Diagonalization by a unitary similarity transformation

In these notes, we will always assume that the vector space V is a complex n-dimensional space.

1. Introduction

A semi-simple matrix A is an $n\times n$ matrix that possesses n linearly independent eigenvectors. (If the corresponding eigenvalues are non-degenerate, then we say that the matrix is simple.) We may therefore use the eigenvalues to construct a basis $\mathcal{B}$. A simple computation[1] shows that the matrix $D \equiv [A]_{\mathcal{B}}$, whose matrix elements are given with respect to the basis $\mathcal{B}$ is diagonal, with

$$P^{-1}AP = D\,, \tag{1}$$

where P is a matrix whose columns are the eigenvectors of A, and D is a diagonal matrix whose diagonal elements are the eigenvalues of A. All diagonalizable matrices are semi-simple.

A *normal* matrix A is defined to be a matrix that commutes with its hermitian conjugate. That is,

$$A \text{ is normal} \qquad \Longleftrightarrow \qquad AA^\dagger = A^\dagger A\,.$$

In the present note, we wish to examine a special case of matrix diagonalization in which the diagonalizing matrix P is unitary. In this case, the basis of eigenvectors $\mathcal{B}$ is orthonormal. We demonstrate below that a matrix A is diagonalizable by a unitary similarity transformation if and only if A is normal.

Before proceeding, we record a few facts about unitary and hermitian matrices. A unitary matrix U is a matrix that satisfies $UU^\dagger = U^\dagger U = \mathbf{I}$. By writing out these matrix equations in terms of the matrix elements, one sees that the columns [or rows] of U, treated as vectors, are orthonormal. That is, if the columns of U are denoted by $\widehat{\boldsymbol{e}}_j$, then the inner product[2] is given by $\langle \widehat{\boldsymbol{e}}_i\,,\,\widehat{\boldsymbol{e}}_j\rangle = \delta_{ij}$. In particular, each column is a vector that is normalized to unity. Note that a unitary matrix is also a normal matrix.

An hermitian matrix satisfies $A^\dagger = A$. Clearly, hermitian matrices are also normal. The eigenvalues of hermitian matrices are necessarily real. To prove this, we use the definition of the hermitian adjoint:

$$\langle A\vec{\boldsymbol{v}}\,,\,\vec{\boldsymbol{w}}\rangle = \langle \vec{\boldsymbol{v}}\,,\,A^\dagger\vec{\boldsymbol{w}}\rangle\,.$$

[1] See the handout entitled, *Coordinates, matrix elements and changes of basis.*

[2] The inner product of two vectors can be expressed, in terms of their coordinates with respect to an orthonormal basis, by $\langle \vec{\boldsymbol{v}}\,,\,\vec{\boldsymbol{w}}\rangle = \sum_k v_k^* w_k$, where $\vec{\boldsymbol{v}} = \sum_k v_k \hat{\boldsymbol{e}}_k$ and $\vec{\boldsymbol{w}} = \sum_k w_k \hat{\boldsymbol{e}}_k$. Note the appearance of the complex conjugate, v_k^*, in the expression for the inner product.

Since $A^\dagger = A$, it follows that

$$\langle A\vec{\boldsymbol{v}}\,,\,w\rangle = \langle\vec{\boldsymbol{v}}\,,\,A\vec{\boldsymbol{w}}\rangle \tag{2}$$

must be true for any vectors $\vec{\boldsymbol{v}}$, $\vec{\boldsymbol{w}} \in V$. In particular, if we choose $\vec{\boldsymbol{v}} = \vec{\boldsymbol{w}}$ to be an eigenvector of A; *i.e.*, $A\vec{\boldsymbol{v}} = \lambda\vec{\boldsymbol{v}}$ for $\vec{\boldsymbol{v}} \neq 0$, then

$$\langle\lambda\vec{\boldsymbol{v}}\,,\,\vec{\boldsymbol{v}}\rangle = \langle\vec{\boldsymbol{v}}\,,\,\lambda\vec{\boldsymbol{v}}\rangle\,.$$

Using the properties of the inner product of a complex vector space,

$$\lambda^*\langle\vec{\boldsymbol{v}}\,,\,\vec{\boldsymbol{v}}\rangle = \lambda\langle\vec{\boldsymbol{v}}\,,\,\vec{\boldsymbol{v}}\rangle\,,$$

where λ^* is the complex conjugate of λ. Since $\vec{\boldsymbol{v}} \neq 0$, it follows that $\langle\vec{\boldsymbol{v}}\,,\,\vec{\boldsymbol{v}}\rangle \neq 0$. Thus, we may conclude that $\lambda^* = \lambda$ so that λ is real. One can also prove that eigenvectors corresponding to non-degenerate eigenvalues of A are orthogonal. To prove this, take $\vec{\boldsymbol{v}}$ and $\vec{\boldsymbol{w}}$ to be eigenvectors of A with corresponding eigenvalues λ and λ'. Then, from eq. (2), it follows that

$$\langle\lambda\vec{\boldsymbol{v}}\,,\,\vec{\boldsymbol{w}}\rangle = \langle\vec{\boldsymbol{v}}\,,\,\lambda'\vec{\boldsymbol{w}}\rangle\,.$$

Once again, we make use of the properties of the inner product to conclude that

$$(\lambda^* - \lambda')\langle\vec{\boldsymbol{v}}\,,\,\vec{\boldsymbol{w}}\rangle = 0\,.$$

Since $\lambda^* = \lambda \neq \lambda'$, one can conclude that $\langle\vec{\boldsymbol{v}}\,,\,\vec{\boldsymbol{w}}\rangle = 0$. If $\lambda = \lambda'$, then this argument does not allow us to deduce anything about $\langle\vec{\boldsymbol{v}}\,,\,\vec{\boldsymbol{w}}\rangle$. However, it is possible to show that the eigenvectors corresponding to degenerate eigenvalues can be chosen to be orthogonal. To prove this requires a more powerful technique that does not care about the degeneracy of the eigenvalues. In fact, the statement that the eigenvectors of an hermitian matrix can be chosen to be orthonormal will be one of the consequences of the analysis that follows.

2. The unitary diagonalization of an hermitian matrix

Let A be an hermitian matrix. Consider the eigenvalue problem $A\vec{\boldsymbol{v}} = \lambda\vec{\boldsymbol{v}}$, where $\vec{\boldsymbol{v}} \neq 0$. All matrices possess at least one eigenvector and corresponding eigenvalue. Thus, we we focus on one of the eigenvalues and eigenvectors of A that satisfies $A\vec{\boldsymbol{v}}_1 = \lambda\vec{\boldsymbol{v}}_1$. We can always normalize $\vec{\boldsymbol{v}}_1$ to unity by dividing out by its norm. We now construct a unitary matrix U_1 as follows. Take the first column of U_1 to be given by (the normalized) $\vec{\boldsymbol{v}}_1$. The rest of the unitary matrix will be called Y, which is an $n \times (n-1)$ matrix. Explicitly,

$$U_1 = \left(\begin{array}{c:c} \vec{\boldsymbol{v}}_1 & Y \end{array}\right)\,,$$

where the vertical dashed line is inserted for the reader's convenience as a reminder that this is a *partitioned* matrix that is $n \times 1$ to the left of the dashed line and

$n \times (n-1)$ to the right of the dashed line. Since the columns of U_1 comprise an orthonormal set of vectors, we can write the matrix elements of Y in the form $Y_{ij} = (\vec{\boldsymbol{v}}_j)_i$, for $i = 1, 2, \ldots, n$ and $j = 2, 3, \ldots$, where $\{\vec{\boldsymbol{v}}_1, \vec{\boldsymbol{v}}_2, \ldots, \vec{\boldsymbol{v}}_n\}$ is an orthonormal set of vectors. Here $(\vec{\boldsymbol{v}}_j)_i$ is the ith coordinate (with respect to a fixed orthonormal basis) of the jth vector of the orthonormal set. It then follows that:

$$\langle \vec{\boldsymbol{v}}_j \, , \, \vec{\boldsymbol{v}}_1 \rangle = \sum_{k=1}^{n} (\vec{\boldsymbol{v}}_j)^*_k (\vec{\boldsymbol{v}}_1)_k = 0 \, , \qquad \text{for} \qquad j = 2 \, , \, 3 \, , \, \ldots \, , \, n \, , \tag{3}$$

where $(\vec{\boldsymbol{v}}_j)^*_k$ is the complex conjugate of the kth component of the vector $\vec{\boldsymbol{v}}_j$. We can rewrite eq. (3) as a matrix product (where $\vec{\boldsymbol{v}}_1$ is an $n \times 1$ matrix) as:

$$Y^\dagger \vec{\boldsymbol{v}}_1 = \sum_{k=1}^{n} Y^*_{kj} (\vec{\boldsymbol{v}}_1)_k = \sum_{k=1}^{n} (\vec{\boldsymbol{v}}_j)^*_k (\vec{\boldsymbol{v}}_1)_k = 0 \, . \tag{4}$$

We now compute the following product of matrices:

$$U_1^\dagger A U_1 = \left(\begin{array}{c} \vec{\boldsymbol{v}}_1^\dagger \\ \hdashline Y^\dagger \end{array} \right) A \left(\vec{\boldsymbol{v}}_1 \,\dashv\, Y \right) = \left(\begin{array}{c:c} \vec{\boldsymbol{v}}_1^\dagger A \vec{\boldsymbol{v}}_1 & \vec{\boldsymbol{v}}_1^\dagger A Y \\ \hdashline Y^\dagger A \vec{\boldsymbol{v}}_1 & Y^\dagger A Y \end{array} \right) \, . \tag{5}$$

Note that the partitioned matrix above has the following structure:

$$\left(\begin{array}{c:c} 1 \times 1 & 1 \times (n-1) \\ \hdashline (n-1) \times 1 & (n-1) \times (n-1) \end{array} \right) \, ,$$

where we have indicated the dimensions (number of rows $\times$ number of columns) of the matrices occupying the four possible positions of the partitioned matrix. In particular, there is one row above the horizontal dashed line and $(n-1)$ rows below; there is one column to the left of the vertical dashed line and $(n-1)$ columns to the right. Using $A\vec{\boldsymbol{v}}_1 = \lambda_1 \vec{\boldsymbol{v}}_1$, with $\vec{\boldsymbol{v}}_1$ normalized to unity (*i.e.*, $\vec{\boldsymbol{v}}_1^\dagger \vec{\boldsymbol{v}}_1 = 1$), we see that:

$$\vec{\boldsymbol{v}}_1^\dagger A \vec{\boldsymbol{v}}_1 = \lambda_1 \vec{\boldsymbol{v}}_1^\dagger \vec{\boldsymbol{v}}_1 = \lambda_1 \, ,$$

$$Y^\dagger A \vec{\boldsymbol{v}}_1 = \lambda_1 Y^\dagger \vec{\boldsymbol{v}}_1 = 0 \, .$$

after making use of eq. (4). Using these result in eq. (5) yields:

$$U_1^\dagger A U_1 = \left(\begin{array}{c:c} \lambda_1 & \vec{\boldsymbol{v}}_1^\dagger A Y \\ \hdashline 0 & Y^\dagger A Y \end{array} \right) \, . \tag{6}$$

At this point, we impose the condition that A is hermitian. Note that $A^\dagger = A$ implies that $(U_1^\dagger A U_1)^\dagger = U_1^\dagger A U_1$ [recall that $(AB)^\dagger = B^\dagger A^\dagger$]. This means that $U_1^\dagger A U_1$ is also hermitian. This latter condition can then be used to deduce that the upper right $1 \times (n-1)$ matrix block in eq. (6) must be zero. We therefore conclude that:

$$U_1^\dagger A U_1 = \left(\begin{array}{c:c} \lambda_1 & 0 \\ \hdashline 0 & Y^\dagger A Y \end{array}\right) . \tag{7}$$

In particular, $(Y^\dagger AY)^\dagger = Y^\dagger AY$. In fact, since $U_1^\dagger AU_1$ is hermitian, it follows that λ_1 is real and $Y^\dagger AY$ is hermitian, as expected.

Thus, we have reduced the problem to the diagonalization of the $(n-1)\times(n-1)$ hermitian matrix $Y^\dagger AY$. In particular, the set of eigenvalues of $Y^\dagger AY$ must coincide with the set of the eigenvalues of A with λ_1 omitted. To prove this, consider the eigenvalue problem $Y^\dagger AY\vec{\boldsymbol{v}} = \lambda\vec{\boldsymbol{v}}$. Multiply both sides by Y and use the fact that $YY^\dagger = \mathbf{I}_n$, where $\mathbf{I}_n$ is the $n \times n$ identity matrix.[3] Hence we end up with $AY\vec{\boldsymbol{v}} = \lambda Y\vec{\boldsymbol{v}}$. Putting $\vec{\boldsymbol{w}} \equiv Y\vec{\boldsymbol{v}}$, we obtain $A\vec{\boldsymbol{w}} = \lambda\vec{\boldsymbol{w}}$. This is the eigenvalue problem for A. Thus, in the solution to $Y^\dagger AY\vec{\boldsymbol{v}} = \lambda\vec{\boldsymbol{v}}$, λ must be one of the eigenvalues of A. However, $Y^\dagger AY$ is an $(n-1)\times(n-1)$-dimensional matrix that possesses $n-1$ eigenvalues (in contrast to the $n \times n$ matrix A, which possesses n eigenvalues). We conclude that the eigenvalues of $Y^\dagger AY$ are a subset of the eigenvalues of A. The one that is missing is clearly λ_1 (since the corresponding eigenvector $\vec{\boldsymbol{v}}_1$ is orthogonal to Y), which has already been accounted for above.

We can now repeat the above analysis starting with $Y^\dagger AY$. Defining a new unitary matrix U_2 whose first column is one of the normalized eigenvectors of $Y^\dagger AY$, we will end up reducing the matrix further. We can keep going until we end up with a fully diagonal matrix. At each step, one is simply multiplying on the left with the inverse of a unitary matrix and on the right with a unitary matrix. Since the product of unitary matrices is unitary (check this!), at the end of the process one arrives at:

$$U^\dagger AU = D \equiv \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} , \tag{8}$$

where the eigenvalues of A are the diagonal elements of D and the eigenvectors of A are the columns of U. The latter can be verified explicitly from the equation $AU = UD$. Thus, we have proven that an hermitian matrix is diagonalizable by a unitary similarity transformation. Note that some of the eigenvalues of A may be degenerate (this imposes no difficulty in the above proof). In fact, one of the consequences of this analysis is that the eigenvectors of an hermitian matrix can be chosen to be orthonormal. These eigenvectors are precisely the columns of the matrix U, which was explicitly constructed in deriving eq. (8).

[3] The result $YY^\dagger = \mathbf{I}_n$ follows from the fact that Y is an $n \times (n-1)$ matrix whose columns [rows] are orthonormal.

One important corollary to the above result involves the case of a real symmetric matrix A (*i.e.* A is a real matrix that satisfies $A = A^{\mathsf{T}}$). Since the eigenvalues of a hermitian matrix are real, it follows that the eigenvalues of a real symmetric matrix are real (since a real symmetric matrix is also hermitian). Thus, it is possible to diagonalize a real symmetric matrix by a real orthogonal similarity transformation:

$$R^T A R = D\,,$$

where R is a real matrix that satisfies $RR^T = R^T R = \mathbf{I}$ (note that a real orthogonal matrix is also unitary). The real orthonormal eigenvectors of A are the columns of R, and D is a diagonal matrix whose diagonal elements are the eigenvalues of A.

3. Simultaneous diagonalization of two commuting hermitian matrices

Two hermitian matrices are *simultaneously* diagonalizable by a unitary similarity transformation if and only if they commute. That is, given two hermitian matrices A and B, we can find a unitary matrix V such that *both* $V^\dagger A V = D_A$ and $V^\dagger B V = D_B$ are diagonal matrices. Note that the two diagonal matrices D_A and D_B are not equal in general. But, since V is a matrix whose columns are the eigenvectors of the both A and B, it must be true that the eigenvectors of A and B coincide.

Since all diagonal matrices commute, it follows that $D_A D_B = D_B D_A$. Hence, if $V^\dagger A V = D_A$ and $V^\dagger B V = D_B$, then $(V^\dagger A V)(V^\dagger B V) = (V^\dagger B V)(V^\dagger A V)$. Using $VV^\dagger = \mathbf{I}$, the previous expression simplifies to $V^\dagger A B V = V^\dagger B A V$. Hence, we conclude that $AB = BA$. To complete the proof, we must prove the converse: if two hermitian matrices A and B commute, then A and B are simultaneously diagonalizable by a unitary similarity transformation.

Suppose that $AB = BA$, where A and B are hermitian matrices. Then, we can find a unitary matrix U such that $U^\dagger A U = D_A$, where D_A is diagonal. Using the same matrix U, we shall define $B' \equiv U^\dagger B U$. Explicitly,

$$U^\dagger A U = D_A = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix}\,, \qquad U^\dagger B U \equiv B' = \begin{pmatrix} b'_{11} & b'_{12} & \cdots & b'_{1n} \\ b'_{21} & b'_{22} & \cdots & b'_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b'_{n1} & b'_{n2} & \cdots & b'_{nn} \end{pmatrix}\,.$$

Note that because B is hermitian, it follows that B' is hermitian as well:

$$(B')^\dagger = (U^\dagger B U)^\dagger = U^\dagger B^\dagger U = U^\dagger B U = B'\,.$$

The relation $AB = BA$ imposes a strong constraint on the form of B'. First, observe that:

$$D_A B' = (U^\dagger A U)(U^\dagger B U) = U^\dagger A B U = U^\dagger B A U = (U^\dagger B U)(U^\dagger A U) = B' D_A\,.$$

Explicitly,

$$D_A B = \begin{pmatrix} a_1 b'_{11} & a_1 b'_{12} & \cdots & a_1 b'_{1n} \\ a_2 b'_{21} & a_2 b'_{22} & \cdots & a_2 b'_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_n b'_{n1} & a_n b'_{n2} & \ldots & a_n b'_{nn} \end{pmatrix}, \qquad B D_A = \begin{pmatrix} a_1 b'_{11} & a_2 b'_{12} & \cdots & a_n b'_{1n} \\ a_1 b'_{21} & a_2 b'_{22} & \cdots & a_n b'_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_1 b'_{n1} & a_2 b'_{n2} & \cdots & a_n b'_{nn} \end{pmatrix}.$$

Then, $D_A B - B D_A = 0$ yields $(a_i - a_j) b'_{ij} = 0$. If all the a_i were distinct, then we would be able to conclude that $b'_{ij} = 0$ for $i \neq j$. That is, B' is diagonal. Thus, let us examine carefully what happens if some of the diagonal elements are equal (*i.e.*, some of the eigenvalues of A are degenerate).

If A has some degenerate eigenvalues, we can order the columns of U so that the degenerate eigenvalues are contiguous along the diagonal. Henceforth, we assume this to be the case. We would then conclude that $b_{ij} = 0$ if $a_i \neq a_j$. One can then write D_A and B' in block matrix form:

$$D_A = \left(\begin{array}{c:c:c:c} \lambda_1 \mathbf{I}_1 & 0 & \cdots & 0 \\ \hdashline 0 & \lambda_2 \mathbf{I}_2 & \cdots & 0 \\ \hdashline \vdots & \vdots & \ddots & \vdots \\ \hdashline 0 & 0 & \cdots & \lambda_k \mathbf{I}_k \end{array}\right), \qquad B' = \left(\begin{array}{c:c:c:c} B'_1 & 0 & \cdots & 0 \\ \hdashline 0 & B'_2 & \cdots & 0 \\ \hdashline \vdots & \vdots & \ddots & \vdots \\ \hdashline 0 & 0 & \cdots & B'_k \end{array}\right), \tag{9}$$

assuming that A possesses k distinct eigenvalues. $\mathbf{I}_j$ indicates the identity matrix whose dimension is equal to the multiplicity of the corresponding eigenvalue λ_j. The corresponding B'_j is an hermitian matrix with the same dimension as $\mathbf{I}_j$. Since B'_j is hermitian, it can be diagonalized by a unitary similarity transformation. In particular, we can find a unitary matrix of the form:

$$U' = \left(\begin{array}{c:c:c:c} U'_1 & 0 & \cdots & 0 \\ \hdashline 0 & U'_2 & \cdots & 0 \\ \hdashline \vdots & \vdots & \ddots & \vdots \\ \hdashline 0 & 0 & \cdots & U'_k \end{array}\right),$$

such that $(U')^\dagger B' U' = D_B$ is diagonal. One can easily see by explicit multiplication of the matrices that $(U')^\dagger D_A U' = D_A$. Hence, we have succeeded in finding an invertible matrix $V = UU'$ such that:

$$V^\dagger A V = D_A\,, \qquad V^\dagger B V = D_B\,.$$

That is, A and B are simultaneously diagonalizable by a unitary similarity transformation. The columns of V are the simultaneous eigenvectors of A and B.

The proof we have just given can be extended to prove a stronger result. Two diagonalizable matrices are *simultaneously* diagonalizable if and only if they commute. That is, given two diagonalizable matrices A and B, we can find one invertible operator S such that *both* $S^{-1}AS = D_A$ and $S^{-1}BS = D_B$ are diagonal matrices. The proof follows the same steps given above. However, one step requires more care. Although B is diagonalizable (which implies that B' is diagonalizable), one must prove that each of the B'_j that appears in eq. (9) is diagonalizable. Details of this proof can be found in *Matrix Analysis* (2nd edition), by Robert A. Horn and Charles R. Johnson (Cambridge University Press, Cambridge, UK, 2013) p. 62. We do not require this stronger version of the theorem in these notes.

4. The unitary diagonalization of a normal matrix

We first prove that if A can be diagonalizable by a unitary similarity transformation, then A is normal. If $U^\dagger AU = D$, where D is a diagonal matrix, then $A = UDU^\dagger$ (using the fact that U is unitary). Then, $A^\dagger = UD^\dagger U^\dagger$ and

$$AA^\dagger = (UDU^\dagger)(UD^\dagger U^\dagger) = UDD^\dagger U^\dagger = UD^\dagger DU^\dagger = (UD^\dagger U^\dagger)(UDU^\dagger) = A^\dagger A\,.$$

In this proof, we use the fact that diagonal matrices commute with each other, so that $DD^\dagger = D^\dagger D$.

Conversely, if A is normal then it can be diagonalizable by a unitary similarity transformation. To prove this, we note that any complex matrix A can be uniquely written in the form:

$$A = B + iC\,, \qquad \text{where } B \text{ and } C \text{ are hermitian matrices}\,.$$

To verify this assertion, we simply identify: $B = \frac{1}{2}(A + A^\dagger)$ and $C = -\frac{1}{2}i(A - A^\dagger)$. One easily checks that $B = B^\dagger$, $C = C^\dagger$ and $A = B + iC$. If we now impose the condition that A is normal (*i.e.*, $AA^\dagger = A^\dagger A$) then

$$0 = AA^\dagger - A^\dagger A = (B + iC)(B - iC) - (B - iC)(B + iC) = 2i(CB - BC)\,.$$

Hence, $BC = CB$. Since the two hermitian matrices B and C commute, they can be simultaneously diagonalized by a unitary similarity transformation. If $V^\dagger BV = D_B$ and $V^\dagger CV = D_C$ where D_B and D_C are diagonal, then

$$V^\dagger AV = V^\dagger(B + iC)V = D_B + iD_C\,, \tag{10}$$

which is a diagonal matrix. Therefore, we have explicitly demonstrated that any normal matrix can be diagonalizable by a unitary similarity transformation. Moreover, as was the case for the hermitian matrix, the eigenvectors of a normal matrix can be chosen to be orthonormal and correspond to the columns of V. However, eq. (10) shows that the eigenvalues of a normal matrix are in general complex (in contrast to the real eigenvalues of an hermitian matrix).

5. The unitary diagonalization of a normal matrix—revisited

In the last section, we used the unitary diagonalization of hermitian matrices and the simultaneous unitary diagonalization of two hermitian matrices to prove that a normal matrix can be diagonalized by a unitary similarity transformation. Nevertheless, one can provide a direct proof of the unitary diagonalization of a normal matrix that does not rely on the diagonalization of hermitian matrices. In particular, the same proof given for the unitary diagonalization of an hermitian matrix can also be applied to the case of a normal matrix with only minor changes. For completeness, we provide the details here.

Our starting point is eq. (6), which is valid for any complex matrix. If A is normal then $U_1^\dagger AU_1$ is normal, since

$$\begin{aligned} U_1^\dagger AU_1(U_1^\dagger AU_1)^\dagger &= U_1^\dagger AU_1U_1^\dagger A^\dagger U_1 = U_1^\dagger AA^\dagger U_1\,, \\ (U_1^\dagger AU_1)^\dagger U_1^\dagger AU_1 &= U_1^\dagger A^\dagger U_1U_1^\dagger AU_1 = U_1^\dagger A^\dagger AU_1\,, \end{aligned}$$

where we have used the fact that U_1 is unitary ($U_1U_1^\dagger = \mathbf{I}$). Imposing $AA^\dagger = A^\dagger A$, we conclude that

$$U_1^\dagger AU_1(U_1^\dagger AU_1)^\dagger = (U_1^\dagger AU_1)^\dagger U_1^\dagger AU_1\,. \tag{11}$$

However, eq. (6) implies that

$$U_1^\dagger AU_1(U_1^\dagger AU_1)^\dagger = \left(\begin{array}{c:c} \lambda_1 & \vec{\boldsymbol{v}}_1^\dagger AY \\ \hdashline 0 & Y^\dagger AY \end{array}\right)\left(\begin{array}{c:c} \lambda_1^* & 0 \\ \hdashline Y^\dagger A^\dagger\vec{\boldsymbol{v}}_1 & Y^\dagger A^\dagger Y \end{array}\right) = \left(\begin{array}{c:c} |\lambda_1|^2 + \vec{\boldsymbol{v}}_1^\dagger AYY^\dagger A^\dagger\vec{\boldsymbol{v}}_1 & \vec{\boldsymbol{v}}_1^\dagger AYY^\dagger A^\dagger Y \\ \hdashline Y^\dagger AYY^\dagger A^\dagger\vec{\boldsymbol{v}}_1 & Y^\dagger AYY^\dagger A^\dagger Y \end{array}\right),$$

$$(U_1^\dagger AU_1)^\dagger U_1^\dagger AU_1 = \left(\begin{array}{c:c} \lambda_1^* & 0 \\ \hdashline Y^\dagger A^\dagger\vec{\boldsymbol{v}}_1 & Y^\dagger A^\dagger Y \end{array}\right)\left(\begin{array}{c:c} \lambda_1 & \vec{\boldsymbol{v}}_1^\dagger AY \\ \hdashline 0 & Y^\dagger AY \end{array}\right) = \left(\begin{array}{c:c} |\lambda_1|^2 & \lambda_1^*\,\vec{\boldsymbol{v}}_1^\dagger AY \\ \hdashline \lambda_1 Y^\dagger A^\dagger\vec{\boldsymbol{v}}_1 & Y^\dagger A^\dagger\vec{\boldsymbol{v}}_1\vec{\boldsymbol{v}}_1^\dagger AY + Y^\dagger A^\dagger YY^\dagger AY \end{array}\right),$$

Imposing the result of eq. (11), we first compare the upper left hand block of the two matrices above. We conclude that:

$$\vec{\boldsymbol{v}}_1^\dagger AYY^\dagger A^\dagger\vec{\boldsymbol{v}}_1 = 0\,. \tag{12}$$

But $Y^\dagger A^\dagger\vec{\boldsymbol{v}}_1$ is an $(n-1)$-dimensional vector, so that eq. (12) is the matrix version of the following equation:

$$\langle Y^\dagger A^\dagger\vec{\boldsymbol{v}}_1\,,\, Y^\dagger A^\dagger\vec{\boldsymbol{v}}_1\rangle = 0\,. \tag{13}$$

Since $\langle\vec{\boldsymbol{w}}\,,\,\vec{\boldsymbol{w}}\rangle = 0$ implies that $\vec{\boldsymbol{w}} = 0$ (and $\vec{\boldsymbol{w}}^\dagger = 0$), we conclude from eq. (13) that

$$Y^\dagger A^\dagger\vec{\boldsymbol{v}}_1 = \vec{\boldsymbol{v}}_1^\dagger AY = 0\,. \tag{14}$$

Using eq. (14) in the expressions for $U_1^\dagger AU_1(U_1^\dagger AU_1)^\dagger$ and $(U_1^\dagger AU_1)^\dagger U_1^\dagger AU_1$ above, we see that eq. (11) requires that eq. (14) and the following condition are both satisfied:

$$Y^\dagger AYY^\dagger A^\dagger Y = Y^\dagger A^\dagger YY^\dagger AY\,.$$

The latter condition simply states that $Y^\dagger AY$ is normal. Using eq. (14) in eq. (6) then yields:

$$U_1^\dagger AU_1 = \left(\begin{array}{c:c} \lambda_1 & 0 \\ \hdashline 0 & Y^\dagger AY \end{array}\right) , \tag{15}$$

where $Y^\dagger AY$ is normal. As in the case of hermitian A, we have succeeded in reducing the original $n \times n$ normal matrix A down to an $(n-1) \times (n-1)$ normal matrix $Y^\dagger AY$, and we can now repeat the procedure again. The end result is once again the unitary diagonalization of A:

$$U^\dagger AU = D \equiv \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} .$$

Moreover, the eigenvalues of A are the diagonal elements of D and the eigenvectors of A are the columns of U. This should be clear from the equation $AU = UD$. Thus, we have proven that a normal matrix is diagonalizable by a unitary similarity transformation.