Discussion Paper No. 152

Fenchel's Hypothesis and the Existence of Recession Directions in Convex Programming

by

Elmor L. Peterson

June 2, 1975

Fenchel's Hypothesis and the Existence of

Recession Directions in Convex Programming*

by

Elmor L. Peterson**

Abstract. Given a consistent (closed) convex programming problem and one of its corresponding dual problems, it is known that Fenchel's hypothesis for the dual problem guarantees both the absence of a duality gap and the existence of an optimal solution to the primal problem. In this paper we show that a negation of Fenchel's hypothesis for the dual problem guarantees the existence of a recession direction for the primal problem. We also show that, except in certain pathological situations, the existence of a recession direction for the primal problem guarantees a negation of Fenchel's hypothesis for the dual problem. We then show that, except in the same pathological situations, Fenchel's hypothesis for the dual problem is actually equivalent to the existence of a nonempty bounded optimal solution set for the primal problem.

Each of these results takes on a different form with different duality formulations, and each provides useful information about various special programming types -- such as linear programming, quadratic programming, posynomial programming, ℓ_p-regression analysis, optimal location, discrete optimal control with linear dynamics, monotone network analysis, chemical equilibrium analysis, and ordinary programming. Although some of this information is not new, none of it seems to be well-known.

---

* The author is indebted to Dr. C.T. Wu of Northern Illinois University, DeKalb, for reminding him about the relationship between the constancy space for a closed convex function and the domain of its conjugate transform -- a critical ingredient for the enclosed Theorem 4.

** 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-73-2516.

TABLE OF CONTENTS

1. <u>Introduction</u>. In convex programming there are at least five different formulations of duality -- the original Fenchel formulation [6,16], the (generalized) geometric programming formulation [3,8,11], the Fenchel-Rockafellar formulation [14,16], the ordinary Lagrangian formulation [17,5,16], and the Rockafellar formulation [15,16]. Although each formulation has its own advantages and disadvantages, each can also be viewed as a special case of each of the other four (by virtue of the specializations given in [8,16]). In particular, a given theorem in one formulation has its counterparts in each of the other four formulations, all of which can be used to supply important information about various special programming types.

For simplicity, we begin by establishing our main results in the context of the most recent unconstrained geometric programming formulation [11]. This also provides a convenient mechanism for translating them into the context of the most recent constrained geometric programming formulation [11], as well as both the Fenchel-Rockafellar formulation [16](including the original Fenchel formulation) and the more recent Rockafellar formulation [16]. In turn, the latter provides a convenient mechanism for translating them into the context of the ordinary Lagrangian formulation [16]. When appropriate, we also consider applications to special programming types.

Familiarity with the usual facts, terminology, and notation from "convex analysis" [16] is assumed. For example, (ri S) denotes the "relative interior" of a set $S \subseteq E_N$ (N-dimensional Euclidean space). Moreover, all cones and functions introduced are assumed to be both "convex" and "closed".

2. <u>The Main Results</u>. Given a closed convex cone $\mathcal{X} \subseteq E_n$ and given a closed convex function $\mathcal{g}$ with (effective) domain $\mathcal{C} \subseteq E_n$, consider the resulting "geometric programming problem" $\mathcal{A}$.

PROBLEM $\mathcal{A}$. Using the feasible solution set

$$\mathcal{S} \triangleq \mathcal{X} \cap \mathcal{C},$$

calculate both the problem infimum

$$\varphi \triangleq \inf_{x \in \mathcal{S}} g(x)$$

and the optimal solution set

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

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

$$\mathcal{Y} \triangleq \{y \in E_n \mid 0 \leq \langle x, y \rangle \text{ for each } x \in \mathcal{X}\}$$

and the "conjugate transform function" h, whose domain

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

and whose functional values

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

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

PROBLEM $\mathcal{B}$. Using the feasible solution set

$$\mathcal{T} \triangleq \mathcal{Y} \cap \mathcal{D},$$

calculate both the problem infimum

$$\psi \triangleq \inf_{y \in \mathcal{T}} h(y)$$

and the optimal solution set

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

The following Rockafellar version of "Fenchel's theorem" [16] is one of the most important theorems in geometric programming [11].

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 $\mathcal{S}^*$.

The existence of a feasible solution $y^\circ \in (\mathrm{ri}\,\mathcal{Y}) \cap (\mathrm{ri}\,\mathcal{D})$ is termed Fenchel's hypothesis for problem $\mathcal{B}$. It is to be intimately related to the existence of a recession direction for problem $\mathcal{A}$, namely, a nonzero (direction) vector $\delta \in \mathcal{X}$ such that for each $x \in \mathcal{C}$ the function $g(x + \cdot\,\delta)$ is defined and monotone nonincreasing on the set of all nonnegative real numbers.

The following theorem brings to light an important duality between the negation of Fenchel's hypothesis for problem $\mathcal{B}$ and the existence of recession directions for problem $\mathcal{A}$. It is the key to the main theorems of this paper.

Theorem 2. If problem $\mathcal{B}$ has no feasible solution $y^\circ \in (\mathrm{ri}\,\mathcal{Y}) \cap (\mathrm{ri}\,\mathcal{D})$, then there exists a nonzero (direction) vector $\delta \in \mathcal{X}$ such that for each $x \in \mathcal{C}$ the function $g(x + \cdot\,\delta)$ is defined and monotone nonincreasing on the set of all nonnegative real numbers (but not conversely). On the other hand, if there exists a nonzero (direction) vector $\delta \in \mathcal{X}$ such that for some $x \in \mathcal{C}$ the function $g(x + \cdot\,\delta)$ is defined and achieves its maximum value on the set of all nonnegative real numbers at zero, and if there is no $(n-1)$-dimensional vector space that contains both $\mathcal{Y}$ and $\mathcal{D}$ (e.g. if either the vector space generated by $\mathcal{Y}$ or the vector space generated by $\mathcal{D}$ is E_n), then problem $\mathcal{B}$ has no feasible solution $y^\circ \in (\mathrm{ri}\,\mathcal{Y}) \cap (\mathrm{ri}\,\mathcal{D})$.

Proof. Since both $\mathcal{Y}$ and $\mathcal{D}$ are convex (by virtue of conjugate transform theory [16]), Theorem 6.2 on page 45 of [16] implies that both $(\mathrm{ri}\,\mathcal{Y})$ and $(\mathrm{ri}\,\mathcal{D})$ are convex. Consequently, our assumption that $(\mathrm{ri}\,\mathcal{Y}) \cap (\mathrm{ri}\,\mathcal{D})$ is empty and (the separation) Theorem 11.3 on page 97 of [16] imply that $\mathcal{Y}$ and $\mathcal{D}$ can be "properly separated" by a hyperplane in E_n. Moreover, since $\mathcal{Y}$ is a cone, Theorem 11.7 on page 100 of [16] shows that there is at least one such hyperplane that passes through the origin of E_n. In particular then, there exists at least one nonzero vector $\delta \in E_n$ such that

$$0 \le \langle \delta, y \rangle \text{ for each } y \in \mathcal{Y} \tag{1}$$

and

$$\langle \delta, y \rangle \le 0 \text{ for each } y \in \mathcal{D}. \tag{2}$$

Now, statement (1) and the well-known symmetry of the duality between $\mathcal{X}$ and $\mathcal{Y}$ imply that $\delta \in \mathcal{X}$. Moreover, elementary considerations show that the validity of statement (2) is not altered when $\mathcal{D}$ is replaced by the closed conical hull $(\mathrm{cl\ cone}\,\mathcal{D})$ of $\mathcal{D}$. Furthermore, the resulting statement (2) and Theorem 14.2 on page 122 of [16] imply that δ is in the "recession cone" for $g{:}\mathcal{C}$, which means that for each $x \in \mathcal{C}$ the function $g(x + \cdot\, \delta)$ is defined and monotone nonincreasing on the set of all nonnegative real numbers. Finally, a counterexample to the converse of the assertion just proved can easily be obtained by letting $\mathcal{X}$ be any nontrivial vector space in E_2, while letting g be any constant function with domain $\mathcal{C} \stackrel{\Delta}{=} E_2$.

Now, according to Theorem 8.7 on page 70 of [16], our assumption that there exists a nonzero vector $\delta \in \mathcal{X}$ such that for some $x \in \mathcal{C}$ the function $g(x + \cdot\ \delta)$ is defined and achieves its maximum value on the set of all nonnegative real numbers at zero means that there exists a nonzero vector δ in both $\mathcal{X}$ and the recession cone for $g{:}\mathcal{C}$. Given such a vector δ, we know from the definition of $\mathcal{Y}$ and Theorem 14.2 on page 122 of [16] that

$$0 \le \langle \delta, y \rangle \text{ for each } y \in \mathcal{Y} \tag{1}$$

and

$$\langle \delta, y \rangle \le 0 \text{ for each } y \in (\text{cl cone}\,\mathcal{D}). \tag{2'}$$

Since $\mathcal{D} \subseteq (\text{cl cone}\,\mathcal{D})$, the validity of statement (2') is not altered when $(\text{cl cone}\,\mathcal{D})$ is replaced by $\mathcal{D}$. Now, statement (1) and the resulting statement (2) along with our assumption that there is no $(n-1)$-dimensional vector space (i.e. no hyperplane through the origin) that contains both $\mathcal{Y}$ and $\mathcal{D}$ clearly imply via Theorem 11.1 on page 95 of [16] that $\mathcal{Y}$ and $\mathcal{D}$ can be properly separated by a hyperplane in E_n. It is then a consequence of Theorem 11.3 on page 97 of [16] that $(\text{ri}\,\mathcal{Y}) \cap (\text{ri}\,\mathcal{D})$ is empty. q.e.d.

It is worth noting that the existence of a recession direction δ for problem $\mathcal{A}$ does not require that problem $\mathcal{A}$ have a nonempty feasible solution set $\mathcal{S}$. However, if $\mathcal{S}$ happens to be nonempty, the existence of δ does obviously imply that the optimal solution set $\mathcal{S}^*$ for problem $\mathcal{A}$ is either empty or a union of half-lines. In either case, the dimensionality of problem $\mathcal{A}$ can be reduced via the procedures described in [1]. In fact, such a "reduction" can frequently be carried to such an extent that Fenchel's hypothesis is satisfied for the geometric dual $\mathcal{B}'$ of the resulting "canonical form" $\mathcal{A}'$ for problem $\mathcal{A}$. For example, see [4,12,7,13,2].

The following theorem brings to light an important duality between Fenchel's hypothesis for problem $\mathcal{B}$ and a certain boundedness property for problem $\mathcal{A}$. It also points out certain important consequences of the boundedness property.

Theorem 3. Suppose that problem $\mathcal{A}$ has a nonempty feasible solution set $\mathcal{S}$. If there is a real number r greater than or equal to the infimum φ for problem $\mathcal{A}$ such that the corresponding objective function "level set"

$$\mathcal{L}_r \overset{\Delta}{=} \{x \in \mathcal{S} \mid g(x) \leq r\}$$

is nonempty and bounded, then

(i) for each $s > \varphi$, the level set $\mathcal{L}_s$ is nonempty, convex and compact,

(ii) the optimal solution set $\mathcal{S}^*$ for problem $\mathcal{A}$ is nonempty, convex and compact,

(iii) problem $\mathcal{B}$ has a feasible solution $y^\circ \in (\mathrm{ri}\,\mathcal{Y}) \cap (\mathrm{ri}\,\mathcal{D})$.

On the other hand, if problem $\mathcal{B}$ has a feasible solution $y^\circ \in (\mathrm{ri}\,\mathcal{Y}) \cap (\mathrm{ri}\,\mathcal{D})$, and if there is no $(n\text{-}1)$-dimensional vector space that contains both $\mathcal{Y}$ and $\mathcal{D}$ (e.g. if either the vector space generated by $\mathcal{Y}$ or the vector space generated by $\mathcal{D}$ is E_n), then for each real number r greater than or equal to the infimum φ for problem $\mathcal{A}$ the corresponding objective function level set $\mathcal{L}_r$ is bounded.

Proof. It is easy to see that the objective function $g:\mathcal{S}$ inherits the convexity and closedness of $\mathcal{X}$ and $g:\mathcal{C}$; and it is then a consequence of Theorem 4.6 on page 28 of [16] and Theorem 7.1 on page 51 of [16] that $\mathcal{L}_s$ is convex and closed for each $s \geq \varphi$.

Now, our assumption that $\mathcal{L}_r$ is nonempty and bounded implies via Theorem

8.4 on page 64 of [16] that the recession cone for $\mathscr{L}_r$ consists solely of the zero vector. It is then a consequence of Theorem 8.7 on page 70 of [16] that the recession cone for $\mathscr{L}_s$ consists solely of the zero vector for each s for which $\mathscr{L}_s$ is nonempty. We now infer from Theorem 8.4 on page 64 of [16] that $\mathscr{L}_s$ is bounded for each $s \geq \varphi$. In summary then, $\mathscr{L}_s$ is convex and compact for each $s \geq \varphi$. Since $\mathscr{L}_s$ is obviously nonempty for each $s > \varphi$, the proof of conclusion (i) is now complete. Moreover, the convexity and compactness of $\mathscr{S}^*$ have also been established, because $\mathscr{S}^*$ is clearly identical to $\mathscr{L}_\varphi$. To show that $\mathscr{S}^*$ is nonempty and hence complete the proof of conclusion (ii), first note from the assumption that $r \geq \varphi$ and the definition of $\mathscr{L}_r$ that the infimum of $g:\mathscr{L}_r$ is just φ, and that the corresonding optimal solution set $\mathscr{L}_r^* = \mathscr{S}^*$. Since it is easy to see that $g:\mathscr{L}_r$ inherits the (previously established) closedness of $g:\mathscr{S}$, and since closedness of $g:\mathscr{L}_r$ implies via Theorem 7.1 on page 51 of [16] that $g:\mathscr{L}_r$ is lower semicontinuous, the (previously established) compactness of $\mathscr{L}_r$ and a standard argument from analysis can now be used to show the existence of an $x^* \in \mathscr{L}_r^* = \mathscr{S}^*$. Finally, conclusion (iii) is now a straightforward consequence of the first assertion of Theorem 2.

To prove the final assertion of Theorem 3, suppose to the contrary that there is an $r \geq \varphi$ such that $\mathscr{L}_r$ is not bounded. Then, Theorem 8.4 on page 64 of [16] shows that the recession cone for $\mathscr{L}_r$ contains a nonzero vector δ. Now, according to Theorem 8.7 on page 70 of [16], δ must also be in the recession cone for $g:\mathscr{S}$. Consequently, Theorem 8.6 on page 68 of [16] implies that for each $x \in \mathscr{S}$ the function $g(x + \cdot\, \delta)$ is defined and achieves its maximum value on the set of all nonnegative real numbers at zero. Moreover, the definition of $\mathscr{S}$ shows that δ is also in the recession cone for $\mathcal{X}$; and it is then a consequence of the conicality of $\mathcal{X}$ and Theorem 8.1 on page 61 of [16] that $\delta \in \mathcal{X}$. The final assertion of Theorem 2 now contradicts our assumption that problem

$\mathcal{B}$ has a feasible solution $y^\circ \in (\mathrm{ri}\,\mathcal{Y}) \cap (\mathrm{ri}\,\mathcal{D})$; so $\mathcal{L}_r$ is bounded for each $r \geq \varphi$. q.e.d.

It should be emphasized that the assumption in the first assertion of Theorem 3 is satisfied when the optimal solution set $\mathcal{S}^*$ for problem $\mathcal{A}$ is nonempty and bounded (because $\mathcal{L}_\varphi = \mathcal{S}^*$). Moreover, it is worth noting that the second assumption in the final assertion of Theorem 3 is needed for the validity of that assertion. In fact, a counterexample to that assertion without its second assumption can be obtained by letting $\mathcal{X}$ be any nontrivial vector space in E_2, while letting g be any constant function with domain $\mathcal{C} \stackrel{\Delta}{=} E_2$.

For the preceding counterexample, it is important to note that problem $\mathcal{A}$ has a nonempty feasible solution set $\mathcal{S}$ and that its objective function $g:\mathcal{S}$ has a "constancy space" [16] with dimension one, namely, $\mathcal{X}$. That is, given an arbitrary $x \in \mathcal{S}$, the vector $x + \gamma \in \mathcal{S}$ and $g(x+\gamma) \equiv g(x)$ for each $\gamma \in \mathcal{X}$. The following theorem shows that the existence of this constancy space with positive dimension is not accidental.

<u>Theorem 4.</u> <u>Suppose that problem $\mathcal{A}$ has a nonempty feasible solution set $\mathcal{S}$. Then, the orthogonal complement $\mathcal{E}^\perp$ of the smallest vector space $\mathcal{E}$ containing both $\mathcal{Y}$ and $\mathcal{D}$ is contained in the constancy space for $g:\mathcal{S}$; that is, given an arbitrary $x \in \mathcal{S}$, the vector $x+\gamma \in \mathcal{S}$ and $g(x+\gamma) \equiv g(x)$ for each $\gamma \in \mathcal{E}^\perp$. In particular, if there is an $(n$-1)-dimensional vector space that contains both $\mathcal{Y}$ and $\mathcal{D}$, then the objective function $g:\mathcal{S}$ has a constancy space with dimension at least one.</u>

Proof. Since $\mathcal{Y} \subseteq \mathcal{E}$, it is clear from the duality of $\mathcal{X}$ and $\mathcal{Y}$ that

$$\gamma \in \mathcal{X} \text{ when } \gamma \in \mathcal{E}^\perp.$$

Moreover, since $\mathcal{D} \subseteq \mathcal{E}$, it is clear that $\langle \gamma, y \rangle \equiv 0$ when $\gamma \in \mathcal{E}^{\perp}$ and $y \in \mathcal{D}$. Hence, $\sup_{y \in \mathcal{D}} [\langle x+\gamma, y\rangle - h(y)] = \sup_{y \in \mathcal{D}} [\langle x, y\rangle - h(y)]$ when $\gamma \in \mathcal{E}^{\perp}$; from which it follows via the conjugacy of $g:\mathcal{C}$ and $h:\mathcal{D}$ that

$$x + \gamma \in \mathcal{C} \text{ and } g(x+\gamma) \equiv g(x) \text{ when } x \in \mathcal{C} \text{ and } \gamma \in \mathcal{E}^{\perp}.$$

Now, the convex conicality of $\mathcal{X}$ and both of the preceding displayed relations obviously imply the first assertion of Theorem 4. Moreover, the second assertion of Theorem 4 is then an immediate consequence of the well-known dimensional complementarity of $\mathcal{E}$ and $\mathcal{E}^{\perp}$. q.e.d.

A consistent problem $\mathcal{A}$ whose objective function $g:\mathcal{S}$ has a constancy space with dimension at least one is clearly pathological -- in that the dimension n of the space E_n in which the problem has been formulated is obviously larger than it needs to be. In fact, it is not difficult to see that E_n can be replaced by $\mathcal{E}$ in such a way that such problems are reduced to equivalent consistent problems whose objective functions have constancy spaces with dimension zero.

The following theorem summarizes the main results of this paper in the context of nonpathological consistent problems $\mathcal{A}$.

<u>Theorem</u> 5. <u>Suppose that problem</u> $\mathcal{A}$ <u>has a nonempty feasible solution set</u> $\mathcal{S}$ <u>and that the constancy space for its objective function</u> $g:\mathcal{S}$ <u>has dimension zero</u>. <u>Then</u>, <u>the following four conditions are equivalent</u>:

(I) <u>there exists no nonzero (direction) vector</u> $\delta \in \mathcal{X}$ <u>such that for some</u> $x \in \mathcal{C}$ <u>the function</u> $g(x + \cdot\, \delta)$ <u>is defined and monotone nonincreasing on the set of all nonnegative real numbers</u>,

(II) <u>the objective function</u> $g:\mathcal{S}$ <u>has a level set</u> $\mathcal{L}_r$ <u>that is nonempty and bounded</u>,

(III) the optimal solution set $\mathscr{S}^*$ is nonempty and bounded,

(IV) problem $\mathcal{B}$ has a feasible solution $y^\circ \in (\mathrm{ri}\,\mathcal{Y}) \cap (\mathrm{ri}\,\mathcal{D})$.

Moreover, if any of these four conditions are satisfied, then

(i) for each $s \geq \varphi$, the level set $\mathcal{L}_s$ is nonempty, convex and compact,

(ii) the optimal solution set $\mathscr{S}^*$ for problem $\mathcal{A}$ is nonempty, convex and compact.

Proof. Simply note that our second hypothesis implies via Theorem 4 that there is no $(n-1)$-dimensional vector space that contains both $\mathcal{Y}$ and $\mathcal{D}$; and then use Theorems 2 and 3 repetitively. q.e.d.

In view of the crucial nature of the second hypothesis of Theorem 5, it is helpful to have the following information about the constancy space for the objective function $g{:}\mathscr{S}$.

Theorem 6. Given that problem $\mathcal{A}$ has a nonempty feasible solution set $\mathscr{S}$, the constancy space for its objective function $g{:}\mathscr{S}$ is the intersect of the largest subspace contained in the cone $\mathcal{X}$ with the constancy space for $g{:}\mathcal{C}$.

Proof. First, note from the definition of $g{:}\mathscr{S}$ and Theorem 8.3 on page 63 of [16] along with Corollary 8.6.1 on page 69 of [16] that the constancy space for $g{:}\mathscr{S}$ is simply the intersect of the recession cone for $\mathcal{X}$ with minus the recession cone for $\mathcal{X}$ and the constancy space for $g{:}\mathcal{C}$. Now, use the conicality of $\mathcal{X}$ and Theorem 8.1 on page 61 of [16] to show that the recession cone for $\mathcal{X}$ is just $\mathcal{X}$ itself; and then observe from Theorem 2.7 on page 15 of [16] that the intersect of $\mathcal{X}$ and minus $\mathcal{X}$ is simply the largest subspace contained in $\mathcal{X}$. q.e.d.

The following corollary is frequently all that is needed to show that the second hypothesis of Theorem 5 is satisfied.

<u>Corollary</u> 6A. <u>Given that problem $\mathcal{A}$ has a nonempty feasible solution set $\mathcal{S}$, if either the largest subspace contained in $\mathcal{X}$ contains only the zero vector</u> (i.e. <u>the cone $\mathcal{X}$ is "pointed") or the constancy space for $g:\mathcal{C}$ contains only the zero vector, then the constancy space for $g:\mathcal{S}$ has dimension zero.</u>

It is worth noting that the symmetry of conical duality (implied by Theorem 14.1 on page 121 of [16]) together with the symmetry of functional conjugacy (asserted in Corollary 12.2.1 on page 104 of [16]) induces a symmetry on the theory that relates problem $\mathcal{A}$ to problem $\mathcal{B}$. In particular, each of the preceding theorems about $\mathcal{A}$ and $\mathcal{B}$ automatically produces an equally valid "dual theorem" about $\mathcal{B}$ and $\mathcal{A}$ (obtained by interchanging the symbols $\mathcal{X}$ and $\mathcal{Y}$, the symbols $\mathcal{C}$ and $\mathcal{D}$, the symbols g and h, the symbols $\mathcal{S}$ and $\mathcal{T}$, the symbols φ and ψ, the symbols $\mathcal{S}^*$ and $\mathcal{T}^*$, and the symbols x and y). To be concise, the statement of each dual theorem is left to the reader.

The most effective applications of the preceding theorems and their duals also employ many of the basic facts about relative interiors (given in section 6 of [16]). In particular, the following basic facts are usually crucial:

(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 derived on page 44 of [16]; fact (B) can be obtained inductively from the formula at the top of page 49 of [16]; and fact (C) is explained on page 44 of [16].

The preceding theory can be applied directly to posynomial programming, ℓ_p-regression analysis, optimal location, discrete optimal control with linear dynamics, and monotone network analysis -- simply by consulting examples 1 through 5 respectively in section 2 of [11]. Needless to say, such applications are actually left to the interested reader.

3. Other Formulations. For each of the other formulations, we simply specify problem $\mathcal{A}$ and then give the resulting dual problem $\mathcal{B}$ along with the corresponding Fenchel hypothesis for problem $\mathcal{B}$. The reader can then easily construct the corresponding counterpart of each of the main results given in section 2.

The following three subsections are pedagogically independent of one another, but the fourth subsection utilizes the third subsection.

3.1. The Constrained Geometric Programming Formulation. First, suppose that:

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 $(x^0, x^I, x^J) \triangleq x$ and $(y^0, y^I, y^J) \triangleq y$ are independent vector variables in E_n, where

$$n \triangleq n_0 + \sum_I n_i + \sum_J n_j,$$

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

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

Now, let

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

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

$$\mathcal{C} \stackrel{\Delta}{=} \{(x^0, x^I, \alpha, x^J, \varkappa) \in E_{\eta} \mid x^0 \in C_0; x^i \in C_i,\ \alpha_i \in E_1, \text{ and }$$
$$g_i(x^i) + \alpha_i \le 0,\ i \in I;\ (x^j, \varkappa_j) \in C_j^+,\ j \in J\},$$

and let

$$\mathcal{g}(x^0, x^I, \alpha, x^J, \varkappa) \stackrel{\Delta}{=} g_0(x^0) + \sum_J g_j^+(x^j, \varkappa_j) \stackrel{\Delta}{=} G(x, \varkappa),$$

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

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

and functional values

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

Then, section 6 of [10] shows that

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

Section 6 of [10] also shows that

$$\mathcal{D} = \{(y^0, y^I, \lambda, y^J, \beta) \in E_n \mid y^0 \in D_0;\ (y^i, \lambda_i) \in D_i^+,\ i \in I;\ y^j \in D_j,$$
$$\beta_j \in E_1,\ \text{and}\ h_j(y^j) + \beta_j \le 0,\ j \in J\},$$

and that

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

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

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

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

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

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

$$(\mathrm{ri}\,\mathcal{D}) = \{(y^0, y^I, \lambda, y^J, \beta) \in E_n \mid y^0 \in (\mathrm{ri}\,D_0);\ \lambda_i > 0 \text{ and } y^i \in \lambda_i(\mathrm{ri}\,D_i),$$
$$i \in I;\ y^j \in (\mathrm{ri}\,D_j),\ \beta_j \in E_1,\ \text{and}\ h_j(y^j) + \beta_j < 0,\ j \in J\},$$

by virtue of both the equation

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

and the equation

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

To derive the latter equation, simply use Theorem 6.8 on page 49 of [16] along with fact (C). To derive the former equation, first consider the point-to-set mapping $Y_i^+ : \Lambda_i^+$ where

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

and

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

Now, Corollary 6.8.1 on page 50 of [16] implies that

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

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

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

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

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

by virtue of Corollary 6.6.1 on page 48 of [16]. Consequently, our derivation of the preceding formula for $(\mathrm{ri}\,\mathcal{D})$ is complete.

In particular then, the corresponding Fenchel hypothesis for problem $\mathcal{B}$ simply asserts that

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

Moreover, it is probably worth noting that this hypothesis is in fact equivalent to the hypothesis

$$\begin{aligned}
&\text{there is a vector } (\bar{y}^0, \bar{y}^I, \bar{\lambda}, \bar{y}^J) \\
&\text{such that } (\bar{y}^0, \bar{y}^I, \bar{y}^J) \in Y;\ \bar{y}^0 \in D_0;\ (\bar{y}^i, \bar{\lambda}_i) \in D_i^+, \\
&i \in I;\ \bar{y}^j \in D_j \text{ and } h_j(\bar{y}^j) < 0,\ j \in J \text{ --- and} \\
&\text{there is a vector } (\tilde{y}^0, \tilde{y}^I, \tilde{\lambda}, \tilde{y}^J) \\
&\text{such that } (\tilde{y}^0, \tilde{y}^I, \tilde{y}^J) \in (\mathrm{ri}\, Y);\ \tilde{y}^0 \in (\mathrm{ri}\, D_0); \\
&\tilde{\lambda}_i > 0 \text{ and } \tilde{y}^i \in \tilde{\lambda}_i (\mathrm{ri}\, D_i),\ i \in I;\ \tilde{y}^j \in (\mathrm{ri}\, D_j),\ j \in J.
\end{aligned}$$

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 [16] and Theorem 7.1 on page 51 of [16] imply that a convex combination $\alpha(\bar{y}^0, \bar{y}^I, \bar{\lambda}, \bar{y}^J) + \beta(\tilde{y}^0, \tilde{y}^I, \tilde{\lambda}, \tilde{y}^J)$ of vectors $(\bar{y}^0, \bar{y}^I, \bar{\lambda}, \bar{y}^J)$ and $(\tilde{y}^0, \tilde{y}^I, \tilde{\lambda}, \tilde{y}^J)$ that satisfy the latter hypothesis will satisfy the former hypothesis for sufficiently small $\beta > 0$. Needless to say, the latter Fenchel hypothesis is generally the easiest to verify. Although the condition $\bar{y}^j \in D_j$ and $h_j(\bar{y}^j) < 0$, $j \in J$

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.

The theory established in section 2 can now be applied to constrained posynomial programming, linear programming, and ordinary programming -- simply by consulting examples 6 through 8 respectively in section 2 of [11]. Moreover, the interested reader should also have no trouble making applications to chemical equilibrium analysis, quadratic programming, constrained ℓ_p-regression analysis, and constrained optimal location -- simply by consulting the references alluded to in section 2 of [11].

3.2. The Fenchel-Rockafellar Formulation. Let

$$\mathcal{X} \overset{\Delta}{=} \text{the column space of } \begin{bmatrix} I_r \\ M \end{bmatrix},$$

where I_r is the $r \times r$ identity matrix, M is an arbitrary $s \times r$ matrix, and $r + s = n$. In addition, let

$$\mathcal{C} \overset{\Delta}{=} C^1 \times C^2 \text{ and } g(x) \overset{\Delta}{=} g^1(x^1) - g^2(x^2),$$

where the closed convex function g^1 has domain $C^1 \subseteq E_r$, and the closed concave function g^2 has domain $C^2 \subseteq E_s$.

Then, section 5 of [8] shows that

$$\mathcal{Y} = \text{the column space of } \begin{bmatrix} M^t \\ -I_s \end{bmatrix},$$

where M^t is the transpose of M, and I_s is the $s \times s$ identity matrix. Section 5 of [8] also shows that

$$\mathcal{D} = D^1 \times \left[-D^2\right] \text{ and } h(y) = h^1(y^1) - h^2(-y^2),$$

where $h^1{:}D^1$ is the (convex) conjugate transform of $g^1{:}C^1$, and $h^2{:}D^2$ is the <u>concave</u> conjugate transform of $g^2:C^2$ obtained by replacing sup with inf in the definition of the conjugate transform. (Actually, section 5 of [8] also shows how to use simple linear transformations to eliminate the troublesome minus signs and hence achieve the true symmetry of the original Fenchel-Rockafellar formulation, but those details need not be of any direct concern here.)

Now, the preceding formulas for $\mathcal{Y}$ and $\mathcal{D}$ along with fact (A) and fact (B) imply that

$$(\text{ri}\,\mathcal{Y}) = \mathcal{Y}$$

and

$$(\text{ri}\,\mathcal{D}) = (\text{ri}\,D^1) \times [-(\text{ri}\,D^2)],$$

by virtue of Corollary 6.6.1 on page 48 of [16].

In particular then, the corresponding Fenchel hypothesis for problem $\mathcal{B}$ simply asserts that

there is a vector $z \in (\text{ri}\,D^2)$ such
that $M^t z \in (\text{ri}\,D^1)$.

Finally, it should be mentioned that the original Fenchel formulation can be obtained by letting $M \triangleq I_r$, in which case $s = r$.

3.3. <u>The Rockafellar Formulation</u>. Let

$$\mathcal{X} \triangleq \text{the column space of } \begin{bmatrix} I_p \\ 0_q \end{bmatrix} ,$$

where I_p is the $p \times p$ identity matrix, 0_q is the $q \times p$ zero matrix, and $p + q = n$.

Then, section 5 of [8] shows that

$$\mathscr{Y} = \text{the column space of } \begin{bmatrix} 0_p \\ I_q \end{bmatrix},$$

where I_q is the $q \times q$ identity matrix and 0_p is the $p \times q$ zero matrix.

Now, the preceding formula for $\mathscr{Y}$ along with fact (A) implies that

$$(\mathrm{ri}\,\mathscr{Y}) = \mathscr{Y}.$$

In particular then, the corresponding Fenchel hypothesis for problem $\mathscr{B}$ simply asserts that

$$\text{there is a vector } \lambda \in E_q \text{ such that } (0,\lambda) \in (\mathrm{ri}\,\mathscr{D}).$$

Moreover, it is probably worth noting that Theorem 6.8 on page 49 of [16] and Theorem 6.2 on page 45 of [16] imply that this hypothesis is in fact equivalent to the hypothesis

$$0 \in (\mathrm{ri}\, M),$$

where

$$M \stackrel{\Delta}{=} \{\mu \mid (\mu,\lambda) \in \mathscr{D} \text{ for some } \lambda\}.$$

Finally, it should be mentioned that the original Rockafellar formulation can be obtained by simple linear transformations of the resulting problem $\mathscr{B}$ into an equivalent maximization problem.

3.4. <u>The Ordinary Lagrangian Formulation</u>. In the context of the preceding Rockafellar formulation, suppose that

$$x \stackrel{\Delta}{=} (z,u) \text{ where } z \in E_p \text{ and } u \in E_q.$$

Then, let

$$\mathscr{C} \stackrel{\Delta}{=} \{(z,u) \mid z \in C \text{ and } G_i(z) + u_i \le 0,\ i \in I\} \text{ and } g(z,u) \stackrel{\Delta}{=} G_0(z),$$

where I is a (possibly empty) positive - integer index set with finite cardinality o(I), and the G_k, $k \in \{0\} \cup I$, are closed convex functions with a common domain $C \subseteq E_p$. Problem A now consists of minimizing $G_0(z)$ over C subject to the constraints $G_i(z) \le 0$, $i \in I$.

Now, suppose that

$$y \overset{\Delta}{=} (\mu, \lambda) \text{ where } \mu \in E_p \text{ and } \lambda \in E_q.$$

Then, an elementary computation (given essentially in section 30 of [16]) shows that

$$\mathcal{D} = \{(\mu, \lambda) \mid \lambda \ge 0 \text{ and } \sup_{z \in C} [\langle \mu, z \rangle - G_0(z) - \sum_I \lambda_i G_i(z)] < +\infty\}$$

and

$$h(\mu, \lambda) = \sup_{z \in C} [\langle \mu, z \rangle - G_0(z) - \sum_I \lambda_i G_i(z)].$$

Problem $\mathcal{B}$ now consists of computing the minimax of the negative of the ordinary Lagrangian $G_0(z) + \sum_I \lambda_i G_i(z)$.

Now, the preceding formula for $\mathcal{D}$ along with the next-to-last paragraph of subsection 3.3 implies that the corresponding Fenchel hypothesis for problem $\mathcal{B}$ simply asserts that

$$0 \in (\mathrm{ri}\, M),$$

where

$$M \overset{\Delta}{=} \{\mu \mid \sup_{z \in C} [\langle \mu, z \rangle - G_0(z) - \sum_I \lambda_i G_i(z)] < +\infty \text{ for some } \lambda \ge 0\}.$$

However, this hypothesis is generally difficult to check because of the complicated nature of M. In fact, it should be contrasted with the relatively simple alternative hypothesis that results from utilizing the constrained geometric dual $\mathcal{B}$ (as indicated toward the end of subsection 3.1). The alternative hypothesis is simply that there is a vector (y^0, y^I, λ) such

that $y^0 + \sum_I y^i = 0$; $y^0 \in (\mathrm{ri}\, D_0)$; $\lambda_i > 0$ and

$y^i \in \lambda_i (\mathrm{ri}\, D_i)$, $i \in I$,

where D_k is the domain of the conjugate transform of $G_k : C$, $k \in \{0\} \cup I$.

Since the preceding functions g and h are different in nature, a formula for $(\mathrm{ri}\, \mathcal{C})$ should also be developed. To do so, note that the preceding formula for $\mathcal{C}$ along with fact (B), Theorem 6.8 on page 49 of [16], and fact (C) implies that

$$(\mathrm{ri}\, \mathcal{C}) = \{(z,u) \mid z \in (\mathrm{ri}\, C) \text{ and } G_i(z) + u_i < 0,\ i \in I\}.$$

In particular then, the corresponding Fenchel hypothesis for problem $\mathcal{Q}$ simply asserts that

there is a vector $z \in (\mathrm{ri}\, C)$ such that $G_i(z) < 0$, $i \in I$,

a condition that is slightly stronger than the well-known "Slater constraint qualification".

References

1. Abrams, R.A., "Projections of Convex Programs with Unattained Infima", SIAM Jour. Control, 13(1975), in press.

2. ___________, "Consistency, Superconsistency, and Dual Degeneracy in Posynomial Geometric Programming", Operations Research, to appear.

3. Duffin, R.J., "Dual Programs and Minimum Cost", SIAM Jour. Appl. Math., 10(1962), 119.

4. Duffin, R.J., Peterson, E.L., and Zener, C., Geometric Programming-Theory and Applications, Wiley, New York (1967)--Russian translation by D. Babayev, Mir, Moscow (1972).

5. Falk, J.E., "Lagrange Multipliers and Nonlinear Programming", Jour. Math. Anal. Appls., 19(1967), 141.

6. Fenchel, W., "Convex Cones, Sets and Functions", Math. Dept. mimeographed lecture notes, Princeton Univ., Princeton, N.J. (1951).

7. Hall, M.A., and Peterson, E.L., "Traffic Equilibria analysed via Geometric Programming", Proc. Intern. Symp. Traffic Equilibrium Methods, Florian (ed.), to appear.

8. Peterson, E.L., "Symmetric Duality for Generalized Unconstrained Geometric Programming", SIAM Jour. Appl. Math., 19(1970), 487.

9. ______________, "Mathematical Foundations of Geometric Programming", Appendix to "Geometric Programming and Some of Its Extensions", Optimization and Design, Avriel, Rijckaert, and Wilde (eds.), Prentice-Hall, Englewood Cliffs, N.J. (1973), 244.

10. ______________, "Generalization and Symmetrization of Duality in Geometric Programming", to appear.

11. ______________, "Geometric Programming", SIAM Rev., to appear.

12. Peterson, E.L., and Ecker, J.G., "Geometric Programming: Duality in Quadratic Programming and ℓ_p-approximation I", Proc. Intern. Sym. Math. Prog., Kuhn (ed.), Princeton Univ. Press, Princeton, N.J. (1970), 445.

13. Peterson, E.L., and Wendell, R.E., "Optimal Location by Geometric Programming", to appear.

14. Rockafellar, R.T., "Duality and Stability in Extremum Problems Involving Convex Functions", Pacific Jour. Math., 21(1967), 167.

15. ________________, "Duality in Nonlinear Programming", Amer. Math. Soc. Lects. Appl. Math., 11, Dantzig and Veinott (eds.), Amer. Math. Soc., Providence, R.I. (1968), 401.

16. ________________, Convex Analysis, Princeton Univ. Press, Princeton, N.J. (1970).

17. Wolfe, P., "A Duality Theorem for Nonlinear Programming", Quart. Appl. Math., 19(1961), 239.