Three.VI Projection

Linear Algebra

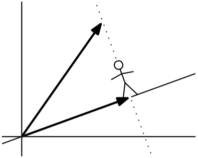
Jim Hefferon

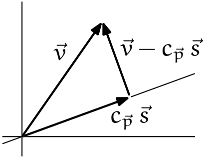
http://joshua.smcvt.edu/linearalgebra

Orthogonal Projection Into a Line

Project a vector into a line

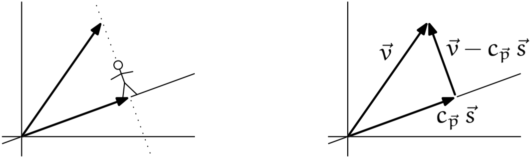
This shows a figure walking out on the line to a point $\vec{p}$ such that the tip of $\vec{v}$ is directly above them, where "above" does not mean parallel to the y-axis but instead means orthogonal to the line.





Project a vector into a line

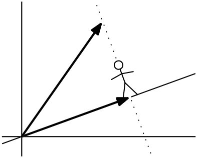
This shows a figure walking out on the line to a point $\vec{p}$ such that the tip of $\vec{v}$ is directly above them, where "above" does not mean parallel to the y-axis but instead means orthogonal to the line.

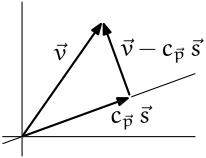


Since the line is the span of some vector $\ell = \{c \cdot \vec{s} \mid c \in \mathbb{R}\}$, we have a coefficient $c_{\vec{p}}$ with the property that $\vec{v} - c_{\vec{p}}\vec{s}$ is orthogonal to $c_{\vec{p}}\vec{s}$.

Project a vector into a line

This shows a figure walking out on the line to a point $\vec{p}$ such that the tip of $\vec{v}$ is directly above them, where "above" does not mean parallel to the y-axis but instead means orthogonal to the line.

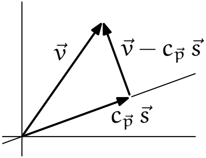




Since the line is the span of some vector $\ell = \{c \cdot \vec{s} \mid c \in \mathbb{R}\}$, we have a coefficient $c_{\vec{p}}$ with the property that $\vec{v} - c_{\vec{p}}\vec{s}$ is orthogonal to $c_{\vec{p}}\vec{s}$.

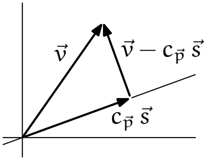
To solve for this coefficient, observe that because $\vec{v} - c_{\vec{p}}\vec{s}$ is orthogonal to a scalar multiple of $\vec{s}$, it must be orthogonal to $\vec{s}$ itself. Then $(\vec{v} - c_{\vec{p}}\vec{s}) \bullet \vec{s} = 0$ gives that $c_{\vec{p}} = \vec{v} \bullet \vec{s} / \vec{s} \bullet \vec{s}$.

We have decomposed $\vec{v}$ into two parts $\vec{v} = (c_{\vec{p}}\vec{s}) + (v - c_{\vec{p}}\vec{s})$.



Intuitively, some of $\vec{v}$ lies with the line and that gives the first part $c_{\vec{p}}\vec{s}$. The part of $\vec{v}$ that lies with a line orthogonal to ℓ is $\vec{v} - c_{\vec{p}}\vec{s}$. What's compelling about pairing these two parts is that they don't interact, in that the projection of one into the line spanned by the other is the zero vector.

We have decomposed $\vec{v}$ into two parts $\vec{v} = (c_{\vec{p}}\vec{s}) + (v - c_{\vec{p}}\vec{s})$.



Intuitively, some of $\vec{v}$ lies with the line and that gives the first part $c_{\vec{p}}\vec{s}$. The part of $\vec{v}$ that lies with a line orthogonal to ℓ is $\vec{v} - c_{\vec{p}}\vec{s}$. What's compelling about pairing these two parts is that they don't interact, in that the projection of one into the line spanned by the other is the zero vector.

Note: We have not given a definition of 'angle' in spaces other than $\mathbb{R}^n$'s, so we will stick here to those spaces. Extending the definitions to other spaces is perfectly possible but we don't need them here.

1.1 *Definition* The *orthogonal projection of $\vec{v}$ into the line spanned by a nonzero $\vec{s}$* is this vector.

$$\operatorname{proj}_{[\vec{s}\,]}(\vec{v}) = \frac{\vec{v} \bullet \vec{s}}{\vec{s} \bullet \vec{s}} \cdot \vec{s}$$

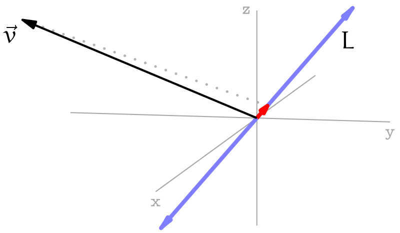
Example The projection of this $\mathbb{R}^3$ vector into the line

$$\vec{v} = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \qquad L = \{c \cdot \vec{s} = c \cdot \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} \mid c \in \mathbb{R}\}$$

is this vector.

$$\frac{\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \bullet \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}}{\begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} \bullet \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}} \cdot \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1/3 \\ 1/6 \\ 1/6 \end{pmatrix}$$

Because $\vec{v}$ is nearly orthogonal to the line L

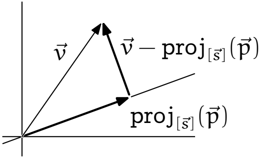


only a small part of $\vec{v}$ lies with the direction of that line, so the projected-to red vector $\mathrm{proj}_{[\vec{s}\,]}(\vec{v})$ is quite short: ($|\vec{v}| = \sqrt{6} \approx 2.45$ while $|\mathrm{proj}_{[\vec{s}\,]}(\vec{v})| = \sqrt{1/6} \approx 0.41$).

Gram-Schmidt Orthogonalization

Mutually orthogonal vectors

The prior subsection suggests that projecting a vector $\vec{v}$ into the line spanned by $\vec{s}$ decomposes $\vec{v}$ into two parts, a part with the line and a part orthogonal to that.

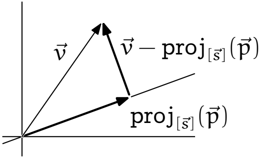


$$\vec{v} = \mathrm{proj}_{[\vec{s}\,]}(\vec{v}) + \left(\vec{v} - \mathrm{proj}_{[\vec{s}\,]}(\vec{v})\right)$$

Because these are orthogonal they are in some sense non-interacting. Here we will develop that.

Mutually orthogonal vectors

The prior subsection suggests that projecting a vector $\vec{v}$ into the line spanned by $\vec{s}$ decomposes $\vec{v}$ into two parts, a part with the line and a part orthogonal to that.



$$\vec{v} = \mathrm{proj}_{[\vec{s}\,]}(\vec{v}) + \left(\vec{v} - \mathrm{proj}_{[\vec{s}\,]}(\vec{v})\right)$$

Because these are orthogonal they are in some sense non-interacting. Here we will develop that.

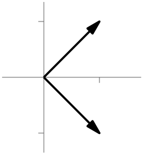
2.1 *Definition* Vectors $\vec{v}_1, \ldots, \vec{v}_k \in \mathbb{R}^n$ are *mutually orthogonal* when any two are orthogonal: if $i \neq j$ then the dot product $\vec{v}_i \cdot \vec{v}_j$ is zero.

Example The vectors of the standard basis $\mathcal{E}_3 \subset \mathbb{R}^3$ are mutually orthogonal.

$$\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}$$

Example These two vectors in $\mathbb{R}^2$ are mutually orthogonal.

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$



The next result makes 'non-interacting' precise.

2.2 *Theorem* If the vectors in a set $\{\vec{v}_1, \ldots, \vec{v}_k\} \subset \mathbb{R}^n$ are mutually orthogonal and nonzero then that set is linearly independent.

The next result makes 'non-interacting' precise.

2.2 *Theorem* If the vectors in a set $\{\vec{v}_1, \ldots, \vec{v}_k\} \subset \mathbb{R}^n$ are mutually orthogonal and nonzero then that set is linearly independent.

Proof Consider $\vec{0} = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_k\vec{v}_k$. For $i \in \{1, .., k\}$, taking the dot product of $\vec{v}_i$ with both sides of the equation $\vec{v}_i \bullet (c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_k\vec{v}_k) = \vec{v}_i \bullet \vec{0}$, which gives $c_i \cdot (\vec{v}_i \bullet \vec{v}_i) = 0$, shows that $c_i = 0$ since $\vec{v}_i \neq \vec{0}$. QED

The next result makes ‘non-interacting’ precise.

2.2 *Theorem* If the vectors in a set $\{\vec{v}_1, \ldots, \vec{v}_k\} \subset \mathbb{R}^n$ are mutually orthogonal and nonzero then that set is linearly independent.

Proof Consider $\vec{0} = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_k\vec{v}_k$. For $i \in \{1, .., k\}$, taking the dot product of $\vec{v}_i$ with both sides of the equation $\vec{v}_i \cdot (c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_k\vec{v}_k) = \vec{v}_i \cdot \vec{0}$, which gives $c_i \cdot (\vec{v}_i \cdot \vec{v}_i) = 0$, shows that $c_i = 0$ since $\vec{v}_i \neq \vec{0}$. QED

2.3 *Corollary* In a k dimensional vector space, if the vectors in a size k set are mutually orthogonal and nonzero then that set is a basis for the space.

The next result makes 'non-interacting' precise.

2.2 *Theorem* If the vectors in a set $\{\vec{v}_1, \ldots, \vec{v}_k\} \subset \mathbb{R}^n$ are mutually orthogonal and nonzero then that set is linearly independent.

Proof Consider $\vec{0} = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_k\vec{v}_k$. For $i \in \{1, .., k\}$, taking the dot product of $\vec{v}_i$ with both sides of the equation $\vec{v}_i \cdot (c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_k\vec{v}_k) = \vec{v}_i \cdot \vec{0}$, which gives $c_i \cdot (\vec{v}_i \cdot \vec{v}_i) = 0$, shows that $c_i = 0$ since $\vec{v}_i \neq \vec{0}$. QED

2.3 *Corollary* In a k dimensional vector space, if the vectors in a size k set are mutually orthogonal and nonzero then that set is a basis for the space.

Proof Any linearly independent size k subset of a k dimensional space is a basis. QED

The next result makes 'non-interacting' precise.

2.2 *Theorem* If the vectors in a set $\{\vec{v}_1, \ldots, \vec{v}_k\} \subset \mathbb{R}^n$ are mutually orthogonal and nonzero then that set is linearly independent.

Proof Consider $\vec{0} = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_k\vec{v}_k$. For $i \in \{1, .., k\}$, taking the dot product of $\vec{v}_i$ with both sides of the equation $\vec{v}_i \bullet (c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_k\vec{v}_k) = \vec{v}_i \bullet \vec{0}$, which gives $c_i \cdot (\vec{v}_i \bullet \vec{v}_i) = 0$, shows that $c_i = 0$ since $\vec{v}_i \neq \vec{0}$. QED

2.3 *Corollary* In a k dimensional vector space, if the vectors in a size k set are mutually orthogonal and nonzero then that set is a basis for the space.

Proof Any linearly independent size k subset of a k dimensional space is a basis. QED

2.5 *Definition* An *orthogonal basis* for a vector space is a basis of mutually orthogonal vectors.

2.7 *Theorem* If $\langle\vec{\beta}_1,\ldots\vec{\beta}_k\rangle$ is a basis for a subspace of $\mathbb{R}^n$ then the vectors

$$\begin{aligned}
\vec{\kappa}_1 &= \vec{\beta}_1 \\
\vec{\kappa}_2 &= \vec{\beta}_2 - \mathrm{proj}_{[\vec{\kappa}_1]}(\vec{\beta}_2) \\
\vec{\kappa}_3 &= \vec{\beta}_3 - \mathrm{proj}_{[\vec{\kappa}_1]}(\vec{\beta}_3) - \mathrm{proj}_{[\vec{\kappa}_2]}(\vec{\beta}_3) \\
&\vdots \\
\vec{\kappa}_k &= \vec{\beta}_k - \mathrm{proj}_{[\vec{\kappa}_1]}(\vec{\beta}_k) - \cdots - \mathrm{proj}_{[\vec{\kappa}_{k-1}]}(\vec{\beta}_k)
\end{aligned}$$

form an orthogonal basis for the same subspace.

The book has the proof. We will instead illustrate.

Example This basis for $\mathbb{R}^2$

$$B = \langle \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \end{pmatrix} \rangle$$

does not have orthogonal vectors. To derive from it a basis $K = \langle \vec{\kappa}_1, \vec{\kappa}_2 \rangle$ that is orthogonal, start by taking the first vector unchanged.

$$\vec{\kappa}_1 = \vec{\beta}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

For $\vec{\kappa}_2$ take the part of $\vec{\beta}_2$ that does not lie with $\vec{\kappa}_1$.

$$\vec{\kappa}_2 = \vec{\beta}_2 - \mathrm{proj}_{[\vec{\kappa}_1]}(\vec{\beta}_2) = \begin{pmatrix} 1 \\ 3 \end{pmatrix} - \frac{\begin{pmatrix} 1 \\ 3 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix}}{\begin{pmatrix} 1 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix}} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} -2/5 \\ 1/5 \end{pmatrix}$$

Note that $\vec{\kappa}_1$ and $\vec{\kappa}_2$ are indeed orthogonal.

Example This is a basis for $\mathbb{R}^3$.

$$B = \langle \begin{pmatrix}1\\1\\2\end{pmatrix}, \begin{pmatrix}-1\\2\\1\end{pmatrix}, \begin{pmatrix}0\\3\\-1\end{pmatrix} \rangle$$

Start the orthogonal basis with $\vec{\beta}_1$.

$$\vec{\kappa}_1 = \begin{pmatrix}1\\1\\2\end{pmatrix}$$

Example This is a basis for $\mathbb{R}^3$.

$$B = \langle \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 3 \\ -1 \end{pmatrix} \rangle$$

Start the orthogonal basis with $\vec{\beta}_1$.

$$\vec{\kappa}_1 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$$

As in the prior slide, the next step is $\vec{\kappa}_2 = \vec{\beta}_2 - \text{proj}_{[\vec{\kappa}_1]}(\vec{\beta}_2)$.

$$\begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix} - \frac{\begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix} \bullet \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}}{\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \bullet \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}} \cdot \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} -3/2 \\ 3/2 \\ 0 \end{pmatrix}$$

The third step is $\vec{\kappa}_3 = \vec{\beta}_3 - \mathrm{proj}_{[\vec{\kappa}_1]}(\vec{\beta}_3) - \mathrm{proj}_{[\vec{\kappa}_2]}(\vec{\beta}_3)$.

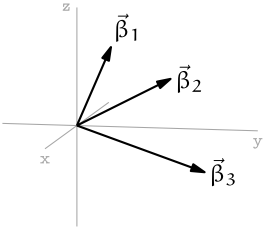
$$\begin{pmatrix}0\\3\\-1\end{pmatrix} - \frac{\begin{pmatrix}0\\3\\-1\end{pmatrix}\cdot\begin{pmatrix}1\\1\\2\end{pmatrix}}{\begin{pmatrix}1\\1\\2\end{pmatrix}\cdot\begin{pmatrix}1\\1\\2\end{pmatrix}}\cdot\begin{pmatrix}1\\1\\2\end{pmatrix} - \frac{\begin{pmatrix}0\\3\\-1\end{pmatrix}\cdot\begin{pmatrix}-3/2\\3/2\\0\end{pmatrix}}{\begin{pmatrix}-3/2\\3/2\\0\end{pmatrix}\cdot\begin{pmatrix}-3/2\\3/2\\0\end{pmatrix}}\cdot\begin{pmatrix}-3/2\\3/2\\0\end{pmatrix} = \begin{pmatrix}4/3\\4/3\\-4/3\end{pmatrix}$$

The third step is $\vec{\kappa}_3 = \vec{\beta}_3 - \mathrm{proj}_{[\vec{\kappa}_1]}(\vec{\beta}_3) - \mathrm{proj}_{[\vec{\kappa}_2]}(\vec{\beta}_3)$.

$$\begin{pmatrix}0\\3\\-1\end{pmatrix} - \frac{\begin{pmatrix}0\\3\\-1\end{pmatrix}\cdot\begin{pmatrix}1\\1\\2\end{pmatrix}}{\begin{pmatrix}1\\1\\2\end{pmatrix}\cdot\begin{pmatrix}1\\1\\2\end{pmatrix}}\cdot\begin{pmatrix}1\\1\\2\end{pmatrix} - \frac{\begin{pmatrix}0\\3\\-1\end{pmatrix}\cdot\begin{pmatrix}-3/2\\3/2\\0\end{pmatrix}}{\begin{pmatrix}-3/2\\3/2\\0\end{pmatrix}\cdot\begin{pmatrix}-3/2\\3/2\\0\end{pmatrix}}\cdot\begin{pmatrix}-3/2\\3/2\\0\end{pmatrix} = \begin{pmatrix}4/3\\4/3\\-4/3\end{pmatrix}$$

The members of B are at odd angles but the members of K are mutually orthogonal.

$$B = \langle \begin{pmatrix}1\\1\\2\end{pmatrix}, \begin{pmatrix}-1\\2\\1\end{pmatrix}, \begin{pmatrix}0\\3\\-1\end{pmatrix} \rangle$$

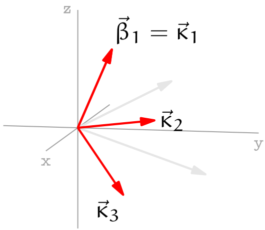


The third step is $\vec{\kappa}_3 = \vec{\beta}_3 - \mathrm{proj}_{[\vec{\kappa}_1]}(\vec{\beta}_3) - \mathrm{proj}_{[\vec{\kappa}_2]}(\vec{\beta}_3)$.

$$\begin{pmatrix}0\\3\\-1\end{pmatrix} - \frac{\begin{pmatrix}0\\3\\-1\end{pmatrix}\cdot\begin{pmatrix}1\\1\\2\end{pmatrix}}{\begin{pmatrix}1\\1\\2\end{pmatrix}\cdot\begin{pmatrix}1\\1\\2\end{pmatrix}}\cdot\begin{pmatrix}1\\1\\2\end{pmatrix} - \frac{\begin{pmatrix}0\\3\\-1\end{pmatrix}\cdot\begin{pmatrix}-3/2\\3/2\\0\end{pmatrix}}{\begin{pmatrix}-3/2\\3/2\\0\end{pmatrix}\cdot\begin{pmatrix}-3/2\\3/2\\0\end{pmatrix}}\cdot\begin{pmatrix}-3/2\\3/2\\0\end{pmatrix} = \begin{pmatrix}4/3\\4/3\\-4/3\end{pmatrix}$$

The members of B are at odd angles but the members of K are mutually orthogonal.

$$K = \langle \begin{pmatrix}1\\1\\2\end{pmatrix}, \begin{pmatrix}-3/2\\3/2\\0\end{pmatrix}, \begin{pmatrix}4/3\\4/3\\-4/3\end{pmatrix} \rangle$$

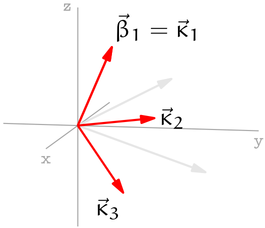


The third step is $\vec{\kappa}_3 = \vec{\beta}_3 - \mathrm{proj}_{[\vec{\kappa}_1]}(\vec{\beta}_3) - \mathrm{proj}_{[\vec{\kappa}_2]}(\vec{\beta}_3)$.

$$\begin{pmatrix} 0 \\ 3 \\ -1 \end{pmatrix} - \frac{\begin{pmatrix} 0 \\ 3 \\ -1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}}{\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}} \cdot \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} - \frac{\begin{pmatrix} 0 \\ 3 \\ -1 \end{pmatrix} \cdot \begin{pmatrix} -3/2 \\ 3/2 \\ 0 \end{pmatrix}}{\begin{pmatrix} -3/2 \\ 3/2 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} -3/2 \\ 3/2 \\ 0 \end{pmatrix}} \cdot \begin{pmatrix} -3/2 \\ 3/2 \\ 0 \end{pmatrix} = \begin{pmatrix} 4/3 \\ 4/3 \\ -4/3 \end{pmatrix}$$

The members of B are at odd angles but the members of K are mutually orthogonal.

$$K = \langle \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} -3/2 \\ 3/2 \\ 0 \end{pmatrix}, \begin{pmatrix} 4/3 \\ 4/3 \\ -4/3 \end{pmatrix} \rangle$$



We could go on to make this basis even more like $\mathcal{E}_3$ by normalizing all of its members to have length 1, making an *orthonormal* basis.