McMaster University

Advanced Optimization Laboratory

Title:

A conic formulation for l_p-norm optimization

Author:

François Glineur and Tamás Terlaky

AdvOl-Report No. 2000/3

May 2000, Hamilton, Ontario, Canada

A conic formulation for l_p-norm optimization*

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.

Tamás Terlaky[‡]
Department of Computing and Software,
McMaster University,
Hamilton, Ontario, L8S 4L7, Canada
Terlaky@mcmaster.ca.

May 2000

Abstract

In this paper, we formulate the l_p-norm optimization problem as a conic optimization problem, derive its standard duality properties and show it can be solved in polynomial time.

We first define an *ad hoc* closed convex cone $\mathcal{L}^p$, study its properties and derive its dual. This allows us to express the standard l_p-norm optimization primal problem as a conic problem involving $\mathcal{L}^p$. Using convex conic duality and our knowledge about $\mathcal{L}^p$, we proceed to derive the dual of this problem and prove the well-known regularity properties of this primal-dual pair, i.e. zero duality gap and primal attainment. Finally, we prove that the class of l_p-norm optimization problems can be solved up to a given accuracy in polynomial time, using the framework of interior-point algorithms and self-concordant barriers.

Keywords. l_p-norm optimization, conic optimization, duality theory, interior-point methods, self-concordant barrier

*This research has been partly carried at the Delft University of Technology, while the first author was visiting the SSOR department, supported by a grant from the Communauté Française de Belgique.

[†]This author is supported by a grant from the F.N.R.S. (Belgian National Fund for Scientific Research).

[‡]This author is partially supported by the grant #OPG0048923 of the National Sciences and Engineering Research Council of Canada.

Contents

1 Introduction

l_p-norm optimization problems form an important class of convex problems, which includes as special cases linear optimization, quadratically constrained quadratic optimization and l_p-norm approximation problems.

A few interesting duality results are known for l_p-norm optimization. Namely, a pair of feasible primal-dual l_p-norm optimization problems satisfies the weak duality property, which is a mere consequence consequence of convexity, but can also be shown to satisfy two additional properties that cannot be guaranteed in the general convex case: the optimum duality gap is equal to zero and at least one feasible solution attains the optimum primal objective. These results were first presented by Peterson and Ecker [7, 6, 8] and later greatly simplified by Terlaky [12], using standard convex duality theory (e.g. the convex Farkas theorem).

The aim of this paper is to derive these results in a completely different setting, using the machinery of conic convex duality. This new approach has the advantage of further simplifying the proofs and giving some insight about the reasons why this class of problems has better properties than a general convex problem. We also show that this class of optimization problems can be solved up to a given accuracy in polynomial time, using the theory of self-concordant barriers in the framework of interior-point algorithms [5, 1].

1.1 Problem definition

The primal l_p-norm optimization problem can be defined as follows. Let $K = \{1, 2, \dots, r\}$, $I = \{1, 2, \dots, n\}$ 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 $B \in \mathbb{R}^{m \times r}$ (whose columns will be denoted by a_i, $i \in I$ and b_k, $k \in K$) and four column vectors $\eta \in \mathbb{R}^m$, $c \in \mathbb{R}^n$, $d \in \mathbb{R}^r$ 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

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

It is readily seen that this formulation is quite general. Indeed,

- ◇ linear optimization problems can be modelled by taking $n = 0$ (and thus $I_k = \emptyset\ \forall k \in K$), which gives

$$\sup \eta^T y \quad \text{s.t.} \quad B^T y \le d \; ,$$

- ◇ l_p-norm approximation problems correspond to the case $b_k = 0\ \forall k \in K$, described in [7, 12] and [5, Section 6.3.2],

- ◇ a convex quadratic constraint $Q(y) \le 0$ with $Q(y) = \frac{1}{2} y^T C y + f^T y + g \le 0$ (where C is positive semidefinite) is equivalent to $\frac{1}{2} \left\| H^T y \right\|^2 \le -f^T y - g$, where H is a $m \times s$ matrix such that $C = HH^T$ (whose columns will be denoted by h_i), and can be modelled as

$$\sum_{i=1}^{s} \frac{1}{2} \left| h_i^T y \right|^2 \le -g - f^T y \; ,$$

which has the same form as the constraints of problem (P) with $p_i = 2$ and $c_i = 0$. This implies that linearly and quadratically constrained convex quadratic optimization problems can be modelled as l_p-norm optimization problems (since a convex quadratic objective can be modelled using an additional variable, a linear objective and a convex quadratic constraint).

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 (P) can be defined as (see e.g. [12])

$$\inf \ \psi(x, z) = c^T x + d^T z + \sum_{\substack{k \in K \\ z_k > 0}} z_k \sum_{i \in I_k} \frac{1}{q_i} \left| \frac{x_i}{z_k} \right|^{q_i} \quad \text{s.t.} \quad \begin{cases} Ax + Bz = \eta \text{ and } z \geq 0 \ , \\ z_k = 0 \Rightarrow x_i = 0 \ \forall i \in I_k \ . \end{cases} \tag{D}$$

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 (P), this problem has a simpler feasible region (mostly defined by linear equalities and nonnegativity constraints) at the price of a highly nonlinear objective.

1.2 Structure of the paper

The rest of this paper is organized as follows. We recall in Section 2 the results of conic optimization theory that are necessary for our purpose. In order to use this setting, we define in Section 3 an appropriate convex cone that will allow us to express l_p-norm optimization problems as conic programs. We also study some aspects of this cone (closedness, interior, dual). We are then in position to formulate the primal-dual pair (P)–(D) using a conic formulation and apply in Section 4 the general duality theory for conic optimization, in order to prove the above-mentioned duality results about l_p-norm optimization. Section 5 deals with algorithmic complexity issues and presents a self-concordant barrier construction for our problem. We conclude with some remarks in Section 6.

2 Conic optimization

In this section, we describe the formulation of a primal-dual pair of conic optimization problems and state some results from the associated duality theory (this material is quite classical and proofs can be found e.g. in [10, 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 has the advantage of providing a very symmetric form for the dual problem and often gives a new insight about its structure.

2.1 Conic problems

The basic ingredient we need is a convex cone.

Definition 2.1 *A set $\mathcal{C}$ is a* cone *if and only if it is closed under nonnegative scalar multiplication, i.e.*

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

Moreover, it is convex *if and only if it is closed under addition, i.e.*

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

Recall that a set is convex if and only if it contains the whole segment joining any two of its points. It is straightforward to prove that this is equivalent to closedness under addition in the special case of cones.

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 $\mathcal{C}$ is* solid *if and only if* $\operatorname{int} \mathcal{C} \neq \emptyset$ *(where* $\operatorname{int} S$ *denotes the interior of set S).*

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

These two properties basically mean that $\mathcal{C}$ 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 $\mathcal{C} \subseteq \mathbb{R}^n$ a pointed, solid, closed convex cone. The (primal) conic optimization problem is defined as

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

where $x \in \mathbb{R}^n$ is the column vector we are optimizing and the problem data is given by cone $\mathcal{C}$, a $m \times n$ matrix A and two column vectors b and c 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 $\mathcal{C}$ is closed, it may happen that the infimum in (CP) is not attained (an example of this situation will be given in Subsection 4.3).

It is well-known that this class of problems is equivalent to the class of convex problems, see e.g. [5]. However, the usual Lagrangean dual of a conic problem can be also 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 $\mathcal{C} \subseteq \mathbb{R}^n$ is defined by*

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

The following theorem stipulates that the dual of a cone is always a closed convex cone.

Theorem 2.1 *If $\mathcal{C}$ is a closed convex cone, its dual $\mathcal{C}^*$ is another closed convex cone. Moreover, the dual $(\mathcal{C}^*)^*$ of $\mathcal{C}^*$ is equal to $\mathcal{C}$.*

Closedness is essential for $(\mathcal{C}^*)^* = \mathcal{C}$ to hold (without the closedness assumption on $\mathcal{C}$, we only have $(\mathcal{C}^*)^* = \operatorname{cl} \mathcal{C}$ where $\operatorname{cl} S$ denotes the closure of set S). The additional notions of solidness and pointedness also behave well when taking the dual of a convex cone (indeed, these two properties are dual to each other).

Theorem 2.2 *If $\mathcal{C}$ is a solid, pointed, closed convex cone, its dual $\mathcal{C}^*$ is another solid, pointed, closed convex cone.*

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

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

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 the translation of the range space of A^T, while primal involved a translation of the null space of A). It is also possible to show that the dual of this dual problem is equivalent to the primal problem, using the fact that $(\mathcal{C}^*)^* = \mathcal{C}$.

2.3 Duality theory

The two conic problems of this primal-dual pair are strongly related to each other, as demonstrated by the following duality theorems. First, we start with the so-called weak duality property: using the linear constraints of problems (CP) and (CD), we can write the following elementary chain of equalities

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

Noting that the last inner product is always nonnegative when $x \in \mathcal{C}$, $s \in \mathcal{C}^*$ (because of Definition 2.4 of the dual cone $\mathcal{C}^*$), we have

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. 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, called the *duality gap*, can be different from zero in the general case. However, under certain circumstances, it will be possible to prove it is equal to zero, which implies then that the optimum values of problems (CP) and (CD) are equal. Before describing the conditions that can guarantee 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 $\mathcal{C}$ (resp. $\mathcal{C}^*$), i.e.*

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

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 (obviously, infeasible and unbounded problems cannot be attained).

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 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 really relies on this hypothesis. 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.

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 solution for a bounded primal problem 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.*

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* $x^i \in \mathcal{C}^i$ *for all* $i \in \{1, 2, \ldots, k\}$*, where* $\mathcal{C}^i \subseteq \mathbb{R}^{n_i}$*, we will simply consider the Cartesian product of these cones* $\mathcal{C} = \mathcal{C}^1 \times \mathcal{C}^2 \times \cdots \times \mathcal{C}^k \subseteq \mathbb{R}^{\sum_{i=1}^k n_i}$ *and express all these constraints simultaneously as* $x \in \mathcal{C}$ *with* $x = (x^1, x^2, \ldots, x^k)$. *The dual cone of* $\mathcal{C}$ *will be expressed by*

$$\mathcal{C}^* = (\mathcal{C}^1)^* \times (\mathcal{C}^2)^* \times \cdots \times (\mathcal{C}^k)^* \subseteq \mathbb{R}^{\sum_{i=1}^k n_i} ,$$

as implied by the following theorem :

Theorem 2.5 *Let* $\mathcal{C}^1$ *and* $\mathcal{C}^2$ *two closed convex cones, and* $\mathcal{C} = \mathcal{C}^1 \times \mathcal{C}^2$ *their Cartesian product. Cone* $\mathcal{C}$ *is also a closed convex cone, and its dual* $\mathcal{C}^*$ *is given by*

$$\mathcal{C}^* = (\mathcal{C}^1)^* \times (\mathcal{C}^2)^* .$$

3 Cones for l_p-norm optimization

Let us now introduce the $\mathcal{L}^p$ cone, which will allow us to give a conic formulation of l_p-norm optimization problems.

3.1 The $\mathcal{L}^p$ cone

Definition 3.1 *Let $n \in \mathbb{N}$ and $p \in \mathbb{R}^n$ with $p_i > 1$. We define the following set*

$$\mathcal{L}^p = \Big\{ (x, \theta, \kappa) \in \mathbb{R}^n \times \mathbb{R}_+ \times \mathbb{R}_+ \mid \sum_{i=1}^{n} \frac{|x_i|^{p_i}}{p_i \theta^{p_i - 1}} \leq \kappa \Big\}$$

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

$$\frac{|x|}{0} = \begin{cases} +\infty & \text{if } x \neq 0 \ , \\ 0 & \text{if } x = 0 \ . \end{cases}$$

This convention means that if $(x, \theta, \kappa) \in \mathcal{L}^p$, $\theta = 0$ implies $x = 0^n$. We start by proving that $\mathcal{L}^p$ is a convex cone.

Theorem 3.1 *$\mathcal{L}^p$ is a convex cone.*

Proof. Let us first introduce the following function

$$f_p : \mathbb{R}^n \times \mathbb{R}_+ \mapsto \mathbb{R}_+ \cup \{+\infty\} : (x, \theta) \mapsto \sum_{i=1}^{n} \frac{|x_i|^{p_i}}{p_i \theta^{p_i - 1}} \ .$$

With the convention mentioned above, its effective domain is $\mathbb{R}^n \times \mathbb{R}_{++} \cup 0^n \times 0$. It is straightforward to check that f_p is positively homogeneous, i.e. $f_p(\lambda x, \lambda \theta) = \lambda f_p(x, \theta)$ for $\lambda \geq 0$. Moreover, f_p is subadditive, i.e. $f_p(x + x', \theta + \theta') \leq f_p(x, \theta) + f_p(x', \theta')$. In order to show it, we only need to prove the following inequality for all $x, x' \in \mathbb{R}$ and $\theta, \theta' \in \mathbb{R}_+$:

$$\frac{|x|^{p_i}}{\theta^{p_i - 1}} + \frac{|x'|^{p_i}}{\theta'^{p_i - 1}} \geq \frac{|x + x'|^{p_i}}{(\theta + \theta')^{p_i - 1}} \ .$$

First observe that this inequality is obviously true if θ or θ' is equal to 0. When θ and θ' are both different from 0, we use the well known fact that x^{p_i} is a convex function on $\mathbb{R}_+$ for $p_i \geq 1$, implying that $\lambda a^{p_i} + \lambda' a'^{p_i} \geq (\lambda a + \lambda' a')^{p_i}$ for any nonnegative a, a', λ and λ' satisfying $\lambda + \lambda' = 1$. Choosing $a = \frac{1}{\theta} |x|$, $a' = \frac{1}{\theta'} |x'|$, $\lambda = \frac{\theta}{\theta + \theta'}$ and $\lambda' = \frac{\theta'}{\theta + \theta'}$, we find that

$$\begin{aligned} \frac{\theta}{\theta + \theta'} \Big(\frac{|x|}{\theta} \Big)^{p_i} + \frac{\theta'}{\theta + \theta'} \Big(\frac{|x'|}{\theta'} \Big)^{p_i} &\geq \Big(\frac{\theta}{\theta + \theta'} \frac{|x|}{\theta} + \frac{\theta'}{\theta + \theta'} \frac{|x'|}{\theta'} \Big)^{p_i} \\ \frac{1}{\theta + \theta'} \left(\frac{|x|^{p_i}}{\theta^{p_i - 1}} + \frac{|x'|^{p_i}}{\theta'^{p_i - 1}} \right) &\geq \left(\frac{|x| + |x'|}{\theta + \theta'} \right)^{p_i} \\ \frac{|x|^{p_i}}{\theta^{p_i - 1}} + \frac{|x'|^{p_i}}{\theta'^{p_i - 1}} &\geq \frac{(|x| + |x'|)^{p_i}}{(\theta + \theta')^{p_i - 1}} \geq \frac{|x + x'|^{p_i}}{(\theta + \theta')^{p_i - 1}} \ . \end{aligned}$$

Positive homogeneity and subadditivity imply that f_p is a convex function. Since $f_p(x, \theta) \geq 0$ for all x and θ, we notice that $\mathcal{L}^p$ is the epigraph of f_p, i.e.

$$\operatorname{epi} f = \Big\{ (x, \theta, \kappa) \in \mathbb{R}^n \times \mathbb{R}_+ \times \mathbb{R} \mid f_p(x, \theta) \leq \kappa \Big\} = \mathcal{L}^p \ .$$

$\mathcal{L}^p$ is thus the epigraph of a convex positively homogeneous function, hence a convex cone. □

In order to characterize strictly feasible points, we would like to identify the interior of this cone.

Theorem 3.2 *The interior of $\mathcal{L}^p$ is given by*

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

Proof. According to Lemma 7.3 in [9] we have

$$\operatorname{int} \mathcal{L}^p = \operatorname{int} \operatorname{epi} f_p = \{(x, \theta, \kappa) \mid (x, \theta) \in \operatorname{int} \operatorname{dom} f_p \text{ and } f_p(x, \theta) < \kappa\} .$$

The stated result then simply follows from the fact that $\operatorname{int} \operatorname{dom} f_p = \mathbb{R}^n \times \mathbb{R}_{++}$. □

Corollary 3.1 *The cone $\mathcal{L}^p$ is solid.*

Proof. It suffices to prove that there exists at least one point that belongs to $\operatorname{int} \mathcal{L}^p$, for example by taking the point $(e, 1, n)$, where e stands for an n-dimensional the all-one vector. Indeed, we have $\sum_{i=1}^{n} \frac{|1|^1}{p_i 1^{p_i - 1}} = \sum_{i=1}^{n} \frac{1}{p_i} < \sum_{i=1}^{n} 1 = n$. □

Note 3.1 *When $n = 0$, our cone $\mathcal{L}^p$ is readily seen to be equivalent to the two-dimensional positive orthant $\mathbb{R}^2_+$. We also notice that in the special case where $p_i = 2$ for all i, our cone $\mathcal{L}^p$ becomes*

$$\mathcal{L}^{(2, \cdots, 2)} = \left\{ (x, \theta, \kappa) \in \mathbb{R}^n \times \mathbb{R}_+ \times \mathbb{R}_+ \mid \sum_{i=1}^{n} x_i^2 \leq 2\theta\kappa \right\} ,$$

which is usually called the hyperbolic *or* rotated *second-order cone [4, 11](it is a simple linear transformation of the usual second-order cone).*

To illustrate our purpose, we provide in Figure 1 the three-dimensional graphs of the boundary surfaces of $\mathcal{L}^{(5)}$ and $\mathcal{L}^{(2)}$ (corresponding to the case $n = 1$).

3.2 The dual cone

We are now going to determine the dual cone of $\mathcal{L}^p$. We first introduce a useful inequality, which lies at the heart of duality for $\mathcal{L}^p$ cones [12, 5]. In order to keep the exposition self-contained, we also include its proof.

Lemma 3.1 *Let $a, b \in \mathbb{R}_+$ and $\alpha, \beta \in \mathbb{R}_{++}$ such that $\frac{1}{\alpha} + \frac{1}{\beta} = 1$. We have the inequality*

$$\frac{a^\alpha}{\alpha} + \frac{b^\beta}{\beta} \geq ab ,$$

with equality holding if and only if $a^\alpha = b^\beta$.

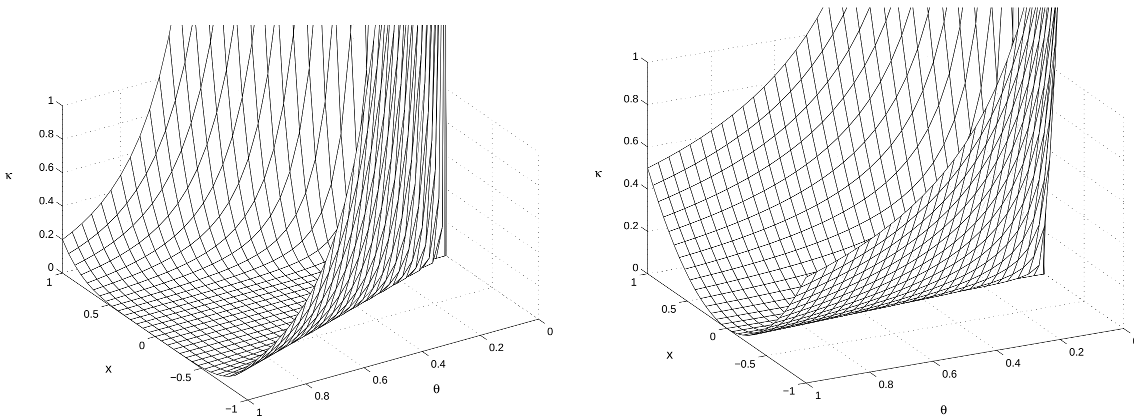

Figure 1: The boundary surfaces of $\mathcal{L}^{(5)}$ and $\mathcal{L}^{(2)}$ (in the case $n = 1$).

Proof. This is an application of the well-known weighted arithmetic-geometric inequality on a^α and b^β with weights $\frac{1}{\alpha}$ and $\frac{1}{\beta}$ (whose sum is equal to one). This gives

$$\frac{a^\alpha}{\alpha} + \frac{b^\beta}{\beta} \geq (a^\alpha)^{1/\alpha}(b^\beta)^{1/\beta} = ab \ .$$

□

For ease of notation, we also introduce the *switched* cone $\mathcal{L}^p_s$ as the $\mathcal{L}^p$ cone with its last two components exchanged, i.e.

$$(x, \theta, \kappa) \in \mathcal{L}^p_s \ \Leftrightarrow \ (x, \kappa, \theta) \in \mathcal{L}^p \ .$$

We are now ready to describe the dual of $\mathcal{L}^p$.

Theorem 3.3 (Dual of $\mathcal{L}^p$) *Let $p, q \in \mathbb{R}^n_{++}$ such that $\frac{1}{p_i} + \frac{1}{q_i} = 1$ for each i. The dual of $\mathcal{L}^p$ is $\mathcal{L}^q_s$.*

Proof. By definition of the dual cone, we have

$$(\mathcal{L}^p)^* = \left\{ v^* \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \mid v^T v^* \geq 0 \text{ for all } v \in \mathcal{L}^p \right\} \ .$$

We start by showing that $\mathcal{L}^q_s \subseteq (\mathcal{L}^p)^*$. Let $v^* = (x^*, \theta^*, \kappa^*) \in \mathcal{L}^q_s$ and $v = (x, \theta, \kappa) \in \mathcal{L}^p$. We are going to prove that $v^T v^* \geq 0$, which will imply the desired inclusion. The case when $\theta = 0$ is easily handled: we have then $x = 0$ implying $v^T v^* = \kappa\kappa^* \geq 0$. Similarly we can eliminate the case where $\kappa^* = 0$. In the remaining cases, we use the definitions of $\mathcal{L}^p$ and $\mathcal{L}^q_s$ to get

$$f_p(x, \theta) = \sum_{i=1}^n \frac{|x_i|^{p_i}}{p_i \theta^{p_i - 1}} \leq \kappa \text{ and } f_q(x^*, \kappa^*) = \sum_{i=1}^n \frac{|x^*_i|^{q_i}}{q_i {\kappa^*}^{q_i - 1}} \leq \theta^* \ .$$

Dividing respectively by θ and κ^* and adding the resulting inequalities we find

$$\sum_{i=1}^n \Big(\frac{|x_i|^{p_i}}{p_i \theta^{p_i}} + \frac{|x^*_i|^{q_i}}{q_i {\kappa^*}^{q_i}} \Big) \leq \frac{\kappa}{\theta} + \frac{\theta^*}{\kappa^*} \ . \tag{3.1}$$

Applying now Lemma 3.1 to each pair $\frac{|x_i|}{\theta}, \frac{|x_i^*|}{\kappa^*}$ we get

$$\sum_{i=1}^{n} \frac{|x_i|}{\theta} \frac{|x_i^*|}{\kappa^*} \leq \frac{\kappa}{\theta} + \frac{\theta^*}{\kappa^*} \ , \tag{3.2}$$

which is equivalent to

$$\sum_{i=1}^{n} |x_i| \, |x_i^*| \leq \kappa\kappa^* + \theta\theta^* \ .$$

Finally, noting that $x_i x_i^* \geq -|x_i| \, |x_i^*|$ we conclude that

$$v^T v^* = x^T x^* + \kappa\kappa^* + \theta\theta^* = \sum_{i=1}^{n} x_i x_i^* + \kappa\kappa^* + \theta\theta^* \geq \sum_{i=1}^{n} - |x_i| \, |x_i^*| + \kappa\kappa^* + \theta\theta^* \geq 0 \ , \tag{3.3}$$

showing that $\mathcal{L}_s^q \subseteq (\mathcal{L}^p)^*$. Let us prove now the reverse inclusion, i.e. $(\mathcal{L}^p)^* \subseteq \mathcal{L}_s^q$. Let $v^* = (x^*, \theta^*, \kappa^*) \in (\mathcal{L}^p)^*$. We have to show that $v^* \in \mathcal{L}_s^q$, using that $v^T v^* \geq 0$ for every $v = (x, \theta, \kappa) \in \mathcal{L}^p$. Choosing $v = (0, 0, 1)$, we first ensure that $v^T v^* = \kappa^* \geq 0$. We distinguish the cases $\kappa^* = 0$ and $\kappa^* > 0$. If $\kappa^* = 0$, we have that $v^T v^* = x^T x^* + \theta\theta^* \geq 0$ for every $v = (x, \theta, \kappa) \in \mathcal{L}^p$. Choosing $\theta = 1$ and $\kappa \geq f_p(x, 1)$ for any $x \in \mathbb{R}^n$, we find that $x^T x^* + \theta^* \geq 0$ for all $x \in \mathbb{R}^n$, which implies $x^* = 0$ and $\theta^* \geq 0$ and thus $v^* \in \mathcal{L}_s^q$. When $\kappa^* > 0$, we can always choose a $v \in \mathcal{L}^p$ such that

$$\frac{|x_i|^{p_i}}{\theta^{p_i}} = \frac{|x_i^*|^{q_i}}{\kappa^{*q_i}} \quad , \quad x_i x_i^* \leq 0 \quad \text{and} \quad f_p(x, \theta) = \sum_{i=1}^{n} \frac{|x_i|^{p_i}}{p_i \theta^{p_i - 1}} = \kappa \ . \tag{3.4}$$

Writing

$$\begin{aligned} 0 \leq \frac{v^T v^*}{\theta\kappa^*} &= \left(\frac{x}{\theta}\right)^T \left(\frac{x^*}{\kappa^*}\right)^T + \frac{\theta^*}{\kappa^*} + \frac{\kappa}{\theta} \\ &= \sum_{i=1}^{n} \frac{x_i}{\theta} \frac{x_i^*}{\kappa^*} + \frac{\theta^*}{\kappa^*} + \frac{\kappa}{\theta} \\ &= \sum_{i=1}^{n} - \frac{|x_i|}{\theta} \frac{|x_i^*|}{\kappa^*} + \frac{\theta^*}{\kappa^*} + \frac{\kappa}{\theta} \ , \end{aligned}$$

using the case of equality of Lemma 3.1 on the pairs $\frac{|x_i|}{\theta}, \frac{|x_i^*|}{\kappa^*}$ and the choice of v in (3.4),

$$\begin{aligned} &= - \sum_{i=1}^{n} \left(\frac{|x_i|^{p_i}}{p_i \theta^{p_i}} + \frac{|x_i^*|^{q_i}}{q_i \kappa^{*q_i}} \right) + \frac{\theta^*}{\kappa^*} + \frac{\kappa}{\theta} \\ &= \frac{\theta^*}{\kappa^*} - \sum_{i=1}^{n} \frac{|x_i^*|^{q_i}}{q_i \kappa^{*q_i}} \ , \end{aligned}$$

and finally multiplying by κ^* leads to

$$\sum_{i=1}^{n} \frac{|x_i^*|^{q_i}}{q_i \kappa^{*q_i - 1}} \leq \theta^* \ ,$$

i.e. $v^* \in \mathcal{L}_s^q$, showing that $(\mathcal{L}^p)^* \subseteq \mathcal{L}_s^q$ and thus $(\mathcal{L}^p)^* = \mathcal{L}_s^q$. □

The dual of a $\mathcal{L}^p$ cone is thus equal, up to a permutation of two variables, to another $\mathcal{L}^p$ cone with a *dual* vector of exponents.

Corollary 3.2 *We also have* $(\mathcal{L}_s^p)^* = \mathcal{L}^q$, $(\mathcal{L}^q)^* = \mathcal{L}_s^p$ *and* $(\mathcal{L}_s^q)^* = \mathcal{L}^p$.

Proof. Obvious considering both the symmetry between $\mathcal{L}^p$ and $\mathcal{L}_s^q$ and the symmetry between p and q. □

Corollary 3.3 $\mathcal{L}^p$ *and* $\mathcal{L}_s^q$ *are solid and pointed.*

Proof. We have already proved that $\mathcal{L}^p$ is solid which, for obvious symmetry reasons, implies that its switched counterpart $\mathcal{L}_s^q$ is also solid. Since pointedness is the property that is dual to solidness (Theorem 2.2), noting that $\mathcal{L}^p = (\mathcal{L}_s^q)^*$ and $\mathcal{L}_s^q = (\mathcal{L}^p)^*$ is enough to prove that $\mathcal{L}^p$ and $\mathcal{L}_s^q$ are also pointed. □

Corollary 3.4 $\mathcal{L}^p$ *and* $\mathcal{L}_s^q$ *are closed.*

Proof. Starting with $(\mathcal{L}^p)^* = \mathcal{L}_s^q$ and taking the dual of both sides, we find $((\mathcal{L}^p)^*)^* = (\mathcal{L}_s^q)^*$. Since $(\mathcal{L}_s^q)^* = \mathcal{L}^p$ by Corollary 3.2 and $((\mathcal{L}^p)^*)^* = \operatorname{cl} \mathcal{L}^p$ [9, page 121], we have $\operatorname{cl} \mathcal{L}^p = \mathcal{L}^p$, hence $\mathcal{L}^p$ is closed. The switched cone $\mathcal{L}_s^q$ is obviously closed as well. □

We can also provide a direct proof of the closedness of $\mathcal{L}^p$: using the fact that it is the epigraph of f_p, it is enough to show that f_p is a lower semicontinuous function [9, Theorem 7.1]. Being convex, f_p is continuous on the interior of its effective domain, i.e. when $\theta > 0$. When $\theta = 0$, we have to prove that

$$\lim_{(x,\theta)\to(x^0,0)} f_p(x,\theta) \geq f_p(x^0,0) \ .$$

On the one hand, if $x_i^0 \neq 0$ for some index i, we have that $f_p(x^0,0) = +\infty$ but also that $\lim_{(x,\theta)\to(x^0,0)} f_p(x,\theta) = +\infty$, since the term $\frac{|x_i|^{p_i}}{p_i \theta^{p_i-1}}$ tends to $+\infty$ when (x_i,θ) tends to $(x_i^0,0)$, hence the inequality is true. On the other hand, if $x^0 = 0$, we have to check that $\lim_{(x,\theta)\to(0,0)} f_p(x,\theta) \geq f_p(0,0) = 0$, which is obviously also true. From this we can conclude that f_p is lower semicontinuous and $\mathcal{L}^p$ is closed.

Note however that f_p is not continuous in $(0,0)$. Choosing an arbitrary positive constant C and defining for example $x_i(\theta) = (Cp_i)^{1/p_i}\theta^{1/q_i}$, so that $x(\theta) \to 0$ when $\theta \to 0$, we have that $\lim_{\theta\to 0} f(x(\theta),\theta) = nC \neq f(0,0) = 0$. The limit of f_p at $(0,0)$ can indeed take any positive value[1].

Note 3.2 *As special cases, we note that when* $n = 0$, $(\mathcal{L}^p)^*$ *is equivalent* $\mathbb{R}_+^2$, *which is the usual dual for* $\mathcal{L}^p = \mathbb{R}_+^2$. *In the case of* $p_i = 2\ \forall i$, *we find*

$$(\mathcal{L}^{(2,\cdots,2)})^2 = \mathcal{L}_s^{(2,\cdots,2)}] = \left\{(x,\theta,\kappa) \in \mathbb{R}^n \times \mathbb{R}_+ \times \mathbb{R}_+ \mid \sum_{i=1}^n x_i^2 \leq 2\theta\kappa\right\} ,$$

which is the expected result. Note that apart from these two special cases, $\mathcal{L}^p$ *is in general not self-dual.*

[1]However, taking $x(\theta)$ proportional to θ, namely $x_i(\theta) = C_i\theta$, we have $\lim_{\theta\to 0} f(x(\theta),\theta) = f(0,0) = 0$, i.e. f_p is continuous on its restrictions to lines passing through the origin.

Note 3.3 (Self-duality of $\mathcal{L}^p$ cones with $n = 1$) *Figure 2, representing $\mathcal{L}^{(\frac{5}{4})}$, illustrates our point: up to a permutation of variables, it is equal to $(\mathcal{L}^{(5)})^*$ (since $\frac{1}{5} + \frac{1}{5/4} = 1$) and is different from $\mathcal{L}^{(5)}$. However, in the particular case where $n = 1$, this difference is not as great as it could be. Namely, one can show easily that $\mathcal{L}^{(p)}$ and its dual are equal up to a simple rescaling of some of the variables. Indeed, we have*

$$(x, \theta, \kappa) \in \mathcal{L}^{(p)} \Leftrightarrow |x|^p \leq p\kappa\theta^{p-1}$$
$$\Leftrightarrow |x|^q \leq p^{\frac{q}{p}} \kappa^{\frac{q}{p}} \theta^{(p-1)\frac{q}{p}}$$

using $\frac{q}{p} = q\frac{1}{p} = q(1 - \frac{1}{q}) = q - 1$ *and* $(p-1)\frac{q}{p} = (1 - \frac{1}{p})q = \frac{1}{q}q = 1$

$$\Leftrightarrow |x|^q \leq p^{q-1}\kappa^{q-1}\theta$$
$$\Leftrightarrow |x|^q \leq q(p\kappa)^{q-1}\frac{\theta}{q} \Leftrightarrow (x, \frac{\theta}{q}, p\kappa) \in \mathcal{L}_s^{(q)} = (\mathcal{L}^{(p)})^* \ .$$

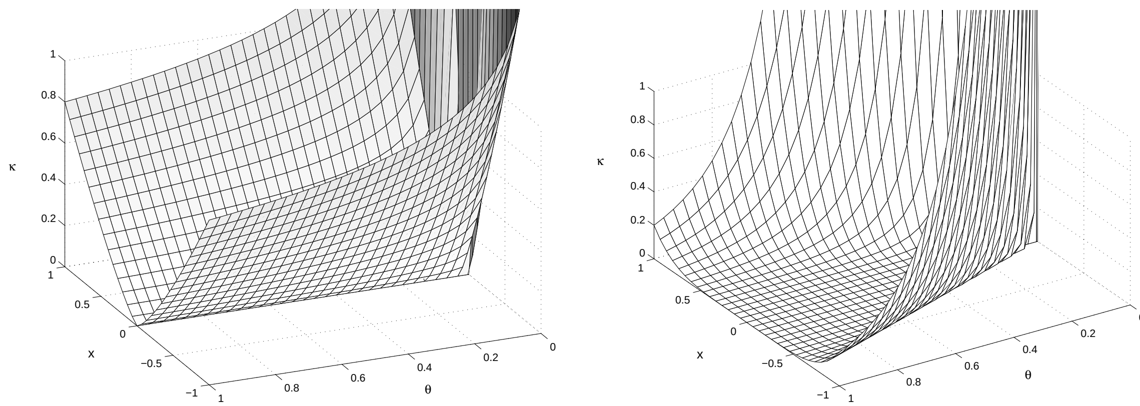

Figure 2: The boundary surface of $\mathcal{L}^{(\frac{5}{4})}$ and $\mathcal{L}^{(5)}$ (in the case $n = 1$).

Our last theorem in this section describes the cases where two vectors from $\mathcal{L}^p$ and $\mathcal{L}_s^q$ are orthogonal to each other, which will be used in the study of the duality properties.

Theorem 3.4 (orthogonality conditions) *Let $v = (x, \theta, \kappa) \in \mathcal{L}^p$ and $v^* = (x^*, \theta^*, \kappa^*) \in \mathcal{L}_s^q$. We have $v^T v^* = 0$ if and only if the following set of conditions holds*

$$\kappa^*(f_p(x, \theta) - \kappa) = 0 \tag{3.5a}$$
$$\theta(f_q(x^*, \kappa^*) - \theta^*) = 0 \tag{3.5b}$$
$$\kappa^* \frac{|x_i|^{p_i}}{\theta^{p_i - 1}} = \theta \frac{|x_i^*|^{q_i}}{\kappa^{* q_i - 1}} \tag{3.5c}$$
$$x_i x_i^* \leq 0 \text{ for all } i \ . \tag{3.5d}$$

Proof. When $\theta > 0$ and $\kappa^* > 0$, a careful reading of the first part of the proof of Theorem 3.3 shows that equality occurs if and only if all conditions in (3.5) are fulfilled. Namely, (3.5a) and (3.5b) are responsible for equality in (3.1), (3.5c) ensures that we are in the case of equality of Lemma 3.1 for inequality (3.2) and the last condition (3.5d) is necessary for equality in (3.3).

When $\theta = 0$ but $\kappa^* > 0$, we have $x = 0$ and thus $v^T v^* = \kappa\kappa^*$. This quantity is zero if and only if $\kappa = 0$, which is equivalent in this case to $f_p(x, \theta) = \kappa$ and occurs if and only if (3.5a) is satisfied (all the other conditions being trivially fulfilled). A similar reasoning takes care of the case $\theta > 0,\ \kappa^* = 0$.

Finally, when $\theta = \kappa^* = 0$, we have $x = x^* = 0$ and $v^T v^* = 0$, while the set of conditions (3.5) is also always satisfied. □

4 Duality for the l_p-norm optimization problem

This is the main section , where we show how a primal-dual pair of l_p-norm optimization problems can be modelled using the $\mathcal{L}^p$ and $\mathcal{L}^q_s$ cones and how this allows us to derive the relevant duality properties.

4.1 Conic formulation

Let us restate here for convenience the definition of the standard primal l_p-norm optimization problem.

$$\sup \eta^T y \quad \text{s.t.} \quad \sum_{i \in I_k} \frac{1}{p_i} \left| c_i - a_i^T y \right|^{p_i} \leq d_k - b_k^T y \quad \forall k \in K \tag{P}$$

(where $K = \{1, 2, \ldots, r\}$, $I = \{1, 2, \ldots, n\}$, $\{I_k\}_{k \in K}$ is a partition of I into r classes, $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{m \times r}$ (whose columns will be denoted by $a_i,\ i \in I$ and $b_k,\ k \in K$), $y \in \mathbb{R}^m$, $\eta \in \mathbb{R}^m$, $c \in \mathbb{R}^n$, $d \in \mathbb{R}^r$ and $p \in \mathbb{R}^n$ such that $p_i > 1\ \forall i \in I$).

Let us now model problem (P) with a conic formulation. The following notation will be used: v_I (resp. M_I) denotes the restriction of column vector v (resp. matrix M) to the components (resp. rows) whose indices belong to set I. We introduce an auxiliary vector of variables $x^* \in \mathbb{R}^n$ to represent the argument of the power functions, namely we let

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

and we also need additional variables $z^* \in \mathbb{R}^r$ for the linear term forming the right-hand side of the inequalities

$$z_k^* = d_k - b_k^T y \text{ for all } k \in K \text{ or, in matrix form, } z^* = d - B^T y \ .$$

Our problem is now equivalent to

$$\sup \eta^T y \quad \text{s.t.} \quad A^T y + x^* = c,\ B^T y + z^* = d \text{ and } \sum_{i \in I_k} \frac{1}{p_i} |x_i^*|^{p_i} \leq z_k^* \quad \forall k \in K \ ,$$

where we can easily plug our definition of the $\mathcal{L}^p$ cone

$$\sup \eta^T y \quad \text{s.t.} \quad A^T y + x^* = c,\ B^T y + z^* = d \text{ and } (x_{I_k}^*, 1, z_k^*) \in \mathcal{L}^{p^k}\ \forall k \in K$$

(where for convenience we defined vectors $p^k = (p_i \mid i \in I_k)$ for $k \in K$). We finally introduce an additional vector of fictitious variables $v^* \in \mathbb{R}^r$ whose components are fixed to 1 by linear constraints to find

$$\sup \eta^T y \quad \text{s.t.} \quad A^T y + x^* = c,\ B^T y + z^* = d,\ v^* = e \text{ and } (x_{I_k}^*, v_k^*, z_k^*) \in \mathcal{L}^{p^k}\ \forall k \in K$$

(where e stands again for the all-one vector). Rewriting the linear constraints with a single matrix equality, we end up with

$$\sup \eta^T y \quad \text{s.t.} \quad \begin{pmatrix} A^T \\ B^T \\ 0 \end{pmatrix} y + \begin{pmatrix} x^* \\ z^* \\ v^* \end{pmatrix} = \begin{pmatrix} c \\ d \\ e \end{pmatrix} \text{ and } (x^*_{I_k}, v^*_k, z^*_k) \in \mathcal{L}^{p^k} \ \forall k \in K \ , \tag{P$_\mathrm{c}$}$$

which is exactly a conic optimization problem in the dual[2] form (CD), using variables $(\tilde{y}, \tilde{s})$, data $(\tilde{A}, \tilde{b}, \tilde{c})$ and a cone $\mathcal{C}^*$ such that

$$\tilde{y} = y, \ \tilde{s} = \begin{pmatrix} x^* \\ z^* \\ v^* \end{pmatrix}, \ \tilde{A} = \begin{pmatrix} A & B & 0 \end{pmatrix}, \ \tilde{b} = \eta, \ \tilde{c} = \begin{pmatrix} c \\ d \\ e \end{pmatrix} \text{ and } \mathcal{C}^* = \mathcal{L}^{p^1} \times \mathcal{L}^{p^2} \times \cdots \times \mathcal{L}^{p^r} \ ,$$

where $\mathcal{C}^*$ 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 $\mathcal{L}^p$ proved in the previous section, it is straightforward to show that $\mathcal{C}^*$ is a solid, pointed, closed convex cone whose dual is

$$(\mathcal{C}^*)^* = \mathcal{C} = \mathcal{L}_s^{q^1} \times \mathcal{L}_s^{q^2} \times \cdots \times \mathcal{L}_s^{q^r} \ ,$$

another solid, pointed, closed convex cone (where we have defined a vector $q \in \mathbb{R}^n$ such that $\frac{1}{p_i} + \frac{1}{q_i} = 1$ for all $i \in I$ and vectors q^k such that $q^k = (q_i \mid i \in I_k)$ for $k \in K$). This allows us to derive a dual problem to (P$_\mathrm{c}$) in a completely mechanical way and find the following conic optimization problem, expressed in the primal form (CP):

$$\inf \begin{pmatrix} c^T & d^T & e^T \end{pmatrix} \begin{pmatrix} x \\ z \\ v \end{pmatrix} \quad \text{s.t.} \quad \begin{pmatrix} A & B & 0 \end{pmatrix} \begin{pmatrix} x \\ z \\ v \end{pmatrix} = \eta \text{ and } (x_{I_k}, v_k, z_k) \in \mathcal{L}_s^{q^k} \text{ for all } k \in K \ ,$$

which is equivalent to

$$\inf c^T x + d^T z + e^T v \quad \text{s.t.} \quad Ax + Bz = \eta \text{ and } (x_{I_k}, v_k, z_k) \in \mathcal{L}_s^{q^k} \text{ for all } k \in K \ , \tag{D$_\mathrm{c}$}$$

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

$$\inf c^T x + d^T z + e^T v \quad \text{s.t.} \quad Ax + Bz = \eta, \ \sum_{i \in I_k} \frac{|x_i|^{q_i}}{p_i z_k^{q_i - 1}} \leq v_k \ \forall k \in K \text{ and } z \geq 0 \ ,$$

keeping in mind the convention on zero denominators that in effect implies $z_k = 0 \Rightarrow x_{I_k} = 0$. Finally, we can remove the v variables from the formulation since they are only constrained by the sum inequalities which have to be tight at any optimal solution. We can thus directly incorporate these sums into the objective function, which leads to

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

Unsurprisingly, the dual formulation (D) we have just found without much effort is exactly the standard form for a dual l_p-norm optimization problem [12].

[2] This is the reason why we added a * superscript to the notation of our additional variables, in order to emphasize the duality correspondence with the associated primal variables.

4.2 Duality properties

We are now able to prove the weak duality property for the l_p-norm optimization problem.

Theorem 4.1 (Weak duality) *If y is feasible for (P) and (x, z) is feasible for (D), we have $\psi(x, z) \geq \eta^T y$. Equality occurs if and only if for all $k \in K$ and $i \in I_k$*

$$z_k\Big(\sum_{i \in I_k} \frac{1}{p_i} \left|c_i - a_i^T y\right|^{p_i} + b_k^T y - d_k\Big) = 0, \quad x_i(c_i - a_i^T y) \leq 0, \quad z_k \left|c_i - a_i^T y\right|^{p_i} = \frac{|x_i|^{q_i}}{z_k^{q_i - 1}} \ . \quad (4.1)$$

Proof. Let y and (x, z) be feasible for (P) and (D). Choosing $v_k = f_{q^k}(x_{I_k}, z_k)$ for all $k \in K$, we have that (x, z, v) is feasible for ($\mathrm{D_c}$) with the same objective function, i.e. with $c^T x + d^T z + e^T v = \psi(x, z)$. Moreover, computing (x^*, z^*, v^*) from y in order to satisfy the linear constraints in ($\mathrm{P_c}$), i.e. according to

$$x_i^* = c_i - a_i^T y, \quad z_k^* = d_k - b_k^T y, \quad v_k^* = 1 \ , \qquad (4.2)$$

we have that (x^*, z^*, v^*, y) is feasible for ($\mathrm{P_c}$). The standard weak duality property for the conic pair ($\mathrm{P_c}$)–($\mathrm{D_c}$) from Theorem 2.3 then states that $c^T x + d^T z + e^T v \geq \eta^T y$, which in turn implies $\psi(x, z) \geq \eta^T y$.

We proceed now to investigate the equality conditions. At the optimum, variables v_k must assume their lower bounds so that we can still assume that $v_k = f_{q^k}(x_{I_k}, z_k)$ holds for all $k \in K$. We also keep variables (x^*, z^*, v^*) defined by (4.2). From Theorem 2.3, we know that equality can only occur if the primal and dual vectors of variables are orthogonal to each other for each conic constraint, i.e. $(x_{I_k}^*, z_k^*, v_k^*)^T (x_{I_k}, z_k, v_k) = 0$ for all $k \in K$.

Having $(x_{I_k}^*, v_k^*, z_k^*)^T \in \mathcal{L}^{p^k}$ and $(x_{I_k}, v_k, z_k) \in \mathcal{L}_s^{q^k}$, Theorem 3.4 gives us the necessary and sufficient conditions for equality to happen

$$z_k(f_{p^k}(x_{I_k}^*, v_k^*) - z_k^*) = 0, \ v_k^*(f_{q^k}(x_{I_k}, z_k) - v_k) = 0, \ z_k \frac{|x_i^*|^{p_i}}{v_k^{*p_i - 1}} = v_k^* \frac{|x_i|^{q_i}}{z_k^{q_i - 1}}, \ x_i x_i^* \leq 0 \qquad (4.3)$$

for all $i \in I_k$ and $k \in K$. The second condition is always satisfied while the other three conditions can be readily simplified using (4.2) to give the announced inequalities (4.1). □

The weak duality property is a rather straightforward consequence of the convexity of the problems, and in fact can be proved without too many difficulties without sophisticated tools from duality theory. However, this is not the case with the next theorem, which deals with a strong duality property.

In the case of a general pair of primal and dual conic problems, the duality gap at the optimum is not always equal to zero, neither are the primal or dual optimum objective values always attained by feasible solutions. However, it is well-known that in the special case of linear optimization, we always have a zero duality gap and attainment of both optimum objective values. The status of l_p-norm optimization lies somewhere between these two situations: the duality gap is always equal zero but attainment of the optimum objective value can only be guaranteed for the primal problem.

In the course of our proof, we will need to use the well-known Goldman-Tucker theorem [3] for linear programming, which we restate here for reference.

Theorem 4.2 (Goldman Tucker) *Let us consider the following primal-dual pair of linear optimization problems in standard form:*

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

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

This partition is called the optimal partition[3]. *Moreover, there exists at least an optimal primal-dual solution* $(\hat{x}, \hat{y}, \hat{s})$ *such that* $\hat{x} + \hat{s} > 0$, *hence satisfying* $\hat{x}_{\mathcal{B}} > 0$ *and* $\hat{s}_{\mathcal{N}} > 0$. *Such a pair is called a* strictly complementary *pair.*

The strong duality theorem for l_p-norm optimization we are about to prove is the following:

Theorem 4.3 (Strong duality) *If both problems* (P) *and* (D) *are feasible, the primal optimal objective value is attained with a zero duality gap, i.e.*

$$\begin{aligned} p^* = \max \eta^T y \quad s.t. \quad & \sum_{i \in I_k} \frac{1}{p_i} \left| c_i - a_i^T y \right|^{p_i} \leq d_k - b_k^T y \quad \forall k \in K \\ = \inf \psi(x, z) \quad s.t. \quad & \left\{ \begin{array}{l} Ax + Bz = \eta \text{ and } z \geq 0 \\ z_k = 0 \Rightarrow x_i = 0 \ \forall i \in I_k \end{array} \right. \qquad = d^* \; . \end{aligned}$$

Proof. Theorem 2.4 tells us that zero duality gap and primal attainment are guaranteed by the existence of a strictly interior dual feasible solution (excluding the case of an unbounded dual). Let (x, z) be a feasible solution for (D). We would like to complement it with a vector v such that the corresponding solution (x, z, v) is strictly feasible for the conic formulation $(\mathrm{D_c})$.

Since cone $\mathcal{C}$ is the cartesian products of the set of cones $\mathcal{L}_s^{q^k}$ for $k \in K$, we need in fact for (x, z, v) to be a strictly feasible solution of $(\mathrm{D_c})$ that $(x_{I_k}, z_k, v_k) \in \operatorname{int} \mathcal{L}_s^{q^k}$ holds for all $k \in K$. Using now Theorem 3.2 to identify the interior of the $\mathcal{L}_s^q$ cones, we see that both conditions $v_k > f_{p^k}(x_{I_k}, z_k)$ and $z_k > 0$ have to be valid for all $k \in K$.

Since vector v contains only free variables and is not constrained by the linear constraints, it is always possible to choose it such that $v_k > f_{p^k}(x_{I_k}, z_k)$ for all $k \in K$. However, the situation is much different for z: it is unfortunately not always possible to find a strictly positive z, since it may happen that the linear constraints combined with the nonnegativity constraint on z force one or more of the components z_k to be equal to zero for all primal feasible solutions. Here is an outline of the three-step strategy we are going to follow:

a. Since some components of z may prevent the existence of a strictly feasible solution to $(\mathrm{D_c})$, we are going to define a *restricted* version of $(\mathrm{D_c})$ where those problematic components of z and the associated variables x have been removed. Hopefully, this restricted problem $(\mathrm{D_r})$ will not behave too differently from the original because the zero components of z and x did not play a crucial role in it.

[3] This optimal partition can be computed in polynomial time by interior-point methods.

b. Since this restricted problem will now admit a strictly feasible solution, its dual problem ($\mathrm{P_r}$) (which is a problem in primal form) has a duality gap equal to zero with its optimal objective value attained by some solution.

c. The last step of our proof will be to convert this optimal solution with a zero duality gap for the restricted primal problem ($\mathrm{P_r}$) into an optimal solution for the original primal problem ($\mathrm{P_c}$).

The whole procedure can be summarized with the following diagram:

(P)	$\equiv$	($\mathrm{P_c}$)	Weak $\longleftrightarrow$	($\mathrm{D_c}$)	$\equiv$	(D)
		c. $\updownarrow$		$\updownarrow$ a.		
		($\mathrm{P_r}$)	Strong (zero gap) $\longleftrightarrow$	($\mathrm{D_r}$)		
		$\downarrow$		$\uparrow$		
		(Attainment)	b.	(Strictly feasible)		

Let us first identify the problematic z_k's that are identically equal to zero for all feasible solutions. This can be done by solving the following linear optimization problem:

$$\min 0 \quad \text{s.t.} \quad Ax + Bz = \eta \text{ and } z \geq 0 \, . \tag{LP}$$

This problem has the same feasible region as our dual problem (D) (actually, its feasible region can be slightly larger from the point of view of the x variables, since the special constraints $z_k = 0 \Rightarrow x_{I_k} = 0$ have been omitted, but this does not have any effect on our reasoning). We are thus looking for components of z that are equal to zero on the whole feasible region of (LP).

Since this problem has a zero objective function, all its feasible solutions are optimal and we can therefore deduce that if a variable z_k is zero for all feasible solutions to problem (LP), it is zero for all optimal solution to problem (LP). In order to use the Goldman-Tucker theorem, we also write the dual[4] of problem (LP):

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

Both (LP) and (DLP) are feasible (the former because (D) is assumed to be feasible, the latter because $(y, z^*) = (0, 0)$ is always a feasible solution), which means 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 z_i that are identically zero on the feasible regions of problems (LP) and (D). We are thus now ready the apply the strategy outlined above.

a. Let us introduce the reduced primal-dual pair of l_p-norm optimization problems where variables z_k and x_{I_k} with $k \in \mathcal{N}$ have been removed. We start with the dual problem

$$\inf c_{I_\mathcal{B}}^T x_{I_\mathcal{B}} + d_\mathcal{B}^T z_\mathcal{B} + e^T v_\mathcal{B} \quad \text{s.t.} \quad A_{I_\mathcal{B}} x_{I_\mathcal{B}} + B_\mathcal{B} z_\mathcal{B} = \eta, \ (x_{I_k}, v_k, z_k) \in \mathcal{L}_s^{q^k} \ \forall k \in \mathcal{B} \, , \tag{$\mathrm{D_r}$}$$

[4]Although problem (LP) is not exactly formulated in the standard form used to state Theorem 4.2, the same results hold in the case of a general linear optimization problem

where $I_{\mathcal{B}}$ stands for $\cup_{k\in\mathcal{B}} I_k$. It is straightforward to check that this problem is completely equivalent to problem ($\mathrm{D_c}$), since the variables $z_{\mathcal{N}}$ and $x_{I_{\mathcal{N}}}$ we removed, being forced to zero for all feasible solutions, had no contribution to the objective or to the linear constraints in ($\mathrm{D_c}$).

The corresponding conic constraints become $(0, v_k, 0) \in \mathcal{L}_s^{q^k} \Leftrightarrow v_k \geq 0$, which imply at the optimum that $v_k = 0$, showing that variables $v_{\mathcal{N}}$ can also be safely removed without changing the optimum objective value. We can thus conclude that $\inf(\mathrm{D_r}) = \inf(\mathrm{D_c}) = \inf(\mathrm{D})$.

b. Because of the second part of the Goldman-Tucker theorem, there is at least one feasible solution to (LP) such that $z_{\mathcal{B}} > 0$. Combining the $(x_{I_{\mathcal{B}}}, z_{\mathcal{B}})$ part of this solution with a vector $v_{\mathcal{B}}$ with sufficiently large components gives us a strictly feasible solution for ($\mathrm{D_r}$) ($z_k > 0$ and $v_k > f_{q^k}(x_{I_k}, z_k)$ for all $k \in \mathcal{B}$), which is exactly what we need to apply our strong duality Theorem 2.4. Let us first write down the dual problem of ($\mathrm{D_r}$), the restricted primal:

$$\sup \eta^T y \quad \text{s.t.} \quad \begin{cases} A_{I_{\mathcal{B}}}^T y + x_{I_{\mathcal{B}}}^* = c_{I_{\mathcal{B}}},\ B_{\mathcal{B}}^T y + z_{\mathcal{B}}^* = d_{\mathcal{B}},\ v_{\mathcal{B}}^* = e, \\ (x_{I_k}^*, v_k^*, z_k^*) \in \mathcal{L}^{p^k}\ \forall k \in \mathcal{B}\ . \end{cases} \tag{$\mathrm{P_r}$}$$

We cannot be in the first case of Theorem 2.4, since unboundedness of ($\mathrm{D_r}$) would imply unboundedness of the original problem (D) which in turn would prevent the existence of a feasible primal solution (simple consequence of the weak duality theorem). We can thus conclude that there exists an optimal solution to ($\mathrm{P_r}$) $(\hat{x}_{I_{\mathcal{B}}}^*, \hat{z}_{\mathcal{B}}^*, \hat{v}_{\mathcal{B}}^*, \hat{y})$ such that $\eta^T \hat{y} = \max(\mathrm{P_r}) = \inf(\mathrm{D_r})$.

c. Combining the results obtained so far, we have proved that $\max(\mathrm{P_r}) = \inf(\mathrm{D})$. The last step we need to perform is to prove that $\max(\mathrm{P}) = \max(\mathrm{P_r})$, i.e. that the optimum objective of (P) is attained and that it is equal to the optimal objective value of ($\mathrm{P_r}$). Unfortunately, the apparently most straightforward way to do this, namely using the optimal solution $\hat{y}$ we have at hand for problem ($\mathrm{P_r}$), does not work since it is not necessarily feasible for problem ($\mathrm{P_c}$). The reason is that ($\mathrm{P_c}$) contains additional conic constraints (the ones corresponding to $k \in \mathcal{N}$) which are not guaranteed to be satisfied by the optimal solution $\hat{y}$ of the restricted problem. We can however overcome this difficulty by perturbing this solution by a suitably chosen vector such that

- ◇ feasibility for the constraints $k \in \mathcal{B}$ is not lost,
- ◇ feasibility for the constraints $k \in \mathcal{N}$ can be gained.

Let us consider $(\bar{x}, \bar{z}, \bar{y}, \bar{z}^*)$, a strictly complementary solution to the primal-dual pair (LP)–(DLP) whose existence is guaranteed by the Goldman-Tucker theorem. We have thus $\bar{z}_{\mathcal{N}}^* > 0$ and $\bar{z}_{\mathcal{B}}^* = 0$. Since all primal solutions have a zero objective, the optimal dual objective value also satisfies $\eta^T \bar{y} = 0$. Summarizing the properties of $\bar{y}$ obtained so far, we can write

$$\eta^T \bar{y} = 0, \quad A^T \bar{y} = 0, \quad B_{\mathcal{B}}^T \bar{y} = -\bar{z}_{\mathcal{B}}^* = 0 \text{ and } B_{\mathcal{N}}^T \bar{y} = -\bar{z}_{\mathcal{N}}^* < 0\ .$$

Let us now consider $y = \hat{y} + \lambda \bar{y}$ with $\lambda \geq 0$ as a solution of ($\mathrm{P_c}$) and compute the value of x^* and z^* given by (4.2), distinguishing the $\mathcal{B}$ and $\mathcal{N}$ parts (we already know that

$v^* = e$):

$$\begin{array}{llllllll} x^*_{I_\mathcal{B}} & = & c_{I_\mathcal{B}} - A^T_{I_\mathcal{B}} y & = & c_{I_\mathcal{B}} - A^T_{I_\mathcal{B}} \hat{y} & = & \hat{x}^*_{I_\mathcal{B}} & (\text{using } A^T_{I_\mathcal{B}} \bar{y} = 0) \\ z^*_\mathcal{B} & = & d_\mathcal{B} - B^T_\mathcal{B} y & = & d_\mathcal{B} - B^T_\mathcal{B} \hat{y} & = & \hat{z}^*_\mathcal{B} & (\text{using } B^T_\mathcal{B} \bar{y} = 0) \\ x^*_{I_\mathcal{N}} & = & c_{I_\mathcal{N}} - A^T_{I_\mathcal{N}} y & = & c_{I_\mathcal{N}} - A^T_{I_\mathcal{N}} \hat{y} & = & \hat{x}^*_{I_\mathcal{N}} & (\text{using } A^T_{I_\mathcal{N}} \bar{y} = 0) \\ z^*_\mathcal{N} & = & d_\mathcal{N} - B^T_\mathcal{N} y & = & d_\mathcal{N} - B^T_\mathcal{N} \hat{y} + \lambda \bar{z}^*_\mathcal{N} & & & (\text{using } -B^T_\mathcal{N} \bar{y} = \bar{z}^*_\mathcal{N}) \; . \end{array}$$

The conic constraints corresponding to $k \in \mathcal{B}$ remain valid for all λ, since the associated variables do not vary with λ. Considering now the constraints for $k \in \mathcal{N}$, we see that $x^*_{I_\mathcal{N}}$ does not depend on λ, while $z^*_\mathcal{N}$ can be made arbitrarily large by increasing λ, due to the fact that $\bar{z}^*_\mathcal{N} > 0$. Choosing a sufficiently large λ, we can force $(x^*_{I_k}, 1, z^*_k) \in \mathcal{L}^{q^k}_s$ for $k \in \mathcal{N}$ and thus make (x^*, v^*, z^*, y) feasible for $(\mathrm{P_c})$. Obviously, we also have that y is feasible for (P) with the same objective value.

Evaluating this objective value, we find that $\eta^T y = \eta^T \hat{y} + \lambda \eta^T \bar{y} = \eta^T \hat{y} = \max(\mathrm{P_r})$, i.e. the feasible solution y we constructed has the same objective value for $(\mathrm{P_c})$ and (P) as $\hat{y}$ for $(\mathrm{P_r})$. This proves that $\max(\mathrm{P_r}) \le \sup(\mathrm{P})$, which combined with our previous results gives $d^* = \inf(\mathrm{D}) = \eta^T \hat{y} = \max(\mathrm{P_r}) \le \sup(\mathrm{P}) = p^*$. Finally, using the weak duality of Theorem 4.1, i.e. $p^* \le d^*$, we obtain $d^* = \inf(\mathrm{D}) = \eta^T \hat{y} = \sup(\mathrm{P}) = p^*$, which implies that $\hat{y}$ is optimum for (P), $\sup(\mathrm{P}) = \max(\mathrm{P})$ and finally the desired result $p^* = \max(\mathrm{P}) = \inf(\mathrm{D}) = d^*$.

□

4.3 Examples

We conclude this section by providing a few examples of the possible situations that can arise for a couple of primal-dual l_p-norm optimization problems. Let us consider the following problem data:

$$r = 1,\ K = \{1\},\ n = 1,\ I_1 = \{1\}, m = 1, A = 1, B = 0, c = 5, d \in \mathbb{R}, \eta = 1, p = 3$$

(d_1 is left unspecified), which translates into the following primal problem:

$$\sup y_1 \quad \text{s.t.} \quad \frac{1}{3} |5 - y_1|^3 \le d_1 \; .$$

Noting $q = \frac{3}{2}$, we can also write down the dual

$$\inf 5x_1 + d_1 z_1 + z_1 \frac{1}{3/2} \left| \frac{x_1}{z_1} \right|^{3/2} \quad \text{s.t.} \quad x_1 = 1,\ z_1 \ge 0, z_1 = 0 \Rightarrow x_1 = 0 \; .$$

This pair of problems can readily be simplified to

$$\sup y_1 \quad \text{s.t.} \quad |5 - y_1| \le \sqrt[3]{3d_1} \qquad \text{and} \qquad \inf 5 + d_1 z_1 + \frac{2}{3\sqrt{z_1}} \quad \text{s.t.} \quad z_1 > 0$$

- ◇ When $d = 9$, our primal constraint becomes $|5 - y_1| \leq 3$, which gives a primal optimum equal to $y_1 = 8$. Looking at the dual, we have

$$9z_1 + \frac{2}{3\sqrt{z_1}} = \frac{1}{3}(27z_1) + \frac{2}{3}(\frac{1}{\sqrt{z_1}}) \leq (27z_1)^{\frac{1}{3}}(\frac{1}{\sqrt{z_1}})^{\frac{2}{3}} = 3$$

(using the weighted arithmetic-geometric mean), which shows that the dual optimum is also equal to 8, and is attained for $(x, z) = (1, \frac{1}{9})$. This is the most common situation: both optimum values are finite and attained, with a zero duality gap.

- ◇ When $d = 0$, our primal constraint becomes $|5 - y_1| \leq 0$, which implies that the only feasible solution is $y_1 = 5$, giving a primal optimum equal to 5. The dual optimum value is then $\inf 5 + \frac{2}{3\sqrt{z_1}} = 5$, equal to the primal but not attained ($z_1 \to +\infty$). This shows that there are problems for which the dual optimum is not attained, i.e. we do not have the perfect duality of linear optimization (one can observe that in this case the primal had no strict interior).

- ◇ Finally, when $d = -1$, the primal becomes infeasible while the dual is unbounded (take again $z \to +\infty$).

5 Complexity

The goal of this section is to prove it is possible to solve an l_p-norm optimization problem up to a given accuracy in polynomial time. According to the theoretical framework of Nesterov and Nemirovksy [5], in order to solve the conic problem from Section 2

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

we only need to find a computable self-concordant barrier for the cone $\mathcal{C}$.

Definition 5.1 *A convex function $f : \operatorname{int} \mathcal{C} \mapsto \mathbb{R}$ is called a self-concordant barrier for the cone $\mathcal{C} \subseteq \mathbb{R}^n$ with parameter θ if and only if f is three times continuously differentiable on* $\operatorname{int} \mathcal{C}$, *$f(x) \to +\infty$ when $x \to \partial\mathcal{C}$ and f satisfies the following two conditions*

$$\begin{aligned} \nabla^3 f(x)[h,h,h] &\leq 2\left(h^T \nabla^2 f(x) h\right)^{\frac{3}{2}} \\ \nabla f(x)^T (\nabla^2 f(x))^{-1} \nabla f(x) &\leq \theta \end{aligned}$$

for all $x \in \operatorname{int} \mathcal{C}$ and $h \in \mathbb{R}^n$.

We have the remarkable result [5]

Theorem 5.1 *Given a self-concordant barrier for the cone $\mathcal{C} \subseteq \mathbb{R}^n$ with parameter θ and a feasible interior starting point $x_0 \in \operatorname{int} \mathcal{C}$, a short-step interior-point algorithm can solve problem* (CP) *up to ϵ accuracy within*

$$O\left(\sqrt{\theta} \log \frac{c^T x_0 - p^*}{\epsilon}\right) \text{ iterations,}$$

such that at each iteration the self-concordant barrier and its first and second derivatives have to be evaluated and a linear system has to be solved in $\mathbb{R}^n$.

We are now going to describe a self-concordant barrier that allows us to solve conic problems involving our $\mathcal{L}^p$ cone (we follow an approach similar to the one used in [13]). The following convex cone

$$\{(x,y) \in \mathbb{R} \times \mathbb{R}_+ \mid |x|^p \leq y\}$$

(with $p > 1$) admits the well-known self-concordant barrier

$$f_p : \mathbb{R} \times \mathbb{R}_{++} \mapsto \mathbb{R} : (x,y) \mapsto -2\log y - \log(y^{2/p} - x^2)$$

with parameter 4 (see [5, Propostion 5.3.1]). Let $n \in \mathbb{N}$, $p \in \mathbb{R}^n$ and $I = \{1, 2, \ldots, n\}$. We have that

$$\{(x,y) \in \mathbb{R}^n \times \mathbb{R}^n_+ \mid |x|_i^p \leq y_i \ \forall i \in I\}$$

admits

$$f_p : \mathbb{R}^n \times \mathbb{R}^n_{++} \mapsto \mathbb{R} : (x,y) \mapsto \sum_{i=1}^n \left(-2\log y_i - \log(y_i^{2/p} - x_i^2)\right)$$

with parameter $4n$ (using [5, Propostion 5.1.2]). This also implies that the set

$$\mathcal{S}_p = \left\{(x,y,\kappa) \in \mathbb{R}^n \times \mathbb{R}^n_+ \times \mathbb{R} \ \mid \ |x|_i^p \leq y_i \ \forall i \in I \text{ and } \kappa = \sum_{i=1}^n \frac{y_i}{p_i}\right\}$$

admits a self-concordant barrier $f'_p(x,y,\kappa) = f_p(x,y)$ with parameter $4n$ (taking the cartesian product with $\mathbb{R}$ essentially leaves the self-concordant barrier unchanged, taking the intersection with an affine subspace does not influence self-concordancy). Finally, we use another result from Nesterov and Nemirovsky to find a self-concordant barrier for the *conic hull* of $\mathcal{S}_p$, which is defined by

$$\begin{aligned}
\mathcal{H}_p &= \text{cl}\left\{(x,t) \in \mathcal{S}_p \times \mathbb{R}_{++} \ \mid \ \frac{x}{t} \in \mathcal{S}_p\right\} \\
&= \text{cl}\left\{(x,y,\kappa,\theta) \in \mathcal{S}_p \times \mathbb{R}_{++} \ \mid \ (\frac{x}{\theta}, \frac{y}{\theta}, \frac{\kappa}{\theta}) \in \mathcal{S}_p\right\} \\
&= \text{cl}\left\{(x,y,\kappa,\theta) \in \mathbb{R}^n \times \mathbb{R}^n_+ \times \mathbb{R} \times \mathbb{R}_{++} \ \mid \ \left|\frac{x_i}{\theta}\right|^p \leq \frac{y_i}{\theta} \ \forall i \in I \text{ and } \frac{\kappa}{\theta} = \sum_{i=1}^n \frac{y_i}{p_i\theta}\right\} \\
&= \text{cl}\left\{(x,y,\kappa,\theta) \in \mathbb{R}^n \times \mathbb{R}^n_+ \times \mathbb{R} \times \mathbb{R}_{++} \ \mid \ \frac{|x_i|^p}{\theta^{p-1}} \leq y_i \ \forall i \in I \text{ and } \kappa = \sum_{i=1}^n \frac{y_i}{p_i}\right\} \\
&= \left\{(x,y,\kappa,\theta) \in \mathbb{R}^n \times \mathbb{R}^n_+ \times \mathbb{R} \times \mathbb{R}_+ \ \mid \ \frac{|x_i|^p}{\theta^{p-1}} \leq y_i \ \forall i \in I \text{ and } \kappa = \sum_{i=1}^n \frac{y_i}{p_i}\right\}
\end{aligned}$$

(to find the last equality, you have to consider accumulation points with $\theta = 0$, which in fact must satisfy $x = 0$, which in turn can be seen to match exactly the convention about zero denominators we chose in Definition 3.1), and find that

$$h_p : \mathbb{R}^n \times \mathbb{R}^n_{++} \times \mathbb{R} \times \mathbb{R}_{++} \mapsto \mathbb{R} : (x,y,\kappa,\theta) \mapsto 400\left(f_p(\frac{x}{\theta}, \frac{y}{\theta}) - 8n\log\theta\right)$$

is a self-concordant barrier for $\mathcal{H}_p$ with parameter $3200n$ (see [5, Proposition 5.1.4]). We now make the following interesting observation linking $\mathcal{H}_p$ to our cone $\mathcal{L}^p$.

Theorem 5.2 *The $\mathcal{L}^p$ cone is equal to the projection of $\mathcal{H}_p$ on the space of (x, κ, θ), i.e.*

$$(x, \theta, \kappa) \in \mathcal{L}^p \quad \Leftrightarrow \quad \exists y \in \mathbb{R}^n_+ \;\mid\; (x, y, \kappa, \theta) \in \mathcal{H}_p \ .$$

Proof. This proof is straightforward. First note that both sets take the same convention in case of a zero denominator. Let $(x, \theta, \kappa) \in \mathcal{L}^p$. Choosing y such that $y_i = \frac{|x_i|}{\theta^{p-1}}$ for all $i \in I$ ensures that

$$\sum_{i=1}^n \frac{y_i}{p_i} = \sum_{i=1}^n \frac{|x_i|^p}{p_i \theta^{p-1}} \leq \kappa$$

(this last inequality because of the definition of $\mathcal{L}^p$). It is now possible to increase y_1 until the equality $\kappa = \sum_{i=1}^n \frac{y_i}{p_i}$ is satisfied, which shows $(x, y, \kappa, \theta) \in \mathcal{H}_p$. For the reverse inclusion, suppose $(x, y, \kappa, \theta) \in \mathcal{H}_p$. This implies that

$$\kappa = \sum_{i=1}^n \frac{y_i}{p_i} \geq \sum_{i=1}^n \frac{|x_i|^p}{p_i \theta^{p-1}} \ ,$$

which is exactly the defining inequality of $\mathcal{L}^p$. □

Suppose now we have now to solve

$$\inf_x c^T x \quad \text{s.t.} \quad Ax = b \text{ and } x \in \mathcal{L}^p \ . \tag{L}$$

In light of the previous theorem, it is equivalent to solve

$$\inf_{(x,y)} c^T x \quad \text{s.t.} \quad Ax = b \text{ and } (x, y) \in \mathcal{H}_p \ ,$$

for which we know a self-concordant barrier with parameter $3200n$. This implies that it is possible to find an approximate solution to problem (L) with accuracy ϵ in $O(\sqrt{n} \log \frac{1}{\epsilon})$ iterations. Moreover, since it is possible to compute in polynomial time the value of h_p and of its first two derivatives, we can conclude that problem (L) is solvable in polynomial time.

This argument is rather easy to generalize to the case of the cartesian product of several $\mathcal{L}^p$ cones dual $\mathcal{L}^q_s$ cones, which shows eventually that any primal or dual l_p-norm optimization can be solved up to a given accuracy in polynomial time.

6 Concluding remarks

In this paper, we have formulated l_p-norm optimization problems in a conic way and applied results from the standard conic duality theory to derive their special duality properties.

This leads in our opinion to clearer proofs, the specificity of the class of problems under study being confined to the convex cone used in the formulation. Moreover, the fundamental reason why this class of optimization problems has better duality properties than a general convex problem becomes clear: this is essentially due to the existence of a strictly interior dual solution (even if a reduction procedure involving an equivalent regularized problem has to be introduced when the original dual lacks a strictly feasible point).

It is also worthy to note that this is an example of nonsymmetric conic duality, i.e. involving cones that are not self-dual, unlike the very well-studied cases of linear, second-order and semidefinite optimization.

Another advantage of this approach is the ease to prove polynomial complexity for our problems: finding a suitable self-concordant barrier is essentially all that is needed.

In the special case where all p_i's are equal, one might think it is possible to derive those duality results with a simpler formulation relying on the standard cone involving p-norms, i.e. the p-cone defined as

$$\mathbb{L}_p^n = \left\{ (x, \kappa) \in \mathbb{R}^n \times \mathbb{R}_+ \mid \|x\|_p \leq \kappa \right\} = \left\{ (x, \kappa) \in \mathbb{R}^n \times \mathbb{R}_+ \mid \sum_{i=1}^n |x_i|^p \leq \kappa^p \right\} .$$

However, we were note able to reach that goal, the reason being that the homogenizing variables θ and κ^* appear to play a significant role in our approach and cannot be avoided.

Finally, we mention that this framework can be easily applied to other classes of structured convex problems, see for example [2] for the case of geometric optimization.

Acknowledgements. The author wishes to thank Professor Kees Roos and the SSOR department at the Delft University of Technology for their kind hospitality.

References

[1] C. Roos D. den Hertog, F. Jarre 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, Guest editors: J.-Ph. Vial and J.-L. Goffin.

[2] Fr. Glineur, *Proving strong duality for geometric optimization using a conic formulation*, IMAGE Technical Report 9903, Faculté Polytechnique de Mons, Mons, Belgium, October 1999.

[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] S. Boyd M. S. Lobo, L. Vandenberghe and H. Lebret, *Applications of second-order cone programming*, Linear Algebra and its Applications **284** (1998), 193–228.

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

[6] 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.

[7] ———, *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.

[8] ———, *Geometric programming: Duality in quadratic programming and l_p approximation III*, Journal on Mathematical Analysis and Applications **29** (1970), 365–383.

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

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

[11] J. F. Sturm, *Primal-dual interior-point approach to semidefinite programming*, High performance optimization (T. Terlaky H. Frenk, C. Roos and S. Zhang, eds.), Applied optimization, vol. 33, Kluwer Academic Press, 2000.

[12] T. Terlaky, *On l_p programming*, European Journal of Operations Research **22** (1985), 70–100.

[13] G. Xue and Y. Ye, *An efficient algorithm for minimizing a sum of p-norms*, Working Paper, Department of Management Sciences 9903, The University of Iowa, Iowa City, U.S.A., September 1997.