Hermitian Inner Products and Unitary Matrices

We would like to have an inner product defined for complex vector spaces, because the concepts of length, orthogonality, and projections are powerful tools for solving certain problems. Moreover, it is a necessary step towards our goal of mimicking orthogonal diagonalization in the complex case.
How are we going to define an inner product on a complex vector space?
We will essentially repeat how we figured out how to define an inner product on a real vector space.

Definition: Let $\vec{w} = \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}, \vec{z} = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} \in \mathbb{C}^n$. We define

$$\vec{z} \cdot \vec{w} = z_1 w_1 + \cdots + z_n w_n$$

Does this define an inner product on $\mathbb{C}^n$?
Let $\vec{z} = \vec{x} + i\vec{y} \in \mathbb{C}^n$, then we have

$$\vec{z} \cdot \vec{z} = z_1^2 + \cdots + z_n^2 = (x_1^2 + \cdots + x_n^2 - y_1^2 - \cdots - y_n^2) + 2i(x_1 y_1 + \cdots + x_n y_n)$$

To match the real case, we will want to define the length of a vector $\vec{z}$ to be the square root of the inner product of $\vec{z}$ with itself. In this case, we observe that $\vec{z} \cdot \vec{z}$ does not even need to be a real number and so the condition $\vec{z} \cdot \vec{z} \geq 0$ does not even make sense.
Thus, **we cannot use the dot product as a rule for defining an inner product on** $\mathbb{C}^n$.

Hermitian Inner Products and Unitary Matrices

Recall that if $z = a + bi \in \mathbb{C}$, then $z\bar{z} = a^2 + b^2$.
Hence, we make the following definition.

Definition: The standard Hermitian inner product for $\mathbb{C}^n$ is defined by

$$\langle \vec{z}, \vec{w} \rangle = \vec{z} \cdot \overline{\vec{w}} = z_1 \overline{w_1} + \cdots + z_n \overline{w_n}, \qquad \text{for } \vec{w}, \vec{z} \in \mathbb{C}^n$$

Note: This is the mathematics definition of the standard Hermitian inner product on $\mathbb{C}^n$. In engineering, they use $\langle \vec{z}, \vec{w} \rangle = \overline{\vec{z}} \cdot \vec{w}$. In this course you are always expected to use the math definition. Most computer programs like Maple and MATLAB use the engineering definition, and so will not necessarily give you correct answers for this course.

Hermitian Inner Products and Unitary Matrices

Example

Let $\vec{z} = \begin{bmatrix} 1+i \\ 2-i \end{bmatrix}, \vec{w} = \begin{bmatrix} -2+i \\ 3+2i \end{bmatrix} \in \mathbb{C}^2$. Determine $\langle \vec{z}, \vec{w} \rangle$ and $\langle \vec{w}, \vec{z} \rangle$.

Solution

We have

$$\langle \vec{z}, \vec{w} \rangle = \vec{z} \cdot \overline{\vec{w}} = \begin{bmatrix} 1+i \\ 2-i \end{bmatrix} \cdot \begin{bmatrix} -2-i \\ 3-2i \end{bmatrix} = (1+i)(-2-i) + (2-i)(3-2i) = 3-10i$$
$$\langle \vec{w}, \vec{z} \rangle = \vec{w} \cdot \overline{\vec{z}} = \begin{bmatrix} -2+i \\ 3+2i \end{bmatrix} \cdot \begin{bmatrix} 1-i \\ 2+i \end{bmatrix} = (-2+i)(1-i) + (3+2i)(2+i) = 3+10i$$

Theorem 11.4.1

If $\vec{v}, \vec{z}, \vec{w} \in \mathbb{C}^n$ and $\alpha \in \mathbb{C}$, then

1. $\langle \vec{z}, \vec{z} \rangle \in \mathbb{R}$, $\langle \vec{z}, \vec{z} \rangle \geq 0$, and $\langle \vec{z}, \vec{z} \rangle = 0$ if and only if $\vec{z} = \vec{0}$.
2. $\langle \vec{z}, \vec{w} \rangle = \overline{\langle \vec{w}, \vec{z} \rangle}$
3. $\langle \vec{v} + \vec{z}, \vec{w} \rangle = \langle \vec{v}, \vec{w} \rangle + \langle \vec{z}, \vec{w} \rangle$
4. $\langle \alpha\vec{z}, \vec{w} \rangle = \alpha\langle \vec{z}, \vec{w} \rangle$

The proof is left as an exercise.

Hermitian Inner Products and Unitary Matrices

Definition: Let $\mathbb{V}$ be a vector space over $\mathbb{C}$. A Hermitian inner product on $\mathbb{V}$ is a function $\langle\ ,\ \rangle : \mathbb{V} \times \mathbb{V} \to \mathbb{C}$ such that for all $\vec{v}, \vec{w}, \vec{z} \in \mathbb{V}$ and $\alpha \in \mathbb{C}$ we have

1. $\langle \vec{z}, \vec{z} \rangle \in \mathbb{R}$, $\langle \vec{z}, \vec{z} \rangle \geq 0$, and $\langle \vec{z}, \vec{z} \rangle = 0$ if and only if $\vec{z} = \vec{0}$
2. $\langle \vec{z}, \vec{w} \rangle = \overline{\langle \vec{w}, \vec{z} \rangle}$
3. $\langle \vec{v} + \vec{z}, \vec{w} \rangle = \langle \vec{v}, \vec{w} \rangle + \langle \vec{z}, \vec{w} \rangle$
4. $\langle \alpha\vec{z}, \vec{w} \rangle = \alpha\langle \vec{z}, \vec{w} \rangle$

A complex vector space with a Hermitian inner product is called a Hermitian inner product space.

Notes:

1. The second property $\langle \vec{z}, \vec{w} \rangle = \overline{\langle \vec{w}, \vec{z} \rangle}$ is called the Hermitian property. We will see throughout the rest of the course, that "Hermitian" is the complex equivalent of "symmetric".
2. Observe that the second property only tells us how to factor out a complex scalar from the first entry of the inner product. Let's figure out how to factor out a scalar from the second component. We have
$$\langle \vec{z}, \alpha\vec{w} \rangle = \overline{\langle \alpha\vec{w}, \vec{z} \rangle} = \overline{\alpha\langle \vec{w}, \vec{z} \rangle} = \overline{\alpha}\,\overline{\langle \vec{w}, \vec{z} \rangle} = \overline{\alpha}\langle \vec{z}, \vec{w} \rangle$$
So, a Hermitian inner product is also not bilinear.
3. Since a Hermitian inner product is not symmetric or bilinear, you may wonder if it is a true generalization of the real inner product. Observe that if t and $\langle \vec{z}, \vec{w} \rangle$ are both real, then we have
$$\langle \vec{z}, \vec{w} \rangle = \overline{\langle \vec{w}, \vec{z} \rangle} = \langle \vec{w}, \vec{z} \rangle \text{ and } \langle \vec{z}, t\vec{w} \rangle = \overline{t}\langle \vec{z}, \vec{w} \rangle = t\langle \vec{z}, \vec{w} \rangle$$
so, it is symmetric and bilinear. In particular, we could actually use this as the definition for a real inner product.

Hermitian Inner Products and Unitary Matrices

Example

Theorem 11.4.1 shows that the standard Hermitian inner product on $\mathbb{C}^n$ is in fact a Hermitian inner product.

Example

Consider the complex vector space $M_{m\times n}(\mathbb{C})$.
How should we define the standard Hermitian inner product on this vector space?
We saw with real vector spaces that we defined $\langle A, B\rangle = \mathrm{tr}(B^T A)$ and that this inner product was equivalent to the dot product on $\mathbb{R}^{mn}$. Of course, we want to define the standard Hermitian inner product on $M_{m\times n}(\mathbb{C})$ in a similar way.
Because of the way the Hermitian inner product is defined on $\mathbb{C}^n$, we see that we also need to take a complex conjugate of the second component.
Thus, we define the standard Hermitian inner product $\langle\ ,\ \rangle$ on $M_{m\times n}(\mathbb{C})$ by

$$\langle Z, W\rangle = \mathrm{tr}(\overline{W}^T Z)$$

for all $Z, W \in M_{m\times n}(\mathbb{C})$.

Note: As we did in the real case, whenever we use $\mathbb{C}^n$ or $M_{m\times n}(\mathbb{C})$ we mean with the standard Hermitian inner product unless specified otherwise.

Hermitian Inner Products and Unitary Matrices

Example

Let $Z = \begin{bmatrix} 2+i & 1 \\ i & 1-i \end{bmatrix}, W = \begin{bmatrix} 3 & 2-3i \\ -2i & 1+2i \end{bmatrix} \in M_{2\times 2}(\mathbb{C})$. Find $\langle Z, W\rangle$.

Solution

We have

$$\langle Z, W\rangle = (2+i)(3) + 1(2+3i) + i(2i) + (1-i)(1-2i) = 5+3i$$

Hermitian Inner Products and Unitary Matrices

We can now define length and orthogonality to match what we did in the real case.

Definition: Let $\mathbb{V}$ be a Hermitian inner product space with Hermitian inner product $\langle\ ,\ \rangle$. For any $\vec{z} \in \mathbb{V}$ we define the length of $\vec{z}$ by

$$\|\vec{z}\| = \sqrt{\langle \vec{z}, \vec{z} \rangle}$$

For any $\vec{z}, \vec{w} \in \mathbb{V}$, we say that $\vec{w}$ and $\vec{z}$ are orthogonal if $\langle \vec{z}, \vec{w} \rangle = 0$.
A set $\mathcal{B} = \{\vec{z}_1, \dots, \vec{z}_\ell\} \subset \mathbb{V}$ is said to be orthogonal if $\langle \vec{z}_j, \vec{z}_k \rangle = 0$ whenever $j \neq k$, and $\mathcal{B}$ is said to be orthonormal if it is orthogonal and $\|\vec{z}_j\| = 1$ for all $1 \leq j \leq \ell$.

Hermitian Inner Products and Unitary Matrices

Theorem 11.4.2

Let $\mathbb{V}$ be a Hermitian inner product space with Hermitian inner product $\langle\ ,\ \rangle$. For any $\vec{z}, \vec{w} \in \mathbb{V}$ and $\alpha \in \mathbb{C}$ we have

$$\|\alpha\vec{z}\| = |\alpha|\|\vec{z}\|$$
$$\|\vec{z} + \vec{w}\| \leq \|\vec{z}\| + \|\vec{w}\|$$

Theorem 11.4.3

If $\{\vec{z}_1, \dots, \vec{z}_k\}$ is an orthonormal set in a Hermitian inner product space, then $\{\vec{z}_1, \dots, \vec{z}_k\}$ is linearly independent and

$$\|\vec{z}_1 + \cdots + \vec{z}_k\|^2 = \|\vec{z}_1\|^2 + \cdots + \|\vec{z}_k\|^2$$

Note: All of our concepts and theory for orthogonal complements, coordinates with respect to an orthonormal basis, and projections are the same except we now must be careful because the inner product is no longer symmetric. That is, we must make sure that we get the vectors in the inner products in the correct order.

Hermitian Inner Products and Unitary Matrices

Example

Use the Gram-Schmidt procedure to transform the basis vectors $\vec{z}_1 = \begin{bmatrix} i \\ 0 \\ 1 \end{bmatrix}, \vec{z}_2 = \begin{bmatrix} 0 \\ 1 \\ i \end{bmatrix}, \vec{z}_3 = \begin{bmatrix} -1 \\ 1 \\ i \end{bmatrix}$ into an orthonormal basis for $\mathbb{C}^3$.

Solution

Let $\vec{w}_1 = \vec{z}_1 = \begin{bmatrix} i \\ 0 \\ 1 \end{bmatrix}$, then

$$\vec{z}_2 - \frac{\langle \vec{z}_2, \vec{w}_1 \rangle}{\|\vec{w}_1\|^2} \vec{w}_1 = \begin{bmatrix} 0 \\ 1 \\ i \end{bmatrix} - \frac{i}{2} \begin{bmatrix} i \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1 \\ i/2 \end{bmatrix}$$

So, for simplicity, we pick $\vec{w}_2 = \begin{bmatrix} 1 \\ 2 \\ i \end{bmatrix}$.

$$\vec{w}_3 = \vec{z}_3 - \frac{\langle \vec{z}_3, \vec{w}_1 \rangle}{\|\vec{w}_1\|^2} \vec{w}_1 - \frac{\langle \vec{z}_3, \vec{w}_2 \rangle}{\|\vec{w}_2\|^2} \vec{w}_2 = \begin{bmatrix} -1 \\ 1 \\ i \end{bmatrix} - \frac{2i}{2} \begin{bmatrix} i \\ 0 \\ 1 \end{bmatrix} - \frac{2}{6} \begin{bmatrix} 1 \\ 2 \\ i \end{bmatrix} = \begin{bmatrix} -1/3 \\ 1/3 \\ -i/3 \end{bmatrix}$$

Normalizing the vectors we get the orthonormal basis $\left\{ \begin{bmatrix} i/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}, \begin{bmatrix} 1/\sqrt{6} \\ 2/\sqrt{6} \\ i/\sqrt{6} \end{bmatrix}, \begin{bmatrix} -1/\sqrt{3} \\ 1/\sqrt{3} \\ -i/\sqrt{3} \end{bmatrix} \right\}$.

Hermitian Inner Products and Unitary Matrices

Definition: A matrix $U \in M_{n\times n}(\mathbb{C})$ is said to be unitary if its columns form an orthonormal basis for $\mathbb{C}^n$.

Notice that if $P \in M_{n\times n}(\mathbb{R})$ is orthogonal, then P is also unitary. So, unitary is the direct extension of orthogonal.

Theorem 11.4.4

If $U \in M_{n\times n}(\mathbb{C})$, then the following are equivalent:

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

The proof is essentially the same as the corresponding theorem for orthogonal matrices and so is omitted.

Hermitian Inner Products and Unitary Matrices

Theorem 11.4.5

If U_1 and U_2 are $n \times n$ unitary matrices, then

1. $\|U_1\vec{z}\| = \|\vec{z}\|$ for all $\vec{z} \in \mathbb{C}^n$,
2. $|\det U_1| = 1$, and
3. U_1U_2 is unitary.

I will prove 1 and leave the other two as exercises.

Proof of 1

For any $\vec{z} \in \mathbb{C}^n$, we have

$$\begin{aligned} \|U_1\vec{z}\|^2 &= \langle U_1\vec{z}, U_1\vec{z}\rangle \\ &= (U_1\vec{z}) \cdot \overline{(U_1\vec{z})} \\ &= (U_1\vec{z})^T\overline{(U_1\vec{z})} \\ &= \vec{z}^T U_1^T \overline{U_1}\,\overline{\vec{z}} \\ &= \vec{z}^T \overline{\overline{U_1^T}}\,\overline{U_1}\,\overline{\vec{z}} \\ &= \vec{z}^T \overline{\overline{U_1}^T U_1}\,\overline{\vec{z}} \\ &= \vec{z}^T \overline{I}\,\overline{\vec{z}} \\ &= \vec{z}^T\overline{\vec{z}} \\ &= \langle \vec{z}, \vec{z}\rangle \\ &= \|\vec{z}\|^2 \end{aligned}$$

□

Also, it is important to remember that we are now using complex numbers. So, the fact that $|\det U| = 1$ means that $\det U$ can be any complex number with absolute value 1 (like 1, -1, i, $-i$, $\frac{1+i}{\sqrt{2}}$, etc.).

Hermitian Inner Products and Unitary Matrices

Definition: Let $A \in M_{m\times n}(\mathbb{C})$. We call $\overline{A}^T$ the conjugate transpose of A and denote it

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

Example

If $A = \begin{bmatrix} 1 & -2 \\ -i & 1+3i \\ 4-2i & 5i \end{bmatrix}$, then $A^* = \begin{bmatrix} 1 & i & 4+2i \\ -2 & 1-3i & -5i \end{bmatrix}$.

Hermitian Inner Products and Unitary Matrices

Notice that if A has all real entries, then $A^* = A^T$.

Thus, A^* is the extension of the transpose to the complex case.

So, it should not be surprising that the conjugate transpose has many of the same properties.

Theorem 11.4.6

If A and B are complex matrices and $\alpha \in \mathbb{C}$, 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. $(AB)^* = B^* A^*$

Proof of 1

We have

$$\langle A\vec{z}, \vec{w}\rangle = (A\vec{z})^T \overline{\vec{w}} = \vec{z}^T A^T \overline{\vec{w}} = \vec{z}^T \overline{\overline{A^T}}\, \overline{\vec{w}} = \vec{z}^T \overline{A^* \vec{w}} = \langle \vec{z}, A^*\vec{w}\rangle$$

□