A conic approach for separable convex optimization

François Glineur

Aspirant F.N.R.S.

Francois.Glineur@fpms.ac.be

Faculté Polytechnique de Mons

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

Han-sur-Lesse, February 2, 2001

Outline

Introduction

- ◇ Conic optimization
- ◇ Geometric optimization

Geometric optimization

- ◇ The geometric cone
- ◇ A conic formulation

Generalization

- ◇ The separable cone
- ◇ Separable convex optimization

Concluding remarks

Introduction

Convex optimization

Let $f_0 : \mathbb{R}^n \mapsto \mathbb{R}$ a convex function and $C \subseteq \mathbb{R}^n$ a convex set

$$\inf_{x \in \mathbb{R}^n} f_0(x) \quad \text{s.t.} \quad x \in C$$

Properties

- ◇ All local optima are global
- ◇ Optimum set is convex
- ◇ Lagrange duality $\Rightarrow$ related dual problem
- ◇ Efficient *interior-point* methods

Conic optimization

Let $\mathcal{C} \subseteq \mathbb{R}^n$ a closed convex cone

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

Equivalent to convex optimization

Primal-dual pair

Dual cone

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

is also a closed convex cone

$\Rightarrow$ symmetrical pair of primal-dual problems

$$\begin{aligned} \inf_{x \in \mathbb{R}^n} c^T x \quad & \text{s.t.} \quad Ax = b \text{ and } x \in \mathcal{C} \\ \sup_{(y,s) \in \mathbb{R}^{m+n}} b^T y \quad & \text{s.t.} \quad A^T y + s = c \text{ and } s \in \mathcal{C}^* \end{aligned}$$

Examples

$\mathcal{C} = \mathbb{R}^n_+ = \mathcal{C}^* \Rightarrow$ linear optimization,
$\mathcal{C} = \mathbb{S}^n_+ = \mathcal{C}^* \Rightarrow$ semidefinite optimization
(self-duality)

Advantages over classical formulation

- ◇ Remarkable primal-dual symmetry
- ◇ Special handling of (*easy*) linear equality constraints

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.

A *posynomial* is a sum of positive monomials

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

defined by data $a_{ij} \in \mathbb{R}$ and $C_i \in \mathbb{R}_{++}$

Example: $G(t_1, t_2, t_3) = 2\frac{t_1^2}{t_2} + 3\sqrt{t_2} + \frac{1}{3}\frac{{t_2^2}^{/3}}{t_1 t_3^3}$

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

$$\inf_{t \in \mathbb{R}^m_{++}} G_0(t) \quad \text{s.t.} \quad G_k(t) \leq 1 \; \forall k \in R$$

where G_k's are *posynomials*: .

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:

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

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)

Dual: (Lagrangean)

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

Results (Duffin, Peterson and Zener, 1967)

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

The geometric cone

Objective: model geometric optimization within the conic framework

DEFINITION *[Gl. 99]*

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

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

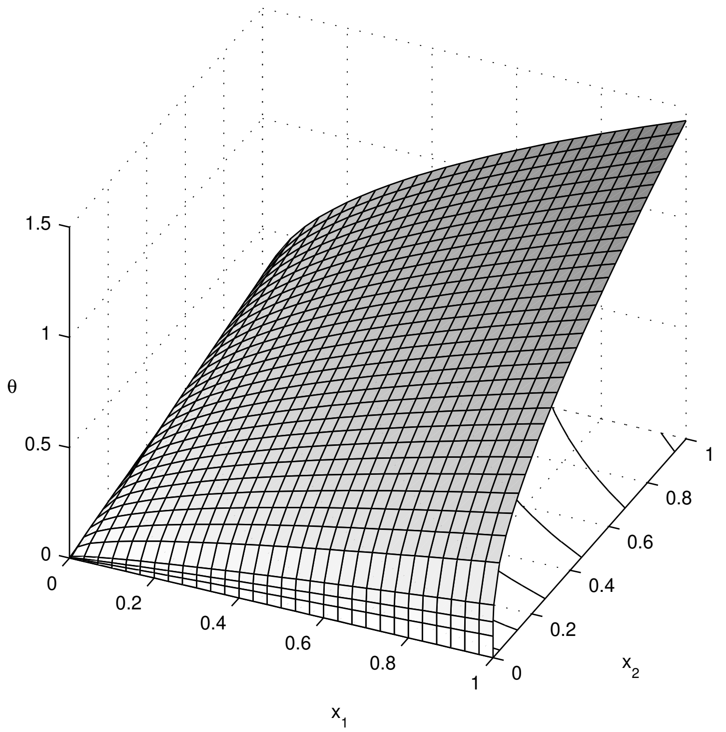
with the convention when $\theta = 0$

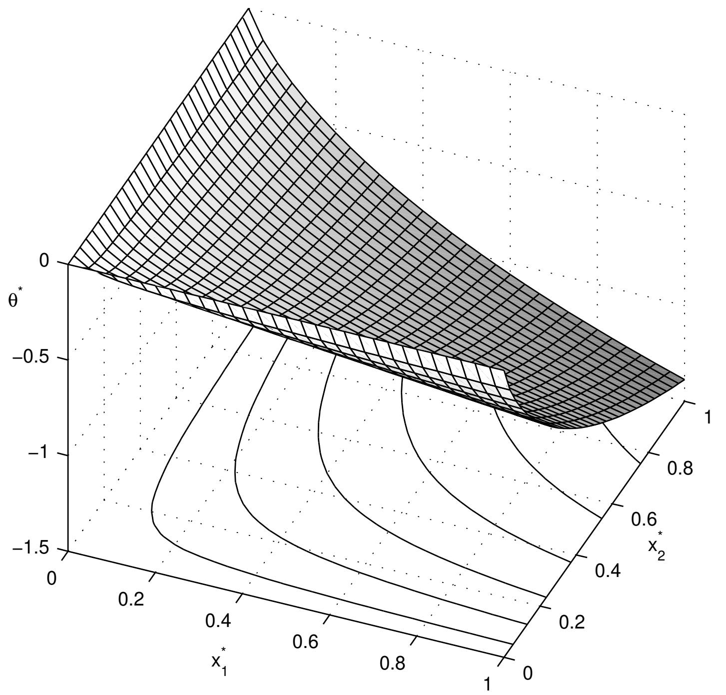
$$e^{-\frac{x_i}{0}} = 0$$

PROPERTIES

- ⋄ Special cases: $\mathcal{G}^0 = \mathbb{R}_+$ and $\mathcal{G}^1 = \mathbb{R}^2_+$
- ⋄ $\mathcal{G}^n$ is a *closed convex cone.*
- ⋄ $(\mathcal{G}^n)^* \subseteq \mathbb{R}^n_+ \times \mathbb{R}$ is

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


1.5
1
θ
0.5
0
0
0.2
0.4
0.6
0.8
1
x_1
0
0.2
0.4
0.6
0.8
1
x_2


0
θ*
−0.5
−1
−1.5
0
0.2
0.4
0.6
0.8
1
x_1^*
0
0.2
0.4
0.6
0.8
1
x_2^*

Geometric optimization

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

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

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

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

$\Updownarrow$

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

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 $\mathcal{C}^*$

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

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

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

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

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

$\Updownarrow$

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

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

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

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

Generalizing our framework

Comparing cones

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

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

Variants

$$\mathcal{G}^n_2 = \left\{ (x, \theta, \kappa) \in \mathbb{R}^n_+ \times \mathbb{R}^2_+ \mid \theta \sum_{i=1}^n e^{-\frac{x_i}{\theta}} \leq \kappa \right\}$$

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

The separable cone *[Gl. 00]*

Consider a set of n scalar closed proper convex functions

$$f_i : \mathbb{R} \mapsto \mathbb{R}$$

and define $\mathcal{K}^f$ as

$$\mathrm{cl}\Big\{(x, \theta, \kappa) \in \mathbb{R}^n \times \mathbb{R}_{++} \times \mathbb{R} \mid \theta \sum_{i=1}^{n} f_i(\frac{x_i}{\theta}) \leq \kappa\Big\}$$

- ◇ $\mathcal{K}^f$ generalizes $\mathcal{L}^p$ and $\mathcal{G}_2^n$
- ◇ $\mathcal{K}^f$ is a closed convex cone
- ◇ The dual $(\mathcal{K}^f)^* \subseteq \mathbb{R}^n \times \mathbb{R}_{++} \times \mathbb{R}$ is

$$\mathrm{cl}\Big\{(x^*, \theta^*, \kappa^*) \mid \kappa^* \sum_{i=1}^{n} f_i^*(-\frac{x_i^*}{\kappa^*}) \leq \theta^*\Big\}$$

using the conjugate functions

$$f_i^* : x^* \mapsto \sup_{x \in \mathbb{R}^n} \{x^T x^* - f_i(x)\}$$

(also closed, proper and convex)

Separable convex optimization

Primal-dual conic pair with $\mathcal{K}^f$ and $(\mathcal{K}^f)^* \rightarrow$

Primal

$$\sup\ b^T y \ \text{ s.t. } \sum_{i \in I_k} f_i(c_i - a_i^T y) \leq d_k - g_k^T y \quad \forall k \in K$$

Dual

$$\inf \quad \Big\{ c^T x + d^T z + \sum_{k \in K \,|\, z_k > 0} z_k \sum_{i \in I_k} f_i^*(-\frac{x_i}{z_k}) - \sum_{k \in K \,|\, z_k = 0} \inf_{x^*_{I_k} \in \mathrm{dom}\, f_{I_k}} x_{I_k}^T x^*_{I_k} \Big\}$$

$$\text{s.t.} \quad Ax + Gz = b \text{ and } z \geq 0 \,.$$

- ◇ Generalizes linear, quadratic, geometric, entropy, l_p-norm optimization and others ...
- ◇ Mix different types of constraints
- ◇ Justification for *conventions* when $\theta = 0$
- ◇ Hints towards a primal-dual setting

Some other types of constraints

◇

$$f_1 : x \mapsto \begin{cases} -\sqrt{a^2 - x^2} & \text{if } |x| \leq a \\ +\infty & \text{if } |x| > a \end{cases}$$

$$f_1^* : x^* \mapsto a\sqrt{1 + x^{*2}}$$

◇

$$f_2 : x \mapsto \begin{cases} -\frac{1}{p} x^p & \text{if } x \geq 0 \\ +\infty & \text{if } x < 0 \end{cases} \quad 0 < p < 1$$

$$f_2^* : x^* \mapsto \begin{cases} -\frac{1}{q}(-x^*)^q & \text{if } x^* < 0 \\ +\infty & \text{if } x^* \geq 0 \end{cases} \quad q < 0$$

◇

$$f_3 : x \mapsto \begin{cases} -\frac{1}{2} - \log x & \text{if } x > 0 \\ +\infty & \text{if } x \leq 0 \end{cases}$$

$$f_3^* : x^* \mapsto \begin{cases} -\frac{1}{2} - \log(-x^*) & \text{if } x^* < 0 \\ +\infty & \text{if } x^* \geq 0 \end{cases}$$

(with property that $f^*(x^*) = f(-x^*)$)

Concluding remarks

- A very wide class of problems can be formulated (links with extended geometric programming *[Peterson 80]*)
- Using this setting, duality properties can be obtained easily (weak and strong duality without Slater condition)
- Finding a computable *self-concordant* barrier for the $\mathcal{K}^f$ cone provides a primal algorithm with polynomial complexity (short-step path-following interior-point method, *[Nesterov & Nemirovsky 83]*)
- Primal-dual formulation $\Rightarrow$ first step towards *true* primal-dual algorithms (*self-regular* barriers, *[Peng et al. 00]*)