Section 6.3:

**Theorem 6.8.** Let $V$ be a finite dimensional inner product space over $F$, and let $g : V \rightarrow F$ be a linear functional. Then, there exists a unique vector $y \in V$ such that $g(x) = \langle x, y \rangle, \forall x \in V$.

**Proof.** Let $\beta = \{v_1, v_2, \ldots, v_n\}$ be an orthonormal basis for $V$ and let

$$ y = \sum_{i=1}^{n} g(v_i)v_i. $$

Then for $1 \leq j \leq n$,

$$ \langle v_j, y \rangle = \langle v_j, \sum_{i=1}^{n} g(v_i)v_i \rangle $$

$$ = \sum_{i=1}^{n} g(v_i)\langle v_j, v_i \rangle $$

$$ = g(v_j)\langle v_j, v_j \rangle $$

$$ = g(v_j) $$

and we have, $\forall x \in V, g(x) = \langle x, y \rangle$.

To show $y$ is unique, suppose $g(x) = \langle x, y' \rangle, \forall x$. Then, $\langle x, y \rangle = \langle x, y' \rangle, \forall x$. By Theorem 6.1(e), we have $y = y'$.

**Example 1.** (2b) Let $V = \mathbb{C}^2$, $g(z_1, z_2) = z_1 - 2z_2$. Then $V$ is an inner-product space with the standard inner product $\langle (x_1, x_2), (y_1, y_2) \rangle = x_1\overline{y_1} + x_2\overline{y_2}$ and $g$ is a linear operator on $V$. Find a vector $y \in V$ such that $g(x) = \langle x, y \rangle, \forall x \in V$.

**Sol:** We need to find $(y_1, y_2) \in \mathbb{C}^2$ such that

$$ g(z_1, z_2) = \langle (z_1, z_2), (y_1, y_2) \rangle, \forall (z_1, z_2) \in \mathbb{C}^2. $$

That is:

$$ z_1 - 2z_2 = z_1\overline{y_1} + z_2\overline{y_2}. $$

Using the standard ordered basis $\{(1, 0), (0, 1)\}$ for $\mathbb{C}^2$, the proof of Theorem 6.8 gives that

$$ y = \sum_{i=1}^{n} g(v_i)v_i. $$

So,

$$ (y_1, y_2) = \frac{g(1, 0)(1, 0) + g(0, 1)(0, 1)}{\overline{y_1}(1, 0) + 2\overline{y_2}(0, 1)} $$

$$ = (1, 0) - 2(0, 1) $$

$$ = (1, -2) $$

Check (1): LHS = $z_1 - 2z_2$ and for $\overline{y_1} = 1$ and $\overline{y_2} = -2$, we have RHS = $z_1 - 2z_2$.  

1
Theorem 6.9. Let $V$ be a finite dimensional inner product space and let $T$ be a linear operator on $V$. There exists a unique operator $T^* : V \to V$ such that $\langle T(x), y \rangle = \langle x, T^*(y) \rangle$, $\forall x, y \in V$. Furthermore, $T^*$ is linear.

Proof. Let $y \in V$

Define $g : V \to F$ as $g(x) = \langle T(x), y \rangle, \forall x \in V$.

Claim: $g$ is linear.

$$g(ax + z) = \langle T(ax + z), y \rangle$$
$$= \langle aT(x) + T(z), y \rangle$$
$$= \langle aT(x), y \rangle + \langle T(z), y \rangle$$
$$= ag(x) + g(z).$$

By Theorem 6.8, there is a unique $y' \in V$ such that $g(x) = \langle x, y' \rangle$.

So we have $\langle T(x), y \rangle = \langle x, y' \rangle, \forall x \in V$.

We define $T^* : V \to V$ by $T^*(y) = y'$. So $\langle T(x), y \rangle = \langle x, T^*(y) \rangle, \forall x \in V$

Claim: $T^*$ is linear. We have for $cy + z \in V$, $T^*(cy + z)$ equals the unique $y'$ such that $\langle T(x), cy + z \rangle = \langle x, T^*(cy + z) \rangle$. But,

$$\langle T(x), cy + z \rangle = \overline{c} \langle T(x), y \rangle + \langle T(x), z \rangle$$
$$= \overline{c} \langle x, T^*(y) \rangle + \langle x, T^*(z) \rangle$$
$$= \langle x, cT^*(y) + T^*(z) \rangle$$

Since $y'$ is unique, $T^*(cy + z) = cT^*(y) + T^*(z)$.

Claim: $T^*$ is unique.

Let $U : V \to V$ be linear such that

$$\langle T(x), y \rangle = \langle x, U(y) \rangle, \forall x, y \in V$$
$$\langle x, T^*(y) \rangle = \langle x, U(y) \rangle, \forall x, y \in V$$
$$= T^*(y) = U(y), \forall y \in V$$

So, $T^* = U$. □

Defn 1. $T^*$ is called the adjoint of the linear operator $T$ and is defined to be the unique operator on $V$ satisfying $\langle T(x), y \rangle = \langle x, T^*(y) \rangle, \forall x, y \in V$.

Fact 1. $\langle x, T(y) \rangle = \langle T^*(x), y \rangle, \forall x, y \in V$.

Proof.

$$\langle x, T(y) \rangle = \overline{\langle T(y), x \rangle}$$
$$= \overline{\langle y, T^*(x) \rangle}$$
$$= \langle T^*(x), y \rangle$$ □
Theorem 6.10. Let $V$ be a finite dimensional inner product space and let $\beta$ be an orthonormal basis for $V$. If $T$ is a linear operator on $V$, then

$$[T^*]_\beta = [T]_\beta^*$$

Proof. Let $A = [T]_\beta$ and $B = [T^*]_\beta$ with $\beta = \{v_1, v_2, \ldots, v_n\}$.

By the corollary to Theorem 6.5,

$$(B)_{i,j} = \langle T^*(v_j), v_i \rangle$$

$$= \langle v_i, T^*(v_j) \rangle$$

$$= \langle T(v_i), v_j \rangle$$

$$= (A)_{j,i}$$

$$= (A^*)_{i,j}$$

□

Cor 1. Let $A$ be an $n \times n$ matrix, then $L_{A^*} = (L_A)^*$.

Proof. Use $\beta$, the standard ordered basis. By Theorem 2.16,

(2) $[L_A]_\beta = A$

(3) $L_{A^*} = A^*$

By Theorem 6.10, $[L_A^*]_\beta = [L_A]^*_\beta$, which equals $A^*$ by (2) and $[L_{A^*}]_\beta$ by (3).

□
Theorem 6.11. Let $V$ be an inner product space, $T, U$ linear operators on $V$. Then

1. $(T + U)^* = T^* + U^*$.
2. $(cT)^* = cT^*$, $\forall c \in F$.
3. $(TU)^* = U^*T^*$ (composition)
4. $(T^*)^* = T$
5. $I^* = I$.

Proof. (1) 
\[
\langle (T + U)(x), y \rangle = \langle x, (T + U)^*(y) \rangle
\]
and
\[
\langle (T + U)(x), y \rangle = \langle T(x), y \rangle + \langle U(x), y \rangle
= \langle x, T^*(y) \rangle + \langle x, U^*(y) \rangle
= \langle x, (T^* + U^*)(y) \rangle
\]
And $(T + U)^*$ is unique, so it must equal $T^* + U^*$.

(2) 
\[
\langle cT(x), y \rangle = \langle x, (cT)^*(y) \rangle
\]
and
\[
\langle cT(x), y \rangle = c\langle T(x), y \rangle
= c\langle x, T^*(y) \rangle
= \langle x, cT^*(y) \rangle
\]

(3) 
\[
\langle TU(x), y \rangle = \langle x, (TU)^*(y) \rangle
\]
and
\[
\langle TU(x), y \rangle = \langle U(T(x)), y \rangle
= \langle U(x), T^*(y) \rangle
= \langle x, U^*(T^*(y)) \rangle
= \langle x, (U^*T^*)(y) \rangle
\]

(4) 
\[
\langle T^*(x), y \rangle = \langle x, (T^*)^*(y) \rangle
\]
by definition and
\[
\langle T^*(x), y \rangle = \langle x, T(y) \rangle
\]
by Fact 1.

(5) 
\[
\langle I^*(x), y \rangle = \langle x, I(y) \rangle = \langle x, y \rangle, \forall x, y \in V
\]
Therefore, $I^*(x) = x$, $\forall x \in V$ and we have $I^* = I$. 

□
Cor 1. Let $A$ and $B$ be $n \times n$ matrices. Then

1. $(A + B)^* = A^* + B^*$.
2. $(cA)^* = \overline{c}A^*, \forall c \in F$.
3. $(AB)^* = B^*A^*$
4. $(A^*)^* = A$
5. $I^* = I$

Proof. Use Theorem 6.11 and Corollary to Theorem 6.10. Or, use below. \hfill \Box

Example 2. (Exercise 5b) Let $A$ and $B$ be $m \times n$ matrices and $C$ an $n \times p$ matrix. Then

1. $(A + B)^* = A^* + B^*$.
2. $(cA)^* = \overline{c}A^*, \forall c \in F$.
3. $(AC)^* = C^*A^*$
4. $(A^*)^* = A$
5. $I^* = I$

Proof. (1) 
\[
(A + B)_{i,j} = \overline{(A + B)_{j,i}} = (A)_{j,i} + (B)_{j,i} = \overline{(A)_{j,i}} + (B)_{j,i}
\]

and
\[
(A^* + B^*)_{i,j} = (A^*)_{i,j} + (B^*)_{i,j} = \overline{(A)_{j,i}} + (B)_{j,i}
\]

(2) Let $c \in F$.
\[
(cA)_{i,j} = \overline{(cA)_{j,i}} = c(A)_{j,i} = \overline{c(A)_{j,i}}
\]

and
\[
(\overline{c}A^*)_{i,j} = \overline{c}(A^*)_{i,j} = \overline{(A)_{j,i}}
\]

(3) 
\[
((AC)^*)_{i,j} = \overline{(AC)_{j,i}}
\]
\[
= \sum_{k=1}^{n} (A)_{j,k}(C)_{k,i}
\]
\[
= \sum_{k=1}^{n} (A^*)_{k,j}(C^*)_{i,k}
\]
\[
= \sum_{k=1}^{n} (C^*)_{i,k}(A^*)_{k,j}
\]
\[
= (C^*A^*)_{i,j}
\]

\hfill \Box
For \( x, y \in F^n \), let \( \langle x, y \rangle_n \) denote the standard inner product of \( x \) and \( y \) in \( F^n \). Recall that if \( x \) and \( y \) are regarded as column vectors, then \( \langle x, y \rangle_n = y^*x \).

**Lemma 1.** Let \( A \in M_{m \times n}(F) \), \( x \in F^n \), and \( y \in F^m \). Then
\[
\langle Ax, y \rangle_m = \langle x, A^*y \rangle_n.
\]

**Proof.**
\[
\langle Ax, y \rangle_m = y^*(Ax) = (y^*A)x = (A^*y)^*x = \langle x, A^*y \rangle_n.
\]

**Lemma 2.** Let \( A \in M_{m \times n}(F) \), \( x \in F^n \). Then \( \text{rank}(A^*A) = \text{rank}(A) \).

**Proof.** \( A^*A \) is an \( n \times n \) matrix. By the Dimension Theorem, \( \text{rank}(A^*A) + \text{nullity}(A^*A) = n \). We also have, \( \text{rank}(A) + \text{nullity}(A) = n \).

We will show that the nullspace of \( A \) equals the nullspace of \( A^*A \). We will show \( A^*Ax = 0 \) if and only if \( Ax = 0 \).

\[
0 = A^*Ax
\]
\[
\iff 0 = \langle A^*Ax, x \rangle_n
\]
\[
\iff 0 = \langle Ax, A^*x \rangle_m
\]
\[
\iff 0 = \langle Ax, Ax \rangle_m
\]
\[
\iff 0 = Ax
\]

**Cor 1.** If \( A \) is an \( m \times n \) matrix such that \( \text{rank}(A) = n \), then \( A^*A \) is invertible.

**Theorem 6.12.** Let \( A \in M_{m \times n}(F) \) and \( y \in F^m \). Then there exists \( x_0 \in F^n \) such that \( (A^*A)x_0 = A^*y \) and \( \|Ax_0 - y\| \leq \|Ax - y\| \) for all \( x \in F^n \). Furthermore, if \( \text{rank}(A) = n \), then \( x_0 = (A^*A)^{-1}A^*y \).

**Proof.** Define \( W = \{Ax : x \in F^n\} = R(L_A) \). By the corollary to Theorem 6.6, there is a unique vector \( u = Ax_0 \) in \( W \) that is “closest” to \( y \). Then, \( \|Ax_0 - y\| \leq \|Ax - y\| \) for all \( x \in F^n \). Also by Theorem 6.6, \( z = y - u \) is in \( W^\perp \). So, \( -z = Ax_0 - y \) is in \( W^\perp \). So,
\[
\langle Ax, Ax_0 - y \rangle_m = 0, \forall x \in F^n.
\]

By Lemma 1,
\[
\langle x, A^*(Ax_0 - y) \rangle_n = 0, \forall x \in F^n.
\]
So, \( A^*(Ax_0 - y) = 0 \). We see that \( x_0 \) is the solution for \( x \) in \( A^*Ax = A^*y \). If, in addition, we know that \( \text{rank}(A) = n \), then by Lemma 2, we have that \( \text{rank}(A^*A) = n \) and is therefore invertible. So,
\[
x_0 = (A^*A)^{-1}A^*y.
\]