Finite Dimensional Optimization
Part II: Sufficiency

Recall from the notes on the Karush-Kuhn-Tucker (KTT) Theorem that, for a feasible point $x^*$, $J$ is the set of indices for which the constraints are binding at $x^*$ ($g_k(x^*) = 0$ for every $k \in J$). Recall also that the KKT condition is that there exist $\lambda_k \geq 0$ for all $k \in J$ such that

$$\nabla f(x^*) = \sum_{k \in J} \lambda_k \nabla g_k(x^*);$$

if no constraints are binding ($J = \emptyset$) then the KKT condition reduces to $\nabla f(x^*) = 0$.

The KKT condition is necessary for a feasible $x^*$ to be a solution, but not always sufficient. For example, for the domain $\mathbb{R}$, the KKT condition holds at the origin for $f(x) = -x^4$, $\hat{f}(x) = x^4$, and $\tilde{f}(x) = x^3$ (i.e., the derivative equals 0 in every case). But while $x^* = 0$ is the unique maximum for $f$, it is the unique minimum for $\hat{f}$ (and, in particular, is not a maximum), and is neither a minimum nor a maximum for $\tilde{f}$.

Say that a function $f$ is a differentiably strictly increasing transformation of a function $\hat{f}$ iff both functions have the same domain and there is a differentiable function $h$, with domain containing the image of $\hat{f}$, such that (a) $f = h \circ \hat{f}$ and (b) $Dh(\hat{f}(x)) > 0$ for every $x$ in the domain of $\hat{f}$. Any concave function is, trivially, a differentiably strictly increasing transformation of a concave function: simply take $f = \hat{f}$ and $h$ to be the identity, $h(y) = y$.

**Theorem 1.** Consider a differentiable MAX problem in standard form with objective function $f$.

1. If $x^*$ is feasible, and the KKT condition holds at $x^*$, then $x^*$ is a solution to the MAX problem if either,

   (a) $f$ is concave (or is a differentiably strictly increasing transformation of a concave function), or

   (b) $f$ is quasi-concave, $\nabla f(x^*) \neq 0$, and any binding constraint functions are quasi-convex.

2. If $f$ is strictly quasi-concave, and if the constraint functions are quasi-convex, then the solution to the MAX problem is unique.

©️ This work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License.
Proof.

1. The proof is by contraposition. Suppose that there is a feasible \( x \) such that \( f(x) > f(x^*) \). Let \( v = x - x^* \). I claim that \( \nabla f(x^*) \cdot v > 0 \).

(a) Suppose that \( f \) is concave. Then \( f(x) \leq \nabla f(x^*) \cdot (x - x^*) + f(x^*) \), hence

\[
0 < f(x) - f(x^*) \leq \nabla f(x^*) \cdot v.
\]

Suppose that \( f \) is a differentiably strictly increasing transformation of \( f^\ast \), with \( f \) concave. Since \( f(x) > f(x^*) \) iff \( f^\ast(x) > f^\ast(x^*) \), the above argument implies \( \nabla f^\ast(x^*) \cdot v > 0 \). Since, by the Chain Rule, \( \nabla f(x^*) = Dh(f^\ast(x)) \nabla f^\ast(x^*) \), it follows that \( \nabla f(x^*) \cdot v = Dh(f^\ast(x)) \nabla f^\ast(x^*) \cdot v > 0 \).

(b) Suppose that \( f \) is merely quasi-concave. Since \( f \) is continuous, there is an \( \varepsilon > 0 \) such that for any \( w \) on the unit sphere in \( \mathbb{R}^N \), \( f(x + \varepsilon w) > f(x^*) \) for any \( \theta \in (0, 1) \), quasi-concavity then implies that \( f(x^*) \leq f(\theta(x + \varepsilon w) + (1 - \theta)x^*) = f(x^* + \theta(x + \varepsilon w - x^*)) \), or

\[
f(x^* + \theta(x + \varepsilon w - x^*)) - f(x^*) \geq 0.
\]

Dividing by \( \theta > 0 \) and taking the limit as \( \theta \downarrow 0 \) implies that the directional derivative of \( f \) at \( x^* \) in the direction \( x + \varepsilon w - x^* \) is non-negative. Since \( f \) is differentiable, this implies that,

\[
\nabla f(x^*) \cdot (x + \varepsilon w - x^*) \geq 0,
\]

hence,

\[
\nabla f(x^*) \cdot v + \varepsilon \nabla f(x^*) \cdot w \geq 0.
\]

This holds for all \( w \) on the unit sphere. Since \( \nabla f(x^*) \neq 0 \), there is a \( w \) such that \( \nabla f(x^*) \cdot w < 0 \). The claim follows.

If \( J = \emptyset \), then \( \nabla f(x^*) \cdot v > 0 \) implies that the KKT condition (which, in this case, is \( \nabla f(x^*) = 0 \)) does not hold, and the proof follows by contraposition.

If \( J \neq \emptyset \) then, by the KKT condition,

\[
0 < \nabla f(x^*) \cdot v = \sum_{k \in J} \lambda_k \nabla g_k(x^*) \cdot v.
\]

which implies that there is at least one \( k \in J \) such that \( \nabla g_k(x^*) \cdot v > 0 \). But since \( g_k \) is differentiable, this implies that the directional derivative of \( g_k \) at \( x^* \) in the direction \( v \) is strictly positive. This then implies that for (all) \( \theta \in (0, 1) \) sufficiently small, \( g_k(x^* + \theta v) > g_k(x^*) \), hence \( g_k(\theta x + (1 - \theta)x^*) > g_k(x^*) \).

But since \( x \) is feasible, \( g_k(x) \leq 0 \), and since \( k \in J \), \( g_k(x^*) = 0 \). Together, these inequalities imply that \( g_k \) is not quasi-convex. Again, the proof follows by contraposition.
2. Suppose that there is a feasible \( x \) with \( f(x) = f(x^*) \). Then, by the definition of strict quasi-concavity, if \( x \neq x^* \) then for any \( \theta \in (0, 1) \), \( f(\theta x + (1 - \theta)x^*) > f(x^*) \). Since \( \theta x + (1 - \theta)x^* \) is feasible (the constraint set is convex if the \( g_k \) are quasi-convex), it follows by contraposition that the maximum is unique.

Theorem 1, in conjunction with the KKT Theorem, implies the following result, which gives a checklist for optimization problems. Recall that the Slater condition is that \( g_k(x) < 0 \) for all \( x \) in the domain of \( f \) (i.e., the constraint set has a non-empty interior).

**Theorem 2.** Consider a differentiable MAX problem in standard form, with objective function \( f \). Let \( x^* \) be feasible. If

1. \( f \) is either (a) concave, or (b) a differentiably strictly increasing transformation of a concave function, or (c) quasi-concave with \( \nabla f(x^*) \neq 0 \),

2. every binding constraint (if any) is either (a) convex or (b) quasi-convex with \( \nabla g_k(x^*) \neq 0 \),

3. the Slater condition holds,

then a necessary and sufficient for \( x^* \) to be a solution is that the KKT condition holds at \( x^* \).

**Proof.** This is an immediate corollary of Theorem 1 and results from the notes on the KKT Theorem. □

**Remark 1.** The companion notes on Convex Optimization establish (a version of) Theorem 2 by a different route. □

**Example 1.** Consider the following problem,

\[
\max_{x \in \mathbb{R}^+_N} \quad \prod_n x_n^{\alpha_n} \quad p \cdot x \leq m
\]

with \( \alpha_n \in (0, 1) \) for all \( n \), \( \sum_n \alpha_n = 1 \), \( p \in \mathbb{R}^N_+ \), \( m \in \mathbb{R}^+_+ \). For interpretation, this is a competitive demand problem with Cobb-Douglas utility, price vector \( p \), and income \( m \).

To apply Theorem 2, note the following.

1. The objective function is actually concave. But rather than show this, note that this objective function is a differentiably strictly increasing transformation of

\[
\hat{f}(x) = \sum_n \alpha_n \ln(x_n),
\]
(set $h(y) = e^y$) and $\hat{f}$ is (differentiably strictly) concave, since it is a positive weighted sum of logs.

These is an issue here in that the original objective function is defined over all of $\mathbb{R}^N_+$ whereas $\hat{f}$ is defined over only $\mathbb{R}^N_{++}$. But the domain is effectively $\mathbb{R}^N_{++}$ even for the original problem, because any solution must be strictly positive: $f(x) = 0$ for any $x \gg 0$ (any $x$ for which $x_n = 0$ for some $n$), whereas $f(x) > 0$ for any $x \gg 0$, and many such $x$ are feasible (simply take $x_n = m/(Np_n)$ for each $n$).

2. The constraints are convex (they are linear).

3. Slater holds: take $x_n = m/(2Np_n)$ for each $n$.

$\square$

Finally, the results here have almost immediate analogs for MIN problems, exchanging “concave” with “convex” and vice versa.