Section 1.5

As A has n distinct eigenvalues and is a square matrix, A is diagonalizable, i.e. there exists an invertible matrix X such that

$$X^{-1}AX = D = \text{diag}[\lambda_1, \dots, \lambda_n]$$

Then

$$A_k = XD_kX^{-1} \qquad (1)$$

Assume the X^{-1} has the following decomposition

$$X^{-1} = LU \qquad (2)$$

Combining (1) and (2) gives

$$A_k = X(D^k L D^{-k}) D_k U$$

If we choose the diagonal elements of L to be 1, the matrix $D^k L D^{-k}$ is lower triangular with diagonal elements equal to 1, and so we have

$$(D^k L D^{-k})_{ij} = \left[\frac{\lambda_i}{\lambda_j}\right]^k L_{ij} \qquad 1 \leq j < i \leq n \qquad (3)$$

Define E_k implicitly by

$$D^k L D^{-k} = I + E_k$$

E_k is a lower triangular matrix which converges to zero, using

(2) and (3):

$$\|E_k\|_\infty \leq C \cdot \underset{1 \leq j \leq n-1}{\text{Maximum}} \left| \frac{\lambda_{j+1}}{\lambda_j} \right|^k , \quad m \geq 1 \qquad (4)$$

for some constant $C > 0$.

We can use the QR factorization on X to achieve

$$X = QR$$

where Q is some orthogonal matrix and R is an invertible upper triangular matrix, which gives

$$\begin{aligned} A^m &= QR(I + E_k) D^k U \\ &= Q(I + RE_k R^{-1}) R D^k U \end{aligned} \qquad (5)$$

Using another QR factorization:

$$I + RE_k R^{-1} = \tilde{Q}_k \tilde{R}_k \qquad (6)$$

We assume the diagonal elements of R_k to be positive, and with this assumption, (6) is unique.

Next, we show that $\tilde{Q}_k, \tilde{R}_k \to I$ as $k \to \infty$. Using (6) and (4),

$$\tilde{R}^T_k \tilde{R}_k - I \to 0 \quad \text{as } k \to \infty$$

The coefficients of $\tilde{R}_k^T \tilde{R}_k$ will show that $\tilde{R}_k \to I$ using the positivity of the diagonal elements. Using (6) in (5)

$$A^m = (Q\tilde{Q}_k)(\tilde{R}_k R D^k U) \qquad (7)$$

It's clear that $Q\tilde{Q}_k$ is orthogonal. Since $\tilde{R}_k, R, U$ are upper triangular and D^k is diagonal, their product is upper triangular. Thus (7) is a QR factorization of A^k. We also know that A^k has the QR factorization

$$A_k = P_k U_k$$

where $P_k = Q_1 \dots Q_k$ is orthogonal and $U_k = R_k \dots R_k$ is upper triangular. Using the uniqueness of the QR factorization, we have

$$P_k = (Q\tilde{Q}_k)\tilde{D}_k \quad , \quad \mathcal{U} = \tilde{D}_k (\tilde{R}_k R D^k U) \qquad (8)$$

for some diagonal matrix $\tilde{D}_k$ with

$$\tilde{D}_k^2 = I \; , \quad k \geq 1$$

Now, we look at the behavior of the sequence $\{A_k\}$ as $k \to \infty$

We know from section 1.3 that A is similar to A_k, and so by transitivity, A_k is similar to A_{k+1}. Hence we have

$$A_{k+1} = P_k^T A_1 P_k \qquad k \geq 1$$

where the matrix P_m is orthogonal and U_m is upper triangular.

Using this with (8) gives:

$$A_{k+1} = \tilde{D}_k \tilde{Q}_k^T Q^T A Q \tilde{Q}_k \tilde{D}_k$$

From $X = QR$ and

$$Q = X R^{-1}$$

$$Q^T = Q^{-1} = R X^{-1}$$

Substituting:

$$A_{k+1} = \tilde{D}_k^T \tilde{Q}_k^T R X^{-1} A X R^{-1} \tilde{Q}_k \tilde{D}_k$$

$$= \tilde{D}_K^T \tilde{Q}_K^T RDR^{-1} \tilde{Q}_K \tilde{D}_K$$

The matrix RDR^{-1} is upper triangular and its diagonal elements are $\{\lambda_1, \dots \lambda_n\}$. Using the fact that $\tilde{Q}_K \to I$ and $\tilde{D}_K^2 = I$, we will have the diagonal elements of A_{K+1} converge to the eigenvalues of A, ordered from larges to smallest in magnitude. Moreone, since RDR^{-1} is upper triangular, the elements below the diagonal in A_{m+1} will converge to 0. This completes the proof.