

# Math 254: Introduction to Linear Algebra

## Notes #5.2 — Gram-Schmidt Process and QR Factorization

Peter Blomgren  
([blomgren@sdsu.edu](mailto:blomgren@sdsu.edu))

Department of Mathematics and Statistics  
Dynamical Systems Group  
Computational Sciences Research Center  
San Diego State University  
San Diego, CA 92182-7720

<http://terminus.sdsu.edu/>

Spring 2022  
(Revised: March 24, 2022)



After this lecture you should know how:

- to perform *The Gram-Schmidt Orthogonalization Process* on a set of vectors, and
- it can be used to compute *The QR-factorization* of a matrix  $A$ :  $A = QR$ 
  - $\Rightarrow$  This builds an orthonormal basis (the columns of  $Q$ ) for the subspace  $V = \text{im}(A)$ , which gives us the means to compute the orthogonal projection  $\text{proj}_V(\vec{x})$  onto  $V$ .
- to orthogonally project onto *any* subspace.



## Outline

- 1 Student Learning Objectives**
  - SLOs: Gram-Schmidt Process and QR Factorization
- 2 Gram-Schmidt Orthogonalization and QR Factorization**
  - The Gram-Schmidt Orthogonalization Process
  - The QR Factorization
  - Observations
- 3 Suggested Problems**
  - Suggested Problems 5.2
  - Lecture–Book Roadmap
- 4 Supplemental Material**
  - Metacognitive Reflection
  - Problem Statements 5.2
  - Why Orthogonal Projections Matter
  - Example:  $V = \{x_1 + x_2 + x_3 + x_4 = 0\} \subset \mathbb{R}^4$
  - Example: 5.2.35 and Beyond — “Live Math” Discussion



## Orthogonal Projection onto a Subspace $V$

From [NOTES#5.1] we have:

### Theorem (Formula for the Orthogonal Projection)

If  $V$  is a subspace of  $\mathbb{R}^n$  with an orthonormal basis  $\vec{u}_1, \dots, \vec{u}_m$ , then

$$\text{proj}_V(\vec{x}) = \vec{x}^{\parallel} = (\vec{u}_1 \cdot \vec{x})\vec{u}_1 + \dots + (\vec{u}_m \cdot \vec{x})\vec{u}_m$$

$$\forall \vec{x} \in \mathbb{R}^n.$$

*How do you project onto a subspace if/when the given basis is not orthonormal?!?*

It turns out that before we compute the projection, we have to find a new — *orthonormal* — basis...



## THIS IS ALL WRONG!!!!

Ponder what happens if we use the formula, but the given basis is **not** orthonormal...

Let's live in  $\mathbb{R}^2$ , let  $V = \mathbb{R}^2$ , with basis  $\mathfrak{B} = (\vec{v}_1, \vec{v}_2)$  defined by

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}; \quad \text{and } \vec{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}; \quad \|\vec{x}\| = \sqrt{13}.$$

Clearly  $\vec{v}_1$  and  $\vec{v}_2$  are linearly independent, and  $\vec{x} = 1\vec{v}_1 + 1\vec{v}_2$ , but the projection formula goes haywire:

$$\text{proj}_V(\vec{x}) = (\vec{v}_1 \cdot \vec{x})\vec{v}_1 + (\vec{v}_2 \cdot \vec{x})\vec{v}_2 = 5\vec{v}_1 + 8\vec{v}_2 = \begin{bmatrix} 13 \\ 21 \end{bmatrix}.$$

... even if we remember to correct for the non-unit length of  $\vec{v}_{1,2}$ :

$$\text{proj}_V(\vec{x}) = \frac{(\vec{v}_1 \cdot \vec{x})}{\|\vec{v}_1\|^2} \vec{v}_1 + \frac{(\vec{v}_2 \cdot \vec{x})}{\|\vec{v}_2\|^2} \vec{v}_2 = \frac{5}{2} \vec{v}_1 + \frac{8}{5} \vec{v}_2 = \begin{bmatrix} 4.1 \\ 5.7 \end{bmatrix}.$$



## Comments

There are other ways to realize the “projection” went awry:

- This is “life in  $\mathbb{R}^2$ ,” and since

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix},$$

are linearly independent  $\rightsquigarrow$  they form a basis for  $\mathbb{R}^2 \rightsquigarrow$  any projection of a vector  $\vec{w} \in \mathbb{R}^2$  onto the subspace  $V = \text{span}(\vec{v}_1, \vec{v}_2) \equiv \mathbb{R}^2$  must be the original vector  $\vec{w}$ .

- Even simpler, the famous *Method of the Eyeball* already showed that  $\vec{v}_1 + \vec{v}_2 = \vec{x}$ :

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}.$$



## Example: Doing it Right...

## Build an Orthonormal Basis

In this case, given a basis of  $\mathbb{R}^2$ , the answer is “obvious.”

Next, we develop (still in  $\mathbb{R}^2$  so we easily can visualize and use our intuition) a method for building an *orthonormal basis* given any starting basis.

Once we have the orthonormal basis, we can use the projection formula...

ⓘ The method will work in the general case: Given  $\vec{x} \in \mathbb{R}^n$ , and  $V = \text{span}(\vec{v}_1, \dots, \vec{v}_m) \subset \mathbb{R}^n$ ; compute  $\text{proj}_V(\vec{x})$ :

We find an orthonormal basis  $\vec{q}_1, \dots, \vec{q}_m$ , so that

$$V = \text{span}(\vec{v}_1, \dots, \vec{v}_m) = \text{span}(\vec{q}_1, \dots, \vec{q}_m);$$

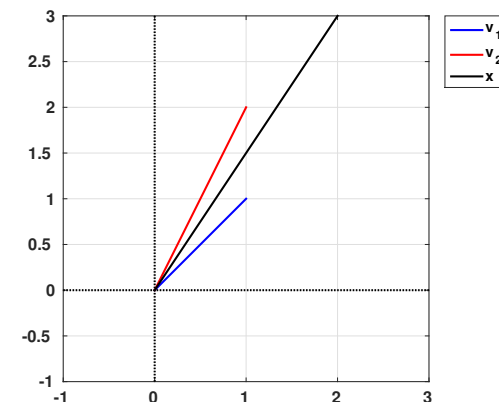
and then use the projection formula.



## Example: Doing it Right...

## Build an Orthonormal Basis

Figure: This is where we start.



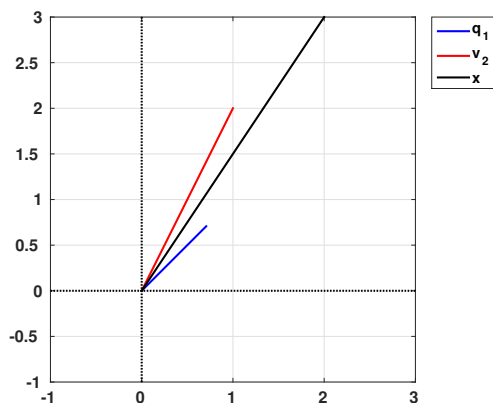
$$\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}; \quad \text{and } \vec{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}.$$



Example: Doing it Right...

(i) Rescale the first vector

Figure: Rescale\*  $\vec{v}_1$  to be norm 1, and call it  $\vec{q}_1$ .



$$\vec{q}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}; \quad \text{and } \vec{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}.$$

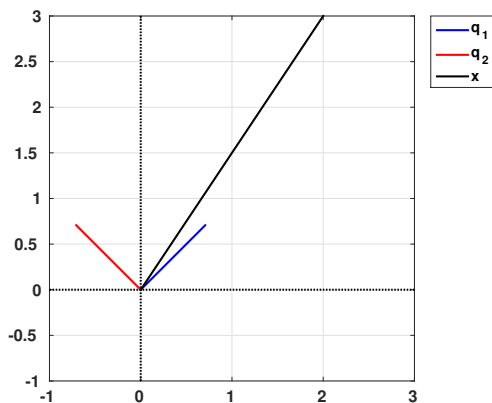
\* Divide by the original norm,  $\sqrt{2}$ .



Example: Doing it Right...

(iii) Rescale the  $\perp$  part, discard  $\parallel$  part

Figure: Next, throw away  $\vec{v}_2^{\parallel \vec{q}_1}$ ; rescale  $\vec{v}_2^{\perp \vec{q}_1}$  to norm 1, and name it  $\vec{q}_2$ .



$$\vec{q}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \vec{q}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}; \quad \text{and } \vec{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix},$$

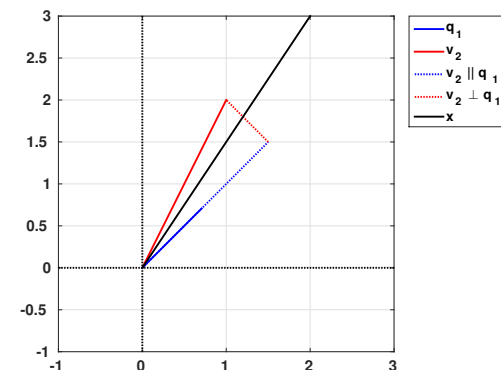
Now we have an orthonormal basis  $\Omega = \langle \vec{q}_1, \vec{q}_2 \rangle$ !



Example: Doing it Right...

(ii) Split the next vector into  $\parallel$  and  $\perp$  parts...

Figure: Next, project\*  $\vec{v}_2$  onto  $\vec{q}_1$  and get  $\vec{v}_2^{\parallel \vec{q}_1}$ , and  $\vec{v}_2^{\perp \vec{q}_1}$ .



$$\vec{q}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \vec{v}_2^{\parallel \vec{q}_1} = \begin{bmatrix} 1.5 \\ 1.5 \end{bmatrix}; \quad \vec{v}_2^{\perp \vec{q}_1} = \begin{bmatrix} -0.5 \\ 0.5 \end{bmatrix}; \quad \text{and } \vec{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}.$$

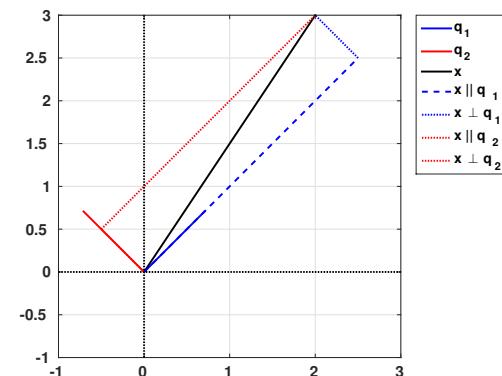
$$* \vec{v}_2^{\parallel \vec{q}_1} = (\vec{q}_1 \cdot \vec{v}_2) \vec{q}_1 = \left(\frac{3}{\sqrt{2}}\right) \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{3}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix}; \quad \vec{v}_2^{\perp \vec{q}_1} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} - \vec{v}_2^{\parallel \vec{q}_1} = \frac{1}{2} \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$



Example: Doing it Right...

(iv) Project using the new orthonormal basis!

Figure: Finally, we can use the projection formula.



$$\text{proj}_V(\vec{x}) = (\vec{q}_1 \cdot \vec{x}) \vec{q}_1 + (\vec{q}_2 \cdot \vec{x}) \vec{q}_2 = \frac{5}{\sqrt{2}} \vec{q}_1 + \frac{1}{\sqrt{2}} \vec{q}_2$$

$$\frac{5}{\sqrt{2}} \left( \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right) + \frac{1}{\sqrt{2}} \left( \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right) = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$



## Example: Doing it Right...

## Coordinates

In the context of [COORDINATES (NOTES#3.4)], we have

$$\text{BASIS: } \Omega = \langle \vec{q}_1, \vec{q}_2 \rangle$$

$$\text{COORDINATES: } [\vec{x}]_{\Omega} = \begin{bmatrix} \frac{5}{\sqrt{2}} \\ 1 \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 5 \\ 1 \\ 1 \end{bmatrix}$$



## Let's Ponder Higher Dimensions

When you have more basis vectors  $\vec{v}_1, \dots, \vec{v}_n$  needing orthogonalization (to make an orthonormal basis):

## Theorem (Gram-Schmidt Process (annotated))

- Start like we did:
  - $\vec{q}_1 = \vec{v}_1 / \|\vec{v}_1\|$
  - $\vec{w}_2 = \vec{v}_2 - (\vec{q}_1 \cdot \vec{v}_2)\vec{q}_1$ , note that this is a vector in the orthogonal complement of  $\text{span}(\vec{q}_1) = \text{span}(\vec{v}_1)$ .
  - $\vec{q}_2 = \vec{w}_2 / \|\vec{w}_2\|$
- Each time we grab a new vector ( $\vec{v}_k$ ), find a "help vector"  $\vec{w}_k$  in the orthogonal complement of the space spanned by the previously computed  $\vec{q}$ -vectors:
  - $\vec{w}_k = \vec{v}_k - (\vec{q}_1 \cdot \vec{v}_k)\vec{q}_1 - (\vec{q}_2 \cdot \vec{v}_k)\vec{q}_2 - \dots - (\vec{q}_{k-1} \cdot \vec{v}_k)\vec{q}_{k-1}$
- Then  $\vec{q}_k = \vec{w}_k / \|\vec{w}_k\|$ .



## Milking the Example for More Details...

We have performed a *Change of Basis*, in this case for the purpose of making the projection onto the subspace easily (after the change of basis, that is) computable.

It is "easy" to see that

$$A = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{2} & \frac{3}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} \end{bmatrix} = QR$$

we have  $A = QR$ , where  $Q$  is the new orthonormal basis, and  $R$  is an upper triangular matrix.

The entries in the  $R$  matrix are —  $\sqrt{2}$ : the original norm of  $\vec{v}_1$ ;  $\frac{3}{\sqrt{2}}$ : the dot product  $(\vec{q}_1 \cdot \vec{v}_2)$ ;  $\frac{1}{\sqrt{2}}$ : the norm of  $\vec{v}_2 - \vec{q}_1$ . Not likely a coincidence...



## The QR Factorization

The Gram-Schmidt process computed a change of basis from the old basis (funky-script-A)

$$\mathfrak{A} = (\vec{v}_1, \dots, \vec{v}_n)$$

to a new *orthonormal* basis (funky-script-Q)

$$\Omega = (\vec{q}_1, \dots, \vec{q}_n).$$

We describe the result using the change-of-basis-Matrix  $R$  from  $\mathfrak{A}$  to  $\Omega$ , writing

$$\underbrace{(\vec{v}_1 \ \dots \ \vec{v}_n)}_A = \underbrace{(\vec{q}_1 \ \dots \ \vec{q}_n)}_Q R$$



## Interpretations and Relations

With  $A = QR$ , we have to following relations:

- $[\vec{x}]_{\Omega} = R[\vec{x}]_{\mathcal{Q}}$ 
  - Multiplication by  $R$  moves us from  $A$ -coordinates to  $Q$ -coordinates.
- $\vec{x} = Q[\vec{x}]_{\Omega} = QR[\vec{x}]_{\mathcal{Q}}$ 
  - Multiplying the  $Q$ -coordinate vector by  $Q$  “builds” the vector  $\vec{x}$ .
- $\vec{x} = A[\vec{x}]_{\mathcal{Q}}$ 
  - Multiplying the  $A$ -coordinate vector by  $A$  “builds” the (same) vector  $\vec{x}$ .

The “burning” question is *how do we construct  $R$* ? It turns out we already have all the pieces, we just need some book-keeping.



## What's in $R$ ?

2 of 3

OK, let's rearrange the previous expression:

$$\vec{v}_k = \underbrace{(\vec{q}_1 \cdot \vec{v}_k)\vec{q}_1 - (\vec{q}_2 \cdot \vec{v}_k)\vec{q}_2 - \cdots - (\vec{q}_{k-1} \cdot \vec{v}_k)\vec{q}_{k-1}}_{\vec{v}_k^{\parallel}} + \underbrace{\vec{w}_k}_{\vec{v}_k^{\perp}}$$

The next thing we do is normalize  $\vec{v}_k^{\perp}$  to be norm 1, and name it  $\vec{q}_k$ ; which means we can write the relation above:

$$\vec{v}_k = \underbrace{(\vec{q}_1 \cdot \vec{v}_k)\vec{q}_1 + (\vec{q}_2 \cdot \vec{v}_k)\vec{q}_2 + \cdots + (\vec{q}_{k-1} \cdot \vec{v}_k)\vec{q}_{k-1}}_{\vec{v}_k^{\parallel}} + \underbrace{\|\vec{v}_k^{\perp}\|\vec{q}_k}_{\vec{v}_k^{\perp}}$$

This is the “recipe” for rebuilding the  $k^{\text{th}}$  column of  $A$  using the first  $k$  columns of  $Q$ . The entries in  $R$  are given by

- $r_{\ell,k} = (\vec{q}_\ell \cdot \vec{v}_k)$ ,  $\ell < k$ ; ( $r_{\ell,k} = 0$ ,  $\ell > k$ ), and
- $r_{k,k} = \|\vec{v}_k^{\perp}\|$ .



## What's in $R$ ?

1 of 3

If we think back to the  $k^{\text{th}}$  step, we compute

$$\underbrace{\vec{w}_k}_{\vec{v}_k^{\perp}} = \vec{v}_k - \underbrace{(\vec{q}_1 \cdot \vec{v}_k)\vec{q}_1 - (\vec{q}_2 \cdot \vec{v}_k)\vec{q}_2 - \cdots - (\vec{q}_{k-1} \cdot \vec{v}_k)\vec{q}_{k-1}}_{\vec{v}_k^{\parallel}}$$

$\vec{v}_k^{\perp}$  is orthogonal to  $V_{k-1} = \text{span}(\vec{q}_1, \dots, \vec{q}_{k-1}) = \text{span}(\vec{v}_1, \dots, \vec{v}_{k-1})$ , and  $\vec{v}_k^{\parallel} \in \text{span}(\vec{q}_1, \dots, \vec{q}_{k-1})$ .

**Note: Subspaces, Orthogonal Complements, and Bases**

We are constructing a sequence of subspace-pairs

$$V_k \oplus V_k^{\perp} = \mathbb{R}^n; \quad \dim(V_k) = k, \quad \dim(V_k^{\perp}) = (n - k); \quad k = 1, \dots, n$$

and orthonormal bases  $\Omega_k = (\vec{q}_1, \dots, \vec{q}_k)$  for each of the  $V_k$ -spaces; and we have  $V_{k-1} \subset V_k$  and  $V_k^{\perp} \subset V_{k-1}^{\perp}$ .

We are explicitly constructing  $V_k$  and  $\Omega_k$ ; whereas we're only concerned with a specific vector  $\vec{v}_k^{\perp} \in V_k^{\perp}$ .



## What's in $R$ ?

3 of 3

$$\vec{v}_k = \underbrace{(\vec{q}_1 \cdot \vec{v}_k)\vec{q}_1 + (\vec{q}_2 \cdot \vec{v}_k)\vec{q}_2 + \cdots + (\vec{q}_{k-1} \cdot \vec{v}_k)\vec{q}_{k-1}}_{\vec{v}_k^{\parallel}} + \underbrace{\|\vec{v}_k^{\perp}\|\vec{q}_k}_{\vec{v}_k^{\perp}}$$

$$\underbrace{\begin{bmatrix} \|\vec{v}_1\| & & & \\ & \ddots & & \\ & & \|\vec{v}_{k-1}\| & \\ & & & 0 \\ & & & & \ddots \\ & & & & & 0 \end{bmatrix}}_A = \underbrace{\begin{bmatrix} \vec{q}_1 & \cdots & \vec{q}_k & \cdots \end{bmatrix}}_Q \underbrace{\begin{bmatrix} r_{1,k} \\ \vdots \\ r_{k-1,k} \\ \|\vec{v}_k^{\perp}\| \\ \vdots \\ 0 \end{bmatrix}}_R$$



## Summarizing $\rightsquigarrow$ The QR-factorization

### Theorem (QR-Factorization)

Consider an  $(n \times m)$  matrix  $A$ , with linearly independent columns,  $\vec{v}_1, \dots, \vec{v}_m \in \mathbb{R}^n$ . Then there exists an  $(n \times m)$  matrix  $Q$  whose columns  $\vec{q}_1, \dots, \vec{q}_m \in \mathbb{R}^n$  are orthonormal, and an upper triangular matrix  $R$  with positive diagonal entries such that  $A = QR$ . This representation is unique.

Further

- $r_{11} = \|\vec{v}_1\|$ ,
- $r_{kk} = \|\vec{v}_k^\perp \text{span}(\vec{q}_1, \dots, \vec{q}_{k-1})\|$ ,  $k \in \{2, \dots, m\}$ , and
- $r_{\ell,k} = (\vec{q}_\ell \cdot \vec{v}_k)$ ,  $\ell \in \{1, \dots, k-1\}$ .

Note that

$$[QR\text{-factorization}] = [\text{Gram-Schmidt}] + [\text{Bookkeeping}].$$



## Suggested Problems 5.2

Available on Learning Glass videos:

5.2 — 3, 7, 13, 31, 32, 33, 35, 39



## Observations $A = [\vec{v}_1 \cdots \vec{v}_m] = QR$ , $A \in \mathbb{R}^{n \times m}$

- Note that  $\text{span}(\vec{q}_1, \dots, \vec{q}_k) = \text{span}(\vec{v}_1, \dots, \vec{v}_k)$ ,  $k = 1, \dots, m$  (that's the point — we are building an orthonormal set of vectors, describing the same subspaces spanned the columns of the matrix  $A$ )
- Let  $V_k = \text{span}(\vec{q}_1, \dots, \vec{q}_k) \equiv \text{span}(\vec{v}_1, \dots, \vec{v}_k)$ ; these subspaces are "nested":

$$V_0 \subset V_1 \subset \cdots \subset V_k,$$

$$\dim(V_0) \leq \dim(V_1) \leq \cdots \leq \dim(V_k),$$

(the maximal dimension is limited by the number of linearly independent vectors in  $\{\vec{v}_1, \dots, \vec{v}_k\}$ )

- **#ProjectionFestival**

$$\text{proj}_{V_k}(\vec{x}) = (\vec{x} \cdot \vec{q}_1)\vec{q}_1 + \cdots + (\vec{x} \cdot \vec{q}_k)\vec{q}_k$$



## Lecture – Book Roadmap

Lecture	Book, [GS5-]
5.1	§4.1, §4.2, §4.4
5.2	§4.1, §4.2, §4.4
5.3	§4.1, §4.2, §4.4



## Metacognitive Exercise — Thinking About Thinking &amp; Learning

I know / learned

Almost there

Huh?!?

Right After Lecture

After Thinking / Office Hours / SI-session

After Reviewing for Quiz/Midterm/Final



(5.2.13), (5.2.31)

(5.2.13) Perform the Gram-Schmidt process on the sequence of vectors given:

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} 0 \\ 2 \\ 1 \\ -1 \end{bmatrix}.$$

(5.2.31) Perform the Gram-Schmidt process on the following basis of  $\mathbb{R}^3$ :

$$\vec{v}_1 = \begin{bmatrix} a \\ 0 \\ 0 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} b \\ c \\ 0 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} d \\ e \\ f \end{bmatrix}.$$



(5.2.3), (5.2.7)

(5.2.3) Perform the Gram-Schmidt process on the sequence of vectors given:

$$\vec{v}_1 = \begin{bmatrix} 4 \\ 0 \\ 3 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 25 \\ 0 \\ -25 \end{bmatrix}.$$

(5.2.7) Perform the Gram-Schmidt process on the sequence of vectors given:

$$\vec{v}_1 = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} 18 \\ 0 \\ 0 \end{bmatrix}.$$



(5.2.33), (5.2.35)

(5.2.33) Find an orthonormal basis for the kernel of the matrix

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \end{bmatrix}.$$

(5.2.35) Find an orthonormal basis for the image of the matrix

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 2 & -2 & 0 \end{bmatrix}.$$



(5.2.39)

(5.2.39) Find an orthonormal basis  $\langle \vec{u}_1, \vec{u}_2, \vec{u}_3 \rangle$  of  $\mathbb{R}^3$ , such that

$$\text{span}(\vec{u}_1) = \text{span} \left( \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right),$$

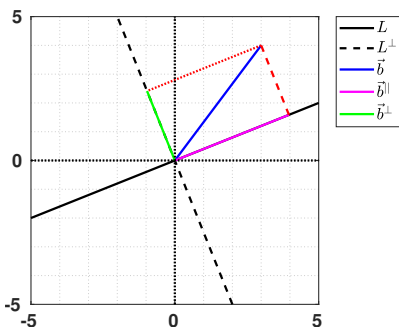
and

$$\text{span}(\vec{u}_1, \vec{u}_2) = \text{span} \left( \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \right),$$



Why Orthogonal Projections Matter  $\rightsquigarrow$  Solving the “Unsolvable”

Recall our old cartoon of orthogonal projections:



where  $\vec{w} \in \mathbb{R}^n$ ,  $L = \{ k\vec{w}, k \in \mathbb{R} \}$  is the (line) subspace of  $\mathbb{R}^n$ .

**Important Note:**

$\vec{w}^{\parallel}$  is the point (in the subspace  $L$ ) which is closest to  $\vec{w}$ .



Why Orthogonal Projections Matter  $\rightsquigarrow$  Solving the “Unsolvable”

Experience shows that at this point, most students tend to be a bit lost...

**Known** We need orthogonal bases to perform (correct) orthogonal projections to higher dimensional ( $n \geq 2$ ) subspaces.

**But** The previous example (projecting from  $\mathbb{R}^2 \rightarrow \mathbb{R}^2$ ) was not very satisfying...

**Mystery** Why are orthogonal projections such a big deal? (Bad reasons include:)

- “The professor said so.” (multiple times)
- “It’ll be on the test.”

The goal of the next example is to give some idea as to why orthogonal projections can be useful... while re-visiting and connecting several “old” ideas.



Why Orthogonal Projections Matter  $\rightsquigarrow$  Solving the “Unsolvable”

Now, let

$$A = \begin{bmatrix} | \\ | \\ \vec{w} \\ | \end{bmatrix} \in \mathbb{R}^{n \times 1},$$

then we are interested in solving the linear system  $A\vec{x} = \vec{b}$ , where  $\vec{x} \in \mathbb{R}^1$  (for now), and  $\vec{b} \in \mathbb{R}^n$ .

The system has a solution **if and only if**  $\vec{b} \in \text{im}(A) = L$ .

When  $\vec{b} \notin \text{im}(A)$  we can either

- say “🔧 you guys, I’m going home!” 🧑🏻, or
- extend the concept of a “solution” to the problem...





## Why Orthogonal Projections Matter $\rightsquigarrow$ Solving the "Unsolvable"

Since this is not a South Park episode, we decide to extend the concept of what it means to "solve" this problem:

We decide to look for a value  $\vec{x}^*$  which makes the **residual**\*

$$r(\vec{x}) = \|A\vec{x} - \vec{b}\|$$

as small as possible.

In our example, that value is  $\vec{x}^* = \left(\frac{\vec{b} \cdot \vec{w}}{\vec{w} \cdot \vec{w}}\right)$ , which makes  $A\vec{x}^* = \vec{b}^\parallel$ , and  $r(\vec{x}^*) = \|\vec{b}^\parallel - \vec{b}\| = \|\vec{b}^\perp\| = \|\vec{b}^\perp\|$ .

It is true in general that the shortest distance between  $\vec{b}$  and a subspace  $L$ , is  $\vec{b}^\perp = \vec{b} - \text{proj}_L(\vec{b})$ .

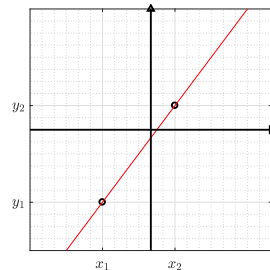
\* think of is as a measure of how far we are from solving the linear system in the "traditional" sense.



## Why Orthogonal Projections Matter $\rightsquigarrow$ Solving the "Unsolvable"

**Case ( $n = 2$ , two distinct points):** In this case we have a unique solution. In our notation the solutions are given by

$$\underbrace{\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \end{bmatrix}}_A \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$



which gives

$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \end{bmatrix}^{-1} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

where the inverse is guaranteed to exist when  $x_1 \neq x_2$ .

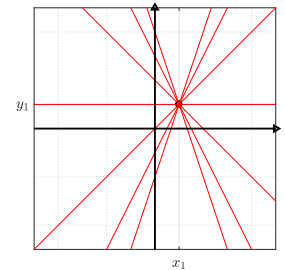


## Why Orthogonal Projections Matter $\rightsquigarrow$ Solving the "Unsolvable"

Next we consider a slightly different category of problems: fitting a straight line  $y = a + bx$  to some number of given points in the  $x$ - $y$ -plane,  $\{(x_k, y_k)\}_{k=1}^n$ .

**Case ( $n = 1$ , a single point):** In this case we have infinitely many solutions. In our notation the solutions are given by

$$\underbrace{\begin{bmatrix} 1 & x_1 \end{bmatrix}}_A \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} y_1 \end{bmatrix}$$



which gives

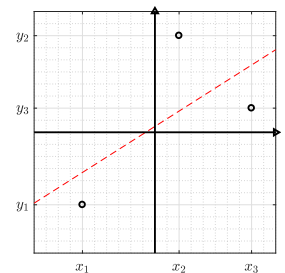
$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} y_1 \\ 0 \end{bmatrix} + s \begin{bmatrix} -x_1 \\ 1 \end{bmatrix}$$



## Why Orthogonal Projections Matter $\rightsquigarrow$ Solving the "Unsolvable"

**Case ( $n = 3$ , three distinct points):** In this case we have no solution. In our notation the solutions would be given by

$$\underbrace{\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \end{bmatrix}}_A \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$



which gives

$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \text{MAGIC} \\ \text{MATRIX} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} ???$$

There is no solution, unless the 3 points are on a common line...

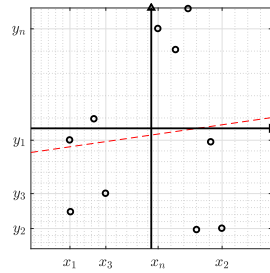


## Why Orthogonal Projections Matter $\rightsquigarrow$ Solving the “Unsolvable”

**Case ( $n =$  large, many (distinct) points):**

In this case we have no solution. In our notation the solutions would be given by

$$\underbrace{\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}}_{A \in \mathbb{R}^{n \times 2}} \begin{bmatrix} a \\ b \end{bmatrix} = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}}_{\vec{y} \in \mathbb{R}^n}$$



which gives

$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \text{MAGIC} \\ \text{MATRIX} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} ???$$

There is no solution, unless the ALL points are on a common line...



## Why Orthogonal Projections Matter $\rightsquigarrow$ Solving the “Unsolvable”

Now, if we are looking for a best-extended-concept-of-solution candidate; we compute  $\text{proj}_P(\vec{y}) \equiv \vec{y}^\parallel$ , and the system

$$\underbrace{\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}}_{A \in \mathbb{R}^{n \times 2}} \underbrace{\begin{bmatrix} a \\ b \end{bmatrix}}_{\vec{c}} = \text{proj}_P(\vec{y})$$

does have a unique solution, call it  $\vec{c}^*$ ; and the residual

$$r(\vec{c}^*) = \|A\vec{c}^* - \vec{y}\| = \|\vec{y}^\parallel - \vec{y}\| = \|\vec{y}^\perp\|$$

is minimized.



## Why Orthogonal Projections Matter $\rightsquigarrow$ Solving the “Unsolvable”

Staying in the general  $n =$  large case, with

$$\underbrace{\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}}_{A \in \mathbb{R}^{n \times 2}} \begin{bmatrix} a \\ b \end{bmatrix} = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}}_{\vec{y} \in \mathbb{R}^n}$$

In our linear algebra language, we “know” that  $P = \text{im}(A)$  is a 2-dimensional subspace of  $\mathbb{R}^n$  (the two columns are different, unless all the  $x_k$ s coincide)...

and, of course, we only have a solution if/when  $\vec{y}$  can be written as a linear combination of the columns of  $A \Leftrightarrow “\vec{y} \in \text{im}(A).”$



## Why Orthogonal Projections Matter $\rightsquigarrow$ Solving the “Unsolvable”

We have defined a new type of “solution” for inconsistent non-square (matrix) problems.

The way we have discussed it, the best name would be a

- “Minimum Residual Solution”

However, the most common mathematical name is the

- “Least Squares Solution”

In many applications (related to statistics), the most common name is the

- “Linear Regression Solution”



$$V = \{x_1 + x_2 + x_3 + x_4 = 0\} \subset \mathbb{R}^4$$

1 of 8

What is your problem?!?

Find an orthonormal basis for the subspace

$$V = \{x_1 + x_2 + x_3 + x_4 = 0\} \subset \mathbb{R}^4,$$

then project the vectors

$$\vec{y}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \text{and} \quad \vec{y}_2 = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$$

onto  $V$ .

$$V = \{x_1 + x_2 + x_3 + x_4 = 0\} \subset \mathbb{R}^4$$

2 of 8

- First, we need a basis for  $V$ ; finding  $\ker([1 \ 1 \ 1 \ 1])$  will do the trick.
- Since  $A = [1 \ 1 \ 1 \ 1]$  already is in rref, we can identify the solutions to  $\vec{A}\vec{x} = 0$ :

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = s \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + u \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix},$$

so our basis is

$$B_V = (\vec{v}_1, \vec{v}_2, \vec{v}_3) = \left( \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right); \quad A = \begin{bmatrix} -1 & -1 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

as an added bonus we will compute the QR-factorization of  $A$ .

$$V = \{x_1 + x_2 + x_3 + x_4 = 0\} \subset \mathbb{R}^4$$

3 of 8

$$\bullet \|\vec{v}_1\| = \sqrt{(-1)^2 + 1^2 + 0^2 + 0^2} = \sqrt{2}$$

$$\bullet \vec{q}_1 = \frac{1}{\|\vec{v}_1\|} \vec{v}_1$$

$$Q = \begin{bmatrix} -1/\sqrt{2} & \times & \times \\ 1/\sqrt{2} & \times & \times \\ 0 & \times & \times \\ 0 & \times & \times \end{bmatrix}, \quad R = \begin{bmatrix} \sqrt{2} & \times & \times \\ 0 & \times & \times \\ 0 & 0 & \times \end{bmatrix}$$

- We move on to  $\vec{v}_2$ ...



$$V = \{x_1 + x_2 + x_3 + x_4 = 0\} \subset \mathbb{R}^4$$

4 of 8

$$\bullet \vec{q}_1 \cdot \vec{v}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \frac{1}{\sqrt{2}} ((-1)^2 + 1 \times 0 + 0 \times 1 + 0 \times 0) = \frac{1}{\sqrt{2}}$$

$$\bullet \vec{v}_2^\perp = \vec{v}_2 - (\vec{q}_1 \cdot \vec{v}_2) \vec{q}_1 = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} - \left( \frac{1}{\sqrt{2}} \right) \left( \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \right) = \frac{1}{2} \begin{bmatrix} -1 \\ -1 \\ 2 \\ 0 \end{bmatrix}$$

$$\bullet \|\vec{v}_2^\perp\| = \frac{1}{2} \sqrt{1+1+4+0} = \frac{\sqrt{6}}{2}$$

$$\bullet \vec{q}_2 = \frac{1}{\|\vec{v}_2^\perp\|} \vec{v}_2^\perp = \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ -1 \\ 2 \\ 0 \end{bmatrix}$$

•

$$Q = \begin{bmatrix} -1/\sqrt{2} & -1/\sqrt{6} & \times \\ 1/\sqrt{2} & -1/\sqrt{6} & \times \\ 0 & 2/\sqrt{6} & \times \\ 0 & 0 & \times \end{bmatrix}, \quad R = \begin{bmatrix} \sqrt{2} & 1/\sqrt{2} & \times \\ 0 & \sqrt{6}/2 & \times \\ 0 & 0 & \times \end{bmatrix}$$



$$V = \{x_1 + x_2 + x_3 + x_4 = 0\} \subset \mathbb{R}^4$$

5 of 8

$$\bullet \vec{q}_1 \cdot \vec{v}_3 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}} ((-1)^2 + 1 \times 0 + 0 \times 1 + 0 \times 0) = \frac{1}{\sqrt{2}}$$

$$\bullet \vec{q}_2 \cdot \vec{v}_3 = \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ -1 \\ 2 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{6}} ((-1)^2 + (-1) \times 0 + 0 \times 1 + 0 \times 1) = \frac{1}{\sqrt{6}}$$

$$\bullet Q = \begin{bmatrix} -1/\sqrt{2} & -1/\sqrt{6} & \times \\ 1/\sqrt{2} & -1/\sqrt{6} & \times \\ 0 & 2/\sqrt{6} & \times \\ 0 & 0 & \times \end{bmatrix}, \quad R = \begin{bmatrix} \sqrt{2} & 1/\sqrt{2} & 1/\sqrt{2} \\ 0 & \sqrt{6}/2 & 1/\sqrt{6} \\ 0 & 0 & \times \end{bmatrix}$$



$$V = \{x_1 + x_2 + x_3 + x_4 = 0\} \subset \mathbb{R}^4$$

6 of 8

$$\bullet \vec{v}_3^\perp = \vec{v}_3 - (\vec{q}_1 \cdot \vec{v}_3)\vec{q}_1 - (\vec{q}_2 \cdot \vec{v}_3)\vec{q}_2:$$

$$\begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix} - \left(\frac{1}{\sqrt{2}}\right) \left(\frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}\right) - \left(\frac{1}{\sqrt{6}}\right) \left(\frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ -1 \\ 2 \\ 0 \end{bmatrix}\right) = \frac{1}{3} \begin{bmatrix} -1 \\ -1 \\ -1 \\ 3 \end{bmatrix}$$

$$\bullet \|\vec{v}_3^\perp\| = \frac{1}{3} \sqrt{1+1+1+9} = \frac{\sqrt{12}}{3}$$

$$\bullet \vec{q}_3 = \frac{1}{\|\vec{v}_3^\perp\|} \vec{v}_3^\perp = \frac{1}{\sqrt{12}} \begin{bmatrix} -1 \\ -1 \\ -1 \\ 3 \end{bmatrix}$$

$$\bullet Q = \begin{bmatrix} -1/\sqrt{2} & -1/\sqrt{6} & -1/\sqrt{12} \\ 1/\sqrt{2} & -1/\sqrt{6} & -1/\sqrt{12} \\ 0 & 2/\sqrt{6} & -1/\sqrt{12} \\ 0 & 0 & 3/\sqrt{12} \end{bmatrix}, \quad R = \begin{bmatrix} \sqrt{2} & 1/\sqrt{2} & 1/\sqrt{2} \\ 0 & \sqrt{6}/2 & 1/\sqrt{6} \\ 0 & 0 & \sqrt{12}/3 \end{bmatrix}$$



$$V = \{x_1 + x_2 + x_3 + x_4 = 0\} \subset \mathbb{R}^4$$

Projections!

7 of 8

$$\bullet \vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$$\bullet \vec{q}_1 \cdot \vec{v}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}} (-1 + 1 + 0 + 0) = 0$$

$$\bullet \vec{q}_2 \cdot \vec{v}_1 = \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ -1 \\ 2 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{6}} (-1 - 1 + 2 + 0) = 0$$

$$\bullet \vec{q}_3 \cdot \vec{v}_1 = \frac{1}{\sqrt{12}} \begin{bmatrix} -1 \\ -1 \\ -1 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{12}} (-1 - 1 - 1 + 3) = 0$$

$$\bullet \text{proj}_V(\vec{v}_1) = \vec{0}$$

• Of course! We constructed  $B_V = (\vec{v}_1, \vec{v}_2, \vec{v}_3)$  by finding all vectors orthogonal to  $\vec{v}_1$  ((Solving  $[1 \ 1 \ 1 \ 1]\vec{x} = \vec{0}$ ))



$$V = \{x_1 + x_2 + x_3 + x_4 = 0\} \subset \mathbb{R}^4$$

Projections!

8 of 8

$$\bullet \vec{v}_2 = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$$

$$\bullet \vec{q}_1 \cdot \vec{v}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} = \frac{1}{\sqrt{2}} (-1 + 2 + 0 + 0) = \frac{1}{\sqrt{2}}$$

$$\bullet \vec{q}_2 \cdot \vec{v}_2 = \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ -1 \\ 2 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} = \frac{1}{\sqrt{6}} (-1 - 2 + 6 + 0) = \frac{3}{\sqrt{6}}$$

$$\bullet \vec{q}_3 \cdot \vec{v}_2 = \frac{1}{\sqrt{12}} \begin{bmatrix} -1 \\ -1 \\ -1 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} = \frac{1}{\sqrt{12}} (-1 - 2 - 3 + 12) = \frac{6}{\sqrt{12}}$$

$$\bullet \text{proj}_V(\vec{v}_2) = \frac{1}{2} \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \frac{3}{6} \begin{bmatrix} -1 \\ -1 \\ 2 \\ 0 \end{bmatrix} + \frac{6}{12} \begin{bmatrix} -1 \\ -1 \\ -1 \\ 3 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} -3 \\ -1 \\ 1 \\ 3 \end{bmatrix}$$



5.2.35 and Beyond

What is your problem?!?

Given  $A$ , find an orthonormal basis for  $\text{im}(A)$ , and the  $QR$ -factorization  $QR = A$ :

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 2 & -2 & 0 \end{bmatrix}, \quad Q = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}, \quad R = \begin{bmatrix} \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot \\ 0 & 0 & \cdot \end{bmatrix}$$

$$\vec{v}_1 :: \|\vec{v}_1\| = \sqrt{1^2 + 2^2 + 2^2} = \sqrt{9} = 3; \quad \vec{q}_1 = \frac{\vec{v}_1}{\|\vec{v}_1\|} = \frac{1}{3} \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$$

$$Q = \begin{bmatrix} 1/3 & \cdot & \cdot \\ 2/3 & \cdot & \cdot \\ 2/3 & \cdot & \cdot \end{bmatrix}, \quad R = \begin{bmatrix} 3 & \cdot & \cdot \\ 0 & \cdot & \cdot \\ 0 & 0 & \cdot \end{bmatrix}$$



5.2.35 and Beyond

$$\begin{aligned} \vec{v}_3^\perp &= \vec{v}_3 - (\vec{q}_1 \cdot \vec{v}_3) \vec{q}_1 - (\vec{q}_2 \cdot \vec{v}_3) \vec{q}_2 \\ &= \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} - \underbrace{\left(\frac{1}{3} \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}\right)}_1 \left(\frac{1}{3} \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}\right) - \underbrace{\left(\frac{1}{3} \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}\right)}_1 \left(\frac{1}{3} \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}\right) \\ &= \frac{1}{3} \begin{bmatrix} 3 \\ 3 \\ 0 \end{bmatrix} - \frac{1}{3} \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} - \frac{1}{3} \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \end{aligned}$$

$$\|\vec{v}_3^\perp\| = 0, \quad \vec{q}_q = \frac{\vec{v}_3^\perp}{\|\vec{v}_3^\perp\|} = \begin{bmatrix} ? \\ ? \\ ? \end{bmatrix}$$

$$Q = \begin{bmatrix} 1/3 & 2/3 & ? \\ 2/3 & 1/3 & ? \\ 2/3 & -2/3 & ? \end{bmatrix}, \quad R = \begin{bmatrix} 3 & 0 & 1 \\ 0 & 3 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$



5.2.35 and Beyond

$$\vec{v}_2 :: \vec{v}_2^\perp = \vec{v}_2 - (\vec{q}_1 \cdot \vec{v}_2) \vec{q}_1 = \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix} - \underbrace{\left(\frac{1}{3} \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}\right)}_0 \left(\frac{1}{3} \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}$$

$$\|\vec{v}_2^\perp\| = \sqrt{2^2 + 1^2 + (-2)^2} = \sqrt{9} = 3; \quad \vec{q}_2 = \frac{\vec{v}_2^\perp}{\|\vec{v}_2^\perp\|} = \frac{1}{3} \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}$$

$$Q = \begin{bmatrix} 1/3 & 2/3 & \cdot \\ 2/3 & 1/3 & \cdot \\ 2/3 & -2/3 & \cdot \end{bmatrix}, \quad R = \begin{bmatrix} 3 & 0 & \cdot \\ 0 & 3 & \cdot \\ 0 & 0 & \cdot \end{bmatrix}$$



5.2.35 and Beyond

- $\vec{v}_3^\perp = 0$  means that  $\vec{v}_3$  is a linear combination of  $\vec{v}_1$  and  $\vec{v}_2$ .
- Therefore  $\text{im}(A) = \text{span}(\vec{v}_1, \vec{v}_2) = \text{span}(\vec{q}_1, \vec{q}_2)$
- We have 2 options for the  $QR$ -factorization:

$$A = \underbrace{\begin{bmatrix} 1/3 & 2/3 \\ 2/3 & 1/3 \\ 2/3 & -2/3 \end{bmatrix}}_{\text{"Economy Size" QR-factorization}} \begin{bmatrix} 3 & 0 & 1 \\ 0 & 3 & 1 \end{bmatrix}, \text{ or } \underbrace{\begin{bmatrix} 1/3 & 2/3 & 2/3 \\ 2/3 & 1/3 & -2/3 \\ 2/3 & -2/3 & 1/3 \end{bmatrix}}_{\text{"Full" QR-factorization}} \begin{bmatrix} 3 & 0 & 1 \\ 0 & 3 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Note that in the second version, we have added a third orthonormal vector to the  $Q$ -matrix, and a row of zeros to the  $R$ -matrix.

