MTH 610:
Introduction to Semidefinite Programming
Final Project Report:
SDP and the Minimum Semidefinite Rank Problem

Sam Reynolds

June 11, 2023

Abstract

A faithful orthogonal representation (FOR) of a simple undirected graph G is an assignment of each vertex i of G to a vector $\boldsymbol{v}_i \in \mathbb{R}^d$ such that $\boldsymbol{v}_i$ is orthogonal to $\boldsymbol{v}_j$ if and only if ij is not an edge of G and $i \neq j$. An elementary result is that G has a FOR in $\mathbb{R}^n$ whenever G has n vertices, but typically G has a FOR in $\mathbb{R}^d$ with $d < n$. A natural question to ask is what the minimum dimension d is for a given graph G. Determining this minimum dimension is equivalent to the minimum semidefinite rank problem (MSR), which is the problem of determining the minimum rank of a positive semidefinite matrix with a given sparsity pattern. Indeed, such a matrix is the Gram matrix of a FOR. Most of the literature seems to focus on a combinatorial approach to the problem, although semidefinite programming can be used to obtain approximate solutions by minimizing the trace of a positive semidefinite matrix that satisfies given sparsity constraints. This project investigates the feasibility and limitations of the latter approach.

1 Introduction

Let G be a simple undirected graph on n vertices. We denote the vertex set of G by $V(G)$ and the edge set of G by $E(G)$, and an edge between veritces $i, j \in V(G)$ will be written as $ij \in E(G)$. The graph complement is denoted by $\overline{G}$. Given $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^d$, we denote their dot product with $\boldsymbol{v} \cdot \boldsymbol{w} = \boldsymbol{v}^\intercal \boldsymbol{w}$ and the Euclidean norm with $\|\boldsymbol{v}\| = \sqrt{\boldsymbol{v} \cdot \boldsymbol{v}}$.

Definition 1.1. Suppose that for each vertex $i \in V(G)$, we assign a vector $\boldsymbol{v}_i \in \mathbb{R}^d$ for some $d \in \mathbb{N}$ with the property that

$$ij \notin E(\overline{G}) \implies \boldsymbol{v}_i \cdot \boldsymbol{v}_j = 0 \ .$$

In this case, we call $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ an **orthogonal representation** of G. An orthogonal representation is called **faithful** if the converse condition holds:

$$\boldsymbol{v}_i \cdot \boldsymbol{v}_j = 0 \implies ij \notin E(\overline{G}) \ .$$

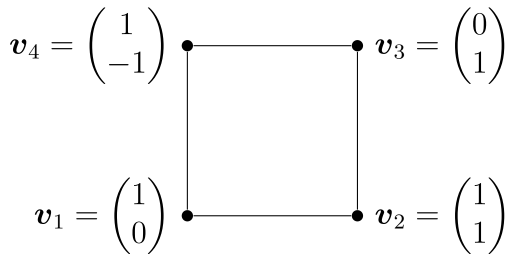

Figure 1: A faithful orthogonal representation of the 4-cycle.

We sometimes use the abbreviation **FOR** for "faithful orthogonal representation."

Example 1.2. The cannonical basis $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ is an orthogonal representation of G.

Example 1.3. Consider the 4-cycle C_4 depicted in Figure 1. Let us choose vertices $1, 3$ that are not adjacent and assign the vectors

$$\boldsymbol{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \ , \quad \boldsymbol{v}_3 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \ .$$

For the other two vectors, pick

$$\boldsymbol{v}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \ , \quad \boldsymbol{v}_4 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \ .$$

Then $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4$ is an orthogonal representation of C_4, since

$$\boldsymbol{v}_1 \cdot \boldsymbol{v}_3 = \boldsymbol{v}_2 \cdot \boldsymbol{v}_4 = 0$$

and this orthogonal representation is faithful because

$$\begin{aligned} \boldsymbol{v}_1 \cdot \boldsymbol{v}_2 &= 1 \neq 0 \ , \\ \boldsymbol{v}_1 \cdot \boldsymbol{v}_4 &= 1 \neq 0 \ , \\ \boldsymbol{v}_2 \cdot \boldsymbol{v}_3 &= 1 \neq 0 \ , \\ \boldsymbol{v}_3 \cdot \boldsymbol{v}_4 &= -1 \neq 0 \ . \end{aligned}$$

Notice that in this example, vectors in $\mathbb{R}^2$ were used to obtain an orthogonal representation of a graph on $n = 4$ vertices.

This project is concerned with determining the minimum dimension d such that G has a FOR in $\mathbb{R}^d$. Determining this minimum dimension is equivalent to minimizing the rank of a matrix with a specified sparsity pattern.

Definition 1.4. A **generalized adjacency matrix** of G is a square matrix $A \in \mathbb{R}^{n \times n}$ such that for every $i \neq j$ it holds that $A_{ij} = 0$ if and only if $ij \notin E(G)$. Note that the diagonal of A is left free. We denote the set of all generalized adjacency matrices of G by $\operatorname{gam}(G)$.

The minimum dimension of a FOR is equivalent to the **minimum semidefinite rank problem**, given by

$$\text{(MSR)} \qquad \operatorname{msr}(G) = \min\{\operatorname{rank}(A) : A \in \operatorname{gam}(G),\ A = A^\intercal,\ A \succeq 0\}\ ,$$

where $A \succeq 0$ denotes that A is positive semidefinite. The connection between MSR and finding the minimum dimension of a FOR can be seen by supposing that the columns of a matrix $B \in \mathbb{R}^{d\times n}$ are a FOR of G, and noting that the Gram matrix $A = B^\intercal B \in \operatorname{gam}(G)$ has $\operatorname{rank}(A) = \operatorname{rank}(B)$.

This project is particularly focused on using semidefinite programming (SDP) to determine $\operatorname{msr}(G)$. There are at least two complications:

1. The sparsity constraints are nonconvex.
2. The objective function $\operatorname{rank}(X)$ is nonlinear, nonsmooth, and nonconvex, even with the domain restricted to the PSD cone.

We will see that the **minimum semidefinite trace** problem

$$\text{(MST)} \qquad \begin{aligned} \min\ & \operatorname{tr}(X) \\ \text{st}\ & X_{ij} = 0 \text{ if } ij \in E\big(\overline{G}\big) \\ & X_{ij} > 0 \text{ if } ij \in E(G) \\ & X \succeq 0 \end{aligned}$$

can be expressed as an SDP (by introducing slack variables), and surprisingly often we find that

$$\operatorname{rank}(X) \approx \operatorname{msr}(G)\ .$$

Section 2 discusses the modifications to the sparsity constraints, and Section 3 discusses the relaxation of the objective function. In Section 4, we conduct numerical experiments to explore when the modified problem (MST) coincides with (MSR).

2 Reformulation of the Sparsity Constraints

Our first step in converting (MST) to an SDP will be to find contraints of the form $\langle A_i, X\rangle = b_i$ that are equivalent to $X \in \operatorname{gam}(G)$, where $\langle\cdot,\cdot\rangle$ denotes the Frobenius inner product.

Recall that, for $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^d$, the trace operator satisfies

$$\operatorname{tr}(\boldsymbol{v}\boldsymbol{w}^\intercal) = \boldsymbol{v}\cdot\boldsymbol{w}\ .$$

Define the symmetric matrix

$$A^{(i,j)} = \frac{1}{2}(\boldsymbol{e}_i\boldsymbol{e}_j^\intercal + \boldsymbol{e}_j\boldsymbol{e}_i^\intercal) \in \mathbb{R}^{n\times n}\ ,$$

where $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ is the cannonical basis of $\mathbb{R}^n$. Then notice that for a symmetric matrix $X = X^\intercal \in \mathbb{R}^{n\times n}$ we have

$$\left\langle A^{(i,j)}, X\right\rangle = \operatorname{tr}(X\, A^{(i,j)}) = \frac{1}{2}(\boldsymbol{e}_j^\intercal X \boldsymbol{e}_i + \boldsymbol{e}_i^\intercal X \boldsymbol{e}_j) = X_{ij} \ .$$

So $X \in \operatorname{gam}(G)$ if and only if

$$\left\langle A^{(i,j)}, X\right\rangle \rightsquigarrow \begin{cases} = 0 & \text{if } ij \notin E(G) \text{ and } i \neq j \ , \\ \neq 0 & \text{if } ij \in E(G) \end{cases}$$

In the first case with ij in the complement of the edge set, we can take $b^{(i,j)} = 0$.

Remark 2.1. The inequality constraints

$$X_{ij} \neq 0 \text{ if } ij \in E(G)$$

imply that the feasible region of (MSR) is nonconvex, and therefore is not a spectrahedron. In what follows, we instead impose the positivity conditions

$$X_{ij} > 0 \text{ if } ij \in E(G) \ .$$

As we will see, the feasible region can be expressed as a spectrahedron by the introduction of slack variables. The following example shows us that working with a more convenient geometry comes at the cost of finding optimal solutions on some graphs.

Example 2.2. Consider once again the 4-cycle depicted in Figure 1. In Figure 2, we have drawn generic vectors which construct a FOR of C_4 in $\mathbb{R}^d$, with $d = \operatorname{msr}(G) = 2$. Inspection of this diagram leads to the conclusion that for every faithful orthogonal representation of C_4, there must be at least one pair of vectors whose dot product is negative. In other words, there is no generalized adjacency matrix of C_4 with nonnegative entries and $A \succeq 0$.

We proceed by imposing the positivity conditions

$$X_{ij} > 0 \text{ if } ij \in E(G) \ .$$

Since the edge set is finite and no upper bounds are placed on the diagonal entries of X, we can fix $\varepsilon > 0$ and write $X_{ij} \geq \varepsilon$. Next, we introduce the slack variables

$$s^{(i,j)} \geq 0 \ , \quad \left\langle A^{(j,j)}, X\right\rangle - s^{(i,j)} = \varepsilon \ , \quad ij \in E(G) \ .$$

Let $S = \operatorname{diag}\big(s^{(i,j)} \mid ij \in E(G)\big)$ be a diagonal matrix with the slack variables. We thereby arrive at the constraints

$$\begin{aligned} &\left\langle A^{(i,j)}, X\right\rangle = 0 \ , \ i \neq j \ , \ ij \notin E(G) \ , \\ &\left\langle A^{(i,j)}, X\right\rangle - s^{(i,j)} = \varepsilon \ , \ ij \in E(G) \ , \\ &\begin{pmatrix} X & 0 \\ 0 & S \end{pmatrix} \succeq 0 \ . \end{aligned}$$

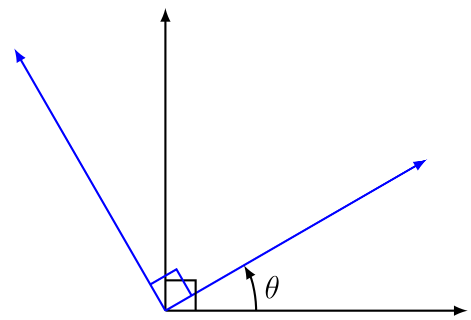

Figure 2: A minimum-dimension FOR of the four-cycle necessarily has at least one pair of vectors $\boldsymbol{v}_i, \boldsymbol{v}_j$ for which $\boldsymbol{v}_i \cdot \boldsymbol{v}_j < 0$. The two black vectors correspond to an independent set of C_4, with the blue vectors corresponding to the other independent set. In order for one of the blue vectors to have a positive dot product with both of the black vectors, the angle θ must be strictly between 0 and $\pi/2$, but this forces the other blue vector to have a negative dot product with one of the black vectors.

Lastly, we can state these constraints using the augmented matrices

$$\widehat{X} = \begin{pmatrix} X & 0 \\ 0 & S \end{pmatrix} \ , \quad \widehat{A}^{(i,j)} = \begin{pmatrix} A^{(i,j)} & 0 \\ 0 & -D^{(i,j)} \end{pmatrix}$$

where $D^{(i,j)} \in R^{e \times e}$ is a matrix of zeros except for a one in the appropriate position on the diagonal for which $S\, D^{(i,j)} = s^{(i,j)}$. Then take

$$b^{(i,j)} = \begin{cases} \varepsilon & \text{if } ij \in E(G) \ , \\ 0 & \text{else,} \end{cases}$$

and at last we can write the constraints in the cannonical form

$$\widehat{X} \succeq 0 \ , \ \left\langle \widehat{A}^{(i,j)}, \widehat{X} \right\rangle = b^{(i,j)} \ , \ 1 \le i < j \le n \ .$$

3 Relaxation of the Objective Function

This section is devoted to justifying the relaxation of the objective function $\text{rank}(X)$ to $\text{tr}(X)$. The results of this section are based on work by Fazel, Hindi, and Boyd [4].

Definition 3.1. Let $X \in \mathbb{R}^{m \times n}$ be a real matrix with singular values $\sigma_1, \ldots, \sigma_r$, with $r = \text{rank}(X)$. The **nuclear norm** of X is given by

$$\|X\|_* := \sum_{k=1}^{r} \sigma_k \ .$$

Remark 3.2. If $X \succeq 0$, then we have that the singular values of X are precisely its eigenvalues. We would therefore have that

$$\|X\|_* = \sum_{k=1}^{r} \lambda_k = \text{tr}(X) \ .$$

Definition 3.3. Let $C \subseteq \mathbb{R}^n$ be a convex set, and let $f : C \to \mathbb{R}$ be a real-valued function. The **convex envelope** of f on C is the largest convex function g for which $g(x) \leq f(x)$ for all $x \in C$.

Theorem 3.4. Consider the convex set

$$C = \left\{ X \in \mathbb{R}^{m \times n} : \|X\| \leq 1 \right\}$$

where $\|X\|$ denotes the Frobenius norm. Then the nuclear norm $X \mapsto \|X\|_*$ is the convex envelope of $X \mapsto \operatorname{rank}(X)$.

In the context of (MST), suppose that we constain X so that

$$\|X\|^2 \leq 1 \ .$$

The sparsity constraints $X_{ij} \geq \varepsilon$ would suggest that

$$\|X\|^2 = \operatorname{tr}(X^2) \geq n\varepsilon^2$$

and it follows that we ought to choose ε such that

$$0 < \varepsilon \leq \frac{1}{\sqrt{n}} \ .$$

4 Numerical Investigation

In this section, we compare the results of numerically solving (MST) to the exact result of solving (MSR) using combinatorial techniques: see Appendix 5 for details of how to determine $\operatorname{msr}(G)$ by hand. Our Python implementation of a solver for (MST) can be found in the GitHub repository

`https://github.com/samreynoldsmath/msr.git`

To determine $\operatorname{rank}(X)$, we compute the singular value decomposition and normalize the singular values with

$$\widehat{\sigma}_i := \frac{\sigma_i}{\sigma_1}$$

where σ_1 is the largest singular value. We then count the number of normalized singular values greater than a fixed tolerance $\tau > 0$. Thoughout all experiments, we use the parameters $\varepsilon = 0.01/\sqrt{n}$ and $\tau = 10^{-4}$.

Example 4.1. In this example we consider five graphs:

1. The four cycle C_4
2. The five cycle C_5

3. The house graph (C_5 with an additional edge)
4. The six cycle with an additional edge such that G has two induced four cycles
5. The six cycle with an additional edge such that G has an induced cycle of lenth 3 and another of length 5

For each graph, we compare the exact value of $\operatorname{msr}(G)$ to the approximation obtained by numerically solving (MST). We also show the normalized singular values of X. The output is:

```
Four cycle:
[1.00e+00 5.00e-01 5.00e-01 1.46e-05]
        Exact    2
        Approx   3
Five cycle:
[1.00e+00 6.18e-01 6.18e-01 3.30e-06 3.30e-06]
        Exact    3
        Approx   3
House:
[1.00e+00 5.54e-01 4.62e-01 4.90e-07 1.26e-07]
        Exact    3
        Approx   3
Even chord:
[1.00e+00 6.00e-01 6.00e-01 4.00e-01 2.00e-01 2.13e-08]
        Exact    4
        Approx   5
Odd chord:
[1.00e+00 6.98e-01 5.76e-01 1.89e-01 4.28e-06 4.34e-07]
        Exact    4
        Approx   4
```

We see that the algorithm is successful in three out of five attempts. It is unsurprising that one of the unsuccessful attempts is the four cycle, in light of the conclusion of Example 2.2. Notice also that the algorithm is unsuccessful in the case with a six cycle with a chord splitting it into two four cycles, but the final case of the chord splitting the six cycle into two odd cycles is successful.

Example 4.2. Up to isomorphism, there are 112 distinct connected graphs on six vertices; see [3] for a complete enumeration. In previous work, computed the exact value of $\operatorname{msr}(G)$ for all 112 of these graphs using the techniques described in Appendix 5.[1] We undertook the tedious process of encoding all 112 of these graphs in order to solve (MST) for each of them, and then compared them with the exact value of $\operatorname{msr}(G)$.

[1]Actually, there are three exceptions in this list that cannot be found using only basic techniques. A more advanced technique, which seems not yet to have appeared in the literature was used for the remaining three.

We found that the $\operatorname{rank}(X)$ differed from $\operatorname{msr}(G)$ in 38 of 112 cases, three of which differed by 2, the remaining 35 differing by 1. Upon careful examination, an interesting phenomenon was found: *in every case where* $\operatorname{rank}(X)$ *differed from* $\operatorname{msr}(G)$*, there was an induced even cycle of* G*.*

5 Conclusion

We have seen that a relaxation of the minimum semidefinite rank problem, both in the constraints and the objective function, leads to a semidefinite program. Surprisingly, despite these relaxations, the solution of the modified problem (MST) very often coincides with the solution to the original problem (MSR). Furthermore, numerical evidence seems to suggest that the cases where the relaxation fails have a common property:

Conjecture 5.1. Suppose that (MST) is solved by X with $\operatorname{rank}(X) > \operatorname{msr}(G)$. Then G has an induced cycle of even order.

Further investigation is required to verify or find a counterexample to this conjecture.

References

[1] M. Booth, P. Hackney, B. Harris, C. R. Johnson, M. Lay, T. D. Lenker, L. H. Mitchell, S. K. Narayan, A. Pascoe, and B. D. Sutton. On the minimum semidefinite rank of a simple graph. *Linear Multilinear Algebra*, 59(5):483–506, 2011.

[2] M. Booth, P. Hackney, B. Harris, C. R. Johnson, M. Lay, L. H. Mitchell, S. K. Narayan, A. Pascoe, K. Steinmetz, B. D. Sutton, and W. Wang. On the minimum rank among positive semidefinite matrices with a given graph. *SIAM J. Matrix Anal. Appl.*, 30(2):731–740, 2008.

[3] D. Cvetković and M. Petrić. A table of connected graphs on six vertices. *Discrete Math.*, 50(1):37–49, 1984.

[4] M. Fazel, H. Hindi, and S. Boyd. A rank minimization heuristic with application to minimum order system approximation. In *Proceedings of the 2001 American Control Conference. (Cat. No.01CH37148)*, volume 6, pages 4734–4739 vol.6, 2001.

Appendix A: Basic Properties of MSR

Recall that an **induced subgraph** of G is a subgraph of G obtained by deleting a subset of the veritces of G and only those edges that are incident to the deleted vertices. An **induced cover** of G is a collection of induced subgraphs $\{G_1, \ldots, G_m\}$ such that $G = \bigcup_{k=1}^{m} G_k$. The **neighborhood** of $i \in V(G)$ is the set of all vertices in G that are adjacent to i, and is denoted by $N(i)$.

Many of the properties (or generalizations of them) can be found in [2, 1], for example.

Proposition .2. Let G be a simple graph on n vertices.

(i) $\operatorname{msr}(G) \leq n-1$.

(ii) If G has no isolated vertices and $\alpha(G)$ is its independence number, then

$$\alpha(G) \leq \operatorname{msr}(G) .$$

(iii) If $i \in V(G)$, then

$$\operatorname{msr}(G-i) \leq \operatorname{msr}(G) .$$

In general, if H is an induced subgraph of G, then $\operatorname{msr}(H) \leq \operatorname{msr}(G)$.

(iv) If $\{G_1, \ldots, G_m\}$ is an induced cover of G, then

$$\operatorname{msr}(G) \leq \sum_{k=1}^{m} \operatorname{msr}(G_k) .$$

Equality holds when, for all $k \neq \ell$, G_k and G_ℓ share at most one vertex in common.

(v) If i is a vertex with exactly one neighbor in G, then

$$\operatorname{msr}(G) = 1 + \operatorname{msr}(G-i) .$$

(vi) If i is a vertex of G with $N(i) = V(G) \setminus \{i\}$, then

$$\operatorname{msr}(G) = \operatorname{msr}(G-i) .$$

(vii) If i, j are adjacent vertices such that $N(i) \setminus \{j\} = N(j) \setminus \{i\} \neq \varnothing$, then

$$\operatorname{msr}(G-i) = \operatorname{msr}(G-j) = \operatorname{msr}(G) .$$

(viii) If G' is obtained by subdividing an edge of G, then

$$\operatorname{msr}(G') = 1 + \operatorname{msr}(G) .$$

(ix) If $ij \in E(G)$, then

$$\operatorname{msr}(G) - 1 \leq \operatorname{msr}(G-ij) \leq \operatorname{msr}(G) + 1 .$$

(x) If ij is a cut-edge of G, then

$$\operatorname{msr}(G) = 1 + \operatorname{msr}(G-ij) .$$

Proposition .3. The minimum semidefinite rank can be computed exactly for the following simple graphs on n veritces:

(i) A clique: $\operatorname{msr}(K_n) = 1$

(ii) A cycle: $\operatorname{msr}(C_n) = n-2$

(iii) A tree: $\operatorname{msr}(T) = n-1$