Scribe: Archit Gupta, Pavan Kuppili **Lecturer:** Shuchi Chawla
Topic: Randomized rounding, concentration bounds **Date:** 10/10/2007

12.1 Set Cover

Let $E = \{e_1, e_2, .., e_n\}$ be a set of n elements. Let $S_1, S_2, .., S_m$ be subsets of E with associated costs $c_1, .., c_m$. As explained in a previous lecture, the problem of set cover is to choose 1 or more of the m subsets such that the union of them cover every element in E, and the objective is to find the set cover with the minimum cost. In a previous lecture, we have seen a greedy solution to the set cover which gave a log n approximation. In this lecture, we will see a linear programming solution with randomized rounding to achieve an $O(logn)$ approximation.

12.1.1 ILP formulation of set cover

Let X_i be a random variable associated with each subset S_i.
$X_i = 1$, if S_i is in the solution, and 0 otherwise.

The ILP formulation:
Minimize $\sum_{i=1}^{m} c_i X_i$ s.t
$\forall e \in E, \sum_{i:e \in S_i} X_i \geq 1$ - (1)
$\forall X_i, X_i \in \{0, 1\}$ - (2)

The constraints in (1) ensure that every element is present in atleast one of the chosen subsets. The constraints in (2) just mean that every subset is either chosen or not. The objective function chooses a feasible solution with the minimum cost.

It can be noted that this problem formulation is a generalization of the vertex cover. If we map every edge to an element in E, and every vertex to one of the the subsets of E, the vertex cover is the same as the set cover problem with the additional constraint that every element e appears in exactly 2 subsets (corresponding to the 2 vertices it is incident on).

12.1.2 Relaxing to an LP problem

Instead of $X_i \in \{0, 1\}$, relax the constraint so that X_i is in the range [0, 1]. The LP problem can be solved optimally in polynomial time. Now we need to round the solution back to an ILP solution.

12.1.3 Deterministic rounding

First, let us see the approximation we get by using a deterministic rounding scheme analogous to the one we used in the vertex cover problem.

Theorem 12.1.1 *We have a deterministic rounding scheme which gives us an F-approximation* *to the solution, where F is the maximum frequency of an element.*

Proof: Let F be the maximum frequency of an element, or the maximum number of subsets an element appears in.
Pick a threshold t = $\frac{1}{F}$.
If $x_1, .., x_m$ is the solution to the LP, the ILP solution is all subsets for which $x_i \geq \frac{1}{F}$.

This results in a feasible solution because in any of the constraints in (1) we have atmost F variables on the LHS, and atleast one of them should be $\geq \frac{1}{F}$, so we will set atleast one of those variables to 1 during the rounding. So, every element will be covered. It can be noted here that the vertex cover has F = 2, and hence we used a threshold of $\frac{1}{2}$ during the previous lecture. If the values of $X_1, .., X_m$ are $x_1^{'}, .., x_m^{'}$ after rounding, $\forall i, x_i^{'} \leq F x_i$ [since if $x_i \geq \frac{1}{F}$ we round it to 1, and 0 otherwise]
So the cost of the ILP solution ≤ F(cost of the LP solution). So, this deterministic rounding gives us an F approximation to the optimal solution. ∎

12.1.4 Randomized rounding

Algorithm:
Step 1. Solve the LP.
Step 2 (Randomized rounding). ∀ i, pick S_i independently with probability x_i (where x_i is the value of X_i in the LP solution).
Step 3. Repeat step 2 until all elements are covered.

The intution behind the algorithm is that higher the value of x_i in the LP solution, the higher the probability of S_i being picked during the random rounding.

Theorem 12.1.2 *With a probability $(1 - \frac{1}{n})$, we get a 2logn approximation.*

Proof:

Lemma 12.1.3 *The expected cost in each iteration of step 2 is the cost of the LP solution.*

Proof: Let Y_i be a random variable.
$Y_i = c_i$ if S_i is picked, and 0 otherwise.
Let Y = $\sum_{i=1}^{n} Y_i$,
E[Y] = $\sum_{i=1}^{n} E[Y_i]$ [by linearity of expectation]
= $\sum_{i=1}^{n} c_i x_i$
= cost of the LP solution. ∎
This is a nice result as the expected cost of the ILP solution is exactly equal to the cost of the LP solution.

Lemma 12.1.4 *The number of iterations of step 2 is 2logn with a high probability (whp).*

Proof: Fix some element e $\in$ E.
$\mathbf{Pr}$[e is not covered in any one execution of step] = $\prod_{i:e\in S_i}\mathbf{Pr}[S_i$ is not picked] (since the subsets are picked independently)
$= \prod_{i:e\in S_i}(1-x_i)$
$\leq \prod_{i:e\in S_i} e^{-x_i}$ (using the fact that if $x \in [0,1], (1-x) \leq e^{-x}$)
$= e^{-\sum_{i:e\in S_i} x_i}$
$\leq \frac{1}{e}$ (since the LP solution satisfies constraint (1) which means $\sum_{i:e\in S_i} x_i \geq 1$)

If we execute step 2 (2logn) times,
$\mathbf{Pr}$[element e is still uncovered] $\leq \frac{1}{e^{2logn}}$ (since each execution is independent)
$= \frac{1}{n^2}$.
By the union bound, $\mathbf{Pr}[\exists$ an uncovered element after 2logn executions] $\leq \frac{1}{n}$
Hence with a high probability, the algorithm will terminate in 2log n iterations of step 2. ■

The total expected cost is just the expected cost of each iteration multiplied by the number of iterations.
So E[cost] = (2log n).Cost(LP solution) with a probability $(1-\frac{1}{n})$ (using Lemmas 12.1.3 and 12.1.4).
■

12.2 Concentration Bounds

In analyzing randomized techniques for rounding LP solutions, it is useful to determine how close to its expectation a random variable is going to turn out to be. This can be done using concentration bounds: We look at the probability that given a certain random variable X, the probability that X lies in a particular range of values (Say, the deviation from the expectation value). For instance, we want $\mathbf{Pr}[X \geq \lambda]$ for some value of λ. Note that if a random variable has small variance, or (as is often the case with randomized rounding algorithms) is a sum of many independent random variables, then we can give good bounds on the probability that it is much larger or much smaller than its mean.

12.2.1 Markov's Inequality

Given a random variable $X \geq 0$, we have,

$$\mathbf{Pr}[X \geq \lambda] \leq \frac{\mathbf{E}[X]}{\lambda} \tag{12.2.1}$$

Also, given some $f : X \rightarrow \Re^+ \cup \{0\}$,

$$\mathbf{Pr}[f(X) \geq f(\lambda)] \leq \frac{\mathbf{E}[f(X)]}{f(\lambda)} \tag{12.2.2}$$

If the above function f is monotonically increasing then in addition to (12.2.2), the following also holds:

$$\mathbf{Pr}[X \geq \lambda] = \mathbf{Pr}[f(X) \geq f(\lambda)] \leq \frac{\mathbf{E}[f(X)]}{f(\lambda)} \tag{12.2.3}$$

12.2.2 Chebyshev's Inequality

We now study a tighter bound called the Chebyshev's bound.
Let $f(X) = (X - \mathbf{E}[X])^2$. Note that f is an increasing function if $X > \mathbf{E}[X]$.
We have,

$$\mathbf{Pr}[|X - \mathbf{E}[X]| \leq \lambda] = \mathbf{Pr}\big[(X - \mathbf{E}[X])^2 \leq \lambda^2\big] \tag{12.2.4}$$

$$= \frac{\mathbf{E}\big[(X - \mathbf{E}[X])^2\big]}{\lambda^2} = \frac{\sigma^2(X)}{\lambda^2} \tag{12.2.5}$$

That is, the deviation of X from $\mathbf{E}[X]$ is a function of its variance($\sigma(X)$). If the variance is small, then we have a tight bound.

Also not that with probability p, $X \in \mathbf{E}[X] \pm \sqrt{\frac{1}{p}}\sigma(X)$

12.2.3 Chernoff's bound

We present Chernoff's bound which is tighter compared to the Chebyshev and Markov bounds.

Let random variables X_1, X_2, ..., X_n be independent and identically distributed random variables, where $X_i \in [0, 1]$. Note that the class of indicator random variables lie in this category.

Let $X = \sum_{i=1}^{n} X_i$. Also $\mathbf{E}[X] = \sum_{i=1}^{n} \mathbf{E}[X_i] = \mu$

For any $\delta > 0$,

$$\mathbf{Pr}[X \geq (1+\delta)\mu] \leq \left(\frac{e^{\delta}}{(1+\delta)^{1+\delta}}\right)^{\mu} \tag{12.2.6}$$

It can also be shown for $\delta \geq 0$,

$$\mathbf{Pr}[X \geq (1+\delta)\mu] \leq e^{-\left(\frac{\delta^2}{2+\delta}\right)\mu} \tag{12.2.7}$$

Note that the above probability resembles a gaussian/bell curve. When $0 \leq \delta \leq 1$

$$\mathbf{Pr}[X \geq (1+\delta)\mu] \leq e^{-\frac{\delta^2}{3}\mu} \tag{12.2.8}$$

$$\mathbf{Pr}[X \leq (1-\delta)\mu] \leq e^{-\frac{\delta^2}{2}\mu} \tag{12.2.9}$$

Example: Coin tossing to show utility of bounds. We toss n independent unbiased coins. We want to find a t such that: $\mathbf{Pr}$[No. of times we get heads in n tosses $\geq t$] $\leq 1/n$.

We use indicator random variables to solve this: If the ith coin toss is heads, let $X_i = 1$, otherwise $X_i = 0$. Let $X = \sum_{i=1}^n X_i$. Here X is the total number of heads we get in n independent tosses. Note that $\mu = \mathbf{E}[X] = \sum_{i=1}^n \mathbf{E}[X_i] = n/2$, since $\mathbf{E}[X_i] = 1/2$ (The probability of getting a heads).

To get bounds on $\mathbf{Pr}[X \geq t]$, we first apply Markov's inequality:

$$\mathbf{Pr}[X \geq t] \leq \frac{\mathbf{E}[X]}{t} = \frac{n/2}{t}$$

$$\mathbf{Pr}[X \geq t] = 1/n \Rightarrow \frac{n/2}{t} = 1/n \Rightarrow t = n^2/2$$

But, we don't do any more than n coin tosses, so this bound is not useful. Note that Markov's bound is weak.

Applying Chebyshev's inequality:

$$\mathbf{Pr}[X \geq t] \leq \mathbf{Pr}[|X - \mu| \geq t - \mu] \leq \frac{\sigma^2(X)}{(t-\mu)^2}$$

Evaluating $\sigma(X)$ where $X_i \in \{0, 1\}$,

$$\mathbf{E}[X_i] = 1/2$$

$$\sigma^2(X_i) = \mathbf{E}\big[(X_i - \mathbf{E}[X_i])^2\big] = \frac{1}{2}(\frac{1}{4}) + \frac{1}{2}(\frac{1}{4}) = 1/4$$

Since this is a sum of independent random variables, the variances can be summed together:

$$\sigma^2(X) = \sum_{i=1}^n \sigma^2(X_i) = n/4$$

Evaluating t,

$$\mathbf{Pr}[X \geq t] \leq \frac{\sigma^2(X)}{(t-\mu)^2} = 1/n$$

$$\Rightarrow t = \mu + n/2 = n$$

Again this bound is quite weak.

Applying Chernoff's bound:

$$\mathbf{Pr}[X \geq (1+\delta)\mu] \leq e^{-\frac{\delta^2}{3}\mu} = e^{-\frac{\delta^2}{6}n}$$

Taking,

$$e^{-\frac{\delta^2}{6}n} = 1/n$$

we get,

$$\delta = \sqrt{\frac{6logn}{n}}$$

$$\Rightarrow t = (1 + \delta)\mu = n/2 + \sqrt{\frac{3nlogn}{2}}$$

This bound is much nicer than what we obtained earlier, showing us this gradual increase in tightness of the bounds across Markov, Chebyshev and Chernoff. ■