

Copositive Approximations Decomposed Structured Subsets

Jared Miller

December 17, 2019

1 Introduction

Copositive and Completely Positive Programming are NP-hard convex conic optimization problems. Some standard cones in general optimization include:

$$\begin{aligned}\mathbb{S}^n &= \{M \in \mathbb{R}^{n \times n} \mid M = M^\top\} \\ \mathbb{S}_+^n &= \{M \in \mathbb{S}^n \mid x^\top M x \geq 0 \ \forall x \in \mathbb{R}^n\} \\ \mathcal{N}_+^n &= \{M \in \mathbb{S}^n \mid M_{ij} \geq 0\}\end{aligned}$$

These are the set of symmetric matrices, positive semidefinite symmetric matrices, and nonnegative matrices. The Copositive cone \mathcal{COP}^n and Completely Positive cone \mathcal{CP}^n are defined as:

$$\begin{aligned}\mathcal{COP}^n &= \{M \in \mathbb{S}^n \mid x^\top M x \geq 0 \ \forall x \in \mathbb{R}_+^n\} \\ \mathcal{CP}^n &= \{M \in \mathbb{S}^n \mid M = UU^\top, U \geq 0\}\end{aligned}$$

Copositive matrices only necessitate nonnegative quadratic forms over the real orthant, and completely positive matrices have nonnegative factorizations (more general than PSD matrices which require $\exists U \mid M = UU^\top$). By these properties, $\mathcal{CP}^n \subset \mathbb{S}_+^n \subset \mathcal{COP}^n$ [?]. The width of U necessary to form this completely positive decomposition is known as the completely positive rank $\text{cp-rank}(M)$. If such a U exists, then its width is bounded above by $n(n+1)/2$. The cones of copositive and completely positive matrices are dual to each other: $\mathcal{COP}^n = (\mathcal{CP}^n)^*$.

Most NP-hard combinatorial problems can be reformulated as copositive or completely positive programs. Polynomial nonnegativity over a (possibly noncompact) region can be certified by the existence of a copositive form (cite Olga's thesis). As an additional example, the stability number of a graph is the optimum value of a copositive program. Let G be a graph with adjacency matrix A_G , I be the identity matrix and $J = 11^\top$ denote the all-ones matrix. The stability number $\alpha(G)$ is:

$$\alpha_G = \max_{\lambda} \lambda \mid \lambda(A_G + I) - J \in \mathcal{COP}^n$$

See [?] on how to derive this copositive program, and [?] for a survey of copositive optimization.

2 Inner and Outer Approximations

Inner and outer approximations of these cones can be developed through semidefinite programming. As a completely positive matrix is both PSD and has nonnegative entries, the Doubly Nonnegative cone $\mathcal{DN}^n = \mathcal{N}^n \cap \mathbb{S}_+^n \supset \mathcal{CP}^n$. By duality, $(\mathcal{DN}^n)^* = \mathcal{N}^n + \mathbb{S}_+^n \subset \mathcal{COP}^n$. Parrilo defined a hierarchy of LMI-representable cones that successively approximate the copositive cone [?], where \mathcal{DN} is the first level of this hierarchy. Given a matrix M with quadratic form $q_M(x) = x^\top M x$, the r -th approximation to the copositive cone is:

$$\mathcal{K}^r = \{M \mid \sum_i^n (x_i^2) q_M(x^{\circ 2}) \in \text{SOS}\}$$

$x^{\circ^2} = x \circ x$ is the elementwise squaring of the entries of x , and SOS is the set of polynomials that can be represented as the sum of squared elements. The dual cones $(K^r)^*$ will then produce successively smaller outer approximations to the completely positive cone. See [?, ?] for additional copositive hierarchies.

In instances of large scale copositive optimization, the PSD constraints in the copositive hierarchy may lead to intractable semidefinite programs. These PSD matrices can be inner-approximated through techniques including structured subsets in optimization. Examples of these sets include (Scaled) Diagonally Dominant matrices, factor width, and block-factor width matrices. See citemajumdar17, zhengblock for more details on structured subsets.

Inner approximations of the completely positive cone can be found by exploiting properties of the \mathcal{DNN} cone. For $k = 1 \dots 4$, $\mathcal{DNN}^k = \mathcal{CP}^k$. When $k \geq 5$, $\mathcal{DNN}^k \supset \mathcal{CP}^k$. A study of matrices in \mathcal{DNN}^5 that are not in \mathcal{CP}^5 is conducted in [?]. This inclusion allows for scaled diagonally dominant nonnegative approximations [?], as well as extensions to factor width $\mathcal{FW}_{k \leq 4}^n$ [?]. Block Factor width 2 matrices with 2×2 blocks can also be used citezheng2019block (I have done this successfully, and it has better performance in time and cost as compared to SDD).

3 Employing Structure

Given the following conic program with variable $X \in \mathbb{S}^n$ sitting inside cone K :

$$p^* = \min_X \langle C, X \rangle$$

$$\langle A_i, X \rangle = b_i, i = 1, \dots, m,$$

$$X \in K^n,$$
(1)

(2)

This program will return an optimum value X^* with objective value p^* . In the usual case of strong duality, the dual objective $d^* = p^*$. This conic program may contain structure (sparse, symmetric, algebraic) that allows for the large cone K^n to be decomposed into a set of smaller cones $K_i^{n_i}$ under possible coupling constraints.

Symmetry and algebraic structure often allow for the block-diagonalization of conic programs. In case the cost and constraint matrices (C, A_i) arise from a common $*$ -algebra, there exists a matrix P that simultaneously block-diagonalizes all (C, A_i) . The allowable block structures and sizes under transformation are detailed by Maschke's theorem. Certain blocks may have nontrivial multiplicities (the contents of the block are repeated), which further reduces the complexity of optimization under algebraic structure. Symmetry is a specific case of algebraic structure, in which (C, A_i) are invariant under group action and the solution X commutes with this action.

The symmetry group in question is typically related to the underlying geometry of the problem. For structural designs, the symmetry group is determined by the shape of the truss (i.e. dihedral symmetry). In the case of stability number computations, the symmetry group is the automorphism group of the graph. Symmetry reduction is possible in copositive programming [?], but likely cannot generally be applied to completely positive programming (if $X = UU^T \in \mathcal{CP}^n$ and X is G -invariant and can be block diagonalized by a matrix P , then PU may not have all nonnegative entries).

Chordal sparsity offers another approach to structure. Let \mathcal{E} be the aggregate sparsity pattern of the matrices C, A_i with maximal cliques $\{C_k\}_{k=1}^p$. A chordal graph is a graph \mathcal{E} in which there exists a perfect elimination ordering; equivalently all cycles or length four or more have a 1-edge shortcut. Chordality can be checked in linear time, and edges can be added to a non-chordal graph to form a chordal embedding (triangulation/chordal extension). Define the cones $\mathbb{S}_+(\mathcal{E}, 0)$ as the set of PSD matrices that are zero outside \mathcal{E} , and $\mathbb{S}_+(\mathcal{E}, ?)$ as the set of matrices with entries on \mathcal{E} that have a PSD completion. These two sets of matrices are dual to each other, and have a characterization based on Agler's theorem and Grone's theorem:

Theorem 1 (Grone [?]). *Let $\mathcal{G}(\mathcal{V}, \mathcal{E})$ be a chordal graph with a set of maximal cliques $\{C_1, C_2, \dots, C_p\}$. Then, $X \in \mathbb{S}_+(\mathcal{E}, ?)$ if and only if*

$$X_k = E_{C_k} X E_{C_k}^T \in \mathbb{S}_+^{|C_k|}, \quad k = 1, \dots, p.$$

Theorem 2 (Agler’s Theorem [?]). *Let $\mathcal{G}(\mathcal{V}, \mathcal{E})$ be a chordal graph with a set of maximal cliques $\{\mathcal{C}_1, \mathcal{C}_2, \dots, \mathcal{C}_p\}$. Then, $Z \in \mathbb{S}_+^n(\mathcal{E}, 0)$ if and only if there exist $Z_k \in \mathbb{S}_+^{|\mathcal{C}_k|}$ where*

$$Z = \sum_{k=1}^p E_{\mathcal{C}_k}^T Z_k E_{\mathcal{C}_k}$$

Grone and Agler’s theorem can be used directly in optimization. As an example:

$$\begin{aligned} \min_X \quad & \langle C, X \rangle \\ \text{subject to} \quad & \langle A_i, X \rangle = b_i, i = 1, \dots, m, \\ & E_{\mathcal{C}_k} X E_{\mathcal{C}_k}^T \in \mathbb{S}_+^{|\mathcal{C}_k|}, \quad k = 1, \dots, p, \\ & \text{rank}(E_{\mathcal{C}_k} X E_{\mathcal{C}_k}^T) \leq t, \quad k = 1, \dots, p, \end{aligned} \tag{3}$$

will have the same objective value as an SDP in (C, A_i) , and solutions are mutually recoverable.

Similar completion results can be applied to copositive optimization. Sparse matrices can be completed into copositive matrices if the sparsity pattern is chordal and each clique submatrix is copositive (cone $\mathcal{COP}(\mathcal{E}, ?)$):

Theorem 3 (Hogben citehogben2005copositive). *Let $\mathcal{G}(\mathcal{V}, \mathcal{E})$ be a chordal graph with a set of maximal cliques $\{\mathcal{C}_1, \mathcal{C}_2, \dots, \mathcal{C}_p\}$. Then, $X \in \mathcal{COP}_+^n(\mathcal{E}, ?)$ if and only if*

$$X_k = E_{\mathcal{C}_k} X E_{\mathcal{C}_k}^T \in \mathcal{COP}^{|\mathcal{C}_k|}, \quad k = 1, \dots, p.$$

Like Grone, this is a completion result that assumes all diagonal elements are specified, but this will always hold as all diagonal entries are optimization variables.

A subset of chordal graphs are block-clique graphs, where two cliques intersect in at most one vertex. These are useful for completely positive completions $\mathcal{CP}^n(\mathcal{E}, ?)$:

Theorem 4 (Berman [?]). *Let $\mathcal{G}(\mathcal{V}, \mathcal{E})$ be a block-clique graph with a set of maximal cliques $\{\mathcal{C}_1, \mathcal{C}_2, \dots, \mathcal{C}_p\}$. Then, $X \in \mathcal{CP}^n(\mathcal{E}, ?)$ if and only if*

$$X_k = E_{\mathcal{C}_k} X E_{\mathcal{C}_k}^T \in \mathcal{CP}^{|\mathcal{C}_k|}, \quad k = 1, \dots, p.$$

By convex duality, the set $\mathcal{COP}(\mathcal{E}, ?)^* = \mathcal{CP}(\mathcal{E}, 0)$ when \mathcal{E} is chordal and $\mathcal{CP}(\mathcal{E}, ?)^* = \mathcal{COP}(\mathcal{E}, 0)$ when \mathcal{E} is block-clique. In this manner, sparsity can be applied to copositive and completely positive optimization. In some cases, conic programs contain symmetry and sparsity, and this will allow for efficient and (relatively more) tractable copositive and completely positive optimization.

4 Copositive Decomposed Structured Subsets

In (our IFAC work) we introduced Decomposed Structured Subsets $\mathcal{K}(\mathcal{E}, ?)$, in which $\mathcal{K} = K_{k=1}^p$ are a set of clique cones. Each clique $E_{\mathcal{C}_k}^T X E_{\mathcal{C}_k} \in K_k$. In the motivating example, a single large clique can be approximated to be Diagonally Dominant while all small cliques remain PSD. This mixing of cones increases the effectiveness of approximations. The same concept can be extended to algebraic structure, forming the cone \mathcal{K}^G where \mathcal{K} are the cones on each symmetric block.

In the copositive case, this can be performed by allowing each clique or symmetric block to lie on different steps of the copositive hierarchy (Parrilo/Resnick/otherwise). As the size of the copositive blocks increase (and the size of the PSD matrices in the hierarchy grow likewise), some of those PSD matrices may be approximated by structured subsets. This process can be accomplished for inner and outer copositive approximations: outer approximations (inner completely positive approximations) will require adding optimization variables to form a block-clique sparsity pattern.

5 Next Steps

Find numerical examples to show improvement. Profile different structured subsets. Clean this up and turn it into a paper. Find an application of copositive programming that has sparse structure