| IOE 691: Approximation Algorithms | Date: 3/20/2017 |
|---|---|
| Lecture Notes: Degree-bounded Minimum Spanning Tree (Iterative LP) | |
| Instructor: Viswanath Nagarajan | Scribe: Chengyu Dai |

# 1 Problem Formulation

The input is a graph $G = (V, E)$ where $V$ and $E$ represent the set of vertices and edges. There is a cost $c_e \geq 0$ on each edge $e \in E$ and a degree bound $b_v \in \mathcal{Z}^+$ for each vertex $v \in V$. The goal is to find a minimum cost spanning tree for $G$ that satisfies all the degree bound constraints. Formally:

$$\begin{aligned} \min \quad & \sum_{e \in \mathcal{T}} c_e \\ s.t. \quad & |\delta_{\mathcal{T}}(v)| \leq b_v, \qquad v \in V, \\ & \mathcal{T} \text{ is a spanning tree in } G. \end{aligned}$$

Above, we use $\delta_{\mathcal{T}}(v)$ to represent the degree of node $v$ in tree $\mathcal{T}$.

This generalizes quite a few of the problems we have studied before. For example, Hamiltonian Path is equivalent to the special case with uniform cost and degree-bound 2 for all vertices; the min-cost spanning tree problem is the special case of having no degree-bounds (or equivalently $b_v = \infty, \forall v \in V$). The NP-hardness of Hamilton path directly shows that any polynomial algorithm for degree-bounded MST must violate degree bounds.

**Theorem 1.1** *There is a polynomial time algorithm that finds a tree $\mathcal{T}$ of cost at most* OPT *that satisfies $|\delta_{\mathcal{T}}(v)| \leq b_v + 1, \forall v \in V$, i.e. the maximal violation to the degree bound is* 1*. Here* OPT *is the minimum cost of any spanning tree that satisfies all the degree bounds.*

# 2 Linear Programming Relaxation and Extreme Points

We consider an LP relaxation with variables $x_e$ corresponding to selecting edges $e \in E$.

$$\begin{aligned} \min_{e \in E} \quad & \sum_{e \in \mathcal{T}} x_e c_e \\ s.t. \quad & \sum_{e \in E(S)} x_e \leq |S| - 1, \quad \forall S \subseteq V \qquad \text{(Spanning Tree constraint)}, \\ & \sum_{e \in E(V)} x_e = |V| - 1, \\ & \sum_{e \in \delta(v)} x_e \leq b_v, \quad \forall v \in V, \\ & x \geq 0. \end{aligned}$$

We use the notation $E(S)$ to denote the edges induced by vertex set $S$, i.e. the edges in $E$ with both end-vertices in $S$.

This LP can be solved in polynomial time via the separation oracle of minimum cut. Also, we can find an optimal extreme point solution in polynomial time.

**Definition 2.1** *Any extreme point $x$ is given by $|E|$ linearly independent constraints being set to equality. The constraints that are satisfied as equality at $x$ are called active (or tight) constraints.*

For any LP problem (bounded and non-empty), the optimal always lies in some extreme point.

For any subset $E' \subseteq E$ of edges we use $x(E') := \sum_{e \in E'} x_e$ and

$$\mathbb{1}_{E'} = \begin{cases} 1 & \text{if } e \in E', \\ 0 & \text{otherwise.} \end{cases}$$

Lemma 2.1 below shows that for any extreme point, we can choose a maximal linearly independent set of tight *spanning tree* constraints in a very structured way.

**Definition 2.2** *A Laminar family is a collection of sets $L \subseteq 2^V$ such that for any $A, B \in L$, we have one of the following be true:* $\begin{cases} A \subseteq B \\ B \subseteq A \\ A \cap B = \phi \end{cases}$

**Lemma 2.1** *For any solution $x > 0$ to the above LP, let $M \subseteq 2^V$ denote the set of tight spanning tree constraints. Then there is a laminar family $L \subseteq M$ such that* $\text{span}(L) = \text{span}(M)$. *Here* $\text{span}(L)$ *is the subspace of $\mathbb{R}^E$ generated by $\{\mathbb{1}_{E(S)} : S \subseteq L\}$, and* $\text{span}(M)$ *is defined similarly.*

To prove this important structural lemma that is the basis of algorithm, we need to prove one more simple observation.

**Observation 2.1** *(uncrossing property) Let $M$ denote the set of all tight spanning tree constraints for some LP solution $x$ and $A, B \in M$. Then we have $A \cap B, A \cup B \in M$. Moreover, if $x > 0$ then $\mathbb{1}(E(A)) + \mathbb{1}(E(B)) = \mathbb{1}(E(A \cap B)) + \mathbb{1}(E(A \cup B))$.*

**Proof:** Since $x$ is a feasible LP solution,

$$(|A| - 1) + (|B| - 1) = (|A \cap B| - 1) + (|A \cup B| - 1) \geq x(E(A \cap B)) + x(E(A \cup B)). \quad (1)$$

Figure 1 shows all the possible types of edges in $E(A \cup B)$. And Table 1 lists which of the subsets $E(A), E(B), E(A \cup B), E(A \cap B)$ each edge lies in. By this edge counting argument,

$$x(E(A \cap B)) + x(E(A \cup B)) \geq x(E(A)) + x(E(B)). \quad (2)$$

And because both $A, B \in M$, so both are tight constraint,

$$x(E(A)) + x(E(B)) = (|A| - 1) + (|B| - 1).$$

All of these can only be true when there is equality in both (1) and (2). As $x(E(A \cup B)) \leq |A \cup B| - 1$ and $x(E(A \cap B)) \leq |A \cap B| - 1$ it follows that both $A \cup B$ and $A \cap B$ are tight spanning tree constraints. It also implies

$$x(E(A)) + x(E(B)) = x(E(A \cap B)) + x(E(A \cup B)).$$

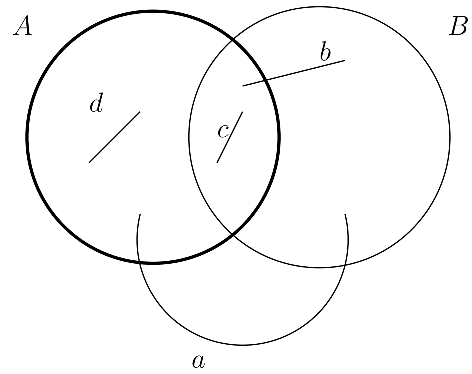


Figure 1: Four possible cases for edges

Table 1: All possible cases for edges: $a, b, c, d \in E(A \cup B)$

| cases for edge | $A$ | $B$ | $A \cup B$ | $A \cap B$ |
|---|---|---|---|---|
| a | | | ✓ | |
| b | | ✓ | ✓ | |
| c | ✓ | ✓ | ✓ | ✓ |
| d | ✓ | | ✓ | |

So when $x > 0$ we must have $\mathbb{1}(E(A)) + \mathbb{1}(E(B)) = \mathbb{1}(E(A \cap B)) + \mathbb{1}(E(A \cup B))$. ■

Now we can prove the lemma!

**Proof:** Let $L$ be any maximal set of linear independent and Laminar constraints from $M$, and we'll prove by contradiction. Suppose that

$$\exists R \in M : R \notin \text{span}(L) \tag{3}$$

Then we pick set $R$ with the minimum number of "crossings" in $L$. We first define **crossing** between sets $A, B$ as a condition that is true only when all of $A \cap B$, $A \setminus B$, $B \setminus A$ are not empty sets. The **number of crossings** for a set $R$ with $L$ is the cardinality of $\{A \in L : A \text{ crosses with } R\}$.

We now pick any $Y \in L$ that crosses with $R$ and use Observation 2.1, so $R \cap Y, R \cup Y \in M$. Note that for any $Z \in L$ (the laminar family), if $R \cap Y$ (or $R \cup Y$) crosses $Z$ then $R$ also crosses $Z$: this uses the fact that both $Y, Z \in L$. Moreover, $R \cap Y$ and $R \cup Y$ *do not* cross $Y$. So the number of crossings of $R \cap Y$ and $R \cup Y$ with $L$ is strictly less than that of $R$.

Now if we can show either $R \cap Y$ or $R \cup Y \notin \text{span}(L)$, then we can get to a contradiction. As $x > 0$, Observation 2.1 also implies

$$\mathbb{1}(E(R)) + \mathbb{1}(E(Y)) = \mathbb{1}(E(R \cap Y)) + \mathbb{1}(E(R \cup Y)).$$

If both $R \cap Y, R \cup Y \in \text{span}(L)$ then we would have $R \in \text{span}(L)$ which contradicts (3). Hence at least one of $R \cap Y, R \cup Y$ is not in the span$(L)$. For the set not in span$(L)$, it has less number of crossing than $R$, again contradicting the choice of $R$. ■

Thus we have shown:

**Theorem 2.1** *For any extreme point $x > 0$ to the above LP there is a subset $W \subseteq V$ and laminar family $L \subseteq 2^V$ such that $x$ is the unique solution to:*

$$x(\delta(v)) = b_v \quad \forall v \in W$$

$$x(E(S)) = |S| - 1 \quad \forall S \in L$$

*In particular, all these constraints are linearly independent and $|E| = |W| + |L|$.*