CS-GY 6763/CS-UY 3943: Lecture 8
Linear programming and relaxations

NYU Tandon School of Engineering, R. Teal Witter

Given:

- n ground elements $[n] = \{1, 2, \ldots, n\}$
- m sets $S_1, S_2, \ldots, S_m$ where $S_j \subseteq [n]$
- non-negative weights $w_j \geq 0$

for $j \in [m]$.

Find:

$$\min_{C \subseteq [m]} \sum_{j \in C} w_j \qquad \text{subject to} \qquad \cup_{j \in C} S_j = [n].$$

Given:

- ground elements are known viruses
- sets are three-byte code sequences that occur in viruses so

$$S_j = \{\text{viruses that contain } j\text{th three-byte sequence}\}$$

- non-negative weights $w_j \geq 0$

for $j \in [m]$.

Which sequences should we use to identify viruses?

Given:

- ground elements are edges
- sets are nodes so

$$S_j = \{\text{edges adjacent to } j\text{th node}\}$$

- $w_j = 1$

for $j \in [m]$.

What vertices should we choose so that all edges are connected to at least one chosen vertex?

Let $\mathsf{x} \in \mathbb{R}^m$ be a vector of decision variables and $\mathsf{c} \in \mathbb{R}^n$ be a vector of constraints. Then the primal linear programs is:

$$\text{Minimize} \quad \mathsf{b}^\mathsf{T}\mathsf{x} \quad \text{subject to} \quad \mathsf{A}\mathsf{x} \geq \mathsf{c}, \quad \mathsf{x} \geq 0 \quad (\textit{Primal})$$

Question: What are the dimensions of A and b?

Let $x_j = 1$ iff $j \in C$.

The objective is to minimize the sum of weights in C:

$$\min_{C \subseteq [m]} \sum_{j \in C} w_j \quad \Leftrightarrow \quad \min_{\mathbf{x}} \sum_{j=1}^{m} w_j x_j \quad \Leftrightarrow \quad \min_{\mathbf{x}} \mathbf{b}^T \mathbf{x}$$

Question: What is $\mathbf{b}$?

The constraint is that C covers the ground elements:

$$\cup_{j \in C} S_j \quad \Leftrightarrow \quad \sum_{j : i \in S_j}^{m} x_j \geq 1 \;\; \forall i \in [n] \quad \Leftrightarrow \quad \mathbf{A}\mathbf{x} \geq \mathbf{c}$$

Question: What are $\mathbf{A}$ and $\mathbf{c}$?

$$\text{Minimize} \quad \mathbf{b}^{\mathsf{T}}\mathbf{x} \quad \text{subject to} \quad \mathbf{A}\mathbf{x} \geq \mathbf{c}$$

$$\text{Minimize} \quad \sum_{j=1}^{m} w_j x_j \quad \text{subject to} \quad \sum_{j:i \in S_j} x_j \geq 1$$

We can solve linear program in polynomial time with e.g. Interior Point, Ellipsoid Method.

But there's a problem …what is it?

Definition (Relaxation)

A linear program (where $\mathbf{x} \in \mathbb{R}^m$) is a *relaxation* of an integer program (where $\mathbf{x} \in \{0,1\}^m$) if

- a feasible solution to the IP is a feasible solution to the LP and
- the value of the feasible solution in the IP has the same value in the LP.

Therefore $OPT_{LP} \leq OPT_{IP} = OPT_{SC}$.

Theorem

Let $\mathbf{x}^*$ be optimal solution to LP. Define

$$f = \max_{i \in [n]} |\{j : i \in S_j\}|.$$

Choose C so that $j \in C$ iff $x_j^* \geq 1/f$.

Then C is feasible and gives an f-approximation to set cover.

Question: What approximation do we get for vertex cover?

Claim: C is feasible.

Fix i. We know

$$\sum_{j:i\in S_j} x_j^* \geq 1.$$

Then

Claim: C gives an f-approximation to set cover.

$$\sum_{j \in C} w_j$$

$$\begin{array}{llll} \text{Minimize} & b^T x & \text{subject to} & Ax \geq c, \quad x \geq 0 \quad (\textit{Primal}) \\ \text{Maximize} & c^T y & \text{subject to} & A^T y \leq b, \quad y \geq 0 \quad (\textit{Dual}) \end{array}$$

Weak duality: $c^T y \leq b^T x$

Proof:

Note: strong duality holds if primal and dual are bounded and feasible.

$$\text{Maximize} \quad \mathbf{c}^T\mathbf{y} \quad \text{subject to} \quad \mathbf{A}^T\mathbf{y} \leq \mathbf{b} \quad (\textit{Dual})$$

$$\text{Maximize} \quad \sum_{i=1}^{n} y_i \quad \text{subject to} \quad \sum_{i:i\in S_j} y_i \leq w_j$$

Intuition: y_i represents how much we pay for ground element i.

Theorem

Let $\mathbf{y}^*$ be optimal solution to dual. Choose C' so that $j \in C'$ iff

$$\sum_{i:i\in S_j} y_i^* = w_j.$$

Then C' is feasible and gives an f-approximation to set cover.

Claim: C' is feasible.

Suppose for contradiction there's an uncovered ground element k.

Then for all subsets S_j containing k, we have $\sum_{i:i \in S_j} y_i^* < w_j$.

Define $\epsilon = \min_{j:k \in S_j}(w_j - \sum_{i:i \in S_j y_i^*})$.

Now let $\mathbf{y}'$ be $\mathbf{y}^*$ except with $y_k' = y_k^* + \epsilon$.

Then

Claim: C' gives an f-approximation to set cover.

$$\sum_{j \in C'} w_j$$

Question: Is C' better than C?

$$\sum_{i=1}^{n} y_i \leq \sum_{i=1}^{n} y_i \sum_{j:i\in S_j} x_j = \sum_{j=1}^{m} x_j \sum_{i:i\in S_j} y_i \leq \sum_{j=1}^{m} x_j w_j$$

Strong duality: the above inequalities are tight for $\mathbf{x}^*$ and $\mathbf{y}^*$.

- If $y_i^* > 0$ then $\sum_{j:i\in S_j} x_j^* = 1$
- If $x_j^* > 0$ then $\sum_{i:i\in S_j} y_i^* = w_j$

We know we can use an LP solver for primal.

Can we do better for dual?