Caleb Logemann
MATH 562 Numerical Analysis II
Homework 3

1. Determine the relative condition number for the following problem. Are there values of x for which the problem is ill-conditioned? Justify your answer.

$$f(x) = \frac{1 - e^{-x}}{1 + e^{-x}}$$

Since f is differentiable the relative condition number of f is given by $\kappa = \frac{\lvert f'(x)\rvert}{\lvert f(x)\rvert / \lvert x\rvert}$. For this problem

$$\begin{aligned} f'(x) &= \frac{(1 + e^{-x})e^{-x} - (1 - e^{-x})(-e^{-x})}{(1 + e^{-x})^2} \\ &= \frac{e^{-x} + e^{-2x} + e^{-x} - e^{-2x}}{(1 + e^{-x})^2} \\ &= \frac{2e^{-x}}{(1 + e^{-x})^2} \end{aligned}$$

Thus the relative condition number for this problem is

$$\begin{aligned} \kappa &= \frac{\lvert f'(x)\rvert}{\lvert f(x)\rvert / \lvert x\rvert} \\ &= \left\lvert \frac{2xe^{-x}}{(1 + e^{-x})^2} \Big/ \frac{1 - e^{-x}}{1 + e^{-x}} \right\rvert \\ &= \left\lvert \frac{2xe^{-x}}{(1 + e^{-x})^2} \times \frac{1 + e^{-x}}{1 - e^{-x}} \right\rvert \\ &= \left\lvert \frac{2xe^{-x}}{(1 + e^{-x})} \times \frac{1}{1 - e^{-x}} \right\rvert \\ &= \left\lvert \frac{2xe^{-x}}{(1 - e^{-2x})} \right\rvert \end{aligned}$$

This problem is not ill-conditioned because for any x this relitive condition number is small. At $x = 0$, this condition number is undefined, but L'Hopital's rule shows that the limit is equal to 1.

$$\begin{aligned}\lim_{x\to 0}(\kappa) &= \lim_{x\to 0}\left(\frac{2e^{-x}-2xe^{-x}}{2e^{-2x}}\right)\\ &= \frac{2e^0}{2e^0}\\ &= 1\end{aligned}$$

As $x \to \infty$, $2xe^{-x} \to 0$ and $1 - e^{-2x} \to 1$, therefore $\kappa \to 0$. As $x \to -\infty$, $1 - e^{-2x} > 2xe^{-x}$, so $\kappa \to 0$. In fact $\kappa \leq 1$ for all x, therefore this problem is not ill-conditioned.

2. Determine whether the calculation $f(x, y) = (1 + x)y^2$ is backward stable by the alogirithm

$$\tilde{f}(x, y) = [1 \oplus fl(x)] \otimes [fl(y) \otimes fl(y)]$$

The algorithm $\tilde{f}$ is backward stable if there exists $\tilde{\mathbf{x}} = (\tilde{x}, \tilde{y})$ such that $\tilde{f}(x, y) = f(\tilde{x}, \tilde{y})$ and $\frac{\|\mathbf{x}-\tilde{\mathbf{x}}\|}{\|\mathbf{x}\|} = O(\epsilon_{machine})$ for all $\mathbf{x}$.

$$\begin{aligned}\tilde{f}(x, y) &= [1 \oplus fl(x)] \otimes [fl(y) \otimes fl(y)]\\ &= [1 \oplus x(1+\epsilon_1)] \otimes [y(1+\epsilon_2) \otimes y(1+\epsilon_3)]\\ &= [1 + x(1+\epsilon_1)](1+\epsilon_4) \otimes [y(1+\epsilon_2) \times y(1+\epsilon_3)](1+\epsilon_5)\\ &= [1 + x(1+\epsilon_1)](1+\epsilon_4) \times [y(1+\epsilon_2) \times y(1+\epsilon_3)](1+\epsilon_5)(1+\epsilon_6)\\ &= [1 + x(1+\epsilon_1)]y^2(1+\epsilon_7)\end{aligned}$$

where $\epsilon_7 = O(\epsilon_{machine})$

$$\begin{aligned}\tilde{f}(x, y) &= [1 + x(1+\epsilon_1)]y^2(1+\epsilon_7)\\ &= [1 + x(1+\epsilon_1)]\left(y\sqrt{1+\epsilon_7}\right)^2\\ &= [1 + x(1+\epsilon_1)](y(1+\epsilon_8))^2\\ &= f(x(1+\epsilon_1), y(1+\epsilon_8))\end{aligned}$$

Therefore $\tilde{x} = x(1+\epsilon_1)$ and $\tilde{y} = y(1+\epsilon_8)$. This does satisfy $\frac{\|\mathbf{x}-\tilde{\mathbf{x}}\|}{\|\mathbf{x}\|} = O(\epsilon_{machine})$, because $\|\mathbf{x} - \tilde{\mathbf{x}}\| = \sqrt{\epsilon_1^2 + \epsilon_8^2} = O(\epsilon_{machine})$. So this algorithm is backward stable.

3. (a) Compute the LU factorization $A = LU$, of

$$A = \begin{bmatrix} 1 & 2 & 4\\ 2 & 3 & 4\\ 2 & 5 & 6\end{bmatrix}$$

Use the factorization to solve the system $A\mathbf{x} = \mathbf{b}$ where $\mathbf{b} = [-1, 1, 1]^T$. Following algorithm 20.1, the LU factorization can be found as follows.

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

$k = 1$

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix} U = \begin{bmatrix} 1 & 2 & 4 \\ 0 & -1 & -4 \\ 0 & 1 & -2 \end{bmatrix}$$

$k = 2$

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

Therefore the LU factorization of A is

$$A = LU = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 4 \\ 0 & -1 & -4 \\ 0 & 0 & -6 \end{bmatrix}$$

Now this system can be solved using forward and backward substitution. Initially we will solve the system $L\mathbf{y} = \mathbf{b}$ where $\mathbf{y} = U\mathbf{x}$ by forward substitution.

$$L\mathbf{y} = \mathbf{b}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 3 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 3 \\ 6 \end{bmatrix}$$

Therefore

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 3 \\ 6 \end{bmatrix}$$

Now we can solve the system $U\mathbf{x} = \mathbf{y}$ by backward substitution.

$$U\mathbf{x} = \mathbf{y}$$

$$\begin{bmatrix} 1 & 2 & 4 \\ 0 & -1 & -4 \\ 0 & 0 & -6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 3 \\ 6 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 4 \\ 0 & -1 & -4 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 3 \\ -1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 4 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$

Therefore

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$

This is also the solution to the system $A\mathbf{x} = \mathbf{b}$.

(b) Solve the system $A\mathbf{x} = \mathbf{b}$ by LU factorization with partial pivoting

The LU factorization using partial pivoting can be found by following algorithm 21.1.

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

$k = 1$

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} U = \begin{bmatrix} 2 & 3 & 4 \\ 0 & 1/2 & 2 \\ 0 & 2 & 2 \end{bmatrix} P = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$k = 2$

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1/2 & 1/4 & 1 \end{bmatrix} U = \begin{bmatrix} 2 & 3 & 4 \\ 0 & 2 & 2 \\ 0 & 0 & 3/2 \end{bmatrix} P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

Now that we have found this LU factorization with pivoting we can solve the system $A\mathbf{x} = \mathbf{b}$ or the equivalent system $PA\mathbf{x} = P\mathbf{b}$, using forward and backward substitution.

$$\begin{aligned} A\mathbf{x} &= \mathbf{b} \\ PA\mathbf{x} &= P\mathbf{b} \\ LU\mathbf{x} &= P\mathbf{b} \end{aligned}$$

Let $\mathbf{y} = U\mathbf{x}$ and solve this system with forward substitution.

$$\begin{gathered} L\mathbf{y} = P\mathbf{b} \\ \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1/2 & 1/4 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \\ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1/2 & 1/4 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \\ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -3/2 \end{bmatrix} \end{gathered}$$

Therefore

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -3/2 \end{bmatrix}$$

Now the system $U\mathbf{x} = \mathbf{y}$ can be solved by backward substitution.

$$\begin{gathered} \begin{bmatrix} 2 & 3 & 4 \\ 0 & 2 & 2 \\ 0 & 0 & 3/2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -3/2 \end{bmatrix} \\ \begin{bmatrix} 2 & 3 & 4 \\ 0 & 2 & 2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \\ \begin{bmatrix} 2 & 3 & 4 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \\ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \end{gathered}$$

This is also the solution to $A\mathbf{x} = \mathbf{b}$ and it is equivalent to the solution found in part (a).

4. Let $A \in \mathbb{C}^{m\times m}$ be nonsingular. Show that A has an LU factorization if and only if for each k, such that $1 \leq k \leq m$, the upper left $(k \times k)$ block $A(1:k, 1:k)$ of A is nonsingular. Show that this LU factorization is unique.

   Proof. Let $A \in \mathbb{C}^{m\times m}$ be nonsingular. Suppose that A has an LU factorization. It is known that $\det(A) = \det(L) \times \det(U)$. Since L and U are triangular the deteminants of L and U are the product of the entries along the diagonal. Since the diagonal of L is all ones, $\det(L) = 1$. Therefore $\det(A) = \det(U)$. Since A is nonsingular, $\det(A) \neq 0$. This implies that all of the diagonal entries of U are nonzero. Also note that $A(1:k,1:k) = L(1:k,1:k)U(1:k,1:k)$ for $1 \leq k \leq m$, so $\det(A(1:k,1:k)) = \det(L(1:k,1:k))\det(U(1:k,1:k))$. As before $\det(L(1:k,1:k)) = 1$. Since all of the diagonal entries of U are nonzero $\det(U(1:k,1:k)) \neq 0$. This implies that $\det(A(1:k,1:k)) \neq 0$ and thus $A(1:k,1:k)$ is nonsingular.

   Now suppose that $A(1:k,1:k)$ is nonsingular for $1 \leq k \leq m$. This implies that $A(1,1) \neq 0$, therefore Gaussian elimination can be applied to the first column. Gaussian Elimination row operations do not change the determinant of a matrix. Therefore L_1A still satisfies the property that $(L_1A)(1:k,1:k)$ is nonsingular. We can now conclude that $(L_1A)(2,2) \neq 0$, because $(L_1A)(1:2,1:2)$ is upper triangular and nonsingular. Now Gaussian elimination can be applied to column 2. As is evident mathematical induction now guarantees that an Gaussian elimination will never fail and therefore A has an LU factorization.

   To show that the LU decomposition is unique assume there exists another LU decomposition of A, $\hat{L}\hat{U}$. This implies that $LU = \hat{L}\hat{U}$, which is equivalent to $L^{-1}\hat{L} = U^{-1}\hat{U}$. L^{-1} exists because L has ones on the diagonal and it is lower triangular. U^{-1} exists because A is nonsingular and $\det A = \det U$. Also U^{-1} is upper triangular. This implies that $L^{-1}\hat{L}$ is lower triangular and $U^{-1}\hat{U}$ is upper triangular. Therefore $L^{-1}\hat{L} = U^{-1}\hat{U}$ if and only if $L^{-1}\hat{L} = I$ and $U^{-1}\hat{U} = I$. Therefore $L = \hat{L}$ and $U = \hat{U}$, so the LU decomposition is unique. □

5. Rank Deficient Least Squares Problem: Let $A \in \mathbb{R}^{m\times n}$ with $m \geq n$, and let $r = \text{rank}(A) < n$. The SVD of A can be written as

$$A = [U_1, U_2]\begin{bmatrix}\Sigma_1 & 0\\ 0 & 0\end{bmatrix}[V_1, V_2]^T = U_1\Sigma_1V_1^T$$

   where Σ_1 is $r \times r$ nonsingular and U_1 and V_1 have r columns. Let $\sigma = \sigma_{min}(\Sigma_1)$, be the smallest nonzero singular value of A. Consider the following rank deficient least squares problem, for some $\mathbf{b} \in \mathbb{R}^m$.

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

Show

#1 all solutions $\mathbf{x}$ can be written as $\mathbf{x} = V_1 \Sigma_1^{-1} U_1^T \mathbf{b} + V_2 \mathbf{z}$ where $\mathbf{z}$ is an arbitrary vector.

It is known that the solution to this problem is a vector $\mathbf{x}$ such that $A\mathbf{x}$ is the projection of $\mathbf{b}$ onto the range of A. Using the singular value decomposition the projector onto the range of A is $U_1 U_1^T$. Therefore we need to find solutions to the system $A\mathbf{x} = U_1 U_1^T \mathbf{b}$. I will replace A with its full SVD decomposition.

$$\begin{aligned} A\mathbf{x} &= U_1 U_1^T \mathbf{b} \\ [U_1, U_2] \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} [V_1, V_2]^T \mathbf{x} &= U_1 U_1^T \mathbf{b} \\ \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} [V_1, V_2]^T \mathbf{x} &= [U_1, U_2]^T U_1 U_1^T \mathbf{b} \\ \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} [V_1, V_2]^T \mathbf{x} &= I_{mr} U_1^T \mathbf{b} \end{aligned}$$

Where I_{mr} is the $m \times r$ matrix with ones on the main diagonal.

$$I_{rm} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} [V_1, V_2]^T \mathbf{x} = U_1^T \mathbf{b}$$

If $m > r$ than $I_{rm} I_{mr} = I_{rr}$.

$$\begin{aligned} [\Sigma_1, 0][V_1, V_2]^T \mathbf{x} &= U_1^T \mathbf{b} \\ \left(\Sigma_1 V_1^T + 0 V_2^T\right) \mathbf{x} &= U_1^T \mathbf{b} \\ \left(V_1^T + 0 V_2^T\right) \mathbf{x} &= \Sigma_1^{-1} U_1^T \mathbf{b} \\ \left(I + V_1 0 V_2^T\right) \mathbf{x} &= V_1 \Sigma_1^{-1} U_1^T \mathbf{b} \end{aligned}$$

Any portion $\mathbf{x}$ that is orthogonal to V_2 will be lost as well so an arbitrary linear combination can be added to $\mathbf{x}$.

$$\mathbf{x} = V_1 \Sigma_1^{-1} U_1^T \mathbf{b} + V_2 \mathbf{z}$$

#2 The solution $\mathbf{x}$ has minimal norm $\|\mathbf{x}\|_2$ when $\mathbf{z} = \mathbf{0}$, and in this case $\|\mathbf{x}\|_2 \leq \|\mathbf{b}\|_2 / \sigma$.

From part 1 we can see that

$$\|\mathbf{x}\|_2 = \left\|V_1\Sigma_1^{-1}U_1^T\mathbf{b} + V_2\mathbf{z}\right\|_2$$

This norm is minimized when $\mathbf{z} = \mathbf{0}$

$$\begin{aligned}&= \left\|V_1\Sigma_1^{-1}U_1^T\mathbf{b}\right\|_2 \\ &\leq \left\|V_1\Sigma_1^{-1}U_1^T\right\|_2\|\mathbf{b}\|_2\end{aligned}$$

Note that $V_1\Sigma_1^{-1}U_1^T$ is the SVD of some other matrix, whose singular values are the reciprocals of the singular values of A. Therefore the 2-norm of this matrix is $1/\sigma$.

$$= \frac{\|\mathbf{b}\|_2}{\sigma}$$

Thus $\|\mathbf{x}\|_2 \leq \|\mathbf{b}\|_2/\sigma$ for the case when $\mathbf{z} = \mathbf{0}$.

6. Consider the matrix

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

(a) Using any method you like, determine reduced and full QR factorizations.

I will use classical Gram-Schmidt

$$\begin{aligned}
r_{11} &= \|a_1\| = \sqrt{2} \\
q_1 &= a_1/r_{11} \\
&= \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix} \\
v_2 &= a_2 \\
&= \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} \\
r_{12} &= q_1^* a_2 \\
&= 2/\sqrt{2} = \sqrt{2} \\
v_2 &= v_2 - r_{12} q_1 \\
&= \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \\
r_{22} &= \|v_2\| = \sqrt{3} \\
q_2 &= v_2/r_{22} \\
&= \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ -1/\sqrt{3} \end{bmatrix}
\end{aligned}$$

Therefore the reduced QR factorization of A is

$$\begin{aligned}
A &= \hat{Q}\hat{R} \\
&= \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{3} \\ 0 & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{3} \end{bmatrix} \begin{bmatrix} \sqrt{2} & \sqrt{2} \\ 0 & \sqrt{3} \end{bmatrix}
\end{aligned}$$

The full QR factorization of A can be found by adding a column to $\hat{Q}$ to make it unitary and adding a row of zeroes to $\hat{R}$. The vector $[-1, 2, 1]$ is orthogonal to both q_1 and q_2. Normalizing this vector results in the last column of Q.

Therefore the full QR factorization of A is

$$\begin{aligned} A &= QR \\ &= \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{3} & -1/\sqrt{6} \\ 0 & 1/\sqrt{3} & 2/\sqrt{6} \\ 1/\sqrt{2} & -1/\sqrt{3} & 1/\sqrt{6} \end{bmatrix} \begin{bmatrix} \sqrt{2} & \sqrt{2} \\ 0 & \sqrt{3} \\ 0 & 0 \end{bmatrix} \end{aligned}$$

(b) Use the QR factorization to solve the linear least square problem

$$\min_{\mathbf{x}} \|A\mathbf{x} - \mathbf{b}\|_2^2$$

with $\mathbf{b} = [110]^T$.

The solution to this problem is equivalent to the solution to the following system $\hat{R}\mathbf{x} = \hat{Q}^*\mathbf{b}$. This system can be solved using backsubstitution.

$$\begin{aligned} \begin{bmatrix} \sqrt{2} & \sqrt{2} \\ 0 & \sqrt{3} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= \begin{bmatrix} 1/\sqrt{2} & 0 & 1\sqrt{2} \\ 1/\sqrt{3} & 1/\sqrt{3} & -1/\sqrt{3} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \\ \begin{bmatrix} \sqrt{2} & \sqrt{2} \\ 0 & \sqrt{3} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= \begin{bmatrix} 1/\sqrt{2} \\ 2/\sqrt{3} \end{bmatrix} \\ \begin{bmatrix} \sqrt{2} & \sqrt{2} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= \begin{bmatrix} 1/\sqrt{2} \\ 2/3 \end{bmatrix} \\ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= \begin{bmatrix} -1/6 \\ 2/3 \end{bmatrix} \end{aligned}$$

Therefore the solution is $x = [-1/6, 2/3]^T$.

(c) Use the QR factorization to solve the linear least squares problem

$$\min_{\mathbf{x}} \|A\mathbf{x} - \mathbf{b}\|_2^2$$

with matrix $A \in \mathbb{R}^{m \times n}$ with rank n and $\mathbf{b} \in \mathbb{R}^m$.

It is known that the solution to this problem is the vector $\mathbf{x}$ such that $A\mathbf{x}$ is the projection of $\mathbf{b}$ onto the range of A. Using the reduced QR factorization the orthogonal projector onto A is $\hat{Q}\hat{Q}^*$. Thus the orthogonal projection of $\mathbf{b}$ onto the range of A is $\hat{Q}\hat{Q}^*\mathbf{b}$. The solution is therefore the solution to the system $A\mathbf{x} = \hat{Q}\hat{Q}^*\mathbf{b}$. Replacing A with its reduced QR decomposition and left multiplying by $\hat{Q}^*$ results in the system $\hat{R}\mathbf{x} = \hat{Q}^*\mathbf{b}$. Since $\hat{R}$ is upper triangular this system can be solved using back-substitution. The solution to this system is the solution to the linear least squares problem.

7. Consider the least-square problem $\min_{\mathbf{x}} \|A\mathbf{x} - b\|_2$, where A is the first 5 columns of the 6×6 inverse Hilbert matrix and

$$b = \begin{bmatrix} 463 \\ -13860 \\ 97020 \\ -258720 \\ 291060 \\ -116424 \end{bmatrix}$$

(a) What are the four conditioning numbers (Theorem 18.1) of the problem? The following script computes the 4 condition numbers.

```
inverseHilbert = invhilb(6);
A = inverseHilbert(:, 1:5);
[m, n] = size(A);
b = [463; -13860; 97020; -258720; 291060; -116424];
xExact = [1; 1/2; 1/3; 1/4; 1/5];
y = A*x;
theta = acos(norm(y)/norm(b));
eta = norm(A)*norm(xExact)/norm(y);
kappa = cond(A);

condby = 1/cos(theta)
condbx = kappa/(eta*cos(theta))
condAy = kappa/cos(theta)
condAx = kappa + kappa^2*tan(theta)/eta
```

```
condby =

    1

condbx =

    1.8263e+05

condAy =

    4.6968e+06

condAx =

    4.6968e+06
```

(b) Use all the algorithms on Pages 138-142 to solve the problem.

i. Householder QR
ii. Householder QR of augmented matrix
iii. Modified Gram-Schmidt QR
iv. Modified Gram-Schmidt QR of augmented matrix
v. Normal Equation
vi. SVD

Check the accuracy of computed solutions as compared to actual solution, and comment on the computed solutions and algorithms used.

```
% Householder QR
[Q, R] = qr(A, 0);
x = R\(Q'*b);
E = norm(x - xExact);
M = {'Householder QR'};

% Householder QR of augmented matrix
[Q, R] = qr([A, b], 0);
Qb = R(1:n, n+1);
R = R(1:n, 1:n);
x = R\Qb;
E = [E; norm(x - xExact)];
M = [M, {'Householder QR of augmented matrix'}];

% Modified Gram-Schmidt QR
[Q, R] = mgs(A);
x = R\(Q'*b);
E = [E; norm(x - xExact)];
M = [M, {'Modified Gram-Schmidt QR'}];

% Modified Gram-Schmidt QR of augmented matrix
[Q, R] = mgs([A, b]);
Qb = R(1:n, n+1);
R = R(1:n, 1:n);
x = R\Qb;
E = [E; norm(x - xExact)];
M = [M, {'Modified Gram-Schmidt QR of augmented matrix'}];

% Normal Equation
x = (A'*A)\(A'*b);
E = [E; norm(x - xExact)];
M = [M, {'Normal Equations'}];

% SVD
[U, S, V] = svd(A, 0);
x = V*(S\(U'*b));
```

```
E = [E; norm(x - xExact)];
M = [M, {'SVD'}];

table(E, 'VariableNames', {'Error'}, 'RowNames', M)
```

```
ans =

                                                   Error
                                                 __________

    Householder QR                               8.3668e-11
    Householder QR of augmented matrix           6.0204e-11
    Modified Gram-Schmidt QR                     4.2067e-06
    Modified Gram-Schmidt QR of augmented matrix  1.253e-12
    Normal Equations                             1.0259e-05
    SVD                                          1.2091e-10
```

Note that the error for all of these methods is reasonably small. Some methods do have error signigicantly larger than the other methods. The modified Gram-Schmidt and Normal equations error is several orders of magnitude larger than the rest of the methods. The normal equations have much larger error because this approach is unstable. Also the accuracy of the normal equations approach is governed by κ^2 which will be larger than κ. The Modified Gram-Schmidt method is also unstable, unless the augmented matrix approach is taken. This is the reason that the error for the Modified Gram-Schmidt is large, while the Modified Gram-Schmidt of the augmented matrix produces lower error. The augmented matrix version is stable.