Approximation Algorithm

Prof. Palash Dey

Department of Computer Science and Engineering

Indian Institute of Technology, Kharagpur

Week – 07

Lecture 35

Lecture 35 : Randomized Rounding Based (1-1/e) Factor Approximation Algorithm for MAX-SAT

welcome. So, in the last class we have seen a $\frac{\sqrt{5}-1}{2}$ factor approximation algorithm for MAX-SAT problem. So, in this class we will see a randomized rounding based algorithm for the same problem with better approximation guarantee. So, today's topic is rounding of LP's for MAX-SAT. So, we first write a linear programming relaxation, first let us write the integer linear program and then we will relax the integrality constraints. So, for each clause C_j we have a variable z_j which takes value 1 if C_j is satisfied.

and 0 otherwise ok. So, with this the objective function is easy to write, hence the objective maximize $\sum w_j z_j$. Next for every variable for every variable of the MAX-SAT formula we have a variable in linear program which will be set to true if the which will be set to 1 if the variable is set to true. So, for each Boolean variable x_i of max set, we have a variable y_i which takes value 1 if x_i is set to true and 0 otherwise.

So, now let us we will write the constraints. So, let us write down the integer linear programming formulation maximize $\sum w_j z_j$ subject to Now consider the j-th variable. So, for each variable we will have a constraint which basically says that is s$z_j$et to 1 only if at least one of its literals are set to true. So, let p_j be the set of variables appearing positively in clause C_j and n_j the set of variables appearing negatively. in clause C_j ok.

So, $C_j$will be satisfied if one of its variable which are appearing positively is set to true or any of its variable which is appearing negatively is set to true set to false. So, then only $z_j$will be allowed to take value 1. So, here is the constraint that I sum over y_i's i in p_j ok and $\sum 1-y_i$ see if any of the variable appearing positively and it is said to true in the corresponding y_i will be value will take value 1 or if any of the variable which are appearing negatively in n_j. are set to false then that y_i will be set to 0 and that $1-y_i$ this value will be 1.

So, this sum will take value 1 if or at least 1 if C_j is satisfied otherwise it take value So, z_j should be allowed to take value 1 if and only if this sum is greater than equal to 1. So, this is z_j ok yeah. Now, what are the y_i's? y_i's belong to $\{0,1\}$. for all $i \in [n]$ ok. Now, here is a little problem I want z_j also to take $\{0,1\}$ values, but you know if say 2 variables which are appearing positively in a clause C_j is said to true then the value of this sum is greater than equal to 2.

that will allow this that will allow to set z_j to value more than 1 and that is fine because that will even increase the objective function which we are trying to maximize. So, to prevent that we can make another constraint that z_j should take value less than equal to 1 and of course, it is greater than equal to 0 ok. So, this is the ILP formulation. Now, to go to the LP formulation, we relax this integrality constraint, we replace this integrality constraints with inequalities instead of demanding y_i to take value 0 or 1, we insist that y i take value in between 0 and 1. So, this is the relaxed LP ok.

So, what we have is opt equal to ILP of integer linear program with the constraints that y_i takes value either 0 and 1 and once we relax it we have a larger search space and hence the maximum value can only increase. So, this is less than equal to LP opt. So, LP opt is a upper bound on opt ok good. Now, what we do? for any rounding based algorithms what we first solve this relaxed LP get the optimal solution and use that as a guide to design a good approximation algorithm. So, solve the relaxed LP, let (y^*, z^*) be an optimal solution.

of this LP and before that before see how it can be used let us recall some inequalities which we will be needing. First I am writing them as fact, first is arithmetic geometric mean inequality AM-GM inequality. It says that for any n non-negative real numbers $a_1, \ldots, a_n$ the geometric mean which is product of a_i, i equal to 1 to n and then you take the nth root geometric mean is less than equal to arithmetic mean $\frac{1}{n}\sum_{i=1}^{n} a_i$ ok.

So, this we will use crucially and another fact if a function $f(x)$. $f(x)$ is concave on the interval closed interval $[0,1]$ what does that mean ok we will see and $f(0)=a$ $f(1)=b+a$, then $f(x) \geq a+bx$ for all $x \in [0,1]$. ok what is concave mean? So, $f(x)$ is concave on say $[0,1]$. The idea is you take an interval suppose this is $[0,1]$ take the function values at 0 take the function value at 1 and you draw this line.

If the graph of the function lies above it then it is concave function if this is a and this is b if $f(x) \geq a+bx$ and this not only for this function. So, this there is some problem with this picture. So, it should look like this. So, what is important is for any two point x_1, x_2 the graph of the function lies above it. So, if correct it if for all $0 \leq x_1 \leq x_2 \leq 1$.

So, for every interval $[x_1, x_2]$ for all $x \in [x_1, x_2]$ $f(x) \geq f(x_1) + (x_2 - x_1) f(x_2)$. So, it basically says that if I take any point x_1, x_2 on this graph and draw this straight line the graph in between x_1 and x_2 the graph of the function $f(x)$ in between x_1 and x_2 should lie above it. An equivalent condition for concavity if f is double differentiable then this double derivative should be less than equal to 0. So, if f is double differentiable, then f is concave if and only if $f^{'}(x)$ is concave in say any range say $[0,1]$, then $f^{''}(x) \leq 0$, in $[0,1]$.

So, with this now we will see a randomized rounding based technique for $1 - \frac{1}{e}$ factor approximation algorithm for MAX-SAT. So, here is theorem. there is a randomized rounding based algorithm for MAX-SAT. which achieves an approximation factor of at least $1 - \frac{1}{e}$. So, in randomized rounding based technique whenever you have variables which take value in between 0 and 1, it is quite natural to treat those variables as probabilities and used as those probabilities effectively.

Here we want to set some values true and false to the variables of the SAT instance. for each such variable x_i I have a variable y_i in linear program programming relaxation which takes value in between 0 and 1. So, what I do I can a natural thing is that I can set x_i to true with probability y_i. So, our algorithm is set x_i to true with probability y_i and false with probability $1 - y_i$ independent of everything else ok. So, what is the now again we will see the same analysis ALG is $w_1 z_1 + ... + w_m z_m$.

So, expectation of ALG is expectation of $w_1 z_1 + ... + w_m z_m$ we apply linearity of expectation $w_1 E[z_1] + ... + w_m E[z_m]$ and again these are indicator random variables their expectations are the probabilities of the events. So, this is w_1 probability that C_1 is satisfied probability that C_m is satisfied. So, we need to give lower bounds on the probability that any clause C_j is satisfied. So, what is the probability that so, I have a $j \in [m]$ probability that C_j is satisfied. Again you see the only way it is it can be unsatisfied is all its positive variables are set to false and all its negative variables are set to true.

So, this is 1 minus probability that C_j is not satisfied ok and this is nothing, but 1 minus product over the positive terms $i \in P_j$ that they are said to false and all are all the negative variables they are said to true. So, this is this is the only way it can be unsatisfied. So, this is greater than equal to 1 minus here I am using arithmetic geometric arithmetic mean AM-GM inequality that arithmetic mean is greater than equal to geometric mean. So, if I replace this geometric means ok.

So, I will apply. So, this is suppose let l_j is number of literals in C_j. So, this is 1 by what is arithmetic mean of these terms $\frac{1}{l_j}\sum_{i\in P_j}(1-y_i^*)$ So, this should be y_i^* because I have started with an optimal solution of the linear program $1-y_i^*$ plus $\sum_{i\in N_j} y_i^*$. So, this is the arithmetic mean and in geometric mean there is a $\frac{1}{l_j}$ term. Now, this $\frac{1}{l_j}$ I raise it to the power here.

This is how I apply arithmetic mean geometric mean a m g m inequality ok good. this is same as $1-(1-\frac{1}{l_j})\sum_{i\in P_j} y_i^*$. plus $\sum_{i\in N_j} 1-y_i^*$. So, if added and subtracted 1 in both side to the power l j ok. Now, you see that this one this sum this is the LP constraint that this is greater than equal to z_j.

So, what we have is because there are 2 negative terms here is 1 and here is 1 the direction of inequality remains same $(1-\frac{1-z_j}{l_j})^{l_j}$ ok. And now you consider the function f this is also z_j^* consider the function z_j^* to be $(1-\frac{1-z_j}{l_j})^{l_j}$ ok. So, this is concave you check this is concave in or on interval $[0,1]$, it is a function of z_j^* for every $l_j\geq 1$, every natural number l_j. So, then we use that fact and we can write that this probability at C_j is satisfied this is greater than equal to this whole thing is greater than equal to. So, if z_j equal to 0 then this is $f(0)$ is what is $f(0)$? $f(0)$ is 0 and $f(1)$ is $1-(1-\frac{1}{l_j})^{l_j}$ ok.

So, $f(z_j^*)\geq f(0)+f(1)z_j^*$. So, this is greater than equal to $(1-\frac{1}{e})z_j^*$. So, this is what we have got probability that C_j is satisfied is greater than equal to $(1-\frac{1}{e})z_j^*$. Now, we put this here expectation of ALG. So, expectation of ALG is then greater than equal to $1-\frac{1}{e}$ and here we have $w_1 z_1^*+\ldots+w_m z_m^*$ ok.

this is LP opt because you see this is the optimization function and z star y star z star is a is an optimal solution. So, this is equal to $\left(1-\frac{1}{e}\right)LP-opt$ which is greater than equal to $\left(1-\frac{1}{e}\right)opt$. hence it is a $1-\frac{1}{e}$ factor randomized approximation algorithm. Again this can be de-randomized using method of conditional expectation I leave that as a homework ok. Thank you.