From relaxations to integral solutions (cont.)

Optimization - 10725

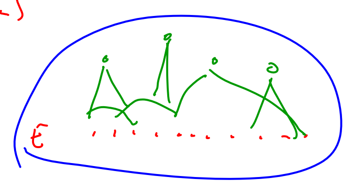
Carlos Guestrin

Carnegie Mellon University

April 23rd, 2008

Relaxations and rounding

- What do we do if we don't get integral solutions?
  - because $P \neq NP$ (probably true)
- E.g., set cover problem
  - Ground elements $v \in V$
  - Set of Sets $S \in \mathcal{S}$ $\quad S \subseteq V$
  - Cost for sets c_S
  - Find cheapest collection of subsets that covers all elements



- Integer program and relaxation:

$$\min_x \sum_S c_S x_S$$

$$x_S \in \{0,1\}$$

$$\forall v \in V \quad \sum_{S: v \in S} x_S \geq 1$$

$S: v \in S$

relax

$$\min_x \sum_S c_S x_S$$

$$\sum_{S: v \in S} x_S \geq 1 \quad \forall v$$

$$0 \leq x_S \leq 1$$

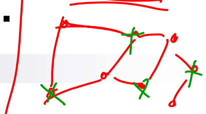
- How can we obtain a good integer (rounded) solution?
  - If we set all nonzero x_S to one, then very bad idea...
  - Smart rounding?

arbitrarily $O(n)$ more expensive solution than OPTIP

Consider a special case…

sets are vertices
elements edges

K=2
vertex cover



- Suppose each element in at most k sets
- From inequality constraint:

$$\forall v \quad \sum_{s: v \in s} x_s \geq 1 \quad \Rightarrow \quad \text{for each } v \in V \quad \exists s,\ x_s \geq \frac{1}{k} \text{ and } s \text{ covers } v$$

- Rounding strategy: pick all s such that $x_s \geq \frac{1}{k}$
  all others get discarded
  $\hat{x}_s$
- Feasibility: with respect to IP:
  $\forall v \in V$ $\exists$ a picked set that covers v
- Cost of rounded solution: $\sum_s c_s \hat{x}_s \geq OPT_{IP}$

$$OPT_{LP} \leq OPT_{IP}$$

relaxation

$$\sum_s c_s \hat{x}_s \leq k\ OPT_{LP} \leq k\ OPT_{IP}$$

k-approximation

Very simple example of randomized rounding

$E[A+B] = E[A] + E[B]$

- Solve set cover relaxation:

$$\min_x \sum_s c_s x_s$$
$$\forall v \quad \sum_{s: v \in s} x_s \geq 1$$
$$x_s \in [0,1]$$

- Randomly pick a collection of subsets $\hat{G}$
  - For each S, add it to $\hat{G}$ with (independent) probability x_s
- What's the expected cost of $\hat{G}$? $E[c(\hat{G})]$
  - I_s indicator of whether set S is in $\hat{G}$

$$E[c(\hat{G})] = E\left[\sum_s I_s c_s\right] = \sum_s c_s E[I_s] = \sum_s c_s x_s = OPT_{LP} \leq OPT_{IP}$$

$\hat{G}$ costs less than OPT_{IP}, because doesn't cover all elements, i.e., $\hat{G}$ is infeasible wrt IP

$\frac{1}{e} \geq (1-\frac{1}{k})^k \quad k \geq 1$

How many elements do we cover?

- Expected cost of G can be lower than OPT_{IP}
  - Must cover fewer elements
- I_v is indicator of whether element v covered by G
- Expected number of elements covered:

v appears in sets $S_1 \ldots S_k$

$E[\text{cover}] = E[\sum_v I_v] = \sum_v \underbrace{E[I_v]}_{\text{prob cover } v \leftarrow P_v}$

$P_v =$ pick at least one of sets where v appears

$= 1 - \prod_{i=1}^{k} (1 - X_{S_i}) \geq 1 - (1-\frac{1}{k})^k \geq 1 - \frac{1}{e}$

worst set of X_{S_i} wrt P_v

$X_{S_i} = \frac{1}{k}$ derivative $= 0$

$E[\text{cover}] \geq \underbrace{(1-\frac{1}{e})}_{\sim 63\%} |V|$

How big can cost get?

Feed back www.cmu.edu/uca please!! :)

Poster session Thur. 3-6pm Atrium NSH

- Expected cost is lower than OPT_{IP}
  - But how big can actual cost get?
  - (a simple bound here, more interesting bounds using more elaborate techniques)
- Markov Inequality: Let Y be a non-negative random variable
  - Then $P(Y \geq t\, E[Y]) \leq \frac{1}{t}$
- In our example: $Y = C(\hat{G}) \geq 0$

$P(C(\hat{G}) \geq 2\, OPT_{IP}) \leq P(C(\hat{G}) \geq 2\, OPT_{LP}) \leq \frac{1}{2}$

because $OPT_{IP} \geq OPT_{LP}$

Randomization & Derandomization

- MAX-3SAT:
  - 3SAT formula:
    - Binary variables $X_1,\ldots,X_n$
    - Conjunction of clauses $C_1,\ldots,C_M$
    - Each clause is a disjunction of three literals on three different variables
  - Want assignment that maximizes number of satisfied formulas

$SAT(X)$ = # of satisfied clauses by assignment x

$$\max_x SAT(X)$$

if formula satisfiable $\max_x SAT(X) = M$

in general $\max_x SAT(X) \leq M$

Randomized algorithm for MAX-3SAT

- Pick assignment for each X_i independently, at random with prob. 0.5

- Expected number of satisfied clauses:

$I_j \leftarrow$ whether clause j is sat.

$$E[SAT(X)] = E_X\left[\sum_{j=1}^{n} I_j\right] = \sum_j \underbrace{E[I_j]}_{\text{prob } j \text{ is sat}}$$

$$= \frac{7}{8} M$$

consider $(X_1 \vee \neg X_2 \vee X_3)$
only $X_1=0, X_2=1, X_3=0$
doesn't satisfy,
only happens with prob $1/8$

Aside: Probabilistic Method

- Expected number of satisfied clauses:

  $E[SAT(x)] \geq \frac{7}{8} M$

- Probabilistic method: for any random var. Y, there exists assignment y such that P(y)>0, y≥E[Y]
  - Almost obvious fact
  - Amazing consequences
- For example, in the context of MAX-3SAT:

  Every 3SAT formula has an assignment that satisfies at least $\frac{7}{8}$ of the clauses

Derandomization

- There exists assignment X that achieves $7/8 M$
- In expectation, we get 7/8.M, but can we get it with prob. 1? Without randomization?
- Derandomization: From a randomized algorithm, obtain a deterministic algorithm with same guarantees
  - Today: method of conditional expectations

Method of conditional expectations

- Conditional expectation: $E[X|y] = \sum_x x \, P(x|y)$
- Expectation of the conditional expectation:
  - $E_y\left[E_x[X|Y]\right] = E[X]$ ✓ $\sum_y P(y)\cdot P(x|y) = P(x)$
- Consider MAX-3SAT:
  - Expectation: $E[SAT(x)] \geq 7/8 M$
  - Expectation of conditional expectation:

$$E[SAT(x)] = E_{x_1}\left[E_{x_2\cdots x_n}[SAT(x)|x_1]\right]$$

$$= P(x_1=1)\, E[SAT(x)|x_1=1] + P(x_1=0)\cdot E[SAT(x)|x_1=0]$$

$$= \frac{1}{2} E[SAT(x)|x_1=1] + \frac{1}{2} E[SAT(x)|x_1=0] \geq \frac{7}{8}M$$

$$\max\left\{E[SAT(x)|x_1=1]\,;\, E[SAT(x)|x_1=0]\right\} \geq \frac{7}{8}M$$

Computing conditional expectation

- Conditioning on $X_1=1$: $E_{x_2\cdots x_n}[SAT(x)|x_1=1] = \sum_{j=1}^{M} \underbrace{E[I_j|x_1=1]}$

$$E[I_j|x_1=1] = \begin{cases} 7/8 & \text{if } x_1 \text{ is not in clause } j \\ 1 & \text{if } x_1 \text{ is in clause } j\text{, and } x_1=1 \text{ makes } C_j=1 \\ 1-\frac{1}{4}=\frac{3}{4} & \text{" " " " " } x_1=1 \text{ doesn't make } C_j=1 \end{cases}$$

- General case: $X_1=v_1,\ldots, X_i=v_i$ $\quad(\neg x_1 \vee x_2 \vee \neg x_7)$
  - Sum over clauses, I_j is indicator clause j is satisfied

$$E[I_j|x_1=v_1,\ldots x_i=v_i] = \begin{cases} 1 & \text{if } x_1=v_1,\ldots,x_i=v_i \text{ makes } C_j \text{ true} \\ 1-\frac{1}{2^k} & \text{if } x_1=v_1,\ldots x_i=v_i \text{ doesn't make } C_j \text{ true and there are } k \text{ unassign variables in } C_j \end{cases}$$

Derandomized algorithm for MAX-3SAT

- For i=1,….,n
  - Try X_i=1
    - Compute $E[SAT(x) \mid X_1 = v_1, \ldots X_{i-1} = v_{i-1}, X_i = 1]$
  - Try X_i=0
    - Compute $E[SAT(x) \mid X_1 = v_1, \ldots X_{i-1} = v_{i-1}, X_i = 0]$
  - Set v_i to best assignment to X_i max
- Deterministic algorithm guaranteed to achieve at least 7/8.M

Most probable explanation (MPE) in a Markov network

- Markov net:
- Most probable explanation:
- In general, NP-complete problem, and hard to approximate

MPE for attractive MNs – 2 classes

- Attractive MN:
  - E.g., image classification
- Finding most probable explanation
- Can be solved by

MPE, Attractive MNs, k classes

- MPE for k classes:
- Multiway cut:
  - Graph G, edge weights w_{ij}
  - Finding minimum cut, separate $s_1,\ldots,s_k$
- Multiway cut problem is

Multiway cut – combinatorial algorithm

- Very simple alg:
  - For each i=1…k
    - Find cut C_i that separates s_i from rest
  - Discard argmax$_i$ w(C_i), return union of rest

- Algorithm achieves 2-2/k approximation
  - OPT cut A* separates graph into k components
    - No advantage in more than k
  - From A* form $A^*_1,\ldots,A^*_k$, where A^*_i separates s_i from rest
  - Each edge in A* appears in
    - Thus

Multiway cut proof

- Thus, for OPT cut A* we have that:
- Each A^*_i separates s_i from rest, thus
- But, can do better, because