CSE 521: Design and Analysis of Algorithms I **Fall 2020**

Lecture 18: Linear Programming Relaxation, Duality and Applications

Lecturer: Shayan Oveis Gharan *12/08/20*

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications.*

1.1 The LP Duality

Before introducing the concept of the duality, a simple example is firstly shown as below.

$$\begin{aligned} \min \quad & 2x_1 + 3x_2 \\ \text{s.t.,} \quad & x_1 + 2x_2 \geq 1 \\ & 3x_1 + 2x_2 \geq 2 \\ & x_1, x_2 \geq 0 \end{aligned} \tag{1.1}$$

An optimum solution of this LP is attained at $x_1 = 1/2$ and $x_2 = 1/4$. This gives an objective value of 7/4. Now, suppose we want to justify that this solution is indeed an OPT. In other words, we want to prove for an adversary that there is no better solution to this LP. How can we do that? We can try to prove a lower-bound on the optimum solution. Let's start: Since $x_1, x_2 \geq 0$ we can write

$$2x_1 + 3x_2 \geq x_1 + 2x_e \geq 1.$$

How can we prove a better upper-bound? Let's use the second inequality:

$$2x_1 + 3x_2 \geq \frac{2}{3}(3x_1 + 2x_2) \geq \frac{2}{3}(2) = 4/3.$$

Another option is to combine both inequalities:

$$2x_1 + 3x_2 \geq \frac{1}{2}(x_1 + 2x_2) + \frac{1}{2}(3x_1 + 2x_2) \geq \frac{1}{2}(1) + \frac{1}{2}(2) \geq 1.5.$$

We are getting very close to 7/4. Can we find coefficients to obtain 7/ as the lower bound? The optimal way is to multiply the first inequality with 5/4 and the second one with 1/4. This gives

$$2x_1 + 3x_2 \geq \frac{5}{4}(x_1 + 2x_2) + \frac{1}{4}(3x_1 + 2x_2) \geq \frac{5}{4}(1) + \frac{1}{4}(2) = 7/4.$$

It turns out that this was not a coincidence that we could prove a lower bound exactly equal to the optimum value that we have found. As we will see below, whenever an optimum solution to a LP exists we can justify that solution following a similar idea. This gives us the dual of a linear program.

Let's formalize what we have done. We were looking for coefficients y_1 and y_2 for the first and the second constraints such that $y_1 + 2y_2$ is maximized subject to $y_1, y_2 \geq 0$ and that $y_1(x_1 + 2x_2) + y_2(3x_1 + 2x_2) \leq (2x_1 + 3x_2)$. Aligning the objective functions and the the summed term, we obtain a dual problem as (1.2).

$$\begin{aligned} \max \quad & y_1 + 2y_2 \\ \text{s.t.,} \quad & y_1 + 3y_2 \leq 2 \\ & 2y_1 + 2y_2 \leq 3 \\ & y_1, y_2 \geq 0 \end{aligned} \tag{1.2}$$

Then we generalize the LP primal-dual problems as follows:

$$\begin{aligned} \min \quad & \langle \mathbf{c}, \mathbf{x} \rangle \\ \text{s.t.,} \quad & A\mathbf{x} \geq \mathbf{b} \\ & \mathbf{x} \geq 0 \end{aligned} \tag{1.3}$$

where $A \in R^{m \times n}$, $\mathbf{c} \in \mathbb{R}^n$, and $\mathbf{b} \in \mathbb{R}^m$. So we can define the dual of the above LP as (1.4), where $\mathbf{y} \in R^m$.

$$\begin{aligned} \max \quad & \langle \mathbf{b}, \mathbf{y} \rangle \\ \text{s.t.,} \quad & A^T\mathbf{y} \leq \quad \mathbf{c} \\ & \mathbf{y} \geq 0 \end{aligned} \tag{1.4}$$

Then, the weak duality theorem and the strong duality theorem are introduced as follows:

Theorem 1.1 (Weak Duality)**.** *If $\boldsymbol{x}$ is a feasible solution of $P = \min\{\langle \boldsymbol{c}, \boldsymbol{x} \rangle | A\boldsymbol{x} \geq \boldsymbol{b}, \boldsymbol{x} \geq 0\}$ and $\boldsymbol{y}$ is a feasible solution of $D = \max\{\langle \boldsymbol{b}, \boldsymbol{y} \rangle | A^T\boldsymbol{y} \leq \boldsymbol{c}, \boldsymbol{y} \geq 0\}$, then*

$$\langle \boldsymbol{c}, \boldsymbol{x} \rangle \geq \langle \boldsymbol{y}, \boldsymbol{b} \rangle.$$

In other word, any feasible solution of the dual gives a lower bound on the optimum solution of the primal.

Proof. Since $\mathbf{y} \geq 0$ and $A\mathbf{x} \geq \mathbf{b}$, we get

$$\langle \mathbf{b}, \mathbf{y} \rangle \leq \langle A\mathbf{x}, \mathbf{y} \rangle. \tag{1.5}$$

Also, since $A^T\mathbf{y} \leq \mathbf{c}$, we have

$$\langle A\mathbf{x}, \mathbf{y} \rangle = \langle A^T\mathbf{y}, \mathbf{x} \rangle \leq \langle \mathbf{c}, \mathbf{x} \rangle. \tag{1.6}$$

Thus, we can get $\langle y, b \rangle \leq \langle c, x \rangle$ and we are done. □

Theorem 1.2 (strong duality)**.** *For any LP and its dual, one of the following holds:*
1. The primal is infeasible and the dual has unbounded optimum.
2. The dual is infeasible and the primal has unbounded optimum.
3. Both of them are infeasible.
4. Both of them are feasible and their optimum value is equal.

1.2 Applications of the LP Duality

In this section, we discuss one important application of duality. It is the Minimax theorem which proves existence of Mixed Nash equilibrium for two-person zero-sum games and proposes an LP to find it. Before stating this, we need a couple of definitions. A two-person game is defined by four sets (X, Y, A, B) where

1. X and Y are the set of strategies of the first and second player respectively.
2. A and B are real-valued functions defined on $X * Y$.

The game is played as follows. Simultaneously, Player (I) chooses $x \in X$ and Player (II) chooses $y \in Y$, each unaware of the choice of the other. Then their choices are made known and (I) wins $A_{i,j}$ and (II) wins $B_{i,j}$. A and B are called utility function for player (I) and (II), and obviously the goal of both players is to maximize their utility. The game is called a zero-sum game if $A = -B$.

A mixed strategy for a player is just a distribution over his/her strategies. The last thing we need to define is mixed Nash equilibrium.

Definition 1.3 (Pure Nash equivalence). *A pair (i^*, j^*) is pure equivalent if nobody wants to deviate, where $\max_j B_{i^*,j} = B_{i^*,j^*}$ and $\max_i A_{i,j^*} = A_{i^*,j^*}$.*

It is proved by Nash that every n-person game has one Nash equilibrium. In general, finding the Nash equilibrium is a very hard problem. However, in the case of two-player zero-sum games there is a polynomial time algorithm to find it. In particular, let (X, Y, A) represents a two-player zero-sum game. If x and y are two mixed strategies for (I) and (II), then one can see the expected utility of (I) is $\mathbf{x}^T A\mathbf{y}$ and for (II) it is $-x^T Ay$. So the player (I) wants to maximize $\mathbf{x}^T A\mathbf{y}$ and (II) wants to minimize it. Then there are the mixed strategies $\mathbf{x}^*, \mathbf{y}^*$ for (I) and (II) satisfying

$$\max_{\mathbf{x}} \min_{\mathbf{y}} \mathbf{x}^T A\mathbf{y} = \min_{\mathbf{y}} \max_{\mathbf{x}} \mathbf{x}^T A\mathbf{y} = \mathbf{x}^{*T} A\mathbf{y}^*, \tag{1.7}$$

then $(\mathbf{x}^*, \mathbf{y}^*)$ is a mixed Nash equilibrium. The following nice result by Neumann guarantees $(\mathbf{x}^*, \mathbf{y}^*)$ exists and gives an LP such that its optimum solution is $\mathbf{x}^*$ and the optimum solution of its dual is $\mathbf{y}^*$. The proof is an application of the strong duality theorem.

Theorem 1.4 (The Minimax Theorem). *For every two-person zero-sum game (X, Y, A) there is a mixed strategy $\boldsymbol{x}^*$ for player I and a mixed strategy $\boldsymbol{y}^*$ for player (II) such that,*

$$\max_{\boldsymbol{x}} \min_{\boldsymbol{y}} \boldsymbol{x}^T A\boldsymbol{y} = \min_{\boldsymbol{y}} \max_{\boldsymbol{x}} \boldsymbol{x}^T A\boldsymbol{y} = \boldsymbol{x}^{*T} A\boldsymbol{y}^*, \tag{1.8}$$

where in the above $\boldsymbol{x}$ and $\boldsymbol{y}$ represent mixed strategies for (I) and (II) respectively. Moreover, $\boldsymbol{x}^*$ and $\boldsymbol{y}^*$ can be found by an LP.

Proof. Let $a_1, ..., a_n$ and $a^1, ..., a^m$ be columns and rows of A respectively. Firstly, observe that for a vector $\mathbf{x}$,

$$\min_{\mathbf{y}} \mathbf{x}^T A\mathbf{y} = \min_i \mathbf{x}^T A\mathbf{1}_i = \min_i < \mathbf{x}, a_i >, \tag{1.9}$$

because Ay is a distribution over $a_1, ..., a_n$. Taking the maximum over all distribution $\mathbf{x}$, we have

$$\max_{\mathbf{x}} \min_{\mathbf{y}} \mathbf{x}^T A\mathbf{y} = \max_{\mathbf{x}} \min_i < \mathbf{x}, a_i > \tag{1.10}$$

Therefore we obtain the following

$$\max_{\mathbf{x}} \min_i < \mathbf{x}, a_i >= \min_{\mathbf{y}} \max_i < a^i, \mathbf{y} >= \mathbf{x}^* A\mathbf{y}^*. \tag{1.11}$$

Both $\max_x \min_i < \mathbf{x}, a_i >$ and $\min_{\mathbf{y}} \max_i < a^i, \mathbf{y} >$ can be formulated by LPs. Then the idea is to show the corresponding LP's are dual of each other and feasible, so they are equal by the strong duality theorem. First, note that $\max_{\mathbf{x}} \min_i < \mathbf{x}, a_i >$ is equivalent to

$$\begin{aligned} &\max \quad t \\ &s.t., \ < \mathbf{x}, a_i >\geq t, \ \ 1 \leq i \leq n \\ &\qquad \sum_{i=1}^{m} x_i = 1 \\ &\qquad x_i \geq 0, \ \ \forall 1 \leq i \leq m \end{aligned} \tag{1.12}$$

We can write the dual of the above LP as follows: We have a dual variable y_i corresponding to each primal constraint $< \mathbf{x}, a_i >\geq t$ and a dual variable w corresponding to the constraint $\sum_{i=1}^m x_i = 1$. Since y_i's correspond to the inequality of constraints in the primal, we need non-negative constraints on y_i's. Since w corresponds to an equality constraint, it will be a free variable. The objective function must be $\min w$, because only the primal constraint corresponding to w has a constant term. In the dual we need to have $m+1$ constraints, one for the primal variable t and the other m constraints are for the x_i's. Since only t appears in the objective of the primal, the constraint corresponding to t has a constant term 1. The dual constraint corresponding to x_i will be as follows:

$$\sum_{j=1}^{m} -a_{i,j} y_i + w \geq 0, \tag{1.13}$$

or equivalently, $w- < \mathbf{y}, a^i > \geq 0$. This gives the following dual formulation:

$$\begin{aligned} & \min \ \ w \\ & s.t., \ w- < \mathbf{y}, a^i > \geq 0, \ \ \forall 1 \leq i \leq m \\ & \sum_{i=1}^{m} y_i = 1 \\ & y_i \geq 0, \ \ \forall 1 \leq i \leq n \end{aligned} \tag{1.14}$$

Now, observe that this is exactly the LP corresponding to $\min_{\mathbf{y}} \max_i < a^i, \mathbf{y} >$. Moreover, let $\mathbf{x}$ and $\mathbf{y}$ be arbitrary distributions and $w = \max_i < \mathbf{y}, a^i >$ and $t = \min_i < \mathbf{x}, a_i >$, shows that both are feasible. So, by the duality theorem,

$$\max_{\mathbf{x}} \min_i < \mathbf{x}, a_i > = \min_{\mathbf{y}} \max_i < a^i, \mathbf{y} > . \tag{1.15}$$

Let x^* and y^* be the optimal solutions of the primal and dual respectively. Then, we have (1.16) and (1.17).

$$\min_i < \mathbf{x}^*, a_i > = \max_i < a^i, \mathbf{y}^* > . \tag{1.16}$$

$$\min_{\mathbf{y}} \mathbf{x}^* A \mathbf{y} = \min_i < \mathbf{x}^*, a_i > = \max_i < a^i, \mathbf{y}^* > = \max_{\mathbf{x}} \mathbf{x}^T A \mathbf{y}^* \tag{1.17}$$

But this means that

$$\mathbf{x}^* A \mathbf{y}^* \leq \max_{\mathbf{x}} \mathbf{x} A \mathbf{y}^* = \min_{\mathbf{y}} \mathbf{x}^* A \mathbf{y} \leq \mathbf{x}^* A \mathbf{y}^*. \tag{1.18}$$

So, all of the above inequalities must be equalities. $\square$

1.3 Optimization

Consider an optimization problem where we are trying to find a solution with minimum cost among a set of feasible solutions. We say an algorithm, ALG, gives an α-approximation for the problem if for any possible input to the problem, we have

$$\frac{\text{cost(ALG)}}{\text{cost(OPT)}} \leq \alpha \tag{1.19}$$

Here, OPT denotes the optimum solution to the problem.

To prove that a given algorithm is an α-approximation, it is sufficient to find a lower-bound for cost(OPT), and then prove that the ratio between cost(ALG) and this lower-bound for any input is upper-bounded by α.

1.3.1 Example: Vertex Cover

Here, we give an application of linear programming in designing an approximation algorithm for a graph problem called vertex cover. We will design a 2-approximation algorithm. This is the best known result for the vertex cover problem. It is a fundamental open problem to beat the factor 2 approximation for the vertex cover problem. In the next lecture we will discuss a generalization of vertex cover called the set cover problem and we see some applications.

Given a graph $G = (V, E)$, we want to find a set $S \subset V$ such that every edge in E is incident to at least one vertex in S. Obviously, we can let $S = V$. But, here among all such sets S we want to choose a one of minimal cost, where $\text{cost}(S)$ is defined as $\sum_{i \in S} c_i$ if every vertex i has associated cost c_i, and $|S|$ if vertices do not have any cost.

In the first step we write a (integer) program which characterizes the optimum solution. Then, we use this program to give a lower bound on the optimum solution. We define this problem with a set of variables $x_i \ \forall i \in V$, where x_i is defined as

$$x_i = \begin{cases} 1 & i \in S \\ 0 & i \notin S \end{cases} \tag{1.20}$$

Our constraint that every edge must be incident to at least one vertex in S can be written as $x_i + x_j \geq 1 \ \forall i \sim j \in E$. So, the question is to find values for all x_i's that minimize the cost of the set S subject to the aforementioned constraint. This can be defined as the following optimization problem

$$\begin{aligned} \min \quad & \sum_{i \in V} c_i x_i \\ \text{s.t.,} \quad & x_i + x_j \geq 1, \ \forall i \sim j \in E \\ & x_i \in \{0, 1\}, \ \forall i \in V \end{aligned} \tag{1.21}$$

Observe that the optimum solution of the above program is exactly equal to the optimum set cover. Note that this is not a linear program, since we have that $x_i \in \{0, 1\}$ for every vertex i, rather than allowing x_i to be a continuous-valued variable. Since the vertex cover problem is NP-hard in general, we do not expect to ever find a general solver to efficiently solve the above integer program. However, there are commercial integer programming solver that work great in practice. They solve a set of linear inequality subject to the each of the underlying variables being 0/1. For many practical applications these program actually find the optimum solution very fast. So, one should always keep them in mind if we are trying to solve an optimization problem in practice.

We can relax the above (integer) program by replacing the integer constraint with the constraint that $0 \leq x_i \leq 1 \ \forall i \in V$. This turns the problem into a linear program. Since this is optimizing over a set of x_i's that includes the optimum set cover, the optimal value of this linear program will be less than or equal to the optimal value of the set cover problem, i.e. OPT LP $\leq$ OPT. The resulting linear program can be written as

$$\begin{aligned} \min \quad & \sum_{i \in V} c_i x_i \\ \text{s.t.,} \quad & x_i + x_j \geq 1, \ \forall i \sim j \in E \\ & 0 \leq x_i \leq 1, \ \forall i \in V \end{aligned} \tag{1.22}$$

Suppose we have an optimal solution of the above program. We want to round this solution into a set cover such that the cost of the cover that we produce is within a small factor of the cost of the LP solution.

The idea is to ue a simple thresholding idea: For each vertex i, if $x_i \geq 0.5$, then we add i to S, otherwise we don't include i in S.

Claim 1.5. *For any solution x of linear program (1.22), the resulting set S, is a vertex cover*

Proof. For a feasible solution x to the linear program, we know that $x_i + x_j \geq 1 \; \forall \, i \sim j \in E$. This means that for every edge $i \sim j$, at least one of x_i, x_j is at least 0.5.Therefore, for any edge $i \sim j$ at least one of i, j is in S. So, S is a vertex cover. □

Claim 1.6. *For any solution x of linear program (1.22) the resulting set S satisfies*

$$\sum_{i \in S} c_i \leq 2 \sum_i c_i x_i = OPT \; LP.$$

This implies that the above algorithm is a 2 approximation for the vertex cover problem.

Proof.

$$\sum_{i \in S} c_i = \sum_{i: x_i \geq 0.5} c_i \leq \sum_{i: x_i \geq 0.5} 2 c_i x_i \leq \sum_i c_i x_i.$$

□

Note that in the worst case $x_i = 0.5$ for all vertices i and the above claim is tight.

1.3.2 Set Cover

Given a set of n elements $V = \{1, 2, ..., n\}$ and a collection of n sets $\{S_1, S_2, ..., S_n\}$ whose union equals the ground set V, the set cover problem is to choose a set $T \subseteq [n]$ with a minimum cost and subject to a constraint that $T \cap S_i \neq \phi, \forall i$. The problem is formulated as (1.23).

$$\begin{aligned} \min \quad & \sum_i x_i c_i \\ \text{s.t.,} \quad & \sum_{i: i \in S_j} x_i \geq 1, \; \forall j. \\ & x_i \in \{0, 1\} \end{aligned} \tag{1.23}$$

Since the problem (1.23) is an NP-hard problem, it can be relaxed via the Linear Programming, where the constraint $x_i \in \{0, 1\}$ is relaxed to $x \geq 0$, to find an optimal point x^*_{lp} such that the optimal value corresponding to x^*_{lp} is a lower bound to the the original problem. Next, a randomized rounding is used, that is

$$Y_i = \begin{cases} 1, & \text{w.p. } \alpha x_i \\ 0, & \text{otherwise} \end{cases} \tag{1.24}$$

The analysis of the randomized rounding

$$\begin{aligned}
\mathbb{P}\left[\sum_{i\in S_j} Y_i = 0\right] &= \mathbb{P}\left[Y_i = 0, \forall i \in S_j\right] \\
&= \prod_{i\in S_j} \mathbb{P}\left[Y_i = 0\right] \\
&= \prod_{i\in S_j} (1-\alpha x_j) \\
&\le \prod_{i\in S_j} e^{-\alpha x_i} \\
&\le e^{-\sum\limits_{i\in S_j} \alpha x_i} \le e^{-\alpha}
\end{aligned}$$

If we choose $\alpha = \log 2m$, we have $\mathbb{P}\left[\sum\limits_{i\in S_j} Y_i = 0\right] \le \frac{1}{2m}$. So, $\mathbb{P}\left[\sum\limits_{i\in S_j} Y_i \ge 0\right] \ge 1 - \frac{1}{2m}$, which means with union bound in every set w.p. $\frac{1}{2}$, we have a probability 1. Furthermore, by the Markov inequality,

$$\mathbb{E}\left[\sum_i c_i Y_i\right] = \alpha \sum_i x_i c_i \le 2\alpha \cdot \text{OPT LP} \le 2\alpha \cdot \text{OPT} \tag{1.25}$$

1.4 Reference

[1] C. Daskalaskis, P. W. Goldberg, and C. H. Papadimitriou. "The complexity of computing an Nash equilibrium". In: SIAM Journal on Computing 39.1 (2009), pp. 195-259.

[2] J. F. Nash et al. "Equilibrium points in n-person games". In: Proc. Nat Acad. Sci. USA 36.1 (1950), pp. 48-49.