Strong Relaxations for Discrete Optimization Problems 13/05/2016

Lecture : Yannakakis' theorem for Cone lifts and Semidefinite rank

Lecturer: Yuri Faenza *Scribes: Christoph Hunkenschröder*

1 Introduction

In the last lecture, we introduced Cone Programming, that aims at maximizing a linear function $\langle c, x\rangle$ under the constraints $Ax - b \in T$ and $x \in K$ for some closed convex cones T and K of some Hilbert space. Recall that Linear Programming arises from Cone Programming by simply setting $L = \{0\}$ and $K = \mathbb{R}^n_{\geq 0}$, and taking the Hilbert space to be the Euclidean space.

For Linear Programming, we studied Yannakakis' Theorem, which links the extension complexity of a polytope to the non-negative rank of its slack matrix. In this lecture, we are interested in finding an equivalent, more general statement for Cone Programming.

2 K-lifts and K-factorizations

Recall Yannakakis' Theorem in a slightly different terminology.

Theorem 1. *Let P be a polytope and S its slack matrix. $rk_+(S) \leq r$ if and only if P has an extension of size at most r, that is, if there exists a pair (π, Q) such that*

- *$Q \subseteq \mathbb{R}^m$ is a polyhedron in standard form, i.e. $Q = \{y : \ Ay = b, \ y \geq 0\}$, and*
- *$\pi : \ \mathbb{R}^m \to \mathbb{R}^n$ is an linear map.*
- *$P = \pi(Q) = \{x \in \mathbb{R}^n : \ \exists y \ s.t. \ x = \pi(y), \ Ay = b, \ y \geq 0\}$.*

Observe that $Ay = b$ defines an affine subspace, whereas $y \geq 0$ forces Q to be contained in the non-negative orthant.

These are the properties we will keep when extending this concept: P will be allowed to any convex body[1], and we allow K to be a cone different from the nonnegativity orthant.

An extension will become a *K-Lift*, the slack matrix will become the *slack operator*, as we will see later.

Definition 2. *Let $C \subset \mathbb{R}^n$ be a convex body, and $K \subseteq \mathbb{R}^m$ a closed, full-dimensional convex cone. A K-*Lift *of C is a pair (π, L) with the following properties.*

- *$L \subseteq \mathbb{R}^m$ is an affine subspace.*
- *$\pi : \ \mathbb{R}^m \to \mathbb{R}^n$ is an linear map.*
- *$C = \{x \in \mathbb{R}^n : \ \exists y \in \mathbb{R}^m \ s.t. \ x = \pi(y), \ y \in K, \ y \in L\} = \pi(K \cap L)$.*

We call the K-Lift proper, *if $L \cap \operatorname{int}(K) \neq \emptyset$.*

In other words, given C and K, we ask whether there is some affine space s.t. C is the (image of the) intersection of K with this affine space. For example, let $C = B^n_r$ (the n-dimensional ball of radius r), and K be the Lorentz cone L^{n+1}. As L, choose the hyperplane $e^T_{n+1}x = r$ and π as the orthogonal projection into the first n coordinates.

[1] Recall that a convex body in $\mathbb{R}^n$ is just a bounded, full-dimensional set that is closed and convex.

It remains to define a more general concept for the slack matrix. Before we can do that, let us briefly recall the concept of *polarity*[2].

Definition 3. *Let $C \subset \mathbb{R}^n$ be a convex set. The polar C° of C is defined as*

$$C^\circ := \{y \in \mathbb{R}^n : \langle y, x\rangle \leq 1 \ \forall x \in C\}.$$

Let us capture some of the properties of the polar.

Lemma 4. *Let C be a convex body containing the origin in its interior. Then:*

1. *C° is a convex body containing the origin in its interior.*
2. *$C = (C^\circ)^\circ$.*

If $C = P$ for some polytope P, given by $Ax \leq b$, it is easy to check that we can define C° by looking at the vertices of P only:

$$C^\circ = \{y : \langle x, y\rangle \leq 1 \ \forall \text{ vertices } x \in P\}.$$

Definition 5. *Let C be a convex body and write $\operatorname{ext}(C)$ for the extreme points of C. The* slack operator *of C is defined as*

$$\begin{aligned} S_C : \operatorname{ext}(C) \times \operatorname{ext}(C^\circ) &\longrightarrow [0, \infty), \\ (x, y) &\longmapsto 1 - \langle x, y\rangle . \end{aligned}$$

For a closed, convex cone K, let $K^\star$ denote its dual cone. We say that S_C is K-factorizable, if there exist two maps

$$\begin{aligned} A &: \operatorname{ext}(C) \to K, \\ B &: \operatorname{ext}(C^\circ) \to K^\star, \end{aligned}$$

such that for all extreme points $x \in \operatorname{ext}(C)$, $y \in \operatorname{ext}(C^\circ)$, we have $S_C(x, y) = \langle A(x), B(y)\rangle$.

Here, the analogy to Linear Programming again becomes clear. If C is a polytope, S_C is simply the (transposed) slack matrix.

3 A generalized Yannakakis' theorem

Theorem 6. *Let C be a convex body containing the origin in its interior and K a full-dimensional closed convex cone. Then the following implications hold.*

$$\begin{aligned} &C \text{ has a proper } K\text{-Lift} \Rightarrow S_C \text{ is } K\text{-factorizable} \\ &C \text{ has a } \qquad\quad K\text{-Lift} \Leftarrow S_C \text{ is } K\text{-factorizable.} \end{aligned}$$

Proof

(⇒) Let C be a convex body, K a closed, full-dimensional convex cone, and (π, L) its proper K-Lift, i.e.

$$C = \pi(L \cap K) = \{x \in \mathbb{R}^n : x = \pi(y), \ y \in K, \ y \in L\}.$$

[2] The polar of a convex body is sometimes also called the *dual* of a convex body. This terminology stands to reason as well, but to avoid confusion with the LP duality, we will avoid it.

As L is an affine subspace, we can rewrite it as the Minkowski sum of a point and a linear space, i.e. $L = w_0 + L_0$, where L_0 is the lineality space of L and $w_0 \in L$ is an interior point of K (since the lift is proper).

In order to show that S_C is K-factorizable, we only need to define the maps A and B, and show $S_C(x, y) = \langle A(x), B(y)\rangle$. We will start with A. For $x \in \text{ext}(C)$, define $A(x) := w \in L \cap K$, where w is any preimage of x. To be precise, it is chosen in $\pi^{-1}(x) \cap L \cap K$.

Let us define map B as well. Now, let $c \in \text{ext}(C^\circ)$. We know $\max\{\langle c, x\rangle, \ x \in C\} = 1$, since $\langle c, x\rangle \le 1$ by the definition of the polar, and equality holds as c is an extreme point[3]. Let M be a full row rank matrix with $\ker(M) = L_0$. Then, the following hold[4].

$$
\begin{aligned}
1 &= \max_{x \in C} \langle c, x\rangle \\
&= \max_{w \in K \cap (w_0 + L_0)} \langle c, \pi(w)\rangle \\
&= \max_{w \in K \cap (w_0 + L_0)} \langle c, Aw\rangle && \text{(since } \pi \text{ is a linear map)} \\
&= \max_{w \in K \cap (w_0 + L_0)} \left\langle A^T c, w\right\rangle \\
&= \max\{\langle \pi^\star(c), w\rangle : \ w \in K \cap (w_0 + L_0), \ Mw = Mw_0\} && \text{(setting } \pi^\star(c) := A^T(c)) \\
&= \min\{\langle Mw_0, y\rangle : \ M^T y - \pi^\star(c) \in K^\star\} && (1) \\
&= \min\{\langle w_0, z\rangle : \ z - \pi^\star(c) \in K^\star, \ z \in L_0^\perp\} && \text{(setting z } = M^T y) \quad (2)
\end{aligned}
$$

Now we are able to define, for every $c \in \text{ext}(C^\circ)$, $B(c) = \bar{z} - \pi^\star(c) \in K^\star$, where $\bar{z}$ is an optimum solution to the dual problem (2).

It remains to show $S_C(x, c) = \langle A(x), B(c)\rangle$. For any $x \in \text{ext}(C)$, $c \in \text{ext}(C^\circ)$, we find

$$\langle x, c\rangle = \langle \pi(w), c\rangle = \langle w, \pi^\star(c)\rangle = \langle w, \bar{z} - B(c)\rangle = \langle w, \bar{z}\rangle - \langle w, B(c)\rangle = 1 - \langle A(x), B(c)\rangle\,,$$

where $\langle w, \bar{z}\rangle = 1$ holds as $\bar{z}$ was chosen as an optimum solution to (2) and strong duality holds.

($\Leftarrow$) Given A, B that K-factorize S_C, we have to define a linear map π and an affine subspace L s.t. $C = \{x = \pi(y) : \ y \in K \cap L\}$. Consider

$$\bar{L} := \{(x, z) \in \mathbb{R}^n \times \mathbb{R}^m : \ \langle x, y\rangle = 1 - \langle z, B(y)\rangle \ \text{ for all } y \in \text{ext}(C^\circ)\}.$$

Clearly, $\bar{L}$ is an affine subspace. We claim that (π, L) is a K-lift of C, where $L := \text{proj}_z(\bar{L})$ and π is an appropriate map. To prove this, we will show the following.

1. L is an affine subspace.
   This is obtained by observing that it is a projection of an affine subspace.
2. $(x, z) \in \bar{L}$ for some $z \in K$ implies $x \in C$.
   Let x be some point such that there exists some $z \in K$ with $(x, z) \in \bar{L}$. Then, for all extreme points $y \in \text{ext}(C^\circ)$, we have $\langle x, y\rangle = 1 - \langle z, B(y)\rangle$ by definition of $\bar{L}$. As $z \in K$ and $B(y) \in K^\star$, $\langle z, B(y)\rangle \ge 0$ by definition of the dual cone, implying $\langle x, y\rangle \le 1$ for all $y \in C^\circ$ by convexity. By definition of the polar, this implies $x \in (C^\circ)^\circ = C$.
3. $\pi(z) = x : (x, z) \in \bar{L}$ is well-defined for $z \in K$.
   Existence of some x for any z is given by the definition. Hence, we have to show uniqueness. Take two points $(x_1, z), (x_2, z) \in \bar{L}$. This implies that the affine hull of these two points has to be contained in $\bar{L}$. Further, the affine hull of x_1, x_2 has to be contained in C, using the previous fact. As C is a convex body (in particular bounded), we conclude $x_1 = x_2$.

[3]If the maximum were strictly smaller, then δc is contained in C for some $\delta > 1$, thus c is a convex combination of δc and 0, contradicting the fact that c is an extreme point of C.

[4]The fact that $w_0 \in \text{int}\, K$ implies that Slater condition holds (see the previous lecture) and allows us to use strong duality in (1).

4. π can be extended to an linear map $\pi : \mathbb{R}^m \to \mathbb{R}^n$.
   This is left to the reader.
5. $\pi(K \cap L) \supseteq C$.
   Take $x \in \text{ext}(C)$. We have $(x, A(x)) \in \bar{L}$, hence $A(x) \in L$ by definition. Since A maps to K, we have $z = A(x) \in K$ as well. Observing $\pi(A(x)) = x$ finishes the proof.

■

This theorem is a good step in the right direction, but the properness still bothers us. We need it in the proof to be able to use strong duality, and it is not clear how to work around it. We now restrict to class of cones where a non-proper lift is another lift of the same class.

Definition 7. *let $\mathfrak{K} = \{K_i\}_{i\in\mathbb{N}}$ be a family of closed convex cones. $\mathfrak{K}$ is called* closed *if, for any $i \in \mathbb{N}$, the following holds. For any face F of K_i, there exists some index $j(F) \leq i$ such that F is isomorphic to $K_{j(F)}$.*

For a convex body C, we define the $\mathfrak{K}$-rank of C as

$$\text{rk}_{\mathfrak{K}}(C) = \min\{i \in \mathbb{N} : \; S_C \textit{ has a } K_i\textit{-factorization.}\}$$

A nice example of a closed family $\mathfrak{K}$ is the family of non-negative orthants, $\{\mathbb{R}^n_{\geq 0}\}_{n\in\mathbb{N}}$.

Theorem 8. *Let $\mathfrak{K}$ be a closed family of closed convex cones, and C a convex body. Then*

$$\text{rk}_{\mathfrak{K}}(C) = \min\{i \in \mathbb{N} : C \textit{ has a } K_i\textit{-Lift.}\}$$

Proof Through the kindness of Theorem 6, there is nothing to show for " $\geq$ ", as we do not need properness. The other direction does not request much effort as well. Take the minimum i s.t. C has a K_i-Lift. If the Lift is not proper, there is some $j < i$ for which C has a K_j-Lift. But this contradicts the minimality of i. Hence, the Lift is proper and we may apply the theorem. ■

4 Positive semidefinite Cone and Positive semidefinite rank

Consider the cone of positive semidefinite matrices X,

$$S^n_+ := \{X \in \mathbb{R}^{n\times n} : \; X \text{ is symmetric, and } a^T X a \geq 0 \; \forall a \in \mathbb{R}^n\}. \tag{3}$$

We write $X \succeq 0$ for a positive semidefinite matrix. The family of all symmetric matrices is a Hilbert space with the inner product $\langle X, Y\rangle = \text{tr}(X^T Y) = \sum_{i,j} X_{i,j} Y_{i,j}$. This is also referred to as Frobenius Product. Note that this is the product we obtain when considering the matrices as vectors in the Euclidean space $\mathbb{R}^{n^2}$. Let us recall some equivalent definitions of positive semidefiniteness.

Lemma 9. *Let $X \in \mathbb{R}^{n\times n}$ be a symmetric matrix. Then the following properties of X are equivalent.*

- *X is positive semidefinite.*
- *All eigenvalues of X are non-negative.*
- *X has a Cholesky Decomposition, i.e. $X = U^T U$ for some $U \in \mathbb{R}^{d\times n}$.*
- *All principle submatrices of X have non-negative determinants.*

Lemma 10. *S^n_+ as defined above is a closed convex cone.*

Proof For all $X, X' \in S_+^n$ $X + X'$ is again symmetric and positive semidefinite, as $a^T(X+X')a = a^T Xa + a^T X' a$. If X is symmetric and positive semidefinite, λX is as well, for all $\lambda \geq 0$. In order to show that S_+^n, we show that its complement is open. If $X \notin S_+^n$ there exists some a with $a^T X a < 0$. As this inequality is strict, you can define some open ε-ball around X not intersecting S_+^n. ■

On the next exercise sheet, we will prove the following Lemma.

Lemma 11. *$\mathfrak{K} = \{S_+^n\}_{n\in\mathbb{N}}$ is a closed family.*

We can define the notion of *positive semidefinite rank* for a polytope P to be the value $\mathrm{rk}_{\mathfrak{K}}(P)$ when $\mathfrak{K} = \{S_+^n\}_{n\in\mathbb{N}}$. We will denote it by $\mathrm{rk}_{sdp}(P)$.

Lemma 12. *The positive semidefinite cone is self-dual, that is, $(S_+^n)^\star = S_+^n$.*

Proof

Let $X, X' \in S_+^n$. Since X is symmetric, we can diagonalize it with an orthogonal matrix; that is, $X = SDS^T$, where D is a diagonal matrix having the eigenvalues λ_i of X on the diagonal. Then:

$$\begin{aligned} \langle X, X' \rangle &= Tr(X^T X') = Tr(SDS^T X') \\ &= Tr(DS^T X' S) \qquad \text{(Since } Tr(A^T B) = Tr(AB^T)\text{)} \\ &= \left\langle D, S^T X' S \right\rangle \\ &= \textstyle\sum_i \lambda_i S_i^T X' S_i \quad \geq 0, \end{aligned}$$

Where we in the last equations we used that D_{ij} is a diagonal matrix, and in the inequality that $\lambda_i \geq 0$ since $X \in S_+^n$, and that $X' \in S_+^n$. This shows $(S_+^n)^\star \supseteq S_+^n$.

Let now $X' \in \mathbb{R}^{n\times n}$, symmetric, so that $\langle X, X' \rangle \geq 0$ for all $X \in S_+^n$. Pick $v \in R^n$. Then $vv^T \in S_+^n$, hence

$$0 \leq \left\langle vv^T, X' \right\rangle = Tr(vv^T X') = v^T X' v, \text{ thus } X' \in S_+^n.$$

This shows $(S_+^n)^\star \subseteq S_+^n$. ■

4.1 Casting a Linear Program as a Semidefinite Program

We want to reformulate the following LP problem

$$\begin{aligned} \max \langle c, x \rangle& \\ Ax &= b \\ x &\geq 0 \end{aligned}$$

as a *Semidefinite Programming problem*, i.e. a linear optimization problem over the intersection of the semidifinite cone with an affine subspace. Define the cost matrix $C \in \mathbb{R}^{n\times n}$ by $C_{ii} = c_i$ and all other entries being 0, and matrices A^ℓ with $\mathbf{A}_{ii}^\ell = A_{\ell i}$, and all other entries being 0. The original linear programming problem is equivalent to the following.

$$\begin{aligned} \max \langle C, X \rangle& \\ \left\langle \mathbf{A^l}, X \right\rangle &= b_l && \forall\ l \text{ within range} \\ X_{ij} &= 0 && \forall\ i \neq j \\ X &\succeq 0 \end{aligned}$$

5 Notes

Definitions, results and proofs from Sections 2 and 3 appeared in:

J. Gouveia, P. A. Parrilo, and R. R. Thomas. Lifts of Convex Sets and Cone Factorizations. Math. Oper. Res. 38(2), pp. 248–264, 2013.