Dual Randomized Coordinate Descent Method for Solving a Class of Nonconvex Problems*

Amir Beck† Marc Teboulle ‡

August 31, 2020

Abstract

We consider a nonconvex optimization problem consisting of maximizing the difference of two convex functions. We present a randomized method that requires low computational effort at each iteration. The described method is a randomized coordinate descent method employed on the so-called Toland-dual problem. We prove subsequence convergence to dual stationarity points, a new notion that we introduce and shown to be tighter than the standard criticality. Almost sure rate of convergence of an optimality measure of the dual sequence is proven. We demonstrate the potential of our results on three Principal Component Analysis (PCA) models resulting in extremely simple algorithms.

1 Introduction

This paper is concerned with nonconvex maximization problems of the form

$$\max_{\mathbf{x}\in\mathbb{R}^d}\{f(\mathbf{A}\mathbf{x})-g(\mathbf{x})\}, \tag{1.1}$$

where f and g are extended real-valued convex functions such that the domain of g is compact and $\mathbf{A}\in\mathbb{R}^{n\times d}$. The above model belongs to the class difference of convex (DC) programming problems, which is a well-studied class of problems encompassing a wide variety of applications, see for example the classical overview paper [4] as well as [9] for recent mathematical developments and references.

Our motivation for looking at the model (1.1) comes from several dimensionality reduction applications such as Principal Component Analysis (PCA), sparse PCA and a novel robust PCA model. The specific details as well as references are provided in Section 4. For now, we just mention the classical PCA problem that has the form

$$\max\left\{\frac{1}{2}\|\mathbf{A}\mathbf{x}\|_2^2 : \|\mathbf{x}\|_2\leq 1\right\}.$$

*The research of A. Beck was partially supported by the Israel Science Foundation Grant 1821/16. The research of M. Teboulle was partially supported by the Israel Science Foundation under ISF Grant 1844/16

†School of Mathematical Sciences, Tel Aviv University, becka@tauex.tau.ac.il

‡School of Mathematical Sciences, Tel Aviv University, teboulle@tauex.tau.ac.il

The above problem fits model (1.1) with

$$f(\mathbf{z}) = \frac{1}{2}\|\mathbf{z}\|_2^2, \quad g(\mathbf{x}) = \begin{cases} 0 & \|\mathbf{x}\|_2 \le 1, \\ \infty & \text{else.} \end{cases}$$

The main objective of the manuscript is to develop algorithms for solving problem (1.1) that are

(A) able to tackle large-scale instances (i.e., involving simple computations);

(B) converge (in some sense) to favorable optimality conditions.

As for objective (A), we will be specifically interested in defining randomized algorithms that require simple operations at each iteration taking into account only a single row of the matrix $\mathbf{A}$. For the specific case of the PCA problem, such algorithms exists. More specifically, several variants of the stochastic gradient method were defined in the literature, starting from the famous Oja's method [8], all of them have the form

$$\mathbf{x}^{k+1} = \frac{\tilde{\mathbf{x}}^{k+1}}{\|\tilde{\mathbf{x}}^{k+1}\|_2}, \text{ where } \tilde{\mathbf{x}}^{k+1} = \mathbf{x}^k + t_k \mathbf{a}_{i_k},$$

where $t_k > 0$ is a stepsize, $\mathbf{a}_1, \ldots, \mathbf{a}_n$ are the rows of $\mathbf{A}$ and $i_k \in [n]$ is a randomly chosen index, see also the more recent study [11]. Our approach for tackling the general model (1.1) is completely different, as it is actually a dual-based approach. Our starting point is the Toland-dual problem [12, 13] associated with (1.1). We then apply a randomized coordinate descent method on the devised dual problem and obtain an algorithm that, much like the above mentioned algorithms for the PCA problem, only involves simple operations related to a *single* row of $\mathbf{A}$ at each iteration. For example, in the specific case of the PCA problem the general update rule of our algorithm has the form $\mathbf{z}^{k+1} = \mathbf{z}^k + s_k \mathbf{a}_{i_k}$ (without normalization!) and s_k is chosen based on data from previous iterations and a solution of a scalar quartic equation. The output of the algorithm will be a normalization of the last iterate (the exact details are in Section 6.1).

The standard necessary optimality condition in DC programming is criticality [4]. We will show that our proposed class of algorithms are able to converge almost surely to *dual stationary* points, which satisfy a stronger condition, and thus our objective (B) will also be accomplished.

Paper Layout. Section 2 reviews well-known optimality conditions such as criticality and stationarity and their relations. Section 3 presents the main model and its assumptions, as well as the Toland-dual problem; the dual-stationarity optimality condition is introduced and studied. The three prototype models - all related to dimensionality reduction - are discussed in Section 4. The randomized dual coordinate descent is studied in Section 5, where almost sure subsequence convergence is proven under a bounded level sets assumption. It is shown how this assumption can be simplified using an analysis based on asymptotic functions. The section ends with an analysis of the case where g is the indicator of the ℓ_2-ball (relevant, for example, in two of the prototype problems). For this case, it is shown how the nondifferentiability of the dual problem can be avoided altogether and as a result, it is possible to establish an $O(1/k)$ rate of convergence for the expected sequence of the squared norms of the gradients of the dual sequence. The paper ends with a specification of the dual randomized coordinate descent method for the three prototype PCA models resulting in extremely simple algorithms.

Notation. Vectors are denoted by boldface lowercase letters, e.g., $\mathbf{y}$, and matrices by boldface uppercase letters, e.g., $\mathbf{B}$. The vectors of all zeros and ones are denoted by $\mathbf{0}$ and $\mathbf{e}$ respectively. The canonical basis of $\mathbb{R}^n$ is denoted by $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$. The n-dimensional unit-simplex set is given by $\Delta_n = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{e}^T\mathbf{x} = 1, \mathbf{x} \geq \mathbf{0}\}$. The closed ℓ_2-norm ball with center $\mathbf{c} \in \mathbb{R}^n$ and center $r > 0$ is denoted by $B_2[\mathbf{c}, r] = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{c}\|_2 \leq r\}$. We use the standard notation $[n] \equiv \{1, 2, \ldots, n\}$ for a positive integer n. For a given function h, $h'(\mathbf{x}; \mathbf{d})$ denotes the directional derivative of h at $\mathbf{x} \in \mathbb{R}^n$ in the direction $\mathbf{d} \in \mathbb{R}^n$. For an extended real-valued convex function h and a point $\mathbf{x} \in \text{dom}(h)$, we denote the subdifferential set of h at $\mathbf{x}$, meaning the set of all subgradients of h at $\mathbf{x}$ as $\partial h(\mathbf{x}) = \{\mathbf{g} \in \mathbb{R}^n : h(\mathbf{y}) \geq h(\mathbf{x}) + \langle \mathbf{g}, \mathbf{y} - \mathbf{x} \rangle \ \forall \mathbf{y} \in \text{dom}(h)\}$. For any extended real-valued function h, the conjugate is defined as $h^*(\mathbf{y}) \equiv \max_{\mathbf{x}} \{\langle \mathbf{x}, \mathbf{y} \rangle - h(\mathbf{x})\}$.

2 Prelude on Optimality Conditions

This section contains several well-known results related to optimality conditions that will be used later on in our analysis.

2.1 Stationarity

One of the most natural first-order optimality conditions is *stationarity*, which in this paper *always* means a point with no feasible descent directions.

Definition 2.1 (stationarity)**.** Suppose that $F : \mathbb{R}^n \to (-\infty, \infty]$ is an extended real-valued function that satisfies (1) $\text{dom}(F)$ is a convex set and (2) it has directional derivatives at all points in its domain in all feasible directions. Then $\bar{\mathbf{x}} \in \text{dom}(F)$ is a **stationary** point of the problem $\min_{\mathbf{x}} F(\mathbf{x})$ if it satisfies:

$$F'(\bar{\mathbf{x}}; \mathbf{y} - \bar{\mathbf{x}}) \geq 0 \text{ for all } \mathbf{y} \in \text{dom}(F).$$

Obviously, stationarity is a necessary optimality condition. A famous example in which the stationarity condition can be written explicitly is the standard additive composite model which consists of minimizing the sum of a differentiable (not necessarily convex) function and a proper closed and convex function.

Lemma 2.1 (stationarity in the standard composite model)**.** *Consider the problem*

$$\min_{\mathbf{x} \in \mathbb{R}^n} F(\mathbf{x}) + G(\mathbf{x}), \tag{2.1}$$

where $F : \mathbb{R}^n \to \mathbb{R}$ is a nonconvex differentiable function and $G : \mathbb{R}^n \to (-\infty, \ \infty]$ is a proper closed and convex function. Then $\bar{\mathbf{x}} \in \text{dom}(G)$ is a stationary point of (2.1) if and only if

$$-\nabla F(\bar{\mathbf{x}}) \in \partial G(\bar{\mathbf{x}}).$$

2.2 Criticality in DC Programming

Consider the optimization problem

$$\text{(M)} \quad \min_{\mathbf{x}\in\mathbb{R}^n} s(\mathbf{x}) - t(\mathbf{x}),$$

where

- $s : \mathbb{R}^n \to (-\infty, \infty]$ is proper closed and convex;
- $t : \mathbb{R}^n \to (-\infty, \infty]$ is proper closed convex and subdifferetiable over its domain.
- $\mathrm{dom}(s) \subseteq \mathrm{dom}(t)$.

Problem (M) is a difference of convex (DC) programming problem, see for example the overview paper [4] for results and applications. To make problem (M) well-defined, we will define $\infty - \infty$ to be equal to ∞. The convention will be assumed only in the context of minimization problems. A natural first-order optimality condition for problem (M) is *criticality.*

Definition 2.2 (criticality)**.** A point $\bar{\mathbf{x}} \in \mathrm{dom}(s)$ is a **critical point** of problem (M) if

$$\partial s(\bar{\mathbf{x}}) \cap \partial t(\bar{\mathbf{x}}) \neq \emptyset.$$

A well-known result states that a necessary optimality condition for problem (M) is criticality. For the sake of completeness, we will provide the extremely short proof of the result.

Theorem 2.1. *Let $\bar{\mathbf{x}}$ be an optimal solution of (M). Then $\bar{\mathbf{x}}$ is a critical point of (M).*

Proof. Suppose that $\bar{\mathbf{x}}$ is an optimal solution of (M). Then necessarily $\bar{\mathbf{x}} \in \mathrm{dom}(s)$, and hence $\bar{\mathbf{x}} \in \mathrm{dom}(t) = \mathrm{dom}(\partial(t))$. By the optimality of $\bar{\mathbf{x}}$, it follows that

$$s(\mathbf{x}) - t(\mathbf{x}) \geq s(\bar{\mathbf{x}}) - t(\bar{\mathbf{x}}) \text{ for all } \mathbf{x} \in \mathrm{dom}(s). \tag{2.2}$$

Since $\bar{\mathbf{x}} \in \mathrm{dom}(\partial(t))$, it follows that there exists $\mathbf{g} \in \partial t(\bar{\mathbf{x}})$. This means in particular that $t(\mathbf{x}) - t(\bar{\mathbf{x}}) \geq \langle \mathbf{g}, \mathbf{x} - \bar{\mathbf{x}} \rangle$, which combined with (2.2) leads to the conclusion that

$$s(\mathbf{x}) \geq s(\bar{\mathbf{x}}) + \langle \mathbf{g}, \mathbf{x} - \bar{\mathbf{x}} \rangle \text{ for all } \mathbf{x} \in \mathrm{dom}(s),$$

meaning that $\mathbf{g} \in \partial s(\bar{\mathbf{x}})$, i.e., that $\mathbf{g} \in \partial s(\bar{\mathbf{x}}) \cap \partial t(\bar{\mathbf{x}})$, thus establishing the desired result. □

Remark 2.1. We note that criticality at a point $\bar{\mathbf{x}}$ is not equivalent to stationarity, meaning the lack of feasible descent directions at $\bar{\mathbf{x}}$, see for example the extensive study of optimality conditions in DC programming presented in [9]. In [3], the case where s is in addition differentiable over $\mathbb{R}^n$ was considered. Criticality in this case takes the form:

$$\nabla s(\bar{\mathbf{x}}) \in \partial t(\bar{\mathbf{x}}).$$

It was shown that in this scenario (where s is differentiable), the stronger condition

$$\{\nabla s(\bar{\mathbf{x}})\} = \partial t(\bar{\mathbf{x}})$$

is equivalent to stationarity. This implies in particular that t is differentiable at any stationary point. It was also shown in [3] that such a point can be generated by either a greedy or randomized coordinate descent method.

3 Problem Formulation and Duality

We consider the nonconvex problem

$$\text{(P)} \quad \max_{\mathbf{x}\in\mathbb{R}^d}\{f(\mathbf{A}\mathbf{x}) - g(\mathbf{x})\},$$

where $\mathbf{A} \in \mathbb{R}^{n\times d}$ and f, g satisfy the following properties:

Assumption 1. • *$f : \mathbb{R}^n \to \mathbb{R}$ is strongly convex;*

- *$g : \mathbb{R}^d \to (-\infty, \infty]$ is proper closed and convex with a compact domain on which it is subdifferentiable;*
- $\mathrm{dom}(g) \subseteq \mathrm{dom}(h)$, *where $h = f \circ \mathbf{A}$ is the function $h(\mathbf{x}) \equiv f(\mathbf{A}\mathbf{x})$.*

Note that here, in the context of a *maximization* problem, our convention is that $\infty-\infty = -\infty$. By the previous section, $\bar{\mathbf{x}} \in \mathrm{dom}(g)$ is a critical point of (P) if

$$\mathbf{A}^T\partial f(\mathbf{A}\bar{\mathbf{x}}) \cap \partial g(\bar{\mathbf{x}}) \neq \emptyset.$$

We first derive the so-called Toland-dual problem of (P) [12, 13]. For that, we note that since f is proper closed and convex, it holds that $f^{**} = f$, and therefore, in particular, for any $\mathbf{x} \in \mathbb{R}^d$, $f(\mathbf{A}\mathbf{x}) = f^{**}(\mathbf{A}\mathbf{x}) = \max_{\mathbf{y}\in\mathbb{R}^n} \{\langle \mathbf{A}\mathbf{x}, \mathbf{y}\rangle - f^*(\mathbf{y})\}$. Consequently, problem (P) can be equivalently written as

$$\max_{\mathbf{x}\in\mathbb{R}^d}\max_{\mathbf{y}\in\mathbb{R}^n}\{\langle \mathbf{A}\mathbf{x}, \mathbf{y}\rangle - f^*(\mathbf{y}) - g(\mathbf{x})\}.$$

Exchanging the order of maximizations yields the equivalent problem

$$\max_{\mathbf{y}\in\mathbb{R}^n}\max_{\mathbf{x}\in\mathbb{R}^d}\{\langle \mathbf{A}\mathbf{x}, \mathbf{y}\rangle - f^*(\mathbf{y}) - g(\mathbf{x})\}.$$

The set of maximizers of the inner problem with respect to $\mathbf{x}$ is $\partial g^*(\mathbf{A}^T\mathbf{y})$, and the maximal value is $g^*(\mathbf{A}^T\mathbf{y})-f^*(\mathbf{y})$. We thus obtain the following problem, known as the "Toland-dual" problem:

$$\text{(D)} \quad \max_{\mathbf{y}\in\mathbb{R}^n}\{q(\mathbf{y}) \equiv g^*(\mathbf{A}^T\mathbf{y}) - f^*(\mathbf{y})\}.$$

Note also that given an optimal solution $\bar{\mathbf{x}}$ of (P), the set of maximizers of the problem $\max_{\mathbf{y}}\{\langle \mathbf{A}\bar{\mathbf{x}}, \mathbf{y}\rangle - f^*(\mathbf{y})\}$, which is exactly $\partial f(\mathbf{A}\bar{\mathbf{x}})$, are all optimal solutions of (D). We summarize the above discussion.

Lemma 3.1. *Suppose that f and g satisfy Assumption 1. If $\bar{\mathbf{y}}$ is an optimal solution of (D), and $\bar{\mathbf{x}} \in \partial g^*(\mathbf{A}^T\bar{\mathbf{y}})$, then $\bar{\mathbf{x}}$ is an optimal solution of (P). If $\bar{\mathbf{x}}$ is an optimal solution of (P), then any $\bar{\mathbf{y}} \in \partial f(\mathbf{A}\bar{\mathbf{x}})$ is an optimal solution of (D).*

By Assumption 1, f^* and g^* are are both real-valued and convex, and in addition f^* is differentiable.

By definition 2.2, a vector $\bar{\mathbf{y}} \in \mathbb{R}^n$ is a *critical* point of (D) if $\nabla f^*(\bar{\mathbf{y}}) \in \mathbf{A}\partial g^*(\mathbf{A}^T\bar{\mathbf{y}})$; by Remark 2.1 it is a *stationary* point of (D) if $\mathbf{A}\partial g^*(\mathbf{A}^T\bar{\mathbf{y}}) = \{\nabla f^*(\bar{\mathbf{y}})\}$. The following result establishes some relations between critical/stationary points of (D) and (P).

Theorem 3.1. *Suppose that Assumption 1 holds.*

(a) Let $\bar{\mathbf{y}}$ *be a critical point of (D). Then there exists* $\bar{\mathbf{x}} \in \partial g^*(\mathbf{A}^T\bar{\mathbf{y}})$ *satisfying* $\nabla f^*(\bar{\mathbf{y}}) = \mathbf{A}\bar{\mathbf{x}}$*, and any such* $\bar{\mathbf{x}}$ *is a critical point of (P).*

(b) Let $\bar{\mathbf{y}}$ *be a stationary point of (D). Then any* $\bar{\mathbf{x}} \in \partial g^*(\mathbf{A}^T\bar{\mathbf{y}})$ *is a critical point of (P).*

Proof. (a) If $\bar{\mathbf{y}}$ is a critical point of (D), then $\nabla f^*(\bar{\mathbf{y}}) \in \mathbf{A}\partial g^*(\mathbf{A}^T\bar{\mathbf{y}})$, which means that there exists $\bar{\mathbf{x}}$ for which

$$\bar{\mathbf{x}} \in \partial g^*(\mathbf{A}^T\bar{\mathbf{y}}), \nabla f^*(\bar{\mathbf{y}}) = \mathbf{A}\bar{\mathbf{x}}. \tag{3.1}$$

The relations (3.1) are equivalent to $\mathbf{A}^T\bar{\mathbf{y}} \in \partial g(\bar{\mathbf{x}})$ and $\bar{\mathbf{y}} \in \partial f(\mathbf{A}\bar{\mathbf{x}})$. Consequently, $\mathbf{A}^T\bar{\mathbf{y}} \in \mathbf{A}^T\partial f(\mathbf{A}\bar{\mathbf{x}}) \cap \partial g(\bar{\mathbf{x}})$, and in particular $\mathbf{A}^T\partial f(\mathbf{A}\bar{\mathbf{x}}) \cap \partial g(\bar{\mathbf{x}}) \neq \emptyset$, establishing that $\bar{\mathbf{x}}$ is a critical point of (P).
(b) If $\bar{\mathbf{y}}$ is a stationary point of (D), and $\bar{\mathbf{x}} \in \partial g^*(\mathbf{A}^T\bar{\mathbf{y}})$, then since $\mathbf{A}\partial g^*(\mathbf{A}^T\bar{\mathbf{y}}) = \{\nabla f^*(\bar{\mathbf{y}})\}$, it follows that $\nabla f^*(\bar{\mathbf{y}}) = \mathbf{A}\bar{\mathbf{x}}$, and the same argument from part (a) shows that $\bar{\mathbf{x}}$ is a critical point of (P). □

Since stationarity is a more restrictive optimality condition than criticality, it is natural to define a notion of "dual-stationarity" condition that is stronger than criticality.

Definition 3.1 (dual-stationarity)**.** A point $\bar{\mathbf{x}} \in \mathrm{dom}(g)$ is called a **dual-stationary** point of (P) if $\bar{\mathbf{x}} \in \partial g^*(\mathbf{A}^T\bar{\mathbf{y}})$ for some stationary point $\bar{\mathbf{y}} \in \mathbb{R}^m$ of (D).

Lemma 3.2 (optimality $\Rightarrow$ dual-stationarity $\Rightarrow$ criticality)**.** *Suppose that Assumption 1 holds.*

(a) If $\bar{\mathbf{x}}$ *is an optimal solution of (P), then it is a dual-stationary point of (P).*

(b) If $\bar{\mathbf{x}}$ *is a dual-stationary point of (P), then it is a critical point of (P).*

Proof. (a) Suppose that $\bar{\mathbf{x}}$ is an optimal solution of problem (P). In particular, by Theorem 2.1, $\bar{\mathbf{x}}$ is a critical point of (P), and thus, $\mathbf{A}^T\partial f(\mathbf{A}\bar{\mathbf{x}}) \cap \partial g(\bar{\mathbf{x}}) \neq \emptyset$. We can therefore pick $\bar{\mathbf{y}} \in \partial f(\mathbf{A}\bar{\mathbf{x}})$ such that $\mathbf{A}^T\bar{\mathbf{y}} \in \partial g(\bar{\mathbf{x}})$. Since $\bar{\mathbf{y}} \in \partial f(\mathbf{A}\bar{\mathbf{x}})$, then by Lemma 3.1, it is also an optimal solution of (D), and thus also a stationary point of (D). Finally, the inclusion $\mathbf{A}^T\bar{\mathbf{y}} \in \partial g(\bar{\mathbf{x}})$ is equivalent to the relation $\bar{\mathbf{x}} \in \partial g^*(\mathbf{A}^T\bar{\mathbf{y}})$, and by definition, this means that $\bar{\mathbf{x}}$ is a dual-stationary point of (P).
(b) Suppose that $\bar{\mathbf{x}}$ is a dual-stationary point. Then

$$\bar{\mathbf{x}} \in \partial g^*(\mathbf{A}^T\bar{\mathbf{y}}) \tag{3.2}$$

for some stationary point $\bar{\mathbf{y}}$ of (D), which means that $\mathbf{A}\partial g^*(\mathbf{A}^T\bar{\mathbf{y}}) = \{\nabla f^*(\bar{\mathbf{y}})\}$, and thus, in particular, $\mathbf{A}\bar{\mathbf{x}} = \nabla f^*(\bar{\mathbf{y}})$, and therefore, $\bar{\mathbf{y}} \in \partial f(\mathbf{A}\bar{\mathbf{x}})$, implying that $\mathbf{A}^T\bar{\mathbf{y}} \in \mathbf{A}^T\partial f(\mathbf{A}\bar{\mathbf{x}})$. This, combined with the fact that (3.2) is the same as $\mathbf{A}^T\bar{\mathbf{y}} \in \partial g(\bar{\mathbf{x}})$, implies that $\mathbf{A}^T\partial f(\mathbf{A}\bar{\mathbf{x}}) \cap \partial g(\bar{\mathbf{x}}) \neq \emptyset$, meaning that $\bar{\mathbf{x}}$ is a critical point. □

The next example shows that dual-stationarity might be a much stronger condition than criticality.

Example 3.1. Consider the problem

$$(P_1) \quad \max_{x_1,x_2} \left\{ \frac{1}{2}(x_1+x_2)^2 : |x_1| \le 1, |x_2| \le 1 \right\}.$$

This problem fits model (P) with $f(t) = \frac{t^2}{2}$, $\mathbf{A} = \begin{pmatrix} 1 & 1 \end{pmatrix}$ and $g(\mathbf{x}) = \delta_{B_\infty[\mathbf{0},1]}(\mathbf{x})$. Note that $f^*(t) = \frac{1}{2}t^2$ and $g^*(\mathbf{x}) = \|\mathbf{x}\|_1$, and thus the one-dimensional dual problem is

$$(D_1) \quad \max_{y} \left\{ 2|y| - \frac{y^2}{2} \right\}.$$

Criticality for the dual problem means that

$$y \in \begin{cases} \{2\} & y > 0, \\ \{-2\} & y < 0, \\ [-2,2] & y = 0, \end{cases}$$

implying that the critical points of (D_1) are $-2, 0, 2$. Stationarity eliminates the point of non-differentiability $y = 0$, and thus the stationary points of (D_1) are $-2, 2$. A plot of the dual objective function is given in Figure 1. As for the primal problem (P_1), the critical points are those satisfying (utilizing the differentiability of f) $\mathbf{A}^T \nabla f(\mathbf{A}\mathbf{x}) \in \partial g(\mathbf{x})$. Substituting the expressions for f, g in this relation, we obtain that it is the same as

$$\begin{pmatrix} x_1 + x_2 \\ x_1 + x_2 \end{pmatrix} \in J(x_1) \times J(x_2), \text{ where } J(t) \equiv \begin{cases} \{0\}, & t \in (-1,1), \\ [0,\infty), & t = 1, \\ (-\infty, 0], & t = -1. \end{cases}$$

The points that satisfy the above relation are

$$\{(x_1,x_2)^T : x_1 + x_2 = 0, |x_1| \le 1, |x_2| \le 1\} \cup \{(-1,-1)^T, (1,1)^T\}.$$

All these points are critical points of (P_1), and actually they are also stationary points of (P_1). Obviously, among these infinite amount of solutions, only $(-1,-1)^T$ and $(1,1)^T$ are optimal solutions of the problem. The two stationary points $y = -2, 2$ of the dual problem (D_1) correspond to the dual-stationary points $(-1,-1)^T$ and $(1,1)^T$ respectively. Meaning that only two critical points out of the infinite amount of critical points are dual-stationary, and in this case, both are global optimal solutions.

4 Three PCA Prototype Problems

We present three PCA-type problems through which we will demonstrate the results of the paper. PCA is a fundamental paradigm for dimensionally reduction. The literature on various PCA models and algorithmic procedures to their solutions is very large, see e.g., the review paper [6] and references therein.

In all examples, we assume that $\mathbf{A} \in \mathbb{R}^{n \times d}$ is a data matrix whose rows $\mathbf{a}_1^T, \mathbf{a}_2^T, \ldots, \mathbf{a}_n^T$ are nonzero d-dimensional vectors (corresponding to d features) which are centered, meaning that $\sum_{i=1}^n \mathbf{a}_i = \mathbf{0}$.

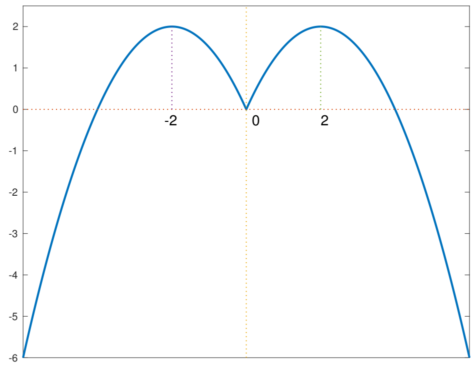

Figure 1: The function $2|y| - \frac{y^2}{2}$ has three critical point $y = -2, 0, 2$. Among them $y = -2, 2$ are stationary points.

4.1 PCA

The PCA problem seeks to find a normalized direction $\mathbf{x} \in \mathbb{R}^d$ for which the projections of the vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ on the subspace spanned by $\mathbf{x}$, meaning $\mathbf{a}_1^T\mathbf{x}, \mathbf{a}_2^T\mathbf{x}, \ldots, \mathbf{a}_n^T\mathbf{x}$, have maximal variation. The problem therefore seeks to find a vector $\mathbf{x} \in \mathbb{R}^d$ that is the solution to

$$\max_{\|\mathbf{x}\|_2=1} \sum_{i=1}^{n} (\mathbf{a}_i^T\mathbf{x})^2. \tag{4.1}$$

The above can also be written as[1]

$$\text{(PCA)} \qquad \max_{\|\mathbf{x}\|_2 \leq 1} \|\mathbf{A}\mathbf{x}\|_2^2.$$

This model fits problem (P) with $f(\cdot) = \frac{1}{2}\|\cdot\|_2^2$, and $g = \delta_{B_2[\mathbf{0},1]}$. The Toland-dual problem is therefore

$$\text{(D-PCA)} \qquad \max_{\mathbf{y} \in \mathbb{R}^n} \left\{ \|\mathbf{A}^T\mathbf{y}\|_2 - \frac{1}{2}\|\mathbf{y}\|_2^2 \right\}.$$

4.2 Sparse PCA

In the *sparse* PCA problem, we further constrain the direction vector to have a bound on the number of nonzero elements of the direction vector. A common formulation of the problem is via the ℓ_0-"norm" defined by $\|\mathbf{x}\|_0 \equiv \#\{i : x_i \neq 0\}$ (see for example [7]) :

$$\text{(SPCA)} \qquad \max\{\|\mathbf{A}\mathbf{x}\|_2^2 : \|\mathbf{x}\|_2 \leq 1, \|\mathbf{x}\|_0 \leq s\},$$

where $s \in [d]$. To show that problem (SPCA) fits model (P), we use the well-known fact that the maximum of a convex function over a nonempty compact and convex set is attained at least at one of its extreme points. Therefore, problem (SPCA) is the same as

$$\max\{\|\mathbf{A}\mathbf{x}\|_2^2 : \mathbf{x} \in \text{conv}(B_2[\mathbf{0}, 1] \cap C_s)\},$$

[1] A simple argument shows that the equality constraint can be replaced by an inequality.

where $B_2[\mathbf{0},1] \equiv \{\mathbf{x} : \|\mathbf{x}\|_2 \leq 1\}$ and $C_s \equiv \{\mathbf{x} : \|\mathbf{x}\|_0 \leq s\}$. The latter formulation of (SPCA) fits model (P) with $f(\cdot) = \frac{1}{2}\|\cdot\|_2^2$ and $g = \delta_{\text{conv}(B_2[\mathbf{0},1]\cap C_s)}$. The conjugate of g is given by

$$\begin{aligned} g^*(\mathbf{z}) &= \max\{\mathbf{x}^T\mathbf{z} : \mathbf{x} \in \text{conv}(B_2[\mathbf{0},1] \cap C_s)\} \\ &= \max\{\mathbf{x}^T\mathbf{z} : \mathbf{x} \in B_2[\mathbf{0},1] \cap C_s\}. \end{aligned} \tag{4.2}$$

We will make the following notation: for $\mathbf{c} \in \mathbb{R}^d$, $T_s(\mathbf{c}) \in \underset{\mathbf{x}}{\text{argmin}}\{\|\mathbf{x}-\mathbf{c}\|_2^2 : \|\mathbf{x}\|_0 \leq s\}$ is an d-length vector that keeps the s values of $\mathbf{c}$ with the largest absolute values and plugs zeros elsewhere. The indices are chosen to be smallest as possible in case of ambiguity. It is easy to show that if $\mathbf{z} \neq \mathbf{0}$, then an optimal solution of (4.2) is $\frac{T_s(\mathbf{z})}{\|T_s(\mathbf{z})\|_2}$, and thus the optimal value is $g^*(\mathbf{z}) = \frac{T_s(\mathbf{z})^T\mathbf{z}}{\|T_s(\mathbf{z})\|_2} = \|T_s(\mathbf{z})\|_2$. Obviously, in the case where $\mathbf{z} = \mathbf{0}$, it holds that $g^*(\mathbf{0}) = 0 = \|T_s(\mathbf{0})\|_2$, so in any case:

$$g^*(\mathbf{z}) = \|T_s(\mathbf{z})\|_2 \text{ for any } \mathbf{z} \in \mathbb{R}^d.$$

Consequently, the Toland-dual problem is given by

$$\text{(D-SPCA)} \quad \max_{\mathbf{y}\in\mathbb{R}^n} \left\{ \|T_s(\mathbf{A}^T\mathbf{y})\|_2 - \frac{1}{2}\|\mathbf{y}\|_2^2 \right\}.$$

4.3 A Model for Robust PCA: The Square-Root PCA

We first consider the alternative interpretation [10] of the PCA problem as the one that seeks a normalized direction $\mathbf{x}$ for which the sum of squares of distances of the points $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ to the line spanned by $\mathbf{x}$ is minimal. The resulting optimization problem is

$$\text{(PCA')} \quad \min_{\|\mathbf{x}\|_2=1} \sum_{i=1}^n \|\mathbf{a}_i - (\mathbf{a}_i^T\mathbf{x})\mathbf{x}\|_2^2.$$

Although problem (PCA') seems very much different than (PCA), it is easy to show that they are equivalent and the optimal sets of both consist of all the normalized leading eigenvectors of the matrix $\mathbf{A}^T\mathbf{A}$.

In the presence of outliers, it is well known that minimizing the sum of distances rather than the sum of squared distances leads to a more robust solution. Thus, we suggest the following formulation of the PCA problem that is robust to outliers:

$$\min_{\|\mathbf{x}\|_2=1} \sum_{i=1}^n \|\mathbf{a}_i - (\mathbf{a}_i^T\mathbf{x})\mathbf{x}\|_2,$$

which is the same as

$$\min_{\|\mathbf{x}\|_2=1} \sum_{i=1}^n \sqrt{\|\mathbf{a}_i\|_2^2 - \langle \mathbf{a}_i, \mathbf{x}\rangle^2}. \tag{4.3}$$

The above problem will be referred to as the *square-root PCA problem.* The objective function in (4.3) is concave, and thus we can replace the equality constraint by an inequality:

$$\min_{\|\mathbf{x}\|_2\leq 1} \sum_{i=1}^n \sqrt{\|\mathbf{a}_i\|_2^2 - \langle \mathbf{a}_i, \mathbf{x}\rangle^2}. \tag{4.4}$$

Finally, to make the objective function smooth, we will consider the following slightly modified formulation:

$$\text{(SRPCA)} \quad \max_{\|\mathbf{x}\|_2 \le 1} -\sum_{i=1}^{n} \sqrt{\|\mathbf{a}_i\|_2^2 - \langle \mathbf{a}_i, \mathbf{x} \rangle^2 + \varepsilon^2}.$$

Problem (SRPCA) fits model (P) with

$$f(\mathbf{z}) = \begin{cases} -\sum_{i=1}^n \sqrt{\|\mathbf{a}_i\|_2^2 + \varepsilon^2 - z_i^2} & |z_i| \le \sqrt{\|\mathbf{a}_i\|_2^2 + \varepsilon^2}, \\ \infty & \text{else}, \end{cases} \qquad g = \delta_{B_2[\mathbf{0},1]}.$$

Note that the underlying assumption $\text{dom}(g) \subseteq \text{dom}(f \circ \mathbf{A})$ holds true here. Denote $f(\mathbf{z}) = \sum_{i=1}^n f_i(z_i)$, where

$$f_i(z_i) \equiv \begin{cases} -\sqrt{\|\mathbf{a}_i\|_2^2 + \varepsilon^2 - z_i^2} & |z_i| \le \sqrt{\|\mathbf{a}_i\|_2^2 + \varepsilon^2}, \\ \infty & \text{else.} \end{cases}$$

By [2, Section 4.4.13], it follows that $f_i^*(y_i) = \sqrt{\|\mathbf{a}_i\|_2^2 + \varepsilon^2}\sqrt{y_i^2 + 1}$, and therefore, by the separability of f,

$$f^*(\mathbf{y}) = \sum_{i=1}^{n} \sqrt{\|\mathbf{a}_i\|_2^2 + \varepsilon^2}\sqrt{y_i^2 + 1}.$$

The Toland-dual problem is therefore,

$$\text{(D-SRPCA)} \quad \max_{\mathbf{y} \in \mathbb{R}^n} \left\{ \|\mathbf{A}^T \mathbf{y}\|_2 - \sum_{i=1}^{n} \sqrt{\|\mathbf{a}_i\|_2^2 + \varepsilon^2}\sqrt{y_i^2 + 1} \right\}.$$

5 The Dual Randomized Coordinate Descent Method

5.1 Definition and Convergence

The algorithm that we will be interested in is a randomized coordinate-descent method applied on the dual problem (D). We begin by describing the randomized coordinate descent method for solving minimization problems.

The RCD Method

Input. $(F, \mathbf{t}^0, r, \mathbf{p})$ where $F : \mathbb{R}^n \to \mathbb{R}, \mathbf{t}^0 \in \mathbb{R}^m, r > 0$ and a positive probability vector $\mathbf{p} \in \mathbb{R}^n_{++}, \mathbf{e}^T \mathbf{p} = 1$.
General Step. For any $k = 0, 1, ...$

(a) pick $i_k \in [n]$ at random according to the probability vector $\mathbf{p}$.

(b) compute $\alpha \in \operatorname*{argmin}_{t \in [-r,r]} F(t_1^k, t_2^k, \ldots, t_{i_k-1}^k, t, t_{i_k+1}^k, \ldots, t_n^k)$;

(c) set $t_j^{k+1} = \begin{cases} \alpha, & j = i_k, \\ t_j^k, & j \in [n] \setminus \{i_k\}. \end{cases}$

Remark 5.1. r can also be chosen as $r = \infty$ and in that case, the minimization in step (c) is over the entire real line.

We will exploit the following convergence result from [3]

Theorem 5.1 ([3, Theorem 4.2])**.** *Suppose that $F = f_1 - f_2$ with $f_1 : \mathbb{R}^n \to \mathbb{R}$ differentiable and convex and $f_2 : \mathbb{R}^n \to \mathbb{R}$ convex. Let $\{\mathbf{y}^k\}_{k\geq 0}$ be the sequence generated by the RCD method with input $(F, \mathbf{y}^0, r, \mathbf{p})$ where $\mathbf{y}^0 \in \mathbb{R}^n, r \in (0, \infty]$ and $\mathbf{p} \in \mathbb{R}^n_{++} \cap \Delta_n$. Then almost surely, all accumulation points of the sequence $\{\mathbf{y}^k\}_{k\geq 0}$ are stationary points of the problem $\min_{\mathbf{y}} F(\mathbf{y})$.*

The next theorem shows that if $\{\mathbf{y}^k\}_{k\geq 0}$ is a sequence generated by the RCD method employed on the problem of minimizing $-q$, and we define a **primal sequence** as any sequence satisfying $\mathbf{x}^k \in \partial g^*(\mathbf{A}^T\mathbf{y}^k)$, then the primal sequence satisfies the property that almost surely all its accumulation points are dual-stationary points of (P).

Theorem 5.2. *Suppose that Assumption 1 holds. Let $\{\mathbf{y}^k\}_{k\geq 0}$ be the sequence generated by the RCD method with input $(-q, \mathbf{y}^0, r, \mathbf{p})$ ($\mathbf{y}^0 \in \mathbb{R}^n, r > 0, \mathbf{p} \in \mathbb{R}^n_{++} \cap \Delta_n$), where $q(\mathbf{y}) = g^*(\mathbf{A}^T\mathbf{y}) - f^*(\mathbf{y})$. Assume that the level sets of $-q$ are bounded. Let $\mathbf{x}^k \in \partial g^*(\mathbf{A}^T\mathbf{y}^k)$. Then $\{\mathbf{x}^k\}_{k\geq 0}$ is bounded and almost surely all its accumulation points are dual-stationary points of problem (P).*

Proof. Since $\mathbf{x}^k \in \partial g^*(\mathbf{A}^T\mathbf{y}^k)$, it follows that $\mathbf{A}^T\mathbf{y}^k \in \partial g(\mathbf{x}^k)$, and thus in particular, $\mathbf{x}^k \in \mathrm{dom}(g)$, which is assumed to be compact. Thus, $\{\mathbf{x}^k\}_{k\geq 0}$ is bounded. We note that by Theorem 5.1 almost surely all accumulation points of the sequence $\{\mathbf{y}^k\}_{k\geq 0}$ are stationary points of (D). Let $\bar{\mathbf{x}}$ be an accumulation point of $\{\mathbf{x}^k\}_{k\geq 0}$. Then there exists a subsequence $\{\mathbf{x}^k\}_{k\in K}$ ($K \subseteq \mathbb{N}$) that converges to $\bar{\mathbf{x}}$, meaning that $\mathbf{x}^k \to \bar{\mathbf{x}}$ as $k \xrightarrow{K} \infty$. Since the sequence of function values $\{-q(\mathbf{y}^k)\}_{k\geq 0}$ is nonincreasing and $-q$ has bounded level sets, it follows that $\{\mathbf{y}^k\}_{k\geq 0}$ is bounded, and in particular, there exists a subsequence of $\{\mathbf{y}^k\}_{k\in K}$, which we denote by $\{\mathbf{y}^k\}_{k\in T}$ with $T \subseteq K$ that converges to some point $\bar{\mathbf{y}}$ which is almost surely a stationary point of (D). Taking $k \xrightarrow{T} \infty$ in the relation $\mathbf{x}^k \in \partial g^*(\mathbf{A}^T\mathbf{y}^k)$, we obtain that

$$\bar{\mathbf{x}} \in \partial g^*(\mathbf{A}^T\bar{\mathbf{y}}).$$

Thus, by Theorem 3.1(b), and the fact that $\bar{\mathbf{y}}$ is almost surely a stationary point of (D), it follows that $\bar{\mathbf{x}}$ is almost surely a dual-stationary point. □

We end this section by presenting the dual RCD method for solving problem (P) explicitly. To avoid expansive matrix/vector multiplications, we keep the vector $\mathbf{z}^k = \mathbf{A}^T\mathbf{y}^k$ whose update is made in linear time in d.

Dual RCD method for solving (P)

Input. $(f, g, \mathbf{A})$ satisfying Assumption 1, $r > 0$ and a positive probability vector $\mathbf{p} \in \mathbb{R}^n_{++} \cap \Delta_n, \mathbf{e}^T\mathbf{p} = 1$.
Initialization. $\mathbf{y}^0 = \mathbf{0} \in \mathbb{R}^n, \mathbf{z}^0 = \mathbf{0} \in \mathbb{R}^d$.
General Step. For any $k = 0, 1, .., K$,

(a) pick $i_k \in [n]$ at random according to the probability vector $\mathbf{p}$;

(b) compute $t_k \in \underset{t\in[-r,r]}{\operatorname{argmin}} \left\{ f^*(\mathbf{y}^k + (t - y^k_{i_k})\mathbf{e}_{i_k}) - g^*(\mathbf{z}^k + (t - y_{i_k})\mathbf{a}_{i_k}) \right\}$;

(c) update $\mathbf{y}^{k+1} = \mathbf{y}^k + (t_k - y^k_{i_k})\mathbf{e}_{i_k}$ and $\mathbf{z}^{k+1} = \mathbf{z}^k + (t_k - y_{i_k})\mathbf{a}_{i_k}$.

Output: $\mathbf{x}_{\text{out}} \in \partial g^*(\mathbf{z}^{K+1})$.

As can be seen from the above description, the general update formula of the method is $\mathbf{z}^{k+1} = \mathbf{z}^k + s_k\mathbf{a}_{i_k}$, where s_k is chosen via the vector $\mathbf{y}^k$ and the solution of the one-dimensional problem in step (b).

5.2 Replacing the Bounded Level Sets Assumption

Theorem 5.2 on the convergence of the dual CD method requires that the level sets of $\mathbf{y} \mapsto f^*(\mathbf{y}) - g^*(\mathbf{A}^T\mathbf{y})$ will be bounded. We will find a simpler condition that implies the bounded level sets assumption, and that is satisfied for all the three prototype problems. For that, we recall some basic notions and properties related to asymptotic functions. For the sake of simplicity, we will only consider real-valued functions (which is the case in the dual setting). For a given real-valued function h, the *asymptotic function* denoted by h_∞ is given by [1, Theorem 2.5.1]

$$h_\infty(\mathbf{d}) \equiv \liminf_{\mathbf{d}'\to\mathbf{d}, t\to\infty} \frac{h(t\mathbf{d}')}{t}.$$

If h is real-valued *convex*, then the following simpler formula for the asymptotic function holds [1, Corollary 2.5.3]:

$$h_\infty(\mathbf{d}) = \lim_{t\to\infty} \frac{h(t\mathbf{d})}{t}. \tag{5.1}$$

Lemma 5.2 below presents a calculus rule for the asymptotic function of the difference of two real-valued functions. The result uses the following simple lemma.

Lemma 5.1. *Let $s, t : \mathbb{R}^p \to \mathbb{R}$ be two real-valued functions and let $\bar{\mathbf{x}} \in \mathbb{R}^p$. Assume that $\lim_{\mathbf{x}\to\bar{\mathbf{x}}} t(\mathbf{x}) = \bar{t}$, where $\bar{t} \in \mathbb{R}$. Then*

$$\liminf_{\mathbf{x}\to\bar{\mathbf{x}}}(s(\mathbf{x}) - t(\mathbf{x})) = \liminf_{\mathbf{x}\to\bar{\mathbf{x}}} s(\mathbf{x}) - \lim_{\mathbf{x}\to\bar{\mathbf{x}}} t(\mathbf{x}).$$

Proof. First,

$$\liminf_{\mathbf{x}\to\bar{\mathbf{x}}}(s(\mathbf{x}) - t(\mathbf{x})) \geq \liminf_{\mathbf{x}\to\bar{\mathbf{x}}} s(\mathbf{x}) + \liminf_{\mathbf{x}\to\bar{\mathbf{x}}}(-t(\mathbf{x})) = \liminf_{\mathbf{x}\to\bar{\mathbf{x}}} s(\mathbf{x}) - \lim_{\mathbf{x}\to\bar{\mathbf{x}}} t(\mathbf{x}). \tag{5.2}$$

On the other hand,

$$\begin{aligned}\liminf_{\mathbf{x}\to\bar{\mathbf{x}}} s(\mathbf{x}) &= \liminf_{\mathbf{x}\to\bar{\mathbf{x}}}(s(\mathbf{x})-t(\mathbf{x})+t(\mathbf{x}))\\ &\geq \liminf_{\mathbf{x}\to\bar{\mathbf{x}}}(s(\mathbf{x})-t(\mathbf{x}))+\liminf_{\mathbf{x}\to\bar{\mathbf{x}}} t(\mathbf{x})\\ &= \liminf_{\mathbf{x}\to\bar{\mathbf{x}}}(s(\mathbf{x})-t(\mathbf{x}))+\lim_{\mathbf{x}\to\bar{\mathbf{x}}} t(\mathbf{x}). \end{aligned}\tag{5.3}$$

Combining (5.2) and (5.3), the desired result follows.

□

Lemma 5.2. *Suppose that u, v are real-valued functions such that v is Lipschitz continuous and convex. Then*

$$(u-v)_\infty = u_\infty - v_\infty.$$

Proof. First note that for any $\mathbf{d}\in\mathbb{R}^n$,

$$(u-v)_\infty(\mathbf{d}) = \liminf_{\mathbf{d}'\to\mathbf{d},t\to\infty}\frac{u(t\mathbf{d}')-v(t\mathbf{d}')}{t} = \liminf_{\mathbf{d}'\to\mathbf{d},t\to\infty}\left(\frac{u(t\mathbf{d}')}{t}-\frac{v(t\mathbf{d}')}{t}\right). \tag{5.4}$$

Since v is convex, it follows (c.f. (5.1)) that

$$v_\infty(\mathbf{d}) = \lim_{t\to\infty}\frac{v(t\mathbf{d})}{t}. \tag{5.5}$$

In addition, since v is Lipschitz, and denoting the Lipschitz constant of v by ℓ_v, we obtain that

$$\left|\frac{v(t\mathbf{d})}{t}\right| = \left|\frac{v(t\mathbf{d})-v(\mathbf{0})+v(\mathbf{0})}{t}\right| \leq \ell_v\|\mathbf{d}\|_2+\frac{v(\mathbf{0})}{t}.$$

Consequently, $v_\infty(\mathbf{d})\in\mathbb{R}$ and satisfies $|v_\infty(\mathbf{d})|\leq \ell_v\|\mathbf{d}\|_2$.

Now, for any $\mathbf{d},\mathbf{d}'\in\mathbb{R}^n$ and $t>0$,

$$\left|\frac{v(t\mathbf{d}')}{t}-v_\infty(\mathbf{d})\right| \leq \left|\frac{v(t\mathbf{d})}{t}-v_\infty(\mathbf{d})\right| + \left|\frac{v(t\mathbf{d}')}{t}-\frac{v(t\mathbf{d})}{t}\right| \leq \left|\frac{v(t\mathbf{d})}{t}-v_\infty(\mathbf{d})\right| + \ell_g\|\mathbf{d}-\mathbf{d}'\|_2, \tag{5.6}$$

where the inequality follows by the fact that v is ℓ_v-Lipschitz continuous. By (5.5), it holds that $\lim_{t\to\infty,\mathbf{d}'\to\mathbf{d}}\left|\frac{v(t\mathbf{d})}{t}-v_\infty(\mathbf{d})\right|=0$; in addition, $\lim_{t\to\infty,\mathbf{d}'\to\mathbf{d}}\|\mathbf{d}-\mathbf{d}'\|_2=0$, and therefore, by (5.6),

$$\lim_{t\to\infty,\mathbf{d}'\to\mathbf{d}}\frac{v(t\mathbf{d}')}{t} = v_\infty(\mathbf{d}),$$

which combined with (5.4) and Lemma 5.1, finally implies

$$\begin{aligned}(u-v)_\infty(\mathbf{d}) &= \liminf_{\mathbf{d}'\to\mathbf{d},t\to\infty}\left(\frac{u(t\mathbf{d}')}{t}-\frac{v(t\mathbf{d}')}{t}\right)\\ &= \liminf_{\mathbf{d}'\to\mathbf{d},t\to\infty}\frac{u(t\mathbf{d}')}{t}-\lim_{\mathbf{d}'\to\mathbf{d},t\to\infty}\frac{v(t\mathbf{d}')}{t} = u_\infty(\mathbf{d})-v_\infty(\mathbf{d}).\end{aligned}$$

□

Based on the above result, we will now show that the bounded level sets assumption in Theorem 5.2 can be replaced by a condition written in terms of $(f^*)_\infty$ and $(g^*)_\infty$.

Corollary 5.1. *Theorem 5.2 still holds if the bounded level sets assumption is replaced by the assumption*

$$(f^*)_\infty(\mathbf{d}) - (g^*)_\infty(\mathbf{A}^T\mathbf{d}) > 0 \textit{ for all } \mathbf{d} \in \mathbb{R}^n \setminus \{\mathbf{0}\}. \tag{5.7}$$

Proof. Set $u = f^*, v(\cdot) = g^*(\mathbf{A}^T \cdot)$. Since g has a compact domain (Assumption 1), it follows that g^* is Lipschitz continuous, and thus the assumptions of Lemma 5.2 hold and we can thus conclude that

$$(-q)_\infty(\mathbf{d}) = (f^*)_\infty(\mathbf{d}) - (g^*)_\infty(\mathbf{A}^T\mathbf{d}),$$

where we used in the above the calculus rule [1, Proposition 2.6.3]. Therefore, condition (5.7) is the same as the condition $(-q)_\infty(\mathbf{d}) > 0$ for any $\mathbf{d} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$. Thus, by [1, Proposition 3.1.3], the level sets of $-q$ are bounded. □

Fortunately, the prototype problems all satisfy condition (5.7). In the first two problems we have $f^*(\mathbf{y}) = \frac{1}{2}\|\mathbf{y}\|_2^2$, and therefore, $(f^*)_\infty(\mathbf{d}) = \infty$, which trivially implies that condition (5.7) holds. The third problem is slightly more complicated. Since $f^*(\mathbf{y}) = \sum_{i=1}^n \sqrt{\|\mathbf{a}_i\|_2^2 + \varepsilon^2}\sqrt{y_i^2+1}$, it follows that $(f^*)_\infty(\mathbf{d}) = \sum_{i=1}^n \sqrt{\|\mathbf{a}_i\|_2^2 + \varepsilon^2}|d_i|$. Also, $g^*(\mathbf{y}) = \|\mathbf{y}\|_2$, and hence, $(g^*)_\infty(\mathbf{d}) = \|\mathbf{d}\|_2$. Thus, the condition that needs to be proven is

$$\sum_{i=1}^n \sqrt{\|\mathbf{a}_i\|_2^2 + \varepsilon^2}|d_i| - \|\mathbf{A}^T\mathbf{d}\|_2 > 0 \text{ for all } \mathbf{d} \neq \mathbf{0}.$$

To prove the above, we first note that by the triangle inequality,

$$\|\mathbf{A}^T\mathbf{d}\|_2 = \left\|\sum_{i=1}^n d_i\mathbf{a}_i\right\|_2 \le \sum_{i=1}^n \|\mathbf{a}_i\|_2|d_i|.$$

Thus, for any $\mathbf{d} \neq \mathbf{0}$,

$$\sum_{i=1}^n \sqrt{\|\mathbf{a}_i\|_2^2 + \varepsilon^2}|d_i| - \|\mathbf{A}^T\mathbf{d}\|_2 \ge \sum_{i=1}^n (\sqrt{\|\mathbf{a}_i\|_2^2 + \varepsilon^2} - \|\mathbf{a}_i\|_2)|d_i| > 0,$$

and we thus established that condition (5.7) holds for all three prototype problems.

5.3 Convergence in the Case $g = \delta_{B_{\|\cdot\|_2}[\mathbf{0},1]}$

In this section we will concentrate on the case where f and g satisfy – in addition to the properties of Assumption 1 – the following:

Assumption 2. $g(\mathbf{x}) = \delta_{B_{\|\cdot\|_2}[\mathbf{0},1]}(\mathbf{x})$ *and* $\operatorname{argmin}_{\mathbf{x}} f(\mathbf{x}) = \{\mathbf{0}\}$.

Thus, under Assumption 2, the primal problem takes the form

$$\max_{\mathbf{x}}\{f(\mathbf{A}\mathbf{x}) : \|\mathbf{x}\|_2 \le 1\}$$

and the Toland-dual problem is thus

$$\max_{\mathbf{y}} \left\{\|\mathbf{A}^T\mathbf{y}\|_2 - f^*(\mathbf{y})\right\}. \tag{5.8}$$

We will often consider the minimization form of the above:

$$\min_{\mathbf{y}}\{f^*(\mathbf{y}) - \|\mathbf{A}^T\mathbf{y}\|_2\}.$$

Note that Assumption 2 is satisfied by the first and third prototype problems. Since the function $\mathbf{y} \mapsto \|\mathbf{A}^T\mathbf{y}\|_2$ is differentiable at $\bar{\mathbf{y}} \in \mathbb{R}^n$ if and only if $\mathbf{A}^T\bar{\mathbf{y}} \neq \mathbf{0}$, it follows by Remark 2.1 that if $\bar{\mathbf{y}}$ is a stationary point, then $\mathbf{A}^T\bar{\mathbf{y}} \neq \mathbf{0}$. The next lemma will be key to showing that non-differentiability can be avoided altogether.

Lemma 5.3. *Suppose that Assumptions 1 and 2 hold. Let $\bar{\mathbf{y}} \in \mathbb{R}^n$ satisfy $q(\bar{\mathbf{y}}) > q(\mathbf{0})$. Then $\mathbf{A}^T\bar{\mathbf{y}} \neq \mathbf{0}$.*

Proof. Since $\operatorname{argmin}_{\mathbf{x}} f(\mathbf{x}) = \{\mathbf{0}\}$, it follows that $\mathbf{0} \in \partial f(\mathbf{0})$, and thus, $\nabla f^*(\mathbf{0}) = \mathbf{0}$, which implies that

$$\mathbf{0} \in \operatorname*{argmin}_{\mathbf{y}} f^*(\mathbf{y}). \tag{5.9}$$

Let $\bar{\mathbf{y}} \in \mathbb{R}^m$ satisfy $q(\bar{\mathbf{y}}) > q(\mathbf{0})$, and assume by contradiction that $\mathbf{A}^T\bar{\mathbf{y}} = \mathbf{0}$. Then

$$q(\bar{\mathbf{y}}) = \|\mathbf{A}^T\bar{\mathbf{y}}\|_2 - f^*(\bar{\mathbf{y}}) \overset{\mathbf{A}^T\bar{\mathbf{y}}=\mathbf{0}}{=} -f^*(\bar{\mathbf{y}}) \overset{(5.9)}{\leq} -f^*(\mathbf{0}) = q(\mathbf{0}),$$

and we arrive at a contradiction to the assumption that $q(\bar{\mathbf{y}}) > q(\mathbf{0})$, proving the desired result that $\mathbf{A}^T\bar{\mathbf{y}} \neq \mathbf{0}$. □

Note that for any $\mathbf{d} \neq \mathbf{0}$,

$$q'(\mathbf{0};\mathbf{d}) = \|\mathbf{A}^T\mathbf{d}\|_2 - \nabla f^*(\mathbf{0})^T\mathbf{d} = \|\mathbf{A}^T\mathbf{d}\|_2 > 0.$$

This means that *any* direction $\mathbf{d}$ such that $\mathbf{A}^T\mathbf{d} \neq \mathbf{0}$ is a descent direction of $-q$. In particular, directions of the form $\pm\mathbf{e}_i$ are descent directions of $-q$ at $\mathbf{0}$ since $\|\mathbf{A}^T(\pm\mathbf{e}_i)\|_2 = \|\mathbf{a}_i\|_2 \neq 0$. Therefore, if we begin the RCD method applied on $-q$ with $\mathbf{y}^0 = \mathbf{0}$, the first iteration will always result with a better objective function, meaning $q(\mathbf{y}^1) > q(\mathbf{0})$, and since the RCD method is a descent method (applied on $-q$), we obtain that for any $k \geq 1$, $q(\mathbf{y}^k) \geq q(\mathbf{y}^1) > 0$. By Lemma 5.3, it follows that $\mathbf{A}^T\mathbf{y}^k \neq 0$ for all $k \geq 1$, meaning in particular that q is differentiable at $\mathbf{y}^k \neq \mathbf{0}$ for any $k \geq 1$, and thus $\|\nabla q(\mathbf{y}^k)\|_2$ is defined for any $k \geq 1$, and can serve as an optimality measure for the k-th iterate vector $\mathbf{y}^k$. We summarize these observations in the following lemma.

Lemma 5.4. *Suppose that Assumptions 1 and 2 hold. Let $\{\mathbf{y}^k\}_{k\geq 0}$ be the sequence generated by the RCD method with input $(-q, \mathbf{0}, r, \mathbf{p})$ ($r > 0, \mathbf{p} \in \mathbb{R}^n_{++} \cap \Delta_n$), where $q(\mathbf{y}) = g^*(\mathbf{A}^T\mathbf{y}) - f^*(\mathbf{y})$. Then*

$$q(\mathbf{y}^k) \geq q(\mathbf{y}^1) > q(\mathbf{0}) \text{ for all } k \geq 1$$

and $\nabla q(\mathbf{y}^k)$ exists for any $k \geq 1$.

Our next task is to show that under the additional assumption that the gradient of f^* is block Lipschitz continuous, it is possible to show a rate of convergence of the expected values of the optimality measure $\|\nabla q(\mathbf{y}^k)\|_2$. We first explicitly write the required property.

Assumption 3 (block Lipschitz-continuity of f^*)**.** *For any* $i \in [n]$*,*

$$\|\nabla_i f^*(\mathbf{y} + h\mathbf{e}_i) - \nabla_i f^*(\mathbf{y})\|_2 \leq L_i|h| \text{ for all } \mathbf{y} \in \mathbb{R}^n, h \in \mathbb{R}$$

for some positive numbers $L_1, L_2, \ldots, L_n$*.*

The rate of convergence of the optimality measure can now be established.

Theorem 5.3. *Suppose that Assumptions 1,2 and 3 hold. Let* $\{\mathbf{y}^k\}_{k\geq 0}$ *be the sequence generated by the RCD method with input* $(-q, \mathbf{0}, \infty, \mathbf{p})$*, where* $q(\mathbf{y}) = g^*(\mathbf{A}^T\mathbf{y}) - f^*(\mathbf{y})$ *and* $\mathbf{p} \in \mathbb{R}^n_{++} \cap \Delta_n$*. Then for any* $k \geq 1$*, it holds that*[2]

$$\min_{k=1,\ldots,N} \mathbb{E}(\|\nabla q(\mathbf{y}^k)\|_2^2) \leq \frac{C}{N}(q_{\text{opt}} - q(\mathbf{0})), \tag{5.10}$$

where $q_{\text{opt}} = \max_{\mathbf{y}\in\mathbb{R}^n} q(\mathbf{y})$ *and* $C = 2\max_{i=1,\ldots,n} \frac{L_i}{p_i}$*.*

Proof. Denote the convex set $S \equiv \text{Lev}(-q, -q(\mathbf{y}^1))$. Then by Lemma 5.4, $\mathbf{y}^k \in S$ for any $k \geq 1$ and q is differentiable over S. By the block descent lemma [2, Lemma 11.8], it follows that for any $i \in [n]$ and $\mathbf{y} \in S$,

$$f^*(\mathbf{y} + h\mathbf{e}_i) \leq f^*(\mathbf{y}) + \nabla_i f^*(\mathbf{y})h + \frac{L_i}{2}h_i^2 \text{ for any } h \in \mathbb{R}. \tag{5.11}$$

Also, by the convexity of $\mathbf{y} \mapsto t(\mathbf{y}) \equiv g^*(\mathbf{A}^T\mathbf{y})$, we conclude that for any $i \in [n]$,

$$-t(\mathbf{y} + h\mathbf{e}_i) \leq -t(\mathbf{y}) - \nabla_i t(\mathbf{y})h \text{ for all } h \in \mathbb{R}, \tag{5.12}$$

where we also used in the above the fact that $t = f^* + q$ is differentiable over S. Adding (5.11) and (5.12), and using the identity $-q = f^* - t$, it follows that

$$-q(\mathbf{y}) + q(\mathbf{y} + h\mathbf{e}_i) \geq \nabla_i q(\mathbf{y})h - \frac{L_i}{2}h^2 \text{ for any } h \in \mathbb{R}. \tag{5.13}$$

Define $\bar{h}_{\mathbf{y}} \equiv \frac{\nabla_i q(\mathbf{y})}{L_i}$. Plugging $h = \bar{h}_{\mathbf{y}}$ in (5.13) yields

$$-q(\mathbf{y}) + q(\mathbf{y} + \bar{h}_{\mathbf{y}}\mathbf{e}_i) \geq \frac{(\nabla_i q(\mathbf{y}))^2}{2L_i} \text{ for any } \mathbf{y} \in S. \tag{5.14}$$

For any $k \geq 1$, it holds that $\mathbf{y}^k \in S$, and thus we can plug $\mathbf{y} = \mathbf{y}^k$ and $i = i_k$ in (5.14), and obtain that for any $k \geq 1$,

$$-q(\mathbf{y}^k) + q(\mathbf{y}^k + \bar{h}_{\mathbf{y}^k}\mathbf{e}_{i_k}) \geq \frac{(\nabla_{i_k} q(\mathbf{y}^k))^2}{2L_{i_k}}. \tag{5.15}$$

By the definition of the process, $-q(\mathbf{y}^{k+1}) \leq -q\left(\mathbf{y}^k + \bar{h}_{\mathbf{y}^k}\mathbf{e}_{i_k}\right)$, which combined with (5.15) implies the inequality

$$-q(\mathbf{y}^k) + q(\mathbf{y}^{k+1}) \geq \frac{1}{2L_{i_k}}(\nabla_{i_k} q(\mathbf{y}^k))^2.$$

[2]The expectation of $\|\nabla q(\mathbf{y}^k)\|_2^2$ is over the random variables $i_0, i_1, \ldots, i_{k-1}$

Taking the expectation over the random variable i_k, we obtain that

$$-q(\mathbf{y}^k) + \mathbb{E}_{i_k}(q(\mathbf{y}^{k+1})) \geq \sum_{i=1}^{n} \frac{p_i}{2L_i}(\nabla_i q(\mathbf{y}^k))^2 \geq \min_{i=1,\ldots,n} \left\{ \frac{p_i}{2L_i} \right\} \|\nabla q(\mathbf{y}^k)\|_2^2. \tag{5.16}$$

For any $k \geq 1$ we denote the multivariate random variable $\xi_k \equiv (i_0, i_1, \ldots, i_{k-1})$. Taking expectation with respect to ξ_k of both sides of (5.16) leads to the following inequality:

$$-\mathbb{E}_{\xi_k}(q(\mathbf{y}^k)) + \mathbb{E}_{\xi_{k+1}}(q(\mathbf{y}^{k+1})) \geq \underbrace{\min_{i=1,\ldots,n} \left\{ \frac{p_i}{2L_i} \right\}}_{\frac{1}{C}} \mathbb{E}_{\xi_k}(\|\nabla q(\mathbf{y}^k)\|_2^2).$$

Summing the above inequality over $k = 1, ..., N$ yields:

$$-\mathbb{E}_{\xi_1}(q(\mathbf{y}^1)) + \mathbb{E}_{\xi_{N+1}}(q(\mathbf{y}^{N+1})) \geq \frac{1}{C} \sum_{k=1}^{N} \mathbb{E}_{\xi_k}(\|\nabla q(\mathbf{y}^k)\|_2^2) \geq \frac{1}{C} N \min_{k=1,\ldots,N} \mathbb{E}_{\xi_k}(\|\nabla q(\mathbf{y}^k)\|_2^2),$$

meaning

$$\min_{k=1,\ldots,N} \mathbb{E}_{\xi_k}(\|\nabla q(\mathbf{y}^k)\|_2^2) \leq \frac{C}{N}(-\mathbb{E}_{\xi_1}(q(\mathbf{y}^1)) + \mathbb{E}_{\xi_{N+1}}(q(\mathbf{y}^{N+1})))$$

By Lemma 5.4, $q(\mathbf{y}^1) > q(\mathbf{0})$, and obviously $q(\mathbf{y}^{N+1}) < q_{\text{opt}}$, implying the desired result (5.10). □

6 Dual RCD for the Three PCA Prototype Problems

This final section shows how the dual RCD method can be applied in each of the three prototype models described in Section 4 to produce extremely simple schemes.

6.1 The PCA Problem

The Toland-dual problem of the PCA problem if given by (see Section 4.1)

$$\max_{\mathbf{y} \in \mathbb{R}^n} \|\mathbf{A}^T \mathbf{y}\|_2 - \frac{1}{2}\|\mathbf{y}\|_2^2.$$

Suppose that we are at iteration k, meaning that we know $\mathbf{y}^k$ and we always keep $\mathbf{z}^k = \mathbf{A}^T \mathbf{y}^k$. We pick at random an index $i_k \in [n]$, and the one-dimensional problem that needs to be solved is

$$\min_t \frac{1}{2}\|\mathbf{y}^k + (t - y_{i_k}^k)\mathbf{e}_{i_k}\|_2^2 - \|\mathbf{z}^k + (t - y_{i_k}^k)\mathbf{a}_{i_k}\|_2,$$

where we use the notation that the rows of $\mathbf{A}$ are $\mathbf{a}_1^T, \ldots, \mathbf{a}_n^T$. Denoting $\tilde{\mathbf{y}}^k = \mathbf{y}^k - \mathbf{y}_{i_k}\mathbf{e}_{i_k}$ and $\tilde{\mathbf{z}}^k = \mathbf{z}^k - y_{i_k}\mathbf{a}_{i_k}$, the problem thus becomes (using the relation $\mathbf{e}_{i_k}^T \tilde{\mathbf{y}}^k = 0$)

$$\min_t \left\{ h(t) \equiv \frac{1}{2}(\|\tilde{\mathbf{y}}^k\|_2^2 + t^2) - \sqrt{\|\tilde{\mathbf{z}}^k\|_2^2 + 2t\mathbf{a}_{i_k}^T \tilde{\mathbf{z}}^k + t^2\|\mathbf{a}_{i_k}\|_2^2} \right\}.$$

The optimality condition $h'(t) = 0$ is the same as

$$t = \frac{\mathbf{a}_{i_k}^T \tilde{\mathbf{z}}^k + t\|\mathbf{a}_{i_k}\|_2^2}{\sqrt{\|\tilde{\mathbf{z}}^k\|_2^2 + 2t\mathbf{a}_{i_k}^T \tilde{\mathbf{z}}^k + t^2\|\mathbf{a}_{i_k}\|_2^2}}$$

A simple and tedious algebraic argument shows that the solutions of the equation $h'(t) = 0$ are all roots of the quartic equation, which admits an analytic solution[3]

$$c_4 t^4 + c_3^3 + c_2 t^2 + c_1 t + c_0 = 0, \tag{6.1}$$

where

$$c_0 = -(\mathbf{a}_{i_k}^T \tilde{\mathbf{z}}^k)^2, c_1 = -2(\mathbf{a}_{i_k}^T \tilde{\mathbf{z}}^k)\|\mathbf{a}_{i_k}\|_2^2, c_2 = \|\tilde{\mathbf{z}}^k\|_2^2 - \|\mathbf{a}_{i_k}\|_2^4, c_3 = 2(\mathbf{a}_{i_k}^T \tilde{\mathbf{z}}^k), c_4 = \|\mathbf{a}_{i_k}\|_2^2. \tag{6.2}$$

The dual RCD method for solving the PCA problem is now explicitly stated.

Dual RCD Method for PCA

Input. $\mathbf{A} \in \mathbb{R}^{n \times d}$ and a positive probability vector $\mathbf{p} \in \mathbb{R}_{++}^n \cap \Delta_n, \mathbf{e}^T\mathbf{p} = 1$, K - number of iterations.
Initialization. $\mathbf{y}^0 = \mathbf{0}, \mathbf{z}^0 = \mathbf{0}$.
General Step. For any $k = 0, 1, ...K$,

(a) pick $i_k \in [n]$ at random according to the probability vector $\mathbf{p}$.

(b) compute $\tilde{\mathbf{z}}^k = \mathbf{z}^k - y_{i_k}^k \mathbf{a}_{i_k}$

(c) find all real roots $r_1, r_2, \ldots, r_c$ of the quartic equation (6.1) whose coefficients are given in (6.2); set $t_k \in \operatorname{argmin}\{h(z) : z \in \{r_1, r_2, \ldots, r_c\}\}$.

(d) update $\mathbf{y}^{k+1} = \mathbf{y}^k + (t_k - y_{i_k}^k)\mathbf{e}_{i_k}, \mathbf{z}^{k+1} = \tilde{\mathbf{z}}^k + t_k \mathbf{a}_{i_k}$.

Output: $\mathbf{x}_{\text{out}} = \frac{\mathbf{z}^{K+1}}{\|\mathbf{z}^{K+1}\|_2}$.

6.2 Square-Root PCA

The Toland-dual problem for the square-root PCA problem is given by (see Section 4.3)

$$\text{(D-SRPCA)} \quad \max_{\mathbf{y}} \left\{ \|\mathbf{A}^T\mathbf{y}\|_2 - \sum_{i=1}^n \sqrt{\|\mathbf{a}_i\|_2^2 + \varepsilon^2} \sqrt{y_i^2 + 1} \right\}.$$

As before, we denote $\mathbf{z}^k = \mathbf{A}^T\mathbf{y}^k$. At iteration k, the one-dimensional problem with respect to y_{i_k} (taking $y_{i_k} = t$) is

$$\min_t \left\{ h_3(t) \equiv \sqrt{\|\mathbf{a}_{i_k}\|_2^2 + \varepsilon^2}\sqrt{t^2 + 1} - \sqrt{\|\tilde{\mathbf{z}}^k\|_2^2 + 2t\mathbf{a}_{i_k}^T \tilde{\mathbf{z}}^k + t^2\|\mathbf{a}_{i_k}\|_2^2} \right\},$$

[3] The algebraic solution of quartic equations was discovered by Lodovico Ferrari and Gerolamo Cardano in its *Ars Magna* (1545); see `https://www.britannica.com/biography/Lodovico-Ferrari`

where $\tilde{\mathbf{z}}^k = \mathbf{z}^k - y_{i_k}\mathbf{a}_{i_k}$. The optimality condition $h_3'(t) = 0$ is the same as

$$\sqrt{\|\mathbf{a}_{i_k}\|_2^2 + \varepsilon^2}\frac{t}{\sqrt{t^2+1}} = \frac{\mathbf{a}_{i_k}^T\tilde{\mathbf{z}}^k + t\|\mathbf{a}_{i_k}\|_2^2}{\sqrt{\|\tilde{\mathbf{z}}^k\|_2^2 + 2t\mathbf{a}_{i_k}^T\tilde{\mathbf{z}}^k + t^2\|\mathbf{a}_{i_k}\|_2^2}},$$

and it can be shown after some algebraic rearrangements that the solutions to the above equations must be root of the quartic problem

$$c_4t^4 + c_3^3 + c_2t^2 + c_1t + c_0 = 0, \tag{6.3}$$

where

$$c_4 = \varepsilon^2\|\mathbf{a}_{i_k}\|_2^2, c_3 = 2\varepsilon^2(\mathbf{a}_{i_k}^T\tilde{\mathbf{z}}^k), c_2 = \|\tilde{\mathbf{z}}^k\|_2^2(\|\mathbf{a}_{i_k}\|_2^2 + \varepsilon^2) - (\|\mathbf{a}_{i_k}\|_2^4 + (\mathbf{a}_{i_k}^T\tilde{\mathbf{z}}^k)^2), \tag{6.4}$$

$$c_1 = -2(\mathbf{a}_{i_k}^T\tilde{\mathbf{z}}^k)\|\mathbf{a}_{i_k}\|_2^2, c_0 = -(\mathbf{a}_{i_k}^T\tilde{\mathbf{z}}^k)^2. \tag{6.5}$$

This leads us to the following algorithm.

Dual RCD Method for SRPCA

Input. $\mathbf{A} \in \mathbb{R}^{n\times d}$ and a positive probability vector $\mathbf{p} \in \mathbb{R}_{++}^n \cap \Delta_n, \mathbf{e}^T\mathbf{p} = 1$, K - number of iterations.
Initialization. $\mathbf{y}^0 = \mathbf{0}, \mathbf{z}^0 = \mathbf{0}$.
General Step. For any $k = 0, 1, ...K$,

(a) pick $i_k \in [m]$ at random according to the probability vector $\mathbf{p}$.

(b) compute $\tilde{\mathbf{z}}^k = \mathbf{z}^k - y_{i_k}^k\mathbf{a}_{i_k}$

(c) find all real roots $r_1, r_2, \ldots, r_c$ of the quartic equation (6.3) whose coefficients are given in (6.4) and (6.5); set $t_k \in \text{argmin}\{h(z) : z \in \{r_1, r_2, \ldots, r_c\}\}$.

(d) update $\mathbf{y}^{k+1} = \mathbf{y}^k + (t_k - y_{i_k}^k)\mathbf{e}_{i_k}, \mathbf{z}^{k+1} = \tilde{\mathbf{z}}^k + t_k\mathbf{a}_{i_k}$.

Output: $\mathbf{x}_{\text{out}} = \frac{\mathbf{z}^{K+1}}{\|\mathbf{z}^{K+1}\|_2}$.

6.3 Sparse PCA

The Toland-dual problem of the sparse PCA problem is (see Section 4.2)

$$\text{(D-PCA)} \quad \max_{\mathbf{y}}\left\{\|T_s(\mathbf{A}^T\mathbf{y})\|_2 - \frac{1}{2}\|\mathbf{y}\|_2^2\right\}.$$

Suppose that we are at iteration k, meaning that we know $\mathbf{y}^k$ and assume that we always keep $\mathbf{z}^k = \mathbf{A}^T\mathbf{y}^k$. We pick at random an index $i_k \in [n]$. The one-dimensional problem that needs to be solved is

$$\min_t\left\{\frac{1}{2}\|\mathbf{y}^k + (t - y_{i_k}^k)\mathbf{e}_{i_k}\|_2^2 - \|T_s(\mathbf{z}^k + (t - y_{i_k}^k)\mathbf{a}_{i_k})\|_2\right\}.$$

Thus, the problem that is solved at each iteration is of the form

$$\min_{t} \left\{ R_{\mathbf{v},\mathbf{w}}(t) \equiv \frac{1}{2}t^2 - \|T_s(\mathbf{v} + t\mathbf{w})\|_2 \right\} \tag{6.6}$$

for some $\mathbf{v}, \mathbf{w} \in \mathbb{R}^d$. The objective function $R_{\mathbf{v},\mathbf{w}}$ is neither convex nor smooth. An illustration can be found in Figure 2.

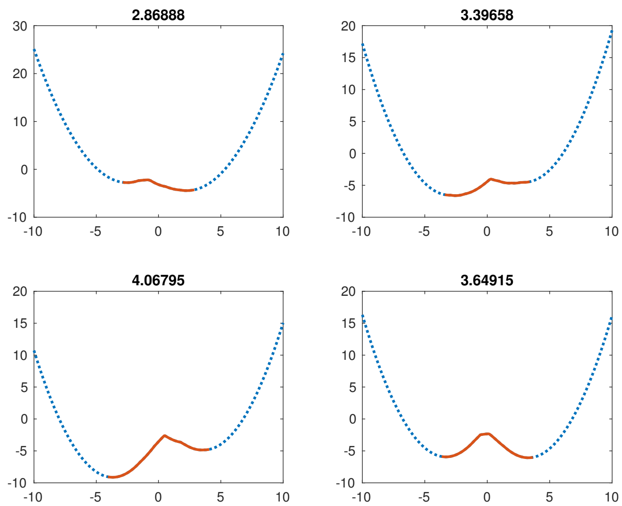

Figure 2: Plot of the function $R_{\mathbf{v},\mathbf{w}}$ for four normally and randomly generated $\mathbf{u}, \mathbf{v}$ with $n = 10, s = 2$. The filled line corresponds to the theoretical interval in which a minimizer of the function is guaranteed to reside.

The next result shows that we can find a compact interval in which an optimal solution is guaranteed to reside and an explicit expression for the Lipschitz constant of the function. This means that it is possible to solve the problem quite effectively using any of the existing solvers for Lipschitz-continuous functions over a compact interval, see for example [5]. Note that the function values on the theoretical compact interval are highlighted in Figure 2.

Theorem 6.1. *Consider problem (6.6) with* $\mathbf{u}, \mathbf{v} \in \mathbb{R}^d$ *and* $s \in [d]$. *Then*

(b) the function $R_{\mathbf{v},\mathbf{w}} : \mathbb{R} \to \mathbb{R}$ *is* $2\|\mathbf{w}\|_2$*-Lipschitz continuous over* $[-\|\mathbf{w}\|_2, \|\mathbf{w}\|_2]$*;*

(b) all minimizers of problem (6.6) reside in the interval $[-\|\mathbf{w}\|_2, \|\mathbf{w}\|_2]$.

Proof. Define $H(\cdot) = \|T_s(\cdot)\|_2$ and $h_{\mathbf{v},\mathbf{w}}(t) = H(\mathbf{v} + t\mathbf{w})$. We begin by noting that

$$H(\mathbf{z}) = \|T_s(\mathbf{z})\|_2 = \max\{\langle \mathbf{x}, \mathbf{z}\rangle : \|\mathbf{x}\|_2 \leq 1, \|\mathbf{x}\|_0 \leq s\} = \sigma_D(\mathbf{z}),$$

where $D = \operatorname{conv}(B_2[\mathbf{0}, 1] \cap C_s)$. Therefore, since D is closed and convex,

$$\partial H(\mathbf{x}) = \operatorname{argmax}_{\mathbf{y}} \{\langle \mathbf{x}, \mathbf{y}\rangle - \sigma_D^*(\mathbf{y})\} = \operatorname{argmax}_{\mathbf{y}} \{\langle \mathbf{x}, \mathbf{y}\rangle - \delta_D(\mathbf{y})\} \subseteq D,$$

and hence,

$$\partial H(\mathbf{x}) \subseteq B_2[\mathbf{0}, 1]. \tag{6.7}$$

Consequently, by (6.7), for any $t \in \mathbb{R}$,

$$|\partial h_{\mathbf{v},\mathbf{w}}(t)| = |\mathbf{w}^T \partial H(\mathbf{v} + t\mathbf{w})| \leq \|\mathbf{w}\|_2. \tag{6.8}$$

In particular, $h_{\mathbf{v},\mathbf{w}}$ is $\|\mathbf{w}\|_2$-Lipschitz continuous. Since $t \mapsto \frac{1}{2}t^2$ is also $\|\mathbf{w}\|_2$-Lipschitz continuous over $[-\|\mathbf{w}\|_2, \|\mathbf{w}\|_2]$, it follows that $R_{\mathbf{v},\mathbf{w}} = \frac{1}{2}(\cdot)^2 - h_{\mathbf{v},\mathbf{w}}$ is $(2\|\mathbf{w}\|_2)$-Lipschitz continuous over $[-\|\mathbf{w}\|_2, \|\mathbf{w}\|_2]$, establishing part (a).

To show part (b), note that by (6.8) it follows that for any $t \in [-\|\mathbf{w}\|_2, \|\mathbf{w}\|_2]$,

$$-\|\mathbf{w}\|_2 \leq (h_{\mathbf{v},\mathbf{w}})'_-(t),\ (h_{\mathbf{v},\mathbf{w}})'_+(t) \leq \|\mathbf{w}\|_2.$$

Therefore, for any $t < -\|\mathbf{w}\|_2$ it holds that

$$\max\{(R_{\mathbf{v},\mathbf{w}})'_-(t), (R_{\mathbf{v},\mathbf{w}})'_+(t)\} = t - \min\{(h_{\mathbf{v},\mathbf{w}})'_-(t), (h_{\mathbf{v},\mathbf{w}})'_+(t)\} < -\|\mathbf{w}\|_2 + \|\mathbf{w}\|_2 = 0,$$

implying that there are no minimizers in the interval $(-\infty, -\|\mathbf{w}\|_2)$. A very similar argument shows that there are no minimizers in the interval $(\|\mathbf{w}\|_2, \infty)$, and part (a) is thus established. □

References

[1] A. Auslender and M. Teboulle. *Asymptotic cones and functions in optimization and variational inequalities*. Springer Monographs in Mathematics. Springer-Verlag, New York, 2003.

[2] A. Beck. *First-order methods in optimization*, volume 25 of *MOS-SIAM Series on Optimization*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Optimization Society, Philadelphia, PA, 2017.

[3] A. Beck and N. Hallak. On the convergence to stationary points of deterministic and randomized feasible descent directions methods. *SIAM Journal on Optimization*, 2019. To appear.

[4] R. Horst and N. V. Thoai. DC Programming: Overview. *Journal of Optimization Theory and Applications*, 103(1):1–43, 1999.

[5] R. Horst and H. Tuy. *Global optimization*. Springer-Verlag, Berlin, second edition, 1993. Deterministic approaches.

[6] R. Luss and M. Teboulle. Conditional gradient algorithms for rank-one matrix approximations with a sparsity constraint. *SIAM Rev.*, 55(1):65–98, 2013.

[7] Baback M., Y. Weiss, and Avidan S. Spectral bounds for sparse pca: Exact and greedy algorithms. In Y. Weiss, B. Schölkopf, and J. C. Platt, editors, *Advances in Neural Information Processing Systems 18*, pages 915–922. MIT Press, 2006.

[8] E. Oja and J. Karhunen. On stochastic approximation of the eigenvectors and eigenvalues of the expectation of a random matrix. *J. Math. Anal. Appl.*, 106(1):69–84, 1985.

[9] J. S. Pang, M. Razaviyayn, and A. Alvarado. Computing B-Stationary Points of Nonsmooth DC Programs. *Mathematics of Operations Research*, 42(1):95–118, jan 2017.

[10] K. Pearson. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2(11):559–572, 1901.

[11] O. Shamir. Convergence of stochastic gradient descent for PCA. Proceedings of the 33rd International Conference on Machine Learning. vol. 48: 257–265 (2016).

[12] J. F. Toland. Duality in nonconvex optimization. *J. Math. Anal. Appl.*, 66(2):399–415, 1978.

[13] J. F. Toland. A duality principle for nonconvex optimisation and the calculus of variations. *Arch. Rational Mech. Anal.*, 71(1):41–61, 1979.