ISSN 1342-2804

Research Reports on Mathematical and Computing Sciences

A Numerical Algorithm for Block-Diagonal Decomposition of Matrix $*$-Algebras

Kazuo Murota, Yoshihiro Kanno,
Masakazu Kojima and Sadayoshi Kojima

September 2007, B–445

Department of
Mathematical and
Computing Sciences
Tokyo Institute of Technology

SERIES B: Operations Research

B-445 **A Numerical Algorithm for Block-Diagonal Decomposition of Matrix $*$-Algebras**[ℵ]

Kazuo Murota[⋆], Yoshihiro Kanno[†], Masakazu Kojima[‡], Sadayoshi Kojima[♯]

September 2007

Abstract.
Motivated by recent interest in group-symmetry in the area of semidefinite programming, we propose a numerical method for finding a finest simultaneous block-diagonalization of a finite number of symmetric matrices, or equivalently the irreducible decomposition of the matrix $*$-algebra generated by symmetric matrices. The method does not require any algebraic structure to be known in advance, whereas its validity relies on matrix $*$-algebra theory. The method is composed of numerical-linear algebraic computations such as eigenvalue computation, and automatically makes the full use of the underlying algebraic structure, which is often an outcome of physical or geometrical symmetry, sparsity, and structural or numerical degeneracy in the given matrices. Numerical examples of truss design are also presented.

Key words.
matrix $*$-algebra, block-diagonalization, group symmetry, sparsity, semidefinite programming

AMS Classifications: 15A21, 90C22

ℵ This manuscript is also issued as METR 2007-52, Department of Mathematical Informatics, University of Tokyo

⋆ Department of Mathematical Informatics, Graduate School of Information Science and Technology, University of Tokyo, Tokyo 113-8656, Japan. murota@mist.i.u-tokyo.ac.jp

† Department of Mathematical Informatics, Graduate School of Information Science and Technology, University of Tokyo, Tokyo 113-8656, Japan. kanno@mist.i.u-tokyo.ac.jp

‡ Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, Tokyo 152-8552, Japan. kojima@is.titech.ac.jp

♯ Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, Tokyo 152-8552, Japan. sadayosi@is.titech.ac.jp

1 Introduction

This paper is motivated by recent studies on group symmetries in semidefinite programs (SDPs) and sum of squares (SOS) and SDP relaxations [1, 4, 5, 6, 7]. A common and essential problem in these studies can be stated as follows: Given a finite set of $n \times n$ real symmetric matrices $A_1, A_2, \ldots, A_m$, find an $n \times n$ orthogonal matrix P that provides them with a simultaneous block-diagonal decomposition, i.e., such that $P^\top A_1 P, P^\top A_2 P, \ldots, P^\top A_m P$ become block-diagonal matrices with a common block-diagonal structure. Here $A_1, A_2, \ldots, A_m$ correspond to data matrices associated with an SDP. We say that the set of given matrices $A_1, A_2, \ldots, A_m$ is decomposed into a set of block-diagonal matrices or that the SDP is decomposed into an SDP with the block-diagonal data matrices. Such a block-diagonal decomposition is not unique in general; for example, any symmetric matrix may trivially be regarded as a one-block matrix. As diagonal-blocks of the decomposed matrices get smaller, so does the amount of input data of the transformed SDP with the decomposed matrices, and the transformed SDP could be solved more efficiently by existing software packages developed for SDPs [2, 13, 14, 17]. Naturally we are interested in a finest decomposition. A more specific account of the decomposition of SDPs will be given in Section 2.1.

There are two different but closely related theoretical frameworks with which we can address our problem of finding a block-diagonal decomposition for a finite set of given $n \times n$ real symmetric matrices. The one is group representation theory [10, 12] and the other matrix $*$-algebra [15]. They are not only necessary to answer the fundamental theoretical question of the existence of such a finest block-diagonal decomposition but also useful in its computation. Both frameworks have been utilized in the literature [1, 4, 5, 6, 7] cited above.

Kanno et al. [7] introduced a class of group symmetric SDPs, which arise from topology optimization problems of trusses, and derived symmetry of central paths which play a fundamental role in the primal-dual interior-point method [16] for solving them. Gatermann and Parrilo [5] investigated the problem of minimizing a group symmetric polynomial. They proposed to reduce the size of SOS and SDP relaxations for the problem by exploiting the group symmetry and decomposing the SDP. On the other hand, de Klerk et al. [3] applied the theory of matrix $*$-algebra to reduce the size of a class of group symmetric SDPs. Instead of decomposing a given SDP by using its group symmetry, their method transforms the problem to an equivalent SDP through a $*$-algebra isomorphism. We also refer to Kojima et al. [8] as a paper where matrix $*$-algebra was studied in connection with SDPs. Jansson et al. [6] brought group symmetries into equality-inequality constrained polynomial optimization problems and their SDP relaxation. More recently, de Klerk and Sotirov [4] dealt with quadratic assignment problems, and showed how to exploit their group symmetries to reduce the size of their SDP relaxations (see Remark 4.3 for more account).

All existing studies [1, 4, 5, 6] on group symmetric SDPs mentioned above assume that the algebraic structure such as group symmetry and matrix $*$-algebra behind a given SDP is known in advance before computing a decomposition of the SDP. Such an algebraic structure arises naturally from the physical or geometrical structure underlying the SDP, so the assumption is certainly practical and reasonable. When we assume symmetry of an SDP (or the data matrices $A_1, A_2, \ldots, A_m$) with reference to a group G, to be specific, we are in fact considering the class of SDPs that enjoy the same group symmetry. As a consequence, the resulting transformation matrix P is universal in the sense that it is valid for the decomposition of all SDPs belonging to the class. Whereas this universality may often be desirable in practice, we should be aware of the obvious fact that the given SDP is just a specific instance in the class. This means that the given problem may possibly satisfy an additional algebraic structure which is not captured by the assumed group symmetry but which can be exploited for a further decomposition. Such an additional algebraic structure is often induced from sparsity of the data matrices of the SDP, as we see in the topology optimization problem of trusses in Section 5. The possibility of a further decomposition due to sparsity will be illustrated in Sections 2.2 and 5.2.

In this paper we propose a numerical method for finding a finest simultaneous block-diagonal decomposition of a finite number of $n \times n$ real symmetric matrices $A_1, A_2, \ldots, A_m$. The method does not require any algebraic structure to be known in advance, and is based on purely linear algebraic computations such as eigenvalue computation. It is free from group representation theory or matrix $*$-algebra during its execution, although its validity relies on matrix $*$-algebra theory. This main feature of our method makes it possible to compute a finest block-diagonal decomposition by taking into account the underlying physical or geometrical symmetry, the sparsity of the given matrices, and some other implicit or overlooked symmetry.

Our method is based on the following ideas. We consider the matrix $*$-algebra $\mathcal{T}$ generated by $A_1, A_2, \ldots, A_m$ with the identity matrix, and make use of a well-known fundamental fact (see Theorem 3.1) about the decomposition of $\mathcal{T}$ into simple components and irreducible components. The key observation is that the decomposition into simple components can be computed from the eigenvalue (or spectral) decomposition of a random symmetric matrix in $\mathcal{T}$. Once the simple components are identified, the decomposition into irreducible components can be obtained by "local" coordinate changes within each eigenspace, to be explained in Section 3.

This paper is organized as follows. Section 2 illustrates our motivation of simultaneous block-diagonalization and the notion of the finest block-diagonal decomposition. Section 3 describes the theoretical background of our algorithm based on matrix $*$-algebra. In Section 4, we present an algorithm for computing the finest simultaneous block-diagonalization, as well as a suggested practical variant thereof. Numerical results are shown in Section 5; SDP problems arising from topology optimization of symmetric trusses in Section 5.1 and illustrative small examples in Section 5.2.

2 Illustration of Motivations

This section is devoted to illustration of the motivations mentioned in Introduction.

2.1 Decomposition of semidefinite programs

It is explained how simultaneous block diagonalization can be utilized in semidefinite programming.

Let $A_p \in \mathcal{S}_n$ $(p = 0, 1, \ldots, m)$ and $b = (b_p)_{p=1}^m \in \mathbb{R}^m$ be constant matrices and a constant vector, where $\mathcal{S}_n$ denotes the set of $n \times n$ symmetric real matrices. The standard form of primal-dual pair of *semidefinite programming* (SDP) problems is formulated as

$$\left.\begin{array}{ll} \min & A_0 \bullet X \\ \text{s.t.} & A_p \bullet X = b_p, \quad p = 1, \ldots, m, \\ & \mathcal{S}_n \ni X \succeq O; \end{array}\right\} \tag{2.1}$$

$$\left.\begin{array}{ll} \max & b^\top y \\ \text{s.t.} & Z + \displaystyle\sum_{p=1}^{m} A_p y_p = A_0, \\ & \mathcal{S}_n \ni Z \succeq O. \end{array}\right\} \tag{2.2}$$

It should be clear that $A \bullet X = \mathrm{tr}(AX)$ for symmetric matrices A and X, $X \succeq O$ means that X is positive semidefinite, and $^\top$ denotes the transpose of a vector or a matrix.

Suppose that $A_0, A_1, \ldots, A_m$ are transformed into block-diagonal matrices by an $n \times n$ orthogonal matrix P as

$$P^\top A_p P = \begin{pmatrix} A_p^{(1)} & O \\ O & A_p^{(2)} \end{pmatrix}, \quad p = 0, 1, \ldots, m,$$

where $A_p^{(1)} \in \mathcal{S}_{n_1}$, $A_p^{(2)} \in \mathcal{S}_{n_2}$, and $n_1 + n_2 = n$. The problems (2.1) and

(2.2) can be reduced to

$$\left.\begin{array}{ll}\min & A_0^{(1)} \bullet X_1 + A_0^{(2)} \bullet X_2 \\ \text{s.t.} & A_p^{(1)} \bullet X_1 + A_p^{(2)} \bullet X_2 = b_p, \quad p = 1, \ldots, m, \\ & \mathcal{S}_{n_1} \ni X_1 \succeq O, \quad \mathcal{S}_{n_2} \ni X_2 \succeq O;\end{array}\right\} \tag{2.3}$$

$$\left.\begin{array}{ll}\max & b^\top y \\ \text{s.t.} & Z_1 + \sum_{p=1}^{m} A_p^{(1)} y_p = A_0^{(1)}, \\ & Z_2 + \sum_{p=1}^{m} A_p^{(2)} y_p = A_0^{(2)}, \\ & \mathcal{S}_{n_1} \ni Z_1 \succeq O, \quad \mathcal{S}_{n_2} \ni Z_2 \succeq O.\end{array}\right\} \tag{2.4}$$

Note that the number of variables of (2.3) is smaller than that of (2.1). The constraint on the $n \times n$ symmetric matrix in (2.2) is reduced to the constraints on the two matrices in (2.4) with smaller sizes.

It is expected that the computational time required by the primal-dual interior-point method is reduced drastically if the problems (2.1) and (2.2) can be reformulated as (2.3) and (2.4). This motivates us to investigate a numerical technique for computing a simultaneous block diagonalization in the form of

$$A_p = \bigoplus_{j=1}^{t} A_p^{(j)}, \quad A_p^{(j)} \in \mathcal{S}_{n_j}, \tag{2.5}$$

where $A_p \in \mathcal{S}_n$ $(p = 0, 1, \ldots, m)$ are given symmetric matrices. Here $\bigoplus$ designates a direct sum of the summand matrices, which contains the summands as diagonal blocks.

2.2 Group symmetry and additional structure due to sparsity

With reference to a concrete example, we illustrate the use of group symmetry and also the possibility of a finer decomposition based on an additional algebraic structure due to sparsity.

Consider an $n \times n$ matrix of the form

$$A = \begin{bmatrix} B & E & E & C \\ E & B & E & C \\ E & E & B & C \\ C^\top & C^\top & C^\top & D \end{bmatrix} \tag{2.6}$$

with $B \in \mathcal{S}_{n_\mathrm{B}}$ and $D \in \mathcal{S}_{n_\mathrm{D}}$. Such a matrix with $E = O$ is sometimes referred to as a *bordered block-diagonal matrix*. Obviously we have $A =$

$A_1 + A_2 + A_3 + A_4$ with

$$A_1 = \begin{bmatrix} B & O & O & O \\ O & B & O & O \\ O & O & B & O \\ O & O & O & O \end{bmatrix}, \quad A_2 = \begin{bmatrix} O & O & O & C \\ O & O & O & C \\ O & O & O & C \\ C^\top & C^\top & C^\top & O \end{bmatrix}, \tag{2.7}$$

$$A_3 = \begin{bmatrix} O & O & O & O \\ O & O & O & O \\ O & O & O & O \\ O & O & O & D \end{bmatrix}, \quad A_4 = \begin{bmatrix} O & E & E & O \\ E & O & E & O \\ E & E & O & O \\ O & O & O & O \end{bmatrix}. \tag{2.8}$$

Let P be an $n \times n$ orthogonal matrix defined by

$$P = \left[\begin{array}{cc|cc} I_{n^{\mathrm{B}}}/\sqrt{3} & O & I_{n^{\mathrm{B}}}/\sqrt{2} & I_{n^{\mathrm{B}}}/\sqrt{6} \\ I_{n^{\mathrm{B}}}/\sqrt{3} & O & -I_{n^{\mathrm{B}}}/\sqrt{2} & I_{n^{\mathrm{B}}}/\sqrt{6} \\ I_{n^{\mathrm{B}}}/\sqrt{3} & O & O & -2I_{n^{\mathrm{B}}}/\sqrt{6} \\ O & I_{n^{\mathrm{D}}} & O & O \end{array}\right], \tag{2.9}$$

where for any n', $I_{n'}$ denotes the $n' \times n'$ identity matrix. With this P the matrices A_p are transformed to block-diagonal matrices as

$$P^\top A_1 P = \left[\begin{array}{cc|c|c} B & O & O & O \\ O & O & O & O \\ \hline O & O & B & O \\ \hline O & O & O & B \end{array}\right] = \begin{bmatrix} B & O \\ O & O \end{bmatrix} \oplus B \oplus B, \tag{2.10}$$

$$P^\top A_2 P = \left[\begin{array}{cc|c|c} O & \sqrt{3}C & O & O \\ \sqrt{3}C^\top & O & O & O \\ \hline O & O & O & O \\ \hline O & O & O & O \end{array}\right] = \begin{bmatrix} O & \sqrt{3}C \\ \sqrt{3}C^\top & O \end{bmatrix} \oplus O \oplus O, \tag{2.11}$$

$$P^\top A_3 P = \left[\begin{array}{cc|c|c} O & O & O & O \\ O & D & O & O \\ \hline O & O & O & O \\ \hline O & O & O & O \end{array}\right] = \begin{bmatrix} O & O \\ O & D \end{bmatrix} \oplus O \oplus O, \tag{2.12}$$

$$P^\top A_4 P = \left[\begin{array}{cc|c|c} 2E & O & O & O \\ O & O & O & O \\ \hline O & O & -E & O \\ \hline O & O & O & -E \end{array}\right] = \begin{bmatrix} 2E & O \\ O & O \end{bmatrix} \oplus (-E) \oplus (-E). \tag{2.13}$$

Note that the partition of P is not symmetric for rows and columns; we have $(n^{\mathrm{B}}, n^{\mathrm{B}}, n^{\mathrm{B}}, n^{\mathrm{D}})$ for row-block sizes and $(n^{\mathrm{B}}, n^{\mathrm{D}}, n^{\mathrm{B}}, n^{\mathrm{B}})$ for column-block sizes. As is shown in (2.10)–(2.13), A_1, A_2, A_3 and A_4 are decomposed simultaneously in the form of (2.5) with $t = 3$, $n_1 = n^{\mathrm{B}} + n^{\mathrm{D}}$, and $n_2 = n_3 = n^{\mathrm{B}}$. Moreover, the second and third blocks coincide, i.e., $A_p^{(2)} = A_p^{(3)}$, for each p.

The decomposition described above conforms with the standard decomposition technique [10, 12] for systems with group symmetry. The matrices A_p above are symmetric with respect to S_3, the symmetric group of degree 3, in that

$$T(g)^\top A_p T(g) = A_p, \quad \forall g \in G, \quad p = 1, \dots, m \tag{2.14}$$

holds for $G = \mathrm{S}_3$ and $m = 4$. Here the family of matrices $T(g)$, indexed by elements of G, is an orthogonal matrix representation of G in general. In the present example, the S_3-symmetry formulated in (2.14) is equivalent to

$$T_i^\top A_p T_i = A_p, \quad i = 1, 2, \quad p = 1, 2, 3, 4$$

with

$$T_1 = \begin{bmatrix} O & I_{n^\mathrm{B}} & O & O \\ I_{n^\mathrm{B}} & O & O & O \\ O & O & I_{n^\mathrm{B}} & O \\ O & O & O & I_{n^\mathrm{D}} \end{bmatrix}, \quad T_2 = \begin{bmatrix} O & I_{n^\mathrm{B}} & O & O \\ O & O & I_{n^\mathrm{B}} & O \\ I_{n^\mathrm{B}} & O & O & O \\ O & O & O & I_{n^\mathrm{D}} \end{bmatrix}.$$

According to group representation theory, a simultaneous block-diagonal decomposition of A_p is obtained through the decomposition of the representation $T(g)$ into irreducible representations. In the present example, we have

$$P^\top T_1 P = \left[\begin{array}{cc|cc} I_{n^\mathrm{B}} & O & O & O \\ O & I_{n^\mathrm{D}} & O & O \\ \hline O & O & -I_{n^\mathrm{B}} & O \\ O & O & O & I_{n^\mathrm{B}} \end{array}\right], \tag{2.15}$$

$$P^\top T_2 P = \left[\begin{array}{cc|cc} I_{n^\mathrm{B}} & O & O & O \\ O & I_{n^\mathrm{D}} & O & O \\ \hline O & O & -I_{n^\mathrm{B}}/2 & \sqrt{3} I_{n^\mathrm{B}}/2 \\ O & O & -\sqrt{3} I_{n^\mathrm{B}}/2 & -I_{n^\mathrm{B}}/2 \end{array}\right], \tag{2.16}$$

where that the first two blocks correspond to the unit representation (with multiplicity $n^\mathrm{B} + n^\mathrm{D}$) and the last two blocks to the two-dimensional irreducible representation (with multiplicity n^B).

The transformation matrix P in (2.9) is universal in the sense that it brings any matrix A satisfying $T_i^\top A T_i = A$ for $i = 1, 2$ into the same block-diagonal form. Put otherwise, the decomposition given in (2.10)–(2.13) is the finest possible decomposition that is valid for the class of matrices enjoying the S_3-symmetry. It is noted in this connection that the underlying group G, as well as its representation $T(g)$, is often evident in practice, reflecting the geometrical or physical symmetry of the problem in question.

The universality of the decomposition explained above is certainly a nice feature of the group-theoretic method, but what we are really interested in is the decomposition of a single specific instance of a set of matrices. As a simplest example, suppose that the given matrix (2.6) is a bordered block-diagonal matrix, with $E = O$. Then the decomposition in (2.10)–(2.13) is not the finest possible, but the last two identical blocks, i.e., $A_p^{(2)}$ and $A_p^{(3)}$, can be decomposed further into diagonal matrices by the eigenvalue (or spectral) decomposition of B. Although this argument is too simple to be convincing, it is sufficient to suggest the possibility that a finer decomposition may possibly be obtained from an additional algebraic structure that is not ascribed to the assumed group symmetry. Such an additional algebraic structure often stems from sparsity, as is the case with the topology optimization problem of trusses treated in Section 5.1.

Mathematically, such an additional algebraic structure could also be described as a group symmetry by introducing a larger group. But this larger group would be difficult to identify in practice, since it is determined as a result of the interaction between the underlying geometrical or physical symmetry and other factors, such as sparsity and parameter dependence. The method of block-diagonalization proposed in this paper will automatically exploit such algebraic structure in the course of numerical computation. Numerical examples in Section 5.2 will demonstrate that the proposed method can cope with different kinds of additional algebraic structures for the matrix (2.6).

3 Mathematical Basis

We introduce some mathematical facts that will serve as a basis for our algorithm to be described in Section 4.

3.1 Matrix $*$-algebras

Let $\mathcal{M}_n$ denote the set of $n \times n$ real matrices. A subset $\mathcal{T}$ of $\mathcal{M}_n$ is said to be a *$*$-subalgebra* (or a *matrix $*$-algebra*) over $\mathbb{R}$ if $I_n \in \mathcal{T}$ and

$$A, B \in \mathcal{T}; \alpha, \beta \in \mathbb{R} \implies \alpha A + \beta B, AB, A^\top \in \mathcal{T}.$$

We say that $\mathcal{T}$ is *simple* if $\mathcal{T}$ has no ideal other than $\{O\}$ and $\mathcal{T}$ itself, where an *ideal* of $\mathcal{T}$ means a subset $\mathcal{I}$ of $\mathcal{T}$ such that

$$A \in \mathcal{T}, B \in \mathcal{I} \implies AB \in \mathcal{I}.$$

A linear subspace W of $\mathbb{R}^n$ is said to be *invariant* with respect to $\mathcal{T}$, or *$\mathcal{T}$-invariant*, if $AW \subseteq W$ for every $A \in \mathcal{T}$. We say that $\mathcal{T}$ is *irreducible* if no $\mathcal{T}$-invariant subspace other than $\{\mathbf{0}\}$ and $\mathbb{R}^n$ exists. If $\mathcal{T}$ is irreducible, it is simple.

In this paper we are particularly interested in a $*$-subalgebra generated by symmetric matrices. From a standard result of the theory of matrix $*$-algebra (e.g., [15, Chapter X], [8, Theorem 5.4]) we can see the following structure theorem for such a $*$-subalgebra. Note that, for an orthogonal matrix P, the set of transformed matrices

$$P^\top \mathcal{T} P = \{P^\top A P \mid A \in \mathcal{T}\}$$

forms another $*$-subalgebra.

Theorem 3.1. Let $\mathcal{T}$ be a $*$-subalgebra of $\mathcal{M}_n$ generated by symmetric matrices.

(A) There exists an orthogonal matrix $\hat{Q} \in \mathcal{M}_n$ and simple $*$-subalgebras $\mathcal{T}_j$ of $\mathcal{M}_{\hat{n}_j}$ for some $\hat{n}_j$ $(j = 1, 2, \ldots, \ell)$ such that

$$\hat{Q}^\top \mathcal{T} \hat{Q} = \{\mathrm{diag}\,(S_1, S_2, \ldots, S_\ell) : S_j \in \mathcal{T}_j \ (j = 1, 2, \ldots, \ell)\}.$$

(B) If $\mathcal{T}$ is simple, there exists an orthogonal matrix $P \in \mathcal{M}_n$ and an irreducible $*$-subalgebra $\mathcal{T}'$ of $\mathcal{M}_{\bar{n}}$ for some $\bar{n}$ such that

$$P^\top \mathcal{T} P = \{\mathrm{diag}\,(B, B, \ldots, B) : B \in \mathcal{T}'\}.$$

(C) If $\mathcal{T}$ is irreducible, then $\mathcal{T} = \mathcal{M}_n$. ■

It follows from the above theorem that, with a single orthogonal matrix P, all the matrices in $\mathcal{T}$ can be transformed simultaneously to a block-diagonal form as

$$P^\top A P = \bigoplus_{j=1}^{\ell} \bigoplus_{i=1}^{\bar{m}_j} B_j = \bigoplus_{j=1}^{\ell} (I_{\bar{m}_j} \otimes B_j) \tag{3.1}$$

with $B_j \in \mathcal{M}_{\bar{n}_j}$, where the structural indices ℓ and $\bar{n}_j$, $\bar{m}_j$ for $j = 1, \ldots, \ell$ are determined by $\mathcal{T}$. It may be noted that $\hat{n}_j$ in Theorem 3.1 (A) is equal to $\bar{m}_j \bar{n}_j$ in the present notation. Conversely, for any choice of $B_j \in \mathcal{M}_{\bar{n}_j}$ for $j = 1, \ldots, \ell$, the matrix of (3.1) belongs to $P^\top \mathcal{T} P$. We denote by

$$\mathbb{R}^n = \bigoplus_{j=1}^{\ell} U_j \tag{3.2}$$

the decomposition of $\mathbb{R}^n$ that corresponds to the simple components. In other words, $U_j = \mathrm{Im}(\hat{Q}_j)$ for the $n \times \hat{n}_j$ submatrix $\hat{Q}_j$ of $\hat{Q}$ that corresponds to $\mathcal{T}_j$ in Theorem 3.1 (A). Although the matrix $\hat{Q}$ is not unique, the subspace U_j is determined uniquely and $\dim U_j = \hat{n}_j = \bar{m}_j \bar{n}_j$ for $j = 1, \ldots, \ell$.

3.2 Simple components from eigenspaces

Let $A_1, \ldots, A_m \in \mathcal{S}_n$ be $n \times n$ symmetric real matrices, and $\mathcal{T}$ be the $*$-subalgebra over $\mathbb{R}$ generated by $\{I_n, A_1, \ldots, A_m\}$. Note that (3.1) holds for every $A \in \mathcal{T}$ if and only if (3.1) holds for $A = A_p$ for $p = 1, \ldots, m$.

A key observation for our algorithm is that the decomposition (3.2) into simple components can be computed from the eigenvalue (or spectral) decomposition of a single matrix A in $\mathcal{T} \cap \mathcal{S}_n$ if it is free from degeneracy in eigenvalues.

Let A be a symmetric matrix in $\mathcal{T}$, and $\alpha_1, \ldots, \alpha_k$ be the distinct eigenvalues of A with multiplicities denoted as $m_1, \ldots, m_k$, and $Q = [Q_1, \ldots, Q_k]$ be an orthogonal matrix consisting of the eigenvectors, where Q_i is an $n \times m_i$ matrix for $i = 1, \ldots, k$. Then we have

$$Q^\top A Q = \mathrm{diag}\,(\alpha_1 I_{m_1}, \ldots, \alpha_k I_{m_k}) = \begin{array}{|c|c|c|c|} \hline \alpha_1 I_{m_1} & O & O & O \\ \hline O & \alpha_2 I_{m_2} & O & O \\ \hline O & O & \ddots & O \\ \hline O & O & O & \alpha_k I_{m_k} \\ \hline \end{array}. \quad (3.3)$$

Put $K = \{1, \ldots, k\}$ and for $i \in K$ define $V_i = \mathrm{Im}(Q_i)$, which is the eigenspace corresponding to α_i.

Let us say that $A \in \mathcal{T} \cap \mathcal{S}_n$ is *generic in eigenvalue structure* (or simply *generic*) if all the matrices $B_1, \ldots, B_\ell$ appearing in the decomposition (3.1) of A are free from multiple eigenvalues and no two of them share a common eigenvalue. For a generic matrix A the number k of distinct eigenvalues is equal to $\sum_{j=1}^{\ell} \bar{n}_j$ and the list (multiset) of their multiplicities $\{m_1, \ldots, m_k\}$ is the union of $\bar{n}_j$ copies of $\bar{m}_j$ over $j = 1, \ldots, \ell$.

The eigenvalue decomposition of a generic A is consistent with the decomposition (3.2) into simple components of $\mathcal{T}$, as follows.

Proposition 3.2. Let $A \in \mathcal{T} \cap \mathcal{S}_n$ be generic in eigenvalue structure. For any $i \in \{1, \ldots, k\}$ there exists $j \in \{1, \ldots, \ell\}$ such that $V_i \subseteq U_j$. Hence there exists a partition of $K = \{1, \ldots, k\}$ into ℓ disjoint subsets:

$$K = K_1 \cup \cdots \cup K_\ell \quad (3.4)$$

such that

$$U_j = \bigoplus_{i \in K_j} V_i, \qquad j = 1, \ldots, \ell. \quad (3.5)$$

■

Note that $m_i = \bar{m}_j$ for $i \in K_j$ and $|K_j| = \bar{n}_j$ for $j = 1, \ldots, \ell$.

The partition (3.4) of K can be determined as follows. Define a binary relation $\sim$ on K by:

$$i \sim i' \quad \Longleftrightarrow \quad \exists p\ (1 \le p \le m) : Q_i^\top A_p Q_{i'} \neq O, \quad (3.6)$$

where $i, i' \in K$. By convention we define $i \sim i$ for any $i \in K$.

Proposition 3.3. The partition (3.4) coincides with the partition of K into equivalence classes of the transitive closure of the binary relation $\sim$.

Proof. This is not difficult to see from the general theory of matrix $*$-algebra, but a proof is given here for completeness. Denote by $\{L_1, \ldots, L_{\ell'}\}$ the equivalence classes with respect to $\sim$.

If $i \sim i'$, then $Q_i^\top A_p Q_{i'} \neq O$ for some p. This means that for any $I \subseteq K$ with $i \in I$ and $i' \in K \setminus I$, the subspace $\bigoplus_{i'' \in I} V_{i''}$ is not invariant under A_p. Hence $V_{i'}$ must be contained in the same simple component as V_i. Therefore each L_j must be contained in some $K_{j'}$.

To show the converse, define a matrix $\tilde{Q}_j = (Q_i \mid i \in L_j)$, which is of size $n \times \sum_{i \in L_j} m_i$, and an $n \times n$ matrix $E_j = \tilde{Q}_j \tilde{Q}_j^\top$ for $j = 1, \ldots, \ell'$. Each matrix E_j belongs to $\mathcal{T}$, as shown below, and it is idempotent (i.e., ${E_j}^2 = E_j$) and $E_1 + \cdots + E_{\ell'} = I_n$. On the other hand, for distinct j and j' we have $\tilde{Q}_j^\top A_p \tilde{Q}_{j'} = O$ for all p, and hence $\tilde{Q}_j^\top M \tilde{Q}_{j'} = O$ for all $M \in \mathcal{T}$. This implies that $E_j M = M E_j$ for all $M \in \mathcal{T}$. Therefore $\mathrm{Im}(E_j)$ is a union of simple components, and hence L_j is a union of some $K_{j'}$'s.

It remains to show that $E_j \in \mathcal{T}$. Since α_i's are distinct, for any real numbers $u_1, \ldots, u_k$ there exists a polynomial f such that $f(\alpha_i) = u_i$ for $i = 1, \ldots, k$. Let f_j be such f for $(u_1, \ldots, u_k)$ defined as $u_i = 1$ for $i \in L_j$ and $u_i = 0$ for $i \in K \setminus L_j$. Then $E_j = \tilde{Q}_j \tilde{Q}_j^\top = Q \cdot f_j(\mathrm{diag}\,(\alpha_1 I_{m_1}, \ldots, \alpha_k I_{m_k})) \cdot Q^\top = Q \cdot f_j(Q^\top A Q) \cdot Q^\top = f_j(A)$. This shows $E_j \in \mathcal{T}$. □

A generic matrix A can be obtained as a random linear combination of generators, as follows. For a real vector $r = (r_1, \ldots, r_m)$ put

$$A(r) = r_1 A_1 + \cdots + r_m A_m.$$

We denote by $\mathrm{span}\{\cdots\}$ the set of linear combinations of the matrices in the braces.

Proposition 3.4. If $\mathrm{span}\{I_n, A_1, \ldots, A_m\} = \mathcal{T} \cap \mathcal{S}_n$, there exists an open dense subset R of $\mathbb{R}^m$ such that $A(r)$ is generic in eigenvalue structure for every $r \in R$.

Proof. Let B_{pj} denote the matrix B_j in the decomposition (3.1) of $A = A_p$ for $p = 1, \ldots, m$. For $j = 1, \ldots, \ell$ define $f_j(\lambda) = f_j(\lambda; r) = \det(\lambda I - (r_1 B_{1j} + \cdots + r_m B_{mj}))$, which is a polynomial in $\lambda, r_1, \ldots, r_m$. By the assumption on the linear span of generators, $f_j(\lambda)$ is free from multiple roots for at least one $r \in \mathbb{R}^m$, and it has a multiple root only if r lies on the algebraic set, say, Σ_j defined by the resultant of $f_j(\lambda)$ and $f_j'(\lambda)$. For distinct j and j', $f_j(\lambda)$ and $f_{j'}(\lambda)$ do not share a common root for at least one $r \in \mathbb{R}^m$, and they have a common root only if r lies on the algebraic set, say, $\Sigma_{jj'}$ defined by the resultant of $f_j(\lambda)$ and $f_{j'}(\lambda)$. Then we can take $R = \mathbb{R}^m \setminus \left[(\cup_j \Sigma_j) \cup \left(\cup_{j,j'} \Sigma_{jj'}\right)\right]$. □

We may assume that the coefficient vector r is normalized, for example, to $\|r\|_2 = 1$, where $\|r\|_2 = \sqrt{\sum_{p=1}^m {r_p}^2}$. Then the above proposition implies that $A(r)$ is generic for almost all values of r, or with probability one if r is chosen at random. It should be clear that we can adopt any natural normalization scheme for this statement.

3.3 Transformation for irreducible components

Once the transformation matrix Q for the eigenvalue decomposition of a generic matrix A is known, the transformation P for $\mathcal{T}$ can be obtained through "local" transformations within eigenspaces corresponding to distinct eigenvalues, followed by a "global" permutation of rows and columns.

Proposition 3.5. Let $A \in \mathcal{T} \cap \mathcal{S}_n$ be generic in eigenvalue structure, and $Q^\top AQ = \mathrm{diag}\,(\alpha_1 I_{m_1}, \ldots, \alpha_k I_{m_k})$ be the eigenvalue decomposition as in (3.3). Then the transformation matrix P in (3.1) can be chosen in the form of

$$P = Q \cdot \mathrm{diag}\,(P_1, \ldots, P_k) \cdot \Pi \tag{3.7}$$

with orthogonal matrices $P_i \in \mathcal{M}_{m_i}$ for $i = 1, \ldots, k$, and a permutation matrix $\Pi \in \mathcal{M}_n$.

Proof. For simplicity of presentation we focus on a simple component, which is tantamount to assuming $P^\top AP = I_{\bar{m}} \otimes B$, where $\bar{m} = m_1 = \cdots = m_k$. Since α_i's are the eigenvalues of B, there exists an orthogonal matrix S such that $S^\top BS = D$, where $D = \mathrm{diag}\,(\alpha_1, \ldots, \alpha_k)$. Hence we have

$$(I_{\bar{m}} \otimes S)^\top (P^\top AP)(I_{\bar{m}} \otimes S) = I_{\bar{m}} \otimes D.$$

With the notation Π for the permutation matrix such that $\Pi(I_{\bar{m}} \otimes M)\Pi^\top = M \otimes I_{\bar{m}}$ for every $k \times k$ matrix M we can rewrite this as

$$\left(P(I_{\bar{m}} \otimes S)\Pi^\top \right)^\top A \left(P(I_{\bar{m}} \otimes S)\Pi^\top \right) = D \otimes I_{\bar{m}}.$$

Comparing this with $Q^\top AQ = D \otimes I_{\bar{m}}$ and noting that α_i's are distinct, we see that

$$P(I_{\bar{m}} \otimes S)\Pi^\top = Q \cdot \mathrm{diag}\,(\tilde{P}_1, \ldots, \tilde{P}_k)$$

for some $\bar{m} \times \bar{m}$ orthogonal matrices $\tilde{P}_i$. Since the left-hand side above is equal to $P\Pi^\top \mathrm{diag}\,(S, \ldots, S)$, this implies (3.7) with $P_i = \tilde{P}_i S^\top$. □

4 Algorithm for Simultaneous Block-Diagonalization

On the basis of the theoretical considerations in Section 3, we propose in this section an algorithm for simultaneous block-diagonalization of given symmetric matrices $A_1, \ldots, A_m \in \mathcal{S}_n$ by an orthogonal matrix P:

$$P^\top A_p P = \bigoplus_{j=1}^{\ell} \bigoplus_{i=1}^{\bar{m}_j} B_{pj} = \bigoplus_{j=1}^{\ell} (I_{\bar{m}_j} \otimes B_{pj}), \quad p = 1, \ldots, m, \tag{4.1}$$

where $B_{pj} \in \mathcal{M}_{\bar{n}_j}$ for $j = 1, \ldots, \ell$ and $p = 1, \ldots, m$. Our algorithm consists of two parts corresponding to (A) and (B) of Theorem 3.1 for the $*$-subalgebra $\mathcal{T}$ generated by $\{I_n, A_1, \ldots, A_m\}$. The former (Section 4.1) corresponds to the decomposition of $\mathcal{T}$ into simple components and the latter (Section 4.2) to the decomposition into irreducible components. A practical variant of the algorithm is described in Section 4.3.

4.1 Decomposition into simple components

We present here an algorithm for the decomposition into simple components. Algorithm 4.1 below does not presume $\text{span}\{I_n, A_1, \ldots, A_m\} = \mathcal{T} \cap \mathcal{S}_n$, although its correctness relies on this condition.

Algorithm 4.1.

Step 1: Generate random numbers $r_1, \ldots, r_m$ (with $\|r\|_2 = 1$), and put $A = \sum_{p=1}^{m} r_p A_p$.

Step 2: Compute the eigenvalues and eigenvectors of A. Let $\alpha_1, \ldots, \alpha_k$ be the distinct eigenvalues of A with their multiplicities denoted by $m_1, \ldots, m_k$. Let $Q_i \in \mathbb{R}^{n \times m_i}$ be the matrix consisting of orthonormal eigenvectors corresponding to α_i, and define the matrix $Q \in \mathbb{R}^{n \times n}$ by $Q = (Q_i \mid i = 1, \ldots, k)$. This means that

$$Q^\top A Q = \text{diag}\,(\alpha_1 I_{m_1}, \ldots, \alpha_k I_{m_k}).$$

Step 3: Put $K = \{1, \ldots, k\}$, and let $\sim$ be a binary relation on K defined by

$$i \sim i' \quad \Longleftrightarrow \quad \exists p \ (1 \leq p \leq m) : Q_i^\top A_p Q_{i'} \neq O, \tag{4.2}$$

where $i, i' \in K$. Let

$$K = K_1 \cup \cdots \cup K_\ell \tag{4.3}$$

be the partition of K consisting of the equivalence classes of the transitive closure of the binary relation $\sim$. Define matrices $Q[K_j]$ by

$$Q[K_j] = (Q_i \mid i \in K_j), \quad j = 1, \ldots, \ell,$$

and set

$$\hat{Q} = (Q[K_1], \ldots, Q[K_\ell]) \,.$$

Compute $\hat{Q}^\top A_p \hat{Q}$ ($p = 1, \ldots, m$), which results in a simultaneous block-diagonalization with respect to the partition (3.4).

Example 4.1. Suppose that the number of the distinct eigenvalues of A is five, i.e., $K = \{1, 2, 3, 4, 5\}$, and that the partition of K is obtained as $K_1 = \{1, 2, 3\}$, $K_2 = \{4\}$, and $K_3 = \{5\}$, where $\ell = 3$. Then $A_1, \ldots, A_m$ are decomposed simultaneously as

$$\hat{Q}^\top A_p \hat{Q} = \begin{array}{|ccc|c|c|} \multicolumn{1}{c}{m_1} & \multicolumn{1}{c}{m_2} & \multicolumn{1}{c}{m_3} & \multicolumn{1}{c}{m_4} & \multicolumn{1}{c}{m_5} \\ \hline * & * & * & O & O \\ * & * & * & O & O \\ * & * & * & O & O \\ \hline O & O & O & * & O \\ \hline O & O & O & O & * \\ \hline \end{array} \tag{4.4}$$

for $p = 1, \ldots, m$. ∎

For the correctness of the above algorithm we have the following.

Proposition 4.2. If the matrix A generated in Step 1 is generic in eigenvalue structure, the orthogonal matrix $\hat{Q}$ constructed by Algorithm 4.1 gives the transformation matrix $\hat{Q}$ in Theorem 3.1 (A) for the decomposition of $\mathcal{T}$ into simple components.

Proof. This follows from Propositions 3.2 and 3.3. □

Proposition 3.4 implies that the matrix A in Step 1 is generic with probability one if $\mathrm{span}\{I_n, A_1, \ldots, A_m\} = \mathcal{T} \cap \mathcal{S}_n$. This condition, however, is not always satisfied by the given matrices $A_1, \ldots, A_m$. In such a case we can generate a basis of $\mathcal{T} \cap \mathcal{S}_n$ as follows. First choose a linearly independent subset, say, $\mathcal{B}_1$ of $\{I_n, A_1, \ldots, A_m\}$. For $k = 1, 2, \cdots$ let $\mathcal{B}_{k+1}$ $(\supseteq \mathcal{B}_k)$ be a linearly independent subset of $\{(AB + BA)/2 \mid A \in \mathcal{B}_1, B \in \mathcal{B}_k\}$. If $\mathcal{B}_{k+1} = \mathcal{B}_k$ for some k, we can conclude that $\mathcal{B}_k$ is a basis of $\mathcal{T} \cap \mathcal{S}_n$. Note that the dimension of $\mathcal{T} \cap \mathcal{S}_n$ is equal to $\sum_{j=1}^{\ell} \bar{n}_j(\bar{n}_j + 1)/2$, which is bounded by $n(n+1)/2$. It is mentioned here that $\mathcal{S}_n$ is a linear space equipped with an inner product $A \bullet B = \mathrm{tr}(AB)$ and the Gram–Schmidt orthogonalization procedure works.

Proposition 4.3. If a basis of $\mathcal{T} \cap \mathcal{S}_n$ is computed in advance, Algorithm 4.1 gives, with probability one, the decomposition of $\mathcal{T}$ into simple components. ∎

4.2 Decomposition into irreducible components

According to Theorem 3.1 (B), the block-diagonal matrices $\hat{Q}^\top A_p \hat{Q}$ obtained by Algorithm 4.1 can further be decomposed. By construction we have $\hat{Q} = Q\hat{\Pi}$ for some permutation matrix $\hat{\Pi}$. In the following we assume $\hat{Q} = Q$ to simplify presentation.

By Proposition 3.5 this finer decomposition can be obtained through a transformation of the form (3.7), which consists of "local coordinate changes" by a family of orthogonal matrices $\{P_1, \ldots, P_k\}$, followed by a permutation by Π.

The orthogonal matrices $\{P_1, \ldots, P_k\}$ should be chosen in such a way that if $i, i' \in K_j$, then

$$P_i^\top Q_i^\top A_p Q_{i'} P_{i'} = b_{ii'}^{(pj)} I_{\bar{m}_j} \tag{4.5}$$

for some $b_{ii'}^{(pj)} \in \mathbb{R}$ for $p = 1, \ldots, m$. Note that the solvability of this system of equations in P_i and $b_{ii'}^{(pj)}$ $(i, i' = 1, \ldots, k;\ j = 1, \ldots, \ell;\ p = 1, \ldots, m)$ is guaranteed by (4.1) and Proposition 3.5. Then with $\tilde{P} = Q \cdot \mathrm{diag}\,(P_1, \ldots, P_k)$ and $B_{pj} = (b_{ii'}^{(pj)} \mid i, i' \in K_j)$ we have

$$\tilde{P}^\top A_p \tilde{P} = \bigoplus_{j=1}^{\ell} (B_{pj} \otimes I_{\bar{m}_j}) \tag{4.6}$$

for $p = 1, \ldots, m$. Finally we apply a permutation of rows and columns to obtain (4.1).

Example 4.2. Recall Example 4.1. We consider the block-diagonalization of the first block $\hat{A}_p = Q[K_1]^\top A_p Q[K_1]$ of (4.4), where $m_1 = m_2 = m_3 = 2$ and $K_1 = \{1, 2, 3\}$. We first compute orthogonal matrices P_1, P_2 and P_3 satisfying

$$\mathrm{diag}\,(P_1, P_2, P_3)^\top \cdot \hat{A}_p \cdot \mathrm{diag}\,(P_1, P_2, P_3) = \begin{array}{|c|c|c|} \hline b_{11}^{(p1)} I_2 & b_{12}^{(p1)} I_2 & b_{13}^{(p1)} I_2 \\ \hline b_{21}^{(p1)} I_2 & b_{22}^{(p1)} I_2 & b_{23}^{(p1)} I_2 \\ \hline b_{31}^{(p1)} I_2 & b_{32}^{(p1)} I_2 & b_{33}^{(p1)} I_2 \\ \hline \end{array}\,.$$

Then a permutation of rows and columns yields a block-diagonal form $\mathrm{diag}\,(B_{p1}, B_{p1})$ with $B_{p1} = \begin{bmatrix} b_{11}^{(p1)} & b_{12}^{(p1)} & b_{13}^{(p1)} \\ b_{21}^{(p1)} & b_{22}^{(p1)} & b_{23}^{(p1)} \\ b_{31}^{(p1)} & b_{32}^{(p1)} & b_{33}^{(p1)} \end{bmatrix}$. ■

The family of orthogonal matrices $\{P_1, \ldots, P_k\}$ satisfying (4.5) can be computed as follows. Recall from (4.2) that for $i, i' \in K$ we have $i \sim i'$ if and only if $Q_i^\top A_p Q_{i'} \neq O$ for some p. It follows from (4.5) that $Q_i^\top A_p Q_{i'} \neq O$ means that it is nonsingular.

Fix j with $1 \leq j \leq \ell$. We consider a graph $G_j = (K_j, E_j)$ with vertex set K_j and edge set $E_j = \{(i, i') \mid i \sim i'\}$. This graph is connected by the definition of K_j. Let T_j be a spanning tree, which means that T_j is a subset of E_j such that $|T_j| = |K_j| - 1$ and any two vertices of K_j are connected by edges in T_j. With each $(i, i') \in T_j$ we can associate some $p = p(i, i')$ such that $Q_i^\top A_p Q_{i'} \neq O$.

To compute $\{P_i \mid i \in K_j\}$, take any $i_1 \in K_j$ and put $P_{i_1} = I_{\bar{m}_j}$. If $(i, i') \in T_j$ and P_i has been determined, then let $\hat{P}_{i'} = (Q_i^\top A_p Q_{i'})^{-1} P_i$ with

$p = p(i,i')$, and normalize it to $P_{i'} = \hat{P}_{i'}/\|q\|$, where q is the first-row vector of $\hat{P}_{i'}$. Then $P_{i'}$ is an orthogonal matrix that satisfies (4.5). By applying the above procedure in an appropriate order of $(i,i') \in T_j$ we can obtain $\{P_i \mid i \in K_j\}$.

Remark 4.1. A variant of the above algorithm for computing $\{P_1, \ldots, P_k\}$ is suggested here. Take a second random vector $r' = (r'_1, \ldots, r'_m)$, independently of r, to form $A(r') = r'_1 A_1 + \cdots + r'_m A_m$. For $i, i' \in K_j$ we have, with probability one, that $(i,i') \in E_j$ if and only if $Q_i^\top A(r') Q_{i'} \neq O$. If P_i has been determined, we can determine $P_{i'}$ by normalizing $\hat{P}_{i'} = (Q_i^\top A(r') Q_{i'})^{-1} P_i$ to $P_{i'} = \hat{P}_{i'}/\|q\|$, where q is the first-row vector of $\hat{P}_{i'}$. ■

Remark 4.2. The proposed method relies on numerical computations to determine block-diagonal structures. As such the method is inevitably faced with numerical noises due to rounding errors. A scaling technique to remedy this difficulty is suggested in Remark 5.1 for truss optimization problems. ■

Remark 4.3. The idea of using a random linear combination in constructing simultaneous block-diagonalization can also be found in a recent paper of de Klerk and Sotirov [4]. Their method, called "block diagonalization heuristic" in Section 5.2 of [4], is different from ours in two major points.

First, the method of [4] assumes explicit knowledge about the underlying group G, and works with the representation matrices, denoted $T(g)$ in (2.14). Through the eigenvalue (spectral) decomposition of a random linear combination of $T(g)$ over $g \in G$, the method finds an orthogonal matrix P such that $P^\top T(g) P$ for $g \in G$ are simultaneously block-diagonalized, just as in (2.15) and (2.16). Then G-symmetric matrices A_p, satisfying (2.14), will also be block-diagonalized.

Second, the method of [4] is not designed to produce the finest possible decomposition of the matrices A_p, as is recognized by the authors themselves. The constructed block-diagonalization of $T(g)$ is not necessarily the irreducible decomposition, and this is why the resulting decomposition of A_p is not guaranteed to be finest possible. We could, however, apply the algorithm of Section 4.2 of the present paper to obtain the irreducible decomposition of the representation $T(g)$. Then the resulting decomposition of A_p will be the finest decomposition that can be obtained by exploiting the G-symmetry (and nothing else). ■

4.3 A practical variant of the algorithm

In Propositions 3.4 we have considered two technical conditions:

1. $\mathrm{span}\{I_n, A_1, \ldots, A_m\} = \mathcal{T} \cap \mathcal{S}_n$,
2. $r \in R$, where R is an open dense set,

to ensure genericity of $A = \sum_{p=1}^{m} r_p A_p$ in eigenvalue structure. The genericity of A guarantees, in turn, that our algorithm yields the finest simultaneous block-diagonalization (see Proposition 4.2). The condition $r \in R$ above can be met with probability one through a random choice of r.

To meet the first condition we could generate a basis of $\mathcal{T} \cap \mathcal{S}_n$ in advance, as is mentioned in Proposition 4.3. However, an explicit computation of a basis seems too heavy to be efficient. It should be understood that the above two conditions are introduced as sufficient conditions to avoid degeneracy in eigenvalues. By no means are they necessary for the success of the algorithm. With this observation we propose the following procedure as a practical variant of our algorithm.

We apply Algorithm 4.1 to the given family $\{A_1, \ldots, A_m\}$ to find an orthogonal matrix Q and a partition $K = K_1 \cup \cdots \cup K_\ell$. In general there is no guarantee that this corresponds to the decomposition into simple components, but in any case Algorithm 4.1 terminates without getting stuck. The algorithm does not hang up either, when a particular choice of r does not meet the condition $r \in R$. Thus we can always go on to the second stage of the algorithm for the irreducible decomposition.

Next, we are to determine a family of orthogonal matrices $\{P_1, \ldots, P_k\}$ that satisfies (4.5). This system of equations is guaranteed to be solvable if A is generic (see Proposition 3.5). In general we may possibly encounter a difficulty of the following kinds:

1. For some $(i, i') \in T_j$ the matrix $Q_i^\top A_p Q_{i'}$ woth $p = p(i, i')$ is not nonsingular and hence $P_{i'}$ cannot be computed. This includes the case of a rectangular matrix, which is demonstrated in Example 4.3 below.

2. For some p and $(i, i') \in E_j$ the matrix $P_i^\top Q_i^\top A_p Q_{i'} P_{i'}$ is not a scalar multiple of an identity matrix.

Such inconsistency is an indication that the decomposition into simple components has not been computed correctly. Accordingly, if either of the above inconsistency is detected, we restart our algorithm by adding some linearly independent matrices of $\mathcal{T} \cap \mathcal{S}_n$ to the current set $\{A_1, \ldots, A_m\}$. It is mentioned that the possibility exists, though with probability zero, that r is chosen badly to yield a nongeneric A even when $\mathrm{span}\{I_n, A_1, \ldots, A_m\} = \mathcal{T} \cap \mathcal{S}_n$ is true.

It is expected that we can eventually arrive at the correct decomposition after a finite number of iterations. With probability one, the number of restarts is bounded by the dimension of $\mathcal{T} \cap \mathcal{S}_n$, which is $O(n^2)$. When it terminates, the modified algorithm always gives a legitimate simultaneous block-diagonal decomposition of the form (4.1).

There is some subtlety concerning the optimality of the obtained decomposition. If a basis of $\mathcal{T} \cap \mathcal{S}_n$ is generated, the decomposition coincides, with probability one, with the canonical finest decomposition of the $*$-algebra $\mathcal{T}$. However, when the algorithm terminates before it generates a basis of $\mathcal{T} \cap \mathcal{S}_n$, there is no theoretical guarantee that the obtained decomposition is the finest possible. Nevertheless, it is very likely in practice that the obtained decomposition coincides with the finest decomposition.

Example 4.3. Here is an example that requires an additional generator to be added. Suppose that we are given

$$A_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 1 & -1 & 0 \end{bmatrix}, \qquad A_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

and let $\mathcal{T}$ be the matrix $*$-algebra generated by $\{I_4, A_1, A_2\}$. It turns out that the structural indices in (4.1) are: $\ell = 2$, $\bar{m}_1 = \bar{m}_2 = 1$, $\bar{n}_1 = 1$ and $\bar{n}_2 = 3$. This means that the list of eigenvalue multiplicities of $\mathcal{T}$ is $\{1, 1, 1, 1\}$. Note also that $\dim(\mathcal{T} \cap \mathcal{S}_4) = \bar{n}_1(\bar{n}_1 + 1)/2 + \bar{n}_2(\bar{n}_2 + 1)/2 = 7$.

For $A(r) = r_1 A_1 + r_2 A_2$ we have

$$A(r) \begin{bmatrix} 1 & 0 \\ 0 & (r_1 - r_2)/c \\ 0 & r_1/c \\ 0 & 0 \end{bmatrix} = (r_1 + r_2) \begin{bmatrix} 1 & 0 \\ 0 & (r_1 - r_2)/c \\ 0 & r_1/c \\ 0 & 0 \end{bmatrix} \tag{4.7}$$

with $c = \sqrt{(r_1 - r_2)^2 + {r_1}^2}$. This shows that $A(r)$ has a multiple eigenvalue $r_1 + r_2$ of multiplicity two, as well as two other simple eigenvalues. Thus for any r the list of eigenvalue multiplicities of $A(r)$ is equal to $\{2, 1, 1\}$, which differs from $\{1, 1, 1, 1\}$ for $\mathcal{T}$.

The discrepancy in the eigenvalue multiplicities cannot be detected during the first stage of our algorithm, in which we will obtain the following. In Step 2 we have $k = 3$, $m_1 = 2$, $m_2 = m_3 = 1$. The orthogonal matrix Q is partitioned into three submatrices Q_1, Q_2 and Q_3, where Q_1 (nonunique) may possibly be the 4×2 matrix shown in (4.7), and Q_2 and Q_3 consist of a single column. Since $Q^\top A_p Q$ is of the form

$$Q^\top A_p Q = \left[\begin{array}{cc|c|c} 1 & 0 & 0 & 0 \\ 0 & * & * & * \\ \hline 0 & * & * & * \\ \hline 0 & * & * & * \end{array}\right]$$

for $p = 1, 2$, we have $\ell = 1$ and $K_1 = \{1, 2, 3\}$ in Step 3. At this moment an inconsistency is detected, since $m_1 \neq m_2$ inspite of the fact that $i = 1$ and $i' = 2$ belong to the same block K_1.

We restart the algorithm, say, with an additional generator

$$A_3 = \frac{1}{2}(A_1 A_2 + A_2 A_1) = \frac{1}{2} \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -2 \end{bmatrix}$$

to consider $\tilde{A}(r) = r_1 A_1 + r_2 A_2 + r_3 A_3$ instead of $A(r) = r_1 A_1 + r_2 A_2$. Then $\tilde{A}(r)$ has four simple eigenvalues for generic values of $r = (r_1, r_2, r_3)$, and accordingly we have $\{1, 1, 1, 1\}$ as the list of eigenvalue multiplicities of $\tilde{A}(r)$, which agrees with that of $\mathcal{T}$.

In Step 2 of Algorithm 4.1 we now have $k = 4$, $m_1 = m_2 = m_3 = m_4 = 1$. The orthogonal matrix Q is partitioned into four 4×1 submatrices, and $Q^\top A_p Q$ is of the form

$$Q^\top A_p Q = \left[\begin{array}{c|c|c|c} 1 & 0 & 0 & 0 \\ \hline 0 & * & * & * \\ \hline 0 & * & * & * \\ \hline 0 & * & * & * \end{array}\right]$$

for $p = 1, 2, 3$, from which we obtain $K_1 = \{1\}$, $K_2 = \{2, 3, 4\}$ with $\ell = 2$ in Step 3. Thus we have arrived at the correct decomposition consisting of a 1×1 block and a 3×3 block. Note that the correct decomposition is obtained in spite of the fact that $\{I_4, A_1, A_2, A_3\}$ does not span $\mathcal{T} \cap \mathcal{S}_4$. ■

5 Numerical Examples

5.1 Optimization of symmetric trusses

Use and significance of our method are illustrated here in the context of semidefinite programming for truss design treated in [11]. Group-symmetry and sparsity arise naturally in truss optimization problems, as is easily imagined from the cubic truss shown in Fig.1. It will be confirmed that the proposed method yields the same decomposition as the group representation theory anticipates (Case 1 below), and moreover, it gives a finer decomposition if the truss structure is endowed with an additional algebraic structure due to sparsity (Case 2 below).

The optimization problem we consider here is as follow. An initial truss configuration is given with fixed locations of nodes and members. Optimal cross-sectional areas, minimizing total volume of the structure, are to be found subject to the constraint that the eigenvalues of vibration are not smaller than a specified value.

To be more specific, let n^{d} and n^{m} denote the number of degrees of freedom of displacements and the number of members of a truss, respectively. The stiffness matrix is denoted by $K \in \mathcal{S}_{n^{\mathrm{d}}}$. Let $M_{\mathrm{S}} \in \mathcal{S}_{n^{\mathrm{d}}}$ and

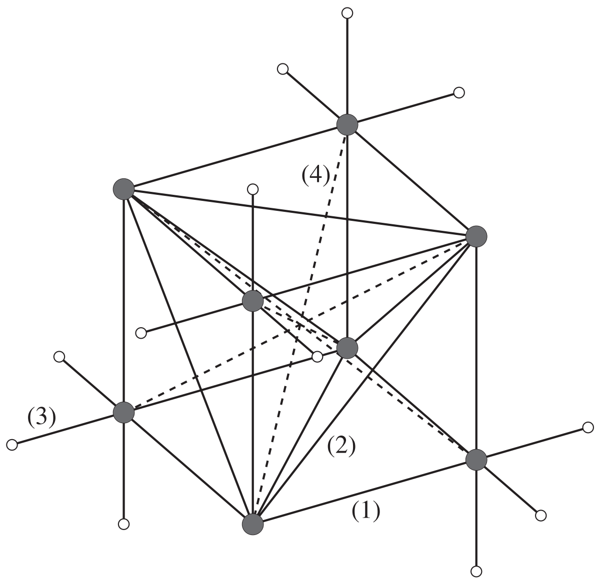

Figure 1: A cubic (or $T_{\rm d}$-symmetric) space truss.

$M_0 \in \mathcal{S}_{n^{\rm d}}$ denote the mass matrices for the structural and nonstructural masses, respectively. The ith eigenvalue Ω_i of vibration and the corresponding eigenvector $\phi_i \in \mathbb{R}^{n^{\rm d}}$ are defined by

$$K\phi_i = \Omega_i(M_{\rm S} + M_0)\phi_i, \quad i = 1, 2, \dots, n^{\rm d}. \tag{5.1}$$

Note that, for a truss, K and $M_{\rm S}$ can be written as

$$K = \sum_{j=1}^{n^{\rm m}} K_j \eta_j, \qquad M_{\rm S} = \sum_{j=1}^{n^{\rm m}} M_j \eta_j,$$

with constant symmetric matrices K_j and M_j, where η_j denotes the cross-sectional area of the jth member. With the notation $l = (l_j) \in \mathbb{R}^{n^{\rm m}}$ for the vector of member lengths and $\bar{\Omega}$ for the specified lower bound of the fundamental eigenvalue, our optimization problem is formulated as

$$\left.\begin{array}{ll} \min & \displaystyle\sum_{j=1}^{n^{\rm m}} l_j \eta_j \\ \text{s.t.} & \Omega_i \geq \bar{\Omega}, \quad i = 1, \dots, n^{\rm d}, \\ & \eta_j \geq 0, \quad j = 1, \dots, n^{\rm m}. \end{array}\right\} \tag{5.2}$$

It is pointed out in [11] that this problem (5.2) can be reduced to the following dual SDP (cf. (2.2)):

$$\left.\begin{array}{ll} \max & -\sum_{j=1}^{n^{\rm m}} l_j \eta_j \\ \mathrm{s.t.} & \sum_{j=1}^{n^{\rm m}} (K_j - \bar{\Omega} M_j)\eta_j - \bar{\Omega} M_0 \succeq O, \\ & \eta_j \geq 0, \quad j = 1, \dots, n^{\rm m}. \end{array}\right\} \tag{5.3}$$

We now consider this SDP for the cubic truss shown in Fig.1. The cubic truss contains 8 free nodes, and hence $n^{\rm d} = 24$. As for the members we consider two cases:

Case 1: $n^{\rm m} = 34$ members including the dotted ones;
Case 2: $n^{\rm m} = 30$ members excluding the dotted ones.

A regular tetrahedron is constructed inside the cube. The lengths of members forming the edges of the cube are 200.0 cm. The lengths of the members outside the cube are 100.0 cm. The same nonstructural mass of 2.1×10^5 kg is located at each node indicated by a filled circle in Fig.1. The lower bound of the eigenvalues is specified as $\bar{\Omega} = 10.0$. All the remaining nodes are pin-supported.

Thus, the geometry, the stiffness distribution, and the mass distribution of this truss are all symmetric with respect to the geometric transformations by elements of tetrahedral group $\mathrm{T_d}$. The $\mathrm{T_d}$-symmetry can be exploited as follows.

First, we divide the index set of members $\{1, \dots, n^{\rm m}\}$ into a family of orbits, say J_p with $p = 1, \dots, m$, where m denotes the number of orbits. We have $m = 4$ in Case 1 and $m = 3$ in Case 2, where representative members belonging to four different orbits are shown as (1)–(4) in Fig.1. It is mentioned in passing that the classification of members into orbits is an easy task for engineers, who may or may not be versed in group representation theory. Indeed, this is nothing but the so-called *variable-linking technique*, which has often been employed in the literature of structural optimization in obtaining symmetric structural designs [9].

Next, with reference to the orbits we aggregate the data matrices as well as the components of vector b in (5.3) to A_p ($p = 0, 1, \dots, m$) and b_p ($p = 1, \dots, m$), respectively, as

$$A_0 = -\bar{\Omega} M_0,$$
$$A_p = \sum_{j \in J_p} (-K_j + \bar{\Omega} M_j), \quad p = 1, \dots, m,$$
$$b_p = \sum_{j \in J_p} l_j, \quad p = 1, \dots, m.$$

Then (5.3) can be reduced to

$$\left.\begin{array}{ll} \max & -\sum_{p=1}^{m} b_p y_p \\ \text{s.t.} & A_0 - \sum_{p=1}^{m} A_p y_p \succeq O, \\ & y_p \geq 0, \quad p = 1, \ldots, m \end{array}\right\} \tag{5.4}$$

as long as we are interested in a symmetric optimal solution [1]. Note that the matrices A_p $(p = 0, 1, \ldots, m)$ are symmetric in the sense of (2.14) for $G = \mathrm{T_d}$. Numerical values of A_p $(p = 1, \ldots, 4)$ are given in Section A. Note that the two cases share the same matrices A_1, A_2, A_3, and A_0 is proportional to the identity matrix.

The proposed method is applied to A_p $(p = 0, 1, \ldots, m)$ for their simultaneous block-diagonalization. The practical variant described in Section 4.3 is employed. In either case it has turned out that additional generators are not necessary, but the random linear combinations of the given matrices A_p $(p = 0, 1, \ldots, m)$ are sufficient to find the block-diagonalization.

In Case 1 we obtain the decomposition into $1 + 2 + 3 + 3 = 9$ blocks, one block of size 2, two identical blocks of size 2, three identical blocks of size 3, and three identical blocks of size 4, as summarized in the left of Table 1. This result conforms with the group-theoretic analysis. The tetrahedral group $\mathrm{T_d}$ has two one-dimensional irreducible representations, one two-dimensional irreducible representation, and two three-dimensional irreducible representations. The block indexed by $j = 1$ corresponds to the unit representation, one of the one-dimensional irreducible representations, while the block for the other one-dimensional irreducible representation is missing. The block with $j = 2$ corresponds to the two-dimensional irreducible representation, hence $\bar{m}_2 = 2$. Similarly, the blocks with $j = 3, 4$ correspond to the three-dimensional irreducible representation, hence $\bar{m}_3 = \bar{m}_4 = 3$.

In Case 2 sparsity plays a role to split the last block into two, as shown in the right of Table 1. We now have 12 blocks in contrast to 9 blocks in Case 1. Recall that the sparsity is due to the lack of the dotted members. It is emphasized that the proposed method successfully captures the additional algebraic structure introduced by sparsity.

Remark 5.1. Typically, actual trusses are constructed by using steel members, where the elastic modulus and the mass density of members are $E = 200.0$ GPa and $\rho = 7.86 \times 10^{-3}$ kg/cm^2, respectively. Note that the matrices K_j and M_j defining the SDP problem (5.4) are proportional to E and ρ, respectively. In order to avoid numerical instability in our block-diagonalization algorithm, E and ρ are scaled as $E = 1.0 \times 10^{-2}$ GPa and $\rho = 100.0$ kg/cm^2, so that the largest eigenvalue in (5.1) becomes sufficiently small. Note that the transformation matrix obtained by our algorithm for

Table 1: Block-diagonalization of cubic truss optimization problem.

	Case 1: $m = 4$ block size $\bar{n}_j$	multiplicity $\bar{m}_j$	Case 2: $m = 3$ block size $\bar{n}_j$	multiplicity $\bar{m}_j$
$j = 1$	2	1	2	1
$j = 2$	2	2	2	2
$j = 3$	2	3	2	3
$j = 4$	4	3	2	3
$j = 5$	—	—	2	3

block-diagonalization of $A_0, A_1, \dots, A_m$ is independent of the values of E and ρ. Hence, it is recommended for numerical stability to compute transformation matrices for the scaled matrices $\tilde{A}_0, \tilde{A}_1, \dots, \tilde{A}_m$ by choosing appropriate fictitious values of E and ρ. Then the obtained transformation matrices can be used to decompose the original matrices $A_0, A_1, \dots, A_m$ defined with the actual material parameters. ■

5.2 Effects of additional algebraic structures

It is demonstrated here that our method automatically reveals inherent algebraic structures due to parameter dependence as well as to group symmetry. The S_3-symmetric matrices $A_1, \dots, A_4$ in (2.7) and (2.8) are considered in three representative cases.

Case 1:

$$B = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad D = \begin{bmatrix} 1 \end{bmatrix}, \quad E = \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix},$$

Case 2:

$$B = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad D = \begin{bmatrix} 1 \end{bmatrix}, \quad E = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix},$$

Case 3:

$$B = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad D = \begin{bmatrix} 1 \end{bmatrix}, \quad E = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}.$$

We have $n^{\mathrm{B}} = 2$ and $n^{\mathrm{D}} = 1$ in the notation of Section 2.2.

Case 1 is a generic case under S_3-symmetry. The simultaneous block-diagonalization is of the form

$$P^\top A_p P = B_{p1} \oplus (I_2 \otimes B_{p2}), \quad p = 1, \dots, 4, \tag{5.5}$$

with $B_{p1} \in \mathcal{M}_3$, $B_{p2} \in \mathcal{M}_2$; i.e., $\ell = 2$, $\bar{m}_1 = 1$, $\bar{m}_2 = 2$, $\bar{n}_1 = 3$, $\bar{n}_2 = 2$ in (4.1). Our implementation of the proposed method yields

$$B_{11} = \begin{bmatrix} -0.99648 & -0.07327 & -0.06501 \\ -0.07327 & 0.54451 & 1.15740 \\ -0.06501 & 1.15740 & 2.45197 \end{bmatrix},$$
$$B_{21} = \begin{bmatrix} -0.10068 & 1.23826 & -0.43048 \\ 1.23826 & -2.73454 & -2.34910 \\ -0.43048 & -2.34910 & 2.83521 \end{bmatrix},$$
$$B_{31} = \begin{bmatrix} 0.00148 & -0.03481 & 0.01647 \\ -0.03481 & 0.81586 & -0.38603 \\ 0.01647 & -0.38603 & 0.18266 \end{bmatrix},$$
$$B_{41} = \begin{bmatrix} 2.95256 & 0.49133 & 0.77219 \\ 0.49133 & 1.31903 & 2.74340 \\ 0.77219 & 2.74340 & 5.72841 \end{bmatrix},$$

and $B_{22} = B_{32} = O$,

$$B_{12} = \begin{bmatrix} -0.99954 & -0.04297 \\ -0.04297 & 2.99954 \end{bmatrix}, \quad B_{42} = \begin{bmatrix} -1.51097 & 0.52137 \\ 0.52137 & -3.48903 \end{bmatrix}.$$

Those matrices are of the same form as (2.10)–(2.13), but have different numerical values, which is not surprising. It can be verified, for example, that

$$\begin{bmatrix} B_{12} & O \\ O & B_{12} \end{bmatrix} = \tilde{P}^\top \begin{bmatrix} B & O \\ O & B \end{bmatrix} \tilde{P}, \quad \begin{bmatrix} B_{42} & O \\ O & B_{42} \end{bmatrix} = \tilde{P}^\top \begin{bmatrix} -E & O \\ O & -E \end{bmatrix} \tilde{P}$$

for an orthogonal matrix $\tilde{P}$ expressed as $\tilde{P} = \begin{bmatrix} \tilde{P}_{11} & \tilde{P}_{12} \\ \tilde{P}_{21} & \tilde{P}_{22} \end{bmatrix}$ with

$$\tilde{P}_{11} = -\tilde{P}_{22} = \begin{bmatrix} 0.12554 & -0.12288 \\ -0.12288 & -0.12554 \end{bmatrix}, \quad \tilde{P}_{12} = \tilde{P}_{21} = \begin{bmatrix} 0.70355 & -0.68859 \\ -0.68859 & -0.70355 \end{bmatrix}.$$

In Case 2 we have a commutativity relation $BE = EB$. This means that B and E can be simultaneously put into a diagonal form, which leads to a further decomposition of the second factor in (5.5). Thus, instead of (5.5) we have

$$P^\top A_p P = B_{p1} \oplus (I_2 \otimes B_{p2}) \oplus (I_2 \otimes B_{p3}), \quad p = 1, \ldots, 4,$$

with $B_{p1} \in \mathcal{M}_3$, $B_{p2} \in \mathcal{M}_1$ and $B_{p3} \in \mathcal{M}_1$; i.e., $\ell = 3$, $\bar{m}_1 = 1$, $\bar{m}_2 = \bar{m}_3 =$

2, $\bar{n}_1 = 3$, $\bar{n}_2 = \bar{n}_3 = 1$ in (4.1). We have obtained

$$B_{11} = \begin{bmatrix} -0.99856 & -0.03968 & 0.00243 \\ -0.03968 & 0.48487 & 1.10589 \\ 0.00243 & 1.10589 & 2.51369 \end{bmatrix},$$

$$B_{21} = \begin{bmatrix} -0.09232 & 1.18464 & -0.37146 \\ 1.18464 & -2.62615 & -2.51217 \\ -0.37146 & -2.51217 & 2.71847 \end{bmatrix},$$

$$B_{31} = \begin{bmatrix} 0.00139 & -0.03415 & 0.01502 \\ -0.03415 & 0.83669 & -0.36807 \\ 0.01502 & -0.36807 & 0.16192 \end{bmatrix},$$

$$B_{41} = \begin{bmatrix} 3.99447 & 0.13106 & -0.07267 \\ 0.13106 & 1.30143 & 2.94623 \\ -0.07267 & 2.94623 & 6.70410 \end{bmatrix},$$

$B_{12} = [\,3.00000\,]$, $B_{42} = [\,-4.00000\,]$, $B_{13} = [\,-1.00000\,]$, $B_{43} = [\,-2.00000\,]$, and $B_{pj} = O$ for $p = 2, 3; j = 2, 3$. Thus the proposed method succeeds in finding the additional algebraic structure caused by $BE = EB$.

Case 3 contains a further degeneracy that the column vector of C is an eigenvector of B and E. This splits the 3×3 block into two, and we have

$$P^\top A_p P = B_{p1} \oplus B_{p4} \oplus (I_2 \otimes B_{p2}) \oplus (I_2 \otimes B_{p3}), \quad p = 1, \ldots, 4,$$

with $B_{p1} \in \mathcal{M}_2$, $B_{pj} \in \mathcal{M}_1$ for $j = 2, 3, 4$; i.e., $\ell = 4$, $\bar{m}_1 = \bar{m}_4 = 1$, $\bar{m}_2 = \bar{m}_3 = 2$, $\bar{n}_1 = 2$, $\bar{n}_2 = \bar{n}_3 = \bar{n}_4 = 1$ in (4.1). We have obtained

$$B_{11} = \begin{bmatrix} 0.48288 & 1.10248 \\ 1.10248 & 2.51712 \end{bmatrix}, \quad B_{21} = \begin{bmatrix} -1.80034 & -1.66096 \\ -1.66096 & 1.80034 \end{bmatrix},$$

$$B_{31} = \begin{bmatrix} 0.83904 & -0.36749 \\ -0.36749 & 0.16096 \end{bmatrix}, \quad B_{41} = \begin{bmatrix} 1.28767 & 2.93994 \\ 2.93994 & 6.71233 \end{bmatrix},$$

$B_{12} = [\,3.00000\,]$, $B_{42} = [\,-4.00000\,]$, $B_{13} = [\,-1.00000\,]$, $B_{43} = [\,-2.00000\,]$, $B_{14} = [\,-1.00000\,]$, $B_{44} = [\,4.00000\,]$, and $B_{pj} = O$ for $p = 2, 3, 4; j = 2, 3$. Also in this case the proposed method works, identifying the additional algebraic structure through numerical computation.

The three cases are compared in Table 2.

Acknowledgments

The authors thank Etienne de Klerk and Dima Pasechnik for discussion. This work is supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

[1] Y. Bai, E. de Klerk, D. V. Pasechnik and R. Sotirov: Exploiting group symmetry in truss topology optimization, CentER Discussion paper

Table 2: Block-diagonalization of S_3-symmetric matrices in (2.7) and (2.8).

	Case 1		Case 2		Case 3	
	$\bar{n}_j$	$\bar{m}_j$	$\bar{n}_j$	$\bar{m}_j$	$\bar{n}_j$	$\bar{m}_j$
$j = 1$	3	1	3	1	2	1
$j = 4$	—	—	—	—	1	1
$j = 2$	2	2	1	2	1	2
$j = 3$	—	—	1	2	1	2

2007-17, Center for Economic Research, Tilburg University, February 2007.

[2] B. Borchers: CSDP 2.3 user's guide, a C library for semidefinite programming, Optimization Methods and Software, Vol. 11 & 12 (1999), pp. 597–611.

[3] E. de Klerk, D.V. Pasechnik and A. Schrijver: Reduction of symmetric semidefinite programs using the regular $*$-representation, Mathematical Programming, Series B, Vol. 109 (2007), pp. 613–624.

[4] E. de Klerk and R. Sotirov: Exploiting group symmetry in semidefinite relaxations of the quadratic assignment, CentER Discussion paper 2007-44, Center for Economic Research, Tilburg University, June 2007.

[5] K. Gatermann and P.A. Parrilo: Symmetry groups, semidefinite programs, and sums of squares, Journal of Pure and Applied Algebra, Vol. 192 (2004), pp. 95–128.

[6] L. Jansson, J.B. Lasserre, C. Riener and T. Theobald: Exploiting symmetries in SDP-relaxations for polynomial optimization, LAAS-report, Toulouse, September 2006.

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

[8] M. Kojima, S. Kojima and S. Hara: Linear algebra for semidefinite programming, Research Report B-290, October 1994; also in RIMS Kokyuroku 1004, Kyoto University, pp. 1–23, 1997.

[9] L. Lamberti and C. Pappalettere: An efficient sequential linear programming algorithm for engineering optimization, Journal of Engineering Design, Vol. 16 (2005), pp. 353–371.

[10] W. Miller, Jr.: Symmetry Groups and Their Applications, Academic Press, New York, 1972.

[11] M. Ohsaki, K. Fujisawa, N. Katoh and Y. Kanno: Semi-definite programming for topology optimization of trusses under multiple eigenvalue constraints, Computer Methods in Applied Mechanics and Engineering, Vol. 180 (1999), pp. 203-217.

[12] J.-P. Serre: Linear Representations of Finite Groups, Springer-Verlag, Berlin, 1977.

[13] J.F. Sturm: Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones, Optimization Methods and Software, Vol. 11 & 12 (1999), pp. 625–653.

[14] K.C. Toh, M.J. Todd and R.H. Tütüncü: SDPT3 — a MATLAB software package for semidefinite programming, version 1.3, Optimization Methods and Software, Vol. 11 & 12 (1999), pp. 545–581.

[15] J.H.M. Wedderburn: Lectures on Matrices, American Mathematical Society, New York, 1934; Dover, Mineola, N.Y., 2005.

[16] H. Wolkowicz, R. Saigal and L. Vandenberghe: Handbook of Semidefinite Programming: Theory, Algorithms, and Applications, Kluwer Academic Publishers, Boston, 2000.

[17] M. Yamashita, K. Fujisawa and M. Kojima: Implementation and evaluation of SDPA6.0 (SemiDefinite Programming Algorithm 6.0), Optimization Methods and Software, Vol. 18 (2003), pp. 491–505.

A Matrices of the example in Section 5.1

For the example of the $\mathrm{T_d}$-symmetric truss in Section 5.1, we show the numerical values of matrices $A_p \in \mathcal{S}_{24}$ $(p = 1, \dots, 4)$. For ease of presentation we suppose that A_p consists of 6×6 matrices $A_{ij}^{(p)}$ $(i, j = 1, \dots, 4)$ as

$$A_p = \begin{bmatrix} A_{11}^{(p)} & A_{12}^{(p)} & A_{13}^{(p)} & A_{14}^{(p)} \\ & A_{22}^{(p)} & A_{23}^{(p)} & A_{24}^{(p)} \\ & & A_{33}^{(p)} & A_{34}^{(p)} \\ & & & A_{44}^{(p)} \end{bmatrix},$$

where the lower-left blocks are omitted by symmetry. The matrices$A_{ij}^{(p)}$ $(p = 1, \dots, 4; i, j = 1, \dots, 4)$ are given as

$$A_{11}^{(1)} = A_{22}^{(1)} = A_{33}^{(1)} = A_{44}^{(1)} = \begin{bmatrix} 199.5 & 0 & 0 & 33.833 & 0 & 0 \\ & 199.5 & 0 & 0 & 33.333 & 0 \\ & & 199.5 & 0 & 0 & 33.333 \\ & & & 199.5 & 0 & 0 \\ & & & & 199.5 & 0 \\ & & & & & 199.5 \end{bmatrix},$$

$$A_{12}^{(1)} = A_{34}^{(1)} = \begin{bmatrix} 0 & 0 & 0 & 33.333 & 0 & 0 \\ 0 & 0 & 0 & 0 & 33.708 & 0.21651 \\ 0 & 0 & 0 & 0 & 0.21651 & 33.458 \\ 33.333 & 0 & 0 & 0 & 0 & 0 \\ 0 & 33.708 & 0.21651 & 0 & 0 & 0 \\ 0 & 0.21651 & 33.458 & 0 & 0 & 0 \end{bmatrix},$$

$$A_{13}^{(1)} = A_{24}^{(1)} = \begin{bmatrix} 33.333 & 0 & 0 & 0 & 0 & 0 \\ 0 & 33.458 & -0.21651 & 0 & 0 & 0 \\ 0 & -0.21651 & 33.708 & 0 & 0 & 0 \\ 0 & 0 & 0 & 33.333 & 0 & 0 \\ 0 & 0 & 0 & 0 & 33.458 & -0.21651 \\ 0 & 0 & 0 & 0 & -0.21651 & 33.708 \end{bmatrix},$$

$A_{14}^{(1)} = A_{23}^{(1)} = O,$

$$A_{11}^{(2)} = \begin{bmatrix} 282.49 & -0.064705 & -0.24148 & 0 & 0 & 0 \\ & 282.64 & -0.088388 & 0 & 0 & 0 \\ & & 282.34 & 0 & 0 & 0 \\ & & & 0 & 0 & 0 \\ & & & & 0 & 0 \\ & & & & & 0 \end{bmatrix},$$

$$A_{12}^{(2)} = \begin{bmatrix} 47.317 & 0.15309 & 0.088388 & 0 & 0 & 0 \\ 0.15309 & 47.273 & 0.076547 & 0 & 0 & 0 \\ 0.088388 & 0.076547 & 47.185 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$A_{13}^{(2)} \begin{bmatrix} 0 & 0 & 0 & 47.317 & -0.088388 & 0.15309 \\ 0 & 0 & 0 & -0.088388 & 47.185 & -0.076547 \\ 0 & 0 & 0 & 0.15309 & -0.076547 & 47.273 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$A_{14}^{(2)} = \begin{bmatrix} 0 & 0 & 0 & 47.14 & 0 & 0 \\ 0 & 0 & 0 & 0 & 47.164 & 0.088388 \\ 0 & 0 & 0 & 0 & 0.088388 & 47.47 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$A_{22}^{(2)} = \begin{bmatrix} 282.49 & -0.24148 & 0.064705 & 0 & 0 & 0 \\ & 282.34 & 0.088388 & 0 & 0 & 0 \\ & & 282.64 & 0 & 0 & 0 \\ & & & 0 & 0 & 0 \\ & & & & 0 & 0 \\ & & & & & 0 \end{bmatrix},$$

$$A_{23}^{(2)} = \begin{bmatrix} 0 & 0 & 0 & 47.14 & 0 & 0 \\ 0 & 0 & 0 & 0 & 47.47 & -0.088388 \\ 0 & 0 & 0 & 0 & -0.088388 & 47.164 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$A_{24}^{(2)} = \begin{bmatrix} 0 & 0 & 0 & 47.317 & 0.088388 & -0.15309 \\ 0 & 0 & 0 & 0.088388 & 47.185 & -0.076547 \\ 0 & 0 & 0 & -0.15309 & -0.076547 & 47.273 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$A_{33}^{(2)} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ & 0 & 0 & 0 & 0 & 0 \\ & & 0 & 0 & 0 & 0 \\ & & & 282.49 & 0.24148 & -0.064705 \\ & & & & 282.34 & 0.088388 \\ & & & & & 282.64 \end{bmatrix},$$

$$A_{34}^{(2)} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 47.317 & -0.15309 & -0.088388 \\ 0 & 0 & 0 & -0.15309 & 47.273 & 0.076547 \\ 0 & 0 & 0 & -0.088388 & 0.076547 & 47.185 \end{bmatrix},$$

$$A_{44}^{(2)} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ & 0 & 0 & 0 & 0 & 0 \\ & & 0 & 0 & 0 & 0 \\ & & & 282.49 & 0.064705 & 0.24148 \\ & & & & 282.64 & -0.088388 \\ & & & & & 282.34 \end{bmatrix},$$

$$A_{11}^{(3)} = A_{22}^{(3)} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ & 0 & 0 & 0 & 0 & 0 \\ & & 0 & 0 & 0 & 0 \\ & & & 99.0 & 0 & 0 \\ & & & & 99.0 & 0 \\ & & & & & 99.0 \end{bmatrix},$$

$$A_{12}^{(3)} = A_{13}^{(3)} = A_{14}^{(3)} = A_{23}^{(3)} = A_{24}^{(3)} = A_{34}^{(3)} = O,$$

$$A_{33}^{(3)} = A_{44}^{(3)} = \begin{bmatrix} 99.0 & 0 & 0 & 0 & 0 & 0 \\ & 99.0 & 0 & 0 & 0 & 0 \\ & & 99.0 & 0 & 0 & 0 \\ & & & 0 & 0 & 0 \\ & & & & 0 & 0 \\ & & & & & 0 \end{bmatrix},$$

$$A_{11}^{(4)} = A_{44}^{(4)} = \begin{bmatrix} 115.37 & -0.035221 & -0.13145 & 0 & 0 & 0 \\ & 115.46 & -0.048113 & 0 & 0 & 0 \\ & & 115.29 & 0 & 0 & 0 \\ & & & 115.37 & 0.035221 & 0.13145 \\ & & & & 115.46 & -0.048113 \\ & & & & & 115.29 \end{bmatrix},$$

$$A_{12}^{(4)} = A_{13}^{(4)} = A_{24}^{(4)} = A_{34}^{(4)} = O,$$

$$A_{14}^{(4)} = \begin{bmatrix} 57.831 & 0.035221 & 0.13145 & 0 & 0 & 0 \\ 0.035221 & 57.748 & 0.048113 & 0 & 0 & 0 \\ 0.13145 & 0.048113 & 57.915 & 0 & 0 & 0 \\ 0 & 0 & 0 & 57.831 & -0.035221 & -0.13145 \\ 0 & 0 & 0 & -0.035221 & 57.748 & 0.048113 \\ 0 & 0 & 0 & -0.13145 & 0.048113 & 57.915 \end{bmatrix},$$

$$A_{22}^{(4)} = A_{33}^{(4)} = \begin{bmatrix} 115.37 & -0.13145 & 0.035221 & 0 & 0 & 0 \\ & 115.29 & 0.048113 & 0 & 0 & 0 \\ & & 115.46 & 0 & 0 & 0 \\ & & & 115.37 & 0.13145 & -0.035221 \\ & & & & 115.29 & 0.048113 \\ & & & & & 115.46 \end{bmatrix},$$

$$A_{23}^{(4)} = \begin{bmatrix} 57.831 & 0.13145 & -0.035221 & 0 & 0 & 0 \\ 0.13145 & 57.915 & -0.048113 & 0 & 0 & 0 \\ -0.035221 & -0.048113 & 57.748 & 0 & 0 & 0 \\ 0 & 0 & 0 & 57.831 & -0.13145 & 0.035221 \\ 0 & 0 & 0 & -0.13145 & 57.915 & -0.048113 \\ 0 & 0 & 0 & 0.035221 & -0.048113 & 57.748 \end{bmatrix}.$$