MA1522 Linear Algebra for Computing

Chapter 5: Orthogonality, Projection, and Least Square Solution

5.1 Orthogonality

Discussion

Let $\mathbf{u} = (u_i)$ and $\mathbf{v} = (v_i)$ be vectors in $\mathbb{R}^n$. Recall that

- The inner product of vectors $\mathbf{u} = (u_i)$ and $\mathbf{v} = (v_i)$ is defined to be

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n.$$

- The norm of a vector $\mathbf{u}$ is

$$\|\mathbf{u}\| = \sqrt{\mathbf{u} \cdot \mathbf{u}} = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2}.$$

- The angle between nonzero vectors $\mathbf{u}, \mathbf{v} \neq \mathbf{0}$ is

$$\theta = \cos^{-1}\left(\frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|}\right), \quad 0 \leq \theta \leq \pi.$$

So suppose $\mathbf{u}, \mathbf{v} \neq \mathbf{0}$ and $\mathbf{u} \cdot \mathbf{v} = 0$, what is the angle between $\mathbf{u}$ and $\mathbf{v}$?

Orthogonal

Definition

Two vectors $\mathbf{u}, \mathbf{v}$ in $\mathbb{R}^n$ are orthogonal if

$$\mathbf{u} \cdot \mathbf{v} = 0.$$

Suppose $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ are orthogonal.

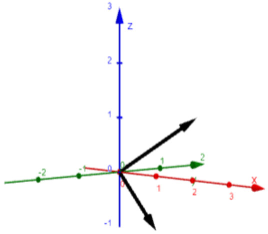
- Case 1: Either $\mathbf{u} = \mathbf{0}$ or $\mathbf{v} = \mathbf{0}$.
- Case 2: Otherwise,

$$\cos(\theta) = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|} = 0$$

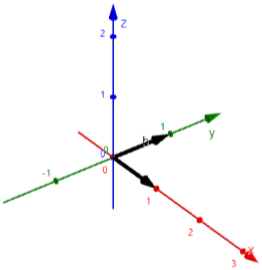
  tells us that $\theta = \frac{\pi}{2}$, that is, $\mathbf{u}$ and $\mathbf{v}$ are perpendicular.

That is, $\mathbf{u}, \mathbf{v}$ are orthogonal if and only if either one of them is the zero vector or they are perpendicular to each other.

Example



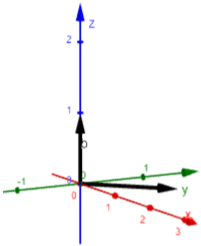
1. $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = (1)(1) + (1)(0) + (1)(-1) = 1 + 0 - 1 = 0.$



2. $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = (1)(0) + (0)(1) + (0)(0) = 0 + 0 + 0 = 0.$

Example

3. $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = (1)(0) + (1)(0) + (0)(1) = 0 + 0 + 0 = 0.$



Question

Suppose $\mathbf{u}$ and $\mathbf{v}$ are orthogonal. Is it true that for any real numbers $s, t \in \mathbb{R}$, $s\mathbf{u}$ and $t\mathbf{v}$ are orthogonal.

Orthogonal and Orthonormal Sets

Definition

A set $S = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k\}$ of vectors in $\mathbb{R}^n$ of vectors is *orthogonal* if $\mathbf{v}_i \cdot \mathbf{v}_j = 0$ for every $i \neq j$, that is, vectors in S are pairwise orthogonal.

The set $S = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k\}$ is *orthonormal* if for all $i, j = 1, ..., k$,

$$\mathbf{v}_i \cdot \mathbf{v}_j = \begin{cases} 0 & \text{if } i \neq j, \\ 1 & \text{if } i = j. \end{cases}$$

That is, S is orthogonal, and all the vectors are unit vectors.

Example

For each of the following sets, decide if it is orthogonal only, orthonormal, or not orthogonal. For the sets that are orthogonal only, normalize it to an orthonormal set, if possible.

1. $S = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$ This set is orthonormal.

2. $S = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\}$ This set is not orthogonal since $\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} = -1 \neq 0$.

Example

3. $S = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \right\}$ This set is orthogonal but not orthonormal. Normalizing, we have $\left\{ \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \right\}$.

4. $S = \left\{ \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix}, \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$ This is an orthonomal set.

5. $S = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}$ This is an orthogonal but not orthonormal set. It cannot be normalized since it contains the zero vector.

Question

(i) Can an orthogonal set can contain the zero vector $\mathbf{0}$?

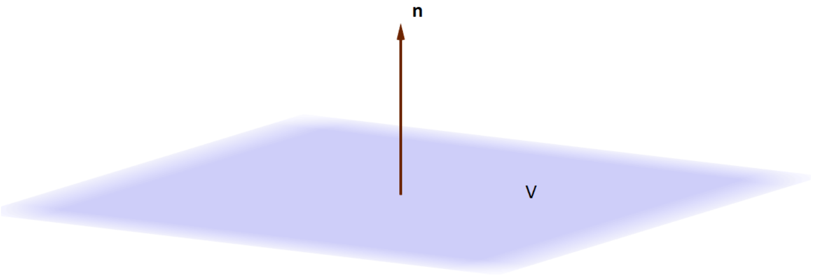
(ii) Can an orthonormal set can contain the zero vector $\mathbf{0}$?

Orthogonal to a Subspace

Definition

Let V be a subspace of $\mathbb{R}^n$. A vector $\mathbf{n}$ is *orthogonal* to V if for every $\mathbf{v}$ in V, $\mathbf{n} \cdot \mathbf{v} = 0$, that is, $\mathbf{n}$ is orthogonal to every vector in V. We will denote it as $\mathbf{n} \perp V$.

$$\mathbf{n} \perp V \quad \Leftrightarrow \quad \mathbf{n} \cdot \mathbf{v} = 0 \text{ for all } \mathbf{v} \in V$$



If $\mathbf{n} \neq \mathbf{0}$.

Remark

The zero vector $\mathbf{0}$ is orthogonal to every subspace, $\mathbf{0} \perp V$.

Example

Let $V = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \;\middle|\; ax + by + cz = 0 \right\}$. Let $\mathbf{n} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$. Given any vector $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ in V,

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = (ax + by + cz) = 0,$$

tells us that $\mathbf{n}$ is orthogonal to V. This shows that we may define a subspace plane by a nonzero orthogonal vector,

$$V = \left\{ \mathbf{v} \in \mathbb{R}^3 \;\middle|\; \mathbf{v} \cdot \mathbf{n} = 0 \right\}$$

for some $\mathbf{n} \in \mathbb{R}^3$, $\mathbf{n} \neq \mathbf{0}$. $\mathbf{n}$ is commonly known as the normal to the plane V.

More generally, a hyperplane V in $\mathbb{R}^n$ can be defined using a nonzero orthogonal vector,

$$V = \left\{ \mathbf{v} \;\middle|\; \mathbf{v} \cdot \mathbf{n} = 0 \right\}, \quad \text{for some } \mathbf{n} \neq \mathbf{0}.$$

Orthogonal to a Subspace

Theorem

Let V be a subspace of $\mathbb{R}^n$ be a subspace and $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ be a spanning set for V, span$(S) = V$. Then a vector $\mathbf{w}$ is orthogonal to V if and only if $\mathbf{w} \cdot \mathbf{u}_i = 0$ for all $i = 1, ..., k$.

Proof.

($\Rightarrow$) Suppose $\mathbf{w}$ is orthogonal to V, $\mathbf{w} \cdot \mathbf{v} = 0$ for all $\mathbf{v} \in V$. Then since $\mathbf{u}_i \in V$ for all i, then $\mathbf{w} \cdot \mathbf{u}_i = 0$ for all i.

($\Leftarrow$) Suppose $\mathbf{w} \cdot \mathbf{u}_i = 0$ for all $i = 1, ..., k$. Now given any $\mathbf{v} \in V$, since S spans V, we can write

$$\mathbf{v} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_k\mathbf{u}_k$$

for some $c_1, c_2, ..., c_k \in \mathbb{R}$. Then

$$\begin{aligned} \mathbf{w} \cdot \mathbf{v} &= \mathbf{w} \cdot (c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_k\mathbf{u}_k) = c_1\mathbf{w} \cdot \mathbf{u}_1 + c_2\mathbf{w} \cdot \mathbf{u}_2 + \cdots + c_k\mathbf{w} \cdot \mathbf{u}_k \\ &= c_1(0) + c_2(0) + \cdots + c_k(0) = 0 \end{aligned}$$

which proves that $\mathbf{w} \cdot \mathbf{v} = 0$ for all $\mathbf{v} \in V$.

$\square$

That is, to check that $\mathbf{w}$ is orthogonal to V, suffice to check that it is orthogonal to every vector in a spanning set.

Example

Let V be a subspace spanned by $S = \left\{ \mathbf{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 2 \end{pmatrix}, \mathbf{u}_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix} \right\}$. A vector $\mathbf{w} = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{pmatrix}$ is orthogonal to V if and only if $\mathbf{w} \cdot \mathbf{u}_1 = w_1 + w_2 + w_3 + 2w_4 = 0$ and $\mathbf{w} \cdot \mathbf{u}_2 = w_2 - w_3 = 0$. That is,

$$\left\{ \begin{array}{ccccccccc} w_1 & + & w_2 & + & w_3 & + & 2w_4 & = & 0 \\ & & w_2 & - & w_3 & & & = & 0 \end{array} \right. \quad \text{or} \quad \begin{pmatrix} 1 & 1 & 1 & 2 \\ 0 & 1 & -1 & 0 \end{pmatrix} \mathbf{w} = \mathbf{0}.$$

Observe that the rows of the coefficient matrix are the vectors in S, $\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 & 2 \\ 0 & 1 & -1 & 0 \end{pmatrix} = \begin{pmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \end{pmatrix}$. Solving the system,

$$\left(\begin{array}{cccc|c} 1 & 1 & 1 & 2 & 0 \\ 0 & 1 & -1 & 0 & 0 \end{array} \right) \xrightarrow{RREF} \left(\begin{array}{cccc|c} 1 & 0 & 2 & 2 & 0 \\ 0 & 1 & -1 & 0 & 0 \end{array} \right)$$

we conclude that $\mathbf{w}$ is orthogonal to V is and only if it is in span $\left\{ \begin{pmatrix} -2 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}$.

Algorithm to check for Orthogonal to a Subspace

Theorem

Let V be a subspace of $\mathbb{R}^n$ and $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ be a spanning set for V. Then $\mathbf{w}$ is orthogonal to V if and only if $\mathbf{w}$ is in the nullspace of $\mathbf{A}^T$, where $\mathbf{A} = \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_k \end{pmatrix}$;

$$\mathbf{w} \perp V \quad \Leftrightarrow \quad \mathbf{w} \in \mathrm{Null}(\mathbf{A}^T)$$

Sketch of Proof.

By previous theorem, $\mathbf{w} \perp V$ if and only if $\mathbf{u}_i^T\mathbf{w} = \mathbf{u}_i \cdot \mathbf{w} = 0$ for all $i = 1, 2, ..., k$. By block multiplication, this is equivalent to

$$\mathbf{A}^T\mathbf{w} = \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_k \end{pmatrix}^T \mathbf{w} = \begin{pmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \vdots \\ \mathbf{u}_k^T \end{pmatrix} \mathbf{w} = \begin{pmatrix} \mathbf{u}_1^T\mathbf{w} \\ \mathbf{u}_2^T\mathbf{w} \\ \vdots \\ \mathbf{u}_k^T\mathbf{w} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$\square$

Example

Let $V = \{(w, x, y, z) \mid w - x + 2y + z = 0\}$. $S = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}$ is a basis for V. Then $\mathbf{w} \perp V$ if and only if $\begin{pmatrix} 1 & 1 & 0 & 0 \\ -2 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix} \mathbf{w} = \mathbf{0}$.

$$\left(\begin{array}{cccc|c} 1 & 1 & 0 & 0 & 0 \\ -2 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 \end{array} \right) \xrightarrow{RREF} \left(\begin{array}{cccc|c} 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & -2 & 0 \end{array} \right)$$

$$\Rightarrow \mathbf{w} \perp V \Leftrightarrow \mathbf{w} \in \text{span} \left\{ \begin{pmatrix} 1 \\ -1 \\ 2 \\ 1 \end{pmatrix} \right\}.$$

Orthogonal Complement

Observe that in all the examples above, the set of vectors that are orthogonal to a subspace V is a subspace. If fact, it is the nullspace of the matrix whose rows are vectors in a spanning set of V.

Definition

Let V be a subspace of $\mathbb{R}^n$. The *orthogonal complement* of V is the set of all vectors that are orthogonal to V, and is denoted as

$$V^{\perp} = \{ \ \mathbf{w} \in \mathbb{R}^n \ \middle| \ \mathbf{w} \cdot \mathbf{v} = 0 \text{ for all } \mathbf{v} \text{ in } V \ \}.$$

Theorem

Let V be a subspace of $\mathbb{R}^n$ and $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ be a spanning set for V. Let $\mathbf{A} = \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_k \end{pmatrix}$. Then the orthogonal complement of V is the nullspace of $\mathbf{A}^T$,

$$V^{\perp} = \text{Null}(\mathbf{A}^T).$$

Challenge

Let $\mathbf{A}$ be a $m \times n$ matrix. Show that the nullspace of $\mathbf{A}$ is the orthogonal complement of the row space of $\mathbf{A}$,

$$\text{Row}(\mathbf{A})^{\perp} = \text{Null}(\mathbf{A}).$$

5.2 Orthogonal and Orthonormal Bases

Examples

Observe that the following orthogonal sets are linearly independent.

1. $S = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$ is linearly independent.
2. $S = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} \right\}$

$$\begin{pmatrix} 1 & 0 & -2 \\ 1 & 1 & 1 \\ 1 & -1 & 1 \end{pmatrix} \xrightarrow{RREF} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Examples

3. $S = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \right\}$

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & -2 \\ 1 & -1 & 1 \end{pmatrix} \xrightarrow{RREF} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

4. $S = \left\{ \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix}, \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$

$$\begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{RREF} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Question

Which of the following statements is/are true?

1. Every orthogonal set is linearly independent.

2. Every orthonormal set is linearly independent.

Orthogonal and Orthonormal Basis

Theorem

Suppose $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ is an orthogonal set of nonzero vectors. Then S is linearly independent.

Proof.

Let $c_1, c_2, ..., c_k$ be coefficients such that $c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_k\mathbf{u}_k = \mathbf{0}$. For each $i = 1, ..., k$,

$$\begin{aligned} \mathbf{0} &= \mathbf{u}_i \cdot \mathbf{0} = \mathbf{u}_i \cdot (c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_k\mathbf{u}_k) = c_1\mathbf{u}_i \cdot \mathbf{u}_1 + \cdots + c_i\mathbf{u}_i \cdot \mathbf{u}_i + \cdots + c_k\mathbf{u}_i \cdot \mathbf{u}_k \\ &= c_1(0) + \cdots + c_i\|\mathbf{u}_i\|^2 + \cdots + c_k(0) \\ &= c_i\|\mathbf{u}_i\|^2, \end{aligned}$$

where the 4th equality follows from the fact that S is orthogonal. Since $\mathbf{u}_i \neq \mathbf{0}$ is nonzero, this means that necessary $c_i = 0$ for all $i = 1, ..., n$. Hence, S is linearly independent. $\square$

Corollary

Every orthonormal set is linearly independent.

Orthogonal and Orthonormal basis

Definition

Let V be a subspace of $\mathbb{R}^n$. A set S is an *orthogonal basis* (resp, *orthonormal basis*) for V if S is a basis of V and S is an orthogonal (resp, orthonormal) set.

Coordinates Relative to an Orthogonal Basis

Let $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ be an orthogonal basis for a subspace V, and let $\mathbf{v} \in V$. Express $\mathbf{v} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_k\mathbf{u}_k$ as a linear combination of vectors in the basis S.

Then since S is orthogonal,

$$\begin{aligned} \mathbf{u}_i \cdot \mathbf{v} &= \mathbf{u}_i \cdot (c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_k\mathbf{u}_k) = c_1\mathbf{u}_i \cdot \mathbf{u}_1 + \cdots + c_i\mathbf{u}_i \cdot \mathbf{u}_i + \cdots + c_k\mathbf{u}_i \cdot \mathbf{u}_k \\ &= c_1(0) + \cdots + c_i\|\mathbf{u}_i\|^2 + \cdots + c_k(0) \\ &= c_i\|\mathbf{u}_i\|^2. \end{aligned}$$

This means that $c_i = \frac{\mathbf{u}_i \cdot \mathbf{v}}{\|\mathbf{u}_i\|^2}$, for all $i = 1, ..., k$. Equivalently, $[\mathbf{v}]_S = \begin{pmatrix} \mathbf{u}_1 \cdot \mathbf{v}/\|\mathbf{u}_1\|^2 \\ \mathbf{u}_2 \cdot \mathbf{v}/\|\mathbf{u}_2\|^2 \\ \vdots \\ \mathbf{u}_k \cdot \mathbf{v}/\|\mathbf{u}_k\|^2 \end{pmatrix}$

Coordinates Relative to an Orthogonal Basis

Theorem

Let $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ *be an orthogonal basis for a subspace* V *of* $\mathbb{R}^n$. *Then for any* $\mathbf{v} \in V$,

$$\mathbf{v} = \left(\frac{\mathbf{v}\cdot\mathbf{u}_1}{\|\mathbf{u}_1\|^2}\right)\mathbf{u}_1 + \left(\frac{\mathbf{v}\cdot\mathbf{u}_2}{\|\mathbf{u}_2\|^2}\right)\mathbf{u}_2 + \cdots + \left(\frac{\mathbf{v}\cdot\mathbf{u}_k}{\|\mathbf{u}_k\|^2}\right)\mathbf{u}_k$$

If further S *is an orthonormal basis, then*

$$\mathbf{v} = (\mathbf{v}\cdot\mathbf{u}_1)\,\mathbf{u}_1 + (\mathbf{v}\cdot\mathbf{u}_2)\,\mathbf{u}_2 + \cdots + (\mathbf{v}\cdot\mathbf{u}_k)\,\mathbf{u}_k.$$

$$\textit{that is } S \textit{ orthogonal, } [\mathbf{v}]_S = \begin{pmatrix} \frac{\mathbf{v}\cdot\mathbf{u}_1}{\|\mathbf{u}_1\|^2} \\ \frac{\mathbf{v}\cdot\mathbf{u}_2}{\|\mathbf{u}_2\|^2} \\ \vdots \\ \frac{\mathbf{v}\cdot\mathbf{u}_k}{\|\mathbf{u}_k\|^2} \end{pmatrix}, \ S \textit{ orthonormal, } [\mathbf{v}]_S = \begin{pmatrix} \mathbf{v}\cdot\mathbf{u}_1 \\ \mathbf{v}\cdot\mathbf{u}_2 \\ \vdots \\ \mathbf{v}\cdot\mathbf{u}_k \end{pmatrix}.$$

Example

Standard basis $E = \{\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n\}$ is an orthonormal basis. For $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{R}^n$,

$$\mathbf{e}_i \cdot \mathbf{v} = v_i.$$

Thus, $[\mathbf{v}]_E = \begin{pmatrix} \mathbf{e}_1 \cdot \mathbf{v} \\ \mathbf{e}_2 \cdot \mathbf{v} \\ \vdots \\ \mathbf{e}_n \cdot \mathbf{v} \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \mathbf{v}.$

Example

$S = \left\{ \mathbf{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \mathbf{u}_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \right\}$ is an orthogonal basis for the hyperplane V in $\mathbb{R}^3$ defined by $z = 0$.

For any $\mathbf{v} = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \in V$, $\mathbf{v} \cdot \mathbf{u}_1 = x + y$, and $\mathbf{v} \cdot \mathbf{u}_2 = x - y$. So,

$$\begin{aligned} \mathbf{v} &= \left(\frac{\mathbf{v} \cdot \mathbf{u}_1}{\|\mathbf{u}_1\|^2} \right) \mathbf{u}_1 + \left(\frac{\mathbf{v} \cdot \mathbf{u}_2}{\|\mathbf{u}_2\|^2} \right) \mathbf{u}_2 \\ &= \frac{x + y}{2} \mathbf{u}_1 + \frac{x - y}{2} \mathbf{u}_2. \end{aligned}$$

That is,

$$[\mathbf{v}]_S = \begin{pmatrix} \frac{x+y}{2} \\ \frac{x-y}{2} \end{pmatrix}.$$

Example

$S = \left\{ \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right\}$ and $V = \mathrm{span}(S)$. S is an orthonormal basis for V. Let $\mathbf{v} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} \in V$. Then

$$\begin{aligned} \mathbf{v} &= \left(\frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} \right) \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \left(\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} \right) \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \\ &= \left(\frac{3}{\sqrt{3}} \right) \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \left(\frac{2}{\sqrt{2}} \right) \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \end{aligned}$$

That is,

$$[\mathbf{v}]_S = \begin{pmatrix} \sqrt{3} \\ \sqrt{2} \end{pmatrix}.$$

Question

Note that this only works if $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ is an orthogonal or orthonormal basis. Example, $S = \left\{ \mathbf{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \mathbf{u}_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\}$, and $\mathbf{w} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$. Then

$$\left(\frac{\mathbf{w} \cdot \mathbf{u}_1}{\|\mathbf{u}_1\|^2} \right) \mathbf{u}_1 + \left(\frac{\mathbf{w} \cdot \mathbf{u}_2}{\|\mathbf{u}_2\|^2} \right) \mathbf{u}_2 = \frac{3}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + 2 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 7 \\ 3 \\ 0 \end{pmatrix} \neq \mathbf{w}.$$

Why is this so?

Challenge

Let V be a subspace of $\mathbb{R}^n$ and S an orthonormal basis of V. Show that for any $\mathbf{u}, \mathbf{v} \in V$,

1. $\mathbf{u} \cdot \mathbf{v} = [\mathbf{u}]_S \cdot [\mathbf{v}]_S$.

2. $\|\mathbf{u} - \mathbf{v}\| = \|[\mathbf{u}]_S - [\mathbf{v}]_S\|$.

Discussion

Let $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ be a set of vectors in $\mathbb{R}^n$. Construct the $n \times k$ matrix $\mathbf{Q} = \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_k \end{pmatrix}$ whose columns are the vectors in S. Consider the product

$$\mathbf{Q}^T\mathbf{Q} = \begin{pmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \vdots \\ \mathbf{u}_k^T \end{pmatrix} \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_k \end{pmatrix} = \begin{pmatrix} \mathbf{u}_1^T\mathbf{u}_1 & \mathbf{u}_1^T\mathbf{u}_2 & \cdots & \mathbf{u}_1^T\mathbf{u}_k \\ \mathbf{u}_2^T\mathbf{u}_1 & \mathbf{u}_2^T\mathbf{u}_2 & \cdots & \mathbf{u}_2^T\mathbf{u}_k \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{u}_k^T\mathbf{u}_1 & \mathbf{u}_k^T\mathbf{u}_2 & \cdots & \mathbf{u}_k^T\mathbf{u}_k \end{pmatrix} = \begin{pmatrix} \mathbf{u}_1\cdot\mathbf{u}_1 & \mathbf{u}_1\cdot\mathbf{u}_2 & \cdots & \mathbf{u}_1\cdot\mathbf{u}_k \\ \mathbf{u}_2\cdot\mathbf{u}_1 & \mathbf{u}_2\cdot\mathbf{u}_2 & \cdots & \mathbf{u}_2\cdot\mathbf{u}_k \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{u}_k\cdot\mathbf{u}_1 & \mathbf{u}_k\cdot\mathbf{u}_2 & \cdots & \mathbf{u}_k\cdot\mathbf{u}_k \end{pmatrix},$$

that is, the (i, j)-entry of the product $\mathbf{Q}^T\mathbf{Q}$ is the inner product $\mathbf{u}_i \cdot \mathbf{u}_j$. Hence,

$$S \text{ is } \begin{cases} \text{orthogonal} \\ \text{orthonormal} \end{cases} \Leftrightarrow \mathbf{Q}^T\mathbf{Q} \text{ is } \begin{cases} \text{a diagonal matrix} \\ \text{the identity matrix} \end{cases}.$$

- In particular, if $k = n$, then $\mathbf{Q}$ is a square matrix and $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}_n$ implies that $\mathbf{Q}^T = \mathbf{Q}^{-1}$.
- That is, S is orthonormal if and only if $\mathbf{Q}^T = \mathbf{Q}^{-1}$.

Orthogonal matrices

Definition

A $n \times n$ square matrix $\mathbf{A}$ is *orthogonal* if $\mathbf{A}^T = \mathbf{A}^{-1}$, equivalently, $\mathbf{A}^T\mathbf{A} = \mathbf{I} = \mathbf{A}\mathbf{A}^T$.

Theorem

Let $\mathbf{A}$ be a square matrix of order n. The following statements are equivalent.

(i) $\mathbf{A}$ *is an orthogonal matrix.*

(ii) *The columns of $\mathbf{A}$ form an orthonormal basis for $\mathbb{R}^n$.*

(iii) *The rows of $\mathbf{A}$ form an orthonormal basis for $\mathbb{R}^n$.*

Remarks

- Note that $\mathbf{A}$ is an orthogonal matrix if and only if the columns/rows form an orthonormal basis for $\mathbb{R}^n$.
- We do not have a name given to matrices whose the rows or columns form an orthogonal basis.
- While some textbooks may define a square matrix whose inverse is the transpose as an orthonormal matrix, this course will follow the majority of the literature and call it an orthogonal matrix instead.

Example

1. $\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$.

$$\mathbf{A}^T\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} \\ 0 & -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

2. $\mathbf{A} = \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 0 & -2/\sqrt{6} \end{pmatrix}$.

$$\mathbf{A}^T\mathbf{A} = \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1/\sqrt{6} & 1/\sqrt{6} & -2/\sqrt{6} \end{pmatrix} \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 0 & -2/\sqrt{6} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Question

Let $\mathbf{A} = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & -1/2 & \sqrt{2}/2 \\ 1/2 & 1/2 & 0 \\ 1/2 & -1/2 & -\sqrt{2}/2 \end{pmatrix}$. Then

$$\mathbf{A}^T\mathbf{A} = \begin{pmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & 1/2 & -1/2 \\ 0 & \sqrt{2}/2 & 0 & -\sqrt{2}/2 \end{pmatrix} \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & -1/2 & \sqrt{2}/2 \\ 1/2 & 1/2 & 0 \\ 1/2 & -1/2 & -\sqrt{2}/2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Is $\mathbf{A}$ an orthogonal matrix?

5.3 Orthogonal Projection

Question

Let $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ be an orthonormal basis for a subspace V of $\mathbb{R}^n$. We have seen that if $\mathbf{v}$ is in V, then

$$\mathbf{v} = (\mathbf{v} \cdot \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v} \cdot \mathbf{u}_2)\mathbf{u}_2 + \cdots + (\mathbf{v} \cdot \mathbf{u}_k)\mathbf{u}_k.$$

Suppose now $\mathbf{w}$ is a vector not in V, $\mathbf{w} \notin \mathbf{V}$. What is

$$\mathbf{w}' = (\mathbf{w} \cdot \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{w} \cdot \mathbf{u}_2)\mathbf{u}_2 + \cdots + (\mathbf{w} \cdot \mathbf{u}_k)\mathbf{u}_k?$$

1. Is it a vector in V?
2. Is it equal to $\mathbf{w}$?
3. Is there a geometrical interpretation of $\mathbf{w}'$?

Example

Let $V = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \;\middle|\; z = 0 \right\}$ and $\mathbf{w} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. $\left\{ \mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}$ is an orthonormal basis for V. Then

$$\mathbf{w}' = (\mathbf{w} \cdot \mathbf{e}_1)\mathbf{e}_1 + (\mathbf{w} \cdot \mathbf{e}_2)\mathbf{e}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

Consider now another orthonormal basis $\left\{ \mathbf{v} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \mathbf{u} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \right\}$ for V. Check that

$$\mathbf{w}' = (\mathbf{w} \cdot \mathbf{v})\mathbf{v} + (\mathbf{w} \cdot \mathbf{u})\mathbf{u} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

It seems that regardless of the choice of orthonormal basis for the result is the same.

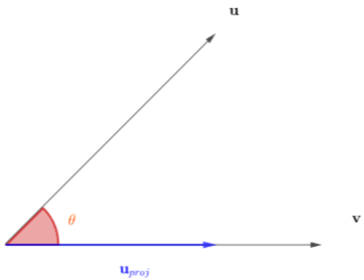
Discussion

Recall that the angle between 2 nonzero vectors **u** and **v** is given by

$$\cos(\theta) = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|}.$$

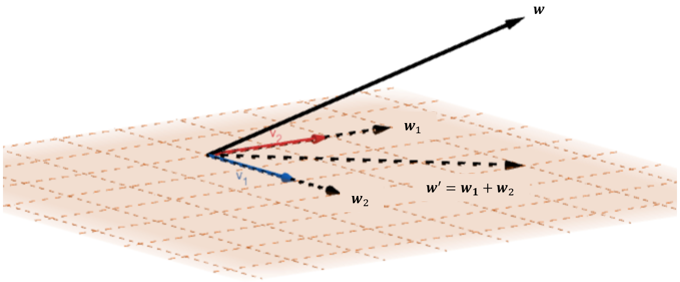
Recall also that the projection of **u** onto the direction of **v** is

$$\|\mathbf{u}\| \cos(\theta) \frac{\mathbf{v}}{\|\mathbf{v}\|} = \|\mathbf{u}\| \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|} \frac{\mathbf{v}}{\|\mathbf{v}\|} = \frac{\mathbf{u} \cdot \mathbf{v}}{\mathbf{v} \cdot \mathbf{v}} \mathbf{v}.$$



Discussion

- Now suppose $S = \{\mathbf{u}_1, \mathbf{u}_2\}$ is an orthogonal basis for a subspace V. Then
- Then $\mathbf{w}' = \frac{\mathbf{w}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{w}\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\mathbf{u}_2$ is the projection of $\mathbf{w}$ onto $\text{span}(S) = V$.
- Moreover, the vector $\mathbf{w}'$ is independent of the choice of orthogonal basis for V. See `https://www.geogebra.org/m/mndafgcp`.



Orthogonal Projection

Theorem (Orthogonal projection theorem)

Let V be a subspace of $\mathbb{R}^n$. Every vector $\mathbf{w}$ in $\mathbb{R}^n$ can be decomposed uniquely as a sum

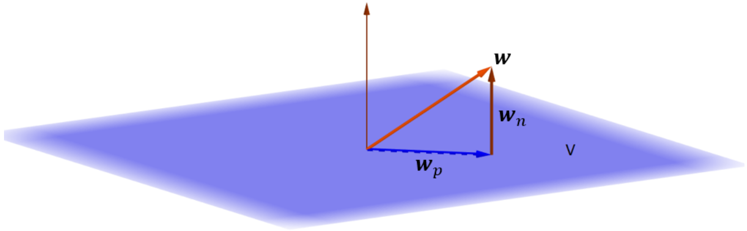
$$\mathbf{w} = \mathbf{w}_p + \mathbf{w}_n,$$

where $\mathbf{w}_n$ is orthogonal to V and $\mathbf{w}_p$ is a vector in V, $\mathbf{w}_n \perp V$, $\mathbf{w}_p \in V$. Moreover, if $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ is an orthogonal basis for V, then

$$\mathbf{w}_p = \frac{\mathbf{w} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{w} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_2 + \cdots + \frac{\mathbf{w} \cdot \mathbf{u}_k}{\mathbf{u}_k \cdot \mathbf{u}_k}\mathbf{u}_k.$$

Definition

Define the vector $\mathbf{w}_p$ in the theorem above as the *orthogonal projection* (or just *projection*) of $\mathbf{w}$ onto the subspace V.



Example

$S = \left\{ \mathbf{v} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \mathbf{u} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \right\}$ is an orthonormal basis subspace V in $\mathbb{R}^3$ defined by $z = 0$. Let $\mathbf{w} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.
Then

$$\mathbf{w}_p = (\mathbf{w} \cdot \mathbf{v})\mathbf{v} + (\mathbf{w} \cdot \mathbf{u})\mathbf{u} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \text{ and } \mathbf{w}_n = \mathbf{w} - \mathbf{w}_p = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

It is clear that $\mathbf{w}_p$ is in V and $\mathbf{w}_n$ is orthogonal to V.

Best Approximation Theorem

Theorem (Best approximation theorem)

Let V be a subspace of $\mathbb{R}^n$ and $\mathbf{w}$ a vector in $\mathbb{R}^n$. Let $\mathbf{w}_p$ be the projection of $\mathbf{w}$ onto V. Then $\mathbf{w}_p$ is vector in V closest to $\mathbf{w}$; that is,

$$\|\mathbf{w} - \mathbf{w}_p\| \leq \|\mathbf{w} - \mathbf{v}\|$$

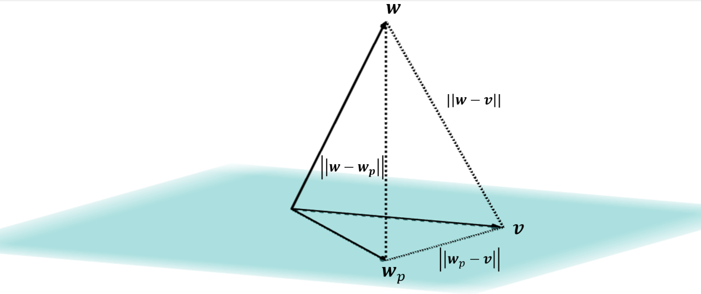
for all $\mathbf{v}$ in V.

Proof.

For any $\mathbf{v}$ in V, $\mathbf{v} - \mathbf{w}_p$ is in V. Thus, $\mathbf{v} - \mathbf{w}_p$ is orthogonal to $\mathbf{w} - \mathbf{w}_p$. Hence, by Pythagorean theorem,

$$\|\mathbf{w} - \mathbf{v}\|^2 = \|\mathbf{w} - \mathbf{w}_p\|^2 + \|\mathbf{v} - \mathbf{w}_p\|^2.$$

Since $\|\mathbf{v} - \mathbf{w}_p\|^2 \geq 0$, the inequality $\|\mathbf{w} - \mathbf{w}_p\| \leq \|\mathbf{w} - \mathbf{v}\|$ is established. $\square$



Question

Let V be a subspace in $\mathbb{R}^n$ and let $\mathbf{w}$ be a vector in V. In the decomposition

$$\mathbf{w} = \mathbf{w}_p + \mathbf{w}_n,$$

where $\mathbf{w}_p$ is the projection of $\mathbf{w}$ onto V and $\mathbf{w}_n$ is orthogonal to V, what is $\mathbf{w}_p$ and $\mathbf{w}_n$?

Challenge

Prove the orthogonal projection theorem.

Theorem (Orthogonal projection theorem)

Let V be a subspace of $\mathbb{R}^n$. Every vector $\mathbf{w}$ in $\mathbb{R}^n$ can be decomposed uniquely as a sum

$$\mathbf{w} = \mathbf{w}_p + \mathbf{w}_n,$$

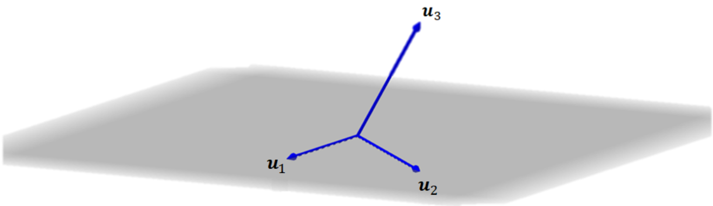
where $\mathbf{w}_n$ is orthogonal to V and $\mathbf{w}_p$ is a vector in V, $\mathbf{w}_n \perp V$, $\mathbf{w}_p \in V$.

Introduction to Gram-Schmidt Process

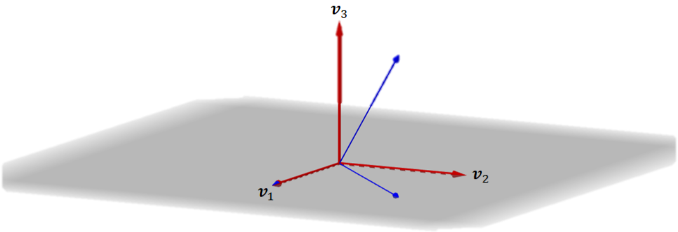
- To compute the projection of a vector $\mathbf{w}$ onto a subspace V, we need to find an orthogonal or orthonormal basis.
- Suppose now S is a basis for a subspace V. The Gram-Schmidt process converts S to an orthonormal basis.
- Given a linear independent set $\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$, we want to find an orthogonal set $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k\}$ such that $\text{span}\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\} = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k\}$.
- Let $\mathbf{v}_1 = \mathbf{u}_1$, then $\text{span}\{\mathbf{u}_1\} = \text{span}\{\mathbf{v}_1\}$.
- If we let $\mathbf{v}_2 = \mathbf{u}_2 - \mathbf{u}_2'$, where $\mathbf{u}_2'$ is the projection of $\mathbf{u}_2$ onto $\text{span}\{\mathbf{v}_1\}$, then $\mathbf{v}_2$ is orthogonal to $\mathbf{v}_1$. Moreover, it should be clear that $\text{span}\{\mathbf{u}_1, \mathbf{u}_2\} = \text{span}\{\mathbf{v}_1, \mathbf{v}_2\}$.
- Continue this process, where we define $\mathbf{v}_{i+1} = \mathbf{u}_{i+1} - \mathbf{u}_{i+1}'$, where $\mathbf{u}_{i+1}'$ is the projection of $\mathbf{u}_{i+1}$ onto $\text{span}\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_i\}$.
- Then $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k\}$ will be an orthogonal set such that $\text{span}\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\} = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k\}$.
- If we normalize $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k\}$, we will have an orthonormal basis.

Gram-Schmidt Process

We will give a visualization of the Gram-Schmidt process in $\mathbb{R}^3$.



$$\Downarrow$$



Gram-Schmidt Process

Let $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ be a linearly independent set. Let

$$\begin{aligned}
\mathbf{v}_1 &= \mathbf{u}_1 \\
\mathbf{v}_2 &= \mathbf{u}_2 - \left(\frac{\mathbf{v}_1 \cdot \mathbf{u}_2}{\|\mathbf{v}_1\|^2}\right)\mathbf{v}_1 \\
\mathbf{v}_3 &= \mathbf{u}_3 - \left(\frac{\mathbf{v}_1 \cdot \mathbf{u}_3}{\|\mathbf{v}_1\|^2}\right)\mathbf{v}_1 - \left(\frac{\mathbf{v}_2 \cdot \mathbf{u}_3}{\|\mathbf{v}_2\|^2}\right)\mathbf{v}_2 \\
&\vdots \\
\mathbf{v}_k &= \mathbf{u}_k - \left(\frac{\mathbf{v}_1 \cdot \mathbf{u}_k}{\|\mathbf{v}_1\|^2}\right)\mathbf{v}_1 - \left(\frac{\mathbf{v}_2 \cdot \mathbf{u}_k}{\|\mathbf{v}_2\|^2}\right)\mathbf{v}_2 - \cdots - \left(\frac{\mathbf{v}_{k-1} \cdot \mathbf{u}_k}{\|\mathbf{v}_{k-1}\|^2}\right)\mathbf{v}_{k-1}.
\end{aligned}$$

Then $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k\}$ is an orthogonal set (of nonzero vectors), and hence,

$$\left\{\frac{\mathbf{v}_1}{\|\mathbf{v}_1\|}, \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|}, ..., \frac{\mathbf{v}_k}{\|\mathbf{v}_k\|}\right\}$$

is an orthonormal set such that $\text{span}\{\mathbf{v}_1, ..., \mathbf{v}_k\} = \text{span}\{\mathbf{u}_k, ..., \mathbf{u}_k\}$.

Example

Convert $S = \left\{ \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \right\}$ into an orthonormal basis for $\mathbb{R}^3$.

$$
\begin{aligned}
\mathbf{v}_1 &= \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \\
\mathbf{v}_2 &= \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \frac{2}{3} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \\
\mathbf{v}_3 &= \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} - \frac{5}{6} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} - \frac{2}{3} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}
\end{aligned}
$$

So $\left\{ \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\}$ is an orthonormal basis for $\mathbb{R}^3$.

Discussion

Let $S = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$ be a set in $\mathbb{R}^4$. After performing the Gram-Schmidt process on S, $\mathbf{v}_4 = \mathbf{0}$, but $\mathbf{v}_3 \neq \mathbf{0}$. What can you conclude?

5.4 QR Factorization

QR Factorization

Suppose now $\mathbf{A}$ is a $m \times n$ matrix with linearly independent columns, i.e. $\text{rank}(\mathbf{A}) = n$. Write

$$\mathbf{A} = \begin{pmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{pmatrix}.$$

Since the set $S = \{\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n\}$ is linearly independent we may apply the Gram-Schmidt process on S to obtain an orthonormal set $\{\mathbf{q}_1, \mathbf{q}_2, ..., \mathbf{q}_n\}$. Set

$$\mathbf{Q} = \begin{pmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_n \end{pmatrix}.$$

Recall that for any $j = 1, 2, ..., n$, $\text{span}\{\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_j\} = \text{span}\{\mathbf{q}_1, \mathbf{q}_2, ..., \mathbf{q}_j\}$. In particular, $\mathbf{a}_j$ is in $\text{span}\{\mathbf{q}_1, \mathbf{q}_2, ..., \mathbf{q}_j\}$. Thus we may write

$$\mathbf{a}_j = r_{1j}\mathbf{q}_1 + r_{2j}\mathbf{q}_2 + \cdots + r_{jj}\mathbf{q}_j + 0\mathbf{q}_{j+1} + \cdots + 0\mathbf{q}_n = \begin{pmatrix} \mathbf{q}_1 & \cdots & \mathbf{q}_j & \cdots & \mathbf{q}_n \end{pmatrix} \begin{pmatrix} r_{1j} \\ \vdots \\ r_{jj} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

QR Factorization

Explicitly,

$$
\begin{aligned}
\mathbf{a}_1 &= r_{11}\mathbf{q}_1 = \begin{pmatrix}\mathbf{q}_1 & \cdots & \mathbf{q}_n\end{pmatrix}\begin{pmatrix} r_{11} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \\
\mathbf{a}_2 &= r_{12}\mathbf{q}_1 + r_{22}\mathbf{q}_2 = \begin{pmatrix}\mathbf{q}_1 & \cdots & \mathbf{q}_n\end{pmatrix}\begin{pmatrix} r_{12} \\ r_{22} \\ \vdots \\ 0 \end{pmatrix} \\
&\vdots \\
\mathbf{a}_n &= r_{1n}\mathbf{q}_1 + r_{2n}\mathbf{q}_2 + \cdots + r_{nn}\mathbf{q}_n = \begin{pmatrix}\mathbf{q}_1 & \cdots & \mathbf{q}_n\end{pmatrix}\begin{pmatrix} r_{1n} \\ r_{2n} \\ \vdots \\ r_{nn} \end{pmatrix}
\end{aligned}
$$

Thus, we may write

$$
\begin{aligned}
\mathbf{A} &= \begin{pmatrix}\mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n\end{pmatrix} \\
&= \begin{pmatrix}\mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_n\end{pmatrix}\begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{nn} \end{pmatrix} \\
&= \mathbf{QR}
\end{aligned}
$$

for some $m \times n$ matrix $\mathbf{Q}$ with orthonormal columns, and a upper triangular $n \times n$ matrix $\mathbf{R}$.

Exercise

1. Prove that $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}_n$.

2. Prove that the diagonal entries of $\mathbf{R}$ are positive, $r_{ii} > 0$ for all $i = 1, ..., n$.

3. Prove that the upper triangular matrix $\mathbf{R} = \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{nn} \end{pmatrix}$ is invertible.

QR Factorization

Theorem (QR Factorization)

Suppose $\mathbf{A}$ *is a* $m \times n$ *matrix with linearly independent columns. Then* $\mathbf{A}$ *can be written as*

$$\mathbf{A} = \mathbf{QR}$$

for some $m \times n$ *matrix* $\mathbf{Q}$ *such that* $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}_n$ *and invertible upper triangular matrix* $\mathbf{R}$ *with positive diagonal entries.*

Definition

The decomposition given in the theorem above is called a *QR factorization* of $\mathbf{A}$.

Example

Find a QR factorization of $\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$.

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad \|\mathbf{v}_1\| = 2$$

$$\mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} - \frac{3}{4}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -3/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{pmatrix}, \quad \|\mathbf{v}_2\| = \frac{\sqrt{3}}{2}$$

$$\mathbf{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} - \frac{1}{6}\begin{pmatrix} -3 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -2/3 \\ 1/3 \\ 1/3 \end{pmatrix}, \quad \|\mathbf{v}_3\| = \sqrt{\frac{2}{3}}$$

So,

$$\mathbf{Q} = \begin{pmatrix} 1/2 & -\sqrt{3}/2 & 0 \\ 1/2 & 1/(2\sqrt{3}) & -\sqrt{2/3} \\ 1/2 & 1/(2\sqrt{3}) & 1/\sqrt{6} \\ 1/2 & 1/(2\sqrt{3}) & 1/\sqrt{6} \end{pmatrix}.$$

Example

To compute $\mathbf{R}$, recall that $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}$. So, premultiplying $\mathbf{Q}^T$ to both sides of the equation $\mathbf{A} = \mathbf{QR}$, we have

$$\mathbf{Q}^T\mathbf{A} = \mathbf{Q}^T\mathbf{QR} = \mathbf{R}.$$

$$\mathbf{R} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{-\sqrt{3}}{2} & \frac{1}{2\sqrt{3}} & \frac{1}{2\sqrt{3}} & \frac{1}{2\sqrt{3}} \\ 0 & -\sqrt{\frac{2}{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 3/2 & 1 \\ 0 & \sqrt{3}/2 & 1/\sqrt{3} \\ 0 & 0 & \sqrt{2/3} \end{pmatrix}.$$

Hence,

$$\mathbf{A} = \begin{pmatrix} 1/2 & -\sqrt{3}/2 & 0 \\ 1/2 & 1/(2\sqrt{3}) & -\sqrt{2/3} \\ 1/2 & 1/(2\sqrt{3}) & 1/\sqrt{6} \\ 1/2 & 1/(2\sqrt{3}) & 1/\sqrt{6} \end{pmatrix} \begin{pmatrix} 2 & 3/2 & 1 \\ 0 & \sqrt{3}/2 & 1/\sqrt{3} \\ 0 & 0 & \sqrt{2/3} \end{pmatrix}$$

Algorithm to QR Factorization

Let $\mathbf{A}$ be a $m \times n$ matrix with linearly independent columns.

1. Perform Gram-Schmidt on the columns of $\mathbf{A} = \begin{pmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{pmatrix}$ to obtain an orthonormal set $\{\mathbf{q}_1, \mathbf{q}_2, ..., \mathbf{q}_n\}$.
2. Set $\mathbf{Q} = \begin{pmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_n \end{pmatrix}$.
3. Compute $\mathbf{R} = \mathbf{Q}^T \mathbf{A}$.

Exercise

Use QR factorization to prove the following

Corollary

Suppose $\mathbf{A}$ *is a* $m \times n$ *matrix with linearly independent columns, i.e.* $\mathrm{rank}(\mathbf{A}) = n$. *Then* $\mathbf{A}^T\mathbf{A}$ *is invertible, and* $\mathbf{A}$ *has a left inverse; that is, there is a* $\mathbf{B}$ *such that*

$$\mathbf{B}\mathbf{A} = \mathbf{I}_n.$$

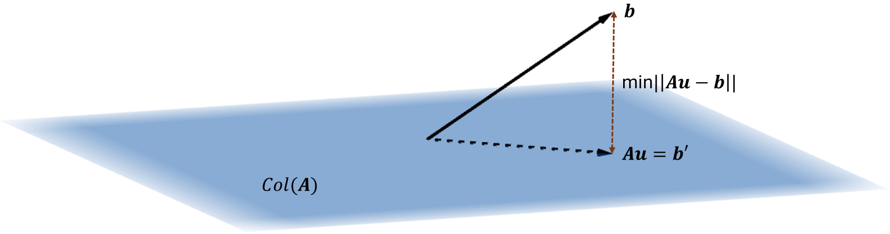
5.5 Least Square Approximation

Introduction

- Recall that the column space of a $m \times n$ matrix $\mathbf{A}$ is the set of all vectors $\mathbf{b}$ such that $\mathbf{Ax} = \mathbf{b}$ is consistent,

$$\text{Col}(\mathbf{A}) = \{\ \mathbf{b} \in \mathbb{R}^m \mid \mathbf{Ax} = \mathbf{b} \text{ is consistent } \} = \{\ \mathbf{Au} \mid \mathbf{u} \in \mathbb{R}^n\ \}.$$

- Now suppose $\mathbf{Ax} = \mathbf{b}$ is inconsistent, that is, $\mathbf{b}$ is not in the column space of $\mathbf{A}$.
- We may ask for a vector $\mathbf{b}'$ in the column of $\mathbf{A}$ that is the closest to $\mathbf{b}$, that is, find a $\mathbf{u}$ such that $\mathbf{Au} = \mathbf{b}'$ is the closest to $\mathbf{b}$.
- This is equivalent to finding a $\mathbf{u}$ in $\mathbb{R}^n$ such that $\|\mathbf{Au} - \mathbf{b}\|$ is minimized.



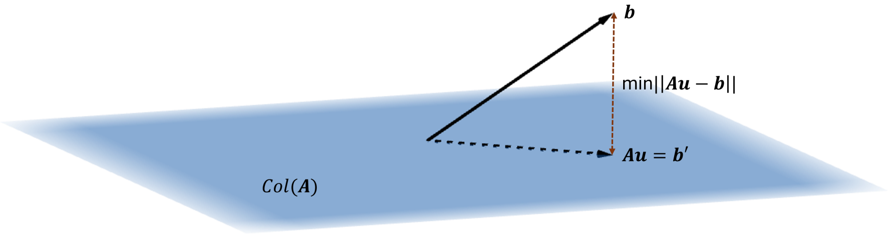
Least Square Approximation

Definition

Let $\mathbf{A}$ be a $m \times n$ matrix and $\mathbf{b}$ a vector in $\in \mathbb{R}^m$. A vector $\mathbf{u}$ in $\mathbb{R}^n$ is a *least square solution* of $\mathbf{Ax} = \mathbf{b}$ if for every vector $\mathbf{v} \in \mathbb{R}^n$,

$$\|\mathbf{Au} - \mathbf{b}\| \leq \|\mathbf{Av} - \mathbf{b}\|.$$

Geometrically, by the best approximation theorem, the vector $\mathbf{b}' = \mathbf{Au}$ in $\mathrm{Col}(\mathbf{A})$ closest to $\mathbf{b}$ is the projection of $\mathbf{b}$ onto $\mathrm{Col}(\mathbf{A})$.



Least Square Approximation

Theorem

Let $\mathbf{A}$ *be a* $m \times n$ *matrix and* $\mathbf{b}$ *a vector in* $\mathbb{R}^m$. *A vector* $\mathbf{u}$ *in* $\mathbb{R}^n$ *is a least square solution to* $\mathbf{Ax} = \mathbf{b}$ *if and only if* $\mathbf{Au}$ *is the projection of* $\mathbf{b}$ *onto the column space of* $\text{Col}(\mathbf{A})$.

Theorem

Let $\mathbf{A}$ *be a* $m \times n$ *matrix and* $\mathbf{b}$ *a vector in* $\mathbb{R}^m$. *A vector* $\mathbf{u}$ *in* $\mathbb{R}^n$ *is a least square solution to* $\mathbf{Ax} = \mathbf{b}$ *if and only if* $\mathbf{u}$ *is a solution to* $\mathbf{A}^T\mathbf{Ax} = \mathbf{A}^T\mathbf{b}$.

Proof.

$\mathbf{Au}$ is the projection of $\mathbf{b}$ onto the column space of $\mathbf{A}$ if and only if $\mathbf{Au} - \mathbf{b}$ is orthogonal to the column space of $\mathbf{A}$. By the orthogonal to a subspace theorem, since the columns of $\mathbf{A}$ spans the column space of $\mathbf{A}$, $\mathbf{Au} - \mathbf{b}$ is orthogonal to the column space of $\mathbf{A}$ if and only if $\mathbf{Au} - \mathbf{b}$ is in the nullspace of $\mathbf{A}^T$,

$$\mathbf{A}^T(\mathbf{Au} - \mathbf{b}) = \mathbf{0}$$

Rearranging the equation, we have

$$\mathbf{A}^T\mathbf{Au} = \mathbf{A}^T\mathbf{b}.$$

□

Examples

Let $\mathbf{A} = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 2 & 1 & 1 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. Is $\mathbf{Ax} = \mathbf{b}$ consistent?

$$\left(\begin{array}{cccc|c} 1 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 1 & 2 & 1 & 1 & 1 \end{array} \right) \xrightarrow{RREF} \left(\begin{array}{cccc|c} 1 & 0 & -1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array} \right)$$

The system is inconsistent, thus $\mathbf{b}$ is not in $\mathrm{Col}(\mathbf{A})$.

Example

Let $\mathbf{A} = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 2 & 1 & 1 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. Find a least square solution of $\mathbf{Ax} = \mathbf{b}$, that is, solve $\mathbf{A}^T\mathbf{Ax} = \mathbf{A}^T\mathbf{b}$.

$$\mathbf{A}^T\mathbf{A} = \begin{pmatrix} 2 & 3 & 1 & 2 \\ 3 & 6 & 3 & 3 \\ 1 & 3 & 2 & 1 \\ 2 & 3 & 1 & 2 \end{pmatrix}, \ \mathbf{A}^T\mathbf{b} = \begin{pmatrix} 2 \\ 4 \\ 2 \\ 2 \end{pmatrix}$$

$$\left(\begin{array}{cccc|c} 2 & 3 & 1 & 2 & 2 \\ 3 & 6 & 3 & 3 & 4 \\ 1 & 3 & 2 & 1 & 2 \\ 2 & 3 & 1 & 2 & 2 \end{array}\right) \xrightarrow{RREF} \left(\begin{array}{cccc|c} 1 & 0 & -1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 2/3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

General solution: $\begin{pmatrix} 0 \\ 2/3 \\ 0 \\ 0 \end{pmatrix} + s\begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$, $s, t \in \mathbb{R}$. This shows that least square solution might not be unique.

Example

Let $\mathbf{A} = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 2 & 1 & 1 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. Find a least square solution of $\mathbf{Ax} = \mathbf{b}$, that is, solve $\mathbf{A}^T\mathbf{Ax} = \mathbf{A}^T\mathbf{b}$.

General solution to $\mathbf{A}^T\mathbf{Ax} = \mathbf{A}^T\mathbf{b}$: $\begin{pmatrix} 0 \\ 2/3 \\ 0 \\ 0 \end{pmatrix} + s\begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$, $s, t \in \mathbb{R}$.

Now for any $s, t \in \mathbb{R}$,

$$\mathbf{A}\left(\begin{pmatrix} 0 \\ 2/3 \\ 0 \\ 0 \end{pmatrix} + s\begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix}\right) = \mathbf{A}\begin{pmatrix} 0 \\ 2/3 \\ 0 \\ 0 \end{pmatrix} = \frac{2}{3}\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}.$$

So, for any choice of least square solution $\mathbf{u}$, the projection $\mathbf{Au}$ is unique.

Example

Find the least square solutions to

$$\begin{cases} x & - & y & + & z & = & 1 \\ -x & + & y & + & z & = & 2 \\ x & & & + & z & = & 2 \\ -x & + & y & - & z & = & -1 \end{cases}$$

$$(\mathbf{A}^T\mathbf{A} \mid \mathbf{A}^T\mathbf{b}) = \left(\begin{array}{ccc|c} 4 & -3 & 2 & 2 \\ -3 & 3 & -1 & 0 \\ 2 & -1 & 4 & 6 \end{array}\right) \xrightarrow{RREF} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1/2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 3/2 \end{array}\right)$$

Unique least square solution: $\begin{pmatrix} 1/2 \\ 1 \\ 3/2 \end{pmatrix}$.

In this case, observe that $\mathbf{A}^T\mathbf{A}$ is invertible. So, could have computed the least square solution by

$$\mathbf{u} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b} = \begin{pmatrix} 11/8 & 5/4 & -3/8 \\ 5/4 & 3/2 & -1/4 \\ -3/8 & -1/4 & 3/8 \end{pmatrix} \begin{pmatrix} 2 \\ 0 \\ 6 \end{pmatrix} = \begin{pmatrix} 1/2 \\ 1 \\ 3/2 \end{pmatrix}.$$

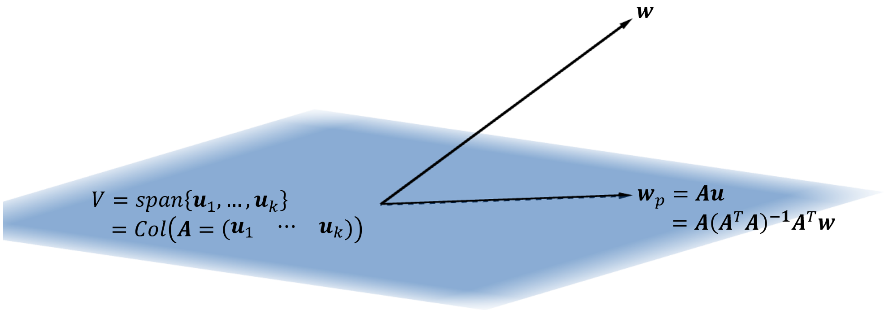
Challenge

Let $\mathbf{A}$ be a $m \times n$ matrix and $\mathbf{b}$ a vector in $\mathbb{R}^m$. Prove that for any choice of least square solution $\mathbf{u}$, that is, for any solution $\mathbf{u}$ of $\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{b}$, the projection $\mathbf{A}\mathbf{u}$ is unique.

Orthogonal Projection (Revisit)

We may use least square solutions to find the projection of a vector onto a subspace.

- Let V be subspace of $\mathbb{R}^n$. Let $S = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ be a basis for V.
- Define $\mathbf{A} = \begin{pmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_k \end{pmatrix}$, by construction, the column space of $\mathbf{A}$ is V, $V = \mathrm{Col}(\mathbf{A})$.
- Let $\mathbf{w}$ be a vector in $\mathbb{R}^n$, and $\mathbf{u}$ a least square solution to $\mathbf{Ax} = \mathbf{w}$. Then $\mathbf{w}_p = \mathbf{Au}$ is the projection of $\mathbf{w}$ onto $\mathrm{Col}(\mathbf{A}) = V$.



Orthogonal Projection (Revisit)

- Now $\mathbf{u}$ is a least square solution to $\mathbf{Ax} = \mathbf{w}$ if and only if it is a solution to $\mathbf{A}^T\mathbf{Ax} = \mathbf{A}^T\mathbf{w}$. But since the columns of $\mathbf{A}$ are linearly independent, $\mathbf{A}^T\mathbf{A}$ is invertible.
- This means that $\mathbf{u} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{w}$.
- Hence, the projection is
$$\mathbf{w}_p = \mathbf{Au} = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{w}.$$

Theorem

Let V be a subspace of $\mathbb{R}^n$ and $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_k\}$ be a basis for V. Then the orthogonal projection of a vector $\mathbf{w}$ onto V is
$$\mathbf{w}_p = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{w},$$
where $\mathbf{A} = \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_k \end{pmatrix}$.

Examples

Let $S = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \right\}$ and $V = \text{span}(S)$. Let $\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 0 & 0 \end{pmatrix}$. Then the orthogonal projection of $\mathbf{w} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ onto V is

$$
\begin{aligned}
\mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{w} &= \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}^{-1} \begin{pmatrix} x+y \\ x-y \end{pmatrix} \\
&= \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x+y \\ x-y \end{pmatrix} \\
&= \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}.
\end{aligned}
$$

Indeed, since V is the xy-plane in $\mathbb{R}^3$.

Question

Suppose the system $\mathbf{Ax} = \mathbf{b}$ is consistent.

1. Suppose $\mathbf{u}$ is a solution to $\mathbf{Ax} = \mathbf{b}$. Is $\mathbf{u}$ a least square solution to $\mathbf{Ax} = \mathbf{b}$?

2. Suppose $\mathbf{u}$ is a least square solution to $\mathbf{Ax} = \mathbf{b}$. Is $\mathbf{u}$ a solution to $\mathbf{Ax} = \mathbf{b}$?

Exercise

Suppose $\mathbf{A}$ is a $m \times n$ matrix with linearly independent columns, i.e. $\text{rank}(\mathbf{A}) = n$. QR factorize $\mathbf{A}$,

$$\mathbf{A} = \mathbf{QR}.$$

Show that the unique least square solution of $\mathbf{Ax} = \mathbf{b}$ is

$$\mathbf{u} = \mathbf{R}^{-1}\mathbf{Q}^T\mathbf{b}.$$

That is, suffice to solve for

$$\mathbf{Rx} = \mathbf{Q}^T\mathbf{b}.$$

This is easy to solve by hand since $\mathbf{R}$ is a upper triangular matrix (i.e. a REF).

Challenge

Let $V \subseteq \mathbb{R}^n$ be a subspace and suppose $\{\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_k\}$ is an orthonormal basis of V. Write

$$\mathbf{Q} = \begin{pmatrix} \mathbf{w}_1 & \mathbf{w}_2 & \cdots & \mathbf{w}_k \end{pmatrix}.$$

Then for any $\mathbf{w} \in \mathbb{R}^n$, the projection of $\mathbf{w}$ onto V is

$$\mathbf{Q}\mathbf{Q}^T\mathbf{w}.$$