6.3. **Convergence of the QR algorithm.** We now state a convergence result for the QR algorithm in the special case that there is only one eigenvalue of a given modulus.

Theorem 22. *Suppose that A has eigenvalues $\lambda_1, \ldots, \lambda_n$ satisfying $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n|$. Let X denote the matrix whose ith column is a eigenvector of A corresponding to λ_i and suppose that X^{-1} has an LU decomposition. Then the subdiagonal elements of the matrices A_s of the basic QR algorithm tend to zero and for $i = 1, 2, \ldots, n$, $(A_s)_{ii} \to \lambda_i$.*

Note that since we are assuming that the eigenvalues are distinct, we know there are a complete set of linearly independent eigenvectors and hence X^{-1} exists.

Without giving a formal proof, let us try to understand what makes this algorithm work. Recall that the QR algorithm sets $A_1 = A$, factors $A_i = Q_i R_i$ and then sets $A_{i+1} = R_i Q_i$. Since $Q_i^{-1} = Q_i^T$, we have $R_i = Q_i^T A_i$ and so

$$A_{i+1} = Q_i^T A_i Q_i.$$

Thus, we have performed a similarity transformation, which preserves the eigenvalues. If we iterate the above equation, we get

$$A_{i+1} = Q_i^T A_i Q_i = Q_i^T Q_{i-1}^T A_{i-1} Q_{i-1} Q_i = \cdots Q_i^T \cdots Q_1^T A_1 Q_1 \cdots Q_i = P_i^T A_1 P_i,$$

where $P_i = Q_1 \cdots Q_i$. Note that P_i is the product of orthogonal matrices and hence is orthogonal.

Now under our assumptions, we can write $A = XDX^{-1}$ where $D = diag(\lambda_1, \ldots, \lambda_n)$ and X is a real matrix of eigenvectors of A. We know there is a factorization of $X = QR$, where Q is orthogonal and R is upper triangular. Then

$$A = QRDR^{-1}Q^{-1} \quad \text{and so} \quad Q^{-1}AQ = RDR^{-1}.$$

Since RDR^{-1} is the product of upper triangular matrices, it is also upper triangular, and hence we know that $Q^{-1}AQ = Q^T AQ$ is upper triangular. Note that the eigenvalues of A will then lie of the diagonal of $Q^T AQ$.

The theorem is proved by showing that $\lim_{i\to\infty} P_i = Q$, since this implies that

$$\lim_{i\to\infty} A_{i+1} = \lim_{i\to\infty} P_i^T A_1 P_i = Q^T AQ.$$

In other words, the matrices A_i are converging to an upper triangular matrix whose diagonal elements are the eigenvalues of A.

To see why $\lim_{i\to\infty} P_i = Q$, we look at the quantity $P_i U_i$, where $U_i = R_i R_{i-1} \cdots R_1$. Then

$$P_i U_i = Q_1 \cdots Q_i R_i \cdots R_1 = Q_1 \cdots Q_{i-1} A_i R_{i-1} \cdots R_1 = P_{i-1} A_i U_{i-1}.$$

But since $A_{i+1} = P_i^T A_1 P_i$, we have $P_i A_{i+1} = A_1 P_i$, or after reducing the indices by one, $P_{i-1} A_i = A_1 P_{i-1}$. Hence,

$$P_i U_i = A_1 P_{i-1} U_{i-1}.$$

If we iterate this identity, we get

$$P_i U_i = (A_1)^{i-1} P_1 U_1 = (A_1)^{i-1} Q_1 R_1 = A^i.$$

Using the fact that $A = XDX^{-1}$, we also have that $A^i = XD^iX^{-1}$. We know that $X = QR$ and by hypothesis $X^{-1} = LU$. Hence,

$$A^i = QRD^iLU = QR(D^iLD^{-i})D^iU.$$

Equating these two expressions for A^i, we get that

$$P_iU_i = QR(D^iLD^{-i})D^iU.$$

The key step in the proof is to show that

$$\lim_{i\to\infty} D^iLD^{-i} = I.$$

Assuming for the moment that this is true, the right hand side becomes QRD^iU. But RD^iU is the product of upper triangular matrices and is therefore upper triangular. Hence, we are essentially able to identify $\lim_{i\to\infty} P_i$ with Q, since both are orthogonal matrices, and $\lim_{i\to\infty} U_i$ with $\lim_{i\to\infty} RD^iU$ since both quantities are upper triangular matrices. We have ignored a subtle point; namely that the QR factorization is only unique if R is chosen to have positive diagonal entries. So we must choose all the decompositions to insure this property so that we can identify the individual pieces of the decomposition. Returning to the key step, we observe that the matrix D^iLD^{-i} is a lower triangular matrix whose j,kth element is given by $l_{jk}(\lambda_j/\lambda_k)^i$, when $j > k$. Since $|\lambda_j/\lambda_k| < 1$ for $j > k$,

$$\lim_{i\to\infty} D^iLD^{-i} = I.$$

So, the convergence is very similar to that of the power method.

Using the similarity to the power method, we expect to be able to improve the convergence by changing the method to correspond to the inverse power method. Thus, we consider the QR method with origin shift, described by the algorithm: Let s_i be any number. (i) Factor $A_i - s_iI = Q_iR_i$ and then (ii) set $A_{i+1} = R_iQ_i + s_iI$. Note that

$$A_{i+1} = Q_i^TQ_iR_iQ_i + Q_i^Ts_iQ_i = Q_i^T(A_i - s_iI)Q_i + Q_i^Ts_iQ_i = Q_i^TA_iQ_i.$$

Hence A_{i+1} is similar to A_i.

Since we expect that $(A_i)_{nn}$ is converging to an eigenvalue of A, we can choose $s_i = (A_i)_{nn}$. We will then get an iteration similar to the Rayleigh quotient iteration.

Remark: In general, the method produces a sequence of matrices that do not converge to a triangular matrix, but rather to a block triangular matrix (for example if we have complex conjugate eigenvalues), from which we can also determine the eigenvalues. In this case, a more complicated shift strategy is used.