Unitary Matrices

Now that we have defined orthogonality, and even used the Gram-Schmidt procedure, the time has come to define an orthogonal matrix.

Definition: An $n \times n$ matrix with complex entries is said to be unitary if its columns form an orthonormal basis for $\mathbb{C}^n$.

The term "unitary" is used instead of "orthogonal" to emphasize that, thanks to the new definition of an inner product, we do not end up with the same properties as we did with orthogonal real matrices. Specifically, we do not have that A is unitary if and only if $A^{-1} = A^T$. Let's look at why.

Let A be an $n \times n$ matrix with complex entries, and let $\vec{a}_1, \ldots, \vec{a}_n$ be the columns of A.

Even in the complex numbers, we have that the jk-th entry of a matrix product BA is the dot product of the j-th row of B with the k-th column of A. If we set B equal to A^T, then we have that the j-th row of A^T is the same as the j-th column of A, and so the jk-th entry of A^TA is $\vec{a}_j \cdot \vec{a}_k$.

Now, back when A had real entries, this dot product was the same as the standard inner product. But now,

$$\vec{a}_j \cdot \vec{a}_k \neq \langle \vec{a}_j, \vec{a}_k \rangle.$$

So, even if A is a unitary matrix, the fact that $\langle \vec{a}_j, \vec{a}_j \rangle = 1$ does not necessarily mean that $\vec{a}_j \cdot \vec{a}_j = 1$, which means that the diagonal entries of A^TA may not be 1.

Unitary Matrices

And also, just because $\langle \vec{a}_j, \vec{a}_k \rangle = 0$ when $j \neq k$, we do not necessarily have that $\vec{a}_j \cdot \vec{a}_k = 0$, which means that A^TA may not have zeros off the main diagonal.

All is not lost, however, because $\langle \vec{a}_j, \vec{a}_k \rangle$ is not that different from $\vec{a}_j \cdot \vec{a}_k$. We are simply looking at $\vec{a}_j \cdot \overline{\vec{a}_k}$ instead. And, as we always take the conjugate of the right hand side, this dot product is the jk-th entry of the matrix $A^T\overline{A}$.

This general fact is true for any matrix A with complex entries, but when A is unitary, then we again have that $\vec{a}_j \cdot \overline{\vec{a}_j} = 1$, so the entries on the main diagonal of $A^T\overline{A}$ are 1, and that $\vec{a}_j \cdot \overline{\vec{a}_k} = 0$ when $j \neq k$, so the entries off the main diagonal of $A^T\overline{A}$ are all zero.

And thus, using a similar argument to the one we used for orthogonal matrices in the reals, we see that A is a unitary matrix if and only if A^T is the inverse of $\overline{A}$.

A different but related fact that is sometimes used is that A is a unitary matrix if and only if $\overline{A}^T$ is the inverse of A. It turns out that the matrix $\overline{A}^T$ is used in complex numbers a lot. This is not a surprise, since it is a blend of the common transpose action from the reals with the common conjugate action necessary in the complex numbers. So, we go ahead and give it its own symbol.

Definition: Let A be an $n \times n$ matrix with complex entries. We define the conjugate transpose A^* of A to be

$$A^* = \overline{A}^T$$

Unitary Matrices

Example

$$A = \begin{bmatrix} 3+7i & -3+i & 4-4i \\ 1-i & 7i & 6+2i \\ 13 & 4+i & 3+9i \end{bmatrix} \Rightarrow \overline{A} = \begin{bmatrix} 3-7i & -3-i & 4+4i \\ 1+i & -7i & 6-2i \\ 13 & 4-i & 3-9i \end{bmatrix}$$

$$\Rightarrow \overline{A}^T = A^* = \begin{bmatrix} 3-7i & 1+i & 13 \\ -3-i & -7i & 4-i \\ 4+4i & 6-2i & 3-9i \end{bmatrix}$$

Note: It does not matter whether we do the conjugate first and then the transpose (as the definition states) or the transpose first and then the conjugate (so we also have $A^* = \overline{A^T}$).

$$A = \begin{bmatrix} 3+7i & -3+i & 4-4i \\ 1-i & 7i & 6+2i \\ 13 & 4+i & 3+9i \end{bmatrix} \Rightarrow A^T = \begin{bmatrix} 3+7i & 1-i & 13 \\ -3+i & 7i & 4+i \\ 4-4i & 6+2i & 3+9i \end{bmatrix}$$

$$\Rightarrow \overline{A^T} = A^* = \begin{bmatrix} 3-7i & 1+i & 13 \\ -3-i & -7i & 4-i \\ 4+4i & 6-2i & 3-9i \end{bmatrix}$$

In fact, once you become comfortable with this process, you usually just do both actions at the same time.

$$B = \begin{bmatrix} 3i & 1-i \\ 2+5i & 6+7i \\ 0 & 3+6i \end{bmatrix} \Rightarrow B^* = \begin{bmatrix} -3i & 2-5i & 0 \\ 1+i & 6-7i & 3-6i \end{bmatrix}$$

Unitary Matrices

Theorem 9.5.3

If U is an $n \times n$ matrix, then the following are equivalent:

1. The columns of U form an orthonormal basis for $\mathbb{C}^n$
2. The rows of U form an orthonormal basis for $\mathbb{C}^n$
3. $U^{-1} = U^*$

Unitary Matrices

Proof

Let's prove '1 if and only if 3' first, and to that end, let's look U^*U.

The jk-th entry of U^*U is the dot product of the j-th row of U^* with the k-th column of U. Since the j-th row of U^* is the conjugate of the j-th column of U, we have that the jk-th entry of $\overline{\vec{u}_j} \cdot \vec{u}_k$.

This is the same as

$$\overline{\vec{u}_j \cdot \vec{u}_k}$$

which is

$$\overline{\langle \vec{u}_j, \vec{u}_k \rangle}$$

If the columns of U form an orthonormal basis, we know that $\langle \vec{u}_j, \vec{u}_k \rangle$ equals either

$$0 \text{ (if } j \neq k) \quad \text{or} \quad 1 \text{ (if } j = k)$$

both of which are real numbers. This means that they equal their conjugate, so we have shown that the jk-th entry of U^*U is

$$1 \text{ if } j = k \text{ and } 0 \text{ if } j \neq k$$

And this means that $U^*U = I$, so $U^{-1} = U^*$. So, we have shown that $U^{-1} = U^*$ if U is unitary.

Unitary Matrices

Proof

To show the reverse direction of our if and only if statement, let's assume that

$$U^{-1} = U^*$$

and show that the columns of U must form an orthonormal basis. Well, we still know that the jk-th entry of U^*U is

$$\overline{\langle \vec{u}_j, \vec{u}_k \rangle}$$

But since we know that

$$U^*U = I$$

we see that

$$\overline{\langle \vec{u}_j, \vec{u}_k \rangle} = 0 \text{ for } j \neq k$$

while

$$\overline{\langle \vec{u}_j, \vec{u}_j \rangle} = 1$$

As these are real numbers, the equal their conjugate, and so we see that

$$\langle \vec{u}_j, \vec{u}_k \rangle = 0 \text{ for } j \neq k$$

while

$$\langle \vec{u}_j, \vec{u}_j \rangle = 1$$

This means that the columns of U form an orthonormal basis for $\mathbb{C}^n$.

Unitary Matrices

Proof

$$\langle \vec{u}_j, \vec{u}_j \rangle = 1$$

We will use this result to show '2 if and only if 1': that the rows of U form and orthonormal basis for $\mathbb{C}^n$ if and only if the columns of U form an orthonormal basis for $\mathbb{C}^n$.

We begin by noting that the rows of U form an orthonormal basis if and only if the columns of U^T form an orthonormal basis. And we now know that the columns of U^T form an orthonormal basis if and only if

$$(U^T)^* = (U^T)^{-1}$$

So, what is the jk-th entry of $U^T(U^T)^*$?

Well, it is the dot product of the j-th row of U^T with the k-th column of $(U^T)^*$. Well, the j-th row of U^T is the j-th column of U, while the k-th column of $(U^T)^*$ is the conjugate of the k-th row of U^T.

Thus, the k-th column of $(U^T)^*$ is the conjugate of the k-th column of U. So, the jk-th entry of $U^T(U^T)^*$ is

$$\vec{u}_j \cdot \overline{\vec{u}_k} = \langle \vec{u}_j, \vec{u}_k \rangle$$

This means that $U^T(U^T)^* = I$ if and only if $\langle \vec{u}_j, \vec{u}_k \rangle$ equals 1 when $j = k$ and equals 0 when $j \neq k$, which means that $(U^T)^* = (U^T)^{-1}$ if and only if the columns of U form an orthonormal basis. So we have shown that U^T is unitary if and only if U is unitary. This proves that the rows of U form an orthonormal basis for C^n if and only if the columns of U form an orthonormal basis.

Unitary Matrices

Example

Let $A = \begin{bmatrix} \frac{i}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-i}{\sqrt{2}} \end{bmatrix}$. To determine if A is unitary, we want to look at the product A^*A.

$$A^*A = \begin{bmatrix} -i/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & i/\sqrt{2} \end{bmatrix} \begin{bmatrix} i/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & -i/\sqrt{2} \end{bmatrix} = \begin{bmatrix} (-i^2+1)/2 & (i-i)/2 \\ (-i+i)/2 & (1-i^2)/2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Since $A^*A = I$, Theorem 9.5.3 tells us that A is unitary.

Let $B = \begin{bmatrix} \frac{i}{\sqrt{3}} & \frac{-1}{\sqrt{3}} \\ \frac{1+i}{\sqrt{3}} & \frac{1+i}{\sqrt{3}} \end{bmatrix}$. To determine if B is unitary, we look at the product B^*B.

$$\begin{aligned} B^*B &= \begin{bmatrix} -i/\sqrt{3} & (1-i)/\sqrt{3} \\ -1/\sqrt{3} & (1-i)/\sqrt{3} \end{bmatrix} \begin{bmatrix} i/\sqrt{3} & -1/\sqrt{3} \\ (1+i)/\sqrt{3} & (1+i)/\sqrt{3} \end{bmatrix} \\ &= \begin{bmatrix} (-i^2+1+i-i-i^2)/3 & (i+1+i-i-i^2)/3 \\ (-i+1+i-i-i^2)/3 & (1+1+i-i-i^2)/3 \end{bmatrix} \\ &= \begin{bmatrix} 1 & (2+i)/3 \\ (2-i)/3 & 1 \end{bmatrix} \end{aligned}$$

Unitary Matrices

Example

Since $B^*B \neq I$, Theorem 9.5.3 tells us that B is not unitary. But we actually know even more. Since the jk-th entry of U^*U is $\overline{\langle \vec{u}_j, \vec{u}_k \rangle}$ for any complex matrix U, we see from B^*B that

$$\overline{\langle \vec{b}_1, \vec{b}_1 \rangle} = 1, \text{ and } \overline{\langle \vec{b}_2, \vec{b}_2 \rangle} = 1$$

so both columns of B are unit vectors. Moreover, we see that

$$\overline{\langle \vec{b}_1, \vec{b}_2 \rangle} \neq 0 \text{ (and } \overline{\langle \vec{b}_2, \vec{b}_1 \rangle} \neq 0)$$

so the reason that the columns of B fail to form an orthonormal basis is that the columns are not orthogonal to each other.

For one last example, let $C = \begin{bmatrix} 1+i & 1 & 1 \\ 0 & 1-i & -1+i \\ -1+i & -i & -i \end{bmatrix}$. Again, to see if C is unitary, we will look at C^*C.

Unitary Matrices

Example

$$\begin{aligned} C^*C &= \begin{bmatrix} 1-i & 0 & -1-i \\ 1 & 1+i & i \\ 1 & -1-i & i \end{bmatrix} \begin{bmatrix} 1+i & 1 & 1 \\ 0 & 1-i & -1+i \\ -1+i & -i & -i \end{bmatrix} \\ &= \begin{bmatrix} 1+i-i-i^2+1-i+i-i^2 & 1-i+i+i^2 & 1-i+i+i^2 \\ 1+i-i+i^2 & 1+1-i+i-i^2-i^2 & 1-2+i-i+i^2-i^2 \\ 1+i-i+i^2 & 1-1+i-i+i^2-i^2 & 1+1-i+i-i^2-i^2 \end{bmatrix} \\ &= \begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix} \end{aligned}$$

Again, C is not unitary, but again, we can use the product C^*C to learn why the columns of C do not form an orthonormal basis. This time, we have that

$$\overline{\langle \vec{c}_j, \vec{c}_k \rangle} = 0$$

when $j \neq k$, so we know that the columns of C are orthogonal. But we have that $\overline{\langle \vec{c}_j, \vec{c}_j \rangle} = 4 \neq 1$ so the columns of C are not unit vectors. But this can be easily changed by simply dividing the columns by their length, which we have already calculated to be $\sqrt{4} = 2$. So, as shown below, the matrix is unitary.

$$D = \begin{bmatrix} \frac{(1+i)}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{(1-i)}{2} & \frac{(-1+i)}{2} \\ \frac{(-1+i)}{2} & \frac{-i}{2} & \frac{-i}{2} \end{bmatrix}$$

Unitary Matrices

Theorem 9.5.2

Let A and B be complex matrices and let $\alpha \in \mathbb{C}$. Then

1. $\langle A\vec{z}, \vec{w}\rangle = \langle \vec{z}, A^*\vec{w}\rangle$ for all $\vec{z}, \vec{w} \in \mathbb{C}^n$
2. $A^{**} = A$
3. $(A+B)^* = A^* + B^*$
4. $(\alpha A)^* = \overline{\alpha} A^*$
5. $(AB)^* = B^*A^*$

I'll prove properties 1 and 2, and leave the others for you to do as a practice problem.

Proof

For Theorem 9.5.2, property 1: Actually, this property is a bit harder to prove than the others. You can expand all the dot products involved to determine the terms you are summing, but we get a neater proof if we think about it differently.

And that is to think of the dot product as a form of matrix multiplication, since for any vectors $\vec{z}$ and $\vec{w}$ in $\mathbb{C}^n$, we have that $\vec{z} \cdot \vec{w}$ is the same as the matrix product $\vec{z}^T\vec{w}$. With that in mind, we see the following:

$$\begin{aligned}\langle A\vec{z}, \vec{w}\rangle &= A\vec{z} \cdot \overline{\vec{w}} = (A\vec{z})^T\overline{\vec{w}} = \vec{z}^T A^T\overline{\vec{w}} \\ &= \vec{z}^T\overline{\overline{A^T}\vec{w}} = \vec{z}^T\overline{(A^T\vec{w})} = \vec{z}^T\overline{A^*\vec{w}} = \vec{z} \cdot \overline{A^*\vec{w}} \\ &= \langle \vec{z}, A^*\vec{w}\rangle\end{aligned}$$

Unitary Matrices

Proof

Note: In this proof, I made use of the fact that for any $n \times n$ matrices A and B with complex entries, we have that

$$\overline{A}\,\overline{B} = \overline{(AB)}$$

This follows from the fact that matrix multiplication simply involves multiplying and adding complex numbers, and we have already seen that for any $z, w \in \mathbb{C}$,

$$\overline{z} + \overline{w} = \overline{z+w} \text{ and } \overline{z}\,\overline{w} = \overline{zw}$$

For the same reason, we also know that

$$\overline{\vec{z}} \cdot \overline{\vec{w}} = \overline{\vec{z} \cdot \vec{w}} \text{ for any vectors } \vec{z}, \vec{w} \in \mathbb{C}^n$$

For Theorem 9.5.2, property 2:

Let $B = A^*$ and $C = B^* = A^{**}$. Then

$$c_{jk} = \overline{b_{kj}} = \overline{\overline{a_{jk}}} = a_{jk}, \text{ for all } 1 \leq j, k \leq n$$

Thus, $C = A$, that is, $A^{**} = A$.