Problem 1. (15 points) Consider the problem of minimizing the function \( f(x) = \|x\|^2 \) subject to the constraint \( x \in S \), where \( S = \{ x \in \mathbb{R}^n : Ax = b \} \). Here \( A \in \mathbb{R}^{m \times n} (m \leq n) \) has \( m \) linearly independent rows. Show that the minimizer \( x^* \) is given by \( x^* = A^\top (AA^\top)^{-1}b \).

Problem 2. (20 points) Consider the hyperplane defined as \( S = \{ z \in \mathbb{R}^n : Az = b \} \), where \( A \in \mathbb{R}^{m \times n} \) has \( m \) linearly independent rows (\( m \leq n \)). Prove that

(a) \( AA^\top \) is invertible.

(b) The projection of \( x \in \mathbb{R}^n \) on \( S \) is given by
\[
z^* = x - A^\top (AA^\top)^{-1}(Ax - b).
\]

Problem 3. (15 points) Use the Lagrange multiplier theorem to solve the following problems of the form
minimize \( f(x) \), subject to \( h(x) = 0 \)

(The \( i \)-th entry of \( x \) is denoted as \( x_i \).)

(a) \( f(x) = \|x\|^2 \) and \( h(x) = \sum_{i=1}^n x_i - 2 \)

(b) \( f(x) = \sum_{i=1}^n x_i \), and \( h(x) = \|x\|^2 - 1 \)

(c) \( f(x) = \|x\|^2 \), and \( h(x) = x^\top Qx - 1 \), where \( Q \) is positive definite.

Problem 4. (15 points) Consider the problem of minimizing the function \( f : \mathbb{R}^2 \to \mathbb{R} \) defined as:
\[
f(x) = f(x_1, x_2) = 2x_1 + x_2 - x_1x_2.
\]
Minimize \( f \) subject to the constraint \( 2x_1 + x_2 = 2 \).

Hint: Don’t forget to check regularity conditions when applying the Lagrange multiplier theorem.

Problem 5. Define \( f : \mathbb{R}^2 \to \mathbb{R} \) as \( f(x_1, x_2) = 2(x_1 - 1)^2 + 4(x_2 - 2)^2 + \sin(x_1) - \cos(x_2) \).

(a) (5 points) Derive the gradient and Hessian of \( f(x_1, x_2) \), and determine if the objective is convex.

(b) (10 points) Write down the update rule of projected Newton method with stepsize \( \eta \) and projection in the unit square, \( S = \{(x_1, x_2) : |x_1| \leq 1, \text{ and } |x_2| \leq 1 \} \). How to compute the projection in this case?

Hint: The projected Newton method just alternates between Newton’s update and the projection step.

(c) (20 points) Implement the above projected Newton method in Python or MATLAB. Run the optimization from \( (x_1, x_2) = (-1, -1) \) with learning rate \( \eta = 1 \). Plot the trajectories in the parameter space, i.e., the space of \( (x_1, x_2) \), along with the values of \( f(x_1, x_2) \) during the optimization. Specifically, you need to plot two figures: (i) scatter plot showing \( (x_1, x_2) \) points for all iterations, where the x-axis if for \( x_1 \) and the y-axis is for \( x_2 \); you need to mark the initial and final iterations in the plot. (ii) plot the curve of \( f(x_1, x_2) \) vs. the number of iterations. Does the algorithm converge?