Self-adjoint operators have orthonormal eigenbases

1. Let T be self adjoint, and let $\vec{v}$ be a nonzero eigenvector of T. Then T maps $\vec{v}^{\perp}$ to itself.
2. Let T be self adjoint, and let $\vec{u}$ and $\vec{v}$ be nonzero eigenvectors of T with distinct eigenvalues α and β. Then $\langle \vec{u}, \vec{v} \rangle = 0$.
3. If T has an orthonormal eigenbasis then T is self-adoint.

1. Let T be self adjoint, and let $\vec{v}$ be a nonzero eigenvector of T. Then T maps $\vec{v}^{\perp}$ to itself.

   Proof: Let $\vec{w}$ be orthogonal to $\vec{v}$. Then $\langle \vec{v}, T(\vec{w})\rangle = \langle T(\vec{v}), \vec{w}\rangle$ $= \langle \lambda\vec{v}, \vec{w}\rangle = \lambda\langle \vec{v}, \vec{w}\rangle = 0$. So $T(\vec{w})$ is also orthogonal to $\vec{v}$.

2. Let T be self adjoint, and let $\vec{u}$ and $\vec{v}$ be nonzero eigenvectors of T with distinct eigenvalues α and β. Then $\langle \vec{u}, \vec{v}\rangle = 0$.

   Proof: We have $\langle T(\vec{v}), \vec{w}\rangle = \langle \vec{v}, T(\vec{w})\rangle$ so $\alpha\langle \vec{v}, \vec{w}\rangle = \beta\langle \vec{v}, \vec{w}\rangle$ and we deduce that $\langle \vec{v}, \vec{w}\rangle = 0$.

3. If T has an orthonormal eigenbasis then T is self-adoint.

   Proof: Let $\vec{v}_1$, $\vec{v}_2$, $\ldots$, $\vec{v}_n$ be the orthonormal eigenbasis with $T(\vec{v}_i) = \lambda_i\vec{v}_i$. Let $\vec{x} = \sum a_i\vec{v}_i$ and $\vec{y} = \sum b_i\vec{v}_i$. Then
   $\langle T(\vec{x}), \vec{y}\rangle = \langle \sum_i a_i\lambda_i\vec{v}_i,\ \sum_j b_j\vec{v}_j\rangle = \sum a_i\lambda_i b_i$ and
   $\langle \vec{x}, T(\vec{y})\rangle = \langle \sum_i a_i\vec{v}_i,\ \sum_j b_j\lambda_j\vec{v}_j\rangle = \sum a_i b_i\lambda_i$ as well.

We now have all the tools necessary to prove the converse of the last statement:

Theorem: Let V be a finite dimensional real vector space with an inner product $\langle\ ,\ \rangle$ and let $T : V \to V$ be a self-adjoint linear transformation. Then T has an orthonormal eigenbasis.

Our main task will be to prove that T has one real eigenvector $\vec{v}$. Once we do this, we will know that T takes $\vec{v}^{\perp} \longrightarrow \vec{v}^{\perp}$ and, by induction, $T|_{\vec{v}^{\perp}}$ will have an orthonormal basis of eigenvectors.

T has a real eigenvector – first proof

The characteristic polynomial $\chi_T(x)$ must have roots in the complex numbers. Let $a + bi$ be one of these roots. We will show that $b = 0$.

T has a real eigenvector – first proof

The characteristic polynomial $\chi_T(x)$ must have roots in the complex numbers. Let $a + bi$ be one of these roots. We will show that $b = 0$.

If not, then there is some (nonzero) complex vector with eigenvalue $a + bi$. Let it be $\vec{x} + i\vec{y}$ with $\vec{x}$ and $\vec{y}$ real. Then $\vec{x} - i\vec{y}$ is an complex eigenvector with eigenvalue $a - bi$.

T has a real eigenvector – first proof

The characteristic polynomial $\chi_T(x)$ must have roots in the complex numbers. Let $a + bi$ be one of these roots. We will show that $b = 0$.

If not, then there is some (nonzero) complex vector with eigenvalue $a + bi$. Let it be $\vec{x} + i\vec{y}$ with $\vec{x}$ and $\vec{y}$ real. Then $\vec{x} - i\vec{y}$ is an complex eigenvector with eigenvalue $a - bi$.

Since $a + bi \neq a - bi$, the vectors $\vec{x} + i\vec{y}$ and $\vec{x} - i\vec{y}$ are orthogonal. But

$$\langle \vec{x}+i\vec{y}, \vec{x}-i\vec{y} \rangle = \langle \vec{x}, \vec{x} \rangle - i\langle \vec{x}, \vec{y} \rangle + i\langle \vec{y}, \vec{x} \rangle + \langle \vec{y}, \vec{y} \rangle = \langle \vec{x}, \vec{x} \rangle + \langle \vec{y}, \vec{y} \rangle > 0.$$

Contradiction, $\Box$.

T has a real eigenvector – second proof

Let S be the unit sphere $\{\vec{v} : \langle \vec{v}, \vec{v} \rangle = 1\}$. There is some $\vec{v} \in S$ where $\langle T(\vec{v}), \vec{v} \rangle$ is maximized. We will show that this $\vec{v}$ is an eigenvector.

T has a real eigenvector – second proof

Let S be the unit sphere $\{\vec{v} : \langle \vec{v}, \vec{v} \rangle = 1\}$. There is some $\vec{v} \in S$ where $\langle T(\vec{v}), \vec{v} \rangle$ is maximized. We will show that this $\vec{v}$ is an eigenvector.

Suppose that $T(\vec{v})$ is not in $\mathbb{R}\vec{v}$. Then $\vec{v}$ and $T(\vec{v})$ span a 2-plane. Let $\vec{v}$ and $\vec{w}$ be an orthonormal basis for this plane, with $\langle T(\vec{v}), \vec{w} \rangle > 0$.

T has a real eigenvector – second proof

Let S be the unit sphere $\{\vec{v} : \langle \vec{v}, \vec{v} \rangle = 1\}$. There is some $\vec{v} \in S$ where $\langle T(\vec{v}), \vec{v} \rangle$ is maximized. We will show that this $\vec{v}$ is an eigenvector.

Suppose that $T(\vec{v})$ is not in $\mathbb{R}\vec{v}$. Then $\vec{v}$ and $T(\vec{v})$ span a 2-plane. Let $\vec{v}$ and $\vec{w}$ be an orthonormal basis for this plane, with $\langle T(\vec{v}), \vec{w} \rangle > 0$.

Consider the circle $(\cos\theta)\vec{v} + (\sin\theta)\vec{w}$. Every vector on this circle is on S. For such a vector, we have

$$\Big\langle T\left((\cos\theta)\vec{v} + (\sin\theta)\vec{w}\right),\ (\cos\theta)\vec{v} + (\sin\theta)\vec{w} \Big\rangle =$$

$$\cos^2\theta \langle T(\vec{v}), \vec{v} \rangle + \cos\theta\sin\theta \Big(\langle T(\vec{v}), \vec{w} \rangle + \langle \vec{v}, T(\vec{w}) \rangle \Big) + \sin^2\theta \langle T(\vec{w}), \vec{w} \rangle$$

$$(\cos^2\theta)\langle T(\vec{v}), \vec{v} \rangle + (2\cos\theta\sin\theta)\langle T(\vec{v}), \vec{w} \rangle + (\sin^2\theta)\langle T(\vec{w}), \vec{w} \rangle$$

T has a real eigenvector – second proof

$$\left\langle T\left((\cos\theta)\vec{v} + (\sin\theta)\vec{w}\right),\ (\cos\theta)\vec{v} + (\sin\theta)\vec{w}\right\rangle =$$

$$\cos^2\theta\langle T(\vec{v}),\vec{v}\rangle + \cos\theta\sin\theta\Big(\langle T(\vec{v}),\vec{w}\rangle + \langle\vec{v},T(\vec{w})\rangle\Big) + \sin^2\theta\langle T(\vec{w}),\vec{w}\rangle$$

$$(\cos^2\theta)\langle T(\vec{v}),\vec{v}\rangle + (2\cos\theta\sin\theta)\langle T(\vec{v}),\vec{w}\rangle + (\sin^2\theta)\langle T(\vec{w}),\vec{w}\rangle$$

So

$$\frac{d}{d\theta}\Big|_{\theta=0}\left\langle T\left((\cos\theta)\vec{v} + (\sin\theta)\vec{w}\right),\ (\cos\theta)\vec{v}+(\sin\theta)\vec{w}\right\rangle = \langle T(\vec{v}),\vec{w}\rangle > 0.$$

This contradicts that $\langle T(\vec{x}),\vec{x}\rangle$ is supposed to be maximized at $\vec{x} = \vec{v}$. $\square$