LECTURE 16

Orthogonality

One of the most useful properties of the standard basis $[\mathbf{e}_1, \dots, \mathbf{e}_n]$ of $\mathbb{R}^n$ is the fact that

$$\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij} \equiv \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \tag{1}$$

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

$$\begin{aligned} \mathbf{v} &= v_1\mathbf{e}_1 + \cdots + v_n\mathbf{e}_n \\ \Longrightarrow \quad & \mathbf{e}_i \cdot \mathbf{v} = \mathbf{e}_i \cdot (v_1\mathbf{e}_1 + \cdots + v_n\mathbf{e}_n) \\ &= v_1\mathbf{e}_i \cdot \mathbf{e}_1 + v_2\mathbf{e}_i \cdot \mathbf{e}_2 + \cdots + v_i\mathbf{e}_i \cdot \mathbf{e}_i + \cdots + v_n\mathbf{e}_i \cdot \mathbf{e}_n \\ &= 0 + 0 + \cdots + 0 + v_i + 0 + \cdots + 0 \\ &= v_i \end{aligned}$$

Of course, this is clear already once we write $\mathbf{v}$ and $\mathbf{e}_i$ in component form

$$\left. \begin{array}{r} \mathbf{v} = [v_1, v_2, \dots, v_i, \dots, v_n] \\ \mathbf{e}_i = [0, \dots, 0, 1, 0, \dots, 0] \end{array} \right\} \quad \Longrightarrow \quad \mathbf{e}_i \cdot \mathbf{v} = v_i$$

However, it is not true for a more general basis. Recall that for a general basis $B = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$, in order to find the constants $c_1, \dots, c_n$ such that

$$\mathbf{v} = c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n$$

you have to solve the linear system

$$\begin{bmatrix} | & & | \\ \mathbf{b}_1 & \cdots & \mathbf{b}_n \\ | & & | \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$$

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 $B = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ that enjoy orthogonality properties just like (1)

$$\mathbf{b}_i \cdot \mathbf{b}_j = \delta_{ij} \equiv \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

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

$$\mathbf{v} = \mathbf{v} = c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n$$

by simply computing inner products

$$c_i = \mathbf{b}_i \cdot \mathbf{v}$$

1. Projections onto Vectors

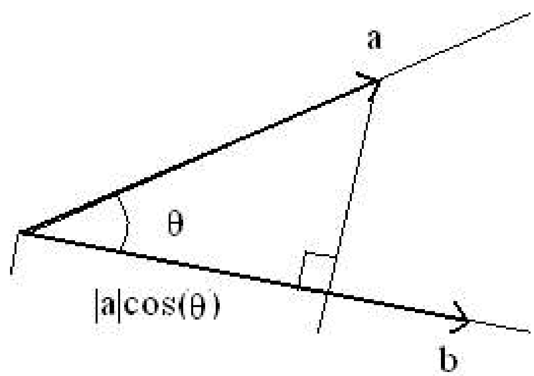
Recall that the inner product $\mathbf{a} \cdot \mathbf{b}$ of two vectors in $\mathbb{R}^n$ has a very concrete geometric interpretation

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

where

$$\begin{aligned} \|\mathbf{a}\| &= \sqrt{\mathbf{a} \cdot \mathbf{a}} \equiv \text{the length of } \mathbf{a} \\ \|\mathbf{b}\| &= \sqrt{\mathbf{b} \cdot \mathbf{b}} \equiv \text{the length of } \mathbf{b} \\ \theta_{\mathbf{ab}} &= \text{ the angle between } \mathbf{a} \text{ and } \mathbf{b} \text{ in the plane spanned by } \mathbf{a} \text{ and } \mathbf{b} \end{aligned}$$

Let's look more closely at the actual geometric situation in the 2-dimensional plane spanned by **a** and **b**.



We see from the diagram above that

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

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

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

we can have the following formula

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

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 **b** with length $\frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{b}\|}$. Now the unit vector in the direction of **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

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

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

$$\mathbf{P}_{\mathbf{a},\mathbf{b}} = \frac{\mathbf{a} \cdot \mathbf{b}}{\mathbf{b} \cdot \mathbf{b}} \mathbf{b}$$

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

$$\mathbf{v} \in W^\perp \quad \Longrightarrow \quad \mathbf{v} \cdot \mathbf{w} = 0 \qquad \text{for every vector } \mathbf{w} \in W \tag{2}$$

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

$$\mathbf{v} \cdot \mathbf{b}_i = 0 \qquad \text{for } i = 1, \ldots, k$$

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

$$\begin{bmatrix} \mathbf{b}_1 \cdot \mathbf{v} \\ \vdots \\ \mathbf{b}_k \cdot \mathbf{v} \end{bmatrix} = \begin{bmatrix} \longleftarrow & \mathbf{b}_1 & \longrightarrow \\ & \vdots & \\ \longleftarrow & \mathbf{b}_k & \longrightarrow \end{bmatrix} \mathbf{v} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

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-dimensional) basis vectors $\mathbf{b}_i$ as rows. Set

$$W^\perp \equiv NullSp \left(\begin{bmatrix} \longleftarrow & \mathbf{b}_1 & \longrightarrow \\ & \vdots & \\ \longleftarrow & \mathbf{b}_k & \longrightarrow \end{bmatrix} \right)$$

Then, we have set things up so that

$$\mathbf{v} \in W^\perp \quad \Longleftrightarrow \quad \mathbf{v} \cdot \mathbf{w} = 0 \quad \text{for all } \mathbf{w} \in W$$

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

$$\mathbf{A}_{W,B} = \begin{bmatrix} \longleftarrow & \mathbf{b}_1 & \longrightarrow \\ & \vdots & \\ \longleftarrow & \mathbf{b}_k & \longrightarrow \end{bmatrix} \tag{3}$$

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

$$n = \#\text{ columns } = rank\left(\mathbf{A}_{W,B}\right) + \dim\left(NullSp\left(\mathbf{A}_{W,B}\right)\right)$$

we have

$$\dim W^\perp = n - k$$

So we can find a basis $B_{W^\perp} = \{\mathbf{v}_1, \dots, \mathbf{v}_{n-k}\}$ for $W^\perp$. Let's write this change notation slightly are write $\{\mathbf{b}_{k+1}, \dots, \mathbf{b}_n\}$ for the $n-k$ basis vector $\{\mathbf{v}_1, \dots, \mathbf{v}_{n-k}\}$.

LEMMA 16.2. *The set $\{\mathbf{b}_1, \dots, \mathbf{b}_k, \mathbf{b}_{k+1}, \dots, \mathbf{b}_n\}$ where $\{\mathbf{b}_1, \dots, \mathbf{b}_k\}$ is our given basis for W and $\{\mathbf{b}_{k+1}, \dots, \mathbf{b}_n\}$ is a basis for the null space of $\mathbf{A}_{W,B}$, is a basis for $\mathbb{R}^n$.*

Proof. Suppose

$$c_1\mathbf{b}_1 + \cdots + c_k\mathbf{b}_k + c_{k+1}\mathbf{b}_{k+1} + \cdots + c_n\mathbf{b}_n = \mathbf{0}$$

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

$$c_1\mathbf{b}_1 + \cdots + c_k\mathbf{b}_k = -c_{k+1}\mathbf{b}_{k+1} - \cdots - c_n\mathbf{b}_n \tag{4}$$

Set

$$\begin{aligned} \mathbf{v}_1 &= c_1\mathbf{b}_1 + \cdots + c_k\mathbf{b}_k \in W \\ \mathbf{v}_2 &= c_{k+1}\mathbf{b}_{k+1} + \cdots + c_n\mathbf{b}_n \in W^\perp \end{aligned}$$

so that (4) becomes

$$\mathbf{v}_1 = -\mathbf{v}_2 \tag{5}$$

Since the basis vectors set $\{\mathbf{b}_1, \dots, \mathbf{b}_k\}$ and $\{\mathbf{b}_{k+1}, \dots, \mathbf{b}_n\}$ are linearly independent, neither $\mathbf{v}_1$ nor $\mathbf{v}_2$ can be $\mathbf{0}$ unless all the coefficients $c_1 \dots, c_n$ are zero, which is a situation that we have excluded from the start. But then if $\mathbf{v}_1 \neq 0$

$$\begin{aligned} 0 &\neq \|\mathbf{v}_1\|^2 = \mathbf{v}_1 \cdot \mathbf{v}_1 \\ &= \mathbf{v}_1 \cdot (-\mathbf{v}_2) \qquad \text{by (5)} \\ &= -(c_1\mathbf{b}_1 + \cdots + c_k\mathbf{b}_k) \cdot c_{k+1}\mathbf{b}_{k+1} - \cdots - c_n\mathbf{b}_n \\ &= -\sum_{i=1}^{k}\sum_{j=\bar{k}1}^{n} c_i c_j \mathbf{b}_i \cdot \mathbf{b}_j \\ &= -\sum_{i=1}^{k}\sum_{j=\bar{k}1}^{n} c_i c_j (0) \\ &= 0 \qquad \text{(contradiction!)} \end{aligned}$$

The fifth step here (setting each $\mathbf{b}_i \cdot \mathbf{b}_j$ 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

$$c_1\mathbf{b}_1 + \cdots + c_k\mathbf{b}_k + c_{k+1}\mathbf{b}_{k+1} + \cdots + c_n\mathbf{b}_n = \mathbf{0}$$

is to have each of the coefficients $c_1, \dots, c_n$ equal to zero. Hence, the vectors $\{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ 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 $\mathbf{v}$ in $\mathbb{R}^n$ has a unique decomposition*

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

with $\mathbf{v}_W \in W$ and $\mathbf{v}_{W^\perp} \in W^\perp$.

Sketch of Proof. We again fix a basis $B_W = \{\mathbf{b}_1, \dots, \mathbf{b}_k\}$ of W and a basis $B_{W^\perp} = \{\mathbf{b}_{k+1}, \dots, \mathbf{b}_n\}$ for $W^\perp$ where

$$W^\perp = NullSp\left(\begin{bmatrix} \longleftarrow & \mathbf{b}_1 & \longrightarrow \\ & \vdots & \\ \longleftarrow & \mathbf{b}_k & \longrightarrow \end{bmatrix}\right)$$

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

$$\begin{aligned} \mathbf{v} &= c_1\mathbf{b}_1 + \cdots + c_k\mathbf{b}_k + c_{k+1}\mathbf{b}_{k+1} + \cdots + c_n\mathbf{b}_n \\ &= (c_1\mathbf{b}_1 + \cdots + c_k\mathbf{b}_k) + (c_{k+1}\mathbf{b}_{k+1} + \cdots + c_n\mathbf{b}_n) \\ &= \mathbf{v}_W + \mathbf{v}_{W^\perp} \end{aligned}$$

where

$$\begin{aligned} \mathbf{v}_W &= c_1\mathbf{b}_1 + \cdots + c_k\mathbf{b}_k \quad \in \quad W \\ \mathbf{v}_{W^\perp} &= c_{k+1}\mathbf{b}_{k+1} + \cdots + c_n\mathbf{b}_n \quad \in \quad W^\perp \end{aligned}$$

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 $W^\perp = NullSp(\mathbf{A}_{W,B})$ (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

$$\left[\begin{array}{ccc|c} | & & | & | \\ \mathbf{b}_1 & \cdots & \mathbf{b}_n & \mathbf{v} \\ | & & | & | \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & \cdots & 0 & | \\ \vdots & \cdots & \vdots & \mathbf{v}_B \\ 0 & & 1 & | \end{array}\right] = [\mathbf{I} \mid \mathbf{v}_B]$$

- Set

$$\begin{aligned} \mathbf{v}_W &= c_1\mathbf{b}_1 + \cdots + c_k\mathbf{b}_k \\ \mathbf{v}_{W^\perp} &= c_{k+1}\mathbf{b}_{k+1} + \cdots + c_n\mathbf{b}_n \end{aligned}$$

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

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

- The two vectors $\mathbf{b}_1 \equiv [1,0,1]$ 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

$$\mathbf{A}_{W,B} = \left[\begin{array}{ccc} 1 & 0 & 1 \\ 0 & 1 & 1 \end{array}\right]$$

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

$$\left.\begin{array}{l} x_1 + x_3 = 0 \\ x_2 + x_3 = 0 \end{array}\right\} \implies \mathbf{x} = x_3 \left[\begin{array}{c} -1 \\ -1 \\ 1 \end{array}\right] \implies B_{W^\perp} = \{[-1,-1,1]\} \equiv \{\mathbf{b}_3\}$$

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

$$\left[\begin{array}{ccc|c} 1 & 0 & -1 & 1 \\ 0 & 1 & -1 & 4 \\ 1 & 1 & 1 & -4 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -3 \end{array}\right]$$

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

$$\begin{aligned} \mathbf{v}_W &= (-2)\,\mathbf{b}_1 + (1)\,\mathbf{b}_2 = [-2,1,-1] \\ \mathbf{v}_{W^\perp} &= (-3)\,\mathbf{b}_3 = [3,3,-3] \end{aligned}$$

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
$$W = \left\{[x_1, x_2, x_3] \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 = 0\right\}$$
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
$$NullSp\left(\begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}\right) = NullSp\left(\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix}\right) = span\left([1,1,1]\right)$$
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 $\mathbb{R}^3$.
$$\left[\begin{array}{ccc|c} -1 & -1 & 1 & 1 \\ 1 & 0 & 1 & 2 \\ 0 & 1 & 1 & 1 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & \frac{2}{3} \\ 0 & 1 & 0 & -\frac{1}{3} \\ 0 & 0 & 1 & \frac{4}{3} \end{array}\right]$$
So
$$\mathbf{v}_W = \frac{2}{3}[-1,1,0] - \frac{1}{3}[-1,0,1] = \left[-\frac{1}{3}, \frac{2}{3}, \frac{-1}{3}\right]$$