EECS 16B Designing Information Devices and Systems II
Fall 2021 Note 2j: Complex Inner Products

1 Real Projections

To understand how we want to define complex inner products, it is useful to first recall how we came to define real inner products. Inner products grow out of our desire to do projections. Projections themselves are intimately connected to the idea of orthogonality.

Let's do as we have done in many other analyses, and start with the simplest case first – namely, that of projecting onto a single vector, or more precisely projecting onto its span.

1.1 Proposing a Real Projection Operator

When projecting a real vector $\vec{v}$ onto the span of another real vector $\vec{u}$, the result needs to be $r\vec{u}$ where r is some real constant. In other words, we want a vector that is linearly dependent with $\vec{u}$ so that it captures all of $\vec{v}$ that is in the direction of $\vec{u}$.

Because the idea of direction is so important, we can first focus on the distilled embodiments of directions themselves – namely unit vectors. Vectors whose length is 1 essentially are all about direction since their magnitude/length is known.

If we have a unit vector $\hat{\vec{u}} \in \mathbb{R}^n$, recall that we can define the following operator:

$$P_{\hat{\vec{u}}} = \hat{\vec{u}}\hat{\vec{u}}^\top \tag{1}$$

that acts on vectors in $\mathbb{R}^n$ and also spits out vectors in $\mathbb{R}^n$. This means, for any vector $\vec{v} \in \mathbb{R}^n$, we can apply the operator $P_{\hat{\vec{u}}}$ to $\vec{v}$ by computing $P_{\hat{\vec{u}}}\vec{v}$:

$$P_{\hat{\vec{u}}}\vec{v} = \hat{\vec{u}}\hat{\vec{u}}^\top \vec{v} \tag{2}$$

$$= \left\langle \vec{v}, \hat{\vec{u}} \right\rangle \hat{\vec{u}}. \tag{3}$$

where $\langle \vec{x}, \vec{y} \rangle = \vec{y}^\top \vec{x} = \sum_{i=1}^n x_i y_i$ is the inner product on $\mathbb{R}^n$ we first found out about in 16A.

The operator we computed above gives us a vector that is in the direction of $\hat{\vec{u}}$ and is a multiple $\left\langle \vec{v}, \hat{\vec{u}} \right\rangle$ of $\hat{\vec{u}}$. So far so good; this at least seems like a plausible projection.

Now, given any vector $\vec{u} \in \mathbb{R}^n$, we can turn it into a unit vector by normalizing it to get $\frac{\vec{u}}{\|\vec{u}\|}$. This motivates our extension of the previous operator to general vectors $\vec{u}$. Indeed, we define the operator

$$P_{\vec{u}} = \left(\frac{\vec{u}}{\|\vec{u}\|}\right)\left(\frac{\vec{u}}{\|\vec{u}\|}\right)^\top = \frac{\vec{u}\vec{u}^\top}{\|\vec{u}\|^2} \tag{4}$$

that acts on vectors in $\mathbb{R}^n$ and also spits out vectors in $\mathbb{R}^n$. Repeating a similar calculation as before, we can apply the operator $P_{\vec{u}}$ to $\vec{v}$ by computing $P_{\vec{u}}\vec{v}$:

$$P_{\vec{u}}\vec{v} = \frac{\vec{u}\vec{u}^\top}{\|\vec{u}\|^2}\vec{v} \tag{5}$$

$$= \frac{\langle \vec{v}, \vec{u} \rangle}{\|\vec{u}\|^2}\vec{u} \tag{6}$$

$$= \frac{\langle \vec{v}, \vec{u} \rangle}{\langle \vec{v}, \vec{v} \rangle}\vec{u}. \tag{7}$$

Concept Check: With our given definition of $\langle \vec{x}, \vec{y} \rangle$, and your knowledge of the norm we use for real vectors $\|\vec{x}\|$, show that the identity we used in the last line holds: namely, that $\|\vec{u}\|^2 = \langle \vec{u}, \vec{u} \rangle$.

1.2 Verifying the Proposed Projection

Now how do we show that the operator we have defined is a projection? Well, in 16A Note 23 we used the Pythagorean Theorem to prove that the point our projection operator $P_{\text{Col}(A)} = (A^\top A)^{-1}A^\top$ applied to our input $\vec{v}$ gave us was the closest point (in the sense of the familiar 2-norm $\|x\| = \sqrt{\sum_{i=1}^n x_i^2}$) to the point $\vec{v}$ in $\text{Col}(A)$.

We can do the same thing here and verify that our real projection works by using the Pythagorean theorem, but our ultimate goal is to work in the complex world. We'd better figure out how to appropriately generalize what we know so far to complex vector spaces, or else we're going to get stuck right out of the gate.

Actually, there is a lot that the Pythagorean theorem implicitly relies on, that holds fine in $\mathbb{R}^n$ but not necessarily in general vector spaces.

- The Pythagorean theorem first needs a notion of *length*, i.e., a *norm* $\|\cdot\|$. In $\mathbb{R}^n$, we use our familiar 2-norm. In the complex world, we don't have anything yet.
- Then the Pythagorean theorem needs a notion of *angle*, i.e., an *inner product* $\langle \cdot, \cdot \rangle$ (from which we can define the angle between $\vec{x}$ and $\vec{y}$ by the identity $\cos(\theta) = \frac{\langle \vec{x}, \vec{y} \rangle}{\|\vec{x}\|\|\vec{y}\|}$).
- Finally, we need the norm and the inner product to be *compatible* in some way, so that we can make connections between length and angle. That is, we require $\|\vec{x}\|^2 = \langle \vec{x}, \vec{x} \rangle$.[1]

If our vector space has these properties, then it becomes much simpler to work with! A lot of our geometric intuition in $\mathbb{R}^n$ carries over, and the Pythagorean theorem (among other results) applies. Thus we can use exactly the same techniques to verify our operators are projections, as we did for least squares.

To that end, one result that we derived from the Pythagorean Theorem in 16A and heavily used in the least squares derivation is that, for a correct projection operator $P_{\text{Col}(A)}$ onto $\text{Col}(A)$, our residual $\vec{x} - P_{\text{Col}(A)}\vec{x}$ has to be orthogonal[2] to every vector in $\text{Col}(A)$.

Now, we may raise the level of our abstraction, so long as we still have the geometric structure afforded by a compatible norm and inner product. In particular, we can replace $\mathbb{R}^n$ with a general vector space V (with an inner product and compatible norm) and $\text{Col}(A)$ with a general subspace U of V. In this form, the result has a well-known name.

[1] This type of norm is called an *induced* norm (induced by the inner product) or a Euclidean norm (because it allows Euclidean geometry to work).

[2] Remember the definition of orthogonality: two vectors $\vec{u}, \vec{v}$ are orthogonal if and only if $\langle \vec{u}, \vec{v} \rangle = 0$.

Theorem (Orthogonality Principle): If U is a subspace of a vector space V with an inner product $\langle \cdot, \cdot \rangle$ and compatible norm $\|\cdot\|$, then $y \in U$ is the closest point to $x \in V$ (i.e., y is the projection of x onto U) in the sense of the norm $\|\cdot\|$ if and only if $\langle x - y, z \rangle = 0$ for all $z \in U$.[3][4]

The orthogonality principle has has broad consequences and leads to deep theory (some of which will be included in the optional section at the end of the note), and yet we have already discovered it in 16AB!

After that small detour, let us return to the problem at hand. Using the notation of the orthogonality principle, we let $V = \mathbb{R}^n$, and also let $U = \operatorname{Span}(\vec{u}) = \{r\vec{u} \colon r \in \mathbb{R}\}$.

To show that $P_{\vec{u}}$ is a projection onto $\operatorname{Span}(\vec{u})$, it remains to check that $\langle P_{\vec{u}}\vec{v} - \vec{v}, r\vec{u} \rangle = 0$ for all $r \in \mathbb{R}$. But by bilinearity of the inner product, $\langle P_{\vec{u}}\vec{v} - \vec{v}, r\vec{u} \rangle = r \langle P_{\vec{u}}\vec{v} - \vec{v}, \vec{u} \rangle$. So it *really* suffices to show that $\langle P_{\vec{u}}\vec{v} - \vec{v}, \vec{u} \rangle = 0$. And we can make sure of that now:

$$\begin{aligned} \langle P_{\vec{u}}\vec{v} - \vec{v}, \vec{u} \rangle &= \left\langle \frac{\langle \vec{v}, \vec{u} \rangle}{\langle \vec{u}, \vec{u} \rangle}\vec{u} - \vec{v}, \vec{u} \right\rangle & (8) \\ &= \left\langle \frac{\langle \vec{v}, \vec{u} \rangle}{\langle \vec{u}, \vec{u} \rangle}\vec{u}, \vec{u} \right\rangle - \langle \vec{v}, \vec{u} \rangle & (9) \\ &= \frac{\langle \vec{v}, \vec{u} \rangle}{\langle \vec{u}, \vec{u} \rangle} \langle \vec{u}, \vec{u} \rangle - \langle \vec{v}, \vec{u} \rangle & (10) \\ &= \langle \vec{v}, \vec{u} \rangle - \langle \vec{v}, \vec{u} \rangle & (11) \\ &= 0. & (12) \end{aligned}$$

Here, we have done nothing except use the bilinearity of the inner product, a property that's also been well-used in 16A.

Hence, the orthogonality principle yields that $P_{\vec{u}}$ is a projection onto $\operatorname{Span}(\vec{u})$, just as we have desired.

2 Defining Complex Inner Product via Projections

We have just re-derived the real projection operator via the real inner product. Now we will do the reverse for complex inner products. That is, we will try to define a complex inner product by coming up with a reasonable projection operator for complex vectors, and then making a complex inner product to fit.

First, we need to discuss what exactly a complex vector is, and some preliminaries we will need before making our argument.

2.1 Fundamental Complex Definitions

An n-dimensional *complex vector* $\vec{v} \in \mathbb{C}^n$ is just like an n-dimensional real vector in $\mathbb{R}^n$ that we all know and love, *but* each of the n entries are complex numbers as opposed to just real numbers.

If $\vec{v}$ is a complex vector, we define its length, or norm, by

$$\|v\|^2 = \sum_{i=1}^{n} |v_i|^2 = \sum_{i=1}^{n} v_i \overline{v_i}. \tag{13}$$

Here, recall that $\overline{v_i}$ is the complex conjugate of the complex number v_i. Note that this is similar *but not*

[3] With some technical caveats that are not relevant to our use case. See the optional section at the end of the note for more.

[4] Why is it called the orthogonality principle? Well, all the theorem is saying is that the residual $x - y$ is orthogonal to anything in U.

exactly equal to the norm on real vectors defined by $\|\vec{v}\|^2 = \sum_{i=1}^n v_i^2$.

Concept Check: Prove that $v_i\overline{v_i} \neq v_i^2$ unless v_i is real. Hence, if $\vec{v}$ is a real vector, then the complex vector norm of $\vec{v}$ is exactly equal to the real vector norm of $\vec{v}$. (This is a common motif for our complex definitions – if we use them in the special case where our complex vector has real entries, they should reduce to the corresponding definitions for real vectors.)

A *complex unit vector* is a vector $\vec{v} \in \mathbb{C}^n$ which has unit norm, i.e., $\|\vec{v}\| = 1$. This is exactly analogous to real vectors with norm 1 being the regular unit vectors we're used to working with.

The *span* of a set of complex vectors, i.e., $\text{Span}(\vec{v}_1, \ldots, \vec{v}_m)$, is the set of complex vectors $\vec{z}$ such that $\vec{z} = c_1\vec{v}_1 + \cdots + c_m\vec{v}_m$ for some complex scalars $c_1, \ldots, c_m \in \mathbb{C}$. This is the same notion of span as with real vectors, except again we are allowed to use complex coefficients.

Finally, a set of complex vectors $\{\vec{v}_1, \ldots, \vec{v}_m\}$ are *linearly dependent* if there are complex scalars $c_1, \ldots, c_m \in \mathbb{C}$, not all zero, such that $c_1\vec{v}_1 + \cdots + c_m\vec{v}_m = \vec{0}$, and *linearly independent* otherwise. Again, this is the same definition as for real linear (in)dependence, but for the fact that we can use complex numbers as the coefficients.

2.2 Starting to Define Complex Projections

Again, let us consider the simplest case – that of projecting onto the span of a single complex vector. We begin, as always, by drawing an analogy with the real case.

We already know that two real vectors $\vec{v}_1$ and $\vec{v}_2$ are linearly dependent if $\vec{v}_1 \in \text{Span}(\vec{v}_2)$. That is, there is a real scalar r such that $\vec{v}_1 = r\vec{v}_2$. In this case, we can see that since $\vec{v}_1 \in \text{Span}(\vec{v}_2)$,

$$\vec{v}_1 = P_{\vec{v}_2}\vec{v}_1 = \frac{\langle \vec{v}_1, \vec{v}_2 \rangle}{\langle \vec{v}_2, \vec{v}_2 \rangle}\vec{v}_2. \tag{14}$$

(All the first equality is saying is that, the projection of a vector to a subspace which the vector is on, is just the vector itself.) Hence $r = \frac{\langle \vec{v}_1, \vec{v}_2 \rangle}{\langle \vec{v}_2, \vec{v}_2 \rangle}$.

Similarly, two complex vectors $\vec{v}_1$ and $\vec{v}_2$ are linearly dependent if $\vec{v}_1 \in \text{Span}(\vec{v}_2)$, i.e., there is a complex scalar $c \in \mathbb{C}$ such that $\vec{v}_1 = c\vec{v}_2$. Even though we haven't defined a projection yet for general complex vectors onto general complex vectors, we already know that c should be the projection coefficient[5] for the projection of $\vec{v}_1$ onto $\text{Span}(\vec{v}_2)$, given that $\vec{v}_1$ and $\vec{v}_2$ are linearly dependent. Note that we can *only* do this right now because $\vec{v}_1$ and $\vec{v}_2$ are linearly dependent, which is an atypical case.

if $\vec{v}_1 = c\vec{v}_2$, then $\vec{v}_2 = \frac{1}{c}\vec{v}_1$, just by dividing by c on both sides. This works for both the cases where $\vec{v}_1, \vec{v}_2$ are real (in which case we have previously used r instead of c in the note), and when they are complex.

For a *real* unit vector $\vec{v}$, the only real unit vectors that are linearly dependent (in the real sense) with $\vec{v}$ are $\vec{v}$ and $-\vec{v}$. The projection coefficients of the projection of $\vec{v}$ onto $-\vec{v}$ and $-\vec{v}$ onto $\vec{v}$ are both $r = -1$. This happens because -1 is its own reciprocal.

However, for a *complex* unit vector $\vec{v}$, there are many complex unit vectors that are linearly dependent (in the complex sense) with $\vec{v}$ – in particular, we know that vectors of the form $e^{j\theta}\vec{v}$, for $\theta \in [0, 2\pi)$, are linearly dependent with $\vec{v}$.

Concept Check: Show that if $\vec{v}$ is a complex unit vector then the *only* unit vectors linearly dependent with $\vec{v}$ are $e^{j\theta}\vec{v}$ for $\theta \in [0, 2\pi)$. [6] This means that we can completely characterize the unit vectors linearly

[5] Remember to not get *projection coefficients* (scalars) and *projections* (vectors) mixed up!

[6] This question is not saying the same as the above paragraph. The above paragraph asserts that vectors of the form $e^{j\theta}\vec{v}$ are unit

dependent with $\vec{v}$, as $c\vec{v}$ where $|c| = 1$.

Concept Check: Show that if c is a complex number with $|c| = 1$, then $\frac{1}{c} = \overline{c}$.

Thus if $\vec{v}_1$ and $\vec{v}_2$ are linearly dependent complex unit vectors with $\vec{v}_1 = c\vec{v}_2$, then $|c| = 1$, so that $\vec{v}_2 = \frac{1}{c}\vec{v}_1 = \overline{c}\vec{v}_1$. In short, the projection coefficients are complex conjugates.

Thus in the real vector case, the projection coefficients of linearly dependent unit vectors onto each other are identical; in the complex vector case they are complex conjugates. This demonstrates an important point – *in the complex world, order matters*.

2.3 Fully Defining a Complex Projection Operator

Now that we do not have symmetry (as to who is being projected onto whom), we need to be more careful in formulating a projection operator for complex vectors.

Consider a complex vector $\vec{u}$ and define the operator

$$P_{\vec{u}} = \frac{\vec{u}\vec{u}^*}{\|\vec{u}\|^2}. \tag{15}$$

Here $\vec{u}^* = \left(\overline{\vec{u}}\right)^\top$ is the conjugate transpose of the vector $\vec{u}$ – in other words, we take the complex conjugate of every entry in $\vec{u}$, and take the transpose of the result. For real vectors $\vec{u}$, this is the same as the projection operator we described above, because the complex conjugate would do nothing.

Concept Check: Show that $\vec{u}^*\vec{u} = \|\vec{u}\|^2$.

We claim that $P_{\vec{u}}$ is reasonable choice of projection operator. We can't use the orthogonality principle yet – to do this, we need a suitable inner product for complex numbers, and our whole goal was to build the inner product from projections. But we can do a couple "smell tests", and check that the projection behaves reasonably in simple cases. Note again that this doesn't actually prove that it's a projection – the way to do that is to use the orthogonality principle.

First, we would like to show that the projection coefficient of $\vec{u}$ onto $\vec{v}$ is the complex conjugate of the projection coefficient of $\vec{v}$ onto $\vec{u}$, for $\vec{u}, \vec{v}$ being unit complex vectors. We can do this by calculating:

$$P_{\vec{u}}\vec{v} = \frac{\vec{u}\vec{u}^*}{\|\vec{u}\|^2}\vec{v} \tag{16}$$

$$= \frac{\vec{u}^*\vec{v}}{\|\vec{u}\|^2}\vec{u} \tag{17}$$

$$= \left(\vec{u}^*\vec{v}\right)\vec{u} \tag{18}$$

$$P_{\vec{v}}\vec{u} = \frac{\vec{v}\vec{v}^*}{\|\vec{v}\|^2}\vec{u} \tag{19}$$

$$= \left(\vec{v}^*\vec{u}\right)\vec{v} \tag{20}$$

So the projection coefficients are $\vec{u}^*\vec{v}$ and $\vec{v}^*\vec{u}$.

Concept Check: Show that $\vec{u}^*\vec{v} = \overline{\vec{v}^*\vec{u}}$, and thus the projection coefficients are complex conjugates.

Note that this simplification only works when $\|\vec{u}\| = \|\vec{v}\| = 1$, i.e., $\vec{u}$ and $\vec{v}$ are unit vectors; otherwise the relationship between projection coefficients gets a little more complicated.

---

vectors that are linearly dependent with $\vec{v}$ – nowhere does it say these are the only such unit vectors. To show this is your task now.

Once we have finished that, another thing we can check is that if $\vec{u}$ and $\vec{v}$ are linearly dependent complex vectors, then $P_{\vec{u}}\vec{v} = \vec{v}$. This is the same idea that we had earlier, i.e., that the projection of a vector onto a subspace containing it should just be the vector itself. We can again do this by computing $P_{\vec{u}}\vec{v}$. To enforce linear dependence, let $\vec{v} = c\vec{u}$ for some $c \in \mathbb{C}$. Then

$$\begin{aligned}
P_{\vec{u}}\vec{v} &= \frac{\vec{u}\vec{u}^*}{\|\vec{u}\|^2}\vec{v} && (21)\\
&= \frac{\vec{u}^*\vec{v}}{\|\vec{u}\|^2}\vec{u} && (22)\\
&= \frac{\vec{u}^*(c\vec{u})}{\|\vec{u}\|^2}\vec{u} && (23)\\
&= c\frac{\vec{u}^*\vec{u}}{\|\vec{u}\|^2}\vec{u} && (24)\\
&= c\frac{\|\vec{u}\|^2}{\|\vec{u}\|^2}\vec{u} && (25)\\
&= c\vec{u} && (26)\\
&= \vec{v}. && (27)
\end{aligned}$$

Here we used the fact we asked you to prove in a concept check earlier – namely $\vec{u}^*\vec{u} = \|\vec{u}\|^2$.

So, our candidate for a projection operator passes the smell tests – in the simplest of cases, it behaves as we would want a projection operator to behave. However, to actually prove that it is a projection operator, we again need to invoke the orthogonality principle. For that, we need a suitable choice of inner product.

2.4 Defining the Inner Product

In the real case, remember that our projection operator has the form

$$P_{\vec{u}}\vec{v} = \frac{\vec{u}^\top\vec{v}}{\|\vec{u}\|^2}\vec{u} = \frac{\langle\vec{v},\vec{u}\rangle}{\langle\vec{u},\vec{u}\rangle}\vec{u}. \tag{28}$$

In the complex case, our proposed projection operator looks like:

$$P_{\vec{u}}\vec{v} = \frac{\vec{u}^*\vec{v}}{\|\vec{u}\|^2}\vec{u}. \tag{29}$$

Do you see anything interesting? Without bringing inner products into play, the only difference in the first expression is the use of transpose ($\top$) versus conjugate transpose ($*$)!

This also reveals one more requirement that our complex inner product should fulfill; in order to keep the theory nice and symmetric, we also would really like *in the complex case* that

$$P_{\vec{u}}\vec{v} = \frac{\langle\vec{v},\vec{u}\rangle}{\langle\vec{u},\vec{u}\rangle}\vec{u}. \tag{30}$$

Given these expressions, a natural way to define the complex inner product is to replace the transpose with the conjugate transpose in the real inner product, i.e., as

$$\langle\vec{v},\vec{u}\rangle = \vec{u}^*\vec{v}. \tag{31}$$

Note the swapping of the order of the arguments! *In the real case*, order doesn't matter – we could have easily defined $\langle \vec{u}, \vec{v} \rangle = \vec{u}^\top \vec{v}$, and that is exactly the same as our real-valued inner product, because if $\vec{u}$ and $\vec{v}$ are real then $\vec{u}^\top \vec{v} = \vec{v}^\top \vec{u}$. In the *complex* case, order *does* matter – $\vec{u}^* \vec{v} = \overline{\vec{v}^* \vec{u}}$. In particular, this means that $\langle \vec{u}, \vec{v} \rangle = \overline{\langle \vec{v}, \vec{u} \rangle}$ – and the complex inner product is no longer *symmetric* (like the real inner product) but is instead *conjugate symmetric*.

One key point to note is that we *could* have picked $\langle \vec{v}, \vec{u} \rangle$ to be equal to $\vec{v}^* \vec{u}$. The choice we make is due to popular convention, but *we are going to stick with our choice* of $\langle \vec{v}, \vec{u} \rangle = \vec{u}^* \vec{v}$. Mixing up the inner product order would mess up whatever calculation we're doing. A helpful mnemonic device to remember our inner product is "to project $\vec{v}$ onto $\vec{u}$, use $\langle \vec{v}, \vec{u} \rangle$" – and putting that with our definition of $P_{\vec{u}} \vec{v} = \frac{\vec{u}^* \vec{v}}{\|\vec{u}\|^2} \vec{u}$ gets you the correct definition of $\langle \vec{v}, \vec{u} \rangle = \vec{u}^* \vec{v}$ as desired.

Concept Check: Show that, by our definition of the complex inner product, if $\vec{u}$ and $\vec{v}$ are real vectors then the complex inner product of $\vec{u}$ and $\vec{v}$ is equal to the real inner product of $\vec{u}$ and $\vec{v}$.

Concept Check: Show that, by our definition of the complex inner product, $P_{\vec{u}} \vec{v} = \frac{\langle \vec{v}, \vec{u} \rangle}{\langle \vec{u}, \vec{u} \rangle} \vec{u}$.

Concept Check: Show that, by our definition of the complex inner product, $\|\vec{u}\|^2 = \langle \vec{u}, \vec{u} \rangle$.

2.5 Verifying Validity of Our Projection

Now that we finally have a complex inner product, we would like to verify the orthogonality principle for our proposed complex projection operator we defined earlier. To do that, we would like to first discuss what orthogonality and orthonormality mean in the complex world.

It turns out, they look more or less the same as their real-valued definitions. That is, vectors $\vec{u}, \vec{v}$ are *orthogonal* if and only if $\langle \vec{u}, \vec{v} \rangle = \vec{0}$. They are *orthonormal* if and only if they are orthogonal and $\|\vec{u}\| = \|\vec{v}\| = 1$, i.e., they are normalized. And we can extend our definitions of orthogonal/orthonormal to sets of vectors in the same way we did for the real case. That is, a set of vectors is orthogonal (orthonormal) if they are pairwise orthogonal (resp. orthonormal), e.g. for every pair of vectors in the set, the pair is orthogonal (resp. orthonormal).

Now we want to check that our operator $P_{\vec{u}} = \frac{\vec{u}\vec{u}^*}{\langle \vec{u}, \vec{u} \rangle}$ satisfies the orthogonality principle. Take $\vec{w} \in \mathrm{Span}(\vec{u})$, i.e., $\vec{w} = c\vec{u}$ for some scalar constant $c \in \mathbb{C}$. We want to show that $0 = \langle \vec{v} - P_{\vec{u}} \vec{v}, \vec{w} \rangle = \langle \vec{v} - P_{\vec{u}} \vec{v}, c\vec{u} \rangle$. But by expanding the definition of the complex inner product and manipulating the algebra, (try it yourself!), we see that $\langle \vec{v} - P_{\vec{u}} \vec{v}, c\vec{u} \rangle = \overline{c} \langle \vec{v} - P_{\vec{u}} \vec{v}, \vec{u} \rangle$.[7] So it *really* only suffices, again, to show that $\langle \vec{v} - P_{\vec{u}} \vec{v}, \vec{u} \rangle = 0$. And we can do this via the same computation:

$$\begin{aligned}
\langle \vec{v} - P_{\vec{u}} \vec{v}, \vec{u} \rangle &= \left\langle \vec{v} - \frac{\langle \vec{v}, \vec{u} \rangle}{\langle \vec{u}, \vec{u} \rangle} \vec{u}, \vec{u} \right\rangle && (32) \\
&= \langle \vec{v}, \vec{u} \rangle - \left\langle \frac{\langle \vec{v}, \vec{u} \rangle}{\langle \vec{u}, \vec{u} \rangle} \vec{u}, \vec{u} \right\rangle && (33) \\
&= \langle \vec{v}, \vec{u} \rangle - \frac{\langle \vec{v}, \vec{u} \rangle}{\langle \vec{u}, \vec{u} \rangle} \langle \vec{u}, \vec{u} \rangle && (34) \\
&= \langle \vec{v}, \vec{u} \rangle - \langle \vec{v}, \vec{u} \rangle && (35) \\
&= 0. && (36)
\end{aligned}$$

Thus, at long last, we have verified that our operator $P_{\vec{u}} = \frac{\vec{u}\vec{u}^*}{\langle \vec{u}, \vec{u} \rangle}$ is a projection operator (onto $\mathrm{Span}(\vec{u})$).

[7] Note the difference from the real case – if this were the real inner product, the coefficient would be not $\overline{c}$ but instead just c.

Hence we have developed, from nothing but our intuition for the real case, a reasonable – and compatible – complex projection and complex inner product. Very exciting!

3 Complex Matrix Computations

Our theory for complex vectors wouldn't be very useful if we couldn't use it in computations involving complex matrices. It turns out that, with very few changes in most of the proofs, one can develop everything from least squares to upper-triangularization using complex vectors/matrices analogously to the proofs we give in class for real vectors/matrices. The key is how similar our real and complex inner products are – if we change transpose to conjugate transpose everywhere (thus turning the real inner product into the complex inner product), our proofs suddenly should hold in full generality in the complex case.

To illustrate the ease of generalization to the complex case, we will do a sample computation which is relevant to the theory of the Discrete Fourier Transform, which we will study later in the class. It involves the manipulation of a particular matrix with orthonormal columns.

More specifically, we would want to show that if the columns of a square matrix M are orthonormal, the (conjugate) transpose of the matrix is its inverse.

Let us first try the real case for the sake of example. Here we are not concerned about conjugates. The proof goes as follows:

Let $M = \begin{bmatrix} \vec{v}_1 & \cdots & \vec{v}_n \end{bmatrix}$ be an $n \times n$ square matrix with orthonormal columns and real entries. Then

$$\begin{aligned}
M^\top M &= \begin{bmatrix} \vec{v}_1^\top \\ \vdots \\ \vec{v}_n^\top \end{bmatrix} \begin{bmatrix} \vec{v}_1 & \cdots & \vec{v}_n \end{bmatrix} \\
&= \begin{bmatrix} \vec{v}_1^\top \vec{v}_1 & \cdots & \vec{v}_1^\top \vec{v}_n \\ \vdots & \ddots & \vdots \\ \vec{v}_n^\top \vec{v}_1 & \cdots & \vec{v}_n^\top \vec{v}_n \end{bmatrix} \\
&= \begin{bmatrix} \langle \vec{v}_1, \vec{v}_1 \rangle & \cdots & \langle \vec{v}_n, \vec{v}_1 \rangle \\ \vdots & \ddots & \vdots \\ \langle \vec{v}_1, \vec{v}_n \rangle & \cdots & \langle \vec{v}_n, \vec{v}_n \rangle \end{bmatrix} \\
&= \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix} \\
&= I_n.
\end{aligned}$$

Here we used the real inner product, the real definition of orthonormality, and so on. Since M is $n \times n$ and has n orthonormal columns, M has full rank, and is hence invertible. Thus, right-multiplying by M^{-1}, we get that $M^\top = M^{-1}$ as desired.

That was a pretty standard proof in the real case. Now let us try the complex case, and you can see how little actually changes.

Let $M = \begin{bmatrix} \vec{v}_1 & \cdots & \vec{v}_n \end{bmatrix}$ be an $n \times n$ square matrix with orthonormal columns. Then

$$\begin{aligned} M^* M &= \begin{bmatrix} \vec{v}_1^* \\ \vdots \\ \vec{v}_n^* \end{bmatrix} \begin{bmatrix} \vec{v}_1 & \cdots & \vec{v}_n \end{bmatrix} \\ &= \begin{bmatrix} \vec{v}_1^* \vec{v}_1 & \cdots & \vec{v}_1^* \vec{v}_n \\ \vdots & \ddots & \vdots \\ \vec{v}_n^* \vec{v}_1 & \cdots & \vec{v}_n^* \vec{v}_n \end{bmatrix} \\ &= \begin{bmatrix} \langle \vec{v}_1, \vec{v}_1 \rangle & \cdots & \langle \vec{v}_n, \vec{v}_1 \rangle \\ \vdots & \ddots & \vdots \\ \langle \vec{v}_1, \vec{v}_n \rangle & \cdots & \langle \vec{v}_n, \vec{v}_n \rangle \end{bmatrix} \\ &= \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix} \\ &= I_n. \end{aligned}$$

Since M is $n \times n$ and has n orthonormal columns, M has full rank, and is hence invertible. Thus, right-multiplying by M^{-1}, we get that $M^* = M^{-1}$ as desired.

See how little changed? Our conceptual understanding stays roughly the same, so the notation does as well. Not all such proofs can be generalized like this, but a large portion of matrix computations can really be taken into the complex domain and solved in greater generality like that.

4 [Optional] More on the Orthogonality Principle

Back when we first discussed the orthogonality principle, we advertised a lot of deep theory and interesting applications, modulo some technical conditions that we said weren't too relevant for our case. We will go over a sampling of those right now.

One application of this is in signal processing. Suppose we are given a signal, i.e., a two-sided infinite sequence y. In particular, we know the value of y_n for all integer n. Further suppose we are trying to find the best linear estimate, given y, of some quantity x (which could even be a signal itself). If y was finite, this would be exactly the least squares problem. But since y is infinite, we have no hope of putting y in a matrix and doing least squares. By using the orthogonality principle, we can get an infinite system of linear equations for x, i.e., if $f(y)$ is our linear projection of y onto the space of x (i.e., $\mathbb{C}$), then we know that $\langle x - f(y), y_n \rangle = 0$ for every n. In general, infinite systems of linear equations may not be efficiently solvable, but in a wide variety of relevant cases, y and x are related in such a way that this problem becomes tractable. This is one of the core problem solving strategies, say, of EE 225A.

Another thing to note is that when we were abstracting away from $\mathbb{R}^n$, we didn't have any restriction on the dimensionality of V. In fact, V can be infinite dimensional! The orthogonality principle in infinite dimensions allows us to find projections in certain nice classes of *function spaces*, or just vector spaces of functions. And this provides a very nice, linear-algebraically sound theory of function estimation, which is a fundamental aspect of machine learning, deep learning, and statistics.

However, in infinite dimensions, there is one additional technical caveat to be obeyed. That is, V needs to be well-behaved in the following sense. If $\vec{v}_1, \vec{v}_2, \ldots$ is a sequence of vectors in $\vec{v}$ such that $\|\vec{v}_n - \vec{v}_m\|$ becomes arbitrarily small as $n, m \to \infty$, then the sequence converges to a limit in V. This is called *completeness* (with respect to $\|\cdot\|$) because it ensures that there are no "gaps" in V.[8] Happily, it is a fact of linear algebra that all finite-dimensional vector spaces with a norm are complete. However, it may be surprising that this condition is not guaranteed in infinite dimensions, but is nonetheless required for a fully general orthogonality principle.

Anyways, this is just a sampler of fun things you can use the orthogonality principle to do, and what you need to be careful of while using it. To really understand the implications and technicalities, you should take the appropriate mathematics classes (say for example Math. 110, Math. 202AB) or signal processing classes (EE 120, EE 123, EE 127, EE 225A, EE 226A).

Contributors:

- Druv Pai.

[8] A vector space equipped with a compatible inner product and norm with respect to which it is complete is called a Hilbert space. Hilbert space theory is one of the cornerstones of signal processing, optimization, etc.. A vector space with just a norm with respect to which it is complete also has a special name; it is called a Banach space.

