lyryx with Open Texts

LINEAR ALGEBRA with Applications

Open Edition

ADAPTABLE | ACCESSIBLE | AFFORDABLE

Adapted for

Emory University

Math 221

Linear Algebra

Sections **1** & **2**

Lectured and adapted by

Le Chen

April 15, 2021

le.chen@emory.edu

Course page

http://math.emory.edu/~lchen41/teaching/2021_Spring_Math221

by W. Keith Nicholson

Contents

8.1 Orthogonal Complements and Projections

If $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ is linearly independent in a general vector space, and if $\mathbf{v}_{m+1}$ is not in $\text{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$, then $\{\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{v}_{m+1}\}$ is independent (Lemma 6.4.1). Here is the analog for *orthogonal* sets in $\mathbb{R}^n$.

Lemma 8.1.1: Orthogonal Lemma

Let $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$ be an orthogonal set in $\mathbb{R}^n$. Given $\mathbf{x}$ in $\mathbb{R}^n$, write

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

Then:

1. $\mathbf{f}_{m+1} \cdot \mathbf{f}_k = 0$ *for* $k = 1, 2, \ldots, m$.
2. *If $\mathbf{x}$ is not in $\text{span}\{\mathbf{f}_1, \ldots, \mathbf{f}_m\}$, then $\mathbf{f}_{m+1} \neq \mathbf{0}$ and $\{\mathbf{f}_1, \ldots, \mathbf{f}_m, \mathbf{f}_{m+1}\}$ is an orthogonal set.*

Proof. For convenience, write $t_i = (\mathbf{x}\cdot\mathbf{f}_i)/\|\mathbf{f}_i\|^2$ for each i. Given $1 \leq k \leq m$:

$$\begin{aligned}\mathbf{f}_{m+1}\cdot\mathbf{f}_k &= (\mathbf{x} - t_1\mathbf{f}_1 - \cdots - t_k\mathbf{f}_k - \cdots - t_m\mathbf{f}_m)\cdot\mathbf{f}_k \\ &= \mathbf{x}\cdot\mathbf{f}_k - t_1(\mathbf{f}_1\cdot\mathbf{f}_k) - \cdots - t_k(\mathbf{f}_k\cdot\mathbf{f}_k) - \cdots - t_m(\mathbf{f}_m\cdot\mathbf{f}_k) \\ &= \mathbf{x}\cdot\mathbf{f}_k - t_k\|\mathbf{f}_k\|^2 \\ &= 0\end{aligned}$$

This proves (1), and (2) follows because $\mathbf{f}_{m+1} \neq \mathbf{0}$ if $\mathbf{x}$ is not in $\text{span}\{\mathbf{f}_1, \ldots, \mathbf{f}_m\}$. □

The orthogonal lemma has three important consequences for $\mathbb{R}^n$. The first is an extension for orthogonal sets of the fundamental fact that any independent set is part of a basis (Theorem 6.4.1).

Theorem 8.1.1

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

1. *Every orthogonal subset $\{\mathbf{f}_1, \ldots, \mathbf{f}_m\}$ in U is a subset of an orthogonal basis of U.*
2. *U has an orthogonal basis.*

Proof.

1. If $\text{span}\{\mathbf{f}_1, \ldots, \mathbf{f}_m\} = U$, it is *already* a basis. Otherwise, there exists $\mathbf{x}$ in U outside $\text{span}\{\mathbf{f}_1, \ldots, \mathbf{f}_m\}$. If $\mathbf{f}_{m+1}$ is as given in the orthogonal lemma, then $\mathbf{f}_{m+1}$ is in U and $\{\mathbf{f}_1, \ldots, \mathbf{f}_m, \mathbf{f}_{m+1}\}$ is orthogonal. If $\text{span}\{\mathbf{f}_1, \ldots, \mathbf{f}_m, \mathbf{f}_{m+1}\} = U$, we are done. Otherwise,

the process continues to create larger and larger orthogonal subsets of U. They are all independent by Theorem 5.3.5, so we have a basis when we reach a subset containing dim U vectors.

2. If $U = \{\mathbf{0}\}$, the empty basis is orthogonal. Otherwise, if $\mathbf{f} \neq \mathbf{0}$ is in U, then $\{\mathbf{f}\}$ is orthogonal, so (2) follows from (1). □

We can improve upon (2) of Theorem 8.1.1. In fact, the second consequence of the orthogonal lemma is a procedure by which *any* basis $\{\mathbf{x}_1, \dots, \mathbf{x}_m\}$ of a subspace U of $\mathbb{R}^n$ can be systematically modified to yield an orthogonal basis $\{\mathbf{f}_1, \dots, \mathbf{f}_m\}$ of U. The $\mathbf{f}_i$ are constructed one at a time from the $\mathbf{x}_i$.

To start the process, take $\mathbf{f}_1 = \mathbf{x}_1$. Then $\mathbf{x}_2$ is not in $\text{span}\{\mathbf{f}_1\}$ because $\{\mathbf{x}_1, \mathbf{x}_2\}$ is independent, so take

$$\mathbf{f}_2 = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{f}_1}{\|\mathbf{f}_1\|^2}\mathbf{f}_1$$

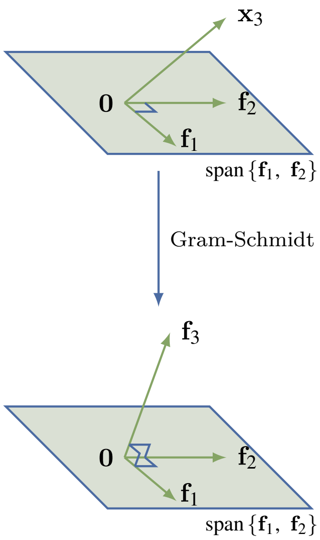
Thus $\{\mathbf{f}_1, \mathbf{f}_2\}$ is orthogonal by Lemma 8.1.1. Moreover, $\text{span}\{\mathbf{f}_1, \mathbf{f}_2\} = \text{span}\{\mathbf{x}_1, \mathbf{x}_2\}$ (verify), so $\mathbf{x}_3$ is not in $\text{span}\{\mathbf{f}_1, \mathbf{f}_2\}$. Hence $\{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3\}$ is orthogonal where

$$\mathbf{f}_3 = \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{f}_1}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{f}_2}{\|\mathbf{f}_2\|^2}\mathbf{f}_2$$

Again, $\text{span}\{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3\} = \text{span}\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$, so $\mathbf{x}_4$ is not in $\text{span}\{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3\}$ and the process continues. At the mth iteration we construct an orthogonal set $\{\mathbf{f}_1, \dots, \mathbf{f}_m\}$ such that

$$\text{span}\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_m\} = \text{span}\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\} = U$$

Hence $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_m\}$ is the desired orthogonal basis of U. The procedure can be summarized as follows.



Theorem 8.1.2: Gram-Schmidt Orthogonalization Algorithm[1]

If $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$ is any basis of a subspace U of $\mathbb{R}^n$, construct $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_m$ in U successively as follows:

$$\begin{aligned}
\mathbf{f}_1 &= \mathbf{x}_1 \\
\mathbf{f}_2 &= \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{f}_1}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 \\
\mathbf{f}_3 &= \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{f}_1}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{f}_2}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 \\
&\vdots \\
\mathbf{f}_k &= \mathbf{x}_k - \frac{\mathbf{x}_k \cdot \mathbf{f}_1}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 - \frac{\mathbf{x}_k \cdot \mathbf{f}_2}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 - \cdots - \frac{\mathbf{x}_k \cdot \mathbf{f}_{k-1}}{\|\mathbf{f}_{k-1}\|^2}\mathbf{f}_{k-1}
\end{aligned}$$

for each $k = 2, 3, \dots, m$. Then

1. *$\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_m\}$ is an orthogonal basis of U.*
2. *$\operatorname{span}\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_k\} = \operatorname{span}\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\}$ for each $k = 1, 2, \dots, m$.*

The process (for $k = 3$) is depicted in the diagrams. Of course, the algorithm converts any basis of $\mathbb{R}^n$ itself into an orthogonal basis.

Example 8.1.1

Find an orthogonal basis of the row space of $A = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 3 & 2 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}$.

Solution. Let $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ denote the rows of A and observe that $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ is linearly independent. Take $\mathbf{f}_1 = \mathbf{x}_1$. The algorithm gives

$$\mathbf{f}_2 = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{f}_1}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 = (3, 2, 0, 1) - \frac{4}{4}(1, 1, -1, -1) = (2, 1, 1, 2)$$
$$\mathbf{f}_3 = \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{f}_1}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{f}_2}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 = \mathbf{x}_3 - \frac{0}{4}\mathbf{f}_1 - \frac{3}{10}\mathbf{f}_2 = \frac{1}{10}(4, -3, 7, -6)$$

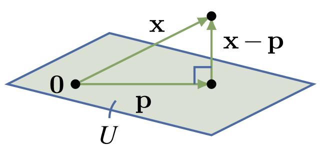
Hence $\{(1, 1, -1, -1), (2, 1, 1, 2), \frac{1}{10}(4, -3, 7, -6)\}$ is the orthogonal basis provided by the algorithm. In hand calculations it may be convenient to eliminate fractions (see the Remark below), so $\{(1, 1, -1, -1), (2, 1, 1, 2), (4, -3, 7, -6)\}$ is also an orthogonal basis for row A.

[1] Erhardt Schmidt (1876–1959) was a German mathematician who studied under the great David Hilbert and later developed the theory of Hilbert spaces. He first described the present algorithm in 1907. Jörgen Pederson Gram (1850–1916) was a Danish actuary.

Remark

Observe that the vector $\frac{\mathbf{x}\cdot\mathbf{f}_i}{\|\mathbf{f}_i\|^2}\mathbf{f}_i$ is unchanged if a nonzero scalar multiple of $\mathbf{f}_i$ is used in place of $\mathbf{f}_i$. Hence, if a newly constructed $\mathbf{f}_i$ is multiplied by a nonzero scalar at some stage of the Gram-Schmidt algorithm, the subsequent **f**s will be unchanged. This is useful in actual calculations.

Projections



Suppose a point $\mathbf{x}$ and a plane U through the origin in $\mathbb{R}^3$ are given, and we want to find the point $\mathbf{p}$ in the plane that is closest to $\mathbf{x}$. Our geometric intuition assures us that such a point $\mathbf{p}$ exists. In fact (see the diagram), $\mathbf{p}$ must be chosen in such a way that $\mathbf{x}-\mathbf{p}$ is *perpendicular* to the plane.

Now we make two observations: first, the plane U is a *subspace of* $\mathbb{R}^3$ (because U contains the origin); and second, that the condition that $\mathbf{x}-\mathbf{p}$ is perpendicular to the plane U means that $\mathbf{x}-\mathbf{p}$ is *orthogonal* to every vector in U. In these terms the whole discussion makes sense in $\mathbb{R}^n$. Furthermore, the orthogonal lemma provides exactly what is needed to find $\mathbf{p}$ in this more general setting.

Definition 8.1 Orthogonal Complement of a Subspace of $\mathbb{R}^n$

*If U is a subspace of $\mathbb{R}^n$, define the **orthogonal complement** $U^\perp$ of U (pronounced "U-perp") by*

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

The following lemma collects some useful properties of the orthogonal complement; the proof of (1) and (2) is left as Exercise 8.1.6.

Lemma 8.1.2

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

1. *$U^\perp$ is a subspace of $\mathbb{R}^n$.*
2. *$\{\mathbf{0}\}^\perp = \mathbb{R}^n$ and $(\mathbb{R}^n)^\perp = \{\mathbf{0}\}$.*
3. *If $U = \text{span}\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\}$, then $U^\perp = \{\mathbf{x} \text{ in } \mathbb{R}^n \mid \mathbf{x}\cdot\mathbf{x}_i = 0 \text{ for } i = 1, 2, \dots, k\}$.*

Proof.

3. Let $U = \text{span}\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\}$; we must show that $U^\perp = \{\mathbf{x} \mid \mathbf{x}\cdot\mathbf{x}_i = 0 \text{ for each } i\}$. If $\mathbf{x}$ is in $U^\perp$ then $\mathbf{x}\cdot\mathbf{x}_i = 0$ for all i because each $\mathbf{x}_i$ is in U. Conversely, suppose that $\mathbf{x}\cdot\mathbf{x}_i = 0$ for all i; we must show that $\mathbf{x}$ is in $U^\perp$, that is, $\mathbf{x}\cdot\mathbf{y} = 0$ for each $\mathbf{y}$ in U. Write $\mathbf{y} = r_1\mathbf{x}_1 + r_2\mathbf{x}_2 + \cdots + r_k\mathbf{x}_k$, where each r_i is in $\mathbb{R}$. Then, using Theorem 5.3.1,

$$\mathbf{x}\cdot\mathbf{y} = r_1(\mathbf{x}\cdot\mathbf{x}_1) + r_2(\mathbf{x}\cdot\mathbf{x}_2) + \cdots + r_k(\mathbf{x}\cdot\mathbf{x}_k) = r_1 0 + r_2 0 + \cdots + r_k 0 = 0$$

as required. □

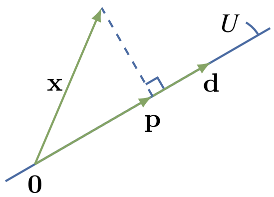
Example 8.1.2

Find $U^\perp$ if $U = \text{span}\{(1,\ -1,\ 2,\ 0),\ (1,\ 0,\ -2,\ 3)\}$ in $\mathbb{R}^4$.

Solution. By Lemma 8.1.2, $\mathbf{x} = (x,\ y,\ z,\ w)$ is in $U^\perp$ if and only if it is orthogonal to both $(1,\ -1,\ 2,\ 0)$ and $(1,\ 0,\ -2,\ 3)$; that is,

$$\begin{aligned} x - y + 2z &= 0 \\ x - 2z + 3w &= 0 \end{aligned}$$

Gaussian elimination gives $U^\perp = \text{span}\{(2,\ 4,\ 1,\ 0),\ (3,\ 3,\ 0,\ -1)\}$.



Now consider vectors $\mathbf{x}$ and $\mathbf{d} \neq \mathbf{0}$ in $\mathbb{R}^3$. The projection $\mathbf{p} = \text{proj}_{\mathbf{d}}\, \mathbf{x}$ of $\mathbf{x}$ on $\mathbf{d}$ was defined in Section 4.2 as in the diagram.

The following formula for $\mathbf{p}$ was derived in Theorem 4.2.4

$$\mathbf{p} = \text{proj}_{\mathbf{d}}\, \mathbf{x} = \left(\frac{\mathbf{x}\cdot\mathbf{d}}{\|\mathbf{d}\|^2}\right)\mathbf{d}$$

where it is shown that $\mathbf{x} - \mathbf{p}$ is orthogonal to $\mathbf{d}$. Now observe that the line $U = \mathbb{R}\mathbf{d} = \{t\mathbf{d} \mid t \in \mathbb{R}\}$ is a subspace of $\mathbb{R}^3$, that $\{\mathbf{d}\}$ is an orthogonal basis of U, and that $\mathbf{p} \in U$ and $\mathbf{x} - \mathbf{p} \in U^\perp$ (by Theorem 4.2.4).

In this form, this makes sense for any vector $\mathbf{x}$ in $\mathbb{R}^n$ and any subspace U of $\mathbb{R}^n$, so we generalize it as follows. If $\{\mathbf{f}_1,\ \mathbf{f}_2,\ \dots,\ \mathbf{f}_m\}$ is an orthogonal basis of U, we define the projection $\mathbf{p}$ of $\mathbf{x}$ on U by the formula

$$\mathbf{p} = \left(\frac{\mathbf{x}\cdot\mathbf{f}_1}{\|\mathbf{f}_1\|^2}\right)\mathbf{f}_1 + \left(\frac{\mathbf{x}\cdot\mathbf{f}_2}{\|\mathbf{f}_2\|^2}\right)\mathbf{f}_2 + \cdots + \left(\frac{\mathbf{x}\cdot\mathbf{f}_m}{\|\mathbf{f}_m\|^2}\right)\mathbf{f}_m \tag{8.1}$$

Then $\mathbf{p} \in U$ and (by the orthogonal lemma) $\mathbf{x} - \mathbf{p} \in U^\perp$, so it looks like we have a generalization of Theorem 4.2.4.

However there is a potential problem: the formula (8.1) for $\mathbf{p}$ must be shown to be independent of the choice of the orthogonal basis $\{\mathbf{f}_1,\ \mathbf{f}_2,\ \dots,\ \mathbf{f}_m\}$. To verify this, suppose that $\{\mathbf{f}'_1,\ \mathbf{f}'_2,\ \dots,\ \mathbf{f}'_m\}$ is another orthogonal basis of U, and write

$$\mathbf{p}' = \left(\frac{\mathbf{x}\cdot\mathbf{f}'_1}{\|\mathbf{f}'_1\|^2}\right)\mathbf{f}'_1 + \left(\frac{\mathbf{x}\cdot\mathbf{f}'_2}{\|\mathbf{f}'_2\|^2}\right)\mathbf{f}'_2 + \cdots + \left(\frac{\mathbf{x}\cdot\mathbf{f}'_m}{\|\mathbf{f}'_m\|^2}\right)\mathbf{f}'_m$$

As before, $\mathbf{p}' \in U$ and $\mathbf{x} - \mathbf{p}' \in U^\perp$, and we must show that $\mathbf{p}' = \mathbf{p}$. To see this, write the vector $\mathbf{p} - \mathbf{p}'$ as follows:

$$\mathbf{p} - \mathbf{p}' = (\mathbf{x} - \mathbf{p}') - (\mathbf{x} - \mathbf{p})$$

This vector is in U (because $\mathbf{p}$ and $\mathbf{p}'$ are in U) and it is in $U^\perp$ (because $\mathbf{x} - \mathbf{p}'$ and $\mathbf{x} - \mathbf{p}$ are in $U^\perp$), and so it must be zero (it is orthogonal to itself!). This means $\mathbf{p}' = \mathbf{p}$ as desired.

Hence, the vector $\mathbf{p}$ in equation (8.1) depends only on $\mathbf{x}$ and the subspace U, and *not* on the choice of orthogonal basis $\{\mathbf{f}_1,\ \dots,\ \mathbf{f}_m\}$ of U used to compute it. Thus, we are entitled to make the following definition:

Definition 8.2 Projection onto a Subspace of $\mathbb{R}^n$

Let U be a subspace of $\mathbb{R}^n$ with orthogonal basis $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$. If $\mathbf{x}$ is in $\mathbb{R}^n$, the vector

$$\text{proj}_U \mathbf{x} = \frac{\mathbf{x}\cdot\mathbf{f}_1}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 + \frac{\mathbf{x}\cdot\mathbf{f}_2}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 + \cdots + \frac{\mathbf{x}\cdot\mathbf{f}_m}{\|\mathbf{f}_m\|^2}\mathbf{f}_m$$

*is called the **orthogonal projection** of $\mathbf{x}$ on U. For the zero subspace $U = \{\mathbf{0}\}$, we define*

$$\text{proj}_{\{\mathbf{0}\}} \mathbf{x} = \mathbf{0}$$

The preceding discussion proves (1) of the following theorem.

Theorem 8.1.3: Projection Theorem

If U is a subspace of $\mathbb{R}^n$ and $\mathbf{x}$ is in $\mathbb{R}^n$, write $\mathbf{p} = \text{proj}_U \mathbf{x}$. Then:

1. *$\mathbf{p}$ is in U and $\mathbf{x} - \mathbf{p}$ is in $U^\perp$.*
2. *$\mathbf{p}$ is the vector in U closest to $\mathbf{x}$ in the sense that*

$$\|\mathbf{x} - \mathbf{p}\| < \|\mathbf{x} - \mathbf{y}\| \quad \text{for all } \mathbf{y} \in U,\ \mathbf{y} \neq \mathbf{p}$$

Proof.

1. This is proved in the preceding discussion (it is clear if $U = \{\mathbf{0}\}$).
2. Write $\mathbf{x} - \mathbf{y} = (\mathbf{x} - \mathbf{p}) + (\mathbf{p} - \mathbf{y})$. Then $\mathbf{p} - \mathbf{y}$ is in U and so is orthogonal to $\mathbf{x} - \mathbf{p}$ by (1). Hence, the Pythagorean theorem gives

$$\|\mathbf{x} - \mathbf{y}\|^2 = \|\mathbf{x} - \mathbf{p}\|^2 + \|\mathbf{p} - \mathbf{y}\|^2 > \|\mathbf{x} - \mathbf{p}\|^2$$

because $\mathbf{p} - \mathbf{y} \neq \mathbf{0}$. This gives (2). □

Example 8.1.3

Let $U = \text{span}\{\mathbf{x}_1, \mathbf{x}_2\}$ in $\mathbb{R}^4$ where $\mathbf{x}_1 = (1, 1, 0, 1)$ and $\mathbf{x}_2 = (0, 1, 1, 2)$. If $\mathbf{x} = (3, -1, 0, 2)$, find the vector in U closest to $\mathbf{x}$ and express $\mathbf{x}$ as the sum of a vector in U and a vector orthogonal to U.

Solution. $\{\mathbf{x}_1, \mathbf{x}_2\}$ is independent but not orthogonal. The Gram-Schmidt process gives an orthogonal basis $\{\mathbf{f}_1, \mathbf{f}_2\}$ of U where $\mathbf{f}_1 = \mathbf{x}_1 = (1, 1, 0, 1)$ and

$$\mathbf{f}_2 = \mathbf{x}_2 - \frac{\mathbf{x}_2\cdot\mathbf{f}_1}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 = \mathbf{x}_2 - \frac{3}{3}\mathbf{f}_1 = (-1, 0, 1, 1)$$

Hence, we can compute the projection using $\{\mathbf{f}_1, \mathbf{f}_2\}$:

$$\mathbf{p} = \text{proj}_U \mathbf{x} = \frac{\mathbf{x}\cdot\mathbf{f}_1}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 + \frac{\mathbf{x}\cdot\mathbf{f}_2}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 = \frac{4}{3}\mathbf{f}_1 + \frac{-1}{3}\mathbf{f}_2 = \frac{1}{3}\left[\begin{array}{cccc} 5 & 4 & -1 & 3 \end{array}\right]$$

Thus, $\mathbf{p}$ is the vector in U closest to $\mathbf{x}$, and $\mathbf{x}-\mathbf{p}=\frac{1}{3}(4,\ -7,\ 1,\ 3)$ is orthogonal to every vector in U. (This can be verified by checking that it is orthogonal to the generators $\mathbf{x}_1$ and $\mathbf{x}_2$ of U.) The required decomposition of $\mathbf{x}$ is thus

$$\mathbf{x}=\mathbf{p}+(\mathbf{x}-\mathbf{p})=\tfrac{1}{3}(5,\ 4,\ -1,\ 3)+\tfrac{1}{3}(4,\ -7,\ 1,\ 3)$$

Example 8.1.4

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

Solution. We write $\mathbb{R}^3$ as rows. The plane is the subspace U whose points $(x,\ y,\ z)$ satisfy $z=2x+y$. Hence

$$U=\{(s,\ t,\ 2s+t)\mid s,\ t \text{ in } \mathbb{R}\}=\text{span}\,\{(0,\ 1,\ 1),\ (1,\ 0,\ 2)\}$$

The Gram-Schmidt process produces an orthogonal basis $\{\mathbf{f}_1,\ \mathbf{f}_2\}$ of U where $\mathbf{f}_1=(0,\ 1,\ 1)$ and $\mathbf{f}_2=(1,\ -1,\ 1)$. Hence, the vector in U closest to $\mathbf{x}=(2,\ -1,\ -3)$ is

$$\text{proj}_U\ \mathbf{x}=\tfrac{\mathbf{x}\cdot\mathbf{f}_1}{\|\mathbf{f}_1\|^2}\mathbf{f}_1+\tfrac{\mathbf{x}\cdot\mathbf{f}_2}{\|\mathbf{f}_2\|^2}\mathbf{f}_2=-2\mathbf{f}_1+0\mathbf{f}_2=(0,\ -2,\ -2)$$

Thus, the point in U closest to $(2,\ -1,\ -3)$ is $(0,\ -2,\ -2)$.

The next theorem shows that projection on a subspace of $\mathbb{R}^n$ is actually a linear operator $\mathbb{R}^n\to\mathbb{R}^n$.

Theorem 8.1.4

Let U be a fixed subspace of $\mathbb{R}^n$. If we define $T:\mathbb{R}^n\to\mathbb{R}^n$ by

$$T(\mathbf{x})=\text{proj}_U\ \mathbf{x}\quad \textit{for all } \mathbf{x} \textit{ in } \mathbb{R}^n$$

1. *T is a linear operator.*
2. *$\text{im}\ T=U$ and $\ker T=U^\perp$.*
3. *$\dim U+\dim U^\perp=n$.*

Proof. If $U=\{\mathbf{0}\}$, then $U^\perp=\mathbb{R}^n$, and so $T(\mathbf{x})=\text{proj}_{\{\mathbf{0}\}}\ \mathbf{x}=\mathbf{0}$ for all $\mathbf{x}$. Thus $T=0$ is the zero (linear) operator, so (1), (2), and (3) hold. Hence assume that $U\neq\{\mathbf{0}\}$.

1. If $\{\mathbf{f}_1,\ \mathbf{f}_2,\ \dots,\ \mathbf{f}_m\}$ is an orthonormal basis of U, then

$$T(\mathbf{x})=(\mathbf{x}\cdot\mathbf{f}_1)\mathbf{f}_1+(\mathbf{x}\cdot\mathbf{f}_2)\mathbf{f}_2+\cdots+(\mathbf{x}\cdot\mathbf{f}_m)\mathbf{f}_m\quad \text{for all } \mathbf{x} \text{ in } \mathbb{R}^n \tag{8.2}$$

by the definition of the projection. Thus T is linear because

$$(\mathbf{x}+\mathbf{y})\cdot\mathbf{f}_i=\mathbf{x}\cdot\mathbf{f}_i+\mathbf{y}\cdot\mathbf{f}_i\quad \text{and}\quad (r\mathbf{x})\cdot\mathbf{f}_i=r(\mathbf{x}\cdot\mathbf{f}_i)\quad \text{for each } i$$

2. We have $\mathrm{im}\, T \subseteq U$ by (8.2) because each $\mathbf{f}_i$ is in U. But if $\mathbf{x}$ is in U, then $\mathbf{x} = T(\mathbf{x})$ by (8.2) and the expansion theorem applied to the space U. This shows that $U \subseteq \mathrm{im}\, T$, so $\mathrm{im}\, T = U$. Now suppose that $\mathbf{x}$ is in $U^\perp$. Then $\mathbf{x} \cdot \mathbf{f}_i = 0$ for each i (again because each $\mathbf{f}_i$ is in U) so $\mathbf{x}$ is in $\ker T$ by (8.2). Hence $U^\perp \subseteq \ker T$. On the other hand, Theorem 8.1.3 shows that $\mathbf{x} - T(\mathbf{x})$ is in $U^\perp$ for all $\mathbf{x}$ in $\mathbb{R}^n$, and it follows that $\ker T \subseteq U^\perp$. Hence $\ker T = U^\perp$, proving (2).

3. This follows from (1), (2), and the dimension theorem (Theorem 7.2.4). □

Exercises for 8.1

Exercise 8.1.1 In each case, use the Gram-Schmidt algorithm to convert the given basis B of V into an orthogonal basis.

a. $V = \mathbb{R}^2$, $B = \{(1, -1), (2, 1)\}$

b. $V = \mathbb{R}^2$, $B = \{(2, 1), (1, 2)\}$

c. $V = \mathbb{R}^3$, $B = \{(1, -1, 1), (1, 0, 1), (1, 1, 2)\}$

d. $V = \mathbb{R}^3$, $B = \{(0, 1, 1), (1, 1, 1), (1, -2, 2)\}$

---

b. $\{(2, 1), \frac{3}{5}(-1, 2)\}$

d. $\{(0, 1, 1), (1, 0, 0), (0, -2, 2)\}$

Exercise 8.1.2 In each case, write $\mathbf{x}$ as the sum of a vector in U and a vector in $U^\perp$.

a. $\mathbf{x} = (1, 5, 7)$, $U = \mathrm{span}\{(1, -2, 3), (-1, 1, 1)\}$

b. $\mathbf{x} = (2, 1, 6)$, $U = \mathrm{span}\{(3, -1, 2), (2, 0, -3)\}$

c. $\mathbf{x} = (3, 1, 5, 9)$,
$U = \mathrm{span}\{(1, 0, 1, 1), (0, 1, -1, 1), (-2, 0, 1, 1)\}$

d. $\mathbf{x} = (2, 0, 1, 6)$,
$U = \mathrm{span}\{(1, 1, 1, 1), (1, 1, -1, -1), (1, -1, 1, -1)\}$

e. $\mathbf{x} = (a, b, c, d)$,
$U = \mathrm{span}\{(1, 0, 0, 0), (0, 1, 0, 0), (0, 0, 1, 0)\}$

f. $\mathbf{x} = (a, b, c, d)$,
$U = \mathrm{span}\{(1, -1, 2, 0), (-1, 1, 1, 1)\}$

---

b. $\mathbf{x} = \frac{1}{182}(271, -221, 1030) + \frac{1}{182}(93, 403, 62)$

d. $\mathbf{x} = \frac{1}{4}(1, 7, 11, 17) + \frac{1}{4}(7, -7, -7, 7)$

f. $\mathbf{x} = \frac{1}{12}(5a - 5b + c - 3d, -5a + 5b - c + 3d, a - b + 11c + 3d, -3a + 3b + 3c + 3d) + \frac{1}{12}(7a + 5b - c + 3d, 5a + 7b + c - 3d, -a + b + c - 3d, 3a - 3b - 3c + 9d)$

Exercise 8.1.3 Let $\mathbf{x} = (1, -2, 1, 6)$ in $\mathbb{R}^4$, and let $U = \mathrm{span}\{(2, 1, 3, -4), (1, 2, 0, 1)\}$.

a. Compute $\mathrm{proj}_U\, \mathbf{x}$.

b. Show that $\{(1, 0, 2, -3), (4, 7, 1, 2)\}$ is another orthogonal basis of U.

c. Use the basis in part (b) to compute $\mathrm{proj}_U\, \mathbf{x}$.

---

a. $\frac{1}{10}(-9, 3, -21, 33) = \frac{3}{10}(-3, 1, -7, 11)$

c. $\frac{1}{70}(-63, 21, -147, 231) = \frac{3}{10}(-3, 1, -7, 11)$

Exercise 8.1.4 In each case, use the Gram-Schmidt algorithm to find an orthogonal basis of the subspace U, and find the vector in U closest to $\mathbf{x}$.

a. $U = \mathrm{span}\{(1, 1, 1), (0, 1, 1)\}$, $\mathbf{x} = (-1, 2, 1)$

b. $U = \mathrm{span}\{(1, -1, 0), (-1, 0, 1)\}$, $\mathbf{x} = (2, 1, 0)$

c. $U = \mathrm{span}\{(1, 0, 1, 0), (1, 1, 1, 0), (1, 1, 0, 0)\}$, $\mathbf{x} = (2, 0, -1, 3)$

d. $U = \text{span}\{(1, -1, 0, 1), (1, 1, 0, 0), (1, 1, 0, 1)\}$, $\mathbf{x} = (2, 0, 3, 1)$

---

b. $\{(1, -1, 0), \frac{1}{2}(-1, -1, 2)\}$; $\text{proj}_U \mathbf{x} = (1, 0, -1)$

d. $\{(1, -1, 0, 1), (1, 1, 0, 0), \frac{1}{3}(-1, 1, 0, 2)\}$; $\text{proj}_U \mathbf{x} = (2, 0, 0, 1)$

Exercise 8.1.5 Let $U = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$, $\mathbf{v}_i$ in $\mathbb{R}^n$, and let A be the $k \times n$ matrix with the $\mathbf{v}_i$ as rows.

a. Show that $U^\perp = \{\mathbf{x} \mid \mathbf{x} \text{ in } \mathbb{R}^n, A\mathbf{x}^T = \mathbf{0}\}$.

b. Use part (a) to find $U^\perp$ if $U = \text{span}\{(1, -1, 2, 1), (1, 0, -1, 1)\}$.

---

b. $U^\perp = \text{span}\{(1, 3, 1, 0), (-1, 0, 0, 1)\}$

Exercise 8.1.6

a. Prove part 1 of Lemma 8.1.2.

b. Prove part 2 of Lemma 8.1.2.

Exercise 8.1.7 Let U be a subspace of $\mathbb{R}^n$. If $\mathbf{x}$ in $\mathbb{R}^n$ can be written in any way at all as $\mathbf{x} = \mathbf{p} + \mathbf{q}$ with $\mathbf{p}$ in U and $\mathbf{q}$ in $U^\perp$, show that necessarily $\mathbf{p} = \text{proj}_U \mathbf{x}$.

Exercise 8.1.8 Let U be a subspace of $\mathbb{R}^n$ and let $\mathbf{x}$ be a vector in $\mathbb{R}^n$. Using Exercise 8.1.7, or otherwise, show that $\mathbf{x}$ is in U if and only if $\mathbf{x} = \text{proj}_U \mathbf{x}$.

Write $\mathbf{p} = \text{proj}_U \mathbf{x}$. Then $\mathbf{p}$ is in U by definition. If $\mathbf{x}$ is U, then $\mathbf{x} - \mathbf{p}$ is in U. But $\mathbf{x} - \mathbf{p}$ is also in $U^\perp$ by Theorem 8.1.3, so $\mathbf{x} - \mathbf{p}$ is in $U \cap U^\perp = \{\mathbf{0}\}$. Thus $\mathbf{x} = \mathbf{p}$.

Exercise 8.1.9 Let U be a subspace of $\mathbb{R}^n$.

a. Show that $U^\perp = \mathbb{R}^n$ if and only if $U = \{\mathbf{0}\}$.

b. Show that $U^\perp = \{\mathbf{0}\}$ if and only if $U = \mathbb{R}^n$.

Exercise 8.1.10 If U is a subspace of $\mathbb{R}^n$, show that $\text{proj}_U \mathbf{x} = \mathbf{x}$ for all $\mathbf{x}$ in U.

Let $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_m\}$ be an orthonormal basis of U. If $\mathbf{x}$ is in U the expansion theorem gives $\mathbf{x} = (\mathbf{x} \cdot \mathbf{f}_1)\mathbf{f}_1 + (\mathbf{x} \cdot \mathbf{f}_2)\mathbf{f}_2 + \cdots + (\mathbf{x} \cdot \mathbf{f}_m)\mathbf{f}_m = \text{proj}_U \mathbf{x}$.

Exercise 8.1.11 If U is a subspace of $\mathbb{R}^n$, show that $\mathbf{x} = \text{proj}_U \mathbf{x} + \text{proj}_{U^\perp} \mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^n$.

Exercise 8.1.12 If $\{\mathbf{f}_1, \dots, \mathbf{f}_n\}$ is an orthogonal basis of $\mathbb{R}^n$ and $U = \text{span}\{\mathbf{f}_1, \dots, \mathbf{f}_m\}$, show that $U^\perp = \text{span}\{\mathbf{f}_{m+1}, \dots, \mathbf{f}_n\}$.

Exercise 8.1.13 If U is a subspace of $\mathbb{R}^n$, show that $U^{\perp\perp} = U$. [*Hint*: Show that $U \subseteq U^{\perp\perp}$, then use Theorem 8.1.4 (3) twice.]

Exercise 8.1.14 If U is a subspace of $\mathbb{R}^n$, show how to find an $n \times n$ matrix A such that $U = \{\mathbf{x} \mid A\mathbf{x} = \mathbf{0}\}$. [*Hint*: Exercise 8.1.13.]

Let $\{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_m\}$ be a basis of $U^\perp$, and let A be the $n \times n$ matrix with rows $\mathbf{y}_1^T, \mathbf{y}_2^T, \dots, \mathbf{y}_m^T, 0, \dots, 0$. Then $A\mathbf{x} = \mathbf{0}$ if and only if $\mathbf{y}_i \cdot \mathbf{x} = 0$ for each $i = 1, 2, \dots, m$; if and only if $\mathbf{x}$ is in $U^{\perp\perp} = U$.

Exercise 8.1.15 Write $\mathbb{R}^n$ as rows. If A is an $n \times n$ matrix, write its null space as $\text{null}\, A = \{\mathbf{x} \text{ in } \mathbb{R}^n \mid A\mathbf{x}^T = \mathbf{0}\}$. Show that:

a) $\text{null}\, A = (\text{row}\, A)^\perp$; b) $\text{null}\, A^T = (\text{col}\, A)^\perp$.

Exercise 8.1.16 If U and W are subspaces, show that $(U + W)^\perp = U^\perp \cap W^\perp$. [See Exercise 5.1.22.]

Exercise 8.1.17 Think of $\mathbb{R}^n$ as consisting of rows.

a. Let E be an $n \times n$ matrix, and let $U = \{\mathbf{x}E \mid \mathbf{x} \text{ in } \mathbb{R}^n\}$. Show that the following are equivalent.

   i. $E^2 = E = E^T$ (E is a **projection matrix**).

   ii. $(\mathbf{x} - \mathbf{x}E) \cdot (\mathbf{y}E) = 0$ for all $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$.

   iii. $\text{proj}_U \mathbf{x} = \mathbf{x}E$ for all $\mathbf{x}$ in $\mathbb{R}^n$. [*Hint*: For (ii) implies (iii): Write $\mathbf{x} = \mathbf{x}E + (\mathbf{x} - \mathbf{x}E)$ and use the uniqueness argument preceding the definition of $\text{proj}_U \mathbf{x}$. For (iii) implies (ii): $\mathbf{x} - \mathbf{x}E$ is in $U^\perp$ for all $\mathbf{x}$ in $\mathbb{R}^n$.]

b. If E is a projection matrix, show that $I - E$ is also a projection matrix.

c. If $EF = 0 = FE$ and E and F are projection matrices, show that $E+F$ is also a projection matrix.

d. If A is $m \times n$ and AA^T is invertible, show that $E = A^T(AA^T)^{-1}A$ is a projection matrix.

---

d. $E^T = A^T[(AA^T)^{-}1]^T(A^T)^T = A^T[(AA^T)^T]^{-1}A = A^T[AA^T]^{-1}A = E$ $E^2 = A^T(AA^T)^{-1}AA^T(AA^T)^{-1}A = A^T(AA^T)^{-1}A = E$

Exercise 8.1.18 Let A be an $n \times n$ matrix of rank r. Show that there is an invertible $n \times n$ matrix U such that UA is a row-echelon matrix with the property that the first r rows are orthogonal. [*Hint*: Let R be the row-echelon form of A, and use the Gram-Schmidt process on the nonzero rows of R from the bottom up. Use Lemma 2.4.1.]

Exercise 8.1.19 Let A be an $(n-1) \times n$ matrix with rows $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{n-1}$ and let A_i denote the $(n-1) \times (n-1)$ matrix obtained from A by deleting column i. Define the vector $\mathbf{y}$ in $\mathbb{R}^n$ by

$$\mathbf{y} = \left[\det A_1 \ -\det A_2 \ \det A_3 \ \cdots \ (-1)^{n+1} \det A_n \right]$$

Show that:

a. $\mathbf{x}_i \cdot \mathbf{y} = 0$ for all $i = 1, 2, \dots, n-1$. [*Hint*: Write $B_i = \left[\begin{array}{c} x_i \\ A \end{array} \right]$ and show that $\det B_i = 0$.]

b. $\mathbf{y} \neq \mathbf{0}$ if and only if $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{n-1}\}$ is linearly independent. [*Hint*: If some $\det A_i \neq 0$, the rows of A_i are linearly independent. Conversely, if the $\mathbf{x}_i$ are independent, consider $A = UR$ where R is in reduced row-echelon form.]

c. If $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{n-1}\}$ is linearly independent, use Theorem 8.1.3(3) to show that all solutions to the system of $n-1$ homogeneous equations

$$A\mathbf{x}^T = \mathbf{0}$$

are given by $t\mathbf{y}$, t a parameter.