SI231B - Matrix Computations, Spring 2022-23

Homework Set #2

Prof. Ziping Zhao

Acknowledgements:

1) Deadline: **2023-03-26 23:59:59**
2) Please submit your assignments via Blackboard.
3) You can write your homework using latex/word or you can write in handwriting and submit the scanned pdf.

Problem 1. (20 points)

1) Given matrix $\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$, compute its QR decomposition using Gram-Schmidt Orthogonality.
2) Please solve the least square problem via QR decomposition: $\min ||\mathbf{Ax} - \mathbf{b}||_2$, where $\mathbf{b} = [1, -1, 0, 1]^T$.

Solution

1) $\mathbf{A} = [a_1, a_2, a_3]$.

$$\tilde{q}_1 = a_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}, q_1 = \frac{\tilde{q}_1}{||\tilde{q}_1||} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}$$

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

$$\tilde{q}_3 = a_3 - q_1^T a_3 q_1 - q_2^T a_3 q_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix} - \frac{3}{\sqrt{6}} \begin{bmatrix} \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

Then we have $\mathbf{Q} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & 0 \\ 0 & \frac{2}{\sqrt{6}} & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

$r_{11} = q_1^T a_1$, $r_{12} = q_1^T a_2$, $r_{13} = q_1^T a_3$

$r_{22} = q_2^T a_2$, $r_{23} = q_2^T a_3$

$r_{33} = q_3^T a_3$

$$\mathbf{R} = \begin{bmatrix} \sqrt{2} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & \frac{\sqrt{3}}{\sqrt{2}} & \frac{\sqrt{3}}{\sqrt{2}} \\ 0 & 0 & 1 \end{bmatrix}$$

2) We can solve the original problem by $\mathbf{R}\mathbf{x} = \mathbf{Q}^T\mathbf{b}$, namely

$$\begin{bmatrix} \sqrt{2} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & \frac{\sqrt{3}}{\sqrt{2}} & \frac{\sqrt{3}}{\sqrt{2}} \\ 0 & 0 & 1 \end{bmatrix} \mathbf{x} = \begin{bmatrix} \sqrt{2} \\ 0 \\ 1 \end{bmatrix}.$$

We can easily derive that $\mathbf{x} = [2, -1, 1]^T$.

Problem 2. (20 points)

Consider two full-column rank matrices $\mathbf{A} \in \mathbb{R}^{m\times n_1}$ and $\mathbf{B} \in \mathbb{R}^{m\times n_2}$ with $n_1 < m$ and $n_2 < m$. Suppose $\mathcal{R}(\mathbf{A})^\perp \cap \mathcal{R}(\mathbf{B})^\perp = \{\mathbf{0}\}$. Find a semi-orthogonal matrix $\mathbf{P}$ based on QR decompositions of $\mathbf{A}$ and $\mathbf{B}$, where the columns of $\mathbf{P}$ form an orthonormal basis for $\mathcal{R}(\mathbf{A}) \cap \mathcal{R}(\mathbf{B})$.

(*Hint:* You may use the orthogonal compliment of $\mathcal{R}(\mathbf{A}) \cap \mathcal{R}(\mathbf{B})$ as $(\mathcal{R}(\mathbf{A}) \cap \mathcal{R}(\mathbf{B}))^\perp = \mathcal{R}(\mathbf{A})^\perp + \mathcal{R}(\mathbf{B})^\perp$.)

Solution:

Denote the subspace

$$\mathcal{T} = (\mathcal{R}(\mathbf{A}) \cap \mathcal{R}(\mathbf{B}))^\perp = \mathcal{R}(\mathbf{A})^\perp + \mathcal{R}(\mathbf{B})^\perp.$$

Then $\mathcal{R}(\mathbf{A}) \cap \mathcal{R}(\mathbf{B}) = \mathcal{T}^\perp$. Let the QR decomposition for $\mathbf{A}$ and $\mathbf{B}$ be

$$\mathbf{A} = \mathbf{Q}^{(\mathbf{A})}\mathbf{R}^{(\mathbf{A})} = \begin{bmatrix} \mathbf{Q}_1^{(\mathbf{A})} & \mathbf{Q}_2^{(\mathbf{A})} \end{bmatrix} \begin{bmatrix} \mathbf{R}_1^{(\mathbf{A})} \\ \mathbf{0} \end{bmatrix}, \quad \mathbf{B} = \mathbf{Q}^{(\mathbf{B})}\mathbf{R}^{(\mathbf{B})} = \begin{bmatrix} \mathbf{Q}_1^{(\mathbf{B})} & \mathbf{Q}_2^{(\mathbf{B})} \end{bmatrix} \begin{bmatrix} \mathbf{R}_1^{(\mathbf{B})} \\ \mathbf{0} \end{bmatrix},$$

where $\mathbf{Q}_2^{(\mathbf{A})} \in \mathbb{R}^{m\times(m-n_1)}$ and $\mathbf{Q}_2^{(\mathbf{B})} \in \mathbb{R}^{m\times(m-n_2)}$ are an orthonormal basis for $\mathcal{R}(\mathbf{A})^\perp$ and $\mathcal{R}(\mathbf{B})^\perp$, respectively.

Define $\mathbf{C} = \begin{bmatrix} \mathbf{Q}_2^{(\mathbf{A})} & \mathbf{Q}_2^{(\mathbf{B})} \end{bmatrix} \in \mathbb{R}^{m\times(m-n_1+m-n_2)}$ and we have $\mathcal{T} = \mathcal{R}(\mathbf{C})$. Since

$$\begin{aligned} \dim(\mathcal{T}) &= \dim(\mathcal{R}(\mathbf{A})^\perp + \mathcal{R}(\mathbf{B})^\perp) \\ &= \dim(\mathcal{R}(\mathbf{A})^\perp) + \dim(\mathcal{R}(\mathbf{B})^\perp) - \dim(\mathcal{R}(\mathbf{A})^\perp \cap \mathcal{R}(\mathbf{B})^\perp) \\ &= (m-n_1) + (m-n_2) - 0\,, \end{aligned}$$

columns of $\mathbf{C}$ are linearly independent and constitute a basis for $\mathcal{T}$. Let QR decomposition for $\mathbf{C}$ be

$$\mathbf{C} = \mathbf{Q}^{(\mathbf{C})}\mathbf{R}^{(\mathbf{C})} = \begin{bmatrix} \mathbf{Q}_1^{(\mathbf{C})} & \mathbf{Q}_2^{(\mathbf{C})} \end{bmatrix} \begin{bmatrix} \mathbf{R}_1^{(\mathbf{C})} \\ \mathbf{0} \end{bmatrix},$$

where $\mathbf{Q}_1^{(\mathbf{C})} \in \mathbb{R}^{m\times(2m-n_1-n_2)}$ is an orthonormal basis for $\mathcal{T}$ and $\mathbf{Q}_2^{(\mathbf{C})}$ is an orthonormal basis for $\mathcal{T}^\perp$, which is the desired $\mathbf{P}$.

Problem 3. (20 points)

For $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{b} \in \mathbb{R}^{n}$, $\mathbf{y} \in \mathbb{R}^{m}$, and $\lambda \in \mathbb{R}^{+}$, derive the optimal solution of

$$\underset{\mathbf{x}\in\mathbb{R}^n}{\mathsf{minimize}} \quad \left\|\mathbf{y}-\mathbf{A}\mathbf{x}\right\|_2^2 + \lambda\left\|\mathbf{b}-\mathbf{x}\right\|_2^2.$$

Solution:

Denote $f(\mathbf{x}) = \left\|\mathbf{y}-\mathbf{A}\mathbf{x}\right\|_2^2 + \lambda\left\|\mathbf{b}-\mathbf{x}\right\|_2^2$, its gradient is computed by

$$\nabla f(\mathbf{x}) = 2\mathbf{A}^T(\mathbf{A}\mathbf{x}-\mathbf{y}) + 2\lambda(\mathbf{x}-\mathbf{b}) = 2\Big(\big(\mathbf{A}^T\mathbf{A}+\lambda\big)\mathbf{x} - \mathbf{A}^T\mathbf{y} - \lambda\mathbf{b}\Big).$$

By setting $\nabla f(\mathbf{x}) = 0$, we can get the optimal solution as

$$\mathbf{x}^{\star} = \big(\mathbf{A}^T\mathbf{A}+\lambda\mathbf{I}\big)^{-1}\big(\mathbf{A}^T\mathbf{y}+\lambda\mathbf{b}\big),$$

note that $\mathbf{A}^T\mathbf{A}+\lambda\mathbf{I}$ is always invertible since $\mathbf{A}^T\mathbf{A}$ is positive semidefinite and $\lambda > 0$.

Problem 4. (20 points)

Given

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

Find a point in the column space of $\mathbf{A}$ to make it closest to point $\mathbf{p} = [1, 0, 2]^T$.

Hints: 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 the $\langle \cdot, \cdot \rangle$ is the inner product of vectors.

Solution:

Let $\alpha_1 = [1, -1, 1]^T$, $\alpha_2 = [1, 2, -1]^T$. Apply the Gram-Schmidt algorithm, we have

$$\beta_1 = \alpha_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

$$\beta_2 = \alpha_2 - \frac{\langle \alpha_2, \beta_1 \rangle}{\langle \beta_1, \beta_1 \rangle}\beta_1 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} - \frac{-2}{3}\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{5}{3} \\ \frac{4}{3} \\ -\frac{1}{3} \end{bmatrix}$$

The closest point is the orthogonal projection of $\mathbf{p}$ onto the column space of $\mathbf{A}$,

$$\text{proj}_{\text{span}\{\beta_1, \beta_2\}}\mathbf{p} = \frac{\langle \mathbf{p}, \beta_1 \rangle}{\langle \beta_1, \beta_1 \rangle}\beta_1 + \frac{\langle \mathbf{p}, \beta_2 \rangle}{\langle \beta_2, \beta_2 \rangle}\beta_2 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} + \frac{3}{14}\begin{bmatrix} \frac{5}{3} \\ \frac{4}{3} \\ -\frac{1}{3} \end{bmatrix} = \frac{1}{14}\begin{bmatrix} 19 \\ -10 \\ 13 \end{bmatrix}$$

Problem 5. (20 points)

Given a matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 4 & 4 & -1 \\ 4 & -2 & 2 & 0 \end{bmatrix}$$

1) Use Householder reflection to give the full QR decomposition of $\mathbf{A}^T$, i.e., $\mathbf{A}^T = \mathbf{QR}$ with $\mathbf{Q}$ being a square and orthogonal matrix.
2) Let $\mathbf{b} = [5, 10, 4]^T$, give one possible solution of linear system $\mathbf{Ax} = \mathbf{b}$ via QR decomposition of $\mathbf{A}^T$.
3) Let $\mathbf{c} = [1, 2, 3, 4]^T$ and W be the kernel space of $\mathbf{A}$. Decompose $\mathbf{c}$ with respect to W as $\mathbf{c} = \mathbf{w} + \mathbf{z}$, where $\mathbf{w} \in W, \mathbf{z} \in W^{\perp}$.

Hints: The orthogonal projector onto $\mathcal{R}(\mathbf{A})$ ($\mathbf{A}$ has full column rank) is $\mathbf{AA}^{\dagger}$.

Solution:

1) Following the steps in Householder QR, we have

$$\mathbf{A}^{(0)} = \mathbf{A}^T = \begin{bmatrix} 1 & -1 & 4 \\ 1 & 4 & -2 \\ 1 & 4 & 2 \\ 1 & -1 & 0 \end{bmatrix} = [\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3]$$

Perform the Householder reflection to the first column of $\mathbf{A}^{(0)}$,

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{v}_1 = \mathbf{a}_1 + \|\mathbf{a}_1\|_2 \mathbf{e}_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$$\mathbf{H}_1 = \mathbf{I} - \frac{2\mathbf{v}_1\mathbf{v}_1^T}{\|\mathbf{v}_1\|_2^2} = -\frac{1}{6}\begin{bmatrix} 3 & 3 & 3 & 3 \\ 3 & -5 & 1 & 1 \\ 3 & 1 & -5 & 1 \\ 3 & 1 & 1 & -5 \end{bmatrix}$$

$$\mathbf{A}^{(1)} = \mathbf{H}_1\mathbf{A}^{(0)} = \frac{1}{3}\begin{bmatrix} -6 & -9 & -6 \\ 0 & 10 & -12 \\ 0 & 10 & 0 \\ 0 & -5 & -6 \end{bmatrix}$$

Next, perform Householder reflection to $\mathbf{A}^{(1)}_{2:4,2}$,

$$\tilde{\mathbf{a}}_2 = \frac{1}{3}\begin{bmatrix} 10 \\ 10 \\ -5 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{v}_2 = \tilde{\mathbf{a}}_2 + \|\tilde{\mathbf{a}}_2\|_2 \mathbf{e}_2 = \frac{1}{3}\begin{bmatrix} 25 \\ 10 \\ -5 \end{bmatrix}$$

$$\tilde{\mathbf{H}}_2 = \mathbf{I} - \frac{2\mathbf{v}_2\mathbf{v}_2^T}{\|\mathbf{v}_2\|_2^2} = \frac{1}{15}\begin{bmatrix} -10 & -10 & 5 \\ -10 & 11 & 2 \\ 5 & 2 & 14 \end{bmatrix}, \quad \mathbf{H}_2 = \frac{1}{15}\begin{bmatrix} 15 & 0 & 0 & 0 \\ 0 & -10 & -10 & 5 \\ 0 & -10 & 11 & 2 \\ 0 & 5 & 2 & 14 \end{bmatrix}$$

$$\mathbf{A}^{(2)} = \mathbf{H}_2\mathbf{A}^{(1)} = -\frac{1}{5}\begin{bmatrix} 10 & 15 & 10 \\ 0 & 25 & -10 \\ 0 & 0 & -12 \\ 0 & 0 & 16 \end{bmatrix}$$

Perform Householder reflection to $\mathbf{A}^{(2)}_{3:4,3}$,

$$\tilde{\mathbf{a}}_3 = \frac{1}{5}\begin{bmatrix} 12 \\ -16 \end{bmatrix}, \quad \mathbf{e}_3 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \mathbf{v}_3 = \tilde{\mathbf{a}}_3 + \|\tilde{\mathbf{a}}_3\|_2 \mathbf{e}_3 = \frac{16}{5}\begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

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

$$\mathbf{A}^{(3)} = \mathbf{H}_3\mathbf{A}^{(2)} = \begin{bmatrix} -2 & -3 & -2 \\ 0 & -5 & 2 \\ 0 & 0 & -4 \\ 0 & 0 & 0 \end{bmatrix}$$

Then let $\mathbf{R} = \mathbf{A}^{(3)}$ and

$$\mathbf{Q} = \mathbf{H}_1\mathbf{H}_2\mathbf{H}_3 = \frac{1}{2}\begin{bmatrix} -1 & 1 & -1 & -1 \\ -1 & -1 & 1 & -1 \\ -1 & -1 & -1 & 1 \\ -1 & 1 & 1 & 1 \end{bmatrix}$$

which satisfies $\mathbf{A}^T = \mathbf{QR}$.

2) We can obtain the thin QR decomposition for A^T,

$$\mathbf{A}^T = \mathbf{QR} = \begin{bmatrix} \mathbf{Q}_1 & \mathbf{Q}_2 \end{bmatrix}\begin{bmatrix} \mathbf{R}_1 \\ 0 \end{bmatrix} = \mathbf{Q}_1\mathbf{R}_1 + \mathbf{Q_2 0}$$

which

$$\mathbf{Q}_1 = \frac{1}{2}\begin{bmatrix} -1 & 1 & -1 \\ -1 & -1 & 1 \\ -1 & -1 & -1 \\ -1 & 1 & 1 \end{bmatrix}, \quad \mathbf{R}_1 = \begin{bmatrix} -2 & -3 & -2 \\ 0 & -5 & 2 \\ 0 & 0 & -4 \end{bmatrix}$$

Then, we have,

$$\mathbf{A}\mathbf{x} = \mathbf{R}_1^T\mathbf{Q}_1^T\mathbf{x} + \mathbf{0}^T\mathbf{Q}_2^T\mathbf{x} = \mathbf{b}$$

As we can see, the $\mathbf{Q}_2^T\mathbf{x}$ could be anything. To get the solution with minimum 2-norm, set it to zero. Thus, one possible soluton is

$$\mathbf{x} = \mathbf{Q}\begin{bmatrix} \mathbf{R}_1^{-T}\mathbf{b} \\ \mathbf{0} \end{bmatrix} = \frac{1}{40}\begin{bmatrix} 8 & -2 & 5 \\ 12 & 2 & -5 \\ -4 & 6 & 5 \\ 24 & -6 & -5 \end{bmatrix}\begin{bmatrix} 5 \\ 10 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 \\ \frac{3}{2} \\ \frac{3}{2} \\ 1 \end{bmatrix}$$

3) First, the pseudo-inverse of $\mathbf{A}^T$ is

$$\begin{aligned}
(\mathbf{A}^T)^\dagger &= (\mathbf{A}\mathbf{A}^T)^{-1}\mathbf{A} \\
&= ((\mathbf{Q}_1\mathbf{R}_1)^T(\mathbf{Q}_1\mathbf{R}_1))^{-1}(\mathbf{Q}_1\mathbf{R}_1)^T \\
&= (\mathbf{R}_1^T\mathbf{Q}_1^T\mathbf{Q}_1\mathbf{R}_1)^{-1}\mathbf{R}_1^T\mathbf{Q}_1^T \\
&= (\mathbf{R}_1^T\mathbf{R}_1)^{-1}\mathbf{R}_1^T\mathbf{Q}_1^T \\
&= \mathbf{R}_1^{-1}\mathbf{R}_1^{-T}\mathbf{R}_1^T\mathbf{Q}_1^T \\
&= \mathbf{R}_1^{-1}\mathbf{Q}_1^T
\end{aligned}$$

As we know that $W^\perp = \mathcal{N}(\mathbf{A})^\perp = \mathcal{R}(\mathbf{A}^T)$, and the orthogonal projector of $\mathbf{A}^T$ is

$$\mathbf{P}_{\mathbf{A}^T} = \mathbf{A}^T(\mathbf{A}^T)^\dagger = \mathbf{Q}_1\mathbf{R}_1\mathbf{R}_1^{-1}\mathbf{Q}_1^T = \mathbf{Q}_1\mathbf{Q}_1^T$$

Now, we have

$$\mathbf{z} = \mathbf{P}_{\mathbf{A}^T}\mathbf{c} = \mathbf{Q}_1\mathbf{Q}_1^T\mathbf{c} = \frac{1}{4}\begin{bmatrix} 3 & -1 & 1 & 1 \\ -1 & 3 & 1 & 1 \\ 1 & 1 & 3 & -1 \\ 1 & 1 & -1 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 2 \\ 3 \end{bmatrix}$$

$$\mathbf{w} = \mathbf{c} - \mathbf{z} = \begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \end{bmatrix}$$

which satisfies $\mathbf{c} = \mathbf{w} + \mathbf{z}$.