

SI231 - Matrix Computations, Fall 2020-21

Homework Set #3

Prof. Yue Qiu and Prof. Ziping Zhao

Name: Xiaoyu Zhu **Major:** Master in CS

Student No.: 2020233243 **E-mail:** zhuxy3@shanghaitech.edu.cn

Acknowledgements:

- 1) Deadline: **2020-11-01 23:59:00**
 - 2) Submit your homework at **Gradescope**. Entry Code: **MY3XBJ**. Homework #3 contains two parts, the theoretical part and the programming part.
 - 3) About the theoretical part:
 - (a) Submit your homework in **Homework 3** in gradescope. Make sure that you have correctly select pages for each problem. If not, you probably will get 0 point.
 - (b) Your homework should be uploaded in the **PDF** format, and the naming format of the file is not specified.
 - (c) You need to use \LaTeX in principle.
 - (d) Use the given template and give your solution in English. Solution in Chinese is not allowed.
 - 4) About the programming part:
 - (a) Submit your codes in **Homework 3 Programming part** in gradescope.
 - (b) Detailed requirements see in Problem 2 and Problem 3.
 - 5) **No late submission is allowed.**
-

I. UNDERSTANDING PROJECTION

Problem 1. (5 points \times 3)

Suppose that $\mathbf{P} \in \mathbb{R}^{n \times n}$ is a projector onto a subspace \mathcal{U} along its orthogonal complement \mathcal{U}^\perp , then it is called the **orthogonal projector** onto \mathcal{U} .

- 1) Prove that an orthogonal projector must be singular if it is not an identity matrix.
- 2) What is the orthogonal projector onto \mathcal{U}^\perp along the subspace \mathcal{U} ?
- 3) Let \mathcal{U} and \mathcal{W} be two subspaces of a vector space \mathcal{V} , and denote $\mathbf{P}_\mathcal{U}$ and $\mathbf{P}_\mathcal{W}$ as the corresponding orthogonal projectors, respectively. Prove that $\mathbf{P}_\mathcal{U}\mathbf{P}_\mathcal{W} = 0$ if and only if $\mathcal{U} \perp \mathcal{W}$.

Solution.

- 1) Since \mathbf{P} is a projector onto a subspace, we have $\mathbf{P}^2 = \mathbf{P}$. It means that $(\mathbf{I} - \mathbf{P})\mathbf{P} = 0$. So we have $\text{rank}(\mathbf{P}) + \text{rank}(\mathbf{I} - \mathbf{P}) \leq \mathbf{n}$. If $\text{rank}(\mathbf{P}) = \mathbf{n}$, we have $\text{rank}((\mathbf{I} - \mathbf{P})) = 0$, it means that \mathbf{P} is an identity matrix. If $\text{rank}(\mathbf{P}) < \mathbf{n}$, it means that \mathbf{P} is singular.
- 2) $\mathbf{I} - \mathbf{P}$ is the orthogonal projector onto \mathcal{U}^\perp along the subspace \mathcal{U} . Since \mathbf{P} is a projector onto a subspace \mathcal{U} , so $\mathbf{P}\mathbf{x} = \mathbf{x}$ for any $x \in \mathcal{U}$ and $\mathbf{P}\mathbf{x} = 0$ for $x \in \mathcal{U}^\perp$. It means that $\mathbf{P}\mathbf{x} - \mathbf{x} = (\mathbf{I} - \mathbf{P})\mathbf{x} = 0$ for any $x \in \mathcal{U}$ and $\mathbf{x} - \mathbf{P}\mathbf{x} = \mathbf{x} - 0 = \mathbf{x}$ for any $x \in \mathcal{U}^\perp$. So $\mathbf{I} - \mathbf{P}$ is the orthogonal projector onto \mathcal{U}^\perp along the subspace \mathcal{U} .
- 3) If $\mathbf{P}_\mathcal{U}\mathbf{P}_\mathcal{W} = 0$, for any $x_\mathcal{U} \in \mathcal{U}$ and $x_\mathcal{W} \in \mathcal{W}$, we have

$$x_\mathcal{U}^T \mathbf{P}_\mathcal{U} \mathbf{P}_\mathcal{W} x_\mathcal{W} = x_\mathcal{U}^T \mathbf{P}_\mathcal{U}^T \mathbf{P}_\mathcal{W} x_\mathcal{W} = (\mathbf{P}_\mathcal{U} x_\mathcal{U})^T (\mathbf{P}_\mathcal{W} x_\mathcal{W}) = x_\mathcal{U}^T x_\mathcal{W} = 0$$

So for any $x_\mathcal{U} \in \mathcal{U}$ and $x_\mathcal{W} \in \mathcal{W}$, $x_\mathcal{U}^T x_\mathcal{W} = 0$, it means that $\mathcal{U} \perp \mathcal{W}$.

If $\mathcal{U} \perp \mathcal{W}$, for any $x_\mathcal{U} \in \mathcal{U}$ and $x_\mathcal{W} \in \mathcal{W}$, we have $x_\mathcal{U}^T x_\mathcal{W} = 0$. Since $\mathbf{P}_\mathcal{U}$ and $\mathbf{P}_\mathcal{W}$ are \mathcal{U} and \mathcal{W} orthogonal projects, we have

$$x_\mathcal{U}^T x_\mathcal{W} = (\mathbf{P}_\mathcal{U} x_\mathcal{U})^T (\mathbf{P}_\mathcal{W} x_\mathcal{W}) = x_\mathcal{U}^T \mathbf{P}_\mathcal{U} \mathbf{P}_\mathcal{W} x_\mathcal{W} = 0$$

so $\mathbf{P}_\mathcal{U} \mathbf{P}_\mathcal{W} = 0$. Above all, $\mathbf{P}_\mathcal{U} \mathbf{P}_\mathcal{W} = 0$ if and only if $\mathcal{U} \perp \mathcal{W}$.

II. LEAST SQUARE (LS) PROGRAMMING.

Problem 2. (10 points + 10 points + 5 points)

Write programs to solve the least square problem with specified methods, any programming language is suitable.

$$\mathbf{x} = \arg \min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}), \quad f(\mathbf{x}) = \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2$$

where $\mathbf{A} \in \mathbb{R}^{m \times n}$ is a matrix representing the predefined data set with m data samples of n dimensions ($m=1000$, $n=210$), and $\mathbf{y} \in \mathbb{R}^m$ represents the labels. The data samples are provided in the "data.txt" file, and the labels are provided in the "label.txt" file, you are supposed to load the data before solving the problem.

- 1) Solve the LS with gradient decent method.

The gradient descent method for solving problem updates \mathbf{x} as

$$\mathbf{x} = \mathbf{x} - \gamma \cdot \nabla_{\mathbf{x}} f(\mathbf{x}),$$

where γ is the step size of the gradient decent methods. We suggest that you can set $\gamma = 1e - 5$.

- 2) Solve the LS by the method of normal equation with Cholesky decomposition and forward/backward substitution.
- 3) Compare two methods above.
 - (a) Basing on the true running results from the program, count the number of "flops";
 - (b) Compare gradient norm and loss $f(\mathbf{x})$ for results $\mathbf{x} = \mathbf{x}_{\text{LS}}$ of above two algorithms.

Notation*: "flop": one flop means one floating point operation, i.e., one addition, subtraction, multiplication, or division of two floating-point numbers, in this problem each floating points operation $+$, $-$, \times , \div , $\sqrt{\cdot}$ counts as one "flop".

Hint for gradient decent programming:

- 1) **Step size selection:** to ensure the convergence of the method, γ is supposed to be selected properly (large step size may accelerate the convergence rate but also may lead to instability, A sufficiently small compensation always ensures that the algorithm converges).
- 2) **Terminal condition:** the gradient decent is an iteration algorithm that need a terminal condition. In this problem, the algorithm can stop when the gradient of the loss function $f(\mathbf{x})$ at current \mathbf{x} is small enough.

Remarks:

- The solution of the two methods should be printed in files named "sol1.txt" and "sol2.txt" and submitted in gradescope. The format should be same as the input file (210 rows plain text, each rows is a dimension of the final solution).
- Make sure that your codes are executable and are consistent with your solutions.

Solution. Denote $m = 1000$, $n = 210$ According to the calculation of matlab ,there are

$$\mathbf{N} = 102$$

cycle in gradient decent method. In each cycle ,there are $flop_i = n * m * n + n^2 + m * n + 3n = n^2 m + n^2 + mn + 3n$ flops. so gradient descent method are $N(n^2 m + n^2 + mn + 3n)$ flops.

For the Cholesky decomposition , there are $\sum_{k=2}^{n-1} ((2k-1)(n-k) + (2k-2)) \approx \frac{1}{3}n^3$ flops. The backward and forward substitution all need about n^2 flops. So the second method need about $\frac{1}{3}n^3$ flops.

III. UNDERSTANDING THE QR FACTORIZATION

Problem 3 [Understanding the Gram-Schmidt algorithm.]. (5 points + 7 points + 6 points + 7 points)

1) Consider the subspace \mathcal{S} spanned by $\{\mathbf{a}_1, \dots, \mathbf{a}_4\}$, where

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \quad \mathbf{a}_2 = \begin{bmatrix} 2 \\ 3 \\ 4 \\ 5 \end{bmatrix}, \quad \mathbf{a}_3 = \begin{bmatrix} 3 \\ 4 \\ 5 \\ 6 \end{bmatrix}, \quad \mathbf{a}_4 = \begin{bmatrix} 3 \\ 5 \\ 7 \\ 11 \end{bmatrix}.$$

Use the **classical** Gram-Schmidt algorithm (See Algorithm ??), find a set of orthonormal basis $\{\mathbf{q}_i\}$ for \mathcal{S} by hand (derivation is expected). Do not use decimals in your answers, fraction and n -th roots of numbers are accepted. Verify the orthonormality of the found basis.

Algorithm 1: Classical Gram-Schmidt algorithm

Input : A collection of linearly independent vectors $\mathbf{a}_1, \dots, \mathbf{a}_n$.

1 **Initialization:** $\tilde{\mathbf{q}}_1 = \mathbf{a}_1, \mathbf{q}_1 = \tilde{\mathbf{q}}_1 / \|\tilde{\mathbf{q}}_1\|_2$

2 **for** $i = 2, \dots, n$ **do**

3 $\tilde{\mathbf{q}}_i = \mathbf{a}_i - \sum_{j=1}^{i-1} (\mathbf{q}_j^T \mathbf{a}_i) \mathbf{q}_j$

4 $\mathbf{q}_i = \tilde{\mathbf{q}}_i / \|\tilde{\mathbf{q}}_i\|_2$

5 **end**

Output: $\mathbf{q}_1, \dots, \mathbf{q}_n$

2) Orthogonal projection of vector \mathbf{a} onto a nonzero vector \mathbf{b} is defined as

$$\text{proj}_{\mathbf{b}}(\mathbf{a}) = \frac{\langle \mathbf{a}, \mathbf{b} \rangle}{\langle \mathbf{b}, \mathbf{b} \rangle} \mathbf{b},$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product of vectors. And for subspace \mathcal{M} with orthonormal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$, the orthogonal projector onto subspace \mathcal{M} is given by

$$\mathbf{P} = \mathbf{U}\mathbf{U}^T, \quad \mathbf{U} = [\mathbf{u}_1 | \dots | \mathbf{u}_k].$$

In the context of **projection of vector** and **projection onto subspace** respectively, can you give another two understandings of the classical Gram-Schmidt algorithm?

3) Consider the subspace \mathcal{S} spanned by $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$,

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ \epsilon \\ \epsilon \end{bmatrix}, \quad \mathbf{a}_2 = \begin{bmatrix} 1 \\ \epsilon \\ 0 \end{bmatrix}, \quad \mathbf{a}_3 = \begin{bmatrix} 1 \\ 0 \\ \epsilon \end{bmatrix},$$

where ϵ is a small real number such that $1 + k\epsilon^2 = 1$ ($k \in \mathbb{N}^+$). First complete the pseudo algorithm in Algorithm ??. Then use the **classical** Gram-Schmidt algorithm and the **modified** Gram-Schmidt algorithm respectively, find two sets of basis for \mathcal{S} by hand (derivation is expected). Are the two sets of basis the same? If not, which one is the desired orthonormal basis? Report what you have found.

Algorithm 2: Modified Gram-Schmidt algorithm

Input : A collection of linearly independent vectors $\mathbf{a}_1, \dots, \mathbf{a}_n$.

1 **Initilization:**

2 *Complete your algorithm here...*

Output: $\mathbf{q}_1, \dots, \mathbf{q}_n$

- 4) **Programming part:** In this part, you are required to code both the **classical Gram-Schmidt** and the **modified Gram-Schmidt** algorithms. For $\epsilon = 1\text{e}-4$ and $\epsilon = 1\text{e}-9$ in sub-problem 2), give the outputs of two algorithms and calculate $\|\mathbf{Q}^T \mathbf{Q} - \mathbf{I}\|_F$, where $\mathbf{Q} = [\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3]$.

Remarks:

- Coding languages are not restricted, but do not use built-in function such as [qr](#).
- When handing in your homework in gradescope, package all your codes into `your_student_id+hw3_code.zip` and upload. In the package, you also need to include a file named `README.txt/md` to clearly identify the function of each file.
- Make sure that your codes can run and are consistent with your solutions.

Solution.

1) According to Algorithm 1

$$\tilde{q}_1 = a_1, \quad q_1 = \frac{\tilde{q}_1}{\|\tilde{q}_1\|} = \frac{\sqrt{30}}{30} \begin{bmatrix} 1 & 2 & 3 & 4 \end{bmatrix}^T$$

$$\tilde{q}_2 = a_2 - (q_1^T a_2)q_1 = \frac{1}{3} \begin{bmatrix} 2 & 1 & 0 & -1 \end{bmatrix}^T, \quad q_2 = \frac{\tilde{q}_2}{\|\tilde{q}_2\|} = \frac{\sqrt{6}}{6} \begin{bmatrix} 2 & 1 & 0 & -1 \end{bmatrix}^T$$

$\tilde{q}_3 = a_3 - (q_1^T a_3)q_1 - (q_2^T a_3)q_2 = 0$, so a_3 can be represented linearly by a_1 and a_2 .

$$\tilde{q}_4 = a_4 - (q_1^T a_4)q_1 - (q_2^T a_4)q_2 = \frac{1}{5} \begin{bmatrix} 2 & -1 & -4 & 3 \end{bmatrix}^T, \quad q_4 = \frac{\tilde{q}_4}{\|\tilde{q}_4\|} = \frac{\sqrt{30}}{30} \begin{bmatrix} 2 & -1 & -4 & 3 \end{bmatrix}^T$$

We can get it easily that $q_1^T q_2 = q_1^T q_4 = q_2^T q_4 = 0$, so set of the basis q_1, q_2, q_4 is orthonormal basis.

2) • projection of vector:

$(q_1^T a_2)q_1$ is the orthogonal projection of vector a_2 onto the vector q_1 , so $\tilde{q}_2 = a_2 - (q_1^T a_2)q_1$ indicates a vector that a_2 eliminated q_1 direction. by parity of reasoning, $\tilde{\mathbf{q}}_i = \mathbf{a}_i - \sum_{j=1}^{i-1} (\mathbf{q}_j^T \mathbf{a}_i) \mathbf{q}_j$ indicates a vector that a_i eliminated $q_1, q_2 \cdots q_{i-1}$ directions.

• projection onto subspace:

First unit the vector a_1 , get vector q_1 , then we have a orthogonal projector matrix $\mathbf{P}_1 = \mathbf{Q}_1 \mathbf{Q}_1^T, \mathbf{Q}_1 = [q_1]$
 $\tilde{q}_2 = a_2 - (q_1^T a_2)q_1 = a_2 - \mathbf{Q}_1 \mathbf{Q}_1^T a_2$ indicates a vector that a_2 eliminated P_1 direction. by parity of reasoning, $\tilde{\mathbf{q}}_i = \mathbf{a}_i - \sum_{j=1}^{i-1} (\mathbf{q}_j^T \mathbf{a}_i) \mathbf{q}_j = \mathbf{a}_i - \mathbf{Q}_i \mathbf{Q}_i^T \mathbf{A}_i$ indicates a vector that a_i eliminated \mathbf{P}_i direction. ($\mathbf{A}_i = [a_1, a_2 \cdots a_i]$)

3) • classical Gram-Schmidt

$$\tilde{q}_1 = a_1, \quad q_1 = \frac{\tilde{q}_1}{\|\tilde{q}_1\|} = \begin{bmatrix} 1 & \epsilon & \epsilon \end{bmatrix}^T$$

$$\tilde{q}_2 = a_2 - (q_1^T a_2)q_1 = \begin{bmatrix} 0 & 0 & -\epsilon \end{bmatrix}^T, \quad q_2 = \frac{\tilde{q}_2}{\|\tilde{q}_2\|} = \begin{bmatrix} 0 & 0 & -1 \end{bmatrix}^T$$

$$\tilde{q}_3 = a_3 - (q_1^T a_3)q_1 - (q_2^T a_3)q_2 = \begin{bmatrix} 0 & -\epsilon & -\epsilon \end{bmatrix}^T, \quad q_3 = \frac{\tilde{q}_3}{\|\tilde{q}_3\|} = \begin{bmatrix} 0 & -\frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{bmatrix}^T$$

• modified Gram-Schmidt

$$\tilde{q}_1 = a_1, \quad q_1 = \frac{\tilde{q}_1}{\|\tilde{q}_1\|} = \begin{bmatrix} 1 & \epsilon & \epsilon \end{bmatrix}^T$$

$$\tilde{q}_2 = a_2 - (q_1^T a_2)q_1 = \begin{bmatrix} 0 & 0 & -\epsilon \end{bmatrix}^T, \quad q_2 = \frac{\tilde{q}_2}{\|\tilde{q}_2\|} = \begin{bmatrix} 0 & 0 & -1 \end{bmatrix}^T$$

$$\tilde{q}_3^{(1)} = a_3 - (q_1^T a_3)q_1 = \begin{bmatrix} 0 & -\epsilon & 0 \end{bmatrix}$$

$$\tilde{q}_3^{(2)} = \tilde{q}_3^{(1)} - (q_2^T \tilde{q}_3^{(1)})q_2 = \begin{bmatrix} 0 & -\epsilon & 0 \end{bmatrix}$$

$$\tilde{q}_3 = \tilde{q}_3^{(2)} = \begin{bmatrix} 0 & -\epsilon & 0 \end{bmatrix}^T, \quad q_3 = \frac{\tilde{q}_3}{\|\tilde{q}_3\|} = \begin{bmatrix} 0 & -1 & 0 \end{bmatrix}^T$$

The two sets of basis are different. The result of classical Gram-Schmidt, q_2 and q_3 are not orthogonal. However, the set of basis derived by modified Gram-Schmidt algorithm is the desired orthogonal basis.

Algorithm 3: Modified Gram-Schmidt algorithm

Input : A collection of linearly independent vectors $\mathbf{a}_1, \dots, \mathbf{a}_n$.

```

1 Initialization:  $\tilde{\mathbf{q}}_1 = \mathbf{a}_1, \mathbf{q}_1 = \tilde{\mathbf{q}}_1 / \|\tilde{\mathbf{q}}_1\|_2$ 
2 for  $i = 2, \dots, n$  do
3    $m = \mathbf{a}_i$ 
4) 4   for  $j = 1, \dots, i - 1$  do
5      $m = m - q_j^T m q_j$ 
6   end
7    $q_i = m / \|m\|_2$ 
8 end

```

Output: $\mathbf{q}_1, \dots, \mathbf{q}_n$

IV. SOLVING LS VIA QR FACTORIZATION AND NORMAL EQUATION

Problem 4 [Understanding the influence of the condition number to the solution.]. (4 points + 5 points + 4 points + 4 points + 3 points points)

Consider such two LS problems:

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{Ax} - \mathbf{b}\|_2^2 \\ \min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{Ax} - (\mathbf{b} + \delta\mathbf{b})\|_2^2 \end{aligned} \quad (1)$$

with $\mathbf{A} \in \mathbb{R}^{m \times n}$. For $\mathbf{b} = \begin{bmatrix} 1 & 3/2 & 3 & 6 \end{bmatrix}^T$ and $\delta\mathbf{b} = \begin{bmatrix} 1/10 & 0 & 0 & 0 \end{bmatrix}^T$,

1) Computing solution to the problem (??) via QR decomposition when

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 5 \\ 3 & 4 & 7 \\ 4 & 5 & 11 \end{bmatrix}.$$

2) For a full-rank matrix \mathbf{A} , consider the equation $\mathbf{Ax} = \mathbf{b}$, after adding some noise $\delta\mathbf{b}$ to \mathbf{b} , we have $\mathbf{A}(\mathbf{x} + \delta\mathbf{x}) = \mathbf{b} + \delta\mathbf{b}$, and then proof

$$\frac{1}{\|\mathbf{A}\| \|\mathbf{A}^\dagger\|} \frac{\|\delta\mathbf{b}\|}{\|\mathbf{b}\|} \leq \frac{\|\delta\mathbf{x}\|}{\|\mathbf{x}\|} \leq \|\mathbf{A}\| \|\mathbf{A}^\dagger\| \frac{\|\delta\mathbf{b}\|}{\|\mathbf{b}\|},$$

and give it a plain interpretation.

3) Computing the solutions to the two LS problems via the normal equation $\mathbf{A}^T \mathbf{Ax}_{LS} = \mathbf{A}^T \mathbf{b}$ when

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \\ 1 & 1 & 0 \end{bmatrix}.$$

4) Computing the solutions to the two LS problems via the normal equation $\mathbf{A}^T \mathbf{Ax}_{LS} = \mathbf{A}^T \mathbf{b}$ when

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \end{bmatrix}.$$

5) Compare the 2-norm condition number $\|\mathbf{A}\| \|\mathbf{A}^\dagger\|$ for \mathbf{A} in 3) and 4) and the influence on the solution to problem (??) resulted by the additional noise $\delta\mathbf{b}$.

Hint: Show the influence on the solution by $\frac{\|\delta\mathbf{x}\|}{\|\mathbf{x}\|}$.

Remarks: You can use MATLAB for some matrix computations (deviation is expected) in 3), 4), 5). Do not use decimals in your answers, fraction and n -th roots of numbers are accepted.

Solution.

1) Denote $\mathbf{A} = \begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix}$, we have

$$\begin{aligned}
 q_1 &= \frac{a_1}{\|a_1\|} = \frac{\sqrt{30}}{30} \begin{bmatrix} 1 & 2 & 3 & 4 \end{bmatrix}^T \\
 \tilde{q}_2 &= a_2 - (q_1^T a_2) q_1 = \frac{1}{3} \begin{bmatrix} 2 & 1 & 0 & -1 \end{bmatrix}^T \quad q_2 = \frac{\tilde{q}_2}{\|\tilde{q}_2\|} = \frac{\sqrt{6}}{6} \begin{bmatrix} 2 & 1 & 0 & -1 \end{bmatrix}^T \\
 \tilde{q}_3 &= a_3 - (q_1^T a_3) q_1 - (q_2^T a_3) q_2 = \frac{1}{5} \begin{bmatrix} 2 & -1 & -4 & 3 \end{bmatrix}^T \quad q_3 = \frac{\tilde{q}_3}{\|\tilde{q}_3\|} = \frac{\sqrt{30}}{30} \begin{bmatrix} 2 & -1 & -4 & 3 \end{bmatrix}^T \\
 r_{11} &= \|\tilde{a}_1\| = \frac{\sqrt{30}}{30} \quad r_{22} = \|\tilde{q}_2\| = \frac{\sqrt{6}}{3} \quad r_{33} = \|\tilde{q}_3\| = \frac{\sqrt{30}}{5} \\
 r_{12} &= q_1^T a_2 = \sqrt{30} \quad r_{13} = q_1^T a_3 = \frac{13\sqrt{30}}{5} \quad r_{23} = q_2^T a_3 = 0
 \end{aligned}$$

$$\mathbf{Q} = \begin{bmatrix} \frac{\sqrt{30}}{30} & \frac{2\sqrt{6}}{6} & \frac{2\sqrt{30}}{30} \\ \frac{2\sqrt{30}}{30} & \frac{\sqrt{6}}{6} & -\frac{\sqrt{30}}{30} \\ \frac{3\sqrt{30}}{30} & 0 & -\frac{4\sqrt{30}}{30} \\ \frac{4\sqrt{30}}{30} & -\frac{\sqrt{6}}{6} & \frac{3\sqrt{30}}{30} \end{bmatrix} \quad \mathbf{R} = \begin{bmatrix} \frac{\sqrt{30}}{30} & \frac{4\sqrt{30}}{3} & \frac{13\sqrt{30}}{5} \\ 0 & \frac{\sqrt{6}}{3} & 0 \\ 0 & 0 & \frac{\sqrt{30}}{5} \end{bmatrix}$$

$$\mathbf{A}\mathbf{x} = \mathbf{b} \rightarrow \mathbf{Q}\mathbf{R}\mathbf{x} = \mathbf{b} \rightarrow \mathbf{R}\mathbf{x} = \mathbf{Q}^T \mathbf{b}$$

solve the equation, we have $\mathbf{x}_{LS} = \begin{bmatrix} \frac{1}{12} \\ -\frac{5}{4} \\ \frac{13}{12} \end{bmatrix}$

2) Since $x + \delta x = \mathbf{A}^\dagger(b + \delta b)$ and $x = \mathbf{A}^\dagger b$, we have $\delta x = \mathbf{A}^\dagger \delta b$, so $\|\delta x\| = \|\mathbf{A}^\dagger \delta b\| \leq \|\mathbf{A}^\dagger\| \|\delta b\|$.

Since $\mathbf{A}x = b$, we have $\frac{1}{\|x\|} \leq \frac{\|b\|}{\|x\|}$, so $\frac{\|\delta x\|}{\|x\|} \leq \|\mathbf{A}\| \|\mathbf{A}^\dagger\| \frac{\|\delta b\|}{\|b\|}$.

Similarly, $\mathbf{A}(x + \delta x) = b + \delta b$, we have $\delta b = \mathbf{A} \delta x$, so $\|\delta b\| \leq \|\mathbf{A}\| \|\delta x\|$. $x = \mathbf{A}^\dagger b$, we have

$$\|x\| \leq \|\mathbf{A}^\dagger\| \|b\|, \text{ so } \frac{1}{\|\mathbf{A}\| \|\mathbf{A}^\dagger\|} \frac{\|\delta b\|}{\|b\|} \leq \frac{\|\delta x\|}{\|x\|}$$

Above all,

$$\frac{1}{\|\mathbf{A}\| \|\mathbf{A}^\dagger\|} \frac{\|\delta b\|}{\|b\|} \leq \frac{\|\delta x\|}{\|x\|} \leq \|\mathbf{A}\| \|\mathbf{A}^\dagger\| \frac{\|\delta b\|}{\|b\|},$$

interpretation:

When we think about the relative error of x , If b changes, the sensitivity of the least square solution depends on $\|\mathbf{A}\| \|\mathbf{A}^\dagger\|$. And if $\|\mathbf{A}\| \|\mathbf{A}^\dagger\|$ is large, though b changes little, the error of x may change large.

3)

$$\mathbf{A}^T \mathbf{A} = \begin{bmatrix} 15 & 16 & 15 \\ 16 & 18 & 17 \\ 15 & 17 & 17 \end{bmatrix} \quad \mathbf{A}^T \mathbf{b} = \begin{bmatrix} 19 \\ 20 \\ 14 \end{bmatrix} \quad \mathbf{A}^T (\mathbf{b} + \delta \mathbf{b}) = \begin{bmatrix} \frac{191}{10} \\ \frac{101}{5} \\ \frac{71}{5} \end{bmatrix}$$

Solve the equation $\mathbf{A}^T \mathbf{A} \mathbf{x}_{LS} = \mathbf{A}^T \mathbf{b}$ and $\mathbf{A}^T \mathbf{A} \mathbf{x}_{LS2} = \mathbf{A}^T (\mathbf{b} + \delta \mathbf{b})$, we have

$$\mathbf{x}_{LS} = \begin{bmatrix} \frac{11}{13} \\ \frac{67}{13} \\ -\frac{66}{13} \end{bmatrix} \quad \mathbf{x}_{LS2} = \begin{bmatrix} \frac{97}{130} \\ \frac{683}{130} \\ -\frac{66}{13} \end{bmatrix}$$

4)

$$\mathbf{A}^T \mathbf{A} = \begin{bmatrix} 4 & 10 & 30 \\ 10 & 30 & 100 \\ 30 & 100 & 357 \end{bmatrix} \quad \mathbf{A}^T \mathbf{b} = \begin{bmatrix} \frac{23}{2} \\ 37 \\ 130 \end{bmatrix} \quad \mathbf{A}^T (\mathbf{b} + \delta \mathbf{b}) = \begin{bmatrix} \frac{58}{5} \\ \frac{371}{10} \\ \frac{1301}{10} \end{bmatrix}$$

Solve the equation $\mathbf{A}^T \mathbf{A} \mathbf{x}_{LS} = \mathbf{A}^T \mathbf{b}$ and $\mathbf{A}^T \mathbf{A} \mathbf{x}_{LS2} = \mathbf{A}^T (\mathbf{b} + \delta \mathbf{b})$, we have

$$\mathbf{x}_{LS} = \begin{bmatrix} \frac{15}{8} \\ -\frac{59}{40} \\ \frac{5}{8} \end{bmatrix} \quad \mathbf{x}_{LS2} = \begin{bmatrix} \frac{21}{10} \\ -\frac{163}{100} \\ \frac{13}{20} \end{bmatrix}$$

5)

$$\mathbf{A}^\dagger = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$$

$$\text{cond}(\mathbf{A}_1) = \frac{5610}{421} \quad \text{cond}(\mathbf{A}_2) = \frac{2653}{36}$$

For \mathbf{A}_1 , $\frac{421}{5610} \frac{\|\delta \mathbf{b}\|}{\|\mathbf{b}\|} \leq \frac{\|\delta \mathbf{x}\|}{\|\mathbf{x}\|} \leq \frac{5610}{421} \frac{\|\delta \mathbf{b}\|}{\|\mathbf{b}\|}$. For \mathbf{A}_2 , $\frac{36}{2653} \frac{\|\delta \mathbf{b}\|}{\|\mathbf{b}\|} \leq \frac{\|\delta \mathbf{x}\|}{\|\mathbf{x}\|} \leq \frac{2653}{36} \frac{\|\delta \mathbf{b}\|}{\|\mathbf{b}\|}$.

So the solution of \mathbf{A}_2 is more sensitive when \mathbf{b} changes.

V. UNDERDETERMINED SYSTEM

Problem 5 [Solving Underdetermined System by QR]. (10 points + 5 points)

Consider the following underdetermined system $\mathbf{Ax} = \mathbf{b}$ with $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $m < n$. Let

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 2 & 0 \\ 0 & -2 & 2 & 1 \\ 2 & 5 & 6 & 1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix},$$

- 1) Use Householder reflection to give the full QR decomposition of tall \mathbf{A}^T , i.e., $\mathbf{A}^T = \mathbf{QR}$ with \mathbf{Q} being a square matrix with orthonormal columns.
- 2) Give one possible solution via QR decomposition of \mathbf{A}^T , write down your solution using \mathbf{b} .

Solution.

1)

$$\mathbf{A}^T = \begin{bmatrix} 1 & 0 & 2 \\ 2 & -2 & 5 \\ 2 & 2 & 6 \\ 0 & 1 & 1 \end{bmatrix}$$

$$x_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \\ 0 \end{bmatrix} \quad \mathbf{v}_1 = \mathbf{x}_1 - \|x_1\| \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \\ 2 \\ 0 \end{bmatrix} \quad \mathbf{H}_1 = \mathbf{I} - 2 \frac{v_1 v_1^T}{v_1^T v_1} = \frac{1}{3} \begin{bmatrix} 1 & 2 & 2 & 0 \\ 2 & 1 & -2 & 0 \\ 2 & -2 & 1 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

$$\mathbf{H}_1 \mathbf{A}^T = \begin{bmatrix} 3 & 0 & 8 \\ 0 & -2 & -1 \\ 0 & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

$$x_2 = \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix} \quad \mathbf{v}_2 = \mathbf{x}_2 + \|x_2\| \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \quad \tilde{\mathbf{H}}_2 = \mathbf{I} - 2 \frac{v_2 v_2^T}{v_2^T v_2} = \frac{1}{3} \begin{bmatrix} 2 & -2 & -1 \\ -2 & -1 & -2 \\ -1 & -2 & 2 \end{bmatrix}$$

$$\mathbf{H}_2 = \frac{1}{3} \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & -2 & -1 \\ 0 & -2 & -1 & -2 \\ 0 & -1 & -2 & 2 \end{bmatrix} \quad \mathbf{H}_2 \mathbf{H}_1 \mathbf{A}^T = \begin{bmatrix} 3 & 0 & 8 \\ 0 & -3 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{H}_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad \mathbf{H}_3 \mathbf{H}_2 \mathbf{H}_1 \mathbf{A}^T = \begin{bmatrix} 3 & 0 & 8 \\ 0 & -3 & -1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\mathbf{Q} = \mathbf{H}_3 \mathbf{H}_2 \mathbf{H}_1^T = \frac{1}{3} \begin{bmatrix} 1 & 0 & -2 & -2 \\ 2 & 2 & 1 & 0 \\ 2 & -2 & 0 & 1 \\ 0 & -1 & 2 & -2 \end{bmatrix} \quad \mathbf{R} = \mathbf{H}_3 \mathbf{H}_2 \mathbf{H}_1 \mathbf{A}^T = \begin{bmatrix} 3 & 0 & 8 \\ 0 & -3 & -1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

2) Denote

$$\mathbf{Q}_1 = \mathbf{Q}(:, 1:3), \mathbf{Q}_2 = \mathbf{Q}(:, 4) \quad \mathbf{R}_1 = \mathbf{R}(1:3, :)$$

$\mathbf{A}^T = \mathbf{Q}\mathbf{R} = \mathbf{Q}_1\mathbf{R}_1 + \mathbf{Q}_2 0$, note $\mathbf{A}\mathbf{x} = \mathbf{R}_1^T \mathbf{Q}_1^T + 0^T \mathbf{Q}_2^T x = \mathbf{b}$, which indicates $\mathbf{Q}_1^T \mathbf{x} = (\mathbf{R}_1^T)^{-1} \mathbf{b}$ and $\mathbf{Q}_2^T x$ can be anything.

$$\text{so } x = \mathbf{Q} \begin{bmatrix} (\mathbf{R}_1^T)^{-1} \mathbf{b} \\ 0 \end{bmatrix} = \frac{1}{9} \begin{bmatrix} 17 & 2 & -6 \\ -6 & -3 & 3 \\ -16 & -1 & 6 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \text{ is one possible solution.}$$

VI. SOLVING LS VIA PROJECTION

Problem 6. (Bonus question, 6 points + 4 points)

Consider the Least Square (LS) problem:

$$\min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 \quad (2)$$

where $\mathbf{A} \in \mathbb{R}^{m \times n}$ ($m > n$) may not be full rank. Denote

$$X_{\text{LS}} = \{\mathbf{x} \in \mathbb{R}^n | \mathbf{A}^T \mathbf{A} \mathbf{x} = \mathbf{A}^T \mathbf{y}\}$$

as the set of all solutions to (??), and

$$\mathbf{x}_{\text{LS}} = \mathbf{A}^\dagger \mathbf{y}$$

where $\mathbf{A}^\dagger \in \mathbb{R}^{n \times m}$ is the *pseudo inverse* of \mathbf{A} satisfies the following properties:

- 1) $\mathbf{A} \mathbf{A}^\dagger \mathbf{A} = \mathbf{A}$.
- 2) $\mathbf{A}^\dagger \mathbf{A} \mathbf{A}^\dagger = \mathbf{A}^\dagger$.
- 3) $(\mathbf{A} \mathbf{A}^\dagger)^T = \mathbf{A} \mathbf{A}^\dagger$.
- 4) $(\mathbf{A}^\dagger \mathbf{A})^T = \mathbf{A}^\dagger \mathbf{A}$.

Answer the following questions:

- 1) Prove that \mathbf{x}_{LS} is a solution to (??) and is of minimum 2-norm in X_{LS} , that is

$$\mathbf{x}_{\text{LS}} = \arg \min_{\mathbf{x} \in X_{\text{LS}}} \|\mathbf{x}\|_2 .$$

Hint. Notice that the orthogonal projection onto $\mathcal{N}(A)$ is given by

$$\mathbf{\Pi}_{\mathcal{N}(A)} = \mathbf{I} - \mathbf{A}^\dagger \mathbf{A}$$

- 2) Prove that $X_{\text{LS}} = \{\mathbf{x}_{\text{LS}}\}$ if and only if $\text{rank}(\mathbf{A}) = n$.

Solution.

1)

$$\mathbf{A}^T \mathbf{A} \mathbf{x}_{\text{LS}} \stackrel{\mathbf{x}_{\text{LS}} = \mathbf{A}^\dagger \mathbf{y}}{=} \mathbf{A}^T \mathbf{A} \mathbf{A}^\dagger \mathbf{y} \stackrel{(\mathbf{A} \mathbf{A}^\dagger)^T = \mathbf{A} \mathbf{A}^\dagger}{=} \mathbf{A}^T (\mathbf{A} \mathbf{A}^\dagger)^T \mathbf{y} = \mathbf{A}^T (\mathbf{A}^\dagger)^T \mathbf{A}^T \mathbf{y} = (\mathbf{A} \mathbf{A}^\dagger \mathbf{A})^T \mathbf{y} \stackrel{\mathbf{A} \mathbf{A}^\dagger \mathbf{A} = \mathbf{A}}{=} \mathbf{A}^T \mathbf{y}$$

We have $\mathbf{x}_{\text{LS}} \in \mathbf{X}_{\text{LS}}$, so \mathbf{x}_{LS} is a solution to the LS problem.

It's analogous to a system of inhomogeneous linear equations, denote $x_{\text{Null}} = \{\mathbf{x} \in \mathbb{R}^n | \mathbf{A}^T \mathbf{A} \mathbf{x} = 0\}$ so $x_{\text{Null}} \in \mathcal{N}(\mathbf{A})$ and any $x \in \mathbf{X}_{\text{LS}}$ can be represented by the addition of \mathbf{x}_{LS} and \mathbf{x}_{Null} .

The 2-norm in \mathbf{X}_{LS} is equal to $\|\mathbf{x}_{\text{Null}} - (-\mathbf{x}_{\text{LS}})\|_2$. when \mathbf{x}_{Null} is the vector which is \mathbf{x}_{LS} projects onto the subspace $\mathcal{N}(\mathbf{A})$ the 2-norm in \mathbf{X}_{LS} is minimum. Since the orthogonal projection onto $\mathcal{N}(\mathbf{A})$ is given by $\Pi_{\mathcal{N}(\mathbf{A})} = \mathbf{I} - \mathbf{A}^\dagger \mathbf{A}$,

$$\mathbf{x}_{\text{Null}} = (\mathbf{I} - \mathbf{A}^\dagger \mathbf{A}) \mathbf{x}_{\text{LS}} = \mathbf{x}_{\text{LS}} - \mathbf{A}^\dagger \mathbf{A} \mathbf{x}_{\text{LS}} = \mathbf{x}_{\text{LS}} - \mathbf{A}^\dagger \mathbf{A} \mathbf{A}^\dagger \mathbf{y} = \mathbf{x}_{\text{LS}} - \mathbf{A}^\dagger \mathbf{y} = 0$$

So $\mathbf{x}_{\text{LS}} = \arg \min_{\mathbf{x} \in \mathbf{X}_{\text{LS}}} \|\mathbf{x}\|_2$.

2) If $\mathbf{X}_{\text{LS}} = \{\mathbf{x}_{\text{LS}}\}$, the solution of equation $\mathbf{A}^T \mathbf{A} \mathbf{x} = \mathbf{A}^T \mathbf{y}$ is unique. So matrix $\mathbf{A}^T \mathbf{A}$ is column full rank.

Since $\mathcal{R}(\mathbf{A}) = \mathcal{R}(\mathbf{A}^T \mathbf{A})$, we have $\text{rank}(\mathbf{A}) = n$.

If $\text{rank}(\mathbf{A}) = n$, $\mathbf{A}^T \mathbf{A}$ is invertible, the solution of equation $\mathbf{A}^T \mathbf{A} \mathbf{x} = \mathbf{A}^T \mathbf{y}$ just is $(\mathbf{A}^T \mathbf{A}^{-1} \mathbf{A}^T)$, so the solution is unique, it means that $\mathbf{X}_{\text{LS}} = \{\mathbf{x}_{\text{LS}}\}$.

Above all, $\mathbf{X}_{\text{LS}} = \{\mathbf{x}_{\text{LS}}\}$ if and only if $\text{rank}(\mathbf{A}) = n$.