SI231 - Matrix Computations, Fall 2020-21

Homework Set #3

Prof. Yue Qiu and Prof. Ziping Zhao

Name: Chengrui Zhang **Major:** Master in CS

Student No.: 2019233183 **E-mail:** zhangchr@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. Please insert your solution here ...

- 1) Since P is a projector, therefore, P satisfies $P^2 = P$, assume that P is a non-singular matrix, then $P * P^{-1} = I$. Thus, $P^2 * P^{-1} = P * P^{-1} = I$, therefore, $P = I$, contradiction. $\therefore$ P must be singular.
- 2) The answer is $I - P$. There are 3 reasons: (1), $(I-P)^2 = I + P^2 - 2P = I - P$, $\therefore (I - P)$ is a projector. (2), the project subspace of $(I - P)$ is $R(I - P) = N(P) = U^{\perp}$, (3), $(I - P)^T = I - P^T$, since P is an orthogonal projector. $\therefore P = P^T, (I - P)^T = I - P$. In general, $(I - P)$ is an orthogonal projector onto $U^{\perp}$.
- 3) if $P_u P_w = 0$, then, suppose $u \in U(U = R(P_u)), w \in W, (W = R(P_w))$, then $P_u P_w \cdot w = P_u \cdot w' = 0, w' \in W$. $\therefore w' \in N(P_u) = U^{\perp}$, since w' is arbitrary vector in W, $\therefore W = U^{\perp}, W \perp U$.

  if $U \perp W$, then $N(P_u) = U^{\perp} = W, N(P_w) = W^{\perp} = U$. Suppose $w \in W, v \in W^{\perp}$, then $P_u P_w \cdot w = P_u \cdot w', w' \in W, \therefore P_u \cdot w' = 0$, and $P_u P_w \cdot v = P_u \cdot (P_w \cdot v) = P_u \cdot 0 = 0$. From above, since every element in V can be written as the sum of a vector in W and a vector in $W^{\perp}$, $\therefore P_u P_w = 0$, Q.E.D.

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. Please insert your solution here ...

- 1) and 2): the codes are shown in **HW3_1.m**
- 3): a. The FLOPS are shown in **HW3_2.m**. For gradient decent method: we have nearly $(2m(n+1)+n) * iteration\_number$ FLOPS.

For normal equation method, we have nearly $n^2(2+m+\frac{1}{3}n)$ FLOPS.

b. Note that here the gradient norm does not times γ

************ Q1 ***********

The gradient norm of normal equation method is: 1.280356e-05

The loss f(x) of gradient descent method is: 2.664261e+04

************ Q2 ***********

The gradient norm of normal equation method is: 3.244141e-09

The loss f(x) of normal equation method is: 2.664261e+04

I find that the loss of two methods is the same, and the normal equation method is more convergence than the gradient decent method.

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 $\{\mathbf{a}_1, \dots, \mathbf{a}_4\}$, 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-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 $\mathbf{a}_1, \dots, \mathbf{a}_n$.

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

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

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

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

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 $\langle , \rangle$ 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

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

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-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 $\mathbf{a}_1, \ldots, \mathbf{a}_n$.

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

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

  $\tilde{p}_i = a_i$

end

for $j = 1, \ldots, n$ **do**

  $p_j = \tilde{p}_j / ||\tilde{p}_j||$

  for $k = j+1, \ldots, n$ **do**

    $\tilde{p}_j = \tilde{p}_j - p_i^T \tilde{p}_j p_i$

  end

end

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)

$$\widetilde{q_1} = a_1 = [1, 2, 3, 4]^T, \qquad q_1 = \frac{\widetilde{q_1}}{\|\widetilde{q_1}\|_2} = \frac{1}{\sqrt{30}}\widetilde{q_1}$$
$$\widetilde{q_2}^{(1)} = a_2 - (q_1^T a_2) q_1 = \frac{1}{3}[2, 1, 0, -1]^T, \qquad q_2 = \frac{\widetilde{q_2}}{\|\widetilde{q_2}\|_2} = \frac{1}{\sqrt{6}} \cdot \begin{bmatrix} 2 & 1 & 0 & -1 \end{bmatrix}^T$$
$$\widetilde{q_3} = a_3 - (q_1^T a_3) q_1 - (q_2^T a_3) q_2 = 0$$

Thus, a_3 cannot be orthogonal to $(\alpha q_1 + \beta q_2)$.

$$\widetilde{q_4} = a_4 - (q_1^T a_4) q_1 - (q_2^T a_4) q_2 = \frac{1}{5}[2, -1, -4, -3]^T, \qquad q_4 = \frac{1}{\sqrt{30}} \cdot \begin{bmatrix} 2 & 1 & -4 & 3 \end{bmatrix}^T$$

- 2) For the projection of vector:

$$
\begin{aligned}
&\tilde{p}_1 = u_1, p_1 = \frac{u_1}{||u_1||} \\
&\tilde{p}_2 = u_2 - proj_{u1}(u_2), p_2 = \frac{u_2}{||u_2||} \\
&\tilde{p}_3 = u_3 - proj_{u1}(u_3) - proj_{u2}(u_3), p_3 = \frac{u_3}{||u_3||} \\
&\vdots \\
&\tilde{p}_k = u_k - \sum_{j=1}^{k-1} proj_{u1}(u_k), p_k = \frac{u_k}{||u_k||}
\end{aligned}
$$

  For the projection of subspace:

  The orthogonal projection of any vector x onto U is the point $p = \sum_{i=1}^{n} \langle x, \hat{u}_i \rangle \hat{u}_i$

- 3)

  MGS:

$$
\widetilde{q_1} = a_1 = [1, \epsilon, \epsilon]^T, \qquad q_1 = \frac{\widetilde{q_1}}{\|\widetilde{q_1}\|_2}
$$

$$
\widetilde{q_2}^{(1)} = a_2 - (q_1^T a_2) q_1 = - \begin{bmatrix} 1 \\ \epsilon \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1, \epsilon, \epsilon \end{bmatrix} \cdot \begin{bmatrix} 1 \\ \epsilon \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ \epsilon \\ \epsilon \end{bmatrix} = - \begin{bmatrix} 0 \\ 0 \\ \epsilon \end{bmatrix} \qquad q_2 = \frac{\widetilde{q_2}}{\|\widetilde{q_2}\|_2} = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}
$$

$$
\widetilde{q_3}^{(1)} = a_3 - (q_1^T a_3) q_1 = \begin{bmatrix} 1 \\ 0 \\ \epsilon \end{bmatrix} - \begin{bmatrix} 1, \epsilon, \epsilon \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ \epsilon \end{bmatrix} \cdot \begin{bmatrix} 1 \\ \epsilon \\ \epsilon \end{bmatrix} = \begin{bmatrix} 0 \\ -\epsilon \\ 0 \end{bmatrix}
$$

$$
\widetilde{q_3}^{(2)} = \widetilde{q_3}^{(2)} - (q_2^T \widetilde{q_3}^{(2)}) q_2 = \begin{bmatrix} 0 \\ -\epsilon \\ 0 \end{bmatrix} - \begin{bmatrix} 0, 0, -1 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ -\epsilon \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ -\epsilon \\ 0 \end{bmatrix}
$$

$$
q_3 = \frac{\widetilde{q_3^{(2)}}}{\|\widetilde{q_3^{(2)}}\|_2} = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}
$$

  CGS

$$\widetilde{q_1} = a_1 = [1, \epsilon, \epsilon]^T, \qquad q_1 = [1, \epsilon, \epsilon]^T$$

$$\widetilde{q_2}^{(1)} = a_2 - (q_1^T a_2) q_1 = -\begin{bmatrix} 0 & 0 & \epsilon \end{bmatrix}^T \qquad q_2 = \begin{bmatrix} 0 & 0 & -1 \end{bmatrix}^T$$

$$\widetilde{q_3} = a_3 - (q_1^T a_3) q_1 - (q_2^T a_3) q_2 = \begin{bmatrix} 1 \\ 0 \\ \epsilon \end{bmatrix} - \begin{bmatrix} 1, \epsilon, \epsilon \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ \epsilon \end{bmatrix} \cdot \begin{bmatrix} 1 \\ \epsilon \\ \epsilon \end{bmatrix} - \begin{bmatrix} 0, 0, -1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ \epsilon \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ -\epsilon \\ -\epsilon \end{bmatrix}$$

$$q_3 = \frac{\widetilde{q_3}}{\|\widetilde{q_3}\|_2} = \begin{bmatrix} 0 & -\frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{bmatrix}^T$$

- 4) The code is shown in **HW3_3.m**.

  The $||Q^T Q - I||$ of CGS method at $eps = 1.000000e^{-04}$ is: $3.988504e^{-09}$;

  The $||Q^T Q - I||$ of MGS method at $eps = 1.000000e^{-04}$ is: $5.640597e^{-13}$

  The $||Q^T Q - I||$ of CGS method at $eps = 1.000000e^{-09}$ is: 1;

  The $||Q^T Q - I||$ of MGS method at $eps = 1.000000e^{-09}$ is: $2e^{-9}$;

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:

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

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

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 $\mathbf{A}\mathbf{x}=\mathbf{b}$, 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 $\mathbf{A}^T\mathbf{A}\mathbf{x}_{LS}=\mathbf{A}^T\mathbf{b}$ 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 $\mathbf{A}^T\mathbf{A}\mathbf{x}_{LS}=\mathbf{A}^T\mathbf{b}$ 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. Please insert your solution here ...

- 1) $Q = \frac{1}{30}\begin{bmatrix} \sqrt{30} & 10\sqrt{6} & 2\sqrt{30} \\ 2\sqrt{30} & 5\sqrt{6} & -\sqrt{30} \\ 3\sqrt{30} & 0 & -4\sqrt{30} \\ 4\sqrt{30} & -5\sqrt{6} & 3\sqrt{30} \end{bmatrix}, R = \frac{1}{15}\begin{bmatrix} 15\sqrt{30} & 20\sqrt{30} & 39\sqrt{30} \\ 0 & 5\sqrt{6} & 0 \\ 0 & 0 & 3\sqrt{30} \end{bmatrix},$

  $x = [\frac{1}{12}, \frac{-4}{5}, \frac{13}{12}], err = ||Ax - b||_2^2 = \frac{1}{6}$
- 2)

$$\textit{Let,} Ax = b, x = A^{\dagger}b, A\tilde{x} = \tilde{b} = b + \delta b, \tilde{x} = A^{\dagger}\tilde{b}$$

$$\textit{Then,} ||\delta x|| = ||x - \tilde{x}|| = ||A^{\dagger}(b - \tilde{b})|| = ||A^{\dagger}\delta b|| \leq ||A^{\dagger}|| \cdot ||\delta b||$$

$$b = Ax \Rightarrow ||b|| \leq ||A|| \cdot ||x||, \therefore \frac{1}{||x||} \leq \frac{||A||}{||b||}, \therefore \frac{||\delta x||}{||x||} \leq ||A|| \cdot ||A^{\dagger}|| \cdot \frac{||\delta b||}{||b||}$$

$$||\delta b|| = ||b - \tilde{b}|| = ||A^{\dagger}(x - \tilde{x})|| = ||A^{\dagger}\delta x|| \leq ||A^{\dagger}|| \cdot ||\delta x||$$

$$x = A^{\dagger}b \Rightarrow ||x|| \leq ||A^{\dagger}|| \cdot ||b||, \therefore \frac{1}{||b||} \leq \frac{||A^{\dagger}||}{||x||}, \therefore \frac{||\delta b||}{||b||} \leq ||A|| \cdot ||A^{\dagger}|| \cdot \frac{||\delta x||}{||x||}$$

$$\therefore \frac{1}{||A|| \cdot ||A^{\dagger}||} \cdot \frac{\delta b}{b} \leq \frac{||\delta x||}{||x||}, Q.E.D$$

- 3) Since A is full column rank, A^tA is a positive definite matrix, therefore it must have inverse matrix.

  $\therefore x_{LS} = (A^TA)^{-1}A^Tb = \begin{bmatrix} -1 & \frac{4}{13} & \frac{6}{13} & 0 \\ 1 & \frac{-4}{13} & \frac{-6}{13} & 1 \\ 0 & \frac{2}{13} & \frac{3}{13} & -1 \end{bmatrix} \cdot b = \begin{bmatrix} \frac{11}{13} & \frac{67}{13} & \frac{-66}{13} \end{bmatrix}^T$

  When there is $b + \delta b$, then

  $\therefore x_{LS} = (A^TA)^{-1}A^T(b + \delta b) = \begin{bmatrix} -1 & \frac{4}{13} & \frac{6}{13} & 0 \\ 1 & \frac{-4}{13} & \frac{-6}{13} & 1 \\ 0 & \frac{2}{13} & \frac{3}{13} & -1 \end{bmatrix} \cdot (b + \delta b) = \begin{bmatrix} \frac{97}{130} & \frac{683}{130} & \frac{-66}{13} \end{bmatrix}^T$

  The code is shown in **HW3_4.m**
- 4) Similar to 3), I get:

  $\therefore x_{LS} = (A^TA)^{-1}A^Tb = \frac{1}{20}\begin{bmatrix} 45 & -15 & -25 & 15 \\ -31 & 23 & 27 & -19 \\ 5 & -5 & -5 & 5 \end{bmatrix} \cdot b = \begin{bmatrix} \frac{15}{8} & \frac{-59}{40} & \frac{5}{8} \end{bmatrix}^T$

  When there is $b + \delta b$, then

  $\therefore x_{LS} = (A^TA)^{-1}A^T(b + \delta b) = \frac{1}{20}\begin{bmatrix} 45 & -15 & -25 & 15 \\ -31 & 23 & 27 & -19 \\ 5 & -5 & -5 & 5 \end{bmatrix} \cdot (b + \delta b) = \begin{bmatrix} \frac{21}{10} & \frac{-163}{100} & \frac{13}{20} \end{bmatrix}^T$

  The code is shown in **HW3_4.m**
- 5) For condition number:

  $k_1 = ||A_3|| \cdot ||A_3^{\dagger}|| = \frac{5610}{421} \approx 13.3254, k_2 = ||A_4|| \cdot ||A_4^{\dagger}|| = \frac{2653}{36} \approx 73.6945$ and the influences are $I1 = \frac{||\tilde{x_3} - x_3||}{||x_3||} = \frac{125}{6438} \approx 0.01942, I2 = \frac{||\tilde{x_4} - x_4||}{||x_4||} = \frac{1231}{11065} \approx 0.11125$

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. Please insert your solution here ...

1)

$$A^T = \begin{bmatrix} 1 & 0 & 2 \\ 2 & -2 & 5 \\ 2 & 2 & 6 \\ 0 & 1 & 1 \end{bmatrix}, ||A_{*1}||_2 = 3, \therefore v_1 = A_{*1} + ||A_{*1}||_2 e_1 = [4,2,2,0]^T$$

$$\therefore H_1 = I - \frac{2}{||V_1||_2^2} \cdot v_1 v_1^T = \frac{1}{3}\begin{bmatrix} -1 & -2 & -2 & 0 \\ -2 & 2 & -1 & 0 \\ -2 & -1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$H_1 A^T = \begin{bmatrix} -3 & 0 & -8 \\ 0 & -2 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix}, ||A_{*2}||_2 = 3, \therefore \tilde{v}_2 = A_{*2} + ||A_{*2}||_2 e_2 = [-5,2,1]^T, v_2 = [0,-5,2,1]^T$$

$$\therefore H_2 = I - \frac{2}{||V_2||_2^2} \cdot v_2 v_2^T = \frac{1}{15}\begin{bmatrix} 15 & 0 & 0 & 0 \\ 0 & -10 & 10 & 5 \\ 0 & 10 & 11 & -2 \\ 0 & 5 & -2 & 14 \end{bmatrix}$$

$$H_2 H_1 A^T = \begin{bmatrix} -3 & 0 & -8 \\ 0 & 3 & 1 \\ 0 & 0 & 0.6 \\ 0 & 0 & 0.8 \end{bmatrix}, ||A_{*3}||_2 = 1, \therefore \tilde{v}_3 = A_{*3} + ||A_{*3}||_2 e_3 = [1.6, 0.8]^T, v_3 = [0,0,1.6,0.8]^T$$

$$\therefore H_3 = I - \frac{2}{||V_3||_2^2} \cdot v_3 v_3^T = \frac{1}{5}\begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & -3 & -4 \\ 0 & 0 & -4 & 3 \end{bmatrix}, H_3H_2H_1A = \begin{bmatrix} -3 & 0 & -8 \\ 0 & 3 & 1 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix} = R$$

$$Q = (H_3H_2H_1)^{-1}\frac{1}{3}\begin{bmatrix} -1 & 0 & 2 & 2 \\ -2 & -2 & -1 & 0 \\ -2 & 2 & 0 & -1 \\ 0 & 1 & -2 & 2 \end{bmatrix}$$

2)

$Ax = R^TQ^Tx = R^T \cdot z = b$, we solve z first. $z_1 = -\frac{1}{3}b_1, z_2 = \frac{1}{3}b_2, z_1 = \frac{8}{3}b_1 + \frac{1}{3}b_2 - b_3, z_4 = *$.

Here we set $z_4 = 0$. Then for $Q^Tx = z, x = Qz$, we get $x = \frac{1}{3}\begin{bmatrix} -1 & 0 & 2 & 2 \\ -2 & -2 & -1 & 0 \\ -2 & 2 & 0 & -1 \\ 0 & 1 & -2 & 2 \end{bmatrix} \cdot z$

Therefore, $x_1 = \frac{17}{9}b_1 + \frac{2}{9}b_2 + \frac{-2}{3}b_3, x_2 = \frac{-2}{3}b_1 + \frac{-1}{3}b_2 + \frac{1}{3}b_3, x_3 = \frac{2}{9}b_1 + \frac{2}{9}b_2, x_4 = \frac{-16}{9}b_1 + \frac{-1}{9}b_2 + \frac{2}{3}b_3$

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

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

as the set of all solutions to (2), 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 $\mathbf{x}_{\text{LS}}$ is a solution to (2) 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. Please insert your solution here ...

- 1) First we prove the $x_{LS}\in X_{LS}$.

$$\because A^TAx_{LS} = A^TAA^\dagger y = A^T(AA^\dagger)^Ty \tag{3}$$

Do transpose in both side of Eq.(2), then

$$y^T(AA^\dagger)A = y^TA, \therefore Eq.(3) = A^Ty$$

$$\therefore x_{LS} \textit{ is a solution of Eq.(2)}$$

- 2) if $rank(A)=n$, then $rank(A^TA) = rank(A) =$ is full rank,

$$\therefore x = (A^TA)^{-1}A^T = A^\dagger y \textit{ is unique}, \therefore X_{LS} = \{x_{LS}\}$$

if $X_{LS} = \{x_{LS}\}$, then $x\in X_{LS}$ is unique, suppose that $rank(A)\leq n$, then $rank(A^TA) = rank(A) < n$ is not full rank, which means that $(A^TA)x = A^Ty = b$ does not have unique solution, contradiction, $\therefore rank(A) = n$, Q.E.D.