Learning Cut Generating Functions for Integer Programming

Hongyu Cheng and Amitabh Basu

Johns Hopkins University

NeruIPS 2024

November 12, 2024

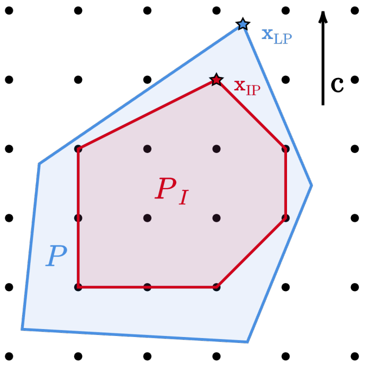
Integer Linear Programming Problems

Integer linear programming (IP) problem in standard form:

$$\begin{aligned} \max \quad & \mathbf{c}^{\mathsf{T}}\mathbf{x} \\ \text{s.t.} \quad & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0} \\ & \mathbf{x} \in \mathbb{Z}^n \end{aligned}$$

where $A \in \mathbb{Z}^{m\times n}, \mathbf{b} \in \mathbb{Z}^m, \mathbf{c} \in \mathbb{R}^n, m, n \in \mathbb{N}_+$.

$P := \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \leq \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$
$P_I := \operatorname{conv}(P \cap \mathbb{Z}^n)$



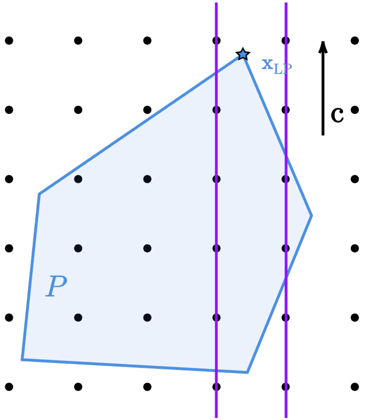
Branch-and-Bound

Select some $i \in \{1, \dots, n\}$ such that $(\mathbf{x}_{\mathrm{LP}})_i \notin \mathbb{Z}$.

$$\begin{aligned} \max \quad & \mathbf{c}^\top \mathbf{x} \\ \text{s.t.} \quad & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0} \\ & \mathbf{x} \in \mathbb{Z}^n \end{aligned}$$

$$\begin{aligned} \max \quad & \mathbf{c}^\top \mathbf{x} \\ \text{s.t.} \quad & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x}_i \leq \lfloor (\mathbf{x}_{\mathrm{LP}})_i \rfloor \\ & \mathbf{x} \geq \mathbf{0} \\ & \mathbf{x} \in \mathbb{Z}^n \end{aligned}$$

$$\begin{aligned} \max \quad & \mathbf{c}^\top \mathbf{x} \\ \text{s.t.} \quad & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x}_i \geq \lceil (\mathbf{x}_{\mathrm{LP}})_i \rceil \\ & \mathbf{x} \geq \mathbf{0} \\ & \mathbf{x} \in \mathbb{Z}^n \end{aligned}$$



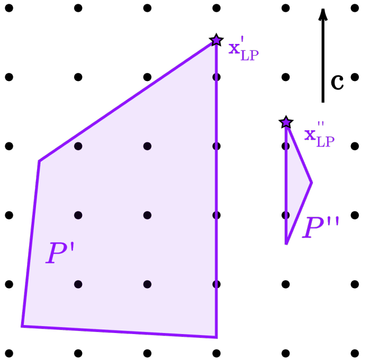
Branch-and-Bound

Select some $i \in \{1, \ldots, n\}$ such that $(\mathbf{x}_{\mathrm{LP}})_i \notin \mathbb{Z}$.

$$\begin{aligned} \max \quad & \mathbf{c}^\mathsf{T}\mathbf{x} \\ \text{s.t.} \quad & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0} \\ & \mathbf{x} \in \mathbb{Z}^n \end{aligned}$$

$$\begin{aligned} \max \quad & \mathbf{c}^\mathsf{T}\mathbf{x} \\ \text{s.t.} \quad & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x}_i \leq \lfloor (\mathbf{x}_{\mathrm{LP}})_i \rfloor \\ & \mathbf{x} \geq \mathbf{0} \\ & \mathbf{x} \in \mathbb{Z}^n \end{aligned}$$

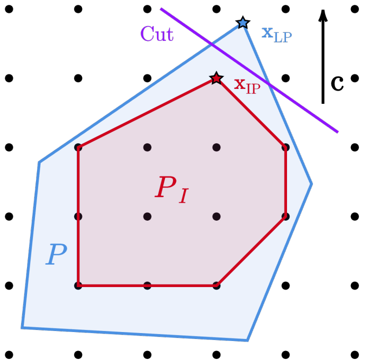
$$\begin{aligned} \max \quad & \mathbf{c}^\mathsf{T}\mathbf{x} \\ \text{s.t.} \quad & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x}_i \geq \lceil (\mathbf{x}_{\mathrm{LP}})_i \rceil \\ & \mathbf{x} \geq \mathbf{0} \\ & \mathbf{x} \in \mathbb{Z}^n \end{aligned}$$



Cutting Plane

Select a halfspace $H := \{\mathbf{x} : \boldsymbol{\alpha}^\mathsf{T}\mathbf{x} \le \beta\}$ such that it contains all the integer feasible points, i.e., $P_I \subseteq H$.

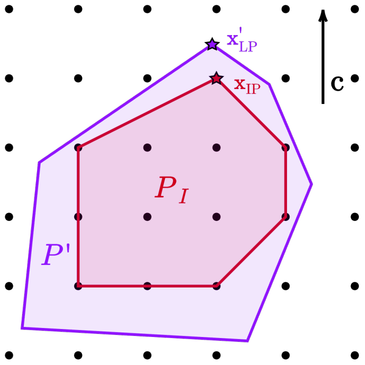
$$\begin{aligned} \max \quad & \mathbf{c}^\mathsf{T}\mathbf{x} \\ \text{s.t.} \quad & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \ge \mathbf{0} \\ & \boldsymbol{\alpha}^\mathsf{T}\mathbf{x} \le \beta \\ & \mathbf{x} \in \mathbb{Z}^n \end{aligned}$$



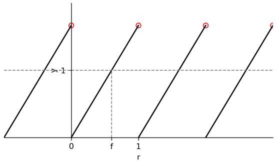
Cutting Plane

Select a halfspace $H := \{\mathbf{x} : \boldsymbol{\alpha}^\mathsf{T}\mathbf{x} \leq \beta\}$ such that it contains all the integer feasible points, i.e., $P_I \subseteq H$.

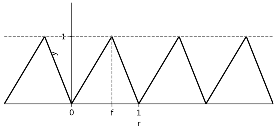
$$\begin{aligned} \max \quad & \mathbf{c}^\mathsf{T}\mathbf{x} \\ \text{s.t.} \quad & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0} \\ & \boldsymbol{\alpha}^\mathsf{T}\mathbf{x} \leq \beta \\ & \mathbf{x} \in \mathbb{Z}^n \end{aligned}$$



CGF Examples



(a) $CG_f(r)$



(b) $GMI_f(r)$

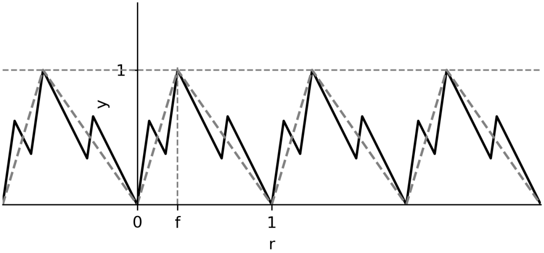
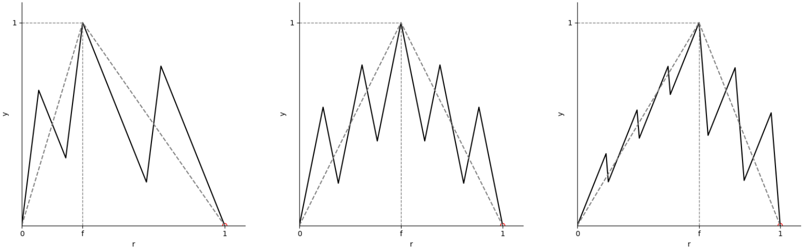


Figure: A one-dimensional cut generating function $\pi^{2,2}_{f,s_1,s_2}$

One-dimensional Extreme CGFs



(a) $\pi^{2,2}_{0.3,-2,7}(r)$ (b) $\pi^{3,3}_{0.5,-5,5}(r)$ (c) $\pi^{4,3}_{0.6,-12.5,2.5}(r)$

Figure: Three examples of the one-dimensional cut generating functions $\pi^{p,q}_{f,s_1,s_2}$ on $[0,1)$.

Learning Theory Tools

- Given $\varepsilon > 0$ and $\delta \in (0, 1)$, the following holds with probability at least $1 - \delta$ for $N \geq N(\varepsilon, \delta)$:

$$\left| \frac{1}{N} \sum_{i=1}^{N} T(I_i, \boldsymbol{\mu}) - \mathbb{E}_{I \sim \mathcal{D}} T(I, \boldsymbol{\mu}) \right| \leq \varepsilon,$$

- Pseudo-dimension can be used to bound $N(\varepsilon, \delta)$ [ABB+99]:

$$N(\varepsilon, \delta) = \mathcal{O}\left(\frac{B^2}{\varepsilon^2} \left(\mathrm{Pdim}(\{T(\cdot, \boldsymbol{\mu}) : \boldsymbol{\mu} \in \mathcal{P}\}) \log\left(\frac{B}{\varepsilon}\right) + \log\left(\frac{1}{\delta}\right) \right) \right),$$

Learnability Results

Theorem (Cheng and Basu)

$$\text{Pdim}\left(\left\{T(\cdot, \mathbf{s}) : \mathbf{s} \in [0,1]^2\right\}\right) = \mathcal{O}\left(n^2 \log((m+n)\varrho)\right).$$

Theorem (Cheng and Basu)

$$\text{Pdim}\left(\left\{T^k(\cdot, \boldsymbol{\mu}) : \boldsymbol{\mu} \in \Delta_k^\tau\right\}\right) = \mathcal{O}\left(kn^2 \log((m+n)\varrho) + k^2 \log(n\tau)\right).$$

Learnability Results

Use neural network to select a good CGF for each instance?

Theorem (Cheng and Basu)

$$\text{Pdim}\left(\left\{h(\cdot,\mathbf{w}) : \mathbf{w} \in \mathbb{R}^W\right\}\right) = \mathcal{O}\left(LW\log(U+2) + n^2W\log((m+n)\varrho)\right),$$

$$\text{Pdim}\left(\left\{h^k(\cdot,\mathbf{w}) : \mathbf{w} \in \mathbb{R}^W\right\}\right) = \mathcal{O}\left(LW\log(U+k) + kW\log(n\tau) + n^2W\log((m+n)\varrho)\right).$$

Numerical Experiments

Table: Average tree sizes on 100 instances, after adding a single type of cut at the root.

Problem Type	GMI	1-row cut	2-row cut	5-row cut	10-row cut	best 1-row cut
Knapsack(20, 1)	158.88	**87.0**	n/a	n/a	n/a	35.54
Knapsack(30, 1)	832.16	**58.84**	n/a	n/a	n/a	13.98
Knapsack(50, 1)	3543.91	**277.01**	n/a	n/a	n/a	125.85
Knapsack(16, 2)	399.86	316.8	**178.68**	234.09	203.66	102.07
Knapsack(30, 3)	4963.91	4311.04	3430.37	**2817.55**	2822.37	3561.36
Packing(15, 30)	389.67	**367.48**	376.86	401.87	391.44	303.72
Packing(20, 40)	1200.55	1123.9	1214.92	**1113.82**	1185.26	738.58

Numerical Experiments

Table: Average tree sizes on 100 instances, after adding a single type of cut at the root.

Problem Type	GMI	1-row cut	2-row cut	5-row cut	10-row cut	best 1-row cut
Knapsack(20, 1)	158.88	**87.0**	n/a	n/a	n/a	35.54
Knapsack(30, 1)	832.16	**58.84**	n/a	n/a	n/a	13.98
Knapsack(50, 1)	3543.91	**277.01**	n/a	n/a	n/a	125.85
Knapsack(16, 2)	399.86	316.8	**178.68**	234.09	203.66	102.07
Knapsack(30, 3)	4963.91	4311.04	3430.37	**2817.55**	2822.37	3561.36
Packing(15, 30)	389.67	**367.48**	376.86	401.87	391.44	303.72
Packing(20, 40)	1200.55	1123.9	1214.92	**1113.82**	1185.26	738.58

References I

Martin Anthony, Peter L Bartlett, Peter L Bartlett, et al., *Neural network learning: Theoretical foundations*, vol. 9, cambridge university press Cambridge, 1999.

Maria-Florina F Balcan, Siddharth Prasad, Tuomas Sandholm, and Ellen Vitercik, *Structural analysis of branch-and-cut and the learnability of gomory mixed integer cuts*, Advances in Neural Information Processing Systems **35** (2022), 33890–33903.