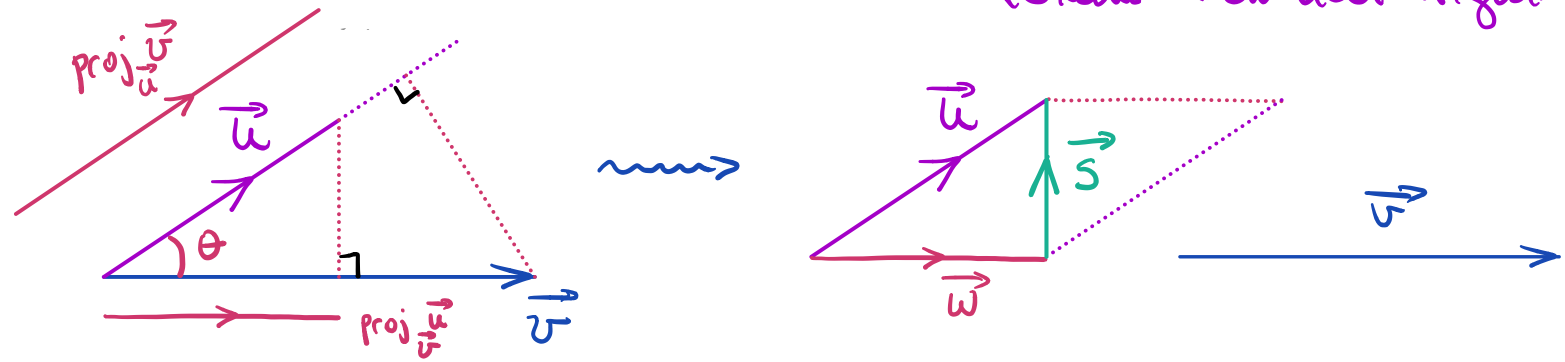
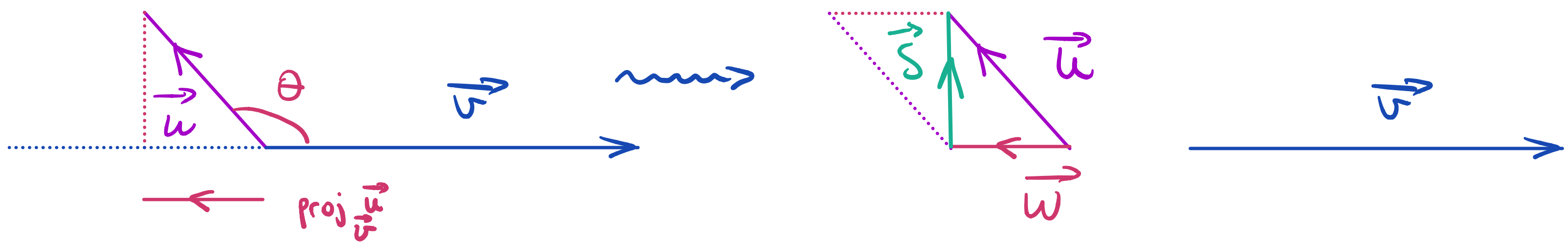
Lecture 18: §3.6 Orthogonal bases for subspaces

TODAY: Discuss inner products (generalizing dot products in $\mathbb{R}^n$) & find bases that are well-behaved with respect to these inner products.
(Gram-Schmidt Algorithm)

$0 \le \theta \le 90^\circ$



$90^\circ < \theta \le 180^\circ$



$$\boxed{\vec u = \underbrace{\text{proj}_{\vec v}\vec u}_{\vec w} + \underbrace{(u - \text{proj}_{\vec v}\vec u)}_{\vec s}}$$

where $\vec w \,/\!/\, \vec v$
$\vec s \perp \vec v$

Key: $\vec w$ & $\vec s$ are the only vectors with these properties. (L11)

Bases: $\{\vec v, \vec u\} \rightsquigarrow \{\vec v, \vec s\}$ $\quad \vec v \perp \vec s$ (like e_1 & e_2)

Inner Products of Subspaces $\mathbb{V}$ of $\mathbb{R}^n$

Def: An inner product for $\mathbb{V}$ is a function $\langle\, , \,\rangle : V \times V \longrightarrow \mathbb{R}$ assigning a number $\langle \vec{u}, \vec{v} \rangle$ to each pair of vectors $\vec{u}, \vec{v}$ in $\mathbb{V}$ & satisfying the following properties:

1. $\langle \vec{u}, \vec{u} \rangle \geq 0$ for all $\vec{u}$; $\langle \vec{u}, \vec{u} \rangle = 0$ if & only if $\vec{u} = \vec{0}$
2. $\langle \vec{u}, \vec{v} \rangle = \langle \vec{v}, \vec{u} \rangle$ (Symmetric)
3. $\langle \alpha \vec{u}, \vec{v} \rangle = \alpha \langle \vec{u}, \vec{v} \rangle = \langle \vec{u}, \alpha \vec{v} \rangle$ for any scalar α.
4. $\langle \vec{u} + \vec{w}, \vec{v} \rangle = \langle \vec{u}, \vec{v} \rangle + \langle \vec{w}, \vec{v} \rangle$
   $\langle \vec{u}, \vec{v} + \vec{w} \rangle = \langle \vec{u}, \vec{v} \rangle + \langle \vec{u}, \vec{w} \rangle$

Main example $\langle \vec{u}, \vec{v} \rangle = \vec{u} \cdot \vec{v} = \underset{1 \times n}{u^T} \underset{n \times 1}{v}$ in $\mathbb{R}^n$ (dot product)

Q: Why? A: Inner products define norms on $\mathbb{R}^n$:

$$\|\vec{v}\|_{\langle , \rangle} = \sqrt{\langle \vec{v}, \vec{v} \rangle} \quad \text{(depends on } \langle , \rangle \text{)}$$

Example 2: Pick Q $n\times n$ symmetric matrix of rank n

(symmetric + invertible Eg $Q = I_n$)

Set $\langle \vec{u}, \vec{v} \rangle := \vec{u}^T Q \vec{v} = \underset{1\times n}{[u_1 \cdots u_n]} \underset{n\times n}{Q} \underset{n\times 1}{\begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}}$ = a number.

Check the 4 properties for an inner product:

3 $\langle \alpha \vec{u}, \vec{v} \rangle = \alpha \langle \vec{u}, \vec{v} \rangle = \langle \vec{u}, \alpha \vec{v} \rangle$ for any scalar α

$(\alpha \vec{u}^T) Q \vec{v} = \alpha (\vec{u}^T Q \vec{v}) = u^T Q \alpha \vec{v}$ ✓

↳ True for matrices

4 $\langle \vec{u} + \vec{w}, \vec{v} \rangle = \langle \vec{u}, \vec{v} \rangle + \langle \vec{w}, \vec{v} \rangle$ & $\langle \vec{u}, \vec{v} + \vec{w} \rangle = \langle \vec{u}, \vec{v} \rangle + \langle \vec{u}, \vec{w} \rangle$

Also true from matrix multiplication properties ✓

2 $\langle \vec{u}, \vec{v} \rangle \overset{?}{=} \langle \vec{v}, \vec{u} \rangle$ (Symmetric)

$\langle \vec{v}, \vec{u} \rangle = \underbrace{\vec{v}^T Q \vec{u}}_{\text{in } \mathbb{R}} = (\vec{v}^T Q \vec{u})^T \underset{\text{transpose rules}}{=} \vec{u}^T Q^T \vec{v} \underset{Q^T = Q}{=} \langle \vec{u}, \vec{v} \rangle$

1 $\langle \vec{u}, \vec{u} \rangle \geq 0$ for all $\vec{u}$; $\langle \vec{u}, \vec{u} \rangle = 0$ if & only if $\vec{u} = \vec{0}$.

⚠ NOT always true!

Ex $Q = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ gives $\left\langle \begin{bmatrix} x \\ y \end{bmatrix}, \begin{bmatrix} x \\ y \end{bmatrix} \right\rangle = x^2 - y^2$

So $\left\langle \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\rangle = 0$ but $\begin{bmatrix} 1 \\ 1 \end{bmatrix} \neq \vec{0}$.

Condition: Q needs to be positive definite
(eg all eigenvalues > 0)

Orthogonal bases for $\mathbb{R}^n$

Fix $\langle\, , \rangle =$ dot product

Def: $\vec{u}$ & $\vec{v}$ in $\mathbb{R}^n$ are <u>orthogonal</u> if $\langle \vec{u}, \vec{v} \rangle = 0$. Write $\vec{u} \perp \vec{v}$

Def: A <u>set</u> of vectors $S = \{\vec{v}_1, \ldots, \vec{v}_p\}$ in $\mathbb{R}^n$ is <u>orthogonal</u> if $\vec{v}_i \perp \vec{v}_j$ for each pair with $i \neq j$

<u>Examples</u>: $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ is orthogonal.

$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}$ ———————. $v_1 \cdot v_2 = v_1 v_3 = v_2 \cdot v_3 = 0$

<u>Q</u>: Why do we care?

<u>Theorem</u>: Orthogonal set <u>NOT</u> containing $\vec{0}$ are <u>ALWAYS l.i</u>.

<u>Why?</u> Write $a_1 \vec{v}_1 + \cdots + a_p \vec{v}_p = \vec{0}$ & show $a_1 = \cdots = a_p = 0$

$$\left.\begin{aligned} 0 = \vec{v}_1 \cdot \vec{0} &= \vec{v}_1 \cdot (a_1 \vec{v}_1 + a_2 \vec{v}_2 + \cdots + a_p \vec{v}_p) \\ &= a_1 \underbrace{\vec{v}_1 \cdot \vec{v}_1}_{\|\vec{v}_1\|^2 \neq 0} + a_2 \underbrace{\vec{v}_1 \cdot \vec{v}_2}_{=0} + \cdots + a_p \underbrace{\vec{v}_1 \cdot \vec{v}_p}_{=0} \end{aligned}\right\}$$

This forces $a_1 = 0$ (but also 0 by the same idea) $\square$

Fix $W \neq \{\vec{0}\}$ a subspace of $\mathbb{R}^n$ with basis $B = \{\vec{w}_1, \ldots, \vec{w}_p\}$

Def We say B is an orthogonal basis if B is an orthogonal set

________ orthonormal basis if B is orthogonal

& $\|\vec{w}_1\| = \|\vec{w}_2\| = \cdots = \|\vec{w}_p\| = 1$ (norm 1 vectors)

Examples ① $W = \mathbb{R}^n$ $B = \left\{ \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \ldots, \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \right\}$ orthonormal basis

② $W = \mathrm{Sp}\left(\underbrace{\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}}_{``\vec{v}_1"}, \underbrace{\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}}_{``\vec{v}_2"} \right)$ = plane $(x - z = 0)$ $\left(\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \times \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -3 \\ 0 \\ 3 \end{bmatrix} \right)$

so $\dim W = 2$

- $B = \{\vec{v}_1, \vec{v}_2\}$ is an orthogonal basis • $\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = 1 - 2 + 1 = 0$

  • basis since it spans W & $\#B = 2 = \dim W$

- B is NOT orthonormal $\|\vec{v}_1\| = \sqrt{1^2 + (-1)^2 + 1^2} = \sqrt{3}$ & $\|\vec{v}_2\| = \sqrt{1^2 + 2^2 + 1^2} = \sqrt{6}$

$\rightsquigarrow$ New orthonormal basis : $\left\{ \frac{1}{\sqrt{3}} \vec{v}_1, \frac{1}{\sqrt{6}} \vec{v}_2 \right\}$

In general : $B = \{\vec{v}_1, \ldots, \vec{v}_p\}$ orthogonal basis for W $\rightsquigarrow$ $B' = \left\{ \frac{\vec{v}_1}{\|\vec{v}_1\|}, \ldots, \frac{\vec{v}_p}{\|\vec{v}_p\|} \right\}$

is orthonormal basis for W.

Main advantage of orthonormal bases: fast to write down coordinates!

Theorem: Fix $B = \{\vec{v}_1, \dots, \vec{v}_p\}$ orthonormal basis for $W \neq \{\vec{0}\}$ in $\mathbb{R}^n$.

Then, for each $\vec{v}$ in W $\quad [\vec{v}]_B = \begin{bmatrix} \vec{v} \cdot \vec{v}_1 \\ \vdots \\ \vec{v} \cdot \vec{v}_p \end{bmatrix}$ in $\mathbb{R}^p$,

meaning $\vec{v} = \underbrace{(\vec{v} \cdot \vec{v}_1)} \vec{v}_1 + \underbrace{(\vec{v} \cdot \vec{v}_2)} \vec{v}_2 + \dots + \underbrace{(\vec{v} \cdot \vec{v}_p)} \vec{v}_p$.

← scalars

Why? Write $\vec{v} = a_1 \vec{v}_1 + a_2 \vec{v}_2 + \dots + a_p \vec{v}_p \qquad a_1, \dots, a_p$ scalars

Then $\vec{v} \cdot \vec{v}_1 = (a_1 \vec{v}_1 + a_2 \vec{v}_2 + \dots + a_p \vec{v}_p) \cdot \vec{v}_1$

$$= a_1 \underbrace{(\vec{v}_1 \cdot \vec{v}_1)}_{= \|\vec{v}_1\|^2 = 1} + a_2 \underbrace{(\vec{v}_2 \cdot \vec{v}_1)}_{= 0 \ (\vec{v}_2 \perp \vec{v}_1)} + \dots + a_p \underbrace{(\vec{v}_p \cdot \vec{v}_1)}_{= 0 \ (\vec{v}_p \perp \vec{v}_1)}$$

So $a_1 = \vec{v} \cdot \vec{v}_1$. Same idea gives $\begin{cases} a_2 = \vec{v} \cdot \vec{v}_2 & (\text{use } \vec{v} \cdot \vec{v}_2) \\ \vdots \\ a_p = \vec{v} \cdot \vec{v}_p & (\text{use } \vec{v} \cdot \vec{v}_p) \end{cases}$

⚠ If B is just an orthogonal basis, then

$$\vec{v} = \boxed{\frac{\vec{v} \cdot \vec{v}_1}{\|\vec{v}_1\|^2}} \vec{v}_1 + \dots + \boxed{\frac{\vec{v} \cdot \vec{v}_p}{\|\vec{v}_p\|^2}} \vec{v}_p .$$

Gram-Schmidt Algorithm

Build orthogonal basis from any basis

- **INPUT:** $B = \{\vec{w}_1, \dots, \vec{w}_p\}$ basis for subspace W of $\mathbb{R}^n$
- **OUTPUT:** $B' = \{\vec{u}_1, \dots, \vec{u}_p\}$ orthogonal basis for W.

Routine: $\vec{u}_1 = \vec{w}_1$

$$\vec{u}_2 = \vec{w}_2 - \operatorname{proj}_{\vec{u}_1} \vec{w}_2 = \vec{w}_2 - \frac{\vec{u}_1 \cdot \vec{w}_2}{\|\vec{u}_1\|^2} \vec{u}_1$$

$$u_3 = \vec{w}_3 - \operatorname{proj}_{\vec{u}_1} \vec{w}_3 - \operatorname{proj}_{\vec{u}_2} \vec{w}_3 = \vec{w}_3 - \frac{\vec{u}_1 \cdot \vec{w}_3}{\|\vec{u}_1\|^2} \vec{u}_1 - \frac{\vec{u}_2 \cdot \vec{w}_3}{\|\vec{u}_2\|^2} \vec{u}_2$$

$\vdots$

In general: once we build $\vec{u}_1, \dots, \vec{u}_{j-1}$, we get $\vec{u}_j$ by

$$\vec{u}_j = \vec{w}_j - \operatorname{proj}_{\vec{u}_1} \vec{w}_j - \operatorname{proj}_{\vec{u}_2} \vec{w}_j - \dots - \operatorname{proj}_{\vec{u}_{j-1}} \vec{w}_j$$

$$= \vec{w}_j - \frac{\vec{u}_1 \cdot \vec{w}_j}{\|\vec{u}_1\|^2} \vec{u}_1 - \frac{\vec{u}_2 \cdot \vec{w}_j}{\|\vec{u}_2\|^2} \vec{u}_2 - \dots - \frac{\vec{u}_{j-1} \cdot \vec{w}_j}{\|\vec{u}_{j-1}\|^2} \vec{u}_{j-1}$$

scalars

$$\vec{u}_j = \vec{w}_j - \frac{\vec{u}_1 \cdot \vec{w}_j}{\|\vec{u}_1\|^2}\vec{u}_1 - \frac{\vec{u}_2 \cdot \vec{w}_j}{\|\vec{u}_2\|^2}\vec{u}_2 - \cdots - \frac{\vec{u}_{j-1} \cdot \vec{w}_j}{\|\vec{u}_{j-1}\|^2}\vec{u}_{j-1}.$$

scalars

Q: Why these scalars?

- $\vec{u}_1 = \vec{w}_1$ so $\vec{w}_1$ in $Sp(\vec{u}_1)$ & $\vec{w}_j$ in $Sp(\vec{u}_1, \vec{u}_2, \dots, \vec{u}_j)$
- $\vec{u}_j$ in $Sp(\vec{w}_j, \vec{u}_1, \dots \vec{u}_{j-1}) = Sp(\vec{w}_j, \vec{w}_1, \vec{w}_2, \dots \vec{w}_{j-1})$
- $\vec{u}_j \cdot \vec{u}_1 = \vec{w}_j \cdot \vec{u}_1 - \frac{\vec{u}_1 \cdot \vec{w}_j}{\cancel{\|\vec{u}_1\|^2}} \cancel{\|\vec{u}_1\|^2} - \frac{\vec{u}_2 \cdot \vec{w}_j}{\|\vec{u}_2\|^2} \underbrace{\vec{u}_2 \cdot \vec{u}_1}_{=0} - \cdots - \frac{\vec{u}_{j-1} \cdot \vec{w}_j}{\|u_{j-1}\|^2} \underbrace{u_{j-1} \cdot \vec{u}_1}_{=0}$

$$= \vec{w}_j \cdot \vec{u}_1 - \vec{u}_1 \cdot \vec{w}_j = 0.$$

Same idea says $\vec{u}_j \cdot \vec{u}_2 = \dots = \vec{u}_j \cdot \vec{u}_{j-1} = 0$

Three consequences: ① $\vec{u}_1, \vec{u}_2, \dots \vec{u}_p$ are all in $\mathbb{W}$

② Each $\vec{w}_1, \dots \vec{w}_p$ is in $Sp(\overbrace{\vec{u}_1, \dots \vec{u}_p}^{=B'})$ so

So $\mathbb{W} = Sp(\vec{w}_1, \dots \vec{w}_p) = Sp(\vec{u}_1, \dots \vec{u}_p)$

B' is a basis ⟸

③ B' is an orthogonal set

$$\boxed{\text{Example}} \qquad V = \mathbb{R}^3$$

$$\text{INPUT}: \left\{ \underbrace{\begin{bmatrix}1\\0\\0\end{bmatrix}}_{=\vec w_1}, \underbrace{\begin{bmatrix}1\\2\\0\end{bmatrix}}_{=\vec w_2}, \underbrace{\begin{bmatrix}1\\-3\\4\end{bmatrix}}_{=\vec w_3} \right\} \qquad \text{OUTPUT} = \{\vec u_1, \vec u_2, \vec u_3\} = \left\{ \begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\2\\0\end{bmatrix}, \begin{bmatrix}0\\0\\4\end{bmatrix} \right\}$$

- $\vec u_1 = \vec w_1 = \begin{bmatrix}1\\0\\0\end{bmatrix}$ $\rightsquigarrow$ $\|\vec u_1\| = \vec u_1 \cdot \vec u_1 = 1^2+0^2+0^2$

  $\vec u_1 \cdot \vec w_2 = \begin{bmatrix}1\\0\\0\end{bmatrix} \cdot \begin{bmatrix}1\\2\\0\end{bmatrix} = 1$

- $\vec u_2 = \vec w_2 - \operatorname{proj}_{\vec u_1} \vec w_2 = \vec w_2 - \dfrac{\vec u_1 \cdot \vec w_2}{\|\vec u_1\|^2} \vec u_1 = \begin{bmatrix}1\\2\\0\end{bmatrix} - \dfrac{1}{1} \begin{bmatrix}1\\0\\0\end{bmatrix} = \begin{bmatrix}0\\2\\0\end{bmatrix}$

- $\vec u_3 = \vec w_3 - \operatorname{proj}_{\vec u_1} \vec w_3 - \operatorname{proj}_{\vec u_2} \vec w_3 = \vec w_3 - \dfrac{\vec u_1 \cdot \vec w_3}{\|\vec u_1\|^2} \vec u_1 - \dfrac{\vec u_2 \cdot \vec w_3}{\|\vec u_2\|^2} \vec u_2$

$$= \begin{bmatrix}1\\-3\\4\end{bmatrix} - \frac{1}{1}\begin{bmatrix}1\\0\\0\end{bmatrix} - \frac{(-6)}{4}\begin{bmatrix}0\\2\\0\end{bmatrix} = \begin{bmatrix}1\\-3\\4\end{bmatrix} - \begin{bmatrix}1\\0\\0\end{bmatrix} + \begin{bmatrix}0\\3\\0\end{bmatrix} = \begin{bmatrix}0\\0\\4\end{bmatrix}$$

$$\|\vec u_2\|^2 = \vec u_2 \cdot \vec u_2 = 0^2 + 2^2 + 0^2 = 4$$

$$\vec u_1 \cdot \vec w_3 = \begin{bmatrix}1\\0\\0\end{bmatrix} \cdot \begin{bmatrix}1\\-3\\4\end{bmatrix} = 1 \quad ; \quad \vec u_2 \cdot \vec w_3 = \begin{bmatrix}0\\2\\0\end{bmatrix} \cdot \begin{bmatrix}1\\-3\\4\end{bmatrix} = -6$$

Q: Orthonormal basis? **A**: $\left\{ \begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix} \right\}$