

Orthogonal Operators

Def. Let T be a linear operator on a real finite dimensional inner product space V . If $\|T(v)\| = \|v\|$ for all $v \in V$, we call T an **orthogonal operator**.

So an orthogonal linear operator T preserves lengths, i.e. $\|T(v)\| = \|v\|$.

If $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is a rotation or a reflection about a line through $(0,0)$, then T is orthogonal as it preserves lengths.

With respect to the standard ordered basis, β , in \mathbb{R}^2 :

$$[T]_{\beta} = A: \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad \text{where } A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Is a rotation of the plane by θ and is thus orthogonal.

$$[T]_{\beta} = A: \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad \text{where } A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Is a reflection about the x -axis and hence orthogonal.

Theorem: Let T be a linear operator on a real, finite dimensional, inner product space V . Then the following statements are equivalent.

- a. $TT^* = T^*T = I$
- b. $\langle T(v), T(w) \rangle = \langle v, w \rangle$ for all $v, w \in V$
- c. If β is an orthonormal basis for V , then so is $T(\beta)$.
- d. There exists an orthonormal basis for V such that $T(\beta)$ is an orthonormal basis for V
- e. $\|T(v)\| = \|v\|$ for all $v \in V$

Since the above 5 statements are equivalent, we could take any of them as the definition of an orthogonal operator.

To prove this theorem we start with the following lemma:

Lemma: Let S be a self-adjoint operator on a real finite dimensional inner product space V . If $\langle v, S(v) \rangle = 0$ for all $v \in V$, then $S = 0$ (the linear transformation that maps all vectors in V to zero).

Proof: By an earlier theorem we know that there exists an orthonormal basis β of V consisting of eigenvectors of S :

$$\beta = \{u_1, \dots, u_n\}, \quad \text{where } u_i \text{ is an eigenvector of } S.$$

Thus, $S(u_i) = \lambda_i u_i$, $1 \leq i \leq n$.

$$\begin{aligned} \text{By assumption: } 0 &= \langle u_i, S(u_i) \rangle \\ &= \langle u_i, \lambda_i u_i \rangle \\ &= \lambda_i \langle u_i, u_i \rangle \\ &= \lambda_i \quad (\text{since } \langle u_i, u_i \rangle = 1). \end{aligned}$$

Hence $S(u_i) = \lambda_i u_i = 0$, $1 \leq i \leq n$.

Since S maps all of the basis vectors to the zero vector, it must map all vectors in V to the zero vector. So $S = 0$.

Proof of theorem: We will show $a \Rightarrow b \Rightarrow c \Rightarrow d \Rightarrow e \Rightarrow a$.

$a \Rightarrow b$: We assume $TT^* = T^*T = I$ and prove

$\langle T(v), T(w) \rangle = \langle v, w \rangle$ for all $v, w \in V$.

$$\begin{aligned} \langle v, w \rangle &= \langle T^*T v, w \rangle \quad (\text{since } TT^* = I) \\ &= \langle T(v), T(w) \rangle. \end{aligned}$$

$b \Rightarrow c$: We assume $\langle T(v), T(w) \rangle = \langle v, w \rangle$ for all $v, w \in V$ and show If β is an orthonormal basis for V , then so is $T(\beta)$.

Let $\beta = \{u_1, \dots, u_n\}$ be an orthonormal basis for V . Then we have:

$$T(\beta) = \{T(u_1), \dots, T(u_n)\}.$$

Since $\langle T(v), T(w) \rangle = \langle v, w \rangle$ we have:

$$\langle T(u_i), T(u_j) \rangle = \langle u_i, u_j \rangle = \delta_{ij}.$$

Therefore, $T(\beta)$ is an orthonormal basis for V .

$c \Rightarrow d$: We assume If β is an orthonormal basis for V , then so is $T(\beta)$ and we show there exists an orthonormal basis for V such that $T(\beta)$ is an orthonormal basis for V .

We know that for any finite dimensional real inner product space V that there exists an orthonormal basis (from the Gram-Schmidt orthonormalization process). Therefore we get d from c .

$d \Rightarrow e$: We assume there exists an orthonormal basis for V such that $T(\beta)$ is an orthonormal basis for V , and we show $\|T(v)\| = \|v\|$ for all $v \in V$.

Let $v \in V$ and $\beta = \{u_1, \dots, u_n\}$ an orthonormal basis for V .

Then we can write: $v = \sum_{i=1}^n a_i u_i$. So we have

$$\begin{aligned} \|v\|^2 &= \langle v, v \rangle = \langle \sum_{i=1}^n a_i u_i, \sum_{j=1}^n a_j u_j \rangle \\ &= \sum_{i=1}^n \sum_{j=1}^n a_i a_j \langle u_i, u_j \rangle \\ &= \sum_{i=1}^n \sum_{j=1}^n a_i a_j \delta_{ij} \quad (\text{because } \beta \text{ is orthonormal}) \\ &= \sum_{i=1}^n |a_i|^2. \end{aligned}$$

Similarly, since $T(\beta) = \{T(u_1), \dots, T(u_n)\}$ is an orthonormal basis for V :

$$\begin{aligned} T(v) &= T\left(\sum_{i=1}^n a_i u_i\right) \\ &= \sum_{i=1}^n a_i T(u_i). \end{aligned}$$

Thus we have:

$$\begin{aligned} \|T(v)\|^2 &= \langle T(v), T(v) \rangle = \left\langle \sum_{i=1}^n a_i T(u_i), \sum_{j=1}^n a_j T(u_j) \right\rangle \\ &= \sum_{i=1}^n \sum_{j=1}^n a_i a_j \delta_{ij} \quad (T(\beta) \text{ is orthonormal}) \\ &= \sum_{i=1}^n |a_i|^2 = \|v\|^2. \end{aligned}$$

Thus $\|T(v)\| = \|v\|$ for all $v \in V$.

$e \Rightarrow a$: We assume $\|T(v)\| = \|v\|$ for all $v \in V$, and we show
 $TT^* = T^*T = I$.

Let $v \in V$ then

$$\begin{aligned} \langle v, v \rangle &= \|v\|^2 = \|T(v)\|^2 \\ &= \langle T(v), T(v) \rangle \\ &= \langle v, T^*T(v) \rangle. \end{aligned}$$

So we have: $\langle v, v \rangle - \langle v, T^*T(v) \rangle = 0$

$$\langle v, (I - T^*T)(v) \rangle = 0 \quad (\text{for all } v \in V).$$

But our lemma says that if $S = I - T^*T$, then

$$0 = \langle v, S(v) \rangle \text{ for all } v \in V \text{ then } S = 0.$$

Thus $I - T^*T = 0 \Rightarrow T^*T = I$.

Since V is a finite dimensional vector space we know that:

$$T^*T = I \Rightarrow TT^* = I.$$

Corollary: Let T be a linear operator on a real, finite dimensional, inner product space V . Then V has an orthonormal basis of eigenvectors of T with corresponding eigenvalues of absolute value one if and only if T is self-adjoint and orthogonal.

Proof. Suppose V has an orthonormal basis $\{u_1, \dots, u_n\}$ such that:

$$T(u_i) = \lambda_i u_i, \quad \text{and} \quad |\lambda_i| = 1, \quad 1 \leq i \leq n.$$

We need to show that T is self-adjoint and orthogonal.

We know by an earlier theorem that T is self-adjoint.

To show that T is orthogonal:

$$\begin{aligned} (TT^*)(u_i) &= T(\lambda_i u_i) \\ &= \lambda_i T(u_i) \\ &= \lambda_i \lambda_i u_i \\ &= \lambda_i^2 u_i \\ &= u_i \quad (\text{since } |\lambda_i| = 1). \end{aligned}$$

So $TT^* = I$ for all of the basis vector and hence for all of V .

Since V is finite dimensional, $TT^* = I \implies T^*T = I$.

So T is orthogonal.

Now assume that T is self-adjoint and orthogonal and let's show that V has an orthonormal basis $\{u_1, \dots, u_n\}$ of eigenvectors of T with corresponding eigenvalues of absolute value one.

By an earlier theorem since T is self-adjoint, V has an orthonormal basis of eigenvectors $\{u_1, \dots, u_n\}$ of T . Thus $T(u_i) = \lambda_i u_i$, $1 \leq i \leq n$.

So we have we have:

$$\begin{aligned} |\lambda_i| \|u_i\| &= \|\lambda_i u_i\| \\ &= \|T(u_i)\| \\ &= \|u_i\| \quad (\text{since } T \text{ is orthogonal}), \end{aligned}$$

So $|\lambda_i| = 1$, for $1 \leq i \leq n$.

Def. A square matrix A with real entries is called **orthogonal** if

$$AA^t = A^t A = I.$$

Ex. Show $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, a rotation by θ , is orthogonal but not self-adjoint (and therefore does not have an orthonormal basis of eigenvectors of T)

With respect to the standard ordered basis, β , on \mathbb{R}^2 ,

$$[T]_\beta = A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

So A^t is: $A^t = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}.$

$$\begin{aligned}
 AA^t &= \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \\
 &= \begin{pmatrix} \cos^2\theta + \sin^2\theta & 0 \\ 0 & \cos^2\theta + \sin^2\theta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.
 \end{aligned}$$

Similarly, $A^tA = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

So T (and A) are orthogonal.

However, $A^t \neq A$, so A is not self-adjoint.

Ex. Show the reflection about the x -axis given by $[T]_{\beta} = A: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, where $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, and β is the standard ordered basis for \mathbb{R}^2 , is both self-adjoint and orthogonal.

$$A^t = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = A, \text{ so } A \text{ (and } T) \text{ are self-adjoint.}$$

$$A^tA = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = AA^t.$$

So A (and T) are orthogonal.

By our theorem \mathbb{R}^2 has an orthonormal basis of eigenvectors of T (in the above example).

However, A is already diagonal with respect to the standard basis, thus $\beta = \{ \langle 1, 0 \rangle, \langle 0, 1 \rangle \}$, is an orthonormal basis of \mathbb{R}^2 with the eigenvectors of T (and A).

Notice that for an orthogonal matrix A , the rows and the columns each form a set of orthonormal vectors. We can see this for 2x2 matrices as follows:

$$AA^t = I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad A^t = \begin{pmatrix} a & c \\ b & d \end{pmatrix}.$$

$$AA^t = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a^2 + b^2 & ac + bd \\ ac + bd & c^2 + d^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Equating the components of the matrices we get:

$$1 = a^2 + b^2 = \|\langle a, b \rangle\|^2 \quad \Rightarrow \quad \|\langle a, b \rangle\| = 1$$

$$1 = c^2 + d^2 = \|\langle c, d \rangle\|^2 \quad \Rightarrow \quad \|\langle c, d \rangle\| = 1$$

$$0 = ac + bd = \langle \langle a, b \rangle, \langle c, d \rangle \rangle.$$

The equations above say that the row vectors of A form an orthonormal set.

A similar result for the column vectors of A follows from $A^t A = I$.

Earlier we saw that the matrix $A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ was orthogonal.

Notice that the row vectors: $\langle \cos\theta, -\sin\theta \rangle, \langle \sin\theta, \cos\theta \rangle$, form an orthonormal set and the column vectors:

$\langle \cos\theta, \sin\theta \rangle, \langle -\sin\theta, \cos\theta \rangle$, form an orthonormal set.

Def. We say a matrix A is orthogonally equivalent to a diagonal matrix D if there exists an orthogonal matrix P such that $A = P^{-1}DP = P^tDP$. (for an orthogonal matrix $PP^t = I$, thus $P^t = P^{-1}$).

Theorem: Let A be a real $n \times n$ matrix. Then A is symmetric (ie self-adjoint) if and only if A is orthogonally equivalent to a real diagonal matrix.

Proof: We already know that if A is a symmetric matrix (i.e. self-adjoint) that there exists an orthonormal basis β of V made up of eigenvectors of A .

The matrix P whose columns are the vectors in β gives us $D = P^{-1}AP$, where D is diagonal and has real numbers (eigenvalues) along the diagonal.

But since the columns of P are orthonormal, it follows that P is orthogonal (ie $P^t = P^{-1}$).

Thus if A is symmetric then it is orthogonally equivalent to a real diagonal matrix.

Now let's assume that A is orthogonally equivalent to a diagonal matrix D and show that A is symmetric.

$$A = P^t D P$$

$$A^t = (P^t D P)^t = P^t D^t P$$

But D is a diagonal matrix so $D^t = D$. So we have:

$$A^t = P^t D P$$

$$= A.$$

So A is symmetric.

Ex. Let $A = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{pmatrix}$. Find an orthogonal matrix P that diagonalizes A ,
ie $P^t A P = D$.

Start by finding the eigenvalues of A .

$$\begin{aligned} \text{Det}(A - \lambda I) &= \det \begin{pmatrix} 4 - \lambda & 2 & 2 \\ 2 & 4 - \lambda & 2 \\ 2 & 2 & 4 - \lambda \end{pmatrix} \\ &= (4 - \lambda)[(4 - \lambda)(4 - \lambda) - 4] - 2[2(4 - \lambda) - 4] + 2[4 - 2(4 - \lambda)] \\ &= (4 - \lambda)[\lambda^2 - 8\lambda + 12] - 8 + 4\lambda - 8 + 4\lambda \\ &= (4 - \lambda)(\lambda - 2)(\lambda - 6) + 8(\lambda - 2) \\ &= (\lambda - 2)[(4 - \lambda)(\lambda - 6) + 8] \\ &= -(\lambda - 2)^2(\lambda - 8). \end{aligned}$$

So the eigenvalues of A are $\lambda = 2, 8$.

Now find an orthonormal basis for the eigenspaces.

$$\lambda_1 = 2: \quad A - 2I = \begin{pmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{pmatrix}; \quad \text{so we solve:}$$

$$\begin{pmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \Rightarrow \quad a_1 = -a_2 - a_3.$$

Thus the eigenspace associated with $\lambda_1 = 2$ is vectors of the form:

$$\langle -a_2 - a_3, a_2, a_3 \rangle = a_2 \langle -1, 1, 0 \rangle + a_3 \langle -1, 0, 1 \rangle.$$

So $\{\langle -1, 1, 0 \rangle, \langle -1, 0, 1 \rangle\}$ span the eigenspace and form a basis.

However, these vectors are neither unit length nor orthogonal.

We can take the first vector and use the Gram-Schmidt process to find a second vector that spans the same space, but are orthogonal. Doing this we

$$\text{find: } v_1 = \langle -1, 1, 0 \rangle, \quad v_2 = \langle 1, 1, -2 \rangle.$$

Dividing each vector by its length we get an orthonormal basis for the eigenspace associated with $\lambda_1 = 2$:

$$u_1 = \frac{1}{\sqrt{2}} \langle -1, 1, 0 \rangle, \quad u_2 = \frac{1}{\sqrt{6}} \langle 1, 1, -2 \rangle.$$

$$\lambda_2 = 8: \quad A - 8I = \begin{pmatrix} -4 & 2 & 2 \\ 2 & -4 & 2 \\ 2 & 2 & -4 \end{pmatrix}; \quad \text{so we solve:}$$

$$\begin{pmatrix} -4 & 2 & 2 \\ 2 & -4 & 2 \\ 2 & 2 & -4 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \Rightarrow \quad a_1 = a_2 = a_3.$$

Thus the eigenspace associated with $\lambda_2 = 8$ is vectors of the form:

$$\langle a_1, a_1, a_1 \rangle = a_1 \langle 1, 1, 1 \rangle.$$

Notice the $\langle 1, 1, 1 \rangle$ is already orthogonal to u_1, u_2 (eigenvectors from different eigenvalues are always orthogonal).

Dividing by the length of $v_3 = \langle 1, 1, 1 \rangle$, we get an orthonormal basis for \mathbb{R}^3 :

$$u_1 = \frac{1}{\sqrt{2}} \langle -1, 1, 0 \rangle, \quad u_2 = \frac{1}{\sqrt{6}} \langle 1, 1, -2 \rangle, \quad u_3 = \frac{1}{\sqrt{3}} \langle 1, 1, 1 \rangle.$$

u_1, u_2, u_3 is an orthonormal basis of eigenvectors of A , so a matrix that diagonalizes A is:

$$P = \begin{pmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix}; \quad \text{and}$$

$$P^t A P = P^{-1} A P = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 8 \end{pmatrix},$$

Where P is orthogonal because its columns are orthonormal.