Proving strong duality for geometric optimization using a conic formulation

François Glineur*
Francois.Glineur@fpms.ac.be
Service de Mathématique et de Recherche Opérationnelle,
Faculté Polytechnique de Mons,
Rue de Houdain, 9, B-7000 Mons, Belgium.

Abstract

Geometric optimization[1] is an important class of problems that has many applications, especially in engineering design. In this article, we provide new simplified proofs for the well-known associated duality theory, using conic optimization. After introducing suitable convex cones and studying their properties, we model geometric optimization problems with a conic formulation, which allows us to apply the powerful duality theory of conic optimization and derive the duality results valid for geometric optimization.

Keywords: geometric optimization, duality theory, conic optimization.

1 Introduction

Geometric optimization forms an important class of problems that enables practitioners to model a large variety of real-world applications, mostly in the field of engineering design.

Although not convex itself, a geometric optimization problem can be easily transformed into a convex problem, for which a Lagrangean dual can be explicitly written. Several duality results are known for this pair of problems, some being mere consequences of convexity (e.g. weak duality), others being specific to this particular class of problems (e.g. the absence of a duality gap).

These properties were first studied in the sixties, and can be found for example in the reference book of Duffin, Peterson and Zener [2]. The aim of this paper is to derive these results using the machinery of duality for conic optimization, which has in our opinion the advantage of simplifying and clarifying the proofs.

In order to use this setting, we start by defining an appropriate convex cone that allows us to express geometric optimization problems as conic programs. The first step we take

*The author is a research fellow supported by a grant from the F.N.R.S. (Belgian National Fund for Scientific Research). This research was partly carried out during a visit to the SSOR department of the Delft University of Technology supported by a grant from the Communauté Française de Belgique.

[1]This class of problems is usually known as *geometric programming*, for historical reasons. However, because of the strong connection of the term "programming" with computer science, we prefer to use the more natural word "optimization".

consists in studying some properties of this cone (e.g. closedness) and determine its dual. We are then in position to apply the general duality theory for conic optimization [9, 10] to our problems and find in a rather seamless way the various well-known duality theorems of geometric optimization.

This paper is organized as follows: Section 2 introduces conic optimization and outlines the associated duality theory, while we define and study in Section 3 the convex cones needed to model geometric optimization. Section 4 constitutes the main part of this article and presents new proofs of several duality theorems based on conic duality. Finally, we provide in Section 5 some hints on how to establish the link between our results and the classical theorems found in the literature, as well as some concluding remarks.

2 Conic optimization

In this section, we describe conic optimization and the associated duality theory. The results are well-known and we will skip most of the proofs (they can be found for example in the Ph.D. thesis of Sturm [10], which has been recently published in [11]). Conic optimization deals with a class of problems that is essentially equivalent to the class of convex problems, i.e. minimization of a convex function over a convex set. However, formulating a convex problem in a conic way often gives a new insight about it, especially when dealing with duality.

2.1 Conic problems

The basic ingredient we need is a convex cone.

Definition 2.1 *A set K is a* cone *if and only if it is closed under nonnegative scalar multiplication, i.e.*

$$x \in K \Rightarrow \lambda x \in K \text{ for all } \lambda \in \mathbb{R}_+ \ .$$

Recall that a set is convex if and only if it contains the whole segment joining any two of its points. Establishing convexity is easier for cones than for general sets, because of the following elementary theorem (see e.g. [7]):

Theorem 2.1 *A cone K is* convex *if and only if it is closed under addition, i.e.*

$$x \in K \text{ and } y \in K \Rightarrow x + y \in K \ .$$

In order to avoid some technical nuisances, the convex cones we are going to consider will be required to be closed, pointed and solid, according to the following definitions:

Definition 2.2 *A cone K is* solid *if and only if* $\operatorname{int} K \neq \emptyset$.

Definition 2.3 *A cone K is* pointed *if and only if* $K \cap -K = \{0\}$.

These two properties basically mean that K is a full-dimensional cone that does not contain any straight line passing through the origin.

We are now in position to define a conic optimization problem: let $K \subseteq \mathbb{R}^n$ a pointed, solid, closed convex cone. The (primal) conic optimization problem is

$$\inf_x c^T x \quad \text{s.t.} \quad Ax = b \text{ and } x \in K \ , \tag{CP}$$

where $x \in \mathbb{R}^n$ is the column vector we are optimizing, A is a $m \times n$ matrix and b, c are column vectors belonging respectively to $\mathbb{R}^m$ and $\mathbb{R}^n$. This problem can be viewed as the minimization of a linear function over the intersection of a convex cone and an affine subspace. At this stage, we would like to emphasize the fact that although our cone K is closed, it may happen that the infimum in (CP) is not attained (some examples of this situation will be given in Subsection 4.4).

It is well-known that this class of problems is equivalent to the class of convex problems, see e.g. [6]. However, the usual Lagrangean dual of a conic problem can be expressed very nicely in a conic form, using the notion of dual cone.

2.2 Dual conic problems

Definition 2.4 *The* dual *of a cone $K \subseteq \mathbb{R}^n$ is defined by*

$$K^* = \left\{ x^* \in \mathbb{R}^n \mid x^T x^* \geq 0 \text{ for all } x \in K \right\} .$$

The dual of a cone is always a closed convex cone. However, we have the stronger theorem

Theorem 2.2 *If K is a solid, pointed, closed convex cone, its dual K^* is another solid, pointed, closed convex cone. Moreover, the dual of K^* is equal to K.*

Indeed, the notions of solidness and pointedness are dual to each other [10, Corollary 2.1], while closedness is essential for $(K^*)^* = K$ to hold (we have in fact $(K^*)^* = \operatorname{cl} K$ without the closedness assumption on K [7, Theorem 14.1]).

The dual of our primal conic problem (CP) is defined by

$$\sup_{(y,s)} b^T y \quad \text{s.t.} \quad A^T y + s = c \text{ and } s \in K^* , \tag{CD}$$

where $y \in \mathbb{R}^m$ and $s \in \mathbb{R}^n$ are the column vectors we are optimizing, the other quantities A, b and c being the same as in (CP). It is immediate to notice that this dual problem has the same kind of structure as the primal problem, i.e. it also involves optimizing a linear function over the intersection of a convex cone and an affine subspace. The only differences are the direction of the optimization (maximization instead of minimization) and the way the affine subspace is described. It is also possible to show that the dual of this dual problem is equivalent to the primal problem, using the fact that $(K^*)^* = K$.

2.3 Duality theory

The two conic problems of this primal-dual pair are strongly related to each other, as demonstrated by the following theorems.

Theorem 2.3 (Weak duality) *Let x a feasible (i.e. satisfying the constraints) solution for* (CP)*, and (y, s) a feasible solution for* (CD)*. We have*

$$b^T y \leq c^T x ,$$

equality occurring if and only if the following orthogonality condition *is satisfied:*

$$x^T s = 0 .$$

This theorem shows that any primal (resp. dual) feasible solution provides an upper (resp. lower) bound for the dual (resp. primal) problem. Its proof is quite easy to obtain: elementary manipulations give

$$c^T x - b^T y = x^T c - (Ax)^T y = x^T (A^T y + s) - x^T A^T y = x^T s \ ,$$

this last inner product being always nonnegative because of $x \in K$, $s \in K^*$ and Definition 2.4 of the dual cone K^*.

Denoting by p^* and d^* the optimum objective values of problems (CP) and (CD), this theorem implies that $p^* - d^* \geq 0$. This nonnegative quantity is called the *duality gap*. Under certain circumstances, it can be proved to be equal to zero, which shows that the optimum values of problems (CP) and (CD) are equal. Before describing the conditions guaranteeing such a situation, called *strong* duality, we need to introduce the notion of strictly feasible point.

Definition 2.5 *A point x (resp. (y, s)) is said to be* strictly feasible *for the primal (resp. dual) problem if and only if it is feasible and belongs to the interior of the cone K (resp. K^*), i.e.*

$$Ax = b \text{ and } x \in \operatorname{int} K \quad (\text{resp. } A^T y + s = c \text{ and } s \in \operatorname{int} K^*)\ .$$

Moreover, we will say that the primal (resp. dual) problem is *unbounded* if $p^* = -\infty$ (resp. $d^* = +\infty$), that it is *infeasible* if there is no feasible solution, i.e. when $p^* = +\infty$ (resp. $d^* = -\infty$), and that it is *attained* if the optimum objective value p^* (resp. d^*) is achieved by at least one feasible primal (resp. dual) solution.

Theorem 2.4 (Strong duality) *If the dual problem* (CD) *admits a strictly feasible solution, we have either*

- ⋄ *an infeasible primal problem* (CP) *if the dual problem* (CD) *is unbounded, i.e. $p^* = d^* = +\infty$*
- ⋄ *a feasible primal problem* (CP) *if the dual problem* (CD) *is bounded. Moreover, in this case, the primal optimum is finite and attained with a zero duality gap, i.e. there is at least an optimal feasible solution x^* such that $c^T x^* = p^* = d^*$.*

The first case in this theorem (see e.g. [10, Theorem 2.7] for a proof) is a simple consequence of Theorem 2.3, and does not really depend on the existence of a strictly feasible solution for the dual, as opposed to the second case which relies on the existence of such a point. It is also worth to mention that boundedness of the dual problem (CD), defining the second case, is implied by the existence of a feasible primal solution, because of the weak duality theorem (however, the converse implication is not true in general, since a bounded dual problem can admit an infeasible primal problem ; an example of this situation is provided in Subsection 4.4).

This theorem is important, because it provides us with way to identify when both the primal and the dual problems have the same optimal value, and when this optimal value is attained by one of the problems. Obviously, this result can be dualized, meaning that the existence of a strictly feasible primal solution implies a zero duality gap and dual attainment. The combination of these two theorems leads to the following well-known and very useful corollary:

Corollary 2.1 *If both the primal and the dual problems admit a strictly feasible point, we have a zero duality gap and attainment for both problems, i.e. the same finite optimum objective value is attained for both problems.*

When the dual problem has no strictly feasible point, nothing can be said about the duality gap (which can happen to be strictly positive) and about attainment of the primal optimum objective value. However, even in this situation, we can prove an alternate version of the strong duality theorem involving the notion of primal problem subvalue. The idea behind this notion is to allow a small constraint violation in the infimum defining the primal problem (CP).

Definition 2.6 *The* subvalue *of primal problem* (CP) *is given by*

$$p^- = \lim_{\epsilon \to 0^+} \left[\inf_x c^T x \quad s.t. \quad \|Ax - b\| < \epsilon \text{ and } x \in K \right]$$

(a similar definition is holding for the dual subvalue d^-).

It is readily seen that this limit always exists (possibly being $+\infty$), because the feasible region of the infimum shrinks as ϵ tends to zero, which implies that its optimum value is a nonincreasing function of ϵ. Moreover, the inequality $p^- \leq p^*$ holds, because all the feasible regions of the infima defining p^- as ϵ tends to zero are larger than the actual feasible region of problem (CP).

The case $p^- = +\infty$, which implies that primal problem (CP) is infeasible (since we have then $p^* \geq p^- = +\infty$), is called primal *strong infeasibility*, and essentially means that the affine subspace defined by the linear constraints $Ax = b$ is strongly separated from cone K. We are now in position to state the following alternate strong duality theorem:

Theorem 2.5 (Strong duality, alternate version) *We have either*

- ⋄ *$p^- = +\infty$ and $d^* = -\infty$ when primal problem* (CP) *is strongly infeasible and dual problem* (CD) *is infeasible.*
- ⋄ *$p^- = d^*$ in all other cases.*

This theorem (see e.g. [10, Theorem 2.6] for a proof) states that there is no duality gap between p^- and d^*, except in the rather exceptional case of primal strong infeasibility and dual infeasibility. Note that the second case covers situations where the primal problem is infeasible but not strongly infeasible (i.e. $p^- < p^* = +\infty$).

To conclude this section, we would like to mention the fact that all the properties and theorems described in this section can be easily extended to the case of several conic constraints involving disjoint sets of variables.

Note 2.1 *Namely, having to satisfy the constraints $x^i \in K^i$ for all $i \in \{1, 2, \ldots, k\}$, where $K^i \subseteq \mathbb{R}^{n_i}$, we will simply consider the Cartesian product of these cones $K = K^1 \times K^2 \times \cdots \times K^k \subseteq \mathbb{R}^{\sum_{i=1}^k n_i}$ and express all these constraints simultaneously as $x \in K$ with $x = (x^1, x^2, \ldots, x^k)$.*

3 Cones for geometric optimization

Let us introduce the geometric cone $\mathcal{G}^n$, which will allow us to give a conic formulation of geometric optimization problems.

3.1 The geometric cone $\mathcal{G}^n$

Definition 3.1 *Let $n \in \mathbb{N}$. The* geometric cone *$\mathcal{G}^n$ is defined by*

$$\mathcal{G}^n = \left\{(x, \theta) \in \mathbb{R}^n_+ \times \mathbb{R}_+ \mid \sum_{i=1}^n e^{-\frac{x_i}{\theta}} \leq 1\right\}$$

using in the case of a zero denominator the following convention:

$$e^{-\frac{x_i}{0}} = 0 \ .$$

We observe that this convention results in $(x, 0) \in \mathcal{G}^n$ for all $x \in \mathbb{R}^n_+$. As special cases, we mention that $\mathcal{G}^0$ is the nonnegative real line $\mathbb{R}_+$, while $\mathcal{G}^1$ is easily shown to be equal to the 2-dimensional nonnegative orthant $\mathbb{R}^2_+$.

In order to use the powerful duality theory of convex optimization outlined in the previous section, we first prove that $\mathcal{G}^n$ is a convex cone. We are going to use the well-known weighted arithmetic-geometric mean inequality, which we recall here as a lemma:

Lemma 3.1 *Let $x \in \mathbb{R}^n_{++}$ and $\delta \in \mathbb{R}^n_{++}$ such that $\sum_{i=1}^n \delta_i = 1$. We have*

$$\prod_{i=1}^n x_i^{\delta_i} \leq \sum_{i=1}^n \delta_i x_i \ ,$$

equality occurring if and only if all x_i's are equal.

This result is easily proved, applying for example Jensen's inequality [7, Theorem 4.3] to the convex function $x \mapsto e^x$.

Theorem 3.1 *$\mathcal{G}^n$ is a convex cone.*

Proof. To prove that a set is a convex cone, it suffices to show that it is closed under addition and nonnegative scalar multiplication (Definition 2.1 and Theorem 2.1). Indeed, if $(x, \theta) \in \mathcal{G}^n$, $(x', \theta') \in \mathcal{G}^n$ and $\lambda \geq 0$, we have

$$\sum_{i=1}^n e^{-\frac{\lambda x_i}{\lambda \theta}} = \begin{cases} \sum_{i=1}^n e^{-\frac{x_i}{\theta}} \leq 1 & \text{if } \lambda > 0 \\ 0 \leq 1 & \text{if } \lambda = 0 \end{cases}$$

which shows that $\lambda(x, \theta) \in \mathcal{G}^n$. Looking now at $(x, \theta) + (x', \theta')$, we first consider the case $\theta > 0$ and $\theta' > 0$ and write

$$\sum_{i=1}^n e^{-\frac{x_i + x'_i}{\theta + \theta'}} = \sum_{i=1}^n \left(e^{-\frac{x_i}{\theta}}\right)^{\frac{\theta}{\theta+\theta'}} \left(e^{-\frac{x'_i}{\theta'}}\right)^{\frac{\theta'}{\theta+\theta'}} \quad ;$$

using now Lemma 3.1 on each term of the sum with vector $(e^{-\frac{x_i}{\theta}}, e^{-\frac{x'_i}{\theta'}})$ and weights $(\frac{\theta}{\theta+\theta'}, \frac{\theta'}{\theta+\theta'})$, satisfying $\frac{\theta}{\theta+\theta'} + \frac{\theta'}{\theta+\theta'} = 1$, we have

$$\begin{aligned}
\sum_{i=1}^{n} e^{-\frac{x_i+x'_i}{\theta+\theta'}} &\leq \sum_{i=1}^{n} \frac{\theta}{\theta+\theta'}(e^{-\frac{x_i}{\theta}}) + \frac{\theta'}{\theta+\theta'}(e^{-\frac{x'_i}{\theta'}}) \\
&= \frac{\theta}{\theta+\theta'} \sum_{i=1}^{n} e^{-\frac{x_i}{\theta}} + \frac{\theta'}{\theta+\theta'} \sum_{i=1}^{n} e^{-\frac{x'_i}{\theta'}} \\
&\leq \frac{\theta}{\theta+\theta'} 1 + \frac{\theta'}{\theta+\theta'} 1 = 1 \ ,
\end{aligned}$$

while in the case of $\theta' = 0$ we have

$$\sum_{i=1}^{n} e^{-\frac{x_i+x'_i}{\theta+\theta'}} = \sum_{i=1}^{n} e^{-\frac{x_i+x'_i}{\theta}} \leq \sum_{i=1}^{n} e^{-\frac{x_i}{\theta}} \leq 1$$

(the case $\theta = 0$ is similar). We have thus shown that $(x + x', \theta + \theta') \in \mathcal{G}^n$ in all cases, and therefore that $\mathcal{G}^n$ is a convex cone. □

We now proceed to prove some properties of the geometric cone $\mathcal{G}^n$.

Theorem 3.2 *$\mathcal{G}^n$ is closed.*

Proof. Let $\{(x^k, \theta^k)\}$ a sequence of points in $\mathbb{R}^{n+1}$ such that $(x^k, \theta^k) \in \mathcal{G}^n$ for all k and $\lim_{k\to\infty}(x^k, \theta^k) = (x^\infty, \theta^\infty)$. In order to prove that $\mathcal{G}^n$ is closed, it suffices to show that $(x^\infty, \theta^\infty) \in \mathcal{G}^n$. Let us distinguish two cases:

- ◇ $\theta^\infty > 0$. Using the easily proven fact that functions $(x_i, \theta) \mapsto e^{-\frac{x_i}{\theta}}$ are continuous on $\mathbb{R}_+ \times \mathbb{R}_{++}$, we have that

$$\sum_{i=1}^{n} e^{-\frac{x_i^\infty}{\theta^\infty}} = \sum_{i=1}^{n} \lim_{k\to\infty} e^{-\frac{x_i^k}{\theta^k}} = \lim_{k\to\infty} \sum_{i=1}^{n} e^{-\frac{x_i^k}{\theta^k}} \leq 1 \ ,$$

  which implies $(x^\infty, \theta^\infty) \in \mathcal{G}^n$.

- ◇ $\theta^\infty = 0$. Since $(x^k, \theta^k) \in \mathcal{G}^n$, we have $x^k \geq 0$ and thus $x^\infty \geq 0$, which implies that $(x^\infty, 0) \in \mathcal{G}^n$.

In both cases, $(x^\infty, \theta^\infty)$ is shown to belong to $\mathcal{G}^n$, which proves the claim. □

In order to use the strong duality theorem, we now proceed to identify the interior of the geometric cone.

Theorem 3.3 *The interior of $\mathcal{G}^n$ is given by*

$$\operatorname{int} \mathcal{G}^n = \left\{(x, \theta) \in \mathbb{R}^n_{++} \times \mathbb{R}_{++} \mid \sum_{i=1}^{n} e^{-\frac{x_i}{\theta}} < 1\right\} .$$

Proof. A point x belongs to the interior of a set S if and only if there exists an open ball centered at x entirely included in S. Let $(x, \theta) \in \mathcal{G}^n$. We first note that $(x, 0)$ cannot belong to $\operatorname{int} \mathcal{G}^n$, because every open ball centered at $(x, 0)$ contains a point with a negative θ component, which does not belong to the cone $\mathcal{G}^n$. Suppose $\theta > 0$ and the inequality in the definition of $\mathcal{G}^n$ is satisfied with equality, i.e.

$$\sum_{i=1}^{n} e^{-\frac{x_i}{\theta}} = 1 .$$

Every open ball centered at (x, θ) contains a point (x', θ') with $x' < x$ and $\theta' > \theta$, which satisfies then

$$\sum_{i=1}^{n} e^{-\frac{x_i'}{\theta'}} > \sum_{i=1}^{n} e^{-\frac{x_i}{\theta}} = 1$$

and is thus outside of $\mathcal{G}^n$, implying $(x, \theta) \notin \operatorname{int} \mathcal{G}^n$. We now show that all the remaining points that do not satisfy one of the two conditions mentioned above, i.e. the points with $\theta > 0$ satisfying the strict inequality, belong to the interior of $\mathcal{G}^n$. Let (x, θ) one of these points, and $\mathcal{B}(\epsilon)$ the open ball centered at (x, θ) with radius ϵ. Restricting ϵ to sufficiently small values (i.e. choosing $\epsilon < \theta$), we have for all points $(x', \theta') \in \mathcal{B}(\epsilon)$

$$x_i - \epsilon \leq x_i' \leq x_i + \epsilon \text{ and } 0 < \theta - \epsilon \leq \theta' \leq \theta + \epsilon ,$$

which implies

$$\frac{x_i'}{\theta'} \geq \frac{x - \epsilon}{\theta + \epsilon} \quad \text{and thus} \quad \sum_{i=1}^{n} e^{-\frac{x_i'}{\theta'}} \leq \sum_{i=1}^{n} e^{-\frac{x_i - \epsilon}{\theta + \epsilon}} \text{ for all } (x', \theta') \in \mathcal{B}(\epsilon) . \tag{3.1}$$

Taking the limit of the last right-hand side when $\epsilon \to 0$, we find

$$\lim_{\epsilon \to 0} \sum_{i=1}^{n} e^{-\frac{x_i - \epsilon}{\theta + \epsilon}} = \sum_{i=1}^{n} e^{-\frac{x_i}{\theta}} < 1$$

(because of the continuity of functions $(x_i, \theta) \mapsto e^{-\frac{x_i}{\theta}}$ on $\mathbb{R}_+ \times \mathbb{R}_{++}$). Therefore we can assume the existence of a value ϵ^* such that

$$\sum_{i=1}^{n} e^{-\frac{x_i - \epsilon^*}{\theta + \epsilon^*}} < 1 ,$$

which because of (3.1) will imply that

$$\sum_{i=1}^{n} e^{-\frac{x_i'}{\theta'}} < 1$$

for all $(x', \theta') \in \mathcal{B}(\epsilon^*)$. This inequality, combined with $\theta' > 0$, is sufficient to prove that the open ball $\mathcal{B}(\epsilon^*)$ is entirely included in $\mathcal{G}^n$, hence that $(x, \theta) \in \operatorname{int} \mathcal{G}^n$. $\square$

Theorem 3.4 *$\mathcal{G}^n$ is solid and pointed.*

Proof. The fact that $0 \in \mathcal{G}^n \subseteq \mathbb{R}_+^{n+1}$ implies that $\mathcal{G}^n \cap -\mathcal{G}^n = \{0\}$, i.e. $\mathcal{G}^n$ is pointed (Definition 2.3). To prove it is solid (Definition 2.2), we simply provide a point belonging to its interior, for example $(e, \frac{1}{n})$ (where e stands for the all-one vector). We have then

$$\sum_{i=1}^{n} e^{-\frac{x_i}{\theta}} = ne^{-n} < 1 \ ,$$

because $e^n > n$ for all $n \in \mathbb{N}$, and therefore $(e, \frac{1}{n}) \in \operatorname{int} \mathcal{G}^n$. $\square$

To summarize, $\mathcal{G}^n$ is a solid pointed close convex cone, hence suitable for conic optimization.

3.2 The dual geometric cone $(\mathcal{G}^n)^*$

In order to express the dual of a conic problem involving the geometric cone $\mathcal{G}^n$, we need to find an explicit description of its dual.

Theorem 3.5 *The dual of $\mathcal{G}^n$ is given by*

$$(\mathcal{G}^n)^* = \left\{ (x^*, \theta^*) \in \mathbb{R}_+^n \times \mathbb{R} \mid \theta^* \geq \sum_{i|x_i^*>0} x_i^* \log \frac{x_i^*}{\sum_{i=1}^n x_i^*} \right\} .$$

Proof. Using Definition 2.4 for the dual cone, we have

$$(\mathcal{G}^n)^* = \left\{ (x^*, \theta^*) \in \mathbb{R}^n \times \mathbb{R} \mid (x, \theta)^T (x^*, \theta^*) \geq 0 \text{ for all } (x, \theta) \in \mathcal{G}^n \right\}$$

(the * superscript on variables x^* and θ^* is a reminder of their dual nature). This condition on (x^*, θ^*) is equivalent to saying that the following infimum

$$\delta(x^*, \theta^*) = \inf x^T x^* + \theta\theta^* \quad \text{s.t.} \quad (x, \theta) \in \mathcal{G}^n \ .$$

has to be nonnegative. Let us distinguish the cases $\theta = 0$ and $\theta > 0$: we have that

$$\delta(x^*, \theta^*) = \min\{\delta_1(x^*, \theta^*), \delta_2(x^*, \theta^*)\}$$

with

$$\left\{ \begin{array}{l} \delta_1(x^*, \theta^*) = \inf \ x^T x^* + \theta\theta^* \quad \text{s.t.} \quad (x, \theta) \in \mathcal{G}^n \text{ and } \theta = 0 \\ \delta_2(x^*, \theta^*) = \inf \ x^T x^* + \theta\theta^* \quad \text{s.t.} \quad (x, \theta) \in \mathcal{G}^n \text{ and } \theta > 0 \end{array} \right. .$$

The first of these infima can be rewritten as

$$\inf x^T x^* \quad \text{s.t.} \quad x \geq 0 \ ,$$

since $(x, 0) \in \mathcal{G}^n \Leftrightarrow x \geq 0$. It is easy to see that this infimum is equal to 0 if $x^* \geq 0$ and to $-\infty$ when $x^* \ngeq 0$. Since we are looking for points with a nonnegative infimum $\delta(x^*, \theta^*)$, we will require in the rest of this proof x^* to be nonnegative and only consider the second infimum, which is equal to

$$\inf \ \theta \left[\frac{x^T x^*}{\theta} + \theta^* \right] \quad \text{s.t.} \quad \sum_{i=1}^{n} e^{-\frac{x_i}{\theta}} \leq 1 \text{ and } (x, \theta) \in \mathbb{R}_+^n \times \mathbb{R}_{++} \ . \tag{3.2}$$

Let us again distinguish two cases. When $x^* = 0$, this infimum becomes

$$\inf \theta\theta^* \quad \text{s.t.} \quad \sum_{i=1}^n e^{-\frac{x_i}{\theta}} \le 1 \text{ and } (x,\theta) \in \mathbb{R}^n_+ \times \mathbb{R}_{++} \ ,$$

which is nonnegative if and only if $\theta^* \ge 0$, since θ can take any value in the open positive interval $]0 + \infty[$. On the other hand, if $x^* \neq 0$, we have $\sum_{i=1}^n x_i^* > 0$ and can define the auxiliary variables w_i^* by

$$w_i^* = \frac{x_i^*}{\sum_{i=1}^n x_i^*}$$

(in order to simplify notations). We write the following chain of inequalities

$$1 \ge \sum_{i=1}^n e^{-\frac{x_i}{\theta}} \ge \sum_{i|w_i^*>0} e^{-\frac{x_i}{\theta}} = \sum_{i|w_i^*>0} w_i^* \left(\frac{e^{-\frac{x_i}{\theta}}}{w_i^*}\right) \ge \prod_{i|w_i^*>0} \left(\frac{e^{-\frac{x_i}{\theta}}}{w_i^*}\right)^{w_i^*} \tag{3.3}$$

The second inequality comes from the fact that each term of the sum is positive (we remove some terms), and the third one uses Lemma 3.1 with weights w_i^*, noting that $\sum_{i|w_i^*>0} w_i^* = \sum_{i=1}^n w_i^* = 1$. From this last inequality we derive successively

$$\begin{aligned}
\prod_{i|w_i^*>0} e^{-\frac{x_i w_i^*}{\theta}} &\le \prod_{i|w_i^*>0} w_i^{*w_i^*} \ , \\
-\sum_{i|w_i^*>0} \frac{x_i w_i^*}{\theta} &\le \sum_{i|w_i^*>0} w_i^* \log w_i^* \quad \text{(taking the logarithms)} \ , \\
\sum_{i=1}^n \frac{x_i x_i^*}{\theta} &\ge -\sum_{i|x_i^*>0} x_i^* \log w_i^* \quad \text{(multiplying by } -\textstyle\sum_{i=1}^n x_i^*) \ , \\
\frac{x^T x^*}{\theta} + \theta^* &\ge \theta^* - \sum_{i|x_i^*>0} x_i^* \log w_i^* \ , \text{ and finally} \\
\inf_{(x,\theta)\in\mathcal{G}^n|\theta>0} \frac{x^T x^*}{\theta} + \theta^* &\ge \theta^* - \sum_{i|x_i^*>0} x_i^* \log w_i^* \ .
\end{aligned}$$

Examining carefully the chain of inequalities in (3.3), we observe that a suitable choice of (x,θ) can lead to attainment of this last infimum: namely, we need to have

- ⋄ $\sum_{i=1}^n e^{-\frac{x_i}{\theta}} = 1$, for the first inequality in (3.3),
- ⋄ $x_i \to +\infty$ for all indices i such that $w_i^* = 0$, in order to have $e^{-\frac{x_i}{\theta}} \to 0$ when $w_i^* = 0$ for the second inequality in (3.3),
- ⋄ all terms $(\frac{e^{-\frac{x_i}{\theta}}}{w_i^*})$ with indices such that $w_i^* > 0$ equal to each other, for the third inequality in (3.3).

These conditions are compatible: summing up the constant terms, we find

$$\frac{e^{-\frac{x_i}{\theta}}}{w_i^*} \text{ (when } w_i^* > 0) = \frac{\sum_{i|w_i^*>0} e^{-\frac{x_i}{\theta}}}{\sum_{i|w_i^*>0} w_i^*} = \sum_{i|w_i^*>0} e^{-\frac{x_i}{\theta}} \to \sum_{i=1}^n e^{-\frac{x_i}{\theta}} = 1 \ ,$$

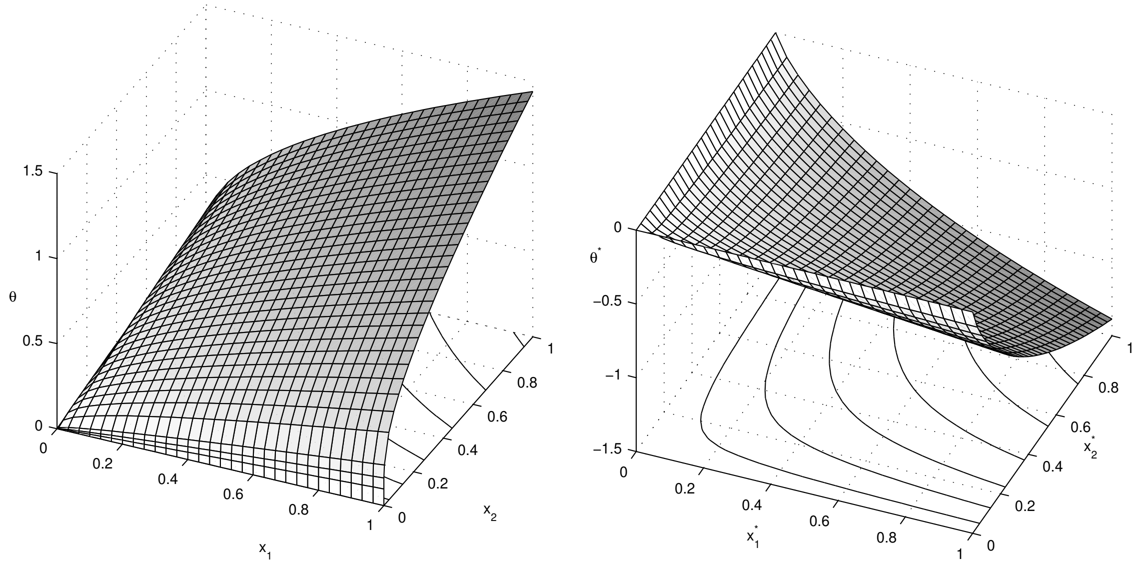

Figure 1: The boundary surfaces of $\mathcal{G}^2$ and its dual cone $(\mathcal{G}^2)^*$.

which gives $e^{-\frac{x_i}{\theta}} = w_i^*$ for all i such that $w_i^* > 0$. Summarizing, we can choose x according to

$$\begin{cases} x_i = -\theta \log w_i^* & \text{when } w_i^* > 0 \\ x_i \to +\infty & \text{when } w_i^* = 0 \end{cases},$$

which proves that

$$\inf_{(x,\theta)\in\mathcal{G}^n \,|\, \theta>0} \frac{x^T x^*}{\theta} + \theta^* = \theta^* - \sum_{i|x_i^*>0} x_i^* \log w_i^* \ . \tag{3.4}$$

Since the additional multiplicative θ in (3.2) doesn't change the sign of this infimum (because $\theta > 0$), we may conclude that it is nonnegative if and only if

$$\theta^* - \sum_{i|x_i^*>0} x_i^* \log w_i^* \geq 0 \ .$$

Combining with the special case $x^* = 0$ and the constraint $x^* \geq 0$ implied by the first infimum, we conclude that the dual cone is given by

$$(\mathcal{G}^n)^* = \left\{ (x^*, \theta^*) \in \mathbb{R}_+^n \times \mathbb{R} \mid \theta^* \geq \sum_{i|x_i^*>0} x_i^* \log w_i^* \right\} ,$$

as announced. □

As special cases, since $\mathcal{G}^0 = \mathbb{R}_+$ and $\mathcal{G}^1 = \mathbb{R}_+^2$, we may check that $(\mathcal{G}^0)^* = (\mathbb{R}_+)^* = \mathbb{R}_+$ and $(\mathcal{G}^1)^* = (\mathbb{R}_+^2)^* = \mathbb{R}_+^2$, as expected. These two cones are thus self-dual, but it is easy to see that geometric cones of higher dimension are not self-dual any more. To illustrate our purpose, we provide in Figure 1 the three-dimensional graphs of the boundary surfaces of $\mathcal{G}^2$ and $(\mathcal{G}^2)^*$.

Note 3.1 *Since we have $0 \leq w_i^* \leq 1$ for all indices i, each logarithmic term appearing in this definition is nonpositive, as well as their sum, which means that $(x^*, \theta^*) \in (\mathcal{G}^n)^*$ as soon as x^* and θ^* are nonnegative. This fact could have been guessed prior to any computation: noticing that $\mathcal{G}^n \subseteq \mathbb{R}_+^{n+1}$ and $(\mathbb{R}_+^{n+1})^* = \mathbb{R}_+^{n+1}$, we immediately have that $(\mathcal{G}^n)^* \supseteq \mathbb{R}_+^{n+1}$, because taking the dual of a set inclusion reverses its direction.*

Finding the dual of $\mathcal{G}^n$ was a little involved, but establishing its properties is straightforward.

Theorem 3.6 *$(\mathcal{G}^n)^*$ is a solid, pointed, closed convex cone. Moreover, $((\mathcal{G}^n)^*)^* = \mathcal{G}^n$.*

The proof of this fact is immediate by Theorem 2.2 since $(\mathcal{G}^n)^*$ is the dual of a solid, pointed, closed convex cone. □

The interior of $(\mathcal{G}^n)^*$ is also rather easy to obtain:

Theorem 3.7 *The interior of $(\mathcal{G}^n)^*$ is given by*

$$\operatorname{int}(\mathcal{G}^n)^* = \left\{ (x^*, \theta^*) \in \mathbb{R}_{++}^n \times \mathbb{R} \mid \theta^* > \sum_{i=1}^n x_i^* \log \frac{x_i^*}{\sum_{i=1}^n x_i^*} \right\} .$$

Proof. We first note that $(\mathcal{G}^n)^*$, a convex set, is the epigraph of the following function

$$f_n : \mathbb{R}_+^n \mapsto \mathbb{R} : x \mapsto \sum_{i \mid x_i^* > 0} x_i^* \log \frac{x_i^*}{\sum_{i=1}^n x_i^*} ,$$

which implies that f_n is convex (by definition of a convex function). Hence we can apply Lemma 7.3 in [7] to get

$$\operatorname{int}(\mathcal{G}^n)^* = \operatorname{int} \operatorname{epi} f_n = \{ (x^*, \theta^*) \in \operatorname{int} \operatorname{dom} f_n \times \mathbb{R} \mid \theta^* > f_n(x^*) \} ,$$

which is exactly our claim since $\operatorname{int} \mathbb{R}_+^n = \mathbb{R}_{++}^n$. □

The last piece of information we need about the pair of cones $(\mathcal{G}^n, (\mathcal{G}^n)^*)$ is its set of orthogonality conditions.

Theorem 3.8 (orthogonality conditions) *Let $v = (x, \theta) \in \mathcal{G}^n$ and $v^* = (x^*, \theta^*) \in (\mathcal{G}^n)^*$. We have $v^T v^* = 0$ if and only if one of these two sets of conditions is satisfied*

$$\begin{array}{lll} \theta = 0 & \textit{and} & x_i x_i^* = 0 \textit{ for all } i \\ \theta > 0 & \textit{and} & \left\{ \begin{array}{rcl} \sum_{i \mid x_i^* > 0} x_i^* \log w_i^* & = & \theta^* \\ (\sum_{i=1}^n x_i^*) e^{-\frac{x_i}{\theta}} & = & x_i^* \textit{ for all } i \end{array} \right. \end{array} .$$

Proof. To prove this fact, we merely have to reread carefully the proof of Theorem 3.5, paying attention to the cases where the infimum is equal to zero. In the first case examined, $\theta = 0$, we have $v^T v^* = x^T x^*$. Since x and x^* are two nonnegative vectors, we have $v^T v^* = 0$ if and only if $x_i x_i^* = 0$ for every index i, which gives the first set of conditions of the theorem.

When $\theta > 0$, we first have the special case $x^* = 0$ which gives $v^T v^* = \theta\theta^*$. This quantity can only be zero if $\theta^* = 0$, i.e. when $(x^*, \theta^*) = 0$. When $x^* \neq 0$, the proof of Theorem 3.5

shows that $v^T v^*$ can only be zero when the infimum (3.4) is equal to zero and attained, which implies $\theta^* = \sum_{i|x_i^*>0} x_i^* \log w_i^*$. However, this infimum is not always attained by a finite vector (x, θ), because of the condition $x_i \to +\infty$ that is required when $w_i^* = 0$. The scalar product $v^T v^*$ is thus equal to zero only if all w_i^*'s are positive, i.e. when all x_i^*'s are positive: in this case, the two sets of equalities $\theta^* = \sum_{i|x_i^*>0} x_i^* \log w_i^*$ (to have a zero infimum) and $e^{-\frac{x_i}{\theta}} = w_i^*$ (to attain the infimum) must be satisfied.

Rephrasing this last equality as $(\sum_{i=1}^n x_i^*)e^{-\frac{x_i}{\theta}} = x_i^*$ to take into account the special case $(x^*, \theta^*) = 0$, we find the second set of conditions of our theorem. □

4 Duality for geometric optimization

In this section, we introduce a form of geometric optimization problems that is suitable to our purpose and prove several duality properties using the previously defined primal-dual pair of convex cones. These results are well-known and can be found e.g. in [2]. However, our presentation differs and handles problems expressed in a slightly different (but equivalent) format, and hence provides results adapted to the formulation we use. We refer the reader to Subsection 5.1 where the connection is made between our results and their classical counterparts.

4.1 Conic formulation

We start with the original formulation of a geometric optimization problem (see e.g. [2]). Let $R = \{0, 1, 2, \ldots, r\}$ and $I = \{1, 2, \ldots, n\}$. Let $\{I_k\}_{k\in R}$ a partition of I into $r+1$ classes, i.e. satisfying

$$\cup_{k\in R} I_k = I \text{ and } I_k \cap I_l = \emptyset \text{ for all } k \neq l \ .$$

The primal geometric optimization problem is the following:

$$\inf G_0(t) \quad \text{s.t.} \quad t \in \mathbb{R}^m_{++} \text{ and } G_k(t) \leq 1 \text{ for all } k \in R \setminus \{0\} \ , \tag{PG}$$

where t is the m-dimensional column vector we want to optimize and the functions G_k defining the objective and the constraints are so-called posynomials, given by

$$G_k : \mathbb{R}^m_{++} \mapsto \mathbb{R}_{++} : t \mapsto \sum_{i\in I_k} C_i \prod_{j=1}^m t_j^{a_{ij}} \ ,$$

where exponents a_{ij} are arbitrary real numbers and coefficients C_i are required to be strictly positive (hence the name *posy*nomial). These functions are very well suited for the formulation of constraints that come from the laws of physics or economics (either directly or using an empirical fit).

Problem (PG) is not convex (take for example $G_0 : t \mapsto t^{1/2}$, which is not a convex function). However, it is well-known that it can be convexified with the following change of variables:

$$t_j = e^{y_j} \text{ for all } j \in \{1, 2, \ldots, m\} \ , \tag{4.1}$$

to become

$$\inf g_0(y) \quad \text{s.t.} \quad g_k(y) \leq 1 \text{ for all } k \in R \setminus \{0\} \ . \tag{TPG}$$

The functions g_k are defined to satisfy $g_k(y) = G_k(t)$ when (4.1) holds, which means

$$g_k : \mathbb{R}^m \mapsto \mathbb{R}_{++} : y \mapsto \sum_{i \in I_k} C_i \prod_{j=1}^{m} (e^{y_j})^{a_{ij}} = \sum_{i \in I_k} e^{-c_i + \sum_{j=1}^{m} y_j a_{ij}} = \sum_{i \in I_k} e^{a_i^T y - c_i} ,$$

where the coefficient vector $c \in \mathbb{R}^n$ is given by $c_i = -\log C_i$ and $a_i = (a_{i1}, a_{i2}, \ldots, a_{im})^T$ is an m-dimensional column vector. Note that unlike the original variables t and coefficients C, variables y and coefficients c are not required to be strictly positive and can take any real value.

It is straightforward to check that functions g_k are now convex, hence that (TPG) is a convex optimization problem. However, we will not establish convexity directly but rather derive it from the fact that problem (TPG) can be cast as a conic optimization problem. Moreover, following others [4, 1], we will not use this formulation but instead work with a slight variation featuring a linear objective:

$$\sup b^T y \quad \text{s.t.} \quad g_k(y) \leq 1 \text{ for all } k \in R , \tag{LPG}$$

where $b \in \mathbb{R}^m$ and 0 has been removed from set R.

It will be shown later that problems in the form (TPG) (and (PG)) can be expressed in this format, and the results we are going to obtain about problem (LPG) will be translated back to these more traditional settings later in Subsection 5.1.

Let us now model problem (LPG) with a conic formulation. In the rest of this article, we will use the following useful convention: v_S (resp. M_S) denotes the restriction of column vector v (resp. matrix M) to the components (resp. rows) whose indices belong to set S. We introduce a vector of auxiliary variables $s \in \mathbb{R}^n$ to represent the exponents used in functions g_k, more precisely we let

$$s_i = c_i - a_i^T y \text{ for all } i \in I \text{ or, in matrix form, } s = c - A^T y ,$$

where A is a $m \times n$ matrix whose columns are a_i. Our problem becomes then

$$\sup b^T y \quad \text{s.t.} \quad s = c - A^T y \text{ and } \sum_{i \in I_k} e^{-s_i} \leq 1 \text{ for all } k \in R ,$$

which is readily seen to be equivalent to the following, using the definition of $\mathcal{G}^n$,

$$\sup b^T y \quad \text{s.t.} \quad A^T y + s = c \text{ and } (s_{I_k}, 1) \in \mathcal{G}^{\# I_k} \text{ for all } k \in R ,$$

and finally to

$$\sup b^T y \quad \text{s.t.} \quad \begin{pmatrix} A^T \\ 0 \end{pmatrix} y + \begin{pmatrix} s \\ v \end{pmatrix} = \begin{pmatrix} c \\ e \end{pmatrix} \text{ and } (s_{I_k}, v_k) \in \mathcal{G}^{n_k} \text{ for all } k \in R , \tag{CPG}$$

where e is the all-one vector in $\mathbb{R}^r$, $n_k = \# I_k$ and an additional vector of fictitious variables $v \in \mathbb{R}^r$ has been introduced, whose components are fixed to 1 by part of the linear constraints. This is exactly a conic optimization problem, in the dual form (CD), using variables $(\tilde{y}, \tilde{s})$, data $(\tilde{A}, \tilde{b}, \tilde{c})$ and a cone K^* such that

$$\tilde{y} = y, \ \tilde{s} = \begin{pmatrix} s \\ v \end{pmatrix}, \ \tilde{A} = \begin{pmatrix} A & 0 \end{pmatrix}, \ \tilde{b} = b, \ \tilde{c} = \begin{pmatrix} c \\ e \end{pmatrix} \text{ and } K^* = \mathcal{G}^{n_1} \times \mathcal{G}^{n_2} \times \cdots \times \mathcal{G}^{n_r} ,$$

where K^* has been defined according to Note 2.1, since we have to deal with multiple conic constraints involving disjoint sets of variables.

Using properties of $\mathcal{G}^n$ and $(\mathcal{G}^n)^*$ proved in the previous section, it is straightforward to show that K^* is a solid, pointed, closed convex cone whose dual is

$$(K^*)^* = K = (\mathcal{G}^{n_1})^* \times (\mathcal{G}^{n_2})^* \times \cdots \times (\mathcal{G}^{n_r})^* \ ,$$

another solid, pointed, closed convex cone, according to Theorem 3.6. This allows us to derive a dual problem to (CPG) in a completely mechanical way and find the following conic optimization problem, expressed in the primal form (CP):

$$\inf \begin{pmatrix} c^T & e^T \end{pmatrix} \begin{pmatrix} x \\ z \end{pmatrix} \quad \text{s.t.} \quad \begin{pmatrix} A & 0 \end{pmatrix} \begin{pmatrix} x \\ z \end{pmatrix} = b \text{ and } (x_{I_k}, z_k) \in (\mathcal{G}^{n_k})^* \text{ for all } k \in R \ , \tag{CDG}$$

where $x \in \mathbb{R}^n$ and $z \in \mathbb{R}^r$ are the vectors we optimize. This problem can be simplified: making the conic constraints explicit, we find

$$\inf c^T x + e^T z \quad \text{s.t.} \quad Ax = b,\ x_{I_k} \geq 0 \text{ and } z_k \geq \sum_{i \in I_k | x_i > 0} x_i \log \frac{x_i}{\sum_{i \in I_k} x_i} \text{ for all } k \in R \ ,$$

which can be further reduced to

$$\inf c^T x + \sum_{k \in R} \sum_{i \in I_k | x_i > 0} x_i \log \frac{x_i}{\sum_{i \in I_k} x_i} \quad \text{s.t.} \quad Ax = b \text{ and } x \geq 0 \ . \tag{LDG}$$

Indeed, since each variable z_k is free except for the inequality coming from the associated conic constraint, these inequalities must be satisfied with equality at each optimum solution and variables z can therefore be removed from the formulation. As could be expected, the dual problem we have just found using conic duality and our primal-dual pair of cones $(\mathcal{G}^n, (\mathcal{G}^n)^*)$ corresponds to the usual dual for problem (LPG) found in the literature [5, 1]. We will also show later in Subsection 5.1 that it also allows us to derive the dual problem in the traditional formulations (PG) and (TPG).

4.2 Duality theory

We are now about to apply the various duality theorems enumerated in the previous section to geometric optimization. Our strategy will be the following: in order to prove results about the pair (LPG)–(LDG), we are going to apply our theorems to the conic primal-dual pair (CPG)–(CDG) and use the equivalence that holds between (CPG) and (LPG) and between (CDG) and (LDG). We start with the weak duality theorem.

Theorem 4.1 (Weak duality) *Let y a feasible solution for primal problem* (LPG) *and x a feasible solution for dual problem* (LDG). *We have*

$$b^T y \leq c^T x + \sum_{k \in R} \sum_{i \in I_k | x_i > 0} x_i \log \frac{x_i}{\sum_{i \in I_k} x_i} \ , \tag{4.2}$$

equality occurring if and only if

$$\Big(\sum_{i \in I_k} x_i \Big) e^{a_i^T y - c_i} = x_i \textit{ for all } i \in I_k, k \in R \ .$$

Proof (the original proof can be found in [8] or [4, §1]). On the one hand, we note that y can be easily converted to a feasible solution (y, s, v) for the conic problem (CPG), simply by choosing vectors s and v according to the linear constraints. On the other hand, x can also be converted to a feasible solution (x, z) for the conic problem (CDG), admitting the same objective value, by choosing

$$z_k = \sum_{i \in I_k | x_i > 0} x_i \log \frac{x_i}{\sum_{i \in I_k} x_i} \text{ for all } k \in R \, . \tag{4.3}$$

Applying now the weak duality theorem to the conic primal-dual pair (CPG)–(CDG) with feasible solutions (x, z) and (y, s, v), we find the announced inequality

$$b^T y \leq c^T x + \sum_{k \in R} \sum_{i \in I_k | x_i > 0} x_i \log \frac{x_i}{\sum_{i \in I_k} x_i} \, ,$$

equality occurring if and only if the orthogonality conditions given in Theorem 3.8 are satisfied for each conic constraint. Since θ corresponds here to v_k, which is always equal to 1 because of the linear constraints, we can rule out the first set of equalities (occurring where $\theta = 0$) and keep only the second set of conditions. The first of these equalities being always satisfied because of our choice of z_k, we finally conclude that equality (4.2) can occur if and only if the following set of remaining equalities is satisfied, namely

$$\Big(\sum_{i \in I_k} x_i \Big) e^{-\frac{s_i}{v_i}} = x_i \text{ for all } i \in I_k, k \in R \, ,$$

which is equivalent to our claim because of the linear constraints on s_i and v_i. □

The following theorem is an application of the strong duality theorem, and requires therefore the existence of a specific primal feasible solution.

Theorem 4.2 *If there exists a feasible solution for the primal problem* (LPG) *satisfying strictly the inequality constraints, i.e. a vector y such that*

$$g_k(y) < 1 \textit{ for all } k \in R$$

we have either

- *an infeasible dual problem* (LDG) *if primal problem* (LPG) *is unbounded*
- *a feasible dual problem* (LDG) *whose optimum objective value is attained by a feasible vector x if primal problem* (LPG) *is bounded. Moreover, the optimum objective values of* (LPG) *and* (LDG) *are equal.*

Proof (a more classical proof can be found in [4, §2]). Choosing again vectors s and v according to the linear constraints, we find a feasible solution (y, s, v) for the primal conic problem (CPG). Moreover, recalling the description of $\operatorname{int} \mathcal{G}^n$ given by Theorem 3.3, the conditions $v_k = 1 > 0$ and $g_k(y) = \sum_{i \in I_k} e^{-s_i} < 1$ ensure that (y, s, v) is a strictly feasible solution for (CPG). Theorem 2.4 implies then that we have either

- ◇ an infeasible dual problem (CDG) if primal problem (CPG) is unbounded: this is equivalent to the first part of our claim, since it is clear that (CPG) is unbounded if and only if (LPG) is unbounded and that (CDG) is infeasible if and only if (LDG) is infeasible (indeed, (x, z) feasible for (CDG) implies x feasible for (LDG), while x feasible for (LDG) implies $(x, 0)$ feasible for (CDG)). This fact could also have been obtained as a simple consequence of Theorem 4.1.
- ◇ a feasible dual problem (CDG) whose optimum objective value is attained by a feasible vector (x, z) if primal problem (CPG) is bounded. Moreover, the optimum objective values of (CPG) and (CDG) are equal. Obviously, the finite optimum objective values of (CPG) and (LPG) are equal. It is also clear that optimal variables z_k in (CDG) must attain the lower bounds defined by the conic constraints, as in (4.3), which implies that vector x is optimum for problem (LDG) and has the same objective value as (x, z) in (CDG). This proves the second part of our claim.

□

Let us note again that a sufficient condition for the second case of this theorem to happen is the existence of a feasible solution for the dual problem (LDG), because of the weak duality property.

The strong duality theorem can also be applied on the dual side.

Theorem 4.3 *If there exists is a strictly positive feasible solution for the dual problem* (LDG), *i.e. a vector x such that*

$$Ax = b \text{ and } x > 0 \, ,$$

we have either

- ◇ *an infeasible primal problem* (LPG) *if dual problem* (LDG) *is unbounded*
- ◇ *a feasible primal problem* (LPG) *whose optimum objective value is attained by a feasible vector y if dual problem* (LDG) *is bounded. Moreover, the optimum objective values of* (LDG) *and* (LPG) *are equal.*

Proof (a traditional proof can be found in [4, §5]). As for the previous theorem, the first part of our claim is a direct consequence of Theorem 4.1, that does not really rely on the existence of a strictly positive x. Let us prove the second part of our claim and suppose that problem (LDG) is bounded. Problem (CDG) cannot be unbounded, because each feasible solution (x, z) for (CDG) leads to a feasible x for (LDG) with a lower objective (because of the conic constraints), which would also lead to an unbounded (LDG). Using the description of $\mathrm{int}(\mathcal{G}^n)^*$ given by Theorem 3.7, we find that a feasible $x > 0$ for (LDG) can be easily converted to a strictly feasible solution (x, z) for (CDG), taking sufficiently large values for variables z_k (letting $z_k = 1$ for example is enough). The strong duality theorem implies thus, since (CDG) has been shown to be bounded, that problem (CPG) is feasible with an optimum objective value attained by a feasible vector (y, s, v) and equal to the dual optimum objective value of (CDG). Obviously, on the one hand, vector y is a feasible optimum solution to problem (LPG), attaining the same objective value as (y, s, v) in (CDG). On the other hand, the finite optimum objective values of (CDG) and (LDG) must be equal, even if no feasible solution is actually optimum (since x feasible for (LDG) implies (x, z) feasible for (CDG) with the same objective value and (x, z) feasible for (CDG) implies x feasible for (LDG) with a smaller or equal objective value). This is enough to prove the second part of our claim. □

To conclude this section, we prove a last theorem that involves the alternate version of the strong duality theorem. Let us introduce the following family of optimization problems, parameterized by a strictly positive parameter δ:

$$\hat{p}(\delta) \quad = \quad \sup b^T y \quad \text{s.t.} \quad g_k(y) \leq e^\delta \text{ for all } k \in R \ . \tag{LPG$_\delta$}$$

It is clear that each of these problems is a (strict) relaxation of problem (LPG), because $e^\delta > 1$ for $\delta > 0$, hence we have $\hat{p}(\delta) \geq p^*$ for all δ. Moreover, since the feasible region of these problems shrinks as δ tends to zero, $\hat{p}(\delta)$ is a nondecreasing function of δ and we can always define the following limit

$$\hat{p} = \lim_{\delta \to 0^+} \hat{p}(\delta) \ ,$$

which we will call the *subvalue* of problem (LPG). We have the following theorem

Theorem 4.4 *If there exists a feasible solution to the dual problem* (LDG)*, the subvalue of the primal problem* (LPG) *is equal to the optimum objective value of the dual problem* (LDG).

Proof. We are going to show in fact that the primal subvalue $\hat{p}$ is equal to the subvalue p^- of the primal conic optimization problem (CPG). Using Theorem 2.5 on the primal-dual conic pair (CPG)–(CDG), we will find that $p^- = d^*$ (the first case of the theorem cannot happen since (LDG), and hence (CDG), is feasible by hypothesis). Noting finally that the optimum objective values of (CDG) and (LDG) are equal (which has been shown in the course of the previous proof) will conclude our proof.

Let us restate the definition of the subvalue p^- for problem (CPG). Defining the following family of problems, parameterized by a strictly positive parameter ϵ,

$$\sup_{(y,s,v)} b^T y \quad \text{s.t.} \quad \left\| \begin{pmatrix} A^T \\ 0 \end{pmatrix} y + \begin{pmatrix} s \\ v \end{pmatrix} - \begin{pmatrix} c \\ e \end{pmatrix} \right\| < \epsilon, \ (s_{I_k}, v_k) \in \mathcal{G}^{n_k} \ \forall k \in R \ , \tag{CPG$_\epsilon$}$$

whose optimum objective values will be denoted by $\bar{p}(\epsilon)$, we have

$$p^- = \lim_{\epsilon \to 0^+} \bar{p}(\epsilon) \ .$$

We first show that for all $\delta > 0$, the inequality $\hat{p}(\delta) \leq \bar{p}(\epsilon)$ holds for some well chosen value of ϵ. Let y a feasible solution for problem (LPG$_\delta$). Using the definition of g_k, constraints $g_k(y) \leq e^\delta$ easily give

$$\sum_{i \in I_k} e^{a_i^T y - c_i - \delta} \leq 1 \ ,$$

which shows that the following choice of vectors s and v

$$s_i = c_i - a_i^T y + \delta \text{ for all } i \in I \text{ and } v_k = 1 \text{ for all } k \in R$$

will be feasible for problem (CPG$_\epsilon$) with $\epsilon = \delta\sqrt{n}$, since we have then $(s_{I_k}, v_k) \in \mathcal{G}^{n_k} \forall k \in R$ and

$$\left\| \begin{pmatrix} A^T \\ 0 \end{pmatrix} y + \begin{pmatrix} s \\ v \end{pmatrix} - \begin{pmatrix} c \\ e \end{pmatrix} \right\| = \left\| \begin{pmatrix} \delta \\ 0 \end{pmatrix} \right\| = \delta\sqrt{n} \ .$$

Since every feasible solution y for (LPG$_\delta$) gives a feasible solution (y,s,v) for (CPG$_\epsilon$) with the same objective value, the latter problem cannot have a smaller optimum objective value and we have $\hat{p}(\delta) \leq \bar{p}(\delta\sqrt{n})$. Taking the limit when $\delta \to 0$, this shows that $\hat{p} \leq p^-$.

Let us now work in the opposite direction and let (y,s,v) a feasible solution to problem (CPG$_\epsilon$). We have thus

$$\sum_{i \in I_k} e^{-\frac{s_i}{v_k}} \leq 1 \text{ for all } k \in R \text{ and } \left\| \begin{pmatrix} A^T y + s - c \\ v - e \end{pmatrix} \right\| < \epsilon \; ,$$

which implies

$$\begin{cases} \left|a_i^T y + s_i - c_i\right| < \epsilon \text{ for all } i \in I \\ |v_k - 1| < \epsilon \text{ for all } k \in R \end{cases} .$$

We write

$$1 \geq \sum_{i \in I_k} e^{-\frac{s_i}{v_i}} > \sum_{i \in I_k} e^{-\frac{c_i - a_i^T y + \epsilon}{1-\epsilon}} \; ,$$

since $v_k > 1 - \epsilon$, $s_i < c_i - a_i^T y + \epsilon$ and $x \mapsto e^{-x}$ is a monotonic decreasing function. Defining $\tilde{y} = \frac{y}{1-\epsilon}$, we have

$$\begin{aligned} \frac{c_i - a_i^T y + \epsilon}{1-\epsilon} &= c_i - a_i^T \tilde{y} + \frac{c_i + \epsilon}{1-\epsilon} - c_i \\ &= c_i - a_i^T \tilde{y} + \frac{\epsilon}{1-\epsilon}(c_i + 1) \\ &\leq c_i - a_i^T \tilde{y} + \frac{\epsilon}{1-\epsilon}(\max c_i + 1) \\ &\leq c_i - a_i^T \tilde{y} + \frac{C\epsilon}{1-\epsilon} \; , \end{aligned}$$

where $C = \max c_i + 1$. We have thus

$$1 > \sum_{i \in I_k} e^{-\frac{c_i - a_i^T y + \epsilon}{1-\epsilon}} \geq \sum_{i \in I_k} e^{a_i^T \tilde{y} - c_i - \frac{C\epsilon}{1-\epsilon}} = e^{-\frac{C\epsilon}{1-\epsilon}} \sum_{i \in I_k} e^{a_i^T \tilde{y} - c_i} \; ,$$

which shows that

$$\sum_{i \in I_k} e^{a_i^T \tilde{y} - c_i} < e^{\frac{C\epsilon}{1-\epsilon}} \; ,$$

i.e. $\tilde{y}$ is a feasible solution to problem (LPG$_\delta$) with $\delta = \frac{C\epsilon}{1-\epsilon}$. Since this solution has an objective value $b^T\tilde{y}$ equal to $b^T y$ divided by $1-\epsilon$, this means that $\bar{p}(\epsilon) \leq (1-\epsilon)\hat{p}(\frac{C\epsilon}{1-\epsilon})$. Taking the limit when $\epsilon \to 0$, this shows that $p^- \leq \hat{p}$, and we may conclude that $p^- = \hat{p}$, as announced. □

4.3 Refined duality

The properties we have proved so far about our pair of primal-dual geometric optimization problems (LPG)–(LDG) are merely more or less direct consequences of their convex nature, hence valid for all convex optimization problems. In this section, we are going further and prove a result that does not hold in the general convex case, namely we show that our pair of primal-dual problems cannot have a strictly positive duality gap.

Theorem 4.5 *If both problems* (LPG) *and* (LDG) *are feasible, their optimum objective values are equal (but not necessarily attained).*

Proof (the original proof can be found in [4, §7]). In Theorem 4.3, we proved the existence of a zero duality gap using some assumption on the dual, namely the existence of a strictly positive feasible vector. What we are going to show here is that if such a point does not exist, i.e. one or more components of vector x are zero for all feasible dual solutions, our primal-dual pair can be reduced to an equivalent pair of problems where these components have been removed, in other words a primal-dual pair with a strictly positive feasible dual solution and a zero duality gap.

In order to use this strategy, we start by identifying the components of x that are identically equal to zero on the dual feasible region. This can be done with the following linear optimization problem:

$$\min 0 \quad \text{s.t.} \quad Ax = b \text{ and } x \geq 0 \ . \tag{PLP}$$

Since this problem has a zero objective function, all feasible solutions are optimal and we deduce that if a variable x_i is zero for all feasible solutions to problem (LDG), it is zero for all optimal solution to problem (PLP). We are going to use the well-known Goldman-Tucker theorem [3].

Theorem 4.6 (Goldman-Tucker) *Let us consider the following primal-dual pair of linear optimization problems:*

$$\min c^T x \quad \textit{s.t.} \quad Ax = b \textit{ and } x \geq 0 \quad \textit{and} \quad \max b^T y \quad \textit{s.t.} \quad A^T y + s = c \textit{ and } s \geq 0 \ .$$

If both problems are feasible, there exists a unique partition $(\mathcal{B}, \mathcal{N})$ *of the index set common to vectors* x *and* s *such that*

- ◇ *every optimal solution* x^* *to the primal problem satisfies* $x^*_{\mathcal{N}} = 0$.
- ◇ *every optimal solution* (y^*, s^*) *to the dual problem satisfies* $s^*_{\mathcal{B}} = 0$.

This partition is called the optimal partition. *Moreover, there exists at least an optimal primal-dual solution* (x^*, y^*, s^*) *such that* $x^* + s^* > 0$*, hence satisfying* $x^*_{\mathcal{B}} > 0$ *and* $s^*_{\mathcal{N}} > 0$. *Such a pair is called a* strictly complementary *pair.*

Writing the dual of problem (PLP)

$$\max b^T y \quad \text{s.t.} \quad A^T y + s = 0 \text{ and } s \geq 0 \ , \tag{DLP}$$

we find that both (PLP) and (DLP) are feasible (the former because (LDG) is feasible, the latter because $(y, s) = (0, 0)$ is always a feasible solution), and thus that the Goldman-Tucker theorem is applicable.

Having now the optimal partition $(\mathcal{B}, \mathcal{N})$ at hand, we observe that the index set $\mathcal{N}$ defines exactly the set of variables x_i that are identically zero on the feasible region of problem (LDG). We are now able to introduce a reduced primal-dual pair of geometric optimization problems, where variables x_i with $i \in \mathcal{N}$ have been removed. We start with the dual problem

$$\inf c_{\mathcal{B}}^T x_{\mathcal{B}} + \sum_{k \in R} \sum_{i \in I_k \cap \mathcal{B} | x_i > 0} x_i \log \frac{x_i}{\sum_{i \in I_k \cap \mathcal{B}} x_i} \quad \text{s.t.} \quad A_{\mathcal{B}} x_{\mathcal{B}} = b \text{ and } x_{\mathcal{B}} \geq 0 \ . \tag{RDG}$$

It is straightforward to check that this problem is completely equivalent to problem (LDG), since the variables we removed had no contribution to the objective or to the constraints in (LDG). Indeed, there is a one-to-one correspondence preserving objective values between feasible solutions $x_{\mathcal{B}}$ for (RDG) and feasible solutions x for (LDG), the latter satisfying always $x_{\mathcal{N}} = 0$. Our primal geometric optimization problem becomes

$$\sup b^T y \quad \text{s.t.} \quad g_k^{\mathcal{B}}(y) \leq 1 \text{ for all } k \in R \ , \tag{RPG}$$

where functions $g_k^{\mathcal{B}}$ are now defined over the sets $I_k \cap \mathcal{B}$, i.e.

$$g_k^{\mathcal{B}} : \mathbb{R}^m \mapsto \mathbb{R}_{++} : y \mapsto \sum_{i \in I_k \cap \mathcal{B}} e^{a_i^T y - c_i} \ .$$

Since the Goldman-Tucker theorem implies the existence of a feasible vector x^* such that $x^*_{\mathcal{B}} > 0$ and $x^*_{\mathcal{N}} = 0$, we find that $x^*_{\mathcal{B}}$ is a strictly positive feasible solution to (RDG), which allows us to apply Theorem 4.3. Knowing that (LPG) is feasible, problems (LDG) and (RDG) must be bounded: we are in the second case of the theorem and can conclude that problem (RPG) attains an optimum objective value equal to the optimum objective value of problem (RDG). The last thing we have to show in order to finish our proof is that the optimum values of primal problem (LPG) and its reduced version (RPG) are equal.

Let us start with $\bar{y}$, one of the optimal solutions to (RPG) that are known to exist. Our goal is thus to prove that problem (LPG) has an optimum objective value equal to $b^T \bar{y}$. Unfortunately, $\bar{y}$ is not always feasible for problem (LPG), since the additional terms in g_k corresponding to indices $i \in \mathcal{N}$ result in $g_k(\bar{y}) > g_k^{\mathcal{B}}(\bar{y})$ and possibly $g_k(\bar{y}) > 1$.

To solve this problem, we are going to perturb $\bar{y}$ with a suitably chosen vector, in order to make it feasible. The existence of this perturbation vector will be again derived from the Goldman-Tucker theorem, in the following manner. Let (x^*, y^*, s^*) a strictly complementary pair for problems (PLP)–(DLP). Since the optimum primal objective value is obviously equal to zero, we also have that the optimum dual objective $b^T y^*$ is equal to zero. Moreover, we have that $A^T y^* + s^* = 0$, which gives

$$A_{\mathcal{B}}^T y^* = -s^*_{\mathcal{B}} = 0 \text{ and } A_{\mathcal{N}}^T y^* = -s^*_{\mathcal{N}} < 0 \ .$$

Considering a vector y defined by $y = \bar{y} + \lambda y^*$, where λ is a positive parameter that is going to tend to $+\infty$, it is easy to check that

$$\begin{aligned} g_k(y) &= g_k^{\mathcal{B}}(y) + g_k^{\mathcal{N}}(y) \\ &= \sum_{i \in I_k \cap \mathcal{B}} e^{a_i^T y - c_i} + \sum_{i \in I_k \cap \mathcal{N}} e^{a_i^T y - c_i} \\ &= \sum_{i \in I_k \cap \mathcal{B}} e^{a_i^T \bar{y} + \lambda a_i^T y^* - c_i} + \sum_{i \in I_k \cap \mathcal{N}} e^{a_i^T \bar{y} + \lambda a_i^T y^* - c_i} \\ &= \sum_{i \in I_k \cap \mathcal{B}} e^{a_i^T \bar{y} - c_i} + \sum_{i \in I_k \cap \mathcal{N}} e^{a_i^T \bar{y} - c_i - \lambda s_i^*} \\ &= g_k^{\mathcal{B}}(\bar{y}) + \sum_{i \in I_k \cap \mathcal{N}} e^{a_i^T \bar{y} - c_i - \lambda s_i^*} \ , \end{aligned}$$

which means that

$$\lim_{\lambda \to +\infty} g_k(y) = g_k^{\mathcal{B}}(\bar{y}) \leq 1 \text{ for all } k \in R \ ,$$

since $s_i^* > 0$ for all $i \in \mathcal{N}$ implies that all the exponents in the second sum are tending to $-\infty$. Moreover, the objective value $b^T y$ is equal to $b^T \bar{y} + \lambda b^T y^* = b^T \bar{y}$ for all values of λ, since $b^T y^* = 0$. However, this vector y is not necessarily feasible for problem (LPG) (we may have $g_k^{\mathcal{B}}(\bar{y}) = 1$ for some k and thus $g_k(y) > 1$ for all λ), and cannot therefore help us in proving that its optimum objective value is equal to $b^T \bar{y}$. We have to use a second trick, namely to "mix" y with a feasible solution to make it feasible.

Let y^0 a feasible solution to problem (LPG). We know thus that

$$g_k(y^0) = g_k^{\mathcal{B}}(y^0) + g_k^{\mathcal{N}}(y^0) \leq 1 ,$$

which implies

$$g_k^{\mathcal{B}}(y^0) < 1$$

since $g_k^{\mathcal{N}}(y^0)$ is strictly positive. Considering now the vector $y = \delta y^0 + (1-\delta)\bar{y} + \lambda y^*$, we may write

$$\begin{aligned} g_k(y) &= g_k^{\mathcal{B}}(y) + g_k^{\mathcal{N}}(y) \\ &= g_k^{\mathcal{B}}(\delta y^0 + (1-\delta)\bar{y} + \lambda y^*) + g_k^{\mathcal{N}}(\delta y^0 + (1-\delta)\bar{y} + \lambda y^*) \\ &= g_k^{\mathcal{B}}(\delta y^0 + (1-\delta)\bar{y}) + g_k^{\mathcal{N}}(\delta y^0 + (1-\delta)\bar{y} + \lambda y^*) , \end{aligned}$$

this last line using again the fact that $A_{\mathcal{B}}^T y^* = 0$. We have thus

$$\lim_{\lambda \to +\infty} g_k(y) = g_k^{\mathcal{B}}(\delta y^0 + (1-\delta)\bar{y})$$

for the same reasons as above (exponents in $g_k^{\mathcal{N}}$ tending to $-\infty$). Since we know that functions g_k are convex, we have that

$$g_k^{\mathcal{B}}(\delta y^0 + (1-\delta)\bar{y}) \leq \delta g_k^{\mathcal{B}}(y^0) + (1-\delta) g_k^{\mathcal{B}}(\bar{y}) < \delta + (1-\delta) = 1 ,$$

which finally implies

$$\lim_{\lambda \to +\infty} g_k(y) < 1 .$$

Taking now a sufficiently large value of λ, we can ensure that $g_k(y) < 1$ for all k, i.e. that y is feasible for problem (LPG). The objective value associated to such a solution is equal to

$$b^T y = \delta b^T y^0 + (1-\delta) b^T \bar{y} + \lambda b^T y^* = \delta b^T y^0 + (1-\delta) b^T \bar{y} .$$

Letting finally δ tend to zero, we obtain a sequence of solutions y, feasible for problem (LPG), whose objective values converge to $b^T \bar{y}$, the optimum objective value of the reduced problem (RPG), itself equal to the optimum objective value of the dual problem (LDG). This is enough to prove that the primal-dual pair of problems (LPG)–(LDG) has a zero duality gap. $\square$

We also have the following corollary about the subvalue p^- of problem (LPG).

Corollary 4.1 *When both problems* (LPG) *and* (LDG) *are feasible, the optimum objective value of problem* (LPG) *is equal to its subvalue.*

Proof. Indeed, we have in general $p^* \leq p^- \leq d^*$. Since the last theorem implies $p^* = d^*$, we obtain $p^* = p^-$. $\square$

4.4 Summary and examples

Let us summarize the possible situations about the primal problem (LPG), and give corresponding examples to show that the results obtained so far cannot be sharpened.

- $\diamond$ In the best possible situation, the dual problem has a strictly positive solution and is bounded: our primal problem is guaranteed by Theorem 4.3 to be feasible and have at least one finite optimal solution with a zero duality gap. Taking for example

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} , \; b = \begin{pmatrix} 1 \\ 1 \end{pmatrix} , \; c = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \text{ and } I_1 = \{1, 2\} \ ,$$

our primal-dual pair (CPG)–(CDG) becomes

$$\begin{array}{lll} \sup\ y_1 + y_2 & \text{s.t.} & e^{y_1} + e^{y_2} \le 1 \\ \inf\ 0 + x_1 \log \dfrac{x_1}{x_1 + x_2} + x_2 \log \dfrac{x_2}{x_1 + x_2} & \text{s.t.} & x_1 = 1, x_2 = 1 \text{ and } x \ge 0 \ . \end{array}$$

The only feasible dual solution is strictly positive, giving a bounded optimum objective value $d^* = 2 \log \frac{1}{2} = -2 \log 2$, and we may easily check (using Lemma 3.1) that $y_1 = y_2 = -\log 2$ is the only optimum primal solution, giving also $p^* = -2 \log 2$.

- $\diamond$ In the case of an unbounded dual, the primal problem has to be infeasible because of the weak duality theorem. Choosing

$$A = \begin{pmatrix} 0 & 1 \end{pmatrix} , \; b = 1, \; c = \begin{pmatrix} -1 \\ 0 \end{pmatrix} \text{ and } I_1 = \{1, 2\} \ ,$$

our primal-dual pair becomes

$$\begin{array}{lll} \sup\ y_1 & \text{s.t.} & e^1 + e^{y_1} \le 1 \\ \inf\ -x_1 + x_1 \log \dfrac{x_1}{x_1 + x_2} + x_2 \log \dfrac{x_2}{x_1 + x_2} & \text{s.t.} & x_2 = 1 \text{ and } x \ge 0 \ . \end{array}$$

The dual is unbounded: the feasible solution $x = (\lambda, 1)$ for all $\lambda > 0$ has an objective value equal to $-\lambda + \lambda \log \frac{\lambda}{\lambda+1} + \log \frac{1}{\lambda+1}$, which is easily shown to tend to $(-\infty - 1 - \infty) = -\infty$ when $\lambda \to +\infty$. The primal problem is obviously infeasible, as expected.

- $\diamond$ When both the primal and the dual problems are feasible but the dual does not have a strictly feasible solution, the duality gap is guaranteed by Theorem 4.5 to be equal to zero with a finite common optimal objective value, but not necessarily with attainment. Adding a third variable to our previous examples

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix} , \; b = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} , \; c = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \text{ and } I_1 = \{1, 2, 3\} \ ,$$

our primal-dual pair becomes

$$\begin{array}{lll} \sup\ y_1 + y_2 & \text{s.t.} & e^{y_1} + e^{y_2} + e^{y_2 + 2y_3 - 1} \le 1 \\ \inf\ x_3 + \displaystyle\sum_{i | x_i > 0} x_i \log \frac{x_i}{\sum_{i=1}^3 x_i} & \text{s.t.} & x_1 = 1, x_2 + x_3 = 1, 2x_3 = 0 \text{ and } x \ge 0 \ . \end{array}$$

The only feasible dual solution $x = (1,1,0)$ has a zero component and gives $d^* = -2\log 2$. It is not too difficult to find a sequence of primal feasible solutions tending to $y = (-\log 2, -\log 2, -\infty)$ that establishes that the supremum of the primal problem is also equal to $p^* = -2\log 2$. However, this value cannot be attained: the primal constraint implies $e^{y_1} + e^{y_2} < 1$, which in turn can be shown to force $y_1 + y_2 < -2\log 2$ using Lemma 3.1.

◇ Our last example will demonstrate the worst situation that can happen: a feasible bounded dual problem with an infeasible primal problem. Taking

$$A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \; b = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \; c = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \text{ and } I_1 = \{1,2\}, \; I_2 = \{3\} ,$$

our primal-dual pair becomes (after some simplifications in the dual objective)

$$\begin{array}{rcl} \sup \; y_1 & \text{s.t.} & e^{y_1 - 1} + e^{y_2} \leq 1 \text{ and } e^{1-y_1} \leq 1 \\ \inf \; x_1 - x_3 + x_1 \log \dfrac{x_1}{x_1 + x_2} & \text{s.t.} & x_1 - x_3 = 1, x_2 = 0 \text{ and } x \geq 0 \,. \end{array}$$

All the feasible dual solution have at least one zero component and it is not difficult to compute that $d^* = 1$ (when $x = (1,0,0)$, for example). It is also easy to check that the primal problem is infeasible: the first constraint implies $e^{y_1-1} < 1$ and thus $y_1 < 1$, while the second constraint forces $y_1 \geq 1$. However, Theorem 4.4 tells us that the primal problem has a subvalue p^- equal to d^*. Indeed, relaxing the primal problem to

$$\sup \; y_1 \quad \text{s.t.} \quad e^{y_1-1} + e^{y_2} \leq e^{\delta} \text{ and } e^{1-y_1} \leq e^{\delta}$$

for any $\delta > 0$, we find $y_1 < 1 + \delta$ and $y_1 \geq 1 - \delta$, implying $1 - \delta \leq \bar{p}(\delta) < 1 + \delta$ and leading to a subvalue p^- equal to 1, as expected.

5 Concluding remarks

5.1 Original formulation

In Subsection 4.1, we presented a conic formulation for the primal-dual pair of geometric optimization problems (LPG)–(LDG) involving linear objective functions, which allowed us to derive several duality theorems. However, the traditional formulation of geometric optimization usually involves a posynomial objective function, as in (PG) or in (TPG), its convexified variant. In this subsection, we show that such problems can be cast as problems with a linear objective, and outline how these duality results can be translated into this traditional formulation.

Let us restate for convenience the convexified problem (TPG)

$$\inf g_0(y) \quad \text{s.t.} \quad g_k(y) \leq 1 \text{ for all } k \in R \setminus \{0\} \; , \tag{TPG}$$

which is readily seen to be equivalent to

$$\inf e^{-y_0} \quad \text{s.t.} \quad g_0(y) \leq e^{-y_0} \text{ and } g_k(y) \leq 1 \text{ for all } k \in R \setminus \{0\} \; ,$$

introducing a new variable y_0 to express the posynomial objective. Noticing that minimizing e^{-y_0} amounts to maximizing y_0, we can rewrite this last problem as

$$\sup y_0 \quad \text{s.t.} \quad e^{y_0} g_0(y) \leq 1 \text{ and } g_k(y) \leq 1 \text{ for all } k \in R \setminus \{0\} \ ,$$

which can now be expressed in the format of (LPG) as

$$\sup \tilde{b}^T \tilde{y} \quad \text{s.t.} \quad \tilde{g}_k(\tilde{y}) \leq 1 \text{ for all } k \in R \ ,$$

where vector of variables $\tilde{y} \in \mathbb{R}^{m+1}$, objective vector $\tilde{b} \in \mathbb{R}^{m+1}$ and posynomials $\tilde{g}_k$ are defined by

$$\tilde{y} = \begin{pmatrix} y_0 \\ y \end{pmatrix} , \ \tilde{b} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} , \ \tilde{g}_0(\tilde{y}) = e^{y_0} g_0(y) \text{ and } \tilde{g}_k(\tilde{y}) = g_k(y) \text{ for all } k \in R \setminus \{0\} \ .$$

This last definition of posynomials $\tilde{g}_k$ corresponds to the following choice of column vectors $\tilde{a}_i$ (constants c_i are left unchanged):

$$\tilde{a}_i = \begin{pmatrix} 1 \\ a_i \end{pmatrix} \text{ for all } i \in I_0 \text{ and } \tilde{a}_i = \begin{pmatrix} 0 \\ a_i \end{pmatrix} \text{ for all } i \in I \setminus I_0 \ .$$

It is now easy to find a dual for problem (TPG), based on the known dual for (LPG) and our special choice of $\tilde{a}_i$ and $\tilde{b}$. Defining a matrix $\tilde{A}$ whose columns are the a_i's, i.e.

$$\tilde{A} = \begin{pmatrix} \overbrace{1, \ldots, 1}^{I_0} & \overbrace{0, \ldots, 0}^{I_k \ \forall k \neq 0} \\ \multicolumn{2}{c}{A} \end{pmatrix} \ ,$$

we find the dual problem

$$\inf c^T x + \sum_{k \in R} \sum_{i \in I_k | x_i > 0} x_i \log \frac{x_i}{\sum_{i \in I_k} x_i} \quad \text{s.t.} \quad \tilde{A} x = \tilde{b} \text{ and } x \geq 0$$

or, equivalently,

$$\inf c^T x + \sum_{k \in R} \sum_{i \in I_k | x_i > 0} x_i \log \frac{x_i}{\sum_{i \in I_k} x_i} \quad \text{s.t.} \quad A x = 0 \ , \sum_{i \in I_0} x_i = 1 \text{ and } x \geq 0 \ .$$

We can manipulate further the second part of the objective function

$$\begin{aligned} \sum_{k \in R} \sum_{i \in I_k | x_i > 0} x_i \log \frac{x_i}{\sum_{i \in I_k} x_i} &= \sum_{k \in R} \sum_{i \in I_k | x_i > 0} \Big(x_i \log x_i - x_i \log \sum_{i \in I_k} x_i \Big) \\ &= \sum_{i \in I} x_i \log x_i - \sum_{k \in R} \Big(\sum_{i \in I_k} x_i \Big) \log \Big(\sum_{i \in I_k} x_i \Big) \ , \end{aligned}$$

with the convention that $0 \log 0 = 0$, and find

$$\inf c^T x + \sum_{i \in I} x_i \log x_i - \sum_{k \in R \setminus \{0\}} \Big(\sum_{i \in I_k} x_i \Big) \log \Big(\sum_{i \in I_k} x_i \Big) \quad \text{s.t.} \quad A x = 0 \ , \sum_{i \in I_0} x_i = 1 \text{ and } x \geq 0$$

(we could remove the term for $k = 0$ in the second sum because of the linear constraint $\sum_{i \in I_0} x_i = 1$). Noting finally that the objective of (TPG) is actually e^{-y_0} and not y_0, we find after some easy transformations the final dual problem (using $c_i = -\log C_i$)

$$\sup \prod_{i \in I} \left(\frac{C_i}{x_i} \right)^{x_i} \prod_{k \in R \setminus \{0\}} \left(\sum_{i \in I_k} x_i \right)^{\sum_{i \in I_k} x_i} \quad \text{s.t.} \quad Ax = 0 \,, \sum_{i \in I_0} x_i = 1 \text{ and } x \geq 0 \,. \qquad \text{(TDG)}$$

This dual problem is identical to the usual formulation that can be found in the literature [2, Chapter III]. To close this discussion, we give a few hints on how to establish links between the classical theory elaborated in [2] and the results presented in Subsections 4.2 and 4.3.

The *main lemma* in [2, Chapter IV] is essentially our weak duality theorem with its associated set of orthogonality conditions. The *first and second duality theorems* from [2, Chapter III] are basically coming from Theorems 4.2 and 4.3, i.e. the application of the strong duality theorem to the primal and the dual problems (note that the hypotheses of the first duality theorem suppose primal attainment while our version only requires primal boundedness, which is a weaker condition). We also note that the notion of *subinfimum* in [2, Chapter VI] for the primal problem is equivalent to our concept of *subvalue*. Finally, the *strong duality* theorems in [2, Chapter VI] are closely related to our Theorem 4.5, stating that a nonzero duality gap cannot occur ; the notion of *canonical* problem that is heavily used in the associated proofs corresponds to the case $\mathcal{N} = \emptyset$ in the optimal partition of problem (PLP), i.e. existence of a strictly feasible dual solution.

5.2 Conclusion

In this article, we have shown how to use the duality theory of conic optimization to derive results about geometric optimization. This process involved the introduction of a dedicated pair of convex cones $\mathcal{G}^n$ and $(\mathcal{G}^n)^*$. We would like to point out that conic optimization had so far been mostly applied to self-dual cones, i.e. to linear, second-order cone and semidefinite optimization. We hope to have demonstrated here that this theory can be equally useful in the case of a less symmetric duality.

The results we obtained can be classified into two distinct categories: most of them are direct consequences of the convex nature of geometric optimization (weak and strong duality theorems), while some of them are specific to this class of problems (absence of a duality gap). The set of problems we studied differed in fact slightly from the classical formulation of geometric optimization, because of the linear objective function.

We would like to point out that this variation in the formulation was necessary since conic optimization cannot be applied directly to geometric optimization problems cast in the traditional form. Indeed, problem (PG) is not convex, which already prevents us from applying Lagrange duality, while the pair of problems (TPG)–(TDG) does not feature a linear objectives and hence is not suitable for a conic formulation. However, extension of our results to the case of a posynomial objective function is straightforward, as outlined in Subsection 5.1. We also consider the results associated to our formulation more natural than their traditional counterparts (looking for example at the structure of the linear constraints in the dual problem (LDG), which does feature the *normalizing* constraints present in (TDG)).

The proofs presented in this article possess in our opinion several advantages over the classical ones: in addition to being shorter, they allow us to confine the specificity of the class of problems under study to the convex cones used in the formulation. Moreover, the reason why geometric optimization has better duality properties than a general conic problem becomes clear: this is essentially due to the existence of a strictly feasible dual solution. Indeed, even if such an interior solution does not always exist, a regularization procedure involving an equivalent reduced problem can always be carried out and allows us to prove the absence of a duality gap in all cases (we note however that the property of primal attainment, satisfied when there exists a strictly feasible dual solution, is lost in this process and is thus no longer valid in the general case).

A last advantage of our conic formulation is that it allows us to benefit with minimal work from the theory of polynomial interior-point methods for convex optimization developed in [6]. Indeed, finding a computable self-concordant barrier for our geometric cone $\mathcal{G}^n$, a topic we leave for further research, is all what is needed to build an algorithm able to solve a geometric optimization problem up to a given accuracy within a polynomial number of arithmetic operations.

Acknowledgements. The author wishes to thank Professor Tamás Terlaky for the discussions they had on this topic and for providing reference [5]. He would also like to thank Professor Kees Roos and the SSOR department at the Delft University of Technology for their kind hospitality and an anonymous referee for providing reference [11].

References

[1] D. den Hertog, F. Jarre, C. Roos, and T. Terlaky, *A sufficient condition for self-concordance, with application to some classes of structured convex programming problems*, Mathematical Programming **69** (1995), 75–88.

[2] R. J. Duffin, E. L. Peterson, and C. Zener, *Geometric programming*, John Wiley & Sons, New York, 1967.

[3] A. J. Goldman and A. W. Tucker, *Theory of linear programming*, Linear Equalities and Related Systems (H. W. Kuhn and A. W. Tucker, eds.), Annals of Mathematical Studies, vol. 38, Princeton University Press, Princeton, New Jersey, 1956, pp. 53–97.

[4] E. Klafszky, *Geometric programming and some applications*, Ph.D. thesis, Tanulmányok, No. 8, 1974.

[5] ———, *Geometric programming*, Seminar Notes, no. 11.976, Hungarian Committee for Systems Analysis, Budapest, 1976.

[6] Y. E. Nesterov and A. S. Nemirovsky, *Interior-point polynomial methods in convex programming*, SIAM Studies in Applied Mathematics, SIAM Publications, Philadelphia, 1994.

[7] R. T. Rockafellar, *Convex analysis*, Princeton University Press, Princeton, N. J., 1970.

[8] ———, *Some convex programs whose duals are linearly constrained*, Non-linear Programming (J. B. Rosen, ed.), Academic Press, 1970.

[9] J. Stoer and Ch. Witzgall, *Convexity and optimization in finite dimensions I*, Springer Verlag, Berlin, 1970.

[10] J. F. Sturm, *Primal-dual interior-point approach to semidefinite programming*, Ph.D. thesis, Erasmus Universiteit Rotterdam, The Netherlands, 1997, published in [11].

[11] ———, *Duality results*, High Performance Optimization (H. Frenk, C. Roos, T. Terlaky, and S. Zhang, eds.), Kluwer Academic Publishers, 2000, pp. 21–60.