LINEAR ALGEBRA I

LECTURE NOTES

Last modified: December 5, 2016

(These Lecture Notes replace neither the Text Book nor the Lectures)

Part 6

- Inner Product, Length, and Orthogonality.
- Orthogonal Bases and Orthonormal Bases of R^n.
- Orthogonal Projection of one Vector onto Another.
- Orthogonal Projection of one Vector onto a Subspace of R^n.
- The Angle between two Vectors.

INNER PRODUCT, LENGTH, and ORTHOGONALITY

Definition: Let $u = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$ and $v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$. The inner product (or a dot product) of u and v is

$$u \cdot v = u^T v = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = u_1v_1 + u_2v_2 + \cdots + u_nv_n.$$

Example: Let $u = \begin{bmatrix} 3 \\ -1 \\ -5 \end{bmatrix}$ and $v = \begin{bmatrix} 6 \\ -2 \\ 3 \end{bmatrix}$. Then

$$u \cdot v = u^T v = [3 \ \ -1 \ \ -5] \begin{bmatrix} 6 \\ -2 \\ 3 \end{bmatrix} = 3 \cdot 6 + (-1) \cdot (-2) + (-5) \cdot 3 = 18 + 2 - 15 = 5,$$

and

$$v \cdot u = v^T u = [6 \ \ -2 \ \ 3] \begin{bmatrix} 3 \\ -1 \\ -5 \end{bmatrix} = 6 \cdot 3 + (-2)(-1) + 3 \cdot (-5) = 18 + 2 - 15 = 5.$$

So, $u \cdot v = v \cdot u$

Let u, v, and w be vectors in R^n and c be a scalar. We have the following properties:

1. $u \cdot v = v \cdot u$

2. $(u + v) \cdot w = u \cdot w + v \cdot w$

3. $(cu) \cdot v = c(u \cdot v) = u \cdot (cv)$

4. $u \cdot u \geq 0$.

5. $u \cdot u = 0 \iff u = 0$.

Definition: Let $v = \begin{bmatrix} v_1 \\ v_2 \\ \cdots \\ v_n \end{bmatrix}$ in R^n. The length (or norm) of v is defined by $||v|| = \sqrt{v \cdot v}$. Thus $||v||^2 = v \cdot v$.

A vector whose length is one unit is called a **unit vector**.

Example: Let $v = \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix}$. We have $||v||^2 = v \cdot v = 2^2 + (-3)^2 + 1^2 = 14$. So, $||v|| = \sqrt{14}$. A unit vector in the direction of v is

$$u = \frac{v}{\|v\|} = \frac{1}{\sqrt{14}} \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix} = \begin{bmatrix} 2/\sqrt{14} \\ -3/\sqrt{14} \\ 1/\sqrt{14} \end{bmatrix}.$$

The process of creating u from v is called **normalizing** v.

Example: Let $W = \text{Span}\left\{ \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} \right\}$. Find a unit vector v which is a basis for W.

Solution: let $u = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$. Then, $||u|| = \sqrt{1^2 + (-1)^2 + 2^2} = \sqrt{6}$.

$$v = \frac{1}{\sqrt{6}} u = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{6} \\ -1/\sqrt{6} \\ 2/\sqrt{6} \end{bmatrix}.$$

Another unit vector: $\quad -v = \begin{bmatrix} -1/\sqrt{6} \\ 1/\sqrt{6} \\ -2/\sqrt{6} \end{bmatrix}.$

Properties of length

1. $\|u\| \geq 0$

2. $\|u\| = 0 \Longleftrightarrow u = 0$

3. $\|ku\| = |k| \; \|u\|$, k is a scalar.

4. $\|u + v\| \leq \|u\| + \|v\|$

Definition: Let u and v be two vectors in R^n. The distance between u and v is the length of the vector $u - v$. That is

$$\text{dist}(u, v) = \|u - v\|.$$

Example: Let $u = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$ and $v = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$ in R^3.

$$\text{dist}(u, v) \;=\; \|u - v\| = \left\| \begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix} \right\| = \sqrt{(-1)^2 + 1^2 + (-2)^2} = \sqrt{6}.$$

Properties of distance

1. $\text{dist}(u, v) \geq 0$

2. $\text{dist}(u, v) = \text{dist}(v, u)$

3. $\text{dist}(u, v) = 0 \Longleftrightarrow u = v$.

4. $\text{dist}(u, v) \leq \text{dist}(u, w) + \text{dist}(w, v)$

Definition: Let u and v be two vectors in R^n. Then u and v are said to be orthogonal to each other ($u \perp v$) if $u \cdot v = 0$.

Note: Zero vector is orthogonal to every vector.

Example (The Phythagorean Theorem): Let u and v in R^n. Show that

$$u \perp v \Longleftrightarrow ||u+v||^2 = ||u||^2 + ||v||^2.$$

Solution: $u \perp v$ if and only if

$$||u+v||^2 = (u+v) \cdot (u+v) = ||u||^2 + ||v||^2 + 2u \cdot v = ||u||^2 + ||v||^2.$$

Example: Let u and v in R^n. Show that

$$u \perp v \Longleftrightarrow ||u+v||^2 = ||u-v||^2.$$

Solution: We have

$$||u+v||^2 = ||u||^2 + ||v||^2 + 2u \cdot v$$

and

$$||u-v||^2 = ||u||^2 + ||v||^2 - 2u \cdot v$$

from which the result follows.

Question: Does $||u+v+z||^2 = ||u||^2 + ||v||^2 + ||z||^2$ imply that $\{u, v, z\}$ is an orthogonal set?

Answer: No. Take $u = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $v = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$ and $z = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$. Then,

$$\|u+v+z\|^2 = \left\| \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} \right\|^2 = 6,$$

$$\|u\|^2 = \|v\|^2 = \|z\|^2 = 2.$$

So $\|u+v+z\|^2 = \|u\|^2 + \|v\|^2 + \|z\|^2$, but $\{u, v, z\}$ is not an orthogonal set.

Definition: Let W be a subspace of R^n.

1. If $z \in R^n$ is orthogonal to every vector in W, then z is said to be orthogonal to W.
2. The set of all z that are orthogonal to W is called the orthogonal complement of W, and denoted by $W^\perp$. That is,

$$W^\perp = \{z \in R^n \mid z \perp W\}.$$

Example: Let $W = \text{Span}\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right\}$. Then, $z = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \perp W$ and $W^\perp = \text{Span}\left\{\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\}$.

Example: Let $W = \text{Span}\left\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right\}$. Then, $W^\perp = \text{Span}\left\{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right\}$.

Example: Let $u_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$, $u_2 = \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}$ and $S = \{u_1,\ u_2\}$. Find $S^\perp$.

Solution:

$$v = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in S^\perp \Longleftrightarrow v \cdot u_1 = 0 \text{ and } v \cdot u_2 = 0.$$

So,

$$x + 2y + z = 0 \quad \text{and } x - y - z = 0.$$

$$\left[\begin{array}{rrr|r} 1 & 2 & 1 & 0 \\ 1 & -1 & -1 & 0 \end{array}\right] \sim \left[\begin{array}{rrr|r} 1 & 2 & 1 & 0 \\ 0 & -3 & -2 & 0 \end{array}\right] \Longrightarrow v = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{1}{3}z \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}.$$

$$S^\perp = \text{Span}\left\{\begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}\right\}$$

Example: Let $b = \begin{bmatrix} 1 \\ -2 \\ 4 \end{bmatrix}$. Find two linearly independent vectors which are orthogonal to b.

Solution: $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \perp \begin{bmatrix} 1 \\ -2 \\ 4 \end{bmatrix} \Longleftrightarrow x_1 - 2x_2 + 4x_3 = 0.$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2x_2 - 4x_3 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -4 \\ 0 \\ 1 \end{bmatrix}.$$

The vectors $u = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$ and $v = \begin{bmatrix} -4 \\ 0 \\ 1 \end{bmatrix}$ are linearly independent and they are both perpendicular to the vector b. In fact, $b \perp \text{Span}\{u, v\}$.

Exercise: Show that if $y \perp u$ and $y \perp v \Longrightarrow y \perp (u + v)$.

Note: $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \perp \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} + \begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix}$ but $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \not\perp \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \not\perp \begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix}$.

Remark: A vector x is in $W^\perp$ if and only if x is orthogonal to every vector in a set that spans W. $W^\perp$ is a subspace of R^n.

Exercise: Show that if x is in both W and $W^\perp$, then $x = 0$. This means that

$$W \cap W^\perp = \{0\}.$$

Solution: If $x \in W$ and $x \in W^\perp$, then $x \perp x$. Thus $x \cdot x = 0 \Longrightarrow x = 0$.

Theorem: Let A be an $m \times n$ matrix. Then,

$$\textbf{(i)}\ (\text{Row}A)^{\perp} = \text{Nul}A \quad \textbf{(ii)}\ (\text{Col}A)^{\perp} = \text{Nul}A^{T}$$

Note that replacing A by A^T in (i) gives $(\text{Row}A^T)^{\perp} = \text{Nul}A^T$, i.e $(\text{Col}A)^{\perp} = \text{Nul}A^T$.

Exampe: Let $A = \begin{bmatrix} 3 & 2 \\ 6 & 4 \end{bmatrix}$. Find $(\text{Col } A)^{\perp}$.

Solution: Let $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in R^2$.

$$\left. \begin{array}{l} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \perp \begin{bmatrix} 3 \\ 6 \end{bmatrix} \Longleftrightarrow 3x_1 + 6x_2 = 0. \\ \\ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \perp \begin{bmatrix} 2 \\ 4 \end{bmatrix} \Longleftrightarrow 2x_1 + 4x_2 = 0. \end{array} \right\} \Longrightarrow x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_2 \begin{bmatrix} -2 \\ 1 \end{bmatrix}.$$

So we have $(\text{Col } A)^{\perp} = \text{Span} \left\{ \begin{bmatrix} -2 \\ 1 \end{bmatrix} \right\}$

- A set S of non-zero vectors is called an **orthogonal set** if each vector in S is orthogonal to other vectors in S, i.e,

  $$S = \{v_1, v_2, ..., v_p\} \text{ is an orthogonal set } \Longleftrightarrow v_i \cdot v_j = 0 \text{ if } i \neq j.$$

- An orthogonal set in which each vector has length 1 is called an **orthonormal set**.
- A basis consisting of orthogonal vectors is called an **orthogonal basis**.
- A basis consisting of orthonormal vectors is called an **orthonormal basis**.

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

is an orthonormal basis for R^3.

$$\mathcal{B}_1 = \left\{ \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}, \begin{bmatrix} -2 \\ 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 5 \\ 4 \\ 1 \end{bmatrix} \right\} \text{ and } \mathcal{B}_2 = \left\{ \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix} \right\}$$

are both orthogonal bases for R^3.

$$\mathcal{B}_1 = \left\{ \begin{bmatrix} -1/3 \\ 2/3 \\ 2/3 \end{bmatrix}, \begin{bmatrix} 2/3 \\ -1/3 \\ 2/3 \end{bmatrix}, \begin{bmatrix} 2/3 \\ 2/3 \\ -1/3 \end{bmatrix} \right\},$$

$$\mathcal{B}_2 = \left\{ \begin{bmatrix} 1/\sqrt{14} \\ -2/\sqrt{14} \\ 3/\sqrt{14} \end{bmatrix}, \begin{bmatrix} -2/\sqrt{12} \\ 2/\sqrt{12} \\ 2/\sqrt{12} \end{bmatrix}, \begin{bmatrix} 5/\sqrt{42} \\ 4/\sqrt{42} \\ 1/\sqrt{42} \end{bmatrix} \right\}.$$

are both orthonormal bases for R^3.

An orthogonal set in R^4: $\left\{ \begin{bmatrix} 2 \\ 2 \\ 4 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 1 \\ 1 \end{bmatrix} \right\}.$

An orthonormal set in R^4: $\left\{ \begin{bmatrix} 2/\sqrt{24} \\ 2/\sqrt{24} \\ 4/\sqrt{24} \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 2/\sqrt{6} \\ -1/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}, \begin{bmatrix} -2/\sqrt{6} \\ 0 \\ 1/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix} \right\}.$

An orthogonal basis for R^4: $\left\{ \begin{bmatrix} 1 \\ 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}.$

An orthonormal basis for R^4: $\left\{ \begin{bmatrix} 1/\sqrt{6} \\ 1/\sqrt{6} \\ 2/\sqrt{6} \\ 0 \end{bmatrix}, \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ -1/\sqrt{3} \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}.$

Definition: An **orthogonal matrix** is a square matrix with orthonormal columns.

Theorem: An $m \times n$ matrix U has orthonormal columns if and only if $U^T U = I$.

Thus, an **orthogonal matrix** is a square invertible matrix U such that $U^{-1} = U^T$.

Example: Let $U = \begin{bmatrix} 1/\sqrt{5} & 0 & 2/\sqrt{5} \\ -2/\sqrt{5} & 0 & 1/\sqrt{5} \\ 0 & 1 & 0 \end{bmatrix}$. U is a square matrix with orthonormal columns. So, it is an orthogonal matrix. The rows of U are orthonormal as well.

$$U^T U = \begin{bmatrix} 1/\sqrt{5} & -2/\sqrt{5} & 0 \\ 0 & 0 & 1 \\ 2/\sqrt{5} & 1/\sqrt{5} & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{5} & 0 & 2/\sqrt{5} \\ -2/\sqrt{5} & 0 & 1/\sqrt{5} \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$U U^T = \begin{bmatrix} 1/\sqrt{5} & 0 & 2/\sqrt{5} \\ -2/\sqrt{5} & 0 & 1/\sqrt{5} \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{5} & -2/\sqrt{5} & 0 \\ 0 & 0 & 1 \\ 2/\sqrt{5} & 1/\sqrt{5} & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$UU^T = U^T U = I \Longrightarrow U^{-1} = U^T.$$

Example: Let $U = \begin{bmatrix} -1/3 & 2/3 & 2/3 \\ 2/3 & -1/3 & 2/3 \\ 2/3 & 2/3 & -1/3 \end{bmatrix}$. It has the following properties:

- $U = U^T$.
- $U^T U = UU = I \Longrightarrow U^{-1} = U$.
- U is an orthogonal matrix.

Example: Let $U = \begin{bmatrix} -1/3 & 2/3 \\ 2/3 & -1/3 \\ 2/3 & 2/3 \end{bmatrix}$. U is a non-square matrix with orthonormal columns. So, it is not an orthogonal matrix.

$$U^T U = \begin{bmatrix} -1/3 & 2/3 & 2/3 \\ 2/3 & -1/3 & 2/3 \end{bmatrix} \begin{bmatrix} -1/3 & 2/3 \\ 2/3 & -1/3 \\ 2/3 & 2/3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I.$$

$$UU^T = \begin{bmatrix} -1/3 & 2/3 \\ 2/3 & -1/3 \\ 2/3 & 2/3 \end{bmatrix} \begin{bmatrix} -1/3 & 2/3 & 2/3 \\ 2/3 & -1/3 & 2/3 \end{bmatrix} = \begin{bmatrix} 5/9 & -4/9 & 2/9 \\ -4/9 & 5/9 & 2/9 \\ 2/9 & 2/9 & 8/9 \end{bmatrix} \neq I.$$

Example: Let $A = \begin{bmatrix} 1 & -2 & 0 \\ 0 & 0 & 1 \\ 2 & 1 & 0 \end{bmatrix}$. The columns of A are orthogonal but not orthonormal. Hence, A is not an orthogonal matrix.

Note that $A^T A = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{bmatrix}$ is a diagonal matrix but not the identity matrix.

Theorem: Let U be an $m \times n$ matrix with orthonormal columns, and let x and y be in R^n. Then,

(a) $||Ux|| = ||x||$

(b) $(Ux) \cdot (Uy) = x \cdot y$

(c) $(Ux) \cdot (Uy) = 0$ if and only if $x \cdot y = 0$.

Proof:

(a) $||Ux||^2 = Ux \cdot Ux = (Ux)^T(Ux) = (x^TU^T)(Ux) = x^TIx = ||x||^2$.

(b) $Ux \cdot Uy = (Ux)^T(Uy) = (x^TU^T)(Uy) = x^TIy = x \cdot y$.

Part (c) follows from part (b).

Theorem: If $S = \{u_1, u_2, ..., u_p\}$ is an orthogonal set of <u>nonzero</u> vectors in R^n, then S is linearly independent.

Proof: Suppose $c_1u_1 + c_2u_2 + \cdots + c_pu_p = 0$ for some scalars $c_1, c_2, ..., c_p$. Then,

$$(c_1u_1 + c_2u_2 + \cdots + c_pu_p) \cdot u_1 = 0 \cdot u_1$$

$$c_1u_1 \cdot u_1 + c_2u_2 \cdot u_1 + \cdots + c_pu_p \cdot u_1 = 0,$$

$$c_1u_1 \cdot u_1 + 0 + \cdots + 0 = 0 \quad (\text{since } u_i \cdot u_j = 0 \text{ if } i \neq j).$$

$$c_1||u_1||^2 = 0 \Longrightarrow c_1 = 0 \quad (\text{since } u_1 \neq 0).$$

Similarly, we can show that $c_2 = 0$,...,$c_p = 0$. Thus, S is linearly independent.

Theorem: Let $\{u_1, u_2, ..., u_n\}$ be an orthogonal basis for R^n.
Then for each $x \in R^n$,

$$x = c_1u_1 + c_2u_2 + \cdots + c_nu_n, \quad \text{where } c_i = \frac{x \cdot u_i}{u_i \cdot u_i}, \quad i = 1, 2, ..., n.$$

Proof: Let $x \in R^n$. Since $\{u_1, u_2, ..., u_n\}$ is an orthogonal basis for R^n, we have

$$x = c_1u_1 + c_2u_2 + \cdots + c_nu_n$$

for some scalars c_i, $1 \leq i \leq n$. For a fixed i, $1 \leq i \leq n$,

$$x \cdot u_i = (c_1u_1 + c_2u_2 + \cdots + c_nu_n) \cdot u_i = c_i(u_i \cdot u_i),$$

which gives

$$c_i = \frac{x \cdot u_i}{u_i \cdot u_i} \qquad (\text{since } u_i \neq 0,\ u_i \cdot u_i \neq 0).$$

Example: Let $u_1 = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}, u_2 = \begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix}, u_3 = \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix}$ and $S = \{u_1,\ u_2,\ u_3\}$.

Express $x = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ as a linear combination of the vectors in S.

Solution: Clearly $u_i \cdot u_j = 0$, when $i \neq j$. Thus, S is an orthogonal set in R^3, and so an orthogonal basis for R^3. Thus using the above theorem we have

$$x = c_1u_1 + c_2u_2 + c_3u_3,$$

where

$$c_1 = \frac{x \cdot u_1}{u_1 \cdot u_1} = \frac{9}{9} = 1, \quad c_2 = \frac{x \cdot u_2}{u_2 \cdot u_2} = \frac{6}{9} = \frac{2}{3}, \quad c_3 = \frac{x \cdot u_3}{u_3 \cdot u_3} = \frac{3}{9} = \frac{1}{3}.$$

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = 1 \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix} + \frac{2}{3} \begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix} + \frac{1}{3} \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix}.$$

Orthogonal Projection of one Vector Onto Another

Let u be a non-zero vector in R^n. We can decompose any vector $y \in R^n$ into a sum

$$y = \hat{y} + z,$$

where

$$\boxed{\hat{y} = \text{proj}_u y = \frac{y \cdot u}{u \cdot u}\, u \quad \text{and} \quad z = y - \hat{y}.}$$

$\hat{y} \in \text{Span}\{u\}$ and $z \perp u$.

$\hat{y}$ is called “the orthogonal projection of y onto u”, and

z is called “the component of y orthogonal to u.”

Example: Let $u = \begin{bmatrix} 7 \\ 1 \end{bmatrix}$. Write the vector $y = \begin{bmatrix} 2 \\ 6 \end{bmatrix}$ as the sum of a vector in Span$\{u\}$ and a vector orthogonal to u.

Solution:

$$\begin{aligned} \hat{y} &= \frac{y \cdot u}{u \cdot u}\, u = \frac{20}{50}\begin{bmatrix} 7 \\ 1 \end{bmatrix} = \begin{bmatrix} 14/5 \\ 2/5 \end{bmatrix} \in \text{Span}\{u\}. \\ z &= y - \hat{y} = \begin{bmatrix} 2 \\ 6 \end{bmatrix} - \begin{bmatrix} 14/5 \\ 2/5 \end{bmatrix} = \begin{bmatrix} -4/5 \\ 28/5 \end{bmatrix} \perp u. \\ y &= \hat{y} + z \Longrightarrow \begin{bmatrix} 2 \\ 6 \end{bmatrix} = \begin{bmatrix} 14/5 \\ 2/5 \end{bmatrix} + \begin{bmatrix} -4/5 \\ 28/5 \end{bmatrix}. \end{aligned}$$

Example: Let $u = \begin{bmatrix} 0 \\ 3 \\ 4 \end{bmatrix}$. Write the vector $y = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ as the sum of a vector in Span$\{u\}$ and a vector orthogonal to u.

Solution:

$$\begin{aligned} \hat{y} &= \frac{y \cdot u}{u \cdot u}\, u = \frac{7}{25}\begin{bmatrix} 0 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 21/25 \\ 28/25 \end{bmatrix}. \\ z &= y - \hat{y} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 \\ 21/25 \\ 28/25 \end{bmatrix} = \begin{bmatrix} 1 \\ 4/25 \\ -3/25 \end{bmatrix}. \end{aligned}$$

$$y = \hat{y} + z \Longrightarrow \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 21/25 \\ 28/25 \end{bmatrix} + \begin{bmatrix} 1 \\ 4/25 \\ -3/25 \end{bmatrix}.$$

Example: Let $x = \begin{bmatrix} -3 \\ 9 \end{bmatrix}$ and $u = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Compute the distance from x to the line through u and origin.

Solution:

$$\hat{x} = \frac{x \cdot u}{u \cdot u}\, u = \frac{15}{5}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}.$$

The distance is

$$\begin{aligned} dist(x, \hat{x}) &= \|x - \hat{x}\| = \left\| \begin{bmatrix} -3 \\ 9 \end{bmatrix} - \begin{bmatrix} 3 \\ 6 \end{bmatrix} \right\| = \left\| \begin{bmatrix} -6 \\ 3 \end{bmatrix} \right\| \\ &= \sqrt{(-6)^2 + 3^2} = \sqrt{45} = 3\sqrt{5}. \end{aligned}$$

Orthogonal Projection of one Vector onto a Subspace of R^n

Definition: Let W be a subspace of R^n with an orthogonal basis $\{u_1, u_2, ..., u_p\}$, and let x be a vector in R^n. The orthogonal projection of x onto W is given by

$$\begin{aligned}\text{proj}_W x &= \text{proj}_{u_1} x + \text{proj}_{u_2} x + \cdots + \text{proj}_{u_p} x \\ &= \frac{x \cdot u_1}{u_1 \cdot u_1} u_1 + \frac{x \cdot u_2}{u_2 \cdot u_2} u_2 + \cdots + \frac{x \cdot u_p}{u_p \cdot u_p} u_p.\end{aligned}$$

Theorem: Let W be a subspace of R^n. Each vector $x \in R^n$ can be written uniquely in the form

$$x = \hat{x} + z,$$

where $\hat{x} = \text{proj}_W x \in W$ and $z = x - \hat{x} \in W^\perp$.

$\hat{x}$ is called "the orthogonal projection of x onto W", and

z is called "the component of x orthogonal to W."

$x = \hat{x} + z$ is called the **orthogonal decomposition** of x.

Example: Let $u_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$, $u_2 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$, $W = \text{Span}\{u_1, u_2\}$, and $x = \begin{bmatrix} 5 \\ 1 \\ 3 \end{bmatrix}$. Write x as the sum of a vector in W and a vector orthogonal to W.

Solution: We first note that $\{u_1, u_2\}$ is an orthogonal basis for W.

$$\hat{x} = \frac{x \cdot u_1}{u_1 \cdot u_1} u_1 + \frac{x \cdot u_2}{u_2 \cdot u_2} u_2 = \frac{4}{2}\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + \frac{12}{6}\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \\ 4 \end{bmatrix} \in W.$$

$$z = x - \hat{x} = \begin{bmatrix} 5 \\ 1 \\ 3 \end{bmatrix} - \begin{bmatrix} 4 \\ 0 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \in W^\perp.$$

Then

$$x = \hat{x} + z \Longrightarrow \begin{bmatrix} 5 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \\ 4 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}.$$

Example: Let $u_1 = \begin{bmatrix} 3 \\ 4 \\ 0 \end{bmatrix}$, $u_2 = \begin{bmatrix} -4 \\ 3 \\ 0 \end{bmatrix}$, $W = \text{span}\{u_1, u_2\}$, and $y = \begin{bmatrix} 6 \\ 3 \\ -2 \end{bmatrix}$.
Write y as the sum of a vector in W and a vector orthogonal to W.

Solution: $\{u_1, u_2\}$ is an orthogonal basis for W.

$$\hat{y} = \frac{y \cdot u_1}{u_1 \cdot u_1} u_1 + \frac{y \cdot u_2}{u_2 \cdot u_2} u_2 = \frac{30}{25} \begin{bmatrix} 3 \\ 4 \\ 0 \end{bmatrix} - \frac{15}{25} \begin{bmatrix} -4 \\ 3 \\ 0 \end{bmatrix} = \begin{bmatrix} 6 \\ 3 \\ 0 \end{bmatrix} \in W.$$

$$z = y - \hat{y} = \begin{bmatrix} 6 \\ 3 \\ -2 \end{bmatrix} - \begin{bmatrix} 6 \\ 3 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -2 \end{bmatrix} \in W^{\perp}.$$

Then,

$$y = \hat{y} + z \Longrightarrow \begin{bmatrix} 6 \\ 3 \\ -2 \end{bmatrix} = \begin{bmatrix} 6 \\ 3 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ -2 \end{bmatrix}.$$

Example: Let $u_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, u_2 = \begin{bmatrix} -1 \\ 3 \\ -2 \end{bmatrix}$, $W = \text{Span}\{u_1,\ u_2\}$, and $y = \begin{bmatrix} -1 \\ 4 \\ 3 \end{bmatrix}$.

a) Find the orthogonal projection of y onto W.

b) Write y as the sum of a vector in W and a vector orthogonal to W.

Soln: a) $\{u_1, u_2\}$ is an orthogonal basis for W.

$$\hat{y} = \frac{y \cdot u_1}{u_1 \cdot u_1}\, u_1 + \frac{y \cdot u_2}{u_2 \cdot u_2}\, u_2 = \frac{6}{3}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \frac{7}{14}\begin{bmatrix} -1 \\ 3 \\ -2 \end{bmatrix} = \begin{bmatrix} 3/2 \\ 7/2 \\ 1 \end{bmatrix}.$$

b)

$$\begin{aligned} z &= y - \hat{y} = \begin{bmatrix} -1 \\ 4 \\ 3 \end{bmatrix} - \begin{bmatrix} 3/2 \\ 7/2 \\ 1 \end{bmatrix} = \begin{bmatrix} -5/2 \\ 1/2 \\ 2 \end{bmatrix}. \\ y &= \hat{y} + z, \text{ where } \hat{y} \in W \text{ and } z \in W^{\perp}. \end{aligned}$$

The Best Approximation Theorem: Let W be a subspace of R^n, and let x be a vector in R^n. Let $\hat{x}$ be the orthogonal projection of x onto W. Then, $\hat{x}$ is the closest point in W to x, i.e,

$$\|x - \hat{x}\| < \|x - v\|$$

for all v in W distinct from $\hat{x}$.

The distance from x to W is defined as the distance from x to the nearest point in W. That is

$$\operatorname{dist}(x, W) = \operatorname{dist}(x, \hat{x}) = \|x - \hat{x}\|.$$

Example: Let $u_1 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}$, $u_2 = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}$, $W = \operatorname{Span}\{u_1, u_2\}$, and $x = \begin{bmatrix} 2 \\ 1 \\ 2 \\ 3 \end{bmatrix}$.

a) Show that $\{u_1, u_2\}$ is an orthogonal set.

b) Find the orthogonal projection of x onto W.

c) What is the nearest point in W to x?

d) Find the distance from x to W.

e) Find an orthonormal basis for W.

Solution: a) $u_1 \cdot u_2 = 0$. Thus, $\{u_1, u_2\}$ is an orthogonal set, and so an orthogonal basis for W.

b) $$\hat{x} = \operatorname{proj}_W x = \frac{x \cdot u_1}{u_1 \cdot u_1} u_1 + \frac{x \cdot u_2}{u_2 \cdot u_2} u_2 = \frac{-2}{4} \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix} + \frac{2}{4} \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}$$

c) The nearest point in W to x is $\hat{x} = \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}$.

d) The distance from x to W:

$$\operatorname{dist}(x, W) = \operatorname{dist}(x, \hat{x}) = \|x - \hat{x}\| = \left\| \begin{bmatrix} 2 \\ 1 \\ 2 \\ 3 \end{bmatrix} - \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix} \right\| = \left\| \begin{bmatrix} 2 \\ 2 \\ 2 \\ 2 \end{bmatrix} \right\| = \sqrt{16} = 4.$$

e) An orthonormal basis for W is $\left\{ \begin{bmatrix} 1/2 \\ 1/2 \\ -1/2 \\ -1/2 \end{bmatrix}, \begin{bmatrix} 1/2 \\ -1/2 \\ -1/2 \\ 1/2 \end{bmatrix} \right\}$.

Example: Let $u_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}$, $u_2 = \begin{bmatrix} -1 \\ 3 \\ 1 \\ -2 \end{bmatrix}$, $u_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 1 \end{bmatrix}$, $W = \text{Span}\{u_1,\ u_2,\ u_3\}$, and $x = \begin{bmatrix} 4 \\ 3 \\ 3 \\ -1 \end{bmatrix}$.

(a) Show that $\{u_1,\ u_2,\ u_3\}$ is an orthogonal set.

(b) Find the orthogonal projection of the vector x onto W.

(c) Write x as the sum of a vector in W and a vector orthogonal to W.

(d) Find the distance from x to W.

Solution: a) $u_1 \cdot u_2 = -1 + 3 + 0 - 2 = 0$, $u_1 \cdot u_3 = -1 + 0 + 0 + 1 = 0$ and $u_2 \cdot u_3 = 1 + 0 + 1 - 2 = 0$. Thus, $\{u_1, u_2, u_3\}$ is an orthogonal set, and so an orthogonal basis for W.

b)
$$\begin{aligned}\hat{x} = \text{proj}_W x &= \frac{x \cdot u_1}{u_1 \cdot u_1}\, u_1 + \frac{x \cdot u_2}{u_2 \cdot u_2}\, u_2 + \frac{x \cdot u_3}{u_3 \cdot u_3}\, u_3 \\ &= \frac{6}{3}\begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix} + \frac{10}{15}\begin{bmatrix} -1 \\ 3 \\ 1 \\ -2 \end{bmatrix} + \frac{-2}{3}\begin{bmatrix} -1 \\ 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 0 \\ 0 \end{bmatrix}\end{aligned}$$

c)

$$z = x - \hat{x} = \begin{bmatrix} 4 \\ 3 \\ 3 \\ -1 \end{bmatrix} - \begin{bmatrix} 2 \\ 4 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 3 \\ -1 \end{bmatrix}.$$

and so,

$$x = \hat{x} + z, \text{ where } \hat{x} \in W \text{ and } z \in W^{\perp}.$$

We note that the nearest point in W to x is $\hat{x} = \begin{bmatrix} 2 \\ 4 \\ 0 \\ 0 \end{bmatrix}$.

d) The distance from x to W:

$$\text{dist}(x, W) = \text{dist}(x, \hat{x}) = \|x - \hat{x}\| = \left\| \begin{bmatrix} 4 \\ 3 \\ 3 \\ -1 \end{bmatrix} - \begin{bmatrix} 2 \\ 4 \\ 0 \\ 0 \end{bmatrix} \right\| = \left\| \begin{bmatrix} 2 \\ -1 \\ 3 \\ -1 \end{bmatrix} \right\| = \sqrt{15}.$$

The Angle between two Vectors

The angle between two non-zero vectors u and v is given by

$$\boxed{\cos\theta = \frac{u \cdot v}{\|u\|\ \|v\|}, \quad 0 \le \theta \le \pi}.$$

Proof in R^2: Consider the vectors $u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ and the triangle with sides of length $\|u\|$, $\|v\|$ and $\|u - v\|$. By the law of cosines,

$$\|u - v\|^2 = \|u\|^2 + \|v\|^2 - 2\|u\|\|v\|\cos\theta.$$

Thus,

$$\begin{aligned}
\|u\|\ \|v\|\cos\theta &= \frac{1}{2}\left[\|u\|^2 + \|v\|^2 - \|u - v\|^2\right] \\
&= \frac{1}{2}\left[u_1^2 + u_2^2 + v_1^2 + v_2^2 - (u_1 - v_1)^2 - (u_2 - v_2)^2\right] \\
&= u_1 v_1 + u_2 v_2 = u \cdot v
\end{aligned}$$

The same formula may be used to define the angle between two vectors in R^n.

Example: Find the angle θ between the pairs of vectors in each part:

a) $u = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, v = \begin{bmatrix} 3 \\ 0 \end{bmatrix} : \cos\theta = \frac{u \cdot v}{\|u\|\ \|v\|} = \frac{3}{\sqrt{2}\sqrt{9}} = \frac{1}{\sqrt{2}} \Longrightarrow \theta = \frac{\pi}{4}.$

b) $u = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, v = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} : \cos\theta = \frac{u \cdot v}{\|u\|\ \|v\|} = \frac{1}{\sqrt{4}\sqrt{1}} = \frac{1}{2} \Longrightarrow \theta = \frac{\pi}{3}.$

c) $u = \begin{bmatrix} 1 \\ -1 \\ -1 \\ -1 \end{bmatrix}, v = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} : \cos\theta = \frac{u \cdot v}{\|u\|\ \|v\|} = \frac{-2}{\sqrt{4}\sqrt{4}} = \frac{-1}{2} \Longrightarrow \theta = \frac{2\pi}{3}.$

d) $u = \begin{bmatrix} 4 \\ -2 \\ 1 \\ 0 \end{bmatrix}, v = \begin{bmatrix} 1 \\ 3 \\ 2 \\ 1 \end{bmatrix} : \cos\theta = \frac{u \cdot v}{\|u\|\ \|v\|} = \frac{0}{\sqrt{15}\sqrt{21}} = 0 \Longrightarrow \theta = \frac{\pi}{2}.$