ADDENDUM TO CORRESPONDENCE # 21:
LOCAL KUHN-TUCKER THEOREM

This is an addendum to Correspondence # 21, where I formulate and provide the solution to the local version of Problem 2 there (which can also be viewed as extension of Problem 1 to constrained optimization with inequality constraints), which I call here Problem 1b. After formulating Problem 1b, I provide some background material on positive cones, useful not only for Problem 1b but also for Problem 2 of Correspondence # 21.

First let me introduce the problem formulation for local optimization with inequality constraints. The problem is formulated as one of minimization, but the development and all results below (in connection with this problem) apply with obvious modifications to one of maximization (essentially replace f with $-f$).

Problem 1b: Let X and Z be linear vector spaces with Z normed (and X not necessarily normed), and $f : X \to \mathbf{R}$, $G : X \to Z$ be two given functions. **Obtain** first-order necessary conditions for a vector $x^\circ \in X$ to minimize $f(x)$ locally subject to the inequality constraint $G(x) \leq \theta$, where θ is the *zero* element of Z and the inequality is with respect to a positive cone P in Z.

To bring concrete meaning to the formulation of *Problem 1b*, I need to make precise the notion of a *positive cone* in Z, as well as its counterpart in the dual space Z^*:

Definition 1b: (*Positive cones*). Let P be a subset of a normed linear vector space Z, which is also a cone with $\theta \in P$. For z_1, z_2, we write $z_1 \geq z_2$ (with respect to P) if $z_1 - z_2 \in P$. The cone P defining this relationship is the positive cone in Z. Now, if P is a positive cone with respect to Z, the set

$$P^\oplus := \{z^* \in Z^* :< z, z^* > \geq 0 \ \forall z \in P\}$$

is a positive cone in the dual space Z^*.

The following two facts follow directly from the definition above.

Facts:

 (i) $P^\oplus$ is closed even if P is not.

(ii) If P is closed, and $z \in Z$ satisfies

$$< z, z^* > \geq 0 \ \ \forall z^* \geq \theta \ (\text{equivalently}, \forall z^* \in \mathrm{P}^\oplus \text{ or } \forall z^* \in \mathrm{Z}^*_{\geq \theta})$$

then $z \geq \theta$ (equivalently, $z \in P$).

Note: Verification of the second result above uses the separating hyperplane theorem; see Text, p. 215.

To provide the answer to *Problem 1b*, we also need the notion of *regularity* of the inequality constraints (as counterpart of Definition 1 of Correspondence # 21):

Definition 1c. *(Regularity of inequality constraints).* *Let the positive cone P, as defined above, have a nonempty interior, and G defining the inequality constraints have a Gateaux differential, linear in the increment. Then, $x^\circ \in X$ is a* **regular point** *of the inequality constraint $G(x) \le \theta$, if*

(i) $G(x^\circ) \le \theta$, and

(ii) $G(x^\circ) + \delta G(x^\circ; h) < \theta$, for some $h \in X$

The solution to Problem 1b (first-order necessary condition) is now given by the following theorem.

Theorem 1b (Generalized (Local) K-T Theorem). *Consider the statement of Problem 1b. Let f and G be Gateaux-differentiable, linear in the increments (i.e., $\delta f(x; h)$ and $\delta G(x; h)$ are both linear in $h \in X$). Let $x^\circ \in X$ solve Problem 1b, and be a regular point of the inequality constraint. Then, there exists $z_o^* \in Z^*$, $z_o^* \ge \theta$, such that the Lagrangian*

$$L(x; z_o^*) := f(x) + <G(x), z_o^*>$$

is stationary at $x = x^\circ$, that is

$$\delta L(x^\circ; z_o^*, h) := \delta f(x^\circ; h) + <\delta G(x^\circ; h), z_o^*> = 0, \quad \forall h \in X,$$

and

$$<G(x^\circ), z_o^*> = 0.$$

If X is normed, and f and G are Fréchet differentiable, then

$$f'(x^\circ) + <G'(x^\circ), z_o^*> = \theta \quad \text{and} \quad <G(x^\circ), z_o^*> = 0$$

Proof. Introduce the two sets

$$A := \{(r, z) : r \ge \delta f(x^\circ; h), \ z \ge G(x^\circ) + \delta G(x^\circ; h) \text{ for some } h \in X\} \subset \mathbf{R} \times Z$$

$$B := \{(r, z) : r \le 0, z \le \theta\} \subset \mathbf{R} \times Z$$

Note that (i) B has a nonempty interior (since P has), (ii) $A \cap \text{inner}(B) = \emptyset$ (since otherwise the optimality of x° is violated), and (iii) both A and B are convex cones. It then follows that there exists a hyperplane that separates A and B, which implies that there exists a triple (r_o, z_o^*, β) such that

$$r_o r + <z, z_o^*> \ge \beta \ \forall (r, z) \in A \qquad \text{and} \qquad r_o r + <z, z_o^*> \le \beta \ \forall (r, z) \in B. \tag{$*$}$$

Since $(0, \theta)$ belongs to both A and B, $\beta = 0$. Then, from the definition of B, $r_o \ge 0$, $z_o^* \ge \theta$.

r_o cannot be zero, because (from regularity) there exists h such that $G(x^\circ) + \delta G(x^\circ; h) < \theta$, implying that $(r_o > 0, z < \theta)$ belongs to A, and hence if $r_o = 0$ we would have $<z, z_o^*> \ge 0$, and also $<z, z_o^*> \le 0$ (since $z_o^* \ge \theta$), which is possible only if $z_o^* = \theta$ – no hyperplane!

Therefore, $r_o > 0$, and without any loss of generality we can take it to be 1. Let $r = \delta f(x^\circ; h)$, $z = G(x^\circ) + \delta G(x^\circ; h)$ in $(*)$ with $\beta = 0$, which leads to

$$\delta f(x^\circ; h) + <G(x^\circ), z_o^*> + <\delta G(x^\circ; h), z_o^*> \ge 0 \ \forall h \in X$$

Let $h = \theta$ in the above, which leads to $<G(x^\circ), z_o^*> = 0$. Since $\delta f(x^\circ; h)$ and $\delta G(x^\circ; h)$ are linear in h, it follows that $\delta f(x^\circ; h) + <\delta G(x^\circ; h), z_o^*> = 0$, and this completes the proof. $\diamond$

$$\diamond \qquad \diamond \qquad \diamond$$