NUMERICAL RECIPES

Webnote No. 16, Rev. 1

Description of the QR Algorithm for Hessenberg Matrices

Here we give the details of the QR algorithm for real Hessenberg matrices. First, the trick to avoid complex arithmetic depends on a result analogous to the lemma we used for implicit shifts in §11.4. The lemma we need here states that if $\mathbf{B}$ is a nonsingular matrix such that

$$\mathbf{B} \cdot \mathbf{Q} = \mathbf{Q} \cdot \mathbf{H} \tag{1}$$

where $\mathbf{Q}$ is orthogonal and $\mathbf{H}$ is upper Hessenberg, then $\mathbf{Q}$ and $\mathbf{H}$ are fully determined by column 0 of $\mathbf{Q}$. (The determination is unique if $\mathbf{H}$ has positive subdiagonal elements.) The lemma can be proved by induction analogously to the proof given for tridiagonal matrices in §11.4.

The lemma is used in practice by taking two steps of the QR algorithm, either with two real shifts k_s and k_{s+1}, or with complex conjugate values k_s and $k_{s+1} = k_s*$. This gives a real matrix $\mathbf{A}_{s+2}$, where

$$\mathbf{A}_{s+2} = \mathbf{Q}_{s+1} \cdot \mathbf{Q}_s \cdot \mathbf{A}_s \cdot \mathbf{Q}_s^T \cdot \mathbf{Q}_{s+1}^T. \tag{2}$$

The $\mathbf{Q}$'s are determined by

$$\mathbf{A}_s - k_s \mathbf{1} = \mathbf{Q}_s^T \cdot \mathbf{R}_s \tag{3}$$

$$\mathbf{A}_{s+1} = \mathbf{Q}_s \cdot \mathbf{A}_s \cdot \mathbf{Q}_s^T \tag{4}$$

$$\mathbf{A}_{s+1} - k_{s+1} \mathbf{1} = \mathbf{Q}_{s+1}^T \cdot \mathbf{R}_{s+1} \tag{5}$$

Using (4), equation (5) can be rewritten

$$\mathbf{A}_s - k_{s+1} \mathbf{1} = \mathbf{Q}_s^T \cdot \mathbf{Q}_{s+1}^T \cdot \mathbf{R}_{s+1} \cdot \mathbf{Q}_s \tag{6}$$

Hence, if we define

$$\mathbf{M} = (\mathbf{A}_s - k_{s+1} \mathbf{1}) \cdot (\mathbf{A}_s - k_s \mathbf{1}) \tag{7}$$

equations (3) and (6) give

$$\mathbf{R} = \mathbf{Q} \cdot \mathbf{M} \tag{8}$$

where

$$\mathbf{Q} = \mathbf{Q}_{s+1} \cdot \mathbf{Q}_s \tag{9}$$

$$\mathbf{R} = \mathbf{R}_{s+1} \cdot \mathbf{R}_s \tag{10}$$

Equation (2) can be rewritten

$$\mathbf{A}_s \cdot \mathbf{Q}^T = \mathbf{Q}^T \cdot \mathbf{A}_{s+2} \tag{11}$$

Thus suppose we can somehow find an upper Hessenberg matrix $\mathbf{H}$ such that

$$\mathbf{A}_s \cdot \overline{\mathbf{Q}}^T = \overline{\mathbf{Q}}^T \cdot \mathbf{H} \tag{12}$$

where $\overline{\mathbf{Q}}$ is orthogonal. If $\overline{\mathbf{Q}}^T$ has the same column 0 as $\mathbf{Q}^T$ (i.e., $\overline{\mathbf{Q}}$ has the same row 0 as $\mathbf{Q}$), then $\overline{\mathbf{Q}} = \mathbf{Q}$ and $\mathbf{A}_{s+2} = \mathbf{H}$.

Row 0 of $\mathbf{Q}$ is found as follows. Equation (8) shows that $\mathbf{Q}$ is the orthogonal matrix that triangularizes the real matrix $\mathbf{M}$. Any real matrix can be triangularized by premultiplying it by a sequence of Householder matrices $\mathbf{P}_1$ (acting on column 0), $\mathbf{P}_2$ (acting on column 1), ..., $\mathbf{P}_{n-1}$. Thus $\mathbf{Q} = \mathbf{P}_{n-1} \cdots \mathbf{P}_2 \cdot \mathbf{P}_1$, and row 0 of $\mathbf{Q}$ is row 0 of $\mathbf{P}_1$ since $\mathbf{P}_i$ is an $(i-1) \times (i-1)$ identity matrix in the top left-hand corner. We now must find $\overline{\mathbf{Q}}$ satisfying (12) whose row 0 is that of $\mathbf{P}_1$.

The Householder matrix $\mathbf{P}_1$ is determined by column 0 of $\mathbf{M}$. Since $\mathbf{A}_s$ is upper Hessenberg, equation (7) shows that this column has the form $[p_1, q_1, r_1, 0, ..., 0]^T$, where

$$\begin{aligned} p_1 &= a_{00}^2 - a_{00}(k_s + k_{s+1}) + k_s k_{s+1} + a_{01}a_{10} \\ q_1 &= a_{10}(a_{00} + a_{11} - k_s - k_{s+1}) \\ r_1 &= a_{10}a_{21} \end{aligned} \tag{13}$$

Hence

$$\mathbf{P}_1 = 1 - 2\mathbf{w}_1 \cdot \mathbf{w}_1^T \tag{14}$$

where $\mathbf{w}_1$ has only its first 3 elements nonzero (cf. equation 11.3.5). The matrix $\mathbf{P}_1 \cdot \mathbf{A}_s \cdot \mathbf{P}_1^T$ is therefore upper Hessenberg with 3 extra elements:

$$\mathbf{P}_1 \cdot \mathbf{A}_1 \cdot \mathbf{P}_1^T = \begin{bmatrix} \times & \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times & \times \\ \mathbf{x} & \times & \times & \times & \times & \times & \times \\ \mathbf{x} & \mathbf{x} & \times & \times & \times & \times & \times \\ & & & \times & \times & \times & \times \\ & & & & \times & \times & \times \\ & & & & & \times & \times \end{bmatrix} \tag{15}$$

This matrix can be restored to upper Hessenberg form without affecting the first row by a sequence of Householder similarity transformations. The first such Householder matrix, $\mathbf{P}_2$, acts on elements 1, 2, and 3 in column 0, annihilating elements 2 and 3. This produces a matrix of the same form as (15), with the 3 extra elements appearing one column over:

$$\begin{bmatrix} \times & \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times & \times \\ & \times & \times & \times & \times & \times & \times \\ & \mathbf{x} & \times & \times & \times & \times & \times \\ & \mathbf{x} & \mathbf{x} & \times & \times & \times & \times \\ & & & & \times & \times & \times \\ & & & & & \times & \times \end{bmatrix} \tag{16}$$

Proceeding in this way up to $\mathbf{P}_{n-1}$, we see that at each stage the Householder matrix $\mathbf{P}_r$ has a vector $\mathbf{w}_r$ that is nonzero only in elements $r-1$, r, and $r+1$. These elements are determined by the elements $r-1$, r, and $r+1$ in column $r-2$ of the current matrix. Note that the preliminary matrix $\mathbf{P}_1$ has the same structure as $\mathbf{P}_2, \ldots, \mathbf{P}_{n-1}$.

The result is that

$$\mathbf{P}_{n-1} \cdots \mathbf{P}_2 \cdot \mathbf{P}_1 \cdot \mathbf{A}_s \cdot \mathbf{P}_1^T \cdot \mathbf{P}_2^T \cdots \mathbf{P}_{n-1}^T = \mathbf{H} \tag{17}$$

where $\mathbf{H}$ is upper Hessenberg. Thus

$$\bar{\mathbf{Q}} = \mathbf{Q} = \mathbf{P}_{n-1} \cdots \mathbf{P}_2 \cdot \mathbf{P}_1 \tag{18}$$

and

$$\mathbf{A}_{s+2} = \mathbf{H} \tag{19}$$

The shifts of origin at each stage are taken to be the eigenvalues of the 2×2 matrix in the bottom right-hand corner of the current $\mathbf{A}_s$. This gives

$$\begin{aligned} k_s + k_{s+1} &= a_{n-2,n-2} + a_{n-1,n-1} \\ k_s k_{s+1} &= a_{n-2,n-2} a_{n-1,n-1} - a_{n-2,n-1} a_{n-1,n-2} \end{aligned} \tag{20}$$

Substituting (20) in (13), we get

$$\begin{aligned} p_1 &= a_{10} \left\{ [(a_{n-1,n-1} - a_{00})(a_{n-2,n-2} - a_{00}) - a_{n-2,n-1} a_{n-1,n-2}]/a_{10} + a_{01} \right\} \\ q_1 &= a_{10}[a_{11} - a_{00} - (a_{n-1,n-1} - a_{00}) - (a_{n-2,n-2} - a_{00})] \\ r_1 &= a_{10} a_{21} \end{aligned} \tag{21}$$

We have judiciously grouped terms to reduce possible roundoff when there are small off-diagonal elements. Since only the ratios of elements are relevant for a Householder transformation, we can omit the factor a_{10} from (21).

In summary, to carry out a double QR step we construct the Householder matrices $\mathbf{P}_r$, $r = 1, \ldots, n-1$. For $\mathbf{P}_1$ we use p_1, q_1, and r_1 given by (21). For the remaining matrices, p_r, q_r, and r_r are determined by the $(r-1, r-2)$, $(r, r-2)$, and $(r+1, r-2)$ elements of the current matrix. The number of arithmetic operations can be reduced by writing the nonzero elements of the $2\mathbf{w} \cdot \mathbf{w}^T$ part of the Householder matrix in the form

$$2\mathbf{w} \cdot \mathbf{w}^T = \begin{bmatrix} (p \pm s)/(\pm s) \\ q/(\pm s) \\ r/(\pm s) \end{bmatrix} \cdot \begin{bmatrix} 1 & q/(p \pm s) & r/(p \pm s) \end{bmatrix} \tag{22}$$

where

$$s^2 = p^2 + q^2 + r^2 \tag{23}$$

(We have simply divided each element by a piece of the normalizing factor; cf. the equations in §11.3.)

If we proceed in this way, convergence is usually very fast. There are two possible ways of terminating the iteration for an eigenvalue. First, if $a_{n-1,n-2}$ becomes "negligible," then $a_{n-1,n-1}$ is an eigenvalue. We can then delete row and column $n-1$ of the matrix and look for the next eigenvalue. Alternatively, $a_{n-2,n-3}$ may become negligible. In this case the eigenvalues of the 2×2 matrix in the lower right-hand corner may be taken to be eigenvalues. We delete rows and columns $n-1$ and $n-2$ of the matrix and continue.

The test for convergence to an eigenvalue is combined with a test for negligible subdiagonal elements that allows splitting of the matrix into submatrices. We find the largest i such that $a_{i,i-1}$ is negligible. If $i = n-1$, we have found a single eigenvalue. If $i = n-2$, we have found two eigenvalues. Otherwise we continue the iteration on the submatrix in rows i to $n-1$ (i being set to zero if there is no small subdiagonal element).

After determining i, the submatrix in rows i to $n-1$ is examined to see if the *product* of any two consecutive subdiagonal elements is small enough that we can work with an even smaller submatrix, starting say in row m. We start with $m = n-3$ and decrement it down to $i+1$, computing p, q, and r according to equations (21) with 0 replaced by m and 1 by $m+1$. If these were indeed the elements of the special "first" Householder matrix in a double QR step, then applying the Householder matrix would lead to nonzero elements in positions $(m+1, m-1)$, $(m+2, m-1)$, and $(m+2, m)$. We require that the first two of these elements be small compared with the local diagonal elements $a_{m-1,m-1}$, a_{mm} and $a_{m+1,m+1}$. A satisfactory approximate criterion is

$$|a_{m,m-1}|(|q| + |r|) \ll |p|(|a_{m+1,m+1}| + |a_{mm}| + |a_{m-1,m-1}|) \tag{24}$$

Very rarely, the procedure described so far will fail to converge. On such matrices, experience shows that if one double step is performed with any shifts that are of order the norm of the matrix, convergence is subsequently very rapid. Accordingly, if ten iterations occur without determining an eigenvalue, the usual shifts are replaced for the next iteration by shifts defined by

$$\begin{aligned} k_s + k_{s+1} &= 1.5 \times (|a_{n-1,n-2}| + |a_{n-2,n-3}|) \\ k_s k_{s+1} &= (|a_{n-1,n-2}| + |a_{n-2,n-3}|)^2 \end{aligned} \tag{25}$$

The factor 1.5 was arbitrarily chosen to lessen the likelihood of an "unfortunate" choice of shifts. This strategy is repeated after 20 unsuccessful iterations. After 30 unsuccessful iterations, the routine reports failure.