2010/75

Low-rank matrix approximation with weights or missing data is NP-hard

Nicolas Gillis and François Glineur

DISCUSSION PAPER

Center for Operations Research
and Econometrics

Voie du Roman Pays, 34
B-1348 Louvain-la-Neuve
Belgium
http://www.uclouvain.be/core

CORE DISCUSSION PAPER
2010/75

Low-rank matrix approximation with weights or missing data is NP-hard

Nicolas GILLIS [1] and François GLINEUR[2]

November 2010

Abstract

Weighted low-rank approximation (WLRA), a dimensionality reduction technique for data analysis, has been successfully used in several applications, such as in collaborative filtering to design recommender systems or in computer vision to recover structure from motion.
In this paper, we study the computational complexity of WLRA and prove that it is NP-hard to find an approximate solution, even when a rank-one approximation is sought. Our proofs are based on a reduction from the maximum-edge biclique problem, and apply to strictly positive weights as well as binary weights (the latter corresponding to low-rank matrix approximation with missing data).

Keywords: low-rank matrix approximation, weighted low-rank approximation, missing data, matrix completion with noise, PCA with missing data, computational complexity, maximum-edge biclique problem.

[1] Université catholique de Louvain, CORE, B-1348 Louvain-la-Neuve, Belgium. E-mail: Nicolas.gillis@uclouvain.be.
[2] Université catholique de Louvain, CORE, B-1348 Louvain-la-Neuve, Belgium. E-mail: francois.glineur@uclouvain.be. This author is also member of ECORE, the association between CORE and ECARES.

We thank Chia-Tche Chang for his helpful comments. The first author is a research fellow of the Fonds de la Recherche Scientifique (f.R.S.-FNRS).

This paper presents research results of the Belgian Program on Interuniversity Poles of Attraction initiated by the Belgian State, Prime Minister's Office, Science Policy Programming. The scientific responsibility is assumed by the authors.

1 Introduction

Approximating a matrix with one of lower rank is a key problem in data analysis and is widely used for linear dimensionality reduction. Numerous variants exist emphasizing different constraints and objective functions, e.g., principal component analysis (PCA) [15], independent component analysis [5], nonnegative matrix factorization [17], ... and other refinements are often imposed on these models, e.g., sparsity to improve interpretability or increase compression [6].

In some cases, it might be necessary to attach a weight to each entry of the data matrix corresponding to its relative importance [7]. This is for example the case in the following situations:

- ◇ The matrix to be approximated is obtained via a sampling procedure and the number of samples and/or the expected variance vary among the entries, e.g., 2-D digital filter design [18], or microarray data analysis [19].
- ◇ Some data is missing/unknown, which can be taken into account assigning zero weights to the missing/unknown entries of the data matrix. This is for example the case in collaborative filtering, notably used to design recommender systems [22] (in particular, the Netflix prize competition has demonstrated the effectiveness of low-rank matrix factorization techniques [16]), or in computer vision to recover structure from motion [23, 14], see also [3]. This problem is often referred to as *PCA with missing data* [23, 12], and can be viewed as a *low-rank matrix completion problem with noise*, i.e., approximate a given noisy data matrix featuring missing entries with a low-rank matrix[1].
- ◇ A greater emphasis must be placed on the accuracy of the approximation on a localized part of the data, a situation encountered for example in image processing [13, Chapter 6].

Finding a low-rank matrix that is the closest to the input matrix according to these weights is an optimization problem called *weighted low-rank approximation* (WLRA). Formally, it can be formulated as follows: first, given an $m \times n$ nonnegative weight matrix $W \in \mathbb{R}^{m\times n}_{+}$, we define the weighted Frobenius norm of an $m \times n$ matrix A as $||A||_W = (\sum_{i,j} W_{ij} A_{ij}^2)^{\frac{1}{2}}$. Then, given an $m \times n$ real matrix $M \in \mathbb{R}^{m\times n}$ and a positive integer $r \leq \min(m, n)$, we seek an $m \times n$ matrix X with rank at most r that approximates M as closely as possible, where the quality of the approximation is measured by the weighted Frobenius norm of the error:

$$p^* = \inf_{X \in \mathbb{R}^{m\times n}} ||M - X||_W^2 \text{ such that } X \text{ has rank at most } r.$$

Since any $m \times n$ matrix with rank at most r can be expressed as the product of two matrices of dimensions $m \times r$ and $r \times n$, we will use the following more convenient formulation featuring two unknown matrices $U \in \mathbb{R}^{m\times r}$ and $V \in \mathbb{R}^{n\times r}$ but no explicit rank constraint:

$$p^* = \inf_{U \in \mathbb{R}^{m\times r}, V \in \mathbb{R}^{n\times r}} ||M - UV^T||_W^2 = \sum_{ij} W_{ij} (M - UV^T)_{ij}^2 \, . \qquad \text{(WLRA)}$$

Even though (WLRA) is suspected to be NP-hard [14, 24], this has never, to the best of our knowledge, been studied formally. In this paper, we analyze the computational complexity in the rank-one case[2] (i.e., for $r = 1$) and prove the following two results.

Theorem 1. *When $M \in \{0,1\}^{m\times n}$, and $W \in]0,1]^{m\times n}$, it is NP-hard to find an approximate solution of rank-one* (WLRA) *with objective function accuracy less than $2^{-11}(mn)^{-6}$.*

[1] In our settings, the rank of the approximation is fixed a priori.

[2] The obtained results can be easily generalized to any fixed rank r, see Remark 1.

Theorem 2. *When $M \in [0,1]^{m\times n}$, and $W \in \{0,1\}^{m\times n}$, it is NP-hard to find an approximate solution of rank-one* (WLRA) *with objective function accuracy less than $2^{-12}(mn)^{-7}$.*

It is then NP-hard to find an approximate solution to the following problems: (1) rank-one (WLRA) with positive weights, and (2) rank-one approximation of a matrix with missing data.

The paper is organized as follows. We first review existing results about the complexity of (WLRA) in Section 2. In Section 3.1, we introduce the maximum-edge biclique problem (MBP), which is NP-hard. In Section 3, we prove both Theorems 1 and 2 using a polynomial-time reduction from MBP. We conclude with a discussion and some open questions.

Notation. The set of real matrices with dimension m-by-n is denoted $\mathbb{R}^{m\times n}$; the set $\mathbb{R}^{m\times n}$ with component-wise nonnegative entries is denoted $\mathbb{R}^{m\times n}_+$; and $\mathbb{R}_0$ is the set of nonzero reals. For $A \in \mathbb{R}^{m\times n}$, we note $A_{:i}$ the i^{th} column of A, $A_{j:}$ the j^{th} row of A, and A_{ij} the entry at position (i,j); for $b \in \mathbb{R}^{m\times 1} = \mathbb{R}^m$, we note b_i the i^{th} entry of b. The transpose of A is A^T. The Frobenius norm of a matrix A is defined as $||A||_F^2 = \sum_{i,j}(A_{ij})^2$, and $||.||_2$ is the usual Euclidean norm with $||b||_2^2 = \sum_i b_i^2$. For $W \in \mathbb{R}^{m\times n}_+$, the weighted Frobenius 'norm' of a matrix A is defined[3] as $||A||_W^2 = \sum_{i,j} W_{ij}(A_{ij})^2$. The m-by-n matrix of all ones is denoted $\mathbf{1}_{m\times n}$, the m-by-n matrix of all zeros $\mathbf{0}_{m\times n}$, and I_n is the identity matrix of dimension n. The smallest integer larger or equal to x is denoted $\lceil x \rceil$.

2 Previous Results

Weighted low-rank approximation is known to be much more difficult than the corresponding unweighted problem (i.e., when W is the matrix of all ones), which is efficiently solved using the singular value decomposition (SVD) [11]. In fact, it has been previously observed that the weighted problem might have several local minima which are not global [24].

Example 1. *Let*

$$M = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \quad \text{and} \quad W = \begin{pmatrix} 1 & 100 & 2 \\ 100 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix}.$$

In the case of a rank-one factorization ($r = 1$) and a nonnegative matrix M, one can impose without loss of generality that $U \geq 0$ and $V \geq 0$. In fact, one can easily check that any solution UV^T is improved by taking its component-wise absolute value $|UV^T| = |U||V|^T$. Moreover, we can impose without loss of generality that $||U||_2 = 1$, so that only two degrees of freedom remain. Indeed, for a given

$$U(x,y) = \begin{pmatrix} x \\ y \\ \sqrt{1-x^2-y^2} \end{pmatrix}, \quad \text{with} \quad \begin{cases} x \geq 0, y \geq 0 \\ x^2 + y^2 \leq 1 \end{cases},$$

the corresponding optimal $V^*(x,y) = \text{argmin}_V ||M - U(x,y)V||_W^2$ can be computed easily (it reduces to a weighted least squares problem). Figure 1 displays the surface of the objective function $||M - U(x,y)V^*(x,y)^T||_W$ with respect to parameters x and y; we distinguish 4 local minima, close to $(\frac{1}{\sqrt{2}}, 0)$, $(0, \frac{1}{\sqrt{2}})$, $(0,0)$ and $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$. We will see later in Section 3 how this example has been generated.

However, if the rank of the weight matrix $W \in \mathbb{R}^{m\times n}_+$ is equal to one, i.e., $W = st^T$ for some

[3] $||.||_W$ is a matrix norm if and only if $W > 0$.

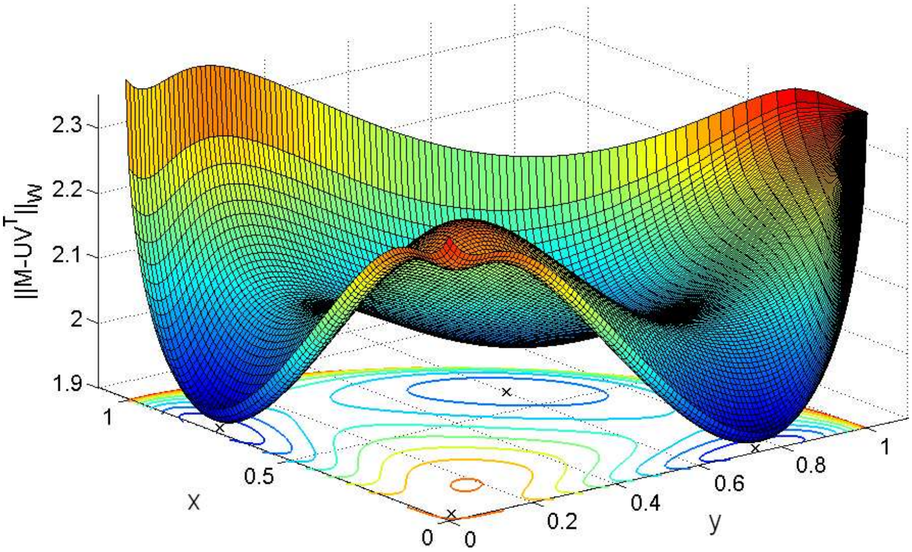

Figure 1: Objective function of (WLRA) with respect to the parameters (x, y).

$s \in \mathbb{R}^m_+$ and $t \in \mathbb{R}^n_+$, (WLRA) can be reduced to an unweighted low-rank approximation. In fact,

$$\begin{aligned} ||M - UV^T||_W^2 &= \sum_{i,j} s_i t_j \, (M - UV^T)^2_{ij} = \sum_{i,j} s_i t_j \, (M - UV^T)^2_{ij} \\ &= \sum_{i,j} \left(\sqrt{s_i t_j} \, M_{ij} - (\sqrt{s_i} \, U_{i:})(\sqrt{t_j} \, V^T_{j:}) \right)^2. \end{aligned}$$

Therefore, if we define a matrix M' such that $M'_{ij} = \sqrt{s_i t_j} \, M_{ij} \; \forall i, j$, an optimal weighted low-rank approximation (U, V) of M can be recovered from the solution (U', V') to the unweighted problem for matrix M' using $U_{i:} = U'_{i:}/\sqrt{s_i} \; \forall i$ and $V_{j:} = V'_{j:}/\sqrt{t_j} \; \forall j$.

When the weight matrix W is binary, WLRA amounts to approximating a matrix with missing data. This problem is closely related to *low-rank matrix completion* (MC), see [2] and references therein, which can be defined as

$$\min_X \; \text{rank}(X) \quad \text{such that } X_{ij} = M_{ij} \text{ for } (i,j) \in \Omega \subset \{1, 2, \ldots, m\} \times \{1, 2, \ldots, n\}, \tag{MC}$$

where Ω is the set of entries for which the values of M are known. (MC) has been shown to be NP-hard [4], and it is clear that an optimal solution X^* of (MC) can be obtained by solving a sequence of (WLRA) problems with the same matrix M, with

$$W_{ij} = \begin{cases} 1 & \text{if } (i,j) \in \Omega \\ 0 & \text{otherwise} \end{cases},$$

and for different values of the target rank ranging from $r = 1$ to $r = \min(m, n)$. The smallest value of r for which the objective function $||M - UV^T||^2_W$ of (WLRA) vanishes provides an optimal solution for (MC). This observation implies that it is NP-hard to solve (WLRA) for each possible value of r (from 1 to $\min(m, n)$) since it would solve (MC). However, this does not imply that (WLRA) is NP-hard when r is fixed, and in particular when $r = 1$. In fact, checking whether (MC) admits a rank-one solution can be done easily[4].

[4]The solution $X = uv^T$ can be constructed observing that the vector u must be multiple of each column of M.

Rank-one (WLRA) can be equivalently reformulated as

$$\inf_A ||M - A||_W^2 \quad \text{such that} \quad \text{rank}(A) \leq 1,$$

and, when W is binary, it is then the problem of finding, if possible, the best rank-one approximation of a matrix with missing entries. To the best of our knowledge, the complexity of this problem has never been studied formally; it will be shown to be NP-hard in the next section.

Another closely related result is the NP-hardness of the structure from motion problem (SFM), in the presence of noise and missing data [20]. Several points of a rigid object are tracked with cameras (we are given the projections of the 3-D points on the 2-D camera planes)[5], and the aim is to recover the structure of the object and the positions of the 3-D points. SFM can be written as a rank-four (WLRA) problem with a binary weight matrix[6] [14]. However, this result does not imply anything on the complexity analysis of rank-one (WLRA).

An important feature of (WLRA) is exposed by the following example.

Example 2. *Let*

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

where ? indicates that an entry is missing, i.e., that the weight associated with this entry is 0 (1 otherwise). Observe that $\forall (u, v) \in \mathbb{R}^m \times \mathbb{R}^n$,

$$\textit{rank}(M) = 2 \ \textit{ and } \ \textit{rank}(uv^T) = 1 \quad \Rightarrow \quad ||M - uv^T||_W > 0.$$

However, we have

$$\inf_{(u,v)\in\mathbb{R}^m\times\mathbb{R}^n} ||M - uv^T||_W = 0.$$

In fact, one can check that with

$$u^{(k)} = \begin{pmatrix} 1 \\ 10^{-k} \end{pmatrix} \ \textit{and} \ v^{(k)} = \begin{pmatrix} 1 \\ 10^{k} \end{pmatrix}, \ \textit{we have} \ \lim_{k\to+\infty} ||M - u^{(k)}v^{(k)^T}||_W = 0.$$

This indicates that when W has zero entries the set of optimal solution of (WLRA) might be empty: there might not exist an optimal solution. In other words, the (bounded) infimum might not be attained. At the other end, the infimum is always attained for $W > 0$ since $||.||_W$ is then a norm.

For this reason, these two cases will be analyzed separately: in Section 3.2, we study the computational complexity of the problem when $W > 0$, and, in Section 3.3, when W is binary (the problem with missing data).

3 Complexity of rank-one (WLRA)

In this section, we use a polynomial-time reduction from the maximum-edge biclique problem to prove Theorems 1 and 2.

[5]Missing data arises because the points might not always be visible by the camera, e.g., in case of rotation.

[6]Except that the last row of V must be all ones, i.e., $V_{r:} = \mathbf{1}_{1\times n}$.

3.1 Maximum-Edge Biclique Problem

A *bipartite graph* is a graph whose vertices can be divided into two disjoint sets such that there is no edge between two vertices in the same set. The maximum-edge biclique problem (MBP) in bipartite graph is the problem of finding a complete bipartite subgraph (a *biclique*) with the maximum number of edges.

Let $M \in \{0,1\}^{m\times n}$ be the biadjacency matrix of a bipartite graph $G_b = (V_1 \cup V_2, E)$ with $V_1 = \{s_1, \dots s_m\}$, $V_2 = \{t_1, \dots t_n\}$ and $E \subseteq (V_1 \times V_2)$, i.e.,

$$M_{ij} = 1 \quad \iff \quad (s_i, t_j) \in E.$$

The cardinality of E will be denoted $|E| = ||M||_F^2 \leq mn$.

For example, Figure 2 displays the graph G_b generated by the matrix M of Example 1.

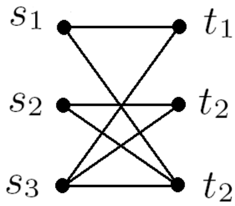

Figure 2: Graph corresponding to the matrix M of Example 1.

With this notation, the maximum-edge biclique problem in a bipartite graph can be formulated as follows [10]

$$\begin{aligned} \min_{u,v} \quad & ||M - uv^T||_F^2 \\ & u_i v_j \leq M_{ij},\ \forall\, i,j \\ & u \in \{0,1\}^m, v \in \{0,1\}^n, \end{aligned} \tag{MBP}$$

where $u_i = 1$ (resp. $v_j = 1$) means that node s_i (resp. t_j) belongs to the solution, $u_i = 0$ (resp. $v_j = 0$) otherwise. The constraint $u_i v_j \leq M_{ij},\ \forall\, i,j$ guarantees feasible solutions of (MBP) to be bicliques of G_b. In fact, it is equivalent to the implication

$$M_{ij} = 0 \quad \Rightarrow \quad u_i = 0 \ \text{ or } \ v_j = 0,$$

i.e., if there is no edge between vertices s_i and t_j, they cannot simultaneously belong to a solution. The objective function minimizes the number of edges outside the biclique, which is equivalent to maximizing the number of edges inside the biclique. Notice that the minimum of (MBP) is $|E| - |E^*|$, where $|E^*|$ denotes the number of edges in an optimal biclique.

The decision version of the MBP problem:

Given K, does G_b contain a biclique with at least K edges?

has been shown to be NP-complete [21] in the usual Turing machine model [8], which is our framework in this paper. Therefore (MBP) is NP-hard.

3.2 Low-Rank Approximation with Positive Weights

In order to prove NP-hardness of rank-one (WLRA) with positive weights ($W > 0$), let us consider the following instance:

$$p^* = \min_{u \in \mathbb{R}^m, v \in \mathbb{R}^n} ||M - uv^T||_W^2, \tag{W-1d}$$

with $M \in \{0,1\}^{m\times n}$ the biadjacency of a bipartite graph $G_b = (V, E)$ and the weight matrix defined as

$$W_{ij} = \begin{cases} 1 & \text{if } M_{ij} = 1 \\ d & \text{if } M_{ij} = 0 \end{cases}, \; 1 \le i \le m, 1 \le j \le n,$$

with $d \ge 1$ a parameter.

Intuitively, increasing the value of d makes the zero entries of M more important in the objective function, which leads them to be approximated by small values. This observation will be used to show that, for d sufficiently large, the optimal value p^* of (W-1d) will be close to the minimum $|E| - |E^*|$ of (MBP) (Lemma 2).

In fact, as the value of parameter d increases, the local minima of (W-1d) get closer to the 'locally' optimal solutions of (MBP), which are binary vectors describing the maximal bicliques in G_b, i.e., bicliques not contained in larger bicliques. Example 1 illustrates the situation: the graph G_b corresponding to matrix M (cf. Figure 2) contains four maximal bicliques $\{s_1, s_3, t_1, t_3\}$, $\{s_2, s_3, t_2, t_3\}$, $\{s_3, t_1, t_2, t_3\}$ and $\{s_1, s_2, s_3, t_3\}$, and the weight matrix W that was used is similar to the case $d = 100$ in problem (W-1d). We now observe that (W-1d) has four local optimal solutions as well (cf. Figure 1) close to $(\frac{1}{\sqrt{2}}, 0)$, $(0, \frac{1}{\sqrt{2}})$, $(0, 0)$ and $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$. There is a one to one correspondence between these solutions and the four maximal bicliques listed above (in this order). For example, for $(x, y) = (\frac{1}{\sqrt{2}}, 0)$ we have $U(x, y) = (\frac{1}{\sqrt{2}}\, 0\, \frac{1}{\sqrt{2}})^T$, $V^*(x, y)$ is approximately equal to $(\sqrt{2}\, 0\, \sqrt{2})^T$, and this solution corresponds to the maximal biclique $\{s_1, s_3, t_1, t_3\}$.

Notice that a similar idea was used in [9] to prove NP-hardness of the rank-one nonnegative factorization problem $\min_{u\in\mathbb{R}^m_+, v\in\mathbb{R}^n_+} ||M - uv^T||_F$, where the zero entries of M were replaced by sufficiently large negative ones.

Let us now prove this formally. It is first observed that for any (u, v) such that $||M - uv^T||^2_W \le |E|$, the absolute value of the row or the column of uv^T corresponding to a zero entry of M must be smaller than a constant inversely proportional to $\sqrt[4]{d}$.

Lemma 1. *Let (i, j) be such that $M_{ij} = 0$, then $\forall (u, v)$ such that $||M - uv^T||^2_W \le |E|$,*

$$\min\Big(\max_{1\le k\le n} |u_i v_k|, \max_{1\le p\le m} |u_p v_j| \Big) \le \sqrt[4]{\frac{4|E|^2}{d}}.$$

Proof. Without loss of generality u and v can be scaled such that $||u||_2 = ||v||_2$ without changing the product uv^T. First, observe that since $||.||_W$ is a norm,

$$||uv^T||_W - \sqrt{|E|} = ||uv^T||_W - ||M||_W \le ||M - uv^T||_W \le \sqrt{|E|}.$$

Since all entries of W are larger than 1 ($d \ge 1$), we have

$$||u||_2||v||_2 = ||uv^T||_F \le ||uv^T||_W \le \sqrt{4|E|},$$

and then $||u||_2 = ||v||_2 \le \sqrt[4]{4|E|}$.

Moreover $d(0 - u_i v_j)^2 \le ||M - uv^T||^2_W \le |E|$, so that $|u_i v_j| \le \sqrt{\frac{|E|}{d}}$ which implies that either $|u_i| \le \sqrt[4]{\frac{|E|}{d}}$ or $|v_j| \le \sqrt[4]{\frac{|E|}{d}}$. Combining above inequalities with the fact that $(\max_{1\le k\le n} |v_k|)$ and $(\max_{1\le p\le m} |u_p|)$ are bounded above by $||u||_2 = ||v||_2 \le \sqrt[4]{4|E|}$ completes the proof. □

Using Lemma 1, we can associate any point (u, v) such that $||M - uv^T||^2_W \le |E|$ with a biclique of G_b, the graph generated by the biadjacency matrix M.

Corollary 1. *For any pair* (u,v) *such that* $||M - uv^T||_W^2 \leq |E|$, *the set*

$$\Omega(u,v) = I \times J, \quad \text{with } I = \{\, i \mid \exists j \text{ s.t. } |u_i v_j| > \alpha \,\} \text{ and } J = \{\, j \mid \exists i \text{ s.t. } |u_i v_j| > \alpha \,\},$$

where $\alpha = \sqrt[4]{\frac{4|E|^2}{d}}$, *defines a biclique of* G_b.

We can now provide lower and upper bounds on the optimal value p^* of (W-1d), and show that it is not too different from the optimal value $|E| - |E^*|$ of (MBP).

Lemma 2. *Let* $0 < \epsilon \leq 1$. *For any value of parameter* d *such that* $d \geq \frac{2^6|E|^6}{\epsilon^4}$, *the optimal value* p^* *of* (W-1d) *satisfies*

$$|E| - |E^*| - \epsilon < p^* \leq |E| - |E^*|.$$

Proof. Let (u,v) be an optimal solution of (W-1d) (there always exists at least one optimal solution, cf. Section 2), and let us note $p = |E| - |E^*| \geq 0$. Since any optimal solution of (MBP) plugged in (W-1d) also achieves an objective function equal to p, we must have

$$p^* = ||M - uv^T||_W^2 \leq p = |E| - |E^*|,$$

which gives the upper bound.

By Corollary 1, the set $\Omega = \Omega(u,v)$ defines a biclique of (MBP) with $|\Omega| \leq |E^*|$ edges. By construction, the entries in M which are not in Ω are approximated by values smaller than α. If $\alpha = \sqrt[4]{\frac{4|E|^2}{d}} \leq 1$, i.e., $d \geq 4|E|^2$ which is satisfied for $0 < \epsilon \leq 1$, the error corresponding to a one entry of M not in the biclique Ω is at least $(1-\alpha)^2$. Since there are at least $p = |E| - |E^*|$ such entries, we have

$$(1-\alpha)^2 p \leq ||M - uv^T||_W^2. \tag{3.1}$$

Moreover

$$(1-\alpha)^2 p > (1-2\alpha)p = p - 2\alpha p \geq p - 2\alpha|E| \geq p - \epsilon,$$

since $2\alpha|E| \leq \epsilon \iff d \geq \frac{2^6|E|^6}{\epsilon^4}$, which gives the lower bound. □

This result implies that for $\epsilon = 1$, i.e., for $d \geq (2|E|)^6$, we have $|E| - |E^*| - 1 < p^* \leq |E| - |E^*|$, and therefore computing p^* exactly would allow to recover $|E^*|$ (since $\lceil p^* \rceil = |E| - |E^*|$), which is NP-hard. Since the reduction from (MBP) to (W-1d) is polynomial (it uses the same matrix M and a weight matrix W whose description has polynomial length), we conclude that solving (W-1d) exactly is NP-hard. The next result shows that even solving (W-1d) approximately is NP-hard.

Corollary 2. *For any* $d > (2mn)^6$, $M \in \{0,1\}^{m\times n}$, *and* $W \in \{1,d\}^{m\times n}$, *it is NP-hard to find an approximate solution of rank-one* (WLRA) *with objective function accuracy less than* $1 - \frac{(2mn)^{3/2}}{d^{1/4}}$.

Proof. Let $d > (2mn)^6$, $0 < \epsilon = \frac{(2mn)^{3/2}}{d^{1/4}} < 1$, and $(\bar{u}, \bar{v})$ be an approximate solution of (W-1d) with objective function accuracy $(1-\epsilon)$, i.e., $p^* \leq \bar{p} = ||M - \bar{u}\bar{v}^T||_W^2 \leq p^* + 1 - \epsilon$. Since $d = \frac{(2mn)^6}{\epsilon^4} \geq \frac{(2|E|)^6}{\epsilon^4}$, Lemma 2 applies and we have

$$|E| - |E^*| - \epsilon \; < \; p^* \leq \; \bar{p} \; \leq \; p^* + 1 - \epsilon \; \leq \; |E| - |E^*| + 1 - \epsilon.$$

We finally observe that $\bar{p}$ allows to recover $|E^*|$, which is NP-hard. In fact, adding ϵ to the above inequalities gives $|E| - |E^*| < \bar{p} + \epsilon \leq |E| - |E^*| + 1$, and therefore

$$|E^*| = |E| - \left\lceil \bar{p} + \epsilon \right\rceil + 1.$$

□

We are now in position to prove Theorem 1, which deals with the hardness of rank-one (WLRA) with bounded weights.

Proof of Theorem 1. Let us use Corollary 2 with $W \in \{1, d\}^{m \times n}$, and define $W' = \frac{1}{d} W \in \{\frac{1}{d}, 1\}^{m \times n}$. Clearly, replacing W by W' in (W-1d) simply amounts to multiplying the objective function by $\frac{1}{d}$, with $||M - uv^T||^2_{W'} = \frac{1}{d} ||M - uv^T||^2_W$. Taking $d^{1/4} = 2(2mn)^{3/2}$ in Corollary 2, we obtain that for $M \in \{0, 1\}^{m \times n}$ and $W \in]0, 1]^{m \times n}$, it is NP-hard to find an approximate solution of rank-one (WLRA) with objective function accuracy less than $\frac{1}{d}\left(1 - \frac{(2mn)^{3/2}}{d^{1/4}}\right) = \frac{1}{2d} = 2^{-11}(mn)^{-6}$. □

Remark 1. *Using the same construction as in [10, Theorem 3], this rank-one NP-hardness result can be generalized to any factorization rank, i.e., approximate* (WLRA) *for any fixed rank r is NP-hard.*

Remark 2. *The bounds on d have been quite crudely estimated, and can be improved. Our goal was only to show existence of a polynomial-time reduction from* (MBP) *to rank-one* (WLRA).

3.3 Low-Rank Matrix Approximation with Missing Data

Unfortunately, the above NP-hardness proof does not include the case when W is binary, corresponding to missing data in the matrix to be approximated (or to low-rank matrix completion with noise). This corresponds to the following problem

$$\inf_{U \in \mathbb{R}^{m \times r}, V \in \mathbb{R}^{n \times r}} \quad ||M - UV^T||^2_W = \sum_{ij} W_{ij}(M - UV^T)^2_{ij}\,, \quad W \in \{0, 1\}^{m \times n}. \tag{LRAMD}$$

In the same spirit as before, we consider the following rank-one version of the problem

$$p^* = \inf_{u \in \mathbb{R}^m, v \in \mathbb{R}^n} ||M - uv^T||^2_W, \tag{MD-1d}$$

with input data matrices M and W defined as follows

$$M = \left(\begin{array}{c|c} M_b & \mathbf{0}_{s \times Z} \\ \hline \mathbf{0}_{Z \times t} & dI_Z \end{array}\right) \text{ and } W = \left(\begin{array}{c|c} \mathbf{1}_{s \times t} & B_1 \\ \hline B_2 & I_Z \end{array}\right),$$

where $M_b \in \{0, 1\}^{s \times t}$ is the biadjacency matrix of the bipartite graph $G_b = (V, E)$, $d > 1$ is a parameter, $Z = st - |E|$ is the number of zero entries in M_b, $m = s + Z$ and $n = t + Z$ are the dimensions of M and W.

Binary matrices $B_1 \in \{0, 1\}^{s \times Z}$ and $B_2 \in \{0, 1\}^{Z \times t}$ are constructed as follows: assume the Z zero entries of M_b can be enumerated as $\{M_b(i_1, j_1), M_b(i_2, j_2), \ldots, M_b(i_Z, j_Z)\}$, and let k_{ij} be the (unique) index k ($1 \leq k \leq Z$) such that $(i_k, j_k) = (i, j)$ (therefore k_{ij} is only defined for pairs (i, j) such that $M_b(i, j) = 0$, and establishes a bijection between these pairs and the set $\{1, 2, \ldots, Z\}$). We now define matrices B_1 and as follows: for every index $1 \leq k_{ij} \leq Z$, we have

$$B_1(i, k_{ij}) = 1, B_1(i', k_{ij}) = 0 \;\forall i' \neq i \text{ and } B_2(k_{ij}, j) = 1, B_2(k_{ij}, j') = 0 \;\forall j' \neq j\,.$$

Equivalently, each column of B_1 (resp. row of B_2) corresponds to a different zero entry $M_b(i, j) = 0$, and contains only zeros except for a one in position i within the column (resp j within the row).

In the case of Example 1, we get

$$M = \left(\begin{array}{ccc|c} 1 & 0 & 1 & \\ 0 & 1 & 1 & \mathbf{0}_{3 \times 2} \\ 1 & 1 & 1 & \\ \hline & \mathbf{0}_{2 \times 3} & & d\,I_2 \end{array}\right) \text{ and } W = \left(\begin{array}{ccc|cc} & & & 1 & 0 \\ & \mathbf{1}_{3 \times 3} & & 0 & 1 \\ & & & 0 & 0 \\ \hline 0 & 1 & 0 & \multicolumn{2}{c}{\multirow{2}{*}{I_2}} \\ 1 & 0 & 0 & & \end{array}\right),$$

i.e., the matrix to be approximated can be represented as

$$\left(\begin{array}{ccc|cc} 1 & 0 & 1 & 0 & ? \\ 0 & 1 & 1 & ? & 0 \\ 1 & 1 & 1 & ? & ? \\ \hline ? & 0 & ? & d & ? \\ 0 & ? & ? & ? & d \end{array}\right).$$

For any feasible solution (u, v) of (MD-1d), we also note

$$u = \left(\begin{array}{c} u_b \\ u_d \end{array} \right) \in \mathbb{R}^m, \; u_b \in \mathbb{R}^s \text{ and } u_d \in \mathbb{R}^Z,$$

$$v = \left(\begin{array}{c} v_b \\ v_d \end{array} \right) \in \mathbb{R}^n, \; v_b \in \mathbb{R}^t \text{ and } v_d \in \mathbb{R}^Z.$$

We will show that this formulation ensures that, as d increases, the zero entries of the matrix M_b (upper left of matrix M, which is the biadjacency matrix of G_b) have to be approximated with smaller values. Hence, as for (W-1d), we will be able to prove that the optimal value p^* of (MD-1d) will have to get close to the minimum $|E| - |E^*|$ of (MBP), implying its NP-hardness.

Intuitively, when d is large, the lower right matrix dI_Z of M will have to be approximated by a matrix with large diagonal entries since they correspond to one entries in the weight matrix W. Hence $u_d(k_{ij})v_d(k_{ij})$ has to be large for all $1 \leq k_{ij} \leq Z$. We then have at least either $u_d(k_{ij})$ or $v_d(k_{ij})$ large for all k_{ij} (recall each k_{ij} corresponds to a zero entry in M at position (i, j), cf. definition of B_1 and B_2 above). By construction, we also have two entries $M(s + k_{ij}, j) = 0$ and $M(i, t + k_{ij}) = 0$ with nonzero weights corresponding to the nonzero entries $B_1(i, k_{ij})$ and $B_2(k_{ij}, j)$, which then have to be approximated by small values. If $u_d(k_{ij})$ (resp. $v_d(k_{ij})$) is large, then $v_b(j)$ (resp. $u_b(i)$) will have to be small since $u_d(k_{ij})v_b(j) \approx 0$ (resp. $u_b(i)v_d(k_{ij}) \approx 0$). Finally, either $u_b(i)$ or $v_b(j)$ has to be small, implying that $M_b(i, j)$ is approximated by a small value, because (u_b, v_b) is bounded independently of the value of d.

We now proceed as in Section 3.2. Let us first give an upper bound for the optimal value p^* of (MD-1d).

Lemma 3. *For $d > 1$, the optimal value p^* of* (MD-1d) *is bounded above by $|E| - |E^*|$, i.e.,*

$$p^* = \inf_{u \in \mathbb{R}^m, v \in \mathbb{R}^n} ||M - uv^T||_W^2 \leq |E| - |E^*|. \tag{3.2}$$

Proof. Let us build the following feasible solution (u, v) of (MD-1d) where (u_b, v_b) is an optimal solution of (MBP) and (u_d, v_d) is defined as

$$u_d(k_{ij}) = \left\{ \begin{array}{lll} d^K & \text{if} & u_b(i) = 0, \\ d^{1-K} & \text{if} & u_b(i) = 1, \end{array} \right. \quad \text{and} \quad v_d(k_{ij}) = \left\{ \begin{array}{lll} d^K & \text{if} & v_b(j) = 0, \\ d^{1-K} & \text{if} & v_b(j) = 1, \end{array} \right.$$

with $K \in \mathbb{R}$ and k_{ij} the index of the column of B_1 and the row of B_2 corresponding to the zero entry (i, j) of M_b (i.e., $(i, j) = (i_{k_{ij}}, j_{k_{ij}})$).

One can check that

$$(uv^T) \circ W = \left(\begin{array}{cc} u_b v_b^T & d^{1-K} B_1 \\ d^{1-K} B_2 & dI_Z \end{array} \right),$$

where $\circ$ is the component-wise (or Hadamard) product between two matrices, so that

$$p^* \leq ||M - uv^T||_W^2 = |E| - |E^*| + \frac{2Z}{d^{2(K-1)}}, \quad \forall K. \tag{3.3}$$

Since $d > 1$, taking the limit $K \to +\infty$ gives the result. $\square$

We now prove a property similar to Lemma 1 for any solution with objective value smaller that $|E|$.

Lemma 4. *Let $d > \sqrt{|E|}$ and (i, j) be such that $M_b(i, j) = 0$, then the following holds for any pair (u, v) such that $||M - uv^T||_W^2 \leq |E|$:*

$$\min\Big(\max_{1\leq k\leq n} |u_i v_k|, \max_{1\leq p\leq m} |u_p v_j|\Big) \leq \frac{\sqrt{2}\,|E|^{\frac{3}{4}}}{\left(d - \sqrt{|E|}\right)^{\frac{1}{2}}}. \tag{3.4}$$

Proof. Without loss of generality we set $||u_b||_2 = ||v_b||_2$ by scaling u and v without changing uv^T. Observing that

$$||u_b||_2||v_b||_2 - \sqrt{|E|} = ||u_b v_b^T||_F - ||M_b||_F \leq ||M_b - u_b v_b^T||_F \leq ||M - uv^T||_W \leq \sqrt{|E|},$$

we have $||u_b||_2||v_b||_2 \leq 2\sqrt{|E|}$, and $||u_b||_2 = ||v_b||_2 \leq \sqrt{2}|E|^{\frac{1}{4}}$.
Assume $M_b(i, j)$ is zero for some pair (i, j) and let $k = k_{ij}$ denote the index of the corresponding column of B_1 and row of B_2 (i.e., such that $B_1(i, k) = B_2(k, j) = 1$). By construction, $u_d(k)v_d(k)$ has to approximate d in the objective function. This implies $(u_d(k)v_d(k) - d)^2 \leq |E|$ and then

$$u_d(k)v_d(k) \geq d - \sqrt{|E|} > 0.$$

Suppose $|u_d(k)|$ is greater than $|v_d(k)|$ (the case $|v_d(k)|$ greater than $|u_d(k)|$ is similar), this implies $|u_d(k)| \geq (d - |E|^{\frac{1}{2}})^{\frac{1}{2}}$. Moreover $u_d(k)v_j$ has to approximate zero in the objective function, since $B_2(k, j) = 1$, implying

$$(u_d(k)v_j - 0)^2 \leq |E| \quad \Rightarrow \quad |u_d(k)v_j| \leq \sqrt{|E|}.$$

Hence

$$|v_j| \leq \frac{\sqrt{|E|}}{|u_d(k)|} \leq \frac{|E|^{\frac{1}{2}}}{\left(d - \sqrt{|E|}\right)^{\frac{1}{2}}}, \tag{3.5}$$

and since $(\max_{1\leq p\leq m} |u_p|)$ is bounded by $||u_b||_2 \leq \sqrt{2}|E|^{\frac{1}{4}}$, the proof is complete. □

One can now associate to any point with objective value smaller than $|E|$ a biclique of G_b, the graph generated by the biadjacency matrix M_b.

Corollary 3. *Let $d > \sqrt{|E|}$, then for any pair (u, v) such that $||M - uv^T||_W^2 \leq |E|$, the set*

$$\Omega(u, v) = I \times J, \quad \text{with} \quad I = \{\, i \mid \exists j \text{ s.t. } |u_i v_j| > \beta \,\} \text{ and } J = \{\, j \mid \exists i \text{ s.t. } |u_i v_j| > \beta \,\}, \tag{3.6}$$

where $\beta = \frac{\sqrt{2}\,|E|^{\frac{3}{4}}}{\left(d-\sqrt{|E|}\right)^{\frac{1}{2}}}$, defines a biclique of G_b.

The next lemma gives a lower bound for the value of p^*.

Lemma 5. *Let $0 < \epsilon \leq 1$. For any value of parameter d that satisfies $d > \frac{8|E|^{\frac{7}{2}}}{\epsilon^2} + |E|^{\frac{1}{2}}$, the infimum p^* of* (MD-1d) *satisfies*

$$|E| - |E^*| - \epsilon < p^*.$$

Proof. If $|E| = |E^*|$, the result is trivial since $p^* = 0$. Otherwise, suppose $p^* \leq |E| - |E^*| - \epsilon$ and let $\beta = \frac{\sqrt{2}\,|E|^{\frac{3}{4}}}{\left(d-\sqrt{|E|}\right)^{\frac{1}{2}}}$. First observe that $d > \frac{8|E|^{\frac{7}{2}}}{\epsilon^2} + |E|^{\frac{1}{2}}$ is equivalent to $2|E|\beta < \epsilon$. Then, by continuity of (MD-1d), for any δ such that $\delta < \epsilon$, there exists a pair (u, v) such that

$$||M_d - uv^T||_W^2 \leq |E| - |E^*| - \delta.$$

In particular, let us take $\delta = 2|E|\beta < \epsilon$. We can now proceed as for Lemma 2. By Corollary 3, $\Omega(u,v)$ corresponds to a biclique of G_b, with at most $|E^*|$ edges. Then, for $\beta \leq 1$, i.e., for $d \geq 2|E|^{\frac{3}{2}} + |E|^{\frac{1}{2}}$ satisfied for $0 < \epsilon \leq 1$,

$$(1-\beta)^2(|E| - |E^*|) \leq ||M - uv^T||_W^2 \leq |E| - |E^*| - \delta.$$

Dividing the above inequalities by $|E| - |E^*| > 0$, we obtain

$$1 - 2\beta < (1-\beta)^2 \leq 1 - \frac{\delta}{|E| - |E^*|} \leq 1 - \frac{\delta}{|E|} \Rightarrow \delta < 2|E|\beta,$$

a contradiction. □

Corollary 4. *For any $d > 8(mn)^{7/2} + \sqrt{mn}$, $M \in \{0,1,d\}^{m\times n}$, and $W \in \{0,1\}^{m\times n}$, it is NP-hard to find an approximate solution of rank-one* (WLRA) *with objective function accuracy $1 - \frac{2\sqrt{2}(mn)^{7/4}}{(d-\sqrt{mn})^{1/2}}$.*

Proof. Let $d > 8(mn)^{7/2} + \sqrt{mn}$, $0 < \epsilon = \frac{2\sqrt{2}(mn)^{7/4}}{(d-\sqrt{mn})^{1/2}} < 1$, and $(\bar{u},\bar{v})$ be an approximate solution of (W-1d) with absolute error $(1-\epsilon)$, i.e., $p^* \leq \bar{p} = ||M - \bar{u}\bar{v}^T||_W^2 \leq p^* + 1 - \epsilon$. Lemma 5 applies because $d = \frac{8(mn)^{7/2}}{\epsilon^2} + \sqrt{mn} \geq \frac{8(st)^{7/2}}{\epsilon^2} + \sqrt{st} \geq \frac{8|E|^{7/2}}{\epsilon^2} + |E|^{1/2}$. Using Lemmas 3 and 5, the rest of the proof is identical as the one of Theorem 1. Since the reduction from (MBP) to (MD-1d) is polynomial (description of matrices W and M has polynomial length, since the increase in matrix dimensions from M_b to M is polynomial), we conclude that finding such an approximate solution for (MD-1d) is NP-hard. □

We can now easily derive Theorem 2, which deals with the hardness of rank-one (WLRA) with a bounded matrix M.

Proof of Theorem 2. Replacing M by $M' = \frac{1}{d}M$ in (MD-1d) gives an equivalent problem with objective function multiplied by $\frac{1}{d^2}$, since $\frac{1}{d^2}||M - uv^T||_W^2 = ||M' - \frac{uv^T}{d}||_W^2$. Taking $d = 2^5(mn)^{7/2} + \sqrt{mn}$ in Corollary 4, we find that it is NP-hard to compute an approximate solution of rank-one (WLRA) for $M \in [0,1]^{m\times n}$ and $W \in \{0,1\}^{m\times n}$, and with objective function accuracy less than $\frac{1}{d^2}\left(1 - \frac{2\sqrt{2}(mn)^{7/4}}{(d-\sqrt{mn})^{1/2}}\right) = \frac{1}{2d^2} \geq 2^{-12}(mn)^{-7}$. □

4 Concluding Remarks

In this paper, we have studied the complexity of the weighted low-rank approximation problem (WLRA), and proved that finding an approximate solution is NP-hard, already in the rank-one case, both for positive and for binary weights (the latter also corresponding to low-rank matrix completion with noise, or PCA with missing data).

Nevertheless, some questions remain open. In particular,

- ◇ When W is the matrix of all ones, WLRA can be solved in polynomial-time. We have shown that, when the ratio between the largest and the smallest entry in W is large enough, the problem is NP-hard (Theorem 1). It would be interesting to investigate the gap between these two facts, i.e., what is the minimum ratio of the entries of W so that WLRA is NP-hard?
- ◇ When $\text{rank}(W) = 1$, WLRA can be solved in polynomial-time (cf. Section 2) while it is NP-hard for general matrix W (with rank up to $\min(m,n)$). But what is the complexity of (WLRA) if the rank of the weight matrix W is fixed and greater than one, e.g., if $\text{rank}(W) = 2$?

◇ When data is missing, the rank-one matrix approximation problem is NP-hard in general. Nevertheless, it has been observed [1] that when the given entries are sufficiently numerous, well distributed in the matrix, and affected by a relatively low level of noise, the original uncorrupted low-rank matrix can be recovered accurately, with a technique based on convex optimization (minimization of the nuclear norm of the approximation, which can be done efficiently). It would then be particularly interesting to analyze the complexity of the problem given additional assumptions on the data matrix, for example on the noise distribution, and deal in particular with situations related to applications.

References

[1] E.J. Candès and Y. Plan, *Matrix Completion with Noise*, in Proceedings of the IEEE, 2009.

[2] E.J. Candès and B. Recht, *Exact Matrix Completion via Convex Optimization*, Foundations of Computational Mathematics, 9 (2009), pp. 717–772.

[3] P. Chen, *Optimization Algorithms on Subspaces: Revisiting Missing Data Problem in Low-Rank Matrix*, International Journal of Computer Vision, 80(1) (2008), pp. 125–142.

[4] A.L. Chistov and D.Yu. Grigoriev, *Complexity of quantifier elimination in the theory of algebraically closed fields*, Proceedings of the 11th Symposium on Mathematical Foundations of Computer Science, Lecture Notes in Computer Science, Springer, 176 (1984), pp. 17–31.

[5] P. Comon, *Independent component analysis, A new concept?*, Signal Processing, 36 (1994), pp. 287–314.

[6] A. d'Aspremont, L. El Ghaoui, M.I. Jordan, and G.R.G. Lanckriet, *A Direct Formulation for Sparse PCA Using Semidefinite Programming*, SIAM Rev., 49(3) (2007), pp. 434–448.

[7] K.R. Gabriel and S. Zamir, *Lower Rank Approximation of Matrices by Least Squares With Any Choice of Weights*, Technometrics, 21(4) (1979), pp. 489–498.

[8] M.R. Garey and D.S. Johnson, *Computers and Intractability: A guide to the theory of NP-completeness*, Freeman, San Francisco, 1979.

[9] N. Gillis and F. Glineur, *Nonnegative Factorization and The Maximum Edge Biclique Problem.* CORE Discussion paper 2008/64, 2008.

[10] ——, *Using underapproximations for sparse nonnegative matrix factorization*, Pattern Recognition, 43(4) (2010), pp. 1676–1687.

[11] G.H. Golub and C.F. Van Loan, *Matrix Computation, 3rd Edition*, The Johns Hopkins University Press Baltimore, 1996.

[12] B. Grung and R. Manne, *Missing values in principal component analysis*, Chemom. and Intell. Lab. Syst., 42 (1998), pp. 125–139.

[13] N.-D. Ho, *Nonnegative Matrix Factorization - Algorithms and Applications*, PhD thesis, Université catholique de Louvain, 2008.

[14] D. Jacobs, *Linear fitting with missing data for structure-from-motion*, Vision and Image Understanding, 82 (2001), pp. 57–81.

[15] I.T. Jolliffe, *Principal Component Analysis*, Springer-Verlag, 1986.

[16] Y. Koren, R. Bell, and C. Volinsky, *Matrix Factorization Techniques for Recommender Systems*, IEEE Computer, 42(8) (2009), pp. 30–37.

[17] D.D. Lee and H.S. Seung, *Learning the Parts of Objects by Nonnegative Matrix Factorization*, Nature, 401 (1999), pp. 788–791.

[18] W.-S. Lu, S.-C. Pei, and P.-H. Wang, *Weighted low-rank approximation of general complex matrices and its application in the design of 2-D digital filters*, IEEE Trans. Circuits Syst. I, 44 (1997), pp. 650–655.

[19] I. Markovsky and M. Niranjan, *Approximate low-rank factorization with structured factors*, Computational Statistics & Data Analysis, 54 (2010), pp. 3411–3420.

[20] D. Nister, F. Kahl, and H. Stewenius, *Structure from Motion with Missing Data is NP-Hard*, in IEEE 11th Int. Conf. on Computer Vision, 2007.

[21] R. Peeters, *The maximum edge biclique problem is NP-complete*, Discrete Applied Mathematics, 131(3) (2003), pp. 651–654.

[22] B.M. Sarwar, G. Karypis, J.A. Konstan, and J. Riedl, *Item-Based Collaborative Filtering Recommendation Algorithms*, in 10th International WorldWideWeb Conference, 2001.

[23] H. Shum, K. Ikeuchi, and R. Reddy, *Principal component analysis with missing data and its application to polyhedral object modeling*, IEEE Trans. Pattern Anal. Mach. Intelligence, 17(9) (1995), pp. 854–867.

[24] N. Srebro and T. Jaakkola, *Weighted Low-Rank Approximations*, in 20th ICML Conference Proceedings, 2004.

Recent titles

CORE Discussion Papers

2010/35. Jean-Jacques DETHIER, Pierre PESTIEAU and Rabia ALI. The impact of a minimum pension on old age poverty and its budgetary cost. Evidence from Latin America.
2010/36. Stéphane ZUBER. Justifying social discounting: the rank-discounting utilitarian approach.
2010/37. Marc FLEURBAEY, Thibault GAJDOS and Stéphane ZUBER. Social rationality, separability, and equity under uncertainty.
2010/38. Helmuth CREMER and Pierre PESTIEAU. Myopia, redistribution and pensions.
2010/39. Giacomo SBRANA and Andrea SILVESTRINI. Aggregation of exponential smoothing processes with an application to portfolio risk evaluation.
2010/40. Jean-François CARPANTIER. Commodities inventory effect.
2010/41. Pierre PESTIEAU and Maria RACIONERO. Tagging with leisure needs.
2010/42. Knud J. MUNK. The optimal commodity tax system as a compromise between two objectives.
2010/43. Marie-Louise LEROUX and Gregory PONTHIERE. Utilitarianism and unequal longevities: A remedy?
2010/44. Michel DENUIT, Louis EECKHOUDT, Ilia TSETLIN and Robert L. WINKLER. Multivariate concave and convex stochastic dominance.
2010/45. Rüdiger STEPHAN. An extension of disjunctive programming and its impact for compact tree formulations.
2010/46. Jorge MANZI, Ernesto SAN MARTIN and Sébastien VAN BELLEGEM. School system evaluation by value-added analysis under endogeneity.
2010/47. Nicolas GILLIS and François GLINEUR. A multilevel approach for nonnegative matrix factorization.
2010/48. Marie-Louise LEROUX and Pierre PESTIEAU. The political economy of derived pension rights.
2010/49. Jeroen V.K. ROMBOUTS and Lars STENTOFT. Option pricing with asymmetric heteroskedastic normal mixture models.
2010/50. Maik SCHWARZ, Sébastien VAN BELLEGEM and Jean-Pierre FLORENS. Nonparametric frontier estimation from noisy data.
2010/51. Nicolas GILLIS and François GLINEUR. On the geometric interpretation of the nonnegative rank.
2010/52. Yves SMEERS, Giorgia OGGIONI, Elisabetta ALLEVI and Siegfried SCHAIBLE. Generalized Nash Equilibrium and market coupling in the European power system.
2010/53. Giorgia OGGIONI and Yves SMEERS. Market coupling and the organization of counter-trading: separating energy and transmission again?
2010/54. Helmuth CREMER, Firouz GAHVARI and Pierre PESTIEAU. Fertility, human capital accumulation, and the pension system.
2010/55. Jan JOHANNES, Sébastien VAN BELLEGEM and Anne VANHEMS. Iterative regularization in nonparametric instrumental regression.
2010/56. Thierry BRECHET, Pierre-André JOUVET and Gilles ROTILLON. Tradable pollution permits in dynamic general equilibrium: can optimality and acceptability be reconciled?
2010/57. Thomas BAUDIN. The optimal trade-off between quality and quantity with uncertain child survival.
2010/58. Thomas BAUDIN. Family policies: what does the standard endogenous fertility model tell us?
2010/59. Nicolas GILLIS and François GLINEUR. Nonnegative factorization and the maximum edge biclique problem.
2010/60. Paul BELLEFLAMME and Martin PEITZ. Digital piracy: theory.
2010/61. Axel GAUTIER and Xavier WAUTHY. Competitively neutral universal service obligations.
2010/62. Thierry BRECHET, Julien THENIE, Thibaut ZEIMES and Stéphane ZUBER. The benefits of cooperation under uncertainty: the case of climate change.
2010/63. Marco DI SUMMA and Laurence A. WOLSEY. Mixing sets linked by bidirected paths.
2010/64. Kaz MIYAGIWA, Huasheng SONG and Hylke VANDENBUSSCHE. Innovation, antidumping and retaliation.

Recent titles

CORE Discussion Papers - continued

2010/65. Thierry BRECHET, Natali HRITONENKO and Yuri YATSENKO. Adaptation and mitigation in long-term climate policies.
2010/66. Marc FLEURBAEY, Marie-Louise LEROUX and Gregory PONTHIERE. Compensating the dead? Yes we can!
2010/67. Philippe CHEVALIER, Jean-Christophe VAN DEN SCHRIECK and Ying WEI. Measuring the variability in supply chains with the peakedness.
2010/68. Mathieu VAN VYVE. Fixed-charge transportation on a path: optimization, LP formulations and separation.
2010/69. Roland Iwan LUTTENS. Lower bounds rule!
2010/70. Fred SCHROYEN and Adekola OYENUGA. Optimal pricing and capacity choice for a public service under risk of interruption.
2010/71. Carlotta BALESTRA, Thierry BRECHET and Stéphane LAMBRECHT. Property rights with biological spillovers: when Hardin meets Meade.
2010/72. Olivier GERGAUD and Victor GINSBURGH. Success: talent, intelligence or beauty?
2010/73. Jean GABSZEWICZ, Victor GINSBURGH, Didier LAUSSEL and Shlomo WEBER. Foreign languages' acquisition: self learning and linguistic schools.
2010/74. Cédric CEULEMANS, Victor GINSBURGH and Patrick LEGROS. Rock and roll bands, (in)complete contracts and creativity.
2010/75. Nicolas GILLIS and François GLINEUR. Low-rank matrix approximation with weights or missing data is NP-hard.

Books

J. GABSZEWICZ (ed.) (2006), La différenciation des produits. Paris, La découverte.
L. BAUWENS, W. POHLMEIER and D. VEREDAS (eds.) (2008), *High frequency financial econometrics: recent developments*. Heidelberg, Physica-Verlag.
P. VAN HENTENRYCKE and L. WOLSEY (eds.) (2007), *Integration of AI and OR techniques in constraint programming for combinatorial optimization problems*. Berlin, Springer.
P-P. COMBES, Th. MAYER and J-F. THISSE (eds.) (2008), *Economic geography: the integration of regions and nations*. Princeton, Princeton University Press.
J. HINDRIKS (ed.) (2008), *Au-delà de Copernic: de la confusion au consensus* ? Brussels, Academic and Scientific Publishers.
J-M. HURIOT and J-F. THISSE (eds) (2009), *Economics of cities*. Cambridge, Cambridge University Press.
P. BELLEFLAMME and M. PEITZ (eds) (2010), *Industrial organization: markets and strategies*. Cambridge University Press.
M. JUNGER, Th. LIEBLING, D. NADDEF, G. NEMHAUSER, W. PULLEYBLANK, G. REINELT, G. RINALDI and L. WOLSEY (eds) (2010), *50 years of integer programming, 1958-2008: from the early years to the state-of-the-art*. Berlin Springer.

CORE Lecture Series

C. GOURIÉROUX and A. MONFORT (1995), Simulation Based Econometric Methods.
A. RUBINSTEIN (1996), Lectures on Modeling Bounded Rationality.
J. RENEGAR (1999), A Mathematical View of Interior-Point Methods in Convex Optimization.
B.D. BERNHEIM and M.D. WHINSTON (1999), Anticompetitive Exclusion and Foreclosure Through Vertical Agreements.
D. BIENSTOCK (2001), Potential function methods for approximately solving linear programming problems: theory and practice.
R. AMIR (2002), Supermodularity and complementarity in economics.
R. WEISMANTEL (2006), Lectures on mixed nonlinear programming.