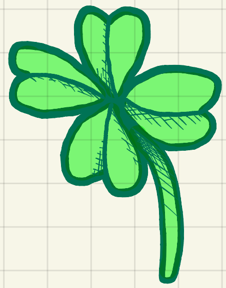
W-3/16

EECS 16B Section 8B



Main Topic: Gram-Schmidt

Administrivia:

- HW 8 deadline extended to Monday (3/28) after Spring Break
- HW 9 released this Saturday (3/19), due Friday (4/1) after Break
- Midterm Redo + Clobber details TBA

Agenda:

- Review of span and orthonormality
- Gram Schmidt Algorithm
    - High-level overview
    - Q1
    - Algorithm
    - Animated Demo
- Orthonormal Matrices
    - Q2

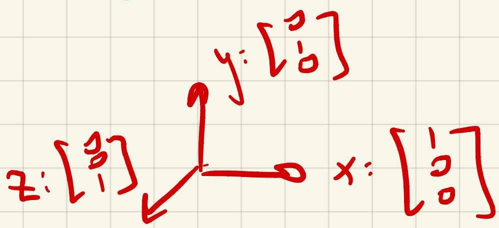
Span and Orthogonality

Consider vectors $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n\}$, $\vec{x}_i \in \mathbb{R}^n$

$\text{span}(\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n\})$ = all vectors $\in \mathbb{R}^n$ that are "reachable", aka representable by a linear combination of $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n$

$\alpha\vec{x}_1 + \beta\vec{x}_2 + \ldots + \gamma\vec{x}_n$

"Orthonormal" = Orthogonal + Normal



Orthogonal: For all $\vec{v}_i, \vec{v}_j$, $i \neq j$,

$\vec{v}_i^T \vec{v}_j = 0$

Normal: For all $\vec{v}_i$,

$\|\vec{v}_i\| = 1$

$(\vec{v}_i^T \vec{v}_i = 1)$

Gram-Schmidt Algorithm

$\text{span}\{\vec{s}_1, \vec{s}_2, \ldots, \vec{s}_n\} = S$

- Linearly Independent*
- Spans vector space S

$\xrightarrow{\text{Gram-Schmidt}}$

$\text{span}\{\vec{q}_1, \vec{q}_2, \ldots, \vec{q}_n\} = S$

- Guaranteed Orthonormal
- Still spans S

Principle: Iteratively generate vectors $\vec{q}_i$

- $\vec{q}_1$: $\|\vec{q}_1\| = 1$ and $\text{span}\{\vec{s}_1\} = \text{span}\{\vec{q}_1\}$
- $\vec{q}_2$: $\|\vec{q}_2\| = 1$ and $\langle \vec{q}_1, \vec{q}_2 \rangle = 0$ and $\text{span}\{\vec{s}_1, \vec{s}_2\} = \text{span}\{\vec{q}_1, \vec{q}_2\}$

$\vdots$

Bottom Line: $\{\vec{q}_i\}$ should be: ① Linearly Independent
② Normal
③ Orthogonal

1. Gram-Schmidt Algorithm

Let's apply Gram-Schmidt orthonormalization to a list of three linearly independent vectors $[\vec{s}_1, \vec{s}_2, \vec{s}_3]$.

(a) Let's say we had two collections of vectors $\{\vec{v}_1, \ldots, \vec{v}_n\}$ and $\{\vec{w}_1, \ldots, \vec{w}_n\}$. **How can we prove that** $\text{Span}(\{\vec{v}_1, \ldots, \vec{v}_n\}) = \text{Span}(\{\vec{w}_1, \ldots, \vec{w}_n\})$**?**

$\vec{v}_i, \vec{w}_i \in \mathbb{R}^n$

Argue that the spans V_1 and W_1 are subsets of one another.

$\text{span}(\{\vec{v}_i\}) = V$

$\text{span}(\{\vec{w}_i\}) = W$

Spans are equal if:

$\forall \vec{v}_i, \ \vec{v}_i \in W$

and $\forall \vec{w}_i, \ \vec{w}_i \in V$

(b) **Find unit vector** $\vec{q}_1$ **such that** $\text{Span}(\{\vec{q}_1\}) = \text{Span}(\{\vec{s}_1\})$**, where** $\vec{s}_1$ **is nonzero.**

$$\boxed{\vec{q}_1 = \frac{\vec{s}_1}{\|\vec{s}_1\|}} \Rightarrow \|\vec{q}_i\| = 1.$$

$\beta \vec{q}_1 = \vec{s},$
$\beta = \|\vec{s}_1\|$

$\text{span}(\{\vec{s}_1\}) = \alpha \vec{s}_1$

$\alpha = \frac{1}{\|\vec{s}_1\|}$, then
$\vec{q}_1 = \alpha \vec{s}_1$

(c) Now, find $\vec{q}_2$.

(c) Let's say that we wanted to write

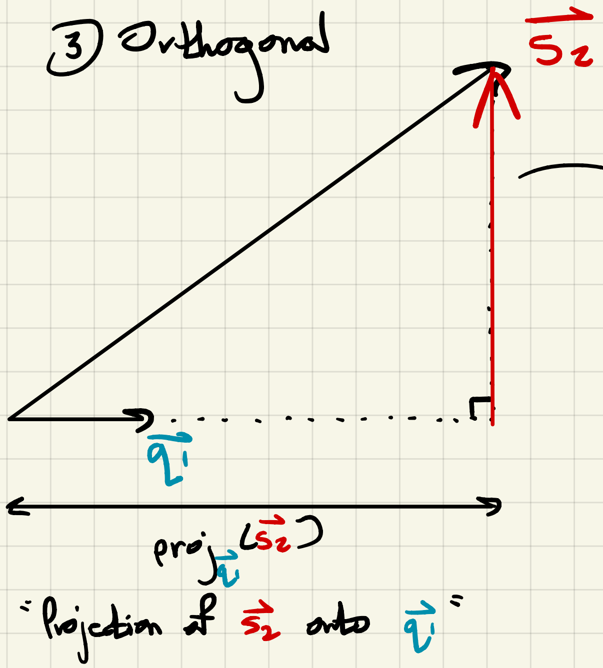
$$\vec{s}_2 = c_1\vec{q}_1 + \vec{z}_2 \tag{1}$$

where $c_1\vec{q}_1$ entirely represents the component of $\vec{s}_2$ in the direction of $\vec{q}_1$, and $\vec{z}_2$ represents the component of $\vec{s}_2$ that is distinctly *not* in the direction of $\vec{q}_1$ (i.e. $\vec{z}_2$ and $\vec{q}_1$ are orthogonal).
Given $\vec{q}_1$ from the previous step, **find** c_1 **as in eq. (1), and use** $\vec{z}_2$ **to find unit vector** $\vec{q}_2$ **such that** $\text{Span}(\{\vec{q}_1, \vec{q}_2\}) = \text{Span}(\{\vec{s}_1, \vec{s}_2\})$ **and** $\vec{q}_2$ **is orthogonal to** $\vec{q}_1$**. Show that** $\text{Span}(\{\vec{q}_1, \vec{q}_2\}) = \text{Span}(\{\vec{s}_1, \vec{s}_2\})$.

Revisit later

Recall:

$\{\vec{q}_i\}$ should be:
1. Linearly Independent ✓
2. Normal ✓
3. Orthogonal



$\vec{s}_2 - \text{proj}_{\vec{q}_1}(\vec{s}_2) = \boxed{\vec{z}_2}$

Projection Formula: $\text{proj}_{\vec{q}_1}(\vec{s}_2) = \dfrac{\overbrace{(\vec{s}_2^{\,T}\vec{q}_1)}^{\text{scalar}}\vec{q}_1}{\|\vec{q}_1\|_2^2}$.

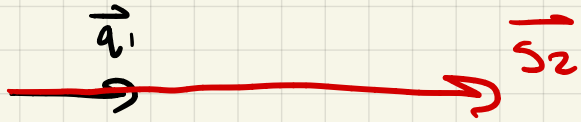
Since $\vec{q}_1$ is normal, $= \underbrace{(\vec{s}_2^{\,T}\vec{q}_1)}_{\text{scalar}}\vec{q}_1$.

$$\vec{z}_2 = \vec{s}_2 - (\vec{s}_2^{\,T}\vec{q}_1)\vec{q}_1$$

$$\vec{q}_2 = \frac{\vec{z}_2}{\|\vec{z}_2\|}$$

(d) **What would happen if $\{\vec{s}_1, \vec{s}_2, \vec{s}_3\}$ were *not* linearly independent, but rather $\vec{s}_1$ were a multiple of $\vec{s}_2$?**

$\vec{s}_2 = \alpha \vec{s}_1$.



Because $\vec{s}_1, \vec{s}_2$ are in the same direction,

$$\vec{z}_2 = \vec{s}_2 - (\vec{s}_2^{\,T}\vec{q}_1)\vec{q}_1 = 0 \qquad \vec{q}_2 = 0.$$

(e) Now given $\vec{q}_1$ and $\vec{q}_2$ in parts 1.b and 1.c, **find $\vec{q}_3$ such that** $\text{Span}(\{\vec{q}_1, \vec{q}_2, \vec{q}_3\}) = \text{Span}(\{\vec{s}_1, \vec{s}_2, \vec{s}_3\})$, **and $\vec{q}_3$ is orthogonal to both $\vec{q}_1$ and $\vec{q}_2$, and finally $\|\vec{q}_3\| = 1$.** You do not have to show that the two spans are equal.

$$\vec{q}_1 = \frac{\vec{s}_1}{\|\vec{s}_1\|}$$

$$\vec{z}_2 = \vec{s}_2 - \text{proj}_{\vec{q}_1}(\vec{s}_2)$$

$$\vec{q}_2 = \frac{\vec{z}_2}{\|\vec{z}_2\|}$$

$$\vec{z}_3 = \vec{s}_3 - \overbrace{(\vec{s}_3^{\,T}\vec{q}_1)\vec{q}_1}^{\text{proj}_{\vec{q}_1}(\vec{s}_3)} - \overset{\text{proj}_{\vec{q}_2}(\vec{s}_3)}{(\vec{s}_3^{\,T}\vec{q}_2)\vec{q}_2}$$

$$\boxed{\vec{q}_3 = \frac{\vec{z}_3}{\|\vec{z}_3\|}}$$

Gram-Schmidt Algorithm Pseudocode:

Base Case

$$\vec{q}_1 = \frac{\vec{s}_1}{\|\vec{s}_1\|}$$

Iterative Case

for $i = 2 \rightarrow n$:

$$\vec{z}_i = \vec{s}_i - \underbrace{\sum_{j=1}^{i-1} \left(\vec{s}_i^{\,T} \vec{q}_j\right) \vec{q}_j}_{\text{Subtract out } i-1 \text{ projections}}$$

$$\underbrace{\vec{q}_i = \frac{\vec{z}_i}{\|\vec{z}_i\|}}_{\text{Normalize}}$$

$$\vec{z}_3 = \vec{s}_3 - \sum_{j=1}^{2} \left(\vec{s}_3^{\,T} \vec{q}_j\right) \vec{q}_j$$

$$\vec{s}_3 - \left(\vec{s}_3^{\,T} \vec{q}_1\right) \vec{q}_1 - \left(\vec{s}_3^{\,T} \vec{q}_2\right) \vec{q}_2$$

2. Orthonormal Matrices and Projections

A matrix A has orthonormal columns, $\vec{a}_i$, if they are:

- Orthogonal (ie. $\langle \vec{a}_i, \vec{a}_j \rangle = \vec{a}_j^\top \vec{a}_i = 0$ when $i \neq j$)
- Normalized (ie. vectors with length equal to 1, $\|\vec{a}_i\| = 1$). This implies that $\|\vec{a}_i\|_2 = \langle \vec{a}_i, \vec{a}_i \rangle = \vec{a}_i^\top \vec{a}_i = 1$.

(a) When $A \in \mathbb{R}^{n \times m}$ and $n \geq m$ (i.e. for tall matrices), **show that if the matrix is orthonormal, then** $A^\top A = I_{m\times m}$.

"Tall"

$$A = \begin{array}{cc} n & \begin{bmatrix} | & & | \\ \vec{a}_1 & \cdots & \vec{a}_m \\ | & & | \end{bmatrix} \\ & m \end{array} \qquad A^T = \begin{array}{cc} m & \begin{bmatrix} — & \vec{a}_1^T & — \\ & \vdots & \\ — & \vec{a}_m^T & — \end{bmatrix} \\ & n \end{array}$$

$$\vec{a}_1^T \vec{a}_2 = 0$$

$$A^T A = \begin{bmatrix} — & \vec{a}_1^T & — \\ & \vdots & \\ — & \vec{a}_m^T & — \end{bmatrix} \begin{bmatrix} | & & | \\ \vec{a}_1 & \cdots & \vec{a}_m \\ | & & | \end{bmatrix}$$

$$= \begin{bmatrix} \vec{a}_1^T\vec{a}_1 & \vec{a}_1^T\vec{a}_2 & \cdots & \vec{a}_1^T\vec{a}_m \\ \vec{a}_2^T\vec{a}_1 & \ddots & & \\ \vdots & & \ddots & \\ \vec{a}_m^T\vec{a}_1 & & & \vec{a}_m^T\vec{a}_m \end{bmatrix} = \begin{bmatrix} 1 & & 0 \\ & 1 & \\ & & \ddots & \\ 0 & & & 1 \end{bmatrix}$$

$$= I_{m\times m}$$

(b) Again, suppose $A \in \mathbb{R}^{n \times m}$ where $n \geq m$ is an orthonormal matrix. **Show that the projection of $\vec{y}$ onto the subspace spanned by the columns of A is now $AA^\top \vec{y}$.**

$A \in \mathbb{R}^{n \times m}$, $n \geq m$, A is tall.

$$\text{proj}_{\text{col}(A)}(\vec{y}) = \underbrace{A \underbrace{(A^T A)^{-1} A^T \vec{y}}_{\text{Least Squares}}}_{A\hat{\vec{x}}}$$

From 2a, $A^T A = I_{m \times m}$.

$$A (A^T A)^{-1} A^T \vec{y}$$

$$A_{n \times m} I_{m \times m}^{-1} A^T_{m \times n} \vec{y}_{n \times 1}$$

$$\Rightarrow AA^T \vec{y}$$

Password: Spring

(Happy Spring Break Everyone!)

links.eecs16b.org/maxwell-dis