Normal and Self-Adjoint Operators

In the section called "Eigenvalues and Eigenvectors" we saw that for a finite dimensional vector space V, a linear operator $T: V \to V$ was diagonalizable if and only if there exists an ordered basis $\beta = \{v_1, \dots v_n\}$ for V of eigenvectors of T. Thus it is natural to ask under what conditions does a finite dimensional inner product space V over $\mathbb{R}$ have an orthonormal basis of eigenvectors?

To determine the conditions for V to have an orthonormal basis of eigenvectors of a linear operator T we start with the following lemma.

Lemma: Let $T: V \to V$ be a linear operator on a real finite dimensional inner product space. If T has an eigenvector then so does T^*.

Proof: Suppose $v \in V$ is an eigenvector of T corresponding to the eigenvalue λ.

Then for any $w \in V$ we have:

$$0 =< 0, w >=< (T - \lambda I)(v),\ w >$$

$$=< v,\ (T - \lambda I)^*(w) >$$

$$=< v,\ (T^* - \lambda I)(w) >.$$

So v is perpendicular to the null space of $T^* - \lambda I$.

So $T^* - \lambda I$ is not onto and hence not 1-1.

Hence $T^* - \lambda I$ has a nonzero null space.

Notice that any nonzero vector in the null space of $T^* - \lambda I$ is an eigenvector of T^* corresponding to λ.

A proof of the following theorem can be done by induction on the dimension of V.

Theorem (Schur's Theorem): Let $T: V \to V$ be a linear operator on a real finite dimensional inner product space V. Suppose that the characteristic polynomial of T splits (i.e., $\det(A = \lambda I) = c(\lambda - a_1) \cdots (\lambda - a_n)$, where $A = [T]$). Then there exists an orthonormal basis β of V such that $[T]_\beta$ is upper triangular.

Our goal is to find an orthonormal basis β for the real inner product space V so that for the linear operator $T: V \to V$, $[T]_\beta$ is diagonal. Notice that if there is an orthonormal basis β such that $[T]_\beta$ is diagonal, then since $[T^*]_\beta = [T]^*{}_\beta$, (from a theorem in the previous section), $[T^*]_\beta$ is also diagonal. In addition, since diagonal matrices commute, T and T^* commute, i.e. $TT^* = T^*T$.

Def. Let V be a real inner product space and $T: V \to V$ a linear operator. We say T is **normal** $TT^* = T^*T$. An nxn real matrix A is called **normal** if $AA^* = A^*A$.

Recall that a real matrix A is called **symmetric** if $A = A^t = A^*$. A real matrix A is called **skew symmetric** if $A = -A^t = -A^*$.

Ex. Show that if a real matrix A is skew symmetric then A is normal.

Since $A = -A^*$ we have: $AA^* = -A^2$, and $A^*A = -A^2$ so

$AA^* = A^*A$, and A is normal.

Ex. Let $T: \mathbb{R}^2 \to \mathbb{R}^2$ be a rotation by θ, $0 < \theta < \pi$. If β is the standard ordered basis then $[T]_\beta = A = \begin{pmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{pmatrix}$. Show that A is normal.

Notice that:

$$AA^* = \begin{pmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{pmatrix}\begin{pmatrix} cos\theta & sin\theta \\ -sin\theta & cos\theta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$A^*A = \begin{pmatrix} cos\theta & sin\theta \\ -sin\theta & cos\theta \end{pmatrix}\begin{pmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

So $AA^* = A^*A = I$, hence A is normal.

Notice that $\begin{pmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{pmatrix}$; $0 < \theta < \pi$, has no eigenvectors since:

$$\det(A - \lambda I) = \det \begin{pmatrix} cos\theta - \lambda & -sin\theta \\ sin\theta & cos\theta - \lambda \end{pmatrix}$$

$$= (cos\theta - \lambda)^2 + \sin^2\theta$$

$$= \lambda^2 - 2\lambda cos\theta + 1.$$

Using the quadratic formula to solve $\lambda^2 - 2\lambda cos\theta + 1 = 0$, we get:

$$\lambda = \frac{2cos\theta \pm \sqrt{4\cos^2\theta - 4}}{2}$$

$$= cos\theta \pm \sqrt{\cos^2\theta - 1}$$

$$= cos\theta \pm \sqrt{\sin^2\theta}$$

$$= cos\theta \pm sin\theta \qquad \text{(since } sin\theta > 0\text{).}$$

If we try to solve for the eigenvectors we get:

$\lambda = cos\theta + sin\theta \implies$

$$A - \lambda I = \begin{pmatrix} cos\theta - (cos\theta + sin\theta) & -sin\theta \\ sin\theta & cos\theta - (cos\theta + sin\theta) \end{pmatrix}$$

$$= \begin{pmatrix} -sin\theta & sin\theta \\ sin\theta & sin\theta \end{pmatrix}$$

$$(A - \lambda I)\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} -sin\theta & sin\theta \\ sin\theta & sin\theta \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = 0$$

$$-a_1 sin\theta + a_2 sin\theta = 0$$

$$a_1 sin\theta + a_2 sin\theta = 0.$$

The only solutions to these equations is $a_1 = a_2 = 0$. So there aren't any eigenvectors associated with $\lambda = cos\theta + sin\theta$.

Similarly for $\lambda = cos\theta - sin\theta$ we get the same conclusion, i.e., there are no eigenvectors associated with $\lambda = cos\theta - sin\theta.$

So we see from this example, if T (and A) are normal that does not guarantee the existence of an orthonormal basis of eigenvectors.

However, the following theorem gives us several properties of normal operators.

Theorem: Let T be a normal operator on a real inner product space V. Then

a. $\|T(v)\| = \|T^*(v)\|$ for all $v \in V$
b. $T - \lambda I$ is normal for every $\lambda \in \mathbb{R}$
c. If v is an eigenvector of T associated with λ then v is an eigenvector of T^* associated with λ.
d. If λ_1 and λ_2 are distinct eigenvalues of T with corresponding eigenvectors v_1 and v_2, then v_1 and v_2 are orthogonal.

Proof:

a. For any $v \in V$ we have:

$$\begin{aligned} \|T(v)\|^2 &=< T(v), T(v) > \\ &=< T^*T(v), v > \\ &=< TT^*(v), v > \quad \text{(since } TT^* = T^*T) \\ &=< T^*(v), T^*(v) > \\ &= \|T^*(v)\|^2 \end{aligned}$$

So $\|T(v)\| = \|T^*(v)\|$ for all $v \in V$.

b. We need to show $(T - \lambda I)(T - \lambda I)^* = (T - \lambda I)^*(T - \lambda I)$ for every $\lambda \in \mathbb{R}$:

$$
\begin{aligned}
(T - \lambda I)(T - \lambda I)^* &= (T - \lambda I)(T^* - \lambda I^*) \\
&= (T - \lambda I)(T^* - \lambda I) \\
&= TT^* - \lambda T^* - \lambda T + \lambda^2 I^2 \\
&= T^*T - \lambda T^* - \lambda T + \lambda^2 I^2 \quad \text{(since } TT^* = T^*T\text{)} \\
&= (T^* - \lambda I^*)(T - \lambda I) \\
&= (T - \lambda I)^*(T - \lambda I).
\end{aligned}
$$

So $T - \lambda I$ is normal for every $\lambda \in \mathbb{R}$.

c. Suppose $T(v) = \lambda v$ for some $v \in V$ and let $S = T - \lambda I$.

S is normal by part b and

$$
\begin{aligned}
S(v) &= (T - \lambda I)(v) \\
&= T(v) - \lambda I(v) \\
&= \lambda v - \lambda v \\
&= 0.
\end{aligned}
$$

So
$$
\begin{aligned}
0 = \|S(v)\| &= \|S^*(v)\| \qquad \text{(by part a)} \\
&= \|(T^* - \lambda I)(v)\| \\
&= \|T^*(v) - \lambda v\|
\end{aligned}
$$

Hence $T^*(v) = \lambda v$, so v is an eigenvector of T^* associated with λ.

d. Let λ_1 and λ_2 be distinct eigenvalues of T corresponding to eigenvectors v_1 and v_2.

$$\lambda_1 < v_1, v_2 >=< \lambda_1 v_1, v_2 >$$
$$=< T(v_1), v_2 >$$
$$=< v_1, T^*(v_2) >$$
$$=< v_1,\ \lambda_2 v_2 > \qquad \text{(by part c)}$$
$$= \lambda_2 < v_1, v_2 > .$$

But $\lambda_1 \neq \lambda_2$, so $< v_1, v_2 >= 0$, and v_1 and v_2 are perpendicular.

Def. Let T be a linear operator on a real inner product space V. We say T is **self-adjoint** if $T = T^*$. An nxn real matrix A is **self-adjoint** if $A = A^* = A^t$

So nxn real matrix A is **self-adjoint** if A is symmetric.

Lemma: Let T be a self-adjoint operator on a finite dimensional real inner product space V. Then:

a. Every eigenvalue of T is real
b. The characteristic polynomial of T splits over $\mathbb{R}$.

Proof:

a. Suppose λ is an eigenvalue of T with eigenvector v. Then we have:

$$\|T(v)\|^2 =< T(v), T(v) >$$
$$=< v, T^*T(v) >$$
$$=< v, TT(v) > \qquad \text{(since } T \text{ is self-adjoint)}$$
$$=< vT(\lambda v) >$$
$$=< v, \lambda^2 v >$$
$$= \lambda^2 < v, v >= \lambda^2 \|v\|^2$$

Thus $\lambda^2 = \|T(v)\|^2/\|v\|^2$.

v is an eigenvector so nonzero and the right hand side is non-negative so λ is a real number.

b. Let β be an orthonormal basis for V. By the fundamental theorem of algebras, the characteristic polynomial for $[T]_\beta$ splits over the complex numbers into linear factors of the form $\lambda - \lambda_i$. But by part a, we know that all of the λ_i's are real. So the characteristic polynomial of T splits over $\mathbb{R}$.

This lemma and Schur's theorem leads us to:

Theorem: Let T be a linear operator on a real finite dimensional inner product space V. Then T is self-adjoint if and only if there exists an orthonormal basis β of V consisting of eigenvectors of T.

Proof: Suppose T is self-adjoint. By the previous lemma, the characteristic polynomial of T splits.

Thus by Schur's theorem, there is an orthonormal basis β for V such that $[T]_\beta$ is upper triangular.

If we let $A = [T]_\beta$ then A is upper triangular.

However, we also have:

$$A^* = [T]_\beta^* = [T^*]_\beta = [T]_\beta = A.$$

So A and A^* are both upper triangular and thus A is diagonal.

That means that β must be made up of eigenvectors of T.

Ex. For the following linear operators, T, determine if T is normal, self-adjoint or neither. If possible, find an orthonormal basis of eigenvectors of T for V.

a. $V = \mathbb{R}^2,\ T(a, b) = ((2a - 2b),\ (-2a + 5b).$
b. $V = \mathbb{R}^3,\ T(a_1, a_2, a_3) = (a_3, a_2, a_1).$

a. With respect to the standard ordered basis β for $\mathbb{R}^2$ we have

$$A = [T]_\beta = \begin{pmatrix} 2 & -2 \\ -2 & 5 \end{pmatrix}$$
$$A^t = [T]^t{}_\beta = \begin{pmatrix} 2 & -2 \\ -2 & 5 \end{pmatrix}$$

$[T]_\beta$ is symmetric so T is self-adjoint and normal.

To find the eigenvalues we take:

$$\det(A - \lambda I) = det\begin{pmatrix} 2-\lambda & -2 \\ -2 & 5-\lambda \end{pmatrix}$$
$$= (2-\lambda)(5-\lambda) - 4$$
$$= \lambda^2 - 7\lambda + 6.$$

$0 = \lambda^2 - 7\lambda + 6 = (\lambda - 1)(\lambda - 6) \quad \Longrightarrow \ \lambda = 1, 6;$ eigenvalues.

To find the eigenvectors we find the null space of $A - \lambda I$.

$\lambda_1 = 1$: $\quad A - \lambda I = A - I = \begin{pmatrix} 1 & -2 \\ -2 & 4 \end{pmatrix}$.

$$\begin{pmatrix} 1 & -2 \\ -2 & 4 \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$
$$a_1 - 2a_2 = 0$$
$$-2a_1 + 4a_2 = 0$$

So, $\quad a_1 = 2a_2$: $\quad v =< 2a_2, a_2 >= a_2 < 2,1 >$.

Thus all eigenvectors associated with $\lambda_1 = 1$ have the form $a_2 < 2,1 >$.

Hence an eigenvector of length one that spans this space is:

$$u_1 = \frac{1}{\sqrt{5}} < 2,1 > .$$

$\lambda_2 = 6$: $\quad A - \lambda I = A - 6I = \begin{pmatrix} -4 & -2 \\ -2 & -1 \end{pmatrix}$.

$$\begin{pmatrix} -4 & -2 \\ -2 & -1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$
$$-4a_1 - 2a_2 = 0$$
$$-2a_1 - a_2 = 0$$

So, $a_2 = -2a_1$: $\quad v =< a_1, -2a_1 >= a_1 < 1, -2 >$.

Thus all eigenvectors associated with $\lambda_2 = 6$ have the form $a_1 < 1, -2 >$.

Hence an eigenvector of length one that spans this space is:

$$u_2 = \frac{1}{\sqrt{5}} < 1, -2 >.$$

By an earlier theorem we know that eigenvectors associated with different eigenvalues will always be perpendicular. Thus the set $\beta = \{u_1, u_2\} = \{\frac{1}{\sqrt{5}} < 2,1 >, \ \frac{1}{\sqrt{5}} < 1, -2 >\}$, is an orthonormal basis for $\mathbb{R}^2$.

b. With respect to the standard ordered basis for $\mathbb{R}^3$ we have

$$T(a_1, a_2, a_3) = (a_3, a_2, a_1)$$

$$A = [T]_\beta = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

$$A^t = {[T]^t}_\beta = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

A is symmetric so T is both self-adjoint and normal.

To find the eigenvalues we take:

$$\det(A - \lambda I) = \begin{pmatrix} -\lambda & 0 & 1 \\ 0 & 1-\lambda & 0 \\ 1 & 0 & -\lambda \end{pmatrix}$$

$$= -(\lambda - 1)^2(\lambda + 1)$$

So the eigenvalues are $\lambda = -1, 1$.

Now let's find the eigenvectors:

$$\lambda_1 = -1: \quad A - \lambda I = A + I = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

So we solve: $\begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$

$$a_1 + a_3 = 0$$
$$2a_2 = 0$$
$$a_1 + a_3 = 0.$$

Thus the null space of $A + I$ is all vectors of the form $a_1 < 1,0,-1 >$.

A unit vector that spans this space is given by: $\frac{1}{\sqrt{2}} < 1,0,-1 >$.

$$\lambda_2 = \lambda_3 = 1: \quad A - \lambda I = A - I = \begin{pmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{pmatrix}$$

To find the null space of $A - I$ we solve:

$$\begin{pmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$-a_1 + a_3 = 0$$

$$a_1 - a_3 = 0.$$

So $a_1 = a_3$ and a_2 is a free variable.

Thus all vectors in this null space are of the form:

$$< a_1, a_2, a_1 >= a_1 < 1,0,1 > + a_2 < 0,1,0 >.$$

So $\{< 1,0,1 >, < 0,1,0 >\}$ span this null space. These vectors are already perpendicular so we just need to normalize them to get an orthonormal basis for this null space. Hence an orthonormal basis for $\mathbb{R}^3$ is given by:

$$u_1 = \frac{1}{\sqrt{2}} < 1,0,-1 >$$

$$u_2 = \frac{1}{\sqrt{2}} < 1,0,1 >$$

$$u_3 =< 0, 1,0 >$$