1 Introduction

In this chapter we study the additional structures that a vector space over field of reals or complex vector spaces have. So, in this chapter, $\mathbb{R}$ will denote the field of reals, $\mathbb{C}$ will denote the field of complex numbers, and $\mathbb{F}$ will denote one of them.

In my view, this is where algebra drifts out to analysis.

1.1 (Definition) Let $\mathbb{F}$ be the field reals or complex numbers and $V$ be a vector space over $\mathbb{F}$. An inner product on $V$ is a function

$$(\cdot, \cdot) : V \times V \longrightarrow \mathbb{F}$$

such that

1. $(ax + by, z) = a(x, z) + b(y, z)$, for $a, b \in \mathbb{F}$ and $x, y, z \in V$.

2. $(x, y) = (y, x)$ for $x, y \in V$.

3. $(x, x) > 0$ for all non-zero $x \in V$

4. Also define $\|x\| = \sqrt{(x, x)}$. This is called norm of $x$.

Comments: Real Case: Assume $\mathbb{F} = \mathbb{R}$. Then

1. Item (2) means $(x, y) = (y, x)$. 


2. Also (1 and 2) means that the inner product is **bilinear**.

**Comments:** **Complex Case:** Assume \( F = \mathbb{C} \). Then

1. Items (1 and 2) means that the \((x, cy + dz) = \overline{c}(x, y) + \overline{d}(x, z)\).

1.2 (Example) On \( \mathbb{R}^n \) we have the standard inner product defined by
\[
(x, y) = \sum_{i=1}^{n} x_i y_i,
\]
where \( x = (x_1, \ldots, x_n) \in \mathbb{R}^n \) and \( y = (y_1, \ldots, y_n) \in \mathbb{R}^n \).

1.3 (Example) On \( \mathbb{C}^n \) we have the standard inner product defined by
\[
(x, y) = \sum_{i=1}^{n} x_i \overline{y_i},
\]
where \( x = (x_1, \ldots, x_n) \in \mathbb{C}^n \) and \( y = (y_1, \ldots, y_n) \in \mathbb{C}^n \).

1.4 (Example) Let \( F = \mathbb{R} \) OR \( \mathbb{C} \) and \( V = M_n(F) \). For \( A = (a_{ij}), B = (b_{ij}) \in V \) define inner product
\[
(A, B) = \sum_{i,j} a_{ij} \overline{b_{ij}}.
\]
Define conjugate transpose \( B^* = (B)^t \). Then
\[
(A, B) = \text{trace}(B^*A)
\]

1.5 (Example: Integration) Let \( F = \mathbb{R} \) OR \( \mathbb{C} \) and \( V \) be the vector space of all \( F \)-valued continuous functions on \([0, 1]\). For \( f, g \in V \) define

\[
(f, g) = \int_0^1 f \overline{g} dt.
\]
This is an inner product on \( V \). In some distant future this will be called \( L^2 \) inner product space. This can be done in any "space" where you have an idea of integration and it will come under Measure Theory.

1.6 (Matrix of Inner Product) Let \( F = \mathbb{R} \) OR \( \mathbb{C} \). Suppose \( V \) is a vector space over \( F \) with an inner product. Let \( e_1, \ldots, e_n \) be a basis of \( V \). Let \( p_{i,j} = (e_i, e_j) \) and \( P = (p_{ij}) \in M_n(F) \). Then for \( v = x_1 e_1 + \cdots + x_n e_n \in V \) and \( W = y_1 e_1 + \cdots + y_n e_n \in V \) we have
\[
(v, w) = \sum x_i \overline{y_j} p_{ij} = (x_1, \ldots, x_n) P \begin{pmatrix} \overline{y_1} \\ \overline{y_2} \\ \vdots \\ \overline{y_n} \end{pmatrix}
\]
1. This matrix $P$ is called the **matrix of the inner product** with respect to the basis $e_1, \ldots, e_n$.

2. **(Definition.)** A matrix $B$ is called **hermitian** if $B = B^*$.

3. So, matrix $P$ in (1) is a hermitian matrix.

4. Since $(v, v) > 0$, for all non-zero $v \in V$, we have

   $$XPX^t > 0 \text{ for all non-zero } X \in \mathbb{F}^n.$$  

5. It also follows that $P$ is non-singular. *(Otherwise $XP = 0$ for some non-zero $X$.)*

6. Conversely, if $P$ is a $n \times n$ hermitian matrix satisfying such that $XPX^t > 0$ for all $X \in \mathbb{F}^n$ then

   $$(X, Y) = XP\overline{Y}^t \text{ for } X \in \mathbb{F}^n.$$  

   defines an inner product on $\mathbb{F}^n$.  


2 Inner Product Spaces

We will do calculus of inner produce.

2.1 (Definition) Let $F = \mathbb{R}$ OR $\mathbb{C}$. A vector space $V$ over $F$ with an inner product $(\cdot, \cdot)$ is said to an inner product space.

1. An inner product space $V$ over $\mathbb{R}$ is also called a Euclidean space.
2. An inner product space $V$ over $\mathbb{C}$ is also called a unitary space.

2.2 (Basic Facts) Let $F = \mathbb{R}$ OR $\mathbb{C}$ and $V$ be an inner product over $F$. For $v, w \in V$ and $c \in F$ we have

1. $\| cv \| = | c | \| v \|$,  
2. $\| v \| > 0$ if $v \neq 0$,  
3. $| (v, w) | \leq \| v \| \| w \|$, Equility holds if and only if $w = \frac{(w, v)}{\| v \|^2} v$. (It is called the Cauchy-Swartz inequality.)  
4. $\| v+w \| \leq \| v \| + \| w \|$. (It is called the triangular inequality.)

Proof. Part 1 and 2 is obvious from the definition. To prove the Part (3), we can assume that $v \neq 0$. Write

$$z = w - \frac{(w, v)}{\| v \|^2} v.$$

Then $(z, v) = 0$ and

$$0 \leq \| z \|^2 = (z, w - \frac{(w, v)}{\| v \|^2} v) = (z, w) = (w - \frac{(w, v)}{\| v \|^2} v, w) = \| w \|^2 - \frac{(w, v)(v, w)}{\| v \|^2}.$$

This establishes (3). We will use Part 3 to prove Part 4, as follows:

$$\| v+w \|^2 = \| v \|^2 + (v, w) + (w, v) + \| w \|^2 = \| v \|^2 + 2Re[(v, w)] + \| w \|^2 \leq \| v \|^2 + 2 | (v, w) | + \| w \|^2 \leq \| v \|^2 + 2 \| v \| \| w \| + \| w \|^2 = (\| v \| + \| w \|)^2.$$

This establishes Part 4.
2.3 (Application of Cauchy-Schwartz inequality) Application of (3) of Facts 2.2 gives the following:

1. Example 1.2 gives

\[ |\sum_{i=1}^{n} x_i y_i| \leq \left( \sum_{i=1}^{n} x_i^2 \right)^{1/2} \left( \sum_{i=1}^{n} y_i^2 \right)^{1/2} \]

for \(x, y \in \mathbb{R}\).

2. Example 1.3, gives

\[ |\sum_{i=1}^{n} x_i y_i| \leq \left( \sum_{i=1}^{n} |x_i|^2 \right)^{1/2} \left( \sum_{i=1}^{n} |y_i|^2 \right)^{1/2} \]

for \(x, y \in \mathbb{C}\).

3. Example 1.4, gives

\[ |\operatorname{trace}(AB^*)| \leq \operatorname{trace}(AA^*)^{1/2} \operatorname{trace}(BB^*)^{1/2} \]

for \(A, B \in \mathbb{M}_n(\mathbb{C})\).

4. Example 1.5, gives

\[ |\int_0^1 f(t)g(t)dt| \leq \left( \int_0^1 |f(t)|^2 dt \right)^{1/2} \left( \int_0^1 |g(t)|^2 dt \right)^{1/2} \]

for any two continuous \(\mathbb{C}\)-valued functions \(f, g\) on \([0, 1]\).
2.1 Orthogonality

2.4 (Definition) Let \( F \) be \( \mathbb{R} \) or \( \mathbb{C} \). Let \( V \) be an inner product space over \( F \).

1. Suppose \( v, w \in V \). We say that \( v \) and \( w \) are mutually orthogonal if the inner product \( (v, w) = 0 \) (OR equivalently if \( (w, v) = 0 \). We use variations of the expression "mutually orthogonal" and sometime we do not mention the word "mutually".)

2. For \( v, w \in V \) we write \( v \perp w \) to mean \( v \) and \( w \) are mutually orthogonal.

3. A subset \( S \subseteq V \) is said to be an orthogonal set if

\[
v \perp w \text{ for all } v, w \in S, \ v \neq w.
\]

4. An orthogonal set \( S \) is said to be an orthonormal set if

\[
\|v\| = 1 \text{ for all } v \in S.
\]

5. (Comment) Note the zero vector is orthogonal to all elements of \( V \).

6. (Comment) Geometrically, \( v \perp w \) means \( v \) is perpendicular to \( w \).

7. (Example) Let \( V = \mathbb{R}^n \) or \( V = \mathbb{C}^n \). Then the standard basis \( E = \{e_1, \ldots, e_n\} \) is an orthonormal set.

8. (Example) In \( \mathbb{R}^2 \) or \( \mathbb{C}^2 \), we have the ordered pairs \( v = (x, y) \) and \( w = (y, -x) \) are orthogonal. (Caution: Notation \( (x, y) \) here.)

2.5 (Example) Consider the example 1.5 over \( \mathbb{R} \). Here \( V \) is the inner product space of all real valued continuous functions on \([0, 1]\). Let

\[
f_n(t) = \sqrt{2} \cos(2\pi nt) \quad g_n(t) = \sqrt{2} \sin(2\pi nt)
\]

Then \( \{1, f_1, g_1, f_2, g_2, \ldots\} \) is an orthonormal set.
Now consider the inner product space $W$ in the same example 1.5 over $\mathbb{C}$. Let
\[ h_n = \frac{f_n + i g_n}{\sqrt{2}} = \exp(2\pi i n). \]
Then
\[ \{ h_n : n = 0, 1, -1, 2, -2, \ldots, \ldots \} \]
is an orthonormal set

2.6 (Theorem) Let $F = \mathbb{R}$ or $F = \mathbb{C}$ and let $V$ be an inner product space over $F$. Let $S$ be an orthogonal set of non-zero vectors. Then $S$ is linearly independent. \((\text{Therefore, cardinality}(S) \leq \text{dim} V. )\)

Proof. Let $v_1, v_2, \ldots, v_n \in S$ and
\[ c_1 v_1 + c_2 v_2 + \cdots + c_n v_n = 0 \]
where $c_i \in F$. We will prove $c_i = 0$. For example, apply inner product $(\ast, v_1)$ to the above equation. We get:
\[ c_1 (v_1, v_1) + c_2 (v_2, v_1) + \cdots + c_n (v_n, v_1) = (0, v_1) = 0. \]
Since $(v_1, v_1) \neq 0$ and $(v_2, v_1) = (v_3, v_1) = \cdots = (v_n, v_1) = 0$, we get $c_1 = 0$. Similarly, $c_i = 0$ for all $i = 1, \ldots, n$.

2.7 (Theorem) Let $F = \mathbb{R}$ or $F = \mathbb{C}$ and let $V$ be an inner product space over $F$. Assume $\{ e_1, e_2, \ldots, e_r \}$ is a set of non-zero orthogonal vectors. Let $v \in V$ and
\[ v = c_1 e_1 + c_2 e_2 + \cdots + c_r e_r \]
where $c_i \in F$. Then
\[ c_i = \frac{(v, e_i)}{\| e_i \|^2} \]
for $i = 1, \ldots, r$.

Proof. For example, apply inner product $(\ast, e_1)$ to the above and get
\[ (v, e_1) = c_1 (e_1, e_1) = c_1 \| e_1 \|^2 \]
So, $c_1$ is as asserted and, similarly, so is $c_i$ for all $i$. 

7
2.8 (Theorem) Let $F = \mathbb{R}$ or $F = \mathbb{C}$ and let $V$ be an inner product space over $F$. Let $v_1, v_2, \ldots, v_r$ be a set of linearly independent set. Then we can construct elements $e_1, e_2, \ldots, e_r \in V$ such that

1. $\{e_1, e_2, \ldots, e_r\}$ is an orthonormal set,

2. $e_k \in \text{Span}(\{v_1, \ldots, v_k\})$

**Proof.** The proof is known as **Gram-Schmidt orthogonalization process.** Note $v_1 \neq 0$. First, let

$$ e_1 = \frac{v_1}{\|v_1\|} $$

Then $\|e_1\| = 1$. Now let

$$ e_2 = \frac{v_2 - (v_2, e_1)e_1}{\|v_2 - (v_2, e_1)e_1\|}. $$

Note that the denominator is non-zero, $e_1 \perp e_2$ and $\|e_2\| = 1$. Now we use the induction. Suppose we already constructed $e_1, \ldots, e_{k-1}$ that satisfies (1) and (2) and $k \leq r$. Let

$$ e_k = \frac{v_k - \sum_{i=1}^{k-1} (v_k, e_i)e_i}{\|v_k - \sum_{i=1}^{k-1} (v_k, e_i)e_i\|}. $$

Note that the denominator is non-zero, $\|e_k\| = 1$ and $e_k \perp e_i$ for $i = 1, \ldots, k - 1$. From construction, we also have

$$ \text{Span}(\{v_1, \ldots, v_k\}) = \text{Span}(\{e_1, \ldots, e_k\}). $$

So, the proof is complete.

2.9 (Corollary) Let $F = \mathbb{R}$ or $F = \mathbb{C}$ and let $V$ be a finite dimensional inner product space over $F$. Then $V$ has an orthonormal basis.

**Proof.** The proof is immediate from the above theorem 2.8.

Examples. Read examples 12 and 13, page 282, for some numerical computations.
2.10 (Definition) Let \( F = \mathbb{R} \) or \( F = \mathbb{C} \) and let \( V \) be a inner product space over \( F \). Let \( W \) be a subspace of \( V \) and \( v \in V \). An element \( w_0 \in W \) is said to be a **best approximation to** \( v \) or **nearest to** \( v \) in \( W \), if

\[
\| v - w_0 \| \leq \| v - w \| \quad \text{for all} \quad w \in W.
\]

(Try to think what it means in \( \mathbb{R}^2 \) or \( \mathbb{C}^n \) when \( W \) is line or a plane through the origin.)

**Remark.** I like the expression ”nearest” and the textbook uses ”best approximation”. I will try to be consistent to the textbook.

2.11 (Theorem) Let \( F = \mathbb{R} \) or \( F = \mathbb{C} \) and let \( V \) be a inner product space over \( F \). Let \( W \) be a subspace of \( V \) and \( v \in V \). Then

1. An element \( w_0 \in W \) is best approximation to \( v \) if and only if \( (v - w_0) \perp w \) for all \( w \in W \).

2. A best approximation \( w_0 \in W \) to \( v \), if exists, is unique.

3. Suppose \( W \) is finite dimensional and \( e_1, e_2, \ldots, e_n \) is an orthonormal basis of \( W \). Then

\[
w_0 = \sum_{k=1}^{n} (v, e_k) e_k
\]

is the best approximation to \( v \) in \( W \). (*The textbook mixes up orthogonal and orthonormal and have a condition the looks complex. We assume orthonormal and so \( \| e_i \| = 1 \)*)

**Proof.** To prove (1) let \( w_0 \in W \) be such that \( (v - w_0) \perp w \) for all \( w \in W \). Then, since \( w - w_0 \in W \), we have

\[
\| v - w \|^2 = \| (v - w_0) + (w_0 - w) \|^2 = \| v - w_0 \|^2 + \| (w_0 - w) \|^2 \geq \| v - w_0 \|^2.
\]

Therefore, \( w_0 \) is nearest to \( v \). Conversely, assume that \( w_0 \in W \) is nearest to \( v \). We will prove that the inner product \( (v - w_0, w) = 0 \) for all \( w \in W \). For convenience, we write \( v_0 = v - w_0 \). So, we have

\[
\| v_0 \|^2 \leq \| v - w \|^2 \quad \text{Eqn } I
\]
for all $w \in W$. Write $v - w = v - w_0 + (w_0 - w) = v_0 + (w_0 - w)$. Since any element in $w$ can be written as $w_0 - w$ for some $w \in W$, Eqn-I can be rewritten as

$$\| v_0 \|^2 \leq \| v_0 + w \|^2$$

for all $w \in W$. So, we have

$$\| v_0 \|^2 \leq \| v_0 + w \|^2 = \| v_0 \|^2 + 2 Re[(v_0, w)] + \| w \|^2$$

and hence

$$0 \leq 2 Re[(v_0, w)] + \| w \|^2 \quad \text{Eqn} - II$$

for all $w \in W$.

Fix $w \in W$ with $w \neq w_0$ and write

$$\tau = -\frac{(v_0, w_0 - w) \cdot (w_0 - w)}{\| w_0 - w \|^2}.$$

Since, $\tau \in W$, we can substitute $\tau$ for $w$ in Eqn-II and get

$$0 \leq \frac{|(v_0, w_0 - w)|^2}{\| w_0 - w \|^2}.$$

Therefore $(v_0, w_0 - w) = 0$ for all $w \in W$ with $w_0 - w \neq 0$. Again, since any non-zero element in $W$ can be written as $w_0 - w$, it follows that $(v_0, w) = 0$ for all $w \in W$. So the proof of (1) is complete.

To prove Part 2, let $w_0, w_1$ be nearest to $v$. Then, by orthogonality (1), we have

$$\| w_0 - w_1 \|^2 = (w_0 - w_1, w_0 - w_1) = (w_0 - w_1, [w_0 - v] + [v - w_1]) = 0.$$

So, Part 2 is established.

Let $w_0$ be given as in Part 3. Now, to prove Part 3, we will prove $(w_0 - v) \perp e_i$ for $i = 1, \ldots, n$. So, for example,

$$(w_0 - v, e_1) = (w_0, e_1) - (v, e_1) = (v, e_1)(e_1, e_1) - (v, e_1) = 0.$$

So, Part 3 is established.
\textbf{2.12 (Definition)} Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$ and let $V$ be an inner product space over $\mathbb{F}$. Let $S$ be a subset of $V$. The \textbf{orthogonal complement} $S^\perp$ of $S$ is the set of all elements in $V$ that are orthogonal to each element of $S$. So,

$$S^\perp = \{v \in V : v \perp w \text{ for all } w \in S\}.$$ 

It is easy to check that

1. $S^\perp$ is a subspace of $V$.
2. $\{0\}^\perp = V$ and
3. $V^\perp = \{0\}$.

\textbf{2.13 (Definition and Facts)} Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$ and let $V$ be an inner product space over $\mathbb{F}$. Let $W$ be a subspace of $V$. Suppose $v \in V$, we say that $w \in W$ is the \textbf{orthogonal projection of $v$ to $W$}, if $w$ is nearest to $v$.

1. There is no guarantee that orthogonal projection exists. But by Part 2 of theorem 2.11, when it exists, orthogonal projection is unique.
2. Also, if $\dim(W)$ is finite, then by Part 3 of theorem 2.11, orthogonal projection always exists.
3. Assume $\dim(W)$ is finite. Define the map

$$\pi_W : V \to V$$

where $\pi(v)$ is the orthogonal projection of $v$ in $W$. The map $\pi_W$ is a linear operator and is called the \textbf{orthogonal projection of $V$ to $W$}. Clearly, $\pi_W^2 = \pi_W$. So, $\pi_W$ is, indeed, a projection.

4. For $v \in V$, let $E(v) = v - \pi_W(v)$. Then $E$ is the orthogonal projection of $V$ to $W^\perp$. \textbf{Proof}. By definition of $\pi_W$, we have, $E(v) = v - \pi_W(v) \in W^\perp$. Now, given $v \in V$ and $w^* \in W^\perp$ we have, $(v - E(v), w^*) = (\pi_W(v), w^*) = 0$. So, by theorem 2.11, $E$ is the projection to $W^\perp$.

\textbf{Example}. Read Example 14, page 286.
2.14 (Theorem) Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$ and let $V$ be an inner product space over $\mathbb{F}$. Let $W$ be a finite dimensional subspace of $V$. Then $\pi_W$ is a projection and $W^\perp$ is the null space of $\pi_W$. Therefore,

$$V = W \oplus W^\perp.$$ 

Proof. Obvious.

2.15 (Theorem: Bessel's inequality) Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$ and let $V$ be an inner product space over $\mathbb{F}$. Suppose $\{v_1, v_2, \ldots, v_n\}$ is an orthogonal set of non-zero vector. Then, for any element $v \in V$ we have

$$\sum_{k=1}^{n} \frac{|(v, v_k)|^2}{\|v_k\|^2} \leq \|v\|^2.$$ 

Also, the equality holds if and only if

$$v = \sum_{k=1}^{n} \frac{(v, v_k)}{\|v_k\|^2} v_k.$$ 

Proof. We write,

$$e_k = \frac{v_k}{\|v_k\|}$$

and prove that

$$\sum_{k=1}^{n} |(v, e_k)|^2 \leq \|v\|^2.$$ 

where $\{e_1, e_2, \ldots, e_n\}$ is an orthonormal set. Write $W = \text{Span}(\{e_1, e_2, \ldots, e_n\})$. By theorem 2.11,

$$w_0 = \sum_{k=1}^{n} (v, e_k) e_k$$

is nearest to $v$ in $W$. So, $(v - w_0, w_0) = 0$. Therefore,

$$\|v\|^2 = \|(v - w_0) + w_0\|^2 = \|(v - w_0)\|^2 + \|w_0\|^2.$$ 

So,

$$\|v\|^2 \geq \|w_0\|^2 = \sum_{k=1}^{n} |(v, e_k)|^2.$$
Also note that the equality holds if and only if \( \| (v - w_0) \|_2 = 0 \) if and only if \( v = w_0 \). So, the proof is complete.

If we apply Bessel’s inequality 2.15 to example 2.5 we get the following inequality.

2.16 (Theorem: Application of Bessel’s inequality) For and \( \mathbb{C} \)-valued continuous function \( f \) on \( [0, 1] \) we have

\[
\sum_{k=-n}^{n} \left| \int_{0}^{1} f(t) \exp(2\pi ikt) dt \right|^2 \leq \int_{0}^{1} |f(t)|^2 dt.
\]

Homework: Exercise 1-7, 9, 11 from page 288-289. These are popular problems for Quals.
3 Linear Functionals and Adjoints

We start with some preliminary comments.

3.1 (Comments and Theorems) Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$ and let $V$ be an inner product space over $\mathbb{F}$.

1. Given an element $v \in V$ we can define a linear functional $f_v : V \to \mathbb{F}$ by $f_v(x) = (x, v)$.

2. The association $F(v) = f_v$ defines a natural linear map $F : V \to V^*$. 

3. In fact, $F$ is injective. **Proof.** Suppose $f_v = 0$ Then $(x, v) = 0$ for all $x \in V$, so, $(v, v) = 0$ and hence $v = 0$.

4. Now assume that $V$ has finite dimension. Then the natural map $F$ is an isomorphism.

5. Again, assume $V$ has finite dimension $n$ and assume $\{e_1, \ldots, e_n\}$ is an orthonormal basis. Let $f \in V^*$. Let 

   $$v = \sum_{k=0}^{n} f(e_i) e_i.$$ 

   Then $f = F(v) = f_v$. That means, 

   $$f(x) = (x, v) \quad \text{for all } x \in V.$$ 

   **Proof.** We will only check $f(e_1) = (e_1, v)$, which is obvious by orthonormality.

6. Assume the same as in Part 5. Then the association 

   $$G(f) = \sum_{k=0}^{n} f(e_i) e_i$$ 

   defines the inverse 

   $$G : V^* \to V$$ 

   of $F$. 

14
7. **Remark.** The map \( F \) fails to be isomorphism if \( V \) is not finite dimensional. Following is an example that shows \( F \) is not on to.

### 3.2 (Example)

Let \( V \) be the vector space of polynomial over \( \mathbb{C} \). For \( f, g \in V \) define inner product:

\[
(f, g) = \int_0^1 f(t)\overline{g(t)}dt.
\]

Note \( \int_0^1 t^j t^k dt = \frac{1}{j+k+1} \). So, if \( f(X) = \sum a_k X^k \) and \( g(X) = \sum b_k X^k \) we have

\[
(f, g) = \sum_{j,k} \frac{1}{j+k+1} a_j b_k.
\]

Fix a complex number \( z \in \mathbb{C} \). By evaluation at \( z \), we define the functional \( L : V \to \mathbb{C} \) as \( L(f) = f(z) \) for any \( f \in V \). We claim that \( L \) is not in the image of the map \( F : V \to V^* \). In other words, there is no polynomial \( g \in V \) such that

\[
f(z) = L(f) = (f, g) = \int_0^1 f(t)\overline{g(t)}dt \quad \text{for all} \quad f \in V.
\]

To prove this suppose there is such a \( g \).

Write \( h = X - z \). Given \( f \in V \), we have \( hf \in V \) and \( 0 = (hf)(z) = L(hf) \).

So,

\[
0 = L(hf) = (hf, g) = \int_0^1 h(t)f(t)\overline{g(t)}dt \quad \text{for all} \quad f \in V.
\]

By substituting \( f = (X - z)g \) we have

\[
0 = \int_0^1 | h(t) |^2 | g(t) |^2 dt.
\]

Since \( h \neq 0 \) it follows \( g = 0 \). But \( L \neq 0 \).
3.1 Adjoint Operator

3.3 (Theorem) Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$ and let $V$ be an finite dimensional inner product space over $\mathbb{F}$. Suppose $T \in L(V, V)$ is a linear operator. Then, there is a unique linear operator

$$T^* : V \rightarrow V$$

such that

$$(T(v), w) = (v, T^*(w)) \text{ for all } v, w \in V.$$  

Definition. This operator $T^*$ is called the Adjoint of $T$.

Proof. Fix and element $w \in V$. Let $\Gamma : V \rightarrow \mathbb{F}$, be defined by the diagram:

$$\begin{array}{ccc}
V & \xrightarrow{T} & V \\
\downarrow{\Gamma} & {\}^{(\ast, w)} & \\
\downarrow{\ast} & & \mathbb{F}
\end{array}$$

That means $\Gamma(v) = (T(v), w)$ for all $v \in V$.

By Part 2 of theorem 3.1, there is an unique element $w'$ such that

$$\Gamma(v) = (v, w') \text{ for all } v \in V.$$  

That means

$$(T(v), w) = (v, w') \text{ for all } v \in V. \quad (Eqn - I)$$

Now, define $T^*(w) = w'$. It is easy to check that $T^*$ is linear (use Eqn-I). Uniqueness also follows from Eqn-I.

3.4 (Theorem) Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$ and let $V$ be an finite dimensional inner product space over $\mathbb{F}$. Suppose $T \in L(V, V)$ is a linear operator. Let $e_1, \ldots, e_n$ be an orthonormal basis of $V$. Suppose $T \in L(V, V)$ be a linear operator of $V$.

1. Write

$$a_{ij} = (T(e_j), e_i) \text{ and } A = (a_{ij}).$$

Then $A$ is the matrix of $T$ with respect to $e_1, \ldots, e_n$. 

16
2. With respect to $e_1, \ldots, e_n$, the matrix of the adjoint $T^*$ is the conjugate transpose of the matrix $A$ of $T$.

**Proof.** To prove Part 1, we need to prove that

$$(T(e_1), \ldots, T(e_n)) = (e_1, \ldots, e_n) \begin{pmatrix} (T(e_1), e_1) & (T(e_2), e_1) & \cdots & (T(e_n), e_1) \\ (T(e_1), e_2) & (T(e_2), e_2) & \cdots & (T(e_n), e_2) \\ \vdots & \vdots & \ddots & \vdots \\ (T(e_1), e_n) & (T(e_2), e_n) & \cdots & (T(e_n), e_n) \end{pmatrix}.$$ 

This follows because, if $T(e_1) = \lambda_1 e_1 + \lambda_2 e_2 + \cdots + \lambda_n e_n$, then $\lambda_i = (T(e_1), e_i)$. So, the proof of Part 1 is complete.

Now, by Part 1, the matrix of the adjoint $T^*$ is

$$B = ((T^*(e_j), e_i)) = ((e_i, T^*(e_j))) = ((T(e_i), e_j)) = A^*.$$ 

This completes the proof of Part 2.

### 3.5 (Theorem: Projection and Adjoint)

Let $F = \mathbb{R}$ or $F = \mathbb{C}$ and let $V$ be an finite dimensional inner product space over $F$. Let $E \in L(V, V)$ be an orthogonal projection. Then $E = E^*$.

**Proof.** For any $x, y \in V$, we have

$$(E(x), y) = (E(x), E(y) + [y - E(y)]) = (E(x), E(y)) \quad \text{because} \quad [y - E(y)] \perp W.$$ 

Also

$$(E(x), E(y)) = (x + [E(x) - x], E(y)) = (x, E(y)) \quad \text{because} \quad [x - E(x)] \perp W.$$ 

Therefore

$$(E(x), y) = (x, E(y)) \quad \text{for all} \ x, y \in V.$$ 

Hence $E = E^*$.

### 3.6 (Remarks and Examples)

Let $V = \mathbb{R}$ and $A \in M_n(\mathbb{R})$ be a symmetric matrix. Let $T \in L(V, V)$ be defined by $A$. Then $T = T^*$. Also note that matrix of $T$ with respect to some other basis may not be symmetric. Read Example 17-21 from page 294-296.
3.7 (Theorem) Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$ and let $V$ be an finite dimensional inner product space over $\mathbb{F}$. Let $T, U \in L(V, V)$ be two linear operator and $c \in \mathbb{F}$. Then

1. $(T + U)^* = T^* + U^*$,
2. $(cT)^* = \overline{c}T^*$,
3. $(TU)^* = U^*T^*$,
4. $(T^*)^* = T$.

Proof. The proof is direct consequence of the definition (theorem 3.3).

The theorem 3.7 can be phrased as the map

$$L(V, V) \to L(V, V)$$

that sends $T \to T^*$ is conjugate-linear, anti-isomorphism of period two.

3.8 (Definition) Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$ and let $V$ be an finite dimensional inner product space over $\mathbb{F}$. Let $T \in L(V, V)$ be a linear operator and $c \in \mathbb{F}$. We say $T$ is self-adjoint or Hermitian if $T = T^*$.

Suppose $E = \{e_1, \ldots, e_n\}$ is an orthonormal basis of $V$. Let $A$ be the matrix of $T$ with respect to $E$. Then $T$ is self-adjoint if and only if $A = A^*$. 
4 Unitary Operators

Let me draw your attention that the expression "isomorphism" means different things in different context - like group isomorphism, vector space isomorphism, module isomorphism. In this section we talk about isomorphisms of inner product spaces.

4.1 (Definition and Facts) Let $F = \mathbb{R}$ or $F = \mathbb{C}$ and let $V, W$ be two inner product spaces over $F$. A linear map $T : V \rightarrow W$ is said to preserve inner product if

$$(T(x), T(y)) = (x, y) \text{ for all } x, y \in V.$$ 

We say $T$ is an an isomorphism of inner product spaces if $T$ preserves inner product and is one to one and onto.

4.2 (Facts) Let $F = \mathbb{R}$ or $F = \mathbb{C}$ and let $V, W$ be two inner product spaces over $F$. Let $T : V \rightarrow W$ be a linear map.

1. If $T$ preserves inner product, then

$$\| T(x) \| = \| x \| \text{ for all } x \in V.$$ 

2. If $T$ preserves inner product, then $T$ is injective (i.e. one to one).

3. If $T$ is an isomorphism of inner product spaces, then $T^{-1}$ is also an isomorphism of inner product spaces.

Proof. Obvious.

4.3 (Theorem) Let $F = \mathbb{R}$ or $F = \mathbb{C}$ and let $V, W$ be two finite dimensional inner product spaces over $F$, with $\dim V = \dim W = n$. Let $T : V \rightarrow W$ be a linear map. The the following are equivalent:

1. $T$ preserves inner product,

2. $T$ is an isomorphism of inner product spaces,

3. If $\{e_1, \ldots, e_n\}$ is an orthonormal basis of $V$ than $\{T(e_1), \ldots, T(e_n)\}$ an orthonormal basis of $W$, 

19
4. There is an orthonormal basis \( \{e_1, \ldots, e_n\} \) of \( V \) such that \( \{T(e_1), \ldots, T(e_n)\} \) also an orthonormal basis of \( W \).

**Proof.** (1 \( \Rightarrow \) 2): Since \( T \) preserves inner product, \( T \) is injective. Also since \( \dim V = \dim W \) it follows Hence \( T \) is also onto. So Part 2 is established.

(2 \( \Rightarrow \) 3): Suppose \( \{e_1, \ldots, e_n\} \) is an orthonormal basis of \( V \). Since also preserves inner product, \( \{T(e_1), \ldots, T(e_n)\} \) is an orthonormal set. Since \( \dim W = n \), and since orthonormal set are independent \( \{T(e_1), \ldots, T(e_n)\} \) is a basis of \( W \). So, Part 3 is established.

(3 \( \Rightarrow \) 4): Since \( V \) has an orthonormal basis Part 4 follows from Part 3.

(4 \( \Rightarrow \) 1): Let \( \{e_1, \ldots, e_n\} \) be an orthonormal basis of \( V \) such that \( \{T(e_1), \ldots, T(e_n)\} \) also an orthonormal basis of \( W \). Since \( (T(e_i), T(e_j)) = (e_i, e_j) \) for all \( i, j \), it follows
\[
(T(x), T(y)) = (x, y) \quad \text{for all} \quad x, y \in V.
\]
So, Part 1 is established.

4.4 (Corollary) Let \( \mathbb{F} = \mathbb{R} \) or \( \mathbb{F} = \mathbb{C} \) and let \( V, W \) be two finite dimensional inner product spaces over \( \mathbb{F} \). Then \( V \) and \( W \) are isomorphic (as inner product spaces) if and only is \( \dim V = \dim W \).

**Proof.** If \( V \) and \( W \) are isomorphic then clearly \( \dim V = \dim W \). Conversely, if \( \dim V = \dim W = n \) then we can find an orthonormal basis \( \{e_1, \ldots, e_n\} \) of \( V \) an an orthonormal basis \( \{E_1, \ldots, E_n\} \) of \( W \). The association
\[
T(e_i) = E_i
\]
defines and isomorphism of \( V \) to \( W \).

4.5 (Example) Let \( \mathbb{F} = \mathbb{R} \) or \( \mathbb{F} = \mathbb{C} \) and let \( V \) be a finite dimensional inner product spaces over \( \mathbb{F} \), with \( \dim V = n \). Fix a basis \( E = \{e_1, \ldots, e_n\} \) of \( V \). Consider the linear map
\[
f : V \to \mathbb{F}^n
\]
given by \( f(a_1 e_1 + \cdots + a_n e_n) = (a_1, \ldots, a_n) \). With respect to the usual inner product on \( \mathbb{F}^n \) this map is an isomorphism of inner product spaces if and only if \( E = \{e_1, \ldots, e_n\} \) is an orthonormal basis.
Proof. Note that $f$ sends $E$ to the standard basis. So, by theorem 4.3, $E$ is orthonormal $f$ is an isomorphism.

Homework: Read Example 23 and 25 from page 301-302.

Question: This refers to Example 25. Let $V$ be the inner product space of all continuous $\mathbb{R}$–valued functions on $[0,1]$, with inner product

$$(f,g) = \int_0^1 f(t)g(t)dt.$$

Let $T : V \rightarrow V$ be any linear operator. What can we say about when $T$ perserves inner product?

4.6 (Theorem) Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$ and let $V,W$ be two inner product spaces over $\mathbb{F}$. Let $T : V \rightarrow W$ be any linear transformation. Then $T$ perserves inner product if and only if

$$\|T(x)\| = \|x\| \text{ for all } x \in V.$$

Proof. If perserves inner product then clearly the condition holds. Now suppose the condition above holds and $x,y \in V$. Then

$$\|T(x+y)\|^2 = \|x+y\|^2.$$

Since

$$\|T(x)\| = \|x\| \text{ and } \|T(y)\| = \|y\|,$$

it follows that

$$(T(x),T(y)) + (T(y),T(x)) = (x,y) + (y,x).$$

Hence

$$\text{Re}[(T(x),T(y))] = \text{Re}[(x,y)].$$

(If $\mathbb{F} = \mathbb{R}$, then the proof is complete. ) Also, since Similar arguments with $x - y$, will give

$$\text{Im}[(w,z)] = \text{Re}[(w,iz)] \text{ for all } w,z \in V \text{ or } W$$

the proof is complete.
4.7 (Definition and Facts) Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$ and let $V$ be an inner product space over $\mathbb{F}$. An isomorphism $T : V \cong V$ of inner product spaces is said to be an **unitary operator** on $V$.

1. Let $\mathcal{U} = \mathcal{U}(n) = \mathcal{U}(V)$ denote the set of all unitary operators on $V$.
2. The identity $I \in \mathcal{U}(V)$.
3. If $U_1, U_2 \in \mathcal{U}(V)$ then $U_1U_2 \in \mathcal{U}(V)$.
4. If $U \in \mathcal{U}(V)$ then $U^{-1} \in \mathcal{U}(V)$.
5. So, $\mathcal{U}(V)$ is a group under composition. It is a subgroup of the group of linear isomorphisms of $V$. Notationally
   $$\mathcal{U}(V) \subseteq GL(V) \subseteq L(V, V).$$
6. If $V$ is finite dimensional then a linear operator $T \in L(V, V)$ is unitary if and only if $T$ preserves inner product.

4.8 (Theorem) Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$ and let $V$ be an inner product space over $\mathbb{F}$. Let $U \in L(V, V)$ be a linear operator. Then $U$ is unitary if and only if the adjoint $U^*$ of $U$ exists and $UU^* = U^*U = I$.

**Proof.** Suppose $U$ is unitary. Then $U$ has an inverse $U^{-1}$. So, for $x, y \in V$ we have

$$\langle U(x), y \rangle = \langle U(x), UU^{-1}(y) \rangle = \langle x, U^{-1}(y) \rangle.$$

So, $U^*$ exists and $U^* = U^{-1}$. Conversely, assume the adjoint $U^*$ exists and $UU^* = U^*U = I$. We need to prove that $U$ preserves inner product. For $x, y \in V$ we have

$$\langle U(x), U(y) \rangle = \langle x, U^*U(y) \rangle = \langle x, y \rangle.$$

So, the proof is complete.

**Homework:** Read example 27, page 304.

4.9 (Theorem) Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$ and $A$ be an $n \times n$ matrix. Let $T : \mathbb{F}^n \to \mathbb{F}^n$ be the linear operator defined by $T(X) = AX$. With usual inner product on $\mathbb{F}^n$, we have $T$ is unitary if and only if $A^*A = I$. 

22
Proof. Suppose $A^*A = I$. Then $A^*A = AA^* = I$. Therefore,

$$(T(x), T(y)) = (Ax, Ay) = y^*A^*Ax = y^*x = (x, y).$$

Conversely, suppose $T$ is unitary. Then, $y^*A^*Ax = y^*x$ for all $x, y \in \mathbb{F}^n$. With appropriate choice of $x, y$ we can show that $A^*A = I$. So, the proof is complete.

4.10 (Definition) An $n \times n$- matrix $A \in \mathcal{M}_n(\mathbb{R})$ is called an orthogonal matrix if $A^tA = I$. The subset $O(n) \subseteq \mathcal{M}_n(\mathbb{R})$ of all orthogonal matrices from a subgroup of $GL_n(\mathbb{R})$.

An $n \times n$- matrix $B \in \mathcal{M}_n(\mathbb{C})$ is called a unitary matrix if $B^*B = I$. The subset $U(n) \subseteq \mathcal{M}_n(\mathbb{C})$ of all unitary matrices from a subgroup of $GL_n(\mathbb{C})$.

4.11 (Theorem) Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$ and let $V$ be a finite dimensional inner product space over $\mathbb{F}$. Let $U \in L(V, V)$ be a linear operator. Then $U$ is unitary if and only if the matrix of $U$ in with respect to some (or every) orthonormal basis is unitary.

Proof. By theorem 4.8, $U$ is unitary if and only if $U^*U = UU^* = I$. Suppose $E = \{e_1, \ldots, e_n\}$ is an orthonormal basis of $V$ and $A$ be the matrix on $U$ with respect to $E$. Assume $U$ is unitary. Since $UU^* = U^*U = I$, we have $AA^* = A^*A = I$. So $A$ is unitary.

Conversely, assume $A$ is unitary. Then $AA^* = A^*A = I$. Write $A = (a_{ij})$. Therefore $(U(e_i), U(e_j)) = \left(\sum_{k=1}^n a_{ki}e_k, \sum_{k=1}^n a_{kj}e_k\right) = \sum_{k=1}^n (a_{ki}e_k, a_{kj}e_k) = \sum_{k=1}^n a_{ki}a_{kj} = \delta_{ij} = (e_i, e_j)$.

So, the proof is complete.

4.12 (Exercise) 1. A matrix $A \in \mathcal{M}_n(\mathbb{R})$ is orthogonal if and only if $A^{-1} = A^t$.

2. A matrix $B \in \mathcal{M}_n(\mathbb{C})$ is unitary if and only if $B^{-1} = B^*$.

4.13 (Theorem) Suppose $A \in GL_n(\mathbb{C})$ be an invertible matrix. Then there is a lower triangular matrix $M \in GL_n(\mathbb{C})$ such that $MA \in U_n(\mathbb{C})$ and diagonal entries of $M$ are positive. Further, such an $M$ is unique.
Proof. Write

\[ A = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \]

where \( v_i \in \mathbb{C}^n \) are the rows of \( A \). Use Gram-Schmidt orthogonalization (theorem 2.8) and define

\[ e_k = \frac{v_k - \sum_{j=1}^{k-1} (v_k, e_j) e_j}{\| v_k - \sum_{j=1}^{k-1} (v_k, e_j) e_j \|} \]

Note \( e_1, \ldots, e_n \) is an orthogonal normal basis of \( \mathbb{C}^n \). Also

\[ e_k = \sum_{j=1}^k c_{kj} v_j \quad \text{with} \quad c_{kj} \in \mathbb{C} \quad \text{and} \quad c_{jj} \neq 0. \]

So, we have

\[
\begin{pmatrix}
  e_1 \\
  e_2 \\
  \vdots \\
  e_n \\
\end{pmatrix} = \begin{pmatrix}
  c_{11} & 0 & \cdots & 0 \\
  c_{21} & c_{22} & \cdots & 0 \\
  \vdots & \vdots & \ddots & \vdots \\
  c_{n1} & c_{n2} & \cdots & c_{nn} \\
\end{pmatrix} \begin{pmatrix}
  v_1 \\
  v_2 \\
  \vdots \\
  v_n \\
\end{pmatrix}
\]

Since

\[
\begin{pmatrix}
  e_1 \\
  e_2 \\
  \vdots \\
  e_n \\
\end{pmatrix}
\]

unitary, the existence is established.

For uniqueness, assume

\[ U = MB \quad \text{and} \quad U_1 = NB \]

where \( U, U_i \in U(n) \) and \( M, N \) are lower triangular with diagonal entries positive. Then \( MN^{-1} = UU_1^{-1} \) is unitary. Note that \( N^{-1} \) is also a lower triangular matrix with positive diagonal entries. Therefore,
1. $MN^{-1} = \Delta \in U((n))$,
2. $\Delta$ is a diagonal matrix,
3. Diagonal entries of $\Delta$ are positive.

Therefore $\Delta = MN^{-1} = I$ and $M = N$. So, the proof is complete.

**Homework:** Read example 28, page 307.

**4.14 (Theorem)** Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$ and let $V$ be a finite dimensional inner product space over $\mathbb{F}$. Suppose $E = \{e_1, \ldots, e_n\}$ and $\mathcal{E} = \{\epsilon_1, \ldots, \epsilon_n\}$ are two orthonormal bases of $V$. Let $(e_1, \ldots, e_n) = (\epsilon_1, \ldots, \epsilon_n)P$ for some matrix $P$. Then $P$ is unitary.

**Proof.** We have

$$I = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix} (e_1, \ldots, e_n) = \begin{bmatrix} P^t \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix} \end{bmatrix} [(\epsilon_1, \ldots, \epsilon_n)P]$$

Sorry, this approach is naive and fails because such products (of matrices and scalars - matrix products and inner products) are not associative. In any case, the correct proof is left as an exercise.

**4.15 (Exercise)** Consider $V = \mathbb{C}^n$, with the usual inner product. Think of the elements of $V$ as row vectors.

Let $v_1, v_2, \ldots, v_n \in V$ and let

$$A = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

1. Prove that $v_1, v_2, \ldots, v_n$ forms a basis if and only if $A$ invertible.
2. Prove that $v_1, v_2, \ldots, v_n$ forms an orthonormal basis if and only if $A$ an unitary matrix (i.e. $A^*A = I$).
3. We can make similar statements about vectors in $\mathbb{R}^n$ and orthogonal matrices.
5 Normal Operators

Let $V$ be an inner product space and $T \in L(V, V)$. Main objective of this section is to find necessary and sufficient conditions for $T$ so that there is an orthonormal basis $\{e_1, \ldots, e_n\}$ of $V$ such that each $e_i$ is an eigen vector of $T$.

5.1 (Theorem) Let $F = \mathbb{R}$ or $F = \mathbb{C}$ and let $V$ be a finite dimensional inner product space over $F$. Let $T \in L(V, V)$ be a linear operator. Suppose $E = \{e_1, \ldots, e_n\}$ is an orthonormal basis of $V$ and each $e_i$ is an eigen vector of $T$

1. So, we have $T(e_i) = c_i e_i$ for some scalars $c_i \in F$.

2. So, the matrix of $T$ with respect to $E$ is the diagonal matrix

$$\Delta = \text{diagonal}(c_1, c_2, \ldots, c_n).$$

3. If $F = \mathbb{R}$ then the matrix of the adjoint operator $T^*$ is

$$\Delta^* = \Delta = \text{diagonal}(c_1, c_2, \ldots, c_n).$$

Therefore, in the real case, a sufficient condition is that $T$ is self-adjoint.

4. If $F = \mathbb{C}$, the matrix of the adjoint operator $T^*$ is

$$\Delta^* = \text{diagonal}(\overline{c_1}, \overline{c_2}, \ldots, \overline{c_n}).$$

Therefore

$$TT^* = T^*T.$$ So, in complex case, a necessary condition is that $T$ commutes with the adjoint $T^*$.

(Compare with theorem 4.8.)

5.2 (Definition) Let $F = \mathbb{R}$ or $F = \mathbb{C}$ and let $V$ be a finite dimensional inner product space over $F$. Let $T \in L(V, V)$ be a linear operator. We say $T$ is a normal operator, if

$$TT^* = T^*T.$$ Therefore, self-adjoint operators are normal.
5.3 **(Theorem)** Let $F = \mathbb{R}$ or $F = \mathbb{C}$ and let $V$ be a finite dimensional inner product space over $F$. Let $T \in L(V, V)$ be a linear operator. If $T$ is self-adjoint then

1. Eigen values of $T$ are real.
2. Eigen vectors associated to distinct eigen values are orthogonal.

**Proof.** Suppose $c \in F$ be an eigen value and $e \in V$ be the corresponding eigen vector. Then $T(e) = ce$ and

$$c(e, e) = (ce, e) = (T(e), e) = (e, T^*(e)) = (e, T(e)) = (e, ce) = \overline{c}(e, e).$$

So, $c = \overline{c}$ and $c$ is real.

Now suppose $T(e) = ce$ and $T(\varepsilon) = d\varepsilon$ where $c \neq d$ scalars and $e, \varepsilon \in V$ be nonzero. Then

$$c(e, \varepsilon) = (ce, \varepsilon) = (T(e), \varepsilon) = (e, T^*(\varepsilon)) = (e, T(\varepsilon)) = (e, d\varepsilon) = \overline{d}(e, \varepsilon).$$

Since $d \neq c$, we have $(e, \varepsilon) = 0$. So, the proof is complete.

5.4 **(Theorem)** Let $F = \mathbb{R}$ or $F = \mathbb{C}$ and let $V$ be a finite dimensional inner product space over $F$. Let $T \in L(V, V)$ be a self-adjoint operator. Assume $\dim V > 0$. Then $T$ has a (non-zero) eigen vector.

**Proof.** Let $E = \{e_1, \ldots, e_n\}$ be an orthonormal basis of $V$. Let $A$ be the matrix of $T$. Since $T$ is self-adjoint, $A = A^*$. Now, we deal with real and complex case seperately.

**Real Case:** Obviously, $A \in M_n(\mathbb{R})$ and $A = A^*$ means $A$ is symmetric. In any case, consider the map

$$U : \mathbb{C}^n \to \mathbb{C}^n \text{ where } U(X) = AX \text{ for } X \in \mathbb{C}^n.$$ 

Since $A = A^*$, by theorem 5.3, $U$ has only real eigen values. So, $\det(XI - A) = 0$ has ONLY real solution. Since $\mathbb{C}$ is algebraically closed, we can pick a real solution $c \in \mathbb{R}$ so that $\det(cI - A) = 0$. Therefore, $(cI - A)X = 0$ has a real non-zero solution $(x_1, \ldots, x_n)^t \in \mathbb{R}^n$.

Write $e = x_1e_1 + \cdots + x_ne_n$ then $e \neq 0$ and $T(e) = ce$.

*(Note, we went upto $\mathbb{C}$ to get a proof in the real case.)*

**Complex Case:** Proof is same, only easier. Here we know $\det(XI - A) = 0$ has a solution $c \in \mathbb{C}$ and the rest of the proof is identical.

27
5.5 (Exercise) Let $V$ be a finite dimensional inner product space over $\mathbb{C}$. Let $T \in L(V, V)$ be a self-adjoint operator. Let $Q(X)$ be the characteristic polynomial of $T$. Then $Q(X) \in \mathbb{R}[X]$.

**Proof.** We can repeat some of the steps of theorem 5.4. Let $E = \{e_1, \ldots, e_n\}$ be an orthonormal basis of $V$. Let $A$ be the matrix of $T$. Since $T$ is self-adjoint, $A = A^*$. The $Q(X) = \det(XI - A)$. Then

$$Q(X) = (X - c_1)(X - c_2) \cdots (X - c_n) \text{ where } c_i \in \mathbb{C}.$$ 

By arguments in theorem 5.4, $c_i \in \mathbb{R}$, for $i = 1, \ldots, n$.

5.6 (Example 29, page 313) Let $V$ be the vector space of continuous $\mathbb{C}$-valued functions on the interval $[0,1]$. As usual, for $f, g \in V$, define inner product

$$(f, g) = \int_0^1 f(t)\overline{g(t)}dt.$$ 

Let $T : V \to V$ be the operator defined by $T(f)(t) = tf(t)$.

1. Then $T$ is self-adjoint. This is true because

$$(tf, g) = (f, tg) \text{ for all } f, g \in V.$$ 

2. $T$ has no non-zero eigen vector. **Proof.** Suppose $f \in V$ and $T(f) = cf$, for some $c \in \mathbb{C}$. Then $tf(t) = cf(t)$ for all $t \in [0,1]$. Since $f$ is continuous, $f = 0$.

3. This example shows that theorem 5.4 fails, if $V$ is infinite dimensional.

5.7 (Theorem) Let $F = \mathbb{R}$ or $F = \mathbb{C}$ and let $V$ be a finite dimensional inner product space over $F$. Let $T \in L(V, V)$ be a linear operator. Suppose $W$ is a $T$-invariant subspace of $V$. Then the orthogonal complement $W^\perp$, of $W$, is invariant under $T^*$.

**Proof.** Let $x \in W$ and $y \in W^\perp$. Since, $T(x) \in W$, we have $(x, T^*(y)) = (T(x), y) = 0$. So, $T^*(y) \in W^\perp$. 


5.8 (Theorem) Let $F = \mathbb{R}$ or $F = \mathbb{C}$ and let $V$ be a finite dimensional inner product space over $F$. Let $T \in L(V, V)$ be a self-adjoint linear operator. Then $V$ has an orthonormal basis $E = \{e_1, \ldots, e_n\}$ such that each $e_i$ is an eigen vector of $T$.

**Proof.** We assume that $\dim V = n > 0$. By theorem 5.4, $T$ has an eigen vector $v$. Let

$$e_1 = \frac{v}{\|v\|}.$$ 

If $\dim V = 1$, we are done. Now we will use induction and assume that the theorem to true for inner product spaces of dimension less than $\dim V$. Write $W = Fe_1$ and $V_1 = W^\perp$. Since $W$ is invariant under $T$, by theorem 5.7, $W^\perp$ is invariant under $T^* = T$. Therefore, $V_1$ has a orthonormal basis $\{e_2, \ldots, e_n\}$ such that $e_2, \ldots, e_n$ are eigen vectors of $T_{V_1}$ hence of $T$. Therefore $\{e_1, e_2, \ldots, e_n\}$ is an orthonormal basis of $V$ and each $e_i$ is eigen vectors of $T$. So, the proof is complete.

5.9 (Theorem) 1. Let $A \in \mathbb{M}_n(\mathbb{C})$ be a Hermitian (self-adjoint) matrix. Then there is unitary matrix $P$ such that $P^{-1}AP = \Delta$ is a diagonal matrix.

2. Let $A \in \mathbb{M}_n(\mathbb{R})$ be a symmetric matrix. Then there is orthogonal real matrix $P$ such that $P^{-1}AP = \Delta$ is a diagonal matrix.

**Proof.** To prove Part 1, consider the operator

$$T : \mathbb{C}^n \to \mathbb{C}^n \quad \text{where} \quad T(X) = AX \quad \text{for} \quad X \in \mathbb{C}^n.$$ 

Since $A$ is Hermitian, so is $T$. By theorem 5.8, there is an orthonormal basis $E = \{e_1, \ldots, e_n\}$ of $\mathbb{C}^n$ such that each $e_i$ is an eigen vector of $T$. So, we have

$$(T(e_1), T(e_2), \ldots, T(e_n)) = (e_1, e_2, \ldots, e_n)\Delta$$

where $\Delta = \text{diagonal}(c_1, \ldots, c_n)$ is a diagonal matrix and $T(e_i) = c_i e_i$. Suppose $e_1, \ldots, e_n$ is the standard basis of $\mathbb{C}^n$, and

$$(e_1, \ldots, e_n) = (e_1, \ldots, e_n)P$$

for some matrix $P \in \mathbb{M}_n(\mathbb{C})$. Then $P$ is unitary. We also have

$$(T(e_1), \ldots, T(e_n)) = (e_1, \ldots, e_n)A.$$
Combining all these, we have

\[ A = P \Delta P^{-1}. \]

So, the proof of Part 1 is complete. The proof of Part 2 is similar.
5.1 Regarding Normal Operators

5.10 (Theorem) Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$ and let $V$ be a finite dimensional inner product space over $\mathbb{F}$. Let $T \in L(V,V)$ be a normal operator.

1. Then $\| T(x) \| = \| T^*(x) \|$ for all $x \in V$.

2. Suppose $v \in V$. Then $v$ is an eigen vector of $T$ with eigen value $c$ if and only if $v$ is an eigen vector of $T^*$ with eigen value $\bar{c}$. In other words,

$$T(v) = cv \iff T^*(v) = \bar{c}v.$$  

Proof. We have $TT^* = T^*T$ and $\| T(x) \|^2 = (T(x), T(x)) = (x, T^*T(x)) = (x, TT^*(x)) = \| T^*(x) \|^2$.

So, Part 1 is established. To prove Part 2, for a $c \in \mathbb{F}$ write $U = T - cI$. So, $U^* = T^* - \bar{c}I$. Since $T^*T = TT^*$, we have $U^*U = UU^*$. Therefore, by 1,

$$\| (T - cI)(v) \| = \| (T^* - \bar{c}I)(v) \|.$$  

Therefore the proof of Part 2 is complete.

5.11 (Definition) A matrix $A \in M_n(\mathbb{C})$ is said to be normal if $AA^* = A^*A$.

5.12 (Theorem) Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$ and let $V$ be a finite dimensional inner product space over $\mathbb{F}$. Let $T \in L(V,V)$ be a linear operator on $V$. Let $E = \{e_1, \ldots, e_n\}$ be an orthonormal basis and let $A$ the matrix of $T$ with respect to $E$. Assume $A$ is upper triangular. Then $T$ is normal if and only if $A$ is diagonal.

Proof. Since $E$ is orthonormal, matrix of $T^*$ is $A^*$. Assume $A$ is diagonal. Then $A^*A = AA^*$. Therefore $T^*T = TT^*$ and $T$ is normal.

Conversely, assume $T$ is normal. So, $T^*T = TT^*$ and hence $A^*A = AA^*$. First, we will assume $n = 2$ and prove $A$ is diagonal. Write

$$A = \begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{pmatrix} \quad \text{and so} \quad A^* = \begin{pmatrix} \bar{a}_{11} & 0 \\ \bar{a}_{12} & \bar{a}_{22} \end{pmatrix}.$$
We have

\[ AA^* = \begin{pmatrix}
|a_{11}|^2 + |a_{12}|^2 & a_{12}a_{22} \\
\overline{a_{22}a_{12}} & |a_{22}|^2
\end{pmatrix} \]

and

\[ A^* A = \begin{pmatrix}
|a_{11}|^2 & \overline{a_{11}a_{12}} \\
a_{12}a_{11} & |a_{12}|^2 + |a_{22}|^2
\end{pmatrix} \]

Since \( AA^* = A^* A \) we have

\[ |a_{11}|^2 + |a_{12}|^2 = |a_{11}|^2. \]

Hence \( a_{12} = 0 \) and \( A \) is diagonal.

For \( n > 2 \), we finish the proof by similar computations. To see this, write

\[ A = \begin{pmatrix}
a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\
0 & a_{22} & a_{23} & \cdots & a_{2n} \\
0 & 0 & a_{33} & \cdots & a_{3n} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & a_{nn}
\end{pmatrix}. \]

Comparing the diagonal entries of the equation \( AA^* = A^* A \) we get,

\[ \sum_{k=i}^{n} |a_{ik}|^2 = |a_{ii}|^2 \]

for \( i = 1, \ldots, n \). So we have \( a_{ik} = 0 \) for all \( k > i \). Therefore, \( A \) is a diagonal matrix.

5.13 (Theorem) Let \( V \) be a finite dimensional inner product space over \( \mathbb{C} \). Let \( T \in L(V, V) \) be a linear operator on \( V \). Then there is an orthonormal basis \( E = \{e_1, \ldots, e_n\} \) such that the matrix of \( T \) with respect to \( E \) is upper triangular.

**Proof.** We will use induction on \( \dim V \). Note that theorem is true when \( \dim V = 1 \).

Since \( \mathbb{C} \) is algebraically closed, the adjoint \( T^* \) has an eigen vector \( v \neq 0 \). Write \( e = v/\|v\| \). The \( e \) is an eigen vector of \( T^* \), and

\[ T^*(e) = ce \quad \text{for some} \quad c \in \mathbb{C}. \]
Let 

\[ W = C e \quad \text{and} \quad V_1 = W^\perp. \]

Since \( W \) is \( T^* \)-invariant, by theorem 5.7, \( V_1 \) is \( T \)-invariant. Let \( T_1 = T_{|V_1} \) be the restriction of \( T \). Then \( T_1 \in L(V_1, V_1) \).

By induction, there is an orthonormal basis \( E = \{e_1, \ldots, e_{n-1}\} \) of \( V_1 \) such that the matrix of \( T_1 \) with respect to \( E \) is upper triangular. Write \( e_n = e \). then the matrix of \( T \) with respect to \( E_0 = \{e_1, \ldots, e_{n-1}, e_n\} \) is upper triangular.

**5.14** Let \( A \in \mathbb{M}_n(\mathbb{C}) \) be any matrix. Then there is an unitary matrix \( U \in \mathcal{U}(n) \subseteq \mathbb{M}_n(\mathbb{C}) \), such that \( U^{-1}AU \) is upper-triangular.

Proof. Proof is an immediate application of theorem 5.13 to the map \( \mathbb{C}^n \to \mathbb{C}^n \) defined by the matrix \( A \).

**5.15 (Theorem)** Let \( V \) be a finite dimensional inner product space over \( \mathbb{C} \). Let \( T \in L(V, V) \) be a normal operator on \( V \). Then there is an orthonormal basis \( E = \{e_1, \ldots, e_n\} \) such that each \( e_i \) is an eigen vector of \( T \).

Proof. By theorem 5.13, \( V \) has an orthonormal basis \( E = \{e_1, \ldots, e_n\} \) such that that matrix \( A \) of \( T \) with respect to \( E \) is upper-triangular.

Since \( T \) is normal, by theorem 5.12, \( A \) is diagonal. Therefore \( e_i \) is an eigen vector of \( T \).

The following is the matrix version of this theorem 5.15.

**5.16** Let \( A \in \mathbb{M}_n(\mathbb{C}) \) be a NORMAL matrix. Then there is an unitary matrix \( U \in \mathcal{U}(n) \subseteq \mathbb{M}_n(\mathbb{C}) \), such that \( U^{-1}AU \) is diagonal.

Proof. Proof is an immediate application of theorem 5.15 to the map \( \mathbb{C}^n \to \mathbb{C}^n \) defined by the matrix \( A \).