(1) Show that Jacobi's method for finding eigenvalues of a real symmetric matrix is ultimately quadratically convergent. Assume that all off-diagonal elements of the matrix A_k are $O(\varepsilon)$,where k enumerates Jacobi sweeps. Show that the all rotations of the next Jacobi sweep are of the form

$$J(i,j) = \begin{array}{cc} \begin{pmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \cdots & 1+O(\varepsilon^2) & \cdots & O(\varepsilon) & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \cdots & O(\varepsilon) & \cdots & 1+O(\varepsilon^2) & \cdots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \end{pmatrix} & \begin{array}{c} \\ \\ i \\ \\ j \\ \\ \\ \end{array} \\ \begin{array}{ccccccc} & & i & & j & & \end{array} & \end{array}$$

Then demonstrate that this implies that, after the sweep, all off-diagonal elements of A_{k+1} are $O(\varepsilon^2)$. Assume that all eigenvalues are non-zero and distinct.

First, let's assume our matrix A has off diagonal entries as $O(\varepsilon)$, that is

$$A = \begin{pmatrix} a_1 & O(\varepsilon) & \cdots & O(\varepsilon) \\ O(\varepsilon) & a_2 & \cdots & O(\varepsilon) \\ \vdots & \vdots & \ddots & \vdots \\ O(\varepsilon) & O(\varepsilon) & \cdots & a_n \end{pmatrix}.$$

Then, for the components of a Jacobi rotation for A at (i,j) can be computed as

$$\tau = \frac{a_j - a_i}{2a_{ij}} = \frac{a_j - a_i}{O(\varepsilon)} = O\left(\frac{1}{\varepsilon}\right)$$

which implies

$$\begin{aligned} \theta &= \arctan\left(\frac{1}{\tau \pm \sqrt{1+\tau}}\right) \\ &= \arctan\left(\frac{1}{O\left(\frac{1}{\varepsilon}\right) \pm \sqrt{1 + O\left(\frac{1}{\varepsilon}\right)}}\right) \\ &= \arctan\left(\frac{1}{O\left(\frac{1}{\varepsilon}\right)}\right) \\ &= \arctan(O(\varepsilon)) \\ &= O(\varepsilon) \end{aligned}$$

and so

$$\begin{aligned} \cos(\theta) &= 1 + \frac{1}{2}\theta^2 + \cdots = 1 + O(\varepsilon^2) \\ \sin(\theta) &= \theta + \cdots = O(\varepsilon). \end{aligned}$$

Then, the Jacobi rotation is given as

$$J(i,j) = \begin{pmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \cdots & \cos(\theta) & \cdots & \sin(\theta) & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \cdots & -\sin(\theta) & \cdots & \cos(\theta) & \cdots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \cdots & 1 + O(\varepsilon^2) & \cdots & O(\varepsilon) & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \cdots & O(\varepsilon) & \cdots & 1 + O(\varepsilon^2) & \cdots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \end{pmatrix}.$$

Next, we can perform a cyclic row sweep of A as

$$A^{(1)} = J(1,2)^* A J(1,2) = \begin{pmatrix} a_1 + O(\varepsilon^2) & 0 & O(\varepsilon) & \cdots & O(\varepsilon) & O(\varepsilon) \\ 0 & a_2 + O(\varepsilon^2) & O(\varepsilon) & \cdots & O(\varepsilon) & O(\varepsilon) \\ O(\varepsilon) & O(\varepsilon) & a_3 & \cdots & O(\varepsilon) & O(\varepsilon) \\ \vdots & \vdots & \vdots & \ddots & & \vdots \\ O(\varepsilon) & O(\varepsilon) & O(\varepsilon) & & a_{n-1} & O(\varepsilon) \\ O(\varepsilon) & O(\varepsilon) & O(\varepsilon) & \cdots & O(\varepsilon) & a_n \end{pmatrix},$$

$$A^{(2)} = J(1,3)^* A^{(1)} J(1,3) = \begin{pmatrix} a_1 + O(\varepsilon^2) & O(\varepsilon^2) & 0 & \cdots & O(\varepsilon) & O(\varepsilon) \\ O(\varepsilon^2) & a_2 + O(\varepsilon^2) & O(\varepsilon) & \cdots & O(\varepsilon) & O(\varepsilon) \\ 0 & O(\varepsilon) & a_3 + O(\varepsilon^2) & \cdots & O(\varepsilon) & O(\varepsilon) \\ \vdots & \vdots & \vdots & \ddots & & \vdots \\ O(\varepsilon) & O(\varepsilon) & O(\varepsilon) & & a_{n-1} & O(\varepsilon) \\ O(\varepsilon) & O(\varepsilon) & O(\varepsilon) & \cdots & O(\varepsilon) & a_n \end{pmatrix},$$

and continuing to the end of the row yields

$$A^{(k-1)} = \begin{pmatrix} a_1 + O(\varepsilon^2) & O(\varepsilon^2) & O(\varepsilon^2) & \cdots & O(\varepsilon^2) & 0 \\ O(\varepsilon^2) & a_2 + O(\varepsilon^2) & O(\varepsilon) & \cdots & O(\varepsilon) & O(\varepsilon) \\ O(\varepsilon^2) & O(\varepsilon) & a_3 + O(\varepsilon^2) & \cdots & O(\varepsilon) & O(\varepsilon) \\ \vdots & \vdots & \vdots & \ddots & & \vdots \\ O(\varepsilon^2) & O(\varepsilon) & O(\varepsilon) & & a_{n-1} + O(\varepsilon^2) & O(\varepsilon) \\ 0 & O(\varepsilon) & O(\varepsilon) & \cdots & O(\varepsilon) & a_n + O(\varepsilon^2) \end{pmatrix}$$

and finally finishing the sweep yields

$$A^{(\text{sweep})} = \begin{pmatrix} a_1 + O(\varepsilon^2) & O(\varepsilon^2) & O(\varepsilon^2) & \cdots & O(\varepsilon^2) & O(\varepsilon^3) \\ O(\varepsilon^2) & a_2 + O(\varepsilon^2) & O(\varepsilon^2) & \cdots & O(\varepsilon^2) & O(\varepsilon^3) \\ O(\varepsilon^2) & O(\varepsilon^2) & a_3 + O(\varepsilon^2) & \cdots & O(\varepsilon^2) & O(\varepsilon^3) \\ \vdots & \vdots & \vdots & \ddots & & \vdots \\ O(\varepsilon^2) & O(\varepsilon^2) & O(\varepsilon^2) & & a_{n-1} + O(\varepsilon^2) & 0 \\ O(\varepsilon^3) & O(\varepsilon^3) & O(\varepsilon^3) & \cdots & 0 & a_n + O(\varepsilon^2) \end{pmatrix}.$$

Thus, we can see that after a sweep, all of the off diagonal entries are at least $O(\varepsilon^2)$ which means that our matrix is moving towards the diagonal matrix quadratically.

(2) Show that A is diagonalizable iff there is a positive definite self-adjoint matrix H such that $H^{-1}AH$ is normal.

Proof:

($\implies$)Suppose a matrix A is diagonalizable. Then we can write

$$A = PDP^{-1} \implies D = P^{-1}AP$$

where D is a diagonal and P is an invertible. Next, let's take the polar decomposition of P as $H = HU$ where H is positive definite self-adjoint and U is unitary. Note, H^{-1} is also positive definite-self-adjoint and $H = PU^*$. Then

$$H^{-1}AH = UP^{-1}APU^* = UDU^*.$$

Then

$$\begin{aligned}(H^*AH)^*(H^*AH) &= (UDU^*)^*(UDU^*)\\ &= (UD^*U^*)(UDU^*)\\ &= UD^*DU^*\\ &= UDD^*U^*\\ &= (UDU^*)(UD^*U^*)\\ &= (UDU^*)(UDU^*)^*\\ &= (H^{-1}AH)(H^{-1}AH)^*\end{aligned}$$

Showing that there exists a positive definite self-adjoint matrix H such that $H^{-1}AH$ is normal.

($\impliedby$) Now, suppose there exists a positive definite self-adjoint matrix H such that $H^{-1}AH$ is normal. Then, because $H^{-1}AH$ is normal, it is diagonalizable by a unitary matrix:

$$H^{-1}AH = UDU^*$$

where D is diagonal and U is unitary. Then, we have

$$D = U^*H^{-1}AHU = P^{-1}AP$$

where $P = HU$. So, A is similar to a diagonal matrix and is thus diagonalizable. □