Orthogonal Projections and Orthonormal Bases

Warm-up (a) *Find the angle between* $\begin{bmatrix}1\\2\\1\end{bmatrix}$ *and* $\begin{bmatrix}1\\-1\\1\end{bmatrix}$.

Solution. Let $\vec{v} = \begin{bmatrix}1\\2\\1\end{bmatrix}$ and $\vec{w} = \begin{bmatrix}1\\-1\\1\end{bmatrix}$, and let θ be the angle between $\vec{v}$ and $\vec{w}$. Then, $\vec{v}\cdot\vec{w} = \|\vec{v}\|\|\vec{w}\|\cos\theta$. We calculate that $\vec{v}\cdot\vec{w} = 1\cdot 1 + 2\cdot -1 + 1\cdot 1 = 0$, so $0 = \|\vec{v}\|\|\vec{w}\|\cos\theta$. Since the lengths $\|\vec{v}\|$ and $\|\vec{w}\|$ are both positive, $\cos\theta = 0$, so $\theta = \frac{\pi}{2}$.

(b) *If* $\vec{v} = \begin{bmatrix}1\\0\\1\\-2\end{bmatrix}$, *find* $\|\vec{v}\|$, *the length of* $\vec{v}$.

Solution. By Definition A.6, $\|\vec{v}\| = \sqrt{\vec{v}\cdot\vec{v}} = \sqrt{6}$.

(c) *Find a unit vector in* $\mathbb{R}^3$ *that is perpendicular to both* $\begin{bmatrix}1\\3\\2\end{bmatrix}$ *and* $\begin{bmatrix}0\\2\\0\end{bmatrix}$.

Solution. We are looking for a vector $\vec{v} = \begin{bmatrix}a\\b\\c\end{bmatrix}$ which has the following three properties:

- It is perpendicular to $\begin{bmatrix}1\\3\\2\end{bmatrix}$, so $0 = \begin{bmatrix}a\\b\\c\end{bmatrix}\cdot\begin{bmatrix}1\\3\\2\end{bmatrix} = a + 3b + 2c$.
- It is perpendicular to $\begin{bmatrix}0\\2\\0\end{bmatrix}$, so $0 = \begin{bmatrix}a\\b\\c\end{bmatrix}\cdot\begin{bmatrix}0\\2\\0\end{bmatrix} = 2b$.
- It is a unit vector, so $\vec{v}\cdot\vec{v} = \|\vec{v}\|^2$ is 1. That is, $a^2 + b^2 + c^2 = 1$.

From the second property, we see that $b = 0$. Plugging this into the other 2 equations, we get that $a + 2c = 0$ and $a^2 + c^2 = 1$. The former says that $a = -2c$; plugging this into the latter gives $5c^2 = 1$, so $c = \pm\frac{1}{\sqrt{5}}$. So, there are two possible answers, $\pm\begin{bmatrix}-2/\sqrt{5}\\0\\1/\sqrt{5}\end{bmatrix}$.

Vector Review:

(a) The length of a vector $\vec{v} = \begin{bmatrix}v_1\\v_2\\\vdots\\v_n\end{bmatrix}$ is $|\vec{v}| = \sqrt{\vec{v}^T\vec{v}} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$

(b) If $|\vec{v}| = 1$, then $\vec{v}$ is called a **unit vector**.

(c) Let α be the angle between two vectors $\vec{v}$ and $\vec{w}$.

$$\cos\alpha = \frac{\vec{v}\cdot\vec{w}}{|\vec{v}||\vec{w}|}$$

Note: The number $\cos\alpha$ is called the correlation coefficient. If it is positive, the vectors are positively correlated, if it is negative they are negatively correlated.

A basis $\{\vec{v}_1, \vec{v}_2, ..., \vec{v}_n\}$ is an **orthonormal basis** of V, if all vectors in the basis are perpendicular to each other and have length 1, which means

$$\vec{v}_i \cdot \vec{v}_j = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases}$$

If they are just orthogonal, they form an **orthogonal basis**

1. Let V be a subspace of $\mathbb{R}^n$. For any vector $\vec{x} \in \mathbb{R}^n$, we can write $\vec{x} = \vec{x}^{\|} + \vec{x}^{\perp}$ where $\vec{x}^{\|}$ is in V and $\vec{x}^{\perp}$ is orthogonal to V.[(1)] Then, $\vec{x}^{\|}$ is called the orthogonal projection of $\vec{x}$ onto V and denoted by $\mathrm{proj}_V(\vec{x})$.

Suppose we have an orthonormal basis $(\vec{u}_1, \ldots, \vec{u}_m)$ of V; that is, $(\vec{u}_1, \ldots, \vec{u}_m)$ is a basis of V with the property that $\vec{u}_1, \ldots, \vec{u}_m$ are orthonormal.

(a) *Explain why* $\mathrm{proj}_V(\vec{x})$ *can be written as* $\mathrm{proj}_V(\vec{x}) = c_1\vec{u}_1 + \cdots + c_m\vec{u}_m$ *for some scalars* $c_1, \ldots, c_m$.

Solution. By definition, $\mathrm{proj}_V(\vec{x})$ is a vector in V. On the other hand, $(\vec{u}_1, \ldots, \vec{u}_m)$ is a basis of V, which means that every vector in V can be expressed as a linear combination of the vectors $\vec{u}_1, \ldots, \vec{u}_m$. That is, every vector in V (including the one we're interested in, $\mathrm{proj}_V(\vec{x})$) can be expressed as $c_1\vec{u}_1 + \cdots + c_m\vec{u}_m$ for some scalars $c_1, \ldots, c_m$.

(b) *Since* $\vec{x} = \vec{x}^{\|} + \vec{x}^{\perp}$, *we can use (a) to write*

$$\vec{x} = (c_1\vec{u}_1 + \cdots + c_m\vec{u}_m) + \vec{x}^{\perp}.$$

Express the coefficient c_k *in terms of* $\vec{x}, \vec{u}_1, \ldots, \vec{u}_m$.

Solution. A very useful technique when working with orthonormal (or even just orthogonal) vectors is to dot with one of them. If we dot the equation $\vec{x} = (c_1\vec{u}_1 + \cdots + c_m\vec{u}_m) + \vec{x}^{\perp}$ with $\vec{u}_k$, we get

$$\vec{x}\cdot\vec{u}_k = (c_1\vec{u}_1 + \cdots + c_m\vec{u}_m + \vec{x}^{\perp})\cdot\vec{u}_k$$

Now, $\vec{u}_k$ is orthogonal to almost all of the vectors in the sum $c_1\vec{u}_1 + \cdots + c_m\vec{u}_m + \vec{x}^{\perp}$. First, it's perpendicular to $\vec{x}^{\perp}$ because $\vec{x}^{\perp}$ is perpendicular to all vectors in V, and $\vec{u}_k \in V$. It's also perpendicular to all of the $c_j\vec{u}_j$ except for $c_k\vec{u}_k$. So, our previous expression simplifies to:

$$\vec{x}\cdot\vec{u}_k = c_k(\vec{u}_k\cdot\vec{u}_k)$$

We know that $\vec{u}_k\cdot\vec{u}_k = \|\vec{u}_k\|^2 = 1$ (the fact that $\vec{u}_1, \ldots, \vec{u}_m$ are orthonormal means they all have length 1), so this simplifies even more to just:

$$\vec{x}\cdot\vec{u}_k = c_k$$

[(1)] When we say $\vec{x}^{\perp}$ is orthogonal to V, we mean that $\vec{x}^{\perp}$ is orthogonal to every vector in V.

(c) *Write a formula for* $\text{proj}_V(\vec{x})$ *in terms of* $\vec{x}, \vec{u}_1, \ldots, \vec{u}_m$.

Solution. We just put together what we did in the previous two parts. We said in (a) that

$$\text{proj}_V(\vec{x}) = c_1\vec{u}_1 + \cdots + c_m\vec{u}_m,$$

and we found in (b) that $c_k = \vec{x} \cdot \vec{u}_k$. So,

$$\boxed{\text{proj}_V(\vec{x}) = (\vec{x} \cdot \vec{u}_1)\vec{u}_1 + (\vec{x} \cdot \vec{u}_2)\vec{u}_2 + \cdots + (\vec{x} \cdot \vec{u}_m)\vec{u}_m}$$

(d) *In coming up with this formula for* $\text{proj}_V(\vec{x})$*, where was it important that* $(\vec{u}_1, \ldots, \vec{u}_m)$ *be an* orthonormal *basis of* V*?*

Solution. The fact that $\vec{u}_1, \ldots, \vec{u}_m$ were orthonormal was key in (b); if the vectors had not been orthonormal, we would not have been able to come up with a simple formula for the coefficients $c_1, \ldots, c_m$.

2. Let V be the plane $2x + 2y + z = 0$, $\vec{u}_1 = \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \end{bmatrix}$, and $\vec{u}_2 = \begin{bmatrix} -2/3 \\ 1/3 \\ 2/3 \end{bmatrix}$. Let $\vec{x} = \begin{bmatrix} 1 \\ 4 \\ 8 \end{bmatrix}$.

(a) *Verify that* $(\vec{u}_1, \vec{u}_2)$ *is an orthonormal basis of* V.

Solution. A basis of V consists of any two non-parallel vectors in V, so $\vec{u}_1$ and $\vec{u}_2$ clearly form a basis of V (they are both in V, and they are not parallel). To check that $\vec{u}_1$ and $\vec{u}_2$ are orthonormal, we compute some dot products:

$$\begin{aligned} \vec{u}_1 \cdot \vec{u}_1 &= 1 \\ \vec{u}_1 \cdot \vec{u}_2 &= 0 \\ \vec{u}_2 \cdot \vec{u}_2 &= 1 \end{aligned}$$

So, $\vec{u}_1$ and $\vec{u}_2$ really are orthonormal.

(b) *Find* $\text{proj}_V(\vec{x})$. *(Check that your answer is reasonable by computing the difference* $\vec{x} - \text{proj}_V(\vec{x})$. *What should be true about this vector?)*

Solution. We are given an orthonormal basis $(\vec{u}_1, \vec{u}_2)$ of V. Therefore, by #1(c),

$$\begin{aligned} \text{proj}_V(\vec{x}) &= (\vec{x} \cdot \vec{u}_1)\vec{u}_1 + (\vec{x} \cdot \vec{u}_2)\vec{u}_2 \\ &= 3\vec{u}_1 + 6\vec{u}_2 \\ &= \boxed{\begin{bmatrix} -3 \\ 0 \\ 6 \end{bmatrix}} \end{aligned}$$

Then, $\vec{x} - \text{proj}_V(\vec{x}) = \begin{bmatrix} 4 \\ 4 \\ 2 \end{bmatrix}$; this should be orthogonal to V (it is what we called $\vec{x}^{\perp}$ in #1), and it is!

Let $\{\vec{v}_1, \vec{v}_2, ..., \vec{v}_n\}$ be an orthonormal basis of V , and Q be the matrix containing the basis vectors as column vectors. Then the projection onto the space V is given by the matrix $P = QQ^T$, where Q^T is the transpose matrix.

3. (T/F) *If $\vec{u}_1, \ldots, \vec{u}_m$ are orthonormal vectors in $\mathbb{R}^n$, then must they be linearly independent.*

   Solution. To decide whether $\vec{u}_1, \ldots, \vec{u}_m$ are linearly independent, we should look for linear relations among them; in particular, we are wondering whether there could be nontrivial linear relations among them. Suppose we have a linear relation among them:

   $$c_1\vec{u}_1 + c_2\vec{u}_2 + \cdots + c_m\vec{u}_m = \vec{0}$$

   A particularly useful technique when working with orthonormal vectors is to dot with one of the vectors, so let's dot this equation with $\vec{u}_k$:

   $$c_1(\vec{u}_1 \cdot \vec{u}_k) + c_2(\vec{u}_2 \cdot \vec{u}_k) + \cdots + c_m(\vec{u}_m \cdot \vec{u}_k) = \vec{0} \cdot \vec{u}_k$$

   The right side is clearly just 0. On the left side, $\vec{u}_k$ is orthogonal to all of the other $\vec{u}_i$, so there is only one non-zero term:

   $$c_k(\vec{u}_k \cdot \vec{u}_k) = 0$$

   Since $\vec{u}_1, \ldots, \vec{u}_m$ are orthonormal, $\vec{u}_k$ has length 1, so $\vec{u}_k \cdot \vec{u}_k = 1$:

   $$c_k = 0$$

   What have we just shown? If we have a linear relation $c_1\vec{u}_1 + \cdots + c_m\vec{u}_m$, then we've shown that all of the c_k are 0; in other words, the only linear relation among $\vec{u}_1, \ldots, \vec{u}_m$ is the trivial one; this exactly says that $\vec{u}_1, \ldots, \vec{u}_m$ are linearly independent.

4. Suppose that we want to fit a line to the data points $(-1, 3)$, $(0, 1)$, and $(1, 1)$.

   (a) *Do you expect the slope of the line to be positive, negative, or zero?*

   Solution. Looking at the data points, we see that the best-fit line should have negative slope.

   (b) *Find the best-fit line. (You will do a similar question in PSet12, but there you need to use a different formula)*

   Solution. We will talk about this again later. For now you just need to know how to use the given formula as in HW12 to find this line.

5. In each part, you are given a subspace V of some $\mathbb{R}^n$. Describe $V^\perp$. (we call $V^\perp$ the orthogonal complement of V)

   (a) $y = 3x$ *in* $\mathbb{R}^2$.

   Solution. This is a line in $\mathbb{R}^2$, so the orthogonal complement is the line through the origin perpendicular to $y = 3x$, or $\boxed{\text{the line } y = -\frac{1}{3}x}$.

(b) $y = 3x$ *in* $\mathbb{R}^3$.

Solution. This is a plane in $\mathbb{R}^3$, so the orthogonal complement is the line through the origin perpendicular to this plane. This plane can be expressed as $3x - y + 0z = 0$, so a normal vector for the plane is $\begin{bmatrix} 3 \\ -1 \\ 0 \end{bmatrix}$. Therefore, the orthogonal complement of the given plane is $\boxed{\text{the line span} \begin{bmatrix} 3 \\ -1 \\ 0 \end{bmatrix}}$.

(c) span $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$.

Solution. This is a line in $\mathbb{R}^3$, so the orthogonal complement is the plane normal to this line, which is exactly $\boxed{\text{the plane } x + 2y + 3z = 0}$.

6. Let V be an m-dimensional subspace of $\mathbb{R}^n$. Consider the linear transformation $\operatorname{proj}_V : \mathbb{R}^n \to \mathbb{R}^n$.

(a) *What is* $\operatorname{im} \operatorname{proj}_V$*? What is* $\operatorname{rank} \operatorname{proj}_V$*?*

Solution. The image of proj_V is simply V itself. We know that the rank of a matrix or linear transformation is the same as the dimension of its image. So, $\operatorname{rank} \operatorname{proj}_V = \dim(\operatorname{im} \operatorname{proj}_V) = \dim V = m$.

(b) *What is* $\ker \operatorname{proj}_V$*? What is its dimension?*

Solution. By definition, $\ker \operatorname{proj}_V$ consists of all vectors $\vec{x}$ in $\mathbb{R}^n$ such that $\operatorname{proj}_V(\vec{x}) = \vec{0}$. That is, it consists of all vectors $\vec{x}$ in $\mathbb{R}^n$ whose projection is $\vec{0}$. So, $\ker \operatorname{proj}_V$ consists of all vectors which are perpendicular to V (i.e., perpendicular to all vectors in V); that is, $\boxed{\ker \operatorname{proj}_V = V^\perp}$.

By the rank-nullity theorem, $\dim(\operatorname{im} \operatorname{proj}_V) + \dim(\ker \operatorname{proj}_V) = n$. Therefore, $\dim(\ker \operatorname{proj}_V) = n - \dim(\operatorname{im} \operatorname{proj}_V) = \boxed{n - m}$.

7. See PSet12 #2 and #3 about expectation, variance, covariance, standard deviation, and the correlation coefficient, and about finding the best linear fit for some given data!

- You should understand how we use the dot product to define geometric ideas like the length of a vector and the angle between two vectors in $\mathbb{R}^n$.
- You should understand what it means for a collection of vectors to be *orthonormal*, and you should be able to determine whether given vectors are orthonormal.
- You should understand the *orthogonal complement* of a subspace of $\mathbb{R}^n$ and be able to visualize it in simple cases (in $\mathbb{R}^2$ or $\mathbb{R}^3$).