Topics in Convex Optimization: Interior-Point Methods, Conic Duality and Approximations

François Glineur

Aspirant F.N.R.S.
Faculté Polytechnique de Mons

Ph.D. dissertation
January 11, 2001

Co-directed by J. Teghem
T. Terlaky

Motivation

Operations research

Model real-life situations to help take the *best* decisions

$$\left.\begin{array}{rcl} \text{Decision} & \leftrightarrow & \text{vector of variables} \\ \text{Best} & \leftrightarrow & \text{objective function} \\ \text{Constraints} & \leftrightarrow & \text{feasible set} \end{array}\right\} \Rightarrow \text{Optimization}$$

Choice of design parameters, scheduling, planification

Two approaches

Solving all problems *efficiently* is impossible in practice!

Optimal method to minimize of Lipschitz-continuous f:
$L = 2$, 10 variables, 1% accuracy $\Rightarrow 10^{20}$ operations

Reaction: two distinct orientations

- ◇ General nonlinear optimization
  Applicable to all problems but no efficiency guarantee
- ◇ Linear, quadratic, semidefinite, . . . optimization
  Restrict set of problems to get efficiency guarantee

Tradeoff generality $\leftrightarrow$ efficiency (algorithmic complexity)

Restrict to which class of problems ?

Linear optimization : $\begin{array}{ll} + & \text{specialized, very fast algorithms} \\ - & \text{too restricted in practice} \end{array}$

$\longrightarrow$ we focus on **Convex optimization**

- ◇ Convex objective and convex feasible set
- ◇ Many problems are convex or can be convexified
- ◇ Efficient algorithms and powerful duality theory
- ◇ Establishing convexity *a priori* is difficult

  $\longrightarrow$ work with specific classes of convex constraints: *Structured* convex optimization (convexity by design)

Reward for a convex formulation is algorithmic efficiency

Overview of the thesis

Interior-point methods

- ◇ Linear optimization survey
- ◇ Self-concordant functions

Conic optimization

- ◇ Formulation and duality
- ◇ Geometric and l_p-norm optimization
- ◇ General framework: separable optimization

Approximations

- ◇ Geometric optimization with l_p-norm optimization
- ◇ Linearizing second-order cone optimization

Overview of this talk

Interior-point methods

- ◇ Linear optimization survey
- ◇ Self-concordant functions

Conic optimization

- ◇ Formulation and duality
- ◇ Geometric and l_p-norm optimization
- ◇ General framework: separable optimization

Approximations

- ◇ Geometric optimization with l_p-norm optimization
- ◇ Linearizing second-order cone optimization

Self-concordant functions: the key to efficient algorithms for convex optimization

(chapter 2)

Interior-point methods

◇ Self-concordant functions

Conic optimization

◇ Formulation and duality

◇ Geometric and l_p-norm optimization

◇ General framework: separable optimization

Approximations

◇ Geometric optimization with l_p-norm optimization

Convex optimization

Let $f_0 : \mathbb{R}^n \mapsto \mathbb{R}$ be a convex function, $C \subseteq \mathbb{R}^n$ be a convex set : optimize a vector $x \in \mathbb{R}^n$

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

Properties

- All local optima are global, optimal set is convex
- Lagrange duality $\longrightarrow$ strongly related dual problem
- Objective can be taken linear w.l.o.g. $(f_0(x) = c^T x)$

Defining a problem

Two distinct approaches

a. **List of convex constraints.**

m convex functions $f_i : \mathbb{R}^n \mapsto \mathbb{R}$, $i = 1, 2, \dots, m$

$$C = \{x \in \mathbb{R}^n \mid f_i(x) \leq 0 \text{ for all } i = 1, 2, \dots, m\}$$

(intersection of convex level sets)

$$\inf_{x \in \mathbb{R}^n} f_0(x) \quad \text{s.t.} \quad f_i(x) \leq 0 \text{ for all } i = 1, 2, \dots, m$$

b. **Use a barrier function.**

Feasible set $\equiv$ domain of a *barrier* function F s.t.

- F is smooth
- F is strongly convex $\operatorname{int} C$
- $F(x) \to +\infty$ when $x \to \partial C$

$$\to \quad C = \operatorname{cl} \operatorname{dom} F = \operatorname{cl} \{x \in \mathbb{R}^n \mid F(x) < +\infty\}$$

Interior-point methods

Principle

Approximate a constrained problem by a *family* of unconstrained problems based on F

Let $\mu \in \mathbb{R}_{++}$ be a parameter and consider

$$\inf_{x \in \mathbb{R}^n} \frac{c^T x}{\mu} + F(x) \qquad (\mathrm{P}_\mu)$$

We have

$$x^*_\mu \to x^* \text{ when } \mu \searrow 0$$

where

- x^*_μ is the (unique) solution of (P_μ) ($\to$ central path)
- x^* is a solution of the original problem (P)

Ingredients

- ◇ A method for unconstrained optimization
- ◇ A barrier function

Interior-point methods rely on

- ◇ *Newton's method* to compute x^*_μ
- ◇ When C is defined with nonlinear functions f_i, one can introduce the *logarithmic* barrier function

$$F(x) = -\sum_{i=1}^n \ln(-f_i(x))$$

Question: What is a good barrier, i.e. a barrier for which Newton's method is efficient ?

Answer: A *self-concordant* barrier

Self-concordant barriers

Definition ***[Nesterov & Nemirovsky, 1988]***

$F : \operatorname{int} C \mapsto \mathbb{R}$ is called (κ, ν)-self-concordant on C iff

- F is convex
- F is three times differentiable
- $F(x) \to +\infty$ when $x \to \partial C$
- the following two conditions hold

$$\nabla^3 F(x)[h,h,h] \leq 2\kappa \left(\nabla^2 F(x)[h,h]\right)^{\frac{3}{2}}$$
$$\nabla F(x)^T (\nabla^2 F(x))^{-1} \nabla F(x) \leq \nu$$

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

Alternative definition

Let $x \in \operatorname{int} C$ and $h \in \mathbb{R}^n$ and define a restriction

$$F_{x,h}(t) : \mathbb{R} \mapsto \mathbb{R} : t \mapsto F(x + th)$$

Replace conditions involving differentials by

$$F'''_{x,h}(0) \leq \kappa F''_{x,h}(0)^{\frac{3}{2}} \text{ and } F'_{x,h}(0)^2 \leq \nu F''_{x,h}(0)$$

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

Scaling and summation

Let $\lambda \in \mathbb{R}_+$ be a positive scalar

$$F \text{ is } (\kappa, \nu)\text{-SC} \Leftrightarrow \lambda F \text{ is } (\frac{\kappa}{\sqrt{\lambda}}, \lambda\nu)\text{-SC}$$

Let F_1 be (κ_1, ν_1)-SC and F_2 be (κ_2, ν_2)-SC

$$F_1 + F_2 \text{ is } (\max\{\kappa_1, \kappa_2\}, \nu_1 + \nu_2)\text{-SC}$$

Complexity result

Summary

Self-concordant barrier $\Rightarrow$ polynomial number of iterations to solve (P) within a given accuracy

Principle of a short-step method

- ◇ Define a *proximity measure* $\delta(x, \mu)$ to central path
- ◇ Choose a starting iterate with a small $\delta(x_0, \mu_0)$
- ◇ While accuracy is not attained

  a. Decrease μ geometrically (δ increases)

  b. Take a Newton step to minimize barrier (δ decreases and is restored)

Geometric interpretation

Two self-concordancy conditions: each has its role

- ◇ First condition bounds the variation of the Hessian $\Rightarrow$ controls the increase of the proximity measure when μ is updated
- ◇ Second condition bounds the size of the Newton step $\Rightarrow$ guarantees that the Newton step restores the initial proximity to the central path

Complexity result

$$\mathcal{O}\left(\kappa\sqrt{\nu}\log\frac{1}{\epsilon}\right)$$

iterations lead a solution with ϵ accuracy on the objective

Optimal complexity result *[Glineur 00]*

Optimal values for two constants

- ◇ (maximum) proximity δ to the central path
- ◇ Constant of decrease of barrier parameter μ

lead to

$$\left\lceil \left(1.03 + 7.15\kappa\sqrt{\nu}\right) \log \frac{1.29\mu_0\kappa\sqrt{\nu}}{\epsilon} \right\rceil$$

iterations for a solution with ϵ accuracy

Two self-concordancy parameters

Complexity $\kappa\sqrt{\nu}$ invariant w.r.t. to scaling of $F \Rightarrow$ one of the constants κ and ν can be arbitrarily fixed

If there exists a (κ, ν)-SC barrier F for C then it can be scaled to get a

- $(\kappa\sqrt{\nu}, 1)$-SC barrier or a
- $(1, \kappa^2\nu)$-SC barrier

Comparison *[Glineur 00]*

When C is defined by f_i's, it is typical to use the first scaling ($\nu = 1$) with the logarithmic barrier

Indeed, if

$$F_i : \mathbb{R}^n \mapsto \mathbb{R} : x \mapsto -\ln(-f_i(x))$$

satisfies the first condition with $\kappa = \kappa_i$ then

$$F_i \text{ is } (\kappa_i, 1)\text{-self-concordant}$$

because the second ν condition is automatically satisfied with $\nu = 1$ if f_i is convex.
This implies in the end that

$$F = \sum_{i=1}^{m} F_i \text{ is } (\kappa, m)\text{-SC with } \kappa = \max_{i=1,\ldots,m} \kappa_i$$

and that the problem can be solved in

$$O(\sqrt{m} \max_{i=1,\ldots,m} \kappa_i) = O(\sqrt{m}\, \|\kappa\|_\infty) \text{ iterations}$$

However, the second scaling ($\kappa = 1$) is superior !

Indeed, we have then that $\kappa_i^2 F_i$ is $(1, \kappa_i^2)$-SC which implies that

$$F = \sum_{i=1}^{m} \kappa_i^2 F_i \text{ is } (1, \nu)\text{-SC with } \nu = \sum_{i=1}^{m} \kappa_i^2$$

and that the problem can be solved in

$$O(\sqrt{\sum_{i=1}^{m} \kappa_i^2}) = O(\|\kappa\|_2) \text{ iterations}$$

which is always better since

$$\|\kappa\|_2 \leq \sqrt{m}\,\|\kappa\|_\infty$$

(strict inequality when κ_i's not all equal)

A useful lemma

Proving self-concordancy not always an easy task
$\Rightarrow$ improved version of lemma by *[Den Hertog et al.]*

Auxiliary functions

Let two increasing functions (see Figure 1)

$$r_1 : \mathbb{R} \mapsto \mathbb{R} : \gamma \mapsto \max\{1, \frac{\gamma}{\sqrt{3 - 2/\gamma}}\}$$

$$r_2 : \mathbb{R} \mapsto \mathbb{R} : \gamma \mapsto \max\{1, \frac{\gamma + 1 + 1/\gamma}{\sqrt{3 + 4/\gamma + 2/\gamma^2}}\}$$

We have $r_1(\gamma) \approx \frac{\gamma}{\sqrt{3}}$ and $r_2(\gamma) \approx \frac{\gamma+1}{\sqrt{3}}$ when $\gamma \to +\infty$.

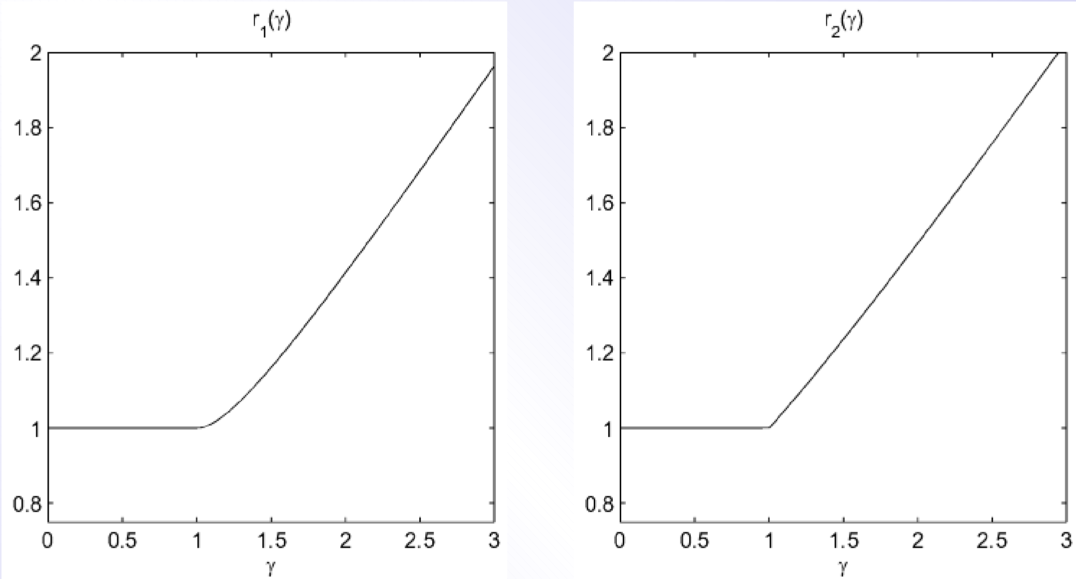

Figure 1: Graphs of functions r_1 and r_2

Lemma's statement *[Glineur 00]*

Let $F : \mathbb{R}^n \mapsto \mathbb{R}$ be a convex function on C.

If there is a constant $\gamma \in \mathbb{R}_+$ such that

$$\nabla^3 F(x)[h,h,h] \leq 3\gamma \nabla^2 F(x)[h,h] \sqrt{\sum_{i=1}^{n} \frac{h_i^2}{x_i^2}}$$

then the following barrier functions

$$F_1 : \mathbb{R}^n \mapsto \mathbb{R} : x \mapsto F(x) - \sum_{i=1}^{n} \ln x_i$$

$$F_2 : \mathbb{R}^n \times \mathbb{R} \mapsto \mathbb{R} : (x,u) \mapsto -\ln(u - F(x)) - \sum_{i=1}^{n} \ln x_i$$

satisfy the first self-concordancy condition with

$$\begin{array}{ll} \kappa_1 = r_1(\gamma) & \text{for } F_1 \text{ on } C \\ \kappa_2 = r_2(\gamma) & \text{for } F_2 \text{ on } \operatorname{epi} F = \{(x,u) \mid F(x) \leq u\} \end{array}$$

A structured convex problem

Extended entropy optimization

$$\min c^T x + \sum_{i=1}^{n} g_i(x_i) \quad \text{s.t.} \quad Ax = b \text{ and } x \geq 0$$

with scalar functions $g_i : \mathbb{R} \mapsto \mathbb{R}$ such that

$$|g_i'''(x)| \leq \kappa_i \frac{g_i''(x)}{x} \; \forall x \geq 0$$

(which implies convexity)

Special case: classical entropy optimization
when $g_i(x) = x \log x \quad \Rightarrow \quad \kappa_i = 1$

Application of the Lemma

Use Lemma with $F(x_i) = g_i(x_i)$ to prove that

$$-\ln\Big(t_i - g_i(x_i)\Big) - \ln(x_i) \;\text{ is }\; \Big(r_2(\frac{\kappa_i}{3}), 2\Big)\text{-SC}$$

Total complexity of EEO is *[Glineur 00]*

$$O\Big(\sqrt{2\sum_{i=1}^{n} r_2(\frac{\kappa_i}{3})^2}\Big) \text{ iterations}$$

or

$$O(\sqrt{2n}) \text{ iterations for entropy optimization}$$

Possible application: *polynomial* g_i's

Conic optimization: an elegant framework to formulate convex problems and study their duality properties

(chapter 3)

Interior-point methods

◇ Self-concordant functions

Conic optimization

◇ Formulation and duality

◇ Geometric and l_p-norm optimization

◇ General framework: separable optimization

Approximations

◇ Geometric optimization with l_p-norm optimization

Conic formulation

Primal problem

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

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

Formulation is equivalent to convex optimization.

Dual problem

Let $\mathcal{C} \subseteq \mathbb{R}^n$ be a *solid, pointed, closed* convex cone.
The 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 convex, solid, pointed and closed $\rightarrow$ dual problem:

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

Primal-dual pair

Symmetrical pair of primal-dual problems

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

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

Optimum values p^* and d^* not necessarily attained !

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

Weak duality

For every feasible x and y $\quad b^T y \leq c^T x$
with equality iff $x^T s = 0$ (*orthogonality* condition)

$\Delta = p^* - d^*$ is the *duality gap* $\Rightarrow$ always nonnegative

Definition: x *strictly feasible* $\Leftrightarrow$ x feasible and $x \in \operatorname{int} \mathcal{C}$

Strong duality (with Slater condition)

a. Strictly feasible dual point $\Rightarrow p^* = d^*$

b. If in addition primal is bounded
$\Rightarrow$ primal optimum is attained $\Leftrightarrow p^* = \min c^T x$

(dualized result obviously holds)

Corollary

Primal and dual Slater $\Rightarrow \min c^T x = p^* = d^* = \max b^T y$

Multiple cones

$x^i \in \mathcal{C}^i$ for all $i \in \{1, 2, \ldots, k\} \Rightarrow \mathcal{C} = \mathcal{C}^1 \times \mathcal{C}^2 \times \cdots \mathcal{C}^k$

Our approach

- ◇ Duality for general convex optimization weaker than for linear optimization (need Slater condition)
- ◇ But some classes of structured convex optimization problems feature better duality properties (i.e. zero duality gap even without Slater condition)

Our goal: prove these duality properties using general theorems for conic optimization $\Rightarrow$ new convex cones

A conic formulation for two well-known classes of problems: geometric and l_p-norm optimization

(chapters 4–5)

Interior-point methods

◇ Self-concordant functions

Conic optimization

◇ Formulation and duality

◇ Geometric and l_p-norm optimization

◇ General framework: separable optimization

Approximations

◇ Geometric optimization with l_p-norm optimization

Geometric optimization

Posynomials

Let $K = \{0, 1, 2, \ldots, r\}$, $I = \{1, 2, \ldots, n\}$;
let $\{I_k\}_{k \in K}$ 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: $G(t_1, t_2, t_3) = 2\frac{t_1^2}{t_2} + 3\sqrt{t_2} + \frac{t_2^{2/3}}{3t_1t_3^3}$
Many applications, especially in engineering
(optimizing design parameters, modelling power laws)

Primal problem

Optimize m variables in vector $t \in \mathbb{R}^m_{++}$

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

Not convex: take $G_0(t) = \sqrt{t_1}$

Convexification

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

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

$\Rightarrow$ we let

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

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

Convexified primal

Free variables $y \in \mathbb{R}^m$, data $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$

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

(Lagrangean) dual

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

Properties *[Duffin, Peterson and Zener, 1967]*

- Convex problem $\Rightarrow$ weak duality
- No duality gap !

The geometric cone

Definition *[Glineur 99]*

Let $n \in \mathbb{N}$. 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 $e^{-\frac{x_i}{0}} = 0$

Our goal: express geometric optimization in a conic form

Properties

- Special cases: $\mathcal{G}^0 = \mathbb{R}_+$ and $\mathcal{G}^1 = \mathbb{R}^2_+$
- $(x, \theta) \in \mathcal{G}^n$, $(x', \theta') \in \mathcal{G}^n$ and $\lambda \geq 0$
  $\Rightarrow \quad \lambda(x, \theta) \in \mathcal{G}^n \quad$ and $\quad (x + x', \theta + \theta') \in \mathcal{G}^n$
  $\Rightarrow \mathcal{G}^n$ 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).

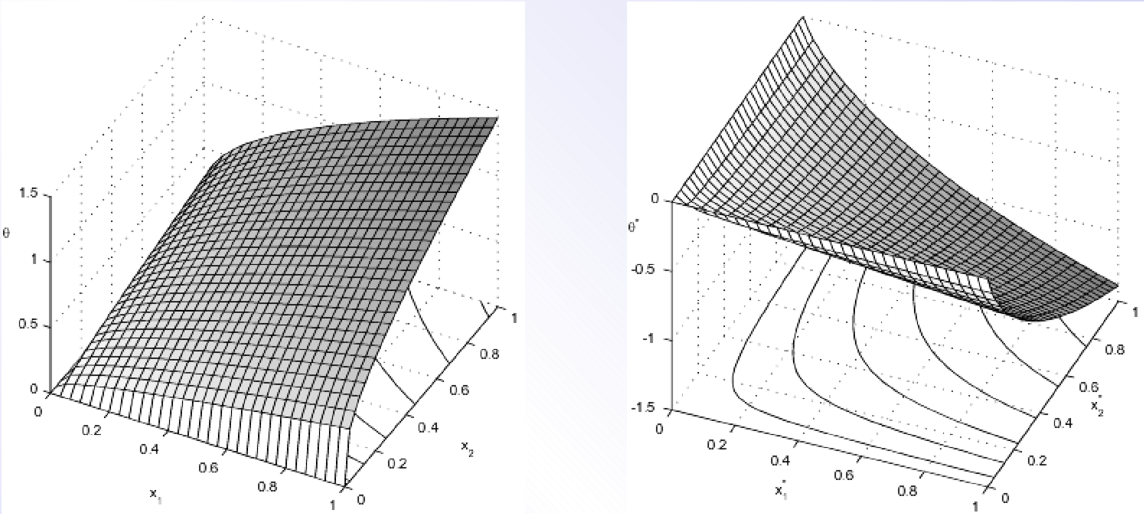

Figure 2: Boundary surfaces of the geometric cone $\mathcal{G}^2$ and its dual cone $(\mathcal{G}^2)^*$

- $\mathbb{R}^{n+1}_+ \subseteq (\mathcal{G}^n)^*$ (since $\mathcal{G}^n \subseteq \mathbb{R}^{n+1}_+$)
- The interior of $(\mathcal{G}^n)^*$ is given by

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

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

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

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

Primal

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

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

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

$\Updownarrow$ (introducing additional v variables)

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

($e \equiv$ all-one vector, $n_k = \#I_k$)
Standard conic problem:
variables $(\tilde{y}, \tilde{s})$, data$(\tilde{A}, \tilde{b}, \tilde{c})$, cone K^* with

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

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

$\Rightarrow$ we can mechanically derive the dual !

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

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

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

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

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

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

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

Weak duality

y feasible for the primal, x is feasible for the dual

$$\Rightarrow \quad b^T y \le c^T x + \sum_{k \in K} \sum_{\substack{i \in I_k \\ x_i > 0}} x_i \log \frac{x_i}{\sum_{i \in I_k} x_i} \; .$$

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

Proof *[Glineur 99]*

Weak duality theorem with conic primal-dual pair $\longrightarrow$ extend objective values to geometric primal-dual pair (easy $\longleftarrow$ convexity)

Strong duality

Primal and dual feasible solutions $\Rightarrow$ zero duality gap (but attainment not guaranteed)

Proof *[Glineur 99]*

Provide a strictly feasible dual point

$$\Leftrightarrow \quad z_k > \sum_{i \in I_k} x_i \log \frac{x_i}{\sum_{i \in I_k} x_i} \textbf{ and } x_i > 0 \; \forall i$$

But the linear constraints $Ax = b$ may force $x_i = 0$ (for some i) at 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 with a 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 $i \in \mathcal{N}$

a. restricted primal-dual conic pair

b. strictly feasible dual solution

c. zero duality gap

There remains to prove that

- ◇ Optimal objective values are equal for restricted and original dual problems (easy)

- ◇ Optimal values are equal for restricted and original primal problems (more difficult). Moreover, attainment is lost in the process.

  Difficulty: restricted posynomials have less terms than in original primal $\Rightarrow$ restricted solution may become infeasible in in original primal

  Solution: perturb the restricted primal solution

Perturbation vector given by *Goldman-Tucker theorem* applied to our detection linear program and its dual

◇ Perturbed restricted solution is asymptotically feasible for the original primal with the same objective value

◇ Another trick (mixing with a feasible solution) leads to a feasible solution with asymptotically the same objective value ($\Rightarrow$ lost attainment)

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

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

l_p-norm optimization

Primal

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

Dual (with $\frac{1}{p_i} + \frac{1}{q_i} = 1$)

$$\inf \ \psi(x, z) \ = \ c^T x + d^T z + \sum_{k=1}^{r} z_k \sum_{i \in I_k} \frac{1}{q_i} \left| \frac{x_i}{z_k} \right|^{q_i}$$

$$\text{s.t.} \ \ Ax + Bz = \eta \text{ and } z \geq 0$$

Properties *[Peterson and Ecker, 1967]*

- ◇ Convex program $\Rightarrow$ weak duality
- ◇ Generalizes linear and convex quadratic optimization
- ◇ No duality gap and primal attainment

Conic optimization approach *[Glineur 99]*

Same approach holds: corresponding cone is

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

with similar properties (closedness, interior, etc.)

Very similar dual cone

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

Same strategy

a. Weak duality is straightforward

b. Strong duality essentially follows from existence of a strictly feasible solution to the (possibly restricted) dual problem

Difference with geometric optimization

Perturbed restricted primal solution is feasible (no additional trick needed) $\Rightarrow$ primal attainment is preserved

Intermezzo:
Approximating geometric optimization with l_p-norm optimization

(chapter 8)

Interior-point methods

- ◇ Self-concordant functions

Conic optimization

- ◇ Formulation and duality
- ◇ Geometric and l_p-norm optimization
- ◇ General framework: separable optimization

Approximations

- ◇ Geometric optimization with l_p-norm optimization

Approximating geometric optimization

Principle *[Glineur 00]*

Geometric constraint is $\sum_{i \in I_k} e^{a_i^T y - c_i} \leq 1$

Relies on exponential function

Let $\alpha \in \mathbb{R}_{++}$ and define

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

We have for all $0 \leq x \leq \alpha$

$$g_\alpha(x) \leq e^{-x} < g_\alpha(x) + \alpha^{-1}$$

which implies

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

Approximated primal

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

becomes for a fixed α

$$\sup\ b^T y \quad \text{s.t.} \quad \sum_{i \in I_k} \big(g_\alpha(c_i - a_i^T y) + \alpha^{-1} \big) \leq 1 \quad (\text{GP}_\alpha)$$

$\Rightarrow$ restriction of (GP) equivalent to

$$\sup\ b^T y \quad \text{s.t.} \quad \sum_{i \in I_k} \frac{1}{\alpha} \left| c_i - \alpha - a_i^T y \right|^\alpha \leq \alpha^{\alpha-1}(1 - n_k \alpha^{-1})$$

$\Rightarrow$ a l_p-norm optimization problem !

- ◇ $\alpha \longrightarrow +\infty \Rightarrow$ approximation $g_\alpha(x) \rightarrow e^{-x}$
- ◇ Solutions of (GP_α) tend to solution of (GP) ?

Duality properties

Dual approximate problem

$$\inf \; c^T x - \alpha e_n^T x + \alpha \sum_{k \in K} (1 - n_k \alpha^{-1})^{\frac{1}{\alpha}} \left\| x_{I_k} \right\|_\beta \text{ s.t. } Ax = b$$

Fixed feasible region, when $\alpha \to +\infty$ objective tends to

$$\inf c^T x + \sum_{k \in K} \sum_{i \in I_k | x_i > 0} x_i \log \frac{x_i}{\sum_{i \in I_k} x_i} \text{ s.t. } Ax = b, \; x \geq 0$$

(hidden constraint $x \geq 0$)

$\Rightarrow$ dual geometric optimization problem

Duality results

Apply l_p-norm duality results to geometric optimization

a. Weak duality

b. Strong duality (attainment lost with the limit)

We note

a. Primal approximation:
same objective, different feasible region (restriction)

b. Dual approximation:
same feasible region, different objective

A general framework for separable convex optimization: Generalizing our conic formulations

(chapters 6–7)

Interior-point methods

◇ Self-concordant functions

Conic optimization

◇ Formulation and duality

◇ Geometric and l_p-norm optimization

◇ General framework: separable optimization

Approximations

◇ Geometric optimization with l_p-norm optimization

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

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

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

The separable cone *[Glineur 00]*

Consider a set of n scalar closed proper convex functions

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

and let

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

- ◇ $\mathcal{K}^f$ generalizes $\mathcal{L}^p$ and $\mathcal{G}_2^n$
- ◇ $\mathcal{K}^f$ is a closed convex cone
- ◇ $\mathcal{K}^f$ is solid and pointed

◇ $(x, \theta, \kappa) \in \operatorname{int} \mathcal{K}^f$ iff

$$x_i \in \operatorname{int} \operatorname{dom} f_i \text{ and } \theta \sum_{i=1}^{n} f_i(\frac{x_i}{\theta}) < \kappa$$

◇ The dual of $(\mathcal{K}^f)^*$ is defined by

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

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 optimization *[Glineur 00]*

Primal

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

Dual

$$\inf\ \psi(x,z) = c^T x + d^T z + \sum_{k \in K | z_k > 0} z_k \sum_{i \in I_k} f_i^*(-\frac{x_i}{z_k})$$

$$- \sum_{k \in K | z_k = 0} \inf_{x^*_{I_k} \in \text{dom } f_{I_k}} x_{I_k}^T x^*_{I_k}$$

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

◇ Justification for conventions when $\theta = 0$

◇ Mix different types of constraints within problems

Some other examples

◇
$$f : x \mapsto \begin{cases} -\sqrt{a^2 - x^2} & \text{if } |x| \leq a \\ +\infty & \text{if } |x| > a \end{cases}$$
$$f^* : x^* \mapsto a\sqrt{1 + x^{*2}}$$
(square roots, circles and ellipses)

◇
$$f : x \mapsto \begin{cases} -\frac{1}{p}x^p & \text{if } x \geq 0 \\ +\infty & \text{if } x < 0 \end{cases} \qquad 0 < p < 1$$
$$f^* : x^* \mapsto \begin{cases} -\frac{1}{q}(-x^*)^q & \text{if } x^* < 0 \\ +\infty & \text{if } x^* \geq 0 \end{cases} \qquad -\infty < q < 0$$
(CES functions in production and consumer theory)

◇

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

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

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

Conclusions
Summary and perspectives

Contributions

Interior-point methods

- ◇ Overview of self-concordancy theory
- ◇ Discussion over different definitions
- ◇ Optimal complexity of short-step method
- ◇ Improvement of useful Lemma

Approximations

- ◇ Approximation of geometric optimization with l_p-norm optimization

Conic optimization

◇ New convex cones to model

a. geometric optimization

b. l_p-norm optimization

◇ Simplified proofs of their duality properties

◇ New framework of separable optimization

Research directions

Interior-point methods

- Replace self-concordancy conditions by single condition involving complexity $\kappa\sqrt{\nu}$

Conic optimization

- Duality properties of separable optimization
- Self-concordant barrier for separable optimization
- Implementation of interior-point methods