Rounding Techniques in Approximation Algorithms

Lecture 11: Iterative Rounding for Network Design

Lecturer: Nathan Klein

1 Iterative Rounding

Before we jump into the problem, let's again recall our key fact for this set of lectures: given any LP with n variables, we can find an optimal solution that is the unique solution to a subsystem of n linearly independent constraints met with equality, $\tilde{A}x = \tilde{b}$.

Last time, we used this in a very basic way. We showed that if there are few non-trivial constraints (i.e. constraints that are not of the form $0 \leq x_i$ or $x_i \leq 1$), then there are few fractional variables. In this lecture, our LP will have exponentially many constraints. At first glance, this looks like an issue for this approach, since it feels like surely we can build a large rank collection of constraints from our exponential set.

The key in this lecture will be that the n linearly independent constraints must all be tight at our vertex and linearly independent. We will show that in our problem, in fact, the rank of the set of tight constraints cannot be too large. This turns out to be the rule rather than the exception, and we can often expect the rank of these systems to not be very large.

We will use this to prove that we can iteratively find a large coordinate i in any vertex and round it to 1. More formally:

Iterative Rounding

Consider a covering problem over $\{0,1\}^n$ with (potentially exponentially many) constraints of the form $a^T x \geq b$ for $a \in \mathbb{R}^n_{\geq 0}$, $b \geq 0$ with a separation oracle for the resulting polytope P. Now, prove that for some k, for any vertex x of P, there is some variable i with $x_i \geq \frac{1}{k}$.

Given this fact, we can iteratively round up an element with $x_i \geq \frac{1}{k}$ to 1, and add the constraint $x_i = 1$ to the LP permanently. We then re-solve our LP with this constraint (we know the LP is still feasible since it's a covering problem) and continue until our solution is integral. This leads to a k-approximation as we prove below.

We are continually paying a factor of k to translate a fractional coordinate to an integral one, so intuitively, our approximation ratio shouldn't be worse than k. Nevertheless, we write out the proof formally.

Lemma 1.1. *So long as we can prove there is some variable with $x_i \geq \frac{1}{k}$ in any vertex x for a covering problem, we can obtain a k-approximation.*

Proof. Let x be the initial vertex solution to the LP. We will obtain a sequence of vertex solutions $x^0, \ldots, x^n$ so that $x^0 = x$ and x_n is integral. At each step j, we will constrain all integral coordinates of x^{j-1} to be integral[1] for x^j as well as add a new constraint $x_i = 1$ for some i. Let $I(x)$ be the cost

[1]Technically, this is not necessary and so is not included in the formal description of iterative rounding. However it can only speed up the algorithm and leads to a slightly more general proof, so we do so here.

of the integral coordinates of a vector x and $F(x)$ be the cost of the fractional coordinates. We will prove that at each step $j \geq 1$,

$$I(x^j) + k \cdot F(x^j) \leq I(x^{j-1}) + k \cdot F(x^{j-1})$$

So, since $F(x^n) = 0$, we will have:

$$c(x^n) = I(x^n) \leq I(x^{n-1}) + k \cdot F(x^{n-1}) \leq \cdots \leq I(x^0) + k \cdot F(x^0) \leq k \cdot c(x)$$

As desired. So let's prove the claim. At step j, we are taking a variable $x_i \geq \frac{1}{k}$ and rounding it to 1. Consider the point y which is simply x^{j-1} with the ith coordinate rounded to 1. Then, $I(y) = I(x^{j-1}) + c_i$ and $F(y) = F(x^{j-1}) - \frac{1}{k}c_i$. So,

$$I(y) + k \cdot F(y) \leq I(x^{j-1}) + c_i + k \cdot \left(F(x^{j-1}) - \frac{1}{k}c_i\right) = I(x^{j-1}) + k \cdot F(x^{j-1})$$

However, $c(x^j) \leq c(y)$ since y was a feasible solution to the LP with the added constraint $x_i = 1$ (as we have a packing problem) and x^j is the cheapest solution to the resulting LP. In addition, $I(x^j) \geq I(y)$, implying $F(x^j) \leq F(y)$. So,

$$I(x^j) + k \cdot F(x^j) \leq I(y) + F(y) + (k-1) \cdot F(x^j) \leq I(y) + k \cdot F(y) \leq I(x^{j-1}) + k \cdot F(x^{j-1}),$$

as desired. □

1.1 Survivable Network Design

As input, we are given a graph $G = (V, E)$, weights $c_e \geq 0$ on each edge, and a connectivity requirement $r_{u,v} \in \mathbb{Z}_{\geq 0}$ for each vertex pair u, v. Our goal is to output the cheapest $F \subseteq E$ so that each pair of vertices u, v has connectivity at least $r_{u,v}$. This is called a "survivable" network since each u, v are still connected after an arbitrary $r_{u,v} - 1$ edge deletions.

Let $f : 2^V \to \mathbb{Z}_{\geq 0}$ be defined as $f(S) = \max_{u \in S, v \notin S}\{r_{u,v}\}$. Then, we can write our polytope for this problem as follows:

$$P_f = \begin{cases} x(\delta(S)) \geq f(S) & \forall S \subseteq V \\ 0 \leq x_e \leq 1 & e \in E \end{cases} \tag{1}$$

It turns out that this function f is *skew supermodular*. This means that for all $S, T \subseteq V$, we have either $f(S) + f(T) \leq f(S \cup T) + f(S \cap T)$ or $f(S) + f(T) \leq f(S \setminus T) + f(T \setminus S)$. This is not difficult to prove, so we leave it as an exercise. Remember that for a supermodular function, the first inequality must always hold. Also, note that $f(S) = k$ for any fixed integer k is skew supermodular. In fact, both inequalities always hold with equality. So, this is a generalization of the k-edge-connectivity problem.

The main theorem of this section is as follows, which leads to a 2-approximation for the survivable network design problem. This was first proved by Jain [Jai01].

Theorem 1.2. *Whenever f is a skew supermodular function, every vertex of P_f has a coordinate e with $x_e \geq \frac{1}{2}$.*

Before we prove this theorem, we will give a description of the set of tight constraints at any vertex. Note that a **laminar** family is a collection of sets $\mathcal{L}$ over a ground set V such that for any $S, T \in \mathcal{L}$, we have $S \cap T \in \{S, T, \emptyset\}$ (this is called **intersecting**). Typically, when we are working with cuts, we will fix a vertex r and have r lie outside all of the sets. This makes the "side" of each cut we are using unambiguous. Two cuts are **crossing** if $S \cap T \neq \emptyset$, $S \cup T \neq V$, $S \setminus T \neq \emptyset$ and $T \setminus S \neq \emptyset$. Excluding a root r is convenient because if S, T intersect and both do not contain r, then they cross.

We will need to use a particular property of the cut function.

Lemma 1.3. *Suppose S, T cross. Then,*

$$\chi(\delta(S)) + \chi(\delta(T)) = \chi(\delta(S \cup T)) + \chi(\delta(S \cap T)) + 2E(S \setminus T, T \setminus S)$$

$$\chi(\delta(S)) + \chi(\delta(T)) = \chi(\delta(S \setminus T)) + \chi(\delta(T \setminus S)) + 2E(S \cap T, V \setminus (S \cup T))$$

You can convince yourself of this with a picture! Now we can prove our structural lemma, which is typically known as *uncrossing*:

Lemma 1.4 (Uncrossing)**.** *Let f be skew supermodular, and let x be a vertex of P_f such that $0 < x_e < 1$ for all $e \in E$. Then, there is a laminar family $\mathcal{L} \subseteq 2^{V \setminus \{r\}}$ (where we exclude an arbitrary root vertex r) such that:*

1. *$x(\delta(S)) = f(S)$ for all $S \in \mathcal{L}$,*
2. *$\{\chi(\delta(S)) \mid S \in \mathcal{L}\}$ spans all tight constraints of x, and*
3. *$|E| = |\mathcal{L}|$*

Proof. Let $\mathcal{L}$ be a maximal laminar family of tight constraints of x. Clearly, (1) holds. We will show that (2) holds, which immediately implies (3) by the rank lemma.

So, take any set S for which $x(\delta(S)) = f(S)$ that is not in the span of $\mathcal{L}$ and minimizes the number of sets $T \in \mathcal{L}$ that it intersects. Since $S \notin \mathcal{L}$, by maximality of $\mathcal{L}$ there must be some $T \in \mathcal{L}$ that it intersects. Now, by skew supermodularity, we have either $f(S) + f(T) \leq f(S \cup T) + f(S \cap T)$ or $f(S) + f(T) \leq f(S \setminus T) + f(T \setminus S)$. Let's assume the former holds as the latter case is similar. So,

$$x(\delta(S)) + x(\delta(T)) = f(S) + f(T) \leq f(S \cup T) + f(S \cap T) \leq x(\delta(S \cup T)) + x(\delta(S \cap T))$$

But by submodularity of the cut function (or just Lemma 1.3, which also clearly holds for weighted graphs), the first term is also *greater than* the first term. So, all of the above must be equalities, in particular implying $x(\delta(S)) + x(\delta(T)) = x(\delta(S \cup T)) + x(\delta(S \cap T))$. However, by Lemma 1.3 this implies that $\chi(\delta(S)) + \chi(\delta(T)) = \chi(\delta(S \cup T)) + \chi(\delta(S \cap T))$.

This means that either $\delta(S \cup T)$ or $\delta(S \cap T)$ is not in the span of $\mathcal{L}$. This is a contradiction, since $S \cup T$ and $S \cap T$ are both tight sets that intersect fewer sets than S. In Fact 1.5 we will show this for $S \cup T$ (the other case is analogous). $\square$

Fact 1.5. *If S intersects $T \in \mathcal{L}$, then $S \cup T$ intersects strictly fewer sets in $\mathcal{L}$.*

Proof. We will show that if a set $R \in \mathcal{L}$ does not intersect S, it also does not intersect $S \cup T$, which would imply the claim since $S \cup T$ does not intersect T. By assumption, $R \cap S \in \{S, R, \emptyset\}$, and since $\mathcal{L}$ is laminar, $R \cap T \in \{T, R, \emptyset\}$. We now want to show that $R \cap (S \cup T) \in \{R, S \cup T, \emptyset\}$.

$R \cap (S \cup T) = (R \cap S) \cup (R \cap T) \in \{S \cup T, S \cup R, R, \emptyset\}$. So we only need to show that we cannot have $R \cap S = S$ and $R \cap T = R$. But this would imply $T \subseteq R \subseteq S$ which contradicts that T intersects S. □

Given this lemma, we can complete the proof of the theorem, where we use an argument from [NRS10] simplifying that of [Jai01].

Theorem 1.2. *Whenever f is a skew supermodular function, every vertex of P_f has a coordinate e with $x_e \geq \frac{1}{2}$.*

Proof. Suppose by way of contradiction that $0 < x_e < \frac{1}{2}$ for all $e \in E$. We will now give a splittable token to each edge and assign tokens to sets in $\mathcal{L}$ such that every set gets one token and there is some token left over. This contradicts (3) of Lemma 1.4.

For an edge $e = (u, v)$, let L be the minimal set in $\mathcal{L}$ containing u and R the minimal set in $\mathcal{L}$ containing v. Assign x_e tokens to each of L and R. Let T be the smallest set containing both u and v, and assign the remaining $1 - 2x_e$ tokens to T. Note that these sets are not necessarily distinct.

Now consider any set S in $\mathcal{L}$. Suppose it has k children $S_1, \ldots, S_k$ (note possibly $k = 0$). Then, $x(\delta(S)) = f(S)$ and $x(\delta(S_i)) = f(S_i)$ for all i. So,

$$x(\delta(S)) - \sum_{i=1}^{k} x(\delta(S_i)) = f(S) - \sum_{i=1}^{k} f(R_i)$$

Let $A = \delta(S)$, let B be the set of edges with one endpoint in exactly one S_i and two endpoints in S, and C be the set of edges with an endpoint in two distinct S_i. Then, the above equation can be rewritten $x(A) - x(B) - 2x(C) = f(S) - \sum_{i=1}^{k} f(S_i)$. Now, $A \cup B \cup C$ cannot be empty as otherwise S is linearly dependent with its children. Now, how many tokens does S receive?

$$x(A) + |B| - x(B) + |C| - 2x(C) = |B| + |C| + f(S) - \sum_{i=1}^{k} f(S_i)$$

So, since every edge assigns a positive number of tokens to each of its sets L, R, T, and $A \cup B \cup C$ is not empty, the number of tokens S gets is some integral non-zero number, so it is at least 1. We get extra tokens since any edge in $\delta(S)$ for a maximal set $S \in \mathcal{L}$ has no corresponding set T. □

We leave one final detail as an exercise: that we can delete the edges with $x_e = 1$ and the function f remains skew supermodular. (**Hint:** a constant is a modular function, and a skew supermodular function minus a modular function is skew supermodular.)

References

[Jai01] Kamal Jain. "A Factor 2 Approximation Algorithm for the Generalized Steiner Network Problem". In: *Combinatorica* 21 (2001), pp. 39–60 (cit. on pp. 2, 4).

[NRS10] Viswanath Nagarajan, R. Ravi, and Mohit Singh. "Simpler analysis of LP extreme points for traveling salesman and survivable network design problems". In: *Operations Research Letters* 38.3 (2010), pp. 156–160. ISSN: 0167-6377. DOI: https://doi.org/10.1016/j.orl.2010.02.005 (cit. on p. 4).