11.4 Eigenvalues of Symmetric Matrices

Householder's Method

Each transformation in Jacobi's method produced two zero off-diagonal elements, but subsequent iterations might make them nonzero. Hence many iterations are required to make the off-diagonal entries sufficiently close to zero. We now develop a method that produces several zero off-diagonal elements in each iteration, and they remain zero in subsequent iterations. We start by developing an important step in the process.

Theorem 11.23 (Householder Reflection). If $\boldsymbol{X}$ and $\boldsymbol{Y}$ are vectors with the same norm, there exists an orthogonal symmetric matrix $\boldsymbol{P}$ such that

$$\boldsymbol{Y} = \boldsymbol{P}\boldsymbol{X}, \tag{1}$$

where

$$\boldsymbol{P} = \boldsymbol{I} - 2\boldsymbol{W}\boldsymbol{W}' \tag{2}$$

and

$$\boldsymbol{W} = \frac{\boldsymbol{X} - \boldsymbol{Y}}{\|\boldsymbol{X} - \boldsymbol{Y}\|_2}. \tag{3}$$

Since $\boldsymbol{P}$ is both orthogonal and symmetric, it follows that

$$\boldsymbol{P}^{-1} = \boldsymbol{P}. \tag{4}$$

Proof. Equation (3) is used and defines $\boldsymbol{W}$ to be the unit vector in the direction $\boldsymbol{X}-\boldsymbol{Y}$; hence

$$\boldsymbol{W}'\boldsymbol{W} = 1 \tag{5}$$

and

$$\boldsymbol{Y} = \boldsymbol{X} + c\boldsymbol{W}, \tag{6}$$

where $c = -\|\boldsymbol{X} - \boldsymbol{Y}\|_2$. Since $\boldsymbol{X}$ and $\boldsymbol{Y}$ have the same norm, the parallelogram rule for vector addition can be used to see that $\boldsymbol{Z} = (\boldsymbol{X}+\boldsymbol{Y})/2 = \boldsymbol{X}+(c/2)\boldsymbol{W}$ is orthogonal to vector $\boldsymbol{W}$ (see Figure 11.4). This implies that

$$\boldsymbol{W}'\left(\boldsymbol{X} + \frac{c}{2}\boldsymbol{W}\right) = 0.$$

Now we can use (5) to expand the preceding equation and get

$$\boldsymbol{W}'\boldsymbol{X} + \frac{c}{2}\boldsymbol{W}'\boldsymbol{W} = \boldsymbol{W}'\boldsymbol{X} + \frac{c}{2} = 0. \tag{7}$$

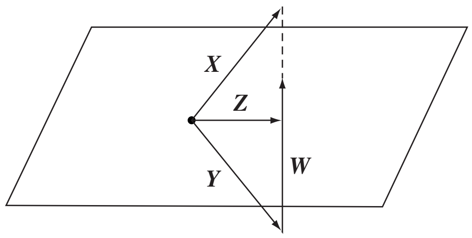

Figure 11.4 The vectors $\boldsymbol{W}$, $\boldsymbol{X}$, $\boldsymbol{Y}$, and $\boldsymbol{Z}$ involved in the Householder reflection.

The crucial step is to use (7) and express c in the form

$$c = -2(\boldsymbol{W}'\boldsymbol{X}). \tag{8}$$

Now (8) can be used in (6) to see that

$$\boldsymbol{Y} = \boldsymbol{X} + c\boldsymbol{W} = \boldsymbol{X} - 2\boldsymbol{W}'\boldsymbol{X}\boldsymbol{W}.$$

Since the quantity $\boldsymbol{W}'\boldsymbol{X}$ is a scalar, the last equation can be written as

$$\boldsymbol{Y} = \boldsymbol{X} - 2\boldsymbol{W}\boldsymbol{W}'\boldsymbol{X} = (\boldsymbol{I} - 2\boldsymbol{W}\boldsymbol{W}')\boldsymbol{X}. \tag{9}$$

Looking at (9), we see that $\boldsymbol{P} = \boldsymbol{I} - 2\boldsymbol{W}\boldsymbol{W}'$. The matrix $\boldsymbol{P}$ is symmetric because

$$\begin{aligned} \boldsymbol{P}' &= (\boldsymbol{I} - 2\boldsymbol{W}\boldsymbol{W}')' = \boldsymbol{I} - 2(\boldsymbol{W}\boldsymbol{W}')' \\ &= \boldsymbol{I} - 2\boldsymbol{W}\boldsymbol{W}' = \boldsymbol{P}. \end{aligned}$$

The following calculation shows that $\boldsymbol{P}$ is orthogonal:

$$\begin{aligned} \boldsymbol{P}'\boldsymbol{P} &= (\boldsymbol{I} - 2\boldsymbol{W}\boldsymbol{W}')(\boldsymbol{I} - 2\boldsymbol{W}\boldsymbol{W}') \\ &= \boldsymbol{I} - 4\boldsymbol{W}\boldsymbol{W}' + 4\boldsymbol{W}\boldsymbol{W}'\boldsymbol{W}\boldsymbol{W}' \\ &= \boldsymbol{I} - 4\boldsymbol{W}\boldsymbol{W}' + 4\boldsymbol{W}\boldsymbol{W}' = \boldsymbol{I}, \end{aligned}$$

and the proof is complete. •

It should be observed that the effect of the mapping $\boldsymbol{Y} = \boldsymbol{P}\boldsymbol{X}$ is to reflect $\boldsymbol{X}$ through the line whose direction is $\boldsymbol{Z}$, hence the name ***Householder reflection***.

Corollary 11.3 (kth Householder Matrix). Let $\boldsymbol{A}$ be an $n \times n$ matrix, and $\boldsymbol{X}$ any vector. If k is an integer with $1 \le k \le n-2$, we can construct a vector $\boldsymbol{W}_k$ and matrix $\boldsymbol{P}_k = \boldsymbol{I} - 2\boldsymbol{W}_k\boldsymbol{W}_k'$ so that

$$\boldsymbol{P}_k\boldsymbol{X} = \boldsymbol{P}_k \begin{bmatrix} x_1 \\ \vdots \\ x_k \\ x_{k+1} \\ x_{k+2} \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_1 \\ \vdots \\ x_k \\ -S \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \boldsymbol{Y}. \tag{10}$$

Proof. The key is to define the value S so that $\|\boldsymbol{X}\|_2 = \|\boldsymbol{Y}\|_2$ and then invoke Theorem 11.23. The proper value for S must satisfy

$$S^2 = x_{k+1}^2 + x_{k+2}^2 + \cdots + x_n^2, \tag{11}$$

which is readily verified by computing the norms of $\boldsymbol{X}$ and $\boldsymbol{Y}$:

$$\begin{aligned}\|\boldsymbol{X}\|_2 &= x_1^2 + x_2^2 + \cdots + x_n^2 \\ &= x_1^2 + x_2^2 + \cdots + x_k^2 + S^2 \\ &= \|\boldsymbol{Y}\|_2 .\end{aligned} \tag{12}$$

The vector $\boldsymbol{W}$ is found by using equation (3) of Theorem 11.23:

$$\begin{aligned}\boldsymbol{W} &= \frac{1}{R}(\boldsymbol{X} - \boldsymbol{Y}) \\ &= \frac{1}{R}\begin{bmatrix}0 & \cdots & 0 & (x_{k+1} + S) & x_{k+2} & \cdots & x_n\end{bmatrix}'.\end{aligned} \tag{13}$$

Less round-off error is propagated when the sign of S is chosen to be the same as the sign of x_{k+1}; hence we compute

$$S = \operatorname{sign}(x_{k+1})(x_{k+1}^2 + x_{k+2}^2 + \cdots + x_n^2)^{1/2}. \tag{14}$$

The number R in (13) is chosen so that $\|\boldsymbol{W}\|_2 = 1$ and must satisfy

$$\begin{aligned}R^2 &= (x_{k+1} + S)^2 + x_{k+2}^2 + \cdots + x_n^2 \\ &= 2x_{k+1}S + S^2 + x_{k+1}^2 + x_{k+2}^2 + \cdots + x_n^2 \\ &= 2x_{k+1}S + 2S^2.\end{aligned} \tag{15}$$

Therefore, the matrix $\boldsymbol{P}_k$ is given by the formula

$$\boldsymbol{P}_k = \boldsymbol{I} - 2\boldsymbol{W}\boldsymbol{W}', \tag{16}$$

and the proof is complete. •

Householder Transformation

Suppose that $\boldsymbol{A}$ is a symmetric $n \times n$ matrix. Then a sequence of $n-2$ transformations of the form $\boldsymbol{PAP}$ will reduce $\boldsymbol{A}$ to a symmetric tridiagonal matrix. Let us visualize the process when $n = 5$. The first transformation is defined to be $\boldsymbol{P}_1\boldsymbol{A}\boldsymbol{P}_1$, where $\boldsymbol{P}_1$ is constructed by applying Corollary 11.3, with the vector $\boldsymbol{X}$ being the first column of the matrix $\boldsymbol{A}$. The general form of $\boldsymbol{P}_1$ is

$$\boldsymbol{P}_1 = \begin{bmatrix}1 & 0 & 0 & 0 & 0 \\ 0 & p & p & p & p \\ 0 & p & p & p & p \\ 0 & p & p & p & p \\ 0 & p & p & p & p\end{bmatrix}, \tag{17}$$

where the letter p stands for some element in $\boldsymbol{P}_1$. As a result, the transformation $\boldsymbol{P}_1\boldsymbol{A}\boldsymbol{P}_1$ does not affect the element a_{11} of $\boldsymbol{A}$:

$$\boldsymbol{P}_1\boldsymbol{A}\boldsymbol{P}_1 = \begin{bmatrix} a_{11} & v_1 & 0 & 0 & 0 \\ u_1 & w_1 & w & w & w \\ 0 & w & w & w & w \\ 0 & w & w & w & w \\ 0 & w & w & w & w \end{bmatrix} = \boldsymbol{A}_1. \tag{18}$$

The element denoted u_1 is changed because of premultiplication by $\boldsymbol{P}_1$, and v_1 is changed because of postmultiplication by $\boldsymbol{P}_1$; since $\boldsymbol{A}_1$ is symmetric, we have $u_1 = v_1$. The changes to the elements denoted w have been affected by both premultiplication and postmultiplication. Also, since $\boldsymbol{X}$ is the first column of $\boldsymbol{A}$, equation (10) implies that $u_1 = -S$.

The second Householder transformation is applied to the matrix $\boldsymbol{A}_1$ defined in (18) and is denoted $\boldsymbol{P}_2\boldsymbol{A}\boldsymbol{P}_2$, where $\boldsymbol{P}_2$ is constructed by applying Corollary 11.3, with the vector $\boldsymbol{X}$ being the second column of the matrix $\boldsymbol{A}_1$. The form of $\boldsymbol{P}_2$ is

$$\boldsymbol{P}_2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & p & p & p \\ 0 & 0 & p & p & p \\ 0 & 0 & p & p & p \end{bmatrix}, \tag{19}$$

where p stands for some element in $\boldsymbol{P}_2$. The 2×2 identity block in the upper-left corner ensures that the partial tridiagonalization achieved in the first step will not be altered by the second transformation $\boldsymbol{P}_2\boldsymbol{A}_1\boldsymbol{P}_2$. The outcome of this transformation is

$$\boldsymbol{P}_2\boldsymbol{A}_1\boldsymbol{P}_2 = \begin{bmatrix} a_{11} & v_1 & 0 & 0 & 0 \\ u_1 & w_1 & v_2 & 0 & 0 \\ 0 & u_2 & w_2 & w & w \\ 0 & 0 & w & w & w \\ 0 & 0 & w & w & w \end{bmatrix} = \boldsymbol{A}_2. \tag{20}$$

The elements u_2 and v_2 were affected by premultiplication and postmultiplication by $\boldsymbol{P}_2$. Additional changes have been introduced to the other elements w by the transformation.

The third Householder transformation, $\boldsymbol{P}_3\boldsymbol{A}_2\boldsymbol{P}_3$, is applied to the matrix $\boldsymbol{A}_2$ defined in (20), where the corollary is used with $\boldsymbol{X}$ being the third column of $\boldsymbol{A}_2$. The form of $\boldsymbol{P}_3$ is

$$\boldsymbol{P}_3 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & p & p \\ 0 & 0 & 0 & p & p \end{bmatrix}. \tag{21}$$

Again, the 3×3 identity block ensures that $\boldsymbol{P}_3\boldsymbol{A}_2\boldsymbol{P}_3$ does not affect the elements of $\boldsymbol{A}_2$, which lie in the upper 3×3 corner, and we obtain

$$\boldsymbol{P}_3\boldsymbol{A}_2\boldsymbol{P}_3 = \begin{bmatrix} a_{11} & v_1 & 0 & 0 & 0 \\ u_1 & w_1 & v_2 & 0 & 0 \\ 0 & u_2 & w_2 & v_3 & 0 \\ 0 & 0 & u_3 & w & w \\ 0 & 0 & 0 & w & w \end{bmatrix} = \boldsymbol{A}_3. \tag{22}$$

Thus it has taken three transformations to reduce $\boldsymbol{A}$ to tridiagonal form.

For efficiency, the transformation $\boldsymbol{PAP}$ is not performed in matrix form. The next result shows that it is more efficiently carried out via some clever vector manipulations.

Theorem 11.24 (Computation of One Householder Transformation). If $\boldsymbol{P}$ is a Householder matrix, the transformation $\boldsymbol{PAP}$ is accomplished as follows. Let

$$\boldsymbol{V} = \boldsymbol{AW} \tag{23}$$

and compute

$$c = \boldsymbol{W}'\boldsymbol{V} \tag{24}$$

and

$$\boldsymbol{Q} = \boldsymbol{V} - c\boldsymbol{W}. \tag{25}$$

Then

$$\boldsymbol{PAP} = \boldsymbol{A} - 2\boldsymbol{W}\boldsymbol{Q}' - 2\boldsymbol{Q}\boldsymbol{W}'. \tag{26}$$

Proof. First, form the product

$$\boldsymbol{AP} = \boldsymbol{A}(\boldsymbol{I} - 2\boldsymbol{W}\boldsymbol{W}') = \boldsymbol{A} - 2\boldsymbol{A}\boldsymbol{W}\boldsymbol{W}'.$$

Using equation (23), this is written as

$$\boldsymbol{AP} = \boldsymbol{A} - 2\boldsymbol{V}\boldsymbol{W}'. \tag{27}$$

Now use (27) and write

$$\boldsymbol{PAP} = (\boldsymbol{I} - 2\boldsymbol{W}\boldsymbol{W}')(\boldsymbol{A} - 2\boldsymbol{V}\boldsymbol{W}'). \tag{28}$$

When this quantity is expanded, the term $2(2\boldsymbol{W}\boldsymbol{W}'\boldsymbol{V}\boldsymbol{W}')$ is divided into two portions and (28) can be rewritten as

$$\boldsymbol{PAP} = \boldsymbol{A} - 2\boldsymbol{W}(\boldsymbol{W}'\boldsymbol{A}) + 2\boldsymbol{W}(\boldsymbol{W}'\boldsymbol{V}\boldsymbol{W}') - 2\boldsymbol{V}\boldsymbol{W}' + 2\boldsymbol{W}(\boldsymbol{W}'\boldsymbol{V})\boldsymbol{W}'. \tag{29}$$

Under the assumption that $\boldsymbol{A}$ is symmetric, we can use the identity $(\boldsymbol{W}'\boldsymbol{A}) = (\boldsymbol{W}'\boldsymbol{A}') = \boldsymbol{V}'$. The tricky part is to observe that $(\boldsymbol{W}'\boldsymbol{V})$ is a scalar quantity; hence it can commute freely about in any term. Another scalar identity, $\boldsymbol{W}'\boldsymbol{V} = (\boldsymbol{W}'\boldsymbol{V})'$, is used to obtain the relation $\boldsymbol{W}'\boldsymbol{V}\boldsymbol{W}' = (\boldsymbol{W}'\boldsymbol{V})\boldsymbol{W}' = \boldsymbol{W}'(\boldsymbol{W}'\boldsymbol{V}) = \boldsymbol{W}'(\boldsymbol{W}'\boldsymbol{V})' = ((\boldsymbol{W}'\boldsymbol{V})\boldsymbol{W})' = (\boldsymbol{W}'\boldsymbol{V}\boldsymbol{W})'$. These results are used in the terms of (29) in parentheses to get

$$\text{(30)} \qquad \boldsymbol{PAP} = \boldsymbol{A} - 2\boldsymbol{W}\boldsymbol{V}' + 2\boldsymbol{W}(\boldsymbol{W}'\boldsymbol{V}\boldsymbol{W})' - 2\boldsymbol{V}\boldsymbol{W}' + 2\boldsymbol{W}'\boldsymbol{V}\boldsymbol{W}\boldsymbol{W}'.$$

Now the distributive law is used in (30) and we obtain

$$\text{(31)} \qquad \boldsymbol{PAP} = \boldsymbol{A} - 2\boldsymbol{W}(\boldsymbol{V}' - (\boldsymbol{W}'\boldsymbol{V}\boldsymbol{W})') - 2(\boldsymbol{V} - \boldsymbol{W}'\boldsymbol{V}\boldsymbol{W})\boldsymbol{W}'.$$

Finally, the definition for $\boldsymbol{Q}$ given in (25) is used in (31) and the outcome is equation (26), and the proof is complete. •

Reduction to Tridiagonal Form

Suppose that $\boldsymbol{A}$ is a symmetric $n \times n$ matrix. Start with

$$\text{(32)} \qquad \boldsymbol{A}_0 = \boldsymbol{A}.$$

Construct the sequence $\boldsymbol{P}_1, \boldsymbol{P}_2, \ldots, \boldsymbol{P}_{n-1}$ of Householder matrices, so that

$$\text{(33)} \qquad \boldsymbol{A}_k = \boldsymbol{P}_k\boldsymbol{A}_{k-1}\boldsymbol{P}_k \quad \text{for } k = 1,\ 2,\ \ldots,\ n-2,$$

where $\boldsymbol{A}_k$ has zeros below the subdiagonal in columns $1, 2, \ldots, k$. Then $\boldsymbol{A}_{n-2}$ is a symmetric tridiagonal matrix that is similar to $\boldsymbol{A}$. This process is called ***Householder's method***.

Example 11.8. Use Householder's method to reduce the following matrix to symmetric tridiagonal form:

$$\boldsymbol{A}_0 = \begin{bmatrix} 4 & 2 & 2 & 1 \\ 2 & -3 & 1 & 1 \\ 2 & 1 & 3 & 1 \\ 1 & 1 & 1 & 2 \end{bmatrix}.$$

The details are left for the reader. The constants $S = 3$ and $R = 30^{1/2} = 5.477226$ are used to construct the vector

$$\boldsymbol{W}' = \frac{1}{\sqrt{30}}\begin{bmatrix} 0 & 5 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 0.000000 & 0.912871 & 0.365148 & 0.182574 \end{bmatrix}.$$

Then matrix multiplication $\boldsymbol{V} = \boldsymbol{A}\boldsymbol{W}$ is used to form

$$\begin{aligned} \boldsymbol{V}' &= \frac{1}{\sqrt{30}}\begin{bmatrix} 0 & -12 & 12 & 9 \end{bmatrix} \\ &= \begin{bmatrix} 0.000000 & -2.190890 & 2.190890 & 1.643168 \end{bmatrix}. \end{aligned}$$

The constant $c = \boldsymbol{W}'\boldsymbol{V}$ is then found to be

$$c = -0.9.$$

Then the vector $\boldsymbol{Q} = \boldsymbol{V} - c\boldsymbol{W} = \boldsymbol{V} + 0.9\boldsymbol{W}$ is formed:

$$\begin{aligned}\boldsymbol{Q}' &= \frac{1}{\sqrt{30}}\begin{bmatrix}0.000000 & -7.500000 & 13.800000 & 9.900000\end{bmatrix}\\ &= \begin{bmatrix}0.000000 & -1.369306 & 2.519524 & 1.807484\end{bmatrix}.\end{aligned}$$

The computation $\boldsymbol{A}_1 = \boldsymbol{A}_0 - 2\boldsymbol{W}\boldsymbol{Q}' - 2\boldsymbol{Q}\boldsymbol{W}'$ produces

$$\boldsymbol{A}_1 = \begin{bmatrix} 4.0 & -3.0 & 0.0 & 0.0 \\ -3.0 & 2.0 & -2.6 & -1.8 \\ 0.0 & -2.6 & -0.68 & -1.24 \\ 0.0 & -1.8 & -1.24 & 0.68 \end{bmatrix}.$$

The final step uses the constants $S = -3.1622777$, $R = 6.0368737$, $c = -1.2649111$ and the vectors

$$\begin{aligned}\boldsymbol{W}' &= \begin{bmatrix}0.000000 & 0.000000 & -0.954514 & -0.298168\end{bmatrix},\\ \boldsymbol{V}' &= \begin{bmatrix}0.000000 & 0.000000 & 1.018797 & 0.980843\end{bmatrix},\\ \boldsymbol{Q}' &= \begin{bmatrix}0.000000 & 0.000000 & -0.188578 & 0.603687\end{bmatrix}.\end{aligned}$$

The tridiagonal matrix $\boldsymbol{A}_2 = \boldsymbol{A}_1 - 2\boldsymbol{W}\boldsymbol{Q}' - 2\boldsymbol{Q}\boldsymbol{W}'$ is

$$\boldsymbol{A}_2 = \begin{bmatrix} 4.0 & -3.0 & 0.0 & 0.0 \\ -3.0 & 2.0 & 3.162278 & 0.0 \\ 0.0 & 3.162278 & -1.4 & -0.2 \\ 0.0 & 0.0 & -0.2 & 1.4 \end{bmatrix}. \qquad \blacksquare$$

Numerical Methods Using Matlab, 4th Edition, 2004

John H. Mathews and Kurtis K. Fink

ISBN: 0-13-065248-2

Prentice-Hall Inc.
Upper Saddle River, New Jersey, USA
http://vig.prenhall.com/