Approximation Algorithms,
M. QUEYRANNE

Lecture 3
August 5, 2009

Vijay VAZIRANI, Approximation Algorithms, Springer 2001

Lecture 1: Chap 1, 3

Lecture 2: Chap 3, 12

Lecture 3: Toshihiro FUJITO, "Approximation Algorithms for Submodular Set Cover with Applications", IEICE TRANS. INF. & SYST. E83-D, 480-487, 2000

see also Vazirani Chaps 2 (except 2.3), 13, 15
Chap 14

Recall the Submodular Set Cover (SSC) problem

given a finite ground set N, costs $c_j > 0 \quad \forall j \in N$

nondecreasing submodular set function $f : 2^N \to \mathbb{R}$

find a subset $S \subseteq N$ with $f(S) = f(N)$ and least cost $c(S) = \sum_{j \in S} c_j$

$$\min \{ c(S) : S \subseteq N \text{ and } f(S) = f(N) \}$$

Integer Programming Formulation

For $S \subseteq N$ the contraction f_S of f onto $N \setminus S$ is $f_S : 2^{N \setminus S} \to \mathbb{R}$

defined by $f_S(X) = f(X \cup S) - f(S) \quad \forall X \subseteq N \setminus S$

let $f_S(j) := f_S(\{j\}) = f(S \cup \{j\}) - f(S)$ marginal increase in f when adding j to S

(since f is submodular $f_S(j) \geq f_T(j) \quad \forall S \subseteq T \subseteq N \setminus \{j\}$)

Let $x_j = \begin{cases} 1 & \text{if we choose } j \text{ in the solution} \\ 0 & \text{o/w} \end{cases}$

$$(IP) \begin{cases} \min & \sum_{j \in N} c_j x_j \\ \text{s.t.} & \sum_{j \in N \setminus S} f_S(j) x_j \geq f_S(N \setminus S) \quad \forall S \subseteq N \\ & 0 \leq x_j \leq 1, \text{ integer} \quad \forall j \in N \end{cases}$$

example: $N = \{1,2\}$

$$\begin{cases} \min & c_1 x_1 + c_2 x_2 \\ \text{s.t.} & f_\emptyset(1) x_1 + f_\emptyset(2) x_2 \geq f_\emptyset(N) \qquad S = \emptyset \\ & \qquad\qquad f_{\{1\}}(2) x_2 \geq f_{\{1\}}(2) \qquad S = \{1\} \\ & f_{\{2\}}(1) x_1 \qquad\qquad \geq f_{\{2\}}(1) \qquad S = \{2\} \\ & 0 x_1 + 0 x_2 \geq 0 \qquad S = \{1,2\} \\ & 0 \leq x_1, x_2 \leq 1, \text{ integer} \end{cases}$$

<u>Lemma:</u> (IP) is a valid formulation for SSC

proof: For any $T \subseteq N$ let x^T be its characteristic vector ($x^T_j = 1$ if $j \in T$, 0 o/w)

If T is a feasible solution to SSC, $f(T) = f(N)$

For any $S \subseteq N$ $\sum_{j \in N \setminus S} f_S(j) x_j^T = \sum_{j \in T \setminus S} f_S(j)$

$\geq f_S(T \setminus S)$ because f is submodular

$= f(T) - f(S)$ by def. of f_S

$= f(N) - f(S)$ because T feasible

$= f_S(N \setminus S)$

so x^T is a feasible solution to (IP)

Conversely, let x^T be a feasible solution to (IP)

$$0 = \sum_{j \in N \setminus T} f_T(j) x_j^T \underset{(IP)}{\geq} f_T(N \setminus T) = f(N) - f(T) \Rightarrow f(T) = f(N)$$

so T is feasible to SSC QED

Let (LP) denote the LP relaxation of (IP), relaxing also $x \leq 1$. Its LP dual is

$$(D) \begin{cases} \max \sum_{S \subseteq N} f_S(N \setminus S) y_S \\ \text{s.t.} \quad \sum_{S: j \notin S} f_S(j) y_S \leq c_j \quad \forall j \in N \\ \qquad y_S \geq 0 \end{cases}$$

Greedy algorithm for SSC (a "Dual Fitting algorithm" in disguise!)

- start with $F = \emptyset$
- add to F elements that are "most attractive":

$$\min_{j \in N \setminus F} c_j / f_F(j) \qquad \text{"bang for buck ratio"}$$

Remark: if $f_F(j) = 0$ then this ratio $= +\infty$ (since $c_j > 0$)

Formally: $F := \emptyset$

while $f(F) < f(N)$ do

- choose $j \in \arg\min \left\{ c_j / f_F(j) : j \in N \setminus F \text{ with } f_F(j) > 0 \right\}$
- $F := F \cup \{j\}$.

Remark: for the Vertex Cover special case of SSC, this Greedy algorithm coincides with the Max Degree Greedy algorithm (since $c_j = 1$ and $f_F(j) =$ degree of node j in current graph induced by $N \setminus F$)

Recall the harmonic numbers $H(n) = \sum_{i=1}^{n} \frac{1}{i}$

Theorem: If f is integer valued, nondecreasing and submodular and $f(\emptyset) = 0$ then the approximation ratio of the Greedy algorithm for SSC is smaller than $H\left(\max_{j \in N} f(j)\right)$

Lemma: Let $0 < a_1 \leq a_2 \leq .. \leq a_T$ be T real numbers
$d_1 \geq d_2 \geq \ldots \geq d_T \geq 1$ be T integers
then $A := a_1 d_1 + \sum_{t=1}^{T-1} (a_{t+1} - a_t) d_{t+1} \leq \left(\max_k a_k d_k\right) H(d_1)$

proof of Lemma: $a_t \leq \left(\max_k a_k d_k\right) \frac{1}{d_t} \quad \forall t = 1 \ldots T$

regrouping $A = \sum_{t=1}^{T-1} a_t (d_t - d_{t+1}) + a_T d_T$

$$\leq \left(\max_k a_k d_k\right) \left[\sum_{t=1}^{T-1} \left(1 - \frac{d_{t+1}}{d_t}\right) + 1 \right]$$

If $d_t > d_{t+1} \geq 1$, $1 - \frac{d_{t+1}}{d_t} = (d_t - d_{t+1})\frac{1}{d_t} \leq \underbrace{\frac{1}{d_{t+1}+1} + \frac{1}{d_{t+1}+2} + \dots + \frac{1}{d_t}}_{\substack{d_t - d_{t+1} \text{ terms} \\ \text{each } \geq \frac{1}{d_t}}}$

therefore $\sum_{t=1}^{T}\left(1 - \frac{d_{t+1}}{d_t}\right) + 1 \leq \frac{1}{d_1} + \frac{1}{d_1+1} + \dots + 1 = H(d_1)$ QED

Proof of Theorem: index the sets F and elements j in the algorithm as $F_0 = \emptyset$ $F_1 = \{j_1\}$ $F_2 = \{j_1, j_2\}$ .... F_T is the SSC returned by the algorithm

for any $j \in N$ the marginal values $f_{F_t}(j) \geq 0$ because f nondecreasing and $(f_{F_{t-1}}(j))_{t=1..T}$ form a nondecreasing sequence because f submodular

Let $\theta_t := \min_{j \in N \setminus F_{t-1}} c_j / f_{F_{t-1}}(j) > 0$

$$\theta_t \leq \min_{j \in N \setminus F_t} c_j / f_{F_{t-1}}(j) \leq \min_{j \in N \setminus F_t} c_j / f_{F_t}(j) = \theta_{t+1}$$

so the $(\theta_t)_{t=1..T}$ form a nondecreasing sequence

By the Lemma, for all $j \in N$

$$A := \theta_1 f_{F_0}(j) + \sum_{t=1}^{T-1}(\theta_{t+1} - \theta_t) f_{F_t}(j)$$

$$\leq \left(\max_k \theta_k f_{F_{k-1}}(j)\right) H(f_{F_0}(j))$$

$$= \left(\qquad \qquad \right) H(f(j))$$

$$\leq c_j H(f(j)) \leq c_j \max_k f(k) \qquad (1)$$

Define a solution y to (D) by $y_{F^0} = \theta^1 / H(\max_k f(k))$

$y_{F^t} = (\theta_{t+1} - \theta_t) / H(\max_k f(k))$ for $t = 1 .. T-1$

all other $y_S = 0$

Then y is a feasible solution to (D) [by (1) divided by $\max_k f(k)$]

The objective value of the Greedy solution F^T is

$$c(F^T) = \sum_{t=1}^{T} c_{j_t} = \sum_{t=1}^{T} \theta^t f_{F^{t-1}}(j_t)$$

$$= \sum_{t=1}^{T} \theta^t \left(f(F^t) - f(F^{t-1}) \right)$$

$$= \theta^1 \left(f(N) - f(F^0) \right) + \sum_{t=2}^{T} \left(\theta^t - \theta^{t-1} \right) \left(f(N) - f(F^{t-1}) \right)$$

$$= H(\max_k f(k)) \left[y_{F^0} f_{F^0}(N \setminus F^0) + \sum_{t=2}^{T} y_{F^{t-1}} (N \setminus F^{t-1}) \right]$$

$$= H(\max_k f(k)) \sum_{S \subseteq N} f_S(N \setminus S)\, y_S$$

$$\leq H(\max_k f(k))\ OPT_{SSC}$$ by weak LP duality and relaxation

QED

Remark:

For IC $\quad H(\max_k f(k)) \leq H(\max_k \sum_{i=1}^{m} a_{ik})$ approximately, the log of the largest column sum in A

since for IC, $f(k) = \sum_{i=1}^{m} \min\{a_{ik}, b_i\} \leq a_{ik}$

For Set Cover: $H(\max_k f(k)) \leq H(\max_{S \in \mathcal{S}} |S|)$ approx. log of the largest cardinality of a given subset

For Vertex Cover $\quad$ ——— $\leq H(\max_{k \in V} \text{degree}(k))$ approx. log. of the maximum degree

Dual Greedy algorithm - a Primal Dual algorithm

Main Phase

- initialize $y := 0$, $F := \phi$, $\ell := 0$
- while $f(F) < f(N)$ do

  $\ell := \ell + 1$

  $$j_\ell := \operatorname{argmin}_{j \in N \setminus F} \frac{c_j - \sum_{\substack{S : j \notin S \\ S \neq F}} f_S(j) y_S}{f_F(j)}$$ (element of N for which the dual constraint becomes binding)

  let $$y_F := \frac{c_{j_\ell} - \sum_{\substack{S : j_\ell \notin S \\ S \neq F}} f_S(j_\ell) y_S}{f_F(j_\ell)}$$ (the maximum value of y_F that keeps y feasible)

  $F := F \cup \{j_\ell\}$

(Optional) Second Phase: ("cleanup", to insure the solution F is minimal)

for $k = \ell$ downto 1 do

if $f(F \setminus \{j_k\}) = f(N)$ then $F := F \setminus \{j_k\}$

Return F

Note that the algorithm returns a SSC F

Recall that for any $S \subseteq N$, $f_S : 2^{N \setminus S} \to \mathbb{R}$ is nondecreasing, submodular with $f(\phi) = 0$

so we can define a SSC problem on $N \setminus S$ with function f_S

$$\min_{T \subseteq N \setminus S} \{ c(T) : f_S(T) = f_S(N) \}$$

which is the remaining SSC problem if we fix S in the solution

(i.e., the best way of extending S to a SSC for the original problem)

<u>Theorem:</u> The cost of the SSC returned by the Dual Greedy algorithm

is at most $$\max_{\substack{S \subseteq N \\ X \text{ a SSC} \\ \text{for } (N \setminus S, f_s)}} \frac{\sum_{j \in X} f_s(j)}{f_s(N \setminus S)} \, OPT$$

- For Integer Cover where $N = \{1..n\}$ set of column of matrix A
  $M = \{1..m\}$ —— row —— —— ——

$$f(S) = \sum_{i=1}^{m} \min\{a_i(S), b_i\} \quad \text{where } a_i(S) = \sum_{j \in S} a_{ij}$$

for any $S \subseteq N$ let $M^S = \{i \in M : a_i(S) < b_i\}$ the set of rows not covered by the partial solution S

$$\text{then } f_s(N \setminus S) = \underbrace{\sum_{i \in M} b_i}_{f(N)} - \underbrace{\left(\sum_{i \in M^S} a_i(S) + \sum_{i \in M} b_i\right)}_{f(S)}$$

$$= \sum_{i \in M^S} (\underbrace{b_i - a_i(S)}_{>0,\ \text{integer}}) \geq |M_S|$$

for any set cover X in $(N \setminus S, f_s)$ and any $j \in N \setminus S$

$$f_s(j) \leq \sum_{i \in M^S} a_{ij}$$

$$\text{so } \sum_{j \in X} f_s(j) \leq \sum_{j \in X} \sum_{i \in M^S} a_{ij} = \sum_{i \in M^S} \sum_{j \in X} a_{ij}$$

$$\underset{\text{all } a_{ij} \geq 0}{\leq} \sum_{i \in M^S} \sum_{j \in N} a_{ij} \leq |M^S| \underbrace{\max_{i \in M} \sum_{j \in N} a_{ij}}_{\text{max } \underline{\text{row}} \text{ sum of } A}$$

the approximation bound in the Theorem

$$\max_{S, X} \frac{\sum_{j \in X} f_s(j)}{f_s(N \setminus S)} \leq \max_{i \in M} \sum_{j \in N} a_{ij}$$ max <u>row sum</u> bound

- For Set Cover, this maximum now sum is the maximum frequency f of an element in the given sets
- For Vertex Cover, $f = 2$ and the Dual Greedy algorithm is a 2-approximation!

Proof of Theorem

Note that when the j_e-constraint becomes binding, and thereafter

$$\sum_{S : j \notin S} f_S(j)\, y_S = c_j$$

The cost of the returned SSC F_T is

$$c(F^T) = \sum_{j \in F_T} c_j = \sum_{j \in F_T} \sum_{S : j \notin S} f_S(j)\, y_S$$

$$= \sum_{S \subseteq N} \left(\sum_{j \in F_T \setminus S} f_S(j) \right) y_S$$

$$= \sum_{t=0}^{T-1} \left(\sum_{j \in F_T \setminus F_t} f_{F_t}(j) \right) y_{F_t}$$

$$\leq \left(\max_{t = 1..T-1} \frac{\sum_{j \in F_T \setminus F_t} f_{F_t}(j)}{f_{F_t}(N \setminus F_t)} \right) \sum_{t=1}^{T-1} f_{F_t}(N \setminus F_t)\, y_{F_t}$$

$$\leq \left(\qquad\qquad \right) OPT$$

$$\leq (\text{max ratio in Thm})\, OPT$$

QED

See Fujito for other covering problems (e.g., with capacity constraints) that can be modeled as SSC problems

Approximation Algorithms based on Solving a Primal LP relaxation

given an integer programming formulation (or relaxation) of a combinatorial optimization problem
- solve the LP relaxation
- use ("round") the LP solution to get a feasible (integer) solution to the original combinatorial optimization problem.

Weighted Set Covering problem

given a ground set M
a collection $\mathcal{S}$ of subsets $S \subseteq N$, with weights (costs) $c_j > 0$

find a covering of M from $\mathcal{S}$ with least total cost

IP formulation: let $x_S = \begin{cases} 1 & \text{if } S \text{ is chosen} \\ 0 & \text{o/w} \end{cases}$

$$SC \begin{cases} \min \sum_{S \in \mathcal{S}} c_S x_S \\ \text{s.t.} \sum_{\substack{S \in \mathcal{S}: \\ i \in S}} x_S \geq 1 \quad \forall i \in M \\ x \geq 0, \text{ integer} \end{cases}$$

Let x^{LP} denote an optimum solution to LP relaxation

- **Rounding Method 1**: round every x_S^{LP} up $\quad x_j^1 = \lceil x_j^{LP} \rceil$

  this defines a feasible solution

  its cost could be much bigger than the optimum...

Recall that f is the maximum frequency of an element i in the given sets $S \subseteq \mathcal{S}$: $\quad f = \max_{i \in M} |\{S \in \mathcal{S} : i \in S\}|$

- **Rounding Method 2**: round every $x_S^{LP} \geq 1/f$ up to $x_S^2 = 1$

  every other x_S^{LP} down to $x_S^2 = 0$

for every i $\quad \underbrace{\sum_{S:i\in S}}_{\text{at most } f \text{ terms}} x_S^{LP} \geq 1$

$\Rightarrow$ at least of these x_S^{LP} (for S containing i) must be $\geq 1/f$

$\Rightarrow$ at least one of these $x_S^2 = 1$

so x^2 is feasible

and $\sum_S c_S x_S^2 \leq \sum_S c_S (f x_S^{LP}) = f \sum_S c_S x_S^{LP} \leq f\,\mathrm{OPT}$

the same performance guarantee as from the Dual Greedy algorithm!

(for Vertex Cover: another 2-approximation!)

Randomized Rounding

$0 \leq x_S^{LP} \leq 1$ can be interpreted as a probability

- for each let $x_S^R = \begin{cases} 1 & \text{with probability } x_S^{LP} \\ 0 & \text{— — } 1 - x_S^{LP} \end{cases}$

independently for all $S \in \mathcal{S}$

let $\mathcal{C}$ be the resulting (random) subset (defined by its characteristic vector x^R)

its expected cost $E\,c(\mathcal{C}) = \sum_S c_S \,\mathrm{Prob}\{S \in \mathcal{C}\} = \sum_S c_S x_S^{LP} = c x^{LP}$

Take $a \in N$ let k be the frequency of a in $\mathcal{S}$

since x^{LP} is feasible $\sum_{S:a\in S} x_S^{LP} \geq 1$

$\mathrm{Prob}\{a \notin \mathcal{C}\} = \mathrm{Prob}\{S \notin \mathcal{C} \;\forall S \text{ s.t. } a \in S\} = \prod_{S:a\in S} \mathrm{Prob}\{S \notin \mathcal{C}\}$

"not covered by"

$= \prod_{S:a\in S} (1 - x_S^{LP}) \leq \left(1 - \frac{1}{k}\right)^k \leq 1/e$

Draw an independent random sample of $d \log n$ such sets $\mathcal{C}$ where d satisfies $\left(1/e\right)^{d \log n} \le 1/4n$ and let $\mathcal{C}'$ be their union

so $\text{Prob}\{a \notin \mathcal{C}'\} \le \left(1/e\right)^{d \log n} \le 1/4n$ for every $a \in M$

$$\text{Prob}\{\text{some } a \notin \mathcal{C}'\} \le \sum_{a \in M} \text{Prob}\{a \notin \mathcal{C}'\} \le 1/4$$

The expected cost

$$E(c(\mathcal{C}')) = c x^{LP} d \log n$$

by Markov inequality ($\text{Prob}\{X > k\,E[X]\} \le 1/k$ for $X \ge 0$ since $c(\mathcal{C}')$ is a nonnegative random variable)

$$\text{Prob}\{c(\mathcal{C}') > 4\, c x^{LP} d \log n\} \le 1/4$$

Therefore $\text{Prob}\{(\text{some } a \notin \mathcal{C}') \text{ or } (c(\mathcal{C}') \ge 4\, c x^{LP} d \log n)\}$

$$\le \text{Prob}\{\text{some } a \notin \mathcal{C}'\} + \text{Prob}\{c(\mathcal{C}') > 4\, c x^{LP} d \log n\}$$

$$\le 1/4 + 1/4 = 1/2$$

so $\text{Prob}(\mathcal{C}' \text{ covers } M \text{ and } c(\mathcal{C}') \le 4\, c x^{LP} d \log n) \ge 1/2$

$\Rightarrow$ with high probability we obtain a cover with cost no more than $(4 d \log n)\,OPT$