数理解析研究所講究録
第 1629 巻 2009 年 15-26

An Algorithm for Decomposition of Matrix $*$-Algebras Generated by Symmetric Matrices

東京大学 数理情報学専攻 室田 一雄 (Kazuo Murota)
東京大学 数理情報学専攻 寒野 善博 (Yoshihiro Kanno)
東京工業大学 数理・計算科学専攻 小島 政和 (Masakazu Kojima)
東京工業大学 数理・計算科学専攻 小島 定吉 (Sadayoshi Kojima)

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, 6, 7, 8]. A common and essential problem in these studies can be stated as follows: Given a finite set of $n\times n$ real symmetric matrices $A_1, \ldots, A_m$, find an $n\times n$ orthogonal matrix P that provides them with a simultaneous block-diagonal decomposition, i.e., such that $P^{\top}A_1P, \ldots, P^{\top}A_mP$ become block-diagonal matrices with a common block-diagonal structure. Here $A_1, \ldots, A_m$ correspond to data matrices associated with an SDP. As diagonal-blocks of the decomposed matrices get smaller, the transformed SDP could be solved more efficiently by existing software packages developed for SDPs [2, 14, 15, 18]. Naturally we are interested in a finest decomposition.

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 [11, 13] and the other matrix $*$-algebra [16]. 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.

Kanno et al. [8] 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 [17] for solving them. Gatermann and Parrilo [6] 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. [9] as a paper where matrix $*$-algebra was studied in connection with SDPs. Jansson et al. [7] 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 3.1).

All existing studies [1, 4, 6, 7] 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 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 4.

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 $A_1, \ldots, A_m$. The method does not require any algebraic structure to be known in advance, and is based on numerical 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 $A_1, \ldots, A_m$ with the identity matrix I_n, and make use of a well-known fundamental fact (see Theorem 2.1) 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 $\mathcal{T}$, where it is mentioned that a similar technique is employed by Eberly and Giesbrecht [5]; see Remark 3.2 for details. 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 2. In this paper we focus on the case where each irreducible component is isomorphic to a full matrix algebra of some order, whereas the other cases, technically more involved, are treated in [10].

This paper is organized as follows. Section 2 describes the theoretical background of our algorithm based on matrix $*$-algebra. An algorithm for computing the finest simultaneous block-diagonalization is presented in Section 3. Numerical results of SDP problems arising from topology optimization of symmetric trusses are shown in Section 4.

2 Mathematical Basis

2.1 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) $I_n \in \mathcal{T}$ and (ii) $A, B \in \mathcal{T}, \alpha, \beta \in \mathbb{R} \implies \alpha A + \beta B, AB, A^\top \in \mathcal{T}$. 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 $*$-subalgebra $\mathcal{I}$ of $\mathcal{T}$ such that $A \in \mathcal{T}, B \in \mathcal{I} \implies AB, BA \in \mathcal{I}$. 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.

From a standard result of the theory of matrix $*$-algebra (e.g., [16, Chapter X]) we can see the following structure theorem for a matrix $*$-subalgebra; see [9, Theorem 5.4] for more details. 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 2.1. *Let $\mathcal{T}$ be a $*$-subalgebra of $\mathcal{M}_n$.*

(A) *There exist 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 exist 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, we have one of the following three cases.*

(i) $\mathcal{T} = \mathcal{M}_n$.

(ii) *There exists an orthogonal matrix $P \in \mathcal{M}_n$ such that*

$$P^\top \mathcal{T} P = \left\{ \begin{bmatrix} C(v_{11}, w_{11}) & \cdots & C(v_{1\check{n}}, w_{1\check{n}}) \\ \vdots & \ddots & \vdots \\ C(v_{\check{n}1} w_{\check{n}1}) & \cdots & C(v_{\check{n}\check{n}}, w_{\check{n}\check{n}}) \end{bmatrix} \right\},$$

where v_{ij} and w_{ij} run over $\mathbb{R}$ for $i, j = 1, \ldots, \check{n} = n/2$, and

$$C(v, w) = \begin{bmatrix} v & w \\ -w & v \end{bmatrix} \quad \text{for } v, w \in \mathbb{R}.$$

(iii) *There exists an orthogonal matrix $P \in \mathcal{M}_n$ such that*

$$P^\top \mathcal{T} P = \left\{ \begin{bmatrix} H(v_{11}, w_{11}, x_{11}, y_{11}) & \cdots & H(v_{1\tilde{n}}, w_{1\tilde{n}}, x_{1\tilde{n}}, y_{1\tilde{n}}) \\ \vdots & \ddots & \vdots \\ H(v_{\tilde{n}1}, w_{\tilde{n}1}, x_{\tilde{n}1}, y_{\tilde{n}1}) & \cdots & H(v_{\tilde{n}\tilde{n}}, w_{\tilde{n}\tilde{n}}, x_{\tilde{n}\tilde{n}}, y_{\tilde{n}\tilde{n}}) \end{bmatrix} \right\},$$

where v_{ij}, w_{ij}, x_{ij} *and* y_{ij} *run over* $\mathbb{R}$ *for* $i, j = 1, \ldots, \check{n} = n/4$ *and*

$$H(v,w,x,y) = \begin{bmatrix} v & -w & -x & -y \\ w & v & -y & x \\ x & y & v & -w \\ y & -x & w & v \end{bmatrix} \quad \textit{for } v, w, x, y \in \mathbb{R}.$$

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

$$P^\top AP = \bigoplus_{j=1}^{\ell} \bigoplus_{i=1}^{\bar{m}_j} B_j = \bigoplus_{j=1}^{\ell} (I_{\bar{m}_j} \otimes B_j) \tag{2.1}$$

with $B_j \in \mathcal{T}'_j$, where $\mathcal{T}'_j$ denotes the irreducible $*$-subalgebra of $\mathcal{M}_{\bar{n}_j}$ corresponding to the simple subalgebra $\mathcal{T}_j$. The structural indices ℓ, $\bar{n}_j$, $\bar{m}_j$ and the algebraic structure of $\mathcal{T}'_j$ for $j = 1, \ldots, \ell$ are uniquely determined by $\mathcal{T}$. It may be noted that $\hat{n}_j$ in Theorem 2.1 (A) is equal to $\bar{m}_j\bar{n}_j$ in the present notation. Conversely, for any choice of $B_j \in \mathcal{T}'_j$ for $j = 1, \ldots, \ell$, the matrix of (2.1) belongs to $P^\top \mathcal{T} P$.

We denote by

$$\mathbb{R}^n = \bigoplus_{j=1}^{\ell} U_j \tag{2.2}$$

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 2.1 (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 $j = 1, \ldots, \ell$.

In this paper we assume that $\mathcal{T}$ is generated by symmetric matrices and that

$$\text{Case (i) always occurs in Theorem 2.1(C).} \tag{2.3}$$

It is mentioned that an algorithm that works without this assumption is given in [10].

2.2 Simple components from eigenspaces

We denote by $\mathcal{S}_n$ the set of $n \times n$ symmetric real matrices. Let $A_1, \ldots, A_m \in \mathcal{S}_n$, and $\mathcal{T}$ be the $*$-subalgebra over $\mathbb{R}$ generated by $\{I_n, A_1, \ldots, A_m\}$. Note that (2.1) holds for every $A \in \mathcal{T}$ if and only if (2.1) holds for $A = A_p$ for $p = 1, \ldots, m$.

A key observation for our algorithm is that the decomposition (2.2) 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 $\mathcal{T}$, $\alpha_1, \ldots, \alpha_k$ 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

$$Q^\top AQ = \mathrm{diag}\,(\alpha_1 I_{m_1}, \ldots, \alpha_k I_{m_k}). \tag{2.4}$$

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 (2.1) 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 (2.2) into simple components of $\mathcal{T}$, as follows.

Proposition 2.2. *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:*

$$K = K_1 \cup \cdots \cup K_\ell \tag{2.5}$$

such that

$$U_j = \bigoplus_{i \in K_j} V_i, \qquad j = 1, \ldots, \ell. \tag{2.6}$$

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

$$i \sim i' \quad \Longleftrightarrow \quad \exists p\ (1 \leq p \leq m) : Q_i^\top A_p Q_{i'} \neq O, \tag{2.7}$$

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

Proposition 2.3. *The partition* (2.5) *coincides with the partition of K into equivalence classes of the transitive closure of the binary relation $\sim$.*

A generic matrix A can be obtained as a random linear combination of generators of $\mathcal{T}$, 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 2.4. *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 $r \in R$.*

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.

2.3 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 2.5. *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* (2.4)*. Then the transformation matrix P in* (2.1) *can be chosen in the form of*

$$P = Q \cdot \mathrm{diag}\,(P_1, \ldots, P_k) \cdot \Pi \tag{2.8}$$

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 that for each $A' \in \mathcal{T}$ we have $P^\top A'P = I_{\bar{m}} \otimes B'$ for some $B' \in \mathcal{M}_k$, where $\bar{m} = m_1 = \cdots = m_k$. Since P may be replaced by $P(I_{\bar{m}} \otimes S)$ for any orthogonal S, it may be assumed further that $P^\top AP = I_{\bar{m}} \otimes D$, where $D = \mathrm{diag}\,(\alpha_1, \ldots, \alpha_k)$, for the particular generic matrix A. Hence $\Pi P^\top AP\Pi^\top = D \otimes I_{\bar{m}}$ for a permutation matrix Π. Comparing this with $Q^\top AQ = D \otimes I_{\bar{m}}$ and noting that α_i's are distinct, we see that $P\Pi^\top = Q \cdot \mathrm{diag}\,(P_1, \ldots, P_k)$for some $\bar{m} \times \bar{m}$ orthogonal matrices $P_1, \ldots, P_k$. □

3 Algorithm for Simultaneous Block-Diagonalization

On the basis of the theoretical considerations in Section 2, 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:

$$P^\top A_p P = \bigoplus_{j=1}^{\ell} \bigoplus_{i=1}^{\bar{m}_j} B_{pj} = \bigoplus_{j=1}^{\ell} (I_{\bar{m}_j} \otimes B_{pj}), \quad p = 1, \ldots, m, \tag{3.1}$$

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 2.1 for the $*$-subalgebra $\mathcal{T}$ generated by $\{I_n, A_1, \ldots, A_m\}$. The former (Section 3.1) corresponds to the decomposition of $\mathcal{T}$ into simple components and the latter (Section 3.2) to the decomposition into irreducible components. Recall that we assume (2.3).

3.1 Decomposition into simple components

We present here an algorithm for the decomposition into simple components. Algorithm 3.1 below does not presume $\mathrm{span}\{I_n, A_1, \ldots, A_m\} = \mathcal{T} \cap \mathcal{S}_n$, although its correctness relies on this condition.

Algorithm 3.1.

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

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 $Q = (Q_i \mid i = 1, \dots, k)$. This means that

$$Q^\top AQ = \mathrm{diag}\,(\alpha_1 I_{m_1}, \dots, \alpha_k I_{m_k}).$$

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

$$i \sim i' \quad \Longleftrightarrow \quad \exists p\ (1 \le p \le m) : Q_i^\top A_p Q_{i'} \neq O, \tag{3.2}$$

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

$$K = K_1 \cup \cdots \cup K_\ell \tag{3.3}$$

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 $Q[K_j] = (Q_i \mid i \in K_j)$ ($j = 1, \dots, \ell$), and set $\hat{Q} = (Q[K_1], \dots, Q[K_\ell])$. Compute $\hat{Q}^\top A_p \hat{Q}$ ($p = 1, \dots, m$), which results in a simultaneous block-diagonalization with respect to the partition (2.5).

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

Proposition 3.2. *If the matrix A generated in Step 1 is generic in eigenvalue structure, the orthogonal matrix $\hat{Q}$ constructed by Algorithm 3.1 gives the transformation matrix $\hat{Q}$ in Theorem 2.1 (A) for the decomposition of $\mathcal{T}$ into simple components.*

Proposition 2.4 implies that the matrix A in Step 1 is generic with probability one if $\mathrm{span}\{I_n, A_1, \dots, A_m\} = \mathcal{T} \cap \mathcal{S}_n$. This condition, however, is not always satisfied by the given matrices $A_1, \dots, A_m$. 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 $\{I_n, A_1, \dots, A_m\}$. 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 3.3. *If a basis of $\mathcal{T} \cap \mathcal{S}_n$ is computed in advance, Algorithm 3.1 gives, with probability one, the decomposition of $\mathcal{T}$ into simple components.*

3.2 Decomposition into irreducible components

According to Theorem 2.1 (B), the block-diagonal matrices $\hat{Q}^\top A_p \hat{Q}$ obtained by Algorithm 3.1 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 the presentation.

By Proposition 2.5 this finer decomposition can be obtained through a transformation of the form (2.8), which consists of "local coordinate changes" by a family of orthogonal matrices $\{P_1, \dots, P_k\}$, 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

$$P_i^\top Q_i^\top A_p Q_{i'} P_{i'} = b_{ii'}^{(pj)} I_{\bar{m}_j} \tag{3.4}$$

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 (3.1) and Proposition 2.5. 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

$$\tilde{P}^\top A_p \tilde{P} = \bigoplus_{j=1}^{\ell} (B_{pj} \otimes I_{\bar{m}_j}) \tag{3.5}$$

for $p = 1, \ldots, m$. Finally we apply a permutation of rows and columns to obtain (3.1).

The family of orthogonal matrices $\{P_1, \ldots, P_k\}$ satisfying (3.4) can be computed as follows. Recall from (3.2) 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 (3.4) 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 (3.4). 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 3.1. 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, say, $T(g)$. 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. Then G-symmetric matrices A_p 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 3.2 of the present paper to obtain the irreducible decomposition of the representation $T(g)$. Then, under the assumption (2.3), the resulting decomposition of A_p will be the finest decomposition that can be obtained by exploiting the G-symmetry.

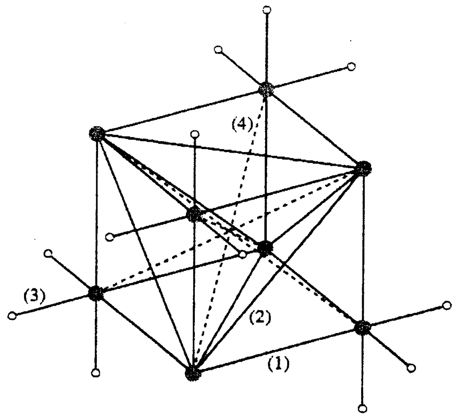

Figure 1: A cubic (or T_d-symmetric) space truss.

Remark 3.2. Eberly and Giesbrecht [5] proposed an algorithm for the simple-component decomposition of a separable matrix algebra (not a $*$-algebra) over an arbitrary infinite field. Their algorithm is closely related to our algorithm in Section 3.1. In particular, their "self-centralizing element" corresponds to our "generic element". Their algorithm, however, is significantly different from ours in two ways: (i) treating a general algebra (not a $*$-algebra) it employs a transformation of the form $S^{-1}AS$ with a nonsingular matrix S instead of an orthogonal transformation, and (ii) it uses companion forms and factorization of minimum polynomials instead of eigenvalue decomposition. The decomposition into irreducible components, inevitably depending on the underlying field, is not treated in [5].

4 Numerical Example: Cubic Trusses

Use and significance of our method are illustrated here in the context of semidefinite programming for truss design treated in [12]. 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 follows. 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 $M_0 \in \mathcal{S}_{n^{\mathrm{d}}}$ denote the mass matrices for the structural and nonstructural masses, respectively. The ith eigenvalue Ω_i of vibration and the corresponding eigenvector $\phi_i \in \mathbb{R}^{n^{\mathrm{d}}}$ are defined by $K\phi_i = \Omega_i(M_{\mathrm{S}} + M_0)\phi_i$ $(i = 1, 2, \ldots, n^{\mathrm{d}})$.

Note that, for a truss, K and M_S can be written as $K = \sum_{j=1}^{n^m} K_j\eta_j$, $M_S = \sum_{j=1}^{n^m} M_j\eta_j$ with constant symmetric matrices K_j and M_j, where η_j denotes the cross-sectional area of the jth member. With the notation $l = (l_j) \in \mathbb{R}^{n^m}$ for the vector of member lengths and $\bar{\Omega}$ for the specified lower bound of the fundamental eigenvalue, our optimization problem is formulated as

$$\min \sum_{j=1}^{n^m} l_j\eta_j \quad \text{s.t.} \quad \Omega_i \geq \bar{\Omega} \ (i = 1,\ldots,n^d), \quad \eta_j \geq 0, \ (j = 1,\ldots,n^m).$$

It is pointed out in [12] that this problem can be reduced to the following dual SDP:

$$\max \ -\sum_{j=1}^{n^m} l_j\eta_j \quad \text{s.t.} \quad \sum_{j=1}^{n^m}(K_j - \bar{\Omega}M_j)\eta_j - \bar{\Omega}M_0 \succeq O, \quad \eta_j \geq 0 \ (j = 1,\ldots,n^m). \qquad (4.1)$$

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

Case 1: $n^m = 34$ members including the dotted ones;

Case 2: $n^m = 30$ members excluding the dotted ones.

A regular tetrahedron is constructed inside the cube. The members outside the cube share the same lengths. A uniform nonstructural mass is located at each node indicated by a filled circle in Fig. 1. 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 (full or achiral) tetrahedral group $\mathrm{T_d}$, which is isomorphic to S_4, the symmetric group of degree 4. The $\mathrm{T_d}$-symmetry can be exploited as follows.

First, we divide the index set of members $\{1,\ldots,n^m\}$ into a family of orbits, say J_p with $p = 1,\ldots,m$, 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.

Next, with reference to the orbits we aggregate the data matrices as well as the components of vector b in (4.1) to A_p $(p = 0,1,\ldots,m)$ and b_p $(p = 1,\ldots,m)$, respectively, as: $A_0 = -\bar{\Omega}M_0$, $A_p = \sum_{j\in J_p}(-K_j + \bar{\Omega}M_j)$ $(p = 1,\ldots,m)$, $b_p = \sum_{j\in J_p} l_j$ $(p = 1,\ldots,m)$. Then (4.1) can be reduced to

$$\max \ -\sum_{p=1}^{m} b_p y_p \quad \text{s.t.} \quad A_0 - \sum_{p=1}^{m} A_p y_p \succeq O, \quad y_p \geq 0 \ (p = 1,\ldots,m) \qquad (4.2)$$

as long as we are interested in a symmetric optimal solution. Note that the matrices A_p $(p = 0,1,\ldots,m)$ are $\mathrm{T_d}$-symmetric. The proposed method is applied to A_p $(p = 0,1,\ldots,m)$ for their simultaneous block-diagonalization. The assumption (2.3) has turned out to be satisfied.

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

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

size 4, as summarized on the left of Table 1. This result conforms with the group-theoretic analysis. In Case 2 sparsity plays a role to split the last block into two, as shown on 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.

Acknowledgments

The authors thank Etienne de Klerk for communicating [5], Dima Pasechnik for discussion, and Takanori Maehara for pointing out a flaw in the original proof of Proposition 2.5. 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 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 programming relaxations of the quadratic assignment problem, CentER Discussion paper 2007-44, Center for Economic Research, Tilburg University, June 2007.

[5] W. Eberly and M. Giesbrecht: Efficient decomposition of separable algebras, Journal of Symbolic Computation, Vol. 37 (2004), pp. 35–81.

[6] 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.

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

[8] 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.

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

[10] T. Maehara and K. Murota: A numerical algorithm for block-diagonal decomposition of matrix $*$-algebras with general irreducible components, METR 2008-26, Department of Mathematical Informatics, University of Tokyo, May 2008.

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

[12] 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.

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

[14] 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.

[15] 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.

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

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

[18] 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.