SI231 - Matrix Computations, Fall 2020-21

Homework Set #3

Prof. Yue Qiu and Prof. Ziping Zhao

Name: GongChenyu **Major:** Master in IE

Student No.: 2020232133 **E-mail:** gongchy@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) Set $\mathbf{u} \in \mathcal{U}, \mathbf{w} \in \mathcal{U}^{\perp}, \mathbf{v} = \mathbf{u} + \mathbf{w} \in \mathbb{R}^{\mathbf{n}}$.

   $\because \mathbf{P} \in \mathbb{R}^{n \times n}$ is a projector onto a subspace $\mathcal{U}$ along its orthogonal complement $\mathcal{U}^{\perp}$

   $\therefore \mathbf{P^2 v} = \mathbf{P} \cdot \mathbf{Pv} = \mathbf{Pu} = \mathbf{u} = \mathbf{Pv}$

   $\therefore \mathbf{P^2} = \mathbf{P} \Rightarrow \mathbf{P}$ is a idempotent matrix. And the eigenvalue of $\mathbf{P}$ must be 0 or 1.

   $\therefore$ det($\mathbf{P}$) = 0 or det($\mathbf{P}$) = 1 If det($\mathbf{P}$) = 1, in other words, $\mathbf{P^{-1}}$ exists.

   $\because \mathbf{P^2} = \mathbf{P}$, then left and right both multiply $\mathbf{P^{-1}}$, we can get $\mathbf{P^2 P^{-1}} = \mathbf{PP^{-1}}$, $\Rightarrow \mathbf{P} = \mathbf{I}$.

   $\therefore$ If $\mathbf{P}$ is not $\mathbf{I}$, $\mathbf{P}$ must be singular.

2) $\because \mathbf{P} \in \mathbb{R}^{n \times n}$ is a orthogonal projector onto a subspace $\mathcal{U}$ along its orthogonal complement $\mathcal{U}^{\perp}$

   $\therefore$ for $\mathbf{x} \in \mathcal{U}$ and $\mathbf{y} \in \mathcal{U}^{\perp}$, $\mathbf{v} = \mathbf{x} + \mathbf{y} \in \mathbb{R}^{\mathbf{n}}$, then $\mathbf{Pv} = \mathbf{x}$.

   Then $\mathbf{v} = \mathbf{Pv} + \mathbf{y} \ \Rightarrow \ (\mathbf{I} - \mathbf{P})\mathbf{v} = \mathbf{y}$, so orthogonal projector onto a subspace $\mathcal{U}^{\perp}$ along its orthogonal complement $\mathcal{U}$ is $(\mathbf{I} - \mathbf{P})$.

3) Firstly,

   If $\mathcal{U} \perp \mathcal{W}, \mathbf{P}_{\mathcal{W}} = \mathbf{I} - \mathbf{P}_{\mathcal{U}}$, then $\mathbf{P}_{\mathcal{U}}\mathbf{P}_{\mathcal{W}} = \mathbf{P}_{\mathcal{U}}(\mathbf{I} - \mathbf{P}_{\mathcal{U}}) = \mathbf{P}_{\mathcal{U}} - \mathbf{P}_{\mathcal{U}}\mathbf{P}_{\mathcal{U}}$,

   $\because \mathbf{P}_{\mathcal{U}}$ is a idempotent matrix,

   $\therefore \mathbf{P}_{\mathcal{U}} = \mathbf{P}_{\mathcal{U}}\mathbf{P}_{\mathcal{U}}$

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

   Secondly,

   $\because \mathbf{P}_{\mathcal{U}}\mathbf{P}_{\mathcal{W}} = \mathbf{0}$,

   $\therefore$ all the column vectors of $\mathbf{P}_{\mathcal{W}}$ must belong to $\mathcal{N}(\mathbf{P}_{\mathcal{U}})$

   $\therefore \mathcal{R}(\mathbf{P}_{\mathcal{W}}) \subseteq \mathcal{N}(\mathbf{P}_{\mathcal{U}})$

   $\because \mathcal{R}(\mathbf{P}_{\mathcal{W}}) = \mathcal{W}$ and $\mathcal{N}(\mathbf{P}_{\mathcal{U}}) = \mathcal{U}^{\perp}$

   $\therefore \mathbf{P}_{\mathcal{U}}\mathbf{P}_{\mathcal{W}} = \mathbf{0} \Rightarrow \mathcal{W} \subseteq \mathcal{U}^{\perp} \Rightarrow \mathcal{U} \perp \mathcal{W}$.

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. Please insert your solution here ...

1) The code is in the file : 'GradientDescent.m'

   The keys are computing loss and updating $\mathbf{x}$ to $\mathbf{x_{LS}}$.

   The loss is the distance between label '$\mathbf{y}$' and predict '$\mathbf{Ax}$':

$$loss = ||\mathbf{y} - \mathbf{Ax}||_\mathbf{2}^\mathbf{2}$$

   We need a $\Delta\mathbf{x}$ to update $\mathbf{x}$:

$$\Delta\mathbf{x} = \nabla_\mathbf{x}\mathbf{f}(\mathbf{x}) = \mathbf{2} \times (\mathbf{A^T A x} - \mathbf{A^T y})$$

$$\mathbf{x} = \mathbf{x} - \lambda\mathbf{\Delta x}$$

2) The code is in the file : 'NormalEquation.m'

   Solving LS via the normal equations

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

   The solving process is below:

   compute the lower triangular portion of $\mathbf{C} = \mathbf{A}^T\mathbf{A}$

   form the matrix-vector product $\mathbf{d} = \mathbf{A}^T\mathbf{y}$

   compute the Cholesky factorization $\mathbf{C} = \mathbf{G}^T\mathbf{G}$

   solve $\mathbf{G^T z} = \mathbf{d}$ and $\mathbf{G}\mathbf{x}_{\text{LS}} = \mathbf{z}$

3) a) count flops ($\mathbf{A} \in \mathcal{R}^{\mathbf{m}\times\mathbf{n}}$):

      As for gradient decent:

      set iteration = $epoch$, $N =$ length($\mathbf{y}$), for every iteration:

      compute $loss = ||\mathbf{y} - \mathbf{Ax}||_\mathbf{2}^\mathbf{2} \Rightarrow cost_1 = m \times (n + n - 1) + m + m + m - 1 + 1 = 2m(n + 1)$

      compute $\Delta\mathbf{x} = \mathbf{2} \times (\mathbf{A^T A x} - \mathbf{A^T y}) \Rightarrow cost_2 = n^2(2m-1) + n(n+n-1) + n = 2mn^2 + 2mn + n^2 - n$

      update $\mathbf{x} = \mathbf{x} - \lambda\mathbf{\Delta x} \Rightarrow cost_3 = n$

      total flops = $epoch \times (cost_1 + cost_2 + cost_3) = epoch \times (2mn^2 + n^2 + 4mn + 2m)$

      As for normal equation:

$$\mathbf{C} = \mathbf{A}^T\mathbf{A} \Rightarrow \text{one output point need m times multiplication and m-1 times addition}$$

$$\Rightarrow cost_1 = n \times n \times (m + m - 1) = n^2(2m - 1) \Rightarrow \mathcal{O}(mn^2)$$

$$\mathbf{d} = \mathbf{A}^T\mathbf{y} \Rightarrow cost_2 = n \times (m + m - 1) = n(2m - 1) \Rightarrow \mathcal{O}(mn)$$

$$
\begin{aligned}
\mathbf{C} = \mathbf{G}^T\mathbf{G} \Rightarrow cost_3 &= 1\cdot n + 3(n-1) + 5(n-2) + \cdots + (2n-3)2 + (2n-1)\cdot 1 \\
&= \sum_i^n (2i-1)(n+1-i) = \sum_i^n [(n+1)(2i-1) - i(2i-1)] \\
&= (n+1)(2\sum_i^n i - \sum_i^n 1) - (2\sum_i^n i^2 - \sum_i^n i) \\
&= (n+1)(2\cdot\frac{n(n+1)}{2} - n) - (2\cdot\frac{n(n+1)(2n+1)}{6} - \frac{n(n+1)}{2}) \\
&= n(n+1)(\frac{2n+1}{6}) \Rightarrow \mathcal{O}\left(n^3/3\right)
\end{aligned}
$$

$$
\mathbf{G^T z} = \mathbf{d} \quad \text{and} \quad \mathbf{G x}_{\text{LS}} = \mathbf{z} \Rightarrow cost_4 = 2\cdot(1+3+5+\cdots+2n-1) = 2\sum_i^n (2i-1) = 2n^2 \Rightarrow \mathcal{O}\left(n^2\right)
$$

total flops = $cost_1 + cost_2 + cost_3 + cost_4 = n^2(2m-1) + n(2m-1) + \frac{n(n+1)(2n+1)}{6} + 2n^2$

complexity: $\mathcal{O}\left(mn^2 + n^3/3\right)$

b) **gradient decent**:

gradient norm: 6.8872e-09

loss:26.6426

normal equation:

gradient norm: 8.2084e-18

loss:26.6426

In the condition of same loss, the gradient norm of normal equation is much smaller.

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 **??**), 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 $\mathbf{a}_1, \ldots, \mathbf{a}_n$.

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

for $i = 2, \ldots, n$ **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: $\mathbf{q}_1, \ldots, \mathbf{q}_n$

---

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 **??**. 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 $\mathbf{a}_1, \ldots, \mathbf{a}_n$.

Initilization:$\mathbf{Q_1 = 0, R_1 = 0, A} = [\mathbf{a_1}, \cdots, \mathbf{a_n}]$

for $i = 1, \cdots, n$

$\quad \mathbf{z} = \mathbf{A}(:, \mathbf{i})$

$\quad \mathbf{R_1}(\mathbf{i}, \mathbf{i}) = ||\mathbf{z}||_\mathbf{2}$

$\quad \mathbf{Q_1}(:, \mathbf{i}) = \mathbf{z}/\mathbf{R_1}(\mathbf{i}, \mathbf{i})$

$\quad \mathbf{R_1}(\mathbf{i}, \mathbf{i+1:n}) = \mathbf{Q_1}(:, \mathbf{i})^\mathbf{T}\mathbf{A}(:, \mathbf{i+1:n})$

$\quad \mathbf{A}(:, \mathbf{i+1:n}) = \mathbf{A}(:, \mathbf{i+1:n}) - \mathbf{Q_1}(:, \mathbf{i})\mathbf{R_1}(\mathbf{i}, \mathbf{i+1:n})$

end

$\mathbf{q_i} = \mathbf{Q_1}(:, \mathbf{i})$

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 $\epsilon = 1\text{e}{-4}$ and $\epsilon = 1\text{e}{-9}$ in sub-problem 2), give the outputs of two algorithms and calculate $\|\mathbf{Q}^T\mathbf{Q} - \mathbf{I}\|_\text{F}$, 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. Please insert your solution here ...

1)

$$\begin{bmatrix}\mathbf{a_1} & \mathbf{a_2} & \mathbf{a_3} & \mathbf{a_4}\end{bmatrix} = \begin{bmatrix}1 & 2 & 3 & 3\\2 & 3 & 4 & 5\\3 & 4 & 5 & 7\\4 & 5 & 6 & 11\end{bmatrix} \Rightarrow \begin{bmatrix}1 & 1 & 1 & 0\\0 & 1 & 2 & 0\\0 & 0 & 0 & 1\\0 & 0 & 0 & 0\end{bmatrix} \Rightarrow \{\mathbf{a_1}, \mathbf{a_2}, \mathbf{a_4}\}\text{are linear independent.}$$

$\mathbf{\hat{q}_1 = a_1 \Rightarrow q_1 = \hat{q}_1/||\hat{q}_1||_2 = \frac{1}{\sqrt{30}}(1, 2, 3, 4)^T}$

$\mathbf{\hat{q}_2 = a_2 - (q_1^Ta_2)q_1 = (2, 3, 4, 5)^T - \frac{40}{30}(1, 2, 3, 4)^T = \frac{1}{3}(2, 1, 0, -1) \Rightarrow q_2 = \frac{1}{\sqrt{6}}(2, 1, 0, -1)^T}$

$\mathbf{\hat{q}_3 = a_4 - (q_1^Ta_4)q_1 - (q_2^Ta_4)q_2 = \frac{1}{5}(2, -1, -4, 3)^T \Rightarrow q_3 = 1/\sqrt{30}(2, -1, -4, 3)^T}$

$$\mathbf{Q} = (\mathbf{q_1}, \mathbf{q_2}, \mathbf{q_3}) = \begin{bmatrix}\frac{\sqrt{30}}{30} & \frac{\sqrt{6}}{3} & \frac{\sqrt{30}}{15}\\ \frac{\sqrt{30}}{15} & \frac{\sqrt{6}}{6} & -\frac{\sqrt{30}}{30}\\ \frac{\sqrt{30}}{10} & 0 & -\frac{2\sqrt{30}}{15}\\ \frac{2\sqrt{30}}{15} & -\frac{\sqrt{6}}{6} & \frac{\sqrt{30}}{10}\end{bmatrix}$$

2) Set the basis of $\mathbb{R}^n$ are $\{\mathbf{a_1}, \mathbf{a_2}, \cdots, \mathbf{a_n}\}$, the classical Gram-Schmidt algorithm is

$$\mathbf{\hat{q}_i} = \mathbf{a_i} - \sum_{\mathbf{j}}^{\mathbf{i-1}} (\mathbf{q_j^T a_i})\mathbf{q_j}, \quad \mathbf{q_i} = \mathbf{\hat{q}_i}/||\mathbf{\hat{q}_i}||_\mathbf{2}$$

In the context of **projection of vector**:

we can rewrite the equation to that

$$\mathbf{\hat{q}_i} = \mathbf{a_i} - \sum_{\mathbf{j}}^{\mathbf{i-1}} \frac{\langle \mathbf{\hat{q}_j}, \mathbf{a_i} \rangle}{||\mathbf{\hat{q}_j}||_\mathbf{2}^\mathbf{2}} \mathbf{\hat{q}_j} = \mathbf{a_i} - \sum_{\mathbf{j}}^{\mathbf{i-1}} \frac{\langle \mathbf{\hat{q}_j}, \mathbf{a_i} \rangle}{\langle \mathbf{\hat{q}_j}, \mathbf{\hat{q}_j} \rangle} \mathbf{\hat{q}_j} = \mathbf{a_i} - \sum_{\mathbf{j}}^{\mathbf{i-1}} \mathrm{proj}_{\mathbf{\hat{q}_j}}(\mathbf{a_i})$$

$\because$ $\mathrm{proj}_\mathbf{b}(\mathbf{a}) = \frac{\langle \mathbf{a}, \mathbf{b} \rangle}{\langle \mathbf{b}, \mathbf{b} \rangle}\mathbf{b}$ is orthogonal projection of vector $\mathbf{a}$ onto vector $\mathbf{b}$

$\therefore$ $\mathbf{\hat{q}_i}$ is the rest vector subtracted by $\{\mathbf{\hat{q}_1}, \mathbf{\hat{q}_2}, \cdots, \mathbf{\hat{q}_{i-1}}\}$. In other words, $\mathbf{\hat{q}_i}$ is orthogonal to the vectors $\{\mathbf{\hat{q}_1}, \mathbf{\hat{q}_2}, \cdots, \mathbf{\hat{q}_{i-1}}\}$.

In the context of **projection onto subspace**:

Set the orthonormal basis matrix $\mathbf{Q} = (\mathbf{q_1}, \mathbf{q_2}, \cdots, \mathbf{q_n})$, so $\mathbf{QQ^T} = \sum_\mathbf{j}^\mathbf{n} \mathbf{q_j q_j^T}$

$$\mathbf{\hat{q}_i} = \mathbf{a_i} - \sum_{\mathbf{j}}^{\mathbf{i-1}} (\mathbf{q_j^T a_i})\mathbf{q_j} = \mathbf{a_i} - \sum_{\mathbf{j}}^{\mathbf{i-1}} \mathbf{q_j q_j^T a_i} = \mathbf{a_i} - \mathbf{QQ^T a_i} = (\mathbf{I} - \mathbf{QQ^T})\mathbf{a_i} = (\mathbf{I} - \mathbf{P})\mathbf{a_i}$$

So we can get that $\mathbf{\hat{q}_i}$ is the orthogonal projection onto $(\mathrm{span}(Q))^\perp$ along $\mathrm{span}(Q)$.

3)

$$(\mathbf{a_1}, \mathbf{a_2}, \mathbf{a_3}) = \mathbf{A} = \mathbf{QR} = \mathbf{Q} \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ 0 & r_{22} & r_{23} \\ 0 & 0 & r_{33} \end{bmatrix} = [\mathbf{q_1}, \mathbf{q_2}, \mathbf{q_3}] \begin{bmatrix} \mathbf{q_1^T a_1} & \mathbf{q_1^T a_2} & \mathbf{q_1^T a_3} \\ 0 & \mathbf{q_2^T a_2} & \mathbf{q_2^T a_3} \\ 0 & 0 & \mathbf{q_3^T a_3} \end{bmatrix}$$

********************The method below is not using** $1 + k\epsilon^2 = 1$ $(k \in \mathbb{N}^+)$******************

Classical GS (could be said that deal with columns):

compute order: $\mathbf{\hat{q}_1} \Rightarrow \mathbf{r_{11}} \Rightarrow \mathbf{q_1} \Rightarrow \mathbf{r_{12}} \Rightarrow \mathbf{\hat{q}_2} \Rightarrow \mathbf{r_{22}} - \mathbf{q_2} \Rightarrow \mathbf{r_{13}}, \mathbf{r_{23}} \Rightarrow \mathbf{\hat{q}_3} \Rightarrow \mathbf{r_{33}} \Rightarrow \mathbf{q_3}$

compute process:

$$\mathbf{\hat{q}_1} = \mathbf{a_1} = \begin{bmatrix} 1 \\ \epsilon \\ \epsilon \end{bmatrix} \Rightarrow \mathbf{q_1} = \frac{\mathbf{\hat{q}_1}}{||\mathbf{\hat{q}_1}||_\mathbf{2}} = \frac{1}{\sqrt{1+2\epsilon^2}} \begin{bmatrix} 1 \\ \epsilon \\ \epsilon \end{bmatrix}$$

$$\mathbf{\hat{q}_2} = \mathbf{a_2} - (\mathbf{q_1^T a_2})\mathbf{q_1} = \begin{bmatrix} 1 \\ \epsilon \\ 0 \end{bmatrix} - \frac{1+\epsilon^2}{1+2\epsilon^2} \begin{bmatrix} 1 \\ \epsilon \\ \epsilon \end{bmatrix} = \begin{bmatrix} \frac{\epsilon^2}{1+2\epsilon^2} \\ \frac{\epsilon^3}{1+2\epsilon^2} \\ -\frac{\epsilon(1+\epsilon^2)}{1+2\epsilon^2} \end{bmatrix} \Rightarrow \begin{bmatrix} \epsilon \\ \epsilon^2 \\ -(1+\epsilon^2) \end{bmatrix} \Rightarrow \mathbf{q_2} = \frac{1}{\sqrt{(1+2\epsilon^2)(1+\epsilon^2)}} \begin{bmatrix} \epsilon \\ \epsilon^2 \\ -(1+\epsilon^2) \end{bmatrix}$$

$$\mathbf{\hat{q}_3} = \mathbf{a_3} - (\mathbf{q_1^T a_3})\mathbf{q_1} - (\mathbf{q_2^T a_3})\mathbf{q_2} = \begin{bmatrix} 1 \\ 0 \\ \epsilon \end{bmatrix} - \frac{1+\epsilon^2}{1+2\epsilon^2} \begin{bmatrix} 1 \\ \epsilon \\ \epsilon \end{bmatrix} - \frac{-\epsilon^3}{(1+2\epsilon^2)(1+\epsilon^2)} \begin{bmatrix} \epsilon \\ \epsilon^2 \\ -(1+\epsilon^2) \end{bmatrix} = \begin{bmatrix} \frac{\epsilon^2}{1+\epsilon^2} \\ -\frac{\epsilon}{1+\epsilon^2} \\ 0 \end{bmatrix} \Rightarrow \begin{bmatrix} \epsilon^2 \\ -\epsilon \\ 0 \end{bmatrix} \Rightarrow$$

$$\mathbf{q_3} = \frac{1}{\sqrt{\epsilon^4+\epsilon^2}} \begin{bmatrix} \epsilon^2 \\ -\epsilon \\ 0 \end{bmatrix}$$

$$\mathbf{Q} = [\mathbf{q_1}, \mathbf{q_2}, \mathbf{q_3}] = \begin{bmatrix} \frac{1}{\sqrt{1+2\epsilon^2}} & \frac{\epsilon}{\sqrt{(1+2\epsilon^2)(1+\epsilon^2)}} & \frac{\epsilon^2}{\sqrt{\epsilon^4+\epsilon^2}} \\ \frac{\epsilon}{\sqrt{1+2\epsilon^2}} & \frac{\epsilon^2}{\sqrt{(1+2\epsilon^2)(1+\epsilon^2)}} & \frac{-\epsilon}{\sqrt{\epsilon^4+\epsilon^2}} \\ \frac{\epsilon}{\sqrt{1+2\epsilon^2}} & \frac{-(1+\epsilon^2)}{\sqrt{(1+2\epsilon^2)(1+\epsilon^2)}} & 0 \end{bmatrix}$$

MGS (could be said that deal with rows):

compute order:$\mathbf{\hat{q}_1} \Rightarrow \mathbf{r_{11}} \Rightarrow \mathbf{q_1} \Rightarrow \mathbf{r_{12}}, \mathbf{r_{13}} \Rightarrow \mathbf{A(1,:)} \Rightarrow \mathbf{\hat{q}_2} \Rightarrow \mathbf{r_{22}} \Rightarrow \mathbf{q_2} \Rightarrow \mathbf{r_{23}} \Rightarrow \mathbf{A(2,:)} \Rightarrow \mathbf{\hat{q}_3} \Rightarrow \mathbf{r_{33}} \Rightarrow \mathbf{q_3}$

compute process:

$$\mathbf{\hat{q}_1} = \mathbf{a_1} \Longrightarrow \mathbf{q_1} = \frac{\mathbf{\hat{q}_1}}{||\mathbf{\hat{q}_1}||_2} = \frac{1}{\sqrt{1+2\epsilon^2}} \begin{bmatrix} 1 \\ \epsilon \\ \epsilon \end{bmatrix}$$

$$r_{12} = \mathbf{q_1^T a_2} = \frac{1}{\sqrt{1+2\epsilon^2}} \begin{bmatrix} 1 & \epsilon & \epsilon \end{bmatrix} \begin{bmatrix} 1 \\ \epsilon \\ 0 \end{bmatrix} = \frac{1+\epsilon^2}{\sqrt{1+2\epsilon^2}}$$

$$r_{13} = \mathbf{q_1^T a_3} = \frac{1}{\sqrt{1+2\epsilon^2}} \begin{bmatrix} 1 & \epsilon & \epsilon \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ \epsilon \end{bmatrix} = \frac{1+\epsilon^2}{\sqrt{1+2\epsilon^2}}$$

$$\mathbf{A}(:, 2:3) = \mathbf{A}(:, 2:3) - (r_{12}\mathbf{q_1}, r_{13}\mathbf{q_1}) = \begin{bmatrix} 1 & 1 \\ \epsilon & 0 \\ 0 & \epsilon \end{bmatrix} - \begin{bmatrix} \frac{1+\epsilon^2}{1+2\epsilon^2} & \frac{1+\epsilon^2}{1+2\epsilon^2} \\ \frac{1+\epsilon^2}{1+2\epsilon^2}\epsilon & \frac{1+\epsilon^2}{1+2\epsilon^2}\epsilon \\ \frac{1+\epsilon^2}{1+2\epsilon^2}\epsilon & \frac{1+\epsilon^2}{1+2\epsilon^2}\epsilon \end{bmatrix} = \begin{bmatrix} \frac{\epsilon^2}{1+2\epsilon^2} & \frac{\epsilon^2}{1+2\epsilon^2} \\ \frac{\epsilon^3}{1+2\epsilon^2} & -\frac{1+\epsilon^2}{1+2\epsilon^2} \\ -\frac{1+\epsilon^2}{1+2\epsilon^2} & \frac{\epsilon^3}{1+2\epsilon^2} \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} \epsilon^2 & \epsilon^2 \\ \epsilon^3 & -(1+\epsilon^2) \\ -(1+\epsilon^2) & \epsilon^3 \end{bmatrix} = (\mathbf{a_2}, \mathbf{a_3})$$

$$\mathbf{\hat{q}_2} = \mathbf{a_2} \Longrightarrow \mathbf{q_2} = \frac{\mathbf{\hat{q}_2}}{||\mathbf{\hat{q}_2}||_2} = \frac{1}{\sqrt{\epsilon^6+\epsilon^4+(1+\epsilon^2)^2}} \begin{bmatrix} \epsilon^2 \\ \epsilon^3 \\ -(1+\epsilon^2) \end{bmatrix}$$

$$r_{23} = \mathbf{q_2^T a_3} = \frac{1}{\sqrt{\epsilon^6+\epsilon^4+(1+\epsilon^2)^2}} \begin{bmatrix} \epsilon^2 & \epsilon^3 & -(1+\epsilon^2) \end{bmatrix} \begin{bmatrix} \epsilon^2 \\ -(1+\epsilon^2) \\ \epsilon^3 \end{bmatrix} = \frac{\epsilon^4 - 2\epsilon^3(1+\epsilon^2)}{\sqrt{\epsilon^6+\epsilon^4+(1+\epsilon^2)^2}}$$

$$\mathbf{A}(:,3) = \mathbf{a_3} = \mathbf{a_3} - \mathbf{q_2} r_{23} = \begin{bmatrix} \epsilon^2 \\ -(1+\epsilon^2) \\ \epsilon^3 \end{bmatrix} - \frac{\epsilon^4 - 2\epsilon^3(1+\epsilon^2)}{\epsilon^6+\epsilon^4+(1+\epsilon^2)^2} \begin{bmatrix} \epsilon^2 \\ \epsilon^3 \\ -(1+\epsilon^2) \end{bmatrix} = \begin{bmatrix} \frac{\epsilon^2(\epsilon^6+2\epsilon^5+\epsilon^4+2\epsilon^3+2\epsilon^2+1)}{\epsilon^6+\epsilon^4+(1+\epsilon^2)^2} \\ \frac{(1+\epsilon^2)(\epsilon^6-2\epsilon^4-2\epsilon^2-1)-\epsilon^7}{\epsilon^6+\epsilon^4+(1+\epsilon^2)^2} \\ \frac{(1+\epsilon^2)(\epsilon^7-\epsilon^5+\epsilon^4-\epsilon^3)}{\epsilon^6+\epsilon^4+(1+\epsilon^2)^2} \end{bmatrix}$$

$$\mathbf{\hat{q}_3} = \mathbf{a_3} \Rightarrow \mathbf{q_3} = \frac{\mathbf{\hat{q}_3}}{||\mathbf{\hat{q}_3}||_2} \Rightarrow$$

$$\frac{1}{\sqrt{(\epsilon^2(\epsilon^6+2\epsilon^5+\epsilon^4+2\epsilon^3+2\epsilon^2+1))^2+((1+\epsilon^2)(\epsilon^6-2\epsilon^4-2\epsilon^2-1)-\epsilon^7)^2+((1+\epsilon^2)(\epsilon^7-\epsilon^5+\epsilon^4-\epsilon^3))^2}} \begin{bmatrix} \epsilon^2(\epsilon^6+2\epsilon^5+\epsilon^4+2\epsilon^3+2\epsilon^2+1) \\ (1+\epsilon^2)(\epsilon^6-2\epsilon^4-2\epsilon^2-1)-\epsilon^7 \\ (1+\epsilon^2)(\epsilon^7-\epsilon^5+\epsilon^4-\epsilon^3) \end{bmatrix}$$

The two results above are obviously different. I find that if ϵ is small even nearly equal to zero , the classical GS algorithm will be more unstable due to computer rounding errors.During classical GS, denominator is more likely to be zero.

********************The method below is using** $1 + k\epsilon^2 = 1$ $(k \in \mathbb{N}^+)$******************

Classical GS (could be said that deal with columns):

compute order: $\mathbf{\hat{q}_1} \Rightarrow \mathbf{r_{11}} \Rightarrow \mathbf{q_1} \Rightarrow \mathbf{r_{12}} \Rightarrow \mathbf{\hat{q}_2} \Rightarrow \mathbf{r_{22}} - \mathbf{q_2} \Rightarrow \mathbf{r_{13}}, \mathbf{r_{23}} \Rightarrow \mathbf{\hat{q}_3} \Rightarrow \mathbf{r_{33}} \Rightarrow \mathbf{q_3}$

compute process:

$$\mathbf{\hat{q}_1} = \mathbf{a_1} = \begin{bmatrix} 1 \\ \epsilon \\ \epsilon \end{bmatrix} \Rightarrow \mathbf{q_1} = \frac{\mathbf{\hat{q}_1}}{||\mathbf{\hat{q}_1}||_2} = \frac{1}{\sqrt{1+2\epsilon^2}} \begin{bmatrix} 1 \\ \epsilon \\ \epsilon \end{bmatrix} = \begin{bmatrix} 1 \\ \epsilon \\ \epsilon \end{bmatrix}$$

$$\mathbf{\hat{q}_2} = \mathbf{a_2} - (\mathbf{q_1^T a_2})\mathbf{q_1} = \begin{bmatrix} 1 \\ \epsilon \\ 0 \end{bmatrix} - \frac{1+\epsilon^2}{1+2\epsilon^2} \begin{bmatrix} 1 \\ \epsilon \\ \epsilon \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -\epsilon \end{bmatrix} \Rightarrow \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}$$

$$\mathbf{\hat{q}_3} = \mathbf{a_3} - (\mathbf{q_1^T a_3})\mathbf{q_1} - (\mathbf{q_2^T a_3})\mathbf{q_2} = \begin{bmatrix} 1 \\ 0 \\ \epsilon \end{bmatrix} - \frac{1+\epsilon^2}{1+2\epsilon^2} \begin{bmatrix} 1 \\ \epsilon \\ \epsilon \end{bmatrix} - (-\epsilon) \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ -\epsilon \\ -\epsilon \end{bmatrix} \Rightarrow \mathbf{q_3} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ -1 \\ -1 \end{bmatrix}$$

$$\mathbf{Q} = [\mathbf{q_1}, \mathbf{q_2}, \mathbf{q_3}] = \begin{bmatrix} 1 & 0 & 0 \\ \epsilon & 0 & -\frac{1}{\sqrt{2}} \\ \epsilon & -1 & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

MGS (could be said that deal with rows):

compute order:$\mathbf{\hat{q}_1} \Rightarrow \mathbf{r_{11}} \Rightarrow \mathbf{q_1} \Rightarrow \mathbf{r_{12}}, \mathbf{r_{13}} \Rightarrow \mathbf{A(1,:)} \Rightarrow \mathbf{\hat{q}_2} \Rightarrow \mathbf{r_{22}} \Rightarrow \mathbf{q_2} \Rightarrow \mathbf{r_{23}} \Rightarrow \mathbf{A(2,:)} \Rightarrow \mathbf{\hat{q}_3} \Rightarrow \mathbf{r_{33}} \Rightarrow \mathbf{q_3}$

compute process:

$$\mathbf{\hat{q}_1} = \mathbf{a_1} \Longrightarrow \mathbf{q_1} = \frac{\mathbf{\hat{q}_1}}{||\mathbf{\hat{q}_1}||_2} = \begin{bmatrix} 1 \\ \epsilon \\ \epsilon \end{bmatrix}, \quad r_{12} = \mathbf{q_1^T a_2} = \frac{1}{\sqrt{1+2\epsilon^2}} \begin{bmatrix} 1 & \epsilon & \epsilon \end{bmatrix} \begin{bmatrix} 1 \\ \epsilon \\ 0 \end{bmatrix} = 1, \quad r_{13} = \mathbf{q_1^T a_3} = 1$$

$$\mathbf{A}(:,2:3) = \mathbf{A}(:,2:3) - (r_{12}\mathbf{q_1}, r_{13}\mathbf{q_1}) = \begin{bmatrix} 1 & 1 \\ \epsilon & 0 \\ 0 & \epsilon \end{bmatrix} - \begin{bmatrix} 1 & 1 \\ \epsilon & \epsilon \\ \epsilon & \epsilon \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & -\epsilon \\ -\epsilon & 0 \end{bmatrix} = (\mathbf{a_2}, \mathbf{a_3})$$

$$\mathbf{\hat{q}_2} = \mathbf{a_2} = \begin{bmatrix} 0 \\ 0 \\ -\epsilon \end{bmatrix} \Rightarrow \mathbf{q_2} = \frac{\mathbf{\hat{q}_2}}{||\mathbf{\hat{q}_2}||_2} = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}, \quad r_{23} = \mathbf{q_2^T a_3} = \begin{bmatrix} 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ -\epsilon \\ 0 \end{bmatrix} = 0$$

$$\mathbf{A}(:,3) = \mathbf{a_3} = \mathbf{a_3} - \mathbf{q_2} r_{23} = \begin{bmatrix} 0 \\ -\epsilon \\ 0 \end{bmatrix} - \mathbf{0} = \begin{bmatrix} 0 \\ -\epsilon \\ 0 \end{bmatrix}, \quad \mathbf{\hat{q}_3} = \mathbf{a_3} \Rightarrow \mathbf{q_3} = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}$$

$$\mathbf{Q} = [\mathbf{q_1}, \mathbf{q_2}, \mathbf{q_3}] = \begin{bmatrix} 1 & 0 & 0 \\ \epsilon & 0 & -1 \\ \epsilon & -1 & 0 \end{bmatrix}$$

4) **Classical GS**:

$$\epsilon = 1e-4 \Rightarrow \mathbf{Q} = \begin{bmatrix} 1 & 0.0001 & 0.0001 \\ 0.0001 & 0 & -1 \\ 0.0001 & -1 & 0 \end{bmatrix}, \|\mathbf{Q^T Q} - \mathbf{I}\|_{\mathrm{F}} = 3.9885\text{e-}9$$

$$\epsilon = 1e-9 \Rightarrow \mathbf{Q} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -0.7071 \\ 0 & -1 & -0.7071 \end{bmatrix}, \|\mathbf{Q^T Q} - \mathbf{I}\|_{\mathrm{F}} = 1$$

MGS:

$$\epsilon = 1e-4 \Rightarrow \mathbf{Q} = \begin{bmatrix} 1 & 0.0001 & 0.0001 \\ 0.0001 & 0 & -1 \\ 0.0001 & -1 & 0 \end{bmatrix}, \|\mathbf{Q^T Q} - \mathbf{I}\|_{\mathrm{F}} = 5.6406\text{e-}13$$

$$\epsilon = 1e-9 \Rightarrow \mathbf{Q} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{bmatrix}, \|\mathbf{Q^T Q} - \mathbf{I}\|_{\mathrm{F}} = 2\text{e-}9$$

The result can support my view in (3).

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 (**??**) 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

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

and give it a plain interpretation.

3) Computing the solutions to the two LS problems via the normal equation $\mathbf{A}^T\mathbf{A}\mathbf{x}_{LS}=\mathbf{A}^T\mathbf{b}$ 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 (**??**) 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)

$$\|\mathbf{Q}\|_2^2 = 1 \Rightarrow \left\|\mathbf{Q}^T\mathbf{z}\right\|_2 = \|\mathbf{z}\|_2$$

$$\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 = \left\|\mathbf{Q}^T\mathbf{y} - \mathbf{Q}^T\mathbf{A}\mathbf{x}\right\|_2^2 = \left\|\mathbf{Q}^T\mathbf{y} - \mathbf{R}\mathbf{x}\right\|_2^2$$

$$= \left\|\begin{bmatrix} \mathbf{Q}_1^T\mathbf{y} \\ \mathbf{Q}_2^T\mathbf{y} \end{bmatrix} - \begin{bmatrix} \mathbf{R}_1\mathbf{x} \\ \mathbf{0} \end{bmatrix}\right\|_2^2 = \left\|\mathbf{Q}_1^T\mathbf{y} - \mathbf{R}_1\mathbf{x}\right\|_2^2 + \left\|\mathbf{Q}_2^T\mathbf{y}\right\|_2^2$$

Then LS problem $\min_{\mathrm{x}\in\mathbb{R}^n} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2$ reduces to solve $\mathbf{R}_1\mathbf{x} = \mathbf{Q}_1^T\mathbf{b}$

$$\mathbf{A} = (\mathbf{a_1}, \mathbf{a_2}, \mathbf{a_3}) = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 7 \\ 4 & 5 & 11 \end{bmatrix} = \mathbf{Q_1 R_1} = \begin{bmatrix} \mathbf{q_1} & \mathbf{q_2} & \mathbf{q_3} \end{bmatrix} \begin{bmatrix} r_{11} & \mathbf{q_1^T a_2} & \mathbf{q_1^T a_3} \\ 0 & r_{22} & \mathbf{q_2^T a_3} \\ 0 & 0 & r_{33} \end{bmatrix}$$

$r_{11} = \|\mathbf{a_1}\|_2 = \sqrt{30}, \mathbf{q_1} = \frac{1}{\sqrt{30}}(1, 2, 3, 4)^T$,

$r_{12} = \mathbf{q_1^T a_2} = \frac{40}{\sqrt{30}}, \hat{\mathbf{q}}_\mathbf{2} = \mathbf{a_2} - (\mathbf{q_1^T a_2})\mathbf{q_1} = \frac{1}{3}(2, 1, 0, -1)^T, r_{22} = \frac{\sqrt{6}}{3}, \mathbf{q_2} = \frac{\hat{\mathbf{q}}_\mathbf{2}}{r_{22}} = \frac{1}{\sqrt{6}}(2, 1, 0, -1)^T$,

$r_{13} = \mathbf{q_1^T a_3} = \frac{78}{\sqrt{30}}, r_{23} = \mathbf{q_2^T a_3} = 0, \hat{\mathbf{q}}_\mathbf{3} = \mathbf{a_3} - r_{13}\mathbf{q_1} - r_{23}\mathbf{q_2} = \frac{1}{5}(2, -1, -4, 3)^T, r_{33} = \frac{\sqrt{30}}{5}$,

$\mathbf{q_3} = \frac{1}{\sqrt{30}}(2, -1, -4, 3)^T$.

$$\mathbf{A} = \mathbf{Q_1 R_1} = \begin{bmatrix} \frac{1}{\sqrt{30}} & \frac{2}{\sqrt{6}} & \frac{2}{\sqrt{30}} \\ \frac{2}{\sqrt{30}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{30}} \\ \frac{3}{\sqrt{30}} & 0 & -\frac{4}{\sqrt{30}} \\ \frac{4}{\sqrt{30}} & -\frac{1}{\sqrt{6}} & \frac{3}{\sqrt{30}} \end{bmatrix} \begin{bmatrix} \sqrt{30} & \frac{40}{\sqrt{30}} & \frac{78}{\sqrt{30}} \\ 0 & \frac{\sqrt{6}}{3} & 0 \\ 0 & 0 & \frac{\sqrt{30}}{5} \end{bmatrix}$$

$$\mathbf{R_1 x} = \mathbf{Q_1^T b} \Rightarrow \begin{bmatrix} \sqrt{30} & \frac{40}{\sqrt{30}} & \frac{78}{\sqrt{30}} \\ 0 & \frac{\sqrt{6}}{3} & 0 \\ 0 & 0 & \frac{\sqrt{30}}{5} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{30}} & \frac{2}{\sqrt{30}} & \frac{3}{\sqrt{30}} & \frac{4}{\sqrt{30}} \\ \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & 0 & -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{30}} & -\frac{1}{\sqrt{30}} & -\frac{4}{\sqrt{30}} & \frac{3}{\sqrt{30}} \end{bmatrix} \begin{bmatrix} 1 \\ \frac{3}{2} \\ 3 \\ 6 \end{bmatrix} = \begin{bmatrix} \frac{37}{\sqrt{30}} \\ -\frac{5}{2\sqrt{6}} \\ \frac{13}{2\sqrt{30}} \end{bmatrix}$$

backward substitution $\Rightarrow x_3 = \frac{13}{12} \Rightarrow x_2 = -\frac{5}{4} \Rightarrow x_1 = \frac{1}{12} \Rightarrow \mathbf{x} = \begin{bmatrix} \frac{1}{12} & -\frac{5}{4} & \frac{13}{12} \end{bmatrix}^T$

2) **Firstly:**

$\because \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{A}(\mathbf{x} + \delta\mathbf{x}) = \mathbf{b} + \delta\mathbf{b} \quad \Rightarrow \quad \mathbf{x} = \mathbf{A}^\dagger\mathbf{b}, (\mathbf{x} + \delta\mathbf{x}) = \mathbf{A}^\dagger(\mathbf{b} + \delta\mathbf{b})$

$\therefore \|\delta\mathbf{x}\| = \|(\mathbf{x} + \delta\mathbf{x}) - \mathbf{x}\| = \|\mathbf{A}^\dagger(\mathbf{b} + \delta\mathbf{b} - \mathbf{b})\| \leq \|\mathbf{A}^\dagger\|\|\delta\mathbf{b}\|$

$\because \mathbf{b} = \mathbf{A}\mathbf{x} \quad \Rightarrow \quad \|\mathbf{b}\| \leq \|\mathbf{A}\|\|\mathbf{x}\| \quad \Rightarrow \quad \frac{\mathbf{1}}{\|\mathbf{x}\|} \leq \frac{\|\mathbf{A}\|}{\|\mathbf{b}\|}$

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

Secondly:

$\because \|\delta\mathbf{b}\| = \|(\mathbf{b} + \delta\mathbf{b}) - \mathbf{b}\| = \|\mathbf{A}(\mathbf{x} + \delta\mathbf{x} - \mathbf{x})\| \leq \|\mathbf{A}\|\|\delta\mathbf{x}\|$

$\because \mathbf{x} = \mathbf{A}^\dagger\mathbf{b} \quad \Rightarrow \quad \|\mathbf{x}\| \leq \|\mathbf{A}^\dagger\|\|\mathbf{b}\| \quad \Rightarrow \quad \frac{\mathbf{1}}{\|\mathbf{b}\|} \leq \frac{\|\mathbf{A}^\dagger\|}{\|\mathbf{x}\|}$

$\therefore \frac{\|\delta\mathbf{b}\|}{\|\mathbf{b}\|} \leq \|\mathbf{A}\|\|\delta\mathbf{x}\|\frac{\|\mathbf{A}^\dagger\|}{\|\mathbf{x}\|} \quad \Rightarrow \quad \frac{\mathbf{1}}{\|\mathbf{A}\|\|\mathbf{A}^\dagger\|}\frac{\|\delta\mathbf{b}\|}{\|\mathbf{b}\|} \leq \frac{\|\delta\mathbf{x}\|}{\|\mathbf{x}\|}$

Then we can get: $\quad \frac{\mathbf{1}}{\|\mathbf{A}\|\|\mathbf{A}^\dagger\|}\frac{\|\delta\mathbf{b}\|}{\|\mathbf{b}\|} \leq \frac{\|\delta\mathbf{x}\|}{\|\mathbf{x}\|} \leq \|\mathbf{A}\|\|\mathbf{A}^\dagger\|\frac{\|\delta\mathbf{b}\|}{\|\mathbf{b}\|}$

3) Using Matlab, the process is below:

as for **problem(1)**, $b = (1, \frac{3}{2}, 3, 6)^T$,

$$\mathbf{C} = \mathbf{A^T A} = \begin{bmatrix} 15 & 16 & 15 \\ 16 & 18 & 17 \\ 15 & 17 & 17 \end{bmatrix}, \mathbf{d} = \mathbf{A^T b} = \begin{bmatrix} 19 \\ 20 \\ 14 \end{bmatrix}$$

$$\text{cholesky decomposition:} \quad \mathbf{C} = \mathbf{G^T G} \Rightarrow \mathbf{G} = \begin{bmatrix} \frac{1921}{496} & \frac{1921}{465} & \frac{1921}{496} \\ 0 & \frac{1624}{1681} & \frac{1681}{1624} \\ 0 & 0 & \frac{1404}{1457} \end{bmatrix}$$

$$\text{Forward and backward substitution:} \quad \mathbf{x_{LS}} = \begin{bmatrix} \frac{11}{13} & \frac{67}{13} & -\frac{66}{13} \end{bmatrix}^T$$

as for **problem(2)**, $b = b + \delta b = (\frac{11}{10}, \frac{3}{2}, 3, 6)^T$,

$$\mathbf{C} = \mathbf{A^T A}\text{is above}, \mathbf{d} = \mathbf{A^T b} = (\frac{191}{10}, \frac{101}{5}, \frac{71}{5})^T, \text{cholesky decomposition is above.} \Rightarrow \mathbf{x_{LS}} = \begin{bmatrix} \frac{97}{130} & \frac{683}{130} & -\frac{66}{13} \end{bmatrix}^T$$

4) using matlab, computing: $\mathbf{x_{LS}} = (A^T A)^{-1} A^T b$

as for problem(1): $\mathbf{x_{LS}} = \begin{bmatrix} \frac{15}{8} & -\frac{59}{40} & \frac{5}{8} \end{bmatrix}^T$

as for problem(2): $\mathbf{x_{LS}} = \begin{bmatrix} \frac{21}{10} & -\frac{163}{100} & \frac{13}{20} \end{bmatrix}^T$

5) for $\mathbf{A}$ in (3):

condition number: $\|A\|\|A^\dagger\| = \frac{5610}{421} \approx 13.32$

influence: $\frac{\|\delta \mathbf{x}\|}{\|\mathbf{x}\|} = \frac{125}{6438} \approx 0.0194$

for $\mathbf{A}$ in (4):

condition number: $\|A\|\|A^\dagger\| = \frac{2653}{36} \approx 73.69$

influence: $\frac{\|\delta \mathbf{x}\|}{\|\mathbf{x}\|} = \frac{1231}{11065} \approx 0.1113$

We can get that: if the condition number is large, even a small error in **b** may cause a large error in **x**. The larger the condition number is , the larger error in **x** will be caused by the same $\delta\mathbf{b}$.

V. Underdetermined System

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

Consider the following underdetermined system $\mathbf{A}\mathbf{x} = \mathbf{b}$ 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., $\mathbf{A}^T = \mathbf{Q}\mathbf{R}$ 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)

$$\mathbf{A}^T = \begin{bmatrix} 1 & 0 & 2 \\ 2 & -2 & 5 \\ 2 & 2 & 6 \\ 0 & 1 & 1 \end{bmatrix} = \mathbf{A_1} = (\mathbf{a_1}, \mathbf{a_2}, \mathbf{a_3})$$

$$\mathbf{u_1} = \mathbf{a_1} + \|\mathbf{a_1}\|_\mathbf{2}\mathbf{e_1} = (4,2,2,0)^T, \quad \mathbf{R_1} = \mathbf{I} - \mathbf{2}\frac{\mathbf{u_1}\mathbf{u_1^T}}{\|\mathbf{u_1}\|_\mathbf{2}^\mathbf{2}}, \quad (\mathbf{A_1})_{*\mathbf{j}} = \mathbf{R_1}(\mathbf{A_1})_{*\mathbf{j}}$$

$$(\mathbf{A_1})_{*\mathbf{1}} = (\mathbf{A_1})_{*\mathbf{1}} - 2\frac{\mathbf{u_1^T}(\mathbf{A_1})_{*\mathbf{j}}}{\|\mathbf{u}_1\|_2^2}\mathbf{u_1} = \begin{bmatrix} -3 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad (\mathbf{A_1})_{*\mathbf{2}} = \begin{bmatrix} 0 \\ -2 \\ 2 \\ 1 \end{bmatrix}, \quad (\mathbf{A_1})_{*\mathbf{3}} = \begin{bmatrix} -8 \\ 0 \\ 1 \\ 1 \end{bmatrix} \Rightarrow \mathbf{R_1}\mathbf{A_1} = \begin{bmatrix} -3 & 0 & -8 \\ 0 & -2 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

Set

$$\mathbf{A_2} = \begin{bmatrix} -2 & 0 \\ 2 & 1 \\ 1 & 1 \end{bmatrix}$$

$$\mathbf{u_2} = \begin{bmatrix} 0 \\ \hat{\mathbf{u}}_\mathbf{2} \end{bmatrix}, \hat{\mathbf{u}}_\mathbf{2} = (\mathbf{A_2})_{*\mathbf{1}} + \|(\mathbf{A_2})_{*\mathbf{1}}\|_\mathbf{2}\mathbf{e_1} = (-5,2,1)^T, \mathbf{R_2} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{R}}_\mathbf{2} \end{bmatrix}, \hat{\mathbf{R}}_\mathbf{2} = \mathbf{I} - \mathbf{2}\frac{\hat{\mathbf{u}}_\mathbf{2}\hat{\mathbf{u}}_\mathbf{2}^\mathbf{T}}{\|\hat{\mathbf{u}}_\mathbf{2}\|_\mathbf{2}^\mathbf{2}}$$

The same as the first step, we can get $\hat{\mathbf{R}}_\mathbf{2}\mathbf{A_2} = \begin{bmatrix} 3 & 1 \\ 0 & \frac{3}{5} \\ 0 & \frac{4}{5} \end{bmatrix}$

Set

$$\mathbf{A_3} = \begin{bmatrix} \frac{3}{5} \\ \frac{4}{5} \end{bmatrix}$$

$$\mathbf{u_3} = \begin{bmatrix} 0 \\ 0 \\ \mathbf{\hat{u}_3} \end{bmatrix}, \mathbf{\hat{u}_3} = (\mathbf{A_3})_{*1} + \|(\mathbf{A_3})_{*1}\|_2 \mathbf{e_1} = (\frac{8}{5}, \frac{4}{5})^T, \mathbf{R_3} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{\hat{R}_3} \end{bmatrix}, \mathbf{\hat{R}_3} = \mathbf{I} - \mathbf{2}\frac{\mathbf{\hat{u}_3}\mathbf{\hat{u}_3^T}}{\|\mathbf{\hat{u}_3}\|_2^2}$$

We can get $\mathbf{\hat{R}_3 A_3} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$

Next we will compute $\mathbf{Q}$ and $\mathbf{R}$:

$$\mathbf{R} \text{ is the upper triangle matrix} \Rightarrow \begin{bmatrix} -3 & 0 & -8 \\ 0 & 3 & 1 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\mathbf{R_1} = \begin{bmatrix} -\frac{1}{3} & -\frac{2}{3} & -\frac{2}{3} & 0 \\ -\frac{2}{3} & \frac{2}{3} & -\frac{1}{3} & 0 \\ -\frac{2}{3} & -\frac{1}{3} & \frac{2}{3} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{R_2} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -\frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ 0 & \frac{2}{3} & \frac{11}{15} & -\frac{2}{15} \\ 0 & \frac{1}{3} & -\frac{2}{15} & \frac{14}{15} \end{bmatrix}, \quad \mathbf{R_3} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -\frac{3}{5} & -\frac{4}{5} \\ 0 & 0 & -\frac{4}{5} & \frac{3}{5} \end{bmatrix}$$

$$\text{orthogonal matrix } \mathbf{Q^{-1}} = \mathbf{R_3 R_2 R_1} = \begin{bmatrix} -\frac{1}{3} & -\frac{2}{3} & -\frac{2}{3} & 0 \\ 0 & -\frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & -\frac{1}{3} & 0 & -\frac{2}{3} \\ \frac{2}{3} & 0 & -\frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

So:

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

2)

$$\mathbf{A^T} = \mathbf{QR} = \begin{bmatrix} \mathbf{Q_1} & \mathbf{Q_2} \end{bmatrix} \begin{bmatrix} \mathbf{R_1} \\ \mathbf{0} \end{bmatrix} = \mathbf{Q_1 R_1} + \mathbf{Q_2 0} \Rightarrow \mathbf{Ax} = (\mathbf{Q_1 R_1} + \mathbf{Q_2 0})^{\mathbf{T}}\mathbf{x} = \mathbf{R_1^T Q_1^T x} + \mathbf{0^T Q_2^T x} = \mathbf{b}$$

$$\therefore \mathbf{Q_1^T x} = \mathbf{R_1^{-T} b} \Rightarrow \begin{bmatrix} \mathbf{Q_1^T} \\ \mathbf{Q_2^T} \end{bmatrix} \mathbf{x} = \mathbf{Q^T x} = \begin{bmatrix} \mathbf{R_1^{-T} b} \\ \mathbf{d} \end{bmatrix} \Rightarrow \therefore \mathbf{x} = \mathbf{Q} \begin{bmatrix} \mathbf{R_1^{-T} b} \\ \mathbf{d} \end{bmatrix} \quad (\mathbf{d} \text{ could be set to } \mathbf{0})$$

$$\mathbf{R_1^{-T}} = \begin{bmatrix} -\frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ \frac{8}{3} & \frac{1}{3} & -1 \end{bmatrix} \Rightarrow \text{so: } \mathbf{x} = \begin{bmatrix} -\frac{1}{3} & 0 & \frac{2}{3} & \frac{2}{3} \\ -\frac{2}{3} & -\frac{2}{3} & -\frac{1}{3} & 0 \\ -\frac{2}{3} & \frac{2}{3} & 0 & -\frac{1}{3} \\ 0 & \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \end{bmatrix} \begin{bmatrix} -\frac{1}{3}b_1 \\ \frac{1}{3}b_2 \\ \frac{8}{3}b_1 + \frac{1}{3}b_2 - b_3 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{17}{9}b_1 + \frac{2}{9}b_2 - \frac{2}{3}b_3 \\ -\frac{2}{3}b_1 - \frac{1}{3}b_2 + \frac{1}{3}b_3 \\ \frac{2}{9}b_1 + \frac{2}{9}b_2 \\ -\frac{5}{3}b_1 - \frac{2}{9}b_2 + \frac{2}{3}b_3 \end{bmatrix}$$

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

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

as the set of all solutions to (**??**), 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 (**??**) 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 ...