Duality and algorithms in convex optimization
Part I - A survey

François Glineur

UCL/FSA/INMA & CORE
glineur@core.ucl.ac.be

HHU Düsseldorf Mathematisches Kolloquium November 26 2004

Motivation

Modelling and decision-making

Help to choose the best decision

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

Use

- Numerous applications in practice
- Resolution methods efficient in practice
- Modelling and solving large-scale problems

Introduction

Applications

- ◇ Planning, management and scheduling

  Supply chain, timetables, crew composition, etc.

- ◇ Design

  Dimensioning, structural optimization, networks

- ◇ Economics and finance

  Portfolio optimization, computation of equilibrium

- ◇ Location analysis and transport

  Facility location, circuit boards, vehicle routing

- ◇ And lots of others ...

Two facets of optimization

- Modelling

  Translate the problem into mathematical language (sometimes trickier than you might think)

$$\Updownarrow$$

Formulation of an optimization problem

$$\Updownarrow$$

- Solving

  Develop and implement algorithms that are efficient in *theory* and in *practice*

Close relationship

- Formulate models that you know how to solve
- Develop methods applicable to real-world problems

Outline of Part I

Convex optimization : models and algorithms

- *Prelude*: the case of linear optimization
- *Motivation*: convex optimization: what and why?
- *Duality*: from linear to conic optimization
- *Algorithms*: the interior-point revolution

Slides available on the web :

`http://www.core.ucl.ac.be/~glineur/`

Algorithmic complexity

Difficulty of a problem depends on the efficiency of methods that can be applied to solve it

$\Rightarrow$ what is a good algorithm ?

- $\diamond$ Solves the problem (approximately)
- $\diamond$ Until the middle of the 20$^{\text{th}}$ century: in finite time (number of elementary operations)
- $\diamond$ Now (computers): in bounded time (depending on the problem size)

  $\longrightarrow$ algorithmic complexity (worst / average case)

Crucial distinction:

polynomial $\leftrightarrow$ exponential complexity

Linear optimization

A simple problem

Consider the linear problem(with m variables y_i)

$$\max \sum_{i=1}^{m} b_i y_i \text{ such that } \sum_{i=1}^{m} a_{ij} y_i \leq c_j \ \forall 1 \leq j \leq n$$

(objective and n linear inequalities), or

$$\max b^{\mathrm{T}} y \text{ such that } A^{\mathrm{T}} y \leq c$$

(matrix notation with $b, y \in \mathbb{R}^m$, $c \in \mathbb{R}^n$ and $A \in \mathbb{R}^{m \times n}$)

All linear problems can be expressed in this format

Duality for linear optimization

$$\max b^{\mathrm{T}} y \text{ such that } A^{\mathrm{T}} y \leq c$$

The following problem, based on the same data

$$\min c^{\mathrm{T}} x \text{ such that } Ax = b \text{ and } x \geq 0$$

is closely linked: it is called the dual

- Weak duality: Inequality $b^{\mathrm{T}} y \leq c^{\mathrm{T}} x$ holds for any x, y such that $Ax = b$, $x \geq 0$ and $A^{\mathrm{T}} y \leq c$
- Strong duality: If x^* is an optimal solution for the primal, there exists an optimal solution y^* for the dual such that $c^{\mathrm{T}} x^* = b^{\mathrm{T}} y^*$

Algorithms for linear optimization

For linear optimization with continuous variables:
very efficient algorithms ($n \approx 10^7$)

- ◇ Simplex algorithm (Dantzig, 1947)
  Exponential complexity but ...
  Very efficient in practice
- ◇ Ellipsoid method (Khachiyan, 1978)
  Polynomial complexity but ...
  Poor practical performance
- ◇ Interior-point methods (Karmarkar, 1985)
  Polynomial complexity and ...
  Very efficient in practice (large-scale problems)

Nonlinear optimization

Motivation

Linear optimization does not permit satisfactory modelling of all situations → let us look again at

$$\min_{x \in \mathbb{R}^n} f(x) \text{ such that } x \in X \subseteq \mathbb{R}^n$$

where X is defined most of the time by

$$X = \{x \in \mathbb{R}^n \mid g_i(x) \leq 0 \text{ and } h_j(x) = 0 \text{ for } i \in I, j \in J\}$$

Back to complexity

Discrete sets X can make the problem difficult (with exponential complexity)
but even continuous problems can be difficult!

Consider a *simple* unconstrained minimization

$$\min f(x_1, x_2, \ldots, x_{10})$$

with smooth f (Lipschitz continuous with $L = 2$):

One can show that exists some functions where at least 10^{20} iterations (function evaluations) are needed to find a solution with accuracy 1% !

Two distinct approaches

- ◇ Tackle all problems without any efficiency guarantee
  - – Traditional **nonlinear** optimization
  - – (Meta)-Heuristic methods
- ◇ Limit the scope to some classes of problems and get in return an efficiency guarantee
  - – Linear optimization
    - ∗ very fast specialized algorithms
    - ∗ but sometimes too limited in practice
  - – **Convex** optimization

Compromise: generality $\leftrightarrow$ efficiency

Convex optimization

Introduction

$$\min f(x) \text{ such that } x \in X$$

A feasible solution x^* is a

- ◇ global minimum iff $f(x^*) \leq f(x)\ \forall x \in X$
- ◇ local minimum iff there exists an open neighborhood $V(x^*)$ such that

$$f(x^*) \leq f(x)\ \forall x \in X \cap V$$

Global minima are (much) more *difficult* to find!

Convexity definitions

- An *optimization* problem is *convex* if it deals with the minimization of a convex function on a convex set
- A set $S \subseteq \mathbb{R}^n$ is convex iff
  $\lambda x + (1-\lambda) y \in S \ \forall x, y \in S, \lambda \in [0\ 1]$
- A function $f : S \mapsto R$ is convex iff
  $f(\lambda x+(1-\lambda)y) \leq \lambda f(x)+(1-\lambda)f(y) \ \forall x, y, \lambda \in [0\ 1]$
  (this imposes that the domain S is convex)
- Equivalently, a function $f : S \subseteq \mathbb{R}^n \mapsto \mathbb{R}$ is convex iff its epigraph is convex

$$\text{epi}\, f = \{(x,t) \in \mathbb{R}^n \times \mathbb{R} \mid x \in S \text{ and } f(x) \leq t\}$$

Examples: convex sets and convex functions

- $\emptyset, \mathbb{R}^n, \mathbb{R}^n_+, \mathbb{R}^n_{++}$
- $\{x \mid \|x - a\| < r\}$ and $\{x \mid \|x - a\| \leq r\}$
- $\{x \mid b^T x < \beta\}$, $\{x \mid b^T x \leq \beta\}$ and $\{x \mid b^T x = \beta\}$
- In $\mathbb{R}$: intervals (open/closed, possibly infinite)
- $x \mapsto c$, $x \mapsto b^T y + \beta_0$, $x \mapsto \|x\|$ and $x \mapsto \|x\|^2$, $x \mapsto x^T Q x$ with $Q \in \mathbb{R}^{n \times n}$ positive semidefinite
- In the case $f : \mathbb{R} \mapsto \mathbb{R}$, we mention $x \mapsto e^x$, $x \mapsto -\log x$, $x \mapsto |x|^p$ with $p \geq 1$.
- f is concave iff $-f$ is convex (i.e. reversing inequalities in the definitions) ; there is no notion of concave set!

Fundamental properties of convex optimization

When dealing with convex optimization problems

- ◇ Every local minimum is global
- ◇ The optimal set is convex
- ◇ Special cases: linear (continuous) optimization, quadratic optimization (with positive semidefinite quadratic forms)
- ◇ Many other problems are convex (or admit equivalent convex reformulations)

Main advantages:

- ◇ efficient (polynomial) interior-point methods
- ◇ Lagrange duality $\longrightarrow$ strongly related dual problem

Conic optimization

Objective

Generalize linear optimization

$$\max b^T y \text{ such that } A^T y \leq c$$

$$\min c^T x \text{ such that } Ax = b \text{ and } x \geq 0$$

while trying to keep the nice properties

◇ duality & efficient algorithms

$\longrightarrow$ change as little as possible

Idea: generalize the inequalities $\leq$ and $\geq$

What are properties of nice inequalities ?

Generalizing $\geq$ and $\leq$

Let $K \subseteq \mathbb{R}^n$. Define

$$a \succeq_K 0 \Leftrightarrow a \in K$$

We also have

$$a \succeq_K b \Leftrightarrow a - b \succeq_K 0 \Leftrightarrow a - b \in K$$

as well as

$$a \preceq_K b \Leftrightarrow b \succeq_K a \Leftrightarrow b - a \succeq_K 0 \Leftrightarrow b - a \in K$$

Let us also impose two sensible properties

$$a \succeq_K 0 \Rightarrow \lambda a \succeq_K 0 \ \forall \lambda \geq 0 \ (K \text{ is a cone})$$

$$a \succeq_K 0 \text{ and } b \succeq_K 0 \Rightarrow a + b \succeq_K 0$$

$$(K \text{ is closed under addition})$$

Properties of admissible sets K

- K is a convex set!
- In fact, if K is a cone, we have

$$K \text{ is closed under addition } \Leftrightarrow K \text{ is convex}$$

Conic optimization

We can then generalize

$$\max b^{\mathrm{T}} y \text{ such that } A^{\mathrm{T}} y \le c$$

to

$$\max b^{\mathrm{T}} y \text{ such that } A^{\mathrm{T}} y \preceq_K c$$

$\Rightarrow$ This problem is convex
The standard linear cases corresponds to $K = \mathbb{R}^n_+$

More requirements for K

- $x \succeq 0$ and $x \preceq 0 \Rightarrow x = 0$
  which means $K \cap (-K) = \{0\}$ (the cone is pointed)
- We define the strict inequality by $a \succ 0 \Leftrightarrow a \in \operatorname{int} K$ (and $a \succ b$ iff $a - b \in \operatorname{int} K$)
  Hence we require $\operatorname{int} K \neq \emptyset$ (the cone is solid)
- Finally, we would like to be able to take limits:
  If $\{x_i\}_{i \to \infty}$ with $x_i \succeq_K 0 \; \forall i$, then $\lim_{i \to \infty} x_i = \bar{x} \Rightarrow \bar{x} \succeq_K 0$
  which is equivalent to saying that K is closed

Example: second-order (or Lorentz or ice-cream) cone

$$\mathbb{L}^n = \{(x_0, \ldots, x_n) \in \mathbb{R}^{n+1} \mid \sqrt{x_1^2 + \cdots + x_n^2} \leq x_0\}$$

Another example: semidefinite cone $K = \mathbb{S}^n_+$ (symmetric positive semidefinite matrices)

Back to conic optimization

A convex cone $K \subseteq \mathbb{R}^n$ that is solid, pointed and closed will be called a proper cone

In the following, we will always consider proper cones

We obtain

$$\max_{y \in \mathbb{R}^m} b^{\mathrm{T}} y \text{ such that } A^{\mathrm{T}} y \preceq_K c$$

or, equivalently,

$$\max_{y \in \mathbb{R}^m} b^{\mathrm{T}} y \text{ such that } c - A^{\mathrm{T}} y \in K$$

with problem data $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$ and $A \in \mathbb{R}^{m \times n}$

Combining several cones

Considering several conic constraints

$$A_1^{\mathrm{T}} y \preceq_{K_1} c_1 \text{ and } A_2^{\mathrm{T}} y \preceq_{K_2} c_2$$

which are equivalent to

$$c_1 - A_1^{\mathrm{T}} y \in K_1 \text{ and } c_2 - A_2^{\mathrm{T}} y \in K_2$$

one introduces the product cone $K = K_1 \times K_2$ to write

$$(c_1 - A_1^{\mathrm{T}} y, c_2 - A_2^{\mathrm{T}} y) \in K_1 \times K_2$$

$$\Leftrightarrow \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} - \begin{pmatrix} A_1^{\mathrm{T}} \\ A_2^{\mathrm{T}} \end{pmatrix} \in K_1 \times K_2 \Leftrightarrow \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} - \begin{pmatrix} A_1^{\mathrm{T}} \\ A_2^{\mathrm{T}} \end{pmatrix} \succeq_{K_1 \times K_2} 0$$

If K_1 and K_2 are proper, $K_1 \times K_2$ is also proper

Equivalence with convex optimization

Conic optimization is clearly a special case of convex optimization: what about the reverse statement ?

$$\min_{x \in \mathbb{R}^n} f(x) \text{ such that } x \in X \subseteq \mathbb{R}^n$$

- ◇ The objective of a convex problem can be assumed w.l.o.g. to be linear w.l.o.g.: $f(x) = c^{\mathrm{T}} x$
- ◇ The feasible region of a convex problem can be assumed w.l.o.g. to be in the conic standard format:

$$X = \{x \in K \text{ and } Ax = b\}$$

$\Rightarrow$ conic optimization equivalent to convex optimization

A linear objective ?

$$\min_{x \in \mathbb{R}^n} f(x) \text{ such that } x \in X \subseteq \mathbb{R}^n$$

$$\Updownarrow$$

$$\min_{(x,t) \in \mathbb{R}^n \times \mathbb{R}} t \text{ such that } x \in X \text{ and } (x,t) \in \operatorname{epi} f$$

$$\Updownarrow$$

$$\min_{(x,t) \in \mathbb{R}^n \times \mathbb{R}} t \text{ such that } x \in X \text{ and } f(x) \leq t$$

$\Rightarrow$ equivalent problem with linear objective

Conic constraints ?

$$K_X = \mathrm{cl}\{(x,u) \in \mathbb{R}^n \times \mathbb{R}_{++} \mid \frac{x}{u} \in X\}$$

is called the (closed) conic hull of X

We have that K_X is a closed convex cone and

$$x \in X \Leftrightarrow (x,u) \in K_X \text{ and } u = 1$$

$$\min_{x \in \mathbb{R}^n} c^{\mathrm{T}} x \text{ such that } x \in X \subseteq \mathbb{R}^n$$

$$\Updownarrow$$

$$\min_{(x,u) \in \mathbb{R}^n \times \mathbb{R}} c^{\mathrm{T}} x \text{ such that } (x,u) \succeq_{K_X} 0 \text{ and } u = 1$$

$\Rightarrow$ equivalent problem with a conic constraint

Duality properties

Since we generalized

$$\max b^T y \text{ such that } A^T y \leq c$$

to

$$\max b^T y \text{ such that } A^T y \preceq_K c$$

it is tempting to generalize

$$\min c^T x \text{ such that } Ax = b \text{ and } x \geq 0$$

to

$$\min c^T x \text{ such that } Ax = b \text{ and } x \succeq_K 0$$

But this is not the right primal-dual pair !

The dual cone

$$K^* = \{z \in \mathbb{R}^n \text{ such that } x^T z \geq 0 \ \forall x \in K\}$$

- For any $x \in K$ and $z \in K^*$, we have $z^T x \geq 0$
- K^* is a convex cone, called the dual cone of K
- K^* is always closed, and if K is closed, $(K^*)^* = K$
- K is pointed (resp. solid) $\Rightarrow$ K^* is solid (resp. pointed)
- Cartesian products: $(K_1 \times K_2)^* = K_1^* \times K_2^*$
- $(\mathbb{R}^n_+)^* = \mathbb{R}^n_+$, $(\mathbb{L}^n)^* = \mathbb{L}^n$, $(\mathbb{S}^n_+)^* = \mathbb{S}^n_+$:
  these cones are self-dual
- But there exists (many) cones that are not self-dual

Primal-dual pair

We can write the primal conic problem

$$\min c^{\mathrm{T}} x \text{ such that } Ax = b \text{ and } x \succeq_K 0$$

and the dual conic problem

$$\max b^{\mathrm{T}} y \text{ such that } A^{\mathrm{T}} y \preceq_{K^*} c$$

(for historical reasons, the min problem is called the primal ; anyway $(K^*)^* = K^*$ holds)

- Very symmetrical formulation
- Computing the dual essentially amounts to finding K^*
- All nonlinearities are confined to the cones K and K^*

Duality properties

- Weak duality: any feasible solution for the primal (resp. dual) provides an upper (resp. lower) bound for the dual (resp. primal)

  (immediate consequence of the dualizing procedure)

- Inequality $b^T y \leq c^T x$ holds for any x, y such that $Ax = b$, $x \succeq_K 0$ and $A^T y \preceq_{K^*} c$ (corollary)

- If the primal (resp. dual) is unbounded, the dual (resp. primal) must be infeasible

  (but the converse is not true!)

Completely similar to the situation for linear optimization

Duality properties (continued)

What about strong duality ?

If y^* is an optimal solution for the dual, does there exist an optimal solution x^* for the primal such that $c^T x^* = b^T y^*$ (in other words: $p^* = d^*$) ?

Consider $K = \mathbb{L}^2$ with

$$A = \begin{pmatrix} -1 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix},\ b = \begin{pmatrix} 0 & -1 \end{pmatrix}^T \text{ and } c = \begin{pmatrix} 0 & 0 & 0 \end{pmatrix}^T$$

We can easily check that

- ◇ the primal is infeasible
- ◇ the dual is bounded and solvable

$\Rightarrow$ strong duality does not hold for conic optimization ...

Other troublesome situations

Let $\lambda \in \mathbb{R}_+$: consider

$$\min \lambda x_3 - 2x_4 \text{ s.t. } \begin{pmatrix} x_1 & x_4 & x_5 \\ x_4 & x_2 & x_6 \\ x_5 & x_6 & x_3 \end{pmatrix} \succeq_{\mathbb{S}^3_+} 0, \ \begin{pmatrix} x_3 + x_4 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

In this case, $p^* = \lambda$ but $d^* = 2$: duality gap!

$$\min x_1 \text{ such that } x_3 = 1 \text{ and } \begin{pmatrix} x_1 & x_3 \\ x_3 & x_2 \end{pmatrix} \succeq_{\mathbb{S}^2_+} 0$$

In this case, $p^* = 0$ but the problem is unsolvable!

In all cases, one can identify the cause for our troubles: the affine subspace defined by the linear constraints is tangent to the cone (it does not intersect its interior)

Rescuing strong duality

A feasible solution to a conic (primal or dual) problem is strictly feasible iff it belongs to the interior of the cone In other words, we must have $Ax = b$ and $x \succ_K 0$ for the primal and/or $A^T y \prec_{K^*} c$ for the dual

Strong duality: If the dual problem admits a strictly feasible solution, we have either

- an unbounded dual, in which case $d^* = +\infty = p^*$ and the primal is infeasible
- a bounded dual, in which case the primal is solvable with $p^* = d^*$ (hence there exists at least one feasible primal solution x^* such that $c^T x^* = p^* = d^*$)

Strong duality (continued)

- If the primal problem admits a strictly feasible solution, we have either
  - an unbounded primal, in which case $p^* = -\infty = d^*$ and the dual is infeasible
  - a bounded primal, in which case the dual is solvable with $d^* = p^*$ (hence there exists at least one feasible dual solution y^* such that $b^T y^* = d^* = p^*$)
- The first case is a mere consequence of weak duality
- Finally, when both problems admit a strictly feasible solution, both problems are solvable and we have

$$c^T x^* = p^* = d^* = b^T y^*$$

Interior-point methods

Back to convex optimization

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

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

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. $(f(x) = c^{\mathrm{T}} x)$

Principle

Approximate a constrained problem by

a *family* of unconstrained problems

Use a barrier function F to replace the inclusion $x \in C$

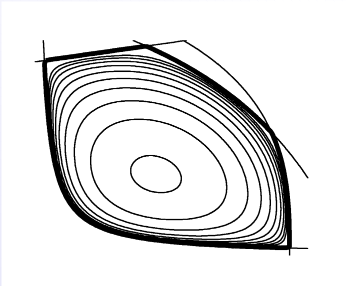
- F is smooth
- F is strictly convex on $\operatorname{int} C$
- $F(x) \to +\infty$ when $x \to \partial C$

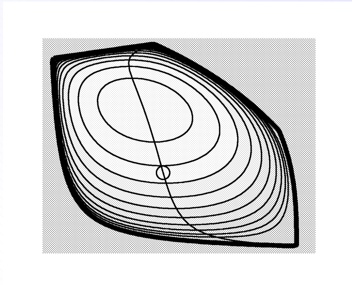
$$\to \quad C = \operatorname{cl} \operatorname{dom} F = \operatorname{cl} \{x \in \mathbb{R}^n \mid F(x) < +\infty\}$$

Central path

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

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





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

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 convex constraints $g_i(x) \leq 0$, one can introduce the logarithmic barrier function

$$F(x) = -\sum_{i=1}^{n} \log(-g_i(x))$$

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 & Nemirovski, 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

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

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

A (simple?) example

For linear optimization, $C = \mathbb{R}^n_+$: take $F(x) = -\sum_{i=1}^n \log x_i$

When $n = 1$, we can choose $(\kappa, \nu) = (1, 1)$

- $\diamond$ $\nabla F(x) = -\frac{1}{x}$ and $\nabla F(x)^{\mathrm{T}} h = -\frac{h}{x}$
- $\diamond$ $\nabla^2 F(x) = \frac{1}{x^2}$ and $\nabla^2 F(x)[h, h] = \frac{h^2}{x^2}$
- $\diamond$ $\nabla^3 F(x) = -2\frac{1}{x^3}$ and $\nabla^3 F(x)[h, h, h] = -2\frac{h^3}{x^3}$

When $n > 1$, we have

- $\diamond$ $\nabla F(x) = (-x_i^{-1})$ and $\nabla F(x)^{\mathrm{T}} h = -\sum h_i x_i^{-1}$
- $\diamond$ $\nabla^2 F(x) = \mathrm{diag}(x_i^{-2})$ and $\nabla^2 F(x)[h, h] = \sum h_i^2 x_i^{-2}$
- $\diamond$ $\nabla^3 F(x) = \mathrm{diag}_3(-2x_i^{-3})$, $\nabla^3 F(x)[h, h, h] = -2 \sum h_i^3 x_i^{-3}$

and one can show that $(\kappa, \nu) = (1, n)$ is valid

Barrier calculus

Two elementary results:

- Scaling:

  F is a (κ, ν)-s.-c. barrier for $\mathcal{C} \subseteq \mathbb{R}^n$ and $\lambda \in \mathbb{R}_{++}$

  $\Rightarrow (\lambda F)$ is a $(\frac{\kappa}{\sqrt{\lambda}}, \lambda\nu)$-s.-c. barrier for $\mathcal{C}$

- Sum:

  F is a (κ_1, ν_1)-s.-c. barrier for $\mathcal{C}_1 \subseteq \mathbb{R}^n$

  G is a (κ_2, ν_2)-s.-c. barrier for $\mathcal{C}_2 \subseteq \mathbb{R}^n$

  $\Rightarrow (F + G)$ is a $(\max\{\kappa_1, \kappa_2\}, \nu_1 + \nu_2)$-s.-c. barrier for the set $\mathcal{C}_1 \cap \mathcal{C}_2$ (if nonempty)

Complexity result

Summary

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

Short-step method: follow the central path

- $\diamond$ Measure distance to the central path with $\delta(x, \mu)$
- $\diamond$ Choose a starting iterate with a small $\delta(x_0, \mu_0) < \tau$
- $\diamond$ While accuracy is not attained
  a. Decrease μ geometrically (δ increases above τ)
  b. Take a Newton step to minimize barrier (δ decreases below τ)

Geometric interpretation

Two self-concordancy conditions: each has its role

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

Summarized complexity result

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

iterations lead a solution with ϵ accuracy on the objective

Complexity result

- Let F be a (κ, ν)-self-concordant barrier for C and let $x_0 \in \operatorname{int} C$ be a starting point,

  a short-step interior-point algorithm can solve problem (P) up to ϵ accuracy within

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

  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 $\mathbb{R}^n$

- Complexity invariant w.r.t. to scaling of F

- Universal bound on complexity parameter: $\kappa\sqrt{\nu} \geq 1$

Corollary

Assume F, ∇F and $\nabla^2 F$ are polynomially computable
$\Rightarrow$ problem (P) can be solved in polynomial time

Existence

There exists a universal SC barrier with parameters

$$\kappa = 1 \text{ and } \nu = \mathcal{O}(n)$$

(but not necessarily efficiently computable)

Examples

- ◇ linear optimization: $(\kappa, \nu) = (1, n) \Rightarrow \mathcal{O}\left(\sqrt{n} \log \frac{1}{\varepsilon}\right)$
- ◇ entropy optimization: $\kappa = 1$ and $\nu = 2n \Rightarrow \mathcal{O}\left(\sqrt{n} \log \frac{1}{\varepsilon}\right)$
  ($\inf c^{\mathrm{T}} x + \sum_i x_i \log x_i$ such that $Ax = b$ and $x \geq 0$)

References for Part I

Convex optimization

- ◇ Convex Analysis, ROCKAFELLAR,
  Princeton University Press, 1980
- ◇ Convex optimization, BOYD and VANDENBERGHE,
  Cambridge University Press, 2004 (on the web)

Convex modelling

- ◇ Lectures on Modern Convex Optimization, Analysis, Algorithms, and Engineering Applications, BEN-TAL and NEMIROVSKI,
  MPS/SIAM Series on Optimization, 2001

Interior-point methods (linear)

- Primal-Dual Interior-Point Methods, Wright
  SIAM, 1997
- Theory and Algorithms for Linear Optimization, Roos, Terlaky, Vial, John Wiley & Sons, 1997

Interior-point methods (convex)

- Interior-point polynomial algorithms in convex programming, Nesterov & Nemirovski, SIAM, 1994
- A Mathematical View of Interior-Point Methods in Convex Optimization, Renegar,
  MPS/SIAM Series on Optimization, 2001

Duality and algorithms in convex optimization

Part II

The case of second-order cone programming

with a single cone

Outline of Part II

Introduction

- ◇ Reminder: Convex, conic and second-order cone optimization
- ◇ Two easy subproblems

Second-order cone feasibility problem

- ◇ Main ideas: homogenization and minimum-norm solution
- ◇ Our algorithm: a three-case discussion

Second-order cone optimization problem

- ◇ Using the feasibility problem as a subproblem
- ◇ What about the dual problem?

Concluding remarks

- ◇ Summary, complexity and generalizations

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

- $\diamond$ Local optima $\Rightarrow$ global, form a convex optimal set
- $\diamond$ Lagrange duality $\Rightarrow$ related (asymmetric) dual problem
- $\diamond$ Efficient *interior-point* methods (*self-concordant barriers*)

Conic optimization

Let $\mathcal{C} \subseteq \mathbb{R}^n$ a *solid, pointed, closed convex cone* :

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

Primal-dual pair

Dual cone is also a solid pointed closed convex cone

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

$\Rightarrow$ pair of primal-dual problems

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

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

Several cones: $x^1 \in \mathcal{C}^1, \ldots, x^r \in \mathcal{C}^r \Leftrightarrow (x^1, \ldots, x^r) \in \mathcal{C}^1 \times \cdots \times \mathcal{C}^r$

Advantages over classical formulation

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

Examples

$\mathcal{C} = \mathbb{R}^n_+ = \mathcal{C}^* \Rightarrow$ linear optimization

$\mathcal{C} = \mathbb{S}^n_+ = \mathcal{C}^* \Rightarrow$ semidefinite optimization

Both cones are *self-dual*.

A *single* $\mathbb{R}^n_+/\mathbb{S}^n_+$ cone can be considered w.l.o.g.

Second-order cone optimization

$$\mathbb{L}^n = \{(x_0, x_1, \ldots, x_n) \in \mathbb{R}_+ \times \mathbb{R}^n \mid \|(x_1, \ldots, x_n)\| \leq x_0\} \subset \mathbb{R}^{n+1}$$

Second-order or Lorentz cone $\mathbb{L}^n$ is self-dual

But a set of several constraints $x^i \in \mathbb{L}^{n_i}, i = 1, \ldots, r$ **cannot** be concatenated into a single second-order cone constraint

Goal of this talk: Study the problem with a single second-order cone

Single-constraint second-order cone problem

$$\inf_{x_0 \in \mathbb{R},\ x \in \mathbb{R}^n} c_0 x_0 + c^{\mathrm{T}} x \quad \text{s.t.} \quad a_0 x_0 + A x = b \text{ and } (x_0, x) \in \mathbb{L}^n$$

with $c_0 \in \mathbb{R}$, $c \in \mathbb{R}^n$; $a_0 \in \mathbb{R}^m$, $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$

Known results

It is *well-known* that this problem can be solved *analytically* *[Alizadeh and Goldfarb 2003]*: solve primal-dual optimality conditions

$$\begin{array}{rcl} a_0 x_0 + A_0 x = b & \text{and} & (x_0, x) \in \mathbb{L}^n \\ a_0^{\mathrm{T}} y + z_0 = c_0,\ A^{\mathrm{T}} y + z = c & \text{and} & (z_0, z) \in \mathbb{L}^n \\ (x_0, x) \circ (z_0, z) = 0 & & (\Leftrightarrow\ c_0 x_0 + c^{\mathrm{T}} x = b^{\mathrm{T}} y) \end{array}$$

where $(x_0, x) \circ (z_0, z) = (x_0 y_0 + x^{\mathrm{T}} y, x_0 y + y_0 x)$

But only works when *strong duality* holds !

Analytical solution of optimality conditions

Define $\bar{A} = (a_0 \; A)$, $\bar{c} = (c_0, c)$, $Q = P_{\text{null}\,\bar{A}} = I - \bar{A}^\dagger \bar{A}$.

Assuming $x_0 > 0$ and $z_0 > 0$, one obtains after some linear algebra

$$
\begin{aligned}
\gamma &= 1 - 2a_0^{\mathrm{T}}(\bar{A}\bar{A}^{\mathrm{T}})^{-1}a_0 \\
\alpha &= \frac{z_0}{x_0} = \sqrt{\frac{-\gamma\bar{c}^{\mathrm{T}}Q\bar{c} + 2(e^{\mathrm{T}}Q\bar{c})^2}{\gamma b^{\mathrm{T}}(\bar{A}\bar{A}^{\mathrm{T}})^{-1}b + 2(a_0^{\mathrm{T}}(\bar{A}\bar{A}^{\mathrm{T}})^{-1}b)^2}} \\
\delta &= \frac{\bar{c}^{\mathrm{T}}Q\bar{c} + \alpha^2 b^{\mathrm{T}}(\bar{A}\bar{A}^{\mathrm{T}})^{-1}b}{e^{\mathrm{T}}Q\bar{c} + \alpha a_0^{\mathrm{T}}(\bar{A}\bar{A}^{\mathrm{T}})^{-1}b} \\
y &= (\bar{A}\bar{A}^{\mathrm{T}})^{-1}(\bar{A}\bar{c} + \alpha b - \delta a_0) \\
(z_0, z) &= \bar{c} - \bar{A}^{\mathrm{T}}y = Q\bar{c} - \bar{A}^\dagger(\alpha b - \delta a_0) \\
(x_0, x) &= (z_0/\alpha, -z/\alpha)
\end{aligned}
$$

Technical derivation but *why* is this possible? Special cases?

Classification of conic convex problems

Feasibility status of conic problem: define $\mathcal{L} = \{x \in \mathbb{R}^n \mid Ax = b\}$

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

◇ Feasible iff $\mathcal{L} \cap \mathcal{C} \neq \emptyset$. Moreover, in this case,

- *Strictly* feasible (s.f.) if $\mathcal{L} \cap \operatorname{int} \mathcal{C} \neq \emptyset$
- *Weakly* feasible (w.f.) otherwise

◇ Infeasible iff $\mathcal{L} \cap \mathcal{C} = \emptyset$. Moreover, in this case,

- *Strictly* infeasible (s.i.) if $\operatorname{dist}(\mathcal{L}, \mathcal{C}) > 0$
- *Weakly* infeasible (w.i.) otherwise, i.e. $\operatorname{dist}(\mathcal{L}, \mathcal{C}) = 0$
  $\Leftrightarrow \exists \{x^k\} \mid x^k \in \mathcal{L}$ and $\lim_{k \to \infty} \operatorname{dist}(x^k, \mathcal{C}) = 0$
  (which implies then that $\{x^k\}$ is unbounded)

All these cases already arise when $\mathcal{C} = \mathbb{L}^2$!

Two easy subproblems

Smallest-norm point on a line

Let $\alpha, \beta \in \mathbb{R}^n$ and $\beta \neq 0$

$$\min_{t \in \mathbb{R}} \phi(t) = \|\alpha - t\beta\|$$

$\phi(t)^2 = t^2 \|\beta\|^2 - 2t\alpha^{\mathrm{T}}\beta + \|\alpha\|^2$ is easily minimized $\rightarrow t_* = \frac{\alpha^{\mathrm{T}}\beta}{\|\beta\|^2}$

Minimum distance is $\phi(t_*)^2 = \|\alpha\|^2 - t_*^2 \|\beta\|^2 = \|\alpha\|^2 - \frac{(\alpha^{\mathrm{T}}\beta)^2}{\|\beta\|^2}$

Intersecting a ball with an affine subspace

Decide the feasibility of the following convex problem

$$\text{Find } x \in \mathbb{R}^n \quad \text{s.t.} \quad Ax = b \text{ and } \|x\| \leq 1$$

with $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$.

Assume that A has full row rank ($\Rightarrow AA^{\mathrm{T}} \succ 0$ and $Ax = b$ is feasible)

Idea: compute minimum-norm solution $\hat{x}$ of $Ax = b$

$$\hat{x} = A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}b$$

($A^{\dagger} = A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}$ is the Moore-Penrose generalized inverse of A)

Discuss according to $\|\hat{x}\|^2 = \left\|A^{\dagger}b\right\|^2 = b^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}b$

- $\|\hat{x}\|^2 > 1 \Rightarrow$ problem is strictly infeasible
- $\|\hat{x}\|^2 = 1 \Rightarrow$ problem is weakly feasible, $\hat{x}$ is unique solution
- $\|\hat{x}\|^2 < 1 \Rightarrow$ problem is strictly feasible, $\hat{x}$ is among the solutions

(obviously problem cannot be weakly infeasible)

Forming $AA^{\mathrm{T}} \in \mathbb{R}^{m \times m} \rightarrow \mathcal{O}\left(m^2 n\right)$ operations
Factorizing $AA^{\mathrm{T}} = LL^{\mathrm{T}}$, $L \in \mathbb{R}^{m \times m}$ triangular $\rightarrow \mathcal{O}\left(m^3\right)$ operations
$\Rightarrow$ Computing $x^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}y$ for any $x, y \in \mathbb{R}^m$ is $\mathcal{O}\left(m^2 n\right)$ operations.

Second-order cone feasibility problem

Our problem

Determine *feasibility* status of the following problem

$$\text{Find } (x_0, x) \in \mathbb{R} \times \mathbb{R}^n \quad \text{s.t.} \quad a_0 x_0 + Ax = b \text{ and } (x_0, x) \in \mathbb{L}^n$$

with $a_0 \in \mathbb{R}^m$, $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. Assume w.l.o.g. that $a_0 x_0 + Ax = b$ is feasible and $(a_0 \ A)$ full row rank $(a_0 a_0^{\mathrm{T}} + AA^{\mathrm{T}} \succ 0)$

Main idea

Use the fact that $\mathbb{L}^n$ and $\mathbb{B}^n$ are strongly related:

$$(x_0, x) \in \mathbb{L}^n \quad \Leftrightarrow \quad (x/x_0) \in \mathbb{B}^n \text{ and } x_0 > 0 \quad \text{or} \quad (x_0, x) = (0, 0)$$

$\rightarrow$ use *homogenization* to get rid of x_0 variable

Our procedure: outline

Let $t > 0$ be a homogenizing variable. Problem

$$\text{Find } (x_0, x) \in \mathbb{R} \times \mathbb{R}^n \quad \text{s.t.} \quad a_0 x_0 + Ax = b \text{ and } (x_0, x) \in \mathbb{L}^n$$

becomes *equivalent* to the problem of finding

$$(x_0, x, t) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}_{++} \quad \text{s.t.} \quad a_0 x_0 + Ax = b\, t \text{ and } (x_0, x) \in \mathbb{L}^n$$

Each point (x_0, x) becomes a ray (tx_0, tx, t), $t > 0$, *except* for $(0, 0)$

Problem is completely *homogeneous* $\rightarrow$ arbitrarily fix $x_0 = 1$

$\Rightarrow$ constraint $(x_0, x) \in \mathbb{L}^n$ becomes equivalent to $x \in \mathbb{B}^n$ $\quad \rightarrow$

$$\text{Find } (x, t) \in \mathbb{R}^n \times \mathbb{R}_{++} \quad \text{s.t.} \quad Ax = b\, t - a_0 \text{ and } x \in \mathbb{B}^n$$

which is the second easy problem with a parameter $t \in \mathbb{R}_{++}$

Soln (x, t) to this problem $\rightarrow$ soln $(1/t, x/t) \in \mathbb{L}^n$ to original problem

Special case

Do the linear constraints $a_0 x_0 + Ax = b$ imply that x_0 is constant?

$$\Leftrightarrow \quad \exists y \mid A^{\mathrm{T}} y = 0 \quad (A \text{ is rank deficient})$$

$\Rightarrow x_0 = y^{\mathrm{T}} b / a_0 = C$ for all feasible solutions

Not a *true* second-order cone problem

- $C < 0 \rightarrow$ problem strictly infeasible
- $C = 0 \rightarrow (0, 0)$ unique potential solution
  - $b = 0 \rightarrow$ problem weakly feasible
  - $b \neq 0 \rightarrow$ problem strictly infeasible
- $C > 0$ problem becomes $A(x/x_0) = b/x_0 - a_0$ with $(x/x_0) \in \mathbb{B}^n$ which is *easy* (look at $\left\| A^{\dagger}(b/x_0 - a_0) \right\| \;\rightarrow\;$ s.i., w.f. or s.f.)

Main case

Assume A has full row rank. Problem is equivalent to

$$\text{Find } (x,t) \in \mathbb{R}^n \times \mathbb{R}_{++} \quad \text{s.t.} \quad Ax = b\,t - a_0 \text{ and } x \in \mathbb{B}^n$$

This suggests to look at $\left\|A^\dagger(b\,t - a_0)\right\|^2 = \|\alpha - t\beta\|^2$ with

$$\alpha = A^\dagger a_0 = A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}a_0 \text{ and } \beta = A^\dagger b = A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}b$$

$(\|\alpha\|^2 = a_0^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}a_0,\ \|\beta\|^2 = b^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}b,\ \alpha^{\mathrm{T}}\beta = a_0^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}b)$

However the possibility $x_0 = 0$ was left out!

In this case, we have $(x_0, x) = (0, 0)$ which implies $b = 0$ and $\beta = 0$

We therefore have to distinguish two cases: $b = 0$ and $b \neq 0$

and add soln $(0, 0)$ to homogenized solutions $(1/t, x/t)$ when $b = 0$

$$\text{Find } (x,t) \in \mathbb{R}^n \times \mathbb{R}_{++} \quad \text{s.t.} \quad Ax = b\,t - a_0 \text{ and } x \in \mathbb{B}^n$$

Case A: $b = 0$

In this case, $\|\alpha - t\beta\|^2 = \|\alpha\|$ does not depend from t

- $\|\alpha\| > 1 \rightarrow$ no solution except $(0,0) \rightarrow$ problem w.f.
- $\|\alpha\| = 1 \rightarrow$ ray $(t, t\alpha), t \in \mathbb{R}_+$ is solution $\rightarrow$ problem w.f.
- $\|\alpha\| < 1 \rightarrow$ interior solutions $\rightarrow$ problem s.f.

Case B: $b \neq 0$

We can here safely ignore solutions with $x_0 = 0$

In theory, minimum value of $\|\alpha - t\beta\|^2$ is attained for $t_* = \frac{\alpha^T\beta}{\|\beta\|^2}$

But t is required to be positive

$\rightarrow$ distinguish whether t_* is positive or not $\Rightarrow$ discuss the *sign* of $\alpha^T\beta$

Geometrically: study intersection of open half-line with ball

$$\{\alpha - \mathbb{R}_{++}\beta\} \cap \mathbb{B}^n$$

CASE B_1: $\alpha^{\mathrm{T}}\beta > 0$

Minimum t_* is achieved. The smallest-norm solution is then

$$t_*\beta - \alpha \text{ with } \|t_*\beta - \alpha\|^2 = \|\alpha\|^2 - t_*^2 \|\beta\|^2 = \|\alpha\|^2 - \frac{(\alpha^{\mathrm{T}}\beta)^2}{\|\beta\|^2} = \delta$$

- $\delta > 1 \longrightarrow$ no solution $\longrightarrow$ problem s.i.
- $\delta < 1 \longrightarrow$ there are interior solutions $\longrightarrow$ problem s.f.
- $\delta = 1 \longrightarrow$ only one solution $(1/t_*, \alpha/t_* - \beta) \longrightarrow$ problem w.f.

  Note this solution is easy to compute (in $\mathcal{O}\left(m^2 n\right)$ operations)

$$(x_0, x) = \left(1/t_*, A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}(b - a_0/t_*)\right) \text{ with } t_* = \frac{a_0^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}b}{b^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}b}$$

Case B_2: $\alpha^T\beta \le 0$

Minimum t_* cannot be reached

$\rightarrow$ actual minimum attained for $t \rightarrow 0^+$

This means we have to look at $\|\alpha\|^2$

- $\|\alpha\| > 1 \rightarrow$ no solution $\rightarrow$ problem s.i.
- $\|\alpha\| < 1 \rightarrow \exists t > 0$ such that $\|\alpha - t\beta\|^2 < 1 \rightarrow$ problem s.f.
- $\|\alpha\| = 1 \rightarrow$ no solution since $t = 0$ is forbidden

  But $\operatorname{dist}(\beta t - \alpha, \mathbb{B}^n) \rightarrow 0$. Let $x^t = (1/t, \beta - \alpha/t)$.

  What about $\operatorname{dist}(x^t, \mathbb{L}^n)$ as $t \rightarrow 0$?

  One has $\operatorname{dist}(x^t, \mathbb{L}^n) \approx (1/t) \operatorname{dist}(\beta t - \alpha, \mathbb{B}^n)$ and thus

  - $\alpha^T\beta = 0 \rightarrow \operatorname{dist}(x^t, \mathbb{L}^n) \rightarrow 0 \rightarrow$ w.i.
  - $\alpha^T\beta < 0 \rightarrow \operatorname{dist}(x^t, \mathbb{L}^n) > \epsilon \rightarrow$ s.i.

Summarizing table

◇ A not full row rank $\rightarrow x_0 = C$

- $C < 0 \rightarrow$ s.i.
- $C = 0$ and $b = 0 \rightarrow$ w.f. ; $C = 0$ and $b \neq 0 \rightarrow$ s.i.
- $C > 0 \rightarrow$ s.f., w.f. or s.i. when $\left\| A^\dagger (b/x_0 - a_0) \right\| <, =, > 1$

◇ A has full row rank, define α, β and $\delta = \|\alpha\|^2 - \frac{(\alpha^T \beta)^2}{\|\beta\|^2}$

	$\|\alpha\| < 1$	$\|\alpha\| = 1$	$\|\alpha\| > 1$
$\alpha^T \beta < 0$	s.f.	s.i.	s.i.
$b \neq 0,\ \alpha^T \beta = 0$	s.f.	w.i.	s.i.
$b = 0,\ \alpha^T \beta = 0$	s.f.	w.f.	w.f.
$\alpha^T \beta > 0$	s.f.	s.f.	*s.f., w.f., s.i.* $(\delta <, =, > 1)$

Notes: "Loose" dependence from a_0 and b ; $\|\beta\|$ in single case

Simplifying notations

$$\text{Find } (x_0, x) \in \mathbb{R} \times \mathbb{R}^n \quad \text{s.t.} \quad a_0 x_0 + Ax = b \text{ and } (x_0, x) \in \mathbb{L}^n$$

$$\rightarrow \quad \text{Find } \bar{x} \in \mathbb{R} \times \mathbb{R}^n \quad \text{s.t.} \quad \bar{A}\bar{x} = b \text{ and } \bar{x} \in \mathbb{L}^n$$

Let

$$P = \bar{A}^{\mathrm{T}} (\bar{A}\bar{A}^{\mathrm{T}})^{-1} \bar{A}$$

(orthogonal projection on range $\bar{A}^{\mathrm{T}}$)

$$(d_0, d) = \bar{d} = \bar{A}^{\mathrm{T}} (\bar{A}\bar{A}^{\mathrm{T}})^{-1} b \in \operatorname{range} \bar{A}^{\mathrm{T}}$$

(minimum-norm solution of $\bar{A}\bar{x} = b$)

$$\rightarrow \quad \text{Find } \bar{x} \in \mathbb{R} \times \mathbb{R}^n \quad \text{s.t.} \quad P\bar{x} = \bar{d} \text{ and } \bar{x} \in \mathbb{L}^n$$

Summarizing table

Let $w = (1\ 0 \cdots 0) \in \mathbb{R}^{n+1}$ and $\lambda = \|Pw\|^2$

We have

$$\|\alpha\|^2 = \frac{\lambda}{1-\lambda},\ \mathrm{sign}(\alpha^{\mathrm{T}}\beta) = \mathrm{sign}\, d_0 \text{ and } \mathrm{sign}(\delta - 1) = \mathrm{sign}(\lambda \left\|\bar{d}\right\|^2 - d_0^2)$$

Main case (not considering rank-deficient A nor $b = 0$)

	$\lambda < 0$	$\lambda = 0$	$\lambda > 0$	
$d_0 < 0$	s.f.	s.i.	s.i.	
$d_0 = 0$	s.f.	w.i.	s.i.	
$d_0 > 0$	s.f.	s.f.	*s.f., w.f., s.i.*	$(\lambda \left\|\bar{d}\right\|^2 <,=,> d_0^2)$

Second-order cone optimization problem

Problem definition

$$\inf_{x_0 \in \mathbb{R},\ x \in \mathbb{R}^n} c_0 x_0 + c^T x \quad \text{s.t.} \quad a_0 x_0 + Ax = b \text{ and } (x_0, x) \in \mathbb{L}^n$$

with $c_0 \in \mathbb{R}$, $c \in \mathbb{R}^n$; $a_0 \in \mathbb{R}^m$, $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. Assume w.l.o.g. that $a_0 x_0 + Ax = b$ is feasible and $(a_0\ A)$ full row rank

Using feasibility problem as subproblem

Idea: add $c_0 x_0 + c^T x = \gamma$ as a constraint with $\gamma \in \mathbb{R}$ as a parameter

$\rightarrow$ test whether γ is a *feasible* objective value

Find $(x_0, x) \in \mathbb{R} \times \mathbb{R}^n$ s.t. $\begin{pmatrix} c_0 \\ a_0 \end{pmatrix} x_0 + \begin{pmatrix} c^T \\ A \end{pmatrix} x = \begin{pmatrix} \gamma \\ b \end{pmatrix}$ and $(x_0, x) \in \mathbb{L}^n$

which is a feasibility second-order cone problem with new data

Data of the subproblem

We have

$$\tilde{a}_0 = \begin{pmatrix} c_0 \\ a_0 \end{pmatrix}, \ \tilde{A} = \begin{pmatrix} c^{\mathrm{T}} \\ A \end{pmatrix}, \tilde{b} = \begin{pmatrix} \gamma \\ b \end{pmatrix}$$

What are the new quantities α and β?

We need to evaluate $(\tilde{A}\tilde{A}^{\mathrm{T}})^{-1} = \begin{pmatrix} c^{\mathrm{T}}c & c^{\mathrm{T}}A^{\mathrm{T}} \\ Ac & AA^{\mathrm{T}} \end{pmatrix}^{-1}$

Possible to compute this as a function of $(AA^{\mathrm{T}})^{-1}$ and c but tedious

Better approach

We can actually suppose *without loss of generality* that $\bar{c} \in \operatorname{null} \bar{A}$

New problem definition

$$\inf_{x_0 \in \mathbb{R},\; x \in \mathbb{R}^n} c_0 x_0 + c^{\mathrm{T}} x \quad \text{s.t.} \quad a_0 x_0 + Ax = b \text{ and } (x_0, x) \in \mathbb{L}^n$$

$$\rightarrow \qquad \inf_{\bar{x} \in \mathbb{R} \times \mathbb{R}^n} \bar{c}^{\mathrm{T}} \bar{x} \quad \text{s.t.} \quad P\bar{x} = \bar{d} \text{ and } \bar{x} \in \mathbb{L}^n$$

with $\bar{d} \in \operatorname{range} \bar{A}^{\mathrm{T}}$ and $\bar{c} \in \operatorname{null} \bar{A}$

Testing whether γ is a feasible objective value:

$$\text{Find } \bar{x} \in \mathbb{R} \times \mathbb{R}^n \quad \text{s.t.} \quad P\bar{x} = \bar{d},\; \bar{c}^{\mathrm{T}}\bar{x} = \gamma \text{ and } \bar{x} \in \mathbb{L}^n$$

Data of the original feasibility problem *becomes*

$$P \rightarrow P + \frac{\bar{c}\bar{c}^{\mathrm{T}}}{\|\bar{c}\|^2}, \quad \lambda \rightarrow \lambda + \frac{c_0^2}{\|\bar{c}\|^2}$$

$$d_0 \rightarrow d_0 + \gamma \frac{c_0}{\|\bar{c}\|^2}, \quad \|\bar{d}\|^2 \rightarrow \|\bar{d}\|^2 + \frac{\gamma^2}{\|\bar{c}\|^2}$$

Special case: $c_0 = 0$

In this case, one can simply read the values from the feasibility table

	$\lambda < 0$	$\lambda = 0$	$\lambda > 0$
$d_0 < 0$	s.f.	s.i.	s.i.
$d_0 = 0$	s.f.	w.i.	s.i.
$d_0 > 0$	s.f.	s.f.	*s.f., w.f., s.i.*

$(\lambda(\|\bar{d}\|^2 + \frac{\gamma^2}{\|\bar{c}\|^2}) <, =, > d_0^2)$

Observations

s.i. $\rightarrow$ s.i, s.f. and w.f. $\rightarrow$ problem unbounded from above and below

w.i. $\rightarrow$ problem asymptotically unbounded from above and below

Exception: Case $(d_0 > 0, \lambda > 0)$: feasible γ satisfy

$$\|\bar{d}\|^2 + \frac{\gamma^2}{\|\bar{c}\|^2} \leq \frac{d_0^2}{\lambda}$$

$\Rightarrow \gamma^2 = \|\bar{c}\|^2 (d_0^2 - \lambda \|\bar{d}\|^2)/\lambda$ defines min and max

General case: $c_0 \neq 0$

Feasibility table becomes

	$\lambda + \frac{c_0^2}{\|\bar{c}\|^2} < 0$	$\lambda + \frac{c_0^2}{\|\bar{c}\|^2} = 0$	$\lambda + \frac{c_0^2}{\|\bar{c}\|^2} > 0$	
$d_0 + \gamma \frac{c_0}{\|\bar{c}\|^2} < 0$	s.f.	s.i.	s.i.	
$d_0 + \gamma \frac{c_0}{\|\bar{c}\|^2} = 0$	s.f.	w.i.	s.i.	
$d_0 + \gamma \frac{c_0}{\|\bar{c}\|^2} > 0$	s.f.	s.f.	*s.f., w.f., s.i.*	(*)

Condition (*) is

$$(\lambda + \frac{c_0^2}{\|\bar{c}\|^2})(\|\bar{d}\|^2 + \frac{\gamma^2}{\|\bar{c}\|^2}) <,=,> (d_0 + \gamma \frac{c_0}{\|\bar{c}\|^2})^2$$

Observations

First column: *unbounded* problems

Second column: problem is not *attained*

General case: $c_0 \neq 0$ (cont.)

	$\lambda + \frac{c_0^2}{\|\bar{c}\|^2} < 0$	$\lambda + \frac{c_0^2}{\|\bar{c}\|^2} = 0$	$\lambda + \frac{c_0^2}{\|\bar{c}\|^2} > 0$	
$d_0 + \gamma \frac{c_0}{\|\bar{c}\|^2} < 0$	s.f.	s.i.	s.i.	
$d_0 + \gamma \frac{c_0}{\|\bar{c}\|^2} = 0$	s.f.	w.i.	s.i.	
$d_0 + \gamma \frac{c_0}{\|\bar{c}\|^2} > 0$	s.f.	s.f.	*s.f., w.f., s.i.*	(*)

Observations (cont.)

Second column: *infimum* and *supremum* for problem are not *attained*

- $c_0 > 0$: infimum at $-\frac{d_0\|\bar{c}\|^2}{c_0}$, unbounded from above
- $c_0 < 0$: supremum at $-\frac{d_0\|\bar{c}\|^2}{c_0}$, unbounded from below

Third column: need to look at condition (*):

equality case $\Leftrightarrow$ w.f. $\Leftrightarrow$ *attained* min and max

General case: $c_0 \neq 0$ (cont.)

When $\lambda + \frac{c_0^2}{\|\bar{c}\|^2} > 0$, one has to solve

$$(\lambda + \frac{c_0^2}{\|\bar{c}\|^2})(\left\|\bar{d}\right\|^2 + \frac{\gamma^2}{\|\bar{c}\|^2}) = (d_0 + \gamma \frac{c_0}{\|\bar{c}\|^2})^2$$

Quadratic equation whose solutions are given by

$$\gamma = \frac{c_0 d_0 \pm \sqrt{(d_0^2 - \lambda \left\|\bar{d}\right\|^2)(c_0^2 + \lambda \|\bar{c}\|^2)}}{\lambda}$$

which define min and max for the problem.

Generalizes correctly previous special cases

Direction of vectors $\bar{c}$ and $\bar{d}$ is *irrelevant*!

Duality

Primal problem

$$p^* = \inf_{\bar{x} \in \mathbb{R} \times \mathbb{R}^n} \bar{c}^{\mathrm{T}} \bar{x} \quad \text{s.t.} \quad P\bar{x} = \bar{d} \text{ and } \bar{x} \in \mathbb{L}^n$$

with $\bar{d} \in \operatorname{range} \bar{A}^{\mathrm{T}}$ and $\bar{c} \in \operatorname{null} \bar{A}$

Dual problem

$$d^* = \inf_{\bar{z} \in \mathbb{R} \times \mathbb{R}^n} \bar{d}^{\mathrm{T}} \bar{z} \quad \text{s.t.} \quad Q\bar{z} = \bar{c} \text{ and } \bar{z} \in \mathbb{L}^n$$

with $Q = I - P$ (orthogonal projection on null $\bar{A}$)

Properties

$p^* + d^* \geq 0$ (weak duality)

$p^* + d^* = 0$ (strong duality) under Slater condition

Primal and dual share the **same** format

Solving the dual problem

Reuse table for the primal while exchanging

$$\bar{c} \leftrightarrow \bar{d},\ P \leftrightarrow Q \text{ and } \lambda \leftrightarrow -\lambda$$

In particular, formula for min and max simply *changes its sign*

$$\gamma = \frac{c_0 d_0 \pm \sqrt{(d_0^2 - \lambda \left\|\bar{d}\right\|^2)(c_0^2 + \lambda \left\|\bar{c}\right\|^2)}}{\lambda}$$

$$\rightarrow \quad \gamma = \frac{d_0 c_0 \pm \sqrt{(c_0^2 + \lambda \left\|\bar{c}\right\|^2)(d_0^2 - \lambda \left\|\bar{d}\right\|^2)}}{-\lambda}$$

$\Rightarrow$ *duality gap* is *equal to* **zero**

Conclusions

Summary and perspectives

- Conic problems with a single second-order cone constraint can be solved analytically

  Bad cases (when strong duality fails) are detected and handled

- Interior-point method complexity for this problem is $\mathcal{O}\left(n^3 \log \frac{1}{\varepsilon}\right)$ this procedure is $\mathcal{O}\left(m^2 n\right) \rightarrow$ dependence on accuracy removed

- Application of the same technique to *dual* problem (study of *gap*)

- Possible *generalizations*
  - Allow additional *free* variables $\rightarrow$ still possible
  - Allow linear *inequality* constraints $\rightarrow$ quadratic programming
  - Use of this result for several second-order cones (*subproblem*?)