



Nonsmooth rank-one matrix factorization landscape

Cédric Josz¹ · Lexiao Lai¹

Received: 8 September 2020 / Accepted: 22 October 2021 / Published online: 5 November 2021
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract

We provide the first positive result on the nonsmooth optimization landscape of robust principal component analysis, to the best of our knowledge. It is the object of several conjectures and remains mostly uncharted territory. We identify a necessary and sufficient condition for the absence of spurious local minima in the rank-one case. Our proof exploits the subdifferential regularity of the objective function in order to eliminate the existence quantifier from the first-order optimality condition known as Fermat's rule.

1 Introduction

Low-rank matrix factorization has received significant attention in the last decade, initiated by several seminal papers [7,8,18,29]. It has various applications in data science and machine learning, which include principal component analysis [4,7], facial recognition [7], video surveillance [5,16], recommender systems [22] and natural language processing [23]. The number of survey papers [10–12,20,26,28] on the subject in the last three years is a testament to the amount of research it has spawned.

While the exact recovery of a low-rank matrix via convex optimization is well understood [7,9], its non-convex counterpart

$$\inf_{(X,Y) \in \mathbb{R}^{m \times r} \times \mathbb{R}^{n \times r}} \|XY^T - M\|_1 \quad (1)$$

remains elusive, where $M \in \mathbb{R}^{m \times n}$ and $\|A\|_1 := \sum_{i=1}^m \sum_{j=1}^n |A_{ij}|$ for any $A \in \mathbb{R}^{m \times n}$. Note that minimizing the Frobenius norm squared instead yields approximate recovery and is better understood [2,14,32]. For a general data matrix $M \in \mathbb{R}^{m \times n}$, finding a

✉ Cédric Josz
cj2638@columbia.edu

Lexiao Lai
ll3352@columbia.edu

¹ Department of Industrial Engineering and Operations Research, Columbia University, Columbia, USA

global minimum to (1) is known to be an NP-hard problem [17, Theorem 3]. However if M is itself a low-rank matrix, then it is conjectured that the objective function is sharp [6] in a neighborhood of the global minima (see [25, Equation (2.11)] and [10, Conjecture 8.7]). At best, this would imply convergence guarantees for local search algorithms when initialized in a neighborhood of the global minima. In order to prove convergence to a global minimum from any random initial point, as observed in [1, 15, 21, 24], it is necessary to analyze the landscape. We do so in the rank-one case and obtain the following theorem.

Theorem 1 *Assume that $\text{rank}(M) \leq 1$. Then the function defined from $\mathbb{R}^m \times \mathbb{R}^n$ to \mathbb{R} by*

$$f(x, y) := \sum_{i=1}^m \sum_{j=1}^n |x_i y_j - M_{ij}| \quad (2)$$

has no spurious local minima if and only if none or all of the entries of M are equal to zero.

Spurious local minima are defined as local minima that are not global minima. The analysis of the landscape of the nonconvex and nonsmooth objective function f poses a significant challenge. We next explain why several standard tools in optimization fall short of providing a way forward. As a first approach, one could consider the following equivalent constrained optimization problem

$$\min_{(x, y, z) \in \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^{m \times n}} \sum_{i=1}^m \sum_{j=1}^n z_{ij} \quad (3a)$$

$$\text{subject to } -z_{ij} \leq x_i y_j - M_{ij} \leq z_{ij}, \quad i = 1, \dots, m, \quad j = 1, \dots, n. \quad (3b)$$

The advantage of such a formulation is, of course, the fact that it is smooth. One may then invoke the Karush-Kuhn-Tucker conditions. However, it is not clear how one should establish constraint qualification. Observe that when $M = (0, 1)^T$, the gradients of the active constraints are linearly dependent at the origin, which is a saddle point. Besides, the Karush-Kuhn-Tucker conditions comprise an exponential numbers of cases as a function of the dimensions m and n , and there is seemingly no way to avoid treating each case separately.

As a second approach, one could try to approximate the nonconvex and nonsmooth objective function (2) by a smooth function. This viewpoint underlies many approaches for solving nonsmooth optimization problems, notably proximal methods [3] and Nesterov's smoothing technique [27]. Along these lines, the first author and co-authors recently established that the uniform limit of a sequence of functions which are devoid of spurious local minima is itself devoid of spurious strict local minima [21, Proposition 2.7]. A spurious strict local minimum is a strict local minimum that is not a global minimum. In our setting, a natural candidate of smooth approximations is given by

the ℓ_p -norm, where $p > 1$:

$$f_p(x, y) := \sum_{i=1}^m \sum_{j=1}^n |x_i y_j - M_{ij}|^p. \quad (4)$$

By letting p converge to 1 from above, f_p converges uniformly to f on any compact set. However, this approach fails to deliver any meaningful results for the problem at hand. Indeed, we know full well that the objective function (2) is devoid of spurious strict local minima. This follows from the invariance $f(\theta x, \theta^{-1} y) = f(x, y)$ for all $\theta \in \mathbb{R} \setminus \{0\}$, whereby any point different from the origin, itself a saddle point, has neighbors with identical function values.

A third approach, and the one that we will follow in this paper, is to use a generalized Fermat rule [13, 2.3.2 Proposition] which holds for locally Lipschitz functions. It states that if $(x, y) \in \mathbb{R}^m \times \mathbb{R}^n$ is a local minimum of f , then $0 \in \partial f(x, y)$ where ∂f is the Clarke subdifferential [13, pp. 25-27]. We will refer to a point satisfying this set inclusion as a *critical* point. The Clarke subdifferential of f is defined for all $(x, y) \in \mathbb{R}^m \times \mathbb{R}^n$ by

$$\partial f(x, y) := \{(s, t) \in \mathbb{R}^m \times \mathbb{R}^n \mid f^\circ(x, y; h, k) \geq s^T h + t^T k, \quad \forall (h, k) \in \mathbb{R}^m \times \mathbb{R}^n\} \quad (5)$$

where one uses the generalized directional derivative

$$f^\circ(x, y; h, k) := \limsup_{\substack{(\bar{x}, \bar{y}) \rightarrow (x, y) \\ t \searrow 0}} \frac{f(\bar{x} + th, \bar{y} + tk) - f(\bar{x}, \bar{y})}{t}. \quad (6)$$

By virtue of [13, 2.3.10 Chain Rule II], we have the simpler form

$$\partial f(x, y) = \left\{ \begin{pmatrix} \Lambda y \\ \Lambda^T x \end{pmatrix} \mid \Lambda \in \text{sign}(xy^T - M) \right\} \quad (7)$$

where

$$\text{sign}(t) := \begin{cases} -1 & \text{if } t < 0, \\ [-1, 1] & \text{if } t = 0, \\ 1 & \text{if } t > 0. \end{cases} \quad (8)$$

Above, $[-1, 1]$ stands for the interval in \mathbb{R} that includes -1 and 1 , and the sign function applies to matrices term by term. It follows from (7) that a point is critical if and only if

$$\exists \Lambda \in \text{sign}(xy^T - M) : \quad \Lambda y = 0 \quad \text{and} \quad \Lambda^T x = 0. \quad (9)$$

In order to study the variation of f in the vicinity of its critical points, we seek to eliminate the quantifier \exists from (9). The Tarski-Seidenberg theorem [30, 31] guarantees

```

In[1]:= M1 = 0; M2 = 1;

In[2]:= Reduce[
  Exists[{L1, L2},
    ((L1 == 1 && x1*y - M1 > 0) || (L1 == -1 && x1*y - M1 < 0) ||
     (L1 ∈ Interval[{-1, 1}] && x1*y - M1 == 0)) &&
    ((L2 == 1 && x2*y - M2 > 0) || (L2 == -1 && x2*y - M2 < 0) ||
     (L2 ∈ Interval[{-1, 1}] && x2*y - M2 == 0)) && (L1*y == 0) && (L2*y == 0) &&
    (L1*x1 + L2*x2 == 0)], Reals]

Out[2]= 
$$\left( x2 < 0 \& \left( (x1 \leq x2 \&& y = 0) \mid\mid \left( x1 = 0 \&& y = \frac{1}{x2} \right) \mid\mid (x1 \geq -x2 \&& y = 0) \right) \mid\mid \right. \\ \left. (x2 = 0 \&& y = 0) \mid\mid \left( x2 > 0 \& \left( (x1 \leq -x2 \&& y = 0) \mid\mid \left( x1 = 0 \&& y = \frac{1}{x2} \right) \mid\mid (x1 \geq x2 \&& y = 0) \right) \right) \right)$$


```

Fig. 1 Elimination of quantifier with Wolfram Mathematica 12

that this is possible for any $M \in \mathbb{R}^{m \times n}$, as long as m and n are fixed. One can thus recover the set of critical points when $M = (0, 1)^T$ using commercial software, as can be seen in Fig. 1. However, no such result is guaranteed for any sizes m and n . In order to overcome this, we analyze step functions that arise in the partial Clarke subdifferentials of f in Lemma 1. (The partial Clarke subdifferential $\partial_{x_1} f(x, y)$ is, by definition [13, p. 48], the Clarke subdifferential of $f(\cdot, x_2, \dots, x_m, y)$ at x_1 .) We are thus able to eliminate the quantifier from (9) with Lemma 2. The new description of the critical points is much more amenable to analysis. It yields a full characterization of the landscape of f in Proposition 1, from which Theorem 1 is deduced.

The reason why we can rely on partial Clarke subdifferentials to analyze the landscape is because the objective function f is subdifferentially regular [13, 2.3.4 Definition]. By definition, this means that its generalized directional derivative agrees with the classical directional derivative, that is to say, we have

$$\limsup_{\substack{(\bar{x}, \bar{y}) \rightarrow (x, y) \\ t \searrow 0}} \frac{f(\bar{x} + th, \bar{y} + tk) - f(\bar{x}, \bar{y})}{t} = \lim_{t \searrow 0} \frac{f(x + th, y + tk) - f(x, y)}{t} \quad (10)$$

for all $(x, y) \in \mathbb{R}^m \times \mathbb{R}^n$ and $(h, k) \in \mathbb{R}^m \times \mathbb{R}^n$, and the limit on the right hand side exists. As shown by Clarke [13, 2.5.2 Example], one should not take this property for granted. The function in Fig. 2a defined from \mathbb{R}^2 to \mathbb{R} by $\varphi(x_1, x_2) := \max\{\min\{x_1, -x_2\}, x_2 - x_1\}$ is not subdifferentially regular, despite being continuous and semi-algebraic, just like f , and hence belonging to the class of tame functions [19]. As a result, its partial Clarke subdifferentials at the origin, which is a critical point, are completely decorrelated from its Clarke subdifferential: $\partial_{x_1} \varphi(0, 0) \times \partial_{x_2} \varphi(0, 0) \not\subset \partial \varphi(0, 0) \not\subset \partial_{x_1} \varphi(0, 0) \times \partial_{x_2} \varphi(0, 0)$, where $\not\subset$ means “is not a subset of”, as can be seen in Fig. 2b. The partial Clarke subdifferentials are hence of no use to analyze the landscape of φ . In contrast, since f is subdifferentially regular, we have

$$\partial f(x, y) \subset \partial_{x_1} f(x, y) \times \dots \times \partial_{x_m} f(x, y) \times \partial_{y_1} f(x, y) \times \dots \times \partial_{y_n} f(x, y) \quad (11)$$

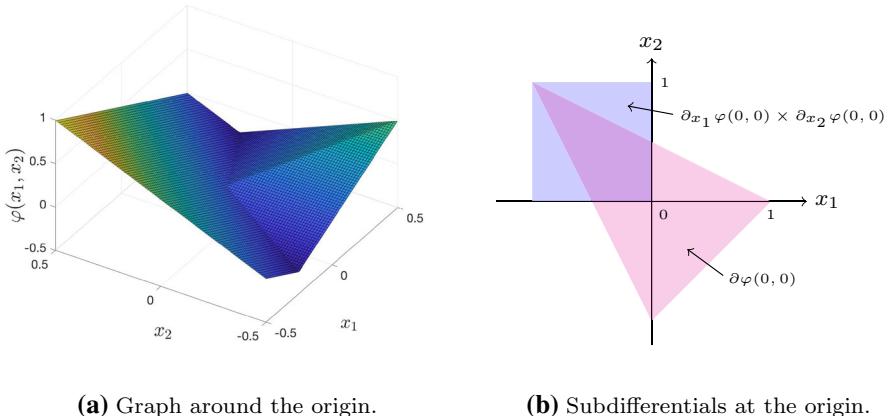


Fig. 2 Continuous semi-algebraic function that is not subdifferentially regular

where \subset means “is a subset of”. The inclusion is strict when

$$x = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad y = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \text{and} \quad M = \begin{pmatrix} 2 & 1 \\ -1 & -1/2 \end{pmatrix} \quad (12)$$

since $0 \notin \partial f(x, y)$ yet $0 \in \partial_{x_1} f(x, y) \times \partial_{x_2} f(x, y) \times \partial_{y_1} f(x, y) \times \partial_{y_2} f(x, y)$. It thus yields a necessary but not sufficient condition for being critical. The proof of Lemma 1 makes extensive use of this necessary condition but ultimately invokes the critical condition to conclude.

2 Proof of Theorem 1

It will be convenient in our analysis to consider a factorization of M whose rank we assume to be less than or equal to one. From now on, let $u \in \mathbb{R}^m$ and $v \in \mathbb{R}^n$ denote vectors such that $M = uv^T$. In order to introduce Lemma 1, observe that

$$f(x, y) = \sum_{i=1}^m \sum_{j=1}^n |x_i y_j - u_i v_j| \quad (13a)$$

$$= \sum_{u_i \neq 0} \sum_{j=1}^n |x_i y_j - u_i v_j| + \sum_{u_i=0} \sum_{j=1}^n |x_i y_j| \quad (13b)$$

$$= \sum_{u_i \neq 0} |u_i| \sum_{j=1}^n |y_j(x_i/u_i) - v_j| + \sum_{j=1}^n |y_j| \sum_{u_i=0} |x_i| \quad (13c)$$

where we use the convention that a sum over an index set which is empty is equal to zero. Similarly,

$$f(x, y) = \sum_{j=1}^n \sum_{i=1}^m |x_i y_j - u_i v_j| \quad (14a)$$

$$= \sum_{v_j \neq 0} \sum_{i=1}^m |x_i y_j - u_i v_j| + \sum_{v_j=0} \sum_{i=1}^m |x_i y_j| \quad (14b)$$

$$= \sum_{v_j \neq 0} |v_j| \sum_{i=1}^m |x_i (y_j/v_j) - u_i| + \sum_{i=1}^m |x_i| \sum_{v_j=0} |y_j|. \quad (14c)$$

As a result,

$$f(x, y) = \sum_{u_i \neq 0} |u_i| \alpha(x_i/u_i) + \|y\|_1 \sum_{u_i=0} |x_i|, \quad (15a)$$

$$f(x, y) = \sum_{v_j \neq 0} |v_j| \beta(y_j/v_j) + \|x\|_1 \sum_{v_j=0} |y_j|, \quad (15b)$$

where the functions α and β are defined from \mathbb{R} to \mathbb{R} by

$$\alpha(t) := \sum_{j=1}^n |y_j t - v_j| \quad \text{and} \quad \beta(t) := \sum_{i=1}^m |x_i t - u_i| \quad (16)$$

and $\|\cdot\|_1$ is the ℓ_1 -norm. The partial Clarke subdifferentials [13, p. 48] of f are

$$\partial_{x_i} f(x, y) = \begin{cases} \text{sign}(u_i) \partial \alpha(x_i/u_i) & \text{if } u_i \neq 0, \\ \text{sign}(x_i) \|y\|_1 & \text{if } u_i = 0, \end{cases} \quad (17a)$$

$$\partial_{y_j} f(x, y) = \begin{cases} \text{sign}(v_j) \partial \beta(y_j/v_j) & \text{if } v_j \neq 0, \\ \text{sign}(y_j) \|x\|_1 & \text{if } v_j = 0, \end{cases} \quad (17b)$$

where

$$\partial \alpha(t) = \sum_{j=1}^n \text{sign}(y_j t - v_j) y_j \quad \text{and} \quad \partial \beta(t) = \sum_{i=1}^m \text{sign}(x_i t - u_i) x_i. \quad (18)$$

Since α and β are convex piecewise affine functions, their subdifferentials $\partial \alpha$ and $\partial \beta$ are non-decreasing step functions, as can be seen in Fig. 3. From the expressions of $\partial \alpha(t)$ and $\partial \beta(t)$ in (18), it follows that the jumps between the steps of $\partial \alpha$ occur at v_j/y_j for all index j such that $y_j \neq 0$, while those of $\partial \beta$ occur at u_i/x_i such that $x_i \neq 0$. Observe that 0 is a root of $\partial \alpha$ and $\partial \beta$ in Fig. 3. This is true whenever (x, y) is a critical point of f and it is not a global minimum of f . Such is the object of Lemma 1 below.

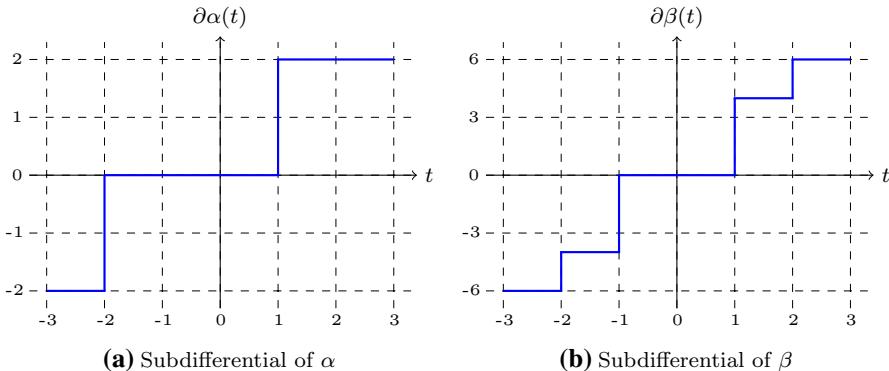


Fig. 3 Step functions at the critical point $x = (2, -1, -1, 1, -1)^T$, $y = (-1, -1/2, -1/2)^T$ where $u = (-2, -1, 2, 1, -2)^T$, $v = (-1, 1, 1)^T$

Lemma 1 $0 \in \partial f(x, y) \implies f(x, y) = 0$ or $0 \in \partial\alpha(0) \cap \partial\beta(0)$.

Proof Assume that $0 \in \partial f(x, y)$. Since f is the composition of a convex function and a continuously differentiable function, by virtue of [13, Corollary p. 32], [13, 2.3.6 Proposition (b)], and [13, 2.3.10 Chain Rule II], it is subdifferentially regular [13, 2.3.4 Definition]. It follows from [13, 2.3.15 Proposition] that $\partial f(x, y) \subset \partial_{x_1} f(x, y) \times \dots \times \partial_{x_m} f(x, y) \times \partial_{y_1} f(x, y) \times \dots \times \partial_{y_n} f(x, y)$. Hence $0 \in \partial_{x_i} f(x, y)$ and $0 \in \partial_{y_j} f(x, y)$ for all indices i and j . Based on the expressions of the partial Clarke subdifferentials in (17a)–(17b), we get that $0 \in \partial\alpha(x_i/u_i) \cap \partial\beta(y_j/v_j)$ for any indices i and j such that $u_i \neq 0$ and $v_j \neq 0$. In other words, whenever they are well-defined, the ratios x_i/u_i and y_j/v_j are roots of $\partial\alpha$ and $\partial\beta$, respectively.

We next reason by contradiction and assume that $f(x, y) > 0$ and $0 \notin \partial\alpha(0) \cap \partial\beta(0)$. If $u = 0$ and $v = 0$, then from (18) we get that $\partial\alpha(t) = \|y\|_1 \text{sign}(t)$ and $\partial\beta(t) = \|x\|_1 \text{sign}(t)$, so that $0 \in \partial\alpha(0) \cap \partial\beta(0)$. As a result, either $u \neq 0$ or $v \neq 0$. Assume that $u = 0$. Based on what was just said, $v \neq 0$. Also, $x \neq 0$ and $y \neq 0$, otherwise $f(x, y) = 0$. Since $u = 0$, from the expression of partial Clarke subdifferential in (17a) we find that $0 \in \partial f_{x_i}(x, y) = \text{sign}(x_i) \|y\|_1$ for all index i . Thus $x = 0$, which is a contradiction. We deduce that $u \neq 0$, and by the same reasoning, $v \neq 0$. Assume that $x = 0$. Then, from (18), we get that $\partial\beta(t) = 0$ for all $t \in \mathbb{R}$, and in particular, $0 \in \partial\beta(0)$. In addition, x_i/u_i is a root of $\partial\alpha$ whenever $u_i \neq 0$, and since $u \neq 0$, 0 is a root of $\partial\alpha$. As a result, $0 \in \partial\alpha(0) \cap \partial\beta(0)$, which is a contradiction. We deduce that $x \neq 0$, and by the same reasoning, $y \neq 0$. In addition, for all indices i and j such that $u_i = 0$ and $v_j = 0$, we have $0 \in \partial f_{x_i}(x, y) = \text{sign}(x_i) \|y\|_1$ and $0 \in \partial f_{y_j}(x, y) = \text{sign}(y_j) \|x\|_1$, whence $x_i = 0$ and $y_j = 0$. To sum up, so far we have shown that $u \neq 0$, $v \neq 0$, $x \neq 0$, $y \neq 0$, $x_i = 0$ if $u_i = 0$, and $y_j = 0$ if $v_j = 0$.

We next analyze the roots and jumps of $\partial\alpha$ and $\partial\beta$. Neither step function has a jump at the origin, otherwise there exists i and j such that $y_j \neq 0$ yet $v_j = 0$, and $x_i \neq 0$ yet $u_i = 0$. As for the values of the step functions at the origin, they cannot both be zero otherwise $0 \in \partial\alpha(0) \cap \partial\beta(0)$. Without loss of generality, we may thus assume from now on that $\partial\alpha(0) \neq 0$. If the non-decreasing function $\partial\alpha$ has a negative and a positive root, then $\partial\alpha(0) = 0$. As a result, without loss of generality, we may

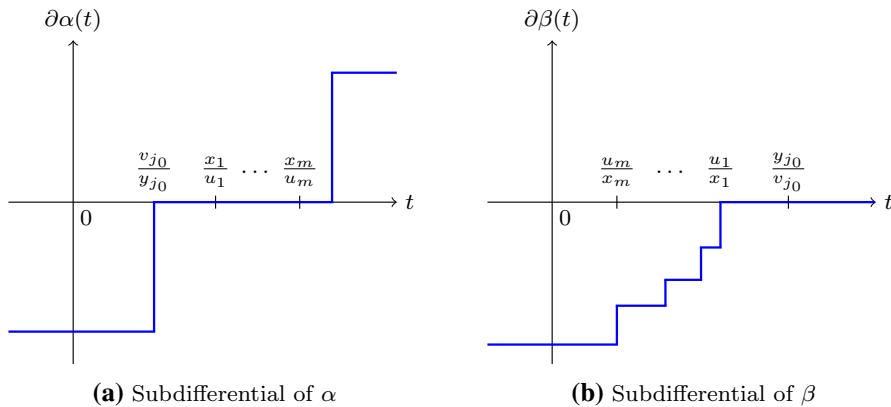


Fig. 4 Visualization of the proof of Lemma 1

assume that $x_i/u_i > 0$ for all index i such that $u_i \neq 0$ (and in particular, $x_i \neq 0$ if $u_i \neq 0$). If $\partial\alpha$ has no positive jump point that is less than or equal to each root x_i/u_i , then $\partial\alpha(0) = 0$. Hence, let v_{j_0}/y_{j_0} be such a jump point, where $y_{j_0} \neq 0$. We thus have $0 < v_{j_0}/y_{j_0} \leq x_i/u_i$ for all index i such that $u_i \neq 0$, as illustrated in Fig. 4a.

We next discuss the repercussions of the jump point v_{j_0}/y_{j_0} of $\partial\alpha$ on the roots and jumps of $\partial\beta$. Inverting yields that $0 < u_i/x_i \leq y_{j_0}/v_{j_0}$ for all index i such that $u_i \neq 0$. As was shown in the second paragraph, $x_i = 0$ if $u_i = 0$, hence the previous sentence is true for all index i such that $x_i \neq 0$. Hence all the jump points of $\partial\beta$ are less than or equal to one of its roots y_{j_0}/v_{j_0} . This is illustrated in Fig. 4b. Assume that $\partial\beta(y_{j_0}/v_{j_0}) \subset (-\infty, 0]$, which is true if y_{j_0}/v_{j_0} is not a jump, as it is represented in Fig. 4b. Then $\partial\beta(t) = 0$ for all $t \geq y_{j_0}/v_{j_0}$. As remarked in parenthesis in the previous paragraph, $u_i = 0$ if $x_i = 0$, hence for all t large enough we have $\text{sign}(x_i t - u_i) = \text{sign}(x_i t)$. Thus $\partial\beta(t) = \sum_{i=1}^m \text{sign}(x_i t) x_i = \sum_{i=1}^m |x_i|$ for all t large enough. It follows that $x = 0$, which is contradiction. As a result, there exists $\epsilon > 0$ such that $[0, \epsilon] \subset \partial\beta(y_{j_0}/v_{j_0})$, and in particular, y_{j_0}/v_{j_0} is a jump point of $\partial\beta$. Since $\partial\beta$ is a non-decreasing step function, it has no roots greater than y_{j_0}/v_{j_0} .

We next consider the case where $[-\epsilon, \epsilon] \subset \partial\beta(y_{j_0}/v_{j_0})$, after possibly reducing $\epsilon > 0$. Since $\partial\beta$ is a non-decreasing step function, it has no roots less than y_{j_0}/v_{j_0} . Since y_j/v_j is a root of $\partial\beta$ whenever $v_j \neq 0$, it follows that $y_j/v_j = y_{j_0}/v_{j_0} > 0$ for all $v_j \neq 0$ (and in particular, we have $y_j \neq 0$ if $v_j \neq 0$). Inverting yields that $v_j/y_j = v_{j_0}/y_{j_0}$ for all $v_j \neq 0$. As was shown in the second paragraph, if $y_j \neq 0$, then $v_j \neq 0$. It follows that $v_j/y_j = v_{j_0}/y_{j_0}$ for all $y_j \neq 0$. Hence v_{j_0}/y_{j_0} is the unique jump point of $\partial\alpha$. Recall that $0 < v_{j_0}/y_{j_0} \leq x_i/u_i$ and x_i/u_i is a root of $\partial\alpha$ for all index i such that $u_i \neq 0$. Assume that there exists an index i_0 such that x_{i_0}/u_{i_0} such that $v_{j_0}/y_{j_0} < x_{i_0}/u_{i_0}$. Then $\partial\alpha(t) = 0$ for all $t \geq x_{i_0}/u_{i_0}$. As remarked in parenthesis in this paragraph, it holds that $v_j = 0$ if $y_j = 0$, and thus for all t large enough we have $\text{sign}(y_{j_0} t - v_j) = \text{sign}(y_{j_0} t)$. Hence $\partial\alpha(t) = \sum_{j=1}^n \text{sign}(y_{j_0} t) y_{j_0} = \sum_{j=1}^n |y_{j_0}|$ for all t large enough. It follows that $y = 0$, which is contradiction. As a result, $v_j/y_j = v_{j_0}/y_{j_0} = x_i/u_i$ for all indices i and j such that $u_i \neq 0$ and $y_j \neq 0$, and hence $x_i y_j - u_i v_j = 0$. Recall that $x_i = 0$ if $u_i = 0$, and $v_j = 0$ if $y_j = 0$. Hence

$x_i y_j - u_i v_j = 0$ for all indices i and j , which implies that $f(x, y) = 0$. This is a contradiction.

It remains to consider the case where $\partial\beta(y_{j_0}/v_{j_0}) = [0, \epsilon]$, possibly after increasing $\epsilon > 0$. Recall that all the jump points of $\partial\beta$, i.e. u_i/x_i for all index i such that $x_i \neq 0$, are less than or equal to y_{j_0}/v_{j_0} . Assume that $u_i/x_i = y_{j_0}/v_{j_0}$ for all index i such that $x_i \neq 0$. Hence y_{j_0}/v_{j_0} is unique jump point of $\partial\beta$. Since $\partial\beta(y_{j_0}/v_{j_0}) = [0, \epsilon]$, we have $\partial\beta(t) = 0$ for all $t \leq y_{j_0}/v_{j_0}$. Since $u_i = 0$ if $x_i = 0$, for all t small enough (i.e., taking large negative values), we have $\text{sign}(x_i t - u_i) = -\text{sign}(x_i)$. Hence $\partial\beta(t) = \sum_{i=1}^m -\text{sign}(x_i)x_i = -\sum_{i=1}^m |x_i|$ for all t small enough. It follows that $x = 0$, which is contradiction. As a result, there exists an index i_0 such that $0 < u_{i_0}/x_{i_0} < y_{j_0}/v_{j_0}$. Recall that y_j/v_j is a root of $\partial\beta$ for all index j such that $v_j \neq 0$. Since $\partial\beta(y_{j_0}/v_{j_0}) = [0, \epsilon]$, we have $0 < u_{i_0}/x_{i_0} \leq y_j/v_j \leq y_{j_0}/v_{j_0}$ for all index j such that $v_j \neq 0$ (and in particular, we have $y_j \neq 0$ if $v_j \neq 0$). Inverting yields $0 < v_{j_0}/y_{j_0} \leq v_j/y_j \leq x_{i_0}/u_{i_0}$ for all index j such that $v_j \neq 0$. As was shown in the second paragraph, if $y_j \neq 0$ then $v_j \neq 0$, hence the previous sentence is true for all index j such that $y_j \neq 0$. Since $\partial\beta(y_{j_0}/v_{j_0}) = [0, \epsilon]$, y_{j_0}/v_{j_0} is a jump point of $\partial\beta$, so there exists i_1 such that $y_{j_0}/v_{j_0} = u_{i_1}/x_{i_1}$. We thus have $0 < x_{i_1}/u_{i_1} = v_{j_0}/y_{j_0} \leq v_j/y_j \leq x_{i_0}/u_{i_0}$ for all index j such that $y_j \neq 0$. If $x_{i_1}/u_{i_1} = v_{j_0}/y_{j_0} < v_j/y_j < x_{i_0}/u_{i_0}$, then v_j/y_j is a jump point of $\partial\alpha$ located strictly between two of its roots, which is impossible. Hence for all index j such that $y_j \neq 0$, either $v_j/y_j = \mu := x_{i_1}/u_{i_1}$ or $v_j/y_j = \nu := x_{i_0}/u_{i_0}$. If $v_j/y_j \neq \nu$ for all index j such that $y_j \neq 0$, then one of the roots of $\partial\alpha$ is greater than all its jump points. Hence $\partial\alpha(t) = 0$ for all t large enough. Together with the fact that $u_i = 0$ if $x_i = 0$, this yields a contradiction. Hence there exists j_1 such that $v_{j_1}/y_{j_1} = x_{i_0}/u_{i_0}$. We next show that the dichotomy $v_j/y_j = \mu$ or $v_j/y_j = \nu$ also holds for x_i/u_i whenever $u_i \neq 0$.

Based on the previous paragraph, we have $0 < x_{i_1}/u_{i_1} = v_{j_0}/y_{j_0} < v_{j_1}/y_{j_1} = x_{i_0}/u_{i_0}$. Inverting yields $u_{i_0}/x_{i_0} = y_{j_1}/v_{j_1} < y_{j_0}/v_{j_0} = u_{i_1}/x_{i_1}$. Recall that y_j/v_j are roots of $\partial\beta$ for all index j such that $v_j \neq 0$, and the jump points of $\partial\beta$ are u_i/x_i for all index i such that $x_i \neq 0$. The non-decreasing function $\partial\beta$ cannot have a jump point strictly between two of its roots, hence for all index i such that $x_i \neq 0$, it holds that $0 < u_i/x_i \leq u_{i_0}/x_{i_0} = y_{j_1}/v_{j_1} < y_{j_0}/v_{j_0} = u_{i_1}/x_{i_1}$ or $0 < u_{i_0}/x_{i_0} = y_{j_1}/v_{j_1} < y_{j_0}/v_{j_0} = u_{i_1}/x_{i_1} \leq u_i/x_i$. Inverting yields $x_{i_1}/u_{i_1} = v_{j_0}/y_{j_0} < v_{j_1}/y_{j_1} = x_{i_0}/u_{i_0} \leq x_i/u_i$ or $x_i/u_i \leq x_{i_1}/u_{i_1} = v_{j_0}/y_{j_0} < v_{j_1}/y_{j_1} = x_{i_0}/u_{i_0}$. If the inequality is strict in the first case, then x_i/u_i is a root of $\partial\alpha$ that is greater than all of its jump points, of which there are two according to the previous paragraph, namely v_{j_1}/y_{j_1} and v_{j_0}/y_{j_0} . If the inequality is strict in the second case, then x_i/u_i is a root of $\partial\alpha$ that is less than all of its jump points. In either case, $\partial\alpha(t) = 0$ for all t either small or large enough. Together with the fact that $v_j = 0$ if $y_j = 0$ (as remarked in parenthesis in the previous paragraph), this yields a contradiction. Hence for all index i such that $u_i \neq 0$, either $x_i/u_i = \mu$ or $x_i/u_i = \nu$, and each case clearly happens for some index (the first with i_1 , the second with i_0). In light of the dichotomy exposed above, consider

$(h, k) \in \mathbb{R}^m \times \mathbb{R}^n$ defined by

$$h_i := \begin{cases} 0 & \text{if } x_i/u_i = \mu, \\ -u_i v & \text{if } x_i/u_i = v, \\ 0 & \text{if } u_i = 0, \end{cases} \quad \text{and} \quad k_j := \begin{cases} 0 & \text{if } y_j/v_j = 1/\mu, \\ v_j/v & \text{if } y_j/v_j = 1/v, \\ 0 & \text{if } v_j = 0. \end{cases} \quad (19)$$

Consider also the function γ defined from $\mathbb{R}^{m \times n}$ to \mathbb{R} by $\gamma(Q) := h^T Q y + x^T Q k$. So far in the proof, we have only used the necessary condition for being critical provided by (11). We next invoke the critical condition, as discussed in the introduction. According to the critical condition (9), there exists $\Lambda \in \text{sign}(xy^T - uv^T)$ such that $\Lambda y = 0$ and $\Lambda^T x = 0$. Thus $\gamma(\Lambda) = h^T \Lambda y + x^T \Lambda k = h^T(\Lambda y) + (\Lambda^T x)^T k = 0$. Yet, observe that the image $\gamma(\text{sign}(xy^T - uv^T)) = \dots$

$$= \sum_{i=1}^m \sum_{j=1}^n \text{sign}(x_i y_j - u_i v_j) (h_i y_j + x_i k_j) \quad (20a)$$

$$= \sum_{\substack{x_i = \mu \\ u_i = \mu}} \sum_{\substack{y_j = \frac{1}{\mu} \\ v_j = \frac{1}{\mu}}} \text{sign}(x_i y_j - u_i v_j) (0 \times y_j + x_i \times 0) + \quad (20b)$$

$$\sum_{\substack{x_i = \mu \\ u_i = \mu}} \sum_{\substack{y_j = \frac{1}{\mu} \\ v_j = \frac{1}{v}}} \text{sign}(x_i y_j - u_i v_j) (0 \times y_j + x_i \times v_j/v) + \quad (20c)$$

$$\sum_{\substack{x_i = \mu \\ u_i = \mu}} \sum_{v_j = 0} \text{sign}(x_i y_j - u_i v_j) (0 \times y_j + x_i \times 0) + \quad (20d)$$

$$\sum_{\substack{x_i = v \\ u_i = v}} \sum_{\substack{y_j = \frac{1}{\mu} \\ v_j = \frac{1}{\mu}}} \text{sign}(x_i y_j - u_i v_j) (-u_i v \times y_j + x_i \times 0) + \quad (20e)$$

$$\sum_{\substack{x_i = v \\ u_i = v}} \sum_{\substack{y_j = \frac{1}{\mu} \\ v_j = \frac{1}{v}}} \text{sign}(x_i y_j - u_i v_j) (-u_i v \times y_j + x_i \times v_j/v) + \quad (20f)$$

$$\sum_{\substack{x_i = v \\ u_i = v}} \sum_{v_j = 0} \text{sign}(x_i y_j - u_i v_j) (-u_i v \times y_j + x_i \times 0) + \quad (20g)$$

$$\sum_{u_i = 0} \sum_{\substack{y_j = \frac{1}{\mu} \\ v_j = \frac{1}{\mu}}} \text{sign}(x_i y_j - u_i v_j) (0 \times y_j + x_i \times 0) + \quad (20h)$$

$$\sum_{u_i = 0} \sum_{\substack{y_j = \frac{1}{\mu} \\ v_j = \frac{1}{v}}} \text{sign}(x_i y_j - u_i v_j) (0 \times y_j + x_i \times v_j/v) + \quad (20i)$$

$$\sum_{u_i = 0} \sum_{v_j = 0} \text{sign}(x_i y_j - u_i v_j) (0 \times y_j + x_i \times 0) \quad (20j)$$

$$= \sum_{\substack{x_i = \mu \\ u_i = \mu}} \sum_{\substack{y_j = \frac{1}{\mu} \\ v_j = \frac{1}{v}}} \text{sign}((\mu/v - 1)u_i v_j)(u_i \mu v_j/v) + \quad (20k)$$

$$\sum_{\substack{x_i \\ u_i \\ u_i=v}} \sum_{\substack{y_j \\ v_j \\ v_j=\frac{1}{\mu}}} \text{sign}((v/\mu - 1)u_i v_j)(-u_i v v_j/\mu) \quad (20l)$$

$$= -\mu/v \sum_{\substack{x_i \\ u_i \\ u_i=\mu}} \sum_{\substack{y_j \\ v_j \\ v_j=\frac{1}{v}}} |u_i v_j| - v/\mu \sum_{\substack{x_i \\ u_i \\ u_i=v}} \sum_{\substack{y_j \\ v_j \\ v_j=\frac{1}{\mu}}} |u_i v_j| < 0. \quad (20m)$$

Above, (20a) follows from the definition of the function γ . We substitute h_i and k_j using their definition in (19), which yields (20b)–(20j). We next substitute x_i and y_j using their expressions below the summation signs and obtain (20k)–(20l). Note that all but two terms cancel out: (20b) cancels out for obvious reasons; (20c) yields (20k); (20d) cancels out for obvious reasons; (20e) yields (20l); (20f) cancels out because $-u_i v \times y_j + x_i \times v_j/v = -u_i v \times v_j/v + u_i v \times v_j/v = 0$; (20g) cancels out because $y_j = 0$ if $v_j = 0$, as shown in the second paragraph; (20h) cancels out for obvious reasons; (20i) cancels out because $x_i = 0$ if $u_i = 0$, as was shown in the second paragraph; (20j) cancels out for obvious reasons. To get from (20k)–(20l) to (20m), we use the fact that $\mu/v - 1 < 0$ and $v/\mu - 1 > 0$ since $0 < \mu < v$. We also use the fact that $\text{sign}(u_i v_j) u_i v_j = |u_i v_j|$. The result in (20m) is negative because the summation takes place over non-empty sets: $x_{i_0}/u_{i_0} = \mu$, $y_{j_0}/v_{j_0} = 1/v$, $x_{i_1}/u_{i_1} = v$, and $y_{j_1}/v_{j_1} = 1/\mu$. (While it may seem strange at first, the image of the set $\text{sign}(xy^T - uv^T)$ via the function γ is actually a singleton.) In particular, $\gamma(\Lambda) < 0$ whereas we had shown above that $\gamma(\Lambda) = 0$. This is a contradiction and terminates the proof. \square

Thanks to Lemma 1, we may now remove the existence quantifier from the critical condition (9). We obtain a finite number of unions and intersections of polynomial equations and inequalities, where the number is independent of the dimensions m and n .

Lemma 2 $(x, y) \in \mathbb{R}^m \times \mathbb{R}^n$ is a critical point of f if and only if

$$x_i y_j = u_i v_j \text{ for } i = 1, \dots, m, \ j = 1, \dots, n, \text{ or} \quad (21a)$$

$$\left| \sum_{u_i \neq 0} \text{sign}(u_i) x_i \right| \leq \sum_{u_i=0} |x_i| \text{ and } y = 0, \text{ or} \quad (21b)$$

$$x = 0 \text{ and } \left| \sum_{v_j \neq 0} \text{sign}(v_j) y_j \right| \leq \sum_{v_j=0} |y_j|, \text{ or} \quad (21c)$$

$$\sum_{u_i \neq 0} \text{sign}(u_i) x_i = \sum_{v_j \neq 0} \text{sign}(v_j) y_j = 0, \quad \frac{x_i y_j}{u_i v_j} \leq 1 \text{ if } u_i v_j \neq 0, \quad (21d)$$

$$x_i = 0 \text{ if } u_i = 0, \text{ and } y_j = 0 \text{ if } v_j = 0. \quad (21e)$$

Proof (\implies) Assume that $0 \in \partial f(x, y)$. If $f(x, y) = 0$, then (21a) holds. Otherwise, we next prove that (21b)–(21e) hold. From the expression of $\partial\alpha(t)$ and $\partial\beta(t)$ in (18),

it follows that

$$\partial\alpha(0) = -\sum_{j=1}^n \text{sign}(v_j)y_j = -\sum_{v_j \neq 0} \text{sign}(v_j)y_j - \sum_{v_j=0} \text{sign}(0)y_j, \quad (22a)$$

$$\partial\beta(0) = -\sum_{i=1}^m \text{sign}(u_i)x_i = -\sum_{u_i \neq 0} \text{sign}(u_i)x_i - \sum_{u_i=0} \text{sign}(0)x_i. \quad (22b)$$

Since $0 \in \partial f(x, y)$ and $f(x, y) \neq 0$, from Lemma 1 we get that $0 \in \partial\alpha(0) \cap \partial\beta(0)$. Together with (22a)–(22b), this yields that

$$\left| \sum_{u_i \neq 0} \text{sign}(u_i)x_i \right| \leq \sum_{u_i=0} |x_i| \quad \text{and} \quad \left| \sum_{v_j \neq 0} \text{sign}(v_j)y_j \right| \leq \sum_{v_j=0} |y_j|. \quad (23)$$

If $y = 0$, then from the inequality on the left hand side of (23), we obtain (21b). Likewise, if $x = 0$, then from the inequality on the right hand side of (23), we obtain (21c). If neither $x = 0$ nor $y = 0$, then for all indices i and j such that $u_i = 0$ and $v_j = 0$, we have $0 \in \partial_{x_i} f(x, y) = \text{sign}(x_i)\|y\|_1$ and $0 \in \partial_{y_j} f(x, y) = \text{sign}(y_j)\|x\|_1$, whence $x_i = 0$ and $y_j = 0$. Hence (21e) is true. In light of this, the inequalities in (23) become equalities and equal to zero, yielding the equalities in (21d). It remains to show the ratio inequalities in (21d), for which we don't need to assume that neither x nor y are equal to zero. Indeed, the ratio inequalities trivially hold in this case. The ratio inequalities are the object of the next paragraph.

Since $0 \in \partial_{x_i} f(x, y)$ and $0 \in \partial_{y_j} f(x, y)$ for all indices i and j , from the expressions of the partial Clarke subdifferentials in (17a)–(17b), it follows that $0 \in \partial\alpha(x_i/u_i) \cap \beta(y_j/v_j)$ whenever $u_i \neq 0$ and $v_j \neq 0$. Recall from Lemma 1 that we also have that $0 \in \partial\alpha(0) \cap \partial\beta(0)$. Observe that the non-decreasing function $\partial\alpha$ cannot contain a jump point between the root 0 and any root x_i/u_i . Hence, for all index j such that $y_j \neq 0$, if the jump point v_j/y_j is positive, then it is greater than or equal to all the roots x_i/u_i , that is to say, $v_j/y_j \geq x_i/u_i$. If the jump point v_j/y_j is negative, then it is less than or equal to all the roots x_i/u_i , that is to say, $v_j/y_j \leq x_i/u_i$. Multiplying both inequalities by y_j/v_j yields $x_i y_j / (u_i v_j) \leq 1$ whenever $u_i v_j \neq 0$.

(\Leftarrow) If (21a) holds, then (x, y) is global minimum of f and hence a critical point according to the generalized Fermat rule [13, 2.3.2 Proposition]. If (21b) holds, then from the expression of $\partial\beta(0)$ in (22b), it follows that $0 \in \partial\beta(0)$. From the expression of $\partial_{y_j} f(x, y)$ in (17b) and $y = 0$, we then get that $0 \in \partial_{y_j} f(x, y)$ for all index j . Since $f(x, \cdot)$ is subdifferentially regular for all $x \in \mathbb{R}^m$, by virtue of [13, 2.3.15 Proposition] it holds that $\partial_y f(x, y) \subset \partial_{y_1} f(x, y) \times \dots \times \partial_{y_n} f(x, y)$. Using [13, 2.3.10 Chain Rule II], we actually find that $\partial_y f(x, y) = \partial_{y_1} f(x, y) \times \dots \times \partial_{y_n} f(x, y) = \{\Lambda^T x \mid \Lambda \in \text{sign}(xy^T - uv^T)\}$. Thus there exists $\Lambda \in \text{sign}(xy^T - uv^T)$ such that $\Lambda^T x = 0$. Since $y = 0$, we naturally also have that $\Lambda y = 0$. Hence the critical condition (9) holds. The same argument applies when (21c) holds. Assume that (21d)–(21e) hold. In order to exhibit a matrix Λ in the critical condition (9), we propose to define the following

function:

$$\operatorname{sgn}(t) := \begin{cases} -1 & \text{if } t < 0, \\ 0 & \text{if } t = 0, \\ 1 & \text{if } t > 0. \end{cases} \quad (24)$$

Like the sign function, the sgn function applies to matrices term by term. Using this new function, we may rewrite the equalities in (21d) as

$$\sum_{j=1}^n \operatorname{sgn}(v_j) y_j = 0, \quad i = 1, \dots, m, \quad (25a)$$

$$\sum_{i=1}^m \operatorname{sgn}(u_i) x_i = 0, \quad j = 1, \dots, n. \quad (25b)$$

After multiplying the above equations by $\operatorname{sgn}(u_i)$ and $\operatorname{sgn}(v_j)$ respectively, we obtain that

$$\sum_{j=1}^n \operatorname{sgn}(u_i v_j) y_j = 0, \quad i = 1, \dots, m, \quad (26a)$$

$$\sum_{i=1}^m \operatorname{sgn}(v_j u_i) x_i = 0, \quad j = 1, \dots, n. \quad (26b)$$

In other words, $\Lambda y = 0$ and $\Lambda^T x = 0$ with $\Lambda := -\operatorname{sgn}(uv^T)$. In order to show that (x, y) is critical, it remains to show that $\Lambda \in \operatorname{sign}(xy^T - uv^T)$. If $u_i v_j = 0$, then from (21e) we get that $x_i y_j = 0$. In that case, $-\operatorname{sgn}(u_i v_j) = 0$ and $\operatorname{sign}(x_i y_j - u_i v_j) = [-1, 1]$. Hence $-\operatorname{sgn}(u_i v_j) \in \operatorname{sign}(x_i y_j - u_i v_j)$. If $u_i v_j \neq 0$, then from the inequalities in (21d) we get that $x_i y_j / (u_i v_j) - 1 \leq 0$. In that case, $\operatorname{sign}(x_i y_j - u_i v_j) = \operatorname{sign}(u_i v_j) \operatorname{sign}(x_i y_j / (u_i v_j) - 1) \ni -\operatorname{sign}(u_i v_j) = -\operatorname{sgn}(u_i v_j)$. \square

Thanks to Lemma 2, we may now classify the critical points according to whether they are global minima, local minima that are not global minima (i.e. spurious local minima), or not local minima (i.e. saddle points).

Proposition 1 *f has the following properties when $u \neq 0$ and $v \neq 0$.*

(1) *The global minima are all $(x, y) \in \mathbb{R}^m \times \mathbb{R}^n$ such that*

$$\exists \theta \in \mathbb{R} \setminus \{0\} : \quad (x, y) = (u\theta, v/\theta) \quad (27)$$

(2) *The spurious local minima are all $(x, y) \in \mathbb{R}^m \times \mathbb{R}^n$ such that*

$$\left| \sum_{u_i \neq 0} \operatorname{sign}(u_i) x_i \right| < \sum_{u_i=0} |x_i| \text{ and } y = 0, \quad \text{or} \quad (28a)$$

$$x = 0 \text{ and } \left| \sum_{v_j \neq 0} \text{sign}(v_j) y_j \right| < \sum_{v_j=0} |y_j|. \quad (28b)$$

(3) The saddle points are all $(x, y) \in \mathbb{R}^m \times \mathbb{R}^n$ such that

$$\left| \sum_{u_i \neq 0} \text{sign}(u_i) x_i \right| = \sum_{u_i=0} |x_i| \text{ and } y = 0, \text{ or} \quad (29a)$$

$$x = 0 \text{ and } \left| \sum_{v_j \neq 0} \text{sign}(v_j) y_j \right| = \sum_{v_j=0} |y_j|, \text{ or} \quad (29b)$$

$$\sum_{u_i \neq 0} \text{sign}(u_i) x_i = \sum_{v_j \neq 0} \text{sign}(v_j) y_j = 0, \quad \frac{x_i y_j}{u_i v_j} \leq 1 \text{ if } u_i v_j \neq 0, \quad (29c)$$

$$x_i = 0 \text{ if } u_i = 0, \text{ and } y_j = 0 \text{ if } v_j = 0. \quad (29d)$$

In addition, for all saddle point (x, y) , there exists a global minimum (x^*, y^*) such that $(x^*, y^*) - (x, y)$ is a direction of descent.

Proof Lemma 2 implies that (27), (28a)–(28b), and (29a)–(29d) form a partition of the set of critical points. It thus suffices to check in each case that the desired property about the local variation of f holds true. (1) Observe that (21a) and (27) are equivalent because $u \neq 0$ and $v \neq 0$. (2) Consider $x \in \mathbb{R}^m$ and $y = 0$. For all $(h, k) \in \mathbb{R}^m \times \mathbb{R}^n$ small enough, we have $f(x + h, y + k) = \dots$

$$= \sum_{i=1}^m \sum_{j=1}^n |(x_i + h_i)k_j - u_i v_j| \quad (30a)$$

$$= \sum_{u_i v_j = 0} |(x_i + h_i)k_j| + \sum_{u_i v_j \neq 0} |(x_i + h_i)k_j - u_i v_j| \quad (30b)$$

$$= \sum_{u_i v_j = 0} |(x_i + h_i)k_j| + \sum_{u_i v_j \neq 0} |u_i v_j| - \text{sign}(u_i v_j)(x_i + h_i)k_j \quad (30c)$$

$$= f(x, y) + \sum_{u_i v_j = 0} |(x_i + h_i)k_j| - \sum_{u_i v_j \neq 0} \text{sign}(u_i v_j)(x_i + h_i)k_j \quad (30d)$$

$$= f(x, y) + \sum_{u_i v_j = 0} |(x_i + h_i)k_j| - \sum_{v_j \neq 0} k_j \text{sign}(v_j) \sum_{u_i \neq 0} \text{sign}(u_i)(x_i + h_i) \quad (30e)$$

$$\geq f(x, y) + \sum_{u_i v_j = 0} |(x_i + h_i)k_j| - \sum_{v_j \neq 0} |k_j| \left| \sum_{u_i \neq 0} \text{sign}(u_i)(x_i + h_i) \right|$$

(equality holds if $\text{sign}(k_j) = \text{sign} \left(v_j \sum_{u_i \neq 0} \text{sign}(u_i)(x_i + h_i) \right)$ when $v_j \neq 0$)

(30f)

$$= f(x, y) + \sum_{v_j=0} |k_j| \sum_{i=1}^n |x_i + h_i| + \dots \quad (30g)$$

$$\sum_{v_j \neq 0} |k_j| \left(\sum_{u_i=0} |x_i + h_i| - \left| \sum_{u_i \neq 0} \text{sign}(u_i)(x_i + h_i) \right| \right) \quad (30h)$$

$$\geq f(x, y) + \sum_{v_j \neq 0} |k_j| \left(\sum_{u_i=0} |x_i + h_i| - \left| \sum_{u_i \neq 0} \text{sign}(u_i)(x_i + h_i) \right| \right)$$

$$(\text{equality holds if } k_j = 0 \text{ when } v_j = 0). \quad (30i)$$

Above, (30a) is a consequence of the definition of f and $y = 0$. (30b) is obtained by splitting the sum according to whether the product $u_i v_j$ is equal to zero. (30c) is valid since k_j is small and $|a + b| = |a| - \text{sign}(a)b$ if $|a| > |b|$. (30d) is due to $f(x, 0) = \sum_{u_i v_j \neq 0} |u_i v_j|$ and $y = 0$. (30e) is obtained by writing that $\text{sign}(u_i v_j)(x_i + h_i)k_j = \text{sign}(u_i)(x_i + h_i)\text{sign}(v_j)k_j$, then factorizing the sum. (30f) is obtained by observing that the last sum in (30e) is less than or equal to its absolute value. We then apply the triangular inequality and the fact that $|k_j \text{sign}(v_j)| = |k_j|$. Equality in the inequality in (30f) is then obtained by taking k_j to be of the same sign as the term it multiplies in the summation over $v_j \neq 0$ in (30e). (30g)–(30h) are obtained by splitting the first sum in (30f) according to whether the v_j is equal to zero. Finally, inequality (30i) holds since $\sum_{v_j=0} |k_j| \sum_{i=1}^n |x_i + h_i| \geq 0$. This term is equal to zero if k_j when $v_j = 0$, in which case we get equality in the inequality in (30i).

Assume that (28a) holds. For all $(h, k) \in \mathbb{R}^m \times \mathbb{R}^n$ small enough, (30i) yields

$$\begin{aligned} f(x + h, y + k) &\geq f(x, y) + \frac{1}{2} \sum_{v_j \neq 0} |k_j| \left(\sum_{u_i=0} |x_i| - \left| \sum_{u_i \neq 0} \text{sign}(u_i)x_i \right| \right) \\ &\geq f(x, y) > 0 \end{aligned} \quad (31a)$$

where the strict equality is due to $y = 0$, $u \neq 0$ and $v \neq 0$. Thus (x, y) is a spurious local minimum. The same argument applies to (28b).

(3) Assume that (29a) holds, namely $|\sum_{u_i \neq 0} \text{sign}(u_i)x_i| = \sum_{u_i=0} |x_i|$ and $y = 0$. If $|\sum_{u_i \neq 0} \text{sign}(u_i)x_i| = 0$, then (29c)–(29d) hold, a case that we will treat later. If $|\sum_{u_i \neq 0} \text{sign}(u_i)x_i| \neq 0$, then take any $\theta \neq 0$ such that $\text{sign}(\theta) = \text{sign}(\sum_{u_i \neq 0} \text{sign}(u_i)x_i)$ and consider the direction $(u\theta, v/\theta) - (x, y)$. It goes from (x, y) towards the global minimum $(u\theta, v/\theta)$. As we next show, taking a small step $t > 0$ in this direction renders the inequalities in (30f) and in (30i) binding, where $(h, k) := t(u\theta - x, v/\theta - y)$. Regarding (30f), observe that, if $v_j \neq 0$, then

$$\text{sign}(k_j) = \text{sign}(tv_j/\theta) \quad (32a)$$

$$= \text{sign} \left(tv_j \sum_{u_i \neq 0} \text{sign}(u_i)x_i \right) \quad (32b)$$

$$= \text{sign} \left((1-t)v_j \sum_{u_i \neq 0} \text{sign}(u_i)x_i \right) \quad (32c)$$

$$= \text{sign} \left((1-t)v_j \underbrace{\sum_{u_i \neq 0} \text{sign}(u_i)x_i}_{\neq 0} + tv_j \sum_{u_i \neq 0} \text{sign}(u_i)u_i\theta \right) \quad (32d)$$

$$= \text{sign} \left(v_j \sum_{u_i \neq 0} \text{sign}(u_i)[(1-t)x_i + tu_i\theta] \right) \quad (32e)$$

$$= \text{sign} \left(v_j \sum_{u_i \neq 0} \text{sign}(u_i)(x_i + h_i) \right). \quad (32f)$$

Above, (32a) is due to $k_j = v_j/\theta - y_j$, $y = 0$, and $t > 0$. (32b) follows from $\text{sign}(\theta) = \text{sign}(\sum_{u_i \neq 0} \text{sign}(u_i)x_i)$. Also, since $t > 0$ is small, multiplying a number by t or $(1-t)$ doesn't change its sign, hence (32c). It is legitimate to add a linear term in function of t in (32d) since it is dominated by the first term (t is small). We get (32e) by factorizing the two terms by v_j . Finally, (32f) is due to the fact that, by definition, $h = t(u\theta - x)$, and hence $tu_i\theta = h_i + tx_i$. Regarding (30i), observe that, if $v_j = 0$, then $k_j = tv_j/\theta = 0$.

In order to show that $(u\theta - x, v/\theta - y)$ is a direction of descent, we pick up the computation in (30a)–(30i) where we left off: $f(x + h, y + k) = \dots$

$$= f(x, y) + \sum_{v_j \neq 0} |tv_j/\theta| \left(\sum_{u_i=0} |x_i + t(u_i\theta - x_i)| + \right. \quad (33a)$$

$$\left. - \left| \sum_{u_i \neq 0} \text{sign}(u_i)[x_i + t(u_i\theta - x_i)] \right| \right) \quad (33b)$$

$$= f(x, y) + t/|\theta| \sum_{v_j \neq 0} |v_j| \left(\sum_{u_i=0} |(1-t)x_i| + \right. \quad (33c)$$

$$\left. - \left| (1-t) \sum_{u_i \neq 0} \text{sign}(u_i)x_i + t\theta \sum_{u_i \neq 0} \text{sign}(u_i)u_i \right| \right) \quad (33d)$$

$$= f(x, y) + t/|\theta| \sum_{j=1}^n |v_j| \left[(1-t) \sum_{u_i=0} |x_i| - (1-t) \left| \sum_{u_i \neq 0} \text{sign}(u_i)x_i \right| + \right. \quad (33e)$$

$$\left. - \text{sign} \left((1-t) \sum_{u_i \neq 0} \text{sign}(u_i)x_i \right) t\theta \sum_{u_i \neq 0} |u_i| \right] \quad (33f)$$

$$= f(x, y) + t/|\theta| \sum_{j=1}^n |v_j| \left[(1-t) \underbrace{\left(\sum_{u_i=0} |x_i| - \left| \sum_{u_i \neq 0} \text{sign}(u_i)x_i \right| \right)}_{=0} + \right. \quad (33g)$$

$$\left. - t|\theta| \sum_{i=1}^m |u_i| \right] \quad (33h)$$

$$= f(x, y) - t^2 \sum_{j=1}^n |v_j| \sum_{i=1}^m |u_i| \quad (33i)$$

$$= (1-t^2) f(x, y). \quad (33j)$$

Above, (33a)–(33b) is due to the definition of (h, k) . We get (33c)–(33d) by factorizing by $t/|\theta|$, canceling out u_i in the summation over $u_i = 0$, and expanding the product in the summation over $u_i \neq 0$. (33e)–(33f) use the fact that $t > 0$ is small and $\sum_{u_i \neq 0} \text{sign}(u_i)x_i \neq 0$. They also use the fact that $|a+b| = |a| + \text{sign}(a)b$ if $|a| > |b|$. (33g)–(33h) are due to $\text{sign}(\theta) = \text{sign}(\sum_{u_i \neq 0} \text{sign}(u_i)x_i)$ and $\text{sign}(\theta)\theta = |\theta|$. It also uses the fact the multiplying a number by $1-t > 0$ doesn't change its sign. The term that cancels out in (33g) is due to (29a). Finally, (33j) uses the definition of f and the fact that $y = 0$. It implies that (h, k) is a direction of descent. We conclude that any point (x, y) satisfying (29a) is a saddle point and that it admits a direction of descent towards the global minimum $(u\theta, v/\theta)$. The same argument applies to (29b), namely $x = 0$ and $|\sum_{v_j \neq 0} \text{sign}(v_j)y_j| = \sum_{v_j=0} |y_j|$.

We now treat the remaining case (29c)–(29d), namely $\sum_{u_i \neq 0} \text{sign}(u_i)x_i = \sum_{v_j \neq 0} \text{sign}(v_j)y_j = 0$, $x_i y_j / (u_i v_j) \leq 1$ if $u_i v_j \neq 0$, $x_i = 0$ if $u_i = 0$, and $y_j = 0$ if $v_j = 0$. Given $\theta \neq 0$, consider the direction $(u\theta, v/\theta) - (x, y)$, which goes from (x, y) towards the global minimum $(u\theta, v/\theta)$. As we explain below, one can choose θ such that, when taking a small step $t > 0$ in this direction, the ratio inequalities in (29c) remain valid. In other words, we have

$$\frac{(x_i + h_i)(y_j + k_j)}{u_i v_j} \leq 1, \quad \text{if } u_i v_j \neq 0. \quad (34)$$

where $(h, k) := t(u\theta - x, v/\theta - y)$ and $t > 0$ is small enough. In order to prove this, observe that (34) is equivalent to

$$\left((1-t) \frac{x_i}{u_i} + t\theta \right) \left((1-t) \frac{y_j}{v_j} + t/\theta \right) \leq 1, \quad \text{if } u_i v_j \neq 0, \quad (35)$$

where we simply use the definition of (h, k) . For all indices i and j where $x_i y_j / (u_i v_j) < 1$, the inequality in (35) holds by continuity for t small enough, regardless of θ . Hence, if there are no binding inequalities $x_i y_j / (u_i v_j) = 1$, then one may choose any $\theta \neq 0$. Otherwise, the set of binding inequalities $(x_i/u_i)(y_j/v_j) = 1$ can be decomposed into two groups: those with positive ratios (i.e. $x_i/u_i > 0$ and $y_j/v_j > 0$) and those with negative ratios (i.e. $x_i/u_i < 0$ and $y_j/v_j < 0$). Without loss

of generality, we may assume that there exists at least one positive binding ratio. As it turns out, all the positive ratios x_i/u_i involved in binding inequalities are equal to one another. Indeed, if $(x_i/u_i)(y_j/v_j) = 1$ and $(x_k/u_k)(y_l/v_l) = 1$, then the inequalities $(x_i/u_i)(y_l/v_l) \leq 1$ and $(x_k/u_k)(y_j/v_j) \leq 1$ yield $x_i/u_i \leq x_k/u_k$ and $x_k/u_k \leq x_i/u_i$, that is to say, $x_k/u_k = x_i/u_i$. Let $\theta > 0$ denote this common ratio. Now consider a binding inequality in (35): if $x_i/u_i > 0$, then $x_i/u_i = \theta$ and $y_j/v_j = 1/\theta$, so that the inequality in (35) readily holds. If $x_i/u_i < 0$, then $x_i/u_i < (1-t)x_i/u_i + t\theta < 0$ and $y_j/v_j < (1-t)y_j/v_j + t\theta < 0$ for $t > 0$ small enough, hence $[(1-t)x_i/u_i + t\theta][(1-t)y_j/v_j + t\theta] \leq x_i y_j / (u_i v_j) \leq 1$. Again, the inequality in (35) holds. As a result, we have found θ for which (35) is true with $t > 0$ small enough. We next make use of this to compute $f(x + h, y + k) = \dots$

$$= \sum_{i=1}^m \sum_{j=1}^n |(x_i + h_i)(y_j + k_j) - u_i v_j| \quad (36a)$$

$$= \sum_{u_i \neq 0} \sum_{v_j \neq 0} |(x_i + h_i)(y_j + k_j) - u_i v_j| \quad (36b)$$

$$= \sum_{u_i \neq 0} \sum_{v_j \neq 0} |u_i v_j| \left| \frac{(x_i + h_i)(y_j + k_j)}{u_i v_j} - 1 \right| \quad (36c)$$

$$= \sum_{u_i \neq 0} \sum_{v_j \neq 0} |u_i v_j| \left(1 - \frac{(x_i + h_i)(y_j + k_j)}{u_i v_j} \right) \quad (36d)$$

$$= \sum_{u_i \neq 0} \sum_{v_j \neq 0} |u_i v_j| - \sum_{u_i \neq 0} \sum_{v_j \neq 0} \text{sign}(u_i v_j)(x_i + h_i)(y_j + k_j) \quad (36e)$$

$$= \sum_{u_i v_j \neq 0} \sum_{v_j \neq 0} |u_i v_j| - \sum_{u_i \neq 0} \text{sign}(u_i)(x_i + h_i) \sum_{v_j \neq 0} \text{sign}(v_j)(y_j + k_j) \quad (36f)$$

$$= \sum_{u_i \neq 0} \sum_{v_j \neq 0} |u_i v_j| - \left((1-t) \sum_{u_i \neq 0} \text{sign}(u_i) x_i + t \sum_{u_i \neq 0} \text{sign}(u_i) u_i \theta \right) \times \quad (36g)$$

$$\left((1-t) \sum_{v_j \neq 0} \text{sign}(v_j) y_j + t \sum_{v_j \neq 0} \text{sign}(v_j) v_j / \theta \right) \quad (36h)$$

$$= \sum_{u_i \neq 0} \sum_{v_j \neq 0} |u_i v_j| - t^2 \sum_{u_i \neq 0} \text{sign}(u_i) u_i \sum_{v_j \neq 0} \text{sign}(v_j) v_j \quad (36i)$$

$$= (1-t^2) f(x, y). \quad (36j)$$

Above, (36a) is a consequence of the definition of f . By definition of h , we have $x_i + h_i = (1-t)x_i + t u_i$. According to (29d), $x_i = 0$ whenever $u_i = 0$, hence we also have that $x_i + h_i = 0$ whenever $u_i = 0$. Likewise $y_j + k_j = 0$ whenever $v_j = 0$, and thus (36b) holds. (36c) is obtained by factorizing each term in the sum by $|u_i v_j|$. (34) implies that the term inside the absolute value is non-positive, hence (36d). (36e) is the result of expanding the product inside the sum and the fact that $\text{sign}(a) = |a|/a$

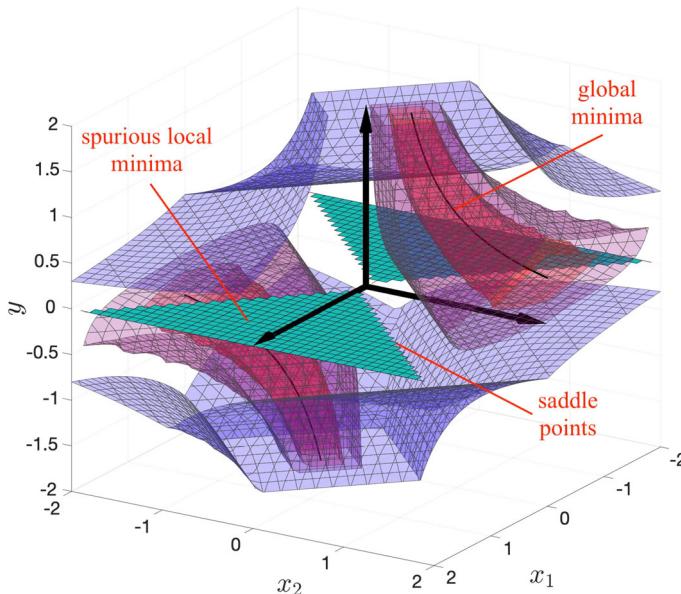


Fig. 5 Landscape of $f(x_1, x_2, y) := |x_1 y| + |x_2 y - 1|$

when $a \neq 0$. (36f) is obtained by factorizing the second term in (36e). (36g)–(36h) uses the definition of (h, k) , namely that $h_i = tu_i\theta - tx_i$ and $k_j = tv_j/\theta - y_j$. (36i) follows from the equalities in (29c), which results in two terms cancelling out. (36j) is due to the fact that when $t = 0$, the resulting expression must be equal to $f(x, y)$ since we are computing $f(x + t(u\theta - x), y + t(v/\theta - y))$. We conclude that (x, y) is a saddle point and that it admits a direction of descent towards the global minimum $(u\theta, v/\theta)$. \square

We can now deduce Theorem 1 from Proposition 1. Recall that we use the convention that a sum over an index set which is empty is equal to zero. For example, if none of the entries of u are zero, then $\sum_{u_i=0} |x_i| = 0$. In this case, (28a) is not feasible since an absolute value cannot be negative. It follows from (28a)–(28b) that there are no spurious local minima if neither of the entries of u nor v are equal to zero. This condition on u and v is equivalent to saying the uv^T has no zero entries. Conversely, if some entries of either u or v are equal to zero, then one can readily see that (28a)–(28b) is feasible. In the case where all the entries of uv^T are equal to zero, it follows from the expression of $f(x, y) = \|x\|_1\|y\|_1$ that every local minimum is a global minimum. Theorem 1 naturally ensues.

We finish this section with an example. When $u = (0, 1)^T$ and $v = 1$, the landscape is represented in Fig. 5. According to (27), the global minima are all $(x_1, x_2, y) \in \mathbb{R}^3$ such that $(x_1, x_2, y) = (0, \theta, 1/\theta)$ where $\theta \neq 0$. This corresponds to the two black hyperbolic branches. The surfaces around the global minima denote level sets of f , where the warmer the color (from blue to red), the smaller the objective value. Among the three level sets in the figure, the one with highest objective value (in blue) has a part which is not represented, namely, its intersection with the positive orthant. This

is done in order to improve visibility. According to (28a)–(28b), the spurious local minima are all $(x_1, x_2, y) \in \mathbb{R}^3$ such that $|x_2| < |x_1|$ and $y = 0$. This corresponds to the area inside the two green triangles. According to (29a)–(29d), the saddle points are all $(x_1, x_2, y) \in \mathbb{R}^3$ such that $|x_2| = |x_1|$ and $y = 0$. This corresponds to the edges of the two green triangles located on the diagonal and anti-diagonal of the (x_1, x_2) -plane. Notice that the union of the global minima, the spurious local minima, and the saddle points in Fig. 5 agrees with the expression of the critical points found with a commercial solver in Fig. 1.

Acknowledgements The authors acknowledge support from DTRA grant 13-1-0021, DARPA grant Lagrange, NSF grant 2023032, and ONR grant N00014-21-1-2282. Data sharing not applicable to this article as no datasets were generated or analysed during the current study

References

1. Anderson, B. G., Sojoudi, S.: Global optimality guarantees for nonconvex unsupervised video segmentation. Annual Allerton Conference on Communication, Control, and Computing, (2019)
2. Baldi, P., Hornik, K.: neural networks and principal component analysis: learning from examples without local minima. *Neural Netw.* **2**(1), 53–58 (1989)
3. Beck, A.: First-order methods in optimization. SIAM, New Delhi (2017)
4. BERTSIMAS, D., Cory-Wright, R., Pauphilet, J.: Solving large-scale sparse pca to certifiable (near) optimality. arXiv preprint [arXiv:2005.05195](https://arxiv.org/abs/2005.05195), (2020)
5. Bouwmans, T., Sobral, A., Javed, S., Jung, S.K., Zahzah, E.-H.: Decomposition into low-rank plus additive matrices for background/foreground separation: A review for a comparative evaluation with a large-scale dataset. *Comput. Sci. Rev.* **23**, 1–71 (2017)
6. Burke, J.V., Ferris, M.C.: Weak sharp minima in mathematical programming. *SIAM J. Control Optim.* **31**(5), 1340–1359 (1993)
7. Candès, E.J., Li, X., Ma, Y., Wright, J.: Robust principal component analysis? *J. ACM (JACM)* **58**(3), 1–37 (2011)
8. Candès, E.J., Tao, T.: The power of convex relaxation: Near-optimal matrix completion. *IEEE Trans. Inf. Theory* **56**(5), 2053–2080 (2010)
9. Chandrasekaran, V., Sanghavi, S., Parrilo, P. A., Willsky, A. S.: Rank-sparsity incoherence for matrix decomposition. *SIOPT*, (2011)
10. Charisopoulos, V., Chen, Y., Davis, D., Díaz, M., Ding, L., Drusvyatskiy, D.: Low-rank matrix recovery with composite optimization: good conditioning and rapid convergence. *Found. Comput. Math.*, pp. 1–89, 2021
11. Chen, Y., Chi, Y.: Harnessing structures in big data via guaranteed low-rank matrix estimation: Recent theory and fast algorithms via convex and nonconvex optimization. *IEEE Signal Process. Mag.* **35**(4), 14–31 (2018)
12. Chi, Y., Lu, Y.M., Chen, Y.: Nonconvex optimization meets low-rank matrix factorization: an overview. *IEEE Trans. Signal Process.* **67**(20), 5239–5269 (2019)
13. Clarke, F. H.: Optimization and nonsmooth analysis. SIAM Classics Appl. Math., 1990
14. Du, Simon S., Hu, Wei, Lee, Jason D.: Algorithmic regularization in learning deep homogeneous models: layers are automatically balanced. *NeurIPS* 2018, (2018)
15. Fattah, S., Sojoudi, S.: Exact guarantees on the absence of spurious local minima for non-negative rank-1 robust principal component analysis. *JMLR*, 2020
16. Garcia-Garcia, B., Bouwmans, T., Silva, A.J.R.: Background subtraction in real applications: challenges, current models and future directions. *Comput. Sci. Rev.* **35**, 100–204 (2020)
17. Gillis, Nicolas, Vavasis, Stephen A.: On the complexity of robust PCA and 11-norm low-rank matrix approximation. *Math. Op. Res.* **43**(4), 1072–1084 (2018)
18. Gross, D.: Recovering low-rank matrices from few coefficients in any basis. *IEEE Trans. Inf. Theory*, (2011)
19. Ioffe, A.D.: An invitation to tame optimization. *SIAM J. Optim.* **19**(4), 1894–1917 (2009)

20. Jain, P., Kar, P.: Non-convex optimization for machine learning. *Found Trends Mach. Learn.*, (2017)
21. Josz, C., Ouyang, Y., Zhang, R. Y., Lavaei, J., Sojoudi, S.: A theory on the absence of spurious solutions for nonconvex and nonsmooth optimization. *NeurIPS*, (2018)
22. Koren, Y., Bell, R., Volinsky, C.: Matrix factorization techniques for recommender systems. *Comput.* (2009)
23. Levy, O., Goldberg, Y.: Neural word embedding as implicit matrix factorization. *NeurIPS*, (2014)
24. Li, X., Zhu, Z., Man-Cho So, A., Vidal, R.: Nonconvex robust low-rank matrix recovery. *SIAM J. Optim.* (2019)
25. Li, Xiao, Zhu, Zhihui, Man-Cho So, Anthony, Lee, Jason D.: Incremental methods for weakly convex optimization. 12th Annual Workshop on Optimization for Machine Learning, (2020)
26. Moitra, A.: Algorithmic aspects of machine learning. Cambridge University Press, Cambridge (2018)
27. Nesterov, Yu.: Smooth minimization of non-smooth functions. *Math. Program.* **103**(1), 127–152 (2005)
28. Nguyen, L.T., Kim, J., Shim, B.: Low-rank matrix completion: a contemporary survey. *IEEE Access* **7**, 94215–94237 (2019)
29. Recht, B.: A simpler approach to matrix completion. *JMLR*, (2011)
30. Seidenberg, A.: A new decision method for elementary algebra. *Annals Math.*, 365–374, (1954)
31. Tarski, A.: A decision method for elementary algebra and geometry: prepared for publication with the assistance of JCC McKinsey. (1951)
32. Valavi, H., Liu, S., Ramadge, P. J.: The landscape of matrix factorization revisited. *AISTATS*, (2020)

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