PRINCETON UNIV. F'20 COS 521: ADVANCED ALGORITHM DESIGN

Lecture 7: Strong Duality and Dual Fitting

Lecturer: *Sahil Singla* Last updated: September 21, 2020

1 Strong Duality

In the last lecture we discussed weak duality: using dual solutions as upper bounds on how good a primal solution could be. In fact, something quite strong is true: there is always a dual witnessing that the optimal primal is optimal. We'll give a proof, but note that most of the intuition (aside from geometry/linear algebra) is provided by Weak Duality. We'll just discuss the "classic" case, the "partial" case is similar and omitted. First recall the primal and the dual LPs.

$$\max \sum_i c_i x_i \qquad \text{(LP1)}$$
$$\sum_i A_{ji} x_i \le b_j, \quad \forall j$$
$$x_i \ge 0, \quad \forall i.$$

$$\min \sum_j w_j \cdot b_j \qquad \text{(LP2)}$$
$$\sum_j w_j \cdot A_{ji} \ge c_i, \quad \forall i$$
$$w_j \ge 0, \quad \forall j.$$

Theorem 1 (Strong LP Duality)**.** *Let LP1 be any maximization LP and LP2 be its dual (a minimization LP). Then:*

- *If the optimum of LP1 is unbounded ($+\infty$), the feasible region of LP2 is empty.*
- *If the feasible region of LP1 is empty, the optimum of LP2 is either unbounded ($-\infty$), or also infeasible.*
- *If optimum of LP1 finite, then the optimum of LP2 is also finite, and they are equal.*

(Proof adapted from Anupam Gupta's scribed lecture notes here).

The key ingredient in the proof will be what's called the Separating Hyperplane Theorem.

Theorem 2 (Separating Hyperplane Theorem)**.** *Let P be a closed, convex region in $\mathbb{R}^n$, and $\vec{x}$ be a point not in P. Then there exists a $\vec{w}$ such that $\vec{x} \cdot \vec{w} > \max_{\vec{y} \in P}\{\vec{y} \cdot \vec{w}\}$.*

Proof. Consider the point $\vec{y} \in P$ closest to $\vec{x}$ (that is, minimizing $||\vec{x} - \vec{y}||_2$ over all $\vec{y} \in P$. As distance is a positive continuous function, and P is a closed region, such a $\vec{y}$ exists. Now consider the vector $\vec{w} = \vec{x} - \vec{y}$. We claim that the chosen $\vec{w}$ is the desired witness.

Observe first that $(\vec{x} - \vec{y}) \cdot \vec{w} = ||\vec{w}||_2^2 > 0$, so indeed $\vec{x} \cdot \vec{w} > \vec{y} \cdot \vec{w}$. We just need to confirm that $\vec{y} = \arg\max_{\vec{z} \in P}\{\vec{z} \cdot \vec{w}\}$ and then we're done.

Assume for contradiction that $\vec{z} \cdot \vec{w} > \vec{y} \cdot \vec{w}$ and $\vec{z} \in P$. Then as P is convex, $\vec{z}_\varepsilon = (1-\varepsilon)\vec{y} + \varepsilon\vec{z} \in P$ as well for all $\varepsilon > 0$. Observe that $||\vec{x} - \vec{z}_\varepsilon||_2^2 = ||\vec{x} - \vec{y} + \varepsilon(\vec{y} - \vec{z})||_2^2 = ||\vec{x}-\vec{y}||_2^2 + 2\varepsilon(\vec{x}-\vec{y})\cdot(\vec{y}-\vec{z}) + \varepsilon^2||\vec{y}-\vec{z}||_2^2 = ||\vec{x}-\vec{y}||_2^2 + 2\varepsilon\vec{w}\cdot(\vec{y}-\vec{z}) + \varepsilon^2||\vec{y}-\vec{z}||_2^2$. By hypothesis, $\vec{w}\cdot(\vec{y}-\vec{z}) < 0$, and $||\vec{y}-\vec{z}||_2^2$ is finite, so for sufficiently small ε, we get $||\vec{x}-\vec{z}_\varepsilon||_2^2 < ||\vec{x}-\vec{y}||_2^2$, a contradiction. □

Now, consider the optimum $\vec{x}$ of LP1. Let S denote the j for which $\sum_i A_{ji}x_i = b_j$, and $\bar{S}$ the constraints for which $\sum_i A_{ji}x_i < b_j$. We claim that $\vec{c}$ can be written as a convex combination of the vectors $\vec{A}_j$, $j \in S$ (up to possible scaling).

Lemma 3. *Let $\vec{x}$ be the optimum of LP1, and let S denote the j for which $\sum_i A_{ji}x_i = b_j$. Then there exist $\{\lambda_j \geq 0\}_{j\in S}$ such that $c_i = \sum_{j\in S} \lambda_j A_{ji}$ for all i.*

Proof. Assume for contradiction that this were not the case. As the space X of all vectors $\vec{y}$ for which there exists $\{\lambda_j \geq 0\}_{j\in S}$ such that $y_i = \sum_{j\in S} \lambda_j A_{ji}$ for all i is clearly closed and convex, we can apply the separating hyperplane theorem. So there would exist some $\vec{\gamma}$ such that $\vec{c} \cdot \vec{\gamma} > \max_{\vec{y}\in X}\{\vec{y} \cdot \vec{\gamma}\}$. Now consider the vector $\vec{x} + \varepsilon\vec{\gamma}$.

We know that for all $j \in S$, $\sum_i A_{ji}\gamma_i \leq 0$. If not, then $\max_{\vec{y}\in X}\{\vec{y} \cdot \vec{\gamma}\} = +\infty$, because we could blow up λ_j. So for all $j \in S$, $\sum_i A_{ji}(x_i + \varepsilon\gamma_i) \leq b_j$. Moreover, for all $i \notin S$, $\sum_i A_{ji}x_i < b_j$, and $\sum_i A_{ji}\gamma_i$ is finite. So there exists a sufficiently small ε so that $\vec{x} + \varepsilon\vec{\gamma}$ is feasible for LP1.

Finally, observe that $\max_{\vec{y}\in X}\{\vec{y} \cdot \vec{\gamma}\} \geq 0$, as $\vec{0} \in X$. So $\vec{c} \cdot \vec{\gamma} > 0$, and we have shown that $\vec{x}$ was not optimal. $\square$

Now with the lemma in hand, we want to show a dual whose value matches $\vec{c} \cdot \vec{x}$. Let $\vec{c} = \sum_{j\in S} \lambda_j \vec{A}_j$ with $\lambda_j \geq 0$ as guaranteed by the lemma. Set $w_j = \lambda_j$ for all $j \in S$, and $w_j = 0$ for all $j \notin S$. First, is it clear that $\vec{w}$ is feasible for LP2, as we have explicitly set w_j so that $c_i = \sum_j w_j A_{ij}$ for all i. Now we just need to evaluate its value:

$$\sum_j b_j w_j = \sum_{j\in S} b_j w_j + \sum_{j\notin S} b_j \cdot 0 = \sum_{j\in S}(\sum_i A_{ji}x_i)w_j = \sum_i \left(\sum_{j\in S} A_{ji}w_j\right) x_i = \sum_i c_i x_i.$$

So its objective value is exactly the same as LP1.

2 Set Cover using Dual Fitting

In this section we will see an example where we use weak LP duality for an approximation algorithm. Interestingly, our algorithm never solves an LP, it will just use an LP and its dual to analyze the algorithm.

In the *min-cost Set Cover problem* we are given a universe $U := \{1, 2, \ldots, n\}$ of n elements and a collection of m subsets $\mathcal{S} := \{S_1, \ldots, S_m\}$ where $\bigcup_k S_k = U$. We are also given a cost function $c : \mathcal{S} \to \mathbb{R}_{\geq 0}$ that assigns a cost to every subset in $\mathcal{S}$. The goal is to find a min-cost subcollection of $\mathcal{S}$ such that its union equals U. This is a classic NP-Hard problem, e.g., it generalizes the Vertex Cover problem that we studied in Lecture 5. Let OPT denote the cost of the optimal solution. We will analyze approximation guarantees of the Greedy Algorithm in Figure 1:

The Greedy Algorithm always returns a valid solution since we assumed $\bigcup_k S_k = U$. We will prove the following result.

Theorem 4. *The Greedy Algorithm gives an H_n approximation to the min-cost Set Cover problem, where $H_n := 1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n} = \Theta(\log n)$.*

Let A represent uncovered elements and let $\mathcal{C}$ represent selected sets.
Initialization: $A = U$ and $\mathcal{C} = \emptyset$.
While $A \neq \emptyset$:

1. Find a set $S_k \in \mathcal{S}$ that maximizes $\alpha = \frac{1}{c(S_k)} \cdot \left(\left|S_k \cap A\right|\right)$.
2. For each newly covered element $e \in S_k \cap A$, set $price(e) = 1/\alpha$.
3. Update $A \leftarrow A \setminus S_k$ and $\mathcal{C} \leftarrow \mathcal{C} \cup k$.

Note that the second step doesn't affect the algorithm. It will be only used for analysis.

Figure 1: The Greedy Algorithm

There are several proofs known for this theorem. (You might want to try a direct combinatorial proof.) We will consider an LP based proof where we consider the following natural LP:

$$\min \sum_{k \in \{1,\ldots,m\}} c(S_k) \cdot x_k$$
$$\sum_{k : S_k \ni e} x_k \geq 1 \qquad \forall e \in \{1,\ldots,n\}$$
$$x_k \geq 0 \qquad \forall k \in \{1,\ldots,m\}$$

Since by setting $x_k = 1$ for sets S_k in the optimal solution, and otherwise setting $x_k = 0$, we get a feasible solution to the LP of cost OPT, we know that the optimal fractional solution x^* to the LP satisfies $\sum_k c(S_k) x_k^* \leq OPT$.

We will actually never compute x^*[1]

and only use the fact that the LP value gives a lower bound on OPT. Instead, we compare the Greedy Algorithm to the following dual LP (since the primal LP was a minimization problem, the dual LP is a maximization problem):

$$\max \sum_{e \in \{1,\ldots,n\}} y_e$$
$$\sum_{e \in S_k} y_e \leq c(S_k) \qquad \forall k \in \{1,\ldots,m\}$$
$$y_e \geq 0 \qquad \forall e \in \{1,\ldots,n\}$$

By weak-duality, we know that any feasible solution to the dual LP gives a lower bound on $\sum_k c(S_k) x_k^*$. Thus, our plan to upper bound the cost of the greedy algorithm is to show a feasible dual solution $\left(z_e\right)_{e \in E}$ which satisfies that the total cost of the Greedy Algorithm

$$\sum_{k \in \mathcal{C}} c(S_k) \leq H_n \cdot \sum_e z_e. \tag{1}$$

[1] Exercise: The randomized rounding algorithm that selects each set S_k independently with probability $\min\{1\ ,\ x_k^* \cdot \log n\}$ is an $O(\log n)$ approximation algorithm.

This will prove Theorem 4 since by weak-duality $\sum_e z_e \le \sum_k c(S_k) x_k^* \le OPT$.

We define $z_e = \dfrac{1}{H_n} \cdot price(e)$, where recall from Figure 1 that $price(e)$ intuitively denotes how much element e pays for S_k when it first gets covered. Proving (1) is easy because the Greedy Algorithm sets an element's price exactly once, and at that point the algorithm exactly distributes $c(S_k)$ between the newly covered elements. Thus the algorithm's total cost $\sum_{k \in \mathcal{C}} c(S_k) = \sum_e price(e) = \frac{1}{H_n} \sum_e z_e$.

Next, we prove that $\left(z_e\right)_{e \in E}$ is a feasible solution to the dual LP. By definition, $z_e \ge 0$. Now consider any set $S \in \mathcal{S}$. We want to show $\sum_{e \in S} z_e \le c(S)$. Start by renaming the elements (for analysis) of S to be $\{e_1, e_2, \ldots, e_{|S|}\}$ in the order they are covered by the Greedy Algorithm (breaking ties arbitrarily). Observe that whenever an element $e_i \in S$ was first covered, the algorithm had the option of selecting S (and may be it even did) at a cost of $c(S)$ and cover $|S| - i + 1$ new elements. Thus, irrespective of which set the Greedy algorithm actually selects, we know that $price(e_i) \le \frac{c(S)}{|S|-i+1}$. This implies

$$\sum_{e \in S} z_e \quad = \quad \frac{1}{H_n} \sum_{e_i \in S} price(e_i) \quad \le \quad \frac{1}{H_n} \sum_{e_i \in S} \frac{c(S)}{|S| - i + 1} \quad = \quad c(S) \cdot \frac{H_{|S|}}{H_n} \quad \le \quad c(S).$$