Three Excursions around Conic Duality

Will Wright

Dept of Mathematics, UC Davis

https://www.math.ucdavis.edu/∼willwright
willwright@math.ucdavis.edu

April 28, 2017

SDP_subspace_test(2000, 5, 'sedumi')

$$
\text{(SDP-P)} \quad \begin{array}{ll} \min\limits_{X} & \langle C, X \rangle \\ \text{s.t.} & A(X) = b \\ & x \in \mathbb{S}_+^n \end{array}
$$

$$
\langle C, X \rangle = \mathrm{tr}(C^T X) \quad A(X) := \begin{bmatrix} \langle A_i, X \rangle \\ \vdots \\ \langle A_m, X \rangle \end{bmatrix}
$$

$\mathbb{S}_+^n$: positive semidefinite matrices

$A_i \in \mathbb{S}_+^{2000}$ for $i = 1, \ldots 5$
$C \in \mathbb{S}_+^{2000}$
All dense

Conic duality

Recall the primal and dual linear programs

$$\text{(LP-P)} \quad \begin{array}{ll} \min\limits_{x} & c^T x \\ \text{s.t.} & Ax = b \\ & x \geq 0 \end{array} \qquad\qquad \text{(LP-D)} \quad \begin{array}{ll} \max\limits_{y,z} & b^T y \\ \text{s.t.} & c - A^T y = z \\ & z \geq 0 \end{array}$$

Question: How can we generalize the inequality $x \geq 0$ and preserve

- symmetry ($x \geq 0$ and $z \geq 0$)?
- barrier properties (interior point tools)?

Intro to conic duality

Definition

A set $\mathcal{K} \subseteq \mathbb{R}^n$ is a **cone** if for all $x \in \mathcal{K}$ and $\alpha \geq 0$, we have $\alpha x \in \mathcal{K}$.

Definition

A cone $\mathcal{K}$ is **proper** if it is closed, pointed $(\mathcal{K} \cap -\mathcal{K} = \{0\})$, and nonempty $(\mathcal{K} + (-\mathcal{K}) = \mathbb{R}^n)$.

Examples

1) $\mathcal{K} = \mathbb{R}_+ = \{x \in \mathbb{R}^n \mid x \geq 0\}$
2) $\mathcal{K} = \mathcal{K}_2 = \{x = (x_0, \bar{x}) \in \mathbb{R} \times \mathbb{R}^{n-1} \mid ||\bar{x}|| \leq x_0,\ x_0 \geq 0\}$ (draw!)
   AKA the Lorentz cone, or "ice cream cone"
3) $\mathcal{K} = \mathbb{S}^n_+ = \{X \in \mathbb{R}^{n \times n} \mid X \succeq 0\}$

Intro to conic duality

Definition

Given a cone $\mathcal{K} \subset \mathbb{R}^n$, the **dual cone** of $\mathcal{K}$ is the set

$$\mathcal{K}^* = \{y \mid x^T y \geq 0 \text{ for all } x \in \mathcal{K}\}$$

Examples

1) $\mathcal{K} = \mathbb{R}_+ \implies \mathcal{K}^* = \mathcal{K}$
2) $\mathcal{K} = \mathcal{K}_2 \implies \mathcal{K}^* = \mathcal{K}$
3) $\mathcal{K} = \mathbb{S}^n_+ \implies \mathcal{K}^* = \mathcal{K}$

Self-dual cones: primal-dual symmetry, great for optimization methods.
Theorem: Every real, self-dual cone is a Cartesian product of $\mathbb{R}_+, \mathcal{K}_2$, and $\mathbb{S}^n_+$.

Conic primal and dual

Let $\mathcal{K}$ be a cone in $\mathbb{R}^n$, $A(\cdot)$ a linear operator, and $\langle \cdot, \cdot \rangle$ an inner product.

$$
\text{(CP-P)} \quad \begin{array}{ll} \min_x & \langle c, x \rangle \\ \text{s.t.} & A(x) = b \\ & x \in \mathcal{K} \end{array}
\qquad\qquad
\text{(CP-D)} \quad \begin{array}{ll} \max_{y,z} & b^T y \\ \text{s.t.} & c - A^*(y) = z \\ & z \in \mathcal{K}^* \end{array}
$$

Conic duality includes:

- (LP) linear programming
- (SOCP) second-order cone programming
- (SDP) semidefinite programming

Second-order cone and semidefinite programming

$$\text{(SOCP-P)} \quad \begin{aligned} \min_{x} \quad & c^T x \\ \text{s.t.} \quad & Ax = b \\ & x \in \mathcal{K}_2 \end{aligned} \qquad \text{(SOCP-D)} \quad \begin{aligned} \max_{y,z} \quad & b^T y \\ \text{s.t.} \quad & c - A^T y = z \\ & z \in \mathcal{K}_2 \end{aligned}$$

$$\text{(SDP-P)} \quad \begin{aligned} \min_{X} \quad & \langle C, X \rangle \\ \text{s.t.} \quad & A(X) = b \\ & X \in \mathbb{S}^n_+ \end{aligned} \qquad \text{(SDP-D)} \quad \begin{aligned} \max_{y,Z} \quad & b^T y \\ \text{s.t.} \quad & C - A^*(y) = Z \\ & Z \in \mathbb{S}^n_+ \end{aligned}$$

$$\text{(LP)} \subset \text{(SOCP)} \subset \text{(SDP)} \subset (CP) \subset \text{convex optimization}$$

Talk Outline

1. Introduction
2. Linear programming and conic duality
   - Lagrangian, finding duals
   - Conic duality theorem
3. Second-order cone programming
   - Jordan algebra, KKT conditions
   - Barrier method, interior point
   - ADMM, 1st order projection method
4. Semidefinite programming
   - KKT conditions
   - New(-ish) subspace method

Recall the (LP) primal and dual

$$\text{(LP-P)} \quad \begin{array}{ll} \min\limits_{x} & c^T x \\ \text{s.t.} & Ax = b \\ & x \geq 0 \end{array} \qquad\qquad \text{(LP-D)} \quad \begin{array}{ll} \max\limits_{y,z} & b^T y \\ \text{s.t.} & c - A^T y = z \\ & z \geq 0 \end{array}$$

Duality

$$\text{(LP-P)} \quad \begin{array}{ll} \min\limits_{x} & c^T x \\ \text{s.t.} & Ax = b \\ & x \geq 0 \end{array} \qquad \text{(CP-P)} \quad \begin{array}{ll} \min\limits_{x} & \langle c, x \rangle \\ \text{s.t.} & A(x) = b \\ & x \in \mathcal{K} \end{array}$$

Questions:

- How to find the dual of (*dualize*) (LP), (CP)? (A: Lagrangian.)
- How do primal and dual feasibility/solvability inform each other?
- Can primal-dual be solved simultaneously? (A: Yes.)
- Why? How? (A: Cone symmetry.)
- Is this advantageous? (A: Yes!)

The Lagrangian

$$
\text{(LP-P)} \quad \begin{array}{ll} \min\limits_{x} & c^T x \\ \text{s.t.} & Ax = b \\ & x \geq 0 \end{array}
$$

Definition

Given the primal linear program (LP-P), the **Lagrangian** is

$$
L(x, y, z) = c^T x - y^T (Ax - b) - x^T z
$$

where y is the **multiplier** (dual variable) for $Ax = b$, and z is the **multiplier** for x.

- Frame primal and dual problems.
- Prove duality results, develop algorithms.
- Show necessary and sufficient conditions for solutions (KKT systems).

(LP) duality via the Lagrangian

$$L(x, y, z) = c^T x - y^T(Ax - b) - x^T z$$

Claim: (LP-P) $= \min_{x} \max_{\substack{y,z \\ z \geq 0}} L(x, y, z)$ and (LP-D) $= \max_{y,z} \min_{x \geq 0} L(x, y, z)$

Define dual function $g(x) = \max_{\substack{y,z \\ z \geq 0}} L(x, y, z)$

$$\begin{aligned} Ax \neq b \quad &\implies \quad g(x) = +\infty \\ &\implies \quad Ax = b \\ &\implies \quad \min_{x} g(x) = \min_{\substack{x \\ Ax=b}} \max_{z \geq 0} c^T x - x^T z \end{aligned}$$

(LP) duality via the Lagrangian

$$L(x, y, z) = c^T x - y^T (Ax - b) - x^T z$$

Claim: (LP-P) $= \min_{x} \max_{\substack{y,z \\ z \geq 0}} L(x, y, z)$ and (LP-D) $= \max_{y,z} \min_{x \geq 0} L(x, y, z)$

Define dual function $g(x) = \max_{\substack{y,z \\ z \geq 0}} L(x, y, z)$

$$\begin{aligned}
\text{Any } x_i < 0 \quad &\Longrightarrow \quad g(x) = +\infty \\
&\Longrightarrow \quad x \geq 0 \\
\text{Any } x_i z_i > 0 \quad &\Longrightarrow \quad \text{inner max not attained} \\
&\Longrightarrow \quad \min_{\substack{x \geq 0 \\ Ax = b}} \max_{z} c^T x - x^T z = \min_{\substack{x \geq 0 \\ Ax = b}} c^T x
\end{aligned}$$

Same idea gives (LP-D) $= \max_{y,z} \min_{x \geq 0} L(x, y, z)$

Interpretation of Lagrange multipliers

$$L(x, y, z) = c^T x - y^T (Ax - b) - x^T z$$

(LP-P) $\min_{x} \max_{y, z \atop z \geq 0} L(x, y, z)$ (LP-D) $\max_{y, z} \min_{x \geq 0} L(x, y, z)$

- Inner $\max_{y_i} -y_i(a_i^T x - b_i)$: “soft” penalty on $a_i^T x - b_i \neq 0$.
- Pointwise infimum implies dual problem is concave even if primal is **not** convex.

Theorem (LP Duality)

Let $\mathbb{R}^n_+$ be the nonnegative orthant in $\mathbb{R}^n$ with the primal-dual pair

$$\text{(LP-P)}\quad \min_x \; c^T x \quad \text{s.t.} \quad Ax = b, \quad x \in \mathbb{R}^n_+ \qquad\qquad \text{(LP-D)}\quad \max_{y,z} \; b^T y \quad \text{s.t.} \quad c - A^T y = z, \quad z \in \mathbb{R}^n_+$$

1) *(duality symmetry): The dual to (LP-D) is (LP-P).*
2) *(weak duality): If x is primal feasible and (y, z) are dual feasible, then $b^T y \leq c^T x$.*

Theorem (LP Duality)

Let $\mathbb{R}^n_+$ be the nonnegative orthant in $\mathbb{R}^n$ with the primal-dual pair

$$\text{(LP-P)} \quad \begin{array}{ll} \min\limits_{x} & c^T x \\ \text{s.t.} & Ax = b \\ & x \in \mathbb{R}^n_+ \end{array} \qquad\qquad \text{(LP-D)} \quad \begin{array}{ll} \max\limits_{y,z} & b^T y \\ \text{s.t.} & c - A^T y = z \\ & z \in \mathbb{R}^n_+ \end{array}$$

3) *The following are equivalent:*
   - i) *(LP-P) is feasible and bounded below.*
   - ii) *(LP-D) is feasible and bounded above.*
   - iii) *(LP-P) is solvable.*
   - iv) *(LP-D) is solvable.*
   - v) *Both (LP-P) and (LP-D) are feasible.*

Key: 2) and 3) give optimality conditions.
$Ax = b$, $c - A^T y = z$, $x, z \in \mathbb{R}^n_+$ and $x^T z = 0$

$$\implies (x, y, z) = (x^*, y^*, z^*)$$

KKT conditions for LPs

Definition

The following are the Karush-Kuhn-Tucker (KKT) optimality conditions for (LP)

$$\begin{array}{rcll} Ax & = & b & \text{primal feasability} \\ x & \geq & 0 & \text{primal feasability} \\ c - A^T y & = & z & \text{dual feasability} \\ z & \geq & 0 & \text{dual feasability} \\ x^T z & = & 0 & \text{complementarity} \end{array}$$

- linear (easy) constraints: $Ax = b$, $c - A^T = z$
- nonlinear (hard) constraints: $x, z \geq 0$, $x^T z = 0$

Coordinate-wise handling of $x, z \geq 0$: Simplex method.

Interior point: Smooth nonlinear constraints with twice-diff'able penalty.

KKT conditions for LPs

Definition

The following are the Karush-Kuhn-Tucker (KKT) optimality conditions for (LP)

$$\begin{array}{rcll} Ax & = & b & \text{primal feasability} \\ x & \geq & 0 & \text{primal feasability} \\ c - A^T y & = & z & \text{dual feasability} \\ z & \geq & 0 & \text{dual feasability} \\ x^T z & = & 0 & \text{complementarity} \end{array}$$

- linear (easy) constraints: $Ax = b$, $c - A^T = z$
- nonlinear (hard) constraints: $x, z \geq 0$, $x^T z = 0$

Question: What other classes of primal-dual pairs offer symmetric duality, nice optimality (KKT) conditions, etc.?

General conic duality

Let $\mathcal{K}$ be a cone in $\mathbb{R}^n$ with the primal-dual pair

$$
\text{(CP-P)} \quad \begin{array}{ll} \min\limits_{x} & \langle c, x \rangle \\ \text{s.t.} & A(x) = b \\ & x \in \mathcal{K} \end{array}
\qquad\qquad
\text{(CP-D)} \quad \begin{array}{ll} \max\limits_{y,z} & b^T y \\ \text{s.t.} & c - A^*(y) = z \\ & z \in \mathcal{K}^* \end{array}
$$

Then we have the Lagrangian

$$
L(x, y, z) = \langle c, x \rangle - y^T(A(x) - b) - \langle x, z \rangle
$$

Recall $\mathcal{K}^* = \{y \mid x^T y \geq 0 \text{ for all } x \in \mathcal{K}\}$

$$
\text{(CP-P)} \quad \min_{x} \max_{\substack{y,z \\ z \in \mathcal{K}^*}} L(x, y, z) \qquad\qquad \text{(CP-D)} \quad \max_{y,z} \min_{x \in \mathcal{K}} L(x, y, z)
$$

Theorem (Conic Duality)

Let $\mathcal{K}$ be a cone in $\mathbb{R}^n$ with the primal-dual pair

$$(CP\text{-}P) \quad \begin{array}{ll} \min\limits_{x} & \langle c, x \rangle \\ \text{s.t.} & A(x) = b \\ & x \in \mathcal{K} \end{array} \qquad\qquad (CP\text{-}D) \quad \begin{array}{ll} \max\limits_{y,z} & b^T y \\ \text{s.t.} & c - A^*(y) = z \\ & z \in \mathcal{K}^* \end{array}$$

1) *(duality symmetry): (CP-D) is conic, and the dual to (CP-D) is (CP-P).*
2) *(weak duality): If x is primal feasible and (y, z) are dual feasible, then $b^T y \leq \langle c, x \rangle$.*

Theorem (Conic Duality)

Let $\mathcal{K}$ be a cone in $\mathbb{R}^n$ with the primal-dual pair

$$\text{(CP-P)} \quad \begin{array}{ll} \min_x & \langle c, x \rangle \\ \text{s.t.} & A(x) = b \\ & x \in \mathcal{K} \end{array} \qquad\qquad \text{(CP-D)} \quad \begin{array}{ll} \max_{y,z} & b^T y \\ \text{s.t.} & c - A^*(y) = z \\ & z \in \mathcal{K}^* \end{array}$$

3) *(strong duality with Slater condition): If (CP-P) is bounded below and strictly feasible ($\exists x$ with $A(x) = b$ and $x \in int(\mathcal{K})$) then (CP-D) is solvable with zero duality gap (and vice versa).*
4) *If (CP-P) is bounded below and strictly feasible, then x is (CP-P) optimal and (y, z) are (CP-D) optimal if and only if both hold*
   a) *(zero duality gap): $b^T y = \langle c, x \rangle$, and*
   b) *(comlementary slackness): $\langle x, z \rangle = 0$.*

Symmetric cone duals

$$
\text{(CP-P)} \quad \begin{array}{ll} \min\limits_{x} & \langle c, x \rangle \\ \text{s.t.} & A(x) = b \\ & x \in \mathcal{K} \end{array}
\qquad\qquad
\text{(CP-D)} \quad \begin{array}{ll} \max\limits_{y,z} & b^T y \\ \text{s.t.} & c - A^*(y) = z \\ & z \in \mathcal{K}^* \end{array}
$$

Goals:

- Apply conic duality results to symmetric cones: $\mathcal{K} = \mathcal{K}_2, \mathcal{K} = \mathcal{S}^n_+$?
- Utilize cone symmetry ($\mathcal{K}^* = \mathcal{K}$) in solver methods.

The second-order cone program (SOCP)

$$
\text{(SOCP-P)}\quad \begin{array}{ll} \min\limits_{x} & c^T x \\ \text{s.t.} & Ax = b \\ & x \in \mathcal{K}_2 \end{array}
\qquad\qquad
\text{(SOCP-D)}\quad \begin{array}{ll} \max\limits_{y,z} & b^T y \\ \text{s.t.} & c - A^T y = z \\ & z \in \mathcal{K}_2 \end{array}
$$

Recall $\mathcal{K}_2 = \{x = (x_0, \bar{x}) \in \mathbb{R} \times \mathbb{R}^{n-1} \mid ||\bar{x}|| \leq x_0,\ x_0 \geq 0\}$
$x \in \mathcal{K}_2$ handles general quadratic constraints:

Examples

1) $||A_i x + b_i|| \leq c_i^T x + d_i \iff \begin{pmatrix} A_i \\ c_i^T \end{pmatrix} x + \begin{pmatrix} b_i \\ d_i \end{pmatrix} \in \mathcal{K}_2$

2) $x^T Q_i x + b_i^T x + c_i \leq 0 \iff$

$$\left\| \begin{matrix} (1 + b_i^T x + c_i)/2 \\ \sqrt{Q_i} x \end{matrix} \right\| \leq (1 - b_i^T x - c_i)/2$$

Application

- filter design
- antenna array weight design
- truss design
- robust estimation
- model predictive control

KKT conditions for SOCPs

$$\begin{aligned} Ax &= b && \text{primal feasability} \\ x &\in \mathcal{K}_2 && \text{primal feasability} \\ c - A^T y &= z && \text{dual feasability} \\ z &\in \mathcal{K}_2 && \text{dual feasability} \\ x^T z &= 0 && \text{complementarity} \end{aligned}$$

Question: How to handle nonsmooth $x, z \in \mathcal{K}_2$, $x^T z = 0$
Answers:

- Jordan algebra with smooth product $x \circ z$
  $\rightarrow$ Barrier/penalty problem and interior point method
- Projection equivalence:
  $x, z \in \mathcal{K}_2$ and $x^T z = 0 \iff \Pi_{\mathcal{K}}(x - z) = x$
  $\rightarrow$ 1^{st}-order problem and ADMM/projection method

Jordan algebra of the second-order cone

Definition

Given $x = (x_0, \bar{x}), z = (z_0, \bar{z}) \in \mathbb{R} \times \mathbb{R}^{n-1}$, the **Jordan product** is

$$x \circ z = \begin{pmatrix} x^T z \\ x_0 \bar{z} + z_0 \bar{x} \end{pmatrix} = \mathrm{Arw}(x) z, \qquad \text{with } \mathrm{Arw}(x) := \begin{bmatrix} x_0 & \bar{x}^T \\ \bar{x} & x_0 I \end{bmatrix}$$

Basic Properties:

- (product identity): $e = (1, 0)$, $x \circ e = (x_0, \bar{x})$
- (commutative): $x \circ z = z \circ x$
- (bilinear): linear in x for fixed z and vice versa
- (non-associative): $x \circ (y \circ z) \neq (x \circ y) \circ z$ in general
- (Jordan associative): $x^2 \circ (z \circ x) = (x^2 \circ z) \circ x$

SOC spectral decomposition

$$x \circ z = \begin{pmatrix} x^T z \\ x_0 \bar{z} + z_0 \bar{x} \end{pmatrix} = \text{Arw}(x) z \qquad \text{Arw}(x) := \begin{bmatrix} x_0 & \bar{x}^T \\ \bar{x} & x_0 I \end{bmatrix}$$

Jordan product $\circ$ induces spectral decomposition of $\mathcal{K}_2$ (like $\mathbb{S}^n_+$)

$$\lambda_{1,2} = x_0 \mp ||\bar{x}||, \quad v_{1,2} = \frac{1}{2} \begin{pmatrix} 1 \\ \mp \bar{v} \end{pmatrix} \text{ s.t.} \begin{cases} \bar{v} = \bar{x}/||\bar{x}|| & \bar{x} \neq 0 \\ \bar{v} \text{ any unit vector} & \bar{x} = 0 \end{cases}$$

Properties: For all $x \in \mathcal{K}_2$,

- $x = \lambda_1 v_1 + \lambda_2 v_2$, with $\lambda_i \geq 0$ and $v_1^T v_2 = 0$, (hence notation $x \succeq_{\mathcal{K}_2} 0$)
- $x \in \text{int}(\mathcal{K}_2) \iff \lambda_i > 0$, (leads to barrier notion)
- $\text{tr}(x) = \lambda_1 + \lambda_2$, $\det(x) = \lambda_1 \lambda_2 = x_0^2 - ||\bar{x}||^2$
- $x^{-1} := \lambda_1^{-1} v_1 + \lambda_2^{-1} v_2$, $(x^{-1} \circ x = e)$
- $x^{1/2} := \lambda_1^{1/2} v_1 + \lambda_2^{1/2} v_2$, $(x^{1/2} \circ x^{1/2} = x)$

Jordan product and complementarity condition

Goal: Handle $x, z \in \mathcal{K}_2$ and $x^T z = 0$ "smoothly".

The following are equivalent:

i) $x, z \in \mathcal{K}_2$ and $x^T z = 0$

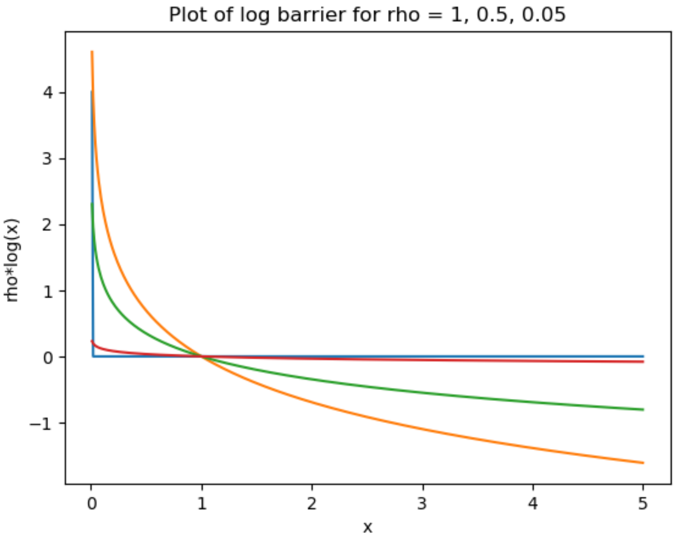
ii) $x, z \in \mathcal{K}_2$ and $x \circ z = 0$

(Proof by picture!)

Great news: Swapping $x^T z = 0$ for $x \circ z = 0$ gives

1. twice-differentiable term $x \circ z$ (for $x, z \in \text{int}(\mathcal{K}_2)$)
2. n constraints, square Newton system

Log barrier: $\phi_{\mathcal{K}}(x) := -\sum_{i=1}^{d} \log \lambda_i$, $\text{dom}(\phi_{\mathcal{K}}) = \text{int}(\mathcal{K})$



SOC log barrier

Log barrier: $\phi_{\mathcal{K}}(x) := -\sum_{i=1}^{d} \log \lambda_i = -\log(x_0^2 - \|\bar{x}\|^2)$
$\nabla \phi_{\mathcal{K}}(x) = -x^{-1} = -(\lambda_1^{-1} v_1 + \lambda_2^{-1} v_2)$
$\nabla^2 \phi_{\mathcal{K}}(x) = Q(x)^{-1} = Q(x^{-1})$
$(Q(x) := 2\mathrm{Arw}^2(x) - \mathrm{Arw}(x^2) = (2xx^T - (x^T J x) J))$
Note, (x, z) complementary if and only if one of the following holds:

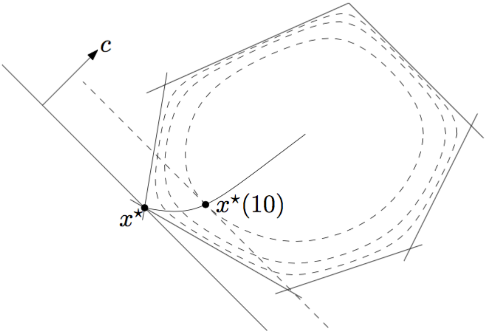
- $x = 0,\ z \in \mathrm{int}(\mathcal{K}_2)$
- $z = 0,\ x \in \mathrm{int}(\mathcal{K}_2)$
- $x, z \in \partial(\mathcal{K}_2)$

Thus $\phi_{\mathcal{K}}(x)$ or $\phi_{\mathcal{K}}(z) \to \infty$, as $(x, z) \to (x^*, z^*)$

SOC central path

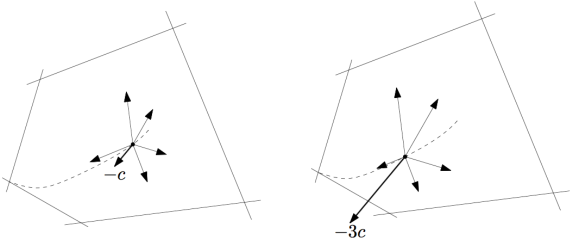
$$(\text{SOCP-P})_\rho \quad \begin{array}{ll} \min_x & c^T x + \rho\phi(x) \\ \text{s.t.} & Ax = b \\ & x \in \mathcal{K}_2 \end{array}$$

(central path): $\{(x(\rho), y(\rho), z(\rho) \mid \rho > 0)\}$



SOC central path

$$(\text{SOCP-P})_\rho \quad \begin{array}{ll} \min_x & c^T x + \rho\phi(x) \\ \text{s.t.} & Ax = b \\ & x \in \mathcal{K}_2 \end{array}$$



Question: How to build a *nice* Newton system?

Ans: Penalize dual z instead.

Barrier KKT conditions and Newton system

$$L_\rho(x, y, z) = c^T x - y^T(Ax - b) - x^T z - \rho\phi(z)$$

$$\nabla_z L_\rho = -x + \rho z^{-1} = 0 \iff x \circ z = \rho e$$

$$(\text{SOCP-KKT})_\rho \quad \begin{bmatrix} c - A^T y - z \\ Ax - b \\ \text{Arw}(x)z - \rho e \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$w^+ = (x^+, y^+, z^+) = (x + \Delta x, y + \Delta y, z + \Delta z), M = \nabla^2 L_\rho(w)$$

$$M\Delta w = \begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ \text{Arw}(z) & 0 & \text{Arw}(x) \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} = \begin{bmatrix} c - A^T y - z \\ b - Ax \\ \rho e - \text{Arw}(x)z \end{bmatrix} = \begin{bmatrix} r_x \\ r_y \\ r_z \end{bmatrix}$$

Barrier KKT conditions and Newton system

$$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ \mathrm{Arw}(z) & 0 & \mathrm{Arw}(x) \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} = \begin{bmatrix} r_x \\ r_y \\ r_z \end{bmatrix}$$

- (iteration): Generally just take one Newton step per ρ
- (factorize and pivot): $Arw(x)$ sparse
- (conditioning): $\mathrm{cond}(M) \sim \mathrm{cond}(\mathrm{Arw}(x))$
- (convergence): Residuals $\approx \mathcal{O}(\sqrt{\epsilon_{\mathrm{mach}}}) = 10^{-8}$

Question: How to handle large problems? ($n >> 1,000$)

ADMM: alternating direction method of multipliers

$$\begin{array}{ll} \min & f(x) + g(z) \\ \text{s.t.} & Ax + Bz = c \end{array}$$

Question: How to apply to KKT conditions on SOCP?

$$\begin{array}{rcll} Ax & = & b & \text{primal feasability} \\ x & \in & \mathcal{K}_2 & \text{primal feasability} \\ c - A^T y & = & z & \text{dual feasability} \\ z & \in & \mathcal{K}_2 & \text{dual feasability} \\ x^T z & = & 0 & \text{complementarity} \end{array}$$

(hint): $x, z \in \mathcal{K}_2$ and $x^T z = 0 \iff \Pi_{\mathcal{K}}(x - z) = x$

ADMM applied to SOCP

Homogeneous embedding of SOCP (self-dual form), $Qu = v$

$$v := \begin{bmatrix} z \\ 0 \\ \kappa \end{bmatrix} = \begin{bmatrix} 0 & A^T & c \\ -A & 0 & b \\ -c^T & -b^T & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ \tau \end{bmatrix} =: Qu$$

- (original variables): $(\hat{x}, \hat{y}, \hat{z}) = (x/\tau, y/\tau, z/\tau)$
- $(\tau, \kappa) = (1, 0)$ recovers standard primal-dual
- (τ, κ) act as primal-dual feasibility certificates

$\mathcal{C} := \mathcal{K} \times \mathbb{R}^n \times \mathbb{R}_+, \quad \mathcal{C}^* = \mathcal{K} \times \{0\}^n \times \mathbb{R}_+$

(indicator): $\delta_S(x) := \begin{cases} 0 & \text{if } x \in S \\ +\infty & \text{else} \end{cases}$

$$\begin{aligned} \min \quad & \delta_{\mathcal{C} \times \mathcal{C}^*}(u, v) + \delta_{Q\tilde{u} = \tilde{v}}(\tilde{u}, \tilde{v}) \\ \text{s.t.} \quad & (u, v) = (\tilde{u}, \tilde{v}) \end{aligned}$$

ADMM applied to SOCP

$$\begin{array}{ll} \min & \delta_{\mathcal{C}\times\mathcal{C}^*}(u,v) + \delta_{Q\tilde{u}=\tilde{v}}(\tilde{u},\tilde{v}) \\ \text{s.t.} & (u,v) = (\tilde{u},\tilde{v}) \end{array}$$

(λ, μ): dual multipliers from ADMM

$$\begin{array}{ll} (\tilde{u}^+, \tilde{v}^+) = & \Pi_{Qu=v}(u+\lambda, v+\mu) \\ u^+ = & \Pi_C(\tilde{u}^+ - \lambda) \\ v^+ = & \Pi_C^*(\tilde{v}^+ - \mu) \\ \lambda^+ = & \lambda - \tilde{u}^+ + u^+ \\ \mu^+ = & \mu - \tilde{v}^+ + v^+ \end{array}$$

- (implementation): Extremely easy, $\mathcal{O}(100)$ lines of code
- (main cost): Single initial factorization of $M = \begin{bmatrix} I & A^T \\ -A & I \end{bmatrix}$
- (iterations): Very cheap, one backsolve and one projection

State Primal and Dual SDP

$$\text{(SDP-P)} \quad \begin{array}{ll} \min\limits_{X} & \langle C, X \rangle \\ \text{s.t.} & A(X) = b \\ & X \in \mathbb{S}^n_+ \end{array} \qquad \text{(SDP-D)} \quad \begin{array}{ll} \max\limits_{y,Z} & b^T y \\ \text{s.t.} & C - A^*(y) = Z \\ & Z \in \mathbb{S}^n_+ \end{array}$$

$$\langle C, X \rangle = \mathrm{tr}(C^T X) \quad A(X) := \begin{bmatrix} \langle A_i, X \rangle \\ \vdots \\ \langle A_m, X \rangle \end{bmatrix} \quad A^*(y) := \sum_{i=1}^{m} y_i A_i$$

SDP Applications

- matrix recovery
- eigenvalue optimization
- anything with a linear matrix inequality ($A_0 + \sum_{i=1}^{m} y_i A_i \succeq 0$)

SDP KKT conditions

$$\begin{array}{rcll} A(X) & = & b & \text{primal feasability} \\ X & \in & \mathbb{S}^n_+ & \text{primal feasability} \\ c - A^*(y) & = & Z & \text{dual feasability} \\ Z & \in & \mathbb{S}^n_+ & \text{dual feasability} \\ \mathrm{tr}(X^T Z) & = & 0 & \text{complementarity} \end{array}$$

(barrier): $XZ = \rho I$, (like SOCP $x \circ z = \rho e$)

$$(\text{SDP-KKT})_\rho \quad \begin{bmatrix} C - A^*(y) - Z \\ A(X) - b \\ XZ - \rho I \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

- (factorize/pivot?): Unlike $\mathrm{Arw}(x)$, $\mathrm{rank}(X)$ unknown
- (question): How to solve large (SDP) with (possibly) low-rank X^*?

SDP subspace method [WW]

Goal: Find "optimal" $k-$dimensional subspace $\mathcal{V}$

- $\mathcal{V}_k^* := \text{span}\{v_1, \ldots, v_k\}$, k largest eigenvalues of X^*
- Optimize over smallest space possible

Key observation:

- $X^*, Z^* \succeq 0, \langle X, Z \rangle = 0$
$$\implies \text{ran}(X) \perp \text{ran}(Z) = \text{ran}(C - A^*(y))$$

Iteration-wise goals:

- Want $\mathcal{V}_k \to \mathcal{V}_k^*$
- $\mathcal{V}^+$: **find** $\lambda(C - A^*(y)) << 0$ and **toss** $\lambda(C - A^*(y)) > 0$
- y^+: cheap update (i.e., smallest subspace SDP solve)

SDP subspace method [WW]

Algorithm

1. Initialize: dual variable y_0, $\mathcal{V}$ subspace of $\mathbb{R}^n$
2. For iter $=$ 1 : iter_max
   - Set $V^+ =$ minimal/nonpositive eigenvectors of $(C - A^*(y))$
   - Toss any $v_i \in V$ with $v_i^T(C - A^*(y))v_i >> 0$
   - Set $V = \text{orth}[V, V^+]$
   - Build subspace problem: $A_i^{\mathcal{V}} = V^T A_i V$, $C^{\mathcal{V}} = V^T C V$
   - Solve tiny SDP: $[X, y] =$ SDP solver$(A^{\mathcal{V}}, b, C^{\mathcal{V}})$
   - Test for convergence

SDP subspace method [WW]

```
SDP_subspace_test(2000, 5, 'subspace')
```

Current method:

- Not tossing bad v_i's
- Not using subspace method for $y \in \mathbb{R}^m$
- Using fixed dimesion update for $\mathcal{V}^+$

Only known reference (I could find): Olivera, 2002

- Only rank 1 updates
- no theoretical results
- no X basis finesse

Thank you!!