EECS16B DIS 8B remote

Any questions before we get started?

Learning objectives

- Gram Schmidt algorithm - how to turn a set of vectors into an orthonormal set of vectors

$\langle \vec{v}_i, \vec{v}_j \rangle = 0$ $\quad |\vec{v}_i| = 1$

- Examples of translating proof ask into something we can work with
- How to express a vector in terms of an orthonormal basis
- Properties of orthonormal matrices

A $n \times m$, $n > m$ (tall) $\quad A^T A = I$ $(m \times m)$ $\quad \underset{(n\times n)}{AA^T} \neq \underset{(n\times n)}{I}$

A $n \times n$ (square) $\quad A^T A = I$ $(n \times n)$, $AA^T = I$, $A^{-1} = A^T$

$\underbrace{\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}}_{\text{set of vectors}}$

$\underbrace{\{\vec{q}_1, \vec{q}_2, \ldots, \vec{q}_n\}}_{\text{set we want to build}}$,

where all $|\vec{q}_i| = 1$

$\vec{q}_i^{\,T} \vec{q}_j = 0$

$i \neq j$

GS algorithm

at every loop

1. take a vector ($\vec{v}_i$)
2. Remove parts of $\vec{v}_i$ that are accounted for by in the set of orthonormal vectors we've already built up.
3. Normalize what's leftover
4. add to our set/dictionary of vectors (next vector $\vec{q}_i$)

EECS 16B Designing Information Devices and Systems II
Fall 2021 Discussion Worksheet Discussion 8B

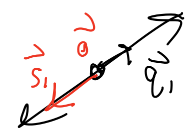
The following notes are useful for this discussion: Note 12.

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

→ $\vec{q}_1$ & $\vec{s}_1$ are in the same direction

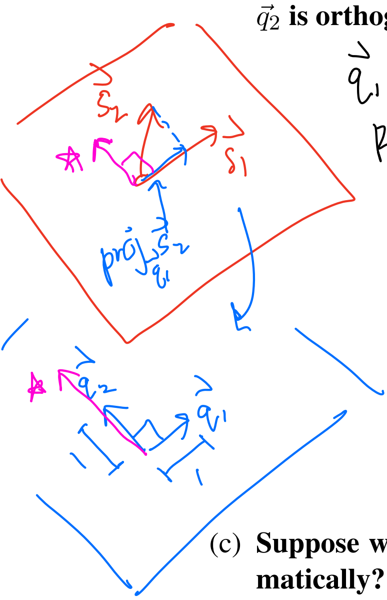


Option 1: Take $\vec{s}_1$ and "normalize it"
→ divide by it's norm

Compute $\boxed{\frac{\vec{s}_1}{\|\vec{s}_1\|} = \vec{q}_1}$ (doesn't work for $\vec{0}$)

Option 2: $\boxed{-\vec{q}_1 = \frac{-\vec{s}_1}{\|\vec{s}_1\|}}$ $\quad \|-\vec{q}_1\| = |-1| \, \|\vec{q}_1\| = 1$

(b) Given $\vec{q}_1$ from the previous step, **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** $\vec{q}_1$.



$\vec{q}_1$ is now part of our set of orthonormalized vectors
Remove from $\vec{s}_2$ parts of it that point along $\vec{q}_1$ (2nd iteration of GS / 2nd loop)

① compute $\text{proj}_{\vec{q}_1} \vec{s}_2 = \vec{q}_1 (\vec{q}_1^T \vec{q}_1)^{-1} \vec{q}_1^T \vec{s}_2$ (vector projection formula) least squares formula

$= \vec{q}_1 \cdot \vec{q}_1^T \vec{s}_2$

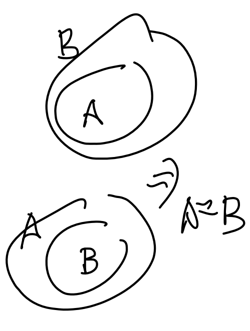
② Remove projections from $\vec{s}_2$

☆ $\vec{s}_2 - (\vec{q}_1^T \vec{s}_2) \vec{q}_1$ = a vector that points in a direction orthogonal to $\vec{q}_1$.

③ Normalize: $\boxed{\vec{q}_2 = \frac{\vec{s}_2 - (\vec{q}_1^T \vec{s}_2)\vec{q}_1}{\|\vec{s}_2 - (\vec{q}_1^T \vec{s}_2)\vec{q}_1\|}}$

(c) **Suppose we want to show that** $\text{Span}(\{\vec{q}_1, \vec{q}_2\}) = \text{Span}(\{\vec{s}_1, \vec{s}_2\})$. **What does this mean mathematically?** *Hint: you cannot use the word span, but must capture the same concept in your translation of the statement we want to show.*

What is span($\{\vec{q}_1, \vec{q}_2\}$)? what object? Collection (of vectors) → it's a set
want to show two sets are equal. equal if all elements in one set are in the other and vice versa



$$\left\{ \begin{array}{l} a\vec{q}_1 + b\vec{q}_2 \in \text{Span}\{\vec{q}_1, \vec{q}_2\} \\ c\vec{s}_1 + d\vec{s}_2 \in \text{Span}\{\vec{s}_1, \vec{s}_2\} \end{array} \right\}$$

for all $\forall a\vec{q}_1 + b\vec{q}_2 \in \text{Span}\{\vec{q}_1, \vec{q}_2\}$
exists $\exists c\vec{s}_1 + d\vec{s}_2 \in \text{Span}\{\vec{s}_1, \vec{s}_2\}$
$a\vec{q}_1 + b\vec{q}_2 = c\vec{s}_1 + d\vec{s}_2$

(says that $\text{Span}\{\vec{q}_1, \vec{q}_2\} \subseteq \text{Span}\{\vec{s}_1, \vec{s}_2\}$) subset of

$\forall c\vec{s}_1 + d\vec{s}_2 \in \text{Span}\{\vec{s}_1, \vec{s}_2\}$
$\exists a\vec{q}_1 + b\vec{q}_2 \in \text{Span}\{\vec{q}_1, \vec{q}_2\}$
$c\vec{s}_1 + d\vec{s}_2 = a\vec{q}_1 + b\vec{q}_2$

If $\vec{s}_2$ & $\vec{q}_1$ point in the same direction

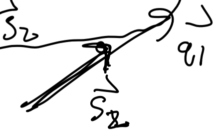
(d) **What would happen if $\{\vec{s}_1, \vec{s}_2, \vec{s}_3\}$ were *not* linearly independent, but rather $\vec{s}_1$ were a multiple of $\vec{s}_2$?**

Ex, look @ $\vec{q}_2$

$$\vec{q}_2 = \frac{\vec{s}_2 - \underbrace{\langle \vec{q}_1, \vec{s}_2 \rangle \vec{q}_1}_{\text{proj}_{\vec{q}_1} \vec{s}_2}}{\| \vec{s}_2 - \langle \vec{q}_1, \vec{s}_2 \rangle \vec{q}_1 \|} = \frac{\vec{s}_2 - \vec{s}_2 \rightarrow \vec{0}}{\| \quad \|}$$

can't compute $\boxed{0}$

$\boxed{\text{Have to throw out } \vec{s}_2}$



(e) Now given $\vec{q}_1$ and $\vec{q}_2$ in parts (a) and (b), **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** $\|\vec{q}_3\| = 1$.

1 Project:

$$\text{proj}_{\text{Span}\{\vec{q}_1, \vec{q}_2\}} \vec{s}_3 = \begin{bmatrix} | & | \\ \vec{q}_1 & \vec{q}_2 \\ | & | \end{bmatrix} \left(\begin{bmatrix} \vec{q}_1 & \vec{q}_2 \end{bmatrix}^T \begin{bmatrix} \vec{q}_1 & \vec{q}_2 \end{bmatrix} \right)^{-1} \begin{bmatrix} \vec{q}_1 & \vec{q}_2 \end{bmatrix}^T \vec{s}_3$$

2 Remove

$$= \underbrace{\langle \vec{q}_1, \vec{s}_3 \rangle \vec{q}_1}_{} + \underbrace{\langle \vec{q}_2, \vec{s}_3 \rangle \vec{q}_2}_{}$$

$$\vec{r} = \vec{s}_3 - \langle \vec{s}_3, \vec{q}_1 \rangle \vec{q}_1 - \langle \vec{q}_2, \vec{s}_3 \rangle \vec{q}_2$$

3 Normalize

$$\vec{q}_3 = \frac{\vec{r}}{\|\vec{r}\|}$$

(f) **[Practice] Confirm that** $\text{Span}(\{\vec{q}_1, \vec{q}_2, \vec{q}_3\}) = \text{Span}(\{\vec{s}_1, \vec{s}_2, \vec{s}_3\})$.

Not solution, but a sketch:

If starting from $\text{Span}\{\vec{q}_1, \vec{q}_2\} = \text{Span}\{\vec{s}_1, \vec{s}_2\}$ we have an easier job: just have to show that $\vec{q}_3$ can be written in terms of $\vec{s}_1, \vec{s}_2, \vec{s}_3$ and $\vec{s}_3$ can be written in terms of $\vec{q}_1, \vec{q}_2, \vec{q}_3$

So that vectors that can be written in terms of $\vec{q}_1, \vec{q}_2, \vec{q}_3$ can be written in terms of $\vec{s}_1, \vec{s}_2, \vec{s}_3$, and vectors that can be written in terms of $\vec{s}_1, \vec{s}_2, \vec{s}_3$ are writable in terms of $\vec{q}_1, \vec{q}_2, \vec{q}_3$

If not assuming anything except that $\{\vec{q}_1, \vec{q}_2, \vec{q}_3\}$ is the output of Gram Schmidt, then have to do this part without just saying it, giving what the coefficients would be.

$Q = \begin{bmatrix} | & | \\ q_1 & q_2 \\ | & | \end{bmatrix}$ ← vectors out of GS

$A = \begin{bmatrix} | & | & & | \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_m \\ | & | & & | \end{bmatrix}$

$\vec{a}_1 = \begin{bmatrix} a_{11} \\ a_{12} \\ \vdots \\ a_{1n} \end{bmatrix}$ $\quad \vec{a}_i \in \mathbb{R}^n$

2. Orthonormal Matrices and Projections

An orthonormal matrix, A, is a matrix whose columns, $\vec{a}_i$, 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 $\|\vec{a}_i\|_2 = \langle \vec{a}_i, \vec{a}_i \rangle = \vec{a}_i^\top \vec{a}_i = 1$.

Gram Schmidt does generate an orthonormal matrix from a set of vectors (provided we throw away $\vec{0}$ vectors or lin. dep. vectors)

(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 $A^\top A = I_{m\times m}$.**

$A^T = \begin{bmatrix} - & \vec{a}_1^T & - \\ - & \vec{a}_2^T & - \\ & \vdots & \\ - & \vec{a}_m^T & - \end{bmatrix}$

$(A^TA)_{ij} = \vec{a}_i^T \vec{a}_j = \begin{cases} \vec{a}_i^T \vec{a}_i = 1^2 & i = j \\ 0 & i \neq j \end{cases}$ ← orthogonal columns of A

A^TA

A^T ($m \times n$), A ($n \times m$)

not an inverse! (when tall)

$A^TA = \begin{bmatrix} - & \vec{a}_1^T & - \\ & \vdots & \\ - & \vec{a}_m^T & - \end{bmatrix} \begin{bmatrix} | & & | \\ \vec{a}_1 & \cdots & \vec{a}_m \\ | & & | \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & \cdots \\ 0 & 1 & 0 & \\ 0 & 0 & 1 & \\ & & & \ddots \end{bmatrix}}_{m \times m}$

(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 $AA^\top \vec{y}$.**

Q: what does this mean?

A: Nonsquare matrices don't have inverses ($AA^T \neq I$) Need both $A^TA, AA^T = I$ to be inverses (only square matrices)

$$\text{proj}_{\text{span}\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m\}} \vec{y} = A(A^TA)^{-1}A^T\vec{y} = A(\underbrace{I}_{m\times m})^{-1}A^T\vec{y} = AIA^T\vec{y} = AA^T\vec{y}$$

$$= A \begin{bmatrix} - & \vec{a}_1^T & - \\ & \vdots & \\ - & \vec{a}_m^T & - \end{bmatrix} \vec{y} = A \begin{bmatrix} \vec{a}_1^T\vec{y} \\ \vec{a}_2^T\vec{y} \\ \vdots \\ \vec{a}_m^T\vec{y} \end{bmatrix} = \sum_{i=1}^{m} (\vec{a}_i^T\vec{y})\vec{a}_i$$

(c) **Show if $A \in \mathbb{R}^{n\times n}$ is an orthonormal matrix then the columns, $\vec{a}_i$, form a basis for $\mathbb{R}^N$.** *Hint: can you use the result of the previous subpart?*

square

① Span ② lin. indep. ← basis (for $\mathbb{R}^n$): "a set of vectors that let you write every vector in $\mathbb{R}^n$ as a unique linear combination of the vectors in the set"

$\{\vec{a}_1, \ldots, \vec{a}_n\}$ ← basis if any $\vec{v} \in \mathbb{R}^n$ can be expressed as: $\sum_{i=1}^{n} \beta_i \vec{a}_i = \vec{v}$ only one possible set of scalars β_i that give us $\vec{v}$.

$AB = I$ if A, B square, then inverses: $A^{-1} = B$, $B^{-1} = A$

What is $\text{proj}_{\text{span}(A)}\vec{y} = AA^T\vec{y}$ when A is square?

→ $A^TA = I$ (true for orthonormal matrix)

$AA^{-1} = I = A^{-1}A$

when A is square and orthonormal $A^{-1} = A^T$

Contributors:

- Regina Eckert.
- Druv Pai.
- Neelesh Ramachandran.

(A, A^T are inverses) $\underbrace{AA^T\vec{y}}_{A \text{ is square}} = \vec{y} = \underbrace{\sum_{i=1}^{n} (\vec{a}_i^T\vec{y})\vec{a}_i}_{\text{from (b)}}$

if $\vec{a}_i$ concrete #'s and a $\vec{y} \in \mathbb{R}^n$ also concrete β_i's chosen $\vec{a}_i^T\vec{y}$ → a # only one value