Consider the following situation. Given a line $L \subset \mathbb{R}^2$, and given any other vector $v \in \mathbb{R}^2$, we can “project” the vector $v$ onto the line $L$ by dropping a perpendicular onto the line. The result is a new vector, which we can denote by $P(v)$, that lies along the line $L$, as in the picture below:

This is the basis of the idea of a projection. We want to generalize this idea and consider projection onto an arbitrary subspace $W$ of a vector space $V$.

In fact, in the above situation we had to drop a perpendicular onto the line $L$, which implicitly relies on an inner product since the notion of being perpendicular, a.k.a. orthogonal, is defined using the inner product. A general vector space does not come equipped with an inner product so in fact “projection onto $W$” is not
well-defined. In order to define a projection in the absence of an inner product, we need to not only specify the subspace $W$ that we are projecting onto, but we also need to specify a complementary subspace $W'$ along which we are projecting.

For the sake of clarity, we split up our discussion of projections into two cases: vector spaces without inner products, and vector spaces equipped with inner products.

Projections in the absence of an inner product

Recall the following definition.

**Definition**: Let $V$ be a vector space and let $W, W' \subset V$ be subspaces. Then we write $V = W \oplus W'$ if $V = W + W'$ and $W \cap W' = \{0\}$. We can describe this situation by saying $V$ is the **direct sum** of $W$ and $W'$. We also say that $W'$ is a **complement** of $W$ in $V$.

Note that a given subspace $W \subset V$ can have many different complements $W'$. Indeed, if $V = \mathbb{R}^2$ and $W \subset V$ is a line, then any other line $W' \subset V$ will be a complement of $W$, satisfying $W \oplus W' = V$.

In an earlier homework problem (2.2.9), we showed that, if $V = W \oplus W'$, then any vector $v \in V$ can be decomposed uniquely as a sum of vectors in $W$ and $W'$, that is, $v = w + w'$ for a unique choice of $w \in W$ and $w' \in W'$. This allows us to define the projection of $V$ onto $W$ along $W'$.

**Definition**: Suppose $V = W \oplus W'$. Then we define the projection of $V$ onto $W$ along $W'$ to be the linear operator $P : V \to V$ given by the formula

$$P(v) = w \text{ where } v = w + w' \text{ with } w \in W, w' \in W'$$

That is, we decompose $v$ as a sum of vectors in $W$ and $W'$ and pick only the component that lies in $W$. This is uniquely defined by the uniqueness of the decomposition.

**Claim**: Suppose $P : V \to V$ is the projection of $V$ onto $W$ along $W'$. Then

a) If $w \in W$, then $P(w) = w$.

b) If $w' \in W'$, then $P(w') = 0$.

c) $W = \text{im}(P)$ and $W' = \ker(P)$.

Proof: a) Indeed, given any $w \in W$ its decomposition into $W$ and $W'$ is simply $w = w + 0$ so we have $P(w) = w$.

b) Similarly, if $w' \in W'$ then its decomposition into $W$ and $W'$ is simply $w' = 0 + w'$ so $P(w') = 0$. 

2
c) By definition, \( \text{im}(P) \subset W \). On the other hand \( W \subset \text{im}(P) \) by part a).

Similarly, \( W' \subset \ker(P) \) by part b). On the other hand, \( \ker(P) \subset W' \) since if \( v \in \ker(P) \), we have \( P(v) = 0 \) so the decomposition of \( v \) into \( W \) and \( W' \) must be of the form \( v = 0 + w' \), so \( v = w' \in W' \).

A consequence of the above claim is that, given a projection \( P : V \to V \), we must have \( V = \ker(P) \oplus \text{im}(P) \).

There is a simple formula describing projection operators.

**Claim:** a linear operator \( T : V \to V \) is a projection if and only if \( T^2 = T \).

**Proof:** We first show that any projection \( P : V \to V \) satisfies the equation \( P^2 = P \). Indeed, say \( P \) is the projection onto \( W = \text{im}(P) \) along \( W' = \ker(P) \). Given any \( v \in V \), we have \( P(v) \in W \). But by part a) of the above claim, \( P(w) = w \) for \( w \in W \) so in particular \( P(P(v)) = P(v) \), so \( P^2(v) = P(V) \) as desired.

On the other hand, suppose \( T : V \to V \) is any operator satisfying \( T^2 = T \). We want to show that \( T \) is the projection onto \( \text{im}(T) \) along \( \ker(T) \). We leave this as an exercise.

Hint: We first need to show that \( V = \text{im}(T) \oplus \ker(T) \). I in fact proved this in the solution to 3.4.10 in HW 5. Look at that proof if you’re interested.

**Example:** Let \( L \subset \mathbb{R}^2 \) be the line given by \( y = x \) and let \( L' \subset \mathbb{R}^2 \) be the line given by \( y = 2x \). We compute the projection \( P \) of \( \mathbb{R}^2 \) onto \( L \) along \( L' \).

The matrix of a projection is always very easy with respect to a certain special basis. Indeed, let \( v_1 \) be a nonzero vector in \( L \) and let \( v_2 \) be a nonzero vector in \( L' \). For example, we can choose \( v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \) and \( v_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \).

Since \( v_1 \in L \) and we are projecting onto \( L \), we have \( P(v_1) = v_1 \). On the other hand, since \( v_2 \in L' \) and we are projecting along \( L' \), we have \( P(v_2) = 0 \). Thus the matrix for \( P \) in the basis \( B = v_1, v_2 \) is \( [P]_B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \).

To figure out the matrix for \( P \) in the standard basis, we just need to do a change of basis. Let \( S = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \). Then \( S \) is the change-of-basis matrix from \( B \)-coordinates to standard coordinates. We can calculate \( S^{-1} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \). Thus
the matrix of $P$ in standard coordinates is
\[ A = S[P]S^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 2 & -1 \end{bmatrix} \]

\[ \square \]

Projections in inner product spaces

In the previous section, we saw that if $W \subset V$, in order to define the projection of $V$ onto $W$ we must specify a complement $W'$ to $W$ along which we project. We also saw that there are many possible choices of a complement. In this section, we see that if our vector space $V$ is equipped with an inner product, then in fact there is a canonical choice of a complement to $W$, called its orthogonal complement. The projection along the orthogonal complement of $W$ will be called the orthogonal projection onto $W$.

**Definition:** Let $V$ be an inner product space and let $W \subset V$. Then we define $W^\perp = \{ v \in V : \langle v, w \rangle = 0 \text{ for all } w \in W \}$. That is, $W^\perp$, called the orthogonal complement of $W$, consists of vectors which are orthogonal to every vector in $W$.

You can check that $W^\perp$ is indeed a subspace.

In order to compute orthogonal projection, we need the concept of an orthonormal basis.

**Definition:** Let $V$ be an inner product space. Then a basis $B$ is orthonormal if $\langle v, v \rangle = 1$ for all $v \in B$ and $\langle v, v' \rangle = 0$ for any two distinct vectors $v, v' \in B$. In other words, a basis is orthonormal if every basis vector has norm (length) 1 and any two distinct basis vectors are orthogonal.

**Example:** 

a) The standard basis $\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ of $\mathbb{R}^2$ is orthonormal.

b) Consider the basis $\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ of $\mathbb{R}^2$. The two vectors are orthogonal to each other, but they do not have norm 1, so they do not form an orthonormal basis.

To get an orthonormal basis, we divide by the norm, getting: 

\[ \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \]

which is an orthonormal basis.
Just as a basis forms a tool for computing things in vector spaces, an orthonormal basis forms a tool for computing things in an inner product space.

(Philosophical remark: just as a basis gives us an isomorphism between a vector space $V$ and $F^n$, an orthonormal basis gives us an isomorphism between $V$ and $F^n$ which preserves inner products, where $F^n$ is equipped with the standard inner product. So using orthonormal bases, we can translate any question about inner products to the concrete setting of the standard inner product in $F^n$).

As an example, we show how an orthogonal basis can be used to compute length.

Claim: Let $V$ be an inner product space over $\mathbb{R}$ or $\mathbb{C}$ with orthonormal basis $B = v_1, \ldots, v_n$. Let $v \in V$ and let $[v]_B = (c_1, \ldots, c_n)$. Then the length of $v$ is given by $||v||^2 = |c_1|^2 + \cdots + |c_n|^2$.

Proof: By definition of coordinates, we can write $v = c_1 v_1 + \cdots + c_n v_n$. Now we have

$||v||^2 = \langle v, v \rangle = \langle c_1 v_1 + \cdots + c_n v_n, c_1 v_1 + \cdots + c_n v_n \rangle = \sum_{i,j} \langle c_i v_i, c_j v_j \rangle = \sum_{i,j} c_i \overline{c_j} \langle v_i, v_j \rangle$

Now, if $i \neq j$, then $\langle v_i, v_j \rangle = 0$ by definition of orthonormal basis so these terms disappear. If $i = j$, then $\langle v_i, v_j \rangle = 1$. So the above expression simplifies to

$||v||^2 = \sum_i c_i \overline{c_i} = \sum_i |c_i|^2$

as desired.

Remark: The above computation reduces to finding the coordinates $[v]_B = (c_1, \ldots, c_n)$ of $v$ in the orthonormal basis. These coordinates in turn can be computed by the formula $c_i = \langle v, v_i \rangle$. That is, the coordinate of $v_i$ in $v$ is precisely the inner product of $v$ with $v_i$.

Orthonormal bases similarly help us compute orthogonal projection.

Claim: Let $V$ be an inner product space and let $W \subset V$ be finite-dimensional. Then

a) $W^\perp$ is indeed a complement of $W$, that is, $V = W \oplus W^\perp$.

b) If $w_1, \ldots, w_n$ is an orthonormal basis of $W$, then the projection $P : V \to V$ onto $W$ along $W^\perp$ is given by the formula

$P(v) = \langle v, w_1 \rangle w_1 + \cdots + \langle v, w_n \rangle w_n$

Proof: We prove a) and b) together.
We first show that \( v - P(v) \in W^\perp \). Indeed, it suffices to show \( v - P(v) \) is orthogonal to each basis element \( w_i \) of \( W \). Since \( w_i \) form an orthonormal basis of \( W \), it follows from the above remark that \( \langle P(v), w_i \rangle \) is the coordinate of \( w_i \) in \( P(v) \), which by definition equals \( \langle v, w_i \rangle \). Using the identity
\[
\langle P(v), w_i \rangle = \langle v, w_i \rangle,
\]
we compute:
\[
\langle (v - P(v)), w_i \rangle = \langle v, w_i \rangle - \langle P(v), w_i \rangle = \langle v, w_i \rangle - \langle v, w_i \rangle = 0
\]
as desired.

Clearly \( P(v) \in W \). It follows that
\[
v = P(v) + (v - P(v))
\]
is a decomposition of \( v \) as a sum of vectors in \( W \) and \( W^\perp \), so indeed \( V = W + W^\perp \).

On the other hand, we show that \( W \cap W^\perp = \{0\} \). Indeed, suppose \( v \in W \cap W^\perp \).

Since \( v \in W^\perp \), \( \langle v, w \rangle = 0 \) for all \( w \in W \). In particular, since \( v \in W \), we have \( \langle v, v \rangle = 0 \). So \( v = 0 \), as desired.

Thus \( V = W \oplus W^\perp \).

Since \( v = P(v) + (v - P(v)) \) is the decomposition of \( v \) into \( W \) and \( W^\perp \), it follows that the projection of \( v \) onto \( W \) along \( W^\perp \) is exactly \( P(v) \).

The transformation \( P \) above, which projects onto \( W \) along the orthogonal complement \( W^\perp \), is called the orthogonal projection onto \( W \).

The above formula for the orthogonal projection \( P \) can be used to explicitly compute orthogonal projection. But first we have to be able to find an orthonormal basis for the given subspace \( W \). This can be done using the process of Gram-Schmidt, which we don’t cover in these notes.

Finally, there is an important inequality concerning orthogonal projections.

**Claim:** Let \( W \) be a finite-dimensional subspace of an inner product space \( V \). Let \( P : V \to V \) be the orthogonal projection of \( V \) onto \( W \). Then, for any \( v \in V \), \( \|P(v)\| \leq \|v\| \), and equality is attained only if \( P(v) \in W \).

**Proof:** Write \( v = P(v) + w' \), where \( w' = v - P(v) \in W^\perp \). In particular, \( w' \) is orthogonal to \( P(v) \), so we compute:
\[
\|v\|^2 = \langle v, v \rangle = \langle P(v) + w', P(v) + w' \rangle
\]
\[
= \langle P(v), P(v) \rangle + \langle P(v), w' \rangle + \langle w', P(v) \rangle + \langle w', w' \rangle = \|P(v)\|^2 + \|w'\|^2
\]
But lengths are non-negative real numbers, so $||w'||^2 \geq 0$ and

$$||v||^2 = ||P(v)||^2 + ||w'||^2 \geq ||P(v)||^2$$

as desired. Furthermore, we see that equality occurs only when $||w'|| = 0$, which means $v = P(v) \in W$. \qed

Note: The above inequality can be rephrased in terms of an orthonormal basis. If $w_1, \ldots, w_n$ is an orthonormal basis of $W$, then $P(v)$ follows the formula $P(v) = \langle v, w_1 \rangle w_1 + \cdots + \langle v, w_n \rangle w_n$. Combining this with the earlier formula for length, we see:

$$||P(v)||^2 = |\langle v, w_1 \rangle|^2 + \cdots + |\langle v, w_n \rangle|^2$$

We obtain Bessel’s inequality:

$$|\langle v, w_1 \rangle|^2 + \cdots + |\langle v, w_n \rangle|^2 \leq ||v||^2$$

whenever $w_1, \ldots, w_n$ are orthonormal vectors (if they were simply orthogonal, we would have to adjust by dividing by length).