MATH 361 - Week 3-5 Notes

Jasraj Sandhu

January 2024

Orthogonality & Orthonormality

Let V be an inner product space.

Definitions

Any $\vec{x}, \vec{y} \in V$ are **orthogonal** if $\langle \vec{x}, \vec{y} \rangle = 0$. Two subsets S and T of V are **orthogonal** if $\langle \vec{x}, \vec{y} \rangle = 0$ for all $\vec{x} \in S$ and $\vec{y} \in T$.

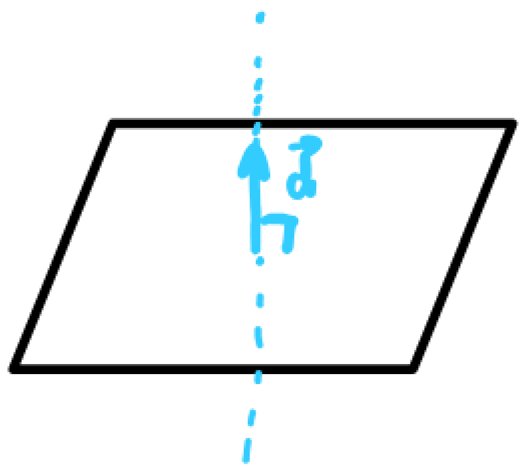
A finite or countable (the latter only applies in infinite dimension) set of non-zero vectors $\{\vec{x}_1, \vec{x}_2, \ldots\}$ is an **orthogonal system** (or **orthogonal**) if $\langle \vec{x}_i, \vec{x}_j \rangle = 0$ for all $i \neq j$. If, in addition, $||\vec{x}_j||^2 = \langle \vec{x}_j, \vec{x}_j \rangle = 1$, we say $\{\vec{x}_1, \vec{x}_2, \ldots\}$ is an **orthonormal system** (or **orthonormal**).

For a finite dimensional V, a basis $\{\vec{b}_1, \ldots, \vec{b}_n\}$ is an **orthogonal basis** if it is an orthogonal system. It is an **orthonormal basis** if it is an orthonormal system.

1. In $\mathbb{R}^3$, the set

$$\left\{ \begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix} \right\}$$

is an orthonormal basis. A line through $(0,0,0)$ with direction $\vec{d}$ is orthogonal to the plane with normal $\vec{d}$.



If $\vec{d} = \begin{pmatrix}a\\b\\c\end{pmatrix}$, the plane is given by $ax + by + cz = 0$, i.e. it consists of all $\begin{pmatrix}x & y & z\end{pmatrix}^T$ so that

$$\begin{pmatrix}a\\b\\c\end{pmatrix} \cdot \begin{pmatrix}x\\y\\z\end{pmatrix} = 0 \ .$$

Proposition (Pythagorean Identity)

If $\{\vec{x}_1, \ldots, \vec{x}_m\}$ is orthogonal, then

$$\left|\left|\sum_{j=1}^{m} \alpha_j \vec{x}_j\right|\right|^2 = \sum_{j=1}^{m} |\alpha_j| \ ||\vec{x}_j||^2 \ .$$

Note that the summation notation here is just expressing the basis vectors as a linear combination.

Special case:

$$||\vec{x} + \vec{y}||^2 = ||\vec{x}||^2 + ||\vec{y}||^2$$

for $\langle \vec{x}, \vec{y} \rangle = 0$ (i.e. when $\vec{x}$ and $\vec{y}$ are orthogonal).

Proof. Let $\{\vec{x}_1, \ldots, \vec{x}_m\}$ be an orthogonal set. Then

$$\begin{aligned}
\left|\left|\sum_{j=1}^{m} \alpha_j \vec{x}_j\right|\right|^2 &= \left(\sqrt{\langle \sum_{j=1}^{m} \alpha_j \vec{x}_j, \sum_{\ell=1}^{m} \alpha_\ell \vec{x}_\ell \rangle} \right)^2 \\
&= \langle \sum_{j=1}^{m} \alpha_j \vec{x}_j, \sum_{\ell=1}^{m} \alpha_\ell \vec{x}_\ell \rangle \\
&= \sum_{j=1}^{m} \alpha_j \langle \vec{x}_j, \sum_{\ell=1}^{m} \alpha_\ell \vec{x}_\ell \rangle \\
&= \sum_{j=1}^{m} \alpha_j \left(\overline{\sum_{\ell=1}^{m} \alpha_\ell} \ \langle \vec{x}_j, \vec{x}_\ell \rangle \right) \\
&= \sum_{j=1}^{m} \alpha_j \left(\sum_{\ell=1}^{m} \overline{\alpha_\ell} \ \langle \vec{x}_j, \vec{x}_\ell \rangle \right) \\
&= \sum_{j=1}^{m} \alpha_j \overline{\alpha_j} \ \langle \vec{x}_j, \vec{x}_j \rangle & (= 0 \text{ if } j \neq \ell) \\
&= \sum_{j=1}^{m} |\alpha_j|^2 \ ||\vec{x}_j||^2 \ .
\end{aligned}$$

□

1. An orthogonal system $\{\vec{x}_1, \ldots, \vec{x}_m\}$ is linearly independent.

   Proof. Let $\{\vec{x}_1, \ldots, \vec{x}_m\}$ be an orthogonal system (so non-zero vectors). Suppose $\alpha_1\vec{x}_1 + \ldots + \alpha_m\vec{x}_m = 0$. Then for $j = 1, \ldots, m$,

$$\begin{aligned} 0 &= \langle 0, \vec{x}_j \rangle \\ &= \langle \alpha_1\vec{x}_1 + \ldots + \alpha_m\vec{x}_m, \vec{x}_j \rangle \\ &= \alpha_1\langle \vec{x}_1, \vec{x}_j \rangle + \ldots + \alpha_m\langle \vec{x}_m, \vec{x}_j \rangle \\ &= \alpha_j\langle \vec{x}_j, \vec{x}_j \rangle \\ &= \alpha_j ||\vec{x}_j||^2 \ , \end{aligned}$$

   where we used the fact that $\langle \vec{x}_i, \vec{x}_j \rangle = 0$ if $i \neq j$ (by definition of the set being orthogonal). Now, we know that $||\vec{x}_j||^2 \neq 0$ since all of the vectors in the orthogonal system are non-zero. So, it must be that $\alpha_j = 0$ in order for $\alpha_j||\vec{x}_j||^2 = 0$ to hold. Thus, since $\alpha_j = 0$, we have shown that all the coefficients of the linear combination are zero, and hence the orthogonal system $\{\vec{x}_1, \ldots, \vec{x}_m\}$ is linearly independent. $\square$

2. Suppose $\{\vec{x}_1, \ldots, \vec{x}_m\}$ are vectors orthogonal to a set S in V. Then $\text{span}\{\vec{x}_1, \ldots, \vec{x}_m\} \perp S$.

Proof. Let $\vec{x}_j \in V$ for $j = 1, \ldots, m$ and let $\vec{y} \in S$. Then since each vector $\vec{x}_j$ is orthogonal to the set S, they are orthogonal to any vector $\vec{y} \in S$. So, it follows from the definition of orthogonality that

$$\langle \vec{x}_j, \vec{y} \rangle = 0 \ .$$

Note that the above also holds from the second part of the definition of orthogonality (i.e. S is a subset of V and V is a subset of itself). Now, we know that

$$\text{span}\{\vec{x}_1, \ldots, \vec{x}_m\} = \{\alpha_1 \vec{x}_1 + \ldots + \alpha_m \vec{x}_m \mid \alpha_1, \ldots, \alpha_m \in \mathbb{C}\} \ .$$

Let $\vec{v} \in \text{span}\{\vec{x}_1, \ldots, \vec{x}_m\}$. Then

$$\vec{v} = \alpha_1 \vec{x}_1 + \ldots + \alpha_m \vec{x}_m \ ,$$

and so

$$\begin{aligned} \langle \vec{v}, \vec{y} \rangle &= \langle \alpha_1 \vec{x}_1 + \ldots + \alpha_m \vec{x}_m, \vec{y} \rangle \\ &= \alpha_1 \langle \vec{x}_1, \vec{y} \rangle + \ldots + \alpha_m \langle \vec{x}_m, \vec{y} \rangle \\ &= \alpha_1 (0) + \ldots + \alpha_m (0) \\ &= 0 + \ldots + 0 \\ &= 0 \ . \end{aligned}$$

So, we have shown that any vector in $\text{span}\{\vec{x}_1, \ldots, \vec{x}_m\}$ is orthogonal to all vectors $\vec{y} \in S$. Thus, $\text{span}\{\vec{x}_1, \ldots, \vec{x}_m\} \perp S$. □

Proposition

For any finite dimensional inner product space V, if $\{\vec{x}_1, \ldots, \vec{x}_n\}$ is an orthogonal basis, then for all $\vec{x} \in V$, we have

$$\vec{x} = \sum_{j=1}^{n} \frac{\langle \vec{x}, \vec{x}_j \rangle}{||\vec{x}_j||^2} \, \vec{x}_j \ .$$

(This is sometimes called the Fourier Expansion of $\vec{x}$, where the $\dfrac{\langle \vec{x}, \vec{x}_j \rangle}{||\vec{x}_j||^2}$ are the Fourier coefficients.)

Proof. Let V be a finite dimensional inner product space with an orthogonal basis $\{\vec{x}_1, \ldots, \vec{x}_n\}$. Suppose $\vec{x} \in V$. Then $\vec{x}$ can be expressed as a linear combination of the basis vectors of V; that is,

$$\vec{x} = \alpha_1 \vec{x}_1 + \ldots + \alpha_n \vec{x}_n = \sum_{j=1}^{n} \alpha_j \vec{x}_j$$

where $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$. So, to solve for α_j, we compute

$$\begin{aligned}
\langle \vec{x}, \vec{x}_j \rangle &= \langle \sum_{\ell=1}^{n} \alpha_\ell \vec{x}_\ell, \vec{x}_j \rangle \\
&= \left(\sum_{\ell=1}^{n} \alpha_\ell \right) \langle \vec{x}_\ell, \vec{x}_j \rangle \\
&= \sum_{\ell=1}^{n} \alpha_\ell \langle \vec{x}_\ell, \vec{x}_j \rangle \\
&= \alpha_1 \langle \vec{x}_1, \vec{x}_j \rangle + \ldots + \alpha_n \langle \vec{x}_n, \vec{x}_j \rangle \\
&= \alpha_j \langle \vec{x}_j, \vec{x}_j \rangle \\
&= \alpha_j ||\vec{x}_j||^2 \ .
\end{aligned}$$

where $\langle \vec{x}_\ell, \vec{x}_j \rangle = 0$ when $\ell \neq j$. Now, solving for α_j gives

$$\alpha_j = \frac{\langle \vec{x}, \vec{x}_j \rangle}{||\vec{x}_j||^2} \ .$$

Thus,

$$\vec{x} = \sum_{j=1}^{n} \alpha_j \vec{x}_j = \sum_{j=1}^{n} \frac{\langle \vec{x}, \vec{x}_j \rangle}{||\vec{x}_j||^2} \, \vec{x}_j \ .$$

$\square$

Orthogonal Projections

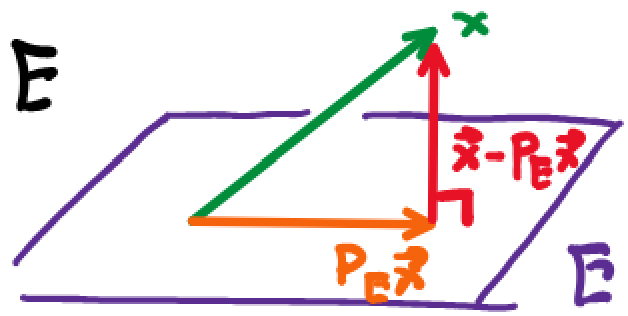
Let E be a subspace of a $\mathbb{C}$ inner product space V. For $\vec{x} \in V$, we wish to construct a vector $P_E\vec{x}$, which we call the **orthogonal projection** of $\vec{x}$ onto E, satisfying

(i) $P_E\vec{x} \in E$.

(ii) $(\vec{x} - P_E\vec{x}) \perp E$.

(iii) The projection is unique.

(iv) $||\vec{x} - P_E\vec{x}|| = \inf ||\vec{x} - \vec{y}||$, where $\vec{y} \in E$ and inf is the **infimum**.



Definition (Orthogonal Projection)

Suppose the subspace $E \subseteq V$ has an **orthogonal** basis $\{\vec{x}_1, \ldots, \vec{x}_m\}$. Define, for any $\vec{x} \in V$,

$$P_E\vec{x} := \sum_{j=1}^{m} \frac{\langle \vec{x}, \vec{x}_j \rangle}{||\vec{x}_j||^2} \; \vec{x}_j$$

with the following properties:

(i) $P_E : V \to E$ (or $V \to V$) is linear, since $\vec{x} \mapsto \langle \vec{x}, \vec{x}_j \rangle \vec{x}_j$ is linear for all j. (This property gives uniqueness).

(ii) $P_E\vec{x} \in E$ for all $\vec{x} \in V$.

(iii) For all $\vec{x} \in V$ and $\vec{x}_j$ of the basis,

$$\begin{aligned}
\langle \vec{x} - P_E\vec{x}, \vec{x}_j \rangle &= \langle \vec{x}, \vec{x}_j \rangle - \langle P_E\vec{x}, \vec{x}_j \rangle \\
&= \langle \vec{x}, \vec{x}_j \rangle - \langle \sum_{\ell=1}^{m} \frac{\langle \vec{x}, \vec{x}_\ell \rangle}{||\vec{x}_\ell||^2} \; \vec{x}_\ell, \vec{x}_j \rangle \\
&= \langle \vec{x}, \vec{x}_j \rangle - \sum_{\ell=1}^{m} \frac{\langle \vec{x}, \vec{x}_\ell \rangle}{||\vec{x}_\ell||^2} \; \langle \vec{x}_\ell, \vec{x}_j \rangle \\
&= \langle \vec{x}, \vec{x}_j \rangle - \left(\frac{\langle \vec{x}, \vec{x}_1 \rangle}{||\vec{x}_1||^2} \; \langle \vec{x}_1, \vec{x}_j \rangle + \ldots + \frac{\langle \vec{x}, \vec{x}_m \rangle}{||\vec{x}_m||^2} \; \langle \vec{x}_m, \vec{x}_j \rangle \right) \\
&= \langle \vec{x}, \vec{x}_j \rangle - \frac{\langle \vec{x}, \vec{x}_j \rangle}{||\vec{x}_j||^2} \; \langle \vec{x}_j, \vec{x}_j \rangle \\
&= \langle \vec{x}, \vec{x}_j \rangle - \frac{\langle \vec{x}, \vec{x}_j \rangle}{||\vec{x}_j||^2} \; ||\vec{x}_j||^2 \\
&= \langle \vec{x}, \vec{x}_j \rangle - \langle \vec{x}, \vec{x}_j \rangle \\
&= 0 \; ,
\end{aligned}$$

where $\langle \vec{x}_\ell, \vec{x}_j \rangle = 0$ when $\ell \neq j$.

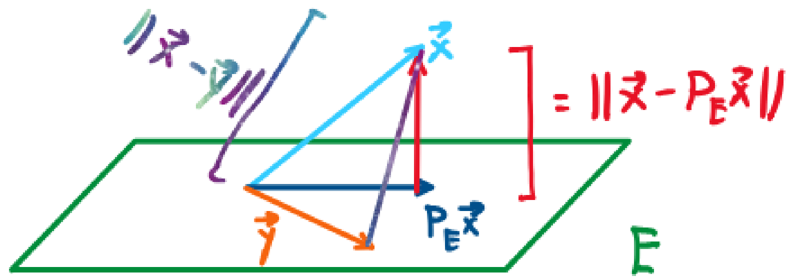
(iv) $||\vec{x} - P_E\vec{x}|| \leq ||\vec{x} - \vec{y}||$ for any $\vec{y} \in E$

Property 4

Let's take a closer look at property (iv) of orthogonal projections:

$$||\vec{x} - P_E\vec{x}|| \le ||\vec{x} - \vec{y}||$$

for any $\vec{y} \in E$.



We have that

$$||\vec{x} - \vec{y}||^2 = ||\vec{x} - P_E\vec{x} + P_E\vec{x} - \vec{y}||^2 \ ,$$

where $\vec{x} - P_E\vec{x} \in E^\perp$ and $P_E\vec{x} - \vec{y} \in E$. Note that

$$\boxed{E^\perp = \{\vec{z} \mid \langle \vec{z}, \vec{w} \rangle = 0 \quad \forall \vec{w} \in E\}}$$

is called the **orthogonal complement**. Now,

$$\begin{aligned} ||\vec{x} - \vec{y}||^2 &= ||\vec{x} - P_E\vec{x} + P_E\vec{x} - \vec{y}||^2 \\ &= ||\vec{x} - P_E\vec{x}||^2 + \underbrace{||P_E\vec{x} - \vec{y}||^2}_{\ge 0} \qquad \text{(Pythag)} \\ &\ge ||\vec{x} - P_E\vec{x}||^2 \ . \end{aligned}$$

Note that Pythagorean's law holds here since $\vec{x} - P_E\vec{x}$ and $P_E\vec{x} - \vec{y}$ are orthogonal (i.e. $\langle \vec{x} - P_E\vec{x}, P_E\vec{x} - \vec{y} \rangle = 0$). In conclusion,

$$\min_{\vec{y} \in E} ||\vec{x} - \vec{y}|| = ||\vec{x} - P_E\vec{x}|| \ .$$

This also implies by the same argument that if $\vec{y} \in E$ and $||\vec{x} - \vec{y}|| = ||\vec{x} - P_E\vec{x}||$, then $\vec{y} = P_E\vec{x}$, which we can show (see next page).

<u>**Remark:**</u> $P_E\vec{x}$ was defined in terms of a specific basis. If $\{\vec{b}_1, \dots, \vec{b}_m\}$ is also a basis for E, is it the case that

$$\sum_{j=1}^m \frac{\langle \vec{x}, \vec{x}_j \rangle}{||\vec{x}_j||^2} \ \vec{x}_j = \sum_{j=1}^m \frac{\langle \vec{x}, \vec{b}_j \rangle}{||\vec{b}_j||^2} \ \vec{b}_j \ ?$$

Both of these satisfy property (iv), so must agree by uniqueness.

1. Show that if $\vec{y} \in E$ and $||\vec{x} - \vec{y}|| = ||\vec{x} - P_E\vec{x}||$, then $\vec{y} = P_E\vec{x}$.

Proof. Let $E \subseteq V$ is a subspace of an inner product space V. Let $\vec{y} \in E$. Suppose $||\vec{x} - \vec{y}|| = ||\vec{x} - P_E\vec{x}||$. We want to show that $\vec{y} = P_E\vec{x}$. It is sufficient to show that $||\vec{y} - P_E\vec{x}|| = 0$. Then

$$\begin{aligned}
||\vec{x} - \vec{y}||^2 &= ||\vec{x} - P_E\vec{x}||^2 \\
&= ||\vec{x} - \vec{y} + \vec{y} - P_E\vec{x}||^2 \\
&= ||\vec{x} - \vec{y}||^2 + ||\vec{y} - P_E\vec{x}||^2 \ .
\end{aligned}$$

From this, we get that

$$\begin{aligned}
||\vec{y} - P_E\vec{x}||^2 + ||\vec{x} - \vec{y}||^2 &= ||\vec{x} - \vec{y}||^2 \\
||\vec{y} - P_E\vec{x}||^2 &= ||\vec{x} - \vec{y}||^2 - ||\vec{x} - \vec{y}||^2 \\
||\vec{y} - P_E\vec{x}||^2 &= 0 \ .
\end{aligned}$$

Thus, $||\vec{y} - P_E\vec{x}|| = 0$, which means that $\vec{y} = P_E\vec{x}$. $\square$

Projections as Transformations/Matrices

- Given $\vec{x}, \vec{y}$ in an inner product space V, we can define a rank 1 transformation as

$$\boxed{\vec{x}\vec{y}^* : \vec{v} \mapsto \langle \vec{v}, \vec{y} \rangle \vec{x}}$$

(any rank 1 linear map on an inner product space is of this form.)

- If $\vec{x}, \vec{y}$ are vectors in $\mathbb{C}^n$, then $\vec{x}\vec{y}^*$ really is the matrix

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \begin{pmatrix} \overline{y_1} & \overline{y_2} & \dots & \overline{y_n} \end{pmatrix} = \begin{pmatrix} x_1\overline{y_1} & x_1\overline{y_2} & \dots & x_1\overline{y_n} \\ \vdots & \vdots & \ddots & \vdots \\ x_n\overline{y_1} & x_n\overline{y_2} & \dots & x_n\overline{y_n} \end{pmatrix} .$$

- P_E is the sum of $m = \dim(E)$ rank 1 maps:

$$P_E = \frac{\vec{x}_1\vec{x}_1^*}{||\vec{x}_1||^2} + \dots + \frac{\vec{x}_m\vec{x}_m^*}{||\vec{x}_m||^2}$$

since

$$\begin{aligned} P_E\vec{x} &= \frac{\langle \vec{x}, \vec{x}_1 \rangle}{||\vec{x}_i||^2}\ \vec{x}_1 + \dots + \dots + \frac{\langle \vec{x}, \vec{x}_m \rangle}{||\vec{x}_m||^2}\ \vec{x}_m \\ &= \frac{\vec{x}_1\vec{x}_1^*}{||\vec{x}_1||^2}\ \vec{x} + \dots + \frac{\vec{x}_m\vec{x}_m^*}{||\vec{x}_m||^2}\ \vec{x}\ , \end{aligned}$$

where from the definition of rank 1 transformations, $\vec{x}_j\vec{x}_j^* = \langle \vec{x}, \vec{x}_j \rangle\ \vec{x}_j$ for $1 \leq j \leq m$. Note that $\vec{x}$ is any vector in V.

Theorem (Gram-Schmidt)

Suppose $\{\vec{x}_1, \ldots, \vec{x}_m\}$ is linearly independent in an inner product space V. Then there is an orthogonal set $\{\vec{y}_1, \ldots, \vec{y}_m\}$ so that

$$\text{span}\{\vec{x}_1, \ldots, \vec{x}_r\} = \text{span}\{\vec{y}_1, \ldots, \vec{y}_r\}$$

for $r = 1, 2, \ldots, m$.

Proof. Set $\vec{y}_1 = \vec{x}_1$ and define $E_r = \text{span}\{\vec{y}_1, \ldots, \vec{y}_r\}$. Also set

$$\begin{aligned}
\vec{y}_2 &= \vec{x}_2 - P_{E_1}\vec{x}_2 \\
\vec{y}_3 &= \vec{x}_3 - P_{E_2}\vec{x}_3 \\
&\vdots \\
\vec{y}_m &= \vec{x}_m - P_{E_{m-1}}\vec{x}_m \ .
\end{aligned}$$

Then

$$\begin{aligned}
&\text{span}\{\vec{y}_1\} = \text{span}\{\vec{x}_1\} \ , \\
&\text{span}\{\vec{y}_1, \vec{y}_2\} = \text{span}\{\vec{x}_1, \vec{x}_2 - P_{E_1}\vec{x}_2\} = \text{span}\{\vec{x}_1, \vec{x}_2\} \\
&\quad \vdots \\
&\text{span}\{\vec{y}_1, \ldots, \vec{y}_m\} = \text{span}\{\vec{x}_1, \ldots, \vec{x}_m\} \ .
\end{aligned}$$

At each stage, by construction, $\vec{y}_r \perp E_{r-1}$, which implies that

$$\vec{y}_r \perp \vec{y}_1, \vec{y}_2, \ldots, \vec{y}_{r-1} \ .$$

This holds for all $r = 1, \ldots, m$. Hence, $\{\vec{y}_1, \ldots, \vec{y}_m\}$ is orthogonal. $\square$

1. Find an orthogonal basis for span $\left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \\ i \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ -i \\ 2 \end{pmatrix} \right\}$ in $\mathbb{C}^4$.

Let $\vec{x}_1 = \begin{pmatrix} 1 & 0 & 1 & i \end{pmatrix}^T$ and $\vec{x}_2 = \begin{pmatrix} -1 & 1 & -i & 2 \end{pmatrix}^T$. Define $E_r = \text{span}\{\vec{y}_1, \vec{y}_2\}$. Set $\vec{y}_1 = \vec{x}_1$. Then

$$\begin{aligned} \vec{y}_2 &= \vec{x}_2 - P_{E_1}\vec{x}_2 \\ &= \vec{x}_2 - \sum_{j=1}^{1} \frac{\langle \vec{x}_2, \vec{x}_j \rangle}{||\vec{x}_1||^2}\ \vec{x}_1 \\ &= \vec{x}_2 - \frac{\langle \vec{x}_2, \vec{x}_1 \rangle}{||\vec{x}_1||^2}\ \vec{x}_1\ . \end{aligned}$$

where $E_1 = \text{span}\{\vec{y}_1\} = \text{span}\{\vec{x}_1\}$. Now, we have that

$$\begin{aligned} \langle \vec{x}_2, \vec{x}_1 \rangle &= \sum_{j=1}^{4} (x_2)_j \overline{(x_1)_j} \\ &= (-1 \cdot \overline{1}) + (1 \cdot \overline{0}) + (-i \cdot \overline{1}) + (2 \cdot \overline{i}) \\ &= (-1 \cdot 1) + (1 \cdot 0) + (-i \cdot 1) + (2 \cdot (-i)) \\ &= -1 + 0 + (-i) + (-2i) \\ &= -1 - 3i\ , \end{aligned}$$

and

$$\begin{aligned} ||\vec{x}_1||^2 &= \langle \vec{x}_1, \vec{x}_1 \rangle \\ &= \sum_{j=1}^{n} (x_1)_j \overline{(x_1)_j} \\ &= (1 \cdot \overline{1}) + (0 \cdot \overline{0}) + (1 \cdot \overline{1}) + (i \cdot \overline{i}) \\ &= (1 \cdot 1) + (0 \cdot 0) + (1 \cdot 1) + (i \cdot (-i)) \\ &= 1 + 0 + 1 + (-i^2) \\ &= 1 + 0 + 1 + (-(-1)) \\ &= 1 + 0 + 1 + 1 \\ &= 3\ . \end{aligned}$$

So,

$$\begin{aligned} \vec{y}_2 &= \vec{x}_2 - \frac{\langle \vec{x}_2, \vec{x}_1 \rangle}{||\vec{x}_1||^2}\ \vec{x}_1 \\ &= \begin{pmatrix} -1 \\ 1 \\ -i \\ 2 \end{pmatrix} - \frac{(-1 - 3i)}{3} \begin{pmatrix} -1 \\ 1 \\ -i \\ 2 \end{pmatrix}\ . \end{aligned}$$

2. Let V be the space of complex coefficient polynomials in real variable t with inner product

$$\langle p, q \rangle = \int_0^1 p(t)\overline{q(t)}\ dt\ .$$

Use Gram-Schmidt to find an orthogonal basis for $\text{span}\{1, t, t^2\}$.

We use the Gram-Schmidt algorithm to find an orthogonal basis for $\text{span}\{1, t, t^2\}$. Let $e_1(t) = 1$, $e_2(t) = t$, and $e_3(t) = t^2$. Define $E_r = \text{span}\{p_1(t), \dots, p_r(t)\}$. Let $p_1(t) = e_1(t) = 1$. Then

$$\begin{aligned} p_2(t) &= e_2(t) - P_{E_1} e_2(t) \\ &= t - \sum_{j=1}^{1} \frac{\langle e_2(t), p_j(t) \rangle}{||p_j(t)||^2}\ p_j(t) \\ &= t - \frac{\langle e_2(t), p_1(t) \rangle}{||p_1(t)||^2}\ p_1(t)\ , \end{aligned}$$

where $E_1 = \text{span}\{p_1(t)\} = \text{span}\{1\}$. So,

$$\begin{aligned} \langle e_2(t), p_1(t) \rangle &= \int_0^1 e_2(t)\overline{p_1(t)}\ dt \\ &= \int_0^1 t \cdot \overline{1}\ dt \\ &= \int_0^1 t \cdot 1\ dt \\ &= \int_0^1 t\ dt \\ &= \left[\frac{t^2}{2} \right]_0^1 \\ &= \frac{1^2}{2} - \frac{0^2}{2} \\ &= \frac{1}{2} - 0 \\ &= \frac{1}{2} \end{aligned}$$

and

$$\begin{aligned} ||p_1(t)||^2 &= \langle p_1(t), p_1(t) \rangle \\ &= \int_0^1 p_1(t)\overline{p_1(t)}\ dt \\ &= \int_0^1 1 \cdot \overline{1}\ dt \end{aligned}$$

$$
\begin{aligned}
&= \int_0^1 1 \cdot 1 \; dt \\
&= \int_0^1 1 \; dt \\
&= \Big[t \Big]_0^1 \\
&= 1 - 0 \\
&= 1 \; .
\end{aligned}
$$

Then

$$
\begin{aligned}
p_2(t) &= t - \frac{\langle e_2(t), p_1(t) \rangle}{||p_1(t)||^2} \; p_1(t) \\
&= t - \frac{1/2}{1} \cdot 1 \\
&= t - \frac{1/2}{1} \\
&= t - \frac{1}{2} \\
&= \frac{2t}{2} - \frac{1}{2} \\
&= \frac{2t-1}{2} \; .
\end{aligned}
$$

Now,

$$
\begin{aligned}
p_3(t) &= e_3(t) - P_{E_2} e_3(t) \\
&= e_3(t) - \sum_{j=1}^{2} \frac{\langle e_3(t), p_j(t) \rangle}{||p_j(t)||^2} \; p_j(t) \\
&= e_3(t) - \left(\frac{\langle e_3(t), p_1(t) \rangle}{||p_1(t)||^2} \; p_1(t) + \frac{\langle e_3(t), p_2(t) \rangle}{||p_2(t)||^2} \; p_2(t) \right) \; ,
\end{aligned}
$$

where $E_2 = \text{span}\{p_1(t), p_2(t)\} = \text{span}\{1, t\}$. So, the evaluating inner products above gives

$$
\begin{aligned}
\langle e_3(t), p_1(t) \rangle &= \int_0^1 e_3(t) \overline{p_1(t)} \; dt \\
&= \int_0^1 t^2 \cdot \overline{1} \; dt \\
&= \int_0^1 t^2 \cdot 1 \; dt \\
&= \int_0^1 t^2 \; dt
\end{aligned}
$$

$$
\begin{aligned}
&= \left[\frac{t^3}{3}\right]_0^1 \\
&= \frac{1^3}{3} - \frac{0^3}{3} \\
&= \frac{1}{3} - 0 \\
&= \frac{1}{3}
\end{aligned}
$$

and

$$
\begin{aligned}
\langle e_3(t), p_2(t)\rangle &= \int_0^1 e_3(t)\overline{p_2(t)}\ dt \\
&= \int_0^1 t^2 \cdot \overline{\left(\frac{2t-1}{2}\right)}\ dt \\
&= \int_0^1 t^2 \cdot \frac{\overline{2t-1}}{\overline{2}}\ dt \\
&= \int_0^1 t^2 \cdot \frac{\overline{2}t-\overline{1}}{2}\ dt \\
&= \int_0^1 t^2 \cdot \frac{2t-1}{2}\ dt \\
&= \int_0^1 \frac{2t^3-t^2}{2}\ dt \\
&= \frac{1}{2}\int_0^1 (2t^3-t^2)\ dt \\
&= \frac{1}{2}\left(\int_0^1 2t^3\ dt - \int_0^1 t^2\ dt\right) \\
&= \frac{1}{2}\left(2\int_0^1 t^3\ dt - \int_0^1 t^2\ dt\right) \\
&= \frac{1}{2}\left(2\left[\frac{t^4}{4}\right]_0^1 - \left[\frac{t^3}{3}\right]_0^1\right) \\
&= \frac{1}{2}\left(2\left(\frac{1}{4}\right) - \frac{1}{3}\right) \\
&= \frac{1}{2}\left(\frac{1}{2} - \frac{1}{3}\right) \\
&= \frac{1}{2}\left(\frac{3}{6} - \frac{2}{6}\right) \\
&= \frac{1}{2}\left(\frac{1}{6}\right) \\
&= \frac{1}{12}\ .
\end{aligned}
$$

Also,

$$
\begin{aligned}
||p_2(t)||^2 &= \langle p_2(t), p_2(t) \rangle \\
&= \int_0^1 p_2(t)\overline{p_2(t)}\ dt \\
&= \int_0^1 \frac{2t-1}{2} \cdot \overline{\left(\frac{2t-1}{2}\right)}\ dt \\
&= \int_0^1 \frac{2t-1}{2} \cdot \frac{\overline{2t-1}}{\overline{2}}\ dt \\
&= \int_0^1 \frac{2t-1}{2} \cdot \frac{\overline{2}t-\overline{1}}{2}\ dt \\
&= \int_0^1 \frac{2t-1}{2} \cdot \frac{2t-1}{2}\ dt \\
&= \int_0^1 \frac{2t-1}{2} \cdot \frac{2t-1}{2}\ dt \\
&= \int_0^1 \frac{4t^2-4t+1}{4}\ dt \\
&= \frac{1}{4}\int_0^1 (4t^2-4t+1)\ dt \\
&= \frac{1}{4}\left(\int_0^1 4t^2\ dt - \int_0^1 4t\ dt + \int_0^1 1\ dt\right) \\
&= \frac{1}{4}\left(4\int_0^1 t^2\ dt - 4\int_0^1 t\ dt + \int_0^1 1\ dt\right) \\
&= \frac{1}{4}\left(4\left[\frac{t^3}{3}\right]_0^1 - 4\left[\frac{t^2}{2}\right]_0^1 + \Big[t\Big]_0^1\right) \\
&= \frac{1}{4}\left(4\left(\frac{1}{3}\right) - 4\left(\frac{1}{2}\right) + 1\right) \\
&= \frac{1}{4}\left(\frac{4}{3} - 2 + 1\right) \\
&= \frac{1}{4}\left(\frac{4}{3} - 1\right) \\
&= \frac{1}{4}\left(\frac{4}{3} - \frac{3}{3}\right) \\
&= \frac{1}{4}\left(\frac{1}{3}\right) \\
&= \frac{1}{12}\ .
\end{aligned}
$$

Then

$$\begin{aligned}
p_3(t) &= e_3(t) - \left(\frac{\langle e_3(t), p_1(t) \rangle}{||p_1(t)||^2}\ p_1(t) + \frac{\langle e_3(t), p_2(t) \rangle}{||p_2(t)||^2}\ p_2(t) \right) \\
&= t^2 - \left(\frac{1/3}{1} \cdot 1 + \frac{1/12}{1/12} \cdot \frac{2t-1}{2} \right) \\
&= t^2 - \left(\frac{1}{3} + \frac{2t-1}{2} \right) \\
&= t^2 - \left(\frac{3}{6} + \frac{6t-3}{6} \right) \\
&= t^2 - \left(\frac{3+6t-3}{6} \right) \\
&= t^2 - \frac{6t}{6} \\
&= t^2 - t\ .
\end{aligned}$$

Therefore, an orthogonal basis for $\{e_1.e_2, e_3\} = \{1, t, t^2\}$ is

$$\{p_1(t), p_2(t), p_3(t)\} = \left\{ 1, \frac{2t-1}{2}, t^2 - t \right\}\ .$$

Of course, the original basis and the newly constructed orthogonal basis have the same span.

Recall that if $S \subseteq V$ (not necessarily a subspace of V), then

$$S^\perp := \{\vec{y} \in V : \langle \vec{x}, \vec{y} \rangle = 0 \text{ for all } \vec{x} \in S\} \ .$$

Proposition

$S^\perp$ is a subspace for any subset S of an inner product space V.

Proof. If $\vec{y}$ and $\vec{z}$ are in $S^\perp$, then for all $\vec{x} \in S$,

$$\begin{aligned}
\langle \vec{y} + \vec{z}, \vec{x} \rangle &= \langle 1\vec{y} + 1\vec{z}, \vec{x} \rangle \\
&= 1\langle \vec{y}, \vec{x} \rangle + 1\langle \vec{z}, \vec{x} \rangle && \text{(sesquilinearity)} \\
&= \langle \vec{y}, \vec{x} \rangle + \langle \vec{z}, \vec{x} \rangle \\
&= 0 + 0 \\
&= 0 \ ,
\end{aligned}$$

where $\langle \vec{y}, \vec{x} \rangle = 0$ and $\langle \vec{z}, \vec{x} \rangle = 0$ since S and $S^\perp$ are orthogonal, and $\vec{y}, \vec{z} \in S^\perp$ while $\vec{x} \in S$. So, $y + z \in S^\perp$. Similarly, if $\vec{y} \in S^\perp$ and $\alpha \in \mathbb{C}$, then

$$\begin{aligned}
\langle \alpha\vec{y}, \vec{x} \rangle &= \alpha\langle \vec{y}, \vec{x} \rangle \\
&= \alpha \cdot 0 \\
&= 0 \ ,
\end{aligned}$$

which means that $\alpha\vec{y} \in S^\perp$. So, $\alpha\vec{y} \in S^\perp$. Finally, $0 \in S^\perp$ because $\langle 0, \vec{x} \rangle = 0$. Thus, by the subspace test, since $0 \in S^\perp$, $S^\perp$ is closed under addition, and $S^\perp$ is closed under scalar multiplication, we can conclude that $S^\perp$ is indeed a subspace. $\square$

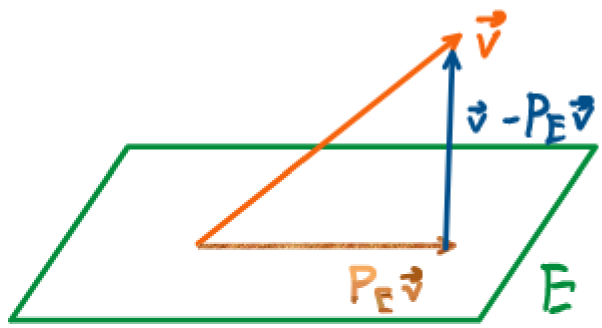
Proposition

If E is a subspace of an inner product space V, then $V = E \oplus E^{\perp}$.

Proof. For any $\vec{v} \in V$, we have

$$\begin{aligned} \vec{v} &= 0 + \vec{v} \\ &= P_E\vec{x} - P_E\vec{x} + \vec{v} \\ &= P_E\vec{x} + \vec{v} - P_E\vec{x} \\ &= P_E\vec{x} + (\vec{v} - P_E\vec{x}) \ , \end{aligned}$$

where $P_E\vec{x} \in E$ and $\vec{v} - P_E\vec{v} \in E^{\perp}$.



For uniqueness, suppose $\vec{v} = \vec{x} + \vec{y}$ with $\vec{x} \in E$ and $\vec{y} \in E^{\perp}$. Then

$$\begin{aligned} P_E\vec{v} &= P_E(\vec{x} + \vec{y}) \\ &= P_E\vec{x} + P_E\vec{y} && \text{(by linearity of } P_E\text{)} \\ &= \vec{x} + 0 \\ &= \vec{x} \ , \end{aligned}$$

where $P_E\vec{x} = \vec{x}$ since $\vec{x} \in E$, and $P_E\vec{y} = 0$ since $\vec{y} \in E^{\perp}$ and $E^{\perp} \perp E$. Hence,

$$\vec{y} = \vec{v} - \vec{x} = \vec{v} - P_E\vec{v} \ .$$

Note also that $E \cap E^{\perp} = \{0\}$. Indeed, if $\vec{x} \in E \cap E^{\perp}$, then

$$||\vec{x}||^2 = \langle \vec{x}, \vec{x} \rangle = 0 \ ,$$

where $\vec{x} \in E$ and $\vec{x} \in E^{\perp}$, and so $\vec{x} = 0$. $\square$

Proposition

Let E be a subspace of an inner product space V. Then

$$P_{E^\perp} = I - P_E \ .$$

Proof. For any $\vec{v} \in V$, we can write

$$\begin{aligned}
\vec{v} &= 0 + \vec{v} \\
&= P_E\vec{v} - P_E\vec{v} + \vec{v} \\
&= P_E\vec{v} + \vec{v} - P_E\vec{v} \\
&= P_E\vec{v} + (\vec{v} - P_E\vec{v}) \\
&= P_E\vec{v} + (I - P_E)\vec{v} \ ,
\end{aligned}$$

where $\vec{v} - P_E\vec{v} \in E^\perp$. Then

$$\begin{aligned}
P_{E^\perp}\vec{v} &= P_{E^\perp}(P_E\vec{v} + (I - P_E)\vec{v}) \\
&= P_{E^\perp}P_E\vec{v} + P_{E^\perp}(I - P_E)\vec{v} \\
&= 0 + (I - P_E)\vec{v} \\
&= (I - P_E)\vec{v} \ ,
\end{aligned}$$

where $P_{E^\perp}P_E\vec{v} = 0$ since

$$(E^\perp)^\perp = E \ ,$$

and

$$P_{E^\perp}(I - P_E)\vec{v} = \ldots = (I - P_E)\vec{v} \ .$$

This is true for any vector $\vec{v}$, so

$$P_{E^\perp} = I - P_E \ .$$

$\square$

Remark: If V has an orthogonal basis (noting that $V = E \oplus E^\perp$)

$$\{ \underbrace{\vec{x}_1, \dots, \vec{x}_m}_{\text{orth. basis for } E} , \underbrace{\vec{y}_{m+1}, \dots, \vec{y}_n}_{\text{orth. basis for } E^\perp} \} ,$$

then

$$\vec{v} = \sum_{j=1}^{m} \frac{\langle \vec{v}, \vec{x}_j \rangle}{||\vec{x}_j||^2} \vec{x}_j + \sum_{j=m+1}^{n} \frac{\langle \vec{v}, \vec{y}_j \rangle}{||\vec{y}_j||^2} \vec{y}_j = P_E \vec{v} + P_{E^\perp} \vec{v}$$

which implies that

$$(I - P_E)\vec{v} = P_{E^\perp} \vec{v} \ . \tag{1}$$

Now, given the basis, what are the matrix representations for these?

$$[P_E] = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{bmatrix} = \begin{bmatrix} I_{m \times m} & 0 \\ 0 & 0 \end{bmatrix} ,$$

$$[I - P_E] = \begin{bmatrix} 0 & 0 \\ 0 & I_{(n-m)\times(n-m)} \end{bmatrix}$$

Question: If A is diagonalizable and has only $\{0, 1\}$ as its eigenvalues, is it an orthogonal projection onto some subspace? No! We require orthogonal diagonalization (more on that later).

The Adjoint of a Linear Transformation

Recall that if $A \in M_{m\times n}(\mathbb{C})$, we defined

$$A^* = [A]^* = [a_{ij}]^* = [\overline{a_{ji}}] \in M_{n\times m}(\mathbb{C}) \ .$$

For the standard inner product $\langle \vec{x}, \vec{y} \rangle = \sum_{j=1}^{n} x_j \overline{y_j}$ in $\mathbb{C}^n$, the adjoint satisfies

$$\underbrace{\langle A\vec{x}, \vec{y} \rangle}_{\substack{\text{inner product} \\ \text{in } \mathbb{C}^m}} \quad = \quad \underbrace{\langle \vec{x}, A^*\vec{y} \rangle}_{\substack{\text{inner product} \\ \text{in } \mathbb{C}^n}}$$

for $\vec{x} \in \mathbb{C}^n$ and $\vec{y} \in \mathbb{C}^m$. Why? We can compute the following. If we let

$$A\vec{x} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + \ldots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n \end{bmatrix} = \begin{bmatrix} w_1 \\ \vdots \\ w_m \end{bmatrix} =: \vec{w} \ ,$$

and

$$A^*\vec{y} = \begin{bmatrix} \overline{a_{11}} & \dots & \overline{a_{m1}} \\ \vdots & \ddots & \vdots \\ \overline{a_{1n}} & \dots & \overline{a_{mn}} \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} \overline{a_{11}}y_1 + \ldots + \overline{a_{m1}}y_m \\ \vdots \\ \overline{a_{1n}}y_1 + \ldots + \overline{a_{mn}}y_m \end{bmatrix} = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} =: \vec{z} \ ,$$

where $\vec{w} \in \mathbb{C}^m$ and $\vec{z} \in \mathbb{C}^n$, then

$$\begin{aligned}
\langle A\vec{x}, \vec{y} \rangle &= \langle \vec{w}, \vec{y} \rangle && \text{(inner prod. in } \mathbb{C}^m) \\
&= \sum_{j=1}^{m} w_j \overline{y_j} \\
&= w_1\overline{y_1} + \ldots + w_m\overline{y_m} \\
&= \overline{y_1}w_1 + \ldots + \overline{y_m}w_m \\
&= \begin{bmatrix} \overline{y_1} & \dots & \overline{y_m} \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_m \end{bmatrix} \\
&= \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}^* \begin{bmatrix} w_1 \\ \vdots \\ w_m \end{bmatrix} \\
&= \vec{y}^*\vec{w} \\
&= \vec{y}^*A\vec{x} \\
&= \vec{y}^*(A^*)^*\vec{x} \\
&= (A^*\vec{y})^*\vec{x} && ((CD)^* = D^*C^*)
\end{aligned}$$

$$
\begin{aligned}
&= \vec{z}^{\,*}\vec{x} \\
&= \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}^* \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \\
&= \begin{bmatrix} \overline{z_1} & \dots & \overline{z_n} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \\
&= \overline{z_1}x_n + \ldots + \overline{z_n}x_n \\
&= x_1\overline{z_1} + \ldots + x_n\overline{z_n} \\
&= \sum_{j=1}^{n} x_j\overline{z_j} \\
&= \langle \vec{x}, \vec{z} \rangle \\
&= \langle \vec{x}, A^*\vec{y} \rangle \ . && \text{(def.)}
\end{aligned}
$$

Proposition

Let V and W be inner product spaces of finite dimension. For any linear transformation $A : V \to W$, there is a unique linear transformation

$$A^* : W \to V$$

that satisfies

$$\langle A\vec{v}, \vec{w}\rangle_W = \langle \vec{v}, A^*\vec{w}\rangle_V \ .$$

Note that $\vec{v} \in V$ and $\vec{w} \in W$.

Proof. For uniqueness, if B and C were both linear maps satisfying

$$\langle A\vec{v}, \vec{w}\rangle = \langle v, B\vec{w}\rangle = \langle v, C\vec{w}\rangle$$

for all $\vec{v}, \vec{w}$, then we get that

$$\langle \vec{v}, (B - C)\vec{w}\rangle = 0$$

for all $\vec{v}, \vec{w}$. Setting $\vec{v} = (B - C)\vec{w}$ gives

$$||(B - C)\vec{w}||^2 = \langle (B - C)\vec{w}, (B - C)\vec{w}\rangle = 0$$

for all $\vec{w}$. So, this implies that

$$(B - C)\vec{w} = 0$$

for all $\vec{w}$ (by the property of non-degeneracy), and hence $B = C$. $\square$

(This proves something about inner product space maps: $\langle A\vec{v}, \vec{w}\rangle = 0$ iff $A = 0$)

- For existence, suppose $A : V \to W$ is linear, $\{\vec{b}_1, \dots, \vec{b}_n\}$ is an orthonormal basis for V, and $\{\vec{d}_1, \dots, \vec{d}_n\}$ is an orthonormal basis for W. If $a_{ij} := \langle A\vec{b}_j, \vec{d}_i \rangle$, then $[a_{ij}]$ is the matrix representation $[A]_{BD}$.

- Define A^* to be the linear map from W to V (i.e, $A^* : W \to V$) so that

$$[A^*]_{DB} := [A]^*_{BD} = [\overline{a_{ji}}] \ . \qquad (\star)$$

By definition,

$$a_{ij} = \langle A\vec{b}_j, \vec{d}_i \rangle \implies \overline{a_{ji}} = \langle \vec{d}_j, A\vec{b}_i \rangle$$

and

$$\overline{a_{ji}} = \langle A^*\vec{d}_j, \vec{b}_i \rangle \ . \qquad \text{(by } (\star))$$

So, we have $\langle \vec{d}_j, A\vec{b}_i \rangle = \langle A^*\vec{d}_j, \vec{b}_i \rangle$ for all i, j. Since these are bases, this gives

$$\langle \vec{w}, A\vec{v} \rangle = \langle A^*\vec{w}, \vec{v} \rangle$$

for all $\vec{v}, \vec{w}$. Equivalently,

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

1. Suppose V = polynomials of degree ≤ 1 and W = polynomials of degree ≤ 2. Suppose both V and W have an inner product such that $\{1, t\}$ and $\{1, t, t^2\}$ are an orthonormal bases for V and W, respectively. For example, if $p(t) = 1$ and $q(t) = 2t + 2$, then

$$\begin{aligned} \langle p, q \rangle &= \langle 1, 2t + 2 \rangle \\ &= 2\langle 1, t \rangle + 2\langle 1, 1 \rangle \\ &= 2(0) + 2(1) \\ &= 2 \ , \end{aligned}$$

where we used the fact that the inner product of any two distinct elements of an orthogonal basis is zero, and the inner product of an element from an orthonormal basis with itself (aka the norm squared) is 1. Now, define a linear map $A : V \to W$ by

$$Ap(t) = \big(p(0) + p(1)\big)t + p(2)t^2 \ .$$

Let $e_1 = 1$ and $e_2 = t$ (the basis vectors of V). Then the basis representation is

$$\begin{aligned} Ae_1 &= A1 = (1 + 1)t + 1t^2 = 2t + t^2 \\ Ae_2 &= At = (0 + 1)t + 2t^2 = t + 2t^2 \ , \end{aligned}$$

and so the matrix representation of A is

$$[A] = \begin{bmatrix} Ae_1 & Ae_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 2 & 1 \\ 1 & 2 \end{bmatrix} \ ,$$

where the first column corresponds to the coefficients of Ae_1 and the second column corresponds to the coefficients of Ae_2. Then

$$[A^*] = [A]^* = \begin{bmatrix} 0 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix} \ ,$$

where the first column corresponds to $[A^*1]$, the second column corresponds to $[A^*t]$, and the third column corresponds to $[A^*t^2]$. Hence, for any polynomial $a + bt + ct^2 \in W$,

$$\begin{aligned} A^*(a + bt + ct^2) &= aA^*1 + bA^*t + cA^*t^2 \\ &= 0 + b(2 + t) + c(1 + 2t) \\ &= (2b + c) + (b + 2c)t \ . \end{aligned}$$

2. Let $A : M_{22}(\mathbb{C}) \to \mathbb{C}^3$, where $M_{22}(\mathbb{C})$ has an inner product defined as

$$\langle C, D\rangle := \mathrm{tr}(CD^*)$$

with orthonormal basis

$$\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$

and $\mathbb{C}^3$ has the standard inner product with c.o.b, be the transformation given by

$$A \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a-b \\ 2c \\ c+d \end{bmatrix} .$$

Applying the transformation to each of the basis vectors of $M_{22}(\mathbb{C})$ gives

$$\begin{aligned}
A \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} &= \begin{bmatrix} 1-0 \\ 2\cdot 0 \\ 0+0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \\
A \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} &= \begin{bmatrix} 0-1 \\ 2\cdot 0 \\ 0+0 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} \\
A \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} &= \begin{bmatrix} 0-0 \\ 2\cdot 1 \\ 1+0 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} \\
A \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} &= \begin{bmatrix} 0-0 \\ 2\cdot 0 \\ 0+1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} .
\end{aligned}$$

It might be worth noting from the above results that

$$\begin{aligned}
A \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} &= \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \vec{e}_1 , \\
A \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} &= \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} = -\vec{e}_1 , \\
A \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} &= \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} = 2\vec{e}_2 + \vec{e}_3 , \\
A \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} &= \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \vec{e}_3 .
\end{aligned}$$

Anyways, we get that the matrix representation of A is

$$[A] = \begin{bmatrix} \vec{e}_1 & -\vec{e}_1 & 2\vec{e}_2 + \vec{e}_3 & \vec{e}_3 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

Then

$$[A^*] = [A]^* = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Hence, for any vector $\begin{bmatrix} a & b & c \end{bmatrix}^T \in \mathbb{C}^3$,

$$\begin{aligned} A^* \begin{bmatrix} a \\ b \\ c \end{bmatrix} &= \begin{bmatrix} 1 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} \\ &= \begin{bmatrix} a + 0b + 0c \\ -a + 0b + 0c \\ 0a + 2b + c \\ 0a + 0b + c \end{bmatrix} \\ &= \begin{bmatrix} a \\ -a \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 2b \\ 0 \end{bmatrix} + \begin{bmatrix} a \\ -a \\ c \\ c \end{bmatrix} \\ &= a \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix} + b \begin{bmatrix} 0 \\ 0 \\ 2 \\ 0 \end{bmatrix} + c \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \end{aligned}$$

Properties of The Adjoint

Proposition (Properties of the Adjoint)

(i) $(A+B)^* = A^* + B^*$

(ii) $(\alpha A)^* = \overline{\alpha} A^*$

(iii) $(AB)^* = B^* A^*$

(iv) $(A^*)^* = A$

Note that for properties (i) and (ii), $A \mapsto A^*$ is a conjugate linear map.

Proof. We prove the above properties.

(i) Let A and B be $m \times n$. Then

$$A + B = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{bmatrix}$$
$$= \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \dots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \dots & a_{2n}+b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \dots & a_{mn}+b_{mn} \end{bmatrix}.$$

Thus,

$$(A+B)^* = \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \dots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \dots & a_{2n}+b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \dots & a_{mn}+b_{mn} \end{bmatrix}^*$$
$$= \begin{bmatrix} \overline{a_{11}+b_{11}} & \overline{a_{21}+b_{21}} & \dots & \overline{a_{m1}+b_{m1}} \\ \overline{a_{12}+b_{12}} & \overline{a_{22}+b_{22}} & \dots & \overline{a_{m2}+b_{m2}} \\ \vdots & \vdots & \ddots & \vdots \\ \overline{a_{1n}+b_{1n}} & \overline{a_{2n}+b_{2n}} & \dots & \overline{a_{mn}+b_{mn}} \end{bmatrix}$$
$$= \begin{bmatrix} \overline{a_{11}} & \overline{a_{21}} & \dots & \overline{a_{m1}} \\ \overline{a_{12}} & \overline{a_{22}} & \dots & \overline{a_{m2}} \\ \vdots & \vdots & \ddots & \vdots \\ \overline{a_{1n}} & \overline{a_{2n}} & \dots & \overline{a_{mn}} \end{bmatrix} + \begin{bmatrix} \overline{b_{11}} & \overline{b_{21}} & \dots & \overline{b_{m1}} \\ \overline{b_{12}} & \overline{b_{22}} & \dots & \overline{b_{m2}} \\ \vdots & \vdots & \ddots & \vdots \\ \overline{b_{1n}} & \overline{b_{2n}} & \dots & \overline{b_{mn}} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}^* + \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{bmatrix}^*$$

$$= A^* + B^* \; .$$

Alternatively, let $A : V \to W$ and $B : V \to W$ be linear operators, where $\dim V = n$ and $\dim W = m$. Then A and B are $m \times n$. So, $A^* : W \to V$ and $B^* : W \to V$ are linear operators. Then for $\vec{v} \in V$ and $\vec{w} \in W$,

$$\begin{aligned} \langle (A+B)\vec{v}, \vec{w} \rangle &= \langle A\vec{v} + B\vec{v}, \vec{w} \rangle \\ &= \langle A\vec{v}, \vec{w} \rangle + \langle B\vec{v}, \vec{w} \rangle \\ &= \langle \vec{v}, A^*\vec{w} \rangle + \langle \vec{v}, B^*\vec{w} \rangle \\ &= \langle \vec{v}, A^*\vec{w} + B^*\vec{w} \rangle \\ &= \langle \vec{v}, (A^* + B^*)\vec{w} \rangle \; , \end{aligned}$$

where $\langle (A+B)\vec{v}, \vec{w} \rangle$ is an inner product in W and $\langle \vec{v}, (A^* + B^*)\vec{w} \rangle$ is an inner product in V. Since these inner products are equal, $(S+T)* = S^* + T^*$.

(ii) Let A be $m \times n$ and let α be a scalar. Then

$$(\alpha A)^* = \left(\alpha \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \right)^*$$

$$= \begin{bmatrix} \alpha(a_{11}) & \alpha(a_{12}) & \dots & \alpha(a_{1n}) \\ \alpha(a_{21}) & \alpha(a_{22}) & \dots & \alpha(a_{2n}) \\ \vdots & \vdots & \ddots & \vdots \\ \alpha(a_{m1}) & \alpha(a_{m2}) & \dots & \alpha(a_{mn}) \end{bmatrix}^*$$

$$= \begin{bmatrix} \overline{\alpha(a_{11})} & \overline{\alpha(a_{21})} & \dots & \overline{\alpha(a_{m1})} \\ \overline{\alpha(a_{12})} & \overline{\alpha(a_{22})} & \dots & \overline{\alpha(a_{m2})} \\ \vdots & \vdots & \ddots & \vdots \\ \overline{\alpha(a_{1n})} & \overline{\alpha(a_{2n})} & \dots & \overline{\alpha(a_{mn})} \end{bmatrix}$$

$$= \begin{bmatrix} \overline{\alpha}\overline{(a_{11})} & \overline{\alpha}\overline{(a_{21})} & \dots & \overline{\alpha}\overline{(a_{m1})} \\ \overline{\alpha}\overline{(a_{12})} & \overline{\alpha}\overline{(a_{22})} & \dots & \overline{\alpha}\overline{(a_{m2})} \\ \vdots & \vdots & \ddots & \vdots \\ \overline{\alpha}\overline{(a_{1n})} & \overline{\alpha}\overline{(a_{2n})} & \dots & \overline{\alpha}\overline{(a_{mn})} \end{bmatrix}$$

$$
\begin{aligned}
&= \overline{\alpha} \begin{bmatrix} \overline{a_{11}} & \overline{a_{21}} & \dots & \overline{a_{m1}} \\ \overline{a_{12}} & \overline{a_{22}} & \dots & \overline{a_{m2}} \\ \vdots & \vdots & \ddots & \vdots \\ \overline{a_{1n}} & \overline{a_{2n}} & \dots & \overline{a_{mn}} \end{bmatrix} \\
&= \overline{\alpha} \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}^* \\
&= \overline{\alpha} A^* \ .
\end{aligned}
$$

(iii) We have that

$$
\begin{aligned}
\langle \vec{v}, (AB)^* \vec{w} \rangle &= \langle (AB)\vec{v}, \vec{w} \rangle && \text{(def.)} \\
&= \langle AB\vec{v}, \vec{w} \rangle \\
&= \langle A(B\vec{v}), \vec{w} \rangle \\
&= \langle B\vec{v}, A^*\vec{w} \rangle && \text{(def.)} \\
&= \langle \vec{v}, B^*A^*\vec{w} \rangle \ . && \text{(def.)}
\end{aligned}
$$

Thus, for all $\vec{v}, \vec{w}$ in some inner product space, $(AB)^* = B^*A^*$.

(iv) We have that

$$
\begin{aligned}
& \langle A\vec{v}, \vec{w} \rangle = \langle \vec{v}, A^*\vec{w} \rangle && \text{(def.)} \\
\iff & \langle \vec{w}, A\vec{v} \rangle = \langle A^*\vec{w}, \vec{v} \rangle && \text{(def.)} \\
\iff & \langle \vec{w}, A\vec{v} \rangle = \langle \vec{w}, (A^*)^*\vec{v} \rangle && \text{(def.)} \\
\iff & A = (A^*)^*
\end{aligned}
$$

$\square$

Proposition

Let V and W be inner product spaces and $A : V \to W$. Then

(i) $\ker A^* = (\operatorname{ran} A)^\perp$ (in W)

(ii) $\operatorname{ran} A^* = (\ker A)^\perp$ (in V)

(iii) $\ker A = (\operatorname{ran} A^*)^\perp$ (in V)

(iv) $\operatorname{ran} A = (\ker A^*)^\perp$ (in W)

Proof. Suppose $A : V \to W$ with V and W as inner product spaces. Then $A^* : W \to V$. We prove the above properties. Note that since all these properties involve equality of sets, we must show they are subsets of each other.

(i) Note that

$$\ker A^* = \{\vec{x} \in W : A^*\vec{x} = 0\}$$

and

$$\operatorname{ran} A = \{A\vec{x} \in W : \vec{x} \in V\} \ .$$

So, since $\operatorname{ran} A$ is a subspace of W (and a subset of W),

$$(\operatorname{ran} A)^\perp = \{\vec{x} \in W : \langle \vec{u}, \vec{x} \rangle = 0 \text{ for all } \vec{u} \in \operatorname{ran} A\} \ .$$

Now, let $\vec{w} \in \ker A^*$. Also, let $\vec{v} \in V$. Since $\vec{w} \in \ker A^*$, we have that $A^*\vec{w} = 0$. Now, since every vector is orthogonal to the zero vector, we get that $\langle \vec{v}, 0 \rangle = 0$. Then since $A^*\vec{w} = 0$, it holds that

$$\langle \vec{v}, A^*\vec{w} \rangle = 0 \ .$$

Now, recall from the proposition of the adjoint that

$$\langle \vec{v}, A^*\vec{w} \rangle_V = \langle A\vec{v}, \vec{w} \rangle_W \ .$$

Then

$$\langle A\vec{v}, \vec{w} \rangle = \langle \vec{v}, A^*\vec{w} \rangle = 0 \ .$$

Since $\langle A\vec{v}, \vec{w} \rangle = 0$, this means that $A\vec{v} \perp \vec{w}$. Note that since $\vec{w} \in \ker A^*$, it also means that $\vec{w} \in W$. Since $A\vec{v} \in W$ (it is an operator from V to W after all), it is in the range of A; that is, $A\vec{v} \in \operatorname{ran} A$. Finally, since $A\vec{v} \in \operatorname{ran} A$ and $\langle A\vec{v}, \vec{w} \rangle = 0$, where $\vec{w} \in W$, we can conclude that $\vec{w} \in (\operatorname{ran} A)^\perp$. The converse of this property is the same argument in reverse.

(ii) Note that

$$\operatorname{ran} A^* = \{A^*\vec{w} \in V : \vec{w} \in W\}$$

and

$$\ker A = \{\vec{v} \in V : A\vec{v} = 0\} \ .$$

So, because $\ker A$ is a subspace of V,

$$(\ker A)^\perp = \{\vec{v} \in V : \langle \vec{u}, \vec{v} \rangle = 0 \ \text{ for all } \vec{u} \in \ker A\} \ .$$

To prove property (ii), it suffices to show that

$$(\operatorname{ran} A^*)^\perp = (\ker A)^{\perp\perp} = \ker A \ .$$

Since $\operatorname{ran} A^*$ is a subspace of V (since $A^* : W \to V$), we have that

$$(\operatorname{ran} A^*)^\perp = \{\vec{v} \in v : \langle \vec{u}, \vec{v} \rangle = 0 \ \text{ for all } \vec{u} \in \operatorname{ran} A^*\} \ .$$

Let $\vec{y} \in (\operatorname{ran} A^*)^\perp$. Then consequently, $\vec{y} \in V$. Let $\vec{u} \in \operatorname{ran} A^*$. Then $\vec{u} = A^*\vec{w} \in V$ for all $\vec{w} \in W$. Since $\vec{y} \in (\operatorname{ran} A^*)^\perp$ and $\vec{u} \in \operatorname{ran} A^*$, we get that

$$\langle A^*\vec{w}, \vec{y} \rangle = \langle \vec{u}, \vec{y} \rangle = 0 \ ,$$

where the inner product is in V. We want to show that $\vec{y} \in \ker A$; that is, $A\vec{y} = 0$. So, we get that

$$0 = \langle A^*\vec{w}, \vec{y} \rangle = \langle \vec{w}, A\vec{y} \rangle \ ,$$

which is an inner product in W, since $\vec{w}, A\vec{y} \in W$. Now, since $\langle \vec{w}, A\vec{y} \rangle = 0$, this means that $A\vec{y}$ is orthogonal to every vector $\vec{w} \in W$. This implies that $A\vec{y} \perp W$, and so $A\vec{y} = 0$ (since $W^\perp = \{0\}$). Thus, since $A\vec{y} = 0$, we can conclude that $\vec{y} \in \ker A$. The converse of this property is the same argument in reverse.

(iii) Same as (i) and (ii) applied to A^* instead of A (using $A^{**} = A$).

(iv) Similar reasoning as (iii).

$\square$

Proposition

Let V be an inner product space and $P : V \to V$. Then P is an orthogonal projection if and only if

$$P^2 = P = P^* \ .$$

Note here that $P^2 = P$ means that P is **idompotent**, and $P = P^*$ means that P is **self-adjoint**.

Proof. Suppose V is an inner product space and E is a subspace of V. Let $P : V \to V$ be a linear operator.

($\Longrightarrow$) Suppose $P_E = \displaystyle\sum_{j=1}^{m} \vec{x}_j \vec{x}_j^*$ is an orthogonal projection. $\square$

Remark: Recall that P_E has a matrix representation $\begin{bmatrix} I_E & 0 \\ 0 & 0 \end{bmatrix}$, which clearly satisfies $P^2 = P = P^*$.

Proposition

Let V and W be inner product spaces and $A : V \to W$. Then

(i) $\ker A = \ker A^*A$

(ii) $\operatorname{rank} A = \operatorname{rank} A^*A$

(iii) If $\ker A = \{0\}$ (i.e. A is one-to-one), then A is left invertible.

(iv) If $\operatorname{ran} A = W$ (i.e. A is onto), then A is right invertible.

Proof. Suppose V and W are inner product spaces. Let $A : V \to W$ be a linear operator. Then $A^* : W \to V$ is a linear operator. We prove the above properties.

(i) We know that

$$\ker A = \{\vec{v} \in V : A\vec{v} = 0_W\} .$$

We also know that A^*A is a composition of linear maps. Since $A : V \to W$ and $A^* : W \to V$, it follows that $A^*A : V \to V$. So,

$$\ker A^*A = \{\vec{v} \in V : A^*A\vec{v} = 0_V\} .$$

($\Longrightarrow$) Let $\vec{x} \in \ker A$. Then $A\vec{x} = 0_W$. (Note that $\vec{x} \in \ker A$ implies $\vec{x} \in V$). Then it also holds that

$$||A\vec{x}||^2 = ||0||^2 = 0 .$$

From the definition of norm, we get that

$$0 = ||A\vec{x}||^2 = \langle A\vec{x}, A\vec{x}\rangle .$$

Then

$$0 = \langle A\vec{x}, A\vec{x}\rangle = \langle A^*A\vec{x}, \vec{x}\rangle .$$

Since $\langle A^*A\vec{x}, \vec{x}\rangle = 0$, this means that $A^*A\vec{x}$ is orthogonal to $\vec{x}$ in V...

($\Longleftarrow$) Let $\vec{x} \in \ker A^*A$. Then $A^*A\vec{x} = 0_V$. So, it holds that

$$\langle A^*A\vec{x}, \vec{x}\rangle = \langle 0, \vec{x}\rangle = 0 .$$

Then

$$0 = \langle A^*A\vec{x}, \vec{x}\rangle = \langle A\vec{x}, (A^*)^*\vec{x}\rangle = \langle A\vec{x}, A\vec{x}\rangle .$$

Now, from the property of non-degeneracy of an inner product, it must be that $A\vec{x} = 0$ for the above equality to hold. Thus, $A\vec{x} = 0$, and so $\vec{x} \in \ker A$.

(ii) By rank-nullity, we have that

$$\begin{aligned}
\operatorname{rank} A^*A &= \dim V - \operatorname{nullity}(A^*A) \\
&= \dim V - \dim(\ker A^*A) \\
&= \dim V - \dim(\ker A) && \text{(by (i))} \\
&= \dim V - \operatorname{nullity} A \\
&= \operatorname{rank} A\ .
\end{aligned}$$

(iii) Suppose $\ker A = \{0\}$. Then A is one-to-one. We know from property (i) that $\ker A = \ker A^*A$. Then

$$\{0\} = \ker A = \ker A^*A\ ,$$

and so $\ker A^*A$ is one-to-one. From property (ii), we know that $\operatorname{rank} A = \operatorname{rank} A^*A$. Since $\ker A = \ker A^*A = \{0\}$, we get from the definition of nullity that $\operatorname{nullity}(A^*A) = 0$. Then by rank-nullity,

$$\begin{aligned}
\operatorname{rank} A^*A &= \dim V - \operatorname{nullity}(A^*A) \\
&= \dim V - 0 \\
&= \dim V\ .
\end{aligned}$$

Note that A^*A is a composition of maps given by $A^*A : V \to V$. Now, $\operatorname{rank} A^*A = \dim V$ means that A is full rank, and so A^*A is invertible. That is, $(A^*A)^{-1}$ exists. So,

$$(A^*A)^{-1}A^*A = I\ .$$

In other words, $(A^*A)^{-1}$ is a left inverse for A.

(iv) Suppose $\operatorname{ran} A = W$. Then A is onto. Let $\dim W = m$. From the definition of rank, we know that

$$\operatorname{rank} A = \dim(\operatorname{Im} A) = \dim(\operatorname{ran} A) = \dim W = m\ .$$

It might be worth pointing out that that the image of A is the same as the column space of A, and so $\dim(\operatorname{Im} A) = \dim(\operatorname{Col} A)$. Now, by rank-nullity, noting that $A : V \to W$,

$$\begin{aligned}
\operatorname{rank} A &= \dim V - \operatorname{nullity} A \\
\operatorname{nullity} A &= \dim V - \operatorname{rank} A \\
\operatorname{nullity} A &= \dim V - \dim W\ .
\end{aligned}$$

Then since $\ker A = \ker A^*A$,

$$\operatorname{nullity} A = \dim(\ker A) = \dim(\ker A^*A) = \operatorname{nullity} A^*A\ .$$

□

Isometries

Definition (Isometry)

A linear map $U : V \to W$, where V and W are inner product spaces, is called an **isometry** (or isometric) if

$$\underbrace{||U\vec{v}||}_{\text{norm on } W} = \underbrace{||\vec{v}||}_{\text{norm on } V}$$

for all $\vec{v} \in V$. This can be thought of as "distance preserving".

1. Let A be the linear map given by

$$\begin{bmatrix} 0 & \frac{1}{\sqrt{2}} \\ 1 & 0 \\ 0 & -\frac{1}{\sqrt{2}} \\ 0 & 0 \end{bmatrix} : \mathbb{C}^2 \to \mathbb{C}^4 .$$

This linear map is an isometry. Indeed, if we take $\vec{v} = \begin{bmatrix} z \\ w \end{bmatrix} \in \mathbb{C}^2$, then

$$\begin{aligned}
\left|\left| \begin{bmatrix} 0 & \frac{1}{\sqrt{2}} \\ 1 & 0 \\ 0 & -\frac{1}{\sqrt{2}} \\ 0 & 0 \end{bmatrix} \vec{v} \right|\right|^2 &= \left|\left| \begin{bmatrix} 0 & \frac{1}{\sqrt{2}} \\ 1 & 0 \\ 0 & -\frac{1}{\sqrt{2}} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} z \\ w \end{bmatrix} \right|\right|^2 \\
&= \left|\left| \begin{bmatrix} w/\sqrt{2} \\ z \\ -w/\sqrt{2} \\ 0 \end{bmatrix} \right|\right|^2 \\
&= \langle \begin{bmatrix} w/\sqrt{2} \\ z \\ -w/\sqrt{2} \\ 0 \end{bmatrix}, \begin{bmatrix} w/\sqrt{2} \\ z \\ -w/\sqrt{2} \\ 0 \end{bmatrix} \rangle \\
&= \frac{w}{\sqrt{2}} \cdot \overline{\left(\frac{w}{\sqrt{2}}\right)} + z \cdot \overline{z} + \left(\frac{-w}{\sqrt{2}}\right) \overline{\left(\frac{-w}{\sqrt{2}}\right)} + 0 \cdot 0 \\
&= \frac{w}{\sqrt{2}} \cdot \frac{\overline{w}}{\overline{\sqrt{2}}} + |z|^2 + \left(\frac{-w}{\sqrt{2}}\right) \left(\frac{\overline{-w}}{\overline{\sqrt{2}}}\right) \\
&= \frac{w}{\sqrt{2}} \cdot \frac{\overline{w}}{\sqrt{2}} + |z|^2 + \left(\frac{-w}{\sqrt{2}}\right) \left(\frac{\overline{-w}}{\sqrt{2}}\right) \\
&= \frac{w\overline{w}}{2} + |z|^2 + \left(\frac{-w\overline{(-w)}}{2}\right)
\end{aligned}$$

$$
\begin{aligned}
&= \frac{|w|^2}{2} + |z|^2 + \frac{|-w|^2}{2} \\
&= \frac{|w|^2}{2} + |z|^2 + \frac{|w|^2}{2} \\
&= |w|^2 + |z|^2 \\
&= |z|^2 + |w|^2 \\
&= z\overline{z} + w\overline{w} \\
&= \langle \begin{bmatrix} z \\ w \end{bmatrix}, \begin{bmatrix} z \\ w \end{bmatrix} \rangle \\
&= \left\| \begin{bmatrix} z \\ w \end{bmatrix} \right\|^2 ,
\end{aligned}
$$

where we used the fact that $|w|^2 = |-w|^2$. Indeed, this can be checked. If $w = a + bi$, then $-w = -(a + bi) = -a - bi$, and so

$$
\begin{aligned}
|-w|^2 &= -w \cdot \overline{(-w)} \\
&= (-a - bi)(-a + bi) \\
&= a^2 - abi + abi - b^2 i^2 \\
&= a^2 - b^2(-1) \\
&= a^2 + b^2 \\
&= a^2 - b^2(-1) \\
&= a^2 - abi + abi - b^2 i^2 \\
&= (a + bi)(a - bi) \\
&= w \cdot \overline{w} \\
&= |w|^2 .
\end{aligned}
$$

Proposition

$U : V \to W$ is an isometry iff $\langle U\vec{v}, U\vec{y}\rangle = \langle \vec{v}, \vec{y}\rangle$ for all $\vec{v}, \vec{y} \in V$.

"Isometries are inner product preserving"

Proof. Let $U : V \to W$ be a linear map.

($\Longrightarrow$) Suppose U is an isometry. Then for any $\vec{v} \in V$, $||U\vec{v}|| = ||\vec{v}||$. From this, it also holds that $||U\vec{v}||^2 = ||\vec{v}||^2$; that is, the squares of their norms are also equal. Then by the definition of norm,

$$\begin{aligned}\langle U\vec{v}, U\vec{v}\rangle &= ||U\vec{v}||^2\\ &= ||\vec{v}||^2\\ &= \langle \vec{v}, \vec{v}\rangle\end{aligned}$$

($\Longleftarrow$) Suppose $\langle U\vec{v}, U\vec{y}\rangle = \langle \vec{v}, \vec{y}\rangle$ for $\vec{v}, \vec{y} \in V$. From the polarization identity, we get that for any two vectors $\vec{v}, \vec{y} \in V$,

$$\begin{aligned}&\langle \vec{v}, \vec{y}\rangle\\ &= \frac{1}{4}\left(||\vec{v}+\vec{y}||^2 - ||\vec{v}-\vec{y}||^2 + i||\vec{v}+i\vec{y}||^2 - i||\vec{v}-i\vec{y}||^2\right)\\ &= \frac{1}{4}\left(\langle \vec{v}+\vec{y}, \vec{v}+\vec{y}\rangle - \langle \vec{v}-\vec{y}, \vec{v}-\vec{y}\rangle + i\langle \vec{v}+i\vec{y}, \vec{v}+i\vec{y}\rangle - i\langle \vec{v}-i\vec{y}, \vec{v}-i\vec{y}\rangle\right)\\ &= \frac{1}{4}\left(\langle U(\vec{v}+\vec{y}), U(\vec{v}+\vec{y})\rangle - \langle U(\vec{v}-\vec{y}), U(\vec{v}-\vec{y})\rangle + i\langle U(\vec{v}+i\vec{y}), U(\vec{v}+i\vec{y})\rangle - i\langle U(\vec{v}-i\vec{y}), U(\vec{v}-i\vec{y})\rangle\right)\\ &= \frac{1}{4}\left(||U(\vec{v}+\vec{y})||^2 - ||U(\vec{v}-\vec{y})||^2 + i||U(\vec{v}+i\vec{y})||^2 - i||U(\vec{v}-i\vec{y})||^2\right)\\ &= \frac{1}{4}\left(||U\vec{v}+U\vec{y}||^2 - ||U\vec{v}-U\vec{y}||^2 + i||U\vec{v}+iU\vec{y}||^2 - i||U\vec{v}-iU\vec{y}||^2\right)\\ &= \langle U\vec{v}, U\vec{y}\rangle\ .\end{aligned}$$

□

Corollary

$U : V \to W$ is an isometry iff $U^*U = I$.

Proof. Let $U : V \to W$ be a linear operator.

($\Longrightarrow$) Suppose U is an isometry. Then for all $\vec{x}, \vec{y} \in V$, $\langle U\vec{x}, U\vec{y}\rangle = \langle \vec{x}, \vec{y}\rangle$. So,

$$\langle \vec{x}, \vec{y}\rangle = \langle U\vec{x}, U\vec{y}\rangle = \langle U^*U\vec{x}, \vec{y}\rangle \ .$$

Then ...

($\Longleftarrow$) Suppose $U^*U = I$. Then for all $\vec{x}, \vec{y} \in V$,

$$\begin{aligned} \langle U\vec{x}, U\vec{y}\rangle &= \langle \vec{x}, U^*U\vec{y}\rangle \\ &= \langle \vec{x}, I\vec{y}\rangle \\ &= \langle \vec{x}, \vec{y}\rangle \ . \end{aligned}$$

Thus, by the previous proposition, U is an isometry. $\square$

Unitaries

Definition (Unitary)

An **isometry** U is called **unitary** if U is invertible.

Remarks:

- If U is invertible, we must have $\dim V = \dim W$. So, a unitary is an **isometric isomorphism** of inner product spaces.
- Isometries have $\{0\}$ kernel since $0 = ||U\vec{x}|| = ||\vec{x}||$ iff $\vec{x} = 0$ (also because invertibility implies one-to-one and onto). By rank-nullity, this means $\dim(\operatorname{ran} U) = \operatorname{rank} U = \dim V$. That is,

$$\begin{aligned}
\dim V &= \dim(\operatorname{ran} U) + \operatorname{nullity} U \\
&= \operatorname{rank} U + \dim(\ker U) \\
&= \operatorname{rank} U + \dim\{0\} \\
&= \operatorname{rank} U + 0 \\
&= \operatorname{rank} U \ .
\end{aligned}$$

  So, U is unitary iff $\dim W = \dim V = \operatorname{rank} U = \dim U$.

1. (Permutation Matrix)

   A **permutation** on the set $\{1,\dots,n\}$ is simply a bijection $\Pi : \{1,\dots,n\} \to \{1,\dots,n\}$. For a given Π, define an $n \times n$ matrix by

$$U_\Pi := \underbrace{\begin{bmatrix} \vec{e}_{\Pi(1)} & \vec{e}_{\Pi(2)} & \dots & \vec{e}_{\Pi(n)} \end{bmatrix}}_{\text{permuting the basis}} ,$$

   where $\{\vec{e}_1,\dots,\vec{e}_n\}$ is the c.o.b of $\mathbb{C}^n$. For example, if

$$\begin{aligned} \Pi(1) &= 2 \\ \Pi(2) &= 3 \\ \Pi(3) &= 1 \end{aligned}$$

   then the permutation matrix is

$$U_\Pi = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

   Notice that U_Π is invertible. In fact, $(U_\Pi)^{-1} = U_{\Pi^{-1}}$. U is an isometry since (using the previous corollary),

$$\begin{aligned} U_\Pi^* U_\Pi &= \begin{bmatrix} \overline{\vec{e}_{\Pi(1)}}^T \\ \overline{\vec{e}_{\Pi(2)}}^T \\ \vdots \\ \overline{\vec{e}_{\Pi(3)}}^T \end{bmatrix} \begin{bmatrix} \vec{e}_{\Pi(1)} & \dots & \vec{e}_{\Pi(n)} \end{bmatrix} \\ &= \left[\langle \vec{e}_{\Pi(j)}, \vec{e}_{\Pi(i)} \rangle \right]_{i,j=1}^n \\ &= \dots \\ &= \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix} . \end{aligned}$$

   If Π_1 and Π_2 are permutations, then

$$U_{\Pi_1 \circ \Pi_2} = U_{\Pi_1} U_{\Pi_2} .$$

   (These matrices are a faithful unitary representation of the permutation group S_n acting on $\mathbb{C}^n$). For any permutation Π, there exists k such that $\Pi^k = \mathrm{id}_{\{1,\dots,n\}}$. So,

$$(U_\Pi)^k = U_{\Pi^k} = U_{\mathrm{id}} = I ,$$

   which means that the U_Π are unipotent.

Proposition (Properties of Unitaries)

Let $U : V \to W$ be a linear map.

(i) U is unitary iff $U^* = U^{-1}$ iff $U^*U = UU^* = I$. (The second "iff" comes from the corollary for isometries.)

(ii) If U is an isometry and $\{\vec{x}_1, \ldots, \vec{x}_m\}$ is orthogonal/orthonormal in V, then $\{U\vec{x}_1, \ldots, U\vec{x}_m\}$ is orthogonal/orthonormal in W.

(iii) The product of isometries is an isometry.

(iv) The columns (of a matrix representation) of an isometric form an orthonormal set.

(v) The columns and rows of a unitary are orthonormal bases of W and V, respectively.

(vi) If U is unitary, then $|\det U| = 1$.

(vii) If λ is an eigenvalue of a unitary, then $|\lambda| = 1$.

Proof. Suppose $U : V \to W$ is a linear map. We prove the above properties.

(i) ($\Longrightarrow$) Suppose U is unitary. Then U is invertible (i.e. U^{-1} exists). Also, since U is unitary, it must be an isometry. So, by the corollary for isometries, it follows that $U^*U = I$. We know that by the property of the inverse of matrix that $U^{-1}U = I$. So,

$$\begin{aligned} U^*U &= U^{-1}U \\ U^*UU^{-1} &= U^{-1}UU^{-1} \\ U^*I &= U^{-1}I \\ U^* &= U^{-1} \ . \end{aligned}$$

Now, since $U^* = U^{-1}$ and $U^*U = I$, we get from the property of the inverse of a matrix

$$UU^* = UU^{-1} = I = U^*U \ .$$

($\Longleftarrow$) Suppose $U^*U = UU^* = I$. From this, we get that

$$U^*U = UU^*$$

$\square$

1. The converse does not hold for properties (vi) and (vii). For example, take $U = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. Here, the columns do not form an orthonormal basis. More specifically,

$$\langle \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \rangle = 1 \cdot \bar{1} + 0 \cdot \bar{1} = 1 \cdot 1 + 0 = 1 \neq 0 \ ,$$

meaning columns vectors are not orthogonal and hence definitely not orthonormal. Also,

$$|\det U| = |(1 \cdot 1) - (1 \cdot 0)| = |1 - 0| = |1| = 1 \ ,$$

but U is not unitary (again from the fact the its columns do not form an orthonormal basis). Also,

$$\begin{aligned} c_U(z) &= \det(zI_2 - U) \\ &= \det\left(\begin{bmatrix} z-1 & -1 \\ 0 & z-1 \end{bmatrix}\right) \\ &= (z-1)^2 \ , \end{aligned}$$

which means that eigenvalues of U are $\lambda_1 = \lambda_2 = 1$. So, although $|\lambda_1| = |1| = 1$, U is not unitary.

2. Compute the eigenvalues for

$$U = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} .$$

We have that

$$\begin{aligned}
c_U(z) &= \det(zI_4 - U) \\
&= \det\left(\begin{bmatrix} z & 0 & 0 & 0 \\ 0 & z & 0 & 0 \\ 0 & 0 & z & 0 \\ 0 & 0 & 0 & z \end{bmatrix} - \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \right) \\
&= \det\left(\begin{bmatrix} z & -1 & 0 & 0 \\ 0 & z & 0 & -1 \\ -1 & 0 & z & 0 \\ 0 & 0 & -1 & z \end{bmatrix} \right) \\
&= z \cdot \det\left(\begin{bmatrix} z & 0 & -1 \\ 0 & z & 0 \\ 0 & -1 & z \end{bmatrix} \right) + \det\left(\begin{bmatrix} 0 & 0 & -1 \\ -1 & z & 0 \\ 0 & -1 & z \end{bmatrix} \right) \\
&= z\left(z(z^2 - 0)\right) + (-1(1-0)) \\
&= z\left(z^3\right) + (-1) \\
&= z^4 - 1 \\
&= (z^2 - 1)(z^2 + 1) \\
&= (z-1)(z+1)(z-i)(z+i) .
\end{aligned}$$

So, the eigenvalues of U are $\lambda_1 = 1$, $\lambda_2 = -1$, $\lambda_3 = i$, and $\lambda_4 = -i$. Then

$$\begin{aligned}
|\lambda_1| &= |1| = 1 \checkmark \\
|\lambda_2| &= |-1| = 1 \checkmark \\
|\lambda_3| &= |i| = |0 + 1i| = \sqrt{0^2 + 1^2} = \sqrt{1} = 1 \checkmark \\
|\lambda_4| &= |-i| = |0 + (-1)i| = \sqrt{0^2 + (-1)^2} = \sqrt{1} = 1 \checkmark ,
\end{aligned}$$

which indeed satisfies property (vii). Furthermore,

$$\begin{aligned}
\det U &= \det\left(\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \right) \\
&= -\det\left(\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \right)
\end{aligned}$$

$$
\begin{aligned}
&= \det\left(\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & 1 & 0\end{bmatrix}\right)\\
&= -\det\left(\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}\right)\\
&= -(1\cdot 1\cdot 1\cdot 1)\\
&= -1\ .
\end{aligned}
$$

So $|\det U| = |-1| = 1$, which satisfies property (vi).

Definition (Unitarily Equivalent & Unitarily Diagonalizable)

Let $A : V \to V$ and $B : V \to V$ be linear maps. A and B are **unitarily** equivalent if there exists a unitary U such that

$$U^* A U = B \ .$$

We write A and B being **unitarily equivalent** as $A \sim_U B$. Note that since $U^* = U^{-1}$ (from property (i) of unitaries), this implies that A is similar to B (i.e. $U^{-1} A U = B$).

A is **unitarily diagonalizable** if $A \sim_U D$ for some diagonal D. That is,

$$U^* A U = D \ .$$

Proposition

A is unitarily diagonalizble iff A has an orthonormal basis of eigenvectors.

Proof. ($\Longleftarrow$) Suppose $\{\vec{u}_1, \ldots, \vec{u}_n\}$ is an orthonormal basis for A and define

$$U = \begin{bmatrix} \vec{u}_1 & \ldots & \vec{u}_n \end{bmatrix} \ .$$

Since U has orthonormal columns, it is an isometry. It's columns form a basis, so it is also invertible and hence unitary. Then

$$\begin{aligned} AU &= \begin{bmatrix} A\vec{u}_1 & \ldots & A\vec{u}_n \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 \vec{u}_1 & \ldots & \lambda_n \vec{u}_n \end{bmatrix} && (\star) \\ &= \begin{bmatrix} \vec{u}_1 & \ldots & \vec{u}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_n \end{bmatrix} \\ &= UD \ . \end{aligned}$$

where in step $(\star)$, we used the fact the $\vec{u}_j$ vectors are eigenvectors. Since $AU = UD$ and U is unitary, we get that

$$\begin{aligned} AU &= UD \\ U^* A U &= U^* U D \\ U^* A U &= ID && \text{(prop. (i))} \\ U^* A U &= D \ . \end{aligned}$$

Thus, $A \sim_U D$; that is, A is unitarily equivalent to D.

($\Longrightarrow$) Suppose A is unitarily diagonalizable. Then there exists a unitary U such that $U^*AU = D$. Then

$$\begin{aligned} U^*AU &= D \\ UU^*AU &= UD \\ IAU &= UD && \text{(prop. (i))} \\ AU &= UD\ . \end{aligned}$$

Now, define $U = \begin{bmatrix} \vec{u}_1 & \dots & \vec{u}_n \end{bmatrix}$.

$$\begin{aligned} AU &= UD \\ A\begin{bmatrix} \vec{u}_1 & \dots & \vec{u}_n \end{bmatrix} &= \begin{bmatrix} \vec{u}_1 & \dots & \vec{u}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{bmatrix} \\ \begin{bmatrix} A\vec{u}_1 & \dots & A\vec{u}_n \end{bmatrix} &= \begin{bmatrix} \lambda_1\vec{u}_1 & \dots & \lambda_n\vec{u}_n \end{bmatrix}\ . \end{aligned}$$

Comparing the entries, we have that $A\vec{u}_j = \lambda_j\vec{u}_j$ for $1 \leq j \leq n$, which means that the $\vec{u}_j$ are eigenvectors. So, we get that U has eigenvectors as its columns. Now, because U is a unitary, we get from property (v) of unitaries that the the column vectors of U form an orthonormal bases (for the codomain of U). $\square$

1. Unitarily diagonalize $A = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ -1 & 2 & 5 \end{bmatrix}$.

We have that

$$\begin{aligned} c_A(z) &= \dots \\ &= z(z-1)(z-6) \ . \end{aligned}$$

So, A has eigenvalues $\lambda_1 = 0$, $\lambda_2 = 1$, and $\lambda_3 = 6$. Now, we obtain the eigenvectors associated with each eigenvalue. For $\lambda_1 = 0$, we solve $(A - 0I)\vec{x} = 0$ for $\vec{x}$.

$$\left[\begin{array}{rrr|r} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \\ -1 & 2 & 5 & 0 \end{array}\right]$$

$$\xrightarrow{1R1+R3} \left[\begin{array}{rrr|r} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 2 & 4 & 0 \end{array}\right]$$

$$\xrightarrow{-2R2+R3} \left[\begin{array}{rrr|r} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

So, the solution to this homogeneous system is

$$\vec{x} = \begin{bmatrix} t \\ -2t \\ t \end{bmatrix} = t \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \ , \ t \in \mathbb{F}$$

We need a unit vector, so choose $\vec{u}_1 = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$. For $\lambda_2 = 1$, we solve $(A - 1I)\vec{x} = 0$ for $\vec{x}$.

$$\left[\begin{array}{rrr|r} 0 & 0 & -1 & 0 \\ 0 & 0 & 2 & 0 \\ -1 & 2 & 4 & 0 \end{array}\right]$$

$$\xrightarrow{R1 \longleftrightarrow R3} \left[\begin{array}{rrr|r} -1 & 2 & 4 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & -1 & 0 \end{array}\right]$$

$$\xrightarrow[-1R3]{-1R1} \left[\begin{array}{rrr|r} 1 & -2 & -4 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right]$$

$$\xrightarrow[-2R3+R2]{4R3+R1}\left[\begin{array}{rrr|r} 1 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right]$$

So, the solution to this homogeneous system is

$$\vec{x} = \begin{bmatrix} 2t \\ t \\ 0 \end{bmatrix} = t \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}.$$

Choose $\vec{u}_2 = \dfrac{1}{\sqrt{5}} \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$. Note that $\vec{u}_1$ and $\vec{u}_2$ are already orthogonal! Finally, for $\lambda_3 = 6$, we solve $(A - 6I)\vec{x} = 0$ for $\vec{x}$.

$$\left[\begin{array}{rrr|r} -5 & 0 & -1 & 0 \\ 0 & -5 & 2 & 0 \\ -1 & 2 & -1 & 0 \end{array}\right]$$

$$\xrightarrow{R1 \longleftrightarrow R3}\left[\begin{array}{rrr|r} -1 & 2 & -1 & 0 \\ 0 & -5 & 2 & 0 \\ -5 & 0 & -1 & 0 \end{array}\right]$$

$$\xrightarrow{-1R1}\left[\begin{array}{rrr|r} 1 & -2 & 1 & 0 \\ 0 & -5 & 2 & 0 \\ -5 & 0 & -1 & 0 \end{array}\right]$$

$$\xrightarrow{5R1+R3}\left[\begin{array}{rrr|r} 1 & -2 & 1 & 0 \\ 0 & -5 & 2 & 0 \\ 0 & -10 & 4 & 0 \end{array}\right]$$

$$\xrightarrow{-2R2+R3}\left[\begin{array}{rrr|r} 1 & -2 & 1 & 0 \\ 0 & -5 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$$\xrightarrow{-\frac{1}{5}R2}\left[\begin{array}{rrr|r} 1 & -2 & 1 & 0 \\ 0 & 1 & -2/5 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$$\xrightarrow{2R2+R1}\left[\begin{array}{rrr|r} 1 & 0 & 1/5 & 0 \\ 0 & 1 & -2/5 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

So, the solution to this homogeneous system is

$$\vec{x} = \begin{bmatrix} -t/5 \\ 2t/5 \\ t \end{bmatrix} = t \begin{bmatrix} -1/5 \\ 2/5 \\ 1 \end{bmatrix}.$$

Choose $\vec{u}_3 = \frac{1}{\sqrt{30}} \begin{bmatrix} -1 \\ 2 \\ 5 \end{bmatrix}$. Thus,

$$D = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 6 \end{bmatrix}$$

and

$$U = \begin{bmatrix} \vec{u}_1 & \vec{u}_2 & \vec{u}_3 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{6} & 2/\sqrt{5} & -1/\sqrt{30} \\ -2/\sqrt{6} & 1/\sqrt{5} & 2/\sqrt{30} \\ 1/\sqrt{6} & 0 & 5/\sqrt{30} \end{bmatrix}.$$

Notice that for **distinct eigenvalues**, the associated eigenvectors were already orthogonal with each other!

If U is unitary with $U\vec{x} = \lambda\vec{x}$ and $U\vec{y} = \mu\vec{y}$, where $\lambda \neq \mu$ and $\vec{x}, \vec{y}$ non-zero, then why is $\langle \vec{x}, \vec{y} \rangle = 0$? (In other words, why are $\vec{x}$ and $\vec{y}$ orthogonal?)

2. Suppose $A \sim_U B$. So, there exists U such that $U^*AU = B$.

   (a) Show $\text{tr}(A^*A) = \text{tr}(B^*B)$.

   We have that

$$\begin{aligned} U^*AU &= B \\ UU^*AU &= UB \\ IAU &= UB \\ AU &= UB \\ AUU^* &= UBU^* \\ AI &= UBU^* \\ A &= UBU^* \ . \end{aligned}$$

   Then it follows that

$$\begin{aligned} A^*A &= (UBU^*)^*(UBU^*) \\ &= U^{**}B^*U^*UBU^* \\ &= UB^*IBU^* \\ &= UB^*BU^* \\ &= UIU^* \\ &= UU^* \\ &= I \end{aligned}$$

   and

$$\begin{aligned} B^*B &= (U^*AU)^*(U^*AU) \\ &= U^*A^*U^{**}U^*AU \\ &= U^*A^*UU^*AU \\ &= U^*A^*IAU \\ &= U^*A^*AU \\ &= U^*IU \\ &= U^*U \\ &= I \ . \end{aligned}$$

   Thus,

$$\text{tr}(A^*A) = \text{tr}(I) = \text{tr}(B^*B) \ .$$

(b) Is $\mathrm{tr}(A) = \mathrm{tr}(B)$?

Yes. Since $B = U^*AU$, we get that

$$\begin{aligned}
\mathrm{tr}(B) &= \mathrm{tr}(U^*AU) \\
&= \mathrm{tr}(U^*(AU)) \\
&= \mathrm{tr}((AU)U^*) \\
&= \mathrm{tr}(AUU^*) \\
&= \mathrm{tr}(AI) \\
&= \mathrm{tr}(A)\ ,
\end{aligned}$$

where we used that fact that $\mathrm{tr}(CD) = \mathrm{tr}(DC)$ for any square matrices C and D.

(c) Are $\begin{bmatrix} 1 & 2 \\ 2 & i \end{bmatrix}$ and $\begin{bmatrix} i & 4 \\ 1 & 1 \end{bmatrix}$ unitarily equivalent?

We have that

$$\operatorname{tr}\left(\begin{bmatrix} 1 & 2 \\ 2 & i \end{bmatrix}\right) = 1 + i = i + 1 = \operatorname{tr}\left(\begin{bmatrix} i & 4 \\ 1 & 1 \end{bmatrix}\right) .$$

However, using part (a),

$$\begin{aligned}
\operatorname{tr}\left(\begin{bmatrix} 1 & 2 \\ 2 & i \end{bmatrix}^* \begin{bmatrix} 1 & 2 \\ 2 & i \end{bmatrix}\right) &= \operatorname{tr}\left(\begin{bmatrix} \overline{1} & \overline{2} \\ \overline{2} & \overline{i} \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & i \end{bmatrix}\right) \\
&= \operatorname{tr}\left(\begin{bmatrix} 1 \cdot \overline{1} + 2 \cdot \overline{2} & 1 \cdot \overline{2} + 2\overline{i} \\ \overline{2} \cdot 1 + \overline{i} \cdot 2 & \overline{2} \cdot 2 + \overline{i}i \end{bmatrix}\right) \\
&= \operatorname{tr}\left(\begin{bmatrix} |1|^2 + |2|^2 & 1 \cdot 2 + 2(-i) \\ 2 \cdot 1 - 2i & |2|^2 + |i|^2 \end{bmatrix}\right) \\
&= |1|^2 + |2|^2 + |2|^2 + |i|^2 \\
&= 1 + 4 + 4 + 1 \\
&= 10 \\
&\neq 19 \\
&= 1 + 1 + 16 + 1 \\
&= |i|^2 + |1|^2 + |4|^2 + |1|^2 \\
&= \operatorname{tr}\left(\begin{bmatrix} |i|^2 + |1|^2 & 4\overline{i} + 1 \cdot \overline{1} \\ \overline{4}i + 1 \cdot \overline{1} & |4|^2 + |1|^2 \end{bmatrix}\right) \\
&= \operatorname{tr}\left(\begin{bmatrix} i\overline{i} + 1 \cdot \overline{1} & 4\overline{i} + 1 \cdot \overline{1} \\ \overline{4}i + 1 \cdot \overline{1} & 4 \cdot \overline{4} + 1 \cdot \overline{1} \end{bmatrix}\right) \\
&= \operatorname{tr}\left(\begin{bmatrix} \overline{i} & \overline{1} \\ \overline{4} & \overline{1} \end{bmatrix} \begin{bmatrix} i & 4 \\ 1 & 1 \end{bmatrix}\right) \\
&= \operatorname{tr}\left(\begin{bmatrix} i & 4 \\ 1 & 1 \end{bmatrix}^* \begin{bmatrix} i & 4 \\ 1 & 1 \end{bmatrix}\right) .
\end{aligned}$$

Thus, these matrices are not unitarily equivalent.

3. Suppose $\ker A = \{0\}$ and $A : V \to W$. Let $E = \operatorname{ran} A$. Show that $P_E = A(A^*A)^{-1}A^*$.

From an earlier proposition, we saw that $\ker A = \ker A^*A$. So, $\ker A^*A = \ker A = \{0\}$, which means that A^*A is one-to-one and has a right inverse (this comes from the properties of adjoints). Note that $A^* : W \to V$, and so $A^*A : V \to V$. So, applying the rank-nullity theorem tells us that

$$\begin{aligned}
\dim V &= \operatorname{rank} A^*A - \operatorname{nullity} A^*A \\
&= \operatorname{rank} A^*A - \dim(\ker A^*A) \\
&= \operatorname{rank} A^*A - \dim(\{0\}) \\
&= \operatorname{rank} A^*A - 0 \\
&= \operatorname{rank} A^*A \ ,
\end{aligned}$$

where $\dim(\{0\}) = 0$ comes from the definition of dimension. Here, we get that $\operatorname{rank} A^*A = \dim V$ (i.e. the rank of A^*A is the same as the dimension of the **codomain**), which means that A^*A is onto. So, because A^*A is one-to-one and onto, it has a left and right inverse. In other words, A^*A is invertible (i.e. $(A^*A)^{-1}$ exists). Now,

$$\begin{aligned}
(A(A^*A)^{-1}A^*)^* &= A^{**}[(A^*A)^{-1}]^*A^* \\
&= A[(A^*A)^*]^{-1}A^* && ((B^{-1})^* = (B^*)^{-1}) \\
&= A(A^*A^{**})^{-1}A^* \\
&= A(A^*A)^{-1}A^*
\end{aligned}$$

and

$$\begin{aligned}
(A(A^*A)^{-1}A^*)^2 &= (A(A^*A)^{-1}A^*)(A(A^*A)^{-1}A^*) \\
&= A(A^*A)^{-1}A^*A(A^*A)^{-1}A^* \\
&= A(I)^{-1}I(A^*A)^{-1}A^* \\
&= AI(A^*A)^{-1}A^* \\
&= A(A^*A)^{-1}A^* \ .
\end{aligned}$$

So, it is a projection.

4. Define $T : M_{2\times 2}(\mathbb{C}) \to \mathbb{P}_2[t]$ by

$$T\begin{bmatrix} a & b \\ c & d \end{bmatrix} = (a+b)t^2 + ct + (a+d)$$

Here, the inner product for $M_{2\times 2}$ is defined as the usual $\langle A, B\rangle = \operatorname{tr}(AB^*)$, and the inner product of $P_2[t]$ is such that $\{1, t, t^2\}$ is an orthonormal basis. Find $\ker T$, $\operatorname{ran} T$, $\ker T^*$, $\operatorname{ran} T^*$, and orthogonal bases for each of these results.

The kernel of T is

$$\begin{aligned}
\ker T &= \{B \in M_{2\times 2}(\mathbb{C}) : T(B) = 0\} \\
&= \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} : (a+b)t^2 + ct + (a+d) = 0 \right\} \\
&= \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} : a+b = 0, c = 0, a+d = 0 \right\} \\
&= \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} : b = -a, c = 0, d = -a \right\} \\
&= \left\{ \begin{bmatrix} a & -a \\ 0 & -a \end{bmatrix} : a \in \mathbb{C} \right\} \\
&= \left\{ a\begin{bmatrix} 1 & -1 \\ 0 & -1 \end{bmatrix} : a \in \mathbb{C} \right\} \\
&= \operatorname{span}\left\{ \begin{bmatrix} 1 & -1 \\ 0 & -1 \end{bmatrix} \right\} \ .
\end{aligned}$$

The range of T is

$$\begin{aligned}
\operatorname{ran} T &= \{T(B) \in \mathbb{P}_2[t] : B \in M_{2\times 2}(\mathbb{C})\} \\
&= \left\{ T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) \in \mathbb{P}_2[t] : \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_{2\times 2}(\mathbb{C}) \right\} \\
&= \{(a+b)^2 + ct + (a+d)\} \ .
\end{aligned}$$

Since

$$\begin{aligned}
\dim(\operatorname{ran} T) &= \operatorname{rank} T \\
&= \dim(M_{2\times 2}(\mathbb{C})) - \operatorname{nullity} T \\
&= \dim(M_{2\times 2}(\mathbb{C})) - \dim(\ker T) \\
&= 4 - 1 \\
&= 3 \\
&= \dim(\mathbb{P}_2[t]) \ ,
\end{aligned}$$

we get that $\operatorname{ran} T = \mathbb{P}_2[t]$, which again has an orthonormal basis $\{1, t, t^2\}$. Now, we know that $\operatorname{ran} T = \operatorname{Im} T$ is a subspace of the codomain $\mathbb{P}_2[t]$. So, noting that $\operatorname{ran} T = \mathbb{P}_2[t]$,

$$\begin{aligned}
\ker T^* &= (\operatorname{ran} T)^{\perp} \\
&= \{\vec{v} \in \mathbb{P}_2[t] : \langle \vec{u}, \vec{v} \rangle = 0 \ \text{ for all } \ \vec{u} \in \operatorname{ran} T\} \\
&= \{\vec{v} \in \mathbb{P}_2[t] : \langle \vec{u}, \vec{v} \rangle = 0 \ \text{ for all } \ \vec{u} \in \mathbb{P}_2[t]\} \\
&= \{0\} \ .
\end{aligned}$$

In other words, only the zero vector (aka the zero polynomial) is orthogonal to every vector (polynomial) in $\mathbb{P}_2[t]$. Finally, since $\ker T$ is a subspace of the domain $M_{2\times 2}(\mathbb{C})$,

$$\begin{aligned}
\operatorname{ran} T^* &= (\ker T)^{\perp} \\
&= \{C \in M_{2\times 2}(\mathbb{C}) : \langle D, C \rangle = 0 \ \text{ for all } D \in \ker T\} \\
&= \{C \in M_{2\times 2}(\mathbb{C}) : \operatorname{tr}(DC^*) = 0 \ \text{ for all } D \in \ker T\} \ .
\end{aligned}$$

Here,

$$\begin{aligned}
\operatorname{tr}(DC^*) &= \operatorname{tr}\left(\begin{bmatrix} a & -a \\ 0 & -a \end{bmatrix} \begin{bmatrix} x & y \\ w & z \end{bmatrix}^* \right) \\
&= \operatorname{tr}\left(\begin{bmatrix} a & -a \\ 0 & -a \end{bmatrix} \begin{bmatrix} \overline{x} & \overline{w} \\ \overline{y} & \overline{z} \end{bmatrix} \right) \\
&= \operatorname{tr}\left(\begin{bmatrix} a\overline{x} - a\overline{y} & a\overline{w} - a\overline{z} \\ -a\overline{y} & -a\overline{z} \end{bmatrix} \right) \\
&= a\overline{x} - a\overline{y} - a\overline{z} \\
&= a\left(\overline{x} - \overline{y} - \overline{z}\right) \ .
\end{aligned}$$

So,

$$\begin{aligned}
\operatorname{ran} T^* &= \{C \in M_{2\times 2}(\mathbb{C}) : \operatorname{tr}(DC^*) = 0 \ \text{ for all } D \in \ker T\} \\
&= \left\{ \begin{bmatrix} x & y \\ w & z \end{bmatrix} : a(\overline{x} - \overline{y} - \overline{z}) = 0 \right\} \\
&=
\end{aligned}$$

Structure of Operators

Definition

An operator $A : V \to V$, where V is an inner product space, is **upper triangular** with respect to an orthonormal basis $\{\vec{u}_1, \ldots, \vec{u}_n\}$ if

$$\begin{aligned} A\vec{u}_1 &\in \operatorname{span}\{\vec{u}_1\} \\ AE_2 &\subseteq E_2 \text{ where } E_2 = \operatorname{span}\{\vec{u}_1, \vec{u}_2\} \\ AE_3 &\subseteq E_3 \text{ where } E_3 = \operatorname{span}\{\vec{u}_1, \vec{u}_2, \vec{u}_3\} \\ &\vdots \\ AE_{n-1} &\subseteq E_{n-1} \text{ where } E_{n-1} = \operatorname{span}\{\vec{u}_1, \ldots, \vec{u}_{n-1}\} \ . \end{aligned}$$

In this case,

$$[A] = \begin{bmatrix} \lambda_1 & * & * & \ldots & * \\ 0 & \lambda_2 & * & \ldots & * \\ 0 & 0 & \lambda_3 & \ldots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & \lambda_n \end{bmatrix} ,$$

where the j-th column is associated with $\vec{u}_j$. We say a matrix A is **upper triangular** if it is expressed in this form.

Note that necessarily for a trianglar A, the **diagonal entries are eigenvalues**. $\vec{u}_1$ is an eigenvector, but $\vec{u}_2, \ldots, \vec{u}_n$ may not be eigenvectors.

We get a **lattice of invariant subspaces** for a triangular A. That is,

$$\{0\} \subset E_1 \subset E_2 \subset \ldots \subset E_n$$

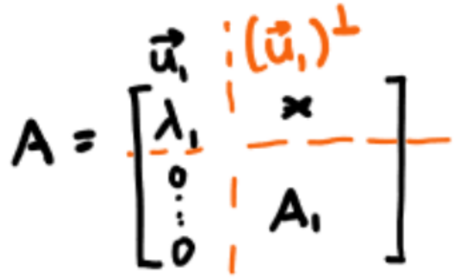
with $\dim(E_j) = j$ and $AE_j \subseteq E_j$.

Schur's Theorem

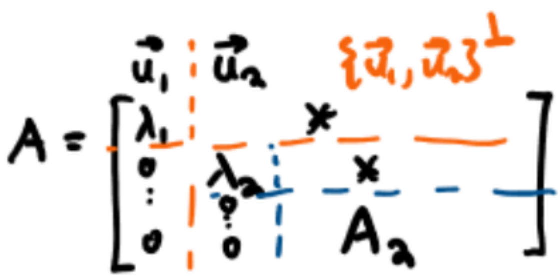
Theorem (Schur)

Suppose $A : V \to V$ is linear, where V is a finite dimension inner product space. There is an orthonormal basis $\{\vec{u}_1 \ldots, \vec{u}_n\}$ such that A is upper triangular with respect to this basis.

Proof. Let λ_1 be an eigenvalue for A. Let $\vec{u}_1$ be a unit eigenvector for λ_1. Then (this is going to difficult to typeset)



for some A. Next, find an eigenvalue λ_2 for the matrix A_1 with unit eigenvector $\vec{u}_2 \in \{\vec{u}_1\}^\perp$. Then



Continue this procedure a total of n times to find an orthonormal basis $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n\}$ with

$$A = \begin{bmatrix} \lambda_1 & * & * & \ldots & * \\ 0 & \lambda_2 & * & \ldots & * \\ 0 & 0 & \lambda_3 & \ldots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & \lambda_n \end{bmatrix},$$

where $\vec{u}_j$ is the j-th column of A. $\square$

Remark: If $U = \begin{bmatrix} \vec{u}_1 & \dots & \vec{u}_n \end{bmatrix}$, this means $U^*AU =: T$ is upper triangular.

Porism

Every $n \times n$ matrix A is unitarily equivalent to an upper triangular matrix. That is, there exists a unitary U such that

$$U^*AU = T \ ,$$

(i.e. $A \sim_U T$) where T is an upper triangular matrix.

Note: A porism is a corollary to the proof of a major result and not the result itself.

Application of Schur's Theorem

(i) The determinant of any A is the product of its eigenvalues.

Write $T = U^*AU$ with T upper triangular and U unitary. Then

$$\begin{aligned} T &= U^*AU \\ UT &= UU^*AU \\ UT &= IAU \\ UT &= AU \\ UTU^* &= AUU^* \\ UTU^* &= AI \\ UTU^* &= A \ . \end{aligned}$$

So,

$$\begin{aligned} \det(A) &= \det(UTU^*) \\ &= \det(U) \cdot \det(T) \cdot \det(U^*) \\ &= \det(U) \cdot \det(T) \cdot \det\bigl(U^{-1}\bigr) \\ &= \det(U) \cdot \det(T) \cdot \frac{1}{\det(U)} \\ &= \det(T) \\ &= \text{product of the diagonals/eigenvalues of } T \\ &= \text{product of the eigenvalues of } A \ . \end{aligned}$$

(ii) $\operatorname{tr} A =$ sum of eigenvalues.

Write $T = U^* A U$. Then

$$\begin{aligned} \lambda_1 + \ldots + \lambda_n &= \operatorname{tr} T \\ &= \operatorname{tr}(U^* A U) \\ &= \operatorname{tr}(A U U^*) \\ &= \operatorname{tr}(A I) \\ &= \operatorname{tr} A \ . \end{aligned}$$

Hence, $\operatorname{tr} T = \operatorname{tr} A$. For example, consider

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 5 & 7 & 9 \end{bmatrix} \ .$$

Here, row 3 = row 1 + row 2, so 0 is an eigenvalue (this can be checked by solving for the characteristic polynomial, of course). Let $\lambda_1 = 0$ and λ_2, λ_3 be the other eigenvalues. Then

$$\begin{aligned} \operatorname{tr} T &= \lambda_1 + \lambda_2 + \lambda_3 \\ &= 0 + \lambda_2 + \lambda_3 \\ &= 15 \\ &= 1 + 5 + 9 \\ &= \operatorname{tr} A \ . \end{aligned}$$

Self-Adjoint

Definition (Self-Adjoint)

$A : V \to V$ is **self-adjoint** if $A = A^*$.

(i) Let $A = \begin{bmatrix} 1 & i \\ -i & 2 \end{bmatrix}$. Then

$$\begin{aligned} A^* &= \begin{bmatrix} 1 & i \\ -1 & 2 \end{bmatrix}^* \\ &= \begin{bmatrix} \overline{1} & \overline{-i} \\ \overline{i} & \overline{2} \end{bmatrix} \\ &= \begin{bmatrix} 1 & i \\ -i & 2 \end{bmatrix} \\ &= A \ . \end{aligned}$$

(ii) If $A = B^*B$ for any $B : V \to W$, then

$$A^* = (B^*B)^* = B^*B^{**} = B^*B = A \ .$$

(iii) A diagonal matrix $D = \begin{bmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{bmatrix}$ is self-adjoint if and only if each $\lambda_j \in \mathbb{R}$, where $1 \leq j \leq n$. This is because

$$D^* = \begin{bmatrix} \overline{\lambda_1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \overline{\lambda_n} \end{bmatrix}$$

and D being self-adjoint would imply that $\lambda_j = \overline{\lambda_j}$ for $1 \leq j \leq n$.

(iv) If A is self-adjoint and U is unitary, then U^*AU is self-adjoint as

$$(U^*AU)^* = U^*A^*U^{**} = U^*A^*U = U^*AU \ .$$

Remark: $A = A^*$ iff $\langle A\vec{v}, \vec{w} \rangle = \langle \vec{v}, A\vec{w} \rangle$ for all $\vec{v}, \vec{w}$.

Proposition (Properties of the Self-Adjoint)

Let $A : V \to W$ and $A^* : W \to V$. If $A = A^*$, then

(i) Eigenvalues of A are real.

(ii) Eigenvectors associated to **distinct** eigenvalues are orthogonal.

(iii) $(\ker A)^\perp = \operatorname{ran} A$. That is, $V = \ker A \oplus \operatorname{ran} A$.

Note that the converse to (i) is false. For example, the matrix $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ has characteristic polynomial

$$C_A(z) = \det\left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right) = (0 \cdot 0) - (1 \cdot 0) = 0 \ ,$$

and so the eigenvalue is 0, which is real. However,

$$A^* = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}^* = \begin{bmatrix} \bar{0} & \bar{0} \\ \bar{1} & \bar{0} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \neq \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = A \ .$$

(See next page for proof of proposition)

Proof. Let V and W be (finite dimensional?) inner product spaces. Suppose $A : V \to W$ and $A^* : W \to V$ are linear operators such that $A = A^*$; that is, A is self-adjoint.

(i) Suppose $A\vec{v} = \lambda\vec{v}$, where $\vec{v} \in V$ and $\vec{v} \neq 0$. Then

$$\begin{aligned}\langle A\vec{v}, \vec{v}\rangle &= \langle \lambda\vec{v}, \vec{v}\rangle \\ &= \lambda\langle \vec{v}, \vec{v}\rangle \\ &= \lambda||\vec{v}||^2\end{aligned}$$

and

$$\begin{aligned}\langle \vec{v}, A^*\vec{v}\rangle &= \langle \vec{v}, A\vec{v}\rangle \\ &= \langle \vec{v}, \lambda\vec{v}\rangle \\ &= \overline{\lambda}\langle \vec{v}, \vec{v}\rangle \\ &= \overline{\lambda}||\vec{v}||^2 \ .\end{aligned}$$

Now, since $A = A^*$, we get that $\langle A\vec{v}, \vec{v}\rangle = \langle \vec{v}, A^*\vec{v}\rangle = \langle \vec{v}, A\vec{v}\rangle$. Thus,

$$\lambda||\vec{v}||^2 = \overline{\lambda}||\vec{v}||^2$$

for all $\lambda \in \mathbb{R}$.

(ii) Suppose λ, μ are distinct eigenvalues of A (i.e. $\lambda \neq \mu$) Suppose $\vec{u}$ and $\vec{v}$ are unit vectors satisfying

$$\begin{aligned}A\vec{u} &= \lambda\vec{u} \\ A\vec{v} &= \mu\vec{v} \ .\end{aligned}$$

Then

$$\begin{aligned}\mu\langle \vec{v}, \vec{u}\rangle &= \langle \mu\vec{v}, \vec{u}\rangle \\ &= \langle A\vec{v}, \vec{u}\rangle \\ &= \langle \vec{v}, A^*\vec{u}\rangle \\ &= \langle \vec{v}, A\vec{u}\rangle && (A = A^*) \\ &= \langle \vec{v}, \lambda\vec{u}\rangle \\ &= \overline{\lambda}\langle \vec{v}, \vec{u}\rangle \\ &= \lambda\langle \vec{v}, \vec{u}\rangle \ ,\end{aligned}$$

where $\overline{\lambda} = \lambda$ since $\lambda \in \mathbb{R}$ (from (i)). Thus, $\langle \vec{v}, \vec{u}\rangle = 0$ since $\lambda \neq \mu$.

(iii) To show that $(\ker A)^\perp = \operatorname{ran} A$, we show that $(\ker A)^\perp \subseteq \operatorname{ran} A$ and $\operatorname{ran} A \subseteq (\ker A)^\perp$.

($\subseteq$) Suppose $\vec{v} \in (\ker A)^\perp$. From the properties of adjoints, we know that $(\ker A)^\perp = \operatorname{ran} A^*$

$\square$

1. Self-adjoints span the space of linear operators on V.

   Let $A : V \to V$. Then since A is self-adjoint (i.e. $A = A^*$),

$$\begin{aligned}
A &= \frac{2A}{2} \\
&= \frac{2A}{2} + 0 \\
&= \frac{2A}{2} + \frac{0}{2} \\
&= \frac{A + A}{2} + \frac{A - A}{2} \\
&= \frac{A + A^*}{2} + \frac{A - A^*}{2} \\
&= \frac{A + A^*}{2} + i\frac{A - A^*}{2i} \ ,
\end{aligned}$$

   where 0 denotes the zero matrix. So, taking the adjoint of both of the resulting terms gives

$$\left(\frac{A + A^*}{2}\right)^* = \frac{1}{2}(A + A^*)^* = \frac{1}{2}(A^* + A^{**}) = \frac{(A^* + A)}{2}$$

   and

$$\begin{aligned}
\left(\frac{A - A^*}{2i}\right)^* &= \left(\frac{1}{2i}(A - A^*)\right)^* \\
&= \overline{\left(\frac{1}{2i}\right)}(A - A^*)^* && \text{(prop. (iii))} \\
&= \left(\frac{\overline{1}}{\overline{2i}}\right)(A^* - A^{**}) \\
&= \frac{1}{-2i}(A^* - A) \\
&= -\frac{1}{2i}(-A + A^*) \\
&= \frac{1}{2i}(A - A^*) \\
&= \frac{(A - A^*)}{2} \ .
\end{aligned}$$

   We see here that these terms are themselves self-adjoint. We call these terms $\frac{A + A^*}{2}$ and $\frac{A - A^*}{2i}$ the **real and imaginary** parts of A, respectively.

   Self adjoints are the "real numbers" inside the set of linear maps.