1. Let $A \in \mathbb{R}^{n \times n}$ be nonsingular and let $b \in \mathbb{R}^n$. We shall use $\| \cdot \|$ to denote both the vector and the matrix norm and require that $\| Mv \| \leq \| M \| \| v \|$ for any $M \in \mathbb{R}^{n \times n}$ and $v \in \mathbb{R}^n$.

(a) Show that given any $\hat{x} \in \mathbb{R}^n$, we have
\[
\frac{1}{\kappa(A)} \| A\hat{x} - b \| \leq \| \hat{x} - A^{-1}b \| \leq \kappa(A) \| A\hat{x} - b \|,
\]
where $\kappa(A) = \| A \| \| A^{-1} \|$. Deduce that if $x = A^{-1}b$ and $\hat{b} = A\hat{x} - b$, then
\[
\frac{1}{\kappa(A)} \| \hat{b} \| \leq \| \hat{x} - x \| \leq \kappa(A) \| \hat{b} \|.
\]

(b) Show that if $\delta A \in \mathbb{R}^{n \times n}$ is any matrix satisfying
\[
\frac{1}{\kappa(A)} \| \delta A \| \| A \| < 1, \tag{1.1}
\]
then $A + \delta A$ must be nonsingular. (Hint: If $A + \delta A$ is singular, then there exists nonzero $v$ such that $(A + \delta A)v = 0$).

2. Let $A \in \mathbb{R}^{n \times n}$ be nonsingular and let $0 \neq b \in \mathbb{R}^n$. Let $x = A^{-1}b \in \mathbb{R}^n$. In the following, $\delta A \in \mathbb{R}^{n \times n}$ and $\delta b \in \mathbb{R}^n$ are some arbitrary matrix and vector.

(a) Suppose $(A + \delta A)(x + \delta x) = b$ and $\hat{x} = x + \delta x$. Show that
\[
\frac{\| \delta x \|}{\| \hat{x} \|} \leq \kappa(A) \frac{\| \delta A \|}{\| A \|}. \tag{2.2}
\]

(b) Suppose $(A + \delta A)(x + \delta x) = b$ and $\hat{x} = x + \delta x$ and (1.1) is satisfied. Show that
\[
\frac{\| \delta x \|}{\| x \|} \leq \frac{\kappa(A) \| \delta A \|}{1 - \kappa(A) \| \delta A \|}.
\]

You may like use the following outline:

(i) Show that
\[
\delta x = -A^{-1} \delta A \hat{x}
\]
and so
\[
\| \delta x \| \leq \kappa(A) \frac{\| \delta A \|}{\| A \|} (\| x \| + \| \delta x \|).
\]

(ii) Rewrite this inequality as
\[
\left( 1 - \kappa(A) \frac{\| \delta A \|}{\| A \|} \right) \| \delta x \| \leq \kappa(A) \frac{\| \delta A \|}{\| A \|} \| x \|
\]
and use (1.1).
(c) Suppose $(A + \delta A)\hat{x} = b + \delta b$ where $\hat{b} = b + \delta b \neq 0$ and $\hat{x} = x + \delta x \neq 0$. Show that
\[
\frac{\|\delta x\|}{\|\hat{x}\|} \leq \kappa(A) \left( \frac{\|\delta A\|}{\|A\|} + \frac{\|\delta b\|}{\|b\|} + \frac{\|\delta A\| \cdot \|\delta b\|}{\|A\| \cdot \|b\|} \right).
\] (2.3)

You may like use the following outline:

(i) Show that
\[
\delta x = A^{-1}(\delta b - \delta A\hat{x})
\]
and so
\[
\frac{\|\delta x\|}{\|\hat{x}\|} \leq \kappa(A) \left( \frac{\|\delta A\|}{\|A\|} + \frac{\|\delta b\|}{\|b\|} \right). \tag{2.4}
\]

(ii) Show that
\[
\frac{1}{\|\hat{x}\|} \leq \frac{\|A\| + \|\delta A\|}{\|b\|}. \tag{2.5}
\]

(iii) Combine (2.4) and (2.5) to get (2.3).

(d) Suppose $(A + \delta A)\hat{x} = b + \delta b$ where $\hat{b} = b + \delta b \neq 0$ and $\hat{x} = x + \delta x \neq 0$ and (1.1) is satisfied. Use the same ideas in (b) to deduce that
\[
\frac{\|\delta x\|}{\|x\|} \leq \frac{\kappa(A) \left( \frac{\|\delta A\|}{\|A\|} + \frac{\|\delta b\|}{\|b\|} \right)}{1 - \kappa(A) \frac{\|\delta A\|}{\|A\|}}.
\]

3. Let $A \in \mathbb{R}^{m \times n}$ where $m \geq n$ and $\text{rank}(A) = n$. Suppose GECP is performed on $A$ to get
\[
\Pi_1 A \Pi_2 = LU
\]
where $L \in \mathbb{R}^{m \times n}$ is unit lower triangular, $U \in \mathbb{R}^{n \times n}$ is upper triangular, and $\Pi_1 \in \mathbb{R}^{m \times m}$, $\Pi_2 \in \mathbb{R}^{n \times n}$ are permutation matrices.

(a) Show that $U$ is nonsingular and that $L$ is of the form
\[
L = \begin{bmatrix} L_1 \\ L_2 \end{bmatrix}
\]
where $L_1 \in \mathbb{R}^{n \times n}$ is nonsingular.

(b) We will see how the $LU$ factorization may be used to solve the least squares problem
\[
\min_{x \in \mathbb{R}^n} \|Ax - b\|_2.
\]

(i) Show that the problem may be solved via
\[
U\tilde{x} = y, \quad L^\top Ly = L^\top \tilde{b},
\]
where $\tilde{b} = \Pi_1 b$ and $\tilde{x} = \Pi_2^\top x$.

(ii) Describe how you would compute the solution $y$ in
\[
L^\top Ly = L^\top \tilde{b}.
\]

4. Let $\varepsilon > 0$. Consider the matrix
\[
A = \begin{bmatrix} 1 & 1 \\ 1 & 1 + \varepsilon \\ 1 & 1 - \varepsilon \end{bmatrix}.
\]

(a) Why is it a bad idea to solve the normal equation associated with $A$, i.e.
\[
A^\top A x = A^\top b
\]
when $\varepsilon$ is small?
(b) Show that the $LU$ factorization of $A$ is

$$A = LU = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & \varepsilon \end{bmatrix}.$$ 

(c) Why is it a much better idea to solve the normal equation associated with $L$, i.e.

$$L^\top Ly = L^\top \tilde{b}?$$

This shows that the method in Problem 3 is a more stable method than using the normal equation in (a) directly.

(d) Show that the Moore-Penrose pseudoinverse of $A$ is

$$A^\dagger = \frac{1}{6} \begin{bmatrix} 2 & 2 - 3\varepsilon^{-1} & 2 + 3\varepsilon^{-1} \\ 0 & 3\varepsilon^{-1} & -3\varepsilon^{-1} \end{bmatrix}.$$ 

(e) Describe a method to compute $A^\dagger$ given $L$ and $U$. Verify that your method is correct by checking it against the expression in (d).

5. We will now discuss an alternative method to solve the least squares problem in Problem 3 that is more efficient when $m - n < n$.

(a) Show the least squares problem in Problem 3 is equivalent to

$$\min_{z \in \mathbb{R}^n} \left\| \begin{bmatrix} I_n \\ S \end{bmatrix} z - \tilde{b} \right\|_2$$

where $S = L_2 L_1^{-1}$ and $L_1 y = z$. Here and below, $I_n$ denotes the $n \times n$ identity matrix.

(b) Write

$$\tilde{b} = \begin{bmatrix} \tilde{b}_1 \\ \tilde{b}_2 \end{bmatrix}$$

where $\tilde{b}_1 \in \mathbb{R}^n$ and $\tilde{b}_2 \in \mathbb{R}^{m-n}$. Show that the solution $z$ is given by

$$z = \tilde{b}_1 + S^\top (I_{m-n} + SS^\top)^{-1}(\tilde{b}_2 - S\tilde{b}_1).$$

(c) Explain why when $m - n < n$, the method in (a) is much more efficient than the method in Problem 3. For example, what happens when $m = n + 1$?

6. Let $\mathbf{c} \in \mathbb{R}^n$ and consider the linearly constrained least squares problem

$$\min \| \mathbf{w} \|_2 \quad \text{s.t.} \quad A^\top \mathbf{w} = \mathbf{c}.$$ 

(a) If we write $\tilde{\mathbf{c}} = \Pi_2^\top \mathbf{c}$ and $\tilde{\mathbf{w}} = \Pi_1 \mathbf{w}$, show that

$$\tilde{\mathbf{w}} = L(L^\top L)^{-1} U^{-\top} \tilde{\mathbf{c}}$$

where $U^{-\top} = (U^{-1})^\top = (U^\top)^{-1}$, a standard notation that we will also use below. (Hint: You’d need to use something that you’ve already determined in an earlier part).

(b) Write

$$\tilde{\mathbf{w}} = \begin{bmatrix} \tilde{\mathbf{w}}_1 \\ \tilde{\mathbf{w}}_2 \end{bmatrix}$$

where $\tilde{\mathbf{w}}_1 \in \mathbb{R}^n$ and $\tilde{\mathbf{w}}_2 \in \mathbb{R}^{m-n}$. Show that

$$\tilde{\mathbf{w}}_1 = L_1^{-\top} U^{-\top} \tilde{\mathbf{c}} - S^\top \tilde{\mathbf{w}}_2.$$
(c) Write \( \mathbf{d} = L_1^{-\top} U^{-\top} \tilde{\mathbf{c}} \). Deduce that \( \tilde{\mathbf{w}}_2 \) may be obtained either as a solution to

\[
\min_{\tilde{\mathbf{w}}_2 \in \mathbb{R}^{m-n}} \left\| \begin{bmatrix} S^\top \\ I_{m-n} \end{bmatrix} \tilde{\mathbf{w}}_2 - \begin{bmatrix} \mathbf{d} \\ \mathbf{0} \end{bmatrix} \right\|_2
\]

or as

\[
\tilde{\mathbf{w}}_2 = (I_{m-n} + SS^\top)^{-1}S\mathbf{d}.
\]

Note that when \( m - n < n \), this method is advantageous for the same reason in Problem 5.

7. So far we have assumed that \( A \) has full column rank. Suppose now that \( \text{rank}(A) = r < \min\{m,n\} \).

(a) Show that the \( LU \) factorization obtained using GECP is of the form

\[
\Pi_1 A \Pi_2 = LU = \begin{bmatrix} L_1 \\ L_2 \end{bmatrix} \begin{bmatrix} U_1 \\ U_2 \end{bmatrix}
\]

where \( L_1, U_1 \in \mathbb{R}^{r \times r} \) are triangular and nonsingular.

(b) Show that the above equation may be rewritten in the form

\[
\Pi_1 A \Pi_2 = \begin{bmatrix} I_r \\ S_1 \end{bmatrix} L_1 U_1 \begin{bmatrix} I_r \\ S_2^\top \end{bmatrix}
\]

for some lower triangular matrices \( S_1 \) and \( S_2 \).

(c) Hence show that the Moore-Penrose inverse of \( A \) is given by

\[
A^\dagger = \Pi_2 \begin{bmatrix} I_r \\ S_2^\top \end{bmatrix}^\dagger U_1^{-1} L_1^{-1} \begin{bmatrix} I_r \\ S_1 \end{bmatrix}^\dagger \Pi_1.
\]

(d) Using the general formula (derived in the lectures) for the Moore-Penrose inverse of a rank-retaining factorization, what do you get for \( A^\dagger \)?