Inner Product Spaces

January 31, 2025

1 Inner Products

A semi-inner-product $\langle u, v\rangle$ on a *real vector space* is a real-valued function of two vectors u, v that is *symmetric*, *linear* in its arguments, and *positive*. That is, for all vectors u, v, w and scalars $\alpha \in \mathbb{R}$,

$$\langle u, v\rangle = \langle v, u\rangle, \quad \langle u + \alpha v, w\rangle = \langle u, w\rangle + \alpha\langle v, w\rangle, \quad \text{and} \quad \langle u, u\rangle \geq 0.$$

An inner product is a semi-inner-product that is *positive definite*, i.e., that satisfies $\langle u, u\rangle = 0$ if and only if $u = 0$. We tend to talk more about inner products than semi-inner products, but there are a few semi-inner products we use often that aren't positive-definite.

Here are some examples of semi-inner products.

- For real scalars, we have the product $\langle u, v\rangle = uv$.
- On finite dimensional vectors $v \in \mathbb{R}^n$, we have the dot product, $\langle u, v\rangle_2 := \sum_{i=1}^n u_i v_i = u^T v$.
- On functions $v(x)$, in terms of a random variable X with distribution P, we have the population inner product $\langle u, v\rangle_{L_2(P)} = \mathrm{E}[u(X)v(X)]$ and the covariance $\mathrm{Cov}_{\mathrm{P}}(u, v) = \mathrm{E}\left[\{u(X) - \mathrm{E}[u(X)]\}\{v(X) - \mathrm{E}[v(X)]\}\right]$.

Exercise 1 *Prove that these examples are semi-inner-products.*

Solution 1 *Let's start with the real scalars.*

- *Symmetry:* $\langle u, v\rangle = uv = vu$.
- *Linearity:* $\langle u + \alpha v, w\rangle = (u + \alpha v)w = uw + \alpha vw = \langle u, w\rangle + \alpha\langle v, w\rangle$.
- *Positivity:* $\langle v, v\rangle = v^2 \geq 0$.

The finite-vector and population two norms are just sums and expectations of these, and since these properties hold for each coordinate v_i or function evaluation $v(x)$, they hold for these sums and expectations. Same for symmetry and linearity of the covariance, and we can show positivity the same way, too: $\mathrm{Cov}_{\mathrm{P}}(v, v) = \mathrm{E}\left[(v(X) - \mathrm{E}\, v(X))^2\right]$ *is the expectation of the square of a real-valued random variable.*

For each of these, there is an associated seminorm $\rho(v) = \sqrt{\langle v, v \rangle}$. In fact, they're all included in the list of examples in the Vector Spaces Homework. To work out which it is, you can write out $\langle v, v \rangle$ for the specific semi-inner-product you're thinking about, then compare to the example seminorms' definitions.

Exercise 2 *For each of these examples of semi-inner-products, what is the corresponding seminorm?*

Solution 2 *For* $\mathrm{Cov}_{\mathrm{P}}$*,* sd_{P}*. For the others, look for matching subscripts.*

Sample Inner Products Just like with seminorms, sometimes when we're working with a sample $X_1 \ldots X_n$, we think of functions as vectors: for functions u and v, $\langle u, v \rangle_2 = \sum_{i=1}^n u(X_i)v(X_i)$. And, as with the seminorms, this is just a scaled version of the population inner product for the *empirical distribution* P_{n}, which is the distribution that puts probability $1/n$ on each of our n observations.

$$\begin{aligned} \langle u, v \rangle_{L_2(\mathrm{P}_{\mathrm{n}})} &= \sum_{x \in X_1 \ldots X_n} P(X = x)\ u(x)v(x) \\ &= \sum_{i=1}^n \frac{1}{n}\ u(X_i)v(X_i) = \frac{1}{n}\langle u, v \rangle_2. \end{aligned}$$

1.1 Cauchy-Schwarz Inequality

The Cauchy-Schwarz inequality is the first tool we reach for when bounding a semi-inner-product. For any semi-inner-product $\langle \cdot, \cdot \rangle$, $|\langle u, v \rangle| \le \rho(u)\rho(v)$ where $\rho(v) = \sqrt{\langle v, v \rangle}$; furthermore, given any u, there is always a vector v of a given 'length' $\rho(v)$ for which this bound is attained.

Exercise 3 *Think about the Cauchy-Schwarz inequality in context of the inner product* $\langle u, v \rangle = uv$ *on scalars, the dot product* $\langle u, v \rangle_2 = u^T v$*, and the covariance inner product* $\mathrm{Cov}_{\mathrm{P}}(u, v)$*. In each context, what does it say? Be as context-specific as you can; repeating the definition three times is not an instructive exercise. A sentence or two will do for each.*

Solution 3 *For scalars, it says the magnitude of a product is less than the product of their magnitudes. In fact, it is the product of their magnitudes in general.*

In the case of finite vectors, it says that the magnitude of the dot product is no larger than the product of their lengths. It is the product of their two lengths times the absolute value of the cosine of the angle between them, equal to the product of lengths when those vectors are pointing in the same or opposite directions.

For functions, it says (using the population inner product) that the average of the product is no larger than the product of their 'typical sizes'—their root-mean-squares. This is equal, as in the vector case, when one is a scalar multiple of the other. And (using the covariance inner product), that their covariance is *no larger than the product of their standard deviations, equal when the functions' deviations from their mean are scalar multiples.*

I will not ask you to prove the Cauchy-Schwarz inequality, but if you're interested, take a look at one of the proofs on Wikipedia.

1.2 Hölder's Inequality

To bound the dot product on vectors in $\mathbb{R}^n$, Hölder's inequality is the second tool we reach for. While this is a fairly general tool, we often use a simple special case that's easy to prove: the one for the dot product, $|\langle u, v\rangle_2| \le \|u\|_1 \|v\|_\infty$.[1]

Exercise 4 *Prove it! If it takes you more than one line, you're doing it wrong.*

Solution 4

$$\left|\sum_i u_i v_i\right| \le \sum_i |u_i||v_i| \le \sum_i |u_i| (\max_i |v_i|) = \|v\|_\infty \|u\|_1.$$

There are also versions for some inner products on functions. We'll want one for sample inner products analogous to the one we have for the dot product on vectors above: $\langle u, v\rangle_{L_2(\mathrm{P_n})} \le \|u\|_{L_1(\mathrm{P_n})} \|v\|_{L_\infty(\mathrm{P_n})}$.

Exercise 5 *Prove it! If you want, you can write a new proof, but it may be more instructive to show that it's implied by the case for vectors in* $\mathbb{R}^n$.

Solution 5 *Given functions $u(x)$ and $v(x)$, consider the n-dimensional vectors with elements $u_i = u(X_i)$ and $v_i = v(X_i)$. The sample inner product is $1/n$ times the dot product of these vectors, and the one and infinity norms of these two functions are $1/n$ and 1 times the one and infinity norms of these vectors. Thus, the left and right sides of the bound for the sample norms are $1/n$ times the left and right sides of the bound for the n-dimensional vectors.*

1.3 Triangle Inequality

When we showed that a few of our examples of seminorms are in fact seminorms in last week's homework, we didn't deal with any examples of seminorms associated with semi-inner-products. Let's do that now. Or the hard part, anyway.

Exercise 6 *Prove that, for any semi-inner product $\langle u, v\rangle$, the seminorm $\rho(v) = \sqrt{\langle v, v\rangle}$ satisfies the triangle inequality.*

Hint. You want to show that $\rho(u+v)^2 \le \{\rho(u) + \rho(v)\}^2$. You know that $\rho(u+v)^2 = \langle u+v, u+v\rangle$. Expand this as the sum of four terms using *linearity*, then see what you can work out using the Cauchy-Schwarz inequality.

Hint. If you are not entirely comfortable with notation $\langle u, v\rangle$ for inner products, use the more familiar notation $u^T v$.

Solution 6

$$\begin{aligned}\rho(u+v)^2 &= \langle u+v, u+v\rangle \\ &= \rho(u)^2 + \rho(v)^2 + (\langle u, v\rangle + \langle v, u\rangle) \\ &\le \rho(u)^2 + \rho(v)^2 + 2\rho(u)\rho(v) \\ &= \{\rho(u) + \rho(v)\}^2.\end{aligned}$$

2 Complex Vector Spaces

This semester, we'll mostly be working with real numbers, real vectors, and real-valued functions. But just like in high school algebra, it's occasionally useful to work with complex ones. Notationally, $x \in \mathbb{C}$ is a complex number, $v \in \mathbb{C}^n$ is a complex vector, and $v : \mathcal{X} \to \mathbb{C}$ is a complex-valued function. If you've forgotten how to work with complex numbers, here's what you'll need to know.

1. A complex number is $x + iy$ where x and y are real numbers and $i = \sqrt{-1}$.
2. These add and multiply as you'd expect: if $z_1 = x_1 + iy_1$ and $z_2 = x_2 + iy_2$, then $z_1 + z_2 = (x_1 + x_2) + i(y_i + y_2)$ and $z_1 z_2 = (x_1 + iy_1)(x_2 + iy_2) = x_1x_2 + i^2 y_1y_2 + i(y_1x_2 + x_1y_2) = x_1x_2 - y_1y_2 + i(y_1x_2 + x_1y_2)$. Think about this: what happened to the i^2 in the last equality?
3. We tend to think of a complex number $z = x + iy$ as a vector in the plane with magnitude $|x+iy| = \sqrt{x^2+y^2}$ and angle $\tan^{-1}(y/x)$. We call $\bar{z} = x - iy$ the complex conjugate of $z = x + iy$, which allows us a convenient expression for the magnitude of z: $|z| = \sqrt{z\bar{z}}$. Note that $\overline{a+b} = \bar{a} + \bar{b}$, and $\overline{ab} = \bar{a}\bar{b}$, and $\bar{\bar{a}} = a$ for $a, b \in \mathbb{C}$. We will also refer to the complex conjugate of a vector or a function, which is interpreted *elementwise*. For a vector $v \in \mathbb{C}^n$ with elements v_i, $\bar{v}$ is the vector with elements $\bar{v}_i$; for a complex-valued function v, $\bar{v}$ is the function with $\bar{v}(x) = \overline{v(x)}$ for all x.

When we're thinking about spaces of vectors $v \in \mathbb{C}^n$ or functions $v : \mathcal{X} \to \mathbb{C}$, we'll want to think of them as elements of a *complex vector space*, i.e., as elements in a vector space where the *scalars* are complex numbers. For example, …

- for a vector $v \in \mathbb{C}^n$ and a scalar $\alpha \in \mathbb{C}$, $u = \alpha v \in \mathbb{C}^n$ will be the vector with elements $u_i = \alpha v_i$.
- for a function $v : \mathcal{X} \to \mathbb{C}$ and a scalar $\alpha \in \mathbb{C}$, $u = \alpha v : \mathcal{X} \to \mathbb{C}$ is the function with $u(x) = \alpha v(x)$ for all $x \in \mathcal{X}$.

To talk about inner products on complex vector spaces, we need to make one small change to our definition of a semi-inner product from Section 1. We have to talk about *conjugate-symmetry* instead of *symmetry*. A semi-inner-product

$\langle u, v\rangle$ on a complex vector space is a complex-valued function of two vectors u, v that is *conjugate-symmetric*, *linear* in its arguments, and *positive*. That is, for all vectors u, v, w and scalars $\alpha \in \mathbb{C}$,

$$\langle u, v\rangle = \overline{\langle v, u\rangle}, \quad \langle u + \alpha v, w\rangle = \langle u, w\rangle + \alpha\langle v, w\rangle, \quad \text{and} \quad \langle u, u\rangle \geq 0.$$

Why conjugate-symmetry? Think about the simplest complex vector space: the complex numbers $\mathbb{C}$. By using the inner product $\langle u, v\rangle = u\bar{v}$ for $u, v \in \mathbb{C}$, we get the magnitude $|v| = \sqrt{v\bar{v}}$ as the norm $\|v\| = \sqrt{\langle v, v\rangle}$. More generally, ...

- For vectors $u, v \in \mathbb{C}^n$, we typically use the inner product $\langle u, v\rangle = u^T\bar{v}$, and we get the norm $\|v\| = \sqrt{\sum_i |v_i|^2}$.
- For functions $u, v : [0, 1] \to \mathbb{C}$, we typically use the inner product $\langle u, v\rangle = \int_0^1 u(x)\overline{v(x)}dx$, and we get the norm $\|v\| = \sqrt{\int_0^1 |v(x)|^2 dx}$.

If you look back at what's written above and in last week's homework, everything still works. Whenever $\langle \cdot, \cdot\rangle$ is a semi-inner product on a complex vector space and $\|v\| = \sqrt{\langle v, v\rangle}$ is the corresponding seminorm, the Cauchy-Schwarz inequality $|\langle u, v\rangle| \leq \rho(u)\rho(v)$ and the triangle inequality $\rho(u + v) \leq \rho(u) + \rho(v)$ hold. And when $\langle \cdot, \cdot\rangle$ is the inner product $\langle u, v\rangle = u^T\bar{v}$ on $\mathbb{C}^n$, Hölder's inequality $|\langle u, v\rangle| \leq \|u\|_\infty \|v\|_1$ holds where $\|u\|_\infty = \max_i |u_i|$ is the maximum of the magnitudes of the elements of u and $\|v\|_1 = \sum_i |v_i|$ is the sum of the magnitudes of the elements of v.

If you want some practice working with complex numbers, try the exercises in Appendix A, where you'll prove a few of these.

3 Self-adjoint Operators

In this problem, we'll generalize of the idea of a symmetric matrix.

You can think of an $n \times n$ matrix $A \in \mathbb{R}^{n\times n}$ as a linear operator on the vector space $\mathbb{R}^n$, i.e. a function from $\mathbb{R}^n$ to $\mathbb{R}^n$ that's linear in the sense that that $A(\alpha u + \beta v) = \alpha Au + \beta Av$ for any $\alpha, \beta \in \mathbb{R}$ and $u, v \in \mathbb{R}^n$. And we can talk about linear operators on other vectors spaces. For example, $\frac{d}{dx}$ is a linear operator on the space of infinitely-differentiable functions, as $\frac{d}{dx}\{\alpha u(x) + \beta v(x)\} = \alpha \frac{d}{dx}u(x) + \beta \frac{d}{dx}v(x)$.

When we're working with an inner product $\langle u, v\rangle$ on our vector space $\mathcal{V}$, we can define the *adjoint* $A^\star$ of a linear operator A to be another linear operator satisfying $\langle A^\star u, v\rangle = \langle u, Av\rangle$ for all vectors u and v. Here are some examples.

3.1 Operators on finite dimensional spaces

When we're working with the dot product $\langle u, v\rangle_2 = u^T v$ on $\mathbb{R}^n$, the adjoint of a matrix $A \in \mathbb{R}^{n\times n}$ is its *transpose* A^T.

$$\langle A^T u, v\rangle_2 = (A^T u)^T v = u^T Av = \langle u, Av\rangle_2.$$

When we're working with the dot product $\langle u, v\rangle_2 = u^T\bar{v}$ on $\mathbb{C}^n$, the adjoint of a matrix $A \in \mathbb{C}^{n\times n}$ is its *conjugate transpose* $\bar{A}^T$. That is, it's the matrix whose elements are the complex conjugates of the elements in A^T.

$$\langle \bar{A}^T u, v\rangle_2 = (\bar{A}^T u)^T\bar{v} = u^T\bar{A}\bar{v} = u^T\overline{Av} = \langle u, Av\rangle.$$

Why we use complex spaces. Even when we really intend to work with real-valued vectors, it's useful to think about matrices as operators on $\mathbb{C}^n$ and think of the dot product $\langle u, v\rangle_2$ as $u^T\bar{v}$. We have to deal with complex numbers in any case, as matrices $A \in \mathbb{R}^{n\times n}$ can have complex eigenvalues and eigenvectors. And the inner product u^Tv we use on $\mathbb{R}^n$ isn't an inner product on complex vectors at all, as the norm $\|v\|^2 = \langle v, v\rangle$ associated with an inner product must be positive and u^Tu will be negative for imaginary vectors.

3.2 Operators on spaces of functions

When we're working with the inner product $\langle u, v\rangle = \int_{-\infty}^{\infty} u(x)v(x)dx$ on the vector space of infinitely-differentiable functions $v : \mathbb{R} \to \mathbb{R}$ with $v(x) \to 0$ as $x \to \pm\infty$, the adjoint of the linear operator $\frac{d}{dx}$ is $-\frac{d}{dx}$. To see this, we integrate by parts.

$$\begin{aligned}
\left\langle u, \frac{d}{dx}v\right\rangle &= \int_{-\infty}^{\infty} u(x)v'(x)dx && \text{by definition} \\
&= u(x)v(x)|_{-\infty}^{\infty} \quad - \quad \int_{-\infty}^{\infty} u'(x)v(x)dx && \text{because } \ (uv)' = u'v + uv' \\
&= 0 \quad - \quad \int_{-\infty}^{\infty} u'(x)v(x)dx && \text{because } \ u(x)v(x) \underset{x\to\pm\infty}{\to} 0 \\
&= \left\langle -\frac{d}{dx}u, v\right\rangle.
\end{aligned}$$

Note that it's important that our vector space includes only functions that go to zero as $x \to \pm\infty$; otherwise our 'boundary term' $u(x)v(x)|_{-\infty}^{\infty}$ would be nonzero and we could not say that $-\frac{d}{dx}$ was the adjoint of $\frac{d}{dx}$.

Specifying the vector space and inner product we're using is more important when talking about operators on spaces of functions than operators on finite-dimensional vectors. We can essentially get away with assuming we're talking about $\mathbb{C}^n$ and $\langle u, v\rangle = u^T\bar{v}$ in the latter case because that's what everyone always does; we don't have unspoken defaults like this for operators on functions.

The Complex Case. The adjoint is still $-\frac{d}{dx}$ if we're thinking about $\frac{d}{dx}$ as a linear operator on the space of *complex-valued* infinitely-differentiable functions with $v(x) \to 0$ as $x \to \pm\infty$ with the inner product $\langle u, v\rangle = \int_{-\infty}^{\infty} u(x)\overline{v(x)}dx$. It's useful to think this way for the same reason it's useful to think about $\mathbb{C}^n$ instead of $\mathbb{R}^n$.

You may not be familiar with derivatives and integrals involving complex-valued functions. That's no big deal.[2] For a complex-valued function $u : \mathbb{R} \to \mathbb{C}$,

$u(x) = u_r(x) + iu_i(x)$ where u_r and u_i are real-valued functions, differentiation and integration are done component-wise. That is, $\frac{d}{dx}u(x) = \frac{d}{dx}u_r(x) + i\frac{d}{dx}u_i(x)$ and $\int u(x)dx = \int u_r(x)dx + i\int u_i(x)dx$. We can show that $-\frac{d}{dx}$ is the adjoint of $\frac{d}{dx}$ by using integration by parts as above on the real and imaginary components separately.

3.3 Self-adjointness

A self-adjoint operator on a vector space $\mathcal{V}$ with an inner product $\langle u, v\rangle$ is, as you would expect, an operator that is its own adjoint. That is, we say an operator A is self-adjoint if $\langle Au, v\rangle = \langle u, Av\rangle$. Symmetric matrices, i.e. matrices A with $A^T = A$, are self-adjoint on $\mathbb{R}^n$ with the usual inner product $\langle u, v\rangle = u^T v$. *Conjugate-symmetric* matrices, i.e. matrices A with $A^T = \bar{A}$, are self-adjoint on $\mathbb{C}^n$ with the usual inner product $\langle u, v\rangle = u^T \bar{v}$.

Now let's talk about self-adjoint operators on spaces of functions. A classic example is the differential operator $-\frac{d^2}{dx^2}$ on the space of 2-periodic complex-valued twice-differentiable functions, $\{v : [-1, 1] \to \mathbb{C} \ : \ v(-1) = v(1)\}$, with inner product $\langle u, v\rangle = (1/2)\int_{-1}^{1} u(x)\overline{v(x)}dx$.[3]

Exercise 7 *Prove that the operator $-\frac{d^2}{dx^2}$ on this space is self-adjoint. That is, prove that $\langle -\frac{d^2}{dx^2}u,\ v\rangle = \langle u, -\frac{d^2}{dx^2}v\rangle$ for periodic functions u and v.*

Hint. Integrate by parts twice. Why is it important that u and v be periodic?

Solution 7 *We can use the integration-by-parts argument from Section 3.2 to show that $-\frac{d}{dx}$ is the adjoint of $\frac{d}{dx}$ on this vector space. To do this, we'd replace integration from $-\infty$ to ∞ with integration from -1 to $+1$; the condition $v(-1) = v(1)$ for $v \in \mathcal{V}$ ensures that the boundary term $u(x)v(x)\ |_{-1}^{1}$ is zero. Because $-\frac{d^2}{dx^2}v = -\frac{d}{dx}\frac{d}{dx}v$, it follows that*

$$\left\langle u, -\frac{d^2}{dx^2}v \right\rangle = \left\langle u, -\frac{d}{dx}\frac{d}{dx}v \right\rangle = \left\langle \frac{d}{dx}u, \frac{d}{dx}v \right\rangle = \left\langle -\frac{d}{dx}\frac{d}{dx}u, v \right\rangle = \left\langle -\frac{d^2}{dx^2}u, v \right\rangle.$$

3.4 Diagonalizing self-adjoint operators

Just like a matrix, a linear operator L has eigenvalues and eigenvectors: scalars λ and vectors v for which $Lv = \lambda v$.[4] In our example, they are defined by the differential equation $-\frac{d^2}{dx^2}v = \lambda v$. And like a symmetric matrix, a self-adjoint linear operator's eigenvalues are real and the eigenvectors corresponding to distinct eigenvalues are orthogonal.

Exercise 8 *Prove that if L is a self-adjoint operator on a complex vector space with an inner product $\langle u, v\rangle$, then its eigenvalues are real and the eigenvectors corresponding to distinct eigenvalues are orthogonal. That is, prove that*

1. *If $Lv = \lambda v$ for some vector v and scalar $\lambda \in \mathbb{C}$, then $\lambda \in \mathbb{R}$.*

2. *If $Lv = \lambda v$ and $Lu = \eta u$ for vectors v and u and $\lambda \neq \eta \in \mathbb{C}$, then $\langle u, v \rangle = 0$.*

Having done this, explain why this implies that, for integers j and k with $j \neq k$,

$$\int_{-1}^{1} \sin(\pi kx)\sin(\pi jx) = \int_{-1}^{1} \cos(\pi kx)\cos(\pi jx) = \int_{-1}^{1} \cos(\pi kx)\sin(\pi jx)dx = 0.$$

Hint. *Recall from Section 2 that if $\langle u, v \rangle$ is an inner product on a complex vector space, then for any vectors u and v,*

$$\langle u, v \rangle = \overline{\langle v, u \rangle}, \quad \langle u + \alpha v, w \rangle = \langle u, w \rangle + \alpha \langle v, w \rangle, \quad \textit{and} \quad \langle u, u \rangle \geq 0.$$

Hint. *What are $\frac{d^2}{dx^2}\sin(\pi kx)$ and $\frac{d^2}{dx^2}\cos(\pi kx)$?*

Solution 8 *Let u and v be eigenvectors of the self-adjoint operator A with corresponding eigenvalues λ and η. First we'll show that these eigenvalues are real. Because A is self-adjoint,*

$$\lambda\langle u, u \rangle = \langle \lambda u, u \rangle = \langle Au, u \rangle = \langle u, Au \rangle = \langle u, \lambda u \rangle = \overline{\langle \lambda u, u \rangle} = \bar{\lambda}\langle u, u \rangle.$$

Furthermore, using this and self-adjointness again, we can show orthogonality of eigenvectors.

$$\lambda\langle u, v \rangle = \langle Au, v \rangle = \langle u, Av \rangle = \eta\langle u, v \rangle.$$

If $\lambda \neq \eta$, this can be true only if $\langle u, v \rangle = 0$.

$\sin(\pi kx)$ *and* $\sin(\pi jx)$ *are eigenvectors of the self-adjoint operator $-\frac{d^2}{dx^2}$ on the space of periodic functions on $[-1, 1]$ with inner product $\langle u, v \rangle = (1/2)\int_{-1}^{1} u(x)v(x)dx$ with eigenvalues $k\pi$ and $j\pi$ respectively. It follows that their inner product $(1/2)\int_{-1}^{1} \sin(\pi kx)\sin(\pi jx)$ must be zero unless $j = k$. Same goes for $\cos(\pi kx)$ and $\cos(\pi jx)$ as well as $\cos(\pi kx)$ and $\sin(\pi jx)$.*

Later on, we'll use the results we've proven to talk about the models defined using the *Sobolev seminorm* $\rho(v) = \sqrt{\int_0^1 |v'(x)|^2 dx}$ and its generalizations.

A Inner Products on Complex Vector Spaces: Exercises

These exercises are optional.

Exercise 9 ***(Optional)***. *Prove that, for any semi-inner product $\langle u, v\rangle$ on a complex vector space, the seminorm $\rho(v) = \sqrt{\langle v, v\rangle}$ satisfies the triangle inequality.*

You may assume that the Cauchy-Schwarz inequality $\langle u, v\rangle \le \rho(u)\rho(v)$ holds.

Tip. Do you need to change your solution to Exercise 6? If so, how?

Solution 9 *Take a look at the solution to Exercise 6. Everything is just arithmetic except for one step: the one where we use the Cauchy-Schwarz inequality. This one.*

$$\rho(u)^2 + \rho(v)^2 + (\langle u, v\rangle + \langle v, u\rangle) \le \rho(u)^2 + \rho(v)^2 + 2\rho(u)\rho(v).$$

Or equivalently, subtracting $\rho(u)^2 + \rho(v)^2$ from both sides,

$$\langle u, v\rangle + \langle v, u\rangle \le 2\rho(u)\rho(v).$$

When we were talking about real vector spaces, this was an obvious consequence of the Cauchy-Schwarz inequality $|\langle u, v\rangle| \le \rho(u)\rho(v)$ (and $|\langle v, u\rangle| \le \rho(v)\rho(u) = \rho(u)\rho(v)$). Why? If x is a real number, then $x \le |x|$; taking $x = \langle u, v\rangle$ we have $\langle u, v\rangle \le |\langle u, v\rangle| \le \rho(u)\rho(v)$. But when x isn't *a real number, it isn't even clear what $x \le |x|$ means. Thankfully, we know that $\langle u, v\rangle + \langle v, u\rangle$ is a real number. This is implied by conjugate-symmetry: writing $\langle u, v\rangle = x + iy$, we have $\langle u, v\rangle + \langle v, u\rangle = (x + iy) + \overline{x + iy} = (x + iy) + (x - iy) = 2x$.*

Now let's assume, for a moment, that the magnitude satisfies a triangle inequality, i.e., that $|z_1 + z_2| \le |z_1| + |z_2|$ for any $z_1, z_2 \in \mathbb{C}$. Applying this to $z_1 = \langle u, v\rangle$ and $z_2 = \langle v, u\rangle$, then using the Cauchy-Schwarz inequality, we get the bound we want. We know that the real number $z_1 + z_2$ is less than or equal to its magnitude $|z_1 + z_2| = |\langle u, v\rangle + \langle v, u\rangle|$, and consequently that …

$$\langle u, v\rangle + \langle v, u\rangle \le |\langle u, v\rangle + \langle v, u\rangle| \le |\langle u, v\rangle| + |\langle v, u\rangle| \le \rho(u)\rho(v) + \rho(v)\rho(u) = 2\rho(u)\rho(v).$$

So all we need to do is show that the magnitude satisfies a triangle inequality. And we've already done that in Exercise 6—we've just done it in disguise. If we think of two complex numbers $z_1 = x_1 + iy_1$ and $z_2 = x_2 + iy_2$ as vectors (x_1, y_1) and (x_2, y_2) in the plane $\mathbb{R}^2$, then the triangle inequality for the magnitude is the triangle inequality for the two-norm of these vectors.

$$|z_1 + z_2| = \|(x_1 + x_2, y_1 + y_2)\|_2 \le \|(x_1, y_1)\|_2 + \|(x_2, y_2)\|_2 = |z_1| + |z_2|.$$

Exercise 10 ***(Optional)***. *Prove Hölder's inequality for $\mathbb{C}^n$.*

Solution 10 *Take a look at the solution to Exercise 4.*

$$\left|\sum_i u_i v_i\right| \leq \sum_i |u_i||v_i| \leq \sum_i |u_i|(\max_i |v_i|) = \|v\|_\infty \|u\|_1.$$

This is still true. But when $u, v \in \mathbb{R}^n$, we could justify the first inequality term-by-term, as a consequence of the fact that $u_i v_i$ is either $|u_i v_i|$ or $-|u_i v_i|$ and therefore $u_i v_i \leq |u_i v_i|$. In the complex case, we can't do this, but we can get there using the triangle inequality for the magnitude repeatedly.

$$\left|\sum_{i=1}^n u_i v_i\right| \leq |u_1 v_1| + \left|\sum_{i=2}^n u_i v_i\right| \leq |u_1 v_1| + |u_2 v_2| + \left|\sum_{i=3}^n u_i v_i\right| \leq \ldots \leq \sum_{i=1}^n |u_i||v_j|.$$

Once we've done this, the rest of the solution is the same.

Notes

[1]If you'd like to get a sense of Hölder's inequality in full generality, take a look at this wikipedia article.

[2]What makes calculus involving complex numbers different is not really dealing with *complex-valued functions*, but rather with *functions of a complex variable.*

[3]If you prefer to think of these as periodic functions from $\mathbb{R} \to \mathbb{C}$ you can, since all a 2-periodic function does outside the interval $[-1, 1]$ is repeat what it does on $[-1, 1]$. Some people like to think of these as functions *on the circle*, too.

[4]The eigenvectors of operators on vector spaces of functions are sometimes called eigenfunctions.