

Orthogonal sets

Math 4A – Scharlemann

9 March 2015



CELL PHONES OFF

Recall: Suppose $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_m\} \subset \mathbb{V}$ is a set of **linearly independent** vectors. Then any **linear combination** can be written in **exactly one** way. In other words:

$$c_1\vec{u}_1 + c_2\vec{u}_2 + \dots + c_m\vec{u}_m = d_1\vec{u}_1 + d_2\vec{u}_2 + \dots + d_m\vec{u}_m \implies \text{each } c_i = d_i.$$

Alternate way of saying this: If $\vec{u} \in \text{Span}\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_m\}$ then there is a **unique** solution $\vec{x} \in \mathbb{R}^m$ to

$$\vec{u} = [\vec{u}_1 \quad \vec{u}_2 \quad \dots \quad \vec{u}_m] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$$

We know how to find the solution, but it can be a lot of work.

Magical special case:

Definition

A set of non-zero vectors $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_m\} \subset \mathbb{R}^n$ is an **orthogonal set** if each pair of them is a perpendicular (orthogonal) pair. That is, each $\vec{u}_i \neq \vec{0}$ and for any pair $\vec{u}_i, \vec{u}_j, i \neq j$ we have $\vec{u}_i \cdot \vec{u}_j = 0$.

Example: The set of vectors

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

is an orthogonal set.

Proof: Need to check three dot products. Here's one of them:

$$\vec{u}_1 \cdot \vec{u}_3 = 1 \cdot 3 + 2 \cdot (-2) + 1 \cdot 1 = 3 - 4 + 1 = 0.$$

iClicker question

Consider the set of vectors

$$\vec{u}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \quad \vec{u}_2 = \begin{bmatrix} 2 \\ 1 \\ -4 \end{bmatrix}, \quad \vec{u}_3 = \begin{bmatrix} 5 \\ -1 \\ -3 \end{bmatrix}.$$

- A) This is an orthogonal set of vectors
- B) This is not an orthogonal set of vectors because $\vec{u}_1 \cdot \vec{u}_1 \neq 0$.
- C) This is not an orthogonal set of vectors because $\vec{u}_2 \cdot \vec{u}_2 \neq 1$.
- D) This is not an orthogonal set of vectors because $\vec{u}_1 \cdot \vec{u}_2 \neq 0$.
- E) This is not an orthogonal set of vectors because $\vec{u}_3 \cdot \vec{u}_2 \neq 0$.

Theorem

If a set of vectors $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_m\} \subset \mathbb{R}^n$ is an *orthogonal set* then it is a *linearly independent* set.

Proof: Remember what we have to show: the only solution to the equation $c_1\vec{u}_1 + c_2\vec{u}_2 + \dots + c_m\vec{u}_m = \vec{0}$ is this: each $c_i = 0$.

So suppose $c_1\vec{u}_1 + c_2\vec{u}_2 + \dots + c_m\vec{u}_m = \vec{0}$ and pick any i . Calculate:

$$\begin{aligned} 0 &= \vec{u}_i \cdot \vec{0} = \vec{u}_i \cdot (c_1\vec{u}_1 + c_2\vec{u}_2 + \dots + c_m\vec{u}_m) = \\ &= c_1(\vec{u}_i \cdot \vec{u}_1) + c_2(\vec{u}_i \cdot \vec{u}_2) + \dots + c_m(\vec{u}_i \cdot \vec{u}_m) \end{aligned}$$

Since each $\vec{u}_i \cdot \vec{u}_j = 0$ when $j \neq i$ this becomes simply

$0 = c_i(\vec{u}_i \cdot \vec{u}_i)$. Since $\vec{u}_i \cdot \vec{u}_i = \|\vec{u}_i\|^2 \neq 0$, get $c_i = 0$.



Magical fact: If $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_m\} \subset \mathbb{R}^n$ is an **orthogonal set** then it's easy to calculate the weights c_i for any $\vec{v} \in \text{Span}\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_m\}$: It's based on the same idea: Suppose you know that $\vec{v} \in \text{Span}\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_m\}$ and you want to find the c_i so that

$$\vec{v} = c_1 \vec{u}_1 + c_2 \vec{u}_2 + \dots + c_m \vec{u}_m.$$

For each $i = 1, \dots, m$, see what happens when you take the dot product of both sides with $\vec{u}_i$:

$$\begin{aligned} \vec{u}_i \cdot \vec{v} &= \vec{u}_i \cdot (c_1 \vec{u}_1 + c_2 \vec{u}_2 + \dots + c_m \vec{u}_m) = \\ &= c_1(\vec{u}_i \cdot \vec{u}_1) + c_2(\vec{u}_i \cdot \vec{u}_2) + \dots + c_m(\vec{u}_i \cdot \vec{u}_m) \end{aligned}$$

Since each $\vec{u}_i \cdot \vec{u}_j = 0$ when $j \neq i$ this becomes simply

$$\vec{u}_i \cdot \vec{v} = c_i(\vec{u}_i \cdot \vec{u}_i).$$

Divide by $\vec{u}_i \cdot \vec{u}_i = \|\vec{u}_i\|^2 \neq 0$ and get

$$c_i = \frac{\vec{u}_i \cdot \vec{v}}{\|\vec{u}_i\|^2} = \frac{\vec{u}_i \cdot \vec{v}}{\vec{u}_i \cdot \vec{u}_i}$$

Example: Begin with the orthogonal set

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

Problem: Find c_1, c_2, c_3 so that

$$\begin{bmatrix} 7 \\ 1 \\ 9 \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 2 \\ 1 \\ -4 \end{bmatrix} + c_3 \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix}$$

The formula says

$$c_1 = \frac{\vec{u}_1 \cdot \vec{v}}{\|\vec{u}_1\|^2}, \quad c_2 = \frac{\vec{u}_2 \cdot \vec{v}}{\|\vec{u}_2\|^2}, \quad c_3 = \frac{\vec{u}_3 \cdot \vec{v}}{\|\vec{u}_3\|^2}$$

Calculations:

$$\vec{u}_1 \cdot \vec{u}_1 = 1^2 + 2^2 + 1^2 = 6 \quad \vec{u}_1 \cdot \vec{v} = 7 + 2 + 9 = 18 \implies c_1 = \frac{18}{6} = 3$$

$$\vec{u}_2 \cdot \vec{u}_2 = 4 + 1 + 16 = 21 \quad \vec{u}_2 \cdot \vec{v} = 14 + 1 - 36 = -21 \implies c_2 = -1$$

$$\vec{u}_3 \cdot \vec{u}_3 = 9 + 4 + 1 = 14 \quad \vec{u}_3 \cdot \vec{v} = 21 - 2 + 9 = 28 \implies c_3 = \frac{28}{14} = 2$$

(Reminder: usually these won't be whole numbers.) Thus

$$\text{Answer : } \begin{bmatrix} 7 \\ 1 \\ 9 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} - 1 \begin{bmatrix} 2 \\ 1 \\ -4 \end{bmatrix} + 2 \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix}$$

Reminder: You'd be nuts not to check this.

iClicker question

Consider this pair of vectors in $\mathbb{R}^4$:

$$\vec{u}_1 = \begin{bmatrix} 1 \\ 2 \\ k \\ 3 \end{bmatrix}, \quad \vec{u}_2 = \begin{bmatrix} 3 \\ k \\ 7 \\ -4 \end{bmatrix}$$

- A) This pair is always an orthogonal set of vectors
- B) This pair is orthogonal when $k = 1$.
- C) This pair is orthogonal when $k = 2$.
- D) This pair is orthogonal when $k = 3$.
- E) This pair is never an orthogonal set of vectors

Discussion

Calculation:

$$\vec{u}_1 \cdot \vec{u}_2 = \begin{bmatrix} 1 \\ 2 \\ k \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ k \\ 7 \\ -4 \end{bmatrix} = 3 + 2k + 7k - 12 = 9k - 9.$$

Thus the pair of vectors is orthogonal $\iff \vec{u}_1 \cdot \vec{u}_2 = 9k - 9 = 0$,
i. e. $k = 1 \implies$ answer (B).

If the $\vec{u}_i$ are orthogonal and **also** have the property that each $\vec{u}_i \cdot \vec{u}_i = 1$ (i. e. $\|\vec{u}_i\| = 1$; each vector is **unit length**) then it's called an orthonormal set.

Example: The standard basis $\vec{e}_1, \dots, \vec{e}_n$ of $\mathbb{R}^n$ is orthonormal. For an orthonormal set $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_m\}$, notice

$$c_i = \frac{\vec{u}_i \cdot \vec{v}}{\|\vec{u}_i\|^2} = \frac{\vec{u}_i \cdot \vec{v}}{1} = \vec{u}_i \cdot \vec{v}$$

The formula $c_i = \vec{u}_i \cdot \vec{v}$ is really easy, so orthonormal sets are great.

Can turn an orthogonal set into an orthonormal set just by dividing each vector by its length.

Example: We have found that

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

is an orthogonal set.

Moreover we know

$$||\vec{u}_1||^2 = \vec{u}_1 \cdot \vec{u}_1 = 6 \quad ||\vec{u}_2||^2 = \vec{u}_2 \cdot \vec{u}_2 = 21 \quad ||\vec{u}_3||^2 = \vec{u}_3 \cdot \vec{u}_3 = 14$$

So this orthogonal set determines the orthonormal set

$$\frac{\vec{u}_1}{||\vec{u}_1||} = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \quad \frac{\vec{u}_2}{||\vec{u}_2||} = \frac{1}{\sqrt{21}} \begin{bmatrix} 2 \\ 1 \\ -4 \end{bmatrix}, \quad \frac{\vec{u}_3}{||\vec{u}_3||} = \frac{1}{\sqrt{14}} \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix}$$

Some fun facts you can check:

- A square matrix U has orthonormal **rows** if and only if it has orthonormal **columns**
- This happens if and only if $U^T U = I$.

Example: We have seen above that the matrix

$$U = \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{21}} & \frac{1}{\sqrt{21}} & \frac{-4}{\sqrt{21}} \\ \frac{3}{\sqrt{14}} & \frac{-2}{\sqrt{14}} & \frac{1}{\sqrt{14}} \end{bmatrix}$$

has orthonormal rows.

Therefore the **columns** are also orthonormal, and $U^T U = I_3$.

Random check, showing first and third columns are orthogonal:

$$\begin{bmatrix} \frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{21}} \\ \frac{3}{\sqrt{14}} \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{\sqrt{6}} \\ \frac{-4}{\sqrt{21}} \\ \frac{1}{\sqrt{14}} \end{bmatrix} = \frac{1}{6} - \frac{8}{21} + \frac{3}{14} = \frac{7}{42} - \frac{16}{42} + \frac{9}{42} = 0$$

$U^T U = I_m$ discussion

Why is this true?

Remember this alternate definition of inner product: $\vec{u} \cdot \vec{v} = \vec{u}^T \vec{v}$.

Suppose $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_m\}$ is an **orthonormal** set of vectors and let

$$U = [\vec{u}_1 \quad \vec{u}_2 \quad \dots \quad \vec{u}_m]$$

be the $m \times m$ matrix with these as columns.

Then U^T is the $m \times m$ matrix whose **rows** are $\{\vec{u}_1^T, \vec{u}_2^T, \dots, \vec{u}_m^T\}$.

In particular,

$$U^T U = \begin{bmatrix} \vec{u}_1^T \\ \vec{u}_2^T \\ \vdots \\ \vec{u}_m^T \end{bmatrix} [\vec{u}_1 \quad \vec{u}_2 \quad \dots \quad \vec{u}_m]$$

$U^T U = I_m$ discussion

$$U^T U = \begin{bmatrix} \vec{u}_1^T \\ \vec{u}_2^T \\ \vdots \\ \vec{u}_m^T \end{bmatrix} \begin{bmatrix} \vec{u}_1 & \vec{u}_2 & \dots & \vec{u}_m \end{bmatrix}$$

The i, j -th entry, $(U^T U)_{i,j}$ is

$$\vec{u}_i^T \vec{u}_j = \vec{u}_i \cdot \vec{u}_j = \begin{cases} 1 & \text{when } i = j \\ 0 & \text{when } i \neq j \end{cases} \text{ since } U \text{ is orthonormal.}$$

But that's the definition of the matrix I_m .

Another way to say this: for an **orthonormal** matrix U , $U^{-1} = U^T$.

iClicker question

$$A = \begin{bmatrix} 0 & 0 & -1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

What is A^{-1} ?

A) $\begin{bmatrix} 0 & 0 & -1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$

B) $\begin{bmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ -1 & 0 & 0 \end{bmatrix}$

C) $\begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$

D) $\begin{bmatrix} 0 & 0 & -1 \\ 0 & -1 & 0 \\ -1 & 0 & 0 \end{bmatrix}$

E) I can't answer this quickly.

A is clearly orthonormal, so its inverse is its transpose.

Second example: The matrix $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ is clearly orthonormal, so $A^{-1} = A^T = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. Another example where $A = -A^{-1}$!

Suppose U is an $m \times m$ orthonormal matrix. Consider the linear transformation $\mathbb{R}^m \rightarrow \mathbb{R}^m$ given by U . What is the effect on the dot product of two vectors?

That is, how do $\vec{u} \cdot \vec{v}$ and $(U\vec{u}) \cdot (U\vec{v})$ compare? Answer:

$$(U\vec{u}) \cdot (U\vec{v}) = (U\vec{u})^T (U\vec{v}) = (\vec{u}^T U^T)(U\vec{v}) = \vec{u}^T (U^T U) \vec{v} = \vec{u}^T \vec{v} = \vec{u} \cdot \vec{v}$$

In other words, U **doesn't change** the dot product.

So U **doesn't change** length, distance, angle, i. e. the **geometry**!

Example:

We can see that **rotation** of $\mathbb{R}^2$ doesn't change angle or distance. Is that because rotation is given by an orthonormal matrix?

Check:

Let $R_\theta = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$. We have seen that this matrix gives rotation by angle θ .

$$R_\theta R_\theta^T = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} =$$

$$\begin{bmatrix} \cos^2 \theta + \sin^2 \theta & \cos \theta \sin \theta - \sin \theta \cos \theta \\ \sin \theta \cos \theta - \cos \theta \sin \theta & \sin^2 \theta + \cos^2 \theta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2$$

Therefore indeed R_θ is orthonormal.