegenerate cases. To a certain amount we shall discuss the determination of eigenvectors, for example, Wilkinson's technique which tries to avoid a dangerous error accumulation. Also Wielandt's method, aiming at an improved determination of approximate eigenvectors, will be treated.

6.1. The power method

We assume that the eigenvalues of A are $\lambda_1, \lambda_2, \dots, \lambda_n$, where $|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|$. Now we let A operate repeatedly on a vector \mathbf{v} , which we express as a linear combination of the eigenvectors

$$\mathbf{v} = c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_n \mathbf{v}_n. \quad (6.1.1)$$

Then we have

$$A\mathbf{v} = c_1 A\mathbf{v}_1 + c_2 A\mathbf{v}_2 + \dots + c_n A\mathbf{v}_n = \lambda_1 \left(c_1 \mathbf{v}_1 + c_2 \frac{\lambda_2}{\lambda_1} \mathbf{v}_2 + \dots + c_n \frac{\lambda_n}{\lambda_1} \mathbf{v}_n \right)$$

and through iteration we obtain

$$A^p \mathbf{v} = \lambda_1^p \left\{ c_1 \mathbf{v}_1 + c_2 \left(\frac{\lambda_2}{\lambda_1} \right)^p \mathbf{v}_2 + \dots + c_n \left(\frac{\lambda_n}{\lambda_1} \right)^p \mathbf{v}_n \right\}. \quad (6.1.2).$$

For large values of p , the vector

$$c_1 \mathbf{v}_1 + c_2 \left(\frac{\lambda_2}{\lambda_1} \right)^p \mathbf{v}_2 + \dots + c_n \left(\frac{\lambda_n}{\lambda_1} \right)^p \mathbf{v}_n$$

will converge toward $c_1 \mathbf{v}_1$, that is, the eigenvector of λ_1 . The eigenvalue is obtained as

$$\lambda_1 = \lim_{p \rightarrow \infty} \frac{(A^{p+1}\mathbf{v})_r}{(A^p \mathbf{v})_r}, \quad r = 1, 2, \dots, n, \quad (6.1.3)$$

where the index r signifies the r th component in the corresponding vector. The rate of convergence is determined by the quotient $|\lambda_2/\lambda_1|$; convergence is faster the

smaller $|\lambda_2/\lambda_1|$ is. For numerical purposes the algorithm just described can be formulated in the following way. Given a vector \mathbf{y}_k , we form two other vectors, \mathbf{y}_{k+1} and \mathbf{z}_{k+1} :

$$\begin{cases} \mathbf{z}_{k+1} = A\mathbf{y}_k, \\ \mathbf{y}_{k+1} = \mathbf{z}_{k+1}/\alpha_{k+1}, \\ \alpha_{k+1} = \max_r |(\mathbf{z}_{k+1})_r|. \end{cases} \quad (6.14)$$

The initial vector \mathbf{y}_0 should be chosen in a convenient way, often one tries a vector with all components equal to 1.

EXAMPLE

$$A = \begin{pmatrix} 1 & -3 & 2 \\ 4 & 4 & -1 \\ 6 & 3 & 5 \end{pmatrix}.$$

Starting from

$$\mathbf{y}_0 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

we find that

$$\mathbf{y}_5 = \begin{pmatrix} 0.3276 \\ 0.0597 \\ 1 \end{pmatrix}, \quad \mathbf{y}_{10} = \begin{pmatrix} 0.3007 \\ 0.0661 \\ 1 \end{pmatrix},$$

$$\mathbf{y}_{15} = \begin{pmatrix} 0.3000 \\ 0.0667 \\ 1 \end{pmatrix}, \quad \text{and} \quad \mathbf{z}_{10} = \begin{pmatrix} 2.0999 \\ 0.4668 \\ 7.0001 \end{pmatrix}.$$

After round-off, we get

$$\lambda_1 = 7 \quad \text{and} \quad \mathbf{v}_1 = \begin{pmatrix} 9 \\ 2 \\ 30 \end{pmatrix}.$$

If the matrix A is Hermitian and all eigenvalues are different, the eigenvectors, as shown before, are orthogonal. Let \mathbf{x} be the vector obtained after p iterations:

$$\mathbf{x} = c_1 \mathbf{v}_1 + c_2 \left(\frac{\lambda_2}{\lambda_1}\right)^p \mathbf{v}_2 + \cdots + c_n \left(\frac{\lambda_n}{\lambda_1}\right)^p \mathbf{v}_n = c_1 \mathbf{v}_1 + \varepsilon_2 \mathbf{v}_2 + \cdots + \varepsilon_n \mathbf{v}_n.$$

We suppose that all \mathbf{v}_i are normalized:

$$\mathbf{v}_i^H \mathbf{v}_k = \delta_{ik}.$$

Then we have

$$A\mathbf{x} = c_1\lambda_1\mathbf{v}_1 + \varepsilon_2\lambda_2\mathbf{v}_2 + \cdots + \varepsilon_n\lambda_n\mathbf{v}_n$$

and

$$\mathbf{x}^H A \mathbf{x} = \lambda_1 |c_1|^2 + \lambda_2 |\varepsilon_2|^2 + \cdots + \lambda_n |\varepsilon_n|^2.$$

Further, $\mathbf{x}^H \mathbf{x} = |c_1|^2 + |\varepsilon_2|^2 + \cdots + |\varepsilon_n|^2$. When p increases, all ε_i tend to zero, and with $\mathbf{x} = \text{const } A^p(\sum c_i \mathbf{v}_i)$, we get Rayleigh's quotient

$$\lambda_1 = \lim_{p \rightarrow \infty} \frac{\mathbf{x}^H A \mathbf{x}}{\mathbf{x}^H \mathbf{x}}. \quad (6.1.5)$$

Example

$$A = \begin{pmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{pmatrix} \quad \text{and} \quad \mathbf{x}_0 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix},$$

With

we obtain for $p = 1, 2,$ and $3,$ $\lambda_1 = 29.75, 30.287,$ and 30.288662 respectively, compared with the correct value 30.28868 . The corresponding eigenvector is

$$\begin{pmatrix} 0.95761 \\ 0.68892 \\ 1 \\ 0.94379 \end{pmatrix}.$$

The quotients of the individual vector components give much slower convergence; for example, $(x_3)_1 / (x_2)_1 = 30.25987$.

The power method can easily be modified in such a way that certain other eigenvalues can also be computed. If, for example, A has an eigenvalue λ , then $A - qI$ has an eigenvalue $\lambda - q$. Using this principle, we can produce the two outermost eigenvalues. Further, we know that λ^{-1} is an eigenvalue of A^{-1} and analogously that $(\lambda - q)^{-1}$ is an eigenvalue of $(A - qI)^{-1}$. If we know that an eigenvalue is close to q , we can concentrate on that, since $(\lambda - q)^{-1}$ becomes large as soon as λ is close to q .

We will now discuss how the absolutely next largest eigenvalue can be calculated if we know the largest eigenvalue λ_1 and the corresponding eigenvector \mathbf{x}_1 . Let \mathbf{a}^T be the first row vector of A and form

$$A_1 = A - \mathbf{x}_1 \mathbf{a}^T. \quad (6.1.6)$$

Here \mathbf{x}_1 is supposed to be normalized in such a way that the first component is 1 . Hence the first row of A_1 is zero. Now let λ_2 and \mathbf{x}_2 be an eigenvalue and the corresponding eigenvector with the first component of \mathbf{x}_2 equal to 1 . Then

we have

$$\begin{aligned} A_1(\mathbf{x}_1 - \mathbf{x}_2) &= A(\mathbf{x}_1 - \mathbf{x}_2) - \mathbf{x}_1 \mathbf{a}^T (\mathbf{x}_1 - \mathbf{x}_2) = \lambda_1 \mathbf{x}_1 - \lambda_2 \mathbf{x}_2 - (\lambda_1 - \lambda_2) \mathbf{x}_1 \\ &= \lambda_2 (\mathbf{x}_1 - \mathbf{x}_2), \end{aligned}$$

since $\mathbf{a}^T \mathbf{x}_1 = \lambda_1$ and $\mathbf{a}^T \mathbf{x}_2 = \lambda_2$ (note that the first component of \mathbf{x}_1 as well as of \mathbf{x}_2 is 1).

Thus λ_2 is an eigenvalue and $\mathbf{x}_1 - \mathbf{x}_2$ is an eigenvector of A_1 . Since $\mathbf{x}_1 - \mathbf{x}_2$ has the first component equal to 0, the first column of A_1 is irrelevant, and in fact we need consider only the $(n-1, n-1)$ -matrix, which is obtained when the first row and first column of A are removed. We determine an eigenvector of this matrix, and by adding a zero as first component, we get a vector \mathbf{z} . Then we obtain \mathbf{x}_2 from the relation

$$\mathbf{x}_2 = \mathbf{x}_1 + c\mathbf{z}.$$

Multiplying with \mathbf{a}^T we find $\mathbf{a}^T \mathbf{x}_2 = \mathbf{a}^T \mathbf{x}_1 + c\mathbf{a}^T \mathbf{z}$, and hence $c = (\lambda_2 - \lambda_1)/\mathbf{a}^T \mathbf{z}$. When λ_2 and \mathbf{x}_2 have been determined, the process, which is called *deflation*, can be repeated.

EXAMPLE

The matrix

$$A = \begin{pmatrix} -306 & -198 & 426 \\ 104 & 67 & -147 \\ -176 & -114 & 244 \end{pmatrix},$$

has an eigenvalue $\lambda_1 = 6$ and the corresponding eigenvector

$$\begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix},$$

or normalized,

$$\mathbf{x}_1 = \begin{pmatrix} 1 \\ -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix}.$$

Without difficulty we find

$$A_1 = \begin{pmatrix} 0 & 0 & 0 \\ -49 & -32 & 66 \\ -23 & -15 & 31 \end{pmatrix}.$$

Now we need consider only

$$B_1 = \begin{pmatrix} -32 & 66 \\ -15 & 31 \end{pmatrix},$$

and we find the eigenvalues $\lambda_2 = -2$ and $\lambda_3 = 1$, which are also eigenvalues of

the original matrix A . The two-dimensional eigenvector belonging to $\lambda_2 = -2$ is

$$\begin{pmatrix} 11 \\ 5 \end{pmatrix},$$

and hence

$$\mathbf{x}_2 = \mathbf{x}_1 + c\mathbf{z} = \begin{pmatrix} 1 \\ -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix} + c \begin{pmatrix} 0 \\ 11 \\ 5 \end{pmatrix}.$$

Since $\mathbf{a}^T \mathbf{z} = -48$, we get $c = \frac{1}{6}$ and

$$\mathbf{x}_2 = \begin{pmatrix} 1 \\ \frac{4}{3} \\ \frac{4}{3} \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 3 \\ 4 \\ 4 \end{pmatrix}.$$

With $\lambda_3 = 1$, we find

$$\mathbf{x}_3 = \begin{pmatrix} 1 \\ -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix} + c \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}$$

and $\mathbf{a}^T \mathbf{z} = 30$. Hence $c = -\frac{1}{6}$ and

$$\mathbf{x}_3 = \begin{pmatrix} 1 \\ -\frac{5}{6} \\ \frac{1}{3} \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 6 \\ -5 \\ 2 \end{pmatrix},$$

and all eigenvalues and eigenvectors are known.

If A is Hermitian, we have $\mathbf{x}_1^H \mathbf{x}_2 = 0$ when $\lambda_1 \neq \lambda_2$. Now suppose that $\mathbf{x}_1^H \mathbf{x}_1 = 1$,

and form

$$A_1 = A - \lambda_1 \mathbf{x}_1 \mathbf{x}_1^H. \quad (6.1.7)$$

It is easily understood that the matrix A_1 has the same eigenvalues and eigenvectors as A except λ_1 , which has been replaced by zero. In fact, we have $A_1 \mathbf{x}_1 = A \mathbf{x}_1 - \lambda_1 \mathbf{x}_1 \mathbf{x}_1^H \mathbf{x}_1 = \lambda_1 \mathbf{x}_1 - \lambda_1 \mathbf{x}_1 = \mathbf{0}$ and $A_1 \mathbf{x}_2 = A \mathbf{x}_2 - \lambda_1 \mathbf{x}_1 \mathbf{x}_1^H \mathbf{x}_2 = \lambda_2 \mathbf{x}_2$, and so on. Then we can again use the power method on the matrix A_1 .

EXAMPLE

$$A = \begin{pmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{pmatrix}; \quad \lambda_1 = 30.288686; \quad \mathbf{x}_1 = \begin{pmatrix} 0.528561 \\ 0.380255 \\ 0.551959 \\ 0.520933 \end{pmatrix};$$

$$A_1 = \begin{pmatrix} 1.53804 & 0.91234 & -0.83654 & -1.33984 \\ 0.91234 & 0.62044 & -0.35714 & -0.99979 \\ -0.83654 & -0.35714 & 0.77228 & 0.29097 \\ -1.33984 & -0.99979 & 0.29097 & 1.78053 \end{pmatrix}.$$

With the starting vector

$$y_0 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix},$$

we find the following values for Rayleigh's quotient: $\lambda_2 = 3.546, 3.8516,$ and 3.85774 compared with the correct value 3.858057 .

If the numerically largest eigenvalue of a real matrix A is complex, $\lambda \cdot e^{i\varphi}$, then $\lambda \cdot e^{-i\varphi}$ must also be an eigenvalue. It is also clear that if x_1 is the eigenvector belonging to $\lambda e^{i\varphi}$, then x_1^* is the eigenvector belonging to $\lambda e^{-i\varphi}$.

Now suppose that we use the power method with a real starting vector x : $x = c_1 x_1 + c_1^* x_1^* + \dots$. Then we form $A^m x$, with m so large that the contributions from all the other eigenvectors can be neglected. Further, a certain component of $A^m x$ is denoted by p_m . Then $p_m \simeq c \lambda^m \xi (e^{mi\varphi+i\theta+i\psi} + e^{-(mi\varphi+i\theta+i\psi)})$, where $c_1 = c e^{i\theta}$ and the initial component of x_1 corresponding to p is $\xi e^{i\psi}$. Hence

$$p_m \simeq 2c \lambda^m \cos(m\varphi + \alpha),$$

where we have put $\theta + \psi = \alpha$. Now we form

$$p_m p_{m+2} - p_{m+1}^2 \simeq 4c^2 \xi^2 \lambda^{2m+2} [\cos(m\varphi + \alpha) \cos((m+2)\varphi + \alpha) - \cos^2((m+1)\varphi + \alpha)] = -4c^2 \xi^2 \lambda^{2m+2} \sin^2 \varphi.$$

Hence

$$\lim_{m \rightarrow \infty} \frac{p_m p_{m+2} - p_{m+1}^2}{p_{m-1} p_{m+1} - p_m^2} = \lambda^2. \quad (6.1.8)$$

Then we easily find

$$\lim_{m \rightarrow \infty} \frac{\lambda^2 p_m + p_{m+2}}{2\lambda p_{m+1}} = \cos \varphi.$$

In particular, if $\varphi = \pi$, that is, if the numerically largest eigenvalues are of the form $\pm \lambda$ with real λ , then we have the simpler formula

$$\lim_{m \rightarrow \infty} \frac{p_{m+2}}{p_m} = \lambda^2. \quad (6.1.10)$$

6.2. Jacobi's method

In many applications we meet the problem of diagonalizing real, symmetric matrices. This problem is particularly important in quantum mechanics.

In Chapter 3 we proved that for a real symmetric matrix A , all eigenvalues are real, and that there exists a real orthogonal matrix O such that $O^{-1} A O$ is diagonal. We shall now try to produce the desired orthogonal matrix as a— product of very special orthogonal matrices. Among the off-diagonal elements

we choose the numerically largest element: $|a_{ik}| = \max$. The elements a_{ii} , a_{ik} , $a_{ki}(=a_{ik})$, and a_{kk} form a $(2, 2)$ -submatrix which can easily be transformed to diagonal form. We put

$$O = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix},$$

and get

$$D = O^{-1}AO = \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} a_{ii} & a_{ik} \\ a_{ik} & a_{kk} \end{pmatrix} \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}, \quad (6.2.1)$$

$$d_{ii} = a_{ii} \cos^2 \varphi + 2a_{ik} \sin \varphi \cos \varphi + a_{kk} \sin^2 \varphi,$$

$$d_{ik} = d_{ki} = -(a_{ii} - a_{kk}) \sin \varphi \cos \varphi + a_{ik}(\cos^2 \varphi - \sin^2 \varphi),$$

$$d_{kk} = a_{ii} \sin^2 \varphi - 2a_{ik} \sin \varphi \cos \varphi + a_{kk} \cos^2 \varphi.$$

Now choose the angle φ such that $d_{ik} = d_{ki} = 0$, that is, $\tan 2\varphi = 2a_{ik}/(a_{ii} - a_{kk})$. This equation gives 4 different values of φ , and in order to get as small rotations as possible we claim $-\pi/4 \leq \varphi \leq \pi/4$. Putting

$$R = \sqrt{(a_{ii} - a_{kk})^2 + 4a_{ik}^2} \quad \text{and} \quad \sigma = \begin{cases} 1 & \text{if } a_{ii} \geq a_{kk}, \\ -1 & \text{if } a_{ii} < a_{kk}, \end{cases}$$

we obtain:

$$\begin{cases} \sin 2\varphi = 2\sigma a_{ik}/R, \\ \cos 2\varphi = \sigma(a_{ii} - a_{kk})/R, \end{cases}$$

since the angle 2φ must belong to the first quadrant if $\tan 2\varphi > 0$ and to the fourth quadrant if $\tan 2\varphi < 0$. Hence we have for the angle φ :

$$\varphi = \frac{1}{2} \arctan (2a_{ik}/(a_{ii} - a_{kk})) \quad \text{if } a_{ii} \neq a_{kk},$$

$$\varphi = \begin{cases} \pi/4 & \text{when } a_{ik} > 0 \\ -\pi/4 & \text{when } a_{ik} < 0 \end{cases} \quad \text{if } a_{ii} = a_{kk},$$

where the value of the arctan-function is chosen between $-\pi/2$ and $\pi/2$. After a few simple calculations we get finally:

$$\begin{cases} d_{ii} = \frac{1}{2}(a_{ii} + a_{kk} + \sigma R), \\ d_{kk} = \frac{1}{2}(a_{ii} + a_{kk} - \sigma R), \\ d_{ik} = d_{ki} = 0. \end{cases} \quad (6.2.2)$$

(Note that $d_{ii} + d_{kk} = a_{ii} + a_{kk}$ and $d_{ii}d_{kk} = a_{ii}a_{kk} - a_{ik}^2$.)

We perform a series of such two-dimensional rotations; the transformation matrices have the form given above in the elements (i, i) , (i, k) , (k, i) , and (k, k) and are identical with the unit matrix elsewhere. Each time we choose such values i and k that $|a_{ik}| = \max$. We shall show that with the notation $P_r = O_1 O_2 \cdots O_r$ the matrix $B_r = P_r^{-1} A P_r$ for increasing r will approach a diagonal

matrix D with the eigenvalues of A along the main diagonal. Then it is obvious that we get the eigenvectors as the corresponding columns of $O = \lim_{r \rightarrow \infty} P_r$, since we have $O^{-1}AO = D$, that is, Let $AO = OD$. x_k be the k th column vector of O and λ_k the k th diagonal element of D . Then we have

$$Ax_k = \lambda_k x_k.$$

If $\sum_{k \neq i} |a_{ik}|$ is denoted by ε_i , we know from Gershgorin's theorem that $|a_{ii} - \lambda| < \varepsilon_i$ for some value of i , and if the process has been brought sufficiently far, every circle defined in this way contains exactly one eigenvalue. Thus it is easy to see when sufficient accuracy has been attained and the procedure can be discontinued.

The convergence of the method has been examined by von Neumann and Goldstine in the following way. We put $\tau^2(A) = \sum_i \sum_{k \neq i} a_{ik}^2 = N^2(A) - \sum_i a_{ii}^2$ and, as before, $B = O^{-1}AO$. The orthogonal transformation affects only the i th row and column and the k th row and column. Taking only off-diagonal elements into account, we find for $r \neq i$ and $r \neq k$ relations of the form

$$\begin{cases} a'_{ir} = a_{ir} \cos \varphi + a_{kr} \sin \varphi, \\ a'_{kr} = -a_{ir} \sin \varphi + a_{kr} \cos \varphi, \end{cases}$$

and hence $a'_{ir}{}^2 + a'_{kr}{}^2 = a_{ir}^2 + a_{kr}^2$. Thus $\tau^2(A)$ will be changed only through the cancellation of the elements a_{ik} and a_{ki} , that is,

$$\tau^2(A') = \tau^2(A) - 2a_{ik}^2.$$

Since a_{ik} was the absolutely largest of all $n(n-1)$ off-diagonal elements, we have

$$a_{ik}^2 \geq \frac{\tau^2(A)}{n(n-1)},$$

and

$$\tau^2(A') \leq \tau^2(A) \cdot \left(1 - \frac{2}{n(n-1)}\right) < \tau^2(A) \cdot \exp\left(-\frac{2}{n(n-1)}\right).$$

Hence we get the final estimate,

$$\tau(A') < \tau(A) \cdot \exp\left(-\frac{1}{n(n-1)}\right).$$

(6.2.3)

After N iterations, $\tau(A)$ has decreased with at least the factor $\exp(-N/n(n-1))$, and for a sufficiently large N we come arbitrarily close to the diagonal matrix containing the eigenvalues.

In a slightly different modification, we go through the matrix row by row performing a rotation as soon as $|a_{ik}| > \varepsilon$. Here ε is a prescribed tolerance which, of course, has to be changed each time the whole matrix has been passed. This modification seems to be more powerful than the preceding one.

The method was first suggested by Jacobi. It has proved very efficient for diagonalization of real symmetric matrices on automatic computers.

Example

$$A = \begin{pmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{pmatrix}.$$

Choosing $i = 3, k = 4$, we obtain, $\tan 2\varphi = 18/(10 - 10) = \infty$ and $\varphi = 45^\circ$.

After the first rotation, we have

$$A_1 = \begin{pmatrix} 10 & 7 & 15/\sqrt{2} & -1/\sqrt{2} \\ 7 & 5 & 11/\sqrt{2} & -1/\sqrt{2} \\ 15/\sqrt{2} & 11/\sqrt{2} & 19 & 0 \\ -1/\sqrt{2} & -1/\sqrt{2} & 0 & 1 \end{pmatrix}.$$

Here we take $i = 1, k = 3$, and obtain $\tan 2\varphi = 15\sqrt{2}/(10 - 19)$ and $\varphi = -33.5051^\circ$. After the second rotation we have

$$A_2 = \begin{pmatrix} 2.978281 & 1.543214 & 0 & -0.589611 \\ 1.543214 & 5 & 10.349806 & -0.707107 \\ 0 & 10.349806 & 26.021719 & -0.390331 \\ -0.589611 & -0.707107 & -0.390331 & 1 \end{pmatrix}$$

and after 10 rotations we have

$$A_{10} = \begin{pmatrix} 3.858056 & 0 & -0.000656 & -0.001723 \\ 0 & 0.010150 & 0.000396 & 0.000026 \\ -0.000656 & 0.000396 & 30.288685 & 0.001570 \\ -0.001723 & 0.000026 & 0.001570 & 0.843108 \end{pmatrix}$$

After 17 rotations the diagonal elements are 3.85805745, 0.01015005, 30.28868533, and 0.84310715, while the remaining elements are equal to 0 to 8 decimals accuracy. The sum of the diagonal elements is 35.99999999 and the product 1.00000015 in good agreement with the exact characteristic equation:

$$\lambda^4 - 35\lambda^3 + 146\lambda^2 - 100\lambda + 1 = 0.$$

Generalization to Hermitian matrices, which are very important in modern physics, is quite natural. As has been proved before, to a given Hermitian matrix H we can find a unitary matrix U such that $U^{-1}HU$ becomes a diagonal matrix. Apart from trivial factors, a two-dimensional unitary matrix has the form

$$U = \begin{pmatrix} \cos \varphi & -\sin \varphi \cdot e^{-i\theta} \\ \sin \varphi \cdot e^{i\theta} & \cos \varphi \end{pmatrix}.$$

A two-dimensional Hermitian matrix

$$H = \begin{pmatrix} a & b - ic \\ b + ic & d \end{pmatrix}$$

is transformed to diagonal form by $U^{-1}HU = D$, where

$$\begin{cases} d_{11} = a \cos^2 \varphi + d \sin^2 \varphi + 2(b \cos \theta + c \sin \theta) \sin \varphi \cos \varphi, \\ d_{22} = a \sin^2 \varphi + d \cos^2 \varphi - 2(b \cos \theta + c \sin \theta) \sin \varphi \cos \varphi, \\ d_{12} = d_{21}^* = (d - a) \sin \varphi \cos \varphi e^{-i\theta} - (b + ic) \sin^2 \varphi e^{-2i\theta} + (b - ic) \cos^2 \varphi. \end{cases}$$

Putting $d_{12} = 0$, we separate the real and imaginary parts and then multiply the resulting equations, first by $\cos \theta$ and $\sin \theta$, then by $-\sin \theta$ and $\cos \theta$, and finally add them together. Using well-known trigonometric formulas, we get

$$\begin{cases} b \sin \theta - c \cos \theta = 0, \\ (a - d) \sin 2\varphi - 2(b \cos \theta + c \sin \theta) \cos 2\varphi = 0. \end{cases} \quad (6.2.4)$$

In principle we obtain θ from the first equation and then φ can be solved from the second. Rather arbitrarily we demand $-\pi/2 \leq \theta \leq \pi/2$ and hence

$$\begin{cases} \sin \theta = c\sigma/r, \\ \cos \theta = b\sigma/r, \end{cases}$$

where

$$\sigma = \begin{cases} 1 & \text{when } b \geq 0, \\ -1 & \text{when } b < 0, \end{cases} \quad \text{and} \quad r = \sqrt{b^2 + c^2}.$$

Since $b \cos \theta + c \sin \theta = \sigma r$ the remaining equation has the solution

$$\begin{cases} \sin 2\varphi = 2\sigma\tau r/R, \\ \cos 2\varphi = \tau(a - d)/R, \end{cases}$$

with $\tau = \pm 1$ and $R = \sqrt{(a - d)^2 + 4r^2} = \sqrt{(a - d)^2 + 4(b^2 + c^2)}$. Now we want to choose φ according to $-\pi/4 \leq \varphi \leq \pi/4$ in order to get as small a rotation as possible which implies

$$\tau = \begin{cases} 1 & \text{for } a \geq d, \\ -1 & \text{for } a < d. \end{cases}$$

The following explicit solution is now obtained (note that b and c cannot both be equal to 0 because then H would already be diagonal):

$$\begin{cases} b \neq 0: & \theta = \arctan(c/b), \\ b = 0: & \theta = \begin{cases} \pi/2 & \text{if } c > 0, \\ -\pi/2 & \text{if } c < 0, \end{cases} \\ a - d \neq 0: & \varphi = \frac{1}{2} \arctan(2\sigma r/(a - d)), \\ a - d = 0: & \varphi = \begin{cases} \pi/4 & \text{if } b \geq 0, \\ -\pi/4 & \text{if } b < 0. \end{cases} \end{cases} \quad (6.2.5)$$

As usual the value of the arctan-function must be chosen between $-\pi/2$ and

$\pi/2$. The element d_{11} can now be written

$$d_{11} = \frac{1}{2}(a + d) + \frac{1}{2}(a - d) \cos 2\varphi + \sigma r \sin 2\varphi$$

and consequently:

$$\begin{cases} d_{11} = \frac{1}{2}(a + d + \tau R) , \\ d_{22} = \frac{1}{2}(a + d - \tau R) . \end{cases} \quad (6.2.6)$$

If $c = 0$ we get $\theta = 0$ and recover the result in Jacobi's method.

This procedure can be used repeatedly on larger Hermitian matrices, where the unitary matrices differ from the unit matrix only in four places. In the places (i, i) , (i, k) , (k, i) , and (k, k) , we introduce the elements of our two-dimensional matrix. The product of the special matrices U_1, U_2, \dots, U_k is a new unitary matrix approaching U when k is increased.

Finally we mention that a normal matrix (defined through $A^H A = A A^H$) can always be diagonalized with a unitary matrix. The process can be performed following a technique suggested by Goldstine and Horwitz [8] which is similar to the method just described for Hermitian matrices. The reduction of an arbitrary complex matrix to normal form can be accomplished through a method given by Patricia Eberlein [10]. In practice, both these processes are performed simultaneously.

6.3. Givens' method

Again we assume that the matrix A is real and symmetric. In Givens' method we can distinguish among three different phases. The first phase is concerned with $\frac{1}{2}(n-1)(n-2)$ orthogonal transformations, giving as result a band matrix with unchanged characteristic equation. In the second phase a sequence of functions $f_i(x)$ is generated, and it is shown that it forms a Sturm sequence, the last member of which is the characteristic polynomial. With the aid of the sign changes in this sequence, we can directly state how many roots larger than the inserted value x the characteristic equation has. By testing for a number of suitable values x , we can obtain all the roots. During the third phase, the eigenvectors are computed. The orthogonal transformations are performed in the following order. The elements a_{22}, a_{23}, a_{32} , and a_{33} define a two-dimensional subspace, and we start by performing a rotation in this subspace. This rotation affects all elements in the second and third rows and in the second and third columns. However, the quantity φ defining the orthogonal matrix

$$O = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}$$

is now determined from the condition $a'_{13} = a'_{31} = 0$ and not, as in Jacobi's method, by $a'_{23} = a'_{32} = 0$. We have $a'_{13} = -a_{12} \sin \varphi + a_{13} \cos \varphi = 0$ and $\tan \varphi = a_{13}/a_{12}$. The next rotation is performed in the $(2, 4)$ -plane with the new

φ determined from $a'_{14} = a'_{41} = 0$, that is, $\tan \varphi = a'_{14}/a'_{12}$. [Note that the (1, 2)-element was changed during the preceding rotation.] Now all elements in the second and fourth rows and in the second and fourth columns are changed, and it should be particularly observed that the element $a'_{13} = 0$ is not affected. In the same way, we make the elements a'_{15}, \dots, a'_{1n} equal to zero by rotations in the (2, 5)-, \dots , (2, n)-planes.

Now we pass to the elements $a'_{24}, a'_{25}, \dots, a'_{2n}$, and they are all set to zero by rotations in the planes (3, 4), (3, 5), \dots , (3, n). During the first of these rotations, the elements in the third and fourth rows and in the third and fourth columns are changed, and we must examine what happens to the elements a'_{13} and a'_{14} which were made equal to zero earlier. We find

$$\begin{aligned} a''_{13} &= a'_{13} \cos \varphi + a'_{14} \sin \varphi = 0, \\ a''_{14} &= -a'_{13} \sin \varphi + a'_{14} \cos \varphi = 0. \end{aligned}$$

Further, we get $a''_{24} = -a'_{23} \sin \varphi + a'_{24} \cos \varphi = 0$ and $\tan \varphi = a'_{24}/a'_{23}$. By now the procedure should be clear, and it is easily understood that we finally obtain a band matrix, that is, such a matrix that $a_{ik} = 0$ if $|i - k| > p$. In this special case we have $p = 1$. Now we put

$$B = \begin{pmatrix} \alpha_1 & \beta_1 & 0 & \dots & 0 \\ \beta_1 & \alpha_2 & \beta_2 & 0 & 0 \\ 0 & \beta_2 & \alpha_3 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \beta_{n-1} \\ 0 & 0 & \dots & \beta_{n-1} & \alpha_n \end{pmatrix}; \quad (6.3.1)$$

B has been obtained from A by a series of orthogonal transformations,

$$\begin{aligned} A_1 &= O_1^{-1} A O_1, \\ A_2 &= O_2^{-1} A_1 O_2 = (O_1 O_2)^{-1} A (O_1 O_2), \\ &\vdots \\ B = A_s &= O_s^{-1} A_{s-1} O_s = (O_1 O_2 \dots O_s)^{-1} A (O_1 O_2 \dots O_s) = O^{-1} A O, \end{aligned}$$

with $s = \frac{1}{2}(n-1)(n-2)$. In Chapter 3 it was proved that A and B have the same eigenvalues and further that, if \mathbf{u} is an eigenvector of A and \mathbf{v} an eigenvector of B (both with the same eigenvalue), then we have $\mathbf{u} = O\mathbf{v}$. Thus the problem has been reduced to the computation of eigenvalues and eigenvectors of the band matrix B .

We can suppose that all $\beta_j \neq 0$ [otherwise $\det(B - \lambda I)$ could be split into two determinants of lower order]. Now we form the following sequence of functions:

$$f_i(\lambda) = (\lambda - \alpha_i) f_{i-1}(\lambda) - \beta_{i-1}^2 f_{i-2}(\lambda), \quad (6.3.2)$$

with $f_0(\lambda) = 1$ and $\beta_0 = 0$. We find at once that $f_1(\lambda) = \lambda - \alpha_1$, which can be interpreted as the determinant of the (1, 1)-element in the matrix $\lambda I - B$.

Analogously, we have $f_2(\lambda) = (\lambda - \alpha_1)(\lambda - \alpha_2) - \beta_1^2$, which is the $(1, 2)$ -minor of $\lambda I - B$. By induction, it is an easy matter to prove that $f_n(\lambda)$ is the characteristic polynomial.

Next we shall examine the roots of the equation $f_i(\lambda) = 0$, $i = 1, 2, \dots, n$. For $i = 1$ we have the only root $\lambda = \alpha_1$. For $i = 2$ we observe that $f_2(-\infty) > 0$; $f_2(\alpha_1) < 0$; $f_2(+\infty) > 0$. Hence we have two real roots σ_1 and σ_2 with, for example, $-\infty < \sigma_1 < \alpha_1 < \sigma_2 < +\infty$. For $i = 3$ we will use a method which can easily be generalized to an induction proof. Then we write $f_2(\lambda) = (\lambda - \sigma_1)(\lambda - \sigma_2)$ and obtain from (6.3.2):

$$f_3(\lambda) = (\lambda - \alpha_3)(\lambda - \sigma_1)(\lambda - \sigma_2) - \beta_2^2(\lambda - \alpha_1).$$

Now it suffices to examine the sign of $f_3(\lambda)$ in a few suitable points:

λ	$-\infty$	σ_1	σ_2	$+\infty$
$\text{sign } [f_3(\lambda)]$	-	+	-	+

We see at once that the equation $f_3(\lambda) = 0$ has three real roots ρ_1, ρ_2 , and ρ_3 such that $-\infty < \rho_1 < \sigma_1 < \rho_2 < \sigma_2 < \rho_3 < +\infty$. In general, if $f_{i-2}(\lambda)$ has the roots $\sigma_1, \sigma_2, \dots, \sigma_{i-2}$ and $f_{i-1}(\lambda) = 0$ the roots $\rho_1, \rho_2, \dots, \rho_{i-1}$, then

$$f_i(\lambda) = (\lambda - \alpha_i)(\lambda - \rho_1)(\lambda - \rho_2) \cdots (\lambda - \rho_{i-1}) - \beta_{i-1}^2(\lambda - \sigma_1)(\lambda - \sigma_2) \cdots (\lambda - \sigma_{i-2}),$$

where

$$-\infty < \rho_1 < \sigma_1 < \rho_2 < \sigma_2 < \cdots < \rho_{i-2} < \sigma_{i-2} < \rho_{i-1} < +\infty.$$

By successively putting $\lambda = -\infty, \rho_1, \rho_2, \dots, \rho_{i-1}$, and $+\infty$, we find that $f_i(\lambda)$ has different signs in two arbitrary consecutive points. Hence $f_i(\lambda) = 0$ has i real roots, separated by the roots of $f_{i-1}(\lambda) = 0$.

We are now going to study the number of sign changes $V(\rho)$ in the sequence $f_0(\rho), f_1(\rho), \dots, f_n(\rho)$. It is evident that $V(-\infty) = n$ and $V(+\infty) = 0$. that a and b are two such real numbers that $f_i(\lambda) \neq 0$ in the closed interval $a \leq \lambda \leq b$. Then obviously $V(a) = V(b)$. First we examine what happens if the equation $f_i(\lambda) = 0$, $1 < i < n$, has a root ρ in the interval. From $f_{i+1}(\lambda) = (\lambda - \alpha_{i+1})f_i(\lambda) - \beta_i^2 f_{i-1}(\lambda)$ it follows for $\lambda = \rho$ that $f_{i+1}(\rho) = -\beta_i^2 f_{i-1}(\rho)$. Hence $f_{i-1}(\rho)$ and $f_{i+1}(\rho)$ have different signs, and clearly this is also true in an interval $\rho - \varepsilon < \lambda < \rho + \varepsilon$. Suppose, for example, that $f_{i-1}(\rho) < 0$; then we may have the following combination of signs:

λ	f_{i-1}	f_i	f_{i+1}
$\rho - \varepsilon$	-	-	+
$\rho + \varepsilon$	-	+	+

Hence, the number of sign changes does not change when we pass through a root of $f_i(\lambda) = 0$ if $i < n$. When $i = n$, however, the situation is different.

Suppose, for example, that n is odd. Denoting the roots of $f_n(\lambda) = 0$ by $\rho_1, \rho_2, \dots, \rho_n$ and the roots of $f_{n-1}(\lambda) = 0$ by $\sigma_1, \sigma_2, \dots, \sigma_{n-1}$, we have

$$\begin{array}{cccccc} \lambda & f_0 & f_1 & \cdots & f_{n-2} & f_{n-1} & f_n \\ \rho_1 - \varepsilon & + & - & \cdots & - & + & - \\ \rho_1 + \varepsilon & + & - & \cdots & - & + & + . \end{array}$$

Then we see that $V(\rho_1 - \varepsilon) - V(\rho_1 + \varepsilon) = 1$. Now we let λ increase until it reaches the neighborhood of σ_1 , where we find the following scheme:

$$\begin{array}{cccc} \lambda & \cdots & f_{n-2} & f_{n-1} & f_n \\ \sigma_1 - \varepsilon & \cdots & - & - & + \\ \sigma_1 + \varepsilon & \cdots & - & + & + . \end{array}$$

Hence $V(\sigma_1 - \varepsilon) - V(\sigma_1 + \varepsilon) = 0$. Then we let λ increase again (now a sign change of f_{n-2} may appear, but, as shown before, this does not affect V) until we reach the neighborhood of ρ_2 , where we have

$$\begin{array}{ccc} \lambda & \cdots & f_{n-1} & f_n \\ \rho_2 - \varepsilon & \cdots & - & + \\ \rho_2 + \varepsilon & \cdots & - & - \end{array}$$

and hence $V(\rho_2 - \varepsilon) - V(\rho_2 + \varepsilon) = 1$. Proceeding in the same way through all the roots ρ_i , we infer that the number of sign changes decreases by one unit each time a root ρ_i is passed. Hence we have proved that if $\varphi(\lambda)$ is the number of eigenvalues of the matrix B which are larger than λ , then

$$V(\lambda) = \varphi(\lambda) . \tag{6.3.3}$$

The sequence $f_i(\lambda)$ is called a *Sturm sequence*. The described technique makes it possible to compute all eigenvalues in a given interval ("telescope method").

For the third phase, computation of the eigenvectors, we shall follow J. H. Wilkinson in [2]. Let λ_1 be an exact eigenvalue of B . Thus we search for a vector x such that $Bx = \lambda_1 x$. Since this is a homogeneous system in n variables, and since $\det(B - \lambda_1 I) = 0$, we can obtain a nontrivial solution by choosing $n - 1$ equations and determine the components of x (apart from a constant factor); the remaining equation must then be automatically satisfied. In practical work it turns out, even for quite well-behaved matrices, that the result to a large extent depends on *which* equation was excluded from the *beginning*. Essentially, we can say that the serious errors which appear on an unsuitable choice of equation to be excluded depend on numerical compensations; thus round-off errors achieve a dominant influence.

Let us assume that the *i*th equation is excluded, while the others are solved by elimination. The solution (supposed to be exact) satisfies the $n - 1$ equations used for elimination but gives an error δ when inserted into the *i*th equation.

Actually, we have solved the system

$$\begin{cases} \beta_{j-1}x_{j-1} + (\alpha_j - \lambda)x_j + \beta_jx_{j+1} = 0, & j \neq i, \\ \beta_{i-1}x_{i-1} + (\alpha_i - \lambda)x_i + \beta_ix_{i+1} = \delta, & \delta \neq 0. \end{cases}$$

(We had to use an approximation λ instead of the exact eigenvalue λ_1 .) Since constant factors may be omitted, this system can be written in a simpler way:

$$(B - \lambda I)\mathbf{x} = \mathbf{e}_i, \quad (6.3.5)$$

where \mathbf{e}_i is a column vector with the i th component equal to 1 and the others equal to 0. If the eigenvectors of B are $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$, this vector \mathbf{e}_i can be expressed as a linear combination, that is,

$$\mathbf{e}_i = \sum_{j=1}^n c_{ij}\mathbf{v}_j, \quad (6.3.6)$$

and from (6.3.5) we get

$$\mathbf{x} = \sum_{j=1}^n c_{ij}(B - \lambda I)^{-1}\mathbf{v}_j = \sum_{j=1}^n c_{ij} \frac{1}{\lambda_j - \lambda} \mathbf{v}_j. \quad (6.3.7)$$

Now let $\lambda = \lambda_1 + \varepsilon$, and we obtain

$$\mathbf{x} = -\frac{c_{i1}}{\varepsilon} \mathbf{v}_1 + \sum_{j=2}^n c_{ij} \frac{1}{\lambda_j - \lambda_1 - \varepsilon} \mathbf{v}_j. \quad (6.3.8)$$

Under the assumption that $c_{i1} \neq 0$, our solution \mathbf{x} approaches \mathbf{v}_1 as $\varepsilon \rightarrow 0$ (apart from trivial factors). However, it may well happen that c_{i1} is of the same order of magnitude as ε (that is, the vector \mathbf{e}_i is almost orthogonal to \mathbf{v}_1), and under such circumstances it is clear that the vector \mathbf{x} in (6.3.8) cannot be a good approximation of \mathbf{v}_1 . Wilkinson suggests that (6.3.5) be replaced by

$$(B - \lambda I)\mathbf{x} = \mathbf{b}, \quad (6.3.9)$$

where we have the vector \mathbf{b} at our disposal. This system is solved by Gaussian elimination, where it should be observed that the equations are permuted properly to make the pivot element as large as possible. The resulting system is written:

$$\begin{cases} p_{11}x_1 + p_{12}x_2 + p_{13}x_3 & = c_1, \\ p_{22}x_2 + p_{23}x_3 + p_{24}x_4 & = c_2, \\ & \vdots \\ p_{n,n}x_n & = c_n. \end{cases} \quad (6.3.10)$$

As a rule, most of the coefficients p_{13}, p_{24}, \dots are zero. Since the c_i have been obtained from the b_i which we had at our disposal, we could as well choose the constants c_i deliberately. It seems to be a reasonable choice to take all c_i

equal to 1; no eigenvector should then be disregarded. Thus we choose

$$c = \sum_{r=1}^n e_r. \quad (6.3.11)$$

The system is solved, as usual, by back-substitution, and last, the vector \mathbf{x} is normalized. Even on rather pathological matrices, good results have been obtained by Givens' method.

6.4. Householder's method

This method, also, has been designed for real, symmetric matrices. We shall essentially follow the presentation given by Wilkinson [4]. The first step consists of reducing the given matrix A to a band matrix. This is done by orthogonal transformations representing *reflections*. The orthogonal matrices, will be denoted by P_r with the general structure

$$P = I - 2\mathbf{w}\mathbf{w}^T. \quad (6.4.1)$$

Here \mathbf{w} is a column vector such that

$$\mathbf{w}^T\mathbf{w} = 1. \quad (6.4.2)$$

It is evident that P is symmetric. Further, we have

$$P^T P = (I - 2\mathbf{w}\mathbf{w}^T)(I - 2\mathbf{w}\mathbf{w}^T) = I - 4\mathbf{w}\mathbf{w}^T + 4\mathbf{w}\mathbf{w}^T\mathbf{w}\mathbf{w}^T = I;$$

that is, P is also orthogonal.

The matrix P acting as an operator can be given a simple geometric interpretation. Let P operate on a vector \mathbf{x} from the left:

$$P\mathbf{x} = (I - 2\mathbf{w}\mathbf{w}^T)\mathbf{x} = \mathbf{x} - 2(\mathbf{w}^T\mathbf{x})\mathbf{w}.$$

In Fig. 6.4 the line L is perpendicular to the unit vector \mathbf{w} in a plane defined by \mathbf{w} and \mathbf{x} . The distance from the endpoint of \mathbf{x} to L is $|\mathbf{x}| \cos(\mathbf{x}, \mathbf{w}) = \mathbf{w}^T\mathbf{x}$, and the mapping P means a reflection in a plane perpendicular to \mathbf{w} .

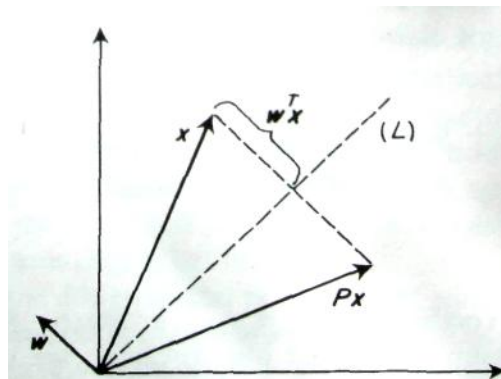


Figure 6.4

Those vectors w which will be used are constructed with the first $(r - 1)$ components zero, or

$$w_r^T = (0, 0, \dots, 0, x_r, x_{r+1}, \dots, x_n).$$

With this choice we form $P_r = I - 2w_r w_r^T$. Further, by (6.4.2) we have

$$x_r^2 + x_{r+1}^2 + \dots + x_n^2 = 1.$$

Now put $A = A_1$ and form successively

$$A_r = P_r A_{r-1} P_r, \quad (6.4.3)$$

$r = 2, 3, \dots, n - 1$. At the first transformation, we get zeros in the positions $(1, 3), (1, 4), \dots, (1, n)$ and in the corresponding places in the first column. The final result will become a band matrix as in Givens' method. The matrix A_{r-1} contains $n - r$ elements in the row, $(r - 1)$, which must be reduced to zero by transformation with P_r ; this gives $n - r$ equations for the $n - r + 1$ elements x_r, x_{r+1}, \dots, x_n , and further we have the condition that the sum of the squares must be 1.

We carry through one step in the computation in an example:

$$A = \begin{pmatrix} a_1 & b_1 & c_1 & d_1 \\ b_1 & b_2 & c_2 & d_2 \\ c_1 & c_2 & c_3 & d_3 \\ d_1 & d_2 & d_3 & d_4 \end{pmatrix},$$

$$w_2^T = (0, x_2, x_3, x_4); \quad x_2^2 + x_3^2 + x_4^2 = 1.$$

The transformation $P_2 A P_2$ must now produce zeros instead of c_1 and d_1 . Obviously, the matrix P_2 has the following form:

$$P_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 - 2x_2^2 & -2x_2x_3 & -2x_2x_4 \\ 0 & -2x_2x_3 & 1 - 2x_3^2 & -2x_3x_4 \\ 0 & -2x_2x_4 & -2x_3x_4 & 1 - 2x_4^2 \end{pmatrix}.$$

Since in the first row of P_2 only the first element is not zero, for example, the $(1, 3)$ -element of $P_2 A P_2$ can become zero only if the corresponding element is zero already in $A P_2$. Putting $p_1 = b_1 x_2 + c_1 x_3 + d_1 x_4$, we find that the first row of $A P_2$ has the following elements:

$$a_1, \quad b_1 - 2p_1 x_2, \quad c_1 - 2p_1 x_3, \quad d_1 - 2p_1 x_4.$$

Now we claim that

$$\begin{cases} c_1 - 2p_1 x_3 = 0, \\ d_1 - 2p_1 x_4 = 0. \end{cases} \quad (6.4.4)$$

Since we are performing an orthogonal transformation, the sum of the squares

of the elements in a row is invariant, and hence

$$a_1^2 + (b_1 - 2p_1x_2)^2 = a_1^2 + b_1^2 + c_1^2 + d_1^2.$$

Putting $S = (b_1^2 + c_1^2 + d_1^2)^{1/2}$, we obtain

$$b_1 - 2p_1x_2 = \pm S. \quad (6.4.5)$$

Multiplying (6.4.5) by x_2 and (6.4.4) by x_3 and x_4 , we get

$$b_1x_2 + c_1x_3 + d_1x_4 - 2p_1(x_2^2 + x_3^2 + x_4^2) = \pm Sx_2.$$

The sum of the first three terms is p_1 and further $x_2^2 + x_3^2 + x_4^2 = 1$. Hence

$$p_1 = \mp Sx_2. \quad (6.4.6)$$

Inserting this into (6.4.5), we find that $x_2^2 = \frac{1}{2}(1 \mp b_1/S)$, and from (6.4.4), $x_3 = \mp \tilde{c}_1/2Sx_2$ and $x_4 = \mp d_1/2Sx_2$.

In the general case, two square roots have to be evaluated, one for S and one for x_2 . Since we have x_2 in the denominator, we obtain the best accuracy if x_2 is large. This is accomplished by choosing a suitable sign for the square-root extraction for S . Thus the quantities ought to be defined as follows:

$$x_2^2 = \frac{1}{2} \left(1 + \frac{b_1 \cdot \text{sign } b_1}{S} \right). \quad (6.4.7)$$

The sign for this square root is irrelevant and we choose plus. Hence we obtain for x_3 and x_4 :

$$x_3 = \frac{c_1 \text{ sign } b_1}{2Sx_2}; \quad x_4 = \frac{d_1 \text{ sign } b_1}{2Sx_2}. \quad (6.4.8)$$

The end result is a band matrix whose eigenvalues and eigenvectors are computed exactly as in Givens' method. In order to get an eigenvector \mathbf{v} of \mathbf{A} , an eigenvector \mathbf{x} of the band matrix has to be multiplied by the matrix $\mathbf{P}_2, \mathbf{P}_3, \dots, \mathbf{P}_{n-1}$; this should be done by iteration:

$$\left\{ \begin{array}{l} \mathbf{x}_{n-1} = \mathbf{P}_{n-1}\mathbf{x}, \\ \mathbf{x}_{n-2} = \mathbf{P}_{n-2}\mathbf{x}_{n-1}, \\ \vdots \\ \mathbf{v} \equiv \mathbf{x}_2 = \mathbf{P}_2\mathbf{x}_3. \end{array} \right. \quad (6.4.9)$$

6.5. Lanczos' method

The reduction of real symmetric matrices to tridiagonal form can be accomplished through methods devised by Givens and Householder. For arbitrary matrices a similar reduction can be performed by a technique suggested by Lanczos. In this method two systems of vectors are constructed, $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ and $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n$, which are *biorthogonal*; that is, for $j \neq k$, we have $\mathbf{x}_j^H \mathbf{y}_k = 0$.

The initial vectors \mathbf{x}_1 and \mathbf{y}_1 can be chosen arbitrarily though in such a way that $\mathbf{x}_1^H \mathbf{y}_1 \neq 0$. The new vectors are formed according to the rules

$$\begin{cases} \mathbf{x}_{k+1} = A\mathbf{x}_k - \sum_{j=1}^k a_{jk} \mathbf{x}_j, \\ \mathbf{y}_{k+1} = A^H \mathbf{y}_k - \sum_{j=1}^k b_{jk} \mathbf{y}_j. \end{cases}$$

The coefficients are determined from the biorthogonality condition, and for $j = 1, 2, \dots, k$, we form:

$$0 = \mathbf{y}_j^H \mathbf{x}_{k+1} = \mathbf{y}_j^H A \mathbf{x}_k - \sum_{r=1}^k a_{rk} \mathbf{y}_j^H \mathbf{x}_r = \mathbf{y}_j^H A \mathbf{x}_k - a_{jk} \mathbf{y}_j^H \mathbf{x}_j.$$

If $\mathbf{y}_j^H \mathbf{x}_j \neq 0$, we get

$$a_{jk} = \frac{\mathbf{y}_j^H A \mathbf{x}_k}{\mathbf{y}_j^H \mathbf{x}_j}.$$

Analogously

$$b_{jk} = \frac{\mathbf{x}_j^H A^H \mathbf{y}_k}{\mathbf{x}_j^H \mathbf{y}_j}.$$

Let us now consider the numerator in the expression for a_{jk} when $j \leq k - 2$:

$$\mathbf{y}_j^H A \mathbf{x}_k = (\mathbf{x}_k^H A^H \mathbf{y}_j)^* = \left\{ \mathbf{x}_k^H \left(\mathbf{y}_{j+1} + \sum_{r=1}^j b_{rj} \mathbf{y}_r \right) \right\}^* = 0,$$

because of the biorthogonality. Hence we have $a_{jk} = 0$ for $j \leq k - 2$, and similarly we also have $b_{jk} = 0$ under the same condition. In this way the following simpler formulas are obtained:

$$\begin{cases} \mathbf{x}_{k+1} = A\mathbf{x}_k - (a_{k-1,k} \mathbf{x}_{k-1} + a_{kk} \mathbf{x}_k), \\ \mathbf{y}_{k+1} = A^H \mathbf{y}_k - (b_{k-1,k} \mathbf{y}_{k-1} + b_{kk} \mathbf{y}_k). \end{cases}$$

If the vectors $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ are considered as columns in a matrix \mathbf{X} and if further a tridiagonal matrix \mathbf{J} is formed from the coefficients $a_{k-1,k}$ and a_{kk} with one's in the remaining diagonal:

$$\mathbf{J} = \begin{pmatrix} a_{11} & a_{12} & 0 & 0 & \dots & 0 & 0 \\ 1 & a_{22} & a_{23} & 0 & \dots & 0 & 0 \\ 0 & 1 & a_{33} & a_{34} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & a_{n-1,n-1} & a_{n-1,n} \\ 0 & 0 & 0 & \dots & 0 & 1 & a_{nn} \end{pmatrix};$$

then we can simply write $A\mathbf{X} = \mathbf{X}\mathbf{J}$, and provided the vectors $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ are linearly independent

$$\mathbf{J} = \mathbf{X}^{-1} A \mathbf{X}.$$

If similar matrices are formed from the vectors $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n$ and from the coefficients $b_{k-1,k}, b_{kk}$, we get

$$\mathbf{K} = \mathbf{Y}^{-1}\mathbf{A}^H\mathbf{Y}.$$

Certain complications may arise, for example, that some \mathbf{x} - or \mathbf{y} -vector may become zero, but it can also happen that $\mathbf{x}_j^H\mathbf{y}_j = 0$ even if $\mathbf{x}_j \neq \mathbf{0}$ and $\mathbf{y}_j \neq \mathbf{0}$. The simplest way out is to choose other initial vectors even if it is sometimes possible to get around the difficulties by modifying the formulas themselves.

Obviously, Lanczos' method can be used also with real symmetric or Hermitian matrices. Then one chooses just *one* sequence of vectors which must form an orthogonal system. For closer details, particularly concerning the determination of the eigenvectors, Lanczos' paper [1] should be consulted; a detailed discussion of the degenerate cases is given by Causey and Gregory [9].

Here we also mention still one method for tridiagonalization of arbitrary real matrices, first given by La Budde. Space limitations prevent us from a closer discussion, and instead we refer to the original paper [11].

6.6. Other methods

Among other interesting methods we mention the LR -method. Starting from a matrix $\mathbf{A} = \mathbf{A}_1$, we split it into two triangular matrices $\mathbf{A}_1 = \mathbf{L}_1\mathbf{R}_1$ with $\tilde{l}_{ii} = 1$, and then we form $\mathbf{A}_2 = \mathbf{R}_1\mathbf{L}_1$. Since $\mathbf{A}_2 = \mathbf{R}_1\mathbf{A}_1\mathbf{R}_1^{-1}$, the new matrix \mathbf{A}_2 has the same eigenvalues as \mathbf{A}_1 . Then we treat \mathbf{A}_2 in the same way as \mathbf{A}_1 , and so on, obtaining a sequence of matrices $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3, \dots$, which in general converges toward an upper triangular matrix. If the eigenvalues are real, they will appear in the main diagonal. Even the case in which complex eigenvalues are present can be treated without serious complications. Closer details are given in [5], where the method is described by its inventor, H. Rutishauser. Here we shall also examine the more general eigenvalue problem,

$$\det(\mathbf{A} - \lambda\mathbf{B}) = 0,$$

where \mathbf{A} and \mathbf{B} are symmetric and, further, \mathbf{B} is positive definite. Then we can split \mathbf{B} according to $\mathbf{B} = \mathbf{L}\mathbf{L}^T$, where \mathbf{L} is a lower triangular matrix. Hence

$$\mathbf{A} - \lambda\mathbf{B} = \mathbf{A} - \lambda\mathbf{L}\mathbf{L}^T = \mathbf{L}(\mathbf{L}^{-1}\mathbf{A}(\mathbf{L}^T)^{-1} - \lambda\mathbf{I})\mathbf{L}^T,$$

and $\det(\mathbf{A} - \lambda\mathbf{B}) = (\det \mathbf{L})^2 \cdot \det(\mathbf{C} - \lambda\mathbf{I})$, where $\mathbf{C} = \mathbf{L}^{-1}\mathbf{A}(\mathbf{L}^{-1})^T$. Since $\mathbf{C}^T = \mathbf{C}$, the problem has been reduced to the usual type treated before.

6.7. Complex matrices

For computing eigenvalues and eigenvectors of arbitrary complex matrices (also, real nonsymmetric matrices fall naturally into this group), we shall first discuss a triangularization method suggested by Lotkin [6] and Greenstadt [7].

The method depends on the lemma by Schur stating that for each square matrix A there exists a unitary matrix U such that $U^{-1}AU = T$, where T is a (lower or upper) triangular matrix (see Section 3.7). In practical computation one tries to find U as a product of essentially two-dimensional unitary matrices, using a procedure similar to that described for Hermitian matrices in Section 6.2. It is possible to give examples for which the method does not converge (the sum of the squares of the absolute values of the subdiagonal elements is not monotonically decreasing, cf. [13]), but in practice convergence is obtained in many cases. We start by examining the two-dimensional case and put

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}; \quad U = \begin{pmatrix} p & -q^* \\ q & p \end{pmatrix} \quad (p \text{ real}). \quad (6.7.1)$$

From $U^H U = I$, we get $p^2 + |q|^2 = 1$. Further, we suppose that $A' = U^{-1}AU$, where

$$A' = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix},$$

and obtain

$$\begin{cases} a' = d + (a - d)p^2 + p(bq + cq^*), \\ b' = bp^2 - cq^{*2} + (d - a)pq^*, \\ c' = cp^2 + (d - a)pq - bq^2, \\ d' = a + (d - a)p^2 - p(bq + cq^*). \end{cases} \quad (6.7.2)$$

Clearly we have $a' + d' = a + d$. Claiming $c' = 0$, we find with $q = \alpha p$,

$$\alpha^2 - \frac{d - a}{b} \alpha = \frac{c}{b},$$

and

$$\alpha = \frac{1}{2b} (d - a \pm \sqrt{(d - a)^2 + 4bc}). \quad (6.7.3)$$

Here we conveniently choose the sign that makes $|\alpha|$ as small as possible; with $p = \cos \theta$ and $q = \sin \theta \cdot e^{i\varphi}$, we get $|\alpha| = \tan \theta$. Hence α is obtained directly from the elements a , b , c , and d . Normally, we must take the square root of a complex number, and this can be done by the formula

$$\sqrt{A + iB} = \pm(1/\sqrt{2})(\sqrt{C + A} + i \text{sign } B\sqrt{C - A}),$$

where $C = \sqrt{A^2 + B^2}$. When α has been determined, we get p and q from

$$\begin{cases} p = (1 + |\alpha|^2)^{-1/2}, \\ q = \alpha p. \end{cases} \quad (6.7.4)$$

Now we pass to the main problem and assume that A is an arbitrary complex matrix (n, n) . We choose that element *below* the main diagonal which is largest

in absolute value, and perform an essentially two-dimensional unitary transformation which makes this element zero. This procedure is repeated until $\sum_{i>k} |a_{ik}|^2$ is less than a given tolerance. Denoting the triangular matrix by T , we have $U^{-1}AU = T$ with $U = U_1 U_2 \dots U_N$, where are the U_1, U_2, \dots, U_N individual essentially two-dimensional unitary matrices. Clearly, the eigenvalues are $t_{rr}, r = 1, 2, \dots, n$.

In order to compute the eigenvectors, we start with the triangular matrix, and we shall restrict ourselves to the case when all eigenvalues are different. We see directly that the vector whose first component is 1 and whose other components are 0 is an eigenvector belonging to the first eigenvalue t_{11} . Next we see that we can determine such a value y_{12} that the vector with first component y_{12} and second component 1 becomes an eigenvector. The condition is $t_{11}y_{12} + t_{12} = t_{22}y_{12}$, from which we can determine y_{12} (we suppose that $t_{11} \neq t_{22}$). We proceed in the same way and collect all eigenvectors to a triangular matrix Y , and further we form a diagonal matrix A with the diagonal elements $t_{rr}, r = 1, 2, \dots, n$. Then we have

$$T = \begin{pmatrix} t_{11} & t_{12} & t_{13} & \dots & t_{1n} \\ 0 & t_{22} & t_{23} & \dots & t_{2n} \\ 0 & 0 & t_{33} & \dots & t_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & t_{nn} \end{pmatrix}; \quad Y = \begin{pmatrix} 1 & y_{12} & y_{13} & \dots & y_{1n} \\ 0 & 1 & y_{23} & \dots & y_{2n} \\ 0 & 0 & 1 & \dots & y_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix};$$

$$A = \begin{pmatrix} t_{11} & 0 & \dots & 0 \\ 0 & t_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & t_{nn} \end{pmatrix}, \quad (6.7.5)$$

and obviously $TY = YA$. The quantities $y_{ik}, k > i$, are computed recursively from the relation

$$y_{ik} = \sum_{r=i+1}^k \frac{t_{ir}y_{rk}}{t_{kk} - t_{ii}}, \quad (6.7.6)$$

valid for $i = 1, 2, \dots, (k-1)$. First we put $y_{kk} = 1$ and then we use (6.7.6) for $i = k-1, k-2, \dots, 1$, and in this way the eigenvector y_k belonging to the eigenvalue t_{kk} can be determined. Last, we obtain the eigenvectors x_k of the original matrix A from

$$x_k = Uy_k. \quad (6.7.7)$$

When the method is used on a computer, we must reserve two memory places for each element (even if A is real). Only in special cases are all results real. The method described here depends on annihilation of a subdiagonal element which, however, does not guarantee that $\tau^2 = \sum_{i>k} |a_{ik}|^2$ decreases. An alternative technique can be constructed aiming at minimization of τ^2 by choosing θ and φ conveniently. The equations become fairly complicated but can be solved numerically and, as a rule, the minimum point need not be established

to a very high degree of accuracy. The value of the method is still difficult to estimate as it has not yet been tried sufficiently in practice.

6.8. Hyman's method

A matrix is said to be of upper Hessenberg form if $a_{ik} = 0$ for $i - k > 1$ and of lower Hessenberg form if $a_{ik} = 0$ for $k - i > 1$. In the following we shall choose to work with the upper Hessenberg form. An arbitrary complex matrix can quite easily be reduced to Hessenberg form which will now be demonstrated. Essentially the reduction goes as in Givens' method demanding $(n - 1)(n - 2)/2$ steps. In the general case the transformations are unitary, but for real matrices we can use real (orthogonal) transformations which is a great advantage.

Starting from an arbitrary complex matrix $C = (c_{ik})$ we perform a two-dimensional rotation in the (i, k) -plane under the condition $c'_{k,i-1} = 0$ ($i = 2, 3, \dots, n - 1; k = i + 1, i + 2, \dots, n$). The first rotation occurs in the $(2, 3)$ -plane with the condition $c'_{31} = 0$, next in the $(2, 4)$ -plane with $c'_{41} = 0$, and so on. In this way all elements in the first column except c'_{11} and c'_{21} are annihilated. After that we rotate in the $(3, 4)$ -plane with $c'_{42} = 0$, and so on. Introduce the notations

$$\begin{cases} c_{i,i-1} = a + bi, \\ c_{k,i-1} = c + di, \end{cases} \quad \text{and} \quad U = \begin{pmatrix} p & -q^* \\ q & p \end{pmatrix},$$

with $p = \cos \theta$, $q = \sin \theta e^{i\varphi}$ making U unitary. Then we get

$$c'_{k,i-1} = -qc_{i,i-1} + pc_{k,i-1} = 0$$

and splitting into real and imaginary parts:

$$\begin{cases} c \cos \theta = a \sin \theta \cos \varphi - b \sin \theta \sin \varphi, \\ d \cos \theta = a \sin \theta \sin \varphi - b \sin \theta \cos \varphi. \end{cases}$$

Squaring and adding we get $(c^2 + d^2) \cos^2 \theta = (a^2 + b^2) \sin^2 \theta$, and a trivial elimination also gives the angle φ :

$$\begin{cases} \tan \theta = \sqrt{(c^2 + d^2)/(a^2 + b^2)}, \\ \tan \varphi = (ad - bc)/(ac + bd). \end{cases}$$

In the real case we have $b = d = 0$ giving $\varphi = 0$ and $\tan \theta = c/a$.

We shall also briefly show how the reduction can be made by use of reflections following Householder (also cf. [13]). Putting $A = A_0$, we shall describe one step in the reduction leading from A_{r-1} to A_r , where

$$A_{r-1} = \left(\begin{array}{c|c} \mathbf{H}_{r-1} & \mathbf{C}_{r-1} \\ \hline \mathbf{0} & \mathbf{B}_{r-1} \end{array} \right) \begin{matrix} r \\ n-r \end{matrix}$$

and

$$H_{r-1} = \begin{pmatrix} * & * & * & * \cdots * & * \\ * & * & * & * \cdots * & * \\ & * & * & * \cdots * & * \\ & & * & * \cdots * & * \\ & & & 0 & * \\ & & & & * \end{pmatrix}$$

is of upper Hessenberg form. The matrix A_r will then be produced through $A_r = P_r A_{r-1} P_r$ leaving H_{r-1} as well as the null-matrix of dimension $(n-r) \times (r-1)$ unchanged while the vector b_{r-1} must be annihilated except for the first component. In this way a new Hessenberg matrix H_r of dimension $(r+1) \times (r+1)$ is formed by moving one new row and one new column to H_{r-1} . Now we choose

$$P_r = \left(\begin{array}{c|c} I & 0 \\ \hline 0 & Q_r \end{array} \right)_{n-r}$$

with $Q_r = I - 2w_r w_r^H$ and $w_r^H w_r = 1$, w_r being a column vector with $(n-r)$ elements. A simple computation gives

$$P_r A_{r-1} P_r = \left(\begin{array}{c|c|c} H_{r-1} & C_{r-1} Q_r & \\ \hline 0 & a_{r-1} & Q_r B_{r-1} Q_r \end{array} \right) \quad \text{with} \quad a_{r-1} = \begin{pmatrix} \alpha_r \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Hence $Q_r b_{r-1} = a_{r-1}$ and $a_{r-1}^H a_{r-1} = b_{r-1}^H Q_r^H Q_r b_{r-1} = b_{r-1}^H b_{r-1}$, that is, $|\alpha_r| = |b_{r-1}|$ (we suppose Euclidean vector norm). Further $b_{r-1} = Q_r^H a_{r-1}$ and $e_1^T b_{r-1} = e_1^T Q_r^H a_{r-1} = (1 - |w_1|^2) \alpha_r$, and since $1 - |w_1|^2$ is real, $\arg \alpha_r = \arg (e_1^T b_{r-1})$. Here e_1 is a vector with the first component $= 1$ and all other components $= 0$. Thus the argument of α_r is equal to the argument of the top element of the vector b_{r-1} . Finally, since

$$Q_r b_{r-1} = (I - 2w_r w_r^H) b_{r-1} = b_{r-1} - (2w_r^H b_{r-1}) w_r = a_{r-1},$$

we get

$$w_r = \frac{b_{r-1} - a_{r-1}}{|b_{r-1} - a_{r-1}|},$$

and A_r is completely determined. If this procedure is repeated we finally reach the matrix A_{n-2} which is of upper Hessenberg form.

After having produced a Hessenberg matrix $A = U^{-1} C U$ with the same eigenvalues as C we now turn to the computation of these. Let x be a vector with unknown components and λ an arbitrary (complex) number. Then form the

following linear system of equations:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = \lambda x_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = \lambda x_2, \\ \vdots \\ a_{n-1,n-2}x_{n-2} + a_{n-1,n-1}x_{n-1} + a_{n-1,n}x_n = \lambda x_{n-1}, \\ a_{n,n-1}x_{n-1} + a_{nn}x_n = \lambda x_n. \end{cases} \quad (6.8.2)$$

We suppose that all elements $a_{i,i-1} \neq 0$, $i = 2, 3, \dots, n$, and as the system is homogeneous we can choose, for example, $x_n = 1$. Then we can solve x_{n-1} as a first-degree polynomial from the last equation. From the next to last equation, x_{n-2} is obtained as a second-degree polynomial in λ , and so on; and finally x_1 is determined from the second equation as a polynomial of degree $n-1$ in λ . If all these values of x_1, x_2, \dots, x_n are inserted into the expression $(\lambda - a_{11})x_1 - a_{12}x_2 - \dots - a_{1n}x_n$, we get as result the characteristic polynomial $f(\lambda) = \det(\lambda I - A)$ apart from a constant factor. It is now possible to compute the eigenvalues from these results by interpolation, first linear and then quadratic.

The method described here is such that the values obtained are *exact* eigenvalues of another matrix differing only slightly from A ("reverse error computation," cf. Section 4.4. and Wilkinson, [18], p. 147). It is quite possible that the errors in the eigenvalues may become large; this means that the eigenvalue problem *per se* is *ill*-conditioned and we can expect numerical difficulties irrespective of the method chosen.

The interpolation can be improved by computing the first and possibly also the second derivative of the characteristic polynomial by iteration. The following formulas are used:

$$\begin{aligned} x_{i-1} &= \frac{(\lambda - a_{ii})x_i - \sum_{k=i+1}^n a_{ik}x_k}{a_{i,i-1}}; & x_n &= 1, & a_{1,0} &= 1; \\ y_{i-1} &= \frac{(\lambda - a_{ii})y_i + x_i - \sum_{k=i+1}^n a_{ik}y_k}{a_{i,i-1}}; & y_n &= 0; \\ z_{i-1} &= \frac{(\lambda - a_{ii})z_i + 2y_i - \sum_{k=i+1}^n a_{ik}z_k}{a_{i,i-1}}; & z_n &= 0. \end{aligned}$$

Denoting the characteristic polynomial by $f(\lambda) (= \lambda^n + \dots)$ and putting

$$p_{00} = \left\{ \prod_{i=2}^n a_{i,i-1} \right\}^{-1},$$

we have

$$f(\lambda) = \frac{x_0}{p_{00}}; \quad f'(\lambda) = \frac{y_0}{p_{00}}; \quad f''(\lambda) = \frac{z_0}{p_{00}}.$$

Then Newton-Raphson's method can be used (possibly a variant independent

of multiplicity):

$$\lambda_{n+1} = \lambda_n - \frac{x_0}{y_0} \quad \text{or} \quad \lambda_{n+1} = \lambda_n - \frac{x_0 y_0}{y_0^2 - x_0 z_0}.$$

Example

$$A = \begin{pmatrix} 6 & 3 & -4 & 2 \\ 2 & 1 & 5 & -3 \\ 0 & 3 & 7 & 1 \\ 0 & 0 & 2 & 5 \end{pmatrix}.$$

Starting with $\lambda_0 = 5$, we compute x_i and y_i :

i	x_i	y_i
4	1	0
3	0	$\frac{1}{2}$
2	$-\frac{1}{3}$	$-\frac{1}{3}$
1	$\frac{5}{6}$	$-\frac{25}{12}$
0	$-\frac{11}{6}$	$\frac{71}{12}$

Hence $\lambda_1 = 5 + \frac{2^2}{7^2} \simeq 5.3$ (correct value 5.374). Let us now once again write down expressions as polynomials in λ for the set x_1, x_2, \dots, x_n which is also extended by an extra variable associated with the first equation of the system:

$$\begin{cases} x_0 = p_{00}\lambda^n + p_{01}\lambda^{n-1} + \dots + p_{0n}, \\ x_1 = p_{11}\lambda^{n-1} + p_{12}\lambda^{n-2} + \dots + p_{1n}, \\ \vdots \\ x_i = p_{ii}\lambda^{n-i} + p_{i,i+1}\lambda^{n-i-1} + \dots + p_{in}, \\ \vdots \\ x_{n-1} = p_{n-1,n-1}\lambda + p_{n-1,n}, \\ x_n = p_{nn} (= 1). \end{cases}$$

Using $Ax = \lambda x$, we get

$$\begin{aligned} x_{i-1} &= \frac{(\lambda - a_{ii})x_i - a_{i,i+1}x_{i+1} - \dots - a_{in}x_n}{a_{i,i-1}} \\ &= \frac{(\lambda - a_{ii}) \sum_{k=i}^n p_{ik}\lambda^{n-k} - a_{i,i+1} \sum_{k=i+1}^n p_{i+1,k}\lambda^{n-k} - \dots - a_{in}p_{nn}}{a_{i,i-1}}. \end{aligned}$$

Comparing powers λ^{n-k} on both sides, we find

$$a_{i,i-1}p_{i-1,k} = p_{i,k+1} - \sum_{r=i}^k a_{ir}p_{rk}$$

(6.8.4)

$$(i = 1, 2, \dots, n; k = i - 1, i, \dots, n; a_{1,0} = 1, p_{i,n+1} = 0).$$

This relation can be written in matrix form; for $n = 4$ we get

$$\begin{pmatrix} 1 & a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{21} & a_{22} & a_{23} & a_{24} \\ 0 & 0 & a_{32} & a_{33} & a_{34} \\ 0 & 0 & 0 & a_{43} & a_{44} \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} p_{00} & p_{01} & p_{02} & p_{03} & p_{04} \\ 0 & p_{11} & p_{12} & p_{13} & p_{14} \\ 0 & 0 & p_{22} & p_{23} & p_{24} \\ 0 & 0 & 0 & p_{33} & p_{34} \\ 0 & 0 & 0 & 0 & p_{44} \end{pmatrix} = \begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} & 0 \\ 0 & p_{22} & p_{23} & p_{24} & 0 \\ 0 & 0 & p_{33} & p_{34} & 0 \\ 0 & 0 & 0 & p_{44} & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \quad (6.8.5)$$

Taking determinants on both sides we find in particular $p_{00} = (a_{21}a_{32}a_{43})^{-1}$, which can easily be generalized to arbitrary n . The characteristic equation is now

$$p_{00}\lambda^n + p_{01}\lambda^{n-1} + \dots + p_{0n} = 0.$$

This method of generating the characteristic equation and then the eigenvalues by some standard technique, as a rule is unsuitable for stability reasons except possibly if the computation of coefficients can be performed within the rational field.

Computation of the eigenvectors from (6.8.2) cannot be recommended since instability will occur. However, it is possible to use an idea by Wielandt, in principle inverse interpolation. Let λ_0 be an approximate eigenvalue ($\lambda_0 \simeq \lambda_1$, where λ_1 is an exact eigenvalue) and \mathbf{u}_0 an approximate eigenvector. Further we suppose that the exact eigenvectors are $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$. Successive iterations are now computed from

$$(\mathbf{A} - \lambda_0 \mathbf{I})\mathbf{u}_{i+1} = s_i \mathbf{u}_i,$$

where s_i is a suitable scale factor compensating for the fact that $\mathbf{A} - \lambda_0 \mathbf{I}$ is almost singular which will cause no trouble otherwise. Now suppose, for example, $\mathbf{u}_0 = \sum \alpha_r \mathbf{v}_r$ which leads to

$$(\mathbf{A} - \lambda_0 \mathbf{I})\mathbf{u}_1 = s_0 \sum \alpha_r \mathbf{v}_r,$$

that is, $\mathbf{u}_1 = s_0 \sum \alpha_r (\mathbf{A} - \lambda_0 \mathbf{I})^{-1} \mathbf{v}_r = s_0 \sum \alpha_r (\lambda_r - \lambda_0)^{-1} \mathbf{v}_r \simeq s_0 \alpha_1 (\lambda_1 - \lambda_0)^{-1} \mathbf{v}_1$ and we have the same effect as when the power method is applied. The solution of the system $(\mathbf{A} - \lambda_0 \mathbf{I})\mathbf{u}_{i+1} = s_i \mathbf{u}_i$ is performed by standard means, for example, following Gauss or Crout.

If the eigenvector problem is ill-conditioned in itself, this technique, of course, will not help. Consider, for example,

$$\mathbf{A} = \begin{pmatrix} 1 & \varepsilon_1 \\ \varepsilon_2 & 1 \end{pmatrix}$$

This relation can be written in matrix form; for $n = 4$ we get

$$\begin{pmatrix} 1 & a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{21} & a_{22} & a_{23} & a_{24} \\ 0 & 0 & a_{32} & a_{33} & a_{34} \\ 0 & 0 & 0 & a_{43} & a_{44} \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} p_{00} & p_{01} & p_{02} & p_{03} & p_{04} \\ 0 & p_{11} & p_{12} & p_{13} & p_{14} \\ 0 & 0 & p_{22} & p_{23} & p_{24} \\ 0 & 0 & 0 & p_{33} & p_{34} \\ 0 & 0 & 0 & 0 & p_{44} \end{pmatrix} = \begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} & 0 \\ 0 & p_{22} & p_{23} & p_{24} & 0 \\ 0 & 0 & p_{33} & p_{34} & 0 \\ 0 & 0 & 0 & p_{44} & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \quad (6.8.5)$$

Taking determinants on both sides we find in particular $p_{00} = (a_{21}a_{32}a_{43})^{-1}$, which can easily be generalized to arbitrary n . The characteristic equation is now

$$p_{00}\lambda^n + p_{01}\lambda^{n-1} + \dots + p_{0n} = 0.$$

This method of generating the characteristic equation and then the eigenvalues by some standard technique, as a rule is unsuitable for stability reasons except possibly if the computation of coefficients can be performed within the rational field.

Computation of the eigenvectors from (6.8.2) cannot be recommended since instability will occur. However, it is possible to use an idea by Wielandt, in principle inverse interpolation. Let λ_0 be an approximate eigenvalue ($\lambda_0 \simeq \lambda_1$, where λ_1 is an exact eigenvalue) and \mathbf{u}_0 an approximate eigenvector. Further we suppose that the exact eigenvectors are $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$. Successive iterations are now computed from

$$(\mathbf{A} - \lambda_0 \mathbf{I})\mathbf{u}_{i+1} = s_i \mathbf{u}_i,$$

where s_i is a suitable scale factor compensating for the fact that $\mathbf{A} - \lambda_0 \mathbf{I}$ is almost singular which will cause no trouble otherwise. Now suppose, for example, $\mathbf{u}_0 = \sum \alpha_r \mathbf{v}_r$ which leads to

$$(\mathbf{A} - \lambda_0 \mathbf{I})\mathbf{u}_1 = s_0 \sum \alpha_r \mathbf{v}_r,$$

that is, $\mathbf{u}_1 = s_0 \sum \alpha_r (\mathbf{A} - \lambda_0 \mathbf{I})^{-1} \mathbf{v}_r = s_0 \sum \alpha_r (\lambda_r - \lambda_0)^{-1} \mathbf{v}_r \simeq s_0 \alpha_1 (\lambda_1 - \lambda_0)^{-1} \mathbf{v}_1$ and we have the same effect as when the power method is applied. The solution of the system $(\mathbf{A} - \lambda_0 \mathbf{I})\mathbf{u}_{i+1} = s_i \mathbf{u}_i$ is performed by standard means, for example, following Gauss or Crout.

If the eigenvector problem is ill-conditioned in itself, this technique, of course, will not help. Consider, for example,

$$\mathbf{A} = \begin{pmatrix} 1 & \varepsilon_1 \\ \varepsilon_2 & 1 \end{pmatrix}$$

and the vector q_k is also determined. Obviously, uniqueness is secured apart from factors $e^{i\theta_k}$ in the diagonal elements of \mathbf{R} , and hence the condition $r_{kk} > 0$ will determine \mathbf{Q} and \mathbf{R} completely.

Now we start from $\mathbf{A} = \mathbf{A}_1$ and form sequences of matrices \mathbf{A}_s , \mathbf{Q}_s , and \mathbf{R}_s by use of the following algorithm:

$$\mathbf{A}_s = \mathbf{Q}_s \mathbf{R}_s; \quad \mathbf{A}_{s+1} = \mathbf{Q}_s^H \mathbf{A}_s \mathbf{Q}_s = \mathbf{Q}_s^H \mathbf{Q}_s \mathbf{R}_s \mathbf{Q}_s = \mathbf{R}_s \mathbf{Q}_s.$$

This means that first \mathbf{A}_s is partitioned in a product of a unitary and an upper triangular matrix, and then \mathbf{A}_{s+1} is computed as $\text{It } \mathbf{R}_s \mathbf{Q}_s$. also means that \mathbf{A}_{s+1} is formed from \mathbf{A}_s through a similarity transformation with a unitary matrix. Next, we put

$$\mathbf{P}_s = \mathbf{Q}_1 \mathbf{Q}_2 \cdots \mathbf{Q}_s \quad \text{and} \quad \mathbf{R}_s \mathbf{R}_{s-1} \cdots \mathbf{R}_1 = \mathbf{U}_s,$$

and so we obtain

$$\mathbf{A}_{s+1} = \mathbf{Q}_s^{-1} \mathbf{A}_s \mathbf{Q}_s = \mathbf{Q}_s^{-1} (\mathbf{Q}_{s-1}^{-1} \mathbf{A}_{s-1} \mathbf{Q}_{s-1}) \mathbf{Q}_s = \cdots = \mathbf{P}_s^{-1} \mathbf{A}_1 \mathbf{P}_s.$$

Then we form

$$\begin{aligned} \mathbf{P}_s \mathbf{U}_s &= \mathbf{Q}_1 \mathbf{Q}_2 \cdots \mathbf{Q}_s \mathbf{R}_s \mathbf{R}_{s-1} \cdots \mathbf{R}_1 = \mathbf{Q}_1 \mathbf{Q}_2 \cdots \mathbf{Q}_{s-1} \mathbf{A}_s \mathbf{R}_{s-1} \cdots \mathbf{R}_1 \\ &= \mathbf{P}_{s-1} \mathbf{A}_s \mathbf{U}_{s-1}. \end{aligned}$$

But $\mathbf{A}_s = \mathbf{P}_{s-1}^{-1} \mathbf{A}_1 \mathbf{P}_{s-1}$ and consequently $\mathbf{P}_{s-1} \mathbf{A}_s = \mathbf{A}_1 \mathbf{P}_{s-1}$ which gives

$$\mathbf{P}_s \mathbf{U}_s = \mathbf{A}_1 \mathbf{P}_{s-1} \mathbf{U}_{s-1} = \cdots = \mathbf{A}_1^{s-1} \mathbf{P}_1 \mathbf{U}_1 = \mathbf{A}_1^{s-1} \mathbf{Q}_1 \mathbf{R}_1 = \mathbf{A}_1^s.$$

Here \mathbf{P}_s is unitary and \mathbf{U}_s upper triangular, and in principle they could be computed from \mathbf{A}_1^s by partition in a product of a unitary and an upper triangular matrix. In this way we would also obtain \mathbf{A}_{s+1} through a similarity transformation:

$$\mathbf{A}_{s+1} = \mathbf{P}_s^{-1} \mathbf{A}_1 \mathbf{P}_s.$$

We now assert that the matrix \mathbf{A}_{s+1} for increasing s more and more will approach an upper triangular matrix. We do not give a complete proof but restrict ourselves to the main points; further details can be found in [12]. The following steps are needed for the proof:

1. $\mathbf{A} = \mathbf{A}_1$ is written $\mathbf{A}_1 = \mathbf{X} \mathbf{D} \mathbf{X}^{-1} = \mathbf{X} \mathbf{D} \mathbf{Y}$ which is always possible if all eigenvalues are different (this restriction can be removed afterward). Further we assume $d_{ii} = \lambda_i$ with $|\lambda_1| > |\lambda_2| > \cdots$
2. Then we also have $\mathbf{P}_s \mathbf{U}_s = \mathbf{A}_1^s = \mathbf{X} \mathbf{D}^s \mathbf{Y}$.
3. \mathbf{X} is partitioned as $\mathbf{X} = \mathbf{Q}_X \mathbf{R}_X$ and \mathbf{Y} as $\mathbf{Y} = \mathbf{L}_Y \mathbf{R}_Y$ where \mathbf{Q}_X is unitary, \mathbf{R}_X and \mathbf{R}_Y upper triangular, and \mathbf{L}_Y lower triangular with ones in the main diagonal (both partitions are unique). For the latter partition a permutation might be necessary.
4. Then we get $\mathbf{P}_s \mathbf{U}_s = \mathbf{Q}_X \mathbf{R}_X \mathbf{D}^s \mathbf{L}_Y \mathbf{D}^{-s} \mathbf{D}^s \mathbf{R}_Y$. The decisive point is that $\mathbf{D}^s \mathbf{L}_Y \mathbf{D}^{-s}$, as is easily shown, in subdiagonal elements will contain quotients

$(\lambda_i/\lambda_1)^s$ so that this matrix actually will approach the identity matrix. If so, we are left with $P_s U_s \simeq Q_X R_X D^s R_Y$.

5. But P_s and Q_X are unitary while U_s and $R_X D^s R_Y$ are upper triangular. Since the partition is unique, we can draw the conclusion that

$$\lim_{s \rightarrow \infty} P_s = Q_X \quad \text{and} \quad U_s \simeq R_X D^s R_Y.$$

6. $A_{s+1} = P_s^H A_1 P_s = P_s^H X D X^{-1} P_s \rightarrow Q_X^H X D X^{-1} Q_X \rightarrow R_X D R_X^{-1}$ (since $X = Q_X R_X$ and $R_X = Q_X^H X$). The matrix $R_X D R_X^{-1}$ is an upper triangular matrix with the same eigenvalues as A .

As already mentioned, the *QR-method* will become too laborious for arbitrary matrices, and instead it is used on special matrices, preferably Hessenberg or symmetric band-matrices. The method has good stability properties and seems to be one of the most promising at present. Several algorithms in ALGOL treating this method have already been published [15, 16].

REFERENCES

- [1] Lanczos: "An Iteration Method for the Solution of the Eigenvalue Problem of Linear Differential and Integral Operators," *J. of Res.*, 45, 255 (1950).
- [2] Wilkinson: "The Calculation of the Eigenvectors of Codiagonal Matrices," *The Computer Journal*, 1, 90-96.
- [3] Givens: *Numerical Computation of the Characteristic Values of a Real Symmetric Matrix*, Oak Ridge National Laboratory, ORNL-1574 (1954).
- [4] Wilkinson: "Householder's Method for the Solution of the Algebraic Eigenproblem," *The Computer Journal*, 23-27 (April 1960).
- [5] National Bureau of Standards: *Appl. Math. Series*, No. 49 (Washington, 1958).
- [6] Lotkin: "Characteristic Values for Arbitrary Matrices," *Quart. Appl. Math.*, 14, 267-275 (1956).
- [7] Greenstadt: *A Method for Finding Roots of Arbitrary Matrices*, MTAC, 9, 47-52 (1955).
- [8] Goldstine, Horwitz: *A Procedure for the Diagonalization of Normal Matrices*; *Jour. Assoc. Comp. Mach.* 6, 1959 pp. 176-195.
- [9] Causey, Gregory: *On Lanczos' Algorithm for Tridiagonalizing Matrices*; *SIAM Review* Vol. 3, No. 4, Oct. 1961.
- [10] Eberlein: *A Jacobi-like Method for the Automatic Computation of Eigenvalues and Eigenvectors of an Arbitrary Matrix*; *J. SIAM* Vol. 10, No. 1, March 1962.
- [11] La Budde: *The Reduction of an Arbitrary Real Square Matrix to Tridiagonal Form Using Similarity Transformations*; *Math. Comp.* 17 (1963) p. 433.
- [12] Francis: *The QR Transformation—A Unitary Analogue to the LR Transformation*; *Comp. Jour.* 4 (1961) pp. 265-271 and 332-345.
- [13] Fröberg: *On Triangularization of Complex Matrices by Two-Dimensional Unitary Transformations*; *BIT* 5: 4 (1965) p. 230.
- [14] Householder, Bauer: *On Certain Methods for Expanding the Characteristic Polynomial*; *Num. Math* 1 (1959) p. 29.
- [15] Businger *Eigenvalues of a real symmetric matrix by the QR method*; *Algorithm* 253, *Comm. ACM* 8: 4, p. 217.

- [16] Ruhe: Eigenvalues of a complex matrix by the QR method; BIT 6 : 4 (1966), p. 350. 17 Hyman: Eigenvalues and Eigenvectors of General Matrices; Ass. Comp. Mach. (June 1957). 18 Wilkinson: The Algebraic Eigenvalue Problem; Oxford Univ. Press 1965. 19 Wilkinson: Convergence of the LR, QR and Related Algorithms; Comp. Jour. 8 (1966), p. 77.

EXERCISES

1. Find the largest eigenvalue of the matrix

$$A = \begin{pmatrix} 25 & -41 & 10 & -6 \\ -41 & 68 & -17 & 10 \\ 10 & -17 & 5 & -3 \\ -6 & 10 & -3 & 2 \end{pmatrix},$$

correct to three places.

2. Find the absolutely smallest eigenvalue and the corresponding eigenvector of the matrix

$$A = \begin{pmatrix} 1 & 2 & -2 & 4 \\ 2 & 12 & 3 & 5 \\ 3 & 13 & 0 & 7 \\ 2 & 11 & 2 & 2 \end{pmatrix},$$

using the fact that λ^{-1} is an eigenvalue of A^{-1} if λ is an eigenvalue of A .

3. Find the largest eigenvalue and the corresponding eigenvector of the Hermitian matrix

$$H = \begin{pmatrix} 8 & -5i & 3 - 2i \\ 5i & 3 & 0 \\ 3 + 2i & 0 & 2 \end{pmatrix}.$$

4. Using the fact that $\lambda - a$ is an eigenvalue of $A - aI$, if λ is an eigenvalue of A , find the highest and the lowest eigenvalue of the matrix

$$A = \begin{pmatrix} 9 & 10 & 8 \\ 10 & 5 & -1 \\ 8 & -1 & 3 \end{pmatrix}.$$

Choose $a = 12$ and the starting vectors

$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix},$$

respectively. (Desired accuracy: two decimal places.)

5. The matrix

$$A = \begin{pmatrix} 14 & 7 & 6 & 9 \\ 7 & 9 & 4 & 6 \\ 6 & 4 & 9 & 7 \\ 9 & 6 & 7 & 15 \end{pmatrix}$$

has an eigenvalue close to 4. Compute this eigenvalue to six places, using the matrix $B = (A - 4I)^{-1}$.

6. A is a matrix with one eigenvalue λ_0 and another $-\lambda_0$ (λ_0 real, > 0); all remaining eigenvalues are such that $|\lambda| < \lambda_0$. Generalize the power method so that it can be used in this case. Use the result for computing λ_0 for the matrix

$$A = \begin{pmatrix} 2.24 & -2.15 & -7.37 \\ -2.15 & 0.75 & -0.87 \\ -7.37 & -0.87 & -1.99 \end{pmatrix} \text{ (two decimals).}$$

7. Find the largest eigenvalue of the modified eigenvalue problem $Ax = \lambda Bx$ when

$$B = \begin{pmatrix} 1 & 2 & -1 & 4 \\ 2 & 5 & 1 & 6 \\ -1 & 1 & 11 & -11 \\ 4 & 6 & -11 & 22 \end{pmatrix} \text{ and } A = \begin{pmatrix} 1 & 6 & 6 & 4 \\ 6 & 37 & 43 & 16 \\ 6 & 43 & 86 & -27 \\ 4 & 16 & -27 & 106 \end{pmatrix}$$

8. Show that

$$A = \begin{pmatrix} a_0 & a_1 & a_2 & a_3 \\ a_3 & a_0 & a_1 & a_2 \\ a_2 & a_3 & a_0 & a_1 \\ a_1 & a_2 & a_3 & a_0 \end{pmatrix}$$

can be written $a_0I + a_1C + a_2C^2 + a_3C^3$, where C is a constant matrix. Also find the eigenvalues and the eigenvectors of A . (A is called *circulant*.)

9. In the matrix A of type (n, n) all diagonal elements are a , while all the others are b . Find such numbers p and q that $A^2 - pA + qI = 0$. Use this relation for finding eigenvalues [one is simple, and one is $(n-1)$ -fold] and the eigenvectors of A .
10. Using the LR-method, find A_1 , A_2 , and A_3 when

$$A = \begin{pmatrix} 7 & 6 \\ 3 & 4 \end{pmatrix}.$$

11. A is a matrix with eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ (all different). Put

$$B_k = \frac{(A - \lambda_1 I)(A - \lambda_2 I) \cdots (A - \lambda_{k-1} I)(A - \lambda_{k+1} I) \cdots (A - \lambda_n I)}{(\lambda_k - \lambda_1)(\lambda_k - \lambda_2) \cdots (\lambda_k - \lambda_{k-1})(\lambda_k - \lambda_{k+1}) \cdots (\lambda_k - \lambda_n)}.$$

Show that $B_k^2 = B_k$.

12. A real, symmetric matrix A has the largest eigenvalue equal to λ_1 and the next largest equal to λ_2 , and both are greater than zero. All the other eigenvalues are considerably less in absolute value than these two. By use of the power method one has obtained a vector of the form $u + \varepsilon v$, where u and v are eigenvectors corresponding, to the eigenvalues λ_1 and λ_2 . From these values a series of consecutive Rayleigh quotients is constructed. Show how λ_1 can be accurately determined from three such quotients R_1, R_2 , and R_3 .

Answers to exercises (page 421)

Chapter 6

1. 98.522

2. 0.0122056; $\begin{pmatrix} -110.595 \\ 24.957 \\ -27.665 \\ 1 \end{pmatrix}$

3. 12.054; $\begin{pmatrix} 1 \\ 0.5522i \\ 0.0995(3 + 2i) \end{pmatrix}$

4. 19.29; -7.08

5. 4.040129

6. 8.00

7. 70.21

8. $\lambda_1 = a_0 + a_1 + a_2 + a_3$
 $\lambda_2 = a_0 - a_1 + a_2 - a_3$
 $\lambda_3 = a_0 + ia_1 - a_2 - ia_3$
 $\lambda_4 = a_0 - ia_1 - a_2 + ia_3$
 $x_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}; x_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}; x_3 = \begin{pmatrix} 1 \\ i \\ -1 \\ -i \end{pmatrix}; x_4 = \begin{pmatrix} 1 \\ -i \\ -1 \\ i \end{pmatrix}$

9. $p = 2a + (n - 2)b; q = (a - b)[a + (n - 1)b]$
 $\lambda_1 = a + (n - 1)b$ (simple); $\lambda_2 = a - b$ [($n - 1$)-fold].

10. $A_1 = \begin{pmatrix} 9.5714 & 6 \\ 0.6122 & 1.4286 \end{pmatrix}; A_2 = \begin{pmatrix} 9.9552 & 6 \\ 0.0668 & 1.0448 \end{pmatrix}; A_3 = \begin{pmatrix} 9.9955 & 6 \\ 0.0067 & 1.0045 \end{pmatrix}$

Exact eigenvalues: 10 and 1.

11. An arbitrary vector \mathbf{v} can be written as a linear combination of the eigenvectors.

12. $\lambda_1 = \frac{R_1 R_3 - R_2^2}{R_1 - 2R_2 + R_3} = R_3 + \frac{R_3 - R_2}{(R_2 - R_1)/(R_3 - R_2) - 1}$.