On the minimal number of measurements in low-rank matrix recovery

Maryia Kabanava, Holger Rauhut, Ulrich Terstiege
RWTH Aachen University
Chair for Mathematics C (Analysis)
Pontdriesch 10, 52062 Aachen
Email: {kabanava,rauhut,terstiege}@mathc.rwth-aachen.de

***Abstract*—In this paper we present a new way to obtain a bound on the number of measurements sampled from certain distributions that guarantee uniform stable and robust recovery of low-rank matrices. The recovery guarantees are characterized by a stable and robust version of the null space property and verifying this condition can be reduced to the problem of obtaining a lower bound for a quantity of the form $\inf_{x \in T} \|Ax\|_2$. Gordon's escape through a mesh theorem provides such a bound with explicit constants for Gaussian measurements. Mendelson's small ball method allows to cover the significantly more general case of measurements generated by independent identically distributed random variables with finite fourth moment.**

I. Introduction

In the present paper we study the problem of uniform stable and robust recovery of low-rank matrices from undersampled measurements. Given noisy data

$$b = \mathcal{A}(X) + w, \quad \|w\|_2 \leq \eta,$$

where $\mathcal{A} : \mathbb{R}^{n_1 \times n_2} \to \mathbb{R}^m$, $m \ll n_1 n_2$, is a linear map and $w \in \mathbb{R}^m$ corresponds to noise, we recover $X \in \mathbb{R}^{n_1 \times n_2}$ by solving the nuclear norm minimization problem

$$\min_{Z \in \mathbb{R}^{n_1 \times n_2}} \|Z\|_* \quad \text{subject to } \|\mathcal{A}(Z) - b\|_2 \leq \eta. \tag{1}$$

It was observed in [1], that the rank-restricted isometry property (rank-RIP) is sufficient for (1) to provide the minimum-rank solution from noiseless observations with $\eta = 0$. The stability and robustness properties of the nuclear norm minimization were addressed in [2], [3] and the obtained error bounds rely on the rank-RIP of the measurement map, which was shown to hold with high probability for certain random measurement ensembles in [1], [3].

Our recovery guarantees are based on a certain property of the null space of the measurement map. It was first mentioned in [4] that the condition called the *null space property* is necessary and sufficient for exact reconstruction of a matrix of the lowest rank via (1) when $\eta = 0$. A slightly stronger version of the null space property also implies stability, when the nuclear norm error of recovering a matrix via the nuclear norm minimization is controlled by the error of its best possible low-rank approximation. We enhance the previous results by strengthening the null space property in order to provide Frobenius norm error estimates and incorporate noise on the measurements.

Since it is difficult to verify the provided theoretical guarantees for deterministic measurements, we pass to random maps and obtain corresponding bounds on the number of Gaussian measurements and measurements with finite fourth moment. In particular, it was unknown before and perhaps surprisingly that only finite fourth moments are required. So, for example, in the recent paper [5] the authors prove the exact recovery of sparse vectors in $\mathbb{R}^n$ with the optimal number of measurements under the assumption of finite $\log n$-th moments.

II. Basic definitions and notation

Let $X \in \mathbb{R}^{n_1 \times n_2}$ and $n := \min\{n_1, n_2\}$. A factorization of X of the form

$$X = U \Sigma V^*,$$

where $U \in \mathbb{R}^{n_1 \times n}$, $V \in \mathbb{R}^{n_2 \times n}$, $U^*U = V^*V = \mathrm{Id}$ and $\Sigma \in \mathbb{R}^{n \times n}$ is diagonal with non-negative non-increasing entries, is called the singular value decomposition of X. The diagonal entries of Σ are the singular values of X and they are collected in the vector $\sigma(X)$. The Schatten p-norm of $X \in \mathbb{R}^{n_1 \times n_2}$ is defined by

$$\|X\|_p = \left(\sum_{j=1}^{n} \sigma_j(X)^p \right)^{1/p}, \quad p \geq 1.$$

It reduces to the nuclear norm $\|\cdot\|_*$ for $p = 1$ and the Frobenius norm $\|\cdot\|_F$ for $p = 2$. The best rank-r approximation to X is given by the matrix $\sum_{j=1}^{r} \sigma_j(X) u_j v_j^*$ and the Schatten p-norm of the approximation error is equal to

$$\inf_{\substack{M \in \mathbb{R}^{n_1 \times n_2} \\ \operatorname{rank} M \leq r}} \|X - M\|_p = \left(\sum_{j=r+1}^{n} \sigma_j(X)^p \right)^{1/p}.$$

The inner product of $X_1 \in \mathbb{R}^{n_1 \times n_2}$ and $X_2 \in \mathbb{R}^{n_1 \times n_2}$ is defined by $\langle X_1, X_2 \rangle = \operatorname{tr} X_2^* X_1$.

III. Main results

A random linear map can be described in terms of random matrices. When $\mathcal{A} : \mathbb{R}^{n_1 \times n_2} \to \mathbb{R}^m$ is linear, there exists $A \in \mathbb{R}^{m \times n_1 n_2}$, such that $\mathcal{A}(X) = A \operatorname{vec}(X)$, where $\operatorname{vec}(X)$

978-1-4673-7353-1/15/$31.00

is obtained by concatenating all the columns of X into a single vector in $\mathbb{R}^{n_1 n_2}$.

Theorem 1: Let $\mathcal{A}: \mathbb{R}^{n_1\times n_2} \to \mathbb{R}^m$ be a linear map whose matrix has independent standard Gaussian entries, $0<\rho<1$, $\kappa>1$ and $0<\varepsilon<1$. If

$$\frac{m^2}{m+1} \geq \frac{r(1+(1+\rho^{-1})^2)\kappa^2}{(\kappa-1)^2}\left[\sqrt{n_1}+\sqrt{n_2} + \sqrt{\frac{2\ln(\varepsilon^{-1})}{r(1+(1+\rho^{-1})^2)}}\right]^2, \tag{2}$$

then with probability at least $1-\varepsilon$ for every $X\in\mathbb{R}^{n_1\times n_2}$ a solution $\hat{X}$ of (1) with $b=\mathcal{A}(X)+w$, $\|w\|_2\leq\eta$, approximates X with the error

$$\|X-\hat{X}\|_2 \leq \frac{2(1+\rho)^2}{(1-\rho)\sqrt{r}}\sum_{j=r+1}^{n}\sigma_j(X) + \frac{2\kappa\sqrt{2}(3+\rho)}{\sqrt{m}(1-\rho)}\eta.$$

Roughly speaking, any matrix of rank r is recovered with high probability from

$$m > 5r(\sqrt{n_1}+\sqrt{n_2})^2$$

Gaussian measurements. Since any $n_1\times n_2$ matrix of rank r is determined by $r(n_1+n_2-r)$ parameters, the provided bound is within a constant of the optimal result. In [1] the authors exploit the rank-RIP to show that any matrix of rank r can be recovered from $O(r(n_1+n_2)\log(n_1 n_2))$ Gaussian measurements. A more refined analysis of the rank-RIP in [3] allowed to get rid of the extra log-factor. Estimates of optimal order based on the null space approach are presented in [6], [7]. However, the advantage of our result is that it provides an explicit constant.

Theorem 1 can be extended to a more general setting. However, the constant in the estimate for the number of measurements is not explicit.

Theorem 2: Let $\mathcal{A}: \mathbb{R}^{n_1\times n_2} \to \mathbb{R}^m$ be a linear map whose matrix has independent identically distributed entries A_{ij} with

$$\mathbb{E}\, A_{ij} = 0, \mathbb{E}\, A_{ij}^2 = 1 \text{ and } \mathbb{E}\, A_{ij}^4 \leq C_4.$$

Let $0<\rho<1$, $\kappa>1$, $0<\varepsilon<1$ and suppose that

$$m \geq \frac{cr(1+(1+\rho^{-1})^2)\kappa^2}{(\kappa-1)^2}\left[\sqrt{\max\{n_1,n_2\}} + \sqrt{\frac{\ln(\varepsilon^{-1})}{r(1+(1+\rho^{-1})^2)}}\right]^2, \tag{3}$$

where $c>0$ depends only on C_4. Then with probability at least $1-\varepsilon$, for any $X\in\mathbb{R}^{n_1\times n_2}$ a solution $\hat{X}$ of (1) with $b=\mathcal{A}(X)+w$, $\|w\|_2\leq\eta$, approximates X with the error

$$\|X-\hat{X}\|_2 \leq \frac{C_\rho}{\sqrt{r}}\sum_{j=r+1}^{n}\sigma_j(X) + D_{\kappa,\rho}\frac{\eta}{\sqrt{m}},\ C_\rho, D_{\kappa,\rho}>0.$$

We prove Theorem 1 and 2 by establishing the Frobenius robust rank null space property (to be defined in Section IV) for given measurements. In the Gaussian setting we rely on Gordon's escape through a mesh theorem [8], [9]. For the case of more general measurements of Theorem 2 we refer to Mendelson's small ball method [10], [11], [12], [13].

IV. Frobenius stable null space property

The Frobenius stability and robustness of the recovery of low-rank matrices via nuclear norm minimization is based on the following property of the measurement map.

Definition 3: We say that $\mathcal{A}: \mathbb{R}^{n_1\times n_2} \to \mathbb{R}^m$ satisfies the Frobenius robust rank null space property of order r with constants $0<\rho<1$ and $\tau>0$ if for all $M\in\mathbb{R}^{n_1\times n_2}$, the singular values of M satisfy

$$\left(\sum_{j=1}^{r}\sigma_j(M)^2\right)^{1/2} \leq \frac{\rho}{\sqrt{r}}\sum_{j=r+1}^{n}\sigma_j(M) + \tau\|\mathcal{A}(M)\|_2.$$

Theorem 4: Let $\mathcal{A}: \mathbb{R}^{n_1\times n_2} \to \mathbb{R}^m$ satisfy the Frobenius robust rank null space property of order r with constants $0<\rho<1$ and $\tau>0$. Then for any $X\in\mathbb{R}^{n_1\times n_2}$ the solution $\hat{X}$ of (1) with $b=\mathcal{A}(X)+w$, $\|w\|_2\leq\eta$, approximates X with the error

$$\|X-\hat{X}\|_2 \leq \frac{2(1+\rho)^2}{(1-\rho)\sqrt{r}}\sum_{j=r+1}^{n}\sigma_j(X) + \frac{2\tau(3+\rho)}{1-\rho}\eta.$$

One of the ways to prove Theorem 4 is to rely on results in [14], which show that if some condition is sufficient for stable and robust recovery of any sparse vector with at most r non-zero entries, then the modification of this condition is sufficient for the stable and robust recovery of any matrix up to rank r.

In order to check whether the measurement map $\mathcal{A}: \mathbb{R}^{n_1\times n_2} \to \mathbb{R}^m$ satisfies the Frobenius robust rank null space property, we introduce the set

$$T_{\rho,r} := \left\{M\in\mathbb{R}^{n_1\times n_2} : \|M\|_2 = 1, \left(\sum_{i=1}^{r}\sigma_i(M)^2\right)^{1/2} > \frac{\rho}{\sqrt{r}}\sum_{i=r+1}^{n}\sigma_i(M)\right\}.$$

If

$$\inf\{\|\mathcal{A}(M)\|_2 : M\in T_{\rho,r}\} > \frac{1}{\tau}, \tag{4}$$

then for any $M\in\mathbb{R}^{n_1\times n_2}$ such that $\|\mathcal{A}(M)\|_2 \leq \frac{\|M\|_2}{\tau}$ it holds

$$\left(\sum_{i=1}^{r}\sigma_i(M)^2\right)^{1/2} \leq \frac{\rho}{\sqrt{r}}\sum_{i=r+1}^{n}\sigma_i(M). \tag{5}$$

For the remaining $M\in\mathbb{R}^{n_1\times n_2}$ with $\|\mathcal{A}(M)\|_2 > \frac{\|M\|_2}{\tau}$ we have

$$\left(\sum_{i=1}^{r}\sigma_i(M)^2\right)^{1/2} \leq \|M\|_2 < \tau\|\mathcal{A}(M)\|_2.$$

Altogether with (5) this leads to

$$\left(\sum_{i=1}^{r}\sigma_i(M)^2\right)^{1/2} \leq \frac{\rho}{\sqrt{r}}\sum_{i=r+1}^{n}\sigma_i(M) + \tau\|\mathcal{A}(M)\|_2.$$

It is natural to expect that the recovery error gets smaller as the number of measurements increases. This can be taken into account by establishing the null space property for $\tau = \frac{\kappa}{\sqrt{m}}$. Then the error bound reads as follows

$$\|X - \hat{X}\|_2 \le \frac{2(1+\rho)^2}{(1-\rho)\sqrt{r}} \sum_{j=r+1}^{n} \sigma_j(X) + \frac{2\kappa(3+\rho)}{\sqrt{m}(1-\rho)}\eta.$$

An important property of the set $T_{\rho,r}$ is that it is embedded in a set with a simple structure.

Lemma 5: Let D be the set defined by

$$D := \operatorname{conv}\left\{M \in \mathbb{R}^{n_1\times n_2} : \|M\|_2 = 1, \operatorname{rank} M \le r\right\}, \quad (6)$$

where conv stands for the convex hull.

1) Then D is the unit ball with respect to the norm

$$\|M\|_D := \sum_{j=1}^{L}\left[\sum_{i\in I_j}(\sigma_i(M))^2\right]^{1/2},$$

where $L = \lceil \frac{n}{r} \rceil$,

$$I_j = \begin{cases} \{r(j-1)+1,\dots,rj\}, & j = 1,\dots,L-1,\\ \{r(L-1)+1,\dots,n\}, & j = L.\end{cases}$$

2) It holds

$$T_{\rho,r} \subset \sqrt{1+(1+\rho^{-1})^2}D. \quad (7)$$

Let us argue briefly why $\|\cdot\|_D$ is a norm. Define $g : \mathbb{R}^n \to [0,\infty)$ by

$$g(x) := \sum_{j=1}^{L}\left(\sum_{i\in I_j}(x_i^*)^2\right)^{1/2},$$

where L and I_j are defined in the same way as in item 1) of Lemma 5 and x^* is obtained by arranging the entries of x in the decreasing order of magnitude. Then g is a symmetric gauge function and $\|M\|_D = g(\sigma(M))$ for any $M \in \mathbb{R}^{n_1\times n_2}$. According to Theorem 7.4.7.2 in [15] $\|\cdot\|_D$ is indeed a norm.

Proof of Lemma 5: 1) Any $M \in D$ can be written as

$$M = \sum_i \alpha_i X_i$$

with

$$\operatorname{rank} X_i \le r,\ \|X_i\|_2 = 1,\ \alpha_i \ge 0,\ \sum_i \alpha_i = 1.$$

Thus

$$\|M\|_D \le \sum_i \alpha_i\|X_i\|_D = \sum_i \alpha_i\|X_i\|_2 = \sum_i \alpha_i = 1.$$

Conversely, suppose that $\|M\|_D \le 1$. Let M have a singular value decomposition $M = U\Sigma V^* = \sum_{j=1}^{L}\sum_{i\in I_j}\sigma_i(M)u_iv_i^*$, where $u_i \in \mathbb{R}^{n_1}$ and $v_i \in \mathbb{R}^{n_2}$ are column vectors of U and V respectively. Set $M_j := \sum_{i\in I_j}\sigma_i(M)u_iv_i^*$ and $\alpha_j := \|M_j\|_2$, $j = 1,\dots,L$. Then each M_j is a sum of r rank-one matrices, so that $\operatorname{rank} M_j \le r$, and we can write M as

$$M = \sum_{j:\alpha_j\ne 0}\alpha_j\left(\frac{1}{\alpha_j}M_j\right)$$

with

$$\sum_{j:\alpha_j\ne 0}\alpha_j = \sum_j \|M_j\|_2 = \|M\|_D \le 1$$

and

$$\|\frac{1}{\alpha_j}M_j\|_2 = \frac{1}{\alpha_j}\|M_j\|_2 = 1.$$

Hence $M \in D$.

2) To prove the embedding of $T_{\rho,r}$ into a blown-up version of D, we estimate the norm of an arbitrary element of $T_{\rho,r}$.

Suppose $M \in T_{\rho,r}$. According to the definition of the $\|\cdot\|_D$-norm

$$\|M\|_D = \sum_{l=1}^{L}\left[\sum_{i\in I_l}(\sigma_i(M))^2\right]^{\frac12} = \left[\sum_{i=1}^{r}(\sigma_i(M))^2\right]^{\frac12} + \left[\sum_{i=r+1}^{2r}(\sigma_i(M))^2\right]^{\frac12} + \sum_{l\ge 3}^{L}\left[\sum_{i\in I_l}(\sigma_i(M))^2\right]^{\frac12}. \quad (8)$$

To bound the last term in the inequality above, we first note that for each $i \in I_l$, $l \ge 3$,

$$\sigma_i(M) \le \frac{1}{r}\sum_{j\in I_{l-1}}\sigma_j(M)$$

and

$$\left[\sum_{i\in I_l}(\sigma_i(M))^2\right]^{1/2} \le \frac{1}{\sqrt{r}}\sum_{j\in I_{l-1}}\sigma_j(M).$$

Summing up over $l \ge 3$ yields

$$\sum_{l\ge 3}^{L}\left[\sum_{i\in I_l}(\sigma_i(M))^2\right]^{\frac12} \le \frac{1}{\sqrt{r}}\sum_{l\ge 2}\sum_{j\in I_l}\sigma_j(M) = \frac{1}{\sqrt{r}}\sum_{j=r+1}^{n}\sigma_j(M)$$

and taking into account the inequality for the singular values of $M \in T_{\rho,r}$

$$\sum_{l\ge 3}^{L}\left[\sum_{i\in I_l}(\sigma_i(M))^2\right]^{\frac12} \le \rho^{-1}\left[\sum_{i=1}^{r}(\sigma_i(M))^2\right]^{\frac12}.$$

Applying the last estimate to (8) we derive that

$$\|M\|_D \le (1+\rho^{-1})\left[\sum_{i=1}^{r}(\sigma_i(M))^2\right]^{\frac12} + \left[\sum_{i=r+1}^{2r}(\sigma_i(M))^2\right]^{\frac12} \le (1+\rho^{-1})\left[\sum_{i=1}^{r}(\sigma_i(M))^2\right]^{\frac12} + \left[1 - \sum_{i=1}^{r}(\sigma_i(M))^2\right]^{\frac12}.$$

Set $a = \left[\sum_{i=1}^{r}(\sigma_i(M))^2\right]^{\frac12}$. The maximum of the function

$$f(a) := (1+\rho^{-1})a + \sqrt{1-a^2}, \quad 0 \le a \le 1,$$

is attained at the point

$$a = \frac{1+\rho^{-1}}{\sqrt{1+(1+\rho^{-1})^2}}$$

and is equal to $\sqrt{1+(1+\rho^{-1})^2}$. Thus for any $M \in T_{\rho,r}$ it holds

$$\|M\|_D \le \sqrt{1+(1+\rho^{-1})^2},$$

which proves (7). ∎

Employing the matrix representation of the measurement map $\mathcal{A}$, the problem of estimating the probability of the event (4) is reduced to the problem of giving a lower bound for the quantities of the form $\inf_{x\in T}\|Ax\|_2$. This is not an easy task for deterministic matrices, but the situation significantly changes for matrices chosen at random.

V. Gaussian measurements

To estimate the probability of the event (4) for Gaussian measurements we employ Gordon's escape through a mesh theorem. First we recall some definitions. Let $g \in \mathbb{R}^m$ be a standard Gaussian random vector, that is, a vector of normal distributed random variables. Then for

$$E_m := \mathbb{E}\,\|g\|_2 = \sqrt{2}\,\frac{\Gamma\left((m+1)/2\right)}{\Gamma\left(m/2\right)}$$

we have

$$\frac{m}{\sqrt{m+1}} \le E_m \le \sqrt{m},$$

see [8], [9]. For a set $T \subset \mathbb{R}^n$ we define its Gaussian width by

$$\ell(T) := \mathbb{E}\sup_{x\in T}\langle x, g\rangle,$$

where $g \in \mathbb{R}^n$ is a standard Gaussian random vector.

Theorem 6 (Gordon's escape through a mesh): Let $A \in \mathbb{R}^{m\times n}$ be a Gaussian random matrix and T be a subset of the unit sphere $\mathbb{S}^{n-1} = \{x \in \mathbb{R}^n : \|x\|_2 = 1\}$. Then, for $t > 0$,

$$\mathbb{P}\left(\inf_{x\in T}\|Ax\|_2 > E_m - \ell(T) - t\right) \ge 1 - e^{-\frac{t^2}{2}}. \quad (9)$$

In order to give a bound on the number of Gaussian measurements, Theorem 6 suggests to estimate from above the Gaussian width of the set $T_{\rho,r}$. As it was pointed out in the previous section, $T_{\rho,r}$ is a subset of a slightly enlarged version of D, which has a relatively simple structure. So instead of evaluating $\ell(T_{\rho,r})$, we consider $\ell(D)$.

Lemma 7: For the set D defined by (6) it holds

$$\ell(D) \le \sqrt{r}(\sqrt{n_1}+\sqrt{n_2}). \quad (10)$$

Proof: Let $\Gamma \in \mathbb{R}^{n_1\times n_2}$ have independent standard normal distributed entries. Then $\ell(D) = \mathbb{E}\sup_{M\in D}\langle \Gamma, M\rangle$. Since a convex continuous real-valued function attains its maximum value at one of the extreme points, it holds

$$\ell(D) = \mathbb{E}\sup_{\substack{\|M\|_2=1\\ \operatorname{rank} M=r}}\langle \Gamma, M\rangle.$$

By Hölder's inequality

$$\begin{aligned}\ell(D) &\le \mathbb{E}\sup_{\substack{\|M\|_2=1\\ \operatorname{rank} M=r}}\|\Gamma\|_\infty\|M\|_1 \le \sqrt{r}\sup_{\substack{\|M\|_2=1\\ \operatorname{rank} M=r}}\|M\|_2\,\mathbb{E}\,\sigma_1(\Gamma)\\ &\le \sqrt{r}(\sqrt{n_1}+\sqrt{n_2}).\end{aligned}$$

The last inequality used a well-known estimate of the expectation of the largest singular value of a standard Gaussian matrix, see [9, Chapter 9.3]. ∎

Lemma 5 and Lemma 7 allow to conclude that

$$\begin{aligned}\ell(T_{\rho,r}) &\le \sqrt{1+(1+\rho^{-1})^2}\,\ell(D)\\ &\le \sqrt{r(1+(1+\rho^{-1})^2)}(\sqrt{n_1}+\sqrt{n_2}).\end{aligned}$$

Proof of Theorem 1: Set $t := \sqrt{2\ln(\varepsilon^{-1})}$. If m is as in (2), then

$$E_m\left(1-\frac{1}{\kappa}\right) \ge \sqrt{r(1+(1+\rho^{-1})^2)}(\sqrt{n_1}+\sqrt{n_2}) + t.$$

Together with (7) and (10) this yields

$$E_m - \ell(T_{\rho,r}) - t \ge \frac{E_m}{\kappa} \ge \frac{1}{\kappa}\sqrt{\frac{m}{2}}.$$

According to Theorem 6

$$\mathbb{P}\left(\inf_{M\in T_{\rho,r}}\|\mathcal{A}(M)\|_2 > \frac{\sqrt{m}}{\kappa\sqrt{2}}\right) \ge 1-\varepsilon,$$

which means that with probability at least $1-\varepsilon$ map $\mathcal{A}$ satisfies the Frobenius robust rank null space property with constants ρ and $\frac{\kappa\sqrt{2}}{\sqrt{m}}$. The error estimate follows by Theorem 4. ∎

VI. Measurements with finite fourth moment

To extend Theorem 1 to a larger class of random maps where we only require finite fourth moments, we apply Mendelson's small ball method.

Theorem 8 ([11], [12], [13]): Fix $E \subset \mathbb{R}^d$ and let $\phi_1, \ldots, \phi_m$ be independent copies of a random vector ϕ in $\mathbb{R}^d$. For $\xi > 0$ let

$$Q_\xi(E;\phi) = \inf_{u\in E}\mathbb{P}\{|\langle \phi, u\rangle| \ge \xi\}$$

and

$$W_m(E;\phi) = \mathbb{E}\sup_{u\in E}\langle h, u\rangle,$$

where $h = \frac{1}{\sqrt{m}}\sum_{j=1}^m \varepsilon_j\phi_j$ with $\{\varepsilon_j\}$ being a Rademacher sequence. Then for any $\xi > 0$ and any $t \ge 0$ with probability at least $1-e^{-2t^2}$

$$\inf_{u\in E}\left(\sum_{i=1}^m |\langle \phi_i, u\rangle|^2\right)^{1/2} \ge \xi\sqrt{m}Q_{2\xi}(E;\phi) - 2W_m(E;\phi) - \xi t.$$

As before we identify $n_1\times n_2$ real-valued matrix with a vector in $\mathbb{R}^{n_1 n_2}$. Suppose that the rows of the matrix A of a linear map $\mathcal{A}: \mathbb{R}^{n_1\times n_2} \to \mathbb{R}^m$ are given by vectors ϕ_i, $i = 1, \ldots, m$. In accordance with Theorem 8 in order to bound

$$\inf_{u\in T_{\rho,r}}\|Au\|_2 = \inf_{u\in T_{\rho,r}}\left(\sum_{i=1}^m |\langle \phi_i, u\rangle|^2\right)^{1/2}$$

we need to estimate $Q_{2\xi}(T_{\rho,r};\phi)$ and $W_m(T_{\rho,r};\phi)$. For a detailed proof of Theorem 2 we refer to [16]. We only mention that a lower bound for $Q_{2\xi}(T_{\rho,r};\phi)$ relies on the fact that the elements of $T_{\rho,r}$ have unit Frobenius norm and an upper bound for $W_m(T_{\rho,r};\phi)$ exploits the inclusion (7).

Acknowledgment

The authors acknowledge support by the European Research Council through the grant StG 258926.

References

[1] B. Recht, M. Fazel, and P. A. Parrilo, "Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization," *SIAM Rev.*, vol. 52, no. 3, pp. 471–501, 2010.

[2] K. Mohan and M. Fazel, "New restricted isometry results for noisy low-rank recovery," in *Proc. International Symposium Information Theory*, 2010.

[3] E. J. Candès and Y. Plan, "Tight oracle inequalities for low-rank matrix recovery from a minimal number of noisy random measurements," *IEEE Trans. Inform. Theory*, vol. 57, no. 4, pp. 2342–2359, 2011.

[4] B. Recht, W. Xu, and B. Hassibi, "Necessary and sufficient conditions for success of the nuclear norm heuristic for rank minimization," in *Proc. 47th IEEE Conference on Decision and Control*, 2008, pp. 3065–3070.

[5] G. Lecuè and S. Mendelson, "Sparse recovery under weak moment assumptions," *http://arxiv.org/abs/1401.2188*, 2014.

[6] B. Recht, W. Xu, and B. Hassibi, "Null space conditions and thresholds for rank minimization," *Math. Program.*, vol. 127, no. 1, Ser. B, pp. 175–202, 2011.

[7] K. Dvijotham and M. Fazel, "A nullspace analysis of the nuclear norm heuristic for rank minimization," *Proc. of ICASSP 2010*, 2010.

[8] Y. Gordon, "On Milman's inequality and random subspaces which escape through a mesh in $\mathbb{R}^n$," in *Geometric aspects of functional analysis (1986/87)*, ser. Lecture Notes in Math. Springer, Berlin, 1988, vol. 1317, pp. 84–106.

[9] S. Foucart and H. Rauhut, *A mathematical introduction to compressive sensing*, ser. Applied and Numerical Harmonic Analysis. Birkhäuser/Springer, New York, 2013.

[10] S. Mendelson, "A remark on the diameter of random sections of convex bodies," *http://arxiv.org/abs/1312.3608*, 2013.

[11] V. Koltchinskii and S. Mendelson, "Bounding the smallest singular value of a random matrix without concentration," *http://arxiv.org/abs/1312.3580*, 2013.

[12] S. Mendelson, "Learning without concentration," *http://arxiv.org/abs/1401.0304*, 2014.

[13] J. A. Tropp, "Convex recovery of a structured signal from independent random linear measurements," *http://arxiv.org/abs/1405.1102*, 2014.

[14] S. Oymak, K. Mohan, M. Fazel, and B. Hassibi, "A simplified approach to recovery conditions for low rank matrices," *http://arxiv.org/abs/1103.1178*, 2011.

[15] R. Horn and C. Johnson, *Topics in matrix analysis*. Cambridge University Press, Cambridge, 1991.

[16] M. Kabanava, R. Kueng, H. Rauhut, and U. Terstiege, "Analysis of low-rank matrix recovery via Mendelson's small ball method," *submitted*, 2015.