Further Mathematical Methods (Linear Algebra) 2002

Solutions For Problem Sheet 3

In this Problem Sheet, we looked at some problems on real inner product spaces. In particular, we saw that *many* different inner products can be defined on a given vector space. We also used the Gram-Schmidt procedure to generate an orthonormal basis.

1. We are asked to verify that the Euclidean inner product of two vectors $\mathbf{x} = [x_1, x_2, \ldots, x_n]^t$ and $\mathbf{y} = [y_1, y_2, \ldots, y_n]^t$ in $\mathbb{R}^n$, i.e.

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n,$$

is indeed an inner product on $\mathbb{R}^n$. To do this, we must verify that it satisfies the definition of an inner product, i.e.

> **Definition:** An *inner product* on a **real** vector space V is a function that associates a real number $\langle \mathbf{u}, \mathbf{v} \rangle$ with each pair of vectors $\mathbf{u}$ and $\mathbf{v}$ in V in such a way that:
>
> **a**. $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$ and $\langle \mathbf{u}, \mathbf{u} \rangle = 0$ iff $\mathbf{u} = \mathbf{0}$.
>
> **b**. $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$.
>
> **c**. $\langle \alpha \mathbf{u} + \beta \mathbf{v}, \mathbf{w} \rangle = \alpha \langle \mathbf{u}, \mathbf{w} \rangle + \beta \langle \mathbf{v}, \mathbf{w} \rangle$.
>
> for all vectors $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and all scalars α and β.

Thus, taking any three vectors $\mathbf{x} = [x_1, x_2, \ldots, x_n]^t$, $\mathbf{y} = [y_1, y_2, \ldots, y_n]^t$ and $\mathbf{z} = [z_1, z_2, \ldots, z_n]^t$ in $\mathbb{R}^n$ and any two scalars α and β in $\mathbb{R}$ we have:

a. $\langle \mathbf{x}, \mathbf{x} \rangle = x_1^2 + x_2^2 + \cdots + x_n^2$ which is the sum of the squares of n real numbers and as such it is real and non-negative. Further, to show that $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ iff $\mathbf{x} = \mathbf{0}$, we note that:

- **LTR:** If $\langle \mathbf{x}, \mathbf{x} \rangle = 0$, then $x_1^2 + x_2^2 + \cdots + x_n^2 = 0$. But, this is the sum of n non-negative numbers and so it must be the case that $x_1 = x_2 = \cdots = x_n = 0$. Thus, $\mathbf{x} = \mathbf{0}$.
- **RTL:** If $\mathbf{x} = \mathbf{0}$, then $x_1 = x_2 = \cdots = x_n = 0$. Thus, $\langle \mathbf{x}, \mathbf{x} \rangle = 0$.

(as required).

b. Obviously, $\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n = y_1 x_1 + y_2 x_2 + \cdots + y_n x_n = \langle \mathbf{y}, \mathbf{x} \rangle$.

c. We note that the vector $\alpha \mathbf{x} + \beta \mathbf{y}$ is given by $[\alpha x_1 + \beta y_1, \alpha x_2 + \beta y_2, \ldots, \alpha x_n + \beta y_n]^t$ and so:

$$\begin{aligned} \langle \alpha \mathbf{x} + \beta \mathbf{y}, \mathbf{z} \rangle &= (\alpha x_1 + \beta y_1) z_1 + (\alpha x_2 + \beta y_2) z_2 + \cdots + (\alpha x_n + \beta y_n) z_n \\ &= \alpha (x_1 z_1 + x_2 z_2 + \cdots + x_n z_n) + \beta (y_1 z_1 + y_2 z_2 + \cdots + y_n z_n) \\ \therefore \quad \langle \alpha \mathbf{x} + \beta \mathbf{y}, \mathbf{z} \rangle &= \alpha \langle \mathbf{x}, \mathbf{z} \rangle + \beta \langle \mathbf{y}, \mathbf{z} \rangle \end{aligned}$$

Consequently, the Euclidean inner product is an inner product on $\mathbb{R}^n$ (as expected).

We are now given n positive real numbers $w_1, w_2, \ldots, w_n$ and vectors $\mathbf{x}$ and $\mathbf{y}$ as given above, and we are asked to verify that the formula

$$\langle \mathbf{x}, \mathbf{y} \rangle = w_1 x_1 y_1 + w_2 x_2 y_2 + \cdots w_n x_n y_n,$$

also defines an inner product on $\mathbb{R}^n$. So, we proceed as above by taking any three vectors $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$ in $\mathbb{R}^n$ and any two scalars α and β in $\mathbb{R}$, and show that this formula also satisfies the definition of an inner product. Thus, as

a. $\langle \mathbf{x}, \mathbf{x} \rangle = w_1 x_1^2 + w_2 x_2^2 + \cdots + w_n x_n^2$ which is the sum of the squares of n real numbers multiplied by the appropriate positive real number and, as such, it is real and non-negative. Further, to show that $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ iff $\mathbf{x} = \mathbf{0}$, we note that:

- **LTR:** If $\langle \mathbf{x}, \mathbf{x} \rangle = 0$, then $w_1 x_1^2 + w_2 x_2^2 + \cdots + w_n x_n^2 = 0$. But, this is the sum of n non-negative numbers multiplied by the appropriate positive real number and so it must be the case that $x_1 = x_2 = \cdots = x_n = 0$. Thus, $\mathbf{x} = \mathbf{0}$.
- **RTL:** If $\mathbf{x} = \mathbf{0}$, then $x_1 = x_2 = \cdots = x_n = 0$. Thus, $\langle \mathbf{x}, \mathbf{x} \rangle = 0$.

(as required).

b. Obviously, $\langle \mathbf{x}, \mathbf{y} \rangle = w_1 x_1 y_1 + w_2 x_2 y_2 + \cdots + w_n x_n y_n = w_1 y_1 x_1 + w_2 y_2 x_2 + \cdots + w_n y_n x_n = \langle \mathbf{y}, \mathbf{x} \rangle$.

c. We note that the vector $\alpha\mathbf{x} + \beta\mathbf{y}$ is given by $[\alpha x_1 + \beta y_1, \alpha x_2 + \beta y_2, \ldots, \alpha x_n + \beta y_n]^t$ and so:

$$\begin{aligned} \langle \alpha\mathbf{x} + \beta\mathbf{y}, \mathbf{z} \rangle &= w_1(\alpha x_1 + \beta y_1) z_1 + w_2(\alpha x_2 + \beta y_2) z_2 + \cdots + w_n(\alpha x_n + \beta y_n) z_n \\ &= \alpha(w_1 x_1 z_1 + w_2 x_2 z_2 + \cdots + w_n x_n z_n) + \beta(w_1 y_1 z_1 + w_2 y_2 z_2 + \cdots + w_n y_n z_n) \\ \therefore \quad \langle \alpha\mathbf{x} + \beta\mathbf{y}, \mathbf{z} \rangle &= \alpha\langle \mathbf{x}, \mathbf{z} \rangle + \beta\langle \mathbf{y}, \mathbf{z} \rangle \end{aligned}$$

this formula defines an inner product too (as required).[1]

2. To derive the vector equation of a plane in $\mathbb{R}^3$ with normal $\mathbf{n}$ and going through the point with position vector $\mathbf{a}$ we refer to the diagram in Figure 1. In this, the vector $\mathbf{r}$ represents any point in the plane, and so the vector formed by $\mathbf{r} - \mathbf{a}$ must lie in the plane. As such, this vector is orthogonal

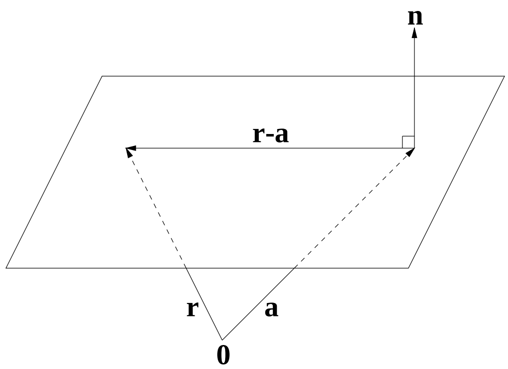

Figure 1: *The plane in* $\mathbb{R}^3$ *with normal* $\mathbf{n}$ *and going through the point with position vector* $\mathbf{a}$.

to the normal[2] $\mathbf{n}$ and so we have

$$\langle \mathbf{r} - \mathbf{a}, \mathbf{n} \rangle = 0.$$

But, by the third property of inner products (see Question 1), this means that

$$\langle \mathbf{r}, \mathbf{n} \rangle = \langle \mathbf{a}, \mathbf{n} \rangle,$$

which is the desired vector equation. To obtain the Cartesian equation of the plane, we write $\mathbf{r} = [x, y, z]^t$ and expand the inner products in the vector equation to get

$$a_1 x + a_2 y + a_3 z = p,$$

where $\mathbf{a} = [a_1, a_2, a_3]^t$ and p is the real number given by $\langle \mathbf{a}, \mathbf{n} \rangle$. If we now stipulate that $\mathbf{n}$ is a unit vector (where we denote this fact by writing $\mathbf{n}$ as $\hat{\mathbf{n}}$), then we have $\|\hat{\mathbf{n}}\| = 1$. This means that when we write the inner product represented by p in terms of the angle between the vectors $\mathbf{a}$ and $\hat{\mathbf{n}}$, we get

$$p = \langle \mathbf{a}, \hat{\mathbf{n}} \rangle = \|\mathbf{a}\| \|\hat{\mathbf{n}}\| \cos\theta = \|\mathbf{a}\| \cos\theta,$$

and looking at Figure 2, we can see that this implies that p represents the *perpendicular distance* from the plane to the origin.

[1] This is often called the *weighted* Euclidean inner product and the n positive real numbers w_1, w_2, ..., w_n are called the *weights*.

[2] As, by definition, the normal is orthogonal to *all* vectors in the plane.

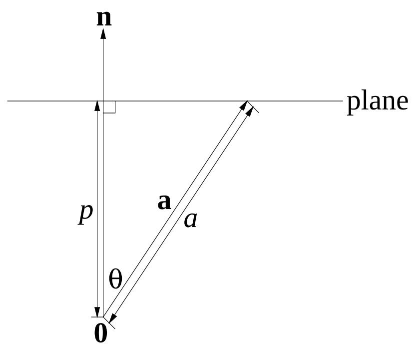

Figure 2: *A 'side-view' of the plane in Figure 1. We write* $\|\mathbf{a}\| = a$ *and notice how simple trigonometry dictates that* $p = a\cos\theta$ *is the perpendicular distance from the plane to the origin.*

Aside: In the last part of this question, some people make the error of supposing that what we have done so far justifies the assertion that p is the *shortest distance* from the plane to the origin. This is true, although a further argument is needed to establish this fact, namely:

> Let p' denote any distance from the plane to the origin. That is, by simple trigonometry (see Figure 3), we can see that $p = p'\cos\theta$. Thus, as $\cos\theta \leq 1$ this gives us $p \leq p'$, i.e. p is always smaller than (or equal to) p'.

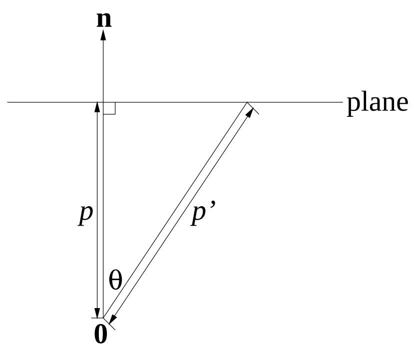

Figure 3: *A 'side-view' of the plane in Figure 1. Notice how simple trigonometry dictates that* $p = p'\cos\theta$*, and this leads us to the conclusion that* p *is also the shortest distance from the origin.*

We are also asked to calculate the quantities considered above for a particular plane, namely the one with normal $[2, 1, 2]^t$ and going through the point with position vector $[1, 2, 1]^t$. The vector equation of this plane is given by:

$$\left\langle \mathbf{r}, \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} \right\rangle = \left\langle \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} \right\rangle,$$

where $\mathbf{r}$ is a vector representing any point in the plane. Now, if we let $\mathbf{r} = [x, y, z]^t$ and evaluate the inner products in the vector equation we get

$$2x + y + 2z = 6,$$

which is the Cartesian equation of the plane in question. Lastly, we normalise the normal vector, i.e. as $\|[2, 1, 2]^t\|^2 = 9$, we set

$$\hat{\mathbf{n}} = \frac{1}{3}\begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix},$$

and so the perpendicular distance from the plane to the origin is

$$p = \langle \mathbf{a}, \hat{\mathbf{n}} \rangle = \frac{1}{3}\left\langle \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} \right\rangle = \frac{6}{3} = 2,$$

where there is no need to calculate the inner product here as it is just the left-hand-side of the Cartesian equation of the plane!

3. We are asked to prove that for all vectors $\mathbf{x}$ and $\mathbf{y}$ in a **real** inner product space the equalities

$$\|\mathbf{x}+\mathbf{y}\|^2 + \|\mathbf{x}-\mathbf{y}\|^2 = 2\|\mathbf{x}\|^2 + 2\|\mathbf{y}\|^2,$$

and

$$\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2 = 4\langle \mathbf{x}, \mathbf{y} \rangle,$$

hold. To do this we note that the norm of a vector $\mathbf{z}$, denoted by $\|\mathbf{z}\|$, is defined by

$$\|\mathbf{z}\|^2 = \langle \mathbf{z}, \mathbf{z} \rangle,$$

and that **real** inner products have the properties given by the Definition stated in Question 1. Thus, firstly, we have

$$\begin{aligned}
\|\mathbf{x}+\mathbf{y}\|^2 + \|\mathbf{x}-\mathbf{y}\|^2 &= \langle \mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y} \rangle + \langle \mathbf{x}-\mathbf{y}, \mathbf{x}-\mathbf{y} \rangle \\
&= \langle \mathbf{x}+\mathbf{y}, \mathbf{x} \rangle + \langle \mathbf{x}+\mathbf{y}, \mathbf{y} \rangle + \langle \mathbf{x}-\mathbf{y}, \mathbf{x} \rangle - \langle \mathbf{x}-\mathbf{y}, \mathbf{y} \rangle \\
&= \langle \mathbf{x}, \mathbf{x} \rangle + \langle \mathbf{y}, \mathbf{x} \rangle + \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{x} \rangle - \langle \mathbf{y}, \mathbf{x} \rangle - \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle \\
&= 2\langle \mathbf{x}, \mathbf{x} \rangle + 2\langle \mathbf{y}, \mathbf{y} \rangle \\
\therefore \quad \|\mathbf{x}+\mathbf{y}\|^2 + \|\mathbf{x}-\mathbf{y}\|^2 &= 2\|\mathbf{x}\|^2 + 2\|\mathbf{y}\|^2
\end{aligned}$$

and secondly, we have

$$\begin{aligned}
\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2 &= \langle \mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y} \rangle - \langle \mathbf{x}-\mathbf{y}, \mathbf{x}-\mathbf{y} \rangle \\
&= \langle \mathbf{x}+\mathbf{y}, \mathbf{x} \rangle + \langle \mathbf{x}+\mathbf{y}, \mathbf{y} \rangle - \langle \mathbf{x}-\mathbf{y}, \mathbf{x} \rangle + \langle \mathbf{x}-\mathbf{y}, \mathbf{y} \rangle \\
&= \langle \mathbf{x}, \mathbf{x} \rangle + \langle \mathbf{y}, \mathbf{x} \rangle + \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle - \langle \mathbf{x}, \mathbf{x} \rangle + \langle \mathbf{y}, \mathbf{x} \rangle + \langle \mathbf{x}, \mathbf{y} \rangle - \langle \mathbf{y}, \mathbf{y} \rangle \\
&= 2\langle \mathbf{y}, \mathbf{x} \rangle + 2\langle \mathbf{x}, \mathbf{y} \rangle \\
\therefore \quad \|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2 &= 4\langle \mathbf{x}, \mathbf{y} \rangle
\end{aligned}$$

as required.

We are also asked to give a geometric interpretation of the significance of the first equality in $\mathbb{R}^3$. To do this, we draw a picture of the plane in $\mathbb{R}^3$ which contains the vectors $\mathbf{x}$ and $\mathbf{y}$ (see Figure 4). Having done this, we see that the vector $\mathbf{x}$ and $\mathbf{y}$ can be used to construct the sides of a parallelogram

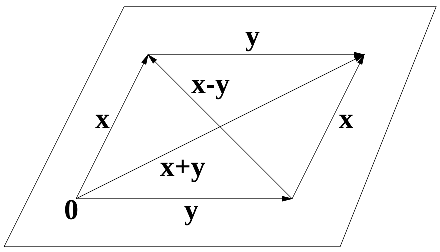

Figure 4: *The plane in* $\mathbb{R}^3$ *which contains the vectors* $\mathbf{x}$ *and* $\mathbf{y}$*. Notice that the vectors* $\mathbf{x}+\mathbf{y}$ *and* $\mathbf{x}-\mathbf{y}$ *form the diagonals of a parallelogram.*

whose diagonals are the vectors $\mathbf{x}+\mathbf{y}$ and $\mathbf{x}-\mathbf{y}$.[3] In this context, the quantities in our first equality measure the lengths of the sides and diagonals of this parallelogram, and so we can interpret it as saying:

> 'The sum of the squares of the sides of a parallelogram is equal to the sum of the squares of the diagonals.'

[3] This is effectively the 'parallelogram rule' for adding two vectors $\mathbf{x}$ and $\mathbf{y}$, or indeed, subtracting them.

Aside: Notice that, if $\mathbf{x}$ and $\mathbf{y}$ are orthogonal, then the shape in Figure 4 would be a rectangle. In this case, the first equality still has the same interpretation, although now, the lengths $\|\mathbf{x}+\mathbf{y}\|$ and $\|\mathbf{x}-\mathbf{y}\|$ are equal since the second equality tells us that $\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2 = 0$. (To my knowledge, there is no 'nice' interpretation of the second equality in the [general] parallelogram case.)

4. We are asked to consider the subspace $\mathbb{P}_n^{\mathbb{R}}$ of $\mathbb{F}^{\mathbb{R}}$ where we let x_0, x_1, $\dots$, x_n be $n+1$ fixed and distinct real numbers. In this case, we are required to show that for all vectors $\mathbf{p}$ and $\mathbf{q}$ in $\mathbb{P}_n^{\mathbb{R}}$ the formula

$$\langle \mathbf{p}, \mathbf{q} \rangle = \sum_{i=0}^{n} p(x_i)q(x_i),$$

defines an inner product on $\mathbb{P}_n^{\mathbb{R}}$. As we are working in a subspace of **real** function space, all of the quantities involved will be real and so all that we have to do is show that this formula satisfies all of the conditions in the Definition given in Question 1. Thus, taking any three vectors $\mathbf{p}$, $\mathbf{q}$ and $\mathbf{r}$ in $\mathbb{P}_n^{\mathbb{R}}$ and any two scalars α and β in $\mathbb{R}$ we have:

a. Clearly, since it is the sum of the squares of $n+1$ real numbers, we have

$$\langle \mathbf{p}, \mathbf{p} \rangle = \sum_{i=0}^{n} [p(x_i)]^2 \geq 0,$$

and it is real too. Further, to show that $\langle \mathbf{p}, \mathbf{p} \rangle = 0$ iff $\mathbf{p} = \mathbf{0}$ (where here, $\mathbf{0}$ is the zero polynomial), we note that:

- **LTR:** If $\langle \mathbf{p}, \mathbf{p} \rangle = 0$, then we have

$$\sum_{i=0}^{n} [p(x_i)]^2 = 0,$$

and as the right-hand-side of this expression is the sum of $n+1$ non-negative numbers, it must be the case that $p(x_i) = 0$ for $i = 0, 1, \dots, n$. However, this does not trivially imply that $\mathbf{p} = \mathbf{0}$ since the x_i could be roots of the polynomial $\mathbf{p}$, i.e. we could conceivably have $p(x_i) = 0$ for $i = 0, 1, \dots, n$ *and* $\mathbf{p} \neq \mathbf{0}$. So, to show that the **LTR** part of our biconditional is satisfied we need to discount this possibility, and this can be done by using the following argument:

> Assume, for contradiction, that $\mathbf{p} \neq \mathbf{0}$. That is, assume that $p(x)$ is a non-zero polynomial (i.e. at least one of the $n+1$ coefficients in $p(x) = a_0 + a_1 x + \cdots + a_n x^n$ is non-zero). Now, we can deduce that:
>
> – Since $p(x_i) = 0$ for $i = 0, 1, \dots, n$, this polynomial has $n+1$ distinct roots given by $x = x_0, x_1, \dots, x_n$.
>
> – Since $p(x)$ is a polynomial of degree at most n, it can have no more than n distinct roots.
>
> But, these two claims are contradictory and so it cannot be the case that $\mathbf{p} \neq \mathbf{0}$.

Thus, $\mathbf{p} = \mathbf{0}$.

- **RTL:** If $\mathbf{p} = \mathbf{0}$, then $\mathbf{p}$ is the polynomial that maps all $x \in \mathbb{R}$ to zero (i.e. it is the zero function) and as such, we have $p(x_i) = 0$ for $i = 0, 1, \dots, n$. Thus, $\langle \mathbf{p}, \mathbf{p} \rangle = 0$.

(as required).

b. It should be clear that $\langle \mathbf{p}, \mathbf{q} \rangle = \langle \mathbf{q}, \mathbf{p} \rangle$ since:

$$\langle \mathbf{p}, \mathbf{q} \rangle = \sum_{i=0}^{n} p(x_i)q(x_i) = \sum_{i=0}^{n} q(x_i)p(x_i) = \langle \mathbf{q}, \mathbf{p} \rangle.$$

c. We note that the vector $\alpha\mathbf{p} + \beta\mathbf{q}$ is just another polynomial and so:

$$\begin{aligned}\langle \alpha\mathbf{p} + \beta\mathbf{q}, \mathbf{r}\rangle &= \sum_{i=0}^{n} \left[\alpha p(x_i) + \beta q(x_i)\right] r(x_i) \\ &= \alpha \sum_{i=0}^{n} p(x_i)r(x_i) + \beta \sum_{i=0}^{n} q(x_i)r(x_i) \\ \therefore \quad \langle \alpha\mathbf{p} + \beta\mathbf{q}, \mathbf{r}\rangle &= \alpha\langle \mathbf{p}, \mathbf{r}\rangle + \beta\langle \mathbf{q}, \mathbf{r}\rangle\end{aligned}$$

as required.

Consequently, the formula given above does define an inner product on $\mathbb{P}_n^{\mathbb{R}}$ (as required).

5. We are asked to prove the following:

Theorem: Let $\mathbf{x}$ and $\mathbf{y}$ be any two non-zero vectors. If $\mathbf{x}$ and $\mathbf{y}$ are orthogonal, then they are linearly independent.

Proof: We are given that the two vectors $\mathbf{x}$ and $\mathbf{y}$ are non-zero and that they are orthogonal, i.e. $\langle \mathbf{x}, \mathbf{y}\rangle = 0$. To establish that they are linearly independent, we shall show that the vector equation

$$\alpha\mathbf{x} + \beta\mathbf{y} = \mathbf{0},$$

only has a trivial solution, namely $\alpha = \beta = 0$. To do this, we note that taking the inner product of this vector equation with $\mathbf{y}$ we get:

$$\begin{aligned}&\langle \alpha\mathbf{x} + \beta\mathbf{y}, \mathbf{y}\rangle = \langle \mathbf{0}, \mathbf{y}\rangle && \\ \therefore\ &\alpha\langle \mathbf{x}, \mathbf{y}\rangle + \beta\langle \mathbf{y}, \mathbf{y}\rangle = 0 && \text{:as } \langle \mathbf{0}, \mathbf{y}\rangle = 0. \\ \therefore\ &\beta\langle \mathbf{y}, \mathbf{y}\rangle = 0 && \text{:as } \mathbf{x} \text{ is orthogonal to } \mathbf{y}. \\ \therefore\ &\beta = 0 && \text{:as } \langle \mathbf{y}, \mathbf{y}\rangle \neq 0 \text{ since } \mathbf{y} \neq \mathbf{0}.\end{aligned}$$

and substituting this into the vector equation above, we get $\alpha\mathbf{x} = \mathbf{0}$ which implies that $\alpha = 0$ since $\mathbf{x} \neq \mathbf{0}$. Thus, our vector equation only has a trivial and so the vectors $\mathbf{x}$ and $\mathbf{y}$ are linearly independent (as required).

To see why the converse of this result, i.e.

> Let $\mathbf{x}$ and $\mathbf{y}$ be any two non-zero vectors. If $\mathbf{x}$ and $\mathbf{y}$ are linearly independent, then they are orthogonal.

does *not* hold we consider the two vectors $[1, 0]^t$ and $[1, 1]^t$ in the vector space $\mathbb{R}^2$. These two vectors are a counter-example to the claim above since they are linearly independent (as one is not a scalar multiple of the other), but they are not orthogonal (as $\langle [1, 0]^t, [1, 1]^t\rangle = 1 + 0 = 1 \neq 0$). Some examples of the use of this result and another counter-example to the converse are given in Question 7.

6. To show that the set of vectors $S = \{\mathbf{1}, \mathbf{x}, \mathbf{x^2}\} \subseteq \mathbb{P}_2^{[0,1]}$ is linearly independent, we note that the Wronskian of these functions is given by

$$W(x) = \begin{vmatrix} 1 & x & x^2 \\ 0 & 1 & 2x \\ 0 & 0 & 2 \end{vmatrix} = 2.$$

Thus, as this is non-zero for all $x \in [0, 1]$, it is non-zero for some $x \in [0, 1]$, and hence this set of functions is linearly independent (as required).

As such, these vectors will form a basis for the vector space $\mathbb{P}_2^{[0,1]}$ and we can use the Gram-Schmidt procedure to construct an orthonormal basis for this space. So, using the formula

$$\langle \mathbf{f}, \mathbf{g}\rangle = \int_0^1 f(x)g(x)dx,$$

to define an inner product on this vector space, we find that

- Taking $\mathbf{v}_1 = \mathbf{1}$, we get
$$\|\mathbf{v}_1\|^2 = \langle \mathbf{v}_1, \mathbf{v}_1 \rangle = \int_0^1 1\, dx = [x]_0^1 = 1,$$
and so we set $\mathbf{e}_1 = \mathbf{1}$.
- Taking $\mathbf{v}_2 = \mathbf{x}$, we construct the vector $\mathbf{u}_2$ where
$$\mathbf{u}_2 = \mathbf{v}_2 - \langle \mathbf{v}_2, \mathbf{e}_1 \rangle \mathbf{e}_1 = \mathbf{x} - \tfrac{1}{2}\,\mathbf{1},$$
since
$$\langle \mathbf{v}_2, \mathbf{e}_1 \rangle = \langle \mathbf{x}, \mathbf{1} \rangle = \int_0^1 x\, dx = \left[\frac{x^2}{2}\right]_0^1 = \frac{1}{2}.$$
Then, we need to normalise this vector, i.e. as
$$\|\mathbf{u}_2\|^2 = \langle \mathbf{u}_2, \mathbf{u}_2 \rangle = \int_0^1 \left(x - \frac{1}{2}\right)^2 dx = \left[\frac{1}{3}\left(x - \frac{1}{2}\right)^3\right]_0^1 = \frac{1}{12},$$
we set $\mathbf{e}_2 = \sqrt{3}(2\mathbf{x} - \mathbf{1})$.
- Taking $\mathbf{v}_3 = \mathbf{x^2}$, we construct the vector $\mathbf{u}_3$ where
$$\mathbf{u}_3 = \mathbf{v}_3 - \langle \mathbf{v}_3, \mathbf{e}_1 \rangle \mathbf{e}_1 - \langle \mathbf{v}_3, \mathbf{e}_2 \rangle \mathbf{e}_2 = \mathbf{x^2} - \mathbf{x} + \tfrac{1}{6}\,\mathbf{1},$$
since
$$\langle \mathbf{v}_3, \mathbf{e}_1 \rangle = \langle \mathbf{x^2}, \mathbf{1} \rangle = \int_0^1 x^2\, dx = \left[\frac{x^3}{3}\right]_0^1 = \frac{1}{3},$$
and,
$$\langle \mathbf{v}_3, \mathbf{e}_2 \rangle = \langle \mathbf{x^2}, \sqrt{3}(2\mathbf{x} - \mathbf{1}) \rangle = \sqrt{3}\int_0^1 x^2(2x-1)\, dx = \sqrt{3}\left[\frac{x^4}{2} - \frac{x^3}{3}\right]_0^1 = \frac{1}{\sqrt{12}}.$$
Then, we need to normalise this vector, i.e. as
$$\begin{aligned}\|\mathbf{u}_3\|^2 = \langle \mathbf{u}_3, \mathbf{u}_3 \rangle &= \int_0^1 \left(x^2 - x + \frac{1}{6}\right)^2 dx \\ &= \int_0^1 \left(x^4 - 2x^3 + \frac{4}{3}x^2 - \frac{1}{3}x + \frac{1}{36}\right) dx \\ &= \left[\frac{x^5}{5} - \frac{x^4}{2} + \frac{4}{9}x^3 - \frac{x^2}{6} + \frac{x}{36}\right]_0^1 \\ \therefore \quad \|\mathbf{u}_3\|^2 &= \frac{1}{180},\end{aligned}$$
we set $\mathbf{e}_3 = \sqrt{5}(6\mathbf{x^2} - 6\mathbf{x} + \mathbf{1})$.

Consequently, the set of vectors,
$$\left\{\mathbf{1}, \sqrt{3}(2\mathbf{x} - \mathbf{1}), \sqrt{5}(6\mathbf{x^2} - 6\mathbf{x} + \mathbf{1})\right\},$$
is an orthonormal basis for $\mathbb{P}_2^{[0,1]}$.

Lastly, we are asked to find a matrix A which will allow us to transform between coordinate vectors that are given relative to these two bases. That is, we are asked to find a matrix A such that, for any vector $\mathbf{v} \in \mathbb{P}_2^{[0,1]}$,
$$[\mathbf{v}]_S = \mathsf{A}[\mathbf{v}]_{S'}$$
where S' is the orthonormal basis. To do this, we note that a general vector given with respect to the basis S would be given by
$$\mathbf{v} = \alpha_1\mathbf{1} + \alpha_2\mathbf{x} + \alpha_3\mathbf{x^2},$$

and we could represent this by a vector in $\mathbb{R}^3$, namely $[\alpha_1, \alpha_2, \alpha_3]^t_S$, which is the coordinate vector of $\mathbf{v}$ relative to the basis S, i.e. $[\mathbf{v}]_S$.[4] Whereas, a general vector given with respect to the basis S' would be given by

$$\mathbf{v} = \alpha'_1 \mathbf{1} + \alpha'_2 \sqrt{3}(2\mathbf{x} - \mathbf{1}) + \alpha'_3 \sqrt{5}(6\mathbf{x^2} - 6\mathbf{x} + \mathbf{1}),$$

and we could also represent this by a vector in $\mathbb{R}^3$, namely $[\alpha'_1, \alpha'_2, \alpha'_3]^t_{S'}$, which is the coordinate vector of $\mathbf{v}$ relative to the basis S', i.e. $[\mathbf{v}]_{S'}$.[5] Thus, the matrix that we seek relates these two coordinate vectors and will therefore be 3×3. To find the numbers which this matrix will contain we note that, ultimately, the vector $\mathbf{v}$ is the same vector regardless of whether it is represented with respect to S or S'. As such, we note that

$$\alpha_1 \mathbf{1} + \alpha_2 \mathbf{x} + \alpha_3 \mathbf{x^2} = \alpha'_1 \mathbf{1} + \alpha'_2 \sqrt{3}(2\mathbf{x} - \mathbf{1}) + \alpha'_3 \sqrt{5}(6\mathbf{x^2} - 6\mathbf{x} + \mathbf{1}),$$

and so equating the coefficients of $\mathbf{1}, \mathbf{x}, \mathbf{x^2}$ on both sides we find that

$$\begin{aligned} \alpha_1 &= \alpha'_1 - \sqrt{3}\,\alpha'_2 + \sqrt{5}\,\alpha'_3, \\ \alpha_2 &= 2\sqrt{3}\,\alpha'_2 - 6\sqrt{5}\,\alpha'_3, \\ \alpha_3 &= 6\sqrt{5}\,\alpha'_3, \end{aligned}$$

respectively. Thus, we can write

$$[\mathbf{v}]_S = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}_S = \begin{bmatrix} 1 & -\sqrt{3} & \sqrt{5} \\ 0 & 2\sqrt{3} & -6\sqrt{5} \\ 0 & 0 & 6\sqrt{5} \end{bmatrix} \begin{bmatrix} \alpha'_1 \\ \alpha'_2 \\ \alpha'_3 \end{bmatrix}_{S'} = \mathsf{A}[\mathbf{v}]_{S'},$$

in terms of the matrix A given in the question.

Other Problems

The Other Problems on this sheet were intended to give you some further insight into what sort of formulae can be used to define inner products on certain subspaces of function space.

7. We are asked to consider the vector space of all smooth functions defined on the interval $[0, 1]$,[6] i.e. $\mathbb{S}^{[0,1]}$. Then, using the inner product defined by the formula

$$\langle \mathbf{f}, \mathbf{g} \rangle = \int_0^1 f(x) g(x)\, dx,$$

we are asked to find the inner products of the following pairs of functions and comment on the significance of these results in terms of the relationship between orthogonality and linear independence established in Question 5.

- The functions $\mathbf{f} : x \to \cos(2\pi x)$ and $\mathbf{g} : x \to \sin(2\pi x)$ have an inner product given by:

  $$\langle \mathbf{f}, \mathbf{g} \rangle = \int_0^1 \cos(2\pi x) \sin(2\pi x)\, dx = \frac{1}{2} \int_0^1 \sin(4\pi x)\, dx = \frac{1}{2} \left[-\frac{\cos(4\pi x)}{4\pi} \right]_0^1 = 0,$$

  where we have used the double-angle formula $\sin(2\theta) = 2 \sin\theta \cos\theta$ to simplify the integral. Further, we note that:

  - As this inner product is zero, the functions $\cos(2\pi x)$ and $\sin(2\pi x)$ are orthogonal.
  - As there is *no* $\alpha \in \mathbb{R}$ such that $\cos(2\pi x) = \alpha \sin(2\pi x)$, the functions $\cos(2\pi x)$ and $\sin(2\pi x)$ are linearly independent.

[4] See Definition 3.10.

[5] Again, see Definition 3.10.

[6] That is, the vector space formed by the set of all functions that are defined in the interval $[0, 1]$ and whose first derivatives exist at all points in this interval.

But, this is what we should expect from the result in Question 5.

- The functions $\mathbf{f} : x \to x$ and $\mathbf{g} : x \to e^x$ have an inner product given by:

$$\langle \mathbf{f}, \mathbf{g} \rangle = \int_0^1 xe^x \, dx = [xe^x]_0^1 - \int_0^1 e^x \, dx = e - [e^x]_0^1 = e - (e-1) = 1,$$

  where we have used integration by parts. Further, we note that:

  - As this inner product is non-zero, the functions x and e^x are *not* orthogonal.
  - As there is *no* $\alpha \in \mathbb{R}$ such that $x = \alpha e^x$, the functions x and e^x are linearly independent.

  Clearly, this is a counter-example to the converse of the result in Question 5.

- The functions $\mathbf{f} : x \to x$ and $\mathbf{g} : x \to 3x$ have an inner product given by:

$$\langle \mathbf{f}, \mathbf{g} \rangle = \int_0^1 3x^2 \, dx = 3 \left[\frac{x^3}{3} \right]_0^1 = 1.$$

  Further, we note that:

  - As this inner product is non-zero, the functions x and $3x$ are *not* orthogonal.
  - As the functions x and $3x$ are scalar multiples of one another, they are linearly dependent.

  But, this is what we should expect from [the contrapositive of] the result in Question 5.

8. We are asked to consider the subspace $\mathbb{P}_2^{\mathbb{R}}$ of $\mathbb{F}^{\mathbb{R}}$. Then, taking two general vectors in this space, say $\mathbf{p}$ and $\mathbf{q}$, where for all $x \in \mathbb{R}$,

$$\mathbf{p}(x) = a_0 + a_1 x + a_2 x^2 \quad \text{and} \quad \mathbf{q}(x) = b_0 + b_1 x + b_2 x^2,$$

respectively, we are required to show that the formula

$$\langle \mathbf{p}, \mathbf{q} \rangle = a_0 b_0 + a_1 b_1 + a_2 b_2,$$

defines an inner product on $\mathbb{P}_2^{\mathbb{R}}$. As we are working in a subspace of **real** function space, all of the quantities involved will be real and so all that we have to do is show that this formula satisfies all of the conditions in the Definition given in Question 1. Thus, taking any three vectors $\mathbf{p}$, $\mathbf{q}$ and $\mathbf{r}$ in $\mathbb{P}_2^{\mathbb{R}}$ and any two scalars α and β in $\mathbb{R}$ we have:

a. $\langle \mathbf{p}, \mathbf{p} \rangle = a_0^2 + a_1^2 + a_2^2$ which is the sum of the squares of three real numbers and as such it is real and non-negative. Further, to show that $\langle \mathbf{p}, \mathbf{p} \rangle = 0$ iff $\mathbf{p} = \mathbf{0}$ (where here, $\mathbf{0}$ is the zero polynomial), we note that:

- **LTR:** If $\langle \mathbf{p}, \mathbf{p} \rangle = 0$, then $a_0^2 + a_1^2 + a_2^2 = 0$. But, this is the sum of the squares of three real numbers and so it must be the case that $a_0 = a_1 = a_2 = 0$. Thus, $\mathbf{p} = \mathbf{0}$.
- **RTL:** If $\mathbf{p} = \mathbf{0}$, then $a_0 = a_1 = a_2 = 0$. Thus, $\langle \mathbf{p}, \mathbf{p} \rangle = 0$.

(as required).

b. Obviously, $\langle \mathbf{p}, \mathbf{q} \rangle = a_0 b_0 + a_1 b_1 + a_2 b_2 = b_0 a_0 + b_1 a_1 + b_2 a_2 = \langle \mathbf{q}, \mathbf{p} \rangle$.

c. We note that the vector $\alpha\mathbf{p} + \beta\mathbf{q}$ is just another quadratic and so:

$$\begin{aligned} \langle \alpha\mathbf{p} + \beta\mathbf{q}, \mathbf{r} \rangle &= (\alpha a_0 + \beta b_0)c_0 + (\alpha a_1 + \beta b_1)c_1 + (\alpha a_2 + \beta b_2)c_2 \\ &= \alpha(a_0 c_0 + a_1 c_1 + a_2 c_2) + \beta(b_0 c_0 + b_1 c_1 + b_2 c_2) \\ \therefore \quad \langle \alpha\mathbf{p} + \beta\mathbf{q}, \mathbf{r} \rangle &= \alpha \langle \mathbf{p}, \mathbf{r} \rangle + \beta \langle \mathbf{q}, \mathbf{r} \rangle \end{aligned}$$

where $\mathbf{r} : x \to c_0 + c_1 x + c_2 x^2$ for all $x \in \mathbb{R}$.

Consequently, the formula given above does define an inner product on $\mathbb{P}_2^{\mathbb{R}}$ (as required).

Harder Problems

The Harder Problems on this sheet were intended to give you some more practice in proving results about inner product spaces. In particular, we justify the assertion made in the lectures that several results that hold in real inner product spaces also hold in the complex case. We also investigate some other results related to the Cauchy-Schwarz inequality.

9. We are asked to verify that the Euclidean inner product of two vectors $\mathbf{x} = [x_1, x_2, \ldots, x_n]^t$ and $\mathbf{y} = [y_1, y_2, \ldots, y_n]^t$ in $\mathbb{C}^n$, i.e.

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1^* + x_2 y_2^* + \cdots x_n y_n^*$$

is indeed an inner product on $\mathbb{C}^n$. To do this, we must verify that it satisfies the definition of an inner product, i.e.

> **Definition:** An *inner product* on a **complex** vector space V is a function that associates a complex number $\langle \mathbf{u}, \mathbf{v} \rangle$ with each pair of vectors $\mathbf{u}$ and $\mathbf{v}$ in V in such a way that:
>
> **a.** $\langle \mathbf{u}, \mathbf{u} \rangle$ is a non-negative real number (i.e. $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$) and $\langle \mathbf{u}, \mathbf{u} \rangle = 0$ iff $\mathbf{u} = \mathbf{0}$.
>
> **b.** $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle^*$.
>
> **c.** $\langle \alpha\mathbf{u} + \beta\mathbf{v}, \mathbf{w} \rangle = \alpha\langle \mathbf{u}, \mathbf{w} \rangle + \beta\langle \mathbf{v}, \mathbf{w} \rangle$.
>
> for all vectors $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and all scalars α and β.

Thus, taking any three vectors $\mathbf{x} = [x_1, x_2, \ldots, x_n]^t$, $\mathbf{y} = [y_1, y_2, \ldots, y_n]^t$ and $\mathbf{z} = [z_1, z_2, \ldots, z_n]^t$ in $\mathbb{C}^n$ and any two scalars α and β in $\mathbb{C}$ we have:

a. Clearly, using the definition of the modulus of a complex number,[7] we have:

$$\langle \mathbf{x}, \mathbf{x} \rangle = x_1 x_1^* + x_2 x_2^* + \cdots + x_n x_n^* = |x_1|^2 + |x_2|^2 + \cdots + |x_n|^2 \geq 0,$$

as it is the sum of n non-negative real numbers and it is real too. Further, to show that $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ iff $\mathbf{x} = \mathbf{0}$, we note that:

- **LTR:** If $\langle \mathbf{x}, \mathbf{x} \rangle = 0$, then

$$x_1 x_1^* + x_2 x_2^* + \cdots + x_n x_n^* = |x_1|^2 + |x_2|^2 + \cdots + |x_n|^2 = 0.$$

  But, this is the sum of n non-negative real numbers and so it must be the case that $|x_1| = |x_2| = \cdots = |x_n| = 0$. However, this means that $x_1 = x_2 = \cdots = x_n = 0$ and so $\mathbf{x} = \mathbf{0}$.
- **RTL:** If $\mathbf{x} = \mathbf{0}$, then $x_1 = x_2 = \cdots = x_n = 0$. Thus, $\langle \mathbf{x}, \mathbf{x} \rangle = 0$.

(as required).

b. Obviously, $\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1^* + x_2 y_2^* + \cdots + x_n y_n^* = (y_1 x_1^* + y_2 x_2^* + \cdots + y_n x_n^*)^* = \langle \mathbf{y}, \mathbf{x} \rangle^*$.

c. We note that the vector $\alpha\mathbf{x} + \beta\mathbf{y}$ is given by $[\alpha x_1 + \beta y_1, \alpha x_2 + \beta y_2, \ldots, \alpha x_n + \beta y_n]^t$ and so:

$$\begin{aligned} \langle \alpha\mathbf{x} + \beta\mathbf{y}, \mathbf{z} \rangle &= (\alpha x_1 + \beta y_1) z_1^* + (\alpha x_2 + \beta y_2) z_2^* + \cdots + (\alpha x_n + \beta y_n) z_n^* \\ &= \alpha(x_1 z_1^* + x_2 z_2^* + \cdots + x_n z_n^*) + \beta(y_1 z_1^* + y_2 z_2^* + \cdots + y_n z_n^*) \\ \therefore\quad \langle \alpha\mathbf{x} + \beta\mathbf{y}, \mathbf{z} \rangle &= \alpha\langle \mathbf{x}, \mathbf{z} \rangle + \beta\langle \mathbf{y}, \mathbf{z} \rangle \end{aligned}$$

[7] The modulus of a complex number z, denoted by $|z|$, is defined by the formula $|z|^2 = zz^*$. Writing $z = a + ib$ (where $a, b \in \mathbb{R}$) we note that $|z|^2 = (a + ib)(a - ib) = a^2 + b^2 \geq 0$. Consequently, we can see that $|z|^2$ is real and non-negative.

Consequently, the Euclidean inner product is an inner product on $\mathbb{C}^n$ (as expected).

Further, we are reminded that the norm of a vector $\mathbf{x} \in \mathbb{C}^n$, denoted by $\|\mathbf{x}\|$, is defined by the formula

$$\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle},$$

and we are asked to use this to prove that the following theorems hold in *any* complex inner product space. Firstly, we have:

Theorem: [The Cauchy-Schwarz Inequality] If $\mathbf{x}$ and $\mathbf{y}$ are vectors in $\mathbb{C}^n$, then

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\| \, \|\mathbf{y}\|.$$

Proof: For any vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{C}^n$ and an arbitrary scalar $\alpha \in \mathbb{C}$, we can write:

$$\begin{aligned}
& \|\mathbf{x} + \alpha \mathbf{y}\|^2 \geq 0 \\
\therefore \; & \langle \mathbf{x} + \alpha \mathbf{y}, \mathbf{x} + \alpha \mathbf{y} \rangle \geq 0 \\
\therefore \; & \langle \mathbf{x}, \mathbf{x} + \alpha \mathbf{y} \rangle + \alpha \langle \mathbf{y}, \mathbf{x} + \alpha \mathbf{y} \rangle \geq 0 \\
\therefore \; & \langle \mathbf{x}, \mathbf{x} \rangle + \alpha^* \langle \mathbf{x}, \mathbf{y} \rangle + \alpha \langle \mathbf{y}, \mathbf{x} \rangle + \alpha \alpha^* \langle \mathbf{y}, \mathbf{y} \rangle \geq 0 \\
\therefore \; & \|\mathbf{x}\|^2 + 2\mathrm{Re}[\alpha^* \langle \mathbf{x}, \mathbf{y} \rangle] + |\alpha|^2 \|\mathbf{y}\|^2 \geq 0
\end{aligned}$$

where we have used the fact that

$$\alpha^* \langle \mathbf{x}, \mathbf{y} \rangle + \alpha \langle \mathbf{y}, \mathbf{x} \rangle = \alpha^* \langle \mathbf{x}, \mathbf{y} \rangle + [\alpha^* \langle \mathbf{x}, \mathbf{y} \rangle]^* = 2\mathrm{Re}[\alpha^* \langle \mathbf{x}, \mathbf{y} \rangle],$$

which is two times the real part of $\alpha^* \langle \mathbf{x}, \mathbf{y} \rangle$.[8] Now, the quantities $\langle \mathbf{x}, \mathbf{y} \rangle$ and α in this expression will be complex numbers and so we can write them in polar form, i.e.

$$\langle \mathbf{x}, \mathbf{y} \rangle = Re^{i\phi} \quad \text{and} \quad \alpha = re^{i\theta},$$

where $R = |\langle \mathbf{x}, \mathbf{y} \rangle|$ and $r = |\alpha|$ are real numbers.[9] So, writing the left-hand-side of our inequality as

$$\Delta = \|\mathbf{x}\|^2 + 2\mathrm{Re}\left[rRe^{i(\phi - \theta)}\right] + r^2 \|\mathbf{y}\|^2,$$

and noting that α was an arbitrary scalar, we can choose α so that its argument (i.e. θ) is such that $\theta = \phi$, i.e. we now have

$$\Delta = \|\mathbf{x}\|^2 + 2rR + r^2 \|\mathbf{y}\|^2.$$

But, this is a real quadratic in r, and so we can see that:

- Either: $\Delta > 0$ for all values of r in which case, the quadratic function represented by Δ never crosses the r-axis and so it will have no real roots (i.e. the roots will be complex),
- **Or:** $\Delta \geq 0$ for all values of r in which case, the quadratic function represented by Δ never crosses the r-axis, but it is tangential to it at some value of r, and so we will have repeated real roots.

So, considering the general real quadratic $ar^2 + br + c$, we know that it has roots given by:

$$r = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a},$$

and this gives no real roots, or repeated real roots, if $b^2 - 4ac \leq 0$. Consequently, our conditions for Δ are equivalent to saying that

$$4R^2 - 4\|\mathbf{x}\|^2 \|\mathbf{y}\|^2 \leq 0,$$

[8] As, if we have a complex number $z = a + ib$ (where $a, b \in \mathbb{R}$), then $z + z^* = (a + ib) + (a - ib) = 2a$ and $a = \mathrm{Re}\, z$.
[9] Since, if $z = re^{i\theta}$, then $|z|^2 = zz^* = (re^{i\theta})(re^{-i\theta}) = r^2$.

which on re-arranging gives

$$|\langle \mathbf{x}, \mathbf{y} \rangle|^2 \leq \|\mathbf{x}\|^2 \|\mathbf{y}\|^2,$$

where we have used the fact that $R = |\langle \mathbf{x}, \mathbf{y} \rangle|$. Thus, taking square roots and noting that both $\|\mathbf{x}\|$ and $\|\mathbf{y}\|$ are non-negative, we get

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\| \, \|\mathbf{y}\|,$$

as required.

and secondly, we have:

Theorem: [The Triangle Inequality] If $\mathbf{x}$ and $\mathbf{y}$ are vectors in $\mathbb{C}^n$, then

$$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|.$$

Proof: For any vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{C}^n$ we can write:

$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\|^2 &= \langle \mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y} \rangle \\ &= \langle \mathbf{x}, \mathbf{x} + \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{x} + \mathbf{y} \rangle \\ &= \langle \mathbf{x}, \mathbf{x} \rangle + \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{x} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle \\ \therefore \quad \|\mathbf{x} + \mathbf{y}\|^2 &= \|\mathbf{x}\|^2 + 2\text{Re}[\langle \mathbf{x}, \mathbf{y} \rangle] + \|\mathbf{y}\|^2 \end{aligned}$$

where we have used the fact that

$$\langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{x} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{y} \rangle^* = 2\text{Re}[\langle \mathbf{x}, \mathbf{y} \rangle],$$

which is two times the real part of $\langle \mathbf{x}, \mathbf{y} \rangle$. Now, the quantity, $\langle \mathbf{x}, \mathbf{y} \rangle$ is complex and so we can write

$$\langle \mathbf{x}, \mathbf{y} \rangle = Re^{i\phi},$$

and, as in the previous proof, this means that $R = |\langle \mathbf{x}, \mathbf{y} \rangle|$. But, then we have

$$\text{Re}[\langle \mathbf{x}, \mathbf{y} \rangle] = \text{Re}[Re^{i\phi}] = R\cos\phi \leq R = |\langle \mathbf{x}, \mathbf{y} \rangle|,$$

since $e^{i\phi} = \cos\phi + i\sin\phi$ (Euler's formula) and $\cos\phi \leq 1$. So, our expression for $\|\mathbf{x} + \mathbf{y}\|^2$ can now be written as

$$\|\mathbf{x} + \mathbf{y}\|^2 \leq \|\mathbf{x}\|^2 + 2|\langle \mathbf{x}, \mathbf{y} \rangle| + \|\mathbf{y}\|^2,$$

which, on applying the Cauchy-Schwarz inequality gives

$$\|\mathbf{x} + \mathbf{y}\|^2 \leq \|\mathbf{x}\|^2 + 2\|\mathbf{x}\|\|\mathbf{y}\| + \|\mathbf{y}\|^2,$$

or indeed,

$$\|\mathbf{x} + \mathbf{y}\|^2 \leq (\|\mathbf{x}\| + \|\mathbf{y}\|)^2.$$

Thus, taking square roots on both sides and noting that both $\|\mathbf{x}\|$ and $\|\mathbf{y}\|$ are non-negative, we get

$$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|,$$

as required.

and thirdly, bearing in mind that two vectors $\mathbf{x}$ and $\mathbf{y}$ are *orthogonal*, written $\mathbf{x} \perp \mathbf{y}$, if $\langle \mathbf{x}, \mathbf{y} \rangle = 0$, we have:

Theorem: [The Generalised Theorem of Pythagoras] If $\mathbf{x}$ and $\mathbf{y}$ are vectors in $\mathbb{C}^n$ and $\mathbf{x} \perp \mathbf{y}$, then $\|\mathbf{x}+\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$.

Proof: For any vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{C}^n$, we know from the previous proof that

$$\|\mathbf{x}+\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + 2\text{Re}[\langle \mathbf{x}, \mathbf{y} \rangle] + \|\mathbf{y}\|^2.$$

Now, if these vectors are orthogonal, then $\langle \mathbf{x}, \mathbf{y} \rangle = 0$, and so this expression becomes

$$\|\mathbf{x}+\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2,$$

as required.

10. We are told to use the Cauchy-Schwarz inequality, i.e.

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\| \|\mathbf{y}\|,$$

for all vectors $\mathbf{x}$ and $\mathbf{y}$ in some vector space to prove that

$$(a \cos\theta + b \sin\theta)^2 \leq a^2 + b^2,$$

for all real values of a, b and θ. As we only need to consider real values of a, b and θ, we shall work in $\mathbb{R}^n$. Further, we shall choose $n = 2$ as the question revolves around spotting that you need to consider the two vectors $[a, b]^t$ and $[\cos\theta, \sin\theta]^t$. So, using these, we find that

$$\left| \left\langle \begin{bmatrix} a \\ b \end{bmatrix}, \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \right\rangle \right| \leq \left\| \begin{bmatrix} a \\ b \end{bmatrix} \right\| \left\| \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \right\|,$$

which gives

$$|a \cos\theta + b \sin\theta| \leq \sqrt{a^2 + b^2} \sqrt{\cos^2\theta + \sin^2\theta},$$

Now noting that $\cos^2\theta + \sin^2\theta = 1$ (trigonometric identity) and squaring both sides we get

$$(a \cos\theta + b \sin\theta)^2 \leq a^2 + b^2,$$

as required.

11. We are asked to prove that the equality in the Cauchy-Schwarz inequality holds iff the vectors involved are linearly dependent, that is, we are asked to prove that

Theorem: For any two vectors $\mathbf{x}$ and $\mathbf{y}$ in a vector space V: The vectors $\mathbf{x}$ and $\mathbf{y}$ are linearly dependent iff $|\langle \mathbf{x}, \mathbf{y} \rangle| = \|\mathbf{x}\| \, \|\mathbf{y}\|$

Proof: Let $\mathbf{x}$ and $\mathbf{y}$ be any two vectors in a vector space V. We have to prove an 'iff' statement and so we have to prove it 'both ways,' i.e.

- **LTR:** If the vectors $\mathbf{x}$ and $\mathbf{y}$ are linearly dependent, then [without loss of generality, we can assume that] there is a scalar α such that $\mathbf{x} = \alpha\mathbf{y}$. As such, we have

  $$|\langle \mathbf{x}, \mathbf{y} \rangle| = |\langle \alpha\mathbf{y}, \mathbf{y} \rangle| = |\alpha \langle \mathbf{y}, \mathbf{y} \rangle| = |\alpha| \|\mathbf{y}\|^2,$$

  and,

  $$\|\mathbf{x}\| \, \|\mathbf{y}\| = \|\alpha\mathbf{y}\| \, \|\mathbf{y}\| = |\alpha| \|\mathbf{y}\|^2.$$

  So, equating these two expressions we get $|\langle \mathbf{x}, \mathbf{y} \rangle| = \|\mathbf{x}\| \, \|\mathbf{y}\|$, as required.

- **RTL:** Assume that the vectors $\mathbf{x}$ and $\mathbf{y}$ are such that

  $$|\langle \mathbf{x}, \mathbf{y} \rangle| = \|\mathbf{x}\| \, \|\mathbf{y}\|.$$

  Considering the proof of the Cauchy-Schwarz inequality in Question 9, we can see that this equality holds in the case where

  $$\|\mathbf{x} + \alpha \mathbf{y}\|^2 \geq 0,$$

  i.e. there exists a scalar α such that

  $$\|\mathbf{x} + \alpha \mathbf{y}\|^2 = \langle \mathbf{x} + \alpha \mathbf{y}, \mathbf{x} + \alpha \mathbf{y} \rangle = 0.$$

  (In particular, this is the scalar $\alpha = re^{i\theta}$ where $\theta = \phi$ and r gives the repeated root of the real quadratic Δ.) But, this equality can only hold if $\mathbf{x} + \alpha \mathbf{y} = \mathbf{0}$ and this implies that $\mathbf{x} = -\alpha \mathbf{y}$, i.e. $\mathbf{x}$ and $\mathbf{y}$ are linearly dependent, as required.

Thus, the vectors $\mathbf{x}$ and $\mathbf{y}$ are linearly dependent iff $|\langle \mathbf{x}, \mathbf{y} \rangle| = \|\mathbf{x}\| \, \|\mathbf{y}\|$, as required.