# Notes for MAT1341A Fall 2023 Part IX

### Chapter 19 - Orthogonality

***Recall.*** If $\mathbf{u} = (u_1, \ldots, u_n), \mathbf{v} = (v_1, \ldots, v_n) \in \mathbb{R}^n$, we defined the dot product

$$\mathbf{u} \cdot \mathbf{v} = \sum_{i=1}^{n} u_i v_i.$$

We say that $\mathbf{u}$ and $\mathbf{v}$ are *orthogonal* if $\mathbf{u} \cdot \mathbf{v} = 0$, we write $\mathbf{u} \perp \mathbf{v}$.

**Definition (19.2.1).** A set of vector $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\} \subseteq \mathbb{R}^n$ is called *orthogonal* if

$$\mathbf{v}_i \cdot \mathbf{v}_j = 0$$

for all $1 \leq i \leq j \leq m$, and $\mathbf{v}_i \neq 0$ for all $1 \leq i \leq m$. That is, every pair of vectors is orthogonal and no vector is zero.

[*E.g.*] Is $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\} \subseteq \mathbb{R}^n$ an orthogonal set?

[*E.g.*] Is $\{(1,0,0), (0,1,0), (0,1,1)\}$ an orthogonal set?

**Definition.** A set of vector $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\} \subseteq \mathbb{R}^n$ is called *orthonormal* if

$$||\mathbf{v}_i|| = 1,\ \forall\ i \ \text{ and }\ \mathbf{v}_i \perp \mathbf{v}_j,\ \forall\ i \neq j.$$

If $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ is orthogonal, then $\{\frac{\mathbf{v}_1}{||\mathbf{v}_1||}, \ldots, \frac{\mathbf{v}_m}{||\mathbf{v}_m||}\}$ is orthonormal.
[*E.g.*] $\{(1,1),(1,-1)\}$ is orthogonal, but not orthonormal.

**Theorem (19.2.4).** Any orthogonal set of vectors is linearly independent.

**Definition.** Let $W$ be a subspace of $\mathbb{R}^n$. We say that a basis is orthogonal (respectively orthonormal) if it consists of a set of orthogonal (respectively orthonormal) basis.

**Theorem (19.2.5).** Suppose $\{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ is an orthogonal basis for a subspace $W$ of $\mathbb{R}^n$. Then any vector $\mathbf{w} \in W$ can be written as

$$\begin{aligned}\mathbf{w} &= \left(\frac{\mathbf{w} \cdot \mathbf{w}_1}{||\mathbf{w}_1||^2}\right)\mathbf{w}_1 + \left(\frac{\mathbf{w} \cdot \mathbf{w}_2}{||\mathbf{w}_2||^2}\right)\mathbf{w}_2 + \ldots + \left(\frac{\mathbf{w} \cdot \mathbf{w}_m}{||\mathbf{w}_m||^2}\right)\mathbf{w}_m \\ &= \mathrm{proj}_{\mathbf{w}_1}(\mathbf{w}) + \mathrm{proj}_{\mathbf{w}_2}(\mathbf{w}) + \ldots + \mathrm{proj}_{\mathbf{w}_m}(\mathbf{w})\end{aligned}$$

We have $\{(1,2,1),(1,0,-1),(1,-1,1)\}$ is an orthogonal set since

$$(1,2,1)\cdot(1,0,-1)=(1,2,1)\cdot(1,-1,1)=(1,0,-1)\cdot(1,-1,1)=0$$

This is an orthogonal basis of $\mathbb{R}^3$, since an orthogonal set is LI, and $\dim(\mathbb{R}^3)=3$.

Given any element of $\mathbb{R}^3$, we can express it as a linear combination of these 3 elements.

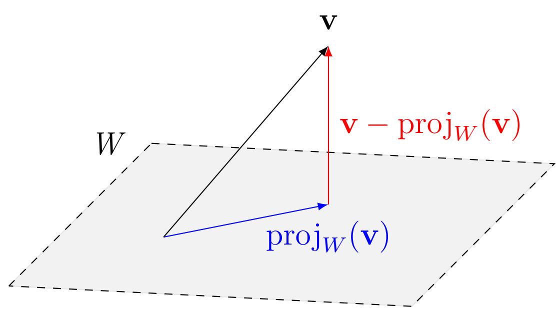
$$\begin{bmatrix}1\\6\\7\end{bmatrix}=\frac{10}{3}\begin{bmatrix}1\\2\\1\end{bmatrix}-3\begin{bmatrix}1\\0\\-1\end{bmatrix}+\frac{2}{3}\begin{bmatrix}1\\-1\\1\end{bmatrix}$$

### Orthogonal projection

Let $W \subseteq \mathbb{R}^n$ a subspace. The orthogonal projection of $\mathbf{v} \in \mathbb{R}^n$ onto $W$, denoted by $\mathrm{proj}_W(\mathbf{v})$, is the vector in $\mathbb{R}^n$ such that

(1) $\mathrm{proj}_W(\mathbf{v}) \in W$

(2) $\mathbf{v} - \mathrm{proj}_W(\mathbf{v}) \perp W$




If $\{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ is an orthogonal basis of $W$, then

$$\mathrm{proj}_W(\mathbf{v}) = \mathrm{proj}_{W_1}(\mathbf{v}) + \ldots + \mathrm{proj}_{W_m}(\mathbf{v})$$

Warning, the basis has to be orthogonal ! Otherwise, the formula does not work.

[*E.g.*] Consider $W = \mathrm{span}\{(0,1,-2,1),(0,0,1,2),(0,-5,-2,1)\}$. Check this is an orthogonal basis. Find $\mathrm{proj}_W(\mathbf{v})$, where $\mathbf{v} = (1,1,1,1)$.

**Theorem (19.3.3 - The Best Approximation Theorem).** Let $W$ be a subspace of $\mathbb{R}^n$ and let $\mathbf{v} \in \mathbb{R}^n$. Then, $\text{proj}_W(\mathbf{v})$ is the best approximation to $\mathbf{v}$ by vectors of $W$, meaning that it's the vector in $W$ whose distance to $\mathbf{v}$ is the smallest.

This turns out that given any subspace of $\mathbb{R}^n$, we may always find an orthogonal basis. In fact, given any basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$, we can convert it to an orthogonal one by the **Gram-Schmidt algorithm**:

- $\mathbf{w}_1 = \mathbf{u}_1$
- $\mathbf{w}_2 = \mathbf{u}_2 - \text{proj}_{\mathbf{w}_1}(\mathbf{u}_2)$
- $\mathbf{w}_3 = \mathbf{u}_3 - \text{proj}_{\mathbf{w}_1}(\mathbf{u}_3) - \text{proj}_{\mathbf{w}_2}(\mathbf{u}_3)$

  $\vdots$
- $\mathbf{w}_m = \mathbf{u}_m - \text{proj}_{\mathbf{w}_1}(\mathbf{u}_m) - \text{proj}_{\mathbf{w}_2}(\mathbf{u}_m) - \ldots - \text{proj}_{\mathbf{w}_{m-1}}(\mathbf{u}_m)$

[*E.g.*] Perform the Gram-Schmidt algorithm on the set

$$\{(1,1,1,1), (6,0,0,2), (-1,-1,2,4)\}.$$

[*E.g.*] Let $W = \mathrm{Null}(A)$, where $A = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 2 & 1 & 0 & 1 \end{bmatrix}$.

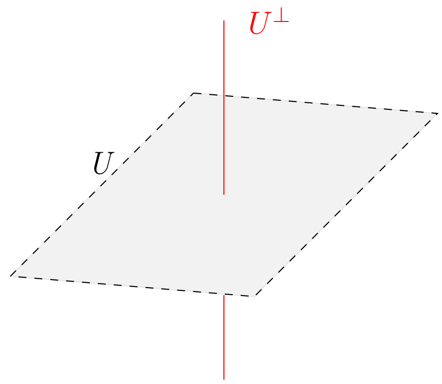
Find $\mathrm{proj}_W(\mathbf{u})$, where $\mathbf{u} = \begin{bmatrix} 0 \\ -1 \\ 1 \\ 1 \end{bmatrix}$.

## Chapter 20 - Orthogonal Complements

**Definition (20.1.1).** Let $U$ be a subspace of $\mathbb{R}^n$. The *orthogonal complement* of $U$ is the set, denoted $U^\perp$ and defined by

$$U^\perp = \{\mathbf{v} \in \mathbb{R}^n \mid \mathbf{u} \cdot \mathbf{v} = 0 \ \forall \mathbf{u} \in U\}$$

If $U$ is a plane in $\mathbb{R}^3$, then $U^\perp$ gives the normal vectors of the plane.



**Theorem (20.1.5 - Properties of the orthogonal complement).**
Let $U$ be a subspace of $R^n$. Then:

(1) $U^{\perp}$ is a subspace of $\mathbb{R}^n$

(2) $(U^{\perp})^{\perp} = U$

(3) $\dim(U) + \dim(U^{\perp}) = n$

[*E.g.*] Consider $W = \{(x, y, z, w) \in \mathbb{R}^4 \mid x - y - w = 0\}$. Find $W^\perp$.

To find $U^\perp$ is the same as finding $\mathrm{Null}(A^\top)$, where the columns of $A$ are basis elements of $U$. In other words, given $U = \mathrm{span}\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$, we form a matrix whose rows are $\mathbf{u}_1, \ldots, \mathbf{u}_m$, then the null space of this matrix is $U^\perp$.

**Orthogonal Projection - an encore**

Suppose we want to find $\text{proj}_W(\mathbf{v})$ where $W = \text{Col}(A)$, $A = \begin{bmatrix}\mathbf{v}_1 & \dots & \mathbf{v}_m\end{bmatrix}$.

$$\text{proj}_W(\mathbf{v}) = \sum_{i=1}^{m} c_i \mathbf{v}_i = A \begin{bmatrix} c_1 \\ \vdots \\ c_m \end{bmatrix}.$$

We can find $c_i$ by solving $A^\top A\mathbf{x} = A^\top \mathbf{v}$. This allows us to find $\text{proj}_W(\mathbf{v})$ without using the projection formula.

[*E.g.*] Find $\text{proj}_W(\mathbf{u})$, where $W = \text{span}\left\{ \begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \end{bmatrix} \right\}$ and $\mathbf{u} = \begin{bmatrix} 2 \\ 1 \\ 0 \\ -1 \end{bmatrix}$.

**Application - Least Squares Method**

Suppose you collect the following data

| $x$ | $y$ |
|---|---|
| 2 | 1 |
| 5 | 2 |
| 7 | 3 |
| 8 | 3 |

These data points don't exactly lie on a parabola, but you think that's experimental error; what is the best-fitting quadratic function through these points?