Еремеев А.В.[1] Юрков А.С.[2]

Об одном приложении групп Ли в квадратичном программировании

[1] Омский филиал Института математики им. С.Л. Соболева СО РАН

[2] Институт радиофизики и физической электроники ОНЦ СО РАН

2023

Symmetries in Mathematical Programming

Solution and analysis of mathematical programming problems may be simplified when these problems are symmetric under appropriate linear transformations. Knowledge of the symmetries may help:

- reduce the problem dimension,
- cut the search space by symmetry-breaking linear cuts
- obtain new local optima from the ones previously found
- develop problem-specific metaheuristics

Previous Research

In the case of **continuous domain**:

- K.Gatermann and P.A. Parrilo "Symmetry groups, semidefinite programs, and sums of squares" (2004)
- A.Costa, P. Hansen and L. Liberti "On the impact of symmetry-breaking constraints on spatial Branch-and-Bound for circle packing in a square" (2013)
- G.Kouyialis, X. Wang, and R. Misener "Symmetry detection for quadratic optimization using binary layered graphs" (2019)
- etc.

In **integer programming**:

- Р. Ю. Симанчев "Линейные симметрии многогранника паросочетаний и автоморфизмы графа" (1996)
- О.В. Червяков "Аффинные симметрии многогранника, системы независимости с единичным сдвигом" (1999)
- M. François "Symmetry in integer linear programming" (2010)
- А. А. Колоколов, Т. Г. Орловска , М. Ф. Рыбалка, "Анализ алгоритмов целочисленного программировани с использованием -разбиени и унимодул рных преобразований" (2012)
- R. Bödi, K. Herr and M. Joswig, M. "Algorithms for highly symmetric linear and integer programs" (2013)
- etc.

... and more papers on symmetries in MIP problems.

The problem of maximizing the radiation of an antenna array in a given direction

$$\begin{cases} \mathsf{x}^T \mathsf{G} \mathsf{x} \to \max, \\ 0 \le \mathsf{x}^T \mathsf{H}^{(1)} \mathsf{x} \le 1, \\ \dots \\ 0 \le \mathsf{x}^T \mathsf{H}^{(n)} \mathsf{x} \le 1, \\ \mathsf{x} \in \mathbb{R}^{2n}. \end{cases} \tag{1}$$

- G is symmetric positive semi-definite,
- each matrix $\mathsf{H}^{(k)}$ is symmetric with
  - two identical positive eigenvalues,
  - two identical eigenvalues less or equal to zero,
  - the rest of the eigenvalues are equal to zero
- all eigenvalues of $\mathbf{H}_{\text{sum}} := \sum_{k=1}^{n} \mathbf{H}^{(k)}$ are positive.

Objective function symmetry: phase-independence of the radiation power flux.
Feasible domain symmetry: phase-independence of real power flow in each feeder of the antenna system.

In terms of (??), this symmetry is invariance w.r.t. rotation by some angle α in each 2-dimensional subspace (x_i, x_{n+i}). It can be utilized to reduce the search space dimensionality, e.g. by fixing $x_{2n} = 0$.

The problem has even more symmetries:

scatter_plot.png

What is symmetry group of the quadratic programming problem

The problem:

$$\begin{cases} \boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x} \to \mathbf{max} \\ \boldsymbol{x}^T \boldsymbol{B}_1 \boldsymbol{x} \le \mathbf{1} \\ \dots \\ \boldsymbol{x}^T \boldsymbol{B}_M \boldsymbol{x} \le \mathbf{1} \end{cases} \tag{2}$$

Tranformations:

$$\boldsymbol{x} \to \boldsymbol{y} = \boldsymbol{P}\boldsymbol{x}\,, \tag{3}$$

constitute group $\boldsymbol{\mathcal{G}}$ of symmetry if

$$\begin{cases} \boldsymbol{y}^T \boldsymbol{A} \boldsymbol{y} \to \mathbf{max} \\ \boldsymbol{y}^T \boldsymbol{B}_1 \boldsymbol{y} \le \mathbf{1} \\ \dots \\ \boldsymbol{y}^T \boldsymbol{B}_M \boldsymbol{y} \le \mathbf{1} \end{cases} \tag{4}$$

Symmetry Group of a Set of Matrices

In some cases, it may also be useful to find the symmetry group of the set of constraints only or the symmetry group of the set of matrices $\boldsymbol{B_i}$.

We denote the **set of given symmetric matrices** $\boldsymbol{B_1, \ldots, B_M}$ by $\boldsymbol{\mathcal{B}}$.

The set of symmetries of the constraints is closely related but **not necessarily identical** to the set of those invertible linear transformations, which map bijectively the feasibility domain of the problem

$$\boldsymbol{\mathcal{D} := \{x \in \mathbb{R}^N \ : \ x^T B_i x \leq 1, \ i = 1, \ldots, M\}}$$

onto itself.

The symmetry group of the set of constraints $\boldsymbol{\mathcal{G}'}$ may be[a] a subgroup in the symmetry group of invertible linear transformations of the domain $\boldsymbol{\mathcal{D}}$.

[a]This happens e.g. if there are several “inactive” constraints.

Definition 1. $\boldsymbol{\mathcal{B}}$ *is called* congruent invariant *(or invariant for short) under the transformation*

$$\boldsymbol{B} \rightarrow \boldsymbol{P^T B P} \tag{5}$$

with a non-degenerate matrix $\boldsymbol{P}$, *if* $\{\boldsymbol{P^T B P} \ : \ \boldsymbol{B} \in \boldsymbol{\mathcal{B}}\} = \{\boldsymbol{\mathcal{B}}\}$.

We denote the group of such matrices as $\boldsymbol{\mathcal{G}_{\mathcal{B}}}$. Clearly, under a transformation $\boldsymbol{B} \rightarrow \boldsymbol{P^T B P}$**,** some of the matrices from $\boldsymbol{\mathcal{B}}$ may be mapped one into another, however not all of the permutations can be obtained this way.

One of the invariants of the transformation (??) is the **inertia of a matrix** $\boldsymbol{B}$, defined as the ordered triple:

1) the number of positive eigenvalues of $\boldsymbol{B}$,
2) the number of negative eigenvalues of $\boldsymbol{B}$, and
3) the number of zero eigenvalues of $\boldsymbol{B}$.

So it is only possible to permute matrices with equal inertia and the whole set $\boldsymbol{\mathcal{B}}$ splits into equivalence classes of matrices with equal inertia.

Definition 2. *An* I-class *is a maximal by inclusion subset* $\boldsymbol{\mathcal{B}_k^I}$ *of* $\boldsymbol{\mathcal{B}}$, *consisting of matrices with equal inertia.*

Definition 3. *A sum of all matrices belonging to one or several I-classes is called an* invariant matrix of I-type.

Obviously for such a matrix $\boldsymbol{B^I}$

$$\boldsymbol{P^T B^I P} = \boldsymbol{B^I} \tag{6}$$

Condition I: *at least one invariant matrix $\boldsymbol{B^I}$ of I-type is positive definite.*

Proposition 1. *If Condition I holds then the group $\boldsymbol{\mathcal{G}_{\mathcal{B}}}$ is isomorphic to some subgroup of the group of orthogonal transformations, and this isomorphism is given by the mapping*

$$\boldsymbol{P} \to \boldsymbol{S P S^{-1}}, \tag{7}$$

where the matrix $\boldsymbol{S}$ is such that $\boldsymbol{B^I} = \boldsymbol{S^T S}$.

Structure of the Symmetry Group

Symmetry implies that

$$\boldsymbol{P}^T \boldsymbol{A} \boldsymbol{P} = \boldsymbol{A}$$

$$\boldsymbol{P}^T \boldsymbol{B}_i \boldsymbol{P} = \sum_{j=1}^{M} \boldsymbol{L}_{ij} \boldsymbol{B}_j \tag{8}$$

where $\boldsymbol{L}_{ij}$ are the elements of a permutation matrix, i.e. matrix $\boldsymbol{L} = (\boldsymbol{L}_{ij})$ has a single "1" in each column and in each row, other elements of $\boldsymbol{L}$ are zeros.

In terms of $\boldsymbol{Q}$ it means that

$$\boldsymbol{Q}^T \tilde{\boldsymbol{A}} \boldsymbol{Q} = \tilde{\boldsymbol{A}}\,, \quad \boldsymbol{Q}^T \tilde{\boldsymbol{B}}_i \boldsymbol{Q} = \sum_{i=1}^{N} \boldsymbol{L}_{ij} \tilde{\boldsymbol{B}}_j \tag{9}$$

where

$$\tilde{\boldsymbol{A}} = \left(\boldsymbol{S}^{-1}\right)^T \boldsymbol{A} \boldsymbol{S}^{-1} \tag{10}$$

$$\tilde{\boldsymbol{B}}_i = \left(\boldsymbol{S}^{-1}\right)^T \boldsymbol{B}_i \boldsymbol{S}^{-1} \tag{11}$$

Or linearly in $\boldsymbol{Q}$:

$$\tilde{\boldsymbol{A}} \boldsymbol{Q} = \boldsymbol{Q} \tilde{\boldsymbol{A}}\,, \quad \tilde{\boldsymbol{B}}_i \boldsymbol{Q} = \sum_{i=1}^{N} \boldsymbol{L}_{ij} \boldsymbol{Q} \tilde{\boldsymbol{B}}_j \tag{12}$$

Continuous and Discrete Subgroups

In non-degenerate case as any Lie group, $\boldsymbol{\mathcal{G}}$ consists of connected components (in the topological sense), only one of which, hereafter denoted as $\boldsymbol{\mathcal{G}_1}$, contains the identity element.

This $\boldsymbol{\mathcal{G}_1}$ is invariant subgroup of $\boldsymbol{\mathcal{G}}$ and further called **the continuous subgroup of symmetries**.

The remaining connected components (not being subgroups) are the cosets of $\boldsymbol{\mathcal{G}_1}$. These cosets make up a **discrete subgroup of symmetries**. Given that $\boldsymbol{\mathcal{G}_1}$ is an invariant subgroup, multiplication of the cosets is determined naturally, and this discrete group is factor group $\boldsymbol{\mathcal{G}/\mathcal{G}_1}$.

The Case of Non-Degenerate Spectra:

Theorem 1. *If Condition I holds and the eigenvalues of at least one matrix* $\boldsymbol{\tilde{B}_i}$ *are all different, then the group* $\boldsymbol{\mathcal{G}_{\mathcal{B}}}$ *is finite and may be found in* $\boldsymbol{O(M!2^{N\mathcal{O}(\log(NM))})}$ *time if all eigenvalues of* $\boldsymbol{\tilde{B}_i}$ *are given.*

Finding Continuous Subgroup of Symmetry

Consider the following system of linear equations in $\boldsymbol{a_n}$:

$$\begin{cases} \boldsymbol{B}_i \left(\sum_n \boldsymbol{a}_n \boldsymbol{G}_n \right) = \left(\sum_n \boldsymbol{a}_n \boldsymbol{G}_n \right) \boldsymbol{B}_i \,, \\ \boldsymbol{A} \left(\sum_n \boldsymbol{a}_n \boldsymbol{G}_n \right) = \left(\sum_n \boldsymbol{a}_n \boldsymbol{G}_n \right) \boldsymbol{A} \,. \end{cases} \tag{13}$$

Let $\hat{\boldsymbol{G}}_{\boldsymbol{n}}$ make up a basis of the space of solutions to the system of linear equations (**??**) in the linear space of the $(\boldsymbol{N} \times \boldsymbol{N})$ skew-symmetric matrices.

Theorem 2

If Condition I holds, the continuous subgroup of symmetries $\boldsymbol{\mathcal{G}_1}$ consists of orthogonal transformations with matrices expressed by the matrix exponential function $\boldsymbol{e}^{\sum_n \boldsymbol{a}_n \hat{\boldsymbol{G}}_n}$, where $\boldsymbol{a_n}$ are any real-valued parameters.

Symmetries Detection via Extrema Enumeration under Condition I

- **Constraint optimization**

$$\min_{\boldsymbol{Q}} \left(||\tilde{\boldsymbol{A}}\boldsymbol{Q} - \boldsymbol{Q}\tilde{\boldsymbol{A}}|| + \sum_{i=1}^{M} ||\tilde{\boldsymbol{B}}_i\boldsymbol{Q} - \boldsymbol{Q}\tilde{\boldsymbol{C}}_i(\boldsymbol{L})|| \ : \ \boldsymbol{Q}\boldsymbol{Q}^T = \boldsymbol{E} \right) \tag{14}$$

where

– $\boldsymbol{Q}$ is a matrix of variables,

– the matrices $\boldsymbol{C}_i(\boldsymbol{L})$ are defined by $\boldsymbol{L}$ as $\tilde{\boldsymbol{C}}_i(\boldsymbol{L}) := \sum_{j=1}^{M} \boldsymbol{L}_{ij}\tilde{\boldsymbol{B}}_j$, and

– $|| \cdot ||$ denotes any matrix norm.

A set of globally optimal solutions (with zero objective value) gives the set of orthogonal symmetry transformations corresponding to a given permutation matrix $\boldsymbol{L}$. The union of $\boldsymbol{M}!$ such sets, taken over all permutation matrices $\boldsymbol{L}$, makes up the whole group $\boldsymbol{\mathcal{G}}$.

- **Unconstraint optimization** One can similarly formulate an optimization problem with respect to the elements of matrix $\boldsymbol{P}$:

$$\min_{\boldsymbol{P}} \left(||\boldsymbol{P}^T\boldsymbol{A}\boldsymbol{P} - \boldsymbol{A}|| + \sum_{i=1}^{M} ||\boldsymbol{P}^T\boldsymbol{B}_i\boldsymbol{P} - \boldsymbol{C}_i(\boldsymbol{L})|| \right), \tag{15}$$

where $\boldsymbol{C}_i(\boldsymbol{L}) := \sum_{j=1}^{M} \boldsymbol{L}_{ij}\boldsymbol{B}_j$.

In the case of trivial continuous subgroup of symmetry and finite discrete symmetry group, these problems have a finite set of optimal solutions, which may be found e.g. by a multi-start of a gradient descent method.

General Case

In general, problems (??) or (??) describe not only the discrete symmetries, but also the continuous ones. Identification of a continuum of solutions by numerical methods is problematic, therefore, before solving these problems, it would be helpful to factor out the continuous symmetries.

Conjecture 1. *Let $\boldsymbol{\mathcal{L}}$ be a subalgebra of the matrix Lie algebra* $\mathbf{so(N)}$ *with basis elements* $\boldsymbol{G_i^{(L)}}$*, corresponding to a closed subgroup* $\boldsymbol{G}$ *in* $\boldsymbol{SO(N)}$*, and let* $\boldsymbol{\mathcal{R}}$ *be the linear complement of* $\boldsymbol{\mathcal{L}}$ *in* $\mathbf{so(N)}$ *with basis elements* $\boldsymbol{G_j^{(\mathcal{R})}}$*. Then any element* $\boldsymbol{g} \in \mathbf{SO(N)}$ *can be represented as:*

$$\boldsymbol{g} = \boldsymbol{e}^{\sum_i \boldsymbol{a_i^{(\mathcal{L})} G_i^{(\mathcal{L})}}} \boldsymbol{e}^{\sum_j \boldsymbol{a_j^{(\mathcal{R})} G_j^{(\mathcal{R})}}}, \tag{16}$$

where $\{\boldsymbol{a_i^{(\mathcal{L})}}\}$ *and* $\{\boldsymbol{a_j^{(\mathcal{R})}}\}$ *are sets of real coefficients.*

Conjecture 2. *Let $\boldsymbol{\mathcal{L}}$ be a subalgebra of the matrix Lie algebra* $\mathbf{so(N)}$ *with basis elements* $\boldsymbol{G_i^{(L)}}$*, corresponding to a closed subgroup* $\boldsymbol{G}$ *in* $\boldsymbol{SO(N)}$*, and let* $\boldsymbol{\mathcal{R}}$ *be the linear complement of* $\boldsymbol{\mathcal{L}}$ *in* $\mathbf{so(N)}$ *with basis elements* $\boldsymbol{G_j^{(\mathcal{R})}}$*. Then any element* $\boldsymbol{g} \in \mathbf{SO(N)}$ *can be represented as:*

$$\boldsymbol{g} = \prod_{\boldsymbol{i}} \boldsymbol{e}^{\boldsymbol{a_i^{(\mathcal{L})} G_i^{(\mathcal{L})}}} \prod_{\boldsymbol{j}} \boldsymbol{e}^{\boldsymbol{a_j^{(\mathcal{R})} G_j^{(\mathcal{R})}}}, \tag{17}$$

Possible Applications in Quadratic Programming

- If some valid cuts are known already for the problem instance, then each linear symmetry of a problem may be used to double the set of valid cuts.
- If some symmetry $\boldsymbol{P} \in \boldsymbol{\mathcal{G}}$ is available, which maps a half-space $\{\boldsymbol{x} : \boldsymbol{a}^{\boldsymbol{T}}\boldsymbol{x} \geq \mathbf{0}\}$ into the half-space $\{\boldsymbol{x} : \boldsymbol{a}^{\boldsymbol{T}}\boldsymbol{x} \leq \mathbf{0}\}$ for some $\boldsymbol{a} \in \mathbb{R}^{\boldsymbol{N}}$, then the constraint $\boldsymbol{a}^{\boldsymbol{T}}\boldsymbol{x} \geq \mathbf{0}$ may be added to the set of problem constraints as a valid cut.
- If $\boldsymbol{\mathcal{G}}$ has a non-trivial continuous subgroup, such that any element of $\boldsymbol{\mathcal{D}}$ may be mapped onto some hyper-plane in $\mathbb{R}^{\boldsymbol{N}}$ by a corresponding $\boldsymbol{P} \in \boldsymbol{\mathcal{G}}$, then the problem dimension may be decreased by one.
- It is not necessary to find all symmetries of a problem to improve performance of solution algorithms, such as the branch and cut method.

Back to Phased Antenna Array Optimization Problem

$$\begin{cases} \boldsymbol{x^T G x \to \max,} \\ \boldsymbol{0 \le x^T H_1 x \le 1,} \\ \boldsymbol{...} \\ \boldsymbol{0 \le x^T H_n x \le 1,} \end{cases} \tag{18}$$

where $\boldsymbol{G}$ and $\boldsymbol{H_i, 0 \le i \le n}$ are $\boldsymbol{2n \times 2n}$ matrices. The sum of matrices $\boldsymbol{H_i}$ is positive definite.

The solutions of phased antenna array optimization problem are equivalent up to a shift of phases in all emitters by an equal angle, which corresponds to a rotation transformation of vector $\boldsymbol{x}$ in terms of problem (??). This symmetry may be taken into account by fixing one of the variables to zero, for example, $\boldsymbol{x_{2n} = 0}$.

The continuous subgroup of symmetries obtained by Theorem 2 consisted of phase shifting symmetries only.

In most of the instances, fixing the variable $\boldsymbol{x_1}$ led to an acceleration of the algorithm, the average acceleration of BARON was 0.95, which indicates the expediency of fixing one of the variables to zero when using this type of solver.

Conclusions

- It is expected that the proposed approach may be generalized to other types of problems in the mathematical programming.
- Conjectures 1 and 2 remain as open questions.
- Technical development of the outlined methods for detection of problem symmetries is also a subject of further research.
- Exploration of other definitions of symmetries of optimization problems, leading to wider groups of symmetries.

The end