DISCUSSION PAPER NO. 270

GEOMETRIC DUALITY VIA ROCKAFELLAR DUALITY

by

Elmor L. Peterson

March 17, 1977

Geometric Duality via Rockafellar Duality

by

Elmor L. Peterson*

Abstract. A specialization of Rockafellar duality to (generalized) geometric duality provides an efficient mechanism for extending to the latter the theory previously developed for the former.

Key Words

Geometric programming

Rockafellar programming

Conjugate transformation

Duality

Parametric programming

Sensitivity analysis

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. Introduction. There are at least five different formulations of duality -- the original Fenchel formulation [3,8], the (generalized) geometric programming formulation [1,4,5], the Fenchel-Rockafellar formulation [6,8], the ordinary Lagrangian formulation [10,2,8], and the Rockafellar formulation [7,8,9]. Although each formulation has its own advantages and disadvantages, each can also be viewed as a special case of each of the other four.

The appropriate specializations have already been carried out in [4,8,9], but only for a very limited geometric programming formulation. Thus, this paper complements those references by specializing the Rockafellar formulation to the most general geometric programming formulation.

Section 2 presents a version of the Rockafellar formulation that facilitates the specialization given in Section 3. This specialization does not require convexity assumptions and uses only elementary real analysis.

2. Rockafellar duality. Suppose that $g:C$ is a (proper) function g with a nonempty (effective) domain $C \subseteq E_N$, and assume that the independent variable (d,p) in C is the Cartesian product of a "decision" (vector) variable d and a "perturbation" (vector) parameter p.

Consider the parameterized family $\mathcal{G}$ that consists of the following optimization problems $A(p)$.

PROBLEM $A(p)$. Using the "feasible solution" set

$$S(p) \triangleq \{d \mid (d,p) \in C\} ,$$

calculate both the "problem infimum"

$$\varphi(p) \triangleq \inf_{d \in S(p)} g(d,p)$$

and the "optimal solution" set

$$S^*(p) \triangleq \{d \in S(p) \mid g(d,p) = \varphi(p)\} .$$

For a given perturbation p, problem $A(p)$ is either "consistent" or "inconsistent", depending on whether the feasible solution set $S(p)$ is nonempty or empty. The (effective) domain of the infimum function φ is the "feasible perturbation" set

$$P \triangleq \{p \mid \text{problem } A(p) \text{ is consistent}\} ,$$

which is obviously identical to $\{p \mid (d,p) \in C \text{ for at least one } d\}$ and hence is not empty. Unlike the function g, the function φ may assume the value $-\infty$. However, for our purposes, it is not advantageous to follow Rockafellar's custom of artificially defining g and φ to be $+\infty$ outside their respective domains C and P.

Now, suppose that $g:C$ has a "conjugate transform" $h:D$; that is, suppose there is a function h with a nonempty domain

$$D \triangleq \{(q,e) \mid \sup_{(d,p)\in C} [\langle q,d\rangle + \langle e,p\rangle - g(d,p)] < +\infty\}$$

and function values

$$h(q,e) \triangleq \sup_{(d,p)\in C} [\langle q,d\rangle + \langle e,p\rangle - g(d,p)] \ .$$

The inner product $\langle q,d\rangle$ associates the "dual perturbation" parameter q with the "primal decision" variable d , and the inner product $\langle e,p\rangle$ associates the "dual decision" variable e with the "primal perturbation" parameter p .

Consider the parameterized family $\mathcal{B}$ that consists of the following optimization problems $B(q)$.

PROBLEM $B(q)$. <u>Using the feasible solution set</u>

$$T(q) \triangleq \{e \mid (q,e) \in D\} \quad ,$$

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

$$\psi(q) \triangleq \inf_{e\in T(q)} h(q,e)$$

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

$$T^*(q) \triangleq \{e \in T(q) | h(q,e) = \psi(q)\} .$$

Needless to say, the domain of the infimum function ψ is the feasible perturbation set

$$Q \triangleq \{q | \text{problem } B(q) \text{ is consistent}\} ,$$

which is clearly not empty.

Due to the known symmetry [3,8] of the conjugate transformation on the class of all closed convex functions $g:C$ (as well as the obvious symmetry of the preceding association of perturbation parameters and decision variables), families $\mathfrak{A}$ and $\mathfrak{B}$ are termed <u>Rockafellar dual families</u>, and problems $A(0)$ and $B(0)$ are termed <u>Rockafellar dual problems</u>. Actually, Rockafellar [7,8,9] formulates $\mathfrak{B}$ as a family of maximization problems by placing minus signs in front of the sup and e in the definition of $h:D$. Although that formulation facilitates specializations [7,8] to (the standard formulations of) linear programming duality and ordinary programming duality, the preceding formulation will facilitate our specialization to geometric programming duality.

To (re)orient the reader toward the preceding formulation, we now summarize Rockafellar's main results in terms of that formulation. In particular, <u>the primal infimum function</u> φ <u>is finite everywhere on its domain</u> P <u>and has a conjugate transform (with a nonempty domain and</u>

finite function values) if and only if the dual problem $B(0)$ *is consistent; in which case*

(i) *the dual objective function* $h(0,\cdot):T(0)$ *is the conjugate transform of* $\varphi:P$,

(ii) *the dual infimum* $\psi(0)$ *is finite if and only if* 0 *is in the domain* P^c *of the closed convex hull* $\varphi^c:P^c$ *of* $\varphi:P$, *in which event*

$$0 = \varphi^c(0) + \psi(0) \quad \text{and} \quad \partial\varphi^c(0) = T^*(0) \ ,$$

(iii) *if the primal problem* $A(0)$ *is also consistent, then* 0 *is in the domain* P^c *and*

$$\varphi^c(0) \le \varphi(0) \ , \ \text{with equality only if} \ \ \partial\varphi^c(0) = \partial\varphi(0) \ ,$$

(iv) *given that* $g:C$ *is convex and closed,* φ *is convex on* P *and can differ from* φ^c *only at relative boundary points of* P .

It is important to note that *the preceding results involve the whole Rockafellar family* $\mathcal{A}$ *but only one problem from the whole Rockafellar dual family* $\mathcal{B}$ -- *the Rockafellar dual problem* $B(0)$.

3. *Geometric duality*. Using the notation given in section 2.2 and subsection 3.3.5 of [5], assume that

$$d \triangleq (x,\kappa) \quad \text{and} \quad p \triangleq (u,\mu) \ ;$$

and suppose that

$$C \triangleq \{(x,\kappa,u,\mu) \mid x^k+u^k \in C_k,\ k\in\{0\}\cup I;\ (x^j+u^j,\kappa_j)\in C_j^+,\ j\in J;$$
$$x\in X;\ \text{and}\ g_i(x^i+u^i)+\mu_i \leq 0,\ i\in I\}$$

and

$$g(x,\kappa,u,\mu) \triangleq G(x+u,\kappa) \triangleq g_0(x^0+u^0) + \sum_J g_j^+(x^j+u^j,\kappa_j) \ .$$

Then, the Rockafellar family $\mathfrak{C}$ is the "geometric programming family" F (described in subsection 3.3.5 of [5]); and the crucial question now is whether the Rockafellar dual family $\mathfrak{B}$ is the "geometric programming dual family" G (described in subsections 3.3.4 and 3.3.5 of [5]). To obtain the answer, we need to compute the conjugate transform $h:D$ of $g:C$ in terms of both the "dual" Y of the given cone X and the conjugate transforms $h_k:D_k$ of the given functions $g_k:C_k, k\in\{0\}\cup I\cup J$.

To compute $h:D$, assume that

$$q \triangleq (v,\nu) \quad \text{and} \quad e \triangleq (z,\lambda) \ ,$$

where v has the same component partitioning as x , and where z has the same component partitioning as u . Then

$$h(v,\nu,z,\lambda) = \sup_{(x,\kappa,u,\mu)} \{ \langle v^0,x^0\rangle + \sum_I \langle v^i,x^i\rangle + \sum_J \langle v^j,x^j\rangle + \sum_J \nu_j\kappa_j$$

$$+ \langle z^0,u^0\rangle + \sum_I \langle z^i,u^i\rangle + \sum_J \langle z^j,u^j\rangle + \sum_I \lambda_i\mu_i - g_0(x^0+u^0)$$

$$- \sum_J g_j^+(x^j+u^j,\kappa_j) \,|\, x^0+u^0 \in C_0; x^i+u^i \in C_i, i\in I; (x^j+u^j,\kappa_j) \in C_j^+, j\in J;$$

$$x \in X; \text{ and } g_i(x^i+u^i) + \mu_i \le 0,\ i \in I\},$$

which is clearly finite only if $\lambda_i \ge 0$, $i \in I$; in which case

$$h(v,\nu,z,\lambda) = \sup_{(x,\kappa,u)} \{ [\langle v^0,x^0\rangle + \langle z^0,u^0\rangle - g_0(x^0+u^0)]$$

$$+ \sum_I [\langle v^i,x^i\rangle + \langle z^i,u^i\rangle - \lambda_i g_i(x^i+u^i)]$$

$$+ \sum_J [\langle v^j,x^j\rangle + \langle z^j,u^j\rangle + \nu_j\kappa_j - g_j^+(x^j+u^j,\kappa_j)] \,|$$

$$x^0+u^0 \in C_0; x^i+u^i \in C_i, i \in I; (x^j+u^j,\kappa_j) \in C_j^+, j \in J; x \in X\}$$

or

$$h(v,\nu,z,\lambda) = \sup_{(x,\kappa,u)} \{ [\langle z^0+v^0,x^0+u^0\rangle - g_0(x^0+u^0)]$$

$$+ \sum_I [\langle z^i+v^i,x^i+u^i\rangle - \lambda_i g_i(x^i+u^i)]$$

$$+ \sum_J [\langle z^j+v^j,x^j+u^j\rangle + \nu_j\kappa_j - g_j^+(x^j+u^j,\kappa_j)]$$

$$- \langle z,x\rangle - \langle v,u\rangle \,|\, x^0+u^0 \in C_0; x^i+u^i \in C_i, i \in I;$$

$$(x^j+u^j,\kappa_j) \in C_j^+, j \in J; x \in X\} .$$

The following lemma provides another condition that is necessary for the finiteness of the preceding expression.

Lemma A. The preceding expression for $h(v,\nu,z,\lambda)$ is finite only if $z - v \in Y$.

Proof. If $z - v \notin Y$, then there is an $\bar{x} \in X$ such that $\langle z-v,\bar{x}\rangle < 0$; in which event we choose $\bar{c}$ so that: $\bar{c}^0 \in C_0$; $\bar{c}^i \in C_i$, $i \in I$; and $(\bar{c}^j,\kappa_j) \in C_j^+$, $j \in J$, for some fixed $\kappa \geq 0$. Letting $x(s) \triangleq s\bar{x}$ and $u(s) \triangleq \bar{c} - s\bar{x}$, we observe that $(x(s),\kappa,u(s))$ satisfies the restrictions on (x,κ,u) for each $s \geq 0$. Thus

$$\{[\langle z^0+v^0,\bar{c}^0\rangle - g_0(\bar{c}^0)] + \sum_I [\langle z^i+v^i,\bar{c}^i\rangle - \lambda_i g_i(\bar{c}^i)]$$

$$+ \sum_J [\langle z^j+v^j,\bar{c}^j\rangle + \nu_j\kappa_j - g_j^+(\bar{c}^j,\kappa_j)] - \langle z,s\bar{x}\rangle - \langle v,\bar{c}-s\bar{x}\rangle\}$$

$$\leq h(v,\nu,z,\lambda)$$

for each $s \geq 0$;

and hence $h(v,\nu,z,\lambda) = +\infty$ because

$$\lim_{s \to +\infty} \{-\langle z,s\bar{x}\rangle - \langle v,\bar{c}-s\bar{x}\rangle\} = \lim_{s \to +\infty} \{-\langle v,\bar{c}\rangle + s\langle v-z,\bar{x}\rangle\} = +\infty$$

by virtue of the property $\langle z-v,\bar{x}\rangle < 0$. q.e.d.

Now, if $z-v \in Y$, then $z = y+v$ for some $y \in Y$; in which case

$$h(v,\nu,z,\lambda) = \sup_{(x,\kappa,u)} \{ [\langle z^0+v^0, x^0+u^0\rangle - g_0(x^0+u^0)]$$

$$+ \sum_I [\langle z^i+v^i, x^i+u^i\rangle - \lambda_i g_i(x^i+u^i)]$$

$$+ \sum_J [\langle z^j+v^j, x^j+u^j\rangle + \nu_j\kappa_j - g_j^+(x^j+u^j,\kappa_j)]$$

$$- \langle y,x\rangle - \langle v,x+u\rangle \,|\, x^0+u^0 \in C_0;\ x^i+u^i \in C_i,\ i \in I\ ;$$

$$(x^j+u^j,\kappa_j) \in C_j^+,\ j \in J\ ;\ x \in X\} \ .$$

Since $0 \le \langle y,x\rangle$ for each $x \in X$, it is easily seen that

$$h(v,\nu,z,\lambda) = \sup_{(\kappa,u)} \{ [\langle z^0+v^0, u^0\rangle - g_0(u^0)]$$

$$+ \sum_I [\langle z^i+v^i, u^i\rangle - \lambda_i g_i(u^i)]$$

$$+ \sum_J [\langle z^j+v^j, u^j\rangle + \nu_j\kappa_j - g_j^+(u^j,\kappa_j)] - \langle v,u\rangle \,|$$

$$u^0 \in C_0;\ u^i \in C_i,\ i \in I;\ (u^j,\kappa_j) \in C_j^+,\ j \in J\}$$

$$= \{ \sup_{u^0 \in C_0} [\langle z^0,u^0\rangle - g_0(u^0)] + \sum_I \sup_{u^i \in C_i} [\langle z^i,u^i\rangle - \lambda_i g_i(u^i)]$$

$$+ \sum_J \sup_{(u^j,\kappa_j)\in C_j^+} [\langle z^j,u^j\rangle + \nu_j\kappa_j - g_j^+(u^j,\kappa_j)]\} \ .$$

Consequently, $(v,\nu,z,\lambda) \in D$ if and only if: $\lambda_i \ge 0$, $i \in I$; $z = y+v$ for some $y \in Y$; and each term on the right-hand side of the preceding equations is finite. Of course, the first term is finite if and only if $z^0 \in D_0$, in which case the first term is equal to $h_0(z^0)$. The finiteness of the remaining terms can be conveniently characterized with two lemmas.

The following lemma characterizes the finiteness of the terms involving the index set I.

Lemma B. Given that $\lambda_i \geq 0$, the $\sup_{u^i \in C_i} [\langle z^i, u^i \rangle - \lambda_i g_i(u^i)]$ is finite if and only if $(z^i, \lambda_i) \in D_i^+$, in which case

$$\sup_{u^i \in C_i} [\langle z^i, u^i \rangle - \lambda_i g_i(u^i)] = h_i^+(z^i, \lambda_i) .$$

Proof. Simply observe that

$$\sup_{u^i \in C_i} [\langle z^i, u^i \rangle - \lambda_i g_i(u^i)] = \begin{cases} \sup_{u^i \in C_i} \langle z^i, u^i \rangle & \text{if } \lambda_i = 0 \\ \lambda_i h_i(z^i/\lambda_i) & \text{if } \lambda_i > 0 \text{ and } z^i \in \lambda_i D_i \\ +\infty & \text{if } \lambda_i > 0 \text{ and } z^i \notin \lambda_i D_i , \end{cases}$$

and then use the defining formula for $h_i^+ : D_i^+$. q.e.d.

The next lemma characterizes the finiteness of the terms involving the index set J .

Lemma C. The $\sup_{(u^j, \kappa_j) \in C_j^+} [\langle z^j, u^j \rangle + \nu_j \kappa_j - g_j^+(u^j, \kappa_j)]$ is finite if and only if both $z^j \in D_j$ and $h_j(z^j) + \nu_j \leq 0$, in which case

$$\sup_{(u^j, \kappa_j) \in C_j^+} [\langle z^j, u^j \rangle + \nu_j \kappa_j - g_j^+(u^j, \kappa_j)] = 0 .$$

Proof. First, observe that

$$\sup_{(u^j,\kappa_j)\in C_j^+} [\langle z^j,u^j\rangle + \nu_j\kappa_j - g_j^+(u^j,\kappa_j)]$$

$$= \sup_{\kappa_j\geq 0} [\sup_{u^j}\{\langle z^j,u^j\rangle + \nu_j\kappa_j - g_j^+(u^j,\kappa_j) | (u^j,\kappa_j)\in C_j^+\}]$$

$$= \sup_{\kappa_j\geq 0} [\nu_j\kappa_j + \sup_{u^j}\{\langle z^j,u^j\rangle - g_j^+(u^j,\kappa_j) | (u^j,\kappa_j)\in C_j^+\}]$$

$$= \sup_{\kappa_j\geq 0} \left[\nu_j\kappa_j + \begin{cases} \sup_{u^j}\{\langle z^j,u^j\rangle - \sup_{d^j\in D_j}\langle u^j,d^j\rangle \mid \sup_{d^j\in D_j}\langle u^j,d^j\rangle < +\infty\} & \text{if } \kappa_j = 0 \\ \sup_{u^j}\{\langle z^j,u^j\rangle - \kappa_j g_j(u^j/\kappa_j) \mid u^j/\kappa_j \in C_j\} & \text{if } \kappa_j > 0 \end{cases}\right]$$

$$= \sup_{\kappa_j\geq 0} \left[\nu_j\kappa_j + \begin{pmatrix} 0 & \text{if } \kappa_j = 0 \text{ and } z^j \in \overline{D}_j \\ +\infty & \text{if } \kappa_j = 0 \text{ and } z^j \notin \overline{D}_j \\ +\infty & \text{if } \kappa_j > 0 \text{ and } z^j \notin D_j \\ \kappa_j h_j(z^j) & \text{if } \kappa_j > 0 \text{ and } z^j \in D_j \end{pmatrix}\right] ,$$

where the final step makes use of the fact that the zero function with domain $\overline{D}_j$ (the topological closure of D_j) is the conjugate transform of the conjugate transform of the zero function with domain D_j . Now, note that the last expression is finite only if $z^j \in D_j$, in which case the last expression clearly

$$= \sup_{\kappa_j\geq 0} [\nu_j\kappa_j + \kappa_j h_j(z^j)] \quad .$$

But this expression is obviously finite if and only if $h_j(z^j) + \nu_j \le 0$, in which case this expression is clearly zero. q.e.d.

We have now shown that

$$D = \{(v,\nu,z,\lambda) \mid \lambda_i \ge 0,\ i \in I;\ z=y+v \text{ for some } y \in Y;\ z^0 \in D_0;\ (z^i,\lambda_i) \in D_i^+, i \in I;\ z^j \in D_j \text{ and } h_j(z^j)+\nu_j \le 0, j \in J\}$$

and

$$h(v,\nu,z,\lambda) = h_0(z^0) + \sum_I h_i^+(z^i,\lambda_i) .$$

Consequently, the Rockafellar dual problem $B(0)$ is the "geometric programming dual problem" B (described in subsection 3.3.4 of [5]). Although the Rockafellar dual family $\mathcal{B}$ is slightly different from the geometric programming dual family G (alluded to in subsection 3.3.5 of [5]), the difference is inconsequential in view of the final paragraph of section 2.

However, the relation $z = y+v$ shows that y can be used instead of z as a dual decision variable -- a change of variables that clearly induces a one-to-one mapping from $\mathcal{B}$ onto G. In particular, this mapping simply translates the (domain of the) dual objective function $h(v,\nu;\cdot)$ through $(-v,0)$ -- a mapping that clearly leaves the problem infimum $\psi(v,\nu)$ invariant while translating the optimal solution set $T^*(v,\nu)$ through $(-v,0)$.

References

[1] R.J. Duffin, "Dual programs and minimum cost", SIAM Journal on Applied Mathematics 10(1962) 119.

[2] J.E. Falk, "Lagrange multipliers and nonlinear programming", Journal of Mathematical Analysis and Applications 19(1967) 141.

[3] W. Fenchel, "Convex cones, sets and functions", Mathematics Department mimeographed lecture notes, Princeton University, Princeton, New Jersey (1951).

[4] E.L. Peterson, "Symmetric duality for generalized unconstrained geometric programming", SIAM Journal on Applied Mathematics 19(1970) 487.

[5] E.L. Peterson, "Geometric programming", SIAM Review 18(1976) 1.

[6] R.T. Rockafellar, "Duality and stability in extremum problems involving convex functions", Pacific Journal of Mathematics 21(1967) 167.

[7] R.T. Rockafellar, "Duality in nonlinear programming", American Mathematical Society Lectures in Applied Mathematics 11, G.B. Dantzig and A.F. Veinott, eds., (American Mathematical Society, Providence, 1968) 401.

[8] R.T. Rockafellar, Convex Analysis, Princeton University Press, Princeton, New Jersey (1970).

[9] R.T. Rockafellar, "Conjugate duality and optimization", Conference Board of the Mathematical Sciences Regional Conference Series in Applied Mathematics, No. 16, (Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania, 1974) 1.

[10] P. Wolfe, "A duality theorem for nonlinear programming", Quarterly of Applied Mathematics 19(1961) 239.