Abraham Charnes; William Wager Cooper; Kenneth O. Kortanek
Semi-infinite programming, differentiability and geometric programming: Part II


Persistent URL: http://dml.cz/dmlcz/103204

Terms of use:
© Institute of Mathematics AS CR, 1969

Institute of Mathematics of the Academy of Sciences of the Czech Republic provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these Terms of use.

This paper has been digitized, optimized for electronic delivery and stamped with digital signature within the project DML-CZ: The Czech Digital Mathematics Library http://dml.cz
SEMI-INFINITE PROGRAMMING, DIFFERENTIABILITY AND GEOMETRIC PROGRAMMING: PART II

A. CHARNES*, W. W. COOPER**, and K. O. KORTANEK***

(Received March 20, 1967)

We propose to specialize the CCK duality theory\(^1\)), which associates as dual problems minimization of an arbitrary convex function over an arbitrary convex set in \(n\)-space with maximization of a linear function in non-negative variables of a generalized finite sequence space subject to a finite system of linear equations, to derive Kuhn-Tucker Theorem\(^2\)) extensions in situations involving (partial) differentiability of objective and constraint functions. There are several ways to procure such generalizations as, for example, by means of non-differentiable analogs of quasi-saddle point conditions or in terms of a saddle point criterion itself. Since we are interested here in exploring extensions which involve some differentiability conditions, we shall proceed via the former course especially since these conditions themselves are analogs of first order conditions of the saddle point criterion.\(^3\))

For our purposes then, let \(f(u)\), and \(G(u) = (g_1(u), g_2(u), \ldots, g_m(u))\) be defined over an open convex set \(K\) in \(R^n\). We shall say that \(f(u)\) is simple piecewise differentiably convex if \(f(u) = \max_{j=1,2,\ldots,N} \{f^{(j)}(u)\}\), where \(f^{(j)}(u)\) is continuously differentiable and convex over \(K\). We shall assume that \(G(u)\) is continuously differentiable and concave, but the extension to simple piecewise concave functions will become apparent during the course of proof for functions of this class.

\(^1\)) See Charnes-Cooper-Kortanek [4] and [5].

\(^2\)) See Kuhn-Tucker [7] and Arrow-Hurwicz-Uzawa [1].

\(^3\)) See Arrow-Hurwicz-Uzawa, ibid., where the authors show that in the case of differentiability the quasi-saddle point condition implies the saddle point condition.
Theorem. (Generalized Quasi-Saddle Point Theorem for Simple Piecewise Differentiably Convex Functions). Let \( f(u) \) and \( G(u) \) have the properties defined above and consider the minimization problem

\[
\min f(u)
\]

subject to

\[
G(u) \geq 0 .
\]

Assume the constraint set \( C = \{ u | G(u) \geq 0 \} \) has an interior point.\(^4\) Then \( u^* \) in \( C \) is an optimal solution to the minimization problem if and only if there exists positive vectors

\[
\eta^* = (\eta^{(1)}_*, \eta^{(2)}_*, \ldots, \eta^{(N)}_*) \quad \text{and} \quad \lambda^* = (\lambda^{(1)}_*, \ldots, \lambda^{(m)}_*)
\]

such that the following properties hold:

\[
\begin{align*}
(1) & \quad - \sum_{j=1}^{N} (\partial f^{(j)}|_{u^*}) \eta^{(j)}_* + \sum_{i=1}^{m} (\partial g^{(i)}|_{u^*}) \lambda^{(i)}_* = 0 \\
(2) & \quad \sum_{j \in J} \eta^{(j)}_* = 1 \\
(3) & \quad G(u^*)^T \lambda^* = 0, \quad \text{and} \quad G(u^*) \geq 0 \quad \text{for} \quad J = \{ j | f^{(j)}(u^*) = f(u^*) \}
\end{align*}
\]

Preliminary Lemmas on Canonical Closure for Differential Systems. By introducing support systems for both objective and constraint functions, we obtain the following equivalent semi-infinite problem (I) with semi-infinite dual (II), which, for the moment, we write in general form.

\[
\begin{align*}
\text{(I)} & \quad \min \ z \\
& \quad z - u^T Q_\alpha \geq d_\alpha, \ \alpha \in A \\
& \quad u^T P_i \geq c_i, \ i \in I
\end{align*}
\]

\[
\begin{align*}
\text{(II)} & \quad \max \sum_{z} d_\alpha \eta_\alpha + \sum_{i} c_i \lambda_i \\
& \quad \sum_{z} \eta_\alpha = 1 \\
& \quad - \sum_{z} Q_\alpha \eta_\alpha + \sum_{i} P_i \lambda_i = 0 \\
& \quad \eta_\alpha, \lambda_i \geq 0.
\end{align*}
\]

\(^4\) This type of constraint qualification has strong intuitive appeal especially in the case of non-differentiability. However, it is known that non-differentiable analogs to the most general constraint qualification for which differentiable Lagrangian techniques are valid (see [6]) involve support systems which are themselves Farkas-Minkowski systems. (See [4] and [5]).

\(^5\) Notationally speaking, \( \partial f|_{u^*} \) is the gradient of \( f \) evaluated at \( u^* \). We use superscripts to correspond to functions and subscripts to correspond to elements in the index set. Thus, \( \partial f^{(j)}_\alpha \) denotes the gradient of \( f^{(j)} \) evaluated at the point \( \alpha \in A \). For convenience, "\( \alpha \in A \) may be identified with "\( u_\alpha \in A \)" when \( A \subseteq R^n \).
Recall that a system of linear inequalities is *canonically closed* if it has interior points and the coefficient set is compact.\(^6\) We need the following lemma.

**Lemma 1.** Suppose that the system is canonically closed and that \(u_*\) solves (I), i.e., the minimum \(z_* = f(u)_*\) is attained. Then in the dual expression, (II), for \(z_*\), the only supports which arise are those passing through the point \((z_*, u_*)\), i.e., the only support planes with \(\eta_a^* = 0\) and \(\lambda_i^* = 0\) are those for which \(z_* = u_*^TQ_x + d_x\) and \(u_*^TP_i = c_i\).

**Proof.** By the extended dual theorem, there exist \(\eta, \lambda\) such that

\[
z_* = \sum_a d_a \eta_a^* + \sum c_i \lambda_i^*.
\]

We must show that if \(\eta_a^* > 0\), then \(z_* = u_*^TQ_x + d_x\) and if \(\lambda_i^* > 0\), then \(u_*^TP_i = c_i\).

First, \(z_* - u_*^TQ_x \geq d_x\), for all \(\eta\). Hence

\[
\sum_a d_a \eta_a^* \leq \sum_a z_a^* \eta_a^* - \sum_a (u_*^TQ_x) \eta_a^* = z_* - \sum_a (u_*^TQ_x) \eta_a^*.
\]

Therefore,

\[
z_* = \sum_a d_a \eta_a^* + \sum_i c_i \lambda_i^* \leq z_* - \sum_a (u_*^TQ_x) \eta_a^* + \sum_i c_i \lambda_i^*.
\]

i.e.,

\[
(A) \quad -\sum_a (u_*^TQ_x) \eta_a^* + \sum_i c_i \lambda_i^* \geq 0.
\]

On the other hand, by dual feasibility,

\[
u_*^T[-\sum_a Q_x \eta_a^* + \sum_i P_i \lambda_i^*] = u_*^T(0) = 0.
\]

However, since \(u_*^TP_i \geq c_i\) for all \(i\), we can rewrite this as follows:

\[
(B) \quad 0 = \sum_a -u_*^TQ_x \eta_a^* + \sum_i u_*^TP_i c_i \geq -\sum_a u_*^TQ_x \eta_a^* + \sum_i c_i \lambda_i^*.
\]

Therefore combining (A) and (B) we have,

\[
\sum_a u_*^TQ_x \eta_a^* = \sum_i c_i \lambda_i^*.
\]

Two conclusions follow:

\[
(C_1) \quad z_* = \sum_a [u_*^TQ_x + d_x] \eta_a^*\), where \(\sum_a \eta_a^* = 1\), \(\eta \geq 0\), and \(z_* \geq u_*^TQ_x + d_x\). Hence \(z_* = u_*^TQ_x + d_x\) for every \(\alpha\) with \(\eta_a^* > 0\).
\]

\[
(C_2) \quad \sum_i u_*^TP_i \lambda_i^* = \sum_i c_i \lambda_i^* \Rightarrow \sum_i (u_*^TP_i - c_i) \lambda_i^* = 0.
\]

Hence \(\lambda_i^* > 0\) implies \(u_*^TP_i = c_i\).

\(^6\) See [4] and [5]. Note that canonical closure is a sufficient condition but not necessary for the validity of the extended dual theorem as pointed out in [4].
Proof of Theorem. With respect to the minimization problem of the Theorem, consider the particular semi-infinite equivalent

\[
\begin{align*}
\min z \\
\text{subject to } z - u^T \partial f_s^{(j)} &\geq f^{(j)}(u_a) - u^T \partial f_s^{(j)}, \quad j = 1, 2, \ldots, N \\
u^T \partial g_s^{(i)} &\geq -g^{(i)}(u_a) + u^T \partial g_s^{(i)}(u_a), \quad i = 1, 2, \ldots, m
\end{align*}
\]

for all \( \alpha \in A \), where \( A \) is some index set in \( \mathbb{R}^n \) (e.g. the convex constraint set \( C \)). Since \( C \) has interior points, it follows that this linear inequality system also does. Form a canonical normalization\(^7\), (i.e., divide each inequality by a positive constant to make the sum of the absolute values of the coefficients sum to 1), to obtain an equivalent system with bounded coefficients and interiority.

\[
\begin{align*}
\min z \\
\text{subject to } \mu_a^{(j)} z - u^T \partial f_s^{(j)} z &\geq f^{(j)}(u_a) \mu_a^{(j)} - u^T \partial f_s^{(j)} \mu_a^{(j)}, \quad \mu_a^{(j)} > 0 \\
u^T \partial g_s^{(i)} v_s^{(i)} &\geq -g^{(i)}(u_a) v_s^{(i)} + u^T \partial g_s^{(i)}(u_a) v_s^{(i)}, \quad v_s^{(i)} > 0
\end{align*}
\]

where \( j = 1, 2, \ldots, m \), and \( \alpha \in A \).

Now form a canonical closure by possibly enlarging the index set to \( \bar{A} \supseteq A \) and adjoining the corresponding limiting inequalities which are of the form:

\[
\begin{align*}
\mu_a^{(j)} z - u^T Q_s^{(j)} &\geq d_s^{(j)} \\
u^T P_s^{(i)} &\geq c_s^{(i)} \quad \text{for } \alpha \in \bar{A} - A.
\end{align*}
\]

Let (I) denote this new canonically closed equivalent (which differs from (I) by only these possibly adjoined inequalities and also has interior points).

Now if \( u^* \) is optimal for (I) it is also optimal for the canonically closed equivalent (I) and lemma 1 applies. However, any of the possibly newly adjoined inequalities which are actively involved in the dual are positive multiples of differential hyperplanes already in the system, for suppose one of them has a \( \lambda^{(j)} > 0 \), say, \( \mu_a^{(j)} z - u^T Q_s^{(j)} \geq d_s^{(j)} \) with \( \alpha \in \bar{A} - A \). Then by lemma 1, the support plane \( \mu_a^{(j)} z - u^T Q_s^{(j)} = d_s^{(j)} \) contains the point \( (u^*, f(u^*)) \) i.e., \( \mu_a^{(j)} f(u^*) = \mu_a^{(j)} z_a = u^T Q_s^{(j)} + d_s^{(j)} \) or equivalently, the plane \( \mu_a^{(j)} z = u^T Q_s^{(j)} + d_s^{(j)} \) is tangent to the surface \( z = f^{(j)}(u) \) at the point \( u^* \). Since \( f^{(j)}(u) \) is continuously differentiable, and since \( \mu_a^{(j)} z \geq u^T Q_s^{(j)} + d_s^{(j)} \) over \( C \), this tangent plane is unique up to a constant positive multiple, and therefore we do not need to adjoin these additional inequalities. A similar argument obviously holds for the constraint functions.

\(^7\) See [5], p. 114.
We now present the semi-infinite dual (II) and derive the conditions of the theorem.

\[
\max \sum_j \sum_a \left[ f^{(j)}(u_a) \mu_a^{(j)} - u_a^T \frac{\partial f^{(j)}}{\partial u_a} \mu_a^{(j)} \right] \bar{\eta}_a^{(j)} + \sum_j \sum_a \left[ -g^{(i)}(u_a) v_a^{(i)} + u_a^T \frac{\partial g^{(i)}}{\partial u_a} v_a^{(i)} \right] \bar{\lambda}_a^{(i)}
\]

subject to

\[
\sum_j \mu_a^{(j)} \bar{\eta}_a^{(j)} = 1
\]

\[
-\sum_j \sum_a \left( \frac{\partial f^{(j)}}{\partial u_a} \mu_a^{(j)} \right) \bar{\eta}_a^{(j)} + \sum_i \sum_a \left( \frac{\partial g^{(i)}}{\partial u_a} v_a^{(i)} \right) \bar{\lambda}_a^{(i)} = 0
\]

and

\[
\bar{\eta}, \bar{\lambda} \geq 0.
\]

By the dual theorem there exists a dual optimal solution \((\bar{\eta}^*, \bar{\lambda}^*)\). By lemma 1 \(\bar{\eta}^*\) has non-zero coordinates corresponding only to support planes passing through the optimum \(u^*, z^*_\), i.e., those gradient tangent planes at this point, one for each function \(f^{(j)}\). This also applies to \(\bar{\lambda}^*\) and constraint functions \(g^{(i)}\), and therefore we may write \(\bar{\eta}^* = (\bar{\eta}^*_1, \ldots, \bar{\eta}^*_N)\) and \(\bar{\lambda}^* = (\bar{\lambda}^*_1, \ldots, \bar{\lambda}^*_m)\). Thus, upon setting \(\eta^{(j)}_a = \mu_a^{(j)} \bar{\eta}_a^{(j)}\) for \(j = 1, \ldots, N\) and \(\lambda^{(i)}_a = v_a^{(i)} \bar{\lambda}_a^{(i)}\) for \(i = 1, \ldots, m\), we obtain the following dual optimal conditions:

\[
(1) \quad -\sum_j \frac{\partial f^{(j)}(u^*)}{\partial u_a} \eta^{(j)}_a + \sum_i \frac{\partial g^{(i)}(u^*)}{\partial u_a} \lambda^{(i)}_a = 0
\]

and

\[
(2) \quad \sum_i \eta^{(i)}_a = 1
\]

where all \(\eta^{(j)}_a\) and \(\lambda^{(i)}_a\) \(\geq 0\).

The equality of dual functionals yields,

\[
f(u^*) = z^*_a = \sum_j f^{(j)}(u^*) \eta^{(j)}_a - \sum_j u^*_a^T \frac{\partial f^{(j)}}{\partial u_a} \eta^{(j)}_a + \sum_i u^*_a^T \frac{\partial g^{(i)}}{\partial u_a} \lambda^{(i)}_a + \sum_i u^*_a^T \frac{\partial g^{(i)}}{\partial u_a} \lambda^{(i)}_a
\]

Since \(f^{(j)}(u^*) \leq f(u^*)\) for all \(j\) and \(g^{(i)}(u^*) \leq g^{(i)}(u^*)\) for all \(i\), it therefore follows that, (3)

\[
\sum_i g^{(i)}(u^*) \lambda^{(i)} = 0.
\]

Furthermore, since \(f(u^*) = \max \{f^{(j)}(u^*)\}\), it follows that \(\eta^{(j)}_a = 0\) whenever \(f^{(j)}(u^*) < f(u^*)\), giving condition (2). Thus, the three conditions of the theorem are proved.

On the other hand, given positive vectors \(\eta^*\) and \(\lambda^*\) satisfying conditions (1), (2), and (3) with respect to \(u^*\), then since \(\mu^{(j)}_a = 0\) and \(\lambda^{(i)}_a = 0\) in the canonically closed system (I), we obtain dual feasible solutions upon setting \(\bar{\eta}^{(j)}_a = \eta^{(j)}_a / \mu^{(j)}_a\) and \(\bar{\lambda}^{(i)}_a = \lambda^{(i)}_a / v^{(i)}\). Furthermore, the dual objective value is \(\sum f^{(j)}(u^*) \eta^{(j)}_a\), and condition (2) implies that \(f(u^*) = \sum f^{(j)}(u^*) \eta^{(j)}_a\) giving dual equality of objective functions, thereby proving that \((f(u^*), u^*)\) is optimal.
Our generalization of the quasi-saddle point version of the Kuhn-Tucker Theorem is not as general as we may possibly get, but it does indicate a unified approach to study these equivalences under more general circumstances. In fact, we are already obtaining results for generalized saddle-point equivalence theorems for arbitrary convex functions over \( \mathbb{R}_n \). This is the subject of another paper and will be reported on elsewhere.

Already these methods have shown that the crucial property of the constraint functions is the Farkas-Minkowski property, which is a property of the functions themselves expressed in terms of finite positive linear combinations of their "gradients". Geometric qualifications are sufficient restrictions on the constraining functions to permit such Farkas-Minkowski expressions. In general however, it may be necessary to go beyond the natural gradient inequalities provided by the constraint functions to obtain strong duality results.

In conclusion, we illustrate this now by constructing a canonically closed equivalent for the one-variable Slater example by adjoining a new variable to the gradient inequality system following the methods of our regularization procedures for semi-infinite programs.\(^8\) Restating the Slater example, we have:

\[
\begin{align*}
(1) \\
\min x \\
\text{subject to } -(1 - x)^2 \geq 0
\end{align*}
\]

with unique optimum \( x^* = 1 \). Introducing a differential system of supports to contain the optimum, we obtain the equivalent problem:

\[
\begin{align*}
(1) \\
\min x \\
\text{subject to } 2(1 - \alpha) x \geq 1 - \alpha^2
\end{align*}
\]

for \( 0 \leq \alpha \leq 2 \).

Let \( M \) and \( V \) be large positive numbers, either real or non-Archimedean, i.e. larger than any real number\(^9\), and construct the following semi-infinite dual regularizations.

\[
\begin{align*}
(I_R) \\
\min Mt + x \\
\text{subject to } t + 2(1 - \alpha) x \geq 1 - \alpha^2, \quad 0 \leq \alpha \leq 2 \\
x \geq -V \\
-x \geq -V
\end{align*}
\]

\(^8\) See Slater [7], [4] p. 216, and [5], p. 119.

(II_R)

\[
\max_x \sum_a (1 - a^2) \lambda_a - V\lambda^+ - V\lambda^-
\]
subject to \( \sum_a \lambda_a = M \)

\[
\sum_a 2(1 - a) \lambda_a + \lambda^+ - \lambda^- = 1
\]

\( \lambda^+ \geq 0 \)

Observe that problem (I_R) is canonically closed and that \( t \geq 0 \) is included in the inequality system and corresponds to the index point \( a = 1 \). As stated above, \( M \) may be viewed as real or non-Archimedean, and therefore we shall derive dual optimal solutions for (I_R) and (II_R) in a manner which is valid for either case.

We know that \( (t, x) = (0, 1) \) is (I_R)-feasible with functional value 1. Thus, we search for a solution \( (t^*, x^*) \) with objective value < 1, if it exists, and therefore we assume \( x^* < 1 \). By lemma 1, this optimum involves only support planes which are tangent to it and therefore involves only its own gradient inequality with index point \( a^* = x^* \). But this implies \( t^* = (1 - a^*)^2 \) yielding (I_R)-objective value \( M(1 - a^*)^2 + a^* \).

Applying the usual differential methods for finding a minimum to this function yields the Taylor expansion,

\[
M(1 - a^*)^2 + a^* = \frac{4M - 1}{4M} + M \left( a^* - \frac{2M - 1}{2M} \right)^2 \text{ for } 0 \leq a^* \leq 2,
\]

an equation which is obviously valid for arbitrary \( M \). This tells us to take \( a^* = \frac{(2M - 1)}{2M} \) to obtain minimum objective value \( \frac{4M - 1}{4M} < 1 \). Furthermore, the point \( (t^*, x^*) = (1/4M^2, (2M - 1)/2M) \) is (I_R)-feasible because

\[
t \geq \frac{1}{4M^2} - \left( a - \frac{2M - 1}{2M} \right)^2 = 1 - a^2 - 2(1 - a)x^* \text{ for } 0 \leq a \leq 2,
\]

which is a restatement of (I_R)-feasibility. But taking \( \lambda_{x^*} = M \), the dual variable associated with the binding constraint, and \( \lambda_a = 0 \) for \( a \neq a^* \) and \( \lambda^+ = \lambda^- = 0 \), yields a dual (II_R)-solution with equality of dual objective functions, and therefore shows that in fact the two solutions form dual optimal solutions for problems (I_R), (II_R) whether \( M \) is viewed as real or non-Archimedean.

Observe that the dual solution, \( \lambda^* \), is an extreme point of the associated generalized finite sequence space\(^{10} \) and as such the non-zero coordinate is linear and homogeneous in \( M \)\(^{11} \), in particular, \( \lambda_{x^*} = M \). Two courses of action with respect to \( M \)

\(^{10} \) See [4] p. 211.

\(^{11} \) See [2], where this statement was first proved for finite linear programming over non-Archimedean ordered fields.
are now open to us. First, if $M$ is real, we may let $M \to \infty$ so that $(t^*, x^*) \to (0, 1)$, the solution to the Slater problem, with corresponding dual variable characterized by $\lambda_{x^*} \to \infty$. Second, viewing $M$ as non-Archimedean, we obtain dual optimal solutions in Hilbert’s field with common objective value $1 - 1/4M$ which in the extended ordering is larger than any real number less than 1, but itself is less than 1.

References


Souhrn

SEMIINFINITNÍ PROGRAMOVÁNÍ, DIFERENCOVATELNOST A GEOMETRICKÉ PROGRAMOVÁNÍ: ČÁST II

A. CHARNES, W. W. COOPER, K. O. KORTANEK

Autoři se v článku zabývají jistou specializací své teorie duality na případ, kdy cílová funkce je spojitě diferencovatelná a konvexní na množině $K$ přípustných řešení a funkce omezení definující $K$ jsou spojitě diferencovatelné a konkávní. V článku je dále ukázaná cesta jak výklad zobecnit na případ, kdy funkce v omezeních problému jsou po částech diferencovatelné a konkávní. Získané podmínky lze chápat jako rozšíření Kuhn-Tucherovy věty.