On the Robust PCA and Weiszfeld’s Algorithm

Sebastian Neumayer\textsuperscript{1} · Max Nimmer\textsuperscript{1} · Simon Setzer\textsuperscript{3} · Gabriele Steidl\textsuperscript{1,2}

Published online: 5 April 2019
© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract
The principal component analysis (PCA) is a powerful standard tool for reducing the dimensionality of data. Unfortunately, it is sensitive to outliers so that various robust PCA variants were proposed in the literature. This paper addresses the robust PCA by successively determining the directions of lines having minimal Euclidean distances from the data points. The corresponding energy functional is non-differentiable at a finite number of directions which we call anchor directions. We derive a Weiszfeld-like algorithm for minimizing the energy functional which has several advantages over existing algorithms. Special attention is paid to carefully handling the anchor directions, where the relation between local minima and one-sided derivatives of Lipschitz continuous functions on submanifolds of $\mathbb{R}^d$ is taken into account. Using ideas for stabilizing the classical Weiszfeld algorithm at anchor points and the Kurdyka–Łojasiewicz property of the energy functional, we prove global convergence of the whole sequence of iterates generated by the algorithm to a critical point of the energy functional. Numerical examples demonstrate the very good performance of our algorithm.

Keywords Robust principal component analysis · PCA · Robust subspace recovery · Weiszfeld algorithm · Kurdyka-Łojasiewicz property

\textsuperscript{1} Department of Mathematics, Technische Universität Kaiserslautern, Paul-Ehrlich-Str. 31, 67663 Kaiserslautern, Germany
\textsuperscript{2} Fraunhofer ITWM, Fraunhofer-Platz 1, 67663 Kaiserslautern, Germany
\textsuperscript{3} Engineers Gate, London, UK
1 Introduction

Principal component analysis (PCA) [48] is an important tool for dimensionality reduction of data which is often applied as a preprocessing step, e.g., for classification or segmentation. The procedure provides dimensionality reduction by projecting the data onto a linear subspace maximizing the variance of the projection or equivalently minimizing the squared Euclidean distance error to the subspace. More precisely, let $N \geq d$ data points $x_1, \ldots, x_N \in \mathbb{R}^d$ be given. By $\| \cdot \|$ we denote the Euclidean norm and by $I_d$ the $d \times d$ identity matrix. PCA finds a $K$-dimensional affine subspace $\{ \hat{A} t + \hat{b} : t \in \mathbb{R}^K \}, 1 \leq K \leq d$, having smallest squared Euclidean distance from the data:

$$
(\hat{A}, \hat{b}) \in \arg \min_{A \in \mathbb{R}^{d,K}, b \in \mathbb{R}^d} \sum_{i=1}^N \min_{t \in \mathbb{R}^K} \| A t + b - x_i \|^2.
$$

(1)

While $\hat{A}$ and $\hat{b}$ in the above minimization problem are not unique, the affine subspace itself is uniquely determined if the empirical covariance matrix only has eigenvalues of multiplicity one, and goes through the offset (bias) $\bar{b} := \frac{1}{N} (x_1 + \ldots + x_N)$. Therefore, we can reduce our attention to data points $y_i := x_i - \bar{b}, i = 1, \ldots, N$, and subspaces through the origin minimizing the squared Euclidean distances to the $y_i, i = 1, \ldots, N$.

Setting the gradient of the inner function in (1) with respect to $t \in \mathbb{R}^K$ to zero and adding the constraint of $A$ being in the Stiefel manifold $\mathbb{S}_{d,K} = \{ A \in \mathbb{R}^{d,K} : A^T A = I_K \}$ to eliminate some redundancies, the problem reduces to

$$
\hat{A} \in \arg \min_{A \in \mathbb{S}_{d,K}} \sum_{i=1}^N \| P^\perp_A y_i \|^2,
$$

where $P^\perp_A := I_d - A A^T$ denotes the orthogonal projection onto $\mathcal{R}(A)^\perp = \mathcal{N}(A^T)$.

One of the most important properties of PCA is the nestedness of the PCA subspaces, i.e., for $K < \tilde{K} \leq d$ the optimal $K$-dimensional PCA subspace is contained in the $\tilde{K}$-dimensional one. In particular, the directions forming the columns of $\hat{A} = (\hat{a}_1 \ldots \hat{a}_K)$ can be found successively by computing for $k = 0, 1, \ldots, K - 1$,

$$
\hat{a}_{k+1} = \arg \min_{\|a\|=1} \sum_{i=1}^N \| P^\perp_{\hat{A}_k} P^\perp_A y_i \|^2
$$

(2)

or equivalently

$$
\hat{a}_{k+1} = \arg \max_{\|a\|=1} \sum_{i=1}^N \langle a, P^\perp_{\hat{A}_k} y_i \rangle^2
$$

(3)

where $\hat{A}_k := (\hat{a}_1 a_2 \ldots \hat{a}_k), k = 1, 2, \ldots, K - 1$, and $P^\perp_{A_0} = I_d$, see, e.g., [13]. The first problem (2) focuses on the minimization of the residual, while the second one (3) underlines the maximization of the variance in the PCA direction.
Fig. 1 Illustration of the effect of outliers on the PCA. The data set consists of 50 points lying approximately on a line from \((0,0)^T\) to \((1,1)^T\) and two outliers. The solid line is the result of PCA without outliers and the dashed line with all points.

Unfortunately, PCA is sensitive to outliers in the data, see Fig. 1. One possibility to circumvent this problem is to remove outliers before computing the principal components. However, in some contexts, outliers are difficult to identify and other data points are incorrectly given outlier status forcing a large number of deletions before a reliable estimate can be found.

Therefore, quite different methods were proposed in the literature to make PCA robust, in particular in robust statistics, see the books [16,32,39]. A recent overview and comparison of many of these methods and corresponding algorithms can be found in [28]. One approach consists in assigning different weights to data points based on their estimated relevance, to get a weighted PCA [21]. The RANSAC algorithm [9] repeatedly estimates the model parameters from a random subset of the data points until a satisfactory result is obtained as indicated by the number of data points within a certain error threshold. In a similar vein, least trimmed squares PCA models [51,53] aim to exclude outliers from the squared error functional, but in a deterministic way. Another possible approach is to minimize the median of the squared errors [40].

The variational model of Candès et al. [5] decomposes the data matrix \(Y = (y_1 \ldots y_N)\) into a low rank and a sparse part by minimizing

\[
\arg \min_{L,S} \|L\|_* + \lambda \|S\|_1 \quad \text{subject to} \quad Y = L + S,
\]

exploiting the nuclear norm \(\|L\|_*\) of \(L\) and the sum of the absolute values of entries \(\|S\|_1\). Then, \(L\) can be considered as robust part, while \(S\) addresses the outliers. Related approaches such as [42,60] separate the low rank component from the column sparse one using similar norms.

Another group of robust PCA approaches replaces the squared \(\ell_2\) norm in PCA by the \(\ell_1\) norm. Then, the minimization of the energy functional can be addressed by linear programming, see, e.g., Ke and Kanade [19]. Unfortunately, this norm is not rotationally invariant.

Mathematically interesting approaches follow (1)–(3), but skip the squares in the Euclidean distances and the inner products to find more robust directions. Taking pure Euclidean distances has several consequences. First of all, the energy functionals become non-differentiable at a finite number of subspaces spanned by matrices \(A\), resp. directions \(a\), which are collected within the so-called anchor set. Further, the offset \(\hat{b}\) in

\[
(\hat{A}, \hat{b}) \in \arg \min_{A \in \mathbb{R}^{d \times K}, b \in \mathbb{R}^d} \sum_{i=1}^N \min_{t \in \mathbb{R}^K} \|At + b - x_i\|
\]
is difficult to determine, see our small discussion in Sect. 5. Let us assume that an offset \( \hat{b} \) is given, so that we can restrict our attention to the data \( y_i = x_i - \hat{b}, \ i = 1, \ldots, N, \) and (4) becomes

\[
\hat{A} \in \arg \min_{A \in S_{d,K}} \sum_{i=1}^{N} \| P_{A} y_i \|.
\] (5)

Even then we lose the nested subspace property of the classical PCA, so that in particular

\[
\hat{a}_{k+1} = \arg \min_{\|a\|=1} \sum_{i=1}^{N} \| P_{a} P_{\hat{A}_k} y_i \|.
\] (6)

and

\[
\hat{a}_{k+1} = \arg \max_{\|a\|=1} \sum_{i=1}^{N} |\langle a, P_{\hat{A}_k} y_i \rangle|.
\] (7)

have in general nothing to do with the columns of the matrix \( \hat{A} \) obtained in (5). Finally, the residual minimizing point of view (6) leads to different results than the variance maximizing one in (7), see Fig. 2. A general theory on the recoverability of subspaces with functionals of the form \( \sum_{i=1}^{N} \| P_{A} y_i \|^p \) for \( p > 0 \) under certain assumptions on the data distributions can be found in [29,30].

The models (5)–(7) were considered in the literature. The maximization of (7) was suggested with more general scalable functions than just the absolute value by Huber [15, p. 203] and studied in detail as PP-PCA by Li and Chen [33]. It was reinvented and tackled with a greedy algorithm in [25], which was made more robust using median computations in [14]. For other methods in this direction, see also [37,38,42,46]. In [14] it was pointed out that the variance maximizing method in [25] lacks a certain robustness since it still involves mean computations. This was already demonstrated in Fig. 2.
Model (5) was treated by Ding et al. [7], where the authors circumvented the anchor set by smoothing the original energy functional. The paper gives no convergence analysis of the proposed algorithm. In [41] the optimization problem is tackled by a geodesic descent approach on the Grassmannian. The energy landscape is analyzed and under certain assumptions on the distribution of so-called inlier and outlier data, local convergence is shown. A tight convex relaxation approach for (5), called REAPER was suggested in [31]. The relaxation replaces the condition that the symmetric positive semidefinite matrix $AA^T$ has eigenvalues in $\{0, 1\}$ by the condition of eigenvalues in $[0, 1]$. This blows the problem size up. Numerically the relaxed problem can be solved via an iteratively re-weighted least squares (IRLS) algorithm. Usually this requires again a smoothing of the relaxed convex, but still non-differentiable functional. Another approach based on IRLS is analyzed in [27], where standard PCA is repeatedly performed on rescaled data points.

In this paper, we are interested in the residual minimizing approach (6). Recently, a minimization algorithm was published by Keeling and Kunisch [20]. Local convergence of their algorithm to a local minimizer was proved if the two parameters within the algorithm are chosen appropriately without a concrete specification of their range. The outcome of the algorithm is very sensitive to the choice of the parameters. We propose a minimization algorithm which is completely different from those in [20]. It is based on ideas of the classical Weiszfeld algorithm [59] for computing the geometric median of points in $\mathbb{R}^d$ and has the advantage that no parameters have to be tuned. In non-anchor directions the algorithm can be considered as gradient descent algorithm on the sphere, where the length of the gradient descent is automatically given. The treatment of anchor directions relies on one-sided directional derivatives of the energy functional. We show that such derivatives can be used to characterize local minima on submanifolds of $\mathbb{R}^d$ of locally Lipschitz continuous functions which is interesting on its own. We prove global convergence of our algorithm to a critical point of the energy functional, where we take special care of the anchor set.

Outline of the Paper In the next Sect. 2, we recall the Weiszfeld algorithm for computing the geometric median of given data points. Properties of the energy function, critical point conditions and the minimization algorithm are developed in Sect. 3. The main part of the paper is the convergence analysis of our algorithm in Sect. 4. Some remarks on the offset are given in Sect. 5. Numerical examples demonstrate the performance of our algorithm in Sect. 6. The paper ends with conclusions and ideas for future work in Sect. 7. The Appendix A provides a criterion for determining local minimizers of locally Lipschitz continuous functions on embedded manifolds in $\mathbb{R}^d$ which is applied in Sect. 3.

2 Weiszfeld’s Algorithm for Geometric Median Computation

We begin with a small review of Weiszfeld’s algorithm with two aims: first, the geometric median usually replaces the mean as offset in robust PCA methods. Second, having the original Weiszfeld algorithm in mind helps to understand the basic intention of our algorithm for minimizing (6).
The geometric median \( \hat{x} \in \mathbb{R}^d \) of pairwise distinct points \( x_i \in \mathbb{R}^d, i = 1, \ldots, N \), which are not aligned, is uniquely determined by

\[
\hat{x} := \arg \min_x \mathcal{E}(x) := \arg \min_x \sum_{i=1}^N \|x - x_i\|.
\]

An efficient algorithm for solving the geometric median problem is the Weiszfeld algorithm which goes back to the Hungarian mathematician A. Vazsonyi (Weiszfeld) [58,59] and can be also seen as a special maximizing-minimizing algorithm, see, e.g., [6]. In [22,23] it was recognized that the original algorithm of Weiszfeld fails if an iterate produced by the algorithm belongs to the so-called anchor set \( \mathcal{A} := \{x_1, \ldots, x_N\} \) consisting of the points where \( \mathcal{E} \) is non-differentiable. For bypassing the anchor points the most natural way is to define an appropriate descent direction of \( \mathcal{E} \) in those points [47,57]. To derive the algorithm recall that the function \( \mathcal{E} \) is convex and by Fermat’s rule the vector \( \hat{x} \in \mathbb{R}^d \) is a minimizer of \( \mathcal{E} \) if and only if

\[
0 \in \partial \mathcal{E}(\hat{x}) = \begin{cases} \nabla \mathcal{E}(\hat{x}) = \sum_{i=1}^N \frac{\hat{x} - x_i}{\|\hat{x} - x_i\|} & \text{if } \hat{x} \not\in \mathcal{A}, \\ \sum_{i=1}^N \frac{\hat{x} - x_i}{\|\hat{x} - x_i\|} + B_1(0) & \text{if } \hat{x} \in \mathcal{A}, \end{cases}
\]

where \( \partial \mathcal{E} \) denotes the subdifferential [52, Sect. 23] of \( \mathcal{E} \) and \( B_1(0) \) the closed Euclidean ball around zero with radius 1. Thus, a minimizer \( \hat{x} \not\in \mathcal{A} \) has to fulfill the fixed point equation

\[
\hat{x} = \left( \sum_{i=1}^N \frac{1}{\|\hat{x} - x_i\|} \right)^{-1} \sum_{i=1}^N \frac{x_i}{\|\hat{x} - x_i\|} = \hat{x} - \left( \sum_{i=1}^N \frac{1}{\|\hat{x} - x_i\|} \right)^{-1} \sum_{i=1}^N \frac{\hat{x} - x_i}{\|\hat{x} - x_i\|}, \quad (8)
\]

while \( \hat{x} \in \mathcal{A} \) is a minimizer if and only if

\[
\left\| \sum_{i=1}^N \frac{\hat{x} - x_i}{\|\hat{x} - x_i\|} \right\|_{\mathcal{A}} \leq 1. \quad (9)
\]

The Weiszfeld algorithm is an iterative algorithm which produces a sequence \( \{x^{(r)}\}_r \) as follows: If \( x^{(r)} \not\in \mathcal{A} \), then we apply the Picard iteration belonging to (8).
\[ x^{(r+1)} = x^{(r)} - \left( \sum_{i=1}^{N} \frac{1}{\|x^{(r)} - x_i\|} \right)^{-1} \sum_{i=1}^{N} \frac{x^{(r)} - x_i}{\|x^{(r)} - x_i\|} \cdot \nabla \mathcal{E}(x^{(r)}) \]

This is a gradient descent step with special step size \( s_{r}^{-1} \). If \( x^{(r)} \in A \), i.e. \( x^{(r)} = x_k \) for some \( k \in \{1, \ldots, N\} \), and fulfills the minimality condition (9), then the algorithm stops; otherwise we perform a descent step in direction of the subgradient in \( \partial \mathcal{E}(x^{(r)}) \) which is closest to zero

\[ x^{(r+1)} := x^{(r)} - \left( \sum_{i=1}^{N} \frac{1}{\|\hat{x} - x_i\|} \right)^{-1} \left( 1 - \frac{1}{\|G_k\|} \right) G_k, \]

where \( G_k := \sum_{i=1}^{N} \frac{x_k - x_i}{\|x_k - x_i\|} \in \partial \mathcal{E}(x_k) \).

Local and asymptotic convergence rates of the Weiszfeld algorithm were given in [18] and a non-asymptotic sublinear convergence rate was proved in [3]. The very good performance of Weiszfeld’s algorithm in comparison with the parallel proximal point algorithm was shown in [55] and a projected Weiszfeld algorithm was established in [44]. Keeling and Kunisch [20] suggested another stable algorithm for finding the geometric mean based on criticizing the behavior of the original Weiszfeld algorithm in anchor points and not taking its stabilized versions into account. A good reference on past and ongoing research in this direction is [3] and the references therein.

## 3 Weiszfeld-Like Algorithm for Robust PCA

Now, the minimization approach (6) is considered. First of all we establish that the direction \( a_{k+1} \) is indeed perpendicular to the previous directions \( \{a_1, \ldots, a_k\} \).

**Proposition 3.1** Let \( \varphi : \mathbb{R}_{\geq 0} \to \mathbb{R} \) be a strictly increasing function. In our application we are interested in \( \varphi(x) = x^{\frac{1}{2}} \). For any \( z_i \in \mathbb{R}^d, i = 1, \ldots, N \), its holds

\[ \arg \min_{\|a\|=1} \sum_{i=1}^{N} \varphi(\|P_{a}^{+} z_i\|^2) \in \text{span}\{z_i : i = 1, \ldots, N\}. \]

**Proof** Every \( a \in \mathbb{R}^d \) with \( \|a\| = 1 \) can be written as

\[ a = \frac{\tilde{a} + \tilde{a}_\perp}{\|\tilde{a} + \tilde{a}_\perp\|}, \]

\( \odot \) Springer
where \( \tilde{a} \in \text{span}\{z_i : i = 1, \ldots, N\} \) and \( \tilde{a}_\perp \) is in the orthogonal complement of \( \text{span}\{z_i : i = 1, \ldots, N\} \). Then, we have for every \( z \in \text{span}\{z_i : i = 1, \ldots, N\} \),
\[
\|P_\perp \tilde{a} z\|^2 = \|z\|^2 - \langle \tilde{a}, z \rangle^2 \geq \|z\|^2 - \frac{\langle \tilde{a}, z \rangle^2}{\|\tilde{a}\|^2 + \|\tilde{a}_\perp\|^2}
\]
with equality if \( \|\tilde{a}_\perp\| = 0 \). Since \( \varphi \) is strictly increasing, any minimizer \( \hat{a} \) must be in \( \text{span}\{z_i : i = 1, \ldots, N\} \).

We have to deal with the function
\[
E(a) := \sum_{i=1}^N E_i(a) = \sum_{i=1}^N \|P_\perp a y_i\|,
\]
which is continuously differentiable on \( \mathbb{R}^d \) except for \( a \in \mathbb{R}^d \) satisfying \( \|P_\perp a y_k\| = \|(I_d - aa^T)y_k\| = 0 \) for some \( k \in \{1, \ldots, N\} \). This is equivalent to \( y_k = a\langle a, y_k \rangle \) and if \( a \in S^{d-1} \), this implies \( a \in \{\pm y_k/\|y_k\|\} \). Let
\[
\mathcal{A} := \left\{ \pm \frac{y_i}{\|y_i\|} : i = 1, \ldots, N \right\}
\]
denote this set of directions on \( S^{d-1} \), where \( E \) is non-differentiable. Similarly as in Weiszfeld’s algorithm, we call it anchor set.

The following theorem collects important properties of \( E \). The third property relies on the relation between one-sided derivatives and local minima of Lipschitz continuous functions on embedded manifolds in \( \mathbb{R}^d \). The definition of one-sided derivatives and a theorem characterizing local minima is given in Appendix A. In our case the embedded manifold is the sphere \( S^{d-1} := \{a \in \mathbb{R}^d : \|a\| = 1\} \).

**Theorem 3.2** Let \( E \) be defined by (11).
1. The function \( E \) is locally Lipschitz continuous on \( \mathbb{R}^d \).
2. For \( a \in S^{d-1} \setminus \mathcal{A} \), it holds
\[
\nabla E(a) = -P_\perp^a C_a a, \quad C_a := \sum_{i=1}^N \frac{1}{\|P_\perp a y_i\|} y_i y_i^T,
\]
and \( \nabla E(a) \) is in the tangent space \( T_a S^{d-1} \) of \( S^{d-1} \) at \( a \).
3. A direction \( a \in \mathcal{A} \) is a local minimizer of \( E \) if
\[
\|G_{a,K}\| < \sum_{k \in K} \|y_k\|,
\]
where \( K := \{k \in \{1, \ldots, N\} : \|P_\perp a y_k\| = 0\} \) and
\[
G_{a,K} := P_\perp^a C_a,K a, \quad C_{a,K} := \sum_{i \notin K} \frac{1}{\|P_\perp a y_i\|} y_i y_i^T.
\]
Proof. 1. It suffices to show the property for the summands $E_i$. For an arbitrary fixed $a \in \mathbb{R}^d$, let $\|a - a_i\| \leq \varepsilon$, $i = 1, 2$. Then, we obtain

$$|E_i(a_1) - E_i(a_2)| = \|P_{a_1}^\perp y_i - P_{a_2}^\perp y_i\| \leq \|a_1 a_1^T - a_2 a_2^T\| F \|y_i\|$$

$$= \frac{1}{2} \|(a_1 - a_2)(a_1^T + a_2^T) + (a_1 + a_2)(a_1^T - a_2^T)\| F \|y_i\|$$

$$\leq (\|a\| + \varepsilon) \|y_i\| \|a_1 - a_2\| F.$$

2. By straightforward computation we obtain for points $a \in \mathbb{R}^d$, where $E$ is differentiable,

$$\nabla E_i(a) = -\frac{1}{\|P_{a_i}^\perp y_i\|} \left( P_{a_i}^\perp y_i a + y_i y_i^T P_{a_i}^\perp a \right).$$

The second summand vanishes for $a \in S^{d-1}$ which yields (12). Since $P_a^\perp$ projects to the space orthogonal to $a$, the gradient $\nabla E(a)$ lies in $T_a S^{d-1}$.

3. For $a \in A$ we have $a \in \{ \pm y_k/\|y_k\| \}$ for $k \in K$. Then, the one-sided directional derivative of $E_k$ at $a \in A$ in direction $h \in T_a S^{d-1}$ reads as

$$D E_k(a; h) = \lim_{\alpha \downarrow 0} \frac{E_k(a + \alpha a) - E_k(a)}{\alpha} = \lim_{\alpha \downarrow 0} \frac{\|(I_d - (a + \alpha h)(a + \alpha h)^T)y_k\|}{\alpha}$$

$$= \lim_{\alpha \downarrow 0} \frac{\|a h^T y_k + a h a^T y_k + \alpha^2 h h^T y_k\|}{\alpha}$$

$$= \|a h^T y_k\| = \|h^T y_k\| = \|h\|\|y_k\|.$$

where the last equality follows from $a \in \{ \pm y_k/\|y_k\| \}$. For $i \notin K$ part 2 of the proof implies

$$D E_i(a; h) = \langle \nabla E_i(a), h \rangle = -\left( P_a^\perp \frac{1}{\|P_{a_i}^\perp y_i\|} y_i y_i^T a, h \right)$$

so that in summary

$$D E(a; h) = \sum_{k \in K} \|y_k\| \|h\| - \langle P_a^\perp C_a, K a, h \rangle.$$

Since $E$ is locally Lipschitz continuous on $\mathbb{R}^d$, we conclude by Theorem A.2 that $a \in A$ is a local minimizer if

$$\langle P_a^\perp C_a, K a, h \rangle < \sum_{k \in K} \|y_k\| \|h\|.$$
for all $h \in T_a S^{d-1}$. Since $P_a^k C_{a,K} a \in T_a S^{d-1}$ this equivalent to

$$\| P_a^k C_{a,K} a \| < \sum_{k \in K} \| y_k \|. \quad \Box$$

The weighted covariance matrix $C_a$ allows a conceptual connection to standard PCA. It is well known that the principal components can be found by computing the (unweighted) empirical covariance matrix $\frac{1}{N-1} \sum_{i=1}^N y_i y_i^T$. In our case, for $a \notin A$, the critical point condition $0 = \nabla E(a) = -P_a^k C_{a,K} a$ implies that

$$C_a a = a^T C_{a,K} a = s_a a, \quad (13)$$

which means that $a$ is an eigenvector of the weighted covariance matrix $C_a$ with eigenvalue $s_a := a^T C_{a,K} a$. Furthermore, the reduced gradient $G_{a,K}$ is the gradient of the summands in $E(a)$ which do not equal zero. This is analogous to $G_k$ in the Weiszfeld iteration in the non-smooth case (10).

To establish a Weiszfeld-like algorithm for the minimization of $E$, we consider two cases:

If $a \notin A$ is a critical point, then $0 = \nabla E(a)$ and (13) can be rewritten as the fixed point equation

$$a = s_a^{-1} C_a a = a + s_a^{-1} (I_d - a a^T) C_a a = a + s_a^{-1} P_a^k C_{a,K} a = a + s_a^{-1} \nabla E(a). \quad (14)$$

This gives rise to the gradient descent step on $S^{d-1}$:

$$a^{(r+1)} = \frac{C_{a^{(r)}} a^{(r)}}{\|C_{a^{(r)}} a^{(r)}\|}.$$  

This step also appears in the algorithm proposed by Ding et al. [7] from another point of view. Note that the factor $s_a^{-1}$ cancels out when projecting on $S^{d-1}$, but can be interpreted as step size in the convergence analysis, see (14).

If $a \in A$ and $\|G_{a,K}\| > \sum_{k \in K} \|y_k\|$, then we suggest to use

$$\left(1 - \frac{\sum_{k \in K} \|y_k\|}{\|G_{a,K}\|}\right) G_{a,K}$$

instead of the gradient as descent direction which results in the iteration

$$a^{(r+\frac{1}{2})} := a^{(r)} + s_{a^{(r)},K}^{-1} \left(1 - \frac{\sum_{k \in K} \|y_k\|}{\|G_{a^{(r)},K}\|}\right) G_{a,K}.$$
with

\[ s_{a,K} := a^T C_{a,K} a = \sum_{i \notin K} \frac{\langle a, y_i \rangle^2}{\| P^*_a y_i \|} \]

and subsequent orthogonal projection onto \( \mathbb{S}^{d-1} \). This is inspired by the iteration step in the non-smooth case (10) in the Weiszfeld algorithm, as the descent direction \( G_{a,K} \) is the sum of the gradients of all smooth summands of \( E \) and the step size incorporates the optimality condition in a similar manner. In summary, we obtain Algorithm 1.

**Algorithm 1** Algorithm for Minimizing \( E \) over \( \mathbb{S}^{d-1} \)

**Input:** \( y_i \in \mathbb{R}^d, i = 1, \ldots, N \), pairwise distinct with positive definite covariance matrix \( a^{(0)} \in \mathbb{S}^{d-1} \)

\( r = 0 \)

repeat
  if \( \| P_{a^{(r)}} y_k \| \neq 0 \) for all \( k \in \{1, \ldots, N\} \) then
    \[ C_{a^{(r)}} := \sum_{i=1}^N \frac{1}{\| P_{a^{(r)}} y_i \|} y_i y_i^T \]
    \[ a^{(r+1)} := \frac{C_{a^{(r)}} a^{(r)}}{\| G_{a^{(r)}} a^{(r)} \|} \]
  else if \( \| P_{a^{(r)}} y_k \| = 0 \) for \( k \in K \subset \{1, \ldots, N\} \) then
    \[ C_{a^{(r)},K} := \sum_{i \notin K} \frac{1}{\| P_{a^{(r)}} y_i \|} y_i y_i^T \]
    \[ G_{a^{(r)},K} := P_{a^{(r)}} C_{a^{(r)},K} a^{(r)} \]
    if \( \| G_{a^{(r)},K} \| \leq \sum_{k \in K} \| y_k \| \) then
      termination
    else
      \[ s_{a^{(r)},K} := \sum_{i \notin K} \frac{\langle a^{(r)}, y_i \rangle^2}{\| P_{a^{(r)}} y_i \|} \]
      \[ a^{(r+\ell)} := a^{(r)} + s_{a^{(r)},K}^{-1} \left( 1 - \frac{\sum_{k \in K} \| y_k \|}{\| G_{a^{(r)},K} \|} \right) G_{a^{(r)},K} \]
      \[ a^{(r+1)} := \frac{a^{(r+\ell)}}{\| a^{(r+\ell)} \|} \]
  \[ r \rightarrow r + 1 \]
until a stopping criterion is reached

**4 Convergence Analysis**

In this section, we show that that sequence generated by the Algorithm 1 converges to a critical point of \( E \), where we say that \( a \in \mathbb{S}^{d-1} \) is a critical point of \( E \) on \( \mathbb{S}^{d-1} \) if one of the following conditions is fulfilled:
(i) \( a \notin \mathcal{A} \) and \(-\nabla E(a) = P_a^+ C_a a = 0\).
(ii) \( a \in \mathcal{A} \) and \( \|G_{a,K}\| \leq \sum_{k \in K} \|y_k\| \).

We need four lemmata and apply a theorem of Attouch, Bolte and Svaiter [2] on the convergence of gradient schemes for functions having the Kurdyka–Łojasiewicz property. First, we establish a connection between the iterative scheme and the critical point conditions. In the second lemma, a descent property of the iterates is shown. Afterwards, we assert that the distance between iterates tends to zero which implies that the set of accumulation points is compact and connected. The result of the fourth lemma implies that if the sequence converges to an anchor direction, it must be a critical point. The section finishes with a convergence theorem for the whole sequence.

**Lemma 4.1** For the sequence \( \{a^{(r)}\}_r \) produced by Algorithm 1 we have \( a^{(r+1)} = a^{(r)} \) if and only if \( a^{(r)} \) is a critical point of \( E \) on \( \mathbb{S}^{d-1} \). If the iteration stops after finitely many steps, then it has reached a critical point.

**Proof**

1. Let \( a^{(r+1)} = a^{(r)} = a \). If \( a \) is not in the anchor set, this implies \( \frac{C_a a}{\|C_a a\|} = a \) and hence \( P_a^+ C_a a = \|C_a a\| P_a^+ a = 0 \). If \( a \) is in the anchor set, then relation in ii) must be fulfilled by the stopping condition.

2. Let \( a^{(r)} \in \mathbb{S}^{d-1} \) be a critical point of \( E \). If \( a^{(r)} \) is not in the anchor set, then by definition \( 0 = P_a^+ C_{a^{(r)}} a^{(r)} = C_{a^{(r)}} a^{(r)} - a^{(r)} (a^{(r)})^T C_{a^{(r)}} a^{(r)} \) so that

\[
a^{(r+1)} = \frac{C_{a^{(r)}} a^{(r)}}{\|C_{a^{(r)}} a^{(r)}\|} = \frac{(a^{(r)})^T C_{a^{(r)}} a^{(r)} a^{(r)}}{\|((a^{(r)})^T C_{a^{(r)}} a^{(r)} a^{(r)})\|} = a^{(r)}.
\]

If \( a^{(r)} \) is in the anchor set, then \( \|G_{a^{(r)},K}\| \leq \sum_{k \in K} \|y_k\| \) and the iteration stops by definition, i.e. \( a^{(r+1)} = a^{(r)} \).

**Lemma 4.2** Let \( \{a^{(r)}\}_r \) be the sequence generated by Algorithm 1. If \( a^{(r+1)} \neq a^{(r)} \), then \( E(a^{(r+1)}) < E(a^{(r)}) \). The sequence \( \{E(a^{(r)})\}_r \) converges to some value \( \hat{E} \geq 0 \).

**Proof**

If the sequence of function values decreases, its convergence follows immediately from the fact that \( E \) is bounded from below by zero. To show the decrease property, we set \( a := a^{(r)} \), \( \tilde{a} := a^{(r+\frac{1}{2})} \) and \( \tilde{a} = a^{(r+1)} \) and abbreviate

\[
G := G_{a^{(r)},K}, \quad C := C_{a^{(r)},K} \quad \text{and} \quad s := s_{a^{(r)},K},
\]

where \( K \) is the empty set if \( a^{(r)} \) is not an anchor direction.

**Case 1** Let \( a \notin \mathcal{A} \) be a non-anchor direction, i.e., \( K = \emptyset \). For \( u \geq 0 \), \( v > 0 \) it holds \( u - v \leq \frac{u^2 - v^2}{2v} \) so that

\[
E(\tilde{a}) - E(a) = \sum_{i=1}^N \left( \|P_{\tilde{a}}^+ y_i\| - \|P_{a}^+ y_i\| \right) \leq \sum_{i=1}^N \frac{\|P_{\tilde{a}}^+ y_i\|^2 - \|P_{a}^+ y_i\|^2}{2\|P_{\tilde{a}}^+ y_i\|} = \sum_{i=1}^N \frac{\|\tilde{a}^T y_i - y_i\|^2 - \|a^T y_i - y_i\|^2}{2\|P_{\tilde{a}}^+ y_i\|}. \]
Using $\|u - v\|^2 - \|w - v\|^2 = 2\langle u - w, u - v \rangle - \|u - w\|^2$ we get

$$E(\tilde{a}) - E(a) \leq \sum_{i=1}^{N} \frac{\langle \tilde{a} \tilde{a}^T y_i - aa^T y_i, \tilde{a} \tilde{a}^T y_i - y_i \rangle}{\| P_{\tilde{a}}^\perp y_i \|} - \sum_{i=1}^{N} \frac{\| \tilde{a} \tilde{a}^T y_i - aa^T y_i \|^2}{2\| P_{\tilde{a}}^\perp y_i \|}$$

$$= \sum_{i=1}^{N} \frac{\langle aa^T y_i, P_{\tilde{a}}^\perp y_i \rangle - \langle \tilde{a} \tilde{a}^T y_i, P_{\tilde{a}}^\perp y_i \rangle}{\| P_{\tilde{a}}^\perp y_i \|} - \sum_{i=1}^{N} \frac{\| \tilde{a} \tilde{a}^T y_i - aa^T y_i \|^2}{2\| P_{\tilde{a}}^\perp y_i \|}$$

$$= \sum_{i=1}^{N} \frac{\langle aa^T y_i, P_{\tilde{a}}^\perp y_i \rangle}{\| P_{\tilde{a}}^\perp y_i \|} - \sum_{i=1}^{N} \frac{\| \tilde{a} \tilde{a}^T y_i - aa^T y_i \|^2}{2\| P_{\tilde{a}}^\perp y_i \|}$$

$$= a^T P_{\tilde{a}}^\perp Ca - \sum_{i=1}^{N} \frac{\| \tilde{a} \tilde{a}^T y_i - aa^T y_i \|^2}{2\| P_{\tilde{a}}^\perp y_i \|}$$

$$= \| Ca \| a^T P_{\tilde{a}}^\perp a - \sum_{i=1}^{N} \frac{\| \tilde{a} \tilde{a}^T y_i - aa^T y_i \|^2}{2\| P_{\tilde{a}}^\perp y_i \|}, \quad (15)$$

which finally implies

$$E(\tilde{a}) - E(a) \leq - \sum_{i=1}^{N} \frac{\| \tilde{a} \tilde{a}^T y_i - aa^T y_i \|^2}{2\| P_{\tilde{a}}^\perp y_i \|}.$$

Since $a, \tilde{a} \in \text{span}(Y)$ the right-hand side is strictly negative except for $\tilde{a} \in \{ \pm a \}$ which was excluded.

**Case 2** Let $a \in A$, i.e., $\| P_{\tilde{a}}^\perp y_k \| = 0$ for $k \in K \neq \emptyset$ and

$$\| G \| > \sum_{k \in K} \| y_k \| =: \alpha.$$

From $P_{\tilde{a}}^\perp y_k = 0, k \in K$, i.e., $y_k = a(a^T y_k)$ we obtain $\| y_k \| = |a^T y_k|$. Since $\tilde{a} = a + s^{-1} \left( 1 - \frac{\alpha}{\| G \|} \right) G$ and $a \perp G$ we have

$$\| \tilde{a} \|^2 = 1 + s^{-2} \left( 1 - \frac{\alpha}{\| G \|} \right)^2 \| G \|^2 > 1$$

$$1 - \frac{1}{\| \tilde{a} \|^2} = \frac{(\| G \| - \alpha)^2}{s^2\| \tilde{a} \|^2} =: \mu^2. \quad (16)$$

We have to estimate

$$E(\tilde{a}) - E(a) = \sum_{i \notin K} (\| P_{\tilde{a}}^\perp y_i \| - \| P_{\tilde{a}}^\perp y_i \|) + \sum_{k \in K} \| P_{\tilde{a}}^\perp y_k \|.$$
First, we get for \( k \in K \),

\[
\| P_{\tilde{a}}^{-1} y_k \|_2^2 = y_k^T \left( I_d - \frac{\tilde{a} \tilde{a}^T}{\| \tilde{a} \|^2} \right) y_k = \| y_k \|^2 - \frac{y_k^T \tilde{a} \tilde{a}^T y_k}{\| \tilde{a} \|^2} = \mu^2 \| y_k \|^2.
\]

so that

\[
\sum_{k \in K} \| P_{\tilde{a}}^{-1} y_k \| = \mu \alpha. \tag{17}
\]

Replacing the sum over \( \{1, \ldots, N\} \) in case 1 of the proof by the one over \( \{1, \ldots, N\} \setminus K \) we get instead of (15)

\[
\sum_{i \notin K} \left( \| P_{\tilde{a}}^{-1} y_i \| - \| P_{\tilde{a}}^{-1} y_i \|_2 \right) \leq a^T \left( I_d - \frac{\tilde{a} \tilde{a}^T}{\| \tilde{a} \|^2} \right) Ca - \sum_{i \notin K} \frac{\| \tilde{a} \tilde{a}^T y_i - aa^T y_i \|_2^2}{2 \| P_{\tilde{a}}^{-1} y_i \|}. \tag{18}
\]

By definition of \( \tilde{a} \) and (16) we can rewrite

\[
a^T \left( I_d - \frac{\tilde{a} \tilde{a}^T}{\| \tilde{a} \|^2} \right) Ca = a^T Ca - \frac{1}{\| \tilde{a} \|^2} \left( a^T Ca + s^{-1} \left( 1 - \frac{\alpha}{\| G \|} \right) G^T Ca \right)
\]

\[
= s - \frac{1}{\| \tilde{a} \|^2} \left( s + s^{-1} \left( 1 - \frac{\alpha}{\| G \|} \right) \| G \|^2 \right)
\]

\[
= \mu^2 s - \frac{1}{\| \tilde{a} \|^2} s^{-1} \left( 1 - \frac{\alpha}{\| G \|} \right) \| G \|^2
\]

\[
= \mu^2 s \left( 1 - \frac{\| G \|}{\| G \| - \alpha} \right) = -\mu^2 s \frac{\alpha}{\| G \| - \alpha}
\]

\[
= -\mu \frac{\alpha}{\| \tilde{a} \|^2}. \tag{19}
\]

For the second sum in (18) we get

\[
- \sum_{i \notin K} \frac{\| \tilde{a} \tilde{a}^T y_i - aa^T y_i \|_2^2}{2 \| P_{\tilde{a}}^{-1} y_i \|} = -\frac{1}{2} \left( \tilde{a}^T Ca - 2\tilde{a}^T a \tilde{a}^T Ca + a^T Ca \right).
\]

Application of \( a^T CG = \| G \|^2 \) and of the definition of \( \tilde{a} \) leads to

\[
\tilde{a}^T Ca = \frac{1}{\| \tilde{a} \|^2} \left( s + 2s^{-1}(\| G \| - \alpha)\| G \| + s^{-2}(\| G \| - \alpha)^2 \right) \frac{1}{\| G \|^2} G^T CG
\]

and

\[
\tilde{a}^T a \tilde{a}^T Ca = \frac{1}{\| \tilde{a} \|^2} \left( s + s^{-1}(\| G \| - \alpha)\| G \| \right).
\]
Hence we obtain
\[
- \sum_{i \not\in \mathcal{K}} \frac{\|\tilde{a}^T y_i - aa^T y_i\|^2}{2\|P_{\tilde{a}} y_i\|} = -\frac{1}{2} \frac{1}{\|\tilde{a}\|^2} \left( (\|\tilde{a}\|^2 - 1)s + s^{-2}(\|G\| - \alpha)^2 \frac{1}{\|G\|^2} G^T C G \right) \\
= -\frac{1}{2} \mu^2 \left( s + \frac{1}{\|G\|^2} G^T C G \right).
\]

Since $C$ is symmetric positive definite we conclude by Young’s inequality
\[
s + \frac{1}{\|G\|^2} G^T C G = \|C \frac{1}{2} a\|^2 + \frac{1}{\|G\|} \|C \frac{1}{2} G\|^2 \geq 2 \frac{1}{\|G\|} a^T C \frac{1}{2} C \frac{1}{2} G = 2\|G\|
\]
so that
\[
- \sum_{i \not\in \mathcal{K}} \frac{\|\tilde{a}^T y_i - aa^T y_i\|^2}{2\|P_{\tilde{a}} y_i\|} \leq -\mu^2 \|G\|.
\]

Combining this equation with (17), (18) and (19), and using that $\|\tilde{a}\| > 1$, we obtain
\[
E(\tilde{a}) - E(a) \leq -\mu \frac{\alpha}{\|\tilde{a}\|} - \mu^2 \|G\| + \mu \alpha = \mu \left( \alpha \left( 1 - \frac{1}{\|\tilde{a}\|} \right) - \mu \|G\| \right) \\
< \mu^2 (\alpha - \|G\|) < 0.
\]

Lemma 4.3 Let $\{a^{(r)}\}_r$ be an infinite sequence generated by Algorithm 1. Then we have
\[
\lim_{r \to \infty} \|a^{(r+1)} - a^{(r)}\| = 0.
\] (20)

The set of accumulation points is compact and connected.

Proof Since the number of anchor directions is finite, we can choose $R$ large enough such that all iterates $a^{(r)}$, $r \geq R$ are not anchor directions. Since the projection $\Pi_{S^{d-1}}$ onto the unit sphere is non-expansive for points not in the interior of the unit ball, we obtain
\[
\|a^{(r+1)} - a^{(r)}\| = \left\| \Pi_{S^{d-1}} \left( a^{(r)} + s_{a^{(r)}}^{-1} P_{a^{(r)}} C a^{(r)} a^{(r)} \right) - \Pi_{S^{d-1}} (a^{(r)}) \right\| \\
\leq \frac{\|P_{a^{(r)}} C a^{(r)} a^{(r)}\|}{s_{a^{(r)}}}.
\]

We show that all accumulation points of $\{\beta_r\}_r$ with $\beta_r := \|a^{(r+1)} - a^{(r)}\|$ are zero. Note that such accumulation points exist, since $S^{d-1}$ is compact so that the sequence is bounded from below and above. Let $\{\beta_{r_j}\}_j$ converge to $\hat{\beta}$ which is then also true for every subsequence. Let $\{\beta_{r_{ji}}\}_i$ be any subsequence for which $\{a^{(r_{ji})}\}_i$ converges.
to an accumulation point \( \hat{a} \). For simplicity of notation, we skip the second index \( i \). We distinguish two cases:

1. Let \( \hat{a} \not\in A \) be a non-anchor direction. Then the update operator \( T(a) = \frac{C_{\hat{a}} a}{\|C_{\hat{a}} a\|} \) of the algorithm is continuous in \( \hat{a} \) so that \( \lim_{j \to \infty} a^{(r_j+1)} = \lim_{j \to \infty} T(a^{(r_j)}) = T(\hat{a}) \). By Lemma 4.2 and continuity of \( E \), we get

\[
\hat{E} = \lim_{j \to \infty} E(a^{(r_j)}) = E(\hat{a})
\]

\[
\hat{E} = \lim_{j \to \infty} E(a^{(r_j+1)}) = E(T(\hat{a}))
\]

so that \( \hat{a} = T(\hat{a}) \). This in turn yields \( P_{\hat{a}}^\perp C_{\hat{a}} \hat{a} = \|C_{\hat{a}} \hat{a}\| P_{\hat{a}}^\perp \hat{a} = 0 \). Since the \( \|P_{a^{(r_j)}}^\perp y_i\| \) are bounded from above, and since \( a^{(r)} \in \mathcal{R}(Y) \) for all \( r \) we conclude that \( s_{a^{(r_j)}} \) is bounded from below. Taking the continuity of the involved operators in \( \hat{a} \) into account, this implies

\[
\lim_{j \to \infty} \frac{\|P_{a^{(r_j)}}^\perp C_{a^{(r_j)}} a^{(r_j)}\|}{s_{a^{(r_j)}}} = 0.
\]

2. Let \( \hat{a} \in A \) be an anchor direction. Then it holds

\[
\lim_{j \to \infty} s_{a^{(r_j)}} = \lim_{j \to \infty} \sum_{i=1}^{N} (y_i^T a^{(r_j)})^2 / \|P_{a^{(r_j)}}^\perp y_i\| = \infty,
\]

while

\[
\|P_{a^{(r_j)}}^\perp C_{a^{(r_j)}} a^{(r_j)}\| = \left\| \sum_{i=1}^{N} y_i^T a^{(r_j)} \frac{P_{a^{(r_j)}}^\perp y_i}{\|P_{a^{(r_j)}}^\perp y_i\|} \right\| \leq \sum_{i=1}^{N} |y_i^T a^{(r_j)}| \leq \sum_{i=1}^{N} \|y_i\|,
\]

so that

\[
\lim_{j \to \infty} \frac{\|P_{a^{(r_j)}}^\perp C_{a^{(r_j)}} a^{(r_j)}\|}{s_{a^{(r_j)}}} = 0.
\]

This proves (20). By Ostrowski’s Theorem, the set of accumulation points of the sequence of iterates is compact and connected.

Lemma 4.4 Let \( \hat{a} \) be an anchor direction. Let \( T \) denote the iteration function of Algorithm 1. Then

\[
\lim_{a \to \hat{a}} \frac{\|T(a) - \hat{a}\|}{\|a - \hat{a}\|} = \frac{\|G_{\hat{a},K}\|}{\sum_{k \in K} \|y_k\|}.
\]
Proof For simplicity of notation, we assume that $K = \{k\}$ and without loss of generality $\hat{a} = y_k / \|y_k\|$. We set

$$T(a) = \frac{C_a a}{\|C_a a\|} = \frac{a + \frac{1}{s_a} P_{a}^{+} C_a a}{\|a + \frac{1}{s_a} P_{a}^{+} C_a a\|} : \frac{T_a}{\|T_a\|}.$$ 

Similarly as in the proof of Lemma 4.3, Case 1, we have that $P_{a}^{+} C_a a$ is bounded from above and $\lim_{a \to \hat{a}} s_a = \infty$ so that $\lim_{a \to \hat{a}} \|T_a\| = 1$. We calculate

$$\frac{\|T(a) - \hat{a}\|^2}{\|a - \hat{a}\|^2} = \frac{\|a - \|T_a\| \hat{a} + \frac{1}{s_a} P_{a}^{+} C_a a\|^2}{\|a - \hat{a}\|^2 \|T_a\|^2}$$

$$= \frac{\|a - \|T_a\| \hat{a} + 2 \langle a - \hat{a}, \frac{1}{s_a} P_{a}^{+} C_a a \rangle + \frac{1}{s_a} \|P_{a}^{+} C_a a\|^2}{\|T_a\|^2 \|a - \hat{a}\|^2} \cdot (21)$$

The first term can be rearranged as

$$\frac{\|a - \|T_a\| \hat{a}\|^2}{\|T_a\|^2 \|a - \hat{a}\|^2} = \frac{1}{\|T_a\|^2} \frac{\|a - \hat{a} + (1 - \|T_a\|) \hat{a}\|^2}{\|a - \hat{a}\|^2}$$

$$= \frac{1}{\|T_a\|^2} \frac{\|a - \hat{a}\|^2 + 2 \langle a - \hat{a}, (1 - \|T_a\|) \hat{a} \rangle + (1 - \|T_a\|)^2 \|\hat{a}\|^2}{\|a - \hat{a}\|^2}$$

$$= \frac{1}{\|T_a\|^2} \left(1 + \frac{2 \langle a, (1 - \|T_a\|) \hat{a} \rangle - 2(1 - \|T_a\|) + (1 - \|T_a\|)^2}{\|a - \hat{a}\|^2}\right)$$

$$= \frac{1}{\|T_a\|^2} \left(1 + \frac{2 \langle a, (1 - \|T_a\|) \hat{a} \rangle - 1 + \|T_a\|^2}{\|a - \hat{a}\|^2}\right)$$

$$= \frac{1}{\|T_a\|^2} \left(1 + \frac{2 \langle a, (1 - \|T_a\|) \hat{a} \rangle + \frac{1}{s_a} \|P_{a}^{+} C_a a\|^2}{\|a - \hat{a}\|^2}\right). \cdot (22)$$

By Taylor approximation of $\sqrt{1 + x}$ at $x = 0$ we get

$$1 - \|T_a\| = 1 - \sqrt{1 + \frac{1}{s_a} \|P_{a}^{+} C_a a\|^2} = -\frac{1}{2s_a} \|P_{a}^{+} C_a a\|^2 + O\left(\frac{1}{s_a^2} \|P_{a}^{+} C_a a\|^4\right)$$

Plugging this into (22) yields

$$\frac{\|a - \|T_a\| \hat{a}\|^2}{\|T_a\|^2 \|a - \hat{a}\|^2} = \frac{1}{\|T_a\|^2} \left(1 + \frac{(1 - \langle a, \hat{a} \rangle) \|P_{a}^{+} C_a a\|^2 + 2 \langle a, \hat{a} \rangle O\left(\frac{1}{s_a^2} \|P_{a}^{+} C_a a\|^4\right)}{s_a \|a - \hat{a}\|^2}\right) \cdot (23)$$

 Springer
In order to calculate the limit of this expression, we first consider

\[
\lim_{a \to \hat{a}} s_a \|a - \hat{a}\| = \lim_{a \to \hat{a}} \sum_{i=1}^{N} \|a - \hat{a}\| \frac{(a^T y_i)^2}{\|P_{\hat{a}}^i y_i\|} = \lim_{a \to \hat{a}} \frac{\|a - \hat{a}\|}{\|P_{\hat{a}}^i \hat{a}\|} \frac{(a^T y_k)^2}{\|y_k\|}
\]

and since

\[
\lim_{a \to \hat{a}} \|a - \hat{a}\| = \lim_{a \to \hat{a}} \frac{2(1 - \langle a, \hat{a} \rangle)}{2(1 - \langle a, \hat{a} \rangle)(1 + \langle a, \hat{a} \rangle)} = \lim_{a \to \hat{a}} \frac{2}{1 + \langle a, \hat{a} \rangle} = 1,
\]

finally

\[
\lim_{a \to \hat{a}} s_a \|a - \hat{a}\| = \|y_k\|.
\]

The remainder of the Taylor approximation converges to zero as \(\|P_{\hat{a}}^i C_a a\|\) is bounded from above, while \(s_a\) goes to infinity. Together with (23) this gives the limit of the first term,

\[
\lim_{a \to \hat{a}} \frac{1}{\|T_a\|^2} \left(1 + \frac{(1 - \langle a, \hat{a} \rangle)\|P_{\hat{a}}^i C_a a\|^2 + O\left(\frac{1}{s_a} \|P_{\hat{a}}^i C_a a\|^4\right)}{s_a^2 \|a - \hat{a}\|^2}\right) = 1
\]

For the second term in (21) we calculate

\[
L = \lim_{a \to \hat{a}} \frac{2 \left\langle a - \|T_a\| \hat{a}, \frac{1}{s_a} P_{\hat{a}}^i C_a a \right\rangle + \frac{1}{s_a} \|P_{\hat{a}}^i C_a a\|^2}{\|T_a\|^2 \|a - \hat{a}\|^2}
\]

\[
= \lim_{a \to \hat{a}} \frac{-2 \|T_a\| s_a \left\langle \hat{a}, G_{a,k} + a^T y_k \frac{P_{\hat{a}}^i y_k}{\|P_{\hat{a}}^i y_k\|} \right\rangle + \|G_{a,k} + a^T y_k \frac{P_{\hat{a}}^i y_k}{\|P_{\hat{a}}^i y_k\|}\|^2}{\|T_a\|^2 s_a^2 \|a - \hat{a}\|^2}
\]

(24)

Now it is straightforward to check that

\[
\lim_{a \to \hat{a}} \left\langle \sum_{i \neq k} (a^T y_i)^2 \frac{\left\langle \hat{a}, G_{a,k} + a^T y_k \frac{P_{\hat{a}}^i y_k}{\|P_{\hat{a}}^i y_k\|} \right\rangle}{\|P_{\hat{a}}^i y_i\|} \right\rangle = 0
\]

so that by definition of \(s_a\) the term (24) becomes

\[
L := \lim_{a \to \hat{a}} \frac{-2 \|T_a\| \left( a^T y_k \frac{P_{\hat{a}}^i y_k}{\|P_{\hat{a}}^i y_k\|} \right) \left\langle \hat{a}, G_{a,k} + a^T y_k \frac{P_{\hat{a}}^i y_k}{\|P_{\hat{a}}^i y_k\|} \right\rangle + \|G_{a,k} + a^T y_k \frac{P_{\hat{a}}^i y_k}{\|P_{\hat{a}}^i y_k\|}\|^2}{\|T_a\|^2 s_a^2 \|a - \hat{a}\|^2}
\]
Using that $G_{a,k} = P_{\delta}^k C_{a,k} a$, $y_k = \hat{a} \|y_k\|$ and $P_{\delta}^k$ is an orthogonal projector, we can simplify
\[
\frac{(a^T y_k)^2}{\|P_{\delta}^k y_k\|} (\hat{a}, G_{a,k}) = \frac{(a^T y_k)^2}{\|y_k\|} \left( \frac{P_{\delta}^k y_k}{\|P_{\delta}^k y_k\|}, G_{a,k} \right),
\]
so that
\[
L = \lim_{a \to \hat{a}} \frac{\|G_{a,k}\|^2 + (a^T y_k)^2 - 2\|T_a\| (a^T y_k)^3 + a^T y_k \left( \frac{P_{\delta}^k y_k}{\|P_{\delta}^k y_k\|}, G_{a,k} \right) (-2\|T_a\| a^T y_k + 2)}{\|T_a\|^2 \|a - \hat{a}\|^2}.
\]
As $(\frac{P_{\delta}^k y_k}{\|P_{\delta}^k y_k\|}, G_{a,k})$ is bounded, $\lim_{a \to \hat{a}} a^T y_k = \|y_k\|$ and $\lim_{a \to \hat{a}} \|T_a\| = 1$ we get
\[
\lim_{a \to \hat{a}} \frac{2(a - \|T_a\| \hat{a}, \frac{1}{s_a} P_{\delta}^k C_a a) + \frac{1}{s_a} \|P_{\delta}^k C_a a\|^2}{\|T_a\|^2 \|a - \hat{a}\|^2} = \frac{\|G_{\hat{a},k}\|^2 - \|y_k\|^2}{\|y_k\|^2}.
\]
Plugging the results into (21) yields the assertion
\[
\lim_{a \to \hat{a}} \frac{\|T(a) - \hat{a}\|^2}{\|a - \hat{a}\|^2} = 1 + \frac{\|G_{\hat{a},k}\|^2 - \|y_k\|^2}{\|y_k\|^2} = \frac{\|G_{\hat{a},k}\|^2}{\|y_k\|^2}.
\]
\[\square\]

Finally, we need the Kurdyka–Łojasiewicz property of functions [1]: The function $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ with limiting Fréchet subdifferential $\partial f$, see [43], is said to have the Kurdyka–Łojasiewicz (KL) property at $x^* \in \text{dom } \partial f$ if there exist $\eta \in (0, +\infty)$, a neighborhood $U$ of $x^*$ and a continuous concave function $\phi : [0, \eta) \to \mathbb{R}_{\geq 0}$ such that
\begin{enumerate}
  \item $\phi(0) = 0$,
  \item $\phi$ is $C^1$ on $(0, \eta)$,
  \item for all $s \in (0, \eta)$ it holds $\phi'(s) > 0$,
  \item for all $x \in U \cup \{ f(x^*) < f < f(x^*) + \eta \}$, the Kurdyka–Łojasiewicz inequality $\phi'(f(x) - f(x^*)) \text{ dist}(0, \partial f(x)) \geq 1$ holds true.
\end{enumerate}
A proper, lower semi-continuous (lsc) function which satisfies the KL property at each point of dom $\partial f$ is called KL-function. Typical examples of KL functions are semi-algebraic functions. Fundamental works on this subject go back to Łojasiewicz [35] and Kurdyka [24].

The next theorem was proved by Bolte et al. [2, Theorem 2.9].

**Theorem 4.5** Let $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be a KL function. Let $\{x^{(r)}\}_{r \in \mathbb{N}}$ be a sequence which fulfills the following conditions:

**C1.** There exists $K_1 > 0$ such that $f(x^{(r+1)}) - f(x^{(r)}) \leq -K_1 \|x^{(r+1)} - x^{(r)}\|^2$ for every $r \in \mathbb{N}$.
C2. There exists $K_2 > 0$ such that for every $r \in \mathbb{N}$ there exists $w_{r+1} \in \partial f(x^{(r+1)})$ with $\|w_{r+1}\| \leq K_2\|x^{(r+1)} - x^{(r)}\|$, where $\partial f$ denotes the Fréchet limiting subdifferential of $f$ \cite{43}.

C3. There exists a convergent subsequence $\{x^{(r_j)}\}_{j \in \mathbb{N}}$ with limit $\hat{x}$ and $f(x^{(r_j)}) \to f(\hat{x})$.

Then the whole sequence $\{x^{(r)}\}_{r \in \mathbb{N}}$ converges to $\hat{x}$ and $\hat{x}$ is a critical point of $f$ in the sense that $0 \in \partial f(x)$. Moreover the sequence has finite length, i.e.,

$$\sum_{r=0}^{\infty} \|x^{(r+1)} - x^{(r)}\| < \infty.$$

Clearly, if $f$ is differentiable at $x$, then $x$ is a critical point of $f$ if and only if $\nabla f(x) = 0$. Similar arguments as used in the proof of the above theorem lead to the next corollary, see \cite[Corollary 2.7]{2}.

**Corollary 4.6** Let $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be a proper, lsc function which satisfies the KL property at $x^*$. Denote by $U$, $\eta$ and $\phi$ the objects appearing in the definition of the KL function. Let $\delta, \rho > 0$ be such that $B(x^*, \delta) \subset U$ with $\rho \in (0, \delta)$. Consider a finite sequence $x^{(r)}$, $r = 0, \ldots, n$, which satisfies the Conditions C1 and C2 of Theorem 4.5 and additionally

C4. $f(x^*) \leq f(x^{(0)}) < f(x^*) + \eta$,

C5. $\|x^* - x^{(0)}\| + 2\sqrt{\frac{f(x^{(0)}) - f(x^*)}{K_1}} + \frac{K_2}{K_1}\phi(f(x^{(0)}) - f(x^*)) \leq \rho$.

If for all $r = 0, \ldots, n$ it holds

$$x^{(r)} \in B(x^*, \rho) \implies x^{(r+1)} \in B(x^*, \delta) \text{ and } f(x^{(r+1)}) \geq f(x^*),$$

then $x^{(r)} \in B(x^*, \rho)$ for all $r = 0, \ldots, n + 1$.

Now we can prove our main convergence theorem.

**Theorem 4.7** The sequence $\{a^{(r)}\}_r$ generated by Algorithm 1 converges to a critical point of $E$.

**Proof** If the sequence is finite, the claim follows from Lemma 4.1. Assume that the algorithm produces an infinite sequence. Since the sequence $(a^{(r)})_{r \in \mathbb{N}}$ on $\mathbb{S}^{d-1}$ is bounded, there exists a convergent subsequence $(a^{(r_j)})_{j \in \mathbb{N}}$ with $\lim_{j \to \infty} a^{(r_j)} = \hat{a}$. Possibly, there exist multiple accumulation points and we distinguish two cases.

1. First assume that no accumulation point is in the anchor set. It is easy to verify that the function $E$ is semi-algebraic on $\mathbb{R}^d$ and hence fulfills the KL property. We will verify that $(a^{(r)})_r$ fulfills the remaining conditions C1 and C2 from Theorem 4.5. From the proof of Lemma 4.2, Case 1, we get

$$E(a^{(r)}) - E(a^{(r+1)}) \geq \sum_{i=1}^{N} \|\langle a^{(r)}, y_i \rangle a^{(r)} - \langle a^{(r+1)}, y_i \rangle a^{(r+1)}\|^2 / \|P_{a^{(r)}} y_i \|^2.$$
Further, it holds \( \| P_a y_i \| \leq \| y_i \| \leq \max_{i=1,...,N} \| y_i \| < \infty \) and there exists \( m > 0 \) such that

\[
\min_{a \in \text{span}(Y)} \max_{i=1,...,N} |\langle a, y_i \rangle| = \min_{a \in \text{span}(Y)} \| y^T a \|_\infty \geq m.
\]

Using that \( a^{(r)} \in \text{span}(Y) \) for all \( r \) and \( \lim_{r \to \infty} \| a^{(r+1)} - a^{(r)} \| = 0 \) by Lemma 4.3, we can find \( i \in \{1, \ldots, N\} \) such that \( |\langle a^{(r)}, y_i \rangle| > \frac{m}{2} \), \( |\langle a^{(r+1)}, y_i \rangle| > \frac{m}{2} \) and both scalar products have the same sign for \( r \) large enough. Hence we can estimate

\[
E(a^{(r)}) - E(a^{(r+1)}) \geq C \left\| a^{(r)} - a^{(r+1)} \right\|_{\infty}^2, \quad C > 0,
\]

where \( \text{w.l.o.g.} \frac{\langle a^{(r+1)}, y_i \rangle}{\langle a^{(r)}, y_i \rangle} \geq 1 \). Using the projection onto the sphere, we can finally estimate

\[
E(a^{(r)}) - E(a^{(r+1)}) \geq C \| a^{(r)} - a^{(r+1)} \|^2.
\]

Next we check the second condition \( C_2 \). Since \( \lim_{r \to \infty} \| a^{(r+1)} - a^{(r)} \| = 0 \) and none of the \( \pm y_i/\| y_i \|, i = 1, \ldots, N \), is an accumulation point, we can find open balls \( B_i \) around every \( y_i \) such that for all \( r \) large enough the connecting line between iterates fulfills \( a^{(r)} a^{(r+1)} \subset \Omega := \mathbb{R}^d \setminus \bigcup_{i=1}^N B_i \). The function \( E \) is smooth on an open set containing the compact set \( \Omega \) and hence there exists \( C > 0 \) such that

\[
\| \nabla E(a^{(r+1)}) - \nabla E(a^{(r)}) \| \leq C \| a^{(r+1)} - a^{(r)} \|
\]

for all \( r \) large enough. Further, note that the sequence \( s_{a^{(r)}} \) is bounded from above on \( \Omega \) which implies

\[
\| \nabla E(a^{(r+1)}) \| \leq \tilde{C} \left( \| a^{(r+1)} - a^{(r)} \| + \frac{\| \nabla E(a^{(r)}) \|}{s_{a^{(r)}}} \right).
\]

Using the iteration law \( a^{(r+1)} = \Pi_{\mathbb{S}^{d-1}} (a^{(r)} - \frac{\nabla E(a^{(r)})}{s_{a^{(r)}}}) \) together with the fact that \( \nabla E(a^{(r)}) \) is in the tangential plane of \( \mathbb{S}^{d-1} \) at \( a^{(r)} \), we get by the law of sines, see Fig. 3,

\[
\frac{\| a^{(r+1)} - a^{(r)} \| s_{a^{(r)}}}{\| \nabla E(a^{(r)}) \|} \geq \sin \left( \frac{\pi}{2} - \angle(a^{(r)}, a^{(r+1)}) \right),
\]

where the right hand side converges to one since \( \angle(a^{(r)}, a^{(r+1)}) \) gets arbitrary small. Hence, the right hand side is larger than \( \frac{1}{2} \) for \( r \) large enough and we can estimate

\[ \square \]
Fig. 3 Sketch of law of sines for our setting (Color figure online)

\[ \| \nabla E(a^{(r+1)}) \| \leq 3 \tilde{C} \| a^{(r+1)} - a^{(r)} \|. \]

Now, by Theorem 4.5, only one accumulation point exists which is also a critical point.

2. It remains to examine the case that some accumulation point is an anchor point \( \hat{a} \) to the vertices \( y_k, k \in K \).
Assume that there exists another accumulation point. Then, by Lemma 4.3, there exists an accumulation point \( \tilde{a} \) which is not an anchor point. We can find a ball \( B(\tilde{a}, \rho) \) around \( \tilde{a} \) which has positive distance to all anchor points. Next, for all the iterates \( a^{(r)} \in B(\tilde{a}, \frac{R}{2}) \) and \( r \) large enough we can reproduce step one of the proof to show that C1 and C2 are fulfilled. Be the continuity of \( f \) and \( \phi \), see also the proof of [2, Theorem 2.9], we can choose a ball \( B(\tilde{a}, \delta) \subset B(\tilde{a}, \frac{R}{2}) \cap U \) (where \( U \) is from the definition of the KL property), \( \rho \in (0, \delta) \) and a starting iterate \( a^{(r_0)} \in B(\tilde{a}, \rho) \) which satisfies C4 and C5 from Corollary 4.6. Since \( \lim_{r \to \infty} \| a^{(r+1)} - a^{(r)} \| = 0 \) and \( \tilde{a} \) is an accumulation point, we can choose \( r_0 \) such that

\[ a^{(r)} \in B(\tilde{a}, \rho) \implies a^{(r+1)} \in B(\tilde{a}, \delta), \quad E(a^{(r+1)}) \geq E(\tilde{a}) \]

for all \( r \geq r_0 \). Either all iterates after \( a^{(r_0)} \) are in \( B(\tilde{a}, \rho) \) or there is a finite sequence \( a^{(r_0)}, a^{(r_0+1)}, \ldots, a^{(r_n)} \) such that \( a^{(r_n+1)} \) is the first element outside \( B(\tilde{a}, \rho) \). But then, by Corollary 4.6, also the iterate \( a^{(r_n+1)} \) is inside \( B(\tilde{a}, \rho) \) and hence all iterates stay in \( B(\tilde{a}, \rho) \) which is an contradiction. Consequently, the whole sequence converges to the anchor point \( \hat{a} \).

It remains to show that the anchor point is critical. By Lemma 4.4 we know that

\[ \lim_{r \to \infty} \frac{\| T(a^{(r)}) - \hat{a} \|}{\| a^{(r)} - \hat{a} \|} = \frac{\| G_{\hat{a}, K} \|}{\sum_{k \in K} \| y_k \|}. \]

If \( \hat{a} \) is not a critical point, i.e. \( \| G_{\hat{a}, K} \| > \sum_{k \in K} \| y_k \| \), then the sequence cannot converge to \( \hat{a} \), which is a contradiction.

\[ \square \]

At this point it should be mentioned that Algorithm 1 may converge to a local minimum as our functional is non-convex. Performing the algorithm multiple times with random initialization \( a^{(0)} \) and comparing the function values of the results increases
the probability to reach a global minimizer. The number of local minimizers and how pronounced they are, depends on the data. In general, with fewer data points and more extreme outliers, we tend to get more pronounced local minima. However, in most applications and also in the numerical part of this paper, this is not an issue since a high number of data points is available.

5 Remarks on the Offset

Finally, we want to address briefly the issue of choosing a suitable offset for the robust PCA model. Although we cannot offer a solution at this point, we want to touch on some aspects of it, as it is not satisfactorily discussed in the existing literature to the best of our knowledge. As already mentioned in the introduction, in classical PCA, solving

$$\arg \min_{A \in \mathbb{S}_d, K, b \in \mathbb{R}^d} \sum_{i=1}^{N} \min_{t \in \mathbb{R}^K} \|A t + b - x_i\|^2 = \arg \min_{A \in \mathbb{S}_d, K, b \in \mathbb{R}^d} \sum_{i=1}^{N} \|P_A^t(x_i - b)\|^2,$$

leads to the unique affine subspace

$$\{\hat{A}t + \hat{b} : t \in \mathbb{R}^K\},$$

where $\hat{b} \in \mathbb{R}^d$ can be chosen as mean (bias) $\bar{b} := \frac{1}{N}(x_1 + \ldots + x_N)$ of the data. For the robust setting, we have assumed so far that the offset $\hat{b}$ is given, e.g., as geometric median of the data. However, for the problem

$$\arg \min_{A \in \mathbb{S}_d, K, b \in \mathbb{R}^d} \sum_{i=1}^{N} \min_{t \in \mathbb{R}^K} \|A t + b - x_i\| = \arg \min_{A \in \mathbb{S}_d, K, b \in \mathbb{R}^d} \sum_{i=1}^{N} \|P_A^t(x_i - b)\|,$$  \hspace{1cm} (25)

the geometric median is in general not a minimizer with respect to $b$.

Lemma 5.1 Let $x_i \in \mathbb{R}^2$, $i = 1, \ldots, N$. Then there exists a minimizing pair

$$(\hat{a}, \hat{b}) \in \arg \min_{a \in \mathbb{S}^1, b \in \mathbb{R}^2} \sum_{i=1}^{N} \min_{t \in \mathbb{R}} \|at + b - x_i\|$$

such that the line $g(t) := \hat{a}t + \hat{b}$ passes through two of the points. If $N$ is odd, then the minimizing line always passes through two points.

Proof Assume that $g$ is an optimal line which does not go through any of the points. Let $N_l$, resp. $N_r$ be the number of points on the left, resp., right hand side of $g$. Then, shifting the line by $\delta > 0$ into the direction of the left, resp. right nearest point changes the distance sum by $(N_r - N_l)\delta$, resp. $(N_l - N_r)\delta$. If $N_l \neq N_r$, then one of the new distance sums becomes smaller than the original minimal one. Hence, $N_l = N_r$, so that
one point has to be on the line if \( N \) is odd and there is a line with smallest distance sum going through one point if \( N \) is even. W.l.o.g., let \( g \) go through \( x_N \). Then, choosing \( b = x_N \) we have to show that \( g \) goes through a second point. Taking polar coordinates \( y_i := x_i - x_N = c_i e^{i\gamma_i}, i \in \{1, \ldots, N - 1\} \), and \( a = e^{i\alpha} \), the distance sum becomes

\[
\sum_{i=1}^{N-1} |t_i e^{i\alpha} - c_i e^{i\gamma_i}| = \sum_{i=1}^{N-1} \min_{t_i} |t_i - c_i e^{(\gamma_i - \alpha)}| = \sum_{i=1}^{N-1} |c_i \sin(\alpha - \gamma_i)| =: \varphi(\alpha).
\]

If \( y_i \notin g \) for all \( i = 1, \ldots, N - 1 \), then \( \varphi \) is smooth and

\[
\varphi''(\alpha) = -\varphi(\alpha) < 0
\]

so that \( \alpha \) cannot be a local minimizer. Consequently, at least one more \( x_i \) must lie on \( g \).

Using the decomposition of \( A \in S_{d,d-1} \) into Givens rotation matrices, the claim can be generalized to hyperplanes of dimension \( d - 1 \) having minimal Euclidean distance from data in \( \mathbb{R}^d \), \( d \geq 2 \), see [54]. However, it would be interesting if in this case also \( d - 1 \) data points can lie within the minimizing hyperplane instead of just two of them.

Based on the lemma, the following example shows that the geometric median is in general not in the solution set of (25).

**Example 5.2** Let \( x_i \in \mathbb{R}^2, i = 1, 2, 3 \) span a triangle with sides \( s_1 = \|x_2 - x_3\|, s_2 = \|x_1 - x_3\|, s_3 = \|x_1 - x_2\| \), where \( s_1 \leq s_2 < s_3 \) and angles smaller than 120°. By Lemma 5.1, the line having minimal Euclidean distance from the three points has to go through two points. Since the height \( h_i \) at side \( s_i, i = 1, 2, 3 \), fulfills

\[
h_i = \frac{2}{s_i} (s(s - s_1)(s - s_2)(s - s_3))^{\frac{1}{2}}, \quad s := \frac{1}{3}(s_1 + s_2 + s_3),
\]

we conclude that the line must go through \( x_1 \) and \( x_2 \) and has distance \( h_3 \) from \( x_3 \). On the other hand, it is easy to check (and known) that the geometric median of the data points is the so-called Steiner point from which the points can be seen under an angle of 120°. Clearly, the minimizing line does not pass the Steiner point.

### 6 Numerical Examples

In this section, we present various numerical examples. We start with numerical results on the convergence behavior of the algorithm in Sect. 6.1 and show examples in image processing in Sect. 6.2.

#### 6.1 Convergence Speed

We demonstrate the convergence of the algorithm for the data set in Fig. 4a consisting of 100 data points which are normally distributed with covariance matrix \( \begin{pmatrix} 0.1 & 0 \\ 0 & 0.5 \end{pmatrix} \) so that

\[\mathcal{S}_{100,50}.\]
they lie near a vertical line through the origin, and 50 outliers which are isotropically normally distributed around the origin with variance 10. For centering, we choose $b$ as the geometric median. From the proof of Lemma 5.1 we know that any local minimizer $a^*$ of $E$ is in the anchor set $A$, so that we can find an exact minimizer (up to machine precision) by iterating until we arrive in an anchor direction which fulfills the third optimality condition in Theorem 3.2. Algorithm 1 is initialized with starting point $1/\sqrt{2}(1, 1)^T$ and converges to a local minimizer $a^* \in A$ which is close to the second unit vector. The difference $E(a^{(r)}) - E(a^*)$ of the function value of the iterates to the minimum $E(a^*)$ and the distance $\|a^{(r)} - a^*\|$ of the iterates to the minimizer can be found in Fig. 4b and c, respectively. Both graphs show a similar behavior. Until about iteration 40, the global function landscape dominates the convergence behavior. Starting at iteration 40, the iterates are already close to the minimizing anchor direction and the graphs suggest an exponential convergence rate in that region.

**6.2 Image Processing Examples**

In this section, we present two application examples and compare the results of (6) with other methods. In particular, we compare the results of our Algorithm 1 (with the geometric median as $b$) with standard PCA and the following methods:

(i) **PC-L1**: the greedy algorithm for minimizing (7) proposed by Kwak [25]. As $b$ we used the geometric median computed by Weiszfeld’s Algorithm 1.

(ii) **TRPCA**: the trimmed PCA of Podosinnikova et al. [51] with default parameters, i.e., the lower bound on the number of true observations is set to $N/2$ and the number of random restarts is 10. Here, $b$ equals the mean of a certain subset of the given data determined within the algorithm.

**6.2.1 Image Sequence with Slightly Varying Background**

We consider an image sequence with slightly varying background as it was used for object detection in various papers. The water front data set, see Fig. 5, was originally considered in [34]. It was used for performance comparisons with several robust PCA
methods including those of Candes et al. [5] in the context of object detection in [51], where TRPCA outperformed the other methods. The data set consists of 633 frames of size $128 \times 160$ of a scenery with water and grass as background. Beginning with frame 481 a person walks into the scene, which we consider as “outlier” frames in the data set. We aim to detect the frames with the person present, and then to separate background (scenery) and foreground (person). It turns out this can be achieved simply by thresholding the Euclidean distances of the vectorized data $x_i \in \mathbb{R}^{20480}$ to their geometric median. The frames with the person in them can be detected from the histogram in Fig. 5c. More precisely, all frames with distance larger than 6 can be considered as outliers which exactly matches the frames with the person present. The foreground in these images can then be extracted as the difference image to the geometric median and subsequent pixelwise thresholding. The difference image for one frame is given in Fig. 5d.

In order to make the task more challenging and simulate a gradual change in lighting conditions, we alter the data as follows. Given the points $x_i, i = 1, \ldots, 633$, we created new data

$$\tilde{x}_i := \frac{3}{4} x_i + \frac{i}{8 \cdot 633} (1 + x_i). \quad (26)$$

Here, outlier frames cannot be found by the previous method since the distance of the frames from their geometric median varies by construction, see Fig. 6 (left). But a model with line fitting, i.e. with $K = 1$, is suitable by the construction of the data. Fig. 6 depicts the histogram of the distances of the frames $\tilde{x}_i$ from the line generated by the standard PCA and by the residual minimizing robust PCA, respectively. The outliers can be better separated by the residual minimizing robust PCA as the frames belonging to the peak between 4 and 5 are less likely to be wrongfully mislabeled as outliers.

Finally, Fig. 7 shows the foreground–background separation in frame $i = 580$ by various methods. We show the projected data (background) $x_{i,\text{rec}} = a_1 a_1^T (x_i - b) + b$ (left) and the residual (person) $x_{i,\text{res}} = x_i - x_{i,\text{rec}}, i = 580$. In the background and foreground of standard PCA, artifacts can be clearly seen at positions where the person rests for a longer time. The PCA-L1 and our approach appear to be more robust here, but the artifacts are more pronounced in the PCA-L1. TRPCA with several restarts gives the best results.
Fig. 6 Distance histograms for the water front data set modified by (26). Left: distance of the frames to their median. Middle: distance of the frames to the line fitted with standard PCA. Right: distance of the frames to the line fitted with our approach.

Fig. 7 Projections on one dimensional subspace (background) and residuals (foreground) of different approaches for the water front data set. Top left: standard PCA. Top right: PCA-L1. Bottom left: TRPCA. Bottom right: our approach.

6.2.2 Face Reconstruction

For images of faces in the same pose but with different lighting, it may be assumed that they lie in a low dimensional subspace [8]. Thus standard PCA is a suitable tool to reduce the dimensionality of such data for classification and other tasks. In practice, however, some of the face images may be occluded resulting in outliers within the data. Here, robust PCA methods appear to be more useful. We test the performance of various approaches with the cropped Extended Yale Face database B [26]. There are 58 images of size $168 \times 192$ of which 12 were altered with a $50 \times 50$ square patch of noise at a random position, see the left image in Fig. 8 for an example.

For noiseless face image data, standard PCA projection onto a subspace of dimension $K = 5$ gives good results as shown on the right of Fig. 8.

In Fig. 9 the projection of the noisy data on the subspace obtained by various approaches are shown. As expected the noisy patches can be clearly seen in the reconstructions by standard PCA. Surprisingly, the result of PCA-L1 looks worse than standard PCA, as the influence of the noisy patches is even worse. The results of TRPCA with several restarts are very similar to those of the standard PCA of the noiseless data except of the right eye in the second image which appears to be too
Fig. 8  Standard PCA with $K = 5$ for noiseless face data. Left two images: Different face images, where a possible corruption is exemplified by the first one. Right two images: projections of the noiseless images onto the 5 dimensional subspace calculated by standard PCA computed from the noiseless data set.

Fig. 9  PCA methods for the noisy face data. Top left two images: standard PCA. Top right two images: PCA-L1. Bottom left two images: TRPCA. Bottom right two images: our approach.

dark. This suggests that the algorithm successfully excluded the outliers and calculated the principal components from a part of the noiseless data. However, it should be mentioned that TRPCA sometimes fails to detect the outliers as it depends on the initial values of the random restarts. The results of residual minimizing robust PCA demonstrate the robustness of the method to outliers although slight artifacts are still visible.

7 Conclusions

We proposed a Weiszfeld-like algorithm to address the robust PCA problem arising from a minimal distance function of lines from points and gave a circumvent convergence analysis of the algorithm. Further extensions of our findings are possible such as the treatment of robust independent component analysis (ICA) and PCA on manifolds, see, e.g., [10,11,17,49,50,56]. Another modification of PCA, the so-called
sparse PCA couples the data term (1) with a sparsity term for $A \in \mathbb{R}^{d,K}$, see, e.g., [12,36] and could be considered under the robustness point of view.

Acknowledgements Funding by the German Research Foundation (DFG) within the Research Training Group 1932, project area P3, is gratefully acknowledged.

Appendix A: One-Sided Derivatives and Minimizers on Embedded Manifolds

The one-sided directional derivative of a function $f : \mathbb{R}^d \to \mathbb{R}$, $d \in \mathbb{N}$, at a point $x \in \mathbb{R}^d$ in direction $h \in \mathbb{R}^d$ is defined by

$$Df(x; h) := \lim_{\alpha \downarrow 0} \frac{f(x + \alpha h) - f(x)}{\alpha}.$$  

Restricting $f$ to a submanifold $M \subseteq \mathbb{R}^d$, we can restrict our considerations to $h \in T_x M$. Recall that $M \subseteq \mathbb{R}^d$ is an $m$-dimensional submanifold of $\mathbb{R}^d$ if for each point $x \in M$ there exists an open neighborhood $U \subseteq \mathbb{R}^d$ as well as an open set $\Omega \subseteq \mathbb{R}^m$ and a so-called parametrization $\varphi \in C^1(\Omega, \mathbb{R}^d)$ of $M$ with the properties

(i) $\varphi(\Omega) = M \cap U$,
(ii) $\varphi^{-1} : M \cap U \to \Omega$ is surjective and continuous, and
(iii) $D\varphi(x)$ has full rank $m$ for all $x \in \Omega$.

To establish the relation between one-sided directional derivatives and local minima of functions on manifolds we need the following lemma. A proof can be found in [45, Lemma B.1].

Lemma A.1 Let $M \subseteq \mathbb{R}^d$ be an $m$-dimensional submanifold of $\mathbb{R}^d$. Then the tangent space $T_x M$ and the tangent cone

$$T_x M := \left\{ u \in \mathbb{R}^d : \exists \text{ sequence } (x_k)_{k \in \mathbb{N}} \subset M \setminus \{x\} \text{ with } x_k \to x \text{ s.t. } \frac{x_k - x}{\|x_k - x\|} \rightarrow \frac{u}{\|u\|} \right\} \cup \{0\}$$

coincide.

The following theorem gives a general necessary and sufficient condition for local minimizers of Lipschitz continuous functions on embedded manifolds using the notion of one-sided derivatives. For the Euclidean setting $M = \mathbb{R}^d$, the first relation of the proposition is trivially fulfilled for any function $f : \mathbb{R}^d \to \mathbb{R}$, while a proof of the sufficient minimality condition in the second part was given in [4]. Moreover, the authors of [4] gave an example that Lipschitz continuity in the second part cannot be weakened to just continuity.

Theorem A.2 Let $M \subseteq \mathbb{R}^d$ be an $m$-dimensional submanifold of $\mathbb{R}^d$ and $f : \mathbb{R}^d \to \mathbb{R}$ a locally Lipschitz continuous function. Then the following holds true:
1. If \( \hat{x} \in \mathcal{M} \) is a local minimizer of \( f \) on \( \mathcal{M} \), then \( Df(\hat{x}; h) \geq 0 \) for all \( h \in T_{\hat{x}}\mathcal{M} \).

2. If \( Df(\hat{x}; h) > 0 \) for all \( h \in T_{\hat{x}}\mathcal{M} \setminus \{0\} \), then \( \hat{x} \) is a strict local minimizer of \( f \) on \( \mathcal{M} \).

A proof can be found in [45, Thm. 6.1] along with an example which demonstrates the necessity of the Lipschitz continuity of \( f \) in the manifold setting in the first part of the theorem. Furthermore, note that \( Df(\hat{x}; h) \geq 0 \) for all \( h \in T_{\hat{x}}\mathcal{M} \setminus \{0\} \) does not imply that \( \hat{x} \) is a local minimizer of \( f \) on \( \mathcal{M} \).

References

1. Attouch, H., Bolte, J., Redont, P., Soubeyran, A.: Proximal alternating minimization and projection methods for nonconvex problems: an approach based on the Kurdyka-Łojasiewicz inequality. Math. Oper. Res. 35(2), 438–457 (2010)
2. Attouch, H., Bolte, J., Svaiter, B.F.: Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward-backward splitting, and regularized Gauss-Seidel methods. Math. Program. 137(1–2, Ser. A), 91–129 (2013)
3. Beck, A., Sabach, S.: Weiszfeld’s method: old and new results. J. Optim. Theory Appl. 164(1), 1–40 (2015)
4. Ben-Tal, A., Zowe, J.: Directional derivatives in nonsmooth optimization. J. Optim. Theory Appl. 47(4), 483–490 (1985)
5. Candès, E.J., Li, X., Ma, Y., Wright, J.: Robust principal component analysis? J. ACM 58(3), Art. 11 (2011)
6. Chouzenoux, E., Idier, J., Moussaoui, S.: A majorize-minimize strategy for subspace optimization applied to image restoration. IEEE Trans. Image Process. 20(6), 1517–1528 (2011)
7. Ding, C., Zhou, D., He, X., Zha, H.: \( R_1 \)-PCA: Rotational invariant \( L_1 \)-norm principal component analysis for robust subspace factorization. In: Proceedings of the 23rd International Conference on Machine Learning, pp. 281–288. ACM (2006)
8. Epstein, R., Hallinan, P., Yuille, A.: 5 ± 2 eigenimages suffice: an empirical investigation of low-dimensional lighting models. In: IEEE Workshop on Physics-Based Vision, pp. 108–116 (1995)
9. Fischler, M.A., Bolles, R.C.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. In: Readings in Computer Vision, pp. 726–740. Elsevier (1987)
10. Fletcher, P.T., Joshi, S.: Principal geodesic analysis on symmetric spaces: statistics of diffusion tensors. In: Computer Vision and Mathematical Methods in Medical and Biomedical Image Analysis, pp. 87–98. Springer (2004)
11. Fletcher, P.T., Lu, C., Pizer, S.M., Joshi, S.: Principal geodesic analysis for the study of nonlinear statistics of shape. IEEE Trans. Med. Imaging 23(8), 995–1005 (2004)
12. Hager, W.W., Phan, D.T., Zhu, J.: Projection algorithms for nonconvex minimization with application to sparse principal component analysis. J. Glob. Optim. 65(4), 657–676 (2016)
13. Hastie, T., Tibshirani, R., Friedman, J.: The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer Series in Statistics. Springer, New York (2001)
14. Hauberg, S., Feragen, A., Black, M.J.: Grassmann averages for scalable robust PCA. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3810–3817 (2014)
15. Huber, P.J.: Robust Statistics. Wiley Series in Probability and Mathematical Statistics. Wiley, New York (1981)
16. Huber, P.J., Ronchetti, E.M.: Robust Statistics, 2nd edn. Wiley Series in Probability and Statistics. Wiley, New York (2009)
17. Huckemann, S., Ziezold, H.: Principal component analysis for Riemannian manifolds with an application to triangular shape spaces. Adv. Appl. Probab. 38(2), 299–319 (2006)
18. Katz, I.N.: Local convergence in Fermat’s problem. Math. Program. 6(1), 89–104 (1974)
19. Ke, Q., Kanade, T.: Robust \( \ell_1 \) norm factorization in the presence of outliers and missing data by alternative convex programming. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005 (CVPR 2005), vol. 1, pp. 739–746. IEEE (2005)
20. Keeling, S.L., Kunisch, K.: Robust $\ell_1$ approaches to computing the geometric median and principal and independent components. J. Math. Imaging Vis. 56(1), 99–124 (2016)
21. Kriegel, H.P., Kröger, P., Schubert, E., Zimek, A.: A general framework for increasing the robustness of PCA-based correlation clustering algorithms. In: Scientific and Statistical Database Management. Lecture Notes in Computer Science. 5069, pp. 418–435 (2008)
22. Kuhn, H.W.: A note on Fermat’s problem. Math. Program. 4, 98–107 (1973)
23. Kriegel, H.P., Kröger, P., Schubert, E., Zimek, A.: A general framework for increasing the robustness of PCA-based correlation clustering algorithms. In: Scientific and Statistical Database Management. Lecture Notes in Computer Science. 5069, pp. 418–435 (2008)
24. Kurdyka, K.: On gradients of functions definable in o-minimal structures. Annales de l’Institut Fourier, vol. 48, pp. 769–783 (1998)
25. Kwak, N.: Principal component analysis based on $\ell_1$-norm maximization. IEEE Trans. Pattern Anal. Mach. Intell. 30(9), 1672–1680 (2008)
26. Lee, K.-C., Ho, J., Kriegman, D.J.: Acquiring linear subspaces for face recognition under variable lighting. IEEE Trans. Pattern Anal. Mach. Intell. 5, 684–698 (2005)
27. Lerman, G., Maunu, T.: Fast, robust and non-convex subspace recovery. Inf. Inference 7(2), 277–336 (2017)
28. Lerman, G., Maunu, T.: An overview of robust subspace recovery. Proc. IEEE 106(8), 1380–1410 (2018)
29. Lerman, G., Zhang, T.: Robust recovery of multiple subspaces by geometric $l_p$ minimization. Ann. Stat. 39(5), 2686–2715 (2011)
30. Lerman, G., Zhang, T.: $l_p$-recovery of the most significant subspace among multiple subspaces with outliers. Constr. Approx. 40(3), 329–385 (2014)
31. Lerman, G., McCoy, M., Tropp, J.A., Zhang, T.: Robust computation of linear models by convex relaxation. Found. Comput. Math. 15(1), 363–410 (2015)
32. Leroy, A.M., Rousseeuw, P.J.: Robust Regression and Outlier Detection. Wiley Series in Probability and Mathematical Statistics. Wiley, Chichester (1987)
33. Li, G., Chen, Z.: Projected-pursuit approach to robust dispersion matrices and principal components: Primary theory and Monte-Carlo. J. Am. Stat. Soc. 80, 759–766 (1985)
34. Li, L., Huang, W., Gu, I.Y.-H., Tian, Q.: Statistical modeling of complex backgrounds for foreground object detection. IEEE Trans. Image Process. 13(11), 1459–1472 (2004)
35. Łojasiewicz, S.: Une propriété topologique des sous-ensembles analytiques réels. In: Les Équations aux Dérivées Partielles (Paris, 1962), pp. 87–89. Éditions du Centre National de la Recherche Scientifique, Paris (1963)
36. Luss, R., Teboulle, M.: Conditional gradient algorithms for rank-one matrix approximations with a sparsity constraint. SIAM Rev. 55(1), 65–98 (2013)
37. Markopoulos, P.P., Karystinos, G.N., Pados, D.A.: Optimal algorithms for $\ell_1$-subspace signal processing. IEEE Trans. Signal Process. 62(19), 5046–5058 (2014)
38. Markopoulos, P.P., Kundo, S., Chamadia, S., Pados, D.A.: Efficient 11-norm principal-component analysis via bit flipping. IEEE Trans. Signal Process. 65(16), 4252–4264 (2017)
39. Maronna, R.A., Martin, R.D., Yohai, V.J.: Robust Statistics: Theory and Methods. Wiley Series in Probability and Statistics. Wiley, Chichester (2006)
40. Massart, D.L., Kaufman, L., Rousseeuw, P.J., Leroy, A.: Least median of squares: a robust method for outlier and model error detection in regression and calibration. Anal. Chim. Acta 187, 171–179 (1986)
41. Maunu, T., Zhang, T., Lerman, G.: A well-tempered landscape for non-convex robust subspace recovery (2017). arXiv:1706.03896
42. McCoy, M., Tropp, J.A.: Two proposals for robust PCA using semidefinite programming. Electron. J. Stat. 5, 1123–1160 (2011)
43. Mordukhovich, B., Nam, N.M., Yen, N.D.: Fréchet subdifferential calculus and optimality conditions in nondifferentiable programming. Optimization 55(5–6), 685–708 (2006)
44. Neumayer, S., Nimmer, M., Steidl, G., Stephani, H.: On a projected Weiszfeld algorithm. In: Lauze, F., Dong, Y., Dahl, A.B. (eds.) Scale Space and Variational Methods in Computer Vision. Lecture Notes in Computer Science, vol. 10302, pp. 486–497. Springer, New York (2017)
45. Neumayer, S., Nimmer, M., Setzer, S., Steidl, G.: On the rotational invariant $L_1$-norm PCA (2019). http://arxiv.org/pdf/1902.03840v1
46. Nie, F., Huang, H., Ding, C., Luo, D., Wang, H.: Robust principal component analysis with non-greedy $\ell_1$-norm maximization. In: IJCAI Proceedings-International Joint Conference on Artificial Intelligence, vol. 22, pp. 1433 (2011)
47. Ostresh Jr., L.M.: On the convergence of a class of iterative methods for solving the Weber location problem. Oper. Res. 26(4), 597–609 (1978)
48. Pearson, K.: On lines and planes of closest fit to systems of points in space. Philos. Mag. 2(11), 559–572 (1901)
49. Pennec, X.: Barycentric subspace analysis on manifolds (2016). arXiv:1607.02833
50. Pennec, X.: Sample-limited $l_p$ barycentric subspace analysis on constant curvature spaces. In: International Conference on Geometric Science of Information, pp. 20–28. Springer (2017)
51. Podosinnikova, A., Setzer, S., Hein, M.: Robust PCA: Optimization of the robust reconstruction error over the Stiefel manifold. In: German Conference on Pattern Recognition, pp. 121–131. Springer (2014)
52. Rockafellar, R.T.: Convex Analysis. Princeton University Press, Princeton (1970)
53. Rousseeuw, P.J., Leroy, A.M.: Robust Regression and Outlier Detection. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. Wiley, New York (1987)
54. Schneck, G.: Robust principal component analysis. Bachelor Thesis, TU Kaiserslautern (2018)
55. Setzer, S., Steidl, G., Teuber, T.: On vector and matrix median computation. J. Comput. Appl. Math. 236(8), 2200–2222 (2012)
56. Sommer, S., Lauze, F., Nielsen, M.: Optimization over geodesics for exact principal component analysis. Adv. Comput. Math. 40, 283–313 (2014)
57. Vardi, Y., Zhang, C.H.: A modified Weiszfeld algorithm for the Fermat-Weber location. Math. Program. 90, 559–566 (2001)
58. Vazsonyi, A.: Pure mathematics and Weiszfeld algorithm. Decis. Line 33(3), 12–13 (2002)
59. Weiszfeld, E.: Sur le point pour lequel les sommes des distances de $n$ points donnés et minimum. Tôhoku Math. J. 43, 355–386 (1937)
60. Xu, H., Caramanis, C., Sanghavi, S.: Robust PCA via outlier pursuit. IEEE Trans. Inf. Theory 58(3), 3047–3064 (2012)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.