Optimization and Approximation	ENS Lyon
Lecture 5	
Lecturer: Alantha Newman	*October 4, 2016*

1 Linear Programming Duality

For a minimization linear program, any feasible solution provides an upper bound on the optimal objective value. But how can we provide a good *lower* bound on the optimal value? In other words, given a solution that is claimed to be optimal, can we prove that this is indeed the minimum solution? How do we prove that there is no solution with a smaller objective value?

In this lecture, we show how to compute a lower bound on the value of an optimal solution for a given linear program corresponding to a minimization problem. (Note that "lower bound" can be replaced by "upper bound" in the case of a maximization problem.) Such lower bounds have several useful applications. Recall (from the first lecture) that the first step in designing an approximation algorithm for a minimization problem is to efficiently compute a "tight" *lower* bound on the optimal value. Moreover, it is often expensive to compute the exact solution for a linear program, and sometimes a lower bound is sufficient.

To demonstrate the idea of duality, we consider the following linear program:

$$\begin{array}{lrcl} \text{min} & 7x+3y & & \\ \text{subject to:} & x+y & \geq & 2 \\ & 3x+y & \geq & 4 \\ & x,y & \geq & 0 \end{array}$$

How can we lower bound the optimal value of this linear program? Here are some easy lower bounds:

$$\begin{aligned} 7x+3y &\geq x+y \geq 2, \\ 7x+3y &\geq 3x+3y \geq 6. \end{aligned}$$

Now consider the following: Multiply the second constraint by 2, and then add it to the first constraint. We obtain the following:

$$\begin{aligned} \underbrace{(x+y)}_{\geq 2} + 2\underbrace{(3x+y)}_{\geq 2(4)} &= 7x+3y \\ &\geq 10. \end{aligned}$$

The above approach shows that 10 is the minimum feasible solution for the considered linear program. In fact the solution $(1,1)$ has objective value 10, so this solution is optimal. The method of linear programming duality generalizes the above approach. The basic idea is to find the multipliers for the constraints that yields the largest lower bound for a minimization linear program.

1.1 Another Example

Consider the following linear program:

$$\begin{aligned} \text{max } & 4x_1 + x_2 + 5x_3 + 3x_4 \\ \text{subject to: } & x_1 - x_2 - x_3 + 3x_4 \leq 1 \\ & 5x_1 + x_2 + 3x_3 + 8x_4 \leq 55 \\ & -x_1 + 2x_2 + 3x_3 - 5x_4 \leq 3 \\ & x_i \geq 0. \end{aligned} \qquad (\mathcal{P}_1)$$

We would like to compute an upper bound on the value of an optimal solution.

- Multiplying the second constraint by $\frac{5}{3}$ gives:

$$z \;\leq\; \frac{25}{3}x_1 + \frac{5}{3}x_2 + 5x_3 + \frac{40}{3}x_4 \;\leq\; \frac{275}{3}.$$

- The sum of the last two constraints gives a better (smaller) upper bound:

$$z \;\leq\; 4x_1 + 3x_2 + 6x_3 + 3x_4 \;\leq\; 58.$$

More generally, one can multiply each constraint C_i by some variable $y_i \geq 0$ and sum them. If:

$y_1 + 5y_2 - y_3 \geq 4$;

$-y_1 + y_2 + 2y_3 \geq 1$;

$-y_1 + 3y_2 + 3y_3 \geq 5$;

$3y_1 + 8y_2 - 5y_3 \geq 3$;

then we have:

$$\begin{aligned} y_1 + 55y_2 + 3y_3 &\geq y_1(x_1 - x_2 - x_3 + 3x_4) + y_2(5x_1 + x_2 + 3x_3 + 8x_4) + y_3(-x_1 + 2x_2 + 3x_3 - 5x_4) \\ &= (y_1 + 5y_2 - y_3)x_1 + (-y_1 + y_2 + 2y_3)x_2 + (-y_1 + 3y_2 + 3y_3)x_3 + (3y_1 + 8y_2 - 5y_3)x_4 \\ &\geq 4x_1 + x_2 + 5x_3 + 3x_4. \end{aligned}$$

This implies that $y_1 + 55y_2 + 3y_3$ is an upper bound on z.

So to obtain the best (smallest) upper bound possible, let us consider the following linear program:

$$\begin{aligned} \min \quad & y_1 + 55y_2 + 3y_3 \\ \text{subject to:} \quad & y_1 + 5y_2 - y_3 \geq 4 \\ & -y_1 + y_2 + 2y_3 \geq 1 \\ & -y_1 + 3y_2 + 3y_3 \geq 5 \\ & 3y_1 + 8y_2 - 5y_3 \geq 3 \\ & y_1, y_2, y_3 \geq 0. \end{aligned} \qquad (\mathcal{D}_1)$$

The linear program $(\mathcal{D}_1)$ is the *dual* of the primal linear program $(\mathcal{P}_1)$, which is in turn the dual of $(\mathcal{D}_1)$.

1.2 Primal($\mathcal{P}$) to Dual($\mathcal{D}$)

We now formally show how to obtain a dual linear program for a given primal linear program. Here, we will consider the case in which the primal linear program is a minimization problem. The case in which the primal linear program is a maximization linear program is analogous. The primal linear program can be written in the following equivalent forms:

$$\begin{aligned} \min \quad & \mathbf{c}^\intercal \mathbf{x} \\ \text{subject to:} \quad & \boldsymbol{A}\mathbf{x} \geq b \\ & \mathbf{x} \geq 0 \end{aligned} \quad \rightarrow \quad \begin{aligned} \min \quad & \sum_{i=1}^{n} c_i x_i \\ \text{subject to:} \quad & \sum_{i=1}^{n} a_{ji} x_i \geq b_j \quad \forall j \leq m \\ & x_i \geq 0 \quad \forall i \leq n \end{aligned} \qquad (\mathcal{P})$$

The dual linear program is obtained as follows:

$$\begin{array}{llll} \max & \mathbf{b}^\intercal \mathbf{y} & & \\ \text{subject to} & \boldsymbol{A^T}\mathbf{y} & \leq & \mathbf{c} \\ & \mathbf{y} & \geq & 0 \end{array} \quad \rightarrow \quad \begin{array}{lllll} \max & \sum_{j=1}^{m} b_j y_j & & & \\ \text{subject to} & \sum_{j=1}^{m} a_{ji} y_j & \leq & c_i & \forall i \leq n \\ & y_j & \geq & 0 & \forall j \leq m \end{array} \qquad (\mathcal{D})$$

1. For each constraint in $(\mathcal{P})$, we assign a multiplier variable y_j.
2. For each variable in $(\mathcal{P})$, we have a corresponding constraint in $(\mathcal{D})$ that upper bounds the value of the coefficient of that variable. (That is, after multiplying each constraint in $(\mathcal{P})$ with its corresponding multiplier and then summing all of these constraints, the coefficient corresponding to each variable in $(\mathcal{P})$ should be upper bounded by its corresponding coefficient in the primal objective function.)
3. The dual objective function maximizes the lower bound on $(\mathcal{P})$ by maximizing $\mathbf{b}^\intercal \mathbf{y}$.

Note that if the j^{th} constraint in the primal is in the form $\geq$ or $\leq$, then the corresponding variable y_j is constrained to be nonnegative. However, if the j^{th} constraint is an equality constraint, then the corresponding variable y_j is unconstrained. To see why this is the case, consider the following example.

$$\begin{aligned} \max \quad & x_1 + x_2 + x_3 \\ \text{subject to:} \quad & -2x_1 + 4x_2 + 5x_3 \leq 3 \\ & 4x_1 + x_2 - 3x_3 = 2 \\ & x_1, x_2, x_3 \geq 0. \end{aligned} \qquad (\mathcal{P}_2)$$

This is equivalent to:

$$\begin{aligned} \max \quad & x_1 + x_2 + x_3 \\ \text{subject to:} \quad & -2x_1 + 4x_2 + 5x_3 \leq 3 \\ & 4x_1 + x_2 - 3x_3 \leq 2 \\ & -4x_1 - x_2 + 3x_3 \leq -2 \\ & x_1, x_2, x_3 \geq 0. \end{aligned} \qquad (\mathcal{P}_2)$$

Now, when we take the dual, we have:

$$\begin{aligned} \min \quad & 3y_1 + 2y_2 - 2y_3 \\ \text{subject to:} \quad & -2y_1 + 4y_2 - 4y_3 \geq 1 \\ & 4y_1 + y_2 - y_3 \geq 1 \\ & 5y_1 - 3y_2 + 3y_3 \geq 1 \\ & y_1, y_2, y_3 \geq 0. \end{aligned} \qquad (\mathcal{D}_2)$$

Let $y_2' = y_2 - y_3$. Then, $(\mathcal{D}_2)$ is equivalent to:

$$\begin{aligned} \min \quad & 3y_1 + 2y_2' \\ \text{subject to:} \quad & -2y_1 + 4y_2' \geq 1 \\ & 4y_1 + y_2' \geq 1 \\ & 5y_1 - 3y_2' \geq 1 \\ & y_1 \geq 0, \end{aligned} \qquad (\mathcal{D}_2)$$

where the variable y_2' is unconstrained.

1.3 Weak and Strong Duality

The "Weak Duality Theorem" states that any feasible solution to the dual linear is a lower bound on the optimal value of the corresponding primal linear program.

Theorem 1. (Weak Duality) *If* $\mathbf{x}$ *is a primal-feasible solution and* $\mathbf{y}$ *is a dual-feasible solution, then* $\mathbf{c}^\intercal\mathbf{x} \geq \mathbf{b}^\intercal\mathbf{y}$.

Proof.

$$\begin{aligned}\mathbf{b}^\intercal\mathbf{y} = \sum_j b_j y_j &\leq \sum_j (\sum_i a_{ji} x_i) y_j \\ &= \sum_i (\sum_j a_{ji} y_j) x_i \\ &\leq \sum_i c_i x_i = \mathbf{c}^\intercal\mathbf{x}.\end{aligned}$$

□

For completeness, we will state the "Strong Duality Theorem". However, there are many applications where Weak Duality suffices.

Theorem 2. (Strong Duality) *If* $\mathbf{x}$ *is an optimal primal-feasible solution and* $\mathbf{y}$ *is an optimal dual-feasible solution, then* $\mathbf{c}^\intercal\mathbf{x} = \mathbf{b}^\intercal\mathbf{y}$, *i.e the optimal objective value of the primal equals the optimal objective value of the dual. Furthermore, if the primal is unbounded (infeasible), then the dual is infeasible (unbounded).*

1.4 Certificates of Optimality

It might seem that there is no benefit in solving the dual linear program; if obtaining such a lower bound requires us to solve a linear program, why not simply solve the primal linear program? However, note that once we have solved the dual linear program, the multipliers obtained (i.e. the y_j values) can be used to prove that an optimal solution to the primal linear program is indeed optimal. Given these multipliers, the process of verifying optimality (by a third party) can be much faster than (re)solving the primal linear program.

In other words, if we find solutions for both the primal and the dual and these solutions have the same values, then this is a *certificate of optimality*. There is no need to re-run the computation (e.g. re-run the simplex algorithm) to convince someone else of the optimality of the proposed solution. For example, the optimal solution for $(\mathcal{P}_1)$ is $(0, 14, 0, 5)$ with objective value 29. The optimal solution for $(\mathcal{D}_1)$ is $(11, 0, 6)$, also with an objective value of 29. This proves that both $(\mathcal{P}_1)$ and $(\mathcal{D}_1)$ have optimal value 29.

These certificates of optimality are closely related to NP certificates. Consider the following decision problems:

1. Does $(\mathcal{P})$ have an optimal solution with value $\leq \gamma$?
2. Does $(\mathcal{P})$ have an optimal solution with value $> \gamma$?

To answer the first question, we can provide an $\mathbf{x}$ such that $\mathbf{x}$ is a feasible solution for $(\mathcal{P})$ and $\mathbf{c}^\intercal\mathbf{x} \leq \gamma$. This shows that the problem of deciding the answer to the first question is in NP. To answer the second question, we can provide a $\mathbf{y}$ such that $\mathbf{y}$ is a feasible solution for $(\mathcal{D})$ and $\mathbf{b}^\intercal\mathbf{y} > \gamma$. This shows that the problem of deciding the answer to the first question is in co-NP.

The theory of linear programming duality is attributed to Von Neumann [vN63] and Gale, Kuhn and Tucker [GKT51]. Thus, they showed that the (decision) linear programming problem is in NP∩co-NP well before these complexity classes were defined! Many people believe that belonging to NP∩co-NP is a good indication that a problem has a polynomial-time algorithm, which was later proved to be the case for linear programming by Khachiyan [Kha80].

2 Computing the Dual Solution via the Simplex Algorithm

Recall the following linear program from Lecture 3:

$$\begin{aligned} \max \quad & 2x_1 + x_2 \\ \text{subject to:} \quad & -2x_1 + x_2 \leq -2 \\ & x_1 - 2x_2 \leq -2 && (1) \\ & x_1 + x_2 \leq 7 && (2) \\ & x_1, x_2 \geq 0. && (\mathcal{P}_3) \end{aligned}$$

The optimal solution $(4, 3)$ has value 11. The dual linear program is:

$$\begin{aligned} \min \quad & -2y_1 - 2y_2 + 7y_3 \\ \text{subject to:} \quad & -2y_1 + y_2 + y_3 \geq 2 \\ & y_1 - 2y_2 + y_3 \geq 1 \\ & y_1, y_2, y_3 \geq 0. && (\mathcal{D}_3) \end{aligned}$$

By duality, there is a solution of $(\mathcal{D}_3)$ with value 11.

Idea: The optimum of $(\mathcal{P}_3)$ is attained at the intersection of constraints (1) and (2), where both of these constraints are tight. One can express the objective function at $(4, 3)$ (i.e. $2x_1 + x_2 = 11$) as a linear combination of the constraints $x_1 - 2x_2 = -2$ and $x_1 + x_2 = 7$. For example,

$$\frac{1}{3}(x_1 - 2x_2 = -2) + \frac{5}{3}(x_1 + x_2 = 7) \quad \Rightarrow \quad (2x_1 + x_2 = 11).$$

In other words, $(0, \frac{1}{3}, \frac{5}{3})$ is an optimal solution for $(\mathcal{D}_3)$.

Crucial Remark: In all dictionaries of the simplex algorithm–and notably in the final dictionary–the objective function z is expressed as a combination of constraints. Therefore, the linear combination of constraints certifying optimality can be read from the last dictionary.

For example, consider the final dictionary (from Lecture 3):

$$\begin{aligned} x_2 &= 3 + \frac{x_4}{3} - \frac{x_5}{3} \\ x_1 &= 4 - \frac{x_4}{3} - \frac{2x_5}{3} \\ x_3 &= 3 - x_4 - x_5 \\ z &= 11 - \frac{x_4}{3} - \frac{5x_5}{3}. && (D_F) \end{aligned}$$

$$\begin{aligned} x_4 &= -2 - x_1 + 2x_2 \\ x_5 &= 7 - x_1 - x_2. \end{aligned}$$

Thus, we have:

$$\begin{aligned} z = 2x_1 + x_2 &= 11 - \frac{x_4}{3} - \frac{5x_5}{3} \quad \Rightarrow \\ 2x_1 + x_2 - 11 &= -\frac{x_4}{3} - \frac{5x_5}{3}. \end{aligned}$$

In summary, when you write your program, you can use the dual solution generated by the last dictionary to check the optimality of the final (primal) solution.

2.1 Certifying Infeasibility

Another interesting case is when the primal linear program has no solution. For example, if that last constraint of our example is replaced by $x_1 + x_2 \leq 3$:

$$\begin{aligned} \max \quad & 2x_1 + x_2 \\ \text{subject to:} \quad & -2x_1 + x_2 \leq -2 \\ & x_1 - 2x_2 \leq -2 \\ & x_1 + x_2 \leq 3 \\ & x_1, x_2 \geq 0. \end{aligned} \qquad (\mathcal{P}_4)$$

Then we have the following auxiliary problem:

$$\begin{aligned} \min \quad & x_0 \quad (= \max -x_0) \\ \text{subject to:} \quad & -2x_1 + x_2 \leq -2 + x_0 \\ & x_1 - 2x_2 \leq -2 + x_0 \\ & x_1 + x_2 \leq 3 + x_0 \\ & x_0, x_1, x_2 \geq 0. \end{aligned} \qquad (\mathcal{P}'_4)$$

The final dictionary of Phase I of the simplex on $(\mathcal{P}'_4)$ is:

$$\begin{aligned} x_0 &= \frac{1}{3} + \frac{x_4}{3} + \frac{x_5}{3} + \frac{x_3}{3} \\ x_1 &= \frac{5}{3} \qquad\quad - \frac{x_5}{3} + \frac{x_3}{3} \\ x_2 &= \frac{5}{3} + \frac{x_4}{3} - \frac{x_5}{3} \\ w &= -\frac{1}{3} - \frac{x_4}{3} - \frac{x_5}{3} - \frac{x_3}{3}. \end{aligned} \qquad (D'_F)$$

This means that the optimal value of $(\mathcal{P}'_4)$ is $-\frac{1}{3}$ and therefore $(\mathcal{P}_4)$ is infeasible. The coefficients in w can be used to certify optimality of $(\mathcal{P}'_4)$ and therefore infeasibility of $(\mathcal{P}_4)$. The left-hand side is:

$$\frac{1}{3}\left((-2x_1 + x_2) + (x_1 - 2x_2) + (x_1 + x_2)\right) = 0.$$

But the right-hand side is:

$$\frac{1}{3}(-2 + -2 + 3) = -\frac{1}{3}.$$

This gives: $0 \leq -\frac{1}{3}$, which proves that $(\mathcal{P}_4)$ is infeasible.

3 Outcomes for Pairs of Primal/Dual Linear Programs

Recall that a linear program has three outcomes: (i) a bounded optimal solution, (ii) unbounded, or (iii) infeasible. Let us consider what happens for the primal and dual:

	$(\mathcal{P})$ OPT	$(\mathcal{P})$ Unbounded	$(\mathcal{P})$ Infeasible
$(\mathcal{D})$ OPT	Yes	No	No
$(\mathcal{D})$ Unbounded	No	No	Yes
$(\mathcal{D})$ Infeasible	No	Yes	Yes

(i) The table is symmetric since the dual of the dual is the primal.

(ii) Strong Duality: Bounded optimal value for $\mathcal{P}$ iff bounded optimal value for $\mathcal{D}$.

(iii) Weak Duality: if $\mathcal{P}$ is unbounded, its value is arbitrarily large. Thus, we have an arbitrarily large lower bound on the value of the dual, which is therefore infeasible.

(iv) Exercise: Find a linear program such that both primal and dual are infeasible!

4 Applications of Duality: Max Flow–Min Cut Theorem

Consider the maximum flow problem. Given a directed graph $G = (V, A)$ with nonnegative capacities on the edges, the goal is to find a maximum flow between two specified vertices $s, t \in V$. For simplicity, let us assume that all of the edge capacities are 1. Then it can be seen that that maximum flow problem is equivalent to finding the maximum number of edge disjoint paths from s to t. Let us formulate this problem as a linear program:

$$\begin{aligned} \max \sum_{su \in A} x_{su} & \\ \text{subject to: } \sum_{uv \in A} x_{uv} - \sum_{vw \in A} x_{vw} &= 0, \quad \forall v \neq s, t, \\ x_{uv} &\leq 1, \quad \forall uv \in A, \\ x_{uv} &\geq 0. \quad \forall uv \in A. \end{aligned} \qquad (\mathcal{P}_{max-flow})$$

To find the dual, we use variables p_v for $v \neq s, t$. For the second set of constraints, we use the variables y_{uv}.

$$\begin{aligned} \min \sum_{uv \in A} y_{uv} & \\ \text{subject to: } y_{uv} + p_v - p_u &\geq 0, \quad u, v \neq s, t, \\ y_{sv} + p_v &\geq 1, \\ y_{vt} - p_u &\geq 0, \\ y_{uv} &\geq 0. \end{aligned}$$

This linear program may look familiar! Let us re-arrange it:

$$\begin{aligned} \min \sum_{uv \in A} & y_{uv} \\ \text{subject to:} \quad y_{uv} &\geq \ p_u - p_v, \quad u, v \neq s, t, \\ y_{sv} &\geq \ 1 - p_v, \\ y_{vt} &\geq \ p_u, \\ y_{uv} &\geq \ 0. \end{aligned} \tag{$\mathcal{D}_{min-cut}$}$$

This is the same linear program that we saw in Lecture 1. We saw that it has an integral optimal value equal to the minimum s-t-cut. (Technically, here we have a directed graph, but the argument for an optimal integral solution is the same as shown in Lecture 1.)

Note that if ($\mathcal{P}_{max-flow}$) were integer, then max-flow equals min-cut would follow immediately from Strong Duality. However, we have not yet proved this.

Lemma 3. *In a given graph $G = (V, A)$ with unit edge capacities, the value of the maximum flow from s to t equals the value of the minimum s-t cut.*

Proof. Let k_{max} denote the value of the maximum flow from s to t in G, and let k_{min} denote the value of the minimum s-t-cut in G.

First, we show that $k_{min} \leq k_{max}$. Since G has unit capacities, k_{min} is clearly an integer and by Menger's Theorem, G has at least k_{min} edge disjoint s-t paths. Since each of these paths can assume one unit of flow, we have $k_{min} \leq k_{max}$.

Now, we want to show that $k_{max} \leq k_{min}$. The value of ($\mathcal{P}_{max-flow}$) on G is at least k_{max}. By weak duality, the value of ($\mathcal{P}_{max-flow}$) is at most the value of ($\mathcal{D}_{min-cut}$), which is at most k_{min}. □

Often, the dual variables have some interpretation offering a new a point of view. For instance, in ($\mathcal{D}_{min-cut}$), if the variables had integral values, then the variables p_v would indicate to which side of the minimum cut vertex v belongs, and the variable y_{uv} would indicate whether or not the edge uv is cut. We will see other such interpretations when we use duality in approximation algorithms.

References

[BT97] Dimitris Bertsimas and John N. Tsitsiklis. *Introduction to linear optimization*. Athena Scientific Belmont, MA, 1997.

[GKT51] David Gale, Harold W Kuhn, and Albert W Tucker. Linear programming and the theory of games. *Activity analysis of production and allocation*, 13:317–335, 1951.

[Kha80] Leonid G. Khachiyan. Polynomial algorithms in linear programming. *USSR Computational Mathematics and Mathematical Physics*, 20(1):53–72, 1980.

[Van01] Robert J. Vanderbei. *Linear programming: Foundations and Extensions*. Kluwer Academic Publishers, Boston, MA, 2001.

[vN63] John von Neumann. Discussion of a maximum problem. In A.H. Taub, editor, *John von Neumann, Collected Works: Theory of Games, Astrophusics, Hydrodynamics and Meteorology*, volume VI, pages 89–95. Pergamon, New York, 1963.

These lecture notes are based in part on the following sources: lecture notes by Stéphan Thomassé from previous versions of the same course, Chapter 4 of [BT97], Chapter 5 of [Van01] and lecture notes from a course on Approximation Algorithms at EPFL (`http://theory.epfl.ch/osven/courses/Approx13`).