Approximation Algorithms

Lecture 2

Max. Matching in P

efficient linear time algorithm that $\frac{1}{2}$-approximates

- Last time:
  - Introduction to approximation algorithms
  - Unweighted SET COVER and MAX COVERAGE: Greedy algorithms with $\ln n + 1$ and $1 - 1/e$ approximation guarantees respectively
  - f-approximation for Weighted SET COVER via deterministic rounding of LP solution
    - f is the maximum number of sets that any element is part of

Why are poly-time algorithms with additive approximation guarantees impossible to obtain in general unless P = NP?
Will upload a short video with a proof for the case of TSP

Today:

- SET COVER continued
  - Dual rounding
  - Primal dual method

Recall

- Weighted SET COVER
  - Given universe U of n elements $e_1, \dots, e_n$ and family $F = \{F_1, \dots F_m\}$ of subsets of U, where F_i has weight w_i for $i \in [m]$, output a min-weight subset of F covering the elements in U

nonnegative

LP:

$$\min \sum_{i \in [m]} w_i x_i$$

$$\sum_{i \,:\, e \in F_i} x_i \geq 1 \qquad \forall e \in U \qquad (P)$$

$$x_i \geq 0$$

- View the SET COVER problem from the following perspective
  - Elements are individuals and sets in $\boldsymbol{F}$ are social groups
  - Weight of a set is the total social status of that group

nonnegative

  - Want to charge elements a price to "be covered" by a set
  - Maximize the total price paid by all elements
  - Sum of prices paid by elements in a set is at most the weight of the set
- We now write another LP related to (P)
  - Let y_e for $e \in U$ denote the price paid by an element

- LP corresponding to the "price charging problem"

Primal objective fn value

Dual obj.

$$\max \sum_{e \in U} y_e$$

(D)

$$\sum_{e : e \in F_j} y_e \leq w_j \quad \forall j \in [m]$$

$$y_e \geq 0 \quad \forall e \in U$$

- This latter LP (D) is called the "dual" of the original LP (P)

- There is a mechanical way to write the dual of any linear program – we will not cover it in this course
- But you will see sufficient examples where we write duals to LPs

  - Reading on LPs: Appendix A of Williamson & Shmoys

- Dual has a variable for each constraint in primal
- Dual has a constraint for each primal variable

- A dual program has interesting connections to its primal program

feasible

- **Weak Duality:** For any solution $\{x_i\}_{i\in[m]}$ to the primal program and any feasible solution $\{y_e\}_{e\in U}$ to the dual program

$$\sum_{e\in U} y_e \leq \sum_{i\in[m]} w_i x_i$$

Proof:

$$\sum_{e\in U} y_e = \sum_{e\in U} y_e \cdot 1 \le \sum_{e\in U} y_e \cdot \Big(\sum_{\substack{i\in[m]:\\ e\in F_i}} x_i\Big) = \sum_{i\in[m]} x_i \underbrace{\sum_{\substack{e\in U:\\ e\in F_i}} y_e}_{\le w_i} \le \sum_{i\in[m]} x_i w_i$$

$$\sum_{\substack{i\in[m]:\\ e\in F_i}} x_i \ge 1$$

Used the fact that we have feasible solutions.

- **Corollary**: Dual optimal value is a lower bound on primal optimal value!

- **Strong Duality Theorem:** If both primal and dual have feasible solutions, then their optimal values are equal.

  If $\{x_i^*\}_{i\in[m]}$ and $\{y_e^*\}_{e\in U}$ are primal optimal and dual optimal solutions, then

$$\underbrace{\sum_{e\in U} y_e^*}_{\text{Value of dual optimal solution}} = \underbrace{\sum_{i\in[m]} w_i x_i^*}_{\text{Value of primal optimal solution.}}$$

f-approximation algorithm from the optimal solution for the dual LP

Algorithm

- Let $\{y_e^*\}_{e \in U}$ be the dual optimal solution
- For $i \in [m]$:
  - if the dual constraint for F_i is such that $\sum_{e \in F_i} y_e^* = w_i$, then add F_i to the cover

- Why does this output a set cover?
- What is the approximation guarantee?

- **Claim:** Output is a valid set cover

$\forall F_j$ s.t. $e \in F_j$

$\rightarrow \sum_{e' \in F_j} y^*_{e'} < w_j$

- **Proof:** by contradiction

- If an element $\boldsymbol{e}$ is not covered, then for every set $\boldsymbol{F_j}$ containing $\boldsymbol{e}$, the dual constraint is not "tight".

- Can increase the value of $\boldsymbol{y_e^*}$ until the first such constraint is tight!

$\sum_{e' \in U} y^*_{e'}$

- This increases the dual solution value without violating any dual constraints

- Contradiction!

- **Claim:** Solution returned is an f-approximation to SET COVER
- Proof:

set as part of solution

F_1, F_{100}, F_{29}

- For each F_i picked, we know that dual constraint is tight

$W_1 + W_{100} + W_{29}$

$$\sum_{e \in F_i} y_e^* = w_i$$

- Weight of solution output = Sum of w_i's for the F_i's picked as part of cover
  = Sum of corresponding y_e^*'s
- Last sum is at most $f \cdot \sum_{e \in U} y_e^*$, since each dual variable belongs to at most f constraints

weak duality

primal LP is a relaxation of SET COVER IP

- By weak duality:

$$f \cdot \sum_{e \in U} y_e^* \leq f \cdot \sum_{i \in [m]} x_i^* \leq f \cdot \text{OPT}$$

$= w_j$

S – cover output

$$\sum_{F_j \in S} \sum_{e \in F_j} y_e^*$$

Complementary Slackness

- We know that:

$$\sum_{e\in U} y_e \le \sum_{e\in U} y_e \cdot \left(\sum_{j:e\in F_j} x_j\right) = \sum_{j\in[m]} x_j \cdot \left(\sum_{e\in F_j} y_e\right) \le \sum_{j\in[m]} x_j w_j$$

dual is feasible

dual obj. fn. value

strong duality holds

primal obj. fn. value

- For optimal solutions $\{x_i^*\}_{i\in[m]}$ and $\{y_e^*\}_{e\in U}$, all these inequalities are equalities!

$$\begin{cases} y_e^* > 0 \implies \sum_{j:e\in F_j} x_j^* = 1 & \forall e \in U \\ x_j^* > 0 \implies \sum_{e\in F_j} y_e^* = w_j & \forall i \in [m] \end{cases}$$

$$\sum_{e \in U} y_e^* = \sum_{e \in U} y_e^* \cdot \left(\sum_{i : e \in F_i} x_i^* \right)$$

Whenever $y_e^* > 0$ $\quad \sum_{i : e \in F_i} x_i^* = 1$

Converse is also true!

If two feasible solns. satisfy complementary slackness, then they are optimal

Complementary Slackness

Converse is also true: Feasible solutions satisfying complementary slackness are optimal

- We know that:

$$\sum_{e\in U} y_e \leq \sum_{e\in U} y_e \cdot \left(\sum_{j:e\in F_j} x_j\right) = \sum_{j\in[m]} x_j \cdot \left(\sum_{e\in F_j} y_e\right) \leq \sum_{j\in[m]} x_j w_j$$

- For optimal solutions $\{x_i^*\}_{i\in[m]}$ and $\{y_e^*\}_{e\in U}$, all these inequalities are equalities!

$$y_e^* > 0 \implies \sum_{j:e\in F_j} x_j^* = 1 \quad \forall e \in U$$

$$x_j^* > 0 \implies \sum_{e\in F_j} y_e^* = w_j \quad \forall i \in [m]$$

- In our algorithm in last class, we added a set F_i to the solution if $x_i^* \geq 1/f$

$x_i^* > 0$

- By complementary slackness, this implies that, for each set F_i added to the set cover (by the algorithm in last class),

$$\sum_{e \in F_i} y_e^* = w_i$$

- So, this same set is also added to the solution by the dual rounding algo as well

Suppose $I \subseteq [m]$ is the collection of sets picked by primal rounding algo.

$I' \subseteq [m]$ " " " " " dual rounding algo

$\boxed{I \subseteq I'}$

- Did we need to first compute a dual optimal solution in the algorithm today?
- Computing the optimal solution to an LP is an expensive operation

~~Only~~ property of dual solution the we used in today's algo:

(i) $\sum_{e \in U} y_e \leq OPT$

(ii) $\{y_e\}_{e \in U}$ is a feasible solution

- Did we need to first compute a dual optimal solution in the algorithm today?

dual feasible solution

1. Begin with $y_e \leftarrow 0$ for $e \in U$ and $C \leftarrow \emptyset$
2. While C is not yet a set cover: $\exists$ uncovered element e
   1. Increase the dual variable y_e for some uncovered element $e \in U$ until some dual constraint, say, for F_i, goes tight, where $e \in F_i$
   2. Add F_i to C
1. Output C

Let $\varepsilon = \min_{j:\, e \in F_j} \left(w_j - \sum_{e' \in F_j} y_{e'} \right)$

$y_e \leftarrow y_e + \varepsilon$

- Did we need to first compute a dual optimal solution in the algorithm today?

1. Begin with $y_e \leftarrow 0$ for $e \in U$ and $C \leftarrow \emptyset$

2. While C is not yet a set cover:
    1. Increase the dual variable y_e for some uncovered element $e \in U$ until some dual constraint, say, for F_i, goes tight, where $e \in F_i$
    2. Add F_i to C

1. Output C

Exercise:

Show that this algorithm has the same guarantees as the dual rounding algorithm

- **Reading Exercise**: Section 1.6 from Williamson & Shmoys

  → Theorem 1.11

- $\ln n$ factor approx. greedy algorithm for weighted SET COVER

  ↓ H_n - n^{th} harmonic number

- Generalization of the greedy algorithm seen in last lecture

- **Next lecture**: Randomized rounding for SET COVER

  Dual fitting