Entropy constrained programs and geometric duality obtained via Fenchel-Lagrange duality approach*

Radu Ioan Boţ† Sorin-Mihai Grad‡ Gert Wanka§

Abstract. Having an optimization problem with linear objective function, linear inequality and maximum entropy inequality constraints, we determine a dual to it. Therefore we use a conjugacy approach which bases on the perturbation theory. As the main results we prove that the geometric dual problems introduced by Peterson for, both, unconstrained and constrained optimization problems can be obtained by using the same perturbation theory. Furthermore, necessary and sufficient optimality conditions are derived using strong duality.

Keywords. Geometric programming; Convex programming; Perturbation theory; Duality; Entropy optimization

2000 Mathematics Subject Classification. 49N15, 54C70, 90C25

1 Introduction

The idea behind this paper was born when we noticed that the dual problems obtained in some works (like [5], [9]) using geometric duality can be obtained also by means of the approach introduced in [1] and [10]. This is a so-called conjugacy approach. We developed it by using the perturbation theory presented in [2].

The first part of the paper deals with the programming problem with a linear objective function, linear inequality and maximum entropy inequality constraints considered in [9]. We get, using the general scheme introduced in [1] and [10], a

*The first author was supported in part by Gottlieb Daimler- and Karl Benz-Stiftung (under 02-48/99). The second author was supported in part by Karl und Ruth Mayer-Stiftung.

†Faculty of Mathematics, Chemnitz University of Technology, D-09107 Chemnitz, Germany, e-mail: bot@mathematik.tu-chemnitz.de.

‡Faculty of Mathematics, Chemnitz University of Technology, D-09107 Chemnitz, Germany, e-mail: grad@mathematik.tu-chemnitz.de.

§Faculty of Mathematics, Chemnitz University of Technology, D-09107 Chemnitz, Germany, e-mail: wanka@mathematik.tu-chemnitz.de.

dual problem to it. This turns out to be the one obtained in the original paper, for which we bring a small, but necessary, correction. Let us note that the authors of [9] used geometric programming duality.

One of the works which deals with duality in geometric programming is the paper of Peterson ([7]). By using some geometric inequalities and the geometric Lagrangean, the author introduced dual problems to the constrained and unconstrained optimization problems.

We consider further the general case, the constrained and unconstrained primal geometric problems from [7]. Our scope is, by using some convenient perturbations, to rediscover in both cases exactly the dual problems introduced there.

This is another important step in the analysis of the relations that can be discovered between different types of duality, alongside papers like [1], or [10].

We can conclude that the perturbation approach introduced in [2] is on a large scale applicable in optimization. The Fenchel dual problem, the Lagrange dual problem (for both see [1], [10]) and the geometric dual problem can be obtained using suitable perturbations.

The structure of this paper is as follows. The second section presents the entropy constrained problem from [9], for which we rediscover and correct the dual. Section 3 is divided into three parts. The first one reminds the method of perturbations (cf. [1], [2], [10]), while the second presents the duality regarding the unconstrained primal geometric problem. The last subsection deals with the constrained primal geometric problem, whose dual is calculated also by using perturbations. On the base of the derived duality results we point out necessary and sufficient optimality conditions for all the considered programming problems.

2 The entropy constrained linear problem

Scott, Jefferson and Jorjani have considered in [9] the following optimization problem

$$(P_E) \qquad \inf c^T x,$$

subject to

$$Ax \geqq b,$$

$$-\sum_{i=1}^{n} x_i \ln x_i \geq H,$$

$$\sum_{i=1}^{n} x_i = 1, x \geqq 0,$$

where $x = (x_1, ..., x_n)^T \in \mathbb{R}^n, c \in \mathbb{R}^n, A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m$. We need to mention that in the maximum entropy optimization (see [5], [6]) it is considered that $0 \ln 0 = 0$, although the logarithm function is defined only on the set of strictly positive real numbers. This assumption is made due to the fact that $\lim_{x\to 0} x\ln x = 0$. We take this into account into our paper and it is also worthy to mention that we denote by "$\geqq$" the partial ordering introduced by the non-negative orthant in the corresponding space.

To find a dual problem to (P_E), the authors used in [9] the method of geometric duality introduced by Peterson in [7], obtaining the following dual to (P_E),

$$(D_{SJ}) \qquad \inf \left\{ -b^T u + \lambda \ln \sum_{i=1}^{n} \exp\left(-\frac{w_i}{\lambda}\right) - H\lambda \right\},$$

subject to

$$A^T u - c + w \leqq 0,$$
$$u \geqq 0, \lambda \geq 0, w \in \mathbb{R}^n.$$

Earlier, Erlander considered the same problem in [4], determining its dual problem by means of Lagrange duality.

Remark 2.1. From the theory of maximum entropy optimization (see [5], [6]) it is known that $-\sum_{i=1}^{n} x_i \ln x_i$, subject to $\sum_{i=1}^{n} x_i = 1, x \geqq 0$, attains its maximal value for $x_i = \frac{1}{n}, i = 1, ..., n$, where it is equal to $\ln n$. Also, its infimal value is 0. So, we can point out that if $H \leq 0$, the problem (P_E) becomes a linear constrained optimization problem, the constraint $-\sum_{i=1}^{n} x_i \ln x_i \geq H$ being redundant, as $-\sum_{i=1}^{n} x_i \ln x_i \geq 0 \geq H, \forall x \geqq 0$.

On the other hand, if $H > \ln n$, the feasible set of (P_E) is empty, so the problem has no solutions.

We can distinguish also the case $H = \ln n$, where it is obvious that only $x_i = \frac{1}{n}, i = 1, ..., n$, satisfy the entropy constraints, so, if x satisfies also the linear constraints (i.e. if it holds $\sum_{i=1}^{n} a_{ji} \geq nb_j, j = 1, ..., m$), the uniform probability distribution is the solution of the problem, otherwise there are no feasible solutions.

To avoid these trivial cases (see also [4]), we consider further that

$$0 < H < \ln n.$$

In this section our purpose is to show that the dual problem (D_{SJ}) can be obtained also by using a different approach. We shall treat the problem (P_E) by means of a conjugacy approach, using the procedure presented in [1] and [10].

To do so, let us rewrite our primal problem in the following way

$$(P_E) \qquad \inf_{\substack{g(x)\leqq 0,\\ x\in X}} f(x),$$

for

$$X = \{x \in \mathbb{R}^n : x \geqq 0\},$$
$$f : \mathbb{R}^n \to \mathbb{R}, \quad f(x) = c^T x,$$

and $g : \mathbb{R}^n \to \mathbb{R}^{m+3}$,

$$g(x) = \left(b - Ax, H + \sum_{i=1}^{n} x_i \ln x_i, 1 - \sum_{i=1}^{n} x_i, \sum_{i=1}^{n} x_i - 1 \right)^T.$$

Using the perturbation approach, in [1] and [10] there is obtained the following dual problem to (P_E), the so-called Fenchel-Lagrange dual,

$$(D_E) \qquad \sup_{\substack{p \in \mathbb{R}^n, \\ (q_1,...,q_m)^T \geqq 0, \\ q_{m+1}, q_{m+2}, q_{m+3} \geq 0}} \left\{ -f^*(p) + \inf_{x \geqq 0} \left[p^T x + q^T g(x) \right] \right\}.$$

For our particular problem, the dual variables are $p = (p_1, ..., p_n)^T \in \mathbb{R}^n$ and $q = (q_1, ..., q_m, q_{m+1}, q_{m+2}, q_{m+3})^T \in \mathbb{R}^{m+3}$. Let us calculate now the terms which appear above. First, it is well-known that the conjugate function of f is

$$f^*(p) = \begin{cases} 0, & \text{if } p = c, \\ +\infty, & \text{otherwise.} \end{cases}$$

So, we have to consider further $p^T x = c^T x = \sum\limits_{i=1}^{n} c_i x_i$. Let us calculate now the last term that appears in (D_E),

$$\begin{aligned} q^T g(x) &= \sum_{j=1}^{m} q_j \left(b_j - \sum_{i=1}^{n} a_{ji} x_i \right) + q_{m+1} \left(H + \sum_{i=1}^{n} x_i \ln x_i \right) \\ &+ q_{m+2} \left(1 - \sum_{i=1}^{n} x_i \right) + q_{m+3} \left(\sum_{i=1}^{n} x_i - 1 \right) \\ &= q_{m+1} \sum_{i=1}^{n} x_i \ln x_i - \sum_{j=1}^{m} \sum_{i=1}^{n} q_j a_{ji} x_i - (q_{m+2} - q_{m+3}) \sum_{i=1}^{n} x_i \\ &+ \sum_{j=1}^{m} q_j b_j + q_{m+1} H + (q_{m+2} - q_{m+3}) \\ &= \sum_{i=1}^{n} \left(q_{m+1} x_i \ln x_i - x_i \left(\sum_{j=1}^{m} q_j a_{ji} + q_{m+2} - q_{m+3} \right) \right) \\ &+ \sum_{j=1}^{m} q_j b_j + q_{m+1} H + (q_{m+2} - q_{m+3}). \end{aligned}$$

The dual problem becomes

$$(D_E) \qquad \sup_{\substack{(q_1,...,q_m)^T \geqq 0,\\ q_{m+1},q_{m+2},q_{m+3} \geq 0}} \left\{ \sum_{j=1}^{m} q_j b_j + q_{m+1} H + (q_{m+2} - q_{m+3}) \right.$$

$$\left. + \inf_{x \geqq 0} \sum_{i=1}^{n} \left(q_{m+1} x_i \ln x_i + x_i \left(c_i - \sum_{j=1}^{m} q_j a_{ji} - q_{m+2} + q_{m+3} \right) \right) \right\}.$$

We need to split now our calculation into two cases, whether q_{m+1} is zero or not.

First case: $q_{m+1} = 0$.
The form of the dual problem is simplified now to

$$(D_E^1) \qquad \sup_{\substack{(q_1,...,q_m)^T \geqq 0,\\ q_{m+2},q_{m+3} \geq 0}} \left\{ \sum_{j=1}^{m} q_j b_j + (q_{m+2} - q_{m+3}) \right.$$

$$\left. + \inf_{x \geqq 0} \sum_{i=1}^{n} x_i \left(c_i - \sum_{j=1}^{m} q_j a_{ji} - q_{m+2} + q_{m+3} \right) \right\}.$$

It is easy to notice that if there exists at least an $i \in \{1, ..., n\}$, such that

$$c_i - \sum_{j=1}^{m} q_j a_{ji} - q_{m+2} + q_{m+3} < 0,$$

the infimum regarding $x \geqq 0$ from (D_E^1) is $-\infty$, that is not desirable to our problem. Otherwise, it is equal to 0. So, considering further

$$c_i - \sum_{j=1}^{m} q_j a_{ji} - q_{m+2} + q_{m+3} \geq 0, i = 1, ..., n,$$

the dual problem becomes

$$(D_E^1) \qquad \sup_{\substack{(q_1,...,q_m)^T \geqq 0,\\ q_{m+2},q_{m+3} \geq 0,\\ c_i - \sum_{j=1}^{m} q_j a_{ji} - q_{m+2} + q_{m+3} \geq 0,\\ i=1,...,n,}} \left\{ \sum_{j=1}^{m} q_j b_j + q_{m+2} - q_{m+3} \right\}.$$

We can renounce the variables q_{m+2} and q_{m+3} and the dual problem turns into

$$(D_E^1) \qquad \sup_{(q_1,...,q_m)^T \geqq 0} \left\{ \sum_{j=1}^{m} q_j b_j + \min_{i=1,...,n} \left(c_i - \sum_{j=1}^{m} q_j a_{ji} \right) \right\}.$$

Second case: $q_{m+1} > 0$.
Let us consider the function $\alpha : \mathbb{R}_+ \to \mathbb{R}, \alpha(x) = vx \ln x + wx$, with $v > 0, w \in \mathbb{R}$. After some standard calculations there follows that α attains its minimum at $e^{-\frac{v+w}{v}}$, and it is

$$\alpha(e^{-\frac{v+w}{v}}) = -ve^{-\frac{v+w}{v}}.$$

Applying the result from above to our problem, for $v = q_{m+1}, x = x_i$ and $w = c_i - \sum\limits_{j=1}^{m} q_j a_{ji} - q_{m+2} + q_{m+3}, i = 1, ..., n$, we obtain

$$\inf_{x \geqq 0} \sum_{i=1}^{n} \left(q_{m+1} x_i \ln x_i + x_i \left(c_i - \sum_{j=1}^{m} q_j a_{ji} - q_{m+2} + q_{m+3} \right) \right) = -q_{m+1} \sum_{i=1}^{n} e^{-\frac{c_i - \sum\limits_{j=1}^{m} q_j a_{ji} - q_{m+2} + q_{m+3} + q_{m+1}}{q_{m+1}}}.$$

The dual problem becomes in this case

$$(D_E^2) \qquad \sup_{\substack{(q_1,...,q_m)^T \geqq 0, \\ q_{m+2}, q_{m+3} \geqq 0, \\ q_{m+1} > 0}} \left\{ -q_{m+1} \sum_{i=1}^{n} e^{-\frac{c_i - \sum\limits_{j=1}^{m} q_j a_{ji} - q_{m+2} + q_{m+3} + q_{m+1}}{q_{m+1}}} + \sum_{j=1}^{m} q_j b_j + q_{m+1} H + (q_{m+2} - q_{m+3}) \right\}.$$

Considering the function $\beta : \mathbb{R} \to \mathbb{R}, \beta(x) = -q_{m+1} w e^{\frac{x}{q_{m+1}}} + x + k$, with $w > 0$ and $k \in \mathbb{R}$, whose supremum is attained at $-q_{m+1} \ln w$, being

$$\beta(-q_{m+1} \ln w) = k - q_{m+1} \ln(ew),$$

for $x = q_{m+2} - q_{m+3}, w = \sum\limits_{i=1}^{n} e^{-\frac{c_i - \sum\limits_{j=1}^{m} q_j a_{ji} + q_{m+1}}{q_{m+1}}}$ and $k = \sum\limits_{j=1}^{m} q_j b_j + q_{m+1} H$, the dual problem has the following simplification

$$(D_E^2) \sup_{\substack{q_j \geq 0, \\ j=1,...,m, \\ q_{m+1} > 0}} \left\{ -q_{m+1} \ln \sum_{i=1}^{n} e^{-\frac{c_i - \sum\limits_{j=1}^{m} q_j a_{ji} + q_{m+1}}{q_{m+1}} + 1} + \sum_{j=1}^{m} q_j b_j + q_{m+1} H \right\},$$

which becomes, for $u = (q_1, ..., q_m)^T$, even

$$(D_E^2) \qquad \sup_{\substack{u \geqq 0, \\ q_{m+1} > 0}} \left\{ -q_{m+1} \ln \sum_{i=1}^{n} e^{-\frac{(c - A^T u)_i}{q_{m+1}}} + b^T u + q_{m+1} H \right\},$$

where $(c - A^T u)_i$ denotes the i-th component of the real-valued vector $c - A^T u$, $i = 1, ..., n$.

Finally, the dual problem derived by us to (P_E) is

$$(D_E) \qquad \max \left\{ \sup_{u \geqq 0} \left\{ b^T u + \min_{i=1,\dots,n} (c - A^T u)_i \right\}, \sup_{\substack{u \geqq 0, \\ q_{m+1} > 0}} \left\{ b^T u - q_{m+1} \ln \sum_{i=1}^{n} e^{-\frac{(c - A^T u)_i}{q_{m+1}}} + q_{m+1} H \right\} \right\}.$$

We can reduce this to a simpler form, using the following result (which is a special case of Lemma 4.1 in [3]).

Lemma 2.2. *For $w_i \in \mathbb{R}, i = 1, ..., n$, there holds*

$$\lim_{\lambda \downarrow 0} \lambda \left(\ln \sum_{i=1}^{n} e^{\frac{w_i}{\lambda}} \right) = \max_{i=1,\dots,n} w_i.$$

Using this, we have

$$\lim_{q_{m+1} \downarrow 0} q_{m+1} \left(H - \ln \sum_{i=1}^{n} e^{-\frac{(c - A^T u)_i}{q_{m+1}}} \right) = \min_{i=1,\dots,n} (c - A^T u)_i,$$

so the second branch in the formula of (D_E) is always greater than or equal to the first one. Using this fact, the dual problem can be simplified, denoting $\lambda = q_{m+1}$, to

$$(D_E) \qquad \sup_{\substack{u \geqq 0, \\ \lambda > 0}} \left\{ b^T u - \lambda \ln \sum_{i=1}^{n} e^{-\frac{(c - A^T u)_i}{\lambda}} + \lambda H \right\},$$

and, transformed into a minimization problem, it becomes

$$(D'_E) \qquad \inf_{\substack{u \geqq 0, \\ \lambda > 0}} \left\{ -b^T u + \lambda \ln \sum_{i=1}^{n} e^{-\frac{(c - A^T u)_i}{\lambda}} - \lambda H \right\}.$$

Remark 2.3.

(i) Here we point out an omission in [9] as $\lambda \geq 0$ appears as constraint in the dual problem determined there, but the case $\lambda = 0$ is just mentioned, not treated.

(ii) As e^{-x} is decreasing, the minimal value of $\lambda \ln \sum_{i=1}^{n} \exp\left(-\frac{w_i}{\lambda}\right)$, subject to $w \leqq c - A^T u, w \in \mathbb{R}^n$, is obtained when $w = c - A^T u$, so the variable w

may be eliminated from (D_{SJ}), that turns into a simpler version. Denoting $w = c - A^T u$, the only difference between (D'_E) and the simplified version of (D_{SJ}) resides in the fact that we have proved that the constraint $\lambda > 0$ is sufficient, while in (D_{SJ}) ([9]) λ is considered greater than or equal to 0. Anyway, our aim to prove that (D_{SJ}) may be obtained by using the perturbation approach has been fulfilled.

We are also interested in determining the conditions when the so-called strong duality between (P_E) and (D_E) occurs. First, let us mention that in [1] and [10] the authors prove that the weak duality assertion between the primal and the dual problem obtained using the perturbation approach always holds, i.e. the optimal objective value of the primal problem is always greater than or equal to the optimal objective value of the dual problem, so this takes place also in the present case. The theory we have applied here allows us to formulate also a strong duality result. We recall here that strong duality means, in the sense we consider here, that the dual problem has an optimal solution and the optimal objective values of the primal and dual coincide. To attain strong duality, a necessary condition is to have a constraint qualification fulfilled. Here, we use the following so-called Slater constraint qualification (cf. [3], [10]), for the vector function $g = (g_1, ..., g_{m+3})^T$, with convex component functions,

$$\exists x' \in \mathrm{ri}(X) : \begin{cases} g_j(x') < 0, & \text{if } g_j \text{ is not affine}, j \in \{1, ..., m+3\}, \\ g_j(x') \leq 0, & \text{if } g_j \text{ is affine}, j \in \{1, ..., m+3\}. \end{cases}$$

In our case, it becomes

$$\exists x' \in \mathrm{int}(\mathbb{R}^n_+) : \begin{cases} H + \sum\limits_{i=1}^{n} x'_i \ln x'_i < 0, \\ b - Ax' \leqq 0, \\ \sum\limits_{i=1}^{n} x'_i = 1. \end{cases} \tag{1}$$

Now we can state the strong duality result regarding our problem (cf. [10]).

Theorem 2.4. *If the constraint qualification* (1) *is satisfied, then the strong duality between* (P_E) *and* (D_E) *holds.*

Finally, using strong duality, we formulate and prove necessary and sufficient optimality conditions.

Theorem 2.5. *(a) If the constraint qualification* (1) *is fulfilled and* $\bar{x}$ *is an optimal solution to* (P_E)*, then the strong duality between* (P_E) *and* (D_E) *holds and the dual problem has a solution* $(\bar{u}, \bar{\lambda})$ *satisfying the following optimality conditions* *(i)* $\bar{u}^T(A\bar{x}-b)=0$,

(ii) $\bar{\lambda}\left(H+\sum\limits_{i=1}^{n}\bar{x}_i\ln\bar{x}_i\right)=0$,

(iii) $\bar{\lambda}\left(\sum\limits_{i=1}^{n}\bar{x}_i\ln\bar{x}_i+\ln\sum\limits_{i=1}^{n}e^{-\frac{(c-A^T\bar{u})_i}{\bar{\lambda}}}\right)=\bar{x}(A^T\bar{u}-c)$.

(b) Having a feasible solution $\bar{x}$ to the primal problem and one, $(\bar{u},\bar{\lambda})$, to the dual satisfying the optimality conditions (i)-(iii), the mentioned feasible solutions turn out to be optimal solutions to the corresponding problems and the strong duality holds.

Proof.

(a) Theorem 2.4 assures that the strong duality holds. So the dual problem has a solution Let it be denoted by $(\bar{u},\bar{\lambda})$. As the optimal values of the primal and dual problem coincide in this case and both of them have solutions, it holds

$$c^T\bar{x}-b^T\bar{u}+\bar{\lambda}\ln\sum_{i=1}^{n}e^{-\frac{(c-A^T\bar{u})_i}{\bar{\lambda}}}-\bar{\lambda}H=0.$$

Adding and subtracting some terms in the left-hand side of this equation, it follows

$$-\bar{\lambda}\left(H+\sum_{i=1}^{n}\bar{x}_i\ln\bar{x}_i\right)+\bar{\lambda}\left(\ln\sum_{i=1}^{n}e^{-\frac{(c-A^T\bar{u})_i}{\bar{\lambda}}}+\sum_{i=1}^{n}\bar{x}_i\ln\bar{x}_i\right)-\bar{u}^T(b-A\bar{x})$$
$$-\bar{u}^T(A\bar{x})+c^T\bar{x}=0,$$

which is equivalent to

$$-\bar{\lambda}\left(H+\sum_{i=1}^{n}\bar{x}_i\ln\bar{x}_i\right)+\bar{u}^T(A\bar{x}-b)+\bar{\lambda}\left(\ln\sum_{i=1}^{n}e^{-\frac{(c-A^T\bar{u})_i}{\bar{\lambda}}}+\sum_{i=1}^{n}\bar{x}_i\ln\bar{x}_i\right)$$
$$+\bar{x}^T(c-A^T\bar{u})=0. \tag{2}$$

From the previous calculations presented in this paper regarding the infimum of the function α, it holds, for every $s\in\mathbb{R}$,

$$\sum_{i=1}^{n}\left(\bar{\lambda}\bar{x}_i\ln\bar{x}_i+\bar{x}_i\big((c-A^T\bar{u})_i-s\big)\right)+s\geq-\bar{\lambda}\sum_{i=1}^{n}e^{-\frac{(c-A^T\bar{u})_i-s+\bar{\lambda}}{\bar{\lambda}}}+s.$$

Because $\sum\limits_{i=1}^{n}\bar{x}_i=1$, the terms containing s from the left-hand side of this expression can be simplified. We have further

$$\sum_{i=1}^{n}\left(\bar{\lambda}\bar{x}_i\ln\bar{x}_i+\bar{x}_i(c-A^T\bar{u})_i\right)\geq\sup_{s\in\mathbb{R}}\left\{s-\bar{\lambda}\sum_{i=1}^{n}e^{-\frac{(c-A^T\bar{u})_i-s+\bar{\lambda}}{\bar{\lambda}}}\right\}.$$

¿From the discussions regarding the function β (introduced earlier in this paper), it follows

$$\sum_{i=1}^{n}\left(\bar{\lambda}\bar{x}_i\ln\bar{x}_i+\bar{x}_i(c-A^T\bar{u})_i\right)\geq-\bar{\lambda}\ln\sum_{i=1}^{n}e^{-\frac{(c-A^T\bar{u})_i}{\bar{\lambda}}}.$$

This can be written as

$$\bar{\lambda}\left(\ln\sum_{i=1}^{n}e^{-\frac{(c-A^T\bar{u})_i}{\bar{\lambda}}}+\sum_{i=1}^{n}\bar{x}_i\ln\bar{x}_i\right)+\bar{x}^T(c-A^T\bar{u})\geq 0. \tag{3}$$

Using (3) and the fact that $\bar{x}$ and $(\bar{u},\bar{\lambda})$ are feasible to (P_E), respectively (D_E), it follows that the left-hand side of (2) is a sum of positive terms whose result is zero, so each of them has to be equal to 0, i.e. the optimality conditions (i)-(iii) are fulfilled.

(b) All the calculations presented above can be carried out in reverse order, so the assertion holds. □

3 The general geometric problem

After proving that the geometric dual of a particular problem can be obtained also via perturbations, we demonstrate further that this can be generalized to any geometric program. Peterson's classical work [7] presents a complete duality treatment for geometric programs. We show further that the duals he introduced there by using the geometric Langrangean and geometric inequalities can be obtained also by using the perturbation approach (see [1], [2], [10]). But first let us give a short description of this approach (cf. [1], [10]).

3.1 The perturbation method

Having an optimization problem

$$(P)\qquad \inf_{x\in\mathbb{R}^n} F(x),$$

with $F:\mathbb{R}^n\to\overline{\mathbb{R}}$, we attach to it a so-called perturbation function $\Phi:\mathbb{R}^n\times\mathbb{R}^m\to\overline{\mathbb{R}}$, with the property that

$$\Phi(x,0)=F(x)\ \ \forall x\in\mathbb{R}^n.$$

We call p the perturbation variable and $\mathbb{R}^m$ is the space of the perturbation variables. For each $p\in\mathbb{R}^m$ we obtain a so-called perturbed optimization problem

$$(P_p)\qquad \inf_{x\in\mathbb{R}^n}\Phi(x,p).$$

We need to consider further the conjugate function of Φ. Let us remind here its definition.

Definition 3.1. *The function*

$$F^*:\mathbb{R}^n \to \overline{\mathbb{R}}, \quad F^*(x^*) = \sup_{x\in\mathbb{R}^n}\{\langle x^*, x\rangle - F(x)\},$$

where $\langle x^*, x\rangle$ denotes the Euclidian scalar product in $\mathbb{R}^n$ between x^* and x, i.e. $\langle x^*, x\rangle = x^{*T}x$, is called the conjugate function of the function F.

So, the conjugate function of Φ, with $\big\langle (x^*, p^*), (x, p)\big\rangle$ the Euclidian scalar product between (x^*, p^*) and (x, p) in $\mathbb{R}^n \times \mathbb{R}^m$, reads as

$$\begin{aligned}\Phi^*(x^*, p^*) &= \sup_{\substack{x\in\mathbb{R}^n,\\ p\in\mathbb{R}^m}}\Big\{\big\langle (x^*, p^*), (x, p)\big\rangle - \Phi(x, p)\Big\} \\ &= \sup_{\substack{x\in\mathbb{R}^n,\\ p\in\mathbb{R}^m}}\Big\{\langle x^*, x\rangle + \langle p^*, p\rangle - \Phi(x, p)\Big\}.\end{aligned}$$

Now we can define the following dual problem to (P) (cf. [2]),

$$(D) \qquad \sup_{p^*\in\mathbb{R}^m}\Big\{ - \Phi^*(0, p^*)\Big\}.$$

Between the primal problem (P) and its dual introduced above, (D), there holds always the so-called weak duality (proved in [2]), i.e. any objective value of the primal problem is greater than or equal to any objective value of the dual problem.

To have strong duality between the primal problem (P) and its dual, (D), we need the problem (P) to be stable, i.e. the so-called infimum value function, $h(p) = \inf\limits_{x\in\mathbb{R}^n} \Phi(x, p), p \in \mathbb{R}^m$, must have a finite value at 0 and has to be subdifferentiable at the same point.

For a finite dimensional convex optimization problem, Rockafellar has presented in Theorem 31.1 in [8] some necessary conditions in order to achieve strong duality between the primal problem and the dual problem. The so-called constraint qualifications introduced by Peterson in [7] in order to attain strong duality base on the same theorem.

3.2 The unconstrained case

First we treat the general unconstrained geometric programming problem. Let be the function $g : C \to \mathbb{R}$, with the domain $C \subseteq \mathbb{R}^n$. There is given also a closed cone $X \subseteq \mathbb{R}^n$. The unconstrained geometric programming problem (here called primal problem) is

$$(A_u) \qquad \inf_{x\in S} g(x),$$

with the feasible set $S = C \cap X$.

In order to introduce the dual problem, we need to introduce the following definition.

Definition 3.2. *For the function $g : C \to \mathbb{R}$ we call its conjugate concerning its domain C the function $h : D \to \mathbb{R}$, where $D = \big\{y \in \mathbb{R}^n : \sup_{x\in C}\{\langle y, x\rangle - g(x)\} < \infty\big\}$, defined by*

$$h(y) = \sup_{x\in C}\{\langle y, x\rangle - g(x)\}.$$

Remark 3.3. We note that the formula of the conjugate concerning its domain C of the function g introduced above coincides with the classical conjugate of the function $\bar{g} : \mathbb{R}^n \to \overline{\mathbb{R}}$ with $\text{dom}(\bar{g}) = C$ and $\bar{g}(x) = g(x) \quad \forall x \in C$, i.e. $\bar{g}(x) = +\infty \quad \forall x \notin C$.

Considering this conjugate concerning the domain of the function g and also the dual cone Y to X (i.e. $Y = X^* = \{p^* \in \mathbb{R}^n : \langle p^*, x\rangle \geq 0 \quad \forall x \in X\}$), Peterson attached in [7] the following dual to the problem (A_u),

$$(B_u) \qquad \inf_{y\in J} h(y),$$

with the feasible set $J = Y \cap D$.

In order to treat the problem (A_u) by means of the conjugacy approach presented before, let us introduce the following extension to the objective function g of the primal problem,

$$F : \mathbb{R}^n \to \overline{\mathbb{R}}, \quad F(x) \quad = \quad \begin{cases} g(x), & \text{if } x \in S, \\ +\infty, & \text{otherwise.} \end{cases}$$

It is clear that the problem (A_u) can be written equivalently

$$(A'_u) \qquad \inf_{x\in\mathbb{R}^n} F(x).$$

Having this equivalent formulation, let us introduce the following perturbation function which plays a central role in our proof,

$$\Phi : \mathbb{R}^n \times \mathbb{R}^n \to \overline{\mathbb{R}}, \Phi(x, p) \quad = \quad \begin{cases} g(x + p), & \text{if } x \in X, x + p \in C, p \in \mathbb{R}^n, \\ +\infty, & \text{otherwise.} \end{cases}$$

The following relation is fulfilled

$$\Phi(x, 0) = F(x) \quad \forall x \in \mathbb{R}^n.$$

So, according to the previous section, the dual problem to (A'_u), so also to (A_u) since they are equivalent, is

$$(D_u) \qquad \sup_{p^*\in\mathbb{R}^n} \{-\Phi^*(0, p^*)\},$$

where we need to calculate first the conjugate function of the perturbation function

$$\begin{aligned}\Phi^*(x^*,p^*) &= \sup_{\substack{x\in\mathbb{R}^n,\\ p\in\mathbb{R}^n}}\{\langle x^*,x\rangle+\langle p^*,p\rangle-\Phi(x,p)\},\\ &= \sup_{\substack{x\in X,\\ p\in\mathbb{R}^n,\\ x+p\in C}}\{\langle x^*,x\rangle+\langle p^*,p\rangle-g(x+p)\},\\ &= \sup_{\substack{x\in X,\\ t\in C}}\{\langle x^*,x\rangle+\langle p^*,t-x\rangle-g(t)\}.\end{aligned}$$

As we have to take $x^*=0$ in order to calculate the dual problem, it follows

$$\begin{aligned}\Phi^*(0,p^*) &= \sup_{\substack{x\in X,\\ t\in C}}\{\langle p^*,t\rangle-\langle p^*,x\rangle-g(t)\}\\ &= \sup_{x\in X}\{-\langle p^*,x\rangle\}+\sup_{t\in C}\{\langle p^*,t\rangle-g(t)\}.\end{aligned}$$

From the definition of the dual cone Y to the cone X it follows

$$\sup_{x\in X}\{-\langle p^*,x\rangle\} = \begin{cases} 0, & \text{if } p^*\in Y,\\ +\infty, & \text{otherwise.}\end{cases}$$

Hence,

$$\sup_{p^*\in\mathbb{R}^n}\{-\Phi^*(0,p^*)\}=\sup_{p^*\in Y}\left\{-\sup_{t\in C}\{\langle p^*,t\rangle-g(t)\}\right\}.$$

Using the definition of the conjugate of the function g concerning its domain C, we have

$$\sup_{p^*\in\mathbb{R}^n}\{-\Phi^*(0,p^*)\}=\sup_{p^*\in Y\cap D}\{-h(p^*)\}=-\inf_{p^*\in J}h(p^*).$$

The dual problem we have obtained is

$$(D_u)\qquad \sup_{p^*\in J}\{-h(p^*)\},$$

and, considering a minimum formulation for it, we obtain the following dual problem to (A_u),

$$(D'_u)\qquad \inf_{p^*\in J}h(p^*),$$

which is exactly the one introduced in [7], (B_u).

As mentioned before, the weak duality regarding the problems (A_u) and (D_u) always holds, while for the strong duality we have the following theorem, obtained from Theorem 31.1 in [8].

Theorem 3.4. *If C is a convex set, g a convex function defined on C, X a convex closed cone and the condition* $\mathrm{ri}(C) \cap \mathrm{ri}(X) \neq \emptyset$ *is fulfilled, then the strong duality between (A_u) and (D_u) holds, i.e. (D_u) has an optimal solution and the optimal objective values of the primal and dual problem coincide.*

Remark 3.5. In [7] the conditions regarding the strong duality are posed on the dual problem, in which case g and X have to be, moreover, closed and the dual problem's infimum must be finite.

Let us also present necessary and sufficient optimality conditions regarding the unconstrained geometric program.

Theorem 3.6. (*a*) *Assume the hypotheses of Theorem 3.4 fulfilled and let $\bar{x}$ be an optimal solution to (A_u). Then the strong duality between the primal problem and its dual holds and there exists an optimal solution $\bar{p}^*$ to (D_u) satisfying the following optimality conditions*

(i) $\bar{p}^* \in \partial g(\bar{x})$,

(ii) $\langle \bar{p}^*, \bar{x} \rangle = 0$.

(b) Let $\bar{x}$ be a feasible solution to (A_u) and $\bar{p}^*$ one to (D_u) satisfying the optimality conditions (i) and (ii). Then $\bar{x}$ turns out to be an optimal solution to the primal problem, $\bar{p}^*$ one to the dual and the strong duality between the two problems holds.

Proof.

(a) From Theorem 3.4 we know that the strong duality holds and the dual problem has an optimal solution Let it be $\bar{p}^* \in J = Y \cap D$. Therefore, it holds

$$g(\bar{x}) + h(\bar{p}^*) = 0.$$

¿From Young's inequality it is known that

$$g(\bar{x}) + h(\bar{p}^*) \geq \langle \bar{p}^*, \bar{x} \rangle,$$

while

$$\langle \bar{p}^*, \bar{x} \rangle \geq 0$$

since $\bar{p}^* \in Y$ and $x \in X$. Hence it holds

$$g(\bar{x}) + h(\bar{p}^*) \geq \langle \bar{p}^*, \bar{x} \rangle \geq 0,$$

but, since we have equality between the first and the last member of the expression above, both inequalities must be fulfilled as equalities. So the equality must hold in the previous two expressions, i.e. the optimality conditions are true, since $g(\bar{x}) + h(\bar{p}^*) = \langle \bar{p}^*, \bar{x} \rangle$ implies $\bar{p}^* \in \partial g(\bar{x})$.

(b) The optimality conditions imply

$$g(\bar{x}) + h(\bar{p}^*) = \langle \bar{p}^*, \bar{x} \rangle = 0.$$

So the assertion holds. □

3.3 The constrained case

In this case, the primal problem becomes more complicated, as some constraints appear and also the objective function is not so simple anymore. The following preliminaries are required.

Let there be the finite index sets I and J. For $t \in \{0\} \cup I \cup J$, the following functions are considered

$$g_t : C_t \to \mathbb{R},$$

with the domains $C_t \subseteq \mathbb{R}^{n_t}$, as well as the independent vector variables $x^t \in \mathbb{R}^{n_t}$. There are also the sets

$$D_t = \left\{ y^t \in \mathbb{R}^{n_t} : \sup_{x^t \in C_t} \{\langle y^t, x^t \rangle - g_t(x^t)\} < +\infty \right\},$$

which are the domains of the conjugate functions regarding the domains C_t of the functions $g_t, t \in \{0\} \cup I \cup J$, respectively, and an independent vector variable $k = (k_1, ..., k_{|J|})^T$. With x^I one denotes the Cartesian product of the vector variables $x^i, i \in I$, while x^J denotes the same thing for $x^j, j \in J$. Hence, $x = (x^0, x^I, x^J)$ is an independent vector variable in $\mathbb{R}^n$, where $n = n_0 + \sum\limits_{i \in I} n_i + \sum\limits_{j \in J} n_j$. Finally, let there be a closed cone $X \subseteq \mathbb{R}^n$, the sets

$$\begin{aligned} C_j^+ = \big\{(x^j, k_j) \quad : \quad & \text{either } k_j = 0 \text{ and } \sup_{d^j \in D_j} \langle d^j, x^j \rangle < \infty \\ & \text{or } k_j > 0 \text{ and } x^j \in k_j C_j \big\}, \quad j \in J, \end{aligned}$$

$$C = \big\{(x, k) : x^t \in C_t, t \in \{0\} \cup I, (x^j, k_j) \in C_j^+, j \in J \big\},$$

and, for $j \in J$, the functions

$$g_j^+(x^j, k_j) \quad = \quad \begin{cases} \sup\limits_{d^j \in D_j} \langle d^j, x^j \rangle, & \text{if } k_j = 0 \text{ and } \sup\limits_{d^j \in D_j} \langle d^j, x^j \rangle < \infty, \\ k_j g_j\left(\frac{x^j}{k_j}\right), & \text{if } k_j > 0 \text{ and } x^j \in k_j C_j. \end{cases}$$

Peterson ([7]) considers the following objective function

$$g : C \to \mathbb{R}, \quad g(x, k) = g_0(x^0) + \sum_{j \in J} g_j^+(x^j, k_j)$$

such that the primal geometric programming problem is

$$(A_c) \qquad \inf_{(x,k) \in S} g(x, k),$$

with the feasible set

$$S = \{(x,k) \in C : x \in X, g_i(x^i) \leq 0, i \in I\}.$$

To introduce a dual problem to it, one has to introduce the dual cone Y to the cone X, the sets

$$D = \big\{(y^0, y^I, y^J, \lambda) : y^t \in D_t, t \in \{0\} \cup J, (y^i, \lambda_i) \in D_i^+, i \in I\big\},$$

$$\begin{aligned} D_i^+ &= \big\{(y^i, \lambda_i) : \text{ either } \lambda_i = 0 \text{ and } \sup_{c^i \in C_i} \langle y^i, c^i \rangle < \infty, \\ &\qquad \text{or } \lambda_i > 0 \text{ and } y^i \in \lambda_i D_i\big\}, \quad i \in I, \end{aligned}$$

and some functions, namely $h_t : D_t \to \mathbb{R}$, the conjugate of the function g_t concerning its domain $C_t, t \in \{0\} \cup I \cup J$, and for $i \in I$ we have also the functions

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

In [7] there is introduced the following dual problem to (A_c),

$$(B_c) \qquad \inf_{(y,\lambda) \in T} h(y, \lambda),$$

with the feasible set

$$T = \{(y, \lambda) \in D : y \in Y, h_j(y^j) \leq 0, j \in J\},$$

where the objective function is

$$h : D \to \mathbb{R}, \quad h(y, \lambda) = h_0(y^0) + \sum_{i \in I} h_i^+(y^i, \lambda_i).$$

In the following part we demonstrate that this dual problem can be developed also by using our method based on perturbations. Like before, we introduce the following extension of the objective function

$$F : \mathbb{R}^n \times \mathbb{R}^{|J|} \to \overline{\mathbb{R}},$$

$$F(x,k) = \begin{cases} g(x,k), & \text{if } (x,k) \in C, x \in X, g_i(x^i) \leq 0, i \in I, \\ +\infty, & \text{otherwise.} \end{cases}$$

So, we can write the problem (A_c) equivalently as

$$(A'_c) \inf_{(x,k) \in \mathbb{R}^n \times \mathbb{R}^{|J|}} F(x,k).$$

Let us introduce now, as in section 3.1, the perturbation function associated to our problem,

$$\Phi : \mathbb{R}^n \times \mathbb{R}^{|J|} \times \mathbb{R}^n \times \mathbb{R}^{|I|} \to \overline{\mathbb{R}},$$

$$\Phi(x,k,p,v) \quad = \quad \begin{cases} g(x+p,k), & \text{if } x \in X, (x+p,k) \in C, \text{ and} \\ & \quad g_i(x^i+p^i) \leq v^i, i \in I, \\ +\infty, & \text{otherwise.} \end{cases}$$

It is obvious that $\Phi(x,k,0,0) = F(x,k) \;\; \forall (x,k) \in \mathbb{R}^n \times \mathbb{R}^{|J|}$, so from section 3.1, the dual problem to (A_c'), so also to (A_c), is

$$(D_c) \qquad \sup_{\substack{p^* \in \mathbb{R}^n, \\ v^* \in \mathbb{R}^{|I|}}} \big\{ -\Phi^*(0,0,p^*,v^*) \big\},$$

where we have

$$\Phi^*(x^*,k^*,p^*,v^*) = \sup_{\substack{x,p \in \mathbb{R}^n, \\ k \in \mathbb{R}^{|J|}, \\ v \in \mathbb{R}^{|I|}}} \Big\{ \big\langle (x^*,k^*,p^*,v^*),(x,k,p,v) \big\rangle - \Phi(x,k,p,v) \Big\}$$

$$= \sup_{\substack{x \in X, k \in \mathbb{R}^{|J|}_+, \\ p \in \mathbb{R}^n, v \in \mathbb{R}^{|I|}, \\ g_i(x^i+p^i) \leq v^i, i \in I, \\ (x+p,k) \in C}} \Big\{ \quad \langle x^*,x \rangle + \langle k^*,k \rangle + \langle p^*,p \rangle + \langle v^*,v \rangle$$

$$-g_0(x^0+p^0) - \sum_{j \in J} g_j^+(x^j+p^j,k_j) \Big\},$$

with the dual variables $x^* = (x^{*0}, x^{*I}, x^{*J})$ and $p^* = (p^{*0}, p^{*I}, p^{*J})$. Introducing the new variables $z = x + p$ and $y = v - g_I(z^I)$, with $g_I(z^I) = (g_i(z^i))^T_{i \in I}$, there follows

$$\begin{aligned} \Phi^*(x^*,k^*,p^*,v^*) \quad &= \sup_{\substack{x \in X, k \in \mathbb{R}^{|J|}_+, \\ (z,k) \in C, y \in \mathbb{R}^{|I|}, \\ y^i \geq 0, i \in I}} \Big\{ \langle x^*,x \rangle + \langle k^*,k \rangle + \langle p^*, z-x \rangle \\ &+ \quad \langle v^*, y + g_I(z^I) \rangle - g_0(x^0) - \sum_{j \in J} g_j^+(z^j,k_j) \Big\} \\ &= \quad \sum_{i \in I} \sup_{y^i \geq 0} \langle v^{*i}, y^i \rangle + \sup_{z^0 \in C_0} \{ \langle p^{*0}, z^0 \rangle - g_0(z^0) \} \\ &+ \quad \sum_{i \in I} \sup_{z^i \in C_i} \{ \langle p^{*i}, z^i \rangle + v^{*i} g_i(z^i) \} + \sup_{x \in X} \langle x^* - p^*, x \rangle \\ &+ \quad \sum_{j \in J} \sup_{(z^j,k_j) \in C_j^+} \Big\{ \langle p^{*j}, z^j \rangle + \langle k_j^*, k_j \rangle - g_j^+(z^j,k_j) \Big\}. \end{aligned}$$

In order to calculate the dual problem (D_c), we must consider further $x^* = 0$ and $k^* = 0$. Also, we use the following results that arise from definitions or simple calculations. The first of them,

$$\begin{aligned}\sup_{z^i \in C_i} \{\langle p^{*i}, z^i \rangle + v^{*i} g_i(z^i)\} &= \begin{cases} \sup\limits_{z^i \in C_i} \langle p^{*i}, z^i \rangle, & \text{if } v^{*i} = 0, \langle p^{*i}, z^i \rangle < \infty, \\ -v^{*i} h_i\left(\frac{p^{*i}}{-v^{*i}}\right), & \text{if } v^{*i} \neq 0, p^{*i} \in -v^{*i} D_i, \end{cases} \\ &= h_i^+(p^{*i}, -v^{*i}), i \in I,\end{aligned}$$

comes directly from the definitions of the conjugate function concerning its domain and of the functions $h_i^+, i \in I$. Then, the same definitions of the conjugate of a function concerning its domain give us the following result

$$\sup_{z^0 \in C_0} \{\langle p^{*0}, z^0 \rangle - g_0(z^0)\} = h_0(p^{*0}).$$

It is also clear that it holds

$$\sup_{y^i \geq 0} \langle v^{*i}, y^i \rangle = \begin{cases} 0, & \text{if } v^{*i} \leq 0, \\ +\infty, & \text{otherwise,} \end{cases}, i \in I,$$

while from the definition of the dual cone we have

$$\sup_{x \in X} \langle -p^*, x \rangle = \begin{cases} 0, & \text{if } p^* \in X^*, \\ +\infty, & \text{otherwise.} \end{cases}$$

Further we calculate the values of the terms summed after $j \in J$ in the last stage of the formula of $\Phi^*(0, 0, p^*, v^*)$, splitting the calculations into two branches. When $k_j > 0$ we have

$$\begin{aligned}\sup_{(z^j, k_j) \in C_j^+} \{\langle p^{*j}, z^j \rangle - g_j^+(z^j, k_j)\} &= \sup_{(z^j, k_j) \in C_j^+} \left\{\langle p^{*j}, z^j \rangle - k_j g_j\left(\frac{z^j}{k_j}\right)\right\}, \\ &= \sup_{k_j > 0} k_j h_j(p^{*j}) \\ &= \begin{cases} 0, & \text{if } h_j(p^{*j}) \leq 0, p^{*j} \in D_j, \\ +\infty, & \text{otherwise.} \end{cases}\end{aligned}$$

When $h_j(p^{*j}) \leq 0$ and $p^{*j} \in D_j$, the case $k_j = 0$ guides us to

$$\sup_{(z^j, k_j) \in C_j^+} \{\langle p^{*j}, z^j \rangle - g_j^+(z^j, k_j)\} = \sup_{(z^j, 0) \in C_j^+} \{\langle p^{*j}, z^j \rangle - \sup_{d^j \in D_j} \langle d^j, z^j \rangle\} = 0,$$

so we can conclude that for every $j \in J$ it holds

$$\sup_{(z^j, k_j) \in C_j^+} \{\langle p^*, z \rangle - g_j^+(z^j, k_j)\} = \begin{cases} 0, & \text{if } h_j(p^{*j}) \leq 0 \text{ and } p^{*j} \in D_j, \\ +\infty, & \text{otherwise.} \end{cases}$$

The dual problem can be simplified, denoting $\lambda = -v^*$, to

$$(D_c) \qquad \sup_{\substack{p^{t*} \in D_t, t \in \{0\} \cup J, \\ (p^{*i}, \lambda^i) \in D_i^+, i \in I, \\ h_j(p^{*j}) \leq 0, j \in J, \\ p^* \in Y}} \left\{ -h_0(p^{*0}) - \sum_{i \in I} h_i^+(p_i^*, \lambda_i) \right\},$$

which, transformed into a minimization problem becomes, using the notations in [7],

$$(D_c') \qquad \inf_{(p^*, \lambda) \in T} \left\{ h_0(p^{*0}) + \sum_{i \in I} h_i^+(p^{*i}, \lambda_i) \right\},$$

which is exactly the dual introduced by Peterson, (B_c).

The results from section 3.1 assure that the weak duality regarding the problems (A_c) and (D_c) always holds, while for strong duality we need to introduce some supplementary conditions.

First, let us consider that the sets C_t, $t \in \{0\} \cup I \cup J$ and the functions g_t, $t \in \{0\} \cup I \cup J$ are convex. The cone X needs to be closed and convex, too. We have to consider also that the sets C_j and the functions g_j, $j \in J$, are closed. This last property, alongside the convexity, assures (cf. [8]) that, for each $j \in J$, the functions g_j and h_j are a pair of conjugate closed convex functions, i.e. each of them is the other's conjugate concerning its domain. This fact allows us to characterize the sets C_j in the following way

$$C_j = \left\{ x^j \in \mathbb{R}^{n_j} : \sup_{d^j \in D_j} \{\langle d^j, c^j \rangle - h_j(d^j)\} < +\infty \right\}, \quad j \in J.$$

Using this characterization, it follows that the functions g_j^+ and the sets C_j^+ are convex, for all $j \in J$.

Then, let us introduce the following constraint qualification

$$\exists (x', k') \in \mathrm{ri}(X) \times \mathrm{int}(\mathbb{R}_+^{|J|}) : \begin{cases} x'^0 \in \mathrm{ri}(C_0), \\ x'^i \in \mathrm{ri}(C_i), \\ g_i(x'^i) < 0, i \in I, \\ x'^j \in k_j' \, \mathrm{ri}(C_j), j \in J. \end{cases} \tag{4}$$

We are now ready to formulate the strong duality theorem.

Theorem 3.7. *If the conditions introduced above regarding the functions g_t, $t \in \{0\} \cup I \cup J$, the sets C_t, $t \in \{0\} \cup I \cup J$, and the cone X are fulfilled and the constraint qualification* (4) *holds, then we have strong duality between (A_c) and (D_c).*

Remark 3.8.

(i) In [7] the constraint qualification regarding the strong duality is posed on the dual problem, while we choose to consider it on the primal problem.

(ii) Let us note that the closeness property of g_j and C_j, $j \in J$, is necessary in order to prove that g_j^+ and C_j^+ are convex, $\forall j \in J$, as Fenchel's duality theorem (Theorem 31.1 in [8]) requires the existence of convexity for all the functions and sets involved in the primal problem.

On the base of this strong duality, we can conclude necessary and sufficient optimality conditions for the geometric programming problem (A_c).

Theorem 3.9. *(a) Assume the hypotheses of Theorem 3.7 fulfilled and let* $(\bar{x}^0, \bar{x}^I, \bar{x}^J, \bar{k})$ *be an optimal solution to* (A_c)*. Then the strong duality between the primal problem and its dual holds and there exists an optimal solution* $(\bar{p}^{*0}, \bar{p}^{*I}, \bar{p}^{*J}, \bar{\lambda})$ *to* (D_c) *satisfying the following optimality conditions*

$$(i) \quad \bar{p}^{*0} \in \partial g_0(\bar{x}^0),$$

$$(ii) \quad \begin{cases} \bar{p}^{*j} \in \partial g_j\big(\frac{\bar{x}^j}{\bar{k}_j}\big) \text{ and } h_j(\bar{p}^{*j}) = 0, & \text{if } \bar{k}_j \neq 0, \\ \sup\limits_{d^j \in D_j} \langle d^j, \bar{x}^j \rangle = \langle \bar{p}^{*j}, \bar{x}^j \rangle, & \text{if } \bar{k}_j = 0, \end{cases} \quad j \in J,$$

$$(iii) \quad \begin{cases} \bar{x}^i \in \partial h_i\big(\frac{\bar{p}^{*i}}{\bar{\lambda}_i}\big) \text{ and } g_i(\bar{x}^i) = 0, & \text{if } \bar{\lambda}_i \neq 0, \\ \sup\limits_{c^i \in C_i} \langle \bar{p}^{*i}, c^i \rangle = \langle \bar{p}^{*i}, \bar{x}^i \rangle, & \text{if } \bar{\lambda}_i = 0, \end{cases} \quad i \in I,$$

$$(iv) \quad \langle \bar{p}^*, \bar{x} \rangle = 0.$$

(b) Let $(\bar{x}^0, \bar{x}^I, \bar{x}^J, \bar{k})$ *be a feasible solution to* (A_c) *and* $(\bar{p}^{*0}, \bar{p}^{*I}, \bar{p}^{*J}, \bar{\lambda})$ *one to* (D_c) *satisfying the optimality conditions (i)-(iv). Then* $(\bar{x}^0, \bar{x}^I, \bar{x}^J, \bar{k})$ *turns out to be an optimal solution to the primal problem,* $(\bar{p}^{*0}, \bar{p}^{*I}, \bar{p}^{*J}, \bar{\lambda})$ *one to the dual and the strong duality between the two problems holds.*

Proof.

(a) From Theorem 3.7 we know that the strong duality holds and the dual problem has an optimal solution. Let it be $(\bar{p}^{*0}, \bar{p}^{*I}, \bar{p}^{*J}, \bar{\lambda})$. Therefore, it holds

$$g_0(\bar{x}^0) + \sum_{j \in J} g_j^+(\bar{x}^j, \bar{k}_j) + h_0(\bar{p}^{*0}) + \sum_{i \in I} h_i^+(\bar{p}^{*i}, \bar{\lambda}_i) = 0,$$

rewritable as

$$g_0(\bar{x}^0) + h_0(\bar{p}^{*0}) + \sum_{\substack{i \in I, \\ \bar{\lambda}_i \neq 0}} \bar{\lambda}_i h_i\left(\frac{\bar{p}^{*i}}{\bar{\lambda}_i}\right) + \sum_{\substack{i \in I, \\ \bar{\lambda}_i = 0}} \sup_{c^i \in C_i} \langle \bar{p}^{*i}, c^i \rangle$$

$$+ \sum_{\substack{j \in J, \\ \bar{k}_j \neq 0}} \bar{k}_j g_j\left(\frac{\bar{x}^j}{\bar{k}_j}\right) + \sum_{\substack{j \in J, \\ \bar{k}_j = 0}} \sup_{d^j \in D_j} \langle d^j, \bar{x}^j \rangle = 0.$$

Adding and subtracting some terms in the left-hand side, we obtain

$$
\begin{aligned}
&[g_0(\bar{x}^0) + h_0(\bar{p}^{*0}) - \langle \bar{p}^{*0}, \bar{x}^0 \rangle] + \sum_{\substack{i \in I, \\ \bar{\lambda}_i \neq 0}} \left[\bar{\lambda}_i h_i \left(\frac{\bar{p}^{*i}}{\bar{\lambda}_i} \right) + \bar{\lambda}_i g_i(\bar{x}^i) - \langle \bar{p}^{*i}, \bar{x}^i \rangle \right] \\
&+ \sum_{\substack{i \in I, \\ \bar{\lambda}_i = 0}} \left[\sup_{c^i \in C_i} \langle \bar{p}^{*i}, c^i \rangle - \langle \bar{p}^{*i}, \bar{x}^i \rangle \right] + \sum_{\substack{j \in J, \\ \bar{k}_j \neq 0}} \left[\bar{k}_j g_j \left(\frac{\bar{x}^j}{\bar{k}_j} \right) + \bar{k}_j h_j(\bar{p}^{*j}) \right. \\
&\left. - \langle \bar{p}^{*j}, \bar{x}^j \rangle \right] + \sum_{\substack{j \in J, \\ \bar{k}_j = 0}} \left[\sup_{d^j \in D_j} \langle d^j, \bar{x}^j \rangle - \langle \bar{p}^{*j}, \bar{x}^j \rangle \right] \\
&+ \langle (\bar{p}^{*0}, \bar{p}^{*I}, \bar{p}^{*J}), (\bar{x}^0, \bar{x}^I, \bar{x}^J) \rangle - \sum_{\substack{i \in I, \\ \bar{\lambda}_i \neq 0}} \bar{\lambda}_i g_i(\bar{x}^i) - \sum_{\substack{j \in J, \\ \bar{k}_j \neq 0}} \bar{k}_j h_j(\bar{p}^{*j}) = 0. \qquad (5)
\end{aligned}
$$

Let us prove now that all the terms summed in the left-hand side in (5) are positive.

Applying Young's inequality, we get

$$g_0(\bar{x}^0) + h_0(\bar{p}^{*0}) \geq \langle \bar{p}^{*0}, \bar{x}^0 \rangle,$$

$$\bar{k}_j g_j \left(\frac{\bar{x}^j}{\bar{k}_j} \right) + \bar{k}_j h_j(\bar{p}^{*j}) \geq \langle \bar{p}^{*j}, \bar{k}_j \frac{\bar{x}^j}{\bar{k}_j} \rangle = \langle \bar{p}^{*j}, \bar{x}^j \rangle, j \in J : \bar{k}_j \neq 0,$$

and

$$\bar{\lambda}_i g_i(\bar{x}^i) + \bar{\lambda}_i h_i \left(\frac{\bar{p}^{*i}}{\bar{\lambda}_i} \right) = \bar{\lambda}_i \langle \frac{\bar{p}^{*i}}{\bar{\lambda}_i}, \bar{x}^i \rangle = \langle \bar{p}^{*i}, \bar{x}^i \rangle, i \in I : \bar{\lambda}_i \neq 0.$$

On the other hand, it is obvious that

$$\sup_{d^j \in D_j} \langle d^j, \bar{x}^j \rangle \geq \langle \bar{p}^{*j}, \bar{x}^j \rangle, j \in J : \bar{k}_j = 0,$$

and

$$\sup_{c^i \in C_i} \langle \bar{p}^{*i}, c^i \rangle \geq \langle \bar{p}^{*i}, \bar{x}^i \rangle, i \in I : \bar{\lambda}_i = 0.$$

Since $\bar{p}^* \in Y = X^*$, it follows also that $\langle \bar{p}^*, \bar{x} \rangle \geq 0$. Moreover, from the feasibility conditions it follows that $g_i(\bar{x}^i) \leq 0, \bar{\lambda}_i \geq 0, i \in I$, so $-\sum_{\substack{i \in I, \\ \bar{\lambda}_i \neq 0}} \bar{\lambda}_i g_i(\bar{x}^i) \geq 0$. Also, $h_j(\bar{p}^{*j}) \leq 0, \bar{k}_j \neq 0, j \in J$, implies $-\sum_{\substack{j \in J, \\ \bar{k}_j \neq 0}} \bar{k}_j h_j(\bar{p}^{*j}) \geq 0$. Therefore follows that the left-hand side in (5) is a sum of greater than or equal to zero terms whose result is zero, so all the terms must be equal to zero, i.e. the inequalities obtained above are fulfilled as equalities. So, (iv) is true, and, by the definition of the subdifferential, the other optimality conditions, (i)-(iii), hold, too.

(b) The calculations above can be carried out in reverse order and the assertion arises easily. □

Remark 3.10. We mention that (b) applies without any convexity assumptions as well as constraint qualifications. So, the sufficiency of the optimality conditions (i)-(iv) is true in the most general case.

References

[1] R.I. Boţ, G. Kassay, G. Wanka, Strong duality for generalized convex optimization problems, submitted for publication (2002).

[2] I. Ekeland, R. Temam, Convex analysis and variational problems, North-Holland, Amsterdam, 1976.

[3] K.-H. Elster, R. Reinhardt, M. Schäuble, G. Donath, Einführung in die nichtlineare Optimierung, Mathematisch-Naturwissenschaftliche Bibliothek, BSB B.G. Teubner Verlagsgesellschaft, Leipzig, 1977.

[4] S. Erlander, Entropy in Linear Programs, Mathematical Programming 21 (1981) 137–151.

[5] S.C. Fang, J.R. Rajasekera, H.-S.J. Tsao, Entropy Optimization and Mathematical Programming, Kluwer Academic Publishers, Boston, 1997.

[6] S. Guiaşu, A. Shenitzer, The Principle of Maximum Entropy, The Mathematical Intelligencer 7 (1985) 42–48.

[7] E.L. Peterson, Geometric programming, SIAM Review 18 (1976) 1–51.

[8] R.T. Rockafellar, Convex Analysis, Princeton University Press, Princeton, 1970.

[9] C.H. Scott, T.R. Jefferson, S. Jorjani, On Duality for Entropy Constrained Programs, Information Sciences 35 (1985) 211–221.

[10] G. Wanka, R.I. Boţ, On the Relations Between Different Dual Problems in Convex Mathematical Programming, in: P. Chamoni, R. Leisten, A. Martin, J. Minnemann, H. Stadtler (Eds.), Operations Research Proceedings 2001, Springer-Verlag, 2002, pp. 255–262.