Symmetry detection in Mixed-Integer Conic Programming

25th International Symposium on Mathematical Programming
July 23th, 2024
Montréal, Canada

Sven Wiese

www.mosek.com

A software package/library for solving:

- Linear and conic problems.
- Convex quadratic and quadratically constrained problems.
- Also mixed-integer versions of the above.

Current version is **MOSEK** 10.2.

(Mixed-Integer) Conic Programming in standard form:

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \in \mathcal{K} \cap \left(\mathbb{Z}^p \times \mathbb{R}^{n-p}\right) \end{array}$$

where $\mathcal{K}$ is a convex cone.

Conic building blocks

Typically, $\mathcal{K} = \mathcal{K}_1 \times \mathcal{K}_2 \times \cdots \times \mathcal{K}_K$ is a product of lower-dimensional cones:

- the linear cone $\mathbb{R}_+^{n_k}$
- the quadratic cone

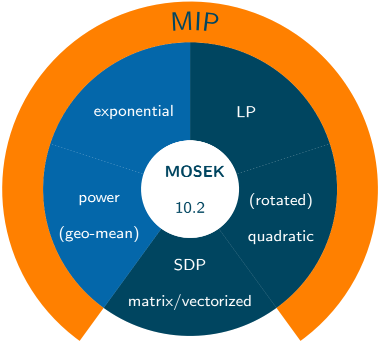
$$\mathcal{Q}^{n_k} = \{x \in \mathbb{R}^{n_k} \mid x_1 \geq \left(x_2^2 + \cdots + x_{n_k}^2\right)^{1/2} = \|x_{2:n_k}\|_2\}$$

- the exponential cone

$$\mathcal{K}_{exp} := \mathrm{cl}\{x \in \mathbb{R}^3 \mid x_1 \geq x_2 \exp(x_3/x_2),\ x_2 > 0\}$$

- ...

What's inside **MOSEK** ?



- Simplex / Interior-point algorithms for continuous problems.
- (Conic) Branch-and-Cut / Outer-approximation algorithms for discrete problems, including heuristics, cuts, and other MIP-solver components...
- ... like symmetry handling!

Symmetry handling in MILP

Focus on formulation symmetries, i.e., permutations!

MILP:

$$\begin{aligned} \text{minimize} \quad & c^T x \\ \text{subject to} \quad & Ax = b \\ & x \in \mathbb{Z}_+^p \times \mathbb{R}_+^{n-p} \end{aligned} \tag{MILP}$$

Definition (see, e.g., Pfetsch & Rehn, 2019)

$\pi \in \mathcal{S}_n$ is a formulation symmetry of (MILP) *iff* $\exists\ \sigma \in \mathcal{S}_m$ *such that*

$$\pi(\{1, \ldots, p\}) = \{1, \ldots, p\} \tag{1}$$

$$\pi(c) = c \tag{2}$$

$$\sigma(b) = b \tag{3}$$

$$A_{\sigma(i),\pi(j)} = A_{ij}. \tag{4}$$

Symmetry handling in MILP

Focus on formulation symmetries, i.e., permutations!

MILP:

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \in \mathbb{Z}_+^p \times \mathbb{R}_+^{n-p} \end{array} \tag{MILP}$$

Definition (see, e.g., Pfetsch & Rehn, 2019)

$\pi \in S$ is a formulation symmetry of (MILP) iff $\exists\, \sigma \in S$ such t

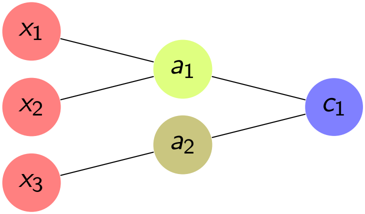
π is a valid permutation of the variables, if

- cost and type are preserved for each variable
- for each constraint, there is an identical one after permuting the variables

Symmetry detection in MILP

For detecting symmetries, (MILP) is represented as a colored graph (the matrix graph):

$$2x_1 + 2x_2 + 3x_3 \leq 4 \quad \longrightarrow$$



- Every (color-invariant) graph automorphism corresponds to a formulation symmetry, and vice-versa!
- Software packages for detecting graph automorphisms are `nauty`, `saucy`, `bliss`, ...

Formulation symmetries in MICP

MICP:

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \in \mathcal{K} \cap \left(\mathbb{Z}^p \times \mathbb{R}^{n-p}\right) \end{array} \tag{P}$$

Definition

$\pi \in \mathcal{S}_n$ is a formulation symmetry of (P) iff $\exists\ \sigma \in \mathcal{S}_m$ such that (1) - (4), and $\pi(\mathcal{K}) = \mathcal{K}$.

- $\pi(\mathcal{K}) = \mathcal{K}$ is rather generic...
- ... but may translate to more concrete conditions when looking at a specific $\mathcal{K}$.

Formulation symmetries in MICP

MICP:

$$
\begin{array}{ll}
\text{minimize} & c^T x \\
\text{subject to} & Ax = b \\
 & x \in \mathcal{K} \cap \left(\mathbb{Z}^p \times \mathbb{R}^{n-p}\right)
\end{array}
\tag{P}
$$

Definition

$\pi \in \mathcal{S}_n$ is a formulation symmetry of (P) iff $\exists\ \sigma \in \mathcal{S}_m$ such that
(

just as in the linear case, plus $\pi(\mathcal{K}) = \mathcal{K}$...

- $\pi(\mathcal{K}) = \mathcal{K}$ is rather generic...
- ... but may translate to more concrete conditions when looking at a specific $\mathcal{K}$.

A cone's symmetry labelings

For the quadratic cone $\mathcal{K} = \mathcal{Q}^n := \{x \in \mathbb{R}^n \mid x_1 \geq \|x_{2:n}\|_2\}$:

$$\pi(\mathcal{K}) = \mathcal{K} \iff \pi(1) = 1$$

In other words,

$$\pi(\{1\}) = \{1\} \text{ and } \pi(\{2, \ldots, n\}) = \{2, \ldots, n\}.$$

Generalize this concept:

Definition

We call a function $h : \{1, \ldots, n\} \mapsto \mathbb{N}$ a symmetry labeling w.r.t. a cone $\mathcal{K} \subseteq \mathbb{R}^n$, iff for any $\pi \in \mathcal{S}_n$ the condition $h(\pi(i)) = h(i)\ \forall i$ implies $\pi(\mathcal{K}) = \mathcal{K}$.

A cone's symmetry labelings

For the quadratic cone $\mathcal{K} = \mathcal{Q}^n := \{x \in \mathbb{R}^n \mid x_1 \geq \|x_{2:n}\|_2\}$:

$$\pi(\mathcal{K}) = \mathcal{K} \iff \pi(1) = 1$$

In other words,

$$\pi(\{1\}) = \{1\} \text{ and } \pi(\{2, \ldots, n\}) = \{2, \ldots, n\}.$$

Generalize this concept:

Definition

We call a function $h : \{1, \ldots, n\} \mapsto \mathbb{N}$ a symmetry labeling w.r.t. a

c

i

Find a partition of the variables in $\mathcal{K}$, and permute only inside the partition cells!

A cone's symmetry labelings (cont.)

- $\mathcal{K} = \mathcal{Q}^n = \{x \in \mathbb{R}^n \mid x_1 \geq \|x_{2:n}\|_2\}$: a labeling is given by

$$h(i) = \begin{cases} 1, & i = 1 \\ 2, & \textit{otherwise}. \end{cases}$$

- $\mathcal{K} = \mathbb{R}^n_+$: a labeling is any constant function,

$$h(i) = c \; \forall i.$$

- $\mathcal{K} = \mathcal{K}_{exp} = \mathrm{cl}\{x \in \mathbb{R}^3 \mid x_1 \geq x_2 \exp(x_3/x_2), x_2 > 0\}$: the only labeling is the identity.

Proposition

$\pi \in \mathcal{S}_n$ is a formulation symmetry of (P) if $\exists\ \sigma \in \mathcal{S}_m$ such that (1) - (4), and $\exists$ a labeling h w.r.t. $\mathcal{K}$ such that $h(\pi(i)) = h(i)\ \forall i$.

A cone's symmetry labelings (cont.)

- $\mathcal{K} = \mathcal{Q}^n = \{x \in \mathbb{R}^n \mid x_1 \geq \|x_{2:n}\|_2\}$: a labeling is given by

$$h(i) = \begin{cases} 1, & i = 1 \\ 2, & \textit{otherwise}. \end{cases}$$

- $\mathcal{K} = \mathbb{R}^n_+$: a labeling is any constant function,

$$h(i) = c \ \forall i.$$

- $\mathcal{K} = \mathcal{K}_{exp} = \mathrm{cl}\{x \in \mathbb{R}^3 \mid x_1 \geq x_2 \exp(x_3/x_2), x_2 > 0\}$: the only labeling is the identity.

Proposition

$\pi \in \mathcal{S}_n$ is a formulation symmetry of (P) *if* $\exists\ \sigma \in \mathcal{S}_m$ *such that* (1) - (4),

a

just as in the linear case, plus the permuted variables have the same label for some labeling w.r.t. $\mathcal{K}$...

A (seemingly) bad example

$$(x_1, x_2, x_3, x_4, x_5, x_6)^T \in \mathcal{K} = \mathcal{Q}^3 \times \mathcal{Q}^3$$

- $\pi_1 = (1,4)(2,5)(3,6)$ is a valid formulation symmetry.
- In a labeling, 1 and 4 would have the same label:

$$h(1) = h(4).$$

- Then also $\pi_2 = (1,4)$ would be formulation symmetry.
- But with $x = (1,0,0,2,1,1) \in \mathcal{K}$, $\pi_2(x) \notin \mathcal{K}$. $\lightning$

Cartesian products are still tractable

Theorem

Let $\mathcal{K} = \mathcal{K}_1 \times \ldots \times \mathcal{K}_K \subseteq \mathbb{R}^n$ with $\mathcal{K}_k \subseteq \mathbb{R}^{n_k}$. Then $\pi \in \mathcal{S}_n$ is a formulation symmetry of (P) *if $\exists\ \sigma \in \mathcal{S}_m$ such that* (1) - (4)*, and $\exists\ \tau \in \mathcal{S}_K$ and labelings $h_k : \{1, \ldots, n_k\} \mapsto \mathbb{N}$ w.r.t. $\mathcal{K}_k$ such that*

$$
\begin{aligned}
\mathcal{K}_{\tau(k)} &= \mathcal{K}_k \\
\pi(\{N_k + 1, \ldots, N_{k+1}\}) &= \{N_{\tau(k)} + 1, \ldots, N_{\tau(k)+1}\} \\
h_{\tau(k)} &= h_k \\
h_{\tau(k)}(\pi(i) - N_{\tau(k)}) &= h_k(i - N_k)\ \forall i \in \{N_k + 1, \ldots, N_{k+1}\},
\end{aligned}
$$

where $N_k = \sum_{l<k} n_l$, $N_{K+1} = n$.

The matrix graph construction in **MOSEK** is based on this.

Cartesian products are still tractable

Theorem

Let $\mathcal{K} = \mathcal{K}_1 \times \ldots \times \mathcal{K}_K \subseteq \mathbb{R}^n$ with $\mathcal{K}_k \subseteq \mathbb{R}^{n_k}$. Then $\pi \in \mathcal{S}_n$ is a formulation symmetry of (P) *if $\exists\ \sigma \in \mathcal{S}_m$ such that* (1) - (4)*, and $\exists\ \tau \in \mathcal{S}_K$ and labelings $h_k : \{1, \ldots, n_k\} \mapsto \mathbb{N}$ w.r.t. $\mathcal{K}_k$ such that*

π is a valid permutation of the variables, if

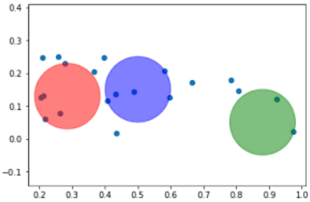
- bounds, cost and type are preserved for each variable
- for each constraint, there is an identical one after permuting the variables
- for each cone, there is another cone of the same type in which the permuted variables have the same label

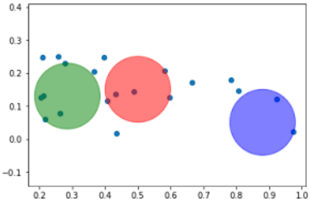
The matrix graph construction in **MOSEK** is based on this.

Relation to MINLP

- Labelings may still fail to capture symmetries for certain cones. Example: SDP (instead refer to Hojny & Pfetsch, 2022!)
- Seem to work well for cones coming from epigraphs of convex functions, i.e., MICP reformulations of convex MINLPs.
- Could apply MINLP techniques (Liberti, 2012) to such MICPs, using the expression graph of an instance...
- ... but the matrix graph construction using symmetry labelings is likely be dominant in terms of graph size.

Computational results: disk covering





- MICP formulations with quadratic cones $\mathcal{Q}^{n_k}$
- 64 models:
  - between 1 and 4 symmetry generators per model
  - roughly 47% - 99% of variables moved
- conic Branch-and-Cut, symmetry exploitation: orbital fixing

Computational results: disk covering (cont.)

- Time spent in symmetry detection:

det. time	min	max	mean
in millisec.	0.3	4.0	1.3
as % of running time	4.1e-5	0.15	3.6e-3

- Overall performance:

		solved		time				nodes	
	models	def	sym-0	def	sym-0*	faster	slower	def	sym-0*
$[0, \infty)$	64	64	54	71.6	3.2	61	2	12108	3.0
$[10, \infty)$	56	56	46	112.6	3.6	54	1	18824	3.4
$[100, \infty)$	40	40	30	314.2	3.9	39	0	43613	3.6
$[1000, \infty)$	21	21	11	954.2	4.6	21	0	113533	5.2

Computational results: f-SPARC

- optimize efficiency in data transmission systems with k communication channels and n users
- data rate requirement for user j: $B \sum_i \log_2(1 + p_{ij}/N) \geq d_j$
- has MICP reformulation with exponential cones $\mathcal{K}_{exp}$
- 62 models
  - between 7 and 14 symmetry generators per model
  - $>$ 99% of variables moved

Computational results: f-SPARC (cont.)

- Time spent in symmetry detection:

det. time	min	max	mean
in millisec.	1.3	9.1	3.2
as % of running time	9.4e-5	1.2	2.8e-2

- Overall performance:

		solved		time				nodes	
	models	def	sym-0	def	sym-0*	faster	slower	def	sym-0*
$[0, \infty)$	62	62	29	17.4	73.5	62	0	254	98.3
$[10, \infty)$	57	57	24	21.9	91.1	57	0	309	147.5
$[100, \infty)$	46	46	13	40.2	124.2	46	0	652	305.7
$[1000, \infty)$	43	43	10	45.7	128.9	43	0	583	450.9

- From CBLIB and QPLIB (convex):

	det. time (1e-3 sec.)	% of vars	# gen.	factors
achtziger_stolpe07-5.2bflowc	1.6	2.78	1	$\mathcal{M}(\mathcal{S}_2, 38)$
netmod_kar1	2.7	2.78	9	1 unknown
netmod_kar2	1.5	100	9	1 unknown
QPLIB_3361	114.9	3.22	1	$\mathcal{M}(\mathcal{S}_2, 64)$
QPLIB_3496	1.3	42.89	6	1 unknown, $(\mathcal{M}(\mathcal{S}_2, 26))^2$, $\mathcal{M}(\mathcal{S}_2, 22)$
QPLIB_3547	21.0	100	30	1 unknown
QPLIB_3643	1.9	49.51	7	1 unknown, $(\mathcal{M}(\mathcal{S}_2, 18))^2$
rsyn0810m04h	2.2	1.69	3	$\mathcal{M}(\mathcal{S}_4, 12)$
syn05m04h	0.2	18.18	3	$\mathcal{M}(\mathcal{S}_4, 12)$

Symmetry in public instance libraries (cont.)

- Symmetry occurence:
  - CBLIB: 12.5% of analyzed instance groups
  - QPLIB (convex): 12.9% of analyzed instances

  - MIPLIB2003: 30%
  - MIPLIB2010: 42.7%

    (numbers taken from Pfetsch & Rehn, 2019)

$\Rightarrow$ Contribute to CBLIB (or to QPLIB, preferably convex...)! ☺
$\rightarrow$ `cblib.zib.de/`

Further information on **MOSEK**

- Documentation at `mosek.com/documentation/`
  - Modeling cook book / cheat sheet.
  - White papers.
  - Manuals for interfaces.
  - Notebook collection.

- Tutorials and more at `github.com/MOSEK/`