April 20 Math 3260 sec. 51 Spring 2022

Section 6.1: Inner Product, Length, and Orthogonality

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

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

Remark: Note that this product produces a scalar. It is sometimes called a *scalar product*.

The Norm

Definition: The **norm** of the vector $\mathbf{v} = (v_1, \ldots, v_n)$ in $\mathbb{R}^n$ is the nonnegative number

$$\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$$

Theorem: For vector $\mathbf{v}$ in $\mathbb{R}^n$ and scalar c

$$\|c\mathbf{v}\| = |c|\|\mathbf{v}\|.$$

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

Given any nonzero vector $\mathbf{v}$ in $\mathbb{R}^n$, *normalizing* the vector means finding the unit vector in the direction of $\mathbf{v}$

$$\mathbf{v}/\|\mathbf{v}\|$$

Distance in $\mathbb{R}^n$

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

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

and is defined by

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

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),

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

Example

Find the distance between the vectors $\mathbf{u} = (4, 0, -1, 1)$ and $\mathbf{v} = (0, 0, 2, 7)$ in $\mathbb{R}^4$.

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

$$\vec{u} - \vec{v} = (4-0,\ 0-0,\ -1-2,\ 1-7) = (4, 0, -3, -6)$$

$$\text{dist}(\vec{u}, \vec{v}) = \sqrt{4^2 + 0^2 + (-3)^2 + (-6)^2} = \sqrt{61}$$

Orthogonality

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

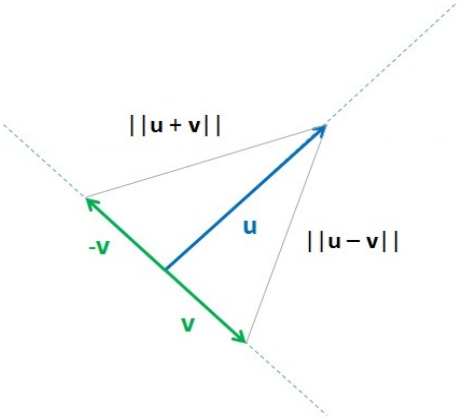

Figure: Note that two vectors are perpendicular if $\|\mathbf{u} - \mathbf{v}\| = \|\mathbf{u} + \mathbf{v}\|$

Orthogonal and Perpendicular

Show that $\|\mathbf{u} - \mathbf{v}\| = \|\mathbf{u} + \mathbf{v}\|$ if and only if $\mathbf{u} \cdot \mathbf{v} = 0$.

Consider

$$\|\vec{u} - \vec{v}\|^2 = (\vec{u} - \vec{v}) \cdot (\vec{u} - \vec{v})$$

$$= \vec{u} \cdot \vec{u} - \vec{v} \cdot \vec{u} - \vec{u} \cdot \vec{v} + \vec{v} \cdot \vec{v}$$

$$= \|\vec{u}\|^2 - 2\vec{u} \cdot \vec{v} + \|\vec{v}\|^2$$

$$\|\vec{u} + \vec{v}\|^2 = (\vec{u} + \vec{v}) \cdot (\vec{u} + \vec{v})$$

$$= \vec{u} \cdot \vec{u} + \vec{v} \cdot \vec{u} + \vec{u} \cdot \vec{v} + \vec{v} \cdot \vec{v}$$

$$= \|\vec{u}\|^2 + 2\vec{u} \cdot \vec{v} + \|\vec{v}\|^2$$

Suppose $\vec{u}\cdot\vec{v} = 0$. Then

$$\|\vec{u}-\vec{v}\|^2 = \|\vec{u}+\vec{v}\|^2$$

$$\Rightarrow \|\vec{u}-\vec{v}\| = \|\vec{u}+\vec{v}\|$$ as these are nonnegative

If $\|\vec{u}-\vec{v}\| = \|\vec{u}+\vec{v}\|$, then

$$\|\vec{u}-\vec{v}\|^2 = \|\vec{u}+\vec{v}\|^2 \Rightarrow$$

$$\|\vec{u}\|^2 - 2\vec{u}\cdot\vec{v} + \|\vec{v}\|^2 = \|\vec{u}\|^2 + 2\vec{u}\cdot\vec{v} + \|\vec{v}\|^2$$

$$\Rightarrow \quad 0 = 4\vec{u}\cdot\vec{v}$$

$$\Rightarrow \quad \vec{u}\cdot\vec{v} = 0$$

The Pythagorean Theorem

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

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

This follows immediately from the observation that

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

The two vectors are defined as being orthogonal precisely when $\mathbf{u}\cdot\mathbf{v} = 0$.

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 $W^\perp$.

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

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

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

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

Let's find a representation for $\vec{x}$ in Nul(A).

$$[A\ \vec{0}] = \begin{bmatrix} 1 & 3 & 2 & 0 \\ -2 & 0 & 4 & 0 \end{bmatrix} \xrightarrow{\text{rref}} \begin{bmatrix} 1 & 0 & -2 & 0 \\ 0 & 1 & \frac{4}{3} & 0 \end{bmatrix}$$

$x_1 = 2x_3$

$x_2 = -\frac{4}{3}x_3$

x_3 - free

$$\vec{x} = x_3 \begin{bmatrix} 2 \\ -\frac{4}{3} \\ 1 \end{bmatrix} = t \begin{bmatrix} 6 \\ -4 \\ 3 \end{bmatrix} \text{ for } t \text{ in } \mathbb{R}$$

We can use the rows of A as a

spanning set for Row(A).

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

$\vec{x}$ in Nul(A) looks like $t\begin{bmatrix}6\\-4\\3\end{bmatrix} = \vec{x}$.

Note

$$\vec{x}\cdot\vec{u}_1 = t\begin{bmatrix}6\\-4\\3\end{bmatrix}\cdot\begin{bmatrix}1\\3\\2\end{bmatrix} = t(6-12+6) = 0$$

$$\vec{x}\cdot\vec{u}_2 = t\begin{bmatrix}6\\-4\\3\end{bmatrix}\cdot\begin{bmatrix}-2\\0\\4\end{bmatrix} = t(-12+0+12) = 0$$

Theorem

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 ~~orthongal~~ 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: Find the orthogonal complement of Col(A)

$$A = \begin{bmatrix} 5 & 2 & 1 \\ -3 & 3 & 0 \\ 2 & 4 & 1 \\ 2 & -2 & 9 \\ 0 & 1 & -1 \end{bmatrix}$$

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

$$A^T = \begin{bmatrix} 5 & -3 & 2 & 2 & 0 \\ 2 & 3 & 4 & -2 & 1 \\ 1 & 0 & 1 & 9 & -1 \end{bmatrix}$$

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

$$\xrightarrow{\text{rref}} \begin{bmatrix} 1 & 0 & 0 & -54 & 7 \\ 0 & 1 & 0 & -146/3 & 19/3 \\ 0 & 0 & 1 & 63 & -8 \end{bmatrix}$$

$$x_1 = 54x_4 - 7x_5$$

$$x_2 = \frac{146}{3}x_4 - \frac{19}{3}x_5$$

$$x_3 = -63x_4 + 8x_5$$

x_4, x_5 are free

$$\vec{x} = x_4 \begin{bmatrix} 54 \\ \frac{146}{3} \\ -63 \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -7 \\ \frac{-19}{3} \\ 8 \\ 0 \\ 1 \end{bmatrix}$$

$$[\mathrm{Col}(A)]^{\perp} = \mathrm{Span}\left\{ \begin{bmatrix} 54 \\ \frac{146}{3} \\ -63 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -7 \\ \frac{-19}{3} \\ 8 \\ 0 \\ 1 \end{bmatrix} \right\}$$

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 $\left\{ \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix} \right\}$ is an orthogonal set.

Call these $\vec{u}_1$, $\vec{u}_2$, $\vec{u}_3$ in the order shown.

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

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

$$\left\{ \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix} \right\}$$

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

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

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

$$c_j = \frac{\mathbf{y} \cdot \mathbf{u}_j}{\mathbf{u}_j \cdot \mathbf{u}_j}.$$

Example

$\left\{ \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix} \right\}$ 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.

$$\vec{y} = c_1\vec{u}_1 + c_2\vec{u}_2 + c_3\vec{u}_3$$

$$c_1 = \frac{\vec{y}\cdot\vec{u}_1}{\|\vec{u}_1\|^2} = \frac{-3}{11}, \quad c_2 = \frac{\vec{y}\cdot\vec{u}_2}{\|\vec{u}_2\|^2} = \frac{8}{6} = \frac{4}{3}$$

$$c_3 = \frac{\vec{y}\cdot\vec{u}_3}{\|\vec{u}_3\|^2} = \frac{-10}{66} = \frac{-5}{33}.$$

$$\begin{bmatrix} -2 \\ 3 \\ 0 \end{bmatrix} = \frac{-3}{11}\begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} + \frac{4}{3}\begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} - \frac{5}{33}\begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix}$$