Optimality Conditions

1. Constrained Optimization

1.1. **First–Order Conditions.** In this section we consider first–order optimality conditions for the constrained problem

$$\begin{array}{lll} \mathcal{P}: & \text{minimize} & f_0(x) \\ & \text{subject to} & x \in \Omega, \end{array}$$

where $f_0 : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable and $\Omega \subset \mathbb{R}^n$ is closed and non-empty. The first step in the analysis of the problem $\mathcal{P}$ is to derive conditions that allow us to recognize when a particular vector $\bar{x}$ is a solution, or local solution, to the problem. For example, when we minimize a function of one variable we first take the derivative and see if it is zero. If it is, then we take the second derivative and check that it is positive. If this is also true, then we know that the point under consideration is a local minimizer of the function. Of course, the presence of constraints complicates this kind of test.

To understand how an optimality test might be derived in the constrained case, let us first suppose that we are at a feasible point x and we wish to find a better point $\tilde{x}$. That is, we wish to find a point $\tilde{x}$ such that $\tilde{x} \in \Omega$ and $f(\tilde{x}) < f(x)$. As in the unconstrained case, one way to do this is to find a direction d in which the directional derivative of f in the direction d is negative: $f'(x; d) < 0$. We know that for such directions we can reduce the value of the function by moving away from the point x in the direction d. However, moving in such a direction may violate feasibility. That is, it may happen that $x + td \notin \Omega$ regardless of how small we take $t > 0$. To avoid this problem, we consider the notion of a *feasible direction*.

Definition 1.1. [Feasible Directions]
Given a subset Ω of $\mathbb{R}^n$ and a point $x \in \Omega$, we say that a direction $d \in \mathbb{R}^n$ is a feasible direction for Ω at x if there is a $\bar{t} > 0$ such that $x + td \in \Omega$ for all $t \in [0, \bar{t}]$.

Theorem 1.1. *If $\bar{x}$ is a local solution to the problem $\mathcal{P}$, then $f'(\bar{x}; d) \geq 0$ for all feasible directions d for Ω at $\bar{x}$ for which $f'(\bar{x}; d)$ exists.*

Proof. The proof is a straightforward application of the definitions. If the result were false, then there would be a direction of descent for f at $\bar{x}$ that is also a feasible direction for Ω at $\bar{x}$. But then moving a little bit in this direction both keeps us in Ω and strictly reduces the value of f. This contradicts the assumption that $\bar{x}$ is a local solution. Therefore, the result must be true. □

Unfortunately, this result is insufficient in many important cases. The insufficiency comes from the dependence on the notion of *feasible direction*. For example, if

$$\Omega = \{(x_1, x_2)^T \,:\, x_1^2 + x_2^2 = 1\},$$

then the only feasible direction at any point of Ω is the zero direction. Hence, regardless of the objective function f and the point $\bar{x} \in \Omega$, we have that $f'(\bar{x}; d) \geq 0$ for every feasible direction to Ω at $\bar{x}$. In this case, Theorem 1.1 has no content.

To overcome this deficiency we introduce a general notion of *tangency* that considers all directions d *pointing into* Ω at $x \in \Omega$ in a limiting sense. Define the *tangent cone* to Ω at a point $x \in \Omega$ to be the set of limiting directions obtained from sequences in Ω that converge to x. Specifically, the tangent cone is given by

$$T_\Omega(x) := \{d \,:\, \exists \tau_i \searrow 0, \text{ and } \{x_i\} \subset \Omega, \text{ with } x_i \to x, \text{ such that } \tau_i^{-1}(x_i - x) \to d\}.$$

EXAMPLES:

(1) If $\Omega = \{x \,:\, Ax = b\}$, where $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$, then $T_\Omega(x) = \mathrm{Nul}\,(A)$ for every $x \in \Omega$.

Reason: Let $x \in \Omega$. Note that if $d \in \mathrm{Nul}\,(A)$, then for every $t \geq 0$ we have $A(x + td) = Ax + tAd = Ax = b$ so that $d \in T_\Omega(x)$. Since $d \in \mathrm{Nul}\,(A)$ was chosen arbitrarily, this implies that $\mathrm{Nul}\,(A) \subset T_\Omega(x)$. Hence we only need to establish the reverse inclusion to verify the equivalence of these sets.
Let $d \in T_\Omega(x)$. Then, by definition, there are sequences $t_i \downarrow 0$ and $\{x^i\} \subset \Omega$ with $x^i \to x$ such that $d^i \to d$ where $d^i = t_i^{-1}(x^i - x),\ i = 1, 2, \ldots$. Note that

$$Ad^i = t_i^{-1}A(x^i - x) = t_i^{-1}[Ax^i - Ax] = t_i^{-1}[b - b] = 0 \quad \forall\, i = 1, 2, \ldots.$$

Therefore, $Ad = \lim_{i\to\infty} Ad^i = 0$ so that $d \in \mathrm{Nul}\,(A)$. Since d was chosen arbitrarily from $T_\Omega(x)$, we have $T_\Omega(x) \subset \mathrm{Nul}\,(A)$ which proves the equivalence.

(2) If $\Omega = \{(x_1, x_2)^T \,:\, x_1^2 + x_2^2 = 1\}$, then $T_\Omega(x) = \{(y_1, y_2) \,:\, x_1y_1 + x_2y_2 = 0\}$.

(3) A convex set is said to be *polyhedral* if it can be represented as the solution set of a finite number of linear equality and /or inequality constraints. Thus, for example te constraint region for an LPs is a convex polyhedron. If it is assumed that Ω is a convex polyhedron, then

$$T_\Omega(x) = \bigcup_{\lambda\geq 0} \lambda(\Omega - x) = \{\lambda(y - x) \;|\lambda \geq 0,\ y \in \Omega\}\,.$$

(4) If Ω is a convex subset of $\mathbb{R}^n$, then

$$T_\Omega(x) = \overline{\bigcup_{\lambda\geq 0} \lambda(\Omega - x)} = \mathrm{cl}\,\{\lambda(y - x) \;|\lambda \geq 0,\ y \in \Omega\}\,.$$

Theorem 1.2. *[*BASIC CONSTRAINED FIRST-ORDER NECESSARY CONDITIONS*]* *Suppose that the function $f_0 : \mathbb{R}^n \to \mathbb{R}$ in $\mathcal{P}$ is continuously differentiable near the point $\bar{x} \in \Omega$. If $\bar{x}$ is a local solution to $\mathcal{P}$, then*

$$f_0'(\bar{x} : d) \geq 0 \quad \textit{for all } d \in T_\Omega(\bar{x}).$$

Proof. The result follows immediately from the fact that

$$f_0'(x : d) = \lim_{\tau\searrow 0} \frac{f_0(\bar{x} + \tau d) - f_0(\bar{x})}{\tau} = \lim_{\substack{s\to d\\ \tau\searrow 0}} \frac{f_0(\bar{x} + \tau s) - f_0(\bar{x})}{\tau}$$

due to the fact that f_0 is continuously differentiable (just apply the Mean–Value Theorem). $\square$

This general result is not particulary useful on its own since it refers the the abstract notion of tangent cone. In order to make it useful, we need to be able to be able to compute the tangent cone once a representation for Ω is given. We now show how this can be done.

We begin by assuming that Ω has the form

$$\Omega := \{x : f_i(x) \le 0, i = 1, \ldots, s, f_i(x) = 0, i = s+1, \ldots, m\}, \tag{1}$$

where each $f_i : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable on $\mathbb{R}^n$. Observe that if $x \in \Omega$ and $d \in T_\Omega(x)$ then there are sequences $\{x_k\} \subset \Omega$ and $\tau_k \searrow 0$ with $x_k \to x$ such that $\tau_k^{-1}(x_k - x) \to d$. Setting $d_k = \tau_k^{-1}(x_k - x)$ for all k we have that

$$f_i'(x : d) = \lim_{k\to\infty} \frac{f_i(x + \tau_k d_k) - f_i(x)}{\tau_k}$$

equals 0 for $i \in \{s+1, \ldots, m\}$ and is less than or equal to 0 for $i \in I(x)$ where

$$I(x) := \{i \,:\, i \in \{1, \ldots, s\},\ f_i(x) = 0\}\ .$$

Consequently,

$$T_\Omega(x) \subset \{d \,:\, \nabla f_i(x)^T d \le 0,\ i \in I(x),\ \nabla f_i(x)^T d = 0, i = s+1, \ldots, m\}\ .$$

The set on the right hand side of this inclusion is a computationally tractable. Moreover, in a certain sense, the cases where these two sets do not coincide are exceptional. For this reason we make the following definition.

Definition 1.2. *[*REGULARITY*]*
We say that the set Ω is regular at $x \in \Omega$ if

$$T_\Omega(x) = \{d \in \mathbb{R}^n : f_i'(x; d) \le 0, i \in I(x), f_i'(x; d) = 0\ i = s+1, \ldots, m\}.$$

But we must always remember that not every set is regular.

EXERCISE: Graph the set

$$\Omega := \{x \in \mathbb{R}^2 | - x_1^3 \le x_2 \le x_1^3\},$$

and show that it is not regular at the origin. This is done by first showing that

$$T_\Omega(0) = \left\{(d_1, d_2)^T \ | d_1 \ge 0, d_2 = 0\right\}.$$

Then set

$$f_1(x_1, x_2) = -x_1^3 - x_2 \qquad \text{and} \qquad f_1(x_1, x_2) = -x_1^3 + x_2,$$

so that $\Omega = \left\{(x_1, x_2)^T \ | f_1(x_1, x_2) \le 0, f_2(x_1, x_2) \le 0\right\}$. Finally, show that

$$\left\{d \,\middle|\, \nabla f_1(0, 0)^T d \le 0, \nabla f_2(0, 0)^T d \le 0\right\} = \left\{(d_1, d_2)^T \ | d_2 = 0\right\} \ne T_\Omega(0).$$

Next let us suppose we are at a given point $x \in \Omega$ and that we wish to obtain a new point $x_+ = x + td$ for which $f(x_+) < f(x)$ for some direction $d \in \mathbb{R}^n$ and steplength $t > 0$. A good candidate for a search direction d is one that minimizes $f'(x; d)$ over all directions that point into Ω up to first-order. That is, we should minimize $\nabla f(x)^T d$ over the set of tangent directions. Remarkably, this search for a *feasible direction of steepest descent* can be posed as the following linear program (assuming regularity):

$$
\begin{array}{lll}
\max & (-\nabla f_0(\bar{x}))^T d & \\
\text{subject to} & \nabla f_i(\bar{x})^T d \le 0 & i \in I(\bar{x}) \\
& \nabla f_i(\bar{x})^T d = 0 & i = s+1, \dots, m.
\end{array}
\tag{2}
$$

The dual of (2) is the linear program

$$
\begin{array}{ll}
\min\ 0 & \\
\text{subject to} & \sum_{i \in I(\bar{x})} u_i \nabla f_i(\bar{x}) + \sum_{i=s+1}^m u_i \nabla f_i(\bar{x}) = -\nabla f_0(\bar{x}) \\
& 0 \le u_i,\ i \in I(\bar{x}).
\end{array}
\tag{3}
$$

If we assume that $\bar{x}$ is a local solution to $\mathcal{P}$, Theorem 1.2 tells us that the maximum in (2) is less than or equal to zero. But $d = 0$ is feasible for (2), hence the maximum value in (2) is zero. Therefore, by the Strong Duality Theorem for Linear Programming, the linear program (3) is feasible, that is, there exist scalars $u_i,\ i \in I(\bar{x}) \cup \{s+1, \dots, m\}$ with $u_i \ge 0$ for $i \in I(\bar{x})$ such that

$$
0 = \nabla f_0(\bar{x}) + \sum_{i \in I(\bar{x})} u_i \nabla f_i(\bar{x}) + \sum_{i=s+1}^m u_i \nabla f_i(\bar{x}).
\tag{4}
$$

This observation yields the following result.

Theorem 1.3. [Constrained First-Order Optimality Conditions]
Let $\bar{x} \in \Omega$ be a local solution to $\mathcal{P}$ at which Ω is regular. Then there exist $u \in \mathbb{R}^m$ such that

(1) $0 = \nabla_x L(\bar{x}, u)$,
(2) $0 = u_i f_i(\bar{x})$ *for* $i = 1, \dots, s$, *and*
(3) $0 \le u_i,\ i = 1, \dots, s$,

where the mapping $L : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is defined by

$$
L(x, u) := f_0(x) + \sum_{i=1}^m u_i f_i(x)
$$

and is called the Lagrangian for the problem $\mathcal{P}$.

Proof. For $i \in I(\bar{x}) \cup \{s+1, \dots, m\}$ let u_i be as given in (4) and for $i \in \{1, \dots, s\} \setminus I(\bar{x})$ set $u_i = 0$. Then this choice of $u \in \mathbb{R}^m$ satisfies (1)–(3) above. $\square$

Definition 1.3. [KKT Conditions]
Let $x \in \mathbb{R}^n$ and $u \in \mathbb{R}^m$. We say that (x, u) is a Karush-Kuhn-Tucker (KKT) pair for $\mathcal{P}$ if

(1) $f_i(x) \le 0\ i = 1, \dots, s,\ f_i(x) = 0\ i = s+1, \dots, m$ *(Primal feasibility),*
(2) $u_i \ge 0$ *for* $i = 1, \dots, s$ *(Dual feasibility),*
(3) $0 = u_i f_i(x)$ *for* $i = 1, \dots, s$ *(complementarity), and*
(4) $0 = \nabla_x L(x, u)$ *(stationarity of the Lagrangian).*

Given $x \in \mathbb{R}^n$, if there is a $u \in \mathbb{R}^m$ such that (x, u) is a Karush-Kuhn-Tucker pair for $\mathcal{P}$, then we say that x is a KKT point for $\mathcal{P}$ (we also refer to such an x as a stationary point for $\mathcal{P}$). ■

2. Regularity and Constraint Qualifications

We now briefly discuss conditions that yield the regularity of Ω at a point $x \in \Omega$. These conditions should be testable in the sense that there is a finitely terminating algorithm that can determine whether they are satisfied or not satisfied. The condition that we will concentrate on is the so called *Mangasarian-Fromovitz constraint qualification* (MFCQ).

Definition 2.1. *[*MFCQ*]*
We say that a point $x \in \Omega$ satisfies the Mangasarian-Fromovitz constraint qualification (or MFCQ) at x if

(1) *there is a $d \in \mathbb{R}^n$ such that*

$$\begin{aligned} &\nabla f_i(x)^T d < 0 \text{ for } i \in I(x), \\ &\nabla f_i(x)^T d = 0 \text{ for } i = s+1, \cdots, m, \end{aligned}$$

and

(2) *the gradients $\{\nabla f_i(x) | i = s+1, \cdots, m\}$ are linearly independent.*

We have the following key result which we shall not prove.

Theorem 2.1. *[*MFCQ $\to$ Regularity*] Let $f_i : \mathbb{R}^n \to \mathbb{R}, \ i = 1, 2, \cdots, m$ be C^1 near $\bar{x} \in \Omega$. If the MFCQ holds at $\bar{x}$, then Ω is regular at $\bar{x}$.*

The MFCQ is algorithmically verifiable. This is seen by considering the LP

$$\begin{array}{lll} \min & 0 & \\ \text{subject to} & \nabla f_i(\bar{x})^T d \le -1 & i \in I(\bar{x}) \\ & \nabla f_i(\bar{x})^T d = 0 & i = s+1, \cdots, m. \end{array} \tag{5}$$

Cleary, the MFCQ is satisfied at $\bar{x}$ if and only if the above LP is feasible and the gradients $\{\nabla f_i(\bar{x}) \mid i = s+1, \cdots, m\}$ are linearly independent. This observation also leads to a *dual* characterization of the MFCQ by considering the dual of the LP (5).

Lemma 2.1. *[*Dual MFCQ*]*
The MFCQ is satisfied at a point $\bar{x} \in \Omega$ if and only if the only solution to the system

$$\begin{aligned} \sum_{i=1}^{m} u_i \nabla f_i(\bar{x}) &= 0, \\ u_i f_i(\bar{x}) &= 0 \quad i = 1, 2, \cdots, s, \text{ and} \\ u_i &\ge 0 \qquad i = 1, 2, \cdots, s, \end{aligned}$$

is $u_i = 0, \ i = 1, 2, \cdots, m$.

Proof. The dual the LP (5) is the LP

$$\begin{array}{ll} \min & \sum_{i \in I(\bar{x})} u_i \\ \text{subject to} & \sum_{i \in I(\bar{x})} u_i \nabla f_i(\bar{x}) + \sum_{i=s+1}^{m} u_i \nabla f_i(\bar{x}) = 0 \\ & 0 \le u_i, \ i \in I(\bar{x}). \end{array} \tag{6}$$

This LP is always feasible, simply take all u_i's equal to zero. Hence, by the Strong Duality Theorem of Linear Programming, the LP (5) is feasible if and only if the LP (6) is finite valued in which case the optimal value in both is zero. That is, the MFCQ holds at $\bar{x}$ if and only if the optimal value in (6) is zero and the gradients $\{\nabla f_i(\bar{x}) \mid i = s+1, \cdots, m\}$ are linearly independent. The latter statement is equvalent to the statement that the only solution to the system

$$\sum_{i=1}^{m} u_i \nabla f_i(\bar{x}) = 0,$$
$$u_i f_i(\bar{x}) = 0 \quad i = 1, 2, \cdots, s, \text{ and}$$
$$u_i \geq 0 \qquad i = 1, 2, \cdots, s,$$

is $u_i = 0, \; i = 1, 2, \cdots, m$. □

Techniques similar to these show that the MFCQ is a local property. That is, if it is satisfied at a point then it must be satisfied on a neighborhood of that point. The MFCQ is a powerful tool in the analysis of constraint systems as it implies many useful properties. One such property is established in the following result.

Theorem 2.2. *[MFCQ → Compact Multiplier Set]*
Let $\bar{x} \in \Omega$ be a local solution to $\mathcal{P}$ at which the set of Karush-Kuhn-Tucker multipliers

$$\text{(7)} \qquad K\text{–}K\text{–}T(\bar{x}) := \left\{ u \in \mathbb{R}^m \;\middle|\; \begin{array}{c} \nabla_x L(\bar{x}, u) = 0 \\ u_i f_i(\bar{x}) = 0, \; i = 1, 2, \cdots, s, \\ 0 \leq u_i, \; i = 1, 2, \cdots, s \end{array} \right\}$$

is non-empty. Then $K\text{–}K\text{–}T(\bar{x})$ is a compact set if and only if the MFCQ is satisfied at $\bar{x}$.

Proof. (⇒) If MFCQ is not satisfied at $\bar{x}$, then from the Strong Duality Theorem for linear programming, Lemma 2.1, and the LP (6) guarentees the existence of a non-zero vector $\bar{u} \in \mathbb{R}^m$ satisfying

$$\sum_{i=1}^{m} u_i \nabla f_i(\bar{x}) = 0 \text{ and } 0 \leq u_i \text{ with } 0 = u_i f_i(\bar{x}) \text{ for } i = 1, 2, \cdots, s.$$

Then for each $u \in$ K–K–T$(\bar{x})$ we have that $u + t\bar{u} \in$ K–K–T$(\bar{x})$ for all $t > 0$. Consequently, K–K–T$(\bar{x})$ cannot be compact.
(⇐) If K–K–T$(\bar{x})$ is not compact, there is a sequence $\{u^j\} \subset$ K–K–T$(\bar{x})$ with $\|u^j\| \uparrow +\infty$. With no loss is generality, we may assume that

$$\frac{u^j}{\|u^j\|} \to u.$$

But then

$$\begin{array}{rll} u_i \geq 0, & i = 1, 2, \cdots, s, & \\ u_i f_i(\bar{x}) = & \lim_{i \to \infty} \frac{u^j}{\|u^j\|} f_i(\bar{x}) = & 0, \; i = 1, 2, \cdots, s, \text{ and} \\ \sum_{i=1}^{m} u_i f_i(\bar{x}) = & \lim_{i \to \infty} \frac{\nabla_x L(\bar{x}, u^j)}{\|u^j\|} = & 0. \end{array}$$

Hence, by Lemma 2.1, the MFCQ cannot be satisfied at $\bar{x}$. □

Before closing this section we introduce one more constraint qualification. This is the so called *LI* condition and is associated with the uniqueness of the multipliers..

Definition 2.2 (Linear Independence Condition)**.** *The LI condition is said to be satisfied at the point $x \in \Omega$ if the constraint gradients*

$$\{\nabla f_i(x) \mid i \in I(x) \cup \{s+1, \cdots, m\}\}$$

are linearly independent.

Clearly, the LI condition implies the MFCQ. However, it is a much stronger condition in the presence of inequality constraints. In particular, the LI condition implies the uniqueness of the multipliers at a local solution to $\mathcal{P}$.

3. Second–Order Conditions

Second–order conditions are introduced by way of the Lagrangian. As is illustrated in the following result, the multipliers provide a natural way to incorporate the curvature of the constraints.

Theorem 3.1. *[Constrained Second-Order Sufficiency]*
Let Ω have representation (1) and suppose that each of the functions f_i, $i = 0, 1, 2, \ldots m$ are C^2. Let $\overline{x} \in \Omega$. If $(\overline{x}, \overline{u}) \in \mathbb{R}^n \times \mathbb{R}^m$ is a Karush-Kuhn-Tucker pair for $\mathcal{P}$ such that

$$d^T \nabla_x^2 L(\overline{x}, \overline{u}) d > 0$$

for all $d \in T_\Omega(\overline{x})$, $d \neq 0$, with $\nabla f_0(\overline{x})^T d = 0$, then there is an $\epsilon > 0$ and $\nu > 0$ such that

$$f_0(x) \geq f_0(\overline{x}) + \nu \|x - \overline{x}\|^2$$

for every $x \in \Omega$ with $\|x - \overline{x}\| \leq \epsilon$, in particular $\overline{x}$ is a strict local solution to $\mathcal{P}$.

Proof. Suppose to the contrary that no such $\epsilon > 0$ and $\nu > 0$ exist, then there exist sequences $\{x_k\} \subset \Omega$, $\{\nu_k\} \subset \mathbb{R}_+$ such that $x_k \to \overline{x}$, $\nu_k \downarrow 0$, and

$$f_0(x_k) \leq f_0(\overline{x}) + \nu_k \|x_k - \overline{x}\|^2$$

for all $k = 1, 2, \ldots$. For every $x \in \Omega$ we know that $\overline{u}^T f(x) \leq 0$ and $0 = \overline{u}^T f(\overline{x})$ where the ith component of $f : \mathbb{R}^n \to \mathbb{R}^m$ is f_i. Hence

$$\begin{aligned} L(x_k, \overline{u}) \leq f_0(x_k) \quad & \leq f_0(\overline{x}) + \nu_k \|x_k - \overline{x}\|^2 \\ & = L(\overline{x}, \overline{u}) + \nu_k \|x_k - \overline{x}\|^2. \end{aligned}$$

Therefore,

$$f_0(\overline{x}) + \nabla f_0(\overline{x})^T (x_k - \overline{x}) + o(\|x_k - \overline{x}\|) \leq f_0(\overline{x}) + \nu_k \|x_k - \overline{x}\|^2 \tag{8}$$

and

$$\begin{aligned} L(\overline{x}, \overline{u}) \quad & + \nabla_x L(\overline{x}, \overline{u})^T (x_k - \overline{x}) \\ & + \tfrac{1}{2} (x_k - \overline{x})^T \nabla_x^2 L(\overline{x}, \overline{u}) (x_k - \overline{x}) + o(\|x_k - \overline{x}\|^2) \\ & \leq L(\overline{x}, \overline{u}) + \nu_k \|x_k - \overline{x}\|^2 \ . \end{aligned} \tag{9}$$

With no loss of generality, we can assume that

$$d_k := \frac{x_k - \overline{x}}{\|x_k - \overline{x}\|} \to \overline{d} \in T_\Omega(\overline{x}).$$

Dividing (8) through by $\|x_k - \overline{x}\|$ and taking the limit we find that $\nabla f_0(x)^T\overline{d} \le 0$. Since

$$T_\Omega(\overline{x}) \subset \{d : \nabla f_i(\overline{x})^T d \le 0,\ i \in I(\overline{x}),\ \nabla f_i(\overline{x})^T d = 0,\ i = s+1, \dots, m\},$$

we have $\nabla f_i(x)^T\overline{d} \le 0,\ i \in I(\overline{x}) \cup \{0\}$ and $\nabla f_i(x)^T\overline{d} = 0$ for $i = s+1, \dots, m$. On the other hand, $(\overline{x}, \overline{u})$ is a Karush-Kuhn-Tucker point so

$$\nabla f_0(\overline{x})^T\overline{d} = -\sum_{i \in I(\overline{x})} \overline{u}_i \nabla f_i(\overline{x})^T\overline{d} \ge 0.$$

Hence $\nabla f_0(\overline{x})^T\overline{d} = 0$, so that

$$\overline{d}^T\nabla_x^2 L(\overline{x}, \overline{u})\overline{d} > 0.$$

But if we divide (9) by $\|x_k - \overline{x}\|^2$ and take the limit, we arrive at the contradiction

$$\frac{1}{2}\overline{d}^T\nabla_x^2 L(\overline{x}, \overline{u})\overline{d} \le 0,$$

whereby the result is established. □

The assumptions required to establish Theorem 3.1 are somewhat strong but they do lead to a very practical and, in many cases, satisfactory second-order sufficiency result. In order to improve on this result one requires a much more sophisticated mathematical machinery. We do not take the time to develop this machinery. Instead we simply state a very general result. The statement of this result employs the entire set of Karush-Kuhn-Tucker multipliers K–K–T($\bar{x}$).

Theorem 3.2 (General Constrained Second-Order Necessity and Sufficiency). *Let $\overline{x} \in \Omega$ be a point at which Ω is regular.*

(1) *If $\overline{x}$ is a local solution to $\mathcal{P}$, then K–K–T($\overline{x}$) $\neq \emptyset$, and for every $d \in T_\Omega(\bar{x})$ there is a $u \in$ K–K–T($\overline{x}$) such that*

$$d^T\nabla_x^2 L(\overline{x}, u)d \ge 0.$$

(2) *If K–K–T($\overline{x}$) $\neq \emptyset$, and for every $d \in T_\Omega(\bar{x})$, $d \neq 0$, for which $\nabla f_0(\overline{x})^T d = 0$ there is a $u \in$ K–K–T($\overline{x}$) such that*

$$d^T\nabla_x^2 L(\overline{x}, u)d > 0,$$

then there is an $\epsilon > 0$ and $\nu > 0$ such that

$$f_0(x) \ge f_0(\overline{x}) + \nu\|x - \overline{x}\|^2$$

for every $x \in \Omega$ with $\|x - \overline{x}\| \le \epsilon$, in particular $\overline{x}$ is a strict local solution to $\mathcal{P}$.

4. Optimality Conditions in the Presence of Convexity

As we saw in the unconstrained case, convexity can have profound implications for optimality and optimality conditions. To begin with, we have the following very powerful result whose proof is identicle to the proof in the unconstrained case.

Theorem 4.1. *[Convexity+Local Optimality→Global Optimality]* *Suppose that $f_0 : \mathbb{R}^n \to \mathbb{R}$ is convex and that $\Omega \subset \mathbb{R}^n$ is a convex set. If $\overline{x} \in \mathbb{R}^n$ is a local solution to $\mathcal{P}$, then $\overline{x}$ is a global solution to $\mathcal{P}$.*

Proof. Suppose there is a $\widehat{x} \in \Omega$ with $f_0(\widehat{x}) < f_0(\overline{x})$. Let $\epsilon > 0$ be such that

$$f_0(\overline{x}) \le f_0(x) \quad \text{whenever} \quad \|x - \overline{x}\| \le \epsilon \text{ and } x \in \Omega,$$

and

$$\epsilon < 2\|\overline{x} - \widehat{x}\| \ .$$

Set $\lambda := \epsilon(2\|\overline{x} - \widehat{x}\|)^{-1} < 1$ and $x_\lambda := \overline{x} + \lambda(\widehat{x} - \overline{x}) \in \Omega$. Then $\|x_\lambda - \overline{x}\| \le \epsilon/2$ and $f_0(x_\lambda) \le (1-\lambda)f_0(\overline{x}) + \lambda f_0(\widehat{x}) < f_0(\overline{x})$. This contradicts the choice of ϵ and so no such $\widehat{x}$ exists. $\square$

We also have the following first-order necessary conditions for optimality. The proof of this result again follows that for the unconstrained case.

Theorem 4.2. *[1$^{\text{ST}}$-Order Necessity and Sufficiency]*
Suppose that $f_0 : \mathbb{R}^n \to \mathbb{R}$ is convex and that $\Omega \subset \mathbb{R}^n$ is a convex set, and let $\bar{x} \in \Omega$. Then the following statements are equivalent.

(i) *$\bar{x}$ is a local solution to $\mathcal{P}$.*
(ii) *$f_0'(\bar{x} : y - \bar{x}) \ge 0$ for all $y \in \Omega$.*
(iii) *$\bar{x}$ is a global solution to $\mathcal{P}$.*

Proof. The implication (i)⇒(ii) follows from Theorem 1.1 since each of the directions $d = y - \bar{x},\ y \in \Omega$ is a feasible direction for Ω at $\bar{x}$ due to the convexity of Ω. To see the implication (ii)⇒(iii), we again resort to the subdifferential inequality. Let y be any other point in Ω. Then $d = y - \bar{x} \in T_\Omega(\bar{x})$ and so by the subdifferential inequality we have

$$f_0(y) \ge f_0(\bar{x}) + f_0'(\bar{x}; y - \bar{x}) \ge f_0(\bar{x}).$$

Since $y \in \Omega$ was arbitrary the implication (ii)⇒(iii) follows. The implication (iii)⇒(i) is trivial. $\square$

The utility of this result again depends on our ability to represent the tangent cone $T_\Omega(\bar{x})$ in a computationally tractable manner. Following the general case, we assume that the set Ω has the representation (1):

$$\Omega := \{x : f_i(x) \le 0, i = 1, \ldots, s, f_i(x) = 0, i = s+1, \ldots, m\}. \tag{10}$$

The first issue we must address is to determine reasonable conditions on the functions f_i that guarentee that the set Ω is convex. We begin with the following elementary facts about convex functions and convex sets whose proofs we leave to the reader.

Lemma 4.1. *If $C_i \subset \mathbb{R}^n,\ i = 1, 2, \ldots, N$, are convex sets, then so is the set $C = \bigcap_{i=1}^N C_i$.*

Lemma 4.2. *If $h : \mathbb{R}^n \to \bar{\mathbb{R}}$ is a convex function, then for every $\alpha \in \mathbb{R}$ the set*

$$\text{lev}_\alpha(h) = \{x \ | h(x) \le \alpha\}$$

is a convex set.

These facts combine to give the following result.

Lemma 4.3. *If the functions $f_i,\ i = 1, 2, \ldots, s$ are convex and the functions $f_i,\ i = s + 1, \ldots, m$ are linear, then the set Ω given by (10) is a convex set.*

Remark: Recall that a function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be linear if there exists $c \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$ such that $f(x) = c^T x + \alpha$.

Proof. Note that

$$\Omega = \left(\bigcap_{i=1}^{m} \mathrm{lev}_0(f_i) \right) \cap \left(\bigcap_{i=s+1}^{m} \mathrm{lev}_0(-f_i) \right),$$

where each of the functions $f_i, i = 1, \ldots, m$ and $-f_i, i = s+1, \ldots, m$ is convex. Therefore, the convexity of Ω follows from Lemmas 4.2 and 4.1. $\square$

In order to make the link to the KKT condition in the presence of convexity, we still require the regularity of the set Ω at the point of interest $\bar{x}$. If the set Ω is a polyhedral convex set, i.e.

$$\Omega = \{x \mid Ax \leq a,\ Bx = b\}$$

for some $A \in \mathbb{R}^{s \times n}, a \in \mathbb{R}^s, B \in \mathbb{R}^{(m-s) \times n}$, and $b \in \mathbb{R}^{(m-s)}$, then the set Ω is everywhere regular (Why?). In the general convex case this may not be true. However, convexity can be used to derive a much simpler test for the regularity of non-polyhedral convex sets.

Definition 4.1 (The Slater Constraint Qualification)**.** *Let $\Omega \subset \mathbb{R}^n$ be as given in (10) with $f_i,\ i = 1, \ldots, s$ convex and $f_i,\ i = s+1, \ldots, m$ linear. We say that Ω satisfies the Slater constraint qualification if there exists $\tilde{x} \in \Omega$ such that $f_i(\tilde{x}) < 0$ for $i = 1, \ldots, s$.*

Theorem 4.3 (Convexity and Regularity)**.** *Suppose $\Omega \subset \mathbb{R}^n$ is as given in (10) with $f_i,\ i = 1, \ldots, s$ convex and $f_i,\ i = s+1, \ldots, m$ linear. If either Ω is polyhedral convex or satisfies the Slater constraint qualification, then Ω is regular at every point $\bar{x} \in \Omega$ at which the function $f_i,\ i = 1, \ldots, s$ are differentiable.*

We do not present the proof of this result as it takes us too far afield of our study. Nonetheless, we make use of this fact in the following result of the KKT conditions.

Theorem 4.4 (Convexity+Regularity→(Optimality⇔ KKT Conditions))**.** *Let $f_0 : \mathbb{R}^n \to \mathbb{R}$ be a differentiable convex function and let Ω be as given in Lemma 4.3 where each of the function $f_i,\ i = 1, \ldots, s$ is differentiable.*

(i) *If $\bar{x} \in \Omega$ is a KKT point for $\mathcal{P}$, then $\bar{x}$ is a global solution to $\mathcal{P}$.*
(ii) *Suppose the functions $f_i,\ i = 0, 1, \ldots, s$ are continuously differentiable. If $\bar{x}$ is a solution to $\mathcal{P}$ at which Ω is regular, then $\bar{x}$ is a KKT point for $\mathcal{P}$.*

Proof. Part (ii) of this theorem is just a restatement of Theorem 1.3 and so we need only prove Part (i).

Since $\bar{x}$ is a KKT point there exists $\bar{y} \in \mathbb{R}^m$ such that $(\bar{x}, \bar{y})$ is a KKT pair for $\mathcal{P}$. Consider the function $h : \mathbb{R}^n \to \mathbb{R}$ given by

$$h(x) = L(x, \bar{y}) = f_0(x) + \sum_{i=1}^{m} \bar{y}_i f_i(x).$$

By construction, the function h is convex with $0 = \nabla h(\bar{x}) = \nabla_x L(\bar{x}, \bar{y})$. Therefore, $\bar{x}$ is a global solution to the problem $\min_{x\in\mathbb{R}^n} h(x)$. Also note that for every $x \in \Omega$ we have

$$\sum_{i=1}^{m} \bar{y}_i f_i(x) \leq 0,$$

since $\bar{y}_i f_i(x) \leq 0$ $i = 1, \ldots, s$ and $\bar{y}_i f_i(x) = 0$ $i = s+1, \ldots, m$. Consequently,

$$\begin{aligned} f_0(\bar{x}) &= h(\bar{x}) \leq h(x) = L(x, \bar{y}) \\ &= f_0(x) + \sum_{i=1}^{m} \bar{y}_i f_i(x) \\ &\leq f_0(x) \end{aligned}$$

for all $x \in \Omega$. This establishes Part (i). □

If all of the functions f_i $i = 0, 1, \ldots, m$ are twice continuously differentiable, then the second-order sufficiency conditions stated in Theorem 3.1 apply. However, in the presence of convexity another kind of second-order condition is possible that does not directly incorporate curvature information about the functions $f_i i = 1, \ldots, m$. These second-order conditions are most appropriate when Ω is polyhedral convex.

Theorem 4.5. *[2ND-Order Optimality Conditions for Polyhedral Constraints] Let $f_0 : \mathbb{R}^n \to \mathbb{R}$ be C^2 and $\overline{x}$ be an element of the convex set Ω.*

(1) *(necessity) If $\overline{x} \in \mathbb{R}^n$ is a local solution to $\mathcal{P}$ with Ω a polyhedral convex set, then $\nabla f_0(\overline{x})^T d \geq 0$ for all $d \in T_\Omega(\bar{x})$ and*

$$d^T \nabla^2 f_0(\overline{x}) d \geq 0$$

for all $d \in T_\Omega(\bar{x})$ with $\nabla f(\overline{x})^T d = 0$.

(2) *(sufficiency) If $\overline{x} \in \mathbb{R}^n$ is such that $\nabla f_0(\overline{x})^T (y - \overline{x}) \geq 0$ for all $d \in T_\Omega(\bar{x})$ and*

$$d^T \nabla^2 f_0(\overline{x}) d > 0$$

for all $d \in T_\Omega(\bar{x}) \backslash \{0\}$ with $\nabla f_0(\overline{x})^T d = 0$, then there exist $\epsilon, \nu > 0$ such that

$$f_0(x) \geq f_0(\overline{x}) + \nu \|x - \overline{x}\|^2$$

for all $x \in \Omega$ with $\|x - \bar{x}\| \leq \epsilon$.

Proof. (1) Since Ω is polyhedral convex, we have $T_\Omega(\bar{x}) = \bigcup_{\lambda \geq 0}(\Omega - \bar{x})$. Therefore, the fact that $\nabla f_0(\overline{x})^T d \geq 0$ for all $d \in T_\Omega(\bar{x})$ follows from Theorem 4.2. Next let $d \in T_\Omega(\bar{x}) = \bigcup_{\lambda \geq 0}(\Omega - \bar{x})$ be such that $\nabla f_0(\overline{x})^T d = 0$. Then there is a $y \in \Omega$, $y \neq \overline{x}$, and a $\lambda_0 > 0$ such that $d = \lambda_0 (y - \overline{x})$. Let $\epsilon > 0$ be such that $f_0(\overline{x}) \leq f_0(x)$ for all $x \in \Omega$ with $\|x - \overline{x}\| \leq \epsilon$. Set $\overline{\lambda} = \min\{\lambda_0, \epsilon(\lambda_0 \|y - \overline{x}\|)^{-1}\} > 0$ so that $\overline{x} + \lambda d \in \Omega$ and $\|\overline{x} - (\overline{x} + \lambda d)\| \leq \epsilon$ for all $\lambda \in [0, \overline{\lambda}]$. By hypothesis, we now have

$$\begin{aligned} f_0(\overline{x}) &\leq f_0(\overline{x} + \lambda d) \\ &= f_0(\overline{x}) + \lambda \nabla f_0(\overline{x})^T (y - \overline{x}) + \tfrac{\lambda^2}{2} d^T \nabla^2 f_0(\overline{x}) d + o(\lambda^2) \\ &= f_0(\overline{x}) + \tfrac{\lambda^2}{2} d^T \nabla^2 f_0(\overline{x}) d + o(\lambda^2), \end{aligned}$$

where the second equality follows from the choice of d ($\nabla f_0(\overline{x})^T d = 0$) Therefore $d^T\nabla^2 f_0(\overline{x})d \ge 0$.

(2) We show that $f_0(\overline{x}) \le f_0(x) - \nu\|x-\overline{x}\|^2$ for some $\nu > 0$ for all $x \in \Omega$ near $\bar{x}$. Indeed, if this were not the case there would exist sequences $\{x_k\} \subset \Omega$, $\{\nu_k\} \subset \mathbb{R}_+$ with $x_k \to \overline{x}$, $\nu_k \downarrow 0$, and

$$f_0(x_k) < f_0(\overline{x}) + \nu_k\|x_k - \overline{x}\|^2$$

for all $k = 1, 2, \ldots$ where, with no loss of generality, $\frac{x_k-\overline{x}}{\|x_k-\overline{x}\|} \to d$. Clearly, $d \in T_\Omega(\bar{x})$. Moreover,

$$\begin{aligned} f_0(\overline{x}) + \nabla f_0(\overline{x})^T(x_k - \overline{x}) \quad & + o(\|x_k - \overline{x}\|) \\ & = f_0(x_k) \\ & \le f_0(\overline{x}) + \nu_k\|x_k - \overline{x}\|^2 \end{aligned}$$

so that $\nabla f_0(\overline{x})^T d = 0$.

Now, since $\nabla f_0(\overline{x})^T(x_k - \overline{x}) \ge 0$ for all $k = 1, 2, \ldots$,

$$\begin{aligned} & f_0(\overline{x}) + \tfrac{1}{2}(x_k - \overline{x})^T\nabla^2 f_0(\overline{x})(x_k - \overline{x}) + o(\|x_k - \overline{x}\|^2) \\ & \quad \le f_0(\overline{x}) + \nabla f_0(\overline{x})^T(x_k - \overline{x}) + \tfrac{1}{2}(x_k - \overline{x})^T\nabla^2 f_0(\overline{x})(x_k - \overline{x}) \\ & \qquad + o(\|x_k - \overline{x}\|^2) \\ & \quad = f_0(x_k) \\ & \quad < f_0(\overline{x}) + \nu_k\|x_k - \overline{x}\|^2. \end{aligned}$$

Hence,

$$\left(\frac{x_k - \overline{x}}{\|x_k - \overline{x}\|}\right)^T \nabla^2 f_0(\overline{x}) \left(\frac{x_k - \overline{x}}{\|x_k - \overline{x}\|}\right) \le \nu_k + \frac{o(\|x_k - \overline{x}\|^2)}{\|x_k - \overline{x}\|^2}$$

Taking the limit in k we obtain the contradiction

$$0 < d^T\nabla^2 f_0(\overline{x})d \le 0,$$

whereby the result is established. $\square$

Although it is possible to weaken the assumption of polyhedrality in Part 1, such weakenings are somewhat artificial as they essentially imply that $T_\Omega(x) = \bigcup_{\lambda \ge 0}(\Omega - x)$. The following example illustrates what can go wrong when the assumption of polyhedrality is dropped.

Example: Consider the problem

$$\begin{array}{ll} \min & \frac{1}{2}(x_2 - x_1^2) \\ \text{subject to} & 0 \le x_2,\ x_1^3 \le x_2^2. \end{array}$$

Observe that the constraint region in this problem can be written as $\Omega := \{(x_1, x_2)^T : |x_1|^{\frac{3}{2}} \le x_2\}$, therefore

$$\begin{aligned} f_0(x) &= \frac{1}{2}(x_2 - x_1^2) \\ &\ge \frac{1}{2}(|x_1|^{\frac{3}{2}} - |x_1|^2) \\ &= \frac{1}{2}|x_1|^{\frac{3}{2}}(1 - |x_1|^{\frac{1}{2}}) > 0 \end{aligned}$$

whenever $0 < |x_1| \leq 1$. Consequently, the origin is a strict local solution for this problem. Nonetheless,

$$T_\Omega(0) \cap [\nabla f_0(0)]^\perp = \{(\delta, 0)^T \,:\, \delta \in \mathbb{R}\},$$

while

$$\nabla^2 f_0(0) = \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix} .$$

That is, even though the origin is a strict local solution, the Hessian of f_0 is not positive semidefinite on $T_\Omega(0)$.

When using the second-order conditions given above, one needs to be careful about the relationship between the Hessian of f_0 and the set $K := T_\Omega(x) \cap [\nabla f_0(x)]^\perp$. In particular, the positive definiteness (or semidefiniteness) of the Hessian of f_0 on the cone K does not necessarily imply the positive definiteness (or semidefiniteness) of the Hessian of f_0 on the subspace spaned by K. This is illustrated by the following example.

Example: Consider the problem

$$\begin{array}{ll} \min & (x_1^2 - \frac{1}{2}x_2^2) \\ \text{subject to} & -x_1 \leq x_2 \leq x_1. \end{array}$$

Clearly, the origin is the unique global solution for this problem. Moreover, the constraint region for this problem, Ω, satisfies

$$T_\Omega(0) \cap [\nabla f(0)]^\perp = T_\Omega(0) = \Omega \ ,$$

with the span of Ω being all of $\mathbb{R}^2$. Now, while the Hessian of f_0 is positive definite on Ω, it is not positive definite on all of $\mathbb{R}^2$.

In the polyhedral case it is easy to see that the sufficiency result in Theorem 4.5 is equivalent to the sufficiency result of Theorem 3.1. However, in the nonpolyhedral case, these results are not comparable. It is easy to see that Theorem 4.5 can handle situations where Theorem 3.1 does not apply even if Ω is given in the form (1). Just let one of the active constraint functions be nondifferentiable at the solution. Similarly, Theorem 3.1 can provide information when Theorem 4.5 does not. This is illustrated by the following example.

Example: Consider the problem

$$\begin{array}{ll} \min & x_2 \\ \text{subject to} & x_1^2 \leq x_2. \end{array}$$

Clearly, $\bar{x} = 0$ is the unique global solution to this convex program. Moreover,

$$\begin{array}{rcl} f_0(\bar{x}) + \dfrac{1}{2}\|x - \bar{x}\|^2 & = & \dfrac{1}{2}(x_1^2 + x_2^2) \\ & \leq & \dfrac{1}{2}(x_2 + x_2^2) \\ & \leq & x_2 = f_0(x) \end{array}$$

for all x in the constraint region Ω with $\|x - \bar{x}\| \leq 1$. It is easily verified that this growth property is predicted by Theorem 4.5.

Exercises

(1) Locate all of the KKT points for the following problems. Can you show that these points are local solutions? Global solutions?

(a)
$$\begin{array}{ll} \text{minimize} & e^{(x_1-x_2)} \\ \text{subject to} & e^{x_1}+e^{x_2}\le 20 \\ & 0\le x_1 \end{array}$$

(b)
$$\begin{array}{ll} \text{minimize} & e^{(-x_1+x_2)} \\ \text{subject to} & e^{x_1}+e^{x_2}\le 20 \\ & 0\le x_1 \end{array}$$

(c)
$$\begin{array}{ll} \text{minimize} & x_1^2+x_2^2-4x_1-4x_2 \\ \text{subject to} & x_1^2\le x_2 \\ & x_1+x_2\le 2 \end{array}$$

(d)
$$\begin{array}{ll} \text{minimize} & \frac{1}{2}\|x\|^2 \\ \text{subject to} & Ax=b \end{array}$$

where $b\in\mathbb{R}^m$ and $A\in\mathbb{R}^{m\times n}$ satisfies $\mathrm{Nul}\,(A^T)=\{0\}$.

(2) Show that the set
$$\Omega := \{x\in\mathbb{R}^2 \mid -x_1^3\le x_2\le x_1^3\}$$
is not regular at the origin. Graph the set Ω.

(3) Construct an example of a constraint region of the form (1) at which the MFCQ is satisfied, but the LI condition is not satisfied.

(4) Suppose $\Omega=\{x\,;\,Ax\le b,\ Ex=h\}$ where $A\in\mathbb{R}^{m\times}$, $E\in\mathbb{R}^{k\times n}$, $b\in\mathbb{R}^m$, and $h\in\mathbb{R}^k$.

(a) Given $x\in\Omega$, show that
$$T_\Omega(x)=\{d\,:\,A_{i\cdot}d\le 0 \text{ for } i\in I(x),\ Ed=0\},$$
where $A_{i\cdot}$ denotes the ith row of the matrix A and $I(x)=\{i\ A_{i\cdot}x=b_i\}$.

(b) Given $x\in\Omega$, show that every $d\in T_\Omega(x)$ is a feasible direction for Ω at x.

(c) Note that parts (a) and (b) above show that
$$T_\Omega(x)=\bigcup_{\lambda>0}\lambda(\Omega-x)$$
whenever Ω is a convex polyhedral set. Why?

(5) Let $C\subset\mathbb{R}^n$ be non-empty, closed and convex. For any $x\in\mathbb{R}^n$ consider the problem of finding the closest point in C to x using the 2-norm:
$$\mathcal{D}\qquad \begin{array}{ll} \text{minimize} & \frac{1}{2}\|x-z\|_2^2 \\ \text{subject to} & x\in C \end{array}.$$

Show that $\bar{z}\in C$ solves this problem if and only if
$$\langle x-\bar{z}, z-\bar{z}\rangle\le 0 \quad \text{for all } z\in C.$$

(6) Let Ω be a non-empty closed convex subset of $\mathbb{R}^n$. The geometric object *dual* to the tangent cone is called the *normal cone*:

$$N_\Omega(x) = \{z\,;\ \langle z, d\rangle \le 0,\ \text{for all } d \in T_\Omega(x)\}.$$

(a) Show that if $\bar{x}$ solves the problem $\min\{f(x) : x \in \Omega\}$ then

$$-\nabla f(\bar{x}) \in N_\Omega(\bar{x}).$$

(b) Show that

$$N_\Omega(\bar{x}) = \{z : \langle z, x - \bar{x}\rangle \le 0,\ \text{for all } x \in \Omega\}.$$

(c) Let $\bar{x} \in \Omega$. Show that $\bar{x}$ solves the problem $\min\{\frac{1}{2}\|x - y\|_2^2 : x \in \Omega\}$ for every $y \in \bar{x} + N_\Omega(\bar{x})$.

(7) Consider the functions

$$f(x) = \frac{1}{2}x^T Q x - c^T x$$

and

$$f_t(x) = \frac{1}{2}x^T Q x - c^T x + t\phi(x),$$

where $t > 0$, $Q \in \mathbb{R}^{n\times n}$ is positive semi-definite, $c \in \mathbb{R}^n$, and $\phi : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is given by

$$\phi(x) = \begin{cases} -\sum_{i=1}^n \ln x_i & \text{, if } x_i > 0,\ i = 1, 2, \ldots, n, \\ +\infty & \text{, otherwise.} \end{cases}$$

(a) Show that ϕ is a convex function.
(b) Show that both f and f_t are convex functions.
(c) Show that the solution to the problem $\min f_t(x)$ always exists and is unique.
(d) Let $\{t_i\}$ be a decreasing sequence of positive real scalars with $t_i \downarrow 0$, and let x^i be the solution to the problem $\min f_{t_i}(x)$. Show that if the sequence $\{x^i\}$ has a cluster point $\bar{x}$, then $\bar{x}$ must be a solution to the problem $\min\{f(x) : 0 \le x\}$. *Hint:* Use the KKT conditions for the QP $\min\{f(x) : 0 \le x\}$.