SI231 - Matrix Computations, Fall 2020-21
Homework Set #3

Prof. Yue Qiu and Prof. Ziping Zhao

Name: Tao Huang **Major:** Undergraduate in CS

Student No.: 2018533172 **E-mail:** huangtao1@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 the and the programming part.

3) About the 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 Probelm 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) If P is not an identity matrix, we assume it is non-singular in such a condition. By the property of non-singular matrix,
$$Px = 0 \Leftrightarrow x = 0.$$

Also, by the definition of orthogonal complement,
$$\forall y \in \mathcal{U}^\perp, Py = 0$$

Thus, the above results imply that
$$\forall y \in \mathcal{U}^\perp \Leftrightarrow y = 0.$$

This means that $\dim(\mathcal{U}^\perp) = 0$ which further implies $\dim(\mathcal{U}) = n$. Based on that, This orthogonal projector projects $\forall x \in R^n$ into $\mathcal{U} = R^n$, that is projecting x to itself. Therefore, $Px = x, \forall x$ gives P is an identity matrix, which contradicts to the fact that P is not an identity matrix. So the assumption is unreasonable, an orthogonal projector must be singular if it is not an identity matrix Q.E.D.

2) The orthogonal projector onto $\mathcal{U}^\perp$ is $I - P$. Next, we give a detailed proof of this conclusion. We denote the orthogonal projector onto $\mathcal{U}^\perp$ as $P^\perp$. By the theorem that $\forall x \in R^n$ can be decomposed into $\mathcal{U}$ and $\mathcal{U}^\perp$,
$$x = Px + P^\perp x$$

This implies that
$$P + P^\perp = I \Leftrightarrow P^\perp = I - P$$

3) • We show $P_{\mathcal{U}}P_{\mathcal{W}} = 0 \Rightarrow \mathcal{U} \perp \mathcal{W}$.

For $\forall x \in \mathcal{U}, \exists x' \in R^n$ s.t. $P_{\mathcal{U}}x' = x$. For $\forall y \in \mathcal{W}, \exists y' \in R^n$ s.t. $P_{\mathcal{W}}y' = y$. Then
$$\begin{aligned}
<x, y> &= (P_{\mathcal{U}}x')^T P_{\mathcal{W}}y' \\
&= x'^T P_{\mathcal{U}}^T P_{\mathcal{W}}y' \\
&= x'^T P_{\mathcal{U}} P_{\mathcal{W}}y' \\
&= 0
\end{aligned}$$

So $\forall x \in \mathcal{U}, y \in \mathcal{W}, <x, y> = 0$, which implies $\mathcal{U} \perp \mathcal{W}$. Q.E.D.

- We show $P_{\mathcal{U}}P_{\mathcal{W}} = 0 \Leftarrow \mathcal{U} \perp \mathcal{W}$.

  From the conclusion in 2),

$$\mathcal{U} \perp \mathcal{W} \Rightarrow P_{\mathcal{W}} = I - P_{\mathcal{U}}$$

  . Therefore

$$\begin{aligned}
P_{\mathcal{U}}P_{\mathcal{W}} &= P_{\mathcal{U}}(I - P_{\mathcal{U}}) \\
&= P_{\mathcal{U}} - P_{\mathcal{U}}^2 \\
&= P_{\mathcal{U}} - P_{\mathcal{U}} \\
&= 0
\end{aligned}$$

  Q.E.D.

So $\mathbf{P}_{\mathcal{U}}\mathbf{P}_{\mathcal{W}} = 0$ 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 $\mathbf{x} = \mathbf{x_{LS}}$ 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.

3) a) **–** Gradient descent:

   * Compute $A^T A$: $n^2 * (m + m - 1) = (2m - 1)n^2 = 88155900$ flops.

* Compute $A^T y$: $n * (m + m - 1) = (2m - 1)n = 419790$ flops.

  These two results could be cached for subsequent use in gradient descent, so they only need to be computed once.

* Complexity of each iteration: $n * (n + n - 1) + n + n + n = 2n^2 + 2n = 88620$ flops.

* The terminal condition was set as: norm(gradient) $< 1e - 5$, which lasted for 103 iteration. The total flops in my implementation is:

$$(2m - 1)n^2 + (2m - 1)n + 103 * (2n^2 + 2n) = 97703550 \text{ flops.}$$

– Cholesky decomposition: In this method, we first need to compute $A^T A$ to get a PSD matrix, and compute $A^T b$ to reconstruct a corresponding label.

* Compute $A^T A$: $A^T A$: $n^2 * (m + m - 1) = (2m - 1)n^2 = 88155900$ flops.

* Compute $A^T y$: $n * (m + m - 1) = (2m - 1)n = 419790$ flops.

* Implement Cholesky decomposition:

$$\sum_{j=1}^{n}(1 + 1 + 2(j - 1)) + \sum_{j=1}^{n}[(n - j) + 2(n - j)(j - 1) + (n - j)]$$

$$= \sum_{j=1}^{n}\left(2j + n - j + 2nj - 2n - 2j^2 + 2j + n - j\right)$$

$$= 2\sum_{j=1}^{n}(j + nj - j^2)$$

$$= n(n + 1) + n^2(n + 1) - \frac{n(n + 1)(2n + 1)}{3}$$

$$= \frac{n^3}{3} + n^2 + \frac{2n}{3}$$

$$= 3131240 \text{ flops.}$$

* Implement forward and backward substitution:

$$(Forward) \ \sum_{i=1}^{n}[(i - 1 + i - 2) + 1 + 1] + (Backward) \ \sum_{i=1}^{n}[(n - i + n - i - 1) + 1 + 1]$$

$$= \sum_{i=1}^{n}(2i - 1 + 2n - 2i + 1)$$

$$= 2\sum_{i=1}^{n}$$

$$= n^2 + n$$

$$= 44310 \text{ flops.}$$

* The total flops in this method is:

$$(2m - 1)n^2 + (2m - 1)n + \frac{n^3}{3} + n^2 + \frac{2n}{3} + n^2 + n = 91751240 \text{ flops.}$$

b) – Gradient 2-norm:

* Gradient descent: $3.177181240593741e - 06$.

* Cholesky decomposition: $1.572877035079361e - 09$.

– Loss:

* Gradient descent: $(1.632256538153266e + 02)^2 = 2.664261406344086e + 04$.

* Cholesky decompositon: $(1.632256538153268e + 02)^2 = 2.664261406344091e + 04$

– To sum up, these two methods perform nearly same in view of time complexity and final loss.

III. UNDERSTANDING THE QR FACTORIZATION

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

1) Consider the subspace $\mathcal{S}$ spaned by $\{\mathbf{a}_1, \ldots, \mathbf{a}_4\}$, 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-Schimidt algorithm (See Algorithm 1), 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 $\mathbf{a}_1, \ldots, \mathbf{a}_n$.

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

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

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

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

5 **end**

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

---

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

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

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 $\mathcal{S}$ spaned 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 2. Then use the **classical** Gram-Schimidt algorithm and the **modified** Gram-Schimidt 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 $\mathbf{a}_1, \ldots, \mathbf{a}_n$.

1 **Initilization:** $\tilde{q}_i = a_i, \forall i$

2 **for** $i = 1, \ldots, n$ **do**

3 $\quad$ $q_i = \tilde{q}_i / \|\tilde{q}_i^2\|$

4 $\quad$ **for** $j = i + 1, ..., n$ **do**

5 $\quad\quad$ $\tilde{q}_j = \tilde{q}_j - (q_i^T \tilde{q}_j)q_i$

6 $\quad$ **end**

7 **end**

8 *Complete your algorithm here...*

$\quad$ **Output:** $\mathbf{q}_1, \ldots, \mathbf{q}_n$

---

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

$$q_1 = \frac{a_1}{\|a_1\|} = \begin{bmatrix} \frac{1}{\sqrt{30}} \\ \frac{2}{\sqrt{30}} \\ \frac{3}{\sqrt{30}} \\ \frac{4}{\sqrt{30}} \end{bmatrix}$$

$$\tilde{q}_2 = a_2 - (q_1^T a_2)q_1$$

$$= \begin{bmatrix} 2 \\ 3 \\ 4 \\ 5 \end{bmatrix} - \begin{bmatrix} \frac{1}{\sqrt{30}} \\ \frac{2}{\sqrt{30}} \\ \frac{3}{\sqrt{30}} \\ \frac{4}{\sqrt{30}} \end{bmatrix}^T \begin{bmatrix} 2 \\ 3 \\ 4 \\ 5 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{30}} \\ \frac{2}{\sqrt{30}} \\ \frac{3}{\sqrt{30}} \\ \frac{4}{\sqrt{30}} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \\ 0 \\ -\frac{1}{3} \end{bmatrix}$$

$$q_2 = \frac{\tilde{q}_2}{||\tilde{q}_2||} = \begin{bmatrix} \frac{\sqrt{6}}{3} \\ \frac{\sqrt{6}}{6} \\ 0 \\ -\frac{\sqrt{6}}{6} \end{bmatrix}$$

$$\tilde{q}_3 = a_3 - (q_1^T a_3)q_1 - (q_2^T a_3)q_2 = 0$$

$$q_3 = 0$$

$$\tilde{q}_4 = a_4 - (q_1^T a_4)q_1 - (q_2^T a_4)q_2 - (q_3^T a_4)q_3$$

$$= \begin{bmatrix} 3 \\ 5 \\ 7 \\ 11 \end{bmatrix} - \begin{bmatrix} \frac{1}{\sqrt{30}} \\ \frac{2}{\sqrt{30}} \\ \frac{3}{\sqrt{30}} \\ \frac{4}{\sqrt{30}} \end{bmatrix}^T \begin{bmatrix} 3 \\ 5 \\ 7 \\ 11 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{30}} \\ \frac{2}{\sqrt{30}} \\ \frac{3}{\sqrt{30}} \\ \frac{4}{\sqrt{30}} \end{bmatrix} - \begin{bmatrix} \frac{\sqrt{6}}{3} \\ \frac{\sqrt{6}}{6} \\ 0 \\ -\frac{\sqrt{6}}{6} \end{bmatrix}^T \begin{bmatrix} 3 \\ 5 \\ 7 \\ 11 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{6}}{3} \\ \frac{\sqrt{6}}{6} \\ 0 \\ -\frac{\sqrt{6}}{6} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{2}{5} \\ -\frac{1}{5} \\ -\frac{4}{5} \\ \frac{3}{5} \end{bmatrix}$$

$$q_4 = \frac{\tilde{q}_4}{||\tilde{q}_4||} = \begin{bmatrix} \frac{\sqrt{30}}{15} \\ -\frac{\sqrt{30}}{30} \\ -\frac{2\sqrt{30}}{15} \\ \frac{\sqrt{30}}{10} \end{bmatrix}$$

So we have the basis:

$$q_1 = \begin{bmatrix} \frac{1}{\sqrt{30}} \\ \frac{2}{\sqrt{30}} \\ \frac{3}{\sqrt{30}} \\ \frac{4}{\sqrt{30}} \end{bmatrix}, q_2 = \begin{bmatrix} \frac{\sqrt{6}}{3} \\ \frac{\sqrt{6}}{6} \\ 0 \\ -\frac{\sqrt{6}}{6} \end{bmatrix}, q_4 = \begin{bmatrix} \frac{\sqrt{30}}{15} \\ -\frac{\sqrt{30}}{30} \\ -\frac{2\sqrt{30}}{15} \\ \frac{\sqrt{30}}{10} \end{bmatrix}$$

By computing $q_1^T q_2 = 0, q_1^T q_3 = 0, q_2^T q_3 = 0$, we verify the orthonormality of the found basis.

2)
- **Perspective from projection of vector:** The key idea under this view is that we find a vector orthogonal to the base vectors we have already found at each time. Specifically, for a_i, we project it to $\{q_1, q_2, ..., q_{i-1}\}$. $\tilde{q}_i = a_i - \sum_{j=1}^{i-1}(q_j^T a_i)q_j$ is the desired vector that is orthogonal to $\{q_1, q_2, ..., q_{i-1}\}$. Then the normalized vector $q_i = \frac{\tilde{q}_i}{||\tilde{q}_i||}$ is added to the bases that we have already found.
- **Perspective from projection onto subspace:** The key idea under this view is that we find a vector orthogonal to the subspace span by the base vectors we have already found at each time. Specifically, for a_i, we project it to the subspace $span\{q_1, q_2, ..., q_{i-1}\}$, which is equivalent to the $Px = UU^T x, U = [q_1|...|q_{i-1}]$. Then $\tilde{p}_i = a_i - Px = a_i - UU^T x = a_i - \sum_{j=1}^{i-1}(q_j^T a_i)q_j$ is the desired vector orthogonal to $\{q_1, q_2, ..., q_{i-1}\}$. Then the normalized vector $q_i = \frac{\tilde{q}_i}{||\tilde{q}_i||}$ is added to the bases that we have already found.

3)
- Classical Gram-Schmidt

$$\tilde{q}_1 = a_1, \quad q_1 = \frac{1}{\sqrt{1 + 2\epsilon^2}}q_1 = \begin{bmatrix} 1 \\ \epsilon \\ \epsilon \end{bmatrix}$$

$$\tilde{q}_2 = a_2 - (q_1^T a_2)q_1$$

$$= \begin{bmatrix} 1 \\ \epsilon \\ 0 \end{bmatrix} - \begin{bmatrix} 1 \\ \epsilon \\ \epsilon \end{bmatrix}^T \begin{bmatrix} 1 \\ \epsilon \\ 0 \end{bmatrix} \begin{bmatrix} 1 \\ \epsilon \\ \epsilon \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ 0 \\ -\epsilon \end{bmatrix}$$

$$q_2 = \frac{\tilde{q}_2}{||\tilde{q}_2||_2} = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}$$

$$\tilde{q}_3 = a_3 - (q_1^T a_3)q_1 - (q_2^T a_3)q_2$$

$$= \begin{bmatrix} 1 \\ 0 \\ \epsilon \end{bmatrix} - \begin{bmatrix} 1 \\ \epsilon \\ \epsilon \end{bmatrix}^T \begin{bmatrix} 1 \\ 0 \\ \epsilon \end{bmatrix} \begin{bmatrix} 1 \\ \epsilon \\ \epsilon \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}^T \begin{bmatrix} 1 \\ 0 \\ \epsilon \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ -\epsilon \\ -\epsilon \end{bmatrix}$$

$$q_3 = \frac{\tilde{q}_3}{||\tilde{q}_3||_2} = \begin{bmatrix} 0 \\ -\frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} \end{bmatrix}$$

So we have:

$$q_1 = \begin{bmatrix} 1 \\ \epsilon \\ \epsilon \end{bmatrix}, \quad q_2 = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}, \quad q_3 = \begin{bmatrix} 0 \\ -\frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} \end{bmatrix}$$

- Modified Gram-Schimdt

$$\tilde{q}_1 = a_1 = \begin{bmatrix} 1 \\ \epsilon \\ \epsilon \end{bmatrix}, \quad \tilde{q}_2 = a_2 = \begin{bmatrix} 1 \\ \epsilon \\ 0 \end{bmatrix}, \quad \tilde{q}_3 = a_3 = \begin{bmatrix} 1 \\ 0 \\ \epsilon \end{bmatrix}$$

a)

$$q_1 = \frac{\tilde{q}_1}{||\tilde{q}_1||_2} = \begin{bmatrix} 1 \\ \epsilon \\ \epsilon \end{bmatrix}$$

$$\tilde{q}_2 = \tilde{q}_2 - q_1^T \tilde{q}_2 q_1 = \begin{bmatrix} 0 \\ 0 \\ -\epsilon \end{bmatrix}$$

$$\tilde{q}_3 = \tilde{q}_3 - q_1^T \tilde{q}_3 q_1 = \begin{bmatrix} 0 \\ -\epsilon \\ 0 \end{bmatrix}$$

b)

$$q_2 = \frac{\tilde{q}_2}{||\tilde{q}_2||_2} = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}$$

$$\tilde{q}_3 = \tilde{q}_3 - q_2^T \tilde{q}_3 q_2 = \begin{bmatrix} 0 \\ -\epsilon \\ 0 \end{bmatrix}$$

c)

$$q_3 = \frac{\tilde{q}_3}{||\tilde{q}_3||_3} = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}$$

So we have:

$$q_1 = \begin{bmatrix} 1 \\ \epsilon \\ \epsilon \end{bmatrix}, \quad q_2 = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}, \quad q_3 = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}$$

We see that the found basis are not same. The basis produced by modified GS is the desired one since such an algorithm is more stable than classic GS when meeting computer rounding errors. This can be verified from the rusults, *i.e.* $q_2^T q_3 = \frac{\sqrt{2}}{2} \neq 0$ which is not acceptable in practice.

4) • Classic GS

    – ϵ=1e-4

$$Q = \begin{bmatrix} 0.999999990000000 & 9.999999878202985e-05 & 9.999999950000002e-05 \\ 9.999999900000002e-05 & 9.999999874516698e-09 & -0.999999995000000 \\ 9.999999900000002e-05 & -0.999999995000000 & 2.820298298412677e-09 \end{bmatrix}$$

$$\|Q^T Q - I\|_F = 3.988504118282294e-09$$

    – ϵ=1e-9

$$Q = \begin{bmatrix} 1 & 0 & 0 \\ 1.000000000000000e-09 & 0 & -1 \\ 1.000000000000000e-09 & -1 & 0 \end{bmatrix}$$

$$\|Q^T Q - I\|_F = 1$$

• Modified GS

    – ϵ=1e-4

$$Q = \begin{bmatrix} 0.999999990000000 & 9.999999878202985e-05 & 9.999999978202983e-05 \\ 9.999999900000002e-05 & 9.999999874516698e-09 & -0.999999995000000 \\ 9.999999900000002e-05 & -0.999999995000000 & 3.188926882126998e-17 \end{bmatrix}$$

$$\|Q^T Q - I\|_F = 5.640596652684282e-13$$

    – ϵ= 1e-9

$$Q = \begin{bmatrix} 1 & 0 & 0 \\ 1.000000000000000e-09 & 0 & -0.707106781186548 \\ 1.000000000000000e-09 & -1 & -0.707106781186548 \end{bmatrix}$$

$$\|Q^T Q - I\|_F = 2.000000000000000e-09$$

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:

$$\min_{\mathbf{x}\in\mathbb{R}^n} \|\mathbf{Ax} - \mathbf{b}\|_2^2 \tag{1}$$

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

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 (1) 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

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

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 (1) 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) Since the column vectors are consistent with which in problem 3, we here inherit the results.

$$q_1 = \frac{a_1}{||a_1||} = \begin{bmatrix} \frac{1}{\sqrt{30}} \\ \frac{2}{\sqrt{30}} \\ \frac{3}{\sqrt{30}} \\ \frac{4}{\sqrt{30}} \end{bmatrix}$$

$$\tilde{q}_2 = a_2 - (q_1^T a_2)q_1$$

$$= \begin{bmatrix} 2 \\ 3 \\ 4 \\ 5 \end{bmatrix} - \begin{bmatrix} \frac{1}{\sqrt{30}} \\ \frac{2}{\sqrt{30}} \\ \frac{3}{\sqrt{30}} \\ \frac{4}{\sqrt{30}} \end{bmatrix}^T \begin{bmatrix} 2 \\ 3 \\ 4 \\ 5 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{30}} \\ \frac{2}{\sqrt{30}} \\ \frac{3}{\sqrt{30}} \\ \frac{4}{\sqrt{30}} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \\ 0 \\ -\frac{1}{3} \end{bmatrix}$$

$$q_2 = \frac{\tilde{q}_2}{||\tilde{q}_2||} = \begin{bmatrix} \frac{\sqrt{6}}{3} \\ \frac{\sqrt{6}}{6} \\ 0 \\ -\frac{\sqrt{6}}{6} \end{bmatrix}$$

$$\tilde{q}_3 = a_3 - (q_1^T a_3)q_1 - (q_2^T a_3)q_2$$

$$= \begin{bmatrix} 3 \\ 5 \\ 7 \\ 11 \end{bmatrix} - \begin{bmatrix} \frac{1}{\sqrt{30}} \\ \frac{2}{\sqrt{30}} \\ \frac{3}{\sqrt{30}} \\ \frac{4}{\sqrt{30}} \end{bmatrix}^T \begin{bmatrix} 3 \\ 5 \\ 7 \\ 11 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{30}} \\ \frac{2}{\sqrt{30}} \\ \frac{3}{\sqrt{30}} \\ \frac{4}{\sqrt{30}} \end{bmatrix} - \begin{bmatrix} \frac{\sqrt{6}}{3} \\ \frac{\sqrt{6}}{6} \\ 0 \\ -\frac{\sqrt{6}}{6} \end{bmatrix}^T \begin{bmatrix} 3 \\ 5 \\ 7 \\ 11 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{6}}{3} \\ \frac{\sqrt{6}}{6} \\ 0 \\ -\frac{\sqrt{6}}{6} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{2}{5} \\ -\frac{1}{5} \\ -\frac{4}{5} \\ \frac{3}{5} \end{bmatrix}$$

$$q_3 = \frac{\tilde{q}_3}{||\tilde{q}_3||} = \begin{bmatrix} \frac{\sqrt{30}}{15} \\ -\frac{\sqrt{30}}{30} \\ -\frac{2\sqrt{30}}{15} \\ \frac{\sqrt{30}}{10} \end{bmatrix}$$

So we have the basis:

$$q_1 = \begin{bmatrix} \frac{1}{\sqrt{30}} \\ \frac{2}{\sqrt{30}} \\ \frac{3}{\sqrt{30}} \\ \frac{4}{\sqrt{30}} \end{bmatrix}, q_2 = \begin{bmatrix} \frac{\sqrt{6}}{3} \\ \frac{\sqrt{6}}{6} \\ 0 \\ -\frac{\sqrt{6}}{6} \end{bmatrix}, q_3 = \begin{bmatrix} \frac{\sqrt{30}}{15} \\ -\frac{\sqrt{30}}{30} \\ -\frac{2\sqrt{30}}{15} \\ \frac{\sqrt{30}}{10} \end{bmatrix}$$

$$A = QR = \begin{bmatrix} \frac{1}{\sqrt{30}} & \frac{\sqrt{6}}{3} & \frac{\sqrt{30}}{15} \\ \frac{2}{\sqrt{30}} & \frac{\sqrt{6}}{6} & -\frac{\sqrt{30}}{30} \\ \frac{3}{\sqrt{30}} & 0 & -\frac{2\sqrt{30}}{15} \\ \frac{4}{\sqrt{30}} & -\frac{\sqrt{6}}{6} & \frac{\sqrt{30}}{10} \end{bmatrix} \begin{bmatrix} \sqrt{30} & \frac{40}{\sqrt{30}} & \frac{78}{\sqrt{30}} \\ 0 & \frac{\sqrt{6}}{3} & 0 \\ 0 & 0 & -\frac{\sqrt{30}}{5} \end{bmatrix}$$

First we solve $z = Q^T b$

$$z = Q^T b = \begin{bmatrix} \frac{1}{\sqrt{30}} & \frac{\sqrt{6}}{3} & \frac{\sqrt{30}}{15} \\ \frac{2}{\sqrt{30}} & \frac{\sqrt{6}}{6} & -\frac{\sqrt{30}}{30} \\ \frac{3}{\sqrt{30}} & 0 & -\frac{2\sqrt{30}}{15} \\ \frac{4}{\sqrt{30}} & -\frac{\sqrt{6}}{6} & \frac{\sqrt{30}}{10} \end{bmatrix}^T \begin{bmatrix} 1 \\ \frac{3}{2} \\ 3 \\ 6 \end{bmatrix} = \begin{bmatrix} \frac{37}{\sqrt{30}} \\ -\frac{5\sqrt{6}}{12} \\ \frac{13}{2\sqrt{30}} \end{bmatrix}$$

Then we solve $Rx = z$ by backward substitution:

$$\begin{bmatrix} \sqrt{30} & \frac{40}{\sqrt{30}} & \frac{78}{\sqrt{30}} \\ 0 & \frac{\sqrt{6}}{3} & 0 \\ 0 & 0 & -\frac{\sqrt{30}}{5} \end{bmatrix} x = \begin{bmatrix} \frac{37}{\sqrt{30}} \\ -\frac{5\sqrt{6}}{12} \\ \frac{13}{2\sqrt{30}} \end{bmatrix} \Rightarrow x = \begin{bmatrix} \frac{1}{12} \\ -\frac{5}{4} \\ \frac{13}{12} \end{bmatrix}$$

2) By the definition of pseudo-inverse, we deduce the following equality:

$$A\delta x = b + \delta b - Ax = \delta b$$

$$\Leftrightarrow A^T A \delta x = A^T \delta b$$

$$\Leftrightarrow (A^T A)^{-1} A^T A \delta x = (A^T A)^{-1} A^T \delta b$$

$$\Leftrightarrow \delta x = A^\dagger \delta b$$

Similarly

$$x = A^\dagger b$$

- We show $\frac{1}{\|\mathbf{A}\|\|\mathbf{A}^\dagger\|} \frac{\|\delta \mathbf{b}\|}{\|\mathbf{b}\|} \leq \frac{\|\delta \mathbf{x}\|}{\|\mathbf{x}\|}$.

$$\frac{\|\delta x\|}{\|x\|} = \frac{\|\delta x\|}{\|A^\dagger b\|}$$

$$\geq \frac{\|\delta x\|}{\|A^\dagger\|\|b\|}$$

Since $\|A\|\|\delta x\| \geq \|A\delta x\| = \|\delta b\| \Rightarrow \|\delta x\| \geq \frac{\|\delta b\|}{\|A\|}$, we further have:

$$\frac{\|\delta x\|}{\|x\|} \geq \frac{\|\delta x\|}{\|A^\dagger\|\|b\|}$$

$$\geq \frac{\|\delta b\|}{\|A\|} \frac{1}{\|A^\dagger\|\|b\|}$$

Q.E.D.

- We show $\frac{\|\delta \mathbf{x}\|}{\|\mathbf{x}\|} \leq \|\mathbf{A}\|\|\mathbf{A}^\dagger\|\frac{\|\delta \mathbf{b}\|}{\|\mathbf{b}\|}$.

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

Since $\|b\| = \|Ax\| \leq \|A\|\|x\| \Rightarrow \|x\| \geq \frac{\|b\|}{\|A\|}$, we further have:

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

Q.E.D.

- To sum up, the influence on the solution $\frac{\|\delta x\|}{\|x\|}$ is bounded by the condition number $\|\mathbf{A}\|\|\mathbf{A}^\dagger\|$ and the noise ration $\frac{\|\delta b\|}{\|b\|}$.

3)

$$\begin{bmatrix} 1 & 2 & 2 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \\ 1 & 1 & 0 \end{bmatrix}^T \begin{bmatrix} 1 & 2 & 2 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \\ 1 & 1 & 0 \end{bmatrix} x_{LS} = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \\ 1 & 1 & 0 \end{bmatrix}^T \begin{bmatrix} 1 \\ \frac{3}{2} \\ 3 \\ 6 \end{bmatrix}$$

$$x_{LS} = \begin{bmatrix} \frac{17}{13} & -\frac{17}{13} & \frac{13}{2} \\ -\frac{17}{13} & \frac{30}{13} & -\frac{15}{13} \\ \frac{2}{13} & -\frac{15}{13} & \frac{14}{13} \end{bmatrix} \begin{bmatrix} 19 \\ 20 \\ 14 \end{bmatrix} = \begin{bmatrix} \frac{11}{13} \\ \frac{67}{13} \\ -\frac{66}{13} \end{bmatrix}$$

4)

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \end{bmatrix}^T \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \end{bmatrix} x_{LS} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \end{bmatrix}^T \begin{bmatrix} 1 \\ \frac{3}{2} \\ 3 \\ 6 \end{bmatrix}$$

$$x_{LS} = \begin{bmatrix} \frac{31}{4} & -\frac{27}{4} & \frac{5}{4} \\ -\frac{27}{4} & \frac{129}{20} & -\frac{5}{4} \\ \frac{5}{4} & -\frac{5}{4} & \frac{1}{4} \end{bmatrix} \begin{bmatrix} \frac{23}{2} \\ 37 \\ 130 \end{bmatrix} = \begin{bmatrix} \frac{15}{8} \\ -\frac{59}{40} \\ -\frac{5}{8} \end{bmatrix}$$

5) - In 3).

$$A^\dagger = (A^T A)^{-1} A^T = \begin{bmatrix} -1 & \frac{4}{13} & \frac{6}{13} & 0 \\ 1 & -\frac{4}{13} & -\frac{6}{13} & 1 \\ 0 & \frac{2}{13} & \frac{3}{13} & -1 \end{bmatrix}$$

$$\delta x = A^\dagger \delta b = \begin{bmatrix} -1 & \frac{4}{13} & \frac{6}{13} & 0 \\ 1 & -\frac{4}{13} & -\frac{6}{13} & 1 \\ 0 & \frac{2}{13} & \frac{3}{13} & -1 \end{bmatrix} \begin{bmatrix} \frac{1}{10} \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{1}{10} \\ \frac{1}{10} \\ 0 \end{bmatrix}$$

The 2-norm condition number:

$$\|A\|\|A^\dagger\| = 6.9824 \times 1.9084 = 13.3254$$

The influence of noise δb:

$$\frac{\|\delta x\|}{\|x\|} = \frac{0.1414}{7.2838} = 0.0194$$

- In 4).

$$A^\dagger = (A^T A)^{-1} A^T = \begin{bmatrix} \frac{9}{4} & -\frac{3}{4} & -\frac{5}{4} & \frac{3}{4} \\ -\frac{31}{20} & \frac{23}{20} & \frac{27}{20} & -\frac{19}{20} \\ \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & \frac{1}{4} \end{bmatrix}$$

$$\delta x = A^\dagger \delta b = \begin{bmatrix} \frac{9}{4} & -\frac{3}{4} & -\frac{5}{4} & \frac{3}{4} \\ -\frac{31}{20} & \frac{23}{20} & \frac{27}{20} & -\frac{19}{20} \\ \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & \frac{1}{4} \end{bmatrix} \begin{bmatrix} \frac{1}{10} \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{9}{40} \\ -\frac{31}{200} \\ \frac{1}{40} \end{bmatrix}$$

The 2-norm condition number:

$$\|A\|\|A^\dagger\| = 3.7558 \times 19.6214 = 73.6945$$

The influence of noise δb:

$$\frac{\|\delta x\|}{\|x\|} = \frac{0.2744}{2.4661} = 0.1113$$

- To sum up, under the same noise ratio $\frac{\|\delta b\|}{\|b\|}$, larger condition number, *i.e.* condition number in 3) is larger than which in 4), leads to a heavier influence on the solution $\frac{\|\delta x\|}{\|x\|}$.

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)

$$v_1 = a_1 - \|a_1\|_2 e_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \\ 0 \end{bmatrix} - 3 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \\ 2 \\ 0 \end{bmatrix}$$

$$H_1 = I - \frac{2}{\|v_1\|_2^2} v_1 v_1^T = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} & 0 \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} & 0 \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$A^{(1)} = H_1 A^T = \begin{bmatrix} 3 & 0 & 8 \\ 0 & -2 & -1 \\ 0 & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

$$v_2 = a_2 - \|a_2\|_2 e_1 = \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix} - 3 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -5 \\ 2 \\ 1 \end{bmatrix}$$

$$H_2 = \begin{bmatrix} I & 0 \\ 0 & \tilde{H}_2 \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I - \frac{2}{\|v_2\|_2^2} v_2 v_2^T \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -\frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ 0 & \frac{2}{3} & \frac{11}{15} & -\frac{2}{15} \\ 0 & \frac{1}{3} & -\frac{2}{15} & \frac{14}{15} \end{bmatrix}$$

$$A^{(2)} = H_2 A^{(1)} = \begin{bmatrix} 3 & 0 & 8 \\ 0 & 3 & 1 \\ 0 & 0 & -\frac{4}{5} \\ 0 & 0 & \frac{3}{5} \end{bmatrix}$$

$$v_3 = a_3 - \|a_3\| e_1 = \begin{bmatrix} -\frac{4}{5} \\ \frac{3}{5} \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{9}{5} \\ \frac{3}{5} \end{bmatrix}$$

$$H_3 = \begin{bmatrix} I & 0 \\ 0 & \tilde{H}_3 \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I - \frac{2}{\|v_3\|_2^2} v_3 v_3^T \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -\frac{4}{5} & \frac{3}{5} \\ 0 & 0 & \frac{3}{5} & \frac{4}{5} \end{bmatrix}$$

$$A^{(3)} = H_3 A^{(2)} = \begin{bmatrix} 3 & 0 & 8 \\ 0 & 3 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$Q = H_1 H_2 H_3 = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} & 0 \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} & 0 \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -\frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ 0 & \frac{2}{3} & \frac{11}{15} & -\frac{2}{15} \\ 0 & \frac{1}{3} & -\frac{2}{15} & \frac{14}{15} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -\frac{4}{5} & \frac{3}{5} \\ 0 & 0 & \frac{3}{5} & \frac{4}{5} \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & 0 & -\frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} & 0 \\ \frac{2}{3} & \frac{2}{3} & 0 & -\frac{1}{3} \\ 0 & \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \end{bmatrix}$$

$$R = A^{(3)} = \begin{bmatrix} 3 & 0 & 8 \\ 0 & 3 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} x$$

2) We have

$$A^T = QR = [Q_1 \quad Q_2] \begin{bmatrix} R_1 \\ 0 \end{bmatrix} = Q_1 R_1 + Q_2 0$$

Note

$$Ax = R_1^T Q_1^T x + 0^T Q_2^T = b$$

which indicates

$$Q_1^T x = R_1^{-T} b$$

and $Q_2^T x$ can be anything, which we set to be d. The we have

$$\begin{bmatrix} Q_1^T x \\ Q_2^T x \end{bmatrix} = Q^T x = \begin{bmatrix} R_1^{-T} b \\ d \end{bmatrix}$$

The solution is

$$x = Q \begin{bmatrix} R_1^{-T} b \\ d \end{bmatrix} = Q_1 R^{-T} b + Q_2 d$$

So one possible solution locates in $d = 0$, which is

$$x = Q_1 R_1^{-T} b = \begin{bmatrix} \frac{1}{3} & 0 & -\frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} & 0 \\ \frac{2}{3} & \frac{2}{3} & 0 & -\frac{1}{3} \\ 0 & \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \end{bmatrix} \begin{bmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ -\frac{8}{3} & -\frac{1}{3} & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} \frac{17}{9} b_1 + \frac{2}{9} b_2 - \frac{2}{3} b_3 \\ -\frac{2}{3} b_1 - \frac{1}{3} b_2 + \frac{1}{3} b_3 \\ \frac{2}{9} b_1 + \frac{2}{9} b_2 \\ -\frac{16}{9} - \frac{1}{9} b_2 + \frac{2}{3} b_3 \end{bmatrix}$$

VI. SOLVING LS VIA PROJECTION

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

Consider the Least Square (LS) problem:

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

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

$$X_{\mathrm{LS}} = \left\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{y}\right\}$$

as the set of all solutions to (2), and

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

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

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

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 $X_{\mathrm{LS}} = \{\mathbf{x}_{\mathrm{LS}}\}$ if and only if $\mathrm{rank}(\mathbf{A}) = n$.

Solution.

1) • We show that x_{LS} is a solution to (2).

$$\begin{aligned}
x_{LS} &= A^\dagger y \\
\Leftrightarrow A^T A x_{LS} &= A^T A A^\dagger y \\
&= A^T (A A^\dagger)^T y \\
&= [(A A^\dagger)A]^T y \\
&= (A A^\dagger A)^T y \\
&= A^T y
\end{aligned}$$

Q.E.D.

- Secondly, we show that x_{LS} is of minimum 2-norm in X_{LS}. The key information we can obtain is that

$$\Pi_{\mathcal{N}(A)} x_{LS} = (I - A^\dagger A) A^\dagger y$$
$$= A^\dagger y - (A^\dagger A A^\dagger) y$$
$$= A^\dagger y - A^\dagger y$$
$$= 0$$

which indicates that x_{LS} is orthogonal to $\mathcal{N}(A)$. Furthermore, we find that $\forall x \in X_{LS}$ satisfies:

$$(x - x_{LS})^T x_{LS} = (x - x_{LS})^T A^\dagger y$$
$$= (x - x_{LS})^T A^\dagger A A^\dagger y$$
$$= (x - x_{LS})^T (A^\dagger A)^T A^\dagger y$$
$$= \left(A^\dagger A (x - x_{LS}) \right)^T A^\dagger y$$
$$= \left(A^\dagger A A^\dagger A (x - x_{LS}) \right)^T A^\dagger y$$
$$= \left(A^\dagger (A^\dagger)^T A^T A (x - x_{LS}) \right)^T A^\dagger y$$
$$= \left(A^\dagger (A^\dagger)^T (y - y) \right)^T A^\dagger y$$
$$= 0$$

which indicates $x - x_{LS} \perp x_{LS}$. Thus we have $\forall x \in X_{LS}$:

$$\|x\|_2^2 = \|x - x_{LS} + x_{LS}\|_2^2$$
$$= \|x - x_{LS}\|_2^2 + \|x_{LS}\|_2^2$$
$$\geq \|x_{LS}\|_2^2$$

which indicates that x_{LS} is of minimum 2-norm in X_{LS}.

- In conclusion, $\mathbf{x}_{\mathrm{LS}}$ is a solution to (2) and is of minimum 2-norm in X_{LS}.

2) - We show $X_{\mathrm{LS}} = \{\mathbf{x}_{\mathrm{LS}}\} \Rightarrow \mathrm{rank}(\mathbf{A}) = n$.

First, we show that $B = \{x - x_{LS} | x \in X_{LS}\} \subseteq \mathcal{N}(A)$. This is easy to prove since

$$A(x - x_{LS}) = A(x_{LS} - x_{LS}) = 0$$

Secondly, we show that $\mathcal{N}(A) \subseteq B = \{x - x_{LS} | x \in X_{LS}\}$. For $\forall y \in \mathcal{N}(A)$

$$A^T A y = A^T 0 = 0$$

which implies

$$A^T A (y + x_{LS}) = A^T A x_{LS} = A^T y$$

So $y + x_{LS} \in X_{LS}$ and it indicates that $y = x - x_{LS}, x \in X_{LS}$. This means $\forall y \in \mathcal{N}(A) \Rightarrow y \in B$. So $\mathcal{N}(A) \subseteq B$. So $\mathcal{N}(A) = B = 0$.

$$\mathcal{N}(A) = 0 \Rightarrow rank(A) = n$$

.

- We show $\text{rank}(\mathbf{A}) = n \Rightarrow X_{\text{LS}} = \{\mathbf{x}_{\text{LS}}\}$. Since $rank(A) = rank(A^T) = n$, we know that $\mathcal{N}(A) = 0$.