OPTIMALITY CRITERIA FOR GENERAL UNCONSTRAINED GEOMETRIC PROGRAMMING PROBLEMS

TIBOR ILLÉS
Computer and Automation Institute, Hungarian Academy of Sciences
H-1052 Budapest XI.Kende u. 13-17. Hungary

(Received May, 1990)

Abstract. This paper presents a possible generalization of geometric programming problems. Such a generalization was proposed by Peterson [6], based on Rockafellar's [8] conjugate function theory. Using their results, we define a slightly different, more symmetric dual pair of general unconstrained geometric programming problems.

In the second chapter the conjugate function is defined and some of its properties are demonstrated. In the third chapter the general unconstrained geometric programming problem and its dual pair are introduced and some of its fundamental properties are proved. The primal optimality criteria is based on Peterson's papers [6,7] and the dual optimality criteria completes our examinations.

Key words: general geometric programming, conjugate function, optimality criteria, stationary point.

1. INTRODUCTION

As far as I know the first significant paper about the generalization of geometric programming problems was written by Peterson [6]. Peterson's generalization is based on the properties of conjugate functions.

The dual objective function is defined on the set of primal feasible solutions as the conjugate function of the primal objective function. It means that the dual objective function depends on the primal feasible solution set. Its converse is false: the primal objective function is independent from the set of dual feasible solutions. Thus the primal-dual pair is not symmetric.

We modified the definition of the general unconstrained geometric programming problems, and hence the dual objective function is also independent from the set of primal feasible solutions as the primal objective function was independent from the set of dual feasible solutions in Peterson's [6,7] papers.

So, our primal-dual problems are more symmetrical than Peterson's. After a careful examination of Peterson's primal optimality condition, we found that this is still valid for our primal problem. A dual optimality condition is developed as well. For this reason it was necessary to assume the differentiability of the conjugate function at the dual critical solution. Finally, some conditions are specified for the differentiability of the conjugate function. These conditions are different from Rockafellar's [8] results.

Let us present (without a proof) a well known and useful result.

Theorem 1.1

(see [1], p. 451) Let a function \( f : H \to \mathbb{R}^p \), \( f(H) \subseteq K \subseteq \mathbb{R}^p \) and function \( g : K \to \mathbb{R}_+ \), \( H \subseteq \mathbb{R}^m \) be given. Let \( a \in \text{int} H \), and \( b = f(a) \in \text{int} K \) be arbitrary points. Assume that function \( f \) is differentiable at point \( a \) and function \( \xi \) at point \( b \). Then the composite function \( h = f \circ g \) is differentiable at point \( a \), and

\[
h_{\xi_i}(a) = \sum_k g_{\eta_k}(a)f_{\xi_i}(a).
\]

holds.

Sets are denoted by capital and vectors or functions are denoted by small Latin letters, their components are denoted by the corresponding Greek letters.
2. CONJUGATE FUNCTIONS

The concept of conjugate functions in connection with convex functions was defined by Fenchel [2]. The theory of conjugate functions was developed by Rockafellar [8,9] and he applied this theory to convex programming problems. Some years later, Peterson [6] generalized the theory for nonconvex functions. By this generalization Rockafellar’s method is applicable for nonconvex programming problems, as well.

Our definition of conjugate functions in the nonconvex case is slightly different from the usual one, but preserves all the important properties of Peterson’s conjugate function. This modification guarantees that the constraints of the primal and dual problems become symmetrical not only in formal sense (as in Peterson’s papers [6,7]), but practically, too.

**Definition 2.1**

Let a function $g : D_g \to \mathbb{R}$ (where $D_g \subseteq \mathbb{R}^n$ is the domain of $g$), be given. Denote

$$h(y) := \sup_{x \in D_g} \{g(x) - <x, y>\},$$

where function $h : \mathbb{R}^n \to \mathbb{R}$ is the conjugate function of $g$ defined on the set

$$D_h := \{y \in \mathbb{R}^n : \sup_{x \in D_g} \{g(x) - <x, y>\} < +\infty\} \subseteq \mathbb{R}^n.$$

Some of the basic properties of the conjugate functions are summarized below:

**Remark 2.1**

Let $A$ be a subset of $D_g$ and

$$\hat{h}(y) := \sup_{x \in A} \{g(x) - <x, y>\},$$

where $\hat{h}$ is the conjugate function of $g$ defined on the set

$$D_{\hat{h}} := \{y \in \mathbb{R}^n : \sup_{x \in A} \{g(x) - <x, y>\} < +\infty\}.$$

Then $D_h \subseteq D_{\hat{h}}$ and $h(y) \geq \hat{h}(y)$ ($\forall y \in D_h$).

Remark 2.1. is an immediate consequence of Definition 2.1. and it shows how the properties of the conjugate function change depending on its domain. In Peterson’s [6] above mentioned paper the conjugate function of a $g$ nonconvex function is defined on the intersection of its domain $D_g$ and a cone. It is obvious from Remark 2.1., that for different cones, different conjugate functions are obtained.

**Remark 2.2**

*(Fenchel’s-inequality)* The following inequality can be derived from Definition 2.1.

$$g(x) - h(y) \leq <x, y> \quad (\forall x \in D_g, \forall y \in D_h).$$

From Definition 2.1. it is easy to see that conjugate function may not exist for all function $g$ since $D_h$ can be the empty set.

**Definition 2.2**

Let a conjugate function $h : \mathbb{R}^n \to \mathbb{R}$ of function $g$ given. The set

$$\{(y, a) : a \geq g(x) - h(y) | y \in D_h, a \in \mathbb{R}, \forall x \in D_g\}$$

is called the epigraph of $h$ and it is denoted by epihh.

The general definition of the epigraph of a function can be found in Rockafellar’s [8] book.
Lemma 2.1

If the conjugate function $h$ exists, then
(i) the set $D_h$ is a convex set and $h$ is a convex function,
(ii) the epigraph of $h$ is a closed set.

Proof. (i) Let $y_1, y_2 \in D_h$ be two arbitrary points and $0 \leq \lambda \leq 1$. Then using the inequality, we find that

$$h(\lambda y_1 + (1 - \lambda) y_2) = \sup_{x \in D_g} \{\lambda [g(x) - < x, y_1 >] + (1 - \lambda) [g(x) - < x, y_2 >]\} \leq \sup_{x \in D_g} \sup_{\lambda \in [0,1]} \{\lambda g(x) - < x, y_1 >\} + (1 - \lambda) \sup_{x \in D_g} \{\lambda g(x) - < x, y_2 >\} = \lambda h(y_1) + (1 - \lambda) h(y_2).$$

Hence $h$ is a convex function, and $D_h$ is a convex set.

(ii) The set $\text{epi}h$ is a closed set because it is an intersection of closed half spaces.

As Rockafellar [9] mentioned, Fenchel was the first who proved this proposition.

3. THE GENERALIZED GEOMETRIC PROGRAMMING PROBLEMS AND THEIR MAIN LEMMA

Generalized geometric programming problems are defined in this chapter. Our definition slightly differs from Peterson's [6, 7]. The dual objective function at Peterson's model - as Remark 2.1. shows - depends on the cone used to define the set of primal feasible solutions. In our construction, the dual objective function is independent from the cone mentioned above because we define the conjugate function on the whole domain of the function $g$. Hence in our case, function $h$ depends only on function $g$. It is clear from above that at Peterson's model, the dual objective function depends on the set of primal feasible solutions (but it was not true conversely). We try to define more symmetrical problem pair in this paper. In the problem pair defined below the dual objective function is independent from the set of feasible solutions of the primal problem and vice versa. The new definition of the conjugate function makes it possible to define a more symmetrical problem pair.

Primal problem:

Let a function $g : D_g \to \mathbb{R}$, be given where $D_g \subseteq \mathbb{R}^n$ is the domain of $g$ and let $K \subseteq \mathbb{R}^n$ be a cone.

$$P := D_g \cap K$$

$$\sup\{g(x) | x \in P\}$$

In the definition of the primal problem we assumed that the set of primal feasible solutions is the intersection of a cone and the domain of function $g$. (The cone may be convex.) If $P$ is empty, let $\sup g(x) = -\infty$. 

\textit{Comment 21.1-8}
Dual problem:

Let \( h \) be the conjugate function of the objective function of the primal problem

\[
D := D_h \cap K^*
\]

\[
\inf\{h(y) | y \in D\}
\]

where \( K^* \) is a polar of \( K \).

It is well known from the Minkowski theorem [8] that \( K^* \) is a cone. It is easy to see that \( K^* \) is always a convex cone, and it is well known from the Farkas theorem [5], that \( K^{**} = \text{conv} K \).

Using Lemma 2.1., we can conclude the following evident remark.

**Remark 3.1**

The objective function \( h \) of the dual problem is a convex function, and the dual feasible solution set is convex, too. Thus the dual problem is a convex programming problem.

**Remark 3.2**

The epigraph of function \( h \) and \( D \) are closed sets. If the dual feasible solution set is bounded and the primal feasible solution set is not empty then, there exists a \( \bar{y} \in D \) such that

\[
\inf_{y \in D} h(y) = h(\bar{y}),
\]

hence the dual problem has an optimal solution.

Let us denote the set of the primal and dual optimal solutions by \( P^* \) and \( D^* \) respectively. The optimum value of the primal and dual objective functions are denoted by \( \varphi \) and \( \chi \).

**Lemma 3.1. (Main Lemma)**

If \( z \in P \) and \( y \in D \) then

\[
g(z) \leq h(y),
\]

with equality if and only if \( h(y) = \bar{h}(y) \) and \( g(z) = \sup_{\bar{z} \in P} \{g(\bar{z}) - < \bar{z}, y >\} \).

**Proof.** Using the definitions of primal and dual problems, and considering the definition of conjugate functions, we obtain the following relations:

\[
h(y) = \sup_{z \in D_y} \{g(z) - < z, y >\} \geq \sup_{x \in P} \{g(x) - < x, y >\} = \bar{h}(y) \geq g(x) - < x, y > \geq g(x)
\]

(3.1)

where \( P \leq D_y \) and the last two inequalities hold, because \( z \in K \) and \( y \in K^* \), imply

\[
< z, y > \leq 0
\]

hence

\[
- < z, y > \geq 0,
\]

and then

\[
h(y) \geq g(x).
\]

If \( g(x) = h(y) \) then \( h(\bar{y}) = \bar{h}(y) \) holds, and the last two inequalities in (3.1) are equalities, thus

\[
g(z) = \sup_{\bar{z} \in P} \{g(\bar{z}) - < \bar{z}, y >\} = \bar{h}(y).
\]
Conversely, the first inequality in (3.1) holds with equality because \( h(y) = \bar{h}(y) \) and the equalities in the last two inequalities follow from the

\[
g(x) = \sup_{\bar{x} \in P} \{ g(\bar{x}) - \langle \bar{x}, y \rangle \}
\]

assumption. Hence

\[
g(x) = h(y).
\]

Some corollaries of the Main Lemma are listed below:

**Corollary 3.1.** (i) If \( P \neq \emptyset \), the objective function of the dual problem is bounded from below.

(ii) If \( D \neq \emptyset \) the objective function of the primal problem is bounded from above.

**Corollary 3.2**

If \( g(x) = h(y) \) for some \( x \in P \) and \( y \in D \), then \( \langle x, y \rangle = 0 \).

**Corollary 3.3.** (Weak Equilibrium)

If \( g(x) = h(y) \) for some \( x \in P \) and \( y \in D \), then \( x \in P^* \) and \( y \in D^* \).

### 4. PRIMAL OPTIMALITY CRITERIA

The stationary point is defined in this chapter and the connection between the stationary point and the optimal solutions are analyzed.

**Definition 4.1**

Let a function \( g \) be differentiable at \( \bar{x} \in P \). The stationary point of the primal problem is a point \( \bar{x} \in P \) that satisfies the following constraints

\[
\langle \nabla g(\bar{x}), \bar{x} \rangle = 0 \quad \text{and} \quad \nabla g(\bar{x}) \in K^*
\]

Peterson [6,7] used the critical solution denomination for the above defined \( \bar{x} \in P \), but he mentioned that some other authors use the stationary or primal solution words instead of critical solution.

Some properties of the class of primal problems restricted to differentiable functions are given below. We would show that the constraint (4.1) is a primal optimality criteria for the restricted class of primal problems. The complementarity criteria of nonlinear programming problems are very similar to the assumption (4.1).

**Theorem 4.1**

Let a function \( g \) be differentiable at point \( x^* \) and let \( K \) be a given cone. If \( x^* \in P^* \) then \( x^* \in P \), where

\[
P_* := \{ \bar{x} \in P | \bar{x} \text{ satisfies (4.1)} \}
\]

is the set of (primal) stationary points. (But not conversely).

**Proof.** From \( x^* \in P^* \) we know that \( x^* \in P \). Considering the optimality of \( x^* \), the differentiability of function \( g \) at \( x^* \) and the convexity of cone \( K \), we obtain that

\[
\langle \nabla g(\bar{x}), \bar{x} \rangle \leq 0 \quad \text{for any} \quad x \in K
\]

\[
\langle \nabla g(\bar{x}), \bar{x} \rangle + 0(|x^* + z|) = g(x^* + z) + g(x^*) \leq 0,
\]

using that if \( x \in K \) then \( x^* + z \in K \) which arise from the convexity of cone \( K \). If \( 0 \leq \lambda \leq 1 \) and \( x^* \in K \) then

\[
x^* + \lambda(-x^*) \in K,
\]
where $\lambda(-x^*)$ is the direction of the derivative. In this case we get that
\[ < \nabla g(x^*), \lambda(-x^*) > = g(x^* + \lambda(-x^*)) + g(x^*) \leq 0, \]
and then
\[ < \nabla g(x^*), -x^* > \leq 0 \quad (4.3) \]
From (4.2) and (4.3) we conclude that
\[ < \nabla g(x^*), \lambda(-x^*) > = 0, \]
and (4.2) are simplified that $\nabla g(x^*) \in K^*$, namely $x^* \in P_s$.

The next theorem seems to possess the same information as the previous one, but about a slightly different class of problems.

**Theorem 4.2**
Let a function $g$ be differentiable at $x^* \in \text{int}P$. If $x^* \in P^*$ then $x^* \in P_s$.

**Proof.** If $x^* \in P^*$ then $g(x^*) > g(x)$ for any $x \in P$, thus $x^*$ is the global maximum point of function $g$ on set $P$. Since function $g$ is differentiable at point $x^*$, the partial derivatives exist at $x^*$. Considering that function $g$ has a global maximum at $x^* \in \text{int}P$, (thus it is the local maximum point, too) and the partial derivatives $g_{\xi_i}$ exist for all $\xi_i$ and
\[ g_{\xi_i}(x^*) = 0, \quad \text{for} \quad i = 1, 2, \ldots, n. \]
Then $\nabla g(x^*) = 0$, so
\[ < \nabla g(x^*), x^* > = 0 \quad \text{and} \quad < \nabla g(x^*), x > = 0 \quad \text{for any} \quad x \in P, \]
from this we get $\nabla g(x^*) \in K^*$.

Summarizing the previous facts we conclude that $x^* \in P_s$.

We provide a sufficient condition for the optimality of the stationary point.

**Theorem 4.3**
Let a function $g$ be differentiable at $\dot{x} \in P$, function $g$ be concave and set $D_g$ be convex. If $\dot{x} \in P^*$, then $\dot{x} \in P^*$.

**Proof.** For the function $g$,
\[ \delta g(\dot{x}, x - \dot{x}) = < \nabla g(\dot{x}), x - \dot{x} > \quad \text{for any} \quad x \in P, \quad (4.4) \]
where directional derivative is denoted by $\delta g$. Since $\dot{x} \in P_s$ then
\[ < \nabla g(\dot{x}), x - \dot{x} > \leq 0 \quad (4.5) \]
Since $g$ is concave, we have
\[ \delta g(\dot{x}, x - \dot{x}) \geq g(\dot{x} + (x - \dot{x})) - g(\dot{x}) = g(x) - g(\dot{x}) \quad (4.6) \]
From (4.4), (4.5), and (4.6), we get that
\[ 0 \geq g(x) - g(\dot{x}), \]
i.e.
\[ g(x) \geq g(\dot{x}) \quad \text{for any} \quad x \in P, \]
thus $\dot{x} \in P^*$. 
5. DUAL OPTIMALITY CRITERIA

We would like to preserve and transform the results of the previous chapter for the dual problem. For this reason we try to find such a system of conditions which guarantees that function $h$ is differentiable. The most general result connected with the differentiability of the conjugate function is made by Rockafellar [8]. But Rockafellar's theorem insures only the existence of a point in domain of conjugate function where the conjugate function is differentiable.

**Definition 5.1**

If function $h$ is differentiable on the set $D$ then any $\hat{y} \in D$ is a stationary point of dual problem which satisfies the following constraints

$$< \nabla h(\hat{y}), \hat{y} > = 0 \quad \text{and} \quad - \nabla h(\hat{y}) \in \text{conv}K$$

(5.1)

Denote

$$D_* := \{ \hat{y} \in D | \hat{y} \text{ satisfies (5.1)} \},$$

the set of (dual) stationary points. Definition 5.1. shows that Rockafellar's [8] theorem does not insure that a (dual) stationary point generally exists.

Constraint (5.1), like (4.1), is called dual optimality criteria.

First of all we consider a result which we can prove only if function $h$ is differentiable.

**Theorem 5.1**

Let the conjugate function $h$ be differentiable on its whole domain. (Function $h$ is the objective function of the dual problem.) The constraints (i) and (ii) are equivalent:

(i) $\hat{y} \in D^*$,

(ii) $\hat{y} \in D_*.$

**Proof.** (i) $\Rightarrow$ (ii) : Since cone $K^*$ is convex and function $h$ is differentiable, then (by Definition 5.1. and Theorem 4.1.) we get $\hat{y} \in D_*.$

(ii) $\Rightarrow$ (i) : In this case the constraints of Theorem 4.3. are satisfied, except that now we have a convex objective function to minimize instead of a concave objective function to maximize, but it is obvious, that these differences do not influence the validity of Theorem 4.3. So $\hat{y} \in D^*.$

A sufficient condition for the differentiability of the conjugate function $h$ is given.

**Lemma 5.1**

Let the conjugate function $h$ be the objective function of the dual problem. Suppose that $a \in \text{int}D_h$, and

(i) the supremum in definition of function $h$ is unambiguous,

(ii) there exist a differentiable function $l : D_h \rightarrow \mathbb{R}^n$ such that $l(a) = x_a \in \text{int}D_g$ and $l(D_h) \subseteq D_g$,

(iii) function $g$ is differentiable at $x_a \in \text{int}D_g$.

Then the function $h$ is differentiable at $a \in \text{int}D_h$ and

$$\nabla h(a) = \nabla g(l(a))[J_l(a)]^T - l(y) + l(a),$$

where $J_l$ is the Jacobi-matrix of function $l$.

**Proof.** It follows from the constraints that

$$h(y) - h(a) = \sup_{x \in D_g} \{ g(x) - < x, y > \} - \sup_{x \in D_g} \{ g(x) - < x, a > \} =$$

$$= g(x_y) - < x_y, y > - g(x_a) + < x_a, a > = g(l(y)) - l(y), y - g(l(a)) =< l(a), a > =$$

$$= g(l(y)) - g(l(a)) - < l(y) - l(a), y - a > .$$

The theorem that was mentioned in the introduction can be applied here that gives the following equality

$$h(y) - h(a) =< \nabla (l(a))[J_l(a)]^T - l(y) + l(a), y - a > + O(|y - a|),$$

i.e.

$$\nabla h(a) = \nabla g(l(a))[J_l(a)]^T - l(y) + l(a).$$

The necessary and sufficient condition for a dual optimal solution to be a dual stationary point is the following:
Theorem 5.2

Let the conjugate function $h$ be the objective function of the dual problem and $y \in \text{int}D_h$. Suppose that the following constraints are satisfied:

(i) the supremum in definition of function $h$ is unambiguous,
(ii) there exists a differentiable function $l : D_h \to \mathbb{R}^n$ such $l(y) = x_g \in \text{int}D_g$ and $l(D_h) \subset D_g$,
(iii) function $g$ is differentiable at $x_g \in \text{int}D_g$.

Then $y$ is an optimal solution of the dual problem if and only if it is the (dual) stationary point.

Proof. It is obvious from Lemma 5.1. and Theorem 5.1.

References

1. Á. Császár, Real Variable Function's Analysis i., [in Hungarian].