Inner Product Spaces

Introduction

Recall in the lecture on vector spaces that geometric vectors (i.e. vectors in two and three-dimensional Cartesian space) have the properties of addition, subtraction, a zero vector, additive inverses, scalability, magnitude (i.e. length), and angle. Moreover, the formal definition of vector space includes all these properties except for the last two: magnitude and angle. In this lecture we introduce a vector operation that allows us to define length and angle for vectors in an arbitrary vector space. This operation is called the "inner product between two vectors", and is a generalization of the dot product that was introduced in the Matrices lecture.

The Dot Product Revisited

Recall that if u and v are vectors in $\mathcal{R}^n$ then the dot product $u \cdot v$ is defined as

$$u \cdot v = u_1 v_1 + \cdots + u_n v_n.$$

Recall also how this operation is used when computing the entry of a the product $C = AB$ of two matrices. In particular $c_{ij} = a_{i,} \cdot b_{,j}$.

We now describe how to use the dot product to obtain the length of a vector, and the angle between two vectors in $\mathcal{R}^2$. Note that the same can also be done for vectors in $\mathcal{R}^1$ and $\mathcal{R}^3$, but it seems more illustrative and straightforward to show it for $\mathcal{R}^2$.

Example 1. Show that the length of the vector $v = (x, y)$ in $\mathcal{R}^2$, denoted $|v|$ is equal to $\sqrt{x^2 + y^2}$.

In conclusion, the length of vector $v = (x, y)$ is given as

$$|v| = \sqrt{x^2 + y^2} = \sqrt{v \cdot v},$$

and so the length of a vector in $\mathcal{R}^2$ can be solely expressed in terms of the dot product.

The next step is to show how the angle between two vectors u and v can be expressed solely in terms of the dot product.

Example 2. Show that the angle between the vectors $u = (x_1, y_1)$ and $v = (x_2, y_2)$ can be given as

$$cos^{-1}(\frac{u \cdot v}{|u||v|}).$$

Conclude that the angle between two vectors can be expressed solely by the dot product.

Example 3. Given $u = (3, 5)$ and $v = (-2, 1)$ determine the lengths of these vectors and the angle between them.

The Inner Product

Our goal in this section is to generalize the concept of dot product so that the generalization may be applied to any vector space. Notice how the dot product relies on a vector having components. But not all vector spaces have vectors that have obvious components. For example, the function space $\mathcal{F}(\mathcal{R}, \mathcal{R})$ has vectors that are functions that do not necessarily have components. Therefore, we cannot simply define the generalized dot product in terms of a formula that involves vector components, as was done for the dot product. The next best idea is to list all of the important algebaric properties of the dot product, and require that the generalized dot product have these properties.

As a first step, notice that the dot product is a function that accepts two vector inputs u and v, and returns a real-valued output $u \cdot v$. Therefore, the generalized dot product should also be a function that takes in two vectors u and v, and returns a real-valued output. But so as not to confuse the generalized dot product with the original dot product, we use the notation $<u, v>$ to denote the real-valued output. Moreover, instead of using the term "generalized dot product", we prefer the term "inner product". Thus, given a vector space $\mathcal{V}$, we say that $<, >$ is an **inner product** on $\mathcal{V}$ iff $<, >$ is a function that takes two vectors $u, v \in \mathcal{V}$ and returns the real number $<u, v>$. Moreover, $<, >$ must satisfy the following properties.

Symmetry $<u, v> = <v, u>$

Additivity $<u, v + w> = <u, v> + <u, w>$

Scalar Associativity $k<u, v> = <ku, v>$

Positivity $<u,u> \geq 0$, and $<u,u> = 0$ iff $u = 0$.

It is left as an exercise to show that for $\mathcal{V} = \mathcal{R}^2$, if $<u,v> = u \cdot v$, then the above four properties are satisfied. In other words, the dot product is itself an innter product. If an inner product is defined on a vector space $\mathcal{V}$, then $\mathcal{V}$ is called an **inner-product space**.

Proposition 1. An inner product on a vector space $\mathcal{V}$ satisfies the following additional properties.

1. $<0,u> = 0$
2. $<u+v,w> = <u,w> + <v,w>$
3. $k<u,v> = <u,kv>$

Proof of Proposition 1. Property 2 is proved by invoking the symmetry and additivity properties, while Property 3 is proved by invoking the symmetry and the scalar associativity properties. As for Property 1,

$$<0,u> = <0u,u> = 0<u,u> = 0.$$

Examples

Example 4. Let

$$A = \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} \text{ and } B = \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix}$$

be real-valued matrices, and hence matrices of vector space $\mathcal{M}_{22}$. We show that $\mathcal{M}_{22}$ is an inner product space by defining

$$<A,B> = a_1a_2 + b_1b_2 + c_1c_2 + d_1d_2.$$

It remains to show that $<,>$ has all the requisite properties.

Symmetry

$$<A,B> = a_1a_2 + b_1b_2 + c_1c_2 + d_1d_2 =$$

$$a_2a_1 + b_2b_1 + c_2c_1 + d_2d_1 = <B,A>.$$

Additivity Let

$$C = \begin{pmatrix} a_3 & b_3 \\ c_3 & d_3 \end{pmatrix}.$$

Then,

$$<A+B,C> = (a_1+a_2)a_3 + (b_1+b_2)b_3 + (c_1+c_2)c_3 + (d_1+d_2)d_3 =$$

$$a_1a_3 + a_2a_3 + b_1b_3 + b_2b_3 + c_1c_3 + c_2c_3 + d_1d_3 + d_2d_3 =$$

$$(a_1a_3 + b_1b_3 + c_1c_3 + d_1d_3) + (a_2a_3 + b_2b_3 + c_2c_3 + d_2d_3) = <A, C> + <B, C>.$$

Scalar Associativity

$$k<A, B> = k(a_1a_2 + b_1b_2 + c_1c_2 + d_1d_2) =$$

$$k(a_1a_2) + k(b_1b_2) + k(c_1c_2) + k(d_1d_2) =$$

$$(ka_1)a_2 + (kb_1)b_2 + (kc_1)c_2 + (kd_1)d_2 = <kA, B>.$$

Positivity

$$<A, A> = a_1^2 + b_1^2 + c_1^2 + d_1^2 \geq 0,$$

and $<A, A> = 0$ implies $A = 0$, since this is only possible when all entries of A are zero.

Example 5. Let $\mathcal{C}[a, b]$ denote the set of all real-valued functions that are continuous on the closed interval $[a, b]$. Since the sum of any two continuous functions is also continuous and f continuous implies (kf) is continuous, for any scalar k, it follows that $\mathcal{C}[a, b]$ is a vector space (the proof of this is exactly the same as the proof that $\mathcal{F}(\mathcal{R}, \mathcal{R})$ is a vector space). Moreover, recall from calculus that any continuous function over a closed interval can be integrated over that interval. Thus, given continuous functions $f, g \in \mathcal{C}[a, b]$, define

$$<f, g> = \int_a^b f(x)g(x)dx.$$

Prove that $<, >$ has all the requisite properties of an inner product.

Example 5 Solution.

Measuring Length and Angle in Inner Product Spaces

Recall that for vectors in $\mathcal{R}^2$ the length of a vector and angle between two vectors can be expressed entirely in terms of the dot product. To review,

$$|v| = \sqrt{v \cdot v},$$

while

$$\theta(u, v) = cos^{-1}(\frac{u \cdot v}{\sqrt{u \cdot u}\sqrt{v \cdot v}}).$$

So it makes sense to use these definitions for length and angle in an inner product space. In particular, if vector space $\mathcal{V}$ has defined inner product $<,>$, then the **length** of vector $v \in \mathcal{V}$ is defined as

$$|v| = \sqrt{<v, v>},$$

while the **angle** between vectors u and v is defined as

$$\theta(u, v) = cos^{-1}(\frac{<u, v>}{\sqrt{<u, u>}\sqrt{<v, v>}}).$$

Do the above definitions make sense? The following are three fundamental properties of length relating to geometrical vectors that ought to also be valid for vectors in an inner product space.

Positivity $|v| \geq 0$ and $|v| = 0$ iff $v = 0$.

Scalar Proportional Length $|kv| = |k||v|$.

Triangle Inequality $|u + v| \leq |u| + |v|$.

It is left as an exercise to prove that the length definition for vectors in any inner product space does indeed satisfy the above properties.

As for the angle definition $\theta(u, v)$ we must verify that the expression

$$\frac{<u, v>}{\sqrt{<u, u>}\sqrt{<v, v>}}$$

is in fact a number between -1 and 1, since this is the valid range of the $\cos^{-1}$ function. To prove this we need the following important theorem.

Theorem 1 (Cauchy-Schwarz-Bunyakovsky Inequality). If u and v are vectors in some inner product space, then

$$<u, v>^2 \leq <u, u><v, v>.$$

Consequently,

$$-1 \leq \frac{<u, v>}{\sqrt{<u, u>}\sqrt{<v, v>}} \leq 1.$$

Proof of Theorem 1. For any scalar t, by several applications of the four properties of inner products, we get

$$0 \leq <tu + v, tu + v> = t^2<u, u> + 2t<u, v> + <v, v>,$$

which may be written as $at^2 + bt + c \geq 0$, where $a = <u, u>$, $b = 2<u, v>$, and $c = <v, v>$. But $at^2 + bt + c \geq 0$ implies that the equation $at^2 + bt + c = 0$ either has no roots, or exactly one root. In other words, we must have

$$b^2 - 4ac \leq 0,$$

which implies

$$4<u, v>^2 \leq 4<u, u><v, v>,$$

or

$$<u, v>^2 \leq <u, u><v, v>.$$

Another way of writing the Cauchy inequality is

$$|<u, v>| \leq |u||v|,$$

which is obtained by taking the square root of both sides and realizing that, e.g., $|u| = \sqrt{<u, u>}$.

Example 6. Let

$$A = \begin{pmatrix} -1 & 3 \\ 2 & -2 \end{pmatrix} \text{ and } B = \begin{pmatrix} 5 & -4 \\ 0 & 2 \end{pmatrix}$$

be vectors of the inner product space from Example 4. Determine the lengths of A and B, and $\theta(A, B)$.

Example 7. Let $f(x) = x$ and $g(x) = \sin x$ be two vectors of the inner product space from Example 5, where we assume $[a, b] = [0, 2\pi]$. Determine the lengths of f and g, and $\theta(f, g)$.

Unit and Orthogonal Vectors

Vector v of an inner product space $\mathcal{V}$ is said to be a **unit vector** iff $|v| = 1$. For example, in $\mathcal{R}^2$, $(0, 1)$ and $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ are examples of unit vectors, since

$$|(0,1)| = \sqrt{0^2 + 1^2} = 1,$$

and

$$|(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})| = \sqrt{(\frac{1}{\sqrt{2}})^2 + (\frac{1}{\sqrt{2}})^2} = \sqrt{1/2 + 1/2} = 1.$$

Note that, for any nonzero vector v, we can always find a unit vector that has the same direction as v. This is done by scaling v with $1/|v|$. Indeed

$$|\frac{1}{|v|}v| = \frac{1}{|v|} \cdot |v| = 1.$$

When a nonzero vector v gets replaced by $\frac{1}{|v|}v$, we say that v has been **normalized**, or that $\frac{1}{|v|}v$ is the normalization of v.

Example 8. Normalize the vectors $u = (3, 4)$, $v = (1, 5, 2)$,

$$A = \begin{pmatrix} -1 & 3 \\ 2 & -2 \end{pmatrix},$$

and $f(x) = x^2$, where the latter is considered a vector of $\mathcal{C}[0, 1]$.

In addition to unit vectors, another important concept is that of perpendicularity. Recall that two geometric vectors u and v are perpendicular provided they make a $90°$ angle. In terms of the dot product, this would imply that

$$\cos 90° = 0 = \frac{u \cdot v}{|u||v|},$$

which implies that $u \cdot v = 0$. This gives us an idea of how to generalize the concept of perpendicularity to an arbitrary inner product space. However, the term "orthogonal" is used instead of the term "perpendicular". To generalize, vectors u and v in some inner product space are said to be **orthogonal** iff

$$<u, v> = 0.$$

Example 8. Show that the following pairs of vectors are orthogonal.

a. $u = (1, -1, 3)$ and $v = (4, -2, -2)$

b. Matrices

$$A = \begin{pmatrix} -1 & 3 \\ 2 & -2 \end{pmatrix} \text{ and } B = \begin{pmatrix} 5 & -1 \\ 0 & -4 \end{pmatrix}$$

c. Functions $f(x) = \cos(2x)$ and $g(x) = \cos(3x)$ that belong to $\mathcal{C}[0, 2\pi]$.

Orthonormal Bases

Recall that a basis for a vector space $\mathcal{V}$ is a linearly independent set of vectors B that spans $\mathcal{V}$. Moreover, basis B is said to be **orthonormal** iff

1. each basis vector is a unit vector, and
2. for any two basis vectors $e_1, e_2 \in B$, $<e_1, e_2> = 0$; i.e. e_1 and e_2 are orthogonal.

If B is an orthonormal basis consisting of n vectors, then we denote them by $e_1, \ldots, e_n$.

As an example, $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, and $e_3 = (0, 0, 1)$ is an orthonormal basis for $\mathcal{R}^3$, since all three vectors are clearly unit vectors (verify!), and, if $i \neq j$,

$$<e_i, e_j> = e_i \cdot e_j = 0,$$

since e_i's only nonzero component is i, while e_j's only nonzero component is j, and $i \neq j$. For example,

$$<e_1, e_3> = e_1 \cdot e_3 = (1)(0) + (0)(0) + (0)(1) = 0.$$

Example 9. Consider the subspace $\mathcal{W}$ of $\mathcal{C}[0, 2\pi]$ that is spanned by the vectors

$$1, \cos x, \cos 2x, \ldots.$$

Verify that these vectors are pairwise orthogonal, and then normalize them to form an orthonormal basis for $\mathcal{W}$.

The following theorem provides one reason for the importance of orthonormal bases. It states that, given an orthonormal basis, one can readily write any vector v as a linear combination of the basis vectors, since the e_i coefficient is none other than $<v, e_i>$.

Theorem 2. Suppose $e_1, \dots, e_n$ is an orthonormal basis for an inner product space $\mathcal{V}$. Then for any vector $v \in \mathcal{V}$,

$$v = <v, e_1>e_1 + \cdots + <v, e_n>e_n.$$

Proof of Theorem 2. Since $e_1, \dots, e_n$ is a basis for $\mathcal{V}$, we have

$$v = c_1 e_1 + \cdots + c_n e_n,$$

for some coefficients $c_1, \dots, c_n$. Then,

$$<v, e_i> = <e_i, v> = <e_i, c_1 e_1 + \cdots + c_n e_n> =$$

$$<e_i, c_1 e_1> + \cdots + <e_i, c_n e_n> = c_1<e_i, e_1> + \cdots + c_n<e_i, e_n>.$$

But all the terms in the last expression are zero (why?) except for possibly

$$c_i<e_i, e_i> = c_i|e_i|^2 = c_i 1^2 = c_i.$$

Therefore, it must be the case that $c_i = <v, e_i>$, and the proof is complete.

Theorem 3. Suppose $v_1, \dots, v_n$ are pairwise orthogonal nonzero vectors; i.e., $<v_i, v_j> = 0$ for all $i \neq j$. Then $v_1, \dots, v_n$ are linearly independent.

Proof of Theorem 3. Suppose

$$c_1 v_1 + \cdots + c_n v_n = 0.$$

Then

$$<v_i, 0> = 0 = <v_i, c_1 v_1 + \cdots + c_n v_n> = <v_i, c_1 v_1> + \cdots + <v_i, c_n v_n> =$$

$$c_1<v_i, v_1> + \cdots + c_n<v_i, v_n>.$$

But all the terms in the last expression are zero (since the entire sum equals 0). In particular

$$c_i<v_i, v_i> = c_i|v_i|^2 = 0.$$

But since $|v_i| \neq 0$, it follows that $c_i = 0$. Therefore, the equation

$$c_1 v_1 + \cdots + c_n v_n = 0$$

has only the trivial solution, and $v_1, \dots, v_n$ are linearly independent.

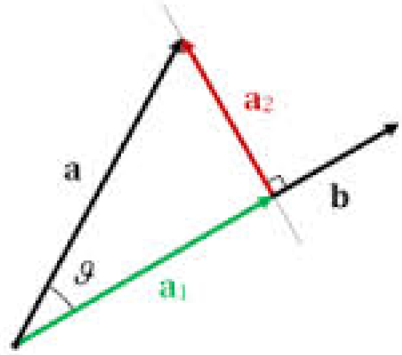

Figure 1: Projection of vector a on to the subspace spanned by b

Gram-Schmidt Algorithm

Suppose $v_1, \ldots, v_n$ is a basis for inner product space $\mathcal{V}$, but not necessarily an orthonormal basis. We now provide an algorithm, called the **Gram-Schmidt Orthonormalization Algorithm**, for converting a basis $v_1, \ldots, v_n$ into an orthonormal basis $e_1, \ldots, e_n$. To describe the algorithm, it is helpful to have the notion of the orthogonal projection of a vector on to a subspace.

Let $\mathcal{W}$ be a finite-dimensional subspace of vector space $\mathcal{V}$, and having orthonormal basis $e_1, \ldots, e_n$. Given $v \in \mathcal{V}$, the **projection of v on to $\mathcal{W}$**, denoted $\text{proj}(v, \mathcal{W})$, is defined as the vector

$$<v, e_1>e_1 + \cdots + <v, e_n>e_n.$$

Later we show that this definition is independent of the given orthonormal basis (i.e. the same projection vector will be computed, regardless of chosen orthonormal basis). We may think of $\text{proj}(v, \mathcal{W})$ as the component of v that lies in $\mathcal{W}$, while $v - \text{proj}(v, \mathcal{W})$ is the component of v that lies outside of $\mathcal{W}$. Figure 1 shows an example of this. Here the subspace $\mathcal{W}$ is the space spanned by vector b. Moreover, a_1 is the component of a that lies in $\mathcal{W}$, while a_2 is the component of a that lies outside of of $\mathcal{W}$. Notice how a_2 looks perpendicular to b. This is no coincidence. It turns out that the "outside" component is always orthogonal to $\mathcal{W}$. This is proven in the following theorem.

Theorem 4. Suppose $\mathcal{W}$ is a finite-dimensional subspace of a vector space $\mathcal{V}$. Then for all $v \in \mathcal{V}$, $v - \text{proj}(v, \mathcal{W})$ is orthogonal to $\mathcal{W}$, meaning that

$$<v - \text{proj}(v, \mathcal{W}), w> = 0,$$

for all $w \in \mathcal{W}$.

Proof of Theorem 4. Let $e_1, \ldots, e_n$ be an orthonormal basis for $\mathcal{W}$ (later we will prove that every finite-dimensional vector space has such a basis). Then it suffices to prove that

$$<v - \text{proj}(v, \mathcal{W}), e_i> = 0,$$

for each $i = 1, \ldots, n$ (see Exercise 14). But

$$<v - \text{proj}(v, \mathcal{W}), e_i> = <v - <v, e_1>e_1 + \cdots + <v, e_n>e_n, e_i> =$$

$$<v, e_i> - <v, e_i><e_i, e_i> = <v, e_i> - <v, e_i> = 0.$$

We are now in a position to describe the Gram-Schmidt algorithm for converting a finite basis $v_1, \ldots, v_n$ to an orthonormal basis $e_1, \ldots, e_n$. The algorithm works in stages.

Stage 1 Let $e_1 = v_1/|v_1|$. Thus $\text{span}(v_1) = \text{span}(e_1)$, since the vectors are multiples of each other.

Stage k: $k \geq 2$ Assume that $e_1, \ldots, e_{k-1}$ have been constructed, where each e_i is a unit vector, and the vectors are pairwise orthogonal. Moreover, assume that

$$\text{span}(v_1, \ldots, v_{k-1}) = \text{span}(e_1, \ldots, e_{k-1}).$$

Then set

$$e_k = v_k - \text{proj}(v_k, \{e_1, \ldots, e_{k-1}\})/|(v_k - \text{proj}(v_k, \{e_1, \ldots, e_{k-1}\})|.$$

In other words, e_k is the normalization of the difference of v_k with its projection on to the subspace generated by $e_1, \ldots, e_{k-1}$. Notice that $v_k - \text{proj}(v_k, \{e_1, \ldots, e_{k-1}\}) \neq 0$, since, otherwise, it would imply that v_k is in

$$\text{span}(e_1, \ldots, e_{k-1}) = \text{span}(v_1, \ldots, v_{k-1}),$$

which contradicts the linear independence of $v_1, \ldots, v_k$. Finally, by Theorem 4, e_k is orthogonal to each of the e_i, $i = 1, \ldots, k-1$.

Example 10. Use the Gram-Schmidt algorithm for converting basis $v_1 = (1, 0, 0)$, $v_2 = (1, 1, 0)$, $v_3 = (1, 1, 1)$, to an orthonormal basis for $\mathcal{R}^3$.

Example 10 Solution. Since $v_1 = (1, 0, 0)$ is already a unit vector, we have $e_1 = v_1$.

Next

$$v_2 - \text{proj}(v_2, \{e_1\}) = v_2 - <v_2, e_1>e_1 = (1, 1, 0) - (1, 0, 0) = (0, 1, 0).$$

Since this vector is already normalized, we have $e_2 = (0, 1, 0)$.

Finally,

$$v_3 - \text{proj}(v_3, \{e_1, e_2\}) = v_3 - <v_3, e_1>e_1 - <v_3, e_2>e_2 = (1, 1, 1) - (1, 0, 0) - (0, 1, 0) = (0, 0, 1).$$

Again, since this vector is already normalized, we have $e_3 = (0, 0, 1)$. Therefore, the orthonormal basis is $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, $e_3 = (0, 0, 1)$.

Example 11. Repeat the previous example, but now assume the basis ordering is $v_1 = (1, 1, 1)$, $v_2 = (1, 1, 0)$, $v_3 = (1, 0, 0)$.

More on Projections

Recall that $\text{proj}(v, \mathcal{W})$ is defined in terms of an orthonormal basis for finite-dimensional subspace $\mathcal{W}$. In this section we first prove that $\text{proj}(v, \mathcal{W})$ is independent of the chosen basis. This is proved in the following theorem.

Theorem 5 (Projection Theorem). Let $\mathcal{W}$ be a finite-dimensional subspace of a vector space $\mathcal{V}$. Then, every $v \in \mathcal{V}$ can be uniquely expressed as

$$v = w + w^{\perp},$$

where $w \in \mathcal{W}$ and $w^{\perp}$ is orthogonal to $\mathcal{W}$.

Proof of Theorem 5. Let $v \in \mathcal{V}$ be arbitrary. Using the Gram-Schmidt algorithm, one can compute an orthonormal basis $e_1, \ldots, e_n$ for $\mathcal{W}$. This basis can then be used to compute $w = \text{proj}(v, \mathcal{W})$ and $w^{\perp} = v - \text{proj}(v, \mathcal{W})$. Then by Theorem 4, $w \in \mathcal{W}$ and $w^{\perp}$ is orthogonal to $\mathcal{W}$. Moreover, $v = w + w^{\perp}$. It remains to show that w and $w^{\perp}$ are unique.

Suppose also that $v = w_2 + w_2^{\perp}$, where $w_2 \in \mathcal{W}$ and $w_2^{\perp}$ is orthogonal to $\mathcal{W}$. Then

$$w + w^{\perp} = w_2 + w_2^{\perp},$$

which implies

$$w - w_2 = w^{\perp} - w_2^{\perp}.$$

Now, since $w^{\perp}$ and $w_2^{\perp}$ are both orthogonal to $\mathcal{W}$, then so is $w^{\perp} - w_2^{\perp}$, since, for any vector $u \in \mathcal{W}$,

$$<w^{\perp} - w_2^{\perp}, u> = <w^{\perp}, u> - <w_2^{\perp}, u> = 0 - 0 = 0.$$

Now, since $w - w_2 = w^{\perp} - w_2^{\perp}$, it follows that $w_1 - w_2$ is orthogonal to every vector in $\mathcal{W}$. But $w_1 - w_2$ is itself in $\mathcal{W}$. Hence,

$$<w_1 - w_2, w_1 - w_2> = 0,$$

which, by the positivity axiom, implies that $w_1 - w_2 = 0$; i.e. $w_1 = w_2$. Therefore, we must also have $w^{\perp} = w_2^{\perp}$, and this proves that w and $w^{\perp}$ are unique with respect to v, and are independent of the chosen orthonormal basis for $\mathcal{W}$.

The next fact to establish about $\text{proj}(v, \mathcal{W})$ is that it is the closest vector in $\mathcal{W}$ to v. In other words, if we want to approximate v with some vector in $\mathcal{W}$, then $\text{proj}(v, \mathcal{W})$ is the best choice, since it minimizes the distance $|v - w|$ amongst all possible vectors $w \in \mathcal{W}$. To prove this we first need to establish the "Pythagorean Theorem" for inner product spaces.

Theorem 6 (Pythagorean Theorem). If $\mathcal{V}$ is an inner product space with inner product $<,>$ and $u, v \in \mathcal{V}$ satisfy $<u, v> = 0$, then

$$|u + v|^2 = |u|^2 + |v|^2.$$

Proof of Theorem 6. Suppose $u, v \in \mathcal{V}$ satisfy $<u, v> = 0$. Then

$$|u+v|^2 = <u+v, u+v> = <u, u> + <v, v> + 2<u, v> = |u|^2 + |v|^2.$$

Theorem 7. Let $\mathcal{W}$ be a finite-dimensional subspace of a vector space $\mathcal{V}$. Then for $v \in \mathcal{V}$

$$\operatorname*{argmin}_{w \in \mathcal{W}} |v - w| = \text{proj}(v, \mathcal{W}),$$

meaning that $w = \text{proj}(v, \mathcal{W})$ minimizes the quantity $|v - w|$, amongst all possible $w \in \mathcal{W}$.

Proof of Theorem 7. Let $v \in \mathcal{V}$ be arbitrary. Then for any $w \in \mathcal{W}$, we may write

$$v - w = (v - \text{proj}(v, \mathcal{W})) + (\text{proj}(v, \mathcal{W}) - w).$$

Now, $v - \text{proj}(v, \mathcal{W})$ is orthogonal to $\mathcal{W}$ by Theorem 4, while $\text{proj}(v, \mathcal{W}) - w$ is a vector in $\mathcal{W}$. Hence, $v - \text{proj}(v, \mathcal{W})$ is orthogonal to $\text{proj}(v, \mathcal{W}) - w$, and, by the Pythagorean Thoerem,

$$|v - w|^2 = |v - \text{proj}(v, \mathcal{W})|^2 + |\text{proj}(v, \mathcal{W}) - w|^2.$$

In the above equation, notice how the first term on the right depends only on v. Thus, if $|v - w|^2$ (and hence $|v - w|$) is to be minimized, then the second term on the right must be made as small as possible. But this term can be made to equal zero by assigning $w = \text{proj}(v, \mathcal{W})$. Therefore, $w = \text{proj}(v, \mathcal{W})$ is the vector in $\mathcal{W}$ that minimizes $|v - w|$.

Exercises

1. Find the angle that is made between each of the following pairs of vectors.

   a. $u = (1, 2)$, $v = (6, 8)$
   b. $u = (-7, -3)$, $v = (0, 1)$
   c. $u = (1, -3, 7)$, $v = (8, -2, -2)$
   d. $u = (-3, 1, 2)$, $v = (4, 2, -5)$

2. Show that for $\mathcal{V} = \mathcal{R}^2$, if $<u, v>$ is defined by $<u, v> = u \cdot v$, then all four inner-product properties are satisfied.

3. For $u, v \in \mathcal{R}^2$, define function $<u, v>$ as
   $$<u, v> = 2u_1v_1 + 6u_2v_2.$$
   Prove that $<, >$ is an inner product on $\mathcal{R}^2$.

4. For $p(x), q(x) \in \mathcal{P}_2$, define the function $<p, q>$ as
   $$<p, q> = p(0)q(0) + p(1/2)q(1/2) + p(1)q(1).$$
   Prove that $<, >$ is an inner product on $\mathcal{P}_2$.

5. For $p(x), q(x) \in \mathcal{P}_2$, define the function $<p, q>$ as
   $$<p, q> = p(0)q(0) + p(1)q(1).$$
   Explain why $<, >$ is *not* an inner product on $\mathcal{P}_2$.

6. Let
   $$A = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \text{ and } B = \begin{pmatrix} 1 & -2 \\ 3 & 0 \end{pmatrix}$$
   be vectors of the inner product space from Example 4. Determine the lengths of A and B, and $\theta(A, B)$.

7. Let
   $$A = \begin{pmatrix} 2 & 6 \\ 1 & -3 \end{pmatrix} \text{ and } B = \begin{pmatrix} 3 & 2 \\ 1 & 0 \end{pmatrix}$$
   be vectors of the inner product space from Example 4. Determine the lengths of A and B, and $\theta(A, B)$.

8. Let $p(x) = x^2$ and $q(x) = x$. Use the inner product from Exercise 4 to determine $|p|$, $|q|$, and $\theta(p, q)$.

9. Let $p(x) = x^2 - x + 2$ and $q(x) = 5x^2 + 1$. Use the inner product from Exercise 4 to determine $|p|$, $|q|$, and $\theta(p, q)$.

10. Use the inner product from Example 5 on the vector space $\mathcal{C}[-1, 1]$ to find $|f|$, $|g|$, and $\theta(f, g)$ for $f(x) = x$ and $g(x) = x^2$.

11. Use the inner product from Example 5 on the vector space $\mathcal{C}[0, 2\pi]$ to find $|f|$, $|g|$, and $\theta(f, g)$ for $f(x) = x$ and $g(x) = \cos x$.

12. Let
$$A = \begin{pmatrix} 2 & -3 \\ 1 & 5 \end{pmatrix} \text{ and } B = \begin{pmatrix} -1 & 3 \\ -3 & 0 \end{pmatrix}$$
be vectors of the inner product space from Example 4. Normalize A and B.

13. Let $p(x) = 4x^2 + 6x + 1$ and $q(x) = 10x - 4$. Use the inner product from Exercise 4 to normalize p and q.

14. Let $e_1, \ldots, e_n$ be an orthonormal basis for a subspace $\mathcal{W}$ of vector space V. Prove that if v is orthogonal to each e_i, then v is orthogonal to every vector $w \in \mathcal{W}$.

15. Is $v_1 = (2/3, -2/3, 1/3)$, $v_2 = (2/3, 1/3, -2/3)$, and $v_3 = (1/3, 2/3, 2/3)$ an orthonormal set of vectors?

16. Is $v_1 = (1, 0, 0)$, $v_2 = (0, 1/\sqrt{2}, 1/\sqrt{2})$, and $v_3 = (0, 0, 1)$ an orthonormal set of vectors?

17. Use the Gram-Schmidt algorithm to transform $v_1 = (1, -3)$, $v_2 = (2, 2)$ into an orthonormal basis.

18. Use the Gram-Schmidt algorithm to transform $v_1 = (1, 0)$, $v_2 = (3, -5)$ into an orthonormal basis.

19. Use the Gram-Schmidt algorithm to transform $v_1 = (1, 1, 1)$, $v_2 = (-1, 1, 0)$ $v_3 = (1, 2, 1)$ into an orthonormal basis.

20. Use the Gram-Schmidt algorithm to transform $v_1 = (1, , 0, 0)$, $v_2 = (3, 7, -2)$ $v_3 = (0, 4, 1)$ into an orthonormal basis.

21. Given $v = (-1, 2, 6, 0)$, write $v = w + w^{\perp}$, where w is in the subspace spanned by $u_1 = (-1, 0, 1, 2)$ and $u_2 = (0, 1, 0, 1)$, and $w^{\perp}$ is orthogonal to this subspace.

22. Verify that $e_1 = (1/\sqrt{10}, -3\sqrt{10})$, $e_2 = (3/\sqrt{10}, (1/\sqrt{10})$ is an orthonormal basis. Determine coefficients c_1 and c_2 for which $v = (-1, 4) = c_1 e_1 + c_2 e_2$

23. Verify that $e_1 = (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})$, $e_2 = (-1/\sqrt{2}, 1/\sqrt{2}, 0)$, $e_3 = (1/\sqrt{6}, 1/\sqrt{6}, -2/\sqrt{6})$ is an orthonormal basis. Determine coefficients c_1, c_2, and c_3 for which $v = (0.5, -2, 3) = c_1 e_1 + c_2 e_2 + c_3 e_3$.

24. Verify that $e_1 = (1, 0, 0)$, $e_2 = (0, 7/\sqrt{53}, -2/\sqrt{53})$, $e_3 = (0, 30/\sqrt{11925}, 105/\sqrt{11925})$ is an orthonormal basis. Determine coefficients c_1, c_2, and c_3 for which $v = (0.5, -2, 3) = c_1 e_1 + c_2 e_2 + c_3 e_3$.

Exercise Solutions

1. a) 10.3°, b) 113.2°, c) 0°, d) 142.8°

2. Let $u = (u_1, u_2)$, $v = (v_1, v_2)$, and $w = (w_1, w_2)$.

   Symmetry:
   $$u \cdot v = u_1v_1 + u_2v_2 = v_1u_1 + v_2u_2 = v \cdot u.$$

   Additivity:
   $$u \cdot (v + w) = u_1(v_1 + w_1) + u_2(v_2 + w_2) = (u_1v_1 + u_1w_1) + (u_2v_2 + u_2w_2) =$$
   $$(u_1v_1 + u_2v_2) + (u_1w_1 + u_2w_2) = u \cdot v + u \cdot w.$$

   Scalar Associativity:
   $$k(u \cdot v) = k(u_1v_1 + u_2v_2) = k(u_1v_1) + k(u_2v_2) = (ku_1)v_1 + (ku_2)v_2 = (ku) \cdot v.$$

   Positivity:
   $$u \cdot u = u_1^2 + u_2^2 \geq 0,$$
   and can only equal zero when $u_1 = u_2 = 0$; i.e. $u = 0$.

3. Let $u = (u_1, u_2)$, $v = (v_1, v_2)$, and $w = (w_1, w_2)$.

   Symmetry:
   $$< u, v > = 2u_1v_1 + 6u_2v_2 = 2v_1u_1 + 6v_2u_2 = < v, u >.$$

   Additivity:
   $$< u, v + w > = 2u_1(v_1 + w_1) + 6u_2(v_2 + w_2) = (2u_1v_1 + 2u_1w_1) + (6u_2v_2 + 6u_2w_2) =$$
   $$(2u_1v_1 + 6u_2v_2) + (2u_1w_1 + 6u_2w_2) = < u, v > + < u, w >.$$

   Scalar Associativity:
   $$k < u, v > = k(2u_1v_1 + 6u_2v_2) = k(2u_1v_1) + k(6u_2v_2) = 2(ku_1)v_1 + 6(ku_2)v_2 = < ku, v >.$$

   Positivity:
   $$< u, u > = 2u_1^2 + 6u_2^2 \geq 0,$$
   and can only equal zero when $u_1 = u_2 = 0$; i.e. $u = 0$.

4. Let p, q, r be polynomials in $\mathcal{P}_2$.

   Symmetry:
   $$< p, q > = p(0)q(0) + p(1/2)q(1/2) + p(1)q(1) = q(0)p(0) + q(1/2)p(1/2) + q(1)p(1) = < q, p >.$$

   Additivity:
   $$< p, q + r > = p(0)(q(0) + r(0)) + p(1/2)(q(1/2) + r(1/2)) + p(1)(q(1) + r(1)) =$$

$$(p(0)q(0) + p(0)r(0)) + (p(1/2)q(1/2) + p(1/2)r(1/2)) + (p(1)q(1) + p(1)r(1)) =$$

$$(p(0)q(0) + p(1/2)q(1/2) + p(1)q(1)) + (p(0)r(0) + p(1/2)r(1/2) + p(1)r(1)) = \langle p, q\rangle + \langle p, r\rangle.$$

Scalar Associativity:

$$k\langle p, q\rangle = k(p(0)q(0)+p(1/2)q(1/2)+p(1)q(1)) = k(p(0)q(0))+k(p(1/2)q(1/2))+k(p(1)q(1)) =$$

$$(kp(0))q(0) + (kp(1/2))q(1/2) + (kp(1))q(1) = \langle kp, q\rangle.$$

Positivity:

$$\langle p, p\rangle = (p(0))^2 + (p(1/2))^2 + (p(1))^2 \geq 0,$$

and can only equal zero when $p(0) = p(1/2) = p(1) = 0$. But the only way an (at most) second-degree polynomial can have three zeros is if the polynomial is constant and eqauls zero for all x; i.e. $p(x) = 0$.

5. The positivity axiom is not satisfied. For if

$$\langle p, p\rangle = (p(0))^2 + (p(1))^2 = 0,$$

then it implies that p has roots at $x = 0$ and $x = 1$. For example, if $p(x) = x^2 - x$, then $\langle p, p\rangle = 0$, but $p \neq 0$.

6. $|A| = \sqrt{4} = 2\sqrt{2}$, $|B| = \sqrt{17}$,

$$\langle A, B\rangle = (1)(2) + (1)(-2) + (1)(3) + (-1)(0) = 3.$$

Therefore,

$$\theta(A, B) = \cos^{-1}(3/(2\sqrt{17}) = 68.7°.$$

7. $|A| = \sqrt{50} = 5\sqrt{2}$, $|B| = \sqrt{14}$,

$$\langle A, B\rangle = (2)(3) + (6)(2) + (1)(1) + (-3)(0) = 19.$$

Therefore,

$$\theta(A, B) = \cos^{-1}(19/(5\sqrt{2}\sqrt{14})) = 44.1°.$$

8. $|p| = \sqrt{17}/4$, $|q| = \sqrt{5}/2$

$$\langle p, q\rangle = (0)(0) + (1/4)(1/2) + (1)(1) = 9/8.$$

Therefore,

$$\theta(p, q) = \cos^{-1}(9/(\sqrt{17}\sqrt{5})) = 12.5°.$$

9. $|p| = \sqrt{156}/4$, $|q| = \sqrt{157}/2$

$$\langle p, q\rangle = (2)(1) + (7/4)(9/4) + (2)(6) = 287/16.$$

Therefore,

$$\theta(p, q) = \cos^{-1}(\frac{(287)(4)(2)}{(16)(\sqrt{156})(\sqrt{157})}) = 23.5°.$$

10.
$$|f|^2 = \int_{-1}^{1} f^2(x)\mathrm{d}x = \int_{-1}^{1} x^2\mathrm{d}x = \frac{x^3}{3}\ |_{-1}^{1} = 2/3.$$

Therefore, $|f| = \sqrt{2/3}$. Similarly,

$$|g|^2 = \int_{-1}^{1} g^2(x)\mathrm{d}x = \int_{-1}^{1} x^4\mathrm{d}x = \frac{x^5}{5}|_{-1}^{1} = 2/5.$$

Therefore, $|g| = \sqrt{2/5}$. Finally,

$$<f,g> = \int_{-1}^{1} f(x)g(x)\mathrm{d}x = \int_{-1}^{1} x^3\mathrm{d}x = \frac{x^4}{4}|_{-1}^{1} = 0$$

Therefore, f and g are orthogonal, and the angle between them is $90°$.

11.
$$|f|^2 = \int_{0}^{2\pi} f^2(x)\mathrm{d}x = \int_{0}^{2\pi} x^2\mathrm{d}x = \frac{x^3}{3}\ |_{0}^{2\pi} = 8\pi^3/3.$$

Therefore, $|f| = \sqrt{8\pi^3/3}$. Similarly,

$$|g|^2 = \int_{0}^{2\pi} g^2(x)\mathrm{d}x = \int_{0}^{2\pi} \cos^2 x\mathrm{d}x =$$

$$\int_{0}^{2\pi} (1+\cos(2x))/2\mathrm{d}x = \pi.$$

Therefore, $|g| = \sqrt{\pi}$. Finally,

$$<f,g> = \int_{0}^{2\pi} f(x)g(x)\mathrm{d}x = \int_{0}^{2\pi} x\cos x\mathrm{d}x =$$

$$x\sin x|_{0}^{2\pi} - \int_{0}^{2\pi} \sin x = \cos x|_{0}^{2\pi} = 0.$$

Therefore, f and g are orthogonal, and the angle between them is $90°$.

12. $|A| = \sqrt{39}$, while $|B| = \sqrt{19}$ Therefore,

$$A/|A| = \begin{pmatrix} 2/\sqrt{39} & -3/\sqrt{39} \\ 1/\sqrt{39} & 5/\sqrt{39} \end{pmatrix} \text{ and } B/|B| = \begin{pmatrix} -1/\sqrt{19} & 3/\sqrt{19} \\ -3/\sqrt{19} & 0 \end{pmatrix}.$$

13.
$$|p| = \sqrt{1^2+5^2+11^2} = \sqrt{147},$$

while

$$|q| = \sqrt{(-4)^2+1^2+6^2} = \sqrt{53}.$$

Therefore, $p/|p| = (4x^2+6x+1)/\sqrt{147}$, while $q/|q| = (10x-4)/\sqrt{53}$.

14. Let $w \in \mathcal{W}$ be arbitrary. Then $w = c_1e_1 + \cdots + c_ne_n$ for some coefficients $c_1, \ldots, c_n$. Then

$$<v, w> = <v, c_1e_1 + \cdots + c_ne_n> =$$

$$c_1<v, e_1> + \cdots + c_n<v, e_n> = 0,$$

since $<v, e_i> = 0$ for all $i = 1, \ldots, n$. Notice that the second-to-last equality follows from repeated uses of the additivity and scalar-associativity axioms for an inner product.

15. Yes, all vectors have unit length and are pairwise orthogonal.

16. No, $<v_2, v_3> = 1/\sqrt{2} \neq 0$.

17. The orthonormal basis is $e_1 = (1/\sqrt{10}, -3\sqrt{10})$, $e_2 = (3/\sqrt{10}, (1/\sqrt{10})$.

18. The orthonormal basis is $e_1 = (1, 0)$, $e_2 = (0, -1)$.

19. The orthonormal basis is $e_1 = (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})$, $e_2 = (-1/\sqrt{2}, 1/\sqrt{2}, 0)$, $e_3 = (1/\sqrt{6}, 1/\sqrt{6}, -2/\sqrt{6})$.

20. The orthonormal basis is $e_1 = (1, 0, 0)$, $e_2 = (0, 7/\sqrt{53}, -2/\sqrt{53})$, $e_3 = (0, 30/\sqrt{11925}, 105/\sqrt{11925})$.

21. $w = 1/4(-5, -1, 5, 9)$ and $w^{\perp} = 1/4(1, 9, 19, -9)$.

22. $c_1 = \langle v, e_1 \rangle = \frac{\sqrt{3}}{2}$, $c_2 = \langle v, e_2 \rangle = \frac{-5}{2\sqrt{2}}$, $c_3 = \langle v, e_3 \rangle = \frac{-15}{2\sqrt{6}}$

23. $c_1 = \langle v, e_1 \rangle = \frac{\sqrt{3}}{2}$, $c_2 = \langle v, e_2 \rangle = \frac{-5}{2\sqrt{2}}$, $c_3 = \langle v, e_3 \rangle = \frac{-15}{2\sqrt{6}}$

24. $c_1 = \langle v, e_1 \rangle = 0.5$, $c_2 = \langle v, e_2 \rangle = \frac{-20}{\sqrt{53}}$, $c_3 = \langle v, e_3 \rangle = \frac{255}{\sqrt{11925}}$