CS 270 Algorithms **Spring 2021**

Lecture 21: Duality, Sum-of-squares proofs

Lecturer: Prasad Raghavendra Scribe: Kevin Ye, Christina Jin

21.1 LP Duality and Proofs

Say we have a linear program with the contraints

$$\langle a_i, x\rangle \leq b_i \quad \text{for } i = 1, \ldots, m. \tag{21.1}$$

The feasible region defined by these constraints will be some polytope $\mathcal{P}$. Given some $y \in \mathcal{P}$, we know that y satisfies

$$\langle a_i, y\rangle \leq b_i \quad \forall i = 1, \ldots, m. \tag{21.2}$$

These inequalities are true by construction, so we may take them as our "axioms".

Now suppose we want to prove some statement about points in $\mathcal{P}$. For example, suppose it is true that $\langle c, y\rangle \leq d$ is true for all $y \in \mathcal{P}$. To prove this from our axioms, we must generate new true statements by taking nonnegative linear combinations of our axioms. For example we can have

$$2(\langle a_1, y\rangle \leq b_1) + 3(\langle a_2, y\rangle \leq b_2) \implies \langle 2a_1 + 3a_2, y\rangle \leq 2b_1 + 3b_2. \tag{21.3}$$

In general, for nonnegative $c_1, \ldots, c_m$ we have

$$\begin{gathered} c_1(\langle a_1, y\rangle \leq b_1) \\ c_2(\langle a_2, y\rangle \leq b_2) \\ \vdots \\ c_m(\langle a_m, y\rangle \leq b_m) \\ \hline \langle \sum_{i=1}^{m} c_i a_i, y\rangle \leq \sum_{i=1}^{m} c_i b_i. \end{gathered}$$

Note that the c_i are the variables of the Dual LP.

21.1.1 Soundness and Completeness

We will take a moment to consider two important properties of proof systems.

Definition 21.1 (Soundness)**.** *A proof system is sound if every statement that can be proved in the system is actually true.*

Definition 21.2 (Completeness)**.** *A proof system is complete if it can prove every statement that is true.*

Ideally, we would want a proof system to be able to prove statements if and only if they are true. Soundness and completeness together give us this biconditional.

Going back to our LP proof system, suppose we want to prove a bound on the value of $\langle d, y\rangle$. We may construct the following LP:

$$\begin{aligned} \max \quad & \langle d, y\rangle \\ \text{s.t.:} \quad & \langle a_i, y\rangle \leq b_i \quad \forall i = 1, \ldots, m \end{aligned}$$

Now consider the dual of this LP

$$\begin{aligned} \min \quad & \sum_{i=1}^{m} c_i b_i \\ \text{s.t.:} \quad & \sum_{i=1}^{m} c_i a_i = d \\ & c_i \geq 0 \quad \forall i = 1, \ldots, m \end{aligned}$$

The dual LP generates proofs of bounds for $\langle d, y\rangle$. By weak duality, each of these bounds must be valid. Hence our LP proof system is sound. By strong duality, the optimal solution to the dual LP is in fact the optimal bound on $\langle d, y\rangle$, meaning that our proof system is complete.

21.2 Polynomial Proofs

21.2.1 Equalities

Now suppose we have a system of polynomial equalities $\mathcal{P} = \{p_i(x) = 0 \mid i = 1, \ldots, m\}$. Given $\mathcal{P}$ suppose it is true that $q(x) = 0$ also, and we want a proof of this fact. Note that proofs are sometimes also called certificates, because they provide a fast way (polynomial time) of verifying that a statement is true. For a system of polynomial equalities, the existence of certificates is given by the following theorem.

Theorem 21.3 (Hilbert's Nullstellensatz)**.** *Given $\mathcal{P} = \{p_i(x) = 0 \mid i = 1, \ldots, m\}$, if $\mathcal{P} \implies (q(x) = 0)$, then there exists an integer $c \geq 1$ and polynomials $r_1(x), \ldots, r_m(x)$ such that*

$$q(x)^c = \sum_{i=1}^{m} r_i(x) p_i(x) \tag{21.4}$$

21.2.2 Inequalities

Now suppose we have polynomial inequalities $\mathcal{P} = \{p_i(x) \geq 0 \mid i = 1, \ldots, m\}$. Given $\mathcal{P}$, suppose that it is true that $q(x) \geq 0$ also, and we want a proof of this fact. $q(x) \geq 0$ would be true if $q(x)$ was a sum of the squares of polynomials, i.e.

$$q(x) = \sum_{i=1}^{N} s_i(x)^2 \geq 0. \tag{21.5}$$

We'll call a polynomial that is the sum of squares of polynomials an SoS polynomial. In particular, multiplying any of $p_i(x)$ by a sum of squares polynomial also gives us a nonnegative polynomial, hence if we can write $q(x)$ as a sum of an SoS polynomial with p_i's multiplied by SoS polynomials, that will prove $q(x) \geq 0$.

Definition 21.4 (SoS Proof)**.** *Given $\mathcal{P} = \{p_i(x) \geq 0 \mid i = 1, \ldots, m\}$, an SoS proof that $\mathcal{P} \implies (q(x) \geq 0)$ is*

$$q(x) = s_0(x) + \sum_{i=1}^{m} p_i(x) s_i(x) \tag{21.6}$$

where $s_0, s_1, \ldots, s_m$ are SoS polynomials. We'll call

$$\max\{\deg(s_0), \deg(p_1 s_1), \ldots, \deg(p_m s_m)\} \tag{21.7}$$

the degree of the proof.

The following theorem guarantees the existence of SoS proofs, but with a slightly altered form.

Theorem 21.5 (Positivstellensatz)**.** *Given $\mathcal{P} = \{p_i(x) \geq 0 \mid i = 1, \ldots, m\}$, if $\mathcal{P} \implies (q(x) \geq 0)$, then there exists an SoS proof of the form*

$$q(x)(1 + r(x)) = s_0(x) + \sum_{i=1}^{m} p_i(x) s_i(x) \tag{21.8}$$

where r and $s_0, \ldots, s_m$ are SoS polynomials.

21.3 Max-Cut SDP

Now we will see how the dual objects are related to SoS proofs. We will first start with Max-Cut SDP to see how to write duals for a simple SDP.

21.3.1 Constructing the dual SDP

Recall Max-Cut SDP can be formulated as follows:

$$\begin{aligned} \max \quad & \langle L, X \rangle = \sum_{i,j} L_{ij} X_{jj} \\ \text{s.t.:} \quad & X_{ii} \leq 1, \quad \forall i = 1, \ldots, n \\ & X \succeq 0 \quad \text{(X is P.S.D.)} \end{aligned}$$

Recall the definition of PSD matrices:
A matrix X is PSD if and only if either of the following is true:

- $\forall v \in \mathbb{R}^n, v^T X v = \langle v v^T, X \rangle \geq 0$.
- $X = \sum_i c_i u_i u_i^T$, for $c_i \geq 0$, $u_i \in \mathbb{R}^n$.

Applying the first definition of PSD matrix, we could rewrite the program as follows:

$$\begin{aligned} \max \quad & \langle L, X\rangle = \sum_{i,j} L_{ij} X_{jj} \\ \text{s.t.:} \quad & X_{ii} \le 1, \quad \forall i = 1, \dots, n && (\alpha_1, \dots, \alpha_n) \\ & \langle vv^T, X\rangle \ge 0, \quad \forall v \in \mathbb{R}^n && (\beta_v) \end{aligned}$$

We define α_i's and β_v as the "dual variables" (one for each constraint). The dual of this program tries to prove an upperbound B for the primal optimal value, i.e.: $B - \langle L, X\rangle \ge 0$, by taking a non-negative linear combination of the axioms (the inequality constraints). Since the non-negative linear combination of non-negative terms are non-negative, ensuring the following identity would suffice to show that $B - \langle L, X\rangle \ge 0$:

$$\begin{aligned} B - \langle L, X\rangle &= \sum_{i=1}^{n} \alpha_i \underbrace{(1 - X_{ii})}_{\ge 0} + \sum_v \beta_v \underbrace{\langle vv^T, X\rangle}_{\ge 0} \\ &= \sum_{i=1}^{n} \alpha_i (1 - X_{ii}) + \left\langle \sum_v \beta_v vv^T, X \right\rangle \\ &= \sum_{i=1}^{n} \alpha_i (1 - X_{ii}) + \langle Y, X\rangle \end{aligned}$$

In the last step we define $Y = \sum_v \beta_v vv^T$. By the second definition of PSD matrices above, $Y \succeq 0$. Now we can write the dual SDP as follows:

$$\begin{aligned} & \text{Find } \alpha_1, \dots \alpha_n \text{ and } Y \\ \text{s.t.:} \quad & B - \langle L, X\rangle = \sum_{i=1}^{n} \alpha_i (1 - X_{ii}) + \langle Y, X\rangle \\ & \alpha_1, \dots \alpha_n \ge 0 \\ & Y \succeq 0 \end{aligned}$$

21.3.2 Converting dual SDP to polynomials

Recall: X is the moment matrix of the solutions to Max-Cut. Also note the correspondence between the dual SDP and the polynomials:

$$\begin{aligned} \langle L, X\rangle &= \tilde{\mathbb{E}} \left[\sum_{(i,j)\in E} (x_i - x_j)^2 \right] \\ X_{ii} &= \tilde{\mathbb{E}}[x_i^2] \\ \langle vv^T, X\rangle &= \sum_{i,j} v_i v_j X_{ij} \\ &= \sum_{i,j} v_i v_j \tilde{\mathbb{E}}[x_i x_j] \\ &= \tilde{\mathbb{E}} \left[\left(\sum_i v_i x_i \right)^2 \right] \end{aligned}$$

As a result, we could convert the dual identity to the polynomial space as follows:

$$B - \langle L, X\rangle = \sum_{i=1}^{n} \alpha_i (1 - X_{ii}) + \sum_v \beta_v \langle vv^T, X\rangle$$

$$\Downarrow$$

$$B - \sum_{(i,j)\in E} (x_i - x_j)^2 = \sum_{i=1}^{n} \alpha_i (1 - x_i^2) + \sum_v \beta_v (\sum_i v_i x_i)^2$$

This is a degree-2 SoS proof. Specifically, by showing each of the term on the right is ≥ 0, we could show that $B - \sum_{(i,j)\in E}(x_i - x_j)^2 \geq 0$.

21.4 Pseudo-expectation and Weak Duality

Definition 21.6 (Pseudo-expectation)**.** *A degree-d pseudo-expectation $\tilde{\mathbb{E}}$ is a linear functional that maps a degree-d polynomial $\to \mathbb{R}$.*

Definition 21.7 (Pseudo-expectation for a polynomial system)**.** *For a polynomial system $\{p_i(x) \geq 0 | i = 1, \ldots, m\}$ and polynomial $s(x)$, the following statements must be true:*

- $\tilde{\mathbb{E}}[p_i(x)s^2(x)] \geq 0, \quad \forall i = 1, \ldots, m$
- $\tilde{\mathbb{E}}[s^2(x)] \geq 0$
- $\tilde{\mathbb{E}}[1] = 1$

Definition 21.8 (Weak Duality for SoS proof)**.** *If $\mathcal{P} \implies (q(x) \geq 0)$ admits a degree-d SoS proof s.t.: $q(x) = s_0(x) + \sum_i p_i(x)s_i(x)$, then for every degree-d pseudo-expectation $\tilde{\mathbb{E}}$, $\tilde{\mathbb{E}}[q(x)] \geq 0$.*

In other words, if we prove a fact using degree-d SoS proof, the fact will also be true for the SDP solution.

In the next lecture, we will look at algorithms based on SDP. To give an idea, let's take the problem of robust linear regression. We will take the following steps:

1. Write a polynomial system $\mathcal{P}$ for robust linear regression. This will likely be NP-complete to solve.
2. Instead we solve a degree-d SoS SDP for $\mathcal{P}$, which returns some pseudo-moments $\tilde{\mathbb{E}}[x_i x_j]$.
3. To prove that these moments actually give the solutions to our original problem, we use low-degree SoS proof to show that any solution to $\mathcal{P}$ is close to the true solution. This implies that the pseudo-momemts generated in step 2 are also close to the true solution.