Math 217 Worksheet 16: Orthogonal Projections and Orthonormal Bases (§5.1)

Definitions:

Two vectors $\vec{v}$, $\vec{w} \in \mathbb{R}^n$ are said to be **orthogonal** (or perpendicular) if $\vec{v} \cdot \vec{w} = 0$.

The **length** of a vector $\vec{v}$ in $\mathbb{R}^n$ is $\|\vec{v}\| = \sqrt{\vec{v} \cdot \vec{v}}$.

Given any set $S \subseteq \mathbb{R}^n$, the **orthogonal complement** $S^\perp$ of S is the set

$$S^\perp = \{\vec{w} \in \mathbb{R}^n \ : \ \vec{w} \cdot \vec{v} = 0 \text{ for all } \vec{v} \in S\}.$$

Problem 1. Examples of Orthogonal Complements.

(a) What is the orthogonal complement of the line $\text{Span}(\vec{e}_2)$ in $\mathbb{R}^3$?

(b) What is the orthogonal complement of the plane $2x - 3y + z = 0$ in $\mathbb{R}^3$?

Solution: (a) The xz-plane in $\mathbb{R}^3$. (b) The line spanned by $\begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix}$.

Problem 2. Orthogonal complements are subspaces.

(a) Prove that for any $S \subseteq \mathbb{R}^n$, $S^\perp$ is a subspace of $\mathbb{R}^n$.

Solution: There are three things to check:

(a) $\vec{0} \in S^\perp$.
(b) If $\vec{x}, \vec{y} \in S^\perp$, then $\vec{x} + \vec{y} \in S^\perp$.
(c) If $\vec{x} \in S^\perp$ and $c \in \mathbb{R}$, then $c\vec{x} \in S^\perp$.

For this, we take arbitrary $\vec{v} \in S$.

(a) $\vec{0} \cdot v = 0$. This is clear since dotting with 0 always gives 0.
(b) If $\vec{x}, \vec{y} \in S^\perp$, we need $(\vec{x} + \vec{y}) \cdot \vec{v} = 0$. But $(\vec{x} + \vec{y}) \cdot \vec{v} = \vec{x} \cdot \vec{v} + \vec{y} \cdot \vec{v} = 0 + 0 = 0$, since both $\vec{x}$ and $\vec{y}$ are in $S^\perp$.
(c) Take $\vec{x} \in S^\perp$ and scalar c. Check $(c\vec{x}) \cdot \vec{v} = c(\vec{x} \cdot \vec{v}) = c0 = 0$ since $\vec{x} \in S^\perp$.

(b) Let $\vec{v} \in \mathbb{R}^n$, let W be any subspace of $\mathbb{R}^n$, and suppose the subset $\{\vec{w}_1, \ldots, \vec{w}_r\} \subseteq W$ is a spanning set for W. Prove that $\vec{v} \in W^\perp$ if and only if $\vec{v} \cdot \vec{w}_i = 0$ for each $1 \leq i \leq r$.

Solution: If $\vec{v} \in W^{\perp}$, then of course $\vec{v} \cdot \vec{w}_i$ for each i since each $\vec{w}_i$ belongs to W. Conversely, suppose $\vec{v} \cdot \vec{w}_i = 0$ for each i, and let $\vec{w} \in W$. Then since $\{\vec{w}_1, \ldots, \vec{w}_r\}$ spans W, we can choose scalars $c_1, \ldots, c_r$ such that $\vec{w} = c_1\vec{w}_1 + \cdots + c_r\vec{w}_r$. Then

$$\vec{v} \cdot \vec{w} \;=\; \vec{v} \cdot (c_1\vec{w}_1 + \cdots + c_r\vec{w}_r) \;=\; c_1(\vec{v} \cdot \vec{w}_1) + \cdots + c_r(\vec{v} \cdot \vec{w}_r) \;=\; 0 + \cdots + 0 \;=\; 0,$$

which shows $\vec{v} \in W^{\perp}$.

Definition A: A set of vectors $\{\vec{v}_1, \ldots, \vec{v}_r\}$ in $\mathbb{R}^n$ is **orthonormal** if

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

In particular, each $\vec{v}_i$ is a *unit vector* and is orthogonal (perpendicular) to the other vectors in the set.
Proposition: Any orthonormal set of vectors is linearly independent. More generally, any set of non-zero vectors $\{\vec{v}_1, \ldots, \vec{v}_r\}$ such $\vec{v}_i \cdot \vec{v}_j = 0$ for $i \neq j$ is linearly independent.

Problem 3: Orthonormal Coordinates.

(a) Is the standard basis for $\mathbb{R}^n$ orthonormal?

(b) Suppose that $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is an orthonormal set. Let $\vec{x} = c_1\vec{v}_1 + \cdots + c_n\vec{v}_n$. Compute $\vec{x} \cdot \vec{v}_i$.

(c) Prove the Proposition above. [First Line: Suppose $c_1\vec{v}_1 + \cdots + c_r\vec{v}_r = 0$ is a relation on $\{\vec{v}_1, \ldots, \vec{v}_r\}$.]

(d) Let $\mathcal{B} = (\vec{v}_1, \ldots, \vec{v}_n)$ be an orthonormal ordered basis for $\mathbb{R}^n$. For arbitrary $\vec{x} \in \mathbb{R}^n$, prove

$$[\vec{x}]_{\mathcal{B}} = \begin{bmatrix} \vec{x} \cdot \vec{v}_1 \\ \vec{x} \cdot \vec{v}_2 \\ \vdots \\ \vec{x} \cdot \vec{v}_n \end{bmatrix}.$$

Discuss one advantage and one disadvantage of working with orthonormal coordinates.

Solution:

(a) Yes.

(b) Write $\vec{x} = c_1\vec{v}_1 + \cdots + c_n\vec{v}_n$. Dot with $\vec{v}_i$ and use the distributive property of dot product: $\vec{x}\vec{v}_i = c_1\vec{v}_1 \cdot \vec{v}_i + \cdots + c_n\vec{v}_n \cdot \vec{v}_i$. Now use the orthonormality: most of the $\vec{v}_j \cdot \vec{v}_i = 0$ so this reduces to $\vec{x} \cdot \vec{v}_i = c_i$.

(c) Consider an arbitrary relation

$$c_1\vec{v}_1 + \cdots + c_r\vec{v}_r = \vec{0}$$

on the set $\{\vec{v}_1, \ldots, \vec{v}_r\}$. Fix one index i in the range $1 \leq i \leq r$. Dotting both sides of the above equation by $\vec{v}_i$, we have

$$0 \;=\; \vec{v}_i \cdot \vec{0} \;=\; \vec{v}_i \cdot (c_1\vec{v}_1 + \cdots + c_r\vec{v}_r) \;=\; c_1(\vec{v}_i \cdot \vec{v}_1) + \cdots + c_r(\vec{v}_i \cdot \vec{v}_r) \;=\; c_i.$$

Since this works for each i, we see that each c_i is zero, which shows that $\{\vec{v}_1, \ldots, \vec{v}_r\}$ is linearly independent.

(d) Write $\vec{x} = c_1\vec{v}_1 + \cdots + c_n\vec{v}_n$ be an arbitrary $\vec{x}$ expressed in the basis $\mathcal{B}$. We use the previous problem to compute the scalar c_i as $\vec{x}\cdot\vec{v}_i$. This means that the $\mathcal{B}$-coordinate column is $\begin{bmatrix} \vec{x}\cdot\vec{v}_1 \\ \vec{x}\cdot\vec{v}_2 \\ \vdots \\ \vec{x}\cdot\vec{v}_n \end{bmatrix}$. It is advantageous because it is easy to compute coordinates in an orthonormal basis by just using dot product by (b). For an arbitrary basis, we would need to solve a large system of equations to find the coordinates. On the other hand, in a numerical example, sometimes finding an orthonormal basis is a pain; we may have fractions and square roots after applying the Gram-Schmidt process..

Definition B: Let W be a subspace of $\mathbb{R}^n$. The **orthogonal projection onto** W is the linear transformation

$$\mathbb{R}^n \xrightarrow{\text{proj}_W} \mathbb{R}^n \qquad \vec{v} \mapsto (\vec{v}\cdot\vec{u}_1)\,\vec{u}_1 + \cdots + (\vec{v}\cdot\vec{u}_d)\,\vec{u}_d$$

where $\vec{u}_1, \ldots, \vec{u}_d$ is an orthonormal basis for W.

(The definition is independent of the choice of orthonormal basis; see Problem 7.)

Problem 4. Dimension of Orthogonal Complement. Fix a line L through the origin in $\mathbb{R}^2$. Consider the map $\text{proj}_L : \mathbb{R}^2 \to \mathbb{R}^2$ projecting orthogonally onto L.

(a) Discuss how we described this map in Chapter 2 and compare to the Definition above.

(b) Use a *geometric argument* to find the kernel and image of proj_L in terms of L and $L^\perp$.

(c) Now let W be any d-dimensional subspace of $\mathbb{R}^n$, and let $\text{proj}_W : \mathbb{R}^n \to \mathbb{R}^n$ be the orthogonal projection onto W. Explain why the kernel of proj_W is $W^\perp$ and the image of proj_W is W.

(d) Prove Theorem C below. [Hint: Rank-Nullity.]

Theorem C: For any subspace $W \subseteq \mathbb{R}^n$, $\dim W + \dim W^\perp = n$.

Solution:

(a) The map is given by sending $\vec{v}$ to the $(\vec{v}\cdot\vec{u})\,\vec{u}$ where $\vec{u}$ is a unit vector in the direction of L. This is exactly what is given by the Definition above, since $\vec{u}$ is an orthonormal basis for L.

(b) Thinking about the projection geometrically, we see every point is mapped to something in L and the points on L are mapped to themselves. So the image is L. The vectors in the kernel are those that project to the origin: this is exactly $L^\perp$. The lines L and $L^\perp$ are perpendicular, crossing at the origin.

(c) To see the image is W, observe first that im $\text{proj}_W \subseteq W$, since by definition, for each $\vec{x} \in \mathbb{R}^n$, $\text{proj}_W(\vec{x})$ is a linear combination of vectors in W, namely $\sum_i (\vec{w}\cdot\vec{u}_i)\vec{u}_i$, where $\{\vec{u}_1, \ldots, \vec{u}_d\}$ is an orthonormal basis for W. But also $W \subseteq$ im proj_W, since given $w \in W$, we have $\text{proj}_W(w) = w$. The kernel is $W^\perp$ since by definition, an element $\vec{v} \in W^\perp$ if and only if $\vec{v}\cdot\vec{u}_i$ for a spanning set (eg, basis) of W.

(d) By rank nullity, since the source of proj_W has dimension n and the image W has dimension d, we know that the kernel $W^\perp$ is dimension $n-d$.

Problem 5. Orthogonal Decomposition with respect to W. Let $W \subseteq \mathbb{R}^n$ be any subspace.

(a) Prove that $W \cap W^\perp = \{\vec{0}\}$. [HINT: Recall $\vec{v}\cdot\vec{v} \geq 0$. When is it zero?]

(b) Prove that each $\vec{v}$ in $\mathbb{R}^n$ decomposes *uniquely* as $\vec{v} = \vec{v}^{||} + \vec{v}^\perp$ where $\vec{v}^{||} \in W$ and $\vec{v}^\perp \in W^\perp$. [HINT: For existence: let $\vec{v}^{||}$ be the projection onto W and let $\vec{v}^\perp$ be $\vec{v} - \vec{v}^{||}$. For uniqueness: Use (a).]

Solution:

(a) Say $\vec{w} \in W \cap W^\perp$. Then $\vec{w}\cdot\vec{w} = 0$, so $\vec{w} = 0$ (by properties of dot product).

(b) Let $\vec{v}^{||} = \text{proj}_W(\vec{v})$ and let $\vec{v}^\perp = \vec{v} - \vec{v}^{||}$. Applying proj_W, we have $\text{proj}_W(\vec{v} - \vec{v}^{||}) = \text{proj}_W(\vec{v}) - \text{proj}_W(\vec{v}^{||}) = \vec{v}^{||} - \vec{v}^{||} = 0$. This says that $\vec{v}^\perp \in \ker \text{proj}_W = W^\perp$. Note that $\vec{v} = \vec{v}^{||} + \vec{v}^\perp$ where $\vec{v}^{||} \in W$ and $\vec{v}^\perp \in W^\perp$, so we have establishhed that such a decomposition *exists*. To check its uniqueness, say we can also write $\vec{v} = \vec{w}_1 + \vec{w}_2$ where $\vec{w}_1 \in W$ and $\vec{w}_2 \in W^\perp$. Then we have $\vec{w}_1 - \vec{v}^{||} = \vec{v}^\perp - \vec{w}_2 \in W \cap W^\perp$. By (a), this is zero, which means $\vec{w}_1 = \vec{v}^{||}$ and $\vec{w}_2 = \vec{v}^\perp$.

Problem 6. The standard Matrix of Orthogonal Projection. Fix a subspace W of $\mathbb{R}^n$, with orthonormal basis $(\vec{u}_1, \dots, \vec{u}_d)$. Let A be the $n \times d$ matrix whose columns are $\vec{u}_1, \dots, \vec{u}_d$. Prove that the standard matrix for the orthogonal projection onto W is $AA^\top$.

[HINT: Compute and compare the i-th column of each. Observe that the i-th column of $A^\top$ is $\begin{bmatrix} \vec{e}_i \cdot \vec{u}_1 \\ \vdots \\ \vec{e}_i \cdot \vec{u}_d \end{bmatrix}$.]

Solution: To find the standard matrix of proj_W, use the Key Theorem to find each column. The i-th column, for each $1 \leq i \leq n$, is

$$\text{proj}_W(\vec{e}_i) \;=\; \sum_{j=1}^{d} (\vec{e}_i \cdot \vec{u}_j)\vec{u}_j.$$

We compare this to the i-th column of $AA^\top$. By definition of matrix multiplication, the i-th column of $AA^\top$ is A times the i-th column of $A^\top$, or equivalently, A times the transpose of the i-th row of A. That is, the i-th column of $AA^\top$ is

$$A \begin{bmatrix} \vec{e}_i \cdot \vec{u}_1 \\ \vdots \\ \vec{e}_i \cdot \vec{u}_d \end{bmatrix} = \begin{bmatrix} \vec{u}_1 & \dots & \vec{u}_d \end{bmatrix} \begin{bmatrix} \vec{e}_i \cdot \vec{u}_1 \\ \vdots \\ \vec{e}_i \cdot \vec{u}_d \end{bmatrix} = \sum_{j=1}^{d} (\vec{e}_i \cdot \vec{u}_j)\vec{u}_j.$$

Thus $AA^\top$ and the standard matrix of proj_W are the same, column by column, and $AA^\top$ must be the standard matrix of proj_W.

Problem 7. Well-defined-ness of orthogonal projection. Let $W \subseteq \mathbb{R}^n$ be any subspace.

*This will be easier after you have more practice with dot products. Try again after Worksheet 18.

(a) Fix an orthonormal basis $(\vec{u}_1, \ldots, \vec{u}_d)$ for W. Verify that the mapping proj_W as defined in Definition B is *linear transformation.*

(b) In Problem 5, you showed that $\mathbf{proj}_W(\vec{x}) = x^{||}$**.** Explain why this implies that the formula for proj_W in Definition B *does not depend on the choice of orthonormal basis.* That is, we get the same value for proj_W using another orthonormal basis $(\vec{w}_1, \ldots, \vec{w}_d)$ instead of $(\vec{u}_1, \ldots, \vec{u}_d)$. [Hint: Look carefully at what you proved in Problem 5(b).]

Solution:

(a) First, take arbitrary $\vec{x}, \vec{y} \in \mathbb{R}^n$. We need to show that $\mathrm{proj}_W(\vec{x}+\vec{y}) = \mathrm{proj}_W(\vec{x}) + \mathrm{proj}(\vec{y})$. This follows from properties of dot product:

$$\mathrm{proj}_W(\vec{x}+\vec{y}) = \sum_{i=1}^{d}((\vec{x}+\vec{y})\cdot\vec{u}_i)\vec{u}_i = \sum_{i=1}^{d}((\vec{x}\cdot\vec{u}_i)u_i + (\vec{y}\cdot\vec{u}_i)u_i) = \sum_{i=1}^{d}\vec{x}\cdot\vec{u}_i + \sum_{i=1}^{d}\vec{y}\cdot\vec{u}_i,$$

which is $\mathrm{proj}_W(\vec{x}) + \mathrm{proj}(\vec{y})$. Next we show that for arbitrary $\vec{x} \in \mathbb{R}^n$ and scalar $c \in \mathbb{R}$, $\mathrm{proj}_W(c\vec{x}) = c\,\mathrm{proj}_W(\vec{x})$. This again follows from properties of dot product:

$$\mathrm{proj}_W(c\vec{x}) = \sum_{i=1}^{d} c\vec{x}\cdot\vec{u}_i = c\sum_{i=1}^{d}\vec{x}\cdot\vec{u}_i = \mathrm{cproj}_W(\vec{x}).$$

Since proj_W respects addition and scalar multiplication, it is a linear transformation.

(b) Let $(\vec{w}_1, \ldots, \vec{w}_d)$ be another orthonormal basis for W. We need to check that for each fixed $\vec{x} \in \mathbb{R}^n$,

$$\sum_{i=1}^{d}(\vec{x}\cdot\vec{u}_i)\vec{u}_i = \sum_{i=1}^{d}(\vec{x}\cdot\vec{w}_i)\vec{w}_i.$$

Let $\vec{y} = \sum_{i=1}^{d}(\vec{x}\cdot\vec{u}_i)\vec{u}_i$, and let $\vec{z} = \sum_{i=1}^{d}(\vec{x}\cdot\vec{w}_i)\vec{w}_i$. We want to show $\vec{y} = \vec{z}$. Both are vectors in W, so can be written as linear combinations of $\{\vec{w}_1, \ldots, \vec{w}_d\}$. We already know that $\vec{z} = \sum_{i=1}^{d}(\vec{x}\cdot\vec{w}_i)\vec{w}_i$, and by Problem 3b, since $(\vec{w}_1, \ldots, \vec{w}_d)$ is orthonormal, we find that

$$\vec{y} = \sum_{i=1}^{d}(\vec{y}\cdot\vec{w}_i)\vec{w}_i.$$

To show that $\vec{y} = \vec{z}$, we can show that their difference

$$\vec{z} - \vec{y} = \sum_{i=1}^{d}(\vec{x}\cdot\vec{w}_i - \vec{y}\cdot\vec{w}_i)\vec{w}_i \tag{1}$$

is zero. The difference $\vec{z} - \vec{y}$ is clearly in W, so it suffices, by Problem 5(a), to show $\vec{z} - \vec{x} \in W^\perp$. For this, we need that for each $i = 1, \ldots, d$,

$$(\vec{z} - \vec{y})\cdot\vec{w}_i = 0.$$

From (1), it suffices if $(\vec{x}\cdot\vec{w}_i - \vec{y}\cdot\vec{w}_i) = 0$, or equivalently, if $(\vec{x} - \vec{y})\cdot\vec{w}_i = 0$ for $i = 1, \ldots d$. This is true by Problem 5b: computing the projection using the basis $(\vec{w}_1, \ldots, \vec{w}_d)$, we have $\vec{x}^{||} = \vec{y}$, so $\vec{x} - \vec{y} = \vec{x} - \vec{x}^{||} = \vec{x}^\perp \in W^\perp$. So $\vec{z} - \vec{y} \in W^\perp \cap W = \{\vec{0}\}$. QED.

***Problem 8.**[†] Show that every subspace of $\mathbb{R}^n$ has an orthonormal basis. [HINT: Induce on $\dim V$. For the inductive step: if V is a $(k+1)$-dimensional subspace of $\mathbb{R}^n$, let $\vec{v}$ be some fixed nonzero vector in V and consider the kernel of the linear transformation $T : V \to \mathbb{R}$ defined by $T(\vec{x}) = \vec{v} \cdot \vec{x}$.]

Solution: Let V be an arbitrary subspace of $\mathbb{R}^n$.

Base case: If V has dimension one, take any non-zero vector $\vec{v} \in V$. Normalize $\vec{v}$ to get $\vec{u} = \frac{\vec{v}}{\|\vec{v}\|}$. Then $\vec{u}$ is a non-zero vector in V, so must span V, and $\{\vec{u}\}$ is a basis.

Inductive assumption: Every subspace of $\mathbb{R}^n$ of dimension $k \geq 1$ has an orthonormal basis.

For the inductive step, suppose V has dimension $k+1 > 1$. Fix any nonzero vector $\vec{v}$ in V, and define the linear transformation $T : V \to \mathbb{R}$ by $T(\vec{x}) = \vec{v} \cdot \vec{x}$. Then $\dim(\mathrm{im}(T)) = 1$, so by Rank-Nullity $\ker(T)$ is an k dimensional subspace of V. Using the inductive hypothesis, we can choose $(\vec{u}_1, \ldots, \vec{u}_k)$ an orthonormal basis of $\ker(T)$. Then

$$\left(\vec{u}_1, \ldots, \vec{u}_k, \frac{\vec{v}}{\|\vec{v}\|}\right)$$

is an orthonormal basis of V, completing the induction.

[†] We will give a different proof on the next worksheet.