Lecture Notes CS:5350 Approximation by Linear Programming (LP) Rounding
Lecture 26: Nov 21, 2019
Scribe: Xiaoyu Xing

1 Generals of LP Rounding

LP rounding to design approximation algorithms is typically used in context of 0-1 optimization problem. Since many combinatorial problems can be encoded as integer programs (IP), solving those IP is in general NP-hard, so we generally relax the integrality constraint into a linear constraint, e.g. non-negativity constraint, then solve the corresponding LP and round the result to integer to get the result of original IP. Here is the overview of LP relaxation and rounding methods:

1. Formulate an optimization problem as an integer program (IP)
2. Relax the integral constraints to turn the IP to an LP
3. Solve LP to obtain an optimal solution x^*
4. Construct a feasible solution $x^{'}$ to IP by rounding from x^* integers

Rounding can be done deterministically or probabilistically, i.e. randomized rounding. As we discussed, the 2-approx algorithm for minimum vertex cover (MVC) is an example of deterministic rounding. What we will discuss today is an example of randomized rounding to SetCover problem by a approximiation factor of $O(\log n)$, other examples include MaxSAT, MaxCut, etc.
Reference: 1, 2, 3, 4.

2 SetCover

As we discussed previously, here is the definition of the problem SetCover.
Input: A ground set X ($|X|$ = m), a collection S_1, S_2, S_3, ..., S_n of subsets of X. Assuming $\bigcup\limits_{i=1}^{n} S_i = X$.
Output: A cover C $\subseteq 1, 2, 3, ..., n$ with smallest size such that $\bigcup\limits_{i:C} S_i = X$.

2.1 IP of SetCover

Let $x_i \in \{0, 1\}$ indicates whether S_i is picked by an algorithm, 1 indicates S_i is picked, otherwise is 0:

$$\min \quad \sum_{i=1}^{n} x_i \tag{1a}$$

$$\text{s.t.} \quad \sum_{i:e\in S_i} x_i \geq 1 \text{ for each } e \in X, \tag{1b}$$

$$x_i \in \{0, 1\} \text{ for i = 1, 2, ..., n} \tag{1c}$$

2.2 LP of SetCover

From what we have in IP for SetCover, to have LP for SetCover, let relax the integrality constraint $x_i \in \{0,1\}$ in IP to non-negativity constrains, $x_i \geq 0$. Still, let $x_i \in \{0,1\}$ indicates whether S_i is picked by an algorithm.

$$\min \quad \sum_{i=1}^{n} x_i \tag{2a}$$

$$\text{s.t.} \quad \sum_{i:e \in S_i} x_i \geq 1 \text{ for each } e \in X, \tag{2b}$$

$$x_i \geq 0 \text{ for i = 1, 2, ..., n} \tag{2c}$$

2.3 Deterministic LP rounding

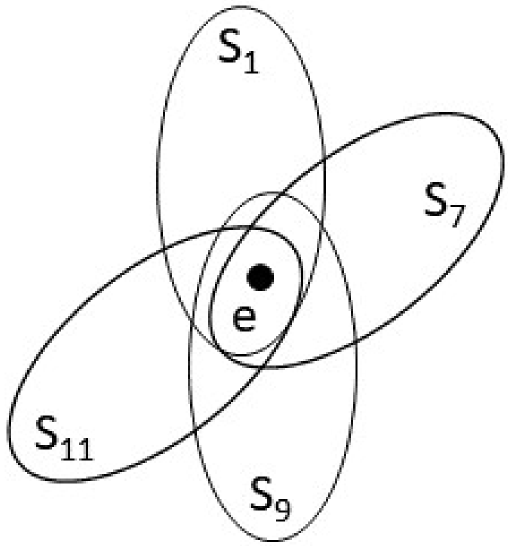
For each element $e \in X$, let $f(e)$ denote number of sets S_i that contains e, i.e. the frequency or counting of S_i covering e. Let $f = \max_{e \in X} f(e)$, here we will discuss a simple f-approx algorithm for SetCover by deterministic rounding. (As an analogy, in MVC, the ground set X would be all the edges, the vertices would be the chosen subset S_i to cover all the edges.)

1: Solve LP relaxation of SetCover to get solution $\{x_i^* | i = 1, 2, ..., n\}$
2: **for** $i \leftarrow 1$ to n **do**
3:     **if** $x_i^* \geq \frac{1}{f}$ **then**
4:         $z_i \leftarrow 1$
5:     **else**
6:         $z_i \leftarrow 0$
7: Output the solution $C = \{i | z_i = 1\}$

Algorithm 1: Approximation Algorithm by Deterministic LP Rounding

Lemma 1. *Solution C is a valid SetCover*

Proof Outline: Let take below graph as an example to illustrate, e is an element in ground set X, S_1, S_7, S_9, S_{11} are the sets covering e. Thus, $f(e) = 4 \leq f$. After solving the LP for this problem, the constraint for e is satisfied as such: $x_1 + x_7 + x_9 + x_1 \geq 1$. Together, it is easy to conclude at least one of $S_i \in [S_1,\ S_7,\ S_9,\ S_{11}] \geq \frac{1}{f(e)} = \frac{1}{4} \geq \frac{1}{f}$. (Reminder: f is the global maximum frequency/counting of sets S_i covering element e.) Following above algorithm line 3, at least one set among $[S_1,\ S_7,\ S_9,\ S_{11}]$ covers element e and set z_i to be 1. Similarly, in general for any element in the ground set X, at least one of the set covers such element and set z_i to be 1. Therefore, C is a valid SetCover.



Lemma 2. *Let OPT be the size of optimized solution to original SetCover problem, then* $|C| \leq f \cdot OPT$

Proof. Since $\{x_i^* | i = 1, 2, ..., n\}$ is an optimal solution to the LP relaxation SetCover problem, we have:

$$\sum_{i=1}^{n} x_i^* \leq OPT$$

(Reminder: because this is a minimization problem and LP has a larger solution domain than IP counterpart, the optimal solution from LP would be at least as good as optimal solution from IP.) Now, note that, following above algorithm line 3 - 6, it is easy to get:

$$z_i \leq f \cdot x_i^*$$

$$\therefore \sum_{i=1}^{n} z_i^* \leq f \cdot \sum_{i=1}^{n} x_i^* \leq f \cdot OPT$$

Following above algorithm line 7, $C = \{i | Z_i = 1\}$, we can conclude:

$$|C| = \sum_{i=1}^{n} z_i^* \leq f \cdot OPT$$

□

2.4 Randomized LP rounding

Now, let's design a LP randomized rounding algorithm that gives an $O(\log n)$-approximation. Here is a randomized LP rounding algorithm, surprisingly using the LP solution as the probability in randomized rounding.

```
Solve LP relaxation of SetCover to get solution {x_i^* | i = 1, 2, ..., n}
for i ← 1 to n do
    z_i ← 1 with probability x_i^*
    z_i ← 0 with probability 1 − x_i^*
Output the solution C = {i | z_i = 1}
```

Algorithm 2: Approximation Algorithm by Randomized LP Rounding

Lemma 3. *Let OPT be the size of optimized solution to original SetCover problem, then*

$$\mathbf{E}[|C|] = \sum_{i=1}^{n} x_i^* \le OPT$$

Proof.

$$\begin{aligned}
\mathbf{E}[|C|] &= \mathbf{E}[\sum_{i=1}^{n} z_i] \\
&= \sum_{i=1}^{n} \mathbf{E}[z_i], \text{ by linearity of expectation} \\
&= \sum_{i=1}^{n} Prob[z_i = 1] \\
&= \sum_{i=1}^{n} x_i^* \\
&\le OPT
\end{aligned} \tag{3}$$

□

Following above proof, we can "surprisingly" see that the expected size of the randomized rounding solution from LP solution is less than the size of optimal solution, though by such randomized rounding, Further, we know that C may not be a feasible solution to the original SetCover problem, i.e. C may only be a partial cover. Thus, naturally we can augment the result by repeating this algorithm t times, getting partial covers $C_1, C_2, C_3, ..., C_t$, then returning $\bigcup_{j=1}^{t} C_j$.

2.5 Analysis of augmented randomized LP rounding

Let us first focus on one iteration in the augmented LP rounding algorithm. Let $e \in X$ be an arbitrary element in ground set X.

$$\begin{aligned}
Prob[e \text{ is covered}] &= 1 - Prob[e \text{ is NOT covered by any of the set } S_i \text{ containing e}] \\
&= 1 - \prod_{i:e \in S_i} Prob[e \text{ is not covered by } S_i], \text{ by independence rounding of } Z_i \\
&= 1 - \prod_{i:e \in S_i} (1 - x_i^*)
\end{aligned} \tag{4}$$

Using $(1 + x) \leq exp(x)$ (Use $exp()$ to denote natural exponential function to prevent confusion with element e), we get

$$
\begin{aligned}
Prob[e \text{ is covered}] &\geq 1 - \prod_{i:e\in S_i} exp(-x_i^*) \\
&= 1 - exp(-\sum_{i:e\in S_i} x_i^*), \text{ by property of exponential function} \\
&\geq 1 - exp(-1), \text{ by constraints of LP, i.e. } \sum_{i:e\in S_i} x_i^* \geq 1
\end{aligned} \tag{5}
$$

Thus, we proved for one iteration of augmenting randomized LP rounding, the probability of an arbitrary element e being covered is at least a constant. In the next class, we will continue to see with after t iterations (t will be substituted by a specific expression), the augmented randomized LP rounding leading to a feasible solution would be with high probability.