Research overview

Imre Pólik

Lehigh University
Department of Industrial & Systems Engineering

COR@L Seminar
September 4, 2008

Outline

1. Nonstandard duality concepts
   - Nonconvex quadratic optimization
   - Duality without regularity condition
   - Duality in non-exact arithmetic

2. Implementation of interior point methods
   - SeDuMi
   - Reimplementation of SeDuMi in Python
   - New input format for mixed LP/SOCP/SDP/etc...
   - Mixed integer conic optimization
   - Rounding procedures for conic optimization
   - Improved preprocessing techniques

The Lagrange-Slater dual

Primal	Dual
$f(x) < 0$	$f(x) + y^T g(x) \geq 0 \quad \forall x \in \mathcal{C}$
$g(x) \leq 0$	$y \geq 0$
$x \in \mathcal{C}$	

If $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^m$, $\mathcal{C} \subseteq \mathbb{R}^n$, and

- $\mathcal{C}$ is a convex set, f and g are convex functions
- there is $\bar{x} \in \operatorname{rint} \mathcal{C}$ such that $g(\bar{x}) < 0$ (Slater point),

Weak duality Primal is solvable $\Longrightarrow$ dual is unsolvable

Strong duality Primal is unsolvable $\Longrightarrow$ dual is solvable

- Why duality?
  - decide solvability
  - characterize optimality
  - drive algorithms

Nonconvex quadratic systems

Theorem (S-lemma, Yakubovich (1972))

$$
\begin{aligned}
x^T A x &< 0 & \qquad A + \lambda B &\succeq 0 \\
x^T B x &\le 0 & \qquad \lambda &\ge 0
\end{aligned}
$$

If $A, B \in \mathbb{R}^{n\times n}$ not necessarily PSD, and

- *there is $\bar{x} \in \mathbb{R}^n$ such that $\bar{x}^T B \bar{x} < 0$ (Slater point), then*

Weak duality *If the primal is solvable then the dual is unsolvable.*

Strong duality *If the primal is unsolvable then the dual is solvable.*

Why does it work? Generalization? Sufficient conditions?

Three approaches

1. Convexity of
$\{(x^T A_1 x, \dots, x^T A_m x) : x \in \mathbb{R}^n, (\|x\| = 1)\}$
   - classical area
   - joint numerical range
   - separation arguments
   - results over complex numbers

2. Low-rank solutions of $\mathcal{A}X = b,\ X \succeq 0$
   - more recent area
   - real: rank-1, complex: rank-2
   - equivalent to convexity

3. Generalized convexity
   - $x \mapsto (x^T A x, x^T B x)$ is König convex
   - modern area
   - easy theorems, hard conditions
   - abstract description

New duality theorem

Theorem

If $A_1, \ldots, A_m$ are all linear combinations of two fixed matrices then the solvability of

$$x^T A_i x \leq h_i, i = 1, \ldots, m$$
$$x \in \mathbb{R}^n$$

is equivalent to the nonsolvability of

$$\sum_{i=1}^{m} y_i A_i \succeq 0$$
$$y^T h < 0$$
$$y \geq 0.$$

(Conjectured in J.F. Sturm, S. Zhang, *On cones of nonnegative quadratic functions*, Maths of OR, 28 (2003), pp. 246–267.)

Lagrange-Slater dual for conic optimization

$$\begin{aligned} \max\ b^T y \qquad & \qquad \min\ c^T x \\ A^T y + s = c \qquad & \qquad Ax = b \\ s \in \mathcal{K} \qquad & \qquad x \in \mathcal{K}^*, \end{aligned}$$

Weak duality x, y, s: $c^T x - b^T y \geq 0$ (duality gap)

Strong duality If one problem is strictly feasible (Slater)

- the other problem is solvable
- zero duality gap at optimality

The regularized problem

- Minimal cone ($\mathcal{K}_{\min}$): spanned by the feasible solutions
- Regularized problem (Borwein-Wolkowicz)

$$\begin{array}{rrr} \max\ b^T y & \max\ b^T y & \min\ c^T x \\ A^T y + s = c & A^T y + s = c & Ax = b \\ s \in \mathcal{K} & s \in \mathcal{K}_{\min} & x \in \mathcal{K}_{\min}^* \end{array}$$

  - equivalent, Slater regular
- What is $\mathcal{K}_{\min}$ and $\mathcal{K}_{\min}^*$?
  - construction, structure, complexity?
- Contribution: $\mathcal{K}_{\min}$, $\mathcal{K}_{\min}^*$ for symmetric cones

Definition (Vinberg, 1963)

$\mathcal{K}$ is *homogeneous* if for all $u, v \in \operatorname{int} \mathcal{K}$ there is a linear map M such that $Mu = v$ and $M\mathcal{K} = \mathcal{K}$.
$\mathcal{K}$ is *symmetric* if it is homogeneous and self dual ($\mathcal{K} = \mathcal{K}^*$).

Exact dual for the symmetric conic case

$$\begin{aligned} \max\ & b^T y \\ A^T y + s &= c \\ s &\in \mathcal{K} \end{aligned}$$

$$\begin{aligned} \min\ & c^T(x + z^L) \\ A(x + z^L) &= b \\ c^T(x^i + z^{i-1}) &= 0,\ i = 1, \dots, L \\ A(x^i + z^{i-1}) &= 0,\ i = 1, \dots, L \\ z^0 &= 0 \\ x^i - B(z^i, z^i) &\in \mathcal{K},\ i = 1, \dots, L \\ x &\in \mathcal{K} \end{aligned}$$

Exact dual for the symmetric conic case

$$
\begin{aligned}
\max\ & b^T y \\
A^T y + s &= c \\
s &\in \mathcal{K}
\end{aligned}
$$

$$
\begin{aligned}
\min\ & c^T(x + z^L) \\
A(x + z^L) &= b \\
c^T(x^i + z^{i-1}) &= 0,\ i = 1, \dots, L \\
A(x^i + z^{i-1}) &= 0,\ i = 1, \dots, L \\
z^0 &= 0 \\
x^i - B(z^i, z^i) &\in \mathcal{K},\ i = 1, \dots, L \\
x &\in \mathcal{K}
\end{aligned}
$$

Exact dual for the symmetric conic case

$$\begin{aligned} \max\ & b^T y \\ A^T y + s &= c \\ s &\in \mathcal{K} \end{aligned}$$

$$\begin{aligned} \min\ & c^T(x + z^L) \\ A(x + z^L) &= b \\ c^T(x^i + z^{i-1}) &= 0,\ i = 1, \dots, L \\ A(x^i + z^{i-1}) &= 0,\ i = 1, \dots, L \\ z^0 &= 0 \\ x^i - B(z^i, z^i) &\in \mathcal{K},\ i = 1, \dots, L \\ x &\in \mathcal{K} \end{aligned}$$

Exact dual for the symmetric conic case

$$\begin{aligned} \max\ & b^T y \\ A^T y + s &= c \\ s &\in \mathcal{K} \end{aligned}$$

$$\begin{aligned} \min\ & c^T(x + z^L) \\ A(x + z^L) &= b \\ c^T(x^i + z^{i-1}) &= 0,\ i = 1, \ldots, L \\ A(x^i + z^{i-1}) &= 0,\ i = 1, \ldots, L \\ z^0 &= 0 \\ x^i - B(z^i, z^i) &\in \mathcal{K},\ i = 1, \ldots, L \\ x &\in \mathcal{K} \end{aligned}$$

Exact dual for the symmetric conic case

$$\max\ b^T y \qquad \min\ c^T(x+z^L)$$

$$A^T y + s = c \qquad A(x+z^L) = b$$

$$s \in \mathcal{K} \qquad c^T(x^i + z^{i-1}) = 0,\ i = 1,\dots,L$$

$$A(x^i + z^{i-1}) = 0,\ i = 1,\dots,L$$

$$z^0 = 0$$

Siegel cone $\quad (x^i, z^i, 1) \in \mathrm{SC}(\mathcal{K}, B),\ i = 1,\dots,L$

$$x \in \mathcal{K}$$

Exact dual for the symmetric conic case

$$
\begin{array}{rr}
\max\ b^T y & \min\ c^T(x+z^L) \\
A^T y + s = c & A(x+z^L) = b \\
s \in \mathcal{K} & c^T(x^i + z^{i-1}) = 0,\ i = 1,\dots,L \\
& A(x^i + z^{i-1}) = 0,\ i = 1,\dots,L \\
& z^0 = 0 \\
\text{Siegel cone} & (x^i, z^i, 1) \in \mathrm{SC}(\mathcal{K}, B),\ i = 1,\dots,L \\
& x \in \mathcal{K}
\end{array}
$$

- Dual uses homogeneous cones
- Zero duality gap
- Primal feasible: primal bounded ⇔ dual feasible
- Cone complexity: $L(\mathrm{rank}(\mathcal{K})+1)+\mathrm{rank}(\mathcal{K})$
- SDP: n^2+2n (improves $2n^2+n$, Ramana (1996))
- SOCP: $8k$ (improves $2nk$, standard result)

Duality in non-exact arithmetic

$$\begin{aligned} &\min\ c^T x && \max\ b^T y \\ &Ax = b && A^T y + s = c \\ &x \in \mathcal{K} && s \in \mathcal{K} \end{aligned}$$

where $x, c, s \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$, $b, y \in \mathbb{R}^m$ and $\mathcal{K} \subset \mathbb{R}^n$ is a closed, convex, pointed, solid, self-dual cone

$$\begin{aligned} \alpha_x = \quad &\inf \|x\| & \beta_u = \quad &\inf \|u\| \\ &Ax = b & &A^T y - u \in -\mathcal{K} \\ &x \in \mathcal{K} & &b^T y = 1, \end{aligned}$$

Theorem (Approximate duality, Sturm, 1998)

$$\alpha_x \beta_u = 1$$

New stopping criteria

Algorithm: IPM with self-dual embedding

Large norm: unboundedness or infeasibility

Theorem

If $b^T y \geq (\tau \|c\| + \theta \|\bar{c}\|) \rho$ *then for every feasible solution* x *of the primal problem we have* $\|x\| \geq \rho$.

Theorem

If $\tau \leq \frac{1-\beta}{1+\rho}$ *then every optimal solution* x^*, (y^*, s^*) *of the original primal-dual problem has*

$$x^{*T} s_0 + s^{*T} x_0 \geq \rho.$$

Theorem

IPM complexity for both cases is $\mathcal{O}(\sqrt{\operatorname{rank}(\mathcal{K})} \log(\rho/\varepsilon))$.

Practical issues about feasibility

- Importance of infeasibility
  - No solution.
    - Why? Certificate!
  - What does it mean?
    - Good news?
    - Wrong model? Wrong data?
    - Numerical problems?
    - Bug in the code?
- Practical problems
  - Not known a priori
  - Feasible but impractical solution
    - Missing constraints
  - Weakly infeasible problems

SeDuMi

- Optimization over symmetric cones
  - linear
  - second order
  - semidefinite
  - complex variables
  - free variables

SeDuMi

- Optimization over symmetric cones
    - linear
    - second order
    - semidefinite
    - complex variables
    - free variables
- Interior point method
    - primal-dual
    - self dual embedding
    - predictor-corrector scheme

SeDuMi

- Optimization over symmetric cones
  - linear
  - second order
  - semidefinite
  - complex variables
  - free variables
- Interior point method
  - primal-dual
  - self dual embedding
  - predictor-corrector scheme
- Open source
  - GPL
  - Matlab, C

SeDuMi

- Optimization over symmetric cones
  - linear
  - second order
  - semidefinite
  - complex variables
  - free variables
- Interior point method
  - primal-dual
  - self dual embedding
  - predictor-corrector scheme
- Open source
  - GPL
  - Matlab, C
- Widely used: both industry and academics
- Active user group, forum

History

late 1997	Jos F. Sturm starts SeDuMi
summer 1998	SeDuMi 1.0
November 2002	SeDuMi 1.05R5 (last version) • robustness, accuracy • general success
November 2003	Jos dies • Who will continue?
October 2004	AdvOL at McMaster takes over • How to continue?
June 2005	SeDuMi 1.1 (new version)
October 2006	SeDuMi 1.1R3
May 2007	Experimental parallel version
April 2008	SeDuMi 1.2 (64 bit version)

Usability

- AMPL, GAMS, AIMMS, MPL,...: no support for SDP
- YALMIP (Johan Löfberg)
- CVX (Michael C. Grant)
- Gloptipoly (Didier Henrion)
- SOSTools (Stephen Prajna et al.)

Strengths

- high numerical accuracy
- robustness
- efficient sparse system handling
- mixed second-order/semidefinite problems
- Matlab

Weaknesses

- large dense problems
- memory requirements
- embeddability
- Matlab

Reimplementation of SeDuMi in Python

- New data structures (sparse/dense)
- Improved performance, memory
- Extended functionality
- Platform independence
- New input format needed

New input format for mixed LP/SOCP/SDP/etc...

- Mixed linear/second-order/semidefinite optimization
- Sparse and dense representations
- General linear operators
- Rank-one/low rank constraint matrices
- Cone intersections
- More cones
- More objectives
- Portability

Mixed integer conic optimization

- Mostly open area
- We have
    - valid linear cuts
- We need
    - quick resolve (warmstart)
    - efficient implementations
    - generating valid conic cuts
    - efficient branching

Rounding procedures for conic optimization

- IPMs provide approximate solutions
- Can we improve them?
- Some hope for SOCP, less for SDP
- Theory: structure of the solution?
- Special cases, applications

Improved preprocessing techniques

- Essential for LP, little done for ConeP
- Simplify problem structure
  - decomposition
  - exploiting symmetry
  - special problem structures (graphs)
- Detecting redundancy in conic optimization
- Treating fixed variables

Research overview

Imre Pólik

Lehigh University
Department of Industrial & Systems Engineering

COR@L Seminar
September 4, 2008