princeton univ. F'14 cos 521: Advanced Algorithm Design

Lecture 18: Duality and MinMax Theorem

Lecturer: *Sanjeev Arora* Scribe:

We are used to the concept of duality in life: yin and yang, Mars and Venus, etc. In mathematics duality refers to the phenomenon whereby two objects that look very different are actually the same in a technical sense.

Today we first see LP duality, which will then be explored a bit more in the homeworks. Duality has several equivalent statements.

1. If K is a polytope and p is a point outside it, then there is a hyperplane separating p from K.

2. The following system of inequalities

$$\begin{array}{rcl} \mathbf{a}_1 \cdot \mathbf{X} & \geq & b_1 \\ \mathbf{a}_2 \cdot \mathbf{X} & \geq & b_2 \\ & \vdots & \\ \mathbf{a}_m \cdot \mathbf{X} & \geq & b_m \\ \mathbf{X} & \geq & 0 \end{array} \tag{1}$$

   is infeasible iff using positive linear combinations of the inequalities it is possible to derive $-1 \geq 0$, i.e. there exist $\lambda_1, \lambda_2, \ldots \lambda_m \geq 0$ such that

$$\sum_{i=1}^{m} \lambda_i \mathbf{a}_i < 0 \quad \text{and} \quad \sum_{i=1}^{m} \lambda_i b_i > 0.$$

   This statement is called *Farkas's Lemma.*

1 Linear Programming and Farkas' Lemma

In courses and texts duality is taught in context of LPs. Say the LP looks as follows:

Given: vectors $\mathbf{c}, \mathbf{a}_1, \mathbf{a}_2, \ldots \mathbf{a}_m \in \mathbf{R}^n$, and real numbers $b_1, b_2, \ldots b_m$.

Objective: find $\mathbf{X} \in \mathbf{R}^n$ to minimize $\mathbf{c} \cdot \mathbf{X}$, subject to:

$$\begin{array}{rcl} \mathbf{a}_1 \cdot \mathbf{X} & \geq & b_1 \\ \mathbf{a}_2 \cdot \mathbf{X} & \geq & b_2 \\ & \vdots & \\ \mathbf{a}_m \cdot \mathbf{X} & \geq & b_m \\ \mathbf{X} & \geq & 0 \end{array} \tag{2}$$

The notation $\mathbf{X} > \mathbf{Y}$ simply means that $\mathbf{X}$ is componentwise larger than $\mathbf{Y}$. Now we represent the system in (2) more compactly using matrix notation. Let

$$A = \begin{pmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \vdots \\ \mathbf{a}_m^T \end{pmatrix} \quad \text{and} \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

Then the Linear Program (LP for short) can be rewritten as:

$$\begin{array}{rl} \min & \mathbf{c}^T\mathbf{X} : \\ A\mathbf{X} & \geq \mathbf{b} \\ \mathbf{X} & \geq 0 \end{array} \tag{3}$$

This form is general enough to represent any possible linear program. For instance, if the linear program involves a linear equality $\mathbf{a} \cdot \mathbf{X} = b$ then we can replace it by two inequalities

$$\mathbf{a} \cdot \mathbf{X} \geq b \quad \text{and} \quad -\mathbf{a} \cdot \mathbf{X} \geq -b.$$

If the variable X_i is unconstrained, then we can replace each occurence by $X_i^+ - X_i^-$ where X_i^+, X_i^- are two new non-negative variables.

2 LP Duality Theorem

With every LP we can associate another LP called its *dual*. The original LP is called the *primal*. If the primal has n variables and m constraints, then the dual has m variables and n constraints. Thus there is a primal variable corresponding to each dual constraint, and a dual variable for each primal constraint.

$$\begin{array}{rl|rl} \multicolumn{2}{c|}{\underline{\text{Primal}}} & \multicolumn{2}{c}{\underline{\text{Dual}}} \\ \min & \mathbf{c}^T\mathbf{X} : & \max & \mathbf{Y}^T\mathbf{b} : \\ A\mathbf{X} & \geq \mathbf{b} & \mathbf{Y}^T A & \leq \mathbf{c}^\mathbf{T} \\ \mathbf{X} & \geq 0 & \mathbf{Y} & \geq 0 \end{array} \tag{4}$$

(Aside: if the primal contains an equality constraint instead of inequality then the corresponding dual variable is unconstrained.)

It is an easy exercise that the dual of the dual is just the primal.

Theorem 1

The Duality Theorem. *If both the Primal and the Dual of an LP are feasible, then the two optima coincide.*

Proof: The proof involves two parts:

1. Primal optimum $\geq$ Dual optimum.
   This is the easy part. Suppose $\mathbf{X}^*, \mathbf{Y}^*$ are the respective optima. This implies that

$$A\mathbf{X}^* \geq \mathbf{b}.$$

Now, since $\mathbf{Y}^* \geq 0$, the product $\mathbf{Y}^* A\mathbf{X}^*$ is a non-negative linear combination of the rows of $A\mathbf{X}^*$, so the inequality

$$\mathbf{Y}^{*T} A\mathbf{X}^* \geq \mathbf{Y}^{*T}\mathbf{b}$$

holds. Again, since $\mathbf{X}^* \geq 0$ and $\mathbf{c}^T \geq \mathbf{Y}^{*T} A$, we obtain the inequality

$$\mathbf{c}^T\mathbf{X}^* \geq (\mathbf{Y}^{*T} A)\mathbf{X}^*.$$

Examining the previous two lines we conclude $\mathbf{c}^T\mathbf{X}^* \geq \mathbf{Y}^{*T}\mathbf{b}$, which completes the proof of this part.

2. Dual optimum $\geq$ Primal optimum.
   Let k be the optimum value of the primal. Since the primal is a minimization problem, the following set of linear inequalities is infeasible for any $\epsilon > 0$:

$$\begin{array}{rl} -\mathbf{c}^T\mathbf{X} & \geq -(k-\epsilon) \\ A\mathbf{X} & \geq \mathbf{b} \\ X & \geq 0 \end{array} \tag{5}$$

   Here, ϵ is a small positive quantity. Therefore, by Farkas' Lemma, there exist $\lambda_0, \lambda_1, \ldots \lambda_m \geq 0$ such that

$$-\lambda_0\mathbf{c} + \sum_{i=1}^{m} \lambda_i \mathbf{a}_i < 0 \tag{6}$$

$$-\lambda_0(k-\epsilon) + \sum_{i=1}^{m} \lambda_i b_i > 0. \tag{7}$$

   Note that $\lambda_0 > 0$ omitting the first inequality in (5) leaves a feasible system by assumption about the primal. Thus, consider the nonnegative vector

$$\Lambda = (\frac{\lambda_1}{\lambda_0}, \ldots \frac{\lambda_m}{\lambda_0})^T.$$

   The inequality (6) implies that $\Lambda^T A \leq \mathbf{c}^T$. So Λ is a feasible solution to the Dual. The inequality (7) implies that $\Lambda^T\mathbf{b} > (k-\epsilon)$, and since the Dual is a maximization problem, this implies that the Dual optimal is bigger than $k-\epsilon$. Since this holds for every $\epsilon > 0$, by compactness we conclude that there is a Dual feasible solution of value k. Thus, this part is proved, too. Hence the Duality Theorem is proved.

□

My thoughts on this business:
(1) Usually textbooks bundle the case of infeasible systems into the statement of the Duality theorem. This muddies the issue for the student. Usually all applications of LPs fall into two cases: (a) We either know (for trivial reasons) that the system is feasible, and are only interested in the value of the optimum or (b) We do not know if the system is feasible and that is precisely what we want to determine. Then it is best to just use Farkas' Lemma.

(2) The proof of the Duality theorem is interesting. The first part shows that for any dual feasible solution $\mathbf{Y}$ the various Y_i's can be used to obtain a *weighted* sum of primal inequalities, and thus obtain a lowerbound on the primal. The second part shows that this method of taking weighted sums of inequalities is *sufficient* to obtain the best possible lowerbound on the primal: there is no need to do anything fancier (e.g., taking products of inequalities or some such thing).

3 Example: Max Flow Min Cut theorem in graphs

The input is a directed graph $G(V, E)$ with one source s and one sink t. Each edge e has a capacity c_e. The flow on any edge must be less than its capacity, and at any node apart from s and t, flow must be conserved: total incoming flow must equal total outgoing flow. We wish to maximize the flow we can send from s to t. The maximum flow problem can be formulated as a Linear Program as follows:

Let $\mathcal{P}$ denote the set of all (directed) paths from s to t. Then the max flow problem becomes:

$$\max \sum_{P \in \mathcal{P}} f_P : \tag{8}$$

$$\forall P \in \mathcal{P} : f_P \geq 0 \tag{9}$$

$$\forall e \in E : \sum_{P : e \in P} f_P \leq c_e \tag{10}$$

Since P could contain exponentially many paths, this is an LP with exponentially many variables. Luckily duality tells us how to solve it using the Ellipsoid method.

Going over to the dual, we get:

$$\min \sum_{e \in E} c_e y_e : \tag{11}$$

$$\forall e \in E : y_e \geq 0 \tag{12}$$

$$\forall P \in \mathcal{P} : \sum_{e \in P} y_e \geq 1 \tag{13}$$

Notice that the dual in fact represents the fractional min $s - t$ cut problem: think of each edge e being picked up to a fraction y_e. The constraints say that a total weight of 1 must be picked on each path. Thus the usual s-t min cut problem simply involves $0 - 1$ solutions to the y_e's in the dual.

Exercise 1 Prove that the optimum solution does have $y_e \in \{0, 1\}$, and thus the solution to the dual is the best s-t min cut.

Thus, LP duality implies max-*st*-flow = (capacity of) min-cut.

Polynomial-time algorithms? The primal has exponentially many variables! (Aside: turns out it is equivalent to a more succinct LP but lets' proceed with this one.) Nevertheless we can use the Ellipsoid method by applying it to the dual, which has m variables and exponentially many constraints. As we saw last time, we only need to show a polynomial-time separation oracle for the dual. Namely, for each candidate vector (y_e) we need to check if it satisfies all the dual constraints. This can be done by creating a weighted version of the graph where the weight on edge e is y_e. Then compute the shortest path from s to t in this weighted graph. If the shortest path has length < 1 then we have found a violated constraint.

Of course, for Max Flow we know of much faster algorithms than the Ellipsoid method (e.g., the algorithms you saw in your undergrad course), but there are other LPs with exponentially many variables for which the only known polynomial time algorithms go via the Ellipsoid method.

4 Game theory and the minmax theorem

In the 1930s, polymath John von Neumann (professor at IAS, now buried in the cemetery close to downtown) was interested in applying mathematical reasoning to understand strategic interactions among people —or for that matter, nations, corporations, political parties, etc. He was a founder of *game theory*, which models rational choice in these interactions as maximization of some payoff function.

A starting point of this theory is the *zero-sum* game. There are two players, 1 and 2, where 1 has a choice of m possible moves, and 2 has a choice of n possible moves. When player 1 plays his ith move and player 2 plays her jth move, the outcome is that player 1 pays A_{ij} to player 2. Thus the game is completely described by an $m \times n$ *payoff* matrix.

-	**scissors**	**paper**	**rock**
rock	1	−1	0
paper	−1	0	1
scissors	0	1	−1

Figure 1: Payoff matrix for Rock/Paper/Scissor

This setting is called *zero sum* because what one player wins, the other loses. By contrast, war (say) is a setting where both parties may lose material and men. Thus their combined worth at the end may be lower than at the start. (Aside: An important stimulus for development of game theory in the 1950s was the US government's desire to behave "strategically "in matters of national defence, e.g. the appropriate tit-for-tat policy for waging war —whether nuclear or conventional or cold.)

von Neumann was interested in a notion of equilibrium. In physics, chemistry etc. an equilibrium is a stable state for the system that results in no further change. In game theory it is a pair of strategies g_1, g_2 for the two players such that each is the optimum response to the other.

Let's examine this for zero sum games. If player 1 announces he will play the ith move, then the *rational* move for player 2 is the move j that maximises A_{ij}. Conversely, if player 2 announces she will play the jth move, player 1 will respond with move i' that minimizes $A_{i'j}$. In general, there may be no *equilibrium* in such announcements: the response of player 1 to player 2's response to his announced move i will not be i in general:

$$\min_i \max_j A_{ij} \neq \max_j \min_i A_{ij}.$$

In fact there is no such equilibrium in Rock/paper/scissors either, as every child knows.

von Neumann realized that this lack of equilibrium disappears if one allows players' announced strategy to be a *distribution* on moves, a so-called *mixed* strategy. Player 1's distribution is $x \in \Re^m$ satisfying $x_i \geq 0$ and $\sum_i x_i = 1$; Player 2's distribution is $y \in \Re^n$ satisfying $y_j \geq 0$ and $\sum_j y_j = 1$. Clearly, the expected payoff from Player 1 to Player 2 then is $\sum_{ij} x_i A_{ij} y_j = x^T Ay$.

But has this fixed the problem about nonexistence of equilibrium? If Player 1 announces first the payoff is $\min_x \max_y x^T Ay$ whereas if Player 2 announces first it is $\max_y \min_x x^T Ay$. The next theorem says that it doesn't matter who announces first; neither player has an incentive to change strategies after seeing the other's announcement.

Theorem 2 (Famous Min-Max Theorem of Von Neumann)
$\min_x \max_y x^T Ay = \max_y \min_x x^T Ay$.

Turns out this result is a simple consequence of LP duality and is equivalent to it. You will explore it further in the homework.

What if the game is not zero sum? Defining an equilibrium for it was an open problem until John Nash at Princeton managed to define it in the early 1950s; this solution is called a Nash equilibrium. We'll return to it in a future lecture. BTW, you can still sometimes catch a glimpse of Nash around campus.