Rank-one matrix completion is solved by the sum-of-squares relaxation of order two

Augustin Cosse
Harvard, IACS and UCL, ICTEAM
School of engineering and applied science
Email: acosse@seas.harvard.edu

Laurent Demanet
Massachusetts Institute of Technology
Department of Mathematics
Email: laurent@math.mit.edu

***Abstract*—This note studies the problem of nonsymmetric rank-one matrix completion. We show that in every instance where the problem has a unique solution, one can recover the original matrix through the second round of the sum-of-squares/Lasserre hierarchy with minimization of the trace of the moments matrix. Our proof system is based on iteratively building a sum of $N-1$ linearly independent squares, where N is the number of monomials of degree at most two, corresponding to the canonical basis $(\mathbf{z}^{\alpha}-\mathbf{z}_0^{\alpha})^2$. Those squares are constructed from the ideal $\mathcal{I}$ generated by the constraints and the monomials provided by the minimization of the trace.**

I. NON SYMMETRIC MATRIX COMPLETION

This paper introduces a deterministic recovery result for non symmetric rank-1 matrix completion by using the Lasserre hierarchy of semidefinite programming relaxations. To our knowledge, the closest result in the current literature is [1] where the authors use the same hierarchy and certify recovery in the case of tensor decomposition. Our paper also shares its deterministic nature with [2] where the authors derive recovery from the spectral properties of a graph Laplacian. Finally, this paper can also be related to [3] in which the authors study noisy tensor completion and use the sixth round of the Lasserre hierarchy to derive probabilistic recovery guarantees.

We will use $\mathcal{M}(r; m\times n)$ to denote the set of matrices of rank r. This set is an algebraic determinantal variety that can be completely characterized through the vanishing of the $(r+1)$-minors. This determinantal variety has dimension $(m+n-r)r$.

The general nonsymmetric rank-1 matrix completion problem consists in recovering an unknown matrix $\boldsymbol{X} \in \mathcal{M}(1; m\times n)$ such that $\boldsymbol{X} = \boldsymbol{x}\boldsymbol{y}^T$, given a fixed subset of its entries [4],

$$\begin{aligned} \text{find} \quad & \boldsymbol{X} \\ \text{subject to} \quad & \operatorname{rank}(\boldsymbol{X}) = 1 \\ & X_{ij} = A_{ij} \quad (i,j) \in \Omega. \end{aligned} \tag{1}$$

As a slight abuse, we also speak of constraints $X_{ij} = A_{ij}$ as belonging to Ω. In relation to problem (1), we introduce the mapping $\mathcal{R}_\Omega : \mathbb{R}^{m\times n} \to \mathbb{R}^{|\Omega|}$ that corresponds to extracting the observed entries of the matrix. We let $\mathcal{R}_\Omega^1$ denote the restriction of $\mathcal{R}_\Omega$ to matrices of rank-1, i.e $\mathcal{R}_\Omega^1 : \mathcal{M}(1; m\times n) \to \mathbb{R}^{|\Omega|}$. Invertibility of this restriction $\mathcal{R}_\Omega^1$ is a natural question. In other words, when can one uniquely recover the matrix $\boldsymbol{X}$ from the knowledge of $\mathcal{R}_\Omega(\boldsymbol{X})$ and the fact that $\boldsymbol{X}$ has rank 1 ?

In particular, this paper considers the completion problem on $\mathcal{M}^*(1, m\times n)$, where $\mathcal{M}^*(1, m\times n)$ denotes the restriction of $\mathcal{M}(1; m\times n)$ to matrices for which none of the entries are zero. The reason for this is that if a rank-1 matrix has a zero element, then the corresponding row or column will be zero, and it is easy to see that the completion problem will generically lack injectivity.

Respectively denote by $\mathcal{V}_1$, $\mathcal{V}_2$ the row and column indices of $\boldsymbol{X}$. We consider the bipartite graph $\mathcal{G}(\mathcal{V}_1, \mathcal{V}_2, \mathcal{E})$ associated to problem (1), where the set of edges in the graph is defined by $(i,j) \in \mathcal{E}$ iff $(i,j) \in \Omega$. The conditions for the recovery of the matrix $\boldsymbol{X}$ from the set Ω are related to the properties of this bipartite graph. In particular, we have the following proposition (see [5]).

Proposition 1 (Rank-1 completion): The mask $\mathcal{R}_\Omega$ is injective on $\mathcal{M}^*(1; m\times n)$ if and only if $\mathcal{G}$ is connected.

Without loss of generality, we will now restrict to matrices given by the product $\boldsymbol{x}\boldsymbol{y}^T$, with $x_i, y_j \neq 0$ for all i, j and where the first element of $\boldsymbol{x}$ has been normalized to unity to enforce unique recovery. Let us then define $\boldsymbol{X}$ as

$$\boldsymbol{X} = \begin{pmatrix} y_1 & \dots & y_n \\ y_1 x_1 & \dots & y_n x_1 \\ \vdots & \ddots & \vdots \\ y_1 x_{m-1} & \dots & y_n x_{m-1} \end{pmatrix} = \boldsymbol{x}\boldsymbol{y}^T \tag{2}$$

The vertices of the bipartite graph $\mathcal{G}$ corresponding to $\boldsymbol{X}$ will thus be labeled by the corresponding row and column indices, and the edges by the elements of $\boldsymbol{X}$ lying in Ω. When we deal with the rank-1 case, an implication of proposition 1 is that for all x_n, y_m, the bipartite graph corresponding to the mask Ω always contains at least one connected path starting with an edge corresponding to an element of the first row and for which the series of existing edges corresponds to running through $\boldsymbol{X}$ according to chains of constraints such as

$$y_{i_1} \to y_{i_1}x_{i_2} \to x_{i_2}y_{i_3} \to \dots y_{i_{L+1}}x_n \qquad (x_n) \tag{3}$$

$$y_{i'_1} \to y_{i'_1}x_{i'_2} \to x_{i'_2}y_{i'_3} \to \dots x_{i'_{L'+1}}y_m \qquad (y_m) \tag{4}$$

In other words, each one of the entries of $\boldsymbol{x}$ and $\boldsymbol{y}$ can always be related to an element of the first row through chains similar to (3) or (4).

Following proposition 1, given Ω, and letting $\boldsymbol{x_0}\boldsymbol{y}_0^T$ denote the optimal solution $\boldsymbol{X}_0$ to problem (1), the rank-1 matrix completion problem can be stated in polynomial form as

$$\begin{aligned} \text{find} \quad & \boldsymbol{x},\ \boldsymbol{y} \\ \text{subject to} \quad & y_\ell = (y_0)_\ell \quad (1,\ell) \in \Omega \\ & x_m y_n = ((x_0)_m (y_0)_n) \quad (m+1, n) \in \Omega. \end{aligned} \tag{5}$$

and can be solved by iteratively propagating the value of the elements of the first row up to any of the elements of $\boldsymbol{x}$ or $\boldsymbol{y}$ through paths like (3) or (4). This "propagation" algorithm is clearly not suited for handling noisy data.

II. Semidefinite programming relaxations

For clarity, we now use $\boldsymbol{z}$ to represent the whole vector of unknowns, $\boldsymbol{z} = (\boldsymbol{x},\ \boldsymbol{y})$.

Let $\mathbb{R}[\mathbf{z}]$ denote the corresponding ring of multivariate polynomials. Let $\mathbb{N}_t^K$ denote the set of n-tuples $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_K) \in \mathbb{N}^K$ such that $|\boldsymbol{\alpha}| \equiv \sum_i \alpha_i \leq t$. A general form for problem (5) is given by

$$\begin{array}{ll} \text{minimize} & p(\boldsymbol{z}) \\ \text{subject to} & h_1(\boldsymbol{z}) = 0, \ldots, h_L(\boldsymbol{z}) = 0, \end{array} \tag{6}$$

where $h_1, \ldots, h_L$ denote polynomials in the variables $x_1, \ldots x_{m-1}$ as well as $y_1, \ldots, y_n$. We will use this general form to introduce the Lasserre sum-of-squares hierarchy.

For a (positive Borel) measure μ on $\mathbb{R}^K$, and a multi-index $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_K)$, one can define the corresponding moment $m_{\boldsymbol{\alpha}} = \int \mathbf{z}^{\boldsymbol{\alpha}} \mu(d\mathbf{z})$, where $\mathbf{z}^{\boldsymbol{\alpha}}$ denotes the product $z_1^{\alpha_1} \ldots z_K^{\alpha_K}$. The sequence of moments for the measure μ is then the sequence $(m_{\boldsymbol{\alpha}})_{\boldsymbol{\alpha} \in \mathbb{N}^K}$ of moments corresponding to the whole set of monomials $\mathbf{z}^{\boldsymbol{\alpha}}$. For any finite $t \in \mathbb{N}$, one can also introduce the truncated sequence of moments $(m_{\boldsymbol{\alpha}})_{\boldsymbol{\alpha} \in \mathbb{N}_t^n}$ defined only for $|\boldsymbol{\alpha}| \leq t$.

Given the sequence of moments $(m_{\boldsymbol{\alpha}})_{\boldsymbol{\alpha} \in \mathbb{N}^K} \in \mathbb{R}^{|\mathbb{N}^K|}$, the corresponding *moment matrix* is defined as the matrix indexed by the K-tuples $\boldsymbol{\alpha}, \boldsymbol{\beta}$ of $\mathbb{N}^K$ and whose $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ entry is defined as $m_{\boldsymbol{\alpha}+\boldsymbol{\beta}}$. This matrix can be truncated as well in the same fashion as the truncated sequence of moments. The resulting finite matrix $\boldsymbol{M}_t(\boldsymbol{m})$ is then defined for $\max(|\boldsymbol{\alpha}|, |\boldsymbol{\beta}|) \leq t$ as $(M_t)_{\boldsymbol{\alpha},\boldsymbol{\beta}} = m_{\boldsymbol{\alpha}+\boldsymbol{\beta}}$. As an example, consider the moments matrix of order 2, $\boldsymbol{M}_2(\boldsymbol{m})$:

$$\boldsymbol{M}_2 = \left[\begin{array}{c|cc|ccc} m_{(00)} & m_{(10)} & m_{(01)} & m_{(11)} & m_{(20)} & m_{(02)} \\ \hline m_{(10)} & m_{(20)} & m_{(11)} & \cdots & & m_{(12)} \\ m_{(01)} & m_{(11)} & & & & \\ \hline m_{(11)} & \vdots & & \ddots & & \vdots \\ m_{(20)} & & & & & \\ m_{(02)} & m_{(12)} & & \cdots & & m_{(22)} \end{array}\right].$$

A polynomial $h \in \mathbb{R}[\mathbf{z}]$ can be represented as a sequence $h = (h_{\boldsymbol{\alpha}})_{\boldsymbol{\alpha}}$, where $h_{\boldsymbol{\alpha}}$ denotes the coefficient of h corresponding to the monomial $\mathbf{z}^{\boldsymbol{\alpha}}$, so that $h(z) = \sum_{\boldsymbol{\alpha}} h_{\boldsymbol{\alpha}} \mathbf{z}^{\boldsymbol{\alpha}}$. We can then define the *shifted moments sequence* hm through the product $hm = \boldsymbol{M}(m)h \in \mathbb{R}^{\mathbb{N}^K}$, i.e., $(hm)_{\boldsymbol{\alpha}} = \sum_{\boldsymbol{\beta}} h_{\boldsymbol{\beta}} m_{\boldsymbol{\alpha}+\boldsymbol{\beta}}$. Such sequences can also be truncated by limiting their index to $|\boldsymbol{\alpha}| \leq t$.

Given these notions, we can introduce the t^{th} round of the Lasserre sum-of-squares hierarchy of relaxations for the general polynomial problem (6), as

$$\begin{aligned} p^* = \inf_{m \in \mathbb{R}^{|\mathbb{N}_{2t}^K|}} p^T m \quad s.t. \quad & m_0 = 1, \quad \boldsymbol{M}_t(m) \succeq 0, \\ & \boldsymbol{M}_{t-d_{h_j}}(h_j m) = 0 \\ & (j = 1, \ldots, L), \end{aligned} \tag{7}$$

where $d_{h_j} := \lceil \deg(h_j)/2 \rceil$. The Lasserre hierarchy thus optimizes over the measures $\mu(d\mathbf{z})$ rather than over $\boldsymbol{z}$, and constraints the moments in order to recover the Dirac measure $\delta(\boldsymbol{z} - \boldsymbol{z}_0)$ leading to $\boldsymbol{z}_0$. The third constraint in Eq. (7) is the general way of encoding all the constraints $p(z)h_\ell(z) = 0$, with $p \in \mathbb{R}[\mathbf{z}]$, of order less than $2t$, in the space of measures. For more on the Lasserre hierarchy, see [6], [7], [8].

III. Main Result and mathematical argument

Since a simple propagation argument can solve the noiseless rank-1 completion problem, it seems reasonable to hope for a robust algorithm with a similar complexity which could also be extended to noisy measurements. The main result of this paper, is that such an algorithm is given by the second round of (7), i.e., $t = 2$, combined with a minimization of the trace norm.

Theorem 1: Consider problem (1) in the context of proposition 1, with $\boldsymbol{X} \in \mathcal{M}^*(1; m \times n)$. Then this problem can be solved exactly through the Lasserre hierarchy of order $t = 2$, under minimization of the trace of $\boldsymbol{M}_2(\boldsymbol{m})$.

To prove theorem 1, we will start with traditional ideas from convex optimization theory. We start by writing problem (7) with a trace objective in the general form

$$\begin{array}{ll} \text{minimize} & \text{Tr}(\boldsymbol{M}) \\ \text{subject to} & \mathcal{A}(\boldsymbol{M}) = \boldsymbol{b}, \\ & \boldsymbol{M} \succeq 0. \end{array} \tag{8}$$

Let $\boldsymbol{m}_0$ be the vector of moments up to order 2 for the 1-atomic measure corresponding to the point $\boldsymbol{z}_0$, i.e, $\boldsymbol{m}_0 = (\boldsymbol{z}_0^{\boldsymbol{\alpha}})_{\boldsymbol{\alpha}}$ where $\boldsymbol{z}_0 = (\boldsymbol{x}_0, \boldsymbol{y}_0)$. Recall that $\boldsymbol{z}_0 = (\boldsymbol{x}_0, \boldsymbol{y}_0)$ defines the optimal solution $\boldsymbol{X}_0 = \boldsymbol{x}_0 \boldsymbol{y}_0^T$ of the matrix completion problem so that the knowledge of $\boldsymbol{m}_0$ gives $\boldsymbol{X}_0$. We will denote the corresponding optimal second order moments matrix as $\boldsymbol{M}_0 = (\boldsymbol{M}_0)_2 = \boldsymbol{m}_0 \boldsymbol{m}_0^T$.

To ensure unique recovery of the matrix $\boldsymbol{M}_0$ from a problem like (8), traditional convex optimization proofs are based on satisfying the first order optimality conditions[1] by exhibiting a dual vector $\tilde{\lambda}$ such that $-\mathcal{A}^* \tilde{\lambda} - \boldsymbol{I} \in \partial \imath_{\mathcal{K}}(\boldsymbol{M}_0)$ where $\imath_{\mathcal{K}}$ denotes the indicator function of the positive semidefinite (psd) cone (see for example [4]). We start by providing the general conditions for the existence of such a certificate in the case of problem (7). We then show how one can construct such a certificate satisfying those conditions for the particular case of problem (5).

Let us reformulate the relaxation (7) of (6) in a more explicit fashion. We first introduce the appropriate matrices $\boldsymbol{B}_{\boldsymbol{\gamma}}$ that encode the monomial $\mathbf{z}^{\boldsymbol{\gamma}}$ as $\boldsymbol{\mu}_{\boldsymbol{z}}^T \boldsymbol{B}_{\boldsymbol{\gamma}} \boldsymbol{\mu}_{\boldsymbol{z}}$, where $\boldsymbol{\mu}_z$ is used to denote the vector of monomials of degree less than t, $\boldsymbol{\mu}_z = (\mathbf{z}^{\boldsymbol{\alpha}})_{\boldsymbol{\alpha} \in \mathbb{N}_t^K}$. (see [7] for more details on the structure of

[1] Note that in the case of convex optimization those conditions are necessary and sufficient.

those matrices), problem (7) can be stated as

$$\begin{array}{ll} \text{minimize} & \mathrm{Tr}(\boldsymbol{M}) \\ \text{subject to} & \sum_{\boldsymbol{\zeta}} \frac{(h_\ell)_{\boldsymbol{\zeta}}}{\|\boldsymbol{B}_{\boldsymbol{\zeta}+\boldsymbol{\kappa}}\|_F^2}\langle \boldsymbol{M}, \boldsymbol{B}_{\boldsymbol{\zeta}+\boldsymbol{\kappa}}\rangle + (h_\ell)_0 = 0. \\ & \text{for} \quad \boldsymbol{\kappa} \in \mathbb{N}^K_{2(t-d_{h_\ell})}, 1 \le \ell \le L \\ & \boldsymbol{M} \succeq 0, \quad \boldsymbol{M} = \sum_{\boldsymbol{\gamma}} m_{\boldsymbol{\gamma}} \boldsymbol{B}_{\boldsymbol{\gamma}} + \boldsymbol{e}_1\boldsymbol{e}_1^T. \end{array} \tag{9}$$

The first sum is taken over all the coefficients of each constraint $h_\ell(x) = 0$, $h_\ell(x) = \sum_{\boldsymbol{\zeta}}(h_\ell)_{\boldsymbol{\zeta}}\mathbf{x}^{\boldsymbol{\zeta}}$ and the second sum is taken over all moments $m_{\boldsymbol{\gamma}}$ of order up to 4.

We now derive the first order optimality conditions $-\mathcal{A}^*\tilde{\lambda} - \boldsymbol{I} \in \partial\imath_{\mathcal{K}}(\boldsymbol{M}_0)$ for problem (9) by writing down the Lagrangian dual function for this problem, and by finding a dual vector $\tilde{\lambda}$ such that $0 \in \partial\mathcal{L}(\boldsymbol{M}_0, \tilde{\lambda})$. Introducing multipliers for each of the polynomial constraints, the Lagrangian dual function can be written as

$$\begin{aligned} \mathcal{L}(\boldsymbol{M}, \boldsymbol{\lambda}, \boldsymbol{\xi}) &= \mathrm{Tr}(\boldsymbol{M}) + \langle \boldsymbol{M} - \sum_{\boldsymbol{\gamma}} m_{\boldsymbol{\gamma}}\boldsymbol{B}_{\boldsymbol{\gamma}} - \boldsymbol{e}_1\boldsymbol{e}_1^T, \boldsymbol{\xi}\rangle \\ &+ \sum_\ell \sum_{\boldsymbol{\kappa}\in\mathbb{N}^K_{2(t-d_{h_\ell})}} \lambda_{\ell,\boldsymbol{\kappa}} \left(\sum_{\boldsymbol{\zeta}} \frac{(h_\ell)_{\boldsymbol{\zeta}}}{\|\boldsymbol{B}_{\boldsymbol{\zeta}+\boldsymbol{\kappa}}\|_F^2}\langle \boldsymbol{B}_{\boldsymbol{\zeta}+\boldsymbol{\kappa}}, \boldsymbol{M}\rangle\right) \\ &+ \imath_{\mathcal{K}}(\boldsymbol{M}). \end{aligned}$$

The multipliers $\lambda_{\ell,\boldsymbol{\kappa}}$ correspond to each of the original and shifted polynomial constraints while $\boldsymbol{\xi}$ encode the redudant structure of the matrix $\boldsymbol{M}$. Usual convex optimization theory states that $\boldsymbol{M}_0 = \boldsymbol{m}_0\boldsymbol{m}_0^T$ is a minimizer for problem (9) iff one can find dual vectors $(\boldsymbol{\xi}, \boldsymbol{\lambda})$ such that $0 \in \partial\mathcal{L}(\boldsymbol{M}_0, \boldsymbol{\lambda}, \boldsymbol{\xi})$. This leads to the following conditions on $\boldsymbol{\xi}, \boldsymbol{\lambda}$. Let

$$T = \left\{\boldsymbol{m}_0\boldsymbol{v}^T + \boldsymbol{v}\boldsymbol{m}_0^T, \quad \boldsymbol{v} \in \mathbb{R}^{|\mathbb{N}_2^K|}\right\},$$

$T^\perp$ being its orthogonal complement, and let Z_T denote the projection of Z onto the subspace T. Then, $\lambda_{\ell,\boldsymbol{\kappa}}$ and $\boldsymbol{\xi}$ combine into a dual certificate Z, and together must obey

1) $Z = -\boldsymbol{I} - \boldsymbol{\xi} - \sum_\ell \sum_{\boldsymbol{\kappa}\in\mathbb{N}^K_{2(t-d_{h_\ell})}} \lambda_{\ell\boldsymbol{\kappa}} \left(\sum_{\boldsymbol{\zeta}} \frac{(h_\ell)_{\boldsymbol{\zeta}}}{\|\boldsymbol{B}_{\boldsymbol{\zeta}+\boldsymbol{\kappa}}\|_F^2}\boldsymbol{B}_{\boldsymbol{\zeta}+\boldsymbol{\kappa}}\right)$
2) $Z_T = 0, \quad Z_{T^\perp} \preceq 0$
3) $\langle \boldsymbol{B}_{\boldsymbol{\gamma}}, \boldsymbol{\xi}\rangle = 0, \qquad \forall \boldsymbol{\gamma} \neq 0.$

The conditions $Z_T = 0$ and $Z_{T^\perp} \preceq 0$ arise from requiring Z to be a subgradient of the indicator of the psd cone (see [9] for more details).

The following proposition guarantees unique recovery.

Proposition 2: To ensure *unique* recovery of $\boldsymbol{M}_0$, in addition to the conditions 1), 2), and 3) mentioned above, it is sufficient to require $Z_{T^\perp} \prec 0$ as well as injectivity on T of the linear constraints $\mathcal{A}(\boldsymbol{M}) = \boldsymbol{b}$ arising from the polynomial constraints $h_\ell(\boldsymbol{z}) = 0$.

Proof: Let us decompose Z into $Z = -\boldsymbol{I} + Z'$

$$\begin{aligned} \mathrm{Tr}(\boldsymbol{M}_0) &= \langle I, \boldsymbol{M}_0\rangle = \langle I_T, \boldsymbol{M}_0\rangle = \langle Z'_T, \boldsymbol{M}_0\rangle \\ &= \langle Z', \boldsymbol{M}_0 - \boldsymbol{M}\rangle + \langle Z', \boldsymbol{M}\rangle = \langle Z', \boldsymbol{M}\rangle \\ &= \langle I_T, \boldsymbol{M}_T\rangle + \langle (Z')_{T^\perp}, \boldsymbol{M}\rangle \\ &= \mathrm{Tr}(\boldsymbol{M}_T) + \langle (Z')_{T^\perp}, \boldsymbol{M}\rangle \\ &< \mathrm{Tr}(\boldsymbol{M}) \qquad \text{for } \boldsymbol{M}_{T^\perp} \neq 0 \end{aligned}$$

The second line comes from the fact that Z' belongs to the range of $\mathcal{A}^*$ and $\mathcal{A}(\boldsymbol{M}) = \mathcal{A}(\boldsymbol{M}_0)$. The last inequality is due to $Z'_{T^\perp} \prec I_{T^\perp}$ which since $\boldsymbol{M} \succeq 0$ implies $\langle (Z')_{T^\perp}, \boldsymbol{M}\rangle < \mathrm{Tr}(\boldsymbol{M}_{T^\perp})$ for $\boldsymbol{M}_{T^\perp} \neq 0$. The last inequality thus implies $\boldsymbol{M}_{T^\perp} = 0$. Finally $\boldsymbol{M}_T = (\boldsymbol{M}_0)_T$ by injectivity of the constraints on T.

■

Note that, to satisfy $Z_{T^\perp} \prec 0$ and $Z_T = 0$, it is sufficient to ask for $\boldsymbol{m}_0 \in \mathrm{Null}(Z)$ and to require $-Z$ to be psd and exact rank $|\mathbb{N}_2^K| - 1$. In the next section, we show how an equivalent polynomial form can help us construct a dual certificate satisfying those conditions.

Equivalent polynomial certificate

We will call sum-of-squares (SOS), any polynomial $p(z)$ for which there exists a decomposition $p(z) = \sum_{j=1}^m s_j^2(z)$ for some polynomials $s_j \in \mathbb{R}[\boldsymbol{z}]$. Introducing a polynomial version of proposition 2, requires the following lemma from [6] relating SOS and semidefinite programming (SDP),

Lemma 1 (Equivalence between SOS and SDP): Let $p \in \mathbb{R}[\boldsymbol{z}]$ with $p = \sum_{\boldsymbol{\alpha}\in\mathbb{N}_{2t}^K} p_{\boldsymbol{\alpha}}\mathbf{z}^{\boldsymbol{\alpha}}$ be a polynomial of degree $\le 2t$, the following assertions are equivalent,

i) p is a sum-of-squares
ii) There exists a positive semidefinite matrix $\boldsymbol{X}$ such that

$$p(z) = \boldsymbol{\mu}_z^T\boldsymbol{X}\boldsymbol{\mu}_z, \tag{10}$$

The conditions of proposition 2, together with lemma 1 imply the following proposition arising from the polynomial nature of problem (6),

Proposition 3 (Polynomial Form): To ensure *unique recovery* of $\boldsymbol{M}_0$, in addition to the injectivity of the constraints on T, it is sufficient to find a sum of $(|\mathbb{N}_2^K| - 1)$ linearly independent squares $s_j^2(z)$ of degree less than or equal to 4, polynomials $\lambda_\ell(z)$ of degree less than or equal to $4 - 2d_{h_\ell}$ and constant ρ such that

$$q(z) = \sum_j s_j^2(z) = \sum_{\boldsymbol{\alpha}\in\mathbb{N}_2^K} \mathbf{z}^{2\boldsymbol{\alpha}} - \rho + \sum_\ell h_\ell(z)\lambda_\ell(z), \tag{11}$$

and such that $q(z_0) = 0$.

Since $q(z)$ is SOS, to satisfy $Z_T = 0$ is suffices to require $s_j(z_0) = 0$. Indeed we have

$$\{s_j(z_0) = 0\} \iff \{\langle \sum_j \boldsymbol{s}_j\boldsymbol{s}_j^T, \boldsymbol{m}_0\boldsymbol{y}^T + \boldsymbol{y}\boldsymbol{m}_0^T\rangle, \boldsymbol{y} \in \mathbb{R}^{|\mathbb{N}_2^K|}\} = 0.$$

The value of the constant ρ is fixed by enforcing $q(z_0) = 0$. The last term on the RHS of (11) is a contribution of degree ≤ 4 from the ideal $\mathcal{I} := \{\sum_{j=1}^L u_jh_j \mid u_1, \ldots, u_L \in \mathbb{R}[\boldsymbol{z}]\}$ generated from the constraints h_j.

IV. Construction of the Certificate

Remember that $\boldsymbol{z}$ is given by the concatenation $\boldsymbol{z} = (\boldsymbol{x},\ \boldsymbol{y})$ of all first order monomials arising in problem (5). Our construction of the certificate is based on generating the canonical squares $(\mathbf{z}^{\boldsymbol{\alpha}} - \mathbf{z}_0^{\boldsymbol{\alpha}})^2$ for all $|\boldsymbol{\alpha}| \leq 2$ from the ideal; the squared monomials arising from the trace norm and the constant ρ.

First let us show that for all monomials $\mathbf{z}^{\boldsymbol{\alpha}}$ with $|\boldsymbol{\alpha}| = 1$ one can build the polynomial $-2\mathbf{z}^{\boldsymbol{\alpha}}\mathbf{z}_0^{\boldsymbol{\alpha}} + 2(\mathbf{z}_0^{\boldsymbol{\alpha}})^2$ by using a decomposition from the ideal of degree at most 3.

- Either the constraint $\mathbf{z}^{\boldsymbol{\alpha}} = \mathbf{z}_0^{\boldsymbol{\alpha}}$ is present explicitly ($\mathbf{z}^{\boldsymbol{\alpha}} = y_\ell$ corresponds to an element of the first row of $\boldsymbol{X}$ and $h_\ell(z) \equiv y_\ell - (y_0)_\ell$ is a constraint in Ω) and one can then just multiply this constraint by $-2(\mathbf{z}_0^{\boldsymbol{\alpha}})$ to get the desired polynomial $-2(\mathbf{z}_0)^{\boldsymbol{\alpha}}\mathbf{z}^{\boldsymbol{\alpha}} + 2(\mathbf{z}_0^{\boldsymbol{\alpha}})^2$
- Or the first order monomial $\mathbf{z}^{\boldsymbol{\alpha}}$, $|\boldsymbol{\alpha}| = 1$, appears in a chain like (3) or (4).

Let $\mathbf{z}^{\boldsymbol{\alpha}} = z_\ell$. Since the graph is connected, there exists a chain

$$z_{i_1} \rightarrow z_{i_1}z_{i_2} \rightarrow z_{i_2}z_{i_3} \rightarrow z_{i_3}z_{i_4} \ldots \rightarrow z_{i_{\ell-1}}z_\ell \tag{12}$$

such that if we denote the corresponding numerical values by $a_{i_1}, a_{i_1}a_{i_2}, \ldots, a_{i_{\ell-1}}a_\ell$, the constraints $z_{i_1} - a_{i_1}, \ldots, z_{i_\ell - 1}z_\ell - a_\ell a_{i_{\ell-1}}$ belong to Ω and thus to the ideal $\mathcal{I}$.

Using (12), one can recursively combine the elements of the chain in the following way,

$$\begin{aligned} a_{i_{\ell-2}}a_{i_{\ell-1}}(z_\ell - a_\ell) =& (z_\ell z_{i_{\ell-1}} - a_\ell a_{i_{\ell-1}})z_{i_{\ell-2}} \\ &- (z_{i_{\ell-2}}z_{i_{\ell-1}} - a_{i_{\ell-2}}a_{i_{\ell-1}})z_\ell \\ &+ a_\ell a_{i_{\ell-1}}(z_{i_{\ell-2}} - a_{i_{\ell-2}}). \end{aligned}$$

This telescoping relation holds for all ℓ throughout the chain until the second element, (z_{i_2}), for which we have $a_{i_1}(z_{i_2} - a_{i_2}) = (z_{i_2}z_{i_1} - a_{i_2}a_{i_1}) - z_{i_2}(z_{i_1} - a_{i_1}) \in \mathcal{I}$. The key here is that one can make use of the bilinear constraints to get a propagation argument which remains degree-3 since the multiplicative factor $a_\ell a_{i_{\ell-1}}$ in front of the propagation term $(z_{i_{\ell-2}} - a_{i_{\ell-2}})$ remains constant.

Now that we can build the polynomials $-2a_k z_k + 2a_k^2$ for all k as degree-3 decompositions from the ideal $\mathcal{I}$, one can just add those polynomials to the frace and constant ρ contributions $(z_k^2 - a_k^2)$ in order to get the squares $(z_k - a_k)^2$. We thus get $|\mathbb{N}_1^K| - 1$ of the required squares.

The remaining $\binom{K}{2}$ decompositions for the second order squared polynomials $(\mathbf{z}^{\boldsymbol{\alpha}} - \mathbf{z}_0^{\boldsymbol{\alpha}})^2$ for $|\boldsymbol{\alpha}| = 2$, are built from the first order decompositions, the trace, and constant ρ as follows. $\forall \boldsymbol{\alpha}, \boldsymbol{\beta}$ with $|\boldsymbol{\alpha}|, |\boldsymbol{\beta}| = 1$,

$$\begin{aligned} (\mathbf{z}^{\boldsymbol{\alpha}}\mathbf{z}^{\boldsymbol{\beta}} - \mathbf{z}_0^{\boldsymbol{\alpha}}\mathbf{z}_0^{\boldsymbol{\beta}})^2 =& (\mathbf{z}^{\boldsymbol{\alpha}}\mathbf{z}^{\boldsymbol{\beta}})^2 - (\mathbf{z}_0^{\boldsymbol{\alpha}}\mathbf{z}_0^{\boldsymbol{\beta}})^2 \\ &- 2\mathbf{z}_0^{\boldsymbol{\alpha}}\mathbf{z}_0^{\boldsymbol{\beta}}(\mathbf{z}^{\boldsymbol{\alpha}}\mathbf{z}^{\boldsymbol{\beta}} - \mathbf{z}_0^{\boldsymbol{\alpha}}\mathbf{z}_0^{\boldsymbol{\beta}}), \end{aligned} \tag{13}$$

where the first two terms arise from the contribution of the trace and ρ, and the third one can be expressed from the ideal $\mathcal{I}$ with degree at most 4, as

$$\begin{aligned} -2\mathbf{z}_0^{\boldsymbol{\alpha}}\mathbf{z}_0^{\boldsymbol{\beta}}(\mathbf{z}^{\boldsymbol{\alpha}}\mathbf{z}^{\boldsymbol{\beta}} - \mathbf{z}_0^{\boldsymbol{\alpha}}\mathbf{z}_0^{\boldsymbol{\beta}}) =& (-2\mathbf{z}^{\boldsymbol{\alpha}}\mathbf{z}_0^{\boldsymbol{\alpha}} + 2(\mathbf{z}_0^{\boldsymbol{\alpha}})^2)(\mathbf{z}_0^{\boldsymbol{\beta}})^2 \\ &+ (\mathbf{z}^{\boldsymbol{\beta}} - \mathbf{z}_0^{\boldsymbol{\beta}})(-2\mathbf{z}^{\boldsymbol{\alpha}}\mathbf{z}_0^{\boldsymbol{\alpha}})\mathbf{z}_0^{\boldsymbol{\beta}} \end{aligned} \tag{14}$$

The first term is of degree at most 3 and the second one is of degree at most 4.

Injectivity of the polynomial constraints on T

To conclude, we show that the linear map $\mathcal{A}$ is injective on T. For this purpose, let us show that the nullspace of $\mathcal{A}$ is empty on T. Let us consider any $\boldsymbol{Z} = \boldsymbol{m}_0\boldsymbol{v}^T + \boldsymbol{v}\boldsymbol{m}_0^T$. Normalization of $\boldsymbol{Z}_{11}$ implies $v_1 = 0$ and reduces $\boldsymbol{Z}$ to a matrix for which the first column equals the first row and is given by $(v_2 \ldots, v_{|\mathbb{N}_2^K|})$.

Then recall that there is a least one constraint setting to zero one of the elements of the first column. So there exists ℓ s.t. $v_\ell = 0$. Accordingly the whole corresponding row and column reduce to $((m_0)_\ell v_k)_{k \leq |\mathbb{N}_2^K|}$. Since $(m_0)_\ell \neq 0$[2], one can then apply the next constraint $z_\ell z_m = 0$ which implies $v_m = 0$. By recursively applying this idea, one can show that the first block of $\boldsymbol{Z}$ corresponding to the monomials of degree at most 2 is zero. The remaining part of the matrix can then be set to 0 as well trough the structural constraints (equality of corresponding monomials) for the first row/column and then using the fact that $\boldsymbol{Z}$ is defined as $\boldsymbol{m}_0\boldsymbol{v}^T + \boldsymbol{v}\boldsymbol{m}_0^T$.

[2]Recall that we assumed $(\boldsymbol{X}_0)_{ij} \neq 0$ for all (i, j)

V. Conclusion

In this note we show that rank-1 matrix completion can be solved exactly from the second round of the Lasserre hierarchy as soon as it meets the minimal necessary conditions on the measurements for solvability. To the best of our knowledge, it is the first time that a deterministic certificate is constructed for a higher level (> 1) of convex relaxation for a completion problem.

Acknowledgment

AC thanks the Belgian National Science Foundation (FNRS) and MISTI Belgium for support. LD thanks ONR, AFOSR, NSF, and Total SA for support.

References

[1] G. Tang and P. Shah, "Guaranteed tensor decomposition: A moment approach," in *Proceedings of The 32nd International Conference on Machine Learning*, 2015, pp. 1491–1500.

[2] L. Demanet and V. Jugnon, "Convex recovery from interferometric measurements," *arXiv preprint arXiv:1307.6864*, 2013.

[3] B. Barak and A. Moitra, "Tensor prediction, rademacher complexity and random 3-xor," *arXiv preprint arXiv:1501.06521*, 2015.

[4] E. J. Candès and B. Recht, "Exact matrix completion via convex optimization," *Foundations of Computational mathematics*, vol. 9, no. 6, pp. 717–772, 2009.

[5] F. Király and R. Tomioka, "A combinatorial algebraic approach for the identifiability of low-rank matrix completion," *arXiv preprint arXiv:1206.6470*, 2012.

[6] M. Laurent, "Sums of squares, moment matrices and optimization over polynomials," in *Emerging applications of algebraic geometry*. Springer, 2009, pp. 157–270.

[7] J. B. Lasserre, "Global optimization with polynomials and the problem of moments," *SIAM Journal on Optimization*, vol. 11, no. 3, pp. 796–817, 2001.

[8] P. A. Parrilo, "Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization," Ph.D. dissertation, Citeseer, 2000.

[9] E. J. Candes, T. Strohmer, and V. Voroninski, "Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming," *Communications on Pure and Applied Mathematics*, vol. 66, no. 8, pp. 1241–1274, 2013.