Strong IP Formulations Need Large Coefficients

Christopher Hojny

TECHNISCHE UNIVERSITÄT DARMSTADT

Technische Universität Darmstadt
Department of Mathematics

Discrete Optimization

Aussois Combinatorial Optimization Workshop 2019

Example I – The Good

TECHNISCHE UNIVERSITÄT DARMSTADT

$$X = \left\{ x \in \{0,1\}^n \ : \ \sum_{i=1}^{n} 2^{n-i} x_i \leq 2^n - 3 \right\}$$

Pro:

- small IP formulation

Con:

- badly scaled coefficients
  $\rightsquigarrow$ numerical instabilities

Question: Is there a better IP formulation?

- tightest IP formulation consists of facet inequalities of conv(X)
- complete linear description [Laurent, Sassano 1992]

$$\left\{ x \in [0,1]^n \ : \ \sum_{i=1}^{n-1} x_i \leq n - 2 \right\}$$

- facet description is numerically stable

Example II – The Bad

$$X = \left\{ x \in \{0,1\}^{2n} \ : \ \sum_{i=1}^{n} 2^{n-i}(x_i + x_{n+i}) \leq 2^n - 1 \right\}$$

Question: Can we use facet description of conv(X)?

- consists of $\Theta(3^n)$ inequalities [Kaibel, Loos 2011]
- separable in linear time [Loos 2010]
- contains badly scaled inequalities, e.g.,

$$x_n + x_{2n} + \sum_{i=1}^{n-1} 2^{n-1-i}(x_i + x_{n+i}) \leq 2^{n-1}$$

- facet description is numerically instable

Questions

Every $X \subseteq \{0,1\}^n$ admits IP formulation with $\{0,\pm 1\}$-inequalities, e.g., via infeasibility cuts

$$\sum_{i=1}^{n} \Big((1-\bar{x}_i)\, x_i + \bar{x}_i\, (1-x_i) \Big) \geq 1 \qquad \forall \bar{x} \in \{0,1\}^n \setminus X,$$

but

1. Are such IP formulations strong?
2. If not, what is the minimum size of coefficients in strong IP formulations?

Outline

TECHNISCHE UNIVERSITÄT DARMSTADT

Discrete Optimization

Measure of Size of Coefficients

TECHNISCHE UNIVERSITÄT DARMSTADT

Inequality $a^\top x \leq \beta$ is numerically more stable the smaller the ratio

$$\rho(a) \coloneqq \max\left\{\frac{|a_i|}{|a_j|} \;:\; a_j \neq 0,\; i \neq j\right\}.$$

Definition

The ρ-value of an IP formulation $Ax \leq b$ is

$$\max\{\rho(a) \;:\; a \text{ is row of } A\}.$$

$\rightsquigarrow$ the smaller the ρ-value the higher the numerical stability

Strong IP Formulations

Question: How to measure strength of IP formulation $Ax \leq b$?

- IP formulation of $X \subseteq \{0,1\}^n$ has to cut off points $x \in \mathbb{Z}^n \setminus X$
- strongest formulation uses facets of $\operatorname{conv}(X)$ cutting off all points in $\mathbb{R}^n \setminus \operatorname{conv}(X)$

Idea: Refine IP formulations by not only cutting of infeasible binary points but also points in a refinement of the integer lattice.

Definition

Let $\lambda \in \mathbb{Z}_{>0}$ and $X \subseteq \{0,1\}^n$. Then $Ax \leq b$ is called $\frac{1}{\lambda}$-relaxation of X if for every $\bar{x} \in \frac{1}{\lambda}\mathbb{Z}^n$ we have $A\bar{x} \leq b$ iff $\bar{x} \in \operatorname{conv}(X)$.

Measuring Strength of IP Formulations

TECHNISCHE UNIVERSITÄT DARMSTADT

Measure of Strength

If $Ax \leq b$ is a $\frac{1}{\lambda}$-relaxation of X, then it is the stronger the larger λ.

In particular, $\frac{1}{\lambda}$-relaxations might have small coefficients, while facets of conv(X) have large coefficients.

But: How to estimate the size of coefficients in $\frac{1}{\lambda}$-relaxations?

Observation

$$
\begin{aligned}
& Ax \leq b,\ x \in [0,1]^n \text{ is } \tfrac{1}{\lambda}\text{-relaxation of } X \\
\Leftrightarrow\quad & Ax \leq \lambda\, b,\ x \in [0,\lambda]^n \text{ is IP formulation of } (\lambda \operatorname{conv}(X)) \cap \mathbb{Z}^n.
\end{aligned}
$$

$\rightsquigarrow$ Find bounds on coefficients in general IP formulations.

Lower Bound on Coefficients

TECHNISCHE UNIVERSITÄT DARMSTADT

Theorem [H. 2018]

Let $P \subseteq \mathbb{R}^n$ be a full-dimensional integral polytope. Let

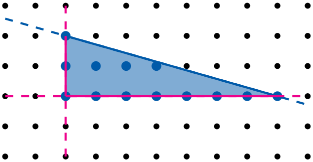
- $\bar{x} \in \mathbb{Z}^n \setminus P$ be not cut off by any valid box constraint.
- $\bar{A}x \leq \bar{b}$ consist of all facet inequalities for P cutting off $\bar{x}$.
- $\bar{s} \coloneqq \bar{A}\bar{x} - \bar{b}$.

If some technical assumptions hold, every valid inequality $c^\top x \leq \delta$ cutting off $\bar{x}$ fulfills

$$\frac{|c_i|}{|c_j|} \geq \frac{\min_k\{|\bar{A}_{ki}| - \bar{s}_k\}}{\max_k\{|\bar{A}_{kj}| + \bar{s}_k\}}.$$

Proof Idea

Consider $P = \text{conv}\{x \in \mathbb{R}^2 \ : \ 2\,x_1 + 7\,x_2 \leq 14, x_1 \geq 0, x_2 \geq 0\}$.



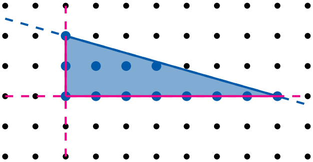
Question: Is there an IP formulation with smaller ρ-value?

Idea:

- select $\bar{x} \in \mathbb{Z}^2 \setminus P$
- derive bounds on coefficients in any valid inequality cutting of $\bar{x}$

Proof Idea

Consider $P = \text{conv}\{x \in \mathbb{R}^2 \;:\; 2\,x_1 + 7\,x_2 \leq 14, x_1 \geq 0, x_2 \geq 0\}$.



Question: Is there an IP formulation with smaller ρ-value?

Idea:

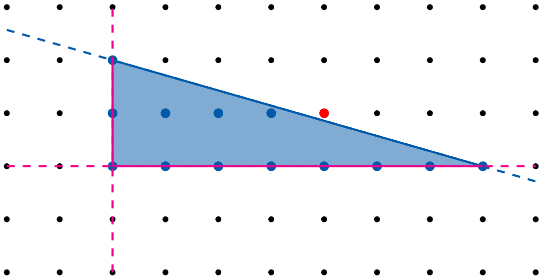
- select $\bar{x} \in \mathbb{Z}^2 \setminus P$
- derive bounds on coefficients in any valid inequality cutting of $\bar{x}$

Lemma

If P is a full-dimensional polytope, every valid inequality for P is a conic combination of facet defining inequalities.

Example

Consider $P = \text{conv}\{x \in \mathbb{R}^2 \ : \ 2\,x_1 + 7\,x_2 \leq 14, x_1 \geq 0, x_2 \geq 0\}$ and select $\bar{x} = \binom{4}{1}$.



Example

TECHNISCHE
UNIVERSITÄT
DARMSTADT

Consider $P = \operatorname{conv}\{x \in \mathbb{R}^2 \ : \ 2\,x_1 + 7\,x_2 \leq 14, x_1 \geq 0, x_2 \geq 0\}$ and select $\bar{x} = \binom{4}{1}$.

- up to scaling, every valid inequality has form

$$c^\top x \leq \delta \qquad :\Leftrightarrow \qquad (2-\alpha)x_1 + (7-\beta)x_2 \leq 14$$

- inserting $\bar{x}$ yields

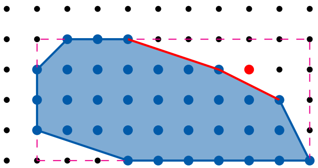
$$(2-\alpha)\bar{x}_1 + (7-\beta)\bar{x}_2 > 14 \qquad \Leftrightarrow \qquad 4\alpha + \beta < 1$$

- together with $\alpha \geq 0$ and $\beta \geq 0$

$$\rho(c) = \frac{7-\beta}{2-\alpha} > \frac{6+4\alpha}{2-\alpha} \geq \frac{6}{2} = 3$$

Generalization to General Integer Polytopes

TECHNISCHE UNIVERSITÄT DARMSTADT



Notation:

- B smallest box containing P with lower bounds ℓ and upper bounds u
- $\bar{x} \in (B \setminus P) \cap \mathbb{Z}^n$
- $\bar{A}x \leq \bar{b}$ facets violated by $\bar{x}$
- $Ax \leq b$ remaining facets

Idea: To find lower bounds on the ρ-value, use the previous lemma to find upper and lower bounds on coefficients in inequalities cutting of given $\bar{x}$.

Problem: Bounds depend on conic multipliers.
$\rightsquigarrow$ Introduce further conditions to get rid of multipliers.

Discrete Optimization

Condition I to Bound $\frac{|c_i|}{|c_j|}$

TECHNISCHE UNIVERSITÄT DARMSTADT

Sign Restrictions:

- For all rows k, k' of $\bar{A}$, we have

$$\operatorname{sgn}(\bar{A}_{kt}) = \operatorname{sgn}(\bar{A}_{k't}) \neq 0 \qquad \forall t \in \{i, j\}.$$

- For every row k of A not corresponding to a box constraint, we have

$$\operatorname{sgn}(A_{kt}) \in \{0, \operatorname{sgn}(\bar{A}_{1t})\} \qquad \forall t \in \{i, j\}.$$

$\bar{x} \in [\ell, u] \setminus P$ $\quad$ $\bar{A}x \leq \bar{b}$ facets violated by $\bar{x}$ $\quad$ $Ax \leq b$ remaining facets $\quad$ $\bar{s} := \bar{A}\bar{x} - \bar{b}$

Condition II to Bound $\frac{|c_i|}{|c_j|}$

TECHNISCHE UNIVERSITÄT DARMSTADT

Box Restrictions:

- $\bar{x}_i > \ell_i$.
- $\bar{x}_j < u_j$.
- For every inequality $a^\top x \leq \beta$ in $Ax \leq b$ not corresponding to a box constraint, we have

$$\sum_{t=1}^{n} a_t \bar{x}_t + \operatorname{sgn}(\bar{A}_{1j})\, e_j \leq \beta.$$

$\bar{x} \in [\ell, u] \setminus P$ $\quad$ $\bar{A}x \leq \bar{b}$ facets violated by $\bar{x}$ $\quad$ $Ax \leq b$ remaining facets $\quad$ $\bar{s} \coloneqq \bar{A}\bar{x} - \bar{b}$

Discrete Optimization

Condition III to Bound $\frac{|c_i|}{|c_j|}$

TECHNISCHE UNIVERSITÄT DARMSTADT

Excess Restrictions:

- $|\bar{A}_{ki}| \geq \bar{s}_k$ for every row k of $\bar{A}$.
- $\bar{x}_j - \ell_j \geq \max_k\{\bar{s}_k\}$.

$\bar{x} \in [\ell, u] \setminus P$ $\quad$ $\bar{A}x \leq \bar{b}$ facets violated by $\bar{x}$ $\quad$ $Ax \leq b$ remaining facets $\quad$ $\bar{s} := \bar{A}\bar{x} - \bar{b}$

Lower Bound on Coefficients

TECHNISCHE UNIVERSITÄT DARMSTADT

Theorem [H. 2018]

Let $P \subseteq \mathbb{R}^n$ be a full-dimensional integral polytope. Let

- $\bar{x} \in \mathbb{Z}^n \setminus P$ be not cut off by any valid box constraint.
- $\bar{A}x \leq \bar{b}$ consist of all facet inequalities for P cutting off $\bar{x}$.
- $\bar{s} \coloneqq \bar{A}\bar{x} - \bar{b}$.

If Conditions I–III hold, every valid inequality $c^\top x \leq \delta$ cutting off $\bar{x}$ fulfills

$$\frac{|c_i|}{|c_j|} \geq \frac{\min_k\{|\bar{A}_{ki}| - \bar{s}_k\}}{\max_k\{|\bar{A}_{kj}| + \bar{s}_k\}}.$$

Applications: IP formulations

Using this theorem, one can show:

- Every IP formulation of

$$\left\{ x \in \mathbb{Z}_+^{n+1} \ : \ \sum_{i=0}^{n} 2^i x_i \leq 2^n \right\}$$

  has ρ-value at least $\frac{2^n - 1}{2}$.
- Consequently, there exist $X \subseteq \mathbb{Z}^n$ that need exponentially large coefficients in any IP formulation.

Applications: $\frac{1}{\lambda}$-relaxation

TECHNISCHE UNIVERSITÄT DARMSTADT

Recall

$Ax \leq b$ is $\frac{1}{\lambda}$-relaxation of X $\Leftrightarrow$ $Ax \leq \lambda b$ is IP formulation of $(\lambda \operatorname{conv}(X)) \cap \mathbb{Z}^n$.

- The theorem implies every $\frac{1}{3}$-relaxation of

$$\Big\{x \in \{0,1\}^{3n+2} : x_{3n+1} + x_{3n+2} + \sum_{i=1}^{n} 2^{3(n-i)+4} x_{3(i-1)+1} + \sum_{i=1}^{n} 2^{3(n-i)+3)} x_{3(i-1)+2} + \sum_{i=1}^{n} (2^{3(n-i)+4} - 2^{3(n-i)+2}) x_{3i} \leq 1 + \sum_{i=1}^{n} (2^{3(n-i)+4} - 2^{3(n-i)+2}) \Big\}$$

  has ρ-value at least $\frac{3^{n/3-2}-1}{2}$.
- result can be generalized to exponentially large class of knapsacks

Consequences

Corollary

There exist sets $X \subseteq \{0,1\}^n$ that need exponentially large coefficients in every $\frac{1}{\lambda}$-relaxation if $\lambda \geq 3$.

Remark: result can be generalized to case $\lambda \geq 2$

Consequences:

- every $X \subseteq \{0,1\}^n$ admits 1-relaxation with $\{0, \pm 1\}$-coefficients.
- $\frac{1}{2}$-relaxations may need large coefficients
- strong IP formulations need exponentially large coefficients in general

Consequences

TECHNISCHE UNIVERSITÄT DARMSTADT

Corollary

There exist sets $X \subseteq \{0, 1\}^n$ that need exponentially large coefficients in every $\frac{1}{\lambda}$-relaxation if $\lambda \geq 3$.

Remark: result can be generalized to case $\lambda \geq 2$

Consequences:

- every $X \subseteq \{0, 1\}^n$ admits 1-relaxation with $\{0, \pm 1\}$-coefficients.
- $\frac{1}{2}$-relaxations may need large coefficients
- strong IP formulations need exponentially large coefficients in general

Thank You For Your Attention