Matrix Completion Problems

Monique Laurent
CWI, Kruislaan 413, 1098 SJ Amsterdam, The Netherlands
monique@cwi.nl
author A0280, article P0550

Matrix completion problems are concerned with determining whether partially specified matrices can be completed to fully specified matrices satisfying certain prescribed properties. In this article we survey some results and provide references about these problems for the following matrix properties: positive semidefinite matrices, Euclidean distance matrices, completely positive matrices, contraction matrices, and matrices of given rank. We treat mainly optimization and combinatorial aspects.

1 Introduction

A *partial matrix* is a matrix whose entries are specified only on a subset of its positions; a *completion* of a partial matrix is simply a specification of the unspecified entries. *Matrix completion problems* are concerned with determining whether or not a completion of a partial matrix exists which satisfies some prescribed property. We consider here the following matrix properties: positive (semi)definite matrices, distance matrices, completely positive matrices, contraction matrices, and matrices of given rank; definitions are recalled below.

In what follows, x^*, A^* denote the conjugate transpose (in the complex case) or transpose (in the real case) of vector x and matrix A. A square real symmetric or complex Hermitian matrix A is *positive semidefinite* (psd) if $x^*Ax \geq 0$ for all vectors x and *positive definite* (pd) if $x^*Ax > 0$ for all vectors $x \neq 0$; then we write: $X \succeq 0$ ($X \succ 0$). Equivalently, A is psd (resp. pd) if and only if all its eigenvalues are nonnegative (resp. positive) and A is psd if and only if $A = BB^T$ for some matrix M. A matrix A is said to be *completely positive* if $A = BB^T$ for some *nonnegative* matrix B. An $n \times n$ real symmetric matrix $D = (d_{ij})$ is a *Euclidean distance matrix* (abbreviated as *distance matrix*) if there exist vectors $v_1, \ldots, v_n \in \mathbb{R}^k$ (for some $k \geq 1$) such that, for all $i, j = 1, \ldots, n$, d_{ij} is equal to the square of the Euclidean distance between v_i and v_j. Finally, a (rectangular) matrix A is a *contraction* matrix if all its singular values (that is, the eigenvalues of A^*A) are less than or equal to 1.

The set of positions corresponding to the specified entries of a partial matrix A is known as the *pattern* of A. If A is an $n \times m$ partial matrix, its pattern can be represented by a bipartite graph with node bipartition $[1, n] \cup [1, m]$ having an edge between nodes $i \in [1, n]$ and $j \in [1, m]$ if and only if entry a_{ij} is specified.

When asking about existence of a psd completion of a partial $n \times n$ matrix A, it is commonly assumed that all diagonal entries of A are specified (which is no loss of generality if we ask for a pd completion); moreover, it can obviously be assumed that A is *partial Hermitian*, which means that entry a_{ji} is specified and equal to a_{ij}^* whenever a_{ij} is specified. Hence, in this case, complete information about the pattern of A is given by the graph $G = ([1, n], E)$ with node set $[1, n]$ and whose edge set E consists of the pairs ij $(1 \leq i < j \leq n)$ for which a_{ij} is a specified entry of A. The same holds when dealing with distance matrix completions (in which case diagonal entries can obviously be assumed to be equal to zero).

An important common feature of the above matrix properties is that they possess an "inheritance structure". Indeed, if a partial matrix A has a psd (pd, completely positive, distance matrix) completion, then every principal specified submatrix of A is psd (pd, completely positive, a distance matrix); similarly, if a partial matrix A admits a completion of rank $\leq k$, then every specified submatrix of A has rank $\leq k$. Hence, having a completion of a certain kind imposes certain "obvious" necessary conditions. This leads to asking which are the patterns for the specified entries that insure that if the obvious necessary conditions are met, then there will be a completion of the desired type; therefore, this introduces a combinatorial aspect into matrix completion problems, as opposed to their analytical nature.

In this article we survey some results and provide references for the various matrix completion problems mentioned above, concerning optimization and combinatorial aspects of the problems. We refer to [Jo90], [La98b] for more detailed surveys on some of the topics treated here.

2 The positive semidefinite completion problem

We consider here the following *positive (semi)definite completion problem (PSD):* Given a partial Hermitian matrix $A = (a_{ij})_{ij \in S}$ whose entries are specified on a subset S of the positions, determine whether A has a psd (or pd) completion; if, yes, find such a completion. (Here, S is generally assumed to contain all diagonal positions.)

This problem belongs to the most studied matrix completion problems. This is due, in particular, to its many applications, e.g., in probability and statistics, systems engineering, geophysics, etc., and also to the fact that positive semidefiniteness is a basic property which is closely related to other matrix properties like being a contraction or distance matrix. Equivalently, (PSD) is the problem of testing feasibility of the following system (in variable $X = (x_{ij})$):

(1) $\quad X \succeq 0, \ x_{ij} = a_{ij} \ (ij \in S).$

Therefore, (PSD) is an instance of the following *semidefinite programming problem (P)*: Given Hermitian matrices $A_1, \ldots, A_m$ and scalars $b_1, \ldots, b_m$, decide whether the following system is feasible:

$$(2) \qquad X \succeq 0, \; A_j \cdot X = b_j \; (j = 1, \ldots, m)$$

(where $A \cdot X := \sum_{i,j=1}^n a_{ij}^* x_{ij}$ for two Hermitian $n \times n$ matrices A and X).

The exact complexity status of problems (PSD) and (P) is not known; in particular, it is not known whether they belong to the complexity class NP. However, it is shown in [Ra97] that (P) is neither NP-complete nor co-NP-complete if NP $\neq$ co-NP. However, the semidefinite programming problem and, thus, problem (PSD) can be solved with an arbitrary precision in polynomial time. This can be done using the ellipsoid method (since one can test in polynomial time whether a rational matrix A is positive semidefinite and, if not, find a vector x such that $x^*Ax < 0$; cf. [GLS88]), or interior-point methods (cf. [NN94], [Al95], [HRVW96]). There has been a growing interest in semidefinite programming in the recent years, which is due, in particular, to its successful application to the approximation of hard combinatorial optimization problems (cf. the survey [Go97]). This has prompted active research on developing interior-point algorithms for solving semidefinite programming problems; as the literature is quite large we refer to the following websites where extensive information can be found:
http://www.zib.de/helmberg/semidef.html and
http://orion.math.uwaterloo.ca:80/~hwolkowi/henry/software/readme.html.
Numerical tests are reported in [JKW98] where an interior-point algorithm is proposed for the approximate psd completion problem; it permits to find exact completions for random instances up to size 110.

Moreover, it is shown in [PK97] that problem (P) can be solved in polynomial time (for rational input data A_j, b_j) if either the number m of constraints, or the order n of the matrices X, A_j in (2) is fixed (cf. also [Ba93]). Moreover, under the same assumption, one can test in polynomial time the existence of an integer solution and find one if it exists [KP97].

Call a partial Hermitian matrix A *partial psd* (resp. *partial pd*) if every principal specified submatrix of A is psd (resp. pd). As mentioned in the Introduction, being partial psd (pd) is an obvious necessary condition for A to have a psd (pd) completion. In general, this condition is not sufficient; for instance, the partial matrix:

$$A = \begin{pmatrix} 1 & 1 & ? & 0 \\ 1 & 1 & 1 & ? \\ ? & 1 & 1 & 1 \\ 0 & ? & 1 & 1 \end{pmatrix}$$

('?' indicates an unspecified entry) is partial psd, yet no psd completion exists; note that the pattern of A is a circuit of length 4. Call a graph *chordal* if it does not contain any circuit of length ≥ 4 as an induced subgraph; chordal graphs occur in particular in connection with the Gaussian elimination process for sparse pd matrices (cf. [Ro70],[Go80]). (An *induced subgraph* of a graph $G = (V, E)$ being of the form $H = (U, F)$ where $U \subseteq V$ and $F := \{ij \in E \mid i, j \in U\}$.) It is shown in [GJSW84] that every partial psd matrix with pattern G has a psd completion if and only if G is a chordal graph; the same holds for pd completions. This extends an earlier result from [DG81] which dealt with 'block-banded' partial matrices; in the Toeplitz case (all entries equal along a band), one finds the classical Carathéodory-Fejer theorem from function theory.

The proof from [GJSW84] is constructive and can be turned into an algorithm with a polynomial running time [La98d]. Moreover, it is shown in [La98d] that (PSD) can be solved in polynomial time when restricted to partial rational matrices whose pattern is a graph having a fixed minimum fill-in; the *minimum fill-in* of a graph being the minimum number of edges needed to be added in order to obtain a chordal graph. This result is based on the above mentioned results from [PK97, KP97] concerning the polynomial-time solvability of (integer) semidefinite programming with a fixed number m of linear constraints in (2).

The result from [GJSW84] on psd completions of partial matrices with a chordal pattern has been generalized in various directions; for instance, considering general inertia possibilities for the completions ([JR84], [ELG88]), or considering completions with entries in a function ring [JR88].

If A is a partial matrix having a pd completion, then A has a unique pd completion with maximum determinant (this unique completion being characterized by the fact that its inverse has zero entries at all unspecified positions of A) [GJSW84]. In the case when the pattern of A is chordal, explicit formulaes for this maximum determinant are given in [BJL89]. The paper [LJ91] considers the more general problem of finding a maximum determinant psd completion satisfying some additionnal linear constraints.

Further necessary conditions are known for the existence of psd completions. Namely, it is shown in [BJT93] that if a partial matrix $A = (a_{ij})$ with pattern G and diagonal entries equal to 1 is completable to a psd matrix, then the associated vector $x := (\frac{1}{\pi} \arccos(a_{ij})_{ij \in E})$ satisfies the inequalities:

$$(3) \qquad \sum_{e \in F} x_e - \sum_{e \in C \setminus F} x_e \leq |F| - 1 \text{ for all } F \subseteq C, \ C \text{ circuit in } G, |F| \text{ odd.}$$

Moreover, any partial matrix with pattern G satisfying (3) is completable to a psd matrix if and only if G does not contain a homeomorph of K_4 as an induced subgraph (then, G is also known as *series-parallel graph*) [La97a]. (Here, K_4 denotes the complete graph on 4 nodes and a *homeomorph of* K_4 is obtained by replacing the edges of K_4 by paths of arbitrary length.) The patterns G for which every partial psd matrix satisfying (3) has a psd completion are characterized in [BJL96]; they are the graphs G which can be made chordal by adding a set of edges in such a way that no new clique of size 4 is created. Although (3) can be checked in polynomial time for rational x [BM86], the complexity of problem (PSD) for series-parallel graphs (or for the subclass of circuits) is not known. A strengthening of condition (3) (involving cuts in graphs) is formulated in [La97a].

Another approach to problem (PSD) is considered in [AHMR88, HPR89], which is based on the study of the cone

$$\mathcal{P}_G := \{X = (x_{ij})_{i,j \in V} \mid X \succeq 0,\ x_{ij} = 0\ \forall i \neq j,\ ij \notin E\}$$

associated to graph $G = (V, E)$. Indeed, it is shown there that a partial matrix A with pattern G has a psd completion if and only if

$$(4) \qquad \sum_{i \in V} a_{ii} x_{ii} + \sum_{i \neq j,\ ij \in E} a_{ij} x_{ij} \geq 0\ \forall X \in \mathcal{P}_G.$$

Obviously, it suffices to check (4) for all X *extremal* in $\mathcal{P}_G$ (i.e., X lying on an extremal ray of the cone $\mathcal{P}_G$). Define the *order* of G as the smallest rank of an extremal matrix in $\mathcal{P}_G$. The graphs of order 1 are precisely the chordal graphs [AHMR88, PPS89] and the graphs of order 2 have been characterized in [La98c]. One might reasonably expect that problem (PSD) is easier for graphs having a small order. This is indeed the case for graphs of order 1; the complexity of (PSD) remains however open for the graphs of order 2 (partial results are given in [La98d]).

3 The Euclidean distance matrix completion problem

We consider here the *Euclidean distance matrix completion problem* (abbreviated as *distance matrix completion problem*) *(EDM)*: Given a graph $G = (V = [1, n], E)$ and a real partial symmetric matrix $A = (a_{ij})$ with pattern G and with zero diagonal entries, determine whether A can be completed to a distance matrix; that is, whether there exist vectors $v_1, \ldots, v_n \in \mathbb{R}^k$ for some $k \geq 1$ such that

$$(5) \qquad a_{ij} = \|v_i - v_j\|^2 \text{ for all } ij \in E.$$

(here, $\|v\| = \sqrt{\sum_{h=1}^k v_h^2}$ denotes the Euclidean norm of $v \in \mathbb{R}^k$.) The vectors $v_1, \ldots, v_n$ are then said to form a *realization* of A. A variant of problem (EDM) is the *graph realization problem (EDMk)*, obtained by letting the dimension k of the space where one searches for a realization of A be part of the input data.

Distance matrices are a central notion in the area of distance geometry; their study was initiated by Cayley last century and it was continued in particular by Menger and Schoenberg in the 1930s. They are, in fact, closely related to psd matrices. The following basic connection was established in [Sc35]. Given a symmetric $n \times n$ matrix $D = (d_{ij})_{i,j=1}^n$ with zero diagonal entries, consider the symmetric $(n-1) \times (n-1)$ matrix $X = (x_{ij})_{i,j=1}^{n-1}$ defined by

$$(6) \qquad x_{ij} := \frac{1}{2}(d_{in} + d_{jn} - d_{ij}) \text{ for all } i, j = 1, \ldots, n-1.$$

Then, D is a distance matrix if and only if X is psd; moreover, D has a realization in the k-space if and only if X has rank $\leq k$. Other characterizations are known for distance matrices. As the literature on this topic is quite large, we refer to the monographs [Bl53], [CH88], [DL97], where further references can be found.

Problems (EDM) and (EDMk) have many important applications; for instance, to multidimensional scaling problems in statistics (cf. [LH82]) and to position-location problems, i.e., problem (EDMk) mostly in dimension $k \leq 3$. A much studied instance of the latter problem is the molecular conformation problem in chemistry; indeed, nuclear magnetic resonance spectroscopy permits to determine some pairwise interatomic distances, the question being then to reconstruct the global shape of the molecule from this partial information (cf. [CH88, KTO93]).

In view of relation (6), problem (EDM) can be formulated as an instance of the semidefinite programming problem (P) and, therefore, it can be solved with an arbitrary precision in polynomial time. Exploiting this fact, some specific algorithms based on interior-point methods are presented in [AKW97] together with numerical tests. Moreover, problem (EDM) can be solved in polynomial time when restricted to partial rational matrices whose pattern is a chordal graph or, more generally, a graph with fixed minimum fill-in [La98d]; as in the psd case, this follows from the fact (mentioned below) that partial matrices that are completable to a distance matrix admit a good characterization when their pattern is a chordal graph.

While the exact complexity of problem (EDM) is not known, it has been shown in [Sa79] that problem (EDMk) is NP-complete if $k = 1$ and NP-hard if $k \geq 2$ (even when restricted to partial matrices with entries in $\{1, 2\}$). Finding ϵ-optimal solutions to the graph realization problem is also NP-hard for small ϵ ([MW96]). The graph realization problem (EDMk) has been much studied, in particular in dimension $k \leq 3$, which is the case most relevant to applications. The problem can be formulated as a nonlinear global optimization problem: $\min f(v)$ such that $v = (v_1, \ldots, v_n) \in \mathbb{R}^{kn}$, where the cost function $f(.)$ can, for instance, be chosen as

$$f(v) = \sum_{ij \in E} (\|v_i - v_j\|^2 - a_{ij})^2.$$

Hence, $f(.)$ is zero precisely when the v_i's provide a realization of the partial matrix A. This optimization problem is hard to solve (as it may have many local optimum solutions). Several algorithms have been proposed in the literature; see, in particular, [CH88], [GHR93], [Ha91], [He90, He95], [KTO93], [MW97], [PL97]. They are based on general techniques for global optimization like tabu and pattern search [PL97], the continuation approach (which consists of transforming the original function $f(.)$ into a smoother function having fewer local optimizers, [MW96, MW97]), or divide-and-conqueer strategies aiming to break the problem into a sequence of smaller or easier subproblems [CH88, He90, He95]. In [He90, He95], the basic step consist of finding principal submatrices having a unique realization, treating each of them separately and then trying to combine the solutions. Thus arises the problem of identifying principal submatrices having a unique realization, which turns out to be NP-hard [Sa79]. However, several necessary conditions for unicity of realization are known, related with connectivity and generic rigidity properties of the graph pattern [Ye79, He92]. Generic rigidity of graphs can be characterized and recognized in polynomial time only in dimension $k \leq 2$ ([La70], [LY82]) (cf. survey [La97b] for more references).

Call a partial matrix A a *partial distance matrix* if every specified principal submatrix of A is a distance matrix. Being a partial distance matrix is obviously a necessary condition for A to be completable to a distance matrix. It is shown in [BJ95] that every partial distance matrix with pattern G is completable to a distance matrix if and only if G is a chordal graph; moreover, if all specified principal submatrices of the partial matrix A have a realization in the k-space, then A admits a completion having a realization in the k-space.

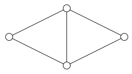
As noted in [JJK95], if a partial matrix A with pattern G is completable to a distance matrix, then the associated vector $x := (\sqrt{a_{ij}})_{ij \in E}$ must satisfy the inequalities:

$$(7) \qquad x_e - \sum_{f \in C \setminus \{e\}} x_f \leq 0 \quad \text{for all } e \in C, \; C \text{ circuit in } G.$$

The graphs G for which every partial matrix (resp. partial distance matrix) A with pattern G for which (7) holds is completable to a distance matrix, are the graphs containing no homeomorph of K_4 as an induced subgraph [La98a] (resp. the graphs that can be made chordal by adding edges in such a way that no new clique of size 4 is created [JJK95]). Note the analogy with the corresponding results for the psd completion problem; some connections between the two problems (EDM) and (PSD) are exposed in [JT95], [La98a].

4 Completion to completely positive and contraction matrices

Call a matrix *doubly nonnegative* if it is psd and entrywise nonnegative. Every completely positive (cp, for short) matrix is obviously doubly nonnegative. The converse implication holds for matrices of order $n \leq 4$ (cf. [GW80]) and for certain patterns of the nonzero entries in A (cf. [KB93]). The cp property is obviously inherited by principal submatrices; call a partial matrix A a *partial cp matrix* if every fully specified principal submatrix of A is cp. It is shown in [DJ98] that every partial cp matrix with graph pattern G is completable to a cp matrix if and only if G is a so-called block-clique graph. A *block-clique graph* being a chordal graph in which any two distinct maximal cliques overlap in at most one node or, equivalently, a chordal graph that does not contain an induced subgraph of the form:



Recall that an $n \times m$ matrix A is a contraction matrix if all eigenvalues of A^*A are less than or equal to 1 or, equivalently, if the matrix

$$\tilde{A} := \begin{pmatrix} I_n & A \\ A^* & I_m \end{pmatrix} \tag{8}$$

is positive semidefinite. Call a partial matrix A a *partial contraction* if all specified submatrices of A are contractions. As every submatrix of a contraction is again a contraction, an obvious necessary condition for a partial matrix A to be completable to a contraction matrix is that A be a partial contraction. Thus arises the question of characterizing the graph patterns G for which every partial contraction with pattern G can be completed to a contraction matrix. As we now deal with rectangular $n \times m$ partial matrices A, their pattern is the bipartite graph G with node set $U \cup V$, where U, V index the rows and columns of A and edges of G correspond to the specified entries of A. We may clearly assume to be dealing with partial matrices whose pattern is a connected graph (as the partial matrices associated with the connected components can be handled separately). Below is an example of a partial matrix A which is a partial contraction, but which is not completable to a contraction matrix:

$$A = \begin{pmatrix} ? & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & ? & \frac{1}{\sqrt{2}} \end{pmatrix}.$$

In fact, the graph pattern displayed in this example is in a sense present in every partial contraction which is not completable to a contraction. Namely, it is shown in [JR86] that the following assertions (i)-(iii) are equivalent for a connected bipartite graph G with node bipartition $U \cup V$: (i) Every partial contraction with pattern G can be completed to a contraction; (ii) G does not contain an induced matching of size 2 (i.e., if $e := uv, e' := u'v'$ are edges in G with $u \neq u' \in U$, $v \neq v' \in V$, then at least one of the pairs uv', $u'v$ is an edge in G; that is, G is *nonseparable* in the terminology of [Go80]); (iii) The graph $\tilde{G}$ obtained from G by adding all edges uu' ($u \neq u' \in V$) and vv' ($v \neq v' \in V$) is chordal. (Note that implication (iii) $\Longrightarrow$ (i) is a consequence of the result on psd completions from [GJSW84] mentioned in Section 2, as $\tilde{G}$ is the graph pattern of the matrix $\tilde{A}$ defined in (8).)

5 Rank completions

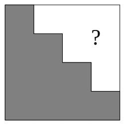
In this section, we consider the problem of determining the possible ranks for the completions of a given partial matrix. For a partial matrix A, let $\mathrm{mr}(A)$ and $\mathrm{MR}(A)$ denote, respectively, the minimum and maximum possible ranks for a completion of A. If B, C are completions of A of respective ranks $\mathrm{mr}(A)$, $\mathrm{MR}(A)$, then changing B into C by changing one entry of B into the corresponding entry of C at a time permits to construct completions realizing all ranks in the range $[\mathrm{mr}(A), \mathrm{MR}(A)]$. Hence, the question is to determine the two extreme values $\mathrm{mr}(A)$ and $\mathrm{MR}(A)$. As we see below, the value $\mathrm{MR}(A)$ can, in fact, be expressed in terms of ranks of fully specified submatrices of A and it can be computed in polynomial time; this constitutes a generalization of the celebrated Frobenius-König theorem (corresponding to the case when specified entries are equal to 0). On the other hand, determining $\mathrm{mr}(A)$ seems to be much a more difficult task.

We first deal with the problem of finding maximum rank completions. Let A be an $n \times m$ partial matrix with graph pattern G, i.e., G is the bipartite graph $(U \cup V, E)$ where U,V index respectively the rows/columns of A and the edges of G correspond to the specified entries of A, and let $\overline{G}$ denote the complementary bipartite graph whose edges correspond to unspecified entries of A. Note that computing $\mathrm{MR}(A)$ amounts to computing the generic rank of A when viewing the unspecified entries of A as independent variables over the field containing the specified entries. For a subset $X \subseteq U \cup V$, let A_X denote the submatrix of A with respective row and column index sets $\{i \in [1,n] \mid u_i \notin X\}$ and $\{j \in [1,m] \mid v_j \notin X\}$. Call X a *cover* of $\overline{G}$ if every edge of $\overline{G}$ has at least one end node in X; that is, if A_X is a fully specified submatrix of A. Clearly, we have: $\mathrm{MR}(A) \le \mathrm{rank}(A_X) + |X|$. In fact, the following equality holds:

$$\mathrm{MR}(A) = \min_{X \text{ cover of } \overline{G}} \mathrm{rank}(A_X) + |X| \tag{9}$$

as shown in [CJRW89]. A determinantal version of the result was given in [HL84]. In the special case when all specified entries of A are equal to 0, then $\mathrm{MR}(A)$ coincides with the maximum cardinality of a matching in $\overline{G}$ and, therefore, the minimax relation (9) reduces to the Frobenius-König's theorem (cf. [LP86] for details on the latter result). Moreover, one can determine $\mathrm{MR}(A)$ and construct a maximum rank completion of A in polynomial time. This was shown in [Mu93] by a reduction to matroid intersection and, more recently, in [Ge] where a simple greedy procedure is presented that solves the problem by perturbing an arbitrary completion.

We now consider minimum rank completions. To start with, note that $\mathrm{mr}(A)$ may depend, in general, on the actual values of the specified entries of A (and not only on the ranks of the specified submatrices of A). Indeed, consider the partial matrix $A = \begin{pmatrix} ? & a & b \\ d & ? & c \\ e & f & ? \end{pmatrix}$ where $a, b, c, d, e, f \neq 0$. Then, $mr(A) = 1$ if $ace = bdf$ and $\mathrm{mr}(A) = 2$ otherwise, while all specified submatrices have rank 1 in both cases. Thus arises the question of identifying the bipartite graphs G for which $\mathrm{mr}(A)$ depends only on the ranks of the specified submatrices of A for every partial matrix A with pattern G; such graphs are called *rank determined.* The graph pattern of the above instance A is the circuit C_6. Hence, C_6 is not rank determined. Call a bipartite graph G *bipartite chordal* if it does not contain a circuit of length ≥ 6 as an induced subgraph. Then, if a bipartite graph is rank determined, it is necessarily bipartite chordal [CJRW89]. It is conjectured there that, conversely, every bipartite chordal graph is rank determined. The conjecture was shown to be true in [Wo87] for the nonseparable bipartite graphs (i.e., the bipartite graphs containing no induced matching of size 2; they are oviously bipartite chordal). Note that a partial matrix A has a nonseparable pattern if and only if it has (up to row/column permutation) the following 'triangular' form:



Then, $\mathrm{mr}(A)$ can be explicitely formulated in terms of the ranks of the specified submatrices of A; in the simplest case, the formula for $\mathrm{mr}(A)$ reads:

$$\mathrm{mr} \begin{pmatrix} B & ? \\ C & D \end{pmatrix} = \mathrm{rank} \begin{pmatrix} B \\ C \end{pmatrix} + \mathrm{rank} \begin{pmatrix} C & D \end{pmatrix} - \mathrm{rank}(C).$$

It is shown in [CJRW89] that the above conjecture holds when the pattern G is a path, or when G is obtained by 'gluing' a collection of circuits of length 4 along a common edge.

References

[AHMR88] J. Agler, J.W. Helton, S. McCullough, and L. Rodman. Positive semidefinite matrices with a given sparsity pattern. *Linear Algebra and its Applications*, 107:101–149, 1988.

[AKW97] A.Y. Alfakih, A. Khandani, and H. Wolkowicz. Solving Euclidean distance matrix completion problems via semidefinite programming. Technical Report CORR 97-9, University of Waterloo, 1997. To appear in *Computational Optimization and Applications.*

[Al95] F. Alizadeh. Interior point methods in semidefinite programming with applications in combinatorial optimization. *SIAM Journal on Optimization*, 5:13–51, 1995.

[BJ95] M. Bakonyi and C.R. Johnson. The Euclidian distance matrix completion problem. *SIAM Journal on Matrix Analysis and Applications*, 16:646–654, 1995.

[BM86] F. Barahona and A.R. Mahjoub. On the cut polytope. *Mathematical Programming*, 36:157–173, 1986.

[BJL96] W.W. Barrett, C.R. Johnson, and R. Loewy. The real positive definite completion problem: cycle completability. *Memoirs of the American Mathematical Society*, Number 584, 69 pages, 1996.

[BJL89] W.W. Barrett, C.R. Johnson, and M. Lundquist. Determinantal formulaes for matrix completions associated with chordal graphs. *Linear Algebra and its Applications*, 121:265–289, 1989.

[BJT93] W. Barrett, C.R. Johnson, and P. Tarazaga. The real positive definite completion problem for a simple cycle. *Linear Algebra and its Applications*, 192:3–31, 1993.

[Ba93] A.I. Barvinok. Feasibility testing for systems of real quadratic equations. *Discrete and Computational Geometry*, 10:1–13, 1993.

[Ba95] A.I. Barvinok. Problems of distance geometry and convex properties of quadratic maps. *Discrete and Computational Geometry*, 13:189–202, 1995.

[Bl53] L.M. Blumenthal. *Theory and Applications of Distance Geometry.* Oxford University Press, Oxford, 1953. (Second edition: Chelsea, New York, 1970.)

[CJRW89] N. Cohen, C.R. Johnson, L. Rodman, and H.J. Woederman. Ranks of completions of partial matrices. In H. Dym et al., editors, *The Gohberg Anniversary Collection, Volume I*, volume 40 of *Operator Theory: Advances and Applications*, pages 165–185. Birkhäuser Verlag, Basel, 1989.

[CH88] G.M. Crippen and T.F. Havel. *Distance Geometry and Molecular Conformation.* Research Studies Press, Taunton, Somerset, Engeland, 1988.

[DL97] M. Deza and M. Laurent. *Geometry of Cuts and Metrics.* Vol. 15 of *Algorithms and Combinatorics*, Springer-Verlag, Berlin, 1997.

[DJ98] J.H. Drew and C.R. Johnson. The completely positive and doubly nonnegative completion problems. *Linear and Multilinear Algebra*, 44:85–92, 1998.

[DG81] H. Dym and I. Gohberg. Extensions of band matrices with band inverses. *Linear Algebra and its Applications*, 36:1–24, 1981.

[ELG88] R.L. Ellis, D.C. Lay, and I. Gohberg. On negative eigenvalues of selfadjoint extensions of band matrices. *Linear and Multilinear Algebra*, 24:15–25, 1988.

[Ge] J. Geelen. Maximum rank matrix completion. *Linear Algebra and its Applications*, to appear.

[GHR93] W. Glunt, T.L. Hayden, and M. Raydan. Molecular conformations from distance matrices. *Journal of Computational Chemistry*, 14:114–120, 1993.

[Go97] M.X. Goemans. Semidefinite programming in combinatorial optimization. *Mathematical Programming*, 79: 143–161, 1997.

[Go80] M.C. Golumbic. *Algorithmic Theory and Perfect Graphs.* Academic Press, New York, 1980.

[GW80] L.J. Gray and D.G. Wilson. Nonnegative factorization of positive semidefinite nonnegative matrices. *Linear Algebra and its Applications*, 31:119–127, 1980.

[GJSW84] R. Grone, C.R. Johnson, E.M. Sá, and H. Wolkowicz. Positive definite completions of partial hermitian matrices. *Linear Algebra and its Applications*, 58:109–124, 1984.

[GLS88] M. Grötschel, L. Lovász and A. Schrijver. *Geometric Algorithms and Combinatorial Optimization*, Springer-Verlag, Berlin, 1988.

[HL84] D.J. Hartfiel and R. Loewy. A determinantal version of the Frobenius-König theorem. *Linear and Multilinear Algebra*, 16:155–165, 1984.

[Ha91] T.F. Havel. An evaluation of computational strategies for use in the determination of protein structure from distance constraints obtained by nuclear magnetic resonance. *Progress in Biophysics and Biophysical Chemistry*, 56:43–78, 1991.

[HRVW96] C. Helmberg, F. Rendl, R.J. Vanderbei, and H. Wolkowicz. An interior-point method for semidefinite programming. *SIAM Journal on Optimization*, 6:342–361, 1996.

[HPR89] J.W. Helton, S. Pierce, and L. Rodman. The ranks of extremal positive semidefinite matrices with given sparsity pattern. *SIAM Journal on Matrix Analysis and its Applications*, 10:407–423, 1989.

[He90] B. Hendrickson. *The Molecule Problem: Determining Conformation from Pairwise Distances.* Ph.D. thesis, Technical Report 90-1159, Cornell University, Department of Computer Science, Ithaca, NY, 1990.

[He92] B. Hendrickson. Conditions for unique graph realizations. *SIAM Journal on Computing*, 21:65–84, 1992.

[He95] B. Hendrickson. The molecule problem: exploiting structure in global optimization. *SIAM Journal on Optimization*, 5:835–857, 1995.

[Jo90] C.R. Johnson. Matrix completion problems: a survey. In C.R. Johnson, editor, *Matrix Theory and Applications*, volume 40 of *Proceedings of Symposia in Applied Mathematics*, pages 171–198. American Mathematical Society, Providence, Rhode Island, 1990.

[JJK95] C.R. Johnson, C. Jones, and B. Kroschel. The distance matrix completion problem: cycle completability. *Linear and Multilinear Algebra*, 39:195–207, 1995.

[JKW98] C.R. Johnson, B. Kroschel, and H. Wolkowicz. An interior-point method for approximate positive semidefinite completions. *Computational Optimization and Applications*, 9:175–190, 1998.

[JR84] C.R. Johnson and L. Rodman. Inertia possibilities for completions of partial Hermitian matrices. *Linear and Multilinear Algebra*, 16:179–195, 1984.

[JR86] C.R. Johnson and L. Rodman. Completion of matrices to contractions. *Journal of Functional Analysis*, 69:260–267, 1986.

[JR88] C.R. Johnson and L. Rodman. Chordal inheritance principles and positive definite completions of partial matrices over function rings. In I. Gohberg et al., editors, *Contributions to Operator Theory and its Applications*, volume 35 of *Operator Theory: Advances and Applications*, pages 107–127. Birkhäuser Verlag, Basel, 1988.

[JT95] C.R. Johnson and P. Tarazaga. Connections between the real positive semidefinite and distance matrix completion problems. *Linear Algebra and its Applications*, 223/224:375–391, 1995.

[KP97] L. Khachiyan and L. Porkolab. Computing integral points in convex semi-algebraic sets. In 38th Annual Symposium on Foundations of Computer Science, pages 162–171, 1997.

[KB93] N. Kogan and A. Berman. Characterization of completely positive graphs. *Discrete Mathematics*, 114:297–304, 1993.

[KTO93] I.D. Kuntz, J.F. Thomason, and C.M. Oshiro. Distance geometry. *Methods in Enzymologie*, 177:159–204, 1993.

[La70] G. Laman. On graphs and rigidity of plane skeletal structures. *Journal of Engineering Mathematics*, 4:331–340, 1970.

[La97a] M. Laurent. The real positive semidefinite completion problem for series-parallel graphs. *Linear Algebra and its Applications*, 252:347–366, 1997.

[La97b] M. Laurent. Cuts, matrix completions and graph rigidity. *Mathematical Programming*, 79: 255–283, 1997.

[La98a] M. Laurent. A connection between positive semidefinite and Euclidean distance matrix completion problems. *Linear Algebra and its Applications*, 273: 9–22, 1998.

[La98b] M. Laurent. A tour d'horizon on positive semidefinite and Euclidean distance matrix completion problems. In P. Pardalos and H. Wolkowicz, editors, *Topics in Semidefinite and Interior-Point Methods,* vol. 18 of *The Fields Institute for research in Mathematical Science, Communications series*, pages 51–76. Providence, Rhode Island, 1998.

[La98c] M. Laurent. On the order of a graph and its deficiency in chordality. CWI Report PNA-R9801, CWI, Amsterdam, 1998.

[La98d] M. Laurent. Polynomial instances of the positive semidefinite and Euclidean distance matrix completion problems. Preprint, 1998.

[LH82] J. de Leeuw and W. Heiser. Theory of multidimensional scaling. In P.R. Krishnaiah and L.N. Kanal, editors, *Handbook of Statistics*, Vol. 2, pages 285–316. North Holland, 1982.

[LP86] L. Lovász and M.D. Plummer. *Matching Theory.* Akadémiai Kiadó, Budapest, 1986.

[LY82] L. Lovász and Y. Yemini. On generic rigidity in the plane. *SIAM Journal on Algebraic and Discrete Methods*, 3:91–98, 1982.

[LJ91] M.E. Lundquist and C.R. Johnson. Linearly constrained positive definite completions. *Linear Algebra and its Applications*, 150:195–207, 1991.

[MW96] J.J. Moré and Z. Wu. ϵ-optimal solutions to distance geometry problems via global continuation. P.M. Pardalos and D. Shalloway, editors, vol. 23 of *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, pages 151–168. American Mathematical Society, 1996.

[MW97] J.J. Moré and Z. Wu. Global continuation for distance geometry problems. *SIAM Journal on Optimization*, 7:814–836, 1997.

[Mu93] K. Murota. Mixed matrices: Irreducibility and decomposition. In R.A. Brualdi et al., editors, *Combinatorial and Graph-Theoretical Problems in Linear Algebra*, volume 50 of *The IMA volumes in Mathematics and its Applications*, pages 39–71. Springer-Verlag, New York, 1993.

[NN94] Y.E. Nesterov and A.S. Nemirovsky. *Interior Point Polynomial Algorithms in Convex Programming: Theory and Algorithms.* SIAM Publications, Philadelphia, 1994.

[PL97] P.M. Pardalos and X. Lin. A tabu based pattern search method for the distance geometry problem. In F. Giannessis et al., editors, *Mathematical Programming*, Kluwer Academic Publishers, 1997.

[PPS89] V.I. Paulsen, S.C. Power, and R.R. Smith. Schur products and matrix completions. *Journal of Functional Analysis*, 85:151–178, 1989.

[PK97] L. Porkolab and L. Khachiyan. On the complexity of semidefinite programs. *Journal of Global Optimization*, 10:351–365, 1997.

[Ra97] M.V. Ramana. An exact duality theory for semidefinite programming and its complexity implications. *Mathematical Programming*, 77:129–162, 1997.

[Ro70] D.J. Rose. Triangulated graphs and the elimination process. *Journal of Mathematical Analysis and Applications*, 32:597–609, 1970.

[Sa79] J.B. Saxe. Embeddability of weighted graphs in k-space is strongly NP-hard. Proceedings of the 17th Allerton Conference in Communications, Control and Computing, pages 480–489, 1979.

[Sc35] I.J. Schoenberg. Remarks to M.Fréchet's article "Sur la définition axiomatique d'une classe d'espaces vectoriels distanciés applicables vectoriellement sur l'espace de Hilbert". *Annals of Mathematics*, 36:724–732, 1935.

[Wo87] H.J. Woerdeman. The lower order of lower triangular operators and minimal rank extensions. *Integral Equations and Operator Theory*, 10:859–879, 1987.

[Ye79] Y. Yemini. Some theoretical aspects of position-location problems. *Proceedings of the 20th Annual Symposium on Foundations of Computer Science*, pages 1–8, 1979.

List of keywords

partial matrix: A partial matrix is a matrix whose entries are specified only on a subset of its positions.

matrix completion problem: This is the problem of determining whether the unspecified entries in a given partial matrix can be determined in order to obtain a full conventional matrix satisfying a prescribed property (e.g., being positive semidefinite, or having a certain rank).

the graph realization problem: Given integers $k, n \geq 1$ and a set of prescribed distances among certain pairs of points in $[1, n]$, determine whether one can find n vectors in the k-dimensional space for which the Euclidean distances among the prescribed pairs are the prescribed distances.

positive (semi)definite matrix: A complex Hermitian or real symmetric matrix A is said to be positive semidefinite if $x^*Ax \geq 0$ for all vectors x and positive definite if $x^*Ax > 0$ for all vectors $x \neq 0$.

completely positive matrix: A real matrix A is said to be completely positive if it can be decomposed as $A = BB^T$ for some nonnegative matrix B.

doubly nonnegative matrix: A matrix is said to be doubly nonnegative if it is positive semidefinite and entrywise nonnegative.

Euclidean distance matrix: An $n \times n$ real symmetric matrix $D = (d_{ij})$ is said to be a Euclidean distance matrix if there exist vectors $v_1, \ldots, v_n \in \mathbb{R}^k$ (for some $k \geq 1$) such that, for all $i, j = 1, \ldots, n$, d_{ij} is equal to the Euclidean distance between v_i and v_j.

contraction matrix: A complex rectangular matrix A is a contraction matrix if the largest eigenvalue of matrix A^*A is less than or equal to 1.

chordal graph: A graph G is said to be chordal if it does not contain any circuit of length ≥ 4 as an induced subgraph.

block-clique graph: A graph G is called a block-clique graph if G is chordal and any two distinct maximal cliques in G have at most one node in common.

minimum fill-in: The minimum fill-in of a graph G is the minimum number of edges needed to be added to G in order to obtain a chordal graph.

separable bipartite graph: A bipartite graph G with node bipartition $U \cup V$ is said to be separable if it contains an induced matching of cardinality 2; that is, if there exist two edges uv, $u'v'$ in G where $u \neq u' \in U$, $v \neq v' \in V$ and the pairs uv' and $u'v$ are not edges in G.

bipartite chordal graph: A bipartite graph is said to be bipartite chordal if it does not contain any circuit of length ≥ 6 as an induced subgraph. (Hence, every nonseparable bipartite graph is bipartite chordal.)

semidefinite programming

polynomial algorithm

NP-hard problem