Lecture 5: Iterated Rounding

6th Feb, 2015

1 Some facts about Linear Programs

A minimization linear program, in its generality, can be denoted as

$$\begin{aligned} \min \quad & c \cdot x \\ \text{subject to} \quad & Ax \geq b \end{aligned} \tag{1}$$

The number of variables is the number of columns of A, the number of constraints is the number of rows of A. The ith column of A is denoted as A_i, the ith row as a_i. Note that constraints of A could also include constraints of the form $x_i \geq 0$. These non-negativity constraints are often considered separately (for a reason). Let's call the other constraints *non-trivial* constraints. For the discussion of this section, let n be the number of variables and let m denote the number of non-trivial constraints. Thus the number of rows in A are $(m+n)$ and $\texttt{rank}(A) = n$ (since it contains an $I_{n\times n}$ as a submatrix).

A solution x is a feasible solution if it satisfies all the constraints. Some constraints may be satisfied with equality, others as a strict inequality. A feasible solution x is called *basic* if the constraints satisfied with equality span the entire space. Let B be the equality rows of A such that $Bx = b_B$, where b_B are the entries of b corresponding to the rows of B. Note that B depends on x. x is a basic feasible solution if $\texttt{rank}(B) = n$.

Fact 1. *There is an optimal solution to every LP which is a basic feasible solution.*

Let x be a basic feasible solution, and let $\texttt{supp}(x) := \{i : x_i > 0\}$. Let B be the basis such that $Bx = b_B$. Note that B contains at most $n - \texttt{supp}(x)$ non-negativity constraints. Therefore, at least $\texttt{supp}(x)$ linearly independent non-trivial constraints which are satisfied with equality. This fact will be used crucially a few classes later.

Fact 2. *There is an optimal solution to every LP with at least* $\texttt{supp}(x)$ *linearly independent non-trivial constraints satisfied with equality.*

A linear program can be solved in polynomial (in the number of variables) time by using the *ellipsoid algorithm* as long as there is a polynomial time *separation oracle*. A separation oracle is the following: given a point x either says it is feasible or returns a violated inequality $a_i \cdot x < b_i$.

Fact 3. *An optimum basic feasible solution to a LP can be found in polynomial time as long as there is a separation oracle.*

2 Iterated Rounding Framework

Recall the LP-methodology for designing approximation algorithms. Write a linear integer program capturing the optimum of the problem, solve the linear programming relaxation, and round the possibly fractional solution of the LP to an integer solution. In the last lecture, we solved the linear program and spent most of out energy rounding the solution thus obtained into a fully integral solution. In iterated rounding, as the name suggests, the rounding of fractional variables is done in iterations. A quick description of the schema is as follows. In each iteration, a set of variables are made integral; this leads rise to a *residual* problem, and more often than not, the residual problem is an instance of the same class of problems. One then considers the LP relaxation on this residual instance, and if one can argue that the drop in the LP-value is at least $\frac{1}{\rho}$ times the cost of the rounded variables (in case of minimization problems), then one gets a ρ-approximation.

Matchings in Bipartite Graphs Consider the min-cost perfect matching problem in bipartite graphs. Given a bipartite graph $G = (A, B, E)$ with costs on edges, we need to find a perfect matching of minimum cost. An LP relaxation for the problem is as follows

$$\begin{aligned} \min \quad & \sum_{e \in E} c_e x_e & \\ \text{subject to} \quad & x(\delta(a)) = 1 & \forall a \in A \\ & x(\delta(b)) = 1 & \forall b \in B \\ & 0 \le x_e \le 1 & \forall e \in E \end{aligned}$$

Let x be a basic feasible solution to the above LP.

Lemma 1. *There is an edge e with $x^*_e = 0$ or $x^*_e = 1$.*

Proof. Suppose not and this all edges e have $0 < x^*_e < 1$. The constraints of the LP imply that each vertex in A (or B) must have at least two edges incident on it; so $|E| \ge 2n$, where $|A| = |B| = n$. Since x^* is basic feasible, we know that $|E|$ is at most the number of linearly independent equalities. These equalities must correspond to the $2n$ constraints shown above because none of the x^*_e's are 0 or 1. However these are not linearly independent – the sum of the rows corresponding to vertices in A equal the sum of rows corresponding to B. This is a contradiction, implying some x^*_e is either 0 or 1. □

If $x^*_e = 0$, we delete e from the graph. If $x^*_e = 1$ then we pick this in our matching and delete the end points of e from the graph. In both cases we end up with another min-cost perfect matching problem, and hence we can apply iterated rounding to obtain an exact algorithm.

Maximum weight bipartite matching. Now let us consider the problem that I botched up in the class. In fact, we were on the right lines, so let me just complete the argument. Here's the LP again.

$$\begin{array}{lll} \max & \sum_{(i,j)\in E} p_{ij}x_{ij} & \\ \text{subject to} & x(\delta(a)) \leq 1 & \forall a \in A \\ & x(\delta(b)) \leq 1 & \forall b \in B \\ & 0 \leq x_e \leq 1 & \forall e \in E \end{array}$$

The following theorem allows us to run iterated rounding to prove that the integrality gap of the LP is $= 1$.

Theorem 1. *Given any bfs x, there exists an edge (i,j) with $x_{ij} = 0$ or $x_{ij} = 1$.*

Proof. Let F be the set of edges, and suppose $0 < x_{ij} < 1$ for all edges. Let r be the number of tight constraints/vertices on the A-side, and s be the number of constraints on the B-side. From F we throw away the edges (i,j) where neither i nor j are tight – as observed in the class, if $p_{ij} > 0$ then such edges can't exist. Abusing notation, we call the remaining edges F. We divide F into two parts: F_{in} and F_{out}, where F_{in} are the edges (i,j) where **both** i and j are tight, and F_{out} are the remaining which have exactly one end point incident on a tight vertex.

Here is the first observation: if $F_{in} = F$, that is if $F_{out} = \emptyset$, then the number of **linearly independent** tight constraints is $\leq (r + s - 1)$. This is because the rows corresponding to tight constraints on the A-side sum up to tight constraints on the B-side. Thus, we get

$$\text{Number of linearly independent tight constraints} = \begin{cases} \leq r + s - 1 & \text{if } |F_{out}| = 0 \\ \leq r + s & \text{if } |F_{out}| \geq 1 \end{cases} \tag{2}$$

Second observation: for any tight vertex, it's degree must be ≥ 2. This is because the sum of x-value son edges incident on it is 1 while each is < 1 (by supposition). Therefore, the sum of degrees of all tight vertices is $\geq 2(r + s)$. Observe now that the sum of degrees of tight vertices is precisely $2|F_{in}| + |F_{out}| = 2|F| - |F_{out}|$ – the F_{in} edges are counted twice while the F_{out} edges are counted once. Therefore, we get

$$2|F| - |F_{out}| \geq 2(r + s) \quad \Rightarrow \quad |F| \geq r + s + \frac{|F_{out}|}{2} = \begin{cases} > r + s - 1 & \text{if } |F_{out}| = 0 \\ > r + s & \text{if } |F_{out}| > 0 \end{cases}$$

Comparing with (2), we get that no matter if $|F_{out}| = 0$ or not, $|F|$ exceeds the number of linearly independent tight constraints. This is a contradiction, and hence our supposition is wrong, proving the theorem. □

3 The Tree Augmentation problem

Given a rooted tree T and a set of non-tree segments F with costs, the goal is to pick $F' \subseteq F$ of minimum cost such that $T + F'$ is 2-edge connected.

We start with an LP relaxation for the problem. First some notation: given any segment $f = (u, v)$, we let P_f denote the **unique** tree path from u to v.

$$\min \quad \sum_{f \in F} c_f x_f \tag{3}$$

$$\text{subject to} \quad \sum_{f: e \in P_f} x_f \geq 1 \qquad \forall e \in T \tag{4}$$

$$x_f \geq 0 \qquad \forall f \in F \tag{5}$$

We show a 2-approximation for the problem. In fact, there are many (at least 3) ways of showing this – today, we will see a beautiful proof using iterated rounding. In fact, we show the following.

Theorem 2. *Let x be any basic feasible solution to* (3). *Then there must exist an f with $x_f \geq 1/2$.*

The above theorem implies a 2-approximation: solve the LP, pick segment f with $x_f \geq 1/2$. Construct the residual tree which **contracts** all the edges covered by f. Rinse and repeat. We leave the details as an exercise. The proof of the above theorem essentially is due to Kamal Jain, but we show a newer proof by Nagarajan, Ravi and Singh.

Proof. Let us assume $x_f > 0$ for all $f \in F$ by just chucking out all the zero-x-segments. For notational purposes, given $e \in T$, let $F_e := \{f : e \in P_f\}$ denote the segments which can potentially cover e. Next, we contract some edges of the tree. That is, we replace (u, v) by a single node and any segment starting at u or v now start or end at that super node. Note that contraction **doesn't** affect $\sum_{f \in F_e} x_f$ for any non-contracted edge e. First we contract every edge e such that $\sum_{f \in F_e} x_f > 1$. Abusing notation let's call the remaining edges E. Next, we contract any edge e such that the set F_e can be decomposed as disjoing unions of $F_{e_1} \cup F_{e_2} \cup \cdots$ for some other edges $e_1, e_2, \ldots$. Note that the row in the constraint matrix of the LP corresponding to e can be then written as the sum of the rows in the constraint matrix corresponding to $e_1, e_2, \ldots$, and therefore in a full rank system the constraint of e won't appear if those of $e_1, e_2, \ldots$ does. Let E (abusing notation once again) be the remaining rows. Since x was a bfs, $|F| \leq |E|$.

We now show if $x_f < 1/2$ for all $f \in F$, $|F| > |E|$ which is an obvious contradiction. This is done by a charging argument. For every segment $f = (u, v)$, we distribute ≤ 1 unit of charge on the edges as follows: the edge $(u, p(u))$ where $p(u)$ is the parent of u gets charge x_f, the edge $(v, p(v))$ gets a charge x_f. Now let a be the lca of u and v. If a is not the root, then the edge $(a, p(a))$ gets a charge $(1 - 2x_f)$.

Claim 1. *The total charge distributed is $< |F|$.*

Proof. It's clear that the total charge is $\leq F$. To see strict inequality note that the root r has at least one child u. Now look at the tree edge $e = (r, u)$ and consider F_e. For any segment $f \in F_e$, the lca of its endpoints is r, the root. Therefore, it gives out $2x_f < 1$ units of charge since $x_f < 1/2$ for all f. □

The following claim along with the previous one shows $|F| > |E|$ which is the contradiction.

Claim 2. *Every edge in E gets at least one unit of charge.*

Proof. Fix a tree edge (u, v) where $v = p(u)$. Let $u_1, \ldots, u_k$ be the $k \geq 1$ children of u and let $T_1, \ldots, T_k$ be the k subtrees rooted at these vertices. The proof that (u, v) gets one unit of charge follows in two steps: first we show that (u, v) gets > 0 charge, and in the next we show the charge on (u, v) must be an integer.

Suppose $e = (u, v)$ gets 0 charge. Then this implies there is no f with endpoint u. Therefore, all segments in F_e have one of their endpoints in one of the T_i's. Furthermore, there is no segment $f = (a, b)$ where $a \in T_i$ and $b \in T_j$ for $i \neq j$. Since otherwise, e would get charge at least $(1 - 2x_f) > 0$. Thus, all segments in F_{e_i} where $e_i := (u, u_i)$ also lie in F_e. But this contradicts the fact that e wasn't removed. In sum, e gets > 0 charge.

Now, let A be the segments with one endpoint in T_i and another in T_j for some $i \neq j$. Let B be the segments with one end point in T_i and which also lies in F_e. C be the segments in F_e which have an endpoint in u. Finally, D be the segments which have one endpoint in T_i and the other endpoint in u. We know the following equalities

$$x(B) + x(C) = 1$$

since $(u, v) \in E$.

$$x(B) + x(D) + 2x(A) = k$$

since each $(u, u_i) \in E$. Now the total charge on the edge (u, v) equals $|A| - 2x(A) + x(C) + |D| - x(D) = |A| + |D| + 1 - k$, which is an integer. □

□