Assignment 2: Structured Vector Spaces and Riesz Representation
Due Tuesday September 13, 2022

John McCuan

October 6, 2022

Problem 1 *Recall the* **Hermitian inner product**

$$\langle \mathbf{z}, \mathbf{w} \rangle = \sum_{j=1}^{n} z_j \overline{w_j}$$

on $\mathbb{C}^n$ *where* $\mathbf{z} = (z_1, z_2, \ldots, z_n)$ *and* $\mathbf{w} = (w_1, w_2, \ldots, w_n)$. *Given a complex vector* $\mathbf{n} \in \mathbb{C}^2$, *describe the orthogonal space*

$$Z = \{\mathbf{z} \in \mathbb{C}^2 : \langle \mathbf{z}, \mathbf{n} \rangle = 0\}.$$

For example, the corresponding answer for a vector $\mathbf{n} \in \mathbb{R}^2$ *with respect to the dot product would be "a two-dimensional plane through the origin orthogonal to the vector* $\mathbf{n}$*," but you should try to avoid resorting to an explanation in terms of (only) complex dimensions.*

Solution: The point of this problem/exercise is to put a certain kind of "picture" in your mind. As pointed out, you should have a picture associated with

$$Z_1 = \{\mathbf{x} \in \mathbb{R}^2 : \langle \mathbf{x}, \mathbf{n} \rangle = 0\}.$$

Let's think about this case as motivation. What does it really mean that you have a picture associated with this set? Initially, one can say the points $\mathbf{x} = (x_1, x_2) \in Z_1$ satisfy

$$n_1 x_1 + n_2 x_2 = 0 \tag{1}$$

for some $n_1, n_2 \in \mathbb{R}$. You're familiar with the fact that that if $n_1 = n_2 = 0$, then $Z_1 = \mathbb{R}^2$ and otherwise, the relation (1) determines a line. We can go further to say that if $n_2 = 0 \neq n_1$, then the line is the x_2-axis, and otherwise assuming $n_2 \neq 0$, the relation can be rewritten as

$$x_2 = -\frac{n_1}{n_2} x_1.$$

Thus, we have a picture in which the **height** is a simple scaling of the value x_1 which may be pictured as the point $(x_1, 0)$ on the x_1-axis. That is, the line is the graph of a function which is a scaling, and which "we" understand pretty well via proportion (or similar triangles) are whatever elementary geometric intuition "we" have at "our" disposal.

As further motivation, let's increase the real dimension by one, and consider the question for

$$Z_2 = \{\mathbf{x} \in \mathbb{R}^3 : \langle \mathbf{x}, \mathbf{n} \rangle = 0\}.$$

Again, if $\mathbf{n} = \mathbf{0}$, then $Z_2 = \mathbb{R}^3$ while if $\mathbf{n} \neq \mathbf{0}$ the relation

$$n_1 x_1 + n_2 x_2 + n_3 x_3 = 0$$

determines a plane (passing through $\mathbf{0} \in \mathbb{R}^3$). More precisely, if $n_3 = 0$ but $\mathbf{n} \neq \mathbf{0}$, this plane is the vertical plane over the injection of the line Z_1 given by

$$Z_1 \ni (x_1, x_2) \mapsto (x_1, x_2, 0) \in \mathbb{R}^3.$$

If $n_3 \neq 0$, then again the plane is a graph

$$Z_2 = \left\{ \left(x_1, x_2, -\frac{n_1}{n_3} x_1 - \frac{n_2}{n_3} x_2 \right) : (x_1, x_2) \in \mathbb{R}^2 \right\} \tag{2}$$

over the x_1, x_2-plane. Notice that we lose a little something here in our picture. The height is not given by a simple scaling; the height is obtained by taking some kind of (relatively complicated) linear combination of the values of x_1 and x_2. In particular, there are many points on the plane at the same height as any given point (p_1, p_2, h), and while we can write down a relation

$$-\frac{n_1}{n_3}x_1 - \frac{n_2}{n_3}x_2 = h$$

determining the other points, they are a little bit hard to "see." In fact, I "know" (2) determines a plane, but it is a little bit difficult to "see" it—in terms of for example similar triangles. Nevertheless, I have this picture: The coordinates of the points $(x_1, x_2, 0)$ in the x_1, x_2-plane are combined in some algebraic/linear way to determine a height, specifically $h = -n_1x_1/n_3 - n_2x_2/n_3$.

The interesting thing about this problem, I think, is that in a certain sense we can get a better picture for Z than we can for Z_2. Let me try to explain. Obviously, we have the case $\mathbf{n} = \mathbf{0} \in \mathbb{C}^2$ in which case $Z = \mathbb{C}^2$ which everybody can see and nobody can see. For the other cases, there is a natural identification of $\mathbb{C}^2$ with $\mathbb{R}^4$ given by

$$\mathbb{C}^2 \ni (x + iy, \xi + i\eta) \longleftrightarrow (x, y, \xi, \eta) \in \mathbb{R}^4$$

which will be somewhat useful both for visualization and for splitting up $\mathbb{C}^2$. You'll note that is what we did with $\mathbb{R}^2$ and $\mathbb{R}^3$ above when we considered graphs. We separated off one "axis" determining a height and then tried to understand the set in question as the graph of a function. Following that approach, we consider next the case $\mathbf{n} = (n_1 + n_2 i, 0 + 0i)$. In this case, we get a kind of canonical picture because it is required for $(z_1, z_2) \in Z$ that

$$z_1(n_1 - n_2 i) = 0 \qquad \text{or} \qquad z_1 = 0 \in \mathbb{C}$$

as long as $n_1 + n_2 \neq 0$. That is,

$$Z = \{0\} \times \mathbb{C} \subset \mathbb{C}^2.$$

This set is obviously identified with the two-dimensional plane

$$\{(0, 0, x_3, x_4) \in \mathbb{R}^4 : (x_3, x_4) \in \mathbb{R}^2\}.$$

Again we have a geometric object (a two-dimensional plane) that everyone can see and no one can see.

We then come to the final (and perhaps most interesting) case:

$$\mathbf{n} = (n_1 + n_2 i, \xi_1 + \xi_2 i) \in \mathbb{C}^2 \qquad \text{with} \qquad \xi_1 + \xi_2 i \neq 0 \in \mathbb{C}.$$

In this case we have the relation

$$(x_1 + x_2 i)(n_1 - n_2 i) + (x_3 + x_4 i)(\xi_1 - \xi_2 i) = 0$$

or

$$x_3 + x_4 i = \frac{n_1 - n_2 i}{\xi_1 - \xi_2 i}(x_1 + x_2 i) = \frac{(n_1 - n_2 i)(\xi_1 + \xi_2 i)}{\xi_1^2 + \xi_2^2}(x_1 + x_2 i).$$

This relation determines a "height" (in the second factor of $\mathbb{C}$) above each point $x_1 + ix_2 \in \mathbb{C}$, and that height admits a very nice geometric interpretation:

$$Z = \{(z, \alpha z) \in \mathbb{C}^2 : z \in \mathbb{C}^2\} \tag{3}$$

where

$$\alpha = \frac{(n_1 - n_2 i)(\xi_1 + \xi_2 i)}{\xi_1^2 + \xi_2^2}$$

is some fixed complex number in $\mathbb{C}$. In particular, scaling $z \in \mathbb{C}$ by a complex number $\alpha = |\alpha| e^{i\theta}$ corresponds to (simple) **scaling** of z by $|\alpha|$ and counterclockwise **rotation** by the angle θ. Taking a column representation in the subspces of $\mathbb{R}^4$ corresponding to the component decomposition in $\mathbb{C} \times \mathbb{C} = \mathbb{C}^2$ mentioned above, the set in $\mathbb{R}^4$ corresponding to the representation (3) is

$$\{(\mathbf{x}, L\mathbf{x}) \in \mathbb{R}^4 : \mathbf{x} \in \mathbb{R}^2\} = \left\{ \left(\left(\begin{array}{c} x_1 \\ x_2 \end{array} \right), L \left(\begin{array}{c} x_1 \\ x_2 \end{array} \right) \right) : \left(\begin{array}{c} x_1 \\ x_2 \end{array} \right) \in \mathbb{R}^2 \right\}$$

and $L : \mathbb{R}^2 \to \mathbb{R}^2$ is a composition of a dilation by $|\alpha|$ and rotation by θ, that is, expressed in terms of matrix multiplication

$$L \left(\begin{array}{c} x_1 \\ x_2 \end{array} \right) = |\alpha| \left(\begin{array}{cc} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{array} \right) \left(\begin{array}{c} x_1 \\ x_2 \end{array} \right).$$

This is clearly a two-dimensional plane. Furthermore the "height," i.e., the component in the second factor of $\mathbb{C}$ in $\mathbb{C}^2$ or the latter two components in $\mathbb{R}^4$ is actually easier to "see" than the real number $h = \alpha_1 x_1 + \alpha_2 x_2$ considered in the case of the two-plane Z_2 above. Incidentally, it is possible that $\alpha = 0 \in \mathbb{C}$ in which case we again obtain a canonical example

$$Z = \mathbb{C} \times \{0\} \longleftrightarrow \{(x_1, x_2, 0, 0) \in \mathbb{R}^4 : (x_1, x_2) \in \mathbb{R}^2\}.$$

That's about the best picture of Z I've got. And it's what I had in mind with the exercise.

Saying and seeing a bit more

Returning to the case of two real dimensions and the set

$$Z_1 = \{\mathbf{x} \in \mathbb{R}^2 : \mathbf{x} \cdot \mathbf{n} = 0\}.$$

Another way to "see" this line is by realizing it as a rotation of one of the canonical examples corresponding to either $n_1 = 0$ (the x_1-axis) or $n_2 = 0$ (the x_2-axis). More precisely, recall that given the condition $n_1 x_1 + n_2 x_2 = 0$ with $n_2 \neq 0$, we have $x_2 = -(n_1/n_2)x_1$ and

$$Z_1 = \{(1, -n_1/n_2)x : x \in \mathbb{R}\}.$$

This gives a parameterization of the line Z_1, that is, this expression for Z_1 suggests implicitly consideration of the mapping $\ell : \mathbb{R} \to \mathbb{R}^2$ by $\ell(x) = (1, -n_1/n_2)x$. The expression for the set Z_1, and the corresponding mapping, can be modified by the inclusion of a nonzero factor and written as

$$Z_1 = \{(n_2, -n_1)x : x \in \mathbb{R}\} \qquad \text{with} \qquad \ell(x) = (n_2, -n_1)x,$$

or

$$Z_1 = \{(b, -a)x : x \in \mathbb{R}\} \qquad \text{with} \qquad \ell(x) = (b, -a)x,$$

and

$$a = \frac{n_1}{\sqrt{n_1^2 + n_2^2}} \qquad \text{and} \qquad b = \frac{n_2}{\sqrt{n_1^2 + n_2^2}}.$$

The last form suggests the rotation suggested at the beginning. First of all, (a, b) is a unit vector and $\{(b, -a), (a, b)\}$ is a positively oriented orthonormal basis. Switching to column notation for vectors to facilitate matrix multiplication, we have

$$\ell(x) = \begin{pmatrix} b \\ -a \end{pmatrix} x$$

which suggests taking the canonical example corresponding to $n_1 = 0$ (the x-axis). That is, $L : \mathbb{R}^2 \to \mathbb{R}^2$ by

$$L \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} b & * \\ -a & * \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

when restricted to $\{(x, 0) : x \in \mathbb{R}\}$ has image Z_1. Finally, we can perhaps "see" Z_1 most clearly if the linear function L is a rotation:

$$L \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} b & a \\ -a & b \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Thus, we see the line Z_1 is the image of the x-axis under rotation of $\mathbb{R}^2$:

$$Z_1 = \left\{ L \begin{pmatrix} x \\ 0 \end{pmatrix} : x \in \mathbb{R} \right\}.$$

Let us attempt to realize the plane Z_2 as a rotation in $\mathbb{R}^3$ of the x, y-plane. That is, we will look for a 3×3 matrix A with columns an orthonormal basis of $\mathbb{R}^3$ such that

$$Z_2 = \left\{ L \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} : (x, y) \in \mathbb{R}^2 \right\}$$

where $L : \mathbb{R}^3 \to \mathbb{R}^3$ by $L\mathbf{x} = A\mathbf{x}$. Again, by homogeneous scaling we can modify the expression in (2) to obtain

$$Z_2 = \{(n_3x_1, n_3x_2, -n_1x_1 - n_2x_2) : (x_1, x_2) \in \mathbb{R}^2\}$$

and

$$Z_2 = \{(cx_1, cx_2, -ax_1 - bx_2) : (x_1, x_2) \in \mathbb{R}^2\}$$

where

$$a = \frac{n_1}{\sqrt{n_1^2 + n_2^2 + n_3^2}}, \quad b = \frac{n_2}{\sqrt{n_1^2 + n_2^2 + n_3^2}}, \quad \text{and} \quad c = \frac{n_3}{\sqrt{n_1^2 + n_2^2 + n_3^2}}.$$

Clearly, we should take $L(\mathbf{e}_3) = \mathbf{v}_3 = (a, b, c)$, but there is no unique choice for the images of $L(\mathbf{e}_1)$ and $L(\mathbf{e}_2)$ in Z_2. One assumes the "geodesic" rotation should be algebraically simplest, but it is not simple.

According to the assumption $(n_1, n_2) \neq \mathbf{0} \in \mathbb{R}^2$ and $n_3 \neq 0$, we can consider the projection $(n_1, n_2, 0)$ or $(a, b, 0)$ and the unit vector

$$\mathbf{u}_2 = \frac{(n_1, n_2, 0)}{\sqrt{n_1^2 + n_2^2}} = \frac{(a, b, 0)}{\sqrt{a^2 + b^2}}$$

along this direction. Taking the cross product $\mathbf{u}_2 \times \mathbf{e}_3$ of this vector with $\mathbf{e}_3$ we obtain the vector

$$\mathbf{u}_1 = \frac{(n_2, -n_1, 0)}{\sqrt{n_1^2 + n_2^2}} = \frac{(b, -a, 0)}{\sqrt{a^2 + b^2}}$$

which the geodesic rotation should fix. In addition $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is a positive orthonormal basis. If $n_3 > 0$, then $\mathbf{u}_2 \times \mathbf{v}_3$ is in the same direction as $\mathbf{u}_1$, and

$$\mathbf{v}_2 = \mathbf{v}_3 \times \mathbf{u}_1 = \frac{(ac, bc, -(a^2 + b^2))}{\sqrt{a^2 + b^2}}$$

gives a vector so that $\{\mathbf{u}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is also a positive orthonormal basis. The geodesic rotation determines the following images:

$$\mathbf{u}_1 = \frac{(b,-a,0)}{\sqrt{a^2+b^2}} \mapsto \frac{(b,-a,0)}{\sqrt{a^2+b^2}} = \mathbf{u}_1 \qquad \text{and} \qquad \mathbf{u}_2 = \frac{(a,b,0)}{\sqrt{a^2+b^2}} \mapsto \frac{(ac,bc,-(a^2+b^2))}{\sqrt{a^2+b^2}} = \mathbf{v}_2.$$

Note that

$$\mathbf{e}_1 = \frac{b}{\sqrt{a^2+b^2}}\mathbf{u}_1 + \frac{a}{\sqrt{a^2+b^2}}\mathbf{u}_2 \qquad \text{and} \qquad \mathbf{e}_2 = -\frac{a}{\sqrt{a^2+b^2}}\mathbf{u}_1 + \frac{b}{\sqrt{a^2+b^2}}\mathbf{u}_2.$$

Therefore, the desired rotation must satisfy

$$\mathbf{e}_1 \mapsto \frac{b}{\sqrt{a^2+b^2}}\mathbf{u}_1 + \frac{a}{\sqrt{a^2+b^2}}\mathbf{v}_2 = \frac{1}{a^2+b^2}(b^2+a^2c, -ab+abc, -a(a^2+b^2)),$$

$$\mathbf{e}_2 \mapsto -\frac{a}{\sqrt{a^2+b^2}}\mathbf{u}_1 + \frac{b}{\sqrt{a^2+b^2}}\mathbf{v}_2 = \frac{1}{a^2+b^2}(-ab+abc, a^2+b^2c, -b(a^2+b^2)),$$

and of course $\mathbf{e}_3 \mapsto \mathbf{v}_3 = (a,b,c)$. Thus, we can write

$$Z_2 = \left\{ L\begin{pmatrix} x \\ y \\ 0 \end{pmatrix} : (x,y) \in \mathbb{R}^2 \right\}$$

where $L : \mathbb{R}^3 \to \mathbb{R}^3$ by

$$L\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \dfrac{b^2+a^2c}{a^2+b^2} & -ab\dfrac{1-c}{a^2+b^2} & a \\ -ab\dfrac{1-c}{a^2+b^2} & \dfrac{a^2+b^2c}{a^2+b^2} & b \\ -a & -b & c \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Exercise 1 *We have not considered the following two cases:*

(a) $(n_1, n_2) = \mathbf{0} \in \mathbb{R}^2$ *with* $n_3 \neq 0$.

(b) $(n_1, n_2) \neq \mathbf{0} \in \mathbb{R}^2$ *with* $n_3 < 0$.

Fill in the details for these three cases. In case **(a)** *choose an appropriate vertical canonical plane, i.e., either the* x,z*-plane or the* y,z*-plane. (Or you can do both.)*

Finally, we attempt a similar visualization with $Z = \{(z_1, z_2) \in \mathbb{C}^2 : z_1\overline{n_1} + z_2\overline{n_2} = 0\}$. We can begin, if necessary, by replacing $\mathbf{n}$ with

$$(a + bi, c + di) = \frac{(n_1, n_2)}{\sqrt{|n_1|^2 + |n_2|^2}}.$$

We then have $a^2 + b^2 + c^2 + d^2 = 1$ and (a, b, c, d) is a unit vector in $\mathbb{R}^4$. We have observed above that

$$Z = \{(-(c - di), a - bi)z : z \in \mathbb{C}\}.$$

Writing $z = x + iy$, this means that under the identification

$$\mathbb{C}^2 \ni (x_1 + ix_2, x_3 + ix_4) \quad \longleftrightarrow \quad (x_1, x_2, x_3, x_4) \in \mathbb{R}^4$$

Z corresponds to

$$\{(-cx - dy, dx - cy, ax + by, -bx + ay) \in \mathbb{R}^4 : (x, y) \in \mathbb{R}^2\}.$$

Note that $(-cx - dy, dx - cy, ax + by, -bx + ay) = x(-c, d, a, -b) + y(-d, -c, b, a)$, and the two vectors $(-c, d, a, -b)$ and $(-d, -c, b, a)$ are elements of an orthonormal basis for a two-dimensional subspace of $\mathbb{R}^4$.

The complex (unit) normal $(a + bi, c + di) \in \mathbb{C}^2$ corresponds apparently to only one unit vector $(a, b, c, d) \in \mathbb{R}^4$. This vector is orthogonal to both $(-c, d, a, -b)$ and $(-d, -c, b, a)$, but we are apparently lacking a fourth vector to obtain a rotation $L : \mathbb{R}^4 \to \mathbb{R}^4$ given by

$$L \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -c & -d & a & * \\ d & -c & b & * \\ a & b & c & * \\ -b & a & d & * \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

and for which Z corresponds to $\{L\mathbf{x} : \mathbf{x} = (x, y, 0, 0)\}$ is the rotation of the x, y-plane corresponding to the canonical example

$$Z_0 = \{(z, 0) \in \mathbb{C}^2 : z \in \mathbb{C}\}$$

obtained when $\mathbf{n} = (0, n_3 + n_4 i) \in \mathbb{C}^2 \backslash \{(0, 0)\}$. If we look more carefully at the orthogonality condition

$$(x_1 + x_2 i)(a - bi) + (x_3 + x_4 i)(c - di) = 0,$$

however, we find this implies two real equations:

$$\begin{aligned} ax_1 + bx_2 + cx_3 + dx_4 &= 0 \\ -bx_1 + ax_2 - dx_3 + cx_4 &= 0. \end{aligned}$$

The first equation gives us the normal vector (a, b, c, d) we expected. But the second equation suggests a second vector $(-b, a, -d, c)$ which is indeed orthogonal to the three vectors we found/observed initially. Thus, by including $(-b, a, -d, c)$ as an image (with the appropriate sign) we get the desired rotation:

$$L\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -c & -d & a & -b \\ d & -c & b & a \\ a & b & c & -d \\ -b & a & d & c \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}. \tag{4}$$

Underlying this problem is the following question: We have a correspondence/identification between $\mathbb{C}^2$ and $\mathbb{R}^4$; if two vectors in $\mathbb{C}^2$ are orthogonal, satisfying

$$\langle \mathbf{z}, \mathbf{n} \rangle = 0,$$

what does that say about the corresponding vectors in $\mathbb{R}^4$? As we have seen above, if $\mathbf{z}$ corresponds to (x_1, x_2, x_3, x_4) and $\mathbf{n}$ corresponds to (a, b, c, d), then (x_1, x_2, x_3, x_4) and (a, b, c, d) are indeed orthogonal, but the orthogonality in $\mathbb{C}^2$ also gives another vector $(-b, a, -d, c)$ orthogonal to both (a, b, c, d) and (x_1, x_2, x_3, x_4).

Finally, it will be noted that the "new" orthogonal vector $(-b, a, -d, c)$ corresponds to

$$(-b + ai, -d + ci) = i(a + bi, c + di) \in \mathbb{C}^2.$$

This prompts another interesting observation: Given any vector $(x_1 + ix_2, x_3 + ix_4) \in \mathbb{C}^2$ corresponding to $(x_1, x_2, x_3, x_4) \in \mathbb{R}^4$, the vector

$$i(x_1 + ix_2, x_3 + ix_4) = (-x_2 + ix_1, -x_4 + ix_3) \in \mathbb{C}^2$$

corresponds to a vector $(-x_2, x_1, -x_4, x_3) \in \mathbb{R}^4$ which is "doubly orthogonal" to the vector (x_1, x_2, x_3, x_4) having first two coordinates given by the counterclockwise rotation (in $\mathbb{R}^2$) of the first two coordinates of (x_1, x_2, x_3, x_4) and similarly second two coordinates the counterclockwise rotation (in $\mathbb{R}^2$) of the second two coordinates of (x_1, x_2, x_3, x_4).

Exercise 2 *Is the matrix in* (4) *really a rotation matrix? (Is the determinant* 1*, or is it* -1*?*

Problem 2 *(norms and norm-like functions; Kreyszig Problems 2.2.3-4; see also Kreyszig Problems 4.3.1-2) Let X be a normed space.*

(a) *If $q : X \to [0, \infty)$ satisfies*

(i) $q(-x) = q(x)$ *for all $x \in X$, and*

(ii) $q(x + y) \leq q(x) + q(y)$ *for all $x, y \in X$.*

then show

$$|q(y) - q(x)| \leq q(y - x) \qquad \text{for all } x, y \in X.$$

In particular,

$$|\ \|y\| - \|x\|\ | \leq \|y - x\|.$$

(b) *If $q : X \to \mathbb{R}$ satisfies $q(ax) = aq(x)$ for all $a \geq 0$ and $x \in X$, then*

$$q(\mathbf{0}) = 0.$$

In particular, $\|\mathbf{0}\| = 0$, even if we only assume "$\|x\| = 0$ implies $x = \mathbf{0}$" instead of "$\|x\| = 0$ if and only if $x = \mathbf{0}$" in the definition of "norm."

(c) *If $p : X \to \mathbb{R}$ satisfies*

(i) $p(\alpha x) = |\alpha| p(x)$ *for all $x \in X$ and $\alpha \in F$, and*

(ii) $p(x + y) \leq p(x) + p(y)$ *for all $x, y \in X$.*

then show

$$p(x) \geq 0 \qquad \text{for all } x \in X.$$

In particular, the condition "$\| \cdot \| : X \to [0, \infty)$" can be relaxed to "$\| \cdot \| : X \to \mathbb{R}$" in the definition of "norm," and the function p itself is a **seminorm** *satisfying $p : X \to [0, \infty)$ and*

SN1 $p(\mathbf{0}) = 0$,

SN2 $p(\alpha x) = |\alpha| p(x)$ *for all $x \in X$, and*

SN3 $p(x + y) \leq p(x) + p(y)$.

See also Kreyszig Problem 2.3.12.

Problem 3 *(inner products and an inner product-like function) Let X be an inner product space. If $h : X \times X \to F$ satisfies*

(i) $h(x, x) \in \mathbb{R}$ *for all* $x \in X$*,*

(ii) $h(\alpha x + \beta y, z) = \alpha h(x, z) + \beta h(y, z)$ *for all* $\alpha, \beta \in F$ *and* $x, y, z \in X$*, and*

(ii) $h(z, \alpha x + \beta y) = \overline{\alpha} h(z, x) + \overline{\beta} h(z, y)$ *for all* $\alpha, \beta \in F$ *and* $x, y, z \in X$*,*

then show the following:

(a) $h(y, x) = \overline{h(x, y)}$ *for all* $x, y \in X$*. In particular, the conjugate symmetry*

$$\langle y, x\rangle = \overline{\langle x, y\rangle} \qquad \text{for all } x, y \in X$$

in the definition of the inner product can be replaced with conjugate linearity in the second argument:

$$\langle z, \alpha x + \beta y\rangle = \overline{\alpha}\langle x, z\rangle + \overline{\beta}\langle z, y\rangle \qquad \text{for all } x, y, z \in X \text{ and } \alpha, \beta \in F.$$

Hint: Consider $h(x + y, x + y)$ *and* $h(x + iy, x + iy)$*.*

(b) *(Kreyszig Problem 3.1.1; Exercise 57 in my notes)* $h(x+y, x+y)+h(x-y, x-y) = 2[h(x, x) + h(y, y)]$ *for all* $x, y \in X$*. In particular, the* **induced norm** *on an inner product space satisfies the* **parallelogram identity***:*

$$\|x + y\|^2 + \|x - y\|^2 = 2[\|x\|^2 + \|y\|^2] \qquad \text{for all } x, y \in X.$$

Problem 4 *(an inner product space; Kreyszig (sub)section 1.2-3) Let* $\ell^2 = \ell^2(F)$ *denote the collection of all sequences* $\{a_j\}_{j=1}^{\infty} \subset F$ *of scalars having the property that*

$$\sum_{j=1}^{\infty} |a_j|^2 < \infty.$$

Such sequences are said to be (absolutely) square summable.

(a) *Show* ℓ^2 *is a vector space by completing the following steps:*

1. *Show* ℓ^2 *is closed under scaling where the scaling* $\alpha\{a_j\}_{j=1}^{\infty}$ *is defined by*

$$\alpha\{a_j\}_{j=1}^{\infty} = \{\alpha a_j\}_{j=1}^{\infty}.$$

2. *Use the triangle inequality for the norm in* F^n *to conclude*

$$\sum_{j=1}^{n} |a_j + b_j|^2 \le \left[\sqrt{\sum_{j=1}^{n} |a_j|^2} + \sqrt{\sum_{j=1}^{n} |b_j|^2} \right]^2.$$

3. *Conclude* ℓ^2 *is closed under addition where the sum* $\{a_j\}_{j=1}^{\infty} + \{b_j\}_{j=1}^{\infty}$ *is defined by*

$$\{a_j + b_j\}_{j=1}^{\infty}.$$

(b) *Show that if* $\{a_j\}_{j=1}^{\infty}, \{b_j\}_{j=1}^{\infty} \in \ell^2$*, then*

$$\left\langle \{a_j\}_{j=1}^{\infty}, \{b_j\}_{j=1}^{\infty} \right\rangle = \sum_{j=1}^{\infty} a_j \overline{b_j}$$

is a well-defined inner product with value in F*. Hint: Use the Cauchy-Schwarz inequality in* F^n *like we used the triangle inequality in* F^n *in the previous part.*

Problem 5 *Give a detailed proof of the triangle inequality for the (induced) norm on* F^n*.*

Problem 6 *(another inner product space) Consider* $C([a,b] \to F)$ *the collection of all continuous scalar valued functions on the closed interval* $[a,b]$*. Note that one can define the integral of a continuous scalar valued function by*

$$\int_a^b f(x)\,dx = \int_a^b \mathrm{Re}[f(x)]\,dx + i\int_a^b \mathrm{Im}[f(x)]\,dx$$

where $\mathrm{Re}[f]$ *and* $\mathrm{Im}[f]$ *are the real and imaginary parts of the value of* f*. Show*

$$\langle f, g\rangle = \int_a^b f(x)\overline{g(x)}\,dx$$

defines an inner product on $C([a,b] \to F)$*.*

Problem 7 *(a normed space; exercises 58-59 in my notes) Show that* $C^0[a,b]$ *with the sup norm*

$$\|f\|_\infty = \max_{a \le x \le b} |f(x)|$$

is a normed space with a norm that is **not** *induced by an inner product. Hint: Consider positive functions* $f \equiv 1$ *and* $g(x) = x + c$*; check the paralellogram identity.*

Problem 8 *Recall that a function* $f : X \to Y$ *where* X *and* Y *are topological spaces is continuous at the point* $p \in X$ *if for every open set* V *in* Y *with* $f(p) \in V$*, there exists an open set* U *in* X *with* $p \in U$ *and*

$$f(U) = \{f(x) : x \in U\} \subset V.$$

Also, the function $f : X \to Y$ *is* **continuous** *if* f *is continuous at each point* $p \in X$*.*

Show this definition of continuity is equivalent to the condition:

The inverse image of every open set is open.

That is, for each open set V *in* Y*, the set*

$$f^{-1}(V) = \{x \in X : f(x) \in V\} \qquad \text{is open in } X.$$

Problem 9 *(Riesz representation) Let X be any inner product space.*

(a) *Show that given any vector $w \in X$, the function defined by*

$$\ell(x) = \langle x, w \rangle$$

is a continuous linear functional.

Definition: Given any inner product space X, let $\mathcal{R}(X \to F)$ denote the **represented linear functionals** on X. That is,

$$\mathcal{R}(X \to F) = \{\phi \in \mathcal{L}(X \to F) : \text{there exists some } w \in X \text{ such that } \phi(x) = \langle x, w \rangle \text{ for all } x \in X\}.$$

Part **(a)** *above asserts $\mathcal{R}(X \to F)$ is a subset of the (continuous) dual space $\daleth(X \to F)$ of bounded linear functionals. The Riesz representation theorem says that if X is a Hilbert space, then $\mathcal{R}(X \to F) = \daleth(X \to F)$. Part (a) also shows the assumption of continuity cannot be left out of the Riesz representation theorem. Parts* **(b)** *and* **(c)** *below give an explicit example.*

(b) *Let ℓ^2 denote the inner product space of square summable real sequences; see Problem 4 above. Let $\mathbf{e}_k = \{\delta_{kj}\}_{j=1}^{\infty} \in \ell^2$ be the sequence with all zeros except for 1 in the k-th entry. Notice that*

$$\{a_j\}_{j=1}^{\infty} = \sum_{j=1}^{\infty} a_j \mathbf{e}_j$$

is a convergent series in ℓ^2 for every $\{a_j\}_{j=1}^{\infty} \in \ell^2$. Consider the inner product space $W = \operatorname{span}\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \ldots\}$. Note that W is a proper subspace of ℓ^2. Show $\ell : W \to \mathbb{R}$ by

$$\ell\left(\{a_j\}_{j=1}^{\infty}\right) = \sum_{j=1}^{\infty} a_j$$

is a well-defined linear functional.

(c) *Show the linear functional ℓ from part* **(b)** *does not admit Riesz representation, that is, there is no sequence $w \in W$ (or even $w \in \ell^2$) for which $\ell(v) = \langle v, w \rangle$ for all $v \in W$. Hint: Show ℓ is discontinuous and use the assertion of part* **(a)** *above.*

Problem 10 *(Riesz representation) Our proof of the Riesz representation theorem, i.e., Riesz' proof, relied very strongly on the fact that we were able to take a nonzero vector in the orthogonal complement of a certain (proper) subspace. More specifically, we needed to find a nonzero vector in in the null space* $\mathcal{N}(\ell)^\perp$ *where* ℓ *was a continuous linear functional on a Hilbert space.*

Recall that given any subspace W *of an inner product space* X*, the* **orthogonal complement** *of* W *is defined by*

$$W^\perp = \{x \in X : \langle x, w \rangle = 0 \text{ for all } w \in W\}.$$

(a) *Show that* $W^\perp$ *is a subspace.*

(b) *Take it as given that* $L^2(a,b)$*, the collection of all square integrable functions* $f : (a,b) \to \mathbb{R}$ *satisfying*

$$\int_{(a,b)} f^2 < \infty,$$

is an inner product space with inner product

$$\langle f, g \rangle = \int_{(a,b)} fg.$$

(We will prove this in some detail later.) If you believe this, then it is clear that $W = C^0[a,b]$ *is a subspace of* $X = L^2(a,b)$*. Show* $W = C^0[a,b]$ *is a* **proper** *subspace of* $X = L^2(a,b)$*, i.e.,* $W \subsetneq X$*, but* $W^\perp = \{\mathbf{0}\}$*. (It may be a little difficult for you to give all the details correctly here, but you should be able to see the main idea.)*

(c) *It was also important in our proof that* $\mathcal{N}(\ell)^\perp$ *was one-dimensional that* $W^{\perp\perp} = W$ *for a certain subspace, namely for* $W = \operatorname{span}\{w\}$ *where* w *was the representing vector. Show that in general*

$$W \subset W^{\perp\perp} \qquad \text{for any subspace } W.$$

(d) *Give an example in which* $W \subsetneq W^{\perp\perp}$*. Hint: Look at part* **(b)** *above.*