Lecture 16 recap

1) Definition of a diagonalizable matrix.

2) A necessary and sufficient condition for a matrix to be diagonalizable.

3) An algorithm to determine if a matrix can be diagonalized, and if so, to find a matrix that diagonalizes it.

4) Several remarks.

5) A sufficient condition for diagonalizability.

6) An example on Fibonacci numbers and two ways to solve it.

LECTURE 17

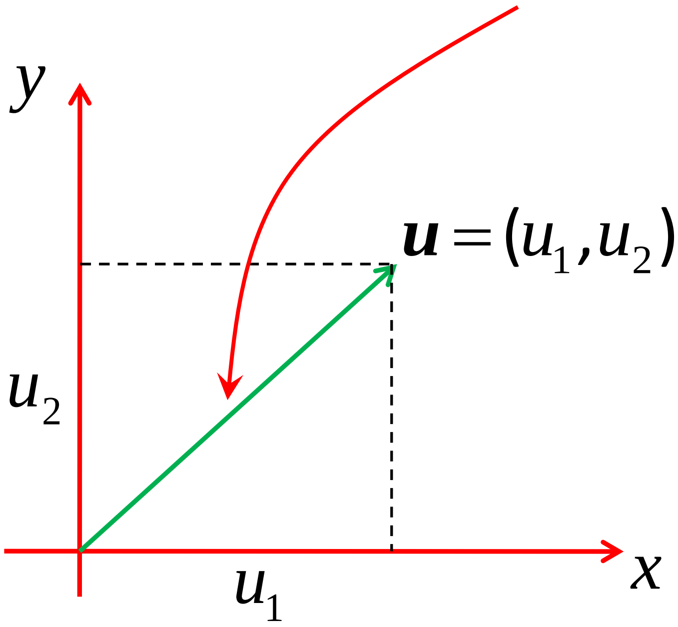
Inner products in $\mathbf{R}^n$

Orthogonal and Orthonormal Bases

What is the length of a vector?

If $\boldsymbol{u}=(u_1,u_2)$ is a vector in $\mathbb{R}^2$, the length of $\boldsymbol{u}$ is defined to be

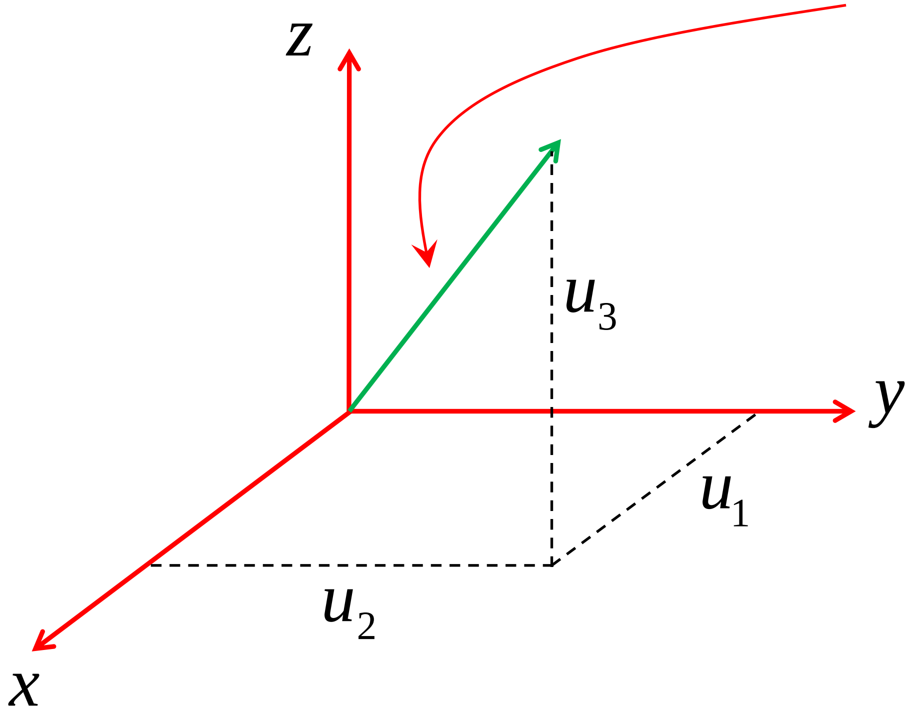
$$\|\boldsymbol{u}\|=\sqrt{u_1^2+u_2^2}$$



What is the length of a vector?

If $\boldsymbol{u} = (u_1, u_2, u_3)$ is a vector in $\mathbb{R}^3$, the length of $\boldsymbol{u}$ is defined to be

$$\|\boldsymbol{u}\| = \sqrt{u_1^2 + u_2^2 + u_3^2}$$



How do you think we should define the length of a vector $\boldsymbol{u}$ in $\mathbb{R}^n$?

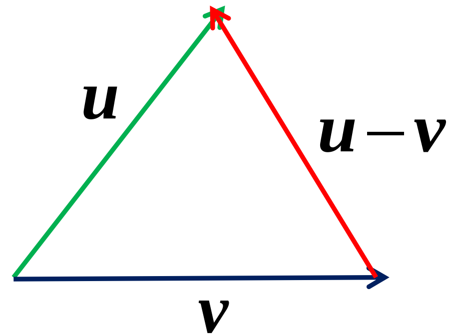
The length of a vector in R^n

If $\boldsymbol{u} = (u_1, u_2, \ldots, u_n)$ is a vector in $\mathbb{R}^n$, the length of $\boldsymbol{u}$ is defined to be

$$\|\boldsymbol{u}\| = \sqrt{u_1^2 + u_2^2 + \ldots + u_n^2}$$

The distance between 2 vectors

If $\boldsymbol{u}$ and $\boldsymbol{v}$ are two vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$, the distance between $\boldsymbol{u}$ and $\boldsymbol{v}$ is defined to be the length of the vector $\boldsymbol{u}-\boldsymbol{v}$.



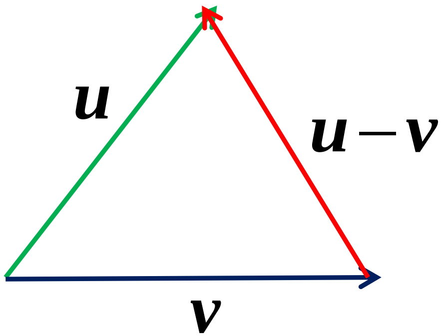
distance between $\boldsymbol{u}$ and $\boldsymbol{v}$

$$= d(\boldsymbol{u},\boldsymbol{v}) = \|\boldsymbol{u}-\boldsymbol{v}\|$$

The definition is similar when $\boldsymbol{u},\boldsymbol{v}$ are vectors in $\mathbb{R}^n$.

The distance between 2 vectors in R^2 or R^3

If $\boldsymbol{u}$ and $\boldsymbol{v}$ are two vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$, the distance between $\boldsymbol{u}$ and $\boldsymbol{v}$ is defined to be the length of the vector $\boldsymbol{u}-\boldsymbol{v}$.



distance between $\boldsymbol{u}$ and $\boldsymbol{v}$

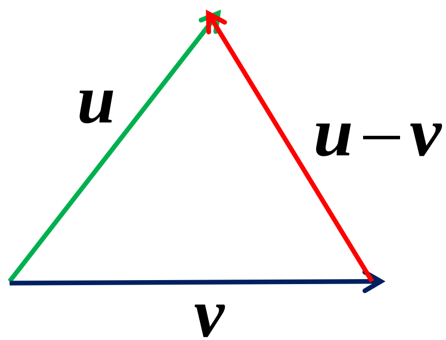
$$= d(\boldsymbol{u},\boldsymbol{v}) = \|\boldsymbol{u}-\boldsymbol{v}\|$$

If $\boldsymbol{u}=(u_1,u_2)$ and $\boldsymbol{v}=(v_1,v_2)$ are vectors in $\mathbb{R}^2$,

$$d(\boldsymbol{u},\boldsymbol{v}) = \|\boldsymbol{u}-\boldsymbol{v}\| = \sqrt{(u_1-v_1)^2+(u_2-v_2)^2}$$

The distance between 2 vectors in R^2 or R^3

If $\boldsymbol{u}$ and $\boldsymbol{v}$ are two vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$, the distance between $\boldsymbol{u}$ and $\boldsymbol{v}$ is defined to be the length of the vector $\boldsymbol{u}-\boldsymbol{v}$.



distance between $\boldsymbol{u}$ and $\boldsymbol{v}$

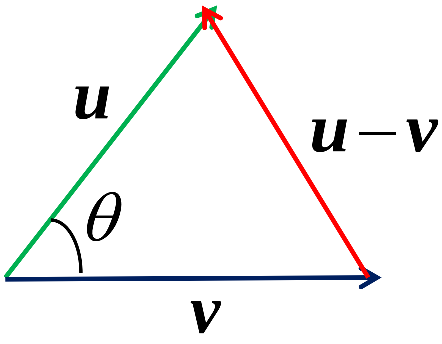
$$= d(\boldsymbol{u},\boldsymbol{v}) = \|\boldsymbol{u}-\boldsymbol{v}\|$$

If $\boldsymbol{u}=(u_1,u_2,u_3)$ and $\boldsymbol{v}=(v_1,v_2,v_3)$ are vectors in $\mathbb{R}^3$,

$$d(\boldsymbol{u},\boldsymbol{v}) = \|\boldsymbol{u}-\boldsymbol{v}\| = \sqrt{(u_1-v_1)^2+(u_2-v_2)^2+(u_3-v_3)^2}$$

Angle between 2 vectors in R^2 or R^3

If $\boldsymbol{u}$ and $\boldsymbol{v}$ are two vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$, let the angle between $\boldsymbol{u}$ and $\boldsymbol{v}$ be θ.



By cosine rule,

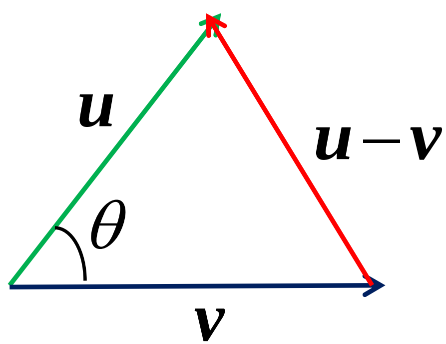
$$\|\boldsymbol{u}-\boldsymbol{v}\|^2 = \|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2 - 2\|\boldsymbol{u}\|\|\boldsymbol{v}\|\cos\theta$$

$$\Rightarrow \cos\theta = \frac{\|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2 - \|\boldsymbol{u}-\boldsymbol{v}\|^2}{2\|\boldsymbol{u}\|\|\boldsymbol{v}\|}$$

$$\Rightarrow \theta = \cos^{-1}\left(\frac{\|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2 - \|\boldsymbol{u}-\boldsymbol{v}\|^2}{2\|\boldsymbol{u}\|\|\boldsymbol{v}\|}\right)$$

Angle between 2 vectors in R^2 or R^3

If $\boldsymbol{u}$ and $\boldsymbol{v}$ are two vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$, let the angle between $\boldsymbol{u}$ and $\boldsymbol{v}$ be θ.



$$\theta = \cos^{-1}\left(\frac{\|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2 - \|\boldsymbol{u}-\boldsymbol{v}\|^2}{2\|\boldsymbol{u}\|\|\boldsymbol{v}\|}\right)$$

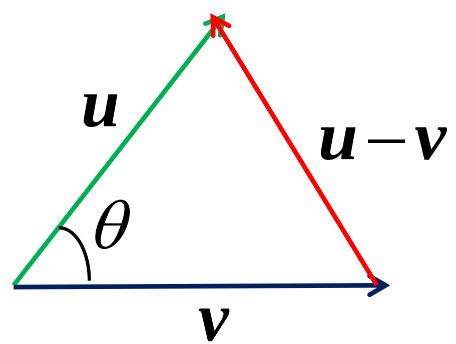
If $\boldsymbol{u} = (u_1, u_2)$ and $\boldsymbol{v} = (v_1, v_2)$ belong to $\mathbb{R}^2$,

$$\theta = \cos^{-1}\left(\frac{u_1^2 + u_2^2 + v_1^2 + v_2^2 - (u_1 - v_1)^2 - (u_2 - v_2)^2}{2\|\boldsymbol{u}\|\|\boldsymbol{v}\|}\right)$$

$$= \cos^{-1}\left(\frac{u_1 v_1 + u_2 v_2}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|}\right)$$

Angle between 2 vectors in R^2 or R^3

If $\boldsymbol{u}$ and $\boldsymbol{v}$ are two vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$, let the angle between $\boldsymbol{u}$ and $\boldsymbol{v}$ be θ.



$$\boxed{\theta = \cos^{-1}\left(\frac{\|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2 - \|\boldsymbol{u}-\boldsymbol{v}\|^2}{2\|\boldsymbol{u}\|\|\boldsymbol{v}\|}\right)}$$

If $\boldsymbol{u} = (u_1, u_2, u_3)$ and $\boldsymbol{v} = (v_1, v_2, v_3)$ belong to $\mathbb{R}^3$,

$$\theta = \cos^{-1}\left(\frac{u_1^2 + u_2^2 + u_3^2 + v_1^2 + v_2^2 + v_3^2 - (u_1 - v_1)^2 - (u_2 - v_2)^2 - (u_3 - v_3)^2}{2\|\boldsymbol{u}\|\|\boldsymbol{v}\|}\right)$$

$$= \cos^{-1}\left(\frac{u_1 v_1 + u_2 v_2 + u_3 v_3}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|}\right)$$

Angle between 2 vectors in R^2 or R^3

If $\boldsymbol{u}$ and $\boldsymbol{v}$ are two vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$, let the angle between $\boldsymbol{u}$ and $\boldsymbol{v}$ be θ.

$$\theta = \cos^{-1}\left(\frac{\|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2 - \|\boldsymbol{u}-\boldsymbol{v}\|^2}{2\|\boldsymbol{u}\|\|\boldsymbol{v}\|}\right)$$

If $\boldsymbol{u} = (u_1, u_2)$ and $\boldsymbol{v} = (v_1, v_2)$ belong to $\mathbb{R}^2$,

$$\theta = \cos^{-1}\left(\frac{\boxed{u_1v_1 + u_2v_2}}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|}\right)$$

If $\boldsymbol{u} = (u_1, u_2, u_3)$ and $\boldsymbol{v} = (v_1, v_2, v_3)$ belong to $\mathbb{R}^3$,

$$\theta = \cos^{-1}\left(\frac{\boxed{u_1v_1 + u_2v_2 + u_3v_3}}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|}\right)$$

Definitions (dot product, norm, distance, angle)

Let $\boldsymbol{u} = (u_1, u_2, \dots, u_n)$ and $\boldsymbol{v} = (v_1, v_2, \dots, v_n)$ be two vectors in $\mathbb{R}^n$.

1) The dot product (or inner product) of $\boldsymbol{u}$ and $\boldsymbol{v}$ is the value

$$\boxed{\boldsymbol{u} \cdot \boldsymbol{v} = u_1 v_1 + u_2 v_2 + \dots + u_n v_n}$$

2) The norm (or length) of $\boldsymbol{u}$ is

$$\boxed{\|\boldsymbol{u}\| = \sqrt{u_1^2 + u_2^2 + \dots + u_n^2}}$$

Vectors of norm 1 are called unit vectors.

Definitions (dot product, norm, distance, angle)

Let $\boldsymbol{u}=(u_1,u_2,\ldots,u_n)$ and $\boldsymbol{v}=(v_1,v_2,\ldots,v_n)$ be two vectors in $\mathbb{R}^n$.

3) The distance between $\boldsymbol{u}$ and $\boldsymbol{v}$ is

$$\boxed{d(\boldsymbol{u},\boldsymbol{v})=\|\boldsymbol{u}-\boldsymbol{v}\|=\sqrt{(u_1-v_1)^2+(u_2-v_2)^2+\ldots+(u_n-v_n)^2}}$$

4) The angle between $\boldsymbol{u}$ and $\boldsymbol{v}$ is

$$\boxed{\cos^{-1}\left(\frac{\boldsymbol{u}\cdot\boldsymbol{v}}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|}\right)}$$

Dot product and matrix product

Let $\boldsymbol{u} = (u_1, u_2, \ldots, u_n)$ and $\boldsymbol{v} = (v_1, v_2, \ldots, v_n)$ be two vectors in $\mathbb{R}^n$ (here $\boldsymbol{u}$ and $\boldsymbol{v}$ are written as row vectors).

$$\boldsymbol{u} \cdot \boldsymbol{v} = u_1 v_1 + u_2 v_2 + \ldots + u_n v_n$$

$$= \begin{pmatrix} u_1 & u_2 & \ldots & u_n \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \boldsymbol{u}\boldsymbol{v}^T$$

dot product

matrix product

Dot product and matrix product

Let $\boldsymbol{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}$ and $\boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$ be two vectors in $\mathbb{R}^n$

(here $\boldsymbol{u}$ and $\boldsymbol{v}$ are written as column vectors).

$$\boldsymbol{u} \cdot \boldsymbol{v} = u_1 v_1 + u_2 v_2 + \ldots + u_n v_n$$

$$= \begin{pmatrix} u_1 & u_2 & \ldots & u_n \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \boldsymbol{u}^T \boldsymbol{v}$$

matrix product

dot product

Example (dot product, norm, distance, angle)

Let $\boldsymbol{u} = (1, -2, 2, -1), \boldsymbol{v} = (1, 0, 2, 0)$.

Compute the following:

$\boldsymbol{u} \cdot \boldsymbol{v}$, $\|\boldsymbol{u}\|$, $\|\boldsymbol{v}\|$, $d(\boldsymbol{u}, \boldsymbol{v})$, angle between $\boldsymbol{u}$ and $\boldsymbol{v}$.

$\boldsymbol{u} \cdot \boldsymbol{v} =$

$\|\boldsymbol{u}\| =$

$\|\boldsymbol{v}\| =$

$d(\boldsymbol{u}, \boldsymbol{v}) =$

angle between $\boldsymbol{u}$ and $\boldsymbol{v} =$

Theorem (some results about dot products)

Let c be a scalar and $\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}$ be vectors in $\mathbb{R}^n$.

1) $\boldsymbol{u}\cdot\boldsymbol{v}=\boldsymbol{v}\cdot\boldsymbol{u}$

2) $(\boldsymbol{u}+\boldsymbol{v})\cdot\boldsymbol{w}=\boldsymbol{u}\cdot\boldsymbol{w}+\boldsymbol{v}\cdot\boldsymbol{w}$

$\boldsymbol{w}\cdot(\boldsymbol{u}+\boldsymbol{v})=\boldsymbol{w}\cdot\boldsymbol{u}+\boldsymbol{w}\cdot\boldsymbol{v}$

3) $(c\boldsymbol{u})\cdot\boldsymbol{v}=\boldsymbol{u}\cdot(c\boldsymbol{v})=c(\boldsymbol{u}\cdot\boldsymbol{v})$

4) $\|c\boldsymbol{u}\|=|c|\|\boldsymbol{u}\|$

5) $\boldsymbol{u}\cdot\boldsymbol{u}\geq 0$ and $\boldsymbol{u}\cdot\boldsymbol{u}=0$ if and only if $\boldsymbol{u}=\boldsymbol{0}$.

The dot product of any vector with itself is non-negative and the only vector whose dot product with itself is zero is the zero vector.

An application of dot products

Do you know that there are applications of linear algebra everywhere?

Not that I am aware of...can you give me an example?

Something we use almost everyday... internet search engine is one such example.

Information retrieval (search)

Problem: Suppose there are six documents that we would like to search through and determine which one of the documents is 'closest' to my search criteria, which is defined in terms of a few keywords.

D_1 D_2 D_3 D_4 D_5 D_6

Five keywords we may want to look for in our documents:

$$k_1, k_2, k_3, k_4, k_5.$$

Information retrieval (search)

Approach: We represent each document D_j by a vector $\boldsymbol{d}_j \in \mathbb{R}^5$ (because there are 5 keywords we are interested in).

D_1 D_2 D_3 D_4 D_5 D_6

$$\boldsymbol{d}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$

Document D_1 contains k_1, k_2, k_4, k_5 but not k_3

Five keywords we may want to look for: k_1, k_2, k_3, k_4, k_5.

Information retrieval (search)

Approach: We represent each document D_j by a vector $\boldsymbol{d}_j \in \mathbb{R}^5$ (because there are 5 keywords we are interested in).

D_1 D_2 D_3 D_4 D_5 D_6

$$\boldsymbol{d}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 1 \end{pmatrix} \quad \boldsymbol{d}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \boldsymbol{d}_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad \boldsymbol{d}_4 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \quad \boldsymbol{d}_5 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \quad \boldsymbol{d}_6 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

Information retrieval (search)

What if we have 207842147823 documents and 12462247621 keywords? NOT EASY (by inspection)!

D_1 D_2 D_3 D_4 D_5 D_6

$$\boldsymbol{d}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 1 \end{pmatrix} \quad \boldsymbol{d}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \boldsymbol{d}_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad \boldsymbol{d}_4 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \quad \boldsymbol{d}_5 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \quad \boldsymbol{d}_6 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

Information retrieval (search)

Let $\boldsymbol{Q}$ be a 5×6 matrix where $\boldsymbol{d}_j, j=1,2,\ldots,6$ is the jth column of the matrix $\boldsymbol{Q}$.

$$\begin{array}{cc} & \begin{array}{cccccc} \boldsymbol{d}_1 & \boldsymbol{d}_2 & \boldsymbol{d}_3 & \boldsymbol{d}_4 & \boldsymbol{d}_5 & \boldsymbol{d}_6 \end{array} \\ \boldsymbol{Q} = & \begin{pmatrix} 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 \end{pmatrix} \end{array}$$

Since we are interested in keywords k_1, k_4 and k_5, let the search vector

$$\boldsymbol{x} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$

Information retrieval (search)

Note that $\boldsymbol{d}_j \cdot \boldsymbol{x}$ is the number of 'matches' there are in document D_j.

How do we put these information together systematically?

$$\begin{array}{cc} & \begin{array}{cccccc} \boldsymbol{d}_1 & \boldsymbol{d}_2 & \boldsymbol{d}_3 & \boldsymbol{d}_4 & \boldsymbol{d}_5 & \boldsymbol{d}_6 \end{array} \\ \boldsymbol{Q} = & \begin{pmatrix} 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 \end{pmatrix} \end{array} \qquad \boldsymbol{x} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$

$$\boldsymbol{d}_1 \cdot \boldsymbol{x} = 3$$

$$\boldsymbol{d}_6 \cdot \boldsymbol{x} = 2$$

Information retrieval (search)

Let $\boldsymbol{y} = \boldsymbol{Q}^T \boldsymbol{x}$.

$$
\begin{array}{cc}
\boldsymbol{d}_1 & \\
& \\
\boldsymbol{Q}^T \boldsymbol{x} = &
\end{array}
\begin{pmatrix}
1 & 1 & 0 & 1 & 1 \\
1 & 1 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & 0 \\
0 & 1 & 0 & 0 & 1 \\
1 & 0 & 0 & 1 & 1 \\
0 & 0 & 1 & 1 & 1
\end{pmatrix}
\begin{pmatrix}
1 \\ 0 \\ 0 \\ 1 \\ 1
\end{pmatrix}
=
\begin{pmatrix}
3 \\ 1 \\ 0 \\ 1 \\ 3 \\ 2
\end{pmatrix}
= \boldsymbol{y}
$$

Conclusion:

D_1 and D_5 matches our search.

Question:

How can we distinguish between D_1 and D_5?

First, we need more information.

Information retrieval (search)

Suppose the entries in each $\boldsymbol{d}_j$ tells us how many times each keyword actually appears in document D_j :

D_1 D_2 D_3 D_4 D_5 D_6

$$\boldsymbol{d}_1 = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 2 \\ 1 \end{pmatrix} \quad \boldsymbol{d}_2 = \begin{pmatrix} 1 \\ 3 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \boldsymbol{d}_3 = \begin{pmatrix} 0 \\ 1 \\ 5 \\ 0 \\ 0 \end{pmatrix} \quad \boldsymbol{d}_4 = \begin{pmatrix} 0 \\ 2 \\ 0 \\ 0 \\ 1 \end{pmatrix} \quad \boldsymbol{d}_5 = \begin{pmatrix} 3 \\ 0 \\ 0 \\ 5 \\ 1 \end{pmatrix} \quad \boldsymbol{d}_6 = \begin{pmatrix} 0 \\ 0 \\ 6 \\ 1 \\ 4 \end{pmatrix}$$

Information retrieval (search)

Note that if the angle between $\boldsymbol{d}_j$ and the search vector $\boldsymbol{x}$ is θ_j, then

$$\cos\theta_j = \frac{\boldsymbol{d}_j \cdot \boldsymbol{x}}{\|\boldsymbol{d}_j\|\|\boldsymbol{x}\|} = \frac{\boldsymbol{d}_j}{\|\boldsymbol{d}_j\|} \cdot \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}$$

Idea: Let's divide each $\boldsymbol{d}_j$ and $\boldsymbol{x}$ by their respective lengths.

$$\boldsymbol{Q}' = \begin{pmatrix} \frac{\boldsymbol{d}_1}{\|\boldsymbol{d}_1\|} & \frac{\boldsymbol{d}_2}{\|\boldsymbol{d}_2\|} & \frac{\boldsymbol{d}_3}{\|\boldsymbol{d}_3\|} & \frac{\boldsymbol{d}_4}{\|\boldsymbol{d}_4\|} & \frac{\boldsymbol{d}_5}{\|\boldsymbol{d}_5\|} & \frac{\boldsymbol{d}_6}{\|\boldsymbol{d}_6\|} \end{pmatrix} \qquad \boldsymbol{x}' = \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}$$

Information retrieval (search)

Note that if the angle between $\boldsymbol{d}_j$ and the search vector $\boldsymbol{x}$ is θ_j, then

$$\cos\theta_j = \frac{\boldsymbol{d}_j \cdot \boldsymbol{x}}{\|\boldsymbol{d}_j\|\|\boldsymbol{x}\|} = \frac{\boldsymbol{d}_j}{\|\boldsymbol{d}_j\|} \cdot \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}$$

Let $\boldsymbol{y}' = \boldsymbol{Q}'^T \boldsymbol{x}'$

cosine of angle between $\boldsymbol{d}_1$ and $\boldsymbol{x}$

$$\boldsymbol{Q}'^T \boldsymbol{x}' = \begin{pmatrix} \boxed{\dfrac{\boldsymbol{d}_1}{\|\boldsymbol{d}_1\|}} \\ \vdots \\ \vdots \\ \vdots \end{pmatrix} \left(\boxed{\dfrac{\boldsymbol{x}}{\|\boldsymbol{x}\|}} \right) = \begin{pmatrix} \dfrac{\boldsymbol{d}_1}{\|\boldsymbol{d}_1\|} \cdot \dfrac{\boldsymbol{x}}{\|\boldsymbol{x}\|} \\ \vdots \\ \vdots \\ \vdots \end{pmatrix} = \begin{pmatrix} \cos\theta_1 \\ \vdots \\ \vdots \\ \vdots \end{pmatrix}$$

Information retrieval (search)

$\cos\theta = 1 \Rightarrow$ angle between the two vectors is 0

$\Rightarrow$ we look for j such that $\cos\theta_j$ is as close to 1 as possible.

$$\cos\theta = \frac{\boldsymbol{d}_j \cdot \boldsymbol{x}}{\|\boldsymbol{d}_j\|\|\boldsymbol{x}\|} = \frac{\boldsymbol{d}_j}{\|\boldsymbol{d}_j\|} \cdot \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}$$

cosine of angle between $\boldsymbol{d}_1$ and $\boldsymbol{x}$

$$\boldsymbol{Q}'^T \boldsymbol{x}' = \begin{pmatrix} \boxed{\dfrac{\boldsymbol{d}_1}{\|\boldsymbol{d}_1\|}} \\ \vdots \\ \vdots \\ \vdots \end{pmatrix} \begin{pmatrix} \boxed{\dfrac{\boldsymbol{x}}{\|\boldsymbol{x}\|}} \end{pmatrix} = \begin{pmatrix} \dfrac{\boldsymbol{d}_1}{\|\boldsymbol{d}_1\|} \cdot \dfrac{\boldsymbol{x}}{\|\boldsymbol{x}\|} \\ \vdots \\ \vdots \\ \vdots \end{pmatrix} = \begin{pmatrix} \cos\theta_1 \\ \vdots \\ \vdots \\ \vdots \end{pmatrix}$$

Information retrieval (search)

For our example:

$$\boldsymbol{y}' = \boldsymbol{Q}'^T \boldsymbol{x}' = \begin{pmatrix} \boxed{0.9130} \\ 0.1826 \\ 0 \\ 0.2582 \\ \boxed{0.8784} \\ 0.6868 \end{pmatrix}$$

So the conclusion is $\boldsymbol{d}_1$ is 'closer' to $\boldsymbol{x}$ compared to $\boldsymbol{d}_5$.

$$\boldsymbol{x} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \qquad \boldsymbol{d}_1 = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 2 \\ 1 \end{pmatrix} \qquad \boldsymbol{d}_5 = \begin{pmatrix} 3 \\ 0 \\ 0 \\ 5 \\ 1 \end{pmatrix}$$

Do you know why?

Definition (orthogonal, orthonormal)

1) Two vectors $\boldsymbol{u}, \boldsymbol{v}$ are said to be orthogonal if $\boldsymbol{u} \cdot \boldsymbol{v} = 0$.

2) A set S of vectors in $\mathbb{R}^n$ is said to be orthogonal if every pair of distinct vectors in S are orthogonal.

$$S = \{\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}, \boldsymbol{x}\} \qquad \boldsymbol{u} \cdot \boldsymbol{v} = 0,\ \boldsymbol{u} \cdot \boldsymbol{w} = 0,\ \boldsymbol{u} \cdot \boldsymbol{x} = 0$$

$$\boldsymbol{v} \cdot \boldsymbol{w} = 0,\ \boldsymbol{v} \cdot \boldsymbol{x} = 0,\ \boldsymbol{w} \cdot \boldsymbol{x} = 0$$

3) A set S of vectors in $\mathbb{R}^n$ is said to be orthonormal if S is orthogonal and every vector in S is a unit vector.

The concept of orthogonality

1) Two vectors $\boldsymbol{u},\boldsymbol{v}$ are said to be orthogonal if $\boldsymbol{u}\cdot\boldsymbol{v}=0$.

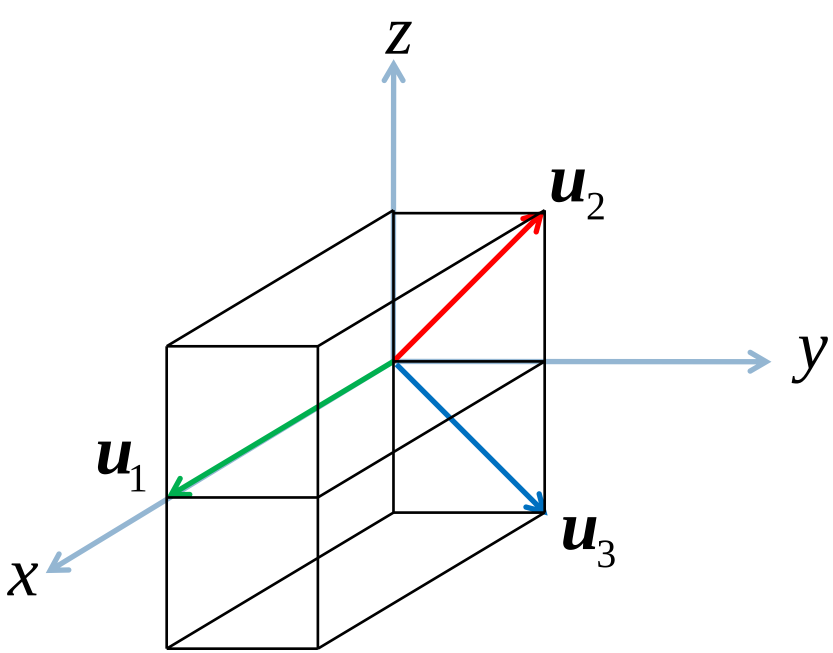
$$\boldsymbol{u}\cdot\boldsymbol{v}=0 \Rightarrow \cos^{-1}\left(\frac{\boldsymbol{u}\cdot\boldsymbol{v}}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|}\right)=\cos^{-1}(0)=\frac{\pi}{2}.$$

Thus the concept of orthogonality is a generalization of perpendicularity that we are familiar in $\mathbb{R}^2$ and $\mathbb{R}^3$.

Example (orthogonal and orthonormal)

Let $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$ where

$\boldsymbol{u}_1 = (2,0,0)$; $\boldsymbol{u}_2 = (0,1,1)$; $\boldsymbol{u}_3 = (0,1,-1)$.



$$\boldsymbol{u}_1 \cdot \boldsymbol{u}_2 = 0$$

$$\boldsymbol{u}_1 \cdot \boldsymbol{u}_3 = 0$$

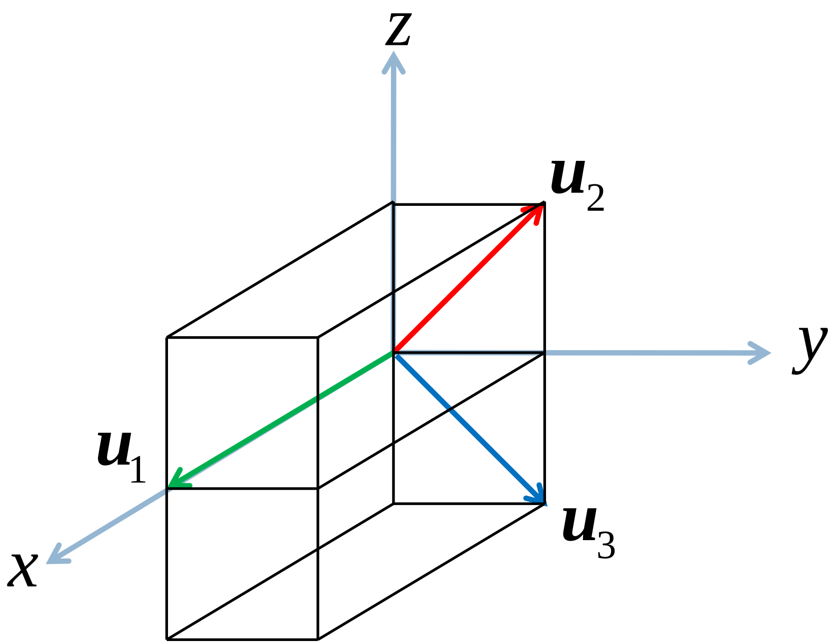
$$\boldsymbol{u}_2 \cdot \boldsymbol{u}_3 = 0$$

S is an orthogonal set

Example (orthogonal and orthonormal)

Converting an orthogonal set to an orthonormal set:

$$\boldsymbol{u}_1 = (2,0,0); \qquad \boldsymbol{u}_2 = (0,1,1); \qquad \boldsymbol{u}_3 = (0,1,-1).$$



$$\boldsymbol{v}_1 = \frac{\boldsymbol{u}_1}{\|\boldsymbol{u}_1\|} = (1,0,0)$$

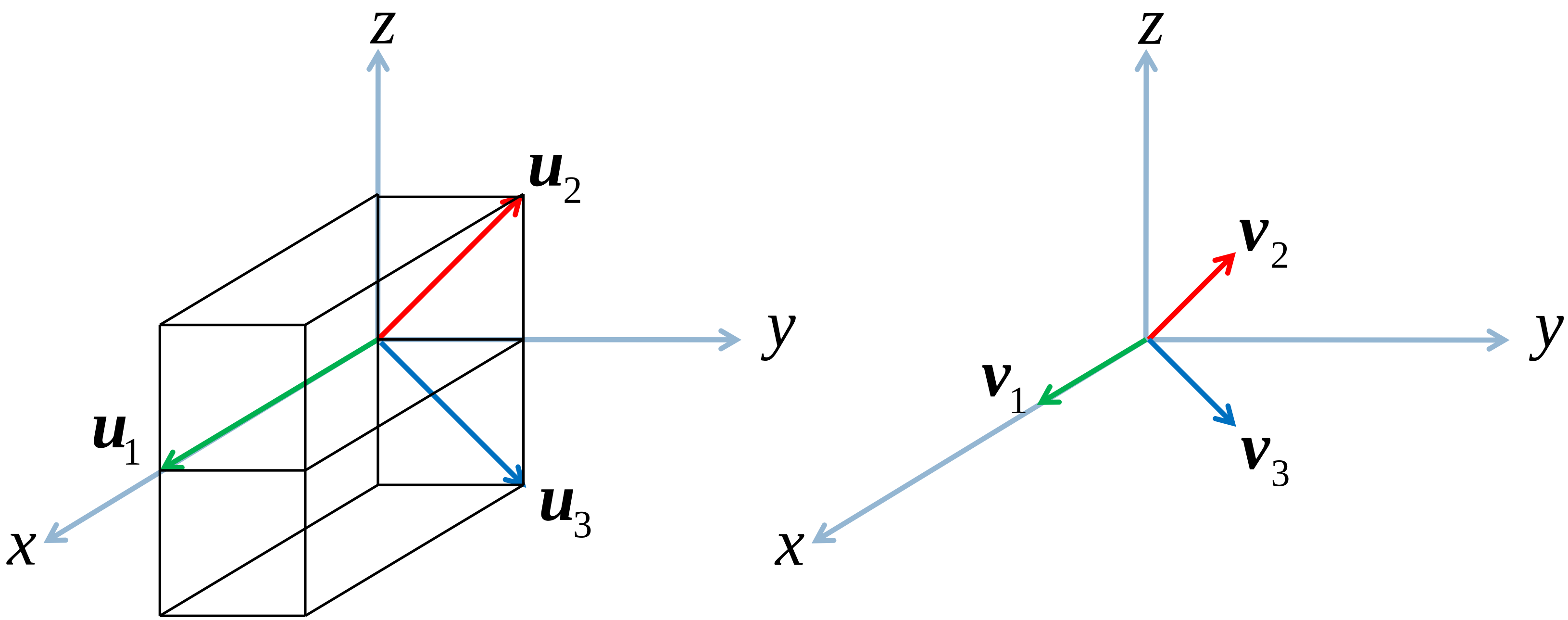
$$\boldsymbol{v}_2 = \frac{\boldsymbol{u}_2}{\|\boldsymbol{u}_2\|} = \frac{1}{\sqrt{2}}(0,1,1)$$

$$\boldsymbol{v}_3 = \frac{\boldsymbol{u}_3}{\|\boldsymbol{u}_3\|} = \frac{1}{\sqrt{2}}(0,1,-1)$$

Example (orthogonal and orthonormal)

Converting an orthogonal set to an orthonormal set:

$$\boldsymbol{v}_1 = \frac{\boldsymbol{u}_1}{\|\boldsymbol{u}_1\|} = (1,0,0) \qquad \boldsymbol{v}_2 = \frac{\boldsymbol{u}_2}{\|\boldsymbol{u}_2\|} = \frac{1}{\sqrt{2}}(0,1,1) \qquad \boldsymbol{v}_3 = \frac{\boldsymbol{u}_3}{\|\boldsymbol{u}_3\|} = \frac{1}{\sqrt{2}}(0,1,-1)$$



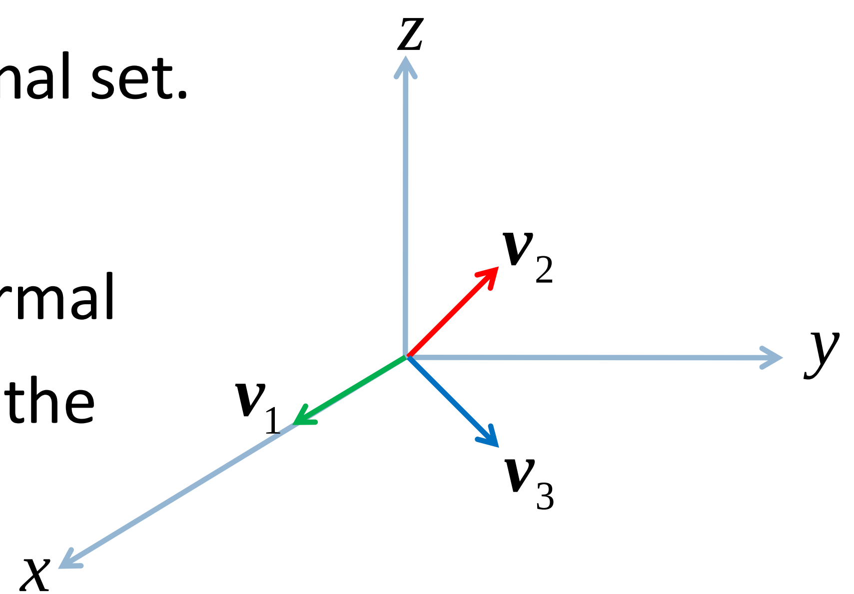
Example (orthogonal and orthonormal)

Converting an orthogonal set to an orthonormal set:

$$\boldsymbol{v}_1 = \frac{\boldsymbol{u}_1}{\|\boldsymbol{u}_1\|} = (1,0,0) \qquad \boldsymbol{v}_2 = \frac{\boldsymbol{u}_2}{\|\boldsymbol{u}_2\|} = \frac{1}{\sqrt{2}}(0,1,1) \qquad \boldsymbol{v}_3 = \frac{\boldsymbol{u}_3}{\|\boldsymbol{u}_3\|} = \frac{1}{\sqrt{2}}(0,1,-1)$$

$S' = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ is an orthonormal set.

This process of converting an orthogonal set to an orthonormal set by dividing each vector in the set by its length is called normalizing.



Example (an orthonormal set)

The standard basis for $\mathbb{R}^n$, $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$ is an orthonormal set.

$$\boldsymbol{e}_1 = (1,0,\ldots,0),\ \boldsymbol{e}_2 = (0,1,\ldots,0),\ldots,\ \boldsymbol{e}_n = (0,0,\ldots,1)$$

1) orthogonal: $\boldsymbol{e}_i \cdot \boldsymbol{e}_j = 0$ if $i \neq j$;

2) unit vectors: $\|\boldsymbol{e}_i\| = \sqrt{0^2 + \ldots + 1^2 + \ldots + 0^2} = 1.$

Theorem

Let S be an orthogonal set of non zero vectors in a vector space. Then S is a linearly independent set.

Definition (orthogonal and orthonormal bases)

1) A basis S for a vector space is an orthogonal basis if S is an orthogonal set.

2) A basis S for a vector space is an orthonormal basis if S is an orthonormal set.

An easy way to check orthogonal basis

If we know that a vector space V has dimension k and

1) S is an orthogonal set of non zero vectors in V (all vectors in S belong to V);

Let S be an orthogonal set of non zero vectors in a vector space. Then S is a linearly independent set.

2) $|S| = k$;

Then we can conclude that S is an orthogonal basis for V.

Example (orthogonal and orthonormal bases)

Let $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$ where

$$\boldsymbol{u}_1 = (2,0,0); \qquad \boldsymbol{u}_2 = (0,1,1); \qquad \boldsymbol{u}_3 = (0,1,-1).$$

S is an orthogonal set of 3 non zero vectors in $\mathbb{R}^3$.

S is an orthogonal basis for $\mathbb{R}^3$.

$\dim(\mathbb{R}^3) = 3$

Let $S' = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ where

$$\boldsymbol{v}_1 = \frac{\boldsymbol{u}_1}{\|\boldsymbol{u}_1\|}; \qquad \boldsymbol{v}_2 = \frac{\boldsymbol{u}_2}{\|\boldsymbol{u}_2\|}; \qquad \boldsymbol{v}_3 = \frac{\boldsymbol{u}_3}{\|\boldsymbol{u}_3\|}.$$

S' is an orthonormal basis for $\mathbb{R}^3$.

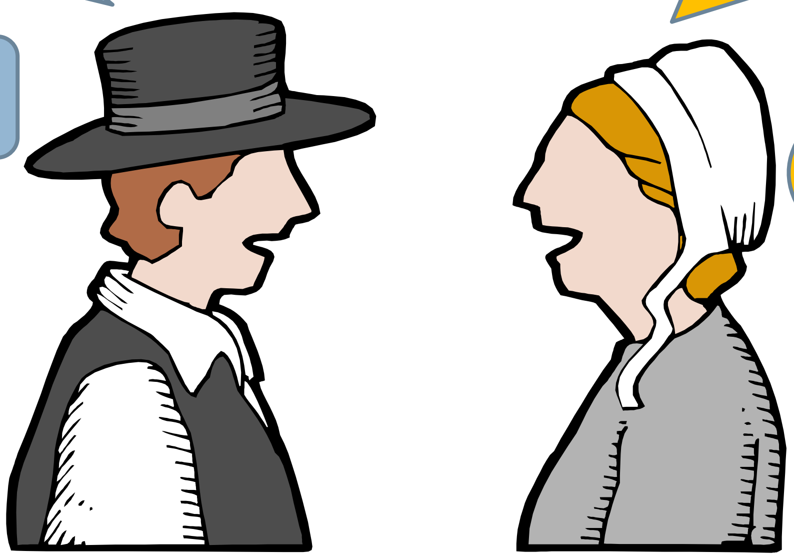
What’s the big deal?

We already know that any vector space has many different bases...

what so special?

orthogonal and orthonormal bases are 'special'...

Let me show you!



How would you answer this question?

Assume that you already know that $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$ where

$$\boldsymbol{u}_1 = (1,2,2,-1),\ \boldsymbol{u}_2 = (1,1,-1,1),\ \boldsymbol{u}_3 = (-1,1,-1,-1)$$

is a basis for a subspace V of $\mathbb{R}^4$.

Give a vector $\boldsymbol{w} = (w_1, w_2, w_3, w_4) \in V$, how do we write $\boldsymbol{w}$ as a linear combination of $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$?

$$a\boldsymbol{u}_1 + b\boldsymbol{u}_2 + c\boldsymbol{u}_3 = \boldsymbol{w}$$

$$\Rightarrow a(1,2,2,-1) + b(1,1,-1,1) + c(-1,1,-1,-1) = (w_1, w_2, w_3, w_4)$$

How would you answer this question?

$$\Rightarrow a(1,2,2,-1)+b(1,1,-1,1)+c(-1,1,-1,-1)=(w_1,w_2,w_3,w_4)$$

$$\Rightarrow \begin{cases} a & + & b & - & c & = & w_1 \\ 2a & + & b & + & c & = & w_2 \\ 2a & - & b & - & c & = & w_3 \\ -a & + & b & - & c & = & w_4 \end{cases}$$

We now solve this linear system for a,b,c.

How would you answer this question?

Assume that you already know that $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$ where

$$\boldsymbol{u}_1 = (1,2,2,-1),\ \boldsymbol{u}_2 = (1,1,-1,1),\ \boldsymbol{u}_3 = (-1,1,-1,-1)$$

is a basis for a subspace V of $\mathbb{R}^4$.

Give a vector $\boldsymbol{w} = (w_1, w_2, w_3, w_4) \in V$, how do we write $\boldsymbol{w}$ as a linear combination of $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$?

What if we know further that S is an orthogonal basis? (is it?)

Theorem

1) If $S=\{\boldsymbol{u}_1,\boldsymbol{u}_2,\ldots,\boldsymbol{u}_k\}$ is an orthogonal basis for a vector space V, then for any vector $\boldsymbol{w}\in V$,

$$\boldsymbol{w}=\left(\frac{\boldsymbol{w}\cdot\boldsymbol{u}_1}{\|\boldsymbol{u}_1\|^2}\right)\boldsymbol{u}_1+\left(\frac{\boldsymbol{w}\cdot\boldsymbol{u}_2}{\|\boldsymbol{u}_2\|^2}\right)\boldsymbol{u}_2+\ldots+\left(\frac{\boldsymbol{w}\cdot\boldsymbol{u}_k}{\|\boldsymbol{u}_k\|^2}\right)\boldsymbol{u}_k$$

WOW!

No need to solve a linear system for these coefficients!

Theorem

1) If $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ is an orthogonal basis for a vector space V, then for any vector $\boldsymbol{w} \in V$,

$$\boldsymbol{w} = \left(\frac{\boldsymbol{w}\cdot\boldsymbol{u}_1}{\|\boldsymbol{u}_1\|^2}\right)\boldsymbol{u}_1 + \left(\frac{\boldsymbol{w}\cdot\boldsymbol{u}_2}{\|\boldsymbol{u}_2\|^2}\right)\boldsymbol{u}_2 + \ldots + \left(\frac{\boldsymbol{w}\cdot\boldsymbol{u}_k}{\|\boldsymbol{u}_k\|^2}\right)\boldsymbol{u}_k$$

So $(\boldsymbol{w})_S = \left(\frac{\boldsymbol{w}\cdot\boldsymbol{u}_1}{\|\boldsymbol{u}_1\|^2}, \frac{\boldsymbol{w}\cdot\boldsymbol{u}_2}{\|\boldsymbol{u}_2\|^2}, \ldots, \frac{\boldsymbol{w}\cdot\boldsymbol{u}_k}{\|\boldsymbol{u}_k\|^2}\right)$ (coordinate vector of $\boldsymbol{w}$ with respect to basis S).

Theorem

2) If $T = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$ is an orthonormal basis for a vector space V, then for any vector $\boldsymbol{w} \in V$,

$$\boldsymbol{w} = (\boldsymbol{w} \cdot \boldsymbol{v}_1)\boldsymbol{v}_1 + (\boldsymbol{w} \cdot \boldsymbol{v}_2)\boldsymbol{v}_2 + \ldots + (\boldsymbol{w} \cdot \boldsymbol{v}_k)\boldsymbol{v}_k$$

So $(\boldsymbol{w})_T = (\boldsymbol{w} \cdot \boldsymbol{v}_1, \boldsymbol{w} \cdot \boldsymbol{v}_2, \ldots \boldsymbol{w} \cdot \boldsymbol{v}_k)$

Example (using orthogonal basis)

Let $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$ where

$$\boldsymbol{u}_1 = (2,0,0); \qquad \boldsymbol{u}_2 = (0,1,1); \qquad \boldsymbol{u}_3 = (0,1,-1).$$

S is an orthogonal basis for $\mathbb{R}^3$.

Express $\boldsymbol{w} = (1,2,3)$ as a linear combination of $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$.

$$\frac{\boldsymbol{w}\cdot\boldsymbol{u}_1}{\|\boldsymbol{u}_1\|^2} = \frac{2}{4} = \boxed{\frac{1}{2}} \qquad \frac{\boldsymbol{w}\cdot\boldsymbol{u}_2}{\|\boldsymbol{u}_2\|^2} = \boxed{\frac{5}{2}} \qquad \frac{\boldsymbol{w}\cdot\boldsymbol{u}_3}{\|\boldsymbol{u}_3\|^2} = \boxed{\frac{-1}{2}}$$

Thus $\boldsymbol{w} = \boxed{\frac{1}{2}}\boldsymbol{u}_1 + \boxed{\frac{5}{2}}\boldsymbol{u}_2 \boxed{-\frac{1}{2}}\boldsymbol{u}_3.$ $\quad (\boldsymbol{w})_S = (\tfrac{1}{2}, \tfrac{5}{2}, -\tfrac{1}{2})$

END OF LECTURE 17

Lecture 18

Orthogonal and Orthonormal Bases (cont'd)

Best Approximations (till Example 5.3.9)