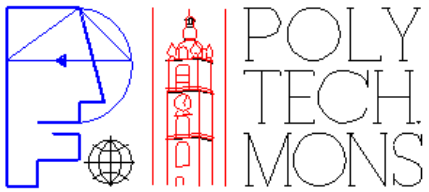
Strong duality in geometric optimization using a conic formulation

François Glineur

Francois.Glineur@fpms.ac.be



Faculté Polytechnique de Mons

Journées du 3ème cycle F.N.R.S.
en programmation mathématique

Han-sur-Lesse, February 4, 2000

Outline

Introduction

◇ Geometric optimization

Conic optimization

◇ Primal-dual pair

◇ Duality theorems

The geometric cone

◇ Definition and properties

◇ Dual cone

Geometric optimization

◇ Formulation with geometric cone

◇ Weak duality

◇ Strong duality

Concluding remarks

Introduction

Convex optimization

Let $f_i : \mathbb{R}^n \mapsto \mathbb{R}$ convex functions, $i = 0, \ldots, m$

$$\inf_{x \in \mathbb{R}^n} f_0(x) \quad \text{s.t.} \quad f_i(x) \leq 0 \text{ for } i = 1, \ldots, m$$

Properties

- ◇ All local optima are global
- ◇ Lagrange duality ⇒ nice dual problem

Conic optimization

Let $K \subseteq \mathbb{R}^n$ be a convex cone

$$\inf_{x \in \mathbb{R}^n} c^T x \quad \text{s.t.} \quad Ax = b \text{ and } x \in K$$

Equivalent to convex optimization.

Nice form for the dual:

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

where K^* is the *dual cone.*

Example: for LO, we have $K = \mathbb{R}^n_+ = K^*$.

Geometric optimization

Let $R = \{0, 1, 2, \ldots, r\}$, $I = \{1, 2, \ldots, n\}$;
$\{I_k\}_{k \in R}$ a partition of I into $r + 1$ classes.

Optimize m variables in vector $t \in \mathbb{R}_{++}$:

$$\inf \; G_0(t) \quad \text{s.t.} \quad G_k(t) \leq 1 \; \forall k \in R$$

where G_k's are *posynomials*:

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

with data $a_{ij} \in \mathbb{R}$ and $C_i \in \mathbb{R}_{++}$.

Example: $G(t_1, t_2, t_3) = 2\frac{t_1^2}{t_2} + 3\sqrt{t_2} + \frac{t_2^{2/3}}{3 t_1 t_3^3}$.

Many applications, especially in engineering (optimizing design parameters).

Generalization of linear optimization.

Not convex (take $G_0(t) = \sqrt{t_1}$) but can be *convexified* with a change of variables.

Convexification

W.l.o.g. we consider a **linear** objective and let

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

Primal:

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

with

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

(let $c_i = -\log C_i \Rightarrow g_k(y) = G_k(t)$ equivalence)

We optimize m free variables $y \in \mathbb{R}^m$.
Problem data is $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$.

Dual: (Lagrangean)

$$\begin{array}{ll} \inf & c^T x + \displaystyle\sum_{k \in R} \sum_{\substack{i \in I_k \\ x_i > 0}} x_i \log \frac{x_i}{\sum_{i \in I_k} x_i} \\ \text{s.t.} & Ax = b \text{ and } x \geq 0 \end{array}$$

Results (Duffin, Peterson and Zener, 1967)

- ◇ Convex problem ⇒ weak duality
- ◇ No duality gap !

Conic optimization

Let $K \subseteq \mathbb{R}^n$ be a *solid, pointed, closed* convex cone. Its dual cone is defined by

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

and is also convex, solid, pointed and closed.

Primal-dual pair

$$\inf_{x \in \mathbb{R}^n} \; c^T x \quad \text{s.t.} \quad Ax = b \text{ and } x \in K$$

$$\sup_{(y,s) \in \mathbb{R}^{m+n}} \; b^T y \quad \text{s.t.} \quad A^T y + s = c \text{ and } s \in K^*$$

Let p^* and d^* be their optimum objective values (not necessarily attained).

Duality theorems

Weak duality. We have $b^T y \leq c^T x$ for every feasible x and y. Equality occurs iff $x^T s = 0$ (*orthogonality* condition).

Hence we have in general $d^* \leq p^* \Rightarrow p^* - d^*$ is called the *duality gap* (always nonnegative).

Definition: x is *strictly feasible* $\Leftrightarrow$ x is feasible and $x \in \operatorname{int} K$

Strong duality (Slater condition)

If the primal has a strictly feasible point, the duality gap is zero, i.e. $p^* = d^*$.

Moreover, if the dual is feasible, its maximum objective value d^* is attained, i.e. $d^* = \max b^T y$

(a dualized result also holds)

Corollary

If both the primal and the dual have a strictly feasible point, the duality gap is zero and both optimal values are attained.

Our approach

Geometric optimization has nice duality properties, somewhere between the general convex case and linear optimization.

We are going to prove these duality properties using these general theorems $\Rightarrow$ we have to find a suitable convex cone.

The geometric cone

Definition

Let $n \in \mathbb{N}$. We define $\mathcal{G}^n$ as

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

with the convention when $\theta = 0$

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

This cone will enable us to express geometric optimization in a conic form.

Useful tool:

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

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

equality occurring iff all x_i's are equal.

(a.k.a. weighted arithmetic-geometric mean inequality)

Properties

- ◇ Special cases: $\mathcal{G}^0 = \mathbb{R}_+$ and $\mathcal{G}^1 = \mathbb{R}^2_+$
- ◇ If $(x, \theta) \in \mathcal{G}^n$, $(x', \theta') \in \mathcal{G}^n$ and $\lambda \geq 0$ we have $\lambda(x, \theta) \in \mathcal{G}^n$ and $(x + x', \theta + \theta') \in \mathcal{G}^n$ $\Rightarrow \mathcal{G}^n$ is a *convex cone*.
- ◇ $\mathcal{G}^n$ is closed, solid and pointed
- ◇ The interior of $\mathcal{G}^n$ is ($\rightarrow$ Slater condition)

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

Dual cone

The dual cone $(\mathcal{G}^n)^*$ is given by

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

It is the epigraph of

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

- ◇ Special cases: $(\mathcal{G}^0)^* = \mathbb{R}_+$ and $(\mathcal{G}^1)^* = \mathbb{R}^2_+$ (but $\mathcal{G}^n$ is not self-dual for $n > 1$)
- ◇ It is also convex, closed, solid and pointed.
- ◇ $((\mathcal{G}^n)^*)^* = \mathcal{G}^n$ (since $\mathcal{G}^n$ is closed).
- ◇ $\mathbb{R}^{n+1}_+ \subseteq (\mathcal{G}^n)^*$ (since $\mathcal{G}^n \subseteq \mathbb{R}^{n+1}_+$)
- ◇ The interior of $(\mathcal{G}^n)^*$ is given by

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

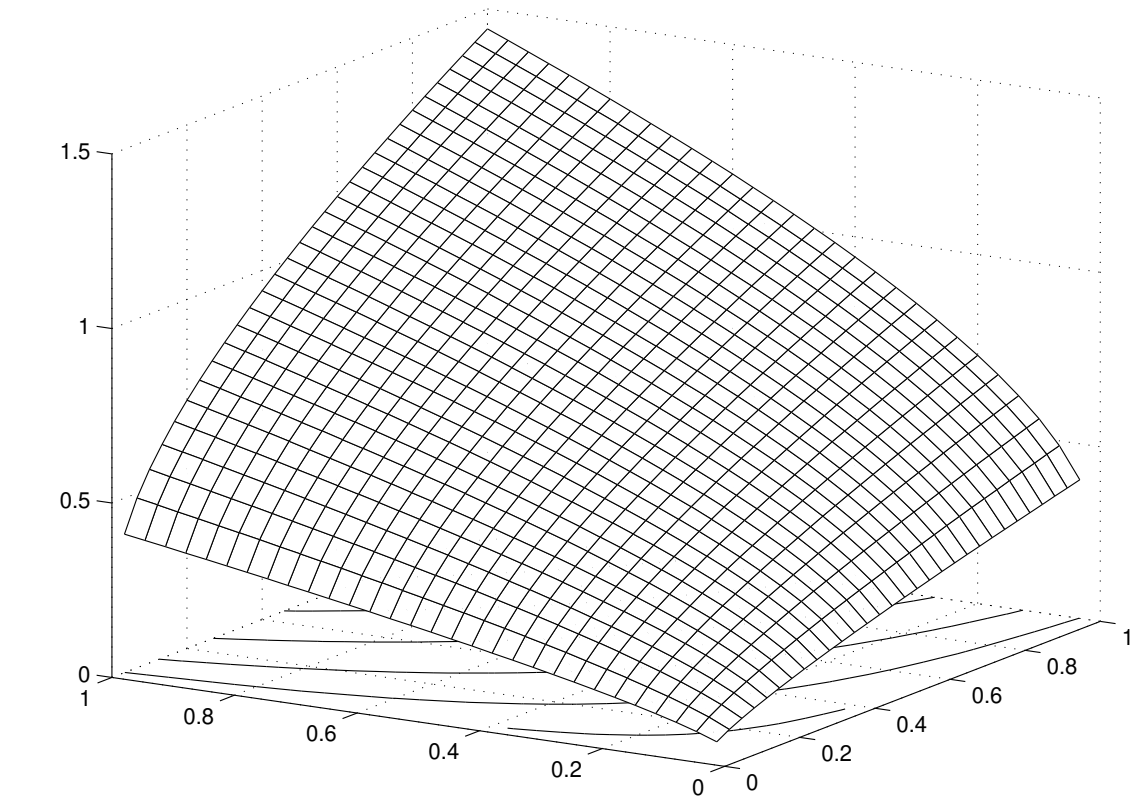
Orthogonality conditions

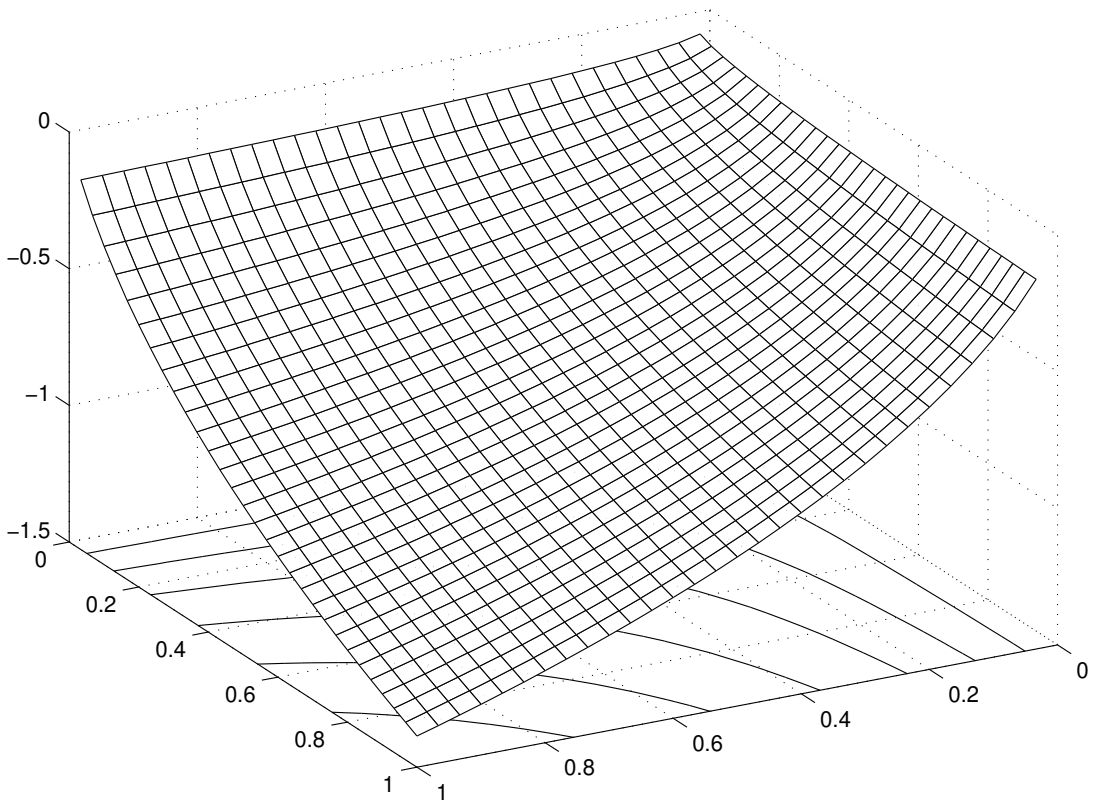
Let $v = (x, \theta) \in \mathcal{G}^n$ and $v^* = (x^*, \theta^*) \in (\mathcal{G}^n)^*$. We always have $v^T v^* \geq 0$, while $v^T v^* = 0$ occurs if and only if

$$\theta = 0 \text{ and } x_i x_i^* = 0 \text{ for all } i$$

or

$$\theta > 0 \text{ and } \begin{cases} \theta^* = \sum_{x_i^* > 0} x_i^* \log w_i^* \\ (\sum_{i=1}^n x_i^*) e^{-\frac{x_i}{\theta}} = x_i^* \text{ for all } i \end{cases}$$


1.5
1
0.5
0
1
0.8
0.6
0.4
0.2
0
0
0.2
0.4
0.6
0.8
1


0
−0.5
−1
−1.5
0
0.2
0.4
0.6
0.8
1
1
0.8
0.6
0.4
0.2
0

We are now ready to apply the general duality theory for conic primal-dual pairs, using our dual cones $\mathcal{G}^n$ and $(\mathcal{G}^n)^*$, to derive the duality properties of the geometric optimization primal-dual pairs.

Notation: v_I (resp.M_I) $\equiv$ restriction of vector v (resp.matrix M) to indices belonging to I.

Strategy diagram

$$
\begin{array}{ccccccc}
(PG) & \equiv & (CPG) & \overset{\text{Weak}}{\longleftrightarrow} & (CDG) & \equiv & (DG) \\
 & & \updownarrow^{*} & & \updownarrow & & \\
 & & (RPG) & \overset{\text{Strong}}{\longleftrightarrow} & (RDG) & & \\
 & & & & \uparrow & & \\
 & & & & \text{(Slater)} & &
\end{array}
$$

Geometric optimization

Formulation with $\mathcal{G}^n$ cone

Primal

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

Introducing variables $s_i = c_i - a_i^T y \ \forall i$ we get

$$\begin{array}{lll} \sup b^T y & \text{s.t.} & s = c - A^T y \\ & \text{and} & \displaystyle\sum_{i \in I_k} e^{-s_i} \leq 1 \text{ for all } k \in R \end{array}$$

$$\Updownarrow$$

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

where e is the all-one vector, $n_k = \#I_k$ and v_i's are fictitious variables.

This is a standard conic problem with variables $(\tilde{y}, \tilde{s})$, data $(\tilde{A}, \tilde{b}, \tilde{c})$ and a cone K^*

Indeed we have

$$\tilde{y} = y,\ \tilde{s} = \begin{pmatrix} s \\ v \end{pmatrix},\ \tilde{A} = \begin{pmatrix} A & 0 \end{pmatrix},\ \tilde{b} = b,$$

$$\tilde{c} = \begin{pmatrix} c \\ e \end{pmatrix} \text{ and } K^* = \mathcal{G}^{n_1} \times \mathcal{G}^{n_2} \times \cdots \times \mathcal{G}^{n_r}\ .$$

$\Rightarrow$ we can mechanically derive the **dual** !

$$\inf \begin{pmatrix} c \\ e \end{pmatrix}^T \begin{pmatrix} x \\ z \end{pmatrix} \qquad \text{s.t.} \qquad \begin{pmatrix} A & 0 \end{pmatrix} \begin{pmatrix} x \\ z \end{pmatrix} = b$$

$$\text{and} \qquad (x_{I_k}, z_k) \in (\mathcal{G}^{n_k})^* \ \forall k$$

$\Updownarrow$

$$\inf c^T x + e^T z \qquad \text{s.t.} \qquad Ax = b,\ x_{I_k} \geq 0$$

$$\text{and} \qquad z_k \geq \sum_{\substack{i \in I_k \\ x_i > 0}} x_i \log \frac{x_i}{\sum_{i \in I_k} x_i}$$

$$\Leftrightarrow \qquad \inf \qquad c^T x + \sum_{k \in R} \sum_{\substack{i \in I_k \\ x_i > 0}} x_i \log \frac{x_i}{\sum_{i \in I_k} x_i}$$

$$\text{s.t.} \qquad Ax = b \text{ and } x \geq 0$$

Weak duality

If y is feasible for the primal and x is feasible for the dual, we have

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

Equality occurs if and only if

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

Proof. Use weak duality theorem on conic optimization primal-dual pair and extend objective values to the geometric optimization primal-dual pair.

- ◇ This proof and the corresponding classical proof are not too difficult.
- ◇ This property is a merely a consequence of convexity

Strong duality

If the primal and the dual are feasible, their optimum objective values are equal (but not necessarily attained).

This strong duality property is not valid for all convex problems but depends on the *specific structure* of this class of problems.

Proof.

Proceed by proving the existence of a strictly feasible point for the dual conic program $\Leftrightarrow z_k > \sum_{i \in I_k} x_i \log \frac{x_i}{\sum_{i \in I_k} x_i}$ **and** $x_i > 0 \ \forall i$.

But the linear constraints $Ax = b$ may force $x_i = 0$ for some i for every feasible solution !

$\Rightarrow$ detect these zero x_i components and form a restricted primal-dual pair without these variables (which had no influence on the objective/constraints anyway).

Detection. Use the following linear problem

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

Define $\mathcal{N}$ = set of indices i such that x_i is identically zero on the feasible region and $\mathcal{B}$ the set of the other indices.

$(\mathcal{B}, \mathcal{N})$ is the optimal partition of this linear problem (Goldman-Tucker theorem).

Strategy. Remove variables x_i for all $i \in \mathcal{N}$
$\Rightarrow$ restricted primal-dual conic pair
$\Rightarrow$ the dual has a strictly feasible solution
$\Rightarrow$ zero duality gap.

There remains to prove that

- $\diamond$ Optimal objective values are equal for restricted and original dual problems (easy)
- $\diamond$ Optimal values are equal for restricted and original primal problems (more difficult). Moreover, attainment is lost in the process.

Restricted problems

Dual

$$\begin{array}{ll} \inf & c_{\mathcal{B}}^T x_{\mathcal{B}} + \displaystyle\sum_{k \in R} \sum_{\substack{i \in I_k \cap \mathcal{B} \\ x_i > 0}} x_i \log \frac{x_i}{\sum_{i \in I_k \cap \mathcal{B}} x_i} \\ \text{s.t.} & A_{\mathcal{B}} x_{\mathcal{B}} = b \text{ and } x_{\mathcal{B}} \geq 0 \end{array}$$

Convert solution to original dual (which has additional variables): let $x_i = 0$ for all $i \in \mathcal{N} \Rightarrow$ same objective.

Primal

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

with

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

Convert solution to original dual is *not straightforward* because feasibility cannot be guaranteed: posynomials have additional terms !

We need to *modify* the optimal solution of the restricted dual.

Strategy

Perturb the optimal solution $\tilde{y}$ of the restricted primal with a well-chosen vector.

The *Goldman-Tucker theorem* implies the existence of a strictly complementary solution to any primal-dual pair of linear problems.

$\Rightarrow$ applied to our detection program and its dual, this guarantees the existence of a vector $\bar{y}$ such that

$$A_{\mathcal{B}}^T \bar{y} = 0 \text{ and } A_{\mathcal{N}}^T \bar{y} < 0 \ .$$

These properties imply that $y = \tilde{y} + \lambda \bar{y}$ is asymptotically feasible for the original primal when λ tends to $+\infty$, while keeping the same objective value as $\tilde{y}$. Combining with another trick (mixing with a feasible solution) leads to the expected result

$\Rightarrow$ the original primal optimum objective value is equal to the original dual optimum objective value.

Concluding remarks

- ◇ These proofs are shorter than the classical ones. All the specificity of the problems is confined to the pair of dual cones
- ◇ Other classical results can also be recovered easily (posynomial objective, notion of subinfimum, etc.)
- ◇ This strategy can be applied to other types of convex problems, i.e. entropy optimization and l_p-norm optimization
- ◇ Finding a computable *self-concordant* barrier for the $\mathcal{G}^n$ cone provides an algorithm with polynomial complexity (short-step path-following interior-point method, Nesterov & Nemirovsky)
- ◇ This *primal-dual pair* of cones might prove helpful in designing efficient algorithms (cf. *self-scaled cones*)