

Article

Stable Analysis of Compressive Principal Component Pursuit

Qingshan You ^{1,2,*} and Qun Wan ²

¹ School of Computer Science, Civil Aviation Flight University of China, Guanghan 618307, China

² School of Electronic Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China; wanqun@uestc.edu.cn; Tel.: +86-028-6183-0358

* Correspondence: youlin_2001@163.com

Academic Editor: Bruno Carpentieri

Received: 5 January 2017; Accepted: 17 February 2017; Published: 21 February 2017

Abstract: Compressive principal component pursuit (CPCP) recovers a target matrix that is a superposition of low-complexity structures from a small set of linear measurements. Previous works mainly focus on the analysis of the existence and uniqueness. In this paper, we address its stability. We prove that the solution to the related convex programming of CPCP gives an estimate that is stable to small entry-wise noise. We also provide numerical simulation results to support our result. Numerical results show that the solution to the related convex program is stable to small entry-wise noise under board condition.

Keywords: matrix completion; low-complexity structure; stability analytic; compressive principal component pursuit

1. Introduction

Recently, there has been a rapidly increasing interest in recovering a target matrix that is a superposition of low-rank and sparse components from a small set of linear measurements. In many cases, this problem is shorted for matrix completion [1–3], which arises in a number of fields, such as medical imaging [4,5], seismology [6], and computer vision [7,8] and Kalman filter [9]. Mathematically, there exists a large-scale data matrix $M = L_0 + S_0$, where L_0 is a low-rank matrix, and S_0 is a sparse matrix. One of the important problems here is how to extract the intrinsic low-dimensional structure from a small set of linear measurements. In a recent paper [10], E. J. Candès et al. proved that most low-rank matrices and the sparse components can be recovered, provided that the rank of the low-rank component is not too large, and that the sparse component is reasonably sparse. It is more important that they proved that these two components can be recovered by solving a simple convex optimization problem. In [11], John Wright et al. generalized this problem to decompose a matrix into multiple incoherent components:

$$\begin{aligned} & \text{minimize} && \sum_i^{\tau} \lambda_i \|X_i\|_{(i)} \\ & \text{subject to} && \sum_i^{\tau} X_i = M, \end{aligned} \tag{1}$$

where $\|X_i\|_{(i)}$ are norms that encourage various types of low-complexity structure. The authors also provide a sufficient condition that can promise the existence and uniqueness theorem of compressive principle component pursuit (CPCP). The result in [11] requires that the components are low-complexity structures.

However, in many applications, the observed measurements are always corrupted by different kinds of noise which may affect every entry of the data matrix. In order to further complete the theory

developed in [11], it is necessary to research the stability of CPCP which can guarantee stable and accurate recovery in the presence of entry-wise noise. In this paper, we make a commendable attempt in this respect. We denote M as the observing matrix which can decompose into multiple incoherent components, and assume that

$$M = \sum_i^{\tau} X_{i,0} + Z_0,$$

where $X_{i,0}$ are corresponding incoherent components and Z_0 is an independent and identically distributed (i.i.d.) noise. We assume that Z_0 is only limited by $\|Z_0\|_F \leq \delta$ for some $\delta > 0$. In order to recover the unknown low-complexity structures, we suggest solving the following relaxed optimization problem.

$$\begin{aligned} & \text{minimize} && \sum_i^{\tau} \lambda_i \|X_i\|_{(i)} \\ & \text{subject to} && \left\| \sum_i^{\tau} X_i - M \right\|_F \leq \delta \end{aligned} \quad (2)$$

In this paper, we prove the solution of (2) is stable to small entry-wise noise. The rest of paper is organized as follows. In Section 2, we show some notations and the main result, which will be proven in Sections 3 and 4. In Section 3, we give two important lemmas which are an important parts of our main result. In Section 4, The proof of Theorem 1 will be given. We further provide numerical results in Section 5 and conclude the paper in Section 6.

2. Notations and Main Results

In this section, we first give some important notions which will be used throughout this paper, and then provide the main result.

2.1. Notations

We denote the operator norm of matrix by $\|X\|$, the Frobenius norm by $\|X\|_F$, and the nuclear norm by $\|X\|_*$, and denote the dual norm of $\|X\|_{(i)}$ by $\|X\|_{(i)}^*$. The Euclidean inner product between two matrices is defined by the formula $\langle X, Y \rangle = \text{trace}(X^*Y)$. Note that $\|X\|_F^2 = \langle X, X \rangle$. The Cauchy–Schwarz inequality gives $\langle X, Y \rangle \leq \|X\|_F \|Y\|_F$, and it is well known that we also have $\langle X, Y \rangle \leq \|X\|_{(i)} \|Y\|_{(i)}^*$ (e.g., [1,12]). $\|X\|_{(i)}$ majorized the Frobenius norm means $\|X\|_{(i)} \geq \|X\|_F$ for all X . Linear transformations which act on the space of matrices are denoted by $\mathcal{P}_T X$. It is easy to see that the operator of \mathcal{P}_T is a high dimension matrix. The operator norm of the operator is denoted by $\|\mathcal{P}_T\|$. It should be noted that $\|\mathcal{P}_T\| = \sup_{\|X\|_F=1} \|\mathcal{P}_T X\|_F$.

For any matrix vector $x = [X_i], i = 1, 2, \dots, \tau$, where $X_i \in \mathbb{R}^{m \times n}$ is i -th matrix. We will consider two norms of this matrix pair, which can define as $\|x\|_{\diamond} := \sum_i^{\tau} \lambda_i \|X_i\|_{(i)}$ and $\|x\|_2 := \sum_i^{\tau} \|X_i\|_F$. In order to simplify the stability analysis of CPCP, we also define the subspaces (the common component) $\gamma := [\Gamma_i]$, $\Gamma_i = (\sum_l^{\tau} X_l)/\tau$ $i = 1, 2, \dots, \tau$, and (the different component) $\gamma^{\perp} := [\Gamma_i^{\perp}]$, $\Gamma_i^{\perp} = X_i - \Gamma_i$ $i = 1, 2, \dots, \tau$. In order to analyze the behavior of special projection operator, we define the projection operator $\mathcal{P}_{T_1} \times \dots \times \mathcal{P}_{T_{\tau}}(x) := [\mathcal{P}_{T_1}(X_1), \dots, \mathcal{P}_{T_{\tau}}(X_{\tau})]$.

we assume that $\|X_i\|_{(i)}$ $i = 1, 2, \dots, \tau$ are decomposable norms. The definition of decomposable norms is below.

Definition 1 (Decomposable Norms). *if there exists a subspace T and a matrix Z satisfying*

$$\partial \|\cdot\| (X) = \{\Lambda | \mathcal{P}_T \Lambda = Z, \|\mathcal{P}_{T^{\perp}} \Lambda\|^* \leq 1\}, \quad (3)$$

where $\|\cdot\|^*$ denotes the dual norm of $\|\cdot\|$ and $\mathcal{P}_{T^{\perp}}$ is nonexpansive with respect to $\|\cdot\|^*$. Then, we say that the norm $\|\cdot\|$ is decomposable at X .

Definition 2 (Inexact Certificate). We say Λ is an (α, β) -inexact certificate for a putative solution $(X_{1,\diamond}, \dots, X_{\tau,\diamond})$ to (1.1) with parameters $(\lambda_1, \dots, \lambda_\tau)$ if for each i , $\|\mathcal{P}_{T_i}\Lambda - \lambda_i Z_i\|_F \leq \alpha$, and $\|\mathcal{P}_{T_i^\perp}\Lambda\|_{(i)}^* < \lambda_i \beta$.

2.2. Main Results

Pertaining to Problem (1), we have the result as follows.

Lemma 1 ([11]). Assume there exists a feasible solution $x = (X_1, \dots, X_\tau)$ to the optimization Problem (1). Suppose that each of the norms $\|\cdot\|_{(i)}$ is decomposable at X_i , and that each of the $\|\cdot\|_{(i)}$ majorizes the Frobenius norm. Then, x is the unique optimal solution if T_1, \dots, T_τ are independent subspaces with

$$\|\mathcal{P}_{T_i}\mathcal{P}_{T_j}\| < \frac{1}{\tau-1} \quad \forall i \neq j,$$

and there exists an (α, β) -inexact certificate $\hat{\Lambda}$, with

$$\beta + \frac{\alpha\sqrt{\tau}}{\sqrt{1 - (\tau-1)\max_{ij}\|\mathcal{P}_{T_i}\mathcal{P}_{T_j}\|}} \times \frac{1}{\min_l \lambda_l} \leq 1.$$

The main contribution of this paper is the stability analysis of the solution of CPCP; the main Theorem of [13] can be regarded as a special case of our result (although the main idea of proof is similar to the paper [13], there are some important differences here). Next, we will provide the proposed related convex programming (2) is stable from small entry-wise noise under board condition. The main result of this paper is provided below.

Theorem 1. Assume $x_\diamond = (X_{1,\diamond}, \dots, X_{\tau,\diamond})$, $\hat{x} = (X_1, \dots, X_\tau)$ are the solutions of the optimization Problems (1) and (2), respectively. Suppose that each of the norms $\|\cdot\|_{(i)}$ is decomposable at $X_{i,\diamond}$, and each of the $\|\cdot\|_{(i)}$ majorizes the Frobenius norm. Then, if T_1, \dots, T_τ are independent subspaces with

$$\|\mathcal{P}_{T_i}\mathcal{P}_{T_j}\| < \frac{1}{\tau-1} \quad \forall i \neq j$$

and there exists an (α, β) -inexact certificate $\hat{\Lambda}$, with

$$\beta + \frac{\alpha\sqrt{\tau}}{\sqrt{1 - (\tau-1)\max_{ij}\|\mathcal{P}_{T_i}\mathcal{P}_{T_j}\|}} \times \frac{1}{\min_l \lambda_l} \leq 1, \quad (4)$$

then for any Z_0 which is limited by $\|Z_0\|_F \leq \delta$, the solution \hat{x} to the convex programming (2) obeys

$$\sum_i \|\hat{x}_i - x_{i,\diamond}\|_2^2 \leq C(n, \tau, \alpha, \beta)\delta^2, \quad (5)$$

where $C(n, \tau, \alpha, \beta)$ is a numerical constant only depending upon n, τ, α, β .

3. Main Lemmas

In this section, we present two main lemmas which are used to obtain Theorem 1. The paper [11] states that:

Lemma 2 ([11]). Suppose T_1, \dots, T_τ are independent subspaces of $\mathbb{R}^{m \times n}$ and $Z_1 \in T_1, \dots, Z_\tau \in T_\tau$, under the other conditions of Lemma 1. Then, the below equations

$$\mathcal{P}_{T_i}\Delta = \lambda_i Z_i - \mathcal{P}_{T_i}\Lambda, \quad i = 1, \dots, \tau$$

have a solution $\Delta \in T_1 + \dots + T_\tau$ obeying

$$\|\Delta\|_F \leq \sqrt{\frac{\alpha^2 \tau}{1 - (\tau - 1) \max_{i \neq j} \|\mathcal{P}_{T_i} \mathcal{P}_{T_j}\|}}.$$

In order to bound the behavior of the norm of $\|x\|_\diamond$, we have the first main lemma that is used to obtain Theorem 1.

Lemma 3. Assume $\|\mathcal{P}_{T_i} \mathcal{P}_{T_j}\| < \frac{1}{\tau-1} \forall i \neq j$. Suppose there exists an (α, β) -inexact certificate $\hat{\Lambda}$ satisfying Lemma 1. Then, for any perturbation $h = [H_i]$ obeying $\sum_i H_i = 0$

$$\|x_0 + h\|_\diamond \geq \|x_0\|_\diamond + \sum_{i=1}^{\tau} (\lambda_i - C_\alpha - \lambda_i \beta) \|\mathcal{P}_{T_i^\perp} H_i\|_{(i)},$$

wherein, let $C_\alpha = \sqrt{\frac{\alpha^2 \tau}{1 - (\tau - 1) \max_{i \neq j} \|\mathcal{P}_{T_i} \mathcal{P}_{T_j}\|}}$. It is easy to see that under the hypothesis of Lemma 1, the coefficients of $\|\mathcal{P}_{T_i^\perp} H_i\|_{(i)}$ satisfy $\lambda_i - C_\alpha - \lambda_i \beta > 0$.

Proof. According to the property of convex function, for any subgradients $z = [Z_i] \in \partial \|x_0\|_\diamond$, we can obtain

$$\|x_0 + h\|_\diamond \geq \|x\|_\diamond + \sum_{i=1}^{\tau} \lambda_i \langle Z_i, H_i \rangle.$$

Now, because the norm of the subgradients is decomposable at X_i , there exists Λ, Z_i, α , and β obeying $\|\mathcal{P}_{T_i} \Lambda - \lambda_i Z_i\|_F \leq \alpha$, and $\|\mathcal{P}_{T_i^\perp} \Lambda\|_{(i)}^* < \lambda_i \beta$. Let $\Delta_i := \mathcal{P}_{T_i} \Delta = \lambda_i Z_i - \mathcal{P}_{T_i} \Lambda \in T_i$ (see Lemma 2). Note that

$$\begin{aligned} \Lambda + \Delta_i + \mathcal{P}_{T_i^\perp}(\lambda_i Z_i - \Lambda) &= \Lambda + \lambda_i Z_i - \mathcal{P}_{T_i} \Lambda + \mathcal{P}_{T_i^\perp} \lambda_i Z_i - \mathcal{P}_{T_i^\perp} \Lambda \\ &= \lambda_i Z_i + \mathcal{P}_{T_i^\perp} \lambda_i Z_i \\ &= \lambda_i Z_i, \end{aligned}$$

where the second equation obeys $Z_i \in T_i$. According to the above equation, we will continue bounding $\sum_{i=1}^{\tau} \lambda_i \langle Z_i, H_i \rangle$.

$$\begin{aligned} \sum_{i=1}^{\tau} \lambda_i \langle Z_i, H_i \rangle &= \sum_{i=1}^{\tau} \langle \Lambda + \Delta_i + \mathcal{P}_{T_i^\perp}(\lambda_i Z_i - \Lambda), H_i \rangle \\ &= \sum_{i=1}^{\tau} \langle \Lambda, H_i \rangle + \sum_{i=1}^{\tau} \langle \mathcal{P}_{T_i} \Delta, H_i \rangle + \sum_{i=1}^{\tau} \langle \mathcal{P}_{T_i^\perp}(\lambda_i Z_i - \Lambda), H_i \rangle \\ &= \langle \Lambda, \sum_{i=1}^{\tau} H_i \rangle + \sum_{i=1}^{\tau} \langle \Delta, \mathcal{P}_{T_i} H_i \rangle + \sum_{i=1}^{\tau} \langle \lambda_i Z_i - \Lambda, \mathcal{P}_{T_i^\perp} H_i \rangle \\ &= \sum_{i=1}^{\tau} \langle \Delta, H_i \rangle - \sum_{i=1}^{\tau} \langle \Delta, \mathcal{P}_{T_i^\perp} H_i \rangle + \sum_{i=1}^{\tau} \langle \lambda_i Z_i - \Lambda, \mathcal{P}_{T_i^\perp} H_i \rangle \\ &\geq \sum_{i=1}^{\tau} \langle \lambda_i Z_i - \Lambda, \mathcal{P}_{T_i^\perp} H_i \rangle - \sum_{i=1}^{\tau} \|\Delta\|_F \|\mathcal{P}_{T_i^\perp} H_i\|_F \\ &\geq \sum_{i=1}^{\tau} \langle \lambda_i Z_i - \Lambda, \mathcal{P}_{T_i^\perp} H_i \rangle - \sum_{i=1}^{\tau} \|\Delta\|_F \|\mathcal{P}_{T_i^\perp} H_i\|_{(i)} \end{aligned}$$

With the definition of duality, there exists $\hat{Z}_i \in \partial \|X_{i,0}\|_{(i)}$ with $\|\hat{Z}_i\|_{(i)}^* \leq 1$ such that $\langle Z_i^*, \mathcal{P}_{T_i^\perp} H_i \rangle = \|\mathcal{P}_{T_i^\perp} H_i\|_{(i)}$. Moreover, with the Cauchy-Schwarz inequality, we have

$$|\langle \Lambda, \mathcal{P}_{T_i^\perp} H_i \rangle| = |\langle \mathcal{P}_{T_i^\perp} \Lambda, \mathcal{P}_{T_i^\perp} H_i \rangle| \leq \|\mathcal{P}_{T_i^\perp} \Lambda\|_{(i)}^* \|\mathcal{P}_{T_i^\perp} H_i\|_{(i)}.$$

Let $Z_i = \hat{Z}_i$. Then, we can obtain:

$$\langle \lambda_i Z_i - \Lambda, \mathcal{P}_{T_i^\perp} H_i \rangle \geq (\lambda_i - \|\mathcal{P}_{T_i^\perp} \Lambda\|_{(i)}^*) \|\mathcal{P}_{T_i^\perp} H_i\|_{(i)}.$$

Combining with the inequalities above, we can obtain

$$\begin{aligned} \|x_0 + h\|_\diamond &\geq \|x_0\|_\diamond + \sum_{i=1}^{\tau} (\lambda_i - \|\Delta\|_F - \|\mathcal{P}_{T_i^\perp} \Lambda\|_{(i)}^*) \|\mathcal{P}_{T_i^\perp} H_i\|_{(i)} \\ &\geq \|x_0\|_\diamond + \sum_{i=1}^{\tau} (\lambda_i - C_\alpha - \lambda_i \beta) \|\mathcal{P}_{T_i^\perp} H_i\|_{(i)}. \end{aligned}$$

The Lemma 3 is established. \square

For bounding the behavior of $\sum_i \|\hat{x}_i - x_{i,\diamond}\|_F^2$, we have to bound the projection operator $\mathcal{P}_{T_1} \times \cdots \times \mathcal{P}_{T_\tau}(x)$. Therefore, we have the second main lemma that will be used to obtain Theorem 1.

Lemma 4. Assume that $\|\mathcal{P}_{T_i} \mathcal{P}_{T_j}\| < \frac{1}{\tau-1} \quad \forall i \neq j$. For any matrix vector $x = [X_i]$, we have

$$\|\mathcal{P}_\gamma(\mathcal{P}_{T_1} \times \cdots \times \mathcal{P}_{T_\tau})(x)\|_F^2 \geq \frac{1 - \max_i \frac{1}{2} \sum_{j \neq i} (\|\mathcal{P}_{T_i} \mathcal{P}_{T_j}\| + \|\mathcal{P}_{T_j} \mathcal{P}_{T_i}\|)}{\tau} \|\mathcal{P}_{T_1} \times \cdots \times \mathcal{P}_{T_\tau}(x)\|_F^2.$$

It is easy to see that under the hypothesis of $\|\mathcal{P}_{T_i} \mathcal{P}_{T_j}\| < \frac{1}{\tau-1} \quad \forall i \neq j$, the constant $\frac{1 - \max_i \frac{1}{2} \sum_{j \neq i} (\|\mathcal{P}_{T_i} \mathcal{P}_{T_j}\| + \|\mathcal{P}_{T_j} \mathcal{P}_{T_i}\|)}{\tau^2}$ is strictly greater than zero.

Proof. With respect to any matrix $x = [X_i]$, we have $\mathcal{P}_\gamma(x) = [\Gamma_i]$, where $\Gamma_i = (\sum_{l=1}^{\tau} X_l) / \tau$. It is easy to see that $\|\mathcal{P}_\gamma(x)\|_F^2 = \frac{1}{\tau} \|\sum_{l=1}^{\tau} X_l\|_F^2$. Then, we have

$$\begin{aligned} \|\mathcal{P}_\gamma(\mathcal{P}_{T_1} \times \cdots \times \mathcal{P}_{T_\tau})(x)\|_F^2 &= \frac{1}{\tau} \left\| \sum_{i=1}^{\tau} \mathcal{P}_{T_i} X_i \right\|_F^2 \\ &= \frac{1}{\tau} \left(\sum_{i=1}^{\tau} \|\mathcal{P}_{T_i} X_i\|_F^2 + \sum_{j \neq i} \langle \mathcal{P}_{T_i} X_i, \mathcal{P}_{T_j} X_j \rangle \right). \end{aligned}$$

Note that

$$\begin{aligned} \langle \mathcal{P}_{T_i} X_i, \mathcal{P}_{T_j} X_j \rangle &= \langle \mathcal{P}_{T_i} X_i, \mathcal{P}_{T_i} \mathcal{P}_{T_j} X_j \rangle \\ &\geq -\|\mathcal{P}_{T_i} \mathcal{P}_{T_j}\| \|\mathcal{P}_{T_i} X_i\|_F \|\mathcal{P}_{T_j} X_j\|_F. \end{aligned}$$

Together with $\|\mathcal{P}_{T_i} \mathcal{P}_{T_j}\| < \frac{1}{\tau-1} \quad \forall i \neq j$, we have

$$\begin{aligned} \|\mathcal{P}_\gamma(\mathcal{P}_{T_1} \times \cdots \times \mathcal{P}_{T_\tau})(x)\|_F^2 &\geq \frac{1}{\tau} \left(\sum_{i=1}^{\tau} \|\mathcal{P}_{T_i} X_i\|_F^2 - \sum_{j \neq i} \|\mathcal{P}_{T_i} \mathcal{P}_{T_j}\| \|\mathcal{P}_{T_i} X_i\|_F \|\mathcal{P}_{T_j} X_j\|_F \right) \\ &\geq \frac{1}{\tau} \left(\sum_{i=1}^{\tau} \|\mathcal{P}_{T_i} X_i\|_F^2 - \sum_{j \neq i} \frac{\|\mathcal{P}_{T_i} \mathcal{P}_{T_j}\|}{2} (\|\mathcal{P}_{T_i} X_i\|_F^2 + \|\mathcal{P}_{T_j} X_j\|_F^2) \right) \\ &= \frac{1}{\tau} \sum_{i=1}^{\tau} \left(1 - \frac{1}{2} \sum_{j \neq i} (\|\mathcal{P}_{T_i} \mathcal{P}_{T_j}\| + \|\mathcal{P}_{T_j} \mathcal{P}_{T_i}\|) \right) \|\mathcal{P}_{T_i} X_i\|_F^2 \\ &\geq \frac{1 - \max_i \frac{1}{2} \sum_{j \neq i} (\|\mathcal{P}_{T_i} \mathcal{P}_{T_j}\| + \|\mathcal{P}_{T_j} \mathcal{P}_{T_i}\|)}{\tau} \left(\sum_i \|\mathcal{P}_{T_i} X_i\|_F^2 \right) \\ &\geq \frac{1 - \max_i \frac{1}{2} \sum_{j \neq i} (\|\mathcal{P}_{T_i} \mathcal{P}_{T_j}\| + \|\mathcal{P}_{T_j} \mathcal{P}_{T_i}\|)}{\tau} \|\mathcal{P}_{T_1} \times \cdots \times \mathcal{P}_{T_\tau}(x)\|_F^2, \end{aligned}$$

where in the second inequality, we have used the inequity that for any x, y , $2xy \leq (x^2 + y^2)$. Therefore, Lemma 4 is established. \square

4. Proof of Theorem 1

In this section, we will provide the proof of Theorem 1. Our main proof is based on two elementary and important properties of \hat{x} , which is the solution of Problem (2). First, note that x_0 is also a feasible solution to Problem (2) and \hat{x} is the optimum solution; therefore, we can obtain $\|\hat{x}\|_\diamond \leq \|x_0\|_\diamond$. Second, according to triangle inequality, we can obtain

$$\begin{aligned} \|\hat{x} - x_0\|_2 &= \|\hat{x} - M - (x_0 - M)\|_2 \\ &\leq \|\hat{x} - M\|_2 + \|x_0 - M\|_2 \\ &\leq 2\delta. \end{aligned} \quad (6)$$

Let $\hat{x} = x_0 + h$, where $h = [H_i]$. According to the definition of subspace of γ , we denote $h^\gamma := \mathcal{P}_\gamma(h)$, $h^{\gamma^\perp} := \mathcal{P}_{\gamma^\perp}(h)$ for short. Our main aim is to bound $\|h\|_2 = \|\hat{x} - x_0\|_2$, which can be rewritten as

$$\begin{aligned} \|h\|_2^2 &= \|h^\gamma\|_2^2 + \|h^{\gamma^\perp}\|_2^2 \\ &= \|h^\gamma\|_2^2 + \|\mathcal{P}_{T_1} \times \cdots \times \mathcal{P}_{T_\tau} h^{\gamma^\perp}\|_2^2 \\ &\quad + \|\mathcal{P}_{T_1^\perp} \times \cdots \times \mathcal{P}_{T_\tau^\perp} h^{\gamma^\perp}\|_2^2. \end{aligned} \quad (7)$$

Combining with (4), we have

$$\|h^\gamma\|_2^2 = \sum_i \left\| \frac{\sum_{j=1}^\tau H_j}{\tau} \right\|_2^2 \leq \frac{4\delta^2}{\tau}$$

Therefore, it is necessary to bound the other two terms on the right-hand-side of (5). We will bound the second and third terms, respectively.

Norm equivalence theorem tells us that *every two norms on a finite dimensional normed space are equivalent*, which implies that there exists two constants $C(n, \tau) \geq c(n, \tau) > 0$ satisfying

$$c(n, \tau)\|x\|_2 \leq \|x\|_\diamond \leq C(n, \tau)\|x\|_2. \quad (8)$$

A. Estimate the third term of (5) Let Λ be a dual certificate obeying Lemma 1. Then, using triangle inequality, we have

$$\|x_0 + h\|_2 \geq \|x_0 + h^{\gamma^\perp}\|_2 - \|h^\gamma\|_2. \quad (9)$$

Combining with Lemma 3, we can obtain

$$\begin{aligned} \|x_0 + h^{\gamma^\perp}\|_2 &\geq \|x_0\|_d + \sum_{i=1}^\tau (\lambda_i - C_\alpha - \lambda_i \beta) \|\mathcal{P}_{T_i^\perp} H_i\|_{(i)} \\ &\geq \|x_0\|_d + (1 - C_\alpha \frac{1}{\min_i \lambda_i} - \beta) \sum_{i=1}^\tau \lambda_i \|\mathcal{P}_{T_i^\perp} H_i^{\Gamma^\perp}\|_{(i)} \\ &\geq \|x_0 + h\|_d + (1 - C_\alpha \frac{1}{\min_i \lambda_i} - \beta) \sum_{i=1}^\tau \lambda_i \|\mathcal{P}_{T_i^\perp} H_i^{\Gamma^\perp}\|_{(i)}, \end{aligned}$$

wherein, to get the third inequality, we used the fact $\|\hat{x}\|_d \leq \|x_0\|_d$. For simplification, let

$$C_1(\alpha, \beta) \triangleq \left(1 - C_\alpha \frac{1}{\min_i \lambda_i} - \beta\right) > 0.$$

Therefore, we have

$$\|x_0 + h^{\gamma^\perp}\|_2 \geq \|x_0 + h\|_d + C_1(\alpha, \beta) \sum_{i=1}^{\tau} \lambda_i \|\mathcal{P}_{T_i^\perp} H_i^{\Gamma^\perp}\|_{(i)}.$$

Combining with (7), we can obtain

$$C_1(\alpha, \beta) \sum_{i=1}^{\tau} \lambda_i \|\mathcal{P}_{T_i^\perp} H_i^{\Gamma^\perp}\|_{(i)} \leq \|h^\gamma\|_2.$$

Then

$$\sum_{i=1}^{\tau} \lambda_i \|\mathcal{P}_{T_i^\perp} H_i^{\Gamma^\perp}\|_{(i)} \leq C_2(\alpha, \beta) \|h^\gamma\|_2, \quad (10)$$

where $C_2(\alpha, \beta) = 1/C_1(\alpha, \beta)$. We will estimate the third term of (5). Using triangle inequality, we have

$$\begin{aligned} \|\mathcal{P}_{T_1^\perp} \times \cdots \times \mathcal{P}_{T_\tau^\perp} h^{\Gamma^\perp}\|_2 &\leq \sum_i \|\mathcal{P}_{T_i^\perp} H_i^{\Gamma^\perp}\|_F \\ &\leq \frac{1}{c(n, \tau)} \sum_{i=1}^{\tau} \lambda_i \|\mathcal{P}_{T_i^\perp} H_i^{\Gamma^\perp}\|_{(i)} \\ &\leq \frac{C_2(\alpha, \beta)}{c(n, \tau)} \|h^\gamma\|_2 \\ &\leq C(n, \tau, \alpha, \beta) \delta, \end{aligned}$$

where $C(n, \tau, \alpha, \beta) := \frac{2C_2(\alpha, \beta)}{c(n, \tau)\sqrt{\tau}}$. The second inequality is set up by (6); the fourth inequality is obtained by (8); the last one is obtained by the fact $\|h^\gamma\|_2 \leq \frac{2\delta}{\sqrt{\tau}}$. Therefore, we can obtain

$$\|\mathcal{P}_{T_1^\perp} \times \cdots \times \mathcal{P}_{T_\tau^\perp} (h^{\Gamma^\perp})\|_2^2 \leq C^2(n, \tau, \alpha, \beta) \delta^2, \quad (11)$$

which implies that the third term of (5) can bound by $C\delta$.

B. Estimate the second term of (5) According to Lemma 4, we can obtain

$$\begin{aligned} \|\mathcal{P}_\gamma(\mathcal{P}_{T_1} \times \cdots \times \mathcal{P}_{T_\tau})(h^{\gamma^\perp})\|_2^2 &\geq \frac{1 - \max_i \frac{1}{2} \sum_{j \neq i} (\|\mathcal{P}_{T_i} \mathcal{P}_{T_j}\| + \|\mathcal{P}_{T_j} \mathcal{P}_{T_i}\|)}{\tau} \\ &\quad \|\mathcal{P}_{T_1} \times \cdots \times \mathcal{P}_{T_\tau} (h^{\gamma^\perp})\|_2^2 \\ &= \hat{C}(\tau, \alpha, \beta) \|\mathcal{P}_{T_1} \times \cdots \times \mathcal{P}_{T_\tau} (h^{\gamma^\perp})\|_2^2, \end{aligned}$$

where $\hat{C}(\tau, \alpha, \beta) := \frac{1 - \max_i \frac{1}{2} \sum_{j \neq i} (\|\mathcal{P}_{T_i} \mathcal{P}_{T_j}\| + \|\mathcal{P}_{T_j} \mathcal{P}_{T_i}\|)}{\tau}$. Note that

$$\begin{aligned} 0 &= \mathcal{P}_\gamma(h^{\gamma^\perp}) \\ &= \mathcal{P}_\gamma \mathcal{P}_{T_1} \times \cdots \times \mathcal{P}_{T_\tau} h^{\gamma^\perp} + \mathcal{P}_\gamma \mathcal{P}_{T_1^\perp} \times \cdots \times \mathcal{P}_{T_\tau^\perp} h^{\gamma^\perp}. \end{aligned}$$

Therefore,

$$\begin{aligned}\|\mathcal{P}_\gamma \mathcal{P}_{T_1} \times \cdots \times \mathcal{P}_{T_\tau} h^{\gamma^\perp}\|_2 &= \|\mathcal{P}_\gamma \mathcal{P}_{T_1^\perp} \times \cdots \times \mathcal{P}_{T_\tau^\perp} h^{\gamma^\perp}\|_2 \\ &\leq \|\mathcal{P}_{T_1^\perp} \times \cdots \times \mathcal{P}_{T_\tau^\perp} h^{\gamma^\perp}\|_2.\end{aligned}$$

Taking the previous two inequalities, we have

$$\begin{aligned}\|\mathcal{P}_{T_1} \times \cdots \times \mathcal{P}_{T_\tau} h^{\gamma^\perp}\|_2 &\leq \frac{\|\mathcal{P}_\gamma \mathcal{P}_{T_1^\perp} \times \cdots \times \mathcal{P}_{T_\tau^\perp} h^{\gamma^\perp}\|_2}{\sqrt{\hat{C}(\tau, \alpha, \beta)}} \\ &\leq \frac{\|\mathcal{P}_{T_1^\perp} \times \cdots \times \mathcal{P}_{T_\tau^\perp} h^{\gamma^\perp}\|_2}{\sqrt{\hat{C}(\tau, \alpha, \beta)}} \\ &\leq C(n, \tau, \alpha, \beta) \delta,\end{aligned}\tag{12}$$

where $C(n, \tau, \alpha, \beta)$ is an appropriate constant. Combining with (9), we can obtain

$$\begin{aligned}\|h\|_2^2 &= \|h^\Gamma\|_2^2 + \|\mathcal{P}_{T_1} \times \cdots \times \mathcal{P}_{T_\tau} h^{\Gamma^\perp}\|_2^2 \\ &\quad + \|\mathcal{P}_{T_1^\perp} \times \cdots \times \mathcal{P}_{T_\tau^\perp} h^{\Gamma^\perp}\|_2^2 \\ &\leq \hat{C}(n, \tau, \alpha, \beta) \delta^2.\end{aligned}$$

Therefore, Theorem 1 is established.

Remark 1. if $\tau = 2$, then Theorem 1 will degrade to the main result of [13].

5. Numerical Results

In this section, numerical experiments with varieties of the value of parameter σ , parameter ρ_s , and rank r are given. For each setting of parameters, we show the average errors over 10 trials. Our implementation was realized with MATLAB. All the computational results were obtained on a desktop computer with a 2.27-GHz CPU (Intel(R) Core(TM) i3) and 2 GB of memory. Without loss of generality, we assume that $\tau = 2$. In [13], the authors certified this result with Accelerated Proximal Gradient (APG) by numerical experiments. In our numerical experiments, we will provide that this result is also proper with Principal Component Pursuit by Alternating Direction Method (PCP-ADM). In our simulations, our matrix is generated by the formulation as: $M = L_0 + S_0 + N_0$, and a rank- r matrix L_0 is a product $L_0 = XY^T$, where X and Y are $m \times r$ and $n \times r$ matrices in which entries are independently sampled from a $\mathcal{N}(0; 1)$ distribution. According to PCP-ADM, we can generate S_0 by choosing a support set of size $k_s = \rho_s mn$ uniformly at random, and set $S_0 = \mathcal{P}_\Omega E$. Noise component N_0 is generated with entries independently sampled from a $\mathcal{N}(0; \sigma)$ distribution. Without loss of generality, we set $m = n = 200$ and $\rho_s = 0.01$, and other parameters which PCP-ADM requires are the same as parameters of PCP-ADM [10]. Here we briefly interpret PCP-ADM. In [10], in order to stably recover $\hat{X} = (\hat{L}; \hat{S})$, the ADM method operates on the *augmented Lagrangian*

$$l(L, S, Y) = \|L\|_* + \lambda \|S\|_1 + \langle Y, M - L - S \rangle + \frac{\mu}{2} \|M - L - S\|_F^2.$$

The details of the PCP-ADM can be found in [14,15].

In our simulations, the stopping criterion of the PCP-ADM algorithm can be

$$\frac{\|L + S - M\|_F}{\|M\|_F} \leq \text{tolerance}$$

or the maximum iteration number ($k_{\max} = 500$). In order to estimate the errors, we use the root-mean-squared (RMS) error as $\|\hat{L} - L_0\|_F/n$, $\|\hat{S} - S_0\|_F/n$ for the low-rank component and the

sparse component, respectively. Figure 1 shows the RMS errors' variation with different values of σ^2 . It is noted that the RMS error grows approximately linearly with the noise level in Figure 1. This phenomenon verifies Theorem 1 by numerical experiments with PCP-ADM (this phenomenon also exists in [13] with APG, which is very different from PCP-ADM in principle).

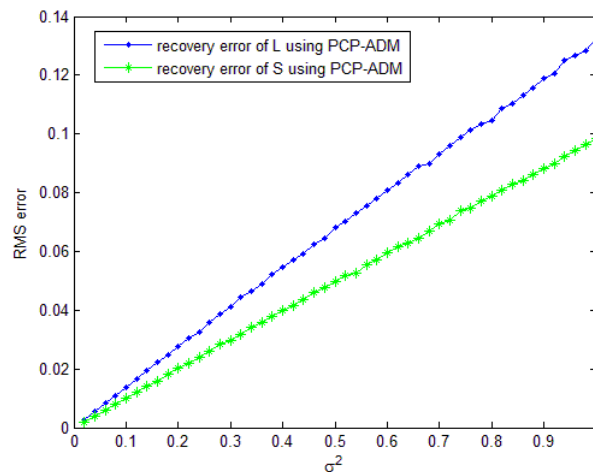


Figure 1. Root-mean-squared (RMS) errors as a function of σ^2 with $r = 10$; $\rho_s = 0.01$; $n = 200$. PCP-ADM: Principal Component Pursuit by Alternating Direction Method.

6. Conclusions

In this paper, we have investigated the the stability of CPCP. Our main contribution is the proof of Theorem 1, which implies the solution to the related convex programming (1.2) is stable to small entrywise noise under board condition. It is an extension of the result in [13], which only allows $\tau = 2$. Moreover, in the numerical experiments, we have investigated the performance of the PCP-ADM algorithm. Numerical results showed that it is stable to small entrywise noise.

Acknowledgments: The author would like to thank the anonymous reviewers for their comments that helped to improve the quality of the paper. This research was supported by the National Natural Science Foundation of China (NSFC) under Grant U1533125, and the Scientific Research Program of the Education Department of Sichuan under Grant 16ZB0032.

Author Contributions: Qingshan You and Qun Wan. contributed reagents/materials/analysis tools; Qingshan You wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Candès, E.J.; Recht, B. Exact matrix completion via convex optimization. *Found. Comput. Math.* **2009**, *9*, 717–772.
2. Candès, E.J.; Plan, Y. Matrix completion with noise. *Proc. IEEE* **2010**, *98*, 925–936.
3. Candès, E.J.; Tao, T. The power of convex relaxation: Near-optimal matrix completion. *IEEE Trans. Inf. Theory* **2010**, *56*, 2053–2080.
4. Ellenberg, J. Fill in the blanks: Using math to turn lo-res datasets into hi-res samples. *Wired* **2010**. Available online: https://www.wired.com/2010/02/ff_algorithm/all/1 (accessed on 26 January 2016).
5. Antonin Chambolle and Pierre-Louis Lions. Image recovery via total variation minimization and related problems. *Numer. Math.* **1997**, *76*, 167–188.
6. Jon, F. Claerbout and Francis Muir. Robust modeling of erratic data. *Geophysics* **1973**, *38*, 826–844, .
7. Zeng, B.; Fu, J. Directional discrete cosine transforms: A new framework for image coding. *IEEE Trans. Circuits Syst. Video Technol.* **2011**, *18*, 305–313.

8. Elad, M.; Aharon, M. Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Trans. Image Process.* **2006**, *15*, 3736–3745.
9. Rodger, J.A. Toward reducing failure risk in an integrated vehicle health maintenance system: A fuzzy multi-sensor data fusion Kalman filter approach for IVHMS. *Expert Syst. Appl.* **2012**, *39*, 9821–9836.
10. Candès, E.J.; Li, X.; Ma, Y.; Wright, J. Robust principal component analysis? *J. ACM* **2011**, doi:10.1145/1970392.1970395.
11. Wright, J.; Ganesh, A.; Min, K.; Ma, Y. Compressive Principal Component Pursuit. Available online: <http://yima.csl.illinois.edu/psfile/CPCP.pdf> (accessed on 9 April 2012).
12. Recht, B.; Fazel, M.; Parrilo, P. Guaranteed minimum rank solutions of matrix equations via nuclear norm minimization. *arXiv* **2007**, arxiv:0706.4138.
13. Zhou, Z.; Li, X.; Wright, J.; Candès, E.J.; Ma, Y. Stable Principal Component Pursuit. *arXiv* **2010**, arXiv:1001.2363v1.
14. Yuan, X.; Yang, J. Sparse and low rank matrix decomposition via alternating direction method. *Pac. J. Optim.* **2009**, *9*, 167–180.
15. Kontogiorgis, S.; Meyer, R. A variable-penalty alternating direction method for convex optimization. *Math. Program.* **1989**, *83*, 29–53.



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).