Honours Algebra - Week 9 - Adjoint and Self-Adjoint Endomorphisms

Antonio León Villares

March 2022

Contents

Based on the notes by Iain Gordon, Section 5.3

1 Adjoint Endomorphisms

1.1 Defining Adjoint Endomorphisms

- **When are 2 endomorphism adjoint?**
  - consider an **inner product space** V with endomorphisms:
  $$T, S : V \to V$$
  - S, T are **adjoint** if:
  $$(T(\underline{v}), \underline{w}) = (\underline{v}, S(\underline{w})), \qquad \forall \underline{v}, \underline{w} \in V$$
  - we write $S = T^*$ to say that "S is the adjoint of T"

1.1.1 Examples

One can think of **adjoints** as **equivalent** to taking **transposes** (see here for more details).

For example,if $V = \mathbb{R}^n$, we can define an endomorphism $T : V \to V$ via matrix multiplication $A\circ$ of $A \in Mat(n, \mathbb{R})$. Recall, we can rewrite the (standard) dot product as:

$$(\underline{v}, \underline{w}) = \underline{v}^T \circ \underline{w}$$

Then:

$$\begin{aligned}(A \circ \underline{v}, \underline{w}) &= (A \circ \underline{v})^T \circ \underline{w} \\ &= \underline{v}^T \circ A^T \circ \underline{w} \\ &= \underline{v}^T \circ (A^T \circ \underline{w}) \\ &= (\underline{v}, A^T \circ \underline{w})\end{aligned}$$

In other words, the **adjoint** of a real matrix is its **transpose**:

$$A^* = A^T$$

Alternatively, if $V = \mathbb{C}^n$, we have that the dot product is:

$$(\underline{v}, \underline{w}) = \underline{v}^T \circ \overline{\underline{w}}$$

So if $A \in Mat(n, \mathbb{C})$:

$$\begin{aligned}(A \circ \underline{v}, \underline{w}) &= (A \circ \underline{v})^T \circ \overline{\underline{w}} \\ &= \underline{v}^T \circ A^T \circ \overline{\underline{w}} \\ &= \underline{v}^T \circ (A^T \circ \overline{\underline{w}}) \\ &= (\underline{v}, \overline{A^T \circ \overline{\underline{w}}}) \\ &= (\underline{v}, \overline{A}^T \circ \underline{w})\end{aligned}$$

In other words, the **adjoint** of a complex matrix is its **conjugate transpose**:

$$A^* = \bar{A}^T$$

1.2 Theorem: Existence of Adjoint

Let V be a ***finite dimensional inner product space.***
Let:

$$T : V \to V$$

be an ***endomorphism****.*
Then, T^* ***exists****.*
That is, we have a ***unique*** *linear mapping:*

$$T^* : V \to V$$

such that:

$$(T(\underline{v}), \underline{w}) = (\underline{v}, T^*(\underline{w}))$$

[Theorem 5.3.4]

Proof. We have 3 steps:

1. Conjecture T^*, dependent on T
2. Shows that T^* is the adjoint of T
3. Show that T^* is linear

Now, since V is finite dimensional, then we know that there exists an **orthonormal basis**:

$$\underline{e}_1, \ldots, \underline{e}_n$$

Then, assuming that T^* exists, we would have to satisfy:

$$T^*(\underline{w}) = \sum_{i=1}^{n} (T^*(\underline{w}), \underline{e}_i)\underline{e}_i$$

T^* should be an adjoint of T, so:

$$T^*(\underline{w}) = \sum_{i=1}^{n} (\underline{w}, T(\underline{e}_i))\underline{e}_i$$

Hence, we claim that this is a valid definition of the adjoint of T.

We now show that it is a valid adjoint:

$$\begin{aligned}
(\underline{v}, T^*(\underline{w})) &= \overline{\left(\underline{v}, \sum_{i=1}^{n} (\underline{w}, T(\underline{e}_i))\underline{e}_i\right)} \\
&= \left(\sum_{i=1}^{n} (\underline{w}, T(\underline{e}_i))\underline{e}_i, \underline{v}\right) \\
&= \left(\sum_{i=1}^{n} \overline{(\underline{w}, T(\underline{e}_i))}\right) \overline{(\underline{e}_i, \underline{v})} \\
&= \left(\sum_{i=1}^{n} (T(\underline{e}_i), \underline{w})\right) (\underline{v}, \underline{e}_i) \\
&= \sum_{i=1}^{n} (T(\underline{e}_i), \underline{w})(\underline{v}, \underline{e}_i) \\
&= \sum_{i=1}^{n} (T((\underline{v}, \underline{e}_i)\underline{e}_i), \underline{w}) \\
&= \left(\sum_{i=1}^{n} T((\underline{v}, \underline{e}_i)\underline{e}_i), \underline{w}\right) \\
&= \left(T\left(\sum_{i=1}^{n} (\underline{v}, \underline{e}_i)\underline{e}_i\right), \underline{w}\right) \\
&= (T(\underline{v}), \underline{w})
\end{aligned}$$

where we have used the fact that since $\underline{e}_i$ are an orthonormal basis for V, we can write:

$$\underline{v} = \sum_{i=1}^{n} (\underline{v}, \underline{e}_i)\underline{e}_i$$

Hence, T^* certainly satisfies the property of an adjoint.

The last step is to show that it is an endomorphism.

Linearity of addition:

$$\begin{aligned}
(\underline{v}, T^*(\underline{w}_1 + \underline{w}_2)) &= (T(\underline{v}), \underline{w}_1 + \underline{w}_2) \\
&= (T(\underline{v}), \underline{w}_1) + (T(\underline{v}), \underline{w}_2) \\
&= (\underline{v}, T^*(\underline{w}_1)) + (\underline{v}, T^*(\underline{w}_2)) \\
&= (\underline{v}, T^*(\underline{w}_1) + T^*(\underline{w}_2))
\end{aligned}$$

Linearity of scalar multiplication:

$$\begin{aligned}
(\underline{v}, T^*(\lambda\underline{w})) &= (T(\underline{v}), \lambda\underline{w}) \\
&= \bar{\lambda}(T(\underline{v}), \underline{w}) \\
&= \bar{\lambda}(\underline{v}, T^*(\underline{w})) \\
&= (\underline{v}, \lambda T^*(\underline{w}))
\end{aligned}$$

Now, the last step is to show that T^* is unique. This is necessary, since:

1. It is part of the proof
2. Without it, our claims of linearity (i.e $T^*(\underline{w}_1 + \underline{w}_2)) = T^*(\underline{w}_1) + T^*(\underline{w}_2))$ don't necessarily hold, since we are making these claims by working over an inner product.

Fortunately, any **endomorphism** has at most 1 adjoint. Indeed, assume that T has 2 adjoints S, S'. Then:

$$(T(\underline{v}), \underline{w}) = (\underline{v}, S(\underline{w})) = (\underline{v}, S'(\underline{w}))$$

and this is true $\forall \underline{v}, \underline{w} \in V$. But then by linearity:

$$(\underline{v}, S(\underline{w}) - S'(\underline{w})) = 0$$

Since this is true $\forall \underline{v}, \underline{w}$, in particular it is true for $\underline{v} = S(\underline{w}) - S'(\underline{w})$, so we must have:

$$(S(\underline{w}) - S'(\underline{w}), S(\underline{w}) - S'(\underline{w})) = 0 \iff S(\underline{w}) - S'(\underline{w}) = 0$$

which implies that $\forall \underline{w} \in V$, we have that S, S' map identically, so $S = S'$, so adjoints are unique (if they exist).

Hence, we have shown the existence and uniqueness of an endomorphism T^*, which is an adjoint of T.

□

1.2.1 Examples

Given a linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$, it can be represented as a 2×2 matrix. We compute its adjoint.

$$T^*(\vec{w}) = \sum_{i=1}^{n} (\vec{w}, T(\vec{e}_i))\vec{e}_i \qquad T: \mathbb{R}^2 \to \mathbb{R}^2 = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

$$\mathbb{R}^2: \left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = x_1y_1 + x_1y_2 + x_2y_1 + 3x_2y_2$$

What is T^*?

$$T(\vec{e}_1) = \begin{pmatrix} a \\ c \end{pmatrix} \quad T(\vec{e}_2) = \frac{1}{\sqrt{2}}\begin{pmatrix} a-b \\ c-d \end{pmatrix}$$

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \vec{e}_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$$(\vec{e}_1, \vec{e}_2) = \frac{1}{\sqrt{2}}[1-1] = 0$$

$$(\vec{e}_2, \vec{e}_2) = \frac{1}{2}[1-1-1+3] = 1$$

$$T^*\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, T(\vec{e}_1)\right)\vec{e}_1 + \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, T(\vec{e}_2)\right)\vec{e}_2$$

$$= \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} a \\ c \end{pmatrix}\right)\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \frac{1}{\sqrt{2}}\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} a-b \\ c-d \end{pmatrix}\right)\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$$= (a+c)\begin{pmatrix}1\\0\end{pmatrix} + \frac{1}{2}(a-b+c-d)\begin{pmatrix}1\\-1\end{pmatrix} = \begin{pmatrix}\frac{3}{2}(a+c)-\frac{1}{2}(b+d)\\ \frac{1}{2}(b+d)-\frac{1}{2}(a+c)\end{pmatrix}$$

$$T^*\left(\begin{pmatrix}0\\1\end{pmatrix}\right) = \left(\begin{pmatrix}0\\1\end{pmatrix}, T(\vec{e}_1)\right)\vec{e}_1 + \left(\begin{pmatrix}0\\1\end{pmatrix}, T(\vec{e}_2)\right)\vec{e}_2$$

$$= (a+3c)\begin{pmatrix}1\\0\end{pmatrix} + \frac{1}{\sqrt{2}}\left(a-b+3(c-d)\right)\frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}$$

$$= \begin{pmatrix}\frac{3}{2}(a+3c)-\frac{1}{2}(b+3d)\\ \frac{1}{2}(b+3d)-\frac{1}{2}(a+3c)\end{pmatrix} \qquad 3($$

Thus, it follows that:

$$T = \begin{pmatrix}a & b\\ c & d\end{pmatrix} \implies T^* = \frac{1}{2}\begin{pmatrix}3(a+c)-(b+d) & 3(a+3c)-(b+3d)\\ (b+d)-(a+c) & (b+3d)-(a+3c)\end{pmatrix}$$

1.2.2 Exercises (TODO)

1. **Show that if T^* is the adjoint of T, then T^* has an adjoint, and:**

$$(T^*)^* = T$$

Consider 2 vectors $\underline{v}, \underline{w}$. Then, assume that the adjoint of T is T^*. Then, $\forall \underline{v}, \underline{w}$:

$$\begin{aligned}(T^*(\underline{v}), \underline{w}) &= \overline{(\underline{w}, T^*(\underline{v}))}\\ &= \overline{(T(\underline{w}), \underline{v})}\\ &= (\underline{v}, T(\underline{w}))\end{aligned}$$

But similarly, by definition:

$$(T^*(\underline{v}), \underline{w}) = (\underline{v}, (T^*)^*(\underline{w}))$$

Hence, $\forall \underline{v}, \underline{w}$:

$$(\underline{v}, T(\underline{w})) = (\underline{v}, T(\underline{w})) \iff (\underline{v}, T(\underline{w}) - (T^*)^*(\underline{w})) = 0$$

and by similar arguments as above, taking $\underline{v} = T(\underline{w}) - (T^*)^*(\underline{w})$, then implies that:

$$T(\underline{w}) - (T^*)^*(\underline{w}) = 0 \implies T = (T^*)^*$$

as required.

2 Self-Adjoint Endomorphism

2.1 Defining Self-Adjoint Endomorphisms

- **When is an endomorphism self-adjoint?**

- consider an **inner product space** V with endomorphism:

$$T : V \to V$$

- T is **self-adjoint** if:

$$T = T^*$$

2.1.1 Examples

Before we showed that:

- for **real** matrices, the **adjoint** is its **tranpose**
- for **complex** matrices, the **adjoint** is its **conjugate transpose**

This then tells us that:

- a **real** matrix is **self-adjoint** if it's **symmetric**:

$$A = A^T$$

- a **complex matrix** is **self-adjoint** if it's **hermitian**:

$$A = \bar{A}^T$$

2.2 Theorem: Properties of Self-Adjoint Endomorphisms

*Let V be an **inner-product** space.*
*Consider a **self-adjoint** linear mapping:*

$$T : V \to V$$

Then:

1. *Every **eigenvalue** of T is **real***
2. *If 2 **eigenvalues** λ, μ are distinct, their corresponding **eigenvectors** $\underline{v}, \underline{w}$ are **orthogonal**:*

$$(\underline{v}, \underline{w}) = 0$$

3. *T has an **eigenvalue***

[Theorem 5.3.7]

Proof. Consider non-zero vectors $\underline{v}, \underline{w} \in V$ and a self-adjoint mapping $T = T^*$.

1. Assume that $\underline{v}$ is an eigenvector of T with eigenvalue $\underline{v}$. Then:

$$T(\underline{v}) = \lambda \underline{v}$$

Now, consider:

$$\lambda(\underline{v}, \underline{v}) = (\lambda \underline{v}, \underline{v}) = (T(\underline{v}), \underline{v}) = (\underline{v}, T^*(\underline{v}))$$

But now, T is self-adjoint, so $(\underline{v}, T^*(\underline{v})) = (\underline{v}, T(\underline{v}))$. Thus:

$$\lambda(\underline{v}, \underline{v}) = (\underline{v}, T(\underline{v})) = (\underline{v}, \lambda\underline{v}) = \bar{\lambda}(\underline{v}, \underline{v})$$

(Notice, using $\bar{\lambda}$ is general, since its a property of both real and complex inner products).

But then we have that:

$$\lambda(\underline{v}, \underline{v}) = \bar{\lambda}(\underline{v}, \underline{v})$$

Since $\underline{v} \neq 0$, then $(\underline{v}, \underline{v}) > 0$. V being an integral domain then implies that:

$$\lambda = \bar{\lambda} \iff \lambda \in \mathbb{R}$$

2. By the above, if λ, μ are eigenvalues, then $\lambda, \mu \in \mathbb{R}$. Thus:

$$\begin{aligned}
\lambda(\underline{v}, \underline{w}) &= (\lambda\underline{v}, \underline{w}) \\
&= (T(\underline{v}), \underline{w}) \\
&= (\underline{v}, T^*(\underline{w})) \\
&= (\underline{v}, T(\underline{w})) \\
&= (\underline{v}, \mu\underline{w}) \\
&= \bar{\mu}(\underline{v}, \underline{w}) \\
&= \mu(\underline{v}, \underline{w})
\end{aligned}$$

By hypothesis, $\lambda \neq \mu$, so:

$$\lambda(\underline{v}, \underline{w}) = \mu(\underline{v}, \underline{w}) \iff (\underline{v}, \underline{w}) = 0$$

so the eigenvectors are orthogonal.

3. We now consider the 2 possible types of inner product spaces.

If V is a complex inner product space, the fact that it has an eigenvalue is no surprise. Since $\mathbb{C}$ is algebraically closed, any characteristic polynomial will have roots in $\mathbb{C}$. This is also Theorem 4.5.4 in the notes.

The interesting case comes wehn we consider a real inner product space; after all, we know there are many endomorphisms (i.e 90^{o} rotations) which don't have **any** real eigenvalue.

To prove this, we use a rather contrived method, which requires analysis (yuck!), but which leads to a pretty nice geometric consequence which we met in ILA.

We now work with V as a finite dimensional, real, inner product space. Define the **Railegh Quotient**:

$$R(\underline{v}) = \frac{(T(\underline{v}), \underline{v})}{(\underline{v}, \underline{v})}, \qquad \underline{v} \in V \setminus \{\underline{0}\}$$

The first thing to note is that we can restrict ourselves to a unit sphere:

$$S = \{\underline{v} \mid \|\underline{v}\| = 1\}$$

This is because:

$$R(\underline{v}) = \frac{(T(\underline{v}), \underline{v})}{(\underline{v}, \underline{v})} = \frac{1}{\|\underline{v}\|^2}(T(\underline{v}), \underline{v}) = \left(T\left(\frac{\underline{v}}{\|\underline{v}\|}\right), \frac{\underline{v}}{\|\underline{v}\|}\right)$$

So R is fully defined by unit vectors.

Since I hate analysis, I'll go out an say that the set S is **closed** (since its complement is an open set: just think of all vectors with length greater than one!) and **bounded** (since it's composed of unit vectors), the **Extreme Value Theorem** states that a continuous function over said interval will achieve a maximum and a minimum. Admittedly, we haven't shown that $R(\underline{v})$ is continuous, but in the videos Iain shows that we can write:

$$(T(\underline{v}), \underline{v}) = \sum_{i,j} \lambda_i \lambda_j (T(\underline{e}_i), \underline{e}_j)$$

where $\lambda_i = (\underline{v}, \underline{e}_i)$ and we have a set of orthonormal basis vectors $\underline{e}_i$. This can be thought as a polynomial in λ_i, λ_j, so as all polynomials, $R(\underline{v})$ is **continuous** and **differentiable**.

Given a symmetric $n \times n$ matrix $A = (a_{ij}) \in \mathrm{Mat}(n; \mathbb{R})$ let $T : \mathbb{R}^n \to \mathbb{R}^n$ be the self-adjoint endomorphism with matrix A defined by

$$T(x_1, x_2, \ldots, x_n) = (\sum_{j=1}^{n} a_{1j}x_j, \sum_{j=1}^{n} a_{2j}x_j, \ldots, \sum_{j=1}^{n} a_{nj}x_j) \in \mathbb{R}^n .$$

Consider the Rayleigh quotient function as in the proof of Theorem 5.3.7

$$R : \mathbb{R}^n \backslash \{\vec{0}\} \to \mathbb{R} ; \vec{x} = (x_1, x_2, \ldots, x_n) \mapsto R(\vec{x}) = \frac{T\vec{x} \bullet \vec{x}}{\vec{x} \bullet \vec{x}} = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j}{\sum_{k=1}^{n} (x_k)^2} .$$

Figure 1: Here we are using the standard inner product.

All this for what, I asked myself as reading (and presuming what I'll think when reading this in the future)? Well, we know there exists some vector $\underline{v}_+$ (and we can assume that $\underline{v}_+$ is a unit vector) such that $R(\underline{v}_+)$ is maximal.

We can now define a function:

$$R_{\underline{w}}(t) = R(\underline{v}_+ + t\underline{w}) = \frac{(T(\underline{v}_+ + t\underline{w}), \underline{v}_+ + t\underline{w})}{(\underline{v}_+ + t\underline{w}, \underline{v}_+ + t\underline{w})}$$

For those of you (aka me in the many futures in which I consult this) which are geometrically inclined:

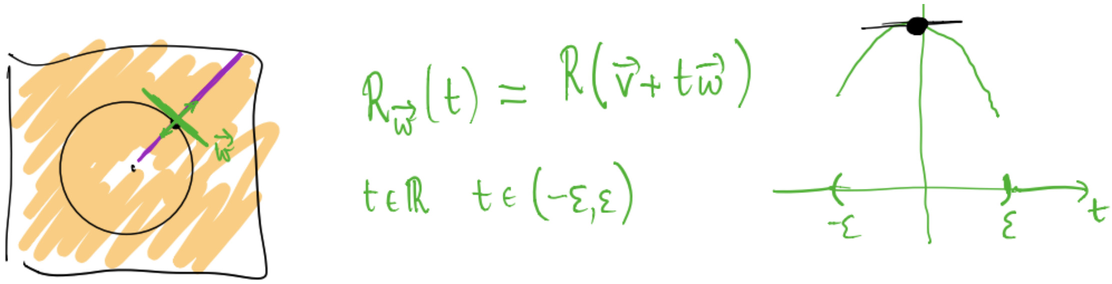


Figure 2: We define $R_{\underline{w}}$ for small t, so that we are still in the unit sphere (I think?). We can see that the maximum is achieved when $t = 0$, since $R(\underline{v}_+)$ is maximal.

Now notice: if we differentiate $R_{\underline{w}}(0)$, we should get 0, since $R(\underline{v}_+)$ is maximal (and differentiation is defined, since R is differentiable). Thus, applying the quotient rule:

$$R'_{\underline{w}}(0) = \frac{(T(\underline{w}), \underline{v}_+) + (T(\underline{v}_+), \underline{w})}{(\underline{v}_+, \underline{v}_+)} - \frac{2(T(\underline{v}_+), \underline{v}_+)(\underline{v}_+, \underline{w})}{(\underline{v}_+, \underline{v}_+)^2}$$

If we then make the smart choice of picking $\underline{w}$ such that:

$$\underline{w} \perp \underline{v}_+$$

Then we have:

$$R'_{\underline{w}}(0) = \frac{(T(\underline{w}), \underline{v}_+) + (T(\underline{v}_+), \underline{w})}{(\underline{v}_+, \underline{v}_+)} = 0 \implies (T(\underline{w}), \underline{v}_+) = -(T(\underline{v}_+), \underline{w})$$

Now, since we are working over a real inner product space:

$$(T(\underline{v}_+), \underline{w}) = (\underline{w}, T(\underline{v}_+))$$

Moreover, T is self-adjoint, so:

$$(T(\underline{w}), \underline{v}_+) = (\underline{w}, T(\underline{v}_+))$$

Hence, we have that:

$$(T(\underline{w}), \underline{v}_+) = -(T(\underline{v}_+), \underline{w}) \implies (\underline{w}, T(\underline{v}_+)) = -(\underline{w}, T(\underline{v}_+))$$

and this is true **if and only if**:

$$(\underline{w}, T(\underline{v}_+)) = 0$$

which means that:

$$\underline{w} \perp T(\underline{v}_+)$$

This means that since:

$$\underline{w} \in (\langle \underline{v}_+ \rangle)^\perp$$

then:

$$T(\underline{v}_+) \in ((\langle \underline{v}_+ \rangle)^\perp)^\perp$$

Proposition 5.2.2 tells us then that:

$$((\langle \underline{v}_+ \rangle)^\perp)^\perp = \langle \underline{v}_+ \rangle$$

so in particular, each $T(\underline{v}_+)$ must be in the span of $\underline{v}_+$, or in other words, $\exists \lambda \in \mathbb{R}$ such that:

$$T(\underline{v}_+) = \lambda \underline{v}_+$$

thus showing that T has an eigenvalue.

□

2.2.1 Example: The Geometric Interpretation of Raleigh Quotient

Consider the transofrmation:

$$T = \begin{pmatrix} 5 & -6 \\ -6 & 13 \end{pmatrix}$$

Since T is a real symmetric matrix, it is self-adjoint.

Now, consider the case of maximising/minimising:

$$R(\underline{v}) = \frac{T(\underline{v}), \underline{v})}{(\underline{v}, \underline{v})}$$

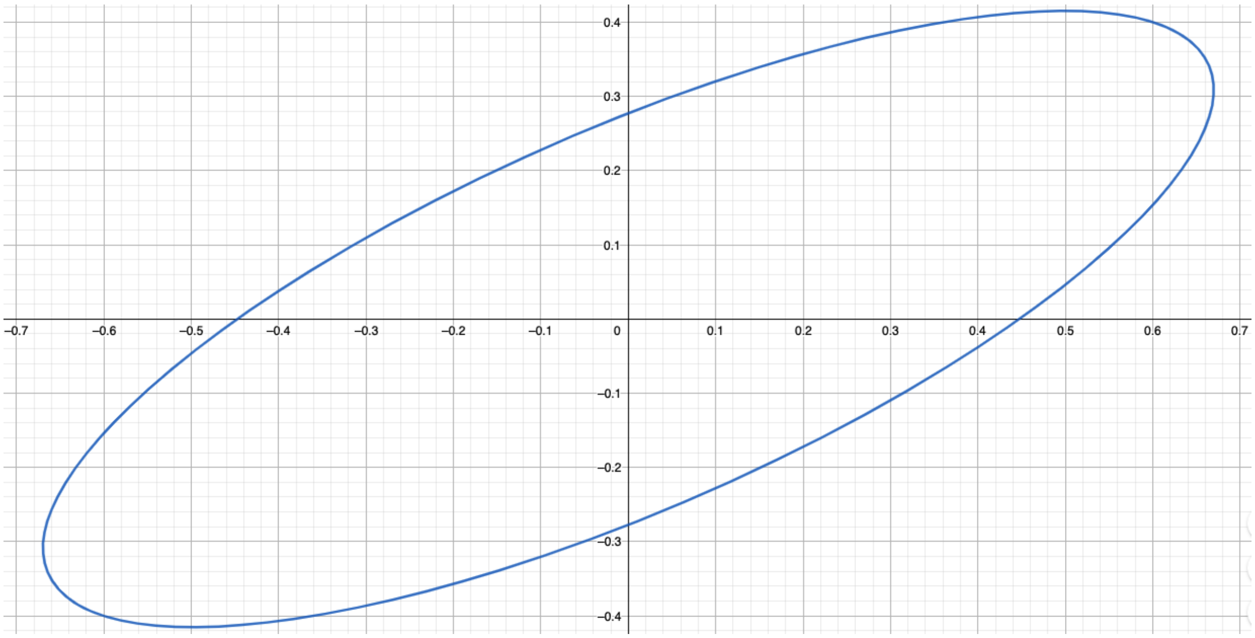
In the above, we maximised R by keeping the denominator constant (the restriction that $\underline{v}$ lies in the unit sphere). However, an alternative is to maximise R by minimising $(\underline{v}, \underline{v})$, given that $T(\underline{v}), \underline{v})$ stays constant.

For example, we can minimise $\|\underline{v}\|$ whilst ensuring that $(T(\underline{v}), \underline{v}) = 1$. Then, using $\underline{v} = (x, y)$:

$$T(\underline{v}) = \begin{pmatrix} 5 & -6 \\ -6 & 13 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 5x - 6y \\ -6x + 13y \end{pmatrix}$$

$$(T(\underline{v}), \underline{v}) = 5x^2 - 6xy - 6xy + 13y^2 = 5x^2 - 12xy + 13y^2$$

We can plot $5x^2 - 12xy + 13y^2 = 1$ and obtain ... an ellipse!



Now, our objective is to **minimise** $\|\underline{v}\|$, where $\underline{v}$ is any of the vectors which define the ellipse. That is we seek the $\underline{v}$ whose distance from the origin to the ellipse is minimal. In other words: **$\underline{\mathbf{v}}$ corresponds to the minor axis of the ellipse above**.

Similarly, if we wanted to **minimise** R, we would seek to maximise $\|\underline{v}\|$, in which case we would have found the **major axis** of the ellipse.

And what do you know? The axes defining the ellipse are eigenvectors of the self-adjoint mapping. And as expected, they are **orthogonal** to each other!

3 The Spectral Theorem

3.1 Theorem: The Spectral Theorme for Self-Adjoint Endomorphisms

Let V be a ***finite dimensional inner product space****.*
Define the self-adjoint endomorphism:

$$T : V \to V$$

Then, V has an ***orthonormal basis****, consisting of* ***eigenvectors*** *of T.*
[Theorem 5.3.9]

Proof. We apply induction on $n = dim(V)$.

① **n = 0,1**

These cases are trivial. If $dim(V) = 0$, then $V = \{\underline{0}\}$, for which there are no eigenvectors. Similarly, if $dim(V) = 1$, then $V = \langle \underline{v} \rangle$, and any endomorphism T must have $\underline{v}$ as its only eigenvector, which is clearly orthonormal to itself (since $\lambda \underline{v} = \underline{0} \iff \lambda = 0$).

② **n = k**

Assume true for $n = k$. That is, if $dim(V) = k$, there exists an orthonormal basis for V.

③ **n = k+1**

Consider a space V with $dim(V) = k + 1$. Since T is self adjoint, we know that T has at least one real eigenvalue λ (2.1.1). Define $\underline{u}$ to be a unit eigenvector of λ, and define a subspace $U = \langle \underline{u} \rangle$.

Now, let $\underline{v} \in U^{\perp}$. Then:

$$(\underline{u}, T(\underline{v})) = (\underline{T(u), v} = (\lambda \underline{u}, \underline{v}) = \lambda(\underline{u}, \underline{v}) = 0$$

where we have used the fact that $\underline{u}$ is an eigenvector of T, and the fact that T is self-adjoint.

Notice, this then means that $\underline{u} \perp T(\underline{v})$, so we must have that:

$$T(U^{\perp}) \subseteq U^{\perp}$$

In particular, we can define an endomorphism $T_{U^{\perp}} : U^{\perp} \to U^{\perp}$ by restricting T to $U^{\perp}$. Now recall, the dimension theorem:

$$dim(A + B) + dim(A \cap B) = dim(A) + dim(B)$$

Moreover, recall that:

$$V = U \oplus U^{\perp}$$

and that this implies that $U \cap U^{\perp} = \emptyset$ and $U + U^{\perp} = V$ (we showed this last week). Then, the dimension theorem tells us that:

$$dim(V) = dim(U) + dim(U^{\perp})$$

so:

$$k + 1 = 1 + dim(U^{\perp}) \implies dim(U^{\perp}) = k$$

Hence, since T being self-adjoint means that $T_{U^{\perp}}$ is also self-adjoint, the induction hypothesis exists, and $U^{\perp}$ has an orthonormal basis (of k elements), call it B.

Then, $B \cup \{\underline{u}\}$ produces an orthonormal basis for V, as required (since its a set of $k+1$ linearly independent vectors, and they will span V, since as discussed above $V = U \oplus U^{\perp}$).

□

3.2 Orthogonal Matrices

- **What is an orthogonal matrix?**
  - let $P \in Mat(n, \mathbb{R})$
  - P is an **orthogonal matrix** if:

$$P^T P = I_n \implies P^T = P^{-1}$$

3.2.1 Examples

Consider the matrix:

$$T = \begin{pmatrix} 5 & -6 \\ -6 & 13 \end{pmatrix}$$

(this was used to define the ellipse above)

We want to create a change of basis matrix which maps from the standard basis to a basis of the eigenvectors of T. The fact that T is self-adjoint (since its real and symmetric) implies that V has an orthonormal basis consisting of eigenvectors of T, by the Spectral Theorem.

We begin by computing the eigenvectors and eigenvalues:

$$\begin{vmatrix} 5-x & -6 \\ -6 & 13-x \end{vmatrix} = (5-x)(13-x) - 36 = x^2 - 18x + 29 = 0$$

Applying the quadratic formula defines 2 real roots (as expected from a self-adjoint transformation):

$$\alpha = 9 + 2\sqrt{13}$$
$$\beta = 9 - 2\sqrt{13}$$

We now find the eigenvectors:

$$\begin{pmatrix} 5-\alpha & -6 \\ -6 & 13-\alpha \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \implies \begin{pmatrix} (5-\alpha)v_1 - 6v_2 \\ -6v_1 + (13-\alpha)v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Then:

$$v_2 = \frac{(5-\alpha)v_1}{6} \implies v_1 = 6,\ v_2 = 5-\alpha$$

So:

$$\underline{v}_1 = \begin{pmatrix} 6 \\ 5-\alpha \end{pmatrix}$$

From identical reasoning it follows that:

$$\underline{v}_2 = \begin{pmatrix} 6 \\ 5-\beta \end{pmatrix}$$

We confirm that these 2 vectors are orthogonal:

$$\begin{aligned} \underline{v}_1 \cdot \underline{v}_2 &= 36 + (5-\alpha)(5-\beta) \\ &= 36 + 25 - 5\alpha - 5\beta + \alpha\beta \\ &= 61 - 5(9+2\sqrt{13}) - 5(9-2\sqrt{13}) + (9+2\sqrt{13})(9-2\sqrt{13}) \\ &= 61 - 90 + 81 - 4(13) \\ &= 142 - 142 \\ &= 0 \end{aligned}$$

We can now normalise $\underline{v}_1, \underline{v}_2$ to obtain unit vectors:

$$\underline{u}_1 = \frac{1}{\sqrt{36 + (5-\alpha)^2}}\begin{pmatrix} 6 \\ 5-\alpha \end{pmatrix}$$

$$\underline{u}_2 = \frac{1}{\sqrt{36+(5-\beta)^2}} \begin{pmatrix} 6 \\ 5-\beta \end{pmatrix}$$

Define a basis of V via $B = \{\underline{u}_1, \underline{u}_2\}$. The change of basis matrix ${}_{S(2)}[id_V]_B$ is constructed by taking the eigenvectors as columns:

$${}_{S(2)}[id_V]_B = P = \begin{pmatrix} \frac{6}{\sqrt{36+(5-\alpha)^2}} & \frac{6}{\sqrt{36+(5-\beta)^2}} \\ \frac{5-\alpha}{\sqrt{36+(5-\alpha)^2}} & \frac{5-\beta}{\sqrt{36+(5-\beta)^2}} \end{pmatrix}$$

To simplify this a bit, we can expand:

$$\begin{aligned} 36+(5-\alpha)^2 &= 36+25-10\alpha+\alpha^2 \\ &= \alpha^2-10\alpha+61 \\ &= (18\alpha-29)-10\alpha+61 \\ &= 8\alpha+32 \\ &= 4(2\alpha+8) \end{aligned}$$

The same applies to $36+(5-\alpha)^2$, since we are just using the fact that α, β satisfy $x^2-18x+29=0$. Thus:

$$P = \begin{pmatrix} \frac{6}{2\sqrt{2\alpha+8}} & \frac{6}{2\sqrt{2\beta+8}} \\ \frac{5-\alpha}{2\sqrt{2\alpha+8}} & \frac{5-\beta}{2\sqrt{2\beta+8}} \end{pmatrix}$$

The change of basis matrix ${}_B[id_V]_{S(2)}$ is nothing but P^{-1}. We can compute this:

$$det(P) = \frac{6(5-\beta)-6(5-\alpha)}{4\sqrt{2\alpha+8}\sqrt{2\beta+8}} = \frac{6(\alpha-\beta)}{4\sqrt{2\alpha+8}\sqrt{2\beta+8}} = \frac{6\sqrt{13}}{\sqrt{2\alpha+8}\sqrt{2\beta+8}}$$

But now notice that:

$$\sqrt{2\alpha+8}\sqrt{2\beta+8} = \sqrt{4\alpha\beta+16\alpha+16\beta+64} = \sqrt{4(81-52)+16(18)+64} = \sqrt{468}$$

$$6\sqrt{13} = \sqrt{36\times 13} = \sqrt{468}$$

It is useful to "save" the fact that:

$$\alpha+\beta = 18 \qquad \alpha\beta = 29$$

(this is just Vieta's Theorem!).

So $det(P) = 1$. Hence

$$P^{-1} = \begin{pmatrix} \frac{5-\beta}{2\sqrt{2\beta+8}} & -\frac{6}{2\sqrt{2\beta+8}} \\ -\frac{5-\alpha}{2\sqrt{2\alpha+8}} & \frac{6}{2\sqrt{2\alpha+8}} \end{pmatrix}$$

At this point, I haven't found a nice algebraic way of showing that $P^{-1} = P^T$ (year spoiler P is orthogonal - prove of this in the exercise below), so I'll go for second best: simply showing that $PP^T = I_2$.

$$
\begin{aligned}
PP^T &= \begin{pmatrix} \frac{6}{2\sqrt{2\alpha+8}} & \frac{6}{2\sqrt{2\beta+8}} \\ \frac{5-\alpha}{2\sqrt{2\alpha+8}} & \frac{5-\beta}{2\sqrt{2\beta+8}} \end{pmatrix} \begin{pmatrix} \frac{6}{2\sqrt{2\alpha+8}} & \frac{5-\alpha}{2\sqrt{2\alpha+8}} \\ \frac{6}{2\sqrt{2\beta+8}} & \frac{5-\beta}{2\sqrt{2\beta+8}} \end{pmatrix} \\
&= \begin{pmatrix} \frac{36}{4(2\alpha+8)} + \frac{36}{4(2\beta+8)} & \frac{6(5-\alpha)}{4(2\alpha+8)} + \frac{6(5-\beta)}{4(2\beta+8)} \\ \frac{6(5-\alpha)}{4(2\alpha+8)} + \frac{6(5-\beta)}{4(2\beta+8)} & \frac{(5-\alpha)^2}{4(2\alpha+8)} + \frac{(5-\beta)^2}{4(2\beta+8)} \end{pmatrix} \\
&= \begin{pmatrix} \frac{9}{2\alpha+8} + \frac{9}{2\beta+8} & \frac{3(5-\alpha)}{2(2\alpha+8)} + \frac{3(5-\beta)}{2(2\beta+8)} \\ \frac{3(5-\alpha)}{2(2\alpha+8)} + \frac{3(5-\beta)}{2(2\beta+8)} & \frac{(5-\alpha)^2}{4(2\alpha+8)} + \frac{(5-\beta)^2}{4(2\beta+8)} \end{pmatrix} \\
&= \begin{pmatrix} \frac{9(2\beta+8+2\alpha+8)}{(2\alpha+8)(2\beta+8)} & \frac{3((5-\alpha)(2\beta+8)+(5-\beta)(2\alpha+8)}{2(2\alpha+8)(2\beta+8)} \\ \frac{3((5-\alpha)(2\beta+8)+(5-\beta)(2\alpha+8)}{2(2\alpha+8)(2\beta+8)} & \frac{(5-\alpha)^2(2\beta+8)+(5-\beta)^2(2\alpha+8)}{4(2\alpha+8)(2\beta+8)} \end{pmatrix} \\
&= \begin{pmatrix} \frac{9(2(\alpha+\beta)+16)}{468} & \frac{3(80+2(\alpha+\beta-2\alpha\beta))}{2(2\alpha+8)(2\beta+8)} \\ \frac{3(80+2(\alpha+\beta-2\alpha\beta))}{2(2\alpha+8)(2\beta+8)} & \frac{(4(2\alpha+8)-36)(2\beta+8)+(4(2\beta+8)-36)(2\alpha+8)}{4(2\alpha+8)(2\beta+8)} \end{pmatrix} \\
&= \begin{pmatrix} \frac{9(2(18)+16)}{468} & \frac{3(80+2((18)-2(29)))}{2(2\alpha+8)(2\beta+8)} \\ \frac{3(80+2((18)-2(29)))}{2(2\alpha+8)(2\beta+8)} & \frac{(4(2\alpha+8)-36)(2\beta+8)+(4(2\beta+8)-36)(2\alpha+8)}{4(2\alpha+8)(2\beta+8)} \end{pmatrix} \\
&= \begin{pmatrix} \frac{468}{468} & \frac{3(80-80)}{2(2\alpha+8)(2\beta+8)} \\ \frac{3(80-80)}{2(2\alpha+8)(2\beta+8)} & \frac{2(2\alpha+8)(2\beta+8)-9(2\alpha+8+2\beta+8)}{(2\alpha+8)(2\beta+8)} \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 \\ 0 & 2 - \frac{9(2(\alpha+\beta)+16)}{468} \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 \\ 0 & 2 - \frac{9(2(18)+16)}{468} \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 \\ 0 & 2 - \frac{9(2(18)+16)}{468} \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 \\ 0 & 2-1 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}
\end{aligned}
$$

So indeed, $P^{-1} = P^T$, and P is an orthogonal matrix!

Our last step is to build the matrix ${}_B[T]_B$. I'll spare you the pain, since we know that this is nothing but the diagonal matrix with eigenvalues at the diagonal:

$$
\begin{aligned}
{}_B[T]_B &= P^{-1}TP \\
&= \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}
\end{aligned}
$$

3.2.2 Exercises (TODO)

1. **Show that the condition:**

$$P^T P = I_n$$

is equivalent to the columns of P forming an orthonormal basis for $\mathbb{R}^n$ with its standard inner product.

Consider an orthonormal basis $E = \{\underline{e}_1, \dots, \underline{e}_n\}$, and let P be a matrix with E as columns:

$$P = (\underline{e}_1 \mid \dots \mid \underline{e}_n)$$

Then:

$$P^T = \begin{pmatrix} \underline{e}_1^T \\ \hline \vdots \\ \hline \underline{e}_n^T \end{pmatrix}$$

Then:

$$\begin{aligned} P^T P &= \begin{pmatrix} \underline{e}_1^T \\ \hline \vdots \\ \hline \underline{e}_n^T \end{pmatrix} (\underline{e}_1 \mid \dots \mid \underline{e}_n) \\ &= \begin{pmatrix} \underline{e}_1^T \underline{e}_1 & \underline{e}_2^T \underline{e}_1 & \dots & \underline{e}_n^T \underline{e}_1 \\ \underline{e}_1^T \underline{e}_2 & \underline{e}_2^T \underline{e}_2 & \dots & \underline{e}_n^T \underline{e}_2 \\ \vdots & \vdots & \ddots & \vdots \\ \underline{e}_1^T \underline{e}_n & \underline{e}_2^T \underline{e}_n & \dots & \underline{e}_n^T \underline{e}_n \end{pmatrix} \\ &= \begin{pmatrix} (\underline{e}_1, \underline{e}_1) & (\underline{e}_2, \underline{e}_1) & \dots & (\underline{e}_n, \underline{e}_1) \\ (\underline{e}_1, \underline{e}_2) & (\underline{e}_2, \underline{e}_2) & \dots & (\underline{e}_n, \underline{e}_2) \\ \vdots & \vdots & \ddots & \vdots \\ (\underline{e}_1, \underline{e}_n) & (\underline{e}_2, \underline{e}_n) & \dots & (\underline{e}_n, \underline{e}_n) \end{pmatrix} \\ &= I_n \end{aligned}$$

by using the orthonormality of $\underline{e}_i$.

2. **Show that the set:**

$$\{P \mid P^T P = I_n, \quad P \in Mat(n, \mathbb{R})\}$$

is a group (this is known as the *orthogonal gorup*, $O(n)$).

We now consider specific cases of self-adjoint matrices, for which we can apply the Spectral Theorem

3.3 Corollary: The Spectral Theorem for Real Symmetric Matrices

Let $A \in Mat(n; \mathbb{R})$, and let A be a ***symmetric*** *matrix:*

$$A = A^T$$

Then:

$$\exists P \in Mat(n; \mathbb{R}) \quad : \quad P^{-1}AP = P^TAP = diag(\lambda_1, \ldots, \lambda_n)$$

where:

- *P is an* ***orthogonal*** *matrix*
- *$\lambda_1, \ldots, \lambda_n$ are the* ***real*** *eigenvalues of A (including those with repeated multiplicity)*

[Corollary 5.3.14]

Proof. Thinking of A as an endomorphism:

$$(A\circ) : \mathbb{R}^n \to \mathbb{R}^n$$

we apply the spectral theorem. This tells us that we have an orthonormal basis of eigenvectors $\{\underline{v}_1, \ldots, \underline{v}_n\}$. We construct the orthogonal matrix P by using this orthonormal basis, as in the exercise above. Then, notice that:

$$\begin{aligned} AP &= A(\underline{v}_1 \mid \ldots \mid \underline{v}_n) \\ &= (A\underline{v}_1 \mid \ldots \mid A\underline{v}_n) \\ &= (\lambda_1\underline{v}_1 \mid \ldots \mid \lambda_n\underline{v}_n) \\ &= Pdiag(\lambda_1, \ldots, \lambda_n) \end{aligned}$$

which directly implies that:

$$P^{-1}AP = diag(\lambda_1, \ldots, \lambda_n)$$

□

3.4 Unitary Matrices

- **What is a unitary matrix?**
  - let $P \in Mat(n; \mathbb{C})$
  - P is a **unitary matrix** if:
    $$\bar{P}^TP = I_n$$
  - alternatively, a **complex** matrix with:
    $$P^{-1} = \bar{P}^T$$

3.4.1 Exercises (TODO)

1. **Show that the condition:**

$$\bar{P}^T P = I_n$$

is equivalent to the columns of P forming an orthonormal basis for $\mathbb{C}$ with its standard inner product.

2. **Show that the set:**

$$\{P \mid \bar{P}^T P = I_n, \quad P \in Mat(n;\mathbb{C})\}$$

is a group (called the unitary group, $U(n)$).

3.5 Theorem: The Spectral Theorem for Hermitian Matrices

*Let $A \in Mat(n;\mathbb{C})$, and let A be a **hermitian** matrix:*

$$A = \bar{A}^T$$

Then:

$$\exists P \in Mat(n;\mathbb{C}) \quad : \quad P^{-1}AP = \bar{P}^T AP = diag(\lambda_1, \ldots, \lambda_n)$$

where:

- *P is an **unitary** matrix*
- *$\lambda_1, \ldots, \lambda_n$ are the **real** eigenvalues of A (including those with repeated multiplicity)*

[Corollary 5.3.15]

Proof. Identical to the one above, but using the fact that $\bar{P}^T P = I_n$. □

4 Workshop

1. **True or false. Consider $\mathbb{R}^2$ equipped with the usual inner product. The orthogonal complement to the set:**

$$U = \{(x, y) \mid xy = 1\} \subset \mathbb{R}^2$$

is $\{(0, 0)\}$

This is true.

Solutions:

- *for any set $U^\perp = \langle U \rangle^\perp$ (that is, the orthogonal complement of a set is the orthogonal complement of its spanning vectors)*
- *the span of U is $\mathbb{R}^2$*
- *the only vector which is orthogonal to all $\mathbb{R}^2$ is $\underline{0}$*

Notice, U contains the following vectors:

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ \frac{1}{2} \end{pmatrix}$$

These vectors span $\mathbb{R}^2$, since they are linearly independent, and $dim(\mathbb{R}^2) = 2$. Then, if $\underline{v} \in U^\perp$, in particular it is orthogonal to these vectors, which means that it is orthogonal to any vector in $\mathbb{R}^2$, since if $\underline{w} \in \mathbb{R}^2$ then $\exists a, b$ such that:

$$\underline{w} = a \begin{pmatrix} 1 \\ 1 \end{pmatrix} + b \begin{pmatrix} 2 \\ \frac{1}{2} \end{pmatrix}$$

so that $(\underline{v}, \underline{w}) = 0$. The only vector which is orthogonal with all of $\mathbb{R}^2$ is $\underline{0}$, so we must have that $\underline{v} = \underline{0}$.

2. **Let V be the finite dimensional real inner product space, and suppose T is an endomorphism of V.**

   (a) **Show that**

   $$\frac{T + T^*}{2}$$

   is self-adjoint.

   We just compute, applying linearity and symmetry of the inner product:

   $$\begin{aligned}\left(\left(\frac{T+T^*}{2}\right)\underline{v}, \underline{w}\right) &= \frac{1}{2}\left(T\underline{v} + T^*\underline{v}, \underline{w}\right) \\ &= \frac{1}{2}\left((T\underline{v}, \underline{w}) + (T^*\underline{v}, \underline{w})\right) \\ &= \frac{1}{2}\left((\underline{v}, T^*\underline{w}) + (\underline{v}, T\underline{w})\right) \\ &= \frac{1}{2}\left(\underline{v}, (T + T^*)\underline{w}\right) \\ &= \left(\underline{v}, \left(\frac{T+T^*}{2}\right)\underline{w}\right)\end{aligned}$$

   Hence, $\frac{T+T^*}{2}$ is self-adjoint.

   (b) **Show that there is an orthonormal basis $\{\underline{v}_1, \dots, \underline{v}_n\}$ of V consisting of eigenvectors of $\frac{T+T^*}{2}$ such that the eigenvalue corresponding to $\underline{v}_i$ is $(T\underline{v}_i, \underline{v}_i)$.**

   By the Spectral Theorem:

   > *Let V be a* ***finite dimensional inner product space****.*
   > *Define the self-adjoint endomorphism:*
   >
   > $$T : V \to V$$
   >
   > *Then, V has an* ***orthonormal basis****, consisting of* ***eigenvectors*** *of T.*
   > *[Theorem 5.3.9]*

   So since $\frac{T+T^*}{2}$ is self-adjoint, its eigenvectors form an orthonormal basis for V. Call them $\{\underline{v}_1, \dots, \underline{v}_n\}$. The eigenvalues satisfy:

   $$\left(\frac{T+T^*}{2}\right)\underline{v}_i = \lambda_i \underline{v}_i$$

If we take the inner product with $\underline{v}_i$ of both sides:

$$\left(\left(\frac{T+T^*}{2}\right)\underline{v}_i, \underline{v}_i\right) = (\lambda_i \underline{v}_i, \underline{v}_i)$$

Now, since we have an orthonormal basis:

$$(\lambda_i \underline{v}_i, \underline{v}_i) = \lambda(\underline{v}_i, \underline{v}_i) = \lambda_i$$

Moreover:

$$\begin{aligned}\left(\left(\frac{T+T^*}{2}\right)\underline{v}_i, \underline{v}_i\right) &= \frac{1}{2}\left(T\underline{v}_i + T^*\underline{v}_i, \underline{v}_i\right) \\ &= \frac{1}{2}\left((T\underline{v}_i, \underline{v}_i) + (T^*\underline{v}_i, \underline{v}_i)\right) \\ &= \frac{1}{2}\left((T\underline{v}_i, \underline{v}_i) + (\underline{v}_i, T\underline{v}_i)\right) \\ &= \frac{1}{2}\left((T\underline{v}_i, \underline{v}_i) + (T\underline{v}_i, \underline{v}_i)\right) \\ &= (T\underline{v}_i, \underline{v}_i)\end{aligned}$$

where we have used the symmetry of a real inner product.
Thus:

$$(T\underline{v}_i, \underline{v}_i) = \lambda_i$$

(c) **What happens to the answers if V is a complex inner product space instead?**
If V were complex, symmetry wouldn't apply in the same way. Then:

$$\begin{aligned}\left(\left(\frac{T+T^*}{2}\right)\underline{v}_i, \underline{v}_i\right) &= \frac{1}{2}\left(T\underline{v}_i + T^*\underline{v}_i, \underline{v}_i\right) \\ &= \frac{1}{2}\left((T\underline{v}_i, \underline{v}_i) + (T^*\underline{v}_i, \underline{v}_i)\right) \\ &= \frac{1}{2}\left((T\underline{v}_i, \underline{v}_i) + (\underline{v}_i, T\underline{v}_i)\right) \\ &= \frac{1}{2}\left((T\underline{v}_i, \underline{v}_i) + \overline{(T\underline{v}_i, \underline{v}_i)}\right) \\ &= \frac{1}{2}\left(2\mathcal{R}[(T\underline{v}_i, \underline{v}_i)]\right) \\ &= \mathcal{R}[(T\underline{v}_i, \underline{v}_i)]\end{aligned}$$

So we'd get that:

$$\mathcal{R}[(T\underline{v}_i, \underline{v}_i)] = \lambda_i$$

Intuitively, this makes sense, since a property of self-adjoint operators is that their eigenvalues are real. (Theorem 5.3.7)

3. (a) **Let $A \in Mat(n; \mathbb{R})$ be an orthogonal matrix. Show that $det(A) \in \{\pm 1\}$.**
Since A is an orthogonal matrix, by definition:

$$A^T = A^{-1} \implies A^T A = I_n$$

So then:

$$det(A^T A) = det(I_n) = 1$$

However:

$$det(A^T A) = det(A^T)det(A) = det(A)det(A) = det(A)^2$$

Thus:

$$det(A)^2 = 1$$

Since we operate over $\mathbb{R}$ this is only possible if:

$$det(A) = \pm 1$$

as required.

(b) **Let $A \in Mat(n; \mathbb{C})$ be a unitary matrix. Show that $det(A)$ lies on the unit circle in $\mathbb{C}$.**
Since A is unitary, by definition:

$$A^{-1} = \bar{A}^T \implies \bar{A}^T A = I_n$$

So then:

$$det(\bar{A}^T A) = det(I_n) = 1$$

However:

$$det(\bar{A}^T A) = det(\bar{A}^T)det(A) = det(\bar{A})det(A) = \overline{det(A)}det(A) = \|det(A)\|^2$$

Thus:

$$\|det(A)\|^2 = 1 \implies \|det(A)\| = 1$$

which is precisely the definition of $det(A)$ lying on the unit circle.

(c) **Find a non-zero nilpotent symmetric matrix $A \in Mat(2; \mathbb{C})$. Can you find one with real entries?**

> *Here, I just used a general matrix:*
>
> $$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$
>
> *, squared it, and sought a, b, c which ensured that entries were 0. However, this won't work: if A is nilpotent, $\exists d$ such that A^d is the 0-matrix - d need not be 2 however. In this particular case, since we consider 2×2 matrices, this will work. However, the method from the solutions is much more robust, so I use that here.*

For this question, we begin by showing that if A is nilpotent, such that if $A^d = 0$, then the characteristic polynomial will be x^d:

$$A\underline{v} = \lambda \underline{v} \implies A^d \underline{v} = \lambda^d \underline{v} \implies \lambda^d = 0$$

In this case, since we have a 2×2 symmetric matrix, the characteristic polynomial will have degree 2, so we expect:

$$p_A(x) = x^2$$

if A is nilpotent. Consider:

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

This has characteristic polynomial:

$$p_A(x) = (x-a)(x-c) - b^2 = x^2 - x(a+c) + (ac - b^2)$$

Thus, we seek:

$$a = -c \qquad ac - b^2 = 0 \implies -a^2 - b^2 = 0 \implies a = \pm b\sqrt{-1}$$

And so, we can pick $a = 1, c = -1, b = \sqrt{-1}$:

$$A = \begin{pmatrix} 1 & \sqrt{-1} \\ \sqrt{-1} & -1 \end{pmatrix}$$

At this stage, I would say that the above already tells us that we can't have this work for real, non-zero matrices, but the solutions give a much more rigorous argument, which is quite neat.

Now, lets assume that we can come up with real a, b, c such that A is nilpotent and symmetric. Notice that by the Corollary of the Spectral Theorem applied to real, symmetric matrices:

*Let $A \in Mat(n; \mathbb{R})$, and let A be a **symmetric** matrix:*

$$A = A^T$$

Then:

$$\exists P \in Mat(n; \mathbb{R}) \quad : \quad P^{-1}AP = P^T AP = diag(\lambda_1, \ldots, \lambda_n)$$

where:

- *P is an **orthogonal** matrix*
- *$\lambda_1, \ldots, \lambda_n$ are the **real** eigenvalues of A (including those with repeated multiplicity)*

[Corollary 5.3.14]

A is diagonalisable, and will have the same characteristic polynomial as its conjugate matrix. The only diagonal matrix with characteristic polynomial x^2 is the 0-matrix, so A is similar to the zero matrix, and only the 0 matrix is conjugate to the 0 matrix, so A will have to be the 0 matrix.

*If V is a **vector space** over a **field** F, we define the **dual vector space** V^* to be the space of **linear mappings**:*

$$V^* = Hom_F(V, F)$$

In Exercise 15, we show that this is a vector space: if $\theta, \phi \in V^*, \underline{v} \in V, \lambda \in F$ then:

$$(\theta + \phi)(\underline{v}) = \theta(\underline{v}) + \phi(\underline{v}) \qquad (\lambda\theta)(\underline{v}) = \lambda(\theta(\underline{v}))$$

4. **Assume that V is an inner product space.**
   (a) **Let $\underline{v} \in V$. Show that the mapping:**

$$(-, \underline{v}) : V \to F$$
$$\underline{w} \to (\underline{w}, \underline{v})$$

is an element of V^*. Call it $\epsilon_{\underline{v}}$.
In simple words, this just wants us to show that $\epsilon_{\underline{v}}$ is a homomorphism of the form $V \to F$.
We check the properties. Let $\underline{a}, \underline{b} \in V$ and $\lambda \in F$. Then:

$$\varepsilon_{\underline{v}}(\underline{a} + \underline{b}) = (\underline{a} + \underline{b}, \underline{v}) = (\underline{a}, \underline{v}) + (\underline{b}, \underline{v}) = \varepsilon_{\underline{v}}(\underline{a}) + \varepsilon_{\underline{v}}(\underline{b})$$
$$\varepsilon_{\underline{v}}(\lambda \underline{a}) = (\lambda \underline{a}, \underline{v}) = \lambda(\underline{a}, \underline{v}) = \lambda \varepsilon_{\underline{v}}(\underline{a})$$

So indeed, $\varepsilon_{\underline{v}}$ is a homomorphism.

   (b) **Show that the mapping:**

$$\Delta : V \to V^*$$
$$\underline{v} \to \epsilon_{\underline{v}}$$

is injective.
If Δ is injective, then $\forall \underline{v}_1, \underline{v}_2 \in V$ if $\Delta(\underline{v}_1) = \Delta(\underline{v}_2)$ then $\underline{v}_1 = \underline{v}_2$.
If $\Delta(\underline{v}_1) = \Delta(\underline{v}_2)$, then:

$$\varepsilon_{\underline{v}_1} = \varepsilon_{\underline{v}_2} \iff \forall \underline{v} \in V,\ \varepsilon_{\underline{v}_1}(\underline{v}) = \varepsilon_{\underline{v}_2}(\underline{v})$$

This is true if and only if $\forall \underline{v} \in V$:

$$\begin{aligned} & (\underline{v}, \underline{v}_1) = (\underline{v}, \underline{v}_2) \\ \implies & \overline{(\underline{v}_1, \underline{v})} = \overline{(\underline{v}_2, \underline{v})} \\ \implies & \overline{(\underline{v}_1 - \underline{v}_2, \underline{v})} = 0 \\ \implies & (\underline{v}, \underline{v}_1 - \underline{v}_2) = 0 \end{aligned}$$

In particular, if we choose $\underline{v} = \underline{v}_1 - \underline{v}_2$ then:

$$(\underline{v}_1 - \underline{v}_2, \underline{v}_1 - \underline{v}_2) = 0 \iff \underline{v}_1 - \underline{v}_2 = 0$$

so indeed, if $\Delta(\underline{v}_1) = \Delta(\underline{v}_2)$ then $\underline{v}_1 = \underline{v}_2$, and so, Δ is injective.

   (c) **Is Δ linear?**
We consider if it satisfies the properties. Let $\underline{v}_1, \underline{v}_2 \in V, \lambda \in F$. Then, $\forall \underline{v} \in V$, consider:

$$\begin{aligned} \Delta(\underline{v}_1 + \underline{v}_2)(\underline{v}) &= \varepsilon_{\underline{v}_1 + \underline{v}_2}(\underline{v}) \\ &= (\underline{v}, \underline{v}_1 + \underline{v}_2) \\ &= \overline{(\underline{v}_1 + \underline{v}_2, \underline{v})} \\ &= \overline{(\underline{v}_1, \underline{v})} + \overline{(\underline{v}_2, \underline{v})} \\ &= (\underline{v}, \underline{v}_1) + (\underline{v}, \underline{v}_2) \\ &= \varepsilon_{\underline{v}_1}(\underline{v}) + \varepsilon_{\underline{v}_2}(\underline{v}) \\ &= \Delta(\underline{v}_1)(\underline{v}) + \Delta(\underline{v}_2)(\underline{v}) \\ &= (\Delta(\underline{v}_1) + \Delta(\underline{v}_2))(\underline{v}) \end{aligned}$$

$$\begin{aligned}\Delta(\lambda \underline{v}_1)(\underline{v}) &= \varepsilon_{\lambda \underline{v}_1}(\underline{v}) \\ &= (\underline{v}, \lambda \underline{v}_1) \\ &= \overline{(\lambda \underline{v}_1, \underline{v})} \\ &= \overline{\lambda(\underline{v}_1, \underline{v})} \\ &= \bar{\lambda}(\underline{v}, \underline{v}_1) \\ &= \bar{\lambda}\varepsilon_{\underline{v}_1}(\underline{v}) \\ &= (\bar{\lambda}\Delta(\underline{v}_1))(\underline{v})\end{aligned}$$

Hence, this means that if the underlying field is $\mathbb{R}$, Δ will be linear; if however it is complex, it won't be linear.

5. **Assume that V is a finite dimensional F-vector space.**
   (a) **Let $\mathcal{A} = (\underline{v}_1, \ldots, \underline{v}_n)$ be an ordered basis of V. Show that $(\theta_1, \ldots, \theta_n))$ is an ordered basis of V^*, where I define:**

$$\theta_i : V \to F$$

$$\theta_i \left(\sum_{j=1}^{n} \lambda_j \underline{v}_j \right) = \lambda_i$$

This is called the *dual basis* to $\mathcal{A}$.

We first show that the set is linearly independent. Indeed, let $\alpha_i \in F$ such that $\forall \underline{v} \in V$ we have:

$$\left(\sum_{i=1}^{n} \alpha_i \theta_i \right) (\underline{v}) = 0$$

so by linearity:

$$\sum_{i=1}^{n} \alpha_i \theta_i(\underline{v}) = 0$$

In particular, if we pick $\underline{v} = \underline{v}_j$:

$$\sum_{i=1}^{n} \alpha_i \theta_i(\underline{v}_j) = 0$$

But by definition of θ_i, we have that:

$$\theta_i(\underline{v}_j) = \delta_{ij}$$

Thus:

$$\sum_{i=1}^{n} \alpha_i \theta_i(\underline{v}_j) = \alpha_j = 0$$

Hence, $\forall j \in [1, n]$ we have that $\alpha_j = 0$, and so, the θ_i are linearly independent.

Now we check if the θ_i are spanning. Consider $\phi \in V^*$ and let $\underline{v} \in V$ such that:

$$\underline{v} = \sum_{j=1}^{n} \alpha_j \underline{v_j}$$

Then:

$$\theta(\underline{v}) = \theta\left(\sum_{j=1}^{n} \alpha_j \underline{v_j}\right) = \sum_{j=1}^{n} \alpha_j \theta(\underline{v_j})$$

Now, consider the linear combination $\sum_{i=1}^{n} \lambda_i \theta_i$. Then:

$$\begin{aligned}\sum_{i=1}^{n} \lambda_i \theta_i(\underline{v}) &= \sum_{i=1}^{n} \lambda_i \theta_i \left(\sum_{j=1}^{n} \alpha_j \underline{v_j}\right) \\ &= \sum_{i=1}^{n} \lambda_i \theta_i \left(\sum_{j=1}^{n} \alpha_j \underline{v_j}\right) \\ &= \sum_{i=1}^{n} \lambda_i \theta_i \left(\sum_{j=1}^{n} \alpha_j \theta_i(\underline{v_j})\right) \\ &= \sum_{i=1}^{n} \lambda_i \alpha_i\end{aligned}$$

In particular, if we let $\lambda_i = \theta_i(\underline{v}_i)$, we get that $\forall \underline{v} \in V$:

$$\theta(\underline{v}) = \left(\sum_{i=1}^{n} \lambda_i \theta_i\right)(\underline{v})$$

so the θ_i are spanning.

(b) **Let W be a finite dimensional F-vector space and $f : V \to W$ a linear mapping. Show that the mapping $f^* : W^* \to V^*$ defined by:**

$$f^*(\theta)(\underline{v}) = \theta(f(\underline{v})), \qquad \theta \in W^*, \underline{v} \in V$$

is linear. This is called the *dual mapping* to f.

We can note (and as is said in the solutions) that $f^* = \theta \circ f$. Since θ, f are homomorphisms, then f^* will be a homomorphism. But this is less fun.

We check the properties of homomorphism. Let $\theta, \psi \in W^*$ and $\lambda \in F$. Then, $\forall \underline{v} \in V$:

$$f^*(\theta + \psi)(\underline{v}) = (\theta + \psi)(f(\underline{v})) = \theta(f(\underline{v})) + \psi(f(\underline{v})) = f^*(\theta)(\underline{v}) + f^*(\psi)(\underline{v}) = (f^*(\theta) + f^*(\psi))(\underline{v})$$

$$f^*(\lambda\theta)(\underline{v}) = (\lambda\theta)(f(\underline{v})) = \lambda(\theta(f(\underline{v})) = (\lambda f^*)(\underline{v})$$

Thus, f^* is linear.

(c) **Let $\mathcal{A} = (\underline{v}_1, \ldots, \underline{v}_n)$ and $\mathcal{B} = (\underline{w}_1, \ldots, \underline{w}_m)$ be ordered bases of V, W respectively. Let $\mathcal{A}^*, \mathcal{B}^*$ be the dual bases of V^*, W^*. Show that:**

$$_{\mathcal{A}^*}[f^*]_{\mathcal{B}^*} = \left(_{\mathcal{B}}[f]_{\mathcal{A}}\right)^T$$

Let:

$$\mathcal{A}^* = (\theta_1, \ldots, \theta_n)$$

$$\mathcal{B}^* = (\psi_1, \ldots, \psi_m)$$

Then, the jth column of ${}_{\mathcal{A}^*}[f^*]_{\mathcal{B}^*}$ is given by the vector $(a_{ij})_{i\in[1,n]}$, where:

$$f^*(\psi_j) = \sum_{i=1}^{n} a_{ij}\theta_i$$

Similarly, the kth column of ${}_{\mathcal{B}}[f]_{\mathcal{A}}$ is given by the vector $(b_{tk})_{t\in[1,m]}$, where:

$$f(\underline{v}_k) = \sum_{t=1}^{m} b_{tk}\underline{w}_t$$

Now, consider applying $f^*(\psi_j)$ to $\underline{v}_k, k \in [1, n]$. Then:

$$f^*(\psi_j)(\underline{v}_k) = \psi_j(f(\underline{v}_k)) = \psi_j\left(\sum_{t=1}^{m} b_{tk}\underline{w}_t\right) = b_{jk}$$

Similarly:

$$\left(\sum_{i=1}^{n} a_{ij}\theta_i\right)(\underline{v}_k) = \sum_{i=1}^{n} a_{ij}\theta_i(\underline{v}_k) = \sum_{i=1}^{n} a_{ij}\delta_{ik} = a_{kj}$$

Hence, we get that the matrices are such that $a_{kj} = b_{jk}$, so it follows that:

$${}_{\mathcal{A}^*}[f^*]_{\mathcal{B}^*} = ({}_{\mathcal{B}}[f]_{\mathcal{A}})^T$$

as required.

All of the following is directly from solutions, since this seemed like such a bizarre and unnecessary question.

6. (a) **Let V be a one-dimensional vector space. Pick a non-zero vector, call it $\underline{v}$. This gives a basis for V, so you can apply 4)a) to get a dual basis vector, $\theta \in V^*$. Show that the mapping $V \to V^*$:**

$$\lambda\underline{v} \to \lambda\theta$$

is an isomorphism.

If we pick $V = \mathbb{R}$, then we can let $\underline{v} = 1$. Then:

$$\theta(\lambda) = \theta(\lambda \cdot 1) = \lambda\theta(1) = \lambda$$

So θ is the identity mapping.

(b) **Do the above question again, this time choosing a different non-zero vector. Is your isomorphism exactly the same as before?**

If we pick $\underline{v} = 2$, then:

$$\theta(\lambda) = \theta(0.5\lambda \cdot 2) = \frac{\lambda}{2}\theta(2) = \frac{\lambda}{2}$$

So θ is the operation of halving.

Thus, we don't get the same isomorphism.

(c) **Now, let V be an arbitrary finite dimensional vector space. Without ever picking a basis, construct an explicit isomorphism $V \to (V^*)^*$**

$(V^*)^*$ is the set of homomorphisms of the form $V^* \to F$. Define a mapping:

$$f : V \to (V^*)^*$$

via:

$$f(\underline{v}) = f_{\underline{v}}$$

where:

$$f_{\underline{v}}(\theta) = \theta(\underline{v})$$

We begin by showing that this is a homomorphism:

$$f_{\underline{v}+\underline{w}}(\theta) = \theta(\underline{v} + \underline{w}) = \theta(\underline{v}) + \theta(\underline{w}) = f_{\underline{v}}(\theta) + f_{\underline{w}}(\theta) = (f_{\underline{v}} + f_{\underline{w}})(\theta)$$

$$f_{\lambda\underline{v}}(\theta) = \theta(\lambda v) = \lambda\theta(\underline{v}) = \lambda f_{\underline{v}}(\theta) = (\lambda f_{\underline{v}})(\theta)$$

Moreover, it is injective. Assume that $f_{\underline{v}} = f_{\underline{w}}$. Then, $\forall\theta \in V^*$:

$$\theta(\underline{v}) = \theta(\underline{w}) \iff \theta(\underline{v} - \underline{w}) = 0$$

If we take θ as the identity mapping, this is true only when $\underline{v} = \underline{w}$, and so, the mapping is injective. Now, since $dim(V) = dim(V^*) = dim((V^*)^*)$, this injective homomorphism must be surjective, and so, it is an isomorphism.