Student Name: Minh Phan
Course: AMATH 352

Homework 6

---

1 Exercises

1. Question 1

   (a) Based on Lecture 16 on LU factorization, we define $A^{(j)}$ as the original matrix of A,

$$\begin{bmatrix} a_{1:} \\ a_{2:} \\ \vdots \\ a_{n:} \end{bmatrix}$$

with the first j^{th} columns' sub-diagonal entries eliminated to 0, and $E^{(j)} \in \mathbb{R}^{n \times n}$ is defined as

$$E_{ik}^{(j)} = \begin{cases} 1, i = k \\ \frac{-a_{ij}^{(j-1)}}{a_{jj}^{(j-1)}}, i > k, k = j \\ 0, otherwise \end{cases}$$

By this definition, we can see that $E^{(j}$'s matrix shape is an identity matrix $\mathbb{R}^{n \times n}$ with only the sub-diagonal entries at j^{th} column filled with $\frac{-a_{ij}^{j-1}}{a_{jj}^{j-1}}$ From the lecture, we can already see that $E^{(1)}A = A^{(1)}$. Generalize the process for an arbitrary j^{th}:

$$E^{(j)}A^{(j-1)} = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & \frac{-a_{j+1,j}^{(j-1)}}{a_{jj}^{(j-1)}} & 1 & \cdots & 0 \\ 0 & 0 & \cdots & \frac{-a_{j+2,j}^{(j-1)}}{a_{jj}^{(j-1)}} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \frac{-a_{n,j}^{(j-1)}}{a_{jj}^{(j-1)}} & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} a_{11}^{(j-1)} & a_{12}^{(j-1)} & \cdots & a_{1,j}^{(j-1)} & a_{1,j+1}^{(j-1)} & \cdots & a_{1,n}^{(j-1)} \\ 0 & a_{2,1}^{(j-1)} & \cdots & a_{2,j}^{(j-1)} & a_{2,j+1}^{(j-1)} & \cdots & a_{2,n}^{(j-1)} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & a_{j,j}^{(j-1)} & a_{j,j+1}^{(j-1)} & \cdots & a_{j,n}^{(j-1)} \\ 0 & 0 & \cdots & a_{j+1,j}^{(j-1)} & a_{j+1,j+1}^{(j-1)} & \cdots & a_{j+1,n}^{(j-1)} \\ 0 & 0 & \cdots & a_{j+2,j}^{(j-1)} & a_{j+2,j+1}^{(j-1)} & \cdots & a_{j+2,n}^{(j-1)} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & a_{n,j}^{(j-1)} & a_{n,j+1}^{(j-1)} & \cdots & a_{n,n}^{(j-1)} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}^{(j)} & a_{12}^{(j)} & \cdots & a_{1,j}^{(j)} & a_{1,j+1}^{(j)} & \cdots & a_{1,n}^{(j)} \\ 0 & a_{2,1}^{(j)} & \cdots & a_{2,j}^{(j)} & a_{2,j+1}^{(j)} & \cdots & a_{2,n}^{(j-1)} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & a_{j,j}^{(j)} & a_{j,j+1}^{(j)} & \cdots & a_{j,n}^{(j)} \\ 0 & 0 & \cdots & 0 & a_{j+1,j+1}^{(j)} & \cdots & a_{j+1,n}^{(j)} \\ 0 & 0 & \cdots & 0 & a_{j+2,j+1}^{(j)} & \cdots & a_{j+2,n}^{(j)} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & a_{n,j+1}^{(j)} & \cdots & a_{n,n}^{(j)} \end{bmatrix} = A^{(j)}$$

(b) To enhance visibility, let $e_{i,j} = \frac{a_{i,j}^{(j-1)}}{a_{jj}^{(j-1)}}$ with $i = j+1, ..., n$.

Define $i_j = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ where j^{th} entry $= 1$, and $e_j = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ e_{j+1,j} \\ \vdots \\ e_{n,j} \end{bmatrix}$ where non-zero entries start at $j+1$.

We can see that $e_k i_k^T = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & \frac{a_{j+1,j}^{(j-1)}}{a_{jj}^{(j-1)}} & 0 & \cdots & 0 \\ 0 & 0 & \cdots & \frac{a_{j+2,j}^{(j-1)}}{a_{jj}^{j-1}} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \frac{a_{n,j}^{(j-1)}}{a_{jj}^{(j-1)}} & 0 & \cdots & 0 \end{bmatrix}$, and therefore $E^j = I - e_j i_j^T$ and

$$(I - e_j i_j^T)(I + e_j i_j^T) = I - e_j i_j^T e_j i_j^T = I$$

since the inner product $e_j i_j^T = 0$ as the only non-zero entry of i_j does not hit and non-zero entry in e_j, making the whole term $-e_j i_j^T e_j i_j^T = 0$. We know that the product of a matrix with its inverse matrix is the identity matrix I, which is what happened in the equation above. We can therefore conclude that $(I + e_j i_j^T)$ is the inverse matrix of $E^{(j)}$, $(E^{(j)})^{-1}$. Since the only difference between $E^{(j)}$ and $(E^{(j)})^{-1}$ is the minus sign:

$$(E^{(j)})^{-1}_{ik} = \begin{cases} 1, i = k \\ \frac{a_{ij}^{(j-1)}}{a_{jj}^{(j-1)}}, i > k, k = j \\ 0, otherwise \end{cases}$$

(c) For two arbitrary $(E^{(j)})^{-1}$, $(E^{(j+1)})^{-1}$, we have the generalized product:

$$\begin{aligned} (E^{(j)})^{-1}(E^{(j+1)})^{-1} &= (I + e_j i_j^T)(I + e_{j+1} i_{j+1}^T) \\ &= I + e_j i_j^T + e_{j+1} i_{j+1}^T + e_j i_j^T e_{j+1} i_{j+1}^T \\ &= I + e_j i_j^T + e_{j+1} i_{j+1}^T \end{aligned}$$

since $i_j^T e_{j+1} = 0$ (same reasoning on why $e_j i_j^T = 0$)

$$
= \begin{bmatrix}
1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\
0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & \cdots & 1 & 0 & \cdots & 0 \\
0 & 0 & \cdots & \frac{a_{j+1,j}^{(j-1)}}{a_{jj}^{(j-1)}} & 1 & \cdots & 0 \\
0 & 0 & \cdots & \frac{a_{j+2,j}^{(j-1)}}{a_{jj}^{(j-1)}} & \frac{a_{j+2,j+1}^{(j)}}{a_{j+1,j+1}^{(j)}} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & \cdots & \frac{a_{n,j}^{(j-1)}}{a_{jj}^{(j-1)}} & \frac{a_{n,j+1}^{(j)}}{a_{j+1,j+1}^{(j)}} & \cdots & 1
\end{bmatrix}
$$

Just like that, repeat the process and multiply the terms up together:

$$
\begin{aligned}
L &= (E^{(1)})^{-1}(E^{(2)})^{-1}\cdots(E^{(n-2)})^{-1}(E^{(n-1)})^{-1} \\
&= I + e_1 i_1^T + e_2 i_2^T + ... + e_{n-2} i_{n-2}^T + e_{n-1} i_{n-1}^T \\
&= \begin{bmatrix}
1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\
e_{2,1} & 1 & \cdots & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
e_{j,1} & e_{j,2} & \cdots & 1 & 0 & \cdots & 0 \\
e_{j+1,1} & e_{j+1,2} & \cdots & e_{j+1,j} & 1 & \cdots & 0 \\
e_{j+2,1} & e_{j+2,2} & \cdots & e_{j+2,j} & e_{j+2,j+1} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
e_{n,1} & e_{n,2} & \cdots & e_{n,j} & e_{n,j+1} & \cdots & 1
\end{bmatrix}
\end{aligned}
$$

(any inner product in this process will reduce to 0).
The result is indeed a lower triangular matrix.

2. Question 2

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

Starting with the identity matrix $I^{5\times5}$ and fill in values derived from first column: $e_{(1,2)} = \frac{a_{2,1}^{(0)}}{a_{1,1}^{(0)}} = -1/2$

$$
E^{(1,2)}A^{(0)} = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 \\
1/2 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
2 & -1 & 0 & 0 & 0 \\
-1 & 2 & -1 & 0 & 0 \\
0 & -1 & 2 & -1 & 0 \\
0 & 0 & -1 & 2 & -1 \\
0 & 0 & 0 & -1 & 2
\end{bmatrix}
= \begin{bmatrix}
2 & -1 & 0 & 0 & 0 \\
0 & 3/2 & -1 & 0 & 0 \\
0 & -1 & 2 & -1 & 0 \\
0 & 0 & -1 & 2 & -1 \\
0 & 0 & 0 & -1 & 2
\end{bmatrix}
$$

Since there are no entries to clear on the rest of the 1st column, $E^{(1)} = E^{(1,2)}$, and $A^{(1)}$ is the product of the above equation.

To column 2, we need to clear the sub-diagonal entries. $e_{(3,2)} = \frac{a_{3,2}^{(1)}}{a_{2,2}^{(1)}} = -2/3$

$$E^{(2,3)}A^{(1)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 2/3 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 & 0 & 0 & 0 \\ 0 & 3/2 & -1 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & -1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & -1 & 0 & 0 & 0 \\ 0 & 3/2 & -1 & 0 & 0 \\ 0 & 0 & 4/3 & -1 & 0 \\ 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & -1 & 2 \end{bmatrix}$$

Since there are no entries to clear on the rest of the 2nd column, $E^{(2)} = E^{(2,3)}$, and $A^{(2)}$ is the product of the above equation.

To column 3, we need to clear the sub-diagonal entries. $e_{(4,3)} = \frac{a^{(2)}_{4,3}}{a^{(2)}_{3,3}} = -3/4$

$$E^{(3,4)}A^{(2)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 3/4 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 & 0 & 0 & 0 \\ 0 & 3/2 & -1 & 0 & 0 \\ 0 & 0 & 4/3 & -1 & 0 \\ 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & -1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & -1 & 0 & 0 & 0 \\ 0 & 3/2 & -1 & 0 & 0 \\ 0 & 0 & 4/3 & -1 & 0 \\ 0 & 0 & 0 & 5/4 & -1 \\ 0 & 0 & 0 & -1 & 2 \end{bmatrix}$$

Since there are no entries to clear on the rest of the 3rd column, $E^{(3)} = E^{(3,4)}$, and $A^{(3)}$ is the product of the above equation.

To column 4, we need to clear the sub-diagonal entries. $e_{(5,4)} = \frac{a^{(3)}_{5,4}}{a^{(3)}_{4,4}} = -4/5$

$$E^{(4,5)}A^{(3)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 4/5 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 & 0 & 0 & 0 \\ 0 & 3/2 & -1 & 0 & 0 \\ 0 & 0 & 4/3 & -1 & 0 \\ 0 & 0 & 0 & 5/4 & -1 \\ 0 & 0 & 0 & -1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & -1 & 0 & 0 & 0 \\ 0 & 3/2 & -1 & 0 & 0 \\ 0 & 0 & 4/3 & -1 & 0 \\ 0 & 0 & 0 & 5/4 & -1 \\ 0 & 0 & 0 & 0 & 6/5 \end{bmatrix}$$

Since there are no entries to clear on the rest of the 3rd column, $E^{(4)} = E^{(4,5)}$, and $A^{(4)}$ is the product of the above equation.

Since $A^{(4)}$ is an upper triangular matrix, it is U. Fill in $e_{(2,1)}, e_{(3,2)}, e_{(4,3)}, e_{(5,4)}$ that we have into the lower triangular matrix model in question 1c:

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ -1/2 & 1 & 0 & 0 & 0 \\ 0 & -2/3 & 1 & 0 & 0 \\ 0 & 0 & -3/4 & 1 & 0 \\ 0 & 0 & 0 & -4/5 & 0 \end{bmatrix}, U = \begin{bmatrix} 2 & -1 & 0 & 0 & 0 \\ 0 & 3/2 & -1 & 0 & 0 \\ 0 & 0 & 4/3 & -1 & 0 \\ 0 & 0 & 0 & 5/4 & -1 \\ 0 & 0 & 0 & 0 & 6/5 \end{bmatrix}$$

The sparsity pattern of L is the non-zero entries in the diagonal and directly below it, and zero everywhere else. The sparsity pattern of U is the non-zero entries in the diagonal and directly above it, and zero everywhere else. L and U therefore inherit a similar structure to the original A matrix but cleared with 0 entries depending on which on the matrix shape.

3. Question 3

(a) $A = \begin{bmatrix} 1 & 6 & 14 \\ 1 & -2 & -8 \\ 1 & 6 & 2 \\ 1 & -2 & 4 \end{bmatrix}$

Let $\underline{a}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, $\underline{a}_2 = \begin{bmatrix} 6 \\ -2 \\ 6 \\ -2 \end{bmatrix}$, and $\underline{a}_3 = \begin{bmatrix} 14 \\ -8 \\ 2 \\ 4 \end{bmatrix}$ as the column vectors of A. Orthonormalize $\underline{a}_1$, we have $\underline{v}_1 = \underline{a}_1$, $||\underline{v}_1||_2 = 2$ and $\underline{q}_1 = \frac{\underline{v}_1}{||\underline{v}_1||} = \begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{bmatrix}$

Using Gram-Schmidt process (discussed in class and homework 3), we orthonormalize $\underline{a}_2$:

$$\underline{v}_2 = \underline{a}_2 - proj_{\underline{q}_1}(\underline{a}_2) = \begin{bmatrix} 6 \\ -2 \\ 6 \\ -2 \end{bmatrix} - 4 \begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{bmatrix} = \begin{bmatrix} 4 \\ -4 \\ 4 \\ -4 \end{bmatrix}, ||\underline{v}_2|| = 8, \underline{q}_2 = \frac{\underline{v}_2}{||\underline{v}_2||} = \begin{bmatrix} 1/2 \\ -1/2 \\ 1/2 \\ -1/2 \end{bmatrix}$$

Orthonormalize $\underline{a}_3$:

$$\underline{v}_3 = \underline{a}_3 - proj_{\underline{q}_1}(\underline{a}_3) - proj_{\underline{q}_2}(\underline{a}_3) = \begin{bmatrix} 14 \\ -8 \\ 2 \\ 4 \end{bmatrix} - 6 \begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{bmatrix} - 10 \begin{bmatrix} 1/2 \\ -1/2 \\ 1/2 \\ -1/2 \end{bmatrix} = \begin{bmatrix} 6 \\ -6 \\ -6 \\ 6 \end{bmatrix}, \underline{q}_3 = \frac{\underline{v}_3}{||\underline{v}_3||} = \begin{bmatrix} 1/2 \\ -1/2 \\ -1/2 \\ 1/2 \end{bmatrix}$$

$||\underline{v}_3|| = 12$

Applying the approach in the lecture 17, We have the reduced

$$\hat{Q}\hat{R} = \begin{bmatrix} 1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & -1/2 \\ 1/2 & 1/2 & -1/2 \\ 1/2 & -1/2 & 1/2 \end{bmatrix} \begin{bmatrix} 2 & 4 & 6 \\ 0 & 8 & 10 \\ 0 & 0 & 12 \end{bmatrix}$$

To find the full rank QR factorization, Q must be a square matrix. Let $\underline{q}_4 = \begin{bmatrix} q_{14} \\ q_{24} \\ q_{34} \\ q_{44} \end{bmatrix}$ Since Q must have orthonormal columns:

$$\begin{cases} \underline{q}_1 \cdot \underline{q}_4 = 0 \\ \underline{q}_2 \cdot \underline{q}_4 = 0 \\ \underline{q}_3 \cdot \underline{q}_4 = 0 \\ \underline{q}_4 \cdot \underline{q}_4 = ||\underline{q}_4||_2^2 = 1 \end{cases}$$

$$\underline{q}_1 \cdot \underline{q}_4 = 0 \rightarrow \frac{1}{2}q_{14} + \frac{1}{2}q_{24} + \frac{1}{2}q_{34} + \frac{1}{2}q_{44} = 0 \tag{1}$$

$$\underline{q}_2 \cdot \underline{q}_4 = 0 \rightarrow \frac{1}{2}q_{14} - \frac{1}{2}q_{24} + \frac{1}{2}q_{34} - \frac{1}{2}q_{44} = 0 \tag{2}$$

$$\underline{q}_3 \cdot \underline{q}_4 = 0 \rightarrow \frac{1}{2}q_{14} - \frac{1}{2}q_{24} - \frac{1}{2}q_{34} + \frac{1}{2}q_{44} = 0 \tag{3}$$

Add (1) and (3) together we have:

$$q_{14} + q_{44} = 0 \tag{4}$$

Add (1) and (2) together we have:

$$q_{14} + q_{34} = 0 \tag{5}$$

Using transitive property on (4) and (5),

$$q_{44} = q_{34} = -q_{14} \tag{6}$$

Apply (6) on (1), we have: $q_{14} = q_{24}$ Let $q_{44} = t$ as a dummy variable. Plug the found values into the $\underline{q}_4$, we have:

$$\underline{q}_4 = t \begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \end{bmatrix}$$

Let $\underline{v}_4$ be the vector before normalized by a factor t Because $\underline{q}_4$ also has to be normalized, we have $||\underline{v}_4||_2 = 2 \rightarrow t = \frac{1}{||\underline{v}_4||_2} = \frac{1}{2}$

Hence $\underline{q}_4 = \frac{1}{2} \begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1/2 \\ -1/2 \\ 1/2 \\ 1/2 \end{bmatrix}$. The full QR factorization for A is:

$$QR = \begin{bmatrix} 1/2 & 1/2 & 1/2 & -1/2 \\ 1/2 & -1/2 & -1/2 & -1/2 \\ 1/2 & 1/2 & -1/2 & 1/2 \\ 1/2 & -1/2 & 1/2 & 1/2 \end{bmatrix} \begin{bmatrix} 2 & 4 & 6 \\ 0 & 8 & 10 \\ 0 & 0 & 12 \\ 0 & 0 & 0 \end{bmatrix}$$

(b) $B = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$

Let $\underline{b}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $\underline{a}_2 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$, and $\underline{a}_3 = \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix}$ as the column vectors of B. Orthonormalize $\underline{b}_1$, we have $\underline{v}_1 = \underline{b}_1$, $||\underline{v}_1|| = \sqrt{14}$ and $q_1 = \frac{\underline{v}_1}{||\underline{v}_1||} = \begin{bmatrix} 1/\sqrt{14} \\ 2/\sqrt{14} \\ 3/\sqrt{14} \end{bmatrix}$ Perform Gram-Schmidt process on $\underline{b}_2, \underline{b}_3$.

Orthonormalize $\underline{b}_2$:

$$\underline{v}_2 = \underline{b}_2 - proj_{\underline{q}_1}(\underline{b}_2) = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} - \frac{32}{\sqrt{14}} \begin{bmatrix} 1/\sqrt{14} \\ 2/\sqrt{14} \\ 3/\sqrt{14} \end{bmatrix} = \begin{bmatrix} 12/7 \\ 3/7 \\ -6/7 \end{bmatrix}, ||\underline{v}_2|| = \frac{3\sqrt{21}}{7}, \underline{q}_2 = \frac{\underline{v}_2}{||\underline{v}_2||} = \begin{bmatrix} 4/\sqrt{21} \\ 1/\sqrt{21} \\ -2/\sqrt{21} \end{bmatrix}$$

Orthonormalize $\underline{b}_3$:

$$\underline{v}_3 = \underline{b}_3 - proj_{\underline{q}_1}(\underline{b}_3) - proj_{\underline{q}_2}(\underline{b}_3) = \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix} - \frac{50}{\sqrt{14}} \begin{bmatrix} 1/\sqrt{14} \\ 2/\sqrt{14} \\ 3/\sqrt{14} \end{bmatrix} - \frac{18}{\sqrt{21}} \begin{bmatrix} 4/\sqrt{21} \\ 1/\sqrt{21} \\ -2/\sqrt{21} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Since the third vector is dependent on the first two vectors, it's not orthonormalizable, and matrix Q has only two columns, each column has 3 entries.

$$\hat{Q} = \begin{bmatrix} 1/\sqrt{14} & 4/\sqrt{21} \\ 2/\sqrt{14} & 1/\sqrt{21} \\ 3/\sqrt{14} & -2/\sqrt{21} \end{bmatrix}, \hat{R} = \begin{bmatrix} \sqrt{14} & 32/\sqrt{14} & 50/\sqrt{14} \\ 0 & \frac{3\sqrt{21}}{7} & 18/\sqrt{21} \\ 0 & 0 & 0 \end{bmatrix}$$

Since the number of columns of the matrix Q is less than the number of columns of the initial matrix, this is not a full QR factorization. We need to find another orthonormal vector in Q. Let $\underline{q}_3 = \begin{bmatrix} q_{13} \\ q_{23} \\ q_{23} \end{bmatrix}$. Since Q must have orthonormal columns:

$$\begin{cases} \underline{q}_1 \cdot \underline{q}_3 = 0 \\ \underline{q}_2 \cdot \underline{q}_3 = 0 \\ \underline{q}_3 \cdot \underline{q}_3 = ||\underline{q}_3||_2^2 = 1 \end{cases}$$

$$\underline{q}_1 \cdot \underline{q}_3 = 0 \to \frac{1}{\sqrt{14}}(q_{13}) + \frac{2}{\sqrt{14}}(q_{23}) + \frac{3}{\sqrt{14}}(q_{33}) = 0 \tag{7}$$

$$\underline{q}_2 \cdot \underline{q}_3 = 0 \to \frac{4}{\sqrt{21}}(q_{13}) + \frac{1}{\sqrt{21}}(q_{23}) - \frac{2}{\sqrt{21}}(q_{33}) = 0 \tag{8}$$

Let $\frac{4\sqrt{6}}{3}(7) - (8)$, we have an equation:

$$\frac{\sqrt{21}}{3}(q_{23}) + \frac{2\sqrt{21}}{3}(q_{33} = 0 \to q_{23} = -2q_{33} \tag{9}$$

Substitute the found relationship between the two variables into equation (7):

$$\frac{1}{\sqrt{14}}(q_{13}) + \frac{2}{\sqrt{14}}(-2q_{33}) + \frac{3}{\sqrt{14}}(q_{33}) = 0 \to \frac{1}{\sqrt{14}}(q_{13}) - \frac{1}{\sqrt{14}}(q_{33}) = 0 \to q_{13} = q_{33}$$

Let $q_{33} = t$ as a dummy variable. Plug the found values, we have:

$$\underline{q}_3 = t \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

Let $\underline{v}_3$ be the vector before normalized by a factor t. Because $\underline{q}_3$ has to be normalized, we have $||\underline{v}_3||_2 = \sqrt{6} \to t = \frac{1}{||\underline{v}_3||_2} = \frac{1}{\sqrt{6}}$

Hence $\underline{q}_3 = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{6} \\ -2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}$. The full QR factorization for matrix B is:

$$Q = \begin{bmatrix} 1/\sqrt{14} & 4/\sqrt{21} & 1/\sqrt{6} \\ 2/\sqrt{14} & 1/\sqrt{21} & -2/\sqrt{6} \\ 3/\sqrt{14} & -2/\sqrt{21} & 1/\sqrt{6} \end{bmatrix}, R = \begin{bmatrix} \sqrt{14} & 32/\sqrt{14} & 50/\sqrt{14} \\ 0 & \frac{3\sqrt{21}}{7} & 18/\sqrt{21} \\ 0 & 0 & 0 \end{bmatrix}$$

4. Question 4

- $\hat{Q}\hat{Q}^T$

  Since $\hat{Q}$ is an orthonormal matrix (a non-square matrix with full rank and all the columns are orthonormal), $\hat{Q}^T\hat{Q} = I$ (page 189, section 10.4 in textbook "Introduction to Applied Linear Algebra")

$$(\hat{Q}\hat{Q}^T)(\hat{Q}\hat{Q}^T) = \hat{Q}(\hat{Q}^T\hat{Q})\hat{Q}^T = \hat{Q}\hat{Q}^T \tag{10}$$

$$(\hat{Q}\hat{Q}^T)^T = (\hat{Q}^T)^T(\hat{Q})^T = \hat{Q}\hat{Q}^T \tag{11}$$

  From (10) and (11) and the definition of orthogonal projection in the question prompt, we can conclude that $\hat{Q}\hat{Q}^T$ is an orthogonal projector

- $I - \hat{Q}\hat{Q}^T$

$$(I - \hat{Q}\hat{Q}^T)^T = I^T - (\hat{Q}\hat{Q}^T)^T \qquad \text{from(11)} \qquad = I - \hat{Q}\hat{Q} \tag{12}$$

$$\begin{aligned}(I - \hat{Q}\hat{Q}^T)(I - \hat{Q}\hat{Q}^T) &= II - I\hat{Q}\hat{Q}^T - (\hat{Q}\hat{Q}^T)I + (\hat{Q}\hat{Q}^T)(\hat{Q}\hat{Q}^T) \\ &= I - \hat{Q}\hat{Q}^T - \hat{Q}\hat{Q}^T + \hat{Q}\hat{Q}^T \qquad \text{from (10)} \\ &= I - \hat{Q}\hat{Q}^T \end{aligned} \tag{13}$$

From (12) and (13) and the definition of orthogonal projection in the question prompt, we can conclude that $I - \hat{Q}\hat{Q}^T$ is an orthogonal projector

5. Question 5

Assume that $\underline{x}$ is a column vector

$$\begin{aligned} f(\underline{x}) =& e^{-\frac{1}{2}\underline{x}^T A \underline{x}} \\ =& e^{-\frac{1}{2}\underline{x}^T \sum_{k=1}^{n} a_{jk}x_k} \\ =& e^{-\frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk}x_j x_k} \end{aligned}$$

For a partial derivative of f with respect to x_i in $\underline{x}$:

$$\begin{aligned} \frac{\partial f}{\partial x_i} =& \frac{\partial}{\partial x_i}\left(e^{-\frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk}x_j x_k}\right) \\ =& \frac{\partial}{\partial x_i}\left(-\frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk}x_j x_k\right)\left(e^{-\frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk}x_j x_k}\right) \\ =& -\frac{1}{2}\left[\sum_{j\neq i}\frac{\partial}{\partial x_i}\left(\sum_{k=1}^{n} a_{jk}x_j x_k\right) + \frac{\partial}{\partial x_i}\sum_{k=1}^{n} a_{ik}x_i x_k\right]\left(e^{-\frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk}x_j x_k}\right) \\ =& -\frac{1}{2}\left[\sum_{j\neq i}\left(\frac{\partial}{\partial x_i}\left(\sum_{k\neq i} a_{jk}x_k x_j\right) + \frac{\partial}{\partial x_i}(a_{ji}x_j x_i)\right) + \left(\sum_{k\neq i}\frac{\partial}{\partial x_i}a_{ik}x_i x_k + \frac{\partial}{\partial x_i}a_{ii}x_i^2\right)\right] \\ & \left(e^{-\frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk}x_j x_k}\right) \\ =& -\frac{1}{2}\left(\sum_{j\neq i} a_{ji}x_j + \sum_{k\neq i} a_{ik}x_k + 2a_{ii}x_i\right)\left(e^{-\frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk}x_j x_k}\right) \\ =& -\frac{1}{2}\left(\sum_{j=1}^{n} a_{ji}x_j + \sum_{k=1}^{n} a_{ik}x_k\right)\left(e^{-\frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk}x_j x_k}\right) \\ =& -\frac{1}{2}((\underline{x}^T A)_i + (A\underline{x})_i)e^{-\frac{1}{2}\underline{x}^T A \underline{x}} \end{aligned} \tag{14}$$

From (14), we can see that $(\underline{x}^T A)_i$ returns the i^{th} component of the row vector of $\underline{x}^T A$ while $(A\underline{x})_i$ returns the i^{th} component of the column vector $A\underline{x}$. Therefore $(A\underline{x})_i^T = (\underline{x}^T A^T)_i$ and

$$(14) \rightarrow \frac{\partial f}{\partial x_i} = -\frac{1}{2}((\underline{x}^T (A + A^T)_i)e^{-\frac{1}{2}\underline{x}^T A \underline{x}}$$

We can therefore conclude that $\nabla f = -\frac{1}{2}\underline{x}^T (A + A^T)e^{-\frac{1}{2}\underline{x}^T A \underline{x}}$