Scribe: Gabriel Kaptchuk

16.1 Min-Cut as an LP

We recall the basic definition of the MIN CUT PROBLEM

$$\begin{aligned} \text{Input:} &\quad \text{Graph } G = (V, E) \\ &\quad \text{Costs } c : E \to \mathbb{R}^+ \\ &\quad \text{Source } s \in V \text{ and Terminal } t \in T \\ \text{Feasible:} &\quad A \subseteq E \text{ s.t. } G - A \text{ has no s-t path} \\ \text{Objective:} &\quad \min \sum_{e \in A} c(e) \end{aligned}$$

We note that this definition of MIN CUT PROBLEM can be written in an equivalent form

$$\begin{aligned} \text{Input:} &\quad \text{Graph } G = (V, E) \\ &\quad \text{Costs } c : E \to \mathbb{R}^+ \\ &\quad \text{Source } s \in V \text{ and Terminal } t \in T \\ \text{Feasible:} &\quad S \subseteq V \text{ s.t. } s \in S \text{ and } t \notin S \\ \text{Objective:} &\quad \min \sum_{e \in E(S, \bar{S})} c(e) \end{aligned}$$

Definition 16.1.1 $\mathbf{P}_{s,t} = \{\textit{all } S - T \textit{ paths}\}$

We define the following LP, whose integer solutions are solutions to the MIN CUT PROBLEM

$$\begin{aligned} \min &\quad \sum_{e \in E} c(e) x_e \\ \text{subject to} &\quad \sum_{e \in p} x_e \geq 1 \quad \forall p \in \mathbf{P}_{s,t} \\ &\quad 0 \leq x_e \leq 1 \qquad \forall e \in E \end{aligned}$$

Intuitively, this states that along each path at least one edge must be in the cut.

Theorem 16.1.2 *This LP can be solved in polytime, even though it has an exponential number of constraints*

Proof: Suppose $\vec{x}$ is not a feasible solution to the LP. There must thus be a path $p \in \mathbf{P}_{s,t}$ s.t. $\sum_{e \in p} x_e < 1$. We now think of x_e as a length assigned to edge $e \in E$. We can easily find a shortest $S - T$ path. Because it is the shortest path, and $\exists p \in \mathbf{P}_{s,t}$ with $\sum_{e \in p} x_e = \sum_{e \in p}$ length$(e) < 1$, we can use shortest path as a separation oracle for the ellipsoid method. ∎

Theorem 16.1.3 *If $\vec{x}$ is a feasible solution to the LP with cost$(\vec{x}) = \sum_{e \in E} c(e) x_e = Z$, then we can find an integral solution with cost Z in polynomial time.*

Proof: We begin by defining a few variables.

Definition 16.1.4 *Let $d(u)$ denote the shortest path distance from s to u under the edge lengths $x_e \in \vec{x}$.*

Definition 16.1.5 *Let $B(s, r) = \{v \in V | d(v) \leq r\}$*

Definition 16.1.6 *If $S \subset V$, let $\delta(S) = E(S, \bar{S})$ = set of edges with one endpoint in S and one endpoint in $\bar{S}$*

We note that using the shortest path metric for all given nodes $v \in V$ turns the graph into a metric space.

Algorithm 1 LP rounding for Min Cut

Input: Graph $G = (V, E)$ and solution to the LP $\vec{x}$
Output: $S \subseteq E$

Choose r uniformly at random in $[0, 1]$
$S \leftarrow B(s, r)$
return $A \leftarrow \delta(S)$

Claim 16.1.7 *Let $e = \{u, v\} \in E$. $Pr[e \in A] \leq x_e$*

Proof: WLOG let $d(u) \leq d(v)$. For $e \in A$ the radius of the ball $B = (s, r)$ must be enough to contain u, but not v. So, $Pr[e \in A] = Pr[r \in [d(u), d(v)]] \leq d(v) - d(u) \leq d(u, v) \leq x_e$, where $d(u, v)$ denotes the length of the shortest $u - v$ path. ∎

Hence by linearity of expectation, By linearity of expectations, $E[c(A)] = \sum_{e \in E} c(e) Pr[e \in A] = \sum_{e \in E} c(e) x_e = Z$. We can take this seemingly randomized algorithm and make it deterministic by simply trying all possibilities. Suppose $\nexists$ node w such that $d(u) \leq d(w) \leq d(v)$. Thus any random $r \in [d(u), d(v)]$ will yield the same cut. Thus there are a linear number of possible cuts, which can each be tested in linear time $\Rightarrow$ the entire algorithm can be run in polynomial time. ∎

16.2 Multiway Cut

We define the MULTIWAY CUT PROBLEM.

Input:	Graph $G = (V, E)$
	Costs $c : E \to \mathbb{R}^+$
	$T = \{s_1, s_2, ..., s_k\} \subseteq V$
Feasible:	$A \subseteq E$ s.t. $G - A$ has no $s_i - s_j$ path $\forall i, j \in \{1, 2, ..., k\}$
Objective:	$\min \sum_{e \in A} c(e)$

Definition 16.2.1 *Let* $\mathbf{P}_{u,v} = \{$*all simple* $u - v$ *paths*$\}$

We define the following LP, whose integer solutions are solutions to the Multiway cut problem

$$
\begin{aligned}
\min \quad & \sum_{e \in E} c(e) x_e \\
\text{subject to} \quad & \sum_{e \in p} x_e \geq 1 \quad \forall i, j \in \{1, 2, ..., k\}, \forall p \in \mathbf{P}_{s_i, s_j} \\
& 0 \leq x_e \leq 1 \qquad\qquad \forall e \in E
\end{aligned}
$$

Note that the LP solution $\vec{x}$ induces a metric d on the nodes through the shortest-path distances, and the constraints guarantee that $d(s_i, s_j) \geq 1$ for all $i, j \in [k]$. We use the same separation oracle as in the above Min Cut problem example, but simply use it to check for each pair of $\{i, j\}$.

Algorithm 2 LP rounding for Multiway Cut

Input: Graph $G = (V, E)$ and solution to the LP $\vec{x}$
Output: $S \subseteq E$

```
A ← ∅
for all i ∈ {1, 2, .., k} do
    Randomly choose r ∈ [0, 1/2]
    A_i ← δ(B(s_i, r))
    A ← A ∪ A_i
end for
return A
```

Theorem 16.2.2 *If* $\vec{x}$ *is a feasible solution to the LP with cost*$(\vec{x}) = \sum_{e \in E} c(e) x_e = Z$, *then there is a polynomial time algorithm to find an integral solution* $\vec{x'}$ *such that cost*$(\vec{x'}) \leq 2Z$ *which can be trivially reduced to* $2(1 - \frac{1}{k})Z$.

Proof: We begin by proving a claim

Claim 16.2.3 $\forall e = \{u, v\} \in E$, $Pr[e \in A] \leq 2x_e$

Proof: Let $w \in V$. By the triangle inequality, $d(s_i, w) + d(w, s_j) \geq d(s_i, s_j) \geq 1$.

Definition 16.2.4 *Let* $C_i = \{v \in V | d(s_i, v) \leq \frac{1}{2}\}$. *Clearly* $C_i \cap C_j = \emptyset \ \forall i, j$

Case 1: $u, v \in C_i$ for some $i \in 1, 2, ..., k$. WLOG we assume $d(s_i, u) \leq d(s_i, v)$.

$$Pr[e \in A] = Pr[e \in A_i] = Pr[r \in [d(s_i, u), d(s_i, v)]] = \frac{d(s_i, v) - d(s_i, u)}{\frac{1}{2}} \leq 2d(u, v) \leq 2x_e$$

Case 2: $u \in C_i$, $v \notin C_i$ and $v \in C_j$ for $i \neq j$

$$Pr[e \in A] \leq Pr[e \in A_i] + Pr[e \in A_j] \leq Pr[r \in [d(s_i, u), \frac{1}{2}]] + Pr[r \in [d(s_j, v), \frac{1}{2}]] =$$

$$\frac{\frac{1}{2} - d(s_i, u)}{\frac{1}{2}} + \frac{\frac{1}{2} - d(s_j, u)}{\frac{1}{2}} = 2(1 - d(s_i, u) - d(s_j, v)) \leq$$
$$2(d(s_i, s_j) - d(s_i, u) - d(s_j, v)) \leq 2d(u, v) \leq 2x_e$$

So in all cases $Pr[e \in A] \leq 2x_e$. ■

By linearity of expectations $E[c(A)] = \sum_{e \in E} c(e) Pr[e \in A] \leq \sum_{e \in E} c(e)(2x_e) \leq 2 \sum_{e \in E} c(e) x_e \leq 2Z$. Thus this is a polynomial time algorithm to find a integral solution given a fractional LP solution $\vec{x}$. In the same way as the previous example, we can test polynomial possibilities in polynomial time. ■

16.2.1 Integrality Gap

Is our analysis tight? We consider the star with k nodes around the outside connected by a single node v. We choose as our k terminal nodes the outside nodes.

The optimal solution OPT is clearly given by cutting all but one edge. So cost is given by $k - 1$.

The worst case LP solution is given by assigning all edges $\frac{1}{2}$. So cost is given by $\frac{k}{2}$

Thus the gap is given by $\frac{OPT}{LP} = \frac{k-1}{\frac{k}{2}} = 2(1 - \frac{1}{k})$. So our analysis is tight.

16.2.2 A better LP

We consider a better solution to the Multiway cut problem. The problem itself stays the same. We create pieces $C_1, C_2, ...C_k$ s.t. $s_i \in C_i$ $\forall i \in 1, 2, ..., k$ and $C_i \cap C_j = \emptyset$ for $i \neq j$. We define

$$X_u^i = \begin{cases} 1 & u \in C_i \\ 0 & else \end{cases}$$

$$Z_e^i = \begin{cases} 1 & e \in \delta(C_i) \\ 0 & else \end{cases}$$

We now define an LP using these indicator variables

$$
\begin{aligned}
\min \quad & \frac{1}{2}\sum_{e\in E}\sum_{i=1}^{k} c(e)Z_e^i \\
\text{subject to} \quad & \sum_{i=1}^{k} x_u^i = 1 && \forall u \in V \\
& Z_e^i \geq X_u^i - X_v^i \qquad \forall e = u,v \in E, && \forall i \in 1,2,...,k \\
& Z_e^i \geq X_v^i - X_u^i \qquad \forall e = u,v \in E, && \forall i \in 1,2,...,k \\
& X_{s_i}^i = 1 && \forall i \in 1,2,...,k \\
& 0 \leq X_u^i \leq 1 \\
& 0 \leq Z_e^i \leq 1
\end{aligned}
$$

It is straightforward to verify that this is a valid relaxation of the multiway cut problem. We will give a more compact way of writing this LP which makes the connection to metrics clear.

Definition 16.2.5 *Let $x, y \in \mathbb{R}^k$ then their ℓ_1-distance is $||x-y||_1 = \sum_{i=1}^{k} |x^i - y^i|$ where x^i is the i^{th} coordinate of the vector x.*

Definition 16.2.6 *Let $\Delta_k = \{x \in \mathbb{R}^k | \sum_{i=1}^{k} x^i = 1, \quad x^i \geq 0 \ \forall i\}$ where x^i is the i^{th} coordinate of the vector x.*

Definition 16.2.7 *Let e_i be a vector with a 1 in the i^{th} coordinate and zeros elsewhere*

Let $X_u = (X_u^1, X_u^2, ..., X_u^k)$ and $X_v = (X_v^1, X_v^2, ..., X_v^k)$. Note that $Z_e^i = |X_u^i - X_v^i|$ in any optimal LP solution because of the constraints on Z_e^i and because we are minimizing the objective function. So, $||X_u - X_v|| = \sum_{i=1}^{k} Z_e^i$. Using all these definitions we can rewrite the new LP as follows

$$
\begin{aligned}
\min \quad & \frac{1}{2}\sum_{e=\{u,v\}\in E} ||X_u - X_v||_1 \\
\text{subject to} \quad & X_{s_i} = e_i && \forall i \in 1,2,...,k \\
& X_u \in \Delta_k
\end{aligned}
$$

Using this LP, there exists a rounding that yields a $\frac{3}{2}$ approximation of the problem