On critical points of quadratic low-rank matrix optimization problems

by

André Uschmajew and Bart Vandereycken

Preprint no.: 58

2018
On critical points of quadratic low-rank matrix optimization problems

André Uschmajew*    Bart Vandereycken†

July 18, 2018

Abstract

The absence of spurious local minima in certain non-convex low-rank matrix recovery problems has been of recent interest in computer science, machine learning, and compressed sensing since it explains the convergence of some low-rank optimization methods to global optima. One such example is low-rank matrix sensing under restricted isometry properties (RIP). It can be formulated as a minimization problem for a quadratic function on a low-rank matrix manifold, with a positive semidefinite Hessian that acts almost like an identity on low-rank matrices. In this work, new estimates for singular values of local minima for such problems are given which lead to improved bounds on RIP constants to ensure absence of non-optimal local minima and sufficiently negative curvature at all other critical points. A geometric viewpoint is taken which is inspired by the fact that the Euclidean distance function to a rank $k$ matrix possesses no critical points on the corresponding embedded submanifold of rank $k$ matrices except for a single global minimum.

1 Introduction

On the space $\mathbb{R}^{m \times n}$ of real $m \times n$ matrices we consider a quadratic function

$$f_{A,B}(X) = \frac{1}{2} \langle A[X], X \rangle_F - \langle B, X \rangle_F$$

(1.1)

with a given symmetric operator $A$ and matrix $B$. The gradient of this function equals

$$\nabla f_{A,B}(X) = A[X] - B$$

and critical (or stationary) points of the function $f_{A,B}$ thus correspond to solutions of the linear matrix equation

$$A[X] = B.$$  

(1.2)

Thus, critical points only exist if $B$ is in the range of $A$. This is, for instance, the case when $A$ is positive definite with respect to the Frobenius inner product, in which case the solution to the matrix equation $A[X] = B$ is also unique.

For $m = n$ and $L$ a symmetric positive definite $n \times n$ matrix, a well-known example of the type described above is the Lyapunov matrix equation

$$LX + XL = B,$$

(1.3)

which has a unique solution $X^*$ for any right-hand side $B$, since the operator $A[X] = LX + XL$ is symmetric positive definite on $\mathbb{R}^{m \times n}$.

*Max Planck Institute for Mathematics in the Sciences, 04103 Leipzig, Germany (uschmajew@mis.mpg.de)
†University of Geneva, Section of Mathematics, 1211 Geneva, Switzerland (bart.vandereycken@unige.ch)
1.1 Rank constrained quadratic problems

In certain applications, the aim is to solve (1.1) or (1.2) with the additional requirement that the solution (or its approximation) is of sufficiently low rank. For instance, when $B$ in the Lyapunov equation (1.3) itself is low rank, it can be proven [29] that $X^*$ has exponentially decaying singular values and, hence, can be approximated well by a low-rank matrix. To obtain such approximations, there exists a few methods that are built from classical solvers in numerical linear algebra, like the ADI and Krylov methods; see [28] for a recent overview.

Let $k \leq \min(m, n)$ and denote by

$$M_k = \{ X \in \mathbb{R}^{m \times n} : \text{rank}(X) = k \}$$

the smooth manifold of fixed rank $k$ matrices, and by

$$M_{\leq k} = \{ X \in \mathbb{R}^{m \times n} : \text{rank}(X) \leq k \}$$

its closure in $\mathbb{R}^{m \times n}$. A natural alternative approach for obtaining low-rank (approximate) solutions to the matrix equation (1.2) is to minimize the quadratic function (1.1) on the set $M_{\leq k}$:

$$\min_{X \in M_{\leq k}} f_{A,B}(X). \quad (1.4)$$

Since the set $M_{\leq k}$ is closed, this problem admits at least one solution if $A$ is positive definite on the cone $M_{\leq 2k}$, that is, $(X, A[X])_F > 0$ for all $X \in M_{\leq 2k}$ (see our Proposition 2.5 later). In fact, in case that $B = A[X^*]$ for some $X^* \in \mathbb{R}^{m \times n}$, it holds

$$f_{A,B}(X) = \frac{1}{2} \| X - X^* \|_F^2 + f_{A,B}(X^*), \quad (1.5)$$

and thus, if, say, $A$ is positive semidefinite, the minimizers of $f_{A,B}$ on $M_{\leq k}$ are the best rank $k$ approximations of $X^*$ in energy (semi)norm.

The constrained optimization problem (1.4) is nonconvex and can be tackled by various methods. A particular approach is based on the bilinear representation $X = UV^T$ for low-rank matrices which allows to optimize the $m \times k$ and $n \times k$ matrices $U$ and $V$ using local search algorithms. This is also known as the Burer–Monteiro factorization in the more general context of low-rank approximations for SDPs [8] and can be very efficient when $k \ll mn$. It is advisable to break the non-uniqueness of the factorization $X = UV^T$ by adding penalty terms or applying alternating least squares; see, e.g., [40, 23, 41]. Another family of methods is based on exploiting that the set $M_k$ is a smooth Riemannian manifold. This allows again the use of local search algorithms but now using techniques from Riemannian optimization [2]. Since these methods directly optimize over $M_k$, they do not require regularization due to non-unique representations. In addition, careful implementations have similar cost per iteration as algorithms based on bilinear factorizations; see, e.g., [38, 27, 37].

There has been considerable interest in the case of positive semidefinite operators $A$. This concerns, for instance, the non-convex formulation of low-rank matrix completion problems [9]

$$\min_{X \in M_k} \frac{1}{2} \langle X, P_\Omega[X] \rangle_F - \langle X, P_\Omega[B] \rangle_F,$$

which corresponds to solutions of

$$P_\Omega[X] = P_\Omega[B].$$

Here, $A = P_\Omega$ is the orthogonal projection on a set of known entries $\Omega$, which means that $A$ is not invertible. Solvers for these problems with the best theoretical guarantees are based on convex relaxation [9, 10], but they can also be treated well by low optimization techniques [10, 37].
which are much less costly. Other problems of this semidefinite type include matrix sensing \[24\]. Here we are faced with the problem

\[
\min \frac{1}{2}\|B[X] - b\|_F^2
\]

(1.6)

where \(B\) is a linear operator from \(\mathbb{R}^{m \times n}\) to \(\mathbb{R}^d\). This problem fits in our symmetric and semidefinite framework using \(A = B^T B\) and \(B = B^T[b]\).

### 1.2 Contributions and existing results

In this paper we focus on problems of the form (1.1) where \(A\) is positive semidefinite. In many applications, it can be observed in numerical experiments that if the true solution \(X^*\) of the corresponding matrix equation has exactly low rank, that is, the global minima \(X^*\) of the function (1.1) are on \(\mathcal{M}_k\), and if this rank \(k\) is known, then local optimization methods for (1.4) do typically recover those global optima \(X^*\). This is somewhat surprising since the problem is non-convex. In addition, when \(X^*\) is only close to but not in \(\mathcal{M}_k\), such algorithms typically return different local minima on \(\mathcal{M}_k\) that are, however, all close to \(X^*\).

As explained in many works, the reason for this fortunate behavior seems to be the relatively benign optimization landscape: given an objective function \(f\) that is sufficiently well-conditioned and convex when restricted to cones of \(\mathcal{M}_{\leq k}\), the local minima always seem to be global minima after restricting \(f\) to \(\mathcal{M}_k\). In other words, other critical points are either saddle points or local maxima, and are hence unlikely to attract sequences generated by minimization algorithms that impose monotonic reduction of the objective. In addition, the saddle points have directions of sufficient large negative curvature (strict saddle property \([12]\)), so that algorithms can escape them sufficiently fast. Such remarkable properties have been rigorously proven under suitable assumptions for different low-rank optimization problems like matrix sensing \([4, 23]\), matrix completion \([34, 14]\), general convex functions on \(\mathcal{M}_{\leq k}\) \([18, 41]\), SDPs \([6, 7]\), and also for some other problems in the context of compressed sensing like phase retrieval \([33]\) and sparse dictionary recovery \([30, 31, 32, 33]\). See also \([13]\) for an overview.

Our aim is to provide here similar results that also include general semidefinite operators that are close to identity when restricted to the cone \(\mathcal{M}_{\leq 2k}\). To this end, we study the critical points of (1.4). In particular, we will show that when \(A\) is a sufficiently small perturbation of identity when restricted to the cone \(\mathcal{M}_{\leq 2k}\), and if a solution of the matrix equation (1.2) lies on \(\mathcal{M}_{\leq k}\), then there are no local minima of rank \(k\) except the global one. Additionally, bounds on the negative eigenvalues of the Riemannian Hessian at other critical points are given. This is important for escaping such critical points in local search methods. These results are in Theorem 3.5 and Corollary 3.6, which, to our knowledge, provide improved and simple conditions on the restricted isometry constants \(\delta_k\) when applied to matrix sensing as compared to those we could find in the literature. For example, we obtain that \(\delta_{4k} \leq 0.3446\) or \(\delta_{2k} \leq 0.2807\) are each sufficient for absence of local minima in the noiseless matrix sensing with non-symmetric matrices. This can be compared to \([29]\) with \(\delta_{4k} \leq 0.0363\) and \([41]\) with \(\delta_{4k} \leq 0.2\), and \([4]\) with \(\delta_{2k} < 0.2\) for symmetric matrices (observe that \(\delta_k \leq \delta_{2k}\) for all integers \(k \leq \ell\)).

The most general version of our analysis is Theorem 3.9 which also deals with the inexact (or noisy) case, where the matrix equation (1.2) admits only an approximate solution on \(\mathcal{M}_{\leq k}\) of some accuracy \(\varepsilon\). In this case, all critical points whose Riemannian Hessians have small or no negative eigenvalues (e.g. local minima) are optimal up to constant. While the statements are in principal easy to use, it might be difficult to get an intuition on the actual values. We therefore provide some concrete examples on the interplay of spectral bounds, negative eigenvalues of the Riemannian Hessian, and \(\varepsilon\) in Section 3.4.
Our strategy to obtain our results is motivated by an interesting observation on the Euclidean distance function $f(X) = \|X - B\|_2^2$, namely that for $B \in M_k$ it has no critical points at all except for the global minimum $X = B$. In order to generalize this rather peculiar behavior of the operator $A = I$ to more general ones, we introduce a certain norm in which we measure the distance of $X - [A[X] - B]$ from the cone $M_{\leq k}$. By comparing upper and lower bounds for this distance, we obtain our results.

Currently, the main results in this paper do not cover the important cases of matrix completion or matrix equations with badly conditioned operators (as they arise in numerical linear algebra), but some of the observations obtained alongside still provide general insight into the problem.

Finally, we hope to make a contribution to the subject by taking a geometric viewpoint on the problem that focuses on the critical points of the constrained problem (1.4), regardless of the method or parametrization used for representing the low-rank matrices. This is in contrast to [23, 41, 18] where an explicit regularization has to be used to cope with the non-uniqueness of the $X = UV^T$ factorization. Also, thanks to the manifold setup, we believe our analysis has potential implications to most local search methods for (1.4).

2 Properties of critical points

In this work, we study the critical points of the function $f_{A,B}$ defined in (1.1) on the smooth manifold $M_k$ of fixed rank $k$ matrices only. We shortly justify this restriction to the smooth part of $M_{\leq k}$ at the end of this section. In general, we neither assume that $B$ is in the range of $A$, nor that $A$ is positive semidefinite. Instead, a so-called restricted positive definiteness on the cones $M_{\leq k}$ will play the crucial role for the main results on absence of local minima of $f_{A,B}$ on $M_k$ in Section 3.

2.1 Tangent space and critical points

A point $X \in M_k$ is called a critical point of $f_{A,X}$ on $M_k$, if $\nabla f_{A,X}(X) = A[X] - B$ is orthogonal to the tangent space $T_X M_k$ at $X$. This tangent space is known to be (see, e.g., [16, Proposition 4.1])

$$T_X M_k = \{CX + XD: C \in \mathbb{R}^{m \times m}, D \in \mathbb{R}^{n \times n}\}. \quad (2.1)$$

Note that $X \in M_k$, $Y \in T_X M_k \Rightarrow Y \in M_{\leq 2k}$.

Let $P^\text{col}_X$ and $P^\text{row}_X$ denote the respective orthogonal projections on the column and row space of a matrix $X$. From (2.1), we see that a matrix $Z$ is orthogonal to $T_X M_k$ if $P^\text{col}_X Z = 0$ and $Z P^\text{row}_X = 0$, or, in other words,

$$Z = (I - P^\text{col}_X)Z(I - P^\text{row}_X).$$

Hence, with $X = U \Sigma V^T$ and $Z = \tilde{U} \tilde{\Sigma} \tilde{V}^T$ two SVDs, we obtain that

$$X + Z = \begin{bmatrix} U & \tilde{U} \end{bmatrix} \begin{bmatrix} \Sigma & 0 \\ 0 & \tilde{\Sigma} \end{bmatrix} \begin{bmatrix} V \\ \tilde{V} \end{bmatrix}^T \quad (2.2)$$

is also an SVD. A main consequence of this is that

$$\text{rank}(X + Z) = \text{rank}(X) + \text{rank}(Z) \quad \text{for } Z \text{ orthogonal to } T_X M_k. \quad (2.3)$$

The seemingly simple observation (2.2) turns out to be quite useful and is the main argument of Lemma 2.3 below. In fact, its immediate consequence (2.3) already has some surprising
implication on the critical points of the Euclidean distance function, that is, when $\mathcal{A}$ is the identity operator.

**Proposition 2.1.** Let $\mathcal{A} = \mathcal{I}$ be the identity. Then for any $B \in \mathcal{M}_k$, the function $f(X) = \frac{1}{2}\|X - B\|^2_F$ has only one critical point on $\mathcal{M}_k$, namely the global minimizer $X^* = B$.

**Proof.** Since $\nabla f(X) = X - B$, the condition for a critical point is that $X - B$ is orthogonal to $T_X\mathcal{M}_k$. Then, by (2.3),

$$k = \text{rank}(X - (X - B)) = \text{rank}(X) + \text{rank}(X - B) = k + \text{rank}(X - B),$$

which implies $X = B$. 

The following equivalent statement is somewhat even more interesting from a geometric point of view. It follows directly from $X - Y$ being the gradient of the function $f(X) = \frac{1}{2}\|X - Y\|^2_F$.

**Proposition 2.2.** Let $X$ and $Y$ be two distinct points on $\mathcal{M}_k$. Then $X - Y$ is not orthogonal to $T_X\mathcal{M}_k$.

Our aim in this paper is to study how far the observation in Proposition 2.1 for the identity operator carries over to the function $f_{\mathcal{A},B}$ when $\mathcal{A}$ is a perturbation of the identity, at least in a restricted sense. For this we will have to quantify the “rank increase” property (2.3). The starting point will be the inequality stated in Lemma 2.3 below, which first require some definitions. By $\sigma_1(Z) \geq \sigma_2(Z) \geq \ldots$ we denote the singular values of a matrix $Z$, with the agreement $\sigma_i(Z) = 0$ for $i \geq \min(m,n)$. We then consider the norm

$$\|Z\|_{\sigma,k} := \sqrt{\sigma_1^2(Z) + \cdots + \sigma_k^2(Z)} = \max_{\|Y\|_F = 1} \langle Y, Z \rangle_F. \quad (2.4)$$

Here the equality of both expressions is a consequence of the fact that truncated singular value decomposition yields best approximations in Frobenius norm on the cone $\mathcal{M}_{\leq k}$, and hence maximizes the orthogonal projection on it. The norm properties of $\|Z\|_{\sigma,k}$ then follow easily from the expression on the right hand side of (2.4). Note that $\|X\|_{\sigma,k} \leq \|X\|_F$ for every matrix $X$.

The norm (2.4) is a unitarily invariant norm, that is, $\|UZV\|_{\sigma,k} = \|Z\|_{\sigma,k}$ for all orthogonal $U$ and $V$. Hence, the truncated singular value decomposition also provides best rank-$k$ approximations in this norm [17, Section 7.4.9]. Therefore, for any fixed $Z$ it holds

$$\text{dist}_{\|\cdot\|_{\sigma,k}}(Z, \mathcal{M}_{\leq k}) := \min_{Y \in \mathcal{M}_{\leq k}} \|Z - Y\|_{\sigma,k} = \sqrt{\sigma_{k+1}^2(Z) + \cdots + \sigma_k^2(Z)}. \quad (2.5)$$

For the case of the identity operator $\mathcal{A} = \mathcal{I}$, we have obtained a contradiction to the existence of two critical points $X \neq X^* = B$ on $\mathcal{M}_k$ from two facts: on the one hand, the matrix $X - (\mathcal{A}[X] - B) = X - \nabla f_{\mathcal{A},B}(X)$ should have a higher rank than $X$, that is a positive distance to $\mathcal{M}_{\leq k}$, while on the other hand it cannot, since it equals $X^*$. For $\mathcal{A}$ close to $\mathcal{I}$ we expect a similar contradiction, but to obtain it, we need both upper and lower bounds for the distance of $X - (\mathcal{A}[X] - B)$ from $\mathcal{M}_{\leq k}$. Our key idea is to obtain such bounds for the distance in $\|\cdot\|_{\sigma,k}$-norm.

**Lemma 2.3.** Let $X$ be a critical point of $f_{\mathcal{A},B}$ on $\mathcal{M}_k$ and

$$\alpha = \text{dist}_{\|\cdot\|_{\sigma,k}}(X - (\mathcal{A}[X] - B), \mathcal{M}_{\leq k}).$$

(i) For any $Y \in \mathcal{M}_k$ it holds

$$\alpha \leq \|([\mathcal{I} - \mathcal{A}][X - Y])\|_{\sigma,k} + \|B - \mathcal{A}[Y]\|_{\sigma,k}.$$
(ii) Let $0 \leq j \leq k$ be the largest number such that $\sigma_i(\mathcal{A}[X] - B) > \sigma_{k-i+1}(X)$ for $1 \leq i \leq j$. Then

$$\alpha^2 \geq \sum_{i=1}^{j} \sigma_{k-i+1}^2(X) + \sum_{i=j+1}^{k} \sigma_i^2(\mathcal{A}[X] - B).$$

**Proof.** Item (i) is immediate from the definition of $\alpha$ and the triangle inequality:

$$\alpha \leq \|X - (\mathcal{A}[X] - B) - Y\|_{\sigma,k} \leq \|X - \mathcal{A}[X] + \mathcal{A}[Y] - Y\|_{\sigma,k} + \|B - \mathcal{A}[Y]\|_{\sigma,k}.$$ 

To show (ii) we use the characterization

$$\alpha = \sqrt{\sigma_{k+1}^2(X - \mathcal{A}[X] + B) + \cdots + \sigma_{2k}^2(X - \mathcal{A}[X] + B)},$$

which holds by \cite{2.5}. Let $\varsigma_i = \sigma_i(X)$ and $s_i = \sigma_i(\mathcal{A}[X] - B)$ for abbreviation. By \cite{2.2} (with $Z = -\mathcal{A}[X] + B$), the largest $2k$ singular values of the matrix $X - \mathcal{A}[X] + B$ are among the $3k$ numbers $\varsigma_1, \ldots, \varsigma_k, s_1, \ldots, s_{2k}$. By definition of $j$, the largest $k$ of these numbers are $\varsigma_1, \ldots, \varsigma_{j-1}, s_1, \ldots, s_j$ (the notation is slightly abusive when $j = k$), and hence $\alpha^2$ is the sum of squares of the largest $k$ remaining ones. In particular, $\alpha^2$ is larger or equal to any sum of squares of $k$ of the remaining singular values, which implies the asserted lower bound.

In agreement to what was pointed out above, the inequalities (i) and (ii) in Lemma 2.3 are contradictory in the case $\mathcal{A} = \mathcal{I}$ and $Y = B \in \mathcal{M}_k$ unless $X = B$. Our strategy is to show that they remain contradictory when $\mathcal{A}$ acts like a perturbation of identity on low-rank matrices. However, different from the case when $\mathcal{A}$ equals the identity, we will have to confine ourselves to local minima on $\mathcal{M}_k$, or at least critical points with almost positive semidefinite Riemannian Hessian (see sec. 2.4), in order to deal with the a-priori unknown singular values of $X$ in the lower bound for $\alpha$.

### 2.2 Restriction to the smooth part $\mathcal{M}_k$

We justify why we are ignoring potentially local minima of $f_{\mathcal{A},B}$ on $\mathcal{M}_{\leq k}$ of rank less than $k$. By the textbook definition (e.g., \cite{25} Theorem 6.12), $X \in \mathcal{M}_{\leq k}$ is a critical point of the nonsmooth problem \cite{14}, if $-\nabla f_{\mathcal{A},B}(X) = -\mathcal{A}[X] - B$ is in the polar of the Bouligand tangent cone of $X$. If $\text{rank}(X) = s \leq k$, the Bouligand tangent cone can be shown to be \cite{15} Example 20.5] (cf. \cite{11, 26})

$$T_X^0 \mathcal{M}_{\leq k} = T_X \mathcal{M}_s + \{Z \in \mathbb{R}^{m \times n} : \text{rank}(Z) \leq k - s\}.$$ 

As it turns out, when $s < k$, the polar cone $(T_X^0 \mathcal{M}_{\leq k})^\circ$ is just the point $\{0\}$, and hence a critical point satisfies $\nabla f_{\mathcal{A},B}(X) = 0$, that is, solves the equation $\mathcal{A}[X] = B$. If we just assume that $\mathcal{A}$ is positive semidefinite, then $X$ is an unconstrained global minimizer of $f_{\mathcal{A},B}$. If we, on the other hand, even assume that the equation $\mathcal{A}[X] = B$ does not have solutions of rank less than $k$ at all, such critical points cannot exist and all critical points of $f_{\mathcal{A},B}$ on $\mathcal{M}_{\leq k}$ in this broader sense in fact lie in $\mathcal{M}_k$. This is for instance the case if $\mathcal{A}$ is positive definite and $\text{rank}(X^*) \geq k$ \cite{26}. For clarity we summarize these considerations.

**Proposition 2.4.** Let $X \in \mathcal{M}_{\leq k}$ be a critical point of $f_{\mathcal{A},B}$ on $\mathcal{M}_{\leq k}$ in the sense $-\nabla f_{\mathcal{A},B}(X) \in (T_X^0 \mathcal{M}_{\leq k})^\circ$. Then either $\text{rank}(X) = k$ and $X$ is a critical point of $f_{\mathcal{A},B}$ on $\mathcal{M}_k$, or $\mathcal{A}[X] = B$. In the latter case, if $\mathcal{A}$ is a positive semidefinite operator, then $X$ is a global minimizer of $f_{\mathcal{A},B}$ on $\mathbb{R}^{m \times n}$.

Based on these facts, all subsequent theorems will be formulated for critical points of the smooth manifold $\mathcal{M}_k$ only. A key challenge, however, is to bound distance of critical points $X \in \mathcal{M}_k$ to $\mathcal{M}_{\leq k-1}$, that is, the smallest singular value $\sigma_k(X)$, from below.
2.3 Restricted spectral bounds

The central tool to analyze the local minima of $f_{A,B}$ on $\mathcal{M}_k$ are the “restricted spectral bounds”, that is, the minima and maxima of the Rayleigh quotient of the symmetric operator $A$ on cones of low-rank matrices. We use the following definitions:

$$\lambda(A,k) = \min_{X \in \mathcal{M}_{\leq k}} \langle X, A[X] \rangle_F, \quad \text{for } \|X\|_F = 1$$

and

$$\Lambda(A,k) = \max_{X \in \mathcal{M}_{\leq k}} \langle X, A[X] \rangle_F.$$

Note that both the minimum and maximum are attained since $\mathcal{M}_{\leq k}$ is closed. Obviously, whenever $k' \geq k$ it holds

$$\lambda(A,k') \leq \lambda(A,k) \leq \Lambda(A,k) \leq \Lambda(A,k').$$

In particular,

$$\lambda(A,k) \leq \Lambda(A,\ell)$$

for all combinations of $k$ and $\ell$.

If $A \neq 0$ is positive semidefinite, it holds $\lambda(A,1) > 0$, since the space $\mathbb{R}^{m \times n}$ possesses an orthonormal basis of rank-one matrices. Furthermore, one can then show that

$$\Lambda(A,k + \ell) \leq \Lambda(A,k) + \Lambda(A,\ell).$$

For this inequality to hold it is sufficient that $\lambda(A,k + \ell) \geq 0$.

We note that the lower spectral bounds provide conditions for the existence of minimizers as follows.

**Proposition 2.5.** Assume $\lambda(A,2k) > 0$. Then the function $f_{A,B}$ has at least one minimizer on $\mathcal{M}_{\leq k}$, that is, the problem (1.3) admits at least one solution.

**Proof.** Fix $Y \in \mathcal{M}_{\leq k}$. Since $\lambda(A,2k) > 0$, the representation

$$f_{A,B}(X) = f_{A,B}(Y) + \langle A[Y] - B, X - Y \rangle_F + \frac{1}{2} \langle X - Y, A[X - Y] \rangle_F$$

easily shows that $f$ is coercive on $\mathcal{M}_{\leq k}$, that is, $f_{A,B}(X) \to \infty$ for $\|X\|_F \to \infty$ on $\mathcal{M}_{\leq k}$. It means that the restriction of $f_{A,B}$ to $\mathcal{M}_{\leq k}$ has bounded sublevel sets, and so the existence of a minimizer follows from the fact that $\mathcal{M}_{\leq k}$ is closed. \qed

We will also need upper estimates for mixed products $\langle Y, A[Z] \rangle_F$ in terms of the restricted spectral bounds. They can be derived using the “parallelogram identity”, similar to [13] Lemma 3.3.

**Lemma 2.6.** Let $\lambda_i \leq \lambda(A,i) \leq \Lambda_i$ for all $i$. Then for any $Y \in \mathcal{M}_{\leq k}$ and $Z \in \mathcal{M}_{\leq \ell}$ it holds

$$\langle Y, A[Z] \rangle_F \leq \frac{1}{4} (\Lambda_{k+\ell} - \lambda_{k+\ell}) (\|Y\|_F^2 + \|Z\|_F^2) + \frac{1}{2} (\Lambda_{k+\ell} + \lambda_{k+\ell}) \langle Y, Z \rangle_F.$$

\footnote{Using SVD, every matrix $Z$ of rank at most $k + \ell$ can be written as $Z = sX + tY$, where $s,t \in \mathbb{R}$ and $X$ and $Y$ are of rank at most $k$ and $\ell$, respectively, and orthonormal with respect to the Frobenius inner product. Consider then the $2 \times 2$ symmetric matrix $G = \begin{pmatrix} (X, A[X])_F & (Y, A[Y])_F \\ (X, A[Y])_F & (Y, A[Y])_F \end{pmatrix}$. With $a = [s, t]^T$, it follows $\langle Z, A[Z] \rangle_F = a^T G a$. Hence, the matrix $G$ is positive semidefinite since $\lambda(A,k + \ell) \geq 0$. From $a^T G a \leq \text{trace}(G) = ((X, A[X])_F + (Y, A[Y])_F) \|a\|_2^2 \leq (\Lambda_k + \Lambda_\ell) \|Z\|_F^2$, one obtains the result.}
Proof. Since \( \mathcal{A} \) is symmetric, it holds
\[
4 \langle Y, \mathcal{A}[Z] \rangle_F = \langle Y + Z, \mathcal{A}[Y + Z] \rangle_F - \langle Y - Z, \mathcal{A}[Y - Z] \rangle_F \\
\leq \Lambda_{k+\ell} \|Y + Z\|_F^2 - \lambda_{k+\ell} \|Y - Z\|_F^2,
\]
which easily yields the asserted bound. \( \square \)

The upper bound (2.8) will be required for the shifted operator \( \mathcal{I} - \mathcal{A} \). Under the assumptions of the lemma, it follows from the definitions that
\[
\Lambda(\mathcal{I} - \mathcal{A}, k) = 1 - \lambda(\mathcal{A}, k) \leq 1 - \lambda_k
\]
and
\[
\lambda(\mathcal{I} - \mathcal{A}, k) = 1 - \Lambda(\mathcal{A}, k) \geq 1 - \Lambda_k
\]
for all \( k \) and \( \mu \in \mathbb{R} \). Therefore, by applying (2.8),
\[
\langle Y, (\mathcal{I} - \mathcal{A})[Z] \rangle_F \leq \frac{1}{4} (\Lambda_{k+\ell} - \lambda_{k+\ell}) (\|Y\|_F^2 + \|Z\|_F^2) + \frac{1}{2} (2 - \Lambda_{k+\ell} - \lambda_{k+\ell}) \langle Y, Z \rangle_F. \tag{2.9}
\]

Let us introduce the constants
\[
\Gamma(\mathcal{A}, k, \ell) = \max_{Y \in \mathcal{M}_{\leq k}, \, Z \in \mathcal{M}_{\leq \ell}, \|Y\|_F = 1, \|Z\|_F = 1} \langle Y, \mathcal{A}[Z] \rangle_F.
\]
They can be related to the \( \| \cdot \|_{\sigma,k} \)-norms introduced in (2.4) in the following way.

**Lemma 2.7.** Let \( Z \) have rank \( \ell \), then \( \|\mathcal{A}[Z]\|_{\sigma,k} \leq \Gamma(\mathcal{A}, k, \ell) \|Z\|_F \).

The proof is immediate from the right hand side of (2.4).

Finally, the scaling behavior of \( \Gamma(\mathcal{A}, k, \ell) \) with respect to the ranks \( k \) and \( \ell \) will turn out to be useful later to relate our results to existing ones.

**Lemma 2.8.** For the positive integers \( p, q \), it holds.
\[
\Gamma(\mathcal{A}, pk, q\ell) \leq \sqrt{pq} \Gamma(\mathcal{A}, k, \ell).
\]

**Proof.** Let \( Y \) and \( Z \) be the maximizers in \( \Gamma(\mathcal{A}, pk, q\ell) \). Using SVD, we can write \( Y = a_1 Y_1 + \cdots + a_p Y_p \), where the matrices \( Y_1, \ldots, Y_p \in \mathcal{M}_{\leq k} \) are pairwise orthogonal and have Frobenius norm one, and the scalars \( a_1, \ldots, a_p \) are not negative. Observe that \( a_1^2 + \cdots + a_p^2 = 1 \). We decompose \( Z = b_1 Z_1 + \cdots + b_q Z_q \) similarly. Hence,
\[
\Gamma(\mathcal{A}, pk, q\ell) = \sum_{i=1}^{p} \sum_{j=1}^{q} a_i b_j \langle Y_i, \mathcal{A}[Z_j] \rangle_F \leq \left( \sum_{i=1}^{p} a_i \right) \left( \sum_{j=1}^{q} b_j \right) \Gamma(\mathcal{A}, k, \ell)
\]
and the result follows from the Cauchy–Schwarz inequality. \( \square \)

### 2.4 Estimates related to the Riemannian Hessian

We provide an abstract bound on the smallest singular values of a critical point \( X \in \mathcal{M}_k \) of \( f_{\mathcal{A}, \mathcal{B}} \) on \( \mathcal{M}_k \) in terms of restricted spectral properties of \( \mathcal{A} \), the singular values of the residual \( \mathcal{A}[X] - \mathcal{B} \), and spectral bounds for the Riemannian Hessian of \( f_{\mathcal{A}, \mathcal{B}} \) at \( X \) (we refer to [2 Ch. 5] for information on the concept of Riemannian Hessian).
Denote by $\mathcal{H}_X$ the Riemannian Hessian of $f_{A,B}$ (restricted to $\mathcal{M}_k$) at $X \in \mathcal{M}_k$. As metric on the submanifold $\mathcal{M}_k$, we choose the restriction of the Frobenius inner product from the ambient space $\mathbb{R}^{m \times n}$. Let $X = U \Sigma V^T$ be an SVD of $X$ with $\Sigma = \text{diag}(\varsigma_1, \ldots, \varsigma_k)$, and set

$$
\tilde{G} = U \Sigma^{1/2}, \quad \tilde{H} = V \Sigma^{1/2}.
$$

Every tangent vector $Z \in T_X \mathcal{M}_k$ can be written as

$$
Z = \partial G \cdot \tilde{H}^T + \tilde{G} \cdot \partial H^T.
$$

With this representation of tangent vectors, we prove in Appendix A that

$$
\mathcal{H}_X[Z, Z] = (Z, A[Z])_F + 2(\partial G \partial H^T, (I - P_X^{\text{col}})(A[X] - B)(I - P_X^{\text{row}}))_F, \quad (2.10)
$$

with $P_X^{\text{col}}$ and $P_X^{\text{row}}$ being the orthogonal projections onto the column and row space of $X$, respectively.

When $X$ is a critical point of $f_{A,B}$ on $\mathcal{M}_k$, then $(I - P_X^{\text{col}})(A[X] - B)(I - P_X^{\text{row}}) = A[X] - B$ (see Section 2.1), and so the Riemannian Hessian at such points reads

$$
\mathcal{H}_X[Z, Z] = (Z, A[Z])_F + 2(\partial G \partial H^T, A[X] - B)_F. \quad (2.11)
$$

If $X$ is a local minimum, the Riemannian Hessian is positive semidefinite, that is, $\mathcal{H}_X[Z, Z] \geq 0$ for all $Z \in T_X \mathcal{M}_k$. In the following proposition, we consider arbitrary critical points $X$ for which the Riemannian Hessian satisfies a lower spectral bound.

**Proposition 2.9.** Let $X$ be a critical point of $f_{A,B}$ on $\mathcal{M}_k$ and $\varsigma_1 \geq \cdots \geq \varsigma_k \geq 0$ denote its singular values. Further, let $s_1 \geq \cdots \geq s_k \geq 0$ denote the $k$ largest singular values (some might be zero) of $A[X] - B$. Assume for some $\mu \geq 0$ that the Riemannian Hessian satisfies

$$
\mathcal{H}_X[Z, Z] \geq -\mu \|Z\|_F^2 \quad \text{for all } Z \in T_X \mathcal{M}_k.
$$

Then for any $j = 1, \ldots, k$ and $\Lambda_{2j} > 0$ with $\lambda(A, 2j) \leq \Lambda_{2j}$ it holds

$$
\sqrt{\varsigma_1^2 + \cdots + \varsigma_{k-j+1}^2} \geq \frac{\sqrt{s_1^2 + \cdots + s_j^2}}{\Lambda_{2j} + \mu} = \frac{\|A[X] - B\|_{\sigma,j}}{\Lambda_{2j} + \mu}.
$$

**Proof.** Let $g_1, \ldots, g_j$ be the normalized dominant $j$ left singular vectors of $A[X] - B$, and $h_1, \ldots, h_j$ the dominant right singular vectors (or some of them if there are equal singular values). We consider

$$
\partial G = \begin{bmatrix} 0 & \cdots & 0 & s_1^{1/2} g_1 & \cdots & s_j^{1/2} g_j \end{bmatrix},
$$

and

$$
\partial H = \begin{bmatrix} 0 & \cdots & 0 & -s_1^{1/2} h_1 & \cdots & -s_j^{1/2} h_j \end{bmatrix}.
$$

Then

$$
(\partial G \partial H^T, A[X] - B)_F = -(s_1^2 + \cdots + s_j^2).
$$

We note that since $X$ is a critical point, that is, $A[X] - B$ is orthogonal to $T_X \mathcal{M}_k$, each vector $g_i$ is orthogonal to the columns $u_1, \ldots, u_k$ of $U$, and each $h_i$ is orthogonal to the columns $v_1, \ldots, v_k$ of $V$ (see Section 2.1), or $s_i = 0$. Therefore, the tangent vector

$$
Z = \partial G \cdot \tilde{H}^T + \tilde{G} \cdot \partial H^T = \sum_{i=1}^j s_i^{1/2} \varsigma_{k-j+i}^2 g_i v_{k-j+i}^T - s_i^{1/2} \varsigma_{k-j+i}^2 u_{k-j+i} h_i^T
$$

This is arguably the most simple metric one can take. It is used in the Riemannian algorithms [25, 27, 37, 39] for low-rank optimization. Other metrics also exist in [19, 21, 20] but they do not always lead to improved bounds in the current context.
is a sum of $2j$ rank-one (or zero) matrices that are pairwise orthogonal with respect to the Frobenius inner product. Thus,

$$\|Z\|_F^2 = 2(s_1 \xi_{k-j+1} + \cdots + s_j \xi_k).$$

In light of (2.11), we obtain

$$H_X[Z, Z] \leq 2\Lambda_2(s_1 \xi_{k-j+1} + \cdots + s_j \xi_k) - 2(s_1^2 + \cdots + s_j^2),$$

which by assumption implies

$$-2(s_1^2 + \cdots + s_j^2) \geq -2\Lambda_2(s_1 \xi_{k-j+1} + \cdots + s_j \xi_k) - 2\mu(s_1 \xi_{k-j+1} + \cdots + s_j \xi_k).$$

Applying the Cauchy–Schwarz inequality we get

$$s_1^2 + \cdots + s_j^2 \leq (\Lambda_2 + \mu) \frac{\sqrt{s_k^2 + \cdots + s_{j+1}^2} \sqrt{s_1^2 + \cdots + s_j^2}}{s_j},$$

which is equivalent to the asserted inequality.

We present two corollaries of the previous estimate that will not be used later, but are of independent interest. They concern the positive semidefinite case.

**Corollary 2.11.** Let $B = A[X^*]$ and $X$ be a critical point of $f_{A, B}$ on $M_k$ at which the Riemannian Hessian $H_X$ is positive semidefinite. Assume $\text{rank}(A[X] - B) = \text{rank}(A[X - X^*]) \leq \ell$ and $\lambda(A, \ell) > 0$. Then the $k$-th singular value of $X$ satisfies the inequality

$$\sigma_k(X) \geq \frac{1}{\sqrt{\ell}} \cdot \frac{\lambda(A, \ell)}{\Lambda(A, 2)} \cdot \|X - X^*\|_F.$$

**Proof.** Let $s_1 \geq \cdots \geq s_\ell$ denote the $\ell$ largest singular values (some might be zero) of $A[X] - B = A[X - X^*]$. Then

$$s_1 \geq \sqrt{\frac{s_1^2 + \cdots + s_\ell^2}{\ell}} = \frac{\|A[X - X^*]\|_{\sigma, \ell}}{\sqrt{\ell}}.$$

By (2.4), since $X - X^* \in M_{\leq \ell}$, we then obtain

$$s_1 \geq \frac{(X - X^*, A[X - X^*])_F}{\sqrt{\ell} \|X - X^*\|_F} \geq \frac{\lambda(A, \ell)}{\sqrt{\ell}} \|X - X^*\|_F.$$

The assertion now follows from the previous proposition with $\mu = 0$, $j = 1$, and $\Lambda_2 = \Lambda(A, 2)$, which is positive by assumption.

Since the $k$th singular value of a rank-$k$ matrix equals its distance to $M_{\leq k-1}$ in Frobenius norm, we can rephrase the previous corollary in the following way.

**Corollary 2.11.** Under the same assumptions as in Corollary 2.10 it holds

$$\text{dist}_F(X, M_{\leq k-1}) \geq \frac{1}{2} \left( \frac{1}{\sqrt{\ell}} \cdot \frac{\lambda(A, \ell)}{\Lambda(A, 2)} \right) \sigma_k(X^*).$$

**Proof.** If $\|X - X^*\|_F \leq \sigma_k(X^*)/2$, then

$$\sigma_k(X) \leq \text{dist}_F(X^*, M_{\leq k-1}) \leq \text{dist}_F(X, M_{\leq k-1}) + \|X - X^*\|_F$$

implies $\text{dist}_F(X, M_{\leq k-1}) \geq \sigma_k(X^*)/2$, which is stronger than the asserted bound (recalling $\lambda(A, \ell) \leq \Lambda(A, 2)$). If $\|X - X^*\|_F > \sigma_k(X^*)/2$, the previous corollary provides the asserted bound since $\text{dist}_F(X, M_{\leq k-1}) = \sigma_k(X)$.

Note that in the case that $X^* \in M_k$, we can choose $\ell = 2k$ and the lower bounds in both corollaries become independent of the size of considered matrices.
3 RPD property and its implications for critical points

We now come to the main results of the paper on the critical points of the function \( f_{A,B} \) for the case that \( A \) almost acts as an identity operator on cones of low-rank matrices. This property is quantified by the restricted positive definiteness (RPD) constants below, which are equivalent to the RIP constants in matrix sensing. The most notable result then is on the case that \( B = A[X^*] \) for some \( X^* \in \mathcal{M}_k \). In this so-called “noiseless scenario”, and under the restricted positive definiteness assumptions, one can show that \( f_{A,B} \) has no local minima on \( \mathcal{M}_k \) except the single global minimum \( X^* \). Moreover, at all other critical points the Riemannian Hessian has sufficiently negative eigenvalues, which is important in optimization methods in order to “escape” such saddle points. The required bounds for the RPD constants for obtaining this conclusion are, to our knowledge, considerably weaker than the ones available in the literature. The results on the noiseless case are stated in Section 3.2. In Section 3.3, the most general version of our result is stated which deals with the case that the equation \( A[X] = B \) admits only an approximate solution \( X \in \mathcal{M}_{\leq k} \). Then local minima or saddle points with small negative curvature may exist, but there distance to \( X \) can be bounded.

3.1 RPD property

We still consider the family \((1.1)\) of quadratic functions \( f_{A,B} \), and make some assumptions on the restricted spectral bounds of \( A \) that quantify deviation from identity.

**Definition 3.1 (RPD property).** Let \( k \geq 1 \). We say that the symmetric operator \( A \) satisfies the \((k, \delta_k)\)-RPD (restricted positive definiteness) property, if there exist \( 0 < \delta_k < 1 \) such that

\[
1 - \delta_k \leq \lambda(A,k) \leq \Lambda(A,k) \leq 1 + \delta_k.
\]

with \( \lambda(A,k) \) and \( \Lambda(A,k) \) defined in (2.6) and (2.7).

**Remark 3.2.** In the context of the matrix sensing problem \((1.6)\) the restricted isometry property (RIP)

\[
(1 - \delta_k)\|X\|_F^2 \leq \|B[X]\|_F^2 \leq (1 + \delta_k)\|X\|_F^2
\]

for all \( X \in \mathcal{M}_{\leq k} \) has been introduced in [24] and used in many subsequent works. We have already mentioned that the matrix sensing problem fits our framework using the operator \( A = B^TB \). So the RPD above is equal to the RIP in this model. In fact, if we assume \( A \) to be positive semidefinite, then we can always find a decomposition \( A = B^TB \) where \( B : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^d \) with \( d = \text{rank}(A) \). Therefore we can write

\[
f_{A,B} = \|B[X] - b\|_F^2 - \|b\|_F^2
\]

with \( b = B^T[B] \), and there is no essential difference between the matrix sensing problem and our setup of finding critical points of (then convex) quadratic functions \( f_{A,B} \) on \( \mathcal{M}_k \). Regarding the existence of operators \( A \) obeying suitable RPD bounds we can therefore rely on the well known results on operators \( B \) satisfying RIP conditions; see [23]. A typical result is, for example, that specifically scaled random Gaussian matrices give rise to \( B \) that satisfy a \((k, \delta_k)\)-RIP with high probability if \( d \geq \delta_k^{-2}k(m + n) \); see [10]. We refer to [1] for more references on the RIP.

**Remark 3.3.** We comment on operators whose restricted spectral bounds are not centered around one. We may say that the symmetric operator \( A \) is positive definite on the cone \( \mathcal{M}_{\leq k} \) if \( \lambda_{A,k} > 0 \). We then may define the restricted condition number as

\[
\kappa(A,k) = \Lambda(A,k) / \lambda(A,k),
\]
and consider the scaled operator $\omega_k A$ with

$$\omega_k = \frac{2}{\lambda(A, k) + \Lambda(A, k)}.$$  

This operator has the spectral bounds

$$\lambda(\omega_k A, k) = \omega_k \lambda(A, k) = 1 - \delta_k, \quad \Lambda(\omega_k A, k) = \omega_k \Lambda(A, k) = 1 + \delta_k,$$

where

$$\delta_k = \frac{\kappa(A, k) - 1}{\kappa(A, k) + 1}.$$  

Obviously, a matrix $X \in \mathcal{M}_k$ is a critical point (local minimum) of $f_{A,B}$ on $\mathcal{M}_k$ if and only if it is a critical point (local minimum) of $\omega f_{A,B} = f_{\omega A, \omega B}$ for any $\omega > 0$. Therefore, the results below on those operators that satisfy RPD conditions translate to more general operators $A$ if the scaled operator $\omega_k A$ satisfies the assumptions. Yet this will mean that $A$ must have a rather small restricted condition number $\kappa(A, 2k)$ or $\kappa(A, 3k)$. We will comment on this issue where appropriate.

We note that with the RPD bounds (3.1), the estimate (2.9) leads in a straightforward way into the following upper bound (see also [23, Proposition 2.1] for essentially the same result).

**Lemma 3.4.** Under the $(k, \delta_k)$-RPD conditions it holds

$$\Gamma(I - A, k, \ell) = \max_{\|Y\|_F = 1, \|Z\|_F = 1} \langle Y, (A - I)[Z] \rangle_F \leq \delta_k + \ell.$$  

### 3.2 Noiseless case

In the noiseless case, we assume that there exist $X^* \in \mathcal{M}_k$ such that $B = A[X^*]$. In other words, it is assumed that a desired low-rank solution $X^*$ to the matrix equation $A[X] = B$ can be found among the critical points of $f_{A,B}$ on $\mathcal{M}_k$ (provided we know $k$), which allows, for example, using Riemannian optimization methods.

For the case that $A$ is positive semidefinite, which is slightly less general than our assumptions below, it has been shown in [23, Remark 1] that for $\delta_{4k} \leq 0.0363$ (this given value arises from estimating the root of cubic polynomial, and $\delta_{4k}$ in fact can be slightly larger) there can be no spurious local minima of $f_{A,B}$ on $\mathcal{M}_k$ except for the global minimum $X^*$. Our result below improves on this bound by an order of magnitude and by putting conditions on $\delta_{2k}$ and $\delta_{3k}$ instead of $\delta_{4k}$.

**Theorem 3.5.** Let $X^* \in \mathcal{M}_k$ such that $B = A[X^*]$ and $\mu \geq 0$. Assume $A$ satisfies RPD properties such that

$$\delta_{3k} < -\left(\frac{1 + \sqrt{2}}{2\sqrt{2}} + \frac{\mu}{2}\right) + \sqrt{\left(\frac{1 + \sqrt{2}}{2\sqrt{2}} + \frac{\mu}{2}\right)^2 + \frac{1}{\sqrt{2}}}.$$  

Then on $\mathcal{M}_k$, $X^*$ is the unique solution of $A[X] = B$ and the unique global minimum of $f_{A,B}$. At all other critical points $X \neq X^*$ of $f_{A,B}$ on $\mathcal{M}_k$ the Riemannian Hessian satisfies

$$\mathcal{H}_X[Z, Z] < -\mu\|Z\|_F^2,$$

for some tangent vector $Z \in T_X \mathcal{M}_k$. 

12
Alternatively, the same statements hold in the case that

\[ \delta_{2k} < -\left(\frac{3}{4} + \frac{\mu}{2}\right) + \sqrt{\left(\frac{3}{4} + \frac{\mu}{2}\right)^2 + \frac{1}{2}}. \]

\textbf{Proof.} The uniqueness statements follow from the representation \[ 1.5 \] together with the RPD assumptions. The statement on the other critical points follows from considering the special case \( \varepsilon = 0 \) in the more general Theorem 3.9 proved below. \( \square \)

As an example, consider the value \( \mu = 1 \). Then under the conditions

\[ \delta_{3k} \leq 0.2399 \quad \text{or} \quad \delta_{2k} \leq 0.1861 \]

all critical points except the global minimum \( X^* \) the Riemannian Hessian has a negative eigenvalue smaller than \(-1\). More values on the relation between \( \mu \) and \( \delta \) are presented in the Tables 1 and 2 below.

We also highlight the case \( \mu = 0 \) separately as it implies the absence of local minima.

\textbf{Corollary 3.6.} Let \( X^* \in \mathcal{M}_k \) and \( \mathcal{A}[X^*] = B \). Assume \( \mathcal{A} \) satisfies RPD properties such that

\[ \delta_{3k} < -\frac{1}{2} \left(1 + \frac{1}{\sqrt{2}}\right) + \sqrt{\frac{1}{4} \left(1 + \frac{1}{\sqrt{2}}\right)^2 + \frac{1}{\sqrt{2}}} \approx 0.3446. \]

Then on \( \mathcal{M}_k \), \( X^* \) is the unique solution of \( \mathcal{A}[X] = B \) and the unique global minimum of \( f_{\mathcal{A},B} \). There exist no other local minima of \( f_{\mathcal{A},B} \) on \( \mathcal{M}_k \).

Alternatively, the same statements hold in the case that

\[ \delta_{2k} < -\frac{1}{2} \left(1 + \frac{1}{\sqrt{2}}\right) + \sqrt{\frac{1}{4} \left(1 + \frac{1}{\sqrt{2}}\right)^2 + \frac{1}{\sqrt{2}}} = \frac{\sqrt{17} - 3}{4} \approx 0.2807. \]

For reference, we also generalize Theorem 3.5 to operators whose restricted spectral bounds are not centered around one. The proof follows according to Remark 3.3 by considering the scaled operators \( \omega_{3k} \mathcal{A} \), respectively \( \omega_{2k} \mathcal{A} \).

\textbf{Corollary 3.7.} Let \( X^* \in \mathcal{M}_k \) and \( B = \mathcal{A}[X^*] \) and \( \mu \geq 0 \). Assume \( \lambda(\mathcal{A},3k) > 0 \) and that

\[ \frac{\kappa(\mathcal{A},3k) - 1}{\kappa(\mathcal{A},3k) + 1} < - \left(\frac{1 + \sqrt{2}}{2\sqrt{2}} + \frac{\mu}{2}\right) + \sqrt{\left(\frac{1 + \sqrt{2}}{2\sqrt{2}} + \frac{\mu}{2}\right)^2 + \frac{1}{\sqrt{2}}}, \]

where \( \kappa(\mathcal{A},3k) = \Lambda(\mathcal{A},3k)/\Lambda(\mathcal{A},3k) \). Then on \( \mathcal{M}_k \), \( X^* \) is the unique solution of \( \mathcal{A}[X] = B \) and the unique global minimum of \( f_{\mathcal{A},B} \). At all other critical points \( X \neq X^* \) of \( f_{\mathcal{A},B} \) on \( \mathcal{M}_k \) the Riemannian Hessian satisfies

\[ \mathcal{H}_X[Z,Z] < -\frac{1}{2} \left(\lambda(\mathcal{A},3k) + \Lambda(\mathcal{A},3k)\right) \mu \|Z\|_F^2 \]

for some tangent vector \( Z \in T_X \mathcal{M}_k \).

Alternatively, the same conclusions hold (with \( \lambda(\mathcal{A},3k), \Lambda(\mathcal{A},3k) \) replaced by \( \lambda(\mathcal{A},2k), \Lambda(\mathcal{A},2k) \)) in the case that \( \lambda(\mathcal{A},2k) > 0 \) and

\[ \frac{\kappa(\mathcal{A},2k) - 1}{\kappa(\mathcal{A},2k) + 1} < -\left(\frac{3}{4} + \frac{\mu}{2}\right) + \sqrt{\left(\frac{3}{4} + \frac{\mu}{2}\right)^2 + \frac{1}{2}}. \]

Speaking for instance of local minima, or \( \mu = 0 \), we get the following sufficient conditions:

\[ \kappa(\mathcal{A},3k) \leq \frac{2}{1 - 0.3446} - 1 \leq 2.0515, \quad \kappa(\mathcal{A},2k) \leq \frac{2}{1 - 0.2807} - 1 \leq 1.7804. \]
3.3 General case

In the general case, the exact solution $X^*$ of the matrix equation $A[X] = B$ may have rank larger than $k$ but still admit good low-rank approximations, that is, it will be close to $M_{\leq k}$. Or, the given matrix $B$ may only satisfy $A[X^*] \approx B$ approximately and may not even belong to the range of $A$. In the matrix sensing problem \((1.6)\) this may correspond to noisy observations $b$.

For linear matrix equations \((1.2)\), this corresponds to a perturbation of the right hand side.

In the main result of this paper we focus on points $X_\epsilon \in M_{\leq k}$ that satisfy
\[
\|A[X_\epsilon] - B\|_{\sigma, 2k} \leq \epsilon,
\]
and estimate the distance of certain other critical points on $M_k$ (including local minima) $X$ to $X_\epsilon$. After giving the proof we calculate some concrete values and make some comments how to interpret this result in the context of the strict saddle point property. Regarding critical points of $f_{A,B}$ on $M_{\leq k}$, we refer once more to Proposition \(2.4\).

We first present a lemma that allows to conveniently state our conditions on the RPD constants in the main result.

**Lemma 3.8.** Let $c, \mu > 0$. Then the restriction of the function
\[
\delta \mapsto K(c, \mu, \delta) = \left( \frac{1 - \delta}{1 + \delta + \mu} - c\delta \right)^{-1} = \frac{1 + \delta + \mu}{1 - (1 + c + \mu)\delta - c\delta^2}.
\]
to the positive axis possesses a single pole and is positive if and only if
\[
0 < \delta < -\left( \frac{1 + c}{2c} + \frac{\mu}{2} \right) + \sqrt{\left( \frac{1 + c}{2c} + \frac{\mu}{2} \right)^2 + \frac{1}{c}}.
\]
It holds $K(c, \delta) \to \infty$ when $\delta$ approaches the bound. The bound is monotonically decreasing with respect to $c$ and $\mu$.

**Proof.** The statement is equivalent to
\[
\delta^2 + \left(1 + \frac{1}{2} + \mu\right)\delta - \frac{1}{2} < 0
\]
under the restriction $\delta > 0$. This open parabola is negative at zero and therefore $\delta$ must lie between zero and the positive root, which is the asserted condition. \(\square\)

We state the main result. Note that we could replace the $\| \cdot \|_{\sigma, 2k}$ norm in the assumptions with the Frobenius norm, since $\|A[X_\epsilon] - B\|_F \leq \epsilon$ would be a stronger assumption.

**Theorem 3.9.** Let $A$, $B$ and $\epsilon > 0$ be given such that there exists $X_\epsilon \in M_{\leq k}$ satisfying
\[
\|A[X_\epsilon] - B\|_{\sigma, 2k} \leq \epsilon. \tag{3.2}
\]
Let $\mu \geq 0$. Consider a critical point $X \in M_k$ of $f_{A,B}$ on $M_k$ satisfying
\[
H_X[Z, Z] \geq -\mu \|Z\|_F^2 \quad \text{for all } Z \in T_X M_k.
\]
The following two statements hold.

(i) If $A$ satisfies RPD properties such that
\[
\delta_2 \leq \delta_{2k} \leq \delta_{3k} < -\left(1 + \frac{\sqrt{2}}{2\sqrt{2}} + \frac{\mu}{2} \right) + \sqrt{\left( \frac{1 + \sqrt{2}}{2\sqrt{2}} + \frac{\mu}{2} \right)^2 + \frac{1}{\sqrt{2}}} \tag{3.3}
\]
(the first two inequalities pose no restriction), then $X$ satisfies the estimate
\[
\|X - X_\epsilon\|_F \leq \left( \sqrt{2} + \frac{1}{1 + \delta_2 + \mu} \right) \left[ \frac{1 + \delta_2 + \mu}{1 - \delta_{2k} - \sqrt{2(1 + \mu)\delta_{3k} - \sqrt{2\delta_{2k}\delta_{3k}}}} \right] \cdot \epsilon. \tag{3.4}
\]
(ii) Alternatively, if
\[ \delta_2 \leq \delta_{2k} < -\left( \frac{3}{4} + \frac{\mu}{2} \right) + \sqrt{\left( \frac{3}{4} + \frac{\mu}{2} \right)^2 + \frac{1}{2}} \]  
(3.5)
(the first inequality is no restriction), then \( X \) satisfies the estimate
\[ \| X - X_\varepsilon \|_F \leq \left( \sqrt{2} + \frac{1}{1 + \delta_2 + \mu} \right) \left[ \frac{1 + \delta_2 + \mu}{1 - (3 + 2\mu)\delta_{2k} - 2\delta_2\delta_{2k}} \right] \varepsilon. \]

**Proof.** We assume \( X \neq X_\varepsilon \), otherwise there is nothing to show. We consider upper and lower bounds for
\[ \alpha = \text{dist}_{\| \cdot \|,k}(X - (A[X] - B), M_{\leq k}). \]

Obviously,
\[ \alpha \leq \| X - A[X] + B - X_\varepsilon \|_{\sigma,k}, \]
and by triangle inequality,
\[ \alpha \leq \| (I - A)[X - X_\varepsilon] \|_{\sigma,k} + \| A[X_\varepsilon] - B \|_{\sigma,k} \leq \| (I - A)[X - X_\varepsilon] \|_{\sigma,k} + \varepsilon. \]

Lemmata 2.7 and 2.8 give
\[ \| (I - A)[X - X_\varepsilon] \|_{\sigma,k} \leq \Gamma(I - A, k, 2k) \| X - X_\varepsilon \|_F \leq \sqrt{2} \Gamma(I - A, k, k) \| X - X_\varepsilon \|_F. \]

Applying Lemma 3.4 to either of these bounds then results in the two estimates
\[ \alpha \leq \delta_{3k} \| X - X_\varepsilon \|_F + \varepsilon \]  
(3.6)
and
\[ \alpha \leq \sqrt{2} \delta_{2k} \| X - X_\varepsilon \|_F + \varepsilon. \]  
(3.7)

We turn to the lower bound on \( \alpha \) provided by Lemma 2.3. Let \( \varsigma_1 \geq \cdots \geq \varsigma_k > 0 \) denote the singular values of \( X \), and \( s_1 \geq \cdots \geq s_{2k} \) the \( 2k \) largest singular values (some might be zero) of \( A[X] - B \). By the lemma, there exists some \( 0 \leq j \leq k \) for which
\[ \alpha^2 \geq \varsigma_k^2 + \cdots + \varsigma_{k-j+1}^2 + s_{j+1}^2 + \cdots + s_{2k}^2. \]  
(3.8)
(If \( j = 0 \) there are no \( \varsigma_i \), while if \( j = k \) there are no \( s_i \).) By Proposition 2.9 (with \( j = 1 \)),
\[ \varsigma_k \geq \frac{s_1}{1 + \delta_2 + \mu}. \]

Due to \( \varsigma_{k-j+1} \geq \cdots \geq \varsigma_k \) and \( s_1 \geq \cdots \geq s_j \), (3.8) then entails
\[ \alpha \geq \left( \frac{1}{1 + \delta_2 + \mu} \right) \sqrt{s_1^2 + \cdots + s_{2k}^2} \geq \frac{1}{\sqrt{2}} \left( \frac{1}{1 + \delta_2 + \mu} \right) \sqrt{s_1^2 + \cdots + s_{2k}^2}. \]  
(3.9)

Using (2.4) twice, we can estimate
\[ \sqrt{s_1^2 + \cdots + s_{2k}^2} \geq \left\langle \frac{X - X_\varepsilon}{\| X - X_\varepsilon \|_F}, A[X] - B \right\rangle_F = \left\langle \frac{X - X_\varepsilon}{\| X - X_\varepsilon \|_F}, A[X - X_\varepsilon] \right\rangle_F + \left\langle \frac{X - X_\varepsilon}{\| X - X_\varepsilon \|_F}, A[X_\varepsilon] - B \right\rangle_F \geq (1 - \delta_{2k}) \| X - X_\varepsilon \|_F - \varepsilon. \]

15
With (3.9) we arrive at the lower bound
\[
\alpha \geq \frac{1}{\sqrt{2}} \left( \frac{1}{1 + \delta_2 + \mu} \right) (1 - \delta_{2k}) \|X - X_\varepsilon\|_F - \frac{1}{\sqrt{2}} \left( \frac{1}{1 + \delta_2 + \mu} \right) \varepsilon. 
\] (3.10)

Taken together and multiplying by \(\sqrt{2}\), the bounds (3.6) and (3.10) yield the inequality
\[
\left[ \frac{1 - \delta_{2k}}{1 + \delta_2 + \mu} - \frac{\sqrt{2} \delta_{3k}}{1 + \delta_2 + \mu} \right] \|X - X_\varepsilon\|_F \leq \left( \sqrt{2} + \frac{1}{1 + \delta_2 + \mu} \right) \varepsilon.
\]

Since \(\delta_2 \leq \delta_{2k} \leq \delta_{3k}\), the term in brackets on the left side is positive under the given condition (3.3) on \(\delta_{3k}\) by Lemma 3.8 (here \(c = \sqrt{2}\)). This leads to the assertion in item (i). Item (ii) is obtained by combining (3.7) and (3.10) instead.

**Remark 3.10.** The theorem is formulated for the distances \(\|X - X_\varepsilon\|_F\) in Frobenius norm, but in applications the difference in function values \(|f_{A,B}(X) - f_{A,B}(X_\varepsilon)|\) may be more relevant, in particular if \(f_{A,B}\) takes the form (1.5) of a shifted squared (semi)-norm for the operator \(A\) up to a constant (one may additionally assume \(f_{A,B}(X_\varepsilon) = 0\)). Using Taylor expansion at \(X_\varepsilon\) it holds
\[
f_{A,B}(X) - f_{A,B}(X_\varepsilon) = \langle A[X_\varepsilon] - B, X - X_\varepsilon \rangle_F + \frac{1}{2} \langle X - X_\varepsilon, A[X - X_\varepsilon]\rangle_F,
\]
and thus, under the assumptions of the theorem,
\[
|f_{A,B}(X) - f_{A,B}(X_\varepsilon)| \leq \varepsilon \|X - X_\varepsilon\|_F + \frac{1 + \delta_{2k}}{2} \|X - X_\varepsilon\|_F^2.
\]

Now the estimates for \(\|X - X_\varepsilon\|_F\) from the theorem can be used.

**Remark 3.11.** Similar to Corollary 3.7 and based on Remark 3.3, the theorem can be generalized to operators whose restricted spectral bounds are not centered around one, but are otherwise well conditioned on \(M \leq 2k\) or \(M \leq 3k\). If the conditions on \(\delta_{2k}\) or \(\delta_{3k}\) in Theorem 3.9 are fulfilled for a scaled operator \(\omega A\) where \(\omega > 0\), the statement of the theorem remains true for the initial operator \(A\) if one replaces \(\mu\) and \(\varepsilon\) by \(\mu/\omega\) and \(\varepsilon/\omega\), respectively.

### 3.4 Some concrete bounds

The upper bounds on \(\|X - X_\varepsilon\|_F\) provided by Theorem 3.9 become arbitrarily large when \(\delta_{3k}\) and \(\delta_{2k}\) approach the critical bounds. Therefore, in order to obtain reasonable estimates one needs considerably smaller values for \(\delta_{3k}\) and \(\delta_{2k}\). To obtain some intuition on the actual numbers we computed for several values of \(\mu\) the guaranteed error bounds for \(\|X - X_\varepsilon\|_F\) when \(A\) satisfies an RPD property where \(\delta_{3k}\) or \(\delta_{2k}\) is 90\% or 50\% of the critical values (3.3) and (3.5), respectively. The values are presented in Tables 1 (for \(\delta_{3k}\)) and Table 2 (for \(\delta_{2k}\)), where \(\varepsilon\) and \(X_\varepsilon\) are as in the theorem. It is clear that smaller RPD constants lead to better estimates.

In the context of so-called strict saddle point properties that have been discussed in related work, one can spell out these results as follows: For given \(\mu > 0\) and assuming \(\delta_{3k}\) (or \(\delta_{2k}\)) satisfies the bound asserted in the table, all critical points \(X\) of \(f_{A,B}\) on \(M_k\) either have the asserted distance \(\|X - X_\varepsilon\|_F\) to a point \(X_\varepsilon \in M \leq k\) satisfying \(|A[X_\varepsilon] - B| \leq \varepsilon\), or the Riemannian Hessian at \(X\) has a negative eigenvalue strictly less than \(-\mu\). In particular, the rows for \(\mu = 0\) in the tables provide bounds on the distance of local minima to the set of all such \(X_\varepsilon\).
Due to size limitation, however, we took called random RPD operators \( A \) for rank \( B \), uses the nearly isometric random matrices from [24]. In particular, \( A = B B^T \) with \( B \in \mathbb{R}^{n \times d} \) and \( B_{ij} \) random Gaussian \( N(0, 1/p) \). For large enough choices of \( n, p \), this \( A \) is RPD w.h.p.; see [24]. Due to size limitation, however, we took \( n = 50, \ k = 5, \ d = 50k \) which does not correspond to \( \delta_{3k} \leq \delta_{3k}^{\text{crit}} \) but still allows us to verify convergence of the algorithms.

In all cases, we generate an “exact” solution \( X^* = G H^T \in M_k \) with \( G, H \) random Gaussian matrices of size \( n \times k \). We then compute \( B = A(X^*) + \varepsilon \cdot N \) with \( N \) random Gaussian matrix, scaled so that \( \|N\|_F = 1 \), and \( \varepsilon \geq 0 \) a noise factor. This guarantees in particular that \( \|A[X_\varepsilon] - B\|_{\varepsilon, 2k} \leq \varepsilon \) as required in Theorem [3]. When \( \varepsilon > 0 \), the global minimizer of \( f_{A,B} \) is unknown and we therefore take \( X_\varepsilon = X^* \).

The methods that minimize \( f_{A,B} \) for rank \( k \) matrices are:

- **Embedded SD**: Riemannian steepest descent on \( M_k \) with the embedded submanifold geometry and Euclidean restricted metric. This is the same geometry as in [38, 27, 37, 39].

### Table 1: Error bounds for different values of \( \mu \). The second column states a sufficient condition on \( \delta_{3k} \), taken as nine tenth of the critical value \( \delta_{3k}^{\text{crit}} \), to obtain the estimate on \( \|X - X_\varepsilon\|_F \) in the third column (choosing \( \delta_2 = \delta_{2k} = \delta_{3k} \) in (3.4)). The fourth and fifth column display the results when \( \delta_{3k} \) is less than one half of the critical value.

<table>
<thead>
<tr>
<th>( \mu )</th>
<th>( \delta_{3k} \leq 0.9 \cdot \delta_{3k}^{\text{crit}} )</th>
<th>( |X - X_\varepsilon|_F )</th>
<th>( \delta_{3k} \leq 0.5 \cdot \delta_{3k}^{\text{crit}} )</th>
<th>( |X - X_\varepsilon|_F )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.3101</td>
<td>24.73 \cdot \varepsilon</td>
<td>0.1723</td>
<td>4.91 \cdot \varepsilon</td>
</tr>
<tr>
<td>0.2</td>
<td>0.2860</td>
<td>27.49 \cdot \varepsilon</td>
<td>0.1589</td>
<td>5.46 \cdot \varepsilon</td>
</tr>
<tr>
<td>0.4</td>
<td>0.2649</td>
<td>30.23 \cdot \varepsilon</td>
<td>0.1472</td>
<td>6.01 \cdot \varepsilon</td>
</tr>
<tr>
<td>0.6</td>
<td>0.2465</td>
<td>32.96 \cdot \varepsilon</td>
<td>0.1369</td>
<td>6.57 \cdot \varepsilon</td>
</tr>
<tr>
<td>0.8</td>
<td>0.2303</td>
<td>35.74 \cdot \varepsilon</td>
<td>0.1279</td>
<td>7.13 \cdot \varepsilon</td>
</tr>
<tr>
<td>1</td>
<td>0.2159</td>
<td>38.52 \cdot \varepsilon</td>
<td>0.1199</td>
<td>7.69 \cdot \varepsilon</td>
</tr>
</tbody>
</table>

### Table 2: Same as Table 1 but with conditions on \( \delta_{2k} \).

<table>
<thead>
<tr>
<th>( \mu )</th>
<th>( \delta_{2k} \leq 0.9 \cdot \delta_{2k}^{\text{crit}} )</th>
<th>( |X - X_\varepsilon|_F )</th>
<th>( \delta_{2k} \leq 0.5 \cdot \delta_{2k}^{\text{crit}} )</th>
<th>( |X - X_\varepsilon|_F )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.2526</td>
<td>24.28 \cdot \varepsilon</td>
<td>0.1403</td>
<td>4.85 \cdot \varepsilon</td>
</tr>
<tr>
<td>0.2</td>
<td>0.2301</td>
<td>27.05 \cdot \varepsilon</td>
<td>0.1278</td>
<td>5.41 \cdot \varepsilon</td>
</tr>
<tr>
<td>0.4</td>
<td>0.2108</td>
<td>29.84 \cdot \varepsilon</td>
<td>0.1171</td>
<td>5.97 \cdot \varepsilon</td>
</tr>
<tr>
<td>0.6</td>
<td>0.1943</td>
<td>32.64 \cdot \varepsilon</td>
<td>0.1079</td>
<td>6.53 \cdot \varepsilon</td>
</tr>
<tr>
<td>0.8</td>
<td>0.18</td>
<td>35.45 \cdot \varepsilon</td>
<td>0.1</td>
<td>7.1 \cdot \varepsilon</td>
</tr>
<tr>
<td>1</td>
<td>0.1675</td>
<td>38.27 \cdot \varepsilon</td>
<td>0.0930</td>
<td>7.66 \cdot \varepsilon</td>
</tr>
</tbody>
</table>

4 Numerical experiments

We now report on numerical experiments that verify our main result in Theorem [3].

We consider \( m = n \) and construct several matrix sensing problems involving two types of RPD operators \( A \) on \( \mathbb{R}^{n \times n} \). The first construction, called deterministic, is of the vectorized form \( A = I + \delta \cdot Q D Q^T \) with \( Q \) a random orthogonal matrix and \( D \) a diagonal matrix with random \( \pm 1 \) on its diagonal. This \( A \) will be RPD with constant \( \delta_k = \delta \) for all \( k \). The other construction, called random, uses the nearly isometric random matrices from [24]. In particular, \( A = B B^T \) with \( B \in \mathbb{R}^{n^2 \times d} \) and \( B_{ij} \) random Gaussian \( N(0, 1/p) \). For large enough choices of \( n, p \), this \( A \) is RPD w.h.p.; see [24].

The Riemannian Hessian at convergence has condition number about 18 which is too large when \( \delta_{3k} < \delta_{3k}^{\text{crit}} \).
• **Embedded CG:** Same as Embedded SD but now with non-linear conjugate gradients. This corresponds to the solver GeomCG from [37] but applied to sensing instead of completion.

• **Quotient SD:** Same as the embedded solvers but using the quotient geometry from [20].

• **ALS:** Alternating least squares with QR stabilization of the iterate to avoid ill-conditioning. This appears, for example, in [40, §2.1] where it is called the Nonlinear Gauss–Seidel method.

All methods except ALS were implemented using Manopt [5] with standard options (in particular, linearized exact line-search and Armijo backtracking). For ALS, the exact solution of the least-square system was computed using Cholesky decomposition. In the figures, we will also display estimated asymptotic convergence rates $\rho^\ell$ for the iterations $\ell = 1, 2, \ldots$. These were computed from the (Riemannian) Hessian $H_X$ at the limit point $X$. In particular, with $\kappa$ the condition number of $H_X$ as computed by Manopt, we used

$$\rho_{SD} = \left(\frac{\kappa - 1}{\kappa + 1}\right)^2, \quad \rho_{CG} = \left(\sqrt{\frac{\kappa - 1}{\kappa + 1}}\right)^2.$$

The rate $\rho_{SD}$ of SD corresponds to the Euclidean case with exact line search and is well known. For SD on Riemannian manifolds, it has been suggested to hold in [35, p. 270]. The rate $\rho_{CG}$ is intended for numerical verification and is rigorous for purely quadratic problems on $\mathbb{R}^{n \times n}$. For ALS, the asymptotic rate is computed from a block triangular decomposition of the Euclidean Hessian; see [22].

We could have compared to many different solvers, but for simplicity, we have restricted ourselves to mainly Riemannian algorithms since they are cheap per iteration and typically perform very well. In addition, many other low-rank optimization methods, like iterative hard thresholding (IHT), have the same asymptotic behavior. Another reason was to verify that the Riemannian Hessian used in Theorem 3.9 indeed captures the correct behavior in the convergence plots.

The convergence plots are visible in Figures 1 and 2. The left panels show convergence of the function value and clearly show a good correspondence of the theoretical asymptotic rates. The right panel is to verify that the error of the local minima obtained for the noisy problem are on the order of $\varepsilon$, as predicted by Theorem 3.9. This bound is explored in more detail in Table 3 for the deterministic case. Compared to Figure 1, we continued the iteration for 500 iterations and took the minimal value of $\|X(\ell) - X_\varepsilon\|_F$ as approximation of the limit point of each iteration. Observe that Table 1 predicts that the error should be bounded by $24.28\varepsilon$, which is certainly the case.

## 5 Conclusions

We have studied some properties of critical points of quadratic functions on manifolds of fixed-rank matrices. In particular, estimates for singular values of local minima have been derived that relate them to the singular values of the gradient at local minima. Then, under certain assumptions on bounds for the Rayleigh quotient of the Hessian on the cones of bounded rank matrices, which generalize the popular RIP conditions for matrix sensing, our estimates imply that there cannot be spurious local minima far away from the global one. In particular, local minima are absent in the noiseless case. The bounds required for the Rayleigh quotient to obtain this result are considerably weaker than in related previous publications.

---

4The rate without the square is easy to prove; see, e.g., [35, Chap. 7, Thm. 4.3].
Figure 1: Deterministic $A$ with $n = 50$, $k = 5$ and $\delta = 0.95 \cdot \delta_{\text{crit}}$. Both panels show in open circles for zero noise ($\varepsilon = 0$), and in closed circles for $\varepsilon = 10^{-5}$. Left panel shows in line the asymptotic convergence rate based on spectrum of Hessian.

Figure 2: Random Gaussian $A$ with $n = 50$, $k = 5$, $d = 5nk$. Both panels show in open circles for zero noise ($\varepsilon = 0$), and in closed circles for $\varepsilon = 10^{-5}$ (every five iterations shown). Left panel shows in line the asymptotic convergence rate based on spectrum of Hessian.

So far, our approach does not cover the important cases of matrix completion or the typical matrix equations in numerical linear algebra, as they do not meet the restricted spectral bounds. However, some of the presented techniques may still useful when studying these cases as well.

References


\[\varepsilon\min_{f=1}^{500} \|X^{(f)} - X_\varepsilon\|_F\]

<table>
<thead>
<tr>
<th>(\varepsilon)</th>
<th>max</th>
<th>mean</th>
<th>min</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>4.021 \cdot 10^{-12}</td>
<td>1.058 \cdot 10^{-12}</td>
<td>1.190 \cdot 10^{-14}</td>
</tr>
<tr>
<td>(10^{-10})</td>
<td>1.165 \cdot 10^{-10}</td>
<td>6.600 \cdot 10^{-11}</td>
<td>2.603 \cdot 10^{-11}</td>
</tr>
<tr>
<td>(10^{-08})</td>
<td>1.135 \cdot 10^{-08}</td>
<td>9.103 \cdot 10^{-09}</td>
<td>2.739 \cdot 10^{-09}</td>
</tr>
<tr>
<td>(10^{-06})</td>
<td>1.131 \cdot 10^{-06}</td>
<td>9.269 \cdot 10^{-07}</td>
<td>7.029 \cdot 10^{-07}</td>
</tr>
<tr>
<td>(10^{-04})</td>
<td>1.153 \cdot 10^{-04}</td>
<td>9.062 \cdot 10^{-05}</td>
<td>6.752 \cdot 10^{-05}</td>
</tr>
<tr>
<td>(10^{-02})</td>
<td>1.148 \cdot 10^{-02}</td>
<td>9.288 \cdot 10^{-03}</td>
<td>6.435 \cdot 10^{-03}</td>
</tr>
</tbody>
</table>

Table 3: Statistics for 20 random realizations of the deterministic \(A\) operator with \(n = 50, k = 5\) and \(\delta = 0.9 \cdot \delta_{\text{crit}}\).


A Proof of the Riemannian Hessian

Take $X = U\Sigma V^T \in M_k$ and let $\text{Exp}_X: B_\varepsilon \to M_k$ be the exponential defined on $B_\varepsilon \subset T_X M_k$, a sufficiently small ball around the zero tangent vector. It was shown in [36 Proposition A1] (see also [3 Proposition 24] for a different proof) that

$$\text{Exp}_X(Z) = X + Z + (I - P^\text{col}_X)Z\Sigma^{-1}U^TZ(I - P^\text{row}_X) + O(\|Z\|^3). \quad (A.1)$$

The Riemannian Hessian $H_X: T_X M_k \to T_X M_k$ of $f_{A,B}$ on $M_k$ is obtained as the standard (Euclidean) Hessian of the pullback $f \circ \text{Exp}_X$ at 0; see [2 Proposition 5.5.4]. Substituting (A.1) into

$$f_{A,B}(X) = \frac{1}{2}\langle A[X], X \rangle_F - \langle B, X \rangle_F,$$

gives the expansion

$$f_{A,B}(\text{Exp}_X(Z)) = f_{A,B}(X) + \langle A[X] - B, Z \rangle_F + \frac{1}{2}\langle A[Z], Z \rangle_F$$
$$+ \langle A[X] - B, (I - P^\text{col}_X)Z\Sigma^{-1}U^TZ(I - P^\text{row}_X) \rangle_F + O(\|Z\|^3).$$
The second and third term on the right-hand side of this expansion are second order in $Z$. The Riemannian Hessian therefore satisfies

$$\mathcal{H}_X[Z, Z] = \langle A[Z], Z \rangle_F + \langle A[X] - B_1(I - P_X^{\text{col}})ZV S^{-1}U^T Z(I - P_X^{\text{row}}) \rangle_F.$$ 

With the particular choices

$$\bar{G} = U\Sigma^{1/2}, \quad \bar{H} = V\Sigma^{1/2}, \quad Z = \partial G \cdot \bar{H}^T + \bar{G} \cdot \partial H^T,$$

we have $(I - P_X^{\text{col}})ZV \Sigma^{-1}U^T Z(I - P_X^{\text{row}}) = (I - P_X^{\text{col}})\partial G \cdot \partial H^T (I - P_X^{\text{row}})$. This establishes our expression (2.10) for the Riemannian Hessian.

23