1. Consider the conjugate direction methods:

$$\alpha_k = -\frac{\mathbf{r}_k^T \mathbf{p}_k}{\mathbf{p}_k^T A \mathbf{p}_k},$$
$$\mathbf{r}_k = A\mathbf{x}_k - \mathbf{b},$$
$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{p}_k.$$

For any starting point $\mathbf{x}_0 \in \mathbb{R}^n$, prove that $\mathbf{x}_k$ is the global minimizer of $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T A\mathbf{x} - \mathbf{b}\mathbf{x}$ over $\mathbf{x}_0 + \text{span}\{\mathbf{p}_0, \mathbf{p}_1, \ldots, \mathbf{p}_{k-1}\}$. (Hint: consider $h(\mathbf{c}) := f(\mathbf{x}_0 + c_0\mathbf{p}_0 + c_1\mathbf{p}_1 + \cdots + c_{k-1}\mathbf{p}_{k-1})$.)

Solution. Define $P = \begin{bmatrix} p_0 & \ldots & p_{k-1} \end{bmatrix}$ and $c^T = \begin{bmatrix} c_0 & \ldots & c_{k-1} \end{bmatrix}$. Consider the function $h(c) = f(x_0 + Pc)$. Taking gradient with respect to c, we see that

$$\nabla h(c) = P^T \nabla f(x_0 + Pc)$$

Letting $c_i = \alpha_i$ as constructed in the conjugate direction method, we see that

$$\nabla h(c) = P^T \nabla f(x_k)$$

Recall that $\nabla f(x_k) = Ax_k - b = r_k$. Then,

$$\nabla h(c) = P^T r_k = 0$$

Since h is a convex function, our choice of c as in the conjugate direction algorithm is a global minimizer of h. Thus, x_k minimizes f over $x_0 + \text{span}\{p_0, \ldots, p_{k-1}\}$.

2. Use the KKT conditions to solve the following constrained optimization problem:

$$\begin{aligned} \min \quad & x_1 x_2 \\ S.t. \quad & x_1^2 + x_2^2 \le 1 \end{aligned}$$

Solution. Note first that x and y are symmetric. That is, each solution listed below is actually a pair of solutions. The KKT conditions for the optimization are

$$x_1^2 + x_2^2 \le 1$$
$$\begin{bmatrix} x_2 \\ x_1 \end{bmatrix} + 2\mu \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$$
$$\mu(1 - x_1^2 - x_2^2) = 0$$

The case $\mu = 0$ admits the solution $x = (0, 0)$. The case $\mu \neq 0$ leaves $(x, \mu) = (\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}, -\frac{1}{2})$ and $(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, \frac{1}{2})$. Comparing the two solutions, it follows that $x = (\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2})$ is the optimal solution (up to symmetry).

3. Consider the following minimization problem:

$$\begin{aligned} \min \quad & f(\mathbf{x}) \\ S.t. \quad & \mathbf{x} \in S \end{aligned} \tag{P}$$

where

$$S := \{ \mathbf{x} \in X \mid g_i(\mathbf{x}) \le 0,\ i = 1, \dots, m \},$$

X is a nonempty open set in $\mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}$ and $g_i : \mathbb{R}^n \to \mathbb{R}, i = 1, \dots, m$ are differentiable functions. Let $\bar{\mathbf{x}}$ be a feasible point and let $I := \{ i \mid g_i(\bar{\mathbf{x}}) = 0 \}$ be the index set of active constraints at $\bar{\mathbf{x}}$. Define

$$F_0 := \{ \mathbf{d} \in \mathbb{R}^n \mid \nabla f(\bar{\mathbf{x}})^T \mathbf{d} < 0 \} \text{ and } G_0' := \{ \mathbf{d} \in \mathbb{R}^n \setminus \{\mathbf{0}\} \mid \nabla g_i(\bar{\mathbf{x}})^T \mathbf{d} \le 0,\ i \in I \}$$

as in class. Prove that $\bar{\mathbf{x}}$ is a KKT point if and only if $F_0 \cap G_0' = \emptyset$. (Hint: Apply Farkas' Lemma.)

Proof. Let $\bar{\mathbf{x}}$ be a feasible point with $F_0 \cap G_0' = \emptyset$ and I be the set of active inequality constraints. It follows that, $\nabla f(\bar{\mathbf{x}})^T d < 0$ and $\nabla g_i(\bar{\mathbf{x}})^T d_i \le 0$ for all $i \in I$ is not solvable. Since this system has no solution, $\exists w \ge 0$ such that the system defined by $\nabla f(\bar{\mathbf{x}})^T d + w \le 0$ and $\nabla g_i(\bar{\mathbf{x}})^T d_i \le 0$ for all $i \in I$ also has no solution. That is, the following has no solution

$$\begin{bmatrix} \nabla f(\bar{\mathbf{x}})^T & 0 & 1 \\ 0 & \nabla g_I(\bar{\mathbf{x}})^T & 0 \end{bmatrix} \begin{bmatrix} d \\ d_I \\ w \end{bmatrix} \le 0, \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} d \\ d_I \\ w \end{bmatrix} \ge 0$$

Applying Farkas's Lemma, it follows that the system

$$A^T y = \begin{bmatrix} \nabla f(\bar{\mathbf{x}}) & 0 \\ 0 & \nabla g_I(\bar{\mathbf{x}})^T \\ 1 & 0 \end{bmatrix} \begin{bmatrix} y \\ y_I \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = c$$

has a solution. That is,

$$\nabla f(\bar{\mathbf{x}})^T + \sum_{i \in I} y_i \nabla g_i(\bar{\mathbf{x}})^T = 0$$

with $y_i \ge 0$ for all $i \in I$. Thus, $\bar{\mathbf{x}}$ satisfies the dual feasibility of the KKT conditions. The complementary slackness conditions are satisfied trivially. So $\bar{\mathbf{x}}$ is a KKT point. □

4. Suppose that linearity constraint qualification is satisfied for problem (P). Prove that if $\bar{\mathbf{x}}$ is a locally optimal solution, then $\bar{\mathbf{x}}$ is a KKT point.

Proof. Define the following sets

$$\begin{aligned} F_0 &:= \{d \in \mathbb{R}^n : \nabla f(\bar{\mathbf{x}})^T d < 0\} \\ D &:= \{d \ne 0 : \bar{\mathbf{x}} + \lambda d \in S \forall \lambda \in (0, \delta) \text{ for some } \delta > 0\} \\ G_0' &:= \{d \in \mathbb{R}^n \backslash \{0\} : \nabla g_i(\bar{\mathbf{x}})^T d \le 0, \forall i \in I\} \end{aligned}$$

Since $\bar{\mathbf{x}}$ is a local optimal solution and f is differentiable at $\bar{\mathbf{x}}$, by Theorem 1 it follows that $F_0 \cap D = \emptyset$. Since each g_i is affine, they must be concave. Thus, by lemma 2, $G_0' = D$. So $F_0 \cap G_0' = \emptyset$ and $\bar{\mathbf{x}}$ is a KKT point follows from the equivalent conditions proved in question 3. □

5. Consider problem (P) and assume that $X = \mathbb{R}^n$. In general, solving a KKT system is not an easy task. However, given a primal solution $\bar{\mathbf{x}}$, it is easy to check whether $\bar{\mathbf{x}}$ is a KKT point. Given the fact that a linear program is easy to solve, explain how to check whether $\bar{\mathbf{x}}$ is a KKT point.

   Solution. Given a primal solution, to determine if $\bar{\mathbf{x}}$ is a KKT point, it suffices to find a solution μ such that μ satisfies complementary slackness and dual feasibility. Any feasbility problem, as the one constructed here, can be modeled as the optimization problem

$$\begin{aligned} \min \quad & 0 \\ S.t. \quad & \mu_i \nabla g_I(\bar{\mathbf{x}}) = 0 \\ & \nabla f(\bar{\mathbf{x}}) + \sum_{i \in I} \mu_i \nabla g_i(\bar{\mathbf{x}}) = 0 \\ & \mu \geq 0 \end{aligned}$$

   Since the program is linear in the decision variables μ_i, determining if $\bar{\mathbf{x}}$ is a KKT point is a simple as solving a linear program, which is easy.

6. Check whether $\bar{\mathbf{x}} = (1, 2, 5)^T$ is a KKT point of

$$\begin{aligned} \min \quad & 2x_1^2 + x_2^2 + 2x_3^2 + x_1x_3 - x_1x_2 + x_1 + 2x_3 \\ S.t. \quad & x_1^2 + x_2^2 - \ \ x_3 \leq 0 \\ & x_1 + x_2 + 2x_3 \leq 16 \\ & x_1 + x_2 \qquad \geq 3 \\ & x_1,\ x_2,\ x_3 \geq 0. \end{aligned}$$

   Solution. Since the first and last constraints are active at $\bar{\mathbf{x}}$, to determine if $\bar{\mathbf{x}}$ is a KKT point, it suffices to check whether or not

$$\nabla f(\bar{\mathbf{x}}) + \mu_1 \nabla g_1(\bar{\mathbf{x}}) + \mu_3 \nabla g_3(\bar{\mathbf{x}}) = 0$$

   has a solution with $\mu_1, \mu_3 \geq 0$. That is, we need to check if

$$\begin{aligned} 8 + 2\mu_1 - \mu_3 &= 0 \\ 3 + 4\mu_1 - \mu_3 &= 0 \\ 23 - \mu_1 &= 0 \end{aligned}$$

   has a solution, and it does not.