Section 6.4:

**Lemma 1.** Let $T$ be a linear operator on a finite-dimensional inner product space $V$. If $T$ has an eigenvector, then so does $T^*$.

**Proof.** Let $v$ be an eigenvector of $T$ corresponding to the eigenvalue $\lambda$. For all $x \in V$, we have

$$0 = \langle 0, x \rangle = \langle (T - \lambda I)(x), x \rangle = \langle v, (T - \lambda I)^*(x) \rangle = \langle v, (T^* - \lambda I)(x) \rangle$$

So $v$ is orthogonal to $(T^* - \lambda I)(x)$ for all $x$. Thus, $v \notin R(T^* - \lambda I)$ and so the nullity of $(T^* - \lambda I)$ is not 0. There exists $x \neq 0$ such that $(T^* - \lambda I)(x) = 0$. Thus $x$ is a eigenvector corresponding to the eigenvalue $\lambda$ of $T^*$.

$\square$
Theorem 6.14. (Schur). Let $T$ be a linear operator on a finite-dimensional inner product space $V$. Suppose that the characteristic polynomial of $T$ splits. Then there exists an orthonormal basis $\beta$ for $V$ such that the matrix $[T]_\beta$ is upper triangular.
Defn 1. Let $V$ be an inner product space, and let $T$ be a linear operator on $V$. We say that $T$ is normal if $TT^* = T^*T$. An $n \times n$ real or complex matrix $A$ is normal if $AA^* = A^*A$.

Theorem 6.15. Let $V$ be an inner product space, and let $T$ be a normal operator on $V$. Then the following statements are true.

(a) $||T(x)|| = ||T^*(x)||$ for all $x \in V$.
(b) $T - cI$ is normal for every $c \in F$.
(c) If $x$ is an eigenvector of $T$, then $x$ is also an eigenvector of $T^*$. In fact, if $T(x) = \lambda x$, then $T^*(x) = \overline{\lambda}x$.
(d) If $\lambda_1$ and $\lambda_2$ are distinct eigenvalues of $T$ with corresponding eigenvectors $x_1$ and $x_2$, then $x_1$ and $x_2$ are orthogonal.
Defn 2. Let $T$ be a linear operator on an inner product space $V$. We say that $T$ is **self-adjoint** (Hermitian) if $T = T^*$. An $n \times n$ real or complex matrix $A$ is **self-adjoint** (Hermitian) if $A = A^*$.

**Theorem 6.16.** Let $T$ be a linear operator on a finite-dimensional complex inner product space $V$. Then $T$ is normal if and only if there exists an orthonormal basis for $V$ consisting of eigenvectors of $T$. 

Lemma 2. Let $T$ be a self-adjoint operator on a finite-dimensional inner product space $V$. Then

(a) Every eigenvalue of $T$ is real.

(b) Suppose that $V$ is a real inner product space. Then the characteristic polynomial of $T$ splits.
Theorem 6.17. Let $T$ be a linear operator on a finite-dimensional real inner product space $V$. Then $T$ is self-adjoint if and only if there exists an orthonormal basis $\beta$ for $V$ consisting of eigenvectors of $T$. 
Section 6.5:

**Defn 3.** Let \( T \) be a linear operator on a finite-dimensional inner product space \( V \) (over \( F \)). If \( ||T(x)|| = ||x|| \) for all \( x \in V \), we call \( T \) a **unitary operator** if \( F = \mathbb{C} \) and an **orthogonal operator** if \( F = \mathbb{R} \).

(Comment on relationship to one-to-one, onto, and invertible.)

**Theorem 6.18.** Let \( T \) be a linear operator on a finite-dimensional inner product space \( V \). Then the following statements are equivalent.

1. \( TT^* = T^*T = I \).
2. \( \langle T(x), T(y) \rangle = \langle x, y \rangle, \forall x, y \in V \).
3. If \( \beta \) is an orthonormal basis for \( V \), then \( T(\beta) \) is an orthonormal basis for \( V \).
4. There exists an orthonormal basis \( \beta \) for \( V \) such that \( T(\beta) \) is an orthonormal basis for \( V \).
5. \( ||T(x)|| = ||x||, \forall x \in V \).

**Proof.** (1) \( \Rightarrow \) (2)

\[
\langle T(x), T(y) \rangle = \langle x, T^*T(y) \rangle = \langle x, y \rangle
\]

(2) \( \Rightarrow \) (3) Let \( v_i, v_j \in \beta \). Then \( 0 = \langle v_i, v_j \rangle = \langle T(v_i), T(v_j) \rangle \), so \( T(\beta) \) is orthogonal.

By Corollary 2 to Theorem 6.3, any orthogonal subset is linearly independent and since \( T(\beta) \) has \( n \) vectors, it must be a basis of \( V \).

Also,

\[
1 = ||v_i||^2 = \langle v_i, v_i \rangle = \langle T(v_i), T(v_i) \rangle.
\]

So, \( T(\beta) \) is an orthonormal basis of \( V \).

(3) \( \Rightarrow \) (4) By Shur’s Theorem, there is an orthonormal basis \( \beta \) for \( V \). By (3), \( T(\beta) \) is orthonormal.

(4) \( \Rightarrow \) (5) Let \( \beta = \{v_1, v_2, \ldots, v_n\} \) be an orthonormal basis for \( V \). Let

\[
x = a_1v_1 + a_2v_2 + \cdots + a_nv_n.
\]

Then

\[
||x||^2 = a_1^2\langle v_1, v_1 \rangle + a_2^2\langle v_2, v_2 \rangle + \cdots + a_n^2\langle v_n, v_n \rangle = a_1^2 + a_2^2 + \cdots + a_n^2
\]

And

\[
||T(x)||^2 = a_1^2\langle T(v_1), T(v_1) \rangle + a_2^2\langle T(v_2), T(v_2) \rangle + \cdots + a_n^2\langle T(v_n), T(v_n) \rangle = a_1^2 + a_2^2 + \cdots + a_n^2
\]

Therefore, \( ||T(x)|| = ||x|| \).

(5) \( \Rightarrow \) (1) We are given \( ||T(x)|| = ||x|| \), for all \( x \).

We know \( \langle x, x \rangle = 0 \) if and only if \( x = 0 \) and \( \langle T(x), T(x) \rangle = 0 \) if and only if \( T(x) = 0 \).

Therefore, \( T(x) = 0 \) if and only if \( x = 0 \).

So \( N(T) = \{0\} \) and therefore, \( T \) is invertible.

We have \( \langle x, x \rangle = \langle T(x), T(x) \rangle = \langle x, T^*T(x) \rangle \). Therefore, \( T^*T(x) = x \) for all \( x \), which implies that \( T^*T = I \). But since \( T \) is invertible, it must be that \( T^* = T^{-1} \) and we have that \( T^*T = TT^* = I \). \( \square \)
Lemma 3. Let $U$ be a self-adjoint operator on a finite-dimensional inner product space $V$. If $\langle x, U(x) \rangle = 0$ for all $x \in V$, then $U = T_0$.

Cor 1. Let $T$ be a linear operator on a finite-dimensional real inner product space $V$. Then $V$ has an orthonormal basis of eigenvectors of $T$ with corresponding eigenvalues of absolute value 1 if and only if $T$ is both self-adjoint and orthogonal.

Cor 2. Let $T$ be a linear operator on a finite-dimensional complex inner product space $V$. Then $V$ has an orthonormal basis of eigenvectors of $T$ with corresponding eigenvalues of absolute value 1 if and only if $T$ is unitary.

Defn 4. A square matrix $A$ is called an orthogonal matrix if $A^t A = A A^t = I$ and unitary if $A^* A = A A^* = I$. We say $B$ is unitarily equivalent to $D$ if there exists a unitary matrix $Q$ such that $D = Q^* B Q$.

Theorem 6.19. Let $A$ be a complex $n \times n$ matrix. Then $A$ is normal if and only if $A$ is unitarily equivalent to a diagonal matrix.

Theorem 6.20. Let $A$ be a real $n \times n$ matrix. Then $A$ is symmetric if and only if $A$ is orthogonally equivalent to a real diagonal matrix.

Theorem 6.21. Let $A \in M_n(F)$ be a matrix whose characteristic polynomial splits over $F$.

1. If $F = \mathbb{C}$, then $A$ is unitarily equivalent to a complex upper triangular matrix.
2. If $F = \mathbb{R}$, then $A$ is orthogonally equivalent to a real upper triangular matrix.

Proof. (1) By Shur’s Theorem there is an orthonormal basis $\beta = \{v_1, v_2, \ldots, v_n\}$ such that $[L_A]_\beta = N$ where $N$ is a complex upper triangular matrix.

Let $\beta'$ be the standard ordered basis. Then $[L_A]_{\beta'} = A$. Let $Q = [I]_{\beta'}$. Then $N = Q^{-1} A Q$. We know that $Q$ is unitary since its columns are an orthonormal set of vectors and so $Q^* Q = I$. □
**Section 6.6**

**Defn 5.** If $V = W_1 \oplus W_2$, then a linear operator $T$ on $V$ is the **projection on $W_1$ along $W_2$** if whenever $x = x_1 + x_2$, with $x_1 \in W_1$ and $x_2 \in W_2$, we have $T(x) = x_1$. In this case,

$$R(T) = W_1 = \{ x \in V : T(x) = x \} \quad \text{and} \quad N(T) = W_2.$$ 

We refer to $T$ as the **projection**. Let $V$ be an inner product space, and let $T : V \to V$ be a projection. We say that $T$ is an **orthogonal projection** if $R(T) \perp = N(T)$ and $N(T)^\perp = R(T)$.

**Theorem 6.24.** Let $V$ be an inner product space, and let $T$ be a linear operator on $V$. Then $T$ is an orthogonal projection if and only if $T$ has an adjoint $T^*$ and $T^2 = T = T^*$.

Compare Theorem 6.24 to Theorem 6.9 where $V$ is finite-dimensional. This is the non-finite dimensional version.

**Theorem 6.25.** *(The Spectral Theorem).* Suppose that $T$ is a linear operator on a finite-dimensional inner product space $V$ over $F$ with the distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$. Assume that $T$ is normal if $F = \mathbb{C}$ and that $T$ is self-adjoint if $F = \mathbb{R}$. For each $i(1 \leq i \leq k)$, let $W_i$ be the eigenspace of $T$ corresponding to the eigenvalue $\lambda_i$, and let $T_i$ be the orthogonal projection of $V$ on $W_i$. Then the following statements are true.

(a) $V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$.

(b) If $W_i'$ denotes the direct sum of the subspaces $W_j$ for $j \neq i$, then $W_i' \perp = W_i'$.

(c) $T_i T_j = \delta_{i,j} T_i$ for $1 \leq i,j \leq k$.

(d) $I = T_1 + T_2 + \cdots + T_k$.

(e) $T = \lambda_1 T_1 + \lambda_2 T_2 + \cdots + \lambda_k T_k$.

**Proof.** Assume $F = \mathbb{C}$.

(a) $T$ is normal. By Theorem 6.16 there exists an orthonormal basis of eigenvectors of $T$. By Theorem 5.10, $V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$.

(b) Let $x \in W_i$ and $y \in W_j$, $i \neq j$.

Then $(x,y) = 0$ and so $W_i' = W_i \perp$.

But from (1),

$$\dim(W_i') = \sum_{j \neq i} \dim(W_j) = \dim(V) - \dim(W_i).$$

By Theorem 6.7(c), we know also that

$$\dim(W_i') = \dim(V) - \dim(W_i).$$

Hence $W_i' = W_i \perp$.

(c) $T_i$ is the orthogonal projection of $T$ on $W_i$. For $i \neq j$, $x \in V$, $x = w_1 + w_2 + \cdots + w_k$, $w_i \in W_i$.

$$T_i(x) = w_i$$

$$T_j T_i(x) = T_j(w_i) = 0$$

$$T_i(w_i) = w_i$$

$$T_i(T_i(x)) = T_i(x)$$
(d) 
\[(T_1 + T_2 + \cdots + T_k)(x) = T_1(x) + T_2(x) + \cdots + T_k(x) = w_1 + w_2 + \cdots + w_k = x\]

(e) Let \(x = T_1(x) + T_2(x) + \cdots + T_k(x)\).

So, \(T(x) = T(T_1(x)) + T(T_2(x)) + \cdots + T(T_k(x))\).

For all \(i\), \(T_i(x) \in W_i\). So
\[
T(T_i(x)) = \lambda_i T_i(x) = \lambda_i T_1(x) + \lambda_2 T_2(x) + \cdots + \lambda_k T_k(x) = (\lambda_1 T_1 + \lambda_2 T_2 + \cdots + \lambda_k T_k)(x).
\]
Defn 6. The set \( \{\lambda_1, \lambda_2, \ldots, \lambda_k\} \) of eigenvalues of \( T \) is called the **spectrum** of \( T \), the sum \( I = T_1 + T_2 + \cdots + T_k \) is called the **resolution of the identity operator** induced by \( T \), and the sum \( T = \lambda_1 T_1 + \lambda_2 T_2 + \cdots + \lambda_k T_k \) is called the **spectral decomposition** of \( T \).

Cor 1. If \( F = \mathbb{C} \), then \( T \) is normal if and only if \( T^* = g(T) \) for some polynomial \( g \).

Cor 2. If \( F = \mathbb{C} \), then \( T \) is unitary if and only if \( T \) is normal and \( |\lambda| = 1 \) for every eigenvalue \( \lambda \) of \( T \).

Cor 3. If \( F = \mathbb{C} \) and \( T \) is normal, then \( T \) is self-adjoint if and only if every eigenvalue of \( T \) is real.

Cor 4. Let \( T \) be as in the spectral theorem with spectral decomposition \( T = \lambda_1 T_1 + \lambda_2 T_2 + \cdots + \lambda_k T_k \). Then each \( T_j \) is a polynomial in \( T \).
We have some preliminary facts:

**Fact 1.** If $F$ is a linear operator on a finite-dimensional vector space $V$ such that $\text{rank}(F) = r$ and $\dim(V) = n$ then since $\text{rank}(F) + \text{nullity}(F) = n$, we have that $\text{nullity}(F) = n - r$. If $r > 0$, then there exists $x \neq 0$ such that $F(x) = 0 = 0 \cdot x$. So that 0 is an eigenvalue of $F$.

**Defn 7.** Let $T : V \to W$ be a linear transformation, where $V$ and $W$ are finite-dimensional inner product spaces with inner products $\langle \cdot , \cdot \rangle_1$ and $\langle \cdot , \cdot \rangle_2$, respectively. A function $T^* : W \to V$ is called an adjoint of $T$ if $\langle T(x), y \rangle_2 = \langle x, T^*(y) \rangle_1$ for all $x \in V$ and $y \in W$.

**Fact 2.** Let $T : V \to W$ be a linear transformation, where $V$ and $W$ are finite-dimensional inner product spaces with inner products $\langle \cdot , \cdot \rangle_1$ and $\langle \cdot , \cdot \rangle_2$, respectively. The following hold.

(a) There is a unique adjoint $T^*$ of $T$, and $T^*$ is linear.
(b) If $\beta$ and $\gamma$ are orthonormal bases for $V$ and $W$, respectively, then $[T^*]_\gamma^\beta = ([T]_\beta^\gamma)^*$.
(c) $\text{rank}(T^*) = \text{rank}(T)$.
(d) $\langle T^*(x), y \rangle_1 = \langle x, T(y) \rangle_2$ for all $x \in W$ and $y \in V$.
(e) For all $x \in V$, $T^*T(x) = 0$ if and only if $T(x) = 0$.

**Proof.**

(a) Suppose

$$\langle x, u \rangle_1 - \langle x, v \rangle_1 = 0$$

Therefore, $T^*(y)$ is unique.

By definition, $\langle T(x), az + y \rangle_2 = \langle x, T^*(az + y) \rangle_1$. But also,

$$\langle T(x), az + y \rangle_2 = \langle T(x), az \rangle_2 + \langle T(x), y \rangle_2$$

$$= a \langle T(x), z \rangle_2 + \langle T(x), y \rangle_2$$

$$= a \langle x, T^*(z) \rangle_1 + \langle x, T^*(y) \rangle_1$$

$$= \langle x, aT^*(z) \rangle_1 + \langle x, T^*(y) \rangle_1$$

So by uniqueness, $T^*(az + y) = aT^*(z) + T^*(y)$.

(d)

$$\langle x, T(y) \rangle_2 = \langle T(y), x \rangle_2$$

$$= \langle y, T^*(x) \rangle_1$$

$$= \langle T^*(x), y \rangle_1$$

(b) Let $A = [T^*]_\gamma^\beta$ and $B^* = ([T]_\beta^\gamma)^*$, where $\gamma = \{w_1, w_2, \ldots w_k\}$ and $\beta = \{v_1, v_2, \ldots, v_n\}$.

Since $W$ is an inner product space, and $\gamma$ is orthonormal, by Theorem 6.3, Corollary 1,

$$T(v_i) = \sum_{j=1}^k \langle T(v_i), w_j \rangle_2 w_j.$$
So that \( B_{j,i} = \langle T(v_i), w_j \rangle_2 \) and \( (B^*)_{i,j} = \langle T(v_i), w_j \rangle_2 = \langle w_j, T(v_i) \rangle_2 \).

Since \( V \) is an inner product space, and \( \beta \) is orthonormal, by Theorem 6.3, Corollary 1,
\[
T^*(w_j) = \sum_{i=1}^{n} \langle T^*(w_j), v_i \rangle_1 v_i.
\]

So that \( A_{i,j} = \langle T^*(w_j), v_i \rangle_2 \).

By (d), \( A_{i,j} = (B^*)_{i,j} \).

(c) \[
\text{rank}(T) = \text{rank}([T]_\beta^\gamma) = \text{rank}(([T]_\beta^\gamma)^*) = \text{rank}([T^*]_\beta^\gamma) = \text{rank}(T^*)
\]

(e) Given \( x \neq 0 \), \( T^*T(x) = 0 \) if and only if \( \langle x, T^*T(x) \rangle_1 = 0 \) which is true if and only if \( \langle T(x), T(x) \rangle_2 = 0 \) which is true if and only if \( T(x) = 0 \).

\[\square\]

**Defn 8.** A linear operator \( S \) on an inner-product space \( V \) is called **positive semidefinite** if \( S = S^* \) and \( \langle S(x), x \rangle \geq 0 \) for all \( x \neq 0 \). Notice that if there exists an \( x \neq 0 \), \( S(x) = \lambda x \), so that \( \lambda \) is an eigenvalue of \( S \), then \( 0 \leq \langle S(x), x \rangle = \langle \lambda x, x \rangle = \lambda \langle x, x \rangle \). We know that \( \langle x, x \rangle \geq 0 \) for all \( x \), therefore \( \lambda \geq 0 \).

**Fact 3.** Let \( V \) and \( W \) be finite-dimensional inner product spaces, and let \( T : V \to W \) be a linear transformation of rank \( r \). The linear operator, \( T^*T : V \to V \) is positive semidefinite and \( \text{rank}(T^*T) = \text{rank}(T) \).

**Proof.** \( (T^*T)^* = T^*(T^*)^* = T^*T \) so \( T^*T \) is self-adjoint and its eigenvalues are real. Also, \( \langle T^*T(x), x \rangle_1 = \langle T(x), T(x) \rangle_2 \geq 0, \forall x \). Therefore, \( T^*T \) is positive semidefinite.

By part (e) above, the nullity\((T^*T) = \text{nullity}(T) = \ell \). So, \( \text{rank}(T^*T) = n - \ell \) and \( \text{rank}(T) = n - \ell \). \[\square\]

**Theorem 6.26.** (Singular Value Theorem for Linear Transformations). Let \( V \) and \( W \) be finite-dimensional inner product spaces, and let \( T : V \to W \) be a linear transformation of rank \( r \). Then there exist orthonormal bases \( \{v_1, v_2, \ldots, v_n\} \) for \( V \) and \( \{u_1, u_2, \ldots, u_m\} \) for \( W \) and positive scalars \( \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r \) such that
\[
T(v_i) = \begin{cases} 
\sigma_i u_i & \text{if } 1 \leq i \leq r \\
0 & \text{if } i > r 
\end{cases}
\]

Conversely, suppose that the preceding conditions are satisfied. Then for \( 1 \leq i \leq n \), \( v_i \) is an eigenvector of \( T^*T \) with corresponding eigenvalue \( \sigma_i^2 \) if \( 1 \leq i \leq r \) and \( 0 \) if \( i > r \). Therefore the scalars \( \sigma_1, \sigma_2, \ldots, \sigma_r \) are uniquely determined by \( T \).

**Proof.** \( T^*T \) is positive semidefinite, its eigenvalues are real, and the non-zero ones are positive. Also, if \( \text{rank}(T^*T) = \text{rank}(T) = n \), so if \( r < n \), \( T^*T \) has zero as an eigenvalue.

By Theorem 6.16, if \( F \) is complex, there exists an orthonormal basis \( \beta \) of eigenvectors of \( T^*T \). And by Theorem 6.17, if \( F \) is real, there exists an orthonormal basis \( \beta \) of eigenvectors of \( T^*T \). So, in either case, we have \( \beta \). Suppose \( \beta = \{v_1, v_2, \ldots, v_n\} \), where, for each \( i \), \( v_i \) corresponds to \( \lambda_i \) and \( \lambda_1, \lambda_2, \ldots, \lambda_n \) are such that \( \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0 \). 


For $1 \leq i \leq r$, we define $\sigma_i = \sqrt{\lambda_i}$ and $u_i = \frac{1}{\sigma_i} T(v_i)$.

Claim: $\{u_1, u_2, \ldots, u_r\}$ is orthonormal.

\[
\langle u_i, u_j \rangle_2 = \langle \frac{1}{\sigma_i} T(v_i), \frac{1}{\sigma_j} T(v_j) \rangle_2 = \frac{1}{\sigma_i \sigma_j} \langle T(v_i), T(v_j) \rangle_2 = \frac{1}{\sigma_i \sigma_j} \langle T^* T(v_i), v_j \rangle_2 = \frac{1}{\sigma_i \sigma_j} \langle \lambda_i v_i, v_j \rangle_2 = \frac{\sigma_i^2}{\sigma_i \sigma_j} \langle v_i, v_j \rangle_2 = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}
\]

We extend this orthonormal set (which is linearly independent by Theorem 6.3, Corollary 2.) to a basis $\{u_1, u_2, \ldots, u_r, u_{r+1}, \ldots, u_m\}$ of $W$.

For $1 \leq i \leq r$, we have $\sigma_i u_i = \sigma_i \frac{1}{\sigma_i} T(v_i) = T(v_i)$.

And for $i > r$, $T^* T(v_i) = 0 = 0 \cdot v_i$ if and only if

\[
\langle T(v_i), T(v_i) \rangle_2 = \langle v_i, T^* T(v_i) \rangle_1 = \langle v_i, 0 \rangle_1 = 0
\]

if and only if $T(v_i) = 0$.

Conversely, suppose $\{v_1, v_2, \ldots, v_n\}$ is an orthonormal basis for $V$, $\{u_1, u_2, \ldots, u_m\}$ is an orthonormal basis for $W$, and $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ are scalars such that

\[
T(v_i) = \begin{cases} \sigma_i u_i & \text{if } 1 \leq i \leq r \\ 0 & \text{if } i > r \end{cases}
\]

Let $1 \leq i \leq m$ and $1 \leq j \leq n$.

\[
\langle T^*(u_i), v_j \rangle_1 = \langle u_i, T(v_j) \rangle_2 = \begin{cases} \langle u_i, \sigma_j u_j \rangle_2 & 1 \leq i \leq r \\ \langle u_i, 0 \rangle_2 & i > r \end{cases} = \begin{cases} \sigma_i & i = j, j \leq r \\ 0 & \text{otherwise} \end{cases}
\]

By Theorem 6.3, Corollary 1,

\[
T^*(u_i) = \sum_{j=1}^n \langle T^*(u_i), v_j \rangle v_j = \sum_{j=1}^r \langle T^*(u_i), v_j \rangle v_j = \begin{cases} \sigma_i v_i & i \leq r \\ 0 & i > r \end{cases}
\]

So for $i \leq r$,

\[
T^* T(v_i) = T^*(\sigma_i u_i) = \sigma_i T^*(u_i) = \sigma_i^2 u_i
\]
and \( i > r \), \[ T^*T(v_i) = T^*(0) = 0. \]

Therefore, \( v_i \) is an eigenvector of \( T^*T \) with corresponding eigenvalues \( \sigma_i^2 \) if \( i \leq r \) and 0 if \( i > r \).

\[ \square \]

**Defn 9.** The unique scalars \( \sigma_1, \sigma_2, \ldots, \sigma_r \) in Theorem 6.26 are called the singular values of \( T \). If \( r \) is less than both \( m \) and \( n \), then the term singular value is extended to include \( \sigma_{r+1} = \cdots = \sigma_k = 0 \), where \( k \) is the minimum of \( m \) and \( n \).

**Defn 10.** Let \( A \) be an \( m \times n \) matrix. We define the singular values of \( A \) to be the singular values of the linear transformation \( L_A \).

**Theorem 6.27.** (Singular Value Decomposition Theorem for Matrices). Let \( A \) be an \( m \times n \) matrix of rank \( r \) with the positive singular values \( \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r \), and let \( \Sigma \) be the \( m \times n \) matrix defined by

\[
\Sigma_{i,j} = \begin{cases} 
\sigma_i & \text{if } i = j \leq r \\
0 & \text{otherwise}
\end{cases}
\]

Then there exists an \( m \times n \) unitary matrix \( U \) and an \( n \times n \) unitary matrix \( V \) such that

\[ A = U \Sigma V^*. \]

**Proof.** Let \( T = L_A : F^n \to F^m \).

By Theorem 6.26, there exist orthonormal bases \( \beta = \{v_1, v_2, \ldots, v_n\} \) of \( F^n \) and \( \gamma = \{u_1, u_2, \ldots, u_m\} \) of \( F^m \) such that

\[ L_A(v_i) = \begin{cases} 
\sigma_i u_i & 1 \leq i \leq r \\
0 & i > r
\end{cases} \]

Let \( U \in M_m(F) \), where the \( j^{th} \) column of \( U \) is the \( j^{th} \) basis vector of \( \gamma \). And let \( V \in M_n(F) \), where the \( j^{th} \) column of \( V \) is the \( j^{th} \) basis vector of \( \beta \). It is easy to see that \( U^*U = UU^* = I \) and \( V^*V = VV^* = I \) so both \( U \) and \( V \) are unitary.

The \( j^{th} \) column of \( AV \) is \( Av_j = \sigma_j u_j \), for \( j \leq r \) and \( Av_j = 0 \) otherwise. The \( j^{th} \) column of \( \Sigma \) is \( \sigma_j e_j \) for \( j \leq r \) and \( 0 \) otherwise. So the \( j^{th} \) column of \( US \) is \( U(\sigma_j e_j) = \sigma_j U e_j = \sigma_j u_j \) for \( j \leq r \) and is \( 0 \) otherwise. Therefore, \( AV = U \Sigma \) which implies \( A = AVV^* = U \Sigma V^* \).

\[ \square \]

**Defn 11.** Let \( A \) be an \( m \times n \) matrix of rank \( r \) with positive singular values \( \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r \).

A factorization \( A = U \Sigma V^* \) where \( U \) and \( V \) are unitary matrices and \( \Sigma \) is the \( m \times n \) matrix defined as in Theorem 6.27 is called a singular value decomposition of \( A \).
**Section 7.1:**

**Defn 1.** A square matrix $A_i$ is called a **Jordan block** corresponding to $\lambda$ if it is of the form:

$$A_i = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{bmatrix}$$

**Defn 2.** A square matrix $[T]_\beta$ is called a **Jordan canonical form** of $T$ if it has the following form, where each $A_i$ is a Jordan block and each $O$ is the zero matrix.

$$[T]_\beta = \begin{bmatrix} A_1 & O & O & \cdots & O & O \\ O & A_2 & O & \cdots & O & O \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ O & O & O & \cdots & A_{k-1} & O \\ O & O & O & \cdots & O & A_k \end{bmatrix}$$

The basis $\beta$ used to form the Jordan canonical form of $T$ is called the **Jordan canonical basis** for $T$.

**Defn 3.** Let $T$ be a linear operator on a vector space $V$, and let $\lambda$ be a scalar. A nonzero vector $x$ in $V$ is called a **generalized eigenvector** of $T$ corresponding to $\lambda$ if $(T - \lambda I)^p(x) = 0$ for some positive integer $p$.

**Defn 4.** Let $T$ be a linear operator on a vector space $V$, and let $\lambda$ be an eigenvalue of $T$. The **generalized eigenspace** of $T$ corresponding to $\lambda$, denoted $K_\lambda$, is the subset of $V$ defined by

$$K_\lambda = \{ x \in V : \exists p \in \mathbb{Z}, (T - \lambda I)^p(x) = 0 \}$$

**Theorem 7.1.** Let $T$ be a linear operator on a vector space $V$, and let $\lambda$ be an eigenvalue of $T$. Then

1. $K_\lambda$ is a $T$-invariant subspace of $V$ containing $E_\lambda$.
2. For any scalar $\mu \neq \lambda$, the restriction of $T - \mu I$ to $K_\lambda$ is one-to-one.

**Proof.** (1) Let $x \in E_\lambda$, $x \neq 0$. $x$ an eigenvector of $T$ corresponding to $\lambda$. We have $(T - \lambda I)(x) = 0$, so $x \in K_\lambda$ and $E_\lambda \subseteq K_\lambda$.

Let $x \in K_\lambda$ then $\exists p \in \mathbb{Z}$ such that $(T - \lambda I)^p(x) = 0$. Let $y \in K_\lambda$ then $\exists q \in \mathbb{Z}$ such that $(T - \lambda I)^q(y) = 0$. Let $a \in F$. Let $s = \max\{p, q\}$. Then $(T - \lambda I)^s(x) = (T - \lambda I)^s(y) = 0$.

So $(T - \lambda I)^s(ax + y) = a(T - \lambda I)^s(x) + (T - \lambda I)^s(y) = 0$. Therefore, $K_\lambda$ is a subspace of $V$.

Now we show $K_\lambda$ is $T$-invariant. Let $x \in K_\lambda$ then $\exists p \in \mathbb{Z}$ such that $(T - \lambda I)^p(x) = 0$. By the Binomial Theorem,

$$(T - \lambda I)^p = \sum_{i=0}^{p} \binom{p}{i} T^{p-i}(-\lambda I)^i$$
Theorem 7.2. Let $\mu \neq \lambda, \mu \in F$.

Let $x, y \in K_\lambda$ such that $(T - \mu I)(x) = (T - \mu I)(y)$. We will show that $x = y$.

Suppose not. Then $x - y \neq 0$ and $(T - \mu I)(x - y) = 0$. So, $x - y \in E_\mu$ and $x - y \not\in E_\lambda$.

But, $x - y \in K_\lambda$, so $p > 1$ be minimum such that $(T - \lambda I)^p(x - y) = 0$.

Let $z = (T - \lambda I)^{p-1}(x - y)$. Then $(T - \lambda I)(z) = 0$ which implies that $z \in E_\lambda$ and $z \neq 0$ implies that $z$ is an eigenvector of $T$ corresponding to $\lambda$.

On the other hand,

$$(T - \mu I)(z) = (T - \mu I)(T - \lambda I)^{p-1}(x - y) = (T - \lambda I)^{p-1}(T - \mu I)(x - y) = 0$$

So $z \in E_\mu$ but this contradicts that $E_\lambda \cap E_\mu = \{0\}$.

\[ \square \]

Theorem 7.2. Let $T$ be a linear operator on a finite-dimensional vector space $V$ such that the characteristic polynomial of $T$ splits. Suppose that $\lambda$ is an eigenvalue of $T$ with multiplicity of $m$. Then

(a) $\dim(K_\lambda) \leq m$.

(b) $K_\lambda = N((T - \lambda I)^m)$.

Proof. (a) By Theorem 7.1, $K_\lambda$ is $T$-invariant. Let $h(t)$ be the characteristic polynomial of $T_{K_\lambda}$. By Theorem 5.21, $h(t)$ divides the characteristic polynomial $f(t)$ of $T$.

By Theorem 5.21, if $\mu \neq \lambda$ and $x \in K_\lambda$ then $(T - \mu I)(x) = 0$ if and only if $x = 0$.

Hence, $\lambda$ is the only eigenvalue of $K_\lambda$ and $h(t) = (-1)^p(t - \lambda)^p$, where $p = \dim K_\lambda$.

Therefore, $p \leq m$.

(b) Clearly, $N((T - \lambda I)^m) \subseteq K_\lambda$.

As before, let $h(t)$ be the characteristic polynomial of $T_{K_\lambda}$ By the Cayley-Hamilton Theorem, $h(T_{K_\lambda}) = 0$ on $K_\lambda$. So,

$$(-1)^p(T_{K_\lambda} - \lambda I)^p(x) = 0, \ \forall x \in K_\lambda$$

But on $K_\lambda$, $T_{K_\lambda} = T$. So we have

$$x \in N((T - \lambda I)^p), \ \forall x \in K_\lambda$$

and since $p \leq m$, we have,

$$x \in N((T - \lambda I)^m), \ \forall x \in K_\lambda.$$

\[ \square \]

Theorem 7.3. Let $T$ be a linear operator on a finite-dimensional vector space $V$ such that the characteristic polynomial of $T$ splits, and let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be the distinct eigenvalues of $T$. Then, for every $x \in V$, there exist vectors $v_i \in K_{\lambda_i}, 1 \leq i \leq k$, such that

$$x = v_1 + v_2 + \cdots + v_k.$$
Theorem 7.4. Let $T$ be a linear operator on a finite-dimensional vector space $V$ such that the characteristic polynomial of $T$ splits, and let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be the distinct eigenvalues of $T$ corresponding to the multiplicities $m_1, m_2, \ldots, m_k$. For $1 \leq i \leq k$, let $\beta_i$ be an ordered basis for $K_{\lambda_i}$. Then the following statements are true.

(a) $\beta_i \cap \beta_j = \emptyset$ for $i \neq j$.
(b) $\beta = \beta_1 \cup \beta_2 \cup \cdots \cup \beta_k$ is an ordered basis for $V$.
(c) $\dim(K_{\lambda_i}) = m_i$, for all $i$. 
Cor 5. Let $T$ be a linear operator on a finite-dimensional vector space $V$ such that the characteristic polynomial of $T$ splits. Then $T$ is diagonalizable if and only if $E_\lambda = K_\lambda$ for every eigenvalue $\lambda$ of $T$. 
**Defn 5.** Let $T$ be a linear operator on a vector space $V$, and let $x$ be a generalized eigenvector of $T$ corresponding to the eigenvalue $\lambda$. Suppose that $p$ is the smallest positive integer for which $(T - \lambda I)^p(x) = 0$. Then the ordered set
\[
\{(T - \lambda I)^{p-1}(x), (T - \lambda I)^{p-2}(x), \ldots, (T - \lambda I)(x)\}
\]
is called a cycle of generalized eigenvectors of $T$ corresponding to $\lambda$. The vectors $(T - \lambda I)^{p-1}(x)$ and $x$ are called the initial vector and the end vector of the cycle, respectively. We say that the length of the cycle is $p$.

**Theorem 7.5.** Let $T$ be a linear operator on a finite-dimensional vector space $V$ whose characteristic polynomial splits, and suppose that $\beta$ is a basis for $V$ such that $\beta$ is a disjoint union of cycles of generalized eigenvectors of $T$. Then the following statements are true.

(a) For each cycle $\gamma$ of generalized eigenvectors contained in $\beta$, $W = \text{span}(\gamma)$ is $T$-invariant, and $[T_W]_{\gamma}$ is a Jordan block.

(b) $\beta$ is a Jordan canonical basis for $V$. 

**Theorem 7.6.** Let $T$ be a linear operator on a vector space $V$, and let $\lambda$ be an eigenvalue of $T$. Suppose that $\gamma_1, \gamma_2, \ldots, \gamma_q$ are cycles of generalized eigenvectors of $T$ corresponding to $\lambda$ such that the initial vectors of the $\gamma_i$’s are distinct and form a linearly independent set. Then the $\gamma_i$’s are disjoint, and their union $\gamma = \bigcup_{i=1}^{q} \gamma_i$ is linearly independent.
Cor 1. Every cycle of generalized eigenvectors of a linear operator is linearly independent.
Theorem 7.7. Let $T$ be a linear operator on a finite-dimensional vector space $V$, and let $\lambda$ be an eigenvalue of $T$. Then $K_\lambda$ has an ordered basis consisting of a union of disjoint cycles of generalized eigenvectors corresponding to $\lambda$. 
**Cor 1.** Let $T$ be a linear operator on a finite-dimensional vector space $V$ whose characteristic polynomial splits. Then $T$ has a Jordan canonical form.

**Defn 6.** Let $A \in M_n(F)$ be such that the characteristic polynomial of $A$ (and hence of $L_A$) splits. Then the Jordan canonical form of $A$ is defined to be the Jordan canonical form of the linear operator $L_A$ on $F^n$.

**Cor 2.** Let $A$ be an $n \times n$ matrix whose characteristic polynomial splits. Then $A$ has a Jordan canonical form $J$, and $A$ is similar to $J$.

**Theorem 7.8.** Let $T$ be a linear operator on a finite-dimensional vector space $V$ whose characteristic polynomial splits. Then $V$ is the direct sum of the generalized eigenspaces of $T$.

(Compare to Theorem 5.11)
Section 7.2. For the purposes of this section, we fix a linear operator $T$ on an $n$-dimensional vector space $V$ such that the characteristic polynomial of $T$ splits. Let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be the distinct eigenvalues of $T$.

By Theorem 7.7, each generalized eigenspace $K_{\lambda_i}$ contains an ordered basis $\beta_i$ consisting of a union of disjoint cycles of generalized eigenvectors corresponding to $\lambda_i$. So by Theorems 7.4(b) and 7.5, the union $\beta = \cup \beta_i$ is a Jordan canonical basis for $T$. For each $i$, let $T_i$ be the restriction of $T$ to $K_{\lambda_i}$, and let $A_i = [T_i]_{\beta_i}$. Then $A_i$ is the Jordan canonical form of $T_i$, and

$$J = [T]_\beta = \begin{bmatrix} A_1 & O & \cdots & O \\ O & A_2 & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & A_k \end{bmatrix}$$

is the Jordan canonical form of $T$. In this matrix, each $O$ is a zero matrix of appropriate size.

**Defn 1.** We adopt the following convention: The basis $\beta_i$ for $K_{\lambda_i}$ will henceforth be ordered in such a way that the cycles appear in order of decreasing length. That is, if $\beta_i$ is a disjoint union of cycles $\gamma_1, \gamma_2, \ldots, \gamma_{n_i}$ and if the length of the cycle $\gamma_j$ is $p_j$, we index the cycles so that $p_1 \geq p_2 \geq \cdots \geq p_{n_i}$. This ordering of the cycles limits the possible orderings of vectors in $\beta_i$, which in turn determines the matrix $A_i$. It is in this sense that $A_i$ is unique. It then follows that the Jordan canonical form for $T$ is unique up to an ordering of the eigenvalues of $T$.

**Defn 2.** Let $T_i$ be the restriction of $T$ to $K_{\lambda_i}$. A dot diagram of $T_i$ will aid in the visualization of the matrix $A_i$. Suppose that $\beta_i$ is a disjoint union of cycles of generalized eigenvectors $\gamma_1, \gamma_2, \ldots, \gamma_{n_i}$ with lengths $p_1 \geq p_2 \geq \cdots \geq p_{n_i}$, respectively. The dot diagram contains one dot for each vector in $\beta_i$, and the dots are configured according to the following rules.

1. The array consists of $n_i$ columns (one column for each cycle.)
2. Counting from left to right, the $j^{th}$ column consists of the $p_j$ dots that correspond to the vectors of $\gamma_j$ starting with the initial vector at the top and continuing down to the end vector.

Denote the end vectors of the cycles by $v_1, v_2, \ldots, v_{n_i}$. In the following dot diagram of $T_i$, each dot is labeled with the name of the vector in $\beta_i$ to which it corresponds.

- $(T - \lambda_i I)^{p_1 - 1}(v_1)$
- $(T - \lambda_i I)^{p_2 - 1}(v_2)$
- \(\vdots\)
- $(T - \lambda_i I)^{p_{n_i} - 1}(v_{n_i})$
- $(T - \lambda_i I)^{p_1 - 2}(v_1)$
- $(T - \lambda_i I)^{p_2 - 2}(v_2)$
- \(\vdots\)
- $(T - \lambda_i I)^{p_{n_i} - 2}(v_{n_i})$
- $(T - \lambda_i I)(v_{n_i})$
- $v_{n_i}$
- $(T - \lambda_i I)(v_2)$
- $v_2$
- $(T - \lambda_i I)(v_1)$
- $v_1$

**Theorem 7.9.** We fix a basis $\beta_i$ of $K_{\lambda_i}$ so that $\beta_i$ is a disjoint union of $n_i$ cycles of generalized eigenvectors with lengths $p_1 \geq p_2 \geq \cdots \geq p_{n_i}$. For any positive integer $r$, the vectors in $\beta_i$ that are associated with the dots in the first $r$ rows of the dot diagram of $T_i$ constitute a basis.
for \( N((T - \lambda_i I)^r) \). Hence the number of dots in the first \( r \) rows of the dot diagram equals nullity\((T - \lambda_i I)^r\).

**Cor 1.** The dimension of \( E_{\lambda_i} \) is \( n_i \). Hence in a Jordan canonical form of \( T \), the number of Jordan blocks corresponding to \( \lambda_i \) equals the dimension of \( E_{\lambda_i} \).

**Theorem 7.10.** Let \( r_i \) denote the number of dots in the \( j \)th row of the dot diagram of \( T_i \), the restriction of \( T \) to \( K_{\lambda_i} \). Then the following statements are true.

(a) \( r_1 = \dim(V) - \text{rank}(T - \lambda_i I) \).

(b) \( r_j = \text{rank}((T - \lambda_i I)^{j-1}) - \text{rank}((T - \lambda_i I)^j) \), if \( j > 1 \).

**Cor 1.** For any eigenvalue \( \lambda_i \) of \( T \), the dot diagram of \( T_i \) is unique. Thus, subject to the convention that the cycles of generalized eigenvectors for the bases of each generalized eigenspace are listed in order of decreasing length, the Jordan canonical form of a linear operator or a matrix is unique up to the ordering of the eigenvalues.

**Theorem 7.11.** Let \( A \) and \( B \) be \( n \times n \) matrices, each having Jordan canonical forms computed according to the conventions of this section. Then \( A \) and \( B \) are similar if and only if they have (up to an ordering of their eigenvalues) the same Jordan canonical form.

**Section 7.3.**

**Defn 1.** Let \( T \) be a linear operator on a finite-dimensional vector space. A polynomial \( p(t) \) is called a **minimal polynomial** of \( T \) if \( p(T) = 0 \).

**Theorem 7.12.** Let \( p(t) \) be a minimal polynomial of linear operator \( T \) on a finite-dimensional vector space \( V \).

(a) For any polynomial \( g(t) \), if \( g(T) = 0 \), then \( p(t) \) divides \( g(t) \). In particular, \( p(t) \) divides the characteristic polynomial of \( T \).

(b) The minimal polynomial of \( T \) is unique.

**Defn 2.** Let \( A \in M_n(F) \). The **minimal polynomial** \( p(t) \) of \( A \) is the monic polynomial of least positive degree for which \( p(A) = 0 \).

**Theorem 7.13.** Let \( T \) be a linear operator on a finite-dimensional vector space \( V \), and let \( \beta \) be an ordered basis for \( V \). Then the minimal polynomial of \( T \) is the same as the minimal polynomial of \([T]_\beta\).

**Cor 1.** For any \( A \in M_n(F) \), the minimal polynomial of \( A \) is the same as the minimal polynomial of \( L_A \).

**Theorem 7.14.** Let \( T \) be a linear operator on a finite-dimensional vector space \( V \), and let \( p(t) \) be the minimal polynomial of \( T \). A scalar \( \lambda \) is an eigenvalue of \( T \) if and only if \( p(\lambda) = 0 \). Hence the characteristic polynomial and the minimal polynomial of \( T \) have the same zeros.

**Cor 1.** Let \( T \) be a linear operator on a finite-dimensional vector space \( V \) with minimal polynomial \( p(t) \) and characteristic polynomial \( f(t) \). Suppose that \( f(t) \) factors as

\[
 f(t) = (\lambda_1 - t)^{m_1}(\lambda_2 - t)^{m_2} \cdots (\lambda_k - t)^{m_k},
\]

where \( \lambda_1, \lambda_2, \ldots, \lambda_k \) are the distinct eigenvalues of \( T \). Then there exist integers \( m_1, m_2, \ldots, m_k \) such that \( 1 \leq m_i \leq n_i \), for all \( i \) and

\[
 p(t) = (t - \lambda_1)^{m_1}(t - \lambda_2)^{m_2} \cdots (t - \lambda_k)^{m_k}.
\]
**Theorem 7.15.** Let $T$ be a linear operator on a finite-dimensional vector space $V$ such that $V$ is a $T$-cyclic subspace of itself. Then the characteristic polynomial $f(t)$ and the minimal polynomial $p(t)$ have the same degree, and hence $f(t) = (-1)^n p(t)$.

**Theorem 7.16.** Let $T$ be a linear operator on a finite-dimensional vector space $V$. Then $G$ is diagonalizable if and only if the minimal polynomial of $T$ is of the form

$$p(t) = (t - \lambda_1)(t - \lambda_2) \cdots (t - \lambda_k)$$

where $\lambda_1, \lambda_2, \ldots, \lambda_k$ are the distinct eigenvalues of $T$. 