Adjoints and Unitaries

This week we will learn about:

- The adjoint of a linear transformation, and
- Unitary transformations and matrices.

Extra reading and watching:

- Sections 1.4.2 and 1.4.3 in the textbook
- Lecture videos 23 and 24 on YouTube
- Unitary matrix at Wikipedia

Extra textbook problems:

$\star$ 1.4.5(b,c,e,f), 1.4.8

$\star\star$ 1.4.3, 1.4.9–1.4.14, 1.4.21, 1.4.22

$\star\star\star$ 1.4.6, 1.4.15, 1.4.18

☠ 1.4.19, 1.4.28

We now introduce the adjoint of a linear transformation, which we can think of as a way of generalizing the transpose of a real matrix to linear transformations between arbitrary inner product spaces.

Definition 6.1 — Adjoint Transformations

Suppose that $\mathcal{V}$ and $\mathcal{W}$ are inner product spaces and $T : \mathcal{V} \to \mathcal{W}$ is a linear transformation. Then a linear transformation $T^* : \mathcal{W} \to \mathcal{V}$ is called the **adjoint** of T if

$$\langle T(\vec{v}), \vec{w} \rangle = \langle \vec{v}, T^*(\vec{w}) \rangle \quad \text{for all } \vec{v} \in \mathcal{V} \text{ and } \vec{w} \in \mathcal{W}.$$

For example, the adjoint of a matrix $A \in \mathcal{M}_{m,n}(\mathbb{R})$ is (i.e., $A : \mathbb{R}^n \to \mathbb{R}^m$)

$$\begin{aligned}(A\vec{v}) \cdot \vec{w} &= \left(\sum_{j=1}^n a_{1,j} v_j, \sum_{j=1}^n a_{2,j} v_j, \ldots, \sum_{j=1}^n a_{m,j} v_j \right) \cdot (w_1, w_2, \ldots, w_m) \\ &= \sum_{i=1}^m \left(\sum_{j=1}^n a_{i,j} v_j \right) w_i = \sum_{j=1}^n \left(\sum_{i=1}^m a_{i,j} w_i \right) v_j \\ &= (v_1, v_2, \ldots, v_n) \cdot \left(\sum_{i=1}^m a_{i,1} w_i, \sum_{i=1}^m a_{i,2} w_i, \ldots, \sum_{i=1}^m a_{i,n} w_i \right) = \vec{v} \cdot (A^T \vec{w})\end{aligned}$$

$\therefore$ The adjoint of $A \in \mathcal{M}_{m,n}(\mathbb{R})$ is its transpose A^T.

Similarly, the adjoint of a matrix $A \in \mathcal{M}_{m,n}(\mathbb{C})$ is

its conjugate transpose A^*. (similar to above)
This is why the conjugate transpose is important in the first place!

So far, we have been a bit careless and referred to "the" adjoint of a matrix (linear transformation), even though it perhaps seems believable that a linear transformation might have more than one adjoint. The following theorem shows that, at least in finite dimensions, this is not actually a problem.

Theorem 6.1 — Existence and Uniqueness of Adjoints

Suppose that $\mathcal{V}$ and $\mathcal{W}$ are finite-dimensional inner product spaces. For every linear transformation $T : \mathcal{V} \to \mathcal{W}$ there exists a unique adjoint transformation $T^* : \mathcal{W} \to \mathcal{V}$. Furthermore, if B and C are orthonormal bases of $\mathcal{V}$ and $\mathcal{W}$ respectively, then

$$[T^*]_{B \leftarrow C} = [T]^*_{C \leftarrow B}.$$

Proof. To prove uniqueness of T^*, suppose that T^* exists, let $\mathbf{v} \in \mathcal{V}$ and $\mathbf{w} \in \mathcal{W}$, and compute $\langle T(\mathbf{v}), \mathbf{w} \rangle$ in two different ways:

$$(1)\quad \langle T(\vec{v}), \vec{w} \rangle = [T(\vec{v})]_C \cdot [\vec{w}]_C = \left([T]_{C\leftarrow B}[\vec{v}]_B\right) \cdot [\vec{w}]_C = [\vec{v}]_B^* [T]^*_{C\leftarrow B}[\vec{w}]_C$$

$$(2)\quad \langle T(\vec{v}), \vec{w} \rangle = \langle \vec{v}, T^*(\vec{w}) \rangle$$
$$= [\vec{v}]_B \cdot [T^*(\vec{w})]_B = [\vec{v}]_B \cdot \left([T^*]_{B\leftarrow C}[\vec{w}]_C\right) = [\vec{v}]_B^* [T^*]_{B\leftarrow C}[\vec{w}]_C$$

$\therefore$ $[\vec{v}]_B^* [T]^*_{C\leftarrow B}[\vec{w}]_C = [\vec{v}]_B^* [T^*]_{B\leftarrow C}[\vec{w}]_C$ for all $\vec{v} \in \mathcal{V}$ and $\vec{w} \in \mathcal{W}$, so $[T]^*_{C\leftarrow B} = [T^*]_{B\leftarrow C}$.

Uniqueness follows from uniqueness of standard matrices.

To prove existence of T, we can choose T^* to be the linear transformation with standard matrix $[T]^*_{C\leftarrow B}$ and then repeat the above argument backward to see that $\langle T(\vec{v}), \vec{w} \rangle = \langle \vec{v}, T^*(\vec{w}) \rangle$ for all $\vec{v} \in \mathcal{V}$ and $\vec{w} \in \mathcal{W}$. ∎

Example. *Show that the adjoint of the transposition map $T : \mathcal{M}_{m,n} \to \mathcal{M}_{n,m}$, with the Frobenius inner product, is also the transposition map.*

Method 1:

We want to show that $\langle A^T, B\rangle = \langle A, B^T\rangle$ for all $A \in \mathcal{M}_{m,n}$, $B \in \mathcal{M}_{n,m}$.

Well, $\langle A^T, B\rangle = tr(\bar{A}B) = tr(B\bar{A}) = tr(A^* B^T) = \langle A, B^T\rangle$. ✓

Method 2: (For $m=n=2$ only)

Recall that if E is the standard basis then $[T]_E = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$. Then $[T]_E^* = [T]_E$, so T is its own adjoint.

The situation presented in the above example, where a linear transformation is its own adjoint, is important enough that we give it a name:

Definition 6.2 — Self-Adjoint Transformations

Suppose that $\mathcal{V}$ is an inner product space. Then a linear transformation $T : \mathcal{V} \to \mathcal{V}$ is called **self-adjoint** if $T^* = T$.

For example, a matrix in $\mathcal{M}_n(\mathbb{R})$ is self-adjoint if and only if it is...

symmetric (i.e., $A^T = A$).

and a matrix in $\mathcal{M}_n(\mathbb{C})$ is self-adjoint if and only if it is...

Hermitian (i.e., $A^* = A$).

Furthermore, a linear transformation is self-adjoint if and only if its standard matrix...

is symmetric/Hermitian (with respect to some orthonormal basis).

if $\mathbb{F} = \mathbb{R}$ if $\mathbb{F} = \mathbb{C}$

Unitary Transformations and Matrices

In situations where the norm of a vector is important, it is often desirable to work with linear transformations that do not alter that norm. We now start investigating these linear transformations.

Definition 6.3 — Unitary Transformations

Let $\mathcal{V}$ and $\mathcal{W}$ be inner product spaces and let $T : \mathcal{V} \to \mathcal{W}$ be a linear transformation. Then T is said to be **unitary** if

$$\|T(\mathbf{v})\| = \|\mathbf{v}\| \quad \text{for all} \quad \mathbf{v} \in \mathcal{V}.$$

We also say that a *matrix* is unitary if it acts as a unitary linear transformation on $\mathbb{F}^n$.

Example. *Show that the matrix* $U = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ *is unitary.*

$$\|U\vec{v}\|^2 = \left\|\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix}\right\|^2 = \left\|\frac{1}{\sqrt{2}}\begin{bmatrix} v_1 - v_2 \\ v_1 + v_2 \end{bmatrix}\right\|^2$$

$$= \left(\frac{1}{\sqrt{2}}(v_1 - v_2)\right)^2 + \left(\frac{1}{\sqrt{2}}(v_1 + v_2)\right)^2$$

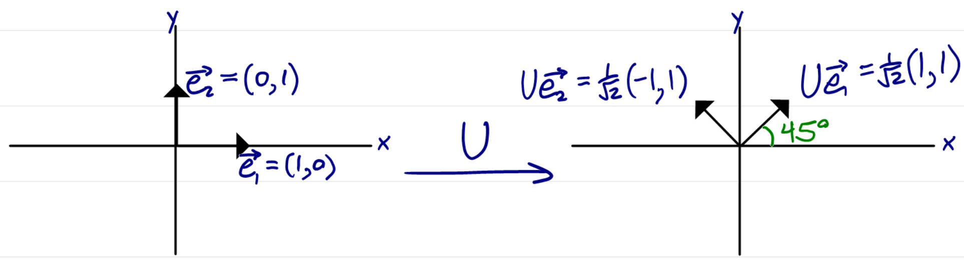
$$= \frac{1}{2}\left(v_1^2 - \cancel{2v_1v_2} + v_2^2\right) + \frac{1}{2}\left(v_1^2 + \cancel{2v_1v_2} + v_2^2\right)$$

$$= v_1^2 + v_2^2$$

$$= \|\vec{v}\|^2 \quad \text{for all} \quad \vec{v} \in \mathbb{R}^2. \checkmark$$

$\therefore$ U is unitary.

Makes sense: U is a rotation matrix.



Fortunately, there is a much simpler method of checking whether or not a matrix (or a linear transformations) is unitary, as demonstrated by the following theorem.

Theorem 6.2 — Characterization of Unitary Matrices

Suppose $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$, and $U \in \mathcal{M}_n(\mathbb{F})$. The following are equivalent:

a) U is unitary,

b) $U^*U = I$,

c) $UU^* = I$,

d) $(U\mathbf{v}) \cdot (U\mathbf{w}) = \mathbf{v} \cdot \mathbf{w}$ for all $\mathbf{v}, \mathbf{w} \in \mathbb{F}^n$,

e) The columns of U are an orthonormal basis of $\mathbb{F}^n$, and

f) The rows of U are an orthonormal basis of $\mathbb{F}^n$.

It is worth comparing these properties to corresponding properties of invertible matrices:

Invertible P	Unitary U
P^{-1} exists	$U^{-1} = U^*$
$\lVert P\vec{v}\rVert \neq 0$ if $\lVert\vec{v}\rVert \neq 0$	$\lVert U\vec{v}\rVert = \lVert\vec{v}\rVert$ for all $\vec{v}$
columns of P are basis	cols. of U are orthonormal basis

Also (d) above says unitaries preserve angles.

Proof of Theorem 6.2. We do not prove all equivalences of this theorem – for that you can see the textbook. But we will demonstrate some of them in order to give an idea of why this theorem is true.

The equivalence of (b) and (c) follows from the fact that

if $A, B \in M_n$ then $AB = I$ if and only if $BA = I$.

To see that (d) $\implies$ (b), note that if we rearrange the equation $(U\mathbf{v}) \cdot (U\mathbf{w}) = \mathbf{v} \cdot \mathbf{w}$ slightly, we get

$\vec{v} \cdot (U^*U\vec{w}) = \vec{v} \cdot \vec{w}$, so $\vec{v} \cdot ((U^*U - I)\vec{w}) = 0$ for all $\vec{v}, \vec{w}$.

If we choose $\vec{v} = (U^*U - I)\vec{w}$ then we see that

$0 = ((U^*U - I)\vec{w}) \cdot ((U^*U - I)\vec{w}) = \|((U^*U - I)\vec{w})\|^2$, so $(U^*U - I)\vec{w} = \vec{0}$.
Since this is true for all $\vec{w}$, we get $U^*U = I$.

To see that (b) implies (a), suppose $U^*U = I$. Then for all $\mathbf{v} \in \mathbb{F}^n$ we have

$\|U\vec{v}\|^2 = (U\vec{v}) \cdot (U\vec{v}) = \vec{v} \cdot (U^*U\vec{v}) = \vec{v} \cdot \vec{v} = \|\vec{v}\|^2$, so
U is unitary.

To see that (b) is equivalent to (e), write U in terms of its columns $U = \left[\, \mathbf{u}_1 \mid \mathbf{u}_2 \mid \cdots \mid \mathbf{u}_n \,\right]$ and then use block matrix multiplication to multiply by U^*:

$$U^*U = \begin{bmatrix} \vec{u}_1^{\,*} \\ \hline \vec{u}_2^{\,*} \\ \hline \vdots \\ \hline \vec{u}_n^{\,*} \end{bmatrix} \left[\vec{u}_1 \mid \vec{u}_2 \mid \cdots \mid \vec{u}_n \right] = \begin{bmatrix} \vec{u}_1^{\,*}\vec{u}_1 & \vec{u}_1^{\,*}\vec{u}_2 & \cdots & \vec{u}_1^{\,*}\vec{u}_n \\ \vec{u}_2^{\,*}\vec{u}_1 & \vec{u}_2^{\,*}\vec{u}_2 & \cdots & \vec{u}_2^{\,*}\vec{u}_n \\ \vdots & \vdots & \ddots & \vdots \\ \vec{u}_n^{\,*}\vec{u}_1 & \vec{u}_n^{\,*}\vec{u}_1 & \cdots & \vec{u}_n^{\,*}\vec{u}_n \end{bmatrix},$$

which equals I if and only if its diagonal entries equal 1 (i.e., $\|\vec{u}_j\| = 1$ for all j) and other entries equal 0 (i.e., $\vec{u}_i \cdot \vec{u}_j = 0$ for all $i \neq j$).

The remaining implications can be proved using similar techniques. ■

Checking whether or not a matrix is unitary is now quite simple, since we just have to check whether or not $U^*U = I$. For example, if we again return to the matrix

$$U = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

from earlier:

$$U^*U = \left(\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\right)\left(\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\right) = \frac{1}{2}\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = I,$$ so U is unitary.

More generally, every rotation matrix and reflection matrix is unitary, as we now demonstrate.

Example. *Show that every rotation matrix $U \in \mathcal{M}_2(\mathbb{R})$ is unitary.*

By geometry, this is obvious. Algebraically:

$$U = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}, \quad \text{so}$$

$$U^*U = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$
$$= \begin{bmatrix} \cos^2(\theta)+\sin^2(\theta) & -\cos(\theta)\sin(\theta)+\sin(\theta)\cos(\theta) \\ -\sin(\theta)\cos(\theta)+\cos(\theta)\sin(\theta) & \sin^2(\theta)+\cos^2(\theta) \end{bmatrix}$$

$= I$, so U is unitary.

across a line

Example. *Show that every reflection matrix $U \in \mathcal{M}_n(\mathbb{R})$ is unitary.*

Again, obvious by geometry. Algebraically:
$U = 2\vec{u}\vec{u}^* - I$, where $\vec{u}$ is a unit vector in the direction of the line

Then
$$\begin{aligned} U^*U &= (2\vec{u}\vec{u}^* - I)^*(2\vec{u}\vec{u}^* - I) \\ &= 4\vec{u}(\vec{u}^*\vec{u})\vec{u}^* - 4\vec{u}\vec{u}^* + I \\ &= 4\vec{u}\vec{u}^* - 4\vec{u}\vec{u}^* + I \\ &= I, \end{aligned}$$
so U is unitary.

In fact, the previous two examples provide exactly the intuition that you should have for unitary matrices—they are the ones that rotate and/or reflect $\mathbb{F}^n$, but do not stretch, shrink, or otherwise "distort" it. They can be thought of as "rigid" linear transformations that leave the size and shape of $\mathbb{F}^n$ in tact, but possibly change its orientation.