MINISTRY OF EDUCATION AND TRAINING QUY NHON UNIVERSITY

NGUYEN THI NGAN

SIMULTANEOUS DIAGONALIZATIONS OF MATRICES AND APPLICATIONS FOR SOME CLASSES OF OPTIMIZATION

Speciality: Algebra and number theory Code: 9 46 01 04

ABSTRACT OF DOCTORAL THESIS IN MATHEMATICS

Binh Dinh - 2024

This thesis was completed at Quy Nhon University

BOARD OF SUPERVISORS:

Dr. Thanh-Hieu Le Prof. Ruey-Lin Sheu

Reviewer 1: Assoc. Prof. Dr. Vu The Khoi

Reviewer 2: Assoc. Prof. Dr. Mai Hoang Bien

Reviewer 3: Assoc. Prof. Dr. Phan Thanh Nam

The thesis shall be defended before the Thesis Assessment Council at Quy Nhon University at:....

The thesis can be found at:

- National Library of Vietnam
- Library of Quy Nhon University

Contents

Introduction 1						
1	Pre	reliminaries				
	1.1	Some prepared concepts for the SDC problems	9			
	1.2	Existing SDC results	10			
2	Solving the SDC problems of Hermitian matrices and real symmetric matrices					
	2.1	The Hermitian SDC problem	12			
		2.1.1 The max-rank method	12			
		2.1.2 The SDP method	15			
	2.2	An alternative solution method for the SDC problem				
		of real symmetric matrices	15			
		2.2.1 $$ The SDC problem of nonsingular collection $$.	15			
		2.2.2 The SDC problem of singular collection	16			
3	Son	ne applications of the SDC results	18			
	3.1 Computing the positive semidefinite interval					

	3.1.1	Computing $I_{\geq}(C_1, C_2)$ when C_1, C_2 are \mathbb{R} -SDC		
			18	
	3.1.2	Computing $I_{\geq}(C_1, C_2)$ when C_1, C_2 are not \mathbb{R}-SDC	20	
3.2	Solvin	g the quadratically constrained quadratic pro-		
	gram	ning	21	
3.3	Applie quotie	m cations for maximizing a sum of generalized Rayle ents	igh 21	
Conclusions Future research				

Introduction

Let $\mathcal{C} = \{C_1, C_2, \dots, C_m\}$ be a collection of $n \times n$ matrices with elements in \mathbb{F} , where \mathbb{F} is the field \mathbb{R} of real numbers or the field $\mathbb C$ of complex numbers. If there is a nonsingular matrix R such that R^*C_iR are all diagonal, the collection \mathcal{C} is then said to be simultaneously diagonalizable via congruence, where R^* is the conjugate transpose of R if C_i are Hermitian and simply the transpose of R if C_i are either complex or real symmetric matrices. Moreover, if there exists a nonsingular matrix S such that $S^{-1}C_iS$ is diagonal for every $i = 1, 2, \ldots, m$ then \mathcal{C} is called *si*multaneously diagonalizable via similarity, shortly SDS. For convenience, throughout the dissertation we use "SDC" to stand for either "simultaneously diagonalizable via congruence" or "simultaneous diagonalization via congruence" if no confusion will arise. The SDS problem is well-known and is completely solved. But the SDC problem is still open in some senses. The SDC of \mathcal{C} implies that a single change of basis x = Ry, makes all the quadratic forms x^*C_ix simultaneously become the canonical forms. Specifically, if $R^*C_iR = \text{diag}(\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{in})$ is the diagonal matrix with diagonal elements $\alpha_{i1}, \alpha_{i2}, \ldots, \alpha_{in}$, then x^*C_ix is transformed to the sum of squares $y^*(R^*C_iR)y = \sum_{j=1}^n \alpha_{ij}|y_j|^2$, for i = 1, 2, ..., m. This is one of the properties connecting the SDC of matrices with many applications such as variational analysis [31], signal processing [14], [52], [62], quantum mechanics [57], medical imaging analysis [2],[13],[67] and many others, please see references therein. Especially, the SDC suggests a promising approach for solving quadratically constrained quadratic programming (QCQP) [17], [74, [5]. In recent studies by Ben-Tal and Hertog [6], Jiang and Li [37], Alizadeh [4], Taati [54], Adachi and Nakatsukasa [1], the SDC of two or three real symmetric matrices has been efficiently applied for solving QCQP with one or two constraints. Ben-Tal and Hertog [6] showed that if the matrices in the objective and constraint

functions are SDC, the QCQP with one constraint can be recast as a convex second-order cone programming (SOCP) problem; the QCQP with two constraints can also be transformed into an equivalent SOCP under the SDC together with additional appropriate assumptions. We know that the convex SOCP is solvable efficiently in polynomial time [4]. Jiang and Li [37] applied the SDC for some classes of QCQP including the generalized trust region subproblem (GTRS), which is exactly the QCQP with one constraint, and its variants. Especially the homogeneous version of QCQP, i.e., when the linear terms in the objective and constraint functions are all zero, is reduced to a linear program if the matrices are SDC. Salahi and Taati [54] derived an efficient algorithm for solving GTRS under the SDC condition. Also under the SDC assumption, Adachi and Nakatsukasa [1] compute the positive definite interval $I_{>}(C_0, C_1) = \{\mu \in \mathbb{R} : C_0 + \mu C_1 > 0\}$ of the matrix pencil and propose an eigenvalue-based algorithm for a definite feasible GTRS, i.e., the GTRS satisfies the Slater condition and $I_{>}(C_0, C_1) \neq \emptyset$.

Those important applications stimulate various studies on the problem, that we call the SDC problem in this dissertation. It is to find conditions on $\{C_1, C_2, \ldots, C_m\}$ ensuring the existence of a congruence matrix R for the SDC problem of real symmetric matrices [70], [27], [41], [65], [37], the SDC problem of complex symmetric matrices [34], [11] and the SDC problem of Hermitian matrices [74], [7], [34]. However, for the real setting, the best SDC results so far can only solve the case of two matrices while the case of more than two matrices is solved under the assumption of a positive semidefinite matrix pencil [37]. On the other hand, for the SDC problem of complex matrices, including the complex symmetric and Hermitian matrices, can be equivalently rephrased as a simultaneous diagonalization via similarity (SDS) problem m [74], [7], [8], [11]. More importantly the obtained results do not include algorithms for finding a congruence matrix R, except for the case of two real symmetric matrices by Jiang and Li [37]. Those unsolved issues inspire us to investigate, in this dissertation, algorithms for determining whether a class C is SDC and compute a congruence matrix R if it indeed is.

The SDC problem was first developed by Weierstrass [70] in 1868. He obtained sufficient SDC conditions for a pair of real symmetric matrices. Since then, several authors have extended those results, including Muth 1905 [45], Finsler 1937 [18], Albert 1938 [3], Hestenes 1940 [28], and various others. See, for example, [12], [27], [29], [30], [34],[44], [65]. The results for two matrices obtained so far can be shortly reviewed as follows. If at least one of the matrices C_1, C_2 is nonsingular, referred to as a nonsingular pair, suppose it is C_1 , then C_1, C_2 are SDC if and only if $C_1^{-1}C_2$ is similarly diagonalizable [27], see also [64], [65]. If the non-singularity is not assumed, the obtained SDC results of C_1, C_2 were only sufficient. Specifically,

- a) if there exist scalars $\mu_1, \mu_2 \in \mathbb{R}$ such that $\mu_1 C_1 + \mu_2 C_2 > 0$, then C_1, C_2 are SDC [30], [65];
- b) if $\{x \in \mathbb{R}^n : x^T C_1 x = 0\} \cap \{x \in \mathbb{R}^n : x^T C_2 x = 0\} = \{0\}$ then C_1, C_2 are SDC [44], [59], [65].

Actually, the classical Finsler theorem [18] in 1937 indicated that these two conditions a) and b) are equivalent whenever $n \ge 3$. It has to wait until Hoi [74] in 1970 and independently Becker [5] in 1980 for a necessary and sufficient SDC condition for a pair of Hermitian matrices. Unfortunately, when more than two matrices are involved, none of those aforementioned results remains true. In 1990 and 1991, Binding [7],[8] provided some equivalent conditions, which link to the generalized eigenvalue problem and numerical range of Hermitian matrices or to the generalized eigenvalue problem, for a finite collection of Hermitian matrices to be SDC by a unitary matrix. However there is still lack of algorithms for finding a congruence matrix R. In 2002, Hiriart-Urruty and M. Torki [29] and then, in 2007, Hiriart-Urruty [30] proposed an open problem to find sensible and "palpable" conditions on C_1, C_2, \ldots, C_m ensuring they are simultaneously diagonalizable via congruence. In 2016 Jiang and Li [37] obtained a necessary and sufficient SDC condition for a pair of real symmetric matrices and proposed an algorithm for finding a congruence matrix R if it exists. Nevertheless, we find that the result of Jiang and Li [37] is not complete. A missing case not considered in their paper is now added to make it up in this dissertation. For more than two matrices, Jiang and Li [37] proposed a necessary and sufficient SDC condition under the existence assumption of a semidefinite matrix pencil. After this result, an open question still remains to be investigated: solving the SDC problem of more than two real symmetric matrices without semidefinite matrix pencil assumption? In 2020, Bustamante et al. [11] proposed a necessary and sufficient condition for a set of complex symmetric matrices to be SDC by equivalently rephrasing the SDC problem as the classical problem of simultaneous diagonalization via similarity (SDS) of a new related set of matrices. A procedure to determine in a finite number of steps whether or not a set of complex symmetric matrices is SDC was also proposed. However, the SDC results of complex symmetric matrices may not hold for the real setting. That is, even the given matrices C_1, C_2, \ldots, C_m are all real, the resulting matrices R and $R^T C_i R$ may have to be complex, please see Example 16 [11] and also in Example 2.1.7. Apparently, the SDC of complex symmetric matrices does also not hold for the Hermitian matrices, please see Theorem 4.5.15 [34], Example 2.1.7.

The dissertation is organized as follows. In Chapter 1 we present some related concepts and obtained results so far of the SDC problem including the SDC of real symmetric matrices, complex symmetric matrices and Hermitian matrices. In Chapter 2 we first focus on solving the SDC problem of Hermitian matrices, i.e., when C_i are all Hermitian. This part is based on the results in [42]. The main contributions of this part are as follows.

• We develop sufficient and necessary conditions (see Theorems

2.1.4 and 2.1.5) for a collection of finitely many Hermitian matrices to be simultaneously diagonalizable via *-congruence. The proofs use only matrix computation techniques;

- Interestingly, one of the conditions shown in Theorem 2.1.5 requires the existence of a positive definite solution of a system of linear equations over Hermitian matrices. This leads to the use of the SDP solvers (for example, SDPT3 [63]) for checking the simultaneous diagonalizability of the initial Hermitian matrices. In case the matrices are SDC, i.e., such a positive definite solution exists, we apply the existing Jacobi-like method in [10], [43] to simultaneously diagonalize the commuting Hermitian matrices that are the images of the initial ones under the congruence defined by the square root of the above positive definite solution. The Hermitian SDC problem is hence completely solved. As a consequence, this solves the long-standing SDC problem for real symmetric matrices mentioned as an open problem in [30], and for arbitrary square matrices since any square matrix is a summation of its Hermitian and skew Hermitian parts (see Theorem 2.1.6);
- In line with giving the equivalent condition that requires the maximum rank of Hermitian pencils (Theorem 2.1.2), we suggest a Schmüdgen-like algorithm for finding such the maximum rank in Algorithm 2. This methodology may also be applied in some other simultaneous diagonalizations, for example, that in [11];
- Finally, we propose corresponding algorithms the most important one of which is Algorithm 6 for solving the Hermitian SDC problem. These are implemented in MATLAB. The main algorithm consists of two stages which are summarized as follows: For $C_1, \ldots, C_m \in \mathbb{H}^n$,
 - STAGE 1: Checking if there is a positive definite matrix P solving an appropriate semidefinite program based on The-

orem 2.1.5 iii). Our main contribution stays in this part.

STAGE 2: If such a P exists then applying Algorithm 5 [10], [43] to find a unitary matrix V that simultaneously diagonalizes the new commuting Hermitian matrices $\sqrt{P}C_i\sqrt{P}, i = 1, ..., m.$

The second part of Chapter 2 is based on [49], which focuses on the SDC problem of the real symmetric matrices, i.e., when C_i are all real symmetric. Although, in Theorem 2.1.5, our results (i)-(iii) on the Hermitian matrices can also apply to the real setting, get we find that the decomposition techniques for two matrices in [37] can be generalized to construct an inductive procedure for the SDC problem of \mathcal{C} with $m \ge 3$. The approach based on [37] may be better than the SDP one, please see Example 2.2.2. To this end the collection \mathcal{C} is divided into two cases: the nonsingular collection, denoted by \mathcal{C}_{ns} , when at least one $C_i \in \mathcal{C}$ is non-singular. Without loss of generality, we always assume that C_1 is non-singular. On the other hand, the singular collection, denoted by C_s , when all $C'_i s$ in C are non-zero but singular. For the nonsingular collection C_{ns} , the arguments first apply to $\{C_1, C_2\}$; if C_1, C_2 are SDC then a matrix $Q^{(1)}$ is constructed at the first iteration such that $C_2^{(1)} := (Q^{(1)})^T C_2 Q^{(1)}$ is a non-homogeneous dilation of $C_1^{(1)} := (Q^{(1)})^T C_1 Q^{(1)}$, while $C_j^{(1)} := (Q^{(1)})^T C_j Q^{(1)}, j \ge 3$ share the same block diagonal structure of $C_1^{(1)}$, please see Lemma 2.2.2 and Remark 2.2.1 below. At the second iteration, $\{C_1^{(1)}, C_3^{(1)}\}$ are checked. If $C_1^{(1)}, C_3^{(1)}$ are SDC, then $Q^{(2)}$ is constructed such that $C_3^{(2)} := (Q^{(2)})^T C_3^{(1)} Q^{(2)}$ and $C_2^{(2)}:=(Q^{(2)})^TC_2^{(1)}Q^{(2)}$ are non-homogeneous dilations of $C_1^{(2)}:=$ $(Q^{(2)})^T C_1^{(1)} Q^{(2)}$. Next, $\{C_1^{(2)}, C_4^{(2)}\}$ are considered at the third step; and so forth. These results are presented in Sect. 2.2.1. For the singular collection C_s , we also begin with $\{C_1, C_2\}$. If C_1, C_2 are SDC, we find a nonsingular matrix U_1 to get

$$\hat{C}_1 := U_1^T C_1 U_1 = \text{diag}((C_{11})_{p_1}, 0_{n-p_1}), p_1 < n,$$

$$\hat{C}_2 := U_1^T C_2 U_1 = \operatorname{diag}((C_{21})_{p_1}, 0_{n-p_1})$$

such that $(C_{11})_{p_1}, (C_{21})_{p_1}$ are SDC and $(C_{21})_{p_1}$ is nonsingular. At the second step, we consider the SDC of \hat{C}_1, \hat{C}_2 and $\hat{C}_3 = U_1^T C_3 U_1$. If they are SDC, we find a nonsingular matrix U_2 to get

$$\begin{split} \check{C}_1 &:= U_2^T \hat{C}_1 U_2 = \operatorname{diag}((C_{11})_{p_2}, 0_{n-p_2}), p_1 \leqslant p_2, \\ \check{C}_2 &:= U_2^T \hat{C}_2 U_2 = \operatorname{diag}((C_{21})_{p_2}, 0_{n-p_2}), \\ \check{C}_3 &:= U_2^T \hat{C}_3 U_2 = \operatorname{diag}((C_{31})_{p_2}, 0_{n-p_2}) \end{split}$$

such that $(C_{11})_{p_2}, (C_{21})_{p_2}, (C_{31})_{p_2}$ are SDC and $(C_{31})_{p_2}$ is nonsingular; and so forth. By this way, we show that if C_s is SDC, we can create a new collection $\tilde{C}_s = {\tilde{C}_1, \tilde{C}_2, \ldots, \tilde{C}_m}$ such that $\tilde{C}_i = \operatorname{diag}((C_{i1})_p, 0_{n-p}), p \leq n$, and $(C_{(m-1)1})_p$ is nonsingular. Importantly, the given collection C_s is SDC if and only if $(C_{11})_p, (C_{21})_p$, $\ldots, (C_{(m-1)1})_p, (C_{m1})_p$ are SDC. Therefore, we move from the SDC of a singular collection to the SDC of a nonsingular collection; please see Theorem 2.2.3 in Sect. 2.2.2.

Chapter 3 is devoted to presenting some applications of the SDC results. We first show how to explore the SDC properties of two real symmetric matrices C_1, C_2 to compute the positive semidefinite interval $I_{\geq}(C_1, C_2) = \{\mu \in \mathbb{R} : C_1 + \mu C_2 \geq 0\}$ of matrix pencil $C_1 + \mu C_2$. Indeed, we show that if C_1, C_2 are not SDC, then $I_{\geq}(C_1, C_2)$ has at most one value μ , while if C_1, C_2 are SDC, $I_{\geq}(C_1, C_2)$ could be empty, a singleton set or an interval. Each case helps to analyze when the GTRS is unbounded from below, has a unique Lagrange multiplier or has an optimal Lagrange multiplier μ^* in a given closed interval. Such a μ^* can be computed by a bisection algorithm. This results follow from [47]. The next application will be for QCQP which takes the following format

(QCQP)
$$\begin{array}{c} \min \quad x^T C_1 x + 2a_1^T x \\ \text{s.t.} \quad x^T C_i x + 2a_i^T x + b_i \leqslant 0, \ i = 2, \dots, m, \end{array}$$

where $a_i \in \mathbb{R}^n, b_i \in \mathbb{R}$. We show that if the matrices C_i in the objective and constraint functions are SDC, the QCQP can be relaxed

to a convex SOCP problem. In general, the ralaxation admits a positive gap. That is, the optimal value of the relaxed SOCP is strictly less than that of the primal QCQP. The cases with a tight ralaxation will be presented in that chapter. Especially, if the matrices C_i are SDC and the QCQP is homogeneous, i.e., $a_i = 0$ for i = 1, 2, ..., m, then QCQP is reduced to a linear programming after two times of changing variables. A special case of the homogeneous QCQP, which minimizes a quadratic form subjective to two homogeneous quadratic constraints over the unit sphere [46], is reduced to a linear programming problem on a simplex if the matrices are SDC. Finally, we show the applications for solving a generalized Rayleigh quotient problem which maximizes a sum of generalized Rayleigh quotients.

Chapter 1

Preliminaries

1.1 Some prepared concepts for the SDC problems

Let us begin with some notations:

- The matrices C₁, C₂,..., C_m ∈ Hⁿ are said to be SDC on C, shortly written as *-SDC, if there exists a nonsingular matrix P ∈ C^{n×n} such that every P*C_iP is diagonal in R^{n×n}.
- The matrices $C_1, C_2, \ldots, C_m \in S^n$ are said to be SDC on \mathbb{R} , shortly written as \mathbb{R} -SDC, if there exists a nonsingular matrix $P \in \mathbb{R}^{n \times n}$ such that every $P^T C_i P$ is diagonal in $\mathbb{R}^{n \times n}$.
- Matrices $C_1, C_2, \ldots, C_m \in S^n(\mathbb{C})$ are said to be SDC on \mathbb{C} if there exists a nonsingular matrix $P \in \mathbb{C}^{n \times n}$ such that every $P^T C_i P$ is diagonal in $\mathbb{C}^{n \times n}$. We also abbreviate this as \mathbb{C} -SDC.

1.2 Existing SDC results

Lemma 1.2.1. ([27], p.255) Two matrices $C_1, C_2 \in S^n$, with C_1 nonsingular, are \mathbb{R} -SDC if and only if $C_1^{-1}C_2$ is real similarly diagonalizable.

Lemma 1.2.6. ([37], Lemma 5) For any two matrices $C_1, C_2 \in S^n$ singular, there always exists a nonsingular matrix U such that

$$\tilde{A} := U^T C_1 U = \begin{pmatrix} A_1 & 0_{p \times (n-p)} \\ 0_{(n-p) \times p} & 0_{n-p} \end{pmatrix}$$
(1.1)

and

$$\tilde{B} := U^{T} C_{2} U = \begin{pmatrix} B_{1} & 0_{p \times q} & B_{2} \\ 0_{q \times p} & B_{3} & 0_{q \times r} \\ B_{2}^{T} & 0_{r \times q} & 0_{r} \end{pmatrix}$$
(1.2)

where $p, q, r \ge 0, p + q + r = n, A_1, B_3$ is a nonsingular diagonal matrix.

Lemma 1.2.8. Let both $C_1, C_2 \in S^n$ be non-zero singular with $rank(C_1) = p < n$. There exists a nonsingular matrix U_1 , such that

$$\tilde{C}_1 = U_1^T C_1 U_1 = \begin{pmatrix} (C_{11})_p & 0\\ invert. \& diag. \\ 0 & 0_{n-p} \end{pmatrix},$$
(1.3)

$$\tilde{C}_2 = U_1^T C_2 U_1 = \begin{pmatrix} (C_{21})_p & C_{22} \\ C_{22}^T & 0_{n-p} \end{pmatrix},$$
(1.4)

or

$$\tilde{C}_{2} = U_{1}^{T} C_{2} U_{1} = \begin{pmatrix} (C_{21})_{p} & 0 & C_{25} \\ 0 & (C_{26})_{s_{1}} & 0 \\ & & \\ & & \\ & & \\ & & \\ C_{25}^{T} & 0 & 0_{n-p-s_{1}} \end{pmatrix}.$$
 (1.5)

where C_{11}, C_{26} are nonsingular diagonal matrices; $s_1 \leq n - p$. If $s_1 = n - p$ then C_{25} does not exist.

Lemma 1.2.9. ([37], Theorem 6) Two singular matrices C_1 and C_2 , which take the forms (1.1) and (1.2), respectively, are \mathbb{R} -SDC if and only if A_1 and B_1 are \mathbb{R} -SDC and B_2 is a zero matrix or $r = n - p - s_1 = 0$ (B_2 does not exist).

Theorem 1.2.1. Let C_1 and C_2 be two symmetric singular matrices of $n \times n$. Let U_1 be the nonsingular matrix that puts $\tilde{C}_1 = U_1^T C_1 U_1$ and $\tilde{C}_2 = U_1^T C_2 U_1$ into the format of (1.3) and (1.4) in Lemma 1.2. Then, \tilde{C}_1 and \tilde{C}_2 are \mathbb{R} -SDC if and only if C_{11} , C_{21} are \mathbb{R} -SDC and $C_{22} = 0_{p \times r}$, where r = n - p.

Chapter 2

Solving the SDC problems of Hermitian matrices and real symmetric matrices

2.1 The Hermitian SDC problem

2.1.1 The max-rank method

Theorem 2.1.1. Let $\mathfrak{C} = \mathfrak{C}(\lambda) \in \mathbb{F}[\lambda]^{n \times n}$ be a Hermitian pencil, i.e, $\mathfrak{C}(\lambda)^* = \mathfrak{C}(\lambda)$ for every $\lambda \in \mathbb{R}^m$. Then there exist polynomial matrices $\mathfrak{X}_+, \mathfrak{X}_- \in \mathbb{F}[\lambda]^{n \times n}$ and polynomials $b, d_j \in \mathbb{R}[\lambda], j = 1, 2, ..., n$ (note that b, d_j are always real even when \mathbb{F} is the complex field) such that

$$\mathfrak{X}_+\mathfrak{X}_- = \mathfrak{X}_-\mathfrak{X}_+ = b^2 I_n \tag{2.1a}$$

$$b^{4}\mathfrak{C} = \mathfrak{X}_{+}\operatorname{diag}(d_{1}, d_{2}, \dots, d_{n})\mathfrak{X}_{+}^{*}, \qquad (2.1b)$$

$$\mathfrak{X}_{-}\mathfrak{C}\mathfrak{X}_{-}^{*} = \operatorname{diag}(d_{1}, d_{2}, \dots, d_{n}).$$
(2.1c)

$$\mathfrak{C}_{k-1} = \begin{pmatrix} \alpha_k & \beta_k \\ \beta_k^* & \widehat{\mathfrak{C}}_k \end{pmatrix}, \mathfrak{C}_k = \alpha_k (\alpha_k \widehat{\mathfrak{C}}_k - \beta_k^* \beta_k), b = \prod_{t=1}^k \alpha_t, \\
\mathfrak{X}_{k+} = \mathfrak{X}_{(k-1)+} \begin{pmatrix} \alpha_k I & 0 \\ 0 & \mathfrak{Y}_{k+} \end{pmatrix}, \mathfrak{X}_{k-} = \begin{pmatrix} \alpha_k I_{k-1} & 0 \\ 0 & \mathfrak{Y}_{k-} \end{pmatrix}, \mathfrak{X}_{(k-1)-}, \\
\mathfrak{X}_{k-} \mathfrak{C} \mathfrak{X}_{k-}^* = \begin{pmatrix} \operatorname{diag}(d_1, d_2, \dots, d_{k-1}, d_k) & 0 \\ 0 & \mathfrak{C}_k \end{pmatrix} := \widetilde{\mathfrak{C}}_k,$$
(2.2)

where
$$\mathfrak{Y}_{k\pm} = \begin{pmatrix} \alpha_k & 0\\ \pm \beta_k^* & \alpha_k I_{n-k} \end{pmatrix}$$
 and
 $d_k = \alpha_k^3, d_j = \alpha_j^3 \prod_{t=j+1}^k \alpha_t^2, j = 1, 2, \dots, k-1.$ (2.3)

Theorem 2.1.2. Use notation as in Theorem 2.1.1, and suppose \mathfrak{C}_k in (2.2) is diagonal but every \mathfrak{C}_t , $t = 0, 1, \ldots, k - 1$, is not so. Consider the modification of (2.2) as

$$\mathfrak{C}_{k-1} = \begin{pmatrix} \alpha_k & \beta_k \\ \beta_k^* & \mathfrak{C}_k \end{pmatrix}, \qquad \mathfrak{C}_k = \alpha_k (\alpha_k \mathfrak{C}_k - \beta_k^* \beta_k), \\
\mathfrak{X}_{k-} = \begin{pmatrix} I_{k-1} & 0 \\ 0 & \mathfrak{Y}_{k-} \end{pmatrix} \mathfrak{X}_{(k-1)-}, \qquad \mathfrak{Y}_{k\pm} = \begin{pmatrix} \alpha_k & 0 \\ \pm \beta_k^* & \alpha_k I_{n-k} \end{pmatrix}, \\
\mathfrak{X}_{k-} \mathfrak{C} \mathfrak{X}_{k-}^* = \begin{pmatrix} \operatorname{diag}(\alpha_1^3, \alpha_2^3, \dots, \alpha_{k-1}^3, \alpha_k^3) & 0 \\ 0 & \mathfrak{C}_k \end{pmatrix} := \mathfrak{C}_k, \\
(2.4)$$

Moreover, let $d_i = \alpha_i^3$, i = 1, 2, ..., k, and $\mathfrak{C}_k = \text{diag}(d_{k+1}, d_{k+2}, ..., d_n)$, $d_j \in \mathbb{R}[\lambda], j = 1, 2, ..., n$, and some of $d_{k+1}, d_{k+2}, ..., d_n$ may be identically zero. The following holds true.

- (i) α_t divides α_{t+1} (and therefore d_t divides d_{t+1}) for every $t \leq k-1$, and if k < n, then α_k divides every $d_j, j > k$.
- (ii) The pencil $\mathfrak{C}(\lambda)$ has the maximum rank r if and only if there exists a permutation such that $\tilde{\mathfrak{C}}(\lambda) = \operatorname{diag}(d_1, \ldots, d_r, 0, \ldots, 0)$, d_j is not identically zero for every $j = 1, 2, \ldots, r$. In addition, the maximum rank of $\mathfrak{C}(\lambda)$ achieves at $\hat{\lambda}$ if and only if $\alpha_k(\hat{\lambda}) \neq 0$ or $(\prod_{t=k+1}^r d_t(\hat{\lambda})) \neq 0$, respectively, depends upon \mathfrak{C}_k being identically zero or not.

Theorem 2.1.3. The matrices $I, C_1, \ldots, C_m \in \mathbb{H}^n, m \ge 1$ are *-SDC if and only if they are commuting. Moreover, when this the case, there are *-SDC by a unitary matrix (resp., orthogonal one) if C_1, \ldots, C_m are complex (resp., all real).

Theorem 2.1.4. Let $0 \neq C_1, C_2, \ldots, C_m \in \mathbb{H}^n$ with $\dim_{\mathbb{C}}(\bigcap_{t=1}^m kerC_t) = q$, (always q < n.)

- 1. If q = 0, then the following hold:
 - (i) If $det \mathfrak{C}(\lambda) = 0$, for all $\lambda \in \mathbb{R}^m$ (over only real m-tuple λ), then C_1, \ldots, C_m are not *-SDC.
 - (ii) Otherwise, there exists $\lambda \in \mathbb{R}^m$ such that $\det \mathfrak{C}(\lambda) \neq 0$. The matrices C_1, \ldots, C_m are *-SDC if and only if $\mathfrak{C}(\lambda)^{-1}C_1, \ldots, \mathfrak{C}(\lambda)^{-1}C_m$ pairwise commute and every $\mathfrak{C}(\lambda)^{-1}C_i, i = 1, 2, \ldots, m$, is similar to a real diagonal matrix.
- 2. If q > 0, then there exists a nonsingular matrix V such that

$$V^*C_iV = \operatorname{diag}(\hat{C}_i, 0_q), \forall i = 1, \dots, m,$$

$$(2.5)$$

where 0_q is the $q \times q$ zero matrix and $\hat{C}_i \in \mathbb{H}^{n-q}$ with $\bigcap_{t=1}^m \ker \hat{C}_t$ = 0. Moreover, C_1, C_2, \ldots, C_m are *-SDC if and only if $\hat{C}_1, \hat{C}_2, \ldots, \hat{C}_m$ are *-SDC.

2.1.2 The SDP method

Theorem 2.1.5. The following conditions are equivalent:

- (i) The matrices $C_1, C_2, \ldots, C_m \in \mathbb{H}^n$ are *-SDC.
- (ii) There exists a nonsingular matrix $P \in \mathbb{C}^{n \times n}$ such that P^*C_1P , ..., P^*C_mP commute.
- (iii) There exists a positive definite $X = X^* \in \mathbb{H}^n$ that solves the following systems:

$$C_i X C_j = C_j X C_i, \quad 1 \le i < j \le m.$$
(2.6)

We note that the theorem is also true for the real setting: If C_i 's are all real then the corresponding matrices P, X in all conditions above can be all picked to be real.

Let $\mathcal{H}(A) = \frac{1}{2}(A + A^*)$, $\mathcal{S}(A) = \frac{1}{2}(A - A^*) = -\mathcal{S}(A)^*$. We further note that both $\mathcal{H}(A)$ and $\mathbf{i}\mathcal{S}(A)$ are Hermitian matrices.

Theorem 2.1.6. (see, e.g., in Section 1.7, Problem 18 [35]) The square matrices $A_1, \ldots, A_m \in \mathbb{F}^{n \times n}$ are *-SDC if and only if so are $\mathcal{H}(A_t), \mathbf{i}\mathcal{S}(A_t), t = 1, \ldots, m$.

2.2 An alternative solution method for the SDC problem of real symmetric matrices

2.2.1 The SDC problem of nonsingular collection

Theorem 2.2.1. Let $C_{ns} = \{C_1, C_2, \ldots, C_m\} \subset S^n, m \ge 3$ be a nonsingular collection with C_1 invertible. Suppose for each *i* the

matrix $C_1^{-1}C_i$ is real similarly diagonalizable. If $C_jC_1^{-1}C_i$ are symmetric for $2 \leq i < j \leq m$, then there always exists a nonsingular real matrix R such that

$$R^{T}C_{1}R = diag(A_{1}, A_{2}, \dots, A_{s}),$$

$$R^{T}C_{2}R = diag(\alpha_{1}^{2}A_{1}, \alpha_{2}^{2}A_{2}, \dots, \alpha_{s}^{2}A_{s}),$$

$$\dots$$

$$R^{T}C_{m}R = diag(\alpha_{1}^{m}A_{1}, \alpha_{2}^{m}A_{2}, \dots, \alpha_{s}^{m}A_{s}),$$

$$(2.7)$$

where $A'_t s$ are nonsingular and symmetric, $\alpha^i_t, t = 1, \ldots, s$, are real numbers.

Theorem 2.2.2. Let $C_{ns} = \{C_1, C_2, \ldots, C_m\} \subset S^n, m \ge 3$ be a nonsingular collection with C_1 invertible. The collection C_{ns} is \mathbb{R} -SDC if and only if for each $2 \le i \le m$, the matrix $C_1^{-1}C_i$ is real similarly diagonalizable and $C_jC_1^{-1}C_i, 2 \le i < j \le m$ are all symmetric.

2.2.2 The SDC problem of singular collection

Theorem 2.2.3. Let $C_s = \{C_1, C_2, \ldots, C_m\} \subset S^n, \ m \ge 3$ be a singular collection in which none is zero. If $C_1, C_2, \ldots, C_{m-1}$ are \mathbb{R} -SDC, then there exist a nonsingular real matrix Q and a positive vector $\mu = (\mu_1, \mu_2, \ldots, \mu_{m-2}, 1) \in \mathbb{R}^{m-1}_{++}$ such that

$$\begin{split} \tilde{C}_1 &= Q^T C_1 Q = diag((C_{11})_p, 0_{n-p}), \ p < n; \\ &\vdots \\ \tilde{C}_{m-1} &= Q^T (\mu_{m-2}(\dots + C_{m-2}) + C_{m-1}) Q = diag((C_{(m-1)1})_p, 0_{n-p}); \end{split}$$

and either

$$\tilde{C}_m = Q^T C_m Q = \begin{pmatrix} (C_{m1})_p & C_{m2} \\ C_{m2}^T & 0_{n-p} \end{pmatrix};$$
(2.8)

or

$$\tilde{C}_m = Q^T C_m Q = \begin{pmatrix} (C_{m1})_p & 0 & C_{m5} \\ 0 & (C_{m6})_s & 0 \\ C_{m5}^T & 0 & 0_{n-p-s} \end{pmatrix}, \quad (2.9)$$

where

- the sub-matrices (C_{i1})_p, i = 1, 2, ..., m-1, are all diagonal of the same size. In particular, (C_{(m-1)1})_p is nonsingular;
- in (2.8), $(C_{m1})_p$ is symmetric;
- in (2.9), (C_{m1})_p is symmetric, (C_{m6})_s is nonsingular diagonal; C_{m5} is either a p×(n−p−s) matrix if s < n−p or does not exist if s = n − p.

Moreover, the following three statements are equivalent.

- (i) all matrices in the collection C_s are \mathbb{R} -SDC;
- (ii) all matrices in the collection $\tilde{C}_s = \{\tilde{C}_1, \dots, \tilde{C}_m\}$ are \mathbb{R} -SDC;
- (iii) either sub-blocks C_{11}, \ldots, C_{m1} with C_{m1} coming from (2.8) are \mathbb{R} -SDC and $C_{m2} = 0$; or sub-blocks C_{11}, \ldots, C_{m1} with C_{m1} coming from (2.9) are \mathbb{R} -SDC and either $C_{m5} = 0$ or C_{m5} does not exist.

Chapter 3

Some applications of the SDC results

3.1 Computing the positive semidefinite interval

3.1.1 Computing $I_{>}(C_1, C_2)$ when C_1, C_2 are \mathbb{R} -SDC

Theorem 3.1.1. Suppose $C_1, C_2 \in S^n$ are \mathbb{R} -SDC and C_2 is nonsingular and $\lambda_1, \lambda_2, \ldots, \lambda_k$ are the k eigenvalues of $C_2^{-1}C_1$, where $\lambda_1 > \lambda_2 > \ldots > \lambda_k$.

- 1. If $C_2 > 0$ then $I_{\geq}(C_1, C_2) = [-\lambda_k, +\infty);$
- 2. If $C_2 < 0$ then $I_{\geq}(C_1, C_2) = (-\infty, -\lambda_1];$
- 3. If C_2 is indefinite then
 - (i) if $B_1, B_2, \dots, B_t > 0$ and $B_{t+1}, \dots, B_k < 0$ for some $t \in \{1, 2, \dots, k\}$, then $I_{\geq}(C_1, C_2) = [-\lambda_t, -\lambda_{t+1}]$,

- (*ii*) if $B_1, B_2, \dots, B_{t-1} > 0$, B_t is indefinite and B_{t+1}, \dots, B_k < 0, then $I_{\geq}(C_1, C_2) = \{-\lambda_t\},$
- (iii) in other cases, that is either B_i, B_j are indefinite for some $i \neq j$ or $B_i < 0, B_j > 0$ for some i < j or B_i is indefinite and $B_j > 0$ for some i < j, then $I_{\geq}(C_1, C_2) = \emptyset$.

Theorem 3.1.2. Suppose $C_1, C_2 \in S^n$ are \mathbb{R} -SDC, C_2 is singular and C_1 is nonsingular. Then

- (i) there always exists a nonsingular matrix U such that $U^T C_2 U = diag(B_1, 0), U^T C_1 U = diag(A_1, A_3)$, where B_1, A_1 are symmetric of the same size, B_1 is nonsingular;
- (ii) if $A_3 > 0$ then $I_{\geq}(C_1, C_2) = I_{\geq}(A_1, B_1)$. Otherwise, $I_{\geq}(C_1, C_2) = \emptyset$.

For any $C_1,C_2\in \mathcal{S}^n,$ there always exists a nonsingular matrix U such that

$$\tilde{C}_2 = U^T C_2 U = \begin{pmatrix} B_1 & 0_{p \times r} \\ 0_{r \times p} & 0_{r \times r} \end{pmatrix}; \\ \tilde{C}_1 = U^T C_1 U = \begin{pmatrix} A_1 & A_2 \\ A_2^T & 0_{r \times r} \end{pmatrix}$$
(3.1)

or

$$\tilde{C}_{1} = U^{T}C_{1}U = \begin{pmatrix} A_{1} & 0_{p \times s} & A_{2} \\ 0_{s \times p} & A_{3} & 0_{s \times (r-s)} \\ A_{2}^{T} & 0_{(r-s) \times s} & 0_{(r-s) \times (r-s)} \end{pmatrix},$$
(3.2)

where A_3 is a nonsingular diagonal matrix; $p, r, s \ge 0, p + r = n$.

Now suppose that C_1, C_2 are \mathbb{R} -SDC, without loss of generality we always assume that \tilde{C}_2, \tilde{C}_1 are already \mathbb{R} -SDC. That is

$$\tilde{C}_2 = U^T C_2 U = \text{diag}(B_1, 0), \tilde{C}_1 = U^T C_1 U = \text{diag}(A_1, 0) \quad (3.3)$$

or

$$\tilde{C}_2 = U^T C_2 U = \operatorname{diag}(B_1, 0), \tilde{C}_1 = U^T C_1 U = \operatorname{diag}(A_1, A_4), \quad (3.4)$$

where A_1, B_1 are of the same size and diagonal, B_1 is nonsingular.

- **Theorem 3.1.3.** (i) If \tilde{C}_2 , \tilde{C}_1 take the form (3.3), then $I_{\geq}(C_1, C_2)$ = $I_{\geq}(A_1, B_1)$;
 - (ii) If \tilde{C}_2, \tilde{C}_1 take the form (3.4), then $I_{\geq}(C_1, C_2) = I_{\geq}(A_1, B_1)$ if $A_4 \geq 0$ and $I_{\geq}(C_1, C_2) = \emptyset$ otherwise.

3.1.2 Computing $I_{\geq}(C_1, C_2)$ when C_1, C_2 are not \mathbb{R} -SDC

Theorem 3.1.4. Let $C_1, C_2 \in S^n$ be as in Lemma 3.1.2 and C_1, C_2 are not \mathbb{R} -SDC. The followings hold.

- (i) if $C_1 \ge 0$ then $I_{\ge}(C_1, C_2) = \{0\};$
- (ii) if $C_1 \geq 0$ and there is a real eigenvalue λ_l of $C_2^{-1}C_1$ such that $C_1 + (-\lambda_l)C_2 \geq 0$ then $I_{\geq}(C_1, C_2) = \{-\lambda_l\};$

(iii) if (i) and (ii) do not occur then $I_{>}(C_1, C_2) = \emptyset$.

Theorem 3.1.5. Let $C_1, C_2 \in S^n$ be not \mathbb{R} -SDC. Suppose C_1 is nonsingular and $C_1^{-1}C_2$ has real Jordan normal form diag $(J_1, \ldots, J_r, J_{r+1}, \ldots, J_m)$, where J_1, \ldots, J_r are Jordan blocks corresponding to real eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_r$ of $C_1^{-1}C_2$ and J_{r+1}, \ldots, J_m are Jordan blocks for pairs of complex conjugate roots $\lambda_i = a_i \pm \mathbf{i} b_i, a_i, b_i \in \mathbb{R}, i = r + 1, r + 2, \ldots, m$ of $C_1^{-1}C_2$.

- (i) If $C_1 \ge 0$ then $I_\ge (C_1, C_2) = \{0\};$
- (ii) If $C_1 \geq 0$ and there is a real eigenvalue $\lambda_l \neq 0$ of $C_1^{-1}C_2$ such that $C_1 + \left(-\frac{1}{\lambda_l}\right)C_2 \geq 0$ then $I_{\geq}(C_1, C_2) = \left\{-\frac{1}{\lambda_l}\right\};$

(iii) If cases (i) and (ii) do not occur then $I_{\geq}(C_1, C_2) = \emptyset$.

Theorem 3.1.6. Given $C_1, C_2 \in S^n$ are not \mathbb{R} -SDC and singular such that \tilde{C}_1 and \tilde{C}_2 take the forms in either (3.1) or (3.2) with $A_2 \neq$ 0. Suppose that $I_{\geq}(A_1, B_1) = [a, b], a < b$. Then, if $a \notin I_{\geq}(C_1, C_2)$ and $b \notin I_{\geq}(C_1, C_2)$ then $I_{\geq}(C_1, C_2) = \emptyset$.

3.2 Solving the quadratically constrained quadratic programming

A QCQP problem with m constraints takes the following format

(P_m)
$$\min_{\substack{f_0(x) = x^T C_0 x + a_0^T x \\ \text{s.t.} \quad f_i(x) = x^T C_i x + a_i^T x + b_i \leq 0, \ i = 1, 2, \dots, m, }$$

where $C_i \in S^n$, $x, a_i \in \mathbb{R}^n$ and $b_i \in \mathbb{R}$. We show that if C_0, C_1, \ldots, C_m are \mathbb{R} -SDC, (P_m) is relaxed to convex second-order cone problem

(SP_m)
$$\begin{array}{ll} \min & f_0(y,z) = \alpha_0^T z + \xi_0^T y \\ \text{s.t.} & f_i(y,z) = \alpha_i^T z + \xi_i^T y + b_i \leqslant 0, \ i = 1, 2, \dots, m, \\ & y_j^2 \leqslant z_j, j = 1, 2, \dots, n. \end{array}$$

$$(3.5)$$

 (SP_m) can be solved in polynomial time by the interior algorithm [21].

3.3 Applications for maximizing a sum of generalized Rayleigh quotients

We consider the following simplest case of the sum:

$$\max_{x \neq 0} \frac{x^T A_1 x}{x^T B_1 x} + \frac{x^T A_2 x}{x^T B_2 x},\tag{3.6}$$

where $B_1 > 0, B_2 > 0$.

By a change of variables, (3.6) is reduced to

$$\max_{\|y\|=1} y^T Dy + \frac{y^T Ay}{y^T By}, \ B > 0.$$
(3.7)

Theorem 3.3.1. ([72]) If A, B, D are \mathbb{R} -SDC by an orthogonal congruence matrix then (3.7) is reduced to a one-dimensional maximization problem over a closed interval.

Conclusions

In this dissertation, the SDC problem of Hermitian matrices and real symmetric matrices has been dealt with. The results obtained in the dissertation are not only theoretical but also algorithmic. On one hand, we proposed necessary and sufficient SDC conditions for a set of arbitrary number of either Hermitian matrices or real symmetric matrices. We also proposed a polynomial time algorithm for solving the Hermitian SDC problem, together with some numerical tests in MATLAB to illustrate for the main algorithm. The results in this part immediately hold for real Hermitian matrices, which is known as a long-standing problem posed in [30]. In addition, the main algorithm in this part can be applied to solve the SDC problem for arbitrarily square matrices by splitting the square matrices up into Hermitian and skew-Hermitian parts. On the other hand, we developed Jiang and Li' technique [37] for two real symmetric matrices to apply for a set of arbitrary number of real symmetric matrices.

- 1. Results on the SDC problem of Hermitian matrices.
 - Proposed an algorithm for solving the SDC problem of commuting Hermitian matrices (Algorithm 3);
 - Solved the SDC problem of Hermitian matrices by maxrank method (please see Theorem 2.1.4, Algorithm 4);
 - Proposed a Schmüdgen-like method to find the maximum rank of a Hermitian matrix-pencil (please see Theorem 2.1.2 and Algorithm 2);
 - Proposed equivalent SDC conditions of Hermitian matrices linked with the existence of a positive definite matrix satisfying a system of linear equations (Theorem 2.1.5);
 - Proposed an algorithm for completely solving the SDC problem of complex or real Hermitian matrices (please

see Algorithm 6).

- 2. Results on the SDC problem of real symmetric matrices.
 - Proposed necessary and sufficient SDC conditions for a collection of real symmetric matrices to be SDC (please see Theorem 2.2.2 for nonsingular collection and Theorem 2.2.3 for singular collection). These results are completeness and generalizations of Jiang and Li's method for two matrices [37];
 - Proposed an inductive method for solving the SDC problem of a singular collection. This method helps to move from study the SDC of a singular collection to study the SDC of a nonsingular collection of smaller dimension as shown in Theorem 2.2.3. Moreover, we realize that a result by Jiang and Li [37] is not complete. A missing case not considered in their paper is now added to make it up in the dissertation, please see Lemma 1.2.7 and Theorem 1.2.1;
 - Proposed algorithms for solving the SDC problems of nonsingular and singular collection (Algorithm 7 and Algorithm 8, respectively).
- 3. We apply above SDC results for dealing with the following problems.
 - Computed the positive semidefinite interval of matrix pencil $C_1 + \mu C_2$ (please see Theorems 3.1.1, 3.1.2, 3.1.3, 3.1.4, 3.1.5 and 3.1.6);
 - Applied the positive semidefinite interval of matrix pencil for completely solving the GTRS (please see Theorems 3.2.1, 3.2.2);
 - Solved the homogeneous QCQP problems, the maximization of a sum of generalized Rayleigh quotients under the SDC of involved matrices.

Future research

The SDC problem has been completely solved on the field of real numbers \mathbb{R} and complex numbers \mathbb{C} . A natural question to aks is whether the obtained SDC results are remained true on a finite field? on a commutative ring with unit? Moreover, as seen, the SDC conditions seem to be very strict. That is, not too many collections can satisfy the SDC conditions. This raises a question that how much disturbance on the matrices such that a not SDC collection becomes SDC? Those unsloved problems suggest our future research as follows.

- 1. Studying the SDC problems on a finite field, on a commutative ring with unit;
- Studying the approximately simultaneous diagonalization via congruence of matrices. This problem can be stated as follows: Suppose the matrices C₁, C₂,..., C_m, are not SDC. Given ε > 0, whether there are matrices E_i with ||E_i|| < ε such that C₁ + E₁, C₂ + E₂,..., C_m + E_m are SDC?

Some results on approximately simultaneously diagonalizable matrices for two real matrices and for three complex matrices can be found in [50], [61], [68].

3. Explore applications of the SDC results.

List of Author's Related Publication

- V. B. Nguyen, T. N. Nguyen, R.L. Sheu (2020), "Strong duality in minimizing a quadratic form subject to two homogeneous quadratic inequalities over the unit sphere", J. Glob. Optim., 76, pp. 121-135.
- T. H. Le, **T. N. Nguyen** (2022), "Simultaneous Diagonalization via Congruence of Hermitian Matrices: Some Equivalent Conditions and a Numerical Solution", SIAM J. Matrix Anal. Appl., 43, Iss. 2, pp. 882-911.
- V. B. Nguyen, T. N. Nguyen (2024), "Positive semidefinite interval of matrix pencil and its applications to the generalized trust region subproblems", Linear Algebra Appl., 680, pp. 371-390.
- T. N. Nguyen, V. B. Nguyen, T. H. Le, R. L. Sheu, "Simultaneous Diagonalization via Congruence of *m* Real Symmetric Matrices and Its Implications in Quadratic Optimization", Preprint.