

Math 221: LINEAR ALGEBRA

Chapter 8. Orthogonality

§8-1. Orthogonal Complements and Projections

Le Chen¹

Emory University, 2021 Spring

(last updated on 01/12/2023)



Creative Commons License
(CC BY-NC-SA)

¹Slides are adapted from those by Karen Seyffarth from University of Calgary.

Linear Algebra with Applications

Lecture Notes

Current Lecture Notes Revision: Version 2018 — Revision B

These lecture notes were originally developed by Karen Seyffarth of the University of Calgary. Edits, additions, and revisions have been made to these notes by the editorial team at Lyryx Learning to accompany their text [Linear Algebra with Applications](#) based on W. K. Nicholson's original text.

In addition we recognize the following contributors. All new content contributed is released under the same license as noted below.

- Ilijas Farah, York University

BE A CHAMPION OF OER!

Contribute suggestions for improvements, new content, or errata:

A new topic

A new example or problem

A new or better proof to an existing theorem

Any other suggestions to improve the material

Contact Lyryx at info@lyryx.com with your ideas.

License



Attribution-NonCommercial-ShareAlike (CC BY-NC-SA)

This license lets others remix, tweak, and build upon your work non-commercially, as long as they credit you and license their new creations under the identical terms.

Copyright

Orthogonal Bases

The Orthogonal Complement U^\perp

Definition of Orthogonal Projection

The Projection Theorem and its Implications

Projection as a Linear Transformation

Copyright

Orthogonal Bases

The Orthogonal Complement U^\perp

Definition of Orthogonal Projection

The Projection Theorem and its Implications

Projection as a Linear Transformation

Orthogonality Basis

Definition (Orthogonality)

- ▶ Let $\vec{x}, \vec{y} \in \mathbb{R}^n$. We say the \vec{x} and \vec{y} are **orthogonal** if $\vec{x} \cdot \vec{y} = 0$.
- ▶ More generally, $X = \{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_k\} \subseteq \mathbb{R}^n$ is an **orthogonal set** if each \vec{x}_i is nonzero, and every pair of distinct vectors of X is orthogonal, i.e., $\vec{x}_i \cdot \vec{x}_j = 0$ for all $i \neq j$, $1 \leq i, j \leq k$.
- ▶ A set $X = \{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_k\} \subseteq \mathbb{R}^n$ is an **orthonormal set** if X is an orthogonal set of unit vectors, i.e., $\|\vec{x}_i\| = 1$ for all i , $1 \leq i \leq k$.



Definition (Linearly Independence)

Let V be a vector space and $S = \{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_k\}$ a subset of V . The set S is **linearly independent** if the following condition holds:

$$s_1 \vec{x}_1 + s_2 \vec{x}_2 + \dots + s_k \vec{x}_k = \vec{0} \quad \Rightarrow \quad s_1 = s_2 = \dots = s_k = 0.$$

Lemma (Independent Lemma)

Let V be a vector space and $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ an independent subset of V . If \mathbf{u} is a vector in V , but $\mathbf{u} \notin \text{span}(S)$, then $S' = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k, \mathbf{u}\}$ is independent.

— v.s. —

Lemma (Orthogonal Lemma)

Suppose $\{\vec{f}_1, \vec{f}_2, \dots, \vec{f}_m\}$ is an orthogonal subset of \mathbb{R}^n , and suppose $\vec{x} \in \mathbb{R}^n$. Define

$$\vec{f}_{m+1} = \vec{x} - \frac{\vec{x} \cdot \vec{f}_1}{\|\vec{f}_1\|^2} \vec{f}_1 - \frac{\vec{x} \cdot \vec{f}_2}{\|\vec{f}_2\|^2} \vec{f}_2 - \dots - \frac{\vec{x} \cdot \vec{f}_m}{\|\vec{f}_m\|^2} \vec{f}_m.$$

Then

1. $\vec{f}_{m+1} \cdot \vec{f}_j = 0$ for all j , $1 \leq j \leq m$.
2. If $\vec{x} \notin \text{span}\{\vec{f}_1, \vec{f}_2, \dots, \vec{f}_m\}$, then $\vec{f}_{m+1} \neq \vec{0}$, and $\{\vec{f}_1, \vec{f}_2, \dots, \vec{f}_m, \vec{f}_{m+1}\}$ is an orthogonal set.

Proof. (of orthogonal lemma)

(1) For any $1 \leq k \leq m$

$$\begin{aligned}\vec{f}_{m+1} \cdot \vec{f}_k &= \left(\vec{x} - \frac{\vec{x} \cdot \vec{f}_1}{\|\vec{f}_1\|^2} \vec{f}_1 - \frac{\vec{x} \cdot \vec{f}_2}{\|\vec{f}_2\|^2} \vec{f}_2 - \dots - \frac{\vec{x} \cdot \vec{f}_m}{\|\vec{f}_m\|^2} \vec{f}_m \right) \cdot \vec{f}_k \\&= \vec{x} \cdot \vec{f}_k - \frac{\vec{x} \cdot \vec{f}_1}{\|\vec{f}_1\|^2} \vec{f}_1 \cdot \vec{f}_k - \frac{\vec{x} \cdot \vec{f}_2}{\|\vec{f}_2\|^2} \vec{f}_2 \cdot \vec{f}_k - \dots - \frac{\vec{x} \cdot \vec{f}_m}{\|\vec{f}_m\|^2} \vec{f}_m \cdot \vec{f}_k \\&= \vec{x} \cdot \vec{f}_k - \frac{\vec{x} \cdot \vec{f}_k}{\|\vec{f}_k\|^2} \vec{f}_k \cdot \vec{f}_k \\&= \vec{x} \cdot \vec{f}_k - \vec{x} \cdot \vec{f}_k = 0.\end{aligned}$$

Proof. (continued)

(2) Since $\{\vec{f}_1, \dots, \vec{f}_m\}$ are independent, by the unique representation theorem, $\vec{x} \in \text{span}\{\vec{f}_1, \vec{f}_2, \dots, \vec{f}_m\}$, iff there exists unique representation for \vec{x}

$$\vec{x} = a_1 \vec{f}_1 + \dots + a_m \vec{f}_m.$$

Using the fact that $\{\vec{f}_1, \dots, \vec{f}_m\}$ is orthogonal, one finds that

$$a_i = \frac{\vec{x} \cdot \vec{f}_i}{\|\vec{f}_i\|^2}.$$

In other words,

$$\vec{x} \in \text{span}\{\vec{f}_1, \dots, \vec{f}_m\} \iff \vec{f}_{m+1} = \vec{x} - \frac{\vec{x} \cdot \vec{f}_1}{\|\vec{f}_1\|^2} \vec{f}_1 - \frac{\vec{x} \cdot \vec{f}_2}{\|\vec{f}_2\|^2} \vec{f}_2 - \dots - \frac{\vec{x} \cdot \vec{f}_m}{\|\vec{f}_m\|^2} \vec{f}_m = \vec{0}.$$

Now, $\vec{x} \notin \text{span}\{\vec{f}_1, \dots, \vec{f}_m\}$ implies that $\vec{f}_{m+1} \neq \vec{0}$.

Finally, $\{\vec{f}_1, \vec{f}_2, \dots, \vec{f}_m, \vec{f}_{m+1}\}$ is orthogonal thanks to (1). ■

Theorem

Let U be a subspace of \mathbb{R}^n .

1. Every orthogonal subset $\{\vec{f}_1, \vec{f}_2, \dots, \vec{f}_m\}$ of U is a subset of an orthogonal basis of U .
2. U has an orthogonal basis.

Proof.

Algorithm 1: Proof of part (1) of Theorem

Input : An orthogonal set $\{\vec{f}_1, \vec{f}_2, \dots, \vec{f}_m\} \subseteq U \subseteq \mathbb{R}^n$

$m \rightarrow n$;

while $\text{span}\{\vec{f}_1, \dots, \vec{f}_n\} \neq U$ do

 Pick up arbitrary $\vec{x} \in U \setminus \text{span}\{\vec{f}_1, \dots, \vec{f}_n\}$;


 Let \vec{f}_{n+1} be given by the Orthogonal Lemma;

 Then $\{\vec{f}_1, \dots, \vec{f}_n, \vec{f}_{n+1}\}$ is an orthogonal set;

$n + 1 \rightarrow n$;

end

Output: An orthogonal basis $\{\vec{f}_1, \dots, \vec{f}_n\}$ of U

(2) If $U = \{\vec{0}\}$, done. Otherwise, find an arbitrary nonzero vector in u and run the algorithm in (1). 

Theorem (Gram-Schmidt Orthogonalization Algorithm)

Let U be a subset of \mathbb{R}^n and let $\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_m\}$ be a basis of U . Let $\vec{f}_1 = \vec{x}_1$, and for each j , $2 \leq j \leq m$, let

$$\vec{f}_j = \vec{x}_j - \frac{\vec{x}_j \cdot \vec{f}_1}{\|\vec{f}_1\|^2} \vec{f}_1 - \frac{\vec{x}_j \cdot \vec{f}_2}{\|\vec{f}_2\|^2} \vec{f}_2 - \dots - \frac{\vec{x}_j \cdot \vec{f}_{j-1}}{\|\vec{f}_{j-1}\|^2} \vec{f}_{j-1}.$$

Then $\{\vec{f}_1, \vec{f}_2, \dots, \vec{f}_m\}$ is an orthogonal basis of U , and

$$\text{span}\{\vec{f}_1, \vec{f}_2, \dots, \vec{f}_j\} = \text{span}\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_j\} \quad \forall j = 1, \dots, m.$$

Algorithm 2: Gram-Schmidt Orthogonalization Algorithm

Input : A basis $\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_m\} \subseteq U \subseteq \mathbb{R}^n$

$\vec{f}_1 \leftarrow \vec{x}_1;$

for $j \leftarrow 2$ to m do

$$\vec{f}_j \leftarrow \vec{x}_j - \frac{\vec{x}_j \cdot \vec{f}_1}{\|\vec{f}_1\|^2} \vec{f}_1 - \frac{\vec{x}_j \cdot \vec{f}_2}{\|\vec{f}_2\|^2} \vec{f}_2 - \dots - \frac{\vec{x}_j \cdot \vec{f}_{j-1}}{\|\vec{f}_{j-1}\|^2} \vec{f}_{j-1}.$$

end

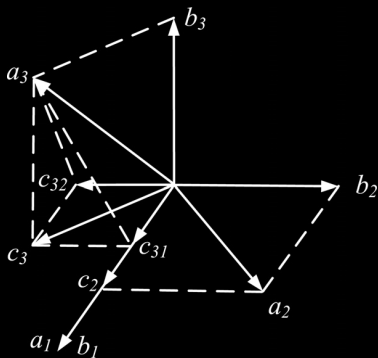
Output: An orthogonal basis $\{\vec{f}_1, \dots, \vec{f}_m\}$ of U s.t.

$$\text{span}\{\vec{f}_1, \vec{f}_2, \dots, \vec{f}_j\} = \text{span}\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_j\}$$

for all $j = 1, \dots, m$.

$$\text{span}\{\vec{a}_1, \vec{a}_2, \vec{a}_3\} = \text{span}\{\vec{b}_1, \vec{b}_2, \vec{b}_3\}$$

basis \rightarrow orthogonal basis



Problem

Let

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

and let $U = \text{span}\{\vec{x}_1, \vec{x}_2, \vec{x}_3\}$. We use the Gram-Schmidt Orthogonalization Algorithm to construct an orthogonal basis B of U .

Proof.

First $\vec{f}_1 = \vec{x}_1$. Next,

$$\vec{f}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} - \frac{2}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Finally,

$$\vec{f}_3 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} - \frac{0}{1} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1 \\ -1/2 \\ 0 \end{bmatrix}.$$


Proof. (continued)

Therefore,

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1/2 \\ 1 \\ -1/2 \\ 0 \end{bmatrix} \right\}$$

is an orthogonal basis of U . However, it is sometimes more convenient to deal with vectors having integer entries, in which case we take

$$B = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ -1 \\ 0 \end{bmatrix} \right\}.$$

(Orthogonality of the set is not affected by multiplying vectors in the set by nonzero scalars.) 

Copyright

Orthogonal Bases

The Orthogonal Complement U^\perp

Definition of Orthogonal Projection

The Projection Theorem and its Implications

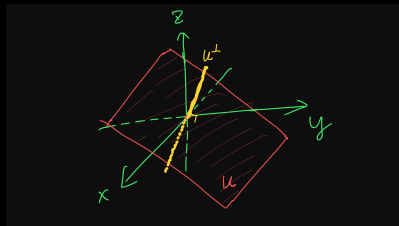
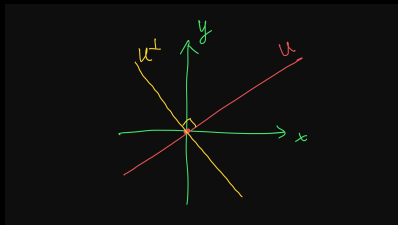
Projection as a Linear Transformation

The Orthogonal Complement U^\perp

Definition

Let U be a subspace of \mathbb{R}^n . The **orthogonal complement** of U , called **U^\perp** , is denoted U^\perp and is defined as

$$U^\perp = \{\vec{x} \in \mathbb{R}^n \mid \vec{x} \cdot \vec{y} = 0 \text{ for all } \vec{y} \in U\}.$$



Example

Let $U = \text{span} \left\{ \begin{bmatrix} -2 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix} \right\}$, and suppose $\vec{v} = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \in U^\perp$. Then

$$-2a + 3b + c = 0 \quad \text{and} \quad 5a - b + 2c = 0.$$

This system of two equations in three variables has solution

$$\vec{v} = \begin{bmatrix} -7 \\ -9 \\ 13 \end{bmatrix} t, \quad \forall t \in \mathbb{R},$$

which is nothing but a line passing through origin and perpendicular with the plane U .

Theorem (Properties of the Orthogonal Complement)

Let U be a subspace of \mathbb{R}^n .

1. U^\perp is a subspace of \mathbb{R}^n .
2. $\{\vec{0}\}^\perp = \mathbb{R}^n$ and $(\mathbb{R}^n)^\perp = \{\vec{0}\}$.
3. If $U = \text{span}\{\vec{y}_1, \vec{y}_2, \dots, \vec{y}_m\}$, then

$$U^\perp = \{\vec{x} \in \mathbb{R}^n \mid \vec{x} \cdot \vec{y}_j = 0 \text{ for } j = 1, 2, \dots, m\}.$$

Proof.

1. This is a standard subspace proof and is left as an exercise.
2. Here, $\vec{0}$ is the zero vector of \mathbb{R}^n . Since $\vec{x} \cdot \vec{0} = 0$ for all $\vec{x} \in \mathbb{R}^n$, $\mathbb{R}^n \subseteq \{\vec{0}\}^\perp$. Since $\{\vec{0}\}^\perp \subseteq \mathbb{R}^n$, the equality follows, i.e., $\{\vec{0}\}^\perp = \mathbb{R}^n$.

Again, since $\vec{x} \cdot \vec{0} = 0$ for all $\vec{x} \in \mathbb{R}^n$, $\vec{0} \in (\mathbb{R}^n)^\perp$, so $\{\vec{0}\} \subseteq (\mathbb{R}^n)^\perp$.

Suppose $\vec{x} \in \mathbb{R}^n$, $\vec{x} \neq \vec{0}$. Since $\vec{x} \cdot \vec{x} = \|\vec{x}\|^2$ and $\vec{x} \neq \vec{0}$, $\vec{x} \cdot \vec{x} \neq 0$, so $\vec{x} \notin (\mathbb{R}^n)^\perp$. Therefore, $\{\vec{0}\}^c \subseteq ((\mathbb{R}^n)^\perp)^c$, or equivalently, $(\mathbb{R}^n)^\perp \subseteq \{\vec{0}\}$.

Thus $(\mathbb{R}^n)^\perp = \{\vec{0}\}$.

Proof. (continued)

3. Let $X = \{\vec{x} \in \mathbb{R}^n \mid \vec{x} \cdot \vec{y}_j = 0 \text{ for } j = 1, 2, \dots, m\}$.

“ $U^\perp \subseteq X$ ”: Suppose that $\vec{v} \in U^\perp$. Then \vec{v} is orthogonal to every vector in U ; in particular, $\vec{v} \cdot \vec{y}_j = 0$ for $j = 1, 2, \dots, m$ since each such \vec{y}_j is in U . Therefore, $\vec{v} \in X$. This proves that $U^\perp \subseteq X$.

“ $X \subseteq U^\perp$ ”: Now suppose that $\vec{v} \in X$ and $\vec{u} \in U$. Then $\vec{u} = a_1\vec{y}_1 + a_2\vec{y}_2 + \dots + a_m\vec{y}_m$ for some $a_1, a_2, \dots, a_m \in \mathbb{R}$, and so

$$\begin{aligned}\vec{v} \cdot \vec{u} &= \vec{v} \cdot (a_1\vec{y}_1 + a_2\vec{y}_2 + \dots + a_m\vec{y}_m) \\ &= \vec{v} \cdot (a_1\vec{y}_1) + \vec{v} \cdot (a_2\vec{y}_2) + \dots + \vec{v} \cdot (a_m\vec{y}_m) \\ &= a_1(\vec{v} \cdot \vec{y}_1) + a_2(\vec{v} \cdot \vec{y}_2) + \dots + a_m(\vec{v} \cdot \vec{y}_m).\end{aligned}$$

Since $\vec{v} \in X$, $\vec{v} \cdot \vec{y}_j = 0$ for all j , $1 \leq j \leq m$. Therefore, $\vec{v} \cdot \vec{u} = 0$, and thus $X \subseteq U^\perp$.

Finally, since $U^\perp \subseteq X$ and $X \subseteq U^\perp$, we see that $U^\perp = X$. ■

Problem

Let

$$U = \text{span} \left\{ \begin{bmatrix} 0 \\ -1 \\ 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \\ 4 \end{bmatrix} \right\}.$$

Find U^\perp by finding a basis of U^\perp .

Solution

$$U^\perp = \left\{ \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \in \mathbb{R}^4 \mid \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \cdot \begin{bmatrix} 0 \\ -1 \\ 3 \\ 2 \end{bmatrix} = 0 \quad \text{and} \quad \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 1 \\ 0 \\ 4 \end{bmatrix} = 0 \right\}.$$

This leads to the system of two equation in four variables

$$\begin{aligned} -b + 3c + 2d &= 0 \\ 2a + b + 4d &= 0 \end{aligned}$$

Solution (continued)

$$\mathbf{A} = \left[\begin{array}{cccc|c} 0 & -1 & 3 & 2 & 0 \\ 2 & 1 & 0 & 4 & 0 \end{array} \right] \rightarrow \cdots \rightarrow \left[\begin{array}{cccc|c} 1 & 0 & 3/2 & 3 & 0 \\ 0 & 1 & -3 & -2 & 0 \end{array} \right]$$

Therefore,

$$U^\perp = \left\{ \left[\begin{array}{c} -\frac{3}{2}s - 3t \\ 3s + 2t \\ s \\ t \end{array} \right] \in \mathbb{R}^4 \mid s, t \in \mathbb{R} \right\} = \text{span} \left\{ \left[\begin{array}{c} -\frac{3}{2} \\ 3 \\ 1 \\ 0 \end{array} \right], \left[\begin{array}{c} 3 \\ 2 \\ 0 \\ 1 \end{array} \right] \right\}.$$

Since the set $B = \left\{ \left[\begin{array}{c} -\frac{3}{2} \\ 3 \\ 1 \\ 0 \end{array} \right], \left[\begin{array}{c} 3 \\ 2 \\ 0 \\ 1 \end{array} \right] \right\}$ is independent and spans U^\perp , B is a

basis of U^\perp . ■

Remark

Notice that $U^\perp = \text{null}(\mathbf{A})$, where \mathbf{A} is the matrix whose rows are a spanning subset of U .

Copyright

Orthogonal Bases

The Orthogonal Complement U^\perp

Definition of Orthogonal Projection

The Projection Theorem and its Implications

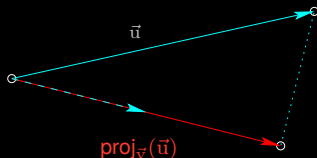
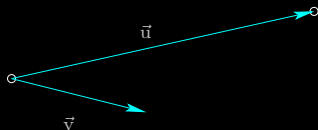
Projection as a Linear Transformation

Definition of Orthogonal Projection

Theorem (Projection Formula)

Suppose \vec{u} and \vec{v} are vectors in \mathbb{R}^3 , $\vec{v} \neq \vec{0}$. Then the **projection of \vec{u} on \vec{v}** , denoted as $\text{proj}_{\vec{v}}(\vec{u})$, is equal to

$$\text{proj}_{\vec{v}}(\vec{u}) = \left(\frac{\vec{u} \cdot \vec{v}}{\|\vec{v}\|^2} \right) \vec{v}.$$



Proof.

Let $\vec{p} = \text{proj}_{\vec{v}}(\vec{u})$; then \vec{p} is parallel to \vec{v} , so $\vec{p} = t\vec{v}$ for some $t \in \mathbb{R}$, and $\vec{u} - \vec{p} = \vec{u} - t\vec{v}$ is orthogonal to \vec{v} , so

$$\begin{aligned}(\vec{u} - t\vec{v}) \cdot \vec{v} &= 0 \\ \vec{u} \cdot \vec{v} - t\vec{v} \cdot \vec{v} &= 0 \\ \vec{u} \cdot \vec{v} &= t\|\vec{v}\|^2\end{aligned}$$

Since $\vec{v} \neq \vec{0}$,

$$t = \frac{\vec{u} \cdot \vec{v}}{\|\vec{v}\|^2}.$$

Therefore,

$$\vec{p} = t\vec{v} = \left(\frac{\vec{u} \cdot \vec{v}}{\|\vec{v}\|^2} \right) \vec{v}.$$



Remark

Note that

- ▶ $\{\vec{v}\}$ is an orthogonal basis of the subspace U of \mathbb{R}^3 consisting of the line through the origin parallel to \vec{v} .
- ▶ $\vec{u} - \vec{p} \in U^\perp$ (since $(\vec{u} - \vec{p}) \cdot \vec{v} = 0$).

Example (Generalizing to \mathbb{R}^n)

Suppose U is a subspace of \mathbb{R}^n , $\vec{x} \in \mathbb{R}^n$, and that $\{\vec{f}_1, \vec{f}_2, \dots, \vec{f}_m\}$ and $\{\vec{g}_1, \vec{g}_2, \dots, \vec{g}_m\}$ are orthogonal bases of U . Define

$$\begin{aligned}\vec{p}_f &= \left(\frac{\vec{x} \cdot \vec{f}_1}{\|\vec{f}_1\|^2} \right) \vec{f}_1 + \left(\frac{\vec{x} \cdot \vec{f}_2}{\|\vec{f}_2\|^2} \right) \vec{f}_2 + \dots + \left(\frac{\vec{x} \cdot \vec{f}_m}{\|\vec{f}_m\|^2} \right) \vec{f}_m \quad \text{and} \\ \vec{p}_g &= \left(\frac{\vec{x} \cdot \vec{g}_1}{\|\vec{g}_1\|^2} \right) \vec{g}_1 + \left(\frac{\vec{x} \cdot \vec{g}_2}{\|\vec{g}_2\|^2} \right) \vec{g}_2 + \dots + \left(\frac{\vec{x} \cdot \vec{g}_m}{\|\vec{g}_m\|^2} \right) \vec{g}_m.\end{aligned}$$

Then $\vec{p}_f, \vec{p}_g \in U$ (since they are linear combinations of vectors of U) and $\vec{x} - \vec{p}_f, \vec{x} - \vec{p}_g \in U^\perp$ (by the Orthogonal Lemma). This implies that $\vec{p}_f - \vec{p}_g \in U$, and $(\vec{x} - \vec{p}_g) - (\vec{x} - \vec{p}_f) \in U^\perp$. However,

$$(\vec{x} - \vec{p}_g) - (\vec{x} - \vec{p}_f) = \vec{p}_f - \vec{p}_g,$$

and thus $\vec{p}_f - \vec{p}_g$ is in both U and U^\perp . This is possible if and only if $\vec{p}_f - \vec{p}_g = \vec{0}$, i.e., $\vec{p}_f = \vec{p}_g$. **This means that the computation of \vec{p}_f and \vec{p}_g does not depend on which orthogonal basis is used.**

Definition

Let $\{\vec{f}_1, \vec{f}_2, \dots, \vec{f}_m\}$ be an orthogonal basis for a subspace U of \mathbb{R}^n , and let $\vec{x} \in \mathbb{R}^n$. The **projection of \vec{x} on U** is defined as

$$\text{proj}_U(\vec{x}) = \left(\frac{\vec{x} \cdot \vec{f}_1}{\|\vec{f}_1\|^2} \right) \vec{f}_1 + \left(\frac{\vec{x} \cdot \vec{f}_2}{\|\vec{f}_2\|^2} \right) \vec{f}_2 + \dots + \left(\frac{\vec{x} \cdot \vec{f}_m}{\|\vec{f}_m\|^2} \right) \vec{f}_m.$$

Remark

1. if $U = \{\vec{0}\}$, then $\text{proj}_{\{\vec{0}\}}(\vec{x}) = \vec{0}$ for any $\vec{x} \in \mathbb{R}^n$;
2. if $\vec{x} \in U$, then $\text{proj}_U(\vec{x})$ is also called the **Fourier Expansion** of \vec{x} .
3. In Orthogonal Lemma

$$\vec{f}_{m+1} = \vec{x} - \underbrace{\left(\frac{\vec{x} \cdot \vec{f}_1}{\|\vec{f}_1\|^2} \vec{f}_1 + \frac{\vec{x} \cdot \vec{f}_2}{\|\vec{f}_2\|^2} \vec{f}_2 + \dots + \frac{\vec{x} \cdot \vec{f}_m}{\|\vec{f}_m\|^2} \vec{f}_m \right)}_{= \text{proj}_U(\vec{x})}.$$

Copyright

Orthogonal Bases

The Orthogonal Complement U^\perp

Definition of Orthogonal Projection

The Projection Theorem and its Implications

Projection as a Linear Transformation

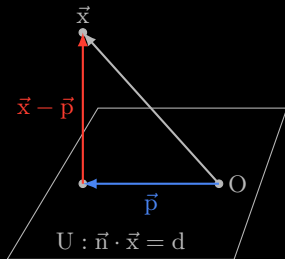
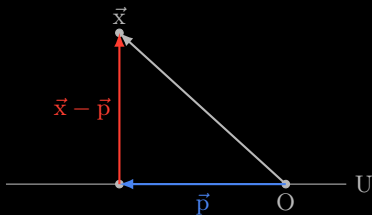
The Projection Theorem and its Implications

Theorem (Projection Theorem)

Let U be a subspace of \mathbb{R}^n , $\vec{x} \in \mathbb{R}^n$, and $\vec{p} = \text{proj}_U(\vec{x})$. Then

1. $\vec{p} \in U$ and $\vec{x} - \vec{p} \in U^\perp$;
2. \vec{p} is the vector in U closest to \vec{x} , meaning that for any $\vec{y} \in U$, $\vec{y} \neq \vec{p}$,

$$\|\vec{x} - \vec{p}\| < \|\vec{x} - \vec{y}\|.$$



Proof.

1. By definition, $\vec{p} \in U$, and by the Orthogonal Lemma, $\vec{x} - \vec{p} \in U^\perp$.
2. Let $\vec{y} \in U$, $\vec{y} \neq \vec{p}$. By the properties of vector addition/subtraction

$$\vec{x} - \vec{y} = (\vec{x} - \vec{p}) + (\vec{p} - \vec{y}).$$

Since $\vec{x} - \vec{p} \in U^\perp$ and $\vec{p} - \vec{y} \in U$,

$$(\vec{x} - \vec{p}) \cdot (\vec{p} - \vec{y}) = 0.$$

Hence, by Pythagoras' Theorem,

$$\|\vec{x} - \vec{y}\|^2 = \|\vec{x} - \vec{p}\|^2 + \|\vec{p} - \vec{y}\|^2.$$

Since $\vec{y} \neq \vec{p}$, $\|\vec{p} - \vec{y}\| > 0$, so

$$\|\vec{x} - \vec{y}\|^2 > \|\vec{x} - \vec{p}\|^2.$$

Taking square roots (since $\|\vec{x} - \vec{y}\|$ and $\|\vec{x} - \vec{p}\|$ are nonnegative),

$$\|\vec{x} - \vec{y}\| > \|\vec{x} - \vec{p}\|.$$



Example

Let

$$\vec{x}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \vec{x}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \vec{x}_3 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad \text{and} \quad \vec{v} = \begin{bmatrix} 4 \\ 3 \\ -2 \\ 5 \end{bmatrix}.$$

We want to find the vector in $U = \text{span}\{\vec{x}_1, \vec{x}_2, \vec{x}_3\}$ closest to \vec{v} .

In a previous example, we used the Gram-Schmidt Orthogonalization Algorithm to construct the orthogonal basis, B , of U :

$$B = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ -1 \\ 0 \end{bmatrix} \right\}.$$

Example (continued)

By the Projection Theorem,

$$\text{proj}_U(\vec{v}) = \frac{2}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + \frac{5}{1} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} + \frac{12}{6} \begin{bmatrix} 1 \\ 2 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \\ -1 \\ 5 \end{bmatrix}$$

is the vector in U closest to \vec{v} .

Problem

Let

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

and let $U = \text{span}\{\vec{x}_1, \vec{x}_2, \vec{x}_3\}$. Find an orthogonal basis of U , and find the vector in U closest to

$$\vec{v} = \begin{bmatrix} 2 \\ 0 \\ -1 \\ 3 \end{bmatrix}.$$

Solution (Outline)

First use the Gram-Schmidt Orthogonalization Algorithm to construct an orthogonal basis of U , and then find the projection of \vec{v} on U .

Solution (continued)

Gram-Schmidt orthogonalization with

$$\begin{aligned}\vec{f}_1 &= \vec{x}_1, \\ \vec{f}_2 &= \vec{x}_2 - \frac{\vec{x}_2 \cdot \vec{f}_1}{\|\vec{f}_1\|^2} \vec{f}_1, \\ \vec{f}_3 &= \vec{x}_3 - \frac{\vec{x}_3 \cdot \vec{f}_1}{\|\vec{f}_1\|^2} \vec{f}_1 - \frac{\vec{x}_3 \cdot \vec{f}_2}{\|\vec{f}_2\|^2} \vec{f}_2\end{aligned}$$

yields an orthogonal basis

$$B = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix} \right\}.$$

Thus the vector in U closest of \vec{v} is

$$\text{proj}_U(\vec{v}) = \frac{1}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + \frac{3}{2} \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ -1 \\ 0 \end{bmatrix}.$$

Problem

Find the point q in the plane $3x + y - 2z = 0$ that is closest to the point $p_0 = (1, 1, 1)$.

Solution

Recall that any plane in \mathbb{R}^3 that contains the origin is a subspace of \mathbb{R}^3 .

1. Find a basis X of the subspace U of \mathbb{R}^3 defined by the equation $3x + y - 2z = 0$.
2. Orthogonalize the basis X to get an orthogonal basis B of U .
3. Find the projection on U of the position vector of the point p_0 .

Solution (continued)

1. $3x + y - 2z = 0$ is a system of one equation in three variables. Putting the augmented matrix in reduced row-echelon form

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

gives general solution $x = \frac{1}{3}s + \frac{2}{3}t$, $y = s$, $z = t$ for any $s, t \in \mathbb{R}$. Then

$$U = \text{span} \left\{ \begin{bmatrix} -\frac{1}{3} \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{2}{3} \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Let

$$X = \left\{ \begin{bmatrix} -1 \\ 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 3 \end{bmatrix} \right\}$$

Then X is linearly independent and $\text{span}(X) = U$, so X is a basis of U .

Solution (continued)

1. Use the Gram-Schmidt Orthogonalization Algorithm to get an orthogonal basis of U:

$$\vec{f}_1 = \begin{bmatrix} -1 \\ 3 \\ 0 \end{bmatrix} \quad \text{and} \quad \vec{f}_2 = \begin{bmatrix} 2 \\ 0 \\ 3 \end{bmatrix} - \frac{-2}{10} \begin{bmatrix} -1 \\ 3 \\ 0 \end{bmatrix} = \frac{1}{5} \begin{bmatrix} 9 \\ 3 \\ 15 \end{bmatrix}.$$

Therefore,

$$B = \left\{ \begin{bmatrix} -1 \\ 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 5 \end{bmatrix} \right\}$$

is an orthogonal basis of U.

Solution (continued)

3. To find the point q on U closest to $p_0 = (1, 1, 1)$, compute

$$\begin{aligned}\text{proj}_U \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} &= \frac{2}{10} \begin{bmatrix} -1 \\ 3 \\ 0 \end{bmatrix} + \frac{9}{35} \begin{bmatrix} 3 \\ 1 \\ 5 \end{bmatrix} \\ &= \frac{1}{7} \begin{bmatrix} 4 \\ 6 \\ 9 \end{bmatrix}.\end{aligned}$$

Therefore, $q = (\frac{4}{7}, \frac{6}{7}, \frac{9}{7})$.



Copyright

Orthogonal Bases

The Orthogonal Complement U^\perp

Definition of Orthogonal Projection

The Projection Theorem and its Implications

Projection as a Linear Transformation

Projection as a Linear Transformation

Definition

Let V and W be vector spaces, and $T : V \rightarrow W$ a linear transformation.

1. The **kernel** of T (sometimes called the null space of T) is defined to be the set

$$\ker(T) = \{\vec{v} \in V \mid T(\vec{v}) = \vec{0}\}.$$

2. The **image** of T is defined to be the set

$$\operatorname{im}(T) = \{T(\vec{v}) \mid \vec{v} \in V\}.$$

Theorem

Let U be a fixed subspace of \mathbb{R}^n , and define $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ by

$$T(\vec{x}) = \operatorname{proj}_U(\vec{x}) \text{ for all } \vec{x} \in \mathbb{R}^n.$$

Then

1. T is a linear operator on \mathbb{R}^n ;
2. $\operatorname{im}(T) = U$ and $\ker(T) = U^\perp$;
3. $\dim(U) + \dim(U^\perp) = n$.

Proof.

If $U = \{\vec{0}\}$, then $U^\perp = \mathbb{R}^n$, so $T(\vec{x}) = \vec{0}$ for all $\vec{x} \in \mathbb{R}^n$. This implies that $T = 0$ (the zero transformation), and the theorem holds.

Now suppose that $U \neq \{\vec{0}\}$. We first prove (3) based on (1) and (2):

3. Since T is a linear transformation – part (1), the **Rank-Nullity Theorem** implies that

$$\dim(\text{im}(T)) + \dim(\ker(T)) = \dim \mathbb{R}^n = n.$$

Applying the result from part (2), we get

$$\dim(U) + \dim(U^\perp) = n.$$

Proof. (continued)

1. Let $B = \{\vec{f}_1, \vec{f}_2, \dots, \vec{f}_m\}$ be an **orthonormal basis** of U . Then by the definition of $\text{proj}_U(\vec{x})$,

$$T(\vec{x}) = (\vec{x} \cdot \vec{f}_1)\vec{f}_1 + (\vec{x} \cdot \vec{f}_2)\vec{f}_2 + \dots + (\vec{x} \cdot \vec{f}_m)\vec{f}_m, \quad (1)$$

(since $\|\vec{f}_i\|^2 = 1$ for each $i = 1, 2, \dots, m$). Let $\vec{x}, \vec{y} \in U$ and $k \in \mathbb{R}$. Then

$$\begin{aligned} T(\vec{x} + \vec{y}) &= ((\vec{x} + \vec{y}) \cdot \vec{f}_1)\vec{f}_1 + ((\vec{x} + \vec{y}) \cdot \vec{f}_2)\vec{f}_2 + \dots + ((\vec{x} + \vec{y}) \cdot \vec{f}_m)\vec{f}_m \\ &= (\vec{x} \cdot \vec{f}_1 + \vec{y} \cdot \vec{f}_1)\vec{f}_1 + (\vec{x} \cdot \vec{f}_2 + \vec{y} \cdot \vec{f}_2)\vec{f}_2 + \\ &\quad \dots + (\vec{x} \cdot \vec{f}_m + \vec{y} \cdot \vec{f}_m)\vec{f}_m \\ &= (\vec{x} \cdot \vec{f}_1)\vec{f}_1 + (\vec{y} \cdot \vec{f}_1)\vec{f}_1 + (\vec{x} \cdot \vec{f}_2)\vec{f}_2 + (\vec{y} \cdot \vec{f}_2)\vec{f}_2 + \\ &\quad \dots + (\vec{x} \cdot \vec{f}_m)\vec{f}_m + (\vec{y} \cdot \vec{f}_m)\vec{f}_m \\ &= [(\vec{x} \cdot \vec{f}_1)\vec{f}_1 + (\vec{x} \cdot \vec{f}_2)\vec{f}_2 + \dots + (\vec{x} \cdot \vec{f}_m)\vec{f}_m] \\ &\quad + [(\vec{y} \cdot \vec{f}_1)\vec{f}_1 + (\vec{y} \cdot \vec{f}_2)\vec{f}_2 + \dots + (\vec{y} \cdot \vec{f}_m)\vec{f}_m] \\ &= T(\vec{x}) + T(\vec{y}). \end{aligned}$$

Thus $\vec{x} + \vec{y} \in U$, so T preserves addition.

Proof. (continued)

1. (continued) Also,

$$\begin{aligned}T(k\vec{x}) &= ((k\vec{x}) \cdot \vec{f}_1)\vec{f}_1 + ((k\vec{x}) \cdot \vec{f}_2)\vec{f}_2 + \cdots + ((k\vec{x}) \cdot \vec{f}_m)\vec{f}_m \\&= (k(\vec{x} \cdot \vec{f}_1))\vec{f}_1 + (k(\vec{x} \cdot \vec{f}_2))\vec{f}_2 + \cdots + (k(\vec{x} \cdot \vec{f}_m))\vec{f}_m \\&= k(\vec{x} \cdot \vec{f}_1)\vec{f}_1 + k(\vec{x} \cdot \vec{f}_2)\vec{f}_2 + \cdots + k(\vec{x} \cdot \vec{f}_m)\vec{f}_m \\&= k[(\vec{x} \cdot \vec{f}_1)\vec{f}_1 + (\vec{x} \cdot \vec{f}_2)\vec{f}_2 + \cdots + (\vec{x} \cdot \vec{f}_m)\vec{f}_m] \\&= kT(\vec{x}).\end{aligned}$$

Thus $k\vec{x} \in U$, so T preserves scalar multiplication.

Therefore, T is a linear transformation.

Proof. (continued)

2. By equation (1), $T(\vec{x}) \in U$ because $T(\vec{x})$ is a linear combination of the elements of $B \subseteq U$, and therefore $\text{im}(T) \subseteq U$. Conversely, suppose that $\vec{x} \in U$. By using Fourier Expansion, $\vec{x} = T(\vec{x})$, and thus $\vec{x} \in \text{im}(T)$. Therefore $U \subseteq \text{im}(T)$. Since $\text{im}(T) \subseteq U$ and $U \subseteq \text{im}(T)$, $\text{im}(T) = U$.

To show that $\ker(T) = U^\perp$, let $\vec{x} \in U^\perp$. Then $\vec{x} \cdot \vec{f}_i = 0$ for each $i = 1, 2, \dots, m$, so $T(\vec{x}) = \vec{0}$, implying $\vec{x} \in \ker(T)$. Thus $U^\perp \subseteq \ker(T)$. Conversely, let $\vec{x} \in \ker(T)$. Then $T(\vec{x}) = \vec{0}$, so $\vec{x} - T(\vec{x}) = \vec{x}$; but, $\vec{x} - T(\vec{x}) \in U^\perp$ (Projection Theorem), so $\vec{x} \in U^\perp$, implying that $\ker(T) \subseteq U^\perp$. Since $U^\perp \subseteq \ker(T)$ and $\ker(T) \subseteq U^\perp$, $\ker(T) = U^\perp$. ■