LECTURE 16

Orthogonality

One of the most useful properties of the standard basis $[e_1, \ldots, e_n]$ of $\mathbb{R}^n$ is the fact that

$$e_i \cdot e_j = \delta_{ij} \equiv \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

(1)

This property for example allows us to easily determine the component of a vector $v$ along the $i^{th}$ basis vector $e_i$ be simply computing its inner product with $e_i$:

$$v = v_1e_1 + \cdots + v_ne_n$$

$$
\implies e_i \cdot v = e_i \cdot (v_1e_1 + \cdots + v_ne_n) \\
= v_1e_i \cdot e_1 + v_2e_i \cdot e_2 + \cdots + v_ie_i \cdot e_i + \cdots + v_ne_i \cdot e_n \\
= 0 + 0 + \cdots + 0 + v_i + 0 + \cdots + 0 \\
= v_i
$$

Of course, this is clear already once we write $v$ and $e_i$ in component form

$$v = [v_1, v_2, \ldots, v_i, \ldots, v_n]$$
$$e_i = [0, \ldots, 0, 1, 0, \ldots, 0]$$

$$\implies e_i \cdot v = v_i$$

However, it is not true for a more general basis. Recall that for a general basis $B = \{b_1, \ldots, b_n\}$, in order to find the constants $c_1, \ldots, c_n$ such that

$$v = c_1b_1 + \cdots + c_nb_n$$

you have to solve the linear system

$$\begin{bmatrix} b_1 & \cdots & b_n \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$$

which is a much harder task.

On the other hand, we have lots and lots of choices of bases for $\mathbb{R}^n$ or for any subspace $W$ of $\mathbb{R}^n$. What we shall be developing in this lecture is a way to contruct bases $B = \{b_1, \ldots, b_n\}$ that enjoy orthogonality properties just like (1)

$$b_i \cdot b_j = \delta_{ij} \equiv \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

For such orthonormal bases, we will be able to rapidly determine the coefficients $c_i$ such that

$$v = v = c_1b_1 + \cdots + c_nb_n$$

by simply computing inner products

$$c_i = b_i \cdot v$$
1. Projections onto Vectors

Recall that the inner product $\mathbf{a} \cdot \mathbf{b}$ of two vectors in $\mathbb{R}^n$ has a very concrete geometric interpretation

$$\mathbf{a} \cdot \mathbf{b} = \|\mathbf{a}\| \|\mathbf{b}\| \cos \theta_{ab}$$

where

$$\|\mathbf{a}\| = \sqrt{\mathbf{a} \cdot \mathbf{a}} \equiv \text{the length of } \mathbf{a}$$

$$\|\mathbf{b}\| = \sqrt{\mathbf{b} \cdot \mathbf{b}} \equiv \text{the length of } \mathbf{b}$$

$$\theta_{ab} = \text{the angle between } \mathbf{a} \text{ and } \mathbf{b} \text{ in the plane spanned by } \mathbf{a} \text{ and } \mathbf{b}$$

Let’s look more closely at the actual geometric situation in the 2-dimensional plane spanned by $\mathbf{a}$ and $\mathbf{b}$.

![Diagram](image)

We see from the diagram above that

$$\|\mathbf{a}\| \cos (\theta_{ab})$$

is the component of the vector $\mathbf{a}$ that runs in the direction of $\mathbf{b}$. We call this the orthogonal projection of $\mathbf{a}$ on $\mathbf{b}$, because if we had a flashlight oriented perpendicularly to the vector $\mathbf{b}$, the “shadow” of the vector $\mathbf{a}$ along $\mathbf{b}$ would be precisely the segment shown above. Since

$$\mathbf{a} \cdot \mathbf{b} = \|\mathbf{a}\| \|\mathbf{b}\| \cos \theta_{ab}$$

we can have the following formula

$$\text{the length of the projection of } \mathbf{a} \text{ along the direction of } \mathbf{b} = \frac{\mathbf{a} \cdot \mathbf{b}}{||\mathbf{b}||}$$

In what follows, however, it is useful to think of this projection not as a length but as the vector that runs in the same direction as $\mathbf{b}$ with length $\frac{\mathbf{a} \cdot \mathbf{b}}{||\mathbf{b}||}$. Now the unit vector in the direction of $\mathbf{b}$ is

$$\frac{\mathbf{b}}{||\mathbf{b}||}$$

so if we multiply this unit vector by the length $\frac{\mathbf{a} \cdot \mathbf{b}}{||\mathbf{b}||}$, we get the vector we want, namely

$$\frac{\mathbf{a} \cdot \mathbf{b}}{||\mathbf{b}||} \frac{\mathbf{b}}{||\mathbf{b}||} = \frac{\mathbf{a} \cdot \mathbf{b}}{\mathbf{b} \cdot \mathbf{b}}$$

**Definition 16.1.** Let $\mathbf{a}$ and $\mathbf{b}$ be two vectors in $\mathbb{R}^n$. Then the projection of $\mathbf{a}$ along the direction of $\mathbf{b}$ is the vector

$$P_{a,b} = \frac{\mathbf{a} \cdot \mathbf{b}}{\mathbf{b} \cdot \mathbf{b}}$$
2. Projections onto Subspaces

Let me now pose a problem that generalizes the construct presented in the last section.

**Problem 16.1.** Given a vector \( \mathbf{v} \in \mathbb{R}^n \) and a subspace \( W \) of \( \mathbb{R}^n \). What component of \( \mathbf{v} \) lies along the directions in \( W \)?

We will in fact show that there are unique vectors \( \mathbf{v}_\perp \) and \( \mathbf{v}_W \) such that

- \( \mathbf{v}_W \in W \)
- \( \mathbf{v}_\perp \) is perpendicular to every vector in \( W \)
- \( \mathbf{v} = \mathbf{v}_W + \mathbf{v}_\perp \)

We will call \( \mathbf{v}_W \) the **orthogonal projection** of \( \mathbf{v} \) **onto** \( W \). It will be exactly the component of \( \mathbf{v} \) that lies in the subspace \( W \).

Let us now suppose that \( W \) is in fact a \( k \)-dimensional subspace with basis \( B_W = \{ \mathbf{b}_1, \ldots, \mathbf{b}_k \} \). The first thing we shall do is construct a subspace \( W^\perp \) of \( \mathbb{R}^n \) that is perpendicular to every vector in \( W \). That is to say, a subspace \( W^\perp \subset \mathbb{R}^n \) such that

\[
\mathbf{v} \in W^\perp \implies \mathbf{v} \cdot \mathbf{w} = 0 \quad \text{for every vector } \mathbf{w} \in W
\]

Since every vector in \( W \) can be written

\[
\mathbf{w} = w_1 \mathbf{b}_1 + w_2 \mathbf{b}_2 + \cdots + w_k \mathbf{b}_k
\]

an easy way to impose the condition \( \mathbf{v} \cdot \mathbf{w} = 0 \) for all vectors \( \mathbf{w} \in W \), would be to demand

\[
\mathbf{v} \cdot \mathbf{b}_i = 0 \quad \text{for } i = 1, \ldots, k
\]

These \( k \) conditions on \( \mathbf{v} \) can then be expressed as a matrix equation

\[
\begin{bmatrix}
\mathbf{b}_1 \cdot \mathbf{v} \\
\vdots \\
\mathbf{b}_k \cdot \mathbf{v}
\end{bmatrix} =
\begin{bmatrix}
\mathbf{b}_1 \\
\vdots \\
\mathbf{b}_k
\end{bmatrix}
\begin{bmatrix}
0 \\
0 \\
0
\end{bmatrix}
\]

In other words, the vector \( \mathbf{v} \) will have to lie in the null space of the \( k \times n \) matrix formed by using the \((n\text{-dimensional})\) basis vectors \( \mathbf{b}_i \) as rows. Set

\[
W^\perp = \operatorname{Null} \operatorname{Sp}
\begin{bmatrix}
\mathbf{b}_1 \\
\vdots \\
\mathbf{b}_k
\end{bmatrix}
\]

Then, we have set things up so that

\[
\mathbf{v} \in W^\perp \iff \mathbf{v} \cdot \mathbf{w} = 0 \quad \text{for all } \mathbf{w} \in W
\]

The space \( W^\perp \) is called the **orthogonal complement** to \( W \) in \( \mathbb{R}^n \).

Next, note that since the vectors \( \mathbf{b}_1, \ldots, \mathbf{b}_k \) form a basis, they must be linearly independent. Therefore the matrix

\[
A_{W,B} =
\begin{bmatrix}
\mathbf{b}_1 \\
\vdots \\
\mathbf{b}_k
\end{bmatrix}
\]

has \( k \) linearly independent row vectors and so has rank \( k \). But then since

\[
n = \# \text{ columns} = \text{rank} (A_{W,B}) + \dim (\text{Null} \operatorname{Sp}(A_{W,B}))
\]

we have

\[
\dim W^\perp = n - k
\]
So we can find a basis $B_{W^\perp} = \{v_1, \ldots, v_{n-k}\}$ for $W^\perp$. Let’s write this change notation slightly are write $\{b_{k+1}, \ldots, b_n\}$ for the $n-k$ basis vector $\{v_1, \ldots, v_{n-k}\}$.

**Lemma 16.2.** The set $\{b_1, \ldots, b_k, b_{k+1}, \ldots, b_n\}$ where $\{b_1, \ldots, b_k\}$ is our given basis for $W$ and $\{b_{k+1}, \ldots, b_n\}$ is a basis for the null space of $A_{W}B$, is a basis for $\mathbb{R}^n$.

**Proof.** Suppose

$$c_1b_1 + \cdots + c_kb_k + c_{k+1}b_{k+1} + \cdots + c_nb_n = 0$$

with not all coefficients $c_i = 0$. Then we’d have

$$c_1b_1 + \cdots + c_kb_k = -c_{k+1}b_{k+1} - \cdots - c_nb_n$$

Set

$$v_1 = c_1b_1 + \cdots + c_kb_k \in W$$
$$v_2 = c_{k+1}b_{k+1} + \cdots + c_nb_n \in W^\perp$$

so that (4) becomes

$$v_1 = -v_2$$

Since the basis vectors set $\{b_1, \ldots, b_k\}$ and $\{b_{k+1}, \ldots, b_n\}$ are linearly independent, neither $v_1$ nor $v_2$ can be 0 unless all the coefficients $c_1, \ldots, c_n$ are zero, which is a situation that we have excluded from the start. But then if $v_1 \neq 0$

$$0 \neq \|v_1\|^2 = v_1 \cdot v_1$$

$$= v_1 \cdot (-v_2) \quad \text{by (5)}$$

$$= -(c_1b_1 + \cdots + c_kb_k) \cdot c_{k+1}b_{k+1} - \cdots - c_nb_n$$

$$= - \sum_{i=1}^{k} \sum_{j=k+1}^{n} c_ic_j b_i \cdot b_j$$

$$= - \sum_{i=1}^{k} \sum_{j=k+1}^{n} c_ic_j (0)$$

$$= 0 \quad \text{(contradiction!)}$$

The fifth step here (setting each $b_i \cdot b_j$ equal to zero) is justified by the fact that we have set things up precisely so that each vector in $W$ is particular to every vector in $W^\perp$. We conclude that the only way we can have

$$c_1b_1 + \cdots + c_kb_k + c_{k+1}b_{k+1} + \cdots + c_nb_n = 0$$

is to have each of the coefficients $c_1, \ldots, c_n$ equal to zero. Hence, the vectors $\{b_1, \ldots, b_n\}$ are linearly independent. But any set of $n$-linearly independent vectors in $\mathbb{R}^n$ will constitute a basis for $\mathbb{R}^n$. The lemma now follows.

**Theorem 16.3.** Let $W$ be a subspace of $\mathbb{R}^n$. Then every vector $v$ in $\mathbb{R}^n$ has a unique decomposition

$$v = v_W + v_{W^\perp}$$

with $v_W \in W$ and $v_{W^\perp} \in W^\perp$.

**Sketch of Proof.** We again fix a basis $B_W = \{b_1, \ldots, b_k\}$ of $W$ and a basis $B_{W^\perp} = \{b_{k+1}, \ldots, b_n\}$ for $W^\perp$ where

$$W^\perp = \text{NullSp} \begin{bmatrix} \leftarrow & b_1 & \rightarrow \\ \vdots \\ \leftarrow & b_k & \rightarrow \end{bmatrix}$$
The preceding lemma tells us that $B_W \cup B_{W^\perp}$ is a basis for $\mathbb{R}^n$. Thus, every vector $\mathbf{v} \in \mathbb{R}^n$ has a unique expression as

$$
\mathbf{v} = c_1 \mathbf{b}_1 + \cdots + c_k \mathbf{b}_k + c_{k+1} \mathbf{b}_{k+1} + \cdots + c_n \mathbf{b}_n
$$

$$
= (c_1 \mathbf{b}_1 + \cdots + c_k \mathbf{b}_k) + (c_{k+1} \mathbf{b}_{k+1} + \cdots + c_n \mathbf{b}_n)
$$

$$
= \mathbf{v}_W + \mathbf{v}_{W^\perp}
$$

where

$$
\mathbf{v}_W = c_1 \mathbf{b}_1 + \cdots + c_k \mathbf{b}_k \quad \in \quad W
$$

$$
\mathbf{v}_{W^\perp} = c_{k+1} \mathbf{b}_{k+1} + \cdots + c_n \mathbf{b}_n \quad \in \quad W^\perp
$$

2.1. Algorithm for Determining $\mathbf{v}_W$ and $\mathbf{v}_{W^\perp}$. We now summarize the algorithm used to in the Lemma and Theorem to obtain the splitting $\mathbf{v} = \mathbf{v}_W + \mathbf{v}_{W^\perp}$.

- Find a basis $B_W = \{\mathbf{b}_1, \ldots, \mathbf{b}_k\}$ for $W$
- Find a basis $B_{W^\perp} = \{\mathbf{b}_{k+1}, \ldots, \mathbf{b}_n\}$ for $W^\perp = \text{NullSp}(A_{W,B})$ (cf. (3)).
- Find the coordinate vector of $\mathbf{v}$ with respect to the basis $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_k, \mathbf{b}_{k+1}, \ldots, \mathbf{b}_n\}$ of $\mathbb{R}^n$
  
  using the row reduction

$$
\begin{bmatrix}
| & | & | \\
\mathbf{b}_1 & \cdots & \mathbf{b}_n & \mathbf{v} \\
| & | & | \\
| & | & | \\
| & | & | \\
| & | & | \\
| & | & | \\
0 & 1 & & \\
1 & \cdots & 0 & \mathbf{v}_B
\end{bmatrix}
\rightarrow
\begin{bmatrix}
1 & \cdots & 0 & | \\
\vdots & \cdots & \vdots & | \\
0 & 1 & | \\
\end{bmatrix} = [I \mid \mathbf{v}_B]
$$

- Set

$$
\mathbf{v}_W = c_1 \mathbf{b}_1 + \cdots + c_k \mathbf{b}_k
$$

$$
\mathbf{v}_{W^\perp} = c_{k+1} \mathbf{b}_{k+1} + \cdots + c_n \mathbf{b}_n
$$

where $c_i$, is the $i^{th}$ component of the coordinate vector $\mathbf{v}_B$.

Example 16.4. Let $W = \text{span}\{[1,0,1],[0,1,1]\} \subset \mathbb{R}^3$. Decompose the vector $\mathbf{v} = [1,4,-4]$ into its components $\mathbf{v}_W \in W$ and $\mathbf{v}_{W^\perp} \in W^\perp$.

- The two vectors $\mathbf{b}_1 = [1,0,1]$ and $\mathbf{b}_2 = [0,1,1]$ are obviously linearly independent and so $B_W = \{\mathbf{b}_1, \mathbf{b}_2\}$ is already a basis for $W$. To get a basis for $W^\perp$, we compute the null space of $\mathbf{A}_{W,B} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$

This matrix is already in reduced row echelon form and so its null space will be the solution set of

$$
x_1 + x_3 = 0
$$

$$
x_2 + x_3 = 0
$$

$$
\implies x = x_3 \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} \implies B_{W^\perp} = \{[-1,-1,1]\} \equiv \{\mathbf{b}_3\}
$$

We now compute the coordinate vector of $\mathbf{v} = [1,2,1]$ with respect to $B = \{[1,0,1],[0,1,1],[-1,-1,1]\}$

$$
\begin{bmatrix}
1 & 0 & -1 \\
0 & 1 & -1 \\
1 & 1 & -4
\end{bmatrix}
\rightarrow
\begin{bmatrix}
1 & 0 & 0 & -2 \\
0 & 1 & 0 & 1 \\
0 & 0 & 1 & -3
\end{bmatrix}
$$

So $\mathbf{v}_B = [-2,1,-3]$. But now

$$
\mathbf{v} = (-2) \mathbf{b}_1 + (1) \mathbf{b}_2 + (-3) \mathbf{b}_3
$$

and so

$$
\mathbf{v}_W = (-2) \mathbf{b}_1 + (1) \mathbf{b}_2 = [-2,1,-1]
$$

$$
\mathbf{v}_{W^\perp} = (-3) \mathbf{b}_3 = [3,3,-3]
$$
Example 16.5. Find the projection of the vector \( \mathbf{v} = [1, 2, 1] \) on the solution set of \( x_1 + x_2 + x_3 = 0 \).

- Let
  \[
  W = \{ [x_1, x_2, x_3] \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 = 0 \}
  \]

  This is obviously spanned by vectors of the form
  \[
  \mathbf{x} = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}
  \]
  and so \( \{[-1, 1, 0], [-1, 0, 1] \} \) is a basis for \( W \). \( W^\perp \) will then be
  \[
  \text{NullSp} \left( \begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \right) = \text{NullSp} \left( \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix} \right) = \text{span} ([1, 1, 1])
  \]

So we need to find the first to component of the coordinate vector of \( \mathbf{v} \) with respect to the basis \( \{[-1, 1, 0], [-1, 0, 1], [1, 1, 1] \} \) of \( \mathbb{R}^3 \).

\[
\begin{bmatrix}
-1 & -1 & 1 & 1 \\
1 & 0 & 1 & 2 \\
0 & 1 & 1 & 1
\end{bmatrix} \quad \longrightarrow \quad
\begin{bmatrix}
1 & 0 & 0 & \frac{2}{3} \\
0 & 1 & 0 & -\frac{2}{3} \\
0 & 0 & 1 & \frac{4}{3}
\end{bmatrix}
\]

So
\[
\mathbf{v}_W = \frac{2}{3} [-1, 1, 0] - \frac{1}{3} [-1, 0, 1] = \begin{bmatrix} -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \end{bmatrix}
\]