Final Exam

(Due: 6pm on June 12, Sunday)

Instruction: Students should provide enough detail of the logical procedure of deriving answers. Answers without sufficient justification will receive partial or no credit. For questions involving MATLAB experiments, provide codes with comments. The maximum possible score is 100.

1. (25 points) Consider a constrained optimization problem:

$$\begin{aligned}
\text{minimize}_{\mathbf{x} \in \mathbb{R}^n} \quad & f(\mathbf{x}) \\
\text{subject to} \quad & g_i(\mathbf{x}) \le 0, \quad i = 1, \ldots, p \\
& h_i(\mathbf{x}) = 0, \quad i = 1, \ldots, m
\end{aligned} \tag{1}$$

Suppose that f, g_i's, and h_i's are twice continuously differentiable. In this question, we will derive a simple way to check whether a given feasible point $\mathbf{x}$ satisfies the KKT conditions.

Let $\bar{\mathbf{x}}$ be a feasible point of (1) and $\mathcal{W}$ denote the subset of indices of inequality constraints that are *active* at $\bar{\mathbf{x}}$, i.e., $\mathcal{W} \subset \{1, \ldots, p\}$ is such that $i \in \mathcal{W}$ if and only if $g_i(\bar{\mathbf{x}}) = 0$. Let q denote the cardinality of $\mathcal{W}$.

(a) Prove that **if there exist real constants $\bar{\lambda}_i, i \in \mathcal{W}$ and $\bar{\mu}_1, \ldots, \bar{\mu}_m$** such that

$$\nabla f(\bar{\mathbf{x}}) + \sum_{i \in \mathcal{W}} \bar{\lambda}_i \nabla g_i(\bar{\mathbf{x}}) + \sum_{i=1}^m \bar{\mu}_i \nabla h_i(\bar{\mathbf{x}}) = \mathbf{0}^T \tag{2}$$

and $\bar{\lambda}_i \ge 0$ for all $i \in \mathcal{W}$, **then $\bar{\mathbf{x}}$ satisfies the KKT conditions of (1)**.

(b) Prove that **if $\bar{\mathbf{x}}$ satisfies the KKT conditions of (1), then** there exist real constants $\bar{\lambda}_i, i \in \mathcal{W}$ and $\bar{\mu}_1, \ldots, \bar{\mu}_m$ such that

$$\nabla f(\bar{\mathbf{x}}) + \sum_{i \in \mathcal{W}} \bar{\lambda}_i \nabla g_i(\bar{\mathbf{x}}) + \sum_{i=1}^m \bar{\mu}_i \nabla h_i(\bar{\mathbf{x}}) = \mathbf{0}^T$$

and $\bar{\lambda}_i \ge 0$ for all $i \in \mathcal{W}$.

(c) Let $\mathbf{A}$ denote a $(q+m)$ by n matrix with its first q rows being $\{\nabla g_i(\bar{\mathbf{x}}) : \ i \in \mathcal{W}\}$ and its remaining m rows being $\nabla h_1(\bar{\mathbf{x}}), \nabla h_2(\bar{\mathbf{x}}), \ldots, \nabla h_m(\bar{\mathbf{x}})$. Suppose that $\mathbf{A}$ has full row rank, i.e., its rows are linearly independent. Show that if there exist $\bar{\lambda}_i, \ i \in \mathcal{W}$, and $\bar{\mu}_1, \ldots, \bar{\mu}_m$ satisfying the equation (2), then they can be obtained as follows:

$$\begin{bmatrix} \bar{\boldsymbol{\lambda}} \\ \bar{\boldsymbol{\mu}} \end{bmatrix} = (\mathbf{A}\mathbf{A}^T)^{-1} \mathbf{A} \cdot (-\nabla f(\bar{\mathbf{x}})^T).$$

where $\bar{\boldsymbol{\lambda}}$ denotes a q-dimensional vector consisting of $\bar{\lambda}_i, i \in \mathcal{W}$, and $\bar{\boldsymbol{\mu}} := (\bar{\mu}_1, \ldots, \bar{\mu}_m)$.

(d) Discuss the implications of your answers to part (a)-(c) on verifying the KKT conditions of (1) for any given point $\bar{\mathbf{x}}$.

2. (25 points) This is Exercise 2 in Chapter 13 of our textbook [Luenberger & Ye]. Consider the constrained problem:

$$\text{minimize}_{\mathbf{x} \in \mathbb{R}^n} \quad f(\mathbf{x})$$
$$\text{subject to} \quad \mathbf{x} \in \mathcal{S}$$
(3)

Let $\{c_k\}$ denote a sequence of positive real numbers satisfying (i) $c_{k+1} > c_k$ for all k, and (ii) $c_k \to \infty$. Let $P(\mathbf{x})$ be a penalty function satisfying (i) $P(\mathbf{x})$ is a continuous function of $\mathbf{x}$, (ii) $P(\mathbf{x}) = 0$ if $\mathbf{x} \in \mathcal{S}$, and (iii) $P(\mathbf{x}) > 0$ if $\mathbf{x} \notin \mathcal{S}$. In addition, let $q(c, \mathbf{x}) := f(\mathbf{x}) + cP(\mathbf{x})$.

Let ϵ be a positive real number, and let $\{\mathbf{x}_k\}$ be a sequence such that each $\mathbf{x}_k$ satisfies

$$q(c_k, \mathbf{x}_k) \leq [\min_{\mathbf{x}} q(c_k, \mathbf{x})] + \epsilon.$$

In other words, $\mathbf{x}_k$ might not be a global minimum point of $q(c_k, \mathbf{x})$; however, it at least satisfies that $q(c_k, \mathbf{x}_k) - \min_{\mathbf{x}} q(c_k, \mathbf{x})$ is bounded above by ϵ (e.g., $\mathbf{x}_k$ could be a local minimum point of $q(c_k, \mathbf{x})$ instead of a global minimum point).

Let $\mathbf{x}^*$ denote a global minimum point of (3). Prove that any limit point $\bar{\mathbf{x}}$ of $\{\mathbf{x}_k\}$ is (i) feasible and (ii) satisfies $f(\bar{\mathbf{x}}) \leq f(\mathbf{x}^*) + \epsilon$.

(Hint: Using the properties of the penalty method discussed in the lectures, first show that

$$f(\mathbf{x}_k) + c_k P(\mathbf{x}_k) \leq f(\mathbf{x}^*) + \epsilon.$$

Let $\bar{\mathbf{x}}$ be an arbitrary limit point of $\{\mathbf{x}_k\}$. Then, by definition, there exists $\mathcal{K} \subset \{1, 2, \ldots, \}$ such that the subsequence $\{\mathbf{x}_k\}_{k \in \mathcal{K}}$ converges to $\bar{\mathbf{x}}$. The above inequality can be used together with the subsequence $\{\mathbf{x}_k\}_{k \in \mathcal{K}}$ to show that $P(\bar{\mathbf{x}}) = 0$.)

3. (50 points) **Penalty Method.** Consider the constrained optimization problem:

$$\min_{\mathbf{x} \in \mathbb{R}^4} \quad \frac{1}{2}\mathbf{x}^T \mathbf{Q}\mathbf{x} + \mathbf{b}^T \mathbf{x}$$
$$\text{subject to} \quad \sum_{i=1}^4 x_i^2 = 1$$
$$x_i \geq 0, i = 1, 2, 3, 4;$$
(4)

where

$$\mathbf{Q} = \begin{bmatrix} 2 & 1 & 0 & 10 \\ 1 & 4 & 3 & 0.5 \\ 0 & 3 & -5 & 6 \\ 10 & 0.5 & 6 & -7 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} -1 \\ 0 \\ -2 \\ 3 \end{bmatrix}. \tag{5}$$

(a) Write the constrained optimization problem (4) in the standard form by properly defining f, h, $g_1, \ldots, g_4$:

$$\min_{\mathbf{x} \in \mathbb{R}^4} \quad f(\mathbf{x})$$

$$\text{subject to} \quad h(\mathbf{x}) = 0$$

$$g_i(\mathbf{x}) \le 0, i = 1, 2, 3, 4;$$

(b) Implement the penalty method with the *quadratic penalty function* to find a local minimum point of this constrained optimization problem. Provide a pseudocode of your implementation that explains the details of main steps clearly and the stopping criteria.

(c) Let $\mathbf{x}_k$ denote the solution at the end of the k-th iteration of the penalty method. And, let $c_k P(\mathbf{x}_k)$ denote the penalty term value at the end of the k-th iteration. Plot (i) $f(\mathbf{x}_k)$ versus k, (ii) $c_k P(\mathbf{x}_k)$ versus k, and (iii) $P(\mathbf{x}_k)$ versus k. And, interpret the plots.

(d) Let $\mathbf{x}_{final}$ denote the solution found by your implementation of the penalty method. Check whether the KKT conditions hold at $\mathbf{x}_{final}$ (you can use the results of Question 1). Check whether the second order necessary condition holds at $\mathbf{x}_{final}$. Also check whether the second order sufficient condition holds at $\mathbf{x}_{final}$.