CSC373S Lecture 23

- We started the class answering questions about the final assignment.

- Lets consider one more IP/LP, namely for the weighted Max-Sat problem (for an arbitrary CNF formula F). Until recently (result presented at a conference in January, 2011), the best approximation ratio (3/4) for Max-Sat was based on the method to be discussed. The recent result provides a randomized greedy algorithm with the same approximation ratio (3/4 in expectation). The best known hardness of approximation is 7/8. Here we will again (as in set cover) use randomization in a natural way when we have a LP relaxation where the fractional variables can be viewed as probabilities. Consider the following IP formulation of (Weighted) Max-Sat:

  Note again: Here we are looking at all CNF formulas as input in contrast to Max-k-Sat and Exact Max-k-Sat.

  maximize $\sum_j w_j \cdot z_j$

  subj to $\sum_{i \in C_j^+} y_i + \sum_{i \in C_j^-} (1 - y_i) \geq z_j$

  $y_i \in \{0, 1\}; z_j \in \{0, 1\}$

  Here the intended meaning of z_j is that clause C_j will be satisfied and the intended meaning of y_i is that the propositional variable x_i is set true (false) if $y_i = 1$ (resp 0).

  C_j^+ (resp C_j^-) is the set of all variables occurring positively (resp negatively) in C_j. e.g. for $C_j = x_1 \wedge \bar{x}_2 \wedge x_3$, we have $C_j^+ = \{x_1, x_3\}; C_j^- = \{x_2\}$

  Since we have forced our fractional solutions to be in [0,1], we can think of each fractional variable as a probability. Then we can do randomized rounding (as we did for the set cover problem). Let $\{y_i^*, z_j^*\}$ be an optimal LP solution. Then we set $\bar{y}_i = 1$ with probability y_i^* to obtain an integral solution. We do not need to round the $\{z_j^*\}$ variables since the desired solution is a truth assignment (which will in turn determine which clauses are satisfied), but we will need to use properties of the LP solution to derive an approximation ratio.

  We can show that this approach will lead to a $1-(1-1/k)^k \geq 1-1/e$ approximation (in expectation) for the contribution of clauses having at least k literals since $(1 - 1/k)^k < 1/e$ (and converges to $1/e$ as k grows). Hence the approx ratio is $\geq 1-1/e > .632$. We will need one further idea to obtain the stated (3/4) ratio. And (as far as I can see) unlike the new randomized greedy result, this method can be derandomized (by th method of conditional expectation) to obtain a deterministic algorithm with the same 3/4 approximation ratio.

  NOTE: Here I am again expressing the approx ratio for Max Sat as a fraction < 1.

This bound is getting worse for large k. On the other hand the approximation ratio from the "naive randomized alg" is $1 - 2^{-k}$ for clauses with exactly k variables. and this gets better for large k. By taking the best of these two algorithms we can guarantee the stated 3/4 approximation ratio.

Let LP-OPT denote the optimal fractional solution value That is, LP-OPT $= \sum_j w_j z_j^*$ And let the rounded solution (a random variable since we are choosing the integral values $\bar{y}_i$ randomly and independently with probability y_i^*.

We want to show that E[rounded solution] $\geq (1 - 1/e)$ LP-OPT.

As stated above, we will show more specifically that for any clause C_j with k literals, the probability that C_j is satisfied (in the rounded solution) is at least $\beta_k z_j^*$ where $\beta_k = 1 - (1 - 1/k)^k$ and then as noted that $\beta_k \geq (1 - 1/e)$ for all k.

This will then imply the desired result by the linearity of expectations.

- Here is the analysis.

  We will need to make use of the arithmetic geometric mean inequality whcih states that for non negative real values,

  $\frac{a_1+a_2+...+a_k}{k} \geq (a_1 \cdot a_2 \ldots \cdot a_k)^{\frac{1}{k}}$
  or equivalently that
  $[\frac{a_1+a_2+...+a_k}{k}]^k \geq (a_1 \cdot a_2 \ldots \cdot a_k)$.

  Let C_j be a k literal clause and by renaming we can assume $C_j = x_1 \vee x_2 ... \vee x_k$.

  C_j is satisfied if not all of the y_i are set to 0 (when we set $y_i = 1$ with probability y_i^*).
  The probability that C_j is satisfied is then $1 - \Pi_{i=1}^k (1 - y_i^*)$.

  By the arithmetic-geometric mean inequality this probability is then at least
  $1 - (\frac{\sum_{i=1}^k (1-y_i^*)}{k})^k$
  $= 1 - (1 - \sum_{i=1}^k \frac{y_i^*}{k})^k$
  $\geq 1 - (1 - \frac{z_j^*}{k})^k$

  where the last inequality is by the LP constraint $\sum_{i \in C_j^+} y_i + \sum_{i \in C_j^-} (1 - y_i) \geq z_j$ (and keeping in mind the variable renaming making all literals positive).

  If one defines $g(z) = 1 - (1 - \frac{z}{k})^k$ then $g(z)$ is a concave function with $g(0) = 0$ and $g(1) = \beta_k$. By concavity, $g(z) \geq \beta_k z$ for all $0 \leq z \leq 1$.

  That ends the proof.

- We can then use the method of conditional expectations to obtain a deterministic algorithm. That is, so far we have a randomized algorithm that has a good expectation. So let $E[F]$ be this expectation. By conditional expectation we have E$[F]$ =E$[F|x_1 = 1]$ Prob $[y_i = 1]$ + E$[F|x_1 = 0]$ Prob $[y_i = 0]$

It follows that one of these two expectations is at least E$[F]$. We can then use the large expectation (or rather use the larger fractional opt we obtain when setting x_1 to 1 and 0) to determine how to set x_1. We can continue to do this so as to set all variables. Note that this entails calling an LP solver $O(n)$ times.

- Finally, we combine this method with the "naive randomized method" to yield a 3/4 approx for Max Sat. Namely, for the naive method, we have for a clause C with k literals E[rounded solution] $\geq \alpha_k w_C \geq \alpha_k w_C z_C^*$ since $z_C^* \leq 1$. So if we choose the naive method with probability 1/2 and the LP randomized method with probability 1/2, we get that the expected value obtained from clause C is $\geq \frac{\alpha_k+\beta_k}{2} w_C z_C^*$

  It is easy to verify that for $k = 1, 2$ that $(\alpha_k + \beta_k) = 3/2$ and for $k \geq 3$, that $(\alpha_k + \beta_k) \geq 7/8 + (1 - 1/e) \geq 3/2$.