An Introduction to Approximation Algorithms

Daya Gaur

Department of Computer Science and Engineering
Indian Institute of Technology Ropar
IGGA Workshop at DAIICT

14 March 2012

approximation algorithms

- trade off accuracy for time.
- for every instance we compute an α approximate solution in polynomial time.

Example

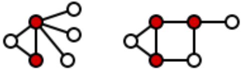
Travelling salesperson: an approximation algorithm returns a tour no more than twice the length of the shortest tour *for every instance*.

Types of approximation algorithms

- Good news: there are problems in NP that admit FPTAS.
- Bad news: there are problems in NP than do not admit any approximation algorithm (unless..)

Inapproximability can be thought of as more refined study of class NP-C.

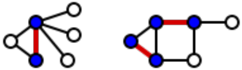
Vertex cover



$$\text{IP}: \text{minimize} \sum_{v \in V} w_v x_v$$
$$x_u + x_v >= 1 \;\forall (u,v) \in E$$
$$x_v \in \{0,1\} \;\forall v \in V$$

$$\text{LP}: \text{minimize} \sum_{v \in V} w_v x_v$$
$$x_u + x_v >= 1 \;\forall (u,v) \in E$$
$$x_v \geq 0 \;\forall v \in V$$

Matching



$$\mathrm{D-LP}: \text{minimize} \sum_{(u,v)\in e} y_{uv}$$

$$\sum_{v\in n(u)} y_{uv} \leq w_u \text{ for all } u \in V$$

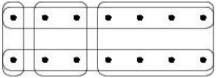
$$y_{uv} \geq 0 \ \forall v \in V$$

$$\mathrm{M}: \text{minimize} \sum_{(u,v)\in e} y_{uv}$$

$$\sum_{v\in n(u)} y_{uv} \leq w_u \text{ for all } u \in V$$

$$y_{uv} \in \{0,1\} \ \forall v \in V$$

Set cover



$$\text{IP : minimize } \sum_{s \in S} w_s x_s$$
$$\sum_{s : v \in s} x_s >= 1 \ \forall v \in U$$
$$x_s \in \{0,1\} \ \forall s \in S$$

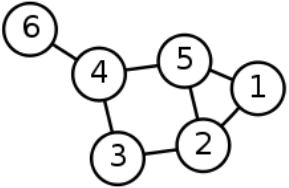
$$\text{LP : minimize } \sum_{v \in v} w_s x_s$$
$$\sum_{s : v \in s} x_s >= 1 \ \forall v \in U$$
$$x_s \geq 0 \ \forall s \in S$$

Set cover: dual

$$\mathrm{LP}: \mathrm{maximize} \sum_{u \in V} y_u$$
$$\sum_{v \in s} y_v \leq w_s \ \forall s \in S$$
$$y_v \geq 0 \ \forall v \in V$$

$$\mathrm{ILP}: \mathrm{maximize} \sum_{u \in V} y_u$$
$$\sum_{v \in s} y_v \leq w_s \ \forall s \in S$$
$$y_v \in \{0,1\} \ \forall v \in V$$

Shortest paths in digraphs



$$\mathrm{ip: minimize} \sum_{e \in e} w_e x_e$$

$$\sum_{e \in n(v)} x_e = \begin{cases} 1 & \text{if } v = s \\ -1 & \text{if } v = t \\ 0 & \text{otherwise} \end{cases}$$

$$x_e \in \{0,1\} \; \forall e \in E$$

vertex, edge incidence matrix. (u,e) is 1 is edge goes out from u, -1 otherwise.

LP relaxation solution is integral (integrality gap is 1)

Constrained shortest paths

$$\mathrm{ip} : \mathrm{minimize} \sum_{e \in e} w_e x_e$$

$$\sum_{e \in n(v)} x_e = \begin{cases} 1 & \text{if } v = s \\ -1 & \text{if } v = t \\ 0 & \text{otherwise} \end{cases}$$

$$\sum_{e \in e} d_e x_e \leq d$$

$$x_e \in \{0,1\} \forall e \in e$$

Constrained shortest paths

$$\mathrm{IP}_1 : \min \sum_{e \in e} w_e x_e + \lambda\left(\sum_{e \in e} d_e x_e - d\right)$$

$$\sum_{e \in n(v)} x_e = \begin{cases} 1 & \text{if } v = s \\ -1 & \text{if } v = t \\ 0 & \text{otherwise} \end{cases}$$

$$\sum_{e \in e} d_e x_e \le d$$

$$x_e \in \{0,1\} \; \forall e \in e$$

Let $x*$ be the optimal integral solution to the constrainted shortest path problem. $v(ip,x*) \ge v(\mathrm{IP}_1,x*)$ for $\lambda \ge 0$.

Constrained shortest paths

$$\mathrm{IP_L : min} \sum_{\mathrm{e \in e}} \mathrm{w_e x_e} + \lambda(\sum_{\mathrm{e \in e}} \mathrm{d_e x_e - d})$$

$$\sum_{\mathrm{e \in n(v)}} \mathrm{x_e} = \begin{cases} 1 & \text{if } \mathrm{v = s} \\ -1 & \text{if } \mathrm{v = t} \\ 0 & \text{otherwise} \end{cases}$$

$$\mathrm{x_e} \in \{0,1\} \; \forall \mathrm{e} \in \mathrm{e}$$

Let $\mathrm{x'}$ be the optimal integral solution to $\mathrm{IP_L}$.
$\mathrm{v(IP_1, x*)} \geq \mathrm{v(IP_L, x')} \; \forall \lambda \geq 0$.

Theorem

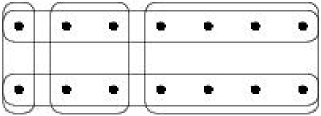
The value of optimal solution to $\mathrm{IP_L}$ is a lower bound on the value of $\mathrm{x*}$.

Lagrangian relaxation

- IP_{L} is the lagrangian relaxation.
- we want the best possible lower bound, therefore find λ such that optimal to IP_{L} maximized.
  - largarangian can be solved using subgradient methods (note that the function might not be differentiable).
  - or using column generation (dantzig-wolfe decomposition).
  - lagrangian bound is atleast as good as the linear programming bound, for integral polytopes the two bounds coincide

Greedy algorithm for set cover

```
r = {}
sol = {}
while r != u
       s : min { w(s)/|s \ r| }
       sol <- sol union s
       r <- r union (s \ r)
       for all e in (s \ r)
           p(e) =  w(s)/|s \ r|
```



Analysis

Proof.

$$\sum_{e \in u} p(e) = w(sol)$$

consider a set $s = (s_1, \ldots, s_k)$

for all $i, p(s_i) \leq \frac{w(s)}{k-i}$

$$\sum_{s_i \in s} p(s_i) \leq w(s)h(k)$$

$p(e)/h(n)$ is feasible in the dual

by weak duality the performance ratio is $h(n)$.

$s_i's$ are ordered in the order they are covered by the greedy.

□

Vertex cover

$$\mathrm{ip: minimize} \sum_{v \in v} w_v x_v$$
$$x_u + x_v >= 1 \ \forall (u,v) \in e$$
$$x_v \in \{0,1\} \ \forall v \in v$$

$$\mathrm{lp: minimize} \sum_{v \in v} w_v x_v$$
$$x_u + x_v >= 1 \ \forall (u,v) \in e$$
$$x_v \geq 0 \ \forall v \in v$$

let x^* be the optimal lp solution.

$$x_v = \begin{cases} 1 & \text{if } x_v^* >= \frac{1}{2} \\ 0 & \text{otherwise.} \end{cases}$$

Vertex cover

- each constraint contains atleast one variable with value $\geq 1/2$ in x^*.
- rounding gives a feasible integral solution.
- value of the integral solution is at most double the value of the optimal lp solution.

Half integrality of vertex cover

Definition

A solution to an LP is an extreme point if it cannot be expressed as a convex combination of two other feasible solutions.

Lemma

Every extreme point solution is half integral i.e., $x_v \in \{0, 1/2, 1\}$.

Proof.

$$V_p = \{v \mid x_v^* > 1/2\} \qquad V_n = \{v \mid x_v^* < 1/2\}$$

$$a_v = \begin{cases} x_v^* + \epsilon & \text{if } v \in V_p \\ x_v^* - \epsilon & \text{if } v \in V_n \\ x_v^* & \text{otherwise.} \end{cases} \qquad b_v = \begin{cases} x_v^* - \epsilon & \text{if } v \in V_p \\ x_v^* + \epsilon & \text{if } v \in V_n \\ x_v^* & \text{otherwise.} \end{cases}$$

Pricing method for vertex cover

Definition

$p_e \geq 0$ is the price associated with each edge e.
w_v is the cost associated with each vertex v.
price p is fair if for every vertex $\sum_{e \text{ on } v} p_e \leq w_v$.

Theorem

A fair price is a lower bound on the cost of any vertex cover.

$$\sum_{e \text{ on } v} p_e \leq w_v$$

$$\sum_{v \in s} \sum_{e \text{ on } v} p_e \leq w(s)$$

Algorithm

Definition

A vertex is saturated if $\sum_{\mathrm{e\ on\ v}} \mathrm{p_e} = \mathrm{w_v}$
an edge is uncovered if neither of its endpoints are in the cover.

```
price of e = 0
   while there exists an
   uncovered edge e
       raise price on e
       without violating
       fairness
s = { v | v is saturated}
```

Analysis

Theorem

Cost of vertex cover produced is at most twice the fair price.

Proof.

Every vertex in the cover is saturated.

$$\sum_{e \text{ on } v} p_e = w_v$$

$$\sum_{v \in s} \sum_{e \text{ on } v} p_e = w(s)$$

$$2 \sum_{e} p_e \geq w(s)$$

□

Complementary slackness

$$\begin{aligned} P : \min \quad & \textstyle\sum_{i=1}^{n} c_i x_i \\ \sum_{i=1}^{n} a_{ji} x_i \geq b_j, \quad & j = 1..m \\ & x_i \geq 0 \end{aligned}$$

$$\begin{aligned} D : \max \quad & \textstyle\sum_{j=1}^{m} b_j y_j \\ \sum_{j=1}^{m} a_{ji} y_j \leq c_i, \quad & i = 1..n \\ & y_j \geq 0 \end{aligned}$$

Complementary slackness

Theorem

Let x, y *be primal and dual feasible solutions.* x, y *are optimal if and only if following conditions are satisfied:*

1. $x_i(\sum_{j=1}^{m} a_{ji} y_j - c_i) = 0$ *for all* $1 \leq i \leq n$
2. $y_j(\sum_{i=1}^{n} a_{ji} x_i - b_j) = 0$ *for all* $1 \leq j \leq m.$

Relaxed complementary slackness

Definition

Let x, y be primal and dual feasible solutions. x, y are said to satisfy relaxed complementary slackness condition if:

1. $x_i > 0 \implies \frac{c_i}{\alpha} \leq \sum_{j=1}^{m} a_{ji} y_j \leq c_i = 0$ for all $1 \leq i \leq n$
2. $y_j > 0 \implies (b_j \leq \sum_{i=1}^{n} a_{ji} x_i \leq \beta b_j) = 0$ for all $1 \leq j \leq m$.

rcsc

Theorem

If x, y *are feasible and satisfy rcsc then* $\sum_{i=1}^{n} c_i x_i \leq \alpha\beta \sum_{j=1}^{m} b_j y_j$.

Proof.

$$\sum_{i=1}^{n} c_i x_i \leq \alpha \sum_{i=1}^{n} \left(\sum_{j=1}^{m} a_{ij} y_j \right) x_i$$

$$\sum_{j=1}^{m} \left(\sum_{i=1}^{n} a_{ij} x_i \right) y_j \leq \beta \sum_{j=1}^{m} b_j y_j$$

□

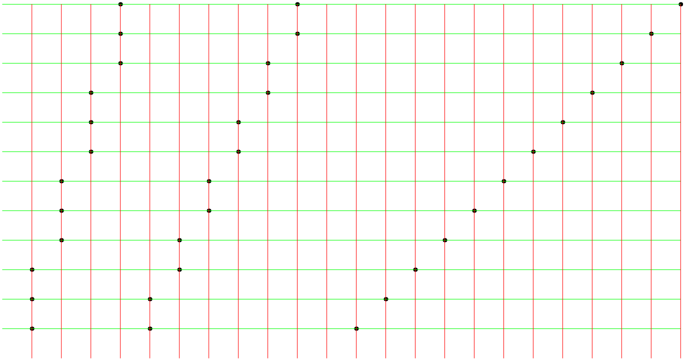
Pricing method for vertex cover revisited

$\alpha = 1, \beta = 2$
$x_v > 0 \implies \sum_{e \text{ on } v} p_e = w_v$
$p_e > 0 \implies x_u + x_v \leq 2$

- primal conditions: pick only saturated vertices in the cover.
- dual conditions: from every edge pick atmost two vertices in the cover.

primal conditions satisfied by pricing algorithm, dual conditions satisfied automatically

Point Cover



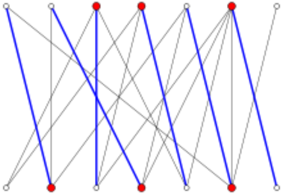
all the horizontal lines (12) comprise the optimal solution, and the greedy algorithm will pick all the vertical lines (22). the example can be generalized.

2D

The problem is equivalent to vertex cover in bipartite graphs in 2-D.

Theorem (König–Egerváry)

The size of the minimum vertex cover is the same as the size of the maximum matching in a bipartite graph.



2-D can be solved optimally.

d dimensions

Let L be the set of all the axis parallel lines, associated with each line $l \in L$ is a binary variable y_l whose value is 1 if the line is picked in the solution, 0 otherwise. $L(x_i)$ is the set of axis parallel lines through point x_i.

$$\text{IP: } \min \sum_{l \in L} y_l \tag{1}$$

$$\sum_{l \,:\, l \in L(x_i)} y_l \geq 1 \quad \forall\, x_i \tag{2}$$

$$y_l \in \{0,1\} \tag{3}$$

The linear programming relaxation LP to the integer program IP, is obtained by replacing constraints of type (3) with non-negativity constraints $y_l \geq 0$. The linear programming dual of LP is:

$$\text{LP-dual: } \max \sum_{i=1}^{n} z_i \tag{4}$$

$$\sum_{i\,:\,l \in L(x_i)} z_i \leq 1 \quad \forall\, l \in L \tag{5}$$

$$z_i \geq 0 \tag{6}$$

Primal-dual d approximation

- while there exists an uncovered point, the algorithm picks all the d axis parallel lines that go through the point.
- set $y_l = 1$ if the line is picked by the algorithm else $y_l = 0$. let x_i be the uncovered point picked in iteration j, then set $z_i = 1$.
- solutions constructed above are feasible, and the value of the primal solution is at most d times the value of the dual solution, i.e. the performance ratio is d.

R. Bar-Yehuda and S. Even, *A linear-time approximation algorithm for the weighted cover problem*. Journal of Algorithms, 2:198–203, 1981.

G. B. Dantzig, L. R. Ford, D. R. Fulkerson, *A primal-dual algorithm for linear programs*, In Linear Inequalities and Related Systems, Editors Kuhn and Tucker, pages 171–181, Princeton University Press, 1956.

T. Gonzalez, editor, *Handbook of Approximation Algorithms and Meta Heuristics*, CRC Press, 2007.

D. Hochbaum, editor, *Approximation Algorithms for NP-hard problems*, PWS Publishing, Boston, 1997.

V. Vazirani, *Approximation Algorithms*, Springer, 2001.

Figures are from Wikipedia.