(https://accounts.coursera.org/i/zendesk/courserahelp?return_to=https://learner.coursera.help/hc)
Lessons

Assignment: Assignment 1

Pass the exercise
(/learn/approximation-algorithms-part-2/peer/QHUvH/assignment-1/submit)

Review 3 classmates
0/3 reviews completed
(/learn/approximation-algorithms-part-2/peer/QHUvH/assignment-1/give-feedback)

Instructions (/learn/approximation-algorithms-part-2/peer/QHUvH/assignment-1)

My submission (/learn/approximation-algorithms-part-2/peer/QHUvH/assignment-1/submit)

Review classmates (/learn/approximation-algorithms-part-2/peer/QHUvH/assignment-1/give-feedback)

Discussions (/learn/approximation-algorithms-part-2/peer/QHUvH/assignment-1/discussions)

This is a preview of your project. Changes have been saved but not submitted.

Your assignment was saved, but please fix the errors below before submitting.

Please fill in a title.

February 6, 2016

A Primal-Dual Algorithm for Set Cover.

In this exercise, we propose to design a primal-dual algorithm for the set cover problem.

The set cover problem is as follows: given a set of elements $E = \{e_1, \ldots, e_n\}$, some subsets of those elements $S_1, S_2, \ldots, S_m \subseteq E$, and a non-negative weight w_j for each subset S_j. The goal is to find a minimum-weight collection of subsets that contains all the elements of E. Namely, we want to find a collection I of subsets that minimizes $\sum_{j \in I} w_j$ and such that subject to $\bigcup_{j \in I} S_j = E$.

Throughout the exercise, we will consider the following linear program LP for the problem.

$\min \sum_{j=1}^{m} x_j \cdot w_j$

subject to,

$\forall i \in \{1, \ldots, n\}, \quad \sum_{j\,:\,e_i \in S_j} x_j \geq 1$

$\forall j \in \{1, \ldots, m\}, \quad x_j \geq 0$

Question 1: What is the dual of this linear program?

We now consider the following primal-dual algorithm.

1. $y \leftarrow 0$
2. $I \leftarrow \emptyset$
3. While I is not a solution (there exists $e_i \notin I$):

- Increase the dual variable y_i of an element e_i that is not covered until there exists an l such that $\sum\limits_{j: e_j \in S_l} y_j = w_l$
- Add the set S_l to I

4. Return I

Correctness.

Question 2: In how many iterations of the while loop can a given dual variable be increased?

Question 3: Using Question 2, argue that the algorithm terminates and so, that the output I is a solution to the problem.

Approximation Ratio. In this section, we assume that each element of the set E appears in at most f sets of $S_1, \ldots, S_m$.

Question 4: Recall a tight lower bound between the value of the optimal fractional solution for the dual val($y*$) and the value of the optimal integral solution for the set cover problem OPT.

Question 5: Argue that the solution y is feasible for the dual.

Question 6: Combining Questions 4 and 5, recall a tight lower bound between the value of the solution y and the value of the optimal integral solution for the set cover problem OPT.

In the following, we want to show

$\sum_{j:S_j \in I} w_j \leq f \cdot \mathrm{val}(y)$.

Question 7: Consider a set $S_j \in I$. What is the relationship between w_j and $\sum_{i:e_i \in S_j} y_i$?

Question 8: Using Question 7, give the relationship between $\sum_{j \in I} w_j$ and the variables y_i.

Question 9: Recall that $|\{j \;:\; e_i \in S_j\}| \leq f$ for all i. Using Question 8, prove that $\sum_{j:S_j \in I} w_j \leq f \cdot \mathrm{val}(y)$.

Question 10: Conclude using Questions 6 and 9.

Answers

1. Dual: max $\sum_{i=1}^{n} y_i$

subject to,

$\forall j \in \{1, \ldots, m\}, \quad \sum_{i:e_i \in S_j} y_i \leq w_j$

$\forall i \in \{1, \ldots, n\}, \quad y_i \geq 0$

2. A given dual variable y_i corresponds to a given element e_i. If e_i is still uncovered, let's consider all the sets $S_k \notin I$ and $e_i \in S_k$ (There must exist at least one such set S_k. If not, then e_i can never be covered by the sets $S_1 \ldots S_m$). Now, in the while loop, when the variable y_i is incremented the value $\sum_{j:e_j \in S_k} y_j$ gets incremented for all those sets S_k (since $y_i \in S_k, \; \forall k$). But each S_k is upper-bounded by a w_k (which is finite, w.l.o.g.), so we can't go on increasing y_i indefinitely while being in the feasible region. Hence, there will be some $l \in k$ for which the constraint $\sum_{i:e_i \in S_k} y_i \leq w_k$ will become tight in the same iteration and it will take just one iteration for a particular element e_i after which the corresponding set S_l will be added to I.

3. Since, as argued in 2, each iteration of the while loop covers an element e_i yet to be covered (and the corresponding set covering the element is added to I), it will take at most n iterations in the worst case to cover all the n elements in E. Hence, the while loop will terminate after at most n iterations with all the n elements covered in the solution .

4. By Weak Duality, we have $val(y*) \leq val(x*)$ where $x*$ is the fractional optimal solution for **Primal** and $y*$ is the fractional optimal solution for the **Dual**. Also, the **OPT** is going to be the optimal integral solution for the **Primal minimization** problem $\Rightarrow val(x*) \leq OPT$. Combining, we have $val(y*) \leq OPT$.

5. Since the while loop guarantees that each of the constraints of the dual are satisfied (all the y_i variables are non-negative and the constraints are at most tight, s.t., $\sum_{i:e_i \in S_j} y_i \leq w_j$ is satisfied $\forall i$, with equality for the sets S_l that are added to I from inside the while loop). Hence the solution y remains feasible.

6.Since y is a feasible solution and $val(y*)$ is the optimal solution for the **Dual maximization** problem, we have $val(y) \leq val(y*)$. Combining with 4 and 5, we have, $val(y) \leq OPT$.

7. For any set $S_j \in I$, we shall have $\sum_{i:e_i \in S_j} y_i = w_j$, as guaranteed by the while loop.

8. Hence, $\sum_{j:S_j \in I} w_j = \sum_{j:S_j \in I} \sum_{i:e_i \in S_j} y_i = \sum_{i=1}^{n} |j : e_i \in S_j| \; y_i$, since all the elements are there in I, with each of them probably present multiple times in multiple sets. Here $|j : e_i \in S_j|$ present the number of times the element e_i was present in I.

9. Also, given, $|j : e_i \in S_j| \leq f$. Hence, we have,
$\sum_{j:S_j \in I} w_j = \sum_{i=1}^{n} |j : e_i \in S_j| \; y_i \leq \sum_{i=1}^{n} f.y_i = f.\sum_{i=1}^{n} y_i = f.val(y) \Rightarrow \sum_{j:S_j \in I} w_j \leq f.val(y)$.

10. Combining $OPT \geq val(y)$ (from 6) and the value of the solution provided by the **Primal-Dual algorithm** $\leq f.val(y)$ (from 9), we get the **approximation ratio** for this algorithm $= f$.

Edit (peer/QHUvH/assignment-1/submit) Submit for review

Help Center