Group-invariant Semidefinite Programming
and
Applications

by

Minyan Ni

An essay
presented to the University of Waterloo
in fulfilment of the
essay requirement for the degree of
Master of Mathematics
in
Combinatorics and Optimization

Waterloo, Ontario, Canada, 2006

Abstract

This essay considers semidefinite programming problems that exhibit a special form of symmetry called group-invariance. We demonstrate the effect of such symmetries on certain path-following interior-point algorithms, and highlight a reduction technique that is particularly useful on certain group-invariant semidefinite programming problems. Two applications of group-invariant semidefinite programming problems–one in truss design and the other in graph theory–are presented.

Acknowledgements

To my supervisor, Dr. Chek Beng Chua, for his wise advice, great encouragement and continuous support during my master's study. To Professor Michael Best for his comments and careful reading of the draft.

To my dear friends who gave me great help and made my life in Waterloo a wonderful experience: H. Xu, S. Pan, G. Kemkes, H. Wei, A. Li, S. Luo, Z. Zhou, etc.

To my dear parents and friends in China for their encouragement and understanding in these two years.

Contents

Chapter 1

Introduction

Semidefinite programming (SDP) can be regarded as an extension of linear programming. In SDP problems, the nonnegative variables are replaced by symmetric, positive semidefinite matrix variables. Research interests in SDP have increased tremendously during the last fifteen years, which is spurred by the development of efficient algorithms and the potential applications in various different areas.

Several efficient polynomial-time algorithms for SDP have been proposed since the late-1980s. These algorithms can be classified into three categories: ellipsoid methods, interior-point methods, and bundle methods. Among the algorithms proposed so far, the primal-dual interior-point algorithms are most popular in practice. See the handbook [1] for an overview of this area.

SDP has various applications in combinatorial optimization, stochastic optimization, engineering, and so forth. In combinatorial optimization, SDP plays an important role in improving the quality of the bounds for some interesting but hard problems. Also, many applications of SDP can be found in engineering. Structural optimization is one important area of these engineering applications.

This essay is organized as follows. We begin by reviewing some fundamentals of SDP. In Chapter 2, we study a special class of SDP problems where each problem is invariant under a group acting on its variables. Several applications of such SDP problems are discussed in Chapters 3 and 4. In Chapter 3, an optimal design of symmetric truss is obtained by formulating it into a group-invariant SDP problem. In Chapter 4, we get an improved lower bound on the crossing number of complete bipartite graphs through an invariant SDP formulation. Chapter 5 gives a summary of this essay and mentions some potential research areas.

1.1 Notation

In this essay, the following notations are used.

$\mathbb{R}^n$:	the space of n-dimensional real vectors
$\mathbb{R}^n_+$:	the space of nonnegative vectors in $\mathbb{R}^n$
$\mathbb{R}^{n\times n}$:	the space of $n \times n$ real matrices
$\mathbb{S}^n$:	the space of symmetric $n \times n$ real matrices
$\mathbb{S}^n_+$:	the cone of symmetric positive semidefinite $n \times n$ real matrices
$\mathbb{S}^n_{++}$:	the cone of symmetric positive definite $n \times n$ real matrices
$\mathcal{N}(\mathcal{A})$:	the null space of the linear operator $\mathcal{A}$
$\mathcal{R}(\mathcal{A})$:	the range of the linear operator $\mathcal{A}$
$\bullet$:	the inner product $(X, S) \in \mathbb{S}^n \bigoplus \mathbb{S}^n \mapsto \text{trace}(XS)$

1.2 Basic Theorems of Semidefinite Programming

We consider SDP problems in the following standard form:

$$(\mathcal{P}) \qquad \begin{array}{lrcl} \inf & C \bullet X & & \\ \text{s.t.} & A^{(i)} \bullet X & = & b_i,\ i = 1, \ldots, m, \\ & X & \in & \mathbb{S}^n_+, \end{array}$$

where $A^{(1)}, \ldots, A^{(m)}, C \in \mathbb{S}^n$ and $b \in \mathbb{R}^m$ are given. Using the linear operator $\mathcal{A} : \mathbb{S}^n \mapsto \mathbb{R}^m$ defined by

$$[\mathcal{A}(X)]_i = A^{(i)} \bullet X \quad \text{for} \quad i = 1, \ldots, m,$$

the first constraint can be written as

$$\mathcal{A}(X) = b.$$

The Lagrange dual problem of $(\mathcal{P})$ is

$$(\mathcal{D}) \qquad \begin{array}{lrcl} \sup & b^T y & & \\ \text{s.t.} & \sum_{i=1}^{m} y_i A^{(i)} + S & = & C, \\ & S & \in & \mathbb{S}^n_+. \end{array}$$

Throughout this essay we make the following two assumptions:

Assumption 1.2.1 The matrices $A^{(1)}, \ldots, A^{(m)}$ are linearly independent. That is, the linear operator $\mathcal{A}$ is surjective.

Assumption 1.2.2 Both $(\mathcal{P})$ and $(\mathcal{D})$ have strictly feasible solutions. Namely, there exist a feasible solution X of $(\mathcal{P})$ and a feasible solution (y, S) of $(\mathcal{D})$ such that $X \in \mathbb{S}^n_{++}$ and $S \in \mathbb{S}^n_{++}$.

The first assumption can be made without loss of generality. If $A^{(1)}, \ldots, A^{(m)}$ are not linearly independent, then either the system $A^{(i)} \bullet X = b_i \, (i = 1, \ldots, m)$ has no solution, or it has some redundant equations. Those redundant equations can be removed without changing the solution set, upon which we get an equivalent problem satisfying our assumption.

Under the first assumption, if (y, S) is a feasible solution of $(\mathcal{D})$, then y is uniquely determined by S via $\sum_{i=1}^m y_i A^{(i)} = C - S$. Thus we may refer to S as the dual feasible solution.

Under the second assumption, both $(\mathcal{P})$ and $(\mathcal{D})$ have finite optimal solutions; see Nesterov and Nemirovskii [9]. This condition, which is necessary for interior-point algorithms, is called the Slater Constraint Qualification.

Suppose X and (y, S) are feasible for $(\mathcal{P})$ and $(\mathcal{D})$, respectively. Then the *duality gap* $(C \bullet X - b^T y)$ is

$$C \bullet X - b^T y = \left(\sum_{i=1}^{m} y_i A^{(i)} + S \right) \bullet X - \sum_{i=1}^{m} (A^{(i)} \bullet X) y_i = X \bullet S.$$

This observation leads to the well-known weak duality relation.

Proposition 1.1 (Weak duality relation) *Let X and (y, S) be feasible for $(\mathcal{P})$ and $(\mathcal{D})$, respectively. Then*

$$C \bullet X \geq b^T y.$$

Proof
Since X and S are positive semidefinite, so is $X^{1/2} S X^{1/2}$. Therefore,

$$C \bullet X - b^T y = X \bullet S = \mathrm{trace}(XS) = \mathrm{trace}(X^{1/2} S X^{1/2}) \geq 0.$$

□

As an immediate consequence of the weak duality relation, we have the following sufficient condition for optimality.

Corollary 1.2 *Let X and (y, S) be feasible for $(\mathcal{P})$ and $(\mathcal{D})$, respectively, and $X \bullet S = 0$. Then X and (y, S) are optimal for their respective problems.*

Note that if $XS = 0$, then

$$X \bullet S = \mathrm{trace}(XS) = 0.$$

On the other hand, if $X \bullet S = 0$, then

$$\mathrm{trace}(X^{1/2} S X^{1/2}) = 0.$$

If X and S are further feasible, then $X^{1/2}SX^{1/2} \in \mathbb{S}^n_+$ implies that all the eigenvalues of $X^{1/2}SX^{1/2}$ are nonnegative. Thus all the eigenvalues of $X^{1/2}SX^{1/2}$ must be zero. Therefore,

$$0 = X^{1/2}SX^{1/2} = (X^{1/2}S^{1/2})(X^{1/2}S^{1/2})^T.$$

It follows that $X^{1/2}S^{1/2} = 0$, and hence

$$XS = X^{1/2}(X^{1/2}S^{1/2})S^{1/2} = 0.$$

Consequently, we may alternatively use the equivalent condition $XS = 0$ in the above corollary.

Unfortunately, the above condition is not necessary in general. It is possible that X and (y, S) are optimal, while the duality gap is strictly positive. However, when both $(\mathcal{P})$ and $(\mathcal{D})$ have strictly feasible solutions, the above condition is necessary.

Theorem 1.3 (Strong duality theorem) *Suppose both* $(\mathcal{P})$ *and* $(\mathcal{D})$ *have strictly feasible solutions. Then both* $(\mathcal{P})$ *and* $(\mathcal{D})$ *have optimal solutions and their optimal objective values are the same.*

The proof of the strong duality theorem is not trivial; see Nesterov and Nemirovskii [9] or Chapter 4 of the handbook [1] for a detailed proof.

From Corollary 1.2 and the strong duality theorem, we know that X and S are optimal for $(\mathcal{P})$ and $(\mathcal{D})$, respectively, if and only if the conditions below are satisfied:

$$\begin{aligned} A^{(i)} \bullet X &= b_i \ (i = 1, \ldots, m), \quad X \in \mathbb{S}^n_+, \\ \sum_{i=1}^m y_i A^{(i)} + S &= C, \quad S \in \mathbb{S}^n_+, \\ XS &= 0. \end{aligned}$$

In subsequent sections, we briefly describe how Newton's method may be employed to approximately solve the above nonlinear system.

1.3 Central Path

Interior-point methods are among the most efficient methods for solving SDP problems. The central path performs an important function in the study of interior-point methods.

Path-following algorithms are an important family of interior-point methods. (Another family of interior-point methods is based on potential functions; see, for example, [1].) The path-following algorithms restrict each iterate to a *neighborhood* of the so-called *central path* and trace the central path to get an approximate optimal solution.

The *central path* is defined as the set of solutions $\{(X(\mu), S(\mu)) : \mu > 0\}$ to

$$(CP_\mu) \qquad \begin{aligned} A^{(i)} \bullet X &= b_i \ (i = 1, \ldots, m), \quad X \in \mathbb{S}^n_+, \\ \sum_{i=1}^m y_i A^{(i)} + S &= C, \quad S \in \mathbb{S}^n_+, \\ XS &= \mu I, \end{aligned}$$

where I is the $n \times n$ identity matrix. We see that (CP_μ) is actually a system of perturbed optimality conditions.

The solution to the above system of equations clearly gives strictly feasible solutions to $(\mathcal{P})$ and $(\mathcal{D})$, since the last equation implies both X and S are positive definite. On the other hand, we will show that if Assumption 1.2.2 holds (i.e., both $(\mathcal{P})$ and $(\mathcal{D})$ have strictly feasible solutions), then (CP_μ) has a unique solution for each $\mu > 0$.

We begin by defining a barrier function for $\mathbb{S}^n_+$. Let $f : \mathbb{S}^n_{++} \mapsto \mathbb{R}$ be defined by

$$f(X) := -\ln \det X.$$

Observe that when X converges to a point on the boundary of $\mathbb{S}^n_+$, the value of the barrier function goes to infinity.

We need to use the notation $\odot$ in the following proof. Here $\odot$ is known as the symmetric Kronecker product and defined by

$$(P \odot Q)U = \frac{1}{2}(PUQ^T + QUP^T),$$

for all $n \times n$ matrices P, Q and U.

Proposition 1.4 *The barrier function f is strictly convex over its domain $\mathbb{S}^n_{++}$.*

Proof

We can verify this by calculating the derivatives of f at each $X \in \mathbb{S}^n_{++}$.

The first derivative of f is $f'(X) = -X^{-1}$, and the second derivative is $f''(X) = X^{-1} \odot X^{-1}$. Since for each $n \times n$ matrix H, we have

$$(f''(X)H) \bullet H = \text{trace}[X^{-1}HX^{-1}H] = \text{trace}[(X^{-1/2}HX^{-1/2})^2] \geq 0$$

with equality holding if and only if $H = 0$, we know f is strictly convex. $\square$

Now consider the following problem for each $\mu > 0$:

$$(P_\mu) \qquad \begin{array}{ll} \inf & C \bullet X + \mu f(X) \\ \text{s.t.} & A^{(i)} \bullet X \;=\; b_i, \; i = 1, \ldots, m. \end{array}$$

In this problem, $X \in \mathbb{S}^n_{++}$ is implied since the the domain of f is $\mathbb{S}^n_{++}$. Also note that the objective function is strictly convex over its domain; thus, (P_μ) has at most one optimal solution for each $\mu > 0$.

Theorem 1.5 *Suppose both $(\mathcal{P})$ and $(\mathcal{D})$ have strictly feasible solutions. Then there exists a unique solution $(X(\mu), S(\mu))$ to the central path equations (CP_μ).*

Proof
First we prove the uniqueness. Note that the Karush-Kuhn-Tucker optimality conditions for (P_μ) are both necessary and sufficient in our case, since the objective function is convex and the constraints are linear. The optimality conditions for (P_μ) are

$$\begin{aligned} A^{(i)} \bullet X &= b_i, \quad i = 1, \dots, m \quad (X \in \mathbb{S}^n_{++}) \\ -\sum_{i=1}^{m} y_i A^{(i)} - \mu X^{-1} + C &= 0. \end{aligned}$$

If we set $S := \mu X^{-1} \in \mathbb{S}^n_{++}$, then the optimality conditions are exactly the central path equations (CP_μ). Therefore, (X, y, S) solves (CP_μ) if and only if X is an optimal solution to $(P_\mu), S = \mu X^{-1}$ and y solves $\sum_{i=1}^m y_i A^{(i)} = C - S$. We have shown that (P_μ) has at most one optimal solution. Thus, this also proves that (CP_μ) has at most one optimal solution.

To prove the existence, we use the Weierstrass theorem. So we aim to reduce (P_μ) to the problem of taking the infimum of a continuous function over a nonempty compact set. However, our current feasible region may be unbounded and relatively open.

Let $\hat{X}$ and $(\hat{y}, \hat{S})$ be strictly feasible solutions for $(\mathcal{P})$ and $(\mathcal{D})$, respectively. Since we have shown that

$$C \bullet X = X \bullet \hat{S} + b^T \hat{y},$$

the original objective function $C \bullet X + \mu f(X)$ differs from $\hat{S} \bullet X + \mu f(X)$ only by the constant $b^T \hat{y}$. Thus, we may replace the objective function in (P_μ) by $\hat{S} \bullet X + \mu f(X)$. Furthermore, we can add the constraint

$$\hat{S} \bullet X + \mu f(X) \leq \hat{S} \bullet \hat{X} + \mu f(\hat{X}),$$

since $\hat{X}$ is a feasible solution for (P_μ). We now show

$$\mathcal{F} := \{X \in \mathbb{S}^n_{++} : \mathcal{A}(X) = b, \quad \hat{S} \bullet X + \mu f(X) \leq \hat{S} \bullet \hat{X} + \mu f(\hat{X})\}$$

is a nonempty compact set, and $\hat{S}\bullet X+\mu f(X)$ is continuous over $\mathcal{F}$. Clearly, it is nonempty since it contains $\hat{X}$. We show it is bounded by showing that the eigenvalues $\lambda_j(X)$ of X on $\mathcal{F}$ are bounded. Suppose $\alpha > 0$ is the smallest eigenvalue of $\hat{S}$. Define

$$\phi(\lambda) = \alpha\lambda - \mu \ln \lambda \quad \text{for} \quad \lambda > 0.$$

Note that ϕ is strictly convex, and the infimum of ϕ is attained at μ/α. Also note that as λ tends to 0 or ∞, ϕ tends to ∞. For each $X \in \mathcal{F}$, we have

$$\begin{aligned}
\sum_j \phi(\lambda_j(X)) &= \sum_j (\alpha\lambda_j(X) - \mu \ln \lambda_j(X)) \\
&= \alpha \sum_j \lambda_j(X) - \mu \ln \prod_j \lambda_j(X) \\
&= \alpha I \bullet X + \mu f(X) \\
&\leq \hat{S} \bullet \hat{X} + \mu f(\hat{X}).
\end{aligned}$$

It follows that $\phi(\lambda_j(X))$ is bounded above on $\mathcal{F}$, and thus for each eigenvalue $\lambda_j(X)$, we have

$$\underline{\lambda} \leq \lambda_j(X) \leq \overline{\lambda}$$

for some positive $\underline{\lambda}$ and some finite $\overline{\lambda}$. The set $\mathcal{F}$ is closed since it is defined by linear equations and an inequality on a function that is continuous on such sets.

The objective function $\hat{S}\bullet X+\mu f(X)$ is continuous on $\mathcal{F}$ since $\mathcal{F}$ contains only positive definite matrices. Hence the Weierstrass theorem applies and the existence follows.

□

We have thus established that the central path is well defined. The following theorem states that the central path in fact converges to optimal solutions of $(\mathcal{P})$ and $(\mathcal{D})$.

Theorem 1.6 *The sequence $X(\mu)$ (resp. $S(\mu)$) always converges to an optimal solution X^* (resp. S^*) of the SDP problem (resp. its dual) as $\mu \downarrow 0$.*

Proof

See Theorem A.3 in [10]. □

If we can find solution pairs $(X(\mu), S(\mu))$ for some decreasing sequence of μ that approaches zero, we will arrive at an optimal solution in the end. That is the idea behind the path-following algorithms. Since the last equation of (CP_μ) is nonlinear, it is expensive to find an exact solution. So the path-following algorithms try to find approximate solutions to (CP_μ).

To measure the quality of approximation, a neighborhood of the central path is defined. The neighborhood is defined either in terms of norms of $(X^{1/2}SX^{1/2} - \mu I)$ or in terms of the eigenvalues of this matrix. See [1] for a detailed discussion.

1.4 Search Directions

At each iteration, the pair $(X(\mu), S(\mu))$ on the central path can be approximated using Newton's method. However, to ensure that the iterates remain strictly feasible, damped Newton steps are usually taken. The Newton step, or solution (dX, dy, dS) of the linearization of (CP_μ), is called a *search direction*.

A straightforward linearization of (CP_μ) yields an overdetermined linear system with $(n(n+1)+m)$ unknowns and $\left(n(n+1)+m+\frac{n(n-1)}{2}\right)$ equations, we usually symmetrize $XS = \mu I$ to reduce the number of equations. The most natural symmetrization is to replace $XS = \mu I$ with $\frac{1}{2}(XS + SX) = \mu I$. Linearizing it gives the following equation:

$$\frac{1}{2}((dX)S + S(dX) + X(dS) + (dS)X) = \mu I - \frac{1}{2}(XS + SX).$$

The resulting search direction is called the Alizadeh-Haeberly-Overton (AHO) direction because it was first introduced and analyzed by Alizadeh, Haeberly and Overton [11].

The second search direction we introduce here is the HRVW/KSH/M direction. This direction was discovered independently by Helmberg, Rendl, Vanderbei, and Wolkowicz [12] and Kojima, Shindoh, and Hara [13]. Later, Monteiro [14] gave another derivation of this direction. The motivation of this direction is that if we view the left-hand side of equations (CP_μ) as a map from $\mathbb{R}^{n\times n} \times \mathbb{R}^m \times \mathbb{S}^n$ into itself, then a direction $(\widetilde{dX}, dy, dS)$ in $\mathbb{R}^{n\times n} \times \mathbb{R}^m \times \mathbb{S}^n$ can be defined. However, we require that all the iterates be in $\mathbb{S}^n \times \mathbb{R}^m \times \mathbb{S}^n$, so we project $\widetilde{dX}$ onto $\mathbb{S}^n$; that is, we take the search direction (dX, dy, dS) with $dX = (\widetilde{dX} + \widetilde{dX}^T)/2$. More specifically, the HRVW/KSH/M direction can be obtained as the solution (dX, dy, dS) to the following system:

$$\begin{aligned} A^{(i)} \bullet dX &= 0, \ (i = 1, \dots, m), \quad dX \in \mathbb{S}^n, \\ \sum_{i=1}^m (dy)_i\, A^{(i)} + dS &= 0, \quad dS \in \mathbb{S}^n, \\ (\widetilde{dX})S + X(dS) &= \mu I - XS, \quad \widetilde{dX} \in \mathbb{R}^{n\times n}, \quad dX = (\widetilde{dX} + \widetilde{dX}^T)/2. \end{aligned}$$

Kojima et al. in [13] also described the dual counterpart of the HRVW/KSH/M direction. The dual direction results from interchanging X and S and correspondingly dX and dS in the third equation above. Namely, it is the solution (dX, dy, dS) to the system

$$\begin{aligned} A^{(i)} \bullet dX &= 0, \ (i = 1, \dots, m), \quad dX \in \mathbb{S}^n, \\ \sum_{i=1}^m (dy)_i\, A^{(i)} + dS &= 0, \quad dS \in \mathbb{S}^n, \\ (dX)S + X(\widetilde{dS}) &= \mu I - XS, \quad \widetilde{dS} \in \mathbb{R}^{n\times n}, \quad dS = (\widetilde{dS} + \widetilde{dS}^T)/2. \end{aligned}$$

All directions described above can be viewed as members of the Monteiro-Zhang family. This family was first introduced by Monteiro in order to give another derivation of the HRVM/KSH/M search direction. Later Zhang generalized Monteiro's work, and the resulting set of directions is called the Monteiro-Zhang family. These directions are the Newton steps for the central path equation (CP_μ) with the last equation replaced by

$$H_P(XS) := [PXSP^{-1} + (PXSP^{-1})^T]/2 = \mu I,$$

where $P \in \mathbb{S}^n_{++}$ is arbitrary but fixed. In other words, the Monteiro-Zhang search direction that corresponds to some positive definite matrix P is the solution (dX, dy, dS) to the system

$$(MZ_P) \qquad \begin{array}{rcl} A^{(i)} \bullet dX &=& 0 \ (i = 1, \ldots, m), \quad dX \in \mathbb{S}^n, \\ \displaystyle\sum_{i=1}^m (dy)_i \, A^{(i)} + dS &=& 0, \quad dS \in \mathbb{S}^n, \\ PX(dS)P^{-1} + P^{-T}(dS)XP^T && \\ +P(dX)SP^{-1} + P^{-T}S(dX)P^T &=& 2\mu I - PXSP^{-1} - P^{-T}SXP^T. \end{array}$$

This system has a unique solution if
(1) $PXSP^{-1} + P^{-T}SXP^T \in \mathbb{S}^n_+$ (see Proposition 2.2 in [21]),
or
(2) $\|S^{1/2}XS^{1/2} - \mu I\|_2 < \mu/\sqrt{2}$ for some $\mu > 0$, where $\|\cdot\|_2$ denotes the operator 2-norm (see Proposition 10.4.2 in [1]).

If we choose $P = I$, then the resulting direction is the AHO direction. When $P = S^{1/2}$ and $P = X^{-1/2}$, it gives the HRVW/KSH/M direction and its dual direction, respectively. If we choose $P = W^{-1/2}$, where $W = X^{1/2}(X^{1/2}SX^{1/2})^{-1/2}X^{1/2} = S^{-1/2}(S^{1/2}XS^{1/2})^{1/2}S^{-1/2}$, then we obtain another widely-used search direction, the Nesterov-Todd (NT) direction; see also [15, 16]. The NT direction (dX, dy, dS) can also be viewed as the solution to the system

$$\begin{array}{rcl} A^{(i)} \bullet dX &=& 0 \ (i = 1, \ldots, m), \\ \displaystyle\sum_{i=1}^m dy_i \, A^{(i)} + dS &=& 0, \\ W^{-1}dXW^{-1} + dS &=& \mu X^{-1} - S. \end{array}$$

The AHO, HRVW/KSH/M, and NT directions are the most commonly used in practice. However, there are several other directions addressed in the literature; see, for example, Tseng [17], Monteiro and Tsuchiya [18]. For long-step path-following algorithms, those based on the NT direction have the lowest possible iteration-complexity bound.

Chapter 2

Group-invariant SDP

Many optimization problems possess symmetries and the SDP formulations inherit them. This chapter describes a representation of these symmetries and their impact on interior-point algorithms.

Let G be a finite group, and let $\sigma : G \mapsto Aut(\mathbb{S}^n_+)$ be a linear representation. An SDP problem is *invariant under* σ if the objective function and the feasible set are invariant under every $\sigma(\pi), \pi \in G$. In most practical instances, the action of $\sigma(\pi)$ on matrix variable X can be described by the linear operator

$$X \mapsto Q(\pi)XQ(\pi)^T,$$

where $Q(\pi)$ is an orthogonal matrix. Throughout this essay, we assume that the linear representation σ takes this form.

The invariance of the objective function implies that for any $\pi \in G$ and any $X \in \mathbb{S}^n$,

$$\begin{aligned} C \bullet X &= C \bullet \sigma(\pi)(X) \\ &= C \bullet Q(\pi)XQ(\pi)^T \\ &= Q(\pi)^T C Q(\pi) \bullet X. \end{aligned}$$

It is easy to see that this equality holds for every $X \in \mathbb{S}^n$ and every $\pi \in G$ if and only if

$$C = Q(\pi)^T C Q(\pi) \quad \text{for all } \pi \in G;$$

i.e., C is invariant under σ.

In the subsequent sections, we discuss the effect of group-invariance on interior-point algorithms.

2.1 Invariance of Search Directions

In this section, we show that if both $(\mathcal{P})$ and $(\mathcal{D})$ are invariant under σ, and the current iterates X, S are invariant under σ, then (dX, dS) is also invariant under σ for some of the search directions discussed in the previous chapter.

Let the null space of $\mathcal{A}$ be denoted by $\mathcal{L}$, and the orthogonal complement of $\mathcal{L}$ by $\mathcal{L}^{\perp}$.

Clearly the constraint $\mathcal{A}(X) = b$ is equivalent to $X \in (\mathcal{L} + \hat{X})$, where $\hat{X}$ is any feasible solution to $(\mathcal{P})$. Suppose the feasible set of $(\mathcal{P})$ is invariant under σ; i.e.,

$$\sigma(\pi)(\mathcal{L} + \hat{X}) = \mathcal{L} + \hat{X} \text{ for all } \pi \in G.$$

Similarly, the constraint $S = C - \sum_{i=1}^{m} y_i A^{(i)}$ is equivalent to

$$S \in (\mathcal{R}(\mathcal{A}^*) + \hat{S}) = (\mathcal{L}^{\perp} + \hat{S}),$$

where $\hat{S}$ is any feasible solution to $(\mathcal{D})$, and $\mathcal{A}^*$ denotes the adjoint of $\mathcal{A}$. Suppose the feasible set of $(\mathcal{D})$ is invariant under σ; i.e.,

$$\sigma(\pi)(\mathcal{L}^{\perp} + \hat{S}) = \mathcal{L}^{\perp} + \hat{S} \text{ for all } \pi \in G.$$

Proposition 2.1 *If $dX \in \mathcal{L}$, then $\sigma(\pi)(dX) \in \mathcal{L}$ for all $\pi \in G$. Similarly, if $dS \in \mathcal{L}^{\perp}$, then $\sigma(\pi)(dS) \in \mathcal{L}^{\perp}$ for all $\pi \in G$.*

Proof
Suppose $dX \in \mathcal{L}$ and $\pi \in G$.

Applying the group action $\sigma(\pi)$ to both sides of $dX \in \mathcal{L}$ gives

$$\sigma(\pi)(dX) \in \sigma(\pi)(\mathcal{L}).$$

Let $\bar{X}$ be a feasible solution of $(\mathcal{P})$. Then the *group average*

$$\hat{X} = \frac{1}{\mid G \mid} \sum_{\pi \in G} \sigma(\pi)(\bar{X})$$

is also feasible for $(\mathcal{P})$. Moreover, $\hat{X}$ is invariant under σ. Therefore

$$\mathcal{L} + \hat{X} = \sigma(\pi)(\mathcal{L} + \hat{X}) = \sigma(\pi)(\mathcal{L}) + \sigma(\pi)(\hat{X}) = \sigma(\pi)(\mathcal{L}) + \hat{X}$$

for all $\pi \in G$, where the first equality follows from the invariance of the feasible set, and the last equality follows from the invariance of $\hat{X}$. So

$$\mathcal{L} = \sigma(\pi)(\mathcal{L}).$$

Thus, we have

$$\sigma(\pi)(dX) \in \sigma(\pi)(\mathcal{L}) = \mathcal{L}.$$

A similar argument shows that

$$\sigma(\pi)(dS) \in \mathcal{L}^{\perp}$$

for all $dS \in \mathcal{L}^{\perp}$ and all $\pi \in G$.

□

Theorem 2.2 *Suppose both $(\mathcal{P})$ and $(\mathcal{D})$ are invariant under σ, and the current iterates X, S are invariant under σ. Then the matrices dX and dS in the Monteiro-Zhang search direction (dX, dy, dS) that corresponds to P are invariant under σ whenever the search direction is well-defined and P is invariant.*

Proof
Recall that $\sigma(\pi) : U \mapsto Q(\pi)UQ(\pi)^T$ for each $\pi \in G$, where $Q(\pi)$ is an orthogonal matrix. We fix $\pi \in G$ and abbreviate $Q(\pi)$ as Q below.

Since both $(\mathcal{P})$ and $(\mathcal{D})$ are invariant, we know from the preceding proposition that $Q(dX)Q^T \in \mathcal{L}$ and $Q(dS)Q^T \in \mathcal{L}^\perp$ for any $dX \in \mathcal{L}$ and any $dS \in \mathcal{L}^\perp$; i.e., $(Q(dX)Q^T, \overline{dy}, Q(dS)Q^T)$ solves the first two equations of (MZ_P) (see Section 1.4) for some $\overline{dy} \in \mathbb{R}^m$.

The third equation of (MZ_P) can be written as

$$PX(dS)P^{-1} + P^{-T}(dS)XP^T + P(dX)SP^{-1} + P^{-T}S(dX)P^T$$
$$= 2\mu I - PXSP^{-1} - P^{-T}SXP^T.$$

Applying $\sigma(\pi)$ to both sides gives

$$QPX(dS)P^{-1}Q^T + QP^{-T}(dS)XP^TQ^T + QP(dX)SP^{-1}QT + QP^{-T}S(dX)P^TQ^T$$
$$= 2\mu QIQ^T - QPXSP^{-1}Q^T - QP^{-T}SXP^TQ^T.$$

Since Q is orthogonal, the above equation is equivalent to

$$QPQ^TQXQ^TQ(dS)Q^TQP^{-1}Q^T + QP^{-T}Q^TQ(dS)Q^TQXQ^TQP^TQ^T$$
$$+QPQ^TQ(dX)Q^TQSQ^TQP^{-1}QT + QP^{-T}Q^TQSQ^TQ(dX)Q^TQP^TQ^T$$
$$= 2\mu I - QPQ^TQXQ^TQSQ^TQP^{-1}Q^T - QP^{-T}Q^TQSQ^TQXQ^TQP^TQ^T.$$

Since the current iterates X, S are invariant under σ, we have $QXQ^T = X$ and $QSQ^T = S$. Suppose P is also invariant under σ. Then $P = QPQ^T$, and thus

$$P^{-1} = Q^{-T}P^{-1}Q^{-1} = QP^{-1}Q^T.$$

Therefore, we have that $(Q(dX)Q^T, \overline{dy}, Q(dS)Q^T)$ also satisfies the third equation of (MZ_P). Consequently, since the search direction is well-defined,

$$Q(dX)Q^T = dX \quad \text{and} \quad Q(dS)Q^T = dS.$$

□

Suppose the assumptions of Theorem 2.2 are satisfied. Then we know the AHO direction is invariant since the matrix I satisfies $I = QIQ^T$ for any orthogonal matrix Q.

Also, since X is invariant under σ, we have

$$X^{-1} = (QXQ^T)^{-1} = QX^{-1}Q^T$$

and

$$(QX^{-1/2}Q^T)^2 = QX^{-1/2}Q^TQX^{-1/2}Q^T = X^{-1} = (X^{-1/2})^2$$

for any $Q = Q(\pi)$, $\pi \in G$. Thus $X^{-1/2}$ is invariant under σ. Similarly, $S^{1/2}$ is invariant under σ. Hence the HRVW/KSH/M direction and its dual direction are invariant as well.

For the NT direction to be invariant, we require the matrix

$$W = X^{1/2}(X^{1/2}SX^{1/2})^{-1/2}X^{1/2}$$

to be invariant under σ. For this, we only need the above observations and the fact that if $Y, Z \in \mathbb{S}^n$ are invariant under σ, then

$$QYZYQ^T = (QYQ^T)(QZQ^T)(QYQ^T) = YZY$$

for any $Q = Q(\pi), \pi \in G$, which shows that YZY is invariant under σ.

As a direct consequence of Theorem 2.2, we have the following corollary.

Corollary 2.3 *Suppose $(\mathcal{P})$ and $(\mathcal{D})$ are invariant under σ, and the feasible solutions X^0 and S^0 are invariant under σ. If we use an interior-point algorithm based on the AHO, HRVM/KSH/M or NT search direction, and start from (X^0, y^0, S^0), then all iterates generated have X and S invariant under σ.*

2.2 Reducing the Size of Invariant SDP

If an SDP problem is invariant under a group representable by permutations, then we have a method to reduce the number of unknowns in the underlying matrix so that we can solve that problem much more efficiently. This method, proposed by E. de Klerk et al. [5], is based on constructing a 'regular $*$-representation' of a matrix $*$-algebra. We introduce it briefly in this section.

This method applies to SDP problems in the standard form $(\mathcal{P})$, and it is particularly effective when the matrices $C, A^{(1)}, \ldots, A^{(m)}$ are invariant under a large group G of symmetric permutations of rows and columns.

Let G be a subgroup of permutations on $\{1, \ldots, n\}$, and for each $\pi \in G$ define

$$(M_\pi)_{i,j} = \begin{cases} 1 & \text{if } \pi(i) = j; \\ 0 & \text{otherwise.} \end{cases}$$

Note that for any $\pi, \pi^{'} \in G$,

$$M_\pi M_{\pi^{'}} = M_{\pi\pi^{'}} \text{ and } M_\pi^T = M_{\pi^{-1}}.$$

Suppose that $C, A^{(1)}, \ldots, A^{(m)}$ are invariant under $X \mapsto M_\pi X M_\pi^T$ for all $\pi \in G$. Since the group average $X^{'} = \frac{1}{|G|} \sum_{\pi \in G} M_\pi(\bar{X}) M_\pi^T$ of an optimal solution $\bar{X}$ is also optimal, we can restrict the problem to the invariant subspace $\{X : M_\pi(\bar{X}) M_\pi^T = X, \ \forall \pi \in G\}$.

Let C be the linear subspace spanned by $\{M_\pi : \pi \in G\}$; i.e.,

$$\mathcal{C} = \left\{ \sum_\pi \lambda_\pi M_\pi : \lambda_\pi \in \mathbb{R}, \pi \in G \right\}.$$

Observe that $\mathcal{C}$ is a *matrix $*$-algebra*; i.e., it is a collection of matrices closed under addition, scalar and matrix multiplication, and transposition. The *commutant* $\mathcal{C}^{'}$ of $\mathcal{C}$ is

$$\mathcal{C}^{'} := \{X \in \mathbb{R}^{n \times n} | XM = MX \text{ for all } M \in \mathcal{C}\}.$$

The commutant is also a matrix $*$-algebra. Note that $X \in \mathbb{S}^n$ is invariant under $\sigma : \pi \in G \mapsto M_\pi \cdot M_\pi^T$ if and only if $X \in \mathcal{C}'$.

The matrix $*$-algebra $\mathcal{C}'$ has a basis of {0,1}-matrices $E_1, ..., E_d$ such that

$$E_1 + \cdots + E_d = J,$$

where J is the all-one matrix. See [19] for the basics of the matrix $*$-algebra. For each i, let B_i be the normalization of E_i ; i.e.,

$$B_i = \text{trace}(E_i^T E_i)^{-\frac{1}{2}} E_i.$$

Then $\{B_i : i = 1, \ldots, d\}$ is an orthonormal basis of $\mathcal{C}'$. For $k = 1, \ldots, d$, let L_k be the $d \times d$ matrix defined by

$$(L_k)_{i,j} = \text{trace}(B_i B_k B_j).$$

Let $\mathcal{K}$ be the linear subspace spanned by $\{L_k : k = 1, \ldots, d\}$. We can prove that the linear operator $\phi : \mathcal{C}' \mapsto \mathcal{K}$ defined by $\phi(B_k) = L_k$ for $k = 1, \ldots, d$ is a $*$-isomorphism; i.e., for any $Y, Z \in \mathcal{C}'$, we have

$$\phi(YZ) = \phi(Y)\phi(Z) \quad \text{and} \quad \phi(Y^T) = \phi(Y)^T.$$

Theorem 2.4 *ϕ is a $*$-isomorphism.*

Proof

See Theorem 1 of [5]. □

A consequence of Theorem 2.4 is that, for any $x_1, \ldots, x_d \in \mathbb{R}$,

$$\sum_{k=1}^{d} x_i B_i \in \mathbb{S}_+^n \quad \text{if and only if} \quad \sum_{k=1}^{d} x_i L_i \in \mathbb{S}_+^d.$$

We can see this result as a well-known fact from matrix $*$-algebra or view it as follows. On the one hand, since ϕ is a $*$-isomorphism, ϕ maintains symmetry of matrices. On the other hand, let $M \in \mathcal{C}'$ be symmetric and let $p(x)$ be the minimal polynomial of M. Because ϕ is an algebra $*$-isomorphism, p is also the minimal polynomial of $\phi(M)$. Therefore, $M \in \mathbb{S}^n_+$ if and only if all roots of p are nonnegative, which is in turn equivalent to $\phi(M) \in \mathbb{S}^n_+$.

Since the order d of the matrices L_i is equal to the number of matrices B_i, we can reduce the number of variables in $(\mathcal{P})$ from $\frac{n(n+1)}{2}$ to the dimension of the subspace of $\mathcal{C}'$ of symmetric matrices.

This reduction technique was successfully applied by E. de Klerk et al. to efficiently compute an SDP lower bound of the crossing number of complete bipartite graphs, which will be introduced in Chapter 4. Another potential application is the truss topology design problem, which will be discussed in Chapter 3.

Chapter 3

Application to Optimal Design of a Symmetric Truss

3.1 Topology Optimization Problem of Trusses

The topology of a truss means an assemblage of nodes connected by members. The optimization problem considered here is to minimize the total structural volume of the trusses under free vibration frequency constraints.

It is well-known in physics that free vibratory systems without damping can be described by differential equations. The equation of motion for such a system can be written in the form of

$$[m]\ddot{u} + [k]u = 0$$

where u is the vector of displacements, $[m]$ is the mass matrix with the mass values at the diagonal entries, and $[k]$ is the stiffness matrix. Upon solving this ordinary differential equation, we obtain general solutions of the form

$$u = \sum_i [C_i \sin(\lambda_i t) + D_i \cos(\lambda_i t)],$$

where λ_i^2 are generalized eigenvalues satisfying

$$[k]u_i = \lambda_i^2 [m]u_i,$$

where u_i is an associated eigenvector. See, for example, [20]. Thus we can use eigenvalue analysis in the study of such problems.

In fact, the eigenvalues of free vibration are an important performance measure of the structures. It is well-known that optimal designs for specified fundamental eigenvalues often have repeated eigenvalues, which gives rise to the difficulties in optimizing such problems. Since the repeated eigenvalues are not differentiable in normal cases, and we can only calculate the directional derivatives with respect to the design variables. No globally convergent algorithm for structural optimization problems with large multiplicity of eigenvalues were proposed before the algorithm presented in [3]. By formulating the topology optimization problem (TOP) of trusses for specified eigenvalues of vibration as an SDP problem, we can make use of some path-following algorithms for semidefinite programming to compute the optimal solution efficiently and accurately.

3.2 SDP Formulation

Consider a truss with fixed locations of nodes and members. Suppose the number of degrees of freedom for displacements is n, and the number of members of a truss is m. Let $y = (y_1, \cdots, y_m) \in \mathbb{R}^m$ denote the vector of member cross-sectional areas. The stiffness matrix and the structural mass matrix are denoted by $K \in \mathbb{S}^n$ and $M_s \in \mathbb{S}^n$, respectively, both of which are linear functions of y. The nonstructural mass matrix is denoted by $M_0 \in \mathbb{S}^n$.

The eigenvalue problem of vibration is formulated as

$$K\Phi_r = \Omega_r (M_s + M_0)\Phi_r \quad \text{for} \quad r = 1, \cdots, n$$

where Ω_r and Φ_r are the rth smallest eigenvalue and associated eigenvector, respectively.

Let $\bar{\Omega}$ denote the lower bound of the eigenvalues. The vector of member length is denoted by $b \in \mathbb{R}^m$. Then the TOP for specified fundamental eigenvalues is formulated as

$$
(TOP) \qquad \begin{array}{lrcl} \min & \displaystyle\sum_{i=1}^{m} b_i y_i & & \\ \text{s.t.} & \Omega_r & \geq & \bar{\Omega}, \ r = 1, \ldots, n, \\ & y_i & \geq & 0, \quad i = 1, \ldots, m. \end{array}
$$

Note that in practice a small positive lower bound $\bar{y}_i$ is often given for y_i to prevent instability of the structure. So we reformulate the TOP as

$$
(TOP') \qquad \begin{array}{lrcl} \min & \displaystyle\sum_{i=1}^{m} b_i y_i & & \\ \text{s.t.} & \Omega_r & \geq & \bar{\Omega}, \quad r = 1, \ldots, n, \\ & y_i & \geq & \bar{y}_i, \quad i = 1, \ldots, m. \end{array}
$$

If the eigenvalue of the optimal design is simple, we can calculate the derivative of Ω_1 with respect to y_i easily and use nonlinear programming techniques to solve it. However, if the eigenvalue is of large multiplicity, only directional derivatives can be calculated. Fortunately, some path-following algorithms for SDP do not need the derivative but only directional derivatives. Now the question is how to formulate (TOP') as an SDP problem.

Consider a structure with $\Omega_1 \geq \bar{\Omega}$. By Rayleigh's principle, it can be shown that

$$
\psi^T [K - \bar{\Omega}(M_s + M_0)]\psi \geq 0
$$

for any ψ; see [4] for more details. This implies that $[K - \bar{\Omega}(M_s + M_0)] \in \mathbb{S}^n_+$, so it is possible to formulate it as a constraint in an SDP problem.

Define K_i and M_i as

$$
K_i = \frac{\partial K}{\partial y_i}, \qquad M_i = \frac{\partial M_s}{\partial y_i} \qquad \text{for} \ \ i = 1, \ldots, m.
$$

Note that $K = \sum_{i=1}^{m} y_i K_i$ and $M_s = \sum_{i=1}^{m} y_i M_i$, because K and M_s are linear functions of y_i. Therefore, (TOP') can be reduced to the SDP problem in dual form:

$$(\mathcal{D}') \qquad \begin{array}{rrcl} \max & -\,b^T y & & \\ \text{s.t.} & -\displaystyle\sum_{i=1}^{m}(K_i - \bar{\Omega} M_i) y_i + S & = & -\bar{\Omega} M_0, \\ & S & \in & \mathbb{S}^n_+, \\ & y & \geq & \bar{y}. \end{array}$$

The corresponding primal problem $(\mathcal{P}')$ is

$$(\mathcal{P}') \qquad \begin{array}{rrcl} \min & -\,\bar{\Omega} M_0 \bullet X - \bar{y}^T \eta & & \\ \text{s.t.} & (K_i - \bar{\Omega} M_i) \bullet X + \eta_i & = & b_i, \ i = 1, \ldots, m, \\ & X & \in & \mathbb{S}^n_+, \\ & \eta & \geq & 0. \end{array}$$

3.3 Exploiting Symmetry

From the practical point of view, a symmetric design of a truss is usually expected. A *symmetric design* for TOP means that the symmetrically located members have the same cross-sectional areas. For a symmetric truss, there exists a transformation group under which the geometry and the mechanical properties of the truss are invariant.

Let G be a finite group such that the geometry, the stiffness and mass distributions and support conditions are all invariant under the action of any $\pi \in G$. More specifically, it has the following properties: the locations of nodes and members are symmetric; the locations and the values of the non-structural masses are symmetric; support conditions are symmetric; and the symmetrically located members have the same material and cross-sectional areas.

Mathematically, such group G can be described as follows. For each $\pi \in G$, let $h_\pi : \{1, \ldots, m\} \mapsto \{1, \ldots, m\}$ denote the permutation of members under the action of π. Define $D(\pi) \in \mathbb{R}^{m\times m}$ as the permutation matrix satisfying

$$D_{i,j}(\pi) = \begin{cases} 1 & \text{if } h_\pi(i) = j; \\ 0 & \text{otherwise.} \end{cases}$$

So the action of $\pi \in G$ on the vector member lengths b can be written as

$$\tilde{b}(\pi) = D(\pi)^T b.$$

The invariance of members under the action of π can thus be expressed as

$$b = \tilde{b} = D(\pi)^T b. \tag{3.1}$$

Similarly, the invariance of the lower bounds on the cross-sectional areas under the action of π means

$$\bar{y} = D(\pi)^T \bar{y}. \tag{3.2}$$

Now consider the action of $\pi \in G$ on nodal transformation. Suppose the vector of the nodal displacements u of the truss corresponding to a mode of vibration is transformed to $\tilde{u}$ by $P(\pi)$, namely,

$$\tilde{u} = P(\pi)^T u.$$

Here, $P(\pi) \in \mathbb{R}^{n\times n}$ is a matrix of the nodal displacements, and it can be obtained as the product of an appropriate linear transformation matrix and permutation matrix for assignment of displacement numbers. Note that $P(\pi)$ is an orthogonal matrix. For any $\pi \in G$, the action of π on K, M_s and M_0 are thus given by

$$\tilde{K}(\pi) = P(\pi)^T K P(\pi),$$
$$\tilde{M}_s(\pi) = P(\pi)^T M_s P(\pi),$$
$$\tilde{M}_0(\pi) = P(\pi)^T M_0 P(\pi).$$

We have

$$\tilde{M}_0(\pi) = M_0 \tag{3.3}$$

for each $\pi \in G$, since the locations and the values of nonstructural masses are invariant under $\pi \in G$. Also,

$$\frac{\partial \tilde{K}(\pi)}{\partial y_{h_\pi(i)}} = \frac{\partial K}{\partial y_i} \quad \text{and} \quad \frac{\partial \tilde{M}_s(\pi)}{\partial y_{h_\pi(i)}} = \frac{\partial M_s}{\partial y_i} \quad \text{for} \quad i = 1, \ldots, m$$

for each $\pi \in G$, because the stiffness and mass distributions are also invariant under $\pi \in G$. Therefore, for each $\pi \in G$,

$$K_i = \frac{\partial}{\partial y_{h_\pi(i)}} \left(P(\pi)^T K P(\pi) \right) = P(\pi)^T K_{h_\pi(i)} P(\pi) \tag{3.4}$$

and

$$M_i = \frac{\partial}{\partial y_{h_\pi(i)}} \left(P(\pi)^T M_s P(\pi) \right) = P(\pi)^T M_{h_\pi(i)} P(\pi) \tag{3.5}$$

for $i = 1, \ldots, m$.

We now show that $(\mathcal{P}')$ and $(\mathcal{D}')$ are both invariant SDP problems under $\sigma : G \mapsto Aut(\mathbb{S}^n_+ \oplus \mathbb{R}^m_+)$ defined by

$$\sigma(\pi) : (x, \eta) \mapsto \left(P(\pi) X P(\pi)^T, D(\pi)^T \eta \right)$$

for each $\pi \in G$. First consider the objective function of $(\mathcal{P}')$. Applying $\sigma(\pi), \pi \in G$ to the pair (x, η), the objective function becomes

$$\begin{aligned} -\bar{\Omega} M_0 \bullet P(\pi) X P(\pi)^T - \bar{y}^T D(\pi)^T \eta &= -\bar{\Omega} P(\pi)^T M_0 P(\pi) \bullet X - (D(\pi)\bar{y})^T \eta \\ &= -\bar{\Omega} M_0 \bullet X - \bar{y}^T \eta, \end{aligned}$$

where the last equality follows from (3.2) and (3.3). Hence, the objective function is invariant under σ.

Now consider the feasible set of $(\mathcal{P}')$. For each $\pi \in G$, applying (3.1), (3.4) and (3.5) to the constraints

$$(K_i - \bar{\Omega} M_i) \bullet X + \eta_i = b_i, \quad i = 1, \ldots, m,$$

gives

$$\left(P(\pi)^T K_{h_\pi(i)} P(\pi) - \bar{\Omega} P(\pi)^T M_{h_\pi(i)} P(\pi) \right) \bullet X + \eta_i = b_{h_\pi(i)}, \quad i = 1, \ldots, m,$$

or equivalently,

$$(K_{h_\pi(i)} - \bar{\Omega} M_{h_\pi(i)}) \bullet P(\pi) X P(\pi)^T + (D(\pi)^T \eta)_{h_\pi(i)} = b_{h_\pi(i)}, \quad i = 1, \ldots, m.$$

Thus, (X, η) is feasible for $(\mathcal{P}^{'})$ if and only if $\sigma(\pi)(X, \eta)$ is feasible for each $\pi \in G$; i.e., the feasible set of $(\mathcal{P}^{'})$ is invariant under σ.

We can show that $(\mathcal{D}^{'})$ is also invariant under σ using the same argument.

We thus established that $(\mathcal{P}^{'})$ and $(\mathcal{D}^{'})$ are invariant SDPs. If we start from an initial invariant solution (X^0, y^0, S^0), and use the search directions described in Theorem 2.3, we will obtain an invariant solution in the end, which is exactly the desired design.

Chapter 4

Application to the Crossing Number of Graphs

4.1 The Crossing Number of Complete Bipartite Graphs

The *crossing number* of a graph G is the minimum number of intersections of edges among drawings of G in the plane. We use $cr(G)$ to denote the crossing number of the graph G.

Zarankiewicz's crossing-number conjecture states that for the complete bipartite graph $K_{m,n}$,

$$cr(K_{m,n}) = Z(m,n),$$

where $Z(m,n) = \lfloor \frac{1}{4}(m-1)^2 \rfloor \lfloor \frac{1}{4}(n-1)^2 \rfloor$.

We can construct drawings of $K_{m,n}$ with exactly $Z(m,n)$ crossings by using the following strategy: arrange the m and n vertices along the $x-$ and $y-$axes, respectively, in each case with as close to half of them on either side of the origin as possible, and join them by mn straight-line segments. Then the number of crossings is

$$
\begin{aligned}
& \binom{\lfloor \frac{n}{2} \rfloor}{2}\binom{\lfloor \frac{m}{2} \rfloor}{2} + \binom{\lfloor \frac{n}{2} \rfloor}{2}\binom{\lceil \frac{m}{2} \rceil}{2} + \binom{\lceil \frac{n}{2} \rceil}{2}\binom{\lfloor \frac{m}{2} \rfloor}{2} + \binom{\lceil \frac{n}{2} \rceil}{2}\binom{\lceil \frac{m}{2} \rceil}{2} \\
= \; & \left\lfloor \frac{n}{2} \right\rfloor \left\lfloor \frac{n-1}{2} \right\rfloor \left\lfloor \frac{m}{2} \right\rfloor \left\lfloor \frac{m-1}{2} \right\rfloor .
\end{aligned}
$$

This shows that $Z(m,n)$ is an upper bound for $cr(K_{m,n})$. We want to find a reasonably good lower bound for this problem. It turns out that the quality of the lower bound can be improved greatly by using SDP.

4.2 SDP Relaxation

For any m, n, suppose $K_{m,n}$ has bipartition {1,2,...,m} and $\{u_1, ..., u_n\}$. For a fixed planar drawing of $K_{m,n}$, we use $\gamma(u_i)$ to denote the cyclic permutation $(i_1, i_2, ..., i_m)$ such that the edges incident with u_i leave u_i and go to $(i_1, i_2, ..., i_m)$ in the clockwise order.

Let $\mathbb{Z}_m$ denote the set of all cyclic permutations of {1,2,...,m}. Note that $\mid \mathbb{Z}_m \mid = m!/m = (m-1)!$. For any $\sigma, \tau \in \mathbb{Z}_m$, let $C_{\sigma,\tau}$ denote the minimum number of crossings when drawing $K_{m,2}$ such that $\gamma(u_1) = \sigma$ and $\gamma(u_2) = \tau$. Then $C_{\sigma,\tau}$ define a matrix $C = (C_{\sigma,\tau})$ in $\mathbb{S}^{\mathbb{Z}_m}$. It was shown in [7] that $C_{\sigma,\sigma} = \lfloor \frac{1}{4}(m-1)^2 \rfloor$. Moreover, other entries in C can be efficiently computed.

Consider a fixed drawing $\mathcal{W}$ of $K_{m,n}$ which has $cr(K_{m,n})$ crossings. For each $\sigma \in \mathbb{Z}_m$, let

$$
x_\sigma = \frac{1}{n} \left| \{ u_i \in (u_1, u_2, ..., u_n) | \, \gamma(u_i) = \sigma \} \right|,
$$

that is, nx_σ is the number of vertices u_i in $\mathcal{W}$ with $\gamma(u_i) = \sigma$. Consider x as the column vector in $\mathbb{R}^{\mathbb{Z}_m}$. Then $e^T x = 1$, where e is the all-one vector. Let $\beta_{i,j}$ denote the number of crossings in $\mathcal{W}$ that involve an edge incident with u_i and an edge incident with u_j. Clearly, $\beta_{i,j} \geq C_{\gamma(u_i),\gamma(u_j)}$ if $i \neq j$. So we have

$$\begin{aligned} (nx)^T C(nx) &= \sum_{i,j=1}^{n} C_{\gamma(u_i),\gamma(u_j)} \\ &\le \sum_{i,j=1, i\neq j}^{n} \beta_{i,j} + \sum_{i=1}^{n} C_{\gamma(u_i),\gamma(u_i)} \\ &= 2cr(K_{m,n}) + n\lfloor \frac{1}{4}(m-1)^2 \rfloor. \end{aligned}$$

This implies that

$$cr(K_{m,n}) \ge \frac{1}{2} n^2 x^T C x - \frac{1}{2} n \lfloor \frac{1}{4}(m-1)^2 \rfloor. \tag{4.1}$$

If we let $\hat{X} := xx^T$, then $e^T x = 1$ is equivalent to $J \bullet \hat{X} = e^T \hat{X} e = (e^T x)(x^T e) = 1$, where J is the all-one matrix in $\mathbb{S}^{\mathbb{Z}_m}$.

Now consider the following SDP problem

$$(P_{m,n}) \qquad \begin{array}{lrcl} \inf & C \bullet X & & \\ \text{s.t.} & J \bullet X & = & 1, \\ & X & \in & \mathbb{S}^n_+. \end{array}$$

Suppose the infimum is attained, and let α_m be the optimal value of $(P_{m,n})$. Since $\hat{X}$ satisfies all the constraints of $(P_{m,n})$, we have $\alpha_m \le C \bullet \hat{X} = x^T C x$.

It follows directly from (4.1) that

Theorem 4.1 *$cr(K_{m,n}) \ge \frac{1}{2} n^2 \alpha_m - \frac{1}{2} n \left\lfloor \frac{1}{4}(m-1)^2 \right\rfloor$ for any m,n.*

As a consequence of Theorem 4.1, we have the following corollary.

Corollary 4.2 *$cr(K_{m,n}) \ge \frac{m(m-1)}{k(k-1)} \left(\frac{1}{2} n^2 \alpha_k - \frac{1}{2} n \lfloor \frac{1}{4}(k-1)^2 \rfloor \right)$ for all n and $k \le m$.*

Proof

Consider an optimal drawing $\mathcal{W}$ of $K_{m,n}$. Since each crossing in $\mathcal{W}$ lies in $\binom{m-2}{k-2}$ distinct $K_{k,n} \subseteq K_{m,n}$, and there are $\binom{m}{k}$ distinct $K_{k,n}$'s in total, we have

$$cr(K_{m,n}) \geq \frac{\binom{m}{k} cr(K_{k,n})}{\binom{m-2}{k-2}} = \frac{m(m-1)}{k(k-1)} \left(\frac{1}{2} n^2 \alpha_k - \frac{1}{2} n \left\lfloor \frac{1}{4}(k-1)^2 \right\rfloor \right).$$

□

Corollary 4.3 $\lim_{n\to\infty} \frac{cr(K_{m,n})}{Z(m,n)} \geq \frac{8\alpha_k}{k(k-1)} \frac{m}{m-1}$ *for all* $k \leq m$.

Proof

$$\begin{aligned}
\lim_{n\to\infty} \frac{cr(K_{m,n})}{Z(m,n)} &\geq \lim_{n\to\infty} \frac{m(m-1)\left(\frac{1}{2}n^2\alpha_k - \frac{1}{2}n\lfloor\frac{1}{4}(k-1)^2\rfloor\right)}{k(k-1)Z(m,n)} \\
&= \lim_{n\to\infty} \frac{m(m-1)\left(\frac{1}{2}n^2\alpha_k - \frac{1}{2}n\lfloor\frac{1}{4}(k-1)^2\rfloor\right)}{k(k-1)\lfloor\frac{1}{4}(m-1)^2\rfloor\lfloor\frac{1}{4}(n-1)^2\rfloor} \\
&= \frac{2\alpha_k}{k(k-1)} \frac{m(m-1)}{\lfloor\frac{1}{4}(m-1)^2\rfloor} \\
&\geq \frac{8\alpha_k}{k(k-1)} \frac{m}{m-1}.
\end{aligned}$$

□

4.3 Inherent Symmetry

It was shown in [6] that there exists a finite group G with orthogonal representation $\sigma : G \mapsto Aut(\mathbb{S}_+^{\mathbb{Z}_m})$ such that the matrix C is invariant under the action of any element in G; namely,

$$C = \sigma(\pi)C \quad \text{for all} \quad \pi \in G.$$

We now describe the group G.

Let $G := Sym(m) \times Sym(2)$, where $Sym(m)$ denotes the symmetric group of degree m (i.e., the group of all permutations of m objects) and $Sym(2) := \{-1, 1\}$. Here $Sym(m)$ acts as a permutation group by conjugation on the $(m-1)!$ elements of $\mathbb{Z}_m$, while $Sym(2)$ acts on $\mathbb{Z}_m$ by switching $\sigma \in \mathbb{Z}_m$ with $\sigma^{-1} \in \mathbb{Z}_m$. Define $h : G \mapsto Sym(\mathbb{Z}_m)$ by

$$h_{\pi,\, i}(\sigma) := \pi \sigma^i \pi^{-1}$$

for $\pi \in Sym(m), i \in Sym(2), \sigma \in \mathbb{Z}_m$. Therefore, for each $(\pi,\, i) \in G, h_{\pi,\, i}$ is an automorphism on $\mathbb{Z}_m$.

For each $(\pi,\, i) \in G$, define $M_{\pi,\, i} \in \mathbb{R}^{\mathbb{Z}_m \times \mathbb{Z}_m}$ by

$$(M_{\pi,\, i})_{\sigma,\tau} = \begin{cases} 1 & \text{if } h_{\pi,\, i}(\sigma) = \tau; \\ 0 & \text{otherwise,} \end{cases}$$

for $\sigma, \tau \in \mathbb{Z}_m$. Then $M_{\pi,i}$ is the $\mathbb{Z}_m \times \mathbb{Z}_m$ permutation matrix corresponding to the permutation $h_{\pi,\, i}$ of $\mathbb{Z}_m$. Moreover, for all $(\pi, i), (\pi^{'},\, i^{'}) \in G$, we have

$$M_{(\pi,\, i),(\pi',\, i')} = M_{\pi,\, i} M_{\pi',\, i'} \quad \text{and} \quad M_{(\pi,\, i)^{-1}} = M^T_{\pi,\, i}.$$

Hence, $\sigma : G \mapsto Aut(\mathbb{S}_+^{\mathbb{Z}_m})$ defined by $\sigma(\pi, i) : X \mapsto M_{\pi,i} X M^T_{\pi,i}$ is an orthogonal representation of G.

It turns out that the cost matrix C is invariant under the action of group G; i.e.,

$$M_{\pi,\, i} C M^T_{\pi,\, i} = C \quad \text{for each} \quad (\pi,\, i) \in G.$$

A short explanation is that the action of $\sigma(\pi, i)$ on C corresponds to a re-labelling of the nodes $\{1, \ldots, m\}$ to $\{\pi(1), \ldots, \pi(m)\}$, together with a change from clockwise order to counter-clockwise order in the definition of γ in the case $i = -1$. See [6] for more details. Also it is easy to check that

$$M_{\pi,\, i} J M^T_{\pi,\, i} = J \quad \text{for each} \quad (\pi,\, i) \in G.$$

Thus the SDP relaxation ($P_{m,n}$) is invariant under σ. Since the $M_{\pi,i}$'s are permutation matrices, we may apply the reduction technique described in section 2.2. The computational results are given in [5].

It was found that $\alpha_9 \approx 7.73521$, so by Theorem 4.1,

$$cr(K_{9,n}) \geq 3.8676n^2 - 8n,$$

For each $m \geq 9$ and n, we have by Corollary 4.2,

$$cr(K_{m,n}) \geq 0.0537m(m-1)n^2 - \frac{1}{9}m(m-1)n,$$

and for each $m \geq 9$, by Corollary 4.3,

$$\lim_{n \to \infty} \frac{cr(K_{m,n})}{Z(m,n)} \geq 0.8594 \frac{m}{m-1}.$$

The best factor previously known was 0.8303 instead of 0.8594.

Chapter 5

Concluding Remarks

In this essay, we introduced SDP, and the basic theorems of SDP. Then the central path and several of the most popular search directions for the primal-dual interior-point methods were discussed. The motivation of these search directions and the relationship among them were also presented.

The special class of group-invariant SDP problems was investigated. It was proved that if both $(\mathcal{P})$ and $(\mathcal{D})$ are invariant under σ, and the current iterates X, S are invariant under σ, then the Monteiro-Zhang search direction (dX, dy, dS) corresponding to P is also invariant under σ whenever P is invariant. This new result is a contribution of this essay. A method for reducing the size of the underlying matrices was presented. This reduction technique is especially effective for a special class of group-invariant SDP problems where the data matrices are invariant under a large group of symmetric permutations of rows and columns.

The powerful application of invariant SDP was demonstrated through two examples. The first one is an engineering problem–the topology optimization problem. The other is a combinatorial problem–the crossing number of complete bipartite graphs. The TOP was formulated as an SDP problem, which is group-invariant if the data possess symmetries. Using the invariance of popular search directions, we deduced that interior-point algorithms based on these directions always produce symmetric designs.

SDP was used to obtain good lower bounds on the crossing number of complete bipartite graphs. These SDP problems inherit symmetries from the bipartite graphs, which were exploited in the reduction of the sizes of the SDP problems. The numerical results of the reduction technique of section 2.2 applied to the lower bound problem were given. We note that the potential application of this technique to the TOP is worth further study.

Bibliography

[1] H. Wolkowicz, R. Saigal, L. Vandenberghe (Eds.), *Handbook of semidefinite programming*, Kluwer, 2000.

[2] L.Tunçel, *Semidefinite Progamming*, Lecture notes, Department of Combinatorics and Optimization, Faculty of Mathematics, University of Waterloo, Waterloo, Ontario, Canada, 2005.

[3] Y.Kanno, M.Ohsaki, K.Murota and N.Katoh, *Group symmetry in interior-point methods for semidefinite program*, Optimization and Engineering **2** (2001), 293-320.

[4] M.Ohsaki, K.Fujisawa, N.Katoh and Y.Kanno, *Semi-definite programming for topology optimization of trusses under multiple eigenvalue constraints*, Comput. Meth. Appl. Mech. Engng. **180** (1999), 203-217.

[5] E. de Klerk, D.V. Pasechnik, and A. Schrijver, *Reduction of symmetric semidefinite programs using the regular $*$-representation*, Mathematical Programming Series B, to appear, preprint 2005.

[6] E. de Klerk, J. Maharry, D.V. Pasechnik, R.B. Richter, G. Salazar, *Improved bounds for the crossing numbers of $K_{m,n}$ and K_n*, SIAM Journals on Discrete Mathematics, preprint 2004. E-print **math.CO/0404142** at arXiv.org.

[7] D.R. Woodall, *Cyclic-order graphs and Zarankiewicz's crossing-number conjecture*, Journal of Graph Theory **17** (1993), no.6, 657-671.

[8] K. Gatermann, P.A. Parrilo, *Symmetry groups, semidefinite programs, and sums of squares*, Journal of Pure and Applied Algebra **192** (2004), 95-128.

[9] Y.E. Nesterov and A.S. Nemirovski, *Interior point polynomial methods in convex programming*, SIAM Publications. SIAM, Philadelphia, USA, 1994.

[10] M. Halická, E. de Klerk and C.Roos, *On the convergence of the central path in semidefinite optimization*, SIAM Journals on Optimization, **12(4)** (2002), 1090-1099.

[11] F. Alizadeh, J.A. Haeberly and M. Overton, *Primal-dual interior-point methods for semidefinite programming: convergence rates, stability and numerical results*, Technical report, Computer Science Department, New York University, New York, 1996.

[12] C. Helmberg, F. Rendl, R.J. Vanderbei and H. Wolkowicz, *An interior-point method for semidefinite programming*, SIAM Journals on Optimization, **6(2)** (1996), 324-361.

[13] M. Kojima, S. Shindoh and S. Hara, *Interior-point methods for the monotone semidefinite linear complementarity problem in symmetric matrices*, SIAM journals on Optimization, **7(1)** (1997), 86-125.

[14] R.D.C.Monteiro, *Primal-dual path-following algorithms for semidefinite programming*, SIAM journals on Optimization, **7(3)** (1997), 663-678.

[15] Y.E. Nesterov and M.J. Todd, *Self-scaled barriers and interior-point methods for convex programming*, Math. Oper. Res., **22(1)** (1997), 1-42.

[16] Y.E. Nesterov and M.J. Todd, *Primal-dual interior-point methods for self-scaled cones*, SIAM journals on Optimzation, **8** (1998), 324-364.

[17] P. Tseng, *Search directions and convergence analysis of some infeasible path-following methods for the monotone semi-definite LCP*, Optim. Mathods Softw., **9** (1998), 245-268.

[18] R.D.C. Monteiro and T. Tsuchiya, *Polynomial convergence of a new family of primal-dual algorithms for semidefinite programming*, SIAM journals on Optimization, **9** (1999), 551-577.

[19] M. Takesaki, *Theory of operator algebras 1*, Springer, New York, 1979.

[20] C.M. Harris(Eds.), *Shock and vibration handbook*, New York ; Toronto : McGraw-Hill, 1988.

[21] R.D.C. Monteiro and Y. Zhang, *A unified analysis for a class of long-step primal-dual path-following interior-point algorithms for semidefinite programming*, Math. Programming, **81(3, Ser. A)** (1998), 281-299.