CSC2429: Proof Complexity **Winter 2018**

Lecture 2: Proof System and Algorithms for LP

Lecturer: Toni Pitassi *Scribe: Lily Li*

2.1 Sherali Adams

2.1.1 Review LP

Consider this as a sound and complete proof system for deriving or refuting a set of linear inequalities over $\mathbb{R}^n$. As always an LP is defined as

$$\max \mathbf{c}^T x \qquad \text{over } \mathbb{R}$$
$$\text{Subject to: } A\mathbf{x} \leq \mathbf{b}$$

The decision version of this problem asks if the feasible region of the LP is empty.

Lemma 2.1 *(Farkas' Lemma.) A set $\{A\mathbf{x} \leq \mathbf{b}\}$ is unsatisfiable over $\mathbb{R}$ if and only if there exists $\mathbf{y} \geq 0$ such that $\mathbf{y}^T A = 0$ and $\mathbf{y}^T \mathbf{b} = -1$.*

Farkas' Lemma should be viewed as a soundness and completeness proof of the decision version and the y^T can be seen as a linear combination of the rows of of A which add to -1.

2.1.2 Duality

(Considered a form of implicational completeness). We want to show that each feasible solution of the dual produces an upper bound of the primal objective function. Consider any $\mathbf{y}^T \geq 0$ such that $\mathbf{y}^T A \geq \mathbf{c}^T$. Then for any feasible solution $\mathbf{x}$ of the primal $A\mathbf{x} \leq \mathbf{b}$,

$$\mathbf{c}^T \leq \mathbf{y}^T A\mathbf{x} \leq \mathbf{y}^T \mathbf{b}$$

where $\mathbf{c}^T\mathbf{x}$ is exactly the primal objective function. Thus we say that $\mathbf{y}$ *witnesses* the upper bound $\mathbf{b}^T\mathbf{y}$.

To see how tight this bound is, we consider following primal and dual LP problems:

(P) Primal:	(D) Dual:
$\max \mathbf{c}^T\mathbf{x}$	$\min \mathbf{b}^T\mathbf{y}$
Such that: $A\mathbf{x} \leq \mathbf{b}$	$A^T\mathbf{y} \geq \mathbf{c}$
$\mathbf{x} \geq 0$	$\mathbf{y} \geq 0$

Theorem 2.2 *(Duality of LP.)* This is implied by Farkas' Lemma. *For an LP with primal solution P and dual solution D, exactly of the following is true:*

1. *Neither (P) nor (D) has a feasible solution.*

2. *(P) has arbitrarily large solutions and (D) is unsatisfiable.*

3. *(P) is unsatisfiable and (D) has an arbitrarily small solution.*

4. *Both (P) and (D) have optimal solutions* $\mathbf{x}^*$ *and* $\mathbf{y}^*$. *Further* $\mathbf{c}^T\mathbf{x}^* = \mathbf{b}^T\mathbf{y}^*$

In the decision version an LP refutation is just the $\mathbf{y}$ from Farkas' Lemma.

An LP "derivation" of

$$\{A\mathbf{x} \leq \mathbf{b}, \mathbf{x} \geq 0\} \implies \mathbf{c}^T\mathbf{x} \leq \mathbf{c_0}$$

is a $\mathbf{y}^* \geq 0$ such that $(\mathbf{y}^*)^T\mathbf{b} = \mathbf{c_0}$.

It is known that LP is in P. The satisfiability of an LP is in NP since we can just guess a $\mathbf{x}$. Further Farkas' lemma shows that the problem is in coNP since we can just guess a $\mathbf{y}$ (there are problems which are in $\mathsf{P} \cap \mathsf{NP}$ but not known to be in P, e.g. factoring and graph isomorphism, but these are uncommon). In 1979 Khachiyan use the Ellipsoid Algorithm to show that $LP \in \mathsf{P}$.

2.1.3 SA Introduction

SA is a sound, lift and project proof system that "tightens" the LP solution to a good approximation of the IP. SA can be interpretations: as (1) LP tightening, solutions to the extended degree d SA LP pseudo-distribution, and as a proof or refutation system.

2.1.4 SA Degree d Tightening

When viewed as LP tightening in degree d, we do the following:

- Add new variables to represent all degree $\leq d$ terms.
- *Lift* the polytope from n dimensions to $n^{o(d)}$ dimensions.
- Project back to $x_1, ..., x_n$ preserves all $0, 1$ solutions (and may remove some fractional solutions).

More formally, let the original LP be

$$A\mathbf{x} \geq \mathbf{b}, \text{ and let } \mathbf{1} \geq \mathbf{0}, \mathbf{0} \leq \mathbf{x} \leq \mathbf{1}.$$

We add new variables:

$$y_S : \qquad \forall S \subseteq [n], |S| \leq d.$$

And new constraints (junta) for all subsets $U, V \subset [n]$ with $U \cap \emptyset$ and $|U \cup V| \leq d$:

$$\prod_{i \in U} x_i \cdot \prod_{j \in V} (1 - x_j) \cdot (\mathbf{a}^T\mathbf{x} - \mathbf{b}) \geq 0$$

for all rows $\mathbf{a}$ of A, for all U, V.

When we translate the above constrains into linear equalities using the new y_S variables and the multi-linear inequalities $x_i^2 = x_i$ we obtain:

$$\begin{aligned} y_\emptyset &= 1 \\ y_i &= x_i \\ 0 \leq y_S &\leq 1 \\ \sum_{V' \subseteq V} (-1)^{|V'|} \cdot \left(\sum_{i=1}^{n} a_i y_{U \cup V' \cup \{i\}} - \mathbf{b} y_{U \cup V'} \right) &\geq 0 \end{aligned}$$

for all rows $\mathbf{a}$ of A.

Further, since we have trivially $\mathbf{1} \geq \mathbf{0}$ as an initial constraint, we can add

$$\prod_{i \in U} x_i \prod_{j \in V} (1 - x_j) \geq 0$$

which translates to

$$\sum_{V' \subseteq V} (-1)^{|V'|} \cdot (y_{U \cup V'}) \geq 0$$

Now we need to defined a pseudo-distribution for a degree d SA LP (this is a real distribution over sets variables of degree at most d and might not work else-where). Let $\mathcal{H} = \{A\mathbf{x} - \mathbf{b} \geq \mathbf{0}, \mathbf{x} \geq \mathbf{0}, \mathbf{1} \geq \mathbf{0}\}$. $\mathcal{E}_d(\mathcal{H})$ is a set of linear functionals $E : \mathbb{R}[x_1, ..., x_n]_d \to \mathbb{R}$ such that $\forall E \in \mathcal{E}_d(\mathcal{H})$:

1. $E(1) = 1$
2. $E(Q) \geq 0$ for all non-negative junta Q with $\deg(Q) \leq d$.
3. $E(PQ) \geq 0$ for $P \in \mathcal{H}$, and non-negative junta Q with $\deg(PQ) \leq d$ where a non-negative junta is of the form:

   $$\alpha \cdot \prod_{i \in U} x_i \prod_{j \in V} (1 - x_j)$$

   for non-negative coefficient α and $U \cup V = \emptyset$. *This is quite useful to show that what you have is actually a a probability distribution (recall the marginals).*

Essentially you should interpret each $E \in \mathcal{E}_d(\mathcal{H})$ as a *pseudo-distribution* for the feasible (maybe? my words here) region of $\mathcal{H}$. Further, feasible solutions to the degree-d SA polytope are exactly the degree-d *pseudo-distributions.*

Consider any feasible solution α. For each variable y_S,$\alpha(y_S)$ is the value in $[0, 1]$ assigned to y_S such that all the linear constraints are satisfied.

The corresponding pseudo-distribution E_α which assigns values to degree-d polynomials is defined in the obvious way.

Example 2.3 *Suppose we have $d = 3$ and the polynomial $f = -x_1x_2x_4 + x_7 - 3x_1x_8$, then*

$$E_\alpha[f] = -\alpha(y_{1,2,4}) + \alpha(y_7) - 3\alpha(y_{1,8}).$$

It so happens that the degree-d SA linear constraints exactly enforce the properties of the pseudo-distribution (what E_α does is assign a $[0, 1]$ value to every set of $\leq d$ of the original variables such that the marginal distributions of each subset $S' \subset S$ is equal to the distribution assigned to each entry of S').

2.1.5 Equivalent view as a Refutation System

Let $\mathcal{F}$ be a set of polynomial equalities of the form $x_i^2 - x_i = 0$ (this set is actually quite important, why?). Let $\mathcal{H}$ be a set of polynomial inequalities $\{A\mathbf{x} - \mathbf{b} \geq 0,\ 1 \geq \mathbf{x} \geq 0$, and $\mathbf{1} \geq \mathbf{0}\}$ our original inequalities. A degree-d SA derivation of -1 from $\mathcal{F}$ and $\mathcal{H}$ — witnessing that no feasible solutions exists — is a set of polynomials $(g_1, ..., g_m, p_1, ..., p_l)$ such that

1. $\sum_{j=1}^{m} g_j f_j + \sum_{i=1}^{l} p_i h_i = -1$
2. The p_i's are non-negative linear combinations of juntas $p_i = \alpha_1 Q_1 + \cdots + \alpha_m Q_m$ (the Q's are the juntas), $h_i \in \mathcal{H}$, $f_j \in \mathcal{F}$, and g_j are arbitrary polynomials — since the f_j's evaluate to 0.
3. $\max\{\deg(g_j f_j), \deg(p_i h_i)\} = d$.

2.1.6 SA as a Derivation System

SA as a derivation system is almost exactly like SA as a refutation system except that the sum of the product of the polynomials $g_j f_j$ and $p_i h_i$ is some other polynomial f which we wish to derive.

Lemma 2.4 *Let $\{A\mathbf{x} - b \geq 0, 1 \geq 0, \mathbf{x} \geq 0\} = \mathcal{H}$. Then the degree-d SA LP has no feasible solution if and only if there is a degree-d SA refutation of $\mathcal{H}$.*

Proof: Let the SA refutation be a sett of polynomials as above. The SA LP has a feasible solution $\mathbf{x}^*$, then we can plug in $\mathbf{x}^*$ into the set of polynomials and evaluate. Since $\mathbf{x}^*$ is feasible the LHS of each inequality will be positive so it is impossible to obtain a sum equal to -1.

Conversely if the SA LP has no feasible solution we can translate it into a refutation of $\mathcal{H}$ as defined in the "SA as a Refutation System" section. ■

Example 2.5 *We are going to apply SA on the maximum independent set problem with the graph $G = C_7$. If you just formulate the standard LP as:*

$$\begin{aligned} &\textit{Maximize: } \sum_{i=1}^{7} x_i \\ &\textit{Subject to: } x_i + x_{i+1} \leq 1 \\ &\qquad\qquad \textit{and } 0 \leq x_i \leq 1 \textit{ for } 1 \leq i \leq 7 \end{aligned}$$

you would get a fractional optimum value of 3.5*. Lifting to $d = 2$ introduces the following constraints:*

$$
\begin{aligned}
x_1(x_1+x_2) \le x_1 &\implies y_{1,2} \le 0 \\
(1-x_1)(x_2+x_3) \le 1-x_1 &\implies y_2+y_3-y_{1,2}-y_{1,3} \le 1-y_1 \\
x_1(x_3+x_4) \le x_1 &\implies y_{1,3}+y_{1,4} \le y_1 \\
(1-x_1)(x_4+x_5) \le 1-x_1 &\implies y_4+y_5-y_{1,4}-y_{1,5} \le 1-y_1 \\
x_1(x_5+x_6) \le x_1 &\implies y_{1,5}+y_{1,6} \le y_1 \\
(1-x_1)(x_6+x_7) \le 1-x_1 &\implies y_6+y_7-y_{1,6}-y_{1,7} \le 1-y_1 \\
x_1(x_1+x_7) \le x_1 &\implies y_{1,7} \le 0
\end{aligned}
$$

(as well as other constraints, but these are all the ones we need to consider). Running the LP here would produce the answer 3 *(which is optimal).*