July 18 Math 3260 sec. 51 Summer 2023

Section 6.1: Inner Product, Length, and Orthogonality

Definition: Inner (a.k.a. Scalar) Product

For vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ we define the **inner product** of $\mathbf{u}$ and $\mathbf{v}$ (also called the **dot product**) by the **matrix product**

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

There are several notations for this including

$$\mathbf{u}^T\mathbf{v},\quad \mathbf{u}\cdot\mathbf{v},\quad \text{and}\quad \langle \mathbf{u},\mathbf{v}\rangle$$

Properties of the Inner Product

Theorem:

For each $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ in $\mathbb{R}^n$ and scalar c

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

ii. $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$

iii. $\mathbf{u} \cdot (c\mathbf{v}) = (c\mathbf{u}) \cdot \mathbf{v} = c(\mathbf{u} \cdot \mathbf{v})$

iv. $\mathbf{u} \cdot \mathbf{u} \geq 0$ and $\mathbf{u} \cdot \mathbf{u} = 0$ if and only if $\mathbf{u} = \mathbf{0}$.

The Norm, Unit Vectors & Normalizing

Definition

The **norm** of the vector $\mathbf{v}$ in $\mathbb{R}^n$ is the nonnegative number $\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}}$.

Definition

A vector $\mathbf{u}$ in $\mathbb{R}^n$ for which $\|\mathbf{u}\| = 1$ is called a **unit vector**.

Theorem

For vector $\mathbf{v}$ in $\mathbb{R}^n$ and scalar c $\|c\mathbf{v}\| = |c|\|\mathbf{v}\|$.

If $\mathbf{v} \neq \mathbf{0}$, then we can **normalize** to get a unit vector. $\dfrac{\mathbf{v}}{\|\mathbf{v}\|}$ is a unit vector parallel to $\mathbf{v}$.

Distance in $\mathbb{R}^n$

Definition:

For vectors **u** and **v** in $\mathbb{R}^n$, the **distance between u and v** is denoted by

$$\text{dist}(\mathbf{u}, \mathbf{v}),$$

and is defined by

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

Remark: This is the same as the traditional formula for distance used in $\mathbb{R}^2$ between points (x_0, y_0) and (x_1, y_1),

$$d = \sqrt{(y_1 - y_0)^2 + (x_1 - x_0)^2}.$$

Orthogonality & The Pythagorean Theorem

Definition:

Two vectors are $\mathbf{u}$ and $\mathbf{v}$ **orthogonal** if $\mathbf{u} \cdot \mathbf{v} = 0$.

The Pythagorean Theorem:

Two vectors $\mathbf{u}$ and $\mathbf{v}$ are orthogonal if and only if

$$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2.$$

Orthogonal Complement

Definition:

Let W be a subspace of $\mathbb{R}^n$. A vector $\mathbf{z}$ in $\mathbb{R}^n$ is said to be **orthogonal to** W if $\mathbf{z}$ is orthogonal to every vector in W. That is, if

$$\mathbf{z} \cdot \mathbf{w} = 0 \quad \text{for every} \quad \mathbf{w} \in W.$$

Definition:

Given a subspace W of $\mathbb{R}^n$, the set of all vectors orthogonal to W is called the **orthogonal complement** of W and is denoted by $W^\perp$ (read as "W perp").

$$W^\perp = \left\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} \cdot \mathbf{w} = 0 \quad \text{for every} \quad \mathbf{w} \in W \right\}$$

Theorem:

Theorem:

If W is a subspace of $\mathbb{R}^n$, then $W^\perp$ is a subspace of $\mathbb{R}^n$.

This is readily proved by appealing to the properties of the inner product. In particular

$$\mathbf{0}\cdot\mathbf{w} = 0 \quad \text{for any vector} \quad \mathbf{w}$$

$$(\mathbf{u}+\mathbf{v})\cdot\mathbf{w} = \mathbf{u}\cdot\mathbf{w}+\mathbf{v}\cdot\mathbf{w} \quad \text{and}$$

$$(c\mathbf{u})\cdot\mathbf{w} = c\mathbf{u}\cdot\mathbf{w}.$$

Example

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

A vector in W has the form

$$\mathbf{w} = x \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + z \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ 0 \\ z \end{bmatrix}.$$

xz-plane

A vector in $\mathbf{v}$ in $W^\perp$ has the form

$$\mathbf{v} = y \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ y \\ 0 \end{bmatrix}.$$

Note that

$$\mathbf{w} \cdot \mathbf{v} = x(0) + 0(y) + z(0) = 0.$$

W is the xz-plane and $W^\perp$ is the y-axis in $\mathbb{R}^3$.

Example

Let $A = \begin{bmatrix} 1 & 3 & 2 \\ -2 & 0 & 4 \end{bmatrix}$. Show that if $\mathbf{x}$ is in Nul(A), then $\mathbf{x}$ is in $[\text{Row}(A)]^{\perp}$.

we can characterize Nul A.

$$A \xrightarrow{\text{rref}} \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 4/3 \end{bmatrix} \quad \text{If } A\vec{x} = \vec{0}$$

$x_1 = 2x_3$

$x_2 = -4/3\, x_3$

x_3 - free

$$\vec{x} = x_3 \begin{bmatrix} 2 \\ -4/3 \\ 1 \end{bmatrix} = \frac{1}{3} x_3 \begin{bmatrix} 6 \\ -4 \\ 3 \end{bmatrix}$$

$$\text{Row } A = \text{Span}\left\{ \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 4 \end{bmatrix} \right\} = \{\vec{v}_1, \vec{v}_2\}$$

Let's show that $\vec{x}\cdot\vec{v}_1 = 0$ and $\vec{x}\cdot\vec{v}_2 = 0$

$$\vec{x}\cdot\vec{v}_1 = x_3\begin{bmatrix}6\\-4\\3\end{bmatrix}\cdot\begin{bmatrix}1\\3\\2\end{bmatrix} = x_3\left(6-12+6\right) = 0$$

$$\vec{x}\cdot\vec{v}_2 = x_3\begin{bmatrix}6\\-4\\3\end{bmatrix}\cdot\begin{bmatrix}-2\\0\\4\end{bmatrix} = x_3\left(-12+0+12\right) = 0$$

For any $\vec{v}$ is Row A,

$\vec{v} = c_1\vec{v}_1 + c_2\vec{v}_2$ so

$$\vec{x}\cdot\vec{v} = c_1\vec{x}\cdot\vec{v}_1 + c_2\vec{x}\cdot\vec{v}_2 = 0+0 = 0$$

So every $\vec{x}$ in Nul A is in

$[\text{Row } A]^{\perp}$.

Note $A\vec{x} = \vec{0} \Rightarrow$

$$\begin{bmatrix} 1 & 3 & 2 \\ -2 & 0 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

The Fundamental Subspaces of a Matrix

Theorem:

Let A be an $m \times n$ matrix. The orthogonal complement of the row space of A is the null space of A. That is

$$[\text{Row}(A)]^{\perp} = \text{Nul}(A).$$

The orthogonal complement of the column space of A is the null space of A^T—i.e.

$$[\text{Col}(A)]^{\perp} = \text{Nul}(A^T).$$

Example

Let $W = \text{Span}\left\{ \begin{bmatrix} 5 \\ -3 \\ 2 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ 9 \\ -1 \end{bmatrix} \right\}$. W is a subspace of $\mathbb{R}^5$. Find a basis for $W^\perp$, the orthogonal complement of W.

Let's use a matrix A having W as its row space.

$$A = \begin{bmatrix} 5 & -3 & 2 & 2 & 0 \\ 1 & 0 & 1 & 9 & -1 \end{bmatrix} \qquad W^\perp = \text{Nul}\, A$$

$$A \xrightarrow{\text{rref}} \begin{bmatrix} 1 & 0 & 1 & 9 & -1 \\ 0 & 1 & 1 & \frac{43}{3} & \frac{-5}{3} \end{bmatrix} \quad A\vec{x} = \vec{0}$$

$$\Rightarrow \quad x_1 = -x_3 - 9x_4 + x_5$$
$$x_2 = -x_3 - \frac{43}{3}x_4 + \frac{5}{3}x_5$$
$$x_3, x_4, x_5 - \text{free}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = x_3 \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -9 \\ \frac{-43}{3} \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} 1 \\ 5/3 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

Using $x_3 = 1$, then $x_4 = 3$ then $x_5 = 3$

A basis for $W^\perp$ is

$$\left\{ \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -27 \\ -43 \\ 0 \\ 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 5 \\ 0 \\ 0 \\ 3 \end{bmatrix} \right\}$$

Section 6.2: Orthogonal Sets

Remark: We know that if $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_p\}$ is a basis for a subspace W of $\mathbb{R}^n$, then each vector $\mathbf{x}$ in W can be realized (uniquely) as a sum

$$\mathbf{x} = c_1\mathbf{b}_2 + \cdots + c_p\mathbf{b}_p.$$

If n is very large, the computations needed to determine the coefficients $c_1, \ldots, c_p$ may require a lot of time (and machine memory).

Question: Can we seek a basis whose nature simplifies this task? And what properties should such a basis possess?

Orthogonal Sets

Definition:

An indexed set $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ in $\mathbb{R}^n$ is said to be an **orthogonal set** provided each pair of distinct vectors in the set is orthogonal. That is, provided

$$\mathbf{u}_i \cdot \mathbf{u}_j = 0 \quad \text{whenever} \quad i \neq j.$$

Example: Show that the set

$$\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\} = \left\{ \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix} \right\}$$ is an orthogonal set.

we have to compute $\vec{u}_1 \cdot \vec{u}_2$,

$\vec{u}_1 \cdot \vec{u}_3$ and $\vec{u}_2 \cdot \vec{u}_3$.

$$\mathbf{u}_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \mathbf{u}_3 = \begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix}$$

$$\vec{u}_1 \cdot \vec{u}_2 = 3(-1) + 1(2) + 1(1) = -3+2+1 = 0$$

$$\vec{u}_1 \cdot \vec{u}_3 = 3(-1) + 1(-4) + 1(7) = -3-4+7 = 0$$

$$\vec{u}_2 \cdot \vec{u}_3 = -1(-1) + 2(-4) + 1(7) = 1-8+7 = 0$$

So $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ is an orthogonal set.

Orthongal Basis

Definition:

An **orthogonal basis** for a subspace W of $\mathbb{R}^n$ is a basis that is also an orthogonal set.

Theorem:

Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ be an orthogonal basis for a subspace W of $\mathbb{R}^n$. Then each vector $\mathbf{y}$ in W can be written as the linear combination

$$\mathbf{y} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_p\mathbf{u}_p, \quad \text{where the weights} \quad c_j = \frac{\mathbf{y}\cdot\mathbf{u}_j}{\mathbf{u}_j\cdot\mathbf{u}_j}.$$

Remark: What's nice about this is how simple the formula for the c's is.

Example

$\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\} = \left\{ \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix} \right\}$ is an orthogonal basis of $\mathbb{R}^3$. Express the vector $\mathbf{y} = \begin{bmatrix} -2 \\ 3 \\ 0 \end{bmatrix}$ as a linear combination of the basis vectors.

$$\vec{y} = c_1\vec{u}_1 + c_2\vec{u}_2 + c_3\vec{u}_3 \qquad c_i = \frac{\vec{y}\cdot\vec{u}_i}{\vec{u}_i\cdot\vec{u}_i}$$

$$\vec{y}\cdot\vec{u}_1 = -3 \qquad \vec{u}_1\cdot\vec{u}_1 = 11$$

$$\vec{y}\cdot\vec{u}_2 = 8 \qquad \vec{u}_2\cdot\vec{u}_2 = 6$$

$$\vec{y}\cdot\vec{u}_3 = -10 \qquad \vec{u}_3\cdot\vec{u}_3 = 66$$

$$\vec{y} = \frac{-3}{11}\vec{u}_1 + \frac{8}{6}\vec{u}_2 - \frac{10}{66}\vec{u}_3$$

$$= \frac{-3}{11}\vec{u}_1 + \frac{4}{3}\vec{u}_2 - \frac{5}{33}\vec{u}_3$$

Observation

Consider the alternatives for this example.

We're writting **y** in the new basis $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$. One option is to use the change of basis matrix and write

$$[\mathbf{y}]_{\mathcal{B}} = \begin{bmatrix} 3 & -1 & -1 \\ 1 & 2 & -4 \\ 1 & 1 & 7 \end{bmatrix}^{-1} \mathbf{y}$$

Or we could set up an augemented matrix and perform row reduction.

$$\begin{bmatrix} 3 & -1 & -1 & -2 \\ 1 & 2 & -4 & 3 \\ 1 & 1 & 7 & 0 \end{bmatrix} \xrightarrow{\text{rref}} \cdots$$

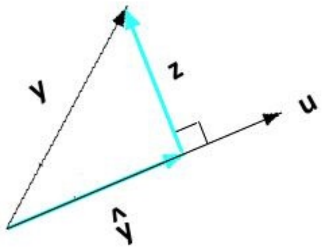
Because of the nice nature of the orthogonal basis, we can find the coefficients **by hand** with very little effort!

Projection

Given a nonzero vector $\mathbf{u}$, suppose we wish to decompose another nonzero vector $\mathbf{y}$ into a sum of the form

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

in such a way that $\hat{\mathbf{y}}$ is parallel to $\mathbf{u}$ and $\mathbf{z}$ is perpendicular to $\mathbf{u}$.



Projection

Since $\hat{\mathbf{y}}$ is parallel to $\mathbf{u}$, there is a scalar α such that

$$\hat{\mathbf{y}} = \alpha\mathbf{u}.$$

Find a formula for this scalar.

$\vec{y} = \hat{y} + \vec{z}$ where $\hat{y} \parallel \vec{u}$ and $\underbrace{\vec{z} \perp \vec{u}}_{\vec{z}\cdot\vec{w} = 0}$

$$\vec{u}\cdot\vec{y} = \vec{u}\cdot(\hat{y} + \vec{z})$$

$$= \vec{u}\cdot(\alpha\vec{u}) + \underbrace{\vec{u}\cdot\vec{z}}_{=0}$$

$$\vec{u}\cdot\vec{y} = \alpha\,\vec{u}\cdot\vec{u} \implies \alpha = \frac{\vec{u}\cdot\vec{y}}{\vec{u}\cdot\vec{u}}$$

$$\hat{y} = \left(\frac{\vec{u}\cdot\vec{y}}{\vec{u}\cdot\vec{u}}\right)\vec{u}$$

Projection onto the subspace $L = \text{Span}\{\mathbf{u}\}$

Projection Notation

We'll use the following notation for the project of a vector $\mathbf{y}$ onto the line $L = \text{Span}\{\mathbf{u}\}$ for nonzero vector $\mathbf{u}$.

$$\hat{\mathbf{y}} = \text{proj}_L \mathbf{y} = \left(\frac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\right)\mathbf{u}.$$

This may also be written as $\text{proj}_{\mathbf{u}}\mathbf{y}$.

This is read as "the projection of $\mathbf{y}$ onto $\mathbf{u}$ (or onto L)."

In writing $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$, there's not a special formula for the $\mathbf{z}$ part. It's just the difference, $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}$.

Example

Let $\mathbf{y} = \begin{bmatrix} 7 \\ 6 \end{bmatrix}$ and $\mathbf{u} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$. Write $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$ where $\hat{\mathbf{y}}$ is in Span$\{\mathbf{u}\}$ and $\mathbf{z}$ is orthogonal to $\mathbf{u}$.

$$\hat{y} = \frac{\vec{y}\cdot\vec{u}}{\vec{u}\cdot\vec{u}}\,\vec{u}$$

$$\vec{y}\cdot\vec{u} = 7(4)+6(2) = 40$$

$$\vec{u}\cdot\vec{u} = 4^2+2^2 = 20$$

$$\hat{y} = \frac{40}{20}\begin{bmatrix} 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$$

$$\vec{y} = \hat{y} + \vec{z} \Rightarrow \vec{z} = \vec{y} - \hat{y}$$

$$= \begin{bmatrix} 7 \\ 6 \end{bmatrix} - \begin{bmatrix} 8 \\ 4 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

$$\begin{bmatrix} 7 \\ 6 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix} + \begin{bmatrix} -1 \\ 2 \end{bmatrix} \qquad \vec{u} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$$

$$\vec{y} \qquad \hat{y} \qquad \vec{z}$$

Note $\vec{u} \cdot \vec{z} = -4 + 4 = 0$ ✓

Example Continued...

Determine the distance between the point $(7, 6)$ and the line Span$\{\mathbf{u}\}$.

The distance is $\|\vec{z}\|$

$$\text{dist}(\vec{y}, L) = \sqrt{(-1)^2 + 2^2} = \sqrt{5}$$

$\uparrow$

distance between $\vec{y}$ and the line $L = \text{Span}\{\vec{u}\}$.

Distance between point and line

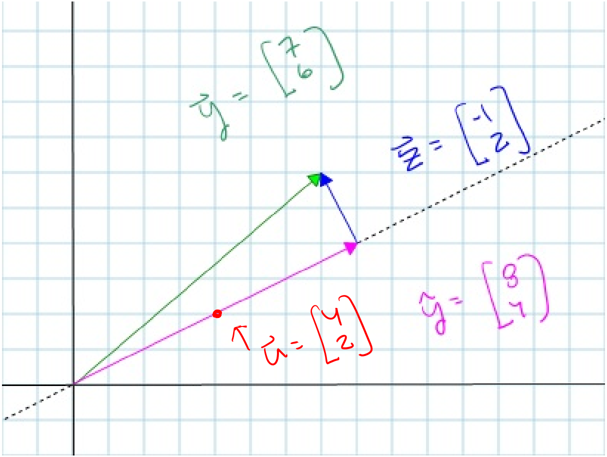


Figure: The distance between the point $(7, 6)$ and the line Span$\{\mathbf{u}\}$ is the norm of $\mathbf{z}$.

Orthonormal Sets

Definition:

A set $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is called an **orthonormal set** if it is an orthogonal set of **unit vectors.**

Definition:

An **orthonormal basis** of a subspace W of $\mathbb{R}^n$ is a basis that is also an orthonormal set.

Remark: So an **orthonormal** set (or basis) is an orthogonal set (or basis) with the extra condition that each vector has norm $\sqrt{\mathbf{u}_i \cdot \mathbf{u}_i} = 1$.

Remark: Any orthogonal set can be normalized to obtain an orthonormal one.

Example:

Show that $\left\{ \overset{\vec{u}_1}{\begin{bmatrix} \frac{3}{5} \\ \frac{4}{5} \end{bmatrix}}, \overset{\vec{u}_2}{\begin{bmatrix} -\frac{4}{5} \\ \frac{3}{5} \end{bmatrix}} \right\}$ is an orthonormal basis for $\mathbb{R}^2$.

They are linearly independent, and there are two of them. So it's a basis.

Note

$$\vec{u}_1 \cdot \vec{u}_2 = \frac{-12}{25} + \frac{12}{25} = 0$$

$$\vec{u}_1 \cdot \vec{u}_1 = \left(\frac{3}{5}\right)^2 + \left(\frac{4}{5}\right)^2 = \frac{9+16}{25} = 1$$

$$\vec{u}_2 \cdot \vec{u}_2 = \left(\frac{-4}{5}\right)^2 + \left(\frac{3}{5}\right)^2 = 1$$

$\{\vec{u}_1, \vec{u}_2\}$ is an orthonormal basis for $\mathbb{R}^2$.

Orthogonal Matrix

Consider the matrix $U = \begin{bmatrix} \frac{3}{5} & -\frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{bmatrix}$ whose columns are the vectors in the last example. Compute the product

$$U^T U = \begin{bmatrix} 3/5 & 4/5 \\ -4/5 & 3/5 \end{bmatrix} \begin{bmatrix} \frac{3}{5} & \frac{-4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} \vec{u}_1 \cdot \vec{u}_1 & \vec{u}_1 \cdot \vec{u}_2 \\ \vec{u}_2 \cdot \vec{u}_1 & \vec{u}_2 \cdot \vec{u}_2 \end{bmatrix}$$

Question: What does this say about U^{-1}?

$$U^{-1} = U^T$$

Orthogonal Matrix

Definition:

A square matrix U is called an **orthogonal matrix** if $U^T = U^{-1}$.

Theorem:

An $n \times n$ matrix U is orthogonal if and only if it's columns form an orthonormal basis of $\mathbb{R}^n$.

Remark: The linear transformation associated to an orthogonal matrix preserves *lengths* and *angles* in the sense of the following theorem.

Theorem: Orthogonal Matrices

Theorem

Let U be an $n \times n$ orthogonal matrix and $\mathbf{x}$ and $\mathbf{y}$ vectors in $\mathbb{R}^n$. Then

(a) $\|U\mathbf{x}\| = \|\mathbf{x}\|$

(b) $(U\mathbf{x}) \cdot (U\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$, in particular

(c) $(U\mathbf{x}) \cdot (U\mathbf{y}) = 0$ if and only if $\mathbf{x} \cdot \mathbf{y} = 0$.

Section 6.3: Orthogonal Projections

Equating points with position vectors, we may wish to find the point $\hat{\mathbf{y}}$ in a subspace W of $\mathbb{R}^n$ that is *closest* to a given point $\mathbf{y}$.

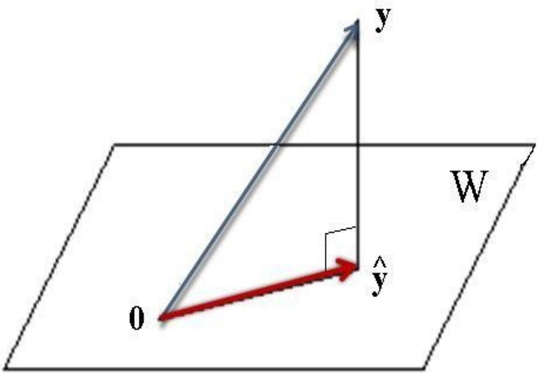


Figure: Illustration of an orthogonal projection. Note that $\text{dist}(\mathbf{y}, \hat{\mathbf{y}})$ is the shortest distance between $\mathbf{y}$ and the points on W.

Orthogonal Decomposition Theorem

Let W be a subspace of $\mathbb{R}^n$. Each vector $\mathbf{y}$ in $\mathbb{R}^n$ can be written uniquely as a sum

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

where $\hat{\mathbf{y}}$ is in W and $\mathbf{z}$ is in $W^\perp$.

If $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is **any orthogonal basis** for W, then

$$\hat{\mathbf{y}} = \sum_{j=1}^{p} \left(\frac{\mathbf{y} \cdot \mathbf{u}_j}{\mathbf{u}_j \cdot \mathbf{u}_j} \right) \mathbf{u}_j, \quad \text{and} \quad \mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}.$$

Remark: The formula for $\hat{\mathbf{y}}$ looks just like the projection onto a line, but with more terms. That is,

$$\hat{\mathbf{y}} = \left(\frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} \right) \mathbf{u}_1 + \left(\frac{\mathbf{y} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2} \right) \mathbf{u}_2 + \cdots + \left(\frac{\mathbf{y} \cdot \mathbf{u}_p}{\mathbf{u}_p \cdot \mathbf{u}_p} \right) \mathbf{u}_p$$

Remarks on the Orthogonal Decomposition Thm.

- Note that the basis must be orthogonal, but otherwise the vector $\hat{\mathbf{y}}$ is **independent** of the particular basis used!

- The vector $\hat{\mathbf{y}}$ is called the **orthogonal projection of y onto** W. We can denote it
$$\mathrm{proj}_W\,\mathbf{y}.$$

- All you really have to do is remember how to project onto a line. Notice that
$$\mathrm{proj}_{\mathbf{u}_1}\,\mathbf{y} = \left(\frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\right)\mathbf{u}_1.$$
If $W = \mathrm{Span}\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ with the $\mathbf{u}$'s orthogonal, then
$$\mathrm{proj}_W\,\mathbf{y} = \mathrm{proj}_{\mathbf{u}_1}\,\mathbf{y} + \mathrm{proj}_{\mathbf{u}_2}\,\mathbf{y} + \cdots + \mathrm{proj}_{\mathbf{u}_p}\,\mathbf{y}.$$

Example

Let $\mathbf{y} = \begin{bmatrix} 4 \\ 8 \\ 1 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}$ and $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$. (a) Verify that $\{\mathbf{u}_1, \mathbf{u}_2\}$ is an orthogonal basis for W.

Note that $\vec{u}_1 \cdot \vec{u}_2 = -4 + 2 + 2 = 0$.

So $\{\vec{u}_1, \vec{u}_2\}$ is orthogonal.

Since they are lin. independent and $W = \text{Span}\{\vec{u}_1, \vec{u}_2\}$, the set is an orthogonal basis.

Example Continued...

$$W = \text{Span}\left\{ \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix} \right\} \quad \text{and} \quad \mathbf{y} = \begin{bmatrix} 4 \\ 8 \\ 1 \end{bmatrix}$$

(b) Find the orthogonal projection of **y** onto W.

$$\hat{y} = \text{proj}_W \vec{y} = \frac{\vec{u}_1 \cdot \vec{y}}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1 + \frac{\vec{u}_2 \cdot \vec{y}}{\vec{u}_2 \cdot \vec{u}_2} \vec{u}_2$$

$\vec{y} \cdot \vec{u}_1 = 18$ $\qquad$ $\vec{u}_1 \cdot \vec{u}_1 = 9$

$\vec{y} \cdot \vec{u}_2 = 9$ $\qquad$ $\vec{u}_2 \cdot \vec{u}_2 = 9$

$$\hat{y} = \frac{18}{9}\begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} + \frac{9}{9}\begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 4 \\ 2 \\ 4 \end{bmatrix} + \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 5 \end{bmatrix}$$

$$\mathrm{Proj}_W \bar{y} = \begin{bmatrix} 2 \\ 4 \\ 5 \end{bmatrix}$$

(c) Find the shortest distance between **y** and the subspace W.

$\vec{y} = \begin{bmatrix} 4 \\ 8 \\ 1 \end{bmatrix}$, $\hat{y} = \begin{bmatrix} 2 \\ 4 \\ 5 \end{bmatrix}$ $\quad \vec{y} = \hat{y} + \vec{z}$, $\vec{z} \in W^{\perp}$

$$\vec{z} = \vec{y} - \hat{y} = \begin{bmatrix} 4 \\ 8 \\ 1 \end{bmatrix} - \begin{bmatrix} 2 \\ 4 \\ 5 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ -4 \end{bmatrix}$$

$$\text{dist}(\vec{y}, W) = \|\vec{z}\| = \sqrt{2^2 + 4^2 + (-4)^2} = 6$$

$$W = \text{Span}\left\{ \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix} \right\}$$

$\vec{z} \cdot \vec{u}_1 = 0$

$\vec{z} \cdot \vec{u}_2 = 0$

Computing Orthogonal Projections

Theorem

If $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an **orthonormal** basis of a subspace W of $\mathbb{R}^n$, and $\mathbf{y}$ is any vector in $\mathbb{R}^n$ then

$$\text{proj}_W \mathbf{y} = \sum_{j=1}^{p} (\mathbf{y} \cdot \mathbf{u}_j)\, \mathbf{u}_j.$$

And, if U is the matrix $U = [\mathbf{u}_1 \quad \cdots \quad \mathbf{u}_p]$, then the above is equivalent to

$$\text{proj}_W \mathbf{y} = UU^T\mathbf{y}.$$

Remark: In general, U is not square; it's $n \times p$. So even though UU^T will be a square matrix, it is not the same matrix as U^TU and it is not the identity matrix.

Example

Let $\mathbf{v}_1 = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}$ and $W = \text{Span}\,\{\mathbf{v}_1, \mathbf{v}_2\}$. Find an orthonormal basis $\{\mathbf{u}_1, \mathbf{u}_2\}$ for W. Then compute the matrices U^TU and UU^T where $U = [\mathbf{u}_1 \;\; \mathbf{u}_2]$.

$\|\vec{v}_1\|^2 = 9 \implies \|\vec{v}_1\| = 3$, $\|\vec{v}_2\|^2 = 9 \implies \|\vec{v}_2\| = 3$

$\vec{u}_1 = \dfrac{\vec{v}_1}{\|\vec{v}_1\|} = \frac{1}{3}\vec{v}_1 = \begin{bmatrix} 2/3 \\ 1/3 \\ 2/3 \end{bmatrix}$, $\vec{u}_2 = \dfrac{\vec{v}_2}{\|\vec{v}_2\|} = \begin{bmatrix} -2/3 \\ 2/3 \\ 1/3 \end{bmatrix}$

$$U = \begin{bmatrix} 2/3 & -2/3 \\ 1/3 & 2/3 \\ 2/3 & 1/3 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 2 & -2 \\ 1 & 2 \\ 2 & 1 \end{bmatrix}$$

$$u^T u = \frac{1}{3}\begin{bmatrix} 2 & 1 & 2 \\ -2 & 2 & 1 \end{bmatrix} \frac{1}{3}\begin{bmatrix} 2 & -2 \\ 1 & 2 \\ 2 & 1 \end{bmatrix} = \frac{1}{9}\begin{bmatrix} 9 & 0 \\ 0 & 9 \end{bmatrix}.$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

$$u u^T = \frac{1}{3}\begin{bmatrix} 2 & -2 \\ 1 & 2 \\ 2 & 1 \end{bmatrix} \frac{1}{3}\begin{bmatrix} 2 & 1 & 2 \\ -2 & 2 & 1 \end{bmatrix}$$

$$= \frac{1}{9}\begin{bmatrix} 8 & -2 & 2 \\ -2 & 5 & 4 \\ 2 & 4 & 5 \end{bmatrix}.$$

$$(u u^T)^T = (u^T)^T u^T = u u^T$$

Example

$$W = \text{Span}\left\{ \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix} \right\} \quad \text{and} \quad \mathbf{y} = \begin{bmatrix} 4 \\ 8 \\ 1 \end{bmatrix}$$

Use the matrix formulation to find $\text{proj}_W \, \mathbf{y}$.

$$U = \frac{1}{3}\begin{bmatrix} 2 & -2 \\ 1 & 2 \\ 2 & 1 \end{bmatrix} \qquad UU^T = \frac{1}{9}\begin{bmatrix} 8 & -2 & 2 \\ -2 & 5 & 4 \\ 2 & 4 & 5 \end{bmatrix}$$

$$\text{proj}_W \vec{y} = UU^T \vec{y} = \frac{1}{9}\begin{bmatrix} 8 & -2 & 2 \\ -2 & 5 & 4 \\ 2 & 4 & 5 \end{bmatrix}\begin{bmatrix} 4 \\ 8 \\ 1 \end{bmatrix}$$

$$= \frac{1}{9} \begin{bmatrix} 18 \\ 36 \\ 45 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 5 \end{bmatrix}$$