1996

Structured eigenvectors, interlacing, and matrix completions

Brenda K. Kroschel
College of William & Mary - Arts & Sciences

STRUCTURED EIGENVECTORS, INTERLACING, AND MATRIX COMPLETIONS

A Dissertation Presented to the

Applied Science Department of

The College of William and Mary

In Partial Fulfillment of the

Requirements for the Degree

Doctor of Philosophy

by

Brenda K. Kroschel

April 1996

UMI Number: 9701100

UMI
300 North Zeeb Road
Ann Arbor, MI 48103

APPROVAL SHEET

This dissertation is submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

Brenda K. Kroschel

Brenda K. Kroschel, Author

Approved, April 1996

Charles R. Johnson

Marc Sher, Physics

Ilya Spitkovsky

David Stanford

Hugo J. Woerdeman

Richard O. Hill
Department of Mathematics
Michigan State University

This dissertation is lovingly dedicated to my husband, Dave, without whose love and support this work would not be possible, and to my daughter, Nicole, who was not around at the beginning of this endeavor, but who makes everything more meaningful.

Contents

ACKNOWLEDGMENTS

I would like to express my appreciation to my advisor, Dr. Charles R. Johnson, whose excitement for Linear Algebra is contagious. I thank him for his patient guidance and careful mentoring throughout my studies. In addition, I would like to thank Dr. Hugo J. Woerdeman for his careful reading of this dissertation and his assistance with documentation. I would also like to thank Dr. Richard O. Hill and Catherine A. Jones with whom some of this dissertation is joint work.

List of Figures

ABSTRACT

This dissertation presents results from three areas of applicable matrix analysis: structured eigenvectors, interlacing, and matrix completion problems. Although these are distinct topics, the structured eigenvector results provide connections.

It is a straightforward matrix calculation that if λ is an eigenvalue of A, x an associated structured eigenvector and α the set of positions in which x has nonzero entries, then λ is also an eigenvalue of the submatrix of A that lies in the rows and columns indexed by α. We present a converse to this statement and apply the results to interlacing and to matrix completion problems. Several corollaries are obtained that lead to results concerning the case of equality in the interlacing inequalities for Hermitian matrices, and to the problem of the relationship among eigenvalue multiplicities in various submatrices.

Classical interlacing for an Hermitian matrix A may be viewed as describing how many eigenvalues of A must be captured by intervals determined by eigenvalues of a principal submatrix of A. We generalize the classical interlacing theorems by using singular values of off-diagonal blocks of A to construct extended intervals that capture a larger number of eigenvalues. The union of pairs of intervals is also discussed, and applications are mentioned.

The matrix completion results that we present include the positive semidefinite cycle completion problem for matrices with data from the complex numbers, distance matrix cycle completability conditions, the P-matrix completion problem, and the totally nonnegative completion problem. We show that the positive semidefinite cycle completion problem for matrices with complex data is a special case of a larger real positive semidefinite completion problem. In addition, we characterize those graphs for which the cycle conditions on all minimal cycles imply that a partial distance matrix has a distance matrix completion. We also prove that every combinatorially symmetric partial P-matrix has a P-matrix completion and we characterize the class of graphs for which every partial totally nonnegative matrix has a totally nonnegative completion. The structured eigenvector results are used to give a new proof of the the maximum minimum eigenvalue completion problem for partial Hermitian matrices with a chordal graph.

Structured Eigenvectors, Interlacing, and Matrix Completions

Chapter 1

Introduction

This dissertation presents results from three areas of matrix theory: structured eigenvectors (chapter 2), interlacing (chapter 3), and matrix completions (chapter 4). Although these are distinct topics in matrix theory, the structured eigenvector results of chapter 2 provide connections. We will begin by providing some notation and background material followed by brief introductions to the chapters. More detailed background and introductions are provided within each of the chapters.

1.1 Notation and Matrix Theoretic Background

The set of all m-by-n matrices with entries from a field F will be denoted by $M_{m,n}(F)$, and if $m = n$, $M_{n,n}(F)$ will be abbreviated to $M_n(F)$. If $F = \mathbb{C}$, the complex numbers, we will often shorten this notation to $M_{m,n}$. For $A \in M_{m,n}(F)$ the notation $A = (a_{ij})$ will indicate that the entries of A are $a_{ij} \in F$, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$.

For $A \in M_{m,n}(F)$, $\alpha \subseteq \{1, 2, \ldots, m\}$, and $\beta \subseteq \{1, 2, \ldots, n\}$, the submatrix of A lying in the rows indexed by α and the columns indexed by β will be denoted

$A[\alpha;\beta]$. Similarly, $A(\alpha;\beta)$ is the matrix that results from the deletion of the rows indexed by α and the columns indexed by β. If $A \in M_n(F)$ and $\alpha = \beta$, then the principal submatrix $A[\alpha;\alpha]$ is abbreviated to $A[\alpha]$ and the complementary principal submatrix is $A(\alpha)$. In the same fashion, for a vector $x \in F^n$, $x[\alpha]$ denotes the entries of x in the positions indexed by α and $x(\alpha)$ denotes the complementary vector. We will often denote the sets $\{1, 2, \ldots, m\}$ and $\{1, 2, \ldots, n\}$ by M and N, respectively. For $\alpha, \beta \subseteq N$ the set difference $\alpha - \beta$ denotes the set of all elements in α that are not in β. The set $N - \alpha$ will also be denoted by α^c and $N - \beta$ by β^c. Note that this means that $A(\alpha;\beta) = A[\alpha^c;\beta^c]$. The notation $\alpha + \{n\}$ will indicate the set that results from adding n to every element of α. For example, if $\alpha = \{1, 3, 4\}$ and $n = 4$, then $\alpha + \{n\} = \{5, 7, 8\}$. The cardinality of a set α will be denoted by $|\alpha|$.

1.1.1 Eigenvectors and Eigenvalues

For an n-by-n matrix A with entries from a field F, the nonzero vector x is a *right eigenvector* of A associated with λ if $Ax = \lambda x$ for some scalar λ. Similarly, $y \neq 0$ is a *left eigenvector* if $y^* A = \lambda y^*$. The scalar λ is an *eigenvalue* of A and $\sigma(A)$ will denote the set of all eigenvalues of A, some of which may lie only in an extension field of F. The subspace of F^n spanned by the set of all eigenvectors of A associated with λ is called the *eigenspace of A associated with* λ. The dimension of the eigenspace of A associated with λ is the *geometric multiplicity* of λ as an eigenvalue of A and is denoted throughout by $g_\lambda(A)$. The *principal of biorthogonality* (see, e.g. theorem 1.4.7 [HJ1]) says that if $\lambda, \mu \in \sigma(A)$ with $\lambda \neq \mu$, then any left eigenvector of A corresponding to μ is orthogonal to any right eigenvector of A corresponding to λ.

1.1.2 Classical Interlacing Inequalities

An important result concerning the eigenvalues of Hermitian matrices is the interlacing eigenvalues theorem for bordered matrices (see, e.g. theorem 4.3.8 [HJ1].) Let $A \in M_n$ be Hermitian and $\alpha \subset N$ be such that $|\alpha| = n-1$ and let the ordered eigenvalues of A be $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ and those of $A[\alpha]$ be $\mu_1 \leq \mu_2 \leq \cdots \leq \mu_{n-1}$. Then

$$\lambda_1 \leq \mu_1 \leq \lambda_2 \leq \cdots \leq \lambda_{n-1} \leq \mu_{n-1} \leq \lambda_n.$$

That is, the eigenvalues of an Hermitian matrix and any of its $(n-1)$-by-$(n-1)$ principal submatrices "interlace." Another way that classical interlacing may be viewed is that each interval $[\lambda_i, \lambda_{i+1}]$, $i = 1, 2, \ldots, n-1$ contains at least one eigenvalue μ_t, $1 \leq t \leq n-1$ from every $(n-1)$-by-$(n-1)$ principal submatrix of A. In addition, for $j \neq i$ the interval $[\lambda_j, \lambda_{j+1}]$ captures μ_s in which $s \neq t$. However, we may also say that every interval $[\mu_i, \mu_{i+1}]$, $i = 1, 2, \ldots, n-2$ contains a different eigenvalue of A (the other two eigenvalues are in $(-\infty, \mu_1]$ and $[\mu_{n-1}, \infty)$). As interlacing is applied to successively smaller principal submatrices we find that, if $\hat{\lambda}_1 \leq \hat{\lambda}_2 \leq \cdots \leq \hat{\lambda}_{n-p}$ are the eigenvalues of a principal $(n-p)$-by-$(n-p)$ submatrix of A, then the interval $\left[\hat{\lambda}_i, \hat{\lambda}_j\right]$, $0 < i \leq j \leq n-p$ contains at least $j-i-p+1$ eigenvalues of A.

1.1.3 The Singular Value Decomposition

Another important result is the singular value decomposition of a matrix (see, e.g. theorem 7.3.5 [HJ1]). For $m \geq n$, if the matrix $A \in M_{m,n}$ then A may be written in the form $A = V\Sigma W^*$ in which $V \in M_m$, $W \in M_n$ are unitary, and

$$\Sigma = \begin{bmatrix} diag(\sigma_1, \sigma_2, \ldots, \sigma_n) \\ 0 \end{bmatrix}$$

($diag(d_1, d_2, \ldots, d_n)$ denotes the diagonal matrix with the d_i on the diagonal). The values σ_i are the *singular values* of A and are the nonnegative square roots of the eigenvalues of AA^*. Usually the singular values are ordered in a nonincreasing fashion. i.e. $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$. Note that when it is convenient we will think of A as having singular values $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq \sigma_{n+1} = \cdots = \sigma_m = 0$. The columns of V are unit eigenvectors of AA^* and are called the *left singular vectors* of A. Similarly, the columns of W are unit eigenvectors of A^*A and are called the *right singular vectors* of A. The case in which $n \geq m$ is analogous for $\Sigma = \begin{bmatrix} diag(\sigma_1, \sigma_2, \ldots, \sigma_m) & 0 \end{bmatrix}$.

The matrix AA^* is Hermitian and since the singular values are the nonnegative square roots of the eigenvalues of AA^* it follows that there is also an interlacing theorem for the singular values of a matrix. Let $A \in M_{m,n}$ be given and let $\hat{A}$ be the matrix obtained by deleting one column from A. For $m \geq n$ let $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$ be the singular values of A and $\hat{\sigma}_1 \geq \hat{\sigma}_2 \geq \cdots \geq \hat{\sigma}_{n-1} \geq 0$ the singular values of $\hat{A}$. Then

$$\sigma_1 \geq \hat{\sigma}_1 \geq \sigma_2 \geq \hat{\sigma}_2 \geq \cdots \geq \hat{\sigma}_{n-1} \geq \sigma_n \geq 0.$$

There are analogous singular value interlacing results for $m < n$ and if a row of A is deleted instead of a column.

Often times there are close analogues to eigenvalue results for the singular values. This is due to the fact that for the matrix $A \in M_{m,n}$ the matrix

$$\begin{bmatrix} 0 & A \\ A^* & 0 \end{bmatrix}$$

is Hermitian and has eigenvalues that are plus and minus the singular values of A (with possibly some extra zeros). This is a fact dating back to Wielandt [Wi] and is used in both chapters 2 and 3 of this dissertation to generalize the main results of these chapters.

1.1.4 Jacobi's Identity

If $A \in M_n(F)$ is nonsingular, then the minors of A^{-1} are related to those of A by Jacobi's identity. Jacobi's identity states (see, e.g. section 0.8.4 [HJ1]) that for $\alpha, \beta \subseteq N$, both nonempty, in which $|\alpha| = |\beta|$

$$\det A^{-1}[\alpha;\beta] = (-1)^{s(\alpha)+s(\beta)} \frac{\det A[\beta^c;\alpha^c]}{\det A} \tag{1.1}$$

in which $s(\alpha) = \sum_{j\in\alpha} j$. Observe that if α and β have cardinality 1, i.e. $\alpha = \{i\}$, $\beta = \{j\}$, $1 \le i, j \le n$, then (1.1) becomes

$$a_{ij}^{-1} = (-1)^{i+j} \frac{\det A[\beta - \{j\}; \alpha - \{i\}]}{\det A}$$

in which a_{ij}^{-1} denotes the i,j entry of A^{-1}. This expression is the adjoint formula for the inverse of a matrix. Thus, Jacobi's identity is a generalization of the adjoint formula.

1.1.5 The Schur Complement

For $\emptyset \neq \alpha \subseteq N$ and $A \in M_n$, if $A[\alpha]$ is nonsingular, then the *Schur complement of* $A[\alpha]$ *in* A is the matrix

$$A[\alpha^c] - A[\alpha^c;\alpha](A[\alpha])^{-1}A[\alpha;\alpha^c].$$

Let A be partitioned as

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

in which A_{11} is nonsingular. Then

$$\begin{bmatrix} I & 0 \\ -A_{21}A_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} I & -A_{11}^{-1}A_{12} \\ 0 & I \end{bmatrix} = \begin{bmatrix} A_{11} & 0 \\ 0 & S \end{bmatrix}$$

in which S is the Schur complement of A_{11} in A. It is easy to see that A is then nonsingular if and only if S is nonsingular and $\det A = \det A_{11} \det S$. Moreover, A is positive (semi)definite if and only if S is positive (semi)definite. For more information on Schur complements see [C].

1.2 Graph Theoretic Background

Many of the graph theoretic definitions and notation used in this dissertation are adopted from [G]. An *undirected graph* is a pair $G = (V, E)$ in which V is a finite set called the *vertex set* and the set, E, called the *edge set*, contains unordered pairs of elements from V called the *edges* of G. For $v_i, v_j \in V$ the edges $\{v_i, v_j\} \in E$ will often be abbreviated to v_iv_j. If the graph is a directed graph the edge set contains ordered pairs. In this dissertation all of the graphs will be undirected. So, *graph* will mean an undirected graph without loops or multiple edges. If $G = (V, E)$ is a graph and $v_iv_j \in E$ then v_i and v_j are said to be *adjacent*. A graph is said to be *complete* if every vertex is adjacent to every other vertex. The complete graph on n vertices is denoted K_n.

If $G = (V, E)$ and $H = (W, F)$ are graphs for which $W \subseteq V$ and $F \subseteq E$, then H is a *subgraph* of G. If $W \subseteq V$ then the subgraph $G_W = (W, E_W)$ in which $E_W = \{v_iv_j \in E : v_i, v_j \in W\}$ is called the *subgraph of G induced by W*. An induced subgraph is called a *clique* if it is a complete graph. A clique that is not properly contained in any other clique is called *maximal*. The graph $H = (V, F)$ is called a *supergraph* of $G = (V, E)$ if $F \supseteq E$.

A *path* in a graph $G = (V, E)$ is a sequence of vertices $(v_1, v_2, \ldots, v_k)$ such that $v_iv_{i+1} \in E$. The graph is said to be *connected* if there is a path from every vertex to every other vertex. If a graph is not connected, then each maximal connected induced subgraph is called a *component* of the graph.

A path that begins and ends with the same vertex, i.e. $(v_1, v_2, \ldots, v_k, v_1)$, $k \geq 3$, is called a *cycle* and a *simple cycle* is a cycle for which the vertices $v_1, v_2, \ldots, v_k$ are distinct. In this dissertation the term cycle will mean a simple cycle. The cycle $(v_1, v_2, \ldots, v_k, v_1)$ is called a *k-cycle* and is denoted by C_k. A *chord* of a cycle C_k, $k \geq 4$, in a graph G is an edge of G between two nonconsecutive vertices of C_k. A *chordless cycle* is a simple cycle that has no chords. A *minimal cycle* in a graph G is an induced subgraph of G that is a chordless cycle. A connected graph with no cycles is called a *tree*. A *k-wheel* is a $(k-1)$-cycle with one additional vertex that is adjacent to every vertex of the cycle and is denoted by W_k.

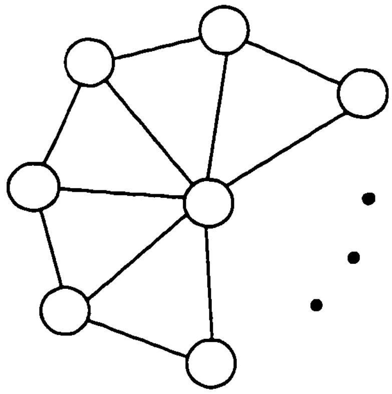
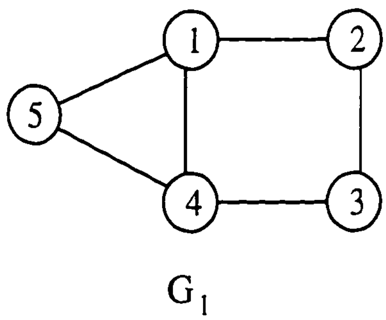


Figure 1.1: Wheel, W_k

If a graph contains no chordless cycles of length 4 or more, then the graph is said to be *chordal.* For example,



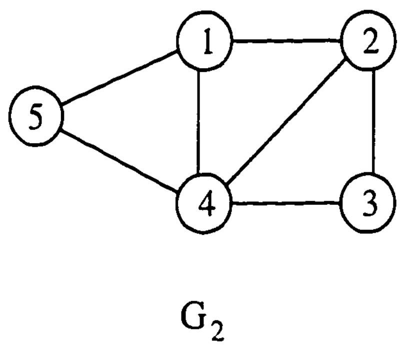


Figure 1.2: Example of a non-chordal (G_1) and a chordal graph (G_2).

The graph G_1 is not chordal since $(1, 2, 3, 4, 1)$ is a cycle of length 4 that does not contain a chord. However, this same cycle in the graph G_2 contains the chord $\{2, 4\}$ and all other cycles of length 4 or more also contain a chord. Chordal graphs have received considerable attention largely due to their importance in the study of perfect elimination schemes for Gaussian elimination [G]. This class of graphs is also very important in the study of matrix completion problems as will be discussed in the next section.

Graphs are often used to represent the zero/nonzero structure of a matrix. For an n-by-n Hermitian (or symmetric) matrix A, the graph $G(A) = (N, E)$ is the graph on n vertices for which $\{i, j\}$ is an edge of $G(A)$ exactly when $a_{ij} \neq 0$ for $i \neq j$ (loops are omitted by convention, so whether or not a_{ii} is zero is irrelevant.) For example, the matrix

$$A = \begin{bmatrix} 1 & 2 & -3 & 0 \\ 2 & -1 & 4 & 0 \\ -3 & 4 & 0 & 2 \\ 0 & 0 & 2 & 3 \end{bmatrix}$$

has the associated graph

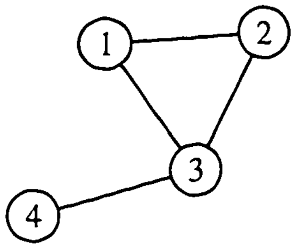


Figure 1.3: $G(A)$

A new graph may be constructed from a given graph G by using one of several operations on the edges and/or vertices of G. An *edge subdivision* of a graph G on n vertices is a graph G' on $n+1$ vertices that results from replacing an edge of G with two edges and a vertex between:

Figure 1.4: Edge subdivision.

A *vertex partition* of G (n vertices) is a graph G' ($n+1$ vertices) in which a vertex (of degree at least 1) in G is replaced by two adjacent vertices that partition the neighbors of the original vertex: e.g.

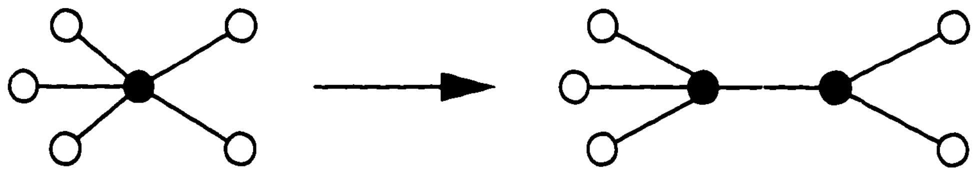


Figure 1.5: Vertex partition.

We say that a graph G_2 is *homeomorphic* to a graph G_1 if G_2 may be obtained from G_1 via a finite sequence of (at least one) edge subdivisions. The graph G_2 is *built from* G_1 if G_2 may be obtained from G_1 via a finite sequence of (at least one) vertex partitions. Note that edge subdivision is a special case of vertex partitioning, so that homeomorphism is a special case of "built from." These operations will be used in chapter 4 in the distance matrix completion problem.

1.3 Partial Matrices

A *partial matrix* is one in which some entries are specified over a field F, while the remainder of the entries are unspecified and free to be chosen from F. It will be assumed throughout this dissertation that the diagonal entries are specified. In many cases this is a natural assumption. A *combinatorially symmetric partial matrix* A is one in which a_{ji} is specified whenever a_{ij} is specified. For example,

$$A = \begin{bmatrix} 1 & 0 & 2 & ? \\ 1 & -1 & 1 & ? \\ -2 & 3 & 0 & 5 \\ ? & ? & 1 & -4 \end{bmatrix}$$

in which '?' indicates an unspecified position, is a combinatorially symmetric partial matrix. A *partial symmetric (Hermitian) matrix* is a combinatorially symmetric partial matrix in which if a_{ij} is specified, then $a_{ji} = a_{ij}$ $(a_{ji} = \bar{a}_{ij})$. The specified positions of an n-by-n combinatorially symmetric partial matrix A may be described by an undirected graph $G(A)$ on n vertices in which there is an edge between vertex i and vertex j exactly when a_{ij} is specified. Loops are omitted by convention.

A *completion* of a partial matrix is a choice of values for the unspecified entries resulting in a conventional matrix. A *matrix completion problem* asks when a given partial matrix has a completion with some desired property.

1.4 Overview

By structured eigenvectors we mean eigenvectors that have 0's in (at least) certain specified positions. If the n-by-n matrix A has a structured eigenvector associated with λ of the form $\begin{bmatrix} x \\ 0 \end{bmatrix}$, $x \in F^k$, $0 < k < n$, then by partitioned matrix multiplication, λ is also an eigenvalue of $A[\alpha]$ in which $\alpha = \{1, 2, \ldots, k\}$, with associated eigenvector x. Note that eigenvalues are invariant under permutation similarity, so, there is no loss of generality in assuming $\alpha = \{1, 2, \ldots, k\}$. In chapter 2 we seek a converse: if λ is an eigenvalue of both $A[\alpha]$ and of A, is there a structured eigenvector of A associated with λ of the form $\begin{bmatrix} x \\ 0 \end{bmatrix}$, $x \in F^k$, $0 < k < n$? As will be shown, the number of linearly independent left eigenvectors of this special form plus the number of linearly independent right eigenvectors with the given structure is at least $g_1 + g_2 - k$ in which g_1 and g_2 are the geometric multiplicities of λ as an eigenvalue of A and

$A[\alpha]$, respectively. This result has remarkably many implications as will be seen in chapters 3 and 4.

The structured eigenvector results of chapter 2 are used in chapter 3 to characterize what we call interlacing diagrams that describe the relationships of the geometric multiplicities of a given eigenvalue among various submatrices. In chapter 3 we also use singular values to extend classical interlacing intervals. If $\lambda_1(B) \leq \lambda_2(B) \leq \cdots \leq \lambda_{n-p}(B)$ are the eigenvalues of the $(n-p)$-by-$(n-p)$ submatrix, B, of the matrix

$$A = \begin{bmatrix} B & C \\ C^* & D \end{bmatrix},$$

then the singular values of the off diagonal block, C, are used to extend the classical interlacing intervals $[\lambda_i(B), \lambda_j(B)]$, $0 < i \leq j \leq n-p$ and possibly capture more than the $j-i-p+1$ eigenvalues of A that classical interlacing insures. The main result of chapter 3 says that the interval $\left[t-\sqrt{\delta^2+\sigma_k^2}, t+\sqrt{\delta^2+\sigma_k^2}\right]$ in which σ_k is the k^{th} largest singular value of C, $t = \frac{\lambda_j(B)+\lambda_i(B)}{2}$, and $\delta = \frac{\lambda_j(B)-\lambda_i(B)}{2}$, captures at least $j-i-k+2$ eigenvalues of A. If B is $(n-1)$-by-$(n-1)$, and $k=2$, then $\sigma_2 = 0$ and the interval becomes a classical interlacing interval. In this case, for $j = i+1$ the results of chapter 3 say that the interval contains at least $(i+1)-i-2+2 = 1$ eigenvalue of A. Thus, classical interlacing is a special case of the main result of chapter 3.

In chapter 4 we present results concerning a variety of matrix completion problems. Among these are the positive definite cycle completion problem for matrices with data from the complex numbers, distance matrix cycle completability conditions, the P-matrix completion problem, and the totally nonnegative completion problem. We show that the positive definite cycle completion problem for matrices with complex data is a special case of a larger real positive definite completion problem. In addition we characterize those graphs for which the cycle conditions on all minimal cycles imply that a partial distance matrix has a distance matrix completion. We also prove that every combinatorially symmetric partial P-matrix has a P-matrix completion and we characterize the class of graphs for which every partial totally nonnegative matrix, the graph of whose specified entries is in the class, has a totally nonnegative completion. The structured eigenvector results of chapter 2 are used to solve the maximum minimum eigenvalue completion problem for partial Hermitian matrices with a chordal graph in a new way.

Chapter 2

Structured Eigenvectors

It is a straightforward partitioned matrix calculation that if λ is an eigenvalue of A, x an associated eigenvector, and α the set of positions in which x has entries not equal to zero, then λ is also an eigenvalue of $A[\alpha]$. Converses to this statement are known in certain special situations. For example, it has been known for some time (and can be deduced from theorem 5 [D1]; see also [JK1]) that if $A \in M_n(\mathbb{C})$ is Hermitian, $|\alpha| = n-1$, and $\lambda \in \mathbb{R}$ is an eigenvalue of both A and $A[\alpha]$, i.e. a case of equality in the interlacing inequalities, then there is an eigenvector $x = (x_1, x_2, \ldots, x_n)^T$ of A associated with λ such that if $i \notin \alpha$ then $x_i = 0$. For a general matrix $A \in M_n(F)$ and λ an eigenvalue of A with geometric multiplicity k, the rank of $A - \lambda I$ is $n-k$. Then for $|\alpha| > n-k$ the rank of $A[\alpha] - \lambda I$ is at most $n-k$ and λ is an eigenvalue of $A[\alpha]$. Moreover, it is implicit in the proof of theorem 1.4.9 in [HJ1] that there is an eigenvector of A associated with λ all of whose components indexed by α^c are zero. It is our purpose here to give a converse to the opening statement that is, in some sense, the most general possible in terms of the data we use. A variety of statements, including those just mentioned, may then be easily recognized as special cases.

These results, as well as some special cases, will be valid over a general field F. For $\lambda \in \sigma(A)$, denote the geometric multiplicity of λ in A by $g_\lambda(A)$. The most optimistic converse to the opening statement would be that if λ is an eigenvalue of both A and $A[\alpha]$, then there is an eigenvector x (of A associated with λ) in which all components of $x(\alpha)$ are zero. However, this is not always the case. Consider

$$A = \left[\begin{array}{cc:cc} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right]$$

and the set $\alpha = \{1, 2\}$. This matrix has zero as an eigenvalue, as does $A[\alpha]$, but any eigenvector of A associated with zero is of the form $\begin{pmatrix} a & 0 & 0 & -a \end{pmatrix}^T$. The converse cannot, therefore, be as general as one might hope.

Before stating a converse, several definitions are needed. The main result will be stated in terms of the dimensions of special subspaces, of the left and right eigenspaces of a general matrix $A \in M_n(F)$ associated with $\lambda \in F$, in which the vectors have support among the components indexed by α. These special subspaces (of the eigenspaces) are defined as follows:

$$\begin{aligned} LE_\alpha^\lambda(A) &= \left\{y \in F^n : y^T A = \lambda y^T, y(\alpha) = 0\right\} \\ RE_\alpha^\lambda(A) &= \{x \in F^n : Ax = \lambda x, x(\alpha) = 0\}. \end{aligned}$$

Similarly, let $LN(A)$ and $RN(A)$ denote the left and right nullspaces of A and define the special subspaces (of the nullspaces) $LN_\alpha(A) = LE_\alpha^0(A)$ and $RN_\alpha(A) = RE_\alpha^0(A)$.

It is clear that the dimensions of all these spaces are permutation similarity invariant. and, by assuming that $\alpha = \{1, 2, \ldots, n-k\}$, this fact will be exploited repeatedly throughout this chapter without further mention. If x is any eigenvector of A associated with λ, then x is an eigenvector of $A - \lambda I$ associated with the eigenvalue zero. For this reason, results concerning the special nullspaces underlie observations concerning the special eigenspaces.

2.1 Main Result

For contrast to the main result, we note some preliminary facts that indicate circumstances under which both the left and right special subspaces are nonempty. The *rank deficiency* of an n-by-n matrix A is $n - rank(A) = g_0(A)$. For general matrices, when the rank deficiency of a principal submatrix is sufficiently large, then the dimensions of the left and right nullspaces are positive (see discussion preceding lemma 4.8.) Suppose that the submatrix $A[\alpha]$ is such that its rank deficiency is greater than the number of rows or columns deleted from A to obtain $A[\alpha]$. That is, for $|\alpha| = n-k$, $g_0(A[\alpha]) > k$. In this case, the rank of $A[\alpha]$ is $n - k - g_0(A[\alpha])$ and the rank of A can be at most $2k$ more than the rank of $A[\alpha]$. But, then, the rank deficiency of A is at least $g_0(A[\alpha]) - k$. Since this number is positive, A is rank deficient and the left and right nullspaces of A are both nonempty. The lemma below states that, in fact, the left and right *special* nullspaces of A are both nonempty.

Lemma 2.1 *Let $A \in M_n(F)$ and let $\alpha \subseteq N$ be such that $|\alpha| = n-k$.*

(i) *If $g_0(A[\alpha]) > k$, then* $\dim(LN_\alpha(A)), \dim(RN_\alpha(A)) \geq g_0(A[\alpha]) - k$.

(ii) *Let $0 \leq g_0 \leq \min\{k, |\alpha|\}$ be given. Then there is a matrix B such that $g_0(B[\alpha]) = g_0$ and* $\dim(LN_\alpha(B)) = \dim(RN_\alpha(B)) = 0$.

Proof: It will be assumed, without loss of generality, that $\alpha = \{1, 2, \ldots, n-k\}$. Then A has partitioned form

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \tag{2.1}$$

in which $A_{11} = A[\alpha]$. In this case, if $x \in RN_\alpha(A)$, it is of the form $x = \begin{bmatrix} \hat{x} \\ 0 \end{bmatrix}$ in which $\hat{x} \in F^{n-k}$ and note that $RN_\alpha(A) = RN \begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix}$. Similarly, any vector $y \in LN_\alpha(A)$ is of the form $y^T = \begin{bmatrix} \hat{y}^T & 0 \end{bmatrix}$ in which $\hat{y} \in F^{n-k}$ and $LN_\alpha(A) = LN \begin{bmatrix} A_{11} & A_{12} \end{bmatrix}$. Assume $g_0(A_{11}) > k$, as in part (i) of the lemma. By elementary linear algebra

$$\dim(RN_\alpha(A)) = \dim\left(RN\begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix}\right) = n - k - rank\begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix}. \tag{2.2}$$

Since A_{21} is k-by-$(n-k)$ the rank of A_{21} is less than or equal to $\min\{k, n-k\}$. Therefore,

$$\begin{aligned} rank\begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix} &\le rank(A_{11}) + rank(A_{21}) \\ &= n - k - g_0(A_{11}) + rank(A_{21}) \\ &\le n - k - g_0(A_{11}) + k \\ &= n - g_0(A_{11}). \end{aligned}$$

But then substituting this in (2.2) gives

$$\begin{aligned}\dim\left(RN\begin{bmatrix}A_{11}\\ A_{21}\end{bmatrix}\right) &\geq n-k-(n-g_0(A_{11}))\\ &= g_0(A_{11})-k.\end{aligned}$$

So, $\dim(RN_\alpha(A)) \geq g_0(A_{11}) - k$. The proof that

$$\dim(LN_\alpha(A)) = \dim\left(LN\begin{bmatrix}A_{11} & A_{12}\end{bmatrix}\right) \geq g_0(A_{11}) - k$$

is analogous and part (i) of the lemma is verified.

For part (ii) consider the matrix

$$B = \begin{bmatrix}B_{11} & B_{12}\\ B_{21} & B_{22}\end{bmatrix} = \begin{bmatrix}0 & 0 & \vdots & I_{g_0} & 0\\ 0 & I_{n-k-g_0} & \vdots & 0 & 0\\ \cdots & \cdots & \vdots & \cdots & \cdots\\ I_{g_0} & 0 & \vdots & * & 0\\ 0 & 0 & \vdots & 0 & \hat{B}_{22}\end{bmatrix} \tag{2.3}$$

in which B_{11} is $(n-k)$-by-$(n-k)$ and $g_0(B_{11}) = g_0$. For this matrix, $0 \leq g_0 \leq k$, but there are no nonzero vectors in either $LN_\alpha(B)$ or $RN_\alpha(B)$, and part (ii) of the lemma is also proved. □

Replacement of A with $A - \lambda I$ in lemma 2.1 gives the following.

Theorem 2.1 *Let $A \in M_n(F)$ and let $\alpha \subseteq N$ be such that $|\alpha| = n - k$.*

(i) *If $g_\lambda(A[\alpha]) > k$, then $\dim(LE_\alpha^\lambda(A)), \dim(RE_\alpha^\lambda(A)) \geq g_\lambda(A[\alpha]) - k$.*

(ii) *Let $0 \leq g_\alpha \leq \min\{k, |\alpha|\}$ be given. Then there is a matrix B such that $g_\lambda(B[\alpha]) = g_\alpha$ and $\dim(LE_\alpha^\lambda(B)) = \dim(RE_\alpha^\lambda(B)) = 0$.*

Statement (i) in theorem 2.1 is best possible when left and right eigenspaces are considered separately. By considering the left and right eigenspaces simultaneously, one arrives at a general converse to the opening statement. This main result will first be stated in terms of the special nullspaces.

Lemma 2.2 *Let $A \in M_n(F)$ and let $\alpha \subseteq N$ be such that $|\alpha| = n - k$.*

(i) $\dim(LN_\alpha(A)) + \dim(RN_\alpha(A)) \geq g_0(A) + g_0(A[\alpha]) - k$.

(ii) *Let g and g_α such that $0 \leq g \leq n$, $0 \leq g_\alpha \leq |\alpha|$, and $|g - g_\alpha| \leq k$ be given. Then, if $g + g_\alpha - k > 0$ there is a matrix B such that $g_0(B) = g$, $g_0(B[\alpha]) = g_\alpha$ and*

$$\dim(LN_\alpha(B)) + \dim(RN_\alpha(B)) = g_0(B) + g_0(B[\alpha]) - k.$$

If $g + g_\alpha - k \leq 0$, then there is a matrix B, with the given parameters, such that

$$\dim(LN_\alpha(B)) = \dim(RN_\alpha(B)) = 0.$$

Proof: Let $A \in M_n(F)$ and partition A as in (2.1). Also, let

$$\begin{aligned} i_r &= rank \begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix} - rank\,(A_{11}) \\ i_c &= rank \begin{bmatrix} A_{11} & A_{12} \end{bmatrix} - rank\,(A_{11}). \end{aligned} \tag{2.4}$$

Then

$$\begin{aligned}\dim\left(RN\begin{bmatrix}A_{11}\\A_{21}\end{bmatrix}\right) &= n-k-rank\begin{bmatrix}A_{11}\\A_{21}\end{bmatrix}\\ &= n-k-(rank\,(A_{11})+i_r)\\ &= g_0\,(A_{11})-i_r.\end{aligned}$$

Similarly, $\dim\left(LN\begin{bmatrix}A_{11} & A_{12}\end{bmatrix}\right) = g_0\,(A_{11}) - i_c$. The sum of the dimensions of these two spaces is then

$$\dim\left(RN\begin{bmatrix}A_{11}\\A_{21}\end{bmatrix}\right)+\dim\left(LN\begin{bmatrix}A_{11} & A_{12}\end{bmatrix}\right) = 2g_0\,(A_{11})-i_r-i_c. \tag{2.5}$$

Choose $S, T \in M_{n-k}\,(F)$ nonsingular matrices such that

$$\begin{bmatrix}S & 0\\0 & I\end{bmatrix}\begin{bmatrix}A_{11} & A_{12}\\A_{21} & A_{22}\end{bmatrix}\begin{bmatrix}T & 0\\0 & I\end{bmatrix} = \begin{bmatrix}0 & 0 & \vdots & Y_1\\0 & I & \vdots & Y_2\\\cdots & \cdots & \cdots & \cdots\\X_1 & X_2 & \vdots & A_{22}\end{bmatrix} \equiv A'.$$

in which the upper left zero block of A' is $g_0(A_{11})$-by-$g_0(A_{11})$. $\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = SA_{12}$, and $\begin{bmatrix} X_1 & X_2 \end{bmatrix} = A_{21}T$. A second equivalence will zero out X_2 and Y_2:

$$\begin{bmatrix} 0 & 0 & \vdots & 0 \\ 0 & I & \vdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & -X_2 & \vdots & I \end{bmatrix} \begin{bmatrix} 0 & 0 & \vdots & Y_1 \\ 0 & I & \vdots & Y_2 \\ \cdots & \cdots & \cdots & \cdots \\ X_1 & X_2 & \vdots & A_{22} \end{bmatrix} \begin{bmatrix} 0 & 0 & \vdots & 0 \\ 0 & I & \vdots & -Y_2 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \vdots & I \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & \vdots & Y_1 \\ 0 & I & \vdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ X_1 & 0 & \vdots & \check{A}_{22} \end{bmatrix} \equiv \check{A}.$$

Neither of these equivalences has disturbed the rank of A, $\begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix}$, or $\begin{bmatrix} A_{11} & A_{12} \end{bmatrix}$.

So,

$$rank \begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix} = rank \begin{bmatrix} 0 & 0 \\ 0 & I \\ \cdots & \cdots \\ X_1 & 0 \end{bmatrix} = rank(A_{11}) + rank(X_1)$$

and

$$rank\begin{bmatrix} A_{11} & A_{12} \end{bmatrix} = rank\begin{bmatrix} 0 & 0 & \vdots & Y_1 \\ 0 & I & \vdots & 0 \end{bmatrix} = rank\,(A_{11}) + rank\,(Y_1)\,.$$

Then,

$$rank\,(X_1) = rank\begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix} - rank\,(A_{11})$$

and

$$rank\,(Y_1) = rank\begin{bmatrix} A_{11} & A_{12} \end{bmatrix} - rank\,(A_{11})\,.$$

Since $rank\begin{bmatrix} 0 & Y_1 \\ X_1 & \tilde{A}_{22} \end{bmatrix} \geq rank\,(Y_1) + rank\,(X_1)$ we have

$$\begin{aligned} rank\,(A) &= rank\left(\tilde{A}\right) = rank\,(A_{11}) + rank\begin{bmatrix} 0 & Y_1 \\ X_1 & \tilde{A}_{22} \end{bmatrix} \\ &\geq rank\,(A_{11}) + rank\,(Y_1) + rank\,(X_1) \\ &= rank\begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix} + rank\begin{bmatrix} A_{11} & A_{12} \end{bmatrix} - rank\,(A_{11})\,. \end{aligned} \tag{2.6}$$

This was previously shown in a more general setting in [Wo] in another way, but we have included a proof here for completeness.

Using the definition of i_r and i_c in (2.4) we have

$$n - g_0\,(A) = rank\,(A) \geq i_r + i_c + rank\,(A_{11}) = i_r + i_c + n - k - g_0\,(A_{11})\,.$$

Rearranging this inequality we have $k - g_0(A) + g_0(A_{11}) \geq i_r + i_c$ and substituting this in (2.5) yields

$$\dim\left(RN\begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix}\right) + \dim\left(LN\begin{bmatrix} A_{11} & A_{12} \end{bmatrix}\right) \geq$$

$$2g_0(A_{11}) + (g_0(A) - g_0(A_{11}) - k)$$

$$= g_0(A) + g_0(A_{11}) - k,$$

and part (i) of lemma 2.2 is proved.

There are two cases to consider in proving part (ii) of lemma 2.2. To begin, consider the case in which $g + g_\alpha - k \leq 0$. Note that for this to be the case, g_α must be less than or equal to k. For the matrix B in (2.3), if $g_\alpha = g_0$, then $g_0(B_{11}) = g_\alpha$ and the submatrix $\hat{B}_{22}$ is $(k - g_\alpha)$-by-$(k - g_\alpha)$. This submatrix can be chosen so that B has rank deficiency, g, from 0 to $k - g_\alpha$. Thus, B has the appropriate parameters, and, as mentioned in the proof of lemma 2.1, B has $\dim(LN_\alpha(B)) = \dim(RN_\alpha(B)) = 0$.

For the case in which $g + g_\alpha - k > 0$, consider

$$B = \begin{bmatrix} 0 & 0 & \vdots & Y \\ 0 & I_{n-k-g_\alpha} & \vdots & 0 \\ \cdots & \cdots & \vdots & \cdots \\ X & 0 & \vdots & 0 \end{bmatrix}.$$

The submatrices Y and X can independently be chosen to have rank from zero to $\min\{g_\alpha, k\}$, inclusive, which gives B a rank deficiency, g, from $|g_\alpha - k|$ to $g_\alpha + k$.

inclusive. Now, note that there is equality in (2.6) and, therefore,

$$\dim(LN_\alpha(B)) + \dim(RN_\alpha(B)) = g_0(B) + g_0(B_{11}) - k,$$

which completes the proof of the lemma. □

Our main result, the proof of which follows from lemma 2.2 by translation, is then:

Theorem 2.2 *Let $A \in M_n(F)$ and let $\alpha \subseteq N$ be such that $|\alpha| = n - k$.*

(i) $\dim(LE_\alpha^\lambda(A)) + \dim(RE_\alpha^\lambda(A)) \geq g_\lambda(A) + g_\lambda(A[\alpha]) - k$.

(ii) *Let g and g_α such that $0 \leq g \leq n$, $0 \leq g_\alpha \leq |\alpha|$, and $|g - g_\alpha| \leq k$ be given. Then, if $g + g_\alpha - k > 0$ there is a matrix B such that $g_\lambda(B) = g$, $g_\lambda(B[\alpha]) = g_\alpha$ and*

$$\dim(LE_\alpha^\lambda(B)) + \dim(RE_\alpha^\lambda(B)) = g_\lambda(B) + g_\lambda(B[\alpha]) - k.$$

If $g + g_\alpha - k \leq 0$, then there is a matrix B, with the given parameters, such that

$$\dim(LE_\alpha^\lambda(B)) = \dim(RE_\alpha^\lambda(B)) = 0.$$

In each of lemmas 2.1 and 2.2 and theorems 2.1 and 2.2, statement (ii) indicates that statement (i) is in some sense the best possible. The restrictions regarding α only avoid logical impossibilities and, otherwise, the situations not covered by statement (i) are covered in statement (ii).

At this point we make two general observations that are direct consequences of theorem 2.2.

(1) If $A \in M_n(F)$ and $|\alpha| = n - 1$, then $\lambda \in \sigma(A) \cap \sigma(A[\alpha])$ if and only if there is either a left or a right eigenvector of A (associated with λ) whose α^c component is zero.

(2) If $A \in M_n(F)$, $\lambda \in \sigma(A)$ and $\alpha \subseteq N$ with $|\alpha| = n-k$ are such that $\dim\left(LE_\alpha^\lambda(A)\right) = \dim\left(RE_\alpha^\lambda(A)\right)$, then each of

$$\dim\left(LE_\alpha^\lambda(A)\right), \dim\left(RE_\alpha^\lambda(A)\right) \geq \frac{g_\lambda(A) + g_\lambda(A[\alpha]) - k}{2}.$$

In this event, if $g_\lambda(A) + g_\lambda(A[\alpha]) > k$, then both $\dim\left(LE_\alpha^\lambda(A)\right)$ and $\dim\left(RE_\alpha^\lambda(A)\right)$ are positive.

Note that statement (1) does not follow from theorem 2.1 and is in some sense tight as, under the given hypothesis, a general matrix may not have both a left special eigenvector and a right special eigenvector. For example,

$$A = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$$

does not have the property assumed in (2) for $0 \in \sigma(A)$, and $g_0(A) = 1$, $g_0(A[\{1,2\}]) = 1$. Thus, as every right null vector of A is a multiple of $\begin{pmatrix} 1 & 1 & 1 \end{pmatrix}^T$, A has no special right eigenvector associated with 0, while it, of course, has a left such eigenvector, e.g. $\begin{pmatrix} 1 & 1 & 0 \end{pmatrix}$, because of (1). Similarly, for many values of $g_\lambda(A)$ and $g_\lambda(A[\alpha])$, the conclusion of (2) does not follow from theorem 2.1.

2.2 Structured Eigenvector Applications

We may now give several specific corollaries to theorem 2.2. First, note that if $A \in M_n$ is normal, then, as $UAU^* = D$, for some U unitary and D diagonal, any left eigenspace of A is the conjugate transpose of a right eigenspace. Thus, the hypothesis of (2) above is satisfied for each λ and α. From this observation we can conclude the following.

Corollary 2.1 *Let $A \in M_n$ be a normal matrix. For $\alpha \subseteq N$ with $|\alpha| = n - k$*

$$\dim(LE_\alpha^\lambda(A)), \dim(RE_\alpha^\lambda(A)) \geq \frac{g_\lambda(A) + g_\lambda(A[\alpha]) - k}{2}.$$

Of course Hermitian matrices are normal so the following is a special case of corollary 2.1.

Corollary 2.2 *Let $A \in M_n$ be Hermitian. For $\alpha \subseteq N$ with $|\alpha| = n - k$*

$$\dim(LE_\alpha^\lambda(A)), \dim(RE_\alpha^\lambda(A)) \geq \frac{g_\lambda(A) + g_\lambda(A[\alpha]) - k}{2}.$$

In the opening paragraph we mentioned that if A is Hermitian, $\lambda \in \sigma(A) \cap \sigma(A[\alpha])$, and $|\alpha| = n - 1$, then there is an eigenvector x (of A associated with λ) in which $x(\alpha) = 0$. But then $g_\lambda(A), g_\lambda(A[\alpha]) \geq 1$ which results in a positive right-hand side in corollary 2.2. In this case, both the left and the right special eigenspaces are nonempty, which proves the following corollary.

Corollary 2.3 *Let $A \in M_n$ be Hermitian, let $\alpha \subseteq N$ be such that $|\alpha| = n - 1$, and let $\lambda \in \mathbb{R}$ be an eigenvalue of A. Then, there is an eigenvector x of A associated with λ such that $x(\alpha) = 0$ if and only if $\lambda \in \sigma(A[\alpha])$.*

Thus, the general scheme adopted here provides an algebraic proof to the statement in the opening paragraph.

2.2.1 Application to Singular Values

As noted in the introduction, there is also an interlacing property for singular values that follows from classical interlacing of eigenvalues of Hermitian matrices. Therefore, it is natural to ask if there are results for singular vectors that are similar to the above theorems. Let A be an m-by-n matrix and let $\sigma_1 > \sigma_2 > \cdots > \sigma_k$, $k \leq min\{m,n\}$ be the set of distinct nonzero singular values of A. Let $m_j(A)$, $j = 1,2,\ldots,k$ denote the multiplicity of σ_j as a singular value of A. Note that $m_j(A) = g_{\sigma_j^2}(A^*A) = g_{\sigma_j^2}(AA^*)$. Then let $A = V\Sigma W^*$ be the singular value decomposition of A. If σ_j has multiplicity $m_j(A)$ then there are right singular vectors, $w_{j,1}, w_{j,2}, \ldots, w_{j,m_j}$, and left singular vectors, $v_{j,1}, v_{j,2}, \ldots, v_{j,m_j}$ such that

$$A\begin{bmatrix} w_{j,1} & w_{j,2} & \cdots & w_{j,m_j} \end{bmatrix} = \sigma_j \begin{bmatrix} v_{j,1} & v_{j,2} & \cdots & v_{j,m_j} \end{bmatrix}.$$

The right singular vector space of A associated with σ_j is

$$span\left\{w_{j,1}, w_{j,2}, \ldots, w_{j,m_j}\right\} \equiv RS^{\sigma_j}(A)$$

and the left singular vector space of A associated with σ_j is

$$span\left\{v_{j,1}, v_{j,2}, \ldots, v_{j,m_j}\right\} \equiv LS^{\sigma_j}(A).$$

Let $\alpha \subseteq N$ and define the special singular vector subspaces

$$\begin{aligned} LS_\alpha^{\sigma_j}(A) &= \{v \in LS^{\sigma_j}(A) : v(\alpha) = 0\} \\ RS_\alpha^{\sigma_j}(A) &= \{w \in RS^{\sigma_j}(A) : w(\alpha) = 0\} \end{aligned}$$

Also, let $M = \{1, 2, \ldots, m\}$ and, as usual, $N = \{1, 2, \ldots, n\}$. Then

Theorem 2.3 *Let $A \in M_{m,n}(\mathbb{C})$, then for $\alpha \subseteq M, \beta \subseteq N$ with $|\alpha| = m - k_r$, $|\beta| = n - k_c$,*

$$\dim(RS_{\alpha}^{\sigma_k}(A)) \oplus \dim(LS_{\alpha}^{\sigma_k}(A)) \geq \frac{m_k(A) + m_k(A[\alpha;\beta]) - k_r - k_c}{2}.$$

Proof: Assume $m \geq n$ (if not, reverse the rolls of m and n in the discussion that follows) and let $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_k$ be the distinct nonzero singular values of A. Let $B = \begin{bmatrix} 0 & A \\ A^* & 0 \end{bmatrix}$. Then B is Hermitian with unique eigenvalues $-\sigma_1, -\sigma_2, \ldots, -\sigma_{k-1}, -\sigma_k, \sigma_k, \sigma_{k-1}, \ldots, \sigma_1$. Note that when multiplicities are considered, B has $(m+n) - 2rank(A)$ eigenvalues that are 0, and $m - n$ of these zero eigenvalues do not correspond to singular values of A. Delete $k_r + k_c$ rows and columns of B in such a way as to delete the k_r rows $M - \alpha$ and the k_c columns $N - \beta$ from A. Also delete the corresponding k_r columns and k_c rows from A^*. The resulting matrix is:

$$\hat{B} = B[\hat{\alpha}] = \begin{bmatrix} 0 & A[\alpha;\beta] \\ A[\alpha;\beta]^* & 0 \end{bmatrix}$$

in which $\hat{\alpha} = \alpha \cup (\beta + \{n\})$ and $\hat{B}$ is $(m + n - k_r - k_c)$-by-$(m + n - k_r - k_c)$. Corollary 2.2 gives

$$\dim\left(LE_{\hat{\alpha}}^{\sigma_k}(B)\right), \dim\left(RE_{\hat{\alpha}}^{\sigma_k}(B)\right) \geq \frac{g_{\sigma_k}(B) + g_{\sigma_k}\left(\hat{B}\right) - k_r - k_c}{2}. \qquad (2.7)$$

The eigenvectors of B are of the form $\begin{bmatrix} w \\ v \end{bmatrix}$ in which w is a right singular vector of A associated with σ_k and v is a left singular vector of A associated with σ_k. This means that the dimension of the eigenspace associated with σ_k in which $w(\alpha) = 0$ and $v(\beta) = 0$ is at least

$$\frac{g_{\sigma_k}(B) + g_{\sigma_k}\left(\hat{B}\right) - k_r - k_c}{2} = \frac{m_k(A) + m_k(A[\alpha;\beta]) - k_r - k_c}{2}$$

since $g_{\sigma_k}(B) = m_k(A)$ and $g_{\sigma_k}\left(\hat{B}\right) = m_k(A[\alpha;\beta])$. Thus,

$$\dim(RS_{\alpha}^{\sigma_k}(A)) \oplus \dim(LS_{\alpha}^{\sigma_k}(A)) \geq \frac{m_k(A) + m_k(A[\alpha;\beta]) - k_r - k_c}{2}.$$

□

Note that if $m_0(A)$ denotes the multiplicity of 0 as a singular value of A, then $g_0(B) = 2m_0(A) + m - n$ and $g_0\left(\hat{B}\right) = 2m_0(A[\alpha;\beta]) + |(m - k_r) - (n - k_c)|$, which when substituted in (2.7) yields

$$\dim(LN_{\hat{\alpha}}(B)), \dim(RN_{\hat{\alpha}}(B)) \geq \left\{ \begin{array}{ll} m_0(A) + m_0(A[\alpha;\beta]) + m - n - k_r & \text{if } m - k_r > n - k_c \\ m_0(A) + m_0(A[\alpha;\beta]) - k_c & \text{if } m - k_r \leq n - k_c \end{array} \right\}.$$

2.3 A Perturbation Result

Corollary 2.3 characterizes the case in which A is Hermitian, $|\alpha| = n - 1$, and there is equality in the interlacing inequalities. We next consider approximate versions of the two equivalent statements in corollary 2.3, i.e. the case in which there is an eigenvalue of $A[\alpha]$ that is "near by" an eigenvalue of A, or there is an eigenvector, x, of A associated with λ that has a component that is "small" relative to the other components of x. Let A_j denote the j^{th} column of the matrix A and note that $\|\bullet\|_2$ is the usual Euclidean norm.

Theorem 2.4 *Let $A \in M_n$ be Hermitian, let $\alpha \subseteq N$ be such that $|\alpha| = n-1$ with $\alpha^c = j$. Then for $\lambda \in \sigma(A)$ and x an associated eigenvector such that $x(j) \neq 0$*

(i) There is a $\mu \in \sigma(A[\alpha])$ such that

$$|\lambda - \mu| \leq \frac{|x_j| \sqrt{\|A_j\|_2^2 - |a_{jj}|^2}}{\sqrt{\|x\|_2 - |x_j|^2}}$$

(ii) If $\lambda \notin \sigma(A[\alpha])$, then

$$\frac{|x_j|}{\|x\|_2} = \frac{1}{\sqrt{\|(A[\alpha] - \lambda I)^{-1} A[\alpha;\alpha^c]\|_2^2 + 1}}$$

Note that $\|x\|_2 - |x_j|^2 = 0$ if and only if $x(j) = 0$.

Proof: Let $x \in \mathbb{C}^n$ be such that $Ax = \lambda x$. To prove (i) let $A_j(j)$ be the vector formed by deleting the j^{th} entry of A_j and note that $A(\emptyset;\{j\})$ $(A(\{j\};\emptyset))$ is the n-by-$(n-1)$ $((n-1)$-by-$n)$ matrix in which the j^{th} column (row) of A has been deleted. By matrix vector multiplication, deleting the j^{th} column of A results in

$$A(\emptyset,\{j\})\,x(j) = Ax - x_j A_j. \tag{2.8}$$

Further deletion of the j^{th} row yields

$$A(j)\,x(j) = A(\{j\};\emptyset)\,x - x_j A_j(j) = \lambda x(j) - x_j A_j(j).$$

Then

$$\|A(j)\,x(j) - \lambda x(j)\|_2 = |x_j|\,\|A_j(j)\|_2 .$$

By [P, theorem 4-5-1], for any scalar λ and any nonzero vector x, there is an eigenvalue γ of an n-by-n matrix B satisfying

$$|\gamma - \lambda| \leq \frac{\|Bx - \lambda x\|_2}{\|x\|_2}.$$

Then, for the scalar λ there is a $\mu \in \sigma(A(j))$ for which

$$\begin{aligned} |\mu - \lambda| &\leq \frac{\|A(j)\,x(j) - \lambda x(j)\|_2}{\|x(j)\|_2} \\ &= \frac{|x_j|\,\|A_j(j)\|_2}{\sqrt{\|x\|_2^2 - |x_j|^2}} = \frac{|x_j|\sqrt{\|A_j\|_2^2 - |a_{jj}|^2}}{\sqrt{\|x\|_2^2 - |x_j|^2}} \end{aligned}$$

which proves (i).

To prove (ii), note that since $Ax = \lambda x$, we have

$$A[\alpha]\,x[\alpha] + A[\alpha;\alpha^c]\,x_j = \lambda x[\alpha]$$

(note that $x(j) = x[\alpha]$, $A(j) = A[\alpha]$, and $x_j = x(\alpha)$) or

$$-(A[\alpha] - \lambda I)\,x[\alpha] = A[\alpha;\alpha^c]\,x_j.$$

If $\lambda \notin \sigma(A[\alpha])$, then $A[\alpha] - \lambda I$ is invertible and

$$x[\alpha] = -(A[\alpha] - \lambda I)^{-1} A[\alpha;\alpha^c]\,x_j. \tag{2.9}$$

By taking the length of both sides in (2.9) we see that

$$\|x\|_2^2 - |x_j|^2 = |x_j|^2 \left\| (A[\alpha] - \lambda I)^{-1} A[\alpha;\alpha^c] \right\|_2^2 . \tag{2.10}$$

The vector $(A[\alpha] - \lambda I)^{-1} A[\alpha;\alpha^c]$ is zero if and only if $A[\alpha;\alpha^c]$ is zero, but then $\lambda \in \sigma(A[\alpha])$, a contradiction. So, $(A[\alpha] - \lambda I)^{-1} A[\alpha;\alpha^c] \neq 0$. Rearranging (2.10) results in

$$\frac{|x_j|^2}{\|x\|_2^2} = \frac{1}{\left\| (A[\alpha] - \lambda I)^{-1} A[\alpha;\alpha^c] \right\|_2^2 + 1}$$

which proves part (ii) of the theorem. □

We note that theorem 2.4 says that corollary 2.3 is valid in the following approximate sense: if x_j is "small" then (i) indicates there is an eigenvalue $\mu \in \sigma(A(j))$ that is "near" some eigenvalue $\lambda \in \sigma(A)$; also, if there is a $\mu \in \sigma(A(j))$ that is "near" some eigenvalue $\lambda \in \sigma(A)$, and $\|A[\alpha;\alpha^c]\| > (n-1)|\mu - \lambda|$, then there is an eigenvector of A associated with λ for which x_j is also "small".

We will illustrate this theorem by an example:

$$A = \begin{bmatrix} -5 & -29 & \vdots & 5 \\ -29 & -95 & \vdots & 0 \\ \cdots & \cdots & \vdots & \cdots \\ 5 & 0 & \vdots & -36 \end{bmatrix}$$

which has $\lambda = -103.56$ as an eigenvalue with associated eigenvector $x = \begin{pmatrix} -.28 \\ -.96 \\ .02 \end{pmatrix}$.

The calculations in (i) indicate that there is a $\mu \in \sigma(A[\{1,2\}])$ such that

$|\mu - \lambda| \leq .78$. The submatrix $A[\{1,2\}]$ has an eigenvalue $\mu = -103.54$ which satisfies this inequality. Similarly, since $\|x\|_2 = 1$, statement (ii) indicates that $|x_3| = .02$ which is the case.

Chapter 3

Interlacing

3.1 Breadth and Depth of Interlacing

Let $A \in M_n(\mathbb{C})$ be Hermitian. Since any principal submatrix of an Hermitian matrix is Hermitian, corollary 2.3 may be applied at each "level" of interlacing. After a few basic observations, careful sequential application of corollary 2.3 will lead to the results below, but first several definitions are needed. Suppose λ is an eigenvalue of A, then A is said to have *interlacing equality at λ of breadth k* if there are exactly k distinct index sets $\alpha_1, \alpha_2, \ldots, \alpha_k \subseteq N$ in which $|\alpha_i| = n - 1$ and $\lambda \in \sigma(A[\alpha_i])$, $i = 1, 2, \ldots, k$. If $g_\lambda(A) \geq 2$, then the breadth of interlacing equality at λ is n (see discussion following corollary 3.3.) As will be seen in the proof of corollary 3.1 if A is such that $g_\lambda(A) = 1$, then, because of corollary 2.3 the breadth of interlacing equality at λ is just the number of zero components in the essentially unique eigenvector. The matrix A is said to have *interlacing equality at λ of depth k* if $\lambda \in \sigma(A[\beta_i])$ for some index sets $\beta_0, \beta_1, \ldots, \beta_k \subseteq N$ such that $\beta_{j+1} \subset \beta_j, j = 0, 1, \ldots, k-1$, $|\beta_j| = n - j$, $j = 0, 1, \ldots, k$ and k is a maximum. If in addition, $g_\lambda(A[\beta_{j+1}]) \geq g_\lambda(A[\beta_j])$,

$j = 0, 1, \ldots, k-1$, then A is said to have *interlacing equality at* λ *of* ***restricted*** *depth k*. Here, k is the number of principal submatrices in the nested sequence for which the geometric multiplicity of λ is nondecreasing, so that the depth of interlacing equality may be greater than the restricted depth. The following corollaries relate these concepts.

Corollary 3.1 *Let* $A \in M_n$ *be Hermitian and be such that* $g_\lambda(A) = 1$. *If* A *has interlacing equality at* λ *of breadth* k, *then* A *has interlacing equality at* λ *of depth at least* k.

Proof: If A has interlacing equality at λ of breadth k, then there are k distinct principal submatrices $A[\alpha_i]$ such that $\lambda \in \sigma(A[\alpha_i])$ and $|\alpha_i| = n-1$. In this case, $g_\lambda(A[\alpha_i]) \geq 1$ and, by assumption, $g_\lambda(A) = 1$. Thus, by corollary 2.3 , for each α_i there is an eigenvector y_i of A associated with λ, such that $y_i(\alpha_i) = 0$. However, since $g_\lambda(A) = 1$, the (right) eigenspace of A associated with λ is one dimensional, so that each of the y_i's may be taken to be the same, x. It follows that $x(\alpha_1 \cap \alpha_2 \cap \cdots \cap \alpha_k) = 0$. By the partitioned calculation mentioned in the opening paragraph of chapter 2 $\beta_0 = N$, and $\beta_i = \alpha_1 \cap \alpha_2 \cap \cdots \cap \alpha_i$, $i = 1, 2, \ldots, k$, exhibit that A has interlacing equality at λ of depth at least k. □

Corollary 3.1 is stated in the Hermitian case for parallelism to the corollaries that follow. However, the concepts of breadth and depth of interlacing may also be thought of simply as coincidence of eigenvalues. In this context, the arguments in the proof of corollary 3.1 are equally valid for normal matrices (using corollary 2.1 in place of corollary 2.3 with an obvious generalization of the definitions of breadth and depth of interlacing). So, corollary 3.1 may be generalized by replacing "Hermitian" in the hypothesis with "normal." On the other hand, the coincidence of eigenvalues indicated by corollary 3.1 is not valid for general matrices, as exhibited by the example

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

in which 0 is an eigenvalue of breadth 2, while its depth is only 1.

The converse to corollary 3.1 does not hold; a counterexample is given by the matrix

$$\begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$$

which has interlacing equality at 0 of depth 3 ($A(\{4\})$, $A(\{3,4\})$, $A(\{2,3,4\})$), but interlacing equality at 0 of breadth only 2 ($A(\{3\}), A(\{4\})$). However, the geometric multiplicities of 0 in the principal submatrices that yield interlacing equality at 0 of depth 3 are

$$\begin{aligned} g_0(A(\{4\})) &= 1 \\ g_0(A(\{3,4\})) &= 2 \\ g_0(A(\{2,3,4\})) &= 1. \end{aligned}$$

In fact, the restricted depth of interlacing equality at 0 is only 2 and this is exactly the breadth of interlacing equality at 0. As indicated in the following corollary, the breadth of interlacing equality at λ must be at least that of the restricted depth.

Corollary 3.2 *Let $A \in M_n$ be Hermitian and suppose $\lambda \in \sigma(A)$. If A has interlacing equality at λ of restricted depth k, then A has interlacing equality at λ of breadth at least k.*

Proof: If $g_\lambda(A) > 1$, the breadth at λ is n (see discussion later, if necessary) and the conclusion is automatically valid. Thus, we suppose $g_\lambda(A) = 1$. If A has interlacing equality at λ of restricted depth k, then there is some nested sequence of $k+1$ principal submatrices $A[\beta_i]$, such that $|\beta_i| = n-i$, $\lambda \in \sigma(A[\beta_i])$, $i = 0, 1, \ldots, k$, and $g_\lambda(A[\beta_{i+1}]) \geq g_\lambda(A[\beta_i])$, $i = 0, 1, \ldots, k-1$. Assume, without loss of generality, that the rows and columns of $A[\beta_i]$ are numbered 1 to $n-i$. Note that $n-i$ is the index of the row and column deleted from $A[\beta_i]$ to obtain $A[\beta_{i+1}]$. By corollary 2.2

$$\begin{aligned}
\dim\left(LE^{\lambda}_{\beta_{i+1}}(A[\beta_i])\right), \dim\left(RE^{\lambda}_{\beta_{i+1}}(A[\beta_i])\right) &\geq \frac{g_\lambda(A[\beta_i]) + g_\lambda(A[\beta_{i+1}]) - 1}{2} \\
&\geq g_\lambda(A[\beta_i]) - \frac{1}{2}
\end{aligned}$$

since $g_\lambda(A[\beta_i]) \leq g_\lambda(A[\beta_{i+1}])$. Both dimensions must be integral; so, the dimensions of the special eigenspaces must both be at least $g_\lambda(A[\beta_i])$. Then, every (left and right) eigenvector of $A[\beta_i]$ associated with λ is in the special (left and right) eigenspace and, thus, component $n-i$ of each of these vectors is 0.

Let x be an eigenvector (essentially unique) of A associated with λ. Since

$$g_\lambda(A) = g_\lambda(A[\beta_0]) = 1 \text{ and } g_\lambda(A[\beta_1]) \geq 1,$$

by corollary 2.3 $x(\beta_1) = 0$. By the preceding paragraph, if $i = 1$ then every eigenvector of $A[\beta_1]$ associated with λ, including $x[\beta_1]$, has a zero in the $n-1$ component. Thus, $x(\beta_1 \cap \beta_2) = x(\beta_2) = 0$.

Continuing in this manner, for each $i = 0, 1, \ldots, k-1$, $x[\beta_i]$ is an eigenvector of $A[\beta_i]$ associated with λ with a zero in the $n-i$ component so that

$$x(\beta_1 \cap \beta_2 \cap \ldots \cap \beta_{i+1}) = x(\beta_{i+1}) = 0.$$

Then, $x(\beta_k) = 0$ and for each $j \notin \beta_k$, $x(\{j\})$ is an eigenvector of $A(\{j\})$ associated with λ. Thus, $\alpha_j = N - \{n+1-j\}$, $j = 1, 2, \ldots, k$, exhibits that A has interlacing equality at λ of breadth at least k. □

Note that the breadth of interlacing equality can be strictly greater than the restricted depth of interlacing equality. For example, the matrix

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

has interlacing equality at 0 of restricted depth 1, but the breadth of interlacing equality at 0 is 2.

If the matrix A is such that $g_\lambda(A[\alpha]) \leq 1$ for every index set $\alpha \subseteq N$, and A has interlacing equality at λ of depth k, then A also has interlacing equality at λ of restricted depth k. In this case, by corollary 3.2, A has interlacing equality at λ of breadth at least k. Combining corollaries 3.1 and 3.2 then yields the following.

Corollary 3.3 *Let $A \in M_n$ be Hermitian and suppose for every index set $\alpha \subseteq N$ that $g_\lambda(A[\alpha]) \leq 1$ with $g_\lambda(A) = 1$. Then, A has interlacing equality at λ of breadth k if and only if A has interlacing equality at λ of depth k.*

Let $A \in M_n$ be Hermitian. Due to classical interlacing, when $g_\lambda(A) > 1$, $\lambda \in \sigma(A[\alpha])$ for any $\alpha \subseteq N$ such that $|\alpha| = n-1$. In addition, when $g_\lambda(A) > 1$ for each such α there is an eigenvector, z, of A associated with λ such that $z(\alpha) = 0$. This may be seen in an elementary way by noting that, given any two linearly independent eigenvectors x, y in the eigenspace, there is a linear combination with a zero in any specified position. Such an A has interlacing equality at λ of breadth n, but may have depth at λ as little as 1. For example, the matrix

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

has interlacing equality at 0 of breadth 3, while the depth at 0 is only 1. Thus, the assumption in corollaries 3.1 and 3.3 that $g_\lambda(A) = 1$ is necessary.

3.2 Interlacing Diagrams

Corollaries 3.1-3.3 indicate restrictions on the values of $g_\lambda(A[\alpha])$ for various α's. This led us to look at the characterization of the relationships between different "levels" of interlacing. These relationship are described via what we call interlacing diagrams. For example, if

$$A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

then A has interlacing diagram

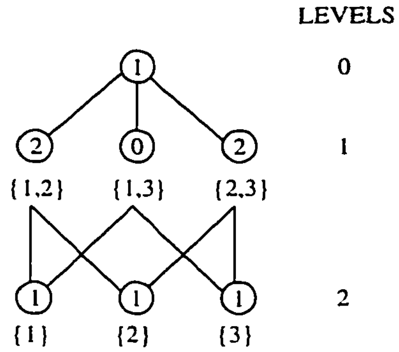


Figure 3.1: Interlacing diagram example.

in which the entry in the cell in level 0 is the geometric multiplicity of 0 as an eigenvalue of A and the entry in each of the other cells is the geometric multiplicity of 0 as an eigenvalue of the principal submatrix of A lying in the rows and columns indicated by the index set below each cell. The general problem is, for an n-by-n matrix A and fixed $\lambda \in \sigma(A)$, which *n-level diagrams* of the form

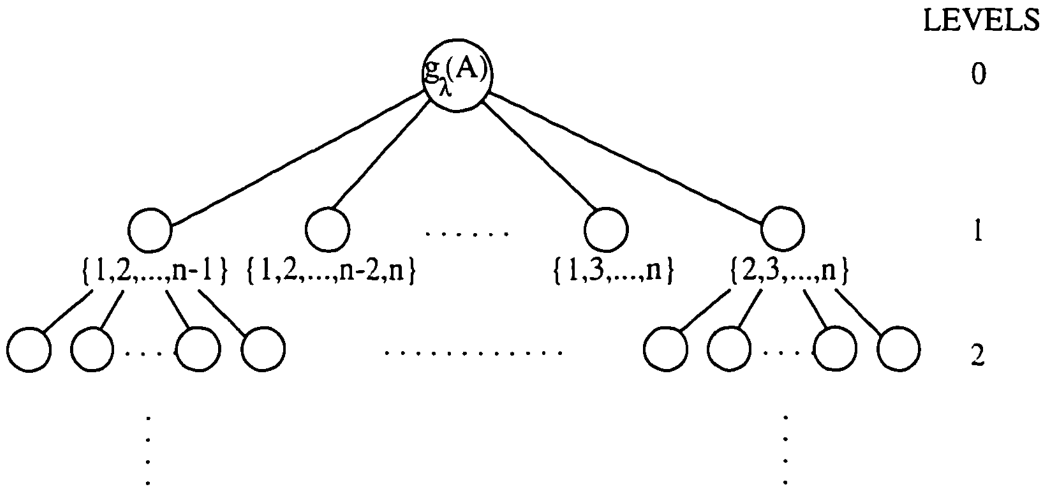


Figure 3.2: General interlacing diagram

in which the geometric multiplicity of λ as an eigenvalue of the principal submatrix found in the index set associated with that cell appears in each cell, can occur?

An obvious condition for such diagrams is

(1) **Dimension:** $g_\lambda(A[\alpha]) \leq |\alpha|$.

So, the entries in the cells of level k can be at most $n - k$. We will be looking primarily at Hermitian matrices. For Hermitian matrices, classical interlacing restrictions require that the geometric multiplicity in a predecessor or successor cell may differ by at most one from that in a given cell. This gives rise to a second condition.

(2) **Interlacing:** Let $A \in M_n$ be Hermitian and $\alpha \subseteq N$. Then for $j \in \alpha$, $|g_\lambda(A[\alpha]) - g_\lambda(A[\alpha - \{j\}])| \leq 1$.

For the remainder of this discussion we will assume that $\lambda = 0$. The general case in which $\lambda \neq 0$ follows easily by translation. For Hermitian matrices, the two obvious necessary conditions given above are not sufficient. That is, there are diagrams that satisfy (1) and (2) for which there is no Hermitian matrix with that diagram. For example, the diagram

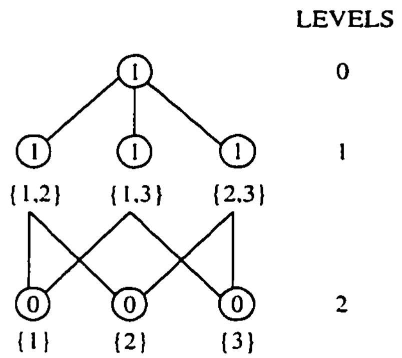


Figure 3.3: Diagram ruled out by biorthogonality

satisfies conditions (1) and (2) but there is no 3-by-3 Hermitian matrix with the interlacing diagram given in figure 3.3. For $\lambda = 0$ level 2 indicates that the diagonal entries of a matrix with the diagram in figure 3.3 are nonzero. Level 1 and level 2 of the diagram, together with the fact that we are considering only Hermitian matrices, indicates that any matrix with the diagram in figure 3.3 must be of the form

$$A = \begin{bmatrix} a^2 & ab & ac \\ ab & b^2 & bc \\ ac & bc & c^2 \end{bmatrix}$$

in which $a, b, c \neq 0$. But $g_0(A) = 2$ and the diagram requires this value to be 1. Therefore, there is no Hermitian matrix with the diagram in figure 3.3.

The structured eigenvector results of chapter 2 and the principal of biorthogonality lead to the following corollary which holds for general diagonalizable matrices.

Corollary 3.4 *If A is a diagonalizable matrix and $g_\lambda(A) = 1$, then $g_\lambda(A(i)) = 0$ for some $1 \leq i \leq n$.*

Proof: Let A be diagonalizable, $g_\lambda(A) = 1$, and assume that $g_\lambda(A(i)) \geq 1$ for all $i = 1, 2, \ldots, n$. Also, let x, y be right and left eigenvectors, respectively, of A associated with λ. Then, by theorem 2.2 at least one of x and y must have a 0 in the i^{th} component for $i = 1, 2, \ldots, n$ so that $y^* x = 0$. By the principal of biorthogonality y must be orthogonal to every right eigenvector of A associated with $\mu \in \sigma(A)$, $\mu \neq \lambda$. Since A is diagonalizable, this means that y is orthogonal to $n - 1$ vectors. But y is also orthogonal to x, and is, therefore, orthogonal to a set of n linearly independent vectors. Then, y must be the zero vector which is a contradiction since y is an eigenvector. Therefore, there must be some i for which $g_\lambda(A(i)) = 0$. □

The diagram in figure 3.3 does not satisfy corollary 3.4 and, therefore, there is no 3-by-3 Hermitian matrix with this diagram. Corollary 3.4 gives a third condition on the interlacing diagrams.

(3) **Biorthogonality:** For diagonalizable matrices, every diagram with a 1 in the cell at level 0 must have at least one 0 in level 1 (corollary 3.4.)

Note that any pattern that is not allowed in an interlacing diagram of size k is not allowed as a subpattern of a diagram of size $n, n > k$. So, condition (3) indicates that if a cell in level k of a diagram contains a 1, then there must be at least one 0 among that cell's successor cells.

Conditions (1)-(3) are still not sufficient. The diagram

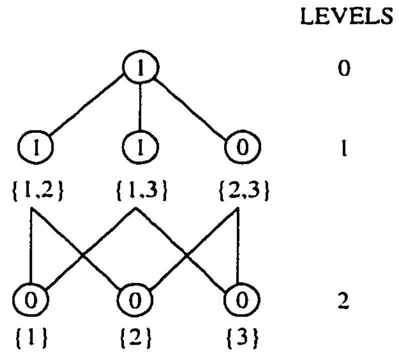


Figure 3.4: Diagram ruled out by breadth and depth corollaries

satisfies conditions (1)-(3), but, due to levels 1 and 2, any Hermitian matrix with this diagram is of the form

$$A = \begin{bmatrix} a^2 & ab & ac \\ ab & b^2 & 0 \\ ac & 0 & c^2 \end{bmatrix}$$

in which $a, b, c \neq 0$. But this matrix is nonsingular and the cell in level 0 of figure 3.4 requires that any matrix with this diagram is singular.

The diagram in figure 3.4 cannot occur because it violates the breadth and depth corollaries of section 3.1. By the proof of corollary 3.1 the 1 in level 0 and the two 1's in level 1 indicate that the first cell in level 2 (corresponding to the 1st diagonal entry) must be a 1 and cannot be 0. The breadth and depth of interlacing corollaries (corollaries 3.1-3.3) give a fourth condition on the interlacing diagrams for Hermitian matrices.

(4) **Breadth and Depth of Interlacing:** corollaries 3.1-3.3.

The following corollary indicates a pattern that cannot occur as a subpattern of any larger diagram.

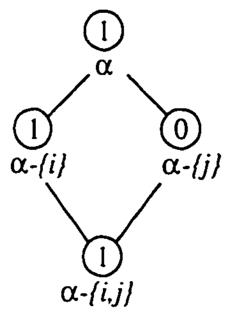


Figure 3.5: Forbidden interlacing sub-diagram

Corollary 3.5 *For n-by-n Hermitian matrices in which $n \geq 3$, the pattern in figure 3.5 cannot occur as a subpattern in any interlacing diagram.*

Proof: Assume the pattern in figure 3.5 can be achieved for the eigenvalue λ of a matrix, A, and that level k corresponds to $A[\alpha], |\alpha| = n - k$. It can be assumed by permutation similarity that $\alpha = \{1, 2, \ldots, n - k\}$. Then, in the $(k + 1)^{st}$ level, the cell containing a 1 corresponds to $A[\alpha - \{i\}]$ and the cell containing a 0 corresponds to $A[\alpha - \{j\}]$, for $i, j \in \alpha, i \neq j$. Let x be an, essentially unique, eigenvector of $A[\alpha]$ associated with λ. Then, by theorem 2.2 $x[\{i\}] = 0$ and, because there is a 1 in the cell in level $k + 2$, $x[\{i, j\}]$ is also 0. But then, by partitioned matrix multiplication, $x(\{j\})$ must be an eigenvector of $A[\alpha - \{j\}]$ associated with λ, contradicting the fact that the diagram indicates that $g_\lambda(A[\alpha - \{j\}]) = 0$. □

Thus, corollary 3.5 adds one more condition to the list that governs the allowable interlacing diagrams of Hermitian matrices:

(5) **Forbidden sub-diagram:** corollary 3.5

For 3-level diagrams in which A is a *real* symmetric matrix and $g_\lambda(A) = 1$, these conditions are enough to characterize the diagrams that can occur (proof is by exhaustion; there are 216 cases) with one exception. In other words, for each 3-level diagram that satisfies conditions (1)-(5) there is an Hermitian matrix for which that diagram represents the geometric multiplicities of λ as an eigenvalue of each of its principal submatrices, with the one exception indicated below. The diagram

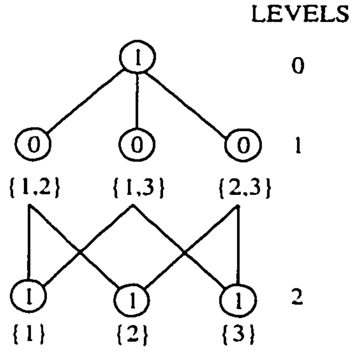


Figure 3.6: 3-level interlacing diagram that is not ruled out

is not covered by the conditions given above and cannot occur for $A \in M_3(\mathbb{R})$, but can occur for $A \in M_3(\mathbb{C})$. For $A \in M_3(\mathbb{R})$ the diagram in figure 3.6 corresponds to a matrix of the form

$$A = \begin{bmatrix} 0 & a & b \\ a & 0 & c \\ b & c & 0 \end{bmatrix}.$$

The determinant of this matrix is $2abc$ which is 0 if and only if a, b, or c is 0. If this is the case, then level 1 of the diagram is violated. However, if $A \in M_3(\mathbb{C})$, then

$$A = \begin{bmatrix} 0 & \alpha & \beta \\ \bar{\alpha} & 0 & \gamma \\ \bar{\beta} & \bar{\gamma} & 0 \end{bmatrix}$$

and $\det(A) = \alpha\bar{\beta}\gamma + \bar{\alpha}\beta\bar{\gamma}$ which is zero if $\alpha\bar{\beta}\gamma = -\bar{\alpha}\beta\bar{\gamma}$.

3.3 Extended Interlacing Intervals

In the introduction to this dissertation it was mentioned that the classical interlacing inequalities may be viewed as saying that if $\hat{\lambda}_1 \leq \hat{\lambda}_2 \leq \cdots \leq \hat{\lambda}_{n-1}$ are the eigenvalues of a principal submatrix $\hat{A} \in M_{n-1}$ of the Hermitian matrix A, then each interval $\left[\hat{\lambda}_i, \hat{\lambda}_{i+1}\right]$, $i = 1, 2, \ldots, n-2$ contains an eigenvalue of the full matrix and if $\left[\hat{\lambda}_i, \hat{\lambda}_{i+1}\right]$ captures λ_s, then $\left[\hat{\lambda}_j, \hat{\lambda}_{j+1}\right]$, $i \neq j$, captures λ_t in which $s \neq t$. In fact, if $\hat{\lambda}_1 \leq \hat{\lambda}_2 \leq \cdots \leq \hat{\lambda}_{n-p}$ are the eigenvalues of a principal $(n-p)$-by-$(n-p)$ submatrix of A, then the interval $\left[\hat{\lambda}_i, \hat{\lambda}_j\right]$, $0 < i \leq j \leq n-p$ contains at least $j-i-p+1$ eigenvalues of A. However, this count may be non-positive, meaning that there may be no eigenvalues of A in the interval. The main result of this section uses the singular values of a non-principal submatrix of A to extend such intervals so as to capture more eigenvalues of A in the interval. Because of the relationship between positivity of singular values and rank, our result provides convenient insights in the event that the matrix structure limits the rank of certain off-diagonal blocks.

3.3.1 Main Result

An immediate consequence of the Courant-Fischer theorem for singular values (see, e.g. theorem 7.3.10 [HJ1]) is the following observation about general matrices.

Lemma 3.1 *The matrix $A \in M_{m,n}$ has k singular values less than or equal to $\delta > 0$ if and only if there exists a k-dimensional subspace, $S \subseteq \mathbb{C}^n$ such that $\|Ax\|_2 \leq \delta$ for all $x \in S$, $\|x\|_2 = 1$.*

If A is Hermitian, then the singular values of A are the absolute values of the eigenvalues of A, so that lemma 3.1 is also a statement about eigenvalues. Translation then yields the following fact for Hermitian matrices.

Lemma 3.2 *Let $A \in M_n$ be Hermitian, and let $t \in \mathbb{R}$. Given $\delta > 0$, there exists a k-dimensional subspace, $S \subseteq \mathbb{C}^n$, such that $\|(A - tI)x\|_2 \leq \delta$ for all $x \in S$, $\|x\|_2 = 1$ if and only if A has k eigenvalues in the interval $[t - \delta, t + \delta]$.*

Proof: By lemma 3.1, there are k singular values of $A - tI$ less than or equal to δ if and only if there is a k-dimensional space, S, for which $\|(A - tI)\, x\|_2 \leq \delta$ for all $x \in S$, $\|x\|_2 = 1$. Since $A - tI$ is Hermitian (A is Hermitian and t is real) its singular values are the values $|\lambda_i - t|$, in which $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ are the eigenvalues of A. But then $|\lambda_i - t| \leq \delta$, for $i = j, j+1, \ldots, j+k-1$, and some j, $1 \leq j \leq n - k + 1$. This means $t - \delta \leq \lambda_i \leq t + \delta$ for these values of i and that there are k eigenvalues of A in the interval $[t - \delta, t + \delta]$. □

If A is Hermitian, the eigenvalues of A are denoted by $\lambda_1 (A) \leq \lambda_2 (A) \leq \cdots \lambda_n (A)$ and those of $A[\alpha]$ by $\lambda_1^\alpha \leq \lambda_2^\alpha \leq \cdots \lambda_{|\alpha|}^\alpha$. The theorem below uses the singular values $\sigma_1^\alpha \geq \sigma_2^\alpha \geq \cdots \sigma_{|\alpha|}^\alpha$ of $A[\alpha^c; \alpha]$ to extend the intervals of the type discussed in the opening paragraph of section 3.3. In the event that σ_k^α is 0, either because $|\alpha|$ is sufficiently larger than $|\alpha^c|$ or because $A[\alpha^c; \alpha]$ has sufficiently low rank, our extended interval reverts to a conventional interval between two, not necessarily consecutive, eigenvalues. Note that, since A is Hermitian, the nonzero singular values of $A[\alpha^c; \alpha]$ are the same as those of $A[\alpha; \alpha^c]$.

Permutation similarity does not change the eigenvalues of A; therefore, for the theorems that follow, it will be assumed, without loss of generality, that $\alpha = \{1, 2, \ldots, |\alpha|\}$. Also, let

$$t \equiv \frac{\lambda_j^\alpha + \lambda_i^\alpha}{2}, \delta \equiv \frac{\lambda_j^\alpha - \lambda_i^\alpha}{2}, \text{ for fixed } i \leq j \leq |\alpha|$$

and define the interval

$$I(i, j, k, \alpha) \equiv \left[t - \sqrt{\delta^2 + (\sigma_k^\alpha)^2}, t + \sqrt{\delta^2 + (\sigma_k^\alpha)^2}\right].$$

Note that $[t - \delta, t + \delta]$ is an interval used in classical interlacing, discussed above, and, thus, this interval must capture at least $j - i - |\alpha^c| + 1$ eigenvalues of A. The main result is:

Theorem 3.1 *Let $A \in M_n$ be Hermitian and let $\alpha \subseteq N$; then, for each k, A has at least $j - i - k + 2$ eigenvalues in the interval $I(i, j, k, \alpha)$.*

Proof: Let $x_i^\alpha, x_{i+1}^\alpha, \ldots x_j^\alpha$ be an orthonormal set of eigenvectors of $A[\alpha]$ corresponding to $\lambda_i^\alpha \leq \lambda_{i+1}^\alpha \leq \cdots \leq \lambda_j^\alpha$, respectively. Also, let $w_k^\alpha, w_{k+1}^\alpha, \ldots w_{|\alpha|}^\alpha$ be a set of (orthonormal) right singular vectors of $A[\alpha^c; \alpha]$ corresponding to $\sigma_k^\alpha \geq \sigma_{k+1}^\alpha \geq \cdots \geq \sigma_{|\alpha|}^\alpha$. Define the subspace $S \subseteq \mathbb{C}^{|\alpha|}$ as

$$S = span\left(x_i^\alpha, x_{i+1}^\alpha, \ldots x_j^\alpha\right) \cap span\left(w_k^\alpha, w_{k+1}^\alpha, \ldots w_{|\alpha|}^\alpha\right).$$

Then $\dim(S) \geq (j-i+1) + (|\alpha| - k + 1) - |\alpha| = j - i - k + 2$. Note that $y \in S, \|y\|_2 = 1$ can be written either as a linear combination of the x^α's or the w^α's. That is, if $y \in S, \|y\|_2 = 1$, then

$$y = \sum_{p=i}^{j} \beta_p x_p^\alpha = \sum_{q=k}^{|\alpha|} \gamma_q w_q^\alpha, \sum_{p=i}^{j} |\beta_p|^2 = \sum_{q=k}^{|\alpha|} |\gamma_q|^2 = 1.$$

A straightforward calculation using the second representation given above for $y \in S, \|y\|_2 = 1$ yields that $\|A[\alpha^c; \alpha]\, y\|_2 \leq \sigma_k^\alpha$. Append zeros to y to obtain vectors of the form $z = \begin{bmatrix} y \\ 0 \end{bmatrix}$ in $\mathbb{C}^n$, $y \in S, \|z\|_2 = \|y\|_2 = 1$. Then

$$\begin{aligned}
\|(A - tI)\, z\|_2^2 &= \|(A[\alpha] - tI)\, y\|_2^2 + \|A[\alpha^c; \alpha]\, y\|_2^2 \\
&= \left\| \sum_{p=i}^{j} \beta_p \left(\lambda_p^\alpha - t\right) x_p^\alpha \right\|_2^2 + \|A[\alpha^c; \alpha]\, y\|_2^2 \\
&\leq \sum_{p=i}^{j} |\beta_p|^2 \left|\lambda_p^\alpha - t\right|^2 + (\sigma_k^\alpha)^2 \leq \left|\lambda_j^\alpha - t\right|^2 + (\sigma_k^\alpha)^2 \\
&= \delta^2 + (\sigma_k^\alpha)^2 .
\end{aligned}$$

The vectors $z = \begin{bmatrix} y \\ 0 \end{bmatrix}$, $y \in S, \|y\|_2 = 1$ together with the zero vector form a subspace, $\hat{S} \subseteq \mathbb{C}^n$ that has dimension at least $j - i - k + 2$, and for which

$$\|(A - tI)\, z\|_2 \leq \sqrt{\delta^2 + (\sigma_k^\alpha)^2}$$

for any $z \in \hat{S}, \|z\|_2 = 1$. Thus, by lemma 3.2, there are at least $j-i-k+2$ eigenvalues of A in the interval $I(i, j, k, \alpha)$. □

Note that for $|\alpha| = n-1$, $A[\alpha^c;\alpha]$ is a row vector and $\sigma_2^\alpha = 0$. Classical interlacing is then the special case of theorem 3.1 in which $j = i+1, |\alpha| = n-1$, and $k = 2$. Another result that has been studied before, but can be seen as a special case of theorem 3.1 will be discussed in the next section.

It is immediate from theorem 3.1, since $\sqrt{\delta^2 + (\sigma_k^\alpha)^2} \le \delta + \sigma_k^\alpha$, $1 \le k \le |\alpha|$, that there are at least $j - i - k + 2$ eigenvalues of A in the interval $\left[\lambda_i^\alpha - \sigma_k^\alpha, \lambda_j^\alpha + \sigma_k^\alpha\right]$. This weaker statement can be proven using an application of classical perturbation theory and interlacing. There are two cases, $|\alpha^c| \ge |\alpha|$ and $|\alpha^c| \le |\alpha|$. For $|\alpha^c| \ge |\alpha|$ let $A[\alpha;\alpha^c] = V\Sigma_\alpha W^*$ be the singular value decomposition of $A[\alpha;\alpha^c]$ in which

$$\Sigma_\alpha = \begin{bmatrix} diag\left(\sigma_1^\alpha, \ldots, \sigma_{|\alpha|}^\alpha\right) & 0 \end{bmatrix}.$$

Then, the following similarity can be performed:

$$\begin{bmatrix} V^* & 0 \\ 0 & W^* \end{bmatrix} A \begin{bmatrix} V & 0 \\ 0 & W \end{bmatrix} = \begin{bmatrix} V^* A[\alpha] V & \Sigma_\alpha \\ \Sigma_\alpha^T & W^* A[\alpha^c] W \end{bmatrix}.$$

From the result, deleting the $k-1$ columns that contain $\sigma_1^\alpha, \sigma_2^\alpha, \ldots, \sigma_{k-1}^\alpha$ (in Σ_α) and the corresponding rows yields

$$\hat{A} \equiv \begin{bmatrix} V^* A[\alpha] V & \hat{\Sigma}_\alpha \\ \hat{\Sigma}_\alpha^T & \hat{A}[\alpha^c] \end{bmatrix} = \begin{bmatrix} V^* A[\alpha] V & 0 \\ 0 & \hat{A}[\alpha^c] \end{bmatrix} + \begin{bmatrix} 0 & \hat{\Sigma}_\alpha \\ \hat{\Sigma}_\alpha^T & 0 \end{bmatrix} \equiv B + C$$

in which

$$\hat{\Sigma}_\alpha = \begin{bmatrix} diag\left(\sigma_k^\alpha, \ldots, \sigma_{|\alpha|}^\alpha\right) & 0 \end{bmatrix}.$$

The eigenvalues of B are the eigenvalues of $A[\alpha]$ together with those of $\hat{A}[\alpha^c]$. As a result, there are indices r and s such that $\lambda_i^\alpha = \lambda_r(B)$ and $\lambda_j^\alpha = \lambda_s(B)$ and then $\left[\lambda_i^\alpha - \sigma_k^\alpha, \lambda_j^\alpha + \sigma_k^\alpha\right] = [\lambda_r(B) - \sigma_k^\alpha, \lambda_s(B) + \sigma_k^\alpha]$. Note that $s - r \geq j - i$ since the eigenvalues of $A[\alpha]$ may not be grouped together as eigenvalues of B; there may be eigenvalues of $\hat{A}[\alpha^c]$ between the eigenvalues $\lambda_r(B)$ and $\lambda_s(B)$. The matrix $\hat{A}$ is a perturbation of B in which the perturbation matrix C has eigenvalues $\pm\sigma_q^\alpha$, $q = k, \ldots, |\alpha|$. Classical perturbation theory then gives

$$\lambda_r(B) - \sigma_k^\alpha \leq \lambda_r\left(\hat{A}\right) \leq \lambda_s\left(\hat{A}\right) \leq \lambda_s(B) + \sigma_k^\alpha$$

which can be proved, for example, by using Weyl's inequalities (see, e.g. theorem 4.3.1 [HJ1].) Since $\hat{A}$ is a principal submatrix of a matrix similar to A, by classical interlacing there are at least $s - r - k + 2$ eigenvalues of A in the interval $\left[\lambda_r\left(\hat{A}\right), \lambda_s\left(\hat{A}\right)\right]$. Then there are at least as many eigenvalues of A in the larger interval $[\lambda_r(B) - \sigma_k^\alpha, \lambda_s(B) + \sigma_k^\alpha]$. But, since $\left[\lambda_i^\alpha - \sigma_k^\alpha, \lambda_j^\alpha + \sigma_k^\alpha\right] = [\lambda_r(B) - \sigma_k^\alpha, \lambda_s(B) + \sigma_k^\alpha]$ and $s - r \geq j - i$ there are certainly at least $j - i - k + 2$ eigenvalues of A in the interval $\left[\lambda_i^\alpha - \sigma_k^\alpha, \lambda_j^\alpha + \sigma_k^\alpha\right]$ (in fact there are at least $s - r - k + 2$.)

For the case in which $|\alpha| \geq |\alpha^c|$, if $1 \leq k \leq |\alpha^c|$ the proof is analogous. If $k > |\alpha^c|$ then $\sigma_k^\alpha = 0$ and the result follows from classical interlacing. This result is immediate from theorem 3.1, however, we know of no direct way to achieve the stronger statement of theorem 3.1 from the usual statement of interlacing.

3.3.2 Rank Deficient Off-Diagonal Blocks

An irreducible Hermitian tridiagonal matrix is of the form

$$T = \begin{bmatrix} a_1 & b_1 & & & \\ \bar{b}_1 & a_2 & b_2 & 0 & \\ & \bar{b}_2 & a_3 & \ddots & \\ & 0 & \ddots & \ddots & b_{n-1} \\ & & & \bar{b}_{n-1} & a_n \end{bmatrix}$$

in which $b_t \neq 0$ for $t = 1, 2, \ldots, n-1$. It is a well known fact (see, e.g. theorem 1 [HP]) that for $\alpha = \{1, 2, \ldots, p\}$, the open interval $\left(\lambda_i^\alpha, \lambda_{i+1}^\alpha\right)$, in which λ_i^α and λ_{i+1}^α are eigenvalues of $T[\alpha]$, contains an eigenvalue of T. The closed interval version of this fact can be seen as a special case of theorem 3.1. Because of the low rank of the off diagonal blocks ($T[\alpha^c; \alpha]$ contains only one nonzero entry) $\sigma_2^\alpha = 0$ and, then, for $k = 2$ and $j = i + 1$, by theorem 3.1 there is at least $(i+1) - i - 2 + 2 = 1$ eigenvalue of T in the interval $\left[\lambda_i^\alpha, \lambda_{i+1}^\alpha\right]$. (An auxiliary argument shows that the closed interval is actually open.) This fact can be generalized to any matrix, A, in which the off diagonal block $A[\alpha^c; \alpha]$ has low rank. If $rank(A[\alpha^c; \alpha]) = k$, then σ_{k+1}^α is zero, and by theorem 3.1, or careful application of classical interlacing as in the discussion above, the interval $\left[\lambda_i^\alpha, \lambda_{i+k}^\alpha\right]$ contains at least $(i+k) - i - (k+1) + 2 = 1$ eigenvalue of A. Thus, we have the following corollary to theorem 3.1, which was also noted in [H].

Corollary 3.6 *Let $A \in M_n$ be Hermitian and let $\alpha \subseteq N$. If $rank(A[\alpha^c; \alpha]) = k$, then there is at least one eigenvalue of A in the interval $\left[\lambda_i^\alpha, \lambda_{i+k}^\alpha\right]$.*

As seen by the Weyl inequality argument in section 3.3.1, if $rank\,(A\,[\alpha^c;\alpha]) = k$, then deletion of the last k rows and columns of $\hat{A}$ results in the direct sum

$$\begin{bmatrix} V^* A\,[\alpha]\,V & 0 \\ 0 & \hat{A}\,[\alpha^c] \end{bmatrix}$$

(since C is the 0 matrix). Corollary 3.6 then follows from usual interlacing.

Recall that the *term rank* of a matrix is the minimum number of lines (rows or columns) of the matrix that cover all of its nonzero entries. It is a simple and classical fact [Ry] that the term rank is an upper bound for the actual rank. (It is tight in the sense that there are matrices, with the same zero pattern, whose rank is the term rank.) The term rank of $A\,[\alpha^c;\alpha]$ is determined by the graph of A. For an n-by-n Hermitian matrix A, the term rank is simply the fewest vertices in $G(A)$ to which all edges between the subgraphs induced by α and α^c are incident. For example, if these two parts ($G(A\,[\alpha])$ and $G(A\,[\alpha^c])$) of $G(A)$ are connected by just one edge (as in the tridiagonal case), then the term rank of $A\,[\alpha^c;\alpha]$ is 1. The term rank will also be 1 if all edges connecting $G(A\,[\alpha])$ to $G(A\,[\alpha^c])$ are incident with just one vertex (in $G(A\,[\alpha^c])$ or $G(A\,[\alpha])$, respectively); see figure 3.7. Based upon this discussion, we make several combinatorially based observations.

Corollary 3.7 *Suppose that $A \in M_n$ is Hermitian, $\alpha \subseteq N$ and that the term rank of $A\,[\alpha^c;\alpha]$ is k. Then, there is at least one eigenvalue of A in the interval* $\left[\lambda_i^\alpha, \lambda_{i+k}^\alpha\right]$.

The special case $k = 1$ of corollary 3.7 yields a large class of generalizations of the tridiagonal fact mentioned at the beginning of this section. In particular, if $G(A)$ is a tree T, there will be many choices of α that generalize the "half" of a path that gives the tridiagonal case. More generally, if $G(A)$ is a tree T, identify any vertex v of T (degree $v = p \geq 2$) as the root and consider the branches

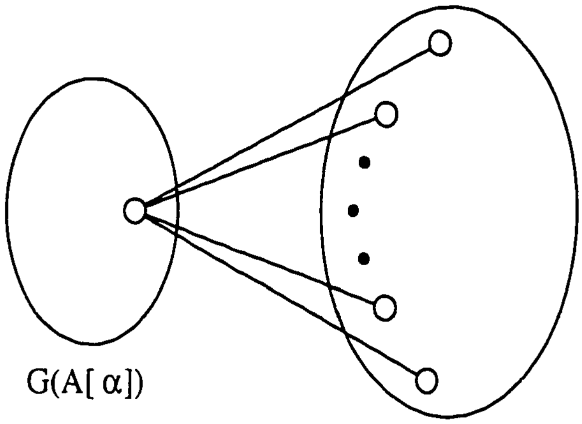


Figure 3.7: Term rank: $G(A)$

$B_1, B_2, \ldots, B_p$ of T emanating from v. Then, choose any proper partition $\{i_1, i_2, \ldots, i_q\} \cup \{j_1, j_2, \ldots, j_r\} = \{1, 2, \ldots, p\}$ and let α be the vertices of $B_{i_1}, B_{i_2}, \ldots, B_{i_q}$ together with v, so that α^c is the vertices of $B_{j_1}, B_{j_2}, \ldots, B_{j_r}$. Then, the term rank of $A[\alpha^c; \alpha]$ is 1 and the interval $[\lambda_i^\alpha, \lambda_{i+1}^\alpha]$ will contain an eigenvalue of A. If $A \in M_n(\mathbb{R})$ is a combinatorially symmetric matrix such that $G(A)$ is a tree T, then there exists a diagonal matrix D for which $B = D^{-1}AD$ is symmetric and $G(B) = T$. Therefore, the above discussion also applies to these *non*-Hermitian matrices.

3.3.3 A Union of Two Intervals

The discussion surrounding theorem 3.1 considers only intervals formed by eigenvalues of the submatrix $A[\alpha]$. Of course, theorem 3.1 equally applies to intervals determined by eigenvalues of $A[\alpha^c]$. This leads naturally to the question: what about the union of two such intervals? If these two intervals are disjoint, their union will contain the sum of the estimated number of eigenvalues of A in each interval. If the intervals are not disjoint, however, there may be coincidence of eigenvalues and the union may not contain the sum of the estimated number of eigenvalues of A in each interval. An example following the next theorem shows that indeed, the union of the two intervals from theorem 3.1 may not capture the expected sum of the number of eigenvalues of A that the two intervals capture separately. However, if each interval is made slightly larger, then, regardless of whether or not the intervals are disjoint, their union contains the desired sum of the estimated number of eigenvalues of A. The two intervals $\left[\lambda_{i_1}^{\alpha} - \sigma_k, \lambda_{j_1}^{\alpha} + \sigma_k\right]$ and $\left[\lambda_{i_2}^{\alpha^c} - \sigma_k, \lambda_{j_2}^{\alpha^c} + \sigma_k\right]$ (in which $\sigma_k = \sigma_k^{\alpha} = \sigma_k^{\alpha^c}$) constructed using eigenvalues from $A[\alpha]$ and $A[\alpha^c]$, respectively, are generally larger than the theorem 3.1 intervals. The Weyl inequality argument shows that there are at least $j_1 - i_1 - k + 2$ eigenvalues of A in the first interval and $j_2 - i_2 - k + 2$ eigenvalues of A in the second. Although the Weyl inequality argument does not easily enumerate the number of eigenvalues of A in the union when the two intervals are not disjoint, the techniques used in the proof of theorem 3.1 may be used to show that the union of these two intervals contains at least $j_1 + j_2 - (i_1 + i_2) - 2k + 4$ eigenvalues of A whether or not the intervals are disjoint.

Theorem 3.2 *Let $A \in M_n$ be Hermitian and let $\alpha \subseteq N$; then*

$$\left[\lambda_{i_1}^{\alpha} - \sigma_k, \lambda_{j_1}^{\alpha} + \sigma_k\right] \cup \left[\lambda_{i_2}^{\alpha^c} - \sigma_k, \lambda_{j_2}^{\alpha^c} + \sigma_k\right]$$

contains at least $j_1 + j_2 - (i_1 + i_2) - 2k + 4$ eigenvalues of A in which σ_k is the k^{th} singular value of both $A[\alpha^c; \alpha]$ and $A[\alpha; \alpha^c]$.

Proof: If the intervals $\left[\lambda_{i_1}^{\alpha} - \sigma_k, \lambda_{j_1}^{\alpha} + \sigma_k\right]$ and $\left[\lambda_{i_2}^{\alpha^c} - \sigma_k, \lambda_{j_2}^{\alpha^c} + \sigma_k\right]$ are disjoint, the conclusion is obviously true by theorem 3.1 and the above comments.

Suppose $\left[\lambda_{i_1}^{\alpha} - \sigma_k, \lambda_{j_1}^{\alpha} + \sigma_k\right] \cap \left[\lambda_{i_2}^{\alpha^c} - \sigma_k, \lambda_{j_2}^{\alpha^c} + \sigma_k\right] \neq \emptyset$. Then

$$\left[\lambda_{i_1}^{\alpha} - \sigma_k, \lambda_{j_1}^{\alpha} + \sigma_k\right] \cup \left[\lambda_{i_2}^{\alpha^c} - \sigma_k, \lambda_{j_2}^{\alpha^c} + \sigma_k\right] = \left[\lambda_{\min} - \sigma_k, \lambda_{\max} + \sigma_k\right]$$

in which $\lambda_{\min} = \min\left\{\lambda_{i_1}^{\alpha}, \lambda_{i_2}^{\alpha^c}\right\}$, and $\lambda_{\max} = \max\left\{\lambda_{j_1}^{\alpha}, \lambda_{j_2}^{\alpha^c}\right\}$. As in the proof of theorem 3.1, define

$$S_1 = span\left(x_{i_1}^{\alpha}, x_{i_1+1}^{\alpha}, \ldots, x_{j_1}^{\alpha}\right) \cap span\left(w_k^{\alpha}, w_{k+1}^{\alpha}, \ldots, w_{|\alpha|}^{\alpha}\right)$$

and

$$S_2 = span\left(x_{i_2}^{\alpha^c}, x_{i_2+1}^{\alpha^c}, \ldots, x_{j_2}^{\alpha^c}\right) \cap span\left(w_k^{\alpha^c}, w_{k+1}^{\alpha^c}, \ldots, w_{|\alpha^c|}^{\alpha^c}\right)$$

in which x_p^{α} $\left(x_p^{\alpha^c}\right)$ are eigenvectors of $A[\alpha]$ $(A[\alpha^c])$ associated 'vith λ_p^{α} $\left(\lambda_p^{\alpha^c}\right)$and w_q^{α} $\left(w_q^{\alpha^c}\right)$ are right singular vectors of $A[\alpha^c;\alpha]$ $(A[\alpha;\alpha^c])$. Recall that, for simplicity of notation, we have assumed $\alpha = \{1, 2, \ldots, |\alpha|\}$. After embedding S_1 and S_2 in $\mathbb{C}^n$ by appending zeros in the appropriate spots, we have $(S_1 \oplus 0) \cap (0 \oplus S_2) = 0$. Let $\hat{S} \equiv S_1 \oplus S_2$. Then, $\dim(S_1) \geq j_1 - i_1 - k + 2$, $\dim(S_2) \geq j_2 - i_2 - k + 2$ as before, and $\dim\left(\hat{S}\right) \geq (j_1 + j_2) - (i_1 + i_2) - 2k + 4$. Let $z \in \hat{S}$, $\|z\|_2 = 1$. Then

$$z = \begin{bmatrix} \beta_1 y_1 \\ \beta_2 y_2 \end{bmatrix}, y_1 \in S_1, y_2 \in S_2, \|y_1\|_2 = \|y_2\|_2 = 1, |\beta_1|^2 + |\beta_2|^2 = 1$$

and

$$\begin{aligned} y_1 &= \sum_{p=i_1}^{j_1} \gamma_p^{\alpha} x_p^{\alpha} = \sum_{q=k}^{|\alpha|} \mu_q^{\alpha} w_q^{\alpha}, \\ y_2 &= \sum_{p=i_2}^{j_2} \gamma_p^{\alpha^c} x_p^{\alpha^c} = \sum_{q=k}^{|\alpha^c|} \mu_q^{\alpha^c} w_q^{\alpha^c}, \\ \sum_{p=i_1}^{j_1} \left|\gamma_p^{\alpha}\right|^2 &= \sum_{q=k}^{|\alpha|} \left|\mu_q^{\alpha}\right|^2 = \sum_{p=i_2}^{j_2} \left|\gamma_p^{\alpha^c}\right|^2 = \sum_{q=k}^{|\alpha^c|} \left|\mu_q^{\alpha^c}\right|^2 = 1. \end{aligned}$$

Define $t \equiv \frac{\lambda_{\min}+\lambda_{\max}}{2}$ and $\delta \equiv \frac{\lambda_{\max}-\lambda_{\min}}{2}$. Then

$$\|(A[\alpha] - tI)\, y_1\|_2^2 = \sum_{p=i_1}^{j_1} \left|\lambda_p^\alpha - t\right|^2 \left|\gamma_p^\alpha\right|^2 \left\|x_p^\alpha\right\|_2^2 \le |\lambda_{\max} - t|^2 = \delta^2$$

and

$$\|A[\alpha^c;\alpha]\, y_1\|_2^2 = \sum_{q=k}^{|\alpha^c|} \sigma_q^2 \left|\mu_q^{\alpha^c}\right|^2 \left\|w_q^\alpha\right\|_2^2 \le \sigma_k^2.$$

Similarly,

$$\|(A[\alpha^c] - tI)\, y_2\|_2^2 \le |\lambda_{\max} - t|^2 = \delta^2 \text{ and } \|A[\alpha;\alpha^c]\, y_2\|_2^2 \le \sigma_k^2.$$

Then

$$\begin{aligned}
\|(A - tI)\, z\|_2^2 &= \|(A[\alpha] - tI)(\beta_1 y_1) + A[\alpha;\alpha^c](\beta_2 y_2)\|_2^2 \\
&\quad + \|(A[\alpha^c] - tI)(\beta_2 y_2) + A[\alpha^c;\alpha](\beta_1 y_1)\|_2^2 \\
&\le (|\beta_1|\, \|(A[\alpha] - tI)\, y_1\|_2 + |\beta_2|\, \|A[\alpha;\alpha^c]\, y_2\|_2)^2 \\
&\quad + (|\beta_2|\, \|(A[\alpha^c] - tI)\, y_2\|_2 + |\beta_1|\, \|A[\alpha^c;\alpha]\, y_1\|_2)^2 \\
&\le (|\beta_1|\, |\lambda_{\max} - t| + |\beta_2|\, \sigma_k)^2 + (|\beta_2|\, |\lambda_{\max} - t| + |\beta_1|\, \sigma_k)^2 \\
&= |\lambda_{\max} - t|^2 + 4\, |\beta_1|\, |\beta_2|\, |\lambda_{\max} - t|\, \sigma_k + \sigma_k^2 \\
&\le |\lambda_{\max} - t|^2 + 2\, |\lambda_{\max} - t|\, \sigma_k + \sigma_k^2 \\
&= (|\lambda_{\max} - t| + \sigma_k)^2 .
\end{aligned}$$

Thus, $\|(A - tI)\, z\|_2 \le \delta + \sigma_k$ for any $z \in \hat{S}$, $\|z\|_2 = 1$. By lemma 3.2, there are at least $(j_1 + j_2) - (i_1 + i_2) - 2k + 4$ eigenvalues of A in the interval $[t - (\delta + \sigma_k), t + (\delta + \sigma_k)] = [\lambda_{\min} - \sigma_k, \lambda_{\max} + \sigma_k]$. □

An example will illustrate that the two interval statement in theorem 3.2 cannot be as tight as the unions of the two smaller intervals in theorem 3.1. The eigenvalues of

$$J = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

are 0, 0, and 3. By theorem 3.1, the interval $I(1,1,1,\{1\}) = \left[1-\sqrt{2}, 1+\sqrt{2}\right]$ captures at least one eigenvalue of J, whereas $I(1,2,1,\{2,3\}) = \left[1-\sqrt{3}, 1+\sqrt{3}\right]$ captures at least two eigenvalues of J. However, the union of the theorem 3.1 intervals,

$$I(1,1,1,\{1\}) \cup I(1,2,1,\{2,3\}) = \left[1-\sqrt{3}, 1+\sqrt{3}\right],$$

captures only the two 0 eigenvalues of J. The interval given in theorem 3.2 is

$$\left[1-\sqrt{2}, 1+\sqrt{2}\right] \cup \left[0-\sqrt{2}, 2+\sqrt{2}\right] = \left[0-\sqrt{2}, 2+\sqrt{2}\right]$$

which does capture all three eigenvalues of J. Thus, a larger interval, such as in theorem 3.2, is necessary to capture the $j_1 + j_2 - (i_1 + i_2) - 2k + 4 = 3$ eigenvalues of J. This shows that theorem 3.2 cannot be improved to the union of the smaller intervals in theorem 3.1.

3.3.4 Applications to Singular Values and Lehmann's Intervals

Application to Singular Values

Once again the similarities between classical eigenvalue interlacing and singular value interlacing indicate that there may be a theorem analogous to theorem 3.1 for singular values. For $A \in M_{m,n}$, define

$$SI\left(A, i, j, k, \{\alpha, \beta\}\right) \equiv \left[\sigma_i\left(A\left[\alpha;\beta\right]\right) - \eta_k, \sigma_j\left(A\left[\alpha;\beta\right]\right) + \eta_k\right]$$

in which η_k is the k^{th} largest singular value of

$$\begin{bmatrix} 0 & A\left[\alpha;\beta^c\right] \\ A^*\left[\beta^c;\alpha\right] & 0 \end{bmatrix}.$$

Then the singular value result is:

Corollary 3.8 *Let $A \in M_{m,n}$ and let $\alpha \subseteq M, \beta \subseteq N$. Then A has at least $j-i-k+2$ singular values in the interval $SI\left(A, i, j, k, \{\alpha, \beta\}\right)$.*

Proof: Assume $m \geq n$ (if not, interchange the rolls of rows and columns in the following discussion). Let C be the Hermitian matrix

$$C \equiv \begin{bmatrix} 0 & A \\ A^* & 0 \end{bmatrix}$$

and let $\tilde{\alpha} = \alpha \cup (\beta + \{m\})$. Then

$$C[\tilde{\alpha}] = \begin{bmatrix} 0 & A[\alpha;\beta] \\ A^*[\beta;\alpha] & 0 \end{bmatrix}.$$

The eigenvalues of C are plus and minus the singular values of A, and those of $C[\tilde{\alpha}]$ are plus and minus the singular values of $A[\alpha;\beta]$ (both with, possibly, some extra 0 singular values). Thus, $\sigma_s(A[\alpha;\beta]) = \lambda_{t-s+1}^{\tilde{\alpha}}$ in which $t = |\alpha| + |\beta|$, and $s = 1, 2, \ldots, min\{|\alpha|, |\beta|\}$. Then

$$I(C, t-i+1, t-j+1, k, \tilde{\alpha}) = SI(A, i, j, k, \{\alpha, \beta\})$$

and by theorem 3.1 there are at least $j-i-k+2$ eigenvalues of C in $I(C, t-i+1, t-j+1, k, \tilde{\alpha})$ and, since $\sigma_s(A) = \lambda_{m+n-s+1}(C)$, $s = 1, 2, \ldots, n$, there are, then, at least $j-i-k+2$ singular values of A in $SI(A, i, j, k, \{\alpha, \beta\})$. □

Lehmann's Intervals

The techniques used in the proof of theorem 3.1 can be used to give an elementary proof of the sufficiency of the optimal inclusion intervals attributed to Lehmann [L].

Theorem 3.3 *Let $A \in M_n$ be Hermitian, $\alpha \subseteq N$, $|\alpha| = k$, and let $x \in \mathbb{R}$ and $0 \leq j \leq k$ be given. Then, an interval of the form $[x-\delta, x+\delta]$ that contains j eigenvalues of A is given by $\delta = \sigma_{k-j+1}(B)$ with* $B = \begin{bmatrix} A[\alpha] - xI \\ A[\alpha^c;\alpha] \end{bmatrix}$.

Proof: Assume, without loss of generality, that $x = 0$ (the general case follows by translation). Let the eigenvalues of A be ordered so that

$\lambda_1^2 \geq \lambda_2^2 \geq \cdots \geq \lambda_n^2$ and let $w_1, w_2, \ldots, w_k$ be the right singular vectors of B corresponding to $\sigma_1(B), \sigma_2(B), \ldots, \sigma_k(B)$, respectively. Then, define $S = span\{v_{w-j+1}, \ldots, w_k\}$ and $z = \begin{bmatrix} y \\ 0 \end{bmatrix} \in C^n$, $y \in S$, $\|y\|_2 = 1$ so that $Az = By$.

By Courant-Fischer min-max conditions for eigenvalues (e.g. theorem 4.2.11 [HJ1]) and singular values (e.g. theorem 7.3.10 [HJ1]) we have:

$$\lambda_{n-j+1}^2 = \min_{\dim(W)=j} \max_{z \in W} (z^* A^* A z) \leq \max_{y \in S} (y^* B^* B y) \leq \sigma_{k-j+1}(B).$$

The middle inequality follows since S has dimension j so that $S \oplus 0_{n-k}$ is a candidate for W. Therefore, there are at least j eigenvalues of A in the interval $[-\sigma_{k-j+1}(B), \sigma_{k-j+1}(B)]$. □

Note that the number δ does not depend on $A[\alpha^c]$ and the interval may be determined whether or not $A[\alpha^c]$ is known. Lehmann has further shown that this value of δ is the smallest one for which j eigenvalues are captured, independent of $A[\alpha^c]$. This is equivalent to

$$\sup_{C=C^*} \sigma_{n-j+1}\left(\begin{bmatrix} A[\alpha]-xI & A[\alpha;\alpha^c] \\ A[\alpha^c;\alpha] & C-xI \end{bmatrix}\right) = \sigma_{k-j+1}\left(\begin{bmatrix} A[\alpha]-xI \\ A[\alpha^c;\alpha] \end{bmatrix}\right).$$

It follows from theorem 2.2 in [GRSW] that when the condition of Hermicity is dropped for the matrix C then

$$\sup_{C} \sigma_{n-j+1}\left(\begin{bmatrix} A[\alpha]-xI & A[\alpha;\alpha^c] \\ A[\alpha^c;\alpha] & C-xI \end{bmatrix}\right) = \sigma_{k-j+1}\left(\begin{bmatrix} A[\alpha]-xI \\ A[\alpha^c;\alpha] \end{bmatrix}\right).$$

Chapter 4

Matrix Completions

Recall that a *matrix completion problem* asks when a given partial matrix has a completion with some desired property. The properties we are interested in are inherited by principal submatrices, i.e. any principal submatrix of a matrix with the given property also has the property. If we are interested in property Φ, then A is said to be a *partial "property Φ" matrix* if every specified principal submatrix of A has property Φ. For example, a partial positive definite matrix is a partial Hermitian matrix in which every fully specified principal submatrix is positive definite.

If property Φ is inherited, then any partial matrix that has a completion with property Φ must be a partial Φ matrix. For a number of interesting properties (e.g. positive semidefinite and distance matrices) it has been shown that this obvious necessary condition on the specified principal submatrices is also sufficient exactly when the graph of the specified entries is chordal. If the graph is not chordal, more information needs to be known about the specified data. For a survey on matrix completion problems see [J].

4.1 Positive Semidefinite Completions

An important class of completion problems that has received considerable attention is the positive semidefinite completion problem. Since positive semidefinite matrices satisfy the inheritance property, it is necessary that every fully specified principal submatrix of the matrix in question is positive semidefinite. It was shown in [GJSW] that this condition is sufficient to ensure a positive semidefinite completion of a partial positive semidefinite matrix A exactly when $G(A)$ is chordal. This chordal result is a generalization of the earlier result by [DG] for banded matrices.

If the graph of the specified entries of a partial positive semidefinite matrix is not chordal more needs to be known about the data. A non-chordal graph contains a chordless cycle of length 4 or more. Necessary and sufficient conditions that ensure a positive semidefinite completion for the n-cycle, $n \geq 4$, were given in [BJT], and [F] is a precursor from a different point of view. For an n-cycle these may be taken to be $\lceil \frac{n}{2} \rceil$ conditions transcendental in the data for an n-cycle to admit a positive semidefinite completion. We will call these the *cycle conditions*. After a diagonal congruence transformation of the data, a partial positive semidefinite matrix A in which $G(A)$ is an n-cycle is of the form

$$A = \begin{bmatrix} 1 & \cos\theta_{i_1} & ? & \cos\theta_{i_n} \\ \cos\theta_{i_1} & 1 & \ddots & ? \\ ? & \ddots & \ddots & \cos\theta_{i_{n-1}} \\ \cos\theta_{i_n} & ? & \cos\theta_{i_{n-1}} & 1 \end{bmatrix}$$

in which $\pi \geq \theta_1 \geq \theta_2 \geq \ldots \geq \theta_n \geq 0$ and $i_1, i_2, \ldots, i_n$ is the permutation of $1, 2, \ldots, n$ in which these θ's occur around the cycle. Note that such a transformation does not change the existence of a positive semidefinite completion. According to corollary 1 in [BJT] the conditions that ensure a positive semidefinite completion of A are,

$$\sum_{i=1}^{k} \theta_i \leq (k-1)\pi + \sum_{i=k+1}^{n} \theta_i$$

for every $1 \leq k \leq n$, k odd. In [BJL] those graphs for which these cycle conditions together with partial positive definiteness are sufficient to ensure a positive semidefinite completion are characterized.

It is interesting to note that the cycle conditions are independent of the order of the θ_i's around the cycle. This can be seen without knowing the [BJT] result and is also true for the case in which the data is from the complex numbers. For the cycle

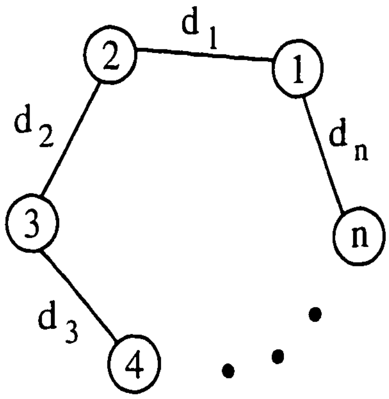


Figure 4.1: Cycle

we will adopt the convention that in the partial matrix A with this graph, $a_{ij} = d_i$ and $a_{ji} = \bar{d}_i$ for $i < j$.

Theorem 4.1 *Let*

$$A = \begin{bmatrix} 1 & d_1 & & & d_n \\ \bar{d}_1 & 1 & d_2 & ? & \\ & \bar{d}_2 & 1 & \ddots & \\ & ? & \ddots & \ddots & d_{n-1} \\ \bar{d}_n & & & \bar{d}_{n-1} & 1 \end{bmatrix}. \tag{4.1}$$

Then A has a positive semidefinite completion if and only if

$$\tilde{A} = \begin{bmatrix} 1 & d_{\sigma(1)} & & & d_{\sigma(n)} \\ \bar{d}_{\sigma(1)} & 1 & d_{\sigma(2)} & ? & \\ & \bar{d}_{\sigma(2)} & 1 & \ddots & \\ & ? & \ddots & \ddots & d_{\sigma(n-1)} \\ \bar{d}_{\sigma(n)} & & & \bar{d}_{\sigma(n-1)} & 1 \end{bmatrix}.$$

has a positive semidefinite completion for any permutation σ of the indices $i = 1, 2, \ldots, n$.

Proof: It suffices to show that the data on two consecutive edges in the cycle can be interchanged and the completability of the corresponding transformed matrix is equivalent to that of the original matrix. Then, the general case in which any permutation of the data results in a corresponding matrix with a positive semidefinite completion is obtained from a finite sequence of consecutive interchanges. Assume that the matrix A as in (4.1) has a positive semidefinite completion. Then $G(A)$ is given by the graph in figure 4.1. Let G' be a triangulation (add edges until the graph becomes chordal and no cliques of size 4 or more are created) of $G(A)$ in which $\{1, 3\}$ is an edge of G'. For example,

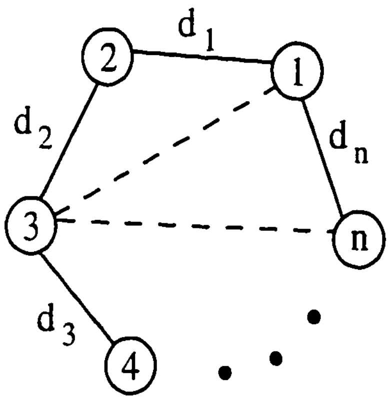


Figure 4.2: Triangulation.

It is a necessary condition that A be partial positive semidefinite. So, there exists a partial positive semidefinite matrix A' such that $G(A') = G'$, $A'_{1n} = A_{1n}$, $A'_{n1} = A_{n1}$, $A'_{i,i+1} = A_{i,i+1}$, and $A'_{i,i+1} = A_{i,i+1}$ for $i = 1, 2, \ldots, n-1$. Assume the value given to the edge $\{1,3\}$ is z. The triangulated graph G' is chordal so that by [GJSW] A' has a positive semidefinite completion.

The principal submatrix $A'[1,2,3]$ is fully specified and has data $\begin{bmatrix} 1 & d_1 & z \\ \bar{d}_1 & 1 & d_2 \\ \bar{z} & \bar{d}_2 & 1 \end{bmatrix}$.

Permuting the first two rows and columns of $A'[1,2,3]$ results in the matrix

$$B = \begin{bmatrix} 1 & \bar{d}_1 & \bar{z} \\ d_1 & 1 & \bar{d}_2 \\ z & d_2 & 1 \end{bmatrix}.$$

which is positive semidefinite since permutation similarity preserves positive semi-definiteness. But then

$$B^T = \begin{bmatrix} 1 & d_2 & z \\ \bar{d}_2 & 1 & d_1 \\ \bar{z} & \bar{d}_1 & 1 \end{bmatrix}$$

is also positive semidefinite. So, replacing $A'[1,2,3]$ by B^T results in a new matrix $\hat{A}$ that is partial positive semidefinite and for which $G(\hat{A})$ is chordal and $G(\hat{A})$ is the same graph as G' except that d_1 and d_2 are interchanged. Since $\hat{A}$ is chordal, by [GJSW], $\hat{A}$ has a positive semidefinite completion. Let $\check{A}$ be the submatrix of $\hat{A}$ for which $G\left(\check{A}\right)$ is given by:

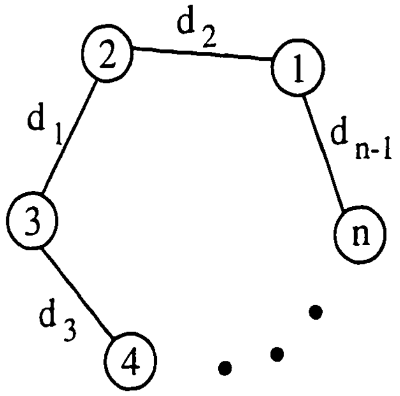


Figure 4.3: $G(A')$.

If $\check{A}$ does not have a positive semidefinite completion, then $\hat{A}$ cannot have a positive semidefinite completion. But $\hat{A}$ is chordal and, thus, has a positive semidefinite completion. It follows that $\check{A}$ has a positive semidefinite completion. Since this argument is symmetric, this completes the proof. □

Note that the proof of theorem 4.1 is equally valid for data from the real or the complex numbers and, therefore, in any list of cycle conditions for matrices with data from the complex numbers, the order of the data around the cycle must be irrelevant.

In [BJT] and in [F] only *real* completions of real partial positive semidefinite matrices the graph of whose specified entries is a cycle are considered. In this case, for each maximal clique that a given unspecified entry completes there is an interval of allowed values for a positive semidefinite completion. This interval becomes a disc in the complex case. In each case we must look at intersections of the sets of allowed values for each maximal clique that the entry completes. In the real case the intersection of intervals is still an interval. As a result, the real case is relatively easy. However, the intersection of two discs is no longer a disc and, therefore, the complex case is much more difficult.

Let $A = B + iC$ be a partial positive semidefinite matrix in which B, C are real and $G(A) = G(B) = G(C)$ is an n-cycle. Then

$$A = \begin{bmatrix} 1 & b_1 & ? & b_n \\ b_1 & 1 & \ddots & ? \\ ? & \ddots & 1 & b_{n-1} \\ b_n & ? & b_{n-1} & 1 \end{bmatrix} + i \begin{bmatrix} 0 & c_1 & ? & c_n \\ -c_1 & 0 & \ddots & ? \\ ? & \ddots & 0 & c_{n-1} \\ -c_n & ? & -c_{n-1} & 0 \end{bmatrix}.$$

If A is positive semidefinite, so is $A^T = \bar{A}$. Then $\frac{A+A^T}{2} = B$ is also positive semidefinite. Thus, the cycle conditions of [BJT] give necessary conditions on the data in B. Let

$$D = \begin{bmatrix} 1 & & & & \\ & -i & & & \\ & & 1 & & \\ & & & -i & \\ & & & & \ddots \end{bmatrix},$$

then

$$Re\left(D^{-1}AD\right) = \begin{bmatrix} 1 & c_1 & ? & c_n \\ -c_1 & 1 & \ddots & ? \\ ? & \ddots & 1 & c_{n-1} \\ -c_n & ? & -c_{n-1} & 1 \end{bmatrix}.$$

Therefore, the conditions in [BJT] also provide necessary conditions on the data in C. This is a start in determining a list of conditions on the data that ensures a completion in the complex cycle case. However, exactly what the complex cycle conditions are remains an open question. The following theorem does not by itself give such a list, but it does indicate that the complex positive semidefinite completion problem is a special case of a larger real completion problem.

Theorem 4.2 *Let G be any graph and $A = B + iC$, in which B, C are real, be a partial positive semidefinite matrix with graph G. Then, A has a positive semidefinite completion if and only if*

$$M = \begin{bmatrix} B & C \\ -C & B \end{bmatrix}$$

has a positive semidefinite real completion.

Proof: Assume that A has a positive semidefinite completion $\tilde{A} = \tilde{B} + i\tilde{C}$ in which $\tilde{B}$ and $\tilde{C}$ are real matrices. Then

$$\tilde{M} = \begin{bmatrix} \tilde{B} & \tilde{C} \\ -\tilde{C} & \tilde{B} \end{bmatrix}$$

is a completion of M. Note that since $\tilde{A}$ is Hermitian, $\tilde{C}$ is such that $\tilde{C}^T = -\tilde{C}$ and $\tilde{M}$

is a real symmetric matrix. If x is an eigenvector of $\tilde{A}$ associated with λ, then $\begin{bmatrix} x \\ ix \end{bmatrix}$ is an eigenvector of $\tilde{M}$ associated with λ. That is, any eigenvalue of $\tilde{A}$ is also an eigenvalue of $\tilde{M}$. Also, $\begin{bmatrix} \bar{x} \\ -i\bar{x} \end{bmatrix}$ is an eigenvector of $\tilde{M}$ associated with $\bar{\lambda}$. Since λ is real, $\bar{\lambda} = \lambda$, and for $x \neq 0$ the vectors $\begin{bmatrix} x \\ ix \end{bmatrix}$ and $\begin{bmatrix} \bar{x} \\ -i\bar{x} \end{bmatrix}$ are linearly independent. Therefore, the eigenvalues of $\tilde{M}$ are the eigenvalues of $\tilde{A}$ (with multiplicities doubled) and are nonnegative. Thus, $\tilde{M}$ is a positive semidefinite completion of M.

Now assume M has a real positive semidefinite completion

$$\tilde{M}_1 = \begin{bmatrix} \tilde{B}_1 & \tilde{C} \\ \tilde{C}^T & \tilde{B}_2 \end{bmatrix}.$$

Note that although the diagonal blocks are equal in the partial matrix, the completion may not have $\tilde{B}_1 = \tilde{B}_2$. Similarly, it may not be the case that $\tilde{C}^T = -\tilde{C}$ as required in order to recover a completion of A from $\tilde{M}_1$. The following similarity leads to another positive semidefinite completion, $\tilde{M}_2$, of M :

$$\begin{bmatrix} 0 & iI \\ -iI & 0 \end{bmatrix} \begin{bmatrix} \tilde{B}_1 & \tilde{C} \\ \tilde{C}^T & \tilde{B}_2 \end{bmatrix} \begin{bmatrix} 0 & iI \\ -iI & 0 \end{bmatrix} = \begin{bmatrix} i\tilde{C}^T & i\tilde{B}_2 \\ -i\tilde{B}_1 & -i\tilde{C} \end{bmatrix} \begin{bmatrix} 0 & iI \\ -iI & 0 \end{bmatrix}$$
$$= \begin{bmatrix} \tilde{B}_2 & -\tilde{C}^T \\ -\tilde{C} & \tilde{B}_1 \end{bmatrix} = \tilde{M}_2.$$

Note that the specified entries of B_1 and B_2 are identical, as are those of C and

$-C = C^T$, so, this is indeed a second completion of M. Then

$$\frac{1}{2}\left(\tilde{M}_1 + \tilde{M}_2\right) = \begin{bmatrix} \frac{1}{2}\left(\tilde{B}_1 + \tilde{B}_2\right) & \frac{1}{2}\left(\tilde{C} - \tilde{C}^T\right) \\ -\frac{1}{2}\left(\tilde{C} - \tilde{C}^T\right) & \frac{1}{2}\left(\tilde{B}_1 + \tilde{B}_2\right) \end{bmatrix} = \tilde{M}$$

which is yet another completion of M. Note that $\tilde{M}$ has the desired form since the diagonal blocks are identical and the off diagonal blocks are the negative of one another. Since $\tilde{M}_1$ and $\tilde{M}_2$ are positive semidefinite, $\tilde{M}$ is also positive semidefinite. Then, the matrix $\tilde{A} = \frac{1}{2}\left(\tilde{B}_1 + \tilde{B}_2\right) + \frac{i}{2}\left(\tilde{C} - \tilde{C}^T\right)$ is a positive semidefinite completion of A. □

4.2 The Euclidean Distance Completion Problem: Cycle Completability

An n-by-n matrix $D = (d_{ij})$ is called a (Euclidean) *distance matrix* if, for some k, there exist points $p_1, p_2, \ldots, p_n \in \mathbb{R}^k$ such that $d_{ij} = \|p_i - p_j\|_2, i, j = 1, 2, \ldots, n$, in which $\|\bullet\|_2$ is the Euclidean vector norm on $\mathbb{R}^k$. A *partial distance matrix* is a partial symmetric matrix in which every specified principal submatrix is a (Euclidean) distance matrix. Generally, it is assumed that all diagonal entries are specified and in the case of distance matrices the diagonal entries are all 0.

Here our interest lies in the general question of determining, for each graph G, conditions on the specified data in a partial distance matrix with graph G that ensure the existence of a distance matrix completion. This interest is motivated in part by the "molecular conformation problem", (see [HC] and [dLH]) in which some measured interatomic distances must be fit to a distance matrix of an entire molecule.

The distance completion problem enjoys the property of "inheritance", so, in order that a partial symmetric matrix A have a distance matrix completion, any principal submatrix of A must have a distance matrix completion. Thus, it is necessary that A be a partial distance matrix.

Recently, it has been shown that, as in the positive semidefinite case, this obvious necessary condition is sufficient exactly when the graph G of the specified entries of A is chordal [BJ]. Otherwise, more need be known about the data. There is a very strong relationship between the distance matrices and the positive (semi-)definite matrices [B]. Unfortunately, these links do not simply extend to the two completion problems [JT], but, nonetheless, it is reasonable to expect strong analogies between completion results.

Here, the focus is upon nonchordal graphs G and, in particular, upon the role of completability conditions for a single full cycle. In the case of the distance completion problem, the conditions for a simple cycle of data are much simpler than in the positive semidefinite case, both to state and to understand. The purpose here is to answer the question, parallel to the one resolved in [BJL] for the positive (semi-)definite case, "for which graphs are the cycle conditions sufficient for a partial distance matrix to have a distance matrix completion?" Importantly, the class of graphs is the same as in the positive (semi-)definite case, but, also importantly, there are notable differences in the details (though not the overall structure) of the proof.

As in [BJL], considerable graph theoretic structure is necessary to carry out this proof and the same notation and definitions as in [BJL] will generally be adopted. Nonstandard concepts and notation will be defined as they arise.

In the case of a partial distance matrix, suppose $d_{12}, d_{23}, \ldots, d_{k-1,k}, d_{k1} \geq 0$ is a k-cycle of specified data. Then, there is a single simple condition that is necessary and sufficient for a distance matrix completion:

$$2\max\{d_{12}, d_{23}, \ldots, d_{k-1,k}, d_{k1}\} \leq d_{12} + d_{23} + \cdots + d_{k-1,k} + d_{k1}.$$

We refer to this inequality as the polygonal inequality. Its necessity follows from repeated application of the triangle inequality, while its sufficiency may be seen inductively, initiated with the case of a triangle, in a variety of ways. For example, two are mentioned here. (1) By adjustment of the angle between the largest distance and one adjacent to it, the k distances $d_{12}, d_{23}, \ldots, d_{k-1,k}, d_{k1}$ may be replaced by $k-1$ distances: a "new" distance lying between the largest plus or minus an adjacent distance, and the $k-2$ other unused distances. The angle may be chosen so that the new set of distances satisfies the polygonal inequality and, thus, is achievable by the induction hypothesis. (2) If $k = 3$, the cycle is a triangle and in this case the cycle condition (i.e. polygonal inequality) is necessary and sufficient for the specified data to be a distance matrix. Therefore, it may be assumed that there are at least four distances. Replace the adjacent pair which has the smallest sum by their sum. If this is not then the largest distance, the polygonal inequality still holds and the induction hypothesis applies, realizing all k distances with the two "small" distances lying on a line. Otherwise, if the sum of the two is then the largest distance, the polygonal inequality still holds, as there are at least four distances originally and the smallest adjacent sum was chosen. The induction hypothesis again applies, realizing the two "small" distances on a line. It should be noted that the cycle of distances may always be realized in a plane, i.e. $\mathbb{R}^2$, and that, as in the positive (semi-)definite case, the order of the data around the cycle is irrelevant.

4.2.1 Statement of Main Result

In order to describe our main result of this section, several definitions are needed. We say that a partial distance matrix A is *distance cycle completable* if every principal submatrix of A, corresponding to a minimal cycle in the graph of the specified entries of A, has a distance matrix completion, i.e. the data for the cycle satisfies the polygonal inequality. We also informally say that the data satisfy the "cycle conditions." Recall that an *edge subdivision* of a graph G on n vertices is a graph G' on $n+1$ vertices that results from replacing an edge of G with two edges and a vertex between, and a *vertex partition* of G (n vertices) is a graph G' ($n+1$ vertices) in which a vertex (of degree at least 1) in G is replaced by two adjacent vertices that partition the neighbors of the original vertex. Also recall, that a graph G_2 is *homeomorphic* to a graph G_1 if G_2 may be obtained from G_1 via a finite sequence of (at least one) edge subdivisions, and the graph G_2 is *built from* G_1 if G_2 may be obtained from G_1 via a finite sequence of (at least one) vertex partitions. Note that edge subdivision is a special case of vertex partitioning, so that homeomorphism is a special case of "built from." The special graph on five vertices that is a single edge subdivision of $W_4 (= K_4)$ is denoted as $\hat{W}_4$:

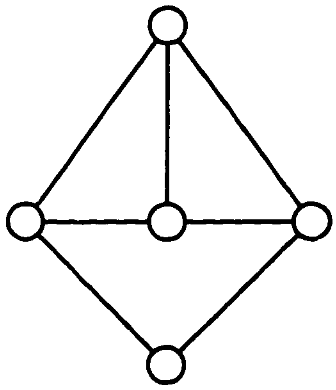


Figure 4.4: $\hat{W}_4$

Our main result may now be stated.

Theorem 4.3 *For an undirected graph G, the following four statements are equivalent:*

(0) *every distance cycle completable, partial distance matrix A, the graph of whose specified entries is G, has a distance matrix completion;*

(1) *no induced subgraph of G is W_k, $k \geq 5$, or can be built from W_k, $k \geq 4$;*

(2) *every induced subgraph of G that contains a homeomorphic image of K_4 also contains an actual copy of K_4; and*

(3) *G has chordal supergraph, in which all edges of any 4-clique are already edges of G.*

If G satisfies condition (3) of theorem 4.3, we say that G has a *3-clique chordal supergraph*. A graph G that satisfies condition (0), the motivating notion, is referred to as *distance cycle completable*. (The other three conditions of theorem 4.3 are purely structural graph theoretic conditions.) So, according to theorem 4.3, the distance cycle completable graphs are exactly those with 3-clique chordal supergraphs. Of course, chordal graphs and minimal cycles (the most *non*chordal graphs) both qualify as distance cycle completable, but many other graphs do also.

The proof of the theorem follows the logic (0)⇒(1)⇒(2)⇒(3)⇒(0). The purely graph theoretic implications (1)⇒(2) and (2)⇒(3) have already been demonstrated in [BJL], which addresses the cycle completable graphs for the positive definite completion problem. Even though these turn out to be the same graphs, we know of no immediate way to deduce the distance completion result from the result of [BJL]. So, section 4.2.2 is devoted to verifying the implications (0)⇒(1) and (3)⇒(0) in the distance case. Some purely graph theoretic technology for the latter implication is also adopted from [BJL].

Though condition (0) only requires the distance cycle conditions on minimal cycles, we shall freely use them for arbitrary cycles. This is justified by the following observation.

Lemma 4.1 *Let A be an n-by-n partial distance matrix, the graph of whose specified entries is G. The data in A satisfies that polygonal inequality for every cycle in G if and only if A is distance cycle completable.*

Proof: Necessity is clear, as "distance cycle completable" formally requires less of the data (i.e. that the polygonal inequality is satisfied by just the minimal cycles.) First we consider the case in which a non-minimal cycle has a single chord:

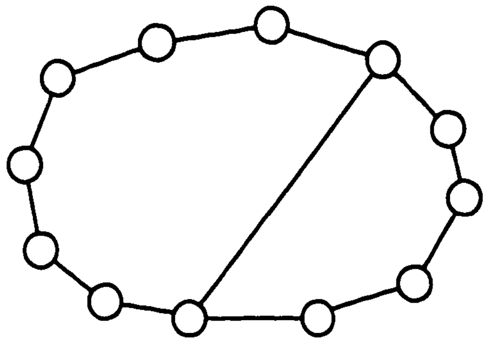


Figure 4.5: Non-minimal cycle

The non-minimal cycle is then the union of two minimal cycles, less the common edge. Suppose the length of the common edge is d and the remaining lengths in the two minimal cycles are $d_1, d_2, \ldots, d_p$ and $d_{p+1}, d_{p+2}, \ldots, d_{p+q}$. Because of the polygonal inequality for the two minimal cycles, we have:

$$\begin{aligned} d &\leq d_1 + d_2 + \cdots + d_p \\ 2d_i &\leq d + d_1 + d_2 + \cdots + d_p \qquad i = 1, 2, \ldots, p \\ d &\leq d_{p+1} + d_{p+2} + \cdots + d_{p+q} \end{aligned}$$

and

$$2d_j \leq d + d_{p+1} + d_{p+2} + \cdots + d_{p+q} \qquad j = p+1, p+2, \ldots, p+q.$$

Addition of the second and third inequalities yields

$$2d_i \leq d_1 + d_2 + \cdots + d_{p+q} \qquad i = 1, 2, \ldots, p.$$

and addition of the first and fourth yields

$$2d_j \leq d_1 + d_2 + \cdots d_{p+q} \qquad j = p+1, p+2, \ldots, p+q.$$

Taken together this is the polygonal inequality for the non-minimal cycle. When a non-minimal cycle has more than one chord, a repeated application of this argument leads to the desired result for such a non-minimal cycle. □

We close this section by noting that, based upon recent work [JM], another graph theoretic description may be added to the list in our theorem. we say that a graph G is a *clique sum* of two graphs G_1 and G_2 if each of G_1 and G_2 contain a copy of K_p for some p and identification of these two copies of K_p results in the single graph G, e.g.

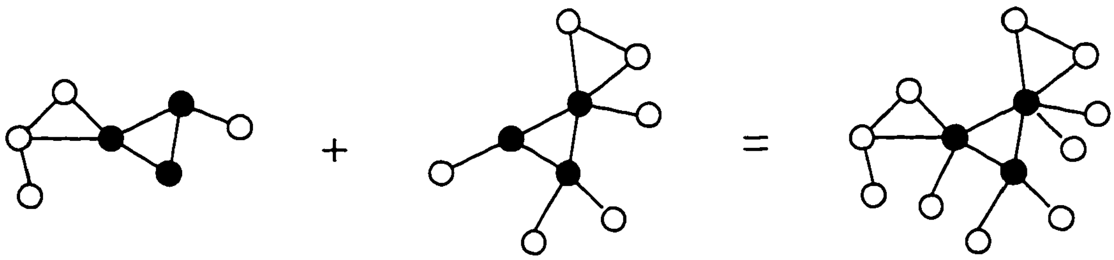


Figure 4.6: Clique sum

It is shown in [JM] that

(4) G is a sequential clique sum of chordal and series-parallel graphs (see [JM]).

is equivalent to the other graph theoretic conditions in the theorem. This is useful. as it shows that distance cycle completable graphs may be efficiently recognized.

4.2.2 Proofs of Necessary Implications

The Implication (0) $\Rightarrow$ (1)

The proof that (0)$\Rightarrow$(1) is by contrapositive. It is shown that if a graph G contains a forbidden subgraph (an induced $W_k, k \geq 5$, or an induced subgraph that is built from a $\hat{W}_k, k \geq 4$), then there is a distance cycle completable, partial distance matrix A, the graph of whose specified entries is G, that has no distance matrix completion. The proof is constructive. First, data of the desired type is exhibited for the basic forbidden subgraphs $\hat{W}_4, W_5, W_6, \ldots$ and then it is shown that if such data exists for G, it also exists for a vertex partition of G. The proof is then completed by showing how to embed such data for a forbidden subgraph in a distance cycle completable, partial distance matrix the graph of whose specified entries is otherwise arbitrary. Because of the inheritance property, the resulting data matrix can have no distance completion and thus provides that data necessary for the proof.

The *strategy* for the proof of this implication is generally the same as it was in the positive semidefinite case [BJL], but the *details* are noticeably (and probably necessarily) different.

Lemma 4.2 *None of the k-wheels W_k, $k \geq 5$, is a distance cycle completable graph.*

Proof: For $k \geq 5$, let A be the k-by-k real partial symmetric matrix given in figure 4.7 below.

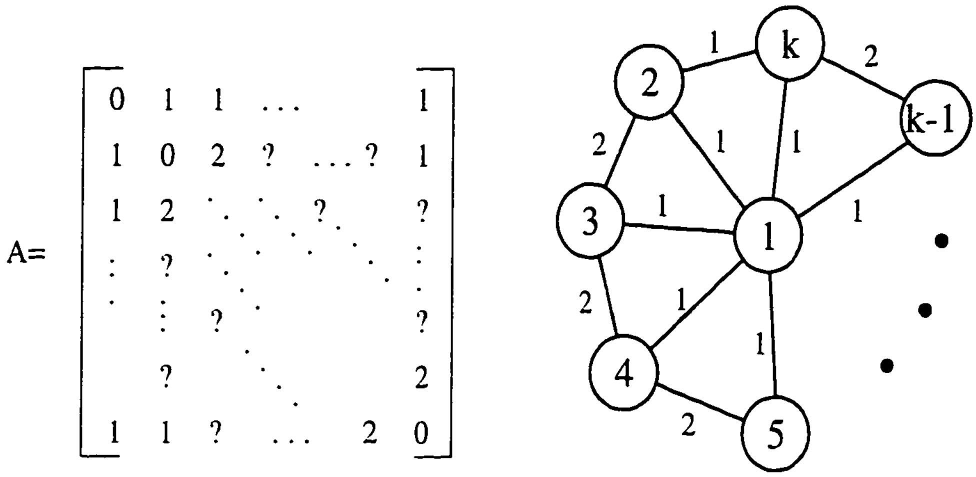


Figure 4.7: Bad data for a k-wheel

The graph of the specified entries of A is W_k and the specified entries satisfy the cycle conditions, as each entry is a 1 or 2, so that no entry is larger than the sum even of only two of the others. There are $k-1$ maximal cliques in W_k all of cardinality three: $\{1,2,3\},\{1,3,4\},\ldots,\{1,k-1,k\}$, and $\{1,2,k\}$. The principal submatrix corresponding to each of these cliques is either

$$\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 2 \\ 1 & 2 & 0 \end{bmatrix} \text{ or } \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

The former is realizable on a line, the latter as an equilateral triangle, so that both are distance matrices. Thus, A is a distance cycle completable, partial distance matrix.

Suppose the data in A were to admit a distance completion. In view of the distances among the vertices 1, 2, and 3, the vertex 1 would be the midpoint of a line segment joining vertices 2 and 3. Similarly, 1 is the midpoint of a line segment (of the same length as the segment joining 2 and 3) joining 3 and 4. It follows that 2 and 4 must coincide. The same reasoning shows that 3 and 5 must coincide, 4 and 6 (and

2), 5 and 7 (and 3), and so on. Thus, vertex k must coincide with vertex 2 or vertex 3 and, then, cannot have distance 1 (as specified) from vertex 2. This contradiction shows that A has no distance completion and completes the proof. □

The graph W_4 is distance cycle completable as it is the complete graph K_4. However, we now show that any graph built from W_k, $k \geq 4$ is not distance cycle completable. The graph $\hat{W}_4$ is built from W_4. To see that $\hat{W}_4$ is not distance cycle completable, consider the following data:

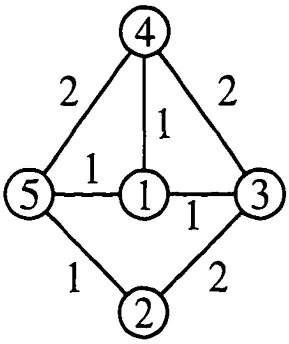


Figure 4.8: Bad data for a $\hat{W}_4$

The specified distances again trivially satisfy the cycle conditions, but vertices $\{4, 1, 5\}$ lie on a line with vertex 1 at the center. Similarly $\{4, 1, 3\}$ lie on a line with vertex 1 at the center. This means that vertex 5 is coincident with vertex 3, which is impossible: vertex 2 cannot be both distance one and distance two from the same point. Thus, $\hat{W}_4$ is not distance cycle completable. The preceding discussion yields:

Lemma 4.3 *The graph $\hat{W}_4$ is not distance cycle completable.*

If a graph can be built from W_4, then it can be built from $\hat{W}_4$. This observation, together with the preceding and following lemma suffice to prove that no graph built from W_k, $k \geq 4$, is distance cycle completable.

Lemma 4.4 *If G is a graph that is not distance cycle completable and G' is obtained from G either by subdividing an edge or by partitioning a vertex, then G' is also not distance cycle completable.*

Proof: If G' is an edge subdivision of G, then G' is also a vertex partition of G, in which one of the neighbor sets has cardinality one. Thus, only the case in which G' is a vertex partition of G need be considered. Without loss of generality, take the partitioned vertex to be n and suppose that $1, 2, \ldots, j$ are the vertices adjacent to n and $j+1, j+2, \ldots, k$ are the vertices adjacent to $n+1$ in G', $1 \le j \le k$.

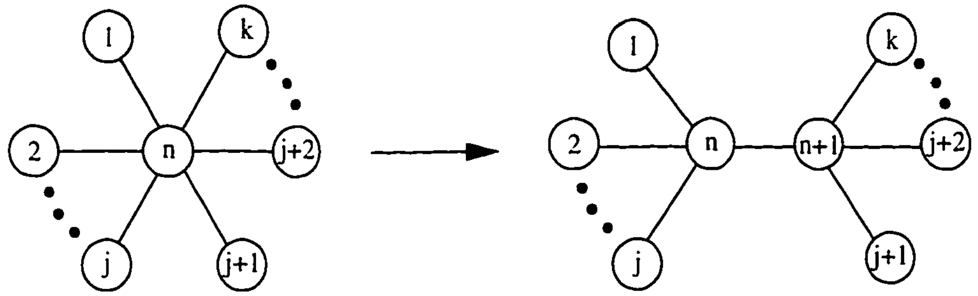


Figure 4.9: Bad data for vertex partitioning

Let $A = (a_{ij})$, the graph of whose specified entries is G, be any n-by-n partial distance matrix that satisfies the cycle conditions, but is not distance cycle completable. Define the $(n+1)$-by-$(n+1)$ partial symmetric matrix $A' = \left(a'_{ij}\right)$ as follows:

$$\begin{aligned} a'_{i,n} &= a'_{n,i} = ? \qquad i = j+1, j+2, \ldots, k \\ a'_{i,n+1} &= a'_{n+1,i} = ? \qquad i = 1, 2, \ldots, j, k+1, \ldots, n \\ a'_{i,n+1} &= a'_{n+1,i} = a_{i,n} \qquad i = j+1, j+2, \ldots, k \\ a'_{n,n+1} &= a'_{n+1,n} = a'_{n+1,n+1} = 0 \end{aligned}$$

and

$$a'_{ij} = a_{ij} \qquad \textit{otherwise.}$$

The graph of A' is G' with a distance of zero on the edge $\{n, n+1\}$.

Any minimal cycles in G' are either present in G or are cycles of G with the new edge $\{n, n+1\}$ inserted. Since the edge $\{n, n+1\}$ has a distance of zero, the cycle conditions hold in A'. The matrix A' is also a partial distance matrix, as the edge $\{n, n+1\}$ is, by construction, a maximal clique of G', and all other cliques in G' correspond to a unique clique in G.

Assume that there is a distance completion $\bar{A}'$ of A'. The vertices n and $n+1$ must coincide spatially and the triangle inequality on the entries in $\bar{A}'$ requires that

$$\begin{aligned}
\bar{a}'_{nk} &= a'_{n+1,k} \\
\bar{a}'_{n,j+1} &= a'_{n+1,j+1} \\
\bar{a}'_{n+1,1} &= a'_{n1} \\
\bar{a}'_{n+1,j} &= a'_{nj}.
\end{aligned}$$

As a result, if A' has a distance completion, then A must have a distance completion. This contradiction completes the proof. □

It follows from lemma 4.4 that if G is not distance cycle completable, then any graph G' built from G is also not distance cycle completable. To complete the proof of the implication (0)⇒(1), data for these forbidden subgraphs must be embedded in data for larger graphs in such a way that the necessary conditions still hold. The larger graphs are then seen not to be distance cycle completable, because, if they were, the smaller ones would be by inheritance.

Proof of the implication (0)⇒(1)

Suppose that G' is a connected graph on n vertices that contains, as an induced subgraph, a graph G that is one of the forbidden subgraphs. If G' were not connected we would need only consider the connected component containing G. For convenience, assume that the vertices of G are $1, 2, \ldots, k$. We assume, without loss of generality, that $k < n$, so that the remaining vertices of G' are $k+1, k+2, \ldots, n$. As in the proofs of lemmas 4.1, 4.2, and 4.3, assign distances to the edges of G. We need to demonstrate data for the remaining edges of G' that does not violate the necessary conditions.

Suppose that the subgraph of G' induced by vertices $k+1, k+2, \ldots, n$ has $p(\geq 1)$ connected components. (Since G' is connected, each of these components has an edge between one of its vertices and a vertex of G.) Within each of these components assign the distance 0 to each edge. For each component, this has the effect of assigning all vertices of that component the same spatial location. (Note that an edge of length 0 does not affect satisfaction of the cycle conditions.) We assume then that $k+p=n$ and treat each component as a single vertex, implicitly consolidating edges as needed. We assign distances to the edges connecting vertices $k+1, k+2, \ldots, k+p$ with the vertices of G as follows. Since G is W_k or built from W_k, G has a natural "center" vertex or cluster of vertices resulting from partitions of the center. In the latter case, all specified distances among vertices in the central cluster are 0; the cluster is connected, and, so, all vertices in the cluster must occupy the same spatial location. We thus identify all vertices in the central cluster and, in both cases, refer to "the" center vertex, which, we suppose, is vertex 1.

Consider now a vertex $k+i, k=1,2,\ldots,p$. If $k+i$ is adjacent to 1 (case one), we assign that edge the distance 0. All other edges connecting vertices in G to $k+i$ are then assigned distances by calculating the shortest path in G from that vertex (i.e. $k+i$) to vertex 1. This insures that the polygonal inequality holds for any cycle containing $k+i$. Any clique in G' that includes vertex $k+i$ has cardinality at most four. If the cardinality is less than four, there is nothing further to check. If it is four, then the clique must include vertex 1, as well as two vertices j and $j+1$ adjacent in the cycle of G. For the vertices j and $j+1$, the distances assigned to the edges $\{j, k+i\}$ and $\{j+1, k+i\}$ must be d_{1j} and $d_{1,j+1}$, the distances from 1 to j and to

$j+1$, respectively The data specified by our prescription for these four points is then

$$\begin{bmatrix} 0 & 0 & d_{1j} & d_{1,j+1} \\ 0 & 0 & d_{1j} & d_{1,j+1} \\ d_{1j} & d_{1j} & 0 & d_{j,j+1} \\ d_{1,j+1} & d_{1,j+1} & d_{j,j+1} & 0 \end{bmatrix}.$$

which is the distance matrix of the triangle of G:

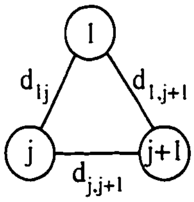


Figure 4.10: Data for a triangle

with vertex 1 repeated.

If $k+i$ is not adjacent to vertex 1 (case 2), choose an arbitrary one of its neighbors in G and assign that edge the distance 0. Again, assign distances to all other edges connecting vertices of G to $k+i$ via shortest path calculations. This ensures the polygonal inequalities for any cycle containing $k+i$. In this case, no clique containing $k+i$ can be of cardinality more than three. Because of the triangle/polygonal inequalities, the data associated with such cliques are distance matrices.

By assigning edge distances throughout G' we now have a partial symmetric matrix A, the graph of whose specified entries is G'. Because we have checked the data associated with cliques of G', A is a partial distance matrix. Furthermore, the polygonal inequalities associated with the cycles of G' are satisfied, so that A is distance cycle completable. (Note that, if there is a minimal cycle in G' containing more than one of the $k+i$, it may also be identified with a cycle of G by our prescription.) If A

had a distance matrix completion $\tilde{A}$, then $\tilde{A}[\{1,2,\ldots,k\}]$ would be a distance matrix completion of $A[\{1,2,\ldots,k\}]$, a contradiction that completes the proof. □

The Implication (3)⇒(0)

To prove (3)⇒(0), we must show that if G is a graph that has a 3-clique chordal supergraph, then every distance cycle completable, partial distance matrix, whose graph is G, has a distance completion. As in the positive definite case, much of the proof is combinatorial in nature. The logic of the proof is, as in [BJL], to sequentially add edges $e_1, e_2, \ldots, e_m$ to a given graph G to obtain a 3-clique chordal supergraph H. Upon addition of the edge e_k, all new minimal cycles are considered. For any two of these minimal cycles, a common distance can be chosen for e_k so as to make each cycle completable. Then, an application of Helly's theorem (see, e.g. [Ro]) produces a common distance for all new cycles.

Because two minimal cycles may intertwine in a variety of ways, difficulties arise in considering a general pair of minimal cycles containing a given edge, e_k. In [BJL] there are a number of graph theoretic lemmas dealing with these difficulties for graphs with 3-clique chordal supergraphs. These lemmas culminate in lemma 9.6 in [BJL], which is stated below as lemma 4.5.

Before this lemma can be stated, another definition is needed. Given a graph G, an edge $\{a,b\}$ of G and a cycle C, we say that G' is obtained from G by *replacing the edge* $\{a,b\}$ *with the cycle* C if, after two of the vertices of C are identified with a and b the graph G' is $(G - \{a,b\}) \cup C$. Pictorially,

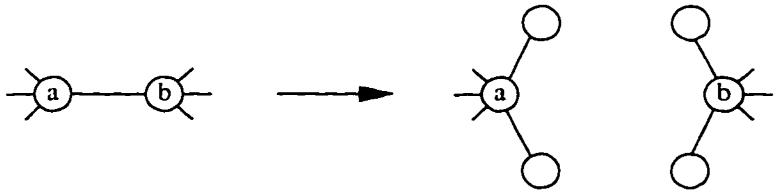


Figure 4.11: Replacing edge a, b with the cycle C

A *necklace* is either a graph that is a cycle plus one additional chord or a graph obtained from a cycle $C_k = [c_0, c_1, \ldots, c_k, c_0]$ by specifying a subset of the edges $\{c_0, c_1\}, \{c_1, c_2\}, \ldots, \{c_{k-1}, c_k\}$ and replacing each edge in the subset with a cycle. Note that the edge $\{c_k, c_0\}$ is excluded. The *base* of a necklace that is a cycle plus an additional chord is the chord; otherwise, it is the edge $\{c_k, c_0\}$. Recall that if $G = (V, E)$ and $W \subseteq V$, then G_W denotes the subgraph of G induced by the vertex set W.

Lemma 4.5 *Let $G = (N, E)$ be a graph that has a 3-clique chordal supergraph H, let $e = \{a, b\}$ be an edge of H that is not in G and let $G' = (N, E \cup \{e\})$. Let $C = (U, E')$ and $D = (V, E'')$ be distinct minimal cycles in G' with common edge e, and express C and D as the union of paths*

$$C = [a_s, a_0] \cup [a_0, a_1] \cup [a_1, a_2] \cup \cdots \cup [a_{s-1}, a_s]$$
$$D = [b_s, b_0] \cup [b_0, b_1] \cup [b_1, b_2] \cup \cdots \cup [b_{s-1}, b_s]$$

in which $a_0 = b_0 = a$, $a_s = b_s = b$, $U \cap V = \{a_0, a_1, \ldots, a_s\} = \{b_0, b_1, \ldots, b_s\}$ and s is a positive integer. Then $G'_{U \cup V}$ is a necklace with base $\{a, b\}$.

One more lemma is required for the proof of (3)⇒(0). A corresponding statement for the positive definite case, lemma 9.7, is found in [BJL]. Since the non-graph theoretic parts of the proof in [BJL] require only a *chordal theorem* and assumed *cycle conditions*, the proof in the distance case is similar.

Lemma 4.6 *Let $G = (N, E)$ be a graph that has a 3-clique chordal supergraph $H = (N, F)$, and let A be a distance cycle completable, partial distance matrix whose graph is G. Let $e = \{r, s\}$ be an edge in F that is not in E and let $G' = (N, E \cup \{e\})$. For any real number x, let A_x be the matrix obtained by specifying the r, s and s, r entries of A to be x. Then, if $C = (U, E')$ and $D = (V, E'')$ are distinct minimal cycles in G' with common edge e, there exists a nonnegative number x such that both $A_x[U]$ and $A_x[V]$ have distance completions.*

Proof: By lemma 4.5, $G'_{U \cup V}$ is a necklace with base $\{r, s\}$. Since $G_{U \cup V}$ is $G'_{U \cup V}$ with the edge $\{r, s\}$ deleted, it is the union of graphs $G_i = (W_i, E'_i)$, $i = 1, 2, \ldots, k$, in which each G_i is either K_2 or a minimal cycle of G. If G_i is a minimal cycle, let $\hat{G}_i$ be the complete graph on W_i and if G_i is K_2, let $\hat{G}_i = G_i$. Let S be the union of the $\hat{G}_i, i = 1, 2, \ldots, k$. Then S is a chordal supergraph of $G_{U \cup V}$ as any cycle in S is a cycle in one of the $\hat{G}_i$. Now if G_i is a minimal cycle, $A[W_i]$ has a distance completion $B[W_i]$ by assumption. Let $B[U \cup V]$ be the matrix obtained from $A[U \cup V]$ by replacing each $A[W_i]$ for which G_i is a minimal cycle by $B[W_i]$. Then the graph of $B[U \cup V]$ is S and for each maximal clique W of S, $B[W]$ is a distance matrix. The graph S is chordal, so, by the chordal theorem [BJ], $B[U \cup V]$ has a distance completion $M[U \cup V]$, which is also a distance completion of $A[U \cup V]$. Let x be the r, s entry of $M[U \cup V]$. Then $A_x[U]$ and $A_x[V]$ have the distance completion $M[U]$ and $M[V]$, respectively, which completes the proof. □

We may now complete the proof of the implication (3)⇒(0). Let $e_1, e_2, \ldots, e_m$ be the edges in a 3-clique chordal supergraph H that are not in G, and let $G_i = (N, E \cup \{e_1, e_2, \ldots, e_i\}), i = 0, 1, \ldots, m$. Then $G_0 = G$ and $G_m = H$. We assume, without loss of generality, that G is connected. Let A be a distance cycle completable partial distance matrix, whose graph is G. We wish to exhibit a distance matrix completion of A.

Consider first G_1. Write $e_1 = \{r, s\}$ and for any real x, let A_x be the matrix obtained by specifying the r, s and s, r entries of A to be x. There are two possibilities regarding the minimal cycles in G_1.

(i) The edge e_1 is contained in exactly one minimal cycle $C = (W, F)$ of G_1. Since the induced subgraph $(G_1)_W$ is a cycle, G_W is a path, and thus a chordal graph. By the chordal theorem [BJ], $A[W]$ has a distance matrix completion. Let x be the r, s entry of this completion. Then $A_x[W]$ has the same completion. Either the edge $\{r, s\}$ or W (if $|W| = 3$) is the only maximal clique of G_1 that is not a maximal clique of G, so A_x is a partial distance matrix. Since C is the only minimal cycle in G_1 that is not a minimal cycle in G, A_x is also distance cycle completable.

(ii) The edge e_1 is contained in two or more minimal cycles of G_1. Let $C_i = (W_i, F_i), i = 1, 2, \ldots, p$ be the minimal cycles in G_1 containing the edge e_1. For $i = 1, 2, \ldots, p$, let $I_i = \{x \in R_+ : A_x[W_i]$ has a distance matrix completion$\}$. If $|W_i| = 3$, we interpret this to mean $A_x[W_i]$ is a distance matrix. By lemma 4.6, $I_i \cap I_j \neq \emptyset$ for all distinct $i, j \in \{1, 2, \ldots, p\}$. Therefore, by Helly's theorem on this line, $\cap_{i=1}^{p} I_i \neq \emptyset$. Let $x \in \cap_{i=1}^{p} I_i$. Then $A_x[W_i]$ has a distance matrix completion for $i = 1, 2, \ldots, p$. By hypothesis, e_1 belongs to no 4-clique in G_1. Thus, the only maximal cliques in G_1 that are not in G are those W_i with $|W_i| = 3$, or only $\{r, s\}$ if $|W_i| \geq 4$ for all $i = 1, 2, \ldots, p$. It follows that A_x is a partial distance matrix and distance cycle completable.

If $m = 1$, we are done. Otherwise, G_1 has a 3-clique chordal supergraph H and A_x is a distance cycle completable, partial distance matrix, the graph of whose specified entries is G_1. Therefore, repeating the above process for $e_2, e_3, \ldots, e_m$ we obtain a partial distance matrix B with chordal graph H. By the chordal theorem [BJ], B

has a distance matrix completion M, which is also a distance completion of A. This completes the proof of the implication (3)⇒(0), which completes the proof of the theorem.

4.3 The Combinatorially Symmetric P-matrix Completion Problem

Positive definite matrices are Hermitian matrices for which all principal minors are positive. If we relax the symmetry required by Hermicity, we get the P-matrices. An n-by-n real matrix is called a P-matrix (P_0-matrix) if all its principal minors are positive (nonnegative) (see, e.g. [BP] or [HJ2]). This class of matrices generalizes many other important classes of matrices (such as M-matrices and totally positive matrices), has useful structure (such as inverse closure, inheritance by principal submatrices, and wedge type eigenvalue restrictions), and arises in applications (such as the linear complementarity problem, and issues of local invertibility of functions).

Here, we consider the *P-matrix completion problem* under the assumptions that the partial matrix is square, all diagonal entries are specified, and the data is *combinatorially symmetric* (the j,i entry is specified if and only if the i,j entry is specified). Further, since the property of being a P-matrix is inherited by principal submatrices, it is necessary that the partial matrix be a *partial P-matrix*, i.e. every fully specified principal submatrix must itself be a P-matrix. Of all these assumptions, the only one that is truly restrictive is combinatorial symmetry and the general case, in which this assumption is relaxed, is commented upon later.

In each of the completion problems: positive definite, M-matrix, inverse M-matrix, and totally positive, there are significant combinatorial restrictions (in addition to the necessity of inheritance) on partial matrices, even when combinatorially symmetry is assumed, in order to ensure a desired completion. For example, as mentioned before, the condition on partial positive definite matrices necessary to ensure a positive definite completion (without further knowledge of the data) is that the undirected graph of the symmetric data be chordal [GJSW]. P-matrices are a generalization of real positive semidefinite matrices in which the matrix is no longer required to be symmetric. Interestingly, it is shown here that, in the case of P-matrix completions, there are no combinatorial restrictions necessary to ensure a P-matrix completion other than the combinatorial symmetry assumption. Every combinatorially symmetric partial P-matrix has a P-matrix completion. However, when the combinatorial symmetry assumption is relaxed, the conclusion no longer holds, and the question of which directed graphs for the specified entries ensure that a partial P-matrix has a P-matrix completion is, in general, open. All 3-by-3 partial P-matrices have P-matrix completions, but we exhibit a 4-by-4 partial P-matrix with just one unspecified entry and no P-matrix completion.

Let A be an n-by-n partial P-matrix with one pair of symmetrically placed unspecified entries. By permutation similarity it can be assumed without loss of generality that the unspecified entries are a_{1n} and a_{n1}. Then, A is of the form:

$$A = \begin{bmatrix} a_{11} & a_{12}^T & ? \\ a_{21} & A_{22} & a_{23} \\ ? & a_{32}^T & a_{33} \end{bmatrix}$$

in which A_{22} is $(n-2)$-by-$(n-2)$ and $a_{12}, a_{21}, a_{23}, a_{32} \in \mathbb{R}^{n-2}$. Define

$$A(x,y) \equiv \begin{bmatrix} a_{11} & a_{12}^T & x \\ a_{21} & A_{22} & a_{23} \\ y & a_{32}^T & a_{33} \end{bmatrix}$$

and denote $A(0,0)$ by A_0. Also define $C \equiv \{\alpha \subseteq N : 1, n \in \alpha\}$ and let $A_\alpha = A(x,y)[\alpha]$. Note that since $1, n \in \alpha$ for all $\alpha \in C$, x and y are unspecified entries in every A_α, $\alpha \in C$.

Lemma 4.7 *Every partial P-matrix with one pair of symmetrically placed unspecified entries has a P-matrix completion.*

Proof: Let A be an n-by-n partial P-matrix with exactly one pair of symmetrically placed unspecified entries. To find a P-matrix completion of A we must find x, y such that $\det A_\alpha > 0$ for all $\alpha \in C$ (the remaining principal minors of A are positive by hypothesis). For each $\alpha \in C$, $\alpha = \left\{i_1 = 1, i_2, \ldots, i_{|\alpha|} = n\right\}$ define

$$\begin{aligned} a_\alpha &= \det A(x,y)\left[\left\{i_2, i_3, \ldots, i_{|\alpha|-1}\right\}\right] \\ b_\alpha &= \det A_0\left[\left\{i_2, i_3, \ldots, i_{|\alpha|}\right\}; \left\{i_1, i_2, \ldots, i_{|\alpha|-1}\right\}\right] \\ c_\alpha &= \det A_0\left[\left\{i_1, i_2, \ldots, i_{|\alpha|-1}\right\}; \left\{i_2, i_3, \ldots, i_{|\alpha|}\right\}\right] \\ d_\alpha &= \det A_0[\alpha]. \end{aligned}$$

By using Sylvester's identity (see, e.g. section 0.8.6 [HJ1]) we see that

$$\det A_\alpha = -a_\alpha xy + (-1)^{|\alpha|-1} b_\alpha x + (-1)^{|\alpha|-1} c_\alpha y + d_\alpha.$$

Since $A(x,y)$ is a partial P-matrix, $a_\alpha > 0$ for all $\alpha \in C$. Then x, y can be chosen so that $xy < 0$ and $-a_\alpha xy > (-1)^{|\alpha|} b_\alpha x + (-1)^{|\alpha|} c_\alpha y - d_\alpha$. In order to have $xy < 0$, choose $y = -x$. The line $y = -x$ intersects the hyperbola $\det A_\alpha = 0$ at the points

$$x_\alpha^{\pm} = \frac{-(b_\alpha - c_\alpha) \pm \sqrt{(b_\alpha - c_\alpha)^2 - 4a_\alpha d_\alpha}}{2a_\alpha}.$$

Define

$$m(\alpha) = \max\left\{\left|x_\alpha^{+}\right|, \left|x_\alpha^{-}\right|\right\}.$$

Then $\det A_\alpha > 0$ for each x such that $|x| > m(\alpha)$. In order to find a P-matrix completion of A we must find a pair x, y that works for all $\alpha \in C$. The matrix $A(x,y)$ is a P-matrix for all x, y such that $|x| > \max_{\alpha \in C} m(\alpha)$ and $y = -x$. □

This lemma can be used sequentially to find a completion of any combinatorially symmetric partial P-matrix. The lemma proves the case in which there is one pair of symmetrically placed unspecified entries. Assume there is a P-matrix completion of every partial P-matrix with $k-1$ pairs of symmetrically placed unspecified entries and let A be a partial P-matrix with k pairs of symmetrically placed unspecified entries. Choose one symmetrically placed pair i, j of unspecified entries of A. Each maximal principal submatrix that this pair completes (there are no other unspecified entries in such a maximal submatrix) is a partial P-matrix by inheritance and, by lemma 4.7, can be completed to a P-matrix. For each maximal principal submatrix $A[\alpha]$ let x_α be the value of the unspecified entry as given by lemma 4.7. Then choosing x so that $|x| > \max\{|x_\alpha| : i, j \in \alpha, \alpha \text{ maximal}\}$ completes each of these maximal principal submatrices. Then, we are left with a partial P-matrix with $k-1$ pairs of symmetrically placed unspecified entries which can be completed to a P-matrix by the induction hypothesis. Note that the order of completion is immaterial (as long as combinatorial symmetry is maintained). This proves our main result.

Theorem 4.4 *Every combinatorially symmetric partial P-matrix has a P-matrix completion.*

As mentioned above, when the combinatorial symmetry assumption is relaxed, the conclusion of the theorem no longer holds. The question of which directed graphs for the specified entries ensures that a partial P-matrix has a P-matrix completion is, in general, open. However, we do know the following.

Proposition 4.1 *Every 3-by-3 partial P-matrix has a P-matrix completion.*

Proof: The combinatorially symmetric case is covered by the lemma above. The only case that remains to be considered is the case in which A is a 3-by-3 partial P-matrix with one unspecified entry. Note that if there are more unspecified entries, values may be assigned to entries making sure the 2-by-2 principal minors are positive, until either one pair of symmetrically placed unspecified entries, or only one unspecified entry remains.

By permutation similarity, it can be assumed without loss of generality that the unspecified entry is in the 3,1 position. It can also be assumed, by positive left diagonal multiplication and diagonal similarity (which both preserve P-matrices), that there are ones on the main diagonal and on the super diagonal. Then A is of the form

$$A = \begin{bmatrix} 1 & 1 & c \\ a & 1 & 1 \\ y & b & 1 \end{bmatrix}$$

in which $a, b, c < 1$ since A is a partial P-matrix. In order to complete A to a P-matrix y must be chosen so that the 1,3 minor is positive (which yields $yc < 1$) and $\det A = 1 + abc - a - b + y(1 - c)$ is positive.

There are several cases to consider. If $c \leq 0$, choose $y > 0$ and large enough to make $\det A$ positive. Similarly, if $c > 1$, choose $y < 0$ and large in absolute value. If $c = 1$, then $\det A = 1 + ab - a - b = (1-a)(1-b)$ which is positive since $a, b < 1$. So, $y = 0$ will give a P-matrix completion of A. All that remains is $0 < c < 1$. In this case, if $a, b \geq 0$ or if $ab \leq 0$ choose $y = ab$. This gives $yc = abc < 1$ (since $a, b, c < 1$) and

$$\begin{aligned} \det A &= 1 + abc - a - b + ab(1-c) \\ &= 1 - a - b + ab \\ &= (1-a)(1-b) > 0. \end{aligned}$$

For $a, b < 0$ the term $1 + abc - a - b$ in the determinant of A is positive. So, $y = 0$ will result in a P-matrix completion of A. Thus, every 3-by-3 partial P-matrix has a P-matrix completion. □

Proposition 4.2 *For every $n \geq 4$, there is a partial P-matrix with exactly one unspecified entry for which there is no P-matrix completion.*

Proof: The matrix

$$A = \begin{bmatrix} 1 & -1 & 1 & 1 \\ 2 & 1 & -1 & 1 \\ 0 & 1 & 1 & 2 \\ y & -10 & -1 & 1 \end{bmatrix}$$

is a partial P-matrix with no P-matrix completion. It is easy to check that all 2-by-2 and 3-by-3 principal minors that do not include both rows 1 and 4 are positive. However, the two 3-by-3 determinants that involve y cannot simultaneously be made positive. In order for the determinant of $A[\{1, 2, 4\}]$ to be positive it must be the case that $y < -\frac{7}{2}$ while a positive determinant for $A[\{1,3,4\}]$ requires that $y > -3$. Thus, there is no P-matrix completion of A. This data can be embedded as a principal submatrix, by putting 1's on the diagonal and 0's in the other specified positions, to produce a partial P-matrix with one unspecified entry and no P-matrix completion for any $n \geq 4$. □

Other completion problems, intermediate between the positive definite completion problem and the combinatorially symmetric P-matrix completion problem are open and may be of interest. For example, when does a sign symmetric, partial P-matrix have a sign symmetric P-matrix completion and what may be said about the infimum of the Frobenius norms of completions of combinatorially symmetric partial P-matrices.

4.4 The Totally Positive Completion Problem

An n-by-n matrix is said to be *totally positive* (nonnegative) if every minor (principal and non-principal) is positive (nonnegative). In particular, this means that every entry of a totally positive (nonnegative) matrix is positive (nonnegative). Further discussion of totally positive matrices may be found in [K] or [A]. This class of matrices arises in many applications including approximation theory, geometric design, and wavelets. It is interesting to note that according to [GM] one can test whether or not a matrix is strictly totally positive in polynomial time. This uses the fact that only the determinants of submatrices consisting of consecutive rows and columns need to be checked [K].

Here we consider the question: for which graphs G does every partial totally nonnegative matrix, the graph of whose specified entries is G, have a totally nonnegative completion. Total nonnegativity is inherited by submatrices. Therefore, it is a necessary condition that every fully specified submatrix be totally nonnegative.

Total nonnegativity is not, however, preserved by permutation similarity. This can easily be seen by noting that the 1,1 entry of a matrix enters positively in every minor in which it occurs. Therefore, total positivity is preserved when this entry is arbitrarily increased. However, if the first and second rows and columns are interchanged, this entry is then in the 2,2 position and enters negatively in some minors; making it larger may then make some of these minors negative.

Because total positivity is not preserved by permutation similarity we will restrict our attention to *labeled* graphs. Labeled graphs are those in which the numbering of the vertices is fixed. Note that two partial totally positive matrices the graphs of whose specified entries are isomorphic may not both have totally positive completions. So, the labeling of the graphs is important.

As mentioned in the introduction a clique is an induced subgraph that is a complete graph. A *block clique graph* is a chordal graph in which every pair of maximal cliques, C_i, C_j, $C_i \neq C_j$ intersect in at most one vertex. That is $|C_i \cap C_j| \leq 1$. A *monotonically labeled block clique graph* is a labeled block clique graph in which for every pair of intersecting cliques such that $C_i \cap C_j = u$, the labeling in the two cliques is such that $\{v : v \in C_i - u\} < u$ and $\{w : w \in C_j - u\} > u$. Then, a monotonically labeled block clique graph is of the form

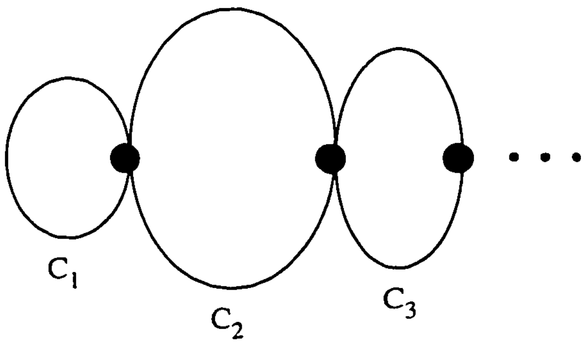


Figure 4.12: Monotonically labeled block clique graph

and a matrix, the graph of whose specified entries is a monotonically labeled block clique graph is of the form

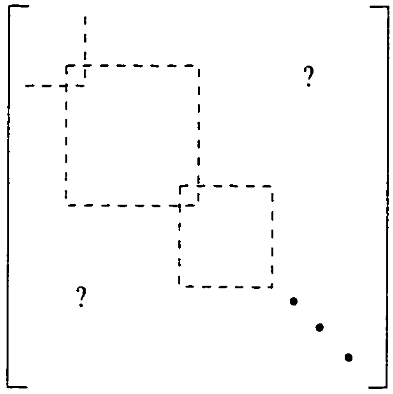


Figure 4.13: Partial matrix with a monotonically labeled block clique graph

In the remaining discussion all graphs will be assumed to be connected. This assumption may be made without loss of generality since totally positive matrices are closed under direct sums. To prove this we will use the classical fact of Frobenius-König (see [Ry] or [S] and its references) that if an n-by-n matrix contains a zero block of size p-by-q in which $p+q \geq n+1$, then the matrix is singular [Ry]. Although we use only this zero block result we prove the more general the result below. Note that this result is a consequence of theorem 2.2 in [CJRW] which may also be viewed as a generalization of the classical Frobenius-König result.

Lemma 4.8 *Suppose that* $A \in M_n(F)$ *has a* p*-by-*q *submatrix of rank* r*. Then* A *is singular whenever* $p+q \geq n+r+1$.

Proof: Without loss of generality, independently permute the rows and columns of A so that the p-by-q submatrix of rank r is in the lower left corner of A. That is

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

in which A_{21} is p-by-q and $rank(A_{21}) = r$. Using row and column operations A_{21} can be reduced so that A becomes

$$\tilde{A} = \begin{bmatrix} B_1 & B_2 & \vdots & \tilde{A}_{21} \\ \cdots & \cdots & \vdots & \cdots \\ 0 & C & \vdots & D_1 \\ 0 & 0 & \vdots & D_2 \end{bmatrix}$$

in which C is r-by-r and nonsingular. Then $\tilde{A}$ is singular if either $\begin{bmatrix} B_1 \\ 0 \\ 0 \end{bmatrix}$ has linearly dependent columns or $\begin{bmatrix} 0 & 0 & D_2 \end{bmatrix}$ has linearly dependent rows. Since B_1 is $(n-p)$-by-$(q-r)$, $\begin{bmatrix} B_1 \\ 0 \\ 0 \end{bmatrix}$ is guaranteed to have linearly dependent columns if $q - r > n - p$. Similarly, since D_2 is $(p-r)$-by-$(n-q)$, D_2 has linearly dependent rows if $p - r > n - q$. Rearranging both of these inequalities we see that $\tilde{A}$ is singular if $p + q > n + r$, that is, if $p + q \geq n + r + 1$. Since $\tilde{A}$ is singular if and only if A is singular, this completes the proof. □

Note that if $p + q < n + r + 1$, it is easy to construct matrices with a p-by-q block of rank r that are nonsingular. So, this lemma is best possible. Lemma 4.8 may be used to prove that the direct sum of totally positive matrices is totally positive.

Lemma 4.9 *The matrices* $A \in M_n(\mathbb{R})$ *and* $B \in M_m(\mathbb{R})$ *are totally nonnegative if and only if*

$$C = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$$

is totally nonnegative.

Proof: If C is a totally nonnegative matrix, then it is obvious that A and B are totally nonnegative since total nonnegativity is inherited by submatrices. To prove the forward implication, assume that A and B are totally nonnegative and we will show that C is also totally nonnegative. Let $\alpha_1, \beta_1 \subseteq N$ and $\alpha_2, \beta_2 \subseteq M$ and define $\tilde{\alpha}_2 = \alpha_2 + \{n\}$, $\tilde{\beta}_2 = \beta_2 + \{n\}$. If $\alpha_1 = \beta_1 = \emptyset$ or $\alpha_2 = \beta_2 = \emptyset$, then $C\left[\alpha_1 \cup \tilde{\alpha}_2; \beta_1 \cup \tilde{\beta}_2\right]$ is a submatrix of B or of A, respectively, and therefore, since A and B are totally nonnegative, has nonnegative determinant. On the other hand, if $\alpha_2 = \beta_1 = \emptyset$ or $\alpha_1 = \beta_2 = \emptyset$, then $C\left[\alpha_1 \cup \tilde{\alpha}_2; \beta_1 \cup \tilde{\beta}_2\right]$ is a zero matrix, and thus has determinant 0. If exactly one of $\alpha_1, \alpha_2, \beta_1$, or β_2 is the empty set, then $C\left[\alpha_1 \cup \tilde{\alpha}_2; \beta_1 \cup \tilde{\beta}_2\right]$ has at least one zero row or column and, therefore, has determinant 0.

So, assume $\alpha_1, \alpha_2, \beta_1, \beta_2 \neq \emptyset$. In this case, if $|\alpha_1| = |\beta_1|$, then, since $|\alpha_1| + |\alpha_2| = |\beta_1| + |\beta_2|$, it must be the case that $|\alpha_2| = |\beta_2|$. Then, $C\left[\alpha_1 \cup \tilde{\alpha}_2; \beta_1 \cup \tilde{\beta}_2\right]$ is a block diagonal matrix in which the diagonal blocks are submatrices of A and B and, thus,

$$\det\left(C\left[\alpha_1 \cup \tilde{\alpha}_2; \beta_1 \cup \tilde{\beta}_2\right]\right) = \det\left(A\left[\alpha_1; \beta_1\right]\right) \det\left(B\left[\alpha_2; \beta_2\right]\right) \geq 0.$$

The last case to check is that in which $\alpha_1, \alpha_2, \beta_1, \beta_2 \neq \emptyset$ and $|\alpha_1| \neq |\beta_1|$ (and $|\alpha_2| \neq |\beta_2|$.) Assume $|\alpha_1| < |\beta_1|$ (the case in which $|\alpha_1| > |\beta_1|$ follows by symmetry). Note that $C\left[\alpha_1 \cup \tilde{\alpha}_2; \beta_1 \cup \tilde{\beta}_2\right]$ contains two blocks of zeros; one of size $|\alpha_1|$-by-$|\beta_2|$ and one of size $|\alpha_2|$-by-$|\beta_1|$. The sum of the dimensions of the zero block in the lower left corner is $|\alpha_2| + |\beta_1|$. Since $|\alpha_1| + 1 \leq |\beta_1|$, the sum of the dimensions satisfies the inequality $|\alpha_2| + |\beta_1| \geq |\alpha_2| + |\alpha_1| + 1$ and, therefore, by lemma 4.8, $C\left[\alpha_1 \cup \tilde{\alpha}_2; \beta_1 \cup \tilde{\beta}_2\right]$ is singular. Thus, $\det\left(C\left[\alpha_1 \cup \tilde{\alpha}_2; \beta_1 \cup \tilde{\beta}_2\right]\right) \geq 0$ for all $\alpha_1, \beta_1 \subseteq N$, $\alpha_2, \beta_2 \subseteq M$ and C is totally nonnegative. □

By the adjoint formula for the inverse of a matrix we see that the inverse of a totally positive matrix has a checkerboard sign pattern. That is, the sign pattern of the inverse is

$$\begin{bmatrix} + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

Therefore, the inverse of a totally positive matrix cannot be totally positive. However, as pointed out in [M, theorem 2.2], the inverse of a totally positive matrix is similar to a totally positive matrix via a signature similarity. The proof of this result is included here for completeness.

Lemma 4.10 *Let* $D = \begin{bmatrix} 1 & & & 0 \\ & -1 & & \\ & & \ddots & \\ 0 & & & (-1)^{n+1} \end{bmatrix}$. *Then a nonsingular matrix* $A \in M_n(\mathbb{R})$ *is totally nonnegative if and only if* $DA^{-1}D^{-1}$ *is totally nonnegative.*

Proof: Let A be a totally nonnegative nonsingular matrix. Since D is diagonal, partitioned matrix multiplication shows that

$$\left(DA^{-1}D^{-1}\right)[\alpha;\beta] = D[\alpha]\,A^{-1}[\alpha;\beta]\,D^{-1}[\beta].$$

By Jacobi's identity for $\alpha, \beta \subseteq N$ such that $|\alpha| = |\beta|$

$$\det\left(A^{-1}[\alpha;\beta]\right) = (-1)^{s(\alpha)+s(\beta)} \frac{\det(A[\beta^c;\alpha^c])}{\det(A)}$$

in which $s(\alpha) = \sum_{j\in\alpha} j$. Then,

$$\begin{aligned} \det\left(\left(DA^{-1}D^{-1}\right)[\alpha;\beta]\right) &= \det(D[\alpha])\det\left(A^{-1}[\alpha;\beta]\right)\det(D[\beta]) \\ &= (-1)^{s(\alpha)+|\alpha|}(-1)^{s(\alpha)+s(\beta)}\frac{\det(A[\beta^c;\alpha^c])}{\det(A)}(-1)^{s(\beta)+|\beta|} \\ &= (-1)^{2(s(\alpha)+s(\beta)+|\alpha|)}\frac{\det(A[\beta^c;\alpha^c])}{\det(A)}. \end{aligned} \tag{4.2}$$

The last equality holds since $|\alpha| = |\beta|$. If A is totally nonnegative, then $\det A[\beta^c;\alpha^c]$ is nonnegative and, since A is also nonsingular, $\det(A)$ is positive. Therefore, $\det((DA^{-1}D^{-1})[\alpha;\beta])$ is nonnegative for all $\alpha, \beta \subseteq N$.

Now assume that $DA^{-1}D^{-1}$ is totally nonnegative and invertible. Then, by the above discussion, $D\left(DA^{-1}D^{-1}\right)^{-1}D^{-1} = A$ is totally nonnegative which proves the lemma. □

We say a partial totally nonnegative matrix is *regular* if every maximal specified principal submatrix is nonsingular and for every pair of maximal specified principal submatrices $A[\alpha], A[\beta]$ such that $\alpha \cap \beta \neq \emptyset$, then $A[\alpha \cap \beta]$ is also nonsingular. The main result of this section is then:

Theorem 4.5 *Let G be a labeled graph on n vertices. Every regular partial totally nonnegative matrix, the labeled graph of whose specified entries is G has a totally nonnegative completion if and only if G is a monotonically labeled block clique graph.*

The forward implication of theorem 4.5 will be proven by contrapositive. This will be done by first "ruling out" non-chordal graphs. That is, it will be shown that there exists partial totally nonnegative matrices the graph of whose specified entries is not chordal for which there is no totally positive completion. After ruling out non-chordal graphs, those chordal graphs that are not block clique graphs will be ruled out. Finally, it will be shown that a block clique graph that is not monotonically labeled is also ruled out. The reverse implication will be proven by showing that every regular partial totally nonnegative matrix the graph of whose specified entries is a monotonically labeled block clique graph has a totally nonnegative completion. This will be done by exhibiting such a completion.

Lemma 4.11 *Let G be a graph on n vertices. If every partial totally nonnegative matrix A, the graph of whose specified entries is G, has a totally nonnegative completion, then G is a monotonically labeled block clique graph.*

Proof: As mentioned above, the proof is by contrapositive. We begin by ruling out non-chordal graphs. Let G be a graph on n vertices that is not chordal. Then G contains a simple cycle of length 4 or more as an induced subgraph. The graph of the specified entries of the k-by-k matrix

$$C = \begin{bmatrix} 1 & 0 & ? & ? & 2 \\ 2 & 1 & 0 & \ddots & ? \\ ? & 2 & 1 & \ddots & ? \\ ? & \ddots & \ddots & \ddots & 0 \\ 0 & ? & ? & 2 & 1 \end{bmatrix}$$

is a simple cycle and C is partial totally nonnegative. In order for the $\{2,3\},\{3,k\}$ minor to be nonnegative, the $2,k$ entry of C must be 0, regardless of the $3,k$ entry. However, if the $2,k$ entry is 0, then the $\{1,2\},\{1,k\}$ minor is -4. Thus, C has no completion to a totally nonnegative matrix. All that remains in order to rule out non-chordal graphs is to embed this data in a larger matrix in such a way that the matrix is a partial totally nonnegative matrix and use the fact that total nonnegativity is inherited by submatrices. This is done by specifying 1's on the diagonals and 0's for any other specified entries. The resulting matrix is totally nonnegative by lemma 4.9 since any fully specified principal submatrix is a direct sum of an identity and a principal submatrix of C (any other fully specified submatrix is a submatrix of one of these principal submatrices). However, because C does not have a totally nonnegative completion, by inheritance the larger matrix also does not have a totally nonnegative completion.

Next we look at chordal graphs that are not block clique graphs. Let G be a graph on n vertices that is not a block clique graph. Then there are two cliques C_1 and C_2 that are induced subgraphs of G and for which $|C_1 \cap C_2| \geq 2$. The simplest such graph is one in which $G = C_1 \cup C_2$, $|C_1| = |C_2| = 3$, and $|C_1 \cap C_2| = 2$. The graph of

$$A = \begin{bmatrix} 1 & 1 & .4 & ? \\ .4 & 1 & 1 & .4 \\ .2 & .8 & 1 & 1 \\ ? & .2 & .4 & 1 \end{bmatrix}$$

is G and A has no totally nonnegative completion. Define

$$A(x, y) = \begin{bmatrix} 1 & 1 & .4 & x \\ .4 & 1 & 1 & .4 \\ .2 & .8 & 1 & 1 \\ y & .2 & .4 & 1 \end{bmatrix}.$$

It is easy to check that all fully specified minors of $A(x, y)$ are positive. However,

$$\det(A(x, y)) = -.0016 - .008x - .328y - .2yx. \qquad (4.3)$$

For x and y nonnegative, (4.3) is always negative and, thus, there is no totally nonnegative completion of $A(x, y)$.

For the graphs that are not block clique graphs there is a little more work than in the non-chordal case in order to embed this data in a larger matrix. If G is not a block clique graph, then there exist cliques C_1 and C_2 that are induced subgraphs of G and for which $|C_1| = |C_2|$, and $|C_1 \cap C_2| = |C_1| - 1$ (note that C_1 and C_2 may not be maximal cliques). Let $B(x, y)$ be a partial matrix with graph G and let $\hat{B}(x, y)$ be the submatrix of $B(x, y)$ such that the graph of $\hat{B}(x, y)$ is $C_1 \cup C_2$. Then, assign data to $\hat{B}(x, y)$ by beginning with $A(x, y)$ and repeating row and column 2 as many times as necessary to complete the clique created by $C_1 \cap C_2$. That is

$$\hat{B}(x, y) = \begin{bmatrix} 1 & 1 & \cdots & 1 & .4 & x \\ .4 & 1 & \cdots & 1 & 1 & .4 \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ .4 & 1 & \cdots & 1 & 1 & .4 \\ .2 & .8 & \cdots & .8 & 1 & 1 \\ y & .2 & \cdots & .2 & .4 & 1 \end{bmatrix}.$$

Since any minor of this matrix is either 0 or is a minor of $A(x, y)$, $\hat{B}(x, y)$ is a partial totally nonnegative matrix. For the remainder of the specified entries of $B(x, y)$ specify 1's on the diagonal and 0's for the off diagonal entries as in the non-chordal case. The resulting matrix is partial totally nonnegative, but has $A(x, y)$ as a principal submatrix and, by inheritance, has no totally nonnegative completion.

Finally, if G is a block clique graph that is not monotonically labeled, then any partial totally nonnegative matrix the graph of whose specified entries is G will have a submatrix of the form

$$\begin{bmatrix} * & * & * \\ * & * & ? \\ * & ? & * \end{bmatrix}.$$

in which '*' indicates a specified position. The matrix

$$M(x,y) = \begin{bmatrix} 2 & 1 & 2 \\ 1 & 2 & x \\ 1 & y & 2 \end{bmatrix}$$

in which x, y are free to be chosen has this form and is partial totally positive. The $\{1,2\},\{3,4\}$ minor of $M(x,y)$ requires that $x \geq 4$ and the $\{3,4\},\{1,2\}$ minor requires that $y \geq 2$. Then, $xy \geq 8$, but the $\{3,4\},\{3,4\}$ minor requires that $xy \leq 4$. Therefore, $M(x,y)$ has no totally nonnegative completion. As usual, this data can be embedded in a larger matrix by specifying 1's on the diagonal and 0's in the other specified positions. The resulting matrix will be a partial totally nonnegative matrix, the graph of whose specified entries is block clique graph that is not linearly labeled, that has no totally nonnegative completion. □

The reverse implication of theorem 4.5 will be proven by induction. We will show that every regular partial totally nonnegative matrix, A, such that $G(A)$ is a monotonically labeled block clique graph with exactly two maximal cliques has a totally nonnegative completion and that completion is nonsingular. Assume the same is true for monotonically labeled block clique graphs with $k-1$ maximal cliques and let A be a regular partial totally nonnegative matrix such that $G(A)$ is a monotonically labeled block clique graph with k maximal cliques. Then, the principal submatrix $A_{C_1 \cup C_2}$ determined by the two maximal cliques C_1 and C_2 is a regular partial totally nonnegative matrix such that $G(A_{C_1 \cup C_2})$ is a monotonically labeled block clique graph containing the two cliques, C_1 and C_2. By the induction hypothesis, $A_{C_1 \cup C_2}$ has a totally positive completion. Let $\tilde{A}$ be the matrix that results when the submatrix $A_{C_1 \cup C_2}$ of A is completed. Then, $\tilde{A}$ is a regular partial totally positive matrix such that $G\left(\tilde{A}\right)$ is a monotonically labeled block clique graph with $k-1$ maximal cliques. By the induction hypothesis, $\tilde{A}$ has a totally nonnegative completion. But then A also has a totally nonnegative completion.

From this discussion we see that the only case that needs to be considered is the one in which A is an n-by-n matrix for which $G(A)$ is a monotonically labeled block clique graph with two maximal cliques. Then A is of the form

$$A = \begin{bmatrix} A_{11} & a_{12} & ? \\ a_{21}^T & a_{22} & a_{23}^T \\ ? & a_{32} & A_{33} \end{bmatrix} \tag{4.4}$$

in which $a_{12}, a_{21} \in \mathbb{R}^p$, $a_{23}, a_{32} \in \mathbb{R}^q$, and $p+q = n-1$. The following is a special case of the chordal result found in [JL].

Lemma 4.12 *Let*

$$A = \begin{bmatrix} A_{11} & a_{12} & a_{12}a_{23}^T \\ a_{21}^T & a_{22} & a_{23}^T \\ a_{32}a_{21}^T & a_{32} & A_{33} \end{bmatrix} \tag{4.5}$$

in which $a_{12}, a_{21} \in \mathbb{R}^p$, $a_{23}, a_{32} \in \mathbb{R}^q$, *and* $p+q = n-1$, $\begin{bmatrix} A_{11} & a_{12} \\ a_{21}^T & a_{22} \end{bmatrix}$ *and* $\begin{bmatrix} a_{22} & a_{23}^T \\ a_{32} & A_{33} \end{bmatrix}$ *are nonsingular, and* $a_{22} \neq 0$. *Then*

$$A^{-1} = \begin{bmatrix} \begin{bmatrix} A_{11} & a_{12} \\ a_{21}^T & a_{22} \end{bmatrix}^{-1} & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & \begin{bmatrix} a_{22} & a_{23}^T \\ a_{32} & A_{33} \end{bmatrix}^{-1} \end{bmatrix} - \begin{bmatrix} 0 & 0 & 0 \\ 0 & a_{22}^{-1} & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Note that the matrix in (4.5) is the completion of (4.4) that gives 0's in the unspecified entries of the inverse. This 0's in the inverse completion is shown in [BJLu]. Interestingly, the same completion works for total nonnegativity.

This result is used in the proof of the following lemma.

Lemma 4.13 *Let G be a monotonically labeled block clique graph. If A is a regular partial totally nonnegative matrix in which $G(A) = G$, then A has a nonsingular totally nonnegative completion.*

Proof: Let A be an n-by-n partial totally nonnegative matrix the graph of whose specified entries is a monotonically labeled block clique graph. The only case that needs to be considered is the case in which there are two cliques since the others follow by induction (see discussion above). Then A is of the form given in (4.4). Since positive left diagonal multiplication preserves total positivity and $a_{22} > 0$, it can be assumed without loss of generality that $a_{22} = 1$. Then, it will be shown that the completion given by

$$\tilde{A} = \begin{bmatrix} A_{11} & a_{12} & a_{12}a_{23}^T \\ a_{21}^T & 1 & a_{23}^T \\ a_{32}a_{21}^T & a_{32} & A_{33} \end{bmatrix} \tag{4.6}$$

is totally nonnegative. The proof will show that $D\tilde{A}^{-1}D^{-1}$ is totally nonnegative for

$$D = \begin{bmatrix} 1 & & & \\ & -1 & & \\ & & \ddots & \\ & & & (-1)^{n+1} \end{bmatrix}.$$

Then, by lemma 4.10, $\tilde{A}$ is totally nonnegative. By theorem 4.12

$$\tilde{A}^{-1} = \begin{bmatrix} \begin{bmatrix} A_{11} & a_{12} \\ a_{21}^T & 1 \end{bmatrix}^{-1} & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & \begin{bmatrix} 1 & a_{23}^T \\ a_{32} & A_{33} \end{bmatrix}^{-1} \end{bmatrix} - \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Define

$$T \equiv D\tilde{A}^{-1}D^{-1} \equiv \begin{bmatrix} B & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & C \end{bmatrix} - \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

in which

$$\begin{aligned} B &\equiv \begin{bmatrix} B_{11} & b_{12} \\ b_{21}^T & b_{22} \end{bmatrix} = D_p \begin{bmatrix} A_{11} & a_{12} \\ a_{21}^T & 1 \end{bmatrix}^{-1} D_p^{-1}, \\ C &\equiv \begin{bmatrix} c_{11} & c_{12}^T \\ c_{21} & C_{22} \end{bmatrix} = D_q \begin{bmatrix} 1 & a_{23}^T \\ a_{32} & A_{33} \end{bmatrix}^{-1} D_q^{-1}, \\ D_p &= D\left[\{1, 2, \ldots, p+1\}\right], \text{ and } D_q = D\left[\{n-q, n-q+1, \ldots, n\}\right]. \end{aligned}$$

Since $\begin{bmatrix} A_{11} & a_{12} \\ a_{21}^T & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & a_{23}^T \\ a_{32} & A_{33} \end{bmatrix}$ are totally nonnegative, B and C are also totally nonnegative by lemma 4.10. Using the adjoint formula for the inverse of $\tilde{A}^{-1}$ we see that

$$1 = \frac{\det(B_{11})\det(C_{22})}{\det\left(\tilde{A}^{-1}\right)}.$$

Then, since B_{11} and C_{22} are principal submatrices of totally positive matrices, $\det\left(\tilde{A}^{-1}\right) > 0$.

Now it must be shown that all minors of T of size $n-1$ and smaller are nonnegative. That is, $\det(T(\gamma;\delta)) \geq 0$ for all $\gamma,\delta \subseteq N, |\gamma| = |\delta|$. Let $\alpha \equiv \{1,2,\ldots,p\}$ and $\beta \equiv \{n-q+1, n-q+2,\ldots,n\}$. Then, the size of $T(\gamma;\delta)$ is $|\alpha| + |\beta| + 1 - |\gamma|$ (or $|\alpha| + |\beta| + 1 - |\delta|$). Many of the minors will be zero. This will be shown by finding a block of zeros large enough to ensure that $T(\gamma;\delta)$ is singular by lemma 4.8. For this to be the case, the sum of the dimensions of the zero block must be at least $|\alpha| + |\beta| - |\gamma| + 2$. There are two blocks of zeros in $T(\gamma;\delta)$ to consider. These blocks have dimensions

$$(|\alpha| - |\alpha \cap \gamma|) \text{-by-} (|\beta| - |\beta \cap \delta|) \tag{4.7}$$

and

$$(|\beta| - |\beta \cap \gamma|) \text{-by-} (|\alpha| - |\alpha \cap \delta|). \tag{4.8}$$

There are several cases to consider in order to show that $\det(T(\gamma;\delta)) \geq 0$ for all $\gamma,\delta \subseteq N$. If $p+1 \in \gamma,\delta$, then $T(\gamma;\delta)$ is a direct sum of submatrices of B and C and

$$|\gamma| - 1 = |\alpha \cap \gamma| + |\beta \cap \gamma| = |\alpha \cap \delta| + |\beta \cap \delta| = |\delta| - 1. \tag{4.9}$$

Since B and C are totally positive, if the submatrices in the direct sum are square, $\det(T(\gamma;\delta))$ is positive. If the submatrices of B and C are not square, then, if $|\alpha \cap \delta| > |\alpha \cap \gamma|$, the sum of the dimensions in (4.7) is:

$$\begin{aligned} |\alpha| + |\beta| - |\alpha \cap \gamma| - |\beta \cap \delta| &\geq |\alpha| + |\beta| - (|\alpha \cap \delta| - 1) - |\beta \cap \delta| \\ &= |\alpha| + |\beta| - (|\alpha \cap \delta| + |\beta \cap \delta|) + 1 \\ &= |\alpha| + |\beta| - |\delta| + 2. \end{aligned} \tag{4.10}$$

By lemma 4.8, this ensures that $T(\gamma;\delta)$ is singular. The case in which $|\alpha\cap\delta| \geq |\alpha\cap\gamma|$ uses (4.8) and is analogous.

There are several possibilities for $p+1 \notin \gamma,\delta$. If $|\alpha\cap\gamma| = |\alpha\cap\delta|$, then $T(\gamma;\delta)$ has the same form as T and, therefore, has positive determinant by the same argument as that for $\det(T)$. The determinant of $T(\gamma;\delta)$ is also positive if $||\alpha\cap\delta| - |\alpha\cap\gamma|| = 1$, because in this case $T(\gamma;\delta)$ is block triangular with submatrices of B and C on the diagonal. For the case in which $||\alpha\cap\delta| - |\alpha\cap\gamma|| \geq 2$ note that, since $p+1 \notin \gamma,\delta$,

$$|\gamma| = |\alpha\cap\gamma| + |\beta\cap\gamma| = |\alpha\cap\delta| + |\beta\cap\delta| = |\delta|. \tag{4.11}$$

Then, arguments similar to those in (4.10) (using (4.7) if $|\alpha\cap\delta| - |\alpha\cap\gamma| \geq 2$ and (4.8) if $|\alpha\cap\gamma| - |\alpha\cap\delta| \geq 2$) show that $T(\gamma;\delta)$ is singular.

It remains to show that the minors are nonnegative for the case in which $p+1$ is in exactly one of γ or δ. Assume $p+1 \in \gamma, k \notin \delta$ (the other case follows by symmetry). In this case,

$$|\alpha\cap\gamma| + |\beta\cap\gamma| = |\gamma| - 1 \tag{4.12}$$

and

$$|\alpha\cap\delta| + |\beta\cap\delta| = |\delta|. \tag{4.13}$$

If $|\alpha\cap\gamma| = |\alpha\cap\delta|$ or $|\alpha\cap\gamma| + 1 = |\alpha\cap\delta|$, then $T(\gamma;\delta)$ is block triangular with submatrices of B and C on the diagonal and $\det(T(\gamma;\delta)) > 0$. If $|\alpha\cap\gamma| = |\alpha\cap\delta| + 1$ then using (4.8) and (4.12) we see by lemma 4.8 that $T(\gamma;\delta)$ is singular. Similarly, if $||\alpha\cap\delta| - |\alpha\cap\gamma|| \geq 2$, $T(\gamma;\delta)$ contains a zero block that is sufficiently large, by lemma 4.8, to ensure that $T(\gamma;\delta)$ is singular. Thus, $\det(T(\gamma;\delta)) \geq 0$ for all $\gamma,\delta \subseteq N$

and $D\tilde{A}^{-1}D^{-1}$ is totally nonnegative. Then, by lemma 4.10, $\tilde{A}$ is totally nonnegative and the lemma is proved. □

Together, lemma 4.11 and lemma 4.13 prove theorem 4.5.

4.5 The Maximum Minimum Eigenvalue Completion Problem

The maximum minimum eigenvalue problem asks: for a partial Hermitian matrix A, what is the largest value that the minimum eigenvalue over all completions of A can attain? By classical interlacing, the minimum eigenvalue of an Hermitian matrix cannot exceed the minimum of the eigenvalues of any principal submatrix. Therefore, in the maximum minimum eigenvalue completion problem, the best that can be hoped for is the minimum of the minimum eigenvalues of the specified principal submatrices. In [GJSW] it was shown that every partial positive semidefinite matrix, the graph of whose specified entries is chordal has a positive semidefinite completion. This work was preceded by [DG] which considered the case of banded matrices. Since each fully specified principal submatrix of a partial positive semidefinite matrix is positive semidefinite, by translation the [GJSW] result is equivalent to saying that, in the chordal case, the maximum minimum eigenvalue over all completions is the minimum of the minimum eigenvalues of the fully specified principal submatrices.

The key observation in the proof in [GJSW] is the case in which there is only one pair of symmetrically placed unspecified entries (the one variable case). This one variable case was done before, and dates back at least to [DG] in which banded patterns were treated. Here we give an entirely different proof of the one variable case. This new proof uses the structured eigenvector results of chapter 2 and, therefore, for the one variable case our approach gives information about the structure of the eigenvectors of the maximum minimum eigenvalue completion that is not present in [GJSW]. This eigenvector information is however present in theorem 2.3 of [D2]. However, we prove both the completion result of [GJSW] and the eigenvector result of [D2] simultaneously and our approach allows us to write down the maximum minimum eigenvalue completion determined by these structured eigenvectors. In the following theorem let $C(A)$ denote the set of all completions of the partial matrix A.

Theorem 4.6 *Let $A \in M_n$ be a partial Hermitian matrix in which only $x = a_{ij}$ (and $\bar{x} = a_{ji}$) is unspecified. Also, let $A_1 = A(i)$, $A_2 = A(j)$ and $\lambda_1 = \min\{\mu : \mu \in \sigma(A_1)\}$, $\lambda_2 = \min\{\mu : \mu \in \sigma(A_2)\}$. Let $\lambda = \min\{\lambda_1, \lambda_2\}$, then*

$$\max\left\{\min_{B \in C(A)} \{\mu : \mu \in \sigma(B)\}\right\} = \lambda.$$

Moreover, there is a maximum minimum eigenvalue completion $\tilde{A}$ of A such that $\lambda = \min\left\{\mu : \mu \in \sigma\left(\tilde{A}\right)\right\}$ and there exists an associated eigenvector z such that if $\lambda = \lambda_1$, $z[i] = 0$ and if $\lambda = \lambda_2$, $z[j] = 0$.

Proof: It suffices to show that there is a completion $\tilde{A}$ of A in which $\lambda = \min\{\lambda_1, \lambda_2\}$ is the minimum eigenvalue of $\tilde{A}$. Assume, without loss of generality that the $1,n$ position is unspecified. Then

$$A = \begin{bmatrix} a_{11} & a_{21}^* & x \\ a_{21} & A_{22} & a_{23} \\ \bar{x} & a_{23}^* & a_{33} \end{bmatrix}$$

in which $A_{22} \in M_{n-2}$ is Hermitian. Then $A_1 = \begin{bmatrix} a_{11} & a_{21}^* \\ a_{21} & A_{22} \end{bmatrix}$ and $A_2 = \begin{bmatrix} A_{22} & a_{23} \\ a_{23}^* & a_{33} \end{bmatrix}$.

There are two cases to consider: $\lambda_1 \neq \lambda_2$ and $\lambda_1 = \lambda_2$. To begin, assume $\lambda_1 < \lambda_2$ so that $\lambda = \lambda_1$ (the case in which $\lambda_2 < \lambda_1$ follows by symmetry). By corollary 2.2 λ is an eigenvalue of a completion $\tilde{A}$ of A if and only if there is an eigenvector of $\tilde{A}$ associated with λ of the form $\begin{bmatrix} y \\ 0 \end{bmatrix}$ in which $y \in \mathbb{C}^{n-1}$ is in the eigenspace of A_1 associated with λ. In order for $\begin{bmatrix} y \\ 0 \end{bmatrix}$ to be an eigenvector of A associated with λ, the following equality must hold:

$$A \begin{bmatrix} y \\ 0 \end{bmatrix} = \begin{bmatrix} A_1 y \\ \bar{x} y[1] + a_{23}^* y(1) \end{bmatrix} = \lambda \begin{bmatrix} y \\ 0 \end{bmatrix}.$$

If $y[1] \neq 0$, then

$$x = \frac{-a_{23}^T \bar{y}(1)}{\bar{y}[1]} \tag{4.14}$$

gives the desired completion. If $y[1] = 0$, then

$$\begin{aligned} A_1 y &= \begin{bmatrix} a_{11} & a_{21}^* \\ a_{21} & A_{22} \end{bmatrix} \begin{bmatrix} 0 \\ y(1) \end{bmatrix} \\ &= \begin{bmatrix} a_{21}^* y(1) \\ A_{22} y(1) \end{bmatrix} = \lambda \begin{bmatrix} 0 \\ y(1) \end{bmatrix} \end{aligned}$$

so that $A_{22}y(1) = \lambda y(1)$ and λ is an eigenvalue of A_{22}. But this is a contradiction since, by interlacing, every eigenvalue of A_{22} must be greater than or equal to λ_2. Therefore, $y[1]$ cannot be 0.

Now, assume $\lambda = \lambda_1 = \lambda_2$, but $g_\lambda(A_1) \neq g_\lambda(A_2)$. After a similarity transformation of A that diagonalizes A_{22} this case can be reduced to the case discussed above. Assume without loss of generality that $g_\lambda(A_2) < g_\lambda(A_1)$ and let $g = g_\lambda(A_2)$. By interlacing, $g_\lambda(A_{22}) = g$ and $g_\lambda(A_1) = g + 1$. Also assume that $\lambda = 0$. We may do this without loss of generality since the general case follows easily by translation. Then, the maximum minimum eigenvalue completion $\check{A}$ of A is such that $\check{A}$ is positive semidefinite. Let U be the unitary matrix that diagonalizes A_{22}. Then,

$$U^* A_{22} U = \begin{bmatrix} 0 & & 0 \\ & \ddots & \\ & 0 & \\ 0 & & D \end{bmatrix}$$

in which $D = diag(\lambda_{g+1}, \ldots, \lambda_{n-2})$ and $\lambda_i > 0$, $i = g+1, \ldots, n-2$ are the nonzero eigenvalues of A_{22}. Then

$$\begin{bmatrix} I & 0 \\ & U^* \\ 0 & I \end{bmatrix} A \begin{bmatrix} I & 0 \\ & U \\ 0 & I \end{bmatrix} = \begin{bmatrix} a_{11} & \check{a}_{21}^* & x \\ \check{a}_{21} & \begin{bmatrix} 0 & \\ & D \end{bmatrix} & \check{a}_{23} \\ \bar{x} & \check{a}_{23}^* & a_{33} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} & 0 & \cdots & 0 & \hat{a}_{21}^* & x \\ 0 & 0 & & & & 0 \\ \vdots & & \ddots & & & \vdots \\ 0 & & & 0 & & 0 \\ \hat{a}_{21} & & & & D & \hat{a}_{23} \\ \bar{x} & 0 & \cdots & 0 & \hat{a}_{23}^* & a_{33} \end{bmatrix} \equiv B. \quad (4.15)$$

The last equality holds because if a positive semidefinite matrix has a zero on the diagonal, then the entire row and column containing that diagonal entry must be 0 [HJ1, pg. 400]. The matrix B in (4.15) is the direct sum of a 0 matrix and

$$\hat{B} \equiv \begin{bmatrix} a_{11} & \hat{a}_{21}^* & x \\ \hat{a}_{21} & D & \hat{a}_{23} \\ \bar{x} & \hat{a}_{23}^* & a_{33} \end{bmatrix}. \quad (4.16)$$

Therefore, if we find a completion of $\hat{B}$ that is positive semidefinite, we will have a completion of B that is also positive semidefinite. Then, since similarity preserves eigenvalues, we have also found a completion of A that is positive semidefinite as desired. Let $\hat{B}_1 = \begin{bmatrix} a_{11} & \hat{a}_{21}^* \\ \hat{a}_{21} & D \end{bmatrix}$ and $\hat{B}_2 = \begin{bmatrix} D & \hat{a}_{23} \\ \hat{a}_{23}^* & a_{33} \end{bmatrix}$. Since $g_\lambda(A_1) = g_\lambda(A_2) + 1$, $\hat{B}_1$ is positive semidefinite and $\hat{B}_2$ is positive definite. Thus, $\hat{B}$ is a partial positive definite matrix in which the smallest eigenvalue of $\hat{B}_1$ is less than the smallest eigenvalue of $\hat{B}_2$ so that we are in the case discussed earlier. For y an eigenvector of $\hat{B}_1$ associated with λ by the calculation as in (4.14) we see that $x = \frac{-\hat{a}_{23}^T \bar{y}(1)}{\bar{y}[1]}$ gives the desired completion.

We next look at the case in which $\lambda = \lambda_1 = \lambda_2$ and $g_\lambda(A_1) = g_\lambda(A_2)$. Let $g = g_\lambda(A_1)$ and $\alpha = \{2, \ldots, n-1\}$. By interlacing, $g_\lambda(A_{22})$ is g or $g-1$. First consider the case in which $g_\lambda(A_{22}) = g-1$. In this case, by corollary 2.2

$$\begin{aligned}\dim\left(LE_\alpha^\lambda(A_1)\right), \dim\left(RE_\alpha^\lambda(A_1)\right) &\geq \frac{g_\lambda(A_1) + g_\lambda(A_{22}) - 1}{2} \\ &= g - 1.\end{aligned}$$

Thus, there are at least $g-1$ linearly independent eigenvectors of A_1 of the form $\begin{bmatrix} 0 \\ u_i \end{bmatrix}$, $u_i \in \mathbb{C}^{n-2}$. By partitioned matrix multiplication each u_i, $i = 1, 2, \ldots, g-1$, is an eigenvector of A_{22} associated with λ. But, since $g_\lambda(A_1) = g$, this means that there is another eigenvector y of A_1 associated with λ for which $\left\{ y, \begin{bmatrix} 0 \\ u_1 \end{bmatrix}, \ldots, \begin{bmatrix} 0 \\ u_{g-1} \end{bmatrix} \right\}$ is a linearly independent set. In addition, it must be the case that $y[1] \neq 0$ since if $y[1] = 0$, then $y(1)$ is another eigenvector of A_{22} which is a contradiction since $g_\lambda(A_{22}) = g - 1$ and we already have the $g-1$ linearly independent vectors u_i of A_{22} associated with λ. The eigenvector y can be used as in (4.14) to find a value for x that gives a completion $\tilde{A}$ of A with λ as an eigenvalue and, then, $\begin{bmatrix} y \\ 0 \end{bmatrix}$ is an eigenvector of $\tilde{A}$ associated with λ. Note that the matrix A_2 is an $(n-1)$-by-$(n-1)$ matrix that we will consider to have entries indexed from 2 to n. To see that λ is the minimum eigenvalue of $\tilde{A}$ note that by corollary 2.2

$$\begin{aligned}dim\left(LE_\alpha^\lambda(A_2)\right), dim\left(RE_\alpha^\lambda(A_2)\right) &\geq \frac{g_\lambda(A_2) + g_\lambda(A_{22}) - 1}{2} \\ &= g - 1.\end{aligned}$$

Then, for each u_i, $i = 1, 2, \ldots, g-1$, the vector $\begin{bmatrix} u_i \\ 0 \end{bmatrix}$ is an eigenvector of A_2 associated with λ. Recall that the vectors $\begin{bmatrix} 0 \\ u_i \end{bmatrix}$, $i = 1, 2, \ldots, g-1$, are eigenvectors of A_1 associated with λ. But then for each $i = 1, 2, \ldots, g-1$, $\begin{bmatrix} 0 \\ u_i \\ 0 \end{bmatrix}$ is an eigenvector of $\tilde{A}$ associated with λ. Then there are at least g linearly independent eigenvectors of $\tilde{A}$ associated with λ so that $g_\lambda\left(\tilde{A}\right) \geq g$ and for $\gamma = \{2, 3, \ldots, n\}$ corollary 2.2 gives

$$\begin{aligned}\dim\left(LE_\gamma^\lambda\left(\tilde{A}\right)\right) \cdot \dim\left(RE_\gamma^\lambda\left(\tilde{A}\right)\right) &\geq \frac{g_\lambda\left(\tilde{A}\right) + g_\lambda\left(A_2\right) - 1}{2} \\ &\geq g - \frac{1}{2}.\end{aligned}$$

Since the dimension of the special eigenspaces must be integer, this means there are at least g linearly independent eigenvectors of $\tilde{A}$ associated with λ with a 0 in the first position. We already know the $g-1$ vectors $\begin{bmatrix} 0 \\ u_i \\ 0 \end{bmatrix}$ are eigenvectors of $\tilde{A}$ associated with λ. So, there must be at least one vector of the form $\begin{bmatrix} 0 \\ z \end{bmatrix}$ in which $z \in \mathbb{C}^{n-1}$ and $z_{n-1} \neq 0$ (otherwise we would contradict the fact that $g_\lambda\left(A_{22}\right) = g - 1$.)

Then $\left\{ \begin{bmatrix} y \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ u_1 \\ 0 \end{bmatrix}, \ldots, \begin{bmatrix} 0 \\ u_{g-1} \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ z \end{bmatrix} \right\}$ is a set of $g+1$ linearly independent eigenvectors of $\tilde{A}$ associated with λ and, therefore, by interlacing, λ is the minimum eigenvalue of $\tilde{A}$.

All that remains is the case in which $g = g_\lambda(A_1) = g_\lambda(A_2)$ and $g_\lambda(A_{22}) = g$. Once again, assume that $\lambda = 0$ and perform the similarity of A as in (4.15). Then, every vector of the form $\begin{bmatrix} 0 \\ w \\ 0 \end{bmatrix}$, $w \in \mathbb{C}^g$ is an eigenvector of B associated with $\lambda = 0$ and, therefore, $\lambda = 0$ is an eigenvalue of A for all values of x. We must find x so that $\lambda = 0$ is the minimum eigenvalue of A. Since $g_\lambda(A_1) = g_\lambda(A_2) = g_\lambda(A_{22})$ the matrix $\hat{B}$ as defined in (4.16) has both $\hat{B}_1$ and $\hat{B}_2$ positive definite. In order for A to be positive semidefinite we must find x so that $\hat{B}$ is positive definite. The Schur complement of D in $\hat{B}$ is

$$\begin{bmatrix} a_{11} & x \\ \bar{x} & a_{33} \end{bmatrix} - \begin{bmatrix} \hat{a}_{12}^* \\ \hat{a}_{23}^* \end{bmatrix} D^{-1} \begin{bmatrix} \hat{a}_{12} & \hat{a}_{23} \end{bmatrix}$$
$$= \begin{bmatrix} a_{11} - \hat{a}_{12}^* D^{-1} \hat{a}_{12} & x - \hat{a}_{12}^* D^{-1} \hat{a}_{23} \\ \bar{x} - \hat{a}_{23}^* D^{-1} \hat{a}_{12} & a_{33} - \hat{a}_{23}^* D^{-1} \hat{a}_{23} \end{bmatrix}. \qquad (4.17)$$

The diagonal entries $a_{11} - \hat{a}_{12}^* D^{-1} \hat{a}_{12}$ and $a_{33} - \hat{a}_{23}^* D^{-1} \hat{a}_{23}$ are the Schur complements of D in $\hat{B}_1$ and $\hat{B}_2$, respectively. Since $\hat{B}_1$ and $\hat{B}_2$ are both positive definite, these diagonal entries are positive. Then, choosing $x = \hat{a}_{12}^* D^{-1} \hat{a}_{23}$ gives a completion of $\hat{B}$ that is positive definite and, therefore, the resulting completion of A is positive semidefinite with $\lambda = 0$ as the minimum eigenvalue. This completes the proof. □

Note that the completion in the last case of this proof is the completion given in [GJSW] and [DG]. However, by this proof we know a little bit more since in this last case we know that the completed matrix has an eigenvector associated with the minimum eigenvalue λ of the form $\begin{bmatrix} 0 \\ y \\ 0 \end{bmatrix}$, in which y is an eigenvector of A_{22} associated with λ.

As shown in [GJSW], the general chordal problem may be solved by sequential application of the one variable problem. So, theorem 4.6 may be applied sequentially to solve the chordal maximum minimum eigenvalue problem. As mentioned above, the approach to the proof of the theorem that we use provides information about the eigenvectors associated with the maximum minimum eigenvalue that the [GJSW] approach does not provide. In addition, the maximum minimum eigenvalue completion as given in the above proof is easy to write down. It is shown in theorem 1.2 of [D2] that, in the chordal case, there exists an eigenvector of the maximum minimum eigenvalue completion with support contained in the entries that correspond to the maximal specified principal submatrix with the smallest minimum eigenvalue of all the maximal specified principal submatrices.

Bibliography

[A] T. Ando, Totally Positive Matrices, *Linear Algebra Appl.* 90 (1987), 165-219.

[B] L.M. Blumenthal, *Theory and Applications of Distance Geometry*, 2nd Ed., New York, Chelsea (1970).

[BJ] M. Bakonyi and C.R. Johnson, *The Euclidean Distance Matrix Completion Problem*, submitted.

[BJL] W. Barrett, C.R. Johnson, and R. Loewy, *The Real Positive Definite Completion Problem: Cycle Completability*, submitted.

[BJLu] W. Barrett, C.R. Johnson, and M. Lundquist, Determinantal Formulae for Matrix Completions Associated with Chordal Graphs, *Linear Algebra Appl.* 121 (1989), 265-290.

[BJT] W. Barret, C.R. Johnson, and P. Tarazaga, The Real Positive Definite Completion Problem for a Simple Cycle, *Linear Algebra Appl.* 192 (1993), 3-31.

[BP] A. Berman, and R.J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, Academic Press, New York, 1979.

[C] D. Carlson, What are Schur Complements Anyway?, *Linear Algebra Appl.* 74 (1986), 257-275.

[CJRW] N. Cohen, C.R. Johnson, L. Rodman, and H.J. Woerdeman, Ranks of Completions of Partial Matrices, *Oper. Theory Adv. Appl.* 40 (1989), 165-185.

[D1] J. Dancis, An Interlacing Theorem for Eigenvalues of Self-Adjoint Operators, *Linear Alg. Appl.* 88/89 (1987), 117-122.

[D2] J. Dancis, Positive Semidefinite Completions of Partial Hermitian Matrices, *Linear Algebra Appl.* 175 (1992), 97-114.

[DG] H. Dym, and I. Gohberg, Extensions of Band Matrices with Banded Inverses, *Linear Algebra Appl.* 36 (1981), 1-24.

[F] M. Fiedler, Matrix Inequalities, *Numerische Mathematik* 9 (1966), 109-119.

[G] M. Golumbic, *Algorithmic Graph Theory and Perfect Graphs*, Academic Press, New York (1980).

[GJSW] R. Grone, C.R. Johnson, E. Sá, and H. Wolkowicz, Positive Definite Completions of Partial Hermitian Matrices, *Linear Algebra Appl.* 58 (1984), 109-124.

[GM] M. Gasca, and G. Muhlbach, A Test for Strict Total Positivity via Neville Elimination, *Current Trends in Matrix Theory* (1987), 225-232.

[GRSW] I. Gohberg, L. Rodman, T. Shalom, and H.J. Woerdeman, Bounds for Eigenvalues and Singular Values of Matrix Completions, *Linear and Multilinear Algebra* 33 (1992), 1-17.

[H] R.O. Hill, Certain Rank $2k$ Updates Act Like Rank k, *Linear Algebra Appl.* 177 (1992), 173-174.

[HC] T. Havel, and G. Crippen, *Distance Geometry and Molecular Conformation*, Wiley, New York, 1988.

[HJ1] R. Horn, and C.R. Johnson, *Matrix Analysis*. Cambridge University Press, Cambridge, 1985.

[HJ2] R. Horn, and C.R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1991.

[HJK] R.O. Hill, C.R. Johnson, and B.K. Kroschel, Extended Interlacing Intervals, *Linear Algebra Appl.*, accepted.

[HP] R. Hill, and B.N. Parlett, Refined Interlacing Properties, *SIAM J. Matrix Anal. Appl.* 13 (1992), 239-247.

[J] C.R. Johnson, Matrix Completion Problems: A Survey, *Proc. of Symposia in Appl. Math.* 40 (1990), 171-198.

[JJK] C.R. Johnson, C.A. Jones, and B.K. Kroschel, The Euclidean Distance Completion Problem: Cycle Completability, *Linear and Multilinear Albebra* 39 (1995), 195-207.

[JK1] C.R. Johnson, and B.K. Kroschel, Principal Submatrices, Geometric Multiplicities, and Structured Eigenvectors, *SIAM J. Matrix Anal. Appl.* 16 (1995), 1004-1012.

[JK2] C.R. Johnson, and B.K. Kroschel, The Combinatorially Symmetric P-matrix Completion Problem, submitted.

[JL] C.R. Johnson, and M. Lundquist, An Inverse Formula for Matrices with Sparse Inverses, in preparation.

[JM] C.R. Johnson, and T. McKee, Structural Conditions for Cycle Completable Graphs, *Discrete Math.*, accepted.

[JT] C.R. Johnson, and P. Tarazaga, Connections Between the Real Positive Semidefinite and Distance Matrix Completion Problems, submitted.

[K] Karlin, *Totally Positive Matrices.* Stanford U.P., Stanford, 1968.

[L] N.J. Lehmann, Optimale Eigenwerteinschiessung, *Numer. Math.* 5 (1963), 246-272.

[dLH] J. de Leeuw, and W. Heiser, Theory of Multidimensional Scaling, Chapter 13 in P.R. Krishnaiah, and L.H. Kanal (eds.), *Handbook of Statistics*, vol. 2, North-Holland, Amsterdam, 1982.

[M] Minc, H.K., *Nonnegative Matrices.* John Wiley & Sons, Inc., New York, 1988.

[P] B. Parlett, *The Symmetric Eigenvalue Problem.* Prentice-Hall, Englewood Cliffs, NJ, 1980.

[Ro] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, NJ, 1970.

[Ry] H.J. Ryser, *Combinatorial Mathematics*, The Mathematical Association of America, Rahway, 1963.

[S] H. Schneider, The Concepts of Irreducibility and Full Indecomposability of a Matrix in the works of Frobenius, König and Markov, *Linear Algebra Appl.* 18 (1977), 139-162.

[Wi] H. Wielandt, An Extremum Property of Sums of Eigenvalues, *Proc. Amer. Math. Soc.* 6 (1955), 106-110.

[Wo] H.J. Woerdeman, The Lower Order of Lower Triangular Operators and Minimal Rank Extensions, Integral Equations, *Operator Theory* 10 (1987), 859-879.

VITA

Brenda K. Kroschel

Born in St. Paul. Minnesota on January 23, 1963. Valedictorian of Mahtomedi High School, 1981. Received a B.A. in Mathematics and Chemistry from St. Olaf college in May 1985 and an M.S. in Mathematical Sciences from Clemson University in May 1987. Worked as a programmer analyst for Control Data Corporation in the Surfacing and Machining group from 1987 to 1989. From 1989 to 1992 worked as a statistical programmer analyst for O'Connor and Associates, an options trading firm in Chicago, and was a member of the technical staff upon leaving O'Connor and Associates. Entered the College of William and Mary Applied Science program as a research assistant in August of 1992 to begin work on a Ph.D. degree.