Lecture #24

- Last time $\rightarrow$ Inner products, Lengths, $\underbrace{\text{Angles}}_{\theta=\cos^{-1}\left(\frac{\vec{u}\cdot\vec{v}}{\|\vec{u}\|\|\vec{v}\|}\right)}$, $\underbrace{\text{Orthogonality}}_{\vec{u}\cdot\vec{v}=0}$

  $\rightarrow$ Orthogonal sets

  Projection of a vector onto a vector / a line

Ex1: Verify that $\vec{u_1}=\begin{pmatrix}1\\3\\2\end{pmatrix}$, $\vec{u_2}=\begin{pmatrix}1\\-1\\1\end{pmatrix}$, $\vec{u_3}=\begin{pmatrix}5\\1\\-4\end{pmatrix}$ is an orthogonal set in $\mathbb{R}^3$.

Recall two main claims from the end of last class:

> **Claim 1:** Let $\{\vec{u_1},\ldots,\vec{u_k}\}$ be an orthogonal basis for a subspace W of $\mathbb{R}^n$. For any $\vec{y}\in W$, we have
> $$\vec{y}=\frac{\vec{y}\cdot\vec{u_1}}{\vec{u_1}\cdot\vec{u_1}}\vec{u_1}+\frac{y\cdot\vec{u_2}}{\vec{u_2}\cdot\vec{u_2}}\vec{u_2}+\ldots+\frac{\vec{y}\cdot\vec{u_k}}{\vec{u_k}\cdot\vec{u_k}}\vec{u_k}$$

> **Claim 2:** The projection of $\vec{y}$ onto a nonzero vector $\vec{u}$ (or a line L containing $\vec{u}$) is given by
> $$\hat{y}=\text{proj}_L\vec{y}=\frac{\vec{y}\cdot\vec{u}}{\vec{u}\cdot\vec{u}}\vec{u}$$

Later today we shall have a result generalizing both Claims. Note that the coefficients in both claims look very alike.

Def: $\{\vec{u_1},\ldots,\vec{u_k}\}$ is an orthonormal set of $\mathbb{R}^n$ iff it is an orthogonal set and each vector is a unit vector (i.e. $\vec{u_i}\cdot\vec{u_j}=1$ for $i=j$ and $\vec{u_i}\cdot\vec{u_j}=0$ for $i\neq j$)

An orthonormal set $\{\vec{u_1},\ldots,\vec{u_k}\}$ is called an orthonormal basis of a subspace W if $W=\text{span}\{\vec{u_1},\ldots,\vec{u_k}\}$.

Ex2: Construct an orthonormal basis $\{\vec{v_1},\vec{v_2},\vec{v_3}\}$ of $\mathbb{R}^3$ with $\vec{v_i}$ in the direction of $\vec{u_i}$ from Ex1.

$\blacktriangleright$ $\vec{v_1}=\begin{pmatrix}1\\3\\2\end{pmatrix}\cdot\frac{1}{\sqrt{14}}$, $\vec{v_2}=\begin{pmatrix}1\\-1\\1\end{pmatrix}\cdot\frac{1}{\sqrt{3}}$, $\vec{v_3}=\begin{pmatrix}5\\1\\-4\end{pmatrix}\cdot\frac{1}{\sqrt{42}}$

Lecture #24

A convenient way to encode k vectors $\{\vec{v}_1, \ldots, \vec{v}_k\}$ in $\mathbb{R}^n$ is by an $n\times k$ matrix $A = \begin{pmatrix} \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_k \end{pmatrix}$.

Claim: $\{\vec{v}_1, \ldots, \vec{v}_k\}$ form an orthonormal set of $\mathbb{R}^n$ iff

$$A^T \cdot A = I_k.$$

! Discuss why!

Claim: In the above setup (i.e. columns of A form orthonormal set), we have:

1) $\|A\vec{x}\| = \|\vec{x}\|$ for any $\vec{x} \in \mathbb{R}^k$.

2) $(A\vec{x})\cdot(A\vec{y}) = \vec{x}\cdot\vec{y}$ for any $\vec{x}, \vec{y} \in \mathbb{R}^k$.

3) $\underbrace{(A\vec{x})\cdot(A\vec{y}) = 0}_{A\vec{x}\ \&\ A\vec{y}\text{ - orthogonal}}$ iff $\underbrace{\vec{x}\cdot\vec{y} = 0}_{\vec{x}\ \&\ \vec{y}\text{ - orthogonal}}$

It is instructive to discuss the proofs (for homework problems)

2) $(A\vec{x}) \underset{\text{dot product}}{\cdot} (A\vec{y}) = (A\vec{x})^T \underset{\text{matrix product}}{\cdot} (A\vec{y}) \underset{(AB)^T = B^T A^T}{=} \vec{x}^T \underbrace{A^T A}_{I_k} \vec{y} = \vec{x}^T \underset{\text{matrix product}}{\cdot} \vec{y} = \vec{x} \underset{\text{dot product}}{\cdot} \vec{y}$

3) follows from 2)

1) follows from 2) by taking $\vec{y} = \vec{x}$

Corollary: In the above setup (columns of A - orthonormal set), the linear map $\mathbb{R}^k \to \mathbb{R}^n$, $\vec{x} \mapsto A\vec{x}$, preserves lengths & inner products, hence, also angles

b/c of $\Theta = \cos^{-1}\left(\frac{A\vec{x}\cdot A\vec{y}}{\|A\vec{x}\|\cdot\|A\vec{y}\|}\right) = \cos^{-1}\left(\frac{\vec{x}\cdot\vec{y}}{\|\vec{x}\|\cdot\|\vec{y}\|}\right)$

Lecture #24

Ex3: Let W be the subspace of $\mathbb{R}^3$ spanned by $\vec{v}_1 = \begin{pmatrix}1\\3\\2\end{pmatrix}$, $\vec{v}_2 = \begin{pmatrix}1\\-1\\1\end{pmatrix}$.
Find its orthogonal complement $W^\perp$.

▶ First, according to Ex1, $\vec{v}_3 = \begin{pmatrix}5\\1\\-4\end{pmatrix}$ is in $W^\perp$, hence, any multiple of $\vec{v}_3$ as well.

Now, let us show that any element in $W^\perp$ is a multiple of $\vec{v}_3$.

Recall: $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ - orthogonal set in $\mathbb{R}^3$ $\Rightarrow$ lin. indep. $\Rightarrow$ basis

Any $\vec{y} \in \mathbb{R}^3$ may be written as

$$\vec{y} = \frac{\vec{y}\cdot\vec{v}_1}{\vec{v}_1\cdot\vec{v}_1}\vec{v}_1 + \frac{\vec{y}\cdot\vec{v}_2}{\vec{v}_2\cdot\vec{v}_2}\vec{v}_2 + \frac{\vec{y}\cdot\vec{v}_3}{\vec{v}_3\cdot\vec{v}_3}\vec{v}_3 \quad \text{(by Claim 1)}$$

Hence, if $\vec{y} \in W^\perp \Rightarrow 1^{st}$ & 2^{nd} summands are zero $\Rightarrow y = \frac{\vec{y}\cdot\vec{v}_3}{\vec{v}_3\cdot\vec{v}_3}\vec{v}_3$.

So: $W^\perp$ is a line containing $\vec{v}_3$

Using the same reasoning, we get:

> Claim: If $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ - orthogonal basis of $\mathbb{R}^n$,
> and $W = \text{span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$, then
> $W^\perp = \text{span}\{\vec{v}_{k+1}, \vec{v}_{k+2}, \ldots, \vec{v}_n\}$.

In the rest of today's class, we shall discuss the projection of $\vec{y} \in \mathbb{R}^n$ onto a subspace W of $\mathbb{R}^n$, denoted $\hat{y}$ or $\text{proj}_W \vec{y}$. In other words, we look for a decomposition

$$\boxed{\vec{y} = \hat{y} + \vec{z} \quad \text{with} \quad \begin{array}{l} 1)\ \vec{z} \in W^\perp \\ 2)\ \hat{y} \in W \end{array}}$$

Lecture #24

Claim: Let W be a subspace of $\mathbb{R}^n$, and let $\{\vec{u_1}, \ldots, \vec{u_k}\}$ be an orthogonal basis of W. Then for any $\vec{y} \in \mathbb{R}^n$, have:

$$\hat{\vec{y}} = \text{proj}_W \vec{y} = \frac{\vec{y}\cdot\vec{u_1}}{\vec{u_1}\cdot\vec{u_1}}\vec{u_1} + \ldots + \frac{\vec{y}\cdot\vec{u_k}}{\vec{u_k}\cdot\vec{u_k}}\vec{u_k}$$

Note: 1) For $k=1$, this recovers Claim 2 from page 1.

2) For $k=n$ (i.e. $W=\mathbb{R}^n$), this recovers Claim 1 from page 1, since $\vec{y} = \text{proj}_W \vec{y}$ whenever $\vec{y} \in W$.

Ex 4: Let W be the subspace of $\mathbb{R}^3$ spanned by $\vec{v_1} = \begin{pmatrix}1\\3\\2\end{pmatrix}, \vec{v_2} = \begin{pmatrix}1\\-1\\1\end{pmatrix}$.

Compute $\text{proj}_W \vec{y}$ for $\vec{y} = \begin{pmatrix}1\\2\\3\end{pmatrix} \in \mathbb{R}^3$.

$$\hat{\vec{y}} = \text{proj}_W \vec{y} = \frac{13}{14}\vec{v_1} + \frac{2}{3}\vec{v_2}$$

Claim ("Best approximation theorem"): If W is a subspace of $\mathbb{R}^n$, $\vec{y} \in \mathbb{R}^n$, and $\hat{\vec{y}}$ is the orthogonal projection of $\vec{y}$ onto W, then

$$\|\vec{y} - \hat{\vec{y}}\| < \|\vec{y} - \vec{w}\| \quad \text{for any } \vec{w} \in W \text{ distinct from } \hat{\vec{y}}$$

Ex 5: In the context of Ex 4, find the distance from $\vec{y}$ to W.

$$\|\vec{y} - \hat{\vec{y}}\| = \left\| \begin{pmatrix}1\\2\\3\end{pmatrix} - \frac{13}{14}\begin{pmatrix}1\\3\\2\end{pmatrix} - \frac{2}{3}\begin{pmatrix}1\\-1\\1\end{pmatrix} \right\| = \ldots \leftarrow \text{compute!}$$

Finally, when $\{\vec{u_1}, \ldots, \vec{u_k}\}$ is not just orthogonal, but orthonormal, get:

Claim: If $\{\vec{u_1}, \ldots, \vec{u_k}\}$ – orthonormal basis of $W \subseteq \mathbb{R}^n$, then

$$\hat{\vec{y}} = \text{proj}_W \vec{y} = (\vec{y}\cdot\vec{u_1})\vec{u_1} + \ldots + (\vec{y}\cdot\vec{u_k})\vec{u_k} = AA^T\vec{y},$$

where $A = \begin{pmatrix} \vec{u_1} & \vec{u_2} & \ldots & \vec{u_k} \end{pmatrix}$

Lecture #24

Two instances of products of A, A^T from today:

(1) if columns of A form an orthonormal set, then

$$\boxed{A^T \cdot A = I}$$

(2) if columns of A form an orthonormal set, and W – subspace spanned by them, then:

$$\boxed{\text{proj}_W \vec{y} = AA^T \vec{y}}$$

Rmk: If A is square, i.e. $A = \begin{pmatrix} \vec{u}_1 & \dots & \vec{u}_n \end{pmatrix}$ ($n \times n$)

whose columns form an orthonormal set $\Rightarrow$ columns form a basis of $\mathbb{R}^n$ $\Rightarrow$ A – invertible.

Then, multiplying $A^T A = I$ by A^{-1} on the right, we get

$\boxed{A^T = A^{-1}}$ $\leftarrow$ such matrices are called orthogonal

Multiplying this by A on the left, we also get

$$\boxed{AA^T = I}$$

Ex6: Let A be an $n \times n$ matrix whose columns form an orthonormal basis of $\mathbb{R}^n$. Show that its rows also form an orthonormal basis of $\mathbb{R}^n$.

Given condition $\Rightarrow A^T \cdot A = I_n \Rightarrow A^{-1} = A^T \Rightarrow \underbrace{AA^T}_{=(A^T)^T \cdot A^T} = I_n \Rightarrow$

$\Rightarrow \underbrace{\text{columns of } A^T}_{=\text{rows of } A}$ form an orthonormal basis $\Rightarrow$ result! ∎