Seminar: Mathematical Optimization and Quantum Information Theory

Introduction: Mathematical Optimization

§1 Conic optimization

E Euclidean space with inner product $\langle x, y \rangle$

$K \subseteq E$ proper convex cone ($\operatorname{int} K \neq \emptyset$, $K \cap (-K) = \{0\}$ (K pointed), $\overline{K} = K$)

gives partial order on E: $x \succeq_K y \Leftrightarrow x - y \in K$.

dual cone: $K^* = \{ y \in E : \langle x, y \rangle \geq 0 \ \forall x \in K \}$

main examples (a) nonnegative orthant $\mathbb{R}^n_{\geq 0} = (\mathbb{R}^n_{\geq 0})^*$, $\langle x, y \rangle = x^T y$.

(b) Lorentz cone $\mathcal{L}^{n+1} = \{ (x, t) \in \mathbb{R}^{n+1} : \|x\|_2 \leq t \}$
$= (\mathcal{L}^{n+1})^*$

(c) PSD cone $S^n_{\succeq 0} = (S^n_{\succeq 0})^*$ in the Euclidean space of sym. matrices S^n with $\langle X, Y \rangle = \operatorname{Tr}(XY)$;
$\dim S^n = \binom{n+1}{2}$.

(d) $CP^n = \operatorname{cone}\{ x x^T : x \in \mathbb{R}^n_{\geq 0} \} \subseteq S^n$

completely positive matrices

$COP^n = (CP^n)^* = \{ X \in S^n : x^T X x \geq 0 \ \forall x \in \mathbb{R}^n_{\geq 0} \}$

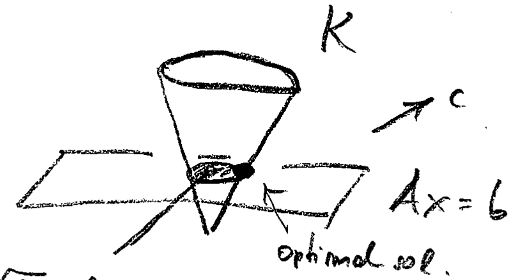
copositive matrices. $[CP^n \subsetneq S^n_{\succeq 0} \subsetneq COP^n]$.

<u>Conic programs</u> E, F Euclidean spaces, $A : E \to F$ linear map.

$c \in E$, $b \in F$

<u>primal</u> (P) $p^* = \sup \langle c, x \rangle$
$x \in K$
$Ax = b$

<u>dual</u> (D) $d^* = \inf \langle y, b \rangle$
$A^T y - c \in K^*$



$\mathcal{F}$ = feasible sol.

[dual of dual $\cong$ primal]

<u>main examples</u>

(a) LP (b) CQP (c) SDP

(d) copositive programming.

<u>duality theory</u> (i) <u>weak duality</u> $p^* \leq d^*$ [trivial]

(ii) <u>strong duality</u> if $d^* > -\infty$ and $\exists$ strictly feasible sol. y (i.e. $A^T y - c \in \operatorname{int} K^*$) $\Rightarrow$ $\exists\, x^*$ feasible for primal with $p^* = c^T x^*$ & $p^* = d^*$

[argument: separating hyperplane]

[similar for dual].

(iii) <u>optimality condition</u> x feasible for (P), y feasible for (D), (P) & (D) strictly feasible. Then x, y optimal $\Longleftrightarrow$

$x^T (A^T y - c) = 0.$

algorithms and complexity

(1) theory (very elegant; ellipsoid method; Grötschel, Lovász, Schrijver 1981)

can find ε-approx. opt. sol. of conic program in poly. time
if (α) know $r, R > 0$, $x_0 \in \mathcal{F}$: $x_0 + r\underbrace{B}_{\text{unit ball}} \subseteq \mathcal{F} \subseteq x_0 + RB$

(encoding of r, R, x_0 poly. size)

(β) can solve separation problem over K in poly. time.

(true for (a)-(c)).

issue with (α) for SDP: $\operatorname{diag}\left[\begin{bmatrix} 1 & x_{i-1} \\ x_{i-1} & x_i \end{bmatrix}, i = 1, \dots, n\right] \in S^{2n}_{\geq 0}$.

Then $x_i \geq x_{i-1}^2$. Set $x_0 = 2$, gives $x_i \geq 2^{2^i}$

computational difficulty

$\underbrace{(a) < (b) < (c)}_{\text{poly. time}} \ll \underbrace{(d)}_{\text{NP-hard}}$ [Motzkin-Straus thm. CLIQUE $\leq_p$ COPOSITIVE PROG.].

(2) practice (interior point method; Nesterov, Nemirovski 1994)

many numerical solvers available (for (a)-(c))

(c): e.g. csdp, sdpa, cvx

§2 Polynomial optimization

$\mathbb{R}[x_1,\dots,x_n]_d = \mathbb{R}[x]_d$ space of polynomials up to deg. d

$\dim \mathbb{R}[x]_d = \binom{n+d}{d}$.

Cone of nonnegative polynomials

$$\mathcal{P}_{n,d} = \{ p \in \mathbb{R}[x]_d : p(x) \geq 0 \quad \forall x \in \mathbb{R}^n \}. \qquad \left[\begin{matrix} \mathcal{P}_{n,d} \text{ is} \\ \text{NP hard} \end{matrix} \right]$$

SOS cone (sum of squares)

$$\Sigma_{n,d} = \{ p \in \mathbb{R}[x]_d : p \text{ is SOS} \}$$

$$\Updownarrow$$

$$\exists\, q_1,\dots,q_m \in \mathbb{R}[x]_d : p = \sum_{i=1}^m q_i^2$$

$\Sigma_{n,d}$ is poly. time, i.e. $\Sigma_{n,2d} = \{ p : \exists\, Q \in S_{\geq 0}^{\binom{n+d}{d}} : [x]_d^T Q [x]_d = p \}$,

where $[x]_d$ is a vector containing a basis of $\mathbb{R}[x]_d$.

Relations: $\Sigma_{n,2d+1} = \Sigma_{n,2d}$, $\mathcal{P}_{n,2d+1} = \mathcal{P}_{n,2d}$

$$\Sigma_{n,d} \subseteq \mathcal{P}_{n,d}$$

Hilbert (1888): $\Sigma_{n,2d} = \mathcal{P}_{n,2d} \iff n=1$ or $d=1$ or $(n=2$ and $d=2)$

Motzkin polynomial: $M(x,y) = x^2y^2(x^2+y^2-3)+1 \in \mathcal{P}_{2,6} \setminus \Sigma_{2,6}$.

polynomial optimization problem (POP)

<u>given</u>: polynomials $f, g_1, \ldots, g_m \in \mathbb{R}[x]$

<u>find</u>: $p_{min} = \inf \{ f(x) : x \in S \}$

with $S = \{ x \in \mathbb{R}^n : g_1(x) \geq 0, \ldots, g_m(x) \geq 0 \}$

* very general & very hard.

<u>dual approach</u>: $\mathcal{P}(S) = \{ p \in \mathbb{R}[x] : p(x) \geq 0 \ \forall x \in S \}$ convex cone.

Then $p_{min} = \sup \{ \lambda : f - \lambda \in \mathcal{P}(S) \}$.

<u>clear</u>: optimization over $\mathcal{P}(S)$ is hard.

<u>relaxation / SDP-SOS hierarchies</u> (Parrilo / Lasserre 2001)

* find sequence of cones $C_1 \subseteq C_2 \subseteq \ldots \subseteq \mathcal{P}(S)$ for which $p_i = \sup \{ \lambda : f - \lambda \in C_i \}$ is easier to solve.
* gives sequence: $p_1 \leq p_2 \leq \ldots \leq p_{min}$.
* SOS relaxation: $C_i = \Sigma_{n,2i} + g_1 \Sigma_{n,2i} + \cdots + g_m \Sigma_{n,2i}$.
* <u>Thm.</u> (Putinar 1993) If S compact and "nicely represented by $g_1, \ldots, g_m$" [the quadratic module $M(g_1, \ldots, g_m)$ is Archimedian], then $p_{min} = \lim_{i \to \infty} p_i$.
* often: $p_{min} = p_i$ for small values of i.