# Solving Sparse Semidefinite Programs by Matrix Completion (Part I)

情報理工学研究科 福田 光浩 (Mituhiro Fukuda) 東京工業大学 工学系研究科 中田 和秀 (Kazuhide Nakata) 東京大学 藤沢 克樹 工学研究科 (Katsuki Fujisawa) 京都大学 情報理工学研究科 小島 政和 (Masakazu Kojima) 東京工業大学 数理解析研究所 室田 一雄 京都大学 (Kazuo Murota)

#### 1 Introduction

In recent year, Semidefinite Programs (SDPs) have been employed in several fields such as in system and control theory, finance theory, architecture, etc., as well as they have been utilized as relaxations of other difficult problems such as combinatorial problems, quadratic programs, etc.. Frequently, the SDP formulation of these problems becomes large-scaled and sparse. The dual interior-point method [1] is preferred to solve some particular class of above problems, since it does not consider the full dense primal positive definite matrix variable. However, this method lacks in accuracy and reliability of the optimal solution if compared to the primal-dual counterpart. The main topic of this article is to resolve this disadvantage of the primal-dual interior-point method for large-scaled and sparse SDPs by applying the matrix completion theory [8] on the primal positive definite matrix variable.

In the following lines, we describe a basic idea which leads to our new algorithms. More details of the present article can be found in [4].

Let  $S^n$  denote the space of  $n \times n$  symmetric matrices with the Frobenius inner product  $X \bullet Y = \sum_{i=1}^n \sum_{j=1}^n X_{ij} Y_{ij}$  for X,  $Y \in S^n$ . We will use the notation  $X \in S^n_+$  to designate that  $X \in S^n$  is positive semidefinite. Given  $A_p \in S^n$  (p = 0, 1, ..., m) and  $b \in \mathbb{R}^m$ , the standard equality form SDP is formulated as

minimize 
$$A_0 \bullet X$$
  
subject to  $A_p \bullet X = b_p \ (p = 1, 2, ..., m), \ X \in \mathcal{S}_+^n,$  (1)

and its dual as

maximize 
$$\sum_{p=1}^{m} b_p z_p$$
subject to 
$$\sum_{p=1}^{m} \mathbf{A}_p z_p + \mathbf{Y} = \mathbf{A}_0, \ \mathbf{Y} \in \mathcal{S}_+^n.$$

We introduce the aggregate sparsity pattern E of the data matrices by

$$E = \{(i, j) \in V \times V : [\mathbf{A}_p]_{ij} \neq 0 \text{ for some } p \in \{0, 1, 2, \dots, m\}\}.$$
 (3)

Here V denotes the set  $\{1, 2, ..., n\}$  of row/column indices of the data matrices  $A_0, A_1, ..., A_m$ , and  $[A_p]_{ij}$  denotes the (i, j)th entry of  $A_p \in S^n$ . It is convenient in the forthcoming discussion to identify the aggregate sparsity pattern E with the aggregate sparsity pattern E matrix E having unspecified nonzero numerical values in E.

Assume that a collection of nonempty subsets  $C_1, C_2, \ldots, C_\ell$  of V satisfies the following two conditions:

(i) 
$$E \subseteq F \equiv \bigcup_{r=1}^{\ell} C_r \times C_r$$
.

(ii) Any partial symmetric matrix X with entries  $X_{ij} = \bar{X}_{ij} \in \mathbb{R}$   $((i,j) \in F)$  has a positive semidefinite matrix completion (i.e., given any  $\bar{X}_{ij} \in \mathbb{R}$   $((i,j) \in F)$ , there exists a positive semidefinite  $X \in \mathcal{S}^n$  such that  $X_{ij} = \bar{X}_{ij} \in \mathbb{R}$   $((i,j) \in F)$ ) if and only if the submatrices  $\bar{X}_{C_rC_r}$   $(r = 1, 2, ..., \ell)$  are all positive semidefinite.

From condition (i), we observe that values of the objective and constraint linear functions  $A_p \bullet X$  (p = 0, 1, ..., m) involved in the SDP (1) are completely determined by values of entries  $X_{ij}$   $((i, j) \notin F)$  and independent of values of entries  $X_{ij}$   $((i, j) \notin F)$ . The remaining entries  $X_{ij}$   $((i, j) \notin F)$  only affect whether X is positive semidefinite. Now we know by condition (ii) whether we can assign some appropriate values to those remaining entries  $X_{ij}$   $((i, j) \notin F)$  so that the resulting whole matrix X becomes positive semidefinite. Therefore the SDP (1) becomes equivalent to

minimize 
$$\sum_{\substack{(i,j)\in F\\\text{subject to}}} [\boldsymbol{A}_0]_{ij} X_{ij}$$
subject to 
$$\sum_{\substack{(i,j)\in F\\\boldsymbol{X}_{C_rC_r} \in \mathcal{S}_+^{C_r} \ (r=1,2,\ldots,\ell).}} [\boldsymbol{A}_p]_{ij} X_{ij} = b_p \ (p=1,2,\ldots,m),$$

$$(4)$$

Here  $\mathcal{S}_{+}^{C_r}$  denotes the set of  $\sharp C_r \times \sharp C_r$  positive semidefinite symmetric matrices with entries specified in  $C_r \times C_r$ , and  $\sharp C_r$  the number of elements of  $C_r$ .

In this article, we will give general theoretical results to obtain the sets  $C_1, C_2, \ldots, C_\ell \subseteq V$  which satisfies (i) and (ii). In fact, a necessary and a sufficient condition to guarantee (i) and (ii) for a given SDP. These results are closely related to chordal graphs, Cholesky factorizations, and minimal fill-in [4] (section 2.2). Once we have determined these sets, we will give explicit formulae to obtain a semidefinite completion  $X \in S^n$  of a symmetric matrix which has only the entries  $X_{ij} \in \mathbb{R}$   $((i,j) \in F)$  specified (section 2.3). These results allow us to propose two kinds of new algorithms based on primal-dual interior-point method which we call the conversion method and the completion method. In section 3, we describe a general procedure to obtain the sets  $C_1, C_2, \ldots, C_\ell$ , and to meet the conditions given in the theorems of section 2 in practice. The conversion method will be exposed in section 4, however, we leave the description of the completion method, as well as the comparative analysis between these two methods and the numerical experiments for the part II of this article.

## 2 Chordal graph and positive semidefinite matrix completion

#### 2.1 Notation

- $S^n(F,?)$ : the set of  $n \times n$  partial symmetric matrices with entries specified in F;
- $S^n(F,0)$ : the set of  $n \times n$  symmetric matrices with vanishing entries outside F; *i.e.*,  $S^n(F,0) = \{ \mathbf{X} \in S^n : X_{ij} = 0 \text{ if } (i,j) \notin F \};$
- $V = \{1, 2, ..., n\}$  and for  $E, F \subseteq V \times V$  in general, we define  $F^{\circ} = F \setminus \{(i, i) : i = 1, 2, ..., n\}$  and  $E^{\bullet} = E \cup \{(i, i) : i = 1, 2, ..., n\}$ ;
- $S^C$ .  $S^C_+$ ,  $S^C_{++}$ : the sets of  $\sharp C \times \sharp C$  symmetric matrices, positive semidefinite symmetric matrices, positive definite symmetric matrices, respectively, with rows and columns indexed by  $C \subseteq V$ , where  $\sharp C$  means the number of elements of C.

#### 2.2 Chordal graph

We denote by G(V, E) an undirected graph with the vertex set V and the edge set  $E \subseteq V \times V$ . It is assumed throughout this paper that a graph has no self-loops.

Notice that it is natural to consider graphs when we are concerned with the structure of sparse matrices, *i.e.*, we can make an one-to-one correspondence between a graph G(V, E) and a partial symmetric matrix  $\bar{X} \in \mathcal{S}^n(E^{\bullet},?)$ .

A graph G(V, E) is said to be *chordal* if every cycle of length  $\geq 4$  has a chord (an edge joining two nonconsecutive vertices of the cycle). Chordal graphs have been studied extensively in many different contexts [2, 5, 7].

Suppose now that  $\bar{X} \in \mathcal{S}^n(E^{\bullet}, 0)$ . An ordering of the rows/columns of  $\bar{X}$  (and therefore, of the vertices V of the graph associated with  $\bar{X}$ ) is called a *perfect elimination ordering* if it does not cause any fill-in when we perform a symbolic Cholesky factorization in  $\bar{X}$  according to this ordering.

Now, we give another characterization of chordal graphs.

**Theorem 2.1** [5] A graph is chordal if and only if it has a perfect elimination ordering.

Let us denote the family of maximal cliques of a graph G(V, E) by  $\{C_r \subseteq V : r = 1, 2, ..., \ell\}$ . It is known that if the graph is chordal, then the number of maximal cliques  $\ell$  is bounded by n. Also, in this case, these maximal cliques can be indexed in such a way that for each  $r = 1, 2, ..., \ell - 1$  it holds that

$$\exists s \ge r+1: \quad C_r \cap (C_{r+1} \cup C_{r+2} \cup \dots \cup C_\ell) \subsetneq C_s. \tag{5}$$

The property (5) is called the running intersection property.

#### 2.3 Positive (semi)definite matrix completion

Suppose we are given a partial symmetric matrix  $\bar{X} \in \mathcal{S}^n(F,?)$ , and let G(V,E) be the associated graph, where  $E = F^{\circ}$ . Denote by  $\{C_r \subseteq V : r = 1, 2, ..., \ell\}$  the family of all maximal cliques of G(V,E). An obvious necessary condition for  $\bar{X}$  to have a positive semidefinite matrix completion is that each  $\bar{X}_{C_rC_r}$  is positive semidefinite, *i.e.*,

$$\bar{X}_{C_r C_r} \in \mathcal{S}_+^{C_r} \qquad (r = 1, 2, \dots, \ell),$$
 (6)

where it is noted that all the entries of the submatrix  $\bar{X}_{C_rC_r}$  are specified. Similarly, an obvious necessary condition for  $\bar{X}$  to have a positive definite matrix completion is that each  $\bar{X}_{C_rC_r}$  is positive definite, *i.e.*,

$$\bar{X}_{C_r C_r} \in \mathcal{S}_{++}^{C_r} \qquad (r = 1, 2, \dots, \ell).$$
 (7)

We refer to (6) and (7) as the *clique-PSD condition* and the *clique-PD condition*, respectively.

The following two theorems are most fundamental concerning the positive (semi)definite matrix completion problem.

**Theorem 2.2** [7] Let G(V, E) be a graph.

- (i) Any partial symmetric matrix  $\bar{X} \in \mathcal{S}^n(E^{\bullet},?)$  satisfying the clique-PSD condition (6) can be completed to a positive semidefinite symmetric matrix X if and only if G(V,E) is chordal.
- (ii) Any partial symmetric matrix  $\bar{X} \in \mathcal{S}^n(E^{\bullet},?)$  satisfying the clique-PD condition (7) can be completed to a positive definite symmetric matrix X if and only if G(V,E) is chordal.

**Theorem 2.3** [7] Suppose that a partial symmetric matrix  $\bar{X} \in \mathcal{S}^n(F,?)$  has a positive definite matrix completion. Then there exists a unique positive definite matrix completion  $X = \hat{X}$  that maximizes the determinant, i.e., such that

$$\det(\hat{X}) = \max\{\det(X) : X \text{ is a positive definite matrix completion of } \bar{X}\}.$$

Moreover, such  $\hat{\boldsymbol{X}}$  is characterized by the condition:

$$[\hat{\boldsymbol{X}}^{-1}]_{ij} = 0 \quad ((i,j) \notin F), \quad i.e., \quad \hat{\boldsymbol{X}}^{-1} \in \mathcal{S}^n(F,0).$$

We refer to the completion  $\hat{X}$  in Theorem 2.3 as the maximum-determinant positive definite matrix completion of  $\bar{X}$ .

Finally, the following result plays an important role in the *completion method* which will be discussed in detail in the part II of this article.

**Lemma 2.4** [4] Let G(V, E) be a chordal graph, and  $\bar{X} \in \mathcal{S}^n(E^{\bullet},?)$  be a partial symmetric matrix satisfying the clique-PD condition (7). Let  $\mathbf{P}$  be a permutation matrix representing a perfect elimination ordering of G(V, E) which is also consistent with the running intersection property in such a way that (1, 2, ..., n) is a perfect elimination ordering for  $\mathbf{P}\bar{X}\mathbf{P}^T$ . Then the maximum-determinant positive definite matrix completion  $\hat{X}$  of  $\bar{X}$  can be expressed in terms of the sparse clique-factorization formula

$$P\hat{X}P^{T} = L_1^{T}L_2^{T}\cdots L_{\ell-1}^{T}DL_{\ell-1}\cdots L_2L_1,$$
(8)

where  $L_r$   $(r = 1, 2, ..., \ell - 1)$  are triangular matrices and D is a positive definite block-diagonal matrix consisting of  $\ell$  diagonal blocks defined as follows:

$$S_r = C_r \setminus (C_{r+1} \cup C_{r+2} \cup \cdots \cup C_{\ell}) \quad (r = 1, 2, \dots, \ell),$$
  

$$U_r = C_r \cap (C_{r+1} \cup C_{r+2} \cup \cdots \cup C_{\ell}) \quad (r = 1, 2, \dots, \ell),$$

$$[\mathbf{L}_r]_{ij} = \begin{cases} 1 & (i=j) \\ [\bar{\mathbf{X}}_{U_r U_r}^{-1} \bar{\mathbf{X}}_{U_r S_r}]_{ij} & (i \in U_r, j \in S_r) \\ 0 & (otherwise) \end{cases}$$
(9)

for  $r = 1, 2, ..., \ell - 1$ , and

$$D = \begin{pmatrix} D_{S_1S_1} & & & \\ & D_{S_2S_2} & & \\ & & \ddots & \\ & & & D_{S_{\ell}S_{\ell}} \end{pmatrix}$$
(10)

with

$$\mathbf{D}_{S_r S_r} = \begin{cases} \bar{\mathbf{X}}_{S_r S_r} - \bar{\mathbf{X}}_{S_r U_r} \bar{\mathbf{X}}_{U_r U_r}^{-1} \bar{\mathbf{X}}_{U_r S_r} & (r = 1, 2, \dots, \ell - 1), \\ \bar{\mathbf{X}}_{S_\ell S_\ell} & (r = \ell). \end{cases}$$
(11)

Due to the above lemma, the inverse of the maximum-determinant positive definite matrix completion  $\hat{X}$  can be expressed in the following form:

$$P\hat{X}^{-1}P^T = WW^T \tag{12}$$

where  $\mathbf{W} \in \mathcal{S}^n(E^{\bullet}, 0)$  is a sparse lower triangular matrix.

## 3 Chordal extension of aggregate sparsity pattern

It is not true in general that given sparse SDPs (1) and (2), the graph  $G(V, E^{\circ})$  defined through the aggregate sparsity pattern E is chordal. Therefore, we can not straightforward utilize the results presented in the previous section to obtain a maximum-determinant positive definite matrix completion of the primal matrix variable  $X \in \mathcal{S}^{n}(E,?)$ .

In this section, we discuss briefly how to obtain a *chordal extension* of a given graph  $G(V, E^{\circ})$ , *i.e.*, a chordal graph  $G(V, F^{\circ})$  such that  $F \supseteq E$ . In the succeeding discussion, we often call the set F as the chordal extension or simply the *extended sparsity pattern* of the aggregate sparsity pattern E.

As we have seen in the previous section, the chordal extension is closely related to the Cholesky factorization. Specifically, the chordal extension that minimizes the total number of edges in  $G(V, F^{\circ})$  is obtained via the Cholesky factorization of the aggregate sparsity pattern matrix A with the minimum fill-in. Therefore it seems reasonable (or at least attractive) in practice to employ various existing heuristic methods, such as the minimum degree ordering for less fill-in, the (nested) dissection ordering for less fill-in or the reverse Cuthill-McKee ordering for reducing bandwidth developed for the Cholesky factorization [6] for this purpose.

Given SDPs (1) and (2), we can construct the desired chordal extension in the following way. Determine the aggregate sparsity pattern matrix A of the SDP, and utilize one of the heuristic algorithms cited above to obtain a row/column ordering which may cause fill-in as less as possible. Then, performing a symbolic Cholesky factorization according to the new ordering to A, we obtain a chordal extension  $G(V, F^{\circ})$  by Theorem 2.1, where F is the extended sparsity pattern corresponding to the nonzero entries of the symbolic Cholesky factorization.

Observe that the ordering of the vertices  $(v_1, v_2, \ldots, v_n)$  we obtained by the heuristic algorithm is a perfect elimination ordering for G(V, F). Let us denote the set of the vertices adjacent to  $v \in V$  by  $Adj(v) = \{u \in V : (v, u) \in F^{\circ}\}.$ 

Now, it can be shown using induction hypothesis that the following algorithm determines all the maximal cliques of  $G(V, F^{\circ})$ .

Once we run this algorithm, we obtain a rooted tree whose nodes are the maximal cliques of G(V, F). This tree is called *clique tree*, and has some nice properties concerning with its structure [2]. After that, the running intersection property can be determined indexing first a leaf of this tree and then removing it from the tree, and so on successively. An ordering of the maximal cliques satisfying the running intersection property (5) induces a perfect elimination ordering of the vertices. Note first that  $S_1 = C_1 \setminus (C_2 \cup C_3 \cup \cdots \cup C_\ell)$  is nonempty. Then we can start a perfect elimination ordering by numbering the vertices in  $S_1$  with  $1, 2, \cdots, |S_1|$ . For each  $r = 1, 2, \ldots, \ell$  in general we number the vertices in  $S_r = C_r \setminus (C_{r+1} \cup \cdots \cup C_\ell)$  with  $\sum_{s=1}^{r-1} |S_s| + 1, \sum_{s=1}^{r-1} |S_s| + 2, \cdots, \sum_{s=1}^{r-1} |S_s| + |S_r|$ . We can thus obtain a perfect elimination ordering of the vertices, in which the vertices in  $S_r$  are given consecutive numbers for each r.

### 4 The conversion method

In the Introduction, we have shown that the SDP (1) is equivalent to the problem (4). This problem involves less variables and smaller size positive semidefinite constraints than the original SDP (1). This feature certainly makes the conversion attractive in practice because such a problem is expected to be solved easier. It should be noted, however, that two distinct positive semidefinite constraints  $X_{C_rC_r} \in \mathcal{S}_+^{C_r}$  and  $X_{C_sC_s} \in \mathcal{S}_+^{C_s}$  in (4) share variables  $X_{ij}$  ( $(i,j) \in (C_r \cap C_s) \times (C_r \cap C_s)$ ) whenever  $C_r \cap C_s \neq \emptyset$ . Hence the problem is not a standard SDP. In this section, we show how to convert the problem to a standard

SDP to which we can apply interior-point methods, and discuss some advantages and disadvantages of the resulting SDP.

For every  $r = 1, 2, \dots, \ell$ , let

$$E_r = \{(i,j) \in C_r \times C_r : (i,j) \in C_s \times C_s \text{ for some } s < r \}.$$

By definition,  $E_1 = \emptyset$ , and if  $(i, j) \in E_r$  then the positive semidefinite constraint  $X_{C_rC_r} \in \mathcal{S}^{C_r}_+$  shares variables  $X_{ij}$   $((i, j) \in E_r)$  with the positive semidefinite constraint  $X_{C_sC_s} \in \mathcal{S}^{C_s}_+$  for some s < r. To make such a pair of dependent positive semidefinite constraints independent, we introduce auxiliary variables  $U^r_{ij}$   $((i, j) \in E_r, r = 2, 3, ..., \ell)$ , and we rewrite the constraint (4) as

$$\sum_{\substack{(i,j)\in F\\ U_{ij}^r = X_{ij} \ ((i,j) \in E_r, \ i \ge j, \ r = 2, 3, \dots, \ell),}} [A_p]_{ij} X_{ij} = b_p \ (p = 1, 2, \dots, m),$$

$$U_{ij}^r = X_{ij} \ ((i,j) \in E_r, \ i \ge j, \ r = 2, 3, \dots, \ell),$$

$$X^r \in \mathcal{S}_+^{C_r} \ (r = 1, 2, \dots, \ell),$$
(13)

where

$$[\mathbf{X}^r]_{ij} = \left\{ egin{array}{ll} U_{ij}^r & ext{if } (i,j) \in E_r, \\ X_{ij} & ext{otherwise.} \end{array} \right.$$

Then we may regard the minimization of the objective function  $\sum_{(i,j)\in F} [A_0]_{ij} X_{ij}$  over the constraint (13) as a standard SDP. In fact, if we further introduce a block-diagonal symmetric matrix variable of the form

$$X' = \left(egin{array}{cccc} X^1 & O & O & \cdots & O \ O & X^2 & O & \cdots & O \ dots & dots & dots & \ddots & dots \ O & O & O & \cdots & X^\ell \end{array}
ight),$$

and if we appropriately rearrange all the coefficients of the linear equality constraints in (13) and the objective function  $\sum_{(i,j)\in F} [A_0]_{ij} X_{ij}$  to reconstruct data matrices with the same block-diagonal structure as X', we obtain an standard equality form SDP.

There are two major advantages of this conversion. First, when the sizes of all positive semidefinite matrix variables in (13) are small, their Cholesky factorizations, computation of their minimum eigenvalues, and matrix multiplications require less CPU time than those of the original positive semidefinite matrix variable X in (1). Second, once we have converted the SDP (1) into the SDP with the block-diagonal positive semidefinite matrix variable X', we can apply effectively any interior-point method incorporating a block-diagonal matrix data structure ([3, etc.]) for SDPs.

We should note, however, that the conversion above from the SDP (1) to the SDP with the block-diagonal symmetric matrix variable X' (13) increases the number of equality constraints from m to the number

$$m' = m + \sum_{r=2}^{\ell} \sharp \{(i,j) \in E_r : i \ge j\}.$$

When we apply interior-point methods to a standard form SDP having m equality constraints, we solve a system of linear equations with a fully dense  $m \times m$  coefficient matrix B to generate a search direction at each iteration. This requires  $\mathcal{O}(m^3)$  arithmetic operations. So the increase in the number of equality constraints in the converted problem may worsen the total computational efficiency. Therefore the reduction in the sizes of positive semidefinite matrix variables should be properly balanced with the increase in the number of equality constraints in (13) when we choose a chordal extension  $G(V, F^{\circ})$  of  $G(V, E^{\circ})$ .

## 5 Concluding remarks

In this article, we described a novel technique to explore the sparsity of SDPs utilizing the ideas of positive definite matrix completion. In particular, the key results are closed related with chordal graphs with were studied in '60 when solving large-scale sparse linear system of equations. We proposed two method to solve sparse SDPs, the *conversion method* and the *completion method*. We leave for the part II of this article the complete description of the completion method as well as the comparative analysis between these two methods and their numerical experiments.

#### References

- [1] S. J. Benson, Y. Ye and X. Zhang, Solving large-scale sparse semidefinite programs for combinatorial optimization, SIAM J. Optim., 10 (2000), pp. 443–461.
- [2] J. R. S. Blair and B. Peyton, An introduction to chordal graphs and clique trees, in Graph Theory and Sparse Matrix Computation, A. George, J. R. Gilbert and J. W. H. Liu, eds., Springer-Verlag, New York, 1993, pp. 1–29.
- [3] K. Fujisawa, M. Kojima and K. Nakata, SDPA (Semidefinite Programming Algorithm) User's Manual —, Technical Report B-308, Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, Japan, December 1995 (revised August 1996). Available at ftp://ftp.is.titech.ac.jp/pub/OpRes/software/SDPA.
- [4] M. Fukuda, M. Kojima, K. Murota and K. Nakata, Exploiting sparsity in semidefinite programming via matrix completion I: general framework, SIAM J. Optim., to appear.
- [5] D. R. Fulkerson and O. A. Gross, *Incidence matrices and interval graphs*, Pacific J. Math., 15 (1965), pp. 835–855.
- [6] A. GEORGE AND J. W. H. Liu, Computer Solution of Large Sparse Positive Definite Systems, Prentice-Hall, Englewood Cliffs, 1981.
- [7] R. GRONE, C. R. JOHNSON, E. M. SÁ AND H. WOLKOWICZ, *Positive definite completions of partial hermitian matrices*, Linear Algebra Appl., 58 (1984), pp. 109–124.
- [8] C. Johnson, Matrix completion problems: a survey, Proc. Sympos. Appl. Math., 40 (1990), pp. 171–198.