A Non-Convex Relaxation for Fixed-Rank Approximation

Carl Olsson[1,2]   Marcus Carlsson[2]   Erik Bylow[2]

[1]Department of Electrical Engineering
Chalmers University of Technology

[2]Centre for Mathematical Sciences
Lund University

Abstract

This paper considers the problem of finding a low rank matrix from observations of linear combinations of its elements. It is well known that if the problem fulfills a restricted isometry property (RIP), convex relaxations using the nuclear norm typically work well and come with theoretical performance guarantees. On the other hand these formulations suffer from a shrinking bias that can severely degrade the solution in the presence of noise.

In this theoretical paper we study an alternative non-convex relaxation that in contrast to the nuclear norm does not penalize the leading singular values and thereby avoids this bias. We show that despite its non-convexity the proposed formulation will in many cases have a single stationary point if a RIP holds. Our numerical tests show that our approach typically converges to a better solution than nuclear norm based alternatives even in cases when the RIP does not hold. [1]

1. Introduction

Low rank approximation is an important tool in applications such as rigid and non rigid structure from motion, photometric stereo and optical flow [30, 5, 31, 14, 2, 13]. The rank of the approximating matrix typically describes the complexity of the solution. For example, in non-rigid structure from motion the rank measures the number of basis elements needed to describe the point motions [5]. Under the assumption of Gaussian noise the objective is typically to solve

$$\min_{\text{rank}(X) \le r} \|X - X_0\|_F^2, \tag{1}$$

where X_0 is a measurement matrix and $\|\cdot\|_F$ is the Frobenius norm. The problem can be solved optimally using SVD [10], but the strategy is limited to problems where all matrix

elements are directly measured. In this paper we will consider low rank approximation problems where linear combinations of the elements are observed. We aim to solve problems of the form

$$\min_X \mathbb{I}(\text{ rank}(X) \le r) + \|\mathcal{A}X - \mathbf{b}\|^2. \tag{2}$$

Here $\mathbb{I}(\text{ rank}(X) \le r)$ is 0 if $\text{ rank}(X) \le r$ and ∞ otherwise. The linear operator $\mathcal{A} : \mathbb{R}^{m \times n} \to \mathbb{R}^p$ is assumed to fulfill a restricted isometry property (RIP) [29]

$$(1 - \delta_q)\|X\|_F^2 \le \|\mathcal{A}X\|^2 \le (1 + \delta_q)\|X\|_F^2, \tag{3}$$

for all matrices with $\text{ rank}(X) \le q$. The standard approach for problems of this class is to replace the rank function with the convex nuclear norm $\|X\|_* = \sum_i x_i$, where x_i, $i = 1, ..., N$ are the singular values of X, [29, 6]. It was first observed that this is the convex envelope of the rank function over the set $\{X; x_1 \le 1\}$ in [12]. Since then a number of generalizations that give performance guarantees for the nuclear norm relaxation have appeared, e.g. [29, 27, 6, 7]. The approach does however suffer from a shrinking bias that can severely degrade the solution in the presence of noise. In contrast to the rank constraint the nuclear norm penalizes both small singular values of X, assumed to stem from measurement noise, and large singular values, assumed to make up the true signal, equally. In some sense the suppression of noise also requires an equal suppression of signal. Non-convex alternatives have been shown to improve performance [26, 23, 19, 24].

In this paper we will consider the relaxation

$$\min_X \mathcal{R}_r(X) + \|\mathcal{A}X - \mathbf{b}\|^2, \tag{4}$$

where

$$\mathcal{R}_r(X) = \max_Z \sum_{i=r+1}^{N} z_i^2 - \|X - Z\|_F^2, \tag{5}$$

and z_i, $i = 1, ..., N$ are the singular values of Z. The minimization over Z does not have any closed form solution, however it was shown in [21, 1] how to efficiently evaluate

and compute its proximal operator. Figure 1 shows a three dimensional illustration of the level sets of the regularizer. In [21, 20] it was shown that
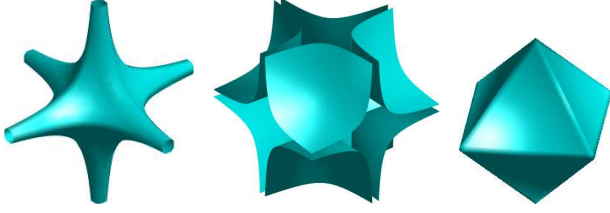

Figure 1: Level set surfaces $\{X \mid \mathcal{R}_r(X) = \alpha\}$ for $X = \operatorname{diag}(x_1, x_2, x_3)$ with $r = 1$ (*Left*) and $r = 2$ (*Middle*). Note that when $r = 1$ the regularizer promotes solutions where only one of x_k is non-zero. For $r = 2$ the regularizer instead favors solutions with two non-zero x_k. For comparison we also include the nuclear norm. (*Right*)

$$\mathcal{R}_r(X) + \|X - X_0\|_F^2, \qquad (6)$$

is the the convex envelope of

$$\mathbb{I}(\operatorname{rank}(X) \leq r) + \|X - X_0\|_F^2. \qquad (7)$$

By itself the regularization term $R_r(X)$ is not convex, but when adding a quadratic term $\|X\|_F^2$ the result is convex. It is shown in [1] that (7) and (6) have the same optimizer as long as the singular values of X_0 are distinct. (When this is not the case the minimizer is not unique.) Assuming that (3) holds $\|\mathcal{A}X\|^2$ will behave roughly like $\|X\|_F$ for matrices of rank less than q, and therefore it seems reasonable that (4) should have some convexity properties. In this paper we study the stationary points of (4). We show that if a RIP property holds it is in many cases possible to guarantee that any stationary point of (4) (with rank r) is unique.

A number of recent works propose to use the related $\|\cdot\|_{r*}$-norm [22, 11, 18, 17] (sometimes referred to as the spectral k-support norm). This is a generalization of the nuclear norm which is obtained when selecting $r = 1$. It can be shown [22] that the extreme points of the unit ball with this norm are rank r matrices. Therefore this choice may be more appropriate than the nuclear norm when searching for solutions of a particular (known) rank. It can be seen (e.g. from the derivations in [18]) that $\|X\|_{r*}$ is the convex envelope of (7) when $X_0 = 0$, which gives $\|X\|_{r*}^2 = \mathcal{R}_r(X) + \|X\|_F^2$. While the approach is convex the extra norm penalty adds a (usually) unwanted shrinking bias similar to what the nuclear norm does. In contrast, our approach avoids this bias since it uses a non-convex regularizer. Despite this non-convexity we are still able to derive strong optimality guarantees for an important class of problem instances.

1.1. Main Results and Contributions

Our main result, Theorem 2.4, shows that if X_s is a stationary point of (4) and the singular values z_i of the matrix $Z = (I - \mathcal{A}^*A)X_s + \mathcal{A}^*\mathbf{b}$ fulfill $z_{r+1} < (1 - 2\delta_{2r})z_r$ then there can not be any other stationary point with rank less than or equal to r. The matrix Z is related to the gradient of the objective function at X_s (see Section 2). The term $\|X - Z\|^2$ can be seen as a local approximation of $\|\mathcal{A}X - \mathbf{b}\|^2 =$

$$\|X\|_F^2 - \langle X, (I - \mathcal{A}^*\mathcal{A})X \rangle - 2\langle X, \mathcal{A}^*\mathbf{b} \rangle + \mathbf{b}^T\mathbf{b}. \quad (8)$$

Replacing $\langle X, (I - \mathcal{A}^*\mathcal{A})X \rangle$ with its first order Taylor expansion $2\langle X, (I - \mathcal{A}^*\mathcal{A})X_s \rangle$ (ignoring the constant term) reduces (8) to $\|X - Z\|^2 + C$, where C is a constant.

If for example there is a rank r matrix X_0 such that $\mathbf{b} = \mathcal{A}X_0$ then it is easy to show that X_0 is a stationary point and the corresponding Z is identical to X_0. Since this means that $z_{r+1} = 0$ our results certify that this is the only stationary point to the problem if $\mathcal{A}$ fulfills (3) with $\delta_{2r} < \frac{1}{2}$. The following lemma clarifies the connection between the stationary point X_s and Z.

Lemma 1.1. *The point X_s is stationary in* $F(X) = \mathcal{R}_r(X) + \|\mathcal{A}X - \mathbf{b}\|^2$ *if and only if* $2Z \in \partial G(X_s)$*, where* $G(X) = \mathcal{R}_r(X) + \|X\|_F^2$*, and if and only if*

$$X_s \in \arg\min_X \mathcal{R}_r(X) + \|X - Z\|_F^2. \qquad (9)$$

(The proof is identical to that of Lemma 3.1 in [25].) In [1] it is shown that as long as $z_r \neq z_{r+1}$ the unique solution of (9) is the best rank r approximation of Z. When there are several singular values that are equal to z_r, (9) will have multiple solutions and some of them will not be of rank r. Note however, that since (9) is the convex envelope of (7) (with $X_0 = Z$) the solution set of (9) is the convex hull of that of (7), and therefore there are still solutions of rank r to (9) even when $z_r = z_{r+1}$.

Stationary points of F will therefore generally be rank r approximations of Z. In this situation the first r singular values of Z coincide with those of X while the remaining ones are related to the residual error. Hence, loosely speaking our results state that if the error residuals are small compared to X_s this stationary point is likely to be unique.

Our work builds on that of [25] which derives similar results for the non-convex regularizer $\mathcal{R}_\mu(X) = \sum_i \mu - \max(\sqrt{\mu} - x_i)^2$, where x_i are the singular values of X. In this case a trade-off between rank and residual error is optimized using the formulation

$$\min_X \mathcal{R}_\mu(X) + \|\mathcal{A}X - \mathbf{b}\|^2. \qquad (10)$$

While it can be argued that (10) and (4) are essentially equivalent since we can iteratively search for a μ that gives

the desired rank, the results of [25] may not rule out the existence of multiple high rank stationary points. In contrast, when using (4) our results imply that if $z_{r+1} < (1-2\delta_{2r})z_r$ then X_s is the unique stationary point of the problem. (To see this, note that if there are other stationary points we can by the preceding discussion assume that at least one is of rank r or less which contradicts our main result in Theorem 2.4). Hence, in this sense our results are stronger than those of [25] and allow for directly searching for a matrix of the desired rank with an essentially parameter free formulation.

In [8] the relationship between minimizers of (4) and (2) is studied. Among other things [8] shows that if $\|\mathcal{A}\| \le 1$ then any local minimizer of (4) is also a local minimizer of (2), and that their global minimizers coincide. Hence results about the stationary points of (4) are also relevant to the original objective (2).

Some very recent papers [4, 28, 15] parametrize X using UV^T, where both U and V have r columns. They show that it is possible to bound the distance between the global and any local minimum in terms of the residual error $\|\mathcal{A}(UV^T)-b\|^2$ if a RIP holds. In contrast to our results this does however not rule out the existence of multiple minima (except in the noise free case).

1.2. Notation

In this section we introduce some preliminary material and notation. In general we will use boldface to denote a vector $\mathbf{x}$ and its ith element x_i. Unless otherwise stated the singular values of a matrix X will be denoted x_i and the vector of singular values $\mathbf{x}$. By $\|\mathbf{x}\|$ we denote the standard euclidean norm $\|\mathbf{x}\| = \sqrt{\mathbf{x}^T\mathbf{x}}$. A diagonal matrix with diagonal elements $\mathbf{x}$ will be denoted $D_{\mathbf{x}}$. For matrices we define the scalar product as $\langle X, Y \rangle = \text{tr}(X^T Y)$, where tr is the trace function, and the Frobenius norm $\|X\|_F = \sqrt{\langle X, X \rangle} = \sqrt{\sum_{i=1}^n x_i^2}$. The adjoint of the linear matrix operator $\mathcal{A}$ is denoted $\mathcal{A}^*$. By $\partial F(X)$ we mean the set of subgradients of the function F at X and by a stationary point we mean a solution to $0 \in \partial F(X)$.

2. Optimality Conditions

Let $F(X) = \mathcal{R}_r(X) + \|\mathcal{A}X - \mathbf{b}\|^2$. We can equivalently write

$$F(X) = G(X) - \delta_q\|X\|_F^2 + H(X) + \|\mathbf{b}\|^2, \quad (11)$$

where $G(X) = \mathcal{R}_r(X) + \|X\|_F^2$ and $H(X) = \delta_q\|X\|_F^2 + \left(\|\mathcal{A}X\|^2 - \|X\|_2^2\right) - 2\langle \mathcal{A}X, \mathbf{b}\rangle$. The function G is convex and sub-differentiable. Any stationary point X_s of F therefore has to fulfill

$$2\delta_q X_s - \nabla H(X_s) \in \partial G(X_s). \quad (12)$$

Computation of the gradient gives the optimality conditions $2Z \in \partial G(X_s)$ where $Z = (I - \mathcal{A}^*\mathcal{A})X_s + \mathcal{A}^*\mathbf{b}$.

2.1. Subgradients of G

For our analysis we need to determine the subdifferential $\partial G(X)$ of the function $G(X)$. Let $\mathbf{x}$ be the vector of singular values of X and $X = UD_{\mathbf{x}}V^T$ be the SVD. Using Von Neumann's trace theorem it is easy to see [21] that the Z that maximizes (5) has to be of the form $Z = UD_{\mathbf{z}}V^T$, where $\mathbf{z}$ are singular values. If we let

$$L(X, Z) = -\sum_{i=1}^r z_i^2 + 2\langle Z, X\rangle, \quad (13)$$

then we have $G(X) = \max_Z L(X, Z)$. The function L is linear in X and concave in Z. Furthermore for any given X the corresponding maximizers can be restricted to a compact set (because of the dominating quadratic term). By Danskin's Theorem, see [3], the subgradients of G are then given by

$$\partial G(X) = \text{convhull}\{\nabla_X L(X, Z), Z \in \mathcal{Z}(X)\}, \quad (14)$$

where $\mathcal{Z}(X) = \arg\max_Z L(X, Z)$. We note that by concavity the maximizing set $\mathcal{Z}(X)$ is convex. Since $\nabla_X L(X, Z) = 2Z$ we get

$$\partial G(X) = 2\arg\max_Z L(X, Z). \quad (15)$$

To find the set of subgradients we thus need to determine all maximizers of L. Since the maximizing Z has the same U and V as X what remains is to determine the singular values of Z. It can be shown [21] that these have the form

$$z_i \in \begin{cases} \max(x_i, s) & i \le r \\ s & i \ge r, \ x_i \ne 0 \ . \\ [0, s] & i > r, \ x_i = 0 \end{cases} \quad (16)$$

for some number $s \ge x_r$. (The case $x_i = 0$, $i > r$ is actually not addressed in [21]. However, it is easy to see that any value in $[0, s]$ works since z_i vanishes from (13) when $x_i = 0$, $i > r$. In fact, any value $[-s, s]$ works, but we use the convention that singular values are positive. Note that the columns of U that correspond to zero singular values of X are not uniquely defined. We can always achieve a decreasing sequence with $z_i \in [0, s]$ by changing signs and switching order.)

For a general matrix X the value of s can not be determined analytically but has to be computed numerically by maximizing a one dimensional concave and differentiable function [21]. If $\text{rank}(X) \le r$ it is however clear that the optimal choice is $s = x_r$. To see this we note that since the optimal Z is of the form $UD_{\mathbf{z}}V^T$ we have

$$L(X, Z) = -\sum_{i=1}^r (z_i - x_i)^2 + \sum_{i=1}^r x_i^2, \quad (17)$$

if $\text{rank}(x) \leq r$. Selecting $s = x_r$ and inserting (16) into (17) gives $L(X, Z) = \sum_{i=1}^{r} x_i^2$, which is clearly the maximum. Hence if $\text{rank}(X) = r$ we conclude that the subgradients of g are given by $2Z = 2UD_\mathbf{z}V^T$ where

$$z_i \in \begin{cases} x_i & i \leq r \\ [0, x_r] & i \geq r \end{cases}. \tag{18}$$

2.2. Growth estimates for the $\partial G(X)$

Next we derive a bound on the growth of the subgradients that will be useful when considering the uniqueness of low rank stationary points.

Let $\mathbf{x}$ and $\mathbf{x}'$ be two vectors both with at most r non-zero (positive) elements, and I and I' be the indexes of the r largest elements of $\mathbf{x}$ and $\mathbf{x}'$ respectively. We will assume that both I and I' contain r elements. If in particular $\mathbf{x}'$ has fewer than r non-zero elements we also include some zero elements in I'. We define the corresponding sequences $\mathbf{z}$ and $\mathbf{z}'$ by

$$z_i \in \begin{cases} x_i & i \in I \\ [0, \overline{s}] & i \notin I \end{cases}, \quad z_i' \in \begin{cases} x_i' & i \in I' \\ [0, \overline{s}'] & i \notin I' \end{cases}, \tag{19}$$

where $\overline{s} = \min_{i \in I} x_i$ and $\overline{s}' = \min_{i \in I} x_i'$. If $\mathbf{x}'$ has fewer than r non-zero elements then $\overline{s}' = 0$. Note that we do not require that the elements of the $\mathbf{x}, \mathbf{x}', \mathbf{z}$ and $\mathbf{z}'$ vectors are ordered in decreasing order. We will see later (Lemma 2.2) that in order to estimate the effects of the U and V matrices we need to be able to handle permutations of the singular values. For our analysis we will also use the quantity $\underline{s} = \max_{i \notin I} z_i$.

Lemma 2.1. *If $\underline{s} < c\overline{s}$, where $0 < c < 1$ then*

$$\langle \mathbf{z}' - \mathbf{z}, \mathbf{x}' - \mathbf{x} \rangle > \frac{1-c}{2}\|\mathbf{x}' - \mathbf{x}\|^2. \tag{20}$$

Proof. Since $z_i = x_i$ when $i \in I$ and $x_i = 0$ otherwise, we can write the inner product $\langle \mathbf{z}' - \mathbf{z}, \mathbf{x}' - \mathbf{x} \rangle$ as

$$\sum_{\substack{i \in I \\ i \in I'}} (x_i - x_i')^2 + \sum_{\substack{i \in I \\ i \notin I'}} x_i(x_i - z_i') + \sum_{\substack{i \notin I \\ i \in I'}} x_i'(x_i' - z_i). \tag{21}$$

Note that $\|\mathbf{x}' - \mathbf{x}\|^2 = $

$$\sum_{\substack{i \in I \\ i \in I'}} (x_i - x_i')^2 + \sum_{\substack{i \in I \\ i \notin I'}} x_i^2 + \sum_{\substack{i \notin I \\ i \in I'}} x_i'^2. \tag{22}$$

Since the second and third sum in (21) have the same number of terms it suffices to show that

$$x_i(x_i - z_i') + x_j'(x_j' - z_j) \geq \frac{1-c}{2}(x_i^2 + x_j'^2), \tag{23}$$

when $i \in I$, $i \notin I'$ and $j \notin I$, $j \in I'$. By the assumption $\underline{s} < c\overline{s}$ we know that $z_j < cx_i$. We further know that $z_i' \leq \overline{s}' \leq x_j'$. This gives

$$x_i z_i' \leq x_i x_j' \leq \frac{x_i^2 + x_j'^2}{2}, \tag{24}$$

$$x_j' z_j < cx_j' x_i \leq c\frac{x_i^2 + x_j'^2}{2}, \tag{25}$$

Inserting these inequalities into the left hand side of (23) gives the desired bound. $\qquad \square$

The above result gives an estimate of the growth of the subdifferential in terms of the singular values. To derive a similar estimate for the matrix elements we need the following lemma:

Lemma 2.2. *Let $\mathbf{x}, \mathbf{x}', \mathbf{z}, \mathbf{z}'$ be fixed vectors with non-increasing and non-negative elements such that $\mathbf{x} \neq \mathbf{x}'$ and $\mathbf{z}$ and $\mathbf{z}'$ fulfill (16) (with $\mathbf{x}$ and $\mathbf{x}'$ respectively). Define $X' = U'D_{\mathbf{x}'}V'^T$, $X = UD_\mathbf{x}V^T$, $Z' = U'D_{\mathbf{z}'}V'^T$, and $Z = UD_\mathbf{z}V^T$ as functions of unknown orthogonal matrices U, V, U' and V'. If*

$$a^* = \min_{U,V,U',V'} \frac{\langle Z' - Z, X' - X \rangle}{\|X' - X\|_F^2} \leq 1 \tag{26}$$

then

$$a^* = \min_{M_\pi} \frac{\langle M_\pi \mathbf{z}' - \mathbf{z}, M_\pi \mathbf{x}' - \mathbf{x} \rangle}{\|M_\pi \mathbf{x}' - \mathbf{x}\|^2}, \tag{27}$$

where M_π belongs to the set of permutation matrices.

The proof is almost identical to that of Lemma 4.1 in [25] and therefore we omit it. While our subdifferential is different to the one studied in [25], for both of them we have that the $\mathbf{z}$ and $\mathbf{x}$ vectors fulfill $z_i \geq x_i \geq 0$ which is the only property that is used in the proof.

Corollary 2.3. *Assume that X is of rank r and $2Z \in \partial G(X)$. If the singular values of the matrix Z fulfill $z_{r+1} < cz_r$, where $0 < c < 1$, then for any $2Z' \in \partial G(X')$ with $\text{rank}(X') \leq r$ we have*

$$\langle Z' - Z, X' - X \rangle > \frac{1-c}{2}\|X' - X\|_F^2, \tag{28}$$

as long as $\|X' - X\|_F \neq 0$.

Proof. We let $\mathbf{x}, \mathbf{x}', \mathbf{z}$ and $\mathbf{z}'$ be the singular values of the matrices X, X', Z and Z' respectively. Our proof essentially follows that of Corollary 4.2 in [25], where a similar result is first proven under the assumption that $\mathbf{x} \neq \mathbf{x}'$ and then generalized to the general case using a continuity argument. For this purpose we need to extend the infeasible interval somewhat. Since $0 < c < 1$ and $z_{r+1} < cz_r$ are open there is an $\epsilon > 0$ such that $z_{r+1} < (c - \epsilon)z_r$ and

$0 < c - \epsilon < 1$. Now assume that $a^* > 1$ in (26), then clearly

$$\langle Z' - Z, X' - X \rangle > \frac{1 - c + \epsilon}{2} \|X' - X\|_F^2, \qquad (29)$$

since $\frac{1-c+\epsilon}{2} < 1$. Otherwise $a^* \leq 1$ and we have

$$\frac{\langle Z' - Z, X' - X \rangle}{\|X' - X\|_F^2} \geq \frac{\langle M_\pi \mathbf{z}' - \mathbf{z}, M_\pi \mathbf{x}' - \mathbf{x} \rangle}{\|M_\pi \mathbf{x}' - \mathbf{x}\|^2}. \qquad (30)$$

According to Lemma 2.1 the right hand side is strictly larger than $\frac{1-c+\epsilon}{2}$, which proves that (29) holds for all X' with $\mathbf{x}' \neq \mathbf{x}$.

It remains to show that

$$\langle Z' - Z, X' - X \rangle \geq \frac{1 - c + \epsilon}{2} \|X' - X\|_F^2, \qquad (31)$$

for the case $\mathbf{x}' = \mathbf{x}$ and $\|X' - X\|_F \neq 0$. Since $\epsilon > 0$ is arbitrary this proves the Corollary. This can be done as in [25] using continuity of the scalar product and the Frobenius norm. Specifically, a sequence $X(t) \to X$, when $t \to 0$, is defined by modifying the largest singular value and letting $\sigma_1(X(t)) = \sigma_1(X) + t$. It is easy to verify that $X(t)$ fulfills (29) for every $t > 0$. Letting $t \to 0$ then proves (31). $\qquad \square$

2.3. Uniqueness of Low Rank Stationary Points

In this section we show that if a RIP (3) holds and the singular values z_r and z_{r+1} are well separated there can only be one stationary point of F that has rank r. We first derive a bound on the gradients of H. We have

$$\nabla H(X) = 2\delta_q X + 2(\mathcal{A}^* \mathcal{A} - I)X - 2\mathcal{A}^* \mathbf{b}. \qquad (32)$$

This gives $\langle \nabla H(X') - \nabla H(X), X' - X \rangle =$

$$2\delta_q \|X' - X\|_F^2 + 2(\|\mathcal{A}(X' - X)\|^2 - \|X' - X\|_F^2). \qquad (33)$$

By (3) $\left| \|\mathcal{A}(X' - X)\|^2 - \|X' - X\|_F^2 \right| \leq \delta_q \|X' - X\|_F^2$, if $\text{rank}(X' - X) \leq q$ which gives

$$\langle \nabla H(X') - \nabla H(X), X' - X \rangle \geq 0 \qquad (34)$$

This leads us to our main result

Theorem 2.4. *Assume that X_s is a stationary point of F, that is, $(I - \mathcal{A}^* \mathcal{A})X_s + \mathcal{A}^* \mathbf{b} = Z$, where $2Z \in \partial G(X_s)$, $\text{rank}(X_s) = r$ and the singular values of Z fulfill $z_{r+1} < (1 - 2\delta_{2r})z_r$. If X_s' is another stationary point then $\text{rank}(X_s') > r$.*

Proof. Assume that $\text{rank}(X_s') \leq r$. Since both X_s and X_s' are stationary we have

$$2\delta_{2r} X_s' - \nabla H(X_s') = 2Z', \qquad (35)$$

$$2\delta_{2r} X_s - \nabla H(X_s) = 2Z, \qquad (36)$$

where $2Z \in \partial G(X_s)$ and $2Z' \in \partial G(X_s')$. Taking the difference between the two equations yields

$$2\delta_{2r}(X_s' - X_s) - \nabla H(X_s') + \nabla H(X_s) = 2Z' - 2Z, \quad (37)$$

which implies

$$2\delta_{2r}\|V\|_F^2 - \langle \nabla H(X_s') - \nabla H(X_s), V \rangle = 2\langle Z' - Z, V \rangle, \qquad (38)$$

where $V = X_s' - X_s$ has $\text{rank}(V) \leq 2r$. By (34) the left hand side is less than $2\delta_{2r}\|V\|_F^2$. However, according to Corollary 2.3 (with $c = 1 - 2\delta_{2r}$) the right hand side is larger than $2\delta_{2r}\|V\|_F^2$ which contradicts $\text{rank}(X_s') \leq r$. $\qquad \square$

Remark. Note that as mentioned in Section 1.1 the existence of a second stationary point of rank larger than r also implies the existence of a second rank r stationary point. Therefore under the conditions of the theorem X_s will be unique over all ranks.

3. Implementation and Experiments

In this section we test the proposed approach on some simple real and synthetic applications (some that fulfill (3) and some that do not). For our implementation we use the GIST approach from [16] because of its simplicity. Given a current iterate X_k this method solves

$$X_{k+1} = \arg\min_X \mathcal{R}_r(X) + \tau_k \|X - M_k\|^2, \qquad (39)$$

where $M_k = X_k - \frac{1}{\tau_k}(\mathcal{A}^* \mathcal{A} X_k - \mathcal{A}^* \mathbf{b})$. Note that if $\tau_k = 1$ then any fixed point of (39) is a stationary point by Lemma 1.1. To solve (39) we use the proximal operator computed in [21].

Our algorithm consists of repeatedly solving (39) for a sequence of $\{\tau_k\}$. We start from a larger value ($\tau_0 = 5$ in our implementation) and reduce towards 1 as long as this results in decreasing objective values. Specifically we set $\tau_{k+1} = \frac{\tau_k - 1}{1.1} + 1$ if the previous step was successful in reducing the objective value. Otherwise we increase τ according to $\tau_{k+1} = 1.5(\tau_k - 1) + 1$. We outline the approach in Algorithm 1.

3.1. Synthetic Data

We first evaluate the quality of the relaxation on a number of synthetic experiments. We compare the two formulations (4) and

$$\min_X \mu\|X\|_* + \|\mathcal{A}X - \mathbf{b}\|^2. \qquad (40)$$

In Figure 2 (a) we tested these two relaxations on a number of synthetic problems with varying noise levels. The data was created so that the operator $\mathcal{A}$ fulfills (3) with $\delta = 0.2$. By column stacking an $m \times n$ matrix X the linear mapping $\mathcal{A}$ can be represented with a matrix A of size
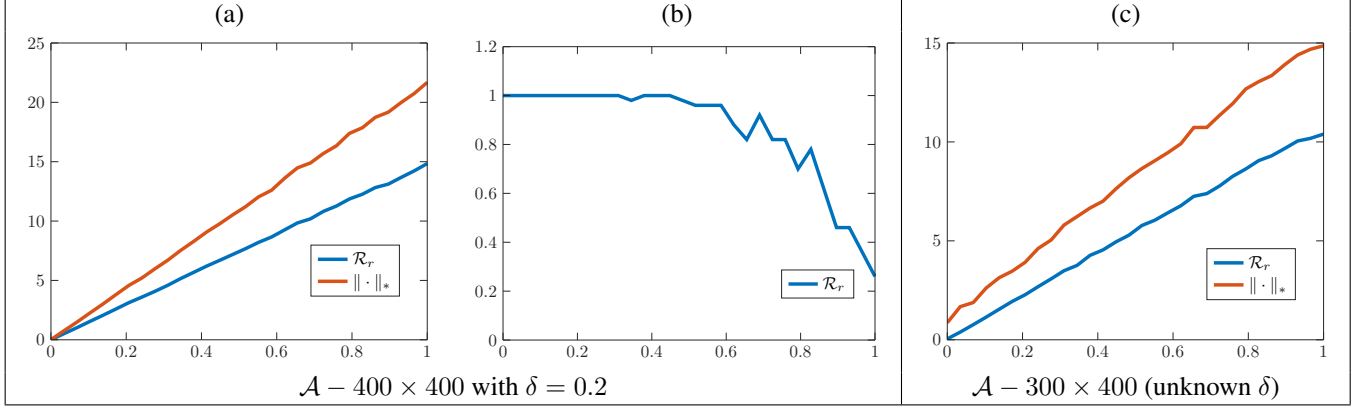
Figure 2: (a) - Noise level (x-axis) vs. data fit $\|\mathcal{A}X - \mathbf{b}\|$ (y-axis) for the solutions obtained with (4) and (40). (b) - Fraction of instances where the solution of (4) could be verified to be globally optimal. (c) - Same as (a). (a) and (b) uses 400×400 $\mathcal{A}$ with $\delta = 0.2$ while (c) uses 300×400 $\mathcal{A}$.

Data: $\mathcal{A}, \mathbf{b}, i_{\max}$
Result: X_k
$X_k := 0, \tau := 5$;
for $i = 1, ..., i_{max}$ **do**
    $M_k := X_k - \frac{1}{\tau}(\mathcal{A}^*\mathcal{A}X_k - \mathcal{A}^*\mathbf{b})$;
    $X_{k+1} = \arg\min_X \mathcal{R}_r(X) + \tau_k \|X - M_k\|^2$;
    if $F(X_{k+1}) > F(X_k)$ **then**
        $\tau := 1.5(\tau - 1) + 1$;
    else
        $\tau := \frac{\tau - 1}{1.1} + 1$;
        $X_k = X_{k+1}$;
    end
end

Algorithm 1: Optimization of (4).

$p \times mn$. It is easy to see that if we let $p = mn$ the term $(1 - \delta_q)$ of (3) will be the same as the smallest singular value of A squared. (In this case the RIP constraint will hold for any rank and we therefore suppress the subscript q and only use δ.) For the data in Figure 3 (a) we selected 400×400 matrices A with random $\mathcal{N}(0,1)$ (gaussian mean 0 and variance 1) entries and modified their singular values. We then generated 20×20 matrices X of rank 5 by sampling 20×5 matrices U and V with $\mathcal{N}(0,1)$ entries and computing $X = UV^T$. The measurement vector $\mathbf{b}$ was created by computing $\mathbf{b} = \mathcal{A}X + \boldsymbol{\epsilon}$, where $\boldsymbol{\epsilon}$ is $\mathcal{N}(0,\sigma^2)$ for varying noise levels σ between 0 and 1. In Figure 2 (a) we plotted the measurement fit $\|\mathcal{A}X - \mathbf{b}\|$ versus the noise level σ for the solutions obtained with (4) and (40). Note that since the formulation (40) does not directly specify the rank of the sought matrix we iteratively searched for the smallest value of μ that gives the correct rank, using a bisection scheme. The reason for choosing the smallest μ is that this reduces

the shrinking bias to a minimum while it still gives the correct rank.

In Figure 2 (b) we computed the Z matrix and plotted the fraction of problem instances where its singular values fulfilled $z_{r+1} < (1 - 2\delta)z_r$, with $\delta = 0.2$. For these instances the obtained stationary points are also globally optimal according to our main results.

In Figure 2 (c) we did the same experiment as in (a) but with an under determined A of size 300×400. It is known [29] that if A is $p \times mn$ and the elements of A are drawn from $\mathcal{N}(0, \frac{1}{p})$ then $\mathcal{A}$ fulfills (3) with high probability. The exact value of δ_q is however difficult to determine and therefore we are not able to verify optimality in this case.

3.2. Non-Rigid Structure from Motion

In this section we consider the problem of Non-Rigid Structure from Motion. We follow the aproach of Dai. *et al.* [9] and let

$$X = \begin{bmatrix} X_1 \\ Y_1 \\ Z_1 \\ \vdots \\ X_F \\ Y_F \\ Z_F \end{bmatrix} \text{ and } X^{\#} = \begin{bmatrix} X_1 & Y_1 & Z_1 \\ \vdots & \vdots & \vdots \\ X_F & Y_F & Z_F \end{bmatrix}, \quad (41)$$

where X_i, Y_i, Z_i are $1 \times m$ matrices containing the x-,y- and z-coordinates of the tracked points in images i. Under the assumption of an orthographic camera the projection of the $3D$ points can be modeled using $M = RX$, where R is a $2F \times 3F$ block diagonal matrix with 2×3 blocks R_i, consisting of two orthogonal rows that encode the camera orientation in image i. The resulting $2F \times m$ measurement
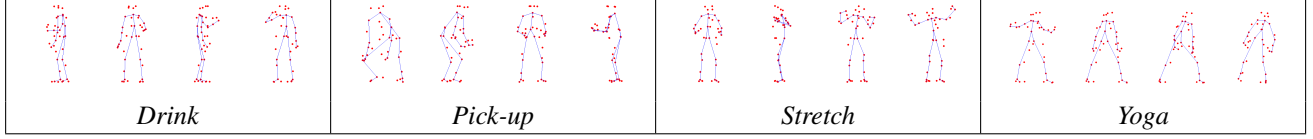
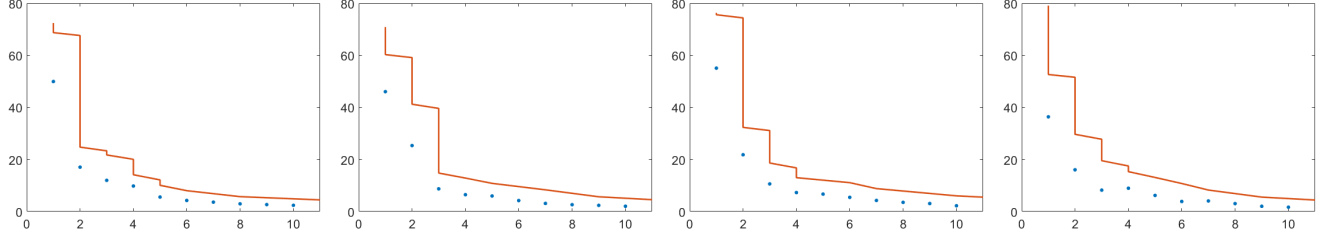Figure 3: Four images from each of the MOCAP data sets.

Figure 4: Results obtained with (44) (blue dots) and (45) (orage curve) for the four sequences. Data fit $\|RX - M\|_F$ (y-axis) versus rank($X^\#$) (x-axis).
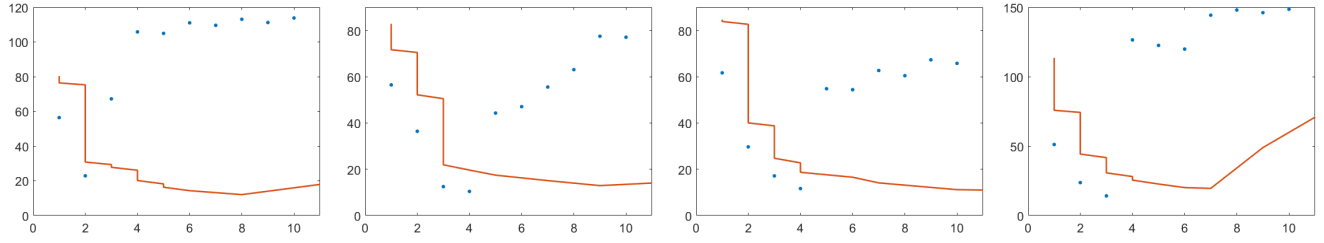

Figure 5: Results obtained with (44) (blue dots) and (45) (orage curve) for the four sequences. Distance to ground truth $\|X - X_{gt}\|_F$ (y-axis) versus rank($X^\#$) (x-axis).
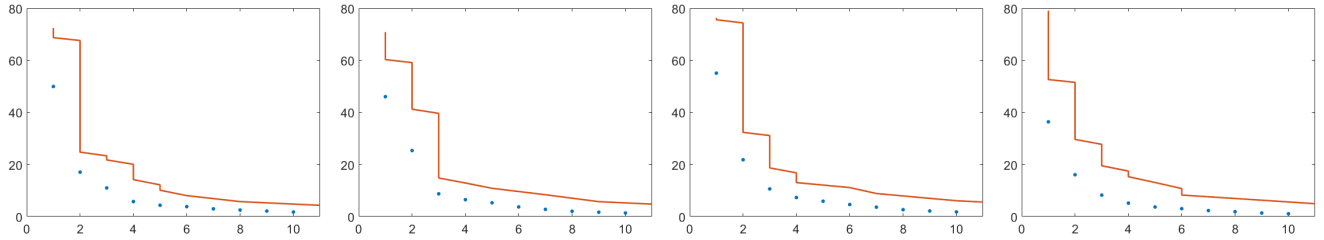

Figure 6: Results obtained with (46) (blue dots) and (47) (orage curve) for the four sequences. Data fit $\|RX - M\|_F$ (y-axis) versus rank($X^\#$) (x-axis).
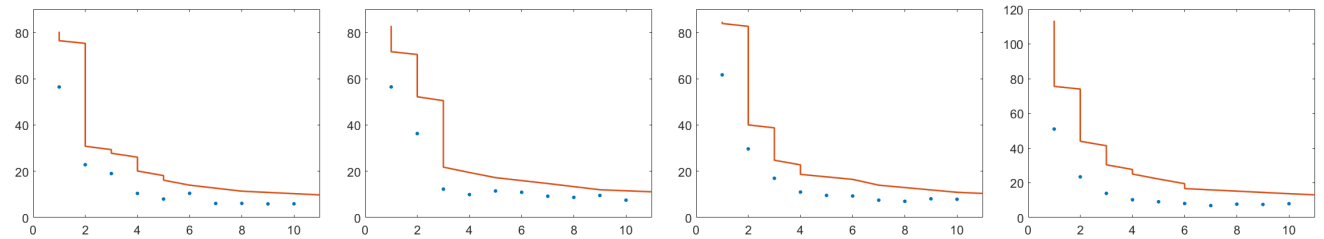

Figure 7: Results obtained with (46) (blue dots) and (47) (orage curve) for the four sequences. Distance to ground truth $\|X - X_{gt}\|_F$ (y-axis) versus rank($X^\#$) (x-axis) is plotted for various regularization strengths.

matrix M consists of the x- and y-image coordinates or the tracked points. Under the assumption of a linear shape basis model [5] with r deformation modes, the matrix $X^{\#}$ can be factorized into $X^{\#} = CB$, where the $r \times 3m$ matrix B contain the basis elements. It is clear that such a factorization is possible when $X^{\#}$ is of rank r. We therefore search for the matrix $X^{\#}$ of rank r that minimizes the residual error $\|PX - M\|_F^2$.

The linear operator defined by $\mathcal{A}(X^{\#}) = RX$ does by itself not obey (3) since there are typically low rank matrices in its nullspace. This can be seen by noting that if N_i is the 3×1 vector perpendicular to the two rows of R_i, that is $R_i N_i = 0$ then

$$\begin{bmatrix} X_i \\ Y_i \\ Z_i \end{bmatrix} = N_i C_i, \qquad (42)$$

where C_i is any $1 \times m$ matrix, is in the null space of R_i. Therefore any matrix of the form

$$N(C) = \begin{bmatrix} n_{11}C_1 & n_{21}C_1 & n_{31}C_1 \\ n_{12}C_2 & n_{22}C_2 & n_{32}C_2 \\ \vdots & \vdots & \vdots \\ n_{1F}C_F & n_{2F}C_F & n_{3F}C_F \end{bmatrix}, \qquad (43)$$

where n_{ij} are the elements of N_i, vanishes under $\mathcal{A}$. Setting everything but the first row of $N(C)$ to zero shows that there is a matrix of rank 1 in the null space of $\mathcal{A}$. Moreover, if the rows of the optimal $X^{\#}$ spans such a matrix it will not be unique since we may add $N(C)$ without affecting the projections or the rank.

In Figure 4 we compare the two relaxations

$$\mathcal{R}_r(X^{\#}) + \|RX - M\|_F^2 \qquad (44)$$

and

$$\mu\|X^{\#}\|_* + \|RX - M\|_F^2 \qquad (45)$$

on the four MOCAP sequences displayed in Figure 3, obtained from [9]. These consist of real motion capture data and therefore the ground truth solution is only approximatively of low rank.

In Figure 4 we plot the rank of the obtained solution versus the datafit $\|RX - M\|_F$. Since (45) does not allow us to directly specify the rank of the sought matrix, we solved the problem for 50 values of μ between 1 and 100 (orange curve) and computed the resulting rank and datafit. Note that even if a change of μ is not large enough to change the rank of the solution it does affect the non-zero singular values. To achieve the best result for a specific rank with (45) we should select the smallest μ that gives the correct rank. Even though (3) does not hold, the relaxation (44) consistently gives better data fit with lower rank than (45). Figure 5 also shows the rank versus the distance to the ground truth solution. For high rank the distance is typically larger

for (44) than (45). A feasible explanation is that when the rank is high it is more likely that the row space of $X^{\#}$ contains a matrix of the type $N(C)$. Loosely speaking, when we allow too complex deformations it becomes more difficult to uniquely recover the shape. The nuclear norm's built in bias to small solutions helps to regularize the problem when the rank constraint is not discriminative enough.

One way to handle the null space of $\mathcal{A}$ is to add additional regularizes that penalize low rank matrices of the type $N(C)$. Dai *et al.* [9] suggested to use the derivative prior $\|DX^{\#}\|_F^2$, where the matrix $D : \mathbb{R}^F \rightarrow \mathbb{R}^{F-1}$ is a first order difference operator. The nullspace of D consists of matrices that are constant in each column. Since this implies that the scene is rigid it is clear that $N(C)$ is not in the nullspace of D. We add this term and compare

$$\mathcal{R}_r(X^{\#}) + \|RX - M\|_F^2 + \|DX^{\#}\|_F^2 \qquad (46)$$

and

$$\mu\|X^{\#}\|_* + \|RX - M\|_F^2 + \|DX^{\#}\|_F^2. \qquad (47)$$

Figures 6 and 7 show the results. In this case both the data fit and the distance to the ground truth is consistently better with (46) than (47). When the rank increases most of the regularization comes from the derivative prior leading to both methods providing similar results.

4. Conclusions

In this paper we studied the local minima of a non-convex rank regularization approach. Our main theoretical result shows that if a RIP property holds then there is often a unique stationary point. Since the proposed relaxation (4) and the original objective (2) is shown to have the same global minimizers if $\|\mathcal{A}\| \leq 1$ in [8] this result is also relevant for the original discontinuous problem. Our experimental evaluation shows that the proposed approach often gives better solutions than standard convex alternatives, even when the RIP constraint does not hold.

References

[1] F. Andersson, M. Carlsson, and C. Olsson. Convex envelopes for fixed rank approximation. *Optimization Letters*, pages 1–13, 2017. 1, 2

[2] R. Basri, D. Jacobs, and I. Kemelmacher. Photometric stereo with general, unknown lighting. *Int. J. Comput. Vision*, 72(3):239–257, May 2007. 1

[3] D. Bertsekas. *Nonlinear Programming*. Athena Scientific, 1999. 3

[4] S. Bhojanapalli, B. Neyshabur, and N. Srebro. Global optimality of local search for low rank matrix recovery. In *Annual Conference in Neural Information Processing Systems (NIPS)*. 2016. 3

[5] C. Bregler, A. Hertzmann, and H. Biermann. Recovering non-rigid 3d shape from image streams. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2000. 1, 8

[6] E. J. Candès, X. Li, Y. Ma, and J. Wright. Robust principal component analysis? *J. ACM*, 58(3):11:1–11:37, 2011. 1

[7] E. J. Candès and B. Recht. Exact matrix completion via convex optimization. *Foundations of Computational Mathematics*, 9(6):717–772, 2009. 1

[8] M. Carlsson. On convexification/optimization of functionals including an l2-misfit term. *arXiv preprint arXiv:1609.09378*, 2016. 3, 8

[9] Y. Dai, H. Li, and M. He. A simple prior-free method for non-rigid structure-from-motion factorization. *International Journal of Computer Vision*, 107(2):101–122, 2014. 6, 8

[10] C. Eckart and G. Young. The approximation of one matrix by another of lower rank. *Psychometrika*, 1(3):211–218, 1936. 1

[11] A. Eriksson, T. Thanh Pham, T.-J. Chin, and I. Reid. The k-support norm and convex envelopes of cardinality and rank. In *The IEEE Conference on Computer Vision and Pattern Recognition*, 2015. 2

[12] M. Fazel, H. Hindi, and S. P. Boyd. A rank minimization heuristic with application to minimum order system approximation. In *American Control Conference*, 2001. 1

[13] R. Garg, A. Roussos, and L. Agapito. A variational approach to video registration with subspace constraints. *Int. J. Comput. Vision*, 104(3):286–314, 2013. 1

[14] R. Garg, A. Roussos, and L. de Agapito. Dense variational reconstruction of non-rigid surfaces from monocular video. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2013. 1

[15] R. Ge, C. Jin, and Y. Zheng. No spurious local minima in nonconvex low rank problems: A unified geometric analysis. *arXiv preprint*, arxiv:1704.00708, 2017. 3

[16] P. Gong, C. Zhang, Z. Lu, J. Huang, and J. Ye. A general iterative shrinkage and thresholding algorithm for non-convex regularized optimization problems. In *International Conference on Machine Learning (ICML)*, pages 37–45, 2013. 5

[17] C. Grussler and A. Rantzer. On optimal low-rank approximation of non-negative matrices. In *2015 54th IEEE Conference on Decision and Control (CDC)*, pages 5278–5283, Dec 2015. 2

[18] C. Grussler, A. Rantzer, and P. Giselsson. Low-rank optimization with convex constraints. *CoRR*, abs/1606.01793, 2016. 2

[19] Y. Hu, D. Zhang, J. Ye, X. Li, and X. He. Fast and accurate matrix completion via truncated nuclear norm regularization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(9):2117–2130, 2013. 1

[20] V. Larsson and C. Olsson. Convex envelopes for low rank approximation. In *International Conference on Energy Minimization Methods in Computer Vision and Pattern Recognition*, 2015. 2

[21] V. Larsson and C. Olsson. Convex low rank approximation. *International Journal of Computer Vision*, 120(2):194–214, 2016. 1, 2, 3, 5

[22] A. M. McDonald, M. Pontil, and D. Stamos. Spectral k-support norm regularization. In *Advances in Neural Information Processing Systems*. 2014. 2

[23] K. Mohan and M. Fazel. Iterative reweighted least squares for matrix rank minimization. In *Annual Allerton Conference on Communication, Control, and Computing*, pages 653–661, 2010. 1

[24] T. H. Oh, Y. W. Tai, J. C. Bazin, H. Kim, and I. S. Kweon. Partial sum minimization of singular values in robust pca: Algorithm and applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(4):744–758, 2016. 1

[25] C. Olsson, M. Carlsson, F. Andersson, and V. Larsson. Non-convex rank/sparsity regularization and local minima. *Proceedings of the International Conference on Computer Vision*, 2017. 2, 3, 4, 5

[26] S. Oymak, A. Jalali, M. Fazel, Y. C. Eldar, and B. Hassibi. Simultaneously structured models with application to sparse and low-rank matrices. *IEEE Transactions on Information Theory*, 61(5):2886–2908, 2015. 1

[27] S. Oymak, K. Mohan, M. Fazel, and B. Hassibi. A simplified approach to recovery conditions for low rank matrices. In *IEEE International Symposium on Information Theory Proceedings (ISIT)*, pages 2318–2322, 2011. 1

[28] D. Park, A. Kyrillidis, C. Caramanis, and S. Sanghavi. Non-square matrix sensing without spurious local minima via the burer-monteiro approach. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, AISTATS*, pages 65–74, 2017. 3

[29] B. Recht, M. Fazel, and P. A. Parrilo. Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM Rev.*, 52(3):471–501, Aug. 2010. 1, 6

[30] C. Tomasi and T. Kanade. Shape and motion from image streams under orthography: A factorization method. *Int. J. Comput. Vision*, 9(2):137–154, 1992. 1

[31] J. Yan and M. Pollefeys. A factorization-based approach for articulated nonrigid shape, motion and kinematic chain recovery from video. *IEEE Trans. Pattern Anal. Mach. Intell.*, 30(5):865–877, 2008. 1