11.3 Jacobi's Method

Jacobi's method is an easily understood algorithm for finding all eigenpairs for a symmetric matrix. It is a reliable method that produces uniformly accurate answers for the results. For matrices of order up to 10, the algorithm is competitive with more sophisticated ones. If speed is not a major consideration, it is quite acceptable for matrices up to order 20.

A solution is guaranteed for all real symmetric matrices when Jacobi's method is used. This limitation is not severe since many practical problems of applied mathematics and engineering involve symmetric matrices. From a theoretical viewpoint, the method embodies techniques that are found in more sophisticated algorithms. For instructive purposes, it is worthwhile to investigate the details of Jacobi's method.

Plane Rotations

We start with some geometrical background about coordinate transformations. Let X denote a vector in n-dimensional space and consider the linear transformation $Y = RX$, where R is an $n \times n$ matrix:

$$R = \begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & & & & & & \vdots \\ 0 & \cdots & \cos\phi & \cdots & \sin\phi & \cdots & 0 \\ \vdots & & & & & & \vdots \\ 0 & \cdots & -\sin\phi & \cdots & \cos\phi & \cdots & 0 \\ \vdots & & & & & & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} \begin{matrix} \\ \\ \leftarrow \text{row } p \\ \\ \leftarrow \text{row } q \\ \\ \\ \end{matrix}$$

$$\begin{matrix} \uparrow & & \uparrow \\ \text{col } p & & \text{col } q \end{matrix}$$

Here all off-diagonal elements of R are zero except for the values $\pm\sin\phi$, and all diagonal elements are 1 except for $\cos\phi$. The effect of the transformation $Y = RX$ is easy to grasp:

$$y_j = x_j \qquad \text{when } j \neq p \text{ and } j \neq q,$$
$$y_p = x_p \cos\phi + x_q \sin\phi,$$
$$y_q = -x_p \sin\phi + x_q \cos\phi.$$

The transformation is seen to be a rotation of n-dimensional space in the $x_p x_q$-plane through the angle ϕ. By selecting an appropriate angle ϕ, we could make either $y_p = 0$ or $y_q = 0$ in the image. The inverse transformation $X = R^{-1}Y$ rotates space in the same $x_p x_q$-plane through the angle $-\phi$. Observe that R is an orthogonal matrix; that is,

$$R^{-1} = R' \quad \text{or} \quad R'R = I.$$

Similarity and Orthogonal Transformations

Consider the eigenproblem

$$(1) \qquad\qquad AX = \lambda X.$$

Suppose that K is a nonsingular matrix and that B is defined by

$$(2) \qquad\qquad B = K^{-1}AK.$$

Multiply both members of (2) on the right side by the quantity $K^{-1}X$. This produces

$$(3) \qquad \begin{aligned} BK^{-1}X &= K^{-1}AKK^{-1}X = K^{-1}AX \\ &= K^{-1}\lambda X = \lambda K^{-1}X. \end{aligned}$$

We define the change of variable

$$(4) \qquad\qquad Y = K^{-1}X \quad \text{or} \quad X = KY.$$

When (4) is used in (3), the new eigenproblem is

$$(5) \qquad\qquad BY = \lambda Y.$$

Comparing (1) and (5), we see that the similarity transformation (2) preserved the eigenvalue λ and that the eigenvectors are different, but are related by the change of variable in (4).

Suppose that the matrix R is an orthogonal matrix (i.e., $R^{-1} = R'$) and that D is defined by

$$(6) \qquad\qquad D = R'AR.$$

Multiply both terms in (6) on the right by $R'X$ to obtain

$$(7) \qquad DR'X = R'ARR'X = R'AX = R'\lambda X = \lambda R'X.$$

We define the change of variable

$$(8) \qquad\qquad Y = R'X \quad \text{or} \quad X = RY.$$

Now use (8) in (7) to obtain a new eigenproblem,

$$(9) \qquad\qquad DY = \lambda Y.$$

As before, the eigenvalues of (1) and (9) are the same. However, for equation (9) the change of variable (8) makes it easier to convert X to Y and Y back into X because $R^{-1} = R'$.

In addition, suppose that A is a symmetric matrix (i.e., $A = A'$). Then we find that

$$(10) \qquad D' = (R'AR)' = R'A(R')' = R'AR = D.$$

Hence D is a symmetric matrix. Therefore, we conclude that if A is a symmetric matrix and R is an orthogonal matrix, the transformation of A to D given by (6) preserves symmetry as well as eigenvalues. The relationship between their eigenvectors is given by the change of variable (8).

Jacobi Series of Transformations

Start with the real symmetric matrix A. Then construct the sequence of orthogonal matrices $R_1, R_2, \ldots, R_n$ as follows:

$$(11) \qquad \begin{aligned} D_0 &= A, \\ D_j &= R'_j D_{j-1} R_j \quad \text{for } j = 1, 2, \ldots. \end{aligned}$$

We will show how to construct the sequence $\{R_j\}$ so that

$$(12) \qquad \lim_{j \to \infty} D_j = D = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n).$$

In practice we will stop when the off-diagonal elements are close to zero. Then we will have

$$(13) \qquad D_n \approx D.$$

The construction produces

$$(14) \qquad D_n = R'_n R'_{n-1} \cdots R'_1 A R_1 R_2 \cdots R_{n-1} R_n.$$

If we define

$$(15) \qquad R = R_1 R_2 \cdots R_{n-1} R_n,$$

then $R^{-1} A R = D_k$, which implies that

$$(16) \qquad A R = R D_k \approx R \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n).$$

Let the columns of R be denoted by the vectors $X_1, X_2, \ldots, X_n$. Then R can be expressed as a row vector of column vectors:

$$(17) \qquad R = \begin{bmatrix} X_1 & X_2 & \cdots & X_n \end{bmatrix}.$$

The columns of the products in (16) now take on the form

$$(18) \qquad \begin{bmatrix} A X_1 & A X_2 & \cdots & A X_n \end{bmatrix} \approx \begin{bmatrix} \lambda_1 X_1 & \lambda_2 X_2 & \cdots & \lambda_n X_n \end{bmatrix}.$$

From (17) and (18) we see that the vector X_j, which is the jth column of R, is an eigenvector that corresponds to the eigenvalue λ_j.

General Step

Each step in the Jacobi iteration will accomplish the limited objective of reduction of the two off-diagonal elements d_{pq} and d_{qp} to zero. Let R_1 denote the first orthogonal matrix used. Suppose that

$$(19) \qquad D_1 = R'_1 A R_1$$

reduces the elements d_{pq} and d_{qp} to zero, where $\boldsymbol{R}_1$ has the form

(20)
$$
\boldsymbol{R}_1 =
\begin{bmatrix}
1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\
\vdots & & & & & & \vdots \\
0 & \cdots & c & \cdots & s & \cdots & 0 \\
\vdots & & & & & & \vdots \\
0 & \cdots & -s & \cdots & c & \cdots & 0 \\
\vdots & & & & & & \vdots \\
0 & \cdots & 0 & \cdots & 0 & \cdots & 1
\end{bmatrix}
\begin{matrix} \\ \\ \leftarrow \text{row } p \\ \\ \leftarrow \text{row } q \\ \\ \\ \end{matrix}
$$
$$\qquad\qquad\qquad\quad \underset{\text{col } p}{\uparrow} \qquad \underset{\text{col } q}{\uparrow}$$

Here all off-diagonal elements of $\boldsymbol{R}_1$ are zero except for the element s located in row p, column q, and the element $-s$ located in row q, column p. Also note that all diagonal elements are 1 except for the element c, which appears at two locations, in row p, column p, and in row q, column q. The matrix is a plane rotation where we have used the notation $c = \cos\phi$ and $s = \sin\phi$.

We must verify that the transformation (19) will produce a change only to rows p and q and columns p and q. Consider postmultiplication of $\boldsymbol{A}$ by $\boldsymbol{R}_1$ and the product $\boldsymbol{B} = \boldsymbol{A}\boldsymbol{R}_1$:

(21)
$$
\boldsymbol{B} =
\begin{bmatrix}
a_{11} & \cdots & a_{1p} & \cdots & a_{1q} & \cdots & a_{1n} \\
a_{p1} & \cdots & a_{pp} & \cdots & a_{pq} & \cdots & a_{pn} \\
a_{q1} & \cdots & a_{qp} & \cdots & a_{qq} & \cdots & a_{qn} \\
a_{n1} & \cdots & a_{np} & \cdots & a_{nq} & \cdots & a_{nn}
\end{bmatrix}
\begin{bmatrix}
1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\
0 & \cdots & c & \cdots & s & \cdots & 0 \\
0 & \cdots & -s & \cdots & c & \cdots & 0 \\
0 & \cdots & 0 & \cdots & 0 & \cdots & 1
\end{bmatrix}
$$

The row by column rule for multiplication applies, and we observe that there is no change to columns 1 to $p-1$ and $p+1$ to $q-1$ and $q+1$ to n. Hence only columns p and q are altered.

(22)
$$
\begin{aligned}
b_{jk} &= a_{jk} & &\text{when } k \neq p \text{ and } k \neq q, \\
b_{jp} &= c a_{jp} - s a_{jq} & &\text{for } j = 1, 2, \ldots, n, \\
b_{jq} &= s a_{jp} + c a_{jq} & &\text{for } j = 1, 2, \ldots, n.
\end{aligned}
$$

A similar argument shows that premultiplication of $\boldsymbol{A}$ by $\boldsymbol{R}_1'$ will only alter rows p and q. Therefore, the transformation

(23)
$$
\boldsymbol{D}_1 = \boldsymbol{R}_1' \boldsymbol{A} \boldsymbol{R}_1
$$

will alter only columns p and q and rows p and q of $\boldsymbol{A}$. The elements d_{jk} of $\boldsymbol{D}_1$ are

computed with the formulas

$$
\begin{aligned}
d_{jp} &= ca_{jp} - sa_{jq} & \text{when } j \neq p \text{ and } j \neq q, \\
d_{jq} &= sa_{jp} + ca_{jq} & \text{when } j \neq p \text{ and } j \neq q, \\
d_{pp} &= c^2 a_{pp} + s^2 a_{qq} - 2csa_{pq}, \\
d_{qq} &= s^2 a_{pp} + c^2 a_{qq} + 2csa_{pq}, \\
d_{pq} &= (c^2 - s^2)a_{pq} + cs(a_{pp} - a_{qq}),
\end{aligned}
$$

(24)

and the other elements of D_1 are found by symmetry.

Zeroing Out d_{pq} and d_{qp}

The goal for each step of Jacobi's iteration is to make the two off-diagonal elements d_{pq} and d_{qp} zero. The obvious strategy would be to observe the fact that

(25)
$$
c = \cos\phi \quad \text{and} \quad s = \sin\phi,
$$

where ϕ is the angle of rotation that produces the desired effect. However, some ingenious maneuvers with trigonometric identities are now required. The identity for $\cot\phi$ is used with (25) to define

(26)
$$
\theta = \cot 2\phi = \frac{c^2 - s^2}{2cs}.
$$

Suppose that $a_{pq} \neq 0$ and we want to produce $d_{pq} = 0$. Then using the last equation in (24), we obtain

(27)
$$
0 = (c^2 - s^2)a_{pq} + cs(a_{pp} - a_{qq}).
$$

This can be rearranged to yield $(c^2 - s^2)/(cs) = (a_{qq} - a_{pp})/a_{pq}$, which is used in (26) to solve for θ:

(28)
$$
\theta = \frac{a_{qq} - a_{pp}}{2a_{pq}}.
$$

Although we can use (28) with formulas (25) and (26) to compute c and s, less round-off error is propagated if we compute $\tan \phi$ and use it in later computations. So we define

(29)
$$t = \tan \phi = \frac{s}{c}.$$

Now divide the numerator and denominator in (26) by c^2 to obtain

$$\theta = \frac{1 - s^2/c^2}{2s/c} = \frac{1 - t^2}{2t},$$

which yields the equation

(30)
$$t^2 + 2t\theta - 1 = 0.$$

Since $t = \tan \phi$, the smaller root of (30) corresponds to the smaller angle of rotation with $|\phi| \le \pi/4$. The special form of the quadratic formula for finding this root is

(31)
$$t = -\theta \pm (\theta^2 + 1)^{1/2} = \frac{\text{sign}(\theta)}{|\theta| + (\theta^2 + 1)^{1/2}},$$

where $\text{sign}(\theta) = 1$ when $\theta \ge 0$ and $\text{sign}(\theta) = -1$ when $\theta < 0$. Then c and s are computed with the formulas

(32)
$$c = \frac{1}{(t^2 + 1)^{1/2}}$$
$$s = ct.$$

Summary of the General Step

We can now outline the calculations required to zero out the element d_{pq}. First, select row p and column q for which $a_{pq} \ne 0$. Second, form the preliminary quantities

(33)
$$\theta = \frac{a_{qq} - a_{pp}}{2a_{pq}},$$
$$t = \frac{\text{sign}(\theta)}{|\theta| + (\theta^2 + 1)^{1/2}},$$
$$c = \frac{1}{(t^2 + 1)^{1/2}},$$
$$s = ct.$$

Third, to construct $D = D_1$, use

$$d_{pq} = 0;$$
$$d_{qp} = 0;$$
$$d_{pp} = c^2 a_{pp} + s^2 a_{qq} - 2cs a_{pq};$$
$$d_{qq} = s^2 a_{pp} + c^2 a_{qq} + 2cs a_{pq};$$

(34)

$$
\begin{aligned}
&\text{for} \quad j = 1 : n \\
&\quad \text{if} \quad (j \sim= p) \quad \text{and} \quad (j \sim= q) \\
&\qquad d_{jp} = c a_{jp} - s a_{jq}; \\
&\qquad d_{pj} = d_{jp}; \\
&\qquad d_{jq} = c a_{jq} + s a_{jp}; \\
&\qquad d_{qj} = d_{jq}; \\
&\quad \text{end} \\
&\text{end}
\end{aligned}
$$

Updating the Matrix of Eigenvectors

We need to keep track of the matrix product $R_1 R_2 \cdots R_n$. When we stop at the nth iteration, we will have computed

(35)
$$V_n = R_1 R_2 \cdots R_n,$$

where V_n is an orthogonal matrix. We need only keep track of the current matrix V_j, for $j = 1, 2, \ldots, n$. Start by initializing $V = I$. Use the vector variables **XP** and **XQ** to store columns p and q of V, respectively. Then for each step perform the calculation

(36)

$$
\begin{aligned}
&\text{for} \quad j = 1 : n \\
&\quad \mathbf{XP}_j = v_{jp}; \\
&\quad \mathbf{XQ}_j = v_{jq}; \\
&\text{end} \\
&\text{for} \quad j = 1 : n \\
&\quad v_{jp} = c\mathbf{XP}_j - s\mathbf{XQ}_j; \\
&\quad v_{jq} = s\mathbf{XP}_j + c\mathbf{XQ}_j; \\
&\text{end}
\end{aligned}
$$

Strategy for Eliminating a_{pq}

The speed of convergence of Jacobi's method is seen by considering the sum of the squares of the off-diagonal elements:

(37)
$$S_1 = \sum_{\substack{j,k=1 \\ k \neq j}}^{n} |a_{jk}|^2$$

(38)
$$S_2 = \sum_{\substack{j,k=1 \\ k \neq j}}^{n} |d_{jk}|^2, \quad \text{where} \quad \boldsymbol{D}_1 = \boldsymbol{R}'\boldsymbol{A}\boldsymbol{R}.$$

The reader can verify that the equations given in (34) can be used to prove that

(39)
$$S_2 = S_1 - 2|a_{pq}|^2.$$

At each step we let S_j denote the sum of the squares of the off-diagonal elements of $\boldsymbol{D}_j$. Then the sequence $\{S_j\}$ decreases monotonically and is bounded below by zero. Jacobi's original algorithm of 1846 selected, at each step, the off-diagonal element a_{pq} of largest magnitude to zero out and involved a search to compute the value

(40)
$$\max\{\boldsymbol{A}\} = \max_{p<q}\{|a_{pq}|\}.$$

This choice will guarantee that $\{S_j\}$ converges to zero. As a consequence, this proves that $\{\boldsymbol{D}_j\}$ converges to $\boldsymbol{D}$ and $\{\boldsymbol{V}_j\}$ converges to the matrix $\boldsymbol{V}$ of eigenvectors.

Jacobi's search can become time consuming since it requires an order of $(n^2-n)/2$ comparisons in a loop. It is prohibitive for larger values of n. A better strategy is the cyclic Jacobi method, where one annihilates elements in a strict order across the rows. A tolerance value ϵ is selected; then a sweep is made throughout the matrix, and if an element a_{pq} is found to be larger than ϵ, it is zeroed out. For one sweep through the matrix the elements are checked in row 1, $a_{12}, a_{13}, \ldots, a_{1n}$; then row 2, $a_{23}, a_{24}, \ldots,$ a_{2n}; and so on. It has been proved that the convergence rate is quadratic for both the original and cyclic Jacobi methods. An implementation of the cyclic Jacobi method starts by observing that the sum of the squares of the diagonal elements increases with each iteration; that is, if

(41)
$$T_0 = \sum_{j=1}^{n} |a_{jj}|^2$$

and

$$T_1 = \sum_{j=1}^{n} |d_{jj}|^2,$$

then

$$T_1 = T_0 + 2|a_{pq}|^2.$$

Consequently, the sequence $\{D_j\}$ converges to the diagonal matrix D. Notice that the average size of a diagonal element can be computed with the formula $(T_0/n)^{1/2}$. The magnitudes of the off-diagonal elements are compared to $\epsilon(T_0/n)^{1/2}$, where ϵ is the preassigned tolerance. Therefore, the element a_{pq} is zeroed out if

(42)
$$|a_{pq}| > \epsilon \left(\frac{T_0}{n} \right)^{1/2}.$$

Another variation of the method, called the *threshold Jacobi method*, is left for the reader to investigate.

Example 11.7. Use Jacobi iteration to transform the following symmetric matrix into diagonal form.

$$\begin{bmatrix} 8 & -1 & 3 & -1 \\ -1 & 6 & 2 & 0 \\ 3 & 2 & 9 & 1 \\ -1 & 0 & 1 & 7 \end{bmatrix}$$

The computational details are left for the reader. The first rotation matrix that will zero out $a_{13} = 3$ is

$$R_1 = \begin{bmatrix} 0.763020 & 0.000000 & 0.646375 & 0.000000 \\ 0.000000 & 0.000000 & 0.000000 & 0.000000 \\ -0.646375 & 0.000000 & 0.763020 & 0.000000 \\ 0.000000 & 0.000000 & 0.000000 & 0.000000 \end{bmatrix}.$$

Calculation reveals that $A_2 = R_1 A_1 R_1$ is

$$A_2 = \begin{bmatrix} 5.458619 & -2.055770 & 0.000000 & -1.409395 \\ -2.055770 & 6.000000 & 0.879665 & 0.000000 \\ 0.000000 & 0.879665 & 11.541381 & 0.116645 \\ -1.409395 & 0.000000 & 0.116645 & 7.000000 \end{bmatrix}.$$

Next, the element $a_{12} = -2.055770$ is zeroed out and we get

$$A_3 = \begin{bmatrix} 3.655795 & 0.000000 & 0.579997 & -1.059649 \\ 0.000000 & 7.802824 & 0.661373 & 0.929268 \\ 0.579997 & 0.661373 & 11.541381 & 0.116645 \\ -1.059649 & 0.929268 & 0.116645 & 7.000000 \end{bmatrix}.$$

After 10 iterations we arrive at

$$A_{10} = \begin{bmatrix} 3.295870 & 0.002521 & 0.037859 & 0.000000 \\ 0.002521 & 8.405210 & -0.004957 & 0.066758 \\ 0.037859 & -0.004957 & 11.704123 & -0.001430 \\ 0.000000 & 0.066758 & -0.001430 & 6.594797 \end{bmatrix}.$$

It will take six more iterations for the diagonal elements to get close to the diagonal matrix

$$\boldsymbol{D} = \text{diag}(3.295699, 8.407662, 11.704301, 6.592338).$$

However, the off-diagonal elements are not small enough, and it will take three more iterations for them to be less than 10^{-6} in magnitude. Then the eigenvectors are the columns of the matrix $\boldsymbol{V} = \boldsymbol{R}_1 \boldsymbol{R}_2 \cdots \boldsymbol{R}_{18}$, which is

$$\boldsymbol{V} = \begin{bmatrix} 0.528779 & -0.573042 & 0.582298 & 0.230097 \\ 0.591967 & 0.472301 & 0.175776 & -0.628975 \\ -0.536039 & 0.282050 & 0.792487 & -0.071235 \\ 0.287454 & 0.607455 & 0.044680 & 0.739169 \end{bmatrix}. \qquad \blacksquare$$