Adjoints, normal and unitary operators, consequences for eigenvalues

Let's recall how we thought about transpose way back when: Given vector spaces V and W and a linear map $A : V \to W$, the dual map $A^* : W^* \to V^*$ is defined by

$$A^*(w^*)(v) = w^*(A(v)).$$

Let's recall how we thought about transpose way back when: Given vector spaces V and W and a linear map $A : V \to W$, the dual map $A^* : W^* \to V^*$ is defined by

$$A^*(w^*)(v) = w^*(A(v)).$$

In coordinates, A^* is given by the transpose matrix.

Let's recall how we thought about transpose way back when: Given vector spaces V and W and a linear map $A : V \to W$, the dual map $A^* : W^* \to V^*$ is defined by

$$A^*(w^*)(v) = w^*(A(v)).$$

In coordinates, A^* is given by the transpose matrix.

If V and W are finite dimensional **real** vector spaces with inner products, then we don't have to keep track of the distinction between V and V^*: We define the transpose map $W \to V$ by

$$\langle A(v), w \rangle = \left\langle v, A^T(w) \right\rangle .$$

Let's recall how we thought about transpose way back when: Given vector spaces V and W and a linear map $A : V \to W$, the dual map $A^* : W^* \to V^*$ is defined by

$$A^*(w^*)(v) = w^*(A(v)).$$

In coordinates, A^* is given by the transpose matrix.

If V and W are finite dimensional **real** vector spaces with inner products, then we don't have to keep track of the distinction between V and V^*: We define the transpose map $W \to V$ by

$$\langle A(v), w \rangle = \left\langle v, A^T(w) \right\rangle .$$

For finite dimensional complex vector spaces, we analogously define $A^\dagger : W \to V$ by

$$\langle A(v), w \rangle = \left\langle v, A^\dagger(w) \right\rangle .$$

For finite dimensional complex vector spaces, we analogously define $A^\dagger : W \to V$ by

$$\langle A(v), w\rangle = \left\langle v, A^\dagger(w)\right\rangle .$$

In terms of orthonormal bases for V and W, the $A^\dagger$ map is given by the complex conjugate transpose matrix. If A is real, then $A^\dagger$ is just A^T.

For finite dimensional complex vector spaces, we analogously define $A^\dagger : W \to V$ by

$$\langle A(v), w \rangle = \left\langle v, A^\dagger(w) \right\rangle .$$

In terms of orthonormal bases for V and W, the $A^\dagger$ map is given by the complex conjugate transpose matrix. If A is real, then $A^\dagger$ is just A^T.

In particular, if $V = W$, then both A and $A^\dagger$ are square matrices.

This gives us some key definitions. Let V be a finite dimensional real or complex vector space with an inner product and let $A : V \to V$ be a linear map.

1. The map A is called ***self-adjoint*** if $A = A^\dagger$. In the real case this is also called ***symmetric***; in the complex case it is also called ***Hermitian***.

2. The map A is called ***unitary*** if $A^{-1} = A^\dagger$. In the real case, we also call this ***orthogonal***.

The map A is called ***unitary*** if $A^{-1} = A^{\dagger}$. In the real case, we also call this ***orthogonal***.

The geometric meaning of a unitary/orthogonal matrix is that it preserves the inner product:

$$\langle A(\vec{v}), A(\vec{w}) \rangle = \langle \vec{v}, A^{\dagger}(A(\vec{w})) \rangle = \langle \vec{v}, \vec{w} \rangle.$$

The map A is called ***unitary*** if $A^{-1} = A^\dagger$. In the real case, we also call this ***orthogonal***.

The geometric meaning of a unitary/orthogonal matrix is that it preserves the inner product:

$$\langle A(\vec{v}), A(\vec{w}) \rangle = \langle \vec{v}, A^\dagger(A(\vec{w})) \rangle = \langle \vec{v}, \vec{w} \rangle.$$

Another way to say this is that the rows/columns of an orthogonal/unitary matrix are an orthonormal basis.

Apology for notation: We really should call these "orthonormal matrices". Sorry.

Let V be a finite dimensional real or complex vector space with an inner product and let $A : V \to V$ be a linear map.

1. The map A is called ***self-adjoint*** if $A = A^\dagger$. In the real case this is also called ***symmetric***; in the complex case it is also called ***Hermitian***.

2. The map A is called ***unitary*** if $A^{-1} = A^\dagger$. In the real case, we also call this ***orthogonal***.

3. There is a useful condition which includes both "self-adjoint" and "unitary": The map A is called ***normal*** if $AA^T = A^T A$.

Here is where we are heading:

Theorem: A matrix is normal if and only if it has an orthonormal basis over $\mathbb{C}$.

A normal matrix is self-adjoint if and only if it has real eigenvalues.

A normal matrix is unitary if and only if it has eigenvalues on the unit circle: $\{\cos\theta + i\sin\theta\}$.

Example:

The matrix

$$\begin{bmatrix} 0.6 & 0.8 \\ -0.8 & 0.6 \end{bmatrix}$$

is unitary (and orthogonal).

Example:

The matrix

$$\begin{bmatrix} 0.6 & 0.8 \\ -0.8 & 0.6 \end{bmatrix}$$

is unitary (and orthogonal). Its characteristic polynomial is

$$\det \begin{bmatrix} x-0.6 & -0.8 \\ 0.8 & x-0.6 \end{bmatrix} = (x-0.6)^2 + (0.8)^2 = x^2 - 1.2x + 1.$$

The roots are $0.6 \pm 0.8i$.

Example:

The matrix

$$\begin{bmatrix} 0.6 & 0.8 \\ -0.8 & 0.6 \end{bmatrix}$$

is unitary (and orthogonal). Its characteristic polynomial is

$$\det \begin{bmatrix} x-0.6 & -0.8 \\ 0.8 & x-0.6 \end{bmatrix} = (x-0.6)^2 + (0.8)^2 = x^2 - 1.2x + 1.$$

The roots are $0.6 \pm 0.8i$.

The $0.6 + 0.8i$ eigenvector is

$$\operatorname{Ker} \begin{bmatrix} (0.6+0.8i)-0.6 & -0.8 \\ 0.8 & (0.6+0.8i)-0.6 \end{bmatrix} = \operatorname{Ker} \begin{bmatrix} 0.8i & -0.8 \\ 0.8 & 0.8i \end{bmatrix} = \mathbb{C} \begin{bmatrix} 1 \\ i \end{bmatrix}.$$

Example:

The matrix

$$\begin{bmatrix} 0.6 & 0.8 \\ -0.8 & 0.6 \end{bmatrix}$$

is unitary (and orthogonal). Its characteristic polynomial is

$$\det \begin{bmatrix} x-0.6 & -0.8 \\ 0.8 & x-0.6 \end{bmatrix} = (x-0.6)^2 + (0.8)^2 = x^2 - 1.2x + 1.$$

The roots are $0.6 \pm 0.8i$.

The $0.6 + 0.8i$ eigenvector is

$$\mathrm{Ker} \begin{bmatrix} (0.6+0.8i)-0.6 & -0.8 \\ 0.8 & (0.6+0.8i)-0.6 \end{bmatrix} = \mathrm{Ker} \begin{bmatrix} 0.8i & -0.8 \\ 0.8 & 0.8i \end{bmatrix} = \mathbb{C} \begin{bmatrix} 1 \\ i \end{bmatrix}.$$

Similarly, the $0.6 - 0.8i$ eigenvector is $\begin{bmatrix} 1 \\ -i \end{bmatrix}$.

Example continued:

The vectors $\begin{bmatrix} 1 \\ i \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -i \end{bmatrix}$ are orthogonal, and we can scale them to be orthonormal. So

$$\begin{bmatrix} 0.6 & 0.8 \\ -0.8 & 0.6 \end{bmatrix} = U \begin{bmatrix} 0.6+0.8i & 0 \\ 0 & 0.6-0.8i \end{bmatrix} U^{-1} = U \begin{bmatrix} 0.6+0.8i & 0 \\ 0 & 0.6-0.8i \end{bmatrix} U^{\dagger}$$

$$\text{for } U = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix}.$$

Theorem: A matrix is normal if and only if it has an orthonormal basis over $\mathbb{C}$. A normal matrix is self-adjoint if and only if it has real eigenvalues. A normal matrix is unitary if and only if it has eigenvalues on the unit circle: $\{\cos\theta + i\sin\theta\}$.

The easy parts: If $A = UDU^\dagger$, then
$A^\dagger = (UDU^\dagger)^\dagger = U^{\dagger\dagger}D^\dagger U^\dagger = U\overline{D}U^\dagger$, so

$$AA^\dagger = UDU^\dagger U\overline{D}U^\dagger = UD\overline{D}U^\dagger = U\overline{D}DU^\dagger = U\overline{D}U^\dagger UDU^\dagger = A^\dagger A.$$

If the eigenvalues are real, we have $\overline{D} = D$, so this shows that
$A^\dagger = U\overline{D}U^\dagger = UDU^\dagger = A$.

If the eigenvalues are on the unit circle, then we have $\overline{D} = D^{-1}$, so this shows that $A^\dagger = U\overline{D}U^\dagger = UD^{-1}U^\dagger = A^{-1}$.

The hard part:

Let V be a finite dimensional complex vector space with an inner product. Let $A : V \to V$ be normal. We want to show that A has an orthonormal basis.

The hard part:

Let V be a finite dimensional complex vector space with an inner product. Let $A : V \to V$ be normal. We want to show that A has an orthonormal basis.

We do know that $\chi_A(x)$ has a root, since every polynomial over the complex numbers has a root. So let $\vec{v}$ be a nonzero eigenvector of A, with eigenvalue λ. We want to show that A takes $\vec{v}^{\perp}$ to $\vec{v}^{\perp}$. Then we will just induct.

Let V be a finite dimensional complex vector space with an inner product. Let $A : V \to V$ be normal ... We want to show that A takes $\vec{v}^{\perp}$ to $\vec{v}^{\perp}$.

Let V be a finite dimensional complex vector space with an inner product. Let $A : V \to V$ be normal ... We want to show that A takes $\vec{v}^\perp$ to $\vec{v}^\perp$.

Take an orthonormal basis $\vec{v}_2, \vec{v}_3, \ldots, \vec{v}_n$ of $\vec{v}^\perp$, so $\vec{v}, \vec{v}_2, \vec{v}_3, \ldots, \vec{v}_n$ is an orthonormal basis of V. In this basis,

$$A = \begin{bmatrix} \lambda & A_{12} & A_{13} & \cdots & A_{1n} \\ 0 & * & * & \cdots & * \\ 0 & * & * & \cdots & * \\ & & & \ddots & \\ 0 & * & * & \cdots & * \end{bmatrix}.$$

Let V be a finite dimensional complex vector space with an inner product. Let $A : V \to V$ be normal ... We want to show that A takes $\vec{v}^{\perp}$ to $\vec{v}^{\perp}$.

Take an orthonormal basis $\vec{v}_2, \vec{v}_3, \ldots, \vec{v}_n$ of $\vec{v}^{\perp}$, so $\vec{v}, \vec{v}_2, \vec{v}_3, \ldots, \vec{v}_n$ is an orthonormal basis of V. In this basis,

$$A = \begin{bmatrix} \lambda & A_{12} & A_{13} & \cdots & A_{1n} \\ 0 & * & * & \cdots & * \\ 0 & * & * & \cdots & * \\ & & & \ddots & \\ 0 & * & * & \cdots & * \end{bmatrix} \text{ and } A^{\dagger} = \begin{bmatrix} \overline{\lambda} & 0 & 0 & \cdots & 0 \\ \overline{A}_{12} & * & * & \cdots & * \\ \overline{A}_{13} & * & * & \cdots & * \\ & & & \ddots & \\ \overline{A}_{1n} & * & * & \cdots & * \end{bmatrix}.$$

Let V be a finite dimensional complex vector space with an inner product. Let $A : V \to V$ be normal ... We want to show that A takes $\vec{v}^{\perp}$ to $\vec{v}^{\perp}$.

Take an orthonormal basis $\vec{v}_2, \vec{v}_3, \ldots, \vec{v}_n$ of $\vec{v}^{\perp}$, so $\vec{v}, \vec{v}_2, \vec{v}_3, \ldots, \vec{v}_n$ is an orthonormal basis of V. In this basis,

$$A = \begin{bmatrix} \lambda & A_{12} & A_{13} & \cdots & A_{1n} \\ 0 & * & * & \cdots & * \\ 0 & * & * & \cdots & * \\ & & & \ddots & \\ 0 & * & * & \cdots & * \end{bmatrix} \text{ and } A^{\dagger} = \begin{bmatrix} \overline{\lambda} & 0 & 0 & \cdots & 0 \\ \overline{A}_{12} & * & * & \cdots & * \\ \overline{A}_{13} & * & * & \cdots & * \\ & & & \ddots & \\ \overline{A}_{1n} & * & * & \cdots & * \end{bmatrix}.$$

The $(1,1)$ entry of $AA^{\dagger}$ is

$$\overline{\lambda}\lambda + A_{12}\overline{A}_{12} + A_{13}\overline{A}_{13} + \cdots + A_{1n}\overline{A}_{1n} = |\lambda|^2 + \sum_{j=2}^{n} |A_{1j}|^2.$$

Let V be a finite dimensional complex vector space with an inner product. Let $A : V \to V$ be normal ... We want to show that A takes $\vec{v}^{\perp}$ to $\vec{v}^{\perp}$.

Take an orthonormal basis $\vec{v}_2, \vec{v}_3, \ldots, \vec{v}_n$ of $\vec{v}^{\perp}$, so $\vec{v}, \vec{v}_2, \vec{v}_3, \ldots, \vec{v}_n$ is an orthonormal basis of V. In this basis,

$$A = \begin{bmatrix} \lambda & A_{12} & A_{13} & \cdots & A_{1n} \\ 0 & * & * & \cdots & * \\ 0 & * & * & \cdots & * \\ & & & \ddots & \\ 0 & * & * & \cdots & * \end{bmatrix} \text{ and } A^{\dagger} = \begin{bmatrix} \overline{\lambda} & 0 & 0 & \cdots & 0 \\ \overline{A}_{12} & * & * & \cdots & * \\ \overline{A}_{13} & * & * & \cdots & * \\ & & & \ddots & \\ \overline{A}_{1n} & * & * & \cdots & * \end{bmatrix}.$$

The $(1,1)$ entry of $AA^{\dagger}$ is

$$\lambda\overline{\lambda} + A_{12}\overline{A}_{12} + A_{13}\overline{A}_{13} + \cdots + A_{1n}\overline{A}_{1n} = |\lambda|^2 + \sum_{j=2}^{n} |A_{1j}|^2.$$

The $(1,1)$ entry of $A^{\dagger}A$ is

$$\overline{\lambda}\lambda + 0 + 0 + \cdots + 0 = |\lambda|^2.$$

The $(1,1)$ entry of $AA^\dagger$ is

$$\lambda\overline{\lambda} + A_{12}\overline{A}_{12} + A_{13}\overline{A}_{13} + \cdots + A_{1n}\overline{A}_{1n} = |\lambda|^2 + \sum_{j=2}^{n} |A_{1j}|^2.$$

The $(1,1)$ entry of $A^\dagger A$ is

$$\overline{\lambda}\lambda + 0 + 0 + \cdots + 0 = |\lambda|^2.$$

So $|\lambda|^2 + \sum_{j=2}^{n} |A_{1j}|^2 = |\lambda|^2$ and $A_{12} = A_{13} = \cdots = A_{1n} = 0$.

The $(1,1)$ entry of $AA^\dagger$ is

$$\lambda\overline{\lambda} + A_{12}\overline{A}_{12} + A_{13}\overline{A}_{13} + \cdots + A_{1n}\overline{A}_{1n} = |\lambda|^2 + \sum_{j=2}^{n} |A_{1j}|^2.$$

The $(1,1)$ entry of $A^\dagger A$ is

$$\overline{\lambda}\lambda + 0 + 0 + \cdots + 0 = |\lambda|^2.$$

So $|\lambda|^2 + \sum_{j=2}^{n} |A_{1j}|^2 = |\lambda|^2$ and $A_{12} = A_{13} = \cdots = A_{1n} = 0$.

So A is of the form

$$\begin{bmatrix} \lambda & 0 & 0 & \cdots & 0 \\ 0 & * & * & \cdots & * \\ 0 & * & * & \cdots & * \\ & & & \ddots & \\ 0 & * & * & \cdots & * \end{bmatrix}.$$

We see that A takes $\vec{v}^\perp$, and now we can induct. **QED**