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 Duality
Rockafellar programming Parametric programming
Conjugate transformation Sensitivity analysis

CONTENTS

1. Introduction. ................................. 2
2. Rockafellar duality ......................... 2
3. Geometric duality ......................... 6
References. ................................ 14


The United States Government is authorized to reproduce and distribute reprints of this paper for Governmental purposes notwithstanding any copyright notation hereon.
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 \( g: \mathbb{R}^n \to \mathbb{R} \) is a (proper) function \( g \) with a nonempty (effective) domain \( \mathcal{C} \subseteq \mathbb{R}^n \), and assume that the independent variable \((d,p)\) in \( \mathcal{C} \) 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

\[ S(p) \triangleq \{ d | (d,p) \in C \} , \]

calculate both the "problem infimum"

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

and the "optimal solution" set

\[ S^o(p) \triangleq \{ d \in S(p) | g(d,p) = \varphi(p) \} . \]

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 \( \varphi \) is the "feasible perturbation" set

\[ P \triangleq \{ p | \text{problem A(p) is consistent} \} , \]

which is obviously identical to \( \{ p | (d,p) \in C \text{ for at least one d} \} \) and hence is not empty. Unlike the function g, the function \( \varphi \) may assume the value \(-\infty\). However, for our purposes, it is not advantageous to follow Rockafellar's custom of artificially defining g and \( \varphi \) 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

\[
D \triangleq \{ (q,e) | \sup_{(e,p) \in C} \langle q,d \rangle + \langle e,p \rangle - g(d,p) \leq \varepsilon \}
\]

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 \( \mathcal{B}(q) \).

**Problem \( \mathcal{B}(q) \).** Using the feasible solution set

\[
\mathcal{T}(q) \triangleq \{ e | (q,e) \in D \}
\]

calculate both the problem infimum

\[
\psi(q) \triangleq \inf_{e \in \mathcal{T}(q)} h(q,e)
\]
and the optimal solution set

\[ T^p(q) \triangleq \{ e \in T(q) | h(q,e) = q(q) \} \]

Needless to say, the domain of the infimum function \( q \) 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 \([5, 6]\) 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 \( G \) and \( B \) are termed Rockafellar dual families, and problems \( A(0) \) and \( B(0) \) are termed Rockafellar dual problems. Actually, Rockafellar \([7, 8, 9]\) formulates \( B \) 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, the primal infimum function \( q \) is finite everywhere on its domain \( P \) and has a conjugate transform (with a nonempty domain and
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 \( \mathcal{X}_0 \) of the closed convex hull \( \varphi_0^+P \) of \( \varphi;P \), in which event

\[
0 = \varphi_0^+(0) + \psi(0) \quad \text{and} \quad \partial \varphi_0^+(0) = T^*(0),
\]

(iii) If the primal problem \( A(0) \) is also consistent, then \( 0 \) is in the domain \( P^c \) and

\[
\varphi_0^+(0) \leq \sigma(0), \quad \text{with equality only if} \quad \partial \varphi_0^+(0) = \partial \sigma(0),
\]

(iv) given that \( g: C \) is convex and closed, \( \varphi \) is convex on \( P \) and can differ from \( \varphi_0^+ \) only at relative boundary points of \( P \).

It is important to note that the preceding results involve the whole Rockafellar family \( G \) but only one problem from the whole Rockafellar dual family \( \Theta \) -- the Rockafellar dual problem \( B(0) \).

3. Geometric duality. Using the notation given in section 2.2 and subsection 5.3.5 of [5], assume that

\[
\hat{c} \triangleq (x,k) \quad \text{and} \quad \hat{p} \triangleq (u,\lambda);
\]
and suppose that

\[ C \triangleq \{ (x, x^1, u, \mu) | x^k + u^k \leq c_k, \mu \in \mathbb{R} \cup \{0\} \}, \quad \{ x^j + u^j, k_j \} \in C^*, \quad j \in J; \]

\[ x \in X; \quad \text{and} \quad g_i(x^j + u^j) + \mu_i \leq 0, \quad i \in I \]

and

\[ g(x, x^1, u, \mu) \triangleq \mathcal{G}(x^1, u^1, \mu) \triangleq g_0(x^0 + u^0) + \sum_j g_j^*(x^j + u^j, k_j). \]

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

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

\[ \sim 7 \]
the same component partitioning as $u$. Then

$$h(v, u, x, \lambda) = \sup_{(x, \lambda, u, \lambda)} \left[ \langle v, x^0 \rangle + \sum_j \langle v_j, x^J \rangle + \sum_j \lambda_j \kappa_j \right.$$  

$$+ \langle x^0, u^0 \rangle + \sum_j \langle x_j, u_j \rangle + \sum_j \lambda_j\beta_j(x^0 + u^0)$$  

$$- \sum_j \beta_j^2(x_j + u_j, \kappa_j) \rangle x^0 + u^0 \in C^0_\nu x^0 + u^0 \in C^\nu_\kappa, x \in \mathbb{X} \right.$$

$$x \in \mathbb{X} \} \text{ and } \beta_j(x^0 + u^0) + \lambda_j < 0, \lambda_j \in \mathbb{R} \},$$

which is clearly finite only if $\lambda_j \geq 0, \lambda_j \in \mathbb{R} \} \text{ in which case}$

$$h(v, u, x, \lambda) = \sup_{(x, \lambda, u, \lambda)} \left[ \langle v, x^0 \rangle + \langle x^0, u^0 \rangle - \sum_j \beta_j^2(x_j + u_j) \right]$$

$$+ \sum_j \left[ \langle v_j, x^j \rangle + \langle x_j, u_j \rangle - \lambda_j \beta_j^2(x_j + u_j) \right]$$  

$$+ \sum_j \left[ \langle v_j, x_j \rangle + \langle x_j, u_j \rangle + \lambda_j \kappa_j - \beta_j^2(x_j + u_j, \kappa_j) \right]$$

$$x^0 + u^0 \in C^0_\nu x^0 + u^0 \in C^\nu_\kappa, x \in \mathbb{X} \} \text{ and } \beta_j(x^0 + u^0) + \lambda_j < 0, \lambda_j \in \mathbb{R} \} \text{ in which case}$$

or

$$h(v, u, x, \lambda) = \sup_{(x, \lambda, u, \lambda)} \left[ \langle x^0, v^0 \rangle + \langle x^0, u^0 \rangle - \sum_j \beta_j^2(x_j + u_j) \right]$$

$$+ \sum_j \left[ \langle x_j, v_j \rangle + \langle x_j, u_j \rangle - \lambda_j \beta_j^2(x_j + u_j) \right]$$

$$+ \sum_j \left[ \langle x_j, v_j \rangle + \langle x_j, u_j \rangle + \lambda_j \kappa_j - \beta_j^2(x_j + u_j, \kappa_j) \right]$$

$$- \langle x, v \rangle x^0 + u^0 \in C^0_\nu x^0 + u^0 \in C^\nu_\kappa, x \in \mathbb{X} \} \text{ and } \beta_j(x^0 + u^0) + \lambda_j < 0, \lambda_j \in \mathbb{R} \} \text{ in which case}$$

$$\langle x_j, \kappa_j \rangle \in C^+_{\kappa, j} x \in \mathbb{X} \} \text{ and } \beta_j(x^0 + u^0) + \lambda_j < 0, \lambda_j \in \mathbb{R} \}$$

- 3 -
The following lemma provides another condition that is necessary for the finiteness of the preceding expression.

**Lemma A.** The preceding expression for \( b(v, u, z, \lambda) \) is finite only if \( z - v \in Y \).

**Proof.** If \( z - v \not\in Y \), then there is an \( x \in X \) such that \( \langle z - v, x \rangle < 0 \); in which event we choose \( \overline{c} \in \mathfrak{c} \) so that: \( c_0 \in c_j \); \( c_k \in c_k \), \( 0 \leq k \leq I \); and \( \langle c_j, x \rangle \in c_j^+ \), \( j \in J \), for some fixed \( k \geq 0 \). Letting \( x(s) \triangleq s \overline{x} \) and \( u(s) \triangleq s \overline{z} \), we observe that \( x(s)x, u(s) \) satisfies the restrictions on \( (x, t, u) \) for each \( s \geq 0 \). Thus

\[
\begin{align*}
[\langle (z - v, \overline{c}^0) - g_0 \overline{c}^0 \rangle] & + \sum_i [\langle z - v, \overline{c}_i^1 \rangle - \lambda_i g_i \overline{c}_i^1 ] \\
& + \sum_j [\langle z - v, \overline{c}_j^1 \rangle - \langle x, \overline{c}_j^2 \rangle - \langle v, \overline{c}_j^2 \rangle] & \leq b(v, u, z, \lambda)
\end{align*}
\]

for each \( s \geq 0 \);

and hence \( b(v, u, z, \lambda) = +\infty \) because

\[
\lim_{s \to +\infty} [\langle z, \overline{x} \rangle - \langle v, \overline{c} \rangle + x\langle v, \overline{z} \rangle] = +\infty
\]

by virtue of the property \( \langle z - v, x \rangle < 0 \).

**q.e.d.**

Now, if \( z - v \not\in Y \), then \( z = y + v \) for some \( y \in Y \); in which case

- 9 -
\[ h(v, u, x, \lambda) = \sup_{x, u, h} \left\{ \langle \zeta^0 + v^0, x^0 + u^0 \rangle - g_0(x^0) \right\} + \sum_i \left\{ \langle \zeta_i^0 + v_i^0, x_i^0 + u_i^0 \rangle - \lambda_i g_4(x_i^0) \right\} + \sum_j \left\{ \langle \zeta_j^0 + v_j^1, x_j^1 + u_j^1 \rangle + v_j k_j^1 - g_j^1(x_j^0, k_j^0) \right\} - \langle v, x \rangle - \langle v, x \rangle x^0 \in C_0; x^0 + u^0 \in C_1; i \in I; \ z \in C_0 \right\} \]

Since \( 0 \leq \langle y, x \rangle \) for each \( x \in X \), it is easily seen that

\[ h(v, u, x, \lambda) = \sup_{x, u, h} \left\{ \langle \zeta^0 + v^0, u^0 \rangle - g_0(u^0) \right\} \]

\[ + \sum_i \left\{ \langle \zeta_i^0 + v_i^0, u_i^0 \rangle - \lambda_i g_4(u_i^0) \right\} \]

\[ + \sum_j \left\{ \langle \zeta_j^0 + v_j^1, u_j^1 \rangle + v_j k_j^1 - g_j^1(u_j^0, k_j^0) \right\} - \langle v, u \rangle \]

\[ u^0 \in C_0; u^1 \in C_1; i \in I; u_j^0 \in C_{j1}; j \in J \}

\[ = \sup_{u^0 \in C_0} \left\{ \langle \zeta^0, u^0 \rangle - g_0(u^0) \right\} + \sum_i \sup_{u_i^0 \in C_{i1}} \left\{ \langle \zeta_i^0, u_i^0 \rangle - \lambda_i g_4(u_i^0) \right\} + \sum_j \sup_{u_j^0 \in C_{j1}} \left\{ \langle \zeta_j^0, u_j^0 \rangle + v_j k_j^1 - g_j^1(u_j^0, k_j^0) \right\} \]

Consequently, \( (v, u, x, \lambda) \in D \) if and only if: \( \lambda_i \geq 0, i \in I \); \( z = y + v \) 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 \( \zeta^0 \in B_0 \), in which case the first term is equal to \( h_0(0^0) \). 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 $\lambda_1 \geq 0$, the $\sup_{u^1 \in \mathcal{C}_1} [(x^1, u^1) - \lambda_1 h^*_1(u^1)]$ is finite if and only if $(x^1, \lambda_1) \in D_1^*$, in which case

$$
\sup_{u^1 \in \mathcal{C}_1} [(x^1, u^1) - \lambda_1 h^*_1(u^1)] = h^*_1(x^1, \lambda_1).
$$

**Proof.** Simply observe that

$$
\begin{align*}
\sup_{u^1 \in \mathcal{C}_1} [(x^1, u^1) - \lambda_1 h^*_1(u^1)] &= \begin{cases} 
\sup_{u^1 \in \mathcal{C}_1} (x^1, u^1) & \text{if } \lambda_1 = 0 \\
\lambda_1 b_1(x^1/\lambda_1) & \text{if } \lambda_1 > 0 \text{ and } x^1 \in \lambda_1 D_1 \\
+ \infty & \text{if } \lambda_1 > 0 \text{ and } x^1 \notin \lambda_1 D_1,
\end{cases}
\end{align*}
$$

and then use the defining formula for $h^*_1$.

$q.e.d.$

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

**Lemma C.** The $\sup_{(u^j, \kappa_j) \in \mathcal{C}_j^*} [(x^j, u^j) + \gamma_j \kappa_j - g^+_j(u^j, \kappa_j)]$ is finite if and only if both $x^j \notin \mathcal{P}_j$ and $h_j(x^j) + \gamma_j \leq 0$, in which case

$$
\sup_{(u^j, \kappa_j) \in \mathcal{C}_j^*} [(x^j, u^j) + \gamma_j \kappa_j - g^+_j(u^j, \kappa_j)] = 0.
$$
Proof. First, observe that

\[
\sup_{(u^j, \kappa_j) \in C_j^*} \{ \langle x^j, u^j \rangle + v_j \kappa_j - g_j^*(u^j, \kappa_j) \}
\]

\[
= \sup_{\kappa_j \geq 0} \sup_{u^j \in \mathbb{R}^j} \{ \langle x^j, u^j \rangle + v_j \kappa_j - g_j^*(u^j, \kappa_j) \}
\]

\[
= \sup_{\kappa_j \geq 0} \left\{ \sup_{u^j \in \mathbb{R}^j} \langle x^j, u^j \rangle - \sup_{d^j \in D_j} \langle u^j, d^j \rangle \right\} \sup_{d^j \in D_j} \langle u^j, d^j \rangle < +\infty \text{ if } \kappa_j = 0 \}
\]

\[
= \sup_{\kappa_j \geq 0} \left\{ \sup_{u^j \in \mathbb{R}^j} \langle x^j, u^j \rangle - \kappa_j \delta_j(u^j, \kappa_j) \right\} \quad \text{if } \kappa_j > 0 \}
\]

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 \( x^j \in D_j \), in which case the last expression clearly

\[
= \sup_{\kappa_j \geq 0} \left\{ v_j \kappa_j + \kappa_j \delta_j(x^j) \right\} \quad \text{if } \kappa_j > 0 \}
\]

- \( \star \) -
But this expression is obviously finite if and only if \( h_j(z^1_j) + v_j \leq 0 \), in which case this expression is clearly zero. \( \text{q.e.d.} \)

We have now shown that

\[
D = \{ (v_j,\nu_i,\lambda_j) \mid \lambda_j \geq 0, \ i \in I; \ z^0 - y + \nu \ \text{for some} \ y \in Y; \ z^0 \in D_0; \\
(z^1_j,\lambda_j) \in D_1^+; \ i \in I; \ z^1_j \in D_j \ \text{and} \ h_j(z^1_j) + v_j \leq 0, j \in J \}
\]

and

\[
h(v,\nu,z,\lambda) - h_0(z^0) = \sum_{j} h_j^*(z^1_j,\lambda_j) .
\]

Consequently, the Rockafellar dual problem \( B(0) \) is the "geometric programming dual problem" \( B \) (described in subsection 3.3.b of [5]). Although the Rockafellar dual family \( B \) is slightly different from the geometric programming dual family \( G \) (alluded to in subsection 3.3.5 of [7]), the difference is inconsequential in view of the final paragraph of section 2.

However, the relation \( z = y + \nu \) 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 \( B \) onto \( G \). In particular, this mapping simply translates the (convexity of the) dual objective function \( h(v,\nu;z) \) through \((-v,0)\) -- a mapping that clearly leaves the problem minimize \( \psi(v,\nu) \) invariant while translating the optimal solution set \( \mathcal{T}^* (v,\nu) \) through \((-v,0)\).
References


