Mathematics 160: Lecture 31

Orthogonality

Dan Sloughter

Furman University

November 30, 2011

Triangle inequality

- For any two vectors $\vec{x}$ and $\vec{y}$ in $\mathbb{R}^n$,

$$\|\vec{x}+\vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|.$$

- Reason: Using the Cauchy-Schwarz inequality,

$$\begin{aligned}\|\vec{x}+\vec{y}\|^2 &= (\vec{x}+\vec{y})\cdot(\vec{x}+\vec{y})\\ &= \vec{x}\cdot\vec{x} + 2\vec{x}\cdot\vec{y} + \vec{y}\cdot\vec{y}\\ &\leq \|\vec{x}\|^2 + 2\|\vec{x}\|\|\vec{y}\| + \|\vec{y}\|^2\\ &= (\|\vec{x}\| + \|\vec{y}\|)^2.\end{aligned}$$

Distance in $\mathbb{R}^n$

- Recall: if $\vec{x}$ and $\vec{y}$ are two points in $\mathbb{R}^n$, then the distance between $\vec{x}$ and $\vec{y}$ is $\|\vec{x}-\vec{y}\|$.
- Triangle inequality: For any three vectors $\vec{x}$, $\vec{y}$, and $\vec{z}$ in $\mathbb{R}^n$,

$$\|\vec{x}-\vec{y}\| = \|(\vec{x}-\vec{z}) + (\vec{z}-\vec{y})\| \leq \|\vec{x}-\vec{z}\| + \|\vec{z}-\vec{y}\|.$$

- That is, the length of any side of a triangle is less than or equal to the sum of the lengths of the other two sides of the triangle.

Definition

- We call a set of nonzero vectors $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\}$ in $\mathbb{R}^n$ *orthogonal* if $\vec{x}_i \cdot \vec{x}_j = 0$ for all $i \neq j$.
- If, in addition, $\|\vec{x}_i\| = 1$ for $i = 1, 2, \ldots, n$, then we say the set is *orthonormal*.
- Note: if $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\}$ is an orthogonal set of vectors in $\mathbb{R}^n$, then

$$\left\{\frac{1}{\|\vec{x}_1\|}\vec{x}_1, \frac{1}{\|\vec{x}_2\|}\vec{x}_2, \ldots, \frac{1}{\|\vec{x}_k\|}\vec{x}_k\right\}$$

is an orthonormal set.

Pythagorean Theorem

- If $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\}$ is an orthogonal set of vectors in $\mathbb{R}^n$, then

$$\|\vec{x}_1 + \vec{x}_2 + \cdots + \vec{x}_k\|^2 = \|\vec{x}_1\|^2 + \|\vec{x}_2\|^2 + \cdots + \|\vec{x}_k\|^2.$$

- Reason: Since $\vec{x}_i \cdot \vec{x}_j = 0$ when $i \neq j$,

$$\begin{aligned}\|\vec{x}_1 + \vec{x}_2 + \cdots + \vec{x}_k\|^2 &= (\vec{x}_1 + \vec{x}_2 + \cdots + \vec{x}_k) \cdot (\vec{x}_1 + \vec{x}_2 + \cdots + \vec{x}_k) \\ &= \sum_{i=1}^{k} \vec{x}_i \cdot \vec{x}_i + 2\sum_{i<j} \vec{x}_i \cdot \vec{x}_j \\ &= \sum_{i=1}^{k} \|\vec{x}_i\|^2.\end{aligned}$$

Theorem

- If $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\}$, is an orthogonal set of vectors in $\mathbb{R}^n$, then $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\}$ is linearly independent.
- Reason:
  - Suppose, for some scalars $t_1, t_2, \ldots, t_k$,

$$t_1\vec{x}_1 + t_2\vec{x}_2 + \cdots + t_k\vec{x}_k = \vec{0}.$$

  - Then for $i = 1, 2, \ldots, k$,

$$\begin{aligned}0 &= \vec{x}_i \cdot \vec{0} \\ &= \vec{x}_i \cdot (t_1\vec{x}_1 + t_2\vec{x}_2 + \cdots + t_k\vec{x}_k) \\ &= t_1(\vec{x}_1 \cdot \vec{x}_i) + t_2(\vec{x}_2 \cdot \vec{x}_i) + \cdots + t_k(\vec{x}_k \cdot \vec{x}_i) \\ &= t_i\|\vec{x}_i\|^2.\end{aligned}$$

  - Hence $t_i = 0$.

Theorem

- Suppose $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\}$ is an orthogonal basis for a subspace U in $\mathbb{R}^n$. If

$$\vec{x} = t_1\vec{x}_1 + t_2\vec{x}_2 + \cdots + t_k\vec{x}_k$$

  is a vector in U, then

$$t_i = \frac{\vec{x} \cdot \vec{x}_i}{\|\vec{x}_i\|^2}.$$

- Reason: the result follows from

$$\begin{aligned}\vec{x} \cdot \vec{x}_i &= (t_1\vec{x}_1 + t_2\vec{x}_2 + \cdots + t_k\vec{x}_k) \cdot \vec{x}_i \\ &= t_1(\vec{x}_1 \cdot \vec{x}_i) + t_2(\vec{x}_2 \cdot \vec{x}_i) + \cdots + t_k(\vec{x}_k \cdot \vec{x}_i) \\ &= t_i\|\vec{x}_i\|^2.\end{aligned}$$

- Note: $\dfrac{\vec{x} \cdot \vec{x}_i}{\|\vec{x}_i\|^2}\vec{x}_i = \operatorname{proj}_{\vec{x}_i} \vec{x}$.

Orthonormal bases

- Note: if $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\}$ is an orthonormal basis for a subspace U in $\mathbb{R}^n$ and $\vec{x}$ is any vector in U, then

$$\vec{x} = (\vec{x} \cdot \vec{x}_1)\vec{x}_1 + (\vec{x} \cdot \vec{x}_2)\vec{x}_2 + \cdots + (\vec{x} \cdot \vec{x}_k) \cdot \vec{x}_k.$$

- Note: $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$ is an orthonormal basis for $\mathbb{R}^n$.
- Note: if $\vec{x} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}^T$, then $\vec{x} \cdot \vec{e}_i = x_i$.
- Hence this result generalizes the representation

$$\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2 + \cdots + x_n\vec{e}_n.$$

Example

- Suppose
$$\vec{x}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \vec{x}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \text{ and } \vec{x}_3 = \begin{bmatrix} 0 \\ -2 \\ 0 \end{bmatrix}.$$
- Then $\{\vec{x}_1, \vec{x}_2, \vec{x}_3\}$ is an orthogonal set, and, hence, an orthogonal basis for $\mathbb{R}^3$.
- Then $\{\vec{z}_1, \vec{z}_2, \vec{z}_3\}$, where
$$\vec{z}_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \vec{z}_2 = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \text{ and } \vec{z}_3 = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix},$$
is an orthonormal basis for $\mathbb{R}^3$.

Example (cont'd)

- Then, for example, if
$$\vec{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix},$$
then
$$\vec{x} = \frac{4}{\sqrt{2}}\vec{z}_1 + \frac{2}{\sqrt{2}}\vec{z}_2 - 2\vec{z}_3 = 2\vec{x}_1 + \vec{x}_2 - \vec{x}_3.$$