6. QR algorithm

We next study the recommended algorithm for finding all the eigenvalues and eigenvectors of a general real nonsymmetric matrix. The basic algorithm is fairly simple. Starting with $A_0 = A$, we define the iteration:

$$A_i = Q_i R_i, \qquad A_{i+1} = R_i Q_i,$$

where Q_i is an orthogonal matrix and R_i is upper triangular. Note that $R_i = Q_i^{-1} A_i$, so $A_{i+1} = Q_i^{-1} A_i Q_i$. Since A_{i+1} is obtained from A_i by a similarity transformation, these matrices have the same eigenvalues. The idea of the method is that the sequence $\{A_i\}$ converges to a matrix from which the eigenvalues can be easily determined. For example, if the eigenvalues are real and distinct, the sequence converges to an upper triangular matrix with the eigenvalues on the diagonal. More generally, we get convergence to a quasi-triangular matrix, from which the eigenvalues are also easily determined.

To use this algorithm, we require that A_i have a QR factorization. This in ensured by the following result.

Theorem 21. *If A is real and nonsingular, then there exists a decomposition $A = QR$ for which Q is orthogonal and R is upper triangular. Furthermore, if the diagonal elements r_{ii} of R are positive, then the decomposition is unique.*

Remark: In general, there can be nonuniqueness of the factorization, since if $A = QR$, then also $A = [QD][D^{-1}R]$, where D is any diagonal matrix with entries of 1 or -1 on the diagonal. Note that QD will still be orthogonal and $D^{-1}R$ upper triangular.

Because of the number of arithmetic operations involved, the QR algorithm is practical only when applied to a matrix in upper Hessenberg form (i.e., $a_{ij} = 0$ for $i > j + 1$). Hence the method consists of two steps. In the first step, we find a similar matrix which is in upper Hessenberg form. The second step is then to apply the QR algorithm to this new matrix. One can show that if A is in upper Hessenberg form, applying the QR algorithm results in a sequence of matrices that are also in this form.

One way of reducing a matrix to upper Hessenberg form is to apply Householder transformations. When A is symmetric, this reduces A to tridiagonal form. For nonsymmetric A, there is another method that is more efficient, using stabilized (elementary) transformations. However, since the factorization of $A = QR$ is also done by Householder transformations, we shall use these transformations for both steps.

6.1. **Reduction to Hessenberg form using Householder transformations.** The idea is to find a sequence of orthogonal matrices P_k such that if $A_k = P_k^T A_{k-1} P_k$, $k = 1, 2, \ldots, n-2$, with $A_0 = A$, then A_{n-2} is upper Hessenberg. We assume that A is a real $n \times n$ matrix. More specifically, we find at the rth step, a matrix P_r that introduces zeroes in the rth column (below the subdiagonal), without affecting zeroes in the previous columns (below the subdiagonal). We illustrate this process in the case $n = 7$, $r = 4$. The configuration immediately before the rth step (in which A_r is computed from A_{r-1}) is given by:

$$A_{r-1} = \left(\begin{array}{cccc|ccc} \times & \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times & \times \\ \hline 0 & 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times & \times \end{array}\right) = \left(\begin{array}{cc|c} \multicolumn{2}{c|}{C_{r-1}} & \begin{array}{c} D_{r-1} \\ \hline d_{r-1}^T \end{array} \\ \hline 0 & b_{r-1} & B_{r-1} \end{array}\right),$$

where b_{r-1} and d_{r-1} are vectors with $n-r$ components, B_{r-1} is a matrix of order $n-r$, C_{r-1} is a matrix of order r, and D_{r-1} is a $r-1 \times n-r$ matrix. We note that the initial configuration is given by

$$A = A_0 = \left(\begin{array}{c|c} c_0 & d_0^T \\ \hline b_0 & B_0 \end{array}\right).$$

The matrix P_r may be expressed in the form

$$P_r = \left(\begin{array}{c|c} I & 0 \\ \hline 0 & Q_r \end{array}\right) = \left(\begin{array}{c|c} I & 0 \\ \hline 0 & I - 2v_r v_r^T \end{array}\right),$$

where Q_r is a matrix of order $n-r$ and v_r is a unit vector having $n-r$ components.

The motivation for choosing P_r of this form is that, as we shall see, $P_r^T A_{r-1} P_r$ leaves the upper $r \times r$ block of A_{r-1} unchanged (which is already upper Hessenberg), and is easily seen to be an orthogonal matrix. Note that

$$P_r P_r^T = P_r P_r = \begin{pmatrix} I & 0 \\ 0 & Q_r \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & Q_r \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & Q_r Q_r \end{pmatrix}$$

and, since $v^T v = 1$,

$$Q_r Q_r = (I - 2vv^T)(I - 2vv^T) = I - 4vv^T + 4vv^T vv^T = I.$$

Hence we have $A_r = P_r^T A_{r-1} P_r = P_r A_{r-1} P_r =$

$$
\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & Q_r \end{array}\right)
\left(\begin{array}{cc|c} \multicolumn{2}{c|}{C_{r-1}} & \begin{array}{c} D_{r-1} \\ \hline d_{r-1}^T \end{array} \\ \hline 0 & \multicolumn{1}{|c|}{b_{r-1}} & B_{r-1} \end{array}\right)
\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & Q_r \end{array}\right)
$$

$$
= \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & Q_r \end{array}\right)
\left(\begin{array}{cc|c} \multicolumn{2}{c|}{C_{r-1}} & \begin{array}{c} D_{r-1}Q_r \\ \hline d_{r-1}^T Q_r \end{array} \\ \hline 0 & \multicolumn{1}{|c|}{b_{r-1}} & B_{r-1}Q_r \end{array}\right)
$$

$$
= \left(\begin{array}{cc|c} \multicolumn{2}{c|}{C_{r-1}} & \begin{array}{c} D_{r-1}Q_r \\ \hline d_{r-1}^T Q_r \end{array} \\ \hline 0 & \multicolumn{1}{|c|}{Q_r b_{r-1}} & Q_r B_{r-1} Q_r \end{array}\right).
$$

If we now choose v_r so that $Q_r b_{r-1}$ is zero except for its first component, then A_r will be upper Hessenberg in its first r columns. To avoid problems with subscripts in the construction of v, we write the matrix

$$P_r = I - 2w_r w_r^T = I - u_r u_r^T/(2K_r^2).$$

Letting $a_{ij}^{(r-1)}$ denote the ijth element of A_{r-1} and setting

$$S_r = \Big(\sum_{i=r+1}^{n} [a_{ir}^{(r-1)}]^2\Big)^{1/2} \operatorname{sgn}(a_{r+1,r}^{(r-1)}),$$

we then choose

$$(u_r)_i = 0, \quad i = 1, \dots, r, \qquad (u_r)_{r+1} = a_{r+1,r}^{(r-1)} + S_r, \qquad (u_r)_i = a_{ir}^{(r-1)}, \quad i = r+2, \dots, n,$$

where $2K_r^2 = S_r^2 + a_{r+1,r}^{(r-1)} S_r$. Computations then show that with these choices, the last $n-r-1$ components of $Q_r b_{r-1} = 0$ and v_r is a unit vector.

At the end of this process, we get the Hessenberg matrix

$$H = P_{n-2} P_{n-1} \cdots P_1 A P_1 P_2 \cdots P_{n-2},$$

which is similar to A.

Remark: If A is symmetric, then since each P_r is symmetric, H will be symmetric and hence tridiagonal.

Remark: One can derive efficient formulas for computing H without performing all the matrix multiplications indicated by the above formula. See the references for details.

6.2. **Implementation of the QR algorithm.** We now assume that the matrix A is in upper Hessenberg form and consider the implementation of the QR algorithm. Recall, we set $A_0 = A$ and then (i) factor $A_i = Q_i R_i$ with Q_i orthogonal and R_i upper triangular and (ii) set $A_{i+1} = R_i Q_i$.

We now show how step (i) can be done using Householder transformations. We set $M_0 = A_i$. We then construct a sequence $\{M_r\}$, $r = 1, 2, \ldots, n-1$, where the matrix M_r will be upper triangular in its first r columns. Thus, there are $n-1$ steps, and just before the rth step, the matrix M_{r-1} (upper triangular in its first $r-1$ columns) will have the form

$$M_{r-1} = \begin{pmatrix} U_{r-1} & | & V_{r-1} \\ - & | & - \\ 0 & | & W_{r-1} \end{pmatrix},$$

where U_{r-1} is an $(r-1)\times(r-1)$ upper triangular matrix. We then set $M_r = P_r M_{r-1}$, where P_r has the form

$$P_r = \begin{pmatrix} I & | & 0 \\ - & | & -- \\ 0 & | & I - 2vv^T \end{pmatrix},$$

where the identity matrix in the upper left corner is of order $r-1$ and $v^T v = 1$. Then

$$M_r = P_r M_{r-1} = \begin{pmatrix} U_{r-1} & | & V_{r-1} \\ - & | & --- \\ 0 & | & (I - 2vv^T)W_{r-1} \end{pmatrix}.$$

We then choose v so that the first column of $(I-2vv^T)W_{r-1}$ is zero, except for its first element. This is essentially what we did previously, except now the first column of $(I - 2vv^T)W_{r-1}$ has $n - r + 1$ components (instead of $n - r$ components).

Let $M_{ij}^{(r-1)}$ denote the i, jth entry of M_{r-1}. Set

$$S_r = \left(\sum_{i=r}^{n} [M_{ir}^{(r-1)}]^2\right)^{1/2} \operatorname{sgn} M_{r,r}^{(r-1)}, \qquad 2K_r^2 = S_r^2 + a_{r,r}^{(r-1)} S_r.$$

To avoid problems with subscripts, we again write

$$P_r = I - 2w_r w_r^T = I - \frac{1}{2K_r^2} u_r r_r^T, \quad \text{where now}$$

$$(u_r)_i = 0,\ i = 1, 2, \ldots, r-1, \qquad (u_r)_r = M_{r,r}^{(r-1)} + S_r, \qquad (u_r)_i = M_{i,r}^{(r-1)},\ i = r+1, \ldots n.$$

Then $M_{n-1} = P_{n-1}P_{n-2}\cdots P_1 M_0$ is upper triangular. Since the P_i are orthogonal, defining $Q^{-1} = P_{n-1}P_{n-2}\cdots P_1$, we see that Q^{-1} is orthogonal and hence $Q = (P_{n-1}P_{n-2}\cdots P_1)^{-1} = (P_{n-1}P_{n-2}\cdots P_1)^T = P_1 P_2 \cdots P_{n-1}$ is orthogonal. Then, taking $R_i = M_{n-1}$ and $Q_i = P_1 P_2 \cdots P_{n-1}$, we get the decomposition $A_i = Q_i R_i$, with Q_i orthogonal and R_i upper triangular.