1 Basic Notions

Definition 1. Let $\vec{u} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$ be two vectors in $\mathbb{R}^n$. We define the *dot product* to be

$$\vec{u} \cdot \vec{v} = a_1 b_1 + a_2 b_2 + \ldots + a_n b_n$$

Example 1. Let $\vec{u} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$

$$\vec{u} \cdot \vec{v} = 1(4) + 2(5) + 3(6)$$

Theorem 1. *Let $\vec{u}, \vec{v}, \vec{w}$ be vectors in $\mathbb{R}^n$ and $c \in \mathbb{R}$.*

- $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$
- $(\vec{u} + \vec{v}) \cdot \vec{w} = \vec{u} \cdot \vec{w} + \vec{v} \cdot \vec{w}$
- $(c\vec{u}) \cdot \vec{v} = c(\vec{u} \cdot \vec{v}) = \vec{u} \cdot c\vec{v}$
- $\vec{u} \cdot \vec{u} \geq 0$ *and* $\vec{u} \cdot \vec{u} = 0$ *if and only if* $\vec{u} = 0$.

Definition 2. We define the *norm* (magnitude, length) of a vector $\vec{u}$ to be $\|\vec{u}\| = \sqrt{\vec{u} \cdot \vec{u}}$. A vector with length 1 is called a *unit vector*. We define the *distance* between two vectors $\vec{v}$ and $\vec{u}$ is defined to be $\|\vec{v} - \vec{u}\|$.

Note that $\dfrac{\vec{v}}{\|\vec{v}\|}$ is always a unit vector.

Definition 3. We say that $\vec{u}$ is *orthogonal* (perpendicular) to $\vec{v}$, written $\vec{u} \perp \vec{v}$, if and only if $\vec{u} \cdot \vec{v} = 0$.

Definition 4. If W is a subspace, we say $\vec{v}$ is orthogonal to W, written $\vec{v} \perp W$ if and only if $\vec{v} \perp \vec{w}$ for all $\vec{w} \in W$. We call the set of all vectors $\vec{v} \perp W$ the *orthogonal complement* of W and denote this set $W^{\perp}$.

Proposition 2. *$W^{\perp}$ is a subspace. Moreover, $\vec{v} \perp W$ if and only if $\vec{v} \perp \vec{w}_i$ for $\imath = 1, \ldots, p$ where $\{\vec{w}_1, \ldots \vec{w}_p\}$ is a basis for W.*

Definition 5. A set of vectors $\{\vec{u}_1, \ldots, \vec{u}_p\}$ is orthogonal if and only if

$$\vec{u}_i \cdot \vec{u}_j = 0 \text{ when } i \neq j.$$

An *orthogonal basis* is a basis which is also an orthogonal set. An *orthonormal basis* is an orthogonal basis consisting of unit vectors.

Theorem 3. *An orthogonal set of vectors is linearly independent. Moreover, if an orthogonal set $\{\vec{u}_1, \ldots \vec{u}_p\}$ spans a subspace U then $\{u_1, \ldots \vec{u}_p\}$ is a basis for U.*

Note that if W is an n dimensional vector space, any set of n orthogonal vectors is automatically a basis for W.

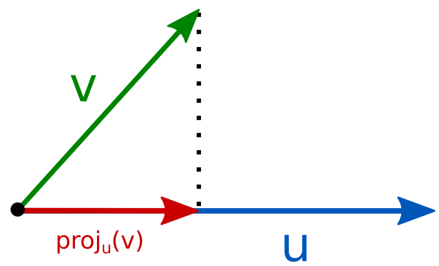
2 Orthogonal Projections

Definition 6. Given two vectors $\vec{v}$ and $\vec{u}$ we define

$$\hat{v} = \left(\frac{\vec{v}\cdot\vec{u}}{\vec{u}\cdot\vec{u}}\right)\vec{u}$$

to be the *orthogonal projection* of $\vec{v}$ onto $\vec{u}$. This is sometimes denoted

$$\hat{v} = \text{proj}_{\vec{u}}(\vec{v}).$$



In the above picture the green vector $\vec{v}$ is being projected onto the blue vector $\vec{u}$ and the resulting red vector is the projection. Notice that the projection points in the same direction as $\vec{u}$.

Moreover,

$$\vec{z} = (\vec{v} - \hat{v}) \perp \vec{u}$$

and the length of $\vec{z} = \vec{v} - \hat{v}$ is represented by the dashed line.

Definition 7. Let $\vec{v}$ and $\vec{u}$ be two vectors let $\hat{v}$ be the projection of $\vec{v}$ onto $\vec{u}$. If we define $\vec{z} = \vec{v} - \hat{v}$ then

$$\vec{v} = \hat{v} + \vec{z}$$

where $\hat{v}$ is parallel to $\vec{u}$ and $\vec{z}$ is perpendicular to $\vec{u}$. We call $\hat{v}$ the component of $\vec{v}$ parallel to $\vec{u}$ and $\vec{z}$ is the component of $\vec{v}$ perpendicular to $\vec{u}$.

Example 2. Let $\vec{v} = \begin{bmatrix} 7 \\ 6 \end{bmatrix}$ and $\vec{u} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$.

$$\vec{v}\cdot\vec{u} = 40, \qquad \vec{u}\cdot\vec{u} = 20, \qquad \hat{v} = \frac{40}{20}\vec{u} = 2\vec{u}$$

Hence we have that,

$$\vec{z} = \vec{v} - \hat{v} = \begin{bmatrix} 7 \\ 6 \end{bmatrix} - 2\begin{bmatrix} 4 \\ 2 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}.$$

Notice that,

$$\vec{z}\cdot\vec{u} = 0$$

and

$$\vec{v} = \begin{bmatrix} 7 \\ 6 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix} + \begin{bmatrix} -1 \\ 2 \end{bmatrix} = \hat{v} + \vec{z}$$

Orthogonal basis are convenient for the following reason.

Theorem 4. *Let $\{\vec{u}_1, \ldots \vec{u}_p\}$ be an orthogonal basis for W. For all $\vec{y} \in W$ we have*

$$\vec{y} = c_1\vec{u}_1 + \ldots + c_p\vec{u}_p \tag{1}$$

and

$$c_j = \left(\frac{\vec{y} \cdot \vec{u}_j}{\vec{u}_j \cdot \vec{u}_j}\right) \text{ for } j = 1, \ldots, p \tag{2}$$

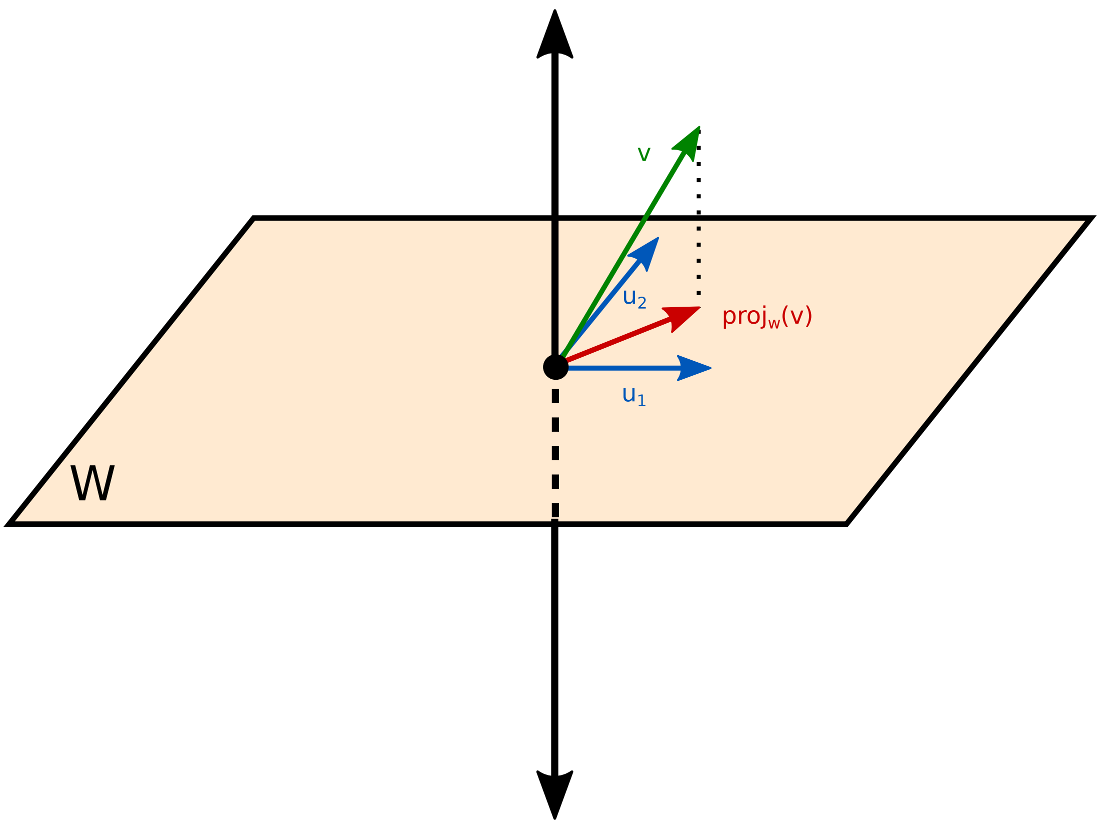
The above formulas (1) and (2) will make a bit more sense after the following definition.

Definition 8. Let W be a subspace of $\mathbb{R}^n$. Then each $\vec{v} \in W$ can be written uniquely in the form

$$\vec{v} = \hat{v} + \vec{z}$$

where $\hat{v} \in W$ and $\vec{z} \in W^{\perp}$. We call $\hat{v}$ the orthogonal projection of $\vec{v}$ onto W, sometimes denoted $\hat{v} = \text{proj}_W(\vec{v})$. Moreover, if $\{\vec{u}_1, \ldots, \vec{u}_p\}$ is an orthogonal basis for W then

$$\hat{v} = \left(\frac{\vec{v} \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1}\right)\vec{u}_1 + \ldots + \left(\frac{\vec{v} \cdot \vec{u}_p}{\vec{u}_p \cdot \vec{u}_p}\right)\vec{u}_p$$



In the above picture, W is a 2 dimensional subspace with basis vectors $\vec{u}_1$ and $\vec{u}_2$ (blue vectors). The red vector,$\hat{v}$, is the projection of the green vector, $\vec{v}$ onto W. Moreover, the dashed line is the length of the vector $\vec{v} - \hat{v}$.

Definition 9. The distance between a vector $\vec{v}$ and a subspace W is given by $\|\vec{z}\| = \|\vec{v} - \hat{v}\|$.

Notice, (1) and (2) makes some more sense in the context of the above definition. If $\vec{y}$ is in the subspace W then it is *equal* to its projection onto W, i.e. $\vec{y} = \hat{y}$.

3 Gram Schmidt

The Gram Schmidt process allows us to get an orthogonalized version of any basis. Essentially, at each step of the process you subtract off the projection onto a subspace. Moreover, we can rescale at each step since if $\vec{v} \perp \vec{u}$ the $c\vec{v} \perp \vec{u}$.

Theorem 5. *Given a basis $\{\vec{x}_1, \ldots, \vec{x}_p\}$ for a subspace W of $\mathbb{R}^n$. Define,*

$$\begin{aligned}
\vec{v}_1 &= \vec{x}_1 \\
\vec{v}_2 &= \vec{x}_2 - \left(\frac{\vec{x}_2 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1}\right)\vec{v}_1 \\
\vec{v}_3 &= \vec{x}_3 - \left(\frac{\vec{x}_3 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1}\right)\vec{v}_1 - \left(\frac{\vec{x}_3 \cdot \vec{v}_2}{\vec{v}_2 \cdot \vec{v}_2}\right)\vec{v}_2 \\
&\vdots \\
\vec{v}_p &= \vec{x}_p - \left(\frac{\vec{x}_p \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1}\right)\vec{v}_1 - \left(\frac{\vec{x}_p \cdot \vec{v}_2}{\vec{v}_2 \cdot \vec{v}_2}\right)\vec{v}_2 - \ldots - \left(\frac{\vec{x}_p \cdot \vec{v}_{p-1}}{\vec{v}_{p-1} \cdot \vec{v}_{p-1}}\right)\vec{v}_{p-1}
\end{aligned}$$

Then $\{\vec{v}_1, \ldots, \vec{v}_p\}$ is an orthogonal basis for W and

$$\text{span}\{\vec{v}_1, \ldots, \vec{v}_k\} = \text{span}\{\vec{x}_1, \ldots, \vec{x}_k\} \qquad \text{for } k = 1, \ldots, p.$$

Example 3. Let

$$\vec{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \vec{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \vec{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$$

Step 1:

Set $\vec{v}_1 = \vec{x}_1$.

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

Step 2:

Set $\vec{v}_2 = \vec{x}_2 - \left(\frac{\vec{x}_2 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1}\right)\vec{v}_1$

$$\vec{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \left(\frac{\vec{x}_2 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1}\right)\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \left(\frac{3}{4}\right)\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -3/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{bmatrix}.$$

Step 3:

Rescale

$$\vec{v}_2 = \begin{bmatrix} -3 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

Step 4:

Set $\vec{v}_3 = \vec{x}_3 - \left(\frac{\vec{x}_3 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1}\right) \vec{v}_1 - \left(\frac{\vec{x}_3 \cdot \vec{v}_2}{\vec{v}_2 \cdot \vec{v}_2}\right) \vec{v}_2$

$$\vec{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} - \left(\frac{2}{4}\right) \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \left(\frac{2}{12}\right) \begin{bmatrix} -3 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -2/3 \\ 1/3 \\ 1/3 \end{bmatrix}.$$

Step 5:

Rescale

$$\vec{v}_3 = \begin{bmatrix} 0 \\ -2 \\ 1 \\ 1 \end{bmatrix}.$$

Result:

We have that

$$\vec{v}_1 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} -3 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \vec{v}_3 = \begin{bmatrix} 0 \\ -2 \\ 1 \\ 1 \end{bmatrix}$$