Advanced Heuristic and Approximation Algorithms	Grenoble-INP (ORCO)

Lecture 2

Lecturer: Alantha Newman

October 8, 2018

1 Linear Programming Duality

For a minimization linear program, any feasible solution provides an upper bound on the optimal objective value of the linear program. 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. For example, 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 optimal solution for a linear program, and sometimes a lower bound suffices.

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

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

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

$$\begin{array}{rcl} 7x + 3y & \geq & x + y \ \geq \ 2, \\ 7x + 3y & \geq & 3x + y \ \geq \ 4, \\ 7x + 3y & \geq & 4x + 2y \ \geq \ 6. \end{array}$$

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:

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

We would like to compute an upper bound on the value z 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:

$$\begin{aligned} y_1 + 5y_2 - y_3 &\geq 4 \\ -y_1 + y_2 + 2y_3 &\geq 1 \\ -y_1 + 3y_2 + 3y_3 &\geq 5 \\ 3y_1 + 8y_2 - 5y_3 &\geq 3, \end{aligned}$$

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:

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

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:

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

The dual linear program is obtained as follows:

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

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.

2 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. These solutions prove that both $(\mathcal{P}_1)$ and $(\mathcal{D}_1)$ have optimal value 29, and one can easily check that these solutions have these objective values.

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$\cap$co-NP well before these complexity classes were defined! Many people believe that belonging to NP$\cap$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].

3 Weak and Strong Duality

The Weak Duality Theorem states that any feasible solution to the dual linear program 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.

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

□

We now 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, then the dual is infeasible.*

3.1 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 the primal $(\mathcal{P})$ iff bounded optimal value for the dual $(\mathcal{D})$.

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

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

3.2 Complementary Slackness

Recall our canonical linear programs, where $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{c} \in \mathbb{R}^n$ are given, and $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{y} \in \mathbb{R}^m$ are the vectors of unknown variables.

$$\begin{aligned} \min\ & \mathbf{c}^\mathsf{T}\mathbf{x} \\ \mathbf{A}\mathbf{x} &\geq \mathbf{b} \\ \mathbf{x} &\geq 0 \end{aligned} \tag{$\mathcal{P}$}$$

$$\begin{aligned} \max\ & \mathbf{b}^\mathsf{T}\mathbf{y} \\ \mathbf{A}^\mathsf{T}\mathbf{y} &\leq \mathbf{c} \\ \mathbf{y} &\geq 0 \end{aligned} \tag{$\mathcal{D}$}$$

Strong duality ensures that if both $(\mathcal{P})$ and $(\mathcal{D})$ have finite optima, they are equal. Assume this is the case. Then, for the optima $\mathbf{x}^*$ and $\mathbf{y}^*$ of the primal and dual programs, respectively, we have

$$\mathbf{c}^\mathsf{T}\mathbf{x}^* = \sum_{i=1}^n c_i x_i^* \geq \sum_{i=1}^n (\sum_{j=1}^m A_{ji} y_j^*) x_i^* = \sum_{j=1}^m (\sum_{i=1}^n A_{ji} x_i^*) y_j^* \geq \sum_{j=1}^m b_j y_j^* = \mathbf{b}^\mathsf{T}\mathbf{y}.$$

Since $\mathbf{c}^\mathsf{T}\mathbf{x}^* = \mathbf{b}^\mathsf{T}\mathbf{y}^*$, the first and second inequalities should hold with equality. This implies that:

$$\text{Primal complementary slackness:} \qquad \text{Either } \sum_{j=1}^m A_{ji} y_j^* = c_i \text{ or } x_i^* = 0.$$

$$\text{Dual complementary slackness:} \qquad \text{Either } \sum_{i=1}^n A_{ji} x_i^* = b_j \text{ or } y_j^* = 0.$$

These conditions are together terms as the (full) *complementary slackness* conditions. They state that either $x_i^* = 0$ or the corresponding dual constraint is tight (or both). Similiarly, either $y_j^* = 0$ or the corresponding primal constraint is tight (or both). In other words, if $x_i^* > 0$, then the corresponding dual constraint is tight. Similarly, if $y^* > 0$, then the corresponding primal constraint is tight. If a dual constraint is *not* tight, then the corresponding primal variable x_i^* must be zero. The same holds for a primal constraint and the respective dual variable.

A nice application of complementary slackness is to derive an optimal dual solution from an optimal primal solution without explicitly solving the dual (e.g., using the simplex algorithm). Let us consider the primal and dual linear programs that we saw previously.

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

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

An optimal solution for $(\mathcal{P}_1)$ is $(0, 14, 0, 5)$. This means that the following dual constraints must be tight.

$$\begin{aligned} -y_1 + y_2 + 2y_3 &= 1 \\ 3y_1 + 8y_2 - 5y_3 &= 3. \end{aligned}$$

From these two equations, we can eliminate two variables.

$$\begin{aligned} y_1 &= 11 - 21y_2, \\ y_3 &= 6 - 11y_2. \end{aligned}$$

Moreover, we note that the second constraint in $(\mathcal{P}_1)$ is not tight. Thus, $y_2 = 0$. From this, we conclude that $(11, 0, 6)$ is an optimal dual solution.

4 König's Theorem for Bipartite Graphs

Strong duality can be used to prove the following theorem.

Theorem 3. (König's Theorem) *For a bipartite graph $G = (V, E)$,*

$$\max\{|M| : M \textit{ is a matching of } G\} = \min\{|S| : S \textit{ is a vertex cover of } G\}.$$

Let $G = (V, E)$ be a bipartite graph and let $V = A \cup B$, where A and B are the two sides of the bipartition of V. Consider the following linear programming relaxation for the matching problem.

$$\begin{aligned} \max \sum_{ij \in E} x_i & \\ \text{subject to:} \quad \sum_{j: ij \in E} x_{ij} \leq 1, & \quad \text{for all } i \in A, \\ \text{subject to:} \quad \sum_{i: ij \in E} x_{ij} \leq 1, & \quad \text{for all } j \in B, \\ x_{ij} \geq 0. & \end{aligned} \tag{LP_M}$$

Now we will take the dual. We use a_i variables for the first set of constraints and b_j variables for the second set of constraints.

$$\begin{aligned} \min \sum_{i \in A} a_i + \sum_{j \in B} b_j & \\ \text{subject to:} \quad a_i + b_j \geq 1, & \quad \text{for all } ij \in E, \\ a_i, b_j \geq 0. & \end{aligned} \tag{LP_{VC}}$$

Note that (LP_{VC}) is a relaxation of the vertex cover problem. Let $M(G)$ denote the size of the maximum matching in G and let $VC(G)$ denote the size of the minimum vertex cover in G. Let $M_{LP(G)}$ and $VC_{LP}(G)$ denote the optimal values of (LP_M) and (LP_{VC}), respectively. By strong duality, we have

$$M(G) \leq M_{LP}(G) = VC_{LP}(G) \leq VC(G).$$

In Lecture 1, we showed that the (LP_M) is an integer polytope (i.e., its extreme points are integral). Therefore, to prove König's Theorem, it suffices to show that (LP_{VC}) is also an integer polytope on bipartite graphs.

Lemma 4. *An optimal solution of (LP_{VC}) is integral.*

Proof. Let $\{a_i^*, b_j^*\}$ denote an optimal solution for (LP_{VC}). Choose this solution so that it maximizes the number of vertices in $A \cup B$ whose values are integral (i.e., either 0 or 1). Let $F \subset A \cup B$ be the set of vertices whose a_i^* or b_j^* values are fractional (i.e., in the open interval $(0, 1)$). Note that if $F = \emptyset$, then we are are done. So assume that $|F| \geq 1$.

Without loss of generality, we assume that $|F \cap A| \geq |F \cap B|$. Let $\epsilon = \min\{a_i^* : i \in F \cap A\}$. We then subtract ϵ to all $a_i^* \in F \cap A$ and add ϵ to all $b_j^* \in F \cap B$. Note that this strictly increases the number of vertices whose values are integral, which is a contradiction (i.e., we can conclude that in fact $F = \emptyset$). It remains to show that this new solution is still feasible. This follows from analyzing the following cases for an edge $ij \in E$.

1. $i \in F \cap A, j \in F \cap B$: a_i^* is decreased and b_j^* is increased the same amount, so the solution remains feasible.
2. $i \in A \setminus F, j \in B \setminus F$: neither endpoint changes value.
3. $i \in A \setminus F, j \in F \cap B$: a_i^* does not change and the value of b_j^* is increased, so the solution remains feasible.
4. $i \in F \cap A, j \in B \setminus F$: the value of a_i^* is decreased, but the value of b_j^* must already be 1 (since the previous solution would not be feasible if it were 0).

□

Lemma 4 can be extended to weighted graphs; we increase the values in $F \cap A$ and decrease those in $F \cap B$ if $w(F \cap A) \geq w(F \cap B)$.

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: Chapter 4 of [BT97], Chapter 5 of [Van01], lecture notes from a course on Approximation Algorithms at EPFL (`http://theory.epfl.ch/osven/courses/Approx13`) and lecture notes from a course on linear and semidefinite programming at CMU (`https://www.cs.cmu.edu/afs/cs.cmu.edu/academic/class/15859-f11/www/notes/lpsdp.pdf`).