CSC 373: Algorithm Design and Analysis
Lecture 21

Allan Borodin

March 11, 2013

Announcements and Outline

Announcements

- I hope to have the term tests available at the tutorials today. Many people had trouble with the NP material and that will be what is discussed today in tutorial.
- Statue of limitations for requests for regrading. We will only entertain requests for regrading (outside of clerical errors) for 2 weeks after the work is returned to the class. There are many assignments and tests that have not been collected.

Today's outline

- Clarifying some remarks on the integrailty gap
- The makespan problem
- Max cut as an IP and a quadratic program
- Duality

Review: the integrality gap for vertex cover

- For the complete (unweighted) graph on n nodes, the optimal IP value is $n-1$, whereas the LP optimum value is $n/2$.
- For a given IP/LP formulation of a minimization problem, the integrality gap is the worst case (over all input instances) ratio of an integral optimum value to a fractional (LP) optimum value.
- The approximation analysis show that this ratio is at most 2 and the previous example shows that it is at least $(2-\frac{1}{n})$.
- There are many ways to "tighten" the IP formulation. But for any known polynomial time approach for adding additional constraints, the integrality gap essentially remains at 2.

Informal but commonly used claim

The integrality gap provides a limit to obtaining an approximation using a particular IP/LP formulation of a problem.

Reflecting on the integrality gap

Informal but commonly used claim

The integrality gap provides a limit to obtaining an approximation using a particular IP/LP formulation of a problem.

- For the n node cycle, the optimum IP solution is $\lceil n/2 \rceil$ and the LP OPT is $n/2$. Note that for n odd, an (optimal) LP solver would not return an integral solution and naive rouding would double the cost of the resulting integral solutuon.
- A naive rounding would then be (at best) a $2 - o(1)$-approximation.
- Note that the integrality gap for a particular IP/LP formulation does not depend on the method of rounding but still in practice it has been a limitation on the approximation ratios obtained.
- It is NP hard to obtain an approximation better than ≈ 1.38 and there is a complexity conjecture (much stronger than $P \neq NP$) that implies that it is not possible to obtain a $2 - \epsilon$ approximation in polynomial time (for any $\epsilon > 0$). .

Makespan for the unrelated machines model

The makespan problem for the unrelated machines model

- The input consists of n jobs $\mathcal{J} = \{J_1, \ldots, J_n\}$ and m machines $M_1, \ldots, M_m$.
- Each job J_j is represented by a vector $\langle p_{1j}, p_{2j}, \ldots, p_{mj} \rangle$ where p_{ij} represents the processing time of job J_j on machine i.
- Without loss of generality, we assume $m \leq n$.
- **Goal:** Minimize the latest finishing time (maximum load) over all machines.

- We will sketch a 2-approximation IP/LP with non naïve rounding algorithm.
- This is the best known poly-time approximation.
- It is known that it is NP-hard to achieve better than $\frac{3}{2}$-approximation even for the special case of the restrictive machines model for which every p_{ij} is either some p_j or ∞.

- In the IP formulation, the problem is:

$$
\begin{aligned}
\textbf{minimize } & t \\
\text{s.t. } & \sum_{i=1}^{m} x_{ij} = 1 && \text{for each job } J_j \\
& \sum_{j=1}^{n} p_{ij} x_{ij} \le t && \text{for each machine } M_i \\
& x_{ij} \in \{0,1\} && \text{for } i = 1, \ldots, m \text{ and } j = 1, \ldots, n
\end{aligned}
$$

- The intended meaning is that $x_{ij} = 1$ iff job J_j is scheduled on machine M_i.
- The LP relaxation is that $0 \le x_{ij}$. The condition $x_{ij} \le 1$ is implied.
- The integrality gap is unbounded! Why?. How do we get around this unbounded integrality gap?

Getting around the integrality gap

- Consider one job with procesing time m, which has $OPT = m$ and $OPT_{LP} = 1$.
- The IP must set $x_{ij} = 0$ if $p_{ij} > t$ whereas the fractional OPT does not have this constraint.
- We want to say for all (i, j): "if $p_{ij} > t$ then $x_{ij} = 0$". But this isn't a linear constraint!
- Since we are only hoping for a good approximation, we can assume all p_{ij} are integral.
- We can then use binary search to find the best LP bound T by solving the search problem $LP(T)$ which eliminates the objective function (i.e. setting it to a constant 1) and removes any x_{ij} having $p_{ij} > T$.
- We clearly have that $OPT_{IP} \geq T$.

Rounding of $LP(T)$ solution

- The rounding here is not naïve. In general, rounding just means converting a rational solution (for an LP relaxation) to some integral solution. This might be done in stages, say first obtaining some integral non-solution and then adjusting to a feasible solution.
- The $LP(T)$ solution x^*_{ij} is the solution of a system of $m+n$ equations over the mn variables x_{ij}.
- LP Theory tells us that when this system has a solution, there is a (so-called basic) solution x^*_{ij} with at most $m+n$ positive values.
- This implies by counting that there are at most m fractional (not integral) values.
- If $x^*_{ij}=1$ then we assign job j to machine i. The remaining part of the proof (using more LP theory) is to show that there is a matching between the fractional x^*_{ij} and the machines.

A non obvious IP representation

- Consider the Max Cut problem.

- We can think of a solution as a choice about which vertices to (say) put into A in an (A, B) cut.

- We could have variables $y_i \in \{+1, -1\}$ with the intended meaning $y_i = 1$ (resp -1) iff vertex $v_i \in A$ (resp. B).

- Then we would want to
  - maximize $\sum_{1 \leq i < j \leq n} \frac{1}{2} w(i,j)(1 - y_i y_j)$
  - subject to $y_i \in \{+1, -1\}$ (i.e. $y_i^2 = 1$)

Problem

While this is a very useful quadratic program (and a .878 approx using a vector program relaxation), it is NOT a linear program.

Max cut as an IP

- Instead we will think of a cut as the edges crossing the $cut(A, B)$
- And have a variable $x_e \in \{0, 1\}$ for every edge $e = (u, v)$ with the intended meaning that $x_e = 1$ iff (u, v) is in the cut.
- Now we need to find inequalities that ensure the set $\{x_e \mid x_e = 1\}$ defines a cut.
- This isn't at all obvious but here is what works.

$$
\begin{aligned}
\text{maximize } & \sum_{e \in E} w_e x_e \\
\text{subject to } & x_e \in \{0,1\} && \text{for every } e \in E \\
& x_{ij} + x_{jk} \geq x_{ik} && \text{for every triangle } (v_i, v_j, v_k) \\
& x_{ij} + x_{ik} + x_{jk} \leq 2
\end{aligned}
$$

Why does this work?

$$
\begin{aligned}
\text{maximize} \quad & \sum_{e \in E} w_e x_e & \\
\text{subject to} \quad & x_e \in \{0,1\} & \text{for every } e \in E \\
& x_{ij} + x_{jk} \geq x_{ik} & \text{for every triangle } (v_i, v_j, v_k) \\
& x_{ij} + x_{ik} + x_{jk} \leq 2 &
\end{aligned}
$$

- You can think of these "triangle inequalities" as saying that the possible sizes of a cut for each triangle are 0 or 2.
- Clearly every cut must satisfy these constraints!

Why does this work? (continued)

$$
\begin{aligned}
\text{maximize} \quad & \sum_{e \in E} w_e x_e \\
\text{subject to} \quad & x_e \in \{0,1\} && \text{for every } e \in E \\
& x_{ij} + x_{jk} \geq x_{ik} && \text{for every triangle } (v_i, v_j, v_k) \\
& x_{ij} + x_{ik} + x_{jk} \leq 2
\end{aligned}
$$

Claim

Every $\{0,1\}$ solution of this IP defines a cut.

- Define a relation $i \sim j$ if $x_{ij} = 0$ or $i = j$.
- **Claim 1:** This is an equivalence relation.
  - Transitivity is the only thing to check.
  - By the triangle condition $x_{ij} = x_{ik} = 0$ implies $x_{jk} = 0$.
- **Claim 2:** There are at most 2 equivalence classes.
  - This follows from the second triangle condition.
  - If i, j, k are in three different classes, then $x_{ij} + x_{ik} + x_{jk} = 3$.
- Hence, the equivalence classes are the cut.

Duality

NOTE:

We will not have time to sufficiently consider duality so that I will not be testing on the following material. But I am leaving the slides for your information as this is an important topic.

- For a primal maximization (resp. minimization) LP in standard form, the dual LP is a minimization (resp. maximization) LP in standard form.
- Specifically, if the primal $\mathcal{P}$ (with n variables and m constraints) is:
  - Minimize $\mathbf{c} \cdot \mathbf{x}$
  - subject to $A \cdot \mathbf{x} \geq \mathbf{b}$
  - $\mathbf{x} \geq 0$
- Then the dual LP $\mathcal{D}$ with m dual variables $\mathbf{y}$ and n constraints is:
  - Maximize $\mathbf{b} \cdot \mathbf{y}$
  - subject to $A^{tr} \cdot \mathbf{y} \leq \mathbf{c}$
  - $\mathbf{y} \geq 0$
- Note that the dual (resp. primal) variables are in 1-1 correspondence to primal (resp. dual) constraints.
- If we consider the dual $\mathcal{D}$ as the primal then its dual is the original

The intuitive idea of duality

Consider the following discussion in V. Vazirani's text:

$$\begin{aligned} \text{minimize} \quad & 7x_1 + x_2 + 5x_3 \\ \text{subject to} \quad & x_1 - x_2 + 3x_3 \geq 10 \\ & 5x_1 + 2x_2 - x_3 \geq 6 \\ & x_1, x_2, x_3 \geq 0 \end{aligned}$$

- It is obvious how to verify (with a "small" certificate being a solution) that $OPT \leq U$ for some given U
- But how can we verify that $OPT \geq L$ for some given L ?
  1. $7x_1 + x_2 + 5x_3 \geq x_1 - x_2 + 3x_3 \geq 10$
  2. Better bound (looking at each coef of x_i):

     $$7x_1 + x_2 + 5x_3 \geq (x_1 - x_2 + 3x_3) + (5x_1 + 2x_2 - x_3) \geq 10 + 6 = 16$$

  3. And even better bound:

     $$7x_1 + x_2 + 5x_3 \geq 2(x_1 - x_2 + 3x_3) + (5x_1 + 2x_2 - x_3) \geq 26$$

What we have learned about this LP

- We just showed that

$$7x_1 + x_2 + 5x_3 \geq 2(x_1 - x_2 + 3x_3) + (5x_1 + 2x_2 - x_3) \geq 26$$

- And setting (x_1, x_2, x_3) to $(\frac{7}{4}, 0, \frac{11}{4})$ we see that the primal is

$$\leq 7 \cdot \left(\frac{7}{4}\right) + 1(0) + 5 \cdot \left(\frac{11}{4}\right) = \frac{104}{4} = 26$$

- Hence the primal and the dual $= 26$.
- The idea was to find the best choice of non-negative multipliers y_1 and y_2 so that

$$7x_1 + x_2 + 5x_3 \geq y_1(x_1 - x_2 + 3x_3) + y_2(5x_1 + 2x_2 - x_3)$$

- The original LP:

$$\begin{aligned} \text{minimize} \quad & 7x_1 + x_2 + 5x_3 \\ \text{subject to} \quad & x_1 - x_2 + 3x_3 \geq 10 \\ & 5x_1 + 2x_2 - x_3 \geq 6 \\ & x_1, x_2, x_3 \geq 0 \end{aligned}$$

- We have shown that solving this problem is the same as finding the best choice of non-negative multipliers y_1 and y_2 so that

$$7x_1 + x_2 + 5x_3 \geq y_1(x_1 - x_2 + 3x_3) + y_2(5x_1 + 2x_2 - x_3) \geq y_1(10) + y_2(6)$$

- This leads to the dual problem:

$$\begin{aligned} \text{maximize} \quad & 10y_1 + 6y_2 \\ \text{subject to} \quad & y_1 + 5y_2 \leq 7 \\ & -y_1 + 2y_2 \leq 1 \\ & 3y_1 - 2y_2 \leq 5 \end{aligned}$$

Duality Theory Overview

- An essential aspect of duality is that a finite optimal value to either the primal or the dual determines an optimal value to both.
- The relation between these two can sometimes be easy to interpret. However, the interpretation of the dual may not always be intuitively meaningful.
- Still, duality is very useful because the duality principle states that optimization problems may be viewed from either of two perspectives and this might be useful as the solution of the dual might be much easier to calculate than the solution of the primal.
- Moreover, the relation between the primal $\mathcal{P}$ and the dual $\mathcal{D}$ will lead to Primal-Dual algorithms and to dual fitting analysis.
- In what follows we will assume the primal is a minimization problem to simplify the exposition.

Strong and Weak Duality

Strong Duality

If x^* and y^* are (finite) optimal primal and resp. dual solutions, then $\mathcal{D}(\mathbf{y}^*) = \mathcal{P}(\mathbf{x}^*)$.

Note

Before it was known that solving LPs was in polynomial time, it was observed that strong duality proves that LP (as a decision problem) is in NP $\cap$ co-NP which strongly suggested that LP was not NP-complete.

Weak Duality

If $\mathbf{x}$ and $\mathbf{y}$ are primal and dual solutions respectively, then $\mathcal{D}(\mathbf{y}) \leq \mathcal{P}(\mathbf{x})$.

Duality was motivated by asking how one can verify that the minimum in the primal is at least some value z.

- To get witnesses, one can explore non-negative scaling factors (i.e. the dual variables) that can be used as multipliers in the constraints.
- The multipliers, however, must not violate the objective (i.e cause any multiplies of a primal variable to exceed the coefficient in the objective) we are trying to bound.

Max flow-min Cut in terms of duality

- While the max flow problem can be naturally formulated as an LP, the natural formulation for min cut is as an IP.
  - However, for this IP, it can be shown that the *extreme point solutions* (i.e. the vertices of the polyhedron defined by the constraints) are all integral $\{0,1\}$ in each coordinate.
  - Moreover (see Vazarani, section 12.2) max flow and min cut can be viewed as dual problems.
- Suppose we formulate max flow in standard LP form so that all flows (i.e. the LP variables) are non-negative.
- And to state the objective as a simple linear function (of the flows) we add an edge of infinite capacity from the terminal t to the source s and hence define a circulation problem.

The max flow LP

maximize $f_{t,s}$

subject to

$$\begin{aligned} & f_{i,j} \leq c_{i,j} && \text{for all } (i,j) \in E \\ & \textstyle\sum_{j:(j,i)\in E} f_{j,i} - \sum_{j:(i,j)\in E} f_{i,j} \leq 0 && \text{for all } i \in V \\ & f_{i,j} \geq 0 && \text{for all } (i,j) \in E \end{aligned}$$

Max flow-min cut duality continued

For the primal edge capacity constraints, introduce dual ("distance") variables $d_{i,j}$ and for the vertex flow conservation constraints, introduce dual ("potential") variables p_i.

The fractional min cut dual

minimize $\sum_{(i,j)\in E} c_{i,j} d_{i,j}$

subject to
$$d_{i,j} - p_i + p_j \geq 0$$
$$p_s - p_t \geq 1$$
$$d_{i,j} \geq 0; p_i \geq 0$$

- Now consider the IP restriction : $d_{i,j}, p_i \in \{0,1\}$ and let $\{(d^*_{i,j}, p^*_i)\}$ be an intergal optimum.
- The $\{0,1\}$ restriction and second constraint forces $p^*_s = 1$ and $p^*_t = 0$.
- The IP optimum defines a cut (S,T) with $S = \{i \mid p^*_i = 1\}$ and $T = \{i \mid p^*_i = 0\}$.
- Suppose (i,j) is in the cut, then $p^*_i = 1, p^*_j = 0$ which by the first constraint forces $d_{i,j} = 1$.
- The optimal $\{0,1\}$ IP solution (of the dual) defines a a min cut.

Another example: the dual of set cover

The set cover problem as an IP/LP

minimize $\sum_j w_j x_j$
subject to $\sum_{j:e_i \in S_j} \geq 1$ for all i
$x_j \in \{0,1\}$ for all j (for LP we use $x_j \geq 0$)

The dual LP

maximize $\sum_i y_i$
subject to $\sum_{i:e_i \in S_j} y_i \leq w_j$ for all j
$y_i \geq 0$ for all i

- If all the parameters in a standard form minimization (resp. maximization) problem are non negative, then the problem is called a covering (resp. packing) problem.
- Note that the set cover problem is a covering problem and its dual is a packing problem.

Solving the f-frequency set cover by a primal dual algorithm

- In the f-frequency set cover problem, each element is contained in at most f sets.
- Clearly, the vertex cover problem is an instance of the 2-frequency set cover.
- As in the vertex cover LP rounding, we can similarly solve the f-frequency cover problem by obtaining an optimal solution $\{x_j^*\}$ to the (primal) LP and then rounding to obtain $\bar{x}_j = 1$ iff $x_j^* \geq \frac{1}{f}$.
  - This is, as noted before, a conceptually simple method but requires solving the LP.
- We know that for a minimization problem, any dual solution is a lower bound on any primal solution.
  - One possible goal in a primal dual method for a minimization problem will be to maintain a fractional feasible dual solution and continue to try improve the dual solution.
  - As dual constraints become tight we then set the corresponding primal variables.

Primal dual for f-frequency set cover continued

Suggestive lemma

Let $\{y_i^*\}$ be an optimal solution to the dual LP and let $\mathcal{C}' = \{S_j \mid \sum_{e_i \in S_j} y_i^* = w_j\}$. Then $\mathcal{C}'$ is a cover.

This suggests the following algorithm:

Primal dual algorithm for set cover

Set $y_i = 0$ for all i
$\mathcal{C}' := \varnothing$
While there exists an e_i not covered by $\mathcal{C}'$
    Increase the dual variable y_i until there is some j :
        $e_i \in S_j$ and $\sum_{\{k: e_i \in S_j\}} y_j = w_j$
    $\mathcal{C}' := \mathcal{C}' \cup \{S_j\}$
End While

Theorem (Approximation bound for f-frequency set cover)

The primal dual solution is an f approximation.

Comments on the primal dual algorithm

- What can be shown is that the integral primal solution is within a factor of f of the dual solution which implies the theorem that the primal dual algorithm is an f-approximation algorithm for the f-frequency set cover problem.

- In fact, what is being shown is that the integraility gap of this IP/LP formulation for f-frequency set cover problem is at most f.

- In terms of implementation we would calculate the minimum ϵ needed to make some constraint tight so as to chose which primal variable to set.
  - This ϵ could be 0 if a previous iteration had more than one constraint that becomes tight simultaneously.
  - This ϵ would then be subtracted from w_j for j such that $e_i \in S_j$.

More comments on primal dual algorithms

- We have just seen an example of a basic form of the primal dual method for a minimization problem.
  - Namely, we start with an infeasible integral primal solution and feasible (fractional) dual.
  - (For a covering primal problem and dual packing problem, the initial dual solution can be the all zero solution.)
  - Unsatisfied primal constraints suggest which dual constraints might be tightened and when one or more dual constraints become tight this determines which primal variable(s) to set.

- Some primal dual algorithms extend this basic form by using a second (reverse delete) stage to achieve minimality.

Note

In the primal dual method we are not solving any LPs. Primal dual algorithms are viewed as "combinatorial algorithms" and in some cases they might even suggest an explicit (and efficient) greedy algorithm.