Discussion Paper No. 252

FENCHEL'S DUALITY THEOREM IN GENERALIZED GEOMETRIC PROGRAMMING

by

Elmor L. Peterson

October, 1976

Fenchel's Duality Theorem in Generalized Geometric Programming

by

Elmor L. Peterson*

Abstract. Fenchel's duality theorem is extended to generalized geometric programming with explicit constraints -- an extension that also generalizes and strengthens Slater's version of the Kuhn-Tucker theorem.

Key words: Fenchel's duality theorem, generalized geometric programming, convex programming, ordinary programming, Slater's constraint qualification, Kuhn-Tucker theorem.

TABLE OF CONTENTS

*Department of Industrial Engineering/Management Sciences and Department of Mathematics, Northwestern University, Evanston, Illinois 60201. Research sponsored by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under Grant No. AFOSR-77-3134.

1. <u>Introduction</u>. Although many implications of this extension have already been discussed in the author's recent survey paper [1], a proof of it is given here for the first time.

This proof utilizes the unconstrained version that has already been established by independent and somewhat different arguments in [2] and [3]. In doing so, it exploits the main result from [4] and also requires some of the convexity theory in [3]--especially the theory having to do with the "relative interior" (ri S) of an arbitrary convex set $S \subseteq E_N$ (N-dimensional Euclidean space).

2. <u>The unconstrained case</u>. We begin with the following notation and hypotheses:

$\mathcal{X}$ is a nonempty closed convex cone in E_n,

g is a (proper) closed convex function with a nonempty (effective) domain $\mathcal{C} \subseteq E_n$.

Now, given $\mathcal{X}$ and g, consider the resulting "geometric programming problem" $\mathcal{A}$.

PROBLEM $\mathcal{A}$. <u>Using the feasible solution set</u>

$$\mathcal{S} \stackrel{\Delta}{=} \mathcal{X} \cap \mathcal{C},$$

<u>calculate both the problem infimum</u>

$$\varphi \stackrel{\Delta}{=} \inf_{x \in \mathcal{S}} g(x)$$

<u>and the optimal solution set</u>

$$\mathcal{S}^* \stackrel{\Delta}{=} \{x \in \mathcal{S} \mid g(x) = \varphi\}.$$

Geometric duality is defined in terms of both the "dual cone"

$$\mathcal{Y} \stackrel{\Delta}{=} \{y \in E_n \mid 0 \le \langle x,y \rangle \text{ for each } x \in \mathcal{X}\}$$

and the "conjugate transform function" h whose (effective) domain

$$\mathcal{D} \stackrel{\Delta}{=} \{y \in E_n \mid \sup_{x \in \mathcal{C}} [\langle y,x \rangle - g(x)] \text{ is finite}\}$$

and whose functional value

$$h(y) \stackrel{\Delta}{=} \sup_{x \in \mathcal{C}} [\langle y,x \rangle - g(x)].$$

In particular, given the geometric programming problem $\mathcal{A}$, consider the resulting "geometric dual problem" $\mathcal{B}$.

PROBLEM $\mathcal{B}$. <u>Using the feasible solution set</u>

$$\mathcal{T} \stackrel{\Delta}{=} \mathcal{Y} \cap \mathcal{D},$$

<u>calculate both the problem infimum</u>

$$\psi \stackrel{\Delta}{=} \inf_{y \in \mathcal{T}} h(y)$$

<u>and the optimal solution set</u>

$$\mathcal{T}^* \stackrel{\Delta}{=} \{y \in \mathcal{T} \mid h(y) = \psi\}.$$

Fenchel's duality theorem in the context of dual problems $\mathcal{A}$ and $\mathcal{B}$ is one of the most important theorems in geometric programming. It can be stated in the following way.

Theorem 1. *If problem $\mathcal{B}$ has both a feasible solution $y^{\circ} \in (\mathrm{ri}\, \mathcal{Y}) \cap (\mathrm{ri}\, \mathcal{D})$ and a finite infimum ψ, then*

(I) *problem $\mathcal{A}$ has both a nonempty feasible solution set $\mathcal{S}$ and a finite infimum φ, and*

$$0 = \varphi + \psi,$$

(II) *problem $\mathcal{A}$ has a nonempty optimal solution set $\mathcal{S}^*$.*

This theorem is established as Theorem 31.4 on page 335 of [3].

The implications of Theorem 1 are given on page 26 of [1]. An important extension of it is established in the next section.

3. *The constrained case.* To incorporate explicit constraints into generalized geometric programming, we introduce the following notation and hypotheses:

I and J are two nonintersecting (possibly empty) positive-integer index sets with finite cardinality $o(I)$ and $o(J)$ respectively;

x^k and y^k are independent vector variables in E_{n_k} for $k \in \{0\} \cup I \cup J$, and x^I and y^I denote the respective Cartesian products of the vector variables x^i, $i \in I$, and y^i, $i \in I$ while x^J and y^J denote the respective Cartesian products of the vector variables x^j, $j \in J$, and y^j, $j \in J$; so the Cartesian products $(x^0, x^I, x^J) \stackrel{\Delta}{=} x$ and $(y^0, y^I, y^J) \stackrel{\Delta}{=} y$ are independent vector variables in E_n, where

$$n \stackrel{\Delta}{=} n_0 + \sum_I n_i + \sum_J n_j;$$

α and λ are independent vector variables with respective components α_i and λ_i for $i \in I$, and β and κ are independent vector variables with respective components β_j and κ_j for $j \in J$;

X and Y are nonempty closed convex dual cones in E_n, and g_k and h_k are (proper) closed convex conjugate functions with respective (effective) domains $C_k \subseteq E_{n_k}$ and $D_k \subseteq E_{n_k}$ for $k \in \{0\} \cup I \cup J$.

Now, let

$$\mathcal{X} \overset{\Delta}{=} \{(x^0, x^I, \alpha, x^J, \kappa) \in E_{\eta} \mid (x^0, x^I, x^J) \in X;\ \alpha = 0;\ \kappa \in E_{o(J)}\},$$

where $n + o(I) + o(J) = \eta$. In addition, let

$$\mathcal{C} \overset{\Delta}{=} \{(x^0, x^I, \alpha, x^J, \kappa) \in E_{\eta} \mid x^0 \in C_0;\ x^i \in C_i,\ \alpha_i \in E_1, \text{ and}$$

$$g_i(x^i) + \alpha_i \le 0,\ i \in I;\ (x^j, \kappa_j) \in C_j^+,\ j \in J\},$$

and let

$$\mathcal{g}(x^0, x^I, \alpha, x^J, \kappa) \overset{\Delta}{=} g_0(x^0) + \sum_J g_j^+(x^j, \kappa_j),$$

where the (closed convex) function g_j^+ has a domain

$$C_j^+ \overset{\Delta}{=} \{(x^j, \kappa_j) \mid \text{either } \kappa_j = 0 \text{ and } \sup_{d^j \in D_j} \langle x^j, d^j \rangle < +\infty, \text{ or } \kappa_j > 0 \text{ and } x^j \in \kappa_j C_j\}$$

and functional values

$$g_j^+(x^j, \kappa_j) \overset{\Delta}{=} \begin{cases} \sup_{d^j \in D_j} \langle x^j, d^j \rangle & \text{if } \kappa_j = 0 \text{ and } \sup_{d^j \in D_j} \langle x^j, d^j \rangle < +\infty \\ \\ \kappa_j g_j(x^j/\kappa_j) & \text{if } \kappa_j > 0 \text{ and } x^j \in \kappa_j C_j. \end{cases}$$

The resulting problem $\mathcal{A}$ can clearly be stated in the following way.

PROBLEM A. Consider the objective function G whose domain

$$C \stackrel{\Delta}{=} \{(x,\kappa) \mid x^k \in C_k,\ k \in \{0\} \cup I, \text{ and } (x^j, \kappa_j) \in C_j^+,\ j \in J\}$$

and whose functional value

$$G(x,\kappa) \stackrel{\Delta}{=} g_0(x^0) + \sum_J g_j^+(x^j, \kappa_j).$$

Using the feasible solution set

$$S \stackrel{\Delta}{=} \{(x,\kappa) \in C \mid x \in X, \text{ and } g_i(x^i) \leq 0,\ i \in I\},$$

calculate both the problem infimum

$$\varphi \stackrel{\Delta}{=} \inf_{(x,\kappa) \in S} G(x,\kappa)$$

and the optimal solution set

$$S^* \stackrel{\Delta}{=} \{(x,\kappa) \in S \mid G(x,\kappa) = \varphi\}.$$

Now, section 3 of [4] shows that

$$\mathcal{Y} = \{(y^0, y^I, \lambda, y^J, \beta) \in E_n \mid (y^0, y^I, y^J) \in Y;\ \beta = 0,\ \lambda \in E_{o(I)}\}.$$

Section 3 of [4] also shows that

$$\mathcal{D} = \{(y^0, y^I, \lambda, y^J, \beta) \in E_n \mid y^0 \in D_0;\ (y^i, \lambda_i) \in D_i^+,\ i \in I;\ y^j \in D_j,$$

$$\beta_j \in E_1, \text{ and } h_j(y^j) + \beta_j \leq 0,\ j \in J\},$$

and that

$$h(y^0, y^I, \lambda, y^J, \beta) = h_0(y^0) + \sum_I h_i^+(y^i, \lambda_i),$$

where the (closed convex) function h_i^+ has a domain

$$D_i^+ \overset{\Delta}{=} \{(y^i,\lambda_i) \mid \text{either } \lambda_i = 0 \text{ and } \sup_{c^i \in C_i} \langle y^i, c^i\rangle < +\infty, \text{ or } \lambda_i > 0 \text{ and } y^i \in \lambda_i D_i\}$$

and functional values

$$h_i^+(y^i,\lambda_i) \overset{\Delta}{=} \begin{cases} \sup_{c^i \in C_i} \langle y^i, c^i\rangle & \text{if } \lambda_i = 0 \text{ and } \sup_{c^i \in C_i} \langle y^i, c^i\rangle < +\infty \\ \lambda_i h_i(y^i/\lambda_i) & \text{if } \lambda_i > 0 \text{ and } y^i \in \lambda_i D_i. \end{cases}$$

The resulting problem $\mathcal{B}$ can clearly be stated in the following way.

PROBLEM B. <u>Consider the objective function</u> H <u>whose domain</u>

$$D \overset{\Delta}{=} \{(y,\lambda) \mid y^k \in D_k,\ k \in \{0\} \cup J, \text{ and } (y^i,\lambda_i) \in D_i^+,\ i \in I\}$$

<u>and whose functional value</u>

$$H(y,\lambda) \overset{\Delta}{=} h_0(y^0) + \sum_I h_i^+(y^i,\lambda_i).$$

<u>Using the feasible solution set</u>

$$T \overset{\Delta}{=} \{(y,\lambda) \in D \mid y \in Y, \text{ and } h_j(y^j) \le 0,\ j \in J\},$$

<u>calculate both the problem infimum</u>

$$\psi \overset{\Delta}{=} \inf_{(y,\lambda) \in T} H(y,\lambda)$$

<u>and the optimal solution set</u>

$$T^* \overset{\Delta}{=} \{(y,\lambda) \in T \mid H(y,\lambda) = \psi\}.$$

It is worth noting that dual problems A and B provide the only completely symmetric duality that is presently known for general (closed) convex programming with explicit constraints. Moreover, [1] and some of the references cited therein show that all other duality in convex programming can be viewed as a special case. For the fundamental relations between geometric duality and ordinary Lagrangian duality see [5].

Fenchel's duality theorem in the context of dual problems A and B is one of the most important theorems, as well as one of the deepest theorems, in geometric programming. It can be stated in the following way.

Theorem 2. If

(i) problem B has a feasible solution (y', λ') such that

$$h_j(y'^j) < 0 \qquad j \in J,$$

(ii) problem B has a finite infimum ψ,

(iii) there exists a vector (y^+, λ^+) such that

$$y^+ \in (\text{ri } Y),$$

$$y^{+k} \in (\text{ri } D_k) \qquad k \in \{0\} \cup J,$$

$$(y^{+i}, \lambda_i^+) \in (\text{ri } D_i^+) \qquad i \in I,$$

then

(I) problem A has both a nonempty feasible solution set S and a finite infimum φ, and

$$0 = \varphi + \psi,$$

(II) problem A has a nonempty optimal solution set S*.

Proof. We obviously need only show that the Fenchel hypothesis in Theorem 1 (i.e. the hypothesis that there exists a vector $\mathcal{y}^{\circ} \in (\mathrm{ri}\,\mathcal{Y}) \cap (\mathrm{ri}\,\mathcal{D})$) is equivalent to hypotheses (i) and (iii) in Theorem 2.

Toward that end, we first use the formulas for $\mathcal{Y}$ and $\mathcal{D}$ to derive comparable formulas for $(\mathrm{ri}\,\mathcal{Y})$ and $(\mathrm{ri}\,\mathcal{D})$ -- two derivations that make crucial use of the following basic facts:

(A) $(\mathrm{ri}\ U) = U$ when U is a vector space,

(B) $(\mathrm{ri}\ V) = \times_1^{\eta} (\mathrm{ri}\ V_k)$ when $V = \times_1^{\eta} V_k$ and the sets V_k are convex,

and

(C) $(\mathrm{ri}\ W) = (\mathrm{int}\ W)$, the "interior" of W, when W is a convex set with the same "dimension" as the space in which it is embedded.

Fact (A) is established on page 44 of [3]; fact (B) can be obtained inductively from the formula at the top of page 49 of [3]; and fact (C) is explained on page 44 of [3].

Now, the formula for $\mathcal{Y}$ along with facts (A) and (B) implies that

$$(\mathrm{ri}\,\mathcal{Y}) = \{(y^0, y^I, \lambda, y^J, \beta) \in E_{\eta} \mid (y^0, y^I, y^J) \in (\mathrm{ri}\ Y);\ \ \lambda \in E_{o(I)};\ \ \beta = 0\}.$$

Moreover, the formula for $\mathcal{D}$ along with facts (A) and (B) implies that

$$(\mathrm{ri}\,\mathcal{D}) = \{(y^0, y^I, \lambda, y^J, \beta) \in E_{\eta} \mid y^0 \in (\mathrm{ri}\ D_0);\ \ \lambda_i > 0 \text{ and } y^i \in \lambda_i (\mathrm{ri}\ D_i),$$

$$i \in I;\ y^j \in (\mathrm{ri}\ D_j),\ \ \beta_j \in E_1, \text{ and } h_j(y^j) + \beta_j < 0,\ \ j \in J\},$$

by virtue of both the equation

$$(\mathrm{ri}\ D_i^+) = \{(y^i, \lambda_i) \mid \lambda_i > 0 \text{ and } y^i \in \lambda_i (\mathrm{ri}\ D_i)\}$$

and the equation

$$(\mathrm{ri}\ \{(y^j, \beta_j) \mid y^j \in D_j \text{ and } h_j(y^j) + \beta_j \le 0\}) =$$

$$\{(y^j, \beta_j) \mid \beta_j \in E_1,\ y^j \in (\mathrm{ri}\ D_j), \text{ and } h_j(y^j) + \beta_j < 0\}.$$

To derive the latter equation, simply use Theorem 6.8 on page 49 of [3] along with fact (C). To derive the former equation, first consider the point-to-set mapping $Y_i^+ : \Lambda_i^+$ where

$$Y_i^+[\lambda_i] \overset{\Delta}{=} \{y^i \mid (y^i, \lambda_i) \in D_i^+\}$$

and

$$\Lambda_i^+ \overset{\Delta}{=} \{\lambda_i \mid Y_i^+[\lambda_i] \text{ is not empty}\}.$$

Now, Corollary 6.8.1 on page 50 of [3] implies that

$$(\mathrm{ri}\ D_i^+) = \{(y^i, \lambda_i) \mid \lambda_i \in (\mathrm{ri}\ \Lambda_i^+) \text{ and } y^i \in (\mathrm{ri}\ Y_i^+[\lambda_i])\}.$$

Moreover, the definition of D_i^+ clearly shows that $\Lambda_i^+ = \{\lambda_i \ge 0\}$, which means of course that

$$(\mathrm{ri}\ \Lambda_i^+) = \{\lambda_i > 0\}.$$

Furthermore, for $\lambda_i > 0$ the definition of D_i^+ clearly shows that $Y_i^+[\lambda_i] = \lambda_i D_i$, which means that

$$(\mathrm{ri}\ Y_i^+[\lambda_i]) \equiv \lambda_i (\mathrm{ri}\ D_i) \text{ for } \lambda_i \in (\mathrm{ri}\ \Lambda_i^+),$$

by virtue of Corollary 6.6.1 on page 48 of [3]. Consequently, our derivation of the preceding formula for $(\mathrm{ri}\,\mathcal{D})$ is complete.

In particular then, the Fenchel hypothesis in Theorem 1 simply asserts that

$$\text{there exists a vector } (y^0, y^I, \lambda, y^J, 0) = y^\circ$$
$$\text{such that } (y^0, y^I, y^J) \in (\mathrm{ri}\ Y);\ \ y^0 \in (\mathrm{ri}\ D_0);$$
$$\lambda_i > 0 \text{ and } y^i \in \lambda_i (\mathrm{ri}\ D_i),\ \ i \in I;\ y^j \in (\mathrm{ri}\ D_j)$$
$$\text{and } h_j(y^j) < 0,\ \ j \in J.$$

To complete our proof, we now show that this hypothesis is in fact equivalent to the hypothesis

$$\text{there exists a vector } (y'^0, y'^I, \lambda', y'^J)$$
$$\text{such that } (y'^0, y'^I, y'^J) \in Y;\ y'^0 \in D_0;$$
$$(y'^i, \lambda'_i) \in D_i^+,\ \ i \in I;\ y'^j \in D_j \text{ and } h_j(y'^j) < 0,\ j \in J$$
$$\text{--- and there exists a vector}$$
$$(y^{+0}, y^{+I}, \lambda^+, y^{+J}) \text{ such that}$$
$$(y^{+0}, y^{+I}, y^{+J}) \in (\mathrm{ri}\ Y);\ \ y^{+0} \in (\mathrm{ri}\ D_0);\ \lambda_i^+ > 0$$
$$\text{and } y^{+i} \in \lambda_i (\mathrm{ri}\ D_i),\ \ i \in I;\ y^{+j} \in (\mathrm{ri}\ D_j),\ \ j \in J.$$

Obviously, a vector (y^0, y^I, λ, y^J) that satisfies the former hypothesis satisfies both parts of the latter hypothesis. On the other hand, Theorem 6.1 on page 45 of [3] and Theorem 7.1 on page 51 of [3] imply that a convex combination $\alpha(y'^{0},y'^{I},\lambda',y'^{J})+\beta(y^{+0},y^{+I},\lambda^{+},y^{+J})$ of vectors $(y'^{0},y'^{I},\lambda',y'^{J})$ and $(y^{+0},y^{+I},\lambda^{+},y^{+J})$ that satisfy the latter hypothesis will satisfy the former hypothesis for sufficiently small $\beta>0$. q.e.d.

Although the condition $h_j(y'^{j})<0$, $j\in J$ in hypothesis (i) of Theorem 2 resembles the well-known "Slater constraint qualification", it is of course to be deleted when J is empty -- which is the situation in most applications. However, the analogous condition $g_i(x'^{i})<0$, $i\in I$ in hypothesis (i) of the (unstated) dual of Theorem 2 (obtained from Theorem 2 by interchanging the symbols A and B, the symbols x and y, the symbols κ and λ, the symbols g and h, the symbols i and j, the symbols I and J, the symbols φ and ψ, the symbols X and Y, the symbols C and D, the symbols S and T, and the symbols S* and T*) is essentially the Slater constraint qualification. In fact, we shall now see that the "ordinary programming" case of the dual of Theorem 2 actually strengthens Slater's version of the "Kuhn-Tucker theorem".

The ordinary programming case occurs when

$$J=\emptyset,$$

$$n_k=m \text{ and } C_k \overset{\Delta}{=} C_o \text{ for some set } C_o\subseteq E_m \qquad k\in\{0\}\cup I,$$

and

$$X\overset{\Delta}{=}\text{column space of }\begin{bmatrix}U\\U\\ \cdot\\ \cdot\\ \cdot\\U\end{bmatrix}\quad\begin{array}{l}\text{where there is a total of } 1+o(I)\\ \\ \\ \\ \\ \text{identity matrices U that are } m\times m.\end{array}$$

In particular, an explicit elimination of the vector space condition $x \in X$ by the linear transformation

$$\begin{pmatrix} x^0 \\ x^I \end{pmatrix} = \begin{bmatrix} U \\ U \\ \cdot \\ \cdot \\ \cdot \\ U \end{bmatrix} z$$

shows that the resulting problem A is equivalent to the very general ordinary programming problem

$$\underline{\text{Minimize}}\ g_0(z)\ \underline{\text{subject to}}$$

$$g_i(z) \le 0 \qquad i \in I$$

$$z \in C_o.$$

Now, the Slater constraint qualification for the preceding problem simply requires the existence of a feasible solution z' such that $g_i(z') < 0$, $i \in I$. Moreover, Slater's version of the Kuhn-Tucker theorem asserts that the existence of such a "Slater solution" z' and the existence of a finite infimum φ are sufficient to guarantee the existence of a Kuhn-Tucker (Lagrange) multiplier vector λ^*.

To strengthen the preceding theorem with the aid of the dual of Theorem 2, first note that the image $x' = (z', z', \ldots, z')$ of a Slater solution z' under the given linear transformation satisfies hypothesis (i) of the dual of Theorem 2. Then, note that the existence of a finite infimum φ is simply hypothesis (ii) of the dual of Theorem 2. Now, the convexity of C_o implies the existence of a vector $z^+ \in (\text{ri } C_o)$, by virtue of Theorem 6.2 on page 45 of [3]. Moreover, its image $x^+ = (z^+, z^+, \ldots, z^+)$ under the given linear transformation clearly satisfies hypothesis (iii) of the dual of Theorem 2 -- because (ri X) = X by virtue of fact (A), and because $J = \phi$. Consequently, the dual of Theorem 2 implies that both T and T* are nonempty and that $0 = \varphi + \psi$. In view of Corollary 7A of [6], we conclude from the nonemptyness of T* that a Kuhn-Tucker (Lagrange) vector λ^* exists. Finally, note that we have also shown the existence of another vector y*; so the Slater version of the Kuhn-Tucker theorem has actually been strengthened.

More significant implications of Theorem 2 are given on page 47 of [1].

References

1. Peterson, E.L., "Geometric Programming", SIAM Review, 18(1976),1.

2. _____________, "Symmetric Duality for Generalized Unconstrained Geometric Programming", SIAM Jour. Appl. Math., 19(1970), 487.

3. Rockafellar, R.T., Convex Analysis, Princeton University Press, Princeton, N.J. (1970).

4. Peterson, E.L., "Constrained Duality via Unconstrained Duality in Generalized Geometric Programming", to appear.

5. _____________, "Geometric Duality vis-a-vis Ordinary Duality", in preparation.

6. _____________, "Saddle Points and Duality in Generalized Geometric Programming", to appear.