Approximating geometric optimization with l_p -norm optimization

IMAGE TECHNICAL REPORT 0008

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

October 2000

^{*}The author is supported by a grant from the F.N.R.S. (Belgian National Fund for Scientific Research).

Abstract

In this article, we demonstrate how to approximate geometric optimization with l_p norm optimization. These two categories of problems are well known in structured convex optimization. We describe a family of l_p -norm optimization problems that can be made arbitrarily close to a geometric optimization problem, and show that the dual problems for these approximations are also approximating the dual geometric optimization problem. Finally, we use these approximations and the duality theory for l_p -norm optimization to derive simple proofs of the weak and strong duality theorems for geometric optimization.

Keywords. geometric optimization, l_p -norm optimization, approximation, duality.

Contents

1	Introduction	3
2	Approximating geometric optimization2.1An approximation of the exponential function2.2An approximation using l_p -norm optimization	4 4 5
3	Deriving duality properties3.1Duality for l_p -norm optimization3.2A dual for the approximate problem3.3Duality for geometric optimization	6 6 8 9
4	Concluding remarks	10

1 Introduction

Let us start by introducing the primal l_p -norm optimization problem [PE70a, Ter85], which is basically a slight modification of a linear optimization problem where the use of l_p -norms applied to linear terms is allowed within the constraints. In order to state its formulation in the most general setting, we need to introduce the following sets: let $K = \{1, 2, \ldots, r\}$, $I = \{1, 2, \ldots, n\}$ and let $\{I_k\}_{k \in K}$ be a partition of I into r classes. The problem data is given by two matrices $A \in \mathbb{R}^{m \times n}$ and $F \in \mathbb{R}^{m \times r}$ (whose columns will be denoted by $a_i, i \in I$ and $f_k, k \in K$) and four column vectors $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$, $d \in \mathbb{R}^r$ and $p \in \mathbb{R}^n$ such that $p_i > 1 \ \forall i \in I$. Our primal problem consists in optimizing a linear function of a column vector $y \in \mathbb{R}^m$ under a set of constraints involving l_p -norms of linear forms, and can be written as

$$\sup b^T y \quad \text{s.t.} \quad \sum_{i \in I_k} \frac{1}{p_i} \left| c_i - a_i^T y \right|^{p_i} \le d_k - f_k^T y \quad \forall k \in K \;. \tag{Pl_p}$$

It is not too difficult to see that this problem is a generalization of several well-known convex optimization problems: linear optimization corresponds to the case where n = 0, while linearly and quadratically constrained convex quadratic problems can be modelled using a reformulation of the quadratic terms as 2-norms [Ter85]. Problems of approximation in l_p -norm can also be easily modelled, such as the problems described in [NN94, Ter85].

The purpose of this note is to show that this category of problems can be used to approximate another famous class of problems known as geometric optimization [DPZ67]. Using the same notations as above for sets K and I_k , $k \in K$, matrix A and vectors b, c and a_i , $i \in I$, this category of problems can be stated as

$$\sup b^T y \quad \text{s.t.} \quad \sum_{i \in I_k} e^{a_i^T y - c_i} \le 1 \ \forall k \in K \ . \tag{PG}$$

This formulation differs slightly from the traditional formulation of geometric optimization with respect to the following two aspects:

♦ The vector of variables that is traditionally considered is not $y \in \mathbb{R}^m$ but rather the strictly positive vector $t \in \mathbb{R}^m_{++}$ defined by $t_j = e^{y_j}, \forall j \in \{1, 2, ..., m\}$. With this convention, the constraints in (PG) can be written as

$$\sum_{i \in I_k} C_i \prod_{j=1}^m t_j^{a_{ij}} \le 1$$

(with $C_j = e^{-c_i} \forall i \in I$), where the expression on the left-hand side is called a *posynomial* (this name comes from the fact that it bears some resemblance with a polynomial whose coefficients are positive, since $C_j > 0$). However, formulation (PG), while being completely equivalent, presents the additional advantage of being convex and is thus much more suitable for analysis (notably from the point of view of duality).

 The traditional formulation deals with a posynomial objective. However, it can be replaced by a linear objective without any loss of generality, as shown in [Kla74, Kla76, Gli99]. We will start by presenting in Section 2 an approximation of the exponential function, which central in the definition of the constraints of a geometric optimization problem. This will allow us to present a family of l_p -norm optimization problems which can be made arbitrarily close to a primal geometric optimization problem. We derive in Section 3 a dual problem for this approximation, and show that the limiting case for these dual approximations is equivalent to the traditional dual geometric optimization problem. Using this family of pairs of primal-dual problems and the weak and strong duality theorems for l_p -norm optimization, we will then show how to derive the corresponding theorems for geometric optimization in a simple manner. Section 4 will conclude and present some topics for further research.

2 Approximating geometric optimization

In this section, we will show how geometric optimization problems can be approximated with l_p -norm optimization.

2.1 An approximation of the exponential function

A key ingredient in our approach is the function that will be used to approximate the exponential terms that arise within the constraints of (PG). Let $\alpha \in \mathbb{R}_{++}$ and let us define

$$g_{\alpha}: \mathbb{R}_{+} \mapsto \mathbb{R}_{+}: x \mapsto \left|1 - \frac{x}{\alpha}\right|^{\alpha}$$
.

We have the following lemma relating $g_{\alpha}(x)$ to e^{-x} :

Lemma 2.1 For any fixed $x \in \mathbb{R}_+$, we have that

$$g_{\alpha}(x) \le e^{-x} \quad \forall \alpha \ge x \quad and \quad e^{-x} < g_{\alpha}(x) + \alpha^{-1} \quad \forall \alpha > 0 ,$$
 (2.1)

equality occurring in the first inequality if and only if x = 0. Moreover, we have

$$\lim_{\alpha \to +\infty} g_{\alpha}(x) = e^{-x} \; .$$

Proof. Let us fix $x \in \mathbb{R}_+$. When $1 < \alpha < x$, we only have to prove the second inequality in (2.1), which is straightforward: we have $e^{-x} < e^{-\alpha} < \alpha^{-1} < g_{\alpha}(x) + \alpha^{-1}$, where we used the obvious inequalities $e^{\alpha} > \alpha$ and $g_{\alpha}(x) > 0$. Assuming $\alpha \ge x$ for the rest of this proof, we define the auxiliary function $h : \mathbb{R}_{++} \to \mathbb{R} : \alpha \to \log f_{\alpha}(x)$. Using the Taylor expansion of $\log(1-x)$ around x = 0

$$\log(1-x) = -\sum_{i=1}^{\infty} \frac{x^i}{i} \quad \text{for all } x \text{ such that } |x| \le 1$$
(2.2)

we have

$$h(\alpha) = \alpha \log \left| 1 - \frac{x}{\alpha} \right| = \alpha \log \left(1 - \frac{x}{\alpha} \right) = -\sum_{i=1}^{\infty} \frac{x^i}{i\alpha^{i-1}} = -x - \sum_{i=2}^{\infty} \frac{x^i}{i\alpha^{i-1}}$$
(2.3)

(where we used the fact that $\frac{x}{\alpha} \leq 1$ to write the Taylor expansion). It is now clear that $h(\alpha) \leq -x$, with equality if and only if x = 0, which in turn implies that $g_{\alpha}(x) \leq e^{-x}$, with equality if and only if x = 0, which is the first inequality in (2.1).

The second inequality is equivalent, after multiplication by e^x , to

$$1 < e^x g_\alpha(x) + e^x \alpha^{-1} \Leftrightarrow 1 - e^x \alpha^{-1} < e^x e^{h_\alpha(x)} \Leftrightarrow 1 - e^x \alpha^{-1} < e^{x + h_\alpha(x)}.$$

This last inequality trivially holds when its left-hand side is negative, i.e. when $\alpha \leq e^x$. When $\alpha > e^x$, we take the logarithm of both sides, use again the Taylor expansion (2.2) and the expression for $h_{\alpha}(x)$ in (2.3) to find

$$\log(1 - e^x \alpha^{-1}) < x + h_\alpha(x) \Leftrightarrow -\sum_{i=1}^\infty \frac{e^{xi}}{i\alpha^i} < -\sum_{i=2}^\infty \frac{x^i}{i\alpha^{i-1}} \Leftrightarrow 0 < \sum_{i=1}^\infty \frac{1}{\alpha^i} \left(\frac{e^{xi}}{i} - \frac{x^{i+1}}{i+1}\right) .$$

This last inequality holds since each of the coefficients between parentheses can be shown to be strictly positive: writing the well-known inequality $e^a > \frac{a^n}{n!}$ for a = xi and n = i + 1, we find

$$e^{xi} > \frac{(xi)^{i+1}}{(i+1)!} \Leftrightarrow \frac{e^{xi}}{i} > \frac{x^{i+1}}{(i+1)} \ \frac{i^i}{i!} \Rightarrow \frac{e^{xi}}{i} > \frac{x^{i+1}}{i+1} \Leftrightarrow \frac{e^{xi}}{i} - \frac{x^{i+1}}{i+1} > 0$$

(where we used $i^i \ge i!$ to derive the third inequality).

To conclude this proof, we note that (2.3) implies that $\lim_{\alpha \to +\infty} h(\alpha) = -x$, which gives $\lim_{\alpha \to +\infty} f_{\alpha}(x) = e^{-x}$, as announced. This last property can also be easily derived from the two inequalities in (2.1).

The first inequality in (2.1) and the limit of $g_{\alpha}(x)$ are well-known, and are sometimes used as definition for the real exponential function, while the second inequality in (2.1) is much less common.

2.2 An approximation using l_p -norm optimization

The formulation of the primal geometric optimization problem (PG) relies heavily on the exponential function. Since Lemma 2.1 shows that it is possible to approximate e^{-x} with increasing accuracy using the function g_{α} , we can consider using this function to formulate an approximation of problem (PG). The key observation we make here is that this approximation can be expressed as a l_p -norm optimization problem.

Indeed, let us fix $\alpha \in \mathbb{R}_{++}$ and write the approximate problem

$$\sup b^T y \quad \text{s.t.} \quad \sum_{i \in I_k} \left(g_\alpha (c_i - a_i^T y) + \alpha^{-1} \right) \le 1 \ \forall k \in K \ . \tag{PG}_\alpha)$$

We note that this problem is a restriction of the original problem (PG), i.e. that any y that is feasible for (PG_{α}) is also feasible for (PG), with the same objective value. This is indeed a direct consequence of the second inequality in (2.1), which implies for any y feasible for (PG_{α})

$$\sum_{i \in I_k} e^{c_i - a_i^T y} < \sum_{i \in I_k} (g_\alpha(c_i - a_i^T y) + \alpha^{-1}) \le 1.$$

We need now to transform the expressions $g_{\alpha}(c_i - a_i^T y) + \alpha^{-1}$ to fit the format of the constraints of a l_p -norm optimization problem. Assuming that $\alpha > 1$ for the rest of this

paper, we write

$$\sum_{i \in I_k} \left(g_\alpha(c_i - a_i^T y) + \alpha^{-1} \right) \le 1 \quad \Leftrightarrow \quad \sum_{i \in I_k} g_\alpha(c_i - a_i^T y) \le 1 - n_k \alpha^{-1}$$
$$\Leftrightarrow \quad \sum_{i \in I_k} \left| 1 - \frac{c_i - a_i^T y}{\alpha} \right|^\alpha \le 1 - n_k \alpha^{-1}$$
$$\Leftrightarrow \quad \sum_{i \in I_k} \left| \alpha - c_i + a_i^T y \right|^\alpha \le \alpha^\alpha (1 - n_k \alpha^{-1})$$
$$\Leftrightarrow \quad \sum_{i \in I_k} \frac{1}{\alpha} \left| c_i - \alpha - a_i^T y \right|^\alpha \le \alpha^{\alpha - 1} (1 - n_k \alpha^{-1})$$

(where n_k is the number of elements in I_k), which allows us to write (PG_{α}) as

$$\sup b^T y \quad \text{s.t.} \quad \sum_{i \in I_k} \frac{1}{\alpha} \left| c_i - \alpha - a_i^T y \right|^{\alpha} \le \alpha^{\alpha - 1} (1 - n_k \alpha^{-1}) \ \forall k \in K \ . \tag{PG}_{\alpha}'$$

This is indeed a l_p -norm optimization in the form (Pl_p) : dimensions m, n and r are the same in both problems, sets I, K and I_k are identical, the vector of exponents p satisfies $p_i = \alpha > 1$ for all $i \in I$, matrix A and vector b are the same for both problems while matrix F is equal to zero. The only difference consists in vectors \tilde{c} and d, which satisfy $\tilde{c}_i = c_i - \alpha$ and $d_k = \alpha^{\alpha-1}(1 - n_k\alpha^{-1})$.

We have thus shown how to approximate a geometric optimization problem with a standard l_p -norm optimization problem. Solving this problem for a fixed value of α will give a feasible solution to the original geometric optimization problem. Letting α tend to $+\infty$, the approximations $g_{\alpha}(c_i - a_i^T y)$ will be more and more accurate, and the corresponding feasible regions will approximate the feasible region of (PG) better and better. We can thus expect the optimal solutions of problems (PG'_{\alpha}) to tend to an optimal solution of (PG). Indeed, this is the most common situation, but it does not happen in all the cases, as will be explained in the next section.

3 Deriving duality properties

The purpose of this section is to study the duality properties of our geometric optimization problem and its approximations. Namely, using the duality properties of l_p -norm optimization problems, we will derive the corresponding properties for geometric optimization, using our family of approximate problems.

3.1 Duality for l_p -norm optimization

Defining a vector $q \in \mathbb{R}^n$ such that $\frac{1}{p_i} + \frac{1}{q_i} = 1$ for all $i \in I$, the dual problem for (Pl_p) consists in finding two vectors $x \in \mathbb{R}^n$ and $z \in \mathbb{R}^r$ that maximize a highly nonlinear objective while satisfying some linear equalities and nonnegativity constraints [Ter85, GT00]:

$$\inf \ \psi(x,z) = c^T x + d^T z + \sum_{\substack{k \in K \\ z_k > 0}} z_k \sum_{i \in I_k} \frac{1}{q_i} \left| \frac{x_i}{z_k} \right|^{q_i} \quad \text{s.t.} \quad \left\{ \begin{array}{l} Ax + Fz = b \text{ and } z \ge 0 \ , \\ z_k = 0 \Rightarrow x_i = 0 \ \forall i \in I_k \ . \end{array} \right.$$
(Dl_p)

We note that a special convention has been taken to handle the case when one or more components of z are equal to zero: the associated terms are left out of the first sum (to avoid a zero denominator) and the corresponding components of x have to be equal to zero. When compared with the primal problem (Pl_p) , this problem has a simpler feasible region at the price of a more complicated (but convex) objective.

We have the following duality properties for the pair of problems $(Pl_p)-(Dl_p)$ (these results were first obtained by Peterson and Ecker [PE70a, PE67, PE70b], see also [GT00] for simpler alternative proofs using conic duality) :

Theorem 3.1 (Weak duality) If y is feasible for (Pl_p) and (x, z) is feasible for (Dl_p) , we have $\psi(x, z) \ge b^T y$.

Theorem 3.2 (Strong duality) If both problems (Pl_p) and (Dl_p) are feasible, the primal optimal objective value is attained with a zero duality gap, i.e.

$$p^* = \max b^T y \quad s.t. \quad \sum_{i \in I_k} \frac{1}{p_i} \left| c_i - a_i^T y \right|^{p_i} \le d_k - f_k^T y \quad \forall k \in K$$
$$= \inf \psi(x, z) \quad s.t. \quad \left\{ \begin{array}{l} Ax + Fz = b \ and \ z \ge 0\\ z_k = 0 \Rightarrow x_i = 0 \ \forall i \in I_k \end{array} \right. = d^* .$$

The weak duality property is a rather straightforward consequence of the convexity of the problem, and is shared by all the convex optimization problems. This contrasts with the absence of a duality gap and the guaranteed attainment of the primal optimal objective value, which are due to the specific structure of l_p -norm optimization problems (i.e. they do not hold for a general convex problem, where a strictly positive duality gap and/or non-attainment for both problems may occur [SW70, Stu00]).

We would like to bring the reader's attention to an interesting special case of dual l_p -norm optimization problem. When all p_i 's are equal to p and when matrix F is identically equal to 0, i.e. when there are no pure linear terms in the constraints, problem (Dl_p) becomes

$$\inf \ \psi(x,z) = c^T x + d^T z + \frac{1}{q} \sum_{\substack{k \in K \\ z_k > 0}} z_k^{1-q} \sum_{i \in I_k} |x_i|^q \quad \text{s.t.} \quad \left\{ \begin{array}{c} Ax = b \text{ and } z \ge 0 ,\\ z_k = 0 \Rightarrow x_i = 0 \ \forall i \in I_k . \end{array} \right.$$
(Dl'_p)

This kind of formulation arises in problems of approximation in l_p -norm, see [NN94, Section 6.3.2] and [Ter85, Section 11, page 98].

Since variables z_k do not appear any more in the linear constraints but only in the objective function $\psi(x, z)$, we may try to find a closed form for their optimal value. Looking at one variable z_k at a time and isolating the corresponding terms in the objective, one finds $d_k z_k + \frac{1}{q} z_k^{1-q} \sum_{i \in I_k} |x_i|^q$, whose derivative is equal to $d_k + \frac{1-q}{q} z_k^{-q} \sum_{i \in I_k} |x_i|^q$. One easily sees that this quantity admits a single maximum when

$$z_k = (p d_k)^{-\frac{1}{q}} \|x_{I_k}\|_q$$

(where $\|\cdot\|_p$ corresponds to the usual *p*-norm defined by $\|x\|_p = (\sum_i |x_i|^p)^{\frac{1}{p}}$ and x_{I_k} denotes the vector made of the components of x whose indices belong to I_k), which always satisfies

the nonnegativity constraint in $(\mathrm{D}l'_p)$ and gives after some straightforward computations a value of

$$d_k z_k + \frac{1-q}{q} z_k^{-q} \sum_{i \in I_k} |x_i|^q = \dots = \left(1 + \frac{p}{q}\right) d_k z_k = p \, d_k z_k = (p \, d_k)^{\frac{1}{p}} \|x_{I_k}\|_q$$

for the two corresponding terms in the objective. Our dual problem (Dl'_p) becomes then

inf
$$\psi(x) = c^T x + \sum_{k \in K} (p \, d_k)^{\frac{1}{p}} \|x_{I_k}\|_q$$
 s.t. $Ax = b$, $(\mathrm{D}l_p'')$

a great simplification when compared to (Dl'_p) . One can check that the special treatment for the case $z_k = 0$ is well handled: indeed, $z_k = 0$ happens when $x_{I_k} = 0$, and the implication that is stated in the constraints of (Dl'_p) is thus satisfied.

3.2 A dual for the approximate problem

We are now going to write the dual for the approximate problem (PG'_{α}) . Since we are in the case where F = 0 and all p_i 's are equal to α , we can use the simplified version of the dual problem (Dl''_p) and write

$$\inf \ \psi_{\alpha}(x) = c^{T}x - \alpha e_{n}^{T}x + \sum_{k \in K} \left(\alpha \ \alpha^{\alpha - 1} (1 - n_{k} \alpha^{-1}) \right)^{\frac{1}{\alpha}} \|x_{I_{k}}\|_{\beta} \quad \text{s.t.} \quad Ax = b$$

(where e_n is a notation for the all-one *n*-dimensional column vector and $\beta > 1$ is a constant such that $\frac{1}{\alpha} + \frac{1}{\beta} = 1$), which can be simplified to give

$$\inf \ \psi_{\alpha}(x) = c^{T} x - \alpha e_{n}^{T} x + \alpha \sum_{k \in K} (1 - n_{k} \alpha^{-1})^{\frac{1}{\alpha}} \|x_{I_{k}}\|_{\beta} \quad \text{s.t.} \quad Ax = b .$$
(DG_{\alpha})

We observe that the constraints and thus the feasible region of this problem are independent from α , which only appears in the objective function $\psi_{\alpha}(x)$. Intuitively, since problems (PG'_{α}) become closer and closer to (PG) as α tends to $+\infty$, the corresponding dual problems (PG'_{α}) should approximate the dual of (PG) better and better. It is thus interesting to write down the limiting case for these problems, i.e. find the limit of ψ_{α} when $\alpha \to +\infty$. Looking first at the terms that are related to single set of indices I_k , we write

$$\begin{split} \psi_{k,\alpha}(x) &= c_{I_k}^T x_{I_k} - \alpha e_{n_k}^T x_{I_k} + \alpha (1 - n_k \alpha^{-1})^{\frac{1}{\alpha}} \|x_{I_k}\|_{\beta} \\ &= c_{I_k}^T x_{I_k} - \alpha e_{n_k}^T x_{I_k} + \alpha \|x_{I_k}\|_1 - \alpha \|x_{I_k}\|_1 + \alpha (1 - n_k \alpha^{-1})^{\frac{1}{\alpha}} \|x_{I_k}\|_{\beta} \\ &= c_{I_k}^T x_{I_k} + \alpha \left[\|x_{I_k}\|_1 - e_{n_k}^T x_{I_k} \right] + \alpha \left[(1 - n_k \alpha^{-1})^{\frac{1}{\alpha}} \|x_{I_k}\|_{\beta} - \|x_{I_k}\|_1 \right] \\ &= c_{I_k}^T x_{I_k} + \alpha \left[\|x_{I_k}\|_1 - e_{n_k}^T x_{I_k} \right] + \frac{\beta}{\beta - 1} \left[(1 - n_k \alpha^{-1})^{\frac{1}{\alpha}} \|x_{I_k}\|_{\beta} - \|x_{I_k}\|_1 \right] \end{split}$$

(where we used at the last line the fact that $\alpha = \frac{\beta}{\beta-1}$). When α tends to $+\infty$ (and thus $\beta \to 1$), we have

$$\lim_{\substack{\alpha \to +\infty \\ \beta \to 1}} \psi_{k,\alpha}(x) = c_{I_k}^T x_{I_k} + \lim_{\alpha \to +\infty} \alpha \left[\|x_{I_k}\|_1 - e_{n_k}^T x_{I_k} \right] + \lim_{\substack{\alpha \to +\infty \\ \beta \to 1}} \frac{\beta}{\beta - 1} \left[(1 - n_k \alpha^{-1})^{\frac{1}{\alpha}} \|x_{I_k}\|_\beta - \|x_{I_k}\|_1 \right] \\
= c_{I_k}^T x_{I_k} + \lim_{\alpha \to +\infty} \alpha \left[\|x_{I_k}\|_1 - e_{n_k}^T x_{I_k} \right] + \lim_{\beta \to 1} \frac{\|x_{I_k}\|_\beta - \|x_{I_k}\|_1}{\beta - 1}$$

The last term in this limit is equal to the derivative of the real function $m_k : \beta \mapsto ||x_{I_k}||_{\beta}$ at the point $\beta = 1$. We can check without difficulties that

$$m_k'(\beta) = \frac{\|x_{I_k}\|_{\beta}^{\frac{1}{\beta}-1}}{\beta^2} \left[\beta \sum_{\substack{i \in I_k \\ x_i > 0}} |x_i|^{\beta} \log |x_i| - \|x_{I_k}\|_1 \log \|x_{I_k}\|_1 \right] \Rightarrow m_k'(1) = \sum_{\substack{i \in I_k \\ x_i > 0}} |x_i| \log \frac{|x_i|}{\|x_{I_k}\|_1} \,,$$

which gives

$$\lim_{\substack{\alpha \to +\infty \\ \beta \to 1}} \psi_{k,\alpha}(x) = c_{I_k}^T x_{I_k} + \lim_{\alpha \to +\infty} \alpha \left[\|x_{I_k}\|_1 - e_{n_k}^T x_{I_k} \right] + \sum_{\substack{i \in I_k \\ x_i > 0}} |x_i| \log \frac{|x_i|}{\|x_{I_k}\|_1}$$

It is easy to see that $||x_{I_k}||_1 - e_{n_k}^T x_{I_k} \ge 0$, with equality if and only if $x_{I_k} \ge 0$. This means that the limit of our objective $\psi_{k,\alpha}(x)$ will be $+\infty$ unless $x_{I_k} \ge 0$. An objective equal to $+\infty$ for a minimization problem can be assimilated to an unfeasible problem, which means that the limit of our dual approximations (DG_{α}) admits the hidden constraint $x_{I_k} \ge 0$. Gathering now all terms in the objective, we eventually find the limit of problems (DG_{α}) when $\alpha \to +\infty$ to be

$$\inf \phi(x) = c^T x + \sum_{\substack{k \in R \\ x_i > 0}} \sum_{\substack{i \in I_k \\ x_i > 0}} x_i \log \frac{x_i}{\sum_{i \in I_k} x_i} \quad \text{s.t.} \quad Ax = b \text{ and } x \ge 0 ,$$
(DG)

which is exactly the dual geometric optimization problem one can find in the literature [Kla74, Kla76, Gli99].

3.3 Duality for geometric optimization

Before we start to prove duality results for geometric optimization, we make a technical assumption on problem (PG), whose purpose will become clear further in this section: we will assume that $n_k \geq 2$ for all $k \in K$, i.e. forbid problems where a constraint is defined with a single exponential term. This can be done without any loss of generality, since a constraint of the form $e^{a_i^T y - c_i} \leq 1$ can be equivalently rewritten as $e^{a_i^T y - c_i - \log 2} + e^{a_i^T y - c_i - \log 2} \leq 1$.

Let us now state the weak duality theorem for geometric optimization:

Theorem 3.3 (Weak duality) If y is feasible for (PG) and x is feasible for (DG), we have $\phi(x) \ge b^T y$.

Proof. Our objective is to prove this theorem using our family of approximate problems (PG'_{α}) – (DG_{α}) . We first note that x is feasible for (DG_{α}) for every α , since the only constraints for this family of problems are the linear constraints Ax = b, which are also present in (DG). The situation is a little different on the primal side: the first inequality in (2.1) and feasibility of y for (PG) imply

$$\sum_{i \in I_k} g_{\alpha}(c_i - a_i^T y) \le \sum_{i \in I_k} e^{a_i^T y - c_i} \le 1 ,$$

with equality if and only if $c_i - a_i^T y = 0$ for all $i \in I_k$. But this cannot happen, since we would have $\sum_{i \in I_k} e^{a_i^T y - c_i} = \sum_{i \in I_k} 1 = n_k > 1$, because of our assumption on n_k , which contradicts the feasibility of y. We can conclude that the following strict inequality holds for all $k \in K$:

$$\sum_{i\in I_k}g_lpha(c_i-a_i^Ty)<1\;.$$

Since the set K is finite, this means that there exists a constant M such that for all $\alpha \geq M$,

$$\sum_{i \in I_k} g_{\alpha}(c_i - a_i^T y) \le 1 - n_k \alpha^{-1} \; \forall k \in K \; ,$$

which in turn implies feasibility of y for problems (PG'_{α}) as soon as $\alpha \geq M$. Feasibility of both y and x for their respective problem allows us to apply the weak duality Theorem 3.1 of l_p -norm optimization to our pair of approximate problems $(PG'_{\alpha})-(DG_{\alpha})$, which implies $\psi_{\alpha}(x) \geq b^T y$ for all $\alpha \geq M$. Taking now the limit of $\psi_{\alpha}(x)$ for α tending to $+\infty$, which is finite and equal to $\phi(x)$ since $x \geq 0$, we find that $\phi(x) \geq b^T y$, which is the announced inequality.

The strong duality theorem for geometric optimization is stated below. We note that contrary to the class of l_p -norm optimization problems, attainment cannot be guaranteed for any of the primal and dual optimum objective values.

Theorem 3.4 If both problems (PG) and (DG) are feasible, their optimum objective values p^* and d^* are equal.

Proof. As shown in the proof of the previous theorem, the existence of a feasible solution for (PG) and (DG) implies that problems (PG'_{α}) and (DG_{α}) are both feasible for all α greater than some constant M. Denoting by p^*_{α} (resp. d^*_{α}) the optimal objective value of problem (PG'_{α}) (resp. (DG_{α})), we can thus apply the strong duality Theorem 3.2 of l_p -norm optimization to these pairs of problems to find that $p_{\alpha}^* = d_{\alpha}^*$ for all $\alpha \geq M$. Since all the dual approximate problems $p^*_{\alpha} = d^*_{\alpha}$ share the same feasible region, it is clear that the optimal value corresponding to the limit of the objective ψ_{α} when $\alpha \to +\infty$ is equal to the limit of the optimal objective values d^*_{α} for $\alpha \to +\infty$. Since the problem featuring this limiting objective has been shown to be equivalent to (DG) in Section 3.2 (including the hidden constraint $x \ge 0$, we must have $d^* = \lim_{\alpha \to +\infty} d^*_{\alpha}$. On the other hand, Theorem 3.2 guarantees for each of the problems (PG'_{α}) the existence of an optimal solution y_{α} that satisfies $b^T y_{\alpha} = p_{\alpha}^*$. Since each of these solutions is also a feasible solution for (PG) (since problems (PG'_{α}) are restrictions of (PG), which shares the same objective function, we have that the optimal objective value of (PG) p^* is at least equal to $b^T y_{\alpha}$ for all $\alpha \geq M$, which implies $p^* \geq \lim_{\alpha \to +\infty} b^T y_\alpha = \lim_{\alpha \to +\infty} p^*_\alpha = \lim_{\alpha \to +\infty} d^*_\alpha = d^*$. Combining this last inequality with the easy consequence of the weak duality Theorem 3.3 that states $d^* \geq p^*$, we end up with the announced equality $p^* = d^*$.

The reason why attainment of the primal optimum objective value cannot be guaranteed is that the sequence y_{α} may not have a finite limit point: indeed, it may happen that one or more components of y_{α} tend to infinity as $\alpha \to +\infty$ (see e.g. the third example in [Gli99, Section 4.4]).

4 Concluding remarks

In this paper, we have shown that the important class of geometric optimization problems can be approximated with l_p -norm optimization.

We have indeed described a parameterized family of primal and dual l_p -norm optimization problems, which can be made arbitrarily close to the geometric primal and dual problems. It is worth to note that the primal approximations are restrictions of the original geometric primal problem, sharing the same objective function, while the dual approximations share essentially the same constraints as the original geometric dual problem (except for the nonnegativity constraints) but feature a different objective.

Another possible approach would be to work with relaxations instead of restrictions on the primal side, using the first inequality in (2.1) instead of the second one, leading to the following problem:

$$\sup \ b^T y \quad ext{s.t.} \quad \sum_{i \in I_k} g_lpha(c_i - a_i^T y) \leq 1 \ orall k \in K \ .$$

However, two problems arise in this setting:

- ♦ the first inequality in (2.1) is only valid when $\alpha \ge x$, which means we would have to add a set of explicit linear inequalities $c_i a_i^T y \le \alpha$ to our approximations, which would make them and their dual problems more difficult to handle,
- ◊ following the same line of reasoning as in the proof of Theorem 3.2, we would end up with another family of optimal solutions y_{α} for the approximate problems; however, since all of these problems are relaxations, we would have no guarantee that any of the optimal vectors y_{α} are feasible for the original primal geometric optimization problem, which would prevent us to conclude that the duality gap is equal to zero. This would only show that there is a family of asymptotically feasible solutions with their objective values tending to the objective value of the dual, a fact that is always true in convex optimization (see e.g. the notion of subinfimum in [SW70, Stu00]).

To conclude, we note that our approximate problems belong to a very special subcategory of l_p -norm optimization problem, since they satisfy F = 0. It might be fruitful to investigate which class of generalized geometric optimization problems can be approximated with general l_p -norm optimization problems, a topic we leave for further research.

References

- [DPZ67] R. J. Duffin, E. L. Peterson, and C. Zener, Geometric programming, John Wiley & Sons, New York, 1967.
- [Gli99] Fr. Glineur, Proving strong duality for geometric optimization using a conic formulation, IMAGE Technical Report 9903, Faculté Polytechnique de Mons, Mons, Belgium, October 1999, to appear in Annals of Operations Research.
- [GT00] Fr. Glineur and T. Terlaky, A conic formulation for lp-norm optimization, IMAGE Technical Report 0005, Faculté Polytechnique de Mons, Mons, Belgium, May 2000, submitted to Journal of Optimization Theory and Applications.
- [Kla74] E. Klafszky, Geometric programming and some applications, Ph.D. thesis, Tanulmányok, No. 8, 1974.
- [Kla76] E. Klafszky, Geometric programming, Seminar Notes, no. 11.976, Hungarian Committee for Systems Analysis, Budapest, 1976.
- [NN94] Y. E. Nesterov and A. S. Nemirovsky, Interior-point polynomial methods in convex programming, SIAM Studies in Applied Mathematics, SIAM Publications, Philadelphia, 1994.

- [PE67] E. L. Peterson and J. G. Ecker, Geometric programming: Duality in quadratic programming and l_p approximation II, SIAM Journal on Applied Mathematics 13 (1967), 317–340.
- [PE70a] E. L. Peterson and J. G. Ecker, Geometric programming: Duality in quadratic programming and l_p approximation I, Proceedings of the International Symposium of Mathematical Programming (Princeton, New Jersey) (H. W. Kuhn and A. W. Tucker, eds.), Princeton University Press, 1970.
- [PE70b] E. L. Peterson and J. G. Ecker, Geometric programming: Duality in quadratic programming and l_p approximation III, Journal on Mathematical Analysis and Applications 29 (1970), 365–383.
- [Stu00] J. F. Sturm, *Duality results*, High Performance Optimization (H. Frenk, C. Roos, T. Terlaky, and S. Zhang, eds.), Kluwer Academic Publishers, 2000, pp. 21–60.
- [SW70] J. Stoer and Ch. Witzgall, *Convexity and optimization in finite dimensions I*, Springer Verlag, Berlin, 1970.
- [Ter85] T. Terlaky, $On l_p$ programming, European Journal of Operations Research **22** (1985), 70–100.