# Supplement: Symmetric and Hermitian Matrices

## A Bunch of Definitions

**Definition**: A real $n \times n$ matrix $A$ is called *symmetric* if $A^T = A$.

**Definition**: A complex $n \times n$ matrix $A$ is called *Hermitian* if $A^* = A$, where $A^* = \overline{A^T}$, the conjugate transpose.

**Definition**: A complex $n \times n$ matrix $A$ is called *normal* if $A^*A = AA^*$, i.e. commutes with its conjugate transpose.

It is quite a surprising result that these three kinds of matrices are *always diagonalizable*; and moreover, one can construct an orthonormal basis (in standard inner product) for $\mathbb{R}^n/\mathbb{C}^n$, consisting of eigenvectors of $A$. Hence the matrix $P$ that gives diagonalization $A = PDP^{-1}$ will be orthogonal/unitary, namely:

**Definition**: An $n \times n$ real matrix $P$ is called *orthogonal* if $P^TP = I_n$, i.e. $P^{-1} = P^T$.

**Definition**: An $n \times n$ complex matrix $P$ is called *unitary* if $P^*P = I_n$, i.e. $P^{-1} = P^*$.

Diagonalization using these special kinds of $P$ will have special names:

**Definition**: A matrix $A$ is called *orthogonally diagonalizable* if $A$ is similar to a diagonal matrix $D$ with an orthogonal matrix $P$, i.e. $A = PDP^T$.
A matrix $A$ is called *unitarily diagonalizable* if $A$ is similar to a diagonal matrix $D$ with a unitary matrix $P$, i.e. $A = PDP^*$.

Then we have the following big theorems:

**Theorem**: Every real $n \times n$ symmetric matrix $A$ is orthogonally diagonalizable

**Theorem**: Every complex $n \times n$ Hermitian matrix $A$ is unitarily diagonalizable.

**Theorem**: Every complex $n \times n$ normal matrix $A$ is unitarily diagonalizable.

To prove the above results, it is convenient to introduce the concept of *adjoint operator*, which allows us to discuss effectively the "transpose" operation in a general inner product space.

## The Adjoint Operator

Let $V$ be an $n$-dimensional inner product space and let $T : V \to V$ be a linear operator. We find out that under the inner product operation, the action of $T : \mathbf{v} \mapsto T(\mathbf{v})$ can be replaced/represented by another inner product action using a suitably chosen vector.

**Lemma 1**: Let $\mathbf{w} \in V$ be a given vector. Then there is a unique vector $\mathbf{w}^* \in V$ such that:

$$<T(\mathbf{v}), \mathbf{w}> = <\mathbf{v}, \mathbf{w}^*>, \quad \text{for every } \mathbf{v} \in V. \qquad (*)$$

**Proof**: Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ be an orthonormal basis for $V$. The following $\mathbf{w}^*$ is what we want:

$$\mathbf{w}^* = \overline{<T(\mathbf{u}_1), \mathbf{w}>}\mathbf{u}_1 + \ldots + \overline{<T(\mathbf{u}_p), \mathbf{w}>}\mathbf{u}_p = \sum_{i=1}^{n} \overline{<T(\mathbf{u}_i), \mathbf{w}>}\mathbf{u}_i.$$

Now, for $j = 1, \ldots, n$, we check $(*)$ for basis vector $\mathbf{u}_j$ first:

$$\begin{aligned} <\mathbf{u}_j, \mathbf{w}^*> &= <\mathbf{u}_j, \sum_{i=1}^{n} \overline{<T(\mathbf{u}_i), \mathbf{w}>}\mathbf{u}_i>= \sum_{i=1}^{n} <T(\mathbf{u}_i), \mathbf{w}><\mathbf{u}_j, \mathbf{u}_i> \\ &= <T(\mathbf{u}_j), \mathbf{w}><\mathbf{u}_j, \mathbf{u}_j>=<T(\mathbf{u}_j), \mathbf{w}> . \end{aligned}$$

So, for a general $\mathbf{v} \in V$, by expressing $\mathbf{v} = c_1\mathbf{u}_1 + \ldots + c_n\mathbf{u}_n = \sum_{j=1}^{n} c_j\mathbf{u}_j$, we have:

$$\begin{aligned} <T(\mathbf{v}), \mathbf{w}> &= <\sum_{j=1}^{n} c_j T(\mathbf{u}_j), \mathbf{w}>= \sum_{j=1}^{n} c_j <T(\mathbf{u}_j), \mathbf{w}> \\ &= \sum_{j=1}^{n} c_j <\mathbf{u}_j, \mathbf{w}^*>=<\sum_{j=1}^{n} c_j\mathbf{u}_j, \mathbf{w}^*>=<\mathbf{v}, \mathbf{w}^*> . \end{aligned}$$

For the uniqueness of $\mathbf{w}^*$, let $\mathbf{w}' \in V$ be another vector with the same property, namely:

$$<T(\mathbf{v}), \mathbf{w}>=<\mathbf{v}, \mathbf{w}^*>=<\mathbf{v}, \mathbf{w}'>, \quad \text{for every } \mathbf{v} \in V.$$

Then we take difference:

$$<\mathbf{v}, \mathbf{w}^* - \mathbf{w}'>= 0, \quad \text{for every } \mathbf{v} \in V.$$

In particular, this equality should be valid for $\mathbf{v} = \mathbf{w}^* - \mathbf{w}' \in V$. Thus we have:

$$<\mathbf{w}^* - \mathbf{w}', \mathbf{w}^* - \mathbf{w}'>= 0 \quad \Rightarrow \quad ||\mathbf{w}^* - \mathbf{w}'|| = 0 \quad \Rightarrow \quad \mathbf{w}^* = \mathbf{w}'$$ □

**Definition**: Let $T : V \to V$ be a linear operator. For each $\mathbf{w} \in V$, we define $T^*(\mathbf{w}) := \mathbf{w}^*$, where $\mathbf{w}^*$ is the unique vector obtained in Lemma 1. This $T^*$ is called the *adjoint* of $T$.

**Lemma 2**: The adjoint operator $T^* : V \to V$ is linear.

**Proof**: Straightforward checking. Let $\mathbf{w}_1, \mathbf{w}_2 \in V$ and $c, d \in \mathbf{C}$. Then for every $\mathbf{v} \in V$, first by definition of $T^*$ we have:

$$<T(\mathbf{v}), (c\mathbf{w}_1 + d\mathbf{w}_2)>=<\mathbf{v}, T^*(c\mathbf{w}_1 + d\mathbf{w}_2)> .$$

But on the other hand:

$$\begin{aligned} <T(\mathbf{v}), (c\mathbf{w}_1 + d\mathbf{w}_2)> &= \bar{c} <T(\mathbf{v}), \mathbf{w}_1> + \bar{d} <T(\mathbf{v}), \mathbf{w}_2> \\ &= \bar{c} <\mathbf{v}, T^*(\mathbf{w}_1)> + \bar{d} <\mathbf{v}, T^*(\mathbf{w}_2)> \\ &= <\mathbf{v}, c\,T^*(\mathbf{w}_1) + d\,T^*(\mathbf{w}_2)> \end{aligned}$$

The above two equalities are valid for every $\mathbf{v} \in V$. So by the same uniqueness proof as in Lemma 1, we obtain:

$$T^*(c\mathbf{w}_1 + d\mathbf{w}_2) = c\,T^*(\mathbf{w}_1) + d\,T^*(\mathbf{w}_2),$$

and thus $T^*$ is linear. □

**Theorem 1**: Let $T, U$ be linear operators on $V$ and $k \in \mathbb{C}$. Then:

(i) $(T+U)^* = T^* + U^*$;

(ii) $(kT)^* = \bar{k}T^*$;

(iii) $(U \circ T)^* = T^* \circ U^*$;

(iv) $(T^*)^* = T$.

**Proof**: Directly from definitions. For example, the checking for (iv):

Let $\mathbf{v} \in V$ be any vector. Then by definition:

$$<(T^*)^*(\mathbf{v}), \mathbf{u}>=<\mathbf{v}, T^*(\mathbf{u})>=<T(\mathbf{v}), \mathbf{u}>, \quad \text{for every } \mathbf{u} \in V.$$

Hence $(T^*)^*(\mathbf{v}) = T(\mathbf{v})$ for any $\mathbf{v} \in V$ and thus $(T^*)^* = T$. □

This adjoint operator $T^*$, when using matrix representation with an orthonormal basis $\mathcal{B}$, has a simple relationship with the original linear operator $T$.

**Theorem 2**: Let $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ be an orthonormal basis of $V$, and let $T$ be a linear operator in $V$. Then the matrix representations of $T$ and $T^*$ relative to the orthonormal basis $\mathcal{B}$ are given by:

$$[T]_{\mathcal{B}} = \Big[ <T(\mathbf{u}_j), \mathbf{u}_i> \Big] \quad \text{and} \quad [T^*]_{\mathcal{B}} = [T]^*_{\mathcal{B}}.$$

**Remark**: $\mathcal{B}$ must be orthonormal!

**Proof**: First we consider the $j$-th column of $[T]_{\mathcal{B}}$, i.e. $[T(\mathbf{u}_j)]_{\mathcal{B}}$. Its entries are the $\mathcal{B}$-coordinates of $T(\mathbf{u}_j)$, which are exactly the coefficients in the linear combination:

$$T(\mathbf{u}_j) = a_{1j}\mathbf{u}_1 + \ldots + a_{nj}\mathbf{u}_j.$$

Since $\mathcal{B}$ is orthonormal, the $i$-th coefficient in the above linear combination can be computed effectively as:

$$<T(\mathbf{u}_j), \mathbf{u}_i>= a_{1j} <\mathbf{u}_i, \mathbf{u}_i> + \ldots + a_{nj} <\mathbf{u}_n, \mathbf{u}_i>= a_{ij}.$$

Thus the $(i,j)$-th entry of $[T]_{\mathcal{B}}$ is given by $a_{ij} =<T(\mathbf{u}_j), \mathbf{u}_i>$.

Similarly the $(i,j)$-th entry of $[T^*]_{\mathcal{B}}$ is given by $<T^*(\mathbf{u}_j), \mathbf{u}_i>$. Using the definition of adjoint operator, we have:

$$<T^*(\mathbf{u}_j), \mathbf{u}_i>= \overline{<\mathbf{u}_i, T^*(\mathbf{u}_j)>} = \overline{<T(\mathbf{u}_i), \mathbf{u}_j>} = \bar{a}_{ji}.$$

So $[T^*]_{\mathcal{B}} = [T]^*_{\mathcal{B}}$ □

**Definition**: A linear operator $T : V \to V$ is called *self-adjoint* if $T^* = T$.

Thus, by Theorem 2, matrix transformation given by a symmetric/Hermitian matrix will be a self-adjoint operator on $\mathbb{R}^n/\mathbb{C}^n$, using the standard inner product.

Next we need to setup some technical lemmas for the proof of the main theorem.

**Lemma 3**: Let $T$ be a self-adjoint operator on $V$. Then every eigenvalue of $T$ must be real.

**Proof**: Let $\mathbf{v} \neq \mathbf{0}$ be an eigenvector of $T$ corresponding to eigenvalue $\lambda$. We consider:

$$<T(\mathbf{v}), \mathbf{v}>=<\lambda\mathbf{v}, \mathbf{v}>= \lambda <\mathbf{v}, \mathbf{v}> .$$

On the other hand, since $T^* = T$, we also have:

$$<T(\mathbf{v}), \mathbf{v}>=<\mathbf{v}, T^*(\mathbf{v})>=<\mathbf{v}, T(\mathbf{v})>=<\mathbf{v}, \lambda\mathbf{v}>= \bar{\lambda} <\mathbf{v}, \mathbf{v}> .$$

As $<\mathbf{v}, \mathbf{v}>\neq 0$, we must have $\lambda = \bar{\lambda}$, i.e. $\lambda$ is real. □

**Lemma 4**: Every self-adjoint operator on $V$ has an eigenvector.

**Proof**: Take an orthonormal basis $\mathcal{B}$ of $V$. Then we get a symmetric/Hermitian matrix $A = [T]_{\mathcal{B}}$. By the fundamental theorem of algebra, $A$ must have an eigenvalue $\lambda \in \mathbb{C}$, and hence a corresponding eigenvector $\mathbf{x} \in \mathbb{C}^n$. In complex case we just send this $\mathbf{x} \in \mathbb{C}^n$ back to $\mathbf{v} \in V$ by inverse $\mathcal{B}$-coordinate mapping, then we will get $T(\mathbf{v}) = \lambda\mathbf{v}$. In real case, we apply Lemma 3 to know that this $\lambda$ must be real. Hence $\mathbf{x} \in \mathbb{R}^n$ and we can send it back to $\mathbf{v} \in V$ to get $T(\mathbf{v}) = \lambda\mathbf{v}$ again. □

**Lemma 5**: Let $W$ be a subspace of $V$ such that $T(W) \subseteq W$, i.e. $T(\mathbf{w}) \in W$ for every $\mathbf{w} \in W$. Then $T^*(W^\perp) \subseteq W^\perp$.

**Proof**: Let $\mathbf{z} \in W^\perp$. Then for $\mathbf{w} \in W$:

$$<\mathbf{w}, T^*(\mathbf{z})>=<T(\mathbf{w}), \mathbf{z}>= 0 \quad \text{as} \quad T(\mathbf{w}) \in W \text{ and } \mathbf{z} \in W^\perp.$$

Since the above is valid for every $\mathbf{w} \in W$, we should have $T^*(\mathbf{z}) \in W^\perp$. □

**Lemma 6**: Let $W$ be a subspace of an $n$-dimensional inner product space $V$. Then:

$$\dim W + \dim W^\perp = n = \dim V.$$

**Proof**: Let $\{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$ and $\{\mathbf{z}_1, \ldots, \mathbf{z}_\ell\}$ be orthogonal bases of $W$ and $W^\perp$ respectively. The lemma is proved if we can show that $S = \{\mathbf{w}_1, \ldots, \mathbf{w}_k, \mathbf{z}_1, \ldots, \mathbf{z}_\ell\}$ forms a basis for $V$.

*Spanning* $V$: For every $\mathbf{v} \in V$, we have the orthogonal decomposition of $\mathbf{v}$ w.r.t. $W$:

$$\mathbf{v} = \text{proj}_W\mathbf{v} + (\mathbf{v} - \text{proj}_W\mathbf{v}), \quad \text{where } \text{proj}_W\mathbf{v} \in W \text{ and } (\mathbf{v} - \text{proj}_W\mathbf{v}) \in W^\perp.$$

Use the bases of $W$ and $W^\perp$ to express $\text{proj}_W\mathbf{v} = \sum_{i=1}^k c_i\mathbf{w}_i$ and $(\mathbf{v} - \text{proj}_W\mathbf{v}) = \sum_{j=1}^\ell d_j\mathbf{z}_j$. Hence $\mathbf{v}$ can be expressed as a linear combination of vectors in $S$.

*Linearly independent*: Consider the vector equation:

$$c_1\mathbf{w}_1 + \ldots + c_k\mathbf{w}_k + d_1\mathbf{z}_1 + \ldots + d_\ell\mathbf{z}_\ell = \mathbf{0}.$$

Take inner product with $\mathbf{w}_1$. As $\{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$ is an orthogonal set, we have $<\mathbf{w}_i, \mathbf{w}_1>= 0$ for $i \neq 1$. On the other hand, since $\mathbf{w}_1 \in W$ and all $\mathbf{z}_j \in W^\perp$, we get $<\mathbf{z}_j, \mathbf{w}_1>= 0$ for all $1 \leq j \leq \ell$. So the above vector equation will become:

$$c_1||\mathbf{w}_1||^2 + 0 + \ldots + 0 =<\mathbf{0}, \mathbf{w}_1>= 0.$$

As $\mathbf{w}_1 \neq \mathbf{0}$, we get $c_1 = 0$. Similarly for other $c_i$ and $d_j$ and they are all zeros. Thus $S$ is also linearly independent. □

Now we are ready to prove the main theorem.

**Diagonalizability of Symmetric and Hermitian Matrices**

**Main Theorem**: Let $T^* = T$ be a self-adjoint linear operator on $V$. Then $V$ has an orthonormal basis consisting of eigenvectors of $T$.

**Proof**: We use induction on $n = \dim V$.

$\underline{n=1}$: Any non-zero vector $\mathbf{v}_1$ will be an eigenvector of $T$ since $V = \text{Span}\{\mathbf{v}_1\}$. After normalization, $\mathbf{u}_1 = \frac{\mathbf{v}_1}{||\mathbf{v}_1||}$, we obtain an orthonormal basis $\{\mathbf{u}_1\}$ of $V$ consisting of eigenvector of $T$.

Now, assume the statement is true for $\dim V = k$. Next consider $\dim V = k+1$.

By Lemma 4, $T$ has an eigenvector $\mathbf{u}_1$ (may assume $||\mathbf{u}_1|| = 1$) corresponding to eigenvalue $\lambda_1$. Let $W = \text{Span}\{\mathbf{u}_1\}$. Note that $T(W) = W$.

By Lemma 5, we have $T^*(W^\perp) \subseteq W^\perp$. Since $T^* = T$, this gives $T(W^\perp) \subseteq W^\perp$. In other words, we can regard $T$ as a linear operator defined on $W^\perp$. Note that Lemma 6 says that $\dim W^\perp = \dim V - \dim W = k$, so by induction hypothesis, there is an orthonormal basis of $W^\perp$ consisting of eigenvectors of $T$, say $\{\mathbf{u}_2, \ldots, \mathbf{u}_{k+1}\}$.

Since $\mathbf{u}_1 \in W$, $||\mathbf{u}_1|| = 1$, and $\{\mathbf{u}_2, \ldots, \mathbf{u}_{k+1}\} \subset W^\perp$, the combined set $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{k+1}\}$ is again orthonormal. This will be an orthonormal basis of $V$ consisting of eigenvectors of $T$.

□

In the case of symmetric (or Hermitian) matrix transformation, by using such an orthonormal basis of eigenvectors to construct the matrix $P$, we will have the diagonalization $A = PDP^{-1}$ with $P^{-1} = P^T$ (or $P^{-1} = P^*$).

**Remark**: To find this $P$, we have a more efficient method than the inductive construction in the proof of main theorem.

**Lemma 7**: Let $T^* = T$. Then eigenvectors of $T$ corresponding to <u>distinct</u> eigenvalues are orthogonal to each other.

**Proof**: Let $T(\mathbf{v}_1) = \lambda_1\mathbf{v}_1$ and $T(\mathbf{v}_2) = \lambda_2\mathbf{v}_2$ with $\lambda_1 \neq \lambda_2$. Consider on the one hand:

$$<T(\mathbf{v}_1), \mathbf{v}_2> = <\lambda_1\mathbf{v}_1, \mathbf{v}_2> = \lambda_1 <\mathbf{v}_1, \mathbf{v}_2>,$$

and on the other hand:

$$<T(\mathbf{v}_1), \mathbf{v}_2> = <\mathbf{v}_1, T^*(\mathbf{v}_2)> = <\mathbf{v}_1, T(\mathbf{v}_2)> = <\mathbf{v}_1, \lambda_2\mathbf{v}_2> = \bar{\lambda}_2 <\mathbf{v}_1, \mathbf{v}_2> .$$

Since $T$ is self-adjoint, $\lambda_2$ must be real, so we obtain:

$$\lambda_1 <\mathbf{v}_1, \mathbf{v}_2> = \lambda_2 <\mathbf{v}_1, \mathbf{v}_2> .$$

As $\lambda_1 \neq \lambda_2$, we must have $<\mathbf{v}_1, \mathbf{v}_2> = 0$.

□

**Corollary**: Let $T^* = T$ and let $\{\mathbf{v}_{1i_1}\}, \ldots, \{\mathbf{v}_{pi_p}\}$ be orthogonal sets of eigenvectors corresponding to distinct eigenvalues $\lambda_1, \ldots, \lambda_p$ of $T$. Then the total collection of eigenvectors $\{\mathbf{v}_{ji_j};\ 1 \leq i \leq p\}$ is again orthogonal.

**Proof**: Exercise.

With Lemma 7 and its corollary, we only need to produce orthonormal basis for each eigenspace, which can be done by a Gram-Schmidt process. Then the total collection will be automatically orthonormal. And it is guaranteed by the main theorem that $A$ must be diagonalizable.

**Remark**: If $\mathbf{v}_1, \mathbf{v}_2$ are eigenvectors of $A$ corresponding to distinct eigenvalues, we know that $\mathbf{v}_1 + \mathbf{v}_2$ can never be an eigenvector of $A$. So Gram-Schmidt process should not be applied across bases for different eigenspaces.

**Example**: Orthogonally diagonalize the following symmetric matrix:

$$A = \begin{bmatrix} 1 & 2 & -4 \\ 2 & -2 & -2 \\ -4 & -2 & 1 \end{bmatrix}.$$

**Solution**: The characteristic equation of $A$ is:

$$\det(A - \lambda I) = -\lambda^3 + 27\lambda + 54 = -(\lambda + 3)^2(\lambda - 6) = 0.$$

So the eigenvalues are $-3$, $-3$, $6$.

For the eigenvalue $\lambda = -3$, we solve for $\mathrm{Nul}\,(A + 3I)$:

$$A + 3I = \begin{bmatrix} 4 & 2 & -4 \\ 2 & 1 & -2 \\ -4 & -2 & 4 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & \frac{1}{2} & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

So $\mathrm{Nul}\,(A+3I)$ has a basis $\{[1 \;\; 0 \;\; 1]^T, [-\frac{1}{2} \;\; 1 \;\; 0]^T\}$. By Gram-Schmidt process, we obtain an orthonormal basis for $\mathrm{Nul}\,(A + 3I)$:

$$\left\{ \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \begin{bmatrix} -\frac{1}{\sqrt{18}} \\ \frac{4}{\sqrt{18}} \\ \frac{1}{\sqrt{18}} \end{bmatrix} \right\}.$$

For the eigenvalue $\lambda = 6$, we solve for $\mathrm{Nul}\,(A - 6I)$:

$$A - 6I = \begin{bmatrix} -5 & 2 & -4 \\ 2 & -8 & -2 \\ -4 & -2 & -5 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix}.$$

So $\mathrm{Nul}\,(A-6I)$ has a basis $\{[-1 \;\; -\frac{1}{2} \;\; 1]^T\}$ and we obtain an orthonormal basis for $\mathrm{Nul}\,(A-6I)$:

$$\left\{ \begin{bmatrix} -\frac{2}{3} \\ -\frac{1}{3} \\ \frac{2}{3} \end{bmatrix} \right\}.$$

We construct the orthogonal matrix $P$ and diagonal matrix $D$ as:

$$P = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{18}} & -\frac{2}{3} \\ 0 & \frac{4}{\sqrt{18}} & -\frac{1}{3} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{18}} & \frac{2}{3} \end{bmatrix}, \quad D = \begin{bmatrix} -3 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 6 \end{bmatrix}.$$

Then one can check that $A = PDP^T$.

**Note**: The diagonalization $A = PDP^T$ is not unique, as one can have different choices of orthonormal bases for those eigenspaces with dimension greater than one. For example, the above $A$ also allows an orthogonal diagonalization $A = QDQ^T$ with:

$$Q = \begin{bmatrix} \frac{1}{3} & -\frac{2}{3} & -\frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \\ \frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \end{bmatrix}.$$

**Diagonalization of Complex Normal Matrices**

**Definition**: A linear operator $T$ on $V$ is called *normal* if $T \circ T^* = T^* \circ T$.

To make the proof of main theorem also work for normal operator, we need the following technical lemma.

**Lemma 8**: Let $T$ be a normal operator on $V$. Then:

(i) $\mathbf{v}$ is an eigenvector of $T$ corresponding to eigenvalue $\lambda$

$\Leftrightarrow$ $\mathbf{v}$ is an eigenvector of $T^*$ corresponding to eigenvalue $\bar{\lambda}$.

(ii) Eigenvectors corresponding to distinct eigenvalues of $T$ are orthogonal to each other.

**Proof**: (i) First we claim that $||T(\mathbf{v})|| = ||T^*(\mathbf{v})||$.

$$\begin{aligned} ||T(\mathbf{v})||^2 &= <T(\mathbf{v}), T(\mathbf{v})>=<\mathbf{v}, T^*T(\mathbf{v})> \\ &= <\mathbf{v}, TT^*(\mathbf{v})>=<T^*(\mathbf{v}), T^*(\mathbf{v})>= ||T^*(\mathbf{v})||^2. \end{aligned}$$

Then for any scalar $\lambda$, note that the operator $U = T - \lambda I$ is also normal with $U^* = T^* - \bar{\lambda} I$, so we have:

$$||(T - \lambda I)(\mathbf{v})|| = ||(T^* - \bar{\lambda} I)(\mathbf{v})||. \qquad (*)$$

Hence:

$$\begin{aligned} & \mathbf{v} \text{ is an eigenvector of } T \text{ corresponding to eigenvalue } \lambda \\ \Leftrightarrow\ & (T - \lambda I)(\mathbf{v}) = \mathbf{0} \\ \Leftrightarrow\ & (T^* - \bar{\lambda} I)(\mathbf{v}) = \mathbf{0} \quad (\text{by } (*)) \\ \Leftrightarrow\ & \mathbf{v} \text{ is an eigenvector of } T^* \text{ corresponding to eigenvalue } \bar{\lambda} \end{aligned}$$

(ii) Now let $\mathbf{v}_1, \mathbf{v}_2$ be eigenvectors of $T$, corresponding to distinct eigenvalues $\lambda_1 \neq \lambda_2$ respectively. Consider on the one hand:

$$<T(\mathbf{v}_1), \mathbf{v}_2>=<\lambda_1 \mathbf{v}_1, \mathbf{v}_2>= \lambda_1 <\mathbf{v}_1, \mathbf{v}_2>;$$

and on the other hand:

$$<T(\mathbf{v}_1), \mathbf{v}_2>=<\mathbf{v}_1, T^*(\mathbf{v}_2)>=<\mathbf{v}_1, \bar{\lambda}_2 \mathbf{v}_2>= \lambda_2 <\mathbf{v}_1, \mathbf{v}_2> .$$

So we again obtain:

$$\lambda_1 <\mathbf{v}_1, \mathbf{v}_2>= \lambda_2 <\mathbf{v}_1, \mathbf{v}_2> .$$

As $\lambda_1 \neq \lambda_2$, we must have $<\mathbf{v}_1, \mathbf{v}_2>= 0$. □

Now, we give the proof of main theorem for normal operators.

**Main Theorem′**: Let $T$ be a normal operator on a complex inner product space $V$. Then $V$ has an orthonormal basis consisting of eigenvectors of $T$.

**Proof**: We use induction on $n = \dim V$.

$n = 1$: Same as before.

Now, assume the statement is true for $\dim V = k$. Next consider $\dim V = k + 1$.

Since $V$ is a complex inner product space, $T$ will have an eigenvector $\mathbf{u}_1$ (may assume $||\mathbf{u}_1|| = 1$) corresponding to eigenvalue $\lambda_1$. (*For real inner product space we might get stuck at this point.*)

By Lemma 8(i), $\mathbf{u}_1$ is also an eigenvector of $T^*$. So if we set $W = \text{Span}\{\mathbf{u}_1\}$, we have $T^*(W) \subseteq W$.

By Lemma 5, we have $(T^*)^*(W^\perp) \subseteq W^\perp$. As $(T^*)^* = T$, this means $T(W^\perp) \subseteq W^\perp$. Then we can continue the inductive argument as in the previous proof of Main Theorem.

□