Vera Xie
Ronald Nap
Andrew Rocco
3L = F 3:30 – 4:20 PM

Problem Set 5
Math 146 Spring 2023
Due: Friday, March 24 , 11:59 PM

**Note: Numbers 1-8 are "pen and paper" problem for which you should show work. Number 9 is a Matlab problem that requires code submissions. All answers should be submitted in a write-up.**

1. (34 points) Consider the following linear system $A\mathbf{x} = \mathbf{b}$:

$$\begin{bmatrix} -1 & 2 \\ 1 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -10 \\ 15 \\ 5 \end{bmatrix} \tag{1}$$

   (a) By performing the Gaussian Elimination, show that the system does not have a solution.
   (b) Construct the normal system and find the least-squares solution $\mathbf{x}_{\text{lsq}}$.
   (c) Compute the residual vector $\mathbf{r} = \mathbf{b} - A\mathbf{x}_{\text{lsq}}$.
   (d) Calculate the relative size of the residual,

$$\epsilon_{\text{rel}} = \frac{\|\mathbf{b} - A\mathbf{x}_{\text{lsq}}\|_2}{\|\mathbf{b}\|_2} \tag{2}$$

   (e) Show that the residual is orthogonal to each column vector of $A = [\mathbf{a}_1, \mathbf{a}_2]$.
   (f) Compute $\|A\mathbf{x}_{\text{lsq}}\|_2^2$ and $\|\mathbf{b}\|_2^2$ and show

$$\|\mathbf{b} - A\mathbf{x}_{\text{lsq}}\|_2^2 + \|A\mathbf{x}_{\text{lsq}}\|_2^2 = \|\mathbf{b}\|_2^2. \tag{3}$$

   (g) With a picture, explain the geometric meaning of the relation (3). Hint: The relation has the same form as the Pythagorean theorem.

2. (13 points) Consider the following vectors:

$$\mathbf{v}_1 = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix}. \tag{4}$$

   (a) Show that $\{\mathbf{v}_1,\ \mathbf{v}_2,\ \mathbf{v}_3\}$ forms an orthogonal basis for $\mathbb{R}^3$.
   (b) By normalizing $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$, construct an orthonormal basis $\{\mathbf{q}_1,\ \mathbf{q}_2,\ \mathbf{q}_3\}$ for $\mathbb{R}^3$.

3. (22 points) Consider the following basis for $\mathbb{R}^2$:

$$\mathbf{a}_1 = \begin{bmatrix} 3 \\ 4 \end{bmatrix}, \quad \mathbf{a}_2 = \begin{bmatrix} -2 \\ 1 \end{bmatrix} \tag{5}$$

   (a) Use the Gram Schmidt orthogonalization process to find the set $\{\mathbf{q}_1, \mathbf{q}_2\}$ that forms an orthonormal basis for $\mathbb{R}^2$:

(b) By plotting $\mathbf{q}_1$ and $\mathbf{q}_2$ in the $\mathbb{R}^2$ plane, explain how geometrically that $\{\mathbf{q}_1,\ \mathbf{q}_2\}$ forms an orthonormal basis for $\mathbb{R}^2$.

(c) For $\mathbf{w} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, compute its coordinates with respect to the basis $\{\mathbf{v}_1,\ \mathbf{v}_2\}$. In other words, compute $\alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}$ satisfying $\mathbf{w} = \alpha_1 \mathbf{q}_1 + \alpha_2 \mathbf{q}_2$.

4. (10 points) Let $\{\mathbf{q}_1,\ \mathbf{q}_2\}$ be an orthonormal basis for $\mathbb{R}^2$. Consider $\mathbf{w} = \alpha_1 \mathbf{q}_1 + \alpha_2 \mathbf{q}_2$.

   (a) Express $\mathbf{w}$ in terms of the matrix $Q = [\mathbf{q}_1\ \mathbf{q}_2]$ and vector $\alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}$.

   (b) By using the result of (a) and the fact that $Q^T Q = Q Q^T = I$, express $\alpha$ in terms of $Q^T$ and $\mathbf{w}$.

5. (10 points) A square matrix $Q$ is orthogonal if $Q^\intercal = Q^{-1}$. Show that the condition number (in the 2-norm) of an orthogonal matrix is 1. Note: This helps to explain why, unlike in using the normal equations, the conditioning of the linear problem is not worsened by using QR Factorization.

6. (10 points) Explain what may happen during the course of the Gram-Schmidt process if the matrix $A$ is rank deficient (i.e. if the columns are not linearly independent). What kind of fix could you put in your code to remedy this?

7. (12 points) In class, we claimed that $P = \widehat{Q}\widehat{Q}^\intercal$ is an orthogonal projector, where $\widehat{Q}$ is defined as the matrix whose columns, $\{\mathbf{q}_1, ...\mathbf{q}_n\}$ ($\mathbf{q}_i \in \mathbb{R}^m$) give an orthonormal basis for the range of matrix $A$. Prove this fact about $P$.

8. (12 points) In class, we defined a reflection defined by $F = I - 2P_v$, where $P_v$ is the orthogonal projection onto vector $\mathbf{v}$. Prove that $F$ is an orthogonal matrix.

9. (a) (25 points) Write functions to perform the reduced QR factorization using the classical Gram-Schmidt algorithm and the modified Gram-Schmidt algorithm. Your code should take as input $m \times n$ matrix $A$ with linearly independent column vectors (Note: this means you do NOT need to include your fix from Number (6)) and return matrices $\widehat{Q}$ and $\widehat{R}$. Also write functions to use these factorizations to solve the least squares problem. These functions should take as input matrices $A$ and vector $\mathbf{b}$ and return the least squares solution to the system $A\mathbf{x} = \mathbf{b}$. Validate your codes using a random matrix of size $50 \times 10$, and explain how you validated them. Note: Matlab's backslash performs least squares if $m > n$ for input matrix $A$.

   20 Bonus points for making a version of the Classical Gram-Schmidt factorization function that includes your fix from Number (6), validating it, and explaining how you

validated it. (Note: If you are checking if something is equal to 0, since there is numerical error, you may instead want to check if it is below some tolerance.)

(b) (12 points) Write a function to solve the least squares problem using the Householder method. For this you may want to simply write one function, taking as input the $m \times n$ matrix $A$ and vector $\mathbf{b}$ and return the least squares solution to the system $A\mathbf{x} = \mathbf{b}$. Validate your code and explain how you validated it.

(c) (25 points) Find an 11th degree polynomial that approximates the function $\cos(4x)$ evaluated at 50 equally spaced points on the interval $[0, 1]$ by solving the discrete least squares problem. You will be finding the 12 coefficients of $p(x) = c_0 + c_1 x + ... + c_{11} x^{11}$. The Vandermonde matrix that arises in the discrete problem is $50 \times 12$. Set up the linear leqst squares problem for the coefficients, and solve it using the methods below:

   i. Form the normal equations and solve them.
   ii. Use a QR factorization generated by classical Gram-Schmidt.
   iii. Use a QR factorization generated by modified Gram-Schmidt.
   iv. Use a QR factorization generated by Householder reflectors.
   v. Use Matlab's least squares solver by `A\b`. This is also based on QR, but it includes pivoting for added stability.

   For each method, give the 2-norm of the residual. Additionally, estimate what the relative error of the coefficients will be using condition numbers of the appropriate matrices. (You can skip this step for the Matlab function.) Lastly, estimate the relative error of the coefficients using the coefficients provided below, and in the file coefficients.mat. Give your results in a table, with one row for each method.

   For reference, using quad precision (128 bits vs the 53 bits of double precision), this problem can be solved with a residual norm of `7.999154576455076e-09`, giving the coefficients provided in the table below, as well as in the file coefficients.mat.

| | |
|---|---|
| $c_0$ | 1.000000000996606e+00 |
| $c_1$ | -4.227430949815150e-07 |
| $c_2$ | -7.999981235683346e+00 |
| $c_3$ | -3.187632625738558e-04 |
| $c_4$ | 1.066943079610163e+01 |
| $c_5$ | -1.382028878048870e-02 |
| $c_6$ | -5.647075625417684e+00 |
| $c_7$ | -7.531602738192263e-02 |
| $c_8$ | 1.693606966623459e+00 |
| $c_9$ | 6.032106743884792e-03 |
| $c_{10}$ | -3.742417027133638e-01 |
| $c_{11}$ | 8.804057595513443e-02 |

(d) (15 points) Comment on the performance of each method. Things to keep in mind: In class we discussed that the classical Gram-Schmidt method is unstable. Additionally, the Householder method is more stable than the modified Gram-Schmidt method. Lastly, using the condition numbers only allows you to *predict/estimate* the relative error of your coefficients. (Is there one error prediction that does not hold up? Which one and why?)

1. (34 points) Consider the following linear system $A\mathbf{x} = \mathbf{b}$:

$$\begin{bmatrix} -1 & 2 \\ 1 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -10 \\ 15 \\ 5 \end{bmatrix} \tag{1}$$

(a) By performing the Gaussian Elimination, show that the system does not have a solution.

(b) Construct the normal system and find the least-squares solution $\mathbf{x}_{\text{lsq}}$.

(c) Compute the residual vector $\mathbf{r} = \mathbf{b} - A\mathbf{x}_{\text{lsq}}$.

(d) Calculate the relative size of the residual,

$$\epsilon_{\text{rel}} = \frac{\|\mathbf{b} - A\mathbf{x}_{\text{lsq}}\|_2}{\|\mathbf{b}\|_2} \tag{2}$$

(e) Show that the residual is orthogonal to each column vector of $A = [\mathbf{a}_1, \mathbf{a}_2]$.

(f) Compute $\|A\mathbf{x}_{\text{lsq}}\|_2^2$ and $\|\mathbf{b}\|_2^2$ and show

$$\|\mathbf{b} - A\mathbf{x}_{\text{lsq}}\|_2^2 + \|A\mathbf{x}_{\text{lsq}}\|_2^2 = \|\mathbf{b}\|_2^2. \tag{3}$$

(g) With a picture, explain the geometric meaning of the relation (3). Hint: The relation has the same form as the Pythagorean theorem.

① linear system : $A\underline{x} = \underline{b}$ $\Longrightarrow \begin{bmatrix} -1 & 2 \\ 1 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -10 \\ 15 \\ 5 \end{bmatrix}$

a.) Show system does NOT have a solution by performing GAUSSIAN ELIMINATION

$$\begin{bmatrix} -1 & 2 \\ 1 & 1 \\ 2 & 1 \end{bmatrix} \begin{matrix} R_2 = R_2 + R_1 \\ R_3 = R_3 + 2R_1 \end{matrix} \begin{bmatrix} -1 & 2 \\ 0 & 3 \\ 0 & 5 \end{bmatrix} \; R_3 = R_3 - \frac{5}{3} R_2 \begin{bmatrix} -1 & 2 \\ 0 & 3 \\ 0 & 0 \end{bmatrix}$$

$$\longrightarrow \begin{bmatrix} -1 & 2 \\ 0 & 3 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -10 \\ 15 \\ 5 \end{bmatrix}$$

$$\begin{aligned} -x_1 + 2x_2 &= -10 \\ 3x_2 &= 15 \\ 0 &\neq 5 \end{aligned}$$

$\therefore$ **The system $A\underline{x} = \underline{b}$ does NOT have a solution by performing Gaussian Elimination**

b.) NORMAL system & LEAST-SQUARES solution $\underline{x}_{lsq}$

$\Longrightarrow$ Normal Eqns. : $\boxed{\underbrace{[A^T A]}_{B}\underline{x} = \underbrace{[A^T \underline{b}]}_{\underline{y}}}$ ; B : symmetric (+) definite, given A has full column rank

$$\longrightarrow A = \begin{bmatrix} -1 & 2 \\ 1 & 1 \\ 2 & 1 \end{bmatrix} \Longrightarrow A^T = \begin{bmatrix} -1 & 1 & 2 \\ 2 & 1 & 1 \end{bmatrix}$$

$$A^T A = \underset{2\times 3}{\begin{bmatrix} -1 & 1 & 2 \\ 2 & 1 & 1 \end{bmatrix}} \underset{3\times 2}{\begin{bmatrix} -1 & 2 \\ 1 & 1 \\ 2 & 1 \end{bmatrix}} = \underset{2\times 2}{\begin{bmatrix} 6 & 1 \\ 1 & 6 \end{bmatrix}}$$

$$A^T \underline{b} = \underset{2\times 3}{\begin{bmatrix} -1 & 1 & 2 \\ 2 & 1 & 1 \end{bmatrix}} \underset{3\times 1}{\begin{bmatrix} -10 \\ 15 \\ 5 \end{bmatrix}} = \underset{2\times 1}{\begin{bmatrix} 35 \\ 0 \end{bmatrix}}$$

$$\longrightarrow [A^TA]\underline{x} = [A^T\underline{b}] \Longrightarrow \boxed{\begin{bmatrix} 6 & 1 \\ 1 & 6 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 35 \\ 0 \end{bmatrix}}$$

[NORMAL SYSTEM]

$6x_1 + x_2 = 35$

$x_1 + 6x_2 = 0 \longrightarrow x_1 = -6x_2 = -6[-1] = 6 \;\therefore\; x_1 = 6$

$6[-6x_2] + x_2 = 35$

$= -36x_2 + x_2 = 35$

$= -35x_2 = 35 \;\therefore\; x_2 = -1$

$$\Longrightarrow \boxed{\underline{x}_{lsq} = \begin{bmatrix} 6 \\ -1 \end{bmatrix}}$$

[LEAST-SQUARES solution]

c.) RESIDUAL Vector : $\boxed{\underline{r} = \underline{b} - A\underline{x}_{lsq}}$

$$\longrightarrow A\underline{x}_{lsq} = \begin{bmatrix} -1 & 2 \\ 1 & 1 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 6 \\ -1 \end{bmatrix} = \begin{bmatrix} -8 \\ 5 \\ 11 \end{bmatrix}$$

$$\longrightarrow \underline{r} = \underline{b} - A\underline{x}_{lsq} \Longrightarrow \underline{r} = \begin{bmatrix} -10 \\ 15 \\ 5 \end{bmatrix} - \begin{bmatrix} -8 \\ 5 \\ 11 \end{bmatrix} = \begin{bmatrix} -2 \\ 10 \\ -6 \end{bmatrix} \Longrightarrow \boxed{\underline{r} = \begin{bmatrix} -2 \\ 10 \\ -6 \end{bmatrix}}$$

[RESIDUAL Vector]

d.) Relative SIZE of the residual : $\boxed{\varepsilon_{rel} = \dfrac{\|\overbrace{\underline{b} - A\underline{x}_{lsq}}^{\underline{r}}\|_2}{\|\underline{b}\|_2}}$

$$\longrightarrow \|\underline{b} - A\underline{x}_{lsq}\|_2 = \|\underline{r}\|_2$$

$$\underline{r} = \begin{bmatrix} -2 \\ 10 \\ -6 \end{bmatrix} \Longrightarrow \underline{r}^T = [-2 \;\; 10 \;\; -6] \;\Big\}\; \underline{r}^T\underline{r} = [-2 \;\; 10 \;\; -6]\begin{bmatrix} -2 \\ 10 \\ -6 \end{bmatrix} = 4 + 100 + 36 = 140 \Longrightarrow \|\underline{r}\|_2 = \sqrt{140}$$

$$\Longrightarrow \|\underline{b} - A\underline{x}_{lsq}\|_2 = \sqrt{140}$$

$$\longrightarrow \underline{b} = \begin{bmatrix} -10 \\ 15 \\ 5 \end{bmatrix} \Longrightarrow \underline{b}^T = [-10 \;\; 15 \;\; 5] \;\Big\}\; \underline{b}^T\underline{b} = [-10 \;\; 15 \;\; 5]\begin{bmatrix} -10 \\ 15 \\ 5 \end{bmatrix} = 100 + 225 + 25 = 350 \Longrightarrow \|\underline{b}\|_2 = \sqrt{350}$$

$$\longrightarrow \varepsilon_{rel} = \frac{\|\underline{b} - A\underline{x}_{lsq}\|_2}{\|\underline{b}\|_2} \Longrightarrow \varepsilon_{rel} = \frac{\sqrt{140}}{\sqrt{350}} = 0.6324555324 \Longrightarrow \boxed{\varepsilon_{rel} = 0.6325}$$

[RELATIVE SIZE of the RESIDUAL]

e.) Show that the residual is ORTHOGONAL to each column vector of $A = [\underline{a}_1\ \underline{a}_2]$

$\Longrightarrow$ 2 vectors ($\underline{u}$ & $\underline{v}$) of the SAME LENGTH are ORTHOGONAL iff $\boxed{\underline{u}^T \underline{v} = 0}$

$\longrightarrow$ $\underline{r} = \begin{bmatrix} -2 \\ 10 \\ -6 \end{bmatrix}$ & $A = \begin{bmatrix} -1 & 2 \\ 1 & 1 \\ 2 & 1 \end{bmatrix}$ ($\underline{a}_1$, $\underline{a}_2$)

$\underline{r}^T \cdot \underline{a}_1 = \begin{bmatrix} -2 & 10 & -6 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} = 2 + 10 - 12 = 0 \quad \therefore \underline{r}^T \cdot \underline{a}_1 = 0$ ✓

$\underline{a}_1^T \cdot \underline{r} = \begin{bmatrix} -1 & 1 & 2 \end{bmatrix} \begin{bmatrix} -2 \\ 10 \\ -6 \end{bmatrix} = 2 + 10 - 12 = 0 \quad \therefore \underline{a}_1^T \cdot \underline{r} = 0$ ✓

$\therefore$ **The residual ($\underline{r}$) is ORTHOGONAL to the 1st column of $A$ ($\underline{a}_1$)**

$\underline{r}^T \cdot \underline{a}_2 = \begin{bmatrix} -2 & 10 & -6 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} = -4 + 10 - 6 = 0 \quad \therefore \underline{r}^T \cdot \underline{a}_2 = 0$ ✓

$\underline{a}_2^T \cdot \underline{r} = \begin{bmatrix} 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} -2 \\ 10 \\ -6 \end{bmatrix} = -4 + 10 - 6 = 0 \quad \therefore \underline{a}_2^T \cdot \underline{r} = 0$ ✓

$\therefore$ **The residual ($\underline{r}$) is ORTHOGONAL to the 2nd column of $A$ ($\underline{a}_2$)**

$\Longrightarrow$ $\therefore$ **The residual is ORTHOGONAL to each column vector of $A = [\underline{a}_1\ \underline{a}_2]$**

f.) Compute $\|A\underline{x}_{lsq}\|_2^2$ & $\|\underline{b}\|_2^2$ and show that $\|\underline{b} - A\underline{x}_{lsq}\|_2^2 + \|A\underline{x}_{lsq}\|_2^2 = \|\underline{b}\|_2^2$

$\longrightarrow$ $A\underline{x}_{lsq} = \begin{bmatrix} -8 \\ 5 \\ 11 \end{bmatrix} \Longrightarrow (A\underline{x}_{lsq})^T = \begin{bmatrix} -8 & 5 & 11 \end{bmatrix}$

$(A\underline{x}_{lsq})^T \cdot A\underline{x}_{lsq} = \begin{bmatrix} -8 & 5 & 11 \end{bmatrix} \begin{bmatrix} -8 \\ 5 \\ 11 \end{bmatrix} = 64 + 25 + 121 = 210$

$\|A\underline{x}_{lsq}\|_2 = \sqrt{210} \Longrightarrow \|A\underline{x}_{lsq}\|_2^2 = 210$

$\longrightarrow$ $\|\underline{b}\|_2 = \sqrt{350} \Longrightarrow \|\underline{b}\|_2^2 = 350$

$\longrightarrow$ $\|\underline{b} - A\underline{x}_{lsq}\|_2 = \sqrt{140} \Longrightarrow \|\underline{b} - A\underline{x}_{lsq}\|_2^2 = 140$

$\longrightarrow$ $\|\underline{b} - A\underline{x}_{lsq}\|_2^2 + \|A\underline{x}_{lsq}\|_2^2 = \|\underline{b}\|_2^2 \Longrightarrow [140] + [210] = [350]$

$\therefore 350 = 350$ ✓

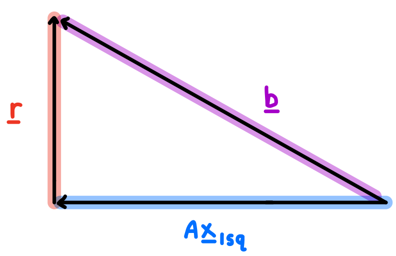
$$\Longrightarrow \therefore \boxed{\|\underline{b} - A\underline{x}_{lsq}\|_2^2 + \|A\underline{x}_{lsq}\|_2^2 = \|\underline{b}\|_2^2}$$

g.) With a PICTURE, explain the GEOMETRIC MEANING of the relation : $\boxed{\|\underline{b} - A\underline{x}_{lsq}\|_2^2 + \|A\underline{x}_{lsq}\|_2^2 = \|\underline{b}\|_2^2}$

HINT : The relation has the same form as the PYTHAGOREAN THEOREM

$\Longrightarrow$ The Pythagorean Theorem states $\boxed{\text{2 vectors } (\underline{u} \,\&\, \underline{v}) \text{ are ORTHOGONAL to each other if } \|\underline{u}+\underline{v}\|^2 = \|\underline{u}\|^2 + \|\underline{v}\|^2}$

$$\|\underbrace{\underline{b} - A\underline{x}_{lsq}}_{\underline{r}}\|_2^2 + \|A\underline{x}_{lsq}\|_2^2 = \|\underline{b}\|_2^2 \quad \Longrightarrow \|\underline{r}\|_2^2 + \|A\underline{x}_{lsq}\|_2^2 = \|\underline{b}\|_2^2$$




As we solved $\underline{x}$ using the normal equations, $\underline{x}$ is the solution to the Least-Squares problem ($\underline{x}_{lsq}$), where the normal equations are defined as $A^TA\underline{x} = A^T\underline{b}$. As the Least-Squares problem aims to minimize the residual, it is most ideal when the residual is 0. We can rearrange the normal equation, such that $A^T\underline{b} - A^TA\underline{x} = 0 \;\therefore\; A^T(\underline{b} - A\underline{x}) = 0$ $\therefore A^T\underline{r} = 0$. Thus, $\underline{b} - A\underline{x}$ OR the residual is ORTHOGONAL to the row space of $A^T$, and hence, the residual is ORTHOGONAL to the column space of $A$. Thus, an optimal solution is determined when the residual is orthogonal to the column space of $A$. $A\underline{x}$ is the point in the column space of $A$ that is the closest distance to $\underline{b}$, where $A\underline{x}$ is the projection of $\underline{b}$ onto the column space of $A$ and $\underline{b}$ is in the column space of $A$. The residual $(\underline{r} = \underline{b} - A\underline{x})$ is orthogonal to each column of $A$, and hence the residual is orthogonal to $A\underline{x}$. Thus, $A\underline{x}$ is the orthogonal projection of $\underline{b}$ onto the column space of $A$. The relation has a similar form as the Pythagorean Theorem. The relation states the sum of the squared lengths of $A\underline{x}$ and $\underline{b} - A\underline{x}$ is equal to the squared length of $\underline{b}$. From the picture, the geometric meaning of the relation shows the components of the relation forms the sides of a right triangle, where $\underline{b}$ is the hypotenuse of the right triangle, and $A\underline{x}$ and $\underline{r}$ are orthogonal to each other.

$$\overbrace{\begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{bmatrix}}^{A^T} \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\left.\begin{aligned} a_{11}r_1 + a_{21}r_2 + a_{31}r_3 &= 0 \\ a_{12}r_1 + a_{22}r_2 + a_{32}r_3 &= 0 \\ a_{13}r_1 + a_{23}r_2 + a_{33}r_3 &= 0 \end{aligned}\right\} \text{ROW space of } A^T = \text{COLUMN space of } A$$

2. (13 points) Consider the following vectors:

$$\mathbf{v}_1 = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix}. \tag{4}$$

(a) Show that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ forms an orthogonal basis for $\mathbb{R}^3$.

(b) By normalizing $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$, construct an orthonormal basis $\{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3\}$ for $\mathbb{R}^3$.

② vectors : $\underline{v}_1 = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}$, $\underline{v}_2 = \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}$, $\underline{v}_3 = \begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix}$

3 vectors

a.) Show that $\{\underline{v}_1, \underline{v}_2, \underline{v}_3\}$ forms an ORTHOGONAL BASIS for $\mathbb{R}^3$

$\Longrightarrow$ ORTHOGONAL Basis $\{\underline{v}_1, \dots, \underline{v}_n\}$ : $\underline{v}_i \cdot \underline{v}_j = 0$ if $i \neq j$

$\hookrightarrow$ $\underline{v}_1 \cdot \underline{v}_2 = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} \bullet \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix} = (2)(1) + (2)(-2) + (1)(2) = 2 - 4 + 2 = 0$ $\quad \therefore \underline{v}_1 \cdot \underline{v}_2 = 0$

$\hookrightarrow$ $\underline{v}_1 \cdot \underline{v}_3 = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix} = (2)(-2) + (2)(1) + (1)(2) = -4 + 2 + 2 = 0$ $\quad \therefore \underline{v}_1 \cdot \underline{v}_3 = 0$

$\hookrightarrow$ $\underline{v}_2 \cdot \underline{v}_3 = \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix} = (1)(-2) + (-2)(1) + (2)(2) = -2 - 2 + 4 = 0$ $\quad \therefore \underline{v}_2 \cdot \underline{v}_3 = 0$

$\Longrightarrow \therefore$ $\boxed{\{\underline{v}_1, \underline{v}_2, \underline{v}_3\} \text{ forms an ORTHOGONAL BASIS for } \mathbb{R}^3}$

b.) By NORMALIZING $\underline{v}_1, \underline{v}_2, \underline{v}_3$, construct an ORTHONORMAL BASIS $\{\underline{q}_1, \underline{q}_2, \underline{q}_3\}$ for $\mathbb{R}^3$

$\Longrightarrow$ NORMALIZING $\underline{v}_i$ : $\underline{q}_i = \dfrac{\underline{v}_i}{\|\underline{v}_i\|}$

[UNIT VECTOR]

$\hookrightarrow$ $\underline{q}_1 = \dfrac{\underline{v}_1}{\|\underline{v}_1\|} \Longrightarrow \|\underline{v}_1\| = \sqrt{(2)^2 + (2)^2 + (1)^2} = \sqrt{4+4+1} = \sqrt{9} = 3$ $\Bigg\}$ $\underline{q}_1 = \dfrac{1}{3} \cdot \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2/3 \\ 2/3 \\ 1/3 \end{bmatrix}$

$\hookrightarrow$ $\underline{q}_2 = \dfrac{\underline{v}_2}{\|\underline{v}_2\|} \Longrightarrow \|\underline{v}_2\| = \sqrt{(1)^2 + (-2)^2 + (2)^2} = \sqrt{1+4+4} = \sqrt{9} = 3$ $\Bigg\}$ $\underline{q}_2 = \dfrac{1}{3} \cdot \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix} = \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \end{bmatrix}$

$\hookrightarrow$ $\underline{q}_3 = \dfrac{\underline{v}_3}{\|\underline{v}_3\|} \Longrightarrow \|\underline{v}_1\| = \sqrt{(-2)^2 + (1)^2 + (2)^2} = \sqrt{4+1+4} = \sqrt{9} = 3$ $\Bigg\}$ $\underline{q}_3 = \dfrac{1}{3} \cdot \begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -2/3 \\ 1/3 \\ 2/3 \end{bmatrix}$

$$\Longrightarrow \boxed{\text{ORTHONORMAL Basis : } \{\underline{q}_1, \underline{q}_2, \underline{q}_3\} = \left\{ \begin{bmatrix} 2/3 \\ 2/3 \\ 1/3 \end{bmatrix}, \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \end{bmatrix}, \begin{bmatrix} -2/3 \\ 1/3 \\ 2/3 \end{bmatrix} \right\}}$$

3. (22 points) Consider the following basis for $\mathbb{R}^2$:

$$\mathbf{a}_1 = \begin{bmatrix} 3 \\ 4 \end{bmatrix}, \quad \mathbf{a}_2 = \begin{bmatrix} -2 \\ 1 \end{bmatrix} \tag{5}$$

(a) Use the Gram Schmidt orthogonalization process to find the set $\{\mathbf{q}_1, \mathbf{q}_2\}$ that forms an orthonormal basis for $\mathbb{R}^2$:

(b) By plotting $\mathbf{q}_1$ and $\mathbf{q}_2$ in the $\mathbb{R}^2$ plane, explain how geometrically that $\{\mathbf{q}_1, \mathbf{q}_2\}$ forms an orthonormal basis for $\mathbb{R}^2$.

(c) For $\mathbf{w} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, compute its coordinates with respect to the basis $\{\mathbf{v}_1, \mathbf{v}_2\}$. In other words, compute $\alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}$ satisfying $\mathbf{w} = \alpha_1 \mathbf{q}_1 + \alpha_2 \mathbf{q}_2$.

**3** Basis in $\mathbb{R}^2$ : $\underline{a}_1 = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ & $\underline{a}_2 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$

a.) GRAM–SCHMIDT ORTHOGONALIZATION process to find the set $\{\underline{q}_1, \underline{q}_2\}$ that forms an ORTHONORMAL BASIS for $\mathbb{R}^2$

$\Longrightarrow$ STEP 1 : $\underline{q}_1 = \dfrac{\underline{a}_1}{\|\underline{a}_1\|_2}$ $\Big\}$ UNIT LENGTH

$\hookrightarrow \underline{a}_1 = \begin{bmatrix} 3 \\ 4 \end{bmatrix} \Longrightarrow \underline{a}_1^T = [3 \quad 4] \Big\} \; \underline{a}_1^T \underline{a}_1 = [3 \quad 4]\begin{bmatrix} 3 \\ 4 \end{bmatrix} = 9 + 16 = 25 \Longrightarrow \|\underline{a}_1\|_2 = \sqrt{25} = 5$

$\longrightarrow \underline{q}_1 = \frac{1}{5} \cdot \begin{bmatrix} 3 \\ 4 \end{bmatrix} \Longrightarrow \underline{q}_1 = \begin{bmatrix} 3/5 \\ 4/5 \end{bmatrix}$

$\Longrightarrow$ STEP 2 : $\underline{v}_2 = \underline{a}_2 - \underbrace{\underline{q}_1(\underline{q}_1^T \underline{a}_2)}_{\text{PROJ. of } \underline{a}_2 \text{ ONTO } \underline{q}_1}$ $\Big\}$ $\therefore$ $\underline{v}_2$ is ORTHOGONAL to $\underline{q}_1$

$\hookrightarrow \underline{q}_1^T \underline{a}_2 = [3/5 \quad 4/5]\begin{bmatrix} -2 \\ 1 \end{bmatrix} = -\frac{6}{5} + \frac{4}{5} = -\frac{2}{5}$

$\hookrightarrow \underline{q}_1(\underline{q}_1^T \underline{a}_2) = \begin{bmatrix} 3/5 \\ 4/5 \end{bmatrix}\left(-\frac{2}{5}\right) = \begin{bmatrix} -6/25 \\ -8/25 \end{bmatrix}$

$$\hookrightarrow \underline{v}_2 = \underline{a}_2 - \underline{q}_1(\underline{q}_1^T \underline{a}_2) = \begin{bmatrix}-2\\1\end{bmatrix} - \begin{bmatrix}-6/25\\-8/25\end{bmatrix}$$

$$= \begin{bmatrix}-50/25\\25/25\end{bmatrix} - \begin{bmatrix}-6/25\\-8/25\end{bmatrix} = \begin{bmatrix}-44/25\\33/25\end{bmatrix} \implies \underline{v}_2 = \begin{bmatrix}-44/25\\33/25\end{bmatrix}$$

$\implies$ STEP 3 : $\underline{q}_2 = \dfrac{\underline{v}_2}{\|\underline{v}_2\|_2}$ } UNIT LENGTH

$$\hookrightarrow \underline{v}_2^T \underline{v}_2 = \begin{bmatrix}-44/25 & 33/25\end{bmatrix}\begin{bmatrix}-44/25\\33/25\end{bmatrix} = \frac{1936}{625} + \frac{1089}{625} = \frac{3025}{625} = 4.84 \implies \|\underline{v}_2\|_2 = \sqrt{4.84}$$

$$= 2.2 = \frac{22}{10} = \frac{11}{5}$$

$$\longrightarrow \underline{q}_2 = \frac{1}{11/5} \cdot \begin{bmatrix}-44/25\\33/25\end{bmatrix} = \frac{5}{11} \cdot \begin{bmatrix}-44/25\\33/25\end{bmatrix}$$

$$= \begin{bmatrix}-220/275\\165/275\end{bmatrix} = \begin{bmatrix}-44/55\\33/55\end{bmatrix} = \begin{bmatrix}-4/5\\3/5\end{bmatrix} \implies \underline{q}_2 = \begin{bmatrix}-4/5\\3/5\end{bmatrix}$$

$$\implies \boxed{\text{ORTHONORMAL Basis} : \{\underline{q}_1, \underline{q}_2\} = \left\{\begin{bmatrix}3/5\\4/5\end{bmatrix}, \begin{bmatrix}-4/5\\3/5\end{bmatrix}\right\}}$$

$$\left\{\underset{\underline{q}_1}{\begin{bmatrix}0.6\\0.8\end{bmatrix}}, \underset{\underline{q}_2}{\begin{bmatrix}-0.8\\0.6\end{bmatrix}}\right\}$$

b.) By PLOTTING $\underline{q}_1$ & $\underline{q}_2$ in the $\mathbb{R}^2$ plane, explain how geometrically that $\{\underline{q}_1, \underline{q}_2\}$ forms an ORTHONORMAL BASIS for $\mathbb{R}^2$

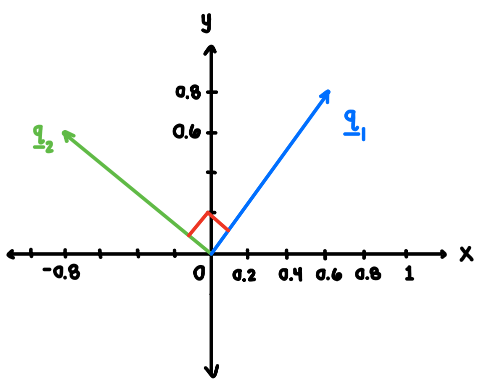
$$|q_1| = \sqrt{(0.6)^2 + (0.8)^2} = \sqrt{0.36 + 0.64} = 1$$

$$|q_2| = \sqrt{(-0.8)^2 + (0.6)^2} = \sqrt{0.36 + 0.64} = 1$$

$$\Downarrow$$

$\therefore |q_1| = 1$
$\therefore |q_2| = 1$ } Each vector has a UNIT LENGTH

$\therefore$ Each vector is NORMALIZED




The 2 vectors ($\underline{q}_1$ & $\underline{q}_2$) are ORTHOGONAL as the vectors formed a RIGHT ANGLE between them. The 2 vectors $\{\underline{q}_1, \underline{q}_2\} = \left\{\begin{bmatrix}3/5\\4/5\end{bmatrix}, \begin{bmatrix}-4/5\\3/5\end{bmatrix}\right\}$ forms an ORTHONORMAL basis as they are both orthogonal to each other and each vector is normalized as a unit vector, such as the magnitude of each vector is 1. Thus, the 2 vectors $\{\underline{q}_1, \underline{q}_2\}$ forms an orthonormal basis in $\mathbb{R}^2$.

c.) For $\underline{w} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, compute its COORDINATES w.r.t. the basis $\{\underline{v}_1, \underline{v}_2\}$

$\llcorner\!\rightarrow$ compute $\underline{\alpha} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}$ satisfying $\underline{w} = \alpha_1 \underline{q}_1 + \alpha_2 \underline{q}_2$

$\longrightarrow \{\underline{q}_1, \underline{q}_2\} = \left\{ \begin{bmatrix} 3/5 \\ 4/5 \end{bmatrix}, \begin{bmatrix} -4/5 \\ 3/5 \end{bmatrix} \right\}$

$\llcorner\!\rightarrow \underline{w} = \alpha_1 \begin{bmatrix} 3/5 \\ 4/5 \end{bmatrix} + \alpha_2 \begin{bmatrix} -4/5 \\ 3/5 \end{bmatrix} \Longrightarrow \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \alpha_1 \begin{bmatrix} 3/5 \\ 4/5 \end{bmatrix} + \alpha_2 \begin{bmatrix} -4/5 \\ 3/5 \end{bmatrix}$

$\frac{3}{5}\alpha_1 - \frac{4}{5}\alpha_2 = 1 \longrightarrow \frac{3}{5}\alpha_1 = 1 + \frac{4}{5}\alpha_2$

$\therefore \alpha_1 = \frac{5}{3} + \frac{4}{3}\alpha_2 = \frac{5}{3} + \frac{4}{3}\left[\frac{2}{5}\right]$

$\frac{4}{5}\alpha_1 + \frac{3}{5}\alpha_2 = 2$

$= \frac{5}{3} + \frac{8}{15}$

$\therefore \frac{4}{5}\left[\frac{5}{3} + \frac{4}{3}\alpha_2\right] + \frac{3}{5}\alpha_2 = 2$

$= \frac{25}{15} + \frac{8}{15} = \frac{33}{15}$

$\therefore \frac{4}{3} + \frac{16}{15}\alpha_2 + \frac{3}{5}\alpha_2 = 2$

$\therefore \alpha_1 = \frac{11}{5}$

$\therefore \frac{4}{3} + \frac{16}{15}\alpha_2 + \frac{9}{15}\alpha_2 = 2 \longrightarrow \frac{4}{3} + \frac{25}{15}\alpha_2 = 2$

$\therefore \frac{5}{3}\alpha_2 = \frac{6}{3} - \frac{4}{3}$

$\therefore \frac{5}{3}\alpha_2 = \frac{2}{3} \quad \therefore \alpha_2 = \frac{2}{5}$

$\Longrightarrow \boxed{\underline{\alpha} = \begin{bmatrix} 11/5 \\ 2/5 \end{bmatrix}}$

$\Longrightarrow$ CHECK :

$\underline{w} = \alpha_1 \begin{bmatrix} 3/5 \\ 4/5 \end{bmatrix} + \alpha_2 \begin{bmatrix} -4/5 \\ 3/5 \end{bmatrix} \Longrightarrow \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \frac{11}{5} \begin{bmatrix} 3/5 \\ 4/5 \end{bmatrix} + \frac{2}{5} \begin{bmatrix} -4/5 \\ 3/5 \end{bmatrix}$

$= \begin{bmatrix} 33/25 \\ 44/25 \end{bmatrix} + \begin{bmatrix} -8/25 \\ 6/25 \end{bmatrix}$

$= \begin{bmatrix} 25/25 \\ 50/25 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ ✓

4. (10 points) Let $\{\mathbf{q}_1, \mathbf{q}_2\}$ be an orthonormal basis for $\mathbb{R}^2$. Consider $\mathbf{w} = \alpha_1\mathbf{q}_1 + \alpha_2\mathbf{q}_2$.

(a) Express $\mathbf{w}$ in terms of the matrix $Q = [\mathbf{q}_1\ \mathbf{q}_2]$ and vector $\alpha = \begin{bmatrix}\alpha_1\\ \alpha_2\end{bmatrix}$.

(b) By using the result of (a) and the fact that $Q^TQ = QQ^T = I$, express $\alpha$ in terms of $Q^T$ and $\mathbf{w}$.

④ Let $\{\underline{q}_1, \underline{q}_2\}$ be an ORTHONORMAL BASIS for $\mathbb{R}^2$

$\{\vec{q}_1, \vec{q}_2\}$ is an orthonormal basis for $\mathbb{R}^2$ $\Rightarrow$ $\begin{cases}\text{Mutually orthogonal}\\ \text{Each vector has magnitude of 1}\end{cases}$

— consider : $\underline{w} = \alpha_1 \underline{q}_1 + \alpha_2 \underline{q}_2$

a.) Express $\underline{w}$ in terms of the matrix $Q = [\underline{q}_1\ \underline{q}_2]$ & vector $\underline{\alpha} = \begin{bmatrix}\alpha_1\\ \alpha_2\end{bmatrix}$

$$\vec{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} \underbrace{\begin{matrix} q_{11} \\ q_{21} \end{matrix}}_{\underline{q}_1} & \underbrace{\begin{matrix} q_{12} \\ q_{22} \end{matrix}}_{\underline{q}_2} \end{bmatrix} \cdot \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} \Rightarrow \vec{w} = [\vec{q}_1\ \vec{q}_2]\vec{\alpha} \Rightarrow \boxed{\vec{w} = Q\vec{\alpha}}$$

b.) By using the result in part (a) & $Q^TQ = QQ^T = I$, express $\underline{\alpha}$ in terms of $Q^T$ and $\underline{w}$

$$Q^TQ = QQ^T = I$$

$$\vec{w} = Q\vec{\alpha} \longrightarrow Q^T\vec{w} = \underbrace{Q^TQ}_{I}\vec{\alpha} \longrightarrow Q^T\vec{w} = I\vec{\alpha} \longrightarrow \vec{\alpha} = Q^T\vec{w}$$

METHOD 2

$$\vec{w} = Q\vec{\alpha} \Longrightarrow [\vec{w} = Q\vec{\alpha}]Q^{-1} \longrightarrow \vec{\alpha} = \underbrace{Q^{-1}}_{Q^{-1} = Q^T}\vec{w} \Longrightarrow \boxed{\vec{\alpha} = Q^T\vec{w}}$$

since $\{\underline{q}_1, \underline{q}_2\}$ is an ORTHONORMAL basis

$\therefore$ Q is ORTHOGONAL

5. (10 points) A square matrix $Q$ is orthogonal if $Q^\intercal = Q^{-1}$. Show that the condition number (in the 2-norm) of an orthogonal matrix is 1. Note: This helps to explain why, unlike in using the normal equations, the conditioning of the linear problem is not worsened by using QR Factorization.

⑤ SQUARE Matrix Q is ORTHOGONAL if $\boxed{Q^T = Q^{-1}}$

— Show that the CONDITION NUMBER (in the 2-NORM) of an ORTHOGONAL Matrix is 1

NOTE: This helps to explain, UNLIKE in using NORMAL EQNS., the CONDITIONING of the linear problem is NOT WORSENED by using QR – FACTORIZATION

$\Longrightarrow$ CONDITION NUMBER of a MATRIX : $\boxed{K(A) = \|A\| \, \|A^{-1}\|}$

want to show

$\Longrightarrow$ WTS : $K_2(A) = 1$

$Qx \Longrightarrow \|Qx\|_2 = \sqrt{(Qx)^T(Qx)} = \sqrt{(x^T \underbrace{Q^T)(Q}_{I} x)} = \sqrt{x^T x} = \|x\|_2 \quad \therefore \|Qx\|_2 = \|x\|_2$

$\Downarrow$

$\therefore \boxed{\|Q\|_2 = 1}$

$Q^{-1}x \Longrightarrow \|Q^{-1}x\|_2 = \sqrt{(Q^{-1}x)^T(Q^{-1}x)} = \sqrt{[x^T(\underbrace{Q^{-1}}_{Q^T})^T(\underbrace{Q^{-1}}_{Q^T}x)} = \sqrt{[x^T\underbrace{(Q^T)^T}_{Q}](Q^Tx)} = \sqrt{(x^T \underbrace{Q)(Q^T}_{I} x)} = \sqrt{x^T x} = \|x\|_2$

$\therefore \|Q^{-1}x\|_2 = \|x\|_2$

$\Downarrow$

$\therefore \boxed{\|Q^{-1}\|_2 = \|Q^T\|_2 = 1}$

$\longrightarrow K_2(A) = \|Q\|_2 \, \|Q^{-1}\|_2$

$= [1][1]$

$= 1 \quad \Longrightarrow \boxed{K_2(A) = 1}$

$\therefore$ The condition number in the 2-norm of an ORTHOGONAL Matrix is 1, which indicates orthogonal matrices have OPTIMAL CONDITIONING (i.e., WELL – CONDITIONED)

6. (10 points) Explain what may happen during the course of the Gram-Schmidt process if the matrix $A$ is rank deficient (i.e. if the columns are not linearly independent). What kind of fix could you put in your code to remedy this?

**6** Explain what may happen during the course of the Gram-Schmidt process to solve a Least-Squares problem if the matrix A is RANK DEFICIENT (i.e., if the columns are NOT L.I.).
What kind of fix could you put in your code to REMEDY this?

6) When the rank of our matrix is deficient this means that our matrix includes linearly dependent columns, this causes problems as our upper triangular matrix "R" will include 0 in our diagonal. Additionally, problems may arise when we compute the columns of Q. If R(j,j) is 0 we will divide by 0 This is because when we compute the value of our perpendicular vector, we subtract two nearly identical vectors. In order to remedy this problem we need to modify our Gram-Schmidt algorithm. We can do this by including an if statement to find the index and save the index of our linearly dependent columns. At the end of our Gram-Schmidt algorithm, we can extract dependent columns from our Q matrix and extract the rows/columns of our R matrix. In the end, we should have a Q that is nxr and an R that is rxr, where n is the columns of A and r is the rank of A.

7. (12 points) In class, we claimed that $P = \widehat{Q}\widehat{Q}^{\mathsf{T}}$ is an orthogonal projector, where $\widehat{Q}$ is defined as the matrix whose columns, $\{\mathbf{q}_1, \ldots \mathbf{q}_n\}$ $(\mathbf{q}_i \in \mathbb{R}^m)$ give an orthonormal basis for the range of matrix $A$. Prove this fact about $P$.

**7** $P = \hat{Q}\hat{Q}^T$ is an ORTHOGONAL PROJECTION, where $\hat{Q}$ is defined as the matrix whose columns $\{\underline{q}_1, \ldots, \underline{q}_n\}$ $(\underline{q}_i \in \mathbb{R}^m)$ give an ORTHONORMAL BASIS for the range of matrix $A$
— Prove this fact about P.

$\Longrightarrow$ ORTHONORMAL PROJECTIONS satisfy: ① $P^2 = P$
② $P^T = P$

$\longrightarrow$ $P^2 = P$

$$P^2 = (\hat{Q}\hat{Q}^T)^2 = (\hat{Q}\underbrace{\hat{Q}^T)(\hat{Q}}_{I}\hat{Q}^T) = \hat{Q}\hat{Q}^T = P \quad \Big\}\ ① \ P^2 = P$$

$\longrightarrow$ $P^T = P$

$$P^T = (\hat{Q}\hat{Q}^T)^T = \underbrace{(\hat{Q}^T)^T}_{Q}\hat{Q}^T = \hat{Q}\hat{Q}^T = P \quad \Big\}\ ② \ P^T = P$$

$\therefore$ $\boxed{P = \hat{Q}\hat{Q}^T \text{ is an ORTHOGONAL PROJECTION}}$

8. (12 points) In class, we defined a reflection defined by $F = I - 2P_v$, where $P_v$ is the orthogonal projection onto vector $\mathbf{v}$. Prove that $F$ is an orthogonal matrix.

(8) REFLECTION : $F = I - 2P_v$ ; $P_v$ : ORTHOGONAL PROJECTION onto $\underline{v}$

— Prove F is an ORTHOGONAL Matrix

$\Longrightarrow$ ORTHOGONAL MATRIX : A real, square matrix Q is ORTHOGONAL if its columns are ORTHOGONAL

$$Q^T Q = I$$

ORTHOGONAL Matrices are LENGTH – PRESERVING : $\|Q\underline{x}\|_2 = \|\underline{x}\|_2$

$\Longrightarrow$ PROJECTION onto $\underline{b}$ : $P_v \underline{b} = \dfrac{\overbrace{\underline{v}^T \underline{b}}^{\text{projecting onto } \underline{b}}}{\underline{v}^T \underline{v}} \underline{v} \Longrightarrow P_v = \dfrac{\underline{v}\,\underline{v}^T}{\underline{v}^T \underline{v}}$ Rearrange (need $v^T$ to operate on $\underline{b}$)

} w/o projecting onto a vector

→ Prove $P_v$ is an ORTHOGONAL PROJECTION

(1) $P_v^2 = P \Longrightarrow P_v^2 = \left(\dfrac{\underline{v}\,\underline{v}^T}{\underline{v}^T\underline{v}}\right)^2$

$$= \left(\frac{\underline{v}\,\underline{v}^T}{\underline{v}^T\underline{v}}\right)\left(\frac{\underline{v}\,\underline{v}^T}{\underline{v}^T\underline{v}}\right) = \frac{\underline{v}\,\underline{v}^T}{\underline{v}^T\underline{v}} = P$$

I

I

} (1) $P_v^2 = P_v$ ✓

(2) $P_v^T = P_v \Longrightarrow P_v^T = \left(\dfrac{\underline{v}\,\underline{v}^T}{\underline{v}^T\underline{v}}\right)^T$

$$= \frac{\overbrace{(\underline{v}^T)^T}^{\underline{v}} \underline{v}^T}{\underline{v}^T \underbrace{(\underline{v}^T)^T}_{\underline{v}}} = \frac{\underline{v}\,\underline{v}^T}{\underline{v}^T\underline{v}} = P$$

} (2) $P_v^T = P_v$ ✓

$\therefore$ $\boxed{P_v = \dfrac{\underline{v}\,\underline{v}^T}{\underline{v}^T\underline{v}} \text{ is an ORTHOGONAL Matrix}}$

→ WTS : $F^T F = I$

$F^T = (I - 2P_v)^T$

$$= [\underbrace{(-2P_v)^T}_{-2P_v \text{ since } P_v^T = P_v} + \underbrace{I^T}_{I}] = (-2P_v + I)$$

$$= (I - 2P_v) = F \Longrightarrow \therefore F^T = F$$

$F^T F = (I - 2P_v)^T (I - 2P_v)$

$$= [(-2P_v)^T + I^T](I - 2P_v) = (-2P_v + I)(I - 2P_v)$$

$$= (I - 2P_v)(I - 2P_v) = I^2 - 2P_v I - 2P_v I + 4P_v^2$$

$$= \underbrace{I^2}_{I} - 4\underbrace{P_v I}_{P_v} + 4\underbrace{P_v^2}_{P_v \text{ since } P_v^2 = P_v} = I - 4P_v + 4P_v = I$$ ✓

$\boxed{F^T F = I \quad \therefore F \text{ is an ORTHOGONAL Matrix}}$

$\longrightarrow$ WTS : $FF^T = I$

$\hookrightarrow FF^T = (I - 2P_v)(I - 2P_v)^T$

$= (I - 2P_v)[(-2P_v)^T + I^T] = (I - 2P_v)(-2P_v + I)$

$= (I - 2P_v)(I - 2P_v) = I^2 - 2P_v I - 2P_v I + 4P_v^2$

$= \underbrace{I^2}_{I} - 4\underbrace{P_v I}_{P_v} + 4\underbrace{P_v^2}_{P_v \text{ since } P_v^2 = P_v} = I - 4P_v + 4P_v = I$ ✓

$\boxed{FF^T = I \quad \therefore F \text{ is an ORTHOGONAL Matrix}}$

9. (a) (25 points) Write functions to perform the reduced QR factorization using the classical Gram-Schmidt algorithm and the modified Gram-Schmidt algorithm. Your code should take as input $m \times n$ matrix $A$ with linearly independent column vectors (Note: this means you do NOT need to include your fix from Number (6)) and return matrices $\widehat{Q}$ and $\widehat{R}$. Also write functions to use these factorizations to solve the least squares problem. These functions should take as input matrices $A$ and vector $\mathbf{b}$ and return the least squares solution to the system $A\mathbf{x} = \mathbf{b}$. Validate your codes using a random matrix of size $50 \times 10$, and explain how you validated them. Note: Matlab's backslash performs least squares if $m > n$ for input matrix $A$.

20 Bonus points for making a version of the Classical Gram-Schmidt factorization function that includes your fix from Number (6), validating it, and explaining how you validated it. (Note: If you are checking if something is equal to 0, since there is numerical error, you may instead want to check if it is below some tolerance.)

(b) (12 points) Write a function to solve the least squares problem using the Householder method. For this you may want to simply write one function, taking as input the $m \times n$ matrix $A$ and vector $\mathbf{b}$ and return the least squares solution to the system $A\mathbf{x} = \mathbf{b}$. Validate your code and explain how you validated it.

(c) (25 points) Find an 11th degree polynomial that approximates the function $\cos(4x)$ evaluated at 50 equally spaced points on the interval $[0, 1]$ by solving the discrete least squares problem. You will be finding the 12 coefficients of $p(x) = c_0 + c_1 x + \ldots + c_{11} x^{11}$. The Vandermonde matrix that arises in the discrete problem is $50 \times 12$. Set up the linear leqst squares problem for the coefficients, and solve it using the methods below:

i. Form the normal equations and solve them.
ii. Use a QR factorization generated by classical Gram-Schmidt.
iii. Use a QR factorization generated by modified Gram-Schmidt.
iv. Use a QR factorization generated by Householder reflectors.
v. Use Matlab's least squares solver by `A\b`. This is also based on QR, but it includes pivoting for added stability.

For each method, give the 2-norm of the residual. Additionally, estimate what the relative error of the coefficients will be using condition numbers of the appropriate matrices. (You can skip this step for the Matlab function.) Lastly, estimate the relative error of the coefficients using the coefficients provided below, and in the file coefficients.mat. Give your results in a table, with one row for each method.

For reference, using quad precision (128 bits vs the 53 bits of double precision), this problem can be solved with a residual norm of `7.999154576455076e-09`, giving the coefficients provided in the table below, as well as in the file coefficients.mat.

| | |
|---|---|
| $c_0$ | 1.000000000996606e+00 |
| $c_1$ | -4.227430949815150e-07 |
| $c_2$ | -7.999981235683346e+00 |
| $c_3$ | -3.187632625738558e-04 |
| $c_4$ | 1.066943079610163e+01 |
| $c_5$ | -1.382028878048870e-02 |
| $c_6$ | -5.647075625417684e+00 |
| $c_7$ | -7.531602738192263e-02 |
| $c_8$ | 1.693606966623459e+00 |
| $c_9$ | 6.032106743884792e-03 |
| $c_{10}$ | -3.742417027133638e-01 |
| $c_{11}$ | 8.804057595513443e-02 |

(d) (15 points) Comment on the performance of each method. Things to keep in mind: In class we discussed that the classical Gram-Schmidt method is unstable. Additionally, the Householder method is more stable than the modified Gram-Schmidt method. Lastly, using the condition numbers only allows you to *predict/estimate* the relative error of your coefficients. (Is there one error prediction that does not hold up? Which one and why?)

9

9a) To validate our Classical Gram-Schmidt algorithm, we generated a 50x10 "A" matrix with linearly independent column vectors and a 50x1 solution vector b. We then solved the system (or least squares problem) using our Classical Gram-Schmidt algorithm to decompose Q and R. We can solve our system by first generating our manipulated solution vector which we have as c in the code. This is done by multiplying the transpose of Q with our solution vector b, afterwards, we can then use backward substitution with R (upper triangular matrix) to solve for our solution. After solving the system with our algorithm, we then utilize Matlab's backslash feature to verify our code. When we compute the relative "error" (relative difference) between our algorithm and Matlab's backslash, we find a relative difference of around 4e-16. This is down at machine epsilon which indicates that our algorithm is working properly. The condition number of our randomly generated matrix was 2.51, this confirms we lose around 1 digit of accuracy which matches what we have. To validate our Modified Gram-Schmidt algorithm, similar to our Classical Gram-Schmidt algorithm we generated a 50x10 "A" matrix and a 50x1 solution vector b. We followed the same process of manipulating Ax=b to utilize Q and R to solve the system. Using our Modified Gram-Schmidt algorithm, we were able to achieve a relative difference of 3.71e-16 when comparing it to Matlab's backslash. The condition number of our randomly generated matrix was 2.01 which confirms that we expect to lose around 1 digit of accuracy which matches what we see.

9Bonus) When our input matrix A includes linearly dependent columns, this causes problems as our upper triangular matrix "R" will include 0 in our diagonal. This is because when we compute the value of our perpendicular vector, we do v=x-R(i, j)*Q(:, i), if x (the column vector) and R(i, j)*Q(:, i) yields the same or near the same value then, when we compute the norm of v, we will get 0. To avoid this situation from occurring, we want to include an if statement checking when we encounter a dependent column, we can then save the index in which the dependent column occurs and eliminate the column for Q and the column/row for R at the very end of our decomposition of Q, R. In the end, we should have a Q that is nxr and an R that is rxr, where n is the columns of A and r is the rank of A. To verify this process, I created a 50x10 "A" matrix with 2 linearly dependent columns and a solution vector of 50x1. Using the reconstructed Gram-Schmidt method, we decompose Q, R, and utilize the fact that Q'*b = y and Rx=y to solve our problem. Doing so yields us a solution of size 8x1 (minus the 2 dependent columns). In order to verify our solution, we must reconstruct our A matrix by removing the dependent columns and utilizing Matlab's backslash. Comparing Matlab's backslash with our reconstructed Gram Schmidt, we achieve a relative difference of 2e-15 and a condition number of 6.71. This confirms that we should expect to lose 1-2 digits of accuracy which matches what we have.

9b) To validate our Householder algorithm, we generated a 50x20 "A" matrix and a 50x1 solution vector b. The Householder algorithm inputs our A matrix and outputs our Q and R matrices. Using the fact that Q'*b=y and Rx=y we can utilize backward substitution to solve for our least squares problem. To verify our algorithm, we can compare our solution with Matlab's backslash feature, doing so we achieve a relative difference of 6.77e-16 and a condition number of 3.71. This is down at machine epsilon which indicates that our householder algorithm is working properly. Since our condition number is 3.71, we should expect to lose 1-2 digits of accuracy which match the relative difference we achieved meaning that our algorithm was successful.

9c.)

| Method | Norm_Residual | Condition_Number | Predicted_Rel_Err |
|---|---|---|---|
| "Normal equations" | 1.17196583566877e-07 | 1.35350699296282e+16 | 1.35350699296282 |
| "Classical Gram-Schmidt" | 4.73441028824497e-05 | 13371061.7367239 | 1.33710617367239e-09 |
| "Modified Gram-Schmidt" | 8.78897830226771e-09 | 117177656.225233 | 1.17177656225233e-08 |
| "Householder" | 7.99915502495182e-09 | 117177656.245018 | 1.17177656245018e-08 |

9d)

When using normal equations to solve our system (Cholesky along with forward/backward substitution) we find A'*A to have a condition number of 1.37e16, this means that we should expect a relative error of around 1.37. This is quite far from machine epsilon and we should only trust around 1 digit of accuracy for our solution. For Cholesky, off-diagonal elements are obtained by dividing by diagonal elements so if a diagonal element is near 0 we may encounter round-off error. Using Matlab's eig() function we indeed find eigenvalues near 0 which means our algorithm will be unstable. For our Classical Gram-Schmidt algorithm, we found R to have a condition number of 1.33e7, this means we should expect a relative error of around 1.33e-9. These results tell us we should expect around 9 digits of accuracy for our solution. However, we shouldn't trust this error because the Classical Gram-Schmidt algorithm is unstable. This is because we subtract off the projections of the jth vector, if the columns are nearly linearly dependent we will encounter round-off error. For the Modified Gram-Schmidt algorithm, we found R to have a condition number of 1.17e8, this means we should expect a relative error of around 1.17e-8. These results tell us we should expect around 8 digits of accuracy for our solution. Here we subtract off the projection onto the first vector then subtract off the projection of that result onto the second vector. This makes the algorithm slightly more stable than the Classical Gram Schmidt but not as stable as the Householder method. In the Householder method, we found a condition number of 1.17e8 meaning we should expect a relative error of around 1.17e-8. These results are nearly identical to the Modified Gram-Schmidt and Matlab algorithm in terms of 2-norm residual and condition numbers. It appears that Matlab may use a similar algorithm to the householder algorithm as the 2-norm residual and condition numbers are basically identical. It appears that Matlab has a slightly smaller 2-norm residual and condition number meaning the Matlab algorithm is the most accurate of the 5 methods.