Partial Resampling to Approximate Covering Integer Programs

Antares Chen[1] David G. Harris[2] Aravind Srinivasan[2]

[1]University of California Berkeley

[2]University of Maryland

SODA 2016

Question Statement

We discuss positive covering integer programs (CIPs)

Covering Constraint Let $A \in \mathbb{R}_+^{m \times n}$, $\mathbf{x} \in \mathbb{Z}_+^n$ and $\mathbf{a} \in \mathbb{R}_+^n$:

$$A\mathbf{x} \geq \mathbf{a}$$

With scaling, all $A_{ki} \in [0, 1]$ and $a_k \geq 1$.

Linear Objective For some $\mathbf{C} \in \mathbb{R}_+^n$, we optimize:

$$\min \mathbf{C} \cdot \mathbf{x}$$

Multiplicity Constraint Optionally for integral values of d_i, restrict:

$$x_i \in \{0, 1, ..., d_i\}$$

Column Sparse CIPs

A bound on the number of appearances a variables has in a constraint.

Two metrics of column sparsity ℓ_0 and ℓ_1-norms:

$$\Delta_0 = \max_i \#k : A_{ki} > 0 \qquad\qquad \Delta_1 = \max_i \sum_k A_{ki}$$

Note that $\Delta_1 \leq \Delta_0$ with possibility that $\Delta_1 \ll \Delta_0$.

Also consider:

$$a_{\min} = \min_k a_k$$

Here the larger $a_{\min}$, the easier approximation task.

Current CIP Algorithms

Current CIP algorithms tend to fall into two design categories:

- Greedy Algorithms
- Linear Relaxation with Randomized Rounding

Initializes all $\mathbf{x} = \mathbf{0}$ and then increments x_i where i is chosen as a local optimum to a residual problem.

Chvátal, Lovász, Johnson (1970s) first develop greedy methods for set cover.

Dobson (1982), Fisher & Wolsey (1982) extend greedy methods to CIP.

Feige (1998) proves that greedy methods are essentially optimal.

Handling multiplicity and multiple objectives with greedy algorithms is cumbersome!

Linear Relaxation

Raghavan & Thompson (1987) Relaxes the CIP problem to the set $\mathbb{R}$, and then randomly round to integral values until covering constraints are satisfied.

Simple analysis using chernoff and union bounds gives this approximation ratio:

$$1 + \mathcal{O}\left(\frac{\log m}{a_{\min}} + \sqrt{\frac{\log m}{a_{\min}}}\right)$$

Linear Relaxation

Srinivasan (2006) uses the FKG inequality to provide a method with approximation ratio:

$$1 + \mathcal{O}\Big(\frac{\log(\Delta_0 + 1)}{a_{\min}} + \sqrt{\frac{log(a_{\min})}{a_{\min}} + \frac{\log(\Delta_0 + 1)}{a_{\min}}}\Big)$$

- Has an exponentially small probability of achieving the desired approximation ratio.
- Can arbitrarily violate multiplicity constraints.
- Can be derandomized efficiently, but is cumbersome and causes loss to approximation ratio.

Kolliopoulos & Young (2005) applied Srinivasan accounting for multiplicity. Provide two algorithms:

Given $\epsilon \in (0, 1]$, violates multiplicity $x_i \leq \lceil (1+\epsilon) d_i \rceil$ with approximation ratio:

$$\mathcal{O}\left(1 + \frac{\log(\Delta_0 + 1)}{a_{\min} \cdot \epsilon^2}\right)$$

Meets $x_i \leq d_i$ exactly with approximation ratio:

$$\mathcal{O}\big(\log(\Delta_0)\big)$$

The Local Lemma

The Lovász Local Lemma is fundamental to probabilistic methods.

Consider the probability space Ω with "bad" events $\mathbb{B} = \{B_1, ..., B_n\}$.

If for all "bad" events B_i affects at most d other events, and $P_\Omega(B_i) \leq p$, then given criterion:

$$ep(d+1) \leq 1$$

There is a positive probability (usually exponentially small) that no B_i occurs.

This is not constructive!

The Moser-Tardos Algorithm

Moser & Tardos (2009) give a fully algorithmic version of the LLL.

Consider all "bad" events are determined by a set of "atomic" variables $X \subseteq \Omega$.

Moser & Tardos give the following algorithm:

- Draw all $X \sim \Omega$
- While there exists a true B_i:
  - Arbitrarily select some true B_i
  - From Ω, resample all variables involved in B_i

Link to CIPs - Consider each covering constraint as a "bad" event.

Primary Motivation: The MT Framework

Harris & Srinivasan (2014) provide a partial resampling variant of the Moser-Tardos algorithm.

Instead of sampling all variables involved in B_i, choose an appropriately random subset.

Many improved algorithmic applications where the classical LLL falls short.

Harris & Srinivasan (2016) Applies the variant of the MT algorithm where LLL is violated.

The Resampling Algorithm

We introduce `RELAXATION`, a linear relaxation algorithm for CIPs

Given the parameters $\alpha > 1$, $\sigma \in [0, 1]$, the system $A\mathbf{x} \geq \mathbf{a}$ and fractional solution $\hat{x}$:

- Draw each $x_i \sim \text{Bernoulli}(\alpha \hat{x}_i)$ independently.
- For k from $1 \to n$
    - While constraint k is unsatisfied $(A_k \cdot \mathbf{x} < a_k)$:
        - Resample all $x_i = 0$ to $\text{Bernoulli}(\sigma A_{ki} \alpha \hat{x}_i)$ independently.

Variables can only increase so termination is guaranteed.

Note: WLOG assume $\hat{x}_i$ small (the hardest case).

Results and Contributions

We can formulate approximation ratio in terms of:

$$\gamma = \frac{\ln(\Delta_1 + 1)}{a_{min}}$$

Our contributions:

1. Formulate all approximation ratios in terms of Δ_1 instead of Δ_0 ($\Delta_1 \ll \Delta_0$).
2. Easily handles multi-criteria CIPs with good results.
3. Achieve asymptotically better results than both Srinivasan and Kolliopoulos & Young.
4. For γ large, we achieve the optimal approximation ratio of γ with correct constant factor.
5. For γ large, we achieve the optimal power of ϵ.

Approximation Ratios

For appropriate σ, α, `RELAXATION` runs in $\mathcal{O}(mn)$ with approximation ratio:

$$1 + \gamma + 4\sqrt{\gamma}$$

For $\epsilon \in [0, 1]$ with $x_i \leq \lceil \hat{x}_i(1 + \epsilon) \rceil$:

$$1 + \frac{4\gamma}{\epsilon} + 4\sqrt{\gamma}$$

Note $\Delta_1 \leq \Delta_0$ with possibility $\Delta_1 \ll \Delta_0$.

We provide critical hardness-bounds for CIPs based on Feige (1998) and Moshkovitz (2012), assuming ETH:

Suppose T is the fractional solution, $a_{\min} \geq a$, c a positive constant.

There exists no poly-time algorithm that can find an integral solution x where:

$$x \leq T\Big(\frac{\ln(\Delta_1 + 1)}{a_{\min}} - \frac{c \ln\ln(\Delta_1 + 1)}{a_{\min}}\Big)$$

But notice for $\gamma \to \infty$:

$$\frac{\ln(\Delta_1 + 1)}{a_{\min}} - \frac{c \ln\ln(\Delta_1 + 1)}{a_{\min}} \;\geq\; \big(1 - o(1)\big)\gamma$$

Optimality - Ignoring Multiplicity

CIP cannot be approximated within:

$$\big(1 - o(1)\big)\gamma$$

By comparison we have (ignoring multiplicity):

$$1 + \gamma + 4\sqrt{\gamma} = \big(1 + o(1)\big)\ \gamma$$

The correct constant factor on γ for γ large.

Hardness - Considering Multiplicity

Any poly-time algorithm within a $1+\epsilon$ factor of multiplicity must have:

For γ large, approximation ratio:

$$\Omega\left(\frac{\gamma}{\epsilon}\right)$$

By comparison:

$$1+\frac{4\gamma}{\epsilon}+4\sqrt{\gamma}=\mathcal{O}\left(\frac{\gamma}{\epsilon}\right)$$

Hardness - For γ Small

Let $\hat{T}$ be the optimal, T the optimal *integral* solution

For linear relaxation algorithms where γ is small, the lowest integrality gap:

$$\frac{T}{\hat{T}} \geq 1 + \Omega(\gamma)$$

By comparison our approximation ratio for γ small:

$$1 + \mathcal{O}\left(\sqrt{\gamma}\right)$$

First hardness result for the case of small γ.

Multi-Criteria CIP and Concentration Bounds

For $C_i \in \mathbb{R}_+^n$, we have $C_1 \cdot \mathbf{x}, \ldots, C_r \cdot \mathbf{x}$ with an over-all objective function.

Negative correlation - property for any subset $R \subseteq [n]$, results of `RELAXATION` satisfy:

$$P\Big(\bigwedge_{i \in R} x_i = 1\Big) \leq \prod_{i \in R} \beta \hat{x}_i$$

β is the approximation ratio.

Can use Chernoff-Hoeffding bounds to derive $P(C_l \cdot \mathbf{x} > t)$

Apply union bound across all linear objectives.

Conclusions

Srinivasan (2006) ignoring multiplicity

$$1 + \mathcal{O}\Big(\frac{\log(\Delta_0 + 1)}{a_{\min}} + \sqrt{\frac{log(a_{\min})}{a_{\min}} + \frac{\log(\Delta_0 + 1)}{a_{\min}}}\Big)$$

We achieve

$$1 + \frac{\ln(\Delta_1 + 1)}{a_{\min}} + \mathcal{O}\Big(\sqrt{\frac{\ln(\Delta_1 + 1)}{a_{\min}}}\Big)$$

Kolliopoulos & Young (2005) respecting multiplicity

$$\mathcal{O}\Big(1 + \frac{\log(\Delta_0 + 1)}{a_{\min} \cdot \epsilon^2}\Big)$$

We achieve

$$1 + \mathcal{O}\Big(\frac{\ln(\Delta_1 + 1)}{a_{\min} \cdot \epsilon} + \sqrt{\frac{\ln(\Delta_1 + 1)}{a_{\min}}}\Big)$$

Conclusions

For γ large considering multiplicity, our approximation ratio achieves the correct power of ϵ:

$$\mathcal{O}\left(\frac{\gamma}{\epsilon}\right)$$

For γ large, our approximation ratio carries γ with the correct constant factor:

$$1 + \gamma + \mathcal{O}\left(\sqrt{\gamma}\right)$$

We show a negative correlation property allowing us to reason with multi-criteria CIPs.

Open Questions

1. Many problems are reduced to CIPs. How can we apply `RELAXATION`?
2. Can we utilize that our algorithm is agnostic to the following:
   - The initial rounding
   - The linear objective
3. Can we improve our approximation ratio?
4. Does there exist an efficient parallel algorithm for this?

Thank You!