Strong relaxations for discrete optimization problems 10/05/16

Lecture: Cone programming. Approximating the Lorentz cone.

Lecturer: Yuri Faenza *Scribes: Igor Malinović*

1 Introduction

Cone programming is a broad generalization of linear programming. We focus on two well studied subclasses: *second-order cone programming* (SOCP) and *semidefinite programing* (SDP). In what follows, we will provide an overview of cone programming, skipping most proof, that can be found e.g. in [2].

As seen in previous classes, there are two equivalent ways to look at a polyhedral cone $K \subseteq \mathbb{R}^n$:

$$K = \{x : Ax \leq 0\} \Longleftrightarrow \exists \{r_1, \ldots, r_m\} \subseteq \mathbb{R}^n : \ K = \{x = \sum_{i=1}^{m} \lambda_i r_i, \ \lambda \geq 0\}.$$

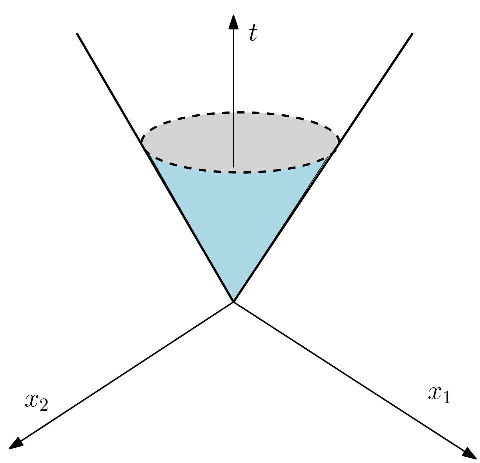
We generalize this notion as follows.

Definition 1. *Let V be a real finite-dimensional Hilbert space and $K \subseteq V$ be a nonempty closed set. K is called a closed convex cone if for any nonnegative scalars α, β and any $x, y \in K$ one has that $\alpha x + \beta y$ belongs to K.*

Examples:

1. $K = \mathbb{R}^n_+ = \{x : \mathbb{I}x \geq 0\}$;
2. Lorentz cones (also called "ice cream cones" or "second-order cones")

$$L^n = \{(x,t) \in \mathbb{R}^n \times \mathbb{R} : \|x\| \leq t\};$$



3. Semidefinite cones

$$S^n_+ = \{X \in \mathbb{R}^{n \times n} : X = X^T, \ v^T X v \geq 0 \quad \forall v \in \mathbb{R}^n\}.$$

The fact that the first two are closed convex cone (with $V = \mathbb{R}^n$) is immediate to check. For the last example, we defer the proof to next lecture. Given closed convex cones L and K, the corresponding *cone programming* (CP) problem is

$$\begin{array}{ll} \max & \langle c, x\rangle \\ \text{s.t.} & Ax - b \in L \\ & x \in K \end{array}$$

where $\langle , \rangle$ is the inner product of the Hilbert space. Example of those problems are:

LP:

$$\begin{array}{ll} \max & \langle c, x\rangle \\ \text{s.t.} & Ax = b \\ & x \geq 0 \end{array}$$

SOCP:

$$\begin{array}{ll} \max & \langle c, x\rangle \\ \text{s.t.} & Ax \geq b \\ & \|A^i x - b^i\| \leq c^i x - d^i \quad i = 1, \ldots, m \end{array}$$

Here SOCP stands for *second-order cone programming*. It is immediate to check that LP can be formulated as a cone programming problem, setting $L = \{0\}$ and $K = \mathbb{R}^n_+$. Standard transformations also show that SOCP can be formulated as a cone programming problem, by taking $L = \{0\}$ and K a Lorentz cone of appropriate size.

In order to define the dual cone program, it is useful to introduce the notion of a dual cone.

Definition 2. *Let $K \subseteq V$ be a closed convex cone. Its dual cone is given by*

$$K^* := \{y \in V : \langle x, y\rangle \geq 0 \quad \forall x \in K\}.$$

Exercise 3. *If K is a closed convex cone then K^* is also a closed convex cone.*

We can develop a duality theory for cone programming in an analogous way to linear programming:

	CP:	LP:
(P)	$\begin{array}{ll} \max & \langle c, x\rangle \\ \text{s.t.} & Ax - b \in L \\ & x \in K \end{array}$	$\begin{array}{ll} \max & \langle c, x\rangle \\ \text{s.t.} & Ax = b \\ & x \geq 0 \end{array}$
(D)	$\begin{array}{ll} \min & \langle b, y\rangle \\ \text{s.t.} & A^T y - c \in K^* \\ & y \in L^* \end{array}$	$\begin{array}{ll} \min & \langle b, y\rangle \\ \text{s.t.} & A^T y \geq c \end{array}$

The difference between general CP and LP is that Strong duality does not necessarily hold, i.e. it can be the case that both the dual and the primal are feasible but their optimal values do not coincide. The most famous sufficient conditions are given bellow, and we will need them in next lecture.

Theorem 4 (Slater condition for Strong duality)**.** *Let (P) be the primal problem with $L = \{0\}$ and suppose there exists a primal feasible point $\bar{x} \in \mathsf{int}(K)$. Then (D) is feasible and the optima have the same value.*

2 Modelling problems as SOCPs

A class of problems that can be remodelled in terms of SOCP are *quadratically constrained convex quadratic programs* (QCCQP):

$$\begin{aligned} \min \quad & \underbrace{x^T Q^T Q x}_{\langle Qx, Qx\rangle = \|Qx\|^2} -2b^T x \\ \text{s.t.} \quad & \underbrace{x^T Q_i^T Q_i x}_{\|Q_i x\|^2} -2b_i^T x \leq c_i \quad i = 1, \ldots, m. \end{aligned}$$

The optimum value of the above program coincides with the optimum value of the SOCP:

$$\begin{aligned} \min \quad & t \\ \text{s.t.} \quad & \left\| \begin{matrix} Qx \\ \frac{2b^T x + t - 1}{2} \end{matrix} \right\| \leq \frac{2b^T x + t + 1}{2}, \\ & \left\| \begin{matrix} Q_i x \\ \frac{2b_i^T x + t - 1}{2} \end{matrix} \right\| \leq \frac{2b_i^T x + t + 1}{2} \quad i = 1, \ldots, m. \end{aligned}$$

It is easy to see that the first constraint can be squared to obtain the following equivalent form:

$$\|Qx\|^2 - \frac{1}{2}(2b^T x + t) \leq \frac{1}{2}(2b^T x + t) \iff \|Qx\|^2 - 2b^T x \leq t$$

and similarly for the others

$$\|Q_i x\|^2 - 2b_i^T x \leq c_i.$$

Hence, if $(t, Q, \{Q_i\}_{i \in [m]})$ is an optimum solution to the SOCP above, then $(Q, \{Q_i\}_{i \in [m]})$ is an optimum solution to the QCQP problem and t its optimum value.

For another interesting application, imagine a linear program with some incertitude on the coefficient vector of the constraints. For each inequality $a^T x \leq b$, we know that a can take any value in the ellipsoid $\mathcal{E}(c, R) := \{c + Ru : \ \|u\| \leq 1\}$. We want therefore our vector x to satisfy

$$b \geq \max_{a \in \mathcal{E}(c,R)} a^T x = c^T x + \max_{u : \|u\| \leq 1} x^T R u.$$

and the optimum is reached for $u = \frac{R^T x}{\|R^T x\|}$ having the value $\langle x^T R, \frac{R^T x}{\|R^T x\|} \rangle = \|R^T x\|$. So finally

$$b \geq c^T x + \|R^T x\|.$$

which is again a second-order constraint.

3 Approximating Second-order cones

Second-order cone programming seems to be much more general then linear programming. Surprisingly, the feasible region of an SOCP program can be well-approximated by a polyhedral cone[1].

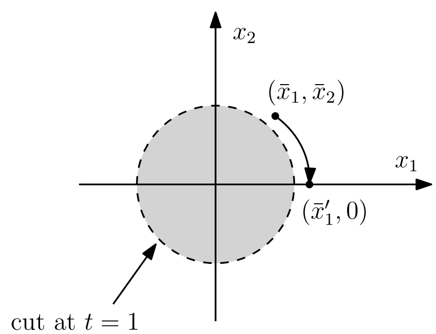
Definition 5. *Let Q be a feasible region of a SOCP, i.e. $Q := \{x : Ax \geq b,\ \|A^i x - b^i\| \leq c^i x - d^i \quad i = 1, \ldots, m\}$ and let $Q_\epsilon := \{x : Ax \geq b,\ \|A^i x - b^i\| \leq (c^i x - d^i)(1 + \epsilon) \quad i = 1, \ldots, m\}$. A polychedral cone P is an ϵ-approximation of Q if $Q \subseteq P \subseteq Q_\epsilon$.*

Our goal here is to find an ϵ-approximation of Q with small extension complexity. We will perform the following two steps:

1. Give an ϵ-approximation of the 3-dimensional Lorentz cone L^2;
2. Show how this approximation can be extended to any Lorentz cone, and then to an arbitrary Q as in Definition 5 above.

3.1 An ϵ-approximation of L^2

Recall that $L^2 = \{(x, t) \in \mathbb{R}^2 \times \mathbb{R} : \|x\| \leq t\}$. Given $\bar{x} = (\bar{x}_1, \bar{x}_2, \bar{t}) \in \mathbb{R}^3$, we would like to understand whether it lies in L^2. Assume wlog $\bar{t} = 1$, $\bar{x} \geq 0$. Ideally we would like to rotate the point $(\bar{x}_1,\ \bar{x}_2,\ 1)$ to $(\bar{x}'_1,\ 0,\ 1)$. Then $\bar{x} \in L^2$ iff $\bar{x}'_1 \leq 1$. Unfortunately, this rotation cannot be done with a linear map, but there is a simple linear transformation that reduces the angle between $\bar{x}$ and the x_1 axis by at least half. Iterating, we can make this angle arbitrarily small. We can now check whether the rotated point $\bar{x}'$ has $\bar{x}'_1 \leq 1$. If not, we know that $\bar{x} \notin L^2$. If yes, then still it may be that $\bar{x} \notin L^2$, but $\bar{x} \in L^2_\epsilon$ for an appropriate small ϵ, since the component $\bar{x}'_2$ is very close to 0.



The intuition above is formalized by the following polyhedron.

$$
\begin{aligned}
T := \{\ (x_1, x_2, \xi_0, \eta_0, \ldots, \xi_\nu, \eta_\nu, t)\ :\ \ & \xi_0 \geq |x_1|, && (1)\\
& \eta_0 \geq |x_2| && (2)\\
& \xi_j = \cos(\frac{\pi}{2^{j+1}})\xi_{j-1} + \sin(\frac{\pi}{2^{j+1}})\eta_{j-1}, \quad j = 1, \ldots, \nu && (3)\\
& \eta_j \geq |-\sin(\frac{\pi}{2^{j+1}})\xi_{j-1} + \cos(\frac{\pi}{2^{j+1}})\eta_{j-1}|, \quad j = 1, \ldots, \nu && (4)\\
& \xi_\nu \leq t, && (5)\\
& \eta_\nu \leq \tan(\frac{\pi}{2^{\nu+1}})\xi_\nu\}. && (6)
\end{aligned}
$$

where (ξ_0, η_0) rotates (x_1, x_2) so that $x_1, x_2 \geq 0$, and we have two additional variables ξ_j, η_j for each iteration $j \in [\nu]$. Note that the fact that $|\cdot|$ only appears in inequalities allows as to deduce via standar techniques that T as a polyhedron.

[1]Note that this does not imply in general that the optimum solution of an SOCP can be well-approximated by an LP. For instance, Q may be empty, but Q_ϵ and P may not.

Theorem 6. *$proj_{x_1,x_2,t}(T)$ is a polyhedral ϵ-approximation to L^2 for $\epsilon = \Omega(\frac{1}{4^\nu})$ and $xc(T) = O(\log\frac{1}{\epsilon})$.*

Proof. The statement on the extension complexity is immediate, so we focus on the rest of the proof.

Part I: $Q \subseteq \text{proj}_{x_1,x_2,t}(T)$.

Let $(\bar{x}_1, \bar{x}_2, \bar{t}) \in Q$, and extend it to a point $(x_1, x_2, \xi_0, \eta_0, \dots, \xi_\nu, \eta_\nu, t)$ by setting to equality constraints (1)–(4). Note that this extension is unique. Before showing that this point is in T, let us study how the sequence of points $(\xi_0, \eta_0), \dots, (\xi_\nu, \eta_\nu,)$ looks like. We set $(\xi_0, \eta_0) = (|\bar{x}_1|, |\bar{x}_2|)$, thus reflecting the initial point to the nonnegative orthant. So the angle between (ξ_0, η_0) and the x_1 axis verifies $\alpha(\xi_0, \eta_0) \le \frac{\pi}{2}$. Standard trigonometry implies that, if there was no modulus in (4), then (ξ_1, η_1) would be exactly the clockwise rotation by $\frac{\pi}{4}$ of (ξ_0, η_0). Through the modulus we are actually reflecting over the x_1-axis if $\eta_1 < 0$. So $\alpha(\xi_1, \eta_1) \le \frac{\pi}{4}$, and we can repeat the argument, deducing $\alpha(\xi_\nu, \eta_\nu) \le \frac{\pi}{2^{\nu+1}}$.

It remains to show that the last two constraints hold after repeating the operation ν times. (5) holds since:

$$\xi_\nu \le \|(\xi_\nu, \eta_\nu)\| = \dots = \|(\xi_0, \eta_0)\| = (\bar{x}_1^2 + \bar{x}_2^2)^{1/2} \le t.$$

Using that $\alpha(\xi_\nu, \eta_\nu) \le \frac{\pi}{2^{\nu+1}}$ and by the definition $\frac{\eta_\nu}{\xi_\nu} = \tan(\alpha(\xi_\nu, \eta_\nu))$ we obtain (6):

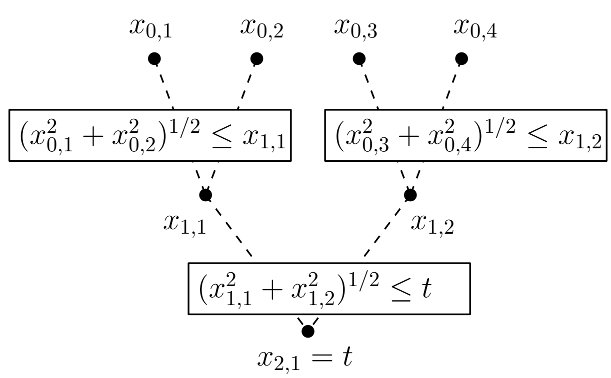
$$\eta_\nu = \tan(\alpha(\xi_\nu, \eta_\nu))\xi_\nu \le \tan(\frac{\pi}{2^{\nu+1}})\xi_\nu.$$

Part II: $\text{proj}_{x_1,x_2,t}(T) \subseteq Q_\epsilon$.

Given $(\bar{x}_1, \bar{x}_2, \bar{\xi}_0, \bar{\eta}_0, \dots, \bar{\xi}_\nu, \bar{\eta}_\nu, \bar{t}) \in T$, we would like to prove that $(\bar{x}_1, \bar{x}_2, \bar{t}) \in Q_\epsilon$. Since in each of (1),(2),(4), we have $\ge$, while (3) is an equation, the norm gets never decreased. We conclude

$$\|(\bar{x}_1, \bar{x}_2)\| \le \dots \le \left\|(\bar{\xi}_\nu, \bar{\eta}_\nu)\right\| = (\bar{\xi}_\nu^2 + \bar{\eta}_\nu^2)^{1/2} \overset{(6)}{\le} \bar{\xi}_\nu(1 + \tan^2(\frac{\pi}{2^{\nu+1}}))^{1/2} \overset{(5)}{\le} t \cdot (1 + \underbrace{\frac{1}{\cos(\frac{\pi}{2^{\nu+1}})} - 1}_{O(\frac{1}{4^\nu})}),$$

where we used the trigonometric identity $1 + \tan^2(\gamma) = 1/\cos^2(\gamma)$. ∎



3.2 Extending the ϵ-approximation

Assume wlog $n = 2^k$ for some $k \in \mathbb{N}$. We show that we can give an extended formulation for

$$L^n \in \{(x_{0,1}, \dots, x_{0,n}) \in \mathbb{R}^n \times \mathbb{R} : \|(x_{0,1}, \dots, x_{0,n})\| \le t\}$$

using $O(n)$ L^2 cones. Rename original variables as $x_{0,1}, \ldots, x_{0,n}$. Note that $(x_{0,1}, \ldots, x_{0,n}, t) \in L^n$ if and only if there exits a value $x_{1,1}$ such that $(x_{1,1}, x_{0,1}, \ldots, x_{0,n}, t,)$ satisfy

$$(x_{1,1}, x_{0,3}, x_{0,4}, \ldots, x_{0,n}, t) \in L^{n-1} \text{ and } (x_{0,1}, x_{0,2}, x_{1,1}) \in L^2.$$

We can now repeat the argument: split all the original variables into pairs $(x_{0,2j-1}, x_{0,2j})$ for $j \in [n/2]$ and associate to each pair a new "1st level" variable $x_{1,j}$. Apply the same procedure to "1st" level variables, introducing "2nd" level variables, etc. We construct therefore a binary tree where each i-th level variable has as children two $(i-1)$-th level variables. The same structure is preserved up to the level $\log_2(n)$ which has a single node $x_{\log_2(n),1} = t$.

At each level of the tree we have a ϵ-approximation of the L^2 cones with the construction from the previous section. This gives in total a $((1+\epsilon)^{\log_2(n)} - 1)$-approximation for L^n.

Last, observe that, if P_n is a polyhedral ϵ-approximation to L^n, then

$$\{x : Ax \geq b, (A^i x, c^i x - b^i) \in P_{k_i} \text{for } i = 1, \ldots, m\}$$

is a polyhedral ϵ-approximation to Q as in Definition 5, where k_i is the number of row of matrix A^i.

4 Notes

Results from Section 3 appeared in [1]. It is also shown there that, by carefully choosing different ϵ for different levels, one can obtain ϵ-approximation of L^n (hence of Q as in Definition 5) with slighty smaller extension complexity.

References

[1] A. Ben-Tal and A. Nemirovski. On polyhedral approximations of the second-order cone. *Math. Oper. Res.*, 26(2):193–205, May 2001.

[2] B. Gärtner and J. Matousek. *Approximation Algorithms and Semidefinite Programming.* Springer-Verlag, Berlin, Heidelberg, 2012.