The $(1-1/e)$ approximation via randomized rounding of an LP

We are taking these notes from Vijay Vazirani's text " Approximation Algorithms".

- **(The algorithm)**

  We consider the weighted Max-Sat problem (for an arbitrary CNF formula F). Here we will use randomization in a natural way when we have an 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 $\sum_{i \in C_j^+} y_i + \sum_{i \in C_j^-} (1 - y_i) \geq z_j$

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

  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. Let $\{y_i^*, z_j^*\}$ be an optimal LP solution. Then we set $y_i' = 1$ with probability y_i^* to obtain an integral solution (and hence truth assignment). 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 do need to use properties of the LP solution to derive an approximation ratio.

  We show that this approach leads to a $1 - (1 - 1/k)^k \geq 1 - 1/e$ approximation (in expectation) for the contribution of clauses having k literals since $(1 - 1/k)^k < 1/e$ (and converges to $1/e$ as k grows). Hence the approx ratio is $\geq 1 - 1/e > .632$.

  This method can be derandomized (by the method of conditional expectation) to obtain a deterministic algorithm with the same $1 - 1/e$ approximation ratio. Note that this derandomization entails calling an LP solver $O(n)$ times. We need one further idea to obtain a (3/4) approximation ratio. Namely, we take the maximum of the $1 - 1/e$ algorithm and the (de-randomized) "naive" algorithm (that sets all variables to true or false randomly with probability = 1/2 .)

Let LP-OPT denote the optimal fractional solution value That is, LP-OPT = $\sum_j w_j z_j^*$ Let Y' be the weight of the rounded solution (a random variable since we are choosing the integral values y_i' randomly and independently with probability y_i^*).

We want to show that i$E[Y'] \geq (1 - 1/e)$ 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.

- (**The analysis**)

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

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

  Let C_j be a k literal clause and by renaming we can simplify the discussion by assuming $C_j = x_1 \vee x_2 \ldots \vee x_k$. Note: we are fixing a particular k literal clause and doing the analysis of its expected contribution. We do this analysis for each clause independently and use linearity of expectations to add up the contributions for each clause.

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